

# Advances in Smart Antennas

Guest Editors: Andreas Czylik, Alex B. Gershman,  
and Thomas Kaiser



EURASIP Journal on Applied Signal Processing

---

## **Advances in Smart Antennas**

EURASIP Journal on Applied Signal Processing

---

## **Advances in Smart Antennas**

Guest Editors: Andreas Czylik, Alex B. Gershman,  
and Thomas Kaiser



---

Copyright © 2004 Hindawi Publishing Corporation. All rights reserved.

This is a special issue published in volume 2004 of "EURASIP Journal on Applied Signal Processing." All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Editor-in-Chief

Marc Moonen, Belgium

## Senior Advisory Editor

K. J. Ray Liu, College Park, USA

## Associate Editors

Gonzalo Arce, USA

Jaakko Astola, Finland

Kenneth Barner, USA

Mauro Barni, Italy

Sankar Basu, USA

Jacob Benesty, Canada

Helmut Bölcskei, Switzerland

Joe Chen, USA

Chong-Yung Chi, Taiwan

M. Reha Civanlar, Turkey

Luciano Costa, Brazil

Satya Dharanipragada, USA

Petar M. Djurić, USA

Jean-Luc Dugelay, France

Touradj Ebrahimi, Switzerland

Frank Ehlers, Germany

Moncef Gabbouj, Finland

Sharon Gannot, Israel

Fulvio Gini, Italy

A. Gorokhov, The Netherlands

Peter Handel, Sweden

Ulrich Heute, Germany

John Homer, Australia

Jiri Jan, Czech

Søren Holdt Jensen, Denmark

Mark Kahrs, USA

Thomas Kaiser, Germany

Moon Gi Kang, Korea

Aggelos Katsaggelos, USA

Walter Kellermann, Germany

Alex Kot, Singapore

C.-C. Jay Kuo, USA

Chin-Hui Lee, USA

Sang Uk Lee, Korea

Geert Leus, The Netherlands

Mark Liao, Taiwan

Yuan-Pei Lin, Taiwan

Bernie Mulgrew, UK

King N. Ngan, Hong Kong

Douglas O'Shaughnessy, Canada

Antonio Ortega, USA

Montse Pardas, Spain

Vincent Poor, USA

Phillip Regalia, France

Markus Rupp, Austria

Hideaki Sakai, Japan

Bill Sandham, UK

Dirk Slock, France

Piet Sommen, The Netherlands

John Sorensen, Denmark

Sergios Theodoridis, Greece

Dimitrios Tzovaras, Greece

Jacques Verly, Belgium

Xiaodong Wang, USA

Douglas Williams, USA

Xiang-Gen Xia, USA

Jar-Ferr Yang, Taiwan

# Contents

**Editorial**, Andreas Czylwik, Alex B. Gershman, and Thomas Kaiser

Volume 2004 (2004), Issue 9, Pages 1189-1190

**Receiver Orientation versus Transmitter Orientation in Linear MIMO Transmission Systems**,

Michael Meurer, Paul Walter Baier, and Wei Qiu

Volume 2004 (2004), Issue 9, Pages 1191-1198

**Spatial-Mode Selection for the Joint Transmit and Receive MMSE Design**, Nadia Khaled,

Claude Desset, Steven Thoen, and Hugo De Man

Volume 2004 (2004), Issue 9, Pages 1199-1211

**Blind Identification of Out-of-Cell Users in DS-CDMA**, Tao Jiang and Nicholas D. Sidiropoulos

Volume 2004 (2004), Issue 9, Pages 1212-1224

**Nonlinear Multiantenna Detection Methods**, Sheng Chen, Lajos Hanzo, and Andreas Wolfgang

Volume 2004 (2004), Issue 9, Pages 1225-1237

**Upper Bounds on the BER Performance of MTCM-STBC Schemes over Shadowed Rician Fading**

**Channels**, M. Uysal and C. N. Georghiades

Volume 2004 (2004), Issue 9, Pages 1238-1245

**Full-Rate Full-Diversity Linear Quasi-Orthogonal Space-Time Codes for Any Number of Transmit**

**Antennas**, Naresh Sharma and Constantinos B. Papadias

Volume 2004 (2004), Issue 9, Pages 1246-1256

**Optimal STBC Precoding with Channel Covariance Feedback for Minimum Error Probability**,

Yi Zhao, Raviraj Adve, and Teng Joon Lim

Volume 2004 (2004), Issue 9, Pages 1257-1265

**A Soft Detector with Good Performance/Complexity Trade-Off for a MIMO System**, Jianhua Liu

and Jian Li

Volume 2004 (2004), Issue 9, Pages 1266-1274

**OFDM versus Single Carrier: A Realistic Multi-Antenna Comparison**, Jan Tubbx,

Liesbet Van der Perre, Marc Engels, Hugo De Man, and Marc Moonen

Volume 2004 (2004), Issue 9, Pages 1275-1287

**Bounds on the Outage-Constrained Capacity Region of Space-Division Multiple-Access Radio**

**Systems**, Haipeng Jin and Anthony Acampora

Volume 2004 (2004), Issue 9, Pages 1288-1298

**A Combined Antenna Arrays and Reverse-Link Synchronous DS-CDMA System over**

**Frequency-Selective Fading Channels with Power Control Error**, Yong-Seok Kim, Seung-Hoon Hwang,

Hyo-Yol Park, and Keum-Chan Whang

Volume 2004 (2004), Issue 9, Pages 1299-1307

**System-Level Performance of Antenna Arrays in CDMA-Based Cellular Mobile Radio Systems,**

Andreas Czylik and Armin Dekorsy

Volume 2004 (2004), Issue 9, Pages 1308-1320

**Gaussian Channel Model for Mobile Multipath Environment,** D. D. N. Bevan, V. T. Ermolayev,

A. G. Flaksman, and I. M. Averin

Volume 2004 (2004), Issue 9, Pages 1321-1329

**Downlink Channel Estimation in Cellular Systems with Antenna Arrays at Base Stations Using Channel Probing with Feedback,** Mehrzad Biguesh and Alex B. Gershman

Volume 2004 (2004), Issue 9, Pages 1330-1339

**Diversity Properties of Multiantenna Small Handheld Terminals,** Wim A. Th. Kotterman,

Gert F. Pedersen, and Kim Olesen

Volume 2004 (2004), Issue 9, Pages 1340-1353

**Multidimensional Rank Reduction Estimator for Parametric MIMO Channel Models,**

Marius Pesavento, Christoph F. Mecklenbräuer, and Johann F. Böhme

Volume 2004 (2004), Issue 9, Pages 1354-1363

**Direction of Arrival Estimation with a Novel Single-Port Smart Antenna,** Chen Sun

and Nemai C. Karmakar

Volume 2004 (2004), Issue 9, Pages 1364-1375

**High-Performance Wireless via the Merger of CI Chip-Shaped DS-CDMA and Oscillating-Beam Smart Antenna Arrays,** Seyed Alireza Zekavat, Carl R. Nassar, and Steve Shattil

Volume 2004 (2004), Issue 9, Pages 1376-1383

**ADAM: A Realistic Implementation for a W-CDMA Smart Antenna,**

Ramón Martínez Rodríguez-Osorio, Laura García García, Alberto Martínez Ollero,

Francisco Javier García-Madrid Velázquez, Leandro de Haro Ariet, and Miguel Calvo Ramón

Volume 2004 (2004), Issue 9, Pages 1384-1406

**A Multiple-Antenna System for ISM-Band Transmission,** J. Rinas, R. Seeger, L. Brötje, S. Vogeler,

T. Haase, and K.-D. Kammeyer

Volume 2004 (2004), Issue 9, Pages 1407-1419

**Exploiting Phase Diversity for CDMA2000 1X Smart Antenna Base Stations,** Seongdo Kim,

Seungheon Hyeon, and Seungwon Choi

Volume 2004 (2004), Issue 9, Pages 1420-1430

## Editorial

### Andreas Czylik

*Department of Communication Systems, University of Duisburg-Essen, 47048 Duisburg, Germany*  
*Email: czylik@sent5.uni-duisburg.de*

### Alex B. Gershman

*Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada L8S 4K1*  
*Email: gershman@ieee.org*

### Thomas Kaiser

*Department of Communication Systems, Faculty of Engineering, University of Duisburg-Essen, 47048 Duisburg, Germany*  
*Email: thomas.kaiser@uni-duisburg.de*

Smart antennas have recently emerged as a key technology for third (3G) and higher generations of wireless communication systems. They are able to add a new *spatial* dimension on top of the currently used *time*, *frequency*, and *code* multiple access schemes. The past of 3G wireless systems licensing process in Europe has shown that spectral bandwidth may cost wireless system providers billions of Euros. In light of this fact, smart antennas offer an elegant and relatively inexpensive opportunity of increasing data rate, system capacity and flexibility, and quality of service. Today developments and progress in this area are still away from cost-efficient practical implementations, and further theoretical and experimental studies are crucial to enable successful applications of smart antennas in the future.

The aim of this special issue is to present recent advances of the smart antenna research. The contributions cover a broad range of topics. We classified them into three major categories, namely, *link level*, *system level*, and *applications*.

The first ten papers of the issue are related to the link-level study of smart antennas. In the paper by Meurer et al., the MIMO transmitter design approach with a priori given structure of the receiver (the so-called receiver-oriented design scheme) is compared to the case when the MIMO receiver design is performed for a fixed transmitter structure (the so-called transmitter-oriented design scheme). The paper by Khaled et al. is devoted to the optimization of the MMSE-based joint MIMO transmitter and receiver design from the BER perspective. In the paper by Jiang and Sidiropoulos, a new algorithm for blind identification of out-of-cell CDMA users is developed for the case when multiple antennas are used at the base station. The paper by Chen et al. investigates a nonlinear detection technique for multi-antenna receivers. In the next paper, Uysal and Georghiades

study the performance of space-time block codes (STBCs) which are concatenated with outer trellis codes. The study of STBC schemes is continued in the paper by Sharma and Papadias, where a class of full-rate full-diversity linear quasi-orthogonal STBCs is proposed and their practical decoding algorithm is studied, and in the paper by Zhao et al., which develops a linear precoder for OSTBCs when the channel covariance matrix is available at the transmitter. In the paper by Liu and Li, a new soft detector for MIMO systems is proposed which combines the advantages of the LS soft detector and sphere decoder. The paper by Tubbax et al. compares the multiple-antenna generalizations of the OFDM and single-carrier schemes, and the paper by Jin and Acampora investigates the outage capacity regions for a multiuser SDMA system with multiple antennas at the base station.

The next four papers consider system level aspects. The paper by J.-S. Kim et al. considers the link quality in a single cell of a CDMA-based system using adaptive antenna arrays and proposes a receiver which combines beamforming and RAKE reception. The analysis of CDMA-based systems is extended towards multicellular systems with adaptive antennas in the paper by Czylik and Dekorsy. A more specific problem of angle-of-arrival estimation at base stations with multibeam antennas is investigated in the paper by Bevan et al. In the next paper, Biguesh and Gershman study down-link channel estimation techniques in cellular systems using training signals.

The next two papers are devoted to multi-antenna channel measurements. Kotterman et al. focus on diversity properties of multiple antennas at hand-held terminals. The paper by Pesavento et al. proposes an algorithm for simultaneous estimation of MIMO channel parameters from channel sounder measurements.

The last five papers are devoted to hardware and applications issues of smart antennas and MIMO technologies. The first paper by Sun and Karmakar studies the use of active single port antennas with surrounding parasitic antennas in order to perform direction-of-arrival estimation. In the second paper by Zekavat et al., an oscillating antenna array pattern with fixed main lobe direction is introduced to provide transmit diversity. This approach is applied to CDMA systems. The next paper by Rodríguez-Osorio et al. deals with a practical implementation of a W-CDMA multiantenna testbed with the main focus on the most relevant digital signal processing operations such as modulation, demodulation, synchronization, and beamforming. Another multiantenna testbed for ISM band transmission is presented in the paper by Rinas et al. In this paper, the hardware concepts are highlighted and the feasibility of current single- and multicarrier MIMO algorithms is studied via measurements. The last paper by S. Kim et al. studies the use of smart antennas in another relevant 3G standard, CDMA2000.

In summary, the twenty-one papers published in this issue cover a broad range of recent advances of smart antenna research and a variety of their applications to existing and future wireless systems.

We thank the authors, the reviewers, the publisher, and the Editor-in-Chief for their efforts which were critical for the success of this special issue.

*Andreas Czylik  
Alex B. Gershman  
Thomas Kaiser*

**Andreas Czylik** studied electrical engineering at the Technical University of Darmstadt, Germany, from 1978 to 1983. In 1988, he received the Ph.D. degree and in 1994 the Habilitation degree, both from the Technical University of Darmstadt and both in the field of optical communications. From 1994 to 2000, he was with the Research and Development Center (Technologiezentrum) at Deutsche Telekom in the Department for Local Area Broadband Radio Systems. He was in charge of several research projects, for example, a broadband radio communication demonstrator based on single-carrier transmission with frequency domain equalization as well as several projects on smart antenna concepts in cellular mobile radio systems. In 2000, he became a Full Professor at the Technical University of Braunschweig, heading the Department for Microcellular Radio Systems. Since 2002, he has been with University of Duisburg-Essen and in charge of the Department of Communication Systems. He published some 70 papers and was the Editor for IEEE Journal on Selected Areas in Communications and IEEE Transactions on Wireless Communications. In numerous national and international scientific conferences, he was member of the technical program committee and organized or chaired sessions, respectively. His research interests are in the fields of adaptive transmission techniques in radio communications such as smart antennas and adaptive modulation and coding techniques.



**Alex B. Gershman** received his Diploma and Ph.D. degrees in radiophysics from the Nizhny Novgorod University, Russia, in 1984 and 1990, respectively. From 1984 to 1989, he was with the Radiotechnical and Radiophysical Institutes, Nizhny Novgorod. From 1989 to 1997, he was with the Institute of Applied Physics, Nizhny Novgorod. From 1997 to 1999, he was a Research Associate at the Department of Electrical Engineering, Ruhr-University Bochum, Germany. In 1999, he joined the Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ontario, Canada, where he is now a Professor. He also held visiting positions at the Swiss Federal Institute of Technology, Lausanne, Ruhr-University Bochum; and Gerhard-Mercator University, Duisburg. His main research interests are in statistical and array signal processing, adaptive beamforming, MIMO systems and space-time coding, multiuser communications, and parameter estimation. He has published over 220 technical papers in these areas. Dr. Gershman was a recipient of the 1993 URSI Young Scientist Award, the 1994 Outstanding Young Scientist Presidential Fellowship (Russia), the 1994 Swiss Academy of Engineering Science Fellowship, and the 1995-1996 Alexander von Humboldt Fellowship (Germany). He received the 2000 Premier's Research Excellence Award, Ontario, Canada, and the 2001 Wolfgang Paul Award, Alexander von Humboldt Foundation, Germany. He was also a recipient of the 2002 Young Explorers Prize from the Canadian Institute for Advanced Research (CIAR), which has honored Canada's top 20 researchers aged 40 or under. He is an Associate Editor for the IEEE Transactions on Signal Processing and EURASIP Journal on Wireless Communications and Networking, as well as a Member of the SAM Technical Committee of the IEEE Signal Processing Society.



**Thomas Kaiser** received the Diploma degree from the Ruhr-University Bochum in 1991, the Ph.D. degree in 1995 with distinction, and the second Ph.D. degree in 2000 (so-called Habilitation) from Gerhard-Mercator-University Duisburg, all in electrical engineering. From 1995 to 1996, he spent a research leave at the University of Southern California, Los Angeles, grant-aided by the German Academic Exchange Service. From April 2000 to March 2001, he has been Head of the Department of Communication Systems at Gerhard-Mercator-University Duisburg, and from April 2001 to March 2002, he has been Head of the Department of Wireless Chips & Systems (WCS) at Fraunhofer Institute of Microelectronic Circuits and Systems. Now he is with the Department of Communication Systems and Coleader of the Smart Antenna Research Team (SmART) at the University of Duisburg-Essen. Dr. Kaiser published more than 75 papers in international journals and conferences and is coeditor of forthcoming books and special issues on *UWB communication systems* and *Smart Antennas*. He belongs to the editorial board of EURASIP Journal on Applied Signal Processing, the advisory board of a European multiantenna project, and is technical program committee member of several international conferences. His current research interest focuses on applied signal processing with emphasis on multiantenna systems, especially its applicability to ultra-wideband systems and on implementation issues.



# Receiver Orientation versus Transmitter Orientation in Linear MIMO Transmission Systems

**Michael Meurer**

*Research Group for RF Communications, University of Kaiserslautern, P.O. Box 3049, 67653 Kaiserslautern, Germany  
Email: meurer@rhrk.uni-kl.de*

**Paul Walter Baier**

*Research Group for RF Communications, University of Kaiserslautern, P.O. Box 3049, 67653 Kaiserslautern, Germany  
Email: baier@rhrk.uni-kl.de*

**Wei Qiu**

*Research Group for RF Communications, University of Kaiserslautern, P.O. Box 3049, 67653 Kaiserslautern, Germany  
Email: wqiu@rhrk.uni-kl.de*

*Received 23 June 2003; Revised 13 February 2004*

In conventional transmission schemes, the transmitter algorithms are a priori given, whereas the algorithms to be used by the receivers have to be a posteriori adapted. Such schemes can be termed transmitter (Tx) oriented and have the potential of simple transmitter implementations. The opposite to Tx orientation would be receiver (Rx) orientation in which the receiver algorithms are a priori given, and the transmitter algorithms have to be a posteriori adapted. An advantage of the rationale Rx orientation is the possibility to arrive at simple receiver structures. In this paper, linear versions of the rationale Tx orientation and Rx orientation are applied to radio transmission systems with multiantennas both at the transmitter and receiver. After the introduction of adequate models for such multiple-input multiple-output (MIMO) systems, different system designs are evaluated by simulations, and recommendations for proper system solutions are given.

**Keywords and phrases:** MIMO systems, transmitter orientation, receiver orientation.

## 1. INTRODUCTION

In conventional transmission schemes the transmitter algorithms are a priori given and made known to the receiver, whereas the algorithms to be used by the receivers have to be a posteriori adapted, possibly under consideration of channel information. For this approach, where the transmitter (Tx) is the master and the receiver (Rx) is the slave, the authors propose the term Tx orientation. The opposite to Tx orientation would be Rx orientation in which the receiver algorithms would be a priori given and made known to the transmitter, and the transmitter algorithms, again possibly under consideration of channel information, have to be a posteriori adapted correspondingly. Since the early times of radio communications, the rationale Tx orientation has been preferred because, seemingly, it has some kind of natural appeal to system designers. It was not before the 1990s that the first ideas of Rx orientation came up (cf. Table 1). It took another couple of years to clearly formulate this rationale in 2000 [1]. From then on, it attracted broader attention so that a systematical study could begin. This late perception of Rx ori-

entation is astonishing because each of the two approaches, depending on the particular field of application, has its distinct pros. In the case of Tx orientation, the transmitter algorithms to be a priori determined can be chosen with a view to arrive at particularly simple transmitter implementations. On the other hand, in the case of Rx orientation, the receiver algorithms can be a priori determined in such a way that the receiver complexity is minimized. If we consider, as an important example of a radio transmission, mobile radio systems, the complexity of the mobile terminals (MT) should be as low as possible, whereas more complicated implementations can be tolerated at the base stations (BS). Having in mind the above-mentioned complexity features of the rationale Tx orientation and Rx orientation, this means that in the uplink (UL), the quasi natural choice would be Tx orientation, which leads to low-cost transmitters at the MTs, whereas in the downlink (DL), the rationale Rx orientation would be the favourite alternative because this results in simple receivers at the MTs. In [1, 2], the application of the rationale Rx orientation to mobile radio DLs is considered.

TABLE 1: Selected early publications on Rx-oriented transmission in chronological order.

References	Type of system, proposed techniques, and further remarks
[3, 4]	SISO, CDMA with spreading at Tx, design of FIR prefilter (MF criterion) $\Rightarrow$ Pre-Rake
[5]	SISO, CDMA with spreading at Tx, pre-decorrelator (ZF criterion)
[6]	SISO, CDMA with spreading at Tx, pre-decorrelator (ZF criterion)
[7]	SISO, CDMA with spreading at Tx, pre-decorrelator (ZF criterion) and pre-MMSE (MMSE criterion)
[8]	MISO, CDMA with spreading at Tx, design of FIR prefilter (MF / ZF / MMSE criterion) $\Rightarrow$ Pre-Rake
[9]	SISO, CDMA with spreading at Tx, design of FIR prefilter (MF criterion) $\Rightarrow$ Pre-Rake
[10]	MIMO, MMSE processing (MMSE criterion)
[11]	MISO, CDMA, joint transmission (ZF criterion) $\Rightarrow$ TxZF
[12]	MISO, CDMA, joint pre-distortion (ZF criterion) $\Rightarrow$ TxZF
[13]	SISO, CDMA with spreading at Tx, design of FIR prefilter (ZF criterion)
[14]	MISO, CDMA, joint transmission (ZF criterion) $\Rightarrow$ TxZF

As mentioned above, in the case of Tx orientation, channel knowledge would be desirable at the MTs, whereas in the case of Rx orientation, such knowledge should be available at the BSs. This means that, in the case of mobile radio systems, the above proposed combination of Tx orientation in the UL and Rx orientation in the DL is particularly easily feasible, if the utilized duplexing scheme is time division duplexing (TDD). In TDD, the UL and the DL use the same frequency in temporally separated periods so that, due to the reciprocity theorem, both links experience the same channel impulse responses as long as the time elapsing between UL and DL transmissions is not too large. Therefore, the channel knowledge needed by the BS receivers in the Tx-oriented UL and obtainable for instance based on the transmission of training signals by the MTs can be used also as the channel knowledge required for the Rx-oriented DL transmission. This approach to exploit channel knowledge available in the BS for DL transmission has the additional advantage that no resources have to be sacrificed for the transmission of training signals in the DL, which is, anyhow, capacity-wise the more critical one of the two links.

An important asset with respect to increasing the spectrum efficiency of radio transmission systems is the use of multiantennas instead of single antennas at both the transmitter and the receiver [15, 16]. Such multi-antenna structures were given the designation multiple input multiple output (MIMO). A series of theoretical results concerning the capacity of MIMO systems [17, 18] and the implementation of such systems [19, 20] came up in recent years. The present paper has the goal to study and compare the rationales Tx

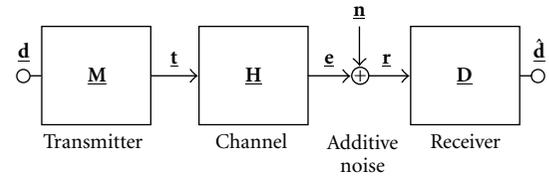


FIGURE 1: Generic model of a linear transmission system.

orientation and Rx orientation and to show some dualities and differences, if linear versions of these schemes are utilized in combination with MIMO antenna structures. Linear systems have, in contrast to nonlinear systems as for instance considered in [21], the advantage of lower complexity [22, 23]. Nevertheless, also in linear systems, a beneficial nonlinear feature can be introduced by operating the linear inner MIMO system in combination with outer FEC coding at the transmitter and FEC decoding at the receiver.

In Section 2, a generic model of linear transmission systems is developed. The topic of Section 3 is the detailed description of the rationales Tx orientation and Rx orientation under inclusion of the linear algorithms to be applied at the transmitters and receivers. In this section, also the quantity signal-to-noise-plus-interference ratio (SNIR) suitable for performance of comparisons of the two rationales is introduced. The generic model developed in Section 2 and the findings of Section 3 are adapted to linear MIMO transmission systems in Section 4. Section 5 presents the results of system simulations; these results help to decide in which cases Tx orientation or Rx orientation should be chosen. Finally, Section 6 summarizes the paper.

The investigations are performed in the time-discrete equivalent low-pass domain under utilization of the vector-matrix representation of signals and system components [24]. Consequently, signals and channel impulse responses are represented by complex vectors or matrices which are printed in bold face. In the analysis,  $[\cdot]_{n,n}$  designates the  $n$ th diagonal element of a square matrix in brackets,  $[\cdot]_n$  stands for the  $n$ th row of a matrix in brackets or the  $n$ th element of a vector in brackets, and  $\|\cdot\|_2$  denotes the Euclidean norm of the vector in brackets. Moreover, the operation  $\text{diag}(\cdot)$  yields a copy of the matrix in brackets with the diagonal elements being set to zero.

## 2. GENERIC MODEL OF LINEAR TRANSMISSION SYSTEMS

Figure 1 shows the generic model of a linear transmission system. In this model, the transmitter, the channel, and the receiver are described by the matrices  $\mathbf{M}$ ,  $\mathbf{H}$ , and  $\mathbf{D}$ , respectively [1].  $\mathbf{M}$ ,  $\mathbf{H}$ , and  $\mathbf{D}$  are termed modulator matrix, channel matrix, and demodulator matrix, respectively. The signals occurring in the structure of Figure 1 are represented by the following column vectors:

- (i)  $\mathbf{d}$ : data signal to be transmitted,
- (ii)  $\mathbf{t}$ : transmit signal,
- (iii)  $\mathbf{e}$ : useful receive signal at the channel output,

TABLE 2: Dimensions of the vectors and matrices used in the structure of Figure 1.

Vector or matrix, respectively	Dimensions
$\underline{\mathbf{d}} = (d_1, \dots, d_N)^T$	$\mathbb{C}^{N \times 1}$
$\underline{\mathbf{M}}$	$\mathbb{C}^{Q \times N}$ , $Q \geq N$
$\underline{\mathbf{t}}$	$\mathbb{C}^{Q \times 1}$
$\underline{\mathbf{H}}$	$\mathbb{C}^{S \times Q}$
$\underline{\mathbf{e}}$	$\mathbb{C}^{S \times 1}$
$\underline{\mathbf{n}}$	$\mathbb{C}^{S \times 1}$
$\underline{\mathbf{r}}$	$\mathbb{C}^{S \times 1}$
$\underline{\mathbf{D}}$	$\mathbb{C}^{N \times S}$
$\hat{\underline{\mathbf{d}}}$	$\mathbb{C}^{N \times 1}$

- (iv)  $\underline{\mathbf{n}}$ : Gaussian noise signal at the receiver input,
- (v)  $\underline{\mathbf{r}}$ : disturbed signal at the receiver input,
- (vi)  $\hat{\underline{\mathbf{d}}}$ : linear estimate of  $\underline{\mathbf{d}}$  at the receiver output.

The dimensions of the vectors and matrices used in the structure of Figure 1 are specified in Table 2.

The elements  $d_n$ ,  $n = 1, \dots, N$ , of  $\underline{\mathbf{d}}$  are the data symbols to be transmitted and are taken from a finite symbol set

$$\mathbb{V} = \{\underline{v}_1 \cdots \underline{v}_M\} \quad (1)$$

of cardinality  $M$ .  $\underline{\mathbf{d}}$  and  $\underline{\mathbf{n}}$  are assumed to be wide-sense stationary with zero mean and the covariance matrices

$$\underline{\mathbf{R}}_{\underline{\mathbf{d}}} = 2\sigma_d^2 \mathbf{I}^{N \times N}, \quad (2)$$

$$\underline{\mathbf{R}}_{\underline{\mathbf{n}}} = 2\sigma_n^2 \mathbf{I}^{S \times S}, \quad (3)$$

respectively. In the system of Figure 1, the estimate  $\hat{\underline{\mathbf{d}}}$  of  $\underline{\mathbf{d}}$  obtained at the receiver output can be expressed as

$$\begin{aligned} \hat{\underline{\mathbf{d}}} &= (\hat{d}_1 \cdots \hat{d}_N)^T = \underline{\mathbf{D}} \underline{\mathbf{r}} = \underline{\mathbf{D}}(\underline{\mathbf{e}} + \underline{\mathbf{n}}) = \underline{\mathbf{D}}(\underbrace{\underline{\mathbf{H}} \underline{\mathbf{t}}}_{\underline{\mathbf{e}}} + \underline{\mathbf{n}}) \\ &= \underline{\mathbf{D}}(\underbrace{\underline{\mathbf{H}} \underline{\mathbf{M}} \underline{\mathbf{d}}}_{\underline{\mathbf{t}}} + \underline{\mathbf{n}}) = \underline{\mathbf{D}} \underline{\mathbf{H}} \underline{\mathbf{M}} \underline{\mathbf{d}} + \underline{\mathbf{D}} \underline{\mathbf{n}}. \end{aligned} \quad (4)$$

$\underline{\mathbf{D}} \underline{\mathbf{H}} \underline{\mathbf{M}}$  is a square matrix of dimension  $N \times N$ . Generally, each data symbol  $d_n$ ,  $n = 1, \dots, N$ , has an influence on all  $Q$  elements of  $\underline{\mathbf{t}}$ . Therefore,  $Q$  can be considered as a spreading factor, where, as we will see in Section 4, spreading can have a temporal and a spatial component.

According to (2) and (4), the mean radiated energy invested for the data symbol  $d_n$  becomes

$$T_n = \frac{1}{2} \|\underline{\mathbf{M}}^T\|_n\|_2^2 2\sigma_d^2, \quad (5)$$

where the factor “1/2” results from the low-pass domain representation used within this contribution [25]. By averaging over all  $N$  data symbols  $d_n$ ,  $n = 1, \dots, N$ , we obtain the mean

radiated energy

$$T = \frac{\sigma_d^2}{N} \sum_{n=1}^N \|\underline{\mathbf{M}}^T\|_n\|_2^2 \quad (6)$$

per data symbol.

The estimate  $\hat{d}_n$  of the transmitted data symbol  $d_n$  consists of the sum of a useful part

$$d_{\text{useful},n} = [\underline{\mathbf{D}} \underline{\mathbf{H}} \underline{\mathbf{M}}]_{n,n} d_n, \quad (7)$$

of an interference part

$$d_{\text{int},n} = [\overline{\text{diag}}(\underline{\mathbf{D}} \underline{\mathbf{H}} \underline{\mathbf{M}})]_n d_n, \quad (8)$$

and of a noise part

$$d_{\text{noise},n} = [\underline{\mathbf{D}} \underline{\mathbf{n}}]_n; \quad (9)$$

see also [24]. In (8) and (9), the terms in brackets are column vectors. A concise and obvious quality measure for the estimates  $\hat{d}_n$  of (4) are the SNIRs  $\gamma_n$  [24]. With (2), (3), (7), (8), and (9), we obtain

$$\begin{aligned} \gamma_n &= \frac{\mathbb{E}\{|d_{\text{useful},n}|^2\}}{\mathbb{E}\{|d_{\text{noise},n}|^2\} + \mathbb{E}\{|d_{\text{int},n}|^2\}} \\ &= \frac{|\underline{\mathbf{D}} \underline{\mathbf{H}} \underline{\mathbf{M}}]_{n,n}|^2 \sigma_d^2}{\|\underline{\mathbf{D}}\|_n\|_2^2 \sigma_n^2 + \|\overline{\text{diag}}(\underline{\mathbf{D}} \underline{\mathbf{H}} \underline{\mathbf{M}})]_n\|_2^2 \sigma_d^2}. \end{aligned} \quad (10)$$

Even though in this paper,  $\gamma_n$  is adopted as the quality measure and quantitatively studied, ultimately the symbol error probabilities would be the proper measure. Fortunately, in many cases, noise plus interference can be modeled as white Gaussian noise with sufficient accuracy. Then, the error probabilities immediately follow from the values  $\gamma_n$ . Otherwise, also the probability density function of noise plus interference has to be taken into account.

### 3. TRANSMITTER ORIENTATION AND RECEIVER ORIENTATION

The a posteriori determination of  $\underline{\mathbf{D}}$  in the case of linear Tx orientation or of  $\underline{\mathbf{M}}$  in the case of linear Rx orientation have to be performed under the consideration of certain criteria. Depending on these criteria, different matrices  $\underline{\mathbf{D}}$  or  $\underline{\mathbf{M}}$ , respectively, result. In what follows, first expressions for determining  $\underline{\mathbf{D}}$  or  $\underline{\mathbf{M}}$ , respectively, are presented, and only then it will be explained which criteria stand behind these expressions. The authors believe that this procedure facilitates the understanding of the presentation, even though the said expressions are consequences of the related criteria.

In the case of Tx orientation,  $\underline{\mathbf{M}}$  and  $\underline{\mathbf{H}}$  are a priori given, whereas  $\underline{\mathbf{D}}$  is a posteriori determined at the Rx based on the knowledge of  $\underline{\mathbf{M}}$  and  $\underline{\mathbf{H}}$ . Well-known approaches for determining  $\underline{\mathbf{D}}$  are the receive matched filter (RxMF), the receive

zero forcer (RxZF), and the receive minimum mean square error estimator (RxMMSE) [24]. In these three cases, the demodulator matrix is a posteriori determined according to [24]

$$\underline{\mathbf{D}} = \begin{cases} (\underline{\mathbf{H}}\underline{\mathbf{M}})^{\mathbf{H}} & \text{(RxMF),} \\ [(\underline{\mathbf{H}}\underline{\mathbf{M}})^{\mathbf{H}}\underline{\mathbf{H}}\underline{\mathbf{M}}]^{-1}(\underline{\mathbf{H}}\underline{\mathbf{M}})^{\mathbf{H}} & \text{(RxZF),} \\ [(\underline{\mathbf{H}}\underline{\mathbf{M}})^{\mathbf{H}}\underline{\mathbf{H}}\underline{\mathbf{M}} + \sigma^2\mathbf{I}^{N \times N}]^{-1}(\underline{\mathbf{H}}\underline{\mathbf{M}})^{\mathbf{H}} & \text{(RxMMSE).} \end{cases} \quad (11)$$

In the case of Rx orientation,  $\underline{\mathbf{H}}$  and  $\underline{\mathbf{D}}$  are a priori given, and  $\underline{\mathbf{M}}$  is a posteriori determined at the Tx based on the knowledge of  $\underline{\mathbf{H}}$  and  $\underline{\mathbf{D}}$ . Approaches meanwhile quite well known to determining  $\underline{\mathbf{M}}$  are the transmit matched filter (TxMF) and the transmit zero forcer (TxZF) [1, 2]. For these, the modulator matrix is a posteriori determined as follows:

$$\underline{\mathbf{M}} = \begin{cases} (\underline{\mathbf{D}}\underline{\mathbf{H}})^{\mathbf{H}}, & \text{(TxMF)} \\ (\underline{\mathbf{D}}\underline{\mathbf{H}})^{\mathbf{H}}[\underline{\mathbf{D}}\underline{\mathbf{H}}(\underline{\mathbf{D}}\underline{\mathbf{H}})^{\mathbf{H}}]^{-1}. & \text{(TxZF)} \end{cases} \quad (12)$$

Other options for Rx orientation are various kinds of transmit minimum mean square error modulators (TxMMSE). In one version, which leads to a closed-form expression for  $\underline{\mathbf{M}}$ , we set out from a given average transmit energy  $T$ , see (6), and, under this condition, determine  $\underline{\mathbf{M}}$  with a real scalar  $k$  according to

$$\underline{\mathbf{M}} = k(\underline{\mathbf{D}}\underline{\mathbf{H}})^{\mathbf{H}} \left[ \underline{\mathbf{D}}\underline{\mathbf{H}}(\underline{\mathbf{D}}\underline{\mathbf{H}})^{\mathbf{H}} + \frac{\sigma^2}{NT} \text{trace}(\underline{\mathbf{D}}\underline{\mathbf{D}}^{\mathbf{H}})\mathbf{I}^{N \times N} \right]^{-1},$$

$$\text{s.t. } \frac{\sigma_d^2}{N} \sum_{n=1}^N \|\underline{\mathbf{M}}^{\mathbf{T}}\|_n^2 \stackrel{!}{=} T \quad \text{by proper choice of } k \quad \text{(TxMMSE).} \quad (13)$$

Equation (13) was first published in [26] in a somewhat different form.

Now we come to the said criteria behind the expressions (11) to (13). The criterion being fulfilled by the Tx-oriented schemes of (11) and the Rx-oriented schemes of (12) is the maximization of  $\gamma_n$  of (10) for a given mean transmit energy  $T_n$  per data symbol  $\underline{d}_n$ , see (5), and under different side conditions, namely [2, 24], the following.

- (1) RxMF, TxMF: the impact of the interference term  $\|[\text{diag}(\underline{\mathbf{D}}\underline{\mathbf{H}}\underline{\mathbf{M}})]_n\|^2\sigma_d^2$  in the denominator on the right-hand side of (10) is neglected.
- (2) RxZF, TxZF: the impact of the interference term  $\|[\text{diag}(\underline{\mathbf{D}}\underline{\mathbf{H}}\underline{\mathbf{M}})]_n\|^2\sigma_d^2$  in the denominator on the right-hand side of (10) is eliminated by forcing this term to zero.
- (3) RxMMSE: an optimum compromise between the impact of the noise term  $\|[\underline{\mathbf{D}}]_n\|^2\sigma^2$  and the interference term  $\|[\text{diag}(\underline{\mathbf{D}}\underline{\mathbf{H}}\underline{\mathbf{M}})]_n\|^2\sigma_d^2$  is brought about.

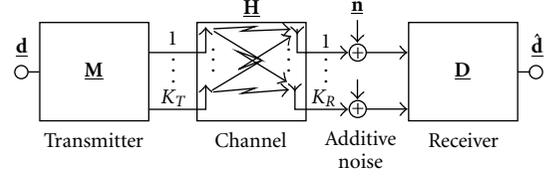


FIGURE 2: Linear MIMO transmission system.

In the case of the TxMMSE of (13), an average SNIR defined as

$$\gamma_{\text{TxMMSE}} = \frac{\sigma_d^2 \sum_{n=1}^N |[\underline{\mathbf{D}}\underline{\mathbf{H}}\underline{\mathbf{M}}]_{n,n}|^2}{\sum_{n=1}^N \{ \|[\underline{\mathbf{D}}]_n\|^2\sigma^2 + \|[\text{diag}(\underline{\mathbf{D}}\underline{\mathbf{H}}\underline{\mathbf{M}})]_n\|^2\sigma_d^2 \}} \quad (14)$$

is maximized for a given mean transmit energy  $T$  of (6) [26].

An important issue when evaluating the transmission schemes of (11) to (13) is the determination of the SNIRs for given mean transmit energies  $T_n$  of (5) or  $T$  of (6). Therefore, the question arises how these energies can be predetermined. In the case of the Tx-oriented schemes of (11), the mean transmit energies  $T_n$  per data symbol can be predetermined based on (5) when a priori establishing  $\underline{\mathbf{M}}$  in a straightforward way. In the case of the TxMF and the TxZF, see (12), the predetermination of  $T_n$  has to be accomplished as follows:

- (i) determine  $\underline{\mathbf{M}}$  by using (12),
- (ii) column-wise scale this  $\underline{\mathbf{M}}$  in such a way that (5) yields the desired mean energies  $T_n$ .

In the case of the TxMMSE, see (13), the mean radiated energy  $T$  per data symbol can again be predetermined in a straightforward way.

The above theory is valid under the implicit understanding that the matrices to be inverted in (11) to (13) are non-singular. This condition is usually fulfilled in reasonably designed systems. However, a closer look at this problem has yet to come.

#### 4. LINEAR MIMO TRANSMISSION SYSTEMS

Figure 2 shows a linear MIMO transmission system with  $K_T$  antennas at the transmitter and  $K_R$  antennas at the receiver. The question is how in the case of such a MIMO system the vectors and matrices introduced in the generic transmission system of Section 2 have to be adjusted in order to make the equations derived in Sections 2 and 3 applicable.

We assume that each data symbol  $\underline{d}_n$  is temporally spread over  $Q_t$  chips [2]. Then, with the  $K_T$  matrices

$$\underline{\mathbf{M}}^{(k_T)} = \begin{pmatrix} \underline{\mathbf{M}}_{1,1}^{(k_T)} & \underline{\mathbf{M}}_{1,2}^{(k_T)} & \cdots & \underline{\mathbf{M}}_{1,N}^{(k_T)} \\ \underline{\mathbf{M}}_{2,1}^{(k_T)} & \underline{\mathbf{M}}_{2,2}^{(k_T)} & \cdots & \underline{\mathbf{M}}_{2,N}^{(k_T)} \\ \vdots & \vdots & \ddots & \vdots \\ \underline{\mathbf{M}}_{Q_t,1}^{(k_T)} & \underline{\mathbf{M}}_{Q_t,2}^{(k_T)} & \cdots & \underline{\mathbf{M}}_{Q_t,N}^{(k_T)} \end{pmatrix} \in \mathbb{C}^{Q_t \times N} \quad (15)$$

termed transmit antenna specific modulator matrices, the (total) modulator matrix takes the form [2]

$$\underline{\mathbf{M}} = \begin{pmatrix} \underline{\mathbf{M}}^{(1)\text{T}} & \underline{\mathbf{M}}^{(2)\text{T}} & \dots & \underline{\mathbf{M}}^{(K_T)\text{T}} \end{pmatrix}^{\text{T}}, \quad (16)$$

$$\underline{\mathbf{M}} \in \mathbb{C}^{(Q_t K_T) \times N}.$$

According to (16), the spreading factor  $Q$  introduced in Table 2 now reads

$$Q = Q_t K_T. \quad (17)$$

This shows that the total spreading quantified by  $Q$  results from a temporal spreading and a spatial spreading represented by  $Q_t$  and  $K_T$ , respectively.

The radio channel between transmit antenna  $k_T$ ,  $k_T = 1, \dots, K_T$ , and receive antenna  $k_R$ ,  $k_R = 1, \dots, K_R$ , can be characterized by the transmit and receive antenna specific impulse response

$$\underline{\mathbf{h}}^{(k_R, k_T)} = \frac{1}{W} \begin{pmatrix} \underline{h}_1^{(k_R, k_T)} & \underline{h}_2^{(k_R, k_T)} & \dots & \underline{h}_W^{(k_R, k_T)} \end{pmatrix}^{\text{T}} \quad (18)$$

of dimension  $W$  [2]. Taking into account that each of the  $K_T$  transmit antennas radiates a signal of dimension  $Q_t \times 1$ , the signal transmission from the transmit antenna  $k_T$ ,  $k_T = 1, \dots, K_T$ , to the receive antenna  $k_R$ ,  $k_R = 1, \dots, K_R$ , can be described by the transmit and receive antenna specific channel matrix

$$\underline{\mathbf{H}}^{(k_R, k_T)} = \begin{pmatrix} \underline{h}_1^{(k_R, k_T)} & 0 & \dots & 0 \\ \underline{h}_2^{(k_R, k_T)} & \underline{h}_1^{(k_R, k_T)} & \ddots & \vdots \\ \vdots & \underline{h}_2^{(k_R, k_T)} & \ddots & 0 \\ \underline{h}_W^{(k_R, k_T)} & \vdots & \ddots & \underline{h}_1^{(k_R, k_T)} \\ 0 & \underline{h}_W^{(k_R, k_T)} & \ddots & \underline{h}_2^{(k_R, k_T)} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \underline{h}_W^{(k_R, k_T)} \end{pmatrix}, \quad (19)$$

$$\underline{\mathbf{H}}^{(k_R, k_T)} \in \mathbb{C}^{(Q_t + W - 1) \times Q_t}.$$

The  $K_R K_T$  transmit and receive antenna specific channel matrices  $\underline{\mathbf{H}}^{(k_R, k_T)}$  of (19) can be stacked to the (total) channel matrix

$$\underline{\mathbf{H}} = \begin{pmatrix} \underline{\mathbf{H}}^{(1,1)} & \underline{\mathbf{H}}^{(1,2)} & \dots & \underline{\mathbf{H}}^{(1, K_T)} \\ \underline{\mathbf{H}}^{(2,1)} & \underline{\mathbf{H}}^{(2,2)} & \dots & \underline{\mathbf{H}}^{(2, K_T)} \\ \vdots & \vdots & \ddots & \vdots \\ \underline{\mathbf{H}}^{(K_R, 1)} & \underline{\mathbf{H}}^{(K_R, 2)} & \dots & \underline{\mathbf{H}}^{(K_R, K_T)} \end{pmatrix}, \quad (20)$$

$$\underline{\mathbf{H}} \in \mathbb{C}^{[(Q_t + W - 1) K_R] \times (Q_t K_T)}.$$

According to (20), the quantity  $S$  introduced in Table 2 can be expressed as

$$S = (Q_t + W - 1) K_R \quad (21)$$

in the case of the considered MIMO system. Therefore, the signals  $\underline{\mathbf{e}}$ ,  $\underline{\mathbf{n}}$ , and  $\underline{\mathbf{r}}$ , see Table 2, have the dimension  $[(Q_t + W - 1) K_R] \times 1$ . Consequently,

$$\underline{\mathbf{D}} \in \mathbb{C}^{N \times [(Q_t + W - 1) K_R]} \quad (22)$$

holds for the demodulator matrix.

With the matrices  $\underline{\mathbf{M}}$ ,  $\underline{\mathbf{H}}$ , and  $\underline{\mathbf{D}}$  defined by (16), (20), and (22), respectively, the different transmission schemes specified by (11), (12), and (13) can be immediately applied to linear MIMO transmission systems.

### 5. SYSTEM EVALUATIONS BY SIMULATIONS

Based on the performance measure SNIR of (10), different versions of linear MIMO transmission systems can be compared and assessed. Questions to be answered by such comparisons concern

- (i) the performance difference of Tx-oriented and Rx-oriented systems,
- (ii) the influence of the antenna numbers  $K_T$  and  $K_R$  on the system performance.

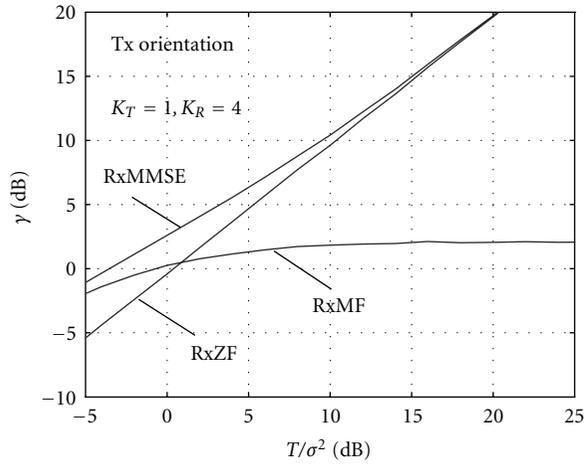
Because a closed-form analysis is not possible, these questions will be addressed by simulations in what follows. Concerning the design of linear MIMO transmission systems, besides the distinction between Tx orientation and Rx orientation, we can choose from a great variety of system parametrizations and channel realizations. In this paper, only a limited selection of such variants can be considered, which, nevertheless, will allow some generally valid statements. In all simulations, we set

$$N = Q_t = W = 4. \quad (23)$$

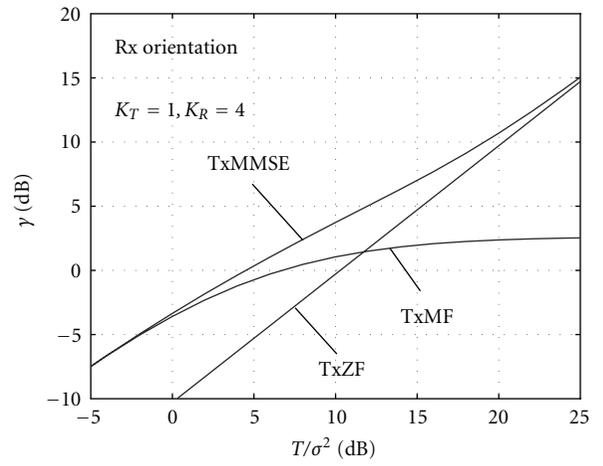
Simulations are performed for different pairs  $K_T$ ,  $K_R$  of antenna numbers. For each such pair, many system realizations are investigated. In each realization, the elements of  $\underline{\mathbf{h}}^{(k_R, k_T)}$  of (18) and—in the case of Tx orientation—the elements of  $\underline{\mathbf{M}}$ , or—in the case of Rx orientation—the elements of  $\underline{\mathbf{D}}$  are chosen as independent realizations of a complex Gaussian random variable with variance 1 of its real and imaginary parts. For a given  $T/\sigma^2$ , by averaging over all  $N$  values  $\gamma_n$  of (10) and all realizations, the mean SNIR  $\gamma$  can be obtained as a function of  $T/\sigma^2$ . Concerning the predetermination of  $T$ , see the last paragraph of Section 3. The determination of  $\underline{\mathbf{h}}^{(k_R, k_T)}$  described above means that all  $K_T K_R$  channel impulse responses are totally uncorrelated. The opposite to this extreme case would be totally correlated channel impulse responses, which, however, are not considered in this paper.

In Figures 3a, 3b, 3c, 3d, 3e, and 3f, the mean SNIR  $\gamma$  is plotted versus  $T/\sigma^2$  for different pairs  $K_T$ ,  $K_R$  and different transmission schemes. The curves in these figures allow the following conclusions.

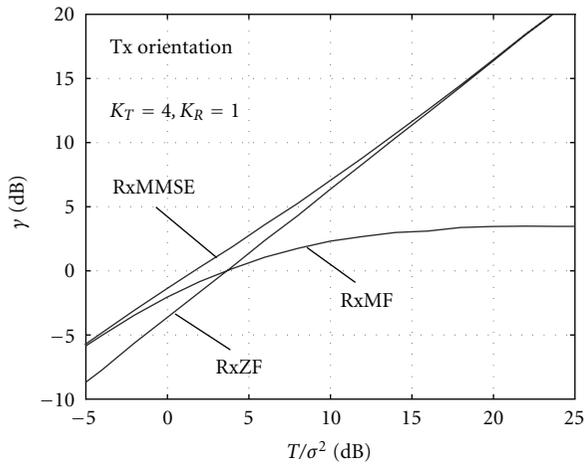
- (1) Both in the case of Tx orientation and Rx orientation, the MF outperforms the ZF for small values of  $T/\sigma^2$ , and the ZF outperforms the MF for large values of  $T/\sigma^2$ . See Figures 3a, 3b, 3c, and 3d.



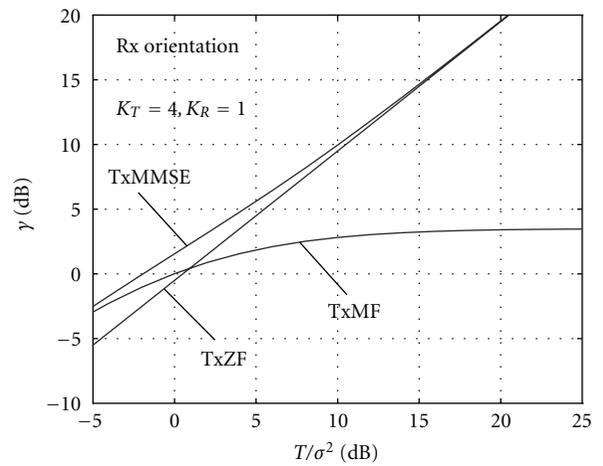
(a)



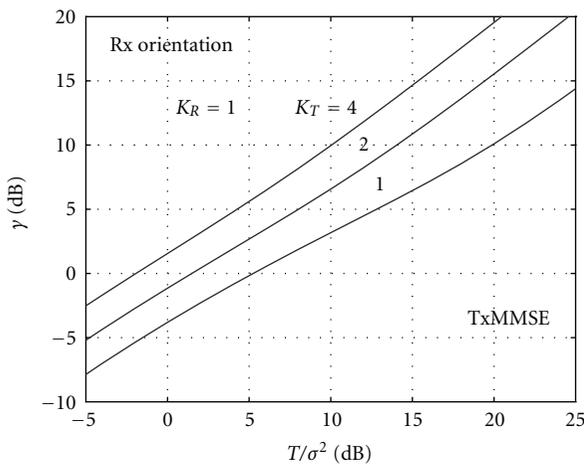
(b)



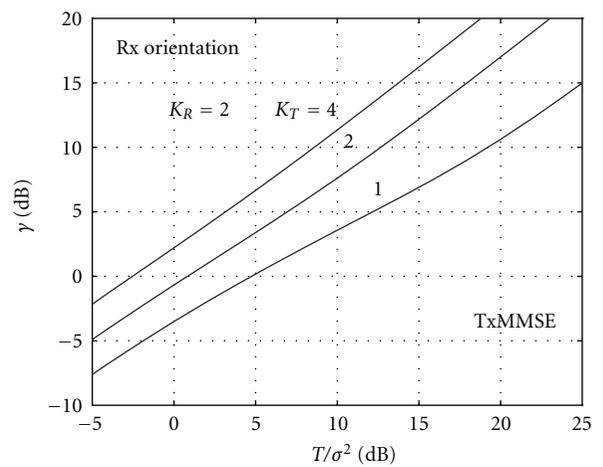
(c)



(d)



(e)



(f)

FIGURE 3: Mean SNIR  $\gamma$  versus  $T/\sigma^2$  for the rationales Tx orientation and Rx orientation and for different combinations  $K_T, K_R$ ;  $N = Q_t = W = 4$ .

- (2) Both in the case of Tx orientation and Rx orientation, the MMSE outperforms the MF and the ZF. For small values of  $T/\sigma^2$ , the performance of the MMSE converges to the performance of the MF, and for large values of  $T/\sigma^2$  to the performance of the ZF. See Figures 3a, 3b, 3c, and 3d.
- (3) If the number  $K_R$  of receive antennas is larger than the number  $K_T$  of transmit antennas, Tx orientation should be chosen because it outperforms Rx orientation. If  $K_R$  is smaller than  $K_T$ , the opposite is true. Compare Figures 3a and 3b, and Figures 3c and 3d.
- (4) The performance is enhanced with growing  $K_T$  and  $K_R$ . See Figures 3e and 3f.

If we compare the Tx-oriented schemes for  $K_T = 1$  and  $K_R = 4$  (see Figure 3a) with the Rx-oriented schemes for  $K_T = 4$  and  $K_R = 1$  (see Figure 3d) or if we compare the Tx-oriented schemes for  $K_T = 4$ ,  $K_R = 1$  (see Figure 3c) with the Rx-oriented schemes for  $K_T = 1$ ,  $K_R = 4$  (see Figure 3b), we can find a very interesting result: if the number of antennas in the two considered schemes both at the a priori given sides and at the a posteriori adapted sides are equal, then the Rx-oriented schemes perform worse than the Tx-oriented schemes. This effect results from the assumption of totally uncorrelated channel impulse responses of dimension  $W$ , which is larger than one.

## 6. SUMMARY

A system model for linear MIMO transmission systems is developed, and this model is worked out for the cases of Tx-oriented and Rx-oriented systems. Based on the system model, performance comparisons and evaluations are made in which the performance measure is the mean SNIR, and the recommendations concerning the system design are given.

## ACKNOWLEDGMENTS

The authors gratefully appreciate the fruitful exchange of ideas with C. A. Jötten, H. Tröger, and T. Weber from the Research Group for RF Communications, University of Kaiserslautern (UKL). The support of individual parts of this work in the framework of the EU-IST-Project FLOWS (Flexible Convergence of Wireless Standards and Services), by DFG, by Siemens AG, and by the supercomputer staff of the central computer facility (RHRK) of the TUKL is highly acknowledged. Thanks are also extended to the anonymous reviewers for their valuable comments and to A. Bruhn and M. Cuntz for, despite all time pressure, carefully typesetting the manuscript in  $\LaTeX$ .

## REFERENCES

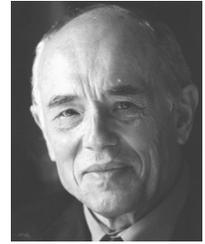
- [1] P. W. Baier, M. Meurer, T. Weber, and H. Tröger, "Joint transmission (JT), an alternative rationale for the downlink of time division CDMA using multi-element transmit antennas," in *Proc. IEEE 6th International Symposium on Spread Spectrum Techniques and Applications (ISSSTA '00)*, vol. 1, pp. 1–5, Parsippany, NJ, USA, September 2000.
- [2] P. W. Baier, W. Qiu, H. Tröger, C. A. Jötten, and M. Meurer, "Modelling and optimization of receiver oriented multi-user MIMO downlinks for frequency selective channels," in *Proc. 10th International Conference on Telecommunications (ICT '03)*, vol. 2, pp. 1547–1554, Papeete, French Polynesia, February 2003.
- [3] R. Esmailzadeh and M. Nakagawa, "Pre-RAKE diversity combination for direct sequence spread spectrum mobile communications systems," *IEICE Transactions on Communications*, vol. 76, no. 8, pp. 1008–1015, 1993.
- [4] R. Esmailzadeh, E. Sourour, and M. Nakagawa, "Pre-RAKE diversity combining in time division duplex CDMA mobile communications," in *Proc. IEEE 6th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '95)*, vol. 2, pp. 431–435, Toronto, Ontario, Canada, September 1995.
- [5] Z. Tang and S. Cheng, "Interference cancellation for DS-CDMA systems over flat fading channels through pre-decorrelating," in *Proc. IEEE 5th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '94)*, vol. 2, pp. 435–438, The Hague, The Netherlands, September 1994.
- [6] H. Matsutani, Y. Sanada, and M. Nakagawa, "A forward link intracell orthogonalization technique using multicarrier pre-decorrelation for CDMA wireless local communication system," in *Proc. IEEE 8th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '97)*, vol. 1, pp. 125–129, Helsinki, Finland, September 1997.
- [7] B. Vojcic and W. M. Jang, "Transmitter precoding in synchronous multiuser communications," *IEEE Transactions on Communications*, vol. 46, no. 10, pp. 1346–1355, 1998.
- [8] G. Montalbano, I. Ghauri, and D. T. M. Slock, "Spatio-temporal array processing for CDMA/SDMA downlink transmission," in *Proc. 32nd Asilomar Conference on Signals, Systems and Computers*, vol. 2, pp. 1337–1341, Pacific Grove, Calif, USA, November 1998.
- [9] A. N. Barreto and G. Fettweis, "On the downlink capacity of TDD CDMA systems using a Pre-RAKE," in *IEEE Global Telecommunications Conference (GLOBECOM '99)*, vol. 1A, pp. 117–121, Rio de Janeiro, Brazil, December 1999.
- [10] H. R. Karimi, M. Sandell, and J. Salz, "Comparison between transmitter and receiver array processing to achieve interference nulling and diversity," in *Proc. IEEE 10th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'99)*, vol. 3, pp. 997–1001, Osaka, Japan, September 1999.
- [11] M. Meurer, P. W. Baier, T. Weber, Y. Lu, and A. Papathanassiou, "Joint transmission: advantageous downlink concept for CDMA mobile radio systems using time division duplexing," *IEE Electronics Letters*, vol. 11, no. 10, pp. 900–901, 2000.
- [12] F. Kowalewski and P. Mangold, "Joint predistortion and transmit diversity," in *Proc. IEEE Global Telecommunications Conference (GLOBECOM '00)*, vol. 1, pp. 245–249, San Francisco, Calif, USA, 2000.
- [13] M. Brandt-Pearce and A. Dharap, "Transmitter-based multiuser interference rejection for the down-link of a wireless CDMA system in a multipath environment," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 3, pp. 407–417, 2000.
- [14] M. Joham and W. Utschick, "Downlink processing for mitigation of intracell interference in DS-CDMA systems," in *Proc. IEEE 6th International Symposium on Spread Spectrum Techniques and Applications (ISSSTA '00)*, vol. 1, pp. 15–19, Parsippany, NJ, USA, September 2000.

- [15] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Transactions on Communications*, vol. 46, no. 3, pp. 357–366, 1998.
- [16] D. Gesbert, M. Shafi, D.-S. Shiu, P. J. Smith, and A. Naguib, "From theory to practice: an overview of MIMO space-time coded wireless systems," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 3, pp. 281–302, 2003.
- [17] G. J. Foschini and M. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, 1998.
- [18] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.
- [19] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, 1996.
- [20] P. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. URSI International Symposium on Signals, Systems, and Electronics (ISSSE '98)*, pp. 295–300, Pisa, Italy, 1998.
- [21] E. Biglieri, G. Taricco, and A. Tulino, "Decoding space-time codes with BLAST architectures," *IEEE Transactions on Signal Processing*, vol. 50, no. 10, pp. 2547–2552, 2002.
- [22] T. Weber and M. Meurer, "Optimum joint transmission: Potentials and dualities," in *Proc. 6th International Symposium on Wireless Personal Multimedia Communications (WPMC '03)*, vol. 1, pp. 79–83, Yokosuka, Japan, October 2003.
- [23] R. Irmer, W. Rave, and G. Fettweis, "Minimum BER transmission for TDD-CDMA in frequency-selective channels," in *Proc. IEEE 14th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '03)*, vol. 2, pp. 1260–1264, Beijing, China, September 2003.
- [24] A. Klein, *Multi-User Detection of CDMA Signals—Algorithms and Their Application to Cellular Mobile Radio*, Number 423 in Fortschrittberichte VDI, Reihe 10. VDI-Verlag, Düsseldorf, Germany, 1996.
- [25] S. Stein and J. J. Jones, *Modern Communication Principles*, McGraw-Hill, New York, NY, USA, 1967.
- [26] M. Joham, K. Kusume, M. H. Gzara, W. Utschick, and J. A. Nossek, "Transmit Wiener filter for the downlink of TDD DS-CDMA systems," in *Proc. IEEE 7th International Symposium on Spread Spectrum Techniques and Applications (ISSSTA '02)*, vol. 1, pp. 9–13, Prague, Czech Republic, September 2002.

**Michael Meurer** was born in Dernbach (Westerwald), Germany, in 1974 and received the diploma in electrical engineering in 1998 and the doctoral degree in 2003, both from the University of Kaiserslautern, Germany. After graduation in October 1998, he joined the Research Group for RF Communications at the University of Kaiserslautern, Germany, as a Research Engineer, where he is presently active as a Senior Research Engineer and Senior Lecturer. His research interests are MIMO systems, receiver-oriented (joint transmission) and channel-oriented (joint transmitter and receiver optimization) transmission concepts, multiuser detection, and statistical signal processing. He is a Member of VDE/ITG and of the IEEE.



**Paul Walter Baier** was born in Backnang, Germany, in 1938, and graduated from the Technical University Munich, Germany. In 1970, he joined Siemens AG, Munich, where he was engaged in various topics of communications engineering. Since 1973, he has been a Professor for electrical communications and Director of the Institute for RF Communications and Fundamentals of Electronic Engineering at the University of Kaiserslautern, Germany. His main research interests are spread spectrum techniques, impulse compression and synthetic aperture radars, mobile radio systems, and adaptive antennas. The basics of the TD-CDMA component of the UMTS Terrestrial Radio Access System (UTRA) agreed upon by 3GPP were developed by him and his coworkers in cooperation with Siemens and in the framework of EU projects. He is a member of VDE/ITG, of the URSI Member Committee Germany, and a Fellow of the IEEE. He was a Scholar of the Japanese Society for the Promotion of Science in 1997 and was awarded the Innovation Prize of the Mannesmann Mobile Radio Foundation in 1999 and the Ring of Honor of VDE Association for Electrical, Electronic & Information Technologies in 2000. Since July 2002, he holds an honorary doctorate of the Technical University Munich.



**Wei Qiu** was born in Jiangsu, China, in 1975. He received his B.E. degree from Tsinghua University, Beijing, China, in 1999, and his M.S. degree from University of Kaiserslautern, Kaiserslautern, Germany, in 2001, both in electrical engineering. Since 2001, he has been a Research Engineer with the Research Group for RF Communications, the University of Kaiserslautern. His research interests are mainly concentrated on mobile radio communications and on MIMO systems. He is a Student Member of IEEE.



# Spatial-Mode Selection for the Joint Transmit and Receive MMSE Design

**Nadia Khaled**

*Interuniversity Micro-Electronics Center (IMEC), Kapeldreef 75, 3001 Leuven, Belgium  
Email: nadia.khaled@imec.be*

**Claude Desset**

*Interuniversity Micro-Electronics Center (IMEC), Kapeldreef 75, 3001 Leuven, Belgium  
Email: claude.desset@imec.be*

**Steven Thoen**

*RF Micro Devices, Technologielaan 4, 3001 Leuven, Belgium  
Email: sthoen@rfmd.com*

**Hugo De Man**

*Interuniversity Micro-Electronics Center (IMEC), Kapeldreef 75, 3001 Leuven, Belgium  
Email: hugo.deman@imec.be*

*Received 28 May 2003; Revised 15 March 2004*

To approach the potential MIMO capacity while optimizing the system bit error rate (BER) performance, the joint transmit and receive minimum mean squared error (MMSE) design has been proposed. It is the optimal linear scheme for spatial multiplexing MIMO systems, assuming a fixed number of spatial streams  $p$  as well as a fixed modulation and coding across these spatial streams. However, state-of-the-art designs arbitrarily choose and fix the value of the number of spatial streams  $p$ , which may lead to an inefficient power allocation strategy and a poor BER performance. We have previously proposed to relax the constraint of fixed number of streams  $p$  and to optimize this value under the constraints of fixed average total transmit power and fixed spectral efficiency, which we referred to as *spatial-mode selection*. Our previous selection criterion was the minimization of the system sum MMSE. In the present contribution, we introduce a new and better spatial-mode selection criterion that targets the minimization of the system BER. We also provide a detailed performance analysis, over flat-fading channels, that confirms that our proposed spatial-mode selection significantly outperforms state-of-the-art joint Tx/Rx MMSE designs for both uncoded and coded systems, thanks to its better exploitation of the MIMO spatial diversity and more efficient power allocation.

**Keywords and phrases:** MIMO systems, spatial multiplexing, joint transmit and receive optimization, selection.

## 1. INTRODUCTION

Over the past few years, multiple-input multiple-output (MIMO) communication systems have prevailed as the key enabling technology for future-generation broadband wireless networks, thanks to their huge potential spectral efficiencies [1]. Such spectral efficiencies are related to the multiple parallel spatial subchannels that are opened through the use of multiple-element antennas at both the transmitter and receiver. These available spatial subchannels can be used to transmit parallel independent data streams, what is referred to as spatial multiplexing (SM) [2, 3]. To enable SM, joint transmit and receive space-time processing has emerged as a powerful and promising design approach for applications,

where the channel is slowly varying such that the channel state information (CSI) can be made available at both sides of the transmission link. In fact, the latter design approach exploits this CSI to optimally allocate resources such as power and bits over the available spatial subchannels so as to either maximize the system's information rate [4] or alternatively reduce the system's bit error rate (BER) [5, 6, 7, 8].

In this contribution, we adopt the second design alternative, namely, optimizing the system BER under the constraints of fixed rate and fixed transmit power. Moreover, among the possible design criteria, we retain the joint transmit and receive minimum mean squared error (joint Tx/Rx MMSE), initially proposed in [5] and further discussed in [7, 8], for it is the optimal linear solution for fixed coding and

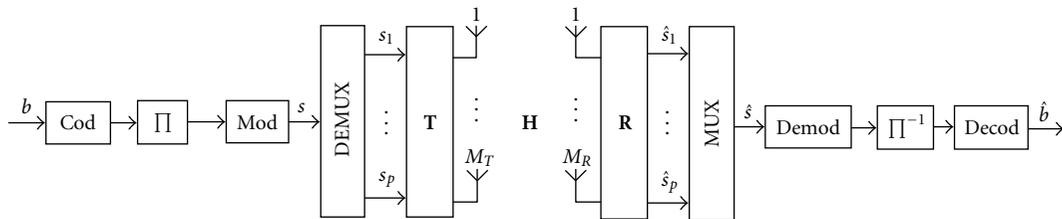


FIGURE 1: The considered  $(M_T, M_R)$  spatial multiplexing MIMO system using linear joint transmit and receive optimization.

symbol constellation across spatial subchannels or modes. The latter constraint is set to reduce the system's complexity and adaptation requirements, in comparison with the optimal yet complex bit loading [9].

Nevertheless, state-of-the-art contributions initially and arbitrarily fix the number of used SM data streams  $p$  [5, 6, 7, 8]. We have previously argued that, compared to their channel-aware power allocation policies, the initial, arbitrary,<sup>1</sup> and static choice of the number of transmit data streams  $p$  is suboptimal [10]. More specifically, we have highlighted the highly inefficient transmit power allocation and poor BER performance this approach may lead to. Consequently, we have proposed to include the number of streams  $p$  as an additional design parameter, rather than a mere arbitrary fixed scalar as in state-of-the-art contributions, to be optimized in order to minimize the joint Tx/Rx MMSE design's BER [10, 11]. A remark in [7] previously raised this issue without pursuing it. The optimization criterion, therein proposed, was the minimization of the sum MMSE and has been also investigated in [10, 11] for flat-fading and frequency-selective fading channels, respectively. The sum MMSE minimization criterion, however, is obviously suboptimal as it equivalently overlooks the joint Tx/Rx MMSE design  $p$  parallel modes as a single one whose BER is minimized. Consequently, it fails to identify the optimal MSEs and BERs on the individual spatial streams that would actually minimize the system average BER. In the present contribution, a better spatial-mode selection criterion is proposed which, on the contrary, examines the BERs on the individual spatial modes in order to identify the optimal number of spatial streams to be used for a minimum system average BER. Finally, spatial-mode selection has also been investigated in the context of space-time coded MIMO systems in presence of imperfect CSI at the transmitter [12, 13]. The therein developed solutions, however, do not apply for spatial multiplexing scenarios, which are the focus of the present contribution.

The rest of the paper is organized as follows. Section 2 provides the system model and describes state-of-the-art joint Tx/Rx MMSE designs. Based on that, Section 3 derives the proposed spatial-mode selection. In Section 4, the BER performance improvements enabled by the proposed spatial-

mode selection are assessed for both uncoded and coded systems. Finally, we draw the conclusions in Section 5.

### Notations

In all the following, normal letters designate scalar quantities, boldface lower case letters indicate vectors, and boldface capitals represent matrices; for instance,  $\mathbf{I}_p$  is the  $p \times p$  identity matrix. Moreover,  $\text{trace}(\mathbf{M})$ ,  $[\mathbf{M}]_{i,j}$ ,  $[\mathbf{M}]_{\cdot,j}$ ,  $[\mathbf{M}]_{\cdot,1:j}$ , respectively, stand for the trace, the  $(i, j)$ th entry, the  $j$ th column, and the  $j$  first columns of matrix  $\mathbf{M}$ .  $[x]^+$  refers to  $\text{Max}(x, 0)$  and  $(\cdot)^H$  denotes the conjugate transpose of a vector or a matrix. Finally,  $\|\mathbf{m}\|_2$  indicates the 2-norm of vector  $\mathbf{m}$ .

## 2. SYSTEM MODEL AND PRELIMINARIES

### 2.1. System model

The SM MIMO wireless communication system under consideration is depicted in Figure 1. It consists of a transmitter and a receiver, both equipped with multiple-element antennas and assumed to have perfect knowledge about the current channel realization. At the transmitter, the input bit stream  $b$  is coded, interleaved, and modulated according to a predetermined symbol constellation of size  $M_p$ . The resulting symbol stream  $s$  is then demultiplexed into  $p \leq \text{Min}(M_R, M_T)$  independent streams. The latter SM operation actually converts the serial symbol stream  $s$  into a higher-dimensional symbol stream where every symbol is a  $p$ -dimensional spatial symbol, for instance,  $\mathbf{s}(k)$  at discrete-time index  $k$ . These spatial symbols are then passed through the linear precoder  $\mathbf{T}$  in order to optimally adapt them to the current channel realization prior to transmission through the  $M_T$ -element transmit antenna. At the receiver, the  $M_R$  symbol-sampled complex baseband outputs from the  $M_R$ -element receive antenna are passed through the linear decoder  $\mathbf{R}$  matched to the precoder  $\mathbf{T}$ . The resulting  $p$  output streams conveying the detected spatial symbols  $\hat{\mathbf{s}}(k)$  are then multiplexed, demodulated, deinterleaved, and decoded to recover the initially transmitted bit stream. For a flat-fading MIMO channel, the global system equation is given by

$$\underbrace{\begin{bmatrix} \hat{s}_1(k) \\ \vdots \\ \hat{s}_p(k) \end{bmatrix}}_{\hat{\mathbf{s}}(k)} = \mathbf{RHT} \underbrace{\begin{bmatrix} s_1(k) \\ \vdots \\ s_p(k) \end{bmatrix}}_{\mathbf{s}(k)} + \mathbf{R} \underbrace{\begin{bmatrix} n_1(k) \\ \vdots \\ n_{M_R}(k) \end{bmatrix}}_{\mathbf{n}(k)}, \quad (1)$$

<sup>1</sup>It is set to either the rank of the MIMO channel matrix [7] or an arbitrary value [6, 8],  $p \leq \text{Min}(M_T, M_R)$ .

where  $\mathbf{n}(k)$  is the  $M_R$ -dimensional receiver noise vector at discrete-time index  $k$ .  $\mathbf{H}$  is the  $M_R \times M_T$  channel matrix whose  $(i, j)$ th entry  $[\mathbf{H}]_{i,j}$  represents the complex channel gain between the  $j$ th transmit antenna element and the  $i$ th receive antenna element. In all the following, the discrete-time index  $k$  is dropped for clarity.

## 2.2. Generic joint Tx/Rx MMSE design

The linear precoder and decoder  $\mathbf{T}$  and  $\mathbf{R}$  represented by an  $M_T \times p$  and  $p \times M_R$  matrix, respectively, are *jointly* designed to minimize the sum mean squared error (MSE) on the spatial symbols  $\mathbf{s}$  subject to fixed average total transmit power  $P_T$  constraint [6] as stated in the following:

$$\begin{aligned} \text{Min}_{\mathbf{R}, \mathbf{T}} E_{\mathbf{s}, \mathbf{n}} \{ \|\mathbf{s} - (\mathbf{R}\mathbf{H}\mathbf{T}\mathbf{s} + \mathbf{R}\mathbf{n})\|_2^2 \} \\ \text{subject to: } E_s \cdot \text{trace}(\mathbf{T}\mathbf{T}^H) = P_T. \end{aligned} \quad (2)$$

The statistical expectation  $E_{\mathbf{s}, \mathbf{n}} \{ \cdot \}$  is carried out over the data symbols  $\mathbf{s}$  and the noise samples  $\mathbf{n}$ . We assume uncorrelated data symbols of average symbol energy  $E_s$  and zero-mean temporally and spatially white complex Gaussian noise samples with covariance matrix  $\sigma_n^2 \mathbf{I}_{M_R}$ .

We introduce the thin [14, page 72] singular value decomposition (SVD) of the MIMO channel matrix  $\mathbf{H}$ :

$$\mathbf{H} = (\mathbf{U}_p \ \mathbf{U}_{\bar{p}}) \begin{pmatrix} \Sigma_p & \mathbf{0} \\ \mathbf{0} & \Sigma_{\bar{p}} \end{pmatrix} (\mathbf{V}_p \ \mathbf{V}_{\bar{p}})^H, \quad (3)$$

where  $\mathbf{U}_p$  and  $\mathbf{V}_p$  are, respectively, the  $M_R \times p$  and  $M_T \times p$  left and right singular vectors associated to the  $p$  strongest singular values or spatial subchannels or modes<sup>2</sup> of  $\mathbf{H}$ , stacked in decreasing order in the  $p \times p$  diagonal matrix  $\Sigma_p$ .  $\mathbf{U}_{\bar{p}}$  and  $\mathbf{V}_{\bar{p}}$  are the left and right singular vectors associated to the remaining  $(\text{Min}(M_R, M_T) - p)$  spatial modes of  $\mathbf{H}$ , similarly stacked in decreasing order in  $\Sigma_{\bar{p}}$ . The optimization problem stated in (2) is solved using the Lagrange multiplier technique which formulates the constrained cost-function as follows:

$$\begin{aligned} C = \text{Min}_{\mathbf{R}, \mathbf{T}} E_{\mathbf{s}, \mathbf{n}} \{ \|\mathbf{s} - (\mathbf{R}\mathbf{H}\mathbf{T}\mathbf{s} + \mathbf{R}\mathbf{n})\|_2^2 \} \\ + \lambda (E_s \cdot \text{trace}(\mathbf{T}\mathbf{T}^H) - P_T), \end{aligned} \quad (4)$$

where  $\lambda$  is the Lagrange multiplier to be calculated to satisfy the transmit power constraint. The optimal linear precoder and decoder pair  $\{\mathbf{T}, \mathbf{R}\}$ , solution to (4), was shown to be [6]

$$\begin{aligned} \mathbf{T} &= \mathbf{V}_p \cdot \Sigma_T \cdot \mathbf{Z}, \\ \mathbf{R} &= \mathbf{Z}^H \cdot \Sigma_R \cdot (\mathbf{U}_p)^H, \end{aligned} \quad (5)$$

where  $\mathbf{Z}$  is an optional  $p \times p$  unitary matrix,  $\Sigma_T$  is the  $p \times p$  diagonal power allocation matrix that determines the transmit power distribution among the available  $p$  spatial modes

<sup>2</sup>We will alternatively use spatial subchannels and spatial modes to refer to the singular values of  $\mathbf{H}$ , as these singular values represent the parallel independent spatial subchannels or modes underlying the flat-fading MIMO channel modeled by  $\mathbf{H}$ .

and is given by

$$\begin{aligned} \Sigma_T^2 &= \left[ \frac{\sigma_n}{\sqrt{E_s \lambda}} \Sigma_p^{-1} - \frac{\sigma_n^2}{E_s} \Sigma_p^{-2} \right]^+ \\ \text{subject to: } \text{trace}(\Sigma_T^2) &= \frac{P_T}{E_s}, \end{aligned} \quad (6)$$

and  $\Sigma_R$  is the  $p \times p$  diagonal complementary equalization matrix given by

$$\Sigma_R = \frac{\sqrt{E_s \lambda}}{\sigma_n} \Sigma_T. \quad (7)$$

The joint Tx/Rx MMSE design of (5) essentially decouples the MIMO channel matrix  $\mathbf{H}$  into its underlying spatial modes and selects the  $p$  strongest ones, represented by  $\Sigma_p$ , to transmit the  $p$  data streams. Among the latter  $p$  spatial modes, only those above a minimum signal-to-noise ratio (SNR) threshold, determined by the transmit power constraint, are the actually allocated power as indicated by  $[\cdot]^+$  in (6). Furthermore, more power is allocated to the weaker ones in an attempt to balance the SNR levels across spatial modes.

## 2.3. Problem statement

The discussed generic joint Tx/Rx MMSE design has been derived for a given number of spatial streams  $p$  which are arbitrarily chosen and fixed [5, 6, 7, 8, 15]. These  $p$  streams will always be transmitted regardless of the power allocation policy that may, as previously highlighted, allocate no power to certain weak spatial subchannels. The data streams assigned to the latter subchannels are then lost, leading to a poor overall BER performance. Furthermore, as the SNR increases, these initially disregarded modes will eventually be given power and will monopolize most of the available transmit power, leading to an inefficient power allocation strategy that detrimentally impacts the strong modes. Finally, it has been shown [16] that the spatial subchannel gains exhibit decreasing diversity orders. This means that the weakest used subchannel sets the spatial diversity order exploited by the joint Tx/Rx MMSE design. The previous remarks highlight the influence of the choice of  $p$  on the transmit power allocation efficiency, the exhibited spatial diversity order, and thus on the joint Tx/Rx MMSE designs' BER performance. Hence, we alternatively propose to include  $p$  as a design parameter to be optimized according to the available channel knowledge for an improved system BER performance, what we subsequently refer to as *spatial-mode selection*.

## 2.4. State-of-the-art joint Tx/Rx MMSE designs

Before proceeding to derive our spatial-mode selection, we first introduce two state-of-the-art designs that instantiate the aforementioned generic joint Tx/Rx MMSE solution and that are the base line for our subsequent optimization proposal. While preserving the joint Tx/Rx MMSE design's core transmission structure  $\{\Sigma_T, \Sigma_p, \Sigma_R\}$ , these two instantiations implement different unitary matrices  $\mathbf{Z}$ . As will be subsequently shown, the latter unitary matrix can be used to

enforce an additional constraint without altering the resulting system's sum MMSE<sub>p</sub>, formally defined in (2). In order to explicit it, we introduce the MSE covariance matrix  $\mathbf{MSE}_p$ , associated with the considered fixed  $p$  data streams and fixed symbol constellation across these streams, defined as follows:

$$\mathbf{MSE}_p = E_{s,n}\{(\mathbf{s} - \hat{\mathbf{s}})(\mathbf{s} - \hat{\mathbf{s}})^H\}. \quad (8)$$

Clearly, the diagonal elements of  $\mathbf{MSE}_p$  represent the MSEs induced on the individual spatial streams. Consequently, their sum would result in the aforementioned sum MMSE<sub>p</sub> when the optimal linear precoder and decoder pair  $\{\mathbf{T}, \mathbf{R}\}$  of (5) is used. In the latter case,  $\mathbf{MSE}_p$  can be straightforwardly expressed as follows:

$$\mathbf{MSE}_p = \mathbf{Z}^H \cdot \left( E_s (\mathbf{I}_p - \boldsymbol{\Sigma}_T \boldsymbol{\Sigma}_p \boldsymbol{\Sigma}_R)^2 + \sigma_n^2 \boldsymbol{\Sigma}_R^2 \right) \cdot \mathbf{Z}. \quad (9)$$

MMSE<sub>p</sub> is then simply given by [6]

$$\begin{aligned} \text{MMSE}_p \\ = \text{trace} \left[ \mathbf{Z}^H \cdot \left( E_s (\mathbf{I}_p - \boldsymbol{\Sigma}_T \boldsymbol{\Sigma}_p \boldsymbol{\Sigma}_R)^2 + \sigma_n^2 \boldsymbol{\Sigma}_R^2 \right) \cdot \mathbf{Z} \right]. \end{aligned} \quad (10)$$

Since the trace of a matrix depends only on its singular values, the unitary matrix  $\mathbf{Z}$ , indeed, does not alter the MMSE<sub>p</sub> that can be reduced to

$$\text{MMSE}_p = \text{trace} \left( E_s (\mathbf{I}_p - \boldsymbol{\Sigma}_T \boldsymbol{\Sigma}_p \boldsymbol{\Sigma}_R)^2 + \sigma_n^2 \boldsymbol{\Sigma}_R^2 \right). \quad (11)$$

#### 2.4.1. Conventional joint Tx/Rx MMSE design

The conventional<sup>3</sup> joint Tx/Rx MMSE design only aims at minimizing the system's sum MSE. Since, as aforementioned, the unitary matrix  $\mathbf{Z}$  does not alter the system's MMSE<sub>p</sub>, this design simply sets it to identity  $\mathbf{Z} = \mathbf{I}_p$  [6, 7, 8]. Nevertheless, this design exhibits nonequal MSEs across the data streams as pointed out in [7, 15]. Thus, its BER performance will be dominated by the weak modes that induce the largest MSEs. To overcome this drawback, the following design has been proposed.

#### 2.4.2. Even-MSE joint Tx/Rx MMSE design

The even-MSE joint Tx/Rx MMSE design enforces equal MSEs on all data streams while maintaining the same overall sum MMSE<sub>p</sub>. This can be achieved by choosing  $\mathbf{Z}$  as the  $p \times p$  IFFT matrix [15] with  $[\mathbf{Z}]_{n,k} = (1/\sqrt{p}) \exp(j2\pi nk/p)$ . In fact, taking advantage of the diagonal structure of the inner matrix in (9), the pair {IFFT, FFT} enforces equal diagonal elements for  $\mathbf{MSE}_p$ ,<sup>4</sup> what amounts to equal MSEs on all data streams. Through balancing the MSEs across the data streams, this design guarantees equal minimum BER on all

streams for the given fixed number of spatial streams  $p$  and fixed constellation across these streams. Nevertheless, the use of the {IFFT, FFT} pair induces additional interstream interference in the case of the even-MSE design.

### 3. SPATIAL-MODE SELECTION

As previously announced, we aim at a spatial-mode selection criterion that minimizes the system's BER. In order to identify such criterion, we subsequently derive the expression of the conventional joint Tx/Rx MMSE design's average BER and analyze the respective contributions of the individual used spatial modes. To do so, we rewrite the input-output system equation (1) for this design, using the optimal linear precoder and decoder solution of (5) and setting  $\mathbf{Z}$  to identity:

$$\hat{\mathbf{s}} = \boldsymbol{\Sigma}_R \boldsymbol{\Sigma}_p \boldsymbol{\Sigma}_T \mathbf{s} + \boldsymbol{\Sigma}_R \mathbf{n}. \quad (12)$$

Remarkably, the conventional joint Tx/Rx MMSE design transmits the  $p$  available data streams on  $p$  parallel independent channel spatial modes. Each of these spatial modes is simply Gaussian with a fixed gain, given by its corresponding entry in  $\boldsymbol{\Sigma}_p \boldsymbol{\Sigma}_T$ , and an additive noise of variance  $\sigma_n^2$ .<sup>5</sup> Consequently, for the used Gray-encoded square QAM constellation of size  $M_p$  and average transmit symbol energy  $E_s$ , the average BER on the  $i$ th spatial mode, denoted by BER<sub>i</sub>, is approximated at high SNRs (see [17, page 280] and [18, page 409]) by

$$\text{BER}_i \approx \frac{4}{\log_2(M_p)} \cdot \left( 1 - \frac{1}{\sqrt{M_p}} \right) \cdot Q \left( \sqrt{3 \frac{\sigma_i^2 \sigma_{T_i}^2 E_s}{(M_p - 1) \sigma_n^2}} \right), \quad (13)$$

where  $\sigma_i$  denotes the  $i$ th diagonal element of  $\boldsymbol{\Sigma}_p$ , which represents the  $i$ th spatial mode gain. Similarly,  $\sigma_{T_i}$  is the  $i$ th diagonal element of  $\boldsymbol{\Sigma}_T$  whose square designates the transmit power allocated to the  $i$ th spatial mode. Since the used square QAM constellation of size  $M_p$  and minimum Euclidean distance  $d_{\min} = 2$  has an average symbol energy  $E_s = 2(M_p - 1)/3$  and  $Q(x)$  can be conveniently written as  $\text{erfc}(x/\sqrt{2})/2$ , BER<sub>i</sub> can be simplified into

$$\text{BER}_i \approx \frac{2}{\log_2(M_p)} \cdot \left( 1 - \frac{1}{\sqrt{M_p}} \right) \cdot \text{erfc} \left( \sqrt{\frac{\sigma_i^2 \sigma_{T_i}^2}{\sigma_n^2}} \right). \quad (14)$$

The argument  $\sigma_i^2 \sigma_{T_i}^2 / \sigma_n^2$  is easily identified as the average symbol SNR normalized to the symbol energy  $E_s$  on the  $i$ th spatial mode. For a given constellation  $M_p$ , the latter average SNR clearly determines the BER on its corresponding spatial mode. The conventional design's average BER performance,

<sup>3</sup>It is the most wide-spread instantiation in the literature, simply referred to as the joint Tx/Rx MMSE design. The term "conventional" has been added here to avoid confusion with the next instantiation.

<sup>4</sup>The common value of these diagonal elements will be shown later to be equal to the arithmetic average of the diagonal elements of the inner diagonal matrix MMSE<sub>p</sub>/p.

<sup>5</sup>Which is calculated according to the actual  $E_b/N_0$  value.

however, depends on the SNRs on all  $p$  spatial modes as follows:

$$\text{BER}_{\text{conv}} \approx \frac{2}{\log_2(M_p)} \cdot \left(1 - \frac{1}{\sqrt{M_p}}\right) \cdot \frac{1}{p} \sum_{i=1}^p \text{erfc} \left( \sqrt{\frac{\sigma_i^2 \sigma_{T_i}^2}{\sigma_n^2}} \right). \quad (15)$$

Consequently, to better characterize the conventional design's BER, we define the  $p \times p$  diagonal SNR matrix  $\mathbf{SNR}_p$  whose diagonal consists of the average SNRs on the  $p$  spatial modes:

$$\mathbf{SNR}_p = \frac{\boldsymbol{\Sigma}_p^2 \cdot \boldsymbol{\Sigma}_T^2}{\sigma_n^2}. \quad (16)$$

Using the expression of the optimal transmit power allocation matrix  $\boldsymbol{\Sigma}_T^2$  formulated in (6), the previous  $\mathbf{SNR}_p$  expression can be further developed into

$$\mathbf{SNR}_p = \left[ \frac{1}{\sigma_n \sqrt{\lambda E_s}} \boldsymbol{\Sigma}_p - \frac{\mathbf{I}_p}{E_s} \right]^+. \quad (17)$$

The latter expression illustrates that the conventional joint Tx/Rx MMSE design induces uneven SNRs on the different  $p$  spatial streams. More importantly, (17) shows that the weaker the spatial mode is, the lower its experienced SNR is. The conventional joint Tx/Rx MMSE BER,  $\text{BER}_{\text{conv}}$ , of (15) can be rewritten as follows:

$$\text{BER}_{\text{conv}} \approx \frac{2}{\log_2(M_p)} \cdot \left(1 - \frac{1}{\sqrt{M_p}}\right) \cdot \frac{1}{p} \sum_{i=1}^p \text{erfc} \left( \sqrt{[\mathbf{SNR}_p]_{ii}} \right). \quad (18)$$

The previous SNR analysis further indicates that the  $p$  spatial modes exhibit uneven BER contributions and that of the weakest  $p$ th mode, corresponding to the lowest SNR  $[\mathbf{SNR}_p]_{p,p}$ , dominates  $\text{BER}_{\text{conv}}$ . Consequently, in order to minimize  $\text{BER}_{\text{conv}}$ , we propose as the optimal number of streams to be used  $p_{\text{opt}}$ , the one that maximizes the SNR on the weakest used mode under a fixed rate  $R$  constraint. The latter proposed spatial-mode selection criterion can be expressed as follows:

$$\begin{aligned} & \text{Max}_p [\mathbf{SNR}_p]_{p,p} \\ & \text{subject to: } p \times \log_2(M_p) = R. \end{aligned} \quad (19)$$

The rate constraint shows that, though the same symbol constellation is used across spatial streams, the selection/adaptation of the optimal number of streams  $p_{\text{opt}}$  requires the joint selection/adaptation of the used constellation size such that  $M_{\text{opt}} = 2^{R/p_{\text{opt}}}$ . Adapting (17) for the considered square QAM constellations (i.e.,  $E_s = 2(M_p - 1)/3$ ), the

spatial-mode selection criterion stated in (19) can be further refined into

$$p_{\text{opt}} = \arg \text{Max}_p \left[ \frac{1}{\sigma_n \sqrt{(2/3)(2^{R/p} - 1)\lambda}} \sigma_p - \frac{1}{(2/3)(2^{R/p} - 1)} \right]^+. \quad (20)$$

The latter spatial-mode selection problem has to be solved for the current channel realization to identify the optimal pair  $\{p_{\text{opt}}, M_{\text{opt}}\}$  that minimizes the system's average BER,  $\text{BER}_{\text{conv}}$ .

We have derived our spatial-mode selection based on the conventional joint Tx/Rx MMSE design because this design represents the core transmission structure on which the even-MSE design is based. Our strategy is to first use our spatial-mode selection to optimize the core transmission structure  $\{\boldsymbol{\Sigma}_T, \boldsymbol{\Sigma}_{p_{\text{opt}}}, \boldsymbol{\Sigma}_R\}$ , the even-MSE, then additionally applies the unitary matrix  $\mathbf{Z}$ , which is now the  $p_{\text{opt}} \times p_{\text{opt}}$  IFFT matrix to further balance the MSEs and the SNRs across the used  $p_{\text{opt}}$  spatial streams.

#### 4. PERFORMANCE ANALYSIS

In this section, we investigate the uncoded and coded BER performance of both conventional and even-MSE joint Tx/Rx MMSE designs when our spatial-mode selection is applied. The goal is manifold. We first assess the BER performance improvement offered by our spatial-mode selection over state-of-the-art full SM conventional and even-MSE joint Tx/Rx MMSE designs. Then, we compare our spatial-mode selection performance and complexity to those of a practical spatial adaptive loading strategy. Last but not least, we evaluate the impact of channel coding on the relative BER performances of all the above-mentioned designs. In all the following, the MIMO channel is stationary Rayleigh flat-fading, modeled by an  $M_R \times M_T$  matrix with i.i.d unit-variance zero-mean complex Gaussian entries. In all the following, the BER figures are averaged over 1000 channel realizations for the uncoded performance and over 100 channels for the coded performance. For each channel, at least 10 bit errors were counted for each  $E_b/N_0$  value, where  $E_b/N_0$  stands for the average receive energy per bit over noise power. A unit average total transmit power was considered,  $P_T = 1$ .

##### 4.1. Uncoded performance

Considering the uncoded system, we first compare the relative BER performance of the conventional and even-MSE joint Tx/Rx MMSE designs when full SM is used. We later apply our spatial-mode selection for improved BER performances, which we further contrast with that of a practical spatial adaptive loading scheme inspired from [19].

##### 4.1.1. Conventional versus even-MSE joint Tx/Rx MMSE

For a fixed number of spatial streams  $p$  and fixed symbol constellation  $M_p$ ,  $\text{BER}_{\text{conv}}$  given by (15) approximates the

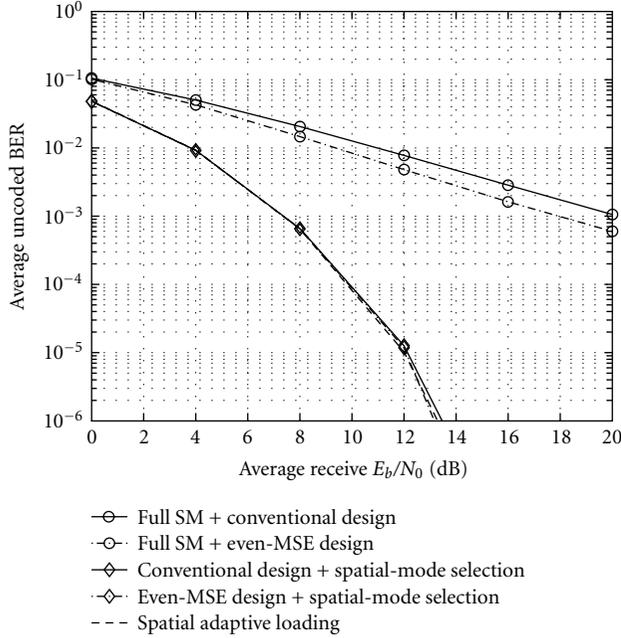


FIGURE 2: Average uncoded BER comparison for a (2, 2) MIMO setup at  $R = 4$  bps/Hz.

conventional joint Tx/Rx MMSE design BER performance in the high SNR region, where the MMSE receiver reduces to a zero-forcing receiver. Associated to this assumption, the conventional design approximately reduces the  $i$ th spatial mode into a Gaussian channel with noise variance equal to  $\sigma_n^2/\sigma_i^2\sigma_{T_i}^2$ . The latter noise variance represents also the equivalent MSE at the output of the  $i$ th spatial mode, which can be denoted by  $[\mathbf{MSE}_p]_{i,i} = 1/[\mathbf{SNR}_p]_{i,i}$ . Hence, using the same zero-forcing assumption, the even-MSE enforces an equal MSE or noise variance across  $p$  streams equal to  $\sum_{i=1}^p(\sigma_n^2/\sigma_i^2\sigma_{T_i}^2)/p = \sum_{i=1}^p(1/[\mathbf{SNR}_p]_{i,i})/p$ ; thus its average BER,  $\text{BER}_{\text{even-MSE}}$ , is approximately given by

$$\text{BER}_{\text{even-MSE}} \approx \frac{2}{\log_2(M_p)} \cdot \left(1 - \frac{1}{\sqrt{M_p}}\right) \text{erfc} \left( \sqrt{\frac{p}{\sum_{i=1}^p 1/[\mathbf{SNR}_p]_{i,i}}} \right). \quad (21)$$

Recalling Jensen's inequality [20, page 25] and the comparison of (18) and (21) where the MSEs ( $[\mathbf{MSE}_p]_{i,i} = 1/[\mathbf{SNR}_p]_{i,i}$ ) would be denoted as variable  $(x_i)_i$ , we can state that

$$\text{BER}_{\text{even-MSE}} \leq \text{BER}_{\text{conv}} \quad (22)$$

when  $f_p(x) = \text{erfc}(1/\sqrt{x})$  is convex. The analysis of the function  $\{f_p(x), x \geq 0\}$ , provided in Appendix A, shows that it is convex for values of  $x$  smaller than a certain  $x_{\text{inf}}$ ; for  $x$  larger than  $x_{\text{inf}}$ , the function turns out to be concave. Since  $x$  stands for the MSEs on the spatial modes, which decrease when the average receive energy per bit over noise power

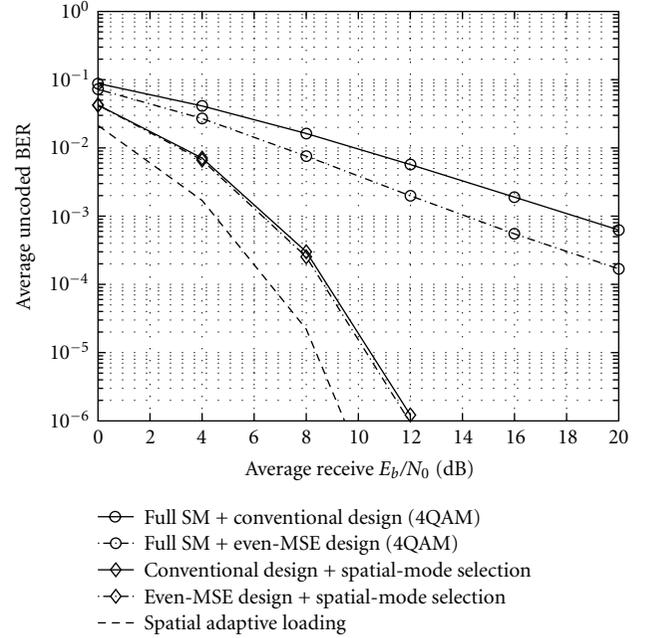


FIGURE 3: Average uncoded BER comparison for a (3, 3) MIMO setup at  $R = 6$  bps/Hz.

( $E_b/N_0$ ) increases, we can relate the convexity of  $f_p(x)$  to the relative BER performance of the conventional and the even-MSE joint Tx/Rx MMSE designs as follows:

$$\text{BER}_{\text{even-MSE}} \leq \text{BER}_{\text{conv}} \quad (23)$$

for  $E_b/N_0 \geq E_b/N_{0\text{inf}} (\text{MSE}_e \leq \text{MSE}_{\text{inf}})$ .

$E_b/N_{0\text{inf}}$  is the  $E_b/N_0$  value needed to reach  $f_p(x)$ 's inflection point  $x_{\text{inf}} = \text{MSE}_{\text{inf}}$ . This BER analysis is further confirmed by the simulated results plotted in Figures 2, 3, and 4. More specifically, the latter figures illustrate that the full SM even-MSE outperforms the full SM conventional design after a certain  $E_b/N_0$  value, previously referred to as  $E_b/N_{0\text{inf}}$ . As it turns out, the latter value occurs before 0 dB for both the (2, 2) MIMO setup at  $R = 4$  bps/Hz and the (3, 3) MIMO setup at  $R = 6$  bps/Hz, respectively, plotted in Figures 2 and 3. For the case of the (3, 3) MIMO setup at  $R = 12$  bps/Hz of Figure 4, however, the even-MSE design surpasses the conventional design only for SNRs larger than  $E_b/N_{0\text{inf}} = 10$  dB. This is due to the fact that, for a given  $(M_T, M_R)$  MIMO system with fixed average total transmit power  $P_T$ , the larger the constellation used and the larger the rate supported, the larger the induced MSEs at a given  $E_b/N_0$  value or alternatively the larger the  $E_b/N_{0\text{inf}}$  needed to fall below  $\text{MSE}_{\text{inf}}$  on the used spatial streams, which is required for the even-MSE design to outperform the conventional one.

#### 4.1.2. Spatial-mode selection versus full spatial multiplexing

Applying our spatial-mode selection to both joint Tx/Rx MMSE designs leads to impressive BER performance

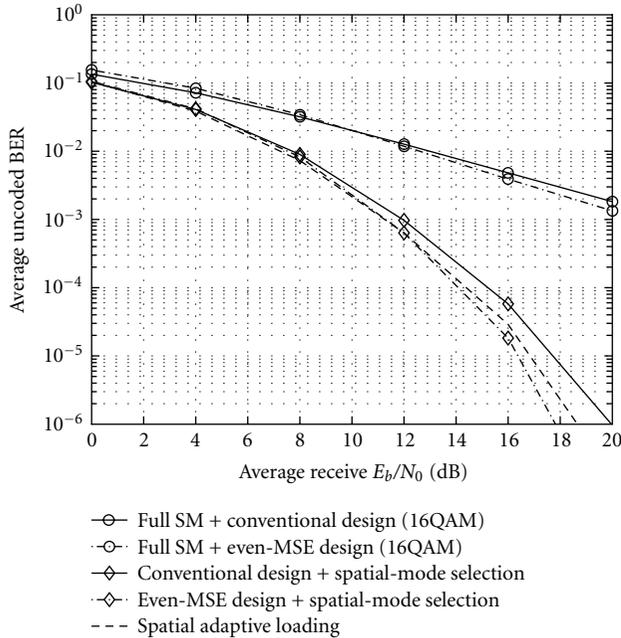


FIGURE 4: Average uncoded BER comparison for a (3,3) MIMO setup at  $R = 12$  bps/Hz.

improvement for various MIMO system dimensions and parameters. Figure 2 illustrates such BER improvement for the case of a (2,2) MIMO setup supporting a spectral efficiency  $R = 4$  bps/Hz. Our proposed spatial-mode selection is shown to provide 12.6 dB and 10.5 dB SNR gain over full SM conventional and even-MSE designs, respectively, at  $\text{BER} = 10^{-3}$ . Figures 3 and 4 confirm similar gains for a (3,3) MIMO setup at spectral efficiency  $R = 6$  bps/Hz and  $R = 12$  bps/Hz, respectively. These significant performance improvements are due to the fact that our spatial-mode selection, depending on the spectral efficiency  $R$ , wisely discards a number of weak spatial modes that exhibit the lowest spatial diversity orders, as argued in [16]. The same weak modes that dominate the performance of both full SM joint Tx/Rx MMSE designs. According to (20), our spatial-mode selection restricts transmission to the  $p_{\text{opt}}$  strongest modes only. The latter  $p_{\text{opt}}$  modes exhibit significantly higher spatial diversity orders and form a more balanced subset<sup>6</sup> over which a more efficient power allocation is possible, leading to higher transmission SNR levels and consequently lower BER figures. Furthermore, it is because the subset of  $p_{\text{opt}}$  selected modes is balanced that the additional effort of the even-MSE joint Tx/Rx MMSE to further average it brings only marginal BER improvement over the conventional joint Tx/Rx MMSE when spatial-mode selection is applied. Clearly, the proposed spatial-mode selection enables a more efficient transmit power allocation and a better exploitation of the available spatial diversity.

<sup>6</sup>The difference between the  $p_{\text{opt}}$  spatial mode gains is reduced.

#### 4.1.3. Spatial-mode selection versus spatial adaptive loading

The spatial adaptive loading, herein considered, is simply the practical Fischer's adaptive loading algorithm [19]. The latter algorithm was initially proposed for multicarrier systems. Nevertheless, it directly applies for a MIMO system where an SVD is used to decouple the MIMO channel into parallel independent spatial modes, which are completely analogous to the orthogonal carriers of a multicarrier system. Hence, the considered spatial adaptive loading setup first performs an SVD that decouples the MIMO channel into parallel independent spatial modes. Fischer's adaptive loading algorithm [19] is then used to determine, using the knowledge of the current channel realization, the optimal assignment for the  $R$  bits on the decoupled spatial modes such that equal minimum symbol-error rate (SER) is achieved on the used modes. Consequently, strong spatial modes are loaded with large constellation sizes, whereas weak modes carry small constellation sizes or are dropped if their gains are below a given threshold. This scheme, indeed, exhibits excellent performance, as shown in Figures 2, 3, and 4, mostly outperforming both joint Tx/Rx MMSE designs even when spatial-mode selection is used. This is due to spatial adaptive loading's additional flexibility of assigning different constellation sizes to different spatial modes. This higher flexibility, however, entails a higher complexity and signaling overhead, as later on highlighted.

When the spectral efficiency is low and there is major discrepancy between available spatial modes, as occurs between the two spatial modes of a (2,2) MIMO system [16], both spatial adaptive loading and spatial-mode selection in conjunction with joint Tx/Rx MMSE designs converge to the same solution, basically single-mode transmission or max-SNR solution [21], as illustrated in Figure 2. Figure 3 illustrates the case of a (3,3) MIMO system when the spectral efficiency is low  $R = 6$  bps/Hz. In this case, the two first channel singular values corresponding to the two strongest spatial modes out of the three available spatial modes have relatively close diversity orders and close gains [16]. Consequently, spatial adaptive loading can optimally distribute the available  $R = 6$  bits between these two strongest modes while using a lower constellation on the second mode to reduce its impact on the BER, whereas spatial-mode selection has to stick to the single-mode transmission with 64 QAM to avoid the weak third mode that would be used by the next possible constellation (4 QAM<sup>7</sup> over all three spatial streams). In this case, spatial-mode selection suffers an SNR penalty of 2 dB compared to spatial adaptive loading at  $\text{BER} = 10^{-3}$ . When the spectral efficiency is further increased to  $R = 12$  bps/Hz, spatial adaptive loading's flexibility margin is reduced and so is its SNR gain over spatial-mode selection, which is now only 0.7 dB at  $\text{BER} = 10^{-3}$  for the conventional joint Tx/Rx MMSE design, as shown in Figure 4.

<sup>7</sup>8 QAM is excluded since, for all designs considered in this contribution, only square QAM constellations {4 QAM, 16 QAM, 64 QAM} have been allowed.

Furthermore, the even-MSE design, when spatial-mode selection is applied, even outperforms spatial adaptive loading for high SNRs. The latter result is related to these two designs' BER minimization strategies. On the one hand, the even-MSE joint Tx/Rx MMSE design guarantees equal minimum MSEs on each stream and hence equal minimum SER and BER since the same constellation is used across streams. On the other hand, spatial adaptive loading enforces equal minimum SER across streams; the BERs on the latter streams, however, are not equal since they bear different constellations. Thus, the weak modes, carrying small constellations, exhibit higher BERs. The latter imbalance explains the fact that the even-MSE design surpasses spatial adaptive loading when spatial-mode selection is applied. For target high data-rate SM systems, the latter regime is particularly relevant and our spatial-mode selection was shown to tightly approach spatial-adaptive-loading optimal BER performance while exhibiting lower complexity and adaptation requirements. The comparison of the complexity required by our spatial-mode selection to that of spatial adaptive loading, assessed in [22, page 67], shows that both techniques exhibit similar complexities when the available number of modes or subchannels is small. When the number of modes increases,<sup>8</sup> however, spatial adaptive loading requires an increased number of iterations to reach the final bits assignment, and consequently, its complexity significantly outgrows that of our spatial-mode selection. More importantly, adaptive loading requires the additional flexibility of assigning different constellations sizes to different modes, whereas our spatial-mode selection assumes a single constellation across modes. This higher flexibility comes at the cost of a higher signaling overhead between the transmitter and receiver.

#### 4.2. Coded performance

In Section 4.1, we established our spatial-mode selection as a diversity technique that successfully exploits the spatial diversity available in MIMO channels to improve the performance of state-of-the-art joint Tx/Rx MMSE designs. In a practical wireless communication system, however, it will not be the only such diversity technique to be present. Indeed, channel coding will also be used, together with the latter state-of-the-art designs, to exploit the same spatial diversity. Therefore, in this section, we undertake a coded system performance analysis to confirm that our spatial-mode selection remains advantageous over the state-of-the-art full SM approach when channel coding is present. We further verify whether our conclusions, concerning the relative performance of all previously discussed schemes, are still valid. We consider a bit-interleaved coded modulation (BICM) system, as shown in Figure 1, with a rate-1/2 convolutional encoder with constraint length  $K = 7$ , generator polynomials  $[133_8, 171_8]$ ,<sup>9</sup> and optimum maximum likelihood sequence estimation (MLSE) decoding using the Viterbi decoder [23].

<sup>8</sup>For instance, when both techniques are applied for multicarrier MIMO systems in presence of frequency-selective fading.

<sup>9</sup>The industry-standard convolutional encoder used in both IEEE 802.11a and ETSI Hiperlan II indoor wireless LAN standards.

##### 4.2.1. Conventional versus even-MSE joint Tx/Rx MMSE

To gain some insight into both designs' coded performances, we derive the *equivalent additive white Gaussian noise (AWGN) channel model* describing the output of the linear equalizer  $\mathbf{R}$  for each of the two designs. Such a model highlights the diversity branches available at the input of the Viterbi decoder and hence the achievable spatial diversity for the corresponding joint Tx/Rx MMSE design. Furthermore, it was used to calculate the bit log-likelihood ratios (LLR), which form the soft inputs for soft-decision Viterbi decoding as in [24].

The output of the linear equalizer  $\mathbf{R}$  for the conventional joint Tx/Rx MMSE design is described in (12). Accordingly, the detected symbol  $\hat{s}_i$  on the  $i$ th spatial mode can be expressed as the output of an equivalent AWGN channel having  $s_i$  as its input:

$$\hat{s}_i = \underbrace{\sigma_{R_i}\sigma_i\sigma_{T_i}}_{\mu_{\text{conv}_i}} s_i + \sigma_{R_i}n_i. \quad (24)$$

The latter equivalent AWGN channel is described by a gain  $\mu_{\text{conv}_i}$  and a zero-mean white complex Gaussian noise of variance  $\sigma_{R_i}^2\sigma_n^2$ . Similarly, the AWGN channel equivalent model for the even-MSE design can be shown to be (See Appendix B)

$$\hat{s}_i = \frac{1}{P} \left( \sum_{i=1}^P \sigma_{R_i}\sigma_i\sigma_{T_i} \right) s_i + \eta_i, \quad (25)$$

where  $\eta_i$  stands for the equivalent zero-mean white complex Gaussian noise of variance  $\sigma_\eta^2$ . In this case, however, the latter equivalent noise contains, in addition to scaled receiver noise, interstream interference induced by the use of the {IFFT, FFT} pair. The equivalent noise variance  $\sigma_\eta^2$  was found to be (See Appendix B)

$$\sigma_\eta^2 = \underbrace{\frac{\sigma_n^2}{P} \sum_{i=1}^P \sigma_{R_i}^2}_{\text{noise contribution}} + \underbrace{\frac{E_s}{P^2} \left( P \sum_{i=1}^P \mu_{\text{conv}_i}^2 - \left( \sum_{i=1}^P \mu_{\text{conv}_i} \right)^2 \right)}_{\text{interstream interference contribution}}. \quad (26)$$

Clearly, the conventional joint Tx/Rx MMSE design provides symbol estimates  $(\hat{s}_i)_{1 \leq i \leq p}$ , and consequently coded bits, that experienced independently fading channels with different diversity orders, which enables the channel coding to exploit the system's spatial diversity, whereas the even-MSE design, through the use of {IFFT, FFT}, creates an equivalent average channel for all  $p$  spatial streams, as shown in (25) and (26). Consequently, the even-MSE design prohibits the channel coding from any diversity combining and only allows for coding gain. In other words, the coded even-MSE design exhibits the same diversity order as the uncoded one. The latter diversity order is the one exhibited, at high  $E_b/N_0$ , by the

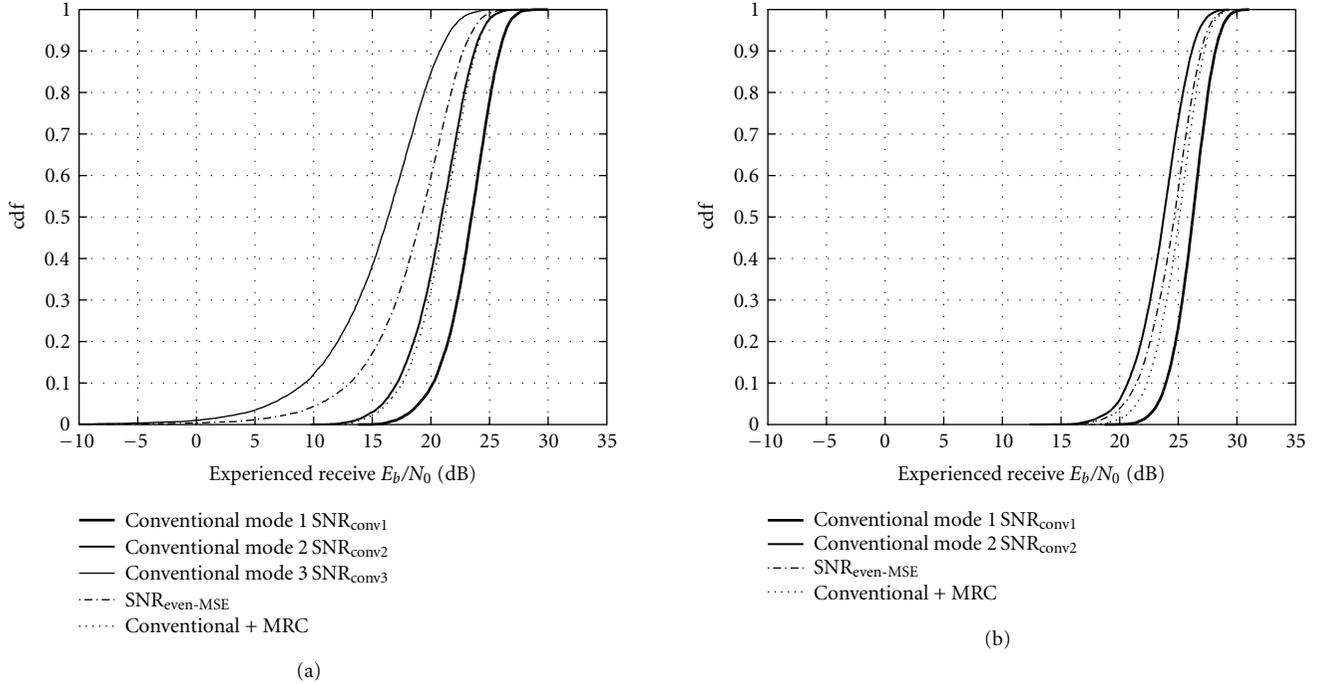


FIGURE 5: Comparison of the diversity orders exhibited by the spatial modes for (a) full SM and (b) spatial-mode selection for a (3, 3) MIMO setup at  $R = 12$  bps/Hz and average receive  $E_b/N_0 = 20$  dB. Conventional mode 3  $\text{SNR}_{\text{conv}3}$  does not appear in (b).

average<sup>10</sup> received bit SNR on the  $p$  spatial streams. At high  $E_b/N_0$ , the MMSE receiver  $\Sigma_R$  reduces to a zero-forcing receiver equal to  $\Sigma_T^{-1}\Sigma_p^{-1}$ . In that case, the average received bit SNR on the  $p$  spatial streams, denoted as  $\text{SNR}_{\text{even-MSE}}$ , can be defined as follows:

$$\text{SNR}_{\text{even-MSE}} = \frac{E_s/\log_2(M_p)}{\sigma_n^2}, \quad (27)$$

where  $\sigma_n^2$  is the asymptotic equivalent noise variance equal to  $(\sigma_n^2/p) \sum_{i=1}^p 1/\sigma_i^2 \sigma_{T_i}^2$ , corresponding to the evaluation of (26) at high  $E_b/N_0$ . Consequently,  $\text{SNR}_{\text{even-MSE}}$  can be developed into

$$\text{SNR}_{\text{even-MSE}} = \frac{p}{\sum_{i=1}^p 1/\sigma_i^2 \sigma_{T_i}^2} \cdot \frac{E_s/\log_2(M_p)}{\sigma_n^2}. \quad (28)$$

The previous  $\text{SNR}_{\text{even-MSE}}$  statistics should be contrasted with those of the average received SNRs on the  $p$  parallel modes of the conventional joint Tx/Rx MMSE design, denoted as  $(\text{SNR}_{\text{conv}i})_i$ . Based on (24), the latter received SNRs are simply given by

$$\text{SNR}_{\text{conv}i} = \sigma_i^2 \sigma_{T_i}^2 \cdot \frac{E_s/\log_2(M_p)}{\sigma_n^2} \quad (1 \leq i \leq p). \quad (29)$$

Furthermore, the spatial diversity exhibited by  $\text{SNR}_{\text{even-MSE}}$  should also be compared to the maximum spatial diversity

order achievable by channel coding,<sup>11</sup> given by maximum-ratio combining (MRC) across the conventional design's  $p$  spatial modes. Since the latter  $p$  spatial modes can be considered independent diversity paths of SNRs  $(\text{SNR}_{\text{conv}i})_i$ , the aforementioned maximum achievable spatial diversity order is described by the statistics of  $\text{SNR}_{\text{MRC}}$  [17, page 780]:

$$\text{SNR}_{\text{MRC}} = \sum_{i=1}^p \sigma_i^2 \sigma_{T_i}^2 \cdot \frac{E_s/\log_2(M_p)}{\sigma_n^2}. \quad (30)$$

Figure 5 provides such a spatial diversity comparison, as it plots the cumulative probability density functions (cdf) of (28), (29), and (30) for a full SM (3, 3) MIMO setup at spectral efficiency  $R = 12$  bps/Hz and average receive  $E_b/N_0 = 20$  dB. The steeper the SNR's cdf is, the higher the diversity order of the corresponding spatial mode or design is. Consequently, Figure 5 confirms the decreasing diversity orders of the conventional design's  $p$  spatial modes. More importantly, it shows that the diversity order exhibited by the even-MSE design is closer to that of the weakest spatial mode, which obviously dominates the even-MSE design's equivalent channel of (25). The even-MSE design's diversity order is also lower than the diversity order achievable by the conventional design when channel coding is applied. The latter observation

<sup>10</sup>Carried out over data symbols and noise samples.

<sup>11</sup>It is assumed that channel coding is able to exploit all the available spatial diversity, based on the assumption that the code's free distance  $d_{\min}$  is large enough [17, page 812]. The latter assumption is fulfilled for the considered (3, 3) MIMO system and convolutional code  $d_{\min} = 10$  [17, page 493].

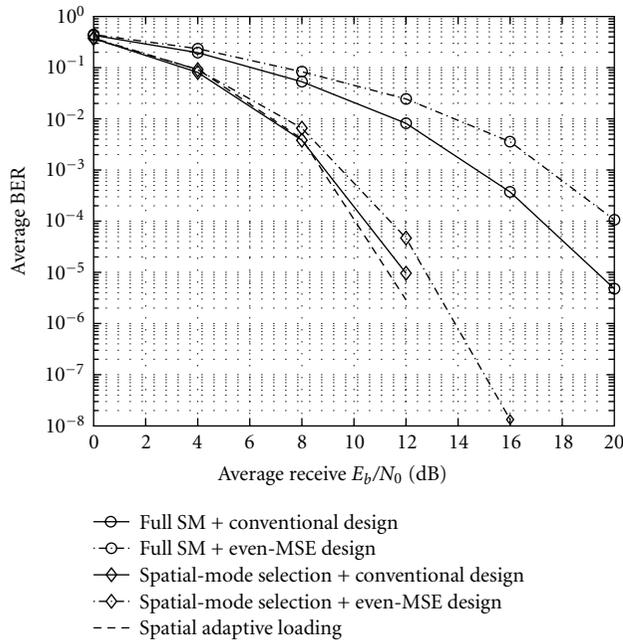


FIGURE 6: Average coded BER comparison for a (3, 3) MIMO setup and  $R = 6$  bps/Hz with hard-decision decoding.

explains the coded BER results of Figures 6 and 7 where, contrarily to the uncoded system, the full SM conventional design now significantly outperforms the SM even-MSE design. Furthermore, comparing Figures 3, 6, and 7 confirms that channel coding, as previously argued, does not improve on the spatial diversity exploited by the even-MSE design, whereas it does significantly improve the performance of the conventional design through exploiting the different diversity branches this design provides.

#### 4.2.2. Spatial-mode selection versus full spatial multiplexing

Figure 5 further depicts the evolution of the previous spatial diversity comparison when our spatial-mode selection is applied. Clearly, only the two highest diversity spatial modes are selected for transmission. As previously explained, these two strong modes form a more balanced subset on which a more efficient power allocation is possible and consequently larger experienced SNR values on the spatial modes are achieved. Moreover, since the weakest mode has been discarded, the even-MSE design now averages the two strongest spatial modes and obviously exhibits a higher equivalent diversity order. However, the latter diversity order is still lower than that achievable through channel coding across the conventional design's two parallel spatial modes. Hence, the coded conventional design still outperforms the coded even-MSE when our spatial-mode selection is applied, as illustrated in Figures 6 and 7. More importantly, our spatial-mode selection still significantly improves the performance of both joint Tx/Rx MMSE designs in presence of channel coding. Figures 6 and 7 report 6 dB and 3.5 dB SNR gains at  $\text{BER} = 10^{-3}$ , respectively, for hard- and soft-decision decoding provided

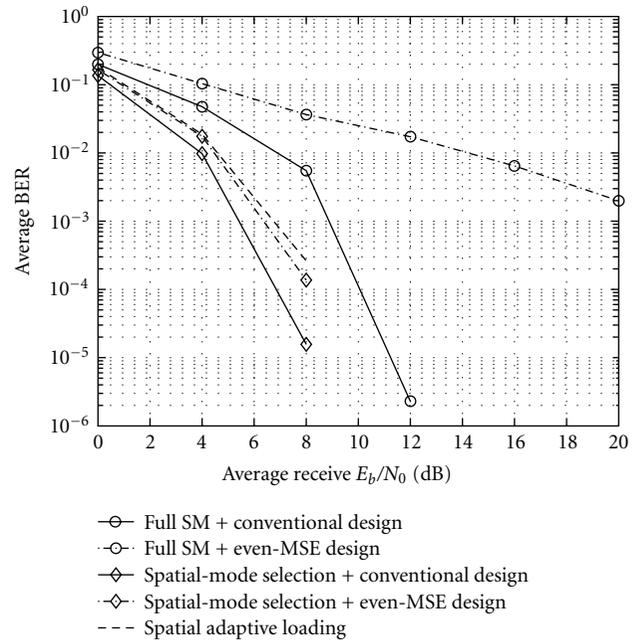


FIGURE 7: Average coded BER comparison for a (3, 3) MIMO setup and  $R = 6$  bps/Hz with soft-decision decoding.

by our spatial-mode selection over full SM for the conventional design. The gains are more dramatic for the even-MSE design, as channel coding is prohibited to access the spatial diversity in the full SM case.

#### 4.2.3. Spatial-mode selection versus spatial adaptive loading

Although our spatial-mode selection significantly improves the BER performance of the uncoded conventional joint Tx/Rx MMSE design, the latter design performance will always be dominated by the weakest mode among the  $p_{\text{opt}}$  selected ones. The latter remark explains the better BER performances of both even-MSE design and spatial adaptive loading in Figure 3. Channel coding and interleaving mitigate this problem as they spread each information bit over several coded bits that are transmitted on all  $p_{\text{opt}}$  spatial modes and eventually optimally combined before detection. Consequently, channel coding suppresses the SNR gap previously observed between the conventional design and spatial adaptive loading, as illustrated in Figure 6. Soft-decision decoding is shown in Figure 7 to further favor the conventional joint Tx/Rx MMSE design as it is the design that provides the more diversity branches at the output of the equalizer  $\mathbf{R}$ . This is because spatial adaptive loading, in order to achieve equal SER across used spatial modes, enforces equal SNR across the latter modes which reduces the equivalent spatial diversity branches it provides to the Viterbi decoder.

## 5. CONCLUSIONS

In this paper, we proposed a novel selection-diversity technique, so-called *spatial-mode selection*, that optimally selects

the number of spatial streams used by the spatial multiplexing joint Tx/Rx MMSE design in order to minimize the system's BER. We assessed the significant improvement in BER performance that our spatial-mode selection provides over the two state-of-the-art full SM joint Tx/Rx MMSE designs, namely, the conventional and even-MSE. Such significant improvements were shown to be due to the more efficient transmit power allocation and the better exploitation of the available spatial diversity achieved by our spatial-mode selection. Furthermore, when our spatial-mode selection is applied, both conventional and even-MSE designs were shown to tightly approach the optimal performance of spatial adaptive loading while exhibiting lower complexity and signaling overhead requirements. Finally, we confirmed that our spatial-mode selection is still advantageous when channel coding is present in the system.

## APPENDICES

### A. CONVEXITY ANALYSIS OF $f_p(x) = \text{erfc}(1/\sqrt{x})$

The function  $f_p(x) = \text{erfc}(1/\sqrt{x})$  for  $x \geq 0$  is explicitly defined as follows:

$$f_p(x) = \frac{2}{\sqrt{\pi}} \int_{1/\sqrt{x}}^{+\infty} \exp(-t^2) dt. \quad (\text{A.1})$$

To determine the convexity of the latter function, we need to evaluate the sign of its second derivative  $f_p''(x)$  for  $x \geq 0$ . To do so, we first calculate the first derivative  $f_p'(x) = d/dx[f_p(x)]$ . For that, we use the identity provided in [25, page 275], which differentiates an integral of the form  $\int_{u(x)}^{v(x)} f(x, t) dt$  with respect to  $x$  as follows:

$$\begin{aligned} \frac{\partial}{\partial x} \int_{u(x)}^{v(x)} f(x, t) dt &= v'(x) f(x, v(x)) - u'(x) f(x, u(x)) \\ &+ \int_{u(x)}^{v(x)} \frac{\partial}{\partial x} f(x, t) dt. \end{aligned} \quad (\text{A.2})$$

Accordingly, the first derivative  $f_p'(x)$  can be easily shown to be

$$f_p'(x) = \frac{1}{\sqrt{\pi}} x^{-3/2} \exp\left(-\frac{1}{x}\right). \quad (\text{A.3})$$

The second derivative  $f_p''(x) = d/dx[f_p'(x)]$  can then be straightforwardly expressed as follows:

$$f_p''(x) = \frac{1}{\sqrt{\pi}} \left(-\frac{3}{2} + \frac{1}{x}\right) x^{-5/2} \exp\left(-\frac{1}{x}\right). \quad (\text{A.4})$$

Consequently, the sign of  $f_p''(x)$  for  $x \geq 0$  is solely determined by the sign of  $(-3/2 + 1/x)$  for  $x \geq 0$ . Accordingly,  $f_p(x)$  is convex ( $f_p''(x) \geq 0$ ) when  $x \leq 3/2$ , whereas it is concave ( $f_p''(x) \leq 0$ ) for  $x \geq 3/2$ .

### B. DERIVATION OF (25) AND (26)

First, we instantiate the input-output system (1) for the even-MSE design using the optimal linear precoder and decoder solution of (5), where  $\mathbf{Z}$  is the  $p \times p$  IFFT matrix with  $\{[\mathbf{Z}]_{n,k} = (1/\sqrt{p}) \exp(j2\pi nk/p); 0 \leq k, n \leq (p-1)\}$ , as follows:

$$\hat{\mathbf{s}} = \mathbf{Z}^H \cdot \boldsymbol{\Sigma}_R \boldsymbol{\Sigma}_p \boldsymbol{\Sigma}_T \cdot \mathbf{Z} + \mathbf{Z}^H \cdot \boldsymbol{\Sigma}_R \mathbf{n}. \quad (\text{B.1})$$

As earlier mentioned, taking advantage of the diagonal structure of the inner matrix  $\boldsymbol{\Sigma}_R \boldsymbol{\Sigma}_p \boldsymbol{\Sigma}_T$ , the {IFFT, FFT} pair enforces equal diagonal elements for  $\mathbf{Z}^H \cdot \boldsymbol{\Sigma}_R \boldsymbol{\Sigma}_p \boldsymbol{\Sigma}_T \cdot \mathbf{Z}$ . Since the {IFFT, FFT} pair is unitary, the trace  $\mathbf{Z}^H \cdot \boldsymbol{\Sigma}_R \boldsymbol{\Sigma}_p \boldsymbol{\Sigma}_T \cdot \mathbf{Z}$  is the trace of the inner diagonal matrix. Consequently, the diagonal elements of  $\mathbf{Z}^H \cdot \boldsymbol{\Sigma}_R \boldsymbol{\Sigma}_p \boldsymbol{\Sigma}_T \cdot \mathbf{Z}$  are equal to  $\sum_{i=1}^p \sigma_{R_i} \sigma_i \sigma_{T_i} / p$ . Hence, the input-output equation (B.1) can be simply developed into

$$\begin{aligned} \begin{bmatrix} \hat{s}_1 \\ \vdots \\ \hat{s}_p \end{bmatrix} &= \frac{1}{p} \sum_{i=1}^p \sigma_{R_i} \sigma_i \sigma_{T_i} \begin{bmatrix} s_1 \\ \vdots \\ s_p \end{bmatrix} \\ &+ \begin{bmatrix} \mathbf{Z}_{:,0}^H \cdot \boldsymbol{\Sigma}_R \boldsymbol{\Sigma}_p \boldsymbol{\Sigma}_T \cdot [\mathbf{Z}]_{:,1:(p-1)} \cdot \mathbf{s}_{1:(p-1)} \\ \vdots \\ [\mathbf{Z}]_{:,p-1}^H \cdot \boldsymbol{\Sigma}_R \boldsymbol{\Sigma}_p \boldsymbol{\Sigma}_T \cdot [\mathbf{Z}]_{:,0:(p-2)} \cdot \mathbf{s}_{0:(p-2)} \end{bmatrix} \\ &+ \mathbf{Z}^H \cdot \boldsymbol{\Sigma}_R \begin{bmatrix} n_1 \\ \vdots \\ n_p \end{bmatrix}. \end{aligned} \quad (\text{B.2})$$

The last two terms, respectively, represent the interstream interference caused by the {IFFT, FFT} pair and the AWGN resulting from the unitary filtering of the receiver noise. To draw the equivalent AWGN channel model of the even-MSE design, these two terms are merged into a single term, denoted  $\boldsymbol{\eta}$ , approximated [24] as a zero-mean white Gaussian noise vector of variance  $\sigma_{\boldsymbol{\eta}}^2$ . Accordingly, the even-MSE design's AWGN-channel equivalent model can be drawn as follows:

$$\hat{\mathbf{s}} = \frac{1}{p} \sum_{i=1}^p \sigma_{R_i} \sigma_i \sigma_{T_i} \mathbf{s} + \boldsymbol{\eta}. \quad (\text{B.3})$$

The evaluation of the previous model for the  $i$ th spatial stream leads to (25). We now calculate the equivalent noise variance  $\sigma_{\boldsymbol{\eta}}^2$ . First, using the statistical independence of the elements of  $\mathbf{n}$  and the effect of the {IFFT, FFT} pair on inner diagonal matrices, it can be easily shown that the filtered noise term of (B.2) has a covariance matrix  $\sigma_n^2 \sum_{i=1}^p \sigma_{R_i}^2 / p \cdot \mathbf{I}_p$ . Second, recalling the Vandermonde structure of  $\mathbf{Z}$  and the fact that for all  $\{k, n\} : [\mathbf{Z}]_{n,k}^p = 1$ , we can show that

$$\begin{aligned} &[\mathbf{Z}]_{:,0}^H \cdot \boldsymbol{\Sigma}_R \boldsymbol{\Sigma}_p \boldsymbol{\Sigma}_T \cdot [\mathbf{Z}]_{:,1:(p-1)} \\ &= [\mathbf{Z}]_{:,j}^H \cdot \boldsymbol{\Sigma}_R \boldsymbol{\Sigma}_p \boldsymbol{\Sigma}_T \cdot \left[ [\mathbf{Z}]_{:, (j+1):(p-1)} \quad [\mathbf{Z}]_{:, 1:(j-1)} \right]; \\ &1 \leq j \leq (p-1). \end{aligned} \quad (\text{B.4})$$

Analyzing the term of interstream interference in (B.2), in light of the latter equality, allows us to see that the variance of the interstream interference on the  $p$  streams is the same. Straightforward calculations on the first stream show that the latter common variance is equal to  $E_s[p \sum_{i=1}^p \mu_{\text{conv}_i}^2 - (\sum_{i=1}^p \mu_{\text{conv}_i})^2]/p^2$ , where  $\mu_{\text{conv}_i}$  stands for  $\sigma_{R_i}\sigma_{T_i}$ . Finally, since the filtered receive noise and the interstream interference are statistically independent, the sum of their above calculated variances coincides with the variance of their sum  $\eta$  as stated in (26).

## REFERENCES

- [1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, 1998.
- [2] A. J. Paulraj and T. Kailath, "Increasing capacity in wireless broadcast systems using distributed transmission/directional reception," US Patent, no. 5,345,599, 1994.
- [3] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, 1996.
- [4] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Trans. Communications*, vol. 46, no. 3, pp. 357–366, 1998.
- [5] J. Yang and S. Roy, "On joint transmitter and receiver optimization for multiple-input-multiple-output (MIMO) transmission systems," *IEEE Trans. Communications*, vol. 42, no. 12, pp. 3221–3231, 1994.
- [6] H. Sampath and A. J. Paulraj, "Joint transmit and receive optimization for high data rate wireless communication using multiple antennas," in *Proc. 33rd IEEE Asilomar Conference on Signals, Systems, and Computers*, vol. 1, pp. 215–219, Pacific Grove, Calif, USA, October 1999.
- [7] H. Sampath, P. Stoica, and A. J. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. Communications*, vol. 49, no. 12, pp. 2198–2206, 2001.
- [8] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Trans. Signal Processing*, vol. 50, no. 5, pp. 1051–1064, 2002.
- [9] D. Hughes-Hartogs, "Ensemble modem structure for imperfect transmission media," US Patent no. 4,833,706, May 1989.
- [10] N. Khaled, S. Thoen, M. Vizzardi, and C. Dessel, "A new joint transmit and receive optimization scheme for OFDM-based MIMO systems," in *Proc. 57th IEEE Semiannual Vehicular Technology Conference, (VTC '03)*, vol. 2, pp. 998–1002, Jeju, South Korea, April 2003.
- [11] N. Khaled, S. Thoen, L. Deneire, C. Dessel, and H. De Man, "Spatial-mode selection for the joint transmit receive MMSE design over flat-fading MIMO channels," in *Proc. IEEE Signal Processing in Wireless Communications (SPAWC '03)*, Rome, Italy, June 2003.
- [12] S. Zhou and G. B. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel mean feedback," *IEEE Trans. Communications*, vol. 50, no. 10, pp. 2599–2613, 2002.
- [13] X. Cai and G. B. Giannakis, "Differential space-time modulation with transmit-beamforming for correlated MIMO fading channels," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP '03)*, vol. 4, pp. 25–28, Hong Kong, China, April 2003.
- [14] G. H. Golub and C. F. Van Loan, *Matrix Computations*, John Hopkins University Press, Baltimore, Md, USA, 1996.
- [15] T. A. Thomas and F. W. Vook, "MIMO strategies for equal-rate data streams," in *Proc. 54th IEEE Vehicular Technology Conference (VTC '01)*, vol. 2, pp. 548–552, Atlantic City, NJ, USA, October 2001.
- [16] J. B. Andersen, "Array gain and capacity for known random channels with multiple element arrays at both ends," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 11, pp. 2172–2178, 2000.
- [17] J. G. Proakis, *Digital Communications*, McGraw-Hill, New York, NY, USA, 3rd edition, 1995.
- [18] B. Sklar, *Digital Communications: Fundamentals and Applications*, Prentice-Hall, Englewoods cliffs, NJ, 1st edition, 1988.
- [19] R. F. H. Fischer and J. B. Huber, "A new loading algorithm for discrete multitone transmission," in *Proc. IEEE Global Telecommunications Conference (GLOBECOM '96)*, vol. 1, pp. 724–728, London, UK, November 1996.
- [20] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, NY, USA, 1991.
- [21] A. J. Paulraj and C. B. Papadias, "Space-time processing for wireless communications," *IEEE Trans. Signal Processing*, vol. 14, no. 6, pp. 49–83, 1997.
- [22] S. Thoen, *Transmit optimization for OFDM/SDMA-based wireless local area networks*, Ph.D. thesis, Katholieke Universiteit Leuven and IMEC, Belgium, May 2002.
- [23] A. J. Viterbi, "Convolutional codes and their performance in communication systems," *IEEE Trans. Communications*, vol. 19, no. 5, pp. 751–772, 1971.
- [24] A. Dejonghe and L. Vandendorpe, "Turbo-equalization for multilevel modulation: an efficient low-complexity scheme," in *Proc. IEEE International Conference on Communications (ICC '02)*, vol. 3, pp. 1863–1867, New York, NY, USA, April 2002.
- [25] W. Kaplan, *Advanced Calculus*, Addison-Wesley, Reading, Mass, 4th edition, 1992.

**Nadia Khaled** was born in Rabat, Morocco, in 1977. She received the M.S. degree in electrical engineering from l'Ecole Nationale Supérieure d'Electrotechnique, d'Electronique, d'Informatique, d'Hydraulique et des Télécommunications (ENSEEHT), Toulouse, France, in 2000. Since the completion of the Katholieke Universiteit Leuven predoctoral examination in May 2001, she has been pursuing her Ph.D. research with the wireless research group of IMEC, Leuven, Belgium as a Ph.D. student at the Katholieke Universiteit Leuven. Her research interests lie in the area of signal processing for wireless communications, particularly MIMO techniques and transmit optimization schemes.



**Claude Dessel** was born in Bastogne, Belgium, in 1974. Graduated (with the highest honors) as an Electrical Engineer from the Katholieke Universiteit Leuven, in 1997, he then received the Ph.D. degree from the same university in 2001, funded by the Belgian National Fund for Scientific Research (FNRS). His doctoral research mainly included joint source-channel coding for image transmissions, focusing on unequal error protection, global optimization of a transmission chain, and



image reconstruction from incomplete data. He is also interested in channel coding, especially bit error rate approximation for error-correcting codes and code selection for specific applications. In 2001, he joined IMEC, Leuven, Belgium, where he is now working as a Senior Researcher in the design of wireless communication systems for higher throughput and quality or lower power consumption and complexity. He is currently focusing on ultra-low-power personal area networks, but also has interests in MIMO communications, link adaptation, and turbo coding/processing.

**Steven Thoen** was born in Leuven, Belgium, in 1974. He received the M.S. degree in electrical engineering and the Ph.D. degree in communications engineering from the Katholieke Universiteit Leuven, Belgium, in 1997 and in 2002, respectively. From October 1997 until May 2002, he was with the wireless systems (WISE) group of IMEC, Leuven, Belgium as a Ph.D. student at the Katholieke Universiteit Leuven, supported by an FWO scholarship. From October 1998 to November 1998, he was a Visiting Researcher at the Information Systems Lab, Stanford University, Palo Alto, USA. In July 2002, he joined Resonext Communications where he worked on the design and implementation of advanced WLAN modems. Currently, he is working as a Staff System Engineer in the WLAN group of RF Micro Devices. His research interests include systems design, signal processing, and digital communications systems, with particular regard to transmit optimization and MIMO transmission.



**Hugo De Man** is a Professor of electrical engineering at the Katholieke Universiteit Leuven, Belgium, since 1976. He was a Visiting Associate Professor at UC Berkeley in 1975. His early research was devoted to mixed-signal, switched-capacitor, and DSP simulation tools. In 1984, he was one of the cofounders of IMEC, which, today, is the largest independent semiconductor research institute in Europe with over 1100 employees. From 1984 to 1995, he was the Vice-President of IMEC, responsible for research in design technology for DSP and telecom applications. In 1995, he became a Senior Research Fellow of IMEC, working on strategies for education and research on design of future post-PC systems. His research at IMEC has led to many novel tools and methods in the area of high-level synthesis, hardware-software codesign, and C++ based design. Many of these tools are now commercialized by spin-off companies like Coware and Target Compilers. In 1999, he received the Technical Achievement Award of the IEEE Signal Processing Society, the Phil Kaufman Award of the EDA Consortium, the Golden Jubilee Medal of the IEEE Circuits and Systems Society, and in 2004, the EDAA Lifetime Achievement Award. Hugo De Man is an IEEE Fellow and a Member of the Royal Academy of Sciences in Belgium.



# Blind Identification of Out-of-Cell Users in DS-CDMA

**Tao Jiang**

*Department of Electrical and Computer Engineering, University of Minnesota, 200 Union Street SE,  
Minneapolis, MN 55455, USA  
Email: jiang@ece.umn.edu*

**Nicholas D. Sidiropoulos**

*Department of Electronic and Computer Engineering, Technical University of Crete, Chania-Crete 73100, Greece  
Department of Electrical and Computer Engineering, University of Minnesota, 200 Union Street SE,  
Minneapolis, MN 55455, USA  
Email: nikos@ece.umn.edu*

*Received 29 May 2003; Revised 2 December 2003*

In the context of multiuser detection for the DS-CDMA uplink, out-of-cell interference is usually treated as Gaussian noise, possibly mitigated by overlaying a long random cell code on top of symbol spreading. Different cells use statistically independent long codes, thereby providing means for statistical out-of-cell interference suppression. When the total number of (in-cell plus out-of-cell) users is less than the spreading gain, subspace identification techniques are applicable. If the base station is equipped with multiple antennas, then completely blind identification is possible via three-dimensional low-rank decomposition. This works with more users than spreading and antennas, but a purely algebraic solution is missing. In this paper, we develop an algebraic solution under the premise that the codes of the in-cell users are known. The codes of out-of-cell users and all array steering vectors are unknown. In this pragmatic scenario, we show that in addition to algebraic solution, better identifiability is possible. Our approach yields the best known identifiability result for three-dimensional low-rank decomposition when one of the three component matrices is partially known, albeit noninvertible. Simulations show that the proposed identification algorithm remains close to the pertinent asymptotic (symbol-independent) Cramér-Rao bound, which is also derived herein.

**Keywords and phrases:** cellular systems, smart antennas, interference mitigation.

## 1. INTRODUCTION

In the context of uplink reception for cellular DS-CDMA systems, interference can be classified as either (i) interchip (ICI) and intersymbol (ISI) self-interference, (ii) in-cell multiuser access interference (commonly referred to as MUI or MAI), or (iii) out-of-cell multiuser access interference. The latter is typically ignored or treated as noise; however, it has been reported [1] that in IS-95 other cells account for a large percentage of the interference relative to the interference coming from within the cell. MUI is usually a side-effect of propagation through dispersive multipath channels. The conceptual difference between in-cell and out-of-cell interference boils down to what the base station (BS) can assume about the nature of interfering signals. Typically, the codes of interfering in-cell users are known to the BS, whereas those of out-of-cell users are not. Specifically, in the presence of ICI, the receive-codes of the in-cell users can be estimated via training or subspace techniques (e.g., cf. [2]), using the fact that the transmit-codes are known. This is not the case for out-of-cell users.

Appealing to the central limit theorem, the total interference from out-of-cell users is usually treated as Gaussian noise. In IS-95, a long random cell-specific code is overlaid on top of symbol spreading, and cell despreading is used at the BS to randomize out-of-cell interference. This helps mitigate out-of-cell interference in a statistical fashion. To see how random cell codes work, consider the simplified synchronous flat-fading baseband-equivalent received data model

$$\mathbf{x} = \mathbf{D}_{\text{in}} \mathbf{C}_{\text{in}} \mathbf{s}_{\text{in}} + \mathbf{D}_{\text{out}} \mathbf{C}_{\text{out}} \mathbf{s}_{\text{out}} + \mathbf{n}, \quad (1)$$

where  $\mathbf{x}$  holds the received data corresponding to one symbol period,  $\mathbf{C}_{\text{in}}$  (resp.  $\mathbf{C}_{\text{out}}$ ) is the spreading code matrix,  $\mathbf{s}_{\text{in}}$  (resp.  $\mathbf{s}_{\text{out}}$ ) is the symbol vector,  $\mathbf{D}_{\text{in}}$  (resp.  $\mathbf{D}_{\text{out}}$ ) is a diagonal matrix that holds a portion of the random cell code for the in-cell (resp. out-of-cell) users, and  $\mathbf{n}$  models receiver noise. For simplicity, assume that the in-cell symbol-periodic codes are orthogonal of length  $P$ , and all codes and symbols are BPSK (+1 or -1). Let  $\mathbf{c}_1$  stand for the code of an in-cell user

of interest. Then

$$\begin{aligned} z_1 &:= \frac{1}{P} \mathbf{c}_1^T \mathbf{D}_{\text{in}} \mathbf{x} \\ &= \mathbf{s}_{\text{in}}(1) + \frac{1}{P} \mathbf{c}_1^T \mathbf{D}_{\text{in}} \mathbf{D}_{\text{out}} \mathbf{C}_{\text{out}} \mathbf{s}_{\text{out}} + \tilde{\mathbf{n}}. \end{aligned} \quad (2)$$

The interference term is zero-mean; under certain conditions, its variance is  $O(1/P)$ . This is easy to see for a single out-of-cell user. It follows that random cell codes work reasonably well in relatively underloaded systems with large spreading gain (e.g., 128 chips/symbol), but performance can suffer from near-far effects, and cell codes cannot help identify out-of-cell transmissions. Although the latter may seem of little concern in commercial applications, it can be important for tracking, handoff, and monitoring.

In a way, a structured approach towards the explicit identification<sup>1</sup> of out-of-cell users is the next logical step beyond in-cell multiuser detection and is motivated by considerations similar to those that stimulated research took from matched filtering to multiuser detection. Note that, unlike the case of in-cell interference, out-of-cell interference cannot be mitigated by power control, simply because the BS does not have the authority to exercise power control over out-of-cell users. For a power-controlled in-cell population, near-far effects may be chiefly due to out-of-cell interference. Unfortunately, out-of-cell detection is compounded by the fact that it has to be blind, since the BS has no control and usually no prior information on out-of-cell users. This places limitations on the number and nature of out-of-cell transmissions that can be identified.

The literature on out-of-cell blind identification is scarce. Assuming that (i) the codes of the in-cell users are known, (ii) the total number of (in-cell plus out-of-cell) users is less than the spreading gain and the combined spreading code matrix is full column rank, and (iii) given the correlation matrix of the vector of chip samples taken over a symbol interval, it is possible to cancel out the effect of out-of-cell users [3], then adopt linear or nonlinear solutions for in-cell detection. This approach is appealing, but it has two drawbacks. First, it can be unrealistic to assume that the *total* number of users is less than the spreading gain. This is especially so in loaded systems and urban areas. Second, in practice one uses sample estimates of the correlation matrix. This yields cancellation errors for finite samples, even in the noiseless case.

Recently, a novel code-blind identification approach has been proposed, exploiting uniqueness of low-rank decomposition of three-way arrays [4]. This requires the use of a BS antenna array, but in return allows the identification of both in-cell and out-of-cell users without requiring knowledge of the code or steering vector of *any* user. More users than spreading and antenna elements can be supported. There are two drawbacks to this approach. First, a direct algebraic solution is generally not possible, thus iterative estimation tech-

niques must be employed. Although these iterative methods generally work very well, they are computationally intensive. Second, in-cell code information, which may be available, is not directly exploited (except numerically, by constraining certain parameters during the iterations). In this paper, we develop an algebraic solution that exploits the fact that the codes of the in-cell users are known. In this scenario, we show that in addition to algebraic solution, better identifiability is possible. Our approach yields the best known identifiability result for three-dimensional low-rank decomposition when one of the three component matrices is partially known, albeit noninvertible.

Note that the group-blind multiuser detection approach of [3] can be easily extended to handle multiple BS antennas, but this requires that the array steering vectors, in addition to the spreading codes<sup>2</sup> of all the in-cell users, are known. Estimating steering vectors is more difficult than estimating codes, partly because they are generally unstructured, but also due to mobility-induced fast fading. Note that the approach developed herein (see also [4]) does not assume any parameterization of the manifold vectors.

For clarity of exposition, we will begin our analysis by assuming that both in-cell and out-of-cell user transmissions are synchronized at the BS. In practice, this can be approximately true in synchronous CDMA systems, like CDMA2000.<sup>3</sup> Quasisynchronism (i.e., timing offsets in the order of a few chips) can be handled by dropping a short chip prefix at the receiver. We will refer to both cases as *synchronous CDMA* for brevity. Synchronization is usually achieved via pilot tones emitted from the BS, or a GPS-derived timing reference for synchronous networks involving multiple cells. Out-of-cell transmissions will typically not be synchronized with in-cell transmissions. Notable exceptions include synchronous microcellular networks for “hotspot” coverage, and calls undergoing hand-off at cell boundaries (hence approximately equidistant from the two base stations). As we will see, when delay spread is small relative to the symbol duration, this can be handled by treating each out-of-cell user as two virtual users. Hence our analysis generalizes to the interesting case of a quasisynchronous in-cell population plus asynchronous out-of-cell interference, as in Wideband CDMA (WCDMA). We will refer to this situation as *asynchronous CDMA*.

The rest of the paper is organized as follows. The main ideas and concepts are exposed in Section 2.1, which treats the idealized case of a synchronous DS-CDMA uplink subject to flat fading. This is then extended to frequency-selective multipath and quasisynchronous transmissions in Section 3, which also discusses a suitable *admission protocol*

<sup>1</sup>Here, by *identification* we mean explicitly modeling and estimating all user signals (as opposed to treating certain user signals as unstructured noise).

<sup>2</sup>In the literature, it is common to use the term “(spreading) codes” for the transmit codes, and “signatures” for the effective receive codes. For brevity and to avoid confusion with spatial signatures, we adopt the term “spreading codes” throughout, with the understanding that in the presence of ICI/ISI, the term “codes” means the receive codes.

<sup>3</sup>CDMA2000 uses (universal coordinated time UTC) system time reference, derived from GPS. Mobile stations use the same system time, offset by the propagation delay from the BS to the mobile station.



$\mathbf{X}_p$  admits the factorization

$$\begin{aligned}\mathbf{X}_p &= \mathbf{A}\mathbf{D}_p(\mathbf{C})\mathbf{S}^T + \mathbf{W}_p \\ &= \mathbf{A}_{\text{in}}\mathbf{D}_p(\mathbf{C}_{\text{in}})\mathbf{S}_{\text{in}}^T + \mathbf{A}_{\text{out}}\mathbf{D}_p(\mathbf{C}_{\text{out}})\mathbf{S}_{\text{out}}^T + \mathbf{W}_p \quad (7) \\ &= \mathbf{X}_p^{\text{in}} + \mathbf{X}_p^{\text{out}} + \mathbf{W}_p,\end{aligned}$$

for  $p = 1, 2, \dots, P$ .

It is also worth mentioning that we can write the above set of matrix equations into more compact form if we introduce the so-called Khatri-Rao product  $\odot$  (column-wise Kronecker product, see [4] and references therein). Stacking the matrices in (7), we obtain

$$\begin{aligned}\mathbf{X}^{KP \times N} &:= \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \vdots \\ \mathbf{X}_P \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{A}\mathbf{D}_1(\mathbf{C}) \\ \mathbf{A}\mathbf{D}_2(\mathbf{C}) \\ \vdots \\ \mathbf{A}\mathbf{D}_P(\mathbf{C}) \end{bmatrix} \mathbf{S}^T + \begin{bmatrix} \mathbf{W}_1 \\ \mathbf{W}_2 \\ \vdots \\ \mathbf{W}_P \end{bmatrix} \quad (8) \\ &= (\mathbf{C} \odot \mathbf{A})\mathbf{S}^T + \mathbf{W}^{KP \times N} \\ &= (\mathbf{C}_{\text{in}} \odot \mathbf{A}_{\text{in}})\mathbf{S}_{\text{in}}^T + (\mathbf{C}_{\text{out}} \odot \mathbf{A}_{\text{out}})\mathbf{S}_{\text{out}}^T + \mathbf{W}^{KP \times N}.\end{aligned}$$

Due to the symmetry of the model (6), we may also recast (8) in the following form

$$\tilde{\mathbf{X}}^{PN \times K} = (\mathbf{S} \odot \mathbf{C})\mathbf{A}^T + \tilde{\mathbf{W}}^{PN \times K}, \quad (9)$$

where  $\tilde{\mathbf{W}}^{PN \times K}$  is a reshuffled AWGN matrix (see [4]).

In what follows, we consider detecting the signal matrix  $\mathbf{S}$  transmitted from *all* active users given only knowledge of  $\mathbf{C}_{\text{in}}$  and  $M$ . As a byproduct, we will be able to recover the steering matrix  $\mathbf{A}$  and the unknown spreading code matrix  $\mathbf{C}_{\text{out}}$  from the received data  $\mathbf{X}$  as well.

## 2.2. Preliminaries

In the sequel, we will need to invoke certain preliminary results in order to prove our main identifiability result in Theorem 1. *Identifiability* means that, *in the absence of noise*, it is possible to recover the sought signals (model parameters) without error; that is, it is possible to pin down the sought parameters exactly. For this reason, we drop noise terms in the discussion that follows. The basic ideas behind preliminary results leading to Theorem 1 are due to Harshman [5]. We begin by recalling the definition of  $k$ -rank.

### 2.2.1. Definition

*Definition 1.* The  $k$ -rank [6] of  $\mathbf{A}$  is equal to  $k_A$  if every  $k_A$  columns drawn from  $\mathbf{A}$  are linearly independent, and either there exists a collection of  $k_A + 1$  linearly dependent columns in  $\mathbf{A}$  or  $\mathbf{A}$  has exactly  $k_A$  columns. Note that  $k_A \leq \text{rank}(\mathbf{A})$ , for all  $\mathbf{A}$ .

### 2.2.2. Eigenanalysis

Consider two matrices  $\mathbf{X}_1 = \mathbf{A}\mathbf{D}_1(\mathbf{C})\mathbf{S}^T$ ,  $\mathbf{X}_2 = \mathbf{A}\mathbf{D}_2(\mathbf{C})\mathbf{S}^T$ , where both  $\mathbf{A} \in \mathbb{C}^{K \times M}$  and  $\mathbf{S} \in \mathbb{C}^{N \times M}$  are full column rank ( $M$ ),  $\mathbf{C} \in \mathbb{C}^{2 \times M}$  contains no zero entry, and all elements on the diagonal of  $\mathbf{D} := \mathbf{D}_2(\mathbf{C})\mathbf{D}_1^{-1}(\mathbf{C})$  are *assumed*<sup>4</sup> distinct. Consider the singular value decomposition (SVD) of the stacked data matrix

$$\begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{A} \\ \mathbf{A}\mathbf{D} \end{bmatrix} \mathbf{D}_1(\mathbf{C})\mathbf{S}^T = \mathbf{U}\Sigma\mathbf{V}^H. \quad (10)$$

The linear space spanned by the columns of  $\mathbf{U}$  is the same as the space spanned by the columns of  $\begin{bmatrix} \mathbf{A} \\ \mathbf{A}\mathbf{D} \end{bmatrix}$  since  $\mathbf{S}\mathbf{D}_1(\mathbf{C})$  has full column rank; hence there exists a nonsingular matrix  $\mathbf{P}$  such that

$$\mathbf{U}\mathbf{P} = \begin{bmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \end{bmatrix} \mathbf{P} = \begin{bmatrix} \mathbf{A} \\ \mathbf{A}\mathbf{D} \end{bmatrix}. \quad (11)$$

Next, construct the auto- and cross-product matrices

$$\begin{aligned}\mathbf{R}_0 &= \mathbf{U}_1^H \mathbf{U}_1 = \mathbf{P}^{-H} \mathbf{A}^H \mathbf{A} \mathbf{P}^{-1} := \mathbf{Q}\mathbf{P}^{-1}, \\ \mathbf{R}_1 &= \mathbf{U}_1^H \mathbf{U}_2 = \mathbf{P}^{-H} \mathbf{A}^H \mathbf{A} \mathbf{D} \mathbf{P}^{-1} := \mathbf{Q}\mathbf{D}\mathbf{P}^{-1}.\end{aligned} \quad (12)$$

Note that since both  $\mathbf{A}$  and  $\mathbf{S}$  are assumed full column rank,<sup>5</sup> the matrices  $\mathbf{R}_0$ ,  $\mathbf{R}_1$ ,  $\mathbf{Q}$ ,  $\mathbf{P}$ , and  $\mathbf{D}$  in (12) are  $M \times M$  full rank matrices. Solving the first equation in (12) for  $\mathbf{Q}$ , then substituting the result into the second, it follows that

$$(\mathbf{R}_0^{-1} \mathbf{R}_1) \mathbf{P} = \mathbf{P} \mathbf{D}, \quad (13)$$

which is a standard eigenvalue problem with distinct eigenvalues.  $\mathbf{P}$  can therefore be determined up to permutation and scaling of columns based on the matrices  $\mathbf{X}_1$  and  $\mathbf{X}_2$ . After that,  $\mathbf{A}$  can be obtained as  $\mathbf{A} = \mathbf{U}_1 \mathbf{P}$ ,  $\mathbf{C}\mathbf{D}_1^{-1}(\mathbf{C})$  can be retrieved with all ones in the first row, and the entire second row taken from the diagonal of  $\mathbf{D}$ , and finally  $\mathbf{S}\mathbf{D}_1(\mathbf{C})$  can be

<sup>4</sup>Note that the columns of  $\mathbf{C}$  correspond to chip-rate samples of the *received* codes (or *signatures*) of the users, that is, the convolution of the transmit codes and the respective multipath channels. Without such multipath, BPSK or other finite-alphabet codes would violate the condition that the diagonal elements of  $\mathbf{D}_2(\mathbf{C})\mathbf{D}_1^{-1}(\mathbf{C})$  are distinct. However, note that we do not advocate using this result for actual separation—it is merely listed here as background needed in the proof of our main result in Theorem 1. Due to the use of the left pseudoinverse of  $\mathbf{C}_{\text{in}}$  employed to bring  $\mathbf{C}$  in canonical form, the  $\mathbf{C}_{\text{out}}$  in Theorem 1 holds *code cross-correlations*, rather than actual codes. For some binary codes, for example, Gold and Kasami codes, the conditions in Theorem 1 hold with high probability. With random multipath taps, the condition can be shown to hold almost surely. Furthermore, the condition can also be sustained with real- or complex-valued spreading codes.

<sup>5</sup>This implies that  $K \geq M$  and  $N \geq M$ , but note again that we do not advocate using this argument as is for separation; we rather present it as a building block to be used later in Theorem 1.

recovered as  $\mathbf{SD}_1(\mathbf{C}) = (\mathbf{A}^\dagger \mathbf{X}_1)^T$ , all under the same permutation and scaling of columns, which carries over from the solution of the eigenvalue problem in (13).

Repeated values along the diagonal of  $\mathbf{D}_2(\mathbf{C})\mathbf{D}_1^{-1}(\mathbf{C})$  give rise to eigenvalues of multiplicity higher than one. In this case, the span of eigenvectors corresponding to each distinct eigenvalue can still be uniquely determined. This will be important when we discuss the case of asynchronous out-of-cell users later in Section 3.

More generally, we have the following claim.

**Claim 1.** Given matrices  $\mathbf{X}_p = \mathbf{AD}_p(\mathbf{C})\mathbf{S}^T$  for  $p = 1, \dots, P \geq 2$ ,  $\mathbf{A}$ ,  $\mathbf{C}$ , and  $\mathbf{S}$  can be found up to permutation and scaling of columns provided that both  $\mathbf{A}$  and  $\mathbf{S}$  are full column rank, and  $k_C \geq 2$ .

Since  $k_C \geq 2$ , we know that the spreading code matrix  $\mathbf{C}$  does not contain any zero columns. Note that  $k_C \geq 2$  does not necessarily imply that there always exists a submatrix of  $\mathbf{C}$  which comprises two rows of  $\mathbf{C}$  such that the  $k$ -rank of this submatrix is 2. For instance, consider

$$\mathbf{C} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 1 \end{bmatrix}. \quad (14)$$

It can be seen that  $r_C = k_C = 3$ , whereas none of the  $2 \times 3$  submatrices of  $\mathbf{C}$  has  $k$ -rank greater than 1. From this example, it is evident that one cannot prove Claim 1 by eigendecomposition applied to a pair of  $\mathbf{X}_p$ 's. For this, we will need the following claim.

**Claim 2.** Given  $\mathbf{C} \in \mathbb{C}^{P \times M}$  with  $k_C \geq 2$ , there always exists a  $2 \times P$  matrix  $\mathbf{G}$  such that the  $k$ -rank of  $\mathbf{GC}$  is two.

For a proof of Claim 2, note that the objective can be easily shown equivalent to proving that there exists a  $2 \times P$  matrix  $\mathbf{G}$  such that the determinants of all  $2 \times 2$  submatrices of  $\mathbf{GC}$  are not zero.  $\mathbf{G}$  is determined by its  $2P$  complex entries. The determinant of each  $2 \times 2$  submatrix of  $\mathbf{GC}$  is a polynomial in those  $2P$  variables, and hence analytic. Since  $k_C \geq 2$ , for each specific  $2 \times 2$  submatrix of  $\mathbf{GC}$ , for instance, the submatrix comprising the first two columns of  $\mathbf{GC}$ , it is not hard to show that there always exists a  $\mathbf{G}_0$  such that the determinant of the corresponding submatrix of  $\mathbf{G}_0\mathbf{C}$  is not zero. Invoking [7, Lemma 2], we conclude that the set of  $\mathbf{G}$ 's which yield zero determinant for any specific submatrix of  $\mathbf{GC}$  constitutes a measure zero set in  $\mathbb{C}^{2P}$ . The number of all  $2 \times 2$  submatrices of  $\mathbf{GC}$  is finite, and any finite union of measure zero sets is of measure zero. The existence of the desired  $\mathbf{G}$  thus follows. Not only does such a  $\mathbf{G}$  exist, but in fact a random  $\mathbf{G}$  drawn from, for example, a Gaussian product distribution, will do with probability one. This establishes Claim 2.

The existence of such  $\mathbf{G}$  implies that the elements on the diagonal of  $\mathbf{D}_2(\mathbf{GC})\mathbf{D}_1^{-1}(\mathbf{GC})$  will be distinct. Therefore, the eigenanalysis steps can be carried through to solve for  $\mathbf{A}$  and  $\mathbf{S}$  from the two mixed slabs  $\mathbf{AD}_1(\mathbf{GC})\mathbf{S}^T$  and  $\mathbf{AD}_2(\mathbf{GC})\mathbf{S}^T$ . With the recovered  $\mathbf{A}$  and  $\mathbf{S}$ ,  $\mathbf{C}$  can be computed from  $\mathbf{X}_p$ . Therefore Claim 1 follows.

### 2.2.3. Lemma

In the proof of our main theorem, we will need the following lemma.

**Lemma 1.** Given

$$\begin{bmatrix} 1 & 0 & * & \cdots & * \\ 0 & 1 & * & \cdots & * \end{bmatrix} \in \mathbb{C}^{2 \times M}, \quad (15)$$

where  $*$  stands for a nonzero entry, it holds that for almost every  $(\mu_1, \mu_2) \in \mathbb{R}^2$  (i.e., except for a set of Lebesgue measure zero), the matrix

$$\begin{aligned} \mathbf{E} &:= \begin{bmatrix} 1 & 1 \\ \mu_1 & \mu_2 \end{bmatrix} \begin{bmatrix} 1 & 0 & * & \cdots & * \\ 0 & 1 & * & \cdots & * \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 & \bullet & \cdots & \bullet \\ \mu_1 & \mu_2 & * & \cdots & * \end{bmatrix} \end{aligned} \quad (16)$$

contains no zero entry in the second row; and the first two elements on the diagonal of  $\mathbf{D}_1(\mathbf{E})\mathbf{D}_2^{-1}(\mathbf{E})$  are distinct and distinct from the remaining elements.

*Proof.* Having a zero entry in the second row occurs when  $(\mu_1, \mu_2)$  lies on the union of  $M$  lines. Since a finite union of lines cannot cover the plane, zeros in the second row are excluded almost surely. The second claim can be proven in the same manner.  $\square$

### 2.3. Main theorem on identifiability

Without loss of generality, we assume that  $\mathbf{C}_{\text{in}}$  is in canonical form. The general case can be reduced to canonical form as explained in the following section.

**Theorem 1.** Given  $\mathbf{X}_p = \mathbf{AD}_p(\mathbf{C})\mathbf{S}^T$ ,  $p = 1, \dots, P$ ,  $2 \leq M_{\text{in}} \leq P$ , where  $\mathbf{A} \in \mathbb{C}^{K \times M}$ ,  $\mathbf{C} \in \mathbb{C}^{P \times M}$ ,  $\mathbf{S} \in \mathbb{C}^{N \times M}$ , and  $\mathbf{C}$  in canonical form

$$\mathbf{C} = [\mathbf{I}_P(1 : M_{\text{in}})\mathbf{C}_{\text{out}}], \quad (17)$$

where  $\mathbf{I}_P(1 : M_{\text{in}})$  denotes the first  $M_{\text{in}}$  columns of  $\mathbf{I}_P$ , if the first  $M_{\text{in}}$  rows of  $\mathbf{C}_{\text{out}}$  contain no zero entries, and  $k_C \geq 2$ ,  $\min\{k_A, k_S\} \geq M_{\text{out}} + 2$ , then the matrices  $\mathbf{A}$ ,  $\mathbf{C}$ , and  $\mathbf{S}$  are unique up to permutation and scaling of columns.

*Proof.* We will show that we can first recover  $\mathbf{A}_{\text{in}}$  and  $\mathbf{S}_{\text{in}}$  up to permutation and scaling of columns from the given  $\mathbf{X}_p$ , and then obtain  $\mathbf{A}_{\text{out}}$ ,  $\mathbf{C}_{\text{out}}$ , and  $\mathbf{S}_{\text{out}}$  afterwards.

We begin by recovering the first two columns of  $\mathbf{A}_{\text{in}}$  and  $\mathbf{S}_{\text{in}}$ . Start from

$$\begin{aligned} \mathbf{X}_1 &= \mathbf{AD}_1(\mathbf{C})\mathbf{S}^T \\ &= \mathbf{A} \text{diag} \left[ 1 \ 0 \ \overbrace{0 \cdots 0}^{M_{\text{in}}-2} \ \overbrace{* \cdots *}^{M_{\text{out}}} \right] \mathbf{S}^T \\ &= \bar{\mathbf{A}} \text{diag} \left[ 1 \ 0 \ * \ \cdots \ * \right] \bar{\mathbf{S}}^T, \\ \mathbf{X}_2 &= \mathbf{AD}_2(\mathbf{C})\mathbf{S}^T \\ &= \mathbf{A} \text{diag} \left[ 0 \ 1 \ 0 \ \cdots \ 0 \ * \ \cdots \ * \right] \mathbf{S}^T \\ &= \bar{\mathbf{A}} \text{diag} \left[ 0 \ 1 \ * \ \cdots \ * \right] \bar{\mathbf{S}}^T. \end{aligned} \quad (18)$$

Recall that  $*$  stands for a nonzero entry;  $\bar{\mathbf{A}}$  ( $\bar{\mathbf{S}}$ ) is a column-reduced submatrix of  $\mathbf{A}$  ( $\mathbf{S}$ ). Invoking Lemma 1, we always can pick a pair  $(\mu_1, \mu_2) \in \mathbf{R}^2$  such that

$$\begin{aligned} \mathbf{E} &:= \begin{bmatrix} 1 & 1 \\ \mu_1 & \mu_2 \end{bmatrix} \begin{bmatrix} 1 & 0 & * & \cdots & * \\ 0 & 1 & * & \cdots & * \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 & \bullet & \cdots & \bullet \\ \mu_1 & \mu_2 & * & \cdots & * \end{bmatrix} \end{aligned} \quad (19)$$

contains no zero entry in the second row; and the first two elements on the diagonal of  $\mathbf{D}_1(\mathbf{E})\mathbf{D}_2^{-1}(\mathbf{E})$  are distinct and distinct from the remaining elements. We also note that both  $\bar{\mathbf{A}}$  and  $\bar{\mathbf{S}}$  have  $M_{\text{out}} + 2$  columns from the original  $\mathbf{A}$  and  $\mathbf{S}$ ; by definition of  $k$ -rank, it follows that

$$\begin{aligned} k_{\bar{\mathbf{A}}} &\geq \min(k_{\mathbf{A}}, M_{\text{out}} + 2), \\ k_{\bar{\mathbf{S}}} &\geq \min(k_{\mathbf{S}}, M_{\text{out}} + 2). \end{aligned} \quad (20)$$

Due to the fact that  $\min\{k_{\bar{\mathbf{A}}}, k_{\bar{\mathbf{S}}}\} \geq M_{\text{out}} + 2$ , both  $\bar{\mathbf{A}}$  and  $\bar{\mathbf{S}}$  are full column rank. Therefore, eigendecomposition as in Section 2.2.2 can be applied to the following mixed slabs,

$$\begin{aligned} \mathbf{Y}_1 &= \mathbf{X}_1 + \mathbf{X}_2 = \bar{\mathbf{A}} \text{diag}[1 \ 1 \ \bullet \ \cdots \ \bullet] \bar{\mathbf{S}}^T, \\ \mathbf{Y}_2 &= \mu_1 \mathbf{X}_1 + \mu_2 \mathbf{X}_2 = \bar{\mathbf{A}} \text{diag}[\mu_1 \ \mu_2 \ \bullet \ \cdots \ \bullet] \bar{\mathbf{S}}^T, \end{aligned} \quad (21)$$

to recover the first two columns of  $\mathbf{A}$  and  $\mathbf{S}^T$  up to permutation and scaling. We can repeat this procedure with  $\mathbf{X}_i$  and  $\mathbf{X}_{i+1}$  to recover the  $i$ th and the  $(i+1)$ th columns of  $\mathbf{A}_{\text{in}}$  and  $\mathbf{S}_{\text{in}}$  for  $i = 2, \dots, M_{\text{in}} - 1$  until both  $\mathbf{A}_{\text{in}}$  and  $\mathbf{S}_{\text{in}}$  are recovered. The matrices  $\mathbf{X}_p^{\text{in}} := \mathbf{A}_{\text{in}} \mathbf{D}_p (\mathbf{I}_P (1 : M_{\text{in}})) \mathbf{S}_{\text{in}}^T$  corresponding to the in-cell users can be constructed, and we thus obtain the matrices  $\mathbf{X}_p^{\text{out}}$  by subtracting  $\mathbf{X}_p^{\text{in}}$  from  $\mathbf{X}_p$  for  $p = 1, \dots, P$ .

$\mathbf{X}_p^{\text{out}}$  is nothing but  $\mathbf{A}_{\text{out}} \mathbf{D}_p (\mathbf{C}_{\text{out}}) \mathbf{S}_{\text{out}}$ . Since  $\mathbf{A}_{\text{out}}$ ,  $\mathbf{C}_{\text{out}}$ , and  $\mathbf{S}_{\text{out}}$  are all  $M_{\text{out}}$ -column submatrices of  $\mathbf{A}$ ,  $\mathbf{C}$ , and  $\mathbf{S}$ , respectively, we have

$$\begin{aligned} k_{\mathbf{A}_{\text{out}}} &\geq \min(k_{\mathbf{A}}, M_{\text{out}}) = M_{\text{out}}, \\ k_{\mathbf{S}_{\text{out}}} &\geq \min(k_{\mathbf{S}}, M_{\text{out}}) = M_{\text{out}}, \\ k_{\mathbf{C}_{\text{out}}} &\geq \min(k_{\mathbf{C}}, M_{\text{out}}) = \min(2, M_{\text{out}}). \end{aligned} \quad (22)$$

The first two inequalities hold due to the condition that  $\min\{k_{\mathbf{A}}, k_{\mathbf{S}}\} \geq M_{\text{out}} + 2$ , and imply that both  $\mathbf{A}_{\text{out}}$  and  $\mathbf{S}_{\text{out}}$  are full column rank matrices.

If  $M_{\text{out}} \geq 2$ , we know that  $k_{\mathbf{C}_{\text{out}}} \geq 2$ ; therefore Claim 1 can be invoked, and eigenanalysis of two mixed slabs can be carried out to recover  $\mathbf{A}_{\text{out}}$ ,  $\mathbf{C}_{\text{out}}$ , and  $\mathbf{S}_{\text{out}}$ , up to permutation and scaling of columns.

When  $M_{\text{out}} = 1$ , it is known that rank-one matrix decomposition is unique up to scaling.  $\square$

*Remark 1.* Note that  $\mathbf{C}$  in Theorem 1 can be a fat matrix. A similar result can be derived for  $M_{\text{in}} = 1$ , with slightly different conditions on  $\mathbf{C}_{\text{out}}$ .

*Remark 2.* The assumption that the first  $M_{\text{in}}$  rows of  $\mathbf{C}_{\text{out}}$  contain no zero entries is posed mainly for simplicity of proof

of Theorem 1. Theorem 1 holds, provided that none of the columns of the submatrix comprising the first  $M_{\text{in}}$  rows of  $\mathbf{C}_{\text{out}}$  is proportional to a column of  $\mathbf{I}_{M_{\text{in}}}$ . We chose to prove the slightly restricted Theorem 1 due to space considerations.

*Remark 3.* The model identifiability conditions of Theorem 1 are usually met in practice deterministically or statistically with proper system parameters. For instance, if we assume that  $\mathbf{A}$  and  $\mathbf{C}$  are drawn from a continuous distribution, and  $\mathbf{S}$  drawn from an i.i.d. BPSK source, it can be shown that  $k_{\mathbf{A}} \geq M_{\text{out}} + 2$ ,  $k_{\mathbf{C}} \geq 2$  holds almost surely, provided  $K \geq M_{\text{out}} + 2$ ,  $P \geq 2$ , while  $k_{\mathbf{S}} \geq M_{\text{out}} + 2$  occurs with high probability provided that  $N$  is moderately higher than  $M$ .

## 2.4. Algorithms

The proof of Theorem 1 is constructive; it directly yields a sequential eigenvalue-based solution that recovers everything exactly in the noiseless case, under only the model identifiability condition in the theorem. In the noisy scenario, this eigenvalue approach can be coupled with an iterative LS-based refinement algorithm that yields good estimation performance for moderate signal-to-noise ratio (SNR) and beyond.

Assuming that  $\mathbf{C}_{\text{in}}$  is known, the two major steps of our algorithm are summarized next.

### (1) Algebraic initialization

Arrange the received noisy data  $x_{k,n,p}$  into a set of matrices,  $\tilde{\mathbf{X}}_k \in \mathbb{C}^{P \times N}$ , for  $k = 1, \dots, K$ . The  $(p, n)$  entry of  $\tilde{\mathbf{X}}_k$  is  $x_{k,n,p}$ . It can be shown that

$$\tilde{\mathbf{X}}_k = \mathbf{C} \mathbf{D}_k (\mathbf{A}) \mathbf{S}^T + \tilde{\mathbf{W}}_k, \quad (23)$$

where  $\tilde{\mathbf{W}}_k$  is the AWGN matrix. Left multiply by the pseudo-inverse of  $\mathbf{C}_{\text{in}}$  to get  $\tilde{\mathbf{Z}}_k \in \mathbb{C}^{M_{\text{in}} \times N}$ :

$$\tilde{\mathbf{Z}}_k = \mathbf{C}_{\text{in}}^\dagger \tilde{\mathbf{X}}_k. \quad (24)$$

Form another set of matrices  $\mathbf{X}_m \in \mathbb{C}^{K \times N}$ , for  $m = 1, \dots, M_{\text{in}}$  such that the  $(k, n)$  entry of  $\mathbf{X}_m$  is equal to the  $(m, n)$  entry of  $\tilde{\mathbf{Z}}_k$ . It can be shown that

$$\mathbf{X}_m = \mathbf{A} \mathbf{D}_m (\mathbf{C}_{\text{in}}^\dagger \mathbf{C}) \mathbf{S}^T + \mathbf{W}_m, \quad (25)$$

where  $\mathbf{W}_m$  is the rearranged Gaussian noise matrix. Note that  $\mathbf{C}_{\text{in}}^\dagger \mathbf{C}$  is in canonical form, and thus we may apply the approach described in the proof of Theorem 1 to estimate  $\mathbf{A}$ ,  $\mathbf{C}_{\text{in}}^\dagger \mathbf{C}_{\text{out}}$ , and  $\mathbf{S}$ .  $\mathbf{C}$  can also be estimated as

$$\mathbf{C} = \left[ \begin{bmatrix} \mathbf{A} \mathbf{D}_1 (\mathbf{S}) \\ \vdots \\ \mathbf{A} \mathbf{D}_N (\mathbf{S}) \end{bmatrix}^\dagger \begin{bmatrix} \bar{\mathbf{X}}_1 \\ \vdots \\ \bar{\mathbf{X}}_N \end{bmatrix} \right]^T, \quad (26)$$

where the  $(k, p)$  element of  $\bar{\mathbf{X}}_n \in \mathbb{C}^{K \times P}$  is given by  $x_{k,n,p}$  (cf. [4] for details).

### (2) Joint constrained Least Squares refinement

Use the  $\mathbf{A}$ ,  $\mathbf{C}_{\text{out}}$ , and  $\mathbf{S}$  obtained in the first step and the known  $\mathbf{C}_{\text{in}}$  as initialization for constrained trilinear alternating least squares (CTALS) regression applied to the original data  $x_{k,n,p}$ . The basic idea behind TALS is to compute a conditional LS update of  $\mathbf{A}$  given  $\mathbf{C}$ ,  $\mathbf{S}$ , then repeat for  $\mathbf{S}$ , and so forth in a circular fashion until convergence [4]. For CTALS, the  $\mathbf{C}_{\text{in}}$  part of  $\mathbf{C}$  is fixed, and only  $\mathbf{C}_{\text{out}}$  is updated in the iterations.

### 3. EXTENSION TO QUASISYNCHRONOUS SYSTEMS AND MULTIPATH CHANNELS

There are two issues that must be addressed in order to establish the usefulness of our algorithm in a realistic cellular CDMA environment. One is synchronization; the other is frequency selectivity.

In so-called quasisynchronous CDMA (QS-CDMA) the symbol timing of the in-cell users may be off by as much as a few chips. This causes ISI, but, as already mentioned, it can be circumvented by dropping a short chip-prefix for each symbol at the receiver—the associated performance degradation is negligible when the prefix is short relative to the spreading gain.

Quasisynchronism is a reasonable assumption for the in-cell user population in the context of 3G systems (e.g., CDMA2000), but much less so for out-of-cell users, who actually attempt to synchronize with a *different* BS. The key here is (5) asynchronous out-of-cell users appear as two virtual synchronous users, with “split” code pieces, and symbol sequences that are offset by one symbol. Note that splitting and offset generally preserve linear independence; however, the steering vectors (spatial responses) will be colinear for each such pair of virtual users. Fortunately, by exchanging the roles of  $\mathbf{A}$  and  $\mathbf{C}$  and invoking the remark on repeated eigenvalues in Section 2.2.2, it can be shown that the parameters of all in-cell users can still be uniquely determined, along with the span of each pair of virtual out-of-cell users.

Frequency selectivity is realistically modeled by convolution with a relatively short chip-rate FIR filter that models the discrete-time baseband-equivalent channel impulse response, including transmit chip pulse-shaping and receive chip-matched filtering. The effective spreading codes seen at the receiver are the convolution of the transmit codes with the corresponding multipath channels. This means that the in-cell receive codes must be estimated before our basic approach developed in the above section can be applied. This estimation is compounded by the cochannel out-of-cell interference, which is not under the control of the BS. In order to deal with the problem of receive-code estimation for the in-cell users, we propose the following *admission protocol* “as new in-cell users come into the system, they are initially treated as out-of-cell: their receive-codes are thereby estimated blindly, and they are subsequently added to the list of in-cell users. Initially, the process is started by solving a blind problem,” as in [4]. In this way, the problem of receive-code estimation for the in-cell users is never explicitly solved.

Once the in-cell receive-codes have been estimated at the BS, the proposed algorithm can be carried over to the quasisynchronous frequency-selective DS-CDMA systems.

### 4. ASYMPTOTIC CRAMÉR-RAO BOUND

In order to benchmark the performance of our estimation algorithm, it is useful to derive pertinent bounds. While low bit error rate (BER) is of primary concern, accurate estimates of the out-of-cell user’s receive-codes and both in-cell and out-of-cell steering vectors are also of interest. CRBs can be developed for the latter, owing to the fact that unlike symbols, steering vectors and receive-codes are continuous parameters.

The conditional CRB for low-rank decomposition of multidimensional arrays has been derived in [8], assuming all matrices are fixed unknowns. In our present context, however, we are more interested in bounds that are independent of the symbol matrix  $\mathbf{S}$ . Towards this end, we can aim for one of two options: computing an averaged (or modified) CRB, or an asymptotic CRB. The former turns out to be far more complicated to derive in closed form; we therefore opt for the latter.

In the appendix, wherein the detailed CRB derivations can be found, we begin by developing a compact form of the conditional CRB in [8]. The new compact form is much simpler to compute than the expression given in [8]. Then, following the approach developed in [9], we work out the asymptotic CRB as the number of symbols,  $N$ , goes to infinity. The key to this computation is that the limit and the CRB operator can be exchanged, since the latter is continuous; and when  $N$  tends to infinity, the sample estimate of the correlation matrix of  $\mathbf{S}$  approaches the exact correlation matrix of  $\mathbf{S}$ . For the sake of brevity, in what follows, we assume that the entries of  $\mathbf{S}$  are drawn from an i.i.d. BPSK source. This implies that

$$E(s_{m_1}(n_1)s_{m_2}(n_2)) = \delta_{m_1,m_2}\delta_{n_1,n_2}. \quad (27)$$

Note that the asymptotic CRB derived in the appendix is valid for arbitrary  $\mathbf{C}$ —it is not necessary to have  $\mathbf{C}_{\text{in}}$  in canonical form. The main limitation of the asymptotic CRB is that it is valid for large enough  $N$ , but for small  $N$  there will be some mismatch.

### 5. SIMULATION RESULTS

In this section, we provide computer simulation results to demonstrate the performance of the proposed algorithm.

As per Theorem 1, scaling ambiguity for all active users and the permutation ambiguity among out-of-cell users is inherent to this blind separation problem. We remove the column scaling ambiguity among the estimated symbol matrix  $\mathbf{S}$  via differential encoding, and assume differentially encoded user signals throughout the simulations. For the purpose of performance evaluation only, the permutation ambiguity among the out-of-cell users is resolved using a greedy least square matching algorithm [4]. This permutation

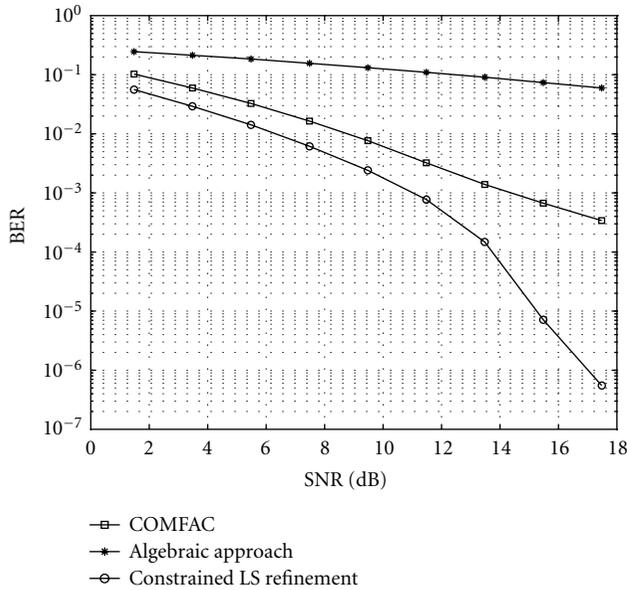


FIGURE 1: No out-of-cell user interference.

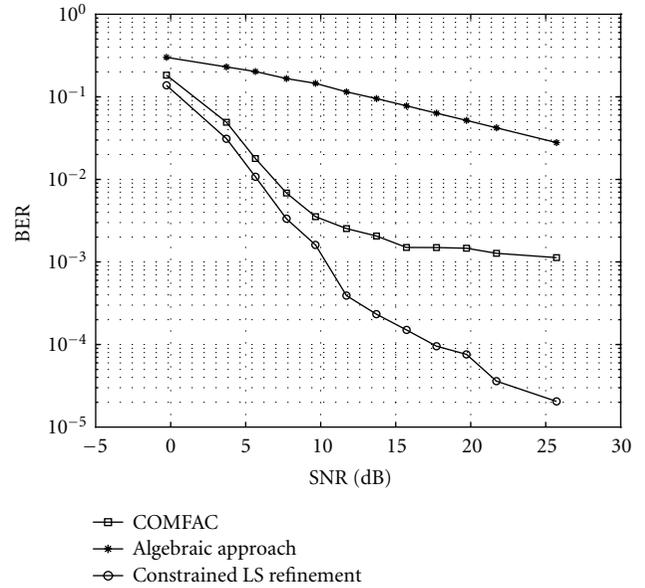


FIGURE 2: More active users than spreading gain.

ambiguity among the out-of-cell users cannot be solved at the BS without additional side information, but this indeterminacy is irrelevant in practice.

Let  $\mathbf{X}_p = \mathbf{A}\mathbf{D}_p(\mathbf{C})\mathbf{S}^T + \mathbf{W}_p$  be the received noisy data, for  $p = 1, \dots, P$ , where  $\mathbf{W}_p$  are the AWGN matrices. We define the sample SNR at the input of the multiuser receiver as

$$\text{SNR} := 10 \log_{10} \frac{\sum_{p=1}^P \|\mathbf{A}\mathbf{D}_p(\mathbf{C})\mathbf{S}^T\|_F^2}{\sum_{p=1}^P \|\mathbf{W}_p\|_F^2} \text{dB}. \quad (28)$$

We first show that the proposed algebraic initialization significantly accelerates the convergence of least square refinement and improves the performance. In order to have a benchmark, we consider cases wherein the TALS-based COMFAC algorithm [4] is also applicable, but note that the approach developed herein can work well when COMFAC fails. When both methods are applicable, our simulations show that the new approach yields better performance.

Figure 1 plots BER versus average SNR, without out-of-cell interference and for  $M_{\text{in}} = 4$ , DE-BPSK,  $K = 2$ ,  $N = 50$ , and  $P = 4$ . Results are averaged over  $10^2$  i.i.d. Rayleigh channels ( $\mathbf{A}$ —no power control is assumed), and  $10^6$  realizations per each Rayleigh channel. Note that total averaging is  $O(10^8)$ . The spreading codes are randomly drawn from a continuous distribution and fixed throughout the simulations. Figure 2 depicts average BER for the in-cell users for  $M_{\text{in}} = 4$ ,  $M_{\text{out}} = 2$ ,  $K = 4$ ,  $N = 50$ ,  $P = 4$ , and otherwise the same simulation setup. Note that in the second experiment, both the number of antennas and spreading gain are less than the number of total active users. It is seen from those figures that, as expected, the proposed algorithm has provided better BER performance than COMFAC; in particular, such

improvement is significant in the high SNR regime. In addition, the proposed algorithm has been observed to converge at least 70 percent faster (in terms of time) than the general TALS with random initialization, and comparably with respect to the computation-efficient TALS-based COMFAC, especially in the high SNR regime.

Next, the performance of the proposed algorithm and that of the linear group-blind decorrelating detector [3] with two different sample sizes is shown in Figure 3. The original group-blind multiuser detector is designed for uplink CDMA with a single receive antenna, but the approach of [3] can be easily extended to handle multiple BS antennas, provided that the array steering vectors, in addition to the spreading codes, of all the in-cell users are known. Estimating steering vectors is more difficult than estimating codes, because the former vary faster due to mobility-induced fast fading. In our simulation, in contrast to the proposed algorithm, the linear group-blind decorrelating detector assumes perfect knowledge of in-cell user's steering matrix  $\mathbf{A}_{\text{in}}$ , that is, we provide the linear group-blind decorrelating detector with perfect knowledge of  $(\mathbf{C}_{\text{in}} \odot \mathbf{A}_{\text{in}})$  in (8). Figure 3 depicts the performance of the two competing detectors for two different sample sizes,  $N = 25$ ,  $N = 50$ . It is observed that the linear group-blind decorrelating detector exhibits an error floor in the high SNR regime due to using sample estimates of the correlation matrix. This yields cancellation errors which persist for any number of finite samples, even in the noiseless case. However, such error floor is acceptable when we use large sample sizes. With 50 snapshots, the linear group-blind decorrelating detector provides better BER performance than the proposed detector in the high SNR regime even though the error floor surfaces at about 24 dB. With a small sample size of  $N = 25$ , the proposed

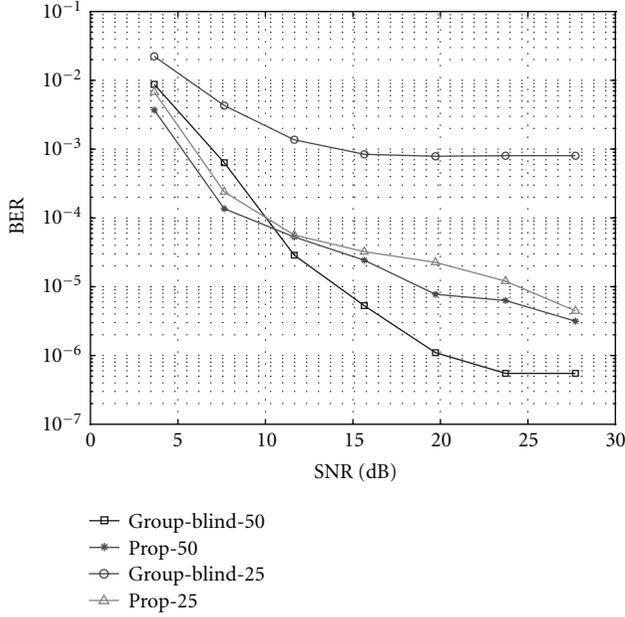


FIGURE 3: Small sample performance compared to the group-blind approach with known in-cell steering ( $K = 4, P = 8, M_{in} = 4, M = 6$ ).

detector clearly outperforms the linear group-blind decorrelating detector, despite the fact that it uses less side information. In both cases, the proposed detector outperforms the linear group-blind decorrelating detector in the low SNR regime. We emphasize that the proposed algorithm performs well even for very small sample sizes (e.g.,  $N = 10$ ) in the high SNR regime, whereas the group-blind approach hits the error floor at very low SNR in this case.

Our proposed detector is also robust to strong out-of-cell interference. We have compared the user 1's BER performance of proposed approach against the usual minimum mean squared error (MMSE) receiver, which assumes exact knowledge of the in-cell user codes and steering vectors, but treats out-of-cell users as Gaussian interference. The soft MMSE solution for  $\mathbf{S}$  is

$$\hat{\mathbf{S}}_{in}^T = \left( (\mathbf{C}_{in} \odot \mathbf{A}_{in})^H (\mathbf{C}_{in} \odot \mathbf{A}_{in}) + \frac{1}{\text{SNR}} \mathbf{I} \right)^{-1} (\mathbf{C}_{in} \odot \mathbf{A}_{in})^H \mathbf{X}^{KP \times N}. \quad (29)$$

Figure 4 shows that as the power of out-of-cell users increases, the performance of the MMSE receiver deteriorates significantly whereas the degradation of the proposed detector is marginal.

The proposed algorithm is capable of accurately estimating the steering matrix of all active users and the code matrix of out-of-cell users. In order to illustrate this, we compare the (mean squared error MSE) performance of the proposed approach against the associated asymptotic CRB. Throughout,

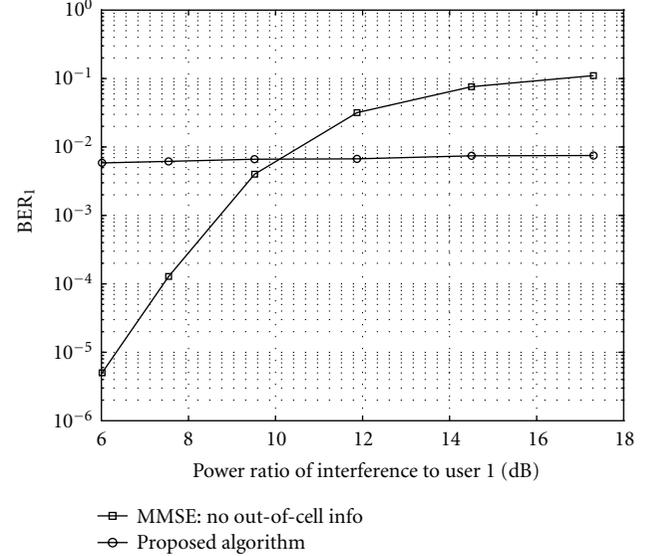


FIGURE 4: Robustness to strong out-of-cell interference ( $\text{SNR}_1 = 8$  dB,  $K = 4, N = 25, P = 8, M_{in} = 3, M = 4$ ).

the asymptotic CRB is first normalized in an elementwise fashion, that is, each unknown parameter's CRB is weighed with weight proportional to the inverse modulus square of respective parameter. The average weighted CRB of all the unknown parameters is then used as a single performance metric. The average MSE for all free model parameters is calculated in the same fashion. The SNR is defined as

$$\text{SNR} := 10 \log_{10} \frac{\|\mathbf{C} \odot \mathbf{A}\|_F^2}{KP\sigma^2} \text{ dB}, \quad (30)$$

which can be shown consistent with the definition (28) when we take the expectation of (28) with respect to  $\mathbf{S}$ .

Figure 5 depicts simulation results comparing TALS performance to this asymptotic CRB for two different snapshots. In this simulation,  $K = 4, P = 4, M = 6$ , and the true parameters were used to initialize TALS. The point here is to measure how tight the asymptotic CRB is for various  $N$ ; for this reason, we use the sought parameters as initialization in order to ensure the best possible scenario for TALS. It can be seen that TALS with good initialization remains very close to the CRB from medium to high SNR and relatively large sample size,  $N = 64$ . Note that  $N = 64$  is a reasonable number of symbol snapshots in practice. When the sample size is relatively small, the MSE performance of TALS is naturally worse than what is predicted by the asymptotic CRB.

Figure 6 presents the average MSE performance of COMFAC and the proposed algorithm against the CRB bound. We note that the performance of the proposed algorithm exceeds that of COMFAC considerably once SNR goes beyond the low SNR regime. This is because the new algebraic approach can provide fairly accurate initializations for CTALS,

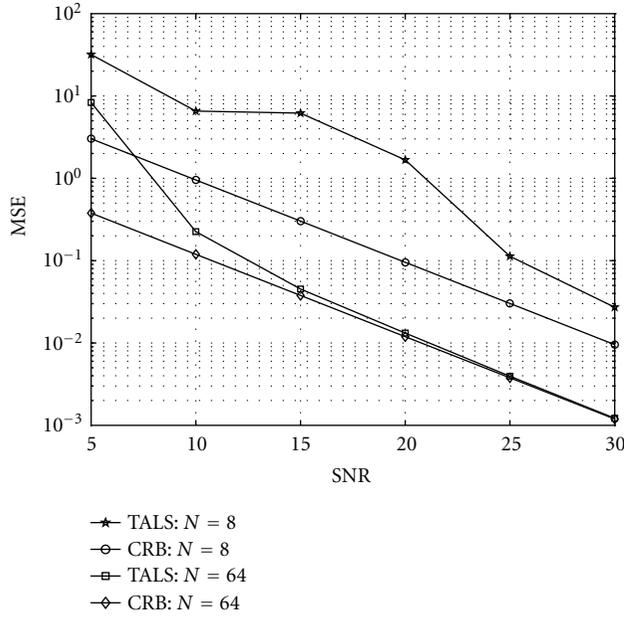
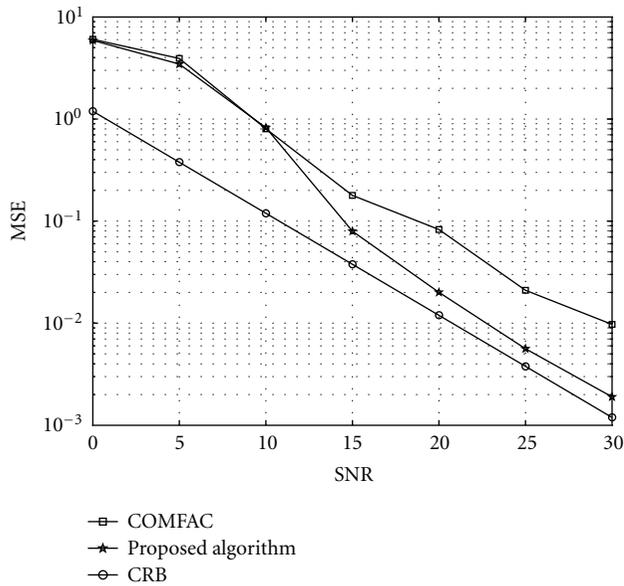


FIGURE 5: TALS versus asymptotic CRB.

FIGURE 6: MSE performance of COMFAC and the proposed algorithm versus asymptotic CRB ( $K = 4, P = 4, M_{in} = 4, M = 6, N = 64$ ).

whereas the COMFAC is forced to use random initializations in this case, wherein no two modes are full column rank. The average MSE of the proposed algorithm deviates from CRB about two to three dB. This is mainly because the initializations the algebraic approach provides are still not perfect, and the pre-specified tolerance threshold used to terminate the iterative refinement algorithm is set higher than in previous simulations, due to complexity considerations.

## 6. CONCLUSIONS

Out-of-cell interference in DS-CDMA systems is usually treated as noise, possibly mitigated using random cell codes. If the total number of in-cell plus out-of-cell users is smaller than the spreading gain, subspace-based suppression of out-of-cell users is possible. The assumption of more spreading than the total number of users can be quite unrealistic, even for moderately loaded cells. Completely blind reception is feasible under certain conditions (even with more users than spreading) with BS antenna arrays. We have proposed a new blind identification procedure that is capable of recovering both in-cell and out-of-cell transmissions, with sole knowledge of the in-cell user codes. The codes of the out-of-cell users and the steering vectors of all users are also recovered. The new procedure remains operational even when completely blind or subspace-based procedures fail. Interestingly, if the in-cell codes are known, then algebraic solution is possible.

## APPENDIX

### ASYMPTOTIC CRB AS $N$ TENDS TO INFINITY

To derive a meaningful CRB, following what has been done in [8], we assume that the first row of  $\mathbf{A}$  and  $\mathbf{S}$  is fixed (or normalized) to  $[1 \cdot \cdot \cdot 1]_{1 \times F}$  (this takes care of scale ambiguity), the first row of  $\mathbf{C}_{out}$  is known and consists of distinct elements (which subsequently resolves the permutation ambiguity) and  $\mathbf{C}_{in}$  is in *canonical form*. In turn, the number of unknown complex parameters is  $(N + K - 2)M + (P - 1)M_{out}$ . Let

$$\boldsymbol{\theta} := [\mathbf{a}_2^T; \dots; \mathbf{a}_K^T; \mathbf{c}_{out_2}^T; \dots; \mathbf{c}_{out_P}^T; \mathbf{s}_2^T; \dots; \mathbf{s}_N^T; \mathbf{a}_2^H; \dots; \mathbf{s}_N^H] \in \mathbb{C}^{(N+K-2)M+(P-1)M_{out} \times 1}, \quad (\text{A.1})$$

where  $\mathbf{a}_k$  denotes the  $k$ th row of  $\mathbf{A}$ ,  $\mathbf{c}_{out_p}$  denotes the  $i$ th row of  $\mathbf{C}_{out}$ , and  $\mathbf{s}_n$  denotes the  $n$ th row of  $\mathbf{S}$ .

It has been shown in [8] that the Fisher information matrix (FIM) is given by

$$\Omega(\boldsymbol{\theta}) = E \left\{ \left( \frac{\partial f(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)^H \left( \frac{\partial f(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) \right\} = \begin{bmatrix} \boldsymbol{\Psi} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Psi}^* \end{bmatrix}, \quad (\text{A.2})$$

where  $f(\boldsymbol{\theta})$  is the log-likelihood function and

$$\boldsymbol{\Psi} = \begin{bmatrix} \boldsymbol{\Psi}_{aa} & \boldsymbol{\Psi}_{ac} & \boldsymbol{\Psi}_{as} \\ \boldsymbol{\Psi}_{ac}^H & \boldsymbol{\Psi}_{cc} & \boldsymbol{\Psi}_{cs} \\ \boldsymbol{\Psi}_{as}^H & \boldsymbol{\Psi}_{cs}^H & \boldsymbol{\Psi}_{ss} \end{bmatrix} \quad (\text{A.3})$$

with obvious notation. In addition,

$$\begin{aligned} & \begin{bmatrix} \text{CRB}_{aa} & \text{CRB}_{ac} \\ \text{CRB}_{ac}^H & \text{CRB}_{cc} \end{bmatrix} \\ &= \left( \begin{bmatrix} \boldsymbol{\Psi}_{aa} & \boldsymbol{\Psi}_{ac} \\ \boldsymbol{\Psi}_{ac}^H & \boldsymbol{\Psi}_{cc} \end{bmatrix} - \begin{bmatrix} \boldsymbol{\Psi}_{as} \\ \boldsymbol{\Psi}_{cs} \end{bmatrix} \boldsymbol{\Psi}_{ss}^{-1} \begin{bmatrix} \boldsymbol{\Psi}_{as}^H & \boldsymbol{\Psi}_{cs}^H \end{bmatrix} \right)^{-1}. \end{aligned} \quad (\text{A.4})$$

The elements of  $\Psi$  can be given<sup>6</sup> as follows

$$\begin{aligned} E \left\{ \frac{\partial f(\boldsymbol{\theta})}{\partial a_{k_1, m_1}^*} \frac{\partial f(\boldsymbol{\theta})}{\partial a_{k_2, m_2}} \right\} &= \frac{1}{\sigma^2} \mathbf{e}_{m_1}^H \left( \sum_{n=1}^N (\mathbf{s}_n \odot \mathbf{C})^H (\mathbf{s}_n \odot \mathbf{C}) \right) \mathbf{e}_{m_2} \delta_{k_1, k_2} \\ &= \frac{1}{\sigma^2} \mathbf{e}_{m_1}^H \left( \sum_{n=1}^N \mathbf{s}_n^H \mathbf{s}_n \right) \diamond (\mathbf{C}^H \mathbf{C}) \mathbf{e}_{m_2} \delta_{k_1, k_2}, \end{aligned} \quad (\text{A.5})$$

$k_1, k_2 = 2, \dots, K, m_1, m_2 = 1, \dots, M,$

where we have used the identity

$$(\mathbf{C}^H \mathbf{C}) \diamond (\mathbf{D}^H \mathbf{D}) = (\mathbf{C} \odot \mathbf{D})^H (\mathbf{C} \odot \mathbf{D}), \quad (\text{A.6})$$

and  $\diamond$  stands for the Hadamard product. Similarly, we have

$$\begin{aligned} E \left\{ \frac{\partial f(\boldsymbol{\theta})}{\partial c_{p_1, m_1}^*} \frac{\partial f(\boldsymbol{\theta})}{\partial c_{p_2, m_2}} \right\} &= \frac{1}{\sigma^2} \mathbf{e}_{m_1}^H (\mathbf{A}^H \mathbf{A}) \diamond \left( \sum_{n=1}^N \mathbf{s}_n^H \mathbf{s}_n \right) \mathbf{e}_{m_2} \delta_{p_1, p_2} \\ & \quad p_1, p_2 = 2, \dots, P; m_1, m_2 = M_{\text{in}} + 1, \dots, M. \end{aligned} \quad (\text{A.7})$$

In addition, we have

$$\begin{aligned} E \left\{ \frac{\partial f(\boldsymbol{\theta})}{\partial s_{n_1, m_1}^*} \frac{\partial f(\boldsymbol{\theta})}{\partial s_{n_2, m_2}} \right\} &= \frac{1}{\sigma^2} \mathbf{e}_{m_1}^H (\mathbf{C} \odot \mathbf{A})^H (\mathbf{C} \odot \mathbf{A}) \mathbf{e}_{m_2} \delta_{n_1, n_2} \\ &= \frac{1}{\sigma^2} \mathbf{e}_{m_1}^H (\mathbf{C}^H \mathbf{C}) \diamond (\mathbf{A}^H \mathbf{A}) \mathbf{e}_{m_2} \delta_{n_1, n_2}, \end{aligned}$$

$n_1, n_2 = 2, \dots, N; m_1, m_2 = 1, \dots, M,$

$$\begin{aligned} E \left\{ \frac{\partial f(\boldsymbol{\theta})}{\partial a_{k, m_1}^*} \frac{\partial f(\boldsymbol{\theta})}{\partial c_{p, m_2}} \right\} &= \frac{1}{\sigma^2} \left( \sum_{n=1}^N \mathbf{s}_n^*(m_1) \mathbf{s}_n(m_2) \right) \mathbf{c}_p^*(m_1) \mathbf{a}_k(m_2), \end{aligned}$$

$$\begin{aligned} E \left\{ \frac{\partial f(\boldsymbol{\theta})}{\partial a_{k, m_1}^*} \frac{\partial f(\boldsymbol{\theta})}{\partial s_{n, m_2}} \right\} &= \frac{1}{\sigma^2} \mathbf{s}_n^*(m_1) \left( \sum_{p=1}^P \mathbf{c}_p^*(m_1) \mathbf{c}_p(m_2) \right) \mathbf{a}_k(m_2), \end{aligned}$$

$$\begin{aligned} E \left\{ \frac{\partial f(\boldsymbol{\theta})}{\partial c_{p, m_1}^*} \frac{\partial f(\boldsymbol{\theta})}{\partial s_{n, m_2}} \right\} &= \frac{1}{\sigma^2} \mathbf{s}_n^*(m_1) \left( \sum_{k=1}^K \mathbf{a}_k^*(m_1) \mathbf{a}_k(m_2) \right) \mathbf{c}_p(m_2). \end{aligned} \quad (\text{A.8})$$

Since we have assumed that

$$E(\mathbf{s}_{n_1}^*(m_1) \mathbf{s}_{n_2}(m_2)) = \delta_{n_1, n_2} \delta_{m_1, m_2}, \quad (\text{A.9})$$

it follows that

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} E \left\{ \frac{\partial f(\boldsymbol{\theta})}{\partial a_{k_1, m_1}^*} \frac{\partial f(\boldsymbol{\theta})}{\partial a_{k_2, m_2}} \right\} &= \frac{1}{\sigma^2} \mathbf{e}_{m_1}^H \left( \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mathbf{s}_n^H \mathbf{s}_n \right) \diamond (\mathbf{C}^H \mathbf{C}) \mathbf{e}_{m_2} \delta_{k_1, k_2} \\ &= \frac{1}{\sigma^2} \mathbf{e}_{m_1}^H \mathbf{I}_M \diamond (\mathbf{C}^H \mathbf{C}) \mathbf{e}_{m_2} \delta_{k_1, k_2}, \\ \lim_{N \rightarrow \infty} \frac{1}{N} E \left\{ \frac{\partial f(\boldsymbol{\theta})}{\partial c_{p_1, m_1}^*} \frac{\partial f(\boldsymbol{\theta})}{\partial c_{p_2, m_2}} \right\} &= \frac{1}{\sigma^2} \mathbf{e}_{m_1}^H (\mathbf{A}^H \mathbf{A}) \diamond \mathbf{I}_M \mathbf{e}_{m_2} \delta_{p_1, p_2} \\ \lim_{N \rightarrow \infty} \frac{1}{N} E \left\{ \frac{\partial f(\boldsymbol{\theta})}{\partial a_{k, m_1}^*} \frac{\partial f(\boldsymbol{\theta})}{\partial c_{p, m_2}} \right\} &= \frac{1}{\sigma^2} \left( \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mathbf{s}_n^*(m_1) \mathbf{s}_n(m_2) \right) \mathbf{c}_p^*(m_1) \mathbf{a}_k(m_2) \\ &= \frac{1}{\sigma^2} \mathbf{c}_p^*(m_1) \mathbf{a}_k(m_2) \delta_{m_1, m_2}, \end{aligned} \quad (\text{A.10})$$

hence

$$\lim_{N \rightarrow \infty} \frac{1}{N} \begin{bmatrix} \Psi_{aa} & \Psi_{ac} \\ \Psi_{ac}^H & \Psi_{cc} \end{bmatrix} = \frac{1}{\sigma^2} \begin{bmatrix} \Psi_{aa\text{limit}} & \Psi_{ac\text{limit}} \\ \Psi_{ac\text{limit}}^H & \Psi_{cc\text{limit}} \end{bmatrix} \quad (\text{A.11})$$

with obvious notation.

From (A.8), we know that

$$\begin{aligned} \Psi_{ss} &= \frac{1}{\sigma^2} \\ & \times \begin{bmatrix} (\mathbf{C}^H \mathbf{C}) \diamond (\mathbf{A}^H \mathbf{A}) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & (\mathbf{C}^H \mathbf{C}) \diamond (\mathbf{A}^H \mathbf{A}) & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & (\mathbf{C}^H \mathbf{C}) \diamond (\mathbf{A}^H \mathbf{A}) \end{bmatrix} \\ & \in \mathbb{C}^{(N-1)M \times (N-1)M}. \end{aligned} \quad (\text{A.12})$$

Let  $\mathbf{H} := ((\mathbf{C}^H \mathbf{C}) \diamond (\mathbf{A}^H \mathbf{A}))^{-1}$ , then

$$\Psi_{ss}^{-1} = \sigma^2 \begin{bmatrix} \mathbf{H} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{H} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{H} \end{bmatrix} \in \mathbb{C}^{(N-1)M \times (N-1)M}. \quad (\text{A.13})$$

Recall

$$\begin{aligned} E \left\{ \frac{\partial f(\boldsymbol{\theta})}{\partial a_{k, m_1}^*} \frac{\partial f(\boldsymbol{\theta})}{\partial s_{n, m_2}} \right\} &= \frac{1}{\sigma^2} \mathbf{s}_n^*(m_1) \left( \sum_{p=1}^P \mathbf{c}_p^*(m_1) \mathbf{c}_p(m_2) \right) \mathbf{a}_k(m_2), \\ E \left\{ \frac{\partial f(\boldsymbol{\theta})}{\partial c_{p, m_1}^*} \frac{\partial f(\boldsymbol{\theta})}{\partial s_{n, m_2}} \right\} &= \frac{1}{\sigma^2} \mathbf{s}_n^*(m_1) \left( \sum_{k=1}^K \mathbf{a}_k^*(m_1) \mathbf{a}_k(m_2) \right) \mathbf{c}_p(m_2), \end{aligned} \quad (\text{A.14})$$

<sup>6</sup>The forms given here can be shown to be mathematically equivalent to those in [8]. The new forms are computationally much simpler.

from which it is not difficult to see that

$$\begin{aligned} \begin{bmatrix} \Psi_{as} \\ \Psi_{cs} \end{bmatrix} &= \frac{1}{\sigma^2} [\mathbf{U}_2 \mathbf{R}, \mathbf{U}_3 \mathbf{R}, \dots, \mathbf{U}_N \mathbf{R}] \\ &\in \mathbb{C}^{((K-1)M+(P-1)M_{\text{out}}) \times (N-1)M}, \end{aligned} \quad (\text{A.15})$$

where

$$\begin{aligned} \mathbf{U}_n &= \text{diag} \left( \left[ \begin{array}{c} \overbrace{\mathbf{s}_n^*(1), \dots, \mathbf{s}_n^*(M)}^1, \dots, \overbrace{\mathbf{s}_n^*(1), \dots, \mathbf{s}_n^*(M)}^{K-1}, \\ \overbrace{\mathbf{s}_n^*(M_{\text{in}}+1), \dots, \mathbf{s}_n^*(M)}^1, \dots, \\ \overbrace{\mathbf{s}_n^*(M_{\text{in}}+1), \dots, \mathbf{s}_n^*(M)}^{P-1} \end{array} \right] \right) \\ &\in \mathbb{C}^{((K-1)M+(P-1)M_{\text{out}}) \times ((K-1)M+(P-1)M_{\text{out}})} \\ \mathbf{R} &= \left[ \begin{array}{c} \left( \left( \sum_{p=1}^P \mathbf{c}_p^*(m_1) \mathbf{c}_p(m_2) \right) \mathbf{a}_k(m_2) \right)_{\{(k-1)M+m_1, m_2\}} \\ \left( \left( \sum_{k=1}^K \mathbf{a}_k^*(m_1) \mathbf{a}_k(m_2) \right) \mathbf{c}_p(m_2) \right)_{\{(p-1)M_{\text{out}}+m_1, m_2\}} \end{array} \right] \\ &\in \mathbb{C}^{((K-1)M+(P-1)M_{\text{out}}) \times M}. \end{aligned} \quad (\text{A.16})$$

Let

$$\begin{aligned} \mathbf{G} &:= \begin{bmatrix} \Psi_{as} \\ \Psi_{cs} \end{bmatrix} \Psi_{ss}^{-1} [\Psi_{as}^H \Psi_{cs}^H] \\ &\in \mathbb{C}^{((K-1)M+(P-1)M_{\text{out}}) \times ((K-1)M+(P-1)M_{\text{out}})}, \end{aligned} \quad (\text{A.17})$$

then

$$\mathbf{G} = \frac{1}{\sigma^2} \sum_{n=2}^N \mathbf{U}_n \mathbf{R} \mathbf{R}^H \mathbf{U}_n^H. \quad (\text{A.18})$$

With  $\mathbf{Z} := \mathbf{R} \mathbf{R}^H$ , we have

$$\mathbf{G} = \frac{1}{\sigma^2} \sum_{n=2}^N \mathbf{U}_n \mathbf{Z} \mathbf{U}_n^H, \quad (\text{A.19})$$

and from

$$E(\mathbf{s}_{n_1}^*(m_1) \mathbf{s}_{n_2}(m_2)) = \delta_{n_1, n_2} \delta_{m_1, m_2}, \quad (\text{A.20})$$

we obtain

$$\lim_{N \rightarrow \infty} \frac{1}{N} \mathbf{G} = \frac{1}{\sigma^2} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=2}^N \mathbf{U}_n \mathbf{Z} \mathbf{U}_n^H = \frac{1}{\sigma^2} \mathbf{Z} \diamond \mathbf{Q}, \quad (\text{A.21})$$

where

$$\mathbf{Q} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=2}^N \text{diag}(\mathbf{U}_n) \text{diag}(\mathbf{U}_n)^H. \quad (\text{A.22})$$

Therefore, we have

$$\begin{aligned} &\begin{bmatrix} \text{CRB}_{aa} & \text{CRB}_{ac} \\ \text{CRB}_{ac}^H & \text{CRB}_{cc} \end{bmatrix}_{\text{limit}} \\ &= \lim_{N \rightarrow \infty} \begin{bmatrix} \text{CRB}_{aa} & \text{CRB}_{ac} \\ \text{CRB}_{ac}^H & \text{CRB}_{cc} \end{bmatrix} \\ &= \frac{1}{N} \left( \lim_{N \rightarrow \infty} \frac{1}{N} \begin{bmatrix} \Psi_{aa} & \Psi_{ac} \\ \Psi_{ac}^H & \Psi_{cc} \end{bmatrix} \right. \\ &\quad \left. - \lim_{N \rightarrow \infty} \frac{1}{N} \begin{bmatrix} \Psi_{as} \\ \Psi_{cs} \end{bmatrix} \Psi_{ss}^{-1} [\Psi_{as}^H \Psi_{cs}^H] \right)^{-1} \\ &= \frac{\sigma^2}{N} \left( \begin{bmatrix} \Psi_{ad\text{limit}} & \Psi_{ac\text{limit}} \\ \Psi_{ac\text{limit}}^H & \Psi_{cc\text{limit}} \end{bmatrix} - \mathbf{Z} \diamond \mathbf{Q} \right)^{-1}. \end{aligned} \quad (\text{A.23})$$

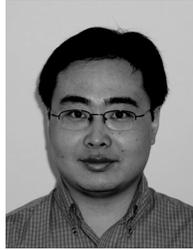
## ACKNOWLEDGMENTS

Preliminary version of part of this paper was presented at *ICASSP 2002*, Orlando, Fla. This work was supported by the Army Research Laboratory (ARL) through participation in the ARL Collaborative Technology Alliance (ARL-CTA) for Communications and Networks under Cooperative Agreement DADD19-01-2-0011, and the National Science Foundation (NSF) under Grants NSF/CAREER CCR-0096165 and NSF/Wireless IT & Networks CCR-0096164.

## REFERENCES

- [1] V. Shtrom, "CDMA vs. OFDM in broadband wireless access—the fundamental characteristics of OFDM make it ideally suited for broadband data," *Broadband Wireless Online*, vol. 3, no. 5, 2002.
- [2] H. Liu and G. Xu, "A subspace method for signature waveform estimation in synchronous CDMA systems," *IEEE Transactions on Communications*, vol. 44, no. 10, pp. 1346–1354, 1996.
- [3] P. Spasojevic, X. Wang, and A. Høst-Madsen, "Nonlinear group-blind multiuser detection," *IEEE Trans. Communications*, vol. 49, no. 9, pp. 1631–1641, 2001.
- [4] N. D. Sidiropoulos, G. B. Giannakis, and R. Bro, "Blind PARAFAC receivers for DS-CDMA systems," *IEEE Transactions on Signal Processing*, vol. 48, no. 3, pp. 810–823, 2000.
- [5] R. Harshman, "Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multi-mode factor analysis," *UCLA Working Papers in Phonetics*, vol. 16, pp. 1–84, 1970.
- [6] J. B. Kruskal, "Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics," *Linear Algebra and Its Applications*, vol. 18, no. 2, pp. 95–138, 1977.
- [7] T. Jiang, N. D. Sidiropoulos, and J. M. F. ten Berge, "Almost-sure identifiability of multidimensional harmonic retrieval," *IEEE Transactions on Signal Processing*, vol. 49, no. 9, pp. 1849–1859, 2001.
- [8] X. Liu and N. D. Sidiropoulos, "Cramer-Rao lower bounds for low-rank decomposition of multidimensional arrays," *IEEE Transactions on Signal Processing*, vol. 49, no. 9, pp. 2074–2086, 2001.
- [9] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramér-Rao bound," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 37, no. 5, pp. 720–741, 1989.

**Tao Jiang** received his B.S. degree from Peking University, Beijing, China, in 1997, and his M.S. degree from University of Minnesota, Minneapolis, in 2000, both in mathematics. He is currently working towards the Ph.D. degree in the Department of Electrical and Computer Engineering at University of Minnesota, Minneapolis. His research interests are in the area of signal processing for communications with focus on wireless communications.



**Nicholas D. Sidiropoulos** received his Diploma in electrical engineering from the Aristotelian University of Thessaloniki, Greece, and M.S. and Ph.D. degrees in electrical engineering from the University of Maryland at College Park (UMCP) in 1988, 1990, and 1992, respectively. He has been a Postdoctoral Fellow (1994–1995) and Research Scientist (1996–1997) at UMCP, Assistant Professor in the Department of Electrical Engineering, University of Virginia (1997–1999), and Associate Professor in the Department of Electrical and Computer Engineering (ECE), University of Minnesota (2000–2002). He is currently a Professor in the Department of Electronic and Computer Engineering, Technical University of Crete, Chania-Crete, Greece, and Adjunct Professor at the University of Minnesota. His current interests are primarily in SP for COM, and multiway analysis. He has published about 40 refereed journal papers, and his research has been funded by the US NSF, DARPA, ONR, ARL, ARO, and the European Commission (EC). He is a Senior Member of IEEE, a Member of the IEEE/SPS SPCOM TC, Associate Editor for IEEE TSP (2000–), and has served as Associate Editor for IEEE SPL (2000–2002). He received the NSF/CAREER Award in June 1998, and an IEEE SPS best paper award in 2001. He is an active consultant for industry in the areas of frequency hopping systems and signal processing for xDSL modems.



# Nonlinear Multiantenna Detection Methods

## Sheng Chen

School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK  
Email: [sqc@ecs.soton.ac.uk](mailto:sqc@ecs.soton.ac.uk)

## Lajos Hanzo

School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK  
Email: [lh@ecs.soton.ac.uk](mailto:lh@ecs.soton.ac.uk)

## Andreas Wolfgang

School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK  
Email: [aw03r@ecs.soton.ac.uk](mailto:aw03r@ecs.soton.ac.uk)

Received 29 May 2003; Revised 16 February 2004

A nonlinear detection technique designed for multiple-antenna assisted receivers employed in space-division multiple-access systems is investigated. We derive the optimal solution of the nonlinear spatial-processing assisted receiver for binary phase shift keying signalling, which we refer to as the Bayesian detector. It is shown that this optimal Bayesian receiver significantly outperforms the standard linear beamforming assisted receiver in terms of a reduced bit error rate, at the expense of an increased complexity, while the achievable system capacity is substantially enhanced with the advent of employing nonlinear detection. Specifically, when the spatial separation expressed in terms of the angle of arrival between the desired and interfering signals is below a certain threshold, a linear beamformer would fail to separate them, while a nonlinear detection assisted receiver is still capable of performing adequately. The adaptive implementation of the optimal Bayesian detector can be realized using a radial basis function network. Two techniques are presented for constructing block-data-based adaptive nonlinear multiple-antenna assisted receivers. One of them is based on the relevance vector machine invoked for classification, while the other on the orthogonal forward selection procedure combined with the Fisher ratio class-separability measure. A recursive sample-by-sample adaptation procedure is also proposed for training nonlinear detectors based on an amalgam of enhanced  $\kappa$ -means clustering techniques and the recursive least squares algorithm.

**Keywords and phrases:** smart antenna, adaptive beamforming, mean square error, bit error rate, Bayesian classification, radial basis function network.

## 1. INTRODUCTION

Spatial processing invoking adaptive antenna arrays has shown real promise in terms of attaining substantial capacity enhancements in mobile communication [1, 2, 3, 4, 5, 6, 7, 8]. Multiple-antenna aided receivers are capable of separating signals transmitted on the same carrier frequency, provided that signals are sufficiently separated in the spatial domain. Classically, beamforming algorithms create a linear combination of the signals received from the different elements of an antenna array. We refer to this classic beamforming principle as *linear* beamforming. A traditional approach to linear beamforming is based on the minimum mean square error (MMSE) principle that minimizes the mean square error (MSE) between the desired output generated from a known reference signal and the actual array output. Adaptive implementations of the linear MMSE (LMMSE) beamforming so-

lution can readily be realized using the well-known family of temporal reference techniques [2, 3, 9, 10, 11, 12, 13]. Specifically, block-data-based beamformer weight adaptation can be achieved using the sample matrix inversion (SMI) algorithm [9, 10], while sample-by-sample based array-weight adaptation can be carried out using the least mean square (LMS) algorithm [11, 12, 13]. Recent work [14, 15] has investigated a linear beamforming technique based directly on minimizing the system's bit error rate (BER) rather than the MSE and developed both block-data-based and sample-by-sample adaptive algorithms for implementing linear minimum BER (LMBER) beamforming. The results of [14, 15] have demonstrated that LMBER beamforming is capable of providing considerable performance gains in terms of a reduced BER over the usual LMMSE beamforming.

In the context of space-division multiple access (SDMA), the spatial separation in angles of arrival between the desired

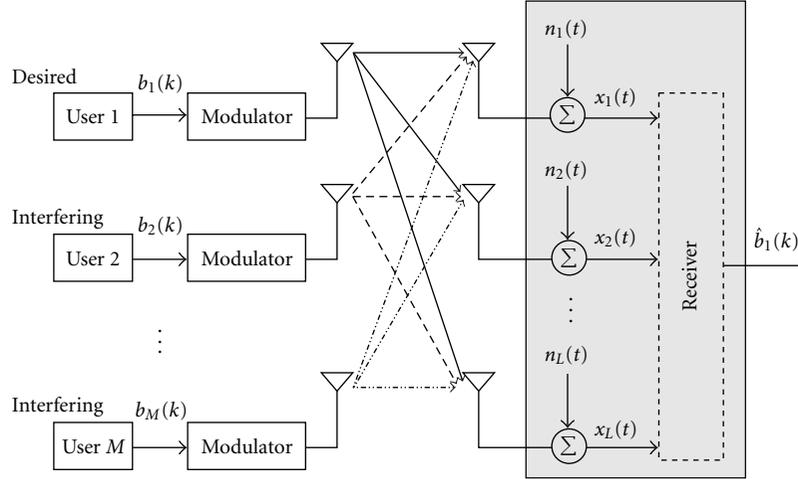


FIGURE 1: Multiantenna receiver configuration for the multiuser space-division multiple-access system.

signal and the closest interfering signal dominates the achievable system performance and hence the system's user capacity. When this angular separation is below a certain threshold, linear beamforming ultimately fails since the signals transmitted by the individual users become linearly inseparable, a situation that has also been observed in the context of single-user channel equalization and multiuser detection designed for code-division multiple access (CDMA) [16, 17, 18, 19, 20]. In fact, it has been observed even in linearly separable scenarios that a nonlinear processing technique is capable of providing a better performance than a linear one, although this is typically achieved at the cost of an increased complexity. In conjunction with nonlinear spatial processing, the achievable system capacity can be significantly increased since an adequate performance can be maintained even in case of a low angular separation compared to linear beamforming. These considerations motivate this study of nonlinear detection techniques contrived for multi-antenna aided systems.

The outline of the paper is as follows. Section 2 introduces the system model, while Section 3 outlines our linear beamforming-based benchmarker. In Section 4, we derive the optimal solution of the nonlinear spatial processing assisted receiver for binary phase shift keying (BPSK) signalling, which is referred to as the Bayesian detection solution. It is shown that this Bayesian solution has an identical form to a radial basis function (RBF) network [17, 21]. In Section 5, two schemes are proposed for realizing block-data-based adaptive RBF detectors. One of them is based on the relevance vector machine (RVM) invoked for classification [22, 23] and the other one is the orthogonal forward selection (OFS) procedure using the Fisher ratio class-separability measure [24]. Finally, in Section 6, an adaptive sample-by-sample implementation of the RBF detector is also considered using an amalgam of the enhanced  $\kappa$ -means clustering and the recursive least squares (CRLS) algorithm [19, 25] before offering our conclusions in Section 7.

## 2. SYSTEM MODEL

We consider the multiple-antenna aided receiver configuration of Figure 1 invoked for assisting the operation of a multiuser SDMA system. It is assumed that the system supports  $M$  users (signal sources), and each user transmits a BPSK modulated signal on the same carrier frequency of  $\omega = 2\pi f$ . Let  $k$  denote the bit instance. Then the baseband signal of user  $i$ , sampled at symbol rate, is given by

$$m_i(k) = A_i b_i(k), \quad 1 \leq i \leq M, \quad (1)$$

where the complex-valued coefficient  $A_i$  models the multiplication of the channel coefficient of user  $i$  with the transmitted signal power of user  $i$  and therefore  $|A_i|^2$  denotes the received signal power for user  $i$ , and  $b_i(k) \in \{\pm 1\}$  is the  $k$ th bit of user  $i$ . Without any loss of generality, source 1 is assumed to be the desired user and the rest of the sources are the interfering users. A linear antenna array is considered which consists of  $L$  uniformly spaced elements, and the signals received by the  $L$ -element antenna array are given by

$$x_l(k) = \sum_{i=1}^M m_i(k) \exp(j\omega t_l(\theta_i)) + n_l(k) = \bar{x}_l(k) + n_l(k) \quad (2)$$

for  $1 \leq l \leq L$ , where  $t_l(\theta_i)$  is the relative time delay at element  $l$  for source  $i$ ,  $\theta_i$  is the direction of arrival for source  $i$ , and  $n_l(k)$  is a complex-valued white Gaussian noise with zero mean and  $E[|n_l(k)|^2] = 2\sigma_n^2$ . The desired user's signal-to-noise ratio is defined as  $\text{SNR} = |A_1|^2/2\sigma_n^2$ , and the desired signal-to-interference ratio with respect to interfering user  $i$  is defined by  $\text{SIR}_i = |A_1|^2/|A_i|^2$  for  $i = 2, \dots, M$ . In vectorial form, the antenna array output  $\mathbf{x}(k) = [x_1(k) \ x_2(k) \ \cdots \ x_L(k)]^T$  can be expressed as

$$\mathbf{x}(k) = \bar{\mathbf{x}}(k) + \mathbf{n}(k) = \mathbf{P}\mathbf{b}(k) + \mathbf{n}(k), \quad (3)$$

where  $\mathbf{n}(k) = [n_1(k) \ n_2(k) \ \cdots \ n_L(k)]^T$  has a covariance matrix of  $E[\mathbf{n}(k)\mathbf{n}^H(k)] = 2\sigma_n^2\mathbf{I}_L$  with  $\mathbf{I}_L$  denoting the  $L \times L$

identity matrix, the system matrix  $\mathbf{P}$  is given by

$$\mathbf{P} = [A_1 \mathbf{s}_1 \ A_2 \mathbf{s}_2 \ \cdots \ A_M \mathbf{s}_M], \quad (4)$$

the steering vector for source  $i$  is formulated as

$$\mathbf{s}_i = [\exp(j\omega t_1(\theta_i)) \ \exp(j\omega t_2(\theta_i)) \ \cdots \ \exp(j\omega t_L(\theta_i))]^T, \quad (5)$$

and the transmitted bit vector is

$$\mathbf{b}(k) = [b_1(k) \ b_2(k) \ \cdots \ b_M(k)]^T. \quad (6)$$

The task of the spatial-processing assisted receiver is to provide an estimate  $\hat{b}_1(k)$  of the desired user's transmitted bit  $b_1(k)$ , given the input  $\mathbf{x}(k)$ . To keep our notations and the associated concepts relatively simple, we have used a BPSK modulation scheme, a narrowband channel model, and narrowband beamforming (space-only processing). The approach can be extended to other modulation schemes and wideband channels that induce intersymbol interference. The same idea can also be applied to broadband beamforming (space-time processing).

### 3. LINEAR BEAMFORMING ASSISTED RECEIVER

The output of the linear beamformer is given by

$$y(k) = \mathbf{w}^H \mathbf{x}(k) = \mathbf{w}^H \bar{\mathbf{x}}(k) + \mathbf{w}^H \mathbf{n}(k) = \bar{y}(k) + e(k), \quad (7)$$

where  $\mathbf{w} = [w_1 \ w_2 \ \cdots \ w_L]^T$  is the complex-valued beamformer weight vector, and  $e(k)$  is Gaussian distributed with a zero mean and a variance  $E[|e(k)|^2] = 2\sigma_n^2 \mathbf{w}^H \mathbf{w}$ . The estimate of the transmitted bit  $b_1(k)$  is given by

$$\hat{b}_1(k) = \text{sgn}(y_R(k)) = \begin{cases} +1, & y_R(k) > 0, \\ -1, & y_R(k) \leq 0, \end{cases} \quad (8)$$

where  $y_R(k) = \Re[y(k)]$  denotes the real part of  $y(k)$ . Classically, the linear beamformer's weight vector is determined by minimizing the MSE term of  $E[|b_1(k) - y(k)|^2]$  between the desired user's transmitted bit and the beamformer's output, which leads to the following LMMSE solution:

$$\mathbf{w}_{\text{LMMSE}} = (\mathbf{P}\mathbf{P}^H + 2\sigma_n^2 \mathbf{I}_L)^{-1} \mathbf{p}_1 \quad (9)$$

with  $\mathbf{p}_1$  being the first column of  $\mathbf{P}$ . Using a temporal reference technique aided approach [7], the LMMSE beamforming solution can be readily realized using the block-data-based SMI algorithm [7], and recursive sample-by-sample adaptation can be performed using the LMS or RLS algorithm [21].

In order to derive the BER formula of the linear beamformer with the weight vector  $\mathbf{w}$ , firstly note that there are  $N_b = 2^M$  possible sequences of  $\mathbf{b}(k)$ , which are denoted as  $\mathbf{b}_q$ ,  $1 \leq q \leq N_b$ . Furthermore, denote the first element of  $\mathbf{b}_q$ , corresponding to the desired user, as  $b_{q,1}$ . As expected, the

noiseless part of the beamformer input signal,  $\bar{\mathbf{x}}(k)$ , assumes encountering values only from the signal set defined as

$$\mathcal{X} \triangleq \{\bar{\mathbf{x}}_q = \mathbf{P}\mathbf{b}_q, \ 1 \leq q \leq N_b\}. \quad (10)$$

This set can be partitioned into two subsets depending on the specific value of  $b_1(k)$  as follows:

$$\mathcal{X}^{(\pm)} \triangleq \{\bar{\mathbf{x}}_q^{(\pm)} \in \mathcal{X} : b_1(k) = \pm 1\}. \quad (11)$$

Similarly,  $\bar{y}(k)$  takes values from the scalar set

$$\mathcal{Y} \triangleq \{\bar{y}_q = \mathbf{w}^H \bar{\mathbf{x}}_q, \ 1 \leq q \leq N_b\} \quad (12)$$

which can be divided into the two subsets defined as

$$\mathcal{Y}^{(\pm)} \triangleq \{\bar{y}_q^{(\pm)} \in \mathcal{Y} : b_1(k) = \pm 1\}. \quad (13)$$

Thus,  $\bar{y}_R(k)$  can only take values from the set

$$\mathcal{Y}_R \triangleq \{\bar{y}_{R,q} = \Re[\bar{y}_q], \ 1 \leq q \leq N_b\} \quad (14)$$

which can be partitioned into the two subsets conditioned on the value of  $b_1(k)$ :

$$\mathcal{Y}_R^{(\pm)} \triangleq \{\bar{y}_{R,q}^{(\pm)} \in \mathcal{Y}_R : b_1(k) = \pm 1\}. \quad (15)$$

It can be readily seen that the conditional probability density function (pdf) of  $y(k)$  given  $b_1(k) = +1$  is a Gaussian mixture given by

$$p(y|+1) = \frac{1}{N_{sb}} \sum_{q=1}^{N_{sb}} \frac{1}{2\pi\sigma_n^2 \mathbf{w}^H \mathbf{w}} \exp\left(-\frac{|y - \bar{y}_q^{(+)}|^2}{2\sigma_n^2 \mathbf{w}^H \mathbf{w}}\right), \quad (16)$$

where  $\bar{y}_q^{(+)} \in \mathcal{Y}^{(+)}$  and  $N_{sb} = N_b/2$  is the number of the points in  $\mathcal{Y}^{(+)}$ . Therefore, the conditional marginal pdf of  $y_R(k)$  given  $b_1(k) = +1$  is formulated as follows:

$$p(y_R|+1) = \frac{1}{N_{sb}} \sum_{q=1}^{N_{sb}} \frac{1}{\sqrt{2\pi\sigma_n^2 \mathbf{w}^H \mathbf{w}}} \exp\left(-\frac{(y_R - \bar{y}_{R,q}^{(+)})^2}{2\sigma_n^2 \mathbf{w}^H \mathbf{w}}\right), \quad (17)$$

where  $\bar{y}_{R,q}^{(+)} \in \mathcal{Y}_R^{(+)}$ . Thus, it can be shown that the BER of the linear beamformer associated with the weight vector  $\mathbf{w}$  is given by [14, 15]

$$P_E(\mathbf{w}) = \frac{1}{N_{sb}} \sum_{q=1}^{N_{sb}} Q(g_{q,+}(\mathbf{w})), \quad (18)$$

where

$$Q(u) = \frac{1}{\sqrt{2\pi}} \int_u^\infty \exp\left(-\frac{v^2}{2}\right) dv, \quad (19)$$

$$g_{q,+}(\mathbf{w}) = \frac{\text{sgn}(b_{q,1}) \bar{y}_{R,q}^{(+)}}{\sigma_n \sqrt{\mathbf{w}^H \mathbf{w}}} = \frac{\text{sgn}(b_{q,1}) \Re[\mathbf{w}^H \bar{\mathbf{x}}_q^{(+)}]}{\sigma_n \sqrt{\mathbf{w}^H \mathbf{w}}}.$$

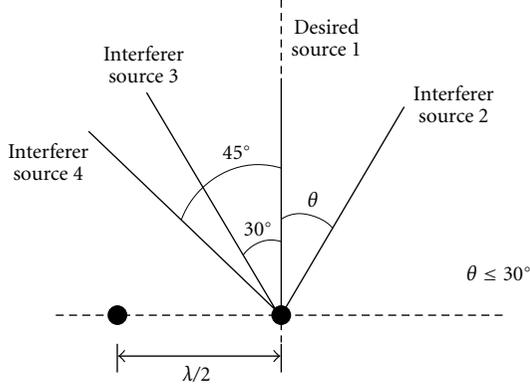


FIGURE 2: Locations of the desired source and the interfering sources with respect to the two-element linear antenna array having  $\lambda/2$  element spacing, where  $\lambda$  is the wavelength.

The LMBER beamforming solution is then defined as follows:

$$\mathbf{w}_{\text{MBER}} = \arg \min_{\mathbf{w}} P_E(\mathbf{w}). \quad (20)$$

Unlike the LMMSE solution (9), there exists no closed-form LMBER solution. In [14, 15], a simplified conjugate gradient method [26, 27] is used to obtain numerical solutions. Both the block-data-based gradient and LMS-style stochastic gradient adaptive algorithms have been derived in [14, 15] to realize the LMBER beamforming solution.

For the linear beamformer to work adequately, the underlying system must be linearly separable. The linear separability means that there exists a weight vector  $\mathbf{w}$  such that  $\mathcal{Y}_R^{(-)}$  and  $\mathcal{Y}_R^{(+)}$  are completely separated by the decision threshold  $y_R = 0$ . When the minimum spatial separation expressed in angles of arrival between the desired user and interfering users is below a certain threshold, the system inevitably becomes linearly inseparable. In such a situation, the linear beamformer will have a high irreducible BER floor, and nonlinear processing has to be adopted for the sake of achieving an adequate BER performance. In general, nonlinear spatial processing is capable of achieving a better performance than a linear receiver, regardless whether the output of the system is linearly separable or not. The limitation of a linear beamforming assisted receiver is illustrated in the following example, which is also used throughout this paper for investigating the proposed nonlinear multiantenna detection techniques.

#### Simulation example

The example consisted of four signal sources and a two-element antenna array. Figure 2 shows the locations of the desired source and the three interfering sources in a graphical form. The simulated channel conditions were  $A_i = 1 + j0$ ,  $1 \leq i \leq 4$ . The desired user and all the three interfering users had equal signal power, and therefore we had  $\text{SIR}_i = 0$  dB for  $i = 2, 3, 4$ . The minimum spatial separation in this example was the difference in angles of arrival between the desired user 1 and the interferer 2, which was  $\theta \leq 30^\circ$ . Figure 3 com-

pare the BERs of the LMMSE and LMBER beamformers for the two cases of  $\theta = 30^\circ$  and  $\theta = 10^\circ$ , respectively. It can be seen from Figure 3a that for  $\theta = 30^\circ$ , the underlying system scenario was linearly separable as was confirmed by the performance of the LMBER beamformer, while the LMMSE beamformer was unable to achieve the linear separability of the signal constellation and hence exhibited a high BER floor. Figure 4 plots the conditional pdfs  $p(y|+1)$ , the conditional marginal pdfs  $p(y_R|+1)$ , and the conditional subsets  $\mathcal{Y}^{(+)}$  and  $\mathcal{Y}_R^{(+)}$  for the LMMSE and LMBER beamformers, given  $\theta = 30^\circ$  and  $\text{SNR} = 10$  dB, which represented a typical condition in Figure 3a. It is clearly seen from Figure 4 that the LMBER beamformer was “smarter” than the LMMSE scheme and hence achieved the desired linear separability. However, when the minimum spatial separation was reduced to  $\theta = 10^\circ$ , the system became inherently linearly inseparable, and any linear beamformer failed to perform adequately as can be seen in Figure 3b. Figure 5 depicts the conditional pdfs  $p(y|+1)$ , the conditional marginal pdfs  $p(y_R|+1)$ , and the conditional subsets  $\mathcal{Y}^{(+)}$  and  $\mathcal{Y}_R^{(+)}$  for the LMMSE and LMBER beamformers, given  $\theta = 10^\circ$  and  $\text{SNR} = 10$  dB, which provided a typical condition in Figure 3b. The results of Figure 5 confirm that the underlying system was linearly inseparable, and it also explains why the LMBER solution did better than the LMMSE scheme, resulting in a lower BER floor. This example clearly demonstrates the need for invoking a nonlinear spatial-processing assisted receiver structure.

#### 4. BAYESIAN DETECTION SCHEME

Given the observation vector  $\mathbf{x}(k)$ , the optimal solution to the multiantenna aided spatial processing problem in terms of the achievable BER is the maximum a posteriori probability solution, which is similar to the case of single-user channel equalization [17, 18], and therefore can readily be formulated. The posterior probabilities or decision variables for  $b_1(k) = \pm 1$  given  $\mathbf{x}(k)$  are given by

$$\eta^{(\pm)}(k) = \sum_{q=1}^{N_{sb}} \frac{\xi_q^{(\pm)}}{(2\pi\sigma_n^2)^L} \exp\left(-\frac{\|\mathbf{x}(k) - \bar{\mathbf{x}}_q^{(\pm)}\|^2}{2\sigma_n^2}\right), \quad (21)$$

where  $\xi_q^{(\pm)}$  are a priori probabilities of  $\bar{\mathbf{x}}_q^{(\pm)}$  and  $\|\mathbf{x}\|^2 = \mathbf{x}^H \mathbf{x}$ . Typically, all the states  $\bar{\mathbf{x}}_q^{(\pm)}$  are equiprobable, and thus we have  $\xi_q^{(\pm)} = 1/N_b$ . The optimal decision regarding the transmitted bit  $b_1(k)$  is given by

$$\hat{b}_1(k) = \begin{cases} +1, & \eta^{(+)}(k) \geq \eta^{(-)}(k), \\ -1, & \text{otherwise.} \end{cases} \quad (22)$$

We redefine a single decision variable as

$$y_B(k) = \sum_{q=1}^{N_b} v_q \exp\left(-\frac{\|\mathbf{x}(k) - \bar{\mathbf{x}}_q\|^2}{2\sigma_n^2}\right), \quad (23)$$

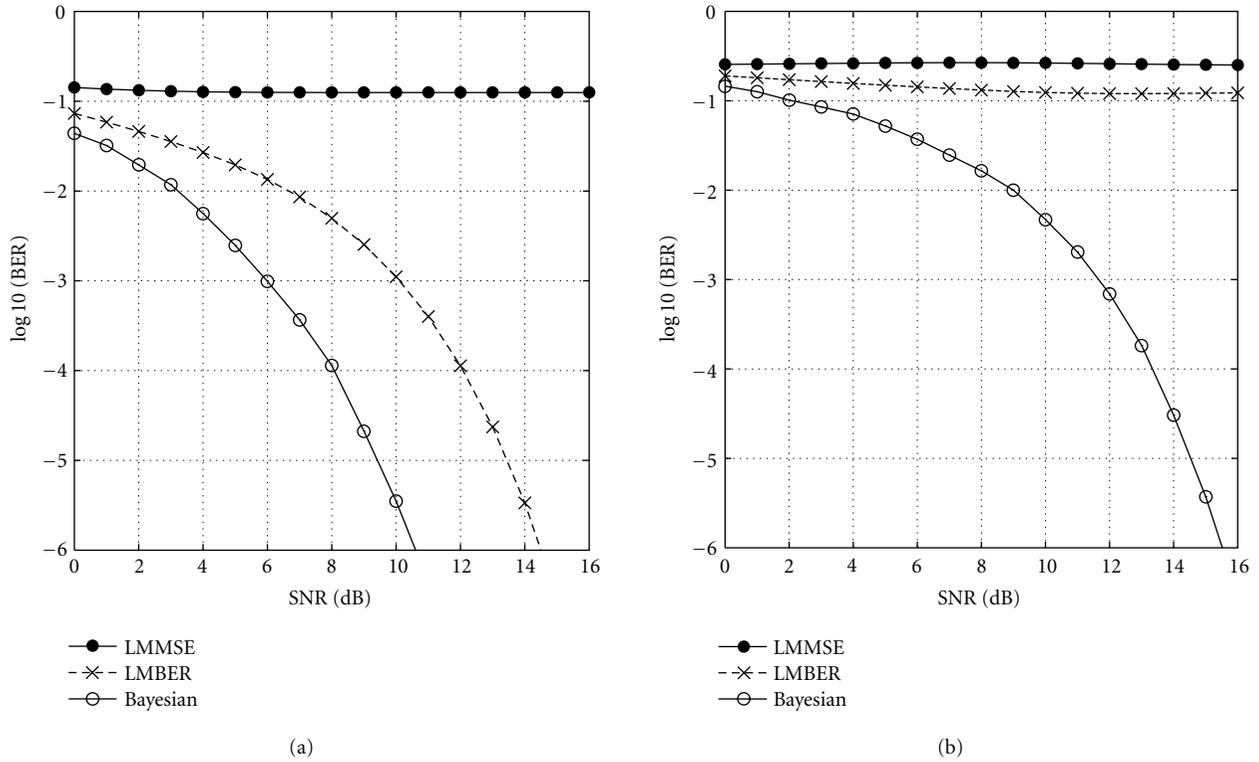


FIGURE 3: Comparison of the bit error rates of three theoretical detection schemes: the LMMSE and LMBER beamformers, and the optimal Bayesian detector. (a)  $\theta = 30^\circ$ . (b)  $\theta = 10^\circ$ .

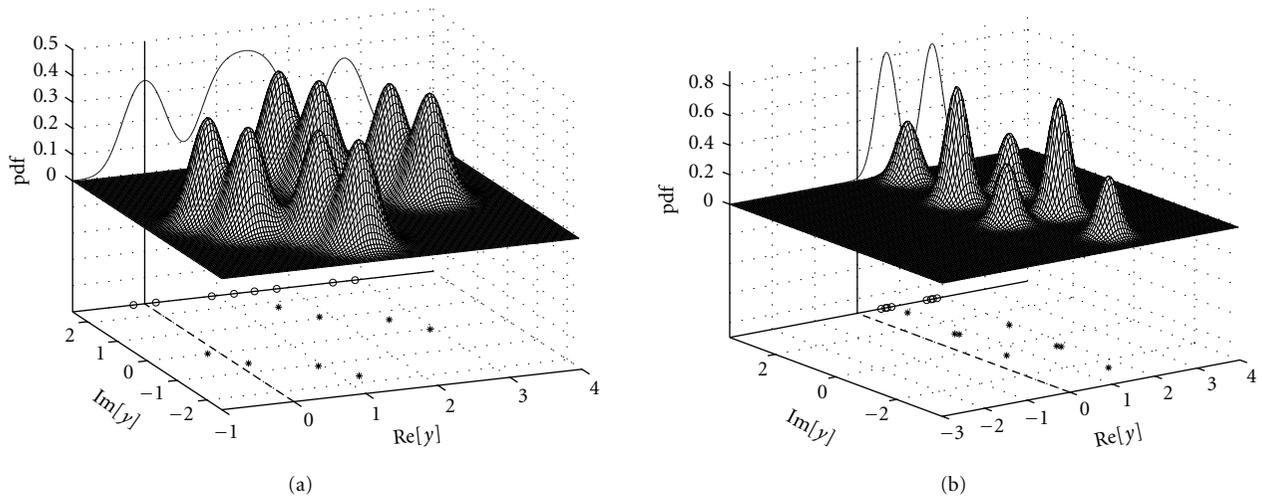


FIGURE 4: Conditional pdfs  $p(y|+1)$  (surface), conditional marginal pdfs  $p(y_R|+1)$  (curve), and conditional subsets  $\mathcal{Y}^{(+)}$  (symbol  $*$ ) and  $\mathcal{Y}_R^{(+)}$  (symbol  $\circ$ ), given  $\theta = 30^\circ$  and SNR = 10 dB. Beamformer weight vector has been normalized to a unit length. (a) LMMSE beamformer. (b) LMBER beamformer.

where

$$v_q = \frac{\text{sgn}(b_{q,1})}{N_b(2\pi\sigma_n^2)^L} \quad (24)$$

Then the optimal decision (22) is equivalent to

$$\hat{b}_1(k) = \text{sgn}(y_B(k)) = \begin{cases} +1, & y_B(k) \geq 0, \\ -1, & y_B(k) < 0. \end{cases} \quad (25)$$

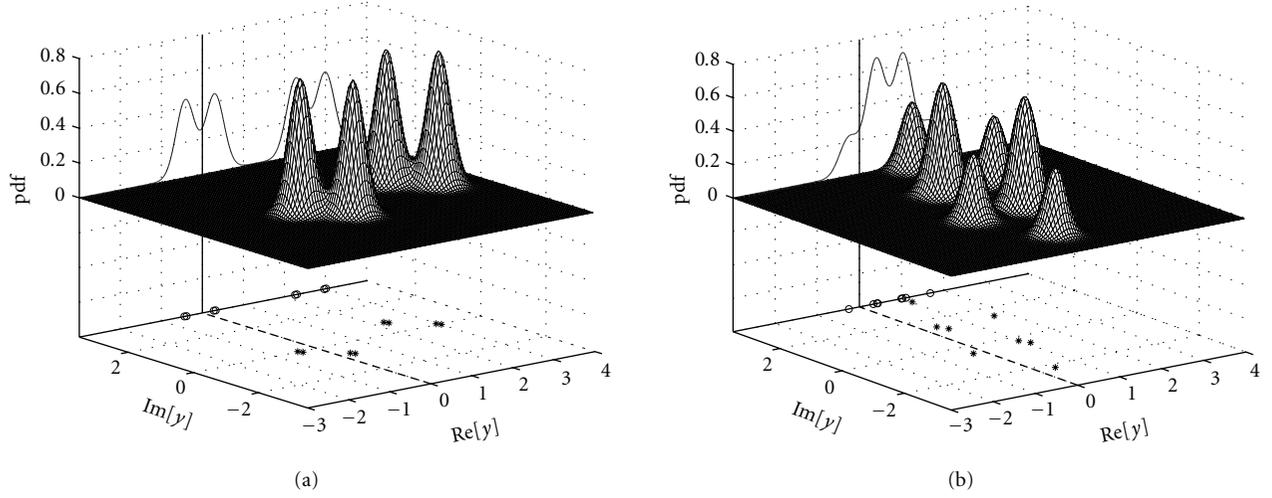


FIGURE 5: Conditional pdfs  $p(y|+1)$  (surface), conditional marginal pdfs  $p(y_R|+1)$  (curve), and conditional subsets  $\mathcal{Y}^{(+)}$  (symbol  $*$ ) and  $\mathcal{Y}_R^{(+)}$  (symbol  $\circ$ ), given  $\theta = 10^\circ$  and SNR = 10 dB. Beamformer weight vector has been normalized to a unit length. (a) LMMSE beamformer. (b) LMBER beamformer.

Note that (23) has the exact form of the RBF network in conjunction with a Gaussian kernel function.

The BER performance of the optimal Bayesian detection scheme was evaluated using the simulation example of Section 3 under the two conditions of having minimum spatial separations of  $\theta = 30^\circ$  and  $\theta = 10^\circ$ , and the results are plotted in Figures 3a and 3b, respectively, in comparison to the BERs of linear beamformers. It can be seen from Figure 3a that the Bayesian detector achieved an SNR improvement of 4 dB at the BER of  $10^{-4}$  over the LMBER beamformer. In the linearly inseparable case, the achievable performance improvement over the linear beamformer was even greater. In particular, Figure 3b shows that the Bayesian spatial processing assisted receiver removed the irreducible BER that was experienced by the linear beamforming aided receiver. The Bayesian detection scheme (23) may be viewed as a nonlinear “beamforming” process, and this nonlinear beamformer is clearly more complex than the simple linear beamformer (7). Therefore, the performance improvement achieved by the Bayesian detection scheme is attained at the expense of considerably increased computational complexity.

## 5. BLOCK-DATA KERNEL-BASED NONLINEAR DETECTOR CONSTRUCTION

In reality, the signal subsets  $\mathcal{X}^{(\pm)}$  are unknown and have to be estimated in order to realize the Bayesian solution. We will adopt a temporal reference technique to construct a nonlinear detector. Given a block of  $N$  training data  $\{\mathbf{x}(k), b_1(k)\}_{k=1}^N$ , consider the nonlinear detector of the form

$$y(\mathbf{x}) = \sum_{l=1}^N \beta_l \phi_l(\mathbf{x}), \quad (26)$$

where  $\beta_l$  represents the real-valued weights and  $\phi_l(\mathbf{x}) = \phi(\mathbf{x}, \mathbf{x}(l))$  are the appropriately chosen kernel basis functions with  $\mathbf{x}(l)$  denoting the  $l$ th training input. In our spatial processing aided application,  $\phi(\cdot, \cdot)$  can be chosen as the Gaussian kernel function of the form

$$\phi(\mathbf{x}, \mathbf{x}(l)) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}(l)\|^2}{2\rho^2}\right), \quad (27)$$

where the kernel variance  $\rho^2$  is an estimate of the noise variance  $\sigma_n^2$ . Define the modelling residual as

$$\epsilon(k) = t(k) - y(k) = b_1(k) - y(\mathbf{x}(k)). \quad (28)$$

Then the kernel model (26) generated for the training data set can be formulated as

$$\mathbf{t} = \Phi\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad (29)$$

where the target vector  $\mathbf{t}$  is defined as

$$\begin{aligned} \mathbf{t} &= [t(1) \ t(2) \ \cdots \ t(N)]^T \\ &= [b_1(1) \ b_1(2) \ \cdots \ b_1(N)]^T, \end{aligned} \quad (30)$$

the kernel weight vector is given by  $\boldsymbol{\beta} = [\beta_1 \ \beta_2 \ \cdots \ \beta_N]^T$ , the residual vector is formulated as  $\boldsymbol{\epsilon} = [\epsilon(1) \ \epsilon(2) \ \cdots \ \epsilon(N)]^T$ , and the regression matrix  $\Phi$  is given by

$$\Phi = [\phi_1 \ \phi_2 \ \cdots \ \phi_N] \quad (31)$$

with

$$\begin{aligned} \phi_i &= [\phi_i(1) \ \phi_i(2) \ \cdots \ \phi_i(N)]^T \\ &= [\phi(\mathbf{x}(1), \mathbf{x}(i)) \ \phi(\mathbf{x}(2), \mathbf{x}(i)) \ \cdots \ \phi(\mathbf{x}(N), \mathbf{x}(i))]^T, \end{aligned} \quad (32)$$

for  $1 \leq i \leq N$ . We adopt two different techniques for constructing a sparse detector model having  $N_{\text{spa}} (\ll N)$  number of terms from the full model (26).

### 5.1. Relevance vector machine for sparse kernel detector construction

The RVM method [22, 23] can readily be applied for constructing a sparse kernel model having  $N_{\text{spa}}$  number of terms from the full model (26). The introduction of an individual hyperparameter  $\alpha_i$  for every weight  $\beta_i$  of the model (26) is the key feature of the RVM, and is ultimately responsible for the sparsity properties of the RVM method [22]. During the optimization process, many of the  $\alpha_i$  coefficients are driven to large values so that the corresponding model weights  $\beta_i$  are effectively pruned out. Thus the corresponding model terms  $\phi_i(\cdot)$  can be removed from the trained model. The construction procedure produces a beamformer having a sparse final kernel structure consisting of  $N_{\text{spa}}$  number of significant terms. The detailed RVM method used is summarized in Appendix A.

The RVM method is known to be able to produce very sparse models while exhibiting excellent generalization capabilities [22]. A drawback of the RVM method is its high computational complexity. The algorithm contains two loops, with the inner loop used for updating the kernel weights and the outer loop for the associated hyperparameters (see Appendix A). Both loops involve “expensive” nonlinear optimization, and therefore converge relatively slowly, while incurring high computational costs. Furthermore, the RVM method starts with the full model set  $\Phi$  and removes those kernel terms that have large values in their associated hyperparameters. In other words, it is based on the backward elimination principle. Since the Hessian matrix  $\mathbf{H}$  associated with the full model set ((A.8) in Appendix A) is typically ill-conditioned and may even be non invertible, the RVM method is inherently ill-conditioned and its iterative procedure may converge at a slow rate, requiring numerous iterations. The threshold  $Lg$  employed by the pruning process (see Appendix A) is problem-dependent and has to be determined empirically. Provided that the value of  $Lg$  is tuned appropriately, the RVM algorithm is in general capable of identifying a sparse detector from the full model (26), which closely approximates the Bayesian performance.

### 5.2. Orthogonal forward selection with Fisher ratio class-separability measure for sparse kernel detector construction

An alternative way of constructing a sparse kernel model from the full model (26) is offered by the OFS procedure based on Fisher ratio class-separability measure [24], which is computationally attractive and numerically very robust. Let an orthogonal decomposition of the regression matrix  $\Phi$  be

$$\Phi = \mathbf{U}\mathbf{D}, \quad (33)$$

where

$$\mathbf{D} = \begin{bmatrix} 1 & d_{1,2} & \cdots & d_{1,N} \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & d_{N-1,N} \\ 0 & \cdots & 0 & 1 \end{bmatrix}, \quad (34)$$

$$\mathbf{U} = [\mathbf{u}_1 \ \mathbf{u}_2 \ \cdots \ \mathbf{u}_N] = \begin{bmatrix} u_{1,1} & u_{1,2} & \cdots & u_{1,N} \\ u_{2,1} & u_{2,2} & \cdots & u_{2,N} \\ \vdots & \vdots & \vdots & \vdots \\ u_{N,1} & u_{N,2} & \cdots & u_{N,N} \end{bmatrix},$$

with orthogonal columns that satisfy  $\mathbf{u}_i^T \mathbf{u}_q = 0$  if  $i \neq q$ . The kernel model (29) can alternatively be expressed as

$$\mathbf{t} = \mathbf{U}\mathbf{g} + \boldsymbol{\epsilon}, \quad (35)$$

where the orthogonal weight vector  $\mathbf{g} = [g_1 \ g_2 \ \cdots \ g_N]^T$  satisfies the triangular system  $\mathbf{D}\mathbf{g} = \mathbf{g}$ .

A sparse  $N_{\text{spa}}$ -term model can be selected by incrementally maximizing a class separability measure in an OFS procedure, as is presented in [24]. Define the two class sets  $\mathbf{X}_{\pm} = \{\mathbf{x}(k) : d(k) = \pm 1\}$ , and let the numbers of points in  $\mathbf{X}_{\pm}$  be  $N_{\pm}$ , respectively, with  $N_+ + N_- = N$ . The means and variances of training samples belonging to class  $\mathbf{X}_+$  and class  $\mathbf{X}_-$  in the direction of basis  $\mathbf{u}_l$  are given by

$$m_{+,l} = \frac{1}{N_+} \sum_{i=1}^N \delta(t(i) - 1) u_{i,l},$$

$$\sigma_{+,l}^2 = \frac{1}{N_+} \sum_{i=1}^N \delta(t(i) - 1) (u_{i,l} - m_{+,l})^2,$$

$$m_{-,l} = \frac{1}{N_-} \sum_{i=1}^N \delta(t(i) + 1) u_{i,l},$$

$$\sigma_{-,l}^2 = \frac{1}{N_-} \sum_{i=1}^N \delta(t(i) + 1) (u_{i,l} - m_{-,l})^2, \quad (36)$$

respectively, where  $\delta(x) = 1$  for  $x = 0$  and  $\delta(x) = 0$  for  $x \neq 0$ . Fisher ratio is defined as the ratio of the interclass difference and the intraclass spread encountered in the direction of  $\mathbf{u}_l$ , which is given by [28]

$$F_l = \frac{(m_{+,l} - m_{-,l})^2}{\sigma_{+,l}^2 + \sigma_{-,l}^2}. \quad (37)$$

Based on this Fisher ratio for class separability measure, significant kernel terms can be selected with the aid of an OFS procedure. At the  $l$ th stage, a term is chosen as the  $l$ th term in the selected model if it produces the largest  $F_l$  among the candidate terms  $\mathbf{u}_i$ ,  $l \leq i \leq N$ . The procedure is terminated with a sparse  $N_{\text{spa}}$ -term model when we have

$$\frac{F_{N_{\text{spa}}}}{\sum_{l=1}^{N_{\text{spa}}} F_l} < \xi, \quad (38)$$

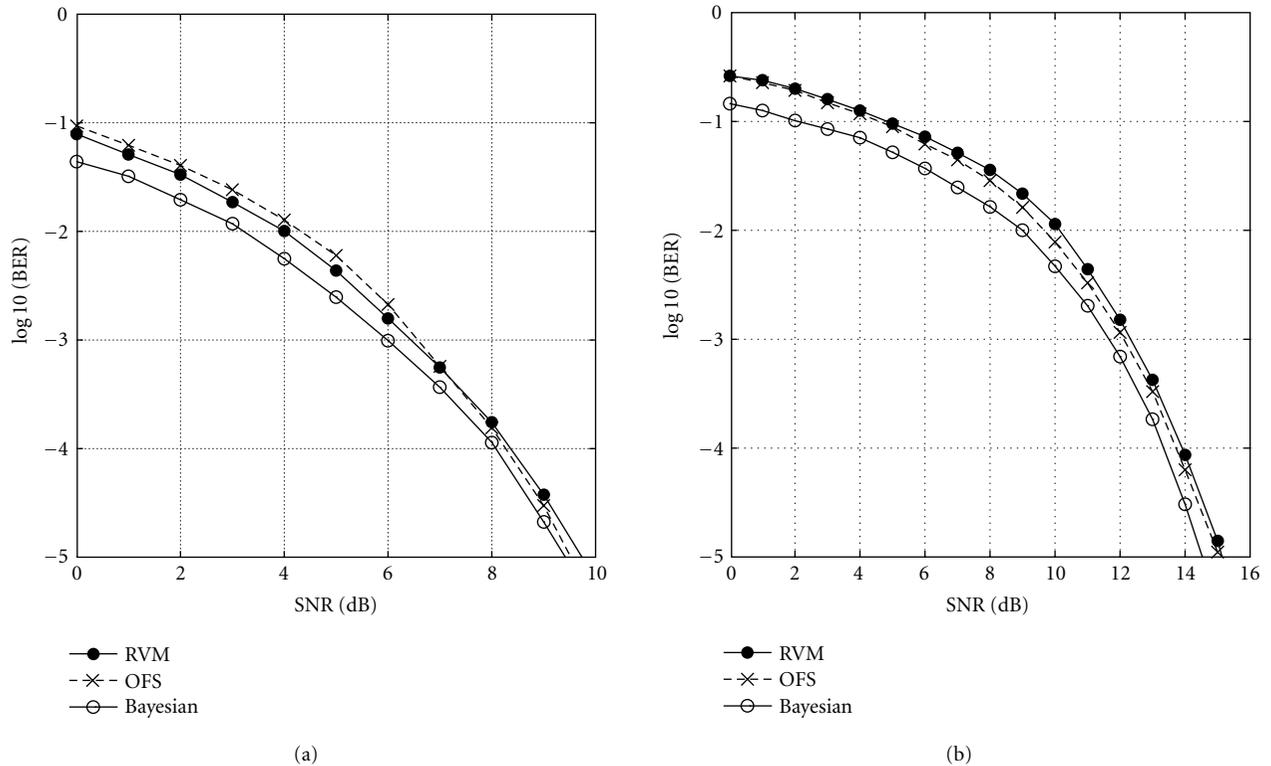


FIGURE 6: Performance comparison of the Bayesian detector with the RBF detectors constructed by the RVM algorithm and the OFS with Fisher ratio, respectively. (a)  $\theta = 30^\circ$ . (b)  $\theta = 10^\circ$ .

where the threshold  $\xi$  determines the sparsity of the selected model. The appropriate value for  $\xi$  depends on the application concerned, and in our spatial processing oriented application, we have found out empirically that the appropriate values for  $\xi$  is in the range of 0.005 to 0.01. The least square solution for the corresponding sparse model weight vector  $\beta_{N_{\text{spa}}}$  is readily available given the least square solution of  $\mathbf{g}_{N_{\text{spa}}}$ .

The detailed construction algorithm is summarized in Appendix B. This algorithm involves only linear optimization and is computationally significantly more attractive compared with the RVM method. In the selection procedure, if  $\mathbf{u}_i^T \mathbf{u}_i$  is too small, this term will not be selected. Thus, any ill-conditioning problem or singular situations are automatically avoided. The construction process is guaranteed to converge and, to arrive at the sparsest possible kernel detector that is also capable of closely approximating the optimum Bayesian performance, the only algorithmic parameter that requires tuning is the threshold  $\xi$ .

### 5.3. Simulation study

The example given in Section 3 was used for testing the two block-data kernel-based construction algorithms. Two conditions of  $\theta = 30^\circ$  and  $\theta = 10^\circ$  were simulated, representing the linearly separable and inseparable cases, respectively. In each case, the OFS algorithm employing the Fisher ratio and the RVM algorithm were used for constructing a RBF detector.

The number of training data used for each SNR value was  $N = 160$ . The Gaussian kernel variance  $\rho^2$  was determined empirically and the appropriate values of  $\rho^2$  were found to be in the range spanning from  $2\sigma_n^2$  to  $10\sigma_n^2$ , depending on the SNR. The number of RBF centers or kernel terms identified by the two algorithms for the given SNR values was similar, ranging from  $N_{\text{spa}} = 14$  to 20, having typical values of  $N_{\text{spa}} = 18$ . The BERs of the RVM and OFS detectors are compared in Figure 6. It can be seen that both kernel-based detectors had a similar performance at a similar model sparsity, and the two RBF detectors constructed from noisy training data closely approximated the optimal Bayesian performance. However, the OFS algorithm based on the Fisher ratio is known to have considerable computational and numerical advantages over the RVM algorithm.

## 6. RECURSIVE ADAPTIVE RBF DETECTOR USING THE COMBINED CLUSTERING AND RLS ALGORITHM

In practice, it is often desirable to update a detector on a recursive sample-by-sample basis. Consider again the RBF detector of the form

$$y(\mathbf{x}(k)) = \sum_{i=1}^{N_c} \beta_i \phi(\mathbf{x}(k), \mathbf{c}_i), \quad (39)$$

where  $\mathbf{c}_i$  are the complex-valued kernel centers and the number of kernel centers  $N_c$  is assumed to be given. We propose

to apply a combined enhanced  $\kappa$ -means clustering and RLS algorithm [19, 25] for a recursive sample-by-sample based adaptation of this RBF detector.

The enhanced  $\kappa$ -means clustering algorithm [29], which recursively updates the RBF centers, is described by

$$\mathbf{c}_i(k) = \mathbf{c}_i(k-1) + \mathcal{M}_i(\mathbf{x}(k))(\bar{g}_c(\mathbf{x}(k) - \mathbf{c}_i(k-1))) \quad (40)$$

for  $1 \leq i \leq N_c$ , where  $0 < \bar{g}_c < 1.0$  defines the learning rate, the membership function  $\mathcal{M}_i(\mathbf{x}(k))$  is defined as follows:

$$\mathcal{M}_i(\mathbf{x}) = \begin{cases} 1, & \text{if } \bar{v}_i \|\mathbf{x} - \mathbf{c}_i\|^2 \leq \bar{v}_l \|\mathbf{x} - \mathbf{c}_l\|^2 \quad \forall l \neq i, \\ 0, & \text{otherwise,} \end{cases} \quad (41)$$

and  $\bar{v}_i$  is the variation of the  $i$ th cluster. In order to estimate the associated variation  $\bar{v}_i$ , the following updating rule is used:

$$\begin{aligned} \bar{v}_i(k) &= \bar{g}_v \bar{v}_i(k-1) \\ &+ (1 - \bar{g}_v) \left( \mathcal{M}_i(\mathbf{x}(k)) \|\mathbf{x}(k) - \mathbf{c}_i(k-1)\|^2 \right), \end{aligned} \quad (42)$$

where  $\bar{g}_v$  is a constant slightly less than 1.0. The initial variations  $\bar{v}_i(0)$ ,  $1 \leq i \leq N_c$ , are set to the same small number. The learning rate  $\bar{g}_c$  can either be set to a fixed small positive number or be self-adjusting, based on an entropy formula [29].

The traditional  $\kappa$ -means clustering algorithm [28] can only achieve a local optimal solution in partitioning the input data set into  $N_c$  clusters, and the solution obtained depends on the initial locations of cluster centers. A consequence of this local optimality is that some initial centers may become trapped in regions of the input domain, which have only a few or no input patterns, and never move to regions where they are needed. This wastes resources and results in an unnecessarily large network. The enhanced  $\kappa$ -means clustering algorithm [29] overcomes the above-mentioned drawback. When using a cluster variation-weighted measure, we always achieve an optimal center configuration in the sense that after convergence, all clusters have an equal cluster variance. The above-mentioned enhanced  $\kappa$ -means clustering algorithm is an unsupervised one. In order to take full advantage of training, the algorithm can be modified in order to create a semisupervised one. Let the RBF center set be divided into the two subsets

$$\begin{aligned} \mathcal{C}^{(+)} &= \left\{ \mathbf{c}_i, 1 \leq i \leq \frac{N_c}{2} \right\}, \\ \mathcal{C}^{(-)} &= \left\{ \mathbf{c}_i, 1 + \frac{N_c}{2} \leq i \leq N_c \right\}, \end{aligned} \quad (43)$$

corresponding to the two classes  $b_1(k) = \pm 1$ . During the training instance  $k$ , the enhanced  $\kappa$ -means clustering algorithm is applied only to the center subset  $\mathcal{C}^{(+)}$  if we have  $b_1(k) = +1$ . Otherwise, it is applied to  $\mathcal{C}^{(-)}$  provided that we have  $b_1(k) = -1$ . This ‘‘semisupervised’’ clustering technique was found to be more effective in dealing with linearly inseparable cases.

The RBF weights  $\beta_i$  are updated using the classic RLS algorithm. Thus the combined CRLS algorithm used for training the RBF detector (39) can readily be summa-

rized as follows. At the instance  $k$ , given the center set  $\{\mathbf{c}_i(k-1), 1 \leq i \leq N_c\}$  and weight vector  $\boldsymbol{\beta}(k-1) = [\beta_1(k-1) \ \beta_2(k-1) \ \cdots \ \beta_{N_c}(k-1)]^T$ , we invoke the following procedure:

*RBF center updating:* use the enhanced  $\kappa$ -means clustering algorithm for obtaining an updated RBF center set  $\{\mathbf{c}_i(k), 1 \leq i \leq N_c\}$ ;

*RBF weight updating:* employ the RLS algorithm for obtaining an updated RBF weight vector  $\boldsymbol{\beta}(k)$ .

The enhanced  $\kappa$ -means clustering process is guaranteed to converge to the optimal center configuration if either the learning rate  $\bar{g}_c$  is self-adjusting based on an entropy formula or it is fixed to a positive constant that is not too large [29]. The convergence properties of the standard RLS algorithm are well known. It is therefore reasonable to believe that the above-mentioned combined  $\kappa$ -means clustering and RLS algorithm is capable of guaranteeing convergence, provided that the algorithmic parameters are set appropriately.

The example given in Section 3 was employed again for investigating the CRLS algorithm used for training the RBF detector of (39). Two conditions associated with  $\theta = 30^\circ$  and  $\theta = 10^\circ$  were simulated. For this example, the number of states that defined the Bayesian detector was  $N_b = 16$ , and  $N_c = 16$  was assumed for the RBF detector. The training data length was  $N = 1000$ . The first  $N_c$  number of samples  $\mathbf{x}(k)$  were used as the initial RBF centers and the two adaptive parameters of the clustering algorithm were set to  $\bar{g}_c = 0.2$  and  $\bar{g}_v = 0.995$ . Half of the RBF weights were set initially to  $+0.001$  and the other half to  $-0.001$ . The initial condition of the RLS algorithm was chosen as  $\boldsymbol{\Psi}(0) = \text{diag}\{1000.0, 1000.0, \dots, 1000.0\}$  with the forgetting factor given by  $\mu = 0.995$ . Figure 7 depicts the achievable BER of the CRLS RBF detector in comparison to the optimal Bayesian performance. For the CRLS RBF detector, the results obtained using the unsupervised and semisupervised clustering algorithms were similar in the linearly separable case ( $\theta = 30^\circ$ ). By contrast, for the linearly inseparable scenario of  $\theta = 10^\circ$ , it was observed that the semisupervised clustering performed better than the unsupervised one. The results given in Figure 7 are those obtained with the aid of semi-supervised clustering. From Figure 7, it can be seen that the performance of the CRLS RBF detector closely matched the optimal Bayesian performance.

## 7. CONCLUSIONS AND DISCUSSIONS

A nonlinear detection technique has been investigated in the context of a multiantenna assisted receiver. The optimal solution of the nonlinear spatial processing aided receiver has been derived for binary phase shift keying signalling. It has been shown that this optimal Bayesian detector significantly outperforms the linear beamformer in terms of a reduced bit error rate, at the expense of an increased complexity. The results presented in this paper have demonstrated the potential system capacity enhancements that may be achieved by employing nonlinear spatial processing. Both block-data-based and recursive sample-by-sample adaptive implementations

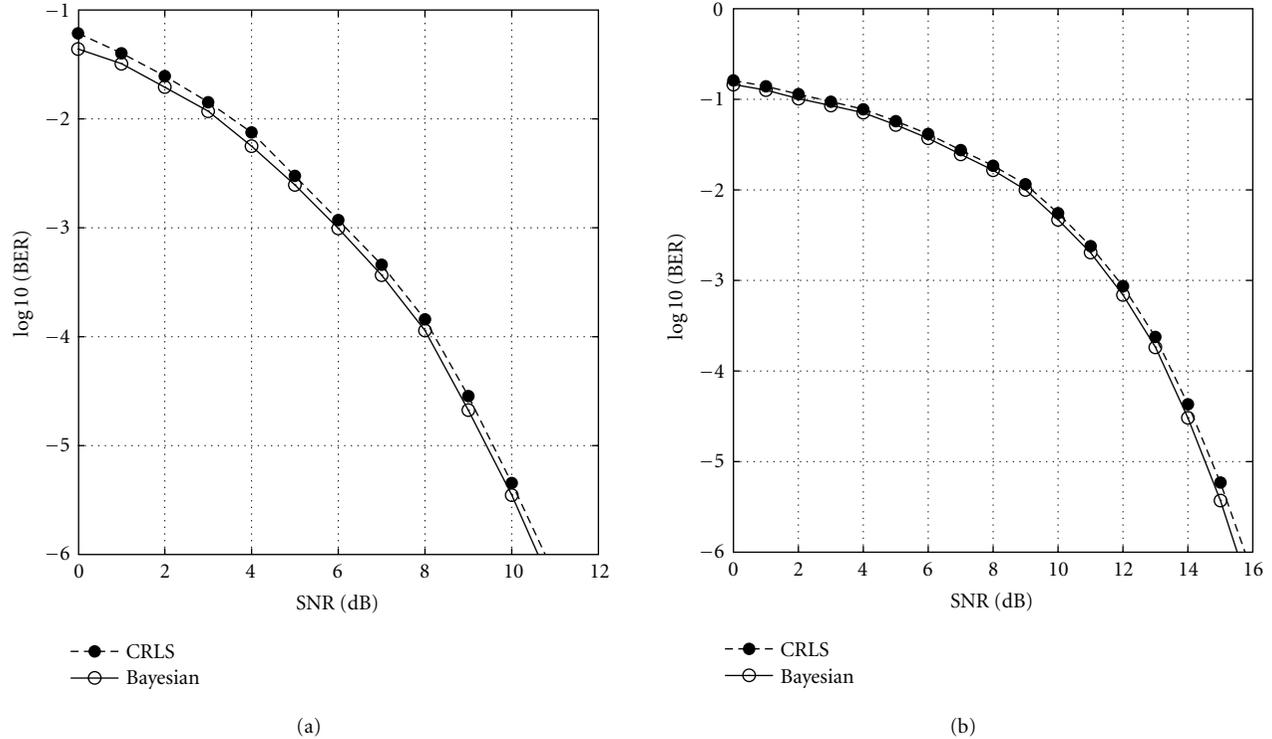


FIGURE 7: Performance comparison of the Bayesian detector with the RBF detector trained by the CRLS algorithm. (a)  $\theta = 30^\circ$ . (b)  $\theta = 10^\circ$ .

of the optimal Bayesian detector have been considered using a radial basis function network. For block-data-based adaptation, both the RVM algorithm and the orthogonal forward selection procedure employing the Fisher-ratio-based class separability measure have been considered. Both algorithms have been shown to produce similarly good performance, but the latter is known to have considerable computational advantages. For recursive sample-by-sample based adaptation, the combination of the enhanced  $\kappa$ -means clustering and the recursive least squares algorithm has been invoked.

The nonlinear detection scheme proposed in this paper is based on what we refer to as a “direct” approach, namely, on estimating the RBF centers directly from received training data contaminated by the channel. Alternatively, an “indirect” approach can be adopted, where the system matrix  $\mathbf{P}$  defined in (4) is first identified and then used for constructing the nonlinear detector. This indirect approach has the advantage of requiring a significantly shorter training time, since estimating the channel matrix needs a shorter training sequence than estimating the noiseless channel states that define RBF centers. This indirect approach is not applicable in the SDMA assisted multiuser downlink, since the receiver in this case only has access to the one desired user’s training sequence. However, this indirect scheme becomes attractive in the uplink, as the receiver has to detect all the users’ data and has access to the training sequences of all the users. Moreover, numerous complexity-reduction schemes can be adopted for the RBF detector [21]. Indeed, it was demonstrated in [21] that the complexity of the RBF detector may be rendered comparable to that of classic linear detectors. For example,

decision feedback can be employed not only to improve the performance significantly but also to reduce the complexity dramatically of the RBF detector, similar to the case of single-user channel equalization [18, 30]. This nonlinear detection scheme designed for the SDMA assisted multiuser uplink is currently under investigation.

## APPENDICES

### A. RELEVANCE VECTOR MACHINE METHOD

The posterior probability of the kernel detector weight vector  $\boldsymbol{\beta}$  is defined by

$$p(\boldsymbol{\beta}|\mathbf{t}, \boldsymbol{\alpha}) = \frac{p(\mathbf{t}|\boldsymbol{\beta}, \boldsymbol{\alpha})p(\boldsymbol{\beta}|\boldsymbol{\alpha})}{p(\mathbf{t}|\boldsymbol{\alpha})}, \quad (\text{A.1})$$

where  $p(\boldsymbol{\beta}|\boldsymbol{\alpha})$  is the prior with  $\boldsymbol{\alpha} = [\alpha_1 \ \alpha_2 \ \dots \ \alpha_N]^T$  denoting the vector of hyperparameters,  $p(\mathbf{t}|\boldsymbol{\beta}, \boldsymbol{\alpha})$  is the likelihood, and  $p(\mathbf{t}|\boldsymbol{\alpha})$  the evidence. Following the Bayesian classification framework [22, 23], the likelihood is expressed as

$$p(\mathbf{t}|\boldsymbol{\beta}, \boldsymbol{\alpha}) = \prod_{l=1}^N (f(y(\mathbf{x}(l))))^{(t(l)+1)/2} (1 - f(y(\mathbf{x}(l))))^{(1-t(l))/2}, \quad (\text{A.2})$$

where

$$f(y) = \frac{1}{1 + \exp(-y)} \quad (\text{A.3})$$

is the logistic sigmoid function. The Gaussian prior is chosen:

$$p(\boldsymbol{\beta}|\boldsymbol{\alpha}) = \prod_{l=1}^N \frac{\sqrt{\alpha_l}}{\sqrt{2\pi}} \exp\left(-\frac{\alpha_l \beta_l^2}{2}\right). \quad (\text{A.4})$$

As the marginal likelihood  $p(\mathbf{t}|\boldsymbol{\alpha})$  cannot be obtained analytically by integrating out the weights from (A.1), an iterative procedure is necessitated [22].

With a fixed given  $\boldsymbol{\alpha}$ , the maximum a posteriori probability (MAP) solution  $\hat{\boldsymbol{\beta}}$  can be obtained by maximizing  $\log(p(\boldsymbol{\beta}|\mathbf{t}, \boldsymbol{\alpha}))$  or, equivalently, by minimizing the following cost function:

$$J(\boldsymbol{\beta}|\mathbf{t}, \boldsymbol{\alpha}) = \sum_{l=1}^N \left( \frac{\alpha_l \beta_l^2}{2} - \frac{t(l)+1}{2} \log(f(y(\mathbf{x}(l)))) - \frac{1-t(l)}{2} \log(1-f(y(\mathbf{x}(l)))) \right). \quad (\text{A.5})$$

The gradient of  $J$  with respect to  $\boldsymbol{\beta}$  is

$$\nabla J = \mathbf{A}\boldsymbol{\beta} + \boldsymbol{\Phi}^T \left( \mathbf{f} - \frac{1}{2}(\mathbf{t} + \mathbf{1}_N) \right), \quad (\text{A.6})$$

where

$$\begin{aligned} \mathbf{A} &= \text{diag}\{\alpha_1, \alpha_2, \dots, \alpha_N\}, \\ \mathbf{f} &= [f(y(\mathbf{x}(1))) \quad f(y(\mathbf{x}(2))) \quad \dots \quad f(y(\mathbf{x}(N)))]^T, \\ \mathbf{1}_N &= [1 \quad 1 \quad \dots \quad 1]^T, \end{aligned} \quad (\text{A.7})$$

and  $\boldsymbol{\Phi}$  is the regression matrix defined in (31). The Hessian of  $J$  is

$$\mathbf{H} = \nabla^2 J = \boldsymbol{\Phi}^T \mathbf{B} \boldsymbol{\Phi} + \mathbf{A}, \quad (\text{A.8})$$

where

$$\mathbf{B} = \text{diag}\{f(y(\mathbf{x}(1)))(1-f(y(\mathbf{x}(1)))) , \dots, f(y(\mathbf{x}(N)))(1-f(y(\mathbf{x}(N))))\}. \quad (\text{A.9})$$

The hyperparameters  $\boldsymbol{\alpha}$  are updated using

$$\alpha_i^{\text{new}} = \frac{1 - \alpha_i^{\text{old}} \gamma_{i,i}}{\hat{\beta}_i^2} \quad (\text{A.10})$$

with  $\gamma_{i,i}$  being the diagonal elements of  $\boldsymbol{\Gamma}$  which is defined by

$$\boldsymbol{\Gamma} = (\mathbf{H}|\hat{\boldsymbol{\beta}})^{-1}. \quad (\text{A.11})$$

The following simple iterative procedure can be adopted to construct a sparse RVM detector.

#### Initialization

The  $N \times N_{\text{spa}}$  kernel matrix  $\boldsymbol{\Phi}$  is initialized with  $N_{\text{spa}} = N$ , that is, every training data point is considered as a candidate kernel. Each weight  $\beta_i$  is initially associated with the same value of the hyperparameter  $\alpha_i$ .

*Step 1.* Given current value  $\boldsymbol{\alpha}$ , find  $\hat{\boldsymbol{\beta}}$  by minimizing the cost function (A.5). A simplified conjugate gradient algorithm [26, 27] is used in our application.

*Step 2.* The hyperparameters are updated using (A.10). If a  $\alpha_i > Lg$ , where  $Lg$  is a preset large positive value,  $N_{\text{spa}} := N_{\text{spa}} - 1$ , the corresponding column in  $\boldsymbol{\Phi}$  is removed, and thus the corresponding weight  $\beta_i$  and model term  $\phi_i(\cdot)$  are pruned out the model.

#### Test

If the hyperparameters  $\boldsymbol{\alpha}$  remain sufficiently unchanged in two successive iterations (no removal of hyperparameters) or a preset maximum iteration number is reached, stop; otherwise, go to Step 1.

## B. ORTHOGONAL FORWARD SELECTION ALGORITHM

The modified Gram-Schmidt orthogonalization procedure [31] calculates the  $\mathbf{D}$  matrix row by row and orthogonalizes  $\boldsymbol{\Phi}$  as follows: at the  $l$ th stage, make the columns  $\boldsymbol{\phi}_i$ ,  $l+1 \leq i \leq N$ , orthogonal to the  $l$ th column and repeat the operation for  $1 \leq l \leq N-1$ . Specifically, denoting  $\boldsymbol{\phi}_i^{(0)} = \boldsymbol{\phi}_i$ ,  $1 \leq i \leq N$ , then

$$\begin{aligned} \mathbf{u}_l &= \boldsymbol{\phi}_l^{(l-1)}, \\ d_{l,i} &= \frac{\mathbf{u}_l^T \boldsymbol{\phi}_i^{(l-1)}}{(\mathbf{u}_l^T \mathbf{u}_l)}, \quad l+1 \leq i \leq N, \\ \boldsymbol{\phi}_i^{(l)} &= \boldsymbol{\phi}_i^{(l-1)} - d_{l,i} \mathbf{u}_l, \quad l+1 \leq i \leq N, \end{aligned} \quad (\text{B.1})$$

where  $l = 1, 2, \dots, N-1$ .

The last stage of the procedure is simply  $\mathbf{u}_N = \boldsymbol{\phi}_N^{(N-1)}$ . The elements of  $\mathbf{g}$  are computed by transforming  $\mathbf{t}^{(0)} = \mathbf{t}$  in a similar way:

$$\begin{aligned} g_l &= \frac{\mathbf{u}_l^T \mathbf{t}^{(l-1)}}{(\mathbf{u}_l^T \mathbf{u}_l)}, \\ \mathbf{t}^{(l)} &= \mathbf{t}^{(l-1)} - g_l \mathbf{u}_l, \end{aligned} \quad (\text{B.2})$$

where  $1 \leq l \leq N$ .

This orthogonalization scheme can be used to derive a simple and efficient algorithm for selecting subset models in a forward-regression manner [31]. First define

$$\boldsymbol{\Phi}^{(l-1)} = [\mathbf{u}_1 \cdots \mathbf{u}_{l-1} \quad \boldsymbol{\phi}_l^{(l-1)} \cdots \boldsymbol{\phi}_N^{(l-1)}]. \quad (\text{B.3})$$

If some of the columns  $\boldsymbol{\phi}_1^{(l-1)}, \dots, \boldsymbol{\phi}_N^{(l-1)}$  in  $\boldsymbol{\Phi}^{(l-1)}$  have been interchanged, this will still be referred to as  $\boldsymbol{\Phi}^{(l-1)}$  for notational convenience. With the notation  $\boldsymbol{\phi}_q^{(l-1)} = [\phi_{1,q}^{(l-1)} \quad \phi_{2,q}^{(l-1)} \quad \dots \quad \phi_{N,q}^{(l-1)}]^T$ , the  $l$ th stage of the selection procedure is given as follows.

Step 1. For  $l \leq q \leq N$ , compute

$$\begin{aligned}
 m_{+,l}^{(q)} &= \frac{1}{N_+} \sum_{i=1}^N \delta(t(i) - 1) \phi_{i,q}^{(l-1)}, \\
 (\sigma_{+,l}^{(q)})^2 &= \frac{1}{N_+} \sum_{i=1}^N \delta(t(i) - 1) (\phi_{i,q}^{(l-1)} - m_{+,l}^{(q)})^2, \\
 m_{-,l}^{(q)} &= \frac{1}{N_-} \sum_{i=1}^N \delta(t(i) + 1) \phi_{i,q}^{(l-1)}, \\
 (\sigma_{-,l}^{(q)})^2 &= \frac{1}{N_-} \sum_{i=1}^N \delta(t(i) + 1) (\phi_{i,q}^{(l-1)} - m_{-,l}^{(q)})^2, \\
 F_l^{(q)} &= \frac{(m_{+,l}^{(q)} - m_{-,l}^{(q)})^2}{(\sigma_{+,l}^{(q)})^2 + (\sigma_{-,l}^{(q)})^2}.
 \end{aligned} \tag{B.4}$$

Step 2. Find

$$F_l = F_l^{(q_l)} = \max \{F_l^{(q)}, l \leq q \leq N\}. \tag{B.5}$$

Then the  $q_l$ th column of  $\Phi^{(l-1)}$  is interchanged with the  $l$ th column of  $\Phi^{(l-1)}$ , and the  $q_l$ th column of  $\mathbf{D}$  is interchanged with the  $l$ th column of  $\mathbf{D}$  up to the  $(l-1)$ th row. This effectively selects the  $q_l$ th candidate as the  $l$ th kernel term in the subset model.

Step 3. Perform the orthogonalization as indicated in (B.1) to derive the  $l$ th row of  $\mathbf{D}$  and to transform  $\Phi^{(l-1)}$  into  $\Phi^{(l)}$ . Calculate  $g_l$  and update  $\mathbf{t}^{(l-1)}$  into  $\mathbf{t}^{(l)}$  in the way shown in (B.2).

The selection is terminated at the  $N_{\text{spa}}$  stage when the criterion (38) is satisfied and this produces a sparse subset model containing  $N_{\text{spa}}$  significant kernel terms.

## REFERENCES

- [1] J. H. Winters, J. Salz, and R. D. Gitlin, "The impact of antenna diversity on the capacity of wireless communication systems," *IEEE Trans. Communications*, vol. 42, no. 2/3/4, pp. 1740–1751, 1994.
- [2] J. Litva and T. K. Y. Lo, *Digital Beamforming in Wireless Communications*, Artech House, London, UK, 1996.
- [3] L. C. Godara, "Applications of antenna arrays to mobile communications. I. Performance improvement, feasibility, and system considerations," *Proceedings of the IEEE*, vol. 85, no. 7, pp. 1031–1060, 1997.
- [4] A. J. Paulraj and C. B. Papadias, "Space-time processing for wireless communications," *IEEE Signal Processing Magazine*, vol. 14, no. 6, pp. 49–83, 1997.
- [5] J. H. Winters, "Smart antennas for wireless systems," *IEEE Personal Communications*, vol. 5, no. 1, pp. 23–27, 1998.
- [6] R. Kohno, "Spatial and temporal communication theory using adaptive antenna array," *IEEE Personal Communications*, vol. 5, no. 1, pp. 28–35, 1998.
- [7] J. S. Blough and L. Hanzo, *Third Generation Systems and Intelligent Wireless Networking: Smart Antenna and Adaptive Modulation*, John Wiley & Sons, Chichester, UK, 2002.
- [8] R. A. Soni, R. M. Buehrer, and R. D. Benning, "Intelligent antenna system for cdma2000," *IEEE Signal Processing Magazine*, vol. 19, no. 4, pp. 54–67, 2002.
- [9] I. S. Reed, J. D. Mallett, and L. E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Trans. on Aerospace and Electronics Systems*, vol. 10, no. 6, pp. 853–863, 1974.
- [10] M. W. Ganz, R. L. Moses, and S. L. Wilson, "Convergence of the SMI and the diagonally loaded SMI algorithms with weak interference [adaptive array]," *IEEE Trans. Antennas and Propagation*, vol. 38, no. 3, pp. 394–399, 1990.
- [11] B. Widrow, P. E. Mantey, L. J. Griffiths, and B. B. Goode, "Adaptive antenna systems," *Proceedings of the IEEE*, vol. 55, no. 12, pp. 2143–2159, 1967.
- [12] L. J. Griffiths, "A simple adaptive algorithm for real-time processing in antenna arrays," *Proceedings of the IEEE*, vol. 57, no. 10, pp. 1696–1704, 1969.
- [13] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, Upper Saddle River, NJ, USA, 3rd edition, 1996.
- [14] S. Chen, L. Hanzo, and N. N. Ahmad, "Adaptive minimum bit error rate beamforming assisted receiver for wireless communications," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, vol. 4, pp. 640–643, Hong Kong, April 2003.
- [15] S. Chen, N. N. Ahmad, and L. Hanzo, "Adaptive minimum bit error rate beamforming," to appear in *IEEE Trans. Wireless Communications*.
- [16] S. Chen, G. J. Gibson, C. F. N. Cowan, and P. M. Grant, "Reconstruction of binary signals using an adaptive radial-basis-function equalizer," *Signal Processing*, vol. 22, no. 1, pp. 77–93, 1991.
- [17] S. Chen, B. Mulgrew, and P. M. Grant, "A clustering technique for digital communications channel equalization using radial basis function networks," *IEEE Transactions on Neural Networks*, vol. 4, no. 4, pp. 570–590, 1993.
- [18] S. Chen, B. Mulgrew, and S. McLaughlin, "Adaptive Bayesian equalizer with decision feedback," *IEEE Trans. Signal Processing*, vol. 41, no. 9, pp. 2918–2927, 1993.
- [19] S. Chen, A. K. Samingan, and L. Hanzo, "Support vector machine multiuser receiver for DS-CDMA signals in multipath channels," *IEEE Transactions on Neural Networks*, vol. 12, no. 3, pp. 604–611, 2001.
- [20] S. Chen and L. Hanzo, "Block-adaptive kernel-based CDMA multiuser detection," in *Proc. IEEE International Conference on Communications*, vol. 2, pp. 682–686, New York, NY, USA, April–May 2002.
- [21] L. Hanzo, C. H. Wong, and M. S. Yee, *Adaptive Wireless Transceivers: Turbo-Coded, Turbo-Equalized and Space-Time Coded TDMA, CDMA, and OFDM Systems*, John Wiley & Sons, New York, NY, USA, 2002.
- [22] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *Journal of Machine Learning Research*, vol. 1, pp. 211–244, 2001.
- [23] S. Chen, S. R. Gunn, and C. J. Harris, "The relevance vector machine technique for channel equalization application," *IEEE Transactions on Neural Networks*, vol. 12, no. 6, pp. 1529–1532, 2001.
- [24] K. Z. Mao, "RBF neural network center selection based on Fisher ratio class separability measure," *IEEE Transactions on Neural Networks*, vol. 13, no. 5, pp. 1211–1217, 2002.
- [25] S. Chen, "Nonlinear time series modelling and prediction using Gaussian RBF networks with enhanced clustering and RLS learning," *Electronics Letters*, vol. 31, no. 2, pp. 117–118, 1995.
- [26] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*, John Wiley & Sons, New York, NY, USA, 2nd edition, 1993.
- [27] S. Chen, A. K. Samingan, B. Mulgrew, and L. Hanzo, "Adaptive minimum-BER linear multiuser detection for DS-CDMA

signals in multipath channels," *IEEE Trans. Signal Processing*, vol. 49, no. 6, pp. 1240–1247, 2001.

- [28] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*, John Wiley & Sons, New York, NY, USA, 1973.
- [29] C. Chinrungrueng and C. H. Sequin, "Optimal adaptive  $\kappa$ -means algorithm with dynamic adjustment of learning rate," *IEEE Transactions on Neural Networks*, vol. 6, no. 1, pp. 157–169, 1995.
- [30] S. Chen, S. McLaughlin, B. Mulgrew, and P. M. Grant, "Bayesian decision feedback equaliser for overcoming co-channel interference," *IEE Proceedings-Communication*, vol. 143, no. 4, pp. 219–225, 1996.
- [31] S. Chen, S. A. Billings, and W. Luo, "Orthogonal least squares methods and their application to non-linear system identification," *International Journal of Control*, vol. 50, no. 5, pp. 1873–1896, 1989.

**Sheng Chen** obtained his B.Eng. degree in control engineering from the East China Petroleum Institute, China, in 1982, and his Ph.D. degree in control engineering from the City University at London in 1986. He joined the Department of Electronics and Computer Science at the University of Southampton, UK, in September 1999. He has previously held research and academic appointments at the Universities of Sheffield, Edinburgh, and Portsmouth, UK. Dr. Chen is a Senior Member of IEEE. His recent research works include adaptive nonlinear signal processing, modeling and identification of nonlinear systems, neural network research, finite-precision digital controller design, and evolutionary computation methods and optimization. He has published over 200 research papers.



**Lajos Hanzo** received his degree in electronics in 1976 and his doctorate in 1983. During his career in telecommunications, he has held various research and academic posts in Hungary, Germany, and the UK. Since 1986, he has been with the Department of Electronics and Computer Science, University of Southampton, UK, where he holds the Chair in telecommunications. He coauthored 10 books totalling 8000 pages on mobile radio communications, published about 450 research papers, organized and chaired conference sessions, presented overview lectures, and has been awarded a number of distinctions. Currently he heads an academic research team, working on a range of research projects in the field of wireless multimedia communications sponsored by industry, the Engineering and Physical Sciences Research Council (EPSRC), UK, the European IST Programme, and the Mobile Virtual Centre of Excellence (VCE), UK. He is an enthusiastic supporter of industrial and academic liaison and he offers a range of industrial courses. L. Hanzo is also an IEEE Distinguished Lecturer at both the Communications as well as the Vehicular Technology Society, a Fellow of the IEE, and a Fellow of the IEEE.



**Andreas Wolfgang** received his Dipl.-Ing. degree in electrical engineering from Karlsruhe University of Technology, Germany, in 2003. Currently he is with the Communications Research Group at the University of Southampton, UK, where he is pursuing the Ph.D. degree. He was a member of the Antenna Group at Chalmers University, Gothenburg, Sweden, where he worked for the development of measurement methods for terminal antennas. His current research interests are in wireless communications with emphasis on nonlinear filter structures designed for multiple antenna systems.



# Upper Bounds on the BER Performance of MTCM-STBC Schemes over Shadowed Rician Fading Channels

**M. Uysal**

*Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada N2L 3G1  
Email: muysal@ece.uwaterloo.ca*

**C. N. Georgiades**

*Department of Electrical Engineering, Texas A&M University, College Station, TX 77843-3128, USA  
Email: georgiades@ee.tamu.edu*

*Received 17 May 2003; Revised 21 October 2003*

Space-time block coding (STBC) provides substantial diversity advantages with a low decoding complexity. However, these codes are not designed to achieve coding gains. Outer codes should be concatenated with STBC to provide additional coding gain. In this paper, we analyze the performance of concatenated trellis-coded STBC schemes over shadowed Rician frequency-flat fading channels. We derive an exact pairwise error probability (PEP) expression that reveals the dominant factors affecting performance. Based on the derived PEP, in conjunction with the transfer function technique, we also present upper bounds on the bit error rate (BER), which are further shown to be tight through a Monte-Carlo simulation study.

**Keywords and phrases:** space-time block coding, trellis-coded modulation, Rician fading channels, shadowing, pairwise error probability.

## 1. INTRODUCTION

Space-time trellis coding was introduced in [1] as an effective transmit diversity technique to combat fading. These codes were designed to achieve maximum diversity gain. However, for a fixed number of transmit antennas, their decoding complexity increases exponentially with the transmission rate. Space-time block coding (STBC) [2] was proposed as an attractive alternative to its trellis counterpart with a much lower decoding complexity. The work in [2] was inspired by Alamouti's early work [3], where a simple two-branch transmit diversity scheme was presented and shown to provide the same diversity order as maximal-ratio receiver combining with two receive antennas. Alamouti's scheme is appealing in terms of its performance and simplicity. Assuming the channel is known at the receiver, it requires a simple maximum-likelihood decoding algorithm based only on linear processing at the receiver. STBC generalizes Alamouti's scheme to an arbitrary number of transmit antennas and is able to provide the full diversity promised by the transmit and receive antennas. However, these codes are not designed to achieve a coding gain. Therefore, outer codes should be concatenated with STBC to achieve additional coding gains. A pioneering work towards this end is presented in [4] where concatenation of trellis-coded modulation (TCM) with STBC is considered. In [4], it is shown that the *free distance* of the trellis code

dominates performance; therefore, the optimal trellis codes designed for additive white Gaussian noise (AWGN) are also optimum for concatenated TCM-STBC over quasistatic Rayleigh fading channels. We studied the same concatenated scheme combined with an interleaver in [5] over Rician fading channels. In this paper, we generalize our work to *shadowed* Rician channels. The shadowed Rician channel [6] is a generalization of the Rician model, where the line-of-sight (LOS) path is subjected to a lognormal transformation due to foliage attenuation or blockage, also referred to as *shadowing*. Specifically, we derive an exact pairwise error probability (PEP) for concatenated TCM-STBC schemes. Our exact evaluation of PEP is based on the moment-generating function technique [7, 8], which has been successfully applied to the analysis of digital communication systems over fading channels. Using the classical transfer function technique based on the exact PEP, we obtain upper bounds on bit error rate (BER) performance, which are further verified through simulation. Our analysis also reveals the selection criteria for trellis codes which should be used in conjunction with STBC.

The organization of the paper is as follows. In Section 2 we explain our system model, where the concatenated TCM-STBC is described and the channel model under consideration is introduced. In Section 3 an exact expression for PEP is derived for the TCM-STBC scheme using the MGF approach. Based on the derived PEP, we discuss the selection criteria

for trellis codes which should be used with space-time codes for optimal performance and compare them with the classical selection criteria for trellis codes over fading channels without transmitter diversity. In Section 4, using the transfer function technique in conjunction with the derived PEP expressions, we obtain upper bounds on the BER performance. Analytical performance results are presented for two example trellis codes, which are further confirmed through Monte-Carlo simulation.

## 2. SYSTEM MODEL

We consider a wireless communication scenario where the transmitter is equipped with  $M$  antennas and the receiver is equipped with  $N$  antennas. The binary data is first encoded by a trellis encoder. After trellis coded symbols are interleaved and mapped to constellation symbols, they are fed to the STBC encoder. An STBC is defined [2] by an  $L \times M$  code matrix, where  $L$  represents the number of time intervals for transmitting  $P$  symbols, resulting in a code rate of  $P/L$ . For Tarokh et al.'s orthogonal space-time block codes [2], the entries of the code matrix are chosen as linear combinations of the transmission symbols and their conjugates. For example, the code matrix for the well-known Alamouti's scheme (i.e., STBC for 2 transmit antennas) is given by

$$\begin{bmatrix} x_1 & x_2 \\ -x_2^* & x_1^* \end{bmatrix} \quad (1)$$

with  $M = P = L = 2$ .

We assume that the transmission frame from each antenna consists of a total of  $FL$  symbols (i.e., consecutive  $F$  smaller inner-frames, each of them having duration  $L$  symbols corresponding to the STBC length). The received signal at receive antenna  $n$  ( $n = 1, 2, \dots, N$ ) at time interval  $l$  of the  $f$ th ( $f = 1, 2, \dots, F$ ) inner-frame is a superposition of  $M$  transmitted signals:

$$r_n^f(l) = \sum_{m=1}^M \alpha_{m,n}^f x_m^f(l) + \eta_n^f(l), \quad (2)$$

where  $x_m^f(l)$  is the modulation symbol transmitted from the  $m$ th transmit antenna at time interval  $l$  of the  $f$ th frame and  $\eta_n^f(l)$  is additive noise, modeled as a complex Gaussian random variable with zero mean and variance  $N_0/2$  per dimension.  $\alpha_{m,n}^f$  represents the fading coefficient modeling the channel from the  $m$ th transmit to the  $n$ th receive antenna during the  $f$ th inner frame and are assumed to be independent and identically distributed (i.i.d.). The fading coefficient is assumed to remain constant over an inner-frame period (i.e.,  $L$  symbol intervals). This assumption is necessary to make use of the orthogonal structure of STBC to guarantee full spatial diversity. The assumption of quasistatic behavior of the channel over an inner-frame period can be justified using an  $L$ -symbol interleaver over a moderately slow varying channel. In our case, the fading amplitude is described by the shadowed Rician fading model. In this model, the

LOS component is not constant but rather a lognormally distributed random variable. The fading coefficient can be expressed (dropping the subscripts and superscripts for notational convenience) as  $\alpha = \mu + \xi_0 + j\xi_1$ , where  $\xi_0$  and  $\xi_1$  are independent Gaussian random variables with zero mean and variance  $\sigma^2$ . Here, the LOS component is given as  $\mu = \exp(\xi_2)$  where  $\xi_2$  is a Gaussian random variable with mean  $m_\mu$  and variance  $\sigma_\mu^2$ , and independent of  $\xi_0$  and  $\xi_1$ . The conditional probability density function of the fading amplitude  $|\alpha|$  is

$$p_{|\alpha||\mu}(|\alpha| | \mu) = \frac{|\alpha|}{\sigma^2} \exp\left(-\frac{|\alpha|^2 + \mu^2}{2\sigma^2}\right) I_0\left(\frac{|\alpha|\mu}{\sigma^2}\right), \quad |\alpha| \geq 0, \quad (3)$$

where  $I_0(\cdot)$  is the zero-order modified Bessel function of the first kind, and the probability density function of the LOS component is given by

$$p_\mu(\mu) = \frac{1}{\sqrt{2\pi}\sigma_\mu\mu} \exp\left(-\frac{(\ln \mu - m_\mu)^2}{2\sigma_\mu^2}\right). \quad (4)$$

The parameters  $\sigma$ ,  $\sigma_\mu$ , and  $m_\mu$  in (3) and (4) specify the degree of shadowing. Denoting by  $\mathbb{C}^{m \times n}$  the vector space of  $m$ -by- $n$  complex matrices, and defining<sup>1</sup>

$$\begin{aligned} \mathbf{r}_n^f &= (r_n^f(1), r_n^f(2), \dots, r_n^f(L))^T \in \mathbb{C}^{L \times 1}, \\ \boldsymbol{\alpha}_n^f &= (\alpha_{1,n}^f, \alpha_{2,n}^f, \dots, \alpha_{M,n}^f)^T \in \mathbb{C}^{M \times 1}, \\ \boldsymbol{\eta}_n^f &= (\eta_n^f(1), \eta_n^f(2), \dots, \eta_n^f(L))^T \in \mathbb{C}^{L \times 1}, \end{aligned} \quad (5)$$

the received signal can be written in matrix notation as

$$\mathbf{r}_n^f = \mathbf{X}^f \boldsymbol{\alpha}_n^f + \boldsymbol{\eta}_n^f, \quad n = 1, 2, \dots, N, \quad f = 1, 2, \dots, F, \quad (6)$$

where  $\mathbf{X}^f \in \mathbb{C}^{L \times M}$  consists of space-time encoded symbols (which have been already trellis encoded) for the  $f$ th inner frame. At the receiver, first the received signal is passed through the space-time decoder, which is essentially based on linear processing for STBC from orthogonal designs [2]. After deinterleaving, the processed sequence is fed to the trellis decoder implemented by a Viterbi algorithm. If a multiple TCM (MTCM) scheme with  $M$  symbols per branch is used (note that the number of transmit antennas is also given as  $M$ ), the decoding steps can be combined in one step with a proper modification of the metric employed in the Viterbi algorithm. In this case, the received signal is just deinterleaved and fed directly to the Viterbi decoder without any further processing.

## 3. DERIVATION OF EXACT PEP

In this section, we analyze the PEP of the concatenated scheme over shadowed Rician fading channels assuming

<sup>1</sup>Throughout this paper, we use  $(\cdot)^T$  and  $(\cdot)^H$  for the transpose and transpose conjugate operations, respectively. Upper case bold face letters represent matrices and lower case bold face letters represent vectors.

perfect channel state information is available at the receiver. Assuming equal transmitted power at all transmit antennas, the conditional PEP of transmitting code matrix  $\mathbf{X}$  (which consists of  $\mathbf{X}^f$ ,  $f = 1, 2, \dots, F$ ) and erroneously deciding in favor of another code matrix  $\hat{\mathbf{X}}$  at the decoder is given by

$$P(\mathbf{X}, \hat{\mathbf{X}} \mid \alpha_{m,n}^f, \mu_{m,n}^f, m = 1, \dots, M, n = 1, \dots, N, f = 1, \dots, F) = Q\left(\sqrt{\sum_{f=1}^F \sum_{n=1}^N (\alpha_n^f)^H \mathbf{A}^f \alpha_n^f}\right), \quad (7)$$

where  $Q(\cdot)$  is the Gaussian Q-function and  $\mathbf{A}^f$  is given by

$$\mathbf{A}^f = \frac{1}{M} \frac{E_s}{2N_0} (\mathbf{X}^f - \hat{\mathbf{X}}^f)^H (\mathbf{X}^f - \hat{\mathbf{X}}^f). \quad (8)$$

Here,  $E_s$  is the total signal power transmitted from all  $M$  transmit antennas and  $N_0/2$  is the noise variance per dimension. In order to find the unconditional PEP, we need to take expectations with respect to  $\alpha_{m,n}^f$  and  $\mu_{m,n}^f$ . The expectation with respect to fading coefficients can be obtained through use of the alternative form of the Gaussian Q-function [8] as

$$P(\mathbf{X}, \hat{\mathbf{X}} \mid \mu_{m,n}^f, m = 1, \dots, M, n = 1, \dots, N, f = 1, \dots, F) = \frac{1}{\pi} \int_0^{\pi/2} \Phi_{\Gamma}\left(-\frac{1}{2 \sin^2 \theta}\right) d\theta, \quad (9)$$

where  $\Phi_{\Gamma}(s)$  is the moment generating function (MGF) of

$$\Gamma = \sum_{f=1}^F \sum_{n=1}^N (\alpha_n^f)^H \mathbf{A}^f \alpha_n^f. \quad (10)$$

$\Gamma$  is a quadratic form of complex Gaussian random variables and its MGF is given as [9, 10]

$$\Phi_{\Gamma}(s) = \prod_{f=1}^F \prod_{n=1}^N \prod_{m=1}^M \frac{1}{1 - s\chi_m} \exp\left(\frac{s\chi_m |d_m|^2}{1 - s\chi_m}\right), \quad (11)$$

where  $\chi_m$  are the eigenvalues of  $\Sigma \mathbf{A}^f$  and  $d_m$  are the elements of  $M$ -length vector  $\mathbf{d} = \boldsymbol{\mu} \boldsymbol{\Sigma}^{-1/2}$ . Here  $\boldsymbol{\mu}$  and  $\boldsymbol{\Sigma}$  represent the mean vector and the covariance matrix of  $\alpha_n^f$ , respectively. Making use of the assumed i.i.d. properties of the fading channel, we obtain  $|d_m|^2 = \mu^2/2\sigma^2$ . Furthermore, in our case,  $\mathbf{A}^f$  is a diagonal matrix due to the orthogonality of STBC and the eigenvalues  $\chi_m$  are simply equal to the diagonal elements of  $\Sigma \mathbf{A}^f$ , that is,

$$\frac{E_s}{2N_0} 2\sigma^2 \frac{\beta}{M} \sum_{p=1}^P |x_p^f - \hat{x}_p^f|^2, \quad (12)$$

where  $\beta = 1$  for  $M = 2$  and  $\beta = 2$  for  $M > 2$  due to the special matrix structure of STBC based on orthogonal designs [2]. Inserting (11) into (9) and using the i.i.d. properties for

fading coefficients, we obtain

$$P(\mathbf{X}, \hat{\mathbf{X}} \mid \mu) = \frac{1}{\pi} \int_0^{\pi/2} \prod_{f=1}^F \left[ \frac{1}{1 + \Omega_f / \sin^2 \theta} \exp\left(-\frac{\mu^2}{2\sigma^2} \frac{\Omega_f / \sin^2 \theta}{1 + \Omega_f / \sin^2 \theta}\right) \right]^{MN} d\theta, \quad (13)$$

where

$$\Omega_f = \frac{E_s}{4N_0} 2\sigma^2 \frac{\beta}{M} \sum_{p=1}^P |x_p^f - \hat{x}_p^f|^2. \quad (14)$$

To find the unconditional PEP, we still need to take an expectation of (13) with respect to  $\mu$ , whose distribution is given by (4). This expectation yields

$$P(\mathbf{X}, \hat{\mathbf{X}}) = \frac{1}{\pi} \int_{\theta=0}^{\pi/2} \prod_{f=1}^F \left[ \frac{1}{1 + \Omega_f / \sin^2 \theta} \frac{1}{\sqrt{2\pi\sigma_\mu}} \times \int_{\mu=0}^{\infty} \frac{1}{\mu} \exp\left(-\frac{\mu^2}{2\sigma^2} \frac{\Omega_f / \sin^2 \theta}{1 + \Omega_f / \sin^2 \theta}\right) \times \exp\left(-\frac{(\ln \mu - m_\mu)^2}{2\sigma_\mu^2}\right) d\mu \right]^{MN} d\theta. \quad (15)$$

Introducing the variable change  $u = (\ln \mu - m_\mu) / \sqrt{2\sigma_\mu^2}$ , (15) can be rewritten as

$$P(\mathbf{X}, \hat{\mathbf{X}}) = \frac{1}{\pi} \int_{\theta=0}^{\pi/2} \prod_{f=1}^F \left[ \frac{1}{1 + \Omega_f / \sin^2 \theta} \frac{1}{\sqrt{\pi}} \times \int_{u=-\infty}^{\infty} \exp(-u^2) \times \exp\left(-\frac{1}{2\sigma^2} \frac{\Omega_f / \sin^2 \theta}{1 + \Omega_f / \sin^2 \theta} \times \exp(2\sqrt{2}\sigma_\mu u + 2m_\mu)\right) du \right]^{MN} d\theta. \quad (16)$$

The inner integral has the form of  $\int_{-\infty}^{\infty} \exp(-u^2) f(u) du$ , which can be expressed in terms of an infinite sum (see the appendix). This yields the final form of the exact PEP as

$$P(\mathbf{X}, \hat{\mathbf{X}}) = \frac{1}{\pi} \int_{\theta=0}^{\pi/2} \prod_{f=1}^F \left\{ \frac{1}{1 + \Omega_f / \sin^2 \theta} \exp(-\Delta_f(\theta)) \times \left[ 1 + \sum_{\substack{k=2 \\ k:\text{even}}}^{\infty} \frac{(k-1)!!}{k!} (2\sigma_\mu)^k \times \sum_{d=1}^k g_{k,d}(\Delta_f(\theta))^d \right] \right\}^{MN} d\theta, \quad (17)$$

where

$$\Delta_f(\theta) = \frac{1}{2\sigma^2} \frac{\Omega_f/\sin^2\theta}{2\sigma^2 + \Omega_f/\sin^2\theta} \exp(2m_\mu) \quad (18)$$

and  $(k-1)!! = 1.3 \cdot \dots \cdot k$  [11, page xlv]. The coefficients  $g_{k,d}$  in (17) can be computed by the recursive equation given in the appendix. It is worth noting that even considering only the first term in the infinite summation in (17) gives a very good approximation for practical values of shadowing. Setting  $k = 2$  and noting that  $g_{2,1} = -1$  and  $g_{2,2} = 1$ , we have

$$P(\mathbf{X}, \hat{\mathbf{X}}) \cong \frac{1}{\pi} \int_{\theta=0}^{\pi/2} \prod_{f=1}^F \left\{ \frac{1}{1 + \Omega_f/\sin^2\theta} \exp(-\Delta_f(\theta)) \times \left[ 1 - 2\sigma_\mu^2 \Delta_f(\theta) + 2\sigma_\mu^2 (\Delta_f(\theta))^2 \right] \right\}^{MN} \quad (19)$$

In our numerical results, taking more terms (i.e.,  $k > 2$ ) did not result in a visible change in the plots.

It is also interesting to point out how (17) relates to the unshadowed case. Assuming there is no shadowing,  $\mu$  is no longer a log-normal random variable, but just given as a constant equal to its mean  $\mu = \exp(2m_\mu)$ . Furthermore, inserting  $\sigma_\mu^2 = 0$  in (17) and using the relationships  $\sigma^2 = 0.5/(1+K)$  and  $\mu = \sqrt{K/(1+K)}$  in terms of the well-known Rician parameter  $K$ , we obtain

$$P(\mathbf{X}, \hat{\mathbf{X}}) = \frac{1}{\pi} \int_{\theta=0}^{\pi/2} \prod_{f=1}^F \left[ \frac{1+K}{1+K + (E_s/4N_0)(\beta/\sin^2\theta) \sum_{p=1}^P |x_p^f - \hat{x}_p^f|^2} \times \exp\left( -\frac{K(E_s/4N_0)(\beta/\sin^2\theta) \sum_{p=1}^P |x_p^f - \hat{x}_p^f|^2}{1+K + (E_s/4N_0)(\beta/\sin^2\theta) \sum_{p=1}^P |x_p^f - \hat{x}_p^f|^2} \right) \right]^{MN} d\theta, \quad (20)$$

which was previously presented in [5]. It is also interesting to note that simply by setting  $\theta = \pi/2$  in (17) and (20), the classical Chernoff bound would be obtained for shadowed and unshadowed Rician channels, respectively.

For sufficiently large signal-to-noise ratios (i.e.,  $E_s/N_0 \gg 1$ ), evaluating the integrand in (17) at  $\theta = \pi/2$ , we obtain a Chernoff-type bound as

$$P(\mathbf{X}, \hat{\mathbf{X}}) \leq \left( \frac{E_s}{4N_0} \right)^{-|\Psi|NM} \prod_{f=1}^{|\Psi|} \left( \frac{\beta}{M} \sum_{p=1}^P |x_p^f - \hat{x}_p^f|^2 \right)^{-NM} [q(\sigma, \sigma_\mu, m_\mu)]^{|\Psi|NM}, \quad (21)$$

where

$$q(\sigma, \sigma_\mu, m_\mu) = \frac{1}{2\sigma^2} \exp\left(-\frac{1}{2\sigma^2} \exp(2m_\mu)\right) \times \left[ 1 + \sum_{\substack{k=2 \\ k:\text{even}}}^{\infty} \frac{(k-1)!!}{k!} (2\sigma_\mu)^k \sum_{d=1}^k g_{k,d} \left( \frac{1}{2\sigma^2} \exp(2m_\mu) \right)^d \right]. \quad (22)$$

Here,  $\Psi$  is the set of inner frames (with a length of  $L$  symbols) at nonzero Euclidean distance summations and  $|\Psi|$  is the number of elements in this set. This can be compared to *effective length* (EL) in TCM schemes [12], which is defined as the smallest number of symbols at nonzero Euclidean distances. Contrary to the symbol-by-symbol count in the definition of EL, frame-by-frame count is considered here as a result of the multidimensional structure of STBC spanning an interval of  $L$  symbols. It should also be noted that symbol-by-symbol interleaving is considered for the single antenna case while an  $L$ -symbol interleaver is employed in our case. In (21), the slope of the performance curve, which yields the diversity order, is determined by  $|\Psi|NM$  and it can be defined as *generalized effective length* (GEL) for multiple antenna systems in an analogy to the effective length for single antenna case.

The second term in (21) contributes to the coding gain, which corresponds to the horizontal shift in the performance curve. Recalling the definition of *product distance* (PD) for the single antenna case (which is given as the product of nonzero branch distances along the error event), we now define the *generalized product distance* (GPD)

$$\prod_{f=1}^{|\Psi|} \left( \frac{\beta}{M} \sum_{p=1}^P |x_p^f - \hat{x}_p^f|^2 \right)^{-NM} \quad (23)$$

which involves the product of nonzero branch distance summations, where the summation is over  $P$  terms based on the STBC used.

The third term in (21) is completely characterized by channel parameters. Since maximization of diversity order is the primary design criterion, the first step in “good” code design is the maximization of  $|\Psi|$ , since  $M$  and  $N$  are already fixed. Once diversity order is optimized, the third term becomes just a constant. This makes us conclude that the GEL and GPD are the appropriate performance criteria in the selection of trellis codes over shadowed Rician channels. This also shows that the trellis codes designed for optimum performance (based on classical effective code length and minimum product distance) over fading channels for the single transmit antenna case are not necessarily optimum for the multiple antenna case.

To derive the upper bound on bit error probability from the exact PEP, we follow the classical transfer function approach. The upper bound is given in terms of the transfer

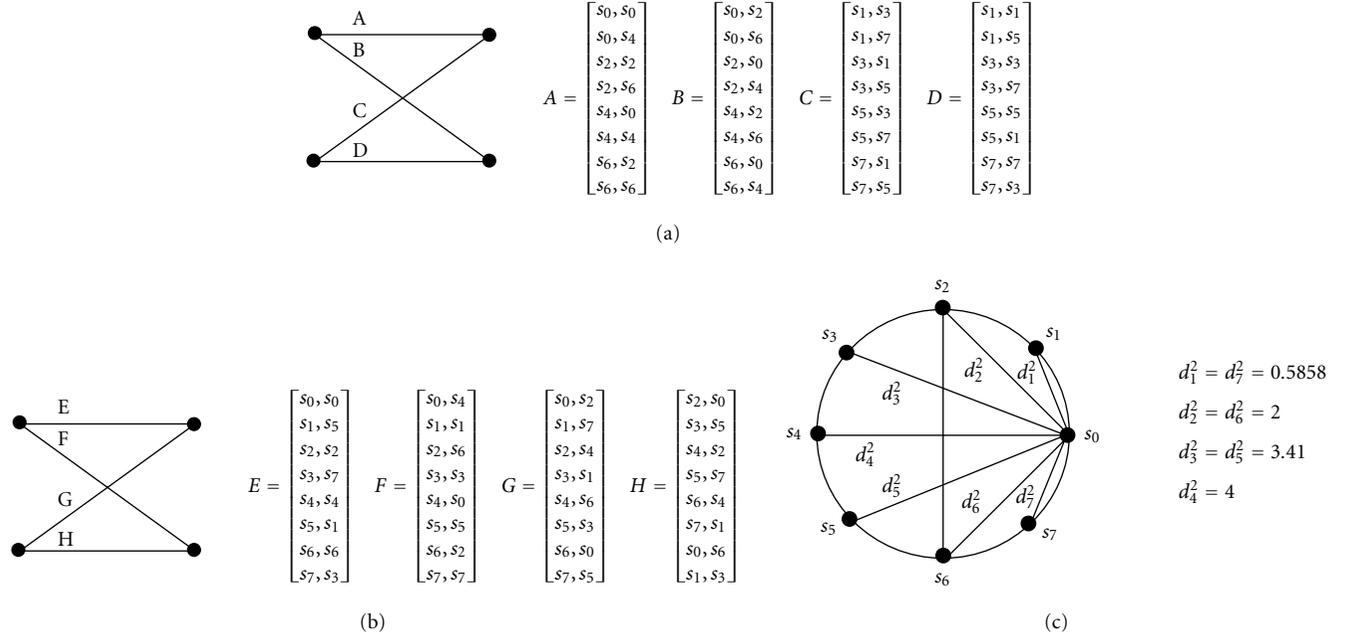


FIGURE 1: (a) Code A2, optimum for AWGN, (b) Code F2, optimum for Rayleigh fading channels with one transmit antenna, (c) 8-PSK signal constellation.

function of the code  $T(D, I)$  by [8, 12]

$$P_b \leq \frac{1}{\pi} \int_0^{\pi/2} \frac{1}{n_b} \frac{\partial}{\partial I} T(D(\theta), I) \Big|_{I=1} d\theta, \quad (24)$$

where  $n_b$  is the number of input bits per transition and  $T(D(\theta), I)$  is the modified transfer function of the code, where  $D(\theta)$ , is given in our case, by

$$D(\theta) = \left(1 + \frac{\Omega_f}{\sin^2 \theta}\right)^{-MN} \exp(-MN \Delta_f(\theta)) \times \left[1 + \sum_{\substack{k=2 \\ k:\text{even}}}^{\infty} \frac{(k-1)!!}{k!} (2\sigma_\mu)^k \sum_{d=1}^k g_{k,d}(\Delta_f(\theta))^d\right]^{MN} \quad (25)$$

based on the derived PEP in (17).

#### 4. EXAMPLES

In this section, we consider two different TCM schemes as outer codes whose trellis diagrams are illustrated in Figure 1. These are 2-state 8-PSK-MTCM codes with 2 symbols per branch, which are optimized for best performance over AWGN and Rayleigh fading channels, respectively [12]. For convenience, we summarize the important parameters of these codes from [12]. The free distance of the code A2 is  $d_{\text{free}}^2 = 3.172$ . Its minimum EL is determined by the error event path of  $\{s_0, s_4\}$ , which differs by one symbol from the correct path (the all-zeros path is assumed to be the correct path based on the uniform properties of the code) achieving

TABLE 1: Parameters for various degrees of shadowing.

Parameter	Light	Average	Heavy
$\sigma^2$	0.158	0.126	0.0631
$m_\mu$	0.115	-0.115	-3.91
$\sigma_\mu$	0.115	0.161	0.806

EL = 1. The corresponding PD is  $d_4^2 = 4$ . On the other hand, the code F2 has a free distance of  $d_{\text{free}}^2 = 2.343$  and it achieves EL = 2 with a product distance of  $d_1^2 \times d_5^2 = 2$ , which is determined by the error event path of  $\{s_1, s_5\}$ . Since EL is the primary factor affecting performance (PD as a secondary factor) over fading channels, F2 is expected to have better performance than A2.

As an example of the shadowed Rician model, we consider the Canadian mobile satellite channel [6]. Table 1 shows the values of shadowing parameters for this channel, which are determined by empirical fit to measured data within Canada. In this table, the terms *light*, *average*, and *heavy* are used to represent an increasing effect of the shadowing.

The upper bounds for both codes with the single transmit antenna are illustrated in Figure 2. No STBC is considered in this case. As expected for the single transmit antenna case, F2 performs better than A2, where the performance is determined by the choices of EL and PD. This observation holds for all considered degrees of shadowing.

In Figure 3, upper bounds for the concatenated scheme are illustrated. Here we use the STBC designed for 2-TX antenna (i.e., Alamouti's code). Based on this code, we have  $P = L = M = 2$  and  $\beta = 1$ . Our results demonstrate that

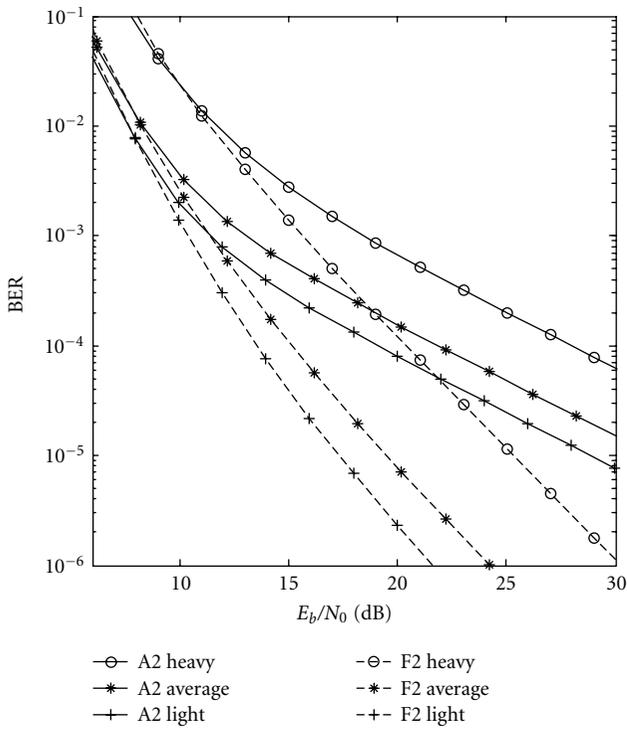


FIGURE 2: Upper bounds for codes A2 and F2 with single transmit antenna over shadowed Rician channels (1-TX and 1-RX antenna).

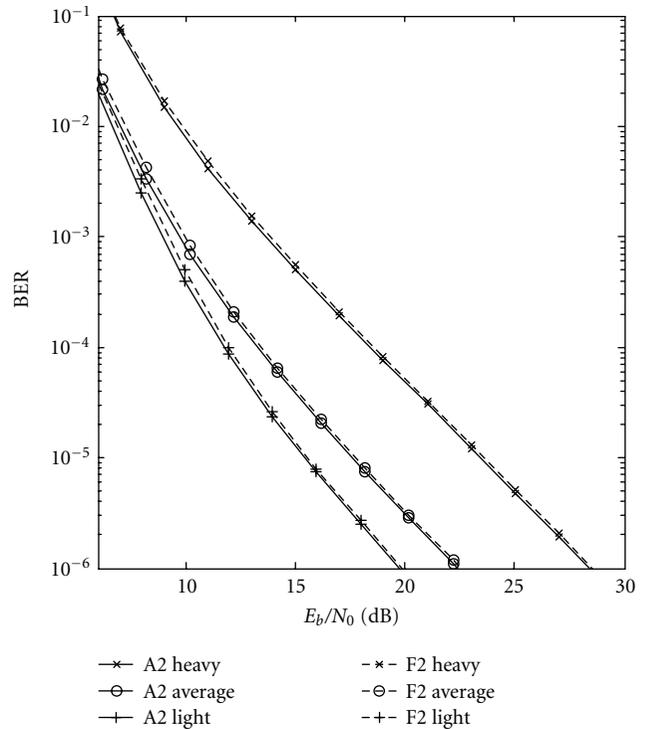


FIGURE 3: Upper bounds for concatenated MTCM-STBC schemes with codes A2 and F2 as outer codes over shadowed Rician channels (2-TX and 1-RX antenna).

the concatenated schemes using A2 and F2 as outer trellis codes achieve roughly the *same* performance. This is a result of the fact that the dominant factors for the single antenna case no longer determine performance. In the 2-TX antenna case, both schemes achieve GEL equal to 2 and GPD equal to 4, that is,  $[(d_1^2 + d_3^2)/2]^2 = 4$  for F2 and  $[(d_0^2 + d_4^2)/2]^2 = 4$  for A2, based on (23). Since both of them have equal GEL and GPD, their performances turn out to be almost identical. This observation holds to be true independent of considered degrees of shadowing.

Comparison between the one- and two-transmit-antenna cases also reveals interesting points on the performance. In both figures, code F2 gives a diversity order of 2 (i.e., slope of the curve), regardless of antenna numbers. Only an additional coding gain (i.e., horizontal shift in the curve) is observed with the use of two antennas. However, this result is somewhat a coincidence because of the particular choice of the parameters characterizing this specific example. For the single transmit antenna case, the code F2 has  $EL = 2$  and the performance curve varies with  $(E_b/N_0)^{-2}$ . On the other hand, for the 2-TX antenna case we have  $|\Psi| = 1$ , since an  $L = 2$ -symbol interleaver is used. However, the overall diversity is determined by GEL (i.e.,  $|\Psi|NM = 1 \cdot 1 \cdot 2 = 2$ ), resulting again in the same slope as in the single transmit antenna case.

To examine the tightness of upper bounds, we also evaluate the performance of codes A2 and F2 through computer simulation, assuming 2-TX antennas. Simulation results for the code F2 are illustrated in Figure 4 with the corresponding

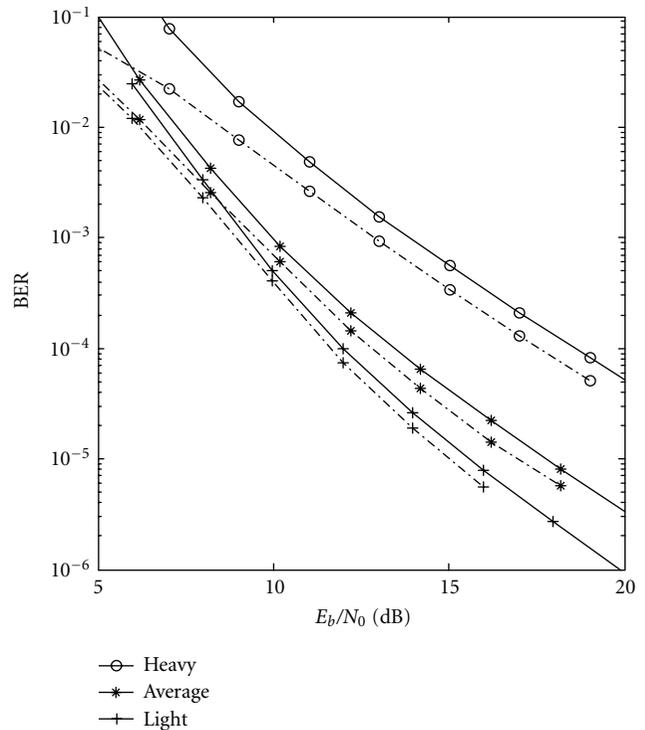


FIGURE 4: Upper bounds versus simulation results for code F2 (solid: upper bounds, dashed: simulation).

upper bounds (plotted as solid lines) computed by (24) and (25). The upper bounds are in very good agreement with simulation results, demonstrating the tightness of the new upper bounds based on the exact PEP. As expected (based on our previous discussion on upper bound expressions), code A2 yields nearly identical simulation results to those of code F2, which we do not include here for brevity.

## 5. CONCLUSION

We analyzed the performance of trellis-coded STBC schemes over shadowed Rician fading channels. Our analysis is based on the derivation of an exact PEP through the moment generating function approach. The derived expression provides insight into the selection criteria for trellis codes which should be used in conjunction with STBC over fading channels. Our results also show that the trellis codes designed for optimum performance over Rician channels with single transmit antenna are not necessarily optimum for the multiple transmit antenna case. Using transfer function techniques based on the new PEP, we present upper bounds on the bit error probability for the concatenated scheme. We also provide simulation results, which seem to be in good agreement with the derived upper bounds.

## APPENDIX

This appendix evaluates the inner integral in (16) in terms of an infinite sum. Defining

$$a = \frac{1}{2\sigma^2} \frac{\Omega_f / \sin^2 \theta}{1 + \Omega_f / \sin^2 \theta}, \quad b = 2\sqrt{2}\sigma_\mu, \quad c = 2m_\mu, \quad (\text{A.1})$$

we can write the inner integral in (16) as

$$\int_{-\infty}^{\infty} \exp(-u^2) f(u) du \quad (\text{A.2})$$

with  $f(u) = \exp(-a \exp(bu + c))$ . Expanding  $f(u)$  in Taylor series, we obtain

$$\int_{-\infty}^{\infty} \exp(-u^2) f(u) du = \sum_{k=0}^{\infty} \frac{f^k(0)}{k!} \int_{-\infty}^{\infty} u^k \exp(-u^2) du, \quad (\text{A.3})$$

where  $f^k(0)$  are the Taylor series coefficients and, in our case, they can be determined as

$$f^k(0) = \exp(-a \exp(c)) b^k \sum_{d=1}^k g_{k,d} (a \exp(c))^d, \quad (\text{A.4})$$

where  $g_{k,d}$  can be computed by the recursive equation

$$g_{k,d} = d g_{k-1,d} - g_{k-1,d-1} \quad \text{with } g_{k,1} = -1 \text{ for } k = 1, 2, \dots, \\ g_{k,d} = 0 \quad \text{for } d > k. \quad (\text{A.5})$$

Using the integral form given by [11, page 382, equation 3.462.1], it can easily be shown that the integral in (A.3) is zero for the odd values of  $k$ . For even values of  $k$ , we can use the result [11, page 382, equation 3.461.4] and express (A.3) as

$$\int_{-\infty}^{\infty} \exp(-u^2) f(u) du = \sqrt{\pi} \sum_{k=0}^{\infty} \frac{f^k(0)}{k!} \frac{(k-1)!!}{2^{k/2}}. \quad (\text{A.6})$$

Replacing (A.2) by (A.6) with  $a$ ,  $b$ , and  $c$  values given as in (A.1), one can obtain the final form for the inner integral of (16) leading to (17).

## ACKNOWLEDGMENT

This paper was presented in part at IEEE Vehicular Technology Conference (VTC-Fall '02), Vancouver, Canada, October 2002.

## REFERENCES

- [1] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744–765, 1998.
- [2] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1456–1467, 1999.
- [3] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, 1998.
- [4] S. M. Alamouti, V. Tarokh, and P. Poon, "Trellis-coded modulation and transmit diversity: design criteria and performance evaluation," in *Proc. IEEE 1998 International Conference on Universal Personal Communications*, vol. 1, pp. 703–707, Florence, Italy, October 1998.
- [5] M. Uysal and C. N. Georghiades, "Analysis of concatenated trellis coded STBC schemes over Rician fading channels," in *Proc. 39th Annual Allerton Conference on Communication, Control and Computing*, Monticello, Ill, USA, October 2001.
- [6] C. Loo, "A statistical model for a land mobile satellite link," *IEEE Trans. Vehicular Technology*, vol. 34, no. 3, pp. 122–127, 1985.
- [7] M. K. Simon and M.-S. Alouini, "A unified approach to the performance analysis of digital communication over generalized fading channels," *Proceedings of the IEEE*, vol. 86, no. 9, pp. 1860–1877, 1998.
- [8] M. K. Simon and M.-S. Alouini, *Digital Communication over Fading Channels: A Unified Approach to Performance Analysis*, John Wiley & Sons, New York, NY, USA, 2000.
- [9] G. L. Turin, "The characteristic function of Hermitian quadratic forms in complex normal variables," *Biometrika*, vol. 47, no. 1/2, pp. 199–201, 1960.
- [10] A. M. Mathai and S. B. Provost, *Quadratic Forms in Random Variables: Theory and Applications*, Marcel Dekker, New York, NY, USA, 1992.
- [11] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*, Academic Press, San Diego, Calif, USA, 5th edition, 1994.
- [12] E. Biglieri, D. Divsalar, P. J. McLane, and M. K. Simon, *Introduction to Trellis-Coded Modulation with Applications*, Macmillan Publishing, New York, NY, USA, 1991.

**M. Uysal** was born in Istanbul, Turkey, in 1973. He received the B.S. and the M.S. degrees in electronics and communication engineering from Istanbul Technical University, Istanbul, Turkey, in 1995 and 1998, respectively, and the Ph.D. degree in electrical engineering from Texas A&M University, Texas, in 2001. From 1995 to 1998, he worked as a Research and Teaching Assistant in the Communication Theory Group



at Istanbul Technical University. From 1998 to 2002, he was affiliated to the Wireless Communication Laboratory, Texas A&M University. During the fall of 2000, he worked as a Research Intern at AT&T Labs-Research, New Jersey. In April 2002, he joined the Department of Electrical and Computer Engineering, University of Waterloo, Canada, as an Assistant Professor. His research interests lie in communications theory with special emphasis on wireless applications. Specific areas include space-time coding, diversity techniques, coding for fading channels, and performance analysis over fading channels. Dr. Uysal currently serves as an Editor for IEEE Transactions on Wireless Communications and as the Guest Coeditor for Special Issue on "MIMO Communications" of Wiley Journal on Wireless Communications and Mobile Computing.

**C. N. Georghiades** received his doctorate in electrical engineering from Washington University in May 1985. Since September 1985 he has been with the Electrical Engineering Department at Texas A&M University where he is a Professor and holder of the Delbert A. Whitaker Endowed Chair. His general interests are in the application of information, communication, and estimation theories to the study of communication systems, with particular interest in wireless and optical systems. Dr. Georghiades has served over the years in several editorial positions with the IEEE Information Theory and Communication Societies and has been involved in organizing a number of conferences. He currently serves as Chair of the Fellow Evaluation Committee of the IEEE Information Theory Society and in the Awards Committee of the IEEE Communications Society. He also serves as General Cochair for the IEEE Information Theory Workshop in San Antonio, Texas, in October 2004. Dr. Georghiades was the recipient of the 1995 Texas A&M University College of Engineering Halliburton Professorship and the 2002 E.D. Brockett Professorship. From 1997 to 2002 he held the J. W. Runyon Jr. Endowed Professorship and in 2002 he became the inaugural recipient of the Delbert A. Whitaker Endowed Chair.



# Full-Rate Full-Diversity Linear Quasi-Orthogonal Space-Time Codes for Any Number of Transmit Antennas

**Naresh Sharma**

*Open Innovations Lab, Lucent Technologies, 67 Whippany Road, Whippany, NJ 07981, USA  
Email: nareshs@bell-labs.com*

**Constantinos B. Papadias**

*Wireless Research Lab, Bell Laboratories, Lucent Technologies, 791 Holmdel-Keyport Road, Holmdel, NJ 07733, USA  
Email: papadias@bell-labs.com*

*Received 31 May 2003; Revised 5 January 2004*

We construct a class of linear quasi-orthogonal space-time block codes that achieve full diversity over quasistatic fading channels for any transmit antennas. These codes achieve a normalized rate of one symbol per channel use. Constellation rotation is shown to be necessary for the full-diversity feature of these codes. When the number of transmit antennas is a power of 2, these codes are also delay “optimal.” The quasi-orthogonal property of the code makes one half of the symbols orthogonal to the other half, and we show that this allows each half to be decoded separately without any loss of performance. We give an iterative construction of these codes with a practical decoding algorithm. Numerical simulations are presented to evaluate the performance of these codes in terms of capacity as well as probability of error versus SNR curves. For some special cases, we compute the pairwise probability of error averaged over all the channel states as a single integral that shows the diversity and coding gain more clearly.

**Keywords and phrases:** multiple antennas, space-time codes, diversity, orthogonal designs, wireless communications.

## 1. INTRODUCTION

Multiple antenna systems have been of great interest in recent times because of their ability to support higher data rates at the same bandwidth and noise conditions; see, for example, [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11] and references therein.

For two transmit antennas, Alamouti’s orthogonal design gave a full-rate space-time block code with full diversity [6, 12]. More general orthogonal designs were later proposed by Tarokh et al. and Tirkkonen that had simple single symbol decoders while offering full diversity [7, 13]. Recently, complex orthogonal designs with maximal rates have been proposed by Liang where the entries are restricted to be the complex modulated symbols or their conjugates with or without a sign change [14]. The upper bounds of the rates of generalized complex orthogonal space-time block codes were given in [15].

One of the key aspects of orthogonal designs has been to ensure diversity for *any* symbol constellation. For more than two transmit antennas and complex constellations, these codes offered on the average a rate of less than one symbol per channel use, where each symbol time period corresponds to a channel use. The highest theoretical code rate for full-

diversity code when the symbols are constrained to be chosen from the *same* constellation was shown to be one symbol per channel use (see [5, Corollary 3.3.1]). (This constraint is relaxed by using rotated constellations and indeed many of the recent papers give space-time codes that offer full diversity for more than one symbols per channel use [16, 17]. We discuss this point further below.)

More recently, a different approach has been attempted to yield the full diversity where the notion of diversity is made specific to a constellation, and this is also referred to as modulation diversity [18]. More specifically, it has been shown that full-rate and full-modulation diversity is achievable with constellation rotation or linear constellation precoding [18, 19], where the transmitted signal is a multiplication of a unitary matrix with a diagonal matrix whose diagonal elements are a function of linearly precoded (or rotated) information symbols. This makes the test of full diversity or the rank criterion trivial by ensuring with proper precoding or constellation rotation that no element in the diagonal becomes zero while taking the difference of two distinct codewords. A similar idea has been presented before in [20] for rotated binary phase shift keying (BPSK) modulation.

The issue of smaller code rate (less than one symbol per channel use) for complex orthogonal designs has been addressed in recent times by the design of quasi-orthogonal codes for achieving higher data rates [21, 22, 23, 24]. The quasi-orthogonal codes were given for 4 transmit antennas with rate 1, and 8 transmit antennas with rate 3/4. These codes sacrificed some orthogonality by making *subsets* of symbols orthogonal to each other instead of making every single symbol orthogonal to any other. Because of this relaxation of constraints, these codes achieve higher code rates that were hitherto not possible with orthogonal codes. It was shown in [25] that performance of the above quasi-orthogonal codes can be improved with constellation rotation. Constellation rotation has also been discussed in [26] as a technique to improve the performance of space-time block codes.

In this paper, we build on earlier work on orthogonal designs and achieving modulation diversity by constellation rotation to propose a quasi-orthogonal structure to iteratively construct full-diversity space-time codes for *any* transmit antennas. These codes have half the symbols orthogonal to the other half, which allows each orthogonal half to be decoded separately without any loss of performance. Hence the decoding complexity of such a code is considerably smaller. We show that these codes achieve full diversity with appropriate constellation rotations. If the transmit antennas are a power of 2, then these codes are also delay “optimal,” that is, the length of block code in symbol periods is same as the number of transmit antennas [27]. We present the numerical results for these codes in terms of probability of error and we also provide a Shannon capacity perspective to these codes.

We use the following notation throughout the paper:  $T$  and  $H$  denote the transpose and conjugate transpose, respectively, of a matrix or a vector;  $\mathbf{I}^M$  and  $\mathbf{0}^M$  are  $M \times M$  identity and null matrices, respectively;  $\|A\|_F$  and  $\text{Tr}(A)$  denote Frobenius norm and trace of matrix  $A$ , respectively;  $Q$ -function is given by  $Q(x) \triangleq \int_x^\infty e^{-u^2/2} du / \sqrt{2\pi}$ ;  $n!$  denotes the factorial of  $n$  for any nonnegative integer  $n$ ;  $\mathbb{C}$  denotes the complex number field;  $\mathbb{C}^P$  denotes a vector of length  $P$  whose elements are taken from  $\mathbb{C}$ ;  $\mathbb{C}^{P \times Q}$  denotes a  $P \times Q$  matrix whose elements are taken from  $\mathbb{C}$ ;  $j$  denotes an integer index or  $\sqrt{-1}$ , where the actual value will be evident from the context;  $\text{Re}(x)$  and  $\text{Im}(x)$  denote the real and imaginary parts of a complex number  $x$  respectively;  $\mathcal{CN}(0, 1)$  indicates a zero mean and circularly symmetric complex Gaussian variable with unit variance;  $\det\{A\}$  denotes the determinant of a square matrix  $A$ .

## 2. SYSTEM MODEL

Consider a system of  $M$  transmit and  $N$  receive antennas that we refer to as  $(M, N)$  system in this paper. The modulated information symbols to be transmitted are taken  $Q$  at a time to form a  $Q \times 1$  vector denoted by  $\mathbf{c} = (c_1, \dots, c_Q)^T$ . This information vector is precoded (i.e., multiplied) by a  $Q \times Q$  unitary rotation matrix denoted by  $\mathcal{R}_Q$ . Let  $\mathbf{s} = (s_1, \dots, s_Q)^T$  and

$$\mathbf{s} = \mathcal{R}_Q \mathbf{c}. \quad (1)$$

This precoded vector  $\mathbf{s}$  is then passed on to a linear space-time block code that generates a  $T \times M$  matrix  $G_Q[\mathbf{s}]$  given by

$$G_Q[\mathbf{s}] = \sum_{q=1}^Q (C_q s_q + D_q s_q^*), \quad (2)$$

where  $C$ 's and  $D$ 's are  $T \times M$  complex matrices, which completely specify the code. This matrix is transmitted in  $T$  channel uses (each channel use is a symbol time period). The average code rate for this system is hence  $Q/T$  symbols per channel use.

For quasistatic fading channel, the received signal is given by

$$X(\mathbf{s}) = \sqrt{\frac{\rho}{M}} G_Q[\mathbf{s}] H + V, \quad (3)$$

where  $X$  and  $V$  are the  $T \times N$  received and noise matrices, and  $H$  is the  $M \times N$  complex channel matrix that is assumed to be constant over  $T$  channel uses and varies independently over the next  $T$  channel uses and so on. The entries of  $H$  and  $V$  are assumed to be mutually independent and  $\mathcal{CN}(0, 1)$ , and  $\rho$  is the average SNR per received antenna. We assume that the channel is perfectly known at the receiver but is unknown at the transmitter.

### 2.1. Design criterion

It has been shown in [5] by examining the pairwise probability of error between two distinct information vectors (say  $\mathbf{c}, \mathbf{e} \in \mathbb{C}^Q$ ) that for full diversity, in quasistatic fading channels,  $G_Q^H[\mathcal{R}_Q(\mathbf{c} - \mathbf{e})] G_Q[\mathcal{R}_Q(\mathbf{c} - \mathbf{e})]$  should have a rank of  $M$  (rank criterion). We assume here that  $T \geq M$ . If for some  $M$ ,  $T = M$ , then the rank criterion could be modified to yield the following: for full diversity, and  $\mathbf{c} \neq \mathbf{e}$ ,

$$\det \{G_Q[\mathcal{R}_Q(\mathbf{c} - \mathbf{e})]\} \neq 0. \quad (4)$$

We will examine this criterion in the context of proposed codes. In addition, we will examine the coding gain for quasistatic fading channels that is defined to be

$$\min_{\mathbf{c}, \mathbf{e}} \left( \prod_{i=1}^M \lambda_i \right)^{1/r}, \quad (5)$$

where  $\lambda_i$ ,  $i = 1, \dots, r$ , are the nonzero eigenvalues of the  $M \times M$  matrix  $G_Q^H[\mathcal{R}_Q(\mathbf{c} - \mathbf{e})] G_Q[\mathcal{R}_Q(\mathbf{c} - \mathbf{e})]$ . For  $T = M$  and for a full-diversity achieving code, the coding gain can be simplified as

$$\min_{\mathbf{c}, \mathbf{e}} |\det \{G_Q[\mathcal{R}_Q(\mathbf{c} - \mathbf{e})]\}|^{2/M}. \quad (6)$$

## 3. LINEAR QUASI-ORTHOGONAL CODES

Partition vector  $\mathbf{s}$  (defined in Section 2) into  $Q/L$  parts where  $L$  divides  $Q$ . These partitions are disjoint and for the purposes of this paper, we will assume that all partitions contain  $L$  symbols. We describe these partitions by a set of functions  $\mathcal{A}_i$ ,  $i = 1, \dots, Q/L$ , where  $\mathcal{A}_i(\mathbf{s})$  is a  $Q$  length vector that has symbols in indices belonging to it and zeros in

all other indices. For example, if the first partition has the first two and the last symbols belonging to it, then  $\mathcal{A}_1(\mathbf{s}) = (s_1, s_2, 0, \dots, 0, s_Q)$ . If the  $k$ th element of the vector denoted by  $\mathcal{A}_i^k(\mathbf{s})$  is nonzero, then  $\mathcal{A}_j^k(\mathbf{s}) = 0$  for all  $j \neq i$ ,  $j = \{1, \dots, Q/L\}$ . This follows since the partitions are disjoint. For disjoint partitions, it follows from linearity that

$$G_Q[\mathbf{s}] = \sum_{i=1}^{Q/L} G_Q[\mathcal{A}_i(\mathbf{s})]. \quad (7)$$

We define a linear quasi-orthogonal code over partitions given by  $\mathcal{A}_i$ ,  $i = 1, \dots, Q/L$ , to be the one that satisfies

$$G_Q^H[\mathbf{s}]G_Q[\mathbf{s}] \triangleq \sum_{i=1}^{Q/L} G_Q^H[\mathcal{A}_i(\mathbf{s})]G_Q[\mathcal{A}_i(\mathbf{s})] \quad \forall \mathbf{s} \in \mathbb{C}^Q. \quad (8)$$

Hence the partitions are completely decoupled from each other when we take this product and this is true for *any* complex vector  $\mathbf{s}$ . Note that the quasi-orthogonal property is defined for any  $\mathbf{s} \in \mathbb{C}^Q$ , while the approach we adopt later to prove full diversity is specific to the choice of modulation constellation.

### 3.1. Properties

**Proposition 1.** *A Linear space-time code is a quasi-orthogonal code if and only if any of the following holds:*

$$G_Q^H[\mathcal{A}_i(\mathbf{s})]G_Q[\mathcal{A}_j(\mathbf{s})] + G_Q^H[\mathcal{A}_j(\mathbf{s})]G_Q[\mathcal{A}_i(\mathbf{s})] = \mathbf{0}^M, \quad i \neq j; \quad (9)$$

$$C_i^H C_j + D_j^H D_i = C_i^H D_j + C_j^H D_i = \mathbf{0}^M, \quad s_i, s_j \notin \mathcal{A}_k(\mathbf{s}) \quad \forall k; \quad (10)$$

$$G_Q^H[\mathbf{s}]G_Q[\mathbf{c}] = \sum_{i=1}^{Q/L} G_Q^H[\mathcal{A}_i(\mathbf{s})]G_Q[\mathcal{A}_i(\mathbf{c})] \quad \forall \mathbf{s}, \mathbf{c} \in \mathbb{C}^Q. \quad (11)$$

*Proof.* Using linearity in (7), the left-hand side of (8) is given by

$$\begin{aligned} & \sum_{i=1}^{Q/L} G_Q^H[\mathcal{A}_i(\mathbf{s})]G_Q[\mathcal{A}_i(\mathbf{s})] \\ & + \sum_{i=1}^{Q/L} \sum_{j=i+1}^{Q/L} G_Q^H[\mathcal{A}_i(\mathbf{s})]G_Q[\mathcal{A}_j(\mathbf{s})] \\ & + G_Q^H[\mathcal{A}_j(\mathbf{s})]G_Q[\mathcal{A}_i(\mathbf{s})]. \end{aligned} \quad (12)$$

Using (9) in the above equation, (8) follows. Suppose that (9) does not hold, then using equation (12), it follows that

$$G_Q^H[\mathbf{s}]G_Q[\mathbf{s}] \neq \sum_{i=1}^{Q/L} G_Q^H[\mathcal{A}_i(\mathbf{s})]G_Q[\mathcal{A}_i(\mathbf{s})], \quad (13)$$

which contradicts (8).

Let  $G_Q[\mathcal{A}_j(\mathbf{s})] = \sum_{l=1}^L C_l s_{l_k} + D_l s_{l_k}^*$  with  $l = i, j$ . Then the left-hand side of (9) is given by

$$\sum_{p,q=1}^L X_1 s_{ip} s_{jq} + (X_1 s_{iq} s_{jq})^H + X_2 s_{ip} s_{jq}^* + (X_2 s_{ip} s_{jq}^*)^H, \quad (14)$$

where  $X_1 = D_{i_p}^H C_{j_q} + D_{j_q}^H C_{i_p}$  and  $X_2 = C_{j_q}^H C_{i_p} + D_{i_p}^H D_{j_q}$ . Using (10),  $X_1 = X_2 = \mathbf{0}^M$ , hence (9) and (8) hold. Conversely, if (10) does not hold, then  $X_1 \neq \mathbf{0}^M$  and  $X_2 \neq \mathbf{0}^M$ , which contradicts (9) and hence also (8).

Define a new vector  $\mathbf{z}$  whose  $i$ th and  $j$ th partitions are the same as  $\mathbf{s}$  and  $\mathbf{c}$  with  $i \neq j$ . Then using (9), we have

$$G_Q^H[\mathcal{A}_i(\mathbf{s})]G_Q[\mathcal{A}_j(\mathbf{c})] + G_Q^H[\mathcal{A}_j(\mathbf{c})]G_Q[\mathcal{A}_i(\mathbf{s})] = \mathbf{0}^M. \quad (15)$$

We can do this over all  $i, j$  with  $i \neq j$ . Then expanding the left-hand side of (11) along similar lines as in (12), (11) follows immediately. Conversely, if (11) is not true, then substituting  $\mathbf{s} = \mathbf{c}$  contradicts (8).  $\square$

**Proposition 2.** *Maximum likelihood (ML) decoding of a linear quasi-orthogonal code with received signal model given by (3) is equivalent to ML decoding of each partitions individually by taking the channel model as*

$$X(\mathcal{A}_i(\mathbf{s})) = \sqrt{\frac{\rho}{M}} G_Q[\mathcal{A}_i(\mathbf{s})]H + V. \quad (16)$$

*Proof.* ML decoding is given by

$$\begin{aligned} \hat{\mathbf{s}} &= \arg \min_{\mathbf{z}} \left\| X(\mathbf{s}) - \sqrt{\frac{\rho}{M}} G_Q[\mathbf{z}]H \right\|_F^2 \\ &= \arg \min_{\mathbf{z}} \text{Tr} \left\{ \frac{\rho}{M} H^H G_Q^H[\mathbf{z}]G_Q[\mathbf{z}]H \right. \\ &\quad \left. - 2\sqrt{\frac{\rho}{M}} \text{Re} [X^H(\mathbf{s})G_Q[\mathbf{z}]H] \right\} \end{aligned} \quad (17a)$$

$$\begin{aligned} &= \arg \min_{\mathbf{z}} \text{Tr} \left\{ \frac{\rho}{M} \sum_{i=1}^{Q/L} H^H G_Q^H[\mathcal{A}_i(\mathbf{z})]G_Q[\mathcal{A}_i(\mathbf{z})]H \right. \\ &\quad \left. - 2\sqrt{\frac{\rho}{M}} \text{Re} \left( \sqrt{\frac{\rho}{M}} H^H G_Q^H[\mathbf{s}]G_Q[\mathbf{z}]H \right. \right. \\ &\quad \left. \left. + V^H G_Q[\mathbf{z}]H \right) \right\} \end{aligned} \quad (17b)$$

$$\begin{aligned} &= \arg \min_{\mathbf{z}} \text{Tr} \left\{ \frac{\rho}{M} \sum_{i=1}^{Q/L} H^H G_Q^H[\mathcal{A}_i(\mathbf{z})]G_Q[\mathcal{A}_i(\mathbf{z})]H \right. \\ &\quad \left. - 2\sqrt{\frac{\rho}{M}} \text{Re} \left[ \sqrt{\frac{\rho}{M}} H^H \sum_{i=1}^{Q/L} G_Q^H[\mathcal{A}_i(\mathbf{s})]G_Q[\mathcal{A}_i(\mathbf{z})]H \right. \right. \\ &\quad \left. \left. + V^H \sum_{i=1}^{Q/L} G_Q[\mathcal{A}_i(\mathbf{z})]H \right) \right\} \end{aligned} \quad (17c)$$

$$\begin{aligned} &= \arg \min_{\mathbf{z}} \sum_{i=1}^{Q/L} \text{Tr} \left\{ \frac{\rho}{M} H^H G_Q^H[\mathcal{A}_i(\mathbf{z})]G_Q[\mathcal{A}_i(\mathbf{z})]H \right. \\ &\quad \left. - 2\sqrt{\frac{\rho}{M}} \text{Re} [H^H X^H(\mathcal{A}_i(\mathbf{s}))G_Q[\mathcal{A}_i(\mathbf{z})]H] \right\} \end{aligned} \quad (17d)$$

$$\begin{aligned} &= \sum_{i=1}^{Q/L} \arg \min_{\mathcal{A}_i(\mathbf{z})} \text{Tr} \left\{ \frac{\rho}{M} H^H G_Q^H[\mathcal{A}_i(\mathbf{z})]G_Q[\mathcal{A}_i(\mathbf{z})]H \right. \\ &\quad \left. - 2\sqrt{\frac{\rho}{M}} \text{Re} [H^H X^H(\mathcal{A}_i(\mathbf{s}))G_Q[\mathcal{A}_i(\mathbf{z})]H] \right\}, \end{aligned} \quad (17e)$$

which is similar to (17a) and hence the effective channel model is given by (16). In (17a), we have used the fact that  $\|A\|_F^2 = \text{Tr}(A^H A)$ ; in (17b), we have used (3) and (8); in (17c), we have used (7) and (11); in (17d), we have used the definition of  $X^H(\mathcal{A}_i(\mathbf{s}))$  from (16); and in (17e) the fact that  $\text{Tr}(\cdot)$  is a linear operation.  $\square$

### 3.2. Construction of a class of linear quasi-orthogonal codes for any $M$

We construct a class of quasi-orthogonal codes that achieve full rate for any transmit antennas. The construction of the code is iterative that ensures its quasi-orthogonal structure. We will first consider the case of  $M$  being a power of 2. A case of other  $M$  is dealt with later in this section.

#### 3.3. $M$ a power of 2

Consider an  $M \times M$  code for  $M$  transmit antennas that encodes  $M$  symbols together and transmits the block code in  $M$  channel uses, where  $M$  is a power of 2. Hence  $Q = T = M$  and the code rate for this code is 1. We will consider quasi-orthogonal codes with two disjoint partitions with  $M/2$  symbols in each of them (i.e.,  $L = M/2$ ) that are orthogonal to each other in the sense of (9). The two partitions for  $M$  transmit antennas are denoted by  $\mathcal{A}_{M,1}(\mathbf{s})$  and  $\mathcal{A}_{M,2}(\mathbf{s})$ , where a subscript  $M$  is added to show that they are for  $M$  transmit antennas.

We first define the code and partitions for a single transmit antenna as

$$G_1[\mathbf{s}] \triangleq s_1 \quad \forall \mathbf{s} \in \mathbb{C}^1, \quad (18)$$

and  $\mathcal{A}_{1,1}(\mathbf{s}) = s_1$  and  $\mathcal{A}_{1,2}(\mathbf{s}) = 0$ , where  $\mathbf{s} \in \mathbb{C}^1$ .

We assume that the following properties are true for any  $M$ , where  $M$  is a power of 2, and for any  $\mathbf{s}, \mathbf{e} \in \mathbb{C}^M$ :

- (P1)  $G_M^H[\mathcal{A}_{M,1}(\mathbf{s})] = G_M[\mathcal{A}_{M,1}(\mathbf{s}^*)]$ ;
- (P2)  $G_M^H[\mathcal{A}_{M,2}(\mathbf{s})] = -G_M[\mathcal{A}_{M,2}(\mathbf{s})]$ ;
- (P3)  $G_M[\mathcal{A}_{M,1}(\mathbf{e})]G_M[\mathcal{A}_{M,1}(\mathbf{s})] = G_M[\mathcal{A}_{M,1}(\mathbf{s})]G_M[\mathcal{A}_{M,1}(\mathbf{e})]$ ;
- (P4)  $G_M[\mathcal{A}_{M,2}(\mathbf{e})]G_M[\mathcal{A}_{M,2}(\mathbf{s})] = G_M[\mathcal{A}_{M,2}(\mathbf{s}^*)]G_M[\mathcal{A}_{M,2}(\mathbf{e}^*)]$ ;
- (P5)  $G_M^H[\mathcal{A}_{M,1}(\mathbf{s})]G_M[\mathcal{A}_{M,2}(\mathbf{s})] + G_M^H[\mathcal{A}_{M,2}(\mathbf{s})]G_M[\mathcal{A}_{M,1}(\mathbf{s})] = 0$ . Note that by using (P1) and (P2), this can be rewritten as  $G_M[\mathcal{A}_{M,1}(\mathbf{s}^*)]G_M[\mathcal{A}_{M,2}(\mathbf{s})] = G_M[\mathcal{A}_{M,2}(\mathbf{s})]G_M[\mathcal{A}_{M,1}(\mathbf{s})]$ .

#### Iterative construction

We construct a code for  $2M$  transmit antennas that takes a  $2M \times 1$  precoded vector  $\mathbf{s}$  as input. For simplicity of notation, we will denote the first  $M$  elements of  $\mathbf{s}$  by  $\mathbf{s}_{M,1}$  and the last  $M$  by  $\mathbf{s}_{M,2}$ . Then the quasi-orthogonal code for  $2M$  antennas is constructed as

$$\mathcal{A}_{2M,1}(\mathbf{s}) = \mathcal{A}_{M,1}(\mathbf{s}_{M,1}) + \mathcal{A}_{M,2}(\mathbf{s}_{M,2}), \quad (19)$$

$$\mathcal{A}_{2M,2}(\mathbf{s}) = \mathcal{A}_{M,2}(\mathbf{s}_{M,1}) + \mathcal{A}_{M,1}(\mathbf{s}_{M,2}), \quad (20)$$

TABLE 1: Indices of the first partition of the code for various  $M$ .

$M$	Indices of first partition, $\mathcal{I}_{M,1}$
2	1
4	$\mathcal{I}_{2,1}, 4$
8	$\mathcal{I}_{4,1}, 6, 7$
16	$\mathcal{I}_{8,1}, 10, 11, 13, 16$
32	$\mathcal{I}_{16,1}, 18, 19, 21, 24, 25, 28, 30, 31$

and the code for each partition is written as

$$G_{2M}[\mathcal{A}_{2M,1}(\mathbf{s})] = \begin{bmatrix} G_M[\mathcal{A}_{M,1}(\mathbf{s}_{M,1})] & G_M[\mathcal{A}_{M,2}(\mathbf{s}_{M,2})] \\ -G_M[\mathcal{A}_{M,2}(\mathbf{s}_{M,2}^*)] & G_M[\mathcal{A}_{M,1}(\mathbf{s}_{M,1}^*)] \end{bmatrix},$$

$$G_{2M}[\mathcal{A}_{2M,2}(\mathbf{s})] = \begin{bmatrix} G_M[\mathcal{A}_{M,2}(\mathbf{s}_{M,1})] & G_M[\mathcal{A}_{M,1}(\mathbf{s}_{M,2})] \\ -G_M[\mathcal{A}_{M,1}(\mathbf{s}_{M,2}^*)] & G_M[\mathcal{A}_{M,2}(\mathbf{s}_{M,1}^*)] \end{bmatrix}. \quad (21)$$

By using the linearity equation (7), we have

$$G_{2M}[\mathbf{s}] = G_{2M}[\mathcal{A}_{2M,1}(\mathbf{s})] + G_{2M}[\mathcal{A}_{2M,2}(\mathbf{s})]$$

$$= \begin{bmatrix} G_M[\mathbf{s}_{M,1}] & G_M[\mathbf{s}_{M,2}] \\ -G_M[\mathbf{s}_{M,2}^*] & G_M[\mathbf{s}_{M,1}^*] \end{bmatrix}. \quad (22)$$

For  $M = 1$ , this gives the Alamouti's code [6]. For  $M = 2$  case, this iterative structure along with some similar ones were presented in [23]. Table 1 gives the indices of the first partition denoted by  $\mathcal{I}_{M,1}$  for  $M = 2, 4, 8, 16$ , and 32. Symbols with the same indices as those given in the table form the first partition for the code. These indices come from the construction above. Note that from (19),  $\mathcal{I}_{M,1}$  is a subset of  $\mathcal{I}_{2M,1}$ . The second partition can be obtained by excluding the indices from the first partition.

**Proposition 3.** *The constructed code for  $2M$  transmit antennas in (19) and (20) satisfies properties (P1)–(P5) for any  $M$ , where  $M$  is a power of 2.*

*Proof.* Omitted.  $\square$

Note that (P1)–(P5) are true for  $M = 1$ . If we assume that they hold for any  $M$  with  $M$  a power of 2, then using Proposition 3, it holds for  $2M$ . It follows from induction that the constructed code satisfies (P1)–(P5) for any  $M$ , where  $M$  is a power of 2.

#### 3.3.1. Properties

**Proposition 4.** *For any  $2M \times 1$  vector  $\mathbf{z}$ , a transformation denoted by  $\hat{\mathbf{z}}$  is defined that interchanges the two halves of  $\mathbf{z}$  with a sign change for the second half, that is,  $\hat{\mathbf{z}} = [-\mathbf{z}(M+1 : 2M)\mathbf{z}(1 : M)]$ . Then for any  $4M \times 1$  vector  $\mathbf{s}$ ,*

$$\det\{G_{4M}[\mathcal{A}_{4M,1}(\mathbf{s})]\} = \det\{G_{2M}[\mathcal{A}_{2M,1}(\mathbf{s}_{2M,1} - \hat{\mathbf{s}}_{2M,2})]\}$$

$$\times \det\{G_{2M}[\mathcal{A}_{2M,1}(\mathbf{s}_{2M,1} + \hat{\mathbf{s}}_{2M,2})]\}, \quad (23)$$

where  $\mathbf{s}_{2M,1} = \mathbf{s}(1 : 2M)$  and  $\mathbf{s}_{2M,2} = \mathbf{s}(2M+1 : 4M)$ .

*Proof.* See the appendix.  $\square$

It can similarly be shown that

$$\det \{G_{4M}[\mathcal{A}_{4M,2}(\mathbf{s})]\} = \det \{G_{2M}[\mathcal{A}_{2M,1}(\mathbf{s}_{2M,2} - \hat{\mathbf{s}}_{2M,1})]\} \\ \times \det \{G_{2M}[\mathcal{A}_{2M,1}(\mathbf{s}_{2M,2} + \hat{\mathbf{s}}_{2M,1})]\}. \quad (24)$$

We omit the proof because of similarity with Proposition 4. We will use Proposition 4 to prove the full diversity. For  $M = 2$ , we obtain by calculation  $\det\{G_2[\mathcal{A}_{2,1}(\mathbf{s})]\} = |s_1|^2$ . For  $4M = 4$ , we use (23) to get  $\det\{G_4[\mathcal{A}_{4,1}(\mathbf{s}_1)]\} = |s_1 - s_4|^2 |s_1 + s_4|^2$  and for  $4M = 8$ , we get

$$\det \{G_8[\mathcal{A}_{8,1}(\mathbf{s}_1)]\} = |s_1 - s_7 + s_4 + s_6|^2 |s_1 - s_7 - s_4 - s_6|^2 \\ \times |s_1 + s_7 + s_4 - s_6|^2 |s_1 + s_7 - s_4 + s_6|^2. \quad (25)$$

**Proposition 5.** Let  $\mathcal{A}_{2M,1}(\mathbf{s}) = \{s_{k_1}, \dots, s_{k_M}\}$  and define a constellation  $\mathcal{C} = \{\sum_{j=1}^M s_{k_j}\}$ . Let  $d_{M,\min}(\mathcal{C})$  denote the minimum distance of this constellation. Then to ensure that the code satisfies the rank criterion with a modulation constellation that is invariant under multiplication with  $\pm 1$ , it suffices to show that there exists a pre-coding  $\mathcal{R}_M$  (defined in (1)) that makes  $d_{M,\min}(\mathcal{C}) > 0$ . Further, the coding gain of such a system is  $d_{M,\min}^2(\mathcal{C})$ .

*Proof.* Firstly, we note that due to quasi-orthogonal structure of the code, we need to prove rank criterion for the partitions instead of the full code. Because of the iterative structure in (23), it is clear that for any  $M \geq 2$  and  $M$  a power of 2,  $\det\{G_{2M}[\mathcal{A}_{2M,1}(\mathbf{s})]\}$  is the product of  $M$  terms of the form

$$\left| \sum_{j=1}^M (-1)^{b_j} s_{k_j} \right|^2, \quad (26)$$

where  $b_j = \{0, 1\}$ . If the modulation constellation used for modulated information symbols in  $\mathbf{c}$  in (1) is invariant under the multiplication with  $\pm 1$ , then modulation constellations for precoded symbols  $\mathbf{s}$  are also invariant under the multiplication with  $\pm 1$ , and hence constellation  $\{\sum_{j=1}^M (-1)^{b_j} s_{k_j}\}$  is the same as the constellation  $\mathcal{C}$  for any choice of  $b_j$ ,  $j = 1, \dots, M$ . If  $d_{M,\min}(\mathcal{C}) > 0$ , then for any difference between two distinct precoded vectors  $\mathbf{s}$  and  $\mathbf{e}$ ,  $\det\{G_{2M}[\mathcal{A}_{2M,1}(\mathbf{s} - \mathbf{e})]\} \neq 0$ , which ensures full rank.

The coding gain denoted by  $\delta_{2M}$  is given by (using (6) for  $2M$  transmit antennas)

$$\delta_{2M} = \min_{\mathbf{s}, \mathbf{e}} \left| \det \{G_{2M}[\mathcal{A}_{2M,1}(\mathbf{s} - \mathbf{e})]\} \right|^{1/M} \\ = d_{M,\min}^2(\mathcal{C}). \quad \square \quad (27)$$

The proof for  $G_{2M}[\mathcal{A}_{2M,2}(\mathbf{s})]$  follows along similar lines. The existence of a precoding to guarantee that  $d_{M,\min}(\mathcal{C}) \neq 0$  is shown in [18, 19, 28, 29, 30] and references therein.

We note here that for  $2M$  transmit antennas,  $M$  symbols are precoded together due to quasi-orthogonal structure, while in [18, 19], all  $2M$  are precoded together. Since minimum distance typically decreases as  $M$  increases, we expect the coding gain to be higher than [18, 19]. From [18, equation (6)], the minimum distance for a class of real constellation rotations is dependent on  $M$  as  $d_{M,\min}^2 \sim (M)^{-M}$ .

### 3.3.2. $M$ not a power of 2

Until now we have dealt with only those number of transmit antennas that are a power of 2. To address this issue, we have the following proposition.

**Proposition 6.** A full-diversity quasi-orthogonal code for  $M$  transmit antennas, where  $M$  is not a power of 2, can be obtained by deleting any  $P - M$  columns of  $G_P$ , where  $P = 2^{\lceil \log_2(M) \rceil}$ .

*Proof.* We first prove that this code is quasi-orthogonal. Assume that the last  $P - M$  columns of  $G_P$  are deleted. Then modified received signal model for this code can be rewritten, without any loss of performance using (3), as

$$X(\mathbf{s}) = \sqrt{\frac{\rho}{M}} G_P[\mathbf{s}] \hat{H} + V, \quad (28)$$

where  $\hat{H}$  is a  $P \times N$  matrix whose first  $M$  rows are the same as that  $M \times N$  matrix  $H$ , and the last  $P - M$  rows are null vectors;  $X$  and  $V$  are  $M \times N$  matrices. Since  $G_P$  is quasi-orthogonal allowing the partitions to be separately decoded for any channel realization, then decoding for any  $M$  can also be accomplished by decoding each partition separately.

It follows from linearity that  $G_M[\mathcal{A}_{M,i}(\mathbf{s})]$  ( $i = 1, 2$ ) is obtained from  $G_P[\mathcal{A}_{M,i}(\mathbf{s})]$  by deleting its last  $P - M$  columns. Since  $G_P[\mathcal{A}_{P,1}(\mathbf{s})]$  is full rank, that is, with rank  $P$ , then deleting  $P - M$  columns makes its rank as  $M$ , which is a full-rank  $P \times M$  matrix and hence has full diversity. This proof is valid if any other  $P - M$  columns of  $G_P$  are deleted instead of the last ones.  $\square$

We note here that if  $M$  is not a power of 2, then the quasi-orthogonal code formed above will require  $P = 2^{\lceil \log_2(M) \rceil}$  channel uses for transmission of one code block. Since  $P > M$ , the code is not delay optimal in this case.

### 3.4. Decoding

While (16) implies that performance of a ML decoder will be the same as that of ML decoding of each partition separately by assuming that only one partition is transmitted, it does not give a practical way of decoding these codes when all the partitions are indeed sent together. We provide a practical way of achieving a low complexity ML decoding done over a single partition. We will do this for  $M$  being a power of 2. If  $M$  is not a power of 2, then one can form a new channel whose rows are a power of 2 as in (28).

We note first that any row of the constructed code either contains the symbols or its conjugates (with possible

sign change). This can be seen from the iterative construction in (22) where this property is preserved. It is trivially true for  $M = 1$  in (18). For any  $\mathbf{h} \in \mathbb{C}^{M \times 1}$ , define a transformation denoted by  $\mathcal{T}$  that takes conjugates of those elements of  $M \times 1$  vector  $G_M[\mathbf{s}]\mathbf{h}$  that contains conjugates of elements of  $\mathbf{s}$ . Hence we can write

$$\mathcal{T}(G_M[\mathcal{A}_{M,i}(\mathbf{s})]\mathbf{h}) = \mathcal{E}_{M,i}(\mathbf{h})v_{M,i}(\mathbf{s}), \quad (29)$$

where  $\mathcal{E}_{M,i}$ 's are  $M \times (M/2)$  matrices dependent only on  $\mathbf{h}$ , and  $v_{M,i}$ 's are  $(M/2) \times 1$  vectors that contain symbols from partition  $i$ , with  $i = 1, 2$ .

**Proposition 7.** For any  $\mathbf{h} \in \mathbb{C}^{M \times 1}$ ,  $\mathcal{E}_{M,1}^H(\mathbf{h})\mathcal{E}_{M,2}(\mathbf{h}) = 0$ .

*Proof.* It follows from (P5) for any  $\mathbf{h}$  that

$$\mathbf{0}^M = \{G_M[\mathcal{A}_{M,1}(\mathbf{s})]\mathbf{h}\}^H \{G_M[\mathcal{A}_{M,2}(\mathbf{s})]\mathbf{h}\} + \{G_M[\mathcal{A}_{M,2}(\mathbf{s})]\mathbf{h}\}^H \{G_M[\mathcal{A}_{M,1}(\mathbf{s})]\mathbf{h}\} \quad (30a)$$

$$= (\mathcal{T}\{G_M[\mathcal{A}_{M,1}(\mathbf{s})]\mathbf{h}\})^H (\mathcal{T}\{G_M[\mathcal{A}_{M,2}(\mathbf{s})]\mathbf{h}\}) + (\mathcal{T}\{G_M[\mathcal{A}_{M,2}(\mathbf{s})]\mathbf{h}\})^H (\mathcal{T}\{G_M[\mathcal{A}_{M,1}(\mathbf{s})]\mathbf{h}\}) \quad (30b)$$

$$= v_{M,1}^H(\mathbf{s})\mathcal{E}_{M,1}^H(\mathbf{h})\mathcal{E}_{M,2}(\mathbf{h})v_{M,2}(\mathbf{s}) + v_{M,2}^H(\mathbf{s})\mathcal{E}_{M,2}^H(\mathbf{h})\mathcal{E}_{M,1}(\mathbf{h})v_{M,1}(\mathbf{s}), \quad (30c)$$

where (30a) follows from (P5), and (30b) follows by noting that taking conjugates of elements at the same indices of any vectors  $M \times 1$   $\mathbf{g}_1$  and  $\mathbf{g}_2$  leaves the product  $\mathbf{g}_1^H \mathbf{g}_2 + \mathbf{g}_2^H \mathbf{g}_1$  unchanged. Note that since the partitions are disjoint, (30c) can be true only if  $\mathcal{E}_{M,1}^H(\mathbf{h})\mathcal{E}_{M,2}(\mathbf{h}) = 0$  for any  $\mathbf{h} \in \mathbb{C}^M$ .  $\square$

By taking conjugates appropriately, we can derive a modified signal model from (3) for receive antenna  $n$  ( $n = 1, \dots, N$ ) as

$$\hat{X}_n(\mathbf{s}) = \sqrt{\frac{\rho}{M}} [\mathcal{E}_{M,1}(H_n)v_{M,1}(\mathbf{s}) + \mathcal{E}_{M,2}(H)v_{M,2}(\mathbf{s})] + \hat{V}_n, \quad (31)$$

where  $H_n$  is the  $n$ th column of  $H$  and  $\hat{X}_n$  and  $\hat{V}_n$  are derived from the  $n$ th column of  $X$  and  $V$ , respectively, by taking the conjugates of some or all their elements. Let the singular value decomposition (SVD) [31] of  $\mathcal{E}_{M,i}(H_n)$  be given by  $\mathcal{E}_{M,i}(H_n) = U_i \hat{S}_i W_i^H$ , where  $U_i$  and  $W_i$  are unitary and  $\hat{S}_i$  is an  $M \times (M/2)$  diagonal matrix. Let  $\hat{S}_i$  be an  $M \times (M/2)$  diagonal matrix whose diagonal elements are the inverse of diagonal elements of  $\hat{S}_i$  and hence

$$\hat{S}_i \hat{S}_i^H = \begin{bmatrix} \mathbf{I}^{M/2} & \mathbf{0}^{M/2} \\ \mathbf{0}^{M/2} & \mathbf{0}^{M/2} \end{bmatrix} \quad (32)$$

and  $\hat{S}_i \hat{S}_i^H S_i = S_i$ . Multiplying both sides of (31) by

$U_i \hat{S}_i W_i^H \mathcal{E}_{M,i}^H(H_n) = U_i \hat{S}_i \hat{S}_i^H U_i^H$ , we get after simplification

$$U_i \hat{S}_i \hat{S}_i^H U_i^H \hat{X}_n(\mathbf{s}) = \sqrt{\frac{\rho}{M}} \mathcal{E}_{M,i}(H_n)v_{M,i}(\mathbf{s}) + U_i \hat{S}_i \hat{S}_i^H U_i^H V_n, \quad (33)$$

where we have used (29) to cancel the contribution of the other partition. Note that using (32), it follows that  $U_i \hat{S}_i \hat{S}_i^H U_i^H V_n$  has the same statistics as  $V_n$ . Using (21), one can iteratively generate the equivalent channels for each partitions as

$$\mathcal{E}_{2M,1}(\mathbf{h}) = \begin{bmatrix} \mathcal{E}_{M,1}(\mathbf{h}_{M,1}) & \mathcal{E}_{M,2}(\mathbf{h}_{M,2}) \\ \mathcal{E}_{M,1}^*(\mathbf{h}_{M,2}) & -\mathcal{E}_{M,2}(\mathbf{h}_{M,1}) \end{bmatrix}, \quad (34)$$

$$\mathcal{E}_{2M,2}(\mathbf{h}) = \begin{bmatrix} \mathcal{E}_{M,2}(\mathbf{h}_{M,1}) & \mathcal{E}_{M,1}(\mathbf{h}_{M,2}) \\ \mathcal{E}_{M,2}^*(\mathbf{h}_{M,2}) & -\mathcal{E}_{M,1}(\mathbf{h}_{M,1}) \end{bmatrix},$$

where  $\mathbf{h}_{M,1} = \mathbf{h}(1 : M)$  and  $\mathbf{h}_{M,2} = \mathbf{h}(M+1 : 2M)$ .

## 4. NUMERICAL RESULTS

In this section, we provide the numerical results for the constructed codes. We provide both the Shannon capacity perspective of these codes along with the probability of error curves for modulated symbols.

### 4.1. Capacity of quasi-orthogonal codes

The capacity of quasi-orthogonal codes is computed by using (33) to get the equivalent channel for the  $n$ th receive antenna. One can write the overall channel matrix taken over all the receive antennas by stacking them as

$$\mathcal{H}_{M,i} = \begin{bmatrix} \mathcal{E}_{M,i}(H_1) \\ \vdots \\ \mathcal{E}_{M,i}(H_N) \end{bmatrix}, \quad (35)$$

which is an  $MN \times (M/2)$  matrix. The channel model in this case is given by

$$\mathcal{X} = \sqrt{\frac{\rho}{M}} \mathcal{H}_{M,i} + \mathcal{V}. \quad (36)$$

Note that elements of  $\mathcal{V}$  are  $\mathcal{CN}(0, 1)$ .

By using the above model, we compute the ergodic capacity of quasi-orthogonal codes and plot this along with open-loop Shannon capacity in Figure 1 for an (8, 1) system. We also plot the capacity of a rate 1/2 complex orthogonal code [7]. As shown in the figure, the proposed quasi-orthogonal codes are quite close to the Shannon capacity. Note that the Shannon capacity is achievable by an ideal rate 1 complex orthogonal code though such a code is known to exist only for  $M = 2$ . In Figure 2, we plot the capacities for an (8, 2) system. The quasi-orthogonal code is not as close to the Shannon capacity in this case though it still performs much better than the orthogonal code.

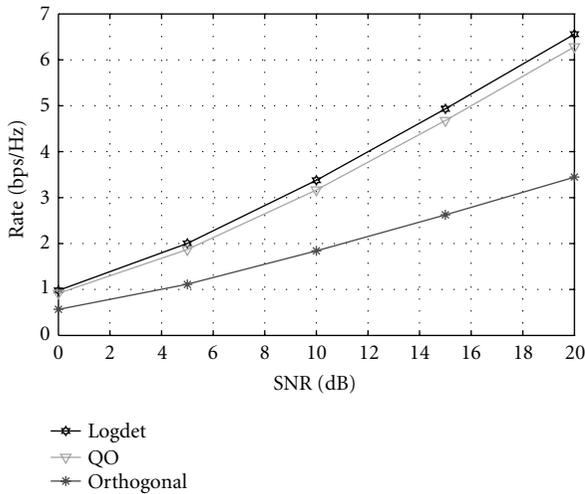


FIGURE 1: Ergodic capacity of quasi-orthogonal codes along with open loop Shannon capacity and that of a rate 1/2 orthogonal code for (8, 1).

#### 4.2. Probability of error

We plot the symbol error rate (SER) versus the average SNR per receive antenna in Figure 3 with QPSK modulation for  $M = 4, 8, 16, 32$  and  $N = 1$ . The elements of  $H$  are assumed to be i.i.d. and  $\mathcal{CN}(0, 1)$ . For  $M = 4$ , we use the rotations described in [25] that were obtained by maximizing the minimum distance of constellation  $\mathcal{C}$  defined in Proposition 5 and the precoding matrix is given by  $\text{diag}[1, \exp(0.52j)]$ . For higher  $M$ , instead of exhaustive search to find the best precoding matrix, we rotate the  $i$ th symbol,  $i = 1, \dots, M/2$ , with a phase of  $(i - 1)\pi/M$ . A better choice is also possible. Hard-decision sphere decoding was done for each partition separately by using (33). For comparison, we also plot the performance of an *ideal* full-rate orthogonal code (though unavailable) that has equivalent channel SNR as  $\|H\|_F^2 \rho/M$  and of uncoded QPSK over a channel with only additive white Gaussian noise and no fading for  $M = N = 1$ .

Note that the performance is better than that given in [18] and [19, Figure 11]. Also note that because of the orthogonality built into the proposed codes, our codes have lower decoding complexity. For a constellation of size  $q$ , the decoding complexity after the preprocessing to separate the two partitions is  $\sim q^{M/2}$  for the proposed codes, while the decoding complexity is  $\sim q^M$  for both [18, 19] under ML decoding. Under sphere decoding [32, 33], the decoding complexity is approximately cubic with the number of symbols that are jointly decoded: the decoding complexity for the proposed codes is  $2O(M^3/8)$ , and for the codes in [18, 19], the decoding complexity is  $O(M^3)$ . Hence there is a significant saving in decoding complexity while there is performance improvement by using the proposed codes.

For higher  $M$ , note that the performance of the proposed codes is very close to the ideal codes. Hence any other full-rate code will offer very marginal gains over the proposed codes for higher transmit antennas.

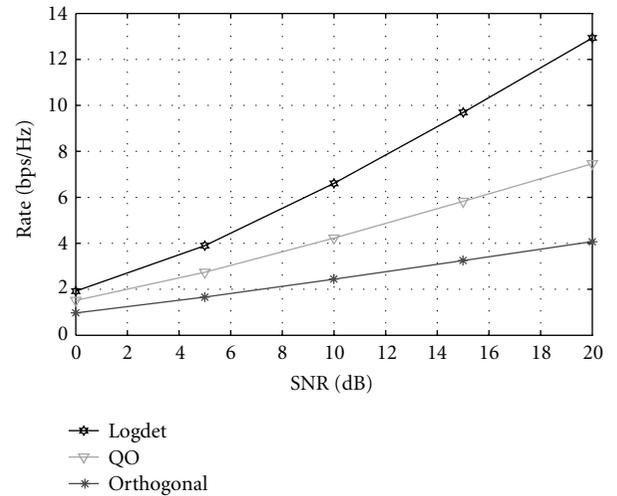


FIGURE 2: Ergodic capacity of quasi-orthogonal codes along with open loop Shannon capacity and that of a rate 1/2 orthogonal code for (8, 2).

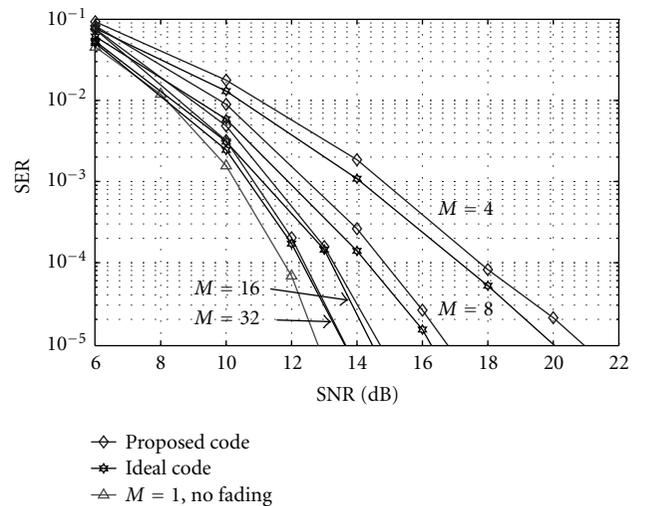


FIGURE 3: Simulated SER versus SNR for various  $M$  and  $N = 1$ , and  $M = 1$  with no fading, for QPSK modulation.

## 5. PERFORMANCE ANALYSIS FOR SELECTED CODES

For  $M = 4$ , the constructed code is the same as given in [23]. The equivalent channel model for the first partition can be written using (29) as

$$\mathcal{E}_{4,1}(\mathbf{h}) = \begin{bmatrix} h_1 & h_4 \\ h_2^* & -h_3^* \\ h_3^* & -h_2^* \\ h_4 & h_1 \end{bmatrix}. \quad (37)$$

By taking SVD of  $\mathcal{E}_{4,1}(\mathbf{h})$  and discarding the last two rows, we get a simpler  $2 \times 2$  receive signal model by discarding the null

rows as

$$\begin{aligned} r_1^i &= \sqrt{\frac{\rho}{M}} \sqrt{\frac{\gamma_i + \alpha_i}{2}} (z_1 + \exp(j\theta)z_2) + n_1^i, \\ r_2^i &= \sqrt{\frac{\rho}{M}} \sqrt{\frac{\gamma_i - \alpha_i}{2}} (z_1 - \exp(j\theta)z_2) + n_2^i, \end{aligned} \quad (38)$$

where

$$\begin{aligned} \gamma_i &= \sum_{k=1}^4 |h_{k,i}|^2, \\ \alpha_i &= 2 \operatorname{Re} (h_{1,i}^* h_{4,i} - h_{3,i}^* h_{2,i}), \end{aligned} \quad (39)$$

and  $\theta$  is the rotation applied to increase the minimum distance of constellation  $\mathcal{C} = z_1 + \exp(j\theta)z_2$  as in Proposition 5 (see also [25] for more details). The symbols  $z_1$  and  $z_2$  are the symbols in the first partition, where the indices are chosen as 1, 2 for convenience.

In addition to this code, it was shown in [25] that the rate 3/4 quasi-orthogonal code for 8 transmit antennas given in [23] has also two interfering signals and its equivalent channel model can also be written like (38) with

$$\begin{aligned} \gamma_i &= \sum_{k=1}^8 |h_{k,i}|^2, \\ \alpha_i &= 2 \operatorname{Re} (h_{1,i}^* h_{5,i} - h_{2,i} h_{6,i}^* - h_{3,i} h_{7,i}^* - h_{4,i}^* h_{8,i}). \end{aligned} \quad (40)$$

While this code does not belong to the class of proposed codes (it is not a full-rate code and the interfering symbols for the proposed code for 8 transmit antennas are 4), we include it here since its analysis is similar to the 4-transmit-antenna code.

We now determine the pairwise probability of error for these two codes by assuming that the transmitted pair  $(z_1, z_2)$  is mistaken as  $(e_1, e_2)$ . The pairwise probability of error for a given  $H$  is given by

$$P_e[(z_1, z_2) \rightarrow (e_1, e_2) | H] = Q\left(\sqrt{\frac{\rho}{4M} \mathcal{D}}\right), \quad (41)$$

where

$$\begin{aligned} \mathcal{D} &= \sum_{i=1}^N \left[ (\gamma_i + \alpha_i) |\delta_1|^2 + (\gamma_i - \alpha_i) |\delta_2|^2 \right] \\ &= \left( |\delta_1|^2 + |\delta_2|^2 \right) \sum_{i=1}^N \gamma_i + \left( |\delta_1|^2 - |\delta_2|^2 \right) \sum_{i=1}^N \alpha_i, \end{aligned} \quad (42)$$

where  $\delta_1 = ((z_1 - e_1) + j \exp(j\theta)(z_2 - e_2))$  and  $\delta_2 = ((z_1 - e_1) - j \exp(j\theta)(z_2 - e_2))$ . We now invoke the clever representation of the  $Q$ -function given in [34] to have

$$\begin{aligned} P_e[(z_1, z_2) \rightarrow (e_1, e_2) | H] \\ = \frac{1}{\pi} \int_0^{\pi/2} \exp\left(-\frac{\rho \mathcal{D}}{8M \sin^2(\theta)}\right) d\phi. \end{aligned} \quad (43)$$

We now wish to average this integral over the channel  $H$ . This may appear to be a formidable exercise, but it can be simplified easily by noting that for some constants  $a_1$  and  $a_2$  with  $a_1 > 0$  and  $(1 + a_1) > a_2$ , and for two independent real Gaussian random variables  $x_1$  and  $x_2$ , each of variance 0.5, we have

$$E_{x_1, x_2} \left\{ \exp(-a_1(x_1^2 + x_2^2) + 2a_2 x_1 x_2) \right\} = \frac{1}{\sqrt{(1 + a_1)^2 - a_2^2}}, \quad (44)$$

where  $E\{\cdot\}$  denotes the expectation. Note that the integrand in the right-hand side of (43) can be decomposed (by using expressions for  $\gamma_i$  and  $\alpha_i$ ) into  $MN/2$  terms of the form  $a_1(|h_{i,k}|^2 + |h_{i,l}|^2) + 2a_2 \operatorname{Re}(h_{i,k}^* h_{i,l})$ , that in turn can be written in two *independent* terms of the form  $a_1(x_1^2 + x_2^2) + 2a_2 x_1 x_2$ , where  $a_1 = \rho(|\delta_1|^2 + |\delta_2|^2)/[8M \sin^2(\phi)]$  and  $a_2 = \rho(|\delta_1|^2 - |\delta_2|^2)/[8M \sin^2(\phi)]$ , and  $x_1, x_2$  are real random variables with the statistics defined above. Hence, we can write (43) averaged over the channel as

$$\begin{aligned} P_e[(z_1, z_2) \rightarrow (e_1, e_2)] \\ = \frac{1}{\pi} \int_0^{\pi/2} \frac{d\phi}{\left[ \left( 1 + \frac{\rho(|\delta_1|^2 + |\delta_2|^2)}{8M \sin^2(\phi)} \right)^2 - \left( \frac{\rho(|\delta_1|^2 - |\delta_2|^2)}{8M \sin^2(\phi)} \right)^2 \right]^{MN/2}} \\ = \frac{1}{\pi} \int_0^{\pi/2} \frac{d\phi}{\left[ 1 + \frac{\rho(|\delta_1|^2 + |\delta_2|^2)}{4M \sin^2(\phi)} + \left( \frac{\rho|\delta_1 \delta_2|}{4M \sin^2(\phi)} \right)^2 \right]^{MN/2}}. \end{aligned} \quad (45)$$

This is a much simpler expression to handle being a single integral. We note that this expression holds true for both  $M = 4$  and  $M = 8$ . Note that we have thus far made no assumptions about the constellations used for  $z_1$  and  $z_2$ . We now consider the following cases.

### Suboptimal constellations

We define the chosen constellations as suboptimal if for any two distinct pairs, that is,  $(z_1, z_2) \neq (e_1, e_2)$ , we have at least one among  $\delta_1$  or  $\delta_2$  to be zero. A simple example for such a case would be for  $\theta = 0$  and  $z_1, z_2$  chosen from the same constellation that is invariant under a rotation of  $\pi$  such as QPSK, 16-QAM, and so forth. We say for the chosen pair,

$\delta_2 = 0$  and  $\delta_1 \neq 0$ ; then

$$P_e[(z_1, z_2) \rightarrow (e_1, e_2)] = \frac{1}{\pi} \int_0^{\pi/2} \frac{(4M)^{MN/2} \sin^{MN}(\phi) d\phi}{[4M \sin^2(\phi) + \rho |\delta_1|^2]^{MN/2}}$$

$$> \frac{(4M)^{MN/2} \Gamma((1+MN)/2)}{2\sqrt{\pi} \Gamma(1+MN/2) (4M + \rho |\delta_1|^2)^{MN/2}}, \quad (46)$$

where  $\Gamma(\cdot)$  denotes the Gamma function and we have used the integral that  $\int_0^{\pi/2} \sin^n(x) dx = \sqrt{\pi} \Gamma((1+n)/2) / 2\Gamma(1+n/2)$ . The diversity of this system is clearly  $MN/2$ .

### Diversity ensuring constellations

We define the chosen constellations to be diversity ensuring if for any two distinct pairs, neither  $\delta_1$  or  $\delta_2$  is zero. The design of such constellations by rotation for the considered cases can be found [25]. In this case, the pairwise probability of error is upper bounded by

$$P_e[(z_1, z_2) \rightarrow (e_1, e_2)] < \frac{1}{\pi} \int_0^{\pi/2} \frac{(4M)^{MN} \sin^{2MN}(\phi) d\phi}{(\rho |\delta_1 \delta_2|)^{MN}}$$

$$= \left( \frac{4M}{\rho |\delta_1 \delta_2|} \right)^{MN} \frac{\Gamma((1+2MN)/2)}{2\sqrt{\pi} \Gamma(1+MN)}, \quad (47)$$

where the inequality follows by taking an upper bound of the integrand in (45). This proves the full diversity of the chosen quasi-orthogonal codes for appropriately designed constellations.

## 6. CONCLUSIONS

A class of linear quasi-orthogonal codes have been constructed that offer full-rate and full diversity with constellation rotation for any transmit antennas. Due to orthogonal structure in the code, two disjoint partitions containing one half of symbols constituting the code can be decoded separately. A practical decoding algorithm is described to utilize the orthogonality. These codes are closer to the Shannon capacity curves for  $(M, 1)$  systems than to the orthogonal codes except for  $M = 2$  in which case the constructed code is the same as an orthogonal code that achieves the Shannon capacity. It may be possible to construct more classes of quasi-orthogonal codes in an iterative fashion as described in this paper.

## APPENDIX

### PROOF OF PROPOSITION 4

We first prove the following Lemma.

Lemma 1. For any  $2M \times 1$  vector  $\mathbf{x}$ ,

$$G_{2M}[\mathcal{A}_{2M,2}(\hat{\mathbf{x}})] G_{2M}[\mathcal{A}_{2M,2}(\mathbf{x}^*)] = -G_{2M}^2[\mathcal{A}_{2M,1}(\mathbf{x})]. \quad (A.1)$$

*Proof.*

$$\text{left hand side} = \begin{bmatrix} -G_M[\mathcal{A}_{M,2}(\mathbf{x}_2)] & G_M[\mathcal{A}_{M,1}(\mathbf{x}_1)] \\ -G_M[\mathcal{A}_{M,1}(\mathbf{x}_1^*)] & -G_M[\mathcal{A}_{M,2}(\mathbf{x}_2^*)] \end{bmatrix}$$

$$\times \begin{bmatrix} -G_M[\mathcal{A}_{M,2}(\mathbf{x}_2^*)] & G_M[\mathcal{A}_{M,1}(\mathbf{x}_1^*)] \\ -G_M[\mathcal{A}_{M,1}(\mathbf{x}_1)] & -G_M[\mathcal{A}_{M,2}(\mathbf{x}_2)] \end{bmatrix}$$

$$= \begin{bmatrix} G_M[\mathcal{A}_{M,1}(\mathbf{x}_1)] & G_M[\mathcal{A}_{M,2}(\mathbf{x}_2)] \\ -G_M[\mathcal{A}_{M,2}(\mathbf{x}_2^*)] & G_M[\mathcal{A}_{M,1}(\mathbf{x}_1^*)] \end{bmatrix}$$

$$\times \begin{bmatrix} -G_M[\mathcal{A}_{M,1}(\mathbf{x}_1)] & -G_M[\mathcal{A}_{M,2}(\mathbf{x}_2)] \\ G_M[\mathcal{A}_{M,2}(\mathbf{x}_2^*)] & -G_M[\mathcal{A}_{M,1}(\mathbf{x}_1^*)] \end{bmatrix}$$

$$= -G_{2M}^2[\mathcal{A}_{M,1}(\mathbf{x})], \quad (A.2)$$

where the second equality follows by interchanging the last  $M$  columns and changing the sign with the first  $M$  columns of the first matrix, and by interchanging the first  $M$  rows and changing the sign with the last  $M$  rows of the second matrix, that leaves the product unchanged.  $\square$

Now we have

$$\det \{G_{4M}[\mathcal{A}_{4M,1}(\mathbf{s})]\}$$

$$= \det \left\{ \begin{bmatrix} G_{2M}[\mathcal{A}_{2M,1}(\mathbf{s}_1)] & G_{2M}[\mathcal{A}_{2M,2}(\mathbf{s}_2)] \\ -G_{2M}[\mathcal{A}_{2M,2}(\mathbf{s}_2^*)] & G_{2M}[\mathcal{A}_{2M,1}(\mathbf{s}_1^*)] \end{bmatrix} \right\} \quad (A.3)$$

$$= \det \{G_{2M}[\mathcal{A}_{2M,1}(\mathbf{s}_1^*)]\}$$

$$\times \det \{G_{2M}[\mathcal{A}_{2M,1}(\mathbf{s}_1)]$$

$$+ G_{2M}[\mathcal{A}_{2M,2}(\mathbf{s}_2)] G_{2M}^{-1}[\mathcal{A}_{2M,1}(\mathbf{s}_1^*)] G_{2M}[\mathcal{A}_{2M,2}(\mathbf{s}_2^*)]\}$$

$$(A.4)$$

$$= \det \{G_{2M}^2[\mathcal{A}_{2M,1}(\mathbf{s}_1)]$$

$$+ G_{2M}[\mathcal{A}_{2M,2}(\mathbf{s}_2)] G_{2M}[\mathcal{A}_{2M,2}(\mathbf{s}_2^*)]\}$$

$$(A.5)$$

$$= \det \{G_{2M}^2[\mathcal{A}_{2M,1}(\mathbf{s}_1)] - G_{2M}^2[\mathcal{A}_{2M,1}(\hat{\mathbf{s}}_2)]\} \quad (A.6)$$

$$= \det \{G_{2M}[\mathcal{A}_{2M,1}(\mathbf{s}_1)] - G_{2M}[\mathcal{A}_{2M,1}(\hat{\mathbf{s}}_2)]\}$$

$$\times \det \{G_{2M}[\mathcal{A}_{2M,1}(\mathbf{s}_1)] + G_{2M}[\mathcal{A}_{2M,1}(\hat{\mathbf{s}}_2)]\} \quad (A.7)$$

$$= \text{RHS of (23)}, \quad (A.8)$$

where (A.4) follows from the relation of the determinant of a block matrix to that of its constituent matrices, (A.5) follows by applying (P5) (which is valid for different vectors since partitions are disjoint) and simplifying, (A.6) follows using (A.1), (A.7) follows by applying (P3), and (A.8) follows from linearity of the code.

## ACKNOWLEDGMENT

The authors wish to thank Dr. Bertrand M. Hochwald whose implementation of the hard-decision sphere decoding algorithm was used for simulations.

## REFERENCES

- [1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple

- antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, 1998.
- [2] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple-element antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, 1996.
- [3] E. Telatar, "Capacity of multiantenna Gaussian channels," Tech. Memo., AT&T Bell Labs, Murray Hill, NJ, USA, June 1995.
- [4] J.-C. Guey, M. P. Fitz, M. R. Bell, and W.-Y. Kuo, "Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels," in *IEEE 46th Vehicular Technology Conference*, vol. 1, pp. 136–140, Atlanta, Ga, USA, April–May 1996.
- [5] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744–765, 1998.
- [6] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, 1998.
- [7] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1456–1467, 1999.
- [8] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Transactions on Information Theory*, vol. 48, no. 7, pp. 1804–1824, 2002.
- [9] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 543–564, 2000.
- [10] A. Shokrollahi, B. Hassibi, B. M. Hochwald, and W. Sweldens, "Representation theory for high-rate multiple-antenna code design," *IEEE Transactions on Information Theory*, vol. 47, no. 6, pp. 2335–2367, 2001.
- [11] D. Agrawal, T. J. Richardson, and R. L. Urbanke, "Multiple-antenna signal constellations for fading channels," *IEEE Transactions on Information Theory*, vol. 47, no. 6, pp. 2618–2626, 2001.
- [12] C. B. Papadias, "On the spectral efficiency of space-time spreading schemes for multiple antenna CDMA systems," in *33rd Asilomar Conference on Signals, Systems, and Computers*, pp. 24–27, Pacific Grove, Calif, USA, October 1999.
- [13] O. Tirkkonen and A. Hottinen, "Square-matrix embeddable space-time block codes for complex signal constellations," *IEEE Transactions on Information Theory*, vol. 48, no. 2, pp. 384–395, 2002.
- [14] X.-B. Liang, "Orthogonal designs with maximal rates," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2468–2503, 2003.
- [15] H. Wang and X.-G. Xia, "Upper bounds of rates of complex orthogonal space-time block codes," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2788–2796, October 2003.
- [16] H. El Gamal and M. O. Damen, "Universal space-time coding," *IEEE Transactions on Information Theory*, vol. 49, no. 5, pp. 1097–1119, 2003.
- [17] X. Ma and G. B. Giannakis, "Full-diversity full-rate complex-field space-time coding," *IEEE Trans. Signal Processing*, vol. 51, no. 11, pp. 2917–2930, 2003.
- [18] M. O. Damen, K. Abed-Meraim, and J.-C. Belfiore, "Diagonal algebraic space-time block codes," *IEEE Transactions on Information Theory*, vol. 48, no. 3, pp. 628–636, 2002.
- [19] Y. Xin, Z. Wang, and G. B. Giannakis, "Space-time diversity systems based on linear constellation precoding," *IEEE Transactions on Wireless Communications*, vol. 2, no. 2, pp. 294–309, 2003.
- [20] V. M. DaSilva and E. S. Sousa, "Fading-resistant modulation using several transmitter antennas," *IEEE Trans. Communications*, vol. 45, no. 10, pp. 1236–1244, 1997.
- [21] C. B. Papadias and G. J. Foschini, "An open loop transmit diversity technique for systems employing four transmitter antennas," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP '01)*, vol. 4, pp. 2481–2484, Salt Lake City, Utah, USA, May 2001.
- [22] C. B. Papadias and G. J. Foschini, "Capacity-approaching space-time codes for systems employing four transmitter antennas," *IEEE Transactions on Information Theory*, vol. 49, no. 3, pp. 726–732, 2003.
- [23] H. Jafarkhani, "A quasi-orthogonal space-time block code," *IEEE Trans. Communications*, vol. 49, no. 1, pp. 1–4, 2001.
- [24] O. Tirkkonen, A. Boariu, and A. Hottinen, "Minimal non-orthogonality rate 1 space-time block code for 3+ Tx antennas," in *Proc. IEEE 6th International Symposium on Spread Spectrum Techniques and Applications*, vol. 2, pp. 429–432, Parsippany, NJ, USA, June–August 2000.
- [25] N. Sharma and C. B. Papadias, "Improved quasi-orthogonal codes through constellation rotation," in *IEEE Trans. Communications*, vol. 51, pp. 332–335, 2003.
- [26] O. Tirkkonen, "Optimizing space-time block codes by constellation rotations," in *Proc. 2nd Finnish Wireless Communications Workshop*, Tampere, Finland, October 2001.
- [27] G. Ganesan and P. Stoica, "Space-time block codes: a maximum SNR approach," *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1650–1656, 2001.
- [28] K. Boullé and J.-C. Belfiore, "Modulation scheme designed for the Rayleigh fading channel," in *Proc. Conference on Information Sciences and Systems*, pp. 288–293, Princeton, NJ, USA, March 1992.
- [29] J. Boutros and E. Viterbo, "Signal space diversity: a power and bandwidth-efficient diversity technique for the Rayleigh fading channel," *IEEE Transactions on Information Theory*, vol. 44, no. 4, pp. 1453–1467, 1998.
- [30] M. Greferath and E. Viterbo, "Algebraic construction of good collision resistant signal sets," in *Proc. Workshop on Coding and Cryptography*, pp. 213–224, Paris, France, January 1999.
- [31] G. H. Golub and C. F. Van Loan, *Matrix Computations*, John Hopkins University Press, Baltimore, Md, USA, 1983.
- [32] E. Viterbo and J. Boutros, "A universal lattice code decoder for fading channels," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1639–1642, 1999.
- [33] M. O. Damen, A. Chkeif, and J.-C. Belfiore, "Lattice code decoder for space-time codes," *IEEE Communications Letters*, vol. 4, no. 5, pp. 161–163, 2000.
- [34] J. W. Craig, "A new, simple and exact result for calculating the probability of error for two-dimensional signal constellations," in *Proc. IEEE Military Communications Conference*, vol. 2, pp. 571–575, McLean, Va, USA, November 1991.

**Naresh Sharma** received his B.S. and M.S. degrees from the Indian Institute of Technology (IIT) and his Ph.D. degree from the University of Maryland at College Park in April 2001 (all in electrical engineering). Since May 2000, he has been with the Communication Theory Group at the Open Innovations Laboratory, Lucent Technologies, Whippany, NJ, USA, where he has worked on third generation (3G) wireless systems with emphasis on physical and MAC layer algorithms for both single and multiantenna systems. His research interests include spread spectrum and multiantenna systems, and error-correcting coding.



Dr. Sharma is a corecipient of Bell Labs President's Gold Award for 2002 for contributions to Bell Labs layered space-time (BLAST) MIMO system, and was awarded the 1997 G. N. Revenkar Prize for the most outstanding performance in EE Graduate School at IIT. Dr. Sharma is a Member of the IEEE.

**Constantinos B. Papadias** was born in Athens, Greece, in 1969. He received his Diploma of electrical engineering from the National Technical University of Athens (NTUA) in 1991 and the Ph.D. degree in signal processing (with highest honors) from the Ecole Nationale Supérieure des Télécommunications (ENST), Paris, France, in 1995. From 1992 to 1995, he was a Teaching and Research Assistant at the Mobile Communications Department, Eurecom, France. From 1995 to 1997, he was a Postdoctoral Researcher at Stanford University's Smart Antennas Research Group. In November 1997, he joined the Wireless Research Laboratory of Bell Labs, Lucent Technologies, Holmdel, NJ, as member of technical staff. He is now Technical Manager in Global Wireless Systems Research Department, Bell Lab's, overseeing several research projects, with an emphasis on space-time and MIMO systems. He has authored several papers, patents, and standard contributions on these topics and he recently received the IEEE Signal Processing Society 2003 Young Author Best Paper Award. He is a member of the Signal Processing for Communications, Technical Committee of the IEEE Signal Processing Society, and Associate Editor for the IEEE Transactions on Signal Processing. Dr. Papadias is a Senior Member of IEEE and a Member of the Technical Chamber of Greece.



# Optimal STBC Precoding with Channel Covariance Feedback for Minimum Error Probability

**Yi Zhao**

*Department of Electrical and Computer Engineering, University of Toronto, 10 King's College Road, Toronto, ON, Canada M5S 3G4  
Email: zhaoyi@comm.utoronto.ca*

**Raviraj Adve**

*Department of Electrical and Computer Engineering, University of Toronto, 10 King's College Road, Toronto, ON, Canada M5S 3G4  
Email: rsadve@comm.utoronto.ca*

**Teng Joon Lim**

*Department of Electrical and Computer Engineering, University of Toronto, 10 King's College Road, Toronto, ON, Canada M5S 3G4  
Email: limtj@comm.utoronto.ca*

*Received 31 May 2003; Revised 15 January 2004*

This paper develops the optimal linear transformation (or precoding) of orthogonal space-time block codes (STBC) for minimizing probability of decoding error, when the channel covariance matrix is available at the transmitter. We build on recent work that stated the performance criterion without solving for the transformation. In this paper, we provide a water-filling solution for multi-input single-output (MISO) systems, and present a numerical solution for multi-input multi-output (MIMO) systems. Our results confirm that eigen-beamforming is optimal at low SNR or highly correlated channels, and full diversity is optimal at high SNR or weakly correlated channels, in terms of error probability. This conclusion is similar to one reached recently from the capacity-achieving viewpoint.

**Keywords and phrases:** MIMO, space-time block coding, beamforming, linear precoding.

## 1. INTRODUCTION

In wireless communications, the adverse effects of channel fading can be mitigated by transmission over a diversity of independent channels. A large, and growing, body of results have firmly established the potential of space-time coding [1, 2, 3] in multi-input multi-output (MIMO) systems, which use antenna arrays at the transmitter and the receiver to provide spatial diversity at both ends of a communications link.

In [3], Tarokh et al. introduced the well-known rank and determinant criteria for the design of space-time codes without channel knowledge at the transmitter. Furthermore, it was argued [2, Section II-C] that these criteria apply to both spatially independent and dependent fading channels. In other words, without channel state information (CSI) at the transmitter, space-time codes should be designed using the rank and determinant criteria, even when the spatial channels are correlated. This result was confirmed by

El Gamal [4, Proposition 7], who proved that with spatial correlation and quasistatic flat fading, full-diversity space-time codes such as orthogonal space-time block codes (OSTBC) extract the maximum diversity gain achievable, without CSI at the transmitter.

While spatial correlation does not affect diversity gain, Shiu and Foschini showed that correlation between spatial channels leads to a loss in capacity [5]. It is also known that spatial correlation results in a smaller coding advantage [2, Section II-C]. This paper explores practical approaches to recover this performance loss. However, given that nothing can improve the performance of current state-of-the-art full-diversity space-time codes without CSI at the transmitter, it is natural to consider performance improvements when this assumption is relaxed.

In this paper, we study the design of a linear precoder for OSTBC in spatially correlated, quasistatic, flat fading channels with knowledge of the channel covariance at the transmitter. The objective is to minimize the probability of

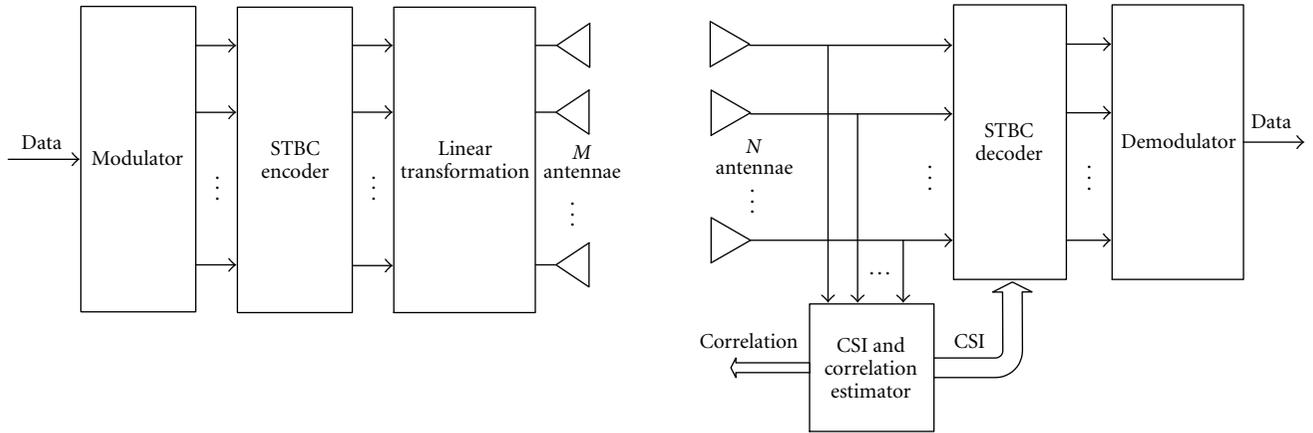


FIGURE 1: Precoded STBC transmitter and receiver block diagrams.

decoding error. The channel covariance information may be fed back from the receiver. Such a system may be considered more practical than the case when true CSI is available at the transmitter, because in that case the feedback channel may be too heavy an overhead on the communication system. Prior work done on this topic developed the optimality criterion [6] to be satisfied by the precoding matrix, but no closed-form or numerical solution was provided. In this paper, a numerical solution is provided for MIMO systems with an arbitrary number of transmit and receive antennae. Furthermore, we derive an exact water-filling solution for MISO systems. Assuming uncorrelated fading at the receiver as in [7], we show that this solution is exact in MIMO systems as well.

This problem setting ties in with recent work on determining the capacity-achieving signal correlation matrix when the channel covariance matrix is available at the transmitter [7, 8, 9]. In contrast, our research is focused on minimizing the error probability, given a linear precoding structure based on orthogonal STBC. Because of the orthogonal structure of the code matrices used, this transmitter has complexity only linear in the number of transmit antennas despite the use of a maximum-likelihood (ML) receiver [10].

The rest of the paper is organized as follows. Section 2 presents the background material needed in the rest of the paper, Section 3 discusses the optimal precoding under various scenarios, while Section 4 introduces three simplified strategies that are shown to result in minimal performance loss. Simulation examples are presented in Section 5. Finally, Section 6 presents conclusions.

## 2. BACKGROUND

Consider a MIMO system with  $M$  transmit and  $N$  receive antennae. OSTBC is used, and a linear transformation unit is applied prior to transmission to take account of the channel covariance information. The transformation matrix  $\mathbf{W} \in \mathbb{C}^{M \times M}$  is to be determined to minimize the maximum pair-

wise error probability (PEP) between codewords in correlated fading. An ML receiver is used. Illustrations of the transmitter and receiver for such a system are shown in Figure 1.

The MIMO channel between the transmitter and the receiver, assumed flat and Rayleigh, is described by the  $N \times M$  matrix

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1M} \\ h_{21} & h_{22} & \cdots & h_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N1} & h_{N2} & \cdots & h_{NM} \end{bmatrix}, \quad (1)$$

where the element  $h_{nm}$  is the fading coefficient between the  $m$ th transmit antenna and the  $n$ th receive antenna. The channel correlation matrix is

$$\begin{aligned} \mathbf{R} &= E[\mathbf{h}\mathbf{h}^\dagger], \\ \mathbf{h} &= \text{vec}(\mathbf{H}), \end{aligned} \quad (2)$$

where  $(\cdot)^\dagger$  denotes Hermitian transpose, and  $\text{vec}(\cdot)$  denotes the vectorization operator which stacks the columns of  $\mathbf{H}$ . Note that this definition is identical to the one in [3].

The STBC encoder organizes data into an  $M \times L$  matrix  $\mathbf{C}$  and successive columns of this matrix are transmitted over  $L$  time indices. The corresponding  $N \times L$  received signal matrix  $\mathbf{X}$  can be written as

$$\mathbf{X} = \mathbf{H}\mathbf{W}\mathbf{C} + \mathbf{E}, \quad (3)$$

where  $\mathbf{E}$  is an  $N \times L$  matrix with i.i.d. complex Gaussian elements representing additive thermal noise. The receiver employs an ML decoder, thus the decoded codeword  $\hat{\mathbf{C}}$  can be expressed as

$$\hat{\mathbf{C}} = \arg \min_{\mathbf{C}} \|\mathbf{X} - \mathbf{H}\mathbf{W}\mathbf{C}\|_F^2, \quad (4)$$

where  $\|\cdot\|_F$  is the Frobenius norm [11]. Note that, because **HW** is equivalent to a modified channel matrix  $\tilde{\mathbf{H}}$ , ML decoding of **C** requires only the simple linear operation described in [10].

It is known that the exact probability of error is hard to compute, so in much of the literature (see e.g. [6]), we work with the maximum PEP, which is the dominating term of the probability of error, and try to minimize a bound on it. This approach was taken in [6] and the result is that the tight upper bound on the Gaussian tail for the maximum PEP is minimized by a transformation matrix **W** that satisfies

$$\begin{aligned} \mathbf{Z}_{\text{opt}} &= \mathbf{W}_{\text{opt}} \mathbf{W}_{\text{opt}}^\dagger \\ &= \arg \max_{\substack{\mathbf{Z} \geq 0, \\ \text{tr}(\mathbf{Z})=M}} \det [(\mathbf{I}_N \otimes \mathbf{Z})\boldsymbol{\eta} + \mathbf{R}^{-1}], \end{aligned} \quad (5)$$

where **Z** has to be positive semidefinite because  $\mathbf{Z} = \mathbf{W}\mathbf{W}^\dagger$ , and the trace constraint is necessary to avoid power amplification.  $\otimes$  denotes the Kronecker product, while  $\boldsymbol{\eta} = \mu_{\min}/4\sigma^2$  with

$$\mu_{\min} = \arg \min_{\mu_{kl}} \left\{ \mu_{kl} \mathbf{I} = (\mathbf{C}_k - \mathbf{C}_l)(\mathbf{C}_k - \mathbf{C}_l)^\dagger \right\}, \quad (6)$$

among all possible codewords. In this paper, we follow this approach as well and solve the optimization problem defined in (5).

### 3. OPTIMAL TRANSFORMATION

#### 3.1. General solution

To solve the optimization problem (5), we begin by introducing a reasonable assumption of the channel correlation: the correlation between two subchannels is equal to the product of the correlation at the transmitter and that at the receiver [12]. In matrix form, letting  $\mathbf{R}_T = (1/M)E\{\mathbf{H}^H\mathbf{H}\}$  denote the correlation between different transmit antennae, and  $\mathbf{R}_R = (1/N)E\{\mathbf{H}\mathbf{H}^H\}$  the correlation between receive antennae, the channel correlation is

$$\mathbf{R} = \mathbf{R}_R \otimes \mathbf{R}_T. \quad (7)$$

It has been shown that the validity of this assumption is supported by measurement results for mobile links [12]. With this assumption, the optimal **Z** matrix is

$$\mathbf{Z}_{\text{opt}} = \arg \max_{\substack{\mathbf{Z} = \mathbf{Z}^* \geq 0, \\ \text{tr}(\mathbf{Z})=M}} \det [(\mathbf{I}_N \otimes \mathbf{Z})\boldsymbol{\eta} + \mathbf{R}_R^{-1} \otimes \mathbf{R}_T^{-1}], \quad (8)$$

since  $\mathbf{R}^{-1} = \mathbf{R}_R^{-1} \otimes \mathbf{R}_T^{-1}$  [13].

The problem is to choose a positive semidefinite matrix **Z** to maximize  $\det[(\mathbf{I}_N \otimes \mathbf{Z})\boldsymbol{\eta} + \mathbf{R}_R^{-1} \otimes \mathbf{R}_T^{-1}]$  subject to the trace constraint  $\text{tr}(\mathbf{Z}) = M$ . Notice that the correlation matrices  $\mathbf{R}_R$  and  $\mathbf{R}_T$  are both positive semidefinite and we can decompose them into

$$\begin{aligned} \mathbf{R}_T &= \mathbf{U}_T \boldsymbol{\Lambda}_T \mathbf{U}_T^\dagger, & \text{where } \mathbf{U}_T \mathbf{U}_T^\dagger &= \mathbf{I}_M, \\ \mathbf{R}_R &= \mathbf{U}_R \boldsymbol{\Lambda}_R \mathbf{U}_R^\dagger, & \text{where } \mathbf{U}_R \mathbf{U}_R^\dagger &= \mathbf{I}_N, \end{aligned} \quad (9)$$

then

$$\begin{aligned} &\det [(\mathbf{I} \otimes \mathbf{Z})\boldsymbol{\eta} + \mathbf{R}_R^{-1} \otimes \mathbf{R}_T^{-1}] \\ &= \det [(\mathbf{I} \otimes \mathbf{Z})\boldsymbol{\eta} + (\mathbf{U}_R \boldsymbol{\Lambda}_R^{-1} \mathbf{U}_R^\dagger) \otimes (\mathbf{U}_T \boldsymbol{\Lambda}_T^{-1} \mathbf{U}_T^\dagger)] \\ &= \det [(\mathbf{I} \otimes \mathbf{Z})\boldsymbol{\eta} + (\mathbf{U}_R \otimes \mathbf{U}_T)(\boldsymbol{\Lambda}_R^{-1} \otimes \boldsymbol{\Lambda}_T^{-1})(\mathbf{U}_R \otimes \mathbf{U}_T)^\dagger] \\ &= \det [(\mathbf{U}_R \otimes \mathbf{U}_T) [(\mathbf{U}_R \otimes \mathbf{U}_T)^\dagger (\mathbf{I} \otimes \mathbf{Z}) (\mathbf{U}_R \otimes \mathbf{U}_T) \\ &\quad + \boldsymbol{\Lambda}_R^{-1} \otimes \boldsymbol{\Lambda}_T^{-1}] (\mathbf{U}_R \otimes \mathbf{U}_T)^\dagger] \\ &= \det [\mathbf{U}_R \otimes \mathbf{U}_T] \det [(\mathbf{U}_R^\dagger \mathbf{I} \mathbf{U}_R) \otimes (\mathbf{U}_T^\dagger \mathbf{Z} \boldsymbol{\eta} \mathbf{U}_T) \\ &\quad + \boldsymbol{\Lambda}_R^{-1} \otimes \boldsymbol{\Lambda}_T^{-1}] \det [\mathbf{U}_R^{-1} \otimes \mathbf{U}_T^{-1}] \\ &= \det [\mathbf{I}_N \otimes \mathbf{B} + \boldsymbol{\Lambda}_R^{-1} \otimes \boldsymbol{\Lambda}_T^{-1}], \end{aligned} \quad (10)$$

where  $\mathbf{B} = \mathbf{U}_T^\dagger \mathbf{Z} \boldsymbol{\eta} \mathbf{U}_T$ . The intermediate steps above come from the fact that [13]

$$\left( \prod_{i=1}^N \mathbf{A}_i \right) \otimes \left( \prod_{i=1}^N \mathbf{B}_i \right) = \prod_{i=1}^N \mathbf{A}_i \otimes \mathbf{B}_i, \quad (11)$$

and  $\det[\mathbf{U}_R \otimes \mathbf{U}_T] = 1$ . The trace constraint becomes

$$\begin{aligned} \text{tr}(\mathbf{B}) &= \text{tr}(\mathbf{U}_T^\dagger \mathbf{Z} \boldsymbol{\eta} \mathbf{U}_T) \\ &= \text{tr}(\mathbf{U}_T \mathbf{U}_T^\dagger \mathbf{Z} \boldsymbol{\eta}) \\ &= \text{tr}(\mathbf{Z} \boldsymbol{\eta}) = \eta M, \end{aligned} \quad (12)$$

since  $\text{tr}(\mathbf{A}\mathbf{B}) = \text{tr}(\mathbf{B}\mathbf{A})$ .

The problem therefore reduces to finding a positive semidefinite matrix

$$\mathbf{B}_{\text{opt}} = \arg \max_{\substack{\mathbf{B} \geq 0, \\ \text{tr}(\mathbf{B})=\eta M}} \det [(\mathbf{I}_N \otimes \mathbf{B}) + \boldsymbol{\Lambda}_R^{-1} \otimes \boldsymbol{\Lambda}_T^{-1}]. \quad (13)$$

Since  $\boldsymbol{\Lambda}_T^{-1}$  and  $\boldsymbol{\Lambda}_R^{-1}$  are both diagonal, **B** must be also diagonal [14]. Let the *i*th diagonal element of  $\boldsymbol{\Lambda}_T$  and **B**, and the *j*th diagonal elements of  $\boldsymbol{\Lambda}_R$  be  $\lambda_{Ti}$ ,  $b_i$ , and  $\lambda_{Rj}$ , respectively. The problem (8) becomes finding a set of nonnegative  $b_i$ 's to maximize

$$\prod_{i=1}^M \prod_{j=1}^N (b_i + \lambda_{Rj}^{-1} \lambda_{Ti}^{-1}) \quad (14)$$

under the trace constraint  $\text{tr}(\mathbf{B}) = \sum_i b_i = \eta M$ . This problem is an extension of the water-filling problem to two parameters (*i* and *j*), so we can view it as a generalized water-filling problem. The closed form solution to this problem is unknown. However, we can find the solution by numerical methods such as sequential quadratic programming (SQP) [15]. Results of the numerical scheme are provided in Section 5.2.

Since  $\mathbf{Z} \boldsymbol{\eta} = \mathbf{U}_T \mathbf{B} \mathbf{U}_T^\dagger$ , the diagonal matrix **B** is actually the eigenvalue matrix of  $\mathbf{Z} \boldsymbol{\eta}$ . Thus **Z** and **W** can be derived from **B** as follows:

$$\begin{aligned} \mathbf{Z} &= \frac{1}{\eta} \mathbf{U}_T \mathbf{B} \mathbf{U}_T^\dagger, \\ \mathbf{W} &= \frac{1}{\sqrt{\eta}} \mathbf{U}_T \sqrt{\mathbf{B}} \boldsymbol{\Phi}, \end{aligned} \quad (15)$$

where  $\Phi$  can be any  $M \times M$  unitary matrix, so  $\mathbf{W}_{\text{opt}}$  is not unique. For simplicity we choose the identity matrix in this paper, that is,  $\Phi = \mathbf{I}_M$ .

### 3.2. Water-filling solution for MISO systems

We now consider the special case of a multi-input single-output (MISO) system, that is, a system with only a single receive antenna ( $N=1$ ). This is a reasonable model for the downlink of mobile communication systems since it may be impractical to employ more than one antenna at the mobile terminal. Under this assumption, the Kronecker product in (13) disappears and we need to solve

$$\mathbf{B}_{\text{opt}} = \arg \max_{\substack{\mathbf{B} \geq 0 \\ \text{tr}(\mathbf{B}) = \eta M}} \det[\mathbf{B} + \Lambda_T^{-1}], \quad (16)$$

where  $\mathbf{B}$  is still a positive semidefinite diagonal matrix. This is identical to the water-filling problem in information theory [14], which has the solution

$$b_i = \max(\nu - \lambda_{T_i}^{-1}, 0), \quad \text{for } i = 1, \dots, M, \quad (17)$$

where  $\nu$  is a constant chosen to satisfy the trace constraint and  $\mathbf{B} = \text{diag}(b_1, \dots, b_M)$ . The optimal transformation matrix is

$$\mathbf{W}_{\text{opt}} = \frac{1}{\sqrt{\eta}} \mathbf{U}_T \sqrt{\mathbf{B}_{\text{opt}}}. \quad (18)$$

With  $\mathbf{W}_{\text{opt}}$  given by (18), the transmitted signal is

$$\begin{aligned} \mathbf{W}\mathbf{x} &= (\mathbf{u}_{t1}, \dots, \mathbf{u}_{tM}) \begin{bmatrix} \sqrt{\frac{b_1}{\eta}} & & \\ & \ddots & \\ & & \sqrt{\frac{b_M}{\eta}} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_M \end{bmatrix} \\ &= \sum_{i=1}^M \mathbf{u}_{ti} \sqrt{\frac{b_i}{\eta}} x_i, \end{aligned} \quad (19)$$

and thus occupies the subspace spanned by the subset of eigenvectors of  $\mathbf{R}_T$  corresponding to nonzero  $b_i$  in (17). Notice that

$$\begin{aligned} \text{rank}(\mathbf{W}\mathbf{C}_k - \mathbf{W}\mathbf{C}_l) \\ = \text{rank}[\mathbf{U}_T \mathbf{B} (\mathbf{C}_k - \mathbf{C}_l)] = \text{rank}(\mathbf{B}), \end{aligned} \quad (20)$$

since both  $\mathbf{U}_T$  and  $(\mathbf{C}_k - \mathbf{C}_l)$  are full rank. Therefore, the dimension of this subspace is equal to the transmit diversity order, as defined in [2].

In the case of very high correlation, only one  $b_i$ —the one corresponding to the principal eigenvector—is nonzero, and we have eigen-beamforming. On the other hand, all the eigenvectors are used when the correlation is low, and we have full diversity. In the uncorrelated channel where  $\mathbf{R} = \mathbf{I}$ , it can easily be shown that  $\mathbf{W} = \mathbf{I}$ , meaning that OSTBC is already optimal, as expected. In between beamforming and full diversity, the water-filling scheme determines the number of active eigenchannels, and distributes the power over them with more power devoted to the stronger ones. In this

transition region, the optimal scheme may be considered to have a partial diversity order. In all cases, the diversity order is equal to the number of nonzero  $b_i$ 's.

### 3.3. Relation to capacity analyses

There has been much interest in the information theory community in MIMO channels with covariance feedback [7, 8, 9]. In those works the goal is to find the input covariance matrix  $\mathbf{S}_{x,\text{opt}}$  necessary to achieve ergodic channel capacity, while in contrast our goal is to find the optimal linear transformation to achieve minimum error probability. Interestingly, the conclusions reached are strikingly similar for both approaches, and warrant some comment.

- (1) *Transmitting over the eigenvectors of the transmit correlation matrix is optimal* assuming only the channel correlation is available at the transmitter. The two schemes both result in allocating transmission power over the eigenvectors of the transmit correlation matrix. The strategy is similar: the stronger eigen-channel gets more power. However, the exact amount allocated to each eigen-channel may differ for the two schemes since different optimization criteria are applied.
- (2) *Beamforming is optimal at high correlation/low SNR*. When the channels are highly correlated, both minimizing error probability and maximizing capacity require transmission over the strongest eigen-channel only. This statement is also true for the low SNR region where the errors are caused mainly by Gaussian noise. Thus focusing all the energy into one particular direction results in maximizing the received SNR. Diversity is not helpful as it is noise, and not fading, that limits performance.
- (3) *Optimal diversity order increases with SNR*. At low SNR, only the strongest eigen-channel is used. As the SNR increases, more eigenchannels come into use, so the diversity order increases until full-diversity order is achieved. However, the SNR points where the diversity order changes may not be the same for the two schemes.
- (4) *Full diversity is optimal in uncorrelated channels*. For the extreme case of an uncorrelated channel, no transformation of STBC is required to minimize error rate, while uncorrelated transmit signals maximize bit rate. Similarly, in the high-SNR region, the optimal scheme should use all the eigen-channels because in this case diversity can be taken advantage of.

Besides these similarities, the transmitter structures of the two schemes are very similar. The channel signals (STBC codewords in our scheme or randomly coded Gaussian signals in capacity-achieving scheme) are first modulated on the eigenvectors of the transmit correlation matrix. Then these vectors are transmitted with different powers, determined by the eigenvalues of the channel correlation matrix. These two steps can be implemented with a linear transformation unit. Therefore, if we replace the STBC encoder with a random encoder and Gaussian signal modulator, the linear transformation structure becomes a capacity-achieving one.

#### 4. SIMPLIFIED SCHEMES

From Section 3 we know that the optimal transformation scheme is not simple to determine. For the general MIMO systems, the computation of the transformation matrix involves complex numerical algorithms. Even for the simpler case of MISO systems, the water-filling solution still requires an iterative process. In this section, we introduce several simplified schemes to reduce the complexity. Simulation results in Section 5 will show that these schemes can achieve performance very similar to the optimal one with much lower complexity.

##### 4.1. Ignoring the receive correlation

Due to differences in their physical surroundings, the transmitter and receiver on the downlink of a mobile network have different correlation properties. The extended “one-ring” model introduced in [5] is a well-known scattering model for channel correlation. If we use this model to simulate the downlink of a mobile connection, the correlation of the fading coefficients between transmit antennae  $p$  and  $q$  and receive antenna  $m$  is

$$[\mathbf{R}_T]_{p,q} = E[h_{mp}h_{mq}^*] \approx J_0\left(\Delta \frac{2\pi}{\lambda} d_T(p, q)\right), \quad (21)$$

where  $\Delta$  is the angle spread, which is defined as the ratio of the radius of the scatterer ring around the receiver and the line-of-sight distance between the transmitter and the receiver,  $\lambda$  is the wavelength,  $d_T(p, q)$  is the distance between the two transmit antennae, and  $J_0(\cdot)$  is the zeroth-order Bessel function of the first kind. The correlation between two receive antennae  $l$  and  $m$  is

$$[\mathbf{R}_R]_{l,m} = E[h_{lp}h_{mp}^*] = J_0\left(\frac{2\pi}{\lambda} d_R(l, m)\right), \quad (22)$$

where  $d_R(l, m)$  is the distance between the two receive antennae.

In practice, the angle spread  $\Delta$  is usually small. As a result, from (21) and (22) we see that the receive correlation is usually small compared to transmit correlation. For instance, if the distance between two transmit antennae equals  $\lambda/2$  and  $\Delta = 0.1$ , the correlation between these two transmit antennae is  $J_0(0.1\pi) = 0.97$ . But the correlation between two receive antennae with the same separation is just  $J_0(\pi) = -0.30$ .

In dealing with receive diversity, a correlation below 0.5 is considered negligible [16]. Therefore we can simplify our algorithm by ignoring the receive correlation. Under this approximation, the rows of  $\mathbf{H}$  become independent and the channel correlation matrix can be written as  $\mathbf{R} = \mathbf{I}_N \otimes \mathbf{R}_T$ . In this case, (13) becomes

$$\begin{aligned} \mathbf{B}_{\text{opt}} &= \arg \max_{\substack{\mathbf{B} \succeq 0 \\ \text{tr}(\mathbf{B}) = \eta M}} \det[\mathbf{I}_N \otimes \mathbf{B} + \mathbf{I}_N \otimes \mathbf{\Lambda}_T^{-1}] \\ &= \arg \max_{\substack{\mathbf{B} \succeq 0 \\ \text{tr}(\mathbf{B}) = \eta M}} \det[\mathbf{B} + \mathbf{\Lambda}_T^{-1}]^N. \end{aligned} \quad (23)$$

Therefore, the solution is exactly the same as in (17), and generalized water filling is avoided.

##### 4.2. Switching between beamforming and STBC

The water-filling scheme in Section 3.2 changes from beamforming to full diversity as a function of SNR. In the transition region, the diversity order is determined by the number of the active eigenchannels, and the optimal power allocation is determined by water filling. This iterative process must be recalculated for each SNR. We can introduce a simplifying scheme to avoid water filling altogether by switching between beamforming ( $\mathbf{W}$  is rank one) and O-STBC ( $\mathbf{W} = \mathbf{I}$ ) at a pre-computed threshold SNR level. This threshold is found by equating the error probability performance with beamforming and O-STBC. In particular, for a MISO system, we want to find the  $\eta$  that solves the equation

$$\det[\mathbf{Z}_{\text{beam}}\eta + \mathbf{R}_T^{-1}] = \det[\eta\mathbf{I}_M + \mathbf{R}_T^{-1}], \quad (24)$$

where  $\mathbf{Z}_{\text{beam}}$  is the  $\mathbf{Z}$  matrix for beamforming, that is,

$$\mathbf{Z}_{\text{beam}} = \frac{1}{\eta} \mathbf{U}_T \text{diag}[M\eta, 0, \dots, 0] \mathbf{U}_T^\dagger = M \mathbf{u}_{t1} \mathbf{u}_{t1}^\dagger, \quad (25)$$

where  $\mathbf{u}_{t1}$  is the eigenvector corresponding to the largest eigenvalue of  $\mathbf{U}_T$ . With the solution of  $\eta$ , the SNR threshold can be set as

$$\text{SNR}_{\text{th}} = \frac{4\eta}{\mu_{\min}}. \quad (26)$$

It is self-evident that the simplified strategy incurs a greater loss in performance relative to the full-complexity scheme when the transition region between beamforming and O-STBC grows. There are however cases when the transition region is so small that no difference in performance is discernible.

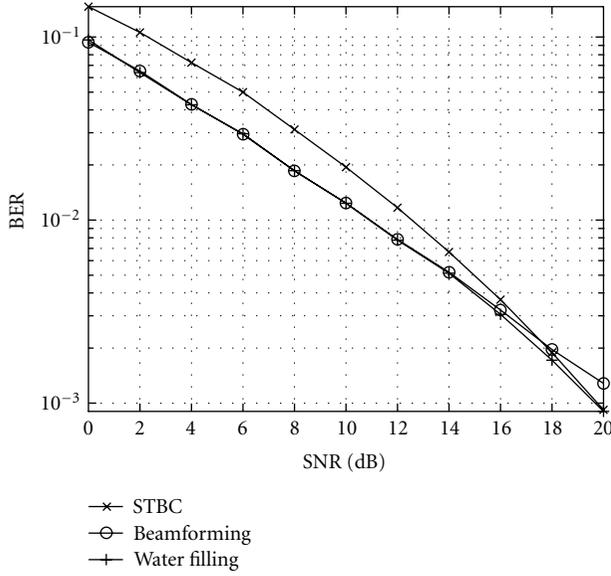
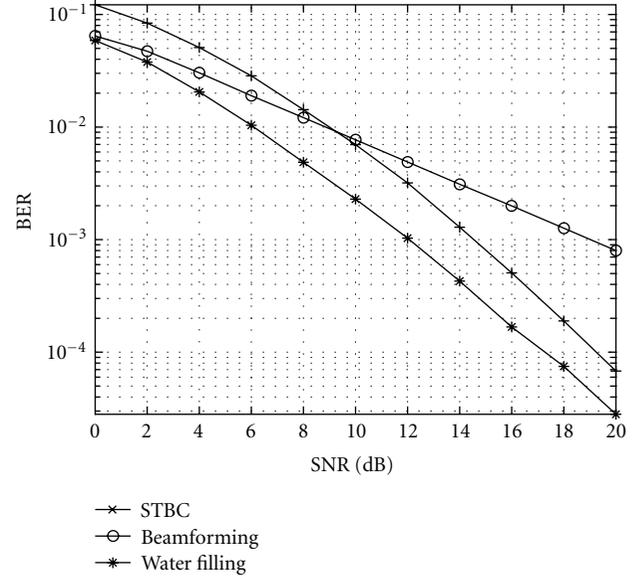
One example is when the correlation between antennae is low. In this case all the eigenvalues are close to 1, so the transition region is small. Another example is when the channel correlations are equal, in which case the eigenvalues of  $\mathbf{R}_T$  take on only two values so that the transition region has zero width. To show this, consider

$$\mathbf{R}_T = \begin{bmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{bmatrix}. \quad (27)$$

This matrix has only two eigenvalues:  $(1 + \rho)$  and  $(1 - \rho)$  (repeated  $(M - 1)$  times). As a result, the water-filling scheme has no transition region. In the low SNR region, only the eigen-channel corresponding to eigenvalue  $(1 + \rho)$  is used, so we have beamforming. All the other  $M - 1$  channels will come into use together when the SNR exceeds the threshold level, so the performance is quite close to STBC. Therefore, the switching scheme can achieve very good performance under this correlation model.

Although the switching scheme is designed for MISO systems to simplify the water-filling process, it can be easily extended to MIMO systems by changing (24) into

$$\begin{aligned} \det[\eta\mathbf{I}_N \otimes \mathbf{Z}_{\text{beam}} + \mathbf{R}_R^{-1} \otimes \mathbf{R}_T^{-1}] \\ = \det[\eta\mathbf{I}_N \otimes \mathbf{I}_M + \mathbf{R}_R^{-1} \otimes \mathbf{R}_T^{-1}]. \end{aligned} \quad (28)$$

FIGURE 2: Water filling with  $M = 2$ ,  $N = 1$ , and BPSK modulation.FIGURE 3: Water filling with  $M = 4$ ,  $N = 1$ , and BPSK modulation.

### 4.3. Equal power allocation (EPA) scheme

The switching scheme cannot guarantee good performance for arbitrary channel correlation since it only provides a diversity order of 1 or  $M$  whereas the optimal scheme may require partial diversity order. As an alternative to the switching scheme, we propose the equal power allocation (EPA) scheme. It automatically chooses the optimal diversity order, and assigns equal power to each active eigen-channel and so numerical water filling is avoided.

Similar to the switching scheme, the first step of EPA is to set SNR thresholds at the points where diversity order changes. These  $M - 1$  thresholds can be found by solving equations similar to (24). The  $i$ th threshold is obtained by solving

$$\begin{aligned} \det[\eta \mathbf{I}_N \otimes \mathbf{Z}_i + \mathbf{R}_R^{-1} \otimes \mathbf{R}_T^{-1}] \\ = \det[\eta \mathbf{I}_N \otimes \mathbf{Z}_{i+1} + \mathbf{R}_R^{-1} \otimes \mathbf{R}_T^{-1}], \end{aligned} \quad (29)$$

where  $\mathbf{Z}_i$  denotes the  $\mathbf{Z}$  matrix corresponding to EPA over the  $i$  strongest eigenchannels, or

$$\mathbf{Z}_i = \frac{M\eta}{i} \mathbf{U}_T \begin{bmatrix} \mathbf{I}_i & \mathbf{0}_{i \times (M-i)} \\ \mathbf{0}_{(M-i) \times i} & \mathbf{0}_{(M-i) \times (M-i)} \end{bmatrix} \mathbf{U}_T^\dagger. \quad (30)$$

The SNR axis is then divided to  $M$  regions, each corresponding to a diversity order. The transmitter can check those thresholds to determine which region the true SNR belongs to. The corresponding diversity order for transmission is used. To reduce the complexity, instead of going through the water-filling process to compute the power distribution, the transmitter now allocates power equally among all the active eigenchannels. We can expect this scheme to have better performance than the switching scheme in Section 4.2, but the complexity is also higher.

## 5. SIMULATION RESULTS

### 5.1. MISO channels

This section examines the performance of the water-filling scheme derived in Section 3.2. Figure 2 shows the performance of the proposed algorithm, O-STBC, and eigenbeamforming when there are two transmit and one receive antennae. The modulation scheme is BPSK and the vertical axis plots the bit error probability (BEP). SNR is defined as the ratio of the transmitted bit energy to power spectral density (i.e.  $E_b/N_0$  at the transmitter). Figure 3 is for the case of four transmit antennae.

For the two simulation examples below, the transmit correlation matrices are chosen to be

$$\mathbf{R}_{T2} = \begin{bmatrix} 1 & 0.9755 \\ 0.9755 & 1 \end{bmatrix}, \quad (31)$$

$$\mathbf{R}_{T4} = \begin{bmatrix} 1 & 0.9755 & 0.9037 & 0.79 \\ 0.9755 & 1 & 0.9755 & 0.9037 \\ 0.9037 & 0.9755 & 1 & 0.9755 \\ 0.79 & 0.9037 & 0.9755 & 1 \end{bmatrix}, \quad (32)$$

respectively. They are obtained by using (21) from the extended ‘‘one-ring’’ model. The distance between two adjacent antennae is  $\lambda/2$ , and the angle spread is  $\Delta = 0.1$  radian.

From the plots, we can see that for very low SNR, the optimal transformation is equivalent to beamforming, as expected. For the other SNR regions, the performance of the optimal scheme is better than both beamforming and STBC. Furthermore, the optimal scheme approaches STBC as SNR increases, again as expected.

Figure 4 shows the performance of the optimal scheme with two transmitters when the channel correlation varies from 0 to 1. The SNR value is fixed at 5 dB. From this plot we

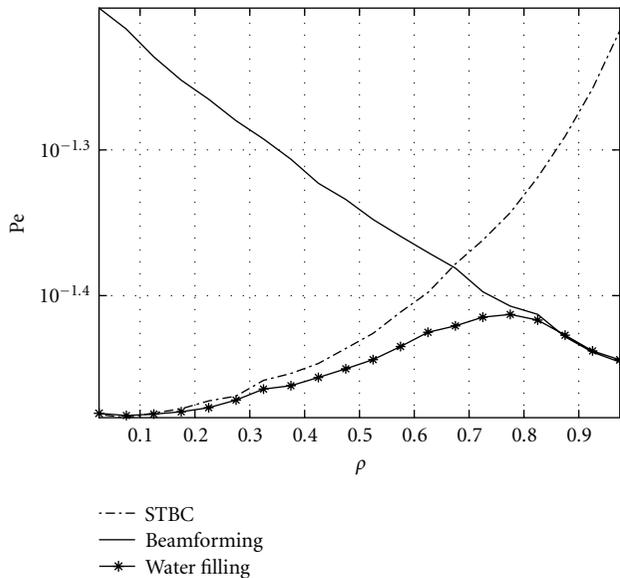


FIGURE 4: BEP versus channel correlation  $\rho$ .  $M = 2, N = 1, \text{SNR} = 5 \text{ dB}$ . Performance of three schemes.

can see that when the correlation coefficient is low ( $\rho < 0.3$ ), the performance of the optimal scheme is a little better than STBC; while with high correlation ( $\rho > 0.8$ ), the optimal scheme is the same as beamforming. In between, a relatively large performance improvement can be achieved by using the optimal scheme. This plot is remarkably similar to the corresponding plot in [17] which deals with a capacity analysis.

**5.2. Numerical solutions for MIMO systems**

As discussed in Section 3.1, the optimal transformation for MIMO system is found through a generalized water-filling problem. No closed-form solution has been found, but numerical methods, such as SQP, can be used to solve (14) with a trace constraint. Here we use the MATLAB function `fmincon` to solve the problem.

Figures 5 and 6 show the performance curves obtained with the optimal transformation. In both cases the receive correlation is set to be

$$\mathbf{R}_{R2} = \begin{bmatrix} 1 & -0.3042 \\ -0.3042 & 1 \end{bmatrix}, \quad (33)$$

which is based on (22), and  $\mathbf{R}_T$  is the same as in MISO cases. It is clear that the same conclusions about the optimality of water filling versus beamforming and O-STBC mentioned in the last section apply in this scenario as well.

**5.3. Simplified schemes**

Figure 7 shows the performance when we ignore receiver correlation. A system with four transmit and two receive antennae is considered. The transmit correlation is given in (32), and at the receiver side, the correlation between the two antennae is set to be a very high value of 0.7. From the figure we

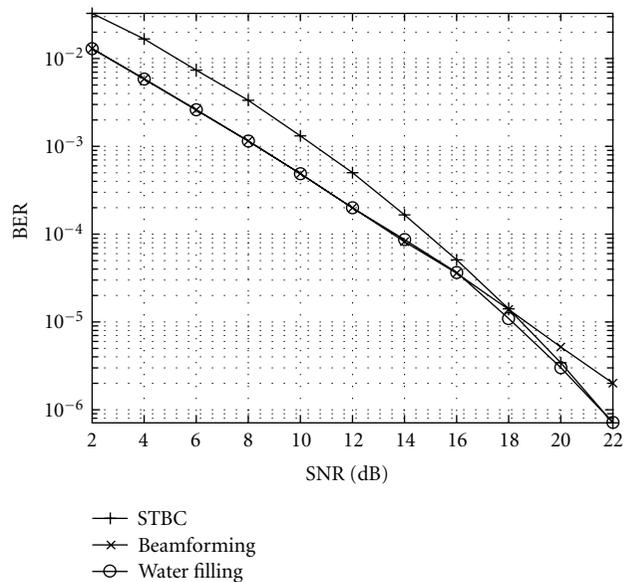


FIGURE 5: Optimal scheme for MIMO system.  $M = 2, N = 2$ .

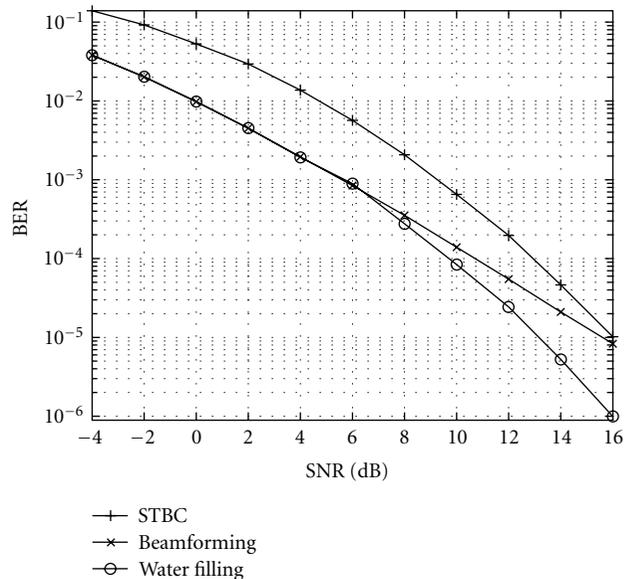


FIGURE 6: Optimal scheme for MIMO system.  $M = 4, N = 2$ .

can find that there is nearly no performance loss when ignoring the receive correlation, even when the correlation is quite large.

Figure 8 shows the performance of the simplified switching scheme compared to the water-filling scheme for MISO systems with two or four transmit antennae. The transmit correlation uses the “all-equal” model and the correlation is set as  $\rho = 0.8$ . For  $M = 4$ , the SNR threshold was found to be 4 dB; for  $M = 2$ , it was 6.5 dB. As analyzed in Section 4.2, the switching scheme achieves the same performance as water-filling in the low SNR region; in high SNR region, it should come very close to water filling. A relatively larger loss

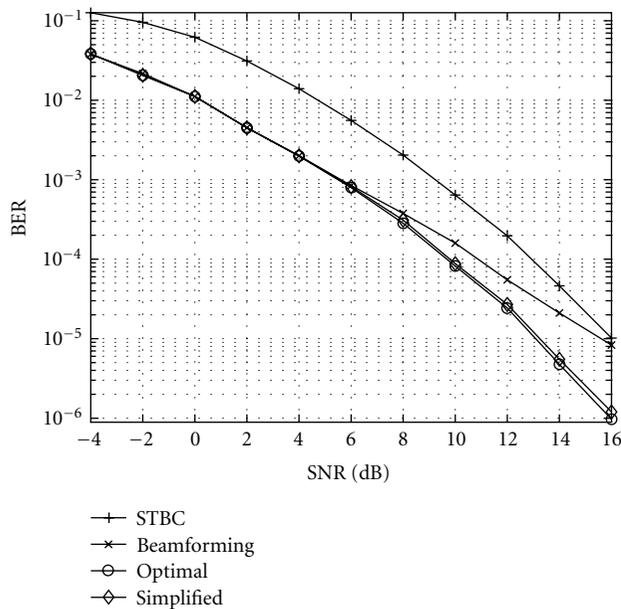


FIGURE 7: BEP curves when receive correlations are ignored.  $M = 4$ ,  $N = 2$ .

occurs in the intermediate SNR region, in the vicinity of the threshold SNR. But considering the much simpler transmitter structure and low computation complexity, the switching scheme can be seen as a good alternative to the water filling scheme, if the SNR is known at the transmitter.

Figure 9 shows the performance of the EPA scheme for a MISO system with 4 transmit antennae. The transmit correlation is again set as in (32). We can see that the switching scheme has a large performance loss in this unequal correlation case, while the EPA scheme performs very close to the optimal water-filling scheme.

## 6. CONCLUSIONS

Orthogonal space-time block codes (OSTBC) are widely used in MIMO systems to achieve diversity gain, but the performance of the conventional OSTBC over correlated fading channels deteriorates rapidly with increasing channel correlation. With feedback of the channel correlation matrix, the transmitter can employ a linear transformation unit following the STBC encoder to improve performance. One such scheme chooses the transformation matrix which minimizes the maximum pairwise error probability.

Based on the performance criterion derived in previous work, we provide a water-filling solution for the optimal transformation matrix for a MISO system. The same scheme is proven to be optimal for a receive-uncorrelated MIMO system. More generally, for arbitrary MIMO systems, we derive a “generalized water-filling” solution which can be found using numerical algorithms such as sequential quadratic programming.

Interestingly, the water-filling scheme to minimize error probability is quite similar to capacity-achieving schemes.

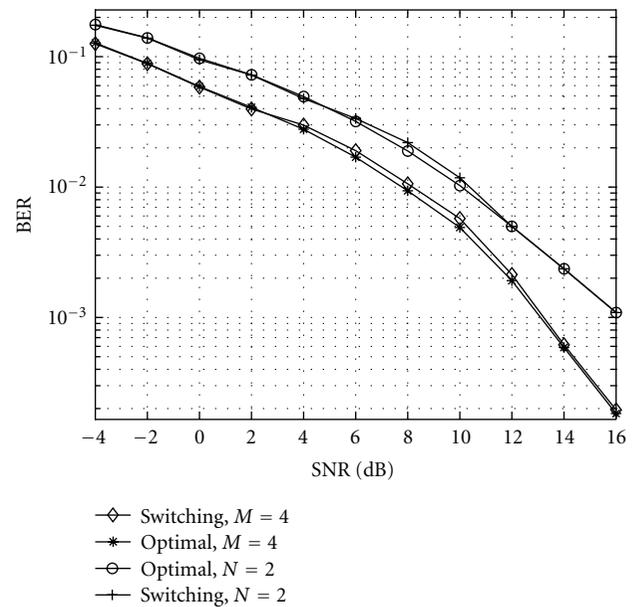


FIGURE 8: Switching scheme versus water filling.

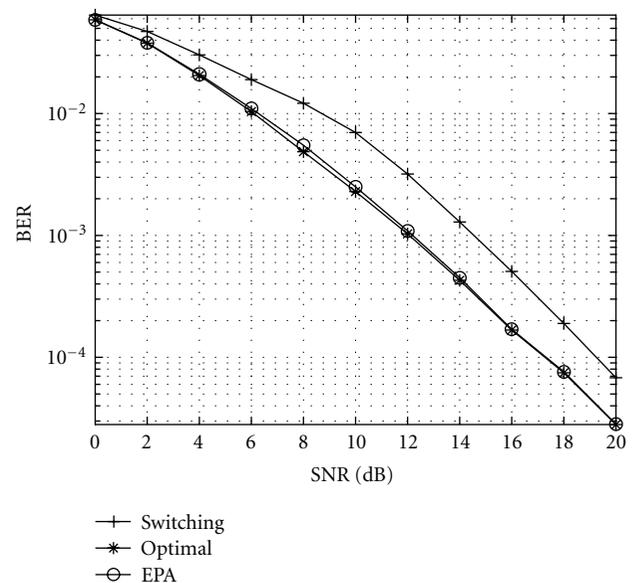


FIGURE 9: Performance of the EPA scheme.  $M = 4$ ,  $N = 1$ .

The best transmission strategy is allocating power over the eigenchannels of the transmit correlation matrices according to their eigenvalues. For both approaches, beamforming is shown to be optimal for low SNR or high correlation, while full diversity is best for high SNR and low correlation.

Based on the “one-ring” model, the correlations between receive antennae are much smaller than those between transmit antennae in the downlink of the cellular system. A simplified scheme for MIMO system is introduced by ignoring the receive correlation and using water-filling scheme with the transmit correlation only. Finally, two schemes are introduced to reduce the complexity of implementing the optimal

technique. The switching scheme uses STBC or beamforming directly based on the SNR level and channel correlation. It reduces the transmitter complexity dramatically. The EPA scheme uses the same diversity order as the optimal one, but all the active eigenchannels have the same power. We show that these schemes suffer from minimal performance loss in realistic scenarios.

## REFERENCES

- [1] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, 1998.
- [2] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1456–1467, 1999.
- [3] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744–765, 1998.
- [4] H. El Gamal, "On the robustness of space-time coding," *IEEE Trans. Signal Processing*, vol. 50, no. 10, pp. 2417–2428, 2002.
- [5] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Communications*, vol. 48, no. 3, pp. 502–513, 2000.
- [6] G. Jongren, M. Skoglund, and B. Ottersten, "Combining beamforming and orthogonal space-time block coding," *IEEE Transactions on Information Theory*, vol. 48, no. 3, pp. 611–627, 2002.
- [7] S. A. Jafar and A. Goldsmith, "On optimality of beamforming for multiple antenna systems with imperfect feedback," in *Proc. IEEE International Symposium on Information Theory (ISIT '01)*, Washington, DC, USA, June 2001.
- [8] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Transactions on Information Theory*, vol. 47, no. 6, pp. 2632–2639, 2001.
- [9] J. H. Kotechal and A. M. Sayeed, "On the capacity of correlated MIMO channels," in *Proc. IEEE International Symposium on Information Theory (ISIT '03)*, Yokohama, Japan, June–July 2003.
- [10] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block coding for wireless communications: performance results," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 3, pp. 451–460, 1999.
- [11] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, New York, NY, USA, 1985.
- [12] J. P. Kermoal, L. Schumacher, K. I. Pedersen, P. E. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 6, pp. 1211–1226, 2002.
- [13] A. Graham, *Kronecker Products and Matrix Calculus with Applications*, Ellis Horwood, Chichester, UK, 1981.
- [14] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, NY, USA, 1991.
- [15] P. E. Gill, W. Murray, and M. H. Wright, *Practical Optimization*, Academic Press, New York, NY, USA, 1981.
- [16] R. Janaswamy, *Radiowave Propagation and Smart Antennas for Wireless Communications*, Kluwer Academic, Boston, Mass, USA, 2001.
- [17] S. H. Simon and A. L. Moustakas, "Optimizing MIMO antenna systems with channel covariance feedback," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 3, pp. 406–417, 2003.

**Yi Zhao** received his B.S. degree from Tsinghua University, Beijing, China, in 2001 and the M.S. degree from the University of Toronto, Canada, in 2003. He is currently working toward the Ph.D. degree in electrical engineering at the University of Toronto. His research interests include space-time coding and space-time processing.



**Raviraj Adve** received his B.Tech. from the Indian Institute of Technology, Bombay, and his Ph.D. from Syracuse University, all in electrical engineering. Between 1997 and 2000 he was with Research Associates for Defense Conversion, Inc. working on knowledge-based space-time adaptive processing, on contract with Air Force Research Laboratory (AFRL), Rome. Since August 2000, he has been an Assistant Professor at the University of Toronto. His current research interests are in the physical layer of wireless communications, sensor networks, and adaptive processing for waveform diverse radar systems.



**Teng Joon Lim** received his B.Eng. degree from the National University of Singapore (NUS) in 1992, and the Ph.D. from Cambridge University in 1996. From 1995 to 2000, he was a member of technical staff at the Centre for Wireless Communications (now known as the Institute for Infocomm Research) in Singapore, where he was the leader of the Digital Communications Group, and an Adjunct Teaching Fellow at the NUS. He held a visiting appointment at Chalmers University in Gothenburg, Sweden, in 2000. Since December 2000, he has been an Assistant Professor at the University of Toronto. His research interests span space-time coding, multiuser system design, multicarrier modulation, and other aspects of broadband wireless communications.



# A Soft Detector with Good Performance/Complexity Trade-Off for a MIMO System

**Jianhua Liu**

*Department of Electrical and Computer Engineering, University of Florida, P.O. Box 116130, Gainesville, FL 32611-6130, USA  
Email: jhliu@dsp.ufl.edu*

**Jian Li**

*Department of Electrical and Computer Engineering, University of Florida, P.O. Box 116130, Gainesville, FL 32611-6130, USA  
Email: li@dsp.ufl.edu*

*Received 30 May 2003; Revised 23 November 2003*

We present a hybrid soft detector that has a good performance/complexity trade-off for a multiple-input multiple-output (MIMO) wireless communication system with known channel information. The new soft detector combines the merits of a simple unstructured least-squares (LS)-based soft detector and a list sphere decoder (LSD)-based soft detector for data bit detection. The former is computationally much more efficient than the latter at the cost of poorer performance. The poor performance of the former occurs mainly when the channel matrix is ill-conditioned. Whenever this happens, we use the LSD-based soft detector in the hybrid soft detector; otherwise, we use the LS-based one. Moreover, we provide a tight radius for a sphere decoder, a hard detector, via using the output of an LS-based hard detector. These two hard detectors are needed to determine if LS or LSD should be used in the hybrid soft detector. As an application example, we consider doubling the maximum data rate of the IEEE 802.11a conformable wireless local area networks by a MIMO system with two transmit and two receive antennas. For this application, the new soft detector is about 10 times faster than the LSD-based one and is about 10 times slower than the LS-based one. Yet the packet error rate due to using the new soft detector is quite close to that of using the LSD-based one.

**Keywords and phrases:** BLAST, MIMO, soft detector, convolutional codes, OFDM, WLAN.

## 1. INTRODUCTION

High transmission data rate is of particular importance for future wireless communication services. One promising way of increasing the transmission data rate is to deploy multiple antennas at both the transmitter and receiver ends to exploit the huge channel capacity offered by such a system in a multipath-rich environment [1, 2]. The corresponding system is referred to as a multiple-input multiple-output (MIMO) wireless system.

In practical communication systems, forward error correction codes, such as the convolutional code, are often used to lower the transmission error rate to an acceptable level [3, 4] by adding redundancy in the transmission. Soft detectors have been preferred to hard detectors since the former can lead to better detection/decoding performance. For the single-input single-output (SISO) systems, soft detectors have been well studied [3, 4]. Lately, much attention has been paid to the soft detectors for the MIMO systems.

The space-time bit-interleaved coded modulation (STBICM) scheme [5, 6] seems to be the best (in terms of performance) soft detector for a Bell-lab layered space-time

(BLAST) system [7, 8, 9], an especially attractive form of the MIMO systems. However, STBICM can only be implemented via the extremely inefficient brute-force search. In practice, soft detectors with good performance/complexity trade-offs are desired.

Among the other existing soft detectors, the following two are particularly attractive. One is the unstructured least-squares (LS)-based soft detector of [10], which focuses more on the computational efficiency side. The other is the list sphere decoder (LSD)-based soft detector of [11], which focuses more on the performance side. The former is very simple since, for example, it decouples a multidimensional QAM symbol detection into multiple one-dimensional QAM symbol detections. However, the performance of this detector can be poor, especially when the channel matrix is ill-conditioned. The latter has a performance close to that of STBICM with a significantly improved computational efficiency; it is based on the STBICM principle but searches in a small sphere, via modifying the sphere decoder (SPD), which is a hard detector [12]. (SPD is an efficient algorithm to implement the computationally expensive maximum-likelihood (ML) hard detector.) However, the LSD-based soft

detector still requires orders of magnitude with more computations than its LS-based counterpart.

In this paper, we combine the merits of the LS- and LSD-based soft detectors to obtain a new soft detector, referred to as the hybrid soft detector, which has a better performance than the LS-based one and a higher computational efficiency than the LSD-based one. The poor performance of the LS-based soft detector is mainly due to providing poor soft information to the Viterbi decoder as a result of the channel matrix being ill-conditioned. Whenever this happens, we use the LSD-based soft detector in the new hybrid soft detector; otherwise, we use the LS-based one. To decide if LS or LSD should be used in the hybrid detector, we check to see whether or not the output of the LS-based hard detector is the same as the output of SPD. If so, we choose LS; otherwise, we use LSD. To further improve the computational efficiency, we provide a tight radius for SPD based on the output of the LS-based hard detector.

As an example, we consider doubling the maximum data rate of the IEEE 802.11a [13] conformable wireless local area networks (WLANs) by a BLAST system with two transmit and two receive antennas. At the receiver, we use soft detectors for data bit detection. We compare the performance of the new hybrid soft detector with that of the LS- and LSD-based soft detectors. The hybrid detector is about 10 times faster than the LSD-based one and is about 10 times slower than the LS-based one. Yet the packet error rate (PER) due to using the hybrid soft detector is quite close to that of using the LSD-based one.

The remainder of this paper is organized as follows. Section 2 describes the channel encoding and decoding for a BLAST system that employs the convolutional encoder. Section 3 gives the data model and formulates the soft information, that is, bit metric. Section 4 presents the proposed new hybrid soft detector. Simulation results are given in Section 5. Finally, we provide our comments and conclusions in Section 6.

## 2. CHANNEL CODING

Consider a BLAST system with  $M$  transmit and  $N$  ( $N \geq M$ ) receive antennas, as shown in Figure 1. Figures 2 and 3, respectively, show the diagrams of the BLAST transmitter and receiver. At the transmitter, a convolutional encoder (CC) is employed, and an interleaver is used to break the memory of *bad channels* of the transmission. At the receiver, a deinterleaver is used before the convolutional (channel) decoder, for example, the Viterbi algorithm, to recover the order of the coded bit sequence. A  $1 : M$  DEMUX and  $M : 1$  MUX pair is used at the transmitter and receiver, respectively, to accommodate the BLAST scheme.

At the transmitter, as shown in Figure 2, the CC, which has a constraint length  $K_C$ , takes a block (also called packet) of bits  $\mathbf{d} = \{d_1, d_2, \dots, d_K\} \in \{-1, +1\}^{1 \times K}$  [with  $(K_C - 1)$   $(-1)$ 's at the tail to reset the CC] as its input and gives a larger block of bits  $\mathbf{u} = C(\mathbf{d}) = \{u_1, u_2, \dots, u_{\tilde{K}}\} \in \{-1, +1\}^{1 \times \tilde{K}}$  as its output, where  $-1$  and  $+1$  denote the binary digits 0

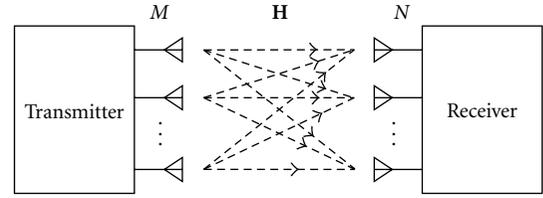


FIGURE 1: Diagram of a MIMO system.

and 1, respectively. The CC coding rate is then defined as  $R_C = K/\tilde{K}$ . We can puncture the CC output block  $\mathbf{u}$  to obtain a smaller block of bits  $\mathbf{v} = \{v_1, v_2, \dots, v_{\tilde{K}}\} \in \{-1, +1\}^{1 \times \tilde{K}}$  ( $\tilde{K} < \tilde{K}$ ) to increase the transmission data rate. The puncturing rate is  $R_p = \tilde{K}/\tilde{K}$ , and the coding rate of the punctured CC is  $R = R_C/R_p = K/\tilde{K}$ . The output  $\mathbf{v}$  of the (punctured) CC is then fed to the interleaver whose output is denoted as  $\check{\mathbf{v}} = \{v^{(1)}, v^{(2)}, \dots, v^{(\tilde{K})}\} \in \{-1, +1\}^{1 \times \tilde{K}}$ . Let  $K' = \tilde{K}/M$  be an integer. Then the outputs of the  $1 : M$  DEMUX are  $M$  independent layers, denoted as  $\check{\mathbf{v}}_m = \{v_m^{(1)}, v_m^{(2)}, \dots, v_m^{(K')}\} \in \{-1, +1\}^{1 \times K'}$ ,  $m = 1, 2, \dots, M$ . The modulator maps each layer of the bits into data symbols through the mapping  $f : \{-1, +1\}^{1 \times B} \rightarrow \mathcal{C}$ , where  $\mathcal{C}$  denotes the data symbol constellation and  $B = \log_2 |\mathcal{C}|$  is the number of bits represented by a data symbol. Let  $\tilde{K} = K'/B$  be an integer, which is the number of data symbols in each layer. Then the outputs of the modulators can be denoted as  $\check{\mathbf{x}}_m = \{x_m^{(1)}, x_m^{(2)}, \dots, x_m^{(\tilde{K})}\}$ ,  $m = 1, 2, \dots, M$ . Finally, the  $M \times 1$  data symbol vector  $\mathbf{x}^{(\tilde{k})} = [x_1^{(\tilde{k})} \ x_2^{(\tilde{k})} \ \dots \ x_M^{(\tilde{k})}]^T$ , where  $(\cdot)^T$  denotes the transpose, is transmitted through the  $M$  transmit antennas at the same time, with  $\tilde{k}$  denoting the time index,  $\tilde{k} = 1, 2, \dots, \tilde{K}$ . The bits corresponding to  $\mathbf{x}^{(\tilde{k})}$  are denoted as a  $BM \times 1$  vector  $\mathbf{b}^{(\tilde{k})} = [b_1^{(\tilde{k})} \ b_2^{(\tilde{k})} \ \dots \ b_{BM}^{(\tilde{k})}]^T$ , with  $b_i^{(\tilde{k})} \in \{-1, +1\}$ ,  $i = 1, 2, \dots, BM$ . Note that  $\mathbf{x}^{(\tilde{k})}$  is a one-to-one map of  $\mathbf{b}^{(\tilde{k})}$ , and if needed it can be written as  $\mathbf{x}^{(\tilde{k})} = \mathbf{x}(\mathbf{b}^{(\tilde{k})})$  to stress its dependence on  $\mathbf{b}^{(\tilde{k})}$ .

At the receiver, as shown in Figure 3, the soft detector first generates the bit metrics  $\{l_1^{(\tilde{k})}, l_2^{(\tilde{k})}, \dots, l_{BM}^{(\tilde{k})}\}$ , with  $l_i^{(\tilde{k})}$  being the bit metric corresponding to  $b_i^{(\tilde{k})}$ ,  $i = 1, 2, \dots, BM$ , at time  $\tilde{k} = 1, 2, \dots, \tilde{K}$ . The soft detector then rearranges the bit metrics to obtain  $\{\hat{v}_m^{([\tilde{k}-1]B+1)}, \hat{v}_m^{([\tilde{k}-1]B+2)}, \dots, \hat{v}_m^{(\tilde{k}B)}\}$  for the bits  $\{v_m^{([\tilde{k}-1]B+1)}, v_m^{([\tilde{k}-1]B+2)}, \dots, v_m^{(\tilde{k}B)}\}$ , which were mapped to the data symbol  $x_m^{(\tilde{k})}$ . Let  $\hat{\mathbf{v}}_m = \{\hat{v}_m^{(1)}, \hat{v}_m^{(2)}, \dots, \hat{v}_m^{(K')}\}$ ,  $m = 1, 2, \dots, M$ , denote the bit metric sequence corresponding to the  $m$ th transmitted layer. The  $M$  bit metric sequences are combined into one longer bit metric sequence  $\hat{\mathbf{v}} = \{\hat{v}^{(1)}, \hat{v}^{(2)}, \dots, \hat{v}^{(\tilde{K})}\}$  by the  $M : 1$  MUX. Passing the above bit metric sequence  $\hat{\mathbf{v}}$  through the deinterleaver, we obtain the deinterleaved bit metric sequence  $\hat{\mathbf{v}} = \{\hat{v}_1, \hat{v}_2, \dots, \hat{v}_{\tilde{K}}\}$ . For the punctured CC codes, we need the bit metric for each punctured bit as well before using the Viterbi algorithm. This can be done easily by using zero as the bit metric for each punctured bit. Once we get the bit metric sequence  $\hat{\mathbf{u}} = \{\hat{u}_1, \hat{u}_2, \dots, \hat{u}_{\tilde{K}}\}$  corresponding to the CC output  $\mathbf{u}$ , we can use the Viterbi algorithm to obtain the estimate  $\hat{\mathbf{d}} = \{\hat{d}_1, \hat{d}_2, \dots, \hat{d}_K\}$  of the source bit sequence  $\mathbf{d}$ .

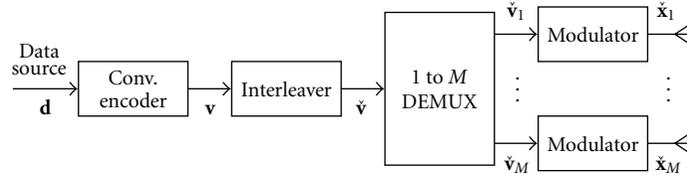


FIGURE 2: Diagram of a BLAST transmitter employing convolutional channel coding.

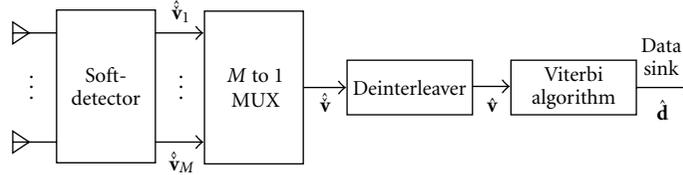


FIGURE 3: Diagram of a BLAST receiver employing Viterbi decoding for convolutional codes.

In the sequel, we focus on the calculation of the bit metrics for the bits in the QAM symbol, due to our WLAN application.

### 3. DATA MODEL AND BIT METRIC

We now give the data model and formulate the bit metric for the BLAST system.

#### 3.1. Data model

The channel matrix of a MIMO time-varying flat Rayleigh-fading channel at time  $\bar{k}$  can be written as

$$\mathbf{H}^{(\bar{k})} = \begin{bmatrix} h_{1,1}^{(\bar{k})} & h_{1,2}^{(\bar{k})} & \cdots & h_{1,M}^{(\bar{k})} \\ h_{2,1}^{(\bar{k})} & h_{2,2}^{(\bar{k})} & \cdots & h_{2,M}^{(\bar{k})} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N,1}^{(\bar{k})} & h_{N,2}^{(\bar{k})} & \cdots & h_{N,M}^{(\bar{k})} \end{bmatrix} \in \mathbb{C}^{N \times M}, \quad (1)$$

where  $h_{n,m}^{(\bar{k})}$  is the gain from the  $m$ th transmit antenna to the  $n$ th receive antenna at time  $\bar{k}$ , which is assumed to be known. With  $\mathbf{x}^{(\bar{k})} = [x_1^{(\bar{k})} \ x_2^{(\bar{k})} \ \cdots \ x_M^{(\bar{k})}]^T$  denoting the  $M \times 1$  QAM symbol vector being sent at time  $\bar{k}$ , the received signal can be written as

$$\mathbf{y}^{(\bar{k})} = \mathbf{H}^{(\bar{k})}\mathbf{x}^{(\bar{k})} + \mathbf{n}^{(\bar{k})} \in \mathbb{C}^{N \times 1}, \quad \bar{k} = 1, 2, \dots, \bar{K}, \quad (2)$$

where  $\mathbf{n}^{(\bar{k})} \sim \mathbf{N}(\mathbf{0}, \sigma_k^2 \mathbf{I}_N)$  is the additive zero-mean white circularly symmetric complex Gaussian noise.

With an appropriate pair of interleaver and deinterleaver, the MIMO channel can be assumed to be block flat Rayleigh fading [14, 15], that is,  $\mathbf{H}^{(\bar{k})}$  is constant at time  $\bar{k}$  for the transmission of  $\mathbf{x}^{(\bar{k})}$  but changes independently from one time index to another. In the sequel, we focus on obtaining the soft information given that we know the channel matrix  $\mathbf{H}^{(\bar{k})}$ , the noise variance  $\sigma_k^2$ , and the received data vector  $\mathbf{y}^{(\bar{k})}$ . For notational convenience, we drop the superscript  $\bar{k}$  in (2) to get

the data model

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \in \mathbb{C}^{N \times 1}, \quad (3)$$

or

$$\mathbf{y} = \mathbf{H}\mathbf{x}(\mathbf{b}) + \mathbf{n}. \quad (4)$$

#### 3.2. Bit metric

The bit metric (also known as the L-value) for the  $i$ th bit,  $i = 1, 2, \dots, BM$ , is defined as

$$l_i = \log \frac{P(b_i = +1 | \mathbf{y}, \mathbf{H})}{P(b_i = -1 | \mathbf{y}, \mathbf{H})}. \quad (5)$$

Assuming equal probability for each data bits and using Bayes' theorem, the bit metric can be written as

$$l_i = \log \frac{\sum_{\mathbf{b} \in \mathcal{B}_{i,+1}} P(\mathbf{y} | \mathbf{b}, \mathbf{H})}{\sum_{\mathbf{b} \in \mathcal{B}_{i,-1}} P(\mathbf{y} | \mathbf{b}, \mathbf{H})}, \quad (6)$$

where  $\mathcal{B}_{i,+1}$  and  $\mathcal{B}_{i,-1}$  are the set of  $2^{BM-1}$  bit vectors  $\mathbf{b}$  with  $b_i$  being +1 and -1, respectively.

With the assumption of additive zero-mean white circularly symmetric complex Gaussian noise for the received data, the above equation can be written as

$$l_i = \log \frac{\sum_{\mathbf{b} \in \mathcal{B}_{i,+1}} e^{-(1/\sigma^2)\|\mathbf{y} - \mathbf{H}\mathbf{b}\|^2}}{\sum_{\mathbf{b} \in \mathcal{B}_{i,-1}} e^{-(1/\sigma^2)\|\mathbf{y} - \mathbf{H}\mathbf{b}\|^2}}, \quad (7)$$

which, by using the max-log approximation [16], can be written as

$$\begin{aligned} l_i &\approx \max_{\mathbf{b} \in \mathcal{B}_{i,+1}} \left\{ -\frac{1}{\sigma^2} \|\mathbf{y} - \mathbf{H}\mathbf{b}\|^2 \right\} \\ &\quad - \max_{\mathbf{b} \in \mathcal{B}_{i,-1}} \left\{ -\frac{1}{\sigma^2} \|\mathbf{y} - \mathbf{H}\mathbf{b}\|^2 \right\} \\ &= \frac{1}{\sigma^2} \left[ \min_{\mathbf{b} \in \mathcal{B}_{i,-1}} \|\mathbf{y} - \mathbf{H}\mathbf{b}\|^2 - \min_{\mathbf{b} \in \mathcal{B}_{i,+1}} \|\mathbf{y} - \mathbf{H}\mathbf{b}\|^2 \right]. \end{aligned} \quad (8)$$

This is in fact the optimal but extremely inefficient STBICM soft detector.

In the sequel, we present the hybrid soft-detector, which has a good performance/complexity trade-off, for calculating the bit metric.

#### 4. THE PROPOSED SOFT DETECTOR

The proposed soft detector is based on the combination of the LSD- and LS-based soft detectors. As a result, before presenting the new soft detector, we summarize and comment on the merits of these two existing detectors, which are different approximations of (8) with different focuses on the performance/complexity trade-off.

##### 4.1. The LSD-based soft detector

The LSD-based soft detector focuses mainly on the performance side of the performance/complexity trade-off. It maintains the framework of the STBICM detector and improves the efficiency of (8) by searching in much smaller subsets  $\tilde{\mathcal{B}}_{i+1} \subset \mathcal{B}_{i+1}$  and  $\tilde{\mathcal{B}}_{i-1} \subset \mathcal{B}_{i-1}$  with  $|\tilde{\mathcal{B}}_{i+1}| \ll 2^{BM-1}$  and  $|\tilde{\mathcal{B}}_{i-1}| \ll 2^{BM-1}$ . The LSD-based soft detector is implemented in the following two steps.

Step SD1. Obtain the set  $\tilde{\mathcal{B}}$  of vectors  $\mathbf{b}$  which satisfies

$$\|\mathbf{y} - \mathbf{H}\mathbf{x}(\mathbf{b})\| \leq d_i, \quad \forall \mathbf{b} \in \tilde{\mathcal{B}}, \quad (9)$$

by using the modified SPD algorithm that has a fixed sphere radius  $d_i$ , determined by the antenna numbers and noise variance [11].

Step SD2. For each  $i = 1, 2, \dots, BM$ , calculate  $\tilde{\mathcal{B}}_{i+1} = \mathcal{B}_{i+1} \cap \tilde{\mathcal{B}}$  and  $\tilde{\mathcal{B}}_{i-1} = \mathcal{B}_{i-1} \cap \tilde{\mathcal{B}}$  and obtain the bit metric by

$$l_i^{(\text{SD})} = \frac{1}{\sigma^2} \left[ \min_{\mathbf{b} \in \tilde{\mathcal{B}}_{i+1}} \|\mathbf{y} - \mathbf{H}\mathbf{x}(\mathbf{b})\|^2 - \min_{\mathbf{b} \in \tilde{\mathcal{B}}_{i-1}} \|\mathbf{y} - \mathbf{H}\mathbf{x}(\mathbf{b})\|^2 \right]. \quad (10)$$

At the cost of some performance degradation, the LSD-based soft detector improves the computational efficiency of the STBICM detector significantly due to limiting the search over the much smaller sets. (We do not know the exact degradation for our WLAN application since the STBICM detector is too slow to make a reasonable comparison.) However, the LSD-based soft detector is not as efficient as SPD due to the following reasons: (a) LSD in Step SD1 uses fixed sphere radius whereas SPD uses changing sphere radius that shrinks with the finding of a new point in the sphere with a shorter distance and (b) the bit metric calculation in Step SD2 needs additional computations.

##### 4.2. The LS-based soft detector

The LS-based soft detector focuses mainly on the computational complexity side of the performance/complexity trade-off. While the LSD-based soft detector improves the efficiency of (8) by limiting the search on smaller sets, the LS-based soft detector decreases the computation of (8) by decoupling the distance  $\|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2$  into  $M$  separate distances,

that is, it decouples a MIMO channel into multiple SISO channels that are processed independently of each other. The LS-based soft detector has the following two main steps.

Step LS1. Ignore the discrete constellation of  $\mathbf{x}$  to obtain an unstructured LS symbol estimate  $\mathbf{x}^{(\text{LS})}$  of  $\mathbf{x}$  as

$$\begin{aligned} \mathbf{x}^{(\text{LS})} &= (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H \mathbf{y} \\ &= \mathbf{x} + (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H \mathbf{n} \\ &\triangleq \mathbf{x} + \mathbf{e}. \end{aligned} \quad (11)$$

Step LS2. For  $j = 1, 2, \dots, B$  and  $m = 1, 2, \dots, M$ , obtain the bit metric for each bit using the scheme (similar to (8), but for the SISO case) given in [17]

$$l_{j,m}^{(\text{LS})} = \frac{1}{\sigma_m^2} \left[ \min_{\mathbf{b}_m \in \mathcal{B}_{m,j,-1}} |x_m^{(\text{LS})} - x(\mathbf{b}_m)|^2 - \min_{\mathbf{b}_m \in \mathcal{B}_{m,j,+1}} |x_m^{(\text{LS})} - x(\mathbf{b}_m)|^2 \right], \quad (12)$$

where  $\mathcal{B}_{m,j,+1}$  and  $\mathcal{B}_{m,j,-1}$  are the set of  $2^{B-1}$  bit vectors  $\mathbf{b}_m \in \{-1, +1\}^{B \times 1}$  with the  $j$ th bit being  $+1$  and  $-1$ , respectively,  $x_m^{(\text{LS})}$  is the  $m$ th element of  $\mathbf{x}^{(\text{LS})}$ ,  $x(\mathbf{b}_m) \in \mathcal{C}$ , and  $\sigma_m^2 = \sigma^2 [(\mathbf{H}^H \mathbf{H})^{-1}]_{m,m}$  with  $[\mathbf{A}]_{m,m}$  denoting the  $(m, m)$ th element of matrix  $\mathbf{A}$ .

We remark that for the SISO systems, we usually consider an ordinary QAM symbol as two PAM symbols (e.g., a 64-QAM symbol can be considered as two 8-PAM symbols) due to the orthogonality between the real and imaginary parts of a QAM symbol as well as the independence between the real and imaginary parts of the additive circularly symmetric Gaussian noise. The same is true for the BLAST systems employing the LS-based soft detector since the real and imaginary parts of  $e_m$ , the  $m$ th element of  $\mathbf{e}$  in (12),  $m = 1, 2, \dots, M$ , are independent of each other, as shown below:

$$\mathbf{E}[\mathbf{e}\mathbf{e}^T] = (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{E}[\mathbf{n}\mathbf{n}^T] \mathbf{H}^* \left[ (\mathbf{H}^H \mathbf{H})^{-1} \right]^T = \mathbf{0}, \quad (13)$$

where we have used the fact that  $\mathbf{E}[\mathbf{n}\mathbf{n}^T] = \mathbf{0}$ .

The LS-based soft detector is orders of magnitude more efficient than the LSD-based soft detector due to the decoupling, as will be analyzed later. However, the performance of the former is worse than the latter (more than 2 dB for the  $M = N = 2$  case for our WLAN application, to be shown by the simulation examples later).

By rounding  $x_m^{(\text{LS})}$ ,  $m = 1, 2, \dots, M$ , to the closest point in the constellation  $\mathcal{C}$ , we obtain the output of the LS-based hard detector, which will be used latter.

Note that the minimum mean-squared error (MMSE)-based soft detector is often deemed to be better than the LS-based one [18]. Although this can be true for the constant-modulus constellations, such as PSK, it is not necessarily true for QAM symbols, as suggested by our simulation results (not provided here) due to the different power levels

of the QAM symbols. Hence the LS-based soft detector is more preferable than the MMSE-based one since the former is slightly more computationally efficient than the latter.

### 4.3. The hybrid soft detector

The above two soft detectors provide different performance/complexity trade-offs for data bit detection, with the LSD-based one focusing on the performance and the LS-based one on the computational efficiency. In practice, it is desirable to have a soft detector that is better than the LS-based one in performance and faster than the LSD-based one in computational complexity. We show that this can be done by combining these two soft detectors, and the corresponding new detector is referred to as the hybrid soft detector.

Now, we examine what hinders the performance of the LS-based soft detector. We can readily see that when  $\mathbf{H}^H\mathbf{H}$  is close to a scaled identity matrix, the bit metrics from the LS-based soft detector will not be worse than those from the LSD-based one. However, when  $\mathbf{H}^H\mathbf{H}$  becomes ill-conditioned, the bit metrics from the former will be much worse than those from the latter, because of the following reasons: (a) some elements of the noise vector  $\mathbf{e}$  in (11) are magnified drastically due to the poor channels and (b) useful information is lost due to the decoupling. Hence, these (bad) bit metrics corresponding to the ill-conditioned channels can be seen as the bottleneck for the performance of the LS-based soft detector. If we can identify these bad bit metrics and replace them by those from the LSD-based soft detector, we can improve the detection performance significantly.

We identify the bad bit metrics by comparing the LS-based hard detector output  $\hat{\mathbf{x}}^{(LS)}$  and the SPD output  $\hat{\mathbf{x}}^{(SPD)}$ . If they are not the same,  $\hat{\mathbf{x}}^{(LS)}$  is more likely to have error(s) since  $\hat{\mathbf{x}}^{(SPD)}$  is an ML estimate, which is better than the former theoretically. In this case, the corresponding bit metrics from the LS-based soft detector are considered bad; otherwise, these bit metrics can be considered reliable.

In view of the above, we have the following steps for the hybrid soft detector.

- Step HY1. Obtain the LS symbol estimate  $\mathbf{x}^{(LS)}$  by using (11) of Step LS1.
- Step HY2. Determine the LS hard detection result  $\hat{\mathbf{x}}^{(LS)}$ .
- Step HY3. Calculate the SPD detection result  $\hat{\mathbf{x}}^{(SPD)}$ .
- Step HY4. Check the hard detection results—if  $\hat{\mathbf{x}}^{(LS)} = \hat{\mathbf{x}}^{(SPD)}$ , then go to Step HY5; otherwise, go to Step HY6.
- Step HY5. Obtain bit metrics by (12) of Step LS2 based on  $\mathbf{x}^{(LS)}$  from Step HY1 and then stop.
- Step HY6: obtain bit metrics by performing Steps SD1 and SD2 and then stop.

The computational complexity of the hybrid soft detector is dominated by SPD and the LSD-based soft detector, that is, Steps HY3 and HY6. To speed up the calculation of SPD in Step HY3, we need to consider the determination of its initial radius, which is a crucial issue for SPD. If the initial radius is too small, there will be no point ( $\mathbf{x}$ ) in the sphere—SPD cannot find the ML solution. On the other hand, if the initial radius is too large, SPD will be very slow due to the un-

necessary additional searches. The number of the additional searches can be reduced by using a modified searching approach given in [19]. However, it complicates the algorithms itself. Here, we give a *tight* sphere radius, based on the LS-based hard-detector output  $\hat{\mathbf{x}}^{(LS)}$ , by using

$$d_r = \|\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}^{(LS)}\| + \epsilon_d, \quad (14)$$

where  $\epsilon_d > 0$  is a very small value. Note that this radius will contain at least one point—the output of the LS-based hard-detector. Note also that, for most cases (98 out of 100 for the signal-to-noise-ratios (SNRs) of interest in our WLAN application, as will be shown by the simulation results in the next section), this radius contains only one point. By using this tight radius, our preliminary simulation results show that SPD can be as efficient as the interference cancellation and nulling algorithm [8] and uses only 5 times as many flops as the LS-based soft detector.

The computational complexity, in terms of flops, for each step of the LSD-based soft detector, can be estimated as follows. (We assume  $M = N$  for convenience.)

- Step HY1:  $\mathcal{O}(M^3)$  for matrix multiplications and inversion. For example, a calculation using Matlab indicates that the number of flops is 444 for the  $M = 2$  case.
- Step HY2: Negligible.
- Step HY3:  $\mathcal{O}(M^3)$  to  $\mathcal{O}(M^6)$  for SPD, depending on the SNR and  $B$  [12, 20]. For example, preliminary calculations using Matlab show that, by using the tight radius, SPD uses only 5 times as many flops as LS in Step HY1 for 64-QAM,  $M = 2$ , and the SNRs of interest.
- Step HY4: Negligible.
- Step HY5: Negligible by table checking for the PAM symbols.
- Step HY6: (a)  $\mathcal{O}(M^3)$  to  $\mathcal{O}(M^6)$  for LSD, which, as shown by simulation results, uses typically 2 to 10 times as many flops as SPD in Step HY3, that is, 10 to 50 times as many flops as LS in Step HY1. (We use the average 25 in the sequel.) (b)  $\mathcal{O}(N_C^2BM)$  for bit metric calculation, where  $N_C$  is the number of candidates in the list for LSD and the operation of finding the minimum is performed by using the conventional bubbling algorithm; for example, for  $M = 2$ ,  $B = 6$ , and  $N_C = 120$  (which is typical for a good performance), this amounts to about 43200 flops (assuming  $|\mathcal{B}_{i+1}| = |\mathcal{B}_{i-1}| = 60$ ,  $i = 1, 2, \dots, 12$ , for convenience), which is about 95 times as many flops as LS in Step HY1.

As will be seen from the simulation results in the next section, less than 2% of the cases have different SPD and LS hard detection results. Hence, we can see that the hybrid soft detector is about

$$\underbrace{1}_{LS} + \underbrace{5}_{SPD} + 0.02 \times \left( \underbrace{25}_{LSD} + 95 \right) = 8.4 \quad (15)$$

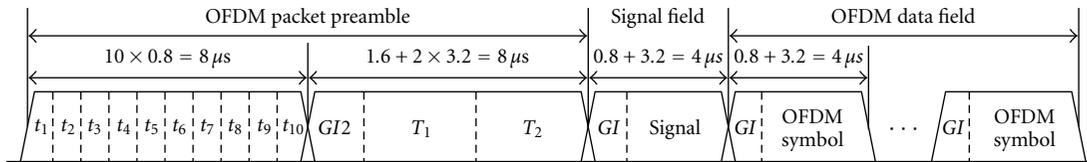


FIGURE 4: Packet structure of the IEEE 802.11a standard.

times as slow as the LS-based soft detector, which indicates that the hybrid soft detector is *about* 10 times slower than the LS-based soft detector. We can also see that the LSD-based soft detector needs 120 times as many flops as the LS-based one, which means that the hybrid soft detector is *about* 10 times faster than the LSD-based one. (This will be elaborated in the next section based on the parameters of our WLAN application.) Note that the new hybrid soft detector is more efficient for high SNRs than for low SNRs since at high SNRs the probabilities of  $\hat{\mathbf{x}}^{(LS)} = \hat{\mathbf{x}}^{(SPD)}$  are high and the chances of using the computationally expensive LSD-based soft detector are low. Note also that the above analysis of the complexity is only intended to give a feeling about the efficiency of the hybrid soft detector and is by no means very accurate. More accurate analysis of the complexities, including those for SPD and LSD, is still an open topic.

We remark that the bad bit metrics can also be identified by using the condition number (CN) of  $\mathbf{H}^H \mathbf{H}$ , and the resulting soft detector can be referred to as the CN-hybrid soft detector. However, the CN-hybrid soft detector is inferior to its hybrid counterpart due to the following reasons. First, it is hard to determine a threshold for the CN. If the threshold is too high, many bad bit metrics from the LS-based soft detector will be used in the hybrid soft detector, which will lead to degraded performance. On the other hand, if the threshold is too low, the computationally expensive LSD-based soft detector will be used too often, which will result in increased computational complexity. Second, a large CN does not necessarily result in detection differences between SPD and the LS-based hard detectors. Neither does a small condition number guarantee the same detection result for the two hard detectors. As will be demonstrated using the simulation results in the next section, for a practical choice of CN, say 100, the CN-hybrid soft detector has a comparable ( $0.06 \times (25 + 95) = 7.2$  times as many flops as LS) complexity with the hybrid soft detector; yet, the performance of the former is inferior to the latter.

## 5. SIMULATION RESULTS

Our results obtained under the flat fading channel condition can be readily extended to the orthogonal frequency-division multiplexing (OFDM)-based WLAN systems operating over frequency-selective fading channels [21]. This is because for each subcarrier the channel is a flat fading one. In our simulations, we follow the IEEE 802.11a 5 GHz band high-speed WLAN standard [13] whenever possible.

The OFDM-based WLAN system, as specified by the IEEE 802.11a standard, uses packet-based transmission. Figure 4 shows the packet structure specified by the standard.

Each packet consists of many OFDM symbols. Each OFDM symbol occupies 64 subcarriers, among which 48 are used for data symbols and 4 for pilot symbols. There are also 12 null subcarriers. The OFDM symbols are obtained via taking the inverse fast Fourier transform (FFT) of the data, pilots, and nulls on these subcarriers. The nominal bandwidth of the OFDM signal is 20 MHz and the I/Q sampling interval is 50 nanoseconds. Due to the fact that the modulation and demodulation are done in the frequency domain, a frequency domain bit-level interleaver is used to segment the encoded bit sequence according to the transmission data rate and to scatter them over the 48 different data-carrying subcarriers. Before interleaving, an industrial standard constraint length 7 and  $R_C = 1/2$  CC is employed to code the source bit sequence. In the IEEE 802.11a standard, the maximum transmission data rate is 54 Mbps; in this case the 64-QAM constellation is used and the channel coding rate is  $R = 3/4$ , which comes from puncturing the  $R_C = 1/2$  convolutionally encoded sequence with the puncturing rate  $R_P = 2/3$ . The channel is assumed to be fixed during the packet transmission.

We consider doubling the maximum 54 Mbps transmission data rate by using a BLAST system with two transmit and two receive antennas, that is,  $M = N = 2$ . This OFDM-based BLAST WLAN system is backward compatible with its SISO counterpart, with the packet structure shown in Figure 5. (See [21] for more detailed description of the MIMO system design.) In our simulations, each of the  $MN = 4$  time domain MIMO channels is generated according to the exponential channel model [22] with the root-mean-square spreading time  $t_{\text{rms}}$  being 50 nanoseconds; the 4 channels are statistically independent of each other. After FFT at the receiver, the channel matrix for each subcarrier has the same form as in (1), with the  $\bar{k}$  being the subcarrier index in this case. This subcarrier index is equivalent to the time index for time-varying flat fading channels since the channel for the OFDM-based WLANs is assumed to be fixed for the entire packet, with the changes across the subcarriers due to the delay time spreading. (Note that the intersymbol interference is avoided in the OFDM-based systems due to using the cyclic prefix [13].) We consider the case of perfect channel knowledge, where the carrier frequency offset, symbol timing, channel response, and noise variance are all known in all our simulations; in practical applications, these parameters can be estimated via applying the channel parameter estimation methods, such as those in [21, 23, 24, 25], to the packet preambles.

Due to the fact that 52 out of 64 subcarriers are used in the OFDM-based SISO WLAN system, the SNR used herein is defined as  $52\sigma^2/64$  for the 64-QAM constellation whose

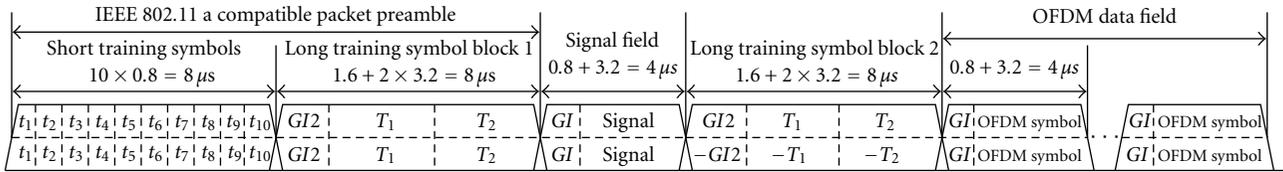


FIGURE 5: Packet structure for the OFDM-based BLAST WLAN system.

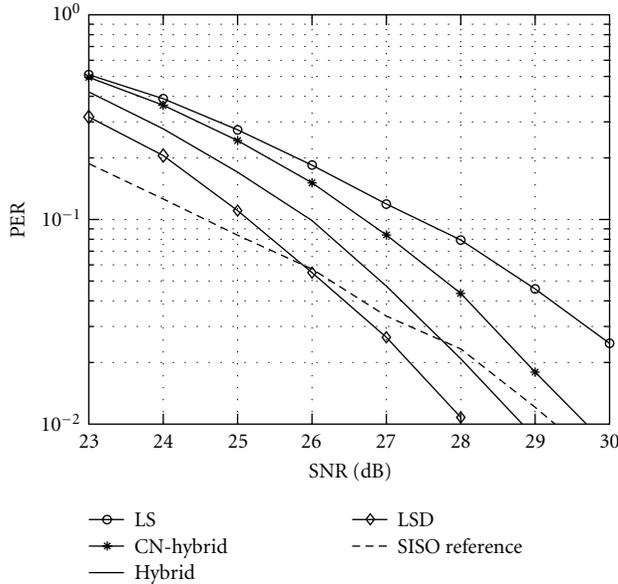


FIGURE 6: PER versus SNR comparisons for the soft detectors.

average energy is normalized to 1. For the OFDM-based BLAST WLAN system, we keep the same total transmission power and maintain the same subcarrier structure as its SISO counterpart.

We give two simulation examples to demonstrate the performance and computational complexity of our hybrid soft detector.

*Example 1 (Performance).* The PER (one packet consists of 1000 bytes, which are contained in 19 OFDM symbols) is an important parameter for the OFDM-based WLAN systems. (In an OFDM-based WLAN system, even if only one error occurs, the entire packet is considered to be wrong.) In Figure 6, we show the PER comparison for the LS-based soft detector, the CN-hybrid soft detector (with CN being 100), the hybrid soft detector, and the LSD-based soft detector as a function of SNR for the OFDM-based BLAST WLAN system at the 108 Mbps data rate. We also give the PER curve of using the soft detector for the SISO system at the 54 Mbps data rate as a reference. We can see from the simulation results that for the OFDM-based BLAST WLAN system, the performance of the hybrid soft detector is close to that of the LSD-based soft detector. We can also see that the hybrid soft detector outperforms its CN-hybrid counterpart. Moreover, the PER curve of the hybrid soft detector has nearly the same slope

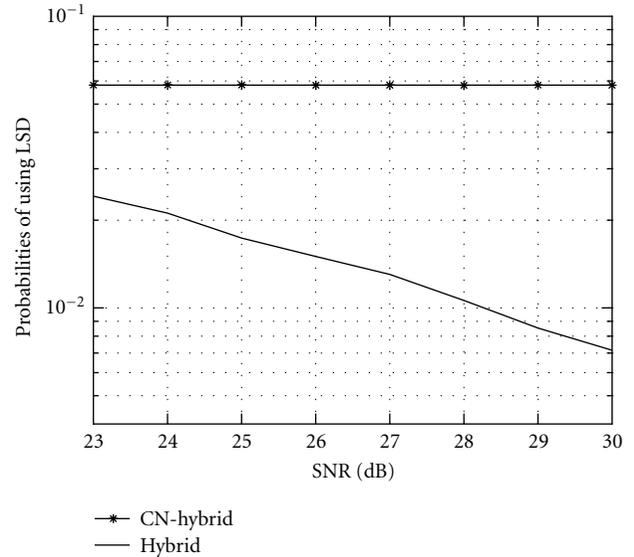


FIGURE 7: Probabilities of using LSD in the hybrid and CN-hybrid soft detectors as SNR varies.

as the LSD-based one, which means that at high SNRs, the hybrid soft detector can offer much better performance than the LS-based one. Also, by comparing the solid line with the dashed line, we can see that if we use the hybrid soft detector at the receiver, we need about 1.5 dB more SNR to keep the same 10% PER (we are currently mostly interested in PERs of 10%) to double the data rate with  $M = N = 2$ . Note that even with the need of this 1.5 dB extra SNR, that is, 1.5 dB more total transmission power, the PER performance of the OFDM-based BLAST WLAN system with the hybrid soft detector is still impressive since even if we wish to double the transmission data rate using two separate SISO systems over two different physical channels by doubling the bandwidth, we still need 3 dB extra SNR or total transmission power. If we consider the case of 1% PER, we can double the data rate with about 0.5 dB less total transmission power.

*Example 2 (Complexity).* To facilitate the analysis of the complexity of the hybrid and CN-hybrid (with CN being 100) soft detectors, we provide a simulation example to demonstrate the probability of using the LSD-based soft detector in these soft detectors. We can see from Figure 7 that for the SNRs of interest, the probability of using the LSD-based soft detector in the CN-hybrid soft detector is about 6% and less than 2% in the hybrid soft detector.

## 6. CONCLUDING REMARKS

We have proposed a hybrid soft detector with a good performance/complexity trade-off by combining the LS- and LSD-based soft detectors. The combination is performed based on comparing the outputs of SPD and the LS-based hard detector. To speed up the computation of SPD, we have also provided a tight sphere radius that can be used to guarantee the finding of at least one solution. Simulation results have shown that the performance of our hybrid soft detector is close to that of the LSD-based soft detector in our WLAN application. The new detector is about 10 times faster than the LSD-based and about 10 times slower than the LS-based soft detectors.

## ACKNOWLEDGMENTS

We would like to thank the anonymous reviewers for their helpful comments which have helped improve the quality of this paper. This work was supported in part by the National Science Foundation Grant CCR-0097114 and the Intersil Corporation Contract 2001056.

## REFERENCES

- [1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, 1998.
- [2] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transaction Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.
- [3] S. Lin and D. J. Costello, *Error Control Coding*, Prentice Hall, Englewood Cliffs, NJ, USA, 1982.
- [4] J. G. Proakis, *Digital Communications*, McGraw-Hill, New York, NY, USA, 3rd edition, 1995.
- [5] A. M. Tonello, "Space-time bit-interleaved coded modulation with an iterative decoding strategy," in *Proc. IEEE VTS-Fall VTC 2000. 52nd*, vol. 1, pp. 473–478, Boston, Mass, USA, 2000.
- [6] J. J. Boutros, F. Boixadera, and C. Lamy, "Bit-interleaved coded modulations for multiple-input multiple-output channels," in *Proc. 6th IEEE International Symposium on Spread Spectrum Techniques and Applications*, vol. 1, pp. 123–126, Parsippany, NJ, USA, September 2000.
- [7] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, 1996.
- [8] G. D. Golden, G. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, "Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture," *Electronics Letters*, vol. 35, no. 1, pp. 14–15, 1999.
- [9] Z. Liu and G. B. Giannakis, "Layered space-time coding for high data rate transmissions," in *Proc. IEEE Military Communications Conference (MILCOM' 01)*, vol. 2, pp. 1295–1299, McLean, VA, USA, October 2001.
- [10] J. Liu and J. Li, "A simple soft-detector for the BLAST system," in *Proc. Sensor Array and Multichannel Signal Processing Workshop (SAM' 02)*, pp. 159–163, Rosslyn, Va, USA, August 2002.
- [11] B. M. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Communications*, vol. 51, no. 3, pp. 389–399, 2003.
- [12] O. Damen, A. Chkeif, and J.-C. Belfiore, "Lattice code decoder for space-time codes," *IEEE Communications Letters*, vol. 4, no. 5, pp. 161–163, 2000.
- [13] IEEE 802.11a-1999, "IEEE Standard for information technology—telecommunications and information exchange between systems—local and metropolitan area networks—specific requirements—part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications—amendment 1: High-speed physical layer in the 5 GHz band," IEEE, 1999.
- [14] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Transactions on Information Theory*, vol. 45, no. 1, pp. 139–157, 1999.
- [15] J.-C. Guey, M. P. Fitz, M. R. Bell, and W.-Y. Kuo, "Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels," *IEEE Trans. Communications*, vol. 47, no. 4, pp. 527–537, 1999.
- [16] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Transactions on Information Theory*, vol. 42, no. 2, pp. 429–445, 1996.
- [17] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Transactions on Information Theory*, vol. 44, no. 3, pp. 927–946, 1998.
- [18] C. Z. W. H. Swetman, J. S. Thompson, B. Mulgrew, and P. M. Grant, "A comparison of the MMSE detector and its BLAST versions for MIMO channels," in *IEE Seminar on MIMO: Communications Systems from Concept to Implementations*, pp. 19/1–19/6, London, UK, December 2001.
- [19] A. M. Chan and I. Lee, "A new reduced-complexity sphere decoder for multiple antenna systems," in *Proc. IEEE International Conference on Communications (ICC' 02)*, vol. 1, pp. 460–464, New York, NY, USA, 2002.
- [20] M. O. Damen, K. Abed-Meraim, and M. S. Lemdani, "Further results on the sphere decoder," in *Proc. 2001 IEEE International Symposium on Information Theory*, pp. 1–333, Washington, DC, USA, June 2001.
- [21] J. Liu and J. Li, "A MIMO system with backward compatibility for OFDM-based WLANs," *EURASIP Journal on Applied Signal Processing*, vol. 2004, no. 5, pp. 696–706, 2004.
- [22] N. Chayat, "Tentative criteria for comparison of modulation methods," *Doc. IEEE 802.11-97/96*, September 1997.
- [23] E. G. Larsson, G. Liu, J. Li, and G. B. Giannakis, "Joint symbol timing and channel estimation for OFDM based WLANs," *IEEE Communications Letters*, vol. 5, no. 8, pp. 325–327, 2001.
- [24] J. Li, G. Liu, and G. B. Giannakis, "Carrier frequency offset estimation for OFDM-based WLANs," *IEEE Signal Processing Letters*, vol. 8, no. 3, pp. 80–82, 2001.
- [25] J. Liu and J. Li, "Channel parameter estimation and error reduction for OFDM-based WLANs," to appear in *IEEE Transactions on Mobile Computing*.

---

**Jianhua Liu** received the B.S. degree in electrical engineering from Dalian Maritime University, Dalian, China, in 1984, the M.S. degree in electrical engineering from the University of Electronic Science and Technology of China, Chengdu, China, in 1987, and the Ph.D. degree in electronic engineering from Tsinghua University, Beijing, China, in 1998. From March 1987 to February 1999, he worked at the Communication, Telemetry and Telecontrol Research Institute, Shijiazhuang, China, where he was an Assistant Engineer, Engineer, Senior Engineer, and Fellow Engineer. From March 1995 to August 1998,



he was also a Research Assistant at Tsinghua University. From February 1999 to June 2000, he worked at Nanyang Technological University, Singapore, as a Research Fellow. Since June 2000, he has been a Research Assistant in the Department of Electrical and Computer Engineering at the University of Florida, Gainesville, working towards a Ph.D. degree majoring in electrical engineering and minoring in statistics. His research interests include wireless communications, statistical signal processing, and sensor array processing.

**Jian Li** received the M.S. and Ph.D. degrees in electrical engineering from The Ohio State University, Columbus, in 1987 and 1991, respectively. From April 1991 to June 1991, she was an Adjunct Assistant Professor with the Department of Electrical Engineering, The Ohio State University, Columbus. From July 1991 to June 1993, she was an Assistant Professor with the Department of Electrical Engineering, University of Kentucky, Lexington. Since August 1993, she has been with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, where she is currently a Professor. Her current research interests include spectral estimation, array signal processing, and their applications. Dr. Li is a Member of Sigma Xi and Phi Kappa Phi. She received the 1994 National Science Foundation Young Investigator Award and the 1996 Office of Naval Research Young Investigator Award. She was an Executive Committee Member of the 2002 International Conference on Acoustics, Speech, and Signal Processing, Orlando, Florida, May 2002. She has been an Associate Editor of the IEEE Transactions on Signal Processing since 1999 and an Associate Editor of the IEEE Signal Processing Magazine since 2003. She is presently a Member of the Signal Processing Theory and Methods (SPTM) Technical Committee of the IEEE Signal Processing Society.



# OFDM versus Single Carrier: A Realistic Multi-Antenna Comparison

## Jan Tubbax

*Interuniversity Micro-Electronics Center (IMEC), Kapeldreef 75, 3001 Leuven, Belgium  
Email: jan.tubbax@imec.be*

## Liesbet Van der Perre

*Interuniversity Micro-Electronics Center (IMEC), Kapeldreef 75, 3001 Leuven, Belgium  
Email: liesbet.vanderperre@imec.be*

## Marc Engels

*Interuniversity Micro-Electronics Center (IMEC), Kapeldreef 75, 3001 Leuven, Belgium  
Email: marc.engels@imec.be*

## Hugo De Man

*Interuniversity Micro-Electronics Center (IMEC), Kapeldreef 75, 3001 Leuven, Belgium  
Email: hugo.deman@imec.be*

## Marc Moonen

*Department of Electrical Engineering, Catholic University of Leuven (KUL), Kasteelpark Arenberg 10, 3001 Leuven, Belgium  
Email: marc.moonen@esat.kuleuven.ac.be*

*Received 20 June 2003; Revised 20 November 2003*

There is an ongoing discussion in the broadband wireless world about the respective benefits of orthogonal frequency division multiplexing (OFDM) and single carrier with frequency domain equalization (SC-FD). SC-FD allows for more relaxed front-end requirements, of which the power amplifier efficiency is very important for battery-driven terminals. OFDM, on the other hand, can yield improved BER performance at low complexity. Both schemes have extensions to multiple antennas to enhance the spectral efficiency and/or the link reliability. Moreover, both schemes have nonlinear versions using decision feedback equalization (DFE) to further improve performance of the linear equalizers. In this paper, we compare these high-performance OFDM and SC-FD schemes using multiple antennas and DFE, while also accounting for the power amplifier efficiency. To make a realistic comparison, we also consider most important digital imperfections, such as channel and noise estimation, transmit and receive filtering, clipping and quantization, as well as link layer impact. Our analysis shows that for frequency-selective channels the relative performance impact of the power amplifier is negligible compared to the frequency diversity impact. The higher frequency diversity exploitation of SC-FD allows it to outperform OFDM in most cases. Therefore, SC-FD is a suitable candidate for broadband wireless communication.

**Keywords and phrases:** single carrier, OFDM, multi-antenna, power amplifier, decision feedback equalization.

## 1. INTRODUCTION

Orthogonal frequency division multiplexing (OFDM) is a popular, standardized technique for broadband wireless systems: it is used for wireless LAN [1, 2], fixed broadband wireless access [3], digital video & audio broadcasting [4, 5] and so forth. OFDM can reach high spectral efficiency at low equalization complexity [6]. In recent years, single carrier with frequency domain equalization (SC-FD) has received

a lot of attention as an alternative technique for broadband wireless communications [7]. Studies [8, 9] show that SC-FD can allow for a more power efficient transmitter, which is a very important aspect for battery-operated mobile terminals.

Multiple antennas allow to increase the spectral efficiency and/or to improve the link reliability. Therefore, more and more systems, both in theory and in practice, make use of multiple antennas. Space division multiple access (SDMA) implements multiuser access spectrally efficiently

[10]. Moreover, [10] shows that decision feedback equalization (DFE) improves the performance of OFDM-SDMA, namely by applying the so-called per-carrier successive interference cancellation (OFDM-pcSIC).

Recently, a single-user SC-FD scheme with frequency domain processing and DFE (SC-FD-DFE) has been introduced [11, 12]. The time domain DFE allows to eliminate intersymbol interference (ISI) based on previous decisions and thus to improve performance over linear equalization. Moreover, it explicitly assumes the feedforward equalization in the frequency domain, which enables a low complexity solution.

These properties make SC-FD-DFE a suitable scheme for a broadband wireless uplink: it can achieve high spectral efficiency, while at the same time it implies a simple and efficient transmitter, which is a real plus for mobile terminals with limited battery and processing power. Moreover, it puts all the processing complexity at the receiver, that is, a base station, where typically more processing power is available.

The authors of [11, 12] derived the SC-FD-DFE for a single-user single-antenna (SISO) case. To compare OFDM-pcSIC fairly to SC-FD, we need to compare it to a multiuser SC-FD-DFE. Therefore, we extend this scheme to incorporate multiple users and multiple antennas (MIMO).

We compare the performance between the SC-FD-DFE and OFDM-pcSIC. Since the introduction of OFDM and SC-FD schemes, the advantages and disadvantages between them have been compared frequently [7, 8, 9, 11, 12, 13]. Recently, this comparison has gained more attention, since both schemes have been included in the IEEE 802.16a standard for fixed broadband wireless access [3].

Two very important points of comparison are the performance in a multipath environment and the impact of the power amplifier (PA). OFDM and SC-FD are designed for transmission over a frequency-selective channel and their cyclic extension is useful only over such a channel. On the other hand, the back-off of the power amplifier determines the power efficiency of the transmitter, a very important aspect for a wireless uplink as the terminal is likely to be battery powered.

Previous comparisons have not investigated the impact of the power amplifier on the bit error rate (BER) performance over a multipath channel. For example, [7, 11, 12, 13] only compare the ideal multipath performance and do not consider the impact of the power amplifier. Struhsaker and Griffin [9] only consider the spectral regrowth due to the power amplifier, but not the impact on the performance, while [8] considers the effect on the BER performance only for AWGN channels.

In this paper, we compare SC-FD and OFDM in a realistic multiuser scenario with DFE. In other words, we compare SC-FD-DFE and OFDM-pcSIC, taking into account the impact of the power amplifier and most important digital imperfections, such as channel and noise estimation, clipping and quantization, transmit and receive filtering. Moreover, we account for coding and retransmission enabling us to compare the useful throughput.

The paper is structured as follows. In Section 2, we briefly introduce OFDM and SC-FD as well as the multi-antenna

DFE concept for both schemes. In Section 3, we derive the multi-antenna SC-FD-DFE scheme and assess its performance. Section 4 evaluates the performance of both SC-FD-DFE and OFDM-pcSIC under realistic multi-antenna conditions. Finally, the conclusions obtained in this paper are presented in Section 5.

## 2. OFDM VERSUS SC-FD

In this section, we briefly compare the basic properties of OFDM and SC-FD and indicate the main differences and similarities. We introduce the decision feedback concept for both schemes.

### 2.1. OFDM

The basic idea of OFDM [6] transmission is to divide the available bandwidth into  $N$  subcarriers. If the number of subcarriers is large enough, the bandwidth per subcarrier is narrow compared to the coherence bandwidth of the channel. Therefore, each subcarrier experiences approximately flat fading. A spectrally and computationally efficient method to modulate the data on the frequency domain subcarriers is by means of an IFFT.

The addition of a cyclic prefix ensures that the channel always appears cyclic and thus the linear convolution with the channel can be considered a circular convolution. This guarantees that the received signal can be equalized by means of a single-tap equalizer per subcarrier. This operation is performed in the frequency domain, thus after the received signal is passed through an FFT.

Uncoded OFDM loses all frequency diversity inherent in the channel: a dip in the channel erases the information data on the subcarriers affected by the dip and this information cannot be recovered from the other carriers. This mechanism results in a poor BER performance. Adding sufficiently strong coding spreads the information over multiple subcarriers. This recovers frequency diversity and improves the BER performance.

### 2.2. SC-FD

SC-FD [7] transmits the data in the time domain. The cyclic extension is added before transmission which ensures the channel appears cyclic at the receiver. This again allows to have the same simple one-tap equalizer in the frequency domain. The decisions have to be taken in the time domain, so after the equalization an IFFT is needed.

Compared to OFDM, SC-FD uses the same building blocks, but with the IFFT moved from the transmitter to the receiver. This also means single carrier transmits the data in the time domain, whereas OFDM puts the data in the frequency domain. This has a very important consequence: the information of each SC-FD symbol is spread out over the complete frequency band. This means that dips in the channel do not wipe out complete symbols, because the information of each symbol can be recovered from the other carriers. Therefore, SC-FD exploits the frequency diversity in the channel and thus has a better uncoded performance than OFDM [7].

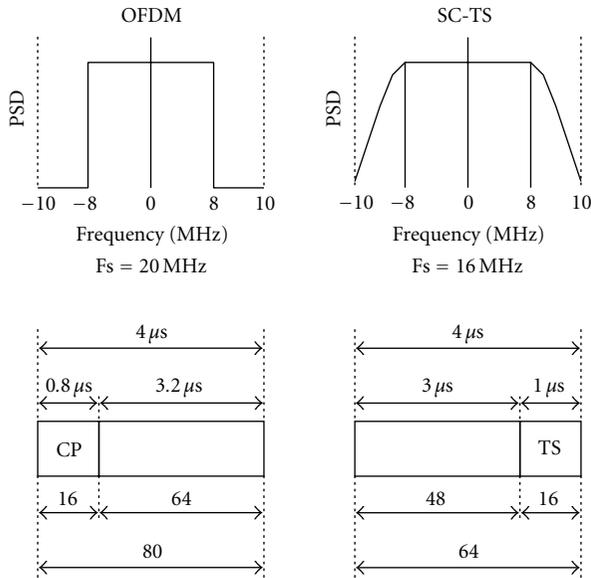


FIGURE 1: OFDM with cyclic prefix versus single carrier with TS.

To obtain a scalar (and thus low complexity) feedforward equalization section in the frequency domain, we need to insert a cyclic extension at the transmitter. We can obtain this by either inserting a cyclic prefix (as in OFDM) or a training sequence (TS). However, only the TS approach yields the same spectral efficiency as OFDM (as indicated in Figure 1) which is essential for a fair comparison: both schemes transmit 48 data symbols in a 4 microseconds period, compliant to the IEEE 802.11a standard. Note that the two schemes require a slightly different sampling rate ( $F_s$ ) to use the same bandwidth. Moreover, the edges of the spectrum are filled with zero carriers for OFDM, while single carrier needs pulse shaping (see Section 4.2).

The TS offers some additional advantages: because it is a known sequence, it can be used for synchronization, tracking, and training [14] and to start up the DFE [12].

Coding improves the SC-FD performance, whereas for OFDM coding is needed to exploit the frequency diversity and improve the performance.

### 2.3. Decision feedback equalization

The knowledge of (part of the) data helps to reconstruct and thus to eliminate (part of) the interference on other data. In turn, this (partial) elimination of interference improves performance.

In this section, we aim to exploit DFE to eliminate interference caused by multipath or by other users.

#### 2.3.1. For OFDM

In single-user (uncoded) OFDM, the data on one subcarrier is not related to the data on any other subcarrier. Since the multipath channel effect is also separated per carrier, DFE reduces to the linear MMSE solution.

In a multiuser scenario, each subcarrier contains a superposition of data from different users. In this case, the knowl-

edge of one user's data (and channel coefficients) on a given subcarrier allows to eliminate this user's interference when making decisions on other users' data on that particular carrier. This basic idea is exploited in the so-called OFDM-SDMA with per-carrier successive-interference cancellation (pcSIC) [10].

In SDMA,  $U$  user terminals communicate with a base station with  $A$  receive antennas. The  $A$  received signals are simultaneously combined (equalized) into the  $U$  user streams.

Vandenameele et al. [10] describe a nonlinear OFDM SDMA equalizer, (OFDM-pcSIC). On a per-carrier basis, the users are sorted according to decreasing power. Per carrier, the most powerful user is detected using MMSE equalization. Using the decision and the channel estimates, the signal transmitted by the strongest user on the given carrier is reconstructed and subtracted from the received signal. This eliminates the interference caused by this user on the given carrier and thus improves performance when detecting the other users. This process is repeated for the second strongest user and so on until all users on all carriers are detected.

pcSIC can considerably improve the MMSE performance with a reasonable complexity increase, for example, [10] presents a case study with  $U = A = 4$  with a 5 dB improvement for a 20% increase in complexity. This clearly illustrates the attractive features of OFDM-pcSIC for multiuser wireless communication.

#### 2.3.2. For SC-FD

For SC-FD, data symbols are transmitted in the time domain. The multipath channel spreads a data symbol across a number of subsequent symbols, equal to the order of the channel. Consequently, knowledge of previous symbols (and channel information) allows to eliminate some interference on subsequent data symbols and thus to improve performance. This idea is exploited in SC-FD-DFE in [11, 12], which describe a DFE method with similar complexity as OFDM pcSIC.

To make a fair comparison with OFDM pcSIC, we need to extend this single-user SC-FD-DFE to the multiuser case: each user data symbol does not only experience interference from its own preceding symbols from that user, but from other users as well. Therefore, previous decisions from all users can be used to cancel interference and to improve performance. We derive the necessary equations for such a scheme in the following section.

## 3. MULTI-ANTENNA SINGLE CARRIER WITH DFE

### 3.1. Scheme

In [11, 12] SC-FD-DFE is presented as in Figure 2, with a feedforward part in the frequency domain and a feedback section with  $L_{dfc}$  taps in the time domain.

Throughout the rest of the paper, matrices are denoted by bold capital letters, vectors have a bold normal font and scalars have a normal font. Frequency domain signals are indicated by a tilde.  $\mathbf{a}$  represents the modulated data,  $\mathbf{H}$  is the channel matrix (it reduces to a vector for a SISO channel),  $\mathbf{n}$  is the noise,  $\mathbf{r}$  is the received time domain signal, while  $\tilde{\mathbf{r}}$  is

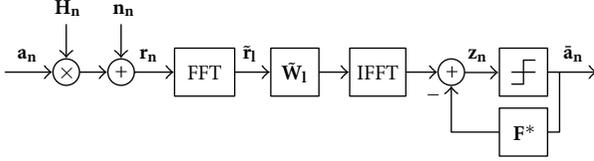


FIGURE 2: Single-carrier scheme with FD-DFE.

the received signal in the frequency domain.  $\mathbf{z}$  is the equalized signal just before the slicer,  $\hat{\mathbf{a}}$  are the decisions and  $\mathbf{F}$  the feedback coefficients matrix (it reduces to a vector in the single-user case).

Since the TS is a known sequence, it can be used to start up the feedback process. Therefore, SC-TS suits the DFE structure better than the cyclic prefix approach.

We now derive the SC-FD-DFE equations for a scenario with  $U$  users and  $A$  receive antennas. The variables are indicated in Figure 2.

### 3.2. Derivation of the coefficients

The equalizer output just before the slicer at time  $n$  is

$$\mathbf{z}_n = \frac{1}{N} \sum_{l=0}^{N-1} \tilde{\mathbf{W}}_1 \cdot \tilde{\mathbf{r}}_1 \cdot \exp\left(j2\pi \frac{nl}{N}\right) - \mathbf{F}^* \cdot \hat{\mathbf{a}}_{n-k}. \quad (1)$$

For each frequency tap  $l$ ,  $\tilde{\mathbf{r}}_1$  is an  $A \times 1$  vector with the received signal in the frequency domain and  $\tilde{\mathbf{W}}_1$  is a  $U \times A$  equalization matrix representing the feedforward part;  $\mathbf{F}$  is a  $U \times (U \cdot L_{\text{dfe}})$  matrix containing the feedback coefficients;  $\hat{\mathbf{a}}_{n-k}$  is a  $(U \cdot L_{\text{dfe}}) \times 1$  vector with the data decisions selected for feedback, for all users.  $N$  is the number of time domain symbols per SC-FD block.

The derivation of the feedforward coefficients  $\tilde{\mathbf{W}}_1$  and the feedback coefficients  $\mathbf{F}$  extends the analysis by [11] to multiple antennas and multiple users and can be found in the appendix. The solution for the optimal feedforward coefficients  $\tilde{\mathbf{W}}_1$  is for every  $l$

$$\tilde{\mathbf{W}}_1 = \tilde{\mathbf{F}}_1 \tilde{\mathbf{H}}_1^* (\tilde{\mathbf{H}}_1 \tilde{\mathbf{H}}_1^* + \sigma^2 \mathbf{I}_A)^{-1} \quad (2)$$

with  $\tilde{\mathbf{H}}_1$  an  $A \times U$  matrix containing the coefficients for the channels from the  $U$  users to the  $A$  receive antennas at carrier  $l$ ,  $\mathbf{I}_A$  the  $A \times A$  identity matrix and  $\sigma^2$  the noise power and

$$\tilde{\mathbf{F}}_1 = \mathbf{I}_U + \mathbf{F}^* \cdot \mathbf{X}_1^*,$$

$$\mathbf{X}_1 = \begin{bmatrix} \exp\left(j2\pi \frac{l\mathbf{k}_1}{N}\right) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \exp\left(j2\pi \frac{l\mathbf{k}_U}{N}\right) \end{bmatrix}, \quad (3)$$

where  $\mathbf{k}_i$  is a  $1 \times L_{\text{dfe}}$  vector containing the selected indices of the feedback taps for user  $i$  and  $\mathbf{I}_U$  the  $U \times U$  identity matrix.

The optimal feedback coefficients  $\mathbf{F}$  are

$$\mathbf{F} = -\mathbf{T}^{-1} \cdot \mathbf{t} \quad (4)$$

with

$$\mathbf{t} = \frac{1}{N} \sum_{l=0}^{N-1} \mathbf{X}_1^* (\tilde{\mathbf{H}}_1^* \tilde{\mathbf{H}}_1 + \sigma^2 \mathbf{I}_U)^{-1},$$

$$\mathbf{T} = \frac{1}{N} \sum_{l=0}^{N-1} \mathbf{X}_1^* (\tilde{\mathbf{H}}_1^* \tilde{\mathbf{H}}_1 + \sigma^2 \mathbf{I}_U)^{-1} \mathbf{X}_1, \quad (5)$$

$\tilde{\mathbf{W}}_1$  is a  $U \times A$  matrix transforming the  $A$  antenna streams into  $U$  user streams.  $\mathbf{F}$  is of size  $U \times (U \cdot L_{\text{dfe}})$  and contains for each user the optimal feedback coefficients to feed back the  $U \cdot L_{\text{dfe}}$  selected decisions,  $L_{\text{dfe}}$  from each user.

### Tap selection

By selecting the number of feedback taps, complexity can be exchanged for performance. With  $L_{\text{dfe}} = 0$ , the SC-FD-DFE scheme reduces to the SC-FD MMSE solution: (2) reduces to the linear equalization coefficients and there is no feedback part. A physically meaningful maximum of  $L_{\text{dfe}} = L - 1$  feedback taps (with  $L$  the order of the channel) can be selected, yielding the largest performance improvement. In this case,  $\mathbf{X}_1$  and  $\mathbf{F}$  are of maximum size and thus this operating point also introduces most processing power.

In the SISO case, the indices of the feedback taps are selected according to the strength of the channel response: the  $L_{\text{dfe}}$  taps that contain most power are used. As we assume a channel of  $L$  equal (average) power taps (see the following section), all channel taps can yield a significant improvement in the feedback process. Therefore, we set  $L_{\text{dfe}} = L - 1$ , which we assume to be known. The tap selection for the multiuser case is done for each user separately. In all multi-antenna simulations, we set the number of feedback taps  $L_{\text{dfe}}$  equal to  $L - 1$  for each user to provide the maximum performance gain. This means the feedback taps to be used are simply the indices from 1 to  $L_{\text{dfe}}$ .

### 3.3. Performance

We apply the multi-antenna SC-FD-DFE to two cases: first, we simulate an uplink SDMA scenario. Secondly, we investigate uplink receive diversity. The simulations presented in this section are performed for uncoded BPSK transmission with perfect channel knowledge, to illustrate basic DFE properties. More real-life results are presented in Section 4.

For all simulations in this paper, we apply a multipath channel with  $L = 4$  Rayleigh fading taps; the taps are assumed to be independent and of equal average power. The channel is considered fixed for the duration of the data burst following the preamble. For the multiple-antenna simulations, we assume uncorrelated channel realizations.

#### 3.3.1. SDMA

We present the results for which each of the  $U$  mobile terminals has one transmit antenna and the base station has a number of receive antennas  $A$  equal to the number of users  $U$ . At the base station, the user streams are received at the  $A$  antennas. This multiuser-multi-antenna scenario has a double effect on performance. On one hand, the  $A$  receive antennas increase the receive diversity, which improves the performance. On the other hand, the  $U$  user streams cause interuser

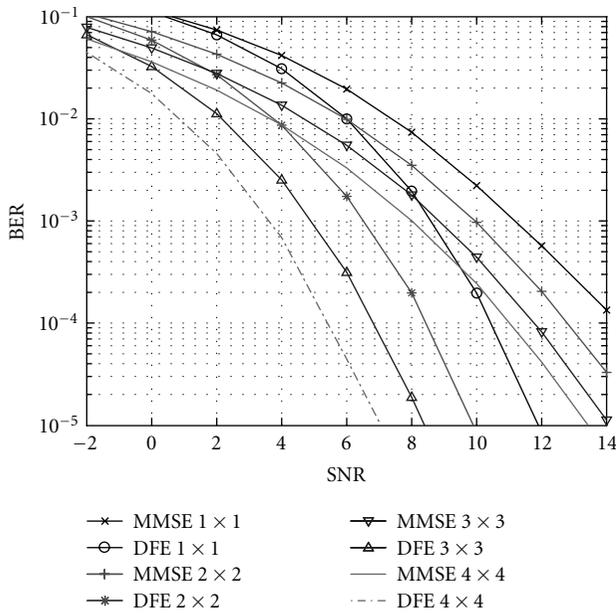


FIGURE 3: Performance of single-carrier SDMA: linear MMSE versus MMSE-DFE.

Interference (IUI) which degrades the performance. Vandenameele et al. [10] show that overall the performance slightly improves as the number of users (equal to the number of antennas) increases for the linear MMSE receiver.

By applying DFE,  $L_{dfc}$  decisions on the user's preceding symbols are used to eliminate the ISI, while  $(U - 1) \cdot L_{dfc}$  decisions on preceding symbols of the other users are used to eliminate the IUI. This nonlinear approach improves the performance over the linear SDMA. Figure 3 shows the performance curves comparing linear SDMA to DFE for 1 up to 4 users.

We observe a 3 to 4 dB gain at a BER of  $10^{-3}$  for the SC-FD-DFE over the conventional linear MMSE. The DFE advantage over the linear SDMA slightly increases as the number of users increases: more users mean more IUI for both schemes, but also mean more decisions available for the DFE to eliminate part of this interference. Therefore, as the number of users increases, so does the interference the DFE can eliminate relative to the linear SDMA and the DFE performance advantage.

### 3.3.2. Receive diversity

In this case, one mobile terminal in the system is active and the base station has  $A$  receive antennas. This means the receive diversity increases with the number of antennas, but there is no IUI, only ISI caused by the frequency selective channel. The DFE can eliminate part of this ISI and thus creates a frequency diversity advantage. However, this frequency diversity advantage becomes relatively less important as the receive diversity increases.

Figure 4 shows that the DFE gain relative to the linear MMSE solution decreases if the number of receive antennas is increased.

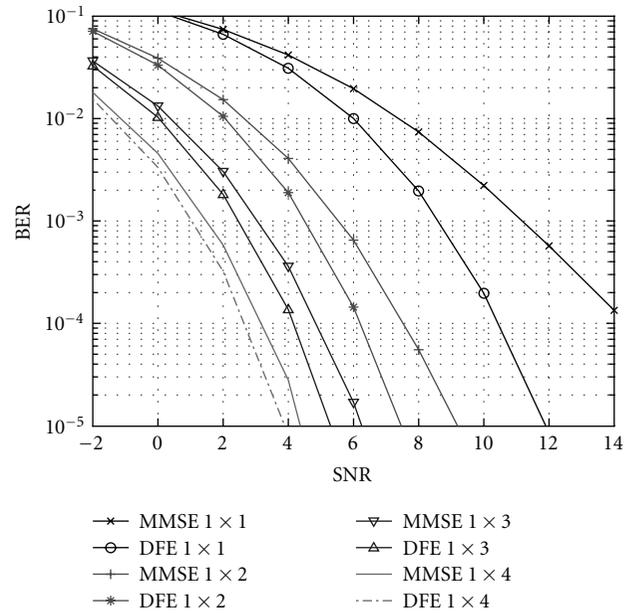


FIGURE 4: Performance of single carrier with receive diversity: linear MMSE versus MMSE-DFE.

### 3.3.3. Summary

The results from the previous sections show that DFE improves the performance over the linear equalizer. Therefore, DFE can offer a targeted performance at lower signal-to-noise ratio (SNR) and thus at lower transmit power. This transmit power reduction comes at the cost of added receiver complexity. For uplink communication, the receiver is a base station and the transmitter is a user terminal. This means DFE allows to reduce terminal power consumption (which is crucial for battery-driven devices) by adding complexity at the base station (where the extra complexity is more easily accommodated). This clearly shows that DFE is a very useful technique for SC-FD uplink communication.

## 4. MULTI-ANTENNA SC-FD-DFE VERSUS OFDM-PCSIC

Section 3 illustrates the performance of SC-FD-DFE under ideal multipath conditions: an ideal analog front end is assumed, with a linear power amplifier; the channel and noise characteristics are assumed to be perfectly known; digital imperfections, such as clipping and quantization as well as transmit and receive filtering, are not taken into account. OFDM and SC-FD react differently to the deviations from these ideal conditions, so they need to be considered in order to make a fair comparison between both schemes.

We discuss the modeling and impact of the power amplifier, channel and noise estimation, clipping and quantization, and transmit and receive filtering. When presenting the performance results, we take the link layer efficiency into account as well.

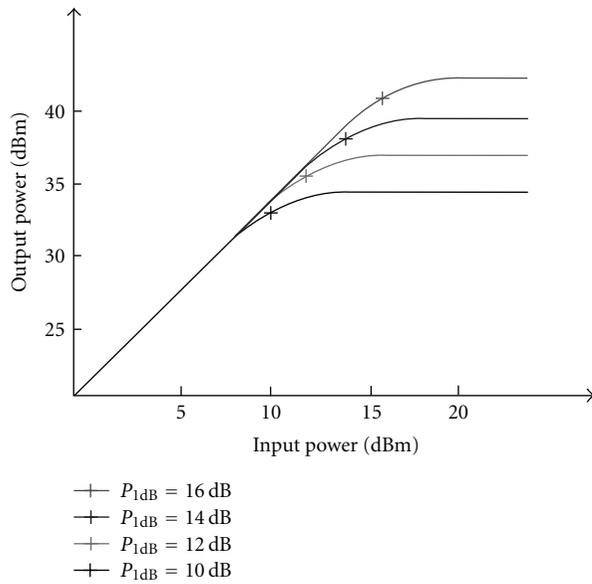


FIGURE 5: Transfer function of a class A power amplifier.

#### 4.1. Power amplifier

For nonconstant envelope signals, a linear power amplifier is needed. We assume a class A power amplifier with back-off because of its linearity. The back-off determines the power consumption of the power amplifier and also its linear dynamic range. Since the linear dynamic range directly relates to the distortion, the back-off also determines the BER.

The linearity of the power amplifier is quantified by the 1-dB-compression point  $P_{1\text{dB}}$ , defined as the input power at which the nonlinearity lowers the output power by 1 dB compared to the ideal amplifier (Figure 5).

Since we want to perform baseband simulations because of simulation speed, we use the baseband representation of the power amplifier. The transfer function of a class A power amplifier with linear amplification  $G$  and a cubic nonlinearity is [15]

$$\mathbf{y} = \mathbf{x} \cdot G \cdot \left(1 - \alpha \frac{3}{4} |\mathbf{x}|^2\right), \quad (6)$$

with  $\mathbf{x}$  the input baseband representation of the signal and  $\mathbf{y}$  the output.

The coefficient  $\alpha$  can be expressed as a function of  $P_{1\text{dB}}$  as

$$\alpha = \frac{4}{3(1 - 10^{-1/20})P_{1\text{dB}}^2}. \quad (7)$$

In our setup, we set the average input power  $P_{\text{in}} = 6$  dBm; the linear gain of the power amplifier is 23 dB, such that we operate at 29 dBm average output power, which is a specified maximum average output power for the 5 GHz band [1]. The higher the  $P_{1\text{dB}}$  compression point, the further the signal is separated from the distortion area of the power amplifier transfer characteristic. The smaller the distortion added by the power amplifier, the smaller the BER performance degradation is. However, the larger the back-off between  $P_{\text{in}}$  and  $P_{1\text{dB}}$  the smaller the power amplifier efficiency, as can be seen

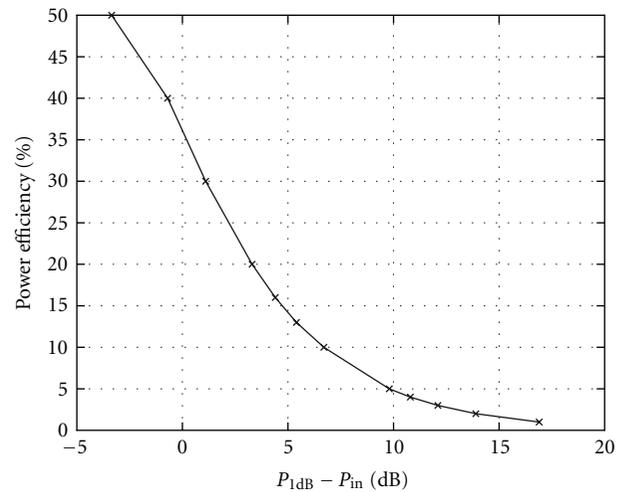


FIGURE 6: Power efficiency of a class A power amplifier as a function of the back-off  $P_{1\text{dB}} - P_{\text{in}}$ .

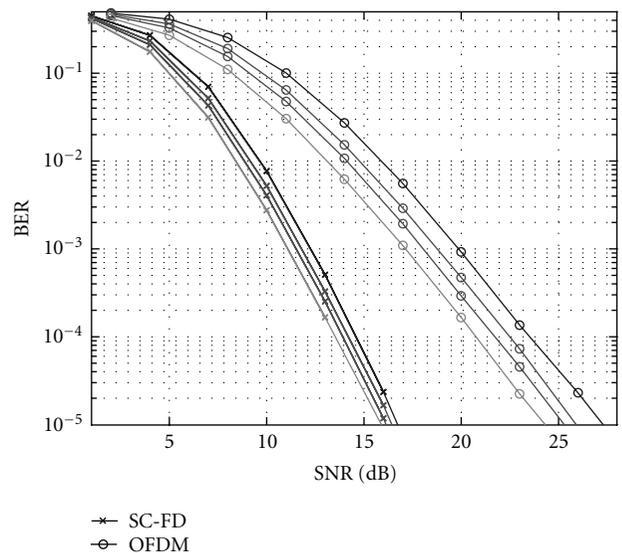


FIGURE 7: The impact of PA back-off  $P_{1\text{dB}} = [\infty 6.4 4.4 2.4]$  dBm on SC-FD and OFDM in multipath with perfect channel knowledge and coding ( $R = 3/4$ ).

in Figure 6. A class A power amplifier has a theoretical maximum efficiency of 50%. This efficiency drops rapidly with increasing back-off.

OFDM has a large dynamic range compared to SC-FD [9]. Therefore, the  $P_{1\text{dB}} - P_{\text{in}}$  back-off needs to be larger for OFDM to accommodate the signal in the linear range of the power amplifier transfer function. Theory and simulations indeed show that in an AWGN channel the performance advantage of SC-FD over OFDM increases as the PA back-off decreases.

However, the comparison in a multipath environment yields different conclusions. Since uncoded OFDM only reaches frequency diversity 1, the additional degradation caused by the power amplifier nonlinearity is much smaller than over an AWGN channel. This is shown in Figure 7 for

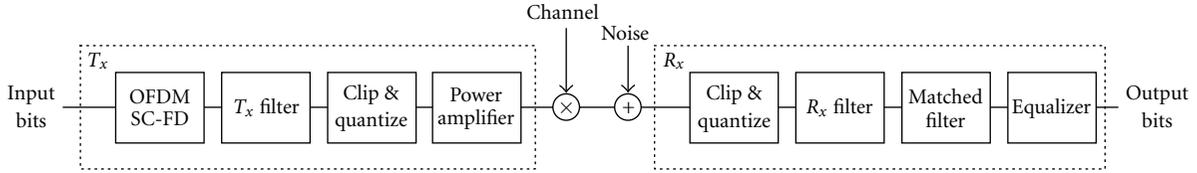


FIGURE 8: The setup for the simulation.

uncoded BPSK transmission over a multipath channel: the OFDM performance curves for decreasing back-off are situated quite close together as are the SC-FD curves. The dominating degradation for OFDM compared to SC-FD is caused by the lack of frequency diversity: with infinite back-off, SC-FD has an 8 dB performance advantage over OFDM at a BER of  $10^{-4}$ . The additional gain for SC-FD because of the power amplifier back-off of [6.4 4.4 2.4] dBm is only [0.6 1 1.6] dB. This clearly illustrates that the exploitation of the frequency diversity is dominant over the power amplifier back-off impact.

This is a noteworthy result, especially since BPSK is the modulation scheme which yields the largest power amplifier advantage for SC-FD over OFDM.

Coding helps to overcome the lack of frequency diversity, bringing the OFDM and SC-FD curves closer together. However, coding does not change the impact of the power amplifier back-off. This means that even on coded OFDM, the impact of the power amplifier back-off is comparable to that of coded SC-FD. The difference in performance between OFDM and SC-FD remains dominated by the difference in frequency diversity; this, in turn, is determined by the code rate. For high code rates (and uncoded transmission) SC-FD outperforms OFDM; for lower code rates, the performance becomes comparable.

These results only indicate the impact of the power amplifier back-off in a single-user multipath scenario with perfect channel knowledge. In the following sections, we introduce other imperfections which need to be taken into account, together with the multi-antenna aspect.

#### 4.2. Digital imperfections

To make a realistic comparison, we include transmit and receive filtering, channel and noise estimation, and clipping and quantization in our simulation model as shown in Figure 8.

##### Transmit and receive filtering

The OFDM and SC-FD symbols are not transmitted as such, but they need to be filtered to limit the out-of-band radiation. Therefore, we apply a square root raised cosine (SRRC) filter as a transmit pulse shaping filter  $\mathbf{g}_T$  and a receive filter  $\mathbf{g}_R$  with a rolloff  $\alpha = 0.2$ , a delay of 5 taps, and an oversampling by 4. Moreover, at the receiver, an optimal channel matched filter is added in the frequency domain before the equalization.

##### Channel and noise estimation

Until now, we assumed perfect channel knowledge. In practice, the channel needs to be estimated and quite often a noise estimate is required as well. We use the estimation methods as described in [16].

The channel estimation is based on the C sequence (BPSK symbol  $\tilde{\mathbf{t}}$ ) of the OFDM-based IEEE 802.11a standard. The following frequency domain estimation can also be directly applied to SC-FD, ensuring a fair comparison. In practice, a time domain sequence will be used for SC-FD channel estimation, which is designed to have a frequency response which is as flat as possible, but that is beyond the scope of this paper.

For a channel with frequency response  $\tilde{\mathbf{c}}$  and noise  $\mathbf{n}$ , we obtain the channel estimate  $\tilde{\mathbf{h}}$  as

$$\begin{aligned} \tilde{\mathbf{h}} &= \tilde{\mathbf{t}} \star (\tilde{\mathbf{g}}_R \star \tilde{\mathbf{c}} \star (\tilde{\mathbf{g}}_T \star \tilde{\mathbf{t}} + \text{FFT}\{\mathbf{n}\})) \\ &= \tilde{\mathbf{g}}_R \star \tilde{\mathbf{c}} \star \tilde{\mathbf{g}}_T + \tilde{\mathbf{n}}', \end{aligned} \quad (8)$$

with  $\star$  the elementwise vector multiplication,  $\tilde{\mathbf{n}}' = \tilde{\mathbf{g}}_R \star \tilde{\mathbf{t}} \star \text{FFT}\{\mathbf{n}\}$  and  $\tilde{\mathbf{g}}_T, \tilde{\mathbf{g}}_R$  the frequency responses of the transmit pulse shaping filter  $\mathbf{g}_T$  and the receive filter  $\mathbf{g}_R$  respectively.

Since the channel as well as the transmit, and receive filters have a limited number of taps in the time domain (assuming the total number of taps to be  $L$ ), we know that all the power in the taps  $\geq L + 1$  can be attributed to noise. This allows us to remove part of the noise on the time domain channel estimate. This has two consequences: first, it allows to have an improved channel estimate, since we remove part of the noise. Second, the power of the noise which is removed is used to obtain a noise estimate. The analysis of these estimates can be found in [16].

In case of multiple transmit and/or receive antennas, we repeat the above process for each antenna pair separately. Vandenameele et al. [10] show that more performant multi-antenna channel estimation schemes exist. However, we do not need the most performant scheme; we just want to make a realistic and fair comparison between the OFDM and SC-FD schemes.

##### Clipping and quantization

OFDM is clipped and quantized to limit the signal's dynamic range or the so-called peak-to-average power ratio (PAPR). This is beneficial since the limited dynamic range enables a smaller back-off. Côme et al. [15] show clipping at  $4\sigma$  and a quantization of 8 bits to be a good solution for a realistic OFDM system. SC-FD needs to be clipped and quantized as well, but because of its limited dynamic range the impact is smaller.

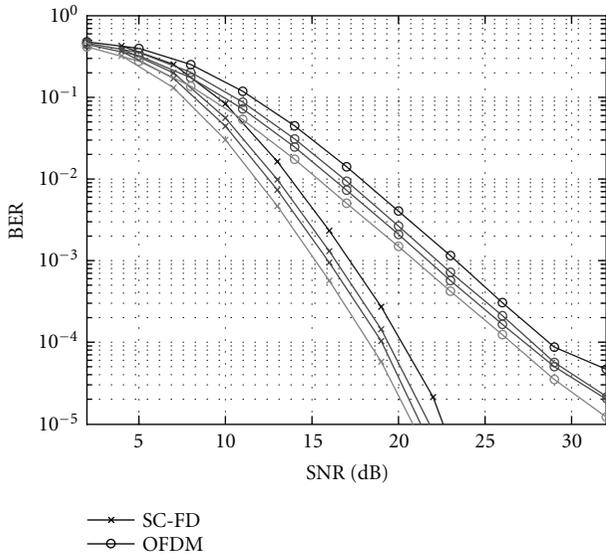


FIGURE 9: The impact of PA back-off  $P_{1\text{dB}} = [\infty 6.4 4.4 2.4]$  dBm on SC-FD and OFDM in multipath with coding ( $R = 3/4$ ), channel/noise estimation, clipping ( $4\sigma$ ) and quantizing, (8 bits) and SRRC filtering ( $\alpha = 0.25$ ).

### 4.3. Performance

We present the simulation results of the OFDM versus SC-FD comparison for a WLAN case study: we use the system parameters for OFDM as described in the IEEE standard [1] and choose the equivalent SC-FD such that the comparison is fair. For both schemes, the results for coded BPSK transmission ( $R = 3/4$ ) are presented.

We first show the SISO results and extend them to the multipleantenna multiple-user case: we compare OFDM to SC-FD for SDMA and receive diversity, both schemes in the linear MMSE and the nonlinear DFE version.

#### 4.3.1. Single user

The single-user single-antenna performance results also apply to a multiuser scenario if the multiple access to a single-antenna base station is provided by TDMA or FDMA.

As shown in Figure 9, the impact of the PA for  $P_{1\text{dB}} = [6.4 4.4 2.4]$  dBm is a  $[0.75 1.1 1.9]$  dB degradation compared to the reference case (with infinite back-off) for SC-FD and  $[0.75 1.2 2.2]$  dB for OFDM at a BER of  $10^{-4}$ . This means that the additional advantage for SC-FD to a decreasing back-off  $P_{1\text{dB}} = [6.4 4.4 2.4]$  dBm is only  $[0 0.1 0.3]$  dB, while the difference in frequency diversity is still 8 dB.

This again clearly illustrates that in a realistic system the advantage of SC-FD over OFDM comes almost completely due to its effective use of the frequency diversity present in the channel. The impact of the PA power efficiency is negligible in the comparison of OFDM and SC-FD in a multipath scenario.

Another way of quantifying the difference in performance is not based on (SNR, BER) curves, but on  $(P_{T_x}, \text{Goodput})$  curves.

Goodput indicates the actual throughput at MAC level. It starts from the raw physical throughput (based on the con-

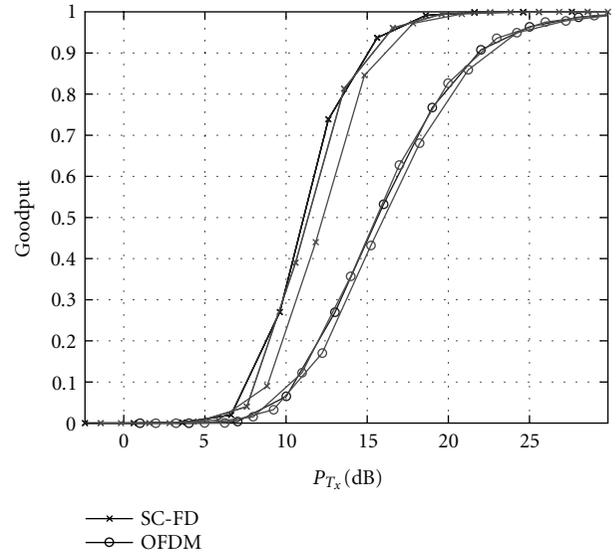


FIGURE 10: The goodput as a function of the total consumed transmitter power  $P_{T_x}$  and the impact of PA back-off  $P_{1\text{dB}} = [6.4 4.4 2.4]$  dBm on SC-FD and OFDM in multipath with coding ( $R = 3/4$ ), channel/noise estimation, clipping ( $4\sigma$ ) and quantizing (8 bits), and SRRC filtering ( $\alpha = 0.25$ ).

stellation size and the sampling time) and takes all overhead up to MAC level into account; it accounts for coding, the number of actual data symbols, the cyclic extension, and protocol overhead. Since we have chosen our system parameters to make a fair comparison between OFDM and SC-FD, both schemes can achieve the same maximum effective throughput or goodput  $R_{\text{max}}$ . Therefore, we can normalize  $R_{\text{max}} = 1$  without loss of generality. Finally, erroneously received packets need to be retransmitted:  $(1-\text{PER})$  accounts for the loss of actual throughput because of retransmissions. This leads to

$$\begin{aligned} \text{Goodput} &= R_{\text{max}} \cdot (1 - \text{PER}) \\ &= 1 - \text{PER}. \end{aligned} \quad (9)$$

The same relative measure can be derived for the total power consumed at the transmitter. If we assume the path loss, noise power, noise figure, and implementation loss are identical for OFDM and SC-FD (which is necessary for a fair comparison), we define the relative consumed power at the transmitter as

$$P_{T_x} = \frac{\text{SNR}}{\mu(P_{1\text{dB}})}. \quad (10)$$

This means the total consumed power at the transmitter is proportional to the SNR (since higher SNR means more transmitted power) and inversely proportional to the PA efficiency  $\mu(P_{1\text{dB}})$ , which is determined by the back-off from the 1 dB-compression point  $P_{1\text{dB}}$  (Figure 6). We assume the power consumption at the transmitter is largely determined by the PA power consumption, thus by the transmitted power and the PA efficiency.

The (BER/PER, SNR) curves as in Figure 9 can be transformed into (Goodput,  $P_{T_x}$ ) curves as in Figure 10 through

TABLE 1: SC-FD advantage (dB) over OFDM for MMSE SDMA as a function of the number of antennas (A).

Number of antennas	1	2	3	4
Goodput = 0.90	6.8	4.8	3.7	2.9
Goodput = 0.95	7.2	5.5	4.5	3.6

TABLE 2: SC-FD advantage (dB) over OFDM for DFE SDMA as a function of the number of antennas (A).

Number of antennas	1	2	3	4
Goodput = 0.90	7.2	1.4	0.8	0.6
Goodput = 0.95	8.0	1.5	0.7	0.5

(9) and (10). These (Goodput,  $P_{T_x}$ ) curves consist of three parts. Below a certain threshold (in Figure 10 around 10 dB) no useful transmission is possible: the SNR is too low, (almost) all packets are erroneously received and an increase in transmission power results in (virtually) no increase in goodput. Above the threshold we observe a ramp-up region: as the SNR increases, packets start to get received correctly, the number of retransmissions decreases and the goodput increases; an increase in transmit power also increases the effective throughput. Beyond another threshold, saturation starts occurring: the PER is very low and decreasing it further (by increasing the transmit power) has a negligible effect on the goodput, because of the  $(1 - \text{PER})$  in (9).

If we target 90% of the maximum achievable goodput, then SC-FD can deliver this at 6.8 dB less consumed transmitter power than OFDM. The back-off  $P_{1\text{dB}} - P_{\text{in}}$  has a double influence on the (goodput,  $P_{T_x}$ ) curves. First, for decreasing back-off, the additional distortion increases and thus the BER/PER increases. This decreases the goodput through (9). On the other hand, a lower back-off results in a higher PA efficiency  $\mu$  (as shown in Figure 6) and a higher efficiency in a lower consumed transmitter power  $P_{T_x}$  through (10). Figure 10 shows the combined effect has a small impact on the effective throughput, compared to the frequency diversity impact.

The larger the goodput we target, the larger the SC-FD advantage becomes: larger goodput requires a smaller PER. Since SC-FD exploits the frequency diversity better than OFDM, their PER/BER curves diverge for lower PER/BER as in Figure 9. Therefore, also their goodput curves diverge for large goodputs. For a goodput of 95% of the maximum, SC-FD is 7.2 dB more efficient.

### 4.3.2. Multiuser SDMA

In this scenario, we consider  $U$  users (each with one antenna) transmitting to a base station with  $A$  receive antennas, equal to the number of users. As the number of receive antennas and the number of users increase, so does the receive diversity and the IUI for both schemes. The receive diversity helps OFDM to overcome the lack of frequency diversity. Therefore, as the number of antennas goes up, the relative diversity difference between OFDM and SC-FD becomes smaller

TABLE 3: SC-FD advantage (dB) over OFDM for MMSE receive diversity as a function of the number of antennas (A).

Number of antennas	1	2	3	4
Goodput = 0.90	6.8	2.6	1.8	1.5
Goodput = 0.95	7.2	3.0	2.1	1.7

TABLE 4: SC-FD advantage (dB) over OFDM for DFE receive diversity as a function of the number of antennas (A).

Number of antennas	1	2	3	4
Goodput = 0.90	7.2	2.9	2.1	1.6
Goodput = 0.95	8.0	3.3	2.3	1.8

and their performances converge. Table 1 shows that the SC-FD advantage gradually decreases as the number of users increases. The number in the following tables are given for  $P_{1\text{dB}} = 4.4$  dBm; as shown in Figure 10, the impact of  $P_{1\text{dB}}$  on the performance comparison is negligible.

Also in a multiuser scenario, the SC-FD advantage increases if a larger goodput is targeted.

DFE allows to eliminate interference. In case of OFDM, pcSIC allows to eliminate precursor and postcursor ISI, whereas in case of SC-FD, SC-FD-DFE only eliminates postcursor ISI [12, 17]. This means OFDM possibly can eliminate the IUI completely, while SC-FD can only partly. Therefore, the SC-FD advantage over OFDM is smaller for DFE (Table 2) than for MMSE (Table 1).

### 4.3.3. Multi-antenna receive diversity

In this scenario, we consider one active user with one antenna while the base station has  $A$  receive antennas.

As the number of receive antennas increases, the receive diversity rapidly increases the performance of both schemes, since there is no extra IUI to counter the diversity benefit. Since the receive diversity also helps to compensate the lack of frequency diversity, both OFDM and SC-FD converge to the same performance. Simulations show that the SC-FD advantage over OFDM, rapidly decreases with increasing number of receive antennas.

In the single-user case, SC-FD performs the DFE through the SC-FD-DFE as described in [11, 12], while for single-user OFDM, the pcSIC algorithm reduces to linear equalization. Therefore, the single-user SC-FD advantage over OFDM increases for the DFE (Table 4) compared to MMSE (Table 3).

Again we remark that the SC-FD advantage increases if larger goodputs are targeted.

## 4.4. Summary of the results

The comparison between OFDM and SC-FD can be summarized as follows. The comparison is dominated by the frequency diversity rather than the power amplifier impact. SISO SC-FD outperforms OFDM by 4–5 dB in our case study, because of its higher frequency diversity exploitation. Adding multiple antennas increases the receive diversity for both schemes and thus decreases the relative diversity gap.

This decreases the SC-FD advantage over OFDM until the linear MMSE performances converge for a large number of antennas. In single-user scenarios (such as SISO and the receive diversity scenario), DFE further increases the SC-FD advantage, while for multi-user SDMA with DFE, the SC-FD advantage decreases with increasing number of users, because of OFDM-pcSIC's better IUI elimination. Targeting a larger goodput always increases the SC-FD advantage over OFDM.

#### 4.5. Remarks

##### Power amplifier

Apart from impact on the BER/PER performance or goodput, we should note that other considerations have to be made. The power amplifier nonlinearity also determines the amount of out-of-band radiation for which specifications exist. This has been studied in [9]. The authors indicate OFDM is indeed more sensitive to PA impact as far as out-of-band radiation is concerned. Therefore, a large back-off might still be required for OFDM, if not to assure link performance, then at least to limit out-of-band radiation.

##### Code rate

The difference in performance between OFDM and SC-FD depends also on the code rate  $R$  and the constellation size. In this paper, we have taken  $R = 3/4$  as specified in both the IEEE802.11a and HIPERLAN-II standards for BPSK transmission. The code  $R = 3/4$  is a frequently used code rate, making a tradeoff between code performance and code overhead. For code rates larger than  $R = 3/4$ , the frequency diversity advantage of SC-FD increases, while for smaller code rates it decreases. For example, Table 5 indicates the performance advantage of SC-FD for BPSK  $R = 1/2$ , which is decreased by about 3 dB compared to the  $R = 3/4$  case.

While the absolute numbers of the respective degradations change, the general conclusions as formulated in Section 4.4 remain valid.

##### Modulation

The BPSK transmission is the modulation scheme with the largest range and applicability area and will be used quite often. Moreover, for BPSK the power amplifier back-off creates the largest impact difference between SC-FD and OFDM. Since we have seen that this impact is small in a multipath environment, BPSK simulations provide the strongest support for this assertion. Therefore, we believe the ( $R = 3/4$ , BPSK) setting is a relevant case study.

##### Multipath channel

As the SC-FD advantage over OFDM is based on frequency diversity exploitation, the channel characteristics are of key importance. The results presented here are obtained for a channel length  $L = 4$ . For longer channel lengths, the inherent channel frequency diversity is larger and thus the SC-FD advantage will be larger as well, while for shorter channel lengths, the frequency diversity and the corresponding SC-FD advantage is smaller.

TABLE 5: SC-FD advantage (dB) over OFDM for SISO BPSK  $R = 1/2$ .

Goodput	0.90	0.95
MMSE	4.5	4.6
DFE	4.7	4.8

## 5. CONCLUSIONS

In this paper, we compared the performance of OFDM and SC-FD as modulation schemes for broadband wireless communication. In high-performant systems, multiple antennas are used to increase spectral efficiency and/or link reliability and DFE is used to improve the linear performance. For OFDM, this results in the so-called OFDM-pcSIC scheme, while for SC-FD we first extended the single-antenna SC-FD-DFE scheme to multiple-antenna scenarios. We have taken into account most important nonidealities to obtain a realistic comparison: the power amplifier of the transmitter is very important for battery-driven terminals; we included most digital imperfections, such as channel and noise estimation, clipping and quantization, transmit and receive filtering, as well as link layer efficiency.

Our analysis shows that for frequency-selective channels the relative impact of the power amplifier on the multipath performance is negligible compared to the frequency diversity impact. Because of its higher frequency diversity exploitation, SISO SC-FD outperforms OFDM by about 7 dB in our case study. Adding multiple antennas decreases the relative diversity gap between both schemes and thus decreases the SC-FD advantage over OFDM until the linear equalizer performances converge for a large number of antennas. DFE increases the SC-FD advantage in single-user scenarios, while it decreases the advantage in a multiuser context.

In summary, the higher frequency diversity exploitation of SC-FD allows it to outperform OFDM in frequency-selective channels. Therefore, SC-FD is a suitable scheme for broadband wireless communication.

## APPENDIX

### DERIVATION OF THE FEEDFORWARD AND FEEDBACK COEFFICIENTS

The equalizer output just before the slicer at time  $n$  is as in (1). As mentioned, for each frequency tap  $l$ ,  $\tilde{\mathbf{r}}_l$  is an  $A \times 1$  vector with the received signal in the frequency domain and  $\tilde{\mathbf{W}}_l$  is a  $U \times A$  equalization matrix representing the feedforward part;  $\mathbf{F}^*$  is a  $U \times (U \cdot L_{\text{dfe}})$  matrix containing the feedback coefficients;  $\tilde{\mathbf{a}}_{n-k}$  is a  $(U \cdot L_{\text{dfe}}) \times 1$  vector with the data decisions selected for feedback, for all users.  $N$  is the number of time domain symbols in an SC-FD block.

For the mathematical tractability of the analysis, we assume perfect feedback (i.e. no decision errors,  $\tilde{\mathbf{a}}_{n-k} = \mathbf{a}_{n-k}$ ); for the simulated performance assessment, we do take decision errors into account.

The derivation of the coefficients extends the analysis by [11] to multiple antennas and multiple users.

Because of the use of a TS the channel and the data symbols appear cyclic at the receiver and thus at each frequency tap  $l$  we can write

$$\tilde{\mathbf{r}}_l = \tilde{\mathbf{H}}_l \tilde{\mathbf{a}}_l + \tilde{\mathbf{v}}_l, \quad (\text{A.1})$$

with

$$\begin{aligned} \tilde{\mathbf{H}}_l &= \sum_{n=0}^{N-1} \mathbf{H}_n \exp\left(-j2\pi \frac{ln}{N}\right), \\ \tilde{\mathbf{a}}_l &= \sum_{n=0}^{N-1} \mathbf{a}_n \exp\left(-j2\pi \frac{ln}{N}\right), \\ \tilde{\mathbf{v}}_l &= \sum_{n=0}^{N-1} \mathbf{n}_n \exp\left(-j2\pi \frac{ln}{N}\right), \end{aligned} \quad (\text{A.2})$$

where  $\tilde{\mathbf{H}}_l$  is an  $A \times U$  matrix containing the coefficients for the channels from the  $U$  users to the  $A$  receive antennas at carrier  $l$ ,  $\mathbf{a}_n$  is the  $U \times 1$  vector with the user data and  $\mathbf{n}_n$  is the  $A \times 1$  noise vector.

The autocorrelation functions associated with the data and the noise are

$$\begin{aligned} E[\tilde{\mathbf{a}}_{l_1} \tilde{\mathbf{a}}_{l_2}^*] &= N\delta(l_1 - l_2)_{\text{mod } N} \mathbf{I}_U, \\ E[\tilde{\mathbf{v}}_{l_1} \tilde{\mathbf{v}}_{l_2}^*] &= N\sigma^2 \cdot \delta(l_1 - l_2)_{\text{mod } N} \mathbf{I}_A, \end{aligned} \quad (\text{A.3})$$

$\mathbf{I}_U$  is the identity matrix of size  $U \times U$ ;  $\mathbf{I}_A$  of size  $A \times A$ . The error just before the slicer can be written as a  $U \times 1$  vector containing the error for each user at each timing instant  $n$ :

$$\mathbf{e}_n = \mathbf{z}_n - \mathbf{a}_n. \quad (\text{A.4})$$

The average error is

$$\begin{aligned} \text{MSE} &= E[\mathbf{e}_n \cdot \mathbf{e}_n^*] = E[(\mathbf{z}_n - \mathbf{a}_n)(\mathbf{z}_n^* - \mathbf{a}_n^*)] \\ &= \frac{1}{N} \sum_{l=0}^{N-1} (\tilde{\mathbf{W}}_l \tilde{\mathbf{H}}_l - \tilde{\mathbf{F}}_l)(\tilde{\mathbf{H}}_l^* \tilde{\mathbf{W}}_l^* - \tilde{\mathbf{F}}_l^*) \\ &\quad + \frac{\sigma^2}{N} \sum_{l=0}^{N-1} \tilde{\mathbf{W}}_l \tilde{\mathbf{W}}_l^*, \end{aligned} \quad (\text{A.5})$$

with (3), where  $\mathbf{k}_i$  is a  $1 \times L_{\text{dfe}}$  vector containing the selected indices of the feedback taps for user  $i$ . Minimizing the MSE from (A.5) with respect to the feedforward coefficients means solving the following equation:

$$\frac{\delta \text{MSE}}{\delta \tilde{\mathbf{W}}_l} = \sum_{l=0}^{N-1} (\tilde{\mathbf{H}}_l \tilde{\mathbf{H}}_l^* + \sigma^2 \mathbf{I}_A) \tilde{\mathbf{W}}_l^* - \tilde{\mathbf{H}}_l \tilde{\mathbf{F}}_l^* = 0. \quad (\text{A.6})$$

The solution for the optimal feedforward coefficients  $\tilde{\mathbf{W}}_l$  is,

for every  $l$ ,

$$\tilde{\mathbf{W}}_l = \tilde{\mathbf{F}}_l \tilde{\mathbf{H}}_l^* (\tilde{\mathbf{H}}_l \tilde{\mathbf{H}}_l^* + \sigma^2 \mathbf{I}_A)^{-1}. \quad (\text{A.7})$$

The matrix inversion lemma states that

$$(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{B}(\mathbf{C}^{-1} + \mathbf{DA}^{-1} \mathbf{B})^{-1} \mathbf{DA}. \quad (\text{A.8})$$

Applying this lemma to our case results in

$$\tilde{\mathbf{H}}_l^* (\tilde{\mathbf{H}}_l \tilde{\mathbf{H}}_l^* + \sigma^2 \mathbf{I}_A)^{-1} \tilde{\mathbf{H}}_l = \mathbf{I}_U - \sigma^2 (\tilde{\mathbf{H}}_l^* \tilde{\mathbf{H}}_l + \sigma^2 \mathbf{I}_U)^{-1}. \quad (\text{A.9})$$

Substituting (A.9) and (A.7) in (A.5) gives us

$$\text{MSE} = \frac{\sigma^2}{N} \sum_{l=0}^{N-1} \tilde{\mathbf{F}}_l (\tilde{\mathbf{H}}_l^* \tilde{\mathbf{H}}_l + \sigma^2 \mathbf{I}_U)^{-1} \tilde{\mathbf{F}}_l^*. \quad (\text{A.10})$$

Minimizing (A.10) with respect to the feedback coefficients  $\mathbf{F}$  means solving

$$\frac{\delta \text{MSE}}{\delta \mathbf{F}} = \frac{\sigma^2}{N} \sum_{l=0}^{N-1} (\mathbf{I}_U + \mathbf{F}^* \mathbf{X}_l^*) (\tilde{\mathbf{H}}_l^* \tilde{\mathbf{H}}_l + \sigma^2 \mathbf{I}_U)^{-1} \mathbf{X}_l = 0. \quad (\text{A.11})$$

This leads to

$$\begin{aligned} &-\frac{1}{N} \sum_{l=0}^{N-1} \underbrace{\mathbf{X}_l^* (\tilde{\mathbf{H}}_l^* \tilde{\mathbf{H}}_l + \sigma^2 \mathbf{I}_U)^{-1}}_{\mathbf{t}} \\ &= \frac{1}{N} \sum_{l=0}^{N-1} \underbrace{\mathbf{X}_l^* (\tilde{\mathbf{H}}_l^* \tilde{\mathbf{H}}_l + \sigma^2 \mathbf{I}_U)^{-1} \mathbf{X}_l \cdot \mathbf{F}}_{\mathbf{T}} \end{aligned} \quad (\text{A.12})$$

from which we get the final result

$$\mathbf{F} = -\mathbf{T}^{-1} \cdot \mathbf{t}. \quad (\text{A.13})$$

Using (A.7) and (A.13) realizes the SC-FD-DFE.

## REFERENCES

- [1] IEEE Standard 802.11a-1999, "Part 11: wireless LAN medium access control (MAC) and physical layer (PHY) specifications—amendment 1: high-speed physical layer in the 5 GHz band," September 1999.
- [2] HIPERLAN/2 standard, "Functional specification data link control (DLC) layer," October 1999.
- [3] I. Koffman and V. Roman, "Broadband wireless access solutions based on OFDM access in IEEE 802.16," *IEEE Communications Magazine*, vol. 40, no. 4, pp. 96–103, 2002.
- [4] U. Reimers, "DVB-T: the COFDM-based system for terrestrial television," *Electronics & Communication Engineering Journal*, vol. 9, no. 1, pp. 28–32, 1997.
- [5] P. Shelswell, "The COFDM modulation system: the heart of digital audio broadcasting," *Electronics & Communication Engineering Journal*, vol. 7, no. 3, pp. 127–136, 1995.
- [6] J. Bingham, "Multicarrier modulation for data transmission: an idea whose time has come," *IEEE Communications Magazine*, vol. 28, no. 5, pp. 5–14, 1990.
- [7] A. Czyliwlik, "Comparison between adaptive OFDM and single carrier modulation with frequency domain equalization," in

- Proc. 47th IEEE Vehicular Technology Conference (VTC '97)*, vol. 2, pp. 865–869, Phoenix, Ariz, USA, May 1997.
- [8] J. Tubbax, B. Côme, L. Van der Perre, L. Deneire, S. Donnay, and M. Engels, "OFDM versus Single Carrier with Cyclic Prefix: a system-based comparison," in *Proc. 54th IEEE Vehicular Technology Conference (VTC '01)*, vol. 2, pp. 1115–1119, Atlantic City, NJ, USA, October 2001.
- [9] P. Struhsaker and K. Griffin, "Analysis of PHY waveform peak to mean ratio and impact on RF amplification," IEEE 802.16 Broadband Wireless Access Working Group, March 2001.
- [10] P. Vandenameele, L. Van der Perre, and M. Engels, *Space Division Multiple Access for Wireless Local Area Networks*, The Kluwer International Series in Engineering and Computer Science, Kluwer Academic, Boston, Mass, USA, 2001.
- [11] D. Falconer, S. L. Ariyavisitakul, A. Benyamin-Seeyar, and B. Eidson, "Frequency domain equalization for single-carrier broadband wireless systems," *IEEE Communications Magazine*, vol. 40, no. 4, pp. 58–66, 2002.
- [12] N. Benvenuto and S. Tomasin, "On the comparison between OFDM and single carrier modulation with a DFE using a frequency-domain feedforward filter," *IEEE Transactions on Communications*, vol. 50, no. 6, pp. 947–955, 2002.
- [13] G. Carron, R. Ness, L. Deneire, L. Van der Perre, and M. Engels, "Comparison of two modulation techniques using frequency domain processing for in-house networks," *IEEE Transactions on Consumer Electronics*, vol. 47, no. 1, pp. 63–72, 2001.
- [14] L. Deneire, B. Gyselinx, and M. Engels, "Training sequence versus cyclic prefix—a new look on single carrier communication," *IEEE Communications Letters*, vol. 5, no. 7, pp. 292–294, 2001.
- [15] B. Côme, R. Ness, S. Donnay, et al., "Impact of front-end non-idealities on bit error rate performance of WLAN-OFDM transceivers," *Microwave Journal*, vol. 44, no. 2, pp. 126–140, 2001.
- [16] S. Thoen, L. Deneire, L. Van der Perre, M. Engels, and H. De Man, "Constrained least squares detector for OFDM/SDMA-based wireless networks," *IEEE Transactions on Wireless Communications*, vol. 2, no. 1, pp. 129–140, 2003.
- [17] A. M. Chan and G. W. Wornell, "A class of block-iterative equalizers for intersymbol interference channels: Fixed channel results," *IEEE Transactions on Communications*, vol. 49, no. 11, pp. 1966–1976, 2001.

**Jan Tubbax** received his Electrical Engineering degree from the Katholieke Universiteit Leuven (KUL) in July 1998. He performed his M.S. thesis on the "Performance Evaluation of IEEE 802.11a and IPv6" in the Design Technology for Integrated Information and Communication Systems (DESICS) division at the Interuniversity Micro-Electronics Center (IMEC) in Leuven. In September 1998, he joined the Wireless Systems (WISE) group of DESICS at IMEC. Within the WISE group, he performed research on synchronization strategies for satellite networks. In the spring of 2000, he started working towards a Ph.D. in the mixed-signal design of wireless local area networks (WLAN) at the Department of Electrical Engineering (ESAT/INSYS), KUL. The goal of this Ph.D. is the design of low-cost and low-power WLAN radios through the development of digital algorithms that allow to relax the analog front-end specifications. During the spring of 2002, he visited the Information Systems Laboratory (ISL) at Stanford University, California, USA,



to perform research on different equalization strategies for WLANs. His main research interests are mixed signal wireless design, digital compensation techniques, modulation techniques, and multiple antenna systems.

**Liesbet Van der Perre** received his M.S. degree from the Katholieke Universiteit Leuven, Leuven, Belgium, in 1992, and his Ph.D. degree from the Katholieke Universiteit Leuven, Belgium, in 1997, both in electrical engineering. She performed her M.S. thesis research at the École Nationale Supérieure des Télécommunications (ENST) in Paris, France. Currently, she is the Director of the wireless program at IMEC, Leuven. Her work in the past focused on system design and digital modems for high-speed wireless communications. She was a system architect in IMEC's OFDM ASICs development and the leader of the turbo coding team. She is also a part-time Professor at the University of Antwerp, Belgium.

**Marc Engels** is currently the General Manager of Flanders' Mechatronics Technology Centre, a new research center that aims at increasing the intelligence of mechatronic systems. Research topics include intelligent sensors, communication, intelligent control, and embedded software. Before, Marc Engels was the Chief Technology Officer (CTO) of LoraNet, a start-up in the field of broadband wireless communication. Previously, he was the Director of the wireless department at IMEC, focused on the implementation of telecommunication systems on a chip. For these systems, he overlooked research on the DSP processing, the mixed-signal RF front-end, and the software protocols. He was also active in design methods and tools for implementing multidisciplinary systems. Under his supervision, several systems have been realized, including a 54 Mbps WLAN terminal, a GPS-GLONASS receiver, a DECT-GSM dual mode phone, a cable modem, and so forth. Previously, Engels performed research at the Katholieke Universiteit Leuven, Belgium, Stanford University, Calif, USA, and the Royal Military School, Brussels, Belgium. Marc Engels received his Engineering degree (1988) and the Ph.D. (1993), both from the Katholieke Universiteit Leuven, Belgium. He is also a Professor at the Department of Electrical Engineering (ESAT/INSYS), KUL.

**Hugo De Man** has been a Professor in electrical engineering at the Katholieke Universiteit Leuven, Belgium, since 1976. He was a Visiting Associate Professor at the University of California at Berkeley in 1975. In 1984, he was a cofounder of IMEC, the largest independent European semiconductor research institute. From 1984 to 1995, he was Vice President of IMEC, responsible for research in design technology for DSP and telecom applications. In 1995, he became a Senior Research Fellow of IMEC, working on strategies for education and research on design of future post PC systems. His research at IMEC led to many novel tools and methods on high-level synthesis, hardware-software codesign and C++ based design. Many tools are commercialized by spin-off companies, like CoWare, Adelante Technologies, Target Compilers. His work and teaching resulted in a cluster of DSP-oriented companies in Leuven, known as DSP Valley, working on design tools, telecommunications, networking, and multimedia integrated system products. Professor De Man received the Technical Achievement Award of the IEEE Signal Processing Society, the Phil Kaufman Award of the EDA Consortium, and the Golden Jubilee Medal of the IEEE Circuits and Systems Society. He is a member of the Royal Academy of Sciences in Belgium.



**Marc Moonen** received his Electrical Engineering degree and his Ph.D. degree in applied sciences from the Katholieke Universiteit Leuven, Leuven, Belgium, in 1986 and 1990, respectively. Since 2000, he has been an Associate Professor at the Electrical Engineering Department of Katholieke Universiteit Leuven, where he is currently heading a research team of 16 Ph.D. candidates and postdocs, working in the area of signal processing for digital communications, wireless communications, DSL, and audio signal processing. He received the 1994 KUL Research Council Award, the 1997 Alcatel Bell (Belgium) Award (with Piet Vandaele), and was a 1997 “Laureate of the Belgium Royal Academy of Science.” He was Chairman of the IEEE Benelux Signal Processing Chapter (1998–2002), and is currently a EURASIP Ad-Com Member (European Association for Signal, Speech and Image Processing from 2000 till now). He is the Editor-in-Chief for the “EURASIP Journal on Applied Signal Processing” from 2003 till now, and a Member of the Editorial Board of “Integration, the VLSI Journal,” “IEEE Transactions on Circuits and Systems II” (2002–2003), “EURASIP Journal on Wireless Communications and Networking,” and “IEEE Signal Processing Magazine.”



# Bounds on the Outage-Constrained Capacity Region of Space-Division Multiple-Access Radio Systems

**Haipeng Jin**

*Center for Wireless Communications, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92097, USA*  
Email: jin@cw.cucsd.edu

**Anthony Acampora**

*Center for Wireless Communications, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92097, USA*  
Email: acampora@ece.ucsd.edu

*Received 30 May 2003; Revised 5 February 2004*

Space-division multiple-access (SDMA) systems that employ multiple antenna elements at the base station can provide much higher capacity than single-antenna-element systems. A fundamental question to be addressed concerns the ultimate capacity region of an SDMA system wherein a number of mobile users, each constrained in power, try to communicate with the base station in a multipath fading environment. In this paper, we express the capacity limit as an outage region over the space of transmission rates  $R_1, R_2, \dots, R_n$  from the  $n$  mobile users. Any particular set of rates contained within this region can be transmitted with an outage probability smaller than some specified value. We find outer and inner bounds on the outage capacity region for the two-user case and extend these to multiple-user cases when possible. These bounds provide yardsticks against which the performance of any system can be compared.

**Keywords and phrases:** outage capacity, space-division multiple-access, capacity region.

## 1. INTRODUCTION

Time-varying multipath fading is a fundamental phenomenon affecting the availability of terrestrial radio systems, and strategies to abate or exploit multipath are crucial. Recent information theoretic research [1, 2] has shown that in most scattering environments, a multiple-antenna-element (MAE) array is a practical and effective technique to exploit the effect of multipath fading and achieve enormous capacity advantages.

Next-generation wireless systems are intended to provide high voice quality and high-rate data services. At the same time, the mobile units must be small and lightweight. It appears that base station complexity is the preferred strategy for meeting the requirements of the next-generation systems. In particular, MAE arrays can be installed at the base stations to provide higher capacity.

A primary question to be addressed is the ultimate capacity limit of a single cell with constrained user power and multipath fading. Since there are multiple users in the cell, the capacity limit is expressed as a region of allowable transmission rates such that information can be reliably transmitted by user 1 at rate  $R_1$ , user 2 at rate  $R_2$ , and so on. For a static channel condition with fixed fade depth at each mobile, the capacity region of the multiple-access channel is then the set

of all rate vectors  $\mathbf{R} = \{R_1, R_2, \dots, R_n\}$  that can be achieved with arbitrarily small error probability [3, 4]. However, when the channel is time varying due to the dynamic nature of the wireless communication environment, the capacity region is characterized differently depending on the delay requirements of the mobiles and the coherence time of the channel fading. Two important notions are used in the dynamic channel case [2, 5]: ergodic capacity and outage capacity. Ergodic capacity is defined for channels with long-term delay constraint, meaning that the transmission time is long enough to reveal the long-term ergodic properties of the fading channel. The ergodic capacity is given by the appropriately averaged mutual information. In practical communication systems operating on fading channels, the ergodic assumption is not necessarily satisfied. For example, in the cases with real-time applications over wireless channels, stringent delay constraints are demanded and the ergodicity requirement cannot be fulfilled. No significant channel variability occurs during a transmission. There may be nonzero outage probability associated with any value of actual transmission rate, no matter how small. Here, we have to consider the information rate that can be maintained under all channel conditions, at least, within a certain outage probability [6, 7, 8, 9]. The maximal rate that can be achieved with a given outage

percentile  $p$  is defined as the  $p$  percent outage capacity. Both the ergodic capacity and outage capacity notions originated from single-user case. They are easily extended to cases with multiple users.

The outage capacity issue for single-user multiple-input multiple-output (MIMO) case was studied extensively. In [1], the authors characterized the outage capacity for a point-to-point MIMO channel subject to flat Rayleigh fading. The cumulative distribution functions for the outage capacity were presented such that given a specific outage probability, we will know at what rate information can be transmitted over the MIMO channel. Biglieri et al. [10] considered the outage capacity of a MIMO system for different delay and transmit power constraints.

However, there are limited results on the outage capacity region of multiple-access system with multiple antennas at the base station. Most previous studies are constrained to either fixed channel condition [11, 12] or ergodic capacity region [13]. The space-division multiple-access (SDMA) capacity regions under fixed channel conditions were considered in [11] with both independent decoding and joint decoding schemes. An iterative algorithm was proposed in [12] to maximize the sum capacity of a time invariant Gaussian MIMO multiple-access channel. The ergodic capacity region for MIMO multiple-access channel with covariance feedback has been studied in [13].

In this paper, we consider the outage capacity region for single-cell flat fading SDMA systems with MAEs at the base station and multiple mobiles, each with a single antenna element. From the outage capacity region, we can determine with what outage probability a certain rate vector can be transmitted with arbitrarily small error probabilities. Specifically, we derive outer and inner bounds on the outage capacity region for the two-user case and explain how the same principles can be extended to multiple-user case.

To find an outer bound on the capacity region, cooperation among the geographically separated mobile stations is assumed to take place via a virtual central processor, and there is a total power constraint. The total capacity is found for all combinations of  $n, n-1, \dots, 2, 1$  mobile stations in the system. For example, if there are two mobiles, the capacity of each, in isolation, is found, along with the combination of both, treated as having a single transmitter with two geographically remote antenna elements.

By definition, any realizable approach for which the capacity region can be found forms an achievable inner bound to the capacity region. Time sharing among users provides a simple inner bound [14]. We also derive a tighter inner bound by allowing users to transmit at the same time while performing joint decoding at the base station. It is noted that construction of the inner bound also provides a method for achieving the inner bound.

Most of our derivations and discussions will be focused on the two-user case since the results in this case can be easily displayed graphically and provide significant physical insight. However, along our way, we will point out wherever the results are extensible to cases involving more than two mobiles.

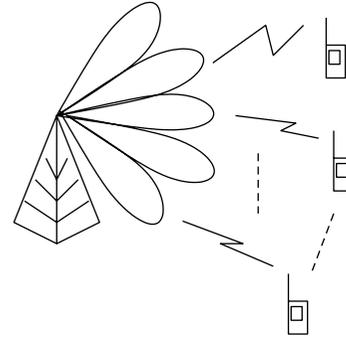


FIGURE 1: Space-division multiple-access systems.

This paper is organized as follows. Section 2 contains a description of the channel model which we use, and also introduces the concept of outage capacity region. Outer and inner bounds on the outage capacity region are derived in Section 3. Numerical results are presented in Section 4, along with discussions.

## 2. CHANNEL MODEL AND DEFINITION OF THE OUTAGE CAPACITY REGION

We consider a single-cell system where a number of geographically separated mobile users communicate with the base station. The system is shown in Figure 1. We consider the reverse link from the mobiles to the base station and model this as a multiple-access system. All mobile are equipped with single antenna element, and the base station is equipped with multiple receive antenna elements to exploit spatial diversity. We assume that the channel between each mobile station and each base station antenna element is subject to flat Rayleigh fading and that the fading between each mobile and each base station antenna element is independent of the fading between other mobile-base station pairs. Also present at each antenna element is additive white Gaussian noise (AWGN). Rayleigh fading between mobile  $j$  and base station antenna element  $i$  is represented by a zero mean, unit variance complex Gaussian random variable  $H_{ij} = N(0, \sigma^2) + \sqrt{-1}N(0, \sigma^2)$ , where  $E|H_{ij}|^2 = 2\sigma^2 = 1$ . The noise components observed at all the receiving antennas are identical, and independent white Gaussian distributed with power  $\sigma_n^2$ . Each mobile's transmitted power is limited such that the average received signal-to-noise ratio at each base station antenna element is  $\rho$  if only one mobile is transmitting.

Suppose there are  $n$  mobile stations and  $m$  antenna elements at the base station. Then the received signal can be represented as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \quad (1)$$

where  $\mathbf{y}$  is an  $m$ -element vector representing the received signals,  $\mathbf{x}$  is a vector with  $n$  elements, each element representing the signal transmitted by one mobile station,  $\mathbf{H}$  is the channel fading matrix with  $m \times n$  complex Gaussian elements, and  $\mathbf{n}$  is the received AWGN noise vector with covariance  $\sigma_n^2 \mathbf{I}_{m \times m}$ .

Consider first the simple case where the channel conditions are fixed, that is,  $\mathbf{H}$  is constant. The capacity region of the multiple-access channel is then the set of all rate vectors  $\mathbf{R} = \{R_1, R_2, \dots, R_n\}$  satisfying [3]

$$\sum_{k \in T} R_k \leq I(\mathbf{y}; x_k, k \in T \mid x_l, l \in \bar{T}), \quad (2)$$

where  $I$  stands for mutual information,  $T$  denotes any subset of  $\{1, 2, \dots, n\}$ ,  $\bar{T}$  its complement, and  $\rho$  denotes the power limitation. In essence, the sum rate of any subset of the mobiles  $\{1, 2, \dots, n\}$  needs to be smaller than the mutual information between the transmitted and received signals if only the mobiles with the subset are transmitting. We can express the capacity region concisely as follows:

$$C(\mathbf{H}, \rho) = \left\{ \mathbf{R} \left\| \sum_{k \in T} R_k \leq I(\mathbf{y}; x_k, k \in T \mid x_l, l \in \bar{T}) \right. \right\}. \quad (3)$$

All the rate vectors  $\mathbf{R}$  within the region  $C(\mathbf{H}, \rho)$  can be achieved with arbitrarily small probability.

Now, since the channel matrix is a set of random variables, the capacity region is random and a key goal is to find the cumulative distribution functions of the regions, from which we can determine the probability that a specific rate vector can be transmitted. With this in mind, we define the outage capacity region as

$$\mathcal{C}(p, \rho) = \{ \mathbf{R} \mid \text{Prob}(\mathcal{N}) \geq p \}, \quad (4)$$

where  $\mathcal{N}$  is defined as the set of the channel conditions under which the rate vector  $\mathbf{R}$  is achievable with arbitrarily small error probability. The set  $\mathcal{N}$  can be written as

$$\mathcal{N} = \{ \mathbf{H} \mid \mathbf{R} \in C(\mathbf{H}, \rho) \}. \quad (5)$$

In (4),  $1 - p$  is the outage probability and  $\rho$  represents the power limitation. The outage capacity region  $\mathcal{C}(p, \rho)$  contains all the rate vectors that can be achieved with a probability greater than or equal to  $p$ . Alternatively speaking, the probability that a rate vector contained in  $\mathcal{C}(p, \rho)$  cannot be achieved is less than  $1 - p$ .

To simplify notation, we write  $\mathcal{C}(p, \rho)$  as  $\mathcal{C}(p)$  and  $C(\mathbf{H}, \rho)$  as  $C(\mathbf{H})$ , with the implication that the power limitation is always specified by  $\rho$  unless otherwise stated.

For the two-mobile case, we define  $\mathbf{h}_1$  and  $\mathbf{h}_2$  as the channel response vectors from mobiles 1 and 2, respectively, to the base station antenna elements, and they can be expressed as  $\mathbf{h}_1 = \{H_{11}, H_{21}, \dots, H_{m1}\}^T$  and  $\mathbf{h}_2 = \{H_{12}, H_{22}, \dots, H_{m2}\}^T$ , where  $T$  denotes the transpose operation. As a result, the channel fading matrix  $\mathbf{H} = [\mathbf{h}_1 \mathbf{h}_2]$ . Then the outage capacity region is given as

$$\mathcal{C}(p) = \{ (R_1, R_2) \mid \text{Prob}(\mathcal{N}) \geq p \}, \quad (6)$$

where  $\mathcal{N}$  is the set of channel response matrices which satisfy the following three conditions:

$$\begin{aligned} R_1 &\leq C_1(\mathbf{h}_1), \\ R_2 &\leq C_2(\mathbf{h}_2), \\ R_1 + R_2 &\leq C_{12}(\mathbf{H}), \end{aligned} \quad (7)$$

where  $C_1(\mathbf{h}_1)$ ,  $C_2(\mathbf{h}_2)$ , and  $C_{12}(\mathbf{H})$  define the capacity region under the fading condition specified by  $\mathbf{H}$ . For convenience, we write this as follows:

$$\mathcal{N} = \left\{ \mathbf{H} \left\| \begin{array}{l} R_1 \leq C_1(\mathbf{h}_1) \\ R_2 \leq C_2(\mathbf{h}_2) \\ R_1 + R_2 \leq C_{12}(\mathbf{H}) \end{array} \right. \right\}. \quad (8)$$

In (6),  $\mathcal{C}(p)$  has the same interpretation as in (4); it consists of the rate pairs  $(R_1, R_2)$  simultaneously achievable with a probability greater than or equal to  $p$ , that is, the outage probability is smaller than or equal to  $1 - p$ .

### 3. OUTAGE CAPACITY BOUNDS

In this section, we derive bounds on the outage capacity region with a focus on the two-mobile case. The base station is equipped with  $m$  antenna elements. As we will see, most of our derivations are not constrained by the number of mobiles, and thus are applicable to cases with an arbitrary number of mobiles.

#### 3.1. Outer bound

To obtain an outer bound on the outage capacity region, we start out by defining the following rate regions:

$$\begin{aligned} \mathcal{B}_1(p) &= \{ (R_1, R_2) \mid \text{Prob}(\mathcal{M}_1) \geq p \}, \\ \mathcal{B}_2(p) &= \{ (R_1, R_2) \mid \text{Prob}(\mathcal{M}_2) \geq p \}, \\ \mathcal{B}_a(p) &= \{ (R_1, R_2) \mid \text{Prob}(\mathcal{M}_a) \geq p \}, \end{aligned} \quad (9)$$

where  $\mathcal{M}_1$ ,  $\mathcal{M}_2$ , and  $\mathcal{M}_a$  are different sets of channel conditions. The three sets are defined by the following conditions, respectively:

$$\begin{aligned} \mathcal{M}_1 &= \{ \mathbf{H} \mid R_1 \leq C_1(\mathbf{h}_1) \}, \\ \mathcal{M}_2 &= \{ \mathbf{H} \mid R_2 \leq C_2(\mathbf{h}_2) \}, \\ \mathcal{M}_a &= \{ \mathbf{H} \mid R_1 + R_2 \leq C_{12}(\mathbf{H}) \}. \end{aligned} \quad (10)$$

The values  $C_1(\mathbf{h}_1)$ ,  $C_2(\mathbf{h}_2)$ , and  $C_{12}(\mathbf{H})$  are the same as those in (8); they jointly define the multiple-access capacity region under the fading condition  $\mathbf{H}$ . As a result of the definition, the rate pairs in  $\mathcal{B}_1(p)$  and  $\mathcal{B}_2(p)$  only satisfy the constraint on the individual rates  $R_1$  and  $R_2$ , respectively; the rate pairs in  $\mathcal{B}_a(p)$  only satisfy the constraint on the sum rate  $R_1 + R_2$ . We can now prove the following:

Claim 1.  $\mathcal{C}(p) \subset \mathcal{B}_1(p)$ ,  $\mathcal{C}(p) \subset \mathcal{B}_2(p)$ , and  $\mathcal{C}(p) \subset \mathcal{B}_a(p)$ .

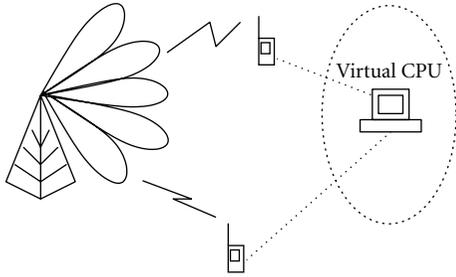


FIGURE 2: Space-division multiple-access systems with coordinated users.

*Proof.* For a given  $(R_1, R_2)$ , if  $\mathbf{H} \in \mathcal{N}$ , where  $\mathcal{N}$  is defined in (8), then  $C_1(\mathbf{h}_1) \geq R_1$ ,  $C_2(\mathbf{h}_2) \geq R_2$ , and  $C_{12}(\mathbf{H}) \geq R_1 + R_2$ . By definition,  $\mathbf{H} \in \mathcal{M}_1$ ,  $\mathbf{H} \in \mathcal{M}_2$ , and  $\mathbf{H} \in \mathcal{M}_a$ . As a result,  $\mathcal{N} \subset \mathcal{M}_1$ . This implies that  $\text{Prob}(\mathcal{M}_1) \geq p$  if  $\text{Prob}(\mathcal{N}) \geq p$ . Therefore,  $\mathcal{C}(p) \subset \mathcal{B}_1(p)$ . Similarly, we can show that  $\mathcal{C}(p) \subset \mathcal{B}_2(p)$  and  $\mathcal{C}(p) \subset \mathcal{B}_a(p)$ .  $\square$

If a set is contained in each of several sets, then it is also contained in the intersection of those sets. Consequently, we can obtain the following outer bound for the outage capacity region.

Claim 2.  $\mathcal{C}(p) \subset \mathcal{U}(p)$ , where  $\mathcal{U}(p) = \mathcal{B}_1(p) \cap \mathcal{B}_2(p) \cap \mathcal{B}_a(p)$ .

In order to obtain the outer bound given in Claim 2, we need to evaluate  $C_1(\mathbf{h}_1)$ ,  $C_2(\mathbf{h}_2)$ , and  $C_{12}(\mathbf{H})$  under every specific fading condition  $\mathbf{H}$ . The sum capacity  $C_{12}(\mathbf{H})$  is usually difficult to find. Fortunately, upper bounds on the sum capacity are easily obtained. We can use these upper bounds to find looser outer bounds on the outage capacity region that are easy to evaluate.

An upper bound  $C'_{12}(\mathbf{H})$  on the sum capacity  $C_{12}(\mathbf{H})$  can be obtained by assuming that both users are connected via some error free channel to a central coordinator as shown in Figure 2. We also assume that the virtual transmitter formed this way has perfect knowledge about the channel; thus singular value decomposition and water-filling techniques [2, 3] can be used to achieve the highest possible capacity. In water-filling, more power is allocated to better subchannels with higher signal-to-noise ratio so as to maximize the sum of data rates in all subchannels. If we define  $\mathcal{N}'$  as

$$\mathcal{N}' = \left\{ \mathbf{H} \left| \begin{array}{l} R_1 \leq C_1(\mathbf{h}_1) \\ R_2 \leq C_2(\mathbf{h}_2) \\ R_1 + R_2 \leq C_w(\mathbf{H}) \end{array} \right. \right\}, \quad (11)$$

where  $C_w(\mathbf{H})$  is the water-filling capacity under channel condition  $\mathbf{H}$  and it is always greater than the actual sum capacity  $C_{12}(\mathbf{H})$ , it follows that the set  $\mathcal{N}$  defined in (8) is always a subset of  $\mathcal{N}'$ . Consequently, we can use  $\mathcal{N}'$  to define the following outer bound on the outage capacity region:

$$\mathcal{C}'(p) = \{(R_1, R_2) \mid \text{Prob}(\mathcal{N}') \geq p\} \supset \mathcal{C}(p). \quad (12)$$

Now we define the following region:

$$\begin{aligned} \mathcal{B}'_a(p) &= \{(R_1, R_2) \mid \text{Prob}(\mathcal{M}'_a) \geq p\}, \\ \mathcal{M}'_a &= \{\mathbf{H} \mid R_1 + R_2 \leq C_w(\mathbf{H})\}. \end{aligned} \quad (13)$$

Region  $\mathcal{B}'_a(p)$  contains all the rate pairs whose sum rates are constrained by the water-filling capacity. Following the steps used to prove Claim 1, the following readily shown.

Claim 3.  $\mathcal{C}'(p) \subset \mathcal{B}_1(p)$ ,  $\mathcal{C}'(p) \subset \mathcal{B}_2(p)$ , and  $\mathcal{C}'(p) \subset \mathcal{B}'_a(p)$ .

Consequently, the following claim provides an outer bound on the outage capacity region.

Claim 4.  $\mathcal{C}(p) \subset \mathcal{C}'(p) \subset \{\mathcal{B}_1(p) \cap \mathcal{B}_2(p) \cap \mathcal{B}'_a(p)\}$ .

The outer bounds given in both Claims 2 and 4 are very easy to evaluate since the outer bounds consist of a set of regions defined by straight lines.

The above derivation can be used to an outer bound for multiple-user outage capacity region. The only difference is that the outer bound on the outage capacity region will be defined by a series of planes rather than straight lines, and is therefore in the shape of a polyhedra instead of a polygon as in the two-user case. Each plane will correspond to an outer bound on the capacity of one combination of users chosen from the entire set of users. For example, if there are three mobile users, then we can find the bounding rate regions for each of the following combinations of the users:  $\{1\}$ ,  $\{2\}$ ,  $\{3\}$ ,  $\{1, 2\}$ ,  $\{1, 3\}$ ,  $\{2, 3\}$ , and  $\{1, 2, 3\}$ , and then take their intersection as the outer bound as we have done in Claims 2 and 4.

In deriving the outer bound in Claim 4, we assumed that the mobiles are coordinated by a central processing unit, and the channel condition is known at both the base station and the virtual coordinated transmitter. Thus, for our outer bound, the SDMA system is reduced to a point-to-point MIMO system. For such a system, it has been shown [2] that the forward and reverse channels are reciprocal and have the same capacity. Therefore, the outage capacity region outer bound given by Claim 4 for the multiple-access channel is also a bound for the broadcast channel.

### 3.2. Time-share bound

We now turn our attention to obtaining inner bounds on the outage capacity region for the two-mobile case. As we have said previously, any realizable approach for which the capacity region can be found forms an achievable inner bound to the capacity region. One such inner bound is the time-share bound, attained by time-sharing the base station between the two mobiles [14]. For every fading state  $(\mathbf{h}_1, \mathbf{h}_2)$ , if only mobile 1 is allowed to transmit, then it can achieve capacity  $C(\mathbf{h}_1)$ . Similarly, if only mobile 2 is allowed to transmit, it can achieve capacity  $C(\mathbf{h}_2)$ . An achievable time-share capacity region is given by  $\{(R_1, R_2) \mid R_1 \leq aC(\mathbf{h}_1), R_2 \leq (1-a)C(\mathbf{h}_2), 0 \leq a \leq 1\}$ .

We define an outage capacity region

$$\mathcal{C}_{\mathcal{T}_s}(p) = \{(R_1, R_2) \mid \text{Prob}(\mathcal{S}) \geq p\}, \quad (14)$$

where

$$\mathcal{S} = \{\mathbf{H} \mid R_1 \leq aC(\mathbf{h}_1), R_2 \leq (1-a)C(\mathbf{h}_2), 0 \leq a \leq 1\}. \quad (15)$$

Then,  $\mathcal{C}_{\mathcal{T}_s}(p)$  contains all the rate pairs that can be achieved by time sharing with an outage probability smaller than  $1-p$ . Thus,  $\mathcal{C}_{\mathcal{T}_s}(p)$  is an inner bound on the outage capacity region defined by (6) since the time-sharing capacity region is an achievable region, and an achievable is, by definition, an inner bound on the actual capacity region. The following claim reiterates this observation.

Claim 5.  $\mathcal{C}_{\mathcal{T}_s}(p) \subset \mathcal{C}(p)$ .

The boundary of the region  $\mathcal{C}_{\mathcal{T}_s}(p)$  is defined by all  $(R_1, R_2)$  pairs that can be achieved with an outage probability exactly equal to  $p$ , as expressed in the following condition:

$$\begin{aligned} p &= \text{Prob}(\{\mathbf{H} \mid R_1 \leq aC(\mathbf{h}_1), \\ &\quad R_2 \leq (1-a)C(\mathbf{h}_2), 0 \leq a \leq 1\}) \\ &= \text{Prob}(\{\mathbf{h}_1 \mid R_1 \leq aC(\mathbf{h}_1)\}) \\ &\quad \times \text{Prob}(\{\mathbf{h}_2 \mid R_2 \leq (1-a)C(\mathbf{h}_2)\}), \end{aligned} \quad (16)$$

since  $\mathbf{h}_1$  and  $\mathbf{h}_2$  are independent random vectors.

We now show how the boundary given in (16) can be derived in closed form. Each value  $C(\mathbf{h}_1)$  and  $C(\mathbf{h}_2)$  is the capacity of an AWGN channel with a single antenna element at the transmitter and  $m$  antenna elements at the receiver. Thus [1],

$$C(\mathbf{h}_i) = \log(1 + \rho |\mathbf{h}_i|^2), \quad (17)$$

where  $|\mathbf{h}_i|^2 = \sum_{j=1}^m |H_{ij}|^2$  is a random variable following chi-square distribution with  $2m$  degrees of freedom, and  $m$  is the number of receive antennas at the base station. The complementary cumulative distribution function  $\bar{F}(\cdot)$  of  $|\mathbf{h}_i|^2$  is given as [15] follows:

$$\bar{F}(t) = \text{Prob}(|\mathbf{h}_i|^2 > t) = \sum_{k=0}^{m-1} \frac{t^k e^{-t/2\sigma^2}}{k!(2\sigma^2)^k} = \sum_{k=0}^{m-1} \frac{t^k e^{-t}}{k!}. \quad (18)$$

We have  $2\sigma^2 = 1$  in the above equation because the Rayleigh fading gain  $H_{ij}$  between mobile  $i$  and base station antenna element  $j$  is zero mean, unit variance complex Gaussian random variable, as specified in Section 2. As a result, the boundary of  $\mathcal{L}_{\mathcal{T}_s}(p)$  is given as follows:

$$\begin{aligned} p &= \left\{ \sum_{k=0}^{m-1} \frac{1}{k!} \left( \frac{e^{R_1/a} - 1}{\rho} \right)^k \exp\left(-\frac{e^{R_1/a} - 1}{\rho}\right) \right\} \\ &\quad \times \left\{ \sum_{l=0}^{m-1} \frac{1}{l!} \left( \frac{e^{R_2/(1-a)} - 1}{\rho} \right)^l \exp\left(-\frac{e^{R_2/(1-a)} - 1}{\rho}\right) \right\} \\ &= \bar{F}\left(\frac{e^{R_1/a} - 1}{\rho}\right) \bar{F}\left(\frac{e^{R_2/(1-a)} - 1}{\rho}\right). \end{aligned} \quad (19)$$

Given a certain probability  $p$ , the maximum  $R_1$  can be found by setting  $R_2$  to zero and  $a$  to 1 in (19). Then for every  $R_1$  between the maximum and zero, we can always sweep out the possible  $R_2$ 's by varying  $a$  between 0 and 1 and solving the equations numerically.

The same principle and derivations can be applied to multiple-user cases to obtain time-sharing bounds. The only difference is that the boundary will be defined by multiple rates and the condition in (19) will be given by the product of multiple complementary cumulative functions.

### 3.3. Joint decoding inner bound

The time-share inner bound may be quite pessimistic since one mobile may transmit at any time. A tighter inner bound may be obtained by allowing the two mobiles to transmit simultaneously, but without the coordinating virtual central processing unit that was used to obtain an outer bound. We now find such an inner bound by allowing the base station to jointly detect the information from both mobile stations. In this way, an achievable capacity region for SDMA under a particular channel condition  $(\mathbf{h}_1, \mathbf{h}_2)$  is given by [3, 11]

$$C'_1(\mathbf{h}_1) = \log |1 + \rho \mathbf{h}_1^H \mathbf{h}_1| = \log |1 + \rho a_1|, \quad (20)$$

$$C'_2(\mathbf{h}_2) = \log |1 + \rho \mathbf{h}_2^H \mathbf{h}_2| = \log |1 + \rho a_2|, \quad (21)$$

$$\begin{aligned} C'_a(\mathbf{H}) &= \log |\mathbf{I} + \rho [\mathbf{h}_1 \ \mathbf{h}_2]^H [\mathbf{h}_1 \ \mathbf{h}_2]| \\ &= \log(1 + \rho a_1 + \rho a_2 + \rho^2 a_1 a_2 - \rho^2 a_1 a_2 \cos^2 \theta), \end{aligned} \quad (22)$$

where

$$\begin{aligned} a_1 &= |\mathbf{h}_1|^2, \\ a_2 &= |\mathbf{h}_2|^2, \\ \theta &= \cos^{-1} \left\{ \frac{|\mathbf{h}_1^H \cdot \mathbf{h}_2|}{|\mathbf{h}_1| \cdot |\mathbf{h}_2|} \right\}. \end{aligned} \quad (23)$$

The scalars  $a_1$  and  $a_2$  are the squared magnitude of the vectors  $\mathbf{h}_1$  and  $\mathbf{h}_2$ , respectively, and  $\theta$  is the angle between  $\mathbf{h}_1$  and  $\mathbf{h}_2$ . The three scalars  $a_1$ ,  $a_2$ , and  $\theta$  are independently distributed random variables.  $a_1$  and  $a_2$  are chi-square distributed with  $2m$  degrees of freedom and  $\theta$  is uniformly distributed over  $(0, 2\pi]$ . Thus, their joint probability density function is given by the product of the individual probability density functions [16]:

$$\text{pdf}(a_1, a_2, \cos 2\theta) = g(a_1)g(a_2)h(\cos 2\theta), \quad (24)$$

where

$$g(x) = \frac{1}{\sigma^{2n} 2^{n/2} \Gamma(n/2)} x^{n/2-1} e^{-x/2\sigma^2}, \quad 0 \leq x < \infty, \quad (25)$$

$$h(x) = \frac{1}{\pi \sqrt{1-x^2}}, \quad |x| \leq 1. \quad (26)$$

In (25), the function  $\Gamma(\cdot)$  is the Gamma function as defined in [15]. As with the time-sharing case, we can define a rate region based on this joint decoding achievable capacity region:

$$\mathcal{C}_{\mathcal{L}\mathcal{B}}(p) = \{(R_1, R_2) \mid \text{Prob}(\mathcal{U}) \geq p\}, \quad (27)$$

where

$$\mathcal{U} = \left\{ \mathbf{H} \left\| \begin{array}{l} R_1 \leq C'_1(\mathbf{h}_1) \\ R_2 \leq C'_2(\mathbf{h}_2) \\ R_1 + R_2 \leq C'_a(\mathbf{H}) \end{array} \right. \right\}. \quad (28)$$

The region  $\mathcal{C}_{\mathcal{L}\mathcal{B}}(p)$  contains all the rate pairs that fall into the achievable region with an outage probability smaller than  $1 - p$ . Again, since it is derived from an achievable region,  $\mathcal{C}_{\mathcal{L}\mathcal{B}}(p)$  forms an inner bound to the actual outage capacity region.

The boundary of the region  $\mathcal{C}_{\mathcal{L}\mathcal{B}}(p)$  is defined by all the rate pairs that can be achieved with a probability  $\text{Prob}(\mathcal{U})$  exactly equal to  $p$ . Next, we show how this probability can be expressed in terms of  $R_1$  and  $R_2$ . First, we define the following dummy variables:

$$\begin{aligned} t_1 &= e^{C'_1(\mathbf{h}_1)} = 1 + \rho a_1, \\ t_2 &= e^{C'_2(\mathbf{h}_2)} = 1 + \rho a_2, \\ t_a &= e^{C'_{12}(\mathbf{H})} = 1 + \rho(a_1 + a_2) + \rho^2 a_1 a_2 - \rho^2 a_1 a_2 \cos^2 \theta \quad (29) \\ &= 1 + \rho(a_1 + a_2) + \frac{\rho^2 a_1 a_2}{2} - \frac{\rho^2 a_1 a_2 \cos 2\theta}{2}. \end{aligned}$$

The values of  $t_1$ ,  $t_2$ , and  $t_a$  are functions of  $a_1$ ,  $a_2$ , and  $\theta$ , and their joint distribution is given by

$$\begin{aligned} f(t_1, t_2, t_a) &= g(a_1)g(a_2)h(\cos 2\theta)|\mathbf{J}| \\ &= g\left(\frac{t_1 - 1}{\rho}\right)g\left(\frac{t_2 - 1}{\rho}\right) \\ &\quad \times h\left(\frac{2(t_1 + t_2 + 1 - t_a)}{(t_1 - 1)(t_2 - 1)} + 1\right) \frac{2}{\rho^2(t_1 - 1)(t_2 - 1)}, \quad (30) \end{aligned}$$

where the square matrix  $\mathbf{J}$  is the Jacobian matrix of the transform from  $(a_1, a_2, \cos 2\theta)$  to  $(t_1, t_2, t_a)$  given by

$$\mathbf{J} = \begin{bmatrix} \frac{1}{\rho} & 0 & 0 \\ 0 & \frac{1}{\rho} & 0 \\ \frac{2(t_a - t_2)}{(t_1 - 1)^2(t_2 - 1)} & \frac{2(t_a - t_1)}{(t_1 - 1)(t_2 - 1)^2} & \frac{-2}{(t_1 - 1)(t_2 - 1)} \end{bmatrix}. \quad (31)$$

As a result, the probability  $\text{Prob}(\mathcal{U})$  defining the boundary of the capacity bound given by (27) can be evaluated as follows:

$$\begin{aligned} \text{Prob}(\mathcal{U}) &= \text{Prob}(R_1 \leq C'_1, R_2 \leq C'_2, R_1 + R_2 \leq C'_a) \\ &= \text{Prob}(e^{R_1} \leq t_1, e^{R_2} \leq t_2, e^{R_1 + R_2} \leq t_a) \\ &= \int_{e^{R_1}}^{\infty} dt_1 \int_{e^{R_2}}^{\infty} dt_2 \int_{e^{R_1 + R_2}}^{\infty} f(t_1, t_2, t_a) dt_a \quad (32) \\ &= T_1 + T_2 + T_3, \end{aligned}$$

where  $T_1$ ,  $T_2$ , and  $T_3$  in the above equation are defined as follows:

$$\begin{aligned} T_1 &= \int_{e^{R_1}}^{a - e^{R_1}} dt_1 \int_{e^{R_2}}^{a - t_1} dt_2 \int_{e^{R_1 + R_2}}^c f(t_1, t_2, t_a) dt_a, \\ T_2 &= \int_{e^{R_1}}^{a - e^{R_2}} dt_1 \int_{a - t_1}^{\infty} dt_2 \int_b^c f(t_1, t_2, t_a) dt_a, \quad (33) \\ T_3 &= \int_{a - e^{R_2}}^{\infty} dt_1 \int_{e^{R_2}}^{\infty} dt_2 \int_b^c f(t_1, t_2, t_a) dt_a. \end{aligned}$$

The variables  $a$ ,  $b$ , and  $c$  in the above equations are defined as follows:

$$\begin{aligned} a &= e^{R_1 + R_2} + 1, \\ b &= t_1 + t_2 - 1, \\ c &= t_1 + t_2 - 1 + (t_1 - 1)(t_2 - 1). \end{aligned} \quad (34)$$

In the appendix, we will show that  $T_2$  and  $T_3$  are fairly easy to evaluate. Following the steps used to prove Claim 5, it can be shown the following.

Claim 6.  $\mathcal{L}(p) = \{(R_1, R_2) \mid T_2 + T_3 \geq p\} \subset \mathcal{C}_{\mathcal{L}\mathcal{B}}(p) \subset \mathcal{C}(p)$ .

The rate region  $\mathcal{L}(p)$  gives another inner bound on the outage capacity region. The boundary of the region  $\mathcal{L}(p)$  is formed by the rate pairs that satisfy the following equation:

$$p = T_2 + T_3. \quad (35)$$

Given every possible  $R_1$ , we can trace out the corresponding  $R_2$  by solving (35) numerically using the expressions obtained for  $T_2$  and  $T_3$  in the appendix.

Although the approach used in this section may, in principle, also be applied to multiple-mobile cases, the increased complexity needed to derive the joint probability distribution functions and evaluate the relevant probabilities prevents us from obtaining closed-form expressions for multiple-mobile cases similar to those shown in (A.2) and (35).

#### 4. NUMERICAL RESULTS

The bounding techniques introduced in the previous section will now be applied to generate numerical results for inner and outer bounds on the outage capacity region for various values of number of antennas, outage percentage, and signal-to-noise ratio. Here, we focus our attention only on the two-mobile case since they are graphically friendly and offer significant physical insight.

Shown in Figure 3 is the outer bound on the outage capacity region with 2 and 16 antenna elements at the base station, plotted for outage probabilities of 1%, 10%, and 50%. In all cases, the signal-to-noise ratio  $\rho$  is 10 dB. Since these are outer bounds on the outage capacity region, no

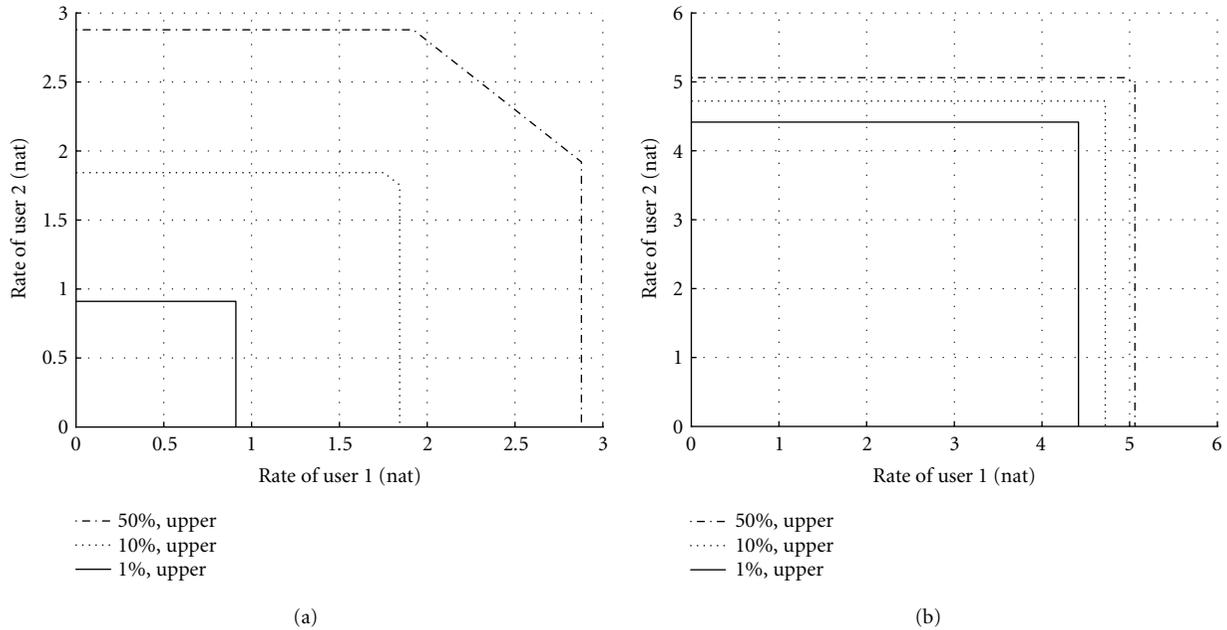


FIGURE 3: Outer bound on outage capacity regions for (a) 2 and (b) 16 antenna elements at the base station with SNR=10 dB for both cases.

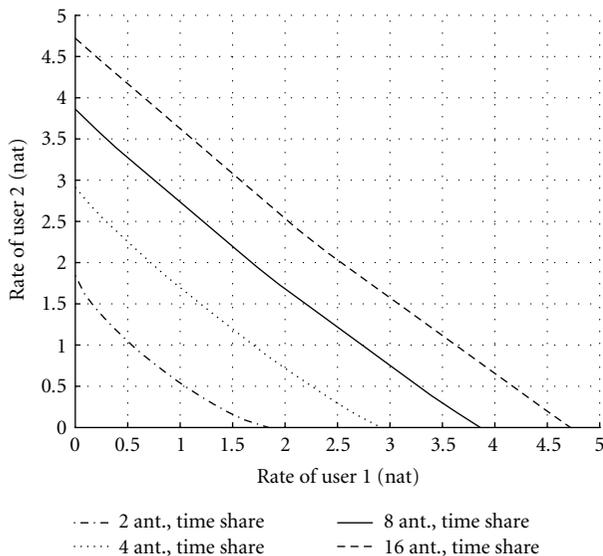


FIGURE 4: Time-share bound on 10% outage capacity region with SNR=10 dB.

rate pair outside the capacity region can ever be achieved with an outage probability smaller than the designated value  $p$ . Rate pairs inside the outer bound region may or may not be achievable with an outage probability smaller than  $p$ . For example, with 2 antenna elements at the base station, any rate pair with one rate higher than 0.9 nat per second (1 nat = 1.44 bytes) has an outage probability greater than 1%. We note that the number of antenna elements has a very significant effect, not only on significantly enlarging

the outer bound on the capacity region but also in reducing the differences between low outage and high outage objectives.

Figure 4 shows the time-share bound for the 10% outage capacity region for base station with 2, 4, 8, and 16 antenna elements, respectively. Since these are inner bounds, all rate pairs within the bound can be achieved with an outage probability smaller than 10%. For example, with 2 antenna elements, both mobiles can transmit at 0.76 nat per second while achieving an outage probability smaller than 10%. With 16 antenna elements at the base station, both mobiles can transmit at 2.25 nat per second while achieving an outage probability of 10%. We notice that all the time-sharing bounds are concave.

When both users are allowed to transmit at the same time and joint decoding is performed at the base station, we get a tighter inner bound on the outage capacity region. Figure 5 shows the joint decoding inner bounds for the 10% and 1% outage capacity regions with 2, 4, 8, and 16 antenna elements at the base station. With 2 antenna elements at the base station, both mobiles can simultaneously transmit at 0.72 nat per second with an outage probability smaller than 1%; with 16 antenna elements at the base station, the mobiles can simultaneously transmit at 2.65 nat per second with an outage probability smaller than 1%. The inner bound given by Claim 6 is obviously much tighter than the time-sharing bound in Claim 5. Unfortunately, this joint decoding bound cannot be easily obtained for more than two mobiles. Once again, we notice that increasing the number of antenna elements greatly reduces the separation between the 1% outage and 10% outage results.

Figure 6 shows both the outer and inner bounds on the

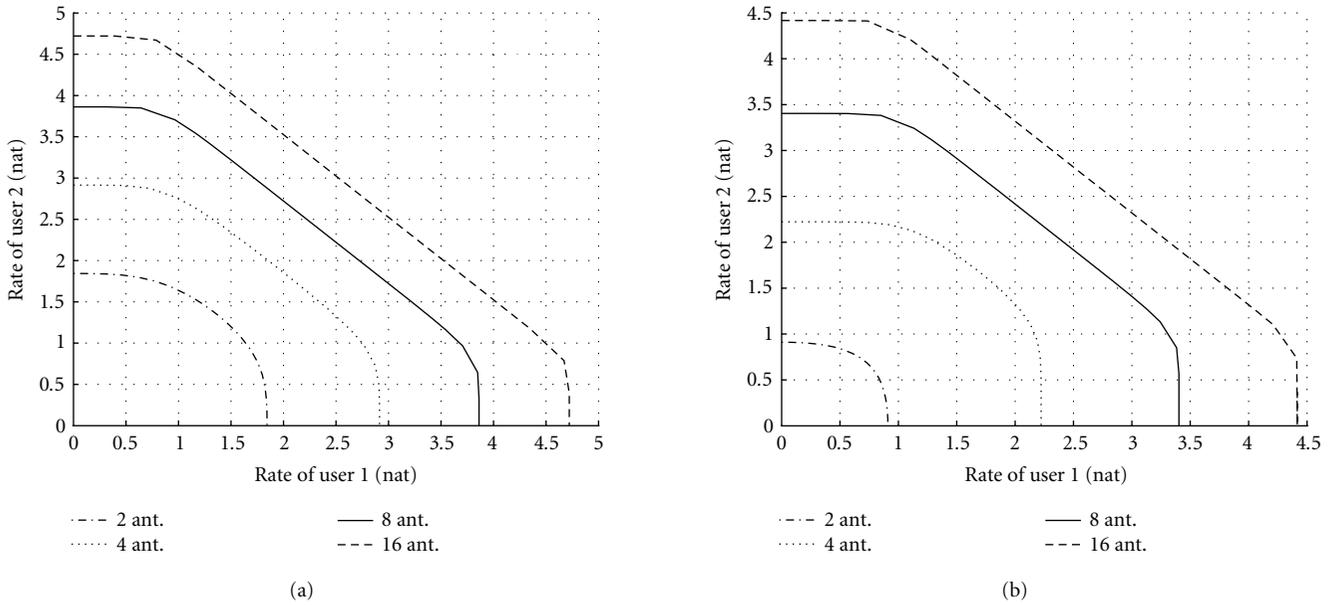


FIGURE 5: Inner bound on (a) 10% and (b) 1% outage capacity regions with different numbers of antenna elements at the base stations and where SNR=10 dB.

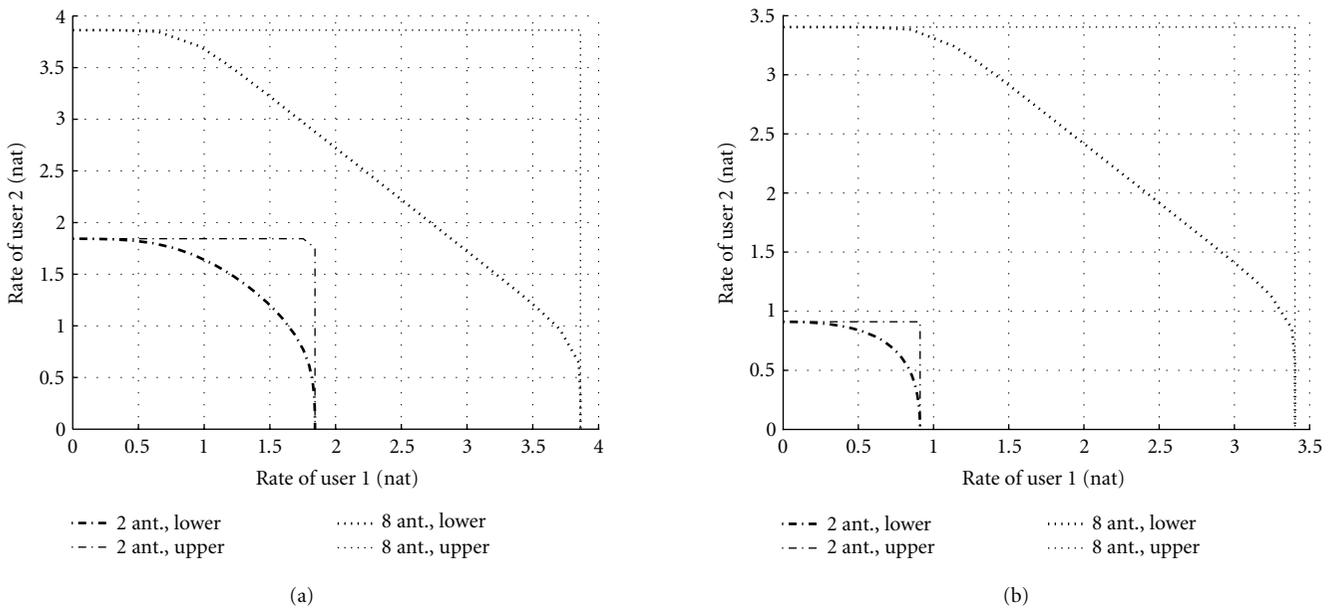


FIGURE 6: Both outer and inner bounds on (a) 10% and (b) 1% outage capacity regions for 2 and 8 antenna elements at the base stations with SNR=10 dB.

10% and 1% outage capacity regions with both 2 and 8 antenna elements. It can be seen from the plot that the inner bound is reasonably tight for the 2 antenna elements case. The bounds are very tight near the corners where one mobile is transmitting at its maximum allowable rate. We also notice that when the required outage probability is low, as in

the 1% case, there is more performance improvement by increasing the number of antenna elements than that when the required outage probability is high, as in the 10% case. This is no surprise since the cumulative distribution function of the allowable rates is much sharper when the number of antenna elements is large.

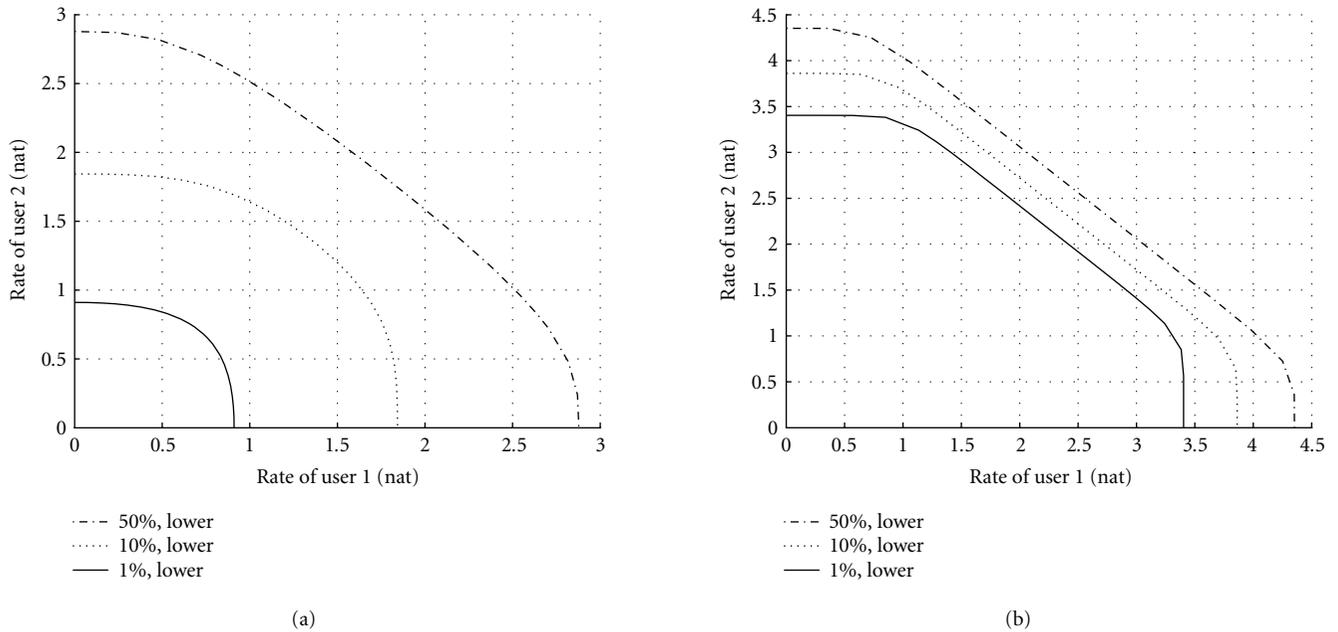


FIGURE 7: Inner bound on capacity region for (a) 2 and (b) 8 antenna elements under different outage probability with SNR=10 dB.

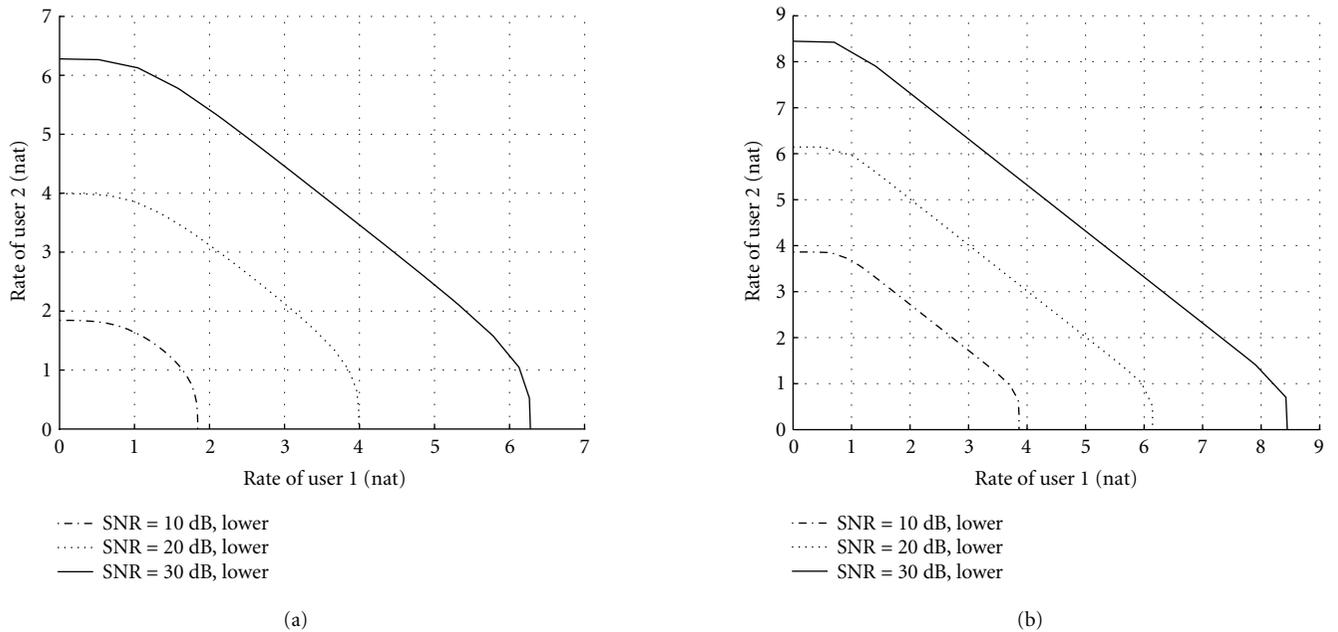


FIGURE 8: Inner bound on 10% outage capacity region under different SNRs for (a) 2 and (b) 8 antenna elements.

Figure 7 shows the tight inner bound for both 2 and 8 antenna elements at the base station with different outage probabilities. We can see from the plot that when the number of antenna elements is large, the outage capacity regions at different outage probabilities are not significantly different. This can be explained by noting that a large number of antenna elements at the base station is very efficient at combating severe fading, thereby keeping the allowable transmitting rates

relatively constant, and producing a sharper cumulative distribution function for the allowable rates.<sup>1</sup>

<sup>1</sup>Note that the capacity depends on chi-square distributed variables with mean  $M$  and variance  $2M$  through a log operation in (17). As a result, the distance between different outage capacities is determined by the ratio of different percentage points along the chi-square CDF. Or, equivalently, the distance is determined by the normalized CDF curve.

Figure 8 shows the tight inner bound for the 10% outage capacity region with 2 or 8 antenna elements at the base station and different SNRs. From (22), we would expect, for large SNR and any fading condition, that the capacity region should increase linearly in each dimension as  $\rho$  increases linearly in dB. As a result, we would also expect the outage capacity region to increase linearly in each dimension as  $\rho$  increases. This trend is confirmed from the regions shown in Figure 8.

## 5. CONCLUSION

In this paper, we have studied the use of MAE array at the base station to increase system capacity. The fundamental question addressed is the ultimate capacity achievable with a MAE array equipped base station communicating with multiple mobile stations. Since there are multiple mobiles, the capacity is expressed as a region over the space of transmission rates from the mobiles. Any particular set of rates contained in the region can be transmitted with a certain outage probability.

We have obtained both outer and inner bounds for the outage capacity regions. The outer bounds indicates what is beyond the capability of the SDMA system while the inner bounds indicates what is achievable. As expected, the use of multiple antennas at the base station can greatly increase the allowable transmission rates from the mobiles. For example, with 2 and 16 antenna elements and joint decoding at the base station, two mobiles can simultaneously transmit at 0.72 nat per second and 2.65 nat per second, respectively, and achieve an outage probability smaller than 1%.

Our results show that although the capacity region can be expanded by allowing higher outage probability, the increase in allowable rates is much greater when the number of antenna elements is small. In cases with a large number of antenna elements at the base station, the strong protection against severe fading provided by the antenna array can keep the maximum allowable rates relatively constant. We also observe that the outage capacity regions can increase dramatically with an increase in the signal-to-noise ratio; the capacity regions expand almost linearly in every dimension as the signal-to-noise ratio increases.

The capacity region bounds derived in this paper provide a yardstick against which the performance of any space-division multiple-access technique can be compared.

## APPENDIX

### EVALUATION OF $T_2$ AND $T_3$ IN SECTION 3.3

Now we examine  $T_2$  and  $T_3$  more closely. It is easily verified that

$$\begin{aligned} & \int_b^c h\left(\frac{2(t_1 + t_2 + 1 - t_a)}{(t_1 - 1)(t_2 - 1)} + 1\right) dt_a \\ &= \int_{-1}^1 h(x) \frac{(t_1 - 1)(t_2 - 1)}{2} dx \quad (\text{A.1}) \\ &= \frac{(t_1 - 1)(t_2 - 1)}{2}. \end{aligned}$$

Using this result and the complementary cumulative distribution function for chi-square random variables given in (18), we can further simplify the expressions for  $T_2$  and  $T_3$ :

$$\begin{aligned} T_3 &= \int_{a-e^{R_2}}^{\infty} dt_1 \int_{e^{R_2}}^{\infty} g\left(\frac{t_1 - 1}{\rho}\right) g\left(\frac{t_2 - 1}{\rho}\right) \frac{1}{\rho^2} dt_2 \\ &= \int_{(a-e^{R_2}-1)/\rho}^{\infty} g(x) dx \int_{(e^{R_2}-1)/\rho}^{\infty} g(y) dy \\ &= \bar{F}\left(\frac{a - e^{R_2} - 1}{\rho}\right) \bar{F}\left(\frac{e^{R_2} - 1}{\rho}\right), \\ T_2 &= \int_{(e^{R_1}-1)/\rho}^{(a-e^{R_2}-1)/\rho} g(x) dx \int_{(a-\rho x-2)/\rho}^{\infty} g(y) dy \\ &= \int_{x_1}^{x_2} g(x) dx \bar{F}\left(\frac{a-2}{\rho} - x\right) \\ &= \int_{x_1}^{x_2} g(x) \sum_{k=0}^{m-1} \frac{((a-2)/\rho - x)^k}{k!} e^{x-(a-2)/\rho} dx \\ &= \sum_{k=0}^{m-1} e^{-(a-2)/\rho} \int_{x_1}^{x_2} \frac{x^{m-1} ((a-2)/\rho - x)^k}{(m-1)!k!} dx \\ &= \sum_{k=0}^{m-1} e^{-(a-2)/\rho} \int_{x_1}^{x_2} \frac{x^{m-1}}{(m-1)!k!} \sum_{l=0}^k \frac{k!}{l!(k-l)!} \left(\frac{a-2}{\rho}\right)^{k-l} (-x)^l dx \\ &= \sum_{k=0}^{m-1} e^{-(a-2)/\rho} \sum_{l=0}^k \left(\frac{a-2}{\rho}\right)^{k-l} \frac{(-1)^l}{l!(k-l)!(m-1)!} \int_{x_1}^{x_2} x^{m+l-1} dx \\ &= \sum_{k=0}^{m-1} e^{-(a-2)/\rho} \sum_{l=0}^k \left(\frac{a-2}{\rho}\right)^{k-l} \frac{(-1)^l}{l!(k-l)!(m-1)!} \frac{(x_2^{n+l} - x_1^{n+l})}{n+l} \\ &= \sum_{l=0}^{m-1} e^{-(a-2)/\rho} \frac{(-1)^l (x_2^{n+l} - x_1^{n+l})}{l!(m-1)!(n+l)} \sum_{k=0}^{m-1} \left(\frac{a-2}{\rho}\right)^{k-l} \frac{1}{(k-l)!} \\ &= \sum_{l=0}^{m-1} e^{-(a-2)/\rho} \frac{(-1)^l}{l!(m-1)!(n+l)} \left(\frac{e^{R_1} - 1}{\rho}\right)^{n+l} (e^{(n+l)R_2} - 1) \\ &\quad \times \sum_{k=0}^{m-1-l} \left(\frac{a-2}{\rho}\right)^k \frac{1}{k!}. \end{aligned} \quad (\text{A.2})$$

## REFERENCES

- [1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, 1998.
- [2] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.
- [3] T. M. Cover and J. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, NY, USA, 1991.
- [4] R. Gallager, *Information Theory and Reliable Communication*, John Wiley & Sons, New York, NY, USA, 1968.
- [5] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: information-theoretic and communications aspects," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2619–2692, 1998.

- [6] D. N. C. Tse and S. V. Hanly, "Part I: Multiaccess fading channels. I. Polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Transactions on Information Theory*, vol. 44, no. 7, pp. 2796–2815, 1998.
- [7] S. V. Hanly and D. N. C. Tse, "Part II: Multiaccess fading channels. II. Delay-limited capacities," *IEEE Transactions on Information Theory*, vol. 44, no. 7, pp. 2816–2831, 1998.
- [8] L. Li and A. J. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels. I. Ergodic capacity," *IEEE Transactions on Information Theory*, vol. 47, no. 3, pp. 1083–1102, 2001.
- [9] L. Li and A. J. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels. II. Outage capacity," *IEEE Transactions on Information Theory*, vol. 47, no. 3, pp. 1103–1127, 2001.
- [10] E. Biglieri, G. Caire, and G. Taricco, "Limiting performance of block-fading channels with multiple antennas," *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1273–1289, 2001.
- [11] B. Suard, G. Xu, H. Liu, and T. Kailath, "Uplink channel capacity of space-division-multiple-access schemes," *IEEE Transactions on Information Theory*, vol. 44, no. 4, pp. 1468–1476, 1998.
- [12] W. Yu, W. Rhee, and J. Cioffi, "Optimal power control in multiple access fading channels with multiple antennas," in *Proc. IEEE International Conference on Communications (ICC '01)*, pp. 575–579, Helsinki, Finland, June 2001.
- [13] S. A. Jafar, S. Vishwanath, and A. Goldsmith, "Vector MAC capacity region with covariance feedback," in *Proc. IEEE International Symposium on Information Theory (ISIT '01)*, p. 54, Washington, DC, USA, June 2001.
- [14] A. S. Acampora, "The ultimate capacity of frequency reuse communication satellites," *Bell System Technical Journal*, vol. 59, no. 7, pp. 1089–1122, 1980.
- [15] J. Proakis, *Digital Communications*, McGraw-Hill, New York, NY, USA, 1995.
- [16] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York, NY, USA, 1991.

**Haipeng Jin** received his B.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 1996. He received his M.S. and Ph.D. degrees in electrical engineering from the University of California, San Diego, in 2000 and 2003, respectively. He is currently working with the Standards Group, Qualcomm Inc., San Diego, Calif, where he participates in the research and development of wireless Internet standards.



**Anthony Acampora** is a Professor of electrical and computer engineering at the University of California, San Diego, and is involved in numerous research projects addressing various issues at the leading edge of telecommunication networks. From 1995 through 1999, he was Director of UCSD's Center for Wireless Communications. Prior to joining the faculty at UCSD in 1995, he was a Professor of electrical engineering at Columbia University and Director of the Center for Telecommunications Research, a national engineering research center. He joined the faculty at Columbia in 1988 following a 20-year career at AT&T Bell Labs, most of which was spent in basic research as



a contributing researcher and research manager. At Columbia, he was involved in research and education programs concerning broadband networks, wireless access networks, network management, optical networks, and multimedia applications. He received his Ph.D. degree in electrical engineering from the Polytechnic Institute of Brooklyn and is a Fellow of the IEEE. Professor Acampora has published over 160 papers, holds 30 patents, and has authored a textbook entitled *An Introduction to Broadband Networks*.

# A Combined Antenna Arrays and Reverse-Link Synchronous DS-CDMA System over Frequency-Selective Fading Channels with Power Control Error

## Yong-Seok Kim

Communication System Lab., School of Electrical and Electronics Engineering, Yonsei University, 134 Sinchon-dong, Seodaemun-gu, Seoul 120-749, Korea  
Email: dragon@yonsei.ac.kr

## Seung-Hoon Hwang

Standardization and System Research Group, Mobile Communication Technology Research Laboratory, LG Electronics, 533 Hogye-dong, Dongan-gu, Anyang-shi, Kyungki-do, Korea  
School of Electronics and Computer Sciences, University of Southampton, Highfield, Southampton, SO17 1BJ, UK  
Email: shwang@ieee.org

## Hyo-Yol Park

Communication System Lab., School of Electrical and Electronics Engineering, Yonsei University, 134 Sinchon-dong, Seodaemun-gu, Seoul 120-749, Korea  
Email: seahog@commsys.yonsei.ac.kr

## Keum-Chan Whang

Communication System Lab., School of Electrical and Electronics Engineering, Yonsei University, 134 Sinchon-dong, Seodaemun-gu, Seoul 120-749, Korea  
Email: kcwhang@yonsei.ac.kr

Received 26 May 2003; Revised 28 January 2004

An improved antenna array (AA) has been introduced, in which reverse-link synchronous transmission technique (RLSTT) is incorporated to effectively make better an estimation of covariance matrices at a beamformer-RAKE receiver. While RLSTT is effective in the first finger at the RAKE receiver in order to reject multiple-access interference (MAI), the beamformer estimates the desired user's complex weights, enhancing its signal and reducing cochannel interference (CCI) from the other directions. In this work, it is attempted to provide a comprehensive analysis of user capacity which reflects several important factors such as the shape of multipath intensity profile (MIP), the number of antennas, and power control error (PCE). Theoretical analysis, confirmed by the simulations, demonstrates that the orthogonality provided by employing RLSTT along with AA may make the DS-CDMA system insensitive to the PCE even with fewer numbers of antennas.

**Keywords and phrases:** antenna arrays, reverse-link synchronous DS-CDMA, frequency-selective fading channel, power control error.

## 1. INTRODUCTION

DS-CDMA systems exhibit a user capacity limit in the sense that there exist a maximum number of users that can simultaneously communicate over multipath fading channels and maintain a specified level of performance per user. This limitation is caused by cochannel interference (CCI) which includes both multiple-access interference (MAI) between the

multiusers, and intersymbol interference (ISI) which arises from the existence of different transmission paths. A promising approach to increase the system capacity is the use of spatial processing with an antenna array (AA) at base station (BS) [1, 2, 3, 4, 5, 6]. Generally, the AA system consists of spatially distributed antennas and a beamformer which generates a weight vector to combine the array output. Several algorithms have been proposed in the spatial signal processing

to design the weights in the beamformer. For example, a new space-time processing framework for the beamforming with AA in DS-CDMA has been proposed in [2], where a code-filtering approach was used in each receiving antenna in order to estimate the optimum weights in the beamformer.

For a terrestrial mobile system, reverse-link synchronous transmission technique (RLSTT) has been proposed to reduce interchannel interference over a reverse link [7]. In the RLSTT, the synchronous transmission in the reverse link can be achieved by adaptively controlling the transmission time in each mobile station (MS). In a similar way to the closed-loop power control technique, the BS computes the time difference between the reference time generated in the BS and the arrival time of the dominant signal transmitted from each MS, and then transmits timing control bits, which order MSs to “advance” or “delay” their transmission times. The considered DS-CDMA system uses orthogonal reverse-link spreading sequences and the timing control algorithm that allows the main paths to be synchronized.

In this paper, an improved AA has been introduced, in which RLSTT is incorporated to effectively make better an estimation of covariance matrices at a Beamformer-RAKE receiver. While RLSTT is effective in the first finger at the RAKE receiver in order to reject MAI, the beamformer estimates the desired user’s complex weights, enhancing its signal and reducing CCI from the other directions. In this work, it is attempted to provide a comprehensive analysis of user capacity which reflects several important factors such as the shape of multipath intensity profile (MIP), the number of antennas, and power control error (PCE). Of particular interest are the trade-offs encountered among parameters such as the number of receiving antennas and PCE. The paper is organized as follows. In Section 2, channel and system models are described. The AA system with RLSTT is introduced and its theoretical analysis is derived to investigate the trade-offs among the system parameters in Section 3. Section 4 shows numerical results mainly focusing on the system capacity. Finally, a concluding remark is given in Section 5.

## 2. CHANNEL AND SYSTEM MODEL

We consider a BPSK-modulated DS-CDMA system over a multipath fading channel. Assuming  $K$  active users ( $k = 1, 2, \dots, K$ ), the low-pass equivalent signal transmitted by user  $k$  is presented as

$$s^{(k)}(t) = \sqrt{2P_k} b^{(k)}(t) g^{(k)}(t) a(t) \cos[\omega_c t + \phi^{(k)}], \quad (1)$$

where  $a(t)$  is a pseudonoise (PN) randomization sequence which is common to all the channels in a cell to maintain the CDMA orthogonality,  $g^{(k)}(t)$  is an orthogonal channelization sequence, and  $b^{(k)}(t)$  is user  $k$ ’s data waveform. In (1),  $P_k$  is the average transmitted power of the  $k$ th user,  $\omega_c$  is the common carrier frequency, and  $\phi^{(k)}$  is the phase angle of the  $k$ th modulator to be uniformly distributed in  $[0, 2\pi)$ . The orthogonal chip duration  $T_g$  and the PN chip interval  $T_c$  is related to data bit interval  $T$  through processing gain

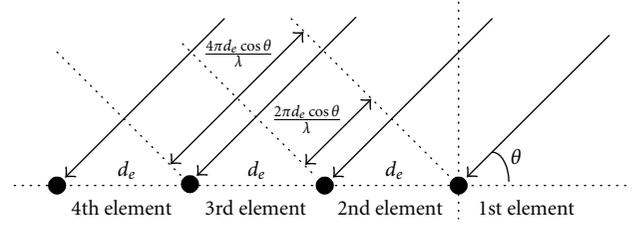


FIGURE 1: Antenna array model geometry.

$N = T/T_c$ . We assume, for simplicity, that  $T_g$  equals  $T_c$ . The complex lowpass impulse response of the vector channel associated with the  $k$ th user may be written as [3]

$$\mathbf{h}_k(\tau) = \sum_{l=0}^{L^{(k)}-1} \beta_l^{(k)} \exp(j\varphi_l^{(k)}) \mathbf{v}(\theta_l^{(k)}) \delta[\tau - \tau_l^{(k)}], \quad (2)$$

where  $\beta_l^{(k)}$  is the Rayleigh fading strength,  $\varphi_l^{(k)}$  is its phase shift, and  $\tau_l^{(k)}$  is the propagation delay. The  $k$ th user’s  $l$ th path array response vector is expressed as

$$\mathbf{v}(\theta_l^{(k)}) = \left[ 1 \exp\left(\frac{-j2\pi d \cos \theta_l^{(k)}}{\lambda}\right) \cdots \exp\left(\frac{-j2(M-1)\pi d \cos \theta_l^{(k)}}{\lambda}\right) \right]^T. \quad (3)$$

Throughout this paper, we consider that the array geometry, which is the parameter of the antenna aperture gain, is a uniform linear array (ULA) of  $M$  identical sensors in Figure 1. All signals from MS arrive at the BS AA with mean angle of arrival (AOA)  $\theta_l^{(k)}$  which is uniformly distributed in  $[0, \pi)$ . Assuming Rayleigh fading, the probability density function (pdf) of signal strength associated with the  $k$ th user’s  $l$ th propagation path,  $l = 0, 1, \dots, L^{(k)} - 1$ , is presented as

$$p(\beta_l^{(k)}) = \frac{2\beta_l^{(k)}}{\Omega_l^{(k)}} \exp\left(-\frac{(\beta_l^{(k)})^2}{\Omega_l^{(k)}}\right), \quad (4)$$

where  $\Omega_l^{(k)}$  is the second moment of  $\beta_l^{(k)}$  with  $\sum_{l=0}^{\infty} \Omega_l = 1$ , and we assume it is related to the second moment of the initial path strength  $\Omega_0^{(k)}$  for exponentially decaying MIP as

$$\Omega_l^{(k)} = \Omega_0^{(k)} \exp(-l\delta), \quad \text{for } 0 < l \leq L^{(k)} - 1, \delta \geq 0, \quad (5)$$

where  $\delta$  reflects the rate at which the decay of average path strength as a function of path delay occurs. Note that a more realistic profile model may be the exponential MIP.

The receiver is a coherent RAKE receiver with AA, where the number of fingers  $L_r$  is a variable less than or equal to  $L^{(k)}$  which is the number of resolvable propagation paths associated with the  $k$ th user. Perfect estimates of the channel

parameters are assumed. The complex received signal is expressed as

$$\begin{aligned} \mathbf{r}(t) = & \sqrt{2P} \sum_{k=1}^K \sqrt{\lambda_k} \sum_{l=0}^{L^{(k)}-1} \beta_l^{(k)} \mathbf{V}(\theta_l^{(k)}) b^{(k)}(t - \tau_l^{(k)}) \\ & \times g^{(k)}(t - \tau_l^{(k)}) a(t - \tau_l^{(k)}) \cos[\omega_c t + \psi_l^{(k)}] + \mathbf{n}(t), \end{aligned} \quad (6)$$

where  $P$  is the average received power and  $\psi_l^{(k)}$  is the phase of the  $l$ th path associated to the  $k$ th carrier.  $\lambda_k$  corresponds to the PCE of the  $k$ th user which is a random variable due to imperfect power control [8]. We consider  $\lambda_k$  to be log-normally distributed with standard deviation  $\sigma_{\lambda_k}$  dB. In other words,  $\lambda_k = 10^{(x/10)}$ , where the variable  $x$  follows a normal distribution.  $\mathbf{n}(t)$  is an  $M \times 1$  spatially and temporally white Gaussian noise vector with a zero mean and covariance which is given by  $E\{\mathbf{n}(t)\mathbf{n}^H(t)\} = \sigma_n^2 \mathbf{I}$ , where  $\mathbf{I}$  is the  $M \times M$  identity matrix,  $\sigma_n^2$  is the antenna noise variance with  $\eta_0/2$ , and the superscript  $H$  denotes the Hermitian-transpose operator. When the received signal is matched to the reference user's code, the  $l$ th multipath matched filter output for the interest user ( $k = 1$ ) can be expressed as

$$\begin{aligned} \mathbf{y}_l^{(1)} = & \int_{\tau_l^{(1)}}^{\tau_l^{(1)}+T} \mathbf{r}(t) \cdot g^{(1)}(t - \tau_l^{(1)}) a(t - \tau_l^{(1)}) \cos[\omega_c t + \psi_l^{(1)}] dt \\ = & \mathbf{S}_l^{(1)} + \mathbf{I}_{l,\text{mai}}^{(1)} + \mathbf{I}_{l,\text{si}}^{(1)} + \mathbf{I}_{l,\text{ni}}^{(1)}. \end{aligned} \quad (7)$$

When a reference signal is not available, a common criterion for optimizing the weight vectors and this criterion is to maximize the signal-to-interference plus noise ratio (SINR). In (7),  $\mathbf{u}_l^{(1)} = \mathbf{I}_{l,\text{si}}^{(1)} + \mathbf{I}_{l,\text{mai}}^{(1)} + \mathbf{I}_{l,\text{ni}}^{(1)}$  is a total interference plus noise for the  $l$ th path of interest user. By solving the following problem, we can obtain the optimal weights to maximize the SINR [9]:

$$\mathbf{W}_{l(\text{opt})}^{(1)} = \max_{\mathbf{W} \neq \mathbf{0}} \frac{\mathbf{W}_l^{(1)H} \mathbf{R}_{yy} \mathbf{W}_l^{(1)}}{\mathbf{W}_l^{(1)H} \mathbf{R}_{uu} \mathbf{W}_l^{(1)}}, \quad (8)$$

where  $\mathbf{R}_{yy}$  and  $\mathbf{R}_{uu}$  are the second-order correlation matrices of the received signal subspace and the interference plus noise subspace, respectively. Here,  $\mathbf{R}_{uu}$  can be estimated by the code-filtering approach in [2], which is presented as

$$\mathbf{R}_{uu} = \frac{N}{N-1} \left( \mathbf{R}_{rr} - \frac{1}{N} \mathbf{R}_{yy} \right), \quad (9)$$

where  $\mathbf{R}_{rr}$  means the covariance matrix of the received signal prior to RAKE. The solution corresponds to the largest eigenvalue ( $\lambda_{\max}$ ) of the generalized eigenvalue problem in the matrix pair  $(\mathbf{R}_{yy}, \mathbf{R}_{uu})$ . Therefore, we can obtain the maximum SINR when the weight vector  $\mathbf{W}_{l(\text{opt})}^{(1)}$  equals the principal eigenvector of the matrix pair, which is presented as

$$\mathbf{R}_{yy} \cdot \mathbf{W}_{l(\text{opt})}^{(1)} = \lambda_{\max} \cdot \mathbf{R}_{uu} \cdot \mathbf{W}_{l(\text{opt})}^{(1)}. \quad (10)$$

From (7) and (8), the corresponding beamformer output for the  $l$ th path of interest user is

$$\begin{aligned} \hat{z}_l^{(1)} = & \mathbf{W}_l^{(1)H} \cdot \mathbf{y}_l^{(1)} \\ = & \hat{S}_l^{(1)} + \hat{I}_{l,\text{mai}}^{(1)} + \hat{I}_{l,\text{si}}^{(1)} + \hat{I}_{l,\text{ni}}^{(1)}, \end{aligned} \quad (11)$$

where

$$\begin{aligned} \hat{S}_l^{(1)} = & \sqrt{P\lambda_1/2} \beta_l^{(1)} C_{ll}^{(1,1)} b_0^{(1)T}, \\ \hat{I}_{l,\text{mai}}^{(1)} = & \sqrt{P/2} \sum_{k=2}^K \sqrt{\lambda_k} \sum_{j=0}^{L^{(k)}-1} \beta_j^{(k)} C_{lj}^{(1,k)} \\ & \times \left\{ b_{-1}^{(k)} RW_{k1}[\tau_j^{(k)}] + b_0^{(k)} \widehat{RW}_{k1}[\tau_j^{(k)}] \right\} \cos[\Psi_{lj}^{(k)}], \\ \hat{I}_{l,\text{si}}^{(1)} = & \sqrt{P\lambda_1/2} \sum_{\substack{j=0 \\ j \neq l}}^{L^{(1)}-1} \beta_j^{(1)} C_{lj}^{(1,1)} \left\{ b_{-1}^{(1)} RW_{11}[\tau_j^{(1)}] \right. \\ & \left. + b_0^{(1)} \widehat{RW}_{11}[\tau_j^{(1)}] \right\} \cos[\Psi_{lj}^{(1)}], \\ \hat{I}_{l,\text{ni}}^{(1)} = & \int_{\tau_l^{(1)}}^{\tau_l^{(1)}+T} \mathbf{W}_l^{(1)H} \cdot \mathbf{n}(t) g^{(1)}(t - \tau_l^{(1)}) \\ & \times a(t - \tau_l^{(1)}) \cos[\omega_c t + \psi_l^{(1)}] dt, \end{aligned} \quad (12)$$

with  $b_0^{(1)}$  being the information bit to be detected,  $b_{-1}^{(1)}$  the preceding bit,  $\tau_{ij}^{(k)} = \tau_j^{(k)} - \tau_l^{(1)}$ , and  $\psi_{ij}^{(k)} = \psi_j^{(k)} - \psi_l^{(1)}$ .  $\mathbf{W}_l^{(1)} = [w_{l,1}^{(1)} w_{l,2}^{(1)} \cdots w_{l,M}^{(1)}]^T$  is the  $M \times 1$  weight vector for the  $l$ th path of the first user.  $C_{lj}^{(1,k)} = \mathbf{W}_l^{(1)H} \cdot \mathbf{V}(\theta_j^{(k)})$  represents the spatial correlation between the array response vector of the  $k$ th user at the  $j$ th multipath and the weight vector of the interest user at the  $l$ th path.  $RW$  and  $\widehat{RW}$  are Walsh-PN continuous partial cross-correlation functions defined by  $RW_{k1}(\tau) = \int_0^T g^{(k)}(t - \tau) a(t - \tau) \cdot g^{(1)}(t) a(t) dt$  and  $\widehat{RW}_{k1}(\tau) = \int_\tau^T g^{(k)}(t - \tau) a(t - \tau) g^{(1)}(t) a(t) dt$ . From (11), we can obtain the Rake receiver output from MRC combining  $\hat{z}^{(1)} = \sum_{l=0}^{L_r} \beta_l^{(1)} \cdot \hat{z}_l^{(1)}$  and see that the outputs of the  $l$ th branch,  $l = 0, 1, \dots, L_r - 1$ , consist of four terms. The first term represents the desired signal component to be detected. The second term represents the MAI from  $(K - 1)$  other simultaneous users in the system. The third term is the self-interference (SI) for the reference user. Finally, the last term is AWGN.

### 3. PERFORMANCE OF AA WITH RLSTT IN RAYLEIGH FADING CHANNEL WITH PCE

In our analysis, the evaluation is carried out for the case in which the arrival time of paths is modeled as synchronous in the first branch (i.e., for main paths) but as asynchronous in the rest of the branches (i.e., for multipaths). With the well-known Gaussian approximation, we model the MAI terms in the first branch and the other branches as a Gaussian process with variances equal to the MAI variances for  $l = 0$  and for  $l \geq 1$ , respectively. Extending the derived results in [7], the

variance of MAI for  $l = 0$ , conditioned on the values of  $\beta_l^{(1)}$  and  $\lambda_k$ , is

$$\bar{\sigma}_{\text{mai},0}^2 = \frac{E_b T (2N - 3)}{12N(N - 1)} \left\{ \beta_0^{(1)} \right\}^2 \sum_{k=2}^K \lambda_k \sum_{j=1}^{L^{(k)}-1} \Omega_j^{(k)} \zeta_{0j}^{(1,k)^2}. \quad (13)$$

Similarly, the variance of MAI for  $l \geq 1$  is

$$\bar{\sigma}_{\text{mai},l}^2 = \frac{E_b T (N - 1)}{6N^2} \left\{ \beta_l^{(1)} \right\}^2 \sum_{k=2}^K \lambda_k \sum_{j=0}^{L^{(k)}-1} \Omega_j^{(k)} \zeta_{lj}^{(1,k)^2}, \quad (14)$$

where  $E_b = PT$  is the signal energy per bit, and  $\zeta_{lj}^{(1,k)^2} = E[\{C_{lj}^{(1,k)}\}^2]$  is the second-order characterization of the spatial correlation between the array response vector of the  $k$ th user at the  $j$ th multipath and the weight vector of interest user at the  $l$ th path, of which more detailed derivation is described in the appendix. The conditional variance of  $\bar{\sigma}_{\text{si},l}^2$  is approximated by [10]:

$$\bar{\sigma}_{\text{si},l}^2 \approx \frac{E_b \lambda_l T}{4N} \left\{ \beta_l^{(1)} \right\}^2 \sum_{\substack{j=0 \\ j \neq l}}^{L^{(l)}-1} \Omega_j^{(l)} \zeta_{lj}^{(1,l)^2}. \quad (15)$$

The variance of the AWGN term, conditioned on the value of  $\beta_l^{(1)}$ , is calculated as

$$\bar{\sigma}_{\text{ni},l}^2 = \frac{T \eta_0 \zeta_{ll}^{(1,1)^2}}{4M} \cdot \left\{ \beta_l^{(1)} \right\}^2. \quad (16)$$

Therefore, the output of the receiver is a Gaussian random process with mean

$$U_s = \sqrt{\frac{E_b \lambda_l T}{2}} \sum_{l=0}^{L_r-1} \left\{ \beta_l^{(1)} \right\}^2 \zeta_{ll}^{(1,1)} \quad (17)$$

and the total variance equal to the sum of the variance of all the interference and noise terms. From (13), (14), (15), and (16), we have

$$\begin{aligned} \bar{\sigma}_T^2 &= \bar{\sigma}_{\text{mai},0}^2 + \sum_{l=1}^{L_r-1} \bar{\sigma}_{\text{mai},l}^2 + \sum_{l=0}^{L_r-1} (\bar{\sigma}_{\text{si},l}^2 + \bar{\sigma}_{\text{ni},l}^2) \\ &= E_b T \Omega_0 \\ &\times \left\{ \frac{(2N - 3) \{q(L_r, \delta) - 1\} \lambda_l \zeta_0^2 \cdot \left\{ \beta_0^{(1)} \right\}^2}{12N(N - 1)} \right. \\ &+ \frac{(N - 1) q(L_r, \delta) \lambda_l \zeta^2 \cdot \sum_{l=1}^{L_r-1} \left\{ \beta_l^{(1)} \right\}^2}{6N^2} \\ &+ \frac{\lambda_l \{q(L_r, \delta) - 1\} \left( \zeta_0^2 \cdot \left\{ \beta_0^{(1)} \right\}^2 + \zeta^2 \cdot \sum_{l=1}^{L_r-1} \left\{ \beta_l^{(1)} \right\}^2 \right)}{4N} \\ &\left. + \frac{\eta_0 \left( \zeta_0'^2 \cdot \left\{ \beta_0^{(1)} \right\}^2 + \zeta'^2 \cdot \sum_{l=1}^{L_r-1} \left\{ \beta_l^{(1)} \right\}^2 \right)}{4ME_b \Omega_0} \right\}. \quad (18) \end{aligned}$$

At the output of the receiver, SNR may be written in a more compact form as  $\gamma_s$ :

$$\begin{aligned} \gamma_s &= \left\{ \frac{(2N - 3) \{q(L_r, \delta) - 1\} \lambda_l \cdot \zeta_0^2 \cdot \left\{ \beta_0^{(1)} \right\}^2}{6N(N - 1) \cdot \zeta_0' \cdot \left\{ \beta_0^{(1)} \right\}^2 + \zeta' \cdot \sum_{l=1}^{L_r-1} \left\{ \beta_l^{(1)} \right\}^2} \right. \\ &+ \frac{(N - 1) q(L_r, \delta) \lambda_l \cdot \zeta^2 \cdot \sum_{l=1}^{L_r-1} \left\{ \beta_l^{(1)} \right\}^2}{3N^2 \cdot \zeta_0' \cdot \left\{ \beta_0^{(1)} \right\}^2 + \zeta' \cdot \sum_{l=1}^{L_r-1} \left\{ \beta_l^{(1)} \right\}^2} \\ &+ \frac{\{q(L_r, \delta) - 1\} \lambda_l \cdot \zeta_0^2 \cdot \left\{ \beta_0^{(1)} \right\}^2 + \zeta^2 \cdot \sum_{l=1}^{L_r-1} \left\{ \beta_l^{(1)} \right\}^2}{2N \cdot \zeta_0' \cdot \left\{ \beta_0^{(1)} \right\}^2 + \zeta' \cdot \sum_{l=1}^{L_r-1} \left\{ \beta_l^{(1)} \right\}^2} \\ &\left. + \frac{\eta_0 \cdot \left( \zeta_0'^2 \cdot \left\{ \beta_0^{(1)} \right\}^2 + \zeta'^2 \cdot \sum_{l=1}^{L_r-1} \left\{ \beta_l^{(1)} \right\}^2 \right)^{-1}}{2M \Omega_0 E_b \cdot \zeta_0' \cdot \left\{ \beta_0^{(1)} \right\}^2 + \zeta' \cdot \sum_{l=1}^{L_r-1} \left\{ \beta_l^{(1)} \right\}^2} \right\} \\ &\times \frac{\lambda_l \left( \zeta_0' \cdot \left\{ \beta_0^{(1)} \right\}^2 + \zeta' \cdot \sum_{l=1}^{L_r-1} \left\{ \beta_l^{(1)} \right\}^2 \right)}{\Omega_0}, \quad (19) \end{aligned}$$

where  $q(L_r, \delta) = \sum_{l=0}^{L_r-1} \exp(-l\delta) = 1 - \exp(-L_r\delta)/1 - \exp(-\delta)$ ,  $\lambda_l = \sum_{k=2}^K \lambda_k$ , and  $\Omega_0 = \Omega_0$ .  $\zeta_{ij}^{(k,m)^2} = \zeta_0^2$  when  $k \neq m$  or  $l \neq j$  for  $l = 0$ ,  $\zeta_{ij}^{(k,m)^2} = \zeta^2$  when  $k \neq m$  or  $l \neq j$  for  $l > 0$ ,  $\zeta_{ij}^{(k,m)^2} = \zeta_0'^2$  when  $k = m$  and  $l = j$  for  $l = 0$ , and  $\zeta_{ij}^{(k,m)^2} = \zeta'^2$  when  $k = m$  and  $l = j$  for  $l > 0$ . In [11], the pdf of  $\lambda_l = \sum_{k=2}^K \lambda_k$  for  $K - 1$  users is an approximately lognormal distribution, with the following logarithmic mean and variance, which is presented as

$$p(\lambda_l) = \frac{1}{\sqrt{2\pi} \sigma_{\lambda_l} \lambda_l} \exp \left[ -\frac{(\ln \lambda_l - m_{\lambda_l})^2}{2\sigma_{\lambda_l}^2} \right], \quad (20)$$

where

$$\begin{aligned} \sigma_{\lambda_l}^2 &= \ln \left( \frac{1}{K - 1} \exp(\sigma_{\lambda_l}^2) + \frac{K - 2}{K - 1} \right), \\ m_{\lambda_l} &= \ln(K - 1) + m + \frac{\sigma_{\lambda_l}^2}{2} \\ &- \frac{1}{2} \ln \left( \frac{K - 2}{K - 1} + \frac{1}{K - 1} \exp(\sigma_{\lambda_l}^2) \right). \quad (21) \end{aligned}$$

This method is valid for a logarithmic standard deviation  $\sigma_{\lambda_l}$  less than 4 dB. To evaluate the average bit error probability,  $P_e^l(\lambda_1, \lambda_l)$ , conditioning on the values of  $\lambda_1$  and  $\lambda_l$  follows as

$$\begin{aligned} P_e^l(\lambda_1, \lambda_l) &= \int_0^\infty \int_0^\infty Q(\sqrt{\gamma_s}) \sum_{k=1}^{L_r-1} \frac{\pi_k}{\Omega_k} \exp(-x/\Omega_k) \cdot \frac{1}{\Omega_0} \exp(-y/\Omega_0) dx dy, \quad (22) \end{aligned}$$

where  $\pi_k = \prod_{i=1, i \neq k}^{L_r-1} (x_k / (x_k - x_i)) = \prod_{i=1, i \neq k}^{L_r-1} (\Omega_k / (\Omega_k - \Omega_i))$ ,  $Q(x) = (1/\sqrt{2\pi}) \int_x^\infty \exp(-u^2/2) du$ . The average bit error

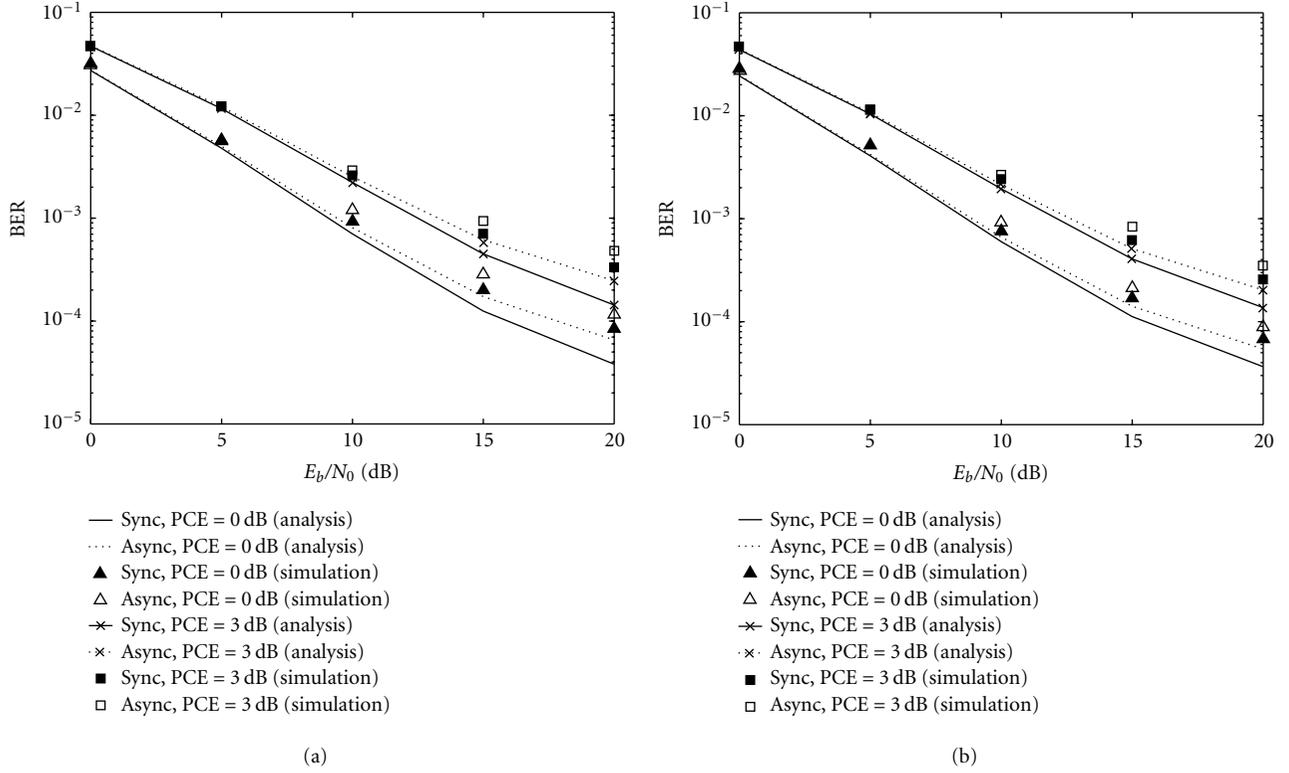


FIGURE 2: Analytical results versus simulation results. (Number of users = 12,  $M = 4$ ,  $L_r = L^{(k)} = 2$ , PCE = 0 and 3 dB.) (a)  $\delta = 1.0$ , (b)  $\delta = 0.2$ .

probability  $P_e$  is calculated as

$$P_e = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} P_e^l(\exp(\sqrt{2}\sigma_{\lambda_1} z_1 + m_{\lambda_1}), \exp(\sqrt{2}\sigma_{\lambda_l} z_l + m_{\lambda_l})) \times \exp[-z_1^2] dz_1 \exp[-z_l^2] dz_l, \quad (23)$$

where  $z_1 = (\ln \lambda_1 - m_{\lambda_1})/\sqrt{2}\sigma_{\lambda_1}$  and  $z_l = (\ln \lambda_l - m_{\lambda_l})/\sqrt{2}\sigma_{\lambda_l}$ . This integration can be easily obtained by using the Hermite polynomial approach, which requires only summation and no integration [12]:

$$P_e = \frac{1}{\pi} \sum_{l=1}^h w_l \sum_{n=1}^h w_n P_e^l(\exp(\sqrt{2}\sigma_{\lambda_1} x_n + m_{\lambda_1}), \exp(\sqrt{2}\sigma_{\lambda_l} x_l + m_{\lambda_l})). \quad (24)$$

#### 4. NUMERICAL RESULTS

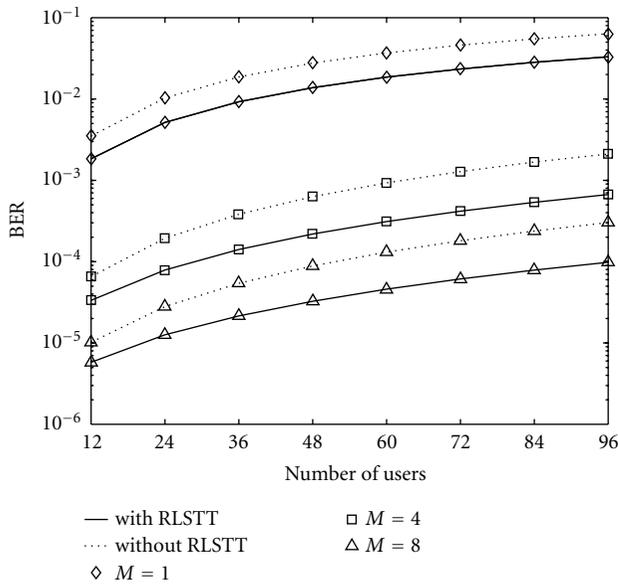
In this section, we have investigated the user capacity of AA system both with RLSTT and without RLSTT, considering several important factors such as the shape of MIP, the number of antennas, and the PCE. In all evaluations, processing gain is assumed to be 128, and the number of paths and taps

in RAKE is assumed to be the same for all users and denoted by two. The decaying factor is considered as 1.0 or 0.2 for the exponential MIP. The sensor spacing is half of the carrier wavelength.

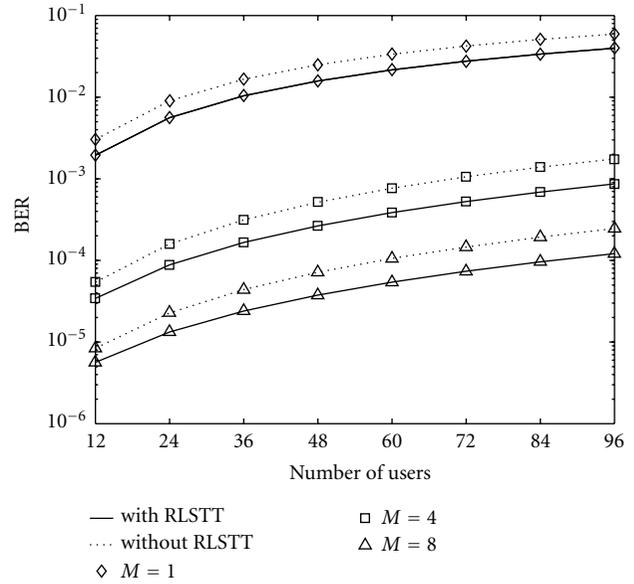
Figure 2 shows uncoded BER performance as a function of  $E_b/N_0$ , when the number of users is twelve and the number of antennas is four in the exponential MIP. Two decay factors are considered, and both perfect power control (PCE = 0 dB) and imperfect power control (PCE = 3 dB) are assumed. The results confirm that the analytical results are well matched to the simulation results. It is noted that using RLSTT together with AA may enhance the performance, since RLSTT tends to make better the estimation of covariance matrices for beamformer-RAKE receiver.

The BER curves are plotted as functions of the number of users in Figure 3 when  $E_b/N_0 = 20$  dB and power control is perfect. The number of antennas is chosen among one, four, or eight. It is shown that AA with RLSTT demonstrates significant performance gain when the number of users increases, even though the performance improvement decreases when the number of antenna increases. For example, in the case of four antennas, while AA without RLSTT supports 60 users, AA with RLSTT supports more than 96 users at a BER of  $10^{-3}$ , showing an enhancement of 50%.

Figure 4 shows the BER system performance as a function of the number of users, when  $M = 4$ ,  $E_b/N_0 = 20$  dB,

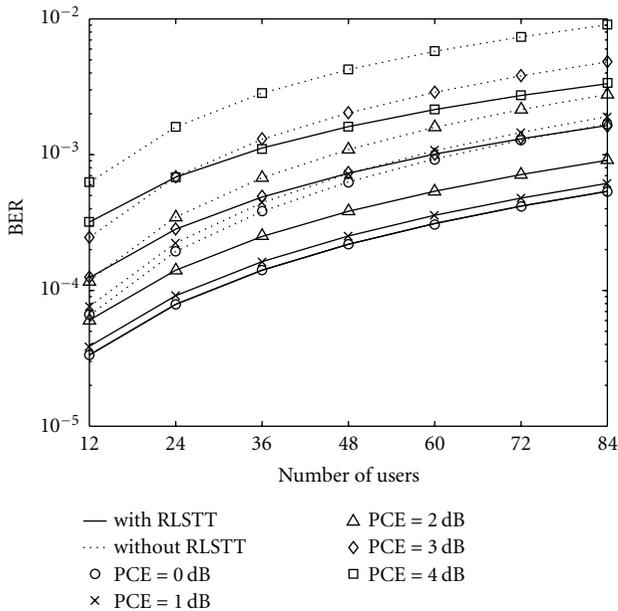


(a)

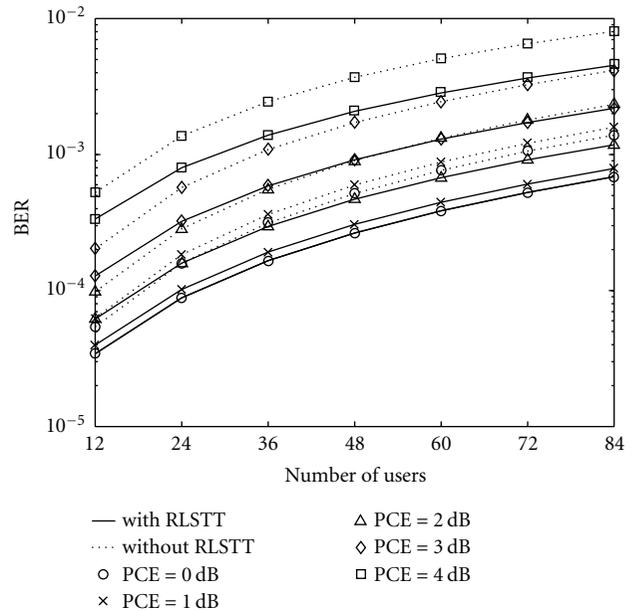


(b)

FIGURE 3: BER versus number of users in AA with RLSTT and AA without RLSTT ( $E_b/N_0 = 20$  dB,  $M = 1, 4,$  and  $8$ ,  $L_r = L^{(k)} = 2$ , PCE = 0 dB). (a)  $\delta = 1.0$ , (b)  $\delta = 0.2$ .



(a)



(b)

FIGURE 4: BER versus number of users in AA with RLSTT and AA without RLSTT ( $E_b/N_0 = 20$  dB,  $M = 4$ ,  $L_r = L^{(k)} = 2$ , PCE = 0, 1, 2, 3, and 4 dB). (a)  $\delta = 1.0$ , (b)  $\delta = 0.2$ .

and power control is imperfect. The curves are parameterized by different PCE values such as PCE = 0, 1, 2, 3, and 4 [dB], and show that RLSTT makes DS-CDMA system with AA insensitive to the PCE and thus increases the achievable overall system capacity. At BER =  $5 \times 10^{-4}$ , AA with RLSTT

when PCE = 2 dB can support even greater number of users, about 35% more than AA without RLSTT when power control is perfect (PCE = 0 dB), even though its capacity when PCE = 2 dB is degraded about 28% in comparison to the perfect power control.

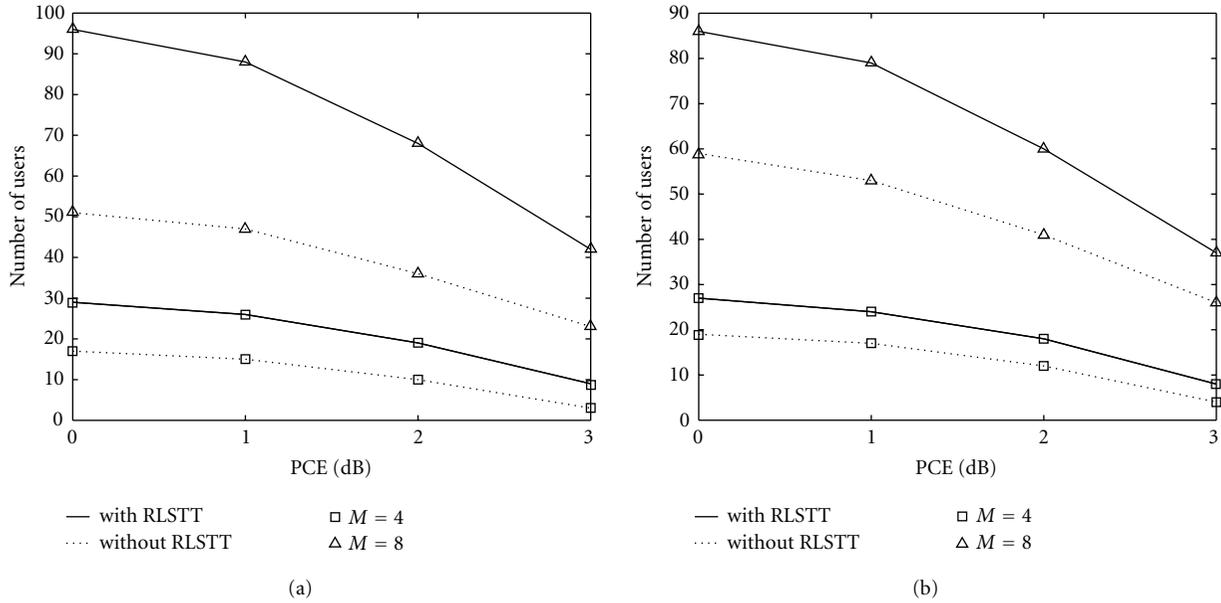


FIGURE 5: Number of users versus PCE in AA with RLSTT and AA without RLSTT ( $E_b/N_0 = 20$  dB,  $M = 4$  and  $8$ ,  $L_r = L^{(k)} = 2$ ,  $\text{BER} = 10^{-4}$ ) (a)  $\delta = 1.0$  (b)  $\delta = 0.2$ .

In Figure 5, the maximum allowable number of users to achieve BER of  $10^{-4}$  is shown as a function of PCE when the number of antenna elements is four or eight. The figure demonstrates while in eight-element AA without RLSTT PCE is required to keep less than 1 dB in order to achieve the user capacity of 50 users, AA with RLSTT may make loose the requirement to 3 dB. The figure can also be used to find the overall system capacity for a given PCE and the number of antenna elements. These results, however, do not take into account effects such as coding and interleaving. Additionally, it is apparent that RLSTT has superior performance and/or reduces the complexity of the system since AA with RLSTT with fewer numbers of antennas can obtain better performance than AA without RLSTT.

## 5. CONCLUSIONS

In this paper, we presented an improved AA, in which RLSTT is incorporated to effectively make better an estimation of covariance matrices at a beamformer-RAKE receiver, and investigated the user capacity and the performance analysis which reflects several important factors such as the shape of multipath intensity profile (MIP), the number of antennas, and power control error (PCE). The results show that the orthogonality provided by employing RLSTT along with AA may make the DS-CDMA system insensitive to the PCE even with fewer numbers of antennas. Additionally, RLSTT has superior performance and/or reduces the complexity of the system since AA with RLSTT with fewer numbers of antennas can obtain better performance than AA without RLSTT. The consideration of estimation technique such as diagonal loading employed in the proposed system may be an interesting issue for future study.

## APPENDIX

### SPATIAL CORRELATION STATISTICS

From (10), we can obtain the optimal beamformer weight presented as

$$\mathbf{W}_l^{(k)} = \xi \cdot \mathbf{R}_{uu,l}^{(k)-1} \mathbf{V}(\theta_l^{(k)}); \quad (\text{A.1})$$

since  $\xi$  does not affect the SINR, we can set  $\xi = 1$ . When the total number of paths is large, a large code length yields  $\mathbf{R}_{uu,l}^{(k)} = \sigma_{s,l}^{(k)^2} \cdot \mathbf{I}$  [2]. However, it means that the total undesired signal vector can be modeled as a spatially white Gaussian random vector. Here,  $\sigma_{s,l}^{(k)^2}$  is the total interference-plus-noise power. From (7), the total interference-plus-noise for the  $l$ th path of the  $k$ th user in the matched filter output is shown as

$$\mathbf{u}_l^{(k)} = \mathbf{I}_{l,\text{si}}^{(k)} + \mathbf{I}_{l,\text{mai}}^{(k)} + \mathbf{I}_{l,\text{ni}}^{(k)}. \quad (\text{A.2})$$

If we assume that the angles of arrival of the multipath components are uniformly distributed over  $[0, \pi]$ , the total interference vector  $\mathbf{I}_{l,\text{si}}^{(k)} + \mathbf{I}_{l,\text{mai}}^{(k)}$  will be spatially white [2, Chapter 6]. In this case, the variance of the undesired signal vector is calculated as

$$\begin{aligned} E[\mathbf{u}_l^{(k)} \cdot \mathbf{u}_l^{(k)H}] &= \sigma_{s,l}^{(k)^2} \cdot \mathbf{I} \\ &= (\sigma_{\text{mai},l}^{(k)^2} + \sigma_{\text{si},l}^{(k)^2} + \sigma_{\text{ni},l}^{(k)^2}) \cdot \mathbf{I}, \end{aligned} \quad (\text{A.3})$$

where  $\sigma_{\text{mai},l}^{(k)^2}$  and  $\sigma_{\text{si},l}^{(k)^2}$  are the noise variance of MAI and SI in one-dimension antenna system. For the RLSTT model [7], all active users are synchronous in the first branch. Therefore,

we can obtain the different variance of the total interference-plus-noise for  $l = 0$  and for  $l \geq 1$ , conditions on the value of  $\lambda_k$ , respectively, expressed as follows:

$$\begin{aligned} \sigma_{s,0}^{(k)^2}(\lambda_1, \lambda_I) &= E_b T \Omega_0 \left( \frac{(2N-3)\lambda_I \{q(L_r, \delta) - 1\}}{12N(N-1)} \right. \\ &\quad \left. + \frac{\lambda_1 \{q(L_r, \delta) - 1\}}{4N} + \frac{\eta_0}{4E_b \Omega_0} \right) \quad \text{for } l = 0, \\ \sigma_{s,l}^{(k)^2}(\lambda_1, \lambda_I) &= E_b T \Omega_0 \left( \frac{(N-1)\lambda_I q(L_r, \delta)}{6N^2} \right. \\ &\quad \left. + \frac{\lambda_1 \{q(L_r, \delta) - 1\}}{4N} + \frac{\eta_0}{4E_b \Omega_0} \right) \quad \text{for } l \geq 1. \end{aligned} \quad (\text{A.4})$$

Using the Hermite polynomial approach, we can evaluate the average total interference-plus-noise power per AA element.

With these assumptions, the optimal beamformer weight of the  $k$ th user at the  $l$ th multipath can be shown to be  $\mathbf{W}_l^{(k)} = \sigma_{s,l}^{(k)^{-2}} \cdot \mathbf{V}(\theta_l^{(k)})$ . Therefore, between the array response vector of the  $m$ th user at the  $h$ th multipath and the weight vector of the  $k$ th user's  $l$ th path, the spatial correlation can be expressed as

$$C_{lh}^{(k,m)} = \frac{\mathbf{V}^H(\theta_l^{(k)}) \mathbf{V}(\theta_h^{(m)})}{\sigma_{s,l}^{(k)^2}} = \frac{CR_{lh}^{(k,m)}}{\sigma_{s,l}^{(k)^2}}, \quad (\text{A.5})$$

where

$$\begin{aligned} CR_{lh}^{(k,m)} &= \sum_{i=0}^{M-1} \exp(j\pi \text{si} \cos(\theta_l^{(k)})) \exp(-j\pi \text{si} \cos(\theta_h^{(m)})), \\ s &= \frac{2d}{\lambda}. \end{aligned} \quad (\text{A.6})$$

The second-order characterization of the spatial correlation is calculated as

$$\zeta_{lh}^{(k,m)^2} = E \left[ \left\{ C_{lh}^{(k,m)} \right\}^2 \right] = \frac{E \left[ \left\{ CR_{lh}^{(k,m)} \right\}^2 \right]}{\sigma_{s,l}^{(k)^4}}, \quad (\text{A.7})$$

where

$$\begin{aligned} \left\{ CR_{lh}^{(k,m)} \right\}^2 &= A(\theta_l^{(k)}, \theta_h^{(m)}) \\ &= \sum_{i=0}^{M-1} (i+1) \exp(j\pi \text{si} \cos \theta_l^{(k)}) \exp(-j\pi \text{si} \cos \theta_h^{(m)}) \\ &\quad + \sum_{i=M}^{2(M-1)} (2M-i-1) \exp(j\pi \text{si} \cos \theta_l^{(k)}) \\ &\quad \times \exp(-j\pi \text{si} \cos \theta_h^{(m)}). \end{aligned} \quad (\text{A.8})$$

The mean angles of arrival  $\theta_l^{(k)}$  and  $\theta_h^{(m)}$  have uniform distribution in  $[0, \pi)$  independently. So,

$$\begin{aligned} E \left[ \left\{ CR_{lh}^{(k,m)} \right\}^2 \right] &= \int_0^\pi \int_0^\pi A(\theta_l^{(k)}, \theta_h^{(m)}) d\theta_l^{(k)} d\theta_h^{(m)} \\ &= \begin{cases} \sum_{i=0}^{M-1} (i+1) J_0(\pi \text{si}) J_0(-\pi \text{si}) \\ + \sum_{i=M}^{2(M-1)} (2M-i-1) J_0(\pi \text{si}) J_0(-\pi \text{si}), & k \neq m \text{ or } l \neq h, \\ M^2, & k = m, l = h, \end{cases} \end{aligned} \quad (\text{A.9})$$

where  $J_0(x)$  is the zero-order Bessel function of the first kind.

## REFERENCES

- [1] L. C. Godara, "Application of antenna arrays to mobile communications. II. Beam-forming and direction-of-arrival considerations," *Proceedings of the IEEE*, vol. 85, no. 8, pp. 1195–1245, 1997.
- [2] A. F. Naguib, *Adaptive antennas for CDMA wireless networks*, Ph.D. dissertation, Electrical Engineering Department, Stanford University, Stanford, Calif, USA, 1996.
- [3] G. Raleigh, S. N. Diggavi, A. F. Naguib, and A. Paulraj, "Characterization of fast fading vector channels for multi-antenna communication systems," in *Proc. 28th Asilomar Conference on Signals, Systems and Computers*, vol. 2, pp. 853–857, Pacific Grove, Calif, USA, October–November 1994.
- [4] A. Stephenne and B. Champagne, "Effective multi-path vector channel simulator for antenna array systems," *IEEE Trans. Vehicular Technology*, vol. 49, no. 6, pp. 2370–2381, 2000.
- [5] J. S. Thompson, P. M. Grant, and B. Mulgrew, "Smart antenna arrays for CDMA systems," *IEEE Personal Communications*, vol. 3, no. 5, pp. 16–25, 1996.
- [6] A. J. Paulraj and C. B. Papadias, "Space-time processing for wireless communications," *IEEE Signal Processing Magazine*, vol. 14, no. 6, pp. 49–83, 1997.
- [7] E.-K. Hong, S.-H. Hwang, K.-J. Kim, and K.-C. Whang, "Synchronous transmission technique for the reverse link in DS-CDMA terrestrial mobile systems," *IEEE Trans. Communications*, vol. 47, no. 11, pp. 1632–1635, 1999.
- [8] W. Ye and A. M. Haimovich, "Performance of cellular CDMA with cell site antenna arrays, Rayleigh fading, and power control error," *IEEE Trans. Communications*, vol. 48, no. 7, pp. 1151–1159, 2000.
- [9] J. Litva and T.-K. Lo, *Digital Beamforming in Wireless Communications*, Artech House Publisher, Boston, Mass, USA, 1996.
- [10] T. Eng and L. B. Milstein, "Coherent DS-CDMA performance in Nakagami multipath fading," *IEEE Trans. Communications*, vol. 43, no. 234, pp. 1134–1143, 1995.
- [11] R. Prasad, *CDMA for Wireless Personal Communications*, Artech House Publisher, Boston, Mass, USA, 1996.
- [12] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, National Bureau of Standards Applied Mathematics Series, Dover Publications, New York, NY, USA, 1965.

**Yong-Seok Kim** was born August, 1970, in Seoul, Korea. He received the B.S. degree in electronic engineering from the Kyung Hee University, Yongin-shi, Korea, in 1998, and the M.S. degree in electrical and computer engineering from Yonsei University, Seoul, Korea, in 2000, and is working toward the Ph.D. degree in electrical and electronic engineering at the same university. His current research interests include multiple antenna system, multiuser communication, multicarrier system, and 4G communication techniques.



Engineering Education of Korea. Currently, he serves as a Member of Korea Communications Commission, a Project Manager of Qualcomm-Yonsei Research Lab, and a Director of Yonsei's IT Research Center. His research interests include spread-spectrum systems, multiuser communications, and 4G communications techniques.

**Seung-Hoon Hwang** received the B.S. degree in electrical engineering and the M.S. and Ph.D. degrees in communication systems from Yonsei University, Seoul, Korea in 1992, 1994, and 1999, respectively. His Ph.D. thesis is entitled "Performance evaluation of a synchronous DS-CDMA system in a mobile radio channel." From 1999 to 2003, he had worked for LG Electronics where he was a Chief Research Engineer in UMTS System Laboratory, LG R&D Center, participating in IMT-2000 physical layer standardization activities. From 2003, he is a Visiting Research Fellow at the School of Electronics and Computer Science in the University of Southampton, UK. His current research interests include interference cancellation techniques for DS-CDMA and various aspects of wideband/broadband CDMA. Dr. Hwang is a recipient of the British Chevening Scholarship awarded by the British Council, UK.



**Hyo-Yol Park** was born October, 1977, in Seoul, Korea. He received the B.S. degree in electronic engineering from the Yonsei University, Seoul, Korea, in 2000, and the M.S. degree in electrical and electronic engineering from Yonsei University, Seoul, Korea, in 2002, and is working toward the Ph.D. degree in electrical and electronic engineering at the same university. His current research interests include space-time coding, hybrid ARQ, turbo coding, multicarrier system, and 4G communication techniques.



**Keum-Chan Whang** was born on July 18, 1944, in Seoul, Korea. He received the B.S. degree in electrical engineering from Yonsei University, Seoul, Korea, in 1967, and the M.S. and Ph.D. degrees from the Polytechnic Institute of New York, in 1975 and 1979, respectively. From 1979 to 1980, he was a Member of Research Staff at the Agency for Defense Development, Korea. Since 1980, he has been Professor of the Department of Electrical and Electronic Engineering, Yonsei University. For the government, he performed various duties such as being a Member of Radio Wave Application Committee, a Member of Korea Information & Communication Standardization Committee, and is an Advisor for the Ministry of Information and Communication's technology fund and a Director of Accreditation Board for



# System-Level Performance of Antenna Arrays in CDMA-Based Cellular Mobile Radio Systems

**Andreas Czylik**

*Department of Communication Systems, University Duisburg-Essen, 47057 Duisburg, Germany  
Email: czylik@sent5.uni-duisburg.de*

**Armin Dekorsy**

*Lucent Technologies GmbH, Bell Labs Innovations, 90411 Nuremberg, Germany  
Email: dekorsy@lucent.com*

*Received 23 June 2003; Revised 1 March 2004*

Smart antennas exploit the inherent spatial diversity of the mobile radio channel, provide an antenna gain, and also enable spatial interference suppression leading to reduced intracell as well as intercell interference. Especially, for the downlink of future CDMA-based mobile communications systems, transmit beamforming is seen as a well-promising smart antenna technique. The main objective of this paper is to study the performance of diverse antenna array topologies when applied for transmit beamforming in the downlink of CDMA-based networks. In this paper, we focus on uniform linear array (ULA) and uniform circular array (UCA) topologies. For the ULA, we consider three-sector base stations with one linear array per sector. While recent research on downlink beamforming is often restricted to one single cell, this study takes into account the important impact of intercell interference on the performance by evaluating complete networks. Especially, from the operator perspective, system capacity and system coverage are very essential parameters of a cellular system so that there is a clear necessity of intensive system level investigations. Apart from delivering assessments on the performance of the diverse antenna array topologies, in the paper also different antenna array parameters, such as element spacing and beamwidth of the sector antennas, are optimized. Although we focus on the network level, fast channel fluctuations are taken into account by including them analytically into the signal-to-interference calculation.

**Keywords and phrases:** cellular system, system level simulation, beamforming, uniform linear array, uniform circular array, sectorized system.

## 1. INTRODUCTION

Mobile radio communication represents a rapidly growing market since the global system for mobile communications (GSM) standard has been established. Since then, third generation mobile radio systems like universal mobile telecommunication system (UMTS) or IMT-2000 have already been standardized [1, 2] and fourth generation systems are currently investigated. They will probably employ code division multiple access (CDMA) as a multiple access technique. In this paper, we focus on a CDMA-based system with frequency division duplex (FDD) like W-CDMA. A fundamental limitation on the capacity as well as coverage of CDMA-based mobile communication systems is the mutual interference among simultaneous users.

Smart antennas exploit the inherent spatial diversity of the mobile radio channel, provide an antenna gain, and also enable spatial interference suppression leading to reduced intracell as well as intercell interference. However, the implementation of this advanced technique in a handset is difficult

with today's hardware due to its limitations in size, cost, and energy storage capability while it is feasible to adopt antenna arrays at base stations.

In such a setting, transmit beamforming at base stations provides a powerful method for increasing downlink capacity [3, 4, 5, 6]. But, full exploitation of the spatial properties of the downlink channel requires meaningful transmit channel information at the base station. Third generation mobile systems are designed only with a low rate feedback information channel [5], hence, we focus in this paper on downlink beamforming strategies which are exclusively based on uplink information. While the instantaneous fading is normally uncorrelated between uplink and downlink, it is known that especially for UMTS, the long-term spatial and fading characteristics of the uplink channel can be used for transmit beamforming.

Recent research on downlink beamforming is either restricted on the direct link between a base station and mobile station or by considering only one single cell with few mobile stations. However, it is well known that especially for

the downlink, the impact of intercell interference on overall system performance plays an important role in CDMA-based systems [5, 7]. Thus, detailed investigations of downlink beamforming on the network level are strongly required. Note that especially from the operator perspective, system capacity and system coverage are very essential also enhancing the necessity of detailed system level investigations.

The main objective of this paper is to study the performance of diverse antenna array topologies when applied for transmit beamforming in the downlink of CDMA-based networks. In literature, some performance comparisons of systems with different array topologies can be found [8, 9, 10, 11, 12], but either no real cellular system is considered or important aspects like downlink transmission, maximum ratio combining at the receivers, or specific array topologies are not taken into account. In order to obtain a clear comparison and work out the performance improvement by transmit beamforming, we study omnidirectional as well as 3-sector networks whereby the latter concept represents the today's standard antenna configuration. Apart from delivering assessments on the performance of different antenna array topologies in a cellular network, the paper also evaluates and optimizes different antenna array parameters. Note that for the parameter optimization, again, we take into account network level aspects rather than only being focused on the arrays itself.

Our investigations are based on the evaluation of the signal-to-interference ratio (SIR) after RAKE reception at a mobile station. Although we are merely interested in system level results we include fast (instantaneous) fading properties in our investigations. Fast fading is analytically included in the calculation of the SIR values at the mobile stations. This analytical method is a new approach in the area of system level investigations. The key parameter of our investigations is the outage probability that is based on the calculation of the cumulative distribution function (CDF) of the SIR values. An outage occurs if the SIR of a mobile falls below a required SIR threshold.

Finally, it has to be mentioned that the results are based on a simulative approach. Thus, the propagation model plays an important role. Within this paper, we applied a quite realistic propagation model also taking into account the probabilistic nature of all parameters.

The paper is structured as follows. First, Section 2 introduces the basic signal model. Section 3 describes the main parameters for transmit beamforming and also gives a first insight on how to perform downlink beamforming by utilizing long-term uplink spatial mobile radio channel properties. Section 4 deals with the evaluation of the downlink path pattern which is composed of the beamformed pattern, the element-specific pattern, and the azimuthal power spectrum of the individual propagation paths. The latter results from the fact that each (macro)path consists of a large number of micropaths which cause an angular spread of each individual path. Within Section 4, we also calculate the SIR values. Next, in Section 5, the simulation model and simulation parameters are described. Section 6 shows extensive simulation results, and, finally, Section 7 concludes the paper.

## 2. SIGNAL MODEL

For the purpose of this paper, either a uniform linear array (ULA) or a uniform circular array (UCA) is considered for the base station, where the number of array elements for both array topologies is  $M$ . Mobile stations use one single antenna for transmission and reception only. For notational clarity, it is assumed that the multipath components of the frequency-selective mobile radio channel can be lumped into spatially or temporally resolvable (macro)paths. The number of resolvable paths is determined by the angular resolution of the antenna array and the angular power distribution of the propagation scenario as well as by the relation of the delay spread to the symbol duration of the signal of interest. It is assumed that the number of resolvable paths is the same for uplink and downlink. Here, the number of resolvable paths between the  $k$ th mobile station and the  $j$ th base station is denoted by  $L_{k,j}$ . The total number of users in the entire network is  $K$  and the number of base stations is  $J$ . Throughout the whole paper, uplink parameters and variables will be denoted by “ $\hat{\cdot}$ ” and correspondingly downlink parameters and variables by “ $\cdot$ ”.

In the following, we focus on uplink transmission at first. The mobile station  $k$  is assigned to the base station  $j(k)$ . At the receiver, the base stations see a sum of resolvable distorted versions of the transmitted signals  $\hat{s}_k(t)$  of users  $k = 0, \dots, K-1$ . The complex baseband representation of the antenna array output signal vector of base station  $j$  is given by

$$\hat{\mathbf{r}}_j(t) = \sum_{k=0}^{K-1} \sqrt{\hat{P}_k} \sum_{l=0}^{L_{k,j}-1} \hat{\mathbf{h}}_{l,k,j} \hat{s}_k(t - \hat{\tau}_{l,k,j}) + \hat{\mathbf{n}}_j(t), \quad (1)$$

where  $\hat{P}_k$  is the transmitted power from the  $k$ th user and  $\hat{\mathbf{h}}_{l,k,j}$  represents the channel vector of length  $M$  of path  $l$  between user  $k$  and base station  $j$ . It is assumed that the channel is quasi time-invariant within the period of interest. The  $k$ th user uplink signal  $\hat{s}_k(t)$  includes the complete baseband signal processing as channel encoding, data modulation, and spreading in case of CDMA transmission and  $\hat{\tau}_{l,k,j}$  is the time delay of the  $l$ th path between user  $k$  and base station  $j$ . Finally,  $\hat{\mathbf{n}}_j(t)$  is a spatially and temporally white Gaussian random process with covariance matrix

$$\hat{\mathbf{R}}_N = \mathbb{E}\{\hat{\mathbf{n}}_j \hat{\mathbf{n}}_j^H\} = \hat{\sigma}_N^2 \mathbf{I} \quad \text{for } j = 0, \dots, J-1, \quad (2)$$

where  $\mathbb{E}\{\cdot\}$  denotes the expectation.

The angular spread of the individual incoming resolvable paths determines the amount of spatial fading seen at an antenna array [4] and the size of the array employed will affect the coherence of the array output signals as well as which detection algorithms are applicable. For the rest of this paper, we assume closely spaced antenna elements yielding highly spatially correlated signals at the array elements. For this case, we can express the channel vector as

$$\hat{\mathbf{h}}_{l,k,j} = \hat{\alpha}_{l,k,j} \hat{\mathbf{a}}(\hat{\theta}_{l,k,j}), \quad (3)$$

where  $\hat{\alpha}_{l,k,j}$  is the channel coefficient which is composed of path loss, log-normal shadow fading as well as fast Rayleigh fading. The vector  $\hat{\mathbf{a}}(\hat{\theta}_{l,k,j})$  denotes the array response or steering vector to a planar wave impinging from an azimuth direction  $\hat{\theta}_{l,k,j}$ . In our model, we assume that the angles of arrival  $\hat{\theta}_{l,k,j}$  with  $l = 0, \dots, L_{k,j} - 1$  are Laplacian-distributed variables with mean  $\theta_{k,j}$ , the line-of-sight direction between user  $k$  and base station  $j$  [13, 14].

With the assumption of planar waves and uniformly located array elements, the frequency-dependent array response of a ULA is given by [13, 15, 16]

$$\mathbf{a}_l(\theta) = \left[ 1, e^{-j2\pi(d/\lambda)\sin(\theta)}, \dots, e^{-j2\pi(M-1)(d/\lambda)\sin(\theta)} \right]^T. \quad (4)$$

The interelement spacing of the antenna array is  $d$ , and  $\lambda$  represents the wavelength of the impinging wave. For the UCA, we have [15]

$$\begin{aligned} \mathbf{a}_c(\theta) \\ = \left[ 1, e^{-j2\pi(R/\lambda)\cos(\theta-2\pi/M)}, \dots, e^{-j2\pi(R/\lambda)\cos(\theta-2\pi(M-1)/M)} \right]^T, \end{aligned} \quad (5)$$

where  $R$  represents the radius of the array.

In order to form a beam for user  $k$  and detect its signal at base station  $j(k)$ , the received vector signal  $\hat{\mathbf{r}}_{j(k)}(t)$  is weighted by the weight vector  $\hat{\mathbf{w}}_k$ ,

$$\hat{y}_k(t) = \hat{\mathbf{w}}_k^H \hat{\mathbf{r}}_{j(k)}(t). \quad (6)$$

These weights depend on the optimization criterion, for example, maximizing the received signal energy (equivalent to SNR), maximizing the SINR, and minimizing the mean squared error between the received signal and some reference signal to be known at the base station [4].

Equation (6) can be rewritten with (1), (3) and either (4) or (5) to

$$\begin{aligned} \hat{y}_k(t) = & \sqrt{\hat{P}_k} \sum_{l=0}^{L_{k,j(k)}-1} \hat{\alpha}_{l,k,j(k)} \hat{\mathbf{w}}_k^H \hat{\mathbf{a}}(\hat{\theta}_{l,k,j(k)}) \hat{s}_k(t - \hat{\tau}_{l,k,j(k)}) \\ & + \sum_{\substack{\kappa=0 \\ \kappa \neq k}}^{K-1} \sqrt{\hat{P}_\kappa} \sum_{l=0}^{L_{\kappa,j(k)}-1} \hat{\alpha}_{l,\kappa,j(k)} \hat{\mathbf{w}}_k^H \hat{\mathbf{a}}(\hat{\theta}_{l,\kappa,j(k)}) \hat{s}_\kappa(t - \hat{\tau}_{l,\kappa,j(k)}) \\ & + \hat{\mathbf{w}}_k^H \hat{\mathbf{n}}_{j(k)}(t). \end{aligned} \quad (7)$$

The first term describes the desired signal, the second term represents the intercell as well as intracell interference, and the last expression describes additive Gaussian noise. Assuming that the data signals  $\hat{s}_k(t - \hat{\tau}_{l,k,j(k)})$  and the additive noise  $\hat{\mathbf{n}}_{j(k)}(t)$  are zero-mean and statistically independent random processes, the total received uplink signal power of the user

of interest at the base station can be expressed in the form

$$\begin{aligned} \hat{P}_{R,k} = & E \left\{ |\hat{y}_k(t)|^2 \right\} \\ = & \hat{P}_k \sum_{l=0}^{L_{k,j(k)}-1} |\hat{\alpha}_{l,k,j(k)}|^2 \cdot |\hat{\mathbf{w}}_k^H \hat{\mathbf{a}}(\hat{\theta}_{l,k,j(k)})|^2 \\ & + \sum_{\substack{\kappa=0 \\ \kappa \neq k}}^{K-1} \hat{P}_\kappa \sum_{l=0}^{L_{\kappa,j(k)}-1} |\hat{\alpha}_{l,\kappa,j(k)}|^2 \cdot |\hat{\mathbf{w}}_k^H \hat{\mathbf{a}}(\hat{\theta}_{l,\kappa,j(k)})|^2 \\ & + E \left\{ |\hat{\mathbf{w}}_k^H \hat{\mathbf{n}}_{j(k)}(t)|^2 \right\} \\ = & \hat{\mathbf{w}}_k^H \hat{\mathbf{R}}_{S,k} \hat{\mathbf{w}}_k + \hat{\mathbf{w}}_k^H \hat{\mathbf{R}}_{I,k} \hat{\mathbf{w}}_k + \hat{\mathbf{w}}_k^H \hat{\mathbf{R}}_N \hat{\mathbf{w}}_k, \end{aligned} \quad (8)$$

where the expectation operation is carried out with respect to the fast varying data signal and the additive noise. Note that the expectation is not carried out with respect to the fast fading processes, since we assume that the channel remains unchanged during a block of data. Here, it has been assumed that also time-delayed versions of the same data signal are uncorrelated. The  $k$ th user signal is normalized by  $E\{|s_k|^2\} = 1$  for  $k = 0, \dots, K-1$ . The essential elements in antenna array beamforming design are the spatial covariance matrices  $\hat{\mathbf{R}}_{S,k}$  for the desired signal as well as the spatial covariance matrices  $\hat{\mathbf{R}}_{I,k}$  for the interference of user  $k$ . Both matrices are instantaneous covariance matrices which are fluctuating according to fast fading. According to (8), these matrices are given by

$$\hat{\mathbf{R}}_{S,k} = \hat{P}_k \sum_{l=0}^{L_{k,j(k)}-1} |\hat{\alpha}_{l,k,j(k)}|^2 \cdot \hat{\mathbf{a}}(\hat{\theta}_{l,k,j(k)}) \hat{\mathbf{a}}(\hat{\theta}_{l,k,j(k)})^H, \quad (9)$$

$$\hat{\mathbf{R}}_{I,k} = \sum_{\substack{\kappa=0 \\ \kappa \neq k}}^{K-1} \hat{P}_\kappa \sum_{l=0}^{L_{\kappa,j(k)}-1} |\hat{\alpha}_{l,\kappa,j(k)}|^2 \cdot \hat{\mathbf{a}}(\hat{\theta}_{l,\kappa,j(k)}) \hat{\mathbf{a}}(\hat{\theta}_{l,\kappa,j(k)})^H. \quad (10)$$

These covariance matrices include all the spatial information necessary for beamforming. They can be measured in the uplink by correlating all antenna array output signals,

$$E \left\{ \hat{\mathbf{r}}_{j(k)} \hat{\mathbf{r}}_{j(k)}^H \right\} = \hat{\mathbf{R}}_{S,k} + \hat{\mathbf{R}}_{I,k} + \hat{\mathbf{R}}_N. \quad (11)$$

The only remaining task is to distinguish between the contribution of the desired signal and the contribution of interference plus noise. This can be accomplished by evaluating user-specific training sequences.

Next, downlink transmission is considered. A mobile terminal receives the desired signal from the base station to which it is connected. But it also receives interference from all other base stations. The received signal is given by

$$\begin{aligned} \check{y}_k(t) = & \sqrt{\check{P}_k} \sum_{l=0}^{L_{k,j(k)}-1} \check{\alpha}_{l,k,j(k)} \check{\mathbf{w}}_k^H \check{\mathbf{a}}(\check{\theta}_{l,k,j(k)}) \check{s}_k(t - \check{\tau}_{l,k,j(k)}) \\ & + \check{i}_k(t) + \check{n}_k(t). \end{aligned} \quad (12)$$

The first term in (12) is the desired signal and the second term  $\check{i}_k(t)$  is interference which is composed from intracell as well as intercell interference. The last term  $\check{n}_k(t)$  is additive

white Gaussian noise which is created from thermal and amplifier noise. Assuming that the data signals for different mobile stations are statistically independent and that also time-delayed versions of the same data signal are uncorrelated, the power of the received signal at mobile station  $k$  yields

$$\begin{aligned}\check{P}_{R,k} &= \mathbb{E}\left\{|\check{y}_k(t)|^2\right\} \\ &= \check{P}_k \sum_{l=0}^{L_{k,j(k)}-1} |\check{\alpha}_{l,k,j(k)}|^2 \cdot |\check{\mathbf{w}}_k^H \check{\mathbf{a}}(\check{\theta}_{l,k,j(k)})|^2 \\ &\quad + \mathbb{E}\left\{|\check{i}_k|^2\right\} + \mathbb{E}\left\{|\check{n}_k|^2\right\} \\ &= \check{\mathbf{w}}_k^H \check{\mathbf{R}}_{S,k} \check{\mathbf{w}}_k + \mathbb{E}\left\{|\check{i}_k|^2\right\} + \mathbb{E}\left\{|\check{n}_k|^2\right\}.\end{aligned}\quad (13)$$

Here,  $\check{\mathbf{R}}_{S,k}$  denotes the downlink covariance matrix for the desired signal component

$$\check{\mathbf{R}}_{S,k} = \check{P}_k \sum_{l=0}^{L_{k,j(k)}-1} |\check{\alpha}_{l,k,j(k)}|^2 \cdot \check{\mathbf{a}}(\check{\theta}_{l,k,j(k)}) \check{\mathbf{a}}(\check{\theta}_{l,k,j(k)})^H. \quad (14)$$

For an FDD system, fast fading processes in uplink and downlink are almost uncorrelated. Therefore, the instantaneous uplink covariance matrix cannot be used directly for downlink beamforming. But on the other hand, measurements have shown that the following spatial transmission characteristics for uplink and downlink are almost the same if the frequency spacing between uplink and downlink bands is not too large (see [17], [18, Section 3.2.2], [19]):

$$\hat{\theta}_{l,k,j} \cong \check{\theta}_{l,k,j}, \quad (15)$$

$$\hat{\tau}_{l,k,j} \cong \check{\tau}_{l,k,j}, \quad (16)$$

$$\mathbb{E}\left\{|\hat{\alpha}_{l,k,j}|^2\right\} \cong \mathbb{E}\left\{|\check{\alpha}_{l,k,j}|^2\right\}. \quad (17)$$

In (17), the expectation is taken over the fast fading processes. The equation implies that fading processes from shadowing are almost the same for uplink and downlink. Because of this reason, a part of the spatial information which is available from the uplink covariance matrices can be utilized also for the downlink.

Since the instantaneous full spatial information is not available for the downlink, downlink beamforming has to be based on averages (with respect to fast fading) of the covariance matrices.

### 3. DOWNLINK BEAMFORMING

The scope of this paper is to investigate different antenna array topologies for downlink beamforming. To fully exploit spatial filtering capabilities, complete downlink spatial information is required at the base station to reduce intercell as well as intracell interference. Complete spatial information comprises the knowledge of the covariance matrices which include the knowledge of instantaneous magnitudes of the channel coefficients  $|\alpha_{l,k,j(k)}|$ , the angles of arrival  $\theta_{l,k,j(k)}$ , and transmitted powers  $P_k$ . The beamforming strategy which will be discussed later in this section is directly based on covariance matrices.

Usually, spatial information is only available for uplink transmission by evaluating user-specific training sequences at base stations. For the downlink, a backward transmission of channel state information from the mobile stations to the base stations would be necessary. Since mobile communication systems are commonly designed with low data rate signalling feedback channels in order to obtain high bandwidth efficiency (e.g., UMTS [5]), neither the instantaneous channel coefficients nor steering vectors are known at the base station. Although the fast fading processes for uplink and downlink are uncorrelated, the averaged (with respect to fast fading) magnitudes of channel coefficients can be assumed to be insensitive to small changes in frequency. Thus, the averaged channel coefficients and angles of arrival can be estimated from the time-averaged uplink covariance matrices. For power control procedures which are controlled by base stations, all transmitted power levels are also known at the base stations.

The following methods can be used to estimate the downlink covariance matrices.

- (i) After estimation of angles of arrival and power transfer factors with high resolution estimation methods [20] from the time-averaged uplink covariance matrices, the downlink covariance matrices are calculated using (14).
- (ii) Alternatively, the covariance matrices are transformed directly from uplink to downlink carrier frequency by linear transformations as proposed in literature [21, 22, 23].
- (iii) Furthermore, it is possible to feedback the averaged downlink covariance matrix which may be measured at the mobile station. But this concept requires a high data rate feedback channel which allows to feedback the analog values of the elements of the covariance matrix. This concept can also be used for interference, but only within the considered cell—the contribution of intercell interference cannot be taken into account.

Of course, estimation errors cause some degradation compared with the ideal case where the covariance matrices are exactly known. For simplicity and in order to estimate the ultimate performance, in this paper we assume perfectly known time-averaged downlink covariance matrices.

The beamforming strategy in the present paper is to maximize the received signal power at mobile station  $k$ . The instantaneous received power at mobile station  $k$  is given by

$$\check{P}_{S,k} = \check{\mathbf{w}}_k^H \check{\mathbf{R}}_{S,k} \check{\mathbf{w}}_k, \quad (18)$$

where  $\check{\mathbf{R}}_{S,k}$  denotes the instantaneous downlink covariance matrix of the desired signal (14). As mentioned before, the instantaneous downlink covariance matrix is not known at the base station. Instead, we are using the time-averaged version which can be calculated with the above described methods. Therefore, the beamforming algorithm is based on the time-averaged downlink covariance matrix  $\check{\mathbf{R}}_{S,k}$  which

corresponds to the expectation

$$\begin{aligned}\tilde{\mathbf{R}}_{S,k} &= E\{\tilde{\mathbf{R}}_{S,k}\} \\ &= \check{P}_k \sum_{l=0}^{L_{k,j(k)}-1} E\left\{|\check{\alpha}_{l,\kappa,j(k)}|^2\right\} \cdot \check{\mathbf{a}}(\check{\theta}_{l,\kappa,j(k)})\check{\mathbf{a}}(\check{\theta}_{l,\kappa,j(k)})^H.\end{aligned}\quad (19)$$

Be aware that the steering vectors have to be determined at downlink frequency. Because we are averaging with respect to Rayleigh fading, the actual beamforming for the downlink is to maximize the average downlink power

$$\tilde{P}_{S,k} = \check{\mathbf{w}}_k^H \tilde{\mathbf{R}}_{S,k} \check{\mathbf{w}}_k, \quad (20)$$

while keeping the average total intracell and intercell interference power  $\tilde{P}_{I,k}$  transmitted from base station  $j(k)$  and received from all undesired mobile stations constant

$$\begin{aligned}\tilde{P}_{I,k} &= \sum_{\substack{\kappa=0 \\ \kappa \neq k}}^{K-1} E\left\{|\check{y}_\kappa(t)|^2\right\} \\ &= \check{P}_k \sum_{\substack{\kappa=0 \\ \kappa \neq k}}^{K-1} \sum_{l=0}^{L_{\kappa,j(k)}-1} E\left\{|\check{\alpha}_{l,\kappa,j(k)}|^2\right\} \cdot |\check{\mathbf{w}}_k^H \check{\mathbf{a}}(\check{\theta}_{l,\kappa,j(k)})|^2 \\ &= \check{\mathbf{w}}_k^H \tilde{\mathbf{R}}_{I,k} \check{\mathbf{w}}_k.\end{aligned}\quad (21)$$

Here,  $\tilde{\mathbf{R}}_{I,k}$  denotes the downlink interference covariance matrix (averaged with respect to the data signals and Rayleigh fading processes):

$$\tilde{\mathbf{R}}_{I,k} = \check{P}_k \sum_{\substack{\kappa=0 \\ \kappa \neq k}}^{K-1} \sum_{l=0}^{L_{\kappa,j(k)}-1} E\left\{|\check{\alpha}_{l,\kappa,j(k)}|^2\right\} \cdot \check{\mathbf{a}}(\check{\theta}_{l,\kappa,j(k)})\check{\mathbf{a}}(\check{\theta}_{l,\kappa,j(k)})^H. \quad (22)$$

Considering an interference-limited system and therefore neglecting the additive noise powers  $E\{|\check{n}_k(t)|^2\}$ , the described beamforming strategy corresponds to maximizing the (virtual) SIR per user, which is given by

$$\text{SIR}_k = \frac{\check{\mathbf{w}}_k^H \tilde{\mathbf{R}}_{S,k} \check{\mathbf{w}}_k}{\check{\mathbf{w}}_k^H \tilde{\mathbf{R}}_{I,k} \check{\mathbf{w}}_k}. \quad (23)$$

Note that the SIR of (23) cannot be measured at any terminal since the denominator contains the sum of interference powers measured at different mobile stations. Therefore, we call it virtual SIR.

The optimization problem to maximize the SIR can mathematically be expressed as

$$\check{\mathbf{w}}_k^{\text{opt}} = \arg \max_{\check{\mathbf{w}}_k} \frac{\check{\mathbf{w}}_k^H \tilde{\mathbf{R}}_{S,k} \check{\mathbf{w}}_k}{\check{\mathbf{w}}_k^H \tilde{\mathbf{R}}_{I,k} \check{\mathbf{w}}_k}, \quad (24)$$

where  $\check{\mathbf{w}}_k^{\text{opt}}$  represents the optimum solution. Since both covariance matrices are positive definite, the maximum SIR criterion is satisfied when the weight vector equals the principal eigenvector of the matrix pair associated with the largest eigenvalue [4, 13, 21], that is,

$$\tilde{\mathbf{R}}_{S,k} \check{\mathbf{w}}_k^{\text{opt}} = \lambda_{\max} \tilde{\mathbf{R}}_{I,k} \check{\mathbf{w}}_k^{\text{opt}}, \quad (25)$$

where  $\lambda_{\max}$  denotes the largest eigenvalue.

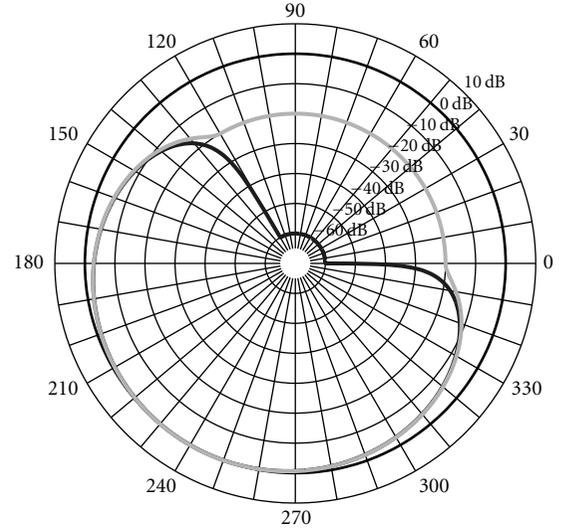


FIGURE 1: Antenna diagram of a single antenna element (main beam direction, 240°), backward attenuation  $a_R = 20$  dB and  $a_R = 60$  dB.

#### 4. DOWNLINK SIR

The total gain of the antenna array is given by [15],

$$\check{G}_k^{\text{tot}}(\theta) = |\check{\mathbf{w}}_k^{\text{opt}} \check{\mathbf{a}}(\theta)|^2 \cdot G^{\text{ele}}(\theta), \quad (26)$$

where the first term is due to the applied beamforming method and dependent on the topology used,  $\check{\mathbf{a}}(\theta)$  is given by (4) or (5), respectively. The second term takes into account the antenna element specific antenna pattern. Typical patterns of base station sector antennas show a smooth behavior within the main beam. Such a characteristic can be modelled quite well with a squared cosine characteristic. Within this paper, we apply antenna elements with squared cosine shapes in the form

$$G^{\text{ele}}(\theta) = \begin{cases} \cos^2\left(\frac{\pi}{2} \cdot \frac{\theta}{\theta_{3\text{dB}}}\right) & \text{for } |\theta| \leq \theta_0, \\ 10^{-a_R/10} & \text{for } |\theta| \geq \theta_0, \end{cases} \quad (27)$$

with  $\theta_0 = \theta_{3\text{dB}} \cdot 2/\pi \cdot \arccos 10^{-a_R/20}$ . In (27), the angle  $\theta_{3\text{dB}}$  is the 3 dB two-sided angular aperture of an antenna element (often termed half-power beamwidth) and  $a_R$  denotes the backward attenuation. By taking very large values for  $\theta_{3\text{dB}}$ , an omnidirectional antenna characteristic can be modelled. The specific shape of the antenna characteristic plays only a subordinate role as is shown later in this paper. Even if the 3 dB angular aperture is changed in a large range, no significant performance difference is found. If not otherwise declared, a 3 dB angular aperture of 120° is used. Figure 1 illustrates the antenna element-specific diagram. For ULAs, Figure 2 shows the orientation of 120° sectors in the cellular system and illustrates the sectorization of cells.

As introduced before, each resolvable path at the base station receiver is composed of micropaths (often modelled by many small scatterers) with slightly different angles of arrival at the antenna arrays. Thus, the power is spread around the

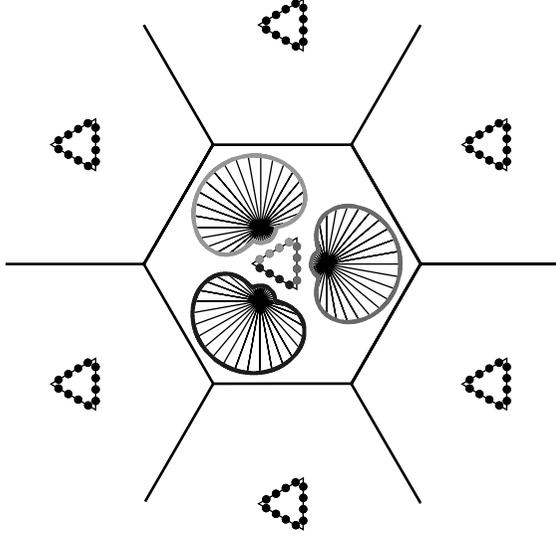


FIGURE 2: Single cell with antenna diagrams of the sector antennas.

average angle of arrival  $\check{\theta}_{l,k,j(k)}$  of each resolvable path and a (path-specific) azimuthal power spectrum has to be incorporated in the calculation of the signal and interference power for downlink transmission. To carry out the calculation we again fall back on the long-term reciprocity of the uplink and the downlink channel, refer to (15), (16), and (17). For the rest of this paper, we assume identical Laplacian-shaped azimuthal power spectra  $p_{l,k,j}(\theta) = p(\theta)$  for all paths in the system [13, 24]. With this assumption, the resulting gain factor seen by the  $l$ th departing path of user  $k$  at base station  $j(k)$  can be evaluated by convolving the total antenna gain diagram (26) with the azimuthal power spectrum,

$$G_k^{\text{path}}(\check{\theta}_{l,k,j}) = \int_{-\pi}^{\pi} \check{G}_k^{\text{tot}}(\theta) p(\theta - \check{\theta}_{l,k,j}) d\theta. \quad (28)$$

Within this paper,  $G_k^{\text{path}}$  is also referred to as path diagram [25].

In the following, we will give an expression for the SIR at a mobile station based on beamformed antenna diagrams at all base stations in the network. We consider CDMA systems with RAKE reception and assume the systems to be interference limited. Thus, the influence of thermal and amplifier noise can be neglected. With these assumptions and with reference on (13), the (instantaneous) postdespreading SIR per path of the user of interest (indexed with  $k$ ) is given by

$$\gamma_{l,k} = \frac{G_S \check{P}_{l,k}}{\check{P}_{l,k}^{\text{cross}} + \check{P}_k^{\text{intra}} + \check{P}_k^{\text{inter}}}, \quad l = 0, \dots, L_{k,j(k)} - 1, \quad (29)$$

with path power

$$\check{P}_{l,k} = \check{P}_k |\check{\alpha}_{l,k,j(k)}|^2 \check{G}_k^{\text{path}}(\check{\theta}_{l,k,j(k)}) \quad (30)$$

and path-crosstalk interference [26]

$$\check{P}_{l,k}^{\text{cross}} = \sum_{\substack{l'=0 \\ l' \neq l}}^{L_{k,j(k)}-1} \check{P}_k |\check{\alpha}_{l',k,j(k)}|^2 \check{G}_k^{\text{path}}(\check{\theta}_{l',k,j(k)}). \quad (31)$$

Here,  $\check{P}_k$  with  $k = 0, \dots, K-1$  denotes the transmitted power to be adjusted by power control [27, 28, 29]. In the present paper, we neglect the effect of power control and therefore assume  $\check{P}_k = \check{P}$  for  $k = 0, \dots, K-1$ . Since we focus on CDMA systems,  $G_S$  denotes the processing gain (despreading gain) [5, 26]. The variable  $\check{\alpha}_{l,k,j(k)}$  is given by (17) and includes signal fading. In implementable CDMA receivers, the number of paths to be evaluated is determined by the applied number of RAKE fingers [26]. Since we are interested in upper bound assessments for beamforming concepts, we neglect this restriction and assume all paths to be exploited by the RAKE receiver. Note that this leads to the highest degree of achievable path diversity in the time domain [26]. The intracell interference power yields

$$\check{P}_k^{\text{intra}} = \sum_{\kappa \in \mathcal{A}_k} \sum_{l=0}^{L_{k,j(k)}-1} \check{P}_\kappa |\check{\alpha}_{l,k,j(k)}|^2 \check{G}_\kappa^{\text{path}}(\check{\theta}_{l,k,j(k)}). \quad (32)$$

The set  $\mathcal{A}_k$  contains intracell interferers of user  $k$ . Note that the intracell interference signals pass through the same mobile channel as the signals of the user of interest, but they are weighted with their corresponding user-specific path diagram  $\check{G}_\kappa^{\text{path}}$ . Finally, the intercell interference power can be expressed as

$$\check{P}_k^{\text{inter}} = \sum_{\kappa \in \mathcal{B}_k} \sum_{l=0}^{L_{k,j(\kappa)}-1} \check{P}_\kappa |\check{\alpha}_{l,k,j(\kappa)}|^2 \check{G}_\kappa^{\text{path}}(\check{\theta}_{l,k,j(\kappa)}), \quad (33)$$

where  $\mathcal{B}_k$ ,  $k = 0, \dots, K-1$ , describes the set of users causing intercell interference seen by the  $k$ th user. The interference signals differ from the signals of interest by the mobile channels as well as path diagrams. Note that a large number of interfering signals arrives at each mobile. Thus, it is valid to approximate the path cross talk interference by including the path of interest, that is,  $\check{P}_{l,k}^{\text{cross}} \approx \sum_l \check{P}_k |\check{\alpha}_{l,k,j(k)}|^2 \check{G}_k^{\text{path}}(\check{\theta}_{l,k,j(k)})$ . This leads to identical interference powers (identical denominators in (29)) for all paths and simplifies the following analysis.

System level simulations often neglect short-term aspects as fast fading. Within this paper, we introduce a new approach which takes fast fading into account. First, it has to be mentioned that combining the resolvable paths is done by maximum ratio combining (MRC). Secondly, rather than explicitly modelling fast fading, we mathematically incorporate it in the evaluation of the SIR distribution when MRC is applied for different path power transfer factors [24, 26].

The key parameter of our investigations is the CDF of the SIR. It is assumed that all channel coefficients  $\check{\alpha}_{l,k,j}$  are complex Gaussian random variables which correspond to Rayleigh fading magnitudes. We furthermore presume that

the channel coefficients  $\check{\alpha}_{l,k,j}$  are statistically independent. The path gain factor  $\check{G}_k^{\text{path}}(\check{\theta}_{l,k,j(k)})$  in (30) depends on the optimum beam pattern (solution of (25)) which changes only very slowly with time since it is based on time-averaged covariance matrices. Because of the large number of terms in the denominator of (29), we can neglect the fluctuations of the denominator. Therefore, the only variables which fluctuate because of the Rayleigh fading are the channel coefficients  $\check{\alpha}_{l,k,j}$ . The Gaussian distribution of channel coefficients results in an exponentially distributed signal power per path (numerator of (29)). Since the interference power and all other terms of (29) (except the coefficients  $\check{\alpha}_{l,k,j}$ ) are assumed to be fixed or very slowly fluctuating, the signal-to-interference power ratios  $\gamma_{l,k}$  per path are distributed according to an exponential distribution [26], that is,

$$f_{\gamma_{l,k}}(\gamma_{l,k}) = \frac{1}{\overline{\gamma_{l,k}}} e^{-(\gamma_{l,k})/(\overline{\gamma_{l,k}})}, \quad (34)$$

where  $\overline{\gamma_{l,k}}$  denotes the average SIR of a single path (ensemble average with respect to fast fading). Assuming that the interference in each path is independent, the SIR after MRC results in

$$\gamma_k = \sum_{l=0}^{L_{k,j(k)}-1} \gamma_{l,k}. \quad (35)$$

Furthermore, it is assumed that the small scale fading of the individual desired paths is statistically independent. Since  $\gamma_k$  is the sum of the random variables  $\gamma_{l,k}$ , the resulting probability density function (PDF) is obtained from convolving the individual PDFs,

$$f_{\gamma_k}(\gamma_k) = f_{\gamma_{1,k}} * f_{\gamma_{2,k}} * f_{\gamma_{3,k}} * \dots * f_{\gamma_{L_{k,m(k)}-1,k}}. \quad (36)$$

Utilizing the characteristic functions of the PDFs, the resulting PDF of  $\gamma_k$  can be found to be [24, 26]

$$f_{\gamma_k}(\gamma_k) = \sum_{l=0}^{L_{k,j(k)}-1} \frac{c_{l,k}}{\overline{\gamma_{l,k}}} e^{-\gamma_k/\overline{\gamma_{l,k}}} \quad (37)$$

with the coefficients

$$c_{l,k} = \prod_{\substack{l'=0 \\ l' \neq l}}^{L_{k,j(k)}-1} \frac{\overline{\gamma_{l,k}}}{\overline{\gamma_{l,k}} - \overline{\gamma_{l',k}}}. \quad (38)$$

In order to compare the different beamforming concepts, the CDF has to be averaged over all mobiles and possibly over several simulations, where different locations for the mobiles and different radio channels are determined. Most information can be extracted from the averaged distribution function of the SIR,

$$F_{\gamma_k} = \int_0^{\gamma_k} E\{f_{\gamma_k}(u)\} du, \quad (39)$$

where the expectation is taken over all mobile stations and snapshots.

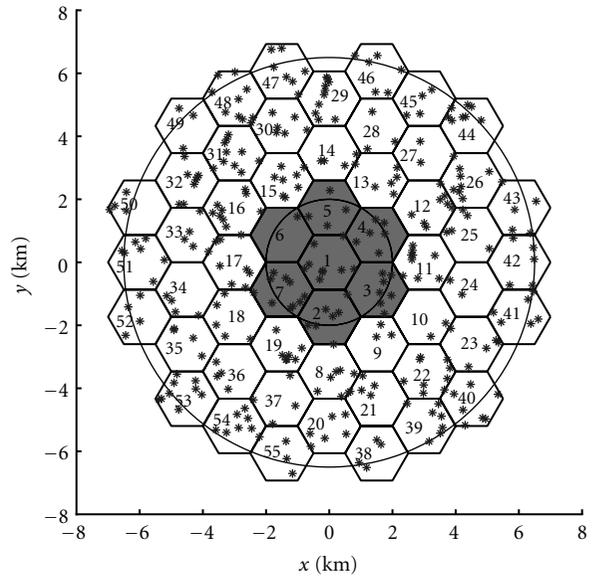


FIGURE 3: Cellular simulation model with reference cells (grey) in the center and randomly distributed mobile stations.

## 5. CELLULAR SIMULATION MODEL AND METHODOLOGY

### 5.1. Simulation model

The simulations are carried out with a regular hexagonal cellular model (see Figure 3). In order to be able to ignore fringe effects, the SIR is calculated only in a central area (reference cells). Mobile stations are randomly distributed in the cellular system according to a spatially uniform distribution. Note that a realistic model of the wave propagation plays an important role for the significance of the simulation results. One common approach, especially in context of downlink beamforming, is to use deterministic propagation scenarios [21, 30] or to apply propagation models which do not take into account the probabilistic nature of all parameters (e.g., the number of paths) [31, 32]. In the present paper, a completely probabilistic propagation model between each base station and each mobile is used which is characterized by the following properties.

The number of resolvable propagation paths is random and exhibits a binomial distribution (according to personal communication with U. Martin at Deutsche Telekom AG, 1999). Shadowing is modelled by a log-normal fading of the total received power [18, Section 3.1.1.2]. The random distribution of the total (log-normal fading) power to individual propagation paths (often denoted as macropaths or paths from scattering clusters) is modelled by applying an additional log-normal fading to the delayed paths with respect to the direct path (line of sight). Furthermore, a basic path attenuation and an extra attenuation that is proportional to the excess delay are taken into account. The basic attenuation is determined by the COST-Hata model [33] and a break point limits the attenuation to a certain minimum value for small distances. The excess delay of reflected paths is exponentially distributed leading to an exponential power delay profile [18, Section 3.1.1.3.3]. As mentioned

TABLE 1: Simulation parameters.

Average number of mobiles per cell	6
Maximum number of mobiles per cell	10
Cell radius	1 km
Carrier frequency	2 GHz
Antenna height of base stations	30 m
Antenna height of mobile stations	1.7 m
Break point that limits the attenuation at small distances	100 m
Standard deviation of slow fading	8 dB
Average number of paths	3
Maximum number of paths	6
Standard deviation of the attenuation of the delayed paths with respect to the direct path	6 dB
Average attenuation of the delayed paths with respect to the direct path	8 dB
Additional attenuation proportional to the excess delay	4 dB/ $\mu$ s
Standard deviation of the DoAs with respect to the direct path	20°
Standard deviation of the angular spread of each individual path	1°

TABLE 2: Antenna arrays.

<i>Circular antenna array</i>	
Number of antenna elements	12
Radius of the array	0.12 m
<i>Uniform linear array</i>	
Number of elements per sector	4
Number of sectors	3
Element spacing	$\lambda/2 = 0.075$ m

before, the directions of arrival which are denoted by  $\hat{\theta}_{l,k,j(k)}$  obey a Laplacian distribution with respect to the direct path (standard deviation = several tens of degrees) [18, Section 3.2.2.1]. Moreover, according to (28), the azimuthal power spectrum of each individual path is also incorporated in the simulations. As mentioned before, the azimuthal power spectra follow also a Laplacian shape (standard deviation in the order of one degree or less) and are identical for the different paths. In order to reduce the computational complexity, fast fading processes are included analytically as described in Section 4.

In the simulations, power control issues are completely neglected for downlink as well as uplink. The downlink transmit power values are assumed to be the same for all mobile stations, that is,  $\check{P}_k = \check{P}$  for  $k = 0, \dots, K - 1$ . It has to be mentioned that the capacity of the system increases when adopting power control since intracell interference is reduced. However, intercell interference is only marginally affected by power control. Finally, no handover issues are considered within this paper.

## 5.2. Simulation methodology and parameters

One main objective of this paper is to compare the performance gain for different smart antenna topologies. The key parameter to express performance is the outage probability for the given antenna concept. An outage occurs if the SIR of

the mobile station after RAKE reception with maximum ratio combining falls below the service dependent required SIR threshold. Thus, the outage probability is given by the CDF of the SIR calculated versus all mobile stations in the reference cells. Since the SIR depends on the spreading gain and the spreading gain is determined by the specific service, we do not take into account the spreading gain. For all following numerical results, we set  $G_S = 1$ . Note that the simulations are based on snapshots with fixed mobiles, where for each snapshot a CDF can be calculated. For each snapshot, we dice the locations of the mobiles as well as all other random variables. The following list gives a short overview of the main simulation steps.

- (1) Based on the uplink transmission and using the reciprocity of uplink and downlink, we calculate the spatial covariances for downlink as well as the optimum beamforming weights.
- (2) In a second step, the path diagrams are evaluated taking into account the beamformed diagram, the element specific diagrams, as well as the azimuthal power distribution of each resolvable path.
- (3) With this, the user-specific SIRs after RAKE reception are known and can be used for CDF calculation.
- (4) Finally, in order to compare the different array topologies, we average the CDF over all mobiles and over several snapshots, where different locations for the mobiles and different radio channels are determined. The averaged CDF allows to directly read the instantaneous outage probability of the downlink transmission.

The main simulation parameters are summarized in Tables 1 and 2. It has to be mentioned that for the system investigations we simulate  $6 \cdot 7$  mobile stations within reference cells in average and 100 snapshots are carried out. Thus, the resulting CDF is calculated by averaging over  $6 \cdot 7 \cdot 100 = 4200$  mobile stations.

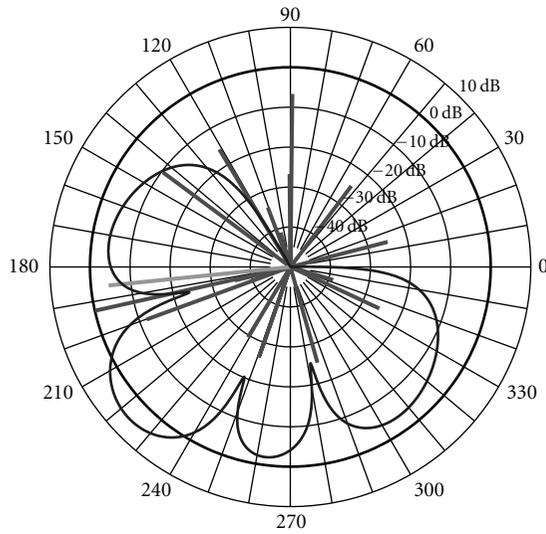


FIGURE 4: Example for an optimized path diagram in a sectorized system for a single sector (main beam direction is  $240^\circ$ ). The ULA consists of 4 elements.

For illustration purposes, Figures 4 and 5 show examples of path diagrams for an identical propagation scenario. A system with three sectors and a ULA with 4 elements per sector (12 antenna elements in total) is compared with a system with circular arrays each of 12 elements. The bars in the diagrams correspond to the gain factors of the individual paths—for the displayed example only one desired path (at beam direction of  $186^\circ$ ) exists.

Figure 4 shows the path diagram for the sectorized system. The backward attenuation of the antenna elements is  $a_R = 60$  dB. It can be observed in the figure that the beam-forming algorithm tries to suppress the undesired paths. Obviously, the four element antenna array does not exhibit sufficient degrees of freedom to generate all required nulls.

For the same propagation scenario, Figure 5 shows the optimization result for the circular array with 12 elements. Due to the larger number of antenna elements, the circular array is much more able to suppress the strong undesired paths.

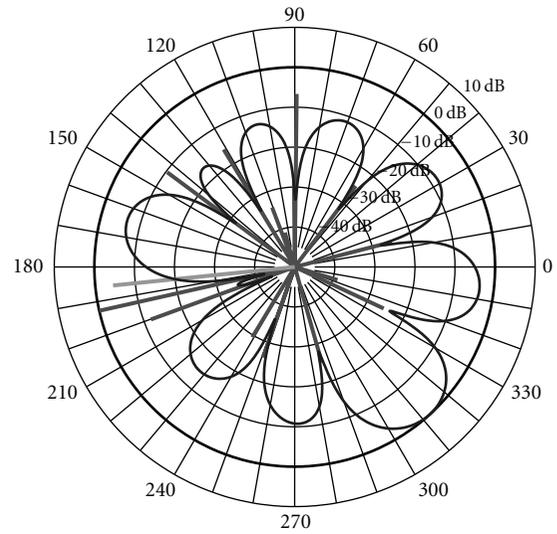


FIGURE 5: Example for an optimized path diagram for a circular antenna array with 12 omnidirectional elements.

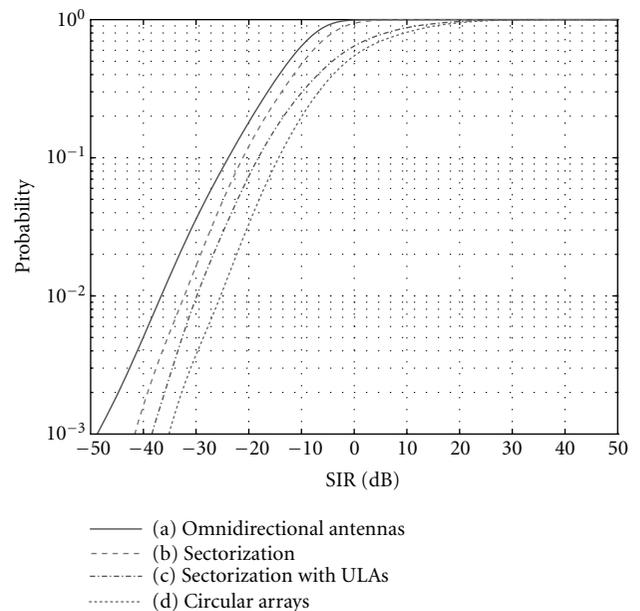


FIGURE 6: Averaged CDF of the instantaneous SIR. Comparison between (a) reference system with omnidirectional antenna elements, (b) sectorized system with a single sector antenna per sector, (c) sectorized system with ULAs in each sector, four antenna elements per sector, and (d) system with circular antenna arrays and 12 omnidirectional antenna elements.

## 6. SIMULATION RESULTS

### Overall performance comparison

Figure 6 shows the different CDFs for the diverse antenna array topologies that are under investigation. The topologies we are interested in are as follows:

- one omnidirectional antenna per base station,
- three-sector base stations with one antenna element per sector and squared cosine characteristic,
- three-sector base stations where we apply one ULA with four elements per sector and squared cosine characteristic,
- one UCA with 12 omnidirectional antenna elements per base station.

The omnidirectional topology is used as reference, while (b) is practically implemented today, and topologies (c) and (d) are under discussion for future implementation.

Figure 6 shows that for an outage probability of  $10^{-2}$ , simple sectorization yields a gain of about 4 dB compared to the omnidirectional configuration. The application of the linear array leads to an additional gain of about 3 dB. The circular array is superior and indicates an extra gain of

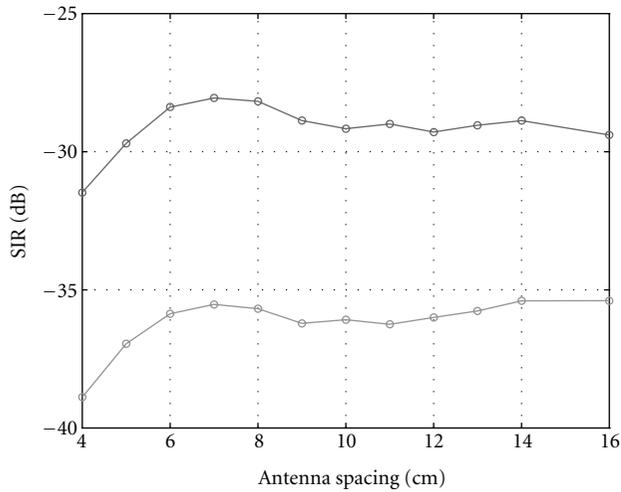


FIGURE 7: SIR for an outage probability of  $10^{-2}$  versus ULA element spacing for sectorized system. Dark curve: 6 mobile stations per cell, light curve: 20 mobile stations per cell.

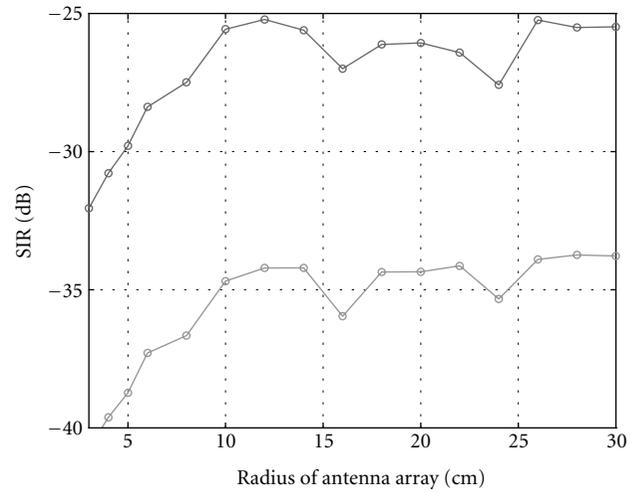


FIGURE 8: SIR for an outage probability of  $10^{-2}$  versus circular array radius. Dark curve: 6 mobile stations per cell, light curve: 20 mobile stations per cell.

approximately 4 dB compared to the linear array topology. The latter gain can be explained as follows.

- (i) The circular array is able to form narrower beams due to the larger number of antenna elements (4 per ULA compared to 12 per UCA). This means that nulls and maxima in the path diagram can be arranged more densely.
- (ii) Due to the larger number of antenna elements, the circular array exhibits more nulls in the diagram. These nulls can be arranged more flexibly in order to perform nulling of the undesired and amplification of the desired paths. For example, if many strong undesired paths are located in a certain angular range, the circular array is more capable to suppress them while the ULA suffers due to its less powerful nulling capability in that range.
- (iii) It is well known [15] that a ULA exhibits a low angular resolution for large angles (with respect to the main beam direction) while for the UCA this is not the case.

It has to be mentioned that the ULA performance is improved by handover between sectors of one base station (softer handover) [5]. But this technology is out of scope for this paper and might be an interesting task for future investigations.

#### Spacing of antenna elements, backward attenuation, and half-power beamwidth

An important parameter of an antenna array is the spacing of its elements. In the following, we discuss the impact of the antenna element spacing on the SIR. For the 3-sector system with ULAs, Figure 7 shows the SIR which is achieved for an average outage probability of  $10^{-2}$  versus the antenna element spacing. We consider system loads of an average number of 6 and 20 mobile stations per cell, respectively. The higher the SIR for a given load the better the performance of

the antenna array, since the array is more capable to suppress the interference. We observe that the antenna spacing should be at least  $\lambda/2 \approx 7.5$  cm independent of the given system load. For larger element spacing, the performance changes only slightly, while for small spacing it extremely degrades. The degradation can be explained by a reduced number of nulls in the path diagram for small antenna distances. A system with circular arrays is analyzed in Figure 8. The radius of the circular array should be at least 12 cm. This value corresponds to an antenna spacing of approximately 6.4 cm which is slightly less than  $\lambda/2$ . Note that for all considered angular spread and spacings between the antenna elements, high correlation between antenna elements is still assumed. Figures 7 and 8 show curves for an average density of 6 and 20 mobiles per cell. It can be observed that the shape of the curves does not depend significantly on the average number of mobiles per cell.

From a practical perspective, antenna arrays with smaller dimensions are easier to adopt. Because of this aspect and because of the results of Figures 7 and 8, it can be concluded that half of the wavelength is the best suitable antenna spacing.

Next, Figure 9 shows the performance of a sectorized system (single antenna and ULA) for different backward attenuations of the antenna elements. No performance difference can be noticed between antenna elements with backward attenuations of 20 and 60 dB. This result indicates that in sectorized systems, the requirements for the backward attenuation are less severe.

Up to here, we assumed a half power beamwidth (3 dB angular aperture) of  $120^\circ$  for sectorized systems. In the following, we study the impact of this design parameter on the system performance. Remember that we consider neither additive noise nor broadcast channels. Thus, the same maximum gain can be used for all antennas independently from the angular aperture. Corresponding to Figures 7 and 8, in

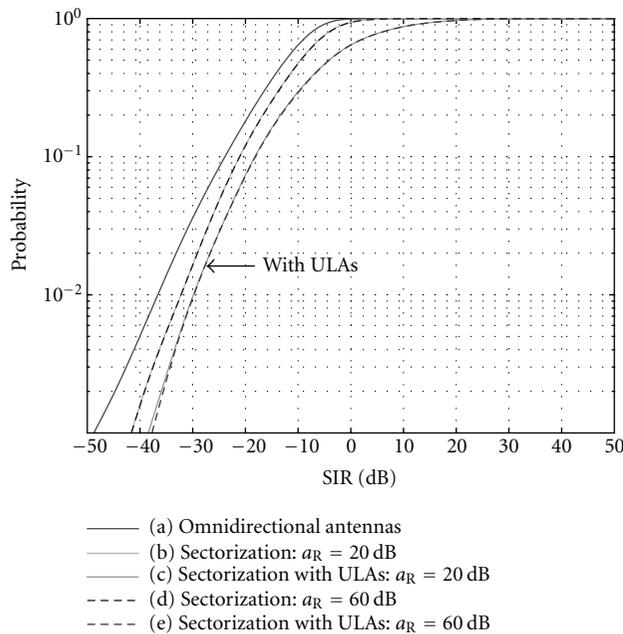


FIGURE 9: Averaged CDF of the instantaneous SIR. Comparison between (a) reference system with omnidirectional antenna elements, (b) sectorized system with a single sector antenna per sector ( $a_R = 20$  dB), (c) sectorized system with ULAs in each sector, four sector antennas per sector ( $a_R = 20$  dB), (d) sectorized system with a single sector antenna per sector ( $a_R = 60$  dB), (e) sectorized system with ULAs in each sector, four sector antennas per sector ( $a_R = 60$  dB).

Figure 10 the SIR for an outage probability of  $10^{-2}$  is shown versus the half-power beamwidth. It can be seen that an angular aperture of  $120^\circ$  is not optimum. The optimum value is of about  $150^\circ$ . But, the optimum reveals to be very wide leading to almost no performance degradation if the angular aperture is in the range  $120^\circ$ – $220^\circ$ .

#### Circular array with sector elements

In our final investigations, we analyze the system performance of a circular array when sector antenna elements are applied instead of elements with omnidirectional antenna patterns. The beam of each antenna element is pointing in radial direction (see Figure 11). Such an antenna array models an array that surrounds an inner mast where the shadowing of the antenna mast cannot be neglected. For simplicity, antenna diagrams described by (27) are applied. Figure 12 depicts the SIR for a given outage probability of  $10^{-2}$  versus the 3 dB beamwidth of the sector antennas. For a 12 element circular array it can be observed that already for small beamwidths of about  $40^\circ$  the optimum performance of omnidirectional antennas is achieved.

The importance of adopting sector antennas in circular antenna arrays has to be emphasized since because of the mutual coupling between antenna elements and even without a mast in the center, it is difficult to develop circular antenna arrays with omnidirectional antenna patterns. An additional

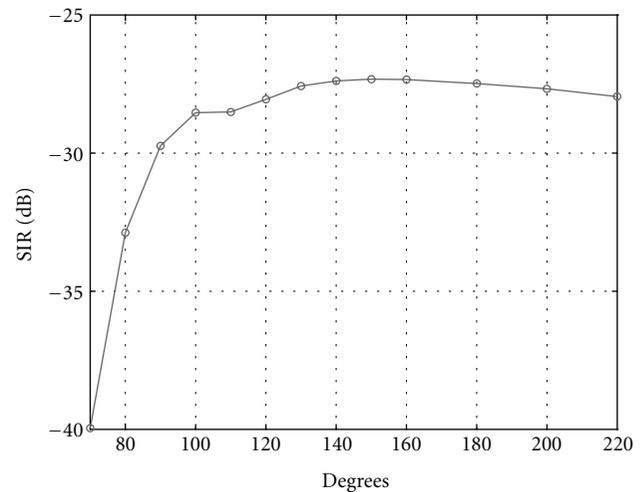


FIGURE 10: SIR for an outage probability of  $10^{-2}$  versus 3 dB beamwidth of sector antennas.

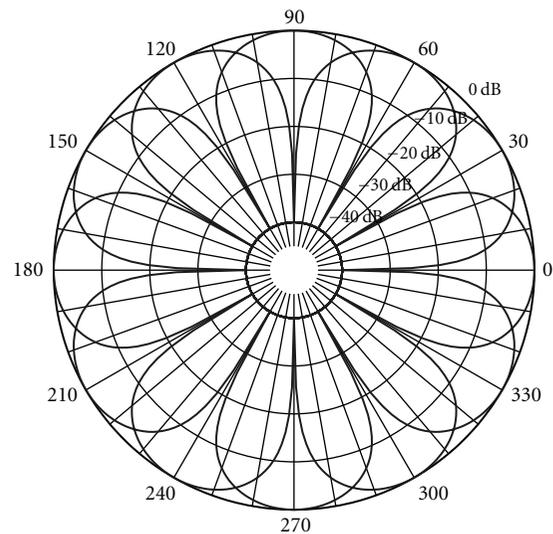


FIGURE 11: Antenna diagrams of all elements of a 12 element antenna array with a beamwidth of  $30^\circ$ . Backward attenuation  $a_R = 40$  dB.

advantage of using sector antennas is that their mutual coupling is weaker than between omnidirectional antenna elements.

## 7. CONCLUSION

A cellular model for system level investigations of antenna arrays has been presented. A new simulation methodology has been applied, which takes into account the gain of path diversity in a realistic manner. With the described assumptions and approximations it was possible to determine upper limits for the SIR gain when smart antennas are applied in CDMA-based mobile radio networks.

The CDF (outage probability) of the SIR after RAKE reception with maximum ratio combing is compared for

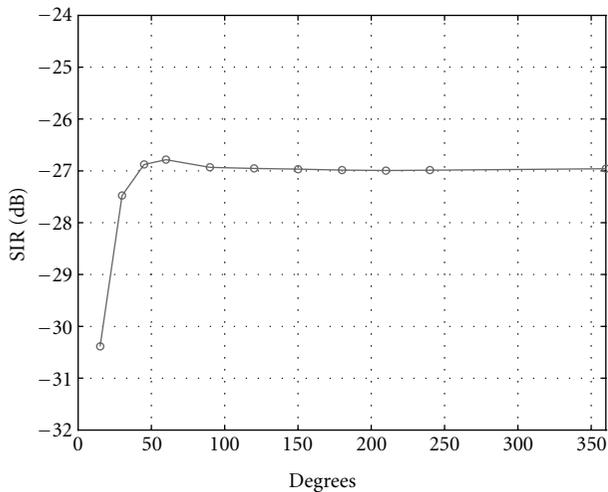


FIGURE 12: SIR for an outage probability of  $10^{-2}$  versus 3 dB beamwidth of sector antennas of a system with circular antenna arrays. For comparison, the beamwidth of  $360^\circ$  corresponds to omnidirectional antenna elements.

networks with and without sectorization, as well as with and without smart antenna arrays. For a fair comparison of diverse smart antenna array topologies, we considered networks with the same number of antenna elements at each base station. The lowest outage probability was found for networks applying circular antenna arrays. The gain with respect to the 3-sector system with one ULA per sector is of about 4 dB.

Furthermore, the parameters of the antenna arrays have been optimized by extensive simulations. The observed results indicate that the element spacing should be approximately half of the wavelength—independently from the antenna topology. Only slight performance changes have been observed for larger element spacings, while for small element distances the performance degrades.

Concerning the backward attenuation of the element-specific antenna diagrams, the results show that the backward attenuation can be as low as 20 dB without any performance degradation. Furthermore, the 3 dB beamwidth is also an uncritical parameter—it may be within the range  $120^\circ$ – $220^\circ$ .

Finally, no performance degradation has been observed for circular arrays if sector antennas with reasonably large beamwidths are used instead of omnidirectional antenna elements.

## REFERENCES

- [1] E. Dahlman, B. Gudmundson, M. Nilsson, and A. Sköld, “UMTS/IMT-2000 based on wideband CDMA,” *IEEE Communications Magazine*, vol. 36, no. 9, pp. 70–80, 1998.
- [2] T. Ojanperä and R. Prasad, “An overview of air interface multiple access for IMT-2000/UMTS,” *IEEE Communications Magazine*, vol. 36, no. 6, pp. 82–95, 1998.
- [3] H. Boche and M. Schubert, “Theoretical and experimental comparison of optimization criteria for downlink beamform-

- ing,” *European Transactions on Telecommunications*, vol. 12, no. 5, pp. 417–426, 2001, Special Issue on Smart Antennas.
- [4] R. M. Buehrer, A. G. Kogiantis, S.-C. Liu, J. Tsai, and D. Uptegrove, “Intelligent antennas for wireless communications - uplink,” *Bell Labs Technical Journal*, vol. 4, no. 3, pp. 73–103, 1999.
- [5] H. Holma and A. Toskala, *WCDMA for UMTS*, John Wiley & Sons, Chichester, UK, 1st edition, 2000.
- [6] A. Yener, R. D. Yates, and S. Ulukus, “Interference management for CDMA systems through power control, multiuser detection, and beamforming,” *IEEE Trans. Communications*, vol. 49, no. 7, pp. 1227–1239, 2001.
- [7] J. Muckenheimer and U. Bernhard, “A framework for load control in 3rd generation CDMA networks,” in *Proc. IEEE Global Telecommunications Conference*, vol. 6, pp. 3738–3742, San Antonio, Tex, USA, November 2001.
- [8] J.-W. Liang and A. J. Paulraj, “On optimizing base station antenna array topology for coverage extension in cellular radio networks,” in *Proc. IEEE 45th Vehicular Technology Conference*, vol. 2, pp. 866–870, Chicago, Ill, USA, July 1995.
- [9] J. Fuhl, D. J. Cichon, and E. Bonek, “Optimum antenna topologies and adaptation strategies for SDMA,” in *Proc. IEEE Global Telecommunications Conference*, vol. 1, pp. 575–580, London, UK, November 1996.
- [10] R. Martinez, D. Trosa, L. de Haro, and M. Calvo, “Smart antennas performance evaluation and capacity increase for WCDMA UMTS,” in *Proc. IEEE 53rd Vehicular Technology Conference*, vol. 1, pp. 147–151, Rhodes, Greece, May 2001.
- [11] A. Osseiran, M. Ericson, J. Barta, B. Goransson, and B. Hagerman, “Downlink capacity comparison between different smart antenna concepts in a mixed service W-CDMA system,” in *Proc. IEEE 54th Vehicular Technology Conference*, vol. 3, pp. 1528–1532, Atlantic City, NJ, USA, October 2001.
- [12] M. Pettersen, L. E. Braten, and A. G. Spilling, “An evaluation of adaptive antennas for UMTS FDD by system simulations,” in *Proc. IEEE Vehicular Technology Conference*, vol. 1, pp. 227–231, Orlando, Fla, USA, October 2003.
- [13] A. Czylik and A. Dekorsy, “System level simulations for downlink beamforming with different array topologies,” in *Proc. IEEE Global Telecommunications Conference*, vol. 5, pp. 3222–3226, San Antonio, Tex, USA, November 2001.
- [14] A. Czylik and A. Dekorsy, “Optimization of downlink beamforming for systems with frequency division duplex,” in *Proc. IEEE International Zurich Seminar on Broadband Communications*, pp. 11-1–11-6, Zurich, Switzerland, February 2002.
- [15] J. Litva and T. K.-Y. Lo, *Digital Beamforming in Wireless Communications*, Artech House, Boston, Mass, USA, 1st edition, 1996.
- [16] M. Schacht, A. Dekorsy, and P. Jung, “Downlink beamforming concepts in UTRA FDD,” in *Kleinheubacher Tagung*, Kleinheubach, Germany, September 2002.
- [17] K. I. Pedersen, P. E. Mogensen, and F. Frederiksen, “Joint-directional properties of uplink and downlink channel in mobile communications,” *Electronics Letters*, vol. 35, no. 16, pp. 1311–1312, 1999.
- [18] L. M. Correia, *Wireless Flexible Personalised Communications: COST 259: European Co-operation in Mobile Radio Research*, John Wiley & Sons, Chichester, UK, 2001.
- [19] K. Hugl, K. Kalliola, and J. Laurila, “Spatial reciprocity of uplink and downlink radio channels in FDD systems,” in *Proc. COST 273 Technical Document TD(02)066*, p. 7, Espoo, Finland, May 2002.
- [20] M. Haardt, *Efficient one-, two-, and multidimensional high-resolution array signal processing*, Ph.D. thesis, Lehrstuhl für Netzwerktheorie und Schaltungstechnik, Technische Universität, Munich, Germany, 1997.

- [21] W. Utschick and J. A. Nossek, "Downlink beamforming for FDD mobile radio systems based on spatial covariances," in *Proc. European Wireless 99 & ITG Mobile Communications*, pp. 65–67, Munich, Germany, October 1999.
- [22] D. Filho, C. Panazio, F. Cavalcanti, and J. Romano, "On downlink beamforming techniques for TDMA/FDD systems," in *Proc. 19th Symposium on Brazilian Telecommunications*, Fortaleza-CE, Brazil, 2001.
- [23] B. Chalise, L. Häring, and A. Czylik, "Robust UL to DL spatial covariance matrix transformation for DL beamforming," in *Proc. IEEE Global Telecommunications Conference (GLOBECOM '03)*, San Francisco, Calif, USA, December 2003.
- [24] A. Czylik, "Downlink beamforming for mobile radio systems with frequency division duplex," in *Proc. IEEE 11th International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 1, pp. 72–76, London, UK, September 2000.
- [25] A. Czylik, "Comparison and optimization of antenna concepts for downlink beamforming," in *Proc. International Conference on Telecommunications*, Papeete, French Polynesia, France, 2003.
- [26] J. G. Proakis, *Digital Communications*, McGraw-Hill, New York, NY, USA, 3rd edition, 1995.
- [27] F. Rashid-Farrokhi, K. J. R. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1437–1450, 1998.
- [28] M. Schubert and H. Boche, "A unifying theory for uplink and downlink multi-user beamforming," in *Proc. International Zurich Seminar on Broadband Communications*, pp. 271–276, Zurich, Switzerland, February 2002.
- [29] E. Visotsky and U. Madhow, "Optimum beamforming using transmit antenna arrays," in *Proc. IEEE 49th Vehicular Technology Conference*, vol. 1, pp. 851–856, Houston, Tex, USA, July 1999.
- [30] C. Farsakh and J. A. Nossek, "Spatial covariance based downlink beamforming in an SDMA mobile radio system," *IEEE Trans. Communications*, vol. 46, no. 11, pp. 1497–1506, 1998.
- [31] G. G. Raleigh, S. N. Diggavi, V. K. Jones, and A. Paulraj, "A blind adaptive transmit antenna algorithm for wireless communication," in *Proc. IEEE International Conference on Communications*, vol. 3, pp. 1494–1499, Seattle, Wash, USA, June 1995.
- [32] H. Asakura and T. Matsumoto, "Cooperative signal reception and down-link beam forming in cellular mobile communications," *IEEE Trans. Vehicular Technology*, vol. 48, no. 2, pp. 333–341, 1999.
- [33] J. S. Lee and L. E. Miller, *CDMA Systems Engineering Handbook*, Artech House, Boston, Mass, USA, 1st edition, 1998.

became a full professor at the Technical University of Braunschweig, heading the Department of Microcellular Radio Systems. Since 2002, he has been with the University Duisburg-Essen and in charge of the Department of Communication Systems. He was an Editor for the IEEE Journal on Selected Areas in Communications and IEEE Transactions on Wireless Communications. His research interests are in the field of adaptive transmission techniques in radio communications, such as smart antennas and adaptive modulation and coding techniques.

**Armin Dekorsy** received his Dipl.-Ing. (FH) (B.S.) degree from Fachhochschule Konstanz, Germany, 1992, his Dipl.-Ing. (M.S.) degree from University of Paderborn, Germany, 1995, and his Ph.D. degree from University of Bremen, Bremen, Germany, 2000, all in electrical engineering. From 2000 to 2001 he was with T-Nova Deutsche Telekom Innovationsgesellschaft mbH, Darmstadt, Germany, where he was leading research projects on smart antenna technologies. In 2001, he joined Lucent Technologies Network Systems GmbH, Nuremberg, Germany. Since October 2003 he has been with Bell Labs Advanced Technologies and is currently conducting research projects on radio resource management algorithms including interference cancellation techniques. He also contributes to marketing strategies, manages government funded research projects, and presents the Bell Labs Advanced Technologies at numerous seminars. His current research interests are mainly smart antenna solutions, interference cancellation techniques, as well as radio resource management algorithms for third-generation mobile standards.



**Andreas Czylik** studied electrical engineering at the Technical University of Darmstadt, Germany, from 1978 to 1983. In 1988, he received his Dr.-Ing. degree and in 1994 his Habilitation degree, both from the Technical University of Darmstadt and both in the field of optical communications. From 1994 to 2000, he was with the research and development center (Technologiezentrum) of Deutsche Telekom in the Department of Local Area Broadband Radio Systems. He was in charge of several research projects, for example, a broadband radio communication demonstrator based on single carrier transmission with frequency domain equalization, as well as several projects on smart antenna concepts in cellular mobile radio systems. In 2000, he



# Gaussian Channel Model for Mobile Multipath Environment

**D. D. N. Bevan**

*Harlow Laboratories, Nortel Networks, Harlow, Essex CM17 9NA, UK  
Email: ddnb@nortelnetworks.com*

**V. T. Ermolayev**

*Communication Systems Research Department, MERA Networks, Nizhny Novgorod 603126, Russia  
Email: ermol@mera.ru*

**A. G. Flaksman**

*Communication Systems Research Department, MERA Networks, Nizhny Novgorod 603126, Russia  
Email: flak@mera.ru*

**I. M. Averin**

*Communication Systems Research Department, MERA Networks, Nizhny Novgorod 603126, Russia  
Email: ave@mera.ru*

*Received 28 May 2003; Revised 5 February 2004*

A model of an angle-spread source is described, termed the “Gaussian channel model” (GCM). This model is used to represent signals transmitted between a user equipment and a cellular base station. It assumes a Gaussian law of the scatterer occurrence probability, depending upon the scatterer distance from the user. The probability density function of the angle of arrival (AoA) of the multipath components is derived for an arbitrary angle spread. The “wandering” of the “centre of gravity” of the scattering source realisation is investigated, which is in turn due to the nonergodicity of the angle-scatter process. Numerical results obtained with the help of the sum-difference bearing method show the dependence of the AoA estimation accuracy on the spread-source model.

**Keywords and phrases:** scattering, angle spread, channel model, angle-of-arrival estimation, multibeam.

## 1. INTRODUCTION

The implementation of smart antennas at macrocellular base stations (BSs) is expected significantly to enhance the capacity of wireless networks [1, 2]. Various algorithms for adaptive array signal processing have been proposed and investigated [2, 3, 4]. The effectiveness of these algorithms depends on the behaviour of the fading channel and in particular on the degree of azimuthal dispersion in the channel. Therefore, accurate statistical channel models are required for the testing of these adaptive algorithms. These models must be realistic and close to real-life channels in order to replicate the angle of arrival (AoA) distribution of the multipath components.

The propagation channel between the BS and the user equipment (UE) is generally held to be reciprocal in most respects. However, the azimuthal angle dispersions seen at

the BS and UE antenna differ significantly from each other. The classical Clarke channel model [5] assumes a uniform probability density function (pdf) of the incoming rays at the UE antenna. However, if the BS antenna array is elevated above the surrounding scatterers, then the rays incoming to the BS are concentrated in some smaller range of azimuth angles than those incoming to the UE. Note also that Clarke’s model provides the well-known “rabbit-ear” characteristic of the classical Doppler spectrum of signals seen both at the BS and at the UE. Some statistical propagation models which include the azimuthal dispersion at the BS have been developed in [6, 7, 8]. For example, the channel model proposed in [7] is based on a geometrical construction, and assumes that scatterers are uniformly distributed within the area of a circle centred at the UE antenna. This means that the AoA of the multipath components at the BS will be restricted to an angular region dependent both upon the circle

radius and upon the distance between BS and user. However, in a real-life channel, the scatterer distribution around the UE can differ significantly from uniform. Therefore, other researchers [9, 10, 11] have proposed other more realistic models based on a Gaussian distribution of scatterer location.

The goal of this paper is to analyse further the Gaussian proposal for the scatterer distribution. We assume that the scatterers can be situated in *any* point in the horizontal plane. In this model, the probability of occurrence of the scatterer location decreases in accordance with a Gaussian law when its distance from the UE antenna increases. Therefore, we call this model the “Gaussian channel model (GCM).” We believe that such an assumption about the scatterer location is closer to the real-life environment than some of the other models mentioned above. Therefore, as we will demonstrate later, the comparison of the obtained pdf of AoA of the multipath for the GCM with the measured results presented in [8] gives very good agreement. Note also that, like Clarke’s model, the proposed GCM also provides the classical Doppler signal spectrum.

It is a likely supplementary requirement for future cellular communication systems that they will be capable of determining the user position within a cell site. One way of doing this is via “triangulation,” whereby the angular bearing of the user is estimated at multiple cellsites (this process is also known as “direction finding”). UE position is estimated as the point where these bearing lines intersect. Thus, in order to carry out triangulation, an estimate of the AoA of the UE signal is required. We consider the “sum-difference bearing method” (SDBM) algorithm for AoA estimation. It was selected from a number of techniques that had been investigated (see, e.g., [12, 13, 14]). The SDBM algorithm is similar to the principle used in monopulse tracking radars, wherein a hybrid junction is used to extract the sum and difference of a received pulse [12]. Note that the tracking radar is able to serve just one user. However, the multibeam antenna arrays at the BS can serve all the users located in the given cell. More details of this SDBM algorithm will be provided later.

One of the major aims of the BS is to achieve a high capacity. To maximise the downlink capacity, it has been proposed elsewhere to use multibeam or beamformed antenna arrays to cover each sector of the cell handled by the BS [15]. Such an antenna array could also be applied to estimate the AoA. Therefore, in this paper, the dependence of the AoA estimation accuracy on the spread source model is also considered for the BS using a multibeam antenna. In this configuration, the beamformer creates three fixed beams per 120°-azimuth sector, generated from a facet containing 6-off  $\lambda/2$ -spaced columns of dual-polar antenna elements. These beams improve the coverage and capacity of the macrocell, and are expected to have greatest application within the urban macrocellular environment, where the need for maximum capacity is the greatest. Simulation results are presented for the case of a Rayleigh fading channel and for this antenna configuration.

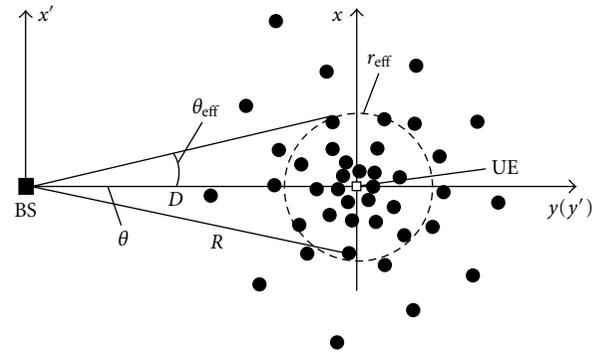


FIGURE 1: Illustration of the Gaussian channel model.

## 2. GAUSSIAN CHANNEL MODEL AND THE PDF OF THE AOAS OF MULTIPATH COMPONENTS SEEN AT THE BASE STATION

The signal received by the BS is a sum of many signals reflected from different scatterers randomly situated around the UE antenna. The AoAs of the multipath signal components are thus various and random. Therefore, the set of the scatterers can be considered collectively as a spread source, and the angle spread is a measure used to determine the angular dispersion of the channel.

Here we present the details of the GCM and derive an analytical expression for the pdf of the AoAs of multipath components as observed at the BS.

First of all, we list the initial assumptions used for creating the channel model. We assume that

- (i) the scattered signals arrive at BS in the horizontal plane, that is, the proposed GCM is two dimensional and the elevation angle is not taken into account;
- (ii) each scatterer is an omnidirectional reradiating element and the plane wave is reflected directly to the BS without influence from other scatterers (i.e., we have only “single-bounce” scattering paths);
- (iii) the direct path from the UE to BS antenna is infinitely attenuated;
- (iv) the reflection coefficient from each scatterer has unity amplitude and random phase;
- (v) the probability of the (random) scatterer location is independent of azimuth angle (from the UE), and decreases if its distance from the UE antenna increases. This dependence has a Gaussian form.

The last of these assumptions distinguishes our channel model from many of the other known models [5, 6, 7].

Thus we can write that

$$p(r, \varphi) = \frac{1}{\pi r_{\text{eff}}^2} \exp\left(-\frac{r^2}{r_{\text{eff}}^2}\right), \quad (1)$$

where  $(r, \varphi)$  is the polar coordinate system centred at the UE,  $r$  is the distance to a given scatterer from the UE antenna, and  $r_{\text{eff}}$  is the radius at which the pdf decreases by a factor of  $e$ , that is,  $p(r_{\text{eff}}, \varphi) = e^{-1}p(0, \varphi)$ . Figure 1 illustrates the GCM,

where  $D$  is the distance between the BS and UE antennas, and  $(x, y)$  are the rectangular coordinates.

In [7], a uniform scatterer distribution within the circle of radius  $r_0$  around the UE was assumed. So for the model of [7], this means that the AoAs of multipath components seen at the BS are limited to the angular region  $[-\theta_{\max} \dots \theta_{\max}]$ , where  $\theta_{\max} = \sin^{-1}(r_0/D)$ . However, for our GCM model, the AoAs of scattered signals as received at the BS are *not* restricted to any constrained angular region.

In order to derive the ensemble pdf of the AoA for the GCM (i.e., averaged over many model realisations), we choose the origin of the system coordinates  $(x', y')$  to be the location of the BS. This means that  $x' = x$  and  $y' = y + D$ . We then transform to the polar coordinates  $(R, \theta)$ , where  $x' = R \sin \theta$ ,  $y' = R \cos \theta$ , and the angle  $\theta$  is measured relative to the line joining the BS and UE antennas. It is straightforward to show that the Jacobian of this transformation is equal to  $R$ . Furthermore, we have

$$r^2 = x^2 + y^2 = x'^2 + (y' - D)^2 = R^2 - 2RD \cos \theta + D^2. \quad (2)$$

As a result of substituting (2) into (1), we obtain that

$$p(R, \theta) = \frac{R}{\pi r_{\text{eff}}^2} \cdot \exp\left(-\frac{D^2}{r_{\text{eff}}^2}\right) \cdot \exp\left(-\frac{R^2 - 2RD \cos \theta}{r_{\text{eff}}^2}\right). \quad (3)$$

In order to derive the one-dimensional pdf of the AoA (i.e., the power angle density) of the multipath components as seen at the BS, an integration over the radius  $R$  must be carried out. Therefore, the pdf is expressed as the following integral:

$$p(\theta) = \int_0^\infty p(R, \theta) dR = \frac{1}{\pi r_{\text{eff}}^2} \cdot \exp\left(-\frac{D^2}{r_{\text{eff}}^2}\right) \int_0^\infty \exp\left(-\frac{R^2 - 2RD \cos \theta}{r_{\text{eff}}^2}\right) R dR. \quad (4)$$

This integral can be calculated analytically and a closed-form solution is obtained. To do this, take into account that (see [16, equation 3.462.1])

$$\int_0^\infty x^{\nu-1} \exp(-\beta x^2 - \gamma x) dx = (2\beta)^{-\nu/2} \Gamma(\nu) \exp\left(\frac{\gamma^2}{8\beta}\right) C_{-\nu}\left(\frac{\gamma}{\sqrt{2\beta}}\right), \quad (5)$$

where  $\text{Re}(\nu, \beta) > 0$ ,  $\Gamma(\nu)$  is the gamma function, and  $C_p(z)$  is the function of the parabolic cylinder. In our case, we have  $\nu = 2$ ,  $\beta = r_{\text{eff}}^{-2}$ , and  $\gamma = -2Dr_{\text{eff}}^{-2} \cos \theta$ . If  $\nu = 2$ , then the function  $C_{-2}(z)$  can be expressed in terms of the probability

integral  $\Phi(z)$  (see [16, equation 9.254.2]<sup>1</sup>), that is,

$$C_{-2}(z) = -\exp\left(\frac{z^2}{4}\right) \sqrt{\frac{\pi}{2}} \left\{ z \left[ 1 - \Phi\left(\frac{z}{\sqrt{2}}\right) \right] - \sqrt{\frac{2}{\pi}} \exp\left(-\frac{z^2}{2}\right) \right\}, \quad (6)$$

where the probability integral  $\Phi(x) = (2/\sqrt{\pi}) \int_0^x \exp(-t^2) dt$ .

Take into account that  $z = -\sqrt{2}Dr_{\text{eff}}^{-1} \cos \theta$ ,  $\Gamma(2) = 1$ , and  $\Phi(z)$  is an odd function of its argument  $z$ . As a result of straightforward transformations, we can obtain from (5) and (6) that the desired one-dimensional pdf  $p(\theta)$  of AoA of the multipath components is given by

$$p(\theta) = \frac{1}{2\pi} \cdot \exp\left(-\frac{D^2}{r_{\text{eff}}^2}\right) \times \left\{ 1 + \sqrt{\pi} \frac{D}{r_{\text{eff}}} \cos \theta \cdot \exp\left(\frac{D^2}{r_{\text{eff}}^2} \cos^2 \theta\right) \cdot \left[ 1 + \Phi\left(\frac{D}{r_{\text{eff}}} \cos \theta\right) \right] \right\}. \quad (7)$$

It is convenient to introduce the angle  $\theta_{\text{eff}} = \sin^{-1}(r_{\text{eff}}/D)$ . Then (7) can be rewritten as

$$p(\theta) = \frac{1}{2\pi} \cdot \exp\left(-\frac{1}{\sin^2 \theta_{\text{eff}}}\right) \times \left\{ 1 + \sqrt{\pi} \frac{\cos \theta}{\sin \theta_{\text{eff}}} \cdot \exp\left(\frac{\cos^2 \theta}{\sin^2 \theta_{\text{eff}}}\right) \cdot \left[ 1 + \Phi\left(\frac{\cos \theta}{\sin \theta_{\text{eff}}}\right) \right] \right\}. \quad (8)$$

Thus the pdf  $p(\theta)$  depends only upon  $\cos \theta$ . The effective angle spread for this pdf can be introduced as  $\Delta = 2\theta_{\text{eff}}$ . The pdf  $p(\theta)$  is an even function of its argument  $\theta$ .

The expression (8) is true in the general case. However, this formula takes a very simple form for the case of small angle spread  $\theta_{\text{eff}} \ll \pi$  when  $\sin \theta \approx \theta$ . In this case, the pdf is approximately given by

$$p(\theta) \approx \frac{1}{\sqrt{\pi} \theta_{\text{eff}}^2} \cdot \exp\left(-\frac{\theta^2}{\theta_{\text{eff}}^2}\right) \quad (9)$$

and described by a (one-dimensional) Gaussian pdf with zero mean and variance  $\sigma^2 = 0.5\theta_{\text{eff}}^2$ .

Figure 2 shows the pdf  $p(\theta)$  of the AoA of the multipath components for the different values  $\theta_{\text{eff}} = 10^\circ, 30^\circ$ , and  $50^\circ$ . The solid and dashed curves correspond to the exact formula (8) and to its Gaussian approximation (9), respectively. We can see that the exact and Gaussian PDFs are very close to each other for a large interval of  $\theta_{\text{eff}}$  up to  $\theta_{\text{eff}} \leq 0.5$  (or  $\theta_{\text{eff}} \leq 30^\circ$ ). Actually, it is quite simple and intuitive to see how the complex pdf of the exact formula (8) should

<sup>1</sup>N.B. There is a minor typographical error (a missing factor of  $-1$ ) in the version of this equation printed in [16], which is corrected within the addenda of the original Russian version.

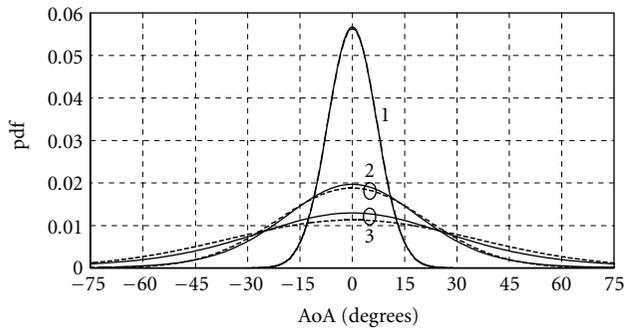


FIGURE 2: The pdf of the AoA of the multipaths at the BS. The angle spread is equal to 20, 60, and 100 degrees (curves 1, 2, 3, respectively). The solid and dashed curves correspond to the exact formula (8) and its Gaussian approximation (9), respectively.

equal a one-dimensional pdf for small angle spreads. At these small angles, the lines bounding different small “slices” of the two-dimensional pdf are nearly parallel, and so it is as if we are calculating the marginal pdf of the two-dimensional spatial pdf along the  $x$ -axis. Since the marginal pdf of a two-dimensional Gaussian distribution is a one-dimensional Gaussian distribution, our approximate result (9) is intuitively of the correct form.

The comparison of the theoretical pdf against real measurement data is of course of interest in order both to validate and to parameterise the GCM. Histograms of the estimated azimuthal power angle density and scatterer occurrence probabilities are presented by the authors of [8]. This measurement data was obtained in Aarhus with a BS antenna located 12 m above the rooftop level. We wish to take this measured data and compare it to the three proposed theoretical channel models: (1) our GCM of (8), (2) the geometrical-based single-bounce model (GBSBM) developed in [7] (in which the scatterers are assumed to be uniformly randomly distributed within the area of a circle), and (3) Clarke’s model [5, 17] (in which the scatterers are assumed to lie on the circumference of a circle).

It was derived in [7] that the pdf of the AOA of the multipath components for GBSBM is given by

$$p(\theta) = \begin{cases} \frac{2 \cos(\theta) \sqrt{\sin^2 \theta_{\max} - \sin^2 \theta}}{\pi \sin^2 \theta_{\max}}, & -\theta_{\max} \leq \theta \leq \theta_{\max}, \\ 0, & \text{otherwise,} \end{cases} \quad (10)$$

where  $\theta_{\max} = \sin^{-1}(r_0/D)$  and  $r_0$  is the radius of the circle within which all the scatterers are uniformly distributed.

Whilst we omit the derivation here, for reasons of brevity, it can be shown that the pdf of the AOA of the multipath components for Clarke’s model is equal to

$$p(\theta) = \frac{1}{\pi} \frac{1}{\cos^2 \theta} \frac{1}{\sqrt{\tan^2 \theta_{\max} - \tan^2 \theta}}, \quad (11)$$

where in this case, when calculating  $\theta_{\max}$ ,  $r_0$  has the meaning

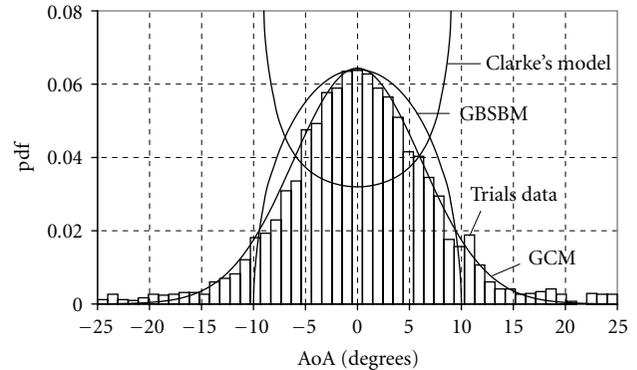


FIGURE 3: The PDFs for the AoA of the multipath components at the BS for GCM, GBSBM, Clarke’s models, and for the measured histograms.

of the radius of the circle periphery on which the scatterers are uniformly distributed.

Figure 3 shows the PDFs for the AoA of the multipath components at the BS for GCM, GBSBM, Clarke’s models, and the measured scatterer occurrence probability histograms taken from [8]. We have chosen the model parameters ( $\theta_{\max}$ ,  $\theta_{\text{eff}}$ ) so that the best agreement was obtained for each model. For both the GBSBM and Clarke’s models, the value chosen was  $\theta_{\max} = 10^\circ$ , and for GCM,  $\theta_{\text{eff}} = 8.8^\circ$ . It can be seen that the GCM ensures the best agreement with real-life results for the whole angular region and especially for the tails of histogram. Clarke’s model produces the worst match to the real-life data.

The measured data and experimental models described above discuss the “ensemble” statistics of the spread source. By ensemble statistics, we mean that these statistics are averaged over a large number of individual measurements or individual model realisations. However, in practice, we would deal with single cases (i.e., in “real-life”) or single-model realisations (i.e., during simulation). It seems reasonable to postulate that the angle-spread behaviour of the source will be nonergodic. That is to say, the statistics of any given realisation (averaged over time) will, in general, be different from the *ensemble* statistics (averaged over *all* realisations and *all* time). So in practice, in any single realisation of the angle-spread model, we will see a limited number of *discrete* scattering centres creating a “lumpy” AoA distribution function, rather than an infinite number of scatterers creating a continuous “smooth” distribution, as observed from the ensemble statistics. If this limited number of discrete scattering centres is particularly small, then their “centre of gravity (CofG)” may “wander” about the true bearing of the UE. The CofG, to be defined in more detail below, is simply a power-weighted average AoA. As an example, in one realisation of the scattering model, *all* of the scattering centres may, purely by chance, be located on the left-hand side of the true UE bearing, which would bias the *apparent* (i.e., estimated) bearing of the UE to the left. Conversely, in another realisation, all of the scattering centres may, again by chance, be located on the right-hand side of the true UE bearing, which would bias the apparent bearing of the UE to the right. So this

apparent change of the UE bearing for different realisations of the scattering model, which we term the “wandering” of the “CofG” is a direct consequence of the nonergodicity of the angle-scattering model. This wandering is more marked when the mean number of scattering sources is low, because if we have a large number of scattering sources, then it would be *extremely* unlikely for *all* of them to be lying on the same side of the UE (assuming that all scatterer locations are independent). In fact, we will show later that this “wandering of the CofG” phenomenon is a significant contributor to the overall estimation error of the UE bearing.

For reasons described above, the variance of the wandering of the CofG depends on the number of scatterers situated around the UE antenna. Let  $N$  be the number of scatterers and  $\theta_1, \theta_2, \dots, \theta_N$  some random values of AoAs of the signal from these scatterers. Assume, for simplicity, that all of the sources have equal power. Then the CofG of the received signal for this particular realisation is equal to

$$\tilde{\theta} = \frac{1}{N}(\theta_1 + \theta_2 + \dots + \theta_N). \quad (12)$$

The expectation of the random value  $\tilde{\theta}$  is equal to zero (i.e.,  $\langle \tilde{\theta} \rangle = 0$ ) and its variance can be obtained from the integral

$$\sigma_{N\theta}^2 = \iint \dots \int \frac{1}{N^2} (\theta_1 + \theta_2 + \dots + \theta_N)^2 \times p(\theta_1, \theta_2, \dots, \theta_N) d\theta_1 d\theta_2 \dots d\theta_N, \quad (13)$$

where  $p(\theta_1, \theta_2, \dots, \theta_N)$  is the joint pdf of the AoAs  $\theta_1, \theta_2, \dots, \theta_N$ . Since these AoAs are assumed to be independent random values, the joint pdf can be presented as the product of individual PDFs, that is,  $p(\theta_1, \theta_2, \dots, \theta_N) = p(\theta_1)p(\theta_2) \dots p(\theta_N)$ , where the function  $p(\theta_i)$  ( $i = 1, 2, \dots, N$ ) is given by formula (8).

The *expected* azimuth angle of each angle-spread source is equal to zero due to the symmetry of the pdf (8) of the multipath component AoAs, that is,  $\langle \theta_i \rangle = 0$ . Thus the  $N$ -dimensional integral (13) can be rewritten as the sum of  $N$  identical one-dimensional integrals, that is,

$$\sigma_{N\theta}^2 = \frac{1}{N^2} \sum_{i=1}^N \int \theta_i^2 p(\theta_i) d\theta_i = \frac{\sigma_{1\theta}^2}{N}, \quad (14)$$

where  $\sigma_{1\theta}^2$  is the variance of the AoA of a single scatterer, equal to

$$\sigma_{1\theta}^2 = \int \theta^2 p(\theta) d\theta \quad (15)$$

and pdf  $p(\theta)$  is defined by formula (8).

So (14) and (15) give the mean squared value for the wandering of the CofG of the spread source when we assume  $N$  scatterers of the same amplitude.

For small  $\theta_{\text{eff}} \ll 1$ , the pdf  $p(\theta)$  has Gaussian form (9). Substituting (9) into (15) and carrying out the integration

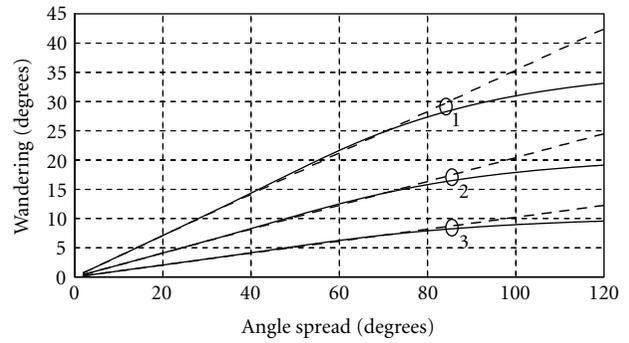


FIGURE 4: The source C of G wandering versus angle spread  $\Delta$  for the different numbers of scatterers  $N = 1, 3, 12$  (curves 1, 2, 3, respectively). The solid and dashed curves correspond to the exact formula (8) and its Gaussian approximation (9), respectively.

in (15), we obtain that  $\sigma_{1\theta} = \theta_{\text{eff}}/\sqrt{2}$ . Hence it can be found from (14) that the wandering of the CofG is equal to

$$\sigma_{N\theta} = \frac{\theta_{\text{eff}}}{\sqrt{2N}}. \quad (16)$$

Figure 4 shows the wandering  $\sigma_{N\theta}$  of the CofG of the source versus angle spread  $\Delta$  for different numbers of scatterers  $N = 1, 3, 12$  (curves 1, 2, 3). The solid and dashed curves correspond to the exact formula (8) and its Gaussian approximation (9), respectively. We can see that the exact and Gaussian PDFs are very close to each other for a large interval of  $\theta_{\text{eff}}$  up to  $\approx 40^\circ$ .

The CofG of the scattering sources gives the best unbiased estimate of the true UE bearing, albeit that it is an estimate with high variance (i.e., high mean squared error) when the number of scattering centres is small. So the aim of our AoA estimation processing is to estimate this CofG from a limited-time snapshot of noisy received signal. The receiver noise will add an additional error term to the final bearing estimation error. However, it can be seen from the foregoing analysis that even using “perfect” CofG estimation algorithms on long samples of high signal-to-noise-ratio (SNR) received signal, there will still be a residual irreducible error if the number of scattering centres is small. This is because of the wandering of the CofG, which in turn is due to the nonergodicity of the spread source.

### 3. AOA ESTIMATION INCORPORATING THE GCM

We have stated above that the best estimate of the true UE bearing is given by estimating the CofG of the received signal (i.e., for a given single realisation of the scattering). However, even using a “perfect” AoA estimation algorithm, we would suffer from irreducible errors due to the “wandering” of the scatterer CofG. For reasons of implementation simplicity, we may well in practice contemplate using a less-than-perfect AoA estimation algorithm if (a) the implementation of this less-than-perfect algorithm is simple, and hence cheap to implement, and (b) the *additional* errors introduced by

the less-than-perfect algorithm (compared to an optimal algorithm) are small compared to the irreducible CofG wandering error which we must allow for in any case. So in this section, we consider just such a simplified AoA estimation process, which we term SDBM. This method was selected from a number of similar techniques which had been investigated because it was found to give the overall most accurate and most robust performance. The mathematical details of the SDBM technique will be presented later. However, the essence of the technique is to measure, average, and compare received signal powers (or amplitudes) received at the BS, as measured in adjacent beams. We assume, for the use of SDBM, that the BS already employs a multibeam antenna (typically with three deep-cusp beams) in each 120°-azimuth sector. The scattered signal from the user is received by each of the beams of the antenna, and the two adjacent beams receiving the highest signal powers are selected. For these beams, a set of functions, which we term “bearing curves,” must be precalculated and stored. The exact form of these bearing curves depends upon the multibeam antenna patterns and upon the expected ensemble angle-spread distribution (which we argued earlier tends to Gaussian form at small angle spreads).

First of all, we determine the dependence of the average received power  $G$  at an arbitrary beam output on the angle location of the source with an angle spread  $\Delta$ . Let  $F(\theta)$  be the reception gain pattern of this beam and  $\theta_0$  be the centre of the spread source (i.e., the “true” UE bearing). Then the function  $G(\theta_0)$  can be presented in form of a mathematical convolution of (i) a function representing the power beam pattern  $|F(\theta_0)|^2$  of this beam as a function of the azimuth angle ( $\theta$ ) and (ii) a function  $p(\theta)$  representing the (ensemble) pdf of the AoAs of signals received by the BS due to reflections from scatterers as a function of azimuth angle ( $\theta$ ), that is,

$$G(\theta_0) = \int_0^\pi |F(\theta)|^2 p(\theta - \theta_0) d\theta. \quad (17)$$

We can refer to the function (17) as a “beam pattern for a spread source,” that is, what we call a “spread” beam pattern. If the spread of signals is a negligibly small quantity ( $\theta_{\text{eff}} \rightarrow 0$ ), then we have a point source, and the pdf  $p(\theta)$  in (8) tends to a delta function (i.e.,  $p(\theta) \rightarrow \delta(\theta - \theta_0)$ ). In this case, the function  $G(\theta_0)$  is given by  $G(\theta_0) = |F(\theta_0)|^2$ , that is, it is simply equal to the power gain pattern of the beam, or to what we will term the “point source” beam pattern.

Now we provide the mathematical definition of what we have termed earlier the “bearing curves.” If  $L$  is the number of the beams generated by the multibeam antenna, then we have a set of beam patterns  $G_i(\theta)$  ( $i = 1, 2, \dots, L$ ) and each beam pattern is oriented in a given direction. The bearing curves  $b_{i+1,i}$  ( $i = 1, 2, \dots, L - 1$ ) for each adjacent beam pair ( $i + 1, i$ ) may be represented by a function  $b_{i+1,i}(\theta)$  of the azimuth angle  $\theta$  of the antenna according to the following equation:

$$b_{i+1,i}(\theta) = \frac{\sqrt{G_i(\theta)} - \sqrt{G_{i+1}(\theta)}}{\sqrt{G_i(\theta)} + \sqrt{G_{i+1}(\theta)}}. \quad (18)$$

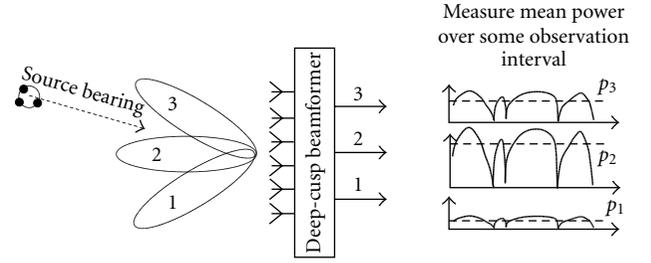


FIGURE 5: Applying the SDBM algorithm.

These bearing curves are precalculated and stored by the network. The precalculation takes place based on equation (17), and hence takes into account both the known multibeam patterns and the *expected* angle-spread distribution of the scattering channel (which we model as Gaussian with a given  $\theta_{\text{eff}}$ ). There is more discussion later about how we determine the expected angle spread.

To estimate the bearing of any given source, the received power from each beam of the antenna is measured over a predetermined observation interval by averaging over a large number of samples. The observation interval should be chosen to be long enough so that the effects of Doppler signal fading do not significantly impact the measured power.

The application of SDBM algorithm is shown in Figure 5. Let  $p_i = |s_i(t) + n_i(t)|^2$  be the mean power measured at the output of the  $i$ th ( $i = 1, 2, \dots, L$ ) antenna beam, where  $s_i(t)$  and  $n_i(t)$  are the useful signal and additive white Gaussian noise (AWGN), respectively. The AWGN variance  $\sigma_0^2$  is assumed to be the same for all of the different antenna beams. The bearing curves, per (18), are produced without regard to AWGN. That is to say, they only take into account ratios of sums and differences of expected *signal* amplitudes (without including noise or interference contributions). Therefore, for a more accurate estimation of AoA based on measured noisy samples, we need to take into account an *expected* noise power contribution for the measured signal, the value of which we subtract from the *measured* power signal of each beam after the averaging. In practice, this means that we use an estimated output signal power equal to  $\tilde{p}_i = |p_i - \sigma_0^2|$ . The estimates  $\tilde{p}_i$  for all  $i = 1, 2, \dots, L$  are compared with each other and the two adjacent beams receiving the highest signal powers are selected. If the  $j$ th and  $(j + 1)$ th beams have the highest output powers, then the sum-difference ratio  $\hat{b}_{j+1,j} = (\sqrt{\tilde{p}_j} - \sqrt{\tilde{p}_{j+1}}) / (\sqrt{\tilde{p}_j} + \sqrt{\tilde{p}_{j+1}})$  is calculated and the AoA is estimated by looking up the bearing  $\theta$  corresponding to this ratio from the corresponding bearing curve  $b_{j+1,j}(\theta)$  of (18).

Now we present simulation results for the SDBM technique in order to estimate the accuracy which can be achieved. Any one of a number of possible multibeam antenna designs could have been assumed for this simulation, but for this work, we have used the “deep-cusp” multibeam antenna design of [15]. The deep-cusp beamformer creates three fixed beams per each 120°-azimuth sector, generated from a facet containing 6-off  $\lambda/2$ -spaced columns of

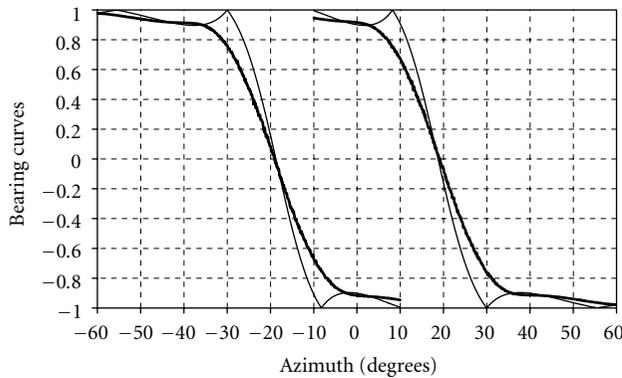


FIGURE 6: Bearing curves  $b_{21}(\theta)$  (left-hand curves) and  $b_{32}(\theta)$  (right-hand curves) for the point ( $\Delta = 0^\circ$ ) and spread ( $\Delta = 17^\circ$ ) sources (thin and thick curves, respectively).

dual-polar antenna elements (although only a single polarisation is considered here). The angular spread of the source will be assumed to be equal to  $17^\circ$ , which corresponds to experimental results obtained in [8]. Two representative cases, for which the number of scatterers is specified as  $N = 3$  and  $N = 12$ , will be simulated. There are two bearing curves  $b_{21}(\theta)$  and  $b_{32}(\theta)$  for the antenna configuration with three beams.

The bearing curves  $b_{21}(\theta)$  and  $b_{32}(\theta)$  for the point ( $\Delta = 0^\circ$ ) and spread ( $\Delta = 17^\circ$ ) sources are presented in Figure 6 (thin and thick curves, respectively). The left-hand curves are  $b_{21}(\theta)$  and the right-hand curves are  $b_{32}(\theta)$ . It can be seen that these bearing curves have the steepest slope at the points where the beams cross. Estimation of the bearing of the point source is possible only in the angle intervals  $[-30^\circ, -10^\circ]$  and  $[10^\circ, 30^\circ]$ . For the spread source, estimation of the bearing is possible over wider angle intervals  $[-35^\circ, 35^\circ]$ . It is assumed, of course, that to estimate the bearing of UEs for angles outside this range, we would construct additional bearing curves relating to the beam at the edge of this sector and its neighbour at the edge of the adjacent sector.

When estimating the AoA, the estimates  $\tilde{p}_1$ ,  $\tilde{p}_2$ , and  $\tilde{p}_3$  of the mean signal power at the output of the  $i$ th ( $i = 1, 2, 3$ ) antenna beam are compared with each other. If  $\tilde{p}_1 > \tilde{p}_3$ , then the ratio  $\hat{b}_{21}$  is calculated and the AoA is estimated using the bearing curve  $b_{21}(\theta)$ . If  $\tilde{p}_1 < \tilde{p}_3$ , then the value  $\hat{b}_{32}$  is calculated and the AoA is estimated according to the bearing curve  $b_{32}(\theta)$ .

Within the simulations, the samples of the complex signals were generated with a sampling period equal to 1 millisecond for three antenna beams. The maximum Doppler frequency  $f_d$  was set equal to 50 Hz. The observation interval was chosen to be 400 milliseconds, that is, approximately 50 times longer than the fading correlation interval. Various SNRs equal to 30, 20, 10 and 0 dB were simulated, where the SNR is defined by what the received SNR is for a point source located at the peak of the central beam. In order to average the results over all source directions, the true source angle  $\theta_{\text{true}}$  was varied from  $-40^\circ$  to  $+40^\circ$  with a step size equal to  $0.5^\circ$ . A thousand experiments were carried out for

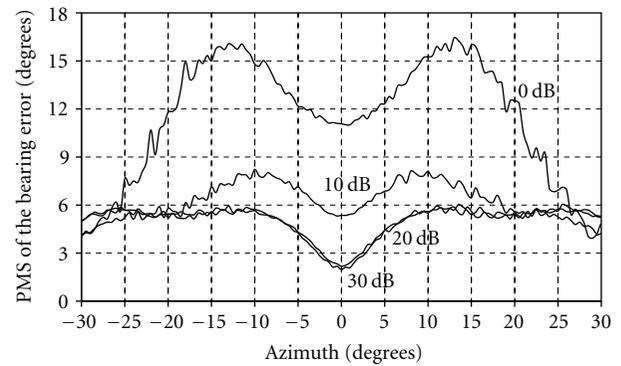


FIGURE 7: The rms of bearing estimation error for various SNRs and for the number of scatterers  $N = 3$ .

each source direction, and different realisations of the (non-ergodic) source model were applied for each of these experiments. For each source position, the root-mean-square (rms)  $\Delta_\theta$  of the bearing estimation error and the cumulative density function (CDF) of absolute value of AoA estimation error  $|\hat{\theta}_j - \theta_{\text{true}}|$  were calculated.

The rms of the bearing estimation error is shown in Figure 7 for the number of scatterers  $N = 3$  and for the given SNRs. We can see that, as expected, the rms of the bearing estimation error decreases when the SNR increases. For large SNRs (20 and 30 dB), the bearing estimation error lies within the range  $2^\circ$  to  $6^\circ$  (depending on the true source bearing) and is solely due to the random wandering of the CofG of the angle-spread source. For the lower SNRs, the bearing estimation error is larger, and depends also on AWGN power. The corresponding CDFs are presented in Figure 8. The CDFs in Figure 8 can be approximated by the CDF of a Gaussian function. Using this Gaussian approximation, we obtain that the standard deviation of the bearing estimation error is  $\approx 4^\circ$  for high SNRs and  $N = 3$ . As can be seen from Figure 4 (curves 2), this standard deviation is approximately equal to the standard deviation of the wandering of the CofG of the source with an angle spread  $\Delta = 2\theta_{\text{eff}} = 17^\circ$  ( $\theta_{\text{eff}} = 8.5^\circ$ ). Thus we can see that the bearing estimation error for high SNRs is conditioned by the nonergodicity of the source model. The highest bearing estimation errors are observed in the crossing area of the antenna beam patterns. This is because the beam gains are lower in this angular region, and so the effective received SNR is also lower in this region compared to what it would be for a source located close to the peak of the central beam. The CDF of the bearing estimation error for a larger number of scatterers  $N = 12$  is also shown in Figure 8. Compared to the results for  $N = 3$ , the standard deviation of the bearing estimation error has decreased by a factor of approximately two for high SNRs, from  $\approx 4^\circ$  to  $\approx 2^\circ$ . Like the results for  $N = 3$ , this also corresponds to Figure 4 and (14).

As is evident from the earlier discussion, the form of the bearing curves is different for different assumed channel angle spreads. This is because the first stage of the generation

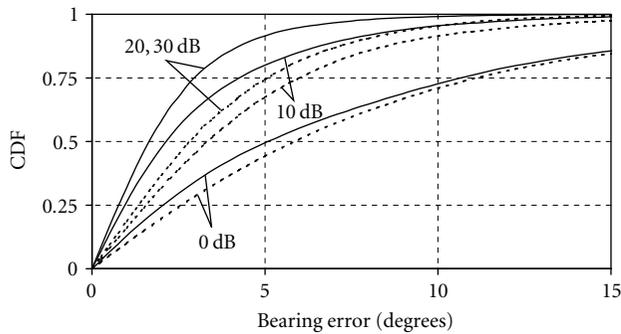


FIGURE 8: The CDFs of the bearing estimation error for various SNRs. The number of scatterers is  $N = 12$  (solid curves) and  $N = 3$  (dashed curves).

of the bearing curves involves a convolution of the *actual* beam pattern with the *assumed* angle-spread ensemble pdf. What if we didn't apply the preconvolution in the generation of the bearing curves, but simply used the bearing curve corresponding to "point source" beam patterns, even when the channel itself *does* exhibit angle spread? To answer this, it is interesting to examine the bearing errors when bearing curves generated for the point source are actually used for estimating AoA in a channel *with* angle spreading. Such comparative simulation results for the CDF of the bearing error are presented in Figure 9 for  $\text{SNR} = 30$  dB and number of scatterers  $N = 12$ . The angle spread in the channel is equal to  $17^\circ$ . We can see that the bearing error has increased significantly due to the use of "nonmatched" bearing curves. In order to generate "matched" bearing curves, we need at least to have a reasonable estimate of the (ensemble) angle spread of the channel. In practice, this would be obtained through examination of published measured angle-spread data such as [8], and by matching the environment in which the multibeam BS is deployed (e.g., urban, suburban, rural) to the expected angle spread of the channel.

#### 4. CONCLUSIONS

In this paper, we have developed a model for an angle-spread source which we term the Gaussian channel model (GCM). This model is suitable for representing the signal seen at the base station (BS) antenna, and assumes that the probability of the scatterer occurrence decreases in accordance with a Gaussian law when its distance from the user equipment (UE) antenna increases. Such an assumption about the scatterer location is closer to the real-life environment than some of the other known models. An analytical expression for the probability density function (pdf) of the multipath angle of arrival (AoA) at the BS has been derived for the general case of an arbitrary angle spread. It is shown that this pdf can be approximated by a Gaussian curve for sources with a small spread. The comparison of the obtained pdf of AoA of the multipath for the GCM with the published experimental results gives a better agreement than for some other known

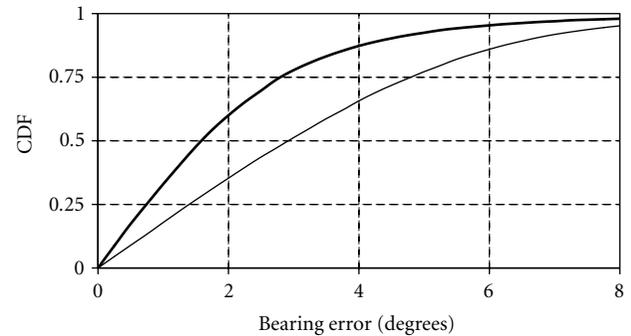


FIGURE 9: The CDF of the bearing estimation error using the "spread" bearing curve (thick curve) and "point source" bearing curve (thin curve) for  $\text{SNR} = 30$  dB, angle spread  $\Delta = 17^\circ$ , and number of scatterers  $N = 12$ .

angle scattering models. However, in a real-life situation, we deal with a single realisation of the angle-spread source, that is, with a fixed finite number of *discrete* scattering centres. If this number is particularly small, then their center of gravity (CofG), defined as a power-weighted average AoA, may "wander" about the true bearing of the UE. The variance of this wandering of the CofG has been obtained. The dependence of the AoA estimation accuracy on the parameters of the spread source model has also been considered for a BS using a multibeam antenna, by carrying out simulations of the so-called sum-difference bearing method (SDBM) AoA estimation algorithm. It has been shown that for high SNRs, the bearing estimation errors are dominated by the wandering of the CofG of the spread source. This wandering is a consequence of the nonergodicity of the angle scattering process and is greater when the number of scattering sources is small.

#### REFERENCES

- [1] J. C. Liberti and T. S. Rappaport, *Smart Antennas for Wireless Communications: IS-95 and Third Generation CDMA Applications*, Prentice Hall, Upper Saddle River, NJ, USA, 1999.
- [2] J. B. Andersen, "Antenna arrays in mobile communications: gain, diversity, and channel capacity," *IEEE Antennas and Propagation Magazine*, vol. 42, no. 2, pp. 12–16, 2000.
- [3] U. Vornefeld, C. Walke, and B. Walke, "SDMA techniques for wireless ATM," *IEEE Communications Magazine*, vol. 37, no. 11, pp. 52–57, 1999.
- [4] R. A. Soni, R. M. Buehrer, and R. D. Benning, "Intelligent antenna system for cdma2000," *IEEE Signal Processing Magazine*, vol. 19, no. 4, pp. 54–67, 2002.
- [5] R. H. Clarke, "A statistical theory of mobile-radio reception," *Bell System Technical Journal*, vol. 47, no. 6, pp. 957–1000, 1968.
- [6] J. C. Liberti and T. S. Rappaport, "A geometrically based model for line-of-sight multipath radio channels," in *Proc. IEEE 46th Vehicular Technology Conference*, vol. 2, pp. 844–848, Atlanta, Ga, USA, April 1996.
- [7] P. Petrus, J. H. Reed, and T. S. Rappaport, "Geometrical-based statistical macrocell channel model for mobile environments," *IEEE Trans. Communications*, vol. 50, no. 3, pp. 495–502, 2002.
- [8] K. I. Pedersen, P. E. Mogensen, and B. H. Fleury, "A stochastic model of the temporal and azimuthal dispersion seen at

the base station in outdoor propagation environments," *IEEE Trans. Vehicular Technology*, vol. 49, no. 2, pp. 437–447, 2000.

- [9] J. Fuhl, A. F. Molisch, and E. Bonek, "Unified channel model for mobile radio systems with smart antennas," *IEE Proceedings Radar, Sonar and Navigation*, vol. 145, no. 1, pp. 32–41, 1998.
- [10] R. M. Buehrer, S. Arunachalam, K. H. Wu, and A. Tonello, "Spatial channel model and measurements for IMT-2000 systems," in *Proc. IEEE Vehicular Technology Conference*, vol. 1, pp. 342–346, Rhodes, Greece, May 2001.
- [11] Lucent Technologies, "Proposal for a spatial channel model in 3GPP RAN1/RAN4," Contribution WG1#20(01)579 of Lucent Technologies to 3GPP-WG1, Busan, May 2001.
- [12] M. I. Skolnik, Ed., *Radar Handbook*, McGraw-Hill, New York, NY, USA, 1970.
- [13] O. Besson, F. Vincent, P. Stoica, and A. B. Gershman, "Approximate maximum likelihood estimators for array processing in multiplicative noise environments," *IEEE Trans. Signal Processing*, vol. 48, no. 9, pp. 2506–2518, 2000.
- [14] S. Valaee, B. Champagne, and P. Kabal, "Parametric localization of distributed sources," *IEEE Trans. Signal Processing*, vol. 43, no. 9, pp. 2144–2153, 1995.
- [15] M. S. Smith, M. Newton, and J. E. Dalley, "Multiple beam antenna," US Patent number 6,480,524, November 2002.
- [16] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals Series and Products*, Academic Press, New York, NY, USA, 1965.
- [17] W. C. Jakes, Ed., *Microwave Mobile Communications*, John Wiley & Sons, New York, NY, USA, 1974.

**I. M. Averin** received his diploma (M.S.) in radiotechnics from Nizhny Novgorod Technical University in 2000. Since then, he has worked in the field of radio technology with the scientific and technical company "Mera," Nizhny Novgorod, Russia. He is also currently a postgraduate student in Nizhny Novgorod Technical University. His research interests include array signal processing, space-time spectral analysis, and wireless communications.



**D. D. N. Bevan** received his M.Eng. in electronic and electrical engineering from Loughborough University of Technology in 1991. Since then, he has worked in the field of radio technology within the Wireless Technology Laboratories of Nortel Networks in Harlow, UK. His research interests include system modelling, array signal processing, and technologies for current and future wide-area and local-area wireless networking.



**V. T. Ermolayev** received his Ph.D. and the Doctor of Science degrees in radiophysics from Nizhny Novgorod State University in 1980 and 1996, respectively. He has worked with the Radiotechnical Institute, State University, and the scientific and technical company "Mera," Nizhny Novgorod, Russia. His research interests include array signal processing, space-time spectral analysis, signal parameter estimation and detection, and wireless communications.



**A. G. Flaksman** received his Ph.D. degree in radiophysics from Nizhny Novgorod State University in 1983. He has worked with the radiotechnical Institute, State University, and the scientific and technical company "Mera," Nizhny Novgorod, Russia. His research interests include array signal processing, space-time spectral analysis, signal parameter estimation and detection, and wireless communications.



# Downlink Channel Estimation in Cellular Systems with Antenna Arrays at Base Stations Using Channel Probing with Feedback

**Mehrzhad Biguesh**

*Department of Communication Systems, University of Duisburg-Essen, Bismarckstrasse 81, 47057 Duisburg, Germany  
Email: biguesh@sent5.uni-duisburg.de*

**Alex B. Gershman**

*Department of Communication Systems, University of Duisburg-Essen, Bismarckstrasse 81, 47057 Duisburg, Germany  
Email: gershman@sent5.uni-duisburg.de*

*Received 21 May 2003; Revised 4 December 2003*

In mobile communication systems with multisensor antennas at base stations, downlink channel estimation plays a key role because accurate channel estimates are needed for transmit beamforming. One efficient approach to this problem is channel probing with feedback. In this method, the base station array transmits probing (training) signals. The channel is then estimated from feedback reports provided by the users. This paper studies the performance of the channel probing method with feedback using a multisensor base station antenna array and single-sensor users. The least squares (LS), linear minimum mean square error (LMMSE), and a new scaled LS (SLS) approaches to the channel estimation are studied. Optimal choice of probing signals is investigated for each of these techniques and their channel estimation performances are analyzed. In the case of multiple LS channel estimates, the best linear unbiased estimation (BLUE) scheme for their linear combining is developed and studied.

**Keywords and phrases:** antenna array, downlink channel, channel estimation, training sequence.

## 1. INTRODUCTION

In recent years, transmit beamforming has been a topic of growing interest [1, 2, 3, 4, 5]. The aim of transmit beamforming is to send desired information signals from the base station array to each user and, at the same time, to minimize undesired crosstalks, that is, to satisfy a certain quality of service constraint for each user. This task becomes very complicated if the transmitter does not have precise knowledge of the downlink channel information for each user. Therefore, the beamforming performance severely depends on the quality of channel estimates and an accurate downlink channel estimation plays a key role in transmit beamforming [6, 7, 8, 9]. One of the most popular approaches to downlink channel estimation is channel probing with user feedback [1, 2]. This approach suggests to probe the downlink channel by transmitting training signals from the base station to each user and then to estimate the channel from feedback reports provided by the users.

In this paper, we study the performance of channel probing with feedback in the case of a multisensor base station antenna array and single-sensor users [2]. We develop three channel estimators which offer different tradeoffs in

terms of performance and a priori required knowledge of the channel statistical parameters. First of all, the traditional least squares (LS) method is considered which does not require any knowledge of the channel parameters. Then, a refined version of the LS estimator is proposed (which is referred to as the scaled LS (SLS) estimator). The SLS estimator offers a substantially improved performance relative to the LS method but requires that the trace of the channel covariance matrix and the receiver noise powers be known a priori. Finally, the linear minimum mean square error (LMMSE) channel estimator is developed and studied. The latter technique is able to outperform both the LS and SLS estimators, but it requires the full a priori knowledge of the channel covariance matrix and the receiver noise powers. For each of the aforementioned techniques, the optimal choices of probing signal matrices for downlink channel measurement are studied and channel estimation errors are analyzed. Moreover, in the case of multiple LS channel estimates, an optimal scheme for their linear combining is proposed using the so-called best linear unbiased estimation (BLUE) approach. The effect of such a combining on the performance of downlink channel estimation is investigated.

## 2. BACKGROUND

We assume a base station array of  $L$  sensors and arbitrary geometry and consider the case of flat block fading<sup>1</sup> [2]. In this case, the signal received by the  $i$ th mobile user can be expressed as follows:

$$\mathbf{r}_i(k) = s(k)\mathbf{w}^H\mathbf{h}_i + n_i(k), \quad (1)$$

where  $s(k)$  is the transmitted signal,  $\mathbf{w}$  is the  $L \times 1$  downlink weight vector,  $\mathbf{h}_i$  is the  $L \times 1$  vector which describes an unknown complex vector channel from the array to the  $i$ th user,  $n_i(k)$  is the user zero-mean white noise, and  $(\cdot)^H$  stands for the Hermitian transpose.

In order to measure the vector channel for each user, the method of [2] suggests to use the so-called *probing mode* to transmit  $N \geq L$  training signals  $s(1), \dots, s(N)$  from the base station antenna array using the beamforming weight vectors  $\mathbf{w}_1, \dots, \mathbf{w}_N$ , respectively. The received signals at the  $i$ th mobile can be expressed as follows:

$$\mathbf{r}_i = \mathbf{W}^H\mathbf{h}_i + \mathbf{n}_i, \quad (2)$$

where

$$\mathbf{W} = [s^*(1)\mathbf{w}_1, s^*(2)\mathbf{w}_2, \dots, s^*(N)\mathbf{w}_N] \quad (3)$$

is the  $L \times N$  *probing matrix*,  $\mathbf{r}_i = [r_i(1), \dots, r_i(N)]^T$ ,  $\mathbf{n}_i = [n_i(1), \dots, n_i(N)]^T$ , and  $(\cdot)^*$  and  $(\cdot)^T$  stand for the complex conjugate and the transpose, respectively.

Then, each receiver (mobile user) employs the *information mode* to feed the data received in the probing mode back to the base station where these data are used to estimate the downlink vector channels. Alternatively (to decrease the amount of feedback bits), channel estimation can be done directly at each receiver. In the latter case, receivers feed the corresponding channel estimates back to the base station.

## 3. LS CHANNEL ESTIMATION

Knowing  $\mathbf{r}_i$ , the downlink vector channel between the base station and the  $i$ th user can be estimated using the least LS approach as [2]

$$\hat{\mathbf{h}}_i = \mathbf{W}^\dagger \mathbf{r}_i, \quad (4)$$

where  $\mathbf{W}^\dagger = (\mathbf{W}\mathbf{W}^H)^{-1}\mathbf{W}$  is the pseudoinverse of  $\mathbf{W}^H$ . Assume that the transmitted power in the probing mode is constrained as:

$$\|\mathbf{W}\|_F^2 = P, \quad (5)$$

where  $P$  is a given power constant. We find  $\mathbf{W}$  which minimizes the channel estimation error for the  $i$ th user subject to the transmitted power constraint (5). This is equivalent to

the optimization problem

$$\min_{\mathbf{W}} E\{\|\mathbf{h}_i - \hat{\mathbf{h}}_i\|^2\} \quad \text{subject to } \|\mathbf{W}\|_F^2 = P, \quad (6)$$

where  $E\{\cdot\}$  is the statistical expectation. Using (2) and (4), we have that  $\mathbf{h}_i - \hat{\mathbf{h}}_i = \mathbf{W}^\dagger \mathbf{n}_i$  and, hence, the objective function in (6) can be rewritten as

$$\begin{aligned} J_{LS} &= E\{\|\mathbf{h}_i - \hat{\mathbf{h}}_i\|^2\} = E\{\|\mathbf{W}^\dagger \mathbf{n}_i\|^2\} \\ &= \sigma_i^2 \text{tr}\{\mathbf{W}^\dagger \mathbf{W}^{\dagger H}\} = \sigma_i^2 \text{tr}\{(\mathbf{W}\mathbf{W}^H)^{-1}\}, \end{aligned} \quad (7)$$

where we use the fact that  $E\{\mathbf{n}_i \mathbf{n}_i^H\} = \sigma_i^2 \mathbf{I}$ . Here,  $\sigma_i^2$  is the noise power of the  $i$ th user,  $\mathbf{I}$  is the identity matrix, and  $\text{tr}\{\cdot\}$  denotes the trace of a matrix.

Using (7), the optimization problem (6) can be equivalently written in the following form:

$$\min_{\mathbf{W}} \text{tr}\{(\mathbf{W}\mathbf{W}^H)^{-1}\} \quad \text{subject to } \text{tr}\{\mathbf{W}\mathbf{W}^H\} = P. \quad (8)$$

We obtain the solution to this problem using the Lagrange multiplier method, that is, via minimizing the function

$$L(\mathbf{W}, \lambda) = \text{tr}\{(\mathbf{W}\mathbf{W}^H)^{-1}\} + \lambda(\text{tr}\{\mathbf{W}\mathbf{W}^H\} - P), \quad (9)$$

where  $\lambda$  is the Lagrange multiplier.

To compute  $\partial L(\mathbf{W}, \lambda)/\partial \mathbf{W}^H$ , the following lemma will be useful.

**Lemma 1.** *If a square matrix  $\mathbf{F}$  is a function of another square matrix  $\mathbf{G} = \mathbf{A} + \mathbf{B}\mathbf{X} + \mathbf{X}^H\mathbf{C}\mathbf{X}$ , then the following chain rule is valid:*

$$\frac{\partial \text{tr}\{\mathbf{F}\}}{\partial \mathbf{X}} = \frac{\partial \text{tr}\{\mathbf{G}\}}{\partial \mathbf{X}} \frac{\partial \text{tr}\{\mathbf{F}\}}{\partial \mathbf{G}}, \quad (10)$$

where  $\mathbf{A}$ ,  $\mathbf{B}$ , and  $\mathbf{C}$  are constant matrices and the dimensions of all the matrices in (10) are assumed to match.

*Proof.* See Appendix A.  $\square$

Furthermore, the following expressions for the matrix derivatives of traces will be used [10]:

$$\frac{\partial \text{tr}\{\mathbf{X}\mathbf{X}^H\}}{\partial \mathbf{X}^H} = \mathbf{X}^T, \quad (11)$$

$$\frac{\partial \text{tr}\{\mathbf{X}^{-1}\}}{\partial \mathbf{X}} = -\mathbf{X}^{-2T}. \quad (12)$$

Inserting  $\mathbf{F} = (\mathbf{W}\mathbf{W}^H)^{-1}$ ,  $\mathbf{X} = \mathbf{W}^H$ , and  $\mathbf{G} = \mathbf{W}\mathbf{W}^H$  into (10), we have

$$\frac{\partial \text{tr}\{(\mathbf{W}\mathbf{W}^H)^{-1}\}}{\partial \mathbf{W}^H} = \frac{\partial \text{tr}\{\mathbf{W}\mathbf{W}^H\}}{\partial \mathbf{W}^H} \frac{\partial \text{tr}\{(\mathbf{W}\mathbf{W}^H)^{-1}\}}{\partial \mathbf{W}\mathbf{W}^H}. \quad (13)$$

<sup>1</sup>The flat fading assumption is valid for narrowband communication systems.

Applying (11) and (12) to (13), we can transform the latter equation as

$$\frac{\partial \text{tr} \{(\mathbf{W}\mathbf{W}^H)^{-1}\}}{\partial \mathbf{W}^H} = -\mathbf{W}^T (\mathbf{W}\mathbf{W}^H)^{-2T}. \quad (14)$$

Using (14) and applying (11) to compute  $\partial \text{tr}\{\mathbf{W}\mathbf{W}^H\}/\partial \mathbf{W}^H$  in the second term of (9), we have that

$$\frac{\partial L(\mathbf{W}, \lambda)}{\partial \mathbf{W}^H} = \mathbf{W}^T (\lambda \mathbf{I} - (\mathbf{W}\mathbf{W}^H)^{-2T}). \quad (15)$$

Setting (15) to zero, we obtain that any probing matrix is the optimal one if it satisfies the equation

$$(\mathbf{W}\mathbf{W}^H)^{-2} = \lambda \mathbf{I}. \quad (16)$$

Since  $\mathbf{W}\mathbf{W}^H$  is Hermitian and positive definite, we can write its eigendecomposition as

$$\mathbf{W}\mathbf{W}^H = \mathbf{Q}\mathbf{\Gamma}\mathbf{Q}^H, \quad (17)$$

where  $\mathbf{\Gamma}$  is a diagonal matrix with positive eigenvalues on the main diagonal. Using the positiveness of the eigenvalues of  $\mathbf{W}\mathbf{W}^H$  and taking into account that  $\mathbf{Q}$  is a unitary matrix ( $\mathbf{Q}^H\mathbf{Q} = \mathbf{Q}\mathbf{Q}^H = \mathbf{I}$ ), we have from (16) that

$$\mathbf{Q}\mathbf{\Gamma}^{-2}\mathbf{Q}^H = \lambda \mathbf{I} \quad (18)$$

and, therefore,

$$\mathbf{\Gamma} = \frac{1}{\sqrt{\lambda}} \mathbf{I}. \quad (19)$$

Inserting (19) into (17) and using the identity  $\mathbf{Q}\mathbf{Q}^H = \mathbf{I}$ , we obtain that  $\mathbf{W}$  is an optimal probing matrix if

$$\mathbf{W}\mathbf{W}^H = \frac{1}{\sqrt{\lambda}} \mathbf{I}. \quad (20)$$

Using the power constraint (5), we can rewrite (20) as

$$\mathbf{W}\mathbf{W}^H = \frac{P}{L} \mathbf{I}. \quad (21)$$

Therefore, any probing matrix with *orthogonal rows* of the same norm  $\sqrt{P/L}$  is an optimal one. Note that the similar fact has been earlier discovered from different points of view in [11, 12]. With such optimal probing, the LS estimator reduces to the simple decorrelator-type estimator.

According to (21), there is an infinite number of choices of the optimal probing matrix. It is also worth noting that each optimal choice of  $\mathbf{W}$  is *user independent*. Therefore, any probing matrix that satisfies (21) is optimal for *all users*.

It should be stressed that additional constraints on  $\mathbf{W}$  may be dictated by particular implementation issues. For example, the peak transmitted power per antenna may be limited. In this case, we have to distribute the transmitted power uniformly over the antennas and, therefore, the additional constraint is that all the elements of the optimal probing matrix should have the same magnitude. To satisfy this con-

straint, a properly normalized submatrix of the DFT matrix can be used, that is,

$$\mathbf{W} = \sqrt{\frac{P}{NL}} \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & W_N & W_N^2 & \cdots & W_N^{N-1} \\ 1 & W_N^2 & W_N^4 & \cdots & W_N^{2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & W_N^{L-1} & W_N^{2(L-1)} & \cdots & W_N^{(L-1)(N-1)} \end{bmatrix}, \quad (22)$$

where  $W_N = e^{j2\pi/N}$ .

Using (21) along with (7), we obtain that the minimum downlink channel mean-square estimation error becomes

$$\min_{\mathbf{W}} J_{LS} = \frac{\sigma_i^2 L^2}{P}. \quad (23)$$

We stress that the error in (23) is proportional to the square of the number of transmit antennas and this may lead to a certain restriction of the dimension of the transmit array. However, one can compensate for this effect by increasing the total transmitted power in the probing mode.

Another interesting observation is that the error in (23) is independent of the channel realization  $\mathbf{h}_i$  and the array geometry.

#### 4. SCALED LS CHANNEL ESTIMATION

Obviously, the LS estimate (4) does not necessarily minimize the channel estimation error because its objective is to minimize the signal estimation error rather than the channel estimation error. Therefore, it may be possible to use an additional scaling factor  $\gamma$  to further reduce this error. Using this idea, applying (2) and (4), and dropping the user index  $i$  for the sake of simplicity, we can write the channel estimation error in the following form:

$$\begin{aligned} E\{\|\mathbf{h} - \gamma \hat{\mathbf{h}}_{LS}\|^2\} &= \text{tr} \{E\{(\mathbf{h} - \gamma \hat{\mathbf{h}}_{LS})(\mathbf{h} - \gamma \hat{\mathbf{h}}_{LS})^H\}\} \\ &= (1 - \gamma)^2 \text{tr} \{\mathbf{R}_h\} + \gamma^2 \sigma^2 \text{tr} \{(\mathbf{W}\mathbf{W}^H)^{-1}\} \\ &= (J_{LS} + \text{tr} \{\mathbf{R}_h\}) \left( \gamma - \frac{\text{tr} \{\mathbf{R}_h\}}{J_{LS} + \text{tr} \{\mathbf{R}_h\}} \right)^2 \\ &\quad + \frac{J_{LS} \text{tr} \{\mathbf{R}_h\}}{J_{LS} + \text{tr} \{\mathbf{R}_h\}}, \end{aligned} \quad (24)$$

where  $\hat{\mathbf{h}}_{LS}$  is the LS channel estimate of (4),  $\mathbf{R}_h = E\{\mathbf{h}\mathbf{h}^H\}$  is the channel correlation matrix, and  $J_{LS}$  is given by (7). Clearly, (24) is minimized with

$$\gamma = \frac{\text{tr} \{\mathbf{R}_h\}}{J_{LS} + \text{tr} \{\mathbf{R}_h\}} \quad (25)$$

and the minimum of (24) with respect to  $\gamma$  is given by

$$\begin{aligned} J_{SLS} &= \min_{\gamma} E\{\|\mathbf{h} - \gamma \hat{\mathbf{h}}_{LS}\|^2\} \\ &= \frac{J_{LS} \text{tr} \{\mathbf{R}_h\}}{J_{LS} + \text{tr} \{\mathbf{R}_h\}} < J_{LS}. \end{aligned} \quad (26)$$

Note that the optimal  $\gamma$  in (25) is a function of the trace of the channel correlation matrix  $\mathbf{R}_h$  and the noise variance  $\sigma^2$ . Therefore, these values have to be known (or preliminary estimated) when using the SLS approach. In practice, the estimate of  $\text{tr}\{\mathbf{R}_h\}$ ,

$$\widehat{\text{tr}\{\mathbf{R}_h\}} = \hat{\mathbf{h}}_{LS}^H \hat{\mathbf{h}}_{LS}, \quad (27)$$

can be used in (25) in lieu of  $\text{tr}\{\mathbf{R}_h\}$ . Assuming that the values of  $\text{tr}\{\mathbf{R}_h\}$  and  $\sigma^2$  are given in advance, defining the SLS channel estimate as

$$\hat{\mathbf{h}}_{SLS} = \gamma \hat{\mathbf{h}}_{LS}, \quad (28)$$

and using (4) and (25), we have

$$\hat{\mathbf{h}}_{SLS} = \frac{\text{tr}\{\mathbf{R}_h\}}{\sigma^2 \text{tr}\{(\mathbf{W}\mathbf{W}^H)^{-1}\} + \text{tr}\{\mathbf{R}_h\}} \mathbf{W}^\dagger \mathbf{r}. \quad (29)$$

The optimal probing matrix for channel estimation using the SLS method can be found by means of solving the following optimization problem:

$$\min_{\mathbf{W}} J_{SLS} \quad \text{subject to } \text{tr}\{\mathbf{W}\mathbf{W}^H\} = P. \quad (30)$$

Since  $\text{tr}\{\mathbf{R}_h\} > 0$ , we see from (26) that  $J_{SLS}$  is a monotonically increasing function of  $J_{LS}$ . Note that  $\text{tr}\{\mathbf{R}_h\}$  is not a function of  $\mathbf{W}$ , and, therefore,  $J_{LS}$  is the only term in (26) which depends on  $\mathbf{W}$ . This means that the optimization problems (6) and (30) are *equivalent*. Therefore, the optimal choice of probing matrix for the SLS channel estimation technique is the same as for the LS approach.

## 5. LMMSE CHANNEL ESTIMATION

In this section, we consider the LMMSE estimator of  $\mathbf{h}$  which is given by [13]

$$\begin{aligned} \hat{\mathbf{h}}_{LMMSE} &= \mathbf{R}_h \mathbf{W} (\mathbf{W}^H \mathbf{R}_h \mathbf{W} + \sigma^2 \mathbf{I})^{-1} \mathbf{r} \\ &= \sigma^{-2} (\mathbf{R}_h^{-1} + \sigma^{-2} \mathbf{W}\mathbf{W}^H)^{-1} \mathbf{W}\mathbf{r}. \end{aligned} \quad (31)$$

The performance of this estimator is characterized by the error  $\mathbf{e} = \mathbf{h} - \hat{\mathbf{h}}_{LMMSE}$  whose mean is zero, and the covariance matrix is given by [13]

$$\mathbf{R}_e = \mathbf{E}\{\mathbf{e}\mathbf{e}^H\} = (\mathbf{R}_h^{-1} + \sigma^{-2} \mathbf{W}\mathbf{W}^H)^{-1}. \quad (32)$$

The LMMSE estimation error is given by

$$J_{LMMSE} = \mathbf{E}\left\{ \|\mathbf{h} - \hat{\mathbf{h}}_{LMMSE}\|^2 \right\} = \text{tr}\{\mathbf{R}_e\}. \quad (33)$$

To minimize (33) subject to the transmitted power constraint  $\text{tr}\{\mathbf{W}\mathbf{W}^H\} = P$ , we can use the Lagrange multiplier method. The problem can be written as follows:

$$L = \text{tr}\left\{ (\mathbf{R}_h^{-1} + \sigma^{-2} \mathbf{W}\mathbf{W}^H)^{-1} \right\} + \lambda \text{tr}\{\mathbf{W}\mathbf{W}^H\}. \quad (34)$$

Using the chain rule (10), it can be readily shown that the optimal probing must satisfy

$$\mathbf{W}\mathbf{W}^H = \frac{\sigma^2}{\sqrt{\lambda}} \mathbf{I} - \sigma^2 \mathbf{R}_h^{-1}. \quad (35)$$

Using the constraint  $\text{tr}\{\mathbf{W}\mathbf{W}^H\} = P$ , (35) can be rewritten as follows:

$$\mathbf{W}\mathbf{W}^H = \frac{1}{L} (P + \sigma^2 \text{tr}\{\mathbf{R}_h^{-1}\}) \mathbf{I} - \sigma^2 \mathbf{R}_h^{-1}. \quad (36)$$

Interestingly, in the high signal-to-noise ratio (SNR) case ( $\sigma^2 \rightarrow 0$ ), (36) transforms to (21). Therefore, in the high SNR domain, the LS, SLS, and LMMSE approaches all have the same condition on optimal probing matrices.

Using (36), we obtain that in the optimal probing case,

$$\mathbf{R}_e = \frac{\sigma^2 L}{P + \sigma^2 \text{tr}\{\mathbf{R}_h^{-1}\}} \mathbf{I}. \quad (37)$$

Therefore,

$$\min_{\mathbf{W}} J_{LMMSE} = \frac{\sigma^2 L^2}{P + \sigma^2 \text{tr}\{\mathbf{R}_h^{-1}\}}. \quad (38)$$

If the channel coefficients are all i.i.d. random variables, we have  $\mathbf{R}_h = \xi^2 \mathbf{I}$ , where  $\xi^2$  can be viewed as the channel attenuation parameter. In this case, (36) transforms to (21) and, therefore, the optimal probing matrix for the LS estimator is also optimal for the LMMSE estimator. Furthermore, in such a situation, the minimum of the channel estimation error is given by

$$\min_{\mathbf{W}} J_{LMMSE} = \frac{\xi^2 \sigma^2 L^2}{\xi^2 P + \sigma^2 L}. \quad (39)$$

Interestingly, if  $\mathbf{R}_h = \xi^2 \mathbf{I}$ , then (26) and (39) are identical which means that the performances of the SLS and LMMSE estimators are similar in this case.

## 6. COMBINING OF MULTIPLE LS CHANNEL ESTIMATES

In Sections 3, 4, and 5, the specific case of a single channel estimate has been considered. In this section, we extend the optimal probing approach to the case of *multiple* LS channel estimates. If there are multiple probing periods available within the channel coherency time, it may be inefficient from the computational and buffering viewpoints to store and process dynamically long amounts of data that are formed by accumulation of multiple received data blocks corresponding to different probing periods. A good alternative here is to obtain a particular channel estimate for each probing period and then to store these estimates dynamically rather than storing the data itself, and to compute the final channel estimate based on a proper combination of such (previously obtained) particular estimates.

Let us have  $K$  estimates  $\hat{\mathbf{h}}_{i,1}, \dots, \hat{\mathbf{h}}_{i,K}$  of the downlink channel corresponding to the  $i$ th user. Let each estimate

be computed using (4) based on some probing matrices  $\mathbf{W}_1, \dots, \mathbf{W}_K$ , respectively. The channel is assumed to be quasistatic (fixed) at the interval of  $K$  probeings, and  $P_k = \|\mathbf{W}_k\|_F^2$  is the transmitted power during the  $k$ th probing.

We aim to improve the performance of downlink channel estimation by combining the estimated values  $\hat{\mathbf{h}}_{i,k}$  for  $k = 1, \dots, K$  in a linear way as follows:

$$\hat{\mathbf{h}}_i = \sum_{k=1}^K \alpha_{i,k} \hat{\mathbf{h}}_{i,k}, \quad (40)$$

where  $\alpha_{i,k}$  are unknown weighting coefficients.

Let us obtain the optimal weighting coefficients by means of minimizing the error in (40). Then, these coefficients can be found by solving the following optimization problem:

$$\min_{\alpha_{i,1}, \dots, \alpha_{i,K}} E \left\{ \left\| \mathbf{h}_i - \sum_{k=1}^K \alpha_{i,k} \hat{\mathbf{h}}_{i,k} \right\|^2 \right\} \quad \text{subject to} \quad \sum_{k=1}^K \alpha_{i,k} = 1, \quad (41)$$

where the constraint in (41) guarantees the unbiasedness of the final channel estimate. This problem corresponds to the so-called BLUE estimator [13].

The solution to (41) is given by the following lemma.

**Lemma 2.** *The optimal weights  $\{\alpha_{i,k}\}_{k=1}^K$  for the  $i$ th user are given by*

$$\alpha_{i,k} = \frac{1}{\text{tr} \left\{ (\mathbf{W}_k \mathbf{W}_k^H)^{-1} \right\} \sum_{n=1}^K 1 / \text{tr} \left\{ (\mathbf{W}_n \mathbf{W}_n^H)^{-1} \right\}}. \quad (42)$$

*Proof.* See Appendix B.  $\square$

It is worth noting that the optimal weighting coefficients  $\alpha_{i,k}$  are user independent (i.e., they are the same for each user).

Choosing optimal orthogonal weighting matrices in each probing, we have

$$\begin{aligned} \text{tr} \left\{ (\mathbf{W}_k \mathbf{W}_k^H)^{-1} \right\} &= \frac{L^2}{P_k}, \\ \sum_{n=1}^K \frac{1}{\text{tr} \left\{ (\mathbf{W}_n \mathbf{W}_n^H)^{-1} \right\}} &= \frac{P_{\text{tot}}}{L^2}, \end{aligned} \quad (43)$$

where

$$P_{\text{tot}} = \sum_{k=1}^K P_k \quad (44)$$

is the total transmitted power during the  $K$  probeings.

Inserting (43) into (42), we obtain that in the case of using optimal orthogonal weighting matrices, the expression for optimal weighting coefficients can be simplified to

$$\alpha_{i,k} = \frac{P_k}{P_{\text{tot}}}. \quad (45)$$

In this case, the downlink channel estimation error is given by

$$\begin{aligned} E \left\{ \|\mathbf{h}_i - \hat{\mathbf{h}}_i\|^2 \right\} &= E \left\{ \left\| \mathbf{h}_i - \sum_{k=1}^K \frac{P_k}{P_{\text{tot}}} \hat{\mathbf{h}}_{i,k} \right\|^2 \right\} \\ &= E \left\{ \left\| \sum_{k=1}^K \frac{P_k}{P_{\text{tot}}} (\mathbf{h}_i - \hat{\mathbf{h}}_{i,k}) \right\|^2 \right\} \\ &= \frac{L^2}{P_{\text{tot}}^2} E \left\{ \left\| \sum_{k=1}^K \mathbf{W}_k \mathbf{n}_{i,k} \right\|^2 \right\} \\ &= \frac{\sigma_i^2 L^2}{P_{\text{tot}}^2} \text{tr} \left\{ \sum_{k=1}^K \mathbf{W}_k \mathbf{W}_k^H \right\} = \frac{\sigma_i^2 L^2}{P_{\text{tot}}}, \end{aligned} \quad (46)$$

where  $\mathbf{n}_{i,k}$  is the zero-mean white noise vector of the  $i$ th user in the  $k$ th probing. When deriving (46), we have used the property  $E \{ \mathbf{n}_{i,k} \mathbf{n}_{i,l}^H \} = \sigma_i^2 \delta_{k,l} \mathbf{I}$ , where  $\delta_{k,l}$  is the Kronecker delta.

We observe that, similar to (23), the error in (46) is independent of the channel realization and the array geometry. Comparing (46) with (23), we see that the optimal linear combining of multiple estimates reduces the estimation error by a factor of  $P_{\text{tot}}/P$ . For example, if each probing has the same power ( $P_k = P$ ,  $K = 1, 2, \dots, K$ ), then  $P_{\text{tot}} = KP$  and the estimation error is reduced by a factor of  $K$ .

## 7. NUMERICAL EXAMPLES

In our simulations, we compare the performance of the LS, SLS, and LMMSE channel estimators in the cases of optimal and nonoptimal choices of probing matrices. Throughout all our simulation examples, we assume that  $N = L$ . The channel coefficients and the receiver noise are assumed to be circular complex Gaussian random variables with the unit variance.

We assume that the base station has a uniform linear array and the downlink channel correlation matrix  $\mathbf{R}_h$  has the following structure:

$$[\mathbf{R}_h]_{n,m} = r^{|n-m|}, \quad 0 \leq r < 1, \quad (47)$$

where  $n$  and  $m$  are the indices of the array sensors. This model of the array covariance is frequently used in the literature; see [14, 15, 16] and references therein.

The elements of  $L \times L$  probing matrices  $\mathbf{W}$  in the case of nonoptimal probing have been drawn independently from a zero-mean complex Gaussian random generator in each simulation run. However, to avoid possible computational inaccuracy of the LS estimator, we have ignored all probing matrices that have resulted in a condition number of  $\mathbf{W}\mathbf{W}^H$  greater than  $10^4$ . Each simulated point is obtained by averaging 5000 independent simulation runs.

In Figure 1, we display the mean of the norm squared of the channel estimation error (MNSE) of the LS channel estimator in the optimal and nonoptimal probing matrix cases. In this figure, MNSEs are plotted versus the probing power

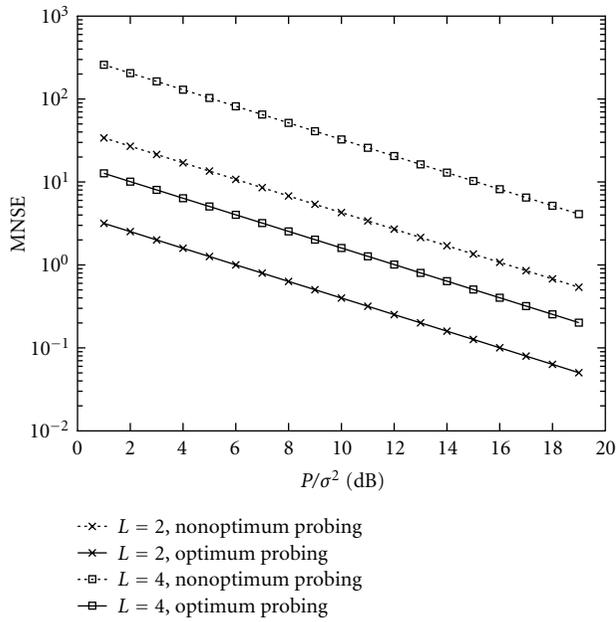


FIGURE 1: MNSEs versus  $P/\sigma^2$  for the LS estimator.

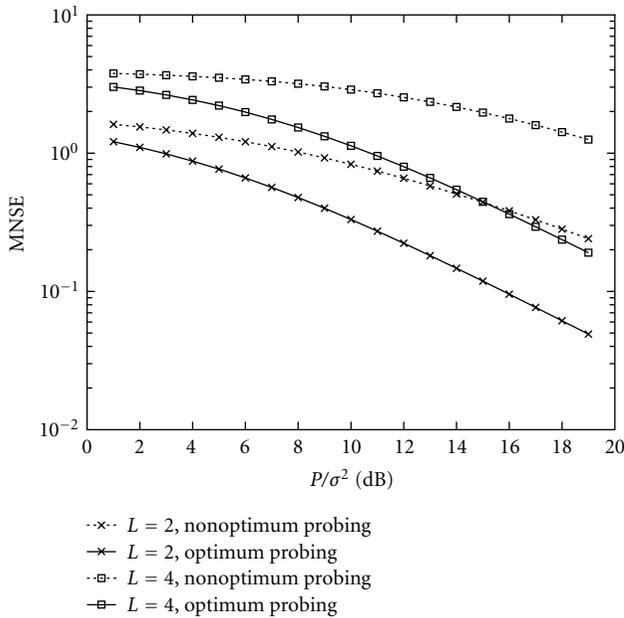


FIGURE 2: MNSEs versus  $P/\sigma^2$  for the SLS estimator.

$P/\sigma^2$ . Note that the performance of the LS estimator is independent of the parameter  $r$ . The parameter  $L$  is varied in this figure.

In Figure 2, the performance of the SLS estimator is tested under the similar conditions. Similar to the LS method, the performance of the LS estimator is independent of the parameter  $r$ .

Figures 3 and 4 display the performance of the LMMSE estimator in the cases of  $r = 0$  and  $r = 0.25$ , respectively.

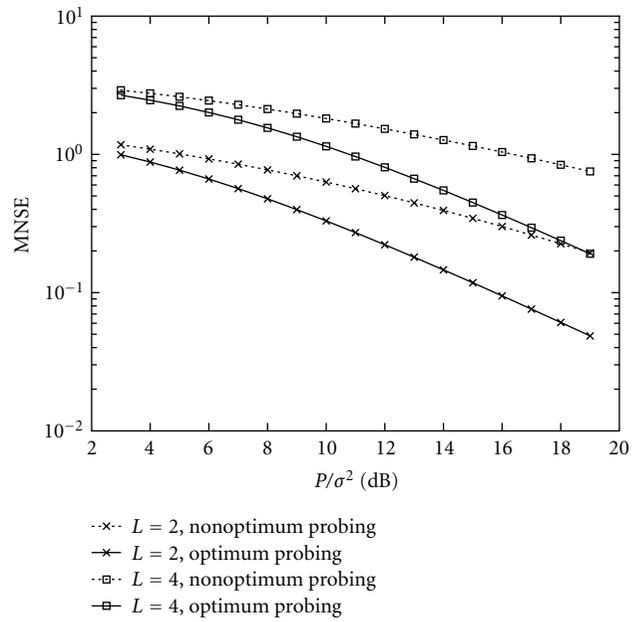


FIGURE 3: MNSEs versus  $P/\sigma^2$  for the LMMSE estimator in the case of uncorrelated channel coefficients ( $r = 0$ ).

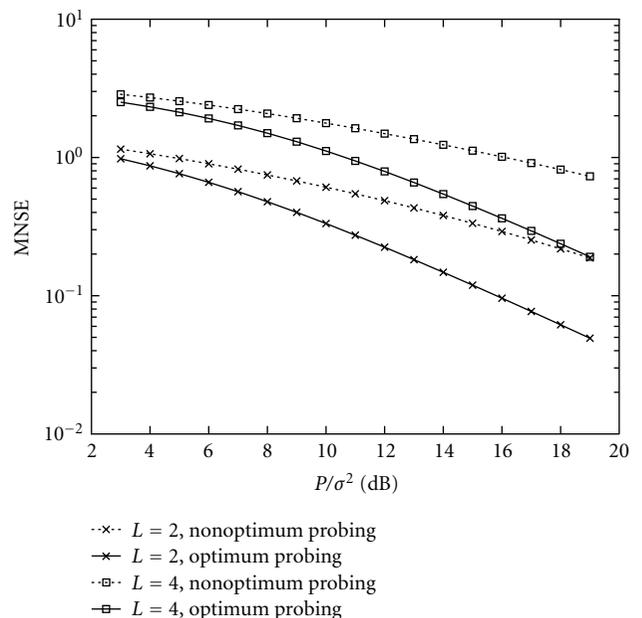
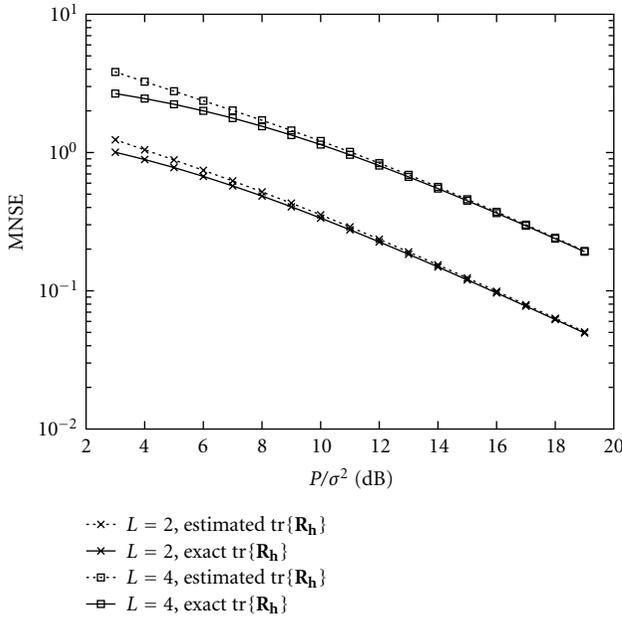


FIGURE 4: MNSEs versus  $P/\sigma^2$  for the LMMSE estimator in the case of correlated channel coefficients ( $r = 0.25$ ).

In both figures, the channel covariance matrix  $\mathbf{R}_h$  is assumed to be known exactly. Other conditions are similar to that of Figures 1 and 2.

From Figures 1, 2, 3, and 4, it can be seen that the optimal probing improves the quality of channel estimation substantially for all methods. Note that this improvement is especially pronounced for large values of  $P/\sigma^2$  if the SLS or LMMSE method is used. Comparing Figures 3 and 4, we also see that these figures give nearly the same results. This means

FIGURE 5: MNSEs versus  $P/\sigma^2$  for the SLS estimator.

that moderate correlation of the channel coefficients does not affect the LMMSE approach.

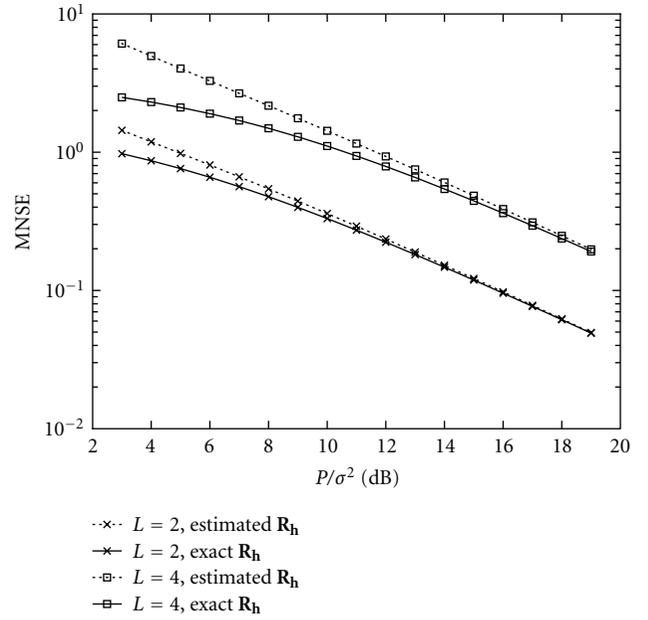
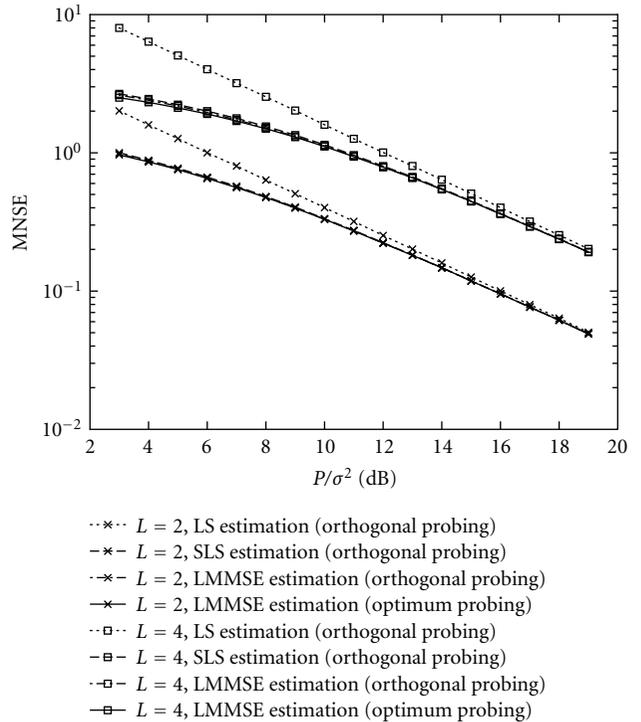
As it has been mentioned before, the SLS channel estimator requires the knowledge of  $\text{tr}\{\mathbf{R}_h\}$ . However, note that the LS estimator can be applied to estimate this parameter using (27). In Figure 5, the MNSEs of the SLS estimator with optimal probing are plotted versus  $P/\sigma^2$  in the cases when the exact and estimated values of  $\text{tr}\{\mathbf{R}_h\}$  are used. In the latter case, the LS method is applied to obtain the estimate of  $\text{tr}\{\mathbf{R}_h\}$  which is then inserted into the SLS estimator. All other conditions are similar to that of the previous figures.

In the LMMSE method, the full knowledge of the channel correlation matrix  $\mathbf{R}_h$  is required either at the base station or at the mobile station to estimate the channel (depending on where the channel estimation is done). Also, the base station transmitter has to know this matrix in order to compute the optimal probing matrix. However, one may use the following rank-one estimate of this matrix:

$$\hat{\mathbf{R}}_h = \hat{\mathbf{h}}_{LS} \hat{\mathbf{h}}_{LS}^H. \quad (48)$$

In Figure 6, the performance of the LMMSE channel estimator is tested versus  $P/\sigma^2$  in the cases when  $\mathbf{R}_h$  is known exactly and when its estimate (48) is used. In the latter case, the optimal LS probing is used (note, however, that in the general case, such a probing is nonoptimal for the LMMSE approach). The value of  $L$  is varied in this figure and  $r = 0.25$  is taken.

From Figures 5 and 6, we see that there are only small performance losses caused by using the estimated values of  $\text{tr}\{\mathbf{R}_h\}$  and  $\mathbf{R}_h$  in the SLS and LMMSE estimators, respectively, in lieu of the exact values of  $\text{tr}\{\mathbf{R}_h\}$  and  $\mathbf{R}_h$ . Also, from Figure 6, we see that the optimal LS probing becomes nearly

FIGURE 6: MNSE versus  $P/\sigma^2$  for the LMMSE estimator in the case of correlated channel coefficients ( $r = 0.25$ ).FIGURE 7: Comparison of the performances of the LS, SLS, and LMMSE estimators versus  $P/\sigma^2$  in the case of correlated channel coefficients ( $r = 0.25$ ).

optimal for the LMMSE approach starting from moderate values of SNR. This observation supports theoretical results of Section 5.

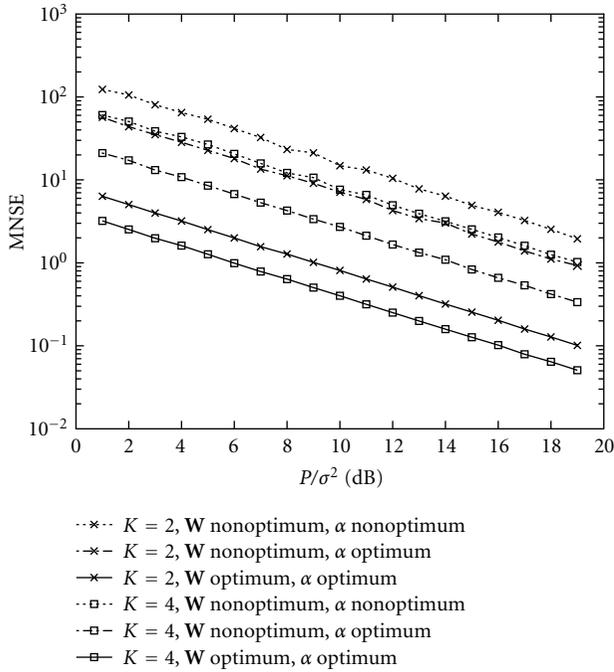


FIGURE 8: MNSE versus  $P/\sigma^2$  for the case of multiple LS channel estimates (the BLUE estimator).

Figure 7 compares the performances of the LS, SLS, and LMMSE estimators versus  $P/\sigma^2$ . In this figure, we assume that  $r = 0.25$ , and two variants of the LMMSE estimator are considered. Both these variants assume that the estimator knows  $\mathbf{R}_h$  exactly, but the first variant uses the optimal probing signal that satisfies (36), while the second one employs the matrix which satisfies (21) and, therefore, is optimal only for LS and SLS estimators and/or for the uncorrelated channel case ( $r = 0$ ). From this figure, we observe that the difference in performance between the first and second variants of the LMMSE estimator is negligible at all the tested values of SNR. Therefore, the LS/SLS probing appears to be suboptimal for the LMMSE estimator.

In the last example, the case of multiple LS channel estimates are assumed. In Figure 8, the parameter  $L = 4$  is chosen and the performance of the BLUE estimator is compared for  $K = 2$  and  $K = 4$ . Three cases are considered in this figure:

- (i) both the probing matrices and the coefficients  $\alpha_{i,k}$  are optimal;
- (ii) the probing matrices are nonoptimal but the coefficients  $\alpha_{i,k}$  are optimal;
- (iii) both the probing matrices and the coefficients  $\alpha_{i,k}$  are nonoptimal.

In the third case, the coefficients  $\alpha_{i,k} = 1/K$  are assumed for all  $i$  and  $k$ .

Figure 8 demonstrates substantial improvements which can be achieved when the BLUE estimator is used in the case

of multiple channel estimates. This figure also shows that the choice of optimal probing matrices and coefficients  $\alpha_{i,k}$  is critical for the estimator performance as nonoptimal choices of one or both of these parameters may cause a severe performance degradation.

## 8. CONCLUSIONS

We have studied the performance of the channel probing method with feedback using a multisensor base station antenna array and single-sensor users. Three channel estimators have been developed which offer different tradeoffs in terms of performance and a priori required knowledge of the channel statistical parameters. First of all, the traditional LS method has been considered. The LS estimator does not require any knowledge of the channel parameters. Then, a new (refined) version of the LS estimator has been proposed. This refined technique is referred to as the SLS estimator. It has been shown to offer a substantially improved channel estimation performance relative to the LS method but requires that the trace of the channel covariance matrix and the receiver noise powers be known a priori. Finally, the LMMSE channel estimator is developed and studied. The latter technique has been shown to potentially outperform both the LS and SLS estimators, but it requires the full a priori knowledge of the channel covariance matrix and the receiver noise powers.

For each of the above mentioned techniques, the optimal choices of probing signal matrices for downlink channel measurement have been studied and channel estimation errors have been analyzed. In the case of multiple LS channel estimates, the BLUE scheme for their linear combining has been developed.

Simulation examples have demonstrated substantial performance improvements that can be achieved using optimal channel probing.

## APPENDICES

### A. PROOF OF LEMMA 1

First of all, we prove the chain rule for the particular case when  $\mathbf{G} = \mathbf{B}\mathbf{X}$ . Writing this equation elementwise, we have  $g_{i,l} = \sum_k b_{i,k}x_{k,l}$  and, therefore,

$$\frac{\partial g_{i,l}}{\partial x_{m,n}} = \delta_{i,n}b_{i,m}, \quad (\text{A.1})$$

where the Wirtinger derivatives for complex variables are used,  $\delta_{i,n}$  is the Kronecker delta, and

$$b_{i,m} = \frac{\partial \text{tr}\{\mathbf{G}\}}{\partial x_{m,i}}. \quad (\text{A.2})$$

Since  $\mathbf{F}$  is a function of  $\mathbf{G}$ , then  $\text{tr}\{\mathbf{F}\}$  can be a function of all elements of  $\mathbf{G}$ . Thus, applying the extended derivative chain

rule ([17, page 99]) and (A.1)-(A.2), we have

$$\begin{aligned} \left[ \frac{\partial \text{tr}\{\mathbf{F}\}}{\partial \mathbf{X}} \right]_{m,n} &= \frac{\partial \text{tr}\{\mathbf{F}\}}{\partial x_{m,n}} = \sum_i \sum_l \frac{\partial \text{tr}\{\mathbf{F}\}}{\partial g_{i,l}} \frac{\partial g_{i,l}}{\partial x_{m,n}} \\ &= \sum_i \frac{\partial \text{tr}\{\mathbf{F}\}}{\partial g_{i,n}} b_{i,m} = \sum_i \frac{\partial \text{tr}\{\mathbf{G}\}}{\partial x_{m,i}} \frac{\partial \text{tr}\{\mathbf{F}\}}{\partial g_{i,n}} \quad (\text{A.3}) \\ &= \left[ \frac{\partial \text{tr}\{\mathbf{G}\}}{\partial \mathbf{X}} \frac{\partial \text{tr}\{\mathbf{F}\}}{\partial \mathbf{G}} \right]_{m,n} \end{aligned}$$

and the proof for the particular case  $\mathbf{G} = \mathbf{B}\mathbf{X}$  is completed.

To extend the proof to the general case  $\mathbf{G} = \mathbf{A} + \mathbf{B}\mathbf{X} + \mathbf{X}^H\mathbf{C}\mathbf{X}$ , we notice that this equation can be rewritten as  $\mathbf{G} = \mathbf{A} + (\mathbf{B} + \mathbf{X}^H\mathbf{C})\mathbf{X}$  and, therefore, the established result for the particular case  $\mathbf{G} = \mathbf{B}\mathbf{X}$  can be applied taking into account that the matrix  $\mathbf{A}$  is constant and that  $\partial \text{tr}\{\mathbf{B} + \mathbf{X}^H\mathbf{C}\}/\partial \mathbf{X} = 0$ . In other words, replacing the matrix  $\mathbf{B}$  by the matrix  $\mathbf{B} + \mathbf{X}^H\mathbf{C}$ , we straightforwardly extend our proof to the general case.

## B. PROOF OF LEMMA 2

To solve (41), we insert (4) into the objective function of (41) and, using (2), rewrite it as

$$\begin{aligned} &E \left\{ \text{tr} \left\{ \left( \sum_{m=1}^K \alpha_{i,m} \mathbf{W}_m^\dagger \mathbf{n}_{i,m} \right) \left( \sum_{n=1}^K \alpha_{i,n} \mathbf{W}_n^\dagger \mathbf{n}_{i,n} \right)^H \right\} \right\} \\ &= \text{tr} \left\{ \left( \sum_{m=1}^K \sum_{n=1}^K \alpha_{i,m} \alpha_{i,n}^* \mathbf{W}_m^\dagger \mathbf{W}_n^H E \{ \mathbf{n}_{i,m} \mathbf{n}_{i,n}^H \} \right) \right\} \quad (\text{B.1}) \\ &= \text{tr} \left\{ \sigma_i^2 \sum_{n=1}^K |\alpha_{i,n}|^2 (\mathbf{W}_n \mathbf{W}_n^H)^{-1} \right\}, \end{aligned}$$

where  $\mathbf{n}_{i,m}$  is the noise vector of the  $i$ th user during the  $m$ th probing interval and the property  $E \{ \mathbf{n}_{i,m} \mathbf{n}_{i,n}^H \} = \delta_{mn} \mathbf{I}$  is used.

To minimize (B.1) subject to the constraint  $\sum_{k=1}^K \alpha_{i,k} = 1$ , we have to find the minimum of the Lagrangian

$$L(\boldsymbol{\alpha}, \lambda) = \text{tr} \left\{ \sigma_i^2 \sum_{k=1}^K |\alpha_{i,k}|^2 (\mathbf{W}_k \mathbf{W}_k^H)^{-1} \right\} - \lambda \left( \sum_{k=1}^K \alpha_{i,k} - 1 \right), \quad (\text{B.2})$$

where the vector  $\boldsymbol{\alpha}$  captures all the coefficients  $\alpha_{i,k}$ .

The gradient of (B.2) is given by

$$\frac{\partial L(\boldsymbol{\alpha}, \lambda)}{\partial \alpha_{i,k}} = 2\sigma_i^2 \alpha_{i,k} \text{tr} \left\{ (\mathbf{W}_k \mathbf{W}_k^H)^{-1} \right\} - \lambda. \quad (\text{B.3})$$

Setting it to zero, we have

$$\alpha_{i,k} = \frac{\lambda}{2\sigma_i^2 \text{tr} \left\{ (\mathbf{W}_k \mathbf{W}_k^H)^{-1} \right\}}. \quad (\text{B.4})$$

Noting that  $\sum_{k=1}^K \alpha_{i,k} = 1$ , we obtain (42).

## ACKNOWLEDGMENTS

A. B. Gershman is on research leave from the Department of Electrical and Computer Engineering, McMaster University, Canada. This work was supported in part by the Wolfgang Paul Award Program, the Alexander von Humboldt Foundation; Premier's Research Excellence Award Program, the Ministry of Energy, Science and Technology (MEST) of Ontario; Natural Sciences and Engineering Research Council (NSERC), Canada; and Communications and Information Technology Ontario (CITO).

## REFERENCES

- [1] D. Gerlach and A. Paulraj, "Adaptive transmitting antenna methods for multipath environments," in *Proc. IEEE Global Telecommunications Conference (GLOBECOM '94)*, vol. 1, pp. 425–429, San Francisco, Calif, USA, November–December 1994.
- [2] D. Gerlach and A. Paulraj, "Adaptive transmitting antenna arrays with feedback," *IEEE Signal Processing Letters*, vol. 1, no. 10, pp. 150–152, 1994.
- [3] D. Gerlach and A. Paulraj, "Base station transmitting antenna arrays for multipath environments," *Signal Processing*, vol. 54, no. 1, pp. 59–73, 1996.
- [4] F. Rashid-Farrokhi, K. J. R. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1437–1450, 1998.
- [5] C. Farsakh and J. A. Nossek, "Spatial covariance based downlink beamforming in an SDMA mobile radio system," *IEEE Trans. Communications*, vol. 46, no. 11, pp. 1497–1506, 1998.
- [6] S. Bhashyam, A. Sabharwal, and U. Mitra, "Channel estimation for multirate DS-CDMA systems," in *Proc. 34th Asilomar Conference on Signals, Systems & Computers*, vol. 2, pp. 960–964, Pacific Grove, Calif, USA, October–November 2000.
- [7] A. Arredondo, K. R. Dandekar, and G. Xu, "Vector channel modeling and prediction for the improvement of downlink received power," *IEEE Trans. Communications*, vol. 50, no. 7, pp. 1121–1129, 2002.
- [8] Y.-C. Liang and F. P. S. Chin, "Downlink channel covariance matrix (DCCM) estimation and its applications in wireless DS-CDMA systems," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 2, pp. 222–232, 2001.
- [9] M. Biguesh, S. Shahbazpanahi, and A. B. Gershman, "Robust power adjustment for transmit beamforming in cellular communication systems," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP '03)*, vol. 5, pp. 105–108, Hong Kong, April 2003.
- [10] H. Lütkepohl, *Handbook of Matrices*, John Wiley & Sons, New York, NY, USA, 1996.
- [11] S. D. Silverstein, "Application of orthogonal codes to the calibration of active phased array antennas for communication satellites," *IEEE Transactions on Signal Processing*, vol. 45, no. 1, pp. 206–218, 1997.
- [12] T. L. Marzetta, "BLAST training: estimating channel characteristics for high capacity space-time wireless," in *Proc. 37th Annual Allerton Conference on Communication, Control, and Computing*, pp. 958–966, Monticello, Ill, USA, September 1999.
- [13] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1993.
- [14] A. B. Gershman, C. F. Mecklenbräuker, and J. F. Böhme, "Matrix fitting approach to direction of arrival estimation with

imperfect spatial coherence of wavefronts," *IEEE Transactions on Signal Processing*, vol. 45, no. 7, pp. 1894–1899, 1997.

- [15] A. B. Gershman, P. Stoica, M. Pesavento, and E. G. Larsson, "Stochastic Cramer-Rao bound for direction estimation in unknown noise fields," *IEEE Proceedings-Radar, Sonar and Navigation*, vol. 149, no. 1, pp. 2–8, 2002.
- [16] J. Ringelstein, A. B. Gershman, and J. F. Böhme, "Direction finding in random inhomogeneous media in the presence of multiplicative noise," *IEEE Signal Processing Letters*, vol. 7, no. 10, pp. 269–272, 2000.
- [17] G. A. Korn and T. M. Korn, *Mathematical Handbook for Scientists and Engineers*, Dover Publications, Mineola, NY, USA, 2000.

top 20 researchers aged 40 or under. He is an Associate Editor for the *IEEE Transactions on Signal Processing* and *EURASIP Journal on Wireless Communications and Networking*, as well as a Member of the SAM Technical Committee of the IEEE SP Society.

**Mehrzad Biguesh** was born in Shiraz, Iran. He received the B.S. degree in electronics engineering from Shiraz University in 1991, and the M.S. and Ph.D. degrees in telecommunications (with honors) from Sharif University of Technology (SUT), Tehran, Iran, in 1994 and 2000, respectively. During his Ph.D. studies, he was appointed at Guilan university and SUT as a Lecturer. From November 1998 to August 1999,



he was with INRS-Telecommunications, University of Quebec, Canada, as a Doctoral Trainee. From 1999 to 2001, he held an appointment at the Iran Telecom Research Center (ITRC), Teheran. From 2000 to 2002, he was with the Electronics Research Center at SUT and held several short-time appointments in the telecommunications industry. Since March 2002, he has been a Postdoctoral Fellow in the Department of Communication Systems, University of Duisburg-Essen, Duisburg, Germany. His research interests include array signal processing, MIMO systems, wireless communications, and radar systems.

**Alex B. Gershman** received his Diploma and Ph.D. degrees in radiophysics from the Nizhny Novgorod University, Russia, in 1984 and 1990, respectively. From 1984 to 1989, he was with the Radiotechnical and Radiophysical Institutes, Nizhny Novgorod. From 1989 to 1997, he was with the Institute of Applied Physics, Nizhny Novgorod. From 1997 to 1999, he was a Research Associate at the Department of Electrical Engineering,



Ruhr University, Bochum, Germany. In 1999, he joined the Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ontario, Canada where he is now a Professor. He also held visiting positions at the Swiss Federal Institute of Technology, Lausanne, Ruhr University, Bochum, and Gerhard-Mercator University, Duisburg. His main research interests are in statistical and array signal processing, adaptive beamforming, MIMO systems and space-time coding, multiuser communications, and parameter estimation. He has published over 220 technical papers in these areas. Dr. Gershman was a recipient of the 1993 URSI Young Scientist Award, the 1994 Outstanding Young Scientist Presidential Fellowship (Russia), the 1994 Swiss Academy of Engineering Science Fellowship, and the 1995–1996 Alexander von Humboldt Fellowship (Germany). He received the 2000 Premiers Research Excellence Award, Ontario, Canada, and the 2001 Wolfgang Paul Award, Alexander von Humboldt Foundation, Germany. He was also a recipient of the 2002 Young Explorers Prize from the Canadian Institute for Advanced Research (CIAR), which has honored Canada's

# Diversity Properties of Multiantenna Small Handheld Terminals

**Wim A. Th. Kotterman**

*Department of Communication Technology (KOM), Aalborg University, 9220 Aalborg Ø, Denmark  
Email: wim@kom.aau.dk*

**Gert F. Pedersen**

*Department of Communication Technology (KOM), Aalborg University, 9220 Aalborg Ø, Denmark  
Email: gfp@kom.aau.dk*

**Kim Olesen**

*Department of Communication Technology (KOM), Aalborg University, 9220 Aalborg Ø, Denmark  
Email: ko@kom.aau.dk*

*Received 23 June 2003; Revised 26 February 2004*

Experimental data are presented on the viability of multiple antennas on small mobile handsets, based on extensive measurement campaigns at 2.14 GHz with multiple base stations, indoors, from outdoor to indoor, and outdoors. The results show medium to low correlation values between antenna branch signals despite small antenna separations down to  $0.16\lambda$ . Amplitude distributions are mainly Rayleigh-like, but for early and late components steeper than Rayleigh. Test users handling the measurement handset caused larger delay spread, increased the variability of the channel, and induced rather large mean branch power differences of up to 10 dB. Positioning of multiple antennas on small terminals should therefore be done with care. The indoor channels were essentially flat fading within 7 MHz bandwidth ( $-6$  dB); the outdoor-to-indoor and outdoor channels, measured with 10 MHz bandwidth, were not. For outdoor-to-indoor and outdoor channels, we found that different taps in the same impulse response are uncorrelated.

**Keywords and phrases:** mobile radio channel, small multiantenna devices, measurement analysis, branch correlation, Doppler spectrum, user influence.

## 1. INTRODUCTION

Research on smart antennas or smart algorithms seem to have focused on base stations (BSs) and fixed terminals with relatively little research being devoted to the benefits of multiple antennas on small mobile terminals. A reason for this surely must be the still frequently expressed opinion that a separation between antennas of at least half a wavelength is needed to get branch correlation coefficients under a threshold of 0.7 needed for exploiting the diversity potential. In this context, one often quotes Jakes [1], but he considered amplitude correlation coefficients for early narrowband mobile systems, whereas for GSM-like systems, it was shown that at least for some forms of diversity, such a threshold does not exist. Diversity gain then increases continuously with decreasing correlation [2]. Moreover, Vaughan and Andersen [3] showed that in the ideal case, the antenna patterns are orthogonal with respect to the incoming wave field, which theoretically can be achieved even at zero separation for par-

ticular environments. This of course implies that the achievable diversity gain depends on both the antenna design and the specific propagation environment. In this respect, spatial separation is merely a factor in decorrelation between antenna signals as are polarisation properties. Experimental confirmation has been documented from the early 1990's onwards [4, 5, 6, 7]. Please note that the overriding importance of handset antennas being small, while efficient and wideband, leaves little room for engineering radiation patterns.

In the framework of a project on smart antennas for small handsets at Aalborg University (AAU), three measurement campaigns were organised in different propagation environments with and without users, as users have a strong influence on the reception by handheld terminals [8, 9]. During these campaigns, we used our proprietary measurement system [10] with our "optical" handset without conducting cables, but using signal transport by optic fibre instead [11]. This paper reports on the findings, with some emphasis placed on the three classical quantities

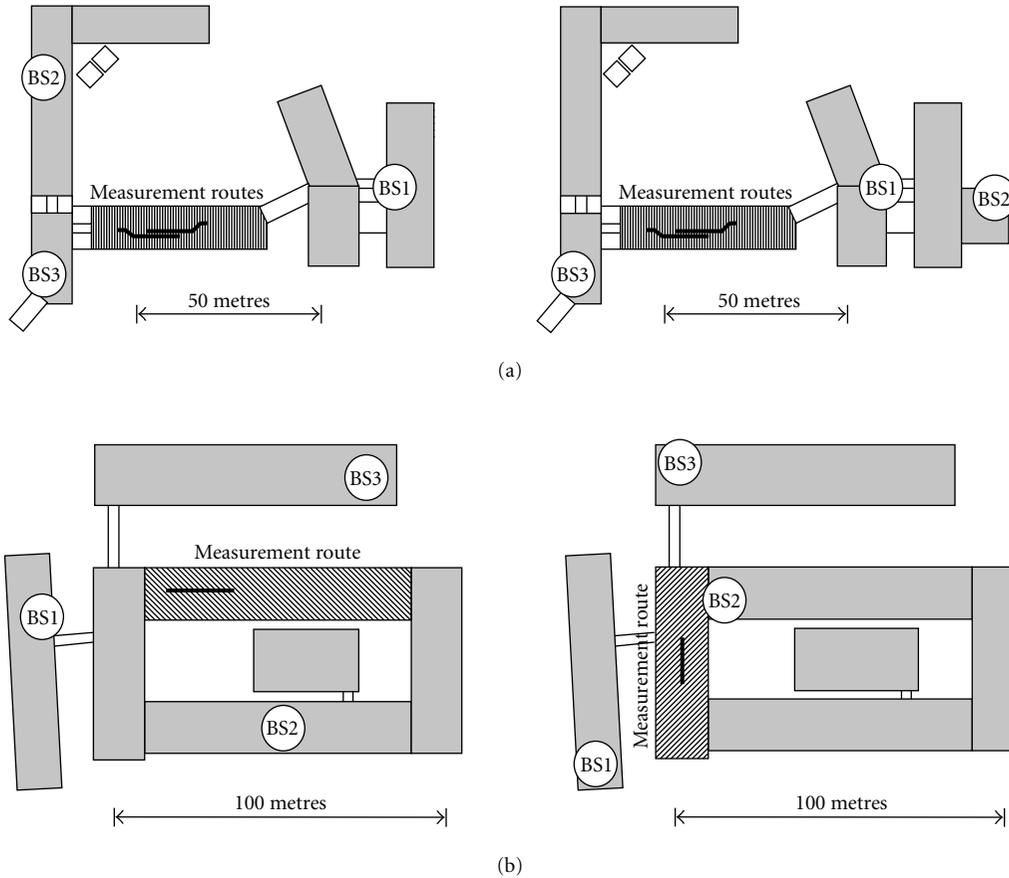


FIGURE 1: Measurement situations and BS configurations for the indoor campaign: (a) star configuration for the new building (left) and inline configuration for the new building (right); (b) star configuration for the old building (left) and inline configuration for the old building (right). The new buildings’ first route is to the right, walked from left to right, while the second route is to the left, walked from right to left.

determining diversity gain: branch correlation coefficients, amplitude distributions, and (mean) branch power differences [12]. The structure of this paper is as follows: first, the measurement setup is discussed with the chosen scenarios, the use of test users, and the equipment. Next, the processing of the data is described, followed by results and discussion. Conclusions form the last section.

## 2. MEASUREMENT SETUP

The measurement campaigns should provide realistic data for channel models to be used for research into smart antennas for small handsets. Therefore, the data should be gathered in a way that reflects typical use of handheld devices and typical handheld devices themselves, including size, antenna types, and locations of major components like display, keypad, and antennas. This means measuring in different cellular scenarios, with users handling the terminal in different ways. Some aspects of the choices made for the campaigns will be treated in the next sections.

### 2.1. Cellular scenarios

Three cellular scenarios were chosen: indoor, outdoor-to-indoor, and outdoor.

For the indoor campaigns, we selected two different buildings as the type of construction determines the propagation regime. One is the university building in downtown Aalborg as example of the early twentieth-century building style: heavy walls with single-sheet windowpanes, favouring penetration through the windows with only limited guiding in corridors. As for the second building, a modern office building at the campus was selected, having a reinforced concrete structure with plasterboard partitioning and metal-coated windows as in Figure 1. Little penetration from outside should be expected as most signals are guided inside.

For the outdoor-to-indoor campaigns, the old university building was selected. In this campaign, the link budget was improved, which allowed placing BSs at more distant and more obstructed locations as in Figure 2. Free-in-air measurements were added too, with the handset on a pole without the user as a form of reference.

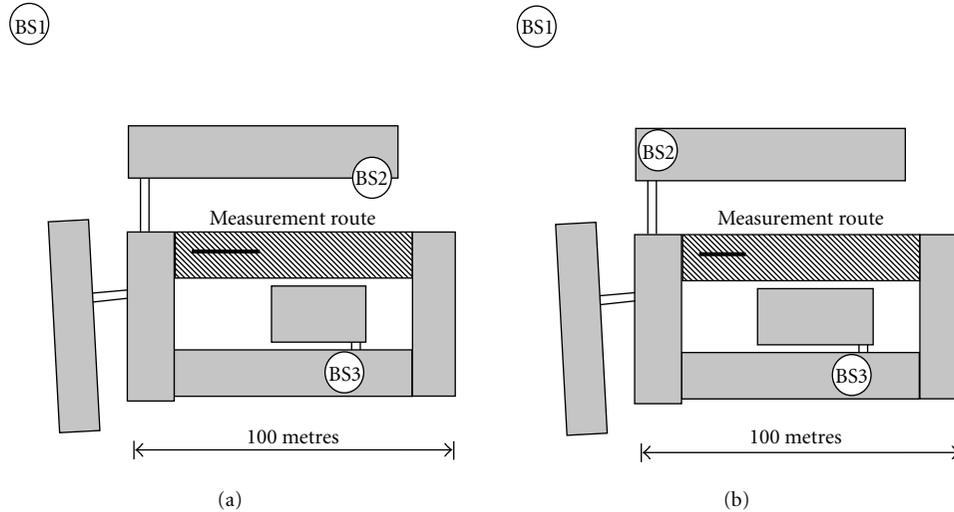


FIGURE 2: Measurement situations and BS configurations for the outdoor-to-indoor campaign: (a) star configuration for the old building and (b) inline configuration for the old building.

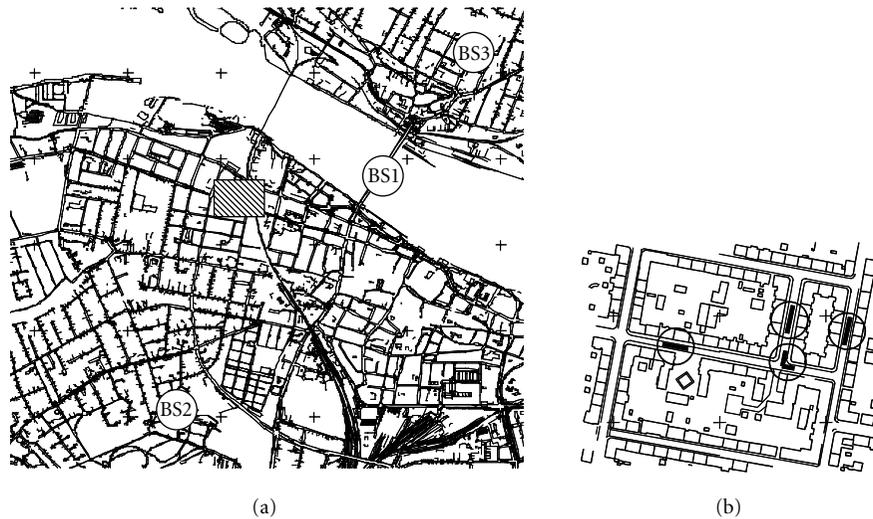


FIGURE 3: Measurement situation for the outdoor campaign: (a) BSs in the centre of Aalborg with the measurement area shaded ( $2.75 \times 2.5 \text{ km}^2$ ) and (b) enlarged outdoor measurement area with the four measurement trajectories encircled ( $245 \times 180 \text{ m}^2$ ).

For the outdoor campaigns, the measurements were aimed at medium size cells in a European downtown area with propagation conditions and path lengths clearly different from the two other environments. Path lengths ranged from 1 to 2 km as in Figure 3. The area in Aalborg with the smallest ratio of street width to rooftop height was chosen and for link budget reasons, relatively high BSs were employed. Here only results will be shown for the handset tied to a torso phantom in a trailer behind the measurement van due to low signal-to-noise ratio (SNR), with the test users inside the van.

## 2.2. Interference situations

The choice for measuring multiple BSs simultaneously is based on the fact that interference certainly is one of the

major aspects of cellular network operation. In CDMA systems, intercell interference may be less important than in TDMA systems, but in CDMA, the best candidate for soft handover/macrodiversity is most likely the strongest interferer.

Two different BS configurations have been chosen, a “star” BS configuration and an “inline” configuration. Figure 1 gives an example of the two configurations for the indoor measurements, and Figure 2 for the outdoor-to-indoor measurements. Of the outdoor measurements, represented in Figure 3, only the inline data is used.

The star configuration imitates the conditions at the edge of a cell, with three BSs surrounding the mobile station at comparable distances. This maximises interference levels but the correlation between interfering signals and the desired

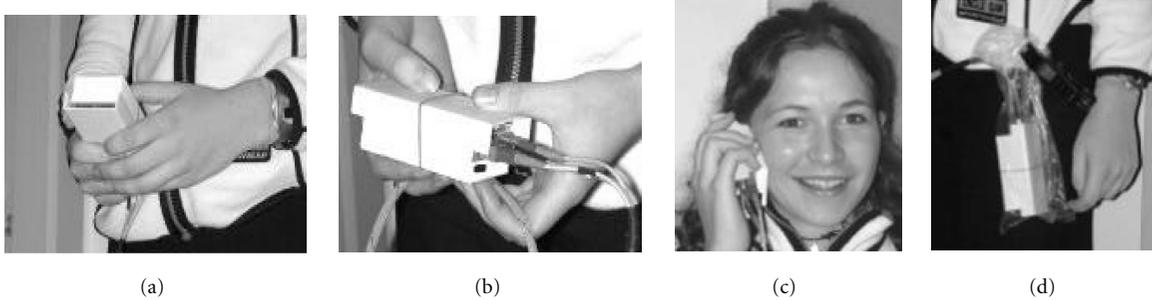


FIGURE 4: The four ways of handling the handset measurement by a test person. (a) portrait, (b) landscape, (c) at the ear, and (d) at the hip.

signal is most likely low. In the inline configuration, the levels of interference differ but the correlation between the interferers and the serving BS could be higher than in the star configuration, especially under waveguiding conditions.

### 2.3. Use of test persons

The use of a number of test persons is based on the experience that the user has a major impact on handset performance [9], for instance, due to body-induced losses (hand, head), due to orientation of the handset, due to specific movements of the user, and so forth. Therefore, we aimed at having at least ten users run the prescribed test route. The users were also asked to hold the handsets in a number of different ways, at the ear and in the hand in two different ways. For the outdoor-to-indoor campaign, enough link budget was available to incorporate placement at the hip too. Figure 4 gives an impression of the various positions.

The position of the terminal in the hand, called “portrait,” imitates the present use of a phone when updating the calendar or SMS directory. The “landscape” mode refers to using the newly developed models with large displays. Carrying the terminal at the hip mainly simulates the idle mode. As mentioned earlier, for the outdoor campaign, only phantom measurement results will be presented.

### 2.4. Equipment

The measurement equipment used was AAU’s proprietary equipment [10], based on a correlating receiver, sampling the received signal in  $I$  and  $Q$  on baseband signal, with correlation of the 511-chips long  $m$ -sequence in postprocessing. Simultaneous sounding of BSs was achieved in the code domain. Throughout the campaigns, we used our optical handset, in two versions, that truly represents a small receiving device without the radiation pattern disturbing effects of conducting cables [11]. The antennas employed on these handsets were chosen to reflect practical implementations and designed to occupy as small volume as possible for the required bandwidth. This leads to monopole-like antennas that act as matching or coupling to the terminal casing that then acts as the main radiator. In this way, very small antennas can show good efficiency and bandwidth compared to the size of the antenna elements because the casing is the main radiator, not the antenna element itself. However, this approach allows

the designer but little control over the antenna radiation patterns and polarisation properties. Also, radiation characteristics are dissimilar for similar antenna elements placed at different positions on the terminal, but this on the other hand contributes significantly to the decorrelation of the antenna signals. We used two different approaches frequently seen with handsets at that time: stubs that are either monopoles or helices, and integrated antennas, in our case planar inverted F antennas (PIFAs), to see whether this would make a difference.

The first version of the handset was used in the indoor campaign with either two monopoles or PIFAs, seen in Figure 5a with monopoles. Chassis dimensions are  $103 \times 48 \times 35 \text{ mm}^3 (h \times w \times t)$ . During the measurements, the wire elements were stabilised with a foam radome. The PIFAs were screwed directly onto the SMA connectors visible at the front. The element size was  $0.1\lambda \times 0.1\lambda \times 0.05\lambda (h \times w \times t)$ , but due to the use of dielectrics, the free in air size was somewhat smaller, making it possible to have a distance between the antennas of only  $0.16\lambda$  centre to centre. The second version was used in the other two campaigns. For outdoor-to-indoor, it was used with both four helices and PIFAs (different from those used indoors) as in Figures 5b and 5c. For the outdoor campaign, the second handset was only equipped with four small (dielectric) PIFAs as in Figure 5d. The first change of antennas was mainly motivated by the mechanic vulnerability of the antenna elements and the wish to have a smooth surface for the second handset. The open PIFA structures used for outdoor-to-indoor proved to be vulnerable too. Consequently, solid dielectric PIFAs were used for the outdoor campaign. Chassis dimensions of the second handset are  $92 \times 51 \times 37 \text{ mm}^3 (h \times w \times t)$ . For protection of the antennas, this handset was used with a plastic lid, visible in Figure 4. The antennas of the second handset have all been measured in the anechoic chamber, spherically, and dual-polarised. Figure 6 shows an example of the radiation patterns for the top two PIFAs antennas in Figure 5c used in the outdoor-to-indoor campaign. For reasons of clarity and due to the limited space, only one plane cut is shown, normal to the faceplate and parallel to the length axis of the handset. Although the patterns are quite similar to each other in both polarisations, the achievable decorrelations are substantial as will be shown in Section 4. Those decorrelations result

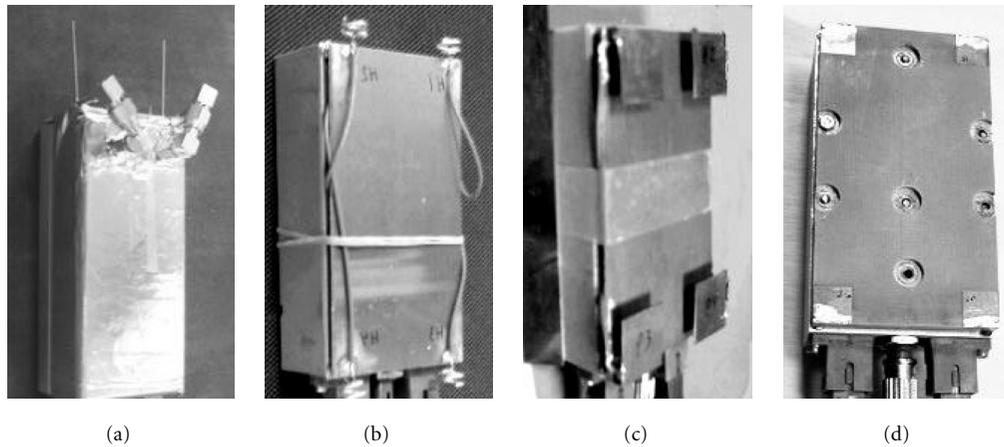


FIGURE 5: Antenna placements on handsets: (a) first handset with monopoles for indoor; (b and c) second handset with helices and PIFAs for outdoor-to-indoor; and (d) second handset with dielectric PIFAs for outdoor. All handsets are shown without radome or protective cover.

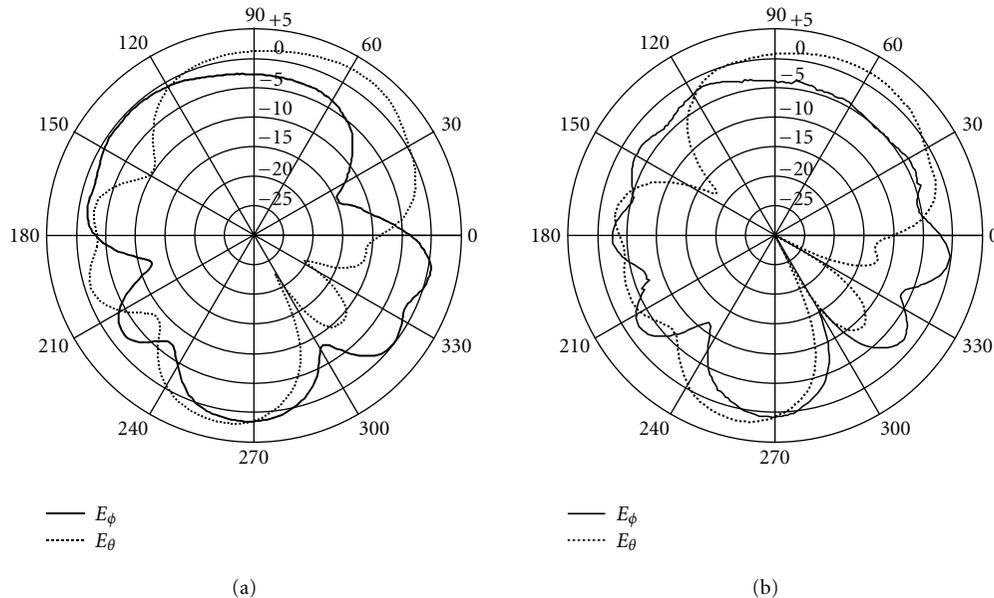


FIGURE 6: Measured radiation patterns (copolar and cross-polar) for two of the PIFAs in the outdoor-to-indoor campaign (Figure 5c) in the plane perpendicular to the faceplate and parallel to the length axis of the terminal: (a) top left antenna and (b) top right antenna. The amplitudes along the radial are in dB.

from the projection of the angular distributions of the incoming wave field onto the radiation patterns (in both polarisations) of the antennas, (see Vaughan [3]). However, seeing the similarities of the antenna patterns, detailed knowledge is required of the angular distributions of the incoming wave field when analysing antenna performance. We did not measure such angular distributions for the environments in these campaigns and considered that to be out of scope for these investigations too. Consequently, we will not expand on the performance of specific antenna types.

Due to a different chip rate, the effective bandwidth was 7 MHz ( $-6$  dB) for indoor campaign and 10 MHz for outdoor-to-indoor and outdoor campaigns. For indoor and outdoor-to-indoor campaigns, the impulse response acquisition was triggered equidistantly in time, and for the outdoor one, equidistantly in distance. All these changes in the equipment resulted from insight gained during the campaigns, spanning more than a year. The main system parameters of the sounding equipment for the different campaigns are summarised in Table 1.

TABLE 1: Main parameters of channel sounding equipment used in the different campaigns.

System parameter	Indoor	Outdoor-to-indoor	Outdoor
PN code length	511	511	511
PN code chip rate (MHz)	3.8325	7.665	7.665
Bandwidth (−6 dB) (MHz)	7	10	10
Baseband sampling (MHz)	15.36	15.36	15.36
IR acquisition rate	30(s <sup>−1</sup> )	25(s <sup>−1</sup> )	1/0.0544(m <sup>−1</sup> )
Carrier frequency (MHz)	2140	2140	2140
Optical handset type	No. 1	No. 2	No. 2
Number of HS antennas	2	4	4
Antenna types (separation)	Monopoles (0.29λ) PIFAs (0.16λ)	Helices (0.21λ/0.51λ) PIFAs (0.21λ/0.51λ)	PIFAs (0.21λ/0.51λ)

### 3. DATA PROCESSING

The purpose of the measurements is to provide data for tapped delay line models. Therefore, the data processing should render suitable tap delays and find the characteristics per tap signal over time or distance measurement. Relations between tap signals should be established too. The characteristics considered are

- (1) amplitude/power distribution per tap;
- (2) cross-correlation between fading patterns of antenna branch signals per tap;
- (3) mean branch power differences;
- (4) Doppler spectrum per tap;
- (5) cross-correlation between fading patterns of BS signals for the same tap and antenna branch;
- (6) cross-correlation between fading patterns of tap signals for the same antenna branch.

The first three are the classical parameters determining diversity gain: the Doppler spectrum determines the evolution of tap signals over time/distance, the cross-correlation between tap signals could influence equalising strategies, and cross-correlation between BSs or interferers influences the gain by both antenna and macrodiversity [13, 14, 15]. The amplitude distribution has also implications on the coverage and the outage performance of system cells; see, for example, [16, 17, 18]. The indoor measurements were essentially flat fading, so only a single tap was used. For the outdoor measurements, no total signal power was computed, so no mean branch power differences were derived.

#### 3.1. General preprocessing

Directly after every campaign, the full set of measurement equipment is taken into a shielded room and calibrated back to back, using attenuators and coaxial cables instead of antennas. The measured data is scaled with the calibration data and correlated with the back-to-back system responses.

##### 3.1.1. Processing specifics for indoor

The indoor responses were essentially single tap. Therefore, the processing consisted of determining the tap delay per BS

and per antenna branch and of separation of slow and fast fading signals from the extracted tap signal. The purpose of using these fading types is to connect to existing modelling schemes in which the fading is modelled as the product of a slow fading term and a fast fading term instead of modelling Nakagami distributions.

The tap excess delay  $\tau_m$  of the single tap was determined per measurement run from the power over all impulse responses  $h(\tau, t_i)$  as  $\tau_m = \operatorname{argmax}_{\tau} \{|h(\tau, t_i)|^2\}$ , with  $t_i \in \{t_1, \dots, t_{512}\}$  the measurement instance. The slow fading power  $p_{\text{slow}}$  was defined as the lowpass filtered output of the received power  $|h(\tau_m, t_i)|^2$  at delay  $\tau_m$ , by convolution with a real-valued Hanning window  $W_H$  of length 48:

$$p_{\text{slow}}(t_i) |h(\tau_m, t_i)|^2 \otimes W_H(t_i) \quad (1)$$

with  $W_H(k) = 0.5 - 0.05 \cdot \cos(2\pi \cdot k/48)$ ;  $k \in \{1, \dots, 48\}$ . The length of the Hanning window was not critical, but the length of 48 rendered fast fading signals that matched Rayleigh distributions quite well, corresponding to 1.6 seconds or a few metres at walking speed. The complex fast fading signal  $h_{\text{fast}}$  is the complex received signal divided by the square root of the slow fading power:

$$h_{\text{fast}}(t_i) = \frac{h(\tau_m, t_i)}{\sqrt{p_{\text{slow}}(t_i)}}. \quad (2)$$

Further processing is done on both the fast fading signal and the (square root of the) slow fading power.

#### 3.1.2. Processing specifics for outdoor and outdoor-to-indoor cases

For the outdoor and outdoor-to-indoor measurement results, tap delays and tap signal characteristics were extracted by using a two-dimensional SAGE algorithm [19]. Based on the rendered estimates, the tap signals (over time for the outdoor-to-indoor case and over distance for the outdoor one) were constructed as described in [20]. The tapped-delay line *structure* is determined by the BS, so it is the same for the different antenna branches and users. This means that each antenna branch and each user signal has the same tap

delays for the response to a particular BS on a particular measurement location, only differing from other branches/users in complex amplitude and Doppler values. For these tap signals, no fast or slow fading signals were extracted. The SAGE estimation process operated on twenty consecutive impulse responses at a time, with the next estimation cycle half overlapping the former. Not always were the estimates available for every tap delay, so on certain measurement intervals, gaps occurred in the constructed tap signals, making the interpretation of slow and fast fading very hard.

### 3.2. Power distributions

Power distributions were derived for indoor data for both the fast and the slow fading power. For outdoor and outdoor-to-indoor data, power distributions were derived for the power in individual tap signals under the constraint that for at least 25% of the tap signal duration, SAGE estimates were available. Data were pooled over measurement runs before determining cumulative distribution functions (CDFs).

### 3.3. Antenna branch correlations

For indoor data, antenna branch correlations for the same BS were determined for both fast and slow fading for the two antenna branches. For outdoor and outdoor-to-indoor data, correlations between each of the six combinations of two out of the four antenna branches were determined for each tap. The correlation per tap was performed over those points where both branches in a combination had (constructed) signal under two constraints: the first being that the tap signal in both branches should have a mean power higher than  $-12$  dB below the highest mean tap power for the respective branch, and the second that the number of common points was larger than 127 (25% of the tap signal duration). The mean power threshold was imposed because of the observed increasing inaccuracy of the SAGE algorithm with decreasing tap powers.

All correlations are complex correlations between variations around the mean. The values given are mean and standard deviation of the magnitude of the correlation coefficients, pooled over users/measurement runs, antenna types, use positions (if applicable), BSs, BS configurations, and antenna branch combinations (for outdoor and outdoor-to-indoor cases).

### 3.4. Mean branch power differences

The mean branch power difference was determined as the difference in the mean power received per branch from a single BS over a single measurement run. For the indoor case, this was the difference in mean values of the slow fading power per antenna branch (fast fading power has mean 1). For the outdoor-to-indoor case, the impulse response powers were integrated over the impulse response duration. For each measurement run, this total received power was averaged per antenna branch. The mean branch power difference per measurement run for each of the six combinations of two out of the four antenna branches was the difference in the respective average total received powers. For the outdoor case,

no mean branch power differences were determined as the computation of the total received power was too sensitive to the influence of noise on the integration interval. As the actual values were often uniformly spread over a large interval symmetric around zero, the mean and standard deviations are given for the absolute values of the differences. The values are pooled over measurement runs, antenna types, use positions, BSs, BS configurations, and antenna branch combinations (for outdoor-to-indoor case).

### 3.5. Doppler spectra

Doppler spectra were made up per measurement run over the full length of each tap signal. For the indoor case, the fast fading signal was used. For plotting purposes, the individual spectra were added powerwise (over measurement runs). The presented results in Table 2 are the average values and the standard deviation of the absolute value of the mean Doppler shift and the Doppler spread determined for each individual spectrum after pooling over users/measurement runs, antenna types, use positions (if applicable), BSs, BS configurations, antenna branches, and taps. Results from tap signals with a mean power lower than  $-12$  dB below the highest mean tap power for the respective branch were discarded. For comparison, the shifts and spreads are normalised with respect to the Nyquist rate of the impulse response acquisition, 15 Hz in the indoor case, 12.5 Hz in the outdoor-to-indoor case, and  $9.2 \text{ m}^{-1}$  in the outdoor case.

### 3.6. Interferer correlation

Interferer correlation was defined as the correlation between two BS signals received on the same antenna branch for a single measurement run. For the indoor case, these (complex) correlations were determined for both antenna branches for all three combinations of two out of three BSs, for both the fast and slow fading signals. For the outdoor-to-indoor case, these correlations have been derived from the total received power. As the power still showed fading in this scenario, the slow fading power was extracted from the total received power by the same smoothing operation as in (1). The fast fading power was defined as the total received power divided by the slow fading power. The interferer correlation was determined as the correlation between either the slow or fast fading powers for all three combinations of two out of three BSs, for all four antenna branches separately. The correlation is of the covariance type. No interferer correlation was determined for the outdoor case. The values given are mean and standard deviation of the absolute value of the correlation coefficients, pooled over measurement runs, antenna types, use positions, BS configurations, antenna branches, and BS combinations.

### 3.7. Intertap correlations

Intertap correlations are the complex correlations between fading patterns of the same tap of the same BS signal on two antenna branches, determined per measurement run. For outdoor and outdoor-to-indoor cases, these correlations were computed for each of the possible combinations (no

TABLE 2: Results of data processing for the different measurement campaigns. Given are the averages of the magnitudes of the considered variable, with standard deviations of the magnitudes in parentheses.

Channel characteristic		Indoor new building	Indoor old building	Outdoor-to-indoor trolley	Outdoor-to-indoor test persons	Outdoor
Amplitude distributions	Fast fading	Rayleigh	Rayleigh	Mainly Rayleigh	Mainly Rayleigh	Mainly Rayleigh
	Slow fading	Lognormal ( $\sigma \sim 3\text{--}7$ dB)	Lognormal ( $\sigma \sim 3\text{--}7$ dB)			
Branch correlations	Fast fading	0.48 (0.26)	0.53 (0.24)	0.33 (0.15)	0.32 (0.16)	0.42 (0.23)
	Slow fading	0.82 (0.16)	0.77 (0.18)			
Mean branch power differences (dB)		2.2 (1.5)	1.8 (1.2)	2.3 (1.6)	4.4 (3.0)	Not determined
Doppler	Mean <sup>‡</sup>	0.25 (0.15)	0.23 (0.17)	0.22 (0.16)	0.41 (0.22)	0.52 (0.29)
	Spread <sup>‡</sup>	0.59 (0.10)	0.67 (0.14)	0.43 (0.07)	0.45 (0.10)	0.34 (0.21)
Interferer correlation	Fast fading	0.14 (0.12)	0.08 (0.05)	0.05 (0.05) <sup>†</sup>	0.05 (0.05) <sup>†</sup>	Not determined
	Slow fading	0.60 (0.23)	0.42 (0.23)	0.31 (0.20) <sup>†</sup>	0.29 (0.20) <sup>†</sup>	
Intertap correlations		N.A.	N.A.	0.19 (0.12)	0.23 (0.14)	0.08 (0.10)

<sup>‡</sup> Values in fractions of Nyquist rate, determined by snapshot repetition rate.

<sup>†</sup> Based on total received power, not on complex signal.

permutations) of two out of all tap signals for a given antenna branch and BS under two constraints: the first being that each tap signal should have a mean power higher than  $-12$  dB below the highest mean tap power for the branch and the second that the tap signals should have at least 127 points in common. For the indoor case with essentially single-tap channels, no intertap correlations were computed. The values given are mean and standard deviation of the magnitude of the correlation coefficients, pooled over measurement runs, antenna types, use positions (if applicable), BSs, BS configurations, antenna branches, and tap combinations.

## 4. RESULTS AND DISCUSSION

The results of the data processing are summarised in Table 2. These results will be discussed in more detail in the following sections.

### 4.1. Power delay profiles

The indoor power delay profiles were the shortest; within the measurement bandwidth, they were factually single tap as mentioned. The tap extraction by the SAGE algorithm rendered two to four taps for the outdoor-to-indoor channels with the largest delay spreads for the outside BS, about 80 nanoseconds. The two other BSs showed delay spreads of around 60 nanoseconds. Differences in use positions or antenna types had no large influence on the spreads or the shape of the power delay profiles. For the outdoors case, widely different results were found from almost single-tap channels to 14-tap channels, with the last number maybe

limited by the fact that the SAGE extraction gave 15 estimates at a time. The effect of test users seen in the outdoor-to-indoor campaign is that users' responses tend to larger delay spread, and so more taps. Also, the variations between responses make it difficult to cluster data from the SAGE algorithm and to arrive at a common tapped-delay representation, especially in cases where the head or body blocks paths to a BS. Therefore, the data for test users of outdoor-to-indoor in Table 2 are for the data terminal portrait use position for BS1 and BS3 only in the star configuration.

### 4.2. Amplitude distributions

The amplitude/power distributions that were found are rather classical. For the indoor campaign, the fast fading showed Rayleigh distributions, while the slow fading power was more or less lognormally distributed. The short measurement runs probably did not allow registering a fully developed slow fading pattern. In the star BS configuration, one BS showed a slow fading pattern with a standard deviation of 6–7 dB, while the other two showed rather low values of 3–4 dB. In the inline configuration, two BSs showed higher standard deviations. For outdoor and outdoor-to-indoor cases, the strongest tap signals were Rayleigh distributed, with the weaker taps before or after strong taps showing some Ricean behaviour; see Figure 7 for a typical example.

Outdoor weak taps could show Ricean distributions with strong dominant components but we are not sure how to interpret this. One explanation is that, for these cases almost always, the very small Doppler spread, and therefore the very slow fading pattern [21], did not allow us to measure a fully developed fading pattern over the measurement

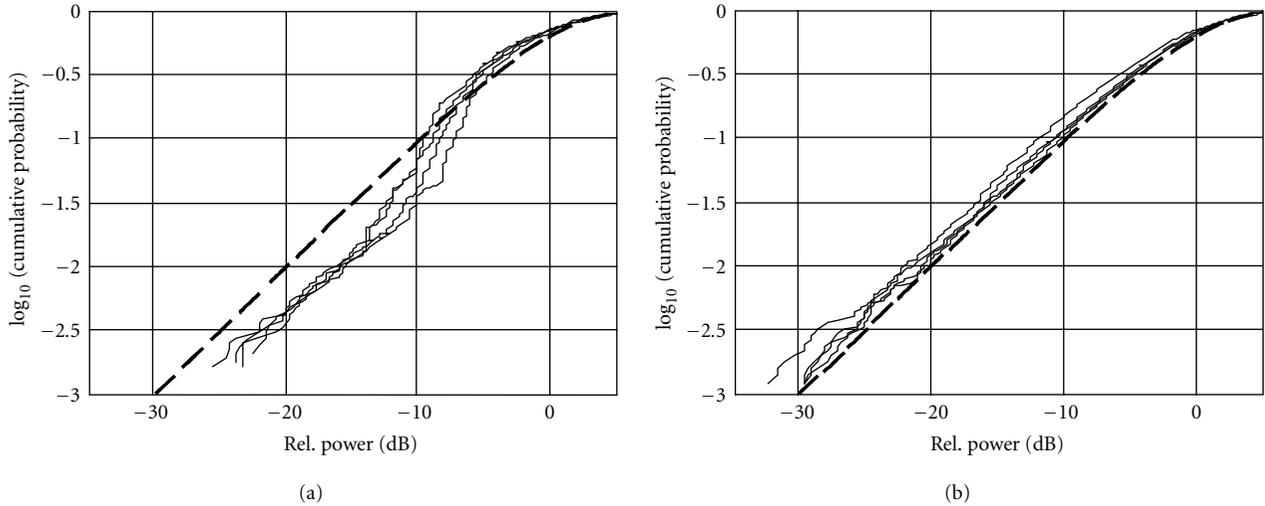


FIGURE 7: Comparison between CDFs of (a) a weak early tap (first tap BS3, average power =  $-9.8$  dB) and (b) a stronger next tap (second tap BS3, excess delay = 73 nanoseconds, average power =  $-1.7$  dB) in the outdoor campaign. Dashed lines indicate CDF of power of Rayleigh distributed process.

TABLE 3: Indoor antenna branch correlations for diverse situations. Given are the averages of the magnitudes of the complex correlation coefficients, standard deviations of the magnitudes in parentheses.

Fading type	Building type	BSs in star		BSs inline	
		Monopole antennas	PIFAs	Monopole antennas	PIFAs
Fast fading	New	0.33(0.16)	0.70(0.19)	0.32(0.17)	0.48(0.26)
	Old	0.39(0.17)	0.70(0.16)	0.33(0.17)	0.72(0.17)
Slow fading	New	0.80(0.16)	0.88(0.15)	0.80(0.16)	0.79(0.18)
	Old	0.72(0.18)	0.80(0.13)	0.74(0.19)	0.81(0.18)

run. Another reason is that the cut-off criterion of  $-30$  dB for the SAGE extraction “cuts the tail” of the distribution of weak components.

### 4.3. Antenna branch signals correlations

As regards the antenna branch correlations, Table 2 shows that differences were found between slow and fast fading. Besides, for the fast fading in the indoor case, apparent differences were found between the antenna types. Table 3 illustrates this fact. The monopole antennas show low correlations for fast fading throughout, of about 0.35 on average. The values for the PIFAs are appreciably higher, on average around 0.75 but at a separation of only  $0.16\lambda$  compared to  $0.29\lambda$  for the monopoles. We have insufficient data to determine what causes this higher cross-correlation: the smaller separation, narrower antenna patterns, better similarity of patterns, a stronger cross-coupling between antennas, or a combination of these.

The slow fading is clearly stronger correlated than the fast fading, with mean values around 0.8. There was little difference between BS configurations, use positions, and antenna types, be it that the PIFAs still had slightly higher correla-

tion values (Table 3). Possible consequences of slow fading correlation coefficients lower than 1 are increased instantaneous branch power differences, as short-term differences in the mean power, even with zero-mean branch power difference, are added to it. As a possible explanation for slow fading not being fully correlated, it has been suggested that it is a coherent propagation effect rather than a result of blocking or shadowing [22, 23].

For outdoors, or for the outdoor-to-indoor case, the values for the antenna branch correlation for the same tap are lower than the values seen indoors, with the lowest values recorded for outdoor-to-indoor, probably due to the larger angular/Doppler spread in this scenario. Outliers for the outdoor scenario were recorded in the middle of the short street, where main contributions to the incoming field showed the smallest Doppler spreads, especially for BS3 (see Section 4.5). In this case, average figures were 0.61 for BS2 and 0.81 for BS3. Line-of-sight connections can be excluded in this street.

In the outdoor-to-indoor case, the helix antennas showed magnitudes of correlation values that on average were 80% of those recorded for the PIFAs, both for free-in-air measurements and with test persons.

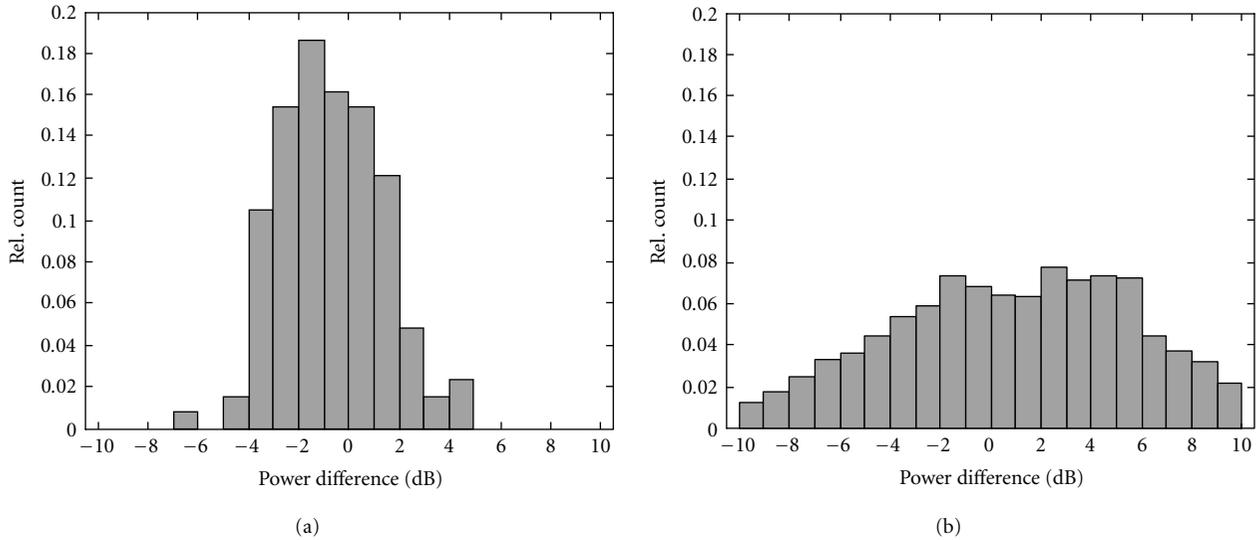


FIGURE 8: Histograms of mean branch power differences for all measurements in the old building for (a) indoor and (b) outdoor-to-indoor with test persons excluding use position “at the hip.”

#### 4.4. Mean branch power differences

Mean branch power differences for the outdoor-to-indoor case are quite large, roughly spanning the interval  $-10$  to  $+10$  dB (Figure 8), confirming results from others [8, 9]. However, during the indoor measurements in the same building, lower values were measured of about half that span. We attribute this to the constructional details of the different handsets used in both campaigns. The first handset used indoors has SMA connectors on the face plate, effectively keeping users’ fingers away from the ground plane of the monopoles, in this way reducing most of the influences on the radiation efficiency. The dielectric PIFAs used indoors are not so sensitive to proximity effects.

Additionally, the distance between the head and antenna elements could be slightly larger in the first handset. During the outdoor-to-indoor campaign, the handset had a fully smooth surface allowing the user more freedom in handling the phone. The types of antennas used in this campaign could also be more sensitive to proximity effects. In Figure 8, use position “at the hip” is excluded as here much lower values were found, showing more or less the same distribution as the indoor values, as did the free-in-air measurements, again a strong indication that the hands and/or fingers of the users are involved.

Note that the instantaneous branch power differences will be larger than the mean value due to the added effect of (uncorrelated) fast fading and partially uncorrelated slow fading on the branches. The values shown here should be regarded as a conservative estimate.

#### 4.5. Doppler spectra

From Table 2, it can be seen that none of the Doppler spectra were symmetric for any of the scenarios. For the indoor

environment, the peak in the spectrum was oriented towards the BS, indicating guiding through the corridors (Figure 9a). The ratio of mean Doppler shift and Doppler spread steadily increases when going from the indoor environment, via outdoor to indoor, to outdoor. For the outdoor environment, this means that signal transport is mainly along street orientation, with low angular/Doppler spread. Figure 9a shows an extreme example for a main tap in the mid of the short street. The guiding effects in the corridors of the indoor environment are less pronounced and the differences between the two buildings are in this respect not as large as anticipated. However, the more “open” old building showed a slightly lower mean Doppler shift with higher Doppler spread due to the larger angular spread of the incoming wave fields.

It is not clear why the ratio of the Doppler shift to the Doppler spread has been increased in the old building, from the indoor campaign to the outdoor-to-indoor one. The BS antennas had narrower antenna beam widths in order to increase the link budget, probably at the expense of the angular spread at the measurement spot. Maybe the receiving antennas were more directional too. It could also be that in the outdoor-to-indoor campaign, we managed better to keep the differences in walking speed between the users small.

The differences between PIFAs and helix antennas are on average small and can often be understood from differences in the radiation patterns. For the outdoor-to-indoor case, a seemingly large difference is shown in Figure 10, where the response of the helices on BS2 has a weak first tap, compared to the PIFAs’ response. However, as the second tap of the helices’ response strongly resembles the PIFAs’ first tap, the most likely explanation is that the helices’ first tap is the obstructed first arrival of BS2 and is not seen at all by the PIFAs. As we did not record absolute delays, we are not able to check this assumption.

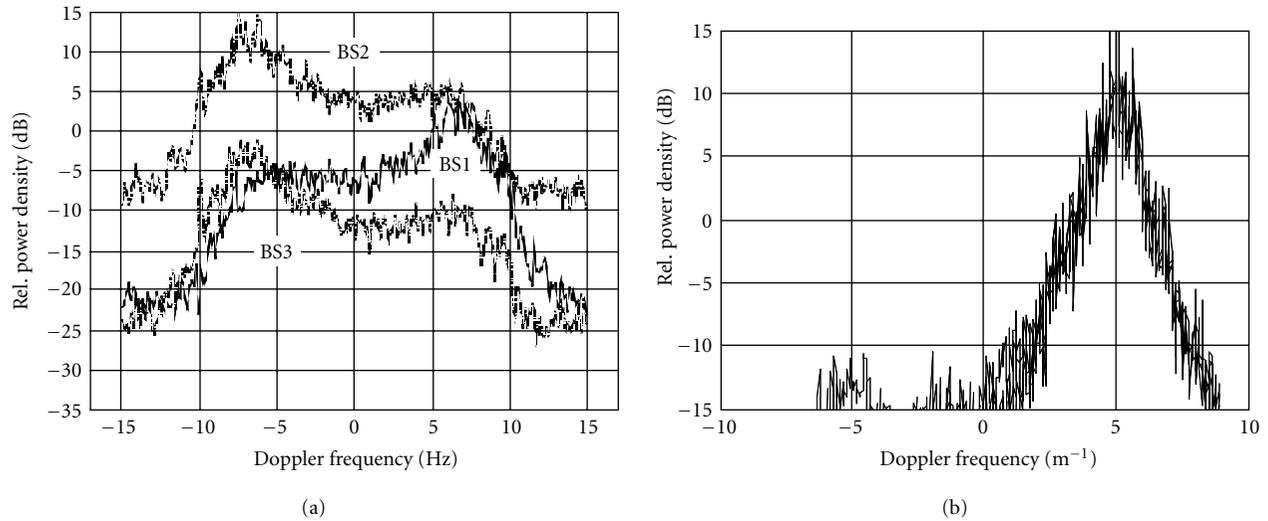


FIGURE 9: (a) Typical Doppler spectra indoor in the new building: first route, BSs in star, monopole antennas, at the ear (curve BS2 offset by +10 dB, curve BS3 offset by -5 dB). (b) Highly directive main tap (tap 2) outdoor for BS3 in the middle of the short street (Figure 3).

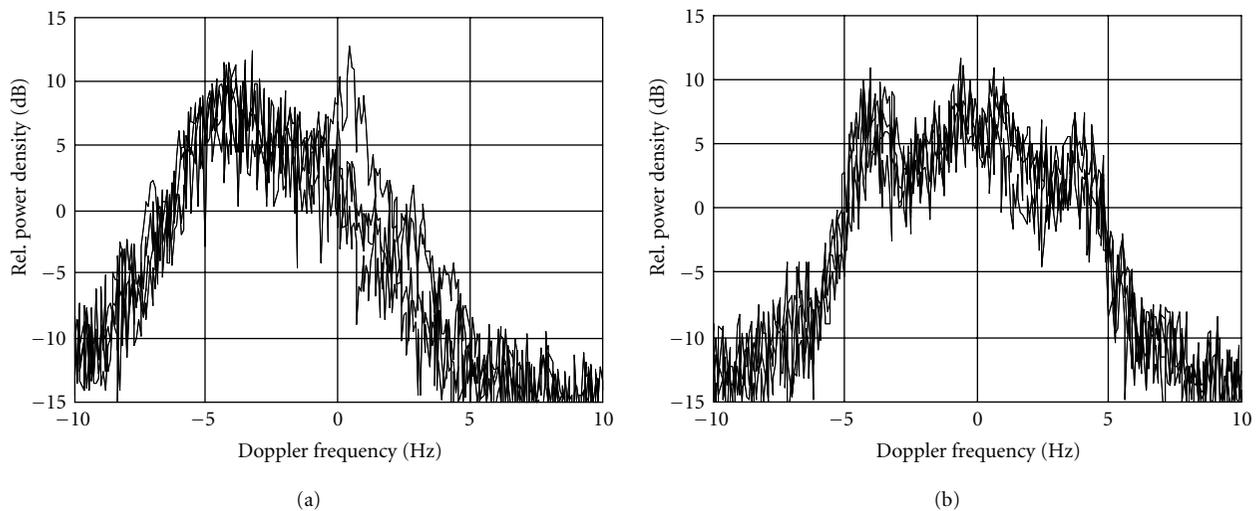


FIGURE 10: Average Doppler spectra outdoor-to-indoor for (a) helix ( $\tau_1 = 0$  nanoseconds,  $p_1 = -10.6$  dB) and (b) PIFAs ( $\tau_1 = 0$  nanoseconds,  $p_1 = 0$  dB): tap 1 of star BS2 configuration, handset free in air, "at the ear."

Test persons' Doppler spectra were generally broader, or smeared out, when compared to those measured free in air, which is reflected in the larger Doppler spread in Table 2. Some influences could be seen in the spectra from shielding by the body or head but the largest influence comes from averaging over ten persons, each walking at a different speed. The effect of different antenna types is comparable to that free in air.

#### 4.6. Interferer correlations

The cross-correlations between BS signals for the same antenna branch (interferer correlation) show higher values for

the slow fading than for the fast fading, just as with the antenna branch correlations. The interferer correlation coefficients are throughout clearly lower than the branch correlations. Fast fading is barely correlated between BSs and slow fading is only in the new building indoors, and is on average moderately correlated. A histogram of all the interferer coefficients measured in the new building indoors reveals a bimodal distribution as in Figure 11a. Note that two real signals are correlated here. The most probable correlation values, around -0.65 and +0.85, are actually not so low. Bimodal distributions in the old building were not found for the star BS configuration (Figure 11b), suggesting more

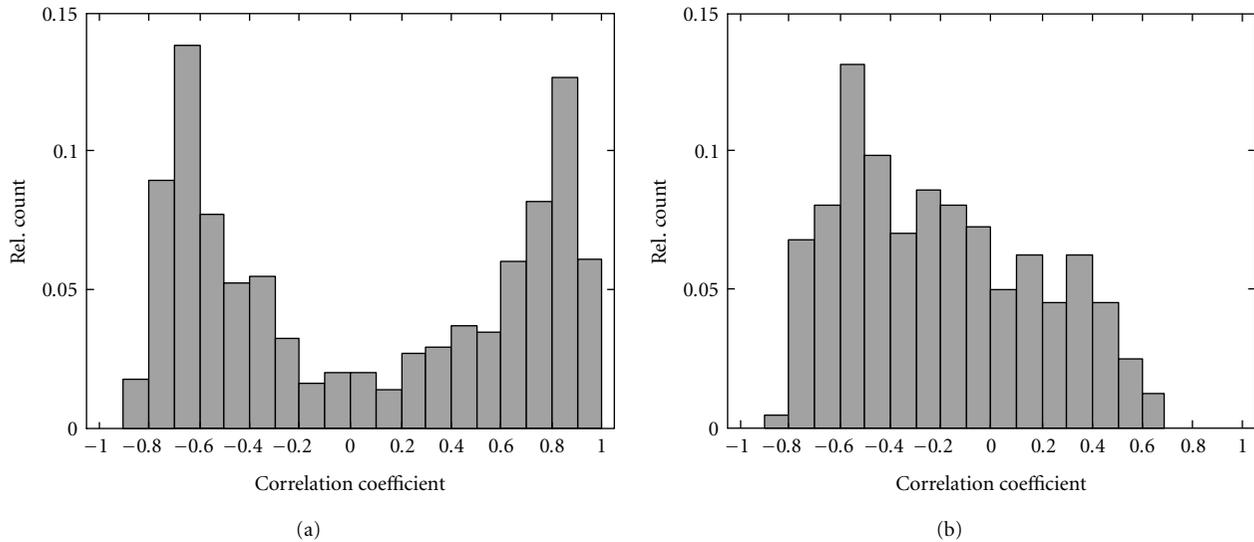


FIGURE 11: (a) Histogram of all (real-valued) slow fading interferer correlation coefficients measured in the new building indoors. (b) Values in star BS configuration in the old building indoors.

similar propagation paths for inline than for star. We have no explanation for the fact that most of the coefficients are negative. The fact that in the new building the distribution of correlation coefficients did not strongly depend on the BS configuration hints on guiding as the main propagation mechanism as opposed to penetration in the old building. Effects of guiding in the new building were even suggested by the fast fading correlation. The combination of BSs that were likely to propagate along the same route to the measurement location had on average three to four times higher correlation coefficients than the other two combinations, irrespective of the antenna type. The maximum average value measured was 0.32 for BS1 and BS2, inline with PIFAs.

#### 4.7. Intertap correlations

Correlations between tap signals, for the same antenna branch and BS, are low, both in the outdoor-to-indoor and the outdoor cases. The highest average value found was 0.65. These values confirm the generally assumed uncorrelated scattering for our measurement environments.

## 5. CONCLUSIONS

We measured a number of characteristics that determine the potential diversity gain of multiple antennas on a small handset such as branch correlations, amplitude/power distributions, Doppler spectra, and mean branch power differences. We measured simultaneously on three base stations for three different typical mobile environments: indoor, outdoor-to-indoor, and outdoor.

The channel characteristics are generally inline with classical assumptions as regards Rayleigh amplitude distribu-

tions and uncorrelated scattering. Doppler spectra, however, are only seldom of classical shape. The branch cross-correlation values are favourably low, especially for the fast fading, down to very small separations between antennas on a mobile handset if the environment allows. In our outdoor scenario, this was not always the case. Interfering base station signals can show moderate to high correlation values, positive or negative, with respect to their slow fading components under guiding conditions as in one of our indoor environments. A handset design optimised for handling by users should take into account the spread in channel characteristics caused by users and especially should seek a solution to the problem of large mean branch power imbalances between the antennas.

## ACKNOWLEDGMENTS

Nokia is acknowledged for financial and technical support of this work. Patrick Eggers supplied the project with valuable background and possible solutions, as did Morten Jeppesen in the first year of the project. Steen Larsen of the E-værksted was responsible for realisation of the measurement hardware and setting up the campaigns. Istvan Kovács and Devendra Prasad are gratefully acknowledged for their realisation of the data acquisition system software and their support during the campaigns. José Klaus Gonzalez implemented the SAGE software.

## REFERENCES

- [1] W. C. Jakes, "New techniques for mobile radio," *Bell Laboratory Rec.*, vol. 48, no. 11, pp. 326–330, 1970.
- [2] P. E. Mogensen and J. Wigard, "On antenna- and frequency

- diversity in GSM related systems (GSM-900, DCS-1800, and PCS1900),” in *Proc. 7th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 3, pp. 1272–1276, Taipei, Taiwan, October 1996.
- [3] R. Vaughan and J. B. Andersen, “Antenna diversity in mobile communications,” *IEEE Trans. Vehicular Technology*, vol. 36, no. 4, pp. 149–172, 1987.
- [4] H. Arai, S. Hosono, and N. Goto, “A flat diversity antenna by disk loaded monopole and notch array,” in *Proc. IEEE Antennas and Propagation Society International Symposium*, vol. 2, pp. 1085–1088, Chicago, Ill, USA, July 1992.
- [5] K. Ogawa and T. Uwano, “A diversity antenna for very small 800-MHz band portable telephones,” *IEEE Trans. Antennas and Propagation*, vol. 42, no. 9, pp. 1342–1345, 1994.
- [6] G. F. Pedersen, S. Widell, and T. Ostervall, “Handheld antenna diversity evaluation in a DCS-1800 small cell,” in *Proc. 8th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 2, pp. 584–588, Helsinki, Finland, September 1997.
- [7] G. F. Pedersen, J. Ø. Nielsen, K. Olesen, and I. Z. Kovacs, “Antenna diversity on a UMTS handheld phone,” in *Proc. 10th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 1, pp. 152–156, Osaka, Japan, September 1999.
- [8] H. Arai, N. Igi, and H. Hanaoka, “Antenna-gain measurement of handheld terminals at 900 MHz,” *IEEE Trans. Vehicular Technology*, vol. 46, no. 3, pp. 537–543, 1997.
- [9] G. F. Pedersen, J. O. Nielsen, K. Olesen, and I. Z. Kovacs, “Measured variation in performance of handheld antennas for a large number of test persons,” in *Proc. 48th IEEE Vehicular Technology Conference*, vol. 1, pp. 505–509, Ottawa, Ontario, Canada, May 1998.
- [10] W. A. Th. Kotterman, D. Prasad, K. Olesen, P. Eggers, I. Z. Kovács, and G. F. Pedersen, “Channel measurement set-up for multi antenna handheld terminal and multiple (interfering) base stations,” in *Proc. International Symposium on Wireless Personal Multimedia Communications*, vol. 1, pp. 153–157, Aalborg, Denmark, September 2001.
- [11] W. A. Th. Kotterman, G. F. Pedersen, K. Olesen, and P. Eggers, “Cable-less measurement set-up for wireless handheld terminals,” in *Proc. 12th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 1, pp. B112–B116, San Diego, Calif, USA, October 2001.
- [12] W. C. Jakes, Ed., *Microwave Mobile Communications*, John Wiley & Sons, New York, NY, USA, 1974.
- [13] L.-C. Wang, G. L. Stüber, and C.-T. Lea, “Effects of Rician fading and branch correlation on a local-mean-based macrodiversity cellular system,” *IEEE Trans. Vehicular Technology*, vol. 48, no. 2, pp. 429–436, 1999.
- [14] J. Zhang and V. Aalo, “Effect of macrodiversity on average-error probabilities in a Rician fading channel with correlated lognormal shadowing,” *IEEE Trans. Communications*, vol. 49, no. 1, pp. 14–18, 2001.
- [15] M.-S. Alouini and M. K. Simon, “Dual diversity over correlated log-normal fading channels,” *IEEE Trans. Communications*, vol. 50, no. 12, pp. 1946–1959, 2002.
- [16] M. D. Austin and G. L. Stüber, “Exact cochannel interference analysis for log-normal shadowed Rician fading channels,” *Electronics Letters*, vol. 30, no. 10, pp. 748–749, 1994.
- [17] M. Zorzi, “Power control and diversity in mobile radio cellular systems in the presence of Ricean fading and log-normal shadowing,” *IEEE Trans. Vehicular Technology*, vol. 45, no. 2, pp. 373–382, 1996.
- [18] M. Pratesi, F. Santucci, F. Graziosi, and M. Ruggieri, “Outage analysis in mobile radio systems with generically correlated log-normal interferers,” *IEEE Trans. Communications*, vol. 48, no. 3, pp. 381–385, 2000.
- [19] B. H. Fleury, M. Tschudin, R. Heddergott, D. Dahlhaus, and K. I. Pedersen, “Channel parameter estimation in mobile radio environments using the SAGE algorithm,” *IEEE JSAC for Wireless Communications Series*, vol. 17, no. 3, pp. 434–450, 1999.
- [20] W. A. Th. Kotterman, “Testing measured outdoor data on correlation properties between scatterers and between antenna branches of a mobile handset,” in *Proc. International Symposium on Wireless Personal Multimedia Communications*, vol. 1, pp. 176–180, Yokosuka, Japan, October 2003.
- [21] R. Vaughan and J. B. Andersen, Eds., *Channels, Propagation and Antennas for Mobile Communications*, IEE, London, UK, 2003.
- [22] S. A. Abbas and A. U. Sheikh, “On understanding the nature of slow fading in LOS microcellular channels,” in *Proc. 47th IEEE Vehicular Technology Conference*, vol. 2, pp. 662–666, Phoenix, Ariz, USA, May 1997.
- [23] J. B. Andersen, “Statistical distributions in mobile communications using multiple scattering,” in *Proc. URSI General Assembly*, Maastricht, Netherlands, August 2002.

**Wim A. Th. Kotterman** graduated from Delft University of Technology, the Netherlands, in physics on acoustic wave field theory. He worked for 13 years at KPN Research Labs, Leidschendam, the Netherlands, as a Scientist in the field of radio network planning and radio system optimisation before becoming an Associate Professor at the Department of Communication Technology, Aalborg University, Denmark, in 2000. His current interests are in channel sounding, the design of channel sounding equipment, and multiantenna handset channel modelling. Wim Kotterman has been a national representative for the Netherlands in the European cooperation projects COST207, COST231, and COST259.



**Gert F. Pedersen** was born in 1965. He received the B.S.E.E. degree in electrical engineering from the College of Technology in Dublin, Ireland, and the M.S.E.E. and Ph.D. degrees from Aalborg University in 1993 and 2003, respectively. He has been employed by Aalborg University, Centre for Personkommunikation, since 1993, where he is currently working as an Associate Professor heading the Antenna group. His research has focused on radio communication for mobile terminals including small antennas, antenna systems, propagation, and biological effects. He has also worked as a Consultant for the development of antennas for mobile terminals including the first internal antenna for mobile phones in 1994 with very low specific absorption rate (SAR), the first internal triple band antenna in 1998 with low SAR and high efficiency, the smallest internal dual band antenna in 2000, and various antenna diversity systems rated as the most efficient in the market. Recently, he has been involved in establishing a method to measure the communication performance for mobile terminals that can be used as a basis for a 3G standard, where measurements also including the antenna will be needed. Further, he is involved in small terminals for 4G including several antennas (MIMO systems) to enhance the data communication.



**Kim Olesen** received his Master's degree in electronic engineering from Aalborg University, Denmark, in 1988. From 1988 to 1993, he was employed in private companies, developing analog radio equipment like maritime radios at very high frequency (VHF) and Nordic mobile telephony (NMT) at ultrahigh frequency (UHF). From 1994 onwards, he has been employed at Aalborg University as Head of the Electronic Workshop in the Department of Communication Technology, Institute of Electronic Systems. His interests are in the design and construction of measurement systems, mainly for research in the field of antennas and propagation. His design activities range from component level to system level, both analog and digital from DC to 6 GHz.



# Multidimensional Rank Reduction Estimator for Parametric MIMO Channel Models

**Marius Pesavento**

*Lehrstuhl für Signaltheorie, Ruhr-Universität Bochum, 44780 Bochum, Germany  
Email: mps@sth.rub.de*

**Christoph F. Mecklenbräuer**

*FTW - Forschungszentrum Telekommunikation Wien, 1220 Wien, Austria  
Email: cfm@ftw.at*

**Johann F. Böhme**

*Lehrstuhl für Signaltheorie, Ruhr-Universität Bochum, 44780 Bochum, Germany  
Email: boehme@sth.rub.de*

*Received 28 May 2003; Revised 25 November 2003*

A novel algebraic method for the simultaneous estimation of MIMO channel parameters from channel sounder measurements is developed. We consider a parametric multipath propagation model with  $P$  discrete paths where each path is characterized by its complex path gain, its directions of arrival and departure, time delay, and Doppler shift. This problem is treated as a special case of the multidimensional harmonic retrieval problem. While the well-known ESPRIT-type algorithms exploit *shift-invariance* between specific partitions of the signal matrix, the rank reduction estimator (RARE) algorithm exploits their internal *Vandermonde structure*. A multidimensional extension of the RARE algorithm is developed, analyzed, and applied to measurement data recorded with the RUSK vector channel sounder in the 2 GHz band.

**Keywords and phrases:** array processing, rank reduction, MIMO, channel sounder, ESPRIT.

## 1. INTRODUCTION

Multidimensional harmonic retrieval problems arise in a large variety of important applications like synthetic aperture radar, image motion estimation, chemistry, and double-directional channel estimation for multiple-input multiple-output (MIMO) communication systems [1]. Also certain signal separation problems can be solved under this framework.

The 4D parameter estimation problem for MIMO channel sounder measurements applies to the following double-directional MIMO channel model in which the signal is assumed to propagate from the transmitter to the receiver over  $P$  discrete propagation paths. Each path ( $p = 1, \dots, P$ ) is characterized by the following parameters: complex path gain  $w_p$ , direction of departure (DOD)  $\theta_p$ , direction of arrival (DOA)  $\phi_p$ , propagation delay  $\tau_p$ , and Doppler shift  $\nu_p$ .

We assume an idealized data acquisition model for MIMO channel sounders. In this model, data consists of simultaneous measurements of the individual complex baseband channel impulse responses between all  $M$  transmit antenna elements and all  $L$  receive antenna elements after ideal lowpass filtering. These are assembled in a 3-way array with

dimensions  $K \times L \times M$ . Such a 3-way array is in the following referred to as a “MIMO snapshot” and consists of  $K$  time samples at sampling period  $T_s$ . A MIMO snapshot is acquired in a duration  $T_a$ . We repeat  $N$  MIMO snapshot measurements consecutively in time and assemble a 4-way array of dimensions  $K \times L \times M \times N$  which we refer to as a “Doppler block.” We assume that all path parameters  $w_p$ ,  $\theta_p$ ,  $\phi_p$ ,  $\tau_p$ ,  $\nu_p$  remain constant within the acquisition time  $NT_a$  of each Doppler block. Individual Doppler blocks are indexed by  $i = 1, \dots, J$ . Between any two Doppler blocks, the complex path gain  $w_p$  may vary arbitrarily while the remaining path parameters are constant for  $p = 1, \dots, P$ . In Section 5.2, we describe the data acquisition with MEDAV’s RUSK-ATM channel sounder [2] (<http://www.medav.de>), (<http://www.channelsounder.de>) which was used for the experiment.

The  $i$ th Doppler block is modelled as

$$x_{k,\ell,m,n}(i) = \sum_{p=1}^P w_p(i) \operatorname{sinc} \left( k - \frac{\tau_p}{T_s} \right) b_p^\ell c_p^m d_p^n + \text{noise}, \quad (1)$$

where

$$\begin{aligned} b_p &= e^{-j(2\pi d_R/\lambda) \cos \phi_p}, & c_p &= e^{-j(2\pi d_T/\lambda) \cos \theta_p}, \\ d_p &= e^{-j(2\pi/N)\nu_p}. \end{aligned} \quad (2)$$

The first index  $k$  represents the time sample, the second index  $\ell$  represents the Rx element number, the third index  $m$  represents the Tx element number, and the fourth index  $n$  represents the Doppler block number. We have assumed uniform linear receive and transmit arrays,  $\lambda$  is the wavelength,  $d_R$  and  $d_T$  denote the elemental spacings of the receive and transmit side, respectively.

After a discrete Fourier transform over the time sample index  $k$ , we obtain

$$\begin{aligned} y_{k,\ell,m,n}(i) &= \sum_{p=1}^P w_p(i) a_p^k b_p^\ell c_p^m d_p^n + \text{noise}, \\ i &= 1, \dots, J, \quad k = 1, \dots, K, \\ \ell &= 1, \dots, L, \quad m = 1, \dots, M, \\ n &= 1, \dots, N, \end{aligned} \quad (3)$$

where

$$\begin{aligned} a_p &= e^{-j(2\pi/K)\tau_p}, & b_p &= e^{-j(2\pi d_R/\lambda) \cos \phi_p}, \\ c_p &= e^{-j(2\pi d_T/\lambda) \cos \theta_p}, & d_p &= e^{-j(2\pi/N)\nu_p}. \end{aligned} \quad (4)$$

We study a joint parameter estimator for the parameters of interest  $\{a_p, b_p, c_p, d_p\}_{p=1}^P$ , where  $|a_p| = |b_p| = |c_p| = |d_p| = 1$ , and  $w_p(i)$  is considered as an unknown nuisance parameter.

Numerous parametric and nonparametric estimation methods have been proposed for the one-dimensional exponential retrieval problem. Only few of these techniques allow a simple extension of the multidimensional case at a reasonable computational load [3]. Simple application of 1D results separately in each dimension is only suboptimal in the sense that it does not exploit the benefits inherent in the multidimensional structure, leading to difficulties in mutually associating the parameter estimates obtained in the various dimensions and over-strict uniqueness conditions [4]. Contrariwise, many parametric high-resolution methods specifically designed for multidimensional frequency estimation require nonlinear and nonconvex optimization so that the computational cost associated with the optimization procedure becomes prohibitively high.

In this paper a novel eigenspace-based estimation method for multidimensional harmonic retrieval problems is proposed. The method can be viewed as an extension to the rank reduction estimator (RARE) [5], originally developed for DOA estimation in partly calibrated arrays. The method is computationally efficient due to its rooting-based implementation, makes explicit use of the rich Vandermonde structure in the measurement data, and therefore shows improved estimation performance compared to conventional search-free methods for multidimensional frequency estimation.

The multidimensional RARE (MD RARE) algorithm estimates the frequencies in the various dimensions sequentially. The dimensionality of the estimation problem and the computational complexity of the estimator is significantly reduced exploiting the Vandermonde structure of the data model. This approach yields high estimation accuracy, moderate identifiability conditions, and automatically associated parameter estimates along the various dimensions. The performance of the algorithm is illustrated at the example of MIMO communication channel estimation based on the double-directional channel model. Numerical examples based on simulated and measured data recorded from the RUSK vector channel sounder at 2 GHz are presented.

## 2. SIGNAL MODEL

For simplicity of notation, we formulate the signal model for the 2D case in detail. Here, the original MIMO channel estimation problem reduces to a single-input multiple-output (SIMO) channel problem, where the parameters of interest are the propagation delays  $\tau_p$  and the DOAs  $\phi_p$  for  $p = 1, \dots, P$ . Extensions of the proposed algorithm to higher numbers of dimensions are straightforward. Consider a superposition of  $P$  discrete-time 2D exponentials corrupted by noise and let  $(a_p, b_p) \in \mathbb{C}^{1 \times 2}$ ,  $|a_p| = |b_p| = 1$ , denote the generator pair corresponding to the  $p$ th discrete 2D harmonic,

$$y_{k,\ell}(i) = \sum_{p=1}^P w_p(i) a_p^k b_p^\ell + n_{k,\ell}(i), \quad (5)$$

$$i = 1, \dots, J, \quad k = 1, \dots, K, \quad \ell = 1, \dots, L.$$

Here,  $a_p = e^{-j(2\pi/K)\tau_p}$ ,  $b_p = e^{-j(2\pi d_R/\lambda) \cos \phi_p}$ ,  $K$  and  $L$  mark the sample support along the  $a$ - and the  $b$ -axis, respectively, and  $J$  is the number of SIMO snapshots available. The Khatri-Rao product (columnwise Kronecker product) of matrix  $\mathbf{U}$  and matrix  $\mathbf{V}$  is defined as,  $\mathbf{U} \circ \mathbf{V} = [\mathbf{u}_1 \otimes \mathbf{v}_1, \mathbf{u}_2 \otimes \mathbf{v}_2, \dots]$ , where  $\mathbf{u}_k \otimes \mathbf{v}_k$  is the Kronecker matrix product of the  $k$ th column  $\mathbf{u}_k$  of  $\mathbf{U}$  and the  $k$ th column  $\mathbf{v}_k$  of  $\mathbf{V}$ . Introducing the vector  $\boldsymbol{\Omega} = [(a_1, b_1), \dots, (a_P, b_P)]$  containing the parameters of interest, and defining the Vandermonde matrices  $[\mathbf{A}]_{i,j} = (a_j)^i$ ,  $\mathbf{A} \in \mathbb{C}^{K \times P}$ , and  $[\mathbf{B}]_{i,j} = (b_j)^i$ ,  $\mathbf{B} \in \mathbb{C}^{L \times P}$ , the 2D harmonic retrieval problem can be stated as follows. Given the measurement data  $\mathbf{y}(i) = [y_{1,1}(i), y_{2,1}(i), \dots, y_{K-1,L}(i), y_{K,L}(i)]^T \in \mathbb{C}^{KL \times 1}$ ,

$$\mathbf{y}(i) = \mathbf{H}(\boldsymbol{\Omega})\mathbf{w}(i) + \mathbf{n}(i), \quad i = 1, \dots, N, \quad (6)$$

determine the parameter vector  $\boldsymbol{\Omega}$  associated with all 2D harmonics. Here, the 2D signal matrix  $\mathbf{H}(\boldsymbol{\Omega})$  is formed as the Khatri-Rao product of the Vandermonde matrices  $\mathbf{B}$  and  $\mathbf{A}$ , that is,

$$\mathbf{H}(\boldsymbol{\Omega}) = \mathbf{B} \circ \mathbf{A} \in \mathbb{C}^{KL \times P}, \quad (7)$$

$\mathbf{y}(i)$  denotes the measurement vector,  $\mathbf{w}(i) = [w_1, \dots, w_P]^T \in \mathbb{C}^{P \times 1}$  stands for the complex envelope of the  $P$  harmonics,

$\mathbf{n}(i)$  is the vector of additive zero-mean complex (circular) Gaussian noise with covariance matrix  $E\{\mathbf{n}(i)\mathbf{n}^H(i)\} = \sigma^2 \mathbf{I}_{KL}$ . In this paper the linear parameters  $\mathbf{w}(i)$  and the noise variance  $\sigma^2$  are treated as nuisance parameters. Once the parameter vector  $\boldsymbol{\Omega}$  is determined the estimation of these parameters is straightforward [6]. Equation (6) describes the 2D harmonic retrieval problem which can be solved by the conventional ESPRIT algorithm [7] and the multidimensional ESPRIT (MD ESPRIT) algorithm [3]. In the following we derive a new search-free eigenspace-based estimation method for the general case in (6) which yields highly accurate estimates of the parameters of interest.

Let the data covariance matrix be given by

$$\mathbf{R} = E\{\mathbf{y}(i)\mathbf{y}^H(i)\} = \mathbf{E}_S \boldsymbol{\Lambda}_S \mathbf{E}_S^H + \mathbf{E}_N \boldsymbol{\Lambda}_N \mathbf{E}_N^H, \quad (8)$$

where  $(\cdot)^H$  denotes the Hermitian transpose, and  $E\{\cdot\}$  stands for statistical expectation. The diagonal matrices  $\boldsymbol{\Lambda}_S \in \mathbb{R}^{(P \times P)}$  and  $\boldsymbol{\Lambda}_N \in \mathbb{R}^{(KL-P) \times (KL-P)}$  contain the signal-subspace and the noise-subspace eigenvalues of  $\mathbf{R}$ , respectively. In turn, the columns of the matrices  $\mathbf{E}_S \in \mathbb{C}^{(KL \times P)}$  and  $\mathbf{E}_N \in \mathbb{C}^{(KL \times (KL-P))}$  denote the corresponding signal-subspace and noise-subspace eigenvectors. The finite sample estimates are given by

$$\hat{\mathbf{R}} = \frac{1}{J} \sum_{i=1}^J \mathbf{y}(i)\mathbf{y}^H(i) = \hat{\mathbf{E}}_S \hat{\boldsymbol{\Lambda}}_S \hat{\mathbf{E}}_S^H + \hat{\mathbf{E}}_N \hat{\boldsymbol{\Lambda}}_N \hat{\mathbf{E}}_N^H. \quad (9)$$

**Definition 1.** We define the two Vandermonde vectors  $\mathbf{a} = (1, a, a^2, \dots, a^{K-1})^T$  and  $\mathbf{b} = (1, b, b^2, \dots, b^{L-1})^T$ . Let  $\mathbf{I}_n$  be the  $n \times n$  identity matrix. We define two ‘‘tall’’ matrices  $\mathbf{T}_a$  and  $\mathbf{T}_b$  via

$$\begin{aligned} \mathbf{T}_a &= \mathbf{I}_L \otimes \mathbf{a} \in \mathbb{C}^{KL \times L}, \\ \mathbf{T}_b &= \mathbf{b} \otimes \mathbf{I}_K \in \mathbb{C}^{KL \times K}. \end{aligned} \quad (10)$$

### 3. THE 2D RARE ALGORITHM

In the derivation of the 2D RARE algorithm, we use the following assumptions.

**Assumption 1.** The number of harmonics does not exceed the smaller of the two numbers  $(K-1)L$  and  $K(L-1)$ , that is,

$$P \leq KL - \max\{K, L\}. \quad (11)$$

**Assumption 2.** The signal matrix  $\mathbf{H}(\boldsymbol{\Omega}) \in \mathbb{C}^{KL \times P}$  (7) has full column rank  $P$ .

**Assumption 3.** The column-reduced signal matrices

$$\begin{aligned} \mathbf{H}_a(\boldsymbol{\Omega}) &= [\mathbf{h}_{a,1}, \dots, \mathbf{h}_{a,P}] = (\mathbf{B} \circ \mathbf{A}_r) \in \mathbb{C}^{(K-1)L \times P}, \\ \mathbf{H}_b(\boldsymbol{\Omega}) &= [\mathbf{h}_{b,1}, \dots, \mathbf{h}_{b,P}] = (\mathbf{B}_r \circ \mathbf{A}) \in \mathbb{C}^{K(L-1) \times P} \end{aligned} \quad (12)$$

with Vandermonde matrices

$$\begin{aligned} [\mathbf{A}_r]_{i,j} &= (a_j)^i \in \mathbb{C}^{(K-1) \times P}, \\ [\mathbf{B}_r]_{i,j} &= (b_j)^i \in \mathbb{C}^{(L-1) \times P} \end{aligned} \quad (13)$$

have full column rank. Note that the matrices  $\mathbf{A}_r$ ,  $\mathbf{B}_r$  can be obtained from  $\mathbf{A}$ ,  $\mathbf{B}$  by deleting the last row.

**Remark 1.** In most realistic applications, Assumptions 2 and 3 hold true *almost surely*, that is, with probability 1. Specifically, it can be shown that if the generators  $\{(a_p, b_p)\}_{p=1}^P$  are drawn from a distribution  $P_{\mathcal{L}}(\mathbb{C}^{2P})$  that is assumed to be continuous with respect to the Lebesgue measure in  $\mathbb{C}^{2P}$ , then the violation of Assumptions 2 and 3 is a probability-zero event [4].

**Remark 2.** Note that Assumption 3 implies that each generator  $a_i$  and  $b_j$  occurs with multiplicity  $M_a < L$  and  $M_b < K$  in the generator sets  $\{a_i\}_{i=1}^P$  and  $\{b_j\}_{j=1}^P$ , respectively.

See Appendix A for the proof.

**Proposition 1.** *Provided that Assumptions 1, 2, and 3 are satisfied, the augmented matrix*

$$\mathbf{G}_a = [\mathbf{T}_a | \mathbf{H}(\boldsymbol{\Omega})] \in \mathbb{C}^{KL \times (L+P)} \quad (14)$$

*has full column rank if and only if  $a \neq a_p$ , for  $p = 1, \dots, P$ ,  $P \leq L(K-1)$ . Similarly, provided that Assumptions 1, 2 and 3 are satisfied, the augmented matrix*

$$\mathbf{G}_b = [\mathbf{T}_b | \mathbf{H}(\boldsymbol{\Omega})] \in \mathbb{C}^{KL \times (K+P)} \quad (15)$$

*has full column rank if and only if  $b \neq b_p$  for  $p = 1, \dots, P$ ,  $P \leq (L-1)K$ .*

See Appendix B for the proof.

With Proposition 1 and provided that  $\{a_1, \dots, a_P\}$  are the true signal generators along the  $a$ -axis, the quadratic form

$$\begin{aligned} F_{R,a}(a) &= \boldsymbol{\gamma}^H \mathbf{H}(\boldsymbol{\Omega})^H (\mathbf{I}_{KL} - \mathbf{T}_a (\mathbf{T}_a^H \mathbf{T}_a)^{-1} \mathbf{T}_a^H) \mathbf{H}(\boldsymbol{\Omega}) \boldsymbol{\gamma} \\ &= 0, \quad \text{for } a \in \{a_1, \dots, a_P\}, \\ &> 0, \quad \text{otherwise,} \end{aligned} \quad (16)$$

for arbitrary vector  $\boldsymbol{\gamma} \in \mathbb{C}^P \setminus \{\mathbf{0}\}$ ,  $|a| = 1$ , and  $P \leq L(K-1)$ . It can readily be verified that the signal matrix  $\mathbf{H}(\boldsymbol{\Omega})$  and the signal-subspace matrix  $\mathbf{E}_S$  span the same subspace [6], that is, there exist a full-rank matrix  $\mathbf{K}$  such that  $\mathbf{H}(\boldsymbol{\Omega}) = \mathbf{E}_S \mathbf{K}$ . From identity (16), we can formulate one of the main results of the paper.

**Proposition 2.** *Provided that  $\{a_1, \dots, a_P\}$  are the true signal generators along the  $a$ -axis, then*

$$\begin{aligned} F_{R,a}(a) &= \tilde{\boldsymbol{\gamma}}^H \mathbf{E}_S^H (\mathbf{I}_{KL} - \mathbf{T}_a (\mathbf{T}_a^H \mathbf{T}_a)^{-1} \mathbf{T}_a^H) \mathbf{E}_S \tilde{\boldsymbol{\gamma}} \\ &= 0, \quad \text{for } a \in \{a_1, \dots, a_P\}, \\ &> 0, \quad \text{otherwise,} \end{aligned} \quad (17)$$

where  $\boldsymbol{\gamma} \in \mathbb{C}^P \setminus \{\mathbf{0}\}$ ,  $\tilde{\boldsymbol{\gamma}} = \mathbf{K}\boldsymbol{\gamma}$ ,  $|a| = 1$ ,  $P \leq L(K-1)$ , and  $\mathbf{K}$  is defined as above.

Since  $\bar{\mathbf{y}}$  is an arbitrary nonzero vector, identity (17) easily translates into an equivalent condition on the harmonic  $a$  given by

$$\begin{aligned} F_{R,a}(a) &= \det \left\{ \mathbf{E}_S^H \left( \mathbf{I}_{KL} - \mathbf{T}_a (\mathbf{T}_a^H \mathbf{T}_a)^{-1} \mathbf{T}_a^H \right) \mathbf{E}_S \right\} \\ &= 0, \quad \text{for } a \in \{a_1, \dots, a_P\}, \\ &> 0, \quad \text{otherwise.} \end{aligned} \quad (18)$$

In other words, the 1D matrix polynomial

$$\mathbf{M}_a(a) \triangleq \mathbf{E}_S^H \left( \mathbf{I}_{KL} - \mathbf{T}_a (\mathbf{T}_a^H \mathbf{T}_a)^{-1} \mathbf{T}_a^H \right) \mathbf{E}_S \in \mathbb{C}^{P \times P} \quad (19)$$

becomes singular (i.e., rank deficient) at exactly  $P$  locations  $a$  with  $|a| = 1$ . These locations  $a$  are the true generators  $\{a_p\}_{p=1}^P$ . In accordance with (18), the fundamental idea of the 2D RARE algorithm consists in determining the  $P$  true harmonics from the roots of the RARE matrix polynomial  $\mathbf{M}_a(a)$  located on the unit circle, that is, the true generators  $\{a_p\}_{p=1}^P$  are given by the solutions of the polynomial equation

$$\begin{aligned} F_{R,a}(a)|_{|a|=1} &= \det \left\{ \mathbf{E}_S^H \left( \mathbf{I}_{KL} - \mathbf{T}_a (\mathbf{T}_a^H \mathbf{T}_a)^{-1} \mathbf{T}_a^H \right) \mathbf{E}_S \right\} \\ &= 0. \end{aligned} \quad (20)$$

Up to now we have considered estimating the generator  $a$  along a single data dimension, that is, the  $a$ -axis. The solution of (20) corresponds to the 1D RARE algorithm for harmonic retrieval originally proposed in [5]. Following similar consideration as above, Proposition 1 reveals that the true generators  $\{b_p\}_{p=1}^P$  are given by the roots of the 1D matrix polynomial in  $b$ ,

$$\mathbf{M}_b(b) \triangleq \mathbf{E}_S^H \left( \mathbf{I}_{KL} - \mathbf{T}_b (\mathbf{T}_b^H \mathbf{T}_b)^{-1} \mathbf{T}_b^H \right) \mathbf{E}_S \in \mathbb{C}^{P \times P}, \quad (21)$$

evaluated on the unit circle. The associated RARE polynomial equation reads

$$\begin{aligned} F_{R,b}(b)|_{|b|=1} &= \det \left\{ \mathbf{E}_S^H \left( \mathbf{I}_{KL} - \mathbf{T}_b (\mathbf{T}_b^H \mathbf{T}_b)^{-1} \mathbf{T}_b^H \right) \mathbf{E}_S \right\} \\ &= 0. \end{aligned} \quad (22)$$

In the finite sample case, the true signal-subspace eigenvectors  $\mathbf{E}_S$  in (20) and (22) are replaced by their finite sample estimates defined in (9). Due to finite sample and noise effects, the signal roots of the RARE polynomial equations are displaced from their ideal positions on the unit circle. In this case the signal roots are determined as the  $P$  roots of (20) and (22) inside the unit circle that are located closest to the unit circle [8].

In the preceding considerations, the estimation criteria provided by (20) and (22) were derived from Proposition 1 to separately determine the generator sets  $\{a_p\}_{p=1}^P$  and  $\{b_p\}_{p=1}^P$ . Interestingly, Proposition 1 can further be exploited to develop a parameter association procedure from which the true parameter pairs  $\{(a_p, b_p)\}_{p=1}^P$  are efficiently obtained.

**Corollary 1.** *Given the true generator sets  $\{a_p\}_{p=1}^P$  and  $\{b_p\}_{p=1}^P$ , we construct the 2D matrix polynomial via the convex linear combination of (19) and (21),*

$$\bar{\mathbf{M}}(a, b) = \alpha \mathbf{M}_a(a) + (1 - \alpha) \mathbf{M}_b(b). \quad (23)$$

*This 2D matrix polynomial becomes singular for real  $0 < \alpha < 1$  if and only if  $(a, b)$  is a true generator pair. Specifically, if  $(a_p, b_p)$  denotes the generator pair of the  $p$ th harmonic, then  $\bar{\mathbf{M}}(a_p, b_p)$  contains exactly one zero eigenvalue ( $\rho_{p,0} = 0$ ) with the associated eigenvector  $\hat{\mathbf{y}}_{p,0} = \mathbf{k}_p$  denoting the  $p$ th column of the full-rank matrix  $\mathbf{K}$  defined through relation  $\mathbf{H}(\boldsymbol{\Omega}) = \mathbf{E}_S \mathbf{K}$ , here equivalence holds up to complex scaling of the columns of  $\mathbf{K}$ .*

See Appendix C for the proof.

Corollary 1 provides a powerful tool for associating the two sets of parameter estimates  $\{\hat{a}_i\}_{i=1}^P$  and  $\{\hat{b}_j\}_{j=1}^P$  that were separately obtained from the RARE criteria (20) and (22) along the  $a$ - and the  $b$ -axis, respectively. For a specific harmonic  $\hat{a}_i$  of the first set, the corresponding harmonic  $\hat{b}_j$  of the second set is given by the element of  $\{\hat{b}_j\}_{j=1}^P$  that minimizes the cost function

$$\begin{aligned} F_{\text{pair},i}(j) &= \lambda_{\min} \{ \bar{\mathbf{M}}(\hat{a}_i, \hat{b}_j) \} \\ &= \lambda_{\min} \{ \alpha \mathbf{M}(\hat{a}_i) + (1 - \alpha) \mathbf{M}(\hat{b}_j) \} \end{aligned} \quad (24)$$

for an appropriately chosen  $\alpha$  between 0 and 1. Here,  $\lambda_{\min} \{ \bar{\mathbf{M}}(\hat{a}_i, \hat{b}_j) \}$  denotes the smallest eigenvalue of  $\bar{\mathbf{M}}(\hat{a}_i, \hat{b}_j)$  (23).

#### 4. IMPLEMENTATION

In this section we provide a short description of the 4D-RARE algorithm for estimating the 4D harmonics associated with the general channel estimation problem in (3) for the finite sample case. Define the generator sets  $\Phi_1 = \{a_1, \dots, a_P\}$ ,  $\Phi_2 = \{b_1, \dots, b_P\}$ ,  $\Phi_3 = \{c_1, \dots, c_P\}$ , and  $\Phi_4 = \{d_1, \dots, d_P\}$  and initialize source index  $S = 0$ .

*Step 1.* Estimate the sample covariance matrix  $\hat{\mathbf{R}}$  and the signal-subspace eigenvectors  $\hat{\mathbf{E}}_S$ , for example, from (9).

*Step 2.* Find the roots of the RARE polynomials along the four dimensions

$$\begin{aligned} F_{R,a}(a) &= \det \left\{ \hat{\mathbf{E}}_S^H \left( \mathbf{I} - \mathbf{T}_a (\mathbf{T}_a^H \mathbf{T}_a)^{-1} \mathbf{T}_a^H \right) \hat{\mathbf{E}}_S \right\} = 0, \\ F_{R,b}(b) &= \det \left\{ \hat{\mathbf{E}}_S^H \left( \mathbf{I} - \mathbf{T}_b (\mathbf{T}_b^H \mathbf{T}_b)^{-1} \mathbf{T}_b^H \right) \hat{\mathbf{E}}_S \right\} = 0, \\ F_{R,c}(c) &= \det \left\{ \hat{\mathbf{E}}_S^H \left( \mathbf{I} - \mathbf{T}_c (\mathbf{T}_c^H \mathbf{T}_c)^{-1} \mathbf{T}_c^H \right) \hat{\mathbf{E}}_S \right\} = 0, \\ F_{R,d}(d) &= \det \left\{ \hat{\mathbf{E}}_S^H \left( \mathbf{I} - \mathbf{T}_d (\mathbf{T}_d^H \mathbf{T}_d)^{-1} \mathbf{T}_d^H \right) \hat{\mathbf{E}}_S \right\} = 0 \end{aligned} \quad (25)$$

for

$$\begin{aligned}
\mathbf{T}_a &= \mathbf{I}_{LMN} \otimes \mathbf{a} \in \mathbb{C}^{KLMN \times LMN}, \\
\mathbf{T}_b &= \mathbf{I}_{MN} \otimes \mathbf{b} \otimes \mathbf{I}_K \in \mathbb{C}^{KLMN \times MNK}, \\
\mathbf{T}_c &= \mathbf{I}_N \otimes \mathbf{c} \otimes \mathbf{I}_{KL} \in \mathbb{C}^{KLMN \times NKL}, \\
\mathbf{T}_d &= \mathbf{d} \otimes \mathbf{I}_{KLM} \in \mathbb{C}^{KLMN \times KLM},
\end{aligned} \tag{26}$$

and we substitute  $\mathbf{T}_a^H \triangleq \mathbf{T}_{1/a}^T$ ,  $\mathbf{T}_b^H \triangleq \mathbf{T}_{1/b}^T$ ,  $\mathbf{T}_c^H \triangleq \mathbf{T}_{1/c}^T$ , and  $\mathbf{T}_d^H \triangleq \mathbf{T}_{1/d}^T$ .

*Step 3.* Determine estimates of the generator sets  $\Phi_1, \Phi_2, \Phi_3$ , and  $\Phi_4$  as the roots located closest to the unit circle of the polynomials in (25) and denoted by  $\Phi_1^{(P)}, \Phi_2^{(P)}, \Phi_3^{(P)}$  and  $\Phi_4^{(P)}$ , respectively. Here, the superscript “(P)” indicates the number of elements in the set. Let  $\hat{u}_{i,k}$  denote the  $k$ th element of the  $i$ th set.

*Step 4.* Set  $S := S + 1$ . Pick a well-separated<sup>1</sup> generator  $\hat{u}_{i,k}$  from any of the sets  $\{\Phi_i^{(P-S+1)}\}$  for  $i = 1, \dots, 4$ .

*Step 5.* For  $j = 1, \dots, 4$  with  $j \neq i$ , find the corresponding root  $\hat{u}_{j,m}$  from the set  $\Phi_j^{(P-S+1)}$  such that the cost function

$$\begin{aligned}
F_{\text{pair},k}^{(i,j)}(\mathbf{m}) &= \lambda_{\min}\{\tilde{\mathbf{M}}(\hat{u}_{i,k}, \hat{u}_{j,m})\} \\
&= \lambda_{\min}\{\alpha \mathbf{M}(\hat{u}_{i,k}) + (1 - \alpha) \mathbf{M}(\hat{u}_{j,m})\}
\end{aligned} \tag{27}$$

is minimized for fixed  $\alpha$  between 0 and 1. Store the solution  $\hat{u}_{j,m}$  in the  $(j, S)$ th entry of the  $(4 \times P)$  association matrix  $\hat{\mathbf{Z}}$  and remove it from the set  $\Phi_j^{(P-S+1)}$ .

*Step 6.* Repeat Steps 4 and 5 until  $S = P$  and all entries of the  $(4 \times P)$  association matrix  $\hat{\mathbf{Z}}$  are filled. Matrix  $\hat{\mathbf{Z}}$  represents the RARE estimate of the true generator matrix  $\mathbf{Z}$ ,

$$\mathbf{Z} = \begin{bmatrix} a_1 & a_2 & \cdots & a_P \\ b_1 & b_2 & \cdots & b_P \\ c_1 & c_2 & \cdots & c_P \\ d_1 & d_2 & \cdots & d_P \end{bmatrix} \tag{28}$$

with mutually associated harmonic estimates along its columns.

*Step 7.* For each 4D harmonic  $(\hat{a}_i, \hat{b}_i, \hat{c}_i, \hat{d}_i)$ ,  $i = 1, \dots, P$ , obtained in the previous step, determine the corresponding delay  $\hat{\tau}_i$ , the direction of arrival  $\hat{\phi}_i$ , the DOD  $\hat{\theta}_i$ , and the Doppler shift  $\hat{\nu}_i$  according to the arguments of the estimates in (4).

The source parameter association procedure in Steps 5, 6, and 7 is based on the pairwise association of all 4D harmonics and stems from the observation that all 4D harmonics are

<sup>1</sup>In order to guarantee uniqueness and best performance in the parameter association, it is recommended to pick a root  $\hat{u}_{i,k}$  (and an associated set  $\Phi_i^{(P-S+1)}$ ) which is well separated in terms of angular distance  $d_i(k, l) = |\arg\{\hat{u}_{i,k}\} - \arg\{\hat{u}_{i,l}\}|$  from the remaining roots  $\{\hat{u}_{i,l}\}_{l \neq k, l=1}^{(P-S+1)}$  in the set.

TABLE 1: Generators of the 3D harmonics used for simulation with synthetic data in Section 5.1.

$P = 3$	$a_p$	$b_p$	$c_p$
$p = 1$	$e^{j0.550\pi}$	$e^{j0.719\pi}$	$e^{j0.906\pi}$
$p = 2$	$e^{j0.410\pi}$	$e^{j0.777\pi}$	$e^{j0.276\pi}$
$p = 3$	$e^{j0.340\pi}$	$e^{j0.906\pi}$	$e^{j0.358\pi}$

separated in at least one dimension. With Corollary 1, this observation facilitates the parameter association in the sense that the general 4D parameter association problem can be reduced to the much simpler pairing problem of multiple 2D harmonics.

## 5. NUMERICAL RESULTS

### 5.1. Simulation with synthetic data

In this section simulation results using synthetic data are presented. Computer simulations are carried out for the 3D case. The signal model is defined in (3), but without the harmonics  $d_p$  and the last dimension  $n$  collapses to a singleton  $n = 1$ . The sample sizes along the  $a$ -,  $b$ -, and  $c$ -axes are chosen as  $K = L = M = 5$  and the  $\mathbf{y}(i)$  vectors have dimension  $5^3$ . The  $(5^3 \times 5^3)$  data covariance matrix is computed from  $J = 10$  independent snapshots and a number of  $P = 3$  equi-powered exponentials is assumed with the generators  $\mathbf{\Omega} = \text{vec}\{\mathbf{Z}\} = [(a_1, b_1, c_1), (a_2, b_2, c_2), (a_3, b_3, c_3)]^T$  given in Table 1. Forward-backward averaging is used to increase the effective snapshot number in order to obtain from (9) a covariance matrix estimate of sufficiently high rank. The simulations are carried out according to the signal model (6) with complex Gaussian  $\mathbf{c}(i)$ , zero mean, with covariance  $E\{\mathbf{c}(i)\mathbf{c}^H(i)\} = \mathbf{I}_3$  and  $E\{\mathbf{c}(i)\mathbf{c}^T(i)\} = \mathbf{0}$ . Complex zero-mean Gaussian noise  $\mathbf{n}(i)$  is added according to (6) with covariance matrix  $E\{\mathbf{n}(i)\mathbf{n}^H(i)\} = \sigma^2 \mathbf{I}_{125}$  and  $E\{\mathbf{n}(i)\mathbf{n}^T(i)\} = \mathbf{0}$ . The root mean squared error (RMSE) of the parameter estimates obtained by the multidimensional RARE algorithm averaged over  $R = 100$  simulation runs are plotted versus the signal to noise ratio (SNR) in Figure 1. We used the following definitions:

$$\begin{aligned}
\text{SNR} &= \frac{1}{\sigma^2}, \\
\text{RMSE}(a) &= \left( \frac{1}{RP} \sum_{r=1}^R \sum_{p=1}^P |\arg((\hat{a}_p)_r) - \arg(a_p)|^2 \right)^{1/2},
\end{aligned} \tag{29}$$

where  $(\hat{a}_p)_r$  denotes the estimate for  $a_p$  obtained in the  $r$ th simulation run (and similarly for the  $b$ - and  $c$ -generators). A comparison to the corresponding Cramer-Rao bound (CRB) [9] and to results obtained from the unitary ESPRIT algorithm [3] reveals that the new method yields estimation performance close to the CRB and clearly outperforms the popular unitary ESPRIT estimator which is based on the joint Schur decomposition.

In Figure 2 we investigate the effect of the weighting parameter  $\alpha$  used in Step 5 on the parameter association

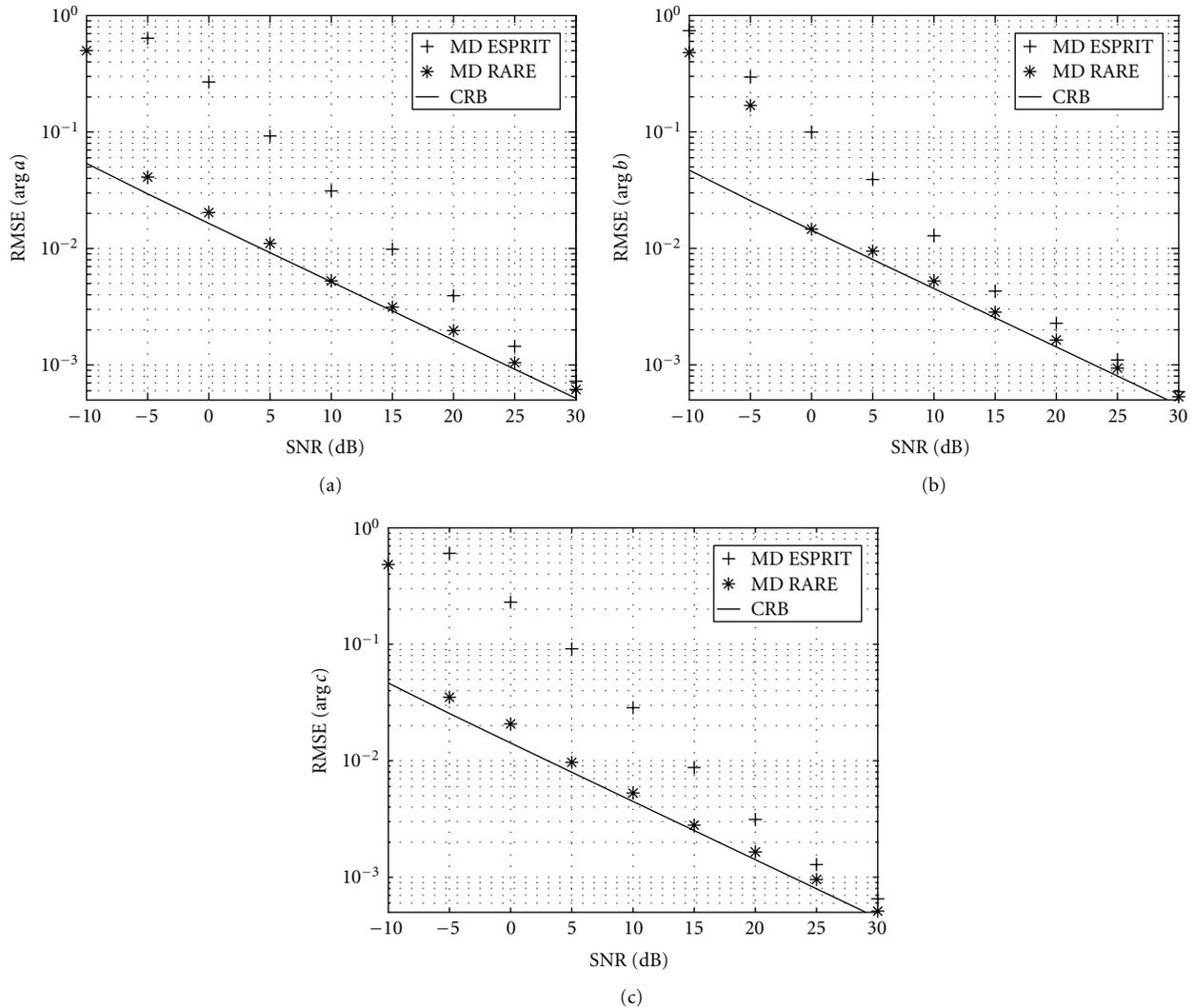


FIGURE 1: Root mean squared error of 3D RARE versus SNR.

performance. For this purpose we sorted the estimates  $\{(\arg \hat{a}_i, \arg \hat{b}_i, \arg \hat{c}_i)\}_{i=1}^3$  obtained by the 3D RARE algorithm according to  $\{\arg \hat{a}_i\}_{i=1}^3$  and plotted the RMSE of the estimates  $\{\arg \hat{b}_i\}_{i=1}^3$  and  $\{\arg \hat{c}_i\}_{i=1}^3$  against the choice of  $\alpha$  for the SNR values  $-5, 0, 5$ , and  $10$  dB. From the simulations, we observe that the proposed parameter association procedure is robust against the choice of  $\alpha$  and performs well for a wide range of  $\alpha$  taken around the intuitively expected uniform weighting factor  $\alpha = 0.5$ . We observe that a particular choice of  $\alpha$  may only affect the performance of the parameter association procedure close to threshold domain while asymptotically the choice of the weighting factor becomes less crucial.

**5.2. Measurement data**

Measurement data were recorded with the RUSK-ATM vector channel sounder, manufactured and marketed by MEDAV [2]. The measurement data used for the numerical experiments in this paper were recorded during a measurement run in Weikendorf, a suburban area in a small town ap-

proximately 50 km north of Vienna, Austria, in autumn 2001 [10, 11]. The measurement area covers one-family houses with private gardens around them. The houses are typically one floor high. A rail-road track is present in the environment which breaks the structure of single placed houses. A small pedestrian tunnel passes below the railway. A map of the environment with the position of the receiver and transmitter is shown in Figure 3.

The sounder operated at a center frequency of 2000 MHz with an output power of 2 Watt and a transmitted signal bandwidth of 120 MHz. The transmitter emitted a periodically repeated signal composed of 384 subcarriers in the band 1940–2060 MHz. The repetition period was 3.2 microseconds. The transmitter was the mobile station and the receiver was at a fixed location. The transmit array had a uniform circular geometry composed of 15 monopoles arranged on a ground plane at an interelement spacing of  $0.43\lambda \approx 6.45$  cm. The mobile transmitter was mounted on top of a small trolley together with the uniform circular array at a height of approximately 1.5 m above ground level. At

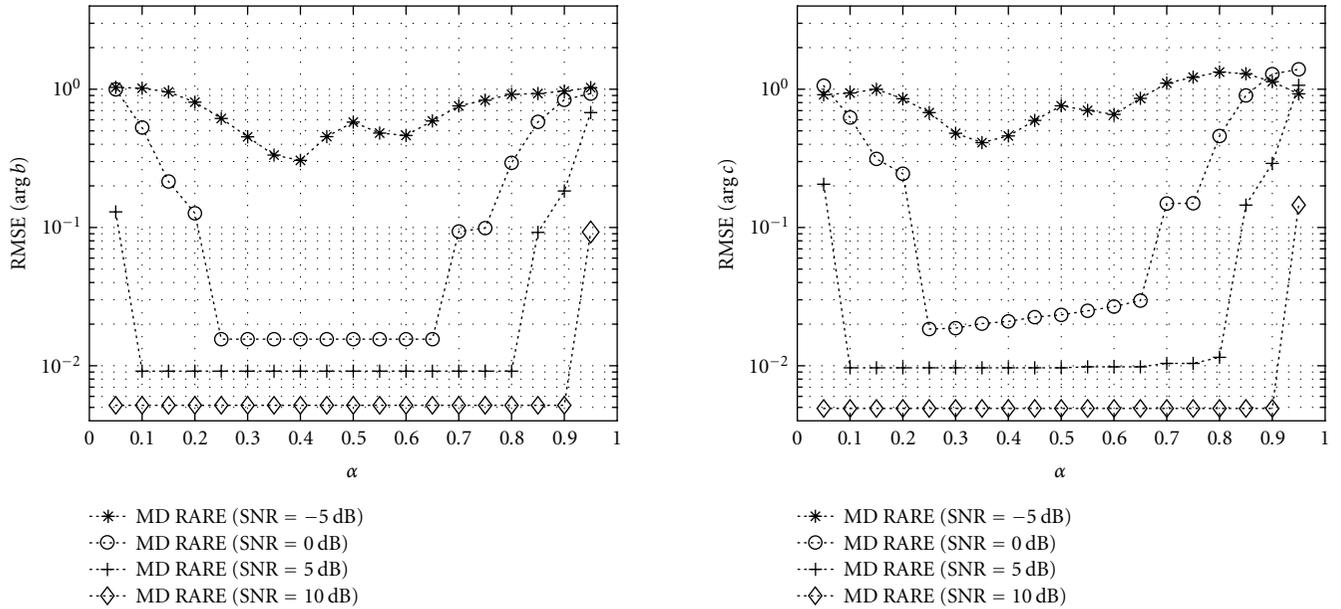
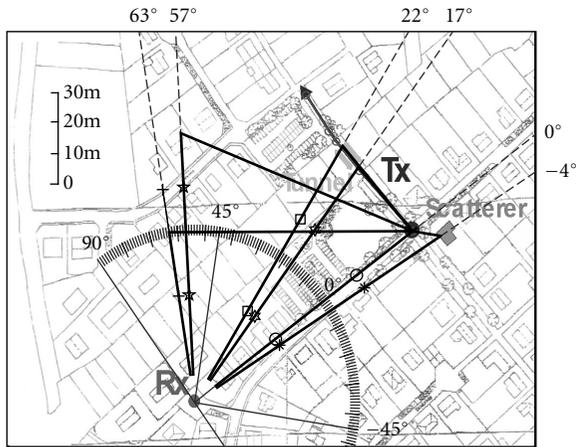
FIGURE 2: Root mean squared error of 3D RARE versus  $\alpha$ .

FIGURE 3: Map of the measurement scenario in Weikendorf.

the receiver site, a uniform linear array<sup>2</sup> composed of 8 elements with half wavelength distance (7.5 cm) between adjacent patch-elements was mounted on a lift in approximately 20 m height.

With this experimental arrangement, consecutive sets of the  $(15 \times 8)$  individual transfer functions, cross-multiplexed in time, were acquired. The receiver calculates the discrete Fourier transform over data blocks of duration 3.2 microseconds and deconvolves the data in the frequency domain with the known transmit signal. The effects from mutual coupling between Rx antenna elements are reduced by multiplying the measurement snapshots  $\mathbf{y}(i)$  with a complex-symmetric cor-

rection matrix [12]. The acquisition period of 3.2 microseconds corresponds to a maximum path length of approximately 1 km. During the measurements the receiver moved at speeds of approx. 5 km/h on the sidewalk. Rx position and Tx position as well as the motion of the transmitter are marked in the site map in Figure 3. The transmitter passed through a pedestrian tunnel approximately between times  $t = 25$  seconds and  $t = 30$  seconds of the measurement run.

We estimated the data covariance matrix from  $J = 10$  consecutive snapshots in time. The measurement system in this experiment differs from the data acquisition model described in the Introduction (1), (2), (3), and (4) in that a uniform circular array instead of a uniform linear array was used at the transmitter side. Therefore we can not simply apply the estimation procedure for the 4D parameter estimation problem described in Section 4 to estimate the directions of departure. In this experiment we only consider the 2D model (5) instead of the general 4D model (1), (2), (3), and (4). In specific we are interested in estimating only the directions of arrival and the time delays. In order to still exploit the complete 4D measurement block that was recorded as described above, we use smoothing over frequency bins and averaging over Tx samples in order to increase the number of snapshots and to obtain a full-rank covariance matrix estimate of reduced variance. Due to the smoothing over frequency bins, the original sample support of  $K = 384$  frequency bins, along the  $a$ -axis is reduced to a sample support of  $K' = 12$ . For further variance reduction we apply forward-backward (FB) averaging [3]. Making use of the notation of the general 4D model in (3) the smoothed FB sample covariance matrix corresponding to (8) reads

$$\hat{\mathbf{R}} = \frac{1}{D} \sum_{i=1}^J \sum_{k=1}^{K-K'} \sum_{m=1}^M (\tilde{\mathbf{y}}_{k,m}(i) \tilde{\mathbf{y}}_{k,m}^H(i) + \mathbf{J} \tilde{\mathbf{y}}_{k,m}^*(i) \tilde{\mathbf{y}}_{k,m}^T(i) \mathbf{J}), \quad (30)$$

<sup>2</sup>A uniform linear array was provided by T-Systems NOVA, Darmstadt, Germany.

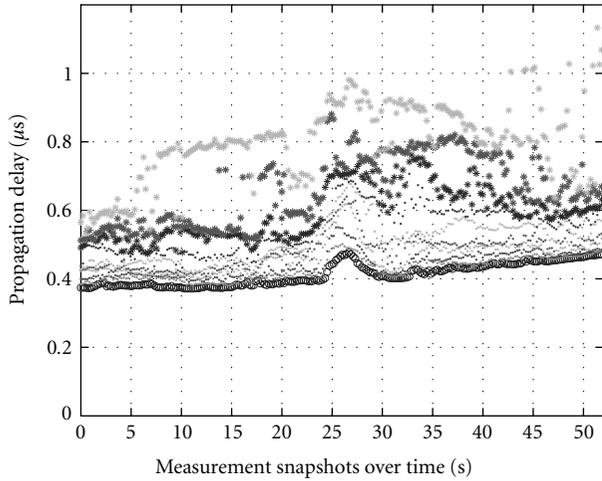


FIGURE 4: Estimates of propagation delay versus snapshots in time.

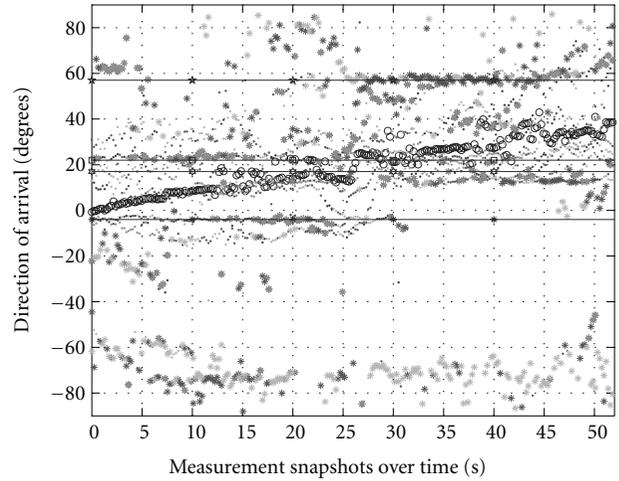


FIGURE 5: Estimates of directions of arrival versus snapshots in time.

where  $D = J(K - K')M$ ,

$$\tilde{\mathbf{y}}_{k,m}(i) = \text{vec} \left\{ \begin{bmatrix} y_{k,1,m,1}(i) & y_{k,2,m,1}(i) & \cdots & y_{k,L,m,1}(i) \\ y_{k+1,1,m,1}(i) & y_{k+1,2,m,1}(i) & \cdots & y_{k+1,L,m,1}(i) \\ \vdots & \vdots & \ddots & \vdots \\ y_{k+K',1,m,1}(i) & y_{k+K',2,m,1}(i) & \cdots & y_{k+K',L,m,1}(i) \end{bmatrix} \right\}, \quad (31)$$

$M = 15$ ,  $L = 8$ ,  $\mathbf{J}$  denotes the  $96 \times 96$  exchange matrix, and  $\mathbf{x} = \text{vec}\{\mathbf{X}\}$  denotes the vectorization operator, stacking the columns of a matrix  $\mathbf{X}$  on top of each other to form a long vector  $\mathbf{x}$ . Propagation delay and DOA estimates obtained with 2D RARE are displayed in Figures 4 and 5 relative to the orientation of the array.<sup>3</sup> We have assumed  $P = 10$  paths and applied 2D RARE for the joint estimation of propagation delay and DOA. In these two figures, the estimates are plotted as colored marks (“.” and “\*”) versus measurement time in seconds. The pairing of the estimates is indicated by the chosen mark and its color. In these figures, the circles (“o”) mark the line of sight path, dots (“.”) mark the consecutive early arrivals whereas the asterisks (“\*”) mark the late ones.

We see that the propagation scenario is dominated by a strong line-of-sight (LOS) component surrounded by local scattering paths from trees and buildings during the first 25 seconds of the experiment (shown with the “o” mark in the figures). The trace of the DOA estimates in Figure 5 and the corresponding propagation delay estimates in Figure 4 match exactly the motion of the transmitter depicted in Figure 3 for the direct path. At time 25 seconds, the trolley reaches the pedestrian tunnel and a second path resulting from scattering

at the building (see Figure 5) appears at a DOA of approximately  $-3^\circ$ . This path corresponds to a significantly larger access delay of approx. 0.55–0.58 microseconds. By the time the Tx moves out of the tunnel, the dominant LOS component with local scattering is newly tracked by the 3D RARE algorithm. In Figure 5 we observe a path emerging at a constant DOA of approx.  $22^\circ$  between snapshot time 0 second and 25 seconds. Similarly, a path emerging at a constant DOA of approx.  $17^\circ$  between time 28 seconds and 52 seconds. These paths are interpreted as contributions from the two ends of the pedestrian tunnel. Furthermore, it is interesting to observe that those propagation paths that show large delay estimates generally yield corresponding DOA estimates with large angular deviations from the line of sight.

## 6. SUMMARY AND CONCLUSIONS

A novel method for  $K$ -dimensional harmonic exponential estimation has been derived as a multidimensional extension of the conventional RARE algorithm. High-resolution frequency parameter estimates are obtained from the proposed method in a search-free procedure at relatively low computational complexity. The parameters in the various dimensions are independently estimated exploiting the rich structure of the multidimensional measurement model and the estimates of the parameters of interest are automatically associated. Simulation results based on synthetic and measured data of a MIMO communication channel underline the strong performance of the new approach. Finally, we conclude that the double-directional parametric MIMO model (3) is very suitable for describing wireless MIMO channels.

## APPENDICES

### A. PROOF OF REMARK 2

We prove by contradiction that  $M_u < L$  is necessary for  $\mathbf{H}_b(\boldsymbol{\Omega})$  to be full rank. Without loss of generality, we assume

<sup>3</sup>An animated movie generated from these results can be downloaded from FTW’s MIMO measurements, <http://www.ftw.at/measurements>.

that  $P = M_a = L$  with  $a_1 = a_2 = \dots = a_L = a$ . In this case we have  $\mathbf{H}_b(\boldsymbol{\Omega}) = (\mathbf{B}_r \circ \mathbf{A}) = (\mathbf{T}_a \mathbf{B}_r)$  where  $\mathbf{T}_a$  is defined according to Definition 1. Due to the orthogonality of the columns of  $\mathbf{T}_a$ , we have  $\text{rank}\{\mathbf{T}_a\} = L$ . Applying Sylvester's inequality yields

$$\begin{aligned} \text{rank}\{\mathbf{T}_a\} + \text{rank}\{\mathbf{B}_r\} - P \\ \leq \text{rank}\{\mathbf{T}_a \mathbf{B}_r\} \leq \min(\text{rank}\{\mathbf{T}_a\}, \text{rank}\{\mathbf{B}_r\}). \end{aligned} \quad (\text{A.1})$$

With  $P = L$ , it is easy to see that in the most general case (i.e., for distinct generators  $\{b_j\}_{j=1}^P$ ), the Vandermonde matrix  $\mathbf{B}_r$  is of rank  $L - 1$ . Equation (A.1) can then be rewritten as

$$(L - 1) \leq \text{rank}\{\mathbf{T}_a \mathbf{B}_r\} \leq (L - 1). \quad (\text{A.2})$$

In other words, the matrix  $\mathbf{H}_b(\boldsymbol{\Omega}) \in \mathbb{C}^{K(L-1) \times L}$  does not have full rank  $\{\text{rank}\{\mathbf{H}_b(\boldsymbol{\Omega})\} = \text{rank}\{\mathbf{T}_a \mathbf{B}_r\} = L - 1 < L$  which contradicts Assumption 3. Similarly we can prove that  $M_b < K$  is necessary for  $\mathbf{H}_a(\boldsymbol{\Omega})$  to be nonsingular. Further it is simple to show that the validity of Assumption 3 implies that also Assumption 2 is satisfied.

## B. PROOF OF PROPOSITION 1

In order to prove that  $\mathbf{G}_a$  has full column rank, it is sufficient to consider the limiting case  $P = L(K - 1)$  where  $\mathbf{G}_a$  becomes a square matrix. The proof is based on the application of appropriate *elementary matrix operations* applied on the rows of  $\mathbf{G}_a$ . More precisely, we exploit that adding a multiple of the row of a matrix to any other row does not change the determinant of the matrix. Similar to the procedure used in Gaussian elimination, we wish to bring the first  $L$  columns of  $\mathbf{G}_a$  into "triangular" form. Towards this aim, we subtract  $a$  times the  $(k - 1)$ th row of  $\mathbf{G}_a$  from the  $k$ th row of  $\mathbf{G}_a$ , for  $k = 2, \dots, K, K + 2, \dots, 2K, 2K + 2, \dots, 3K, \dots, (L - 1)K, (L - 1)K + 2, \dots, LK$ , that is, for all  $k \in \{1, \dots, KL\}$  such that  $(k)_K \neq 1$ , where  $(k)_K$  denotes  $k$  modulo  $K$ . The  $k$ th row of the resulting matrix denoted by  $\tilde{\mathbf{G}}_a$  is given by

$$\left[ \underbrace{0, \dots, 0}_L \mid \underbrace{b_1^{\lfloor k/K \rfloor} a_1^{(k)_K - 2} (a_1 - a), \dots, b_p^{\lfloor k/K \rfloor} a_p^{(k)_K - 2} (a_p - a)}_p \right] \quad (\text{B.1})$$

for  $(k)_K \neq 1$ . For  $(k)_K = 1$ , the rows of  $\tilde{\mathbf{G}}_a$  remain unchanged and identical to the corresponding rows of  $\mathbf{G}_a$ . Note that  $\det\{\tilde{\mathbf{G}}_a\} = \det\{\mathbf{G}_a\}$ . It can readily be verified that each of the  $L$  first columns of  $\tilde{\mathbf{G}}_a$  contain only a single nonzero element. These columns form a matrix  $\mathbf{T}_0 = \mathbf{T}_{a|a=0} = [\mathbf{e}_1, \mathbf{e}_{K+1}, \mathbf{e}_{2K+1}, \dots, \mathbf{e}_{(L-1)K+1}]$  where  $\mathbf{e}_k$  denotes the  $k$ th column of a  $KL \times KL$  identity matrix  $\mathbf{I}_{KL}$ . Making use of a well-known expansion rule for determinants, it is immediate to show that

$$\begin{aligned} \det\{\mathbf{G}_a\} &= \det\{\tilde{\mathbf{G}}_a\} = \det\{[\mathbf{T}_0 \mid \mathbf{H}_a(\boldsymbol{\Omega}) \Delta_a]\} \\ &= \pm \det\{\mathbf{H}_a(\boldsymbol{\Omega})\} \det\{\Delta_a\} \\ &= \pm \det\{\mathbf{H}_a(\boldsymbol{\Omega})\} \prod_{p=1}^P (a_p - a), \end{aligned} \quad (\text{B.2})$$

where  $\Delta_a = \text{diag}\{(a_1 - a), \dots, (a_P - a)\}$  and " $\pm$ " indicates that equality holds up to " $+$ " or " $-$ " sign. Provided that  $\mathbf{H}_a(\boldsymbol{\Omega})$  has full rank, we observe from (B.2) that for  $a \neq a_p$ , ( $p = 1, \dots, P$ ,  $P \leq L(K - 1)$ ) the determinant  $\det\{\mathbf{G}_a\} \neq 0$  and  $\det\{\mathbf{G}_a\} = 0$ , otherwise. For  $\mathbf{G}_b$  the proof follows in a similar manner from (B.2).

## C. PROOF OF COROLLARY 1

Without loss of generality, we assume that  $a = a_1, \dots, a_{M_a}$  is a true generator of multiplicity  $M_a \leq K$  that is associated with the first  $M_a$  harmonics, that is, the first  $M_a$  columns of  $\mathbf{H}(\boldsymbol{\Omega})$  (See Remark 2). From (B.2) we conclude that matrix  $\mathbf{H}(\boldsymbol{\Omega})^H (\mathbf{I}_{KL} - \mathbf{T}_a (\mathbf{T}_a^H \mathbf{T}_a)^{-1} \mathbf{T}_a^H) \mathbf{H}(\boldsymbol{\Omega})$  in (16) has exactly  $M_a$  zero eigenvalues  $\mu_{1,0} = \dots, \mu_{M_a,0} = 0$ . Furthermore, the eigenvectors corresponding to the zero eigenvalues are equivalent to the first  $M_a$  columns of a  $P \times P$  identity matrix. The last property follows from the fact that (B.2) and consequently (16) hold true for any choice of harmonics with  $P \leq L(K - 1)$  including the single harmonic case, where  $P = 1$  and  $\mathbf{H}(\boldsymbol{\Omega}) = \mathbf{h}(a_1, b_1) = \mathbf{b}_1 \circ \mathbf{a}_1$ . This observation implies that in the multiharmonic case, and with  $\mathbf{h}(a_p, b_p)$  denoting the  $p$ th column of the signal matrix  $\mathbf{H}(\boldsymbol{\Omega})$  identity,

$$\begin{aligned} F_{R,a}(a) &= \mathbf{h}^H(a_p, b_p) (\mathbf{I}_{KL} - \mathbf{T}_{a_p} (\mathbf{T}_{a_p}^H \mathbf{T}_{a_p})^{-1} \mathbf{T}_{a_p}^H) \mathbf{h}(a_p, b_p) \\ &= \mathbf{e}_p^H \mathbf{H}(\boldsymbol{\Omega})^H (\mathbf{I}_{KL} - \mathbf{T}_{a_p} (\mathbf{T}_{a_p}^H \mathbf{T}_{a_p})^{-1} \mathbf{T}_{a_p}^H) \mathbf{H}(\boldsymbol{\Omega}) \mathbf{e}_p \\ &= 0 \end{aligned} \quad (\text{C.1})$$

holds true for  $p = 1, \dots, M_a$ . That is for  $a = a_1, \dots, a_{M_a}$  and  $M_a \leq K$ , the unit vectors  $\{\mathbf{e}_p\}_{p=1}^{M_a}$  form an orthogonal basis for the nullspace of  $\mathbf{H}(\boldsymbol{\Omega})^H (\mathbf{I}_{KL} - \mathbf{T}_a (\mathbf{T}_a^H \mathbf{T}_a)^{-1} \mathbf{T}_a^H) \mathbf{H}(\boldsymbol{\Omega})$ . With  $\mathbf{H}(\boldsymbol{\Omega}) = \mathbf{E}_S \mathbf{K}$ , it is immediate that the vectors  $\{\tilde{\mathbf{y}}_{p,0} = \mathbf{K} \mathbf{e}_p = \mathbf{k}_p\}_{p=1}^{M_a}$  span the nullspace of  $\mathbf{M}(a)$  (19) denoted by  $\mathcal{N}\{\mathbf{M}(a)\}$ .

Similarly, assuming  $b = b_1, \dots, b_{M_b}$  to denote a true generator of multiplicity  $M_b \leq L$ , we obtain that the vectors  $\{\tilde{\mathbf{y}}_{p,0} = \mathbf{K} \mathbf{e}_p = \mathbf{k}_p\}_{p=1}^{M_b}$  span the nullspace  $\mathcal{N}\{\mathbf{M}(b)\}$  (21). Since by Assumption 2 all 2D harmonics can uniquely be recovered from (8), at least one of the generators  $a_p$  and  $b_p$  of a specific generator pair  $(a_p, b_p)$  is of multiplicity one. Hence, we conclude that for a true generator pair  $(a_p, b_p)$ , the associated nullspaces  $\mathcal{N}\{\mathbf{M}(a_p)\}$  and  $\mathcal{N}\{\mathbf{M}(b_p)\}$  share exactly one common nullspace vector given, for example, by  $\mathbf{k}_p$ . Moreover, the two nullspaces do not intersect if  $a_p$  and  $b_p$  solve the individual RARE polynomial equations (20) and (22) but  $(a_p, b_p)$  does not correspond to a true generator pair. That is, for a true generator pair  $(a_p, b_p)$ , the vector  $\mathbf{k}_p$  marks the intersection of the nullspaces  $\mathcal{N}\{\mathbf{M}(a_p)\}$  and  $\mathcal{N}\{\mathbf{M}(b_p)\}$  while the nullspaces do not intersect otherwise. It immediately follows that Corollary 1 holds true for arbitrary  $0 < \alpha < 1$ .

## ACKNOWLEDGMENTS

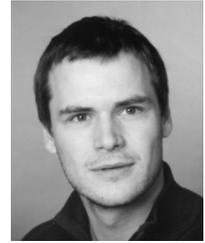
The authors would like to thank Steffen Paul, Infineon Technologies, for valuable comments on this work and Ernst Bonek for continuous encouragement and support. The Weikendorf measurements were carried out under the supervision of Helmut Hofstetter, FTW Part of this work was carried out with funding from Kplus in FTW Project C3 “Smart Antennas for UMTS Frequency Division Duplex” together with Infineon Technologies and Austrian Research Centers (ARCS), Seibersdorf.

## REFERENCES

- [1] M. Steinbauer, D. Hampicke, G. Sommerkorn, et al., “Array measurement of the double-directional mobile radio channel,” in *Proc. IEEE 51st Annual Vehicular Technology Conference*, vol. 3, pp. 1656–1662, Tokyo, Japan, May 2000.
- [2] R. S. Thomä, D. Hampicke, A. Richter, G. Sommerkorn, A. Schneider, and U. Trautwein, “Identification of time-variant directional mobile radio channels,” in *Proc. IEEE 16th Instrumentation and Measurement Technology Conference*, vol. 1, pp. 176–181, Venice, Italy, May 1999.
- [3] M. Haardt and J. A. Nossek, “Simultaneous Schur decomposition of several nonsymmetric matrices to achieve automatic pairing in multidimensional harmonic retrieval problems,” *IEEE Trans. Signal Processing*, vol. 46, no. 1, pp. 161–169, 1998.
- [4] T. Jiang, N. D. Sidiropoulos, and J. M. F. ten Berge, “Almost-identifiability of multidimensional harmonic retrieval,” *IEEE Trans. Signal Processing*, vol. 49, no. 9, pp. 1849–1859, 2001.
- [5] M. Pesavento, A. B. Gershman, and K. M. Wong, “Direction finding in partly calibrated sensor arrays composed of multiple subarrays,” *IEEE Trans. Signal Processing*, vol. 50, no. 9, pp. 2103–2115, 2002.
- [6] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part IV: Optimum Array Processing*, John Wiley & Sons, New York, NY, USA, 2002.
- [7] R. Roy and T. Kailath, “ESPRIT-estimation of signal parameters via rotational invariance techniques,” *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 37, no. 7, pp. 984–995, 1989.
- [8] M. Pesavento, A. B. Gershman, K. M. Wong, and J. F. Böhme, “Direction finding in partly calibrated arrays composed of nonidentical subarrays: a computationally efficient algorithm for the rank reduction (RARE) estimator,” in *Proc. IEEE 11th Statistical Signal Processing Workshop*, pp. 536–539, Orchid Country Club, Singapore, August 2001.
- [9] C. P. Mathews and M. D. Zoltowski, “Eigenstructure techniques for 2-D angle estimation with uniform circular arrays,” *IEEE Trans. Signal Processing*, vol. 42, no. 9, pp. 2395–2407, 1994.
- [10] H. Hofstetter, I. Viering, and W. Utschick, “Evaluation of suburban measurements by eigenvalue statistics,” in *Proc. 1st COST 273 Workshop and 4th Management Committee Meeting*, Espoo, Finland, May 2002.
- [11] H. Hofstetter, C. F. Mecklenbräuker, R. Müller, et al., “Description of wireless MIMO measurements at 2 GHz in selected environments,” in *COST-273 5th Meeting, TD(02)135*, Lisbon, Portugal, September 2002.
- [12] G. Sommerkorn, D. Hampicke, R. Klukas, A. Richter, A. Schneider, and R. S. Thomä, “Uniform rectangular antenna array design and calibration issues for 2-D ESPRIT applica-

tion,” in *Proc. 4th European Personal Mobile Communications Conference*, Vienna, Austria, February 2001.

**Marius Pesavento** was born in Werl, Germany, in 1973. He received the Dipl.-Ing. degree in electrical engineering from Ruhr-Universität Bochum, Germany, in 1999. From 1999 to 2000, he was with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ontario, Canada, where he received his M.Eng. degree (with highest honors). He is currently with the Department of Electrical Engineering, Ruhr-Universität Bochum, where he is pursuing his Ph.D. degree. His research interests include statistical signal and array processing, adaptive beamforming, and parameter estimation. Mr. Pesavento was a recipient of the 2001 Outstanding Thesis Research Award from McMaster University and the 2003 ITG best paper award from the Association of Electrical Engineering, Electronics, and Information Technologies (VDE).



**Christoph F. Mecklenbräuker** was born in Darmstadt, Germany, in 1967. He received the Dipl.-Ing. degree in electrical engineering from Vienna University of Technology in 1992 and the Dr.-Ing. degree from Ruhr-Universität Bochum in 1998, respectively. His doctoral thesis on matched field processing was awarded the Gert Massenbergr Prize. He worked for the Mobile Networks Radio Department of Siemens AG where he participated in the European framework of ACTS 90 “FRAMES.” He was a delegate to the Third Generation Partnership Project (3GPP) and engaged in the standardization of the radio access network for UMTS. Since 2000, he has been holding a senior research position at the Telecommunications Research Center Vienna (FTW) in the field of mobile communications. Currently, he gives a course at the Vienna Technical University on 3G mobile networks. He has authored around 60 papers in international journals and conferences, for which he has also served as a reviewer and holds 8 patents in the field of mobile cellular networks. His current research interests include antenna-array- and MIMO-signal processing for mobile communications.



**Johann F. Böhme** was born in Senftenberg, Germany on January 26, 1940. He received the Diploma degree in mathematics in 1966, the Dr.-Ing. in 1970, and the Habilitation in 1977, all in computer science, from the Technical University of Hanover, Germany, the University of Erlangen, Germany, and the University of Bonn, Germany, respectively. From 1967 to 1974, he was with the sonar-research laboratory of Krupp Atlas Elektronik in Bremen, Germany. He then joined the University of Bonn until 1978 and the FGAN in Wachtberg-Werthhoven. Since 1980, he has been Professor of signal theory in the Department of Electrical Engineering and Information Sciences at Ruhr-Universität Bochum, Germany. His research interests are in the domain of statistical signal processing and its applications. He is a Fellow of the Institution of Electrical and Electronic Engineers and an Elected Member of the North Rhine-Westphalian Academy of Sciences.



# Direction of Arrival Estimation with a Novel Single-Port Smart Antenna

**Chen Sun**

*School of Electrical and Electronic Engineering, Nanyang Technological University, Nanyang Avenue, Singapore 639798  
Email: csun@ieee.org*

**Nemai C. Karmakar**

*Department of Electrical and Computer Systems Engineering, Monash University, Clayton, VIC 3800, Australia  
Email: nemai.karmakar@eng.monash.edu.au*

*Received 19 May 2003; Revised 13 November 2003*

A novel direction of arrival (DOA) estimation technique that uses the conventional multiple-signal classification (MUSIC) algorithm with periodic signals is applied to a single-port smart antenna. Results show that the proposed method gives a high-resolution (1 degree) DOA estimation in an uncorrelated signal environment. The novelty lies in that the MUSIC algorithm is applied to a simplified antenna configuration. Only 1 analogue-to-digital converter (ADC) is used in this antenna, which features low power consumption, low cost, and ease of fabrication. Modifications to the conventional MUSIC algorithm do not bring much additional complexity. The proposed technique is also free from the negative influence by the mutual coupling among antenna elements. Therefore, it offers an economical way to extensively implement smart antennas into the existing wireless mobile communications systems, especially at the power consumption limited mobile terminals such as laptops in wireless networks.

**Keywords and phrases:** smart antennas, direction finding, multiple-signal classification, digital beamforming array antennas, parasitic array antenna.

## 1. INTRODUCTION

The wireless industry is growing 6 times faster than the fixed line services. The number of the mobile subscribers has grown from 0.5 billion in 1999 to 1.3 billion in 2003. The global revenue from the wireless market has been doubled from 1999 to 2003 as more than 400 billion dollars [1]. Wireless communication systems are evolving from the second generation systems to the third and fourth generation systems, which will provide high data rate multimedia services as video transmission. New value-added services such as the position location (PL) services for emergency calls, the fraud detection, intelligent transportation systems, and so forth, are also coming into reality [2, 3, 4].

Efforts have been focused on the developing coding, protocol, and modulation of the second generation system, while the role of the antennas has been overlooked. Recently, in the form of adaptive array or smart antennas, they attract interests among researchers. Deployed at the base station of the existing wireless infrastructures, they bring an outstanding capacity improvement to the frequency resource limited radio communication systems by an efficient frequency reuse scheme, for example, space division multiple-access (SDMA) scheme [5]. The direction-finding ability of the smart anten-

nas is important to the PL services. Furthermore, they also benefit the design of the routing protocol in recently thriving ad hoc networks [6].

Various beamforming and direction of arrival (DOA) estimation algorithms have been designed [1]. Simulations and experiments carried out by many researchers have shown the abilities of these algorithms. Most of these algorithms are designed based on the digital beamforming (DBF) array antennas (Figure 1). Signals received by the individual antenna elements are converted down into baseband signals, digitized and fed into a digital signal processing (DSP) chip where the algorithms reside. However, RF circuit branches connected to the array elements, analog to digital converters (ADCs) and the baseband DSP chip consume a considerable amount of DC power. Furthermore, each channel connected to the array elements has the same structure, so the cost of fabrication increases with the number of array elements [7]. All these factors make the DBF antenna arrays unsuitable for low power consumption and low cost systems and hinder the mass implementation of the smart antenna technologies. It could be too costly to equip a DBF array antenna at battery-powered laptops or mobile-computing terminals within a wireless network.

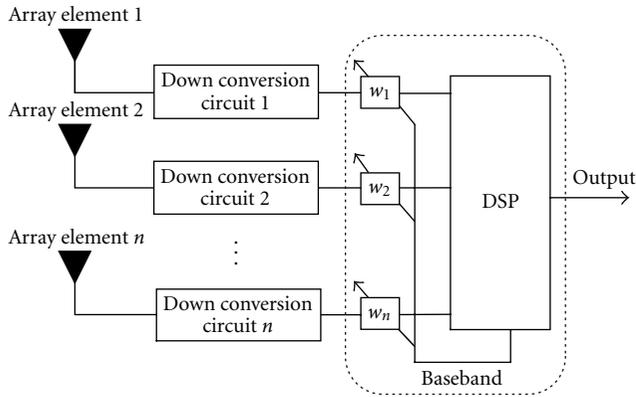


FIGURE 1: Functional block diagram of the DBF array antennas.

To circumvent the problems with the DBF array antennas, we propose a novel smart antenna structure, which has only 1 RF port. It is a parasitic array antenna [8]. One central element connected to the sole RF port and a number of surrounding switched and reactively loaded parasitic elements form the array (Figure 2). We call the antenna “switched and reactively loaded parasitic array (SARLPA).” Since the system has only 1 RF port connected to the active central element and 1 subsequent down conversion circuit, it obviously consumes much less power than a DBF antenna array. Beam steering of parasitic array antennas is achieved by either selecting different sets of parasitic elements with a switching control circuitry in a manner similar to switched beam antennas (called “switched parasitic antennas” [9, 10, 11, 12]), or by controlling the reactive loading at the parasitic elements to steer the beam continuously (called “reactively controlled directive arrays” [13] or electronically steerable parasitic array radiator (ESPAR) [14, 15, 16, 17, 18]). The SARLPA antenna is half wavelength in diameter. It may not be suitable for mobile phones, but could be easily mounted on mobile terminals in wireless networks. For example, the physical dimension of the array operating at 2.4 GHz is 62.25 mm, which is suitable for mounting on laptops.

In this paper, a high-resolution direction finding using the multiple-signal classification (MUSIC) algorithm with periodic signals is employed on this single-port smart antenna. Activating different beam patterns by switching on and off different parasitic elements, while the signal is periodically transmitted, the received scalar signal from the RF port can be stacked into a vector over one period. The beam pattern is shifted with a predefined angle. So, for the application of MUSIC, it is sufficient to know these angular shifts, without knowing exactly the response to one subset. This is similar to a uniform circular or linear antenna array in which we know the phase shift between consecutive antenna responses, without having to know those responses exactly. The MUSIC algorithm is carried out with the constructed signal vector over 1 period and the modified steering vector based on the beam pattern angle shift. Simulation justifies the proposal. Our heuristic study demonstrates the feasibility

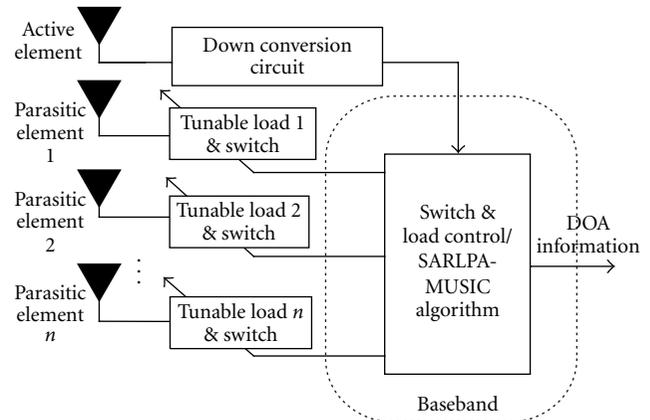


FIGURE 2: Functional block diagram of the SARLPA.

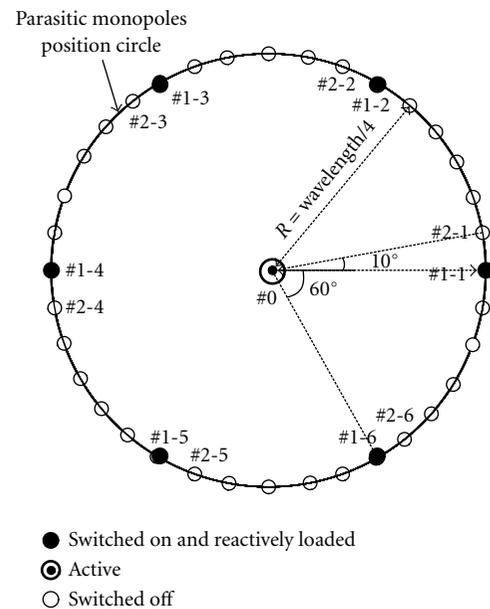


FIGURE 3: Top view of a 37-element SARLPA.

ity of employing the eigenstructure-based direction-finding algorithm with a parasitic array that has only 1 RF port.

The paper is organized as follows: Section 2 presents the configuration and the working principle of the SARLPA. The algorithm for DOA estimation is presented in Section 3. Results are produced in Section 4, followed by conclusions and recommendations in Section 5.

## 2. DESIGN OF THE ANTENNA

Figure 2 shows the functional block diagram of a SARLPA antenna. Figure 3 gives the top view of the antenna. One active central element (monopole) is surrounded by 36 parasitic elements on a circle of radius  $R$  on the circular grounded baseplate. The length of each monopole ( $L$ ) and the radius ( $R$ ) are one quarter wavelength ( $\lambda/4$ ) of the transmitting RF signal. The baseplate transforms the monopoles with their

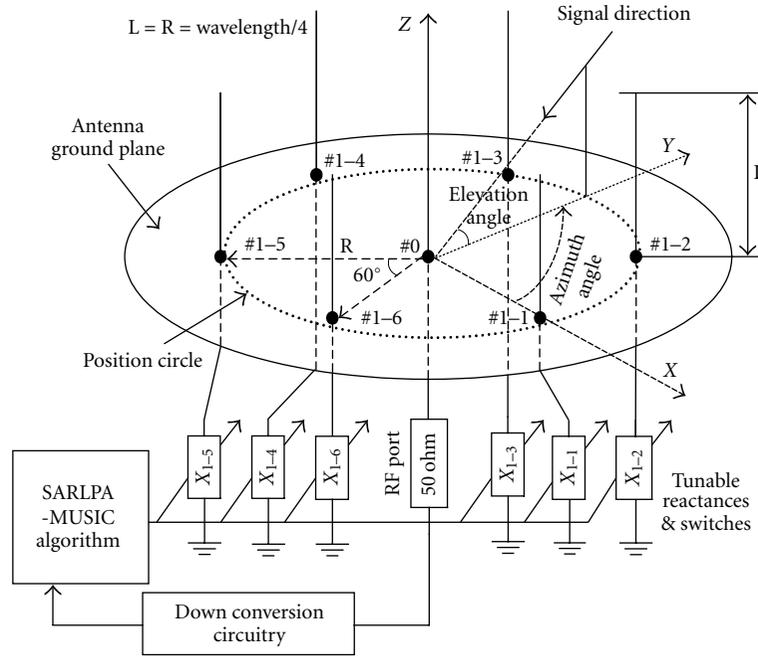


FIGURE 4: Equivalent representation of an SARLPA when subset #1 is activated. The DOA in the azimuth plane of the impinging wave is also shown in the figure.

images to dipoles with a length of  $2L$ . The central monopole is connected to an RF receiver and each parasitic monopole is loaded with a tunable reactor and is controlled by a switch. The circuitry could be simply implanted as in [19]. Since both the switches and the tunable reactances are implemented at each individual parasitic elements, the antenna provides great flexibility of operation. It can be reconfigured into a switched parasitic array or a reactively controlled directive array antenna.

As shown in Figure 3, the 36 parasitic monopoles are grouped into 6 subsets. Each subset consists of 6 parasitic elements equispaced on the circle of radius  $R$ . The central monopole is designated as #0, whereas the parasitic element # $m$ - $n$  ( $m = 1, \dots, 6$ ;  $n = 1, \dots, 6$ ) refers to the  $n$ th element of the subset # $m$ . The antenna is operated in such a way that at any time only 1 subset is switched on and the rest of the subsets are switched off. The switched-off elements do not contribute to beam forming and can be viewed as “transparent” to other switched-on elements. This is realized either by switching the parasitic monopoles to open circuit as in [9, 10], or by tuning the load reactance value to make the element invisible electrically [20]. The central element is always activated as signals pass through the sole RF port. Therefore, at any time, only 6 parasitic monopoles from the selected subset and 1 active central monopole are in operation. Figure 4 gives an equivalent representation of each subset comprised of the active central monopole and the 6 parasitic monopoles.

During operation, the first subset (parasitic monopoles #1-1, #1-2, #1-3, #1-4, #1-5, and #1-6) is selected. After a certain period, the subset #2 (parasitic monopoles #2-1, #2-

2, #2-3, #2-4, #2-5, and #2-6) is selected. As this process continues, the antenna array is virtually rotating anticlockwise on its vertical axis. The loads at parasitic monopoles are tuned in the same manner for all subsets. We have selected 36 parasitic and 1 active monopoles to satisfy the modified MUSIC algorithm as applied to the proposed SARLPA configuration. For practical applications, the number of the monopoles could be chosen with the tradeoff between the complexity and the performance (number of waves to be estimated, beam pattern directivity, etc.).

At any configuration, the antenna has the form of a reactively controlled directive array or an ESPAR. We explain the working principle based on the equivalent structure as the working principle of a SARLPA is different from that for a DBF. The antenna generates a directional beam based on tuning load reactances ( $x_1, x_2, \dots, x_6$ ) on the parasitic monopoles. Signals received or transmitted from the central RF port excite the passive parasitic monopoles with substantial induced mutual currents on them. In the following discussion, we assume that the antenna is working in the transmitting mode. Vectors  $\bar{I}$  and  $\bar{V}$  represent the currents and the voltages on the monopoles, respectively:

$$\begin{aligned} \bar{I} &= [i_0 \ i_1 \ i_2 \ i_3 \ i_4 \ i_5 \ i_6]^T, \\ \bar{V} &= [v_0 \ v_1 \ v_2 \ v_3 \ v_4 \ v_5 \ v_6]^T. \end{aligned} \quad (1)$$

Superscript  $T$  represents the transpose. Mutual admittances are represented by the matrix  $Y$  with each entity  $y_{ij}$  denotes mutual admittance between two elements [16]. There, the

induced mutual currents are represented by mutual admittances,

$$\bar{\mathbf{I}} = \mathbf{Y}\bar{\mathbf{V}} = \begin{bmatrix} y_{00} & y_{01} & y_{02} & y_{03} & y_{04} & y_{05} & y_{06} \\ y_{10} & y_{11} & y_{12} & y_{13} & y_{14} & y_{15} & y_{16} \\ y_{20} & y_{21} & y_{22} & y_{23} & y_{24} & y_{25} & y_{26} \\ y_{30} & y_{31} & y_{32} & y_{33} & y_{34} & y_{35} & y_{36} \\ y_{40} & y_{41} & y_{42} & y_{43} & y_{44} & y_{45} & y_{46} \\ y_{50} & y_{51} & y_{52} & y_{53} & y_{54} & y_{55} & y_{56} \\ y_{60} & y_{61} & y_{62} & y_{63} & y_{64} & y_{65} & y_{66} \end{bmatrix} \times \begin{bmatrix} v_0 \\ v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \end{bmatrix}. \quad (2)$$

The voltages at monopoles are

$$v_0 = v_s - z_0 i_0 \quad (3)$$

(for the active central element.)  $z_0 = 50\Omega$  is the impedance at the RF port.

$$v_l = -jx_l i_l \quad (4)$$

(for parasitic elements  $l = 1, \dots, 6$ ),  $v_s$  represents the transmitted voltage signal source with amplitude and phase from the driving RF port at the central element #0. We represent (3) and (4) with a matrix form,

$$\bar{\mathbf{V}} = \begin{bmatrix} v_s & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}^T - \mathbf{X}\bar{\mathbf{I}} = v_s \mathbf{U}_1 - \mathbf{X}\bar{\mathbf{I}}, \quad (5)$$

and define

$$\mathbf{X} = \text{diag}[z_0, jx_1, \dots, jx_6], \quad (6)$$

$$\mathbf{U}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T.$$

Combining (2) and (5), we get

$$\bar{\mathbf{I}} = \mathbf{Y}\bar{\mathbf{V}} = \mathbf{Y}(v_s \mathbf{U}_1 - \mathbf{X}\bar{\mathbf{I}}). \quad (7)$$

After a simple mathematical manipulation, we obtain the equation as

$$\bar{\mathbf{I}} = v_s (\mathbf{I}_{(7)} + \mathbf{YX})^{-1} \mathbf{YU}_1 = v_s E_1. \quad (8)$$

In (8),  $(\mathbf{I}_{(7)} + \mathbf{YX})^{-1} \mathbf{YU}_1$  is represented by a  $(7 \times 1)$ -dimensional vector  $E_1$ .  $\mathbf{I}_{(7)}$  is a  $(7 \times 7)$ -dimensional identity matrix. The far field radiation pattern is formed by the superposition of radiation patterns of all monopoles on the antenna ground plane [8, 21]. The far field current signal with its amplitude and phase in direction  $\phi_k$  is

$$y_{\text{far}}(t) = \alpha(\phi_k) \bar{\mathbf{I}} = v_s \alpha(\phi_k) E_1. \quad (9)$$

We assume that signals arrive in the azimuth plane.  $k$  is the signal sources index.  $\alpha(\phi_k)$  is the steering vector corresponding to the signal impinging direction  $\phi_k$  and is given based on the array geometry (Figure 3). Please note that only one

parasitic set and the active element are in operation,

$$\alpha(\phi_k) = \begin{bmatrix} 1 & e_1(\phi_k) & e_2(\phi_k) & e_3(\phi_k) & e_4(\phi_k) & e_5(\phi_k) & e_6(\phi_k) \end{bmatrix}$$

$$e_i(\phi_k) = \exp \left\{ j \frac{2\pi}{\lambda} R \cos \left( \phi_k - \frac{2\pi}{6} (i-1) \right) \right\}$$

$$= \exp \left\{ j \frac{\pi}{2} \cos \left( \phi_k - \frac{\pi}{3} (i-1) \right) \right\} \quad (i = 1, \dots, 6). \quad (10)$$

Setting  $v_s$  in (9) as unity, we can get the array factor  $\hat{\alpha}(\phi_k)$  of the SARLPA antenna when any one subset is activated as

$$\hat{\alpha}(\phi_k) = \alpha(\phi_k) E_1. \quad (11)$$

Thus, the far field signal in (9) is represented as

$$y_{\text{far}}(t) = \hat{\alpha}(\phi_k) v_s. \quad (12)$$

According to the reciprocity theory for radiation patterns [22], the received voltage signal  $u(t)$  at the RF port is

$$u(t) = \hat{\alpha}(\phi_k) s_k(t) + n(t), \quad (13)$$

$s_k(t)$  represents the  $k$ th far field incident current wave with amplitude and phase in the azimuth plane.  $n(t)$  is the additive white Gaussian noise (AWGN) with the power of  $\sigma^2$ . When there are totally  $K$  signal sources, the output from the RF port is

$$u(t) = \sum_{k=1}^K \hat{\alpha}(\phi_k) s_k(t) + n(t). \quad (14)$$

Since  $\hat{\alpha}(\phi_k)$  is dependent on the reactance matrix  $\mathbf{X}$ , the desired antenna beam patterns are formed by adjusting the reactance values. In practice, the variable reactance can be easily achieved by changing the bias voltage of Schottky [22] or varactor diodes loaded at parasitic monopoles.

In our proposed DOA estimation technique, the antenna is not used for adaptive beamforming, though it still has the ability to do so. Desired beam patterns are formed to explore the spatial information of the signal sources.

### 3. FORMING OF THE CORRELATION MATRIX WITH SINGLE-PORT SARLPA

The MUSIC algorithm, proposed by Schmidt [23], is a relatively simple and high-resolution eigenstructure approach. It has been widely used as a model-based approach for DOA estimation and source location problems. There are also many improved forms, such as the root MUSIC algorithm and the beamspace MUSIC algorithm [24]. All these algorithms require snapshots of signals from antenna linear or circular array elements to form the signal correlation matrix.

It is obvious that we cannot apply the conventional MUSIC algorithm to the SARLPA antenna, because it has only 1 RF port. Forming the signal sample vector is impossible at one time. This is the distinction between the MUSIC algorithm based on a single-port smart antenna and the

conventional MUSIC algorithm based on a DBF array antenna. However, we can see from the explanation in the following part that the received signals from the RF port contain signal sources direction information, and the signal correlation matrix can still be formed.

In operation, 6 subsets of parasitic elements are selected sequentially as indicated in Figure 3. The beam pattern rotates  $10^\circ$  for 6 times consecutively. Therefore, the antenna is virtually rotated. The sample signal vector  $U(t) = [u_1(t), \dots, u_m(t), \dots, u_6(t)]^T$  is constructed with signal samples from the RF output. Each entity  $u_m(t)$ ,  $m = 1, \dots, 6$  is obtained in the sampling period when the subset  $\#m$  is in operation. The maximum number of sampling periods  $M$  equals the number of subsets. For the SARLPA antenna with 6 subsets, after rotating 6 + 1 times, the beam pattern returns to the original status. The load reactance values of parasitic monopoles are the same for all the sampling periods.  $E_1$ , defined in (8), remains constant as the antenna rotates virtually. However, the signal directions with respect to the antenna  $X$ - $Y$  coordinate (Figure 4) change as the antenna is virtually rotated. According to (14), the received signal sample at each sampling period is represented as:

$$u_m(t) = \sum_{k=1}^K \hat{\alpha} \left( \phi_k - \frac{2\pi}{36} (m-1) \right) s_k(t) + n_m(t), \quad (15)$$

$$m = 1, 2, \dots, 6.$$

For the simplicity, we define

$$\hat{\alpha}_m(\phi_k) = \hat{\alpha} \left( \phi_k - \frac{2\pi}{36} (m-1) \right), \quad (16)$$

$\hat{\alpha}_m(\phi_k)$  contains the  $k$ th signal source direction information  $\phi_k$ .  $n_m(t)$  is the noise component at the  $m$ th sampling period. Suppose each individual signal source transmits periodic signals and the length of 1 signal period is the same as the sampling period. Then received signal samples are the same for all sampling periods. Therefore, the received signal vector  $U(t)$  after 6 sampling periods can be written into the following matrix multiplication form:

$$U(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \\ u_3(t) \\ u_4(t) \\ u_5(t) \\ u_6(t) \end{bmatrix} = \begin{bmatrix} \hat{\alpha}_1(\phi_1) & \hat{\alpha}_1(\phi_2) & \cdots & \hat{\alpha}_1(\phi_K) \\ \hat{\alpha}_2(\phi_1) & \hat{\alpha}_2(\phi_2) & \cdots & \hat{\alpha}_2(\phi_K) \\ \hat{\alpha}_3(\phi_1) & \hat{\alpha}_3(\phi_2) & \cdots & \hat{\alpha}_3(\phi_K) \\ \hat{\alpha}_4(\phi_1) & \hat{\alpha}_4(\phi_2) & \cdots & \hat{\alpha}_4(\phi_K) \\ \hat{\alpha}_5(\phi_1) & \hat{\alpha}_5(\phi_2) & \cdots & \hat{\alpha}_5(\phi_K) \\ \hat{\alpha}_6(\phi_1) & \hat{\alpha}_6(\phi_2) & \cdots & \hat{\alpha}_6(\phi_K) \end{bmatrix} \times \begin{bmatrix} s_1(t) \\ s_2(t) \\ \vdots \\ s_K(t) \end{bmatrix} + \begin{bmatrix} n_1(t) \\ n_2(t) \\ \vdots \\ n_6(t) \end{bmatrix}. \quad (17)$$

Equation (17) has the same form as the output from an array antenna with 6 elements. In the first matrix on the right-hand side of the equation, we define each column as

vector  $\vec{\alpha}(\phi_k)$ . It corresponds to the steering vector [25] for a 6-element array antenna as reported in array signal processing literature. But each entity is not a phase delay of a source signal induced when the plane passes the array with an impinging angle relative to the antenna. We call  $\vec{\alpha}(\phi_k)$  "SARLPA-direction vector" to differentiate it from the "steering vector" for conventional antenna arrays. The signal correlation matrix at the antenna output is  $\mathbf{R}_{UU}$ :

$$\mathbf{R}_{UU} = E(U(t)U^H(t)) = \frac{1}{Ns} UU^H. \quad (18)$$

$E(\cdot)$  is the expected value operator.  $U$  without a time argument represents the discrete samples with sample block length  $Ns$ . Superscript  $H$  denotes conjugate transpose. We assume that all the noise components  $n_m(t)$ ,  $m = 1, \dots, 6$  at each sampling period  $\#m$  are independent identically distributed (i.i.d.) AWGN with the power of  $\sigma^2$ .  $\mathbf{R}_{UU}$  is then represented as

$$\mathbf{R}_{UU} = \begin{bmatrix} \vec{\alpha}(\phi_1) & \vec{\alpha}(\phi_2) & \cdots & \vec{\alpha}(\phi_K) \end{bmatrix} \times \mathbf{R}_{SS} \times \begin{bmatrix} \vec{\alpha}(\phi_1) & \vec{\alpha}(\phi_2) & \cdots & \vec{\alpha}(\phi_K) \end{bmatrix}^H + \sigma^2 \mathbf{I}_{(6)} \quad (19)$$

$\mathbf{I}_{(6)}$  is a  $(6 \times 6)$ -dimensional identity matrix. The source signal correlation matrix  $\mathbf{R}_{SS}$  is

$$\mathbf{R}_{SS} = E \left( \begin{bmatrix} s_1(t)s_2(t) & \cdots & s_k(t) \end{bmatrix} \times \begin{bmatrix} s_1(t)s_2(t) & \cdots & s_k(t) \end{bmatrix}^H \right). \quad (20)$$

Based on (18), the MUSIC algorithm is executed in a similar procedure as the conventional MUSIC algorithm. Figure 5 gives the flowchart. The SARLPA direction vector is used to obtain the MUSIC spectrum. The algorithm requires the signals to be periodic. That means proposed technique cannot be applied when mobile terminals are transmitting message signals, because message signals are quite indeterministic. It could, however, be implemented in a pilot or a command channel or by assigning periodic codes at message channels. Therefore, we call it "MUSIC algorithm with periodic signals." Its procedure is described below.

- (i) Construct the signal sample vector given by  $U(t) = [u_1(t), \dots, u_m(t), \dots, u_6(t)]$ .  $u_m(t)$ ,  $m = 1, \dots, 6$ , is the signal sample when the antenna is at sampling period  $\#m$ . With  $U(t)$ , we form the signal correlation matrix  $\mathbf{R}_{UU}$ .
- (ii) Eigendecompose the signal correlation matrix  $\mathbf{R}_{UU}$ , and form the noise subspace  $\mathbf{E}_N$  with eigenvectors corresponding to the small eigenvalues.
- (iii) Evaluate the MUSIC spectrum  $p_{MU}$  versus the signal direction  $\phi$ ,

$$p_{MU} = \frac{1}{|\mathbf{E}_N^H \vec{\alpha}(\phi)|^2}, \quad (21)$$

$\vec{\alpha}(\phi)$  is the SARLPA direction vector corresponding to the azimuthal looking direction  $\phi$ .

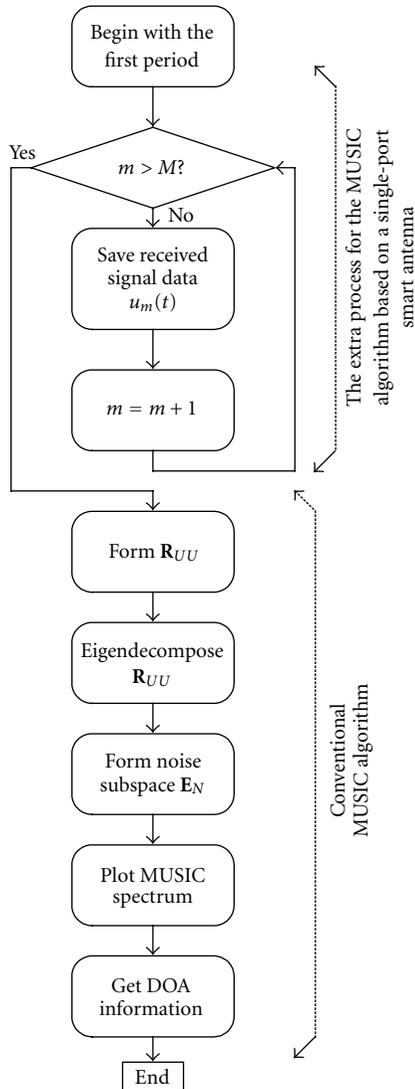


FIGURE 5: Flowchart of the conventional and the proposed MUSIC algorithm and the proposed MUSIC algorithm based on a single-port smart antenna.

- (iv) Finally, extract the impinging waves direction information.

Besides the additional procedure for forming the signal sample vector, the proposed technique follows same procedure as that for the conventional MUSIC algorithm. An extra amount of time is required for storing the data to form the signal sample vector.

#### 4. SIMULATION

In our simulation, 3 binary phase shift keying (BPSK) signals with equal power levels are randomly generated. The signal impinging directions are set arbitrarily to  $120^\circ$ ,  $150^\circ$ , and  $300^\circ$  in the azimuth plane, respectively. During each of the 6

antenna sampling periods, the length of signal samples is 16 bits. This indicates that  $N_s$  is 16 in (18).

The monopoles on the ground plane of the SARLPA antenna are oriented vertically. The impinging signals are assumed copolarized with the antenna. The cross-polarization coupling typically produced by scatters and reflectors in multipath propagation environment [26] is not considered.

#### 4.1. Influence of parasitic load reactance

We study four different reactance load settings to investigate their influence on the performance of the DOA estimation. Table 1 summarizes the reactance values for these 4 cases.

MUSIC spectra with different reactive load settings in uncorrelated signal environments where the signal-to-noise-ratio (SNR) is 40 dB for each individual signal source are presented and explained as follows.

##### 4.1.1. Case 1. Equal load reactance of $30\Omega$

In Case 1, we set all the parasitic load reactances to  $30\Omega$ . The antenna exhibits a nearly omnidirectional beam pattern. It could not be virtually rotated in a way as described previously. The MUSIC spectrum is plotted in Figure 6. From the figure, we can observe, that the signal direction could not be extracted.

##### 4.1.2. Case 2. Ordered but unequal load reactances of $0\Omega$ and $30\Omega$

In Case 2, we set the load to  $x_1 = 0\Omega$ ,  $x_2 = 30\Omega$ ,  $x_3 = 0\Omega$ ,  $x_4 = 30\Omega$ ,  $x_5 = 0\Omega$ , and  $x_6 = 30\Omega$ . The beam pattern for rotation is shown in Figure 7. The MUSIC spectrum is plotted in Figure 8. It also cannot give the direction information. From the simulation studies in Cases 1 and 2, we conclude that we should not set the loads “ordered,” which will result in insufficient beam patterns for rotation.

##### 4.1.3. Case 3. Random load reactances of $0\Omega$ and $30\Omega$

In Case 3, we arbitrarily set  $x_1 = 0\Omega$ ,  $x_2 = 30\Omega$ ,  $x_3 = 0\Omega$ ,  $x_4 = 30\Omega$ ,  $x_5 = 30\Omega$ , and  $x_6 = 30\Omega$ . The beam pattern is shown in Figure 9. The signals DOA information is clearly resolved from the MUSIC spectrum as shown in Figure 10.

##### 4.1.4. Case 4. Random load reactances

Finally, we randomly set the reactance values to  $x_1 = 0\Omega$ ,  $x_2 = 10\Omega$ ,  $x_3 = 0\Omega$ ,  $x_4 = -30\Omega$ ,  $x_5 = 70\Omega$ , and  $x_6 = 30\Omega$ . The MUSIC spectrum is plotted in Figure 10. From the figure we can see that if the reactance is not set ordered, they only slightly influence the MUSIC spectrum but not the performance. All the spectra reach the peak values at the respective signal directions. The beam pattern for this case is shown in Figure 11. Note that the proposed technique is used for DOA estimation and not for beam forming. So the beam pattern does not necessarily correspond to the signals directions.

From the above investigations, we conclude that the manner of load settings greatly influences the algorithm’s DOAs estimation performances. When setting the reactive loads, we should avoid setting them ordered, which produces

TABLE 1: Setting of six parasitic loads of each subset at the sampling period  $\#m$  ( $m = 1, \dots, 6$ ). Once the setting is decided, the values do not change during the process of the DOA estimation.

Case	Load reactances ( $\Omega$ )						MUSIC spectrum	DOA information
	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$		
1	30	30	30	30	30	30	Figure 6	Not shown in the spectrum
2	0	30	0	30	0	30	Figure 8	Not shown in the spectrum
3	0	30	0	30	30	30	Figure 10	Shown in the spectrum
4	0	10	0	-30	70	30	Figure 10	Shown in the spectrum

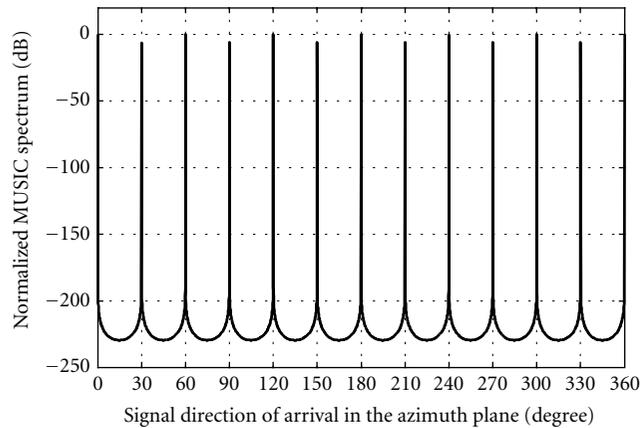


FIGURE 6: MUSIC spectrum for Case 1 is  $x_l = 30\Omega$  ( $l = 1, \dots, 6$ ), SNR = 40 dB, and the signal directions are  $120^\circ$ ,  $150^\circ$ , and  $300^\circ$  in the azimuth plane.

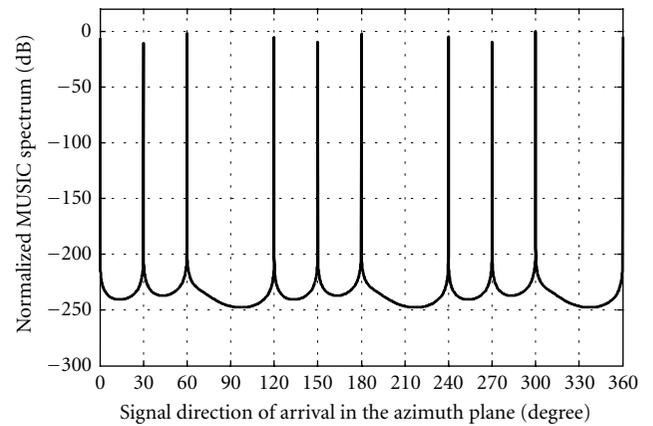


FIGURE 8: MUSIC spectrum for Case 2 is  $x_1 = 0\Omega$ ,  $x_2 = 30\Omega$ ,  $x_3 = 0\Omega$ ,  $x_4 = 30\Omega$ ,  $x_5 = 0\Omega$ , and  $x_6 = 30\Omega$ , SNR = 40 dB, and the signal directions are  $120^\circ$ ,  $150^\circ$ , and  $300^\circ$  in the azimuth plane.

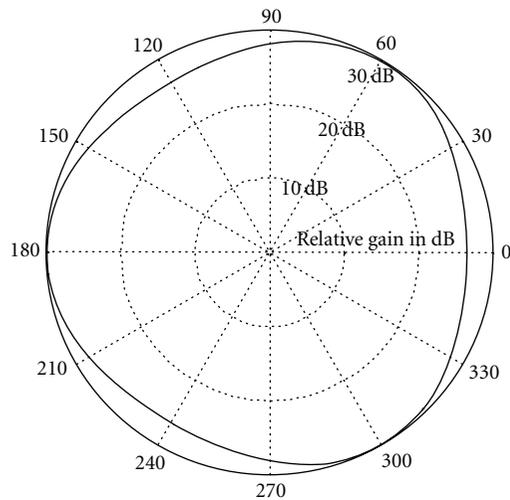


FIGURE 7: Beam pattern in the azimuth plane for the antenna sampling periods #1 to #6. Load setting is  $x_1 = 0\Omega$ ,  $x_2 = 30\Omega$ ,  $x_3 = 0\Omega$ ,  $x_4 = 30\Omega$ ,  $x_5 = 0\Omega$ , and  $x_6 = 30\Omega$ .

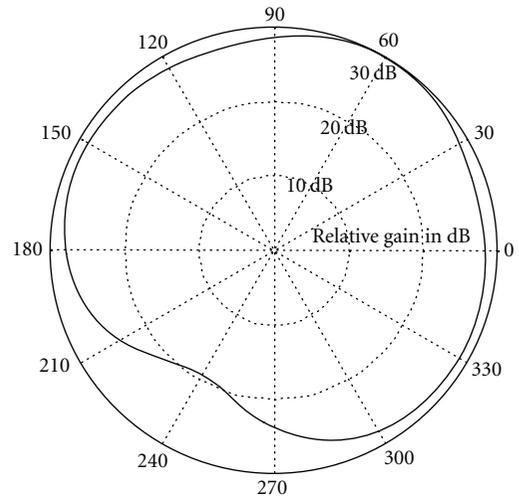


FIGURE 9: Beam pattern in the azimuth plane for antenna sampling periods #1 to #6. The load setting is  $x_1 = 0\Omega$ ,  $x_2 = 30\Omega$ ,  $x_3 = 0\Omega$ ,  $x_4 = 30\Omega$ ,  $x_5 = 30\Omega$ , and  $x_6 = 30\Omega$ .

symmetric azimuthal beam patterns. With a symmetric pattern, signals impinging from different directions may produce the same output at the RF port. Analogous to linear array antenna, image peaks appear. Rather, we can set the reactance values arbitrarily to precisely estimate the DOA information.

Forming directive beam patterns is not important to the performance. The technique is dependent of the angular shifting of the beam pattern. If only the pattern is not symmetric, the DOA can be estimated correctly. However, the load setting that gives deep nulls in radiation patterns in certain directions should be avoided for the DOA estimation,

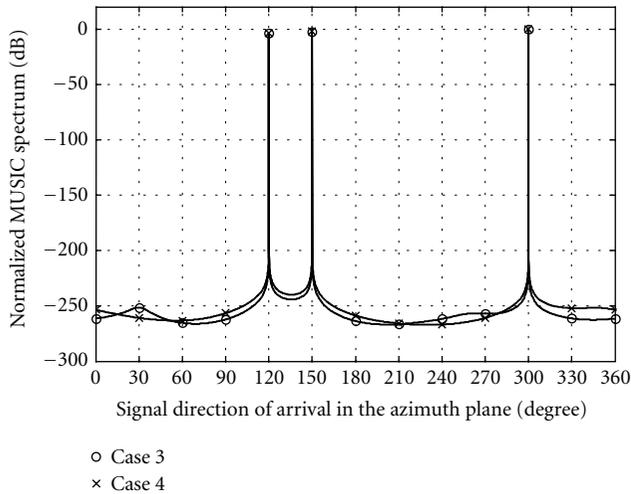


FIGURE 10: MUSIC spectra for Cases 3 and 4. SNR = 40 dB. The load settings are listed in Table 1. Signal directions are 120°, 150°, and 300° in the azimuth plane.

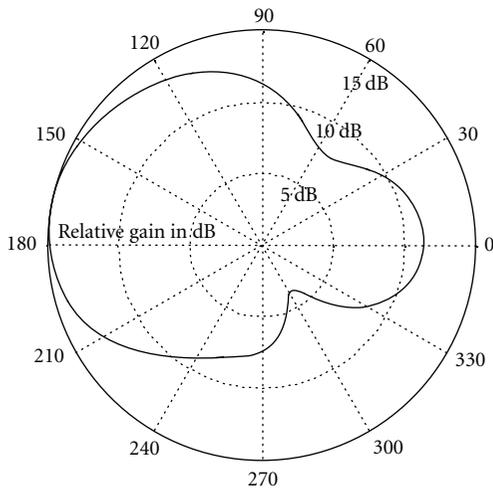


FIGURE 11: Beam pattern in the azimuth plane for antenna sampling periods #1 to #6. The load reactance is  $x_1 = 0\Omega$ ,  $x_2 = 10\Omega$ ,  $x_3 = 0\Omega$ ,  $x_4 = -30\Omega$ ,  $x_5 = 70\Omega$ , and  $x_6 = 30\Omega$ .

because signals coming from that regions could be received with low power levels, or even be nulled out. Consequently, the DOA information of signal in this null region could not be extracted.

#### 4.2. Performance in different SNRs environments

We also simulated the MUSIC spectra in signal environments with different SNRs. We suppose that all signal sources are totally uncorrelated and have the same power. The individual SNR for each signal source is set equal. The MUSIC spectra with 3 different SNRs from 0 dB to 40 dB are shown in Figure 12. The performance of our proposed MUSIC algorithm degrades as SNR decreases, but all three MUSIC spectra reach nearly the same peak values at the signal directions.

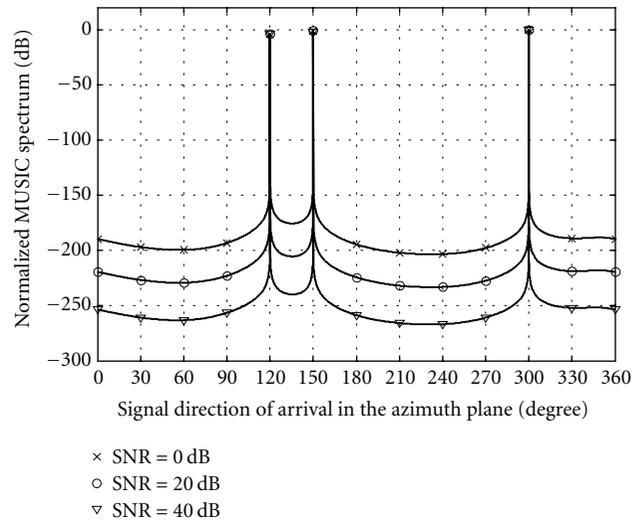


FIGURE 12: MUSIC spectra in different SNRs signal environments. Signal directions are 120°, 150°, and 300° in the azimuth plane. The load setting is same as in Case 4.

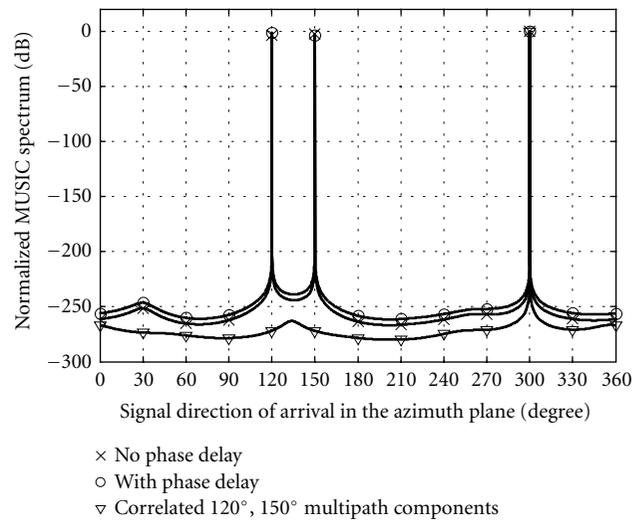


FIGURE 13: MUSIC spectra. The load setting is the same as for Case 4. SNR = 30 dB. Signal directions are 120°, 150°, and 300° in the azimuth plane.

#### 4.3. Performance in a multipath propagation environment

Our study assumes a line of sight (LOS) propagation environment (Ricean channel). In a macroscopic area, where PL services, such as intelligent transportation services, are employed, base station antennas are mounted well above the roof of the buildings to provide LOS propagations between base stations and mobile terminals [27].

Ideally, signals arrive at the antenna without phase delays. The plotted MUSIC spectrum is shown in Figure 13 with a "×"-marked solid line. In a practical wireless communication environment, multipath propagation delays of the

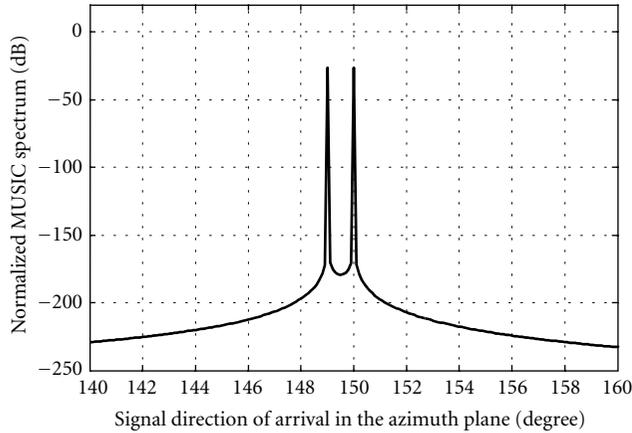


FIGURE 14: MUSIC spectrum. The load setting is the same as for Case 3. SNR = 30 dB. Two signal directions are 149° and 150° in the azimuth plane, respectively.

signals are expected. Received signals contain the LOS component and the multipath components from each far field signal source. We assume that the LOS components are uncorrelated with the multipath components [28] and that the LOS components have much higher signal power level, we could still reasonably represent the received signals at the antennas as a summation of only the LOS components from different directions with phase differences induced by the propagation delays. Signal directions are set as 120°, 150°, and 300° in the azimuth plane. The MUSIC spectrum for delayed signals is shown in Figure 13 with a circle-marked solid line. The signal direction information is obtained from the spectrum.

In a microscopic area, the base station antennas are mounted below the rooftop level to confine the signal coverage into a small local area [29]. In this case, the base stations are located in a rich scattering environment. There are no LOS components. Signals received at the base station antenna are a collection of multipath components with angular spreads, which are possible to be correlated. Figure 13 shows the resulting MUSIC spectrum with a “∇”-marked solid line. Two multipath components from 1 far field signal source arriving at the antenna from 120° and 150° in the azimuth plane are correlated. The MUSIC algorithm could not resolve correlated signals. To circumvent the rich multipath problem, the mobile terminal’s spatial signature [30] could be used to locate the mobile terminal’s location.

#### 4.4. Resolution and limitation in number of resolvable signals

The proposed technique is capable to estimate the DOAs of two signal sources with 1° angular separation. As seen from Figure 14, 2 signals impinging from 149° and 150° in the azimuth plane could be clearly resolved from the spectrum. The SNR is 30 dB.

The maximum number of uncorrelated waves to be estimated is  $M-1$ .  $M$  is the number of subsets, the number of sampling periods, and is also the dimension of the formed signal correlation matrix  $\mathbf{R}_{UU}$ . The scenario is similar to the

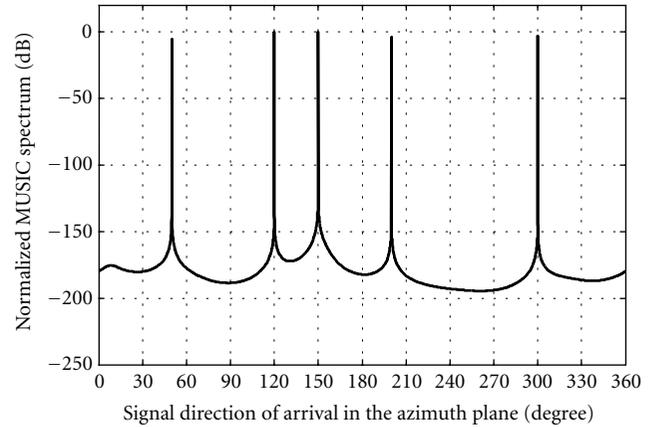


FIGURE 15: MUSIC spectrum. Load setting is the same as for Case 3. SNR = 30 dB. Signal directions are 50°, 120°, 150°, 200°, and 300° in the azimuth plane.

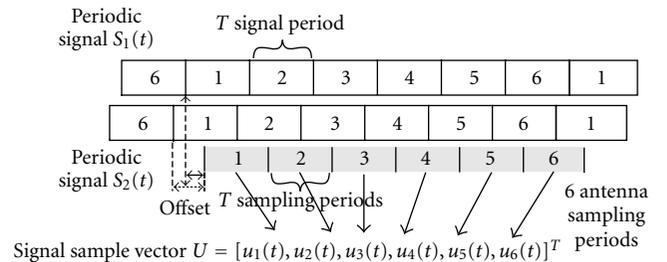


FIGURE 16: Periodic signal time slots and antenna virtual sampling periods in the proposed DOA estimation technique. Each signal contains 6 slots. The antenna rotates 6 sampling periods. Time offsets are expected in a multipath environment.

conventional MUSIC algorithm, in which  $\mathbf{R}_{UU}$  with a dimension of  $M$  could produce up to  $M-1$  signals [23, 31]. In our design, there are 6 parasitic monopoles, thus we can estimate up to 5 DOAs as shown in Figure 15. Signal directions are set to be 50°, 120°, 150°, 200°, and 300° in azimuth plane.

#### 4.5. Sampling period and periodicity of transmitted signals

Finally, we discuss the requirement of the periodicity of the transmitted signals for employing the proposed DOA estimation technique. In the derivation of  $\mathbf{R}_{UU}$ , we require that each individual signal source transmits periodic signals. In a practical wireless communication environment, time offsets among periodic signals and the antenna sampling periods, induced by different propagation path delays are expected.

However, the signal period does not need to be aligned with the sampling period (Figure 16). This is true because the signal source transmits signals longer than the length of total 6 sampling periods; and each signal period is the same as the antenna sampling period  $T$ . The received signals at the antenna are still periodic, therefore, (17) is still valid. The requirement means that when doing DOA estimation, mobile terminals will transmit periodic signals instead of message signals. With the a priori knowledge of the location of

wireless base stations, where the antennas are mounted, and with the database of the map of the local area, PL service could be provided at the expense of a short message period. This is applicable in cases when mobile subscribers are in dangerous situations and need to make a rescue calls, such as the E999 call in Singapore and the E911 call in America. For applications in ad hoc networks where the DOA information is required for data package routing, the periodical signal could be assigned during the network setup period.

#### 4.6. Speed of DOA estimation

In our design, 16 signal bits are stored at each sampling period. A minimum length of  $16 \times 6 = 96$  bits of the periodic signals for 6 antenna sampling periods is required. For time-division multiple-access (TDMA) IS-136 System, 1 frame comprises 6 time slots with the length of 6.67 millisecond each. In each time slot, 260 data bits are transmitted. This indicates that the DOA estimation is possible to be carried out within 1 TDMA time slot. The tuning of the reactance could be done in nanoseconds [22], and the off-the-shelf diodes can realize state switching in a few nanoseconds.

### 5. CONCLUSION

In this paper, a novel approach for executing the high-resolution MUSIC algorithm based on a single RF port smart antenna has been proposed. After presenting the configuration and the working principle of the antenna, the performance of the proposed technique for various aspects have been studied. The results justified that the technique is capable for a high-resolution DOA estimation of 1 degree, which is comparable with a conventional MUSIC algorithm based on DBF array antennas. The proposed technique for DOA estimation based on a signal RF port parasitic antenna has many advantages over that on a DBF. They are as follows.

- (1) In operation, only six parasitic elements are switched on. The switches consume a few milliwatt of power and the leakage current of a reverse biased diode is negligible. The SARLPA antenna features low power consumption and low complexity. It is suitable for commercial implementations at mobile terminals such as laptops in ad hoc wireless networks [7, 19] to provide direction information. It can also be easily mounted on vehicles to realize position-based services.
- (2) Contrary to the DOA estimation based on a DBF, the mutual coupling is utilized to steer the beam [13]. Thus the modified MUSIC algorithm is free from the negative influences of the mutual coupling among antenna elements. In a DBF, this influence has to be compensated for the enhancement of system performance in a high-resolution DOA estimation as studied in [32, 33].
- (3) The technique gives high-quality performance in low SNR signal environments. The performance in an SNR = 0 dB environment is shown in Figure 12.
- (4) The proposed technique is very flexible in operation. The parasitic elements are reactively loaded and con-

trolled with switches. The antenna can be configured into a reactively controlled directive array with a certain amount of reactance-loaded parasitic elements. Adaptive beamforming algorithm could be employed at the baseband to tune the beam direction [16] as an adaptive directional antenna helping to reduce the channel interferences and improve the channel capacity. If the directions of the impinging waves are known, the antenna could be reconfigured into a switched parasitic antenna. Beam patterns could be controlled by simply switching on and off different parasitic elements to function as a switched beam antenna, or sector antenna [2].

- (5) Parasitic elements are arranged to form a circular array antenna. When used for DOA estimation, it could give the signal direction information covering  $360^\circ$  in the azimuth plane.

Simulation studies have also pointed out that further improvements are needed to enhance system performances to make it more applicable to practical wireless communication systems. The areas of improvements are as follows.

- (1) The proposed technique works in an uncorrelated signal environment. New methods should explore signal sources spatial signatures in a correlated signal environment for PL services.
- (2) The proposed antenna will not suffer from the negative influence from the mutual coupling between antenna elements. However, the calibration of the antenna aperture over DOA, frequency and temperature, weather environment, and fabrication error are still unavoidable and they influence the antenna performances.
- (3) The Doppler effect of the fast moving mobile terminals needs to be investigated.
- (4) The antenna having 36 parasitic monopoles equispaced on the circle surrounding the active central element has been used to theoretically demonstrate the idea of eigenstructure-based direction finding algorithm with a single RF port smart antenna. For practical implementations, using less parasitic antenna elements could reduce the complexity. Furthermore, the proposed SARLPA with elements in the form of microstrip patch also renders a much compacter design.

The smart antennas, a critical technology for the third and beyond generation wireless communications, are under extensive studies. Many testbeds and field trials have been built to justify their potential benefits. Employed at the base stations, they could bring a higher-system capacity, coverage extension, and new services such as the cellular infrastructure-based PL services. At the wireless mobile terminals, they could bring the benefits as lower power consumption, reduced interference level. This paper offers an economical approach to implement smart antennas into the existing terrestrial wireless mobile communications system to serve as an essential and fundamental technique.

## REFERENCES

- [1] A. Reinhardt, "The wireless challenge," *Business Week*, pp. 48–58, October 2003.
- [2] T. S. Rappaport, J. H. Reed, and B. D. Woerner, "Position location using wireless communications on highways of the future," *IEEE Communications Magazine*, vol. 34, no. 10, pp. 33–41, 1996.
- [3] G. V. Tsoulos, "Smart antennas for mobile communication systems: benefits and challenges," *IEE Electronic & Communication Engineering Journal*, vol. 11, no. 2, pp. 84–94, 1999.
- [4] K. J. Krizman, T. E. Biedka, and T. S. Rappaport, "Wireless position location: fundamentals, implementation strategies, and sources of error," in *Proc. IEEE Vehicular Technology Conference*, vol. 2, pp. 919–923, Phoenix, Ariz, USA, May 1997.
- [5] W. C. Y. Lee, *Mobile Communications Engineering*, McGraw-Hill, New York, NY, USA, 1997.
- [6] R. Ramanathan and J. Redi, "A brief overview of ad hoc networks: challenges and directions," *IEEE Communications Magazine*, vol. 40, no. 5, pp. 20–22, 2002.
- [7] T. Ohira, "Adaptive array antenna beamforming architectures as viewed by a microwave circuit designer," in *Proc. Asia-Pacific Microwave Conference*, pp. 828–833, Sydney, Australia, December 2000.
- [8] D. V. Thiel and S. Smith, *Switched Parasitic Antennas for Cellular Communications*, Artech House, Norwood, Mass, USA, 2001.
- [9] S. L. Preston, D. V. Thiel, T. A. Smith, S. G. O'Keefe, and J. W. Lu, "Base-station tracking in mobile communications using a switched parasitic antenna array," *IEEE Trans. Antennas and Propagation*, vol. 46, no. 6, pp. 841–844, 1998.
- [10] R. Vaughan, "Switched parasitic elements for antenna diversity," *IEEE Trans. Antennas and Propagation*, vol. 47, no. 2, pp. 399–405, 1999.
- [11] D. V. Thiel, S. G. O'Keefe, and J. W. Lu, "Electronic beam steering in wire and patch antenna systems using switched parasitic elements," in *Proc. IEEE Antennas and Propagation Society International Symposium*, vol. 1, pp. 534–537, Baltimore, Md, USA, July 1996.
- [12] S. L. Preston, D. V. Thiel, J. W. Lu, S. G. O'Keefe, and T. S. Bird, "Electronic beam steering using switched parasitic patch elements," *Electronics Letters*, vol. 33, no. 1, pp. 7–8, 1997.
- [13] R. F. Harrington, "Reactively controlled directive arrays," *IEEE Trans. Antennas and Propagation*, vol. AP-26, no. 3, pp. 390–395, 1978.
- [14] T. Ohira and K. Gyoda, "Hand-held microwave direction-of-arrival finder based on varactor-tuned analog aerial beamforming," in *Proc. Asia-Pacific Microwave Conference*, vol. 2, pp. 585–588, Taipei, Taiwan, December 2001.
- [15] K. Yang and T. Ohira, "ESPAR antennas-based signal processing for DS-CDMA signal waveforms in ad hoc network systems," in *Proc. IEEE 3rd Workshop on Signal Processing Advances in Wireless Communications*, pp. 130–133, Taoyuan, Taiwan, March 2001.
- [16] J. Cheng, Y. Kamiya, and T. Ohira, "Adaptive beamforming of ESPAR antenna based on steepest gradient algorithm," *IEICE Transactions on Communications*, vol. E84-B, no. 7, pp. 1790–1800, 2001.
- [17] C. Sun and N. C. Karmakar, "Adaptive beamforming of ESPAR antenna based on simultaneous perturbation stochastic approximation theory," in *Proc. Asia-Pacific Microwave Conference*, vol. 1, pp. 192–195, Kyoto, Japan, November 2002.
- [18] C. Sun, A. Hirata, T. Ohira, and N. C. Karmakar, "Fast beamforming of electronically steerable parasitic array radiator antennas: theory and experiment," *IEEE Trans. Antennas and Propagation*, vol. 52, no. 7, pp. 1819–1832, 2004.
- [19] N. C. Karmakar and M. E. Bialkowski, "A beam-forming network for a circular switched-beam phased array antenna," *IEEE Microwave and Wireless Components Letters*, vol. 11, no. 1, pp. 7–9, 2001.
- [20] K. Iigusa and T. Ohira, "Reconfigurable array antenna with reactively controlled electronically invisible parasitic radiators," IEICE Tech. Rep. AP2002-122, The Institute of Electronics, Information and Communication Engineers, 2003, pp. 19–24.
- [21] S. A. Leonov and A. I. Leonov, *Handbook of Computer Simulation in Radio Engineering, Communications, and Radar*, Artech House, Norwood, Mass, USA, 2001.
- [22] S. Y. Liao, *Microwave Solid-State Devices*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1985.
- [23] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas and Propagation*, vol. AP-34, no. 3, pp. 276–280, 1986.
- [24] L. C. Godara, "Application of antenna arrays to mobile communications. II. Beam-forming and direction-of-arrival considerations," *Proc. IEEE*, vol. 85, no. 8, pp. 1195–1245, 1997.
- [25] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques*, Prentice-Hall PTR, Englewood Cliffs, NJ, USA, 1993.
- [26] D. C. Cox, R. R. Murray, H. W. Arnold, A. Norris, and M. Wozowicz, "Cross-polarization coupling measured for 800 MHz radio transmission in and around houses and large buildings," *IEEE Trans. Antennas and Propagation*, vol. AP-34, no. 1, pp. 83–87, 1986.
- [27] W. C. Jakes, *Microwave Mobile Communications*, IEEE Press, Piscataway, NJ, USA, 1994.
- [28] C. Tepedelenlioglu, *Modeling and Mitigation of Time- and Frequency-Selective Fading in Single- and Multi-Carrier Communications*, Ph.D. thesis, Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, Minn, USA, March 2001.
- [29] J. Parsons, *The Mobile Radio Propagation Channel*, Halsted Press, New York, NY, USA, 1992.
- [30] S.-S. Jeng, G. Xu, H.-P. Lin, and W. J. Vogel, "Experimental studies of spatial signature variation at 900 MHz for smart antenna systems," *IEEE Trans. Antennas and Propagation*, vol. 46, no. 7, pp. 953–962, 1998.
- [31] R. O. Schmidt and R. E. Franks, "Multiple source DF signal processing: An experimental system," *IEEE Trans. Antennas and Propagation*, vol. AP-34, no. 3, pp. 281–290, 1986.
- [32] Y. Inoue, K. Mori, and H. Arai, "DOA error estimation using 2.6 GHz DBF array antenna," in *Proc. Asia-Pacific Microwave Conference*, vol. 2, pp. 701–704, Taipei, Taiwan, December 2001.
- [33] K. R. Dandekar, H. Ling, and G. Xu, "Effect of mutual coupling on direction finding in smart antenna applications," *Electronics Letters*, vol. 36, no. 22, pp. 1889–1891, 2000.

**Chen Sun** is now pursuing the Ph.D. degree at Nanyang Technological University, Singapore. His research interests include adaptive beamforming, array signal processing, MIMO communications, and array antenna design. From November 2002 to March 2003, he was with ATR Adaptive Communications Research Laboratories, Kyoto, Japan as a student intern, working on personal wireless links for wireless ad hoc networks. He was an Invited Session (Active and Adaptive Array Antennas) Cochair in Asia-Pacific Microwave Conference 2002, Kyoto, Japan.



**Nemai C. Karmakar** obtained his M.S. and Ph.D. in electrical engineering from the University of Saskatchewan, Canada, and University of Queensland, Australia in 1991 and 1999, respectively. In August 1990, he was as a Research Assistant at the Communications Research Group, the University of Saskatchewan. From 1992 to 1995 he worked as a Microwave Design Engineer at Mitec Ltd., Australia, where he contributed



significantly to the development of land mobile satellite antennas for the Australian Mobilesat. He taught at Queensland University of Technology, Australia in 1995–1996. From September 1998 to March 1999 he worked as a research engineer within the Radar Laboratory, Nanyang Technological University, Singapore. Since March 1999 he is an Assistant Professor and Graduate Advisor in the Division of Communications, the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. He has published more than 100 referred journal and conference papers and three book chapters. He is a senior member of IEEE. He served as a member of the executive committee of Singapore IEEE Section in 2000 and is now an executive committee member of IEEE APS/MTT Chapter. He is a Senior Lecturer in the Department of Electrical and Computer Systems Engineering, Monash University, Clayton Campus, Monash.

# High-Performance Wireless via the Merger of CI Chip-Shaped DS-CDMA and Oscillating-Beam Smart Antenna Arrays

**Seyed Alireza Zekavat**

*Department of Electrical and Computer Engineering, Michigan Technological University, Houghton, MI 49931, USA  
Email: rezaz@mtu.edu*

**Carl R. Nassar**

*Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, CO 80523-1373, USA  
Email: carln@colostate.edu*

**Steve Shattil**

*Idris Communications, 1500 Cherry St. Suite L, Louisville, CO 80027, USA  
Email: steve@ciansystems.com*

*Received 29 May 2003; Revised 15 October 2003*

We introduce a novel merger of direct sequence code division multiple access (DS-CDMA) and smart antenna arrays. With regard to the DS-CDMA scheme, we employ carrier interferometry DS-CDMA (CI/DS-CDMA), a novel implementation of DS-CDMA where chips are decomposable into  $N$  narrowband frequency components. With regard to the antenna array, we deploy the oscillating-beam smart array. Here, applying proper time-varying phases to the array elements, we create small movement (oscillation) in the antenna array's pattern, while steering the antenna pattern main lobe to the position of the intended user. The oscillating antenna pattern creates a time-varying channel with a controllable coherence time. This, in turn, provides transmit diversity in the form of a time diversity gain at the mobile receiver side. At the receiver, three stages of combining are available: combining time components of the received signal within symbol duration  $T_S$  (each experiencing a different fade) to enhance performance via time diversity; combining frequency components which make up the CI/DS-CDMA chip to enhance the performance via frequency diversity; and combining across chips to eliminate the interfering users on the system. Merging CI/DS-CDMA with the oscillating-beam smart antenna at the base station, we achieve very high capacity via the merger of SDMA (available through directionality of the antenna array) and code division multiple access (inherent in CI/DS-CDMA), and very high performance via the construction of receivers that exploit both transmit diversity and frequency diversity. We present the performance gains of the proposed merger.

**Keywords and phrases:** smart antennas, antenna arrays, DS-CDMA systems, transmit diversity, carrier interferometry.

## 1. INTRODUCTION

Antenna arrays located at the base station (BS) enhance wireless communication systems via (1) directionality, which supports space division multiple access (SDMA); or (2) more recently, a transmit diversity benefit, that is, a diversity scheme that uses the antenna array at the BS to exploit diversity at the mobile (see, e.g., [1, 2]). Recently, the authors have introduced a new antenna array scheme in [3, 4, 5, 6] which offers both (1) high capacity via SDMA and (2) excellent probability-of-error performance at the mobile via its transmit diversity benefits. Both benefits are available while maintaining low mobile receiver complexity.

In the authors' proposed antenna array of [3, 4, 5, 6], a unique, carefully controlled time-varying phase shift is applied to each antenna array element, sweeping the beam pattern directed to the mobile such that (1) the beam pattern maintains a constant large scale fade for the symbol duration  $T_S$ ; (2) the beam pattern ensures  $L$  independent fades within each  $T_S$ ; (3) after each  $T_S$ , the antenna beam returns to its initial position, and sweeps same area of space over  $T_S$  (leading to an oscillating antenna pattern and easing parameter estimation); (4) the movement of the beam pattern, as a percentage of half-power beamwidth (HPBW), is small, allowing the beam pattern to maintain directionality; and (5) the bandwidth expansion due to beam pattern movement is

negligible. In this paper, we merge this novel antenna array technique with DS-CDMA systems.

Direct-sequence code division multiple access (DS-CDMA) [7] is the world's most popular CDMA architecture. In DS-CDMA, each user's bit is multiplied by a sequence of  $N$  chips (short pulses of duration  $T_C$ ), where each chip has amplitude  $+1$  or  $-1$ . By careful selection of  $+1$  and  $-1$  values (spreading sequences), the receiver can separate users one from another. To enhance performance via path diversity (e.g., [8]), most DS-CDMA systems employ RAKE receivers, which attempt to separate and linearly recombine the multiple paths.

Recently, a novel chip shape referred to as the CI (carrier interferometry) chip shape was introduced to DS-CDMA [9, 10, 11, 12, 13]. Here, each chip is decomposable into  $N$  orthogonal carrier components. As a result, when applying these chip shapes, (1) the DS-CDMA receiver: achieves a frequency diversity benefit (rather than a path diversity benefit) by decomposing chips into carrier components and frequency combining; and (2) the use of frequency combining in place of path combining (as done in RAKE receivers) leads to a significantly improved performance via the ability to avoid interpath interference [9, 10, 11, 12, 13].

In this work, we innovatively apply the oscillating-beam antenna arrays of [3, 4, 5, 6] to DS-CDMA systems with CI chip shapes (CI/DS-CDMA) of [9, 10, 11, 12, 13]. This enables (1) very high capacity via the merger of SDMA (directionality of the antenna array) and CDMA (inherent in CI/DS-CDMA); and (2) very high performance via the construction of receivers that exploit both transmit diversity and frequency diversity. We focus on the performance benefits of the proposed merger.

In this work, we assume carrier frequency ( $f_0$ ) much larger than system bandwidth (BW) (e.g.,  $f_0 > 100 \cdot (BW)$ ), a reasonable assumption in today's mobile systems. Hence, the antenna pattern is identical for the entire transmit bandwidth. With this in mind, the CI/DS-CDMA signal is fed into a single  $M$ -element smart antenna array. By carefully designing the phase shifts applied to antenna array elements, the resulting beam pattern corresponds to an oscillating beam pattern similar to that in [3, 4, 5, 6]. This leads to a time-varying channel with a controllable coherence time. The controllable coherence time is used by the mobile to exploit time diversity and enhance performance.

The benchmark for comparison in this work is a CI/DS-CDMA system employed in conjunction with a conventional smart antenna array (an antenna array which creates an adaptive beam pattern directed toward the intended user, leading to increased capacity via SDMA, but, unlike the proposed scheme, offers no improvements in the performance of CI/DS-CDMA system). This work highlights the performance benefits that can be achieved by small oscillations in the beam pattern of the smart antenna array.

Receivers are constructed to exploit both the transmit diversity, which corresponds to an induced time diversity provided by the antenna array, as well as the diversity inherent in the CI/DS-CDMA system (an exploitable frequency diversity). Thus, at the receiver, three stages of combining are

present: (1) a combining of the time components with different fades (to exploit time diversity), (2) a combining across frequency components (to exploit frequency diversity), and (3) a combining across chips (to eliminate users in a traditional DS-CDMA manner). We can apply the combining first on the frequency components or first on the time components, that is, the first and the second combining stages can be interchanged.

Assuming (a) rich scattering environment (where, up to 7-fold time diversity is achievable via beam pattern movement [5, 6]), (b) fully loaded CI/DS-CDMA with a processing gain of  $N = 32$  (i.e.,  $K = 32$  orthogonal users are available in the system), and (c) 4-fold frequency diversity over the entire bandwidth, simulation results demonstrate that the proposed system achieves 14 dB gain over a CI/DS-CDMA system with a conventional smart antenna array at a probability of error of  $10^{-3}$ . Performance gains are even more impressive when the proposed system is compared to a traditional DS-CDMA system with a conventional smart antenna array. (These performance benefits are, in addition to the usual network capacity gains, provided via SDMA.)

Section 2 introduces the merger of the beam-sweeping smart antenna arrays and CI/DS-CDMA. Section 3 presents receiver structures employing equal gain combining (EGC) across frequency components followed by minimum mean square error combining (MMSEC) across time domain components. Section 4 presents simulated performance results, while Section 5 presents a conclusion.

## 2. THE MERGER OF CI/DS-CDMA AND BEAM-SWEEPING ANTENNA ARRAYS

### 2.1. The CI/DS-CDMA system

In DS-CDMA, considering a binary phase shift keying (BPSK) modulation, a unique time sequence ( $N$  chips, each with amplitude  $+1$  or  $-1$ ) is assigned to each user. Hence, the  $k$ th user's data bit,  $b_k$ , is sent as

$$s_k(t) = \text{Re} \{ b_k \cdot C_k(t) \cdot e^{j2\pi f_0 t} \}, \quad (1)$$

where  $b_k$  is  $+1$  or  $-1$ ,  $f_0$  is the center or carrier frequency, and  $C_k(t)$  is the  $k$ th user spreading code, corresponding to

$$C_k(t) = \sum_{i=0}^{N-1} c_k^i \cdot h(t - iT_C) \cdot g(t). \quad (2)$$

Here,  $c_k^i \in \{-1, +1\}$  is the  $i$ th element of user  $k$ 's spreading code,  $T_C = T_S/N$  is the chip duration, and  $g(t)$  is a rectangular waveform limiting the chip shape to duration  $T_S$ . In the proposed CI/DS-CDMA system of [9, 10, 11, 12, 13], the chip shape  $h(t)$  corresponds to a multicarrier signal. Specifically, the chip shape  $h(t)$  is a superpositioning of  $N$  narrowband subcarriers equally spaced in frequency by  $\Delta f$ :

$$h(t) = \sum_{n=0}^{N-1} e^{j2\pi n \Delta f t}. \quad (3)$$

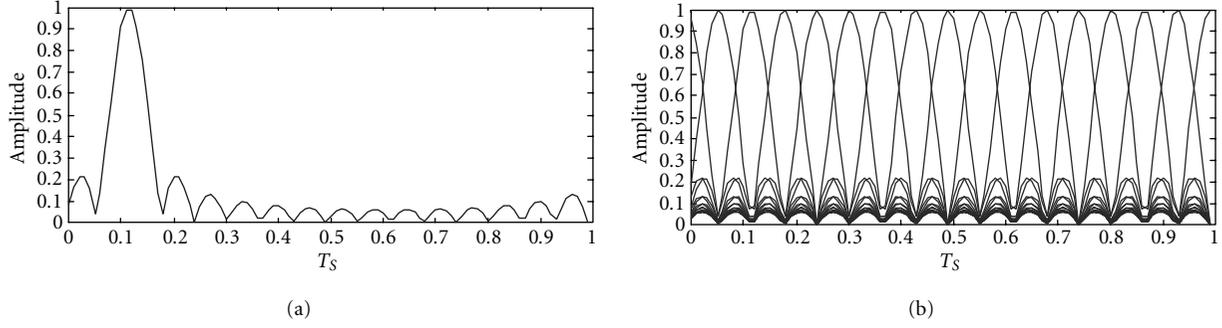


FIGURE 1: (a)  $h(t)$  and (b)  $\{h(t - iT_C), i = 0, 1, \dots, N - 1\}$  ( $N = 16$ ).

Here,  $\Delta f \geq 1/T_S$  to ensure orthogonality (separability) between subcarriers of the chip shape. The magnitude of  $h(t)$  is shown in the Figure 1a assuming  $\Delta f = 1/T_S$  and the set  $\{h(t - iT_C), i = 0, 1, \dots, N - 1\}$  is shown in Figure 1b. It is important to note that  $h(t - jT_C)$  and  $h(t - kT_C)$  ( $k \neq j$ ) are orthogonal to one another, and hence chips are separable. Referring to (3) and Figure 1, the CI chip shape corresponds to a frequency sampled version of the sinc( $\cdot$ ) shape.

## 2.2. Proposed antenna array structure

The CI/DS-CDMA signal characterized by (1), (2), and (3) is fed into a single  $M$ -element antenna array (see Figure 2). The  $m$ th array element applies the phase shift  $m\theta(t, \phi)$ ,  $m \in [0, M - 1]$ .

The beam pattern oscillation is created by careful selection of the antenna element's phase offset,  $\theta(t, \phi)$ . The antenna array's beam pattern movement will be designed to ensure (1) constant large scale fading, that is, the mean of the Rayleigh fading is constant over the symbol duration  $T_S$ ; and (2) that  $L$  independent fades are generated within each partition  $T_S$ . In other words, the antenna array beam pattern is swept in a manner which ensures constant large scale fading over symbol duration  $T_S$  while ensuring  $L$  independent fades within each  $T_S$ .

### Criterion 1: Constant large-scale fading

To ensure constant large-scale fading over each symbol time duration  $T_S$ , the beam pattern must remain in the antenna array HPBW. That is [4]

$$\left| T_S \cdot \frac{d\phi}{dt} \right| = \kappa \cdot \text{HPBW}, \quad 0 < \kappa < 1, \quad (4)$$

where  $\phi$  is the azimuth angle,  $d\phi/dt$  is the rate of antenna pattern movement, and  $T_S \cdot (d\phi/dt)$  is the amount of antenna pattern movement in  $T_S$ . The parameter  $\kappa$ , selected such that  $0 < \kappa < 1$ , guarantees that the received antenna pattern amplitude is within the 3 dB beamwidth for the entire symbol duration. This parameter  $\kappa$  is referred to throughout as the *antenna array control parameter* as it determines (restricts) the amount of beam pattern movement permitted in time duration  $T_S$ . Equation (4) corresponds to selecting the phase

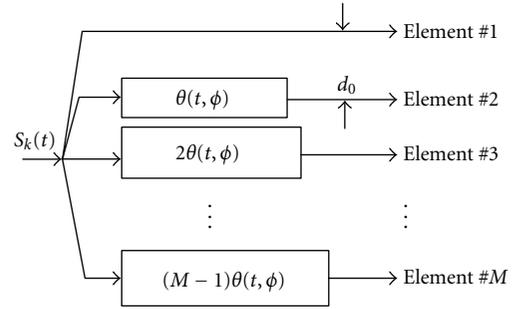


FIGURE 2: The antenna array structure over which the CI/DS-CDMA signal is sent.

offset applied to the antenna elements according to [4]

$$\theta(t, \phi) = \kappa \cdot \frac{2\pi d_0 \cdot |\sin(\phi)| \cdot \text{HPBW}}{\lambda_0 T_S} \cdot \left( t - \frac{T_S}{2} \right), \quad t \in [0, T_S], \quad (5)$$

where  $d_0$  is the distance between the adjacent antenna elements as shown in Figure 2,  $\lambda_0$  is the average wavelength applied to the antenna array, and, for ease in presentation, we have assumed a mobile located at  $\phi_0 \approx \pi/2$ . After each time duration  $T_S$ ,  $\theta(t, \phi)$  returns to its  $t = 0$  value (returning the beam pattern to its original position) and  $\theta(t, \phi)$  then recreates an identical spatial movement over the next  $T_S$  duration. Assuming a small HPBW, (5) can be simplified to

$$\theta(t) \cong \kappa \cdot \frac{2\pi d_0 \cdot \text{HPBW}}{\lambda_0 T_S} \cdot \left( t - \frac{T_S}{2} \right), \quad t \in [0, T_S]. \quad (6)$$

### Criterion 2: Independent fades

Movement of the antenna array beam pattern based on the time-varying antenna array phases in (5) or (6) results in a time varying channel, and rate of variation of the channel is measured by coherence time,  $T_C$ . Coherence time, in turn, determines the number of independent fades over duration  $T_S$ . Computation of the coherence time requires modeling a

channel in the presence of a moving beam pattern. In [5, 6] a linear time-varying impulse response model is introduced to characterize the channel with a beam pattern oscillation based on (6). The channel impulse response was characterized using the so-called geometric-based stochastic channel model (GSCM) [14, 15]. Simulation results with the antenna array control parameter restricted to  $0.0005 < \kappa < 0.05$  demonstrated that the channel coherence time due to beam pattern movement leads to an available diversity gain of up to  $L \approx 7$  [5, 6].

Specifically, based on our earlier work in [5, 6], we demonstrated how a 7-fold time diversity benefit is achieved (by beam pattern movement) in a mid-sized city center assuming three scatterers with an average size of 20 m in every 1000 m<sup>2</sup>,  $\kappa = 0.05$  (i.e., beam pattern movement corresponds to 5% of the HPBW), BS-mobile distance  $x_0 = 1000$  m, and HPBW = 0.3 radian (i.e., HPBW  $\approx 17^\circ$ ). (To create HPBW  $\approx 17^\circ$  requires  $M = 6$  antenna array elements—hence, the number of elements may be a small value.)

### 3. RECEIVER DESIGN

User  $k$ 's signal, input to the antenna array of Figure 2, corresponds to (using (1) and (2))

$$s_k(t) = \text{Re} \left\{ b_k \cdot e^{j2\pi f_0 t} \cdot \sum_{i=0}^{N-1} c_k^i \cdot h(t - iT_C) \cdot g(t) \right\}, \quad (7)$$

which, using  $h(t)$  in (3), leads to

$$s_k(t) = \text{Re} \left\{ b_k \cdot e^{j2\pi f_0 t} \cdot \sum_{i=0}^{N-1} c_k^i \cdot \sum_{n=0}^{N-1} e^{j2\pi n \Delta f (t - iT_C)} \cdot g(t) \right\},$$

$$s_k(t) = b_k \cdot g(t) \cdot \left( \sum_{i=0}^{N-1} c_k^i \cdot \sum_{n=0}^{N-1} \cos(2\pi(f_0 + n\Delta f) \cdot t - \beta_n^i) \right),$$

$$t \in [0, T_S], \quad (8)$$

where  $\beta_n^i = 2\pi i \cdot n\Delta f T_C$ . The output of the  $m$ th element of the antenna array, after application of phase offset  $m\theta(t)$ , is simply

$$s_k^m(t) = b_k \cdot g(t) \cdot \left( \sum_{i=0}^{N-1} c_k^i \cdot \sum_{n=0}^{N-1} \cos(2\pi(f_0 + n\Delta f) \cdot t - \beta_n^i + m\theta(t)) \right). \quad (9)$$

The presence of  $\theta(t)$  creates a frequency offset; however, with  $\theta(t)$  selected according to (6) (and considering  $\kappa < 0.05$ ), and assuming the distance between the antenna elements  $d = \lambda/2$ , antenna HPBW = 0.3, and  $M = 6$  elements, it is easily shown that the frequency offset induced by  $\theta(t)$  is less than 5% of a 1 MHz bandwidth. Hence, we ignore this frequency offset in our presentation. The total downlink transmitted signal, considering all antenna elements (all  $m$ )

and all users (all  $k$ ) in  $t \in [0, T_S]$ , is (from (9))

$$s(t) = \sum_{k=1}^K b_k \cdot g(t) \cdot \left( \sum_{i=0}^{N-1} c_k^i \cdot \sum_{n=0}^{N-1} \frac{1}{M} \cdot \sum_{m=0}^{M-1} \cos(2\pi(f_0 + n\Delta f) \cdot t - \beta_n^i + m\theta(t)) \right), \quad t \in [0, T_S], \quad (10)$$

where  $1/M$  is a normalization factor compensating for transmission over  $M$  array elements.

At the receiver side, the transmit diversity (due to antenna array movement generated by  $\theta(t)$ ) corresponds to an  $L$ -fold time diversity. Hence, the received signal in duration  $[0, T_S]$  can be divided into time slots  $[lT_S/L, (l+1)T_S/L]$ , where  $l \in [0, L-1]$ , and each time slot contains a signal with an independent frequency-selective fade. The received signal corresponds to

$$r^l(t) = \sum_{k=1}^K b_k \cdot g(t) \cdot \left( \sum_{i=0}^{N-1} c_k^i \cdot \sum_{n=0}^{N-1} \alpha_n^l \cdot AF(t, \phi) \cdot \cos(2\pi(f_0 + n\Delta f) \cdot t - \beta_n^i + \frac{M-1}{2}\gamma(t, \phi) + \xi_n^l) \right) + n^l(t),$$

$$t \in \left[ \frac{lT_S}{L}, \frac{(l+1)T_S}{L} \right], \quad l = 0, 1, \dots, L-1. \quad (11)$$

First, we explain the  $\alpha_n^l$  and  $\xi_n^l$  terms in (11). Because each chip shape is a multicarrier signal, the frequency selectivity of the fade is resolved by the multicarrier components (as in OFDM [16] and MC-CDMA [17]). That is, each carrier  $n$ ,  $n \in \{0, 1, \dots, N-1\}$ , that makes up the chip shape, experiences a unique flat fade.  $\alpha_n^l$  is the fade on the  $n$ th carrier in the  $l$ th time slot (due to fading) and  $\xi_n^l$  is the phase offset in the  $n$ th carrier and  $l$ th time slot (due to fading) (hereafter, this phase is assumed to be tracked and removed). The fades  $\alpha_n^l$  over the subcarriers that make up each CI chip, that is,  $\{\alpha_0^l, \alpha_1^l, \dots, \alpha_{N-1}^l\}$  are correlated Rayleigh random variables with correlation coefficient between the  $p$  subcarrier fade and the  $q$  subcarrier fade characterized by [18]

$$\rho_{p,q} = \frac{1}{1 + ((p-q) \cdot (\Delta f / (\Delta f)_C))^2}, \quad (12)$$

where  $(\Delta f)_C$  is the coherence bandwidth of the channel. In addition, in (11), the  $n^l(t)$  term represents the white Gaussian noise in the  $l$ th time slot, and the antenna array introduces the phase offset  $\gamma(t, \phi)$  corresponding to

$$\gamma(t, \phi) = \left( \frac{2\pi d_0}{\lambda_0} \right) \cdot \cos \phi + \theta(t). \quad (13)$$

Here,  $(2\pi d_0 / \lambda_0) \cdot \cos \phi$  represents the phase offset due to the difference in distance between antenna array elements and the mobile (assuming the smart antenna array is mounted horizontally). Moreover, in (11), the antenna array also

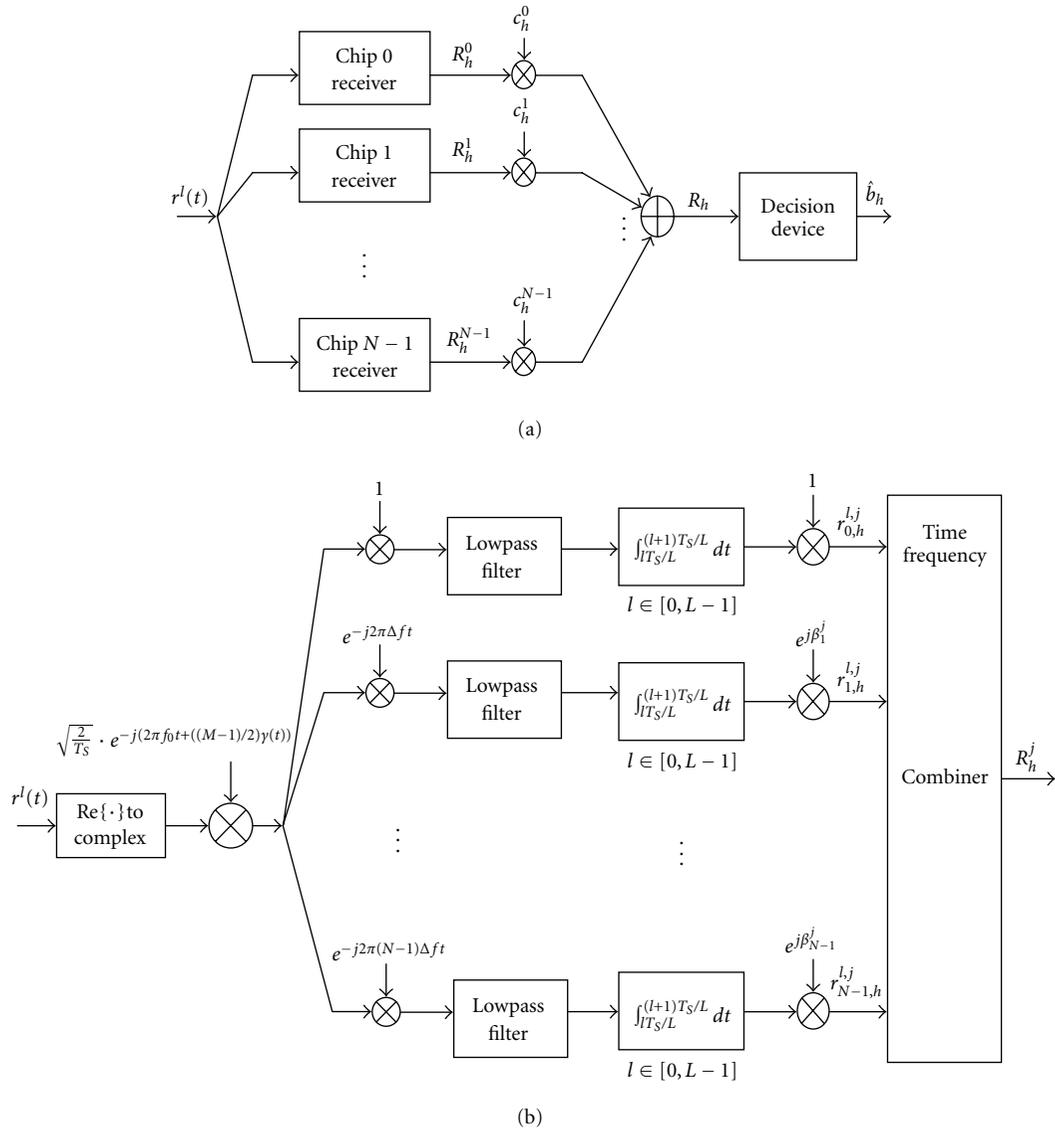


FIGURE 3: User  $h$  (a) mobile receiver and (b) chip  $j$  receiver.

introduces the normalized gain  $AF(t, \phi)$ , corresponding to

$$AF(t, \phi) = \frac{1}{M} \cdot \left[ \frac{\sin((M/2)\gamma(t, \phi))}{\sin((1/2)\gamma(t, \phi))} \right]. \quad (14)$$

Assuming the mobile is located on the antenna beam main axis ( $\phi_0 = \pi/2$ ), and assuming a small beam pattern movement ( $\kappa < 0.05$ ), we can simplify (11) by assuming  $\gamma(t, \phi) \cong \gamma(t) = \theta(t)$  and  $AF(t, \phi) \cong 1$  at the mobile's position for all  $t \in [0, T_s]$ .

The CI/DS-CDMA receiver is shown in Figures 3a and 3b, where Figure 3a shows the overall receiver structure for user  $h$ , and Figure 3b details the block entitled “chip  $j$  receiver” (in Figure 3a). In other words, the receiver operates as follows: first, the received signal is processed through a total of  $N$  chip receivers, where chip  $j$ 's receiver (a) decomposes its

chip into  $N$  carrier components, and (b) recombines across the carrier components to recreate the chip while achieving a frequency diversity benefit. In addition, because a time diversity benefit is available (via transmit diversity), chip  $j$ 's receiver also (c) combines across time components to recreate the chip with a frequency-time diversity gain. Next, once each chip is recreated with an enhanced diversity benefit, the receiver of Figure 3a performs a combining across chips in a usual DS-CDMA manner to eliminate interfering users' signals.

Mathematically, the receiver operates as follows. First, the received signal enters chip  $j$ 's receiver. Here, the carrier component is removed from the incoming signal, and the signal is split into  $N$  branches (one per carrier component). On the  $n$ th branch, the  $n$ th carrier is returned to baseband and separated from other carriers by application of a lowpass filter.

(To ensure perfect separability of the carriers (that make up the  $j$ th chip) via filtering, we select  $\Delta f = 2/T_S$ .)

Each baseband signal (one per carrier) is integrated over each interval over which the fade is constant, that is, over  $t \in [lT_S/L, (l+1)T_S/L]$ ,  $l \in [0, L-1]$ . After applying phase offsets to the  $N$  frequency components (phase offsets corresponding to the delay  $jT_C$ , separating the  $j$ th chip from other chips), the signal in the  $j$ th chip's receiver, for each carrier  $n$ ,  $n \in [0, N-1]$ , and time interval  $l$ ,  $l \in [0, L-1]$ , corresponds to

$$\begin{aligned} r_{n,h}^{l,j} &= \frac{1}{L} \sqrt{E_S} \cdot \alpha_n^l \cdot b_h \cdot c_h^j + \frac{1}{L} \sqrt{E_S} \cdot \alpha_n^l \cdot b_h \cdot \sum_{\substack{i=0 \\ i \neq j}}^{N-1} c_h^i \cdot \rho_n^{i,j} \\ &+ \frac{1}{L} \sqrt{E_S} \cdot \alpha_n^l \cdot \sum_{\substack{k=1 \\ k \neq h}}^K c_k^j \cdot b_k \\ &+ \frac{1}{L} \sqrt{E_S} \cdot \alpha_n^l \cdot \left( \sum_{\substack{k=1 \\ k \neq h}}^K b_k \cdot \sum_{\substack{i=0 \\ i \neq j}}^{N-1} c_k^i \cdot \rho_n^{i,j} \right) + n_{n,h}^{l,j}, \end{aligned} \quad (15)$$

where  $\sqrt{E_S} = \sqrt{T_S/2}$ . In (15), the first term represents the desired  $n$ th frequency component and  $l$ th time component of chip  $j$  for the desired user (user  $h$ ); the second term is the interchip interference due to other chips from the same user (where  $\rho_n^{i,j} = \cos(\beta_n^i - \beta_n^j)$  is the correlation between the  $i$ th chip and the  $j$ th chip in carrier  $n$ ); the third term is the interference due to the same chip from other users, and the fourth term represents the interference from different chips of different users. Moreover,  $n_{n,h}^{l,j}$  is a zero mean Gaussian random variable with variance  $(N_0/2)/(N^2 \cdot L)$ , independent across different carriers  $n$  and different time slots  $l$ , but correlated across chips, with correlation  $\rho_n^{i,j}$  between the  $i$ th chip noise and the  $j$ th chip noise.

It is also important to note the factor of  $1/L$  in the first term (desired term) in (15), which is a direct result of the division of the received signal interval into  $L$  partitions (to create  $L$ -fold time diversity) (i.e., a direct result of the  $L$ -fold oversampling strategy).

Following the decomposition of the  $j$ th chip into its time and frequency components, each with a unique fade, a linear combining strategy is employed to recreate the  $j$ th chip with a joint time-frequency diversity benefit. Using the linear combining scheme discussed in the next paragraph, we combine the  $r_{n,h}^{l,j}$  over time components ( $l$ ) and frequency components ( $n$ ) (with  $L \times N$  diversity components,  $L$  over time and  $N$  over frequency) to simultaneously reduce the interchip interference and the noise, and achieve high diversity gains. This leads to the output,  $R_h^j$ , which (referring to Figure 3a) is combined across the  $N$  chips in the usual DS-CDMA manner to eliminate other users interference (term 3 in (15)). The chip combiner output for user  $h$  corresponds to

$$R_h = \sum_{j=0}^{N-1} c_h^j \cdot R_h^j. \quad (16)$$

This  $R_h$  term enters a hard decision device which generates the final decision,  $\hat{b}_h$ . The time-frequency combiner recreating the  $j$ th chip from its time-frequency components (in Figure 3b) is designed using EGC-MMSEC, that is, EGC across  $L$  time components followed by MMSEC across  $N$  frequency components. Applying EGC in time, then the Wiener filter principle [19] to determine the MMSEC across carriers, the decision variable corresponds to

$$R_h^j = \sum_{n=0}^{N-1} \frac{\sqrt{E_S} \cdot \alpha_n/L}{P \cdot K \cdot (\alpha_n)^2 + N_{0n}'/2} \cdot r_{n,h}^j, \quad (17)$$

where

$$\begin{aligned} r_{n,h}^j &= \sum_{l=0}^{L-1} r_{n,h}^{l,j} = \frac{1}{L} \sqrt{E_S} \cdot b_h \cdot c_h^j \cdot \alpha_n \\ &+ \frac{1}{L} \sqrt{E_S} \cdot b_h \cdot \alpha_n \cdot \sum_{\substack{i=0 \\ i \neq j}}^{N-1} c_h^i \cdot \rho_n^{i,j} \\ &+ \frac{1}{L} \sqrt{E_S} \cdot \alpha_n \cdot \sum_{\substack{k=1 \\ k \neq h}}^K c_k^j \cdot b_k \end{aligned} \quad (18)$$

$$\begin{aligned} &+ \frac{1}{L} \sqrt{E_S} \cdot \alpha_n \cdot \left( \sum_{\substack{k=1 \\ k \neq h}}^K b_k \cdot \sum_{\substack{i=0 \\ i \neq j}}^{N-1} c_k^i \cdot \rho_n^{i,j} \right) + n_{n,h}^j, \\ \alpha_n &= \sum_{l=0}^{L-1} \alpha_n^l, \end{aligned} \quad (19)$$

$$n_{n,h}^j = \sum_{l=0}^{L-1} n_{n,h}^{l,j}, \quad (20)$$

$$P = \frac{E_S}{L^2} \cdot \begin{cases} N & \text{for } n = 0 \text{ or } \frac{N}{2}, \\ \frac{N}{2} & \text{else,} \end{cases} \quad (21)$$

and  $N_{0n}'/2$  is the noise variance of  $n_{n,h}^j$  in (20), that is,  $(N_0/2)/N^2$ .

#### 4. SIMULATED PERFORMANCE

For simulation purposes, we consider (1) CI/DS-CDMA with a processing gain of  $N = 32$ ; (2) each CI/DS-CDMA chip is composed of  $N = 32$  carriers (see (3)); (3) the CI/DS-CDMA system is fully loaded with  $K = 32$  orthogonal users employing Hadamard-Walsh codes; (4) the frequency selectivity of the channel results in 4-fold frequency diversity over the entire bandwidth, that is,  $(\Delta f)_C/BW = 0.25$ , and (5) beam pattern movement results in  $L = 7$  independent fades in the duration  $T_S$  (see Section 2 and [5]).

In Figure 4, a typical simulation result is provided for the proposed CI/DS-CDMA—oscillating beam antenna-array merger. The simulation results in Figure 4 are compared with those of CI/DS-CDMA with a conventional smart antenna at the BS. Here, MMSEC is applied to the subcarriers of the received CI/DS-CDMA signal [9, 10, 11, 12, 13]. It is observed that the introduction of a smart antenna array with

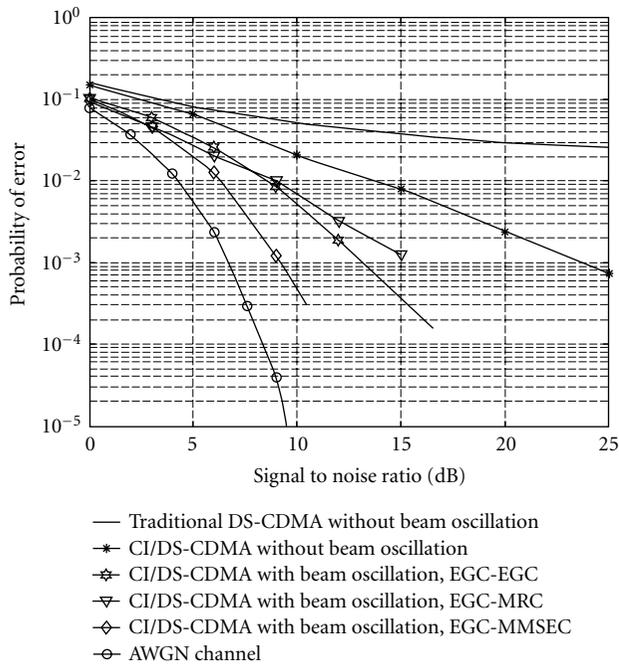


FIGURE 4: Simulation results.

beam-pattern oscillation at the BS introduces an improvement of more than 14 dB at a probability of error of  $10^{-3}$  (at the mobile via EGC-MMSEC combining technique) compared to CI/DS-CDMA with traditional smart antenna arrays. When compared to traditional DS-CDMA with RAKE reception (e.g., [8]) combined with a conventional smart antenna array, even larger performance gains are achieved, as shown in Figure 4. In Figure 4, EGC-MMSEC technique is compared with EGC-EGC and EGC-MRC (maximal ratio combining) schemes. EGC-MMSEC achieves excellent performance relative to other combining options.

## 5. DISCUSSION AND CONCLUSION

CI/DS-CDMA signals are sent via a single antenna array at the BS, and received by a single antenna at the mobile station. The phase shifts introduced to the BS antenna array elements are designated to control the antenna pattern movement (oscillation) such that it achieves directionality and transmit diversity.

Receivers employ diversity combining in the frequency and time domains, and significant performance improvement is shown when compared to the CI/DS-CDMA system with a conventional antenna array. This performance gain highlights the significance of small beam pattern movement in smart-antenna-array CDMA systems. This performance leads to high network capacity in terms of number of users.

## ACKNOWLEDGMENT

This work was supported by NASA Phase II SBIR Grant: "Development of a Wireless Communication System to Support Airport Surface Operations."

## REFERENCES

- [1] A. Hiroike, F. Adachi, and N. Nakajima, "Combined effects of phase sweeping transmitter diversity and channel coding," *IEEE Trans. Veh. Technol.*, vol. 41, no. 2, pp. 170–176, 1992.
- [2] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 744–765, 1998.
- [3] S. A. Zekavat, C. R. Nassar, and S. Shattil, "Combined directionality and transmit diversity via smart antenna spatial sweeping," in *Proc. 38th Annual Allerton Conference on Communication, Control, and Computing*, pp. 203–211, University of Illinois at Urbana-Champaign, Ill, USA, October 2000.
- [4] S. A. Zekavat, C. R. Nassar, and S. Shattil, "Oscillating-beam smart antenna arrays and multicarrier systems: achieving transmit diversity, frequency diversity, and directionality," *IEEE Trans. Veh. Technol.*, vol. 51, no. 5, pp. 1030–1039, 2002.
- [5] S. A. Zekavat and C. R. Nassar, "Geometric-based stochastic channel modeling for adaptive antennas with oscillating beam patterns," in *Proc. 12th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 1, pp. 130–134, San Diego, Calif, USA, September 2001.
- [6] S. A. Zekavat and C. R. Nassar, "Smart antenna arrays with oscillating beam patterns: characterization of transmit diversity in semi-elliptic coverage," *IEEE Trans. Communications*, vol. 50, no. 10, pp. 1549–1556, 2002.
- [7] A. J. Viterbi, *CDMA. Principles of Spread Spectrum Communication*, Addison-Wesley Publishing, Reading, Mass, USA, 1995.
- [8] U. S. Goni and A. M. D. Turkmani, "BER performance of a direct-sequence CDMA system in multipath fading mobile radio channels with Rake reception," in *IEEE 44th Vehicular Technology Conference*, vol. 2, pp. 747–751, Stockholm, Sweden, June 1994.
- [9] C. R. Nassar, B. Natarajan, and Z. Wu, "Multi-carrier technology platform for wireless communications. Part 1: High-performance, high-throughput TDMA and DS-CDMA via multi-carrier implementations," to appear in *Wireless Communications and Mobile Computing*.
- [10] Z. Wu, C. R. Nassar, and S. Shattil, "High performance DS-CDMA via carrier interferometry," in *Wireless 2001*, pp. 564–569, Calgary, Alberta, July 2001.
- [11] Z. Wu, C. R. Nassar, and S. Shattil, "Chip shaping advances for high capacity DS-CDMA," in *2001 International Conference on Third Generation Wireless and Beyond*, pp. 928–932, San Francisco, Calif, May–June 2001.
- [12] Z. Wu and C. R. Nassar, "MMSE frequency combining for CI/DS-CDMA," in *IEEE Radio and Wireless Conference*, pp. 103–106, Denver, Colo, USA, 2000.
- [13] C. R. Nassar and Z. Wu, "High performance broadband DS-CDMA via carrier interferometry chip shaping," in *International Symposium on Advanced Radio Technologies*, Boulder, Colo, USA, September 2000.
- [14] R. B. Ertel, P. Cardieri, K. W. Sowerby, T. S. Rappaport, and J. H. Reed, "Overview of spatial channel models for antenna array communication systems," *IEEE Personal Communications*, vol. 5, no. 1, pp. 10–22, 1998.
- [15] A. F. Molisch, A. Kuchar, J. Laurila, K. Hugl, and E. Bonek, "Efficient implementation of a geometry-based directional model for mobile radio channels," in *IEEE 50th Vehicular Technology Conference*, vol. 3, pp. 1449–1453, Amsterdam, Netherlands, 1999.
- [16] R. van Nee and P. Ramjee, *OFDM for Wireless Multimedia Communications*, Artech House Publishers, Boston, Mass, USA, 2000.

- [17] S. Hara and R. Prasad, "Overview of multicarrier CDMA," *IEEE Communications Magazine*, vol. 35, no. 12, pp. 126–133, 1997.
- [18] W. C. Jakes, *Microwave Mobile Communications*, John Wiley & Sons, New York, NY, USA, 1974.
- [19] S. Haykin, *Adaptive filter theory*, Prentice-Hall, Englewood Cliffs, NJ, USA, 2nd edition, 1991.

**Seyed Alireza Zekavat** received his Ph.D. degree in electrical and computer engineering from Colorado State University, Colo, USA in 2002. From 1993 to 1998 he was with Civil Aviation College of Technology, Tehran, Iran. Since August 2002 he has been with Michigan Technological University (MTU), Houghton, Mich, USA. His research interests include wireless communications, radar theory, statistical modeling, adaptive beam forming, and neural networks. He is the founder and the director of the laboratory for wireless communication research in MTU. He has published about 40 journal and conference papers and has coauthored the book *Multi-Carrier Technologies for Wireless Communications*.



**Carl R. Nassar** received his B.S., M.S., and Ph.D. degrees from McGill University in 1989, 1990, and 1997, respectively. Between his M.S. and Ph.D. degrees, Dr. Nassar worked for a time as a design engineer at CAE Electronics in Montreal, Canada. Upon completion of his Ph.D., he spent a year as an Assistant Professor at McGill University. Soon thereafter, he headed for the hills of Colorado, where he has been an assistant professor for the past five years. Dr. Nassar is the Director of Colorado State University's RAWCom (Research in Advanced Wireless) laboratory, a position which enables him to pursue his interest in wireless telecommunications. With funding from the NSF, NASA, industry, and the State of Colorado, Dr. Nassar's research focuses on the design of high network capacity, high QOS multiple-access technologies. In particular, Dr. Nassar focuses on advances in multicarrier technologies such as MC-CDMA and OFDM, and seeks a common multicarrier platform for all wireless multiple access based on his proposed CI technology. Dr. Nassar's work has been published in over 90 international conference proceedings and journal articles, and he is the author of two books: *Telecommunications Demystified*, a friendly engineering look at telecommunication systems, and *Multi-Carrier Technologies for Future Generation Wireless*.



**Steve Shattil** is the Chief Scientist at CIAN Systems Inc., where he is leading development in carrier interferometry and other coding technologies. Mr. Shattil has over 15-year experience in the wireless industry. Prior to founding CIAN, he was Founder and Chief Technical Officer of Genesis Telecom. Genesis invented highly bandwidth-efficient antenna array technologies for broadband wireless communications. Over the last decade, Mr. Shattil has led the development of baseband and RF processors for broadband products, and is world renowned in both academia and industry for his innovation, leadership, and



the advancement of wireless communications. Mr. Shattil has authored dozens of domestic and international patents. He holds an M.E. in electrical engineering from University of Colorado where he pioneered advances in high-data-rate signal processing. Mr. Shattil also holds an M.S. in physics from Colorado School of Mines, where he advanced the field of laser physics and built the first laser that generates carrier interferometry signals. Mr. Shattil holds a B.S. in physics from Rensselaer Polytechnic Institute.

# ADAM: A Realistic Implementation for a W-CDMA Smart Antenna

## Ramón Martínez Rodríguez-Osorio

*Department of Signals, Systems and Radiocommunications, Polytechnic University of Madrid, 28040 Madrid, Spain*  
Email: ramon@gr.ssr.upm.es

## Laura García García

*Department of Signals, Systems and Radiocommunications, Polytechnic University of Madrid, 28040 Madrid, Spain*  
Email: lgg@gr.ssr.upm.es

## Alberto Martínez Ollero

*Department of Signals, Systems and Radiocommunications, Polytechnic University of Madrid, 28040 Madrid, Spain*  
Email: alberto@gr.ssr.upm.es

## Francisco Javier García-Madrid Velázquez

*Department of Signals, Systems and Radiocommunications, Polytechnic University of Madrid, 28040 Madrid, Spain*  
Email: javiergmv@gr.ssr.upm.es

## Leandro de Haro Ariet

*Department of Signals, Systems and Radiocommunications, Polytechnic University of Madrid, 28040 Madrid, Spain*  
Email: leandro@gr.ssr.upm.es

## Miguel Calvo Ramón

*Department of Signals, Systems and Radiocommunications, Polytechnic University of Madrid, 28040 Madrid, Spain*  
Email: miguel@gr.ssr.upm.es

*Received 30 May 2003; Revised 28 November 2003*

Adaptive-type smart antennas do not usually operate on the deployed universal mobile telecommunication system (UMTS) scenarios, although UTRA (UMTS terrestrial radio access) foresees their operation and they would improve capacity especially in mixed-service environments. This paper describes the implementation of a software radio-based version of an adaptive antenna, named ADAM, that can be used with any standard Node B, both in the up- and downlinks. This transparent operational feature has been made possible by the partial cancelation algorithm applied in the uplink by means of a common beamforming vector. Firstly, a general description of the system as well as the theory of its operation are described. Next, the hardware architecture is presented, showing the real implementation. Also a complete software description is done. Finally, results are presented, obtained from both simulation and real implementation, showing the improvement obtained with the adaptive antenna as compared with a typical sectored one. Performance results obtained in the initial tests show that ADAM prototype provides an SINR increase of 12.5 and 6.5 dB over a conventional sectored antenna in the uplink and downlink, respectively. System-level simulation results are presented, showing the throughput increase obtained with ADAM. These findings provide evidence of the capacity improvement achieved with the ADAM prototype.

**Keywords and phrases:** smart antenna prototype, beamforming, wireless communications, synchronization, DSP, UMTS.

## 1. INTRODUCTION

The smart antenna concept is applied to several kinds of antenna arrays. Phased arrays, switched multibeam antennas, and adaptive array antennas are usually included under the smart antenna concept with the only condition of includ-

ing the possibility to somehow control the radiation pattern. Great advantages have been reported for the smart antenna implementation in base stations for mobile telephone communications, but this kind of antenna has not been extensively applied to those systems yet.

If capabilities of phased array, switched-beam array, and adaptive array antennas are compared, the last type shows considerable advantages over the others [1]. Not only can adaptive arrays improve antenna gain in the user direction but they can also cancel interferences inside the angular range of control. This ability implies an increase of the signal-to-interference-plus-noise ratio (SINR) for each user. For code division multiple access (CDMA) systems, an increase of sector capacity is obtained for those cells with base stations equipped with smart antennas. The capacity increase is higher in cells with high interference levels, usually produced by high bit rate users.

Adaptive antenna systems can be implemented using a space or time reference-based algorithm. In spatial reference adaptive arrays, interference directions are computed and the array weights are obtained to cancel or minimize them. In time reference adaptive arrays, time series from the input signal at each array element are processed to form the array vector of weights. The array factor implemented for each user increases the SINR and improves the energy per bit to noise density ratio ( $E_b/N_0$ ) due to the correlation of the received signals. This strategy is appropriate for CDMA signals since a time reference can be obtained applying the user code. In the particular case of universal mobile telecommunication system (UMTS), the physical layer has been designed to work with adaptive antennas both in uplink and downlink [2].

A significant research effort has taken place in the last years to introduce smart antenna systems in cellular scenarios. However, the deployment of these antenna systems has not become a reality yet due to their cost and complexity. In practice, only switched-beam antennas for second generation (2G) systems have been commercially deployed [3, 4, 5, 6, 7, 8]. This is due to the complexity of adaptive antennas in third generation (3G) systems. In contrast to 2G systems, where beamforming can be done in radio frequency (RF), beamforming in 3G must be applied after demodulating the CDMA signal so that adaptive antenna functions need to be integrated into the (digital and intermediate frequency (IF)) baseband-processing sections of the base station. Therefore, the implementation of adaptive antennas in 3G base stations requires a reconfigurable and flexible architecture. These features can be obtained using software radio platforms [9, 10, 11].

Many of the existing smart antenna solutions for 3G have been developed for a unique base station equipment manufacturer [12, 13]. This fact makes the deployment of smart antenna systems unfeasible for mobile communications operators due to the high associated cost and manufacturer dependency. A plug and play smart antenna solution, appropriate for any base station from any manufacturer, has not been developed yet.

This paper details a practical implementation of an adaptive plug and play smart antenna for 3G mobile communication systems based on wideband-CDMA (W-CDMA) like UMTS [14, 15]. Unlike currently existing adaptive antenna arrays, the implementation described here implies an easy deployment over any base station, not only on those specifically developed to be used with smart antennas [16]. ADAM

stands for “adaptive antenna for multioperator scenarios,” as it can be connected to any base station site even shared by several operators.

As a plug and play functionality is demanded, the UMTS signals are demodulated and remodulated again, allowing a direct connection between the smart antenna outputs and the base station inputs [16]. Due to this process, in the uplink, only those interferences common to the intracellular users and all the extracellular interferences are canceled. The relationship between the extracellular and intracellular interferences is called the extracellular interference factor  $F$  and has a value between 0.4 to 1.4 depending on the environment and the service [15]. This implies that more than 50% of the interferences are canceled on average as the common intracellular interferences should also be taken into account.

This antenna will take profit of hot spots, improving the capacity in the vicinity of high occupied cells. In these situations, mainly higher power external interferences from multimedia services are canceled by ADAM prototype, as it is demonstrated by simulation in this paper. In these situations, the antenna would help the cells in the vicinity of a hot spot to expand their coverage and to compensate the “cell breathing” of high occupied cells. Moreover, in mixed and asymmetric services scenarios, typical of 3G systems, ADAM will increase the capacity in terms of total throughput.

According to the software radio concept, the analog-to-digital converters (ADCs) and digital-to-analog converters (DACs) are located just before the analog RF-to-IF chains, hence working with IF signals instead of the typical baseband signal. This allows most of the system modules to be implemented in software, which is a great advantage with respect to pure hardware implementations because the system can be easily reconfigured and updated with more advanced versions. Therefore, a great flexibility is achieved with this structure.

The beamforming module has been implemented just before the W-CDMA modulation. In the uplink, classical beamforming algorithms have been adapted to the special extracellular cancellation scheme implemented [17, 18]. Although different beamforming algorithms can be used, the normalized least mean squares (NLMS) algorithm has been selected initially due to its reduced computational complexity. In the downlink, beamforming aims to cancel all intra- and extracellular interferences, thus a full cancellation algorithm has been selected.

Apart from NLMS, some tests have been done using the recursive least squares (RLS) algorithm in order to study the performance improvement obtained in the convergence speed and final SINR.

It is important to remark on the implementation of the synchronization algorithms in UMTS [19, 20, 21]. This problem has been solved using a two-step approach, initially doing a coarse synchronization that is followed by a continuous fine synchronization. The implemented algorithm has been intensively optimised.

As the smart antenna should be transparent for the base station, it should not implement the base stations physical procedures, such as power control and handover, which are



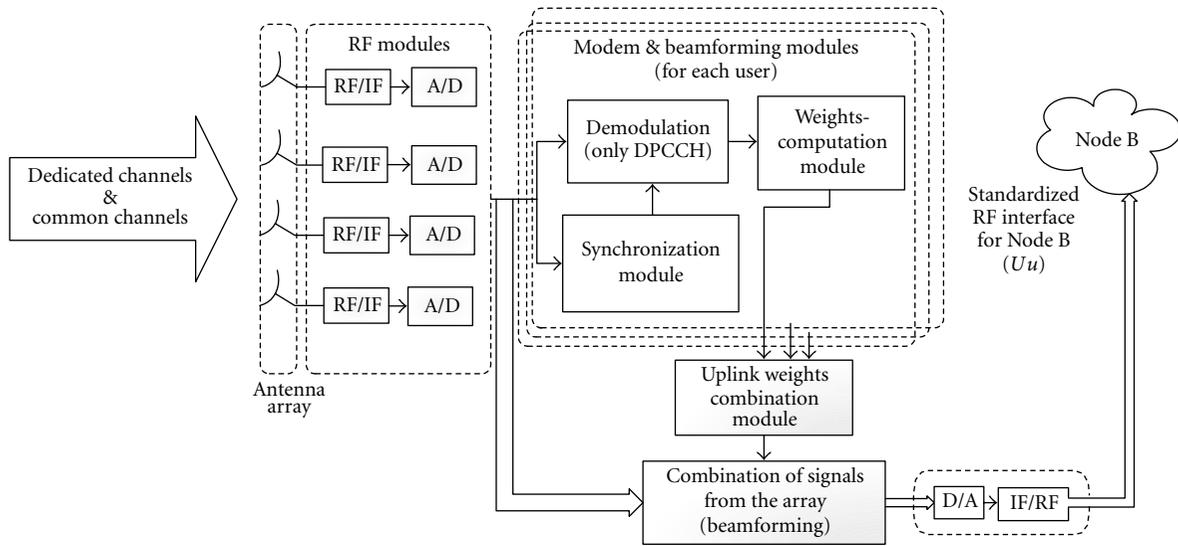


FIGURE 2: Uplink block diagram (1 polarization).

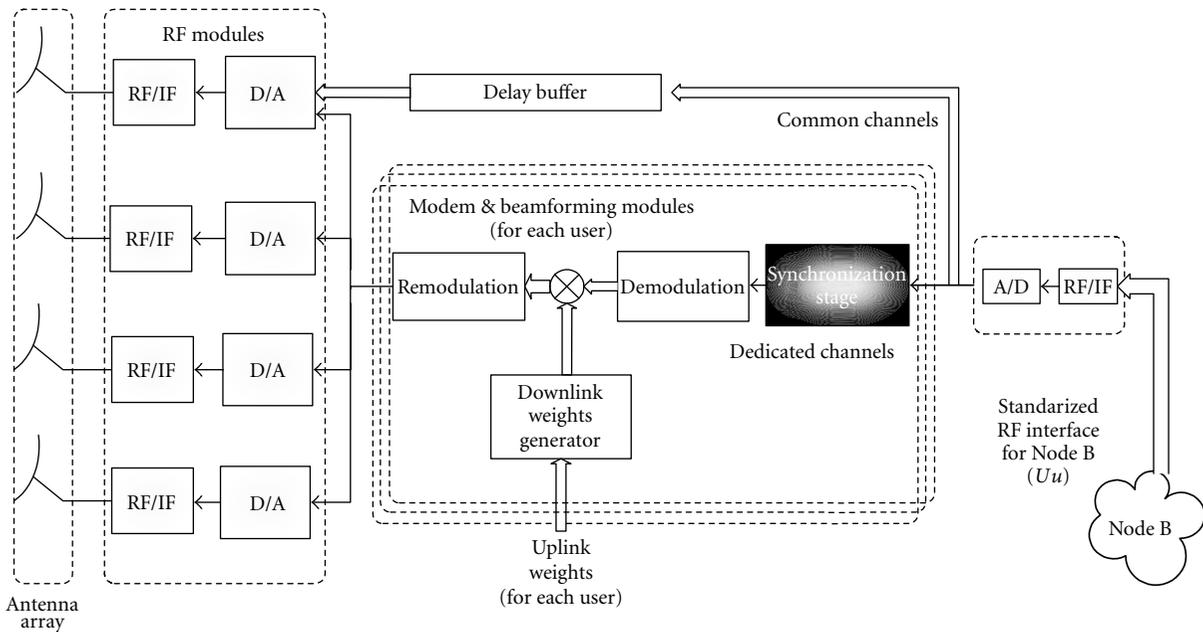


FIGURE 3: Downlink block diagram (1 polarization).

antenna array geometry, adaptive algorithm that controls the beamforming process, and propagation and interference environment. Those issues have been studied by simulation and are presented in Section 6. The ADAM array prototype uses four commercial sectored antennas for the UMTS band, each with a  $-3$  dB beamwidth of  $65^\circ$  and  $\pm 45^\circ$  polarization ports [22]. The individual antennas are put together in a uniform linear array structure, as shown in Figure 4. With this configuration, interelement separation is 15 cm (wide dimension of each sectored antenna), which is equivalent to  $0.975\lambda$  and  $1.070\lambda$  at the uplink and downlink frequencies, respectively.

### 3. HARDWARE ARCHITECTURE

The overall system proposed in this paper is formed by several hardware devices. Their characteristics, as well as the final selected hardware architecture, are presented below for both the uplink and downlink. Although a general description of the adaptive antenna has been made in Section 2, we focus here on the specific selected hardware solutions.

In the uplink, the received analog signal is downconverted by the RF-to-IF chains and digitalized. Afterwards, it is processed in the digital signal processing module, where

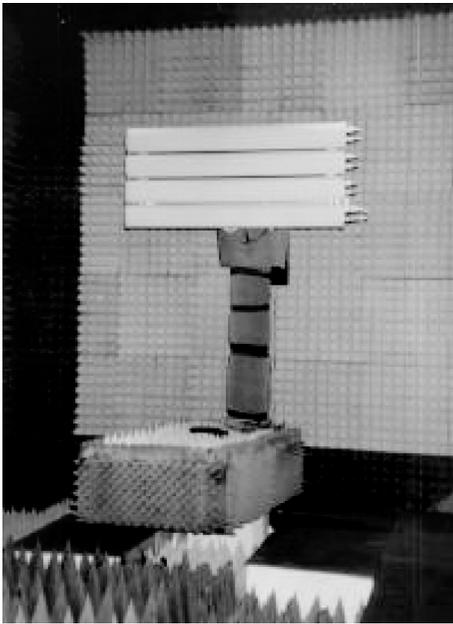


FIGURE 4: ADAM prototype: antenna array structure.

several digital signal processors (DSPs) work in parallel. The processed baseband signal is then analog-converted again, sent to an IF-to-RF chain and then to the Node-B RF input port. Conversely, the signal received from the Node-B output port follows similar steps in the downlink (RF-to-IF conversion, digitalization, digital processing, analog conversion, and RF upconversion), being finally transmitted through the antenna array.

Figure 5 shows a general architecture of the hardware implementation, where the blocks for the two polarizations are identical. The digital processing module, formed by several processors, is common to both polarizations. Analog RF-to-IF and IF-to-RF chains are not thoroughly explained here since it is out of the scope of this paper, mainly focused on the digital signal processing stages. Figure 6a shows the development system for software radio modules, whereas Figure 6b shows the test equipment.

Due to the software radio implementation, the IF frequency value offered to the rest of the modules must be carefully selected. A high IF would simplify the design of the analog chains, especially the filtering of the image frequency, but it would increment the processing capacity requirements. Also the current state of the art in ADCs and DACs should be taken into account since there is a tradeoff between the vertical resolution and sample frequency that can be achieved. With this in mind, an IF of 44 MHz was selected as a compromise solution.

Several aspects were taken into consideration to properly select the ADCs and DACs. The first one was the vertical resolution (or number of bits in conversion) required for this application. The quantification noise is lower with a high vertical resolution, but the available maximum sampling frequency decreases as the number of bits in conver-

sion are incremented. The recommended number of bits to use in a UMTS application is at least 12 [1]. As for the maximum sampling frequency  $f_{s,max}$ , it should be high enough to correctly receive or transmit the desired signal without loss of information. Also related to the  $f_{s,max}$ , we have to take into account the conversion bandwidth parameter. Finally, the dynamic range of the input voltage should be considered, especially in the analog chains design, to properly adjust its gain to the ADC input and DAC output levels.

After the ADC, the signal must be downconverted to baseband by means of an IQ demodulator. One possibility could be to implement it directly in a general DSP. But due to the high UMTS sampling rate, the required computational capacity to accomplish that operation would make the implementation unfeasible. Another interesting solution would be to use on-chip IQ demodulators or broadband downconverters, usually called *front-ends*. These devices can process the signal independently of the general DSPs, which can be used then to do the subsequent processing. The latter option has been chosen to implement the downconversion to baseband; so a general-purpose receiver has been selected from the commercially available devices. The selected receiver boards<sup>1</sup> consist of two broadband IQ demodulators plus two ADCs so that two identical receiver channels per receiver board are available [23]. The vertical resolution for the ADCs is 12 bits, and its maximum sampling frequency is 80 MHz. The ADC sampling frequency must be carefully selected. It has to be a multiple of the UMTS baseband signal rate 3.84 Mchip/s, multiplied by the number of samples per chip, which is  $N_{spc} = 4$  in this prototype. Neither 15.36 MHz nor 30.72 MHz can be used as sampling frequencies since it would cause aliasing in the sampled signal. On the other side, the ADC features restrict the possible sampling frequency to a maximum of 100 MHz. Thus,  $f_s = 61.44$  MHz has been chosen. Since  $f_s$  does not meet the Nyquist theorem ( $f_s$  is lower than  $2 \cdot IF$ ), the resulting signal is undersampled. This does not involve a loss of information because the signal is bandlimited to 5 MHz. A diagram of the main parts of one receiving channel is shown in Figure 7.

Similarly, an IQ modulator is required before each DAC. Also the *front-end* solution has been adopted here. The selected digital upconversion boards<sup>2</sup> provide two identical and independent broadband channels [23]. The DAC accepts 12-bit digital signal as input, and its maximum sampling frequency is 200 MHz. A block diagram of one channel can be seen in Figure 8.

Once the signal has been digitally converted and IQ demodulated, it has to be processed by the synchronization and beamforming modules, which are implemented in general-purpose digital processors. A few characteristics have been considered to select the DSPs that have been used to implement the software modules. The most important features are the arithmetic type, the clock rate and, in connection with

<sup>1</sup>Pentek 6235-board.

<sup>2</sup>Pentek 6229-board.

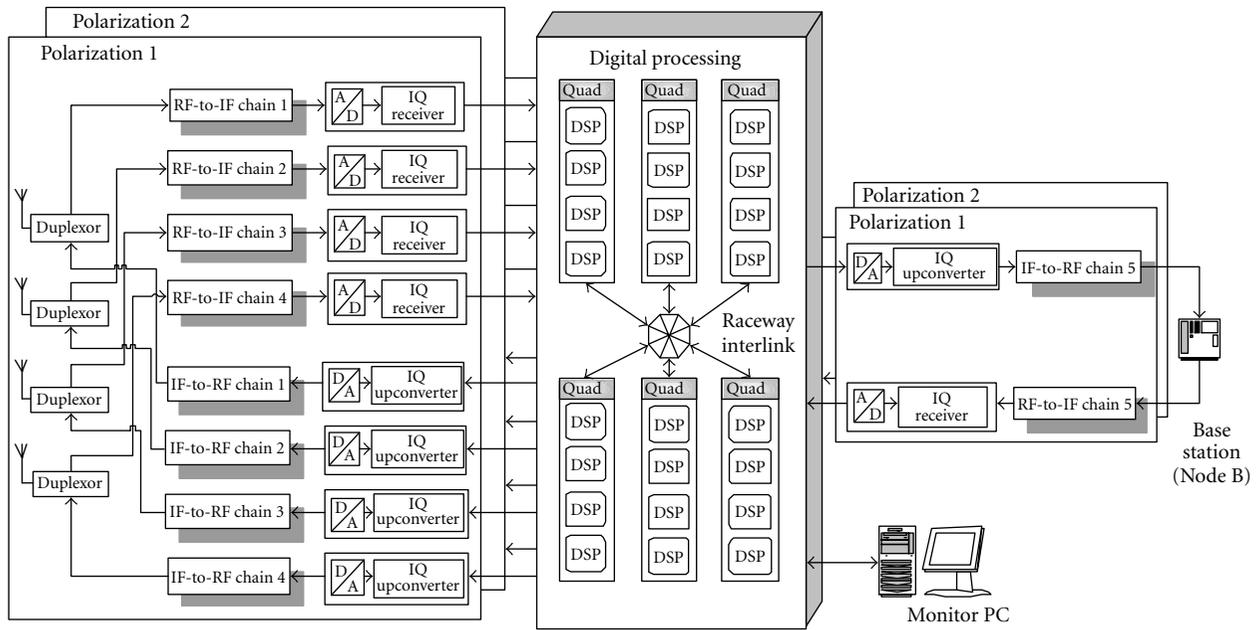


FIGURE 5: General hardware structure.

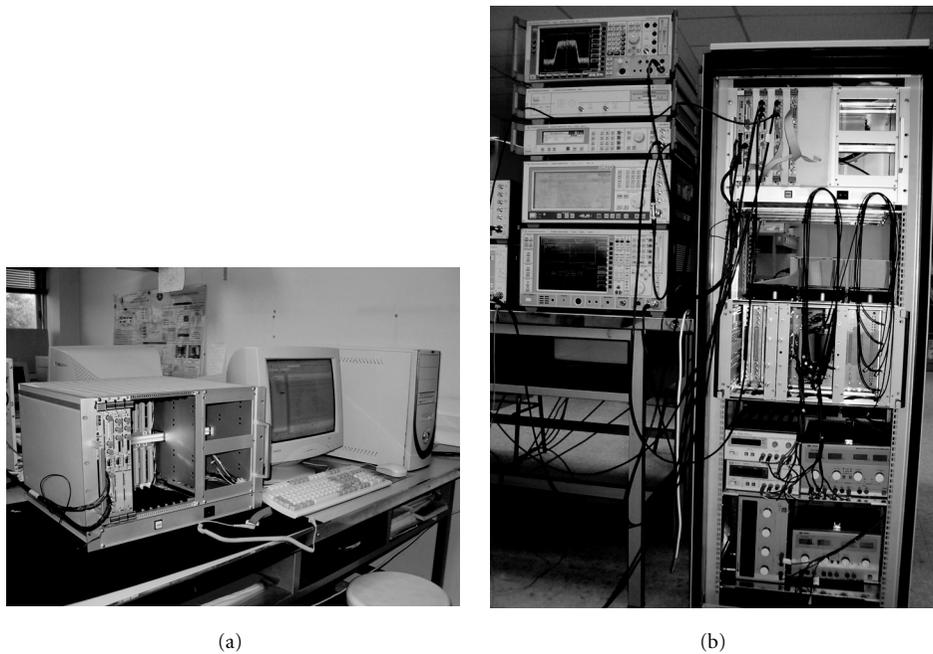


FIGURE 6: Hardware modules of ADAM prototype and test equipment. (a) Development system. (b) Measurement and test system.

this, the computational capacity. Fixed-point arithmetic is preferred instead of floating-point arithmetic since a higher speed processing for linear operations, like the ones required in this application, can be achieved. As regards the clock rate, the higher it is, the greater the number of instructions per second that can be executed, and the higher the computational capacity that can be obtained. In order to increase the computational capacity, a structure of various DSPs in paral-

lel can be used. The selected digital processing structure consists of six 4-DSP boards,<sup>3</sup> referred to as Quads [23]. Each Quad is formed by four 300-MHz fixed-point DSPs along with other interfaces between DSPs. Every Quad is capable of

<sup>3</sup>Pentek 4292-Quad VME board, with four Texas instrument TMS30C6203 processors.

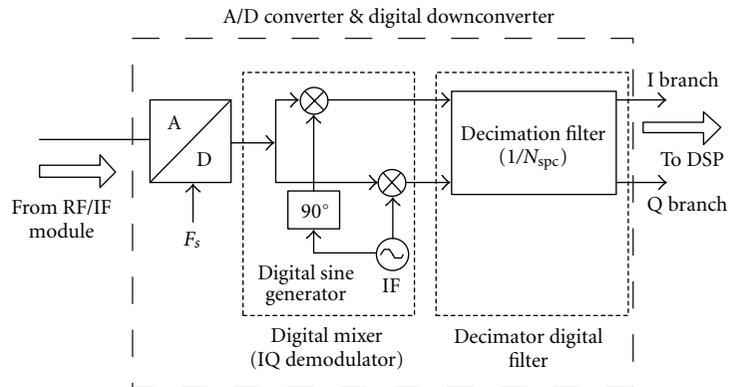


FIGURE 7: ADC and IQ demodulator.

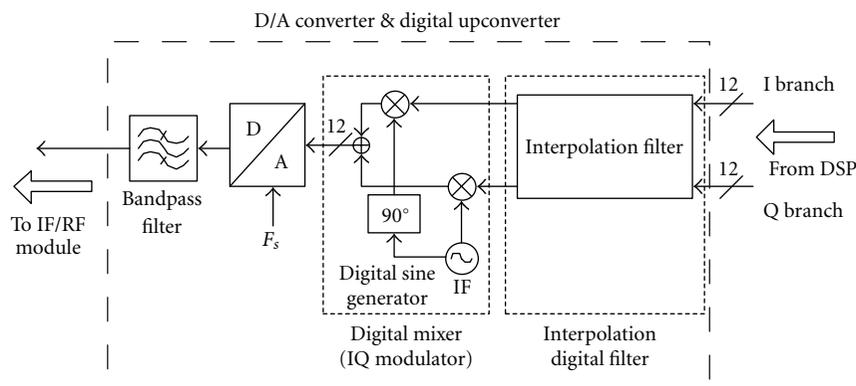


FIGURE 8: DAC and IQ modulator.

delivering a combined peak processing power of 9600 MIPS (millions of instructions per second).

In order to increase the data transfer rate between Quads, a high-speed data bus has been used<sup>4</sup> [23, 24]. This device is a high-speed backplane fabric capable of delivering 32-bit word transfers between versa module eurocard (VME) boards, such as the Quads presented previously. It provides multiple, simultaneous high-speed communication paths between DSPs which make the bus a valuable asset to real-time applications. The bus is capable of communicating up to eight VME boards at a data transfer rate of 267 MBps, which means an aggregate transfer rate up to 1068 GBps.

For monitoring tasks, a personal computer can be connected to the digital processing module to control the process and allow viewing of key variables and parameters.

#### 4. PRINCIPLES AND IMPLEMENTATION OF SOFTWARE RADIO MODULES

The software implementation has been divided into two main submodules: the set-up, synchronization and modem module, and the adaptive beamforming module. They are thoroughly explained below.

##### 4.1. Set-up, synchronization, and MODEM stages

As it is known [2], each physical channel in W-CDMA is spread combining two types of codes with complementary properties: orthogonal variable spreading factor (OVSF) channelization codes and scrambling codes (Gold codes, with excellent correlation properties). Basic information needed in a W-CDMA process is the used codes and, like any spread-spectrum technique, the timing reference [25]. The function of the *set-up* stage is to find the essential data needed before the demodulation process in uplink and downlink.

##### 4.1.1. Set-up procedure

Basic synchronization algorithms employed in the modem will be detailed in Section 4.1.2, and they are common for uplink and downlink. The main difference between uplink and downlink synchronization stages lies in which physical channels are used as reference signals.

In the downlink, all the physical channels (common signalling channels and dedicated user channels) use the same synchronization reference, that is, if the synchronization of one channel is known, the timing of the other channels is automatically known. The procedure to find the common timing reference for all downlink channels is called *cell search procedure*. Typically, cell search procedure is completed after three steps: slot synchronization, frame synchronization,

<sup>4</sup>Pentek 8251 Race++ interlink modules.

TABLE 3: Number of clock cycles and acquisition time for the coarse synchronization algorithm.

	Branches	Clock cycles/bit	Acquisition time (number of frames)
Serial search	1	3000	1
	2	5700	0.5
Parallel search	<b>3</b>	<b>8700</b>	<b>0.33</b>
	4	10800	0.25
	10	27300	0.1

code-group identification, and finally scrambling code identification [2]. Common signalling channels needed in this stage are the synchronization channel (SCH) and the P-CPICH.

The first and second steps use SCH codes. During the first step, the cell slot synchronization is acquired; it can be done by correlating the received signal of the base station with the primary SCH codes, employing the coarse synchronization algorithm, as it will be explained in Section 4.1.2.1. After the cell slot timing is achieved, the frame synchronization procedure is initiated. In this second step, the secondary SCH codes must be used. Once the combination of secondary SCH codes used by the base station is identified, it is possible to acquire the general frame synchronization for downlink and the primary code group of cell simultaneously.

Finally, the exact primary scrambling code used by the cell is determined in the third step. This search is limited to the set of eight different scrambling codes determined by the primary code group. The reference channel employed in this step is the P-CPICH, which is transmitted continuously over the entire cell. The P-CPICH is an unmodulated code channel, which is scrambled with the cell-specific primary scrambling code of the cell. The P-CPICH is unique for each cell. After the primary synchronization code has been identified, the cell search procedure is finished and it is possible to apply the general fine synchronization algorithm in downlink with the P-CPICH channel. At the same time, the P-CCPCH is demodulated in order to extract the specific parameters necessary for user's demodulation, which are the channelization code, spreading factor, and the specific timing delay, for the downlink, and the scrambling and channelization codes, spreading factor, and DPCCH format, for the uplink. The combination of the cell search procedure and extraction of user's specific information is denoted as *set-up* stage of the modem.

Unlike downlink, each user has a specific synchronization reference in the uplink. If the modem knows the parameters of active users for uplink (obtained in the downlink set-up stage), the synchronization scheme is very simple. For each user, the timing reference is extracted from the DPCCH, applying the coarse and fine synchronization algorithms directly.

#### 4.1.2. Synchronization algorithms

The timing information of the transmitted frame is essential in order to properly demodulate the despread signal. Even

if there is a single chip duration error, the received spread spectrum signal cannot be properly demodulated.

Once the used codes in physical channels have been obtained, the appropriate timing reference is extracted. This synchronization issue is resolved following a two-step approach [20]. Firstly, coarse synchronization or initial code acquisition accomplishes the synchronization of the received signal and the corresponding code, with an uncertainty of half a chip period ( $\pm T_c/2$ ). Secondly, fine synchronization or code tracking performs and maintains the synchronization between the received signal and the code with a precision always lower than half a chip period.

To perform the synchronization, the scrambling code properties are used. These codes have an autocorrelation function that reaches its maximum when the code and the received signal are aligned.

##### 4.1.2.1. Coarse synchronization

As stated before, the objective of the coarse code synchronization is to achieve an initial code acquisition between the received signal and the corresponding scrambling code. This is equivalent to matching the phase of the spreading signal with the code.

There are different general acquisition techniques [19, 20, 21]. In the *serial search*, all the possible phases are tested one by one sequentially. The complexity for this method is quite low but the associated acquisition time is high. In the *parallel search*, all the possible phases are tested simultaneously. The complexity is higher but the acquisition time is much lower than in the serial search. An intermediate approach between the serial and parallel search strategies has been implemented in order to achieve the coarse synchronization with a moderate computational load, considering the complexity versus acquisition trade-off. A study of the computational load required by the different implementation approaches is shown in Table 3.

Considering the capacity of the used DSP's, the three-branches serial-parallel approach has been implemented. The block diagram of the coarse synchronization stage is shown in Figure 9.

In the figure, several blocks can be distinguished: correlators, thresholds generator, signal control modules, and a scrambling code generator. The received match-filtered signal is correlated with different cycle-delayed code versions. The maximum correlation value from the branches is compared with the first threshold  $\gamma_1$  which is obtained taking into account the second maximum correlation value. In order to

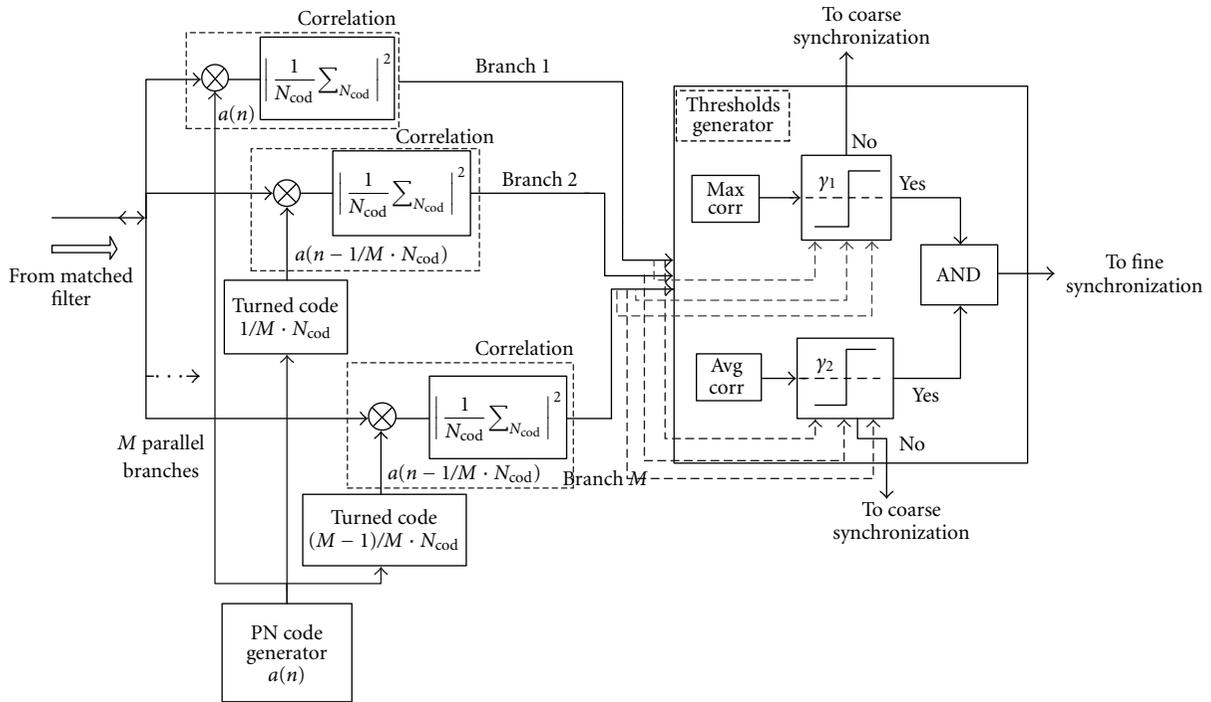


FIGURE 9: Block diagram of coarse synchronization.

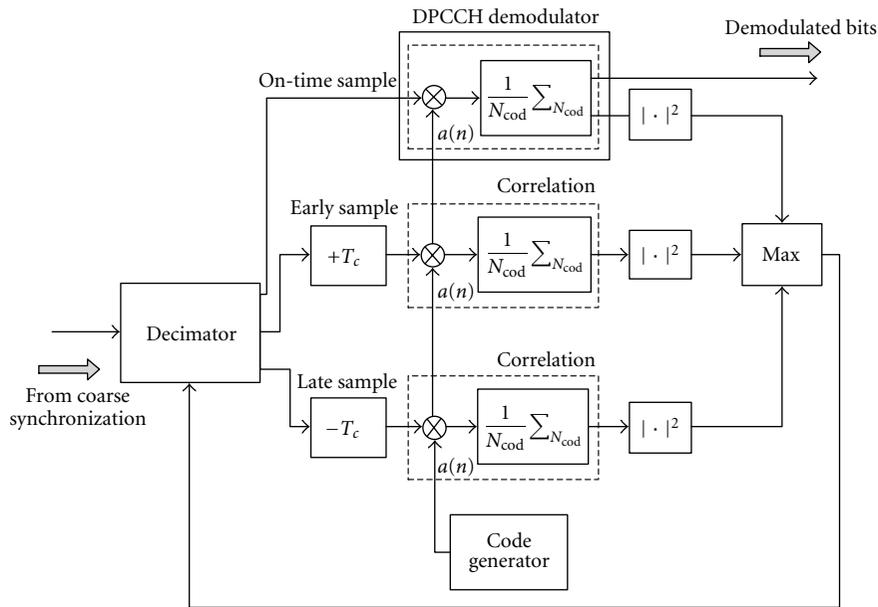


FIGURE 10: Block diagram of fine synchronization.

avoid situations in which the background noise may cause a wrong correlation which exceeds the first threshold, it is necessary to set another threshold to minimize this effect. This second threshold  $\gamma_2$  is calculated from the average of all the correlations except the maximum value. If the input signal surpasses both thresholds, then it is coarse-synchronized and fine synchronization is triggered.

#### 4.1.2.2. Fine synchronization

The purpose of code tracking is to perform and maintain the synchronization. Code tracking starts its operation only after coarse synchronization has been achieved. After coarse synchronization, a small phase error is still present. In order to correct this error, the loop structure shown in Figure 10 is used [19].

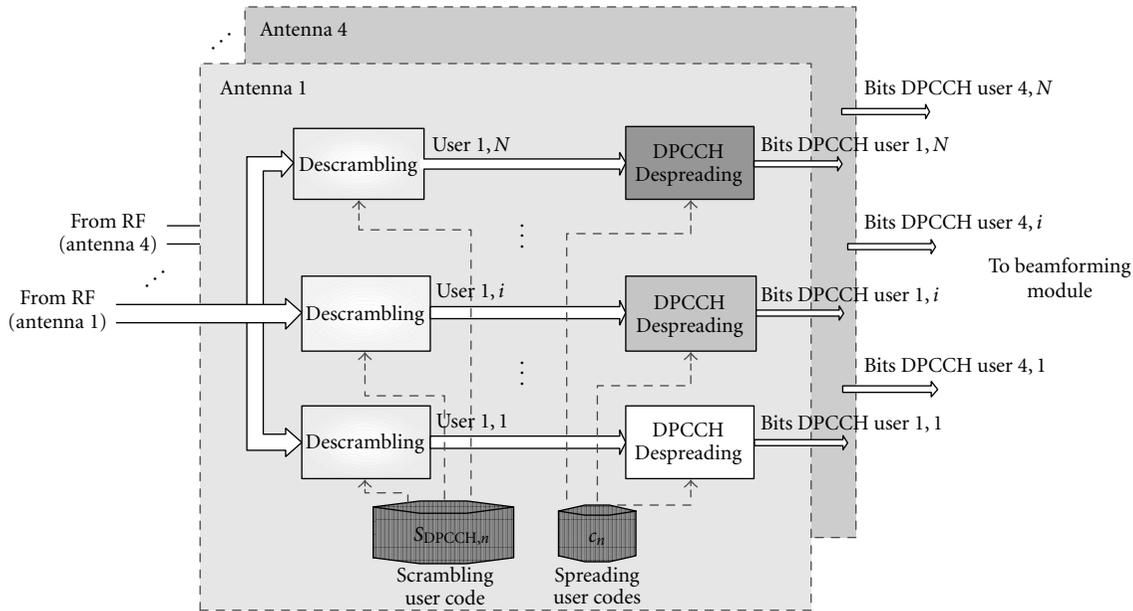


FIGURE 11: Uplink demodulator diagram.

The first block is a decimator that selects the correct sample at the right time, depending on the correlation value. In the second step, the decimated signal is delayed or advanced half a chip period, creating the late, early, and on-time branches. These three signals are correlated with the locally generated scrambling code, and the maximum absolute value of the correlations is selected. According to this selection, the timing information is updated.

**4.1.3. Demodulation in uplink and downlink**

Once the timing information and scrambling and channelization codes are determined, any UMTS physical channel can be demodulated.

In the uplink, DPCCH is demodulated for each user in order to extract the pilot bits that will be used as the reference signal in the beamforming process. To complete this task, two operations must be carried out: the complex-valued signal is descrambled by a complex-valued scrambling code  $S_{DPCCH,n}$  which identifies a user, and the signal is despread using the channelization code  $c_n$  which identifies the DPCCH channel. This process is shown in Figure 11.

In the downlink, the dedicated physical channel (DPCH) is demodulated. Firstly, the signal from Node B is descrambled by a complex-value scrambling code  $S_{dl,n}$  which identifies the cell and afterwards, the signal is despread through the correlation with a real-valued channelization code  $c_{ch,SF,n}$  which identifies the user in the downlink. Both time-multiplexed DPCCH and DPDCH (dedicated physical data channel) bits are obtained after this operation. Once the DPCH bits for every user have been demodulated and beamformed, the spreading operation is performed with  $c_{ch,SF,n}$  and scrambled with  $S_{dl,n}$ . The block diagrams of the modem for the downlink are shown in Figures 12a and 12b.

**4.2. Adaptive beamformer**

Immediately after the synchronization has been achieved, the following stage is the adaptive beamforming. The aim of this module is to calculate the set of array weights that make the array output signal satisfy an optimization criterion. Apart from this computation, the beamforming module adequately combines the received signal vector in order to produce a spatially filtered W-CDMA signal in the array output.

In the downlink, the base station transmits a separate beam pointing at the direction of each user, along with the broadcast channels, which are transmitted to the whole sector.

In this section, beamforming principles and implementation aspects are thoroughly explained. Moreover, theoretical expressions for the SINR are given for the operation of ADAM in uplink and downlink. In CDMA systems, this parameter is used for the estimation of capacity, throughput, and quality of service. Performance results will be shown in Section 6.1.

**4.2.1. Uplink operation and implementation**

Let  $\underline{x}(t)$  be the complex envelope representation for the vector of received signals in the array elements. For a situation with  $K$  mobile users and one interfering source  $i(t)$ , the vector  $\underline{x}(t)$  can be expressed as follows:

$$\underline{x}(t) = \sum_{k=1}^K \sqrt{P_k} \sum_{l=1}^{L_k} \alpha_{kl}^U(t) \underline{a}^U(\theta_{kl}) s_k(t - \tau_{kl}) + \sqrt{P_{int}} \underline{a}^U(\theta_{int}) i(t) + \underline{n}(t), \tag{1}$$

where  $P_k$  is the power transmitted from user  $k$ ,  $\alpha_{kl}$  and  $\tau_{kl}$  are the complex channel gains and delay of the  $l$ -path of the

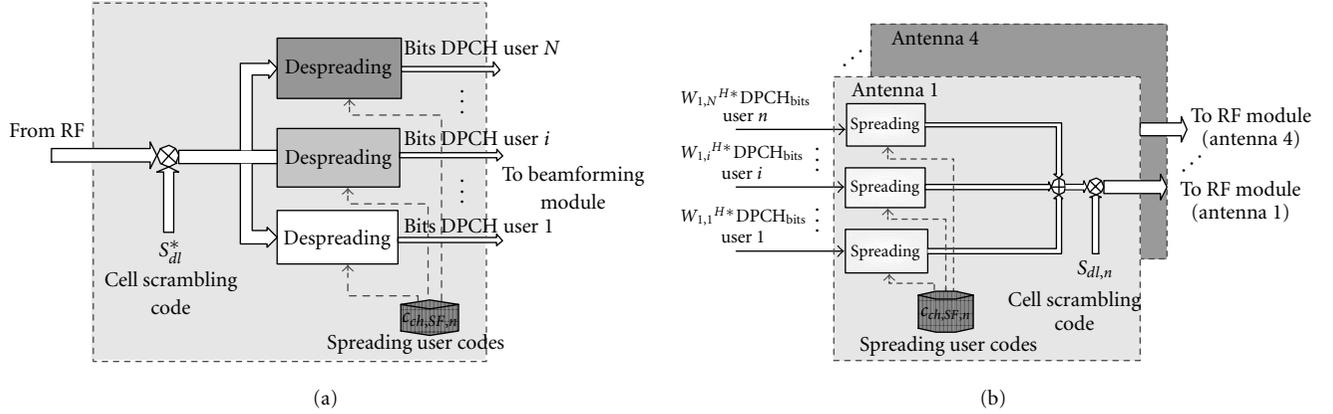


FIGURE 12: Downlink demodulator and modulator diagrams. (a) Downlink demodulator. (b) Downlink modulator.

$k$ th user, and  $\underline{a}(\theta) = g(\theta) \cdot \{\exp(j(2\pi/\lambda)d(l-1)\cos(\theta))\}$ ,  $l = 1, \dots, L$  is the response of a uniform linear array with  $L$  antenna elements and an interelement separation  $d$ , such as ADAM, to a wave impinging from an azimuth direction  $\theta$ , including the element antenna pattern  $g(\theta)$  [26]. The  $k$ th user signal  $s_k(t)$  includes modulation, data, and spreading.  $P_{\text{int}}$  is the power transmitted from the external interference source. Superscript  $U$  stands for the uplink. Finally,  $\underline{n}(t)$  is an  $L$ -dimensional complex Gaussian vector with independent and identically distributed (i.i.d.) components of zero mean and variance given by the corresponding signal-to-noise ratio (SNR).

In the uplink operation, two alternatives can be considered. The first one consists in performing a *total cancellation* of interfering sources for each user, including the contributions from other mobile users. Let  $\underline{w}_k$  be the uplink beamforming vector for each particular user in the total cancellation scheme. With this approach, if  $K$  users are present in the cell, then a separate beamformed signal  $y_k(t) = \underline{w}_k^H \underline{x}(t)$ ,  $k = 1, \dots, K$ , should be transferred to Node B. Therefore,  $K$  separate input channels would have to interface with Node B, and ADAM operation would lose its transparent behavior.

The other alternative is to apply a common beamforming weight vector  $\underline{w}$  to the composite received signal  $\underline{x}(t)$  (mobile user signals plus interference sources). The approach applied to ADAM is to use a linear combination of  $\underline{w}_k$  weights to perform the common beamforming operation that is required in the uplink. All individual beamforming vectors have a common feature, namely, the cancellation of interfering sources external to the system. Following this technique, the array output can be expressed as follows:

$$\begin{aligned}
 y(t) &= \sum_{k=1}^K \underline{w}_k^H \underline{x}(t) = \left\{ \sum_{k=1}^K \underline{w}_k^H \right\} \underline{x}(t) = \underline{w}^H \underline{x}(t) \\
 &= \underline{w}^H \left\{ \sum_{k=1}^K \sqrt{P_k} \sum_{l=1}^{L_k} \alpha_{kl}^U(t) \underline{a}^U(\theta_{kl}) s_k(t - \tau_{kl}) \right\} \\
 &\quad + \sqrt{P_{\text{int}}} \underline{w}^H \underline{a}^U(\theta_{\text{int}}) i(t) + \underline{w}^H \underline{n}(t).
 \end{aligned} \quad (2)$$

This scheme is called *partial interference cancellation* because only common interfering sources will be canceled after applying common beamforming weights. The uplink SINR in the array output for user  $k$  is therefore given by

$$\begin{aligned}
 \text{SINR}_k^{UL} &= \frac{\underline{w}^H E \left\{ \left| \sqrt{P_k} \sum_{l=1}^{L_k} \alpha_{kl}^U(t) \underline{a}^U(\theta_{kl}) s_k(t - \tau_{kl}) \right|^2 \right\} \underline{w}}{\left\{ \underline{w}^H E \left\{ \left| \sum_{i=1, i \neq k}^{K-1} \sqrt{P_i} \sum_{l=1}^{L_i} \alpha_{il}^U(t) \underline{a}^U(\theta_{il}) s_i(t - \tau_{il}) \right|^2 \right\} \underline{w} \right\} + \underline{w}^H E \left\{ \left| \sqrt{P_{\text{int}}} \underline{a}^U(\theta_{\text{int}}) i(t) \right|^2 \right\} \underline{w} + \underline{w}^H E \left\{ \left| \underline{n}(t) \right|^2 \right\} \underline{w} \right\}}.
 \end{aligned} \quad (3)$$

The second term in the denominator represents the common interference contribution that appears in the array output. The level of common interference cancellation is given by the magnitude of  $|\underline{w}^H \underline{a}^U(\theta_{\text{int}})|^2$ .

In both alternatives, the calculation of individual beamforming weights  $\underline{w}_k$  fulfils the minimum mean square error (MMSE) criterion in the array output. The optimum solution is given by the Wiener-Hopf equation as  $\underline{w}_k = R_k^{-1} \underline{p}_k$ , where  $R_k = E\{\underline{x}_k(n) \underline{x}_k^H(n)\}$  and  $\underline{p}_k = E\{\underline{x}_k(n) d_k^*(n)\}$ ,  $\underline{x}_k(n)$  and  $d_k(n)$  being the vector of demodulated pilot bits in the antenna array and the reference pilot bits, respectively. This equation does not represent a practical solution so that a sub-optimum set of weights must be calculated by means of adaptive algorithms. This procedure is based on the iterative estimation of  $\underline{w}_k$  each time a new pilot bit is demodulated. In this way, the antenna is capable of adapting its radiation pattern to a fast varying environment.

Two well-known adaptive algorithms have been considered, namely, NLMS and RLS, whose update equations are shown in Figures 13a and 13b, respectively. The first one is the NLMS, which is based on the instantaneous estimation of  $R_k$  and  $\underline{p}_k$ , and only vector operations must be performed. Due to its simplicity and reduced computational complexity of  $O(L)$ , NLMS is very suitable to a practical implementation that must comply with real-time requirements.

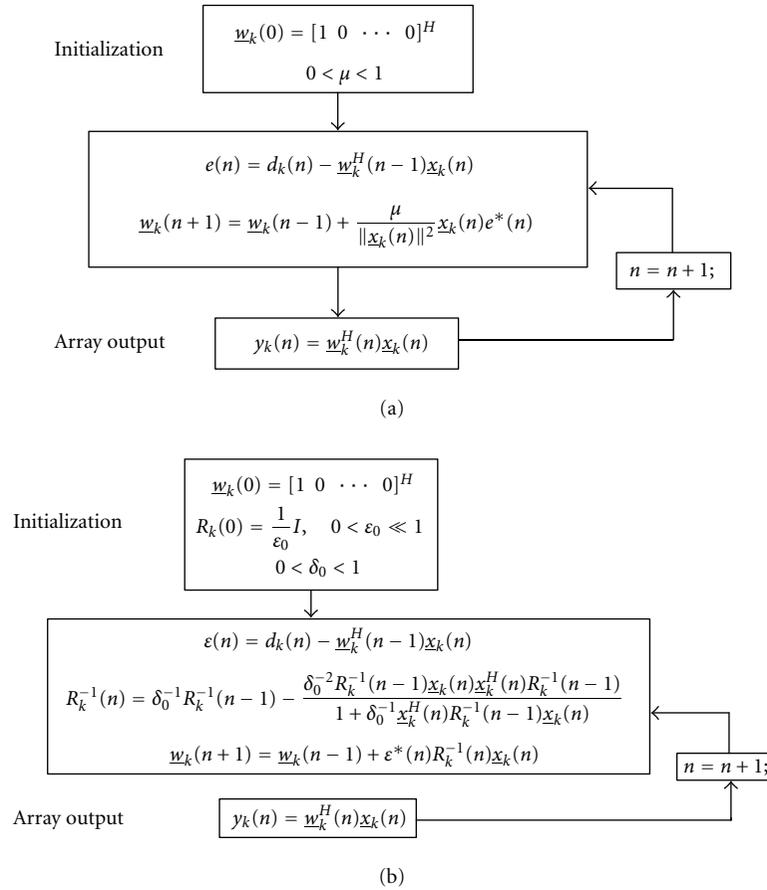


FIGURE 13: Update weight equations. (a) NLMS algorithm. (b) RLS algorithm.

On the other hand, RLS is based on the iterative estimation of the autocorrelation matrix  $R_k$ , which imposes a higher computational load than NLMS, although its convergence speed is faster. Matrix operations make RLS unfeasible for real-time implementations, being the complexity of  $O(L^2)$ . Further information on these algorithms can be found in [17, 18].

Regarding the implementation aspects, Figure 14 shows the beamforming structure used in the uplink. Demodulated DPCCH bits from the modem and the received signal vector  $\underline{x}(t)$  are the inputs of the beamforming module. These bits are used to acquire the slot synchronization, which is necessary to obtain the correct pilot bits that have to be used as the reference signal in the weight computation process. After that, the single-user weights are computed, and the common beamforming weight vector is calculated. The output of the beamforming process is obtained by multiplying the received signal vector by the common weight vector and combining the resulting signal vector, without the need of demodulating each user's data channels. It must be noticed that the signal vector impinging in the array antenna is composed of contributions from several users so that both the slot synchronization and the single-user weight computation must be executed in parallel for every serviced user.

#### 4.2.2. Downlink operation and implementation

Optimal downlink beamforming will minimize the interference received by other users and will enhance the useful signal power received by the desired user. Due to the frequency translation that appears in a frequency division duplex (FDD) system, uplink and downlink communication scenarios are different. As a consequence of the lack of downlink channel information, the calculation of transmission weights is all but an easy task.

One of the methods for estimating downlink weights from uplink channel information is the use of the uplink spatial covariance matrix [27]. However, this approach involves the use of many computational resources.

In order to reduce the complexity of the implementation, ADAM uses individual uplink weights as transmission vectors. Due to frequency translation ( $\Delta f = 190$  MHz), the transmitting array factor is similar but not the same as in the uplink. Therefore, the synthesized beam will not be pointing exactly in the direction of the mobile user, and the interference received by other user will be reduced but not minimized.

For downlink, after joint transmission of the weighted signals bounded for the  $K$  users from the base station, the

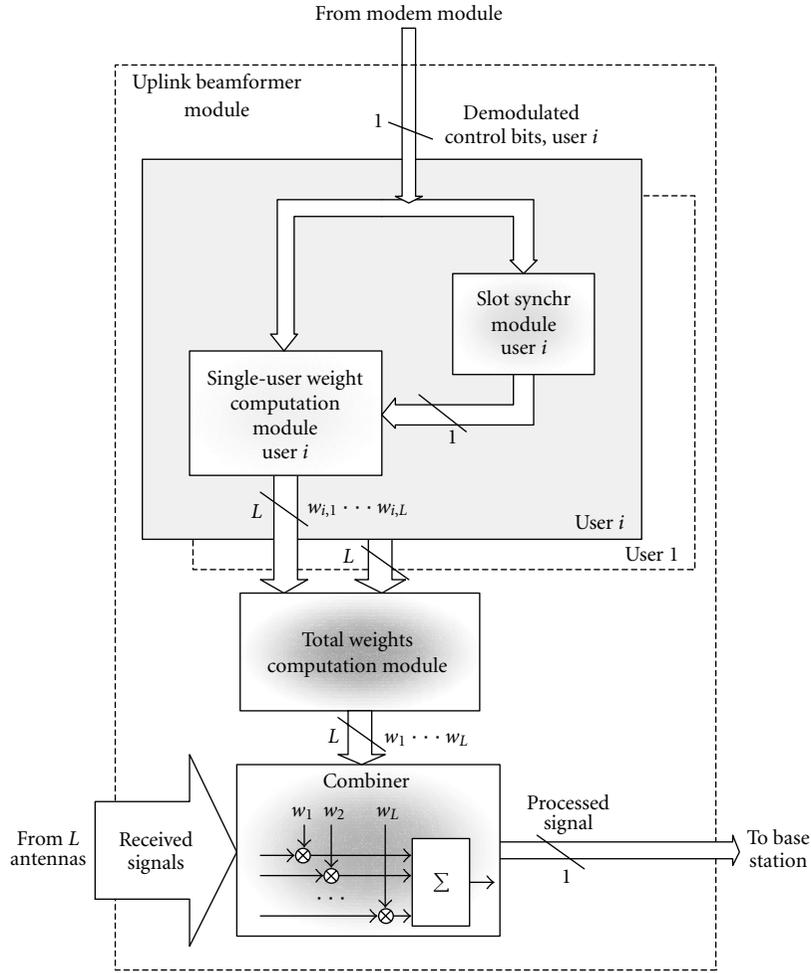


FIGURE 14: Beamformer structure: uplink.

baseband signal received by the mobile  $k$ ,  $x_k(t)$ , is given by

$$x_k(t) = \sum_{k=1}^K \sqrt{P_k} \mathbf{w}_k^H \sum_{l=1}^{L_k} \alpha_{kl}^D(t) \mathbf{a}^D(\theta_{kl}) s_k(t - \tau_{kl}) + n_k(t), \quad (4)$$

where  $\mathbf{w}_k$  is the transmission beamforming vector for user  $k$ ,  $P_k$  is the power assigned to the user  $k$  signal, and  $n_k(t)$  is a complex white Gaussian process that represents the thermal noise contribution in the mobile user equipment. The other elements of (4) have the same meaning as in (1). In an FDD system, uplink ( $\alpha_{kl}^U$ ) and downlink ( $\alpha_{kl}^D$ ) fading coefficients are uncorrelated, and in the simulations, they have been generated from independent Rayleigh fading processes.

The SINR perceived by the user  $k$  can be expressed as follows:

$$\text{SINR}_k^{DL} = \frac{\mathbf{w}_k^H E \left\{ \left| \sqrt{P_k} \sum_{l=1}^{L_k} \alpha_{kl}^D(t) \mathbf{a}^D(\theta_{kl}) s_k(t - \tau_{kl}) \right|^2 \right\} \mathbf{w}_k}{E \left\{ \left| \sum_{i=1, i \neq k}^{K-1} \mathbf{w}_i^H \sqrt{P_i} \sum_{l=1}^{L_k} \alpha_{kl}^D(t) \mathbf{a}^D(\theta_{kl}) s_i(t - \tau_{kl}) + n_k(t) \right|^2 \right\}}. \quad (5)$$

In contrast to the uplink, a complete user separation can be performed in the downlink direction. Because of that, in the proposed downlink structure shown in Figure 15, single-user weight vectors calculated in the uplink are applied as transmit beamforming weights to each user separately. Therefore, downlink beamforming is much simpler than the uplink one since the adaptive weight calculation is not required.

However, demodulation of data and control bits for each user is required, and downlink beamforming is applied at the bit level. This fact results in a considerable reduction in the computational load, as far as the multiplier submodule is concerned, in comparison to the equivalent module for the uplink. Moreover, a total cancelation of interferences is achieved thanks to the individual user separation. Nonetheless, this scheme increases the complexity of the modulation and demodulation module, as it must be performed for every user in every element of the antenna array.

In contrast to dedicated channels, broadcast information conveyed by common transport channels must be received by all the users in the cell. Therefore, the associated

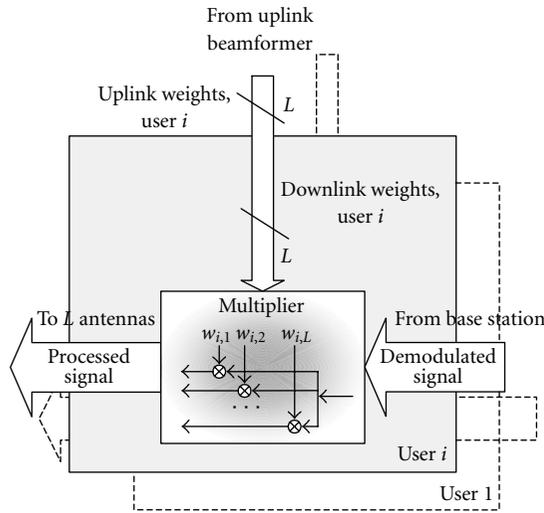


FIGURE 15: Beamformer structure: downlink.

physical channels are transmitted to the whole sector through one of the antenna array elements, which is equivalent to the radiation pattern of a conventional sectored antenna.

## 5. CODE OPTIMIZATION AND LOAD DISTRIBUTION

### 5.1. Maximum number of instruction per DSP

The analog received signal is sampled in the ADC to four samples per chip rate, obtaining packets of 1024 samples. Therefore, the number of samples per DPCCH bit can be calculated as follows:

$$\begin{aligned} \text{Number of samples per DPCCH bit} \\ &= 1 \text{ bit} \cdot 256 \text{ chips/bit} \cdot 4 \text{ samples/bit} \quad (6) \\ &= 1024 \text{ samples/bit.} \end{aligned}$$

For real-time execution, the allowable time for processing each packet of 1024 samples is one DPCCH bit period, that is, 66.67 microseconds.

Each of the used DSP has a performance capability of up to 2400 MIPS on pipeline. Their architecture has eight highly independent functional units (six ALUs (arithmetic and logical units) of 32-/40-bits and two 16-bit multipliers). Therefore, eight 32-bit instructions per cycle can be executed. The clock rate is 300 MHz. As a result, the number of DSP cycles for processing one bit is

$$\text{Num(cycles/bit)}_{\text{MAX}} = \text{MIPS} \cdot T_b = 300 \cdot 66.67 \mu\text{s} = 20000. \quad (7)$$

Therefore, the number of clock cycles in all the modules should be lower than 20000 cycles per bit. Table 4 shows the number of clock cycles per bit of each module. These values are always lower than the maximum number of clock cycles per bit.

### 5.2. Code optimization and module load

For code optimization, the following steps have been followed [28]:

- (1) C-code writing;
- (2) obtaining a maximum instruction reduction;
- (3) use of DSP intrinsic operations.

The flow diagram in Figure 16 summarizes the previous steps.

These intrinsic operations belong to a special C62x DSPLIB library. It consists of some optimized functions for fixed-point DSPs. These functions are especially used in real-time applications since their execution time is much lower than the C equivalent code. The main disadvantage is that they can only be used under certain restrictive conditions.

After code optimization and programming, module load has been measured. Table 5 shows the complexity of each optimization step in clock cycles, time, and clock cycles per second for the module of coarse synchronization. As it can be seen, the complexity of coarse synchronization module has been reduced two orders of magnitude after the third optimization step.

Table 4 illustrates the computational load of the other modules when the three optimization steps have been applied. The reduction in clock cycles of the other modules is, on average, two orders of magnitude too.

### 5.3. Load distribution in processors

According to the required computational capacity for each module after code optimization, the distribution of load and tasks between DSPs must be carried out.

As presented in the Section 3, six Quads boards have been used for signal processing, with four DSPs each. In order to properly design the load distribution between DSPs, several questions have been considered. To begin with, it has been taken into account that two independent polarizations should be processed, so the number of available DSPs for each one is 12. Despite this independence, the load cannot be divided into three Quads per polarization. This is due to the need of five broadband receiver channels plus ADC and five broadband transmitter channels plus DAC per polarization. Receivers and transmitters boards consist of two channels each, which have to be associated to two DSPs in the same Quad. As a result, receiving and transmitting channels must be considered as pairs so that the possibility of using three independent Quads per polarization is eliminated.

Another point to take into consideration is the association of a task per DSP as far as possible. In this way, not only the load distribution but also the data exchange between DSPs is more easily understood. As regards the data exchange between DSPs and Quads, the tasks and load distribution has been designed aiming at reducing the number of data transfers between processors as much as possible. This makes the interconnection between DSPs simpler since less synchronization for data exchange is needed. Data transfers between DSPs are preferred to those between Quads due to the higher complexity of transfer and synchronization in the last ones.

TABLE 4: Computational load of synchronization, demodulation, and beamforming modules.

Module	Clock cycles/bit	Time ( $\mu\text{s}$ )	Million of clock cycles per second	DSP capacity used (%)
Coarse synchronization	8700	29	131	43.5%
Fine synchronization and demodulation	6600	22	99	33%
Slot synchronization	890	3	13	4.43%
Single-user weight computation	830	3	13	4.17%
Common weight computation	5200	17	78	26%
Combiner	11500	38	174	58%

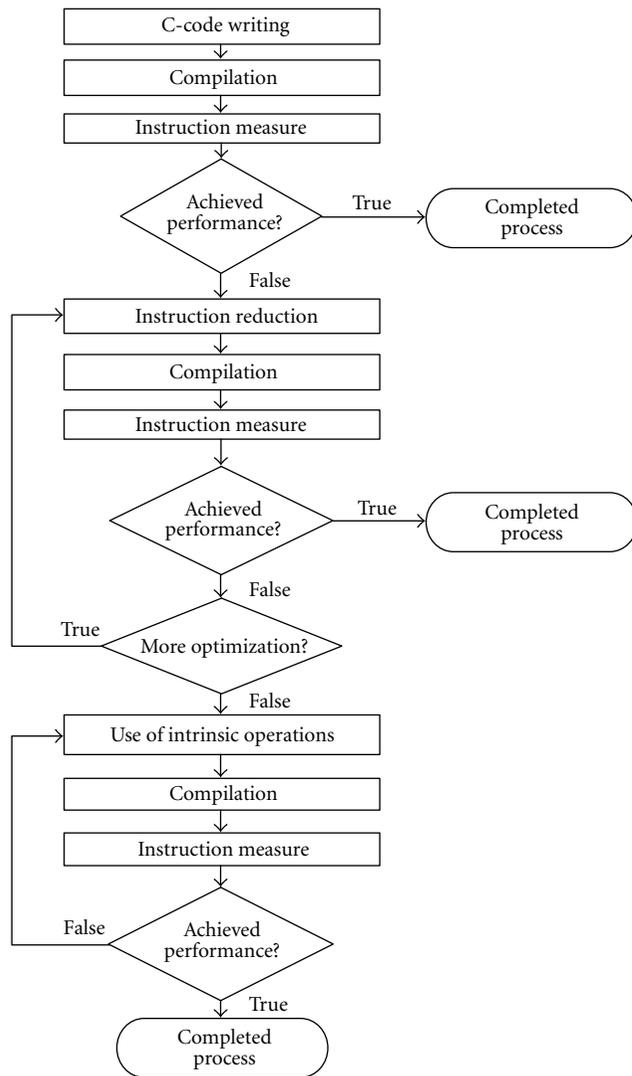


FIGURE 16: Optimization steps flow diagram.

On this basis, the scheme in Figure 17 is proposed. As it can be observed, the synchronization and demodulation of received signals from the four antennas in the uplink are processed in the same Quad, hence avoiding the extra data

exchange and the difficulties of making a correct synchronization if the signals would be received in different Quads. Similarly, all the transmitted data are obtained and sent to the antennas in a single Quad, for the downlink. Due to computational cost restrictions, only three users can be processed with this hardware implementation. A higher number of users could be processed with more DSPs or with higher computational capacity ones.

## 6. RESULTS

### 6.1. Simulations results

Simulations have been conducted to obtain uplink and downlink performance results. Several aspects and characteristics have been varied in order to study different possible implementations.

Concerning adaptive algorithms, the performances achieved with two of them have been studied. These algorithms are NLMS and RLS, which have been widely used in adaptive array processing applications [17]. However, due to computational load restrictions, only NLMS has been implemented in the first version of the prototype.

As it was said in the introduction, the operation of ADAM must be completely transparent to Node B. This approach has an impact on the performance achieved with the adaptive beamformer. In the first group of simulations, a single-cell scenario with a variable number of mobile users is studied, including the effect of external interference on system performance. Afterwards, system-level simulation results show the capacity increase obtained with ADAM, compared to a conventional sector antenna.

#### 6.1.1. Uplink simulation results

As explained in Section 4.2, two cancellation schemes have been considered in the uplink: *total interference cancellation* and *partial interference cancellation*. Performance obtained with both schemes has been studied by means of simulation.

Figure 18 shows the performance achieved by both cancellation schemes when an external interference source is present.

As it can be observed, both array factors cancel the external interference contribution. However, if the total cancellation scheme is used, contributions from other mobile users

TABLE 5: Reduction of the number of clock cycles in the coarse synchronization module.

Module	Optimization step	Clock cycles/bit	Time ( $\mu s$ )	Millions of clock cycles per second
Coarse synchronization	1	2571296	8571	38500
	2	65732	219	986
	3	<b>8700</b>	<b>29</b>	<b>131</b>

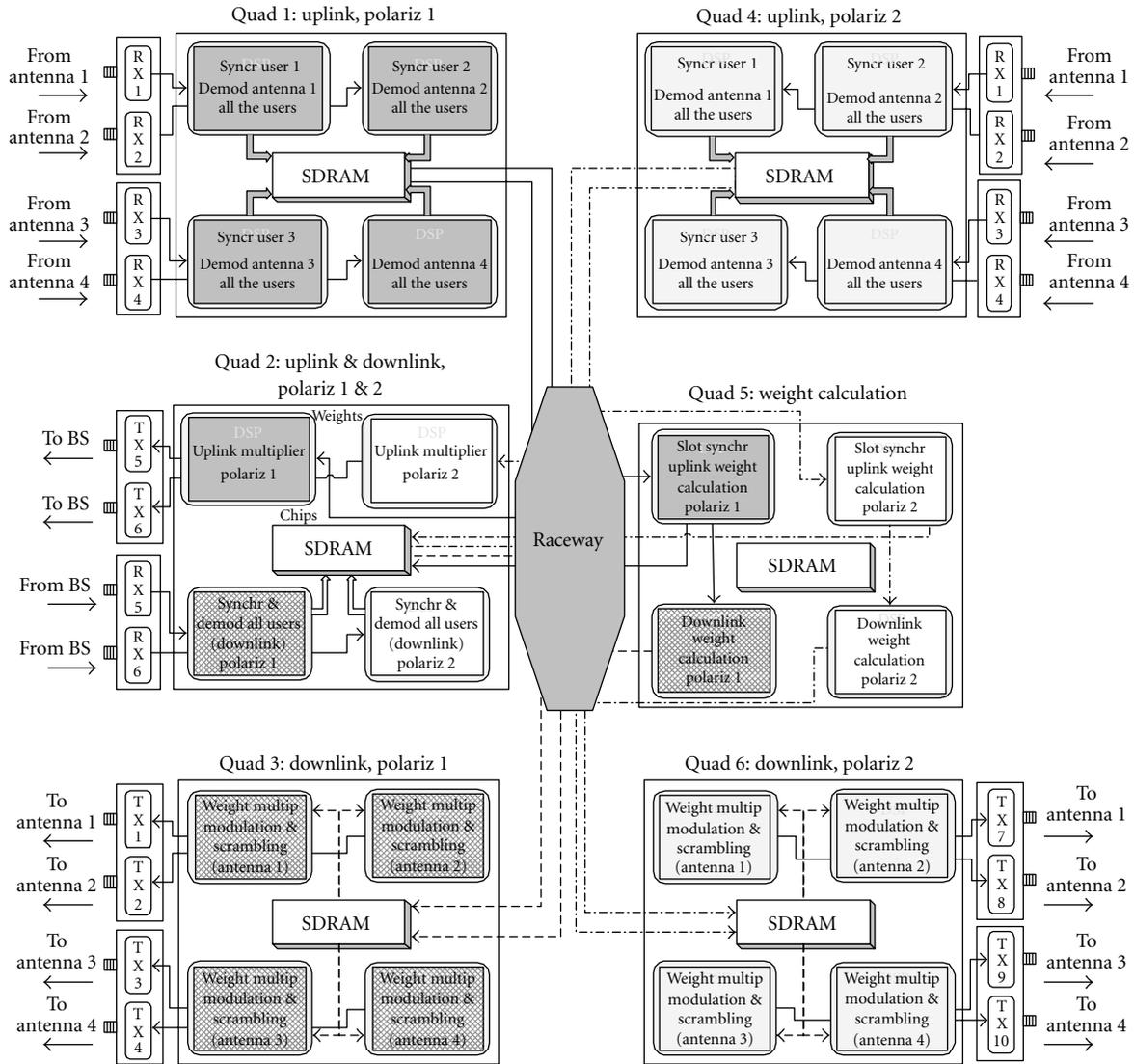


FIGURE 17: Load distribution.

will also be canceled, whereas with partial cancellation, a simultaneous pointing in the directions of mobile users appears as a result of the linear combination of  $w_k$ . As the number of users uniformly distributed within the cell is increased, the final uplink radiation pattern tends to provide a sectored coverage.

A simulation environment with a uniform distribution of mobile pedestrian users has been studied. Mobile speed is

3 km/h, and multipath fading is given by the two-path profile proposed in [2]. In (1),  $\theta_{kl}$  is characterized by a Laplacian azimuth spectrum along with a Gaussian distribution for each user, with an angular spread of  $10^\circ$ . The number of rays impinging on the array per user is found as a Poisson random variable with a mean value of 25 [29]. Each user transmits only one data channel, with a spreading factor of 64.

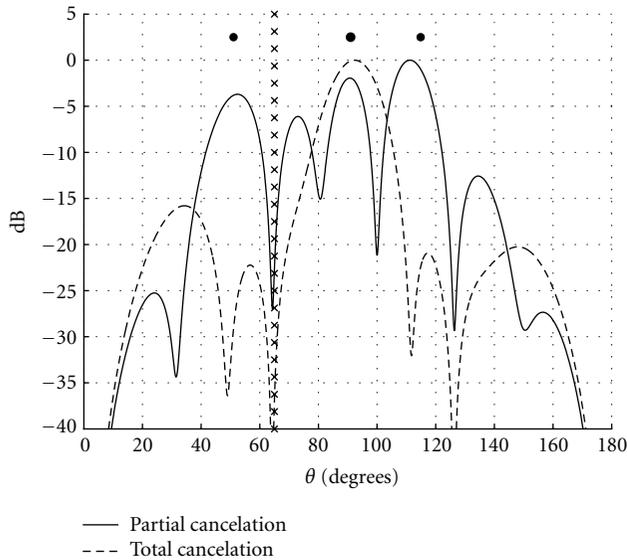


FIGURE 18: Normalized array factor for total and partial cancellation schemes (•: mobile users, ×: external interference).

In the simulations, a perfect power control algorithm is assumed for mobile users, that is,  $P_k = P$ , and external interference power  $P_{\text{int}}$  is set to  $F$  dB over  $P$ .

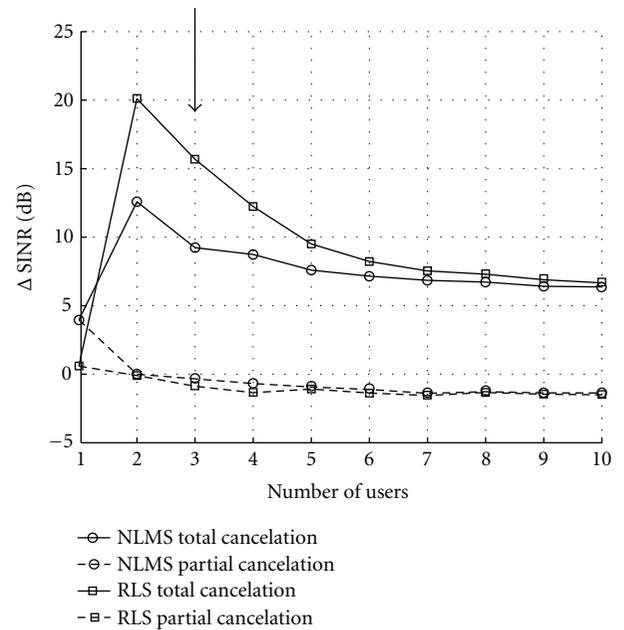
Figure 19 shows the average uplink SINR increase obtained with ADAM with respect to a typical sectored antenna in two scenarios:  $F = -150$  dB (only mobile users are present in the cell), and one external interference with  $F = 20$  dB. In the first scenario, the SINR improvement converges to 6 dB when the total cancellation scheme is used. With partial cancellation, ADAM will provide the same performance as the individual sectored antenna. However, in the second scenario, the partial cancellation scheme outperforms the sectored antenna in more than 5 dB.

As it can be observed in Figure 19, RLS provides better performance than NLMS, although their behavior converges as the number of users increases. In the case of  $F = 20$  dB, the difference between both algorithms is mainly due to the fact that RLS provides a higher cancellation level for the external interference source.

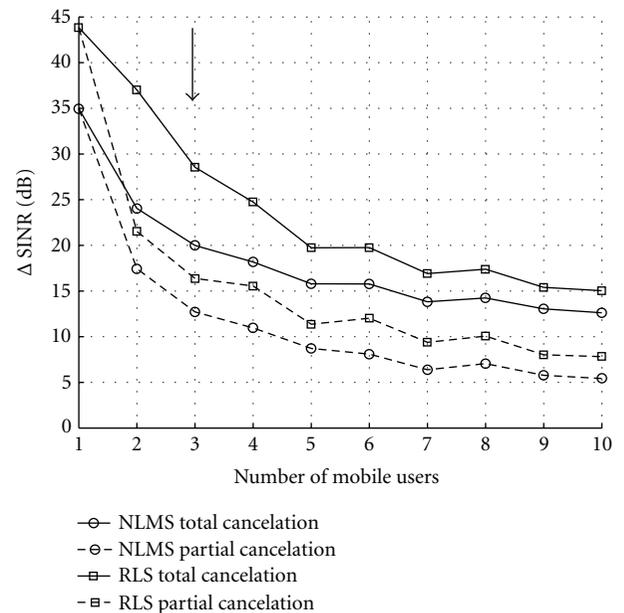
Table 6 shows the reduction of the interference power in the array output as a function of  $F$ . It can be observed that as  $F$  increases, both algorithms provide a more significant interference reduction. The inclusion of the strong external interference produces a spatial coloured covariance matrix because the most significant part of interfering power is concentrated around the same angular direction. As a consequence, when the number of users increases and  $F$  is reduced, interference is uniformly distributed in the cell, and the level of interference cancellation is low.

### 6.1.2. Downlink simulation results

In the downlink, only the total interference cancellation scheme has been considered. Figure 20 shows the average SINR increase experimented by the mobile user when the



(a)



(b)

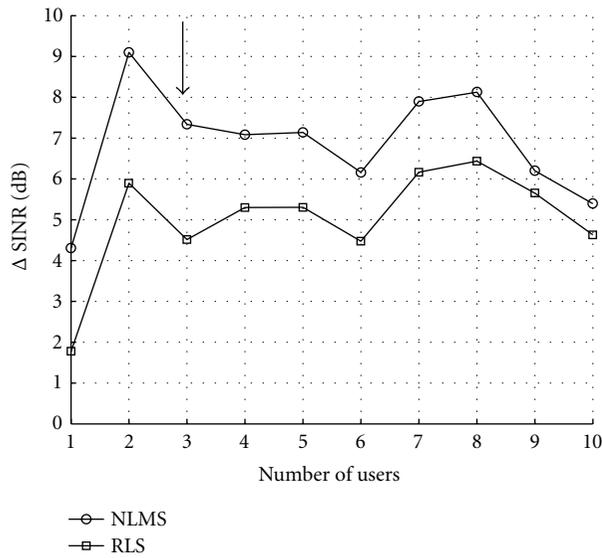
FIGURE 19: Uplink SINR increase as a function of the number of users. (a)  $F = -150$  dB. (b)  $F = 20$  dB.

proposed downlink beamforming algorithm is used. In the downlink, NLMS provides a higher SINR increase than RLS because the improvement obtained with RLS is mainly due to the cancellation of the external interfering source, which does not influence downlink performance.

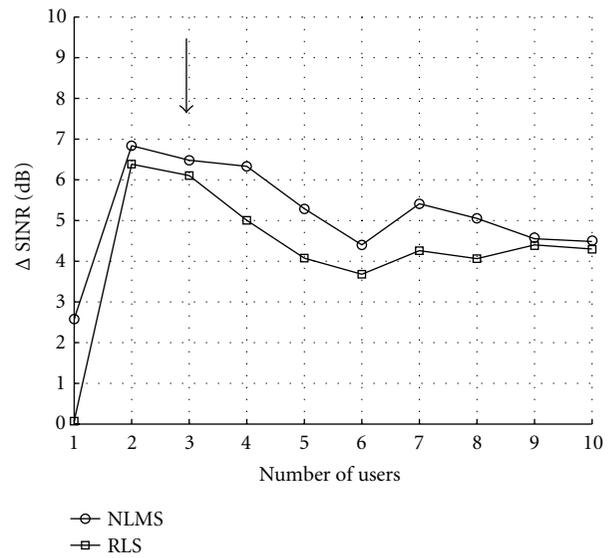
In Figures 19 and 20, the performance is studied for up to ten mobile users in the system. However, and as it was

TABLE 6: Cancellation level of the external interference source with the partial cancellation scheme.

$F$ (dB) $10 \log_{10} (P_{\text{int}}/P)$	Interference cancellation level (dB) $10 \log_{10} ( \underline{w}^H \underline{a}^U(\theta_{\text{int}}) ^2)$			
	$K = 3$ users		$K = 10$ users	
	NLMS	RLS	NLMS	RLS
-150	-5.71	-5.52	-5.36	-6.38
0	-9.29	-24.36	-7.83	-8.58
10	-15.47	-37.19	-10.31	-13.02
20	-23.58	-50.83	-16.00	-24.54
30	-23.88	-60.53	-19.53	-40.83



(a)



(b)

FIGURE 20: Downlink SINR increase seen by the mobile user. (a)  $F = -150$  dB. (b)  $F = 20$  dB.

explained in Section 5.3, available processing hardware allows for processing a maximum of three mobile users. As well, due to the demanding computational load required by the RLS algorithm, the beamforming process is controlled by NLMS. From the above results, it can be concluded that, for a situation with three users and an external interfering source of  $F = 20$  dB, ADAM prototype provides a SINR increase of 12.5 and 6.5 dB over a conventional sectored antenna in the uplink and downlink, respectively.

### 6.1.3. Performance in a typical user scenario

In contrast to existing 2G networks, which are dominated by voice traffic, UMTS networks will provide a mixture of voice and data services with different specifications of bit rate and quality of service [30]. In addition, unlike voice, most data services are asymmetric in nature, meaning that people download more information than they send. This is typical of web browsing and streaming media services.

In order to model the mixed-service and asymmetric characteristics of UMTS networks, three different subscriber profiles are considered. The first one corresponds to a conventional voice service, with a symmetric bit rate of 12.2 kbps. The second group of subscribers deals with an asymmetric data service of 12.2/64 kbps, and the third group demands a 12.2/144 kbps asymmetric data service. In these conditions, the base-to-mobile link will limit the capacity of the system. However, this limitation can be overcome using ADAM prototype thanks to the total interference cancellation achieved in the downlink.

The actual capacity increase achieved using the ADAM prototype must be estimated through system-level simulations. A scenario with 19 sites and 57 sectors is considered. The distance between adjacent sites is 3000 m. In the simulations, 2000 users have been uniformly distributed within the region of interest. Regarding the service distribution, 1000 subscribers demand a voice service, 500 users demand the low bit rate data service, and the other 500 users demand a

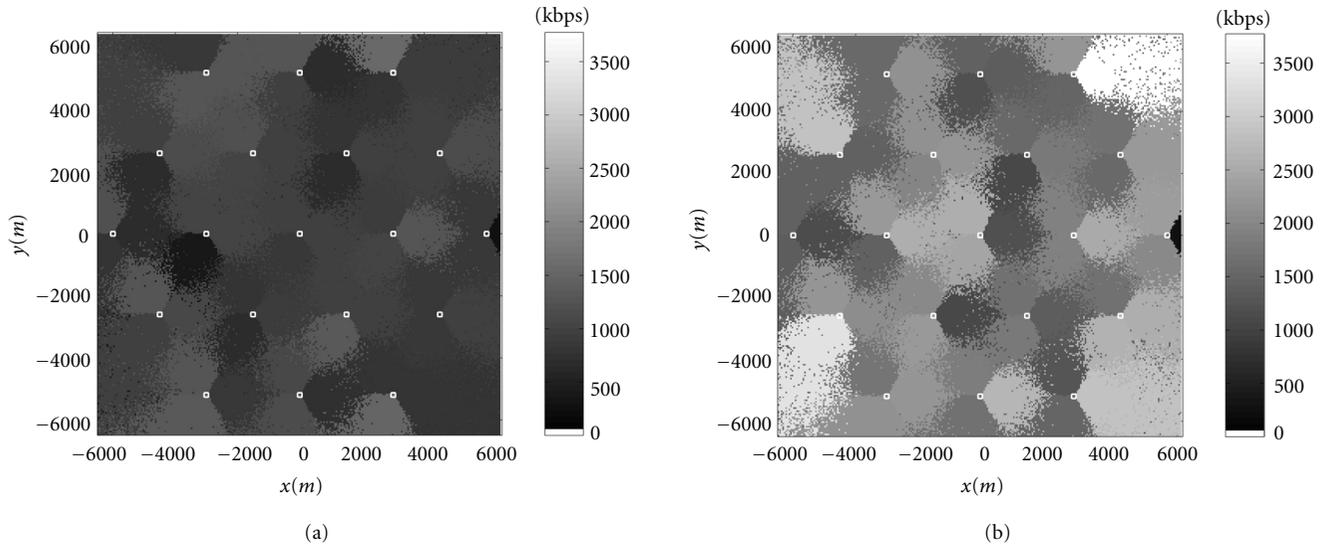


FIGURE 21: Total throughput per cell in the downlink ( $\square$ : base station site). (a) Sector antenna. (b) ADAM prototype.

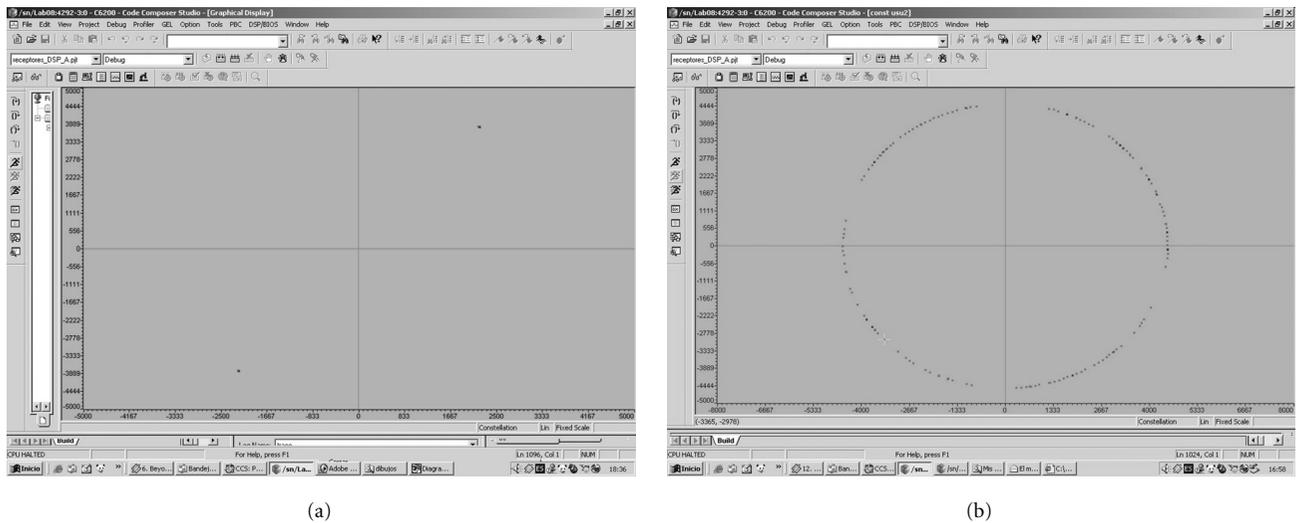


FIGURE 22: Constellation of demodulated symbols with phase and frequency offset. (a)  $\omega_0 = 0$ ,  $\varphi_0 = 5^\circ$ . (b)  $\omega_0 = 8.5$  Hz,  $\varphi_0 = 5^\circ$ .

high bit rate data service. Simulations have been performed with the network planning tool presented in [31], complemented with the incorporation of smart antennas in the scenario.

Figure 21 shows the total throughput per cell in the downlink. Using the ADAM prototype, the throughput is increased by a factor of 2 in each sector, in relation to the situation with sector antennas. This capacity increase comes from the lower number of users put to outage when the adaptive antenna is used.

## 6.2. Implementation results

One of the most important effects for the implementation is the carrier frequency error between transmitter and receiver signals. In general, the carrier error  $\varphi(n)$  consists of two com-

ponents [32]:

$$\varphi(n) = \omega_0 nT + \varphi_0, \quad (8)$$

where  $\omega_0$  is the frequency offset,  $\varphi_0$  is the constant phase offset, and  $T$  is the symbol period.

The phase offset stays constant during the reception so that it can be compensated during the set-up stage. In contrast, the frequency offset produces the most damaging effect. Depending on its value, the total error  $\varphi_0(n)$  varies faster. This effect is corrected modifying the received symbol phase to  $\pm 90$  degrees since all the DPCCH symbol information is transmitted through Q channel.

The constellation of demodulated DPCCH symbols in the uplink is shown in Figure 22. Firstly, (a) illustrates a

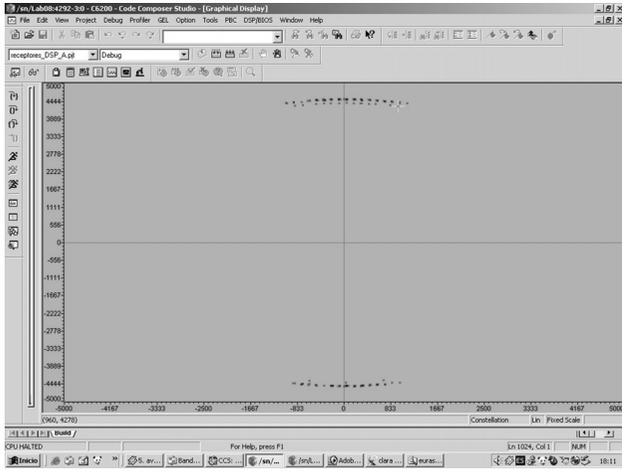


FIGURE 23: The constellation of demodulated symbols after phase error compensation.

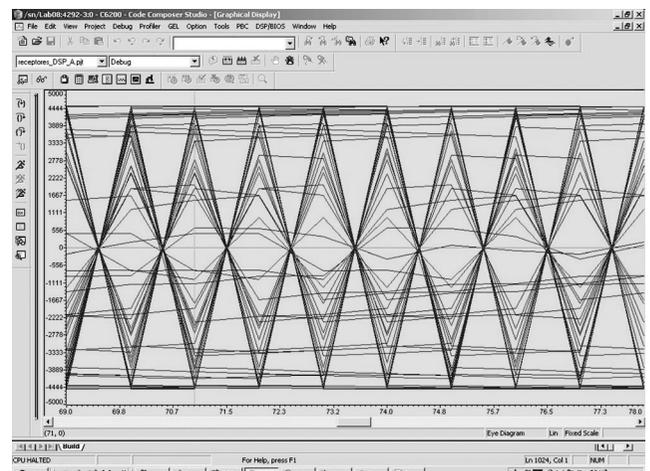
situation where the phase error is 5 degrees and the frequency error is zero. In contrast, when the frequency error is not zero (Figure 22b) and no correction algorithm is applied, the phase of the received signal rotates  $2\pi$  radians each 1765 bits, that is, 117.6726 milliseconds equivalent to 11 radio frames.

Figure 23 shows the constellation after the phase correction algorithm has been applied. The phases of the symbols are not exactly 90 degrees because the implementation of the phase correction algorithm has been realized with finite resolution (sine and cosine tabulated functions). In particular, a  $30^\circ$  resolution is considered in Figure 23 so that the maximum phase error is  $30^\circ$ . If the functions were implemented with a finer resolution, the uncertainty would be smaller. A resolution increment of the algorithm does not imply higher complexity, but it would require more memory size since more data must be stored. Depending on the memory resources, more precision could be added.

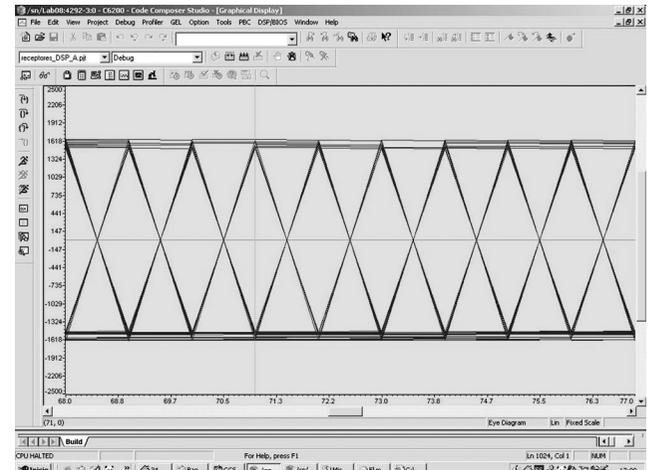
Eye diagrams are another representation for analyzing quality of demodulated symbols. Figures 24a and 24b show the eye diagrams of noncorrected and corrected received signal, respectively. The received signal can be sampled better in the corrected case due to the improvement achieved with the phase compensation algorithm.

As an example of modulation and beamforming performance, Figure 25 shows the IF spectrum (44 MHz) in the output of the broadband transmitters (uplink and downlink paths). As it can be seen, the spectrum corresponds to that of a W-CDMA signal, proving the transparent operation of ADAM. This signal is then upconverted to the UMTS band in the RF stages, and transferred to Node B (uplink) or to the antenna array (downlink).

Finally, the synthesized radiation pattern is plotted in Figure 26. Calculated weights are transferred in real time from the DSP blocks to the monitor-PC using LabView. This application is used in the monitoring stage to test the final radiation patterns in uplink and downlink, and to control the



(a)



(b)

FIGURE 24: Eye diagrams of demodulated symbols. (a) Without phase correction. (b) With phase correction.

correct performance of the overall system, from RF parameters to SINR.

Currently, integration tests and measurements of ADAM prototype are being performed in order to characterize the behaviour of the complete system (antenna array, RF-to-IF chains, and DSP stages) in uplink and downlink.

### 7. CONCLUSIONS

A novel and real implementation of an adaptive antenna prototype for UMTS has been presented. Its main features are flexibility, modularity, and transparency in operation, which make it suitable to be connected to any existing Node B, regardless of whether it is prepared to work with an adaptive antenna or not. The focus of the contribution has been on the implementation aspects of the digital signal processing stages

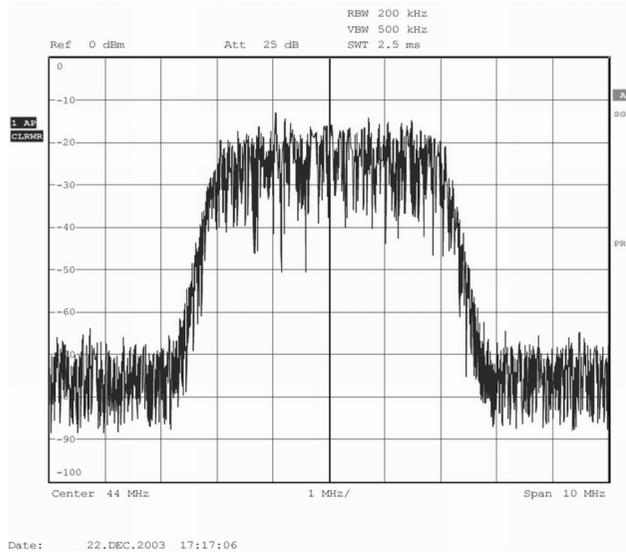


FIGURE 25: IF spectrum (uplink path).

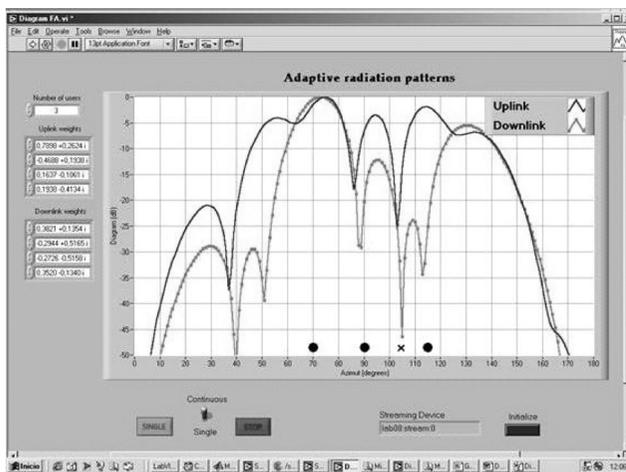


FIGURE 26: Synthesized beamformed radiation patterns (•: mobile users, ×: external interference). Downlink diagram is shown for user in azimuth 70°.

(modem, synchronization, and beamforming), although an overall description of the system has been presented.

Simulation and measurement results show the feasibility and performance achieved, which outperforms the capacity obtained with a typical sectored antenna, especially in a mixed-service scenario. In the case of asymmetric traffic scenarios, ADAM prototype provides in average a 100% total throughput increase per sector in comparison with the performance of a conventional sectored antenna, as it is shown in the results of Section 6.1.3.

Further research work is currently being done on integration with RF stages. Also, the analysis and definition of measurement procedures are currently being carried out. These

procedures will provide a characterization of ADAM performance in anechoic chambers and outdoor environments in connection with real cellular base stations. Moreover, a similar prototype for global system for mobile communications (GSM) standard (under the framework of Enhanced-GSM adaptive Antenna (EVA) project) is currently under development because our final objective is the design of a dual smart antenna for UMTS and GSM systems.

## ACKNOWLEDGMENTS

The authors wish to thank Dragados Sistemas, S.A., and Vodafone for their support in the development of the ADAM project. They would also like to thank the Spanish Ministry of Science and Technology (MCYT) for the financial support of the MAIN project and all the professors and researchers involved in the ADAM project, especially, Professor Manuel Sierra-Pérez.

## REFERENCES

- [1] L. C. Godara, "Application of antenna arrays to mobile communications. I. Performance improvement, feasibility, and system considerations," *Proceedings of the IEEE*, vol. 85, no. 7, pp. 1031–1060, 1997.
- [2] Third Generation Partnership Project (3GPP) Specifications, <ftp://ftp.3gpp.org/Specs/latest>.
- [3] A. O. Boukalov and S.-G. Häggman, "System aspects of smart-antenna technology in cellular wireless communications—an overview," *IEEE Transactions on Microwave Theory and Techniques*, vol. 48, no. 6, pp. 919–929, 2000.
- [4] M. Beach, C. Simmonds, P. Howard, and P. Darwood, "European smart antenna test-bed—field trial results," *IEICE Transactions on Communications*, vol. E84-B, no. 9, pp. 2348–2356, 2001.
- [5] R. Arnott and S. Ponnekanti, "Experimental investigation of adaptive antennas in a real DCS-1800 mobile network," in *Proc. 3rd ACTS Mobile Communications Summit*, pp. 535–543, Rhodes, Greece, June 1998.
- [6] ArrayComm Corporation, [www.arraycomm.com](http://www.arraycomm.com).
- [7] Kyocera Corporation, [www.kyocera.com](http://www.kyocera.com).
- [8] S. Anderson, B. Hagerman, H. Dam, et al., "Adaptive antennas for GSM and TDMA systems," *IEEE Personal Communications*, vol. 6, no. 3, pp. 74–86, 1999.
- [9] J. Mitola III, *Software Radio Architecture: Object-Oriented Approaches to Wireless Systems Engineering*, John Wiley & Sons, New York, NY, USA, 2000.
- [10] E. Buracchini, "The software radio concept," *IEEE Communications Magazine*, vol. 38, no. 9, pp. 138–143, 2000.
- [11] G. Finlay, "Understanding SDR requirements," *Wireless Systems Design*, vol. 6, no. 7, pp. 43–50, 2001.
- [12] Metawave Communications Corporation, [www.metawave.com](http://www.metawave.com).
- [13] Marconi Corporation, [www.marconi.com](http://www.marconi.com).
- [14] M. Calvo, V. Burillo, L. de Haro, and J. M. Hernando, *Sistemas de Comunicaciones Móviles de 3ª Generación (UMTS)*, Fundación Airtel Vodafone, Madrid, Spain, 2002.
- [15] H. Holma and A. Toskala, *WCDMA for UMTS*, John Wiley & Sons, Chichester, UK, 2nd edition, 2001.
- [16] M. Sierra, M. Calvo, L. de Haro, et al., "Modular smart antenna multi-standard for multi-operator cellular communications scenarios," Patent no. P200102780, Spain, 2001.

- [17] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, Englewood Cliffs, NJ, USA, 3rd edition, 1996, Chapters 9 and 13.
- [18] L. C. Godara, "Application of antenna arrays to mobile communications. II. Beam-forming and direction-of-arrival considerations," *Proceedings of the IEEE*, vol. 85, no. 8, pp. 1195–1245, 1997.
- [19] H. Lim and K. Cheun, "Analysis of decimator-based full-digital delay-locked PN code tracking loops for bandlimited direct-sequence spread-spectrum signals in AWGN," *IEICE Transactions on Communications*, vol. E81-B, no. 10, pp. 1903–1911, 1998.
- [20] T. F. Wong, "Spreading code acquisition and tracking," in *Spread spectrum and CDMA (EEL6503)*, University of Florida, Fla, USA, 2002, Chapter 5, <http://wireless.ece.ufl.edu/eel6503/>.
- [21] T.-Y. Cheng and K.-C. Chen, "Joint synchronization and data detection digital receiver for direct sequence spread spectrum systems," in *IEEE 47th Vehicular Technology Conference*, vol. 3, pp. 2103–2107, Phoenix, Ariz, USA, May 1997.
- [22] Sistemas Radiantes F. Moyano., *UMTS (1900-2200 MHz) BTS Antennas*, [www.moyano.com](http://www.moyano.com).
- [23] Pentek operating manuals, [www.pentek.com](http://www.pentek.com).
- [24] Mercury Computer Systems, "RACE++ Series, Eight-Port crossbar," [www.mc.com](http://www.mc.com).
- [25] T. Ojanperä and R. Prasad, *Wideband CDMA for Third Generation Mobile Communications*, Artech-House Publishers, Boston, Mass, USA, 1998.
- [26] C. A. Balanis, *Antenna Theory: Analysis and Design*, John Wiley & Sons, New York, NY, USA, 2nd edition, 1997.
- [27] H. Dai and L. Mailaender, "Performance comparison of CDMA downlink transmission techniques with power control," Technical memorandum, Bell Labs, Lucent Technologies, 2000.
- [28] Texas Instruments, *TMS320C6000 Optimizing Compiler User's Guide*, [www.ti.com](http://www.ti.com).
- [29] K. I. Pedersen, P. E. Mogensen, and B. H. Fleury, "A stochastic model of the temporal and azimuthal dispersion seen at the base station in outdoor propagation environments," *IEEE Trans. Vehicular Technology*, vol. 49, no. 2, pp. 437–447, 2000.
- [30] A. Capone, M. Cesana, G. D'Onofrio, and L. Fratta, "Mixed traffic in UMTS downlink," *IEEE Microwave and Wireless Components Letters*, vol. 13, no. 8, pp. 299–301, 2003.
- [31] J. Laiho, A. Wacker, and T. Novosad, Eds., *Radio Network Planning and Optimisation for UMTS*, John Wiley & Sons, Chichester, UK, 2002.
- [32] R. L. Cupo and R. D. Gitlin, "Adaptive carrier recovery systems for digital data communications receivers," *IEEE Journal on Selected Areas in Communications*, vol. 7, no. 9, pp. 1328–1339, 1989.

**Ramón Martínez Rodríguez-Osorio** was born in Madrid, Spain, in 1975. He received the Ingeniero de Telecomunicación degree in 1999 from the E.T.S.I. Telecomunicación, Universidad Politécnica de Madrid (UPM). In 1999, he joined the Departamento de Señales, Sistemas y Radiocomunicaciones at the same University, where he works as Assistant Professor and is currently working towards the Doctor Ingeniero de Telecomunicación (Ph.D.) degree. He participated in COST286 (European Project), working on the simulation of communication networks, and in other public and private funding research projects. He has also been involved in several other research projects related to DSP-based modems and smart antennas for UMTS, and in the



development simulation platforms for mobile communications systems. He has also participated in field measurement campaigns in actual TV interference scenarios. His other research activities focus on the area of adaptive beamformer implementation for UMTS, and on the impulsive noise interference modelling and channel emulation, studying its impact on communications systems. He has published chapters in two books, and he has also contributed in a number of international conferences and journals.

**Laura García García** was born in Toledo, Spain, in 1978. She received the Ingeniero de Telecomunicación degree in 2001 from the E.T.S.I. Telecomunicación, Universidad Politécnica de Madrid (UPM). In that year, she joined the Departamento de Señales, Sistemas y Radiocomunicaciones at the same University, where she is currently working towards the Doctor Ingeniero de Telecomunicación (Ph.D.) degree. She has worked in several research projects related to smart antennas and their DSP-based implementation. Her other research interests are in the study and implementation of efficient algorithms for smart antennas, specially focused on DSP-based real-time systems.



**Alberto Martínez Ollero** was born in Madrid, Spain, in 1978. He received the Ingeniero de Telecomunicación degree from the E.T.S.I. Telecomunicación, Universidad Politécnica de Madrid (UPM) in 2002. In that year, he joined the Departamento de Señales, Sistemas y Radiocomunicaciones, UPM, where he is currently working in several research projects. His interests include the development of signal processing algorithms for smart antennas and W-CDMA systems, implementation of software radio modules on DSP platforms for real-time applications and, antenna design for UMTS and DVB-S systems.



**Francisco Javier García-Madrid Velázquez** was born in Ciudad Real, Spain, on May 7, 1979. He received the Ingeniero de Telecomunicación degree from the Universidad Politécnica de Madrid (UPM) in 2002. In that year, he joined the Departamento de Señales, Sistemas y Radiocomunicaciones, UPM, where he is currently working toward the Ph.D. degree at the Escuela Técnica Superior de Ingenieros de Telecomunicación (E.T.S.I. Telecomunicación). His research interests include the implementation of signal processing algorithms for smart antennas, studies of impulsive noise effects in communications systems, and fuzzy logic applied to channel estimation and beamforming.



**Leandro de Haro Ariet** received the Ingeniero de Telecomunicación degree in 1986 and the Doctor Ingeniero de Telecomunicación degree (*Apto cum laude*) in 1992, both from the E.T.S.I. Telecomunicación, Universidad Politécnica de Madrid (UPM) (Departamento de Señales, Sistemas y Radiocomunicaciones). Since 1990, he has developed his professional career in the Departamento de Señales, Sistemas y Radiocomunicaciones as Profesor Titular de Universidad in the signal



theory and communications area. His research activity covers the following topics: antenna design for satellite communications (earth stations and satellite on board); study and design of satellite communication systems; and study and design of digital TV communication systems. He has been actively involved in several official projects and with private companies (national and international). He has also been involved in several European projects (RACE, ACTS, COST). The results of his research activity may be found in several presentations in national and international conferences as well as in published papers.

**Miguel Calvo Ramón** was born in Pueyo de Jaca, Huesca, Spain, on June 10, 1949. He received the Ingeniero de Telecomunicación degree from the E.T.S.I. Telecomunicación, Universidad Politécnica de Madrid (UPM) in 1974 and the Doctor Ingeniero de Telecomunicación degree (PhD) from the same University in 1979. He works as a Catedrático (Full Professor) in the Departamento de Señales Sistemas y Radiocomunicaciones since 1986. Since his incorporation to UPM in 1974, he has worked in a number of projects related to numerical methods in electromagnetics, EMC, antennas and testbed simulators for satellite communications (Hispasat). He has worked in the coordination procedures of the Spanish Hispasat satellite system with Intelsat, and part-time as a Technical Director of the Space Division at RYMSA. He has also worked as an Evaluator of Proposals in the framework of the European IST program. He was a Research Visitor at Queen Mary College, London University, in 1983 and Technical Visitor at Nichols Centre, Kansas University in Lawrence, in 1993. He has coauthored a number of papers in technical journals and contributed in a number of international conferences. He wrote a chapter in the book *Reflector and Lens Antennas: Analysis and Design Using Personal Computers* (C. J. Sletten, Ed., Artech House Publishers, 1988).



# A Multiple-Antenna System for ISM-Band Transmission

## J. Rinas

*Department of Communications Engineering, University of Bremen, 28359 Bremen, Germany*  
Email: rinas@ant.uni-bremen.de

## R. Seeger

*Department of Communications Engineering, University of Bremen, 28359 Bremen, Germany*  
Email: seeger@ant.uni-bremen.de

## L. Brötje

*Department of Communications Engineering, University of Bremen, 28359 Bremen, Germany*  
Email: broetje@ant.uni-bremen.de

## S. Vogeler

*Department of Communications Engineering, University of Bremen, 28359 Bremen, Germany*  
Email: vogeler@ant.uni-bremen.de

## T. Haase

*Department of Communications Engineering, University of Bremen, 28359 Bremen, Germany*  
Email: haase@zarm.uni-bremen.de

## K.-D. Kammeyer

*Department of Communications Engineering, University of Bremen, 28359 Bremen, Germany*  
Email: kammeyer@ant.uni-bremen.de

Received 23 June 2003; Revised 19 December 2003

We present a multiple antenna system for industrial, scientific, and medical (ISM)-band transmission (MASI). The hardware demonstrator was developed and realized at our institute. It enables multiple-input multiple-output (MIMO)-communication applications and is capable of transmitting arbitrary signals using 8 transmit and 8 receive antennas in parallel. It operates in the 2.4 GHz ISM-band. The hardware concept is introduced and some design specifications are discussed. Using this transmission system, we present some measurement results to show the feasibility of MIMO concepts currently under discussion. The applications include transmit and receive diversity for single carrier and OFDM as well as blind source separation (BSS) techniques.

**Keywords and phrases:** hardware demonstrator, MIMO, OFDM, Alamouti, blind source separation.

## 1. INTRODUCTION

One impetus to build a MIMO hardware demonstrator is that the assumptions made about real channels may be incorrect, and the behavior of MIMO systems should be investigated under realistic conditions. Therefore it is sufficient to transmit and receive over a real channel and process the received data off-line at the workstation environment. This basic idea roots in [1] where a single antenna system was realized at the University of Bremen. Furthermore, off-line processing significantly reduces the complexity of the simulator. In contrast to a real-time simulator, which is based on sub-optimal frontend processing (due to strict timing constraints

in connection with limited performance of DSP or FPGA chips) [2, 3, 4, 5], this concept has enabled us to freely investigate optimal and suboptimal algorithm implementations.<sup>1</sup> On the other hand, we do not claim to substitute a MIMO channel sounder [6]. A channel sounder is a highly accurate measurement system to precisely acquire the (MIMO) channel parameters. This requires extraordinary effort on, for example, calibrated and synchronized time bases at the trans-

---

<sup>1</sup>Assuming that we have an optimal algorithm in idiosyncratic sense, we can neglect implementation issues (quantization errors) on a double-precision machine.

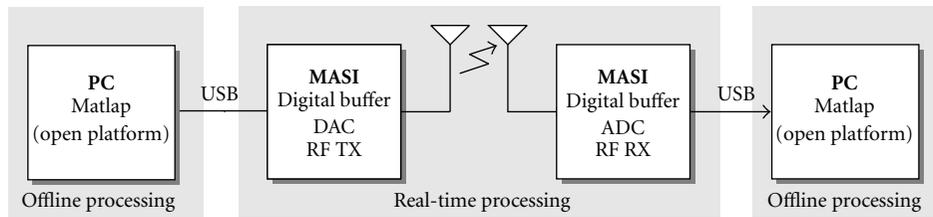


FIGURE 1: Principal block diagram.

mitter and receiver, highly linear frontend amplifiers, and calibrated antenna arrays. In contrast, the objective of our demonstrator is to evaluate MIMO algorithms under non-idealized environments deploying common hardware components. Moreover, thanks to selectable frontend processing, we can handle arbitrary radio interface standards, such as single carrier, multicarrier, and spread spectrum MIMO systems.

## 2. HARDWARE CONCEPT

### 2.1. Top-level system description

The top-level system is diagrammed in Figure 1. At the workstation environment, in-phase and quadrature (I/Q) data, for example, Hiperlan/2 or UMTS frames, are generated by the simulation system of choice. The impulse shaping is done in the digital domain. The data is scaled and quantized to meet the hardware demonstrator concerns and finally stored into a file. Due to its wide distribution, the USB interface is chosen to connect the hardware demonstrator with the workstation. To transfer the I/Q data via the USB interface, we use a customized application software which allows us to set several parameters, like sample rate (from external or internal clock), local oscillator (LO) frequency tuning value, and assignment of data files to corresponding antennas. Furthermore, in a Matlab environment, we can directly access the demonstrator by calling a Matlab function [7]. This is useful for fully automated measurements. Inside the demonstrator, the I/Q data is stored into digital buffers which are addressed in a circular manner: the increment pointers for memory accesses wrap to the beginning of the buffer when its end is reached. The currently addressed I/Q words are fed to a digital-to-analog converter (DAC), whose analog baseband output signals drive the radio frequency (RF) stage, which performs up-conversion to the desired RF band.

At the receiver, the RF passband signal is down-converted to the complex baseband and undergoes analog-to-digital conversion. A snapshot is stored into a digital buffer. Because frame synchronization is not implemented in hardware, the receive buffer has doubled length of the transmit buffer to ensure that at least one complete frame is captured. The sample rate is adjustable up to 80 MHz and may be chosen from a set of internally predefined frequencies or an external source. The request for extensibility of the hardware demonstrator led to a full modular architecture; for each antenna, the connected transmitter or receiver hardware has its own plug-in

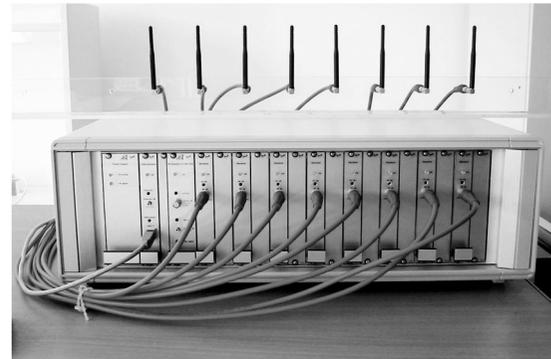


FIGURE 2: The multiple antenna receiver for ISM-band transmission. Currently, the receiver and transmitter are equipped with 8 modules.

module (see Figure 2). The digital clock and LO signal is provided to all modules by a central clock base to ensure inter-module synchronization of sample rate and carrier phase.

Low-cost software radios are the main driver for modern radio architectures (universal receivers that can accommodate many different standards). Consequently, this type of receiver gains increased attention. An all-digital receiver performs all its operations in the digital domain, except the frontend baseband translation and antialiasing filtering. Its ADC sampling clock is not synchronized to the transmitter symbol clock. Therefore, many analog components, such as the voltage-controlled oscillator (VCO), are not required. Thus, it can be smaller, more robust, and less expensive. However, as a fixed sampling clock is used which is not synchronized to the transmitter clock, symbol timing and carrier recovery have to be accomplished in the digital domain. In order to reduce analog component count in the RF stage, the direct conversion (or homodyne) architecture is implemented, which performs passband-to-baseband translation and vice versa directly without intermediate frequency (IF) stages. Traditionally, the direct conversion architecture was considered impractical due to severe realization problems. So far, it was hardly possible to fulfill all requirements like exceptionally linear low-noise amplifier (LNA) and mixer circuits, as well as the LO isolation resulting in a lower sensitivity compared to heterodyne receivers [8]. However, recent advances in chip technology enabled robust direct conversion frontends. In the next section, we will discuss the employed components and some important parameters in a more detailed manner.

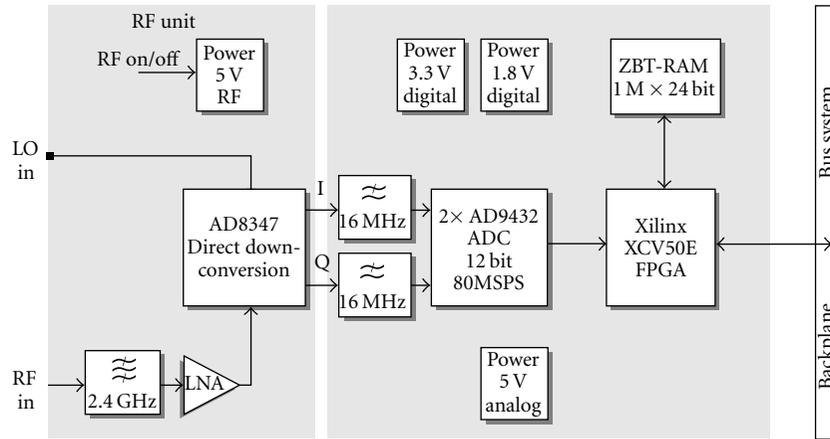


FIGURE 3: Receiver module.

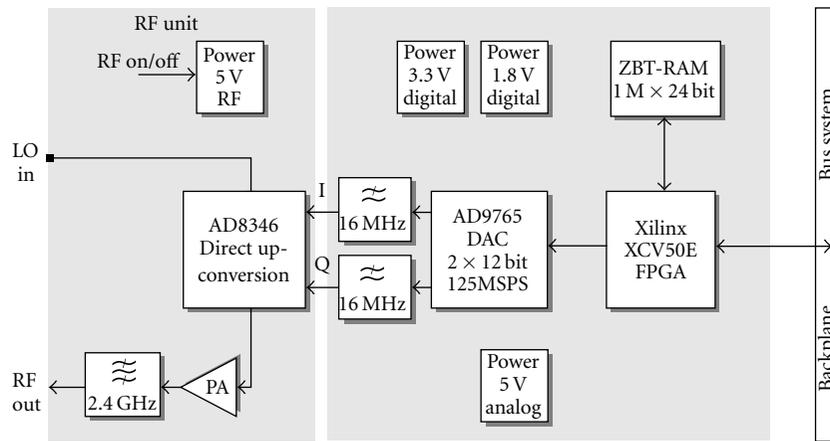


FIGURE 4: Transmitter module.

**2.2. Detailed description of components**

The direct conversion architecture leads to very simple RF designs (Figures 3 and 4). Extra IF stages with amplifiers, passive bandpass filters, and oscillators are omitted, as this simplifies the board design and reduces power dissipation. Furthermore, due to zero IF, the image rejection problem does not exist.<sup>2</sup> All subsequent processing can take place at the lowest possible frequency which makes the direct conversion scheme amenable to integrated circuit (IC) implementation. Applying this architecture, we are restricted to complex baseband processing which halves the signal bandwidth but doubles the component count in comparison to a passband scheme.

**2.2.1. Low-noise amplifier**

The first stage of the receiver is an LNA, whose main function is to provide enough gain to overcome the noise of sub-

sequent stages (such as the mixer). Aside from this providing gain while adding as little noise as possible, an LNA should accommodate large signals without distortion. It must also present an impedance of 50Ω to the input source since the transfer function of the preceding filter is quite sensitive to the quality of termination. The employed LNA chip has a gain of 22 dB and a noise figure (NF) of 1.6 dB at 2.4 GHz. A relatively high 1 dB compression point (the input power at which the gain is 1 dB less than expected) of 4.2 dBm and a high third-order intercept point (IP3) ensures wide range linear operation.

**2.2.2. Mixer**

Since, for direct conversion architectures, the LO frequency lies in the desired frequency band, the LO signal, which normally has much more power than the received signal, can leak into the RF input of the mixer or possibly find its way to the antenna. The self-mixed LO signal results in a time-invariant DC baseband component, which can drive subsequent stages into saturation. In addition, any even-order distortion produces a DC offset that is signal-dependent, so the second-order intercept point (IP2) is a very important parameter for direct conversion schemes. The employed IC quadrature demodulator has two integrated Gilbert (or four-quadrant) cell

<sup>2</sup>In a heterodyne receiver, the first IF is normally chosen relatively high to move the image far away from the desired signal in order to relax the frontend bandpass filter requirements. A direct conversion receiver does not need a frontend filter, however, it is practically needed to avoid out-of-band interferers overloading the frontend [8].

mixers. This mixer style provides reasonable conversion gain (IF power output with respect to the RF power input), as well as good rejection at the RF and LO input ports and the IF output port due to the complete differential design.

External amplifiers are omitted due to integrated RF and baseband AGC amplifiers, which provide about 70 dB gain control. A high dynamic range is indispensable for wireless application. The baseband I/Q output ports allow direct connection to the ADCs.

### 2.2.3. Analog-to-digital conversion

The analog-to-digital converter (ADC) converts the continuous-time stimuli signals to discrete-time binary-code form. For communications applications, the dynamic measures of an ADC, such as signal-to-noise ratio (SNR), spurious-free dynamic range (SFDR), and two-tone intermodulation distortion (IMD), are figures of merit [9]. The effective number of bits (accuracy) depends strongly on these dynamic measurements. High-speed ADCs are extremely sensitive to the quality of the sampling clock. The internal track-and-hold circuit is essentially a mixer. Any noise, distortion, or timing jitter on the clock signal will be combined with the desired signal at the ADC output in addition to internal timing error sources (aperture jitter). A phase-locked loop (PLL)-based synthesizer normally exhibits a higher phase noise value than a fixed frequency clock generator. However, to provide several customized sample rates, a set of stable crystal-controlled oscillator circuits is used. Furthermore, an external clock input up to 80 MHz is available. The chosen 12 bit ADC chip delivers good dynamic measurements, a low-aperture jitter, and was available at small quantities. The digital outputs (I and Q branches) are directly connected to the digital buffer circuit.

### 2.2.4. Digital buffer

The digital buffer stores the raw data, delivered by the ADC (receiver) or provided by the USB controller (transmitter). At the transmitter, the digital buffer serves as a circular buffer. Once the data is completely stored, the buffer is linearly addressed; when the last address is reached, the address counter wraps around to the first address and counts up again, whereas at the receiver, only one frame is captured when the trigger event occurs. Because large FIFO chips are very expensive and hardly obtainable at small quantities, the digital buffer circuit is realized by a field-programmable gate array (FPGA) and static RAM (SRAM). In contrast to dynamic RAM, SRAM does not need refresh cycles and offers a considerably simpler interface. The employed zero bus turnaround (ZBT) RAMs are fast synchronous SRAM chips which are directly connected to the FPGA. Providing interleaved read/write without wasteful turnaround cycles, the ZBT RAM is predestined for capturing applications. Once primed with an address, it can read/write one word of data per clock cycle. Up to  $2^{20}$  samples can be captured per in-phase and quadrature branch. The FPGA connects all digital busses and provides several control signals. Due to the ability of reconfiguration, it offers a high degree of flexibility. It also provides enough resources to hold optional customized

frontend processing logic, like frame detection algorithms.<sup>3</sup> The logic blocks are described at a high abstraction level using VHDL.<sup>4</sup>

## 3. MEASUREMENTS AND APPLICATIONS

The measurements were performed in an indoor environment, that is, we transmitted between two adjacent office rooms of approximately 20 m<sup>2</sup> size each. The total transmit power was 17 dBm (50 mW).

### 3.1. Frame synchronization

Our system works without any wired connection between the transmitting and receiving ends. Therefore we have to synchronize both sides. We transmit periodically repeated frames with  $L_t$  samples. In order to get at least one complete frame, we sample  $L_r = 2 \cdot L_t$  values at the receiving side.

The first task is the detection of the starting point of one complete frame within these  $L_r$  samples. Therefore we apply a simple power detection scheme, which presents a pragmatic approach to our measurement system, because it is mostly independent from impairments like frequency offset and frequency-selective channels, and can be used with any modulation scheme. For the power detection, we normally consider about  $L_z = 1000$  samples within one frame of length  $L_t$ . The high variation of the envelope of the signal is unproblematic since we are using a very slow AGC.

Our synchronization approach is a sliding power detection. We detect the current power of the received signal  $r(k)$  (one channel) by averaging over  $L_z$  successive samples of both gaps (Figure 5):

$$\begin{aligned} k_{\text{start}} &= \arg \min_k p(k) \\ &= \arg \min_k \sum_{\kappa=k}^{k+L_z-1} \frac{1}{2L_z} \left( |r(\kappa)|^2 + |r(\kappa + L_t)|^2 \right) \end{aligned} \quad (1)$$

with  $k = 0 \dots L_t - L_z$ .

This approach for a coarse frame synchronization is not necessarily limited to MIMO setups but can also be used for single input single output (SISO) channels. An example for this scheme is presented in Figure 5, where you can see time series of a measurement including the detection of the complete frame.

### 3.2. Frequency responses

In this section, we will present a setup for measuring the frequency response of the MIMO transmission channel, which we always consider from the digital domain at the transmitter to the digital domain at the receiver—including all effects

<sup>3</sup>The physical memory (ZBT RAM) has identical size, but the address logic of the circular buffer is programmed according to user settings. Notional frame synchronization could be implemented in hardware. Thus, the full physical buffer size could be used at transmitter, however, with the drawback of a fixed preamble or frame structure.

<sup>4</sup>Very high-speed integrated circuit (VHSIC) hardware description language (VHDL).

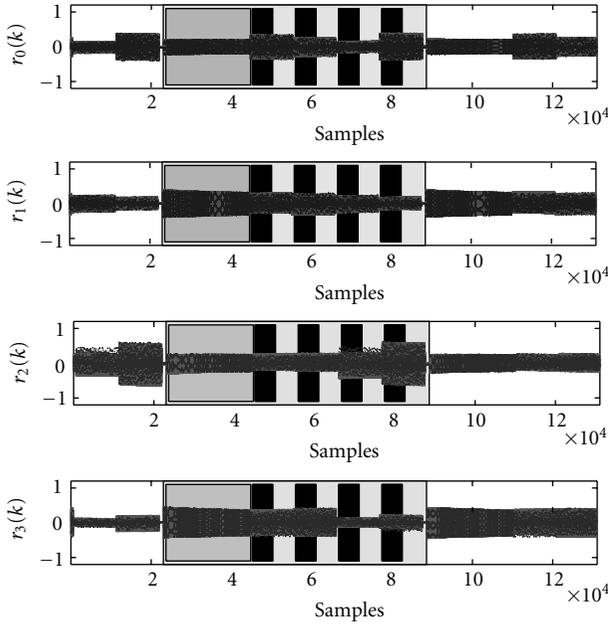


FIGURE 5: Measured signals with frame synchronization.  $f_{\text{off}} = 762.9$  Hz

of the system components. We have to emphasize that it is not our intention to do systematic channel measurements.

For measurements, we apply a chirp-like signal, whereas only one transmitter is sending at a time, in order to measure the complete matrix of frequency responses (Figure 6).

This signal is designed in the frequency domain as

$$M(n) = e^{-j(\pi/N_{\text{DFT}})n^2} \quad \text{for } n = 0 \dots N_{\text{DFT}} - 1, \quad (2)$$

because this guarantees an exactly flat magnitude. Processing the IDFT, we get the time-domain signal

$$m(k) = \text{IDFT}_{N_{\text{DFT}}}\{M(n)\} \quad (3)$$

which is inherently periodic. We exploit this property and send  $m(k)$  in a periodic way so that only a coarse synchronization is necessary.

The quadratic phase increment leads to a small crest factor<sup>5</sup> of the signal. In our case, with  $N_{\text{FFT}} = 128$ , the crest factor for the imaginary part of the signal  $m(k)$  is

$$c_{\text{imag}} = \frac{\max \text{imag}\{m(k)\}}{\sqrt{(1/N_{\text{DFT}}) \sum_{k=0}^{N_{\text{DFT}}-1} \text{imag}\{m(k)\}^2}} \approx 1.47. \quad (4)$$

We can measure the frequency response, up to a linear phase uncertainty, by using a fractional part of the received time signal with  $N_{\text{DFT}}$  samples and calculating

$$\begin{aligned} R(n) &= \text{DFT}_{N_{\text{DFT}}}\{r(k_{\text{offset}} + k)\}, \quad k = 0 \dots N_{\text{DFT}} - 1, \\ H(n) &= \frac{R(n)}{M(n)}. \end{aligned} \quad (5)$$

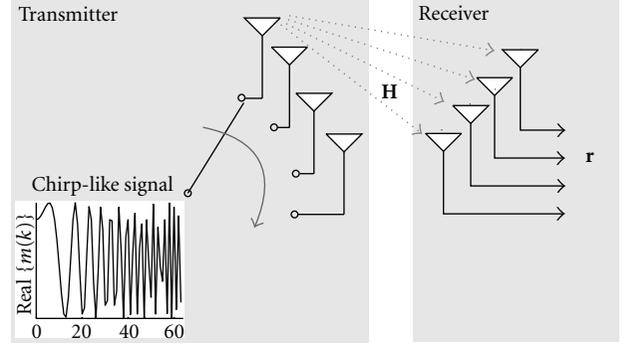


FIGURE 6: Multiplexing for channel measurement.

Figure 5 shows the time series of one measurement. Notice the different amplitudes of the signal that correspond to one constellation of the multiplexing scheme (Figure 6). Since this method is sensitive regarding the frequency offset, we added a pilot sequence to our measurement frame in order to estimate and correct the offset.

The advantage of this approach is that we only need a coarse synchronization and not a high-precision time reference (like in channel sounding setups). Therefore the starting position  $k_{\text{offset}}$  may be slightly inaccurate. This circular<sup>6</sup> time shift of the starting position will result in a linear phase term, but it does not influence the shape of the magnitude response:

$$\begin{aligned} H_{\text{shift}}(n) &= \frac{\text{DFT}_{N_{\text{DFT}}}\{r(k + k_{\text{shift}})\}}{M(n)} \\ &= H(n)e^{j(2\pi/N_{\text{DFT}})nk_{\text{shift}}}. \end{aligned} \quad (6)$$

Figure 7 depicts three different frequency response measurements using 4 transmit and 4 receive antennas. Uniform linear arrays (ULAs) with  $\lambda/2$ -spaced elements are used. The sampling frequency was set to  $f_s = 50$  MHz.

One can directly see the filter influence of our transmissions system, which limits the signal to the 3 dB range of approximately  $\pm 16$  MHz. In addition, there are some notches in the spectrum which arise from a frequency-selective channel. Our measurements already revealed that a small change of the position may have a strong impact on the frequency response.

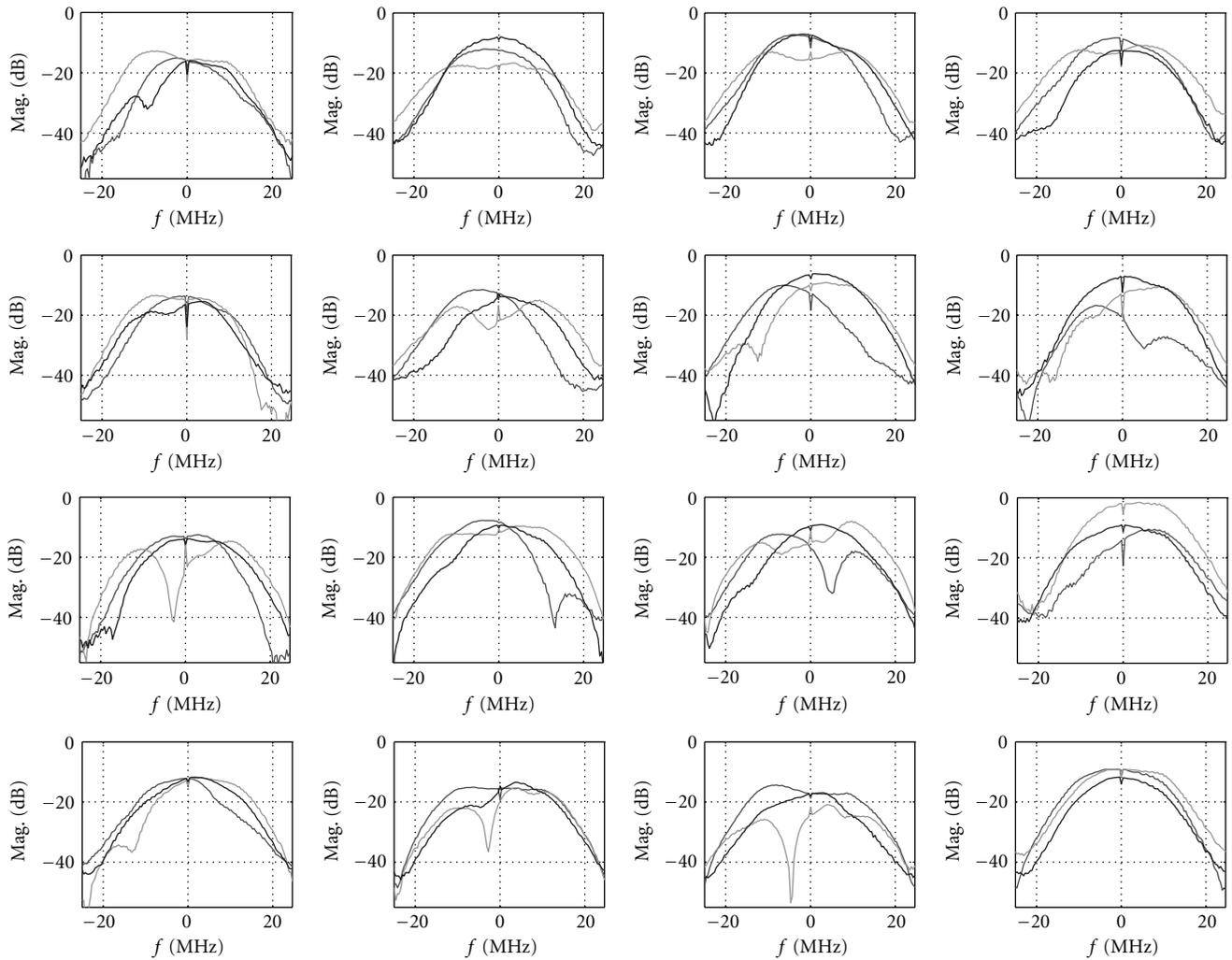
### 3.3. Diversity techniques

There are two principal approaches to get a performance gain from an antenna array. One approach uses the known geometric constellation of the antennas for beamforming. The other approach is independent of the array constellation and increases the diversity of the system.

In this section, we focus on diversity techniques. Diversity through a multiantenna setup can be attained at the receiving as well as at the transmitting end.

<sup>5</sup>The peak-to-rms voltage ratio of an alternating current (AC) signal.

<sup>6</sup>Since we are using a periodic repeated signal, we can interpret a time shift as a periodic time shift.

FIGURE 7: Frequency responses for a  $4 \times 4$  setup.

### 3.3.1. Receive diversity

In order to achieve a diversity gain at the receiving side, we can expand a SISO setup and use multiple receive antennas. The diversity combining can be done in a blind way by using the spatial covariance of the received signal streams. Timing offset is estimated after combining using the approach presented in [10].

For combining, we have to take into account that our system has independent AGCs in each channel. Therefore we have to estimate the noise level, which is done by exploiting the power gap in the frame of the received signal.

In order to show the gain of a combining, we sent one QPSK signal and received it with multiple antennas. Figure 8 depicts the signal constellations of a measurement. On the left-hand side, you can see the signal constellations received from each antenna, while the right-hand side depicts the combining of the signals received by antenna 1 up to 4. The rising SNRs for increasing number of signals involved in the combining process indicate the combining gain.

SNR estimation is done using the approach presented in [11], because it does not suffer from wrong symbol decisions and is suitable without modification for all PSK schemes. The SNR of a single data stream  $y$  is calculated by

$$p = \sqrt{2 - \frac{E\{|y|^4\}}{E\{|y|^2\}^2}}, \quad (7)$$

$$\text{SNR} = \frac{p}{1-p}.$$

### 3.3.2. Transmit diversity

In theory, receive diversity and transmit diversity are interchangeable. In the following, we will discuss transmit diversity schemes, especially the so-called orthogonal space-time block codes (OSTBC), under more realistic conditions. Channel estimation and carrier offset estimation are essential tasks in coherent receivers. However, they are also some kind of error sources due to the imperfectness of the employed algorithm.

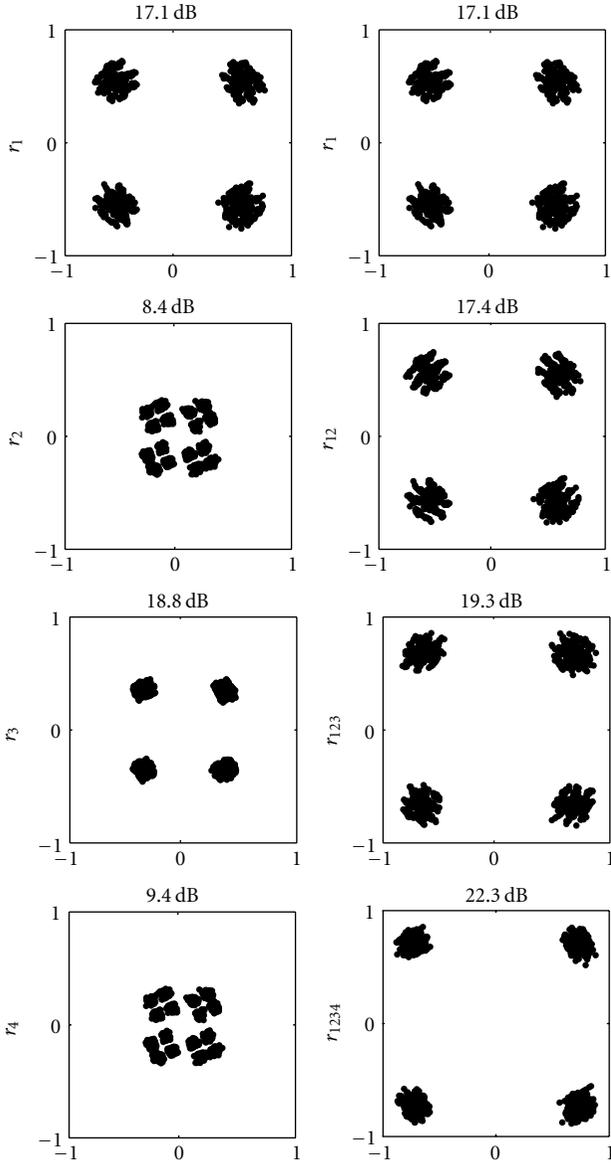


FIGURE 8: Combining gain with estimated SNR.

Alamouti [12] discovered a remarkable transmit diversity scheme for transmission with two antennas. This scheme supports maximum-likelihood detection based only on linear processing at the receiver and is able to achieve full diversity provided by the number of transmit and receive antennas. The input symbols to the ST block encoder are divided into groups of two symbols each,  $\{s_1, s_2\}$ . At a given symbol period,  $s_1$  and  $-s_2^*$  are transmitted from antenna 1 and 2, respectively, and at the consecutive symbol period,  $s_2$  and  $s_1^*$  are transmitted from antenna 1 and 2, respectively. Let  $h_1$  and  $h_2$  be the channel coefficients from the first and second transmit antennas, respectively. It is assumed that  $h_1$  and  $h_2$  are constant over two consecutive symbol periods. Consider a receiver with one receiver antenna and denote the received signals over two consecutive symbol periods as  $r_1$  and

$r_2$ . Defining the code symbol vector  $\mathbf{s} = [s_1 s_2^*]^T$  and the received vector  $\mathbf{r} = [r_1 r_2^*]^T$ , we get

$$\mathbf{r} = \mathbf{H}\mathbf{s} + \boldsymbol{\eta}, \quad (8)$$

where the channel matrix

$$\mathbf{H} = \begin{bmatrix} h_1 & -h_2 \\ h_2^* & h_1^* \end{bmatrix} \quad (9)$$

and the noise vector  $\boldsymbol{\eta} = [\eta_1 \eta_2]^T$  are used. The AWGN is represented by  $\eta_1$  and  $\eta_2$  which are modelled as i.i.d complex Gaussian random variables with zero mean and power spectral density  $N_0/2$  per dimension. Hence  $\boldsymbol{\eta}$  is a Gaussian random vector with zero mean and covariance  $N_0\mathbf{I}$ .

The decoding procedure consists of a simple multiplication with the Hermitian channel matrix  $\hat{\mathbf{H}}^H$ , hence

$$\hat{\mathbf{r}} = \hat{\mathbf{H}}^H \mathbf{H}\mathbf{s} + \hat{\mathbf{H}}^H \boldsymbol{\eta}, \quad (10)$$

where  $\hat{\mathbf{H}}$  is the *estimated* channel matrix. Considering imperfect channel estimation with an estimation error [13]

$$\Delta h = \Delta h_{\text{noise}} + \Delta h_{\text{Doppler}}, \quad (11)$$

it follows that

$$\hat{\mathbf{H}} = \begin{bmatrix} h_1 + \Delta h_1 & -(h_2 + \Delta h_2) \\ h_2^* + \Delta h_2^* & h_1^* + \Delta h_1^* \end{bmatrix}. \quad (12)$$

The (soft) decoded symbol-vector  $\hat{\mathbf{r}} = [\hat{r}_1 \hat{r}_2^*]^T$  can be obtained using (10) and (12):

$$\begin{aligned} \hat{\mathbf{r}} = & \underbrace{\begin{bmatrix} |h_1|^2 + |h_2|^2 & 0 \\ 0 & |h_1|^2 + |h_2|^2 \end{bmatrix}}_{\text{desired}} \mathbf{s} \\ & + \underbrace{\begin{bmatrix} h_1 \Delta h_1^* + h_2^* \Delta h_2 & -h_2 \Delta h_1^* + h_1^* \Delta h_2 \\ -h_1 \Delta h_2^* + h_2^* \Delta h_1 & h_2 \Delta h_2^* + h_1^* \Delta h_1 \end{bmatrix}}_{\text{influence of estimation errors}} \mathbf{s} \\ & + \underbrace{\hat{\mathbf{H}}^H \boldsymbol{\eta}}_{\text{noise}}. \end{aligned} \quad (13)$$

From (13), it is clear that channel estimation errors lead to spatial intersymbol interference (ISI) if the estimated channel matrix  $\hat{\mathbf{H}}$  is not unitary (12).

Another major task for coherent receivers is the carrier frequency offset estimation and correction. Consider two consecutive received symbols  $r_1$  and  $r_2$ . The frequency offset can be modeled by the time-domain multiplication with the two phasors  $e^{j\varphi_1}$  and  $e^{j\varphi_2}$ , respectively. Using the system model (8), it can be stated that

$$\begin{bmatrix} e^{j\varphi_1} & 0 \\ 0 & e^{-j\varphi_2} \end{bmatrix} \begin{bmatrix} r_1 \\ r_2^* \end{bmatrix} = \begin{bmatrix} e^{j\varphi_1} & 0 \\ 0 & e^{-j\varphi_2} \end{bmatrix} (\mathbf{H}\mathbf{s} + \boldsymbol{\eta}). \quad (14)$$

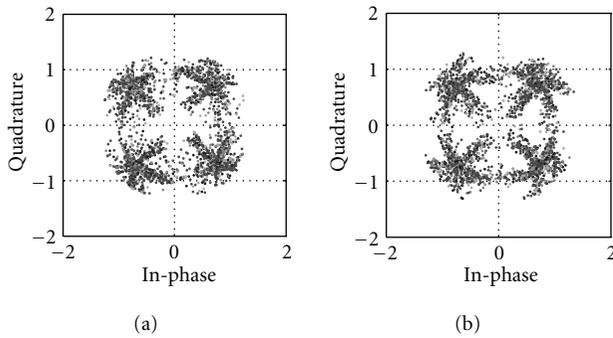


FIGURE 9: QPSK signal constellations at the STBC decoder output: (a) simulated signal, (b) measured signal.

Assuming perfect channel estimation conditions, that is,  $\hat{\mathbf{H}} = \mathbf{H}$ , and neglecting the noise term in (14), we obtain

$$\begin{aligned} & \begin{bmatrix} h_1^* & h_2 \\ -h_2^* & h_1 \end{bmatrix} \begin{bmatrix} e^{j\varphi_1} & 0 \\ 0 & e^{-j\varphi_2} \end{bmatrix} \begin{bmatrix} h_1 & -h_2 \\ h_2^* & h_1^* \end{bmatrix} \mathbf{s} \\ &= \begin{bmatrix} |h_1|^2 e^{j\varphi_1} + |h_2|^2 e^{-j\varphi_2} & -h_1^* h_2 e^{j\varphi_1} + h_2 h_1^* e^{-j\varphi_2} \\ -h_2^* h_1 e^{j\varphi_1} + h_1 h_2^* e^{-j\varphi_2} & |h_2|^2 e^{j\varphi_1} + |h_1|^2 e^{-j\varphi_2} \end{bmatrix} \mathbf{s}. \end{aligned} \quad (15)$$

In contrast to a single transmit antenna system, the loss of orthogonality due to a (residual) frequency offset leads to *magnitude* variations. A comparison of a simulated and a measured signal constellation with channel estimation and frequency offset is depicted by Figure 9.

### 3.4. OFDM transmission

Our simulation tool for OFDM transmission is based on the IEEE 802.11a WLAN standard [14], except for the carrier frequency of 2.4 GHz (instead of 5.2 GHz).

#### 3.4.1. Synchronization

##### Timing Synchronization

First of all, a coarse frame synchronization according to the method for single carrier systems already described (Section 3.1) is carried out. For OFDM transmission, there is no need to find the starting point of the burst exactly, because afterwards the position of the FFT window is adjusted in a second synchronization step.

Due to the cyclic prefix (CP) in every OFDM symbol, the exact position of the FFT window can be found by correlation over the received signal. This results in a well-defined maximum value for each OFDM symbol; the correct FFT window start position is  $N_{\text{guard}}$  samples later. Averaging OFDM symbols to suppress noise may be reasonable.

##### Carrier Frequency Synchronization

The correction of carrier frequency offsets (CFO) in OFDM systems can be carried out in two steps. A synchronization in time domain (before processing the FFT at the receiver) is absolutely necessary, because severe ISI occurs in frequency

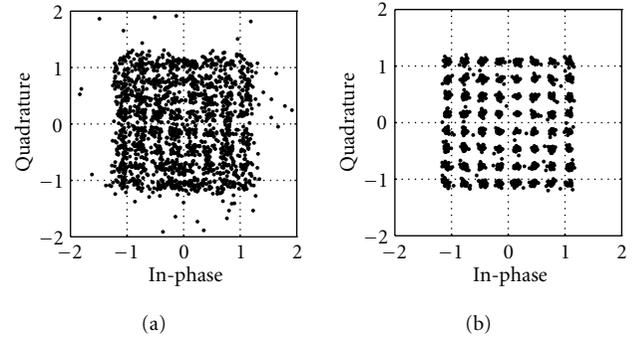


FIGURE 10: Impact of a DC offset (7 dB): signal constellation diagram (uncorrected/corrected).

domain if a CFO is not corrected before. In case of small frequency offsets (compared with the subcarrier spacing), the main effect after processing the FFT is a rotation of symbols regarding their signal constellation, so that in time domain a coarse synchronization is sufficient. This coarse estimation is accomplished by calculating the phase deviation between the two preamble C symbols [14].

A fine carrier frequency synchronization in frequency domain is based on the four pilot symbols which are included in every OFDM symbol and whose carrier positions are symmetric to the carrier with frequency  $f = 0$ . The pilot carriers are BPSK-modulated. To estimate the CFO from the pilot symbols, the channel coefficients according to these carriers have to be known. Because every OFDM symbol carries the pilot information, a tracking of the CFO estimation can easily be performed. For further details on our synchronization methods, see [15].

#### 3.4.2. Impact of a DC offset

A DC offset, for example, resulting from self-mixing of the oscillators as described in Section 2.2, is a serious problem when using direct conversion concepts. With regard to a transmission, we have to take into account different aspects. On the one hand, the coarse burst synchronization fails with signals having high DC offsets. Therefore, it is necessary to average the whole received sequence to get an estimate for the DC offset and to subtract it afterwards. On the other hand, assuming a correct synchronization, the impact of the DC offset at the receiver in frequency domain is, due to the rectangular windowing of the FFT, the same as an addition of a sinc function with the maximum at  $f = 0$  and zero crossing at all other subcarrier frequencies. That is why the DC subcarrier is unassigned in IEEE 802.11a. In fact, this is no solution, because, in combination with a carrier frequency offset, the DC offset affects *all* subcarriers. In this case, the sinc function's maximum is shifted by the value of the CFO and additionally the zero crossings move *between* the sampling points of the subcarriers.

Figure 10a shows the signal constellation diagram of a 54 Mbit/s data burst transmitted with our system at 2.4 GHz after equalization. The received signal contains a DC offset of

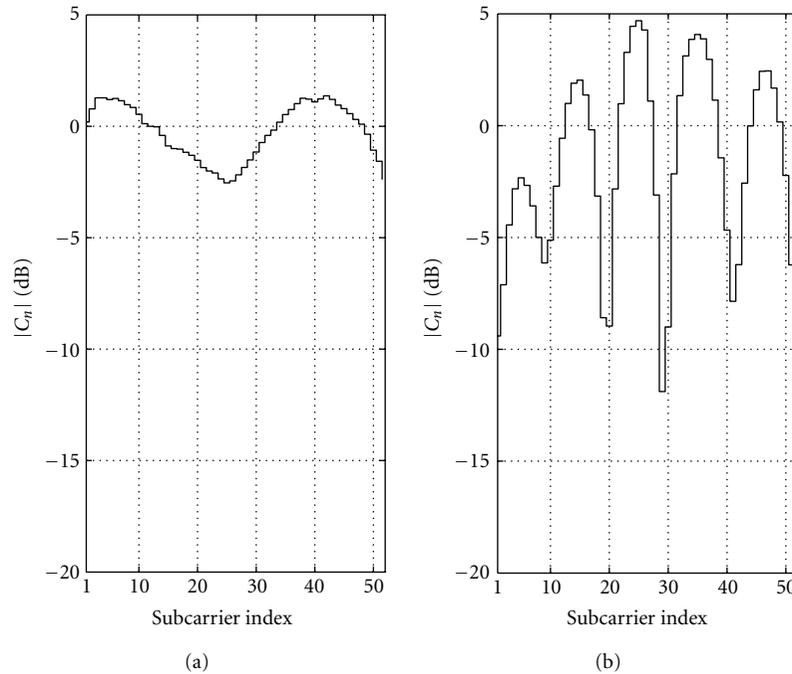


FIGURE 11: Measured channel transfer functions; (a) without CDD, (b) CDD: 2TX, delay 0.3 microsecond.

approximately 7 dB (ratio of DC magnitude to rms<sup>7</sup> of signal without DC) and a CFO of approximately 104 kHz.

When correcting the DC offset, an estimation based on averaging a certain number of preamble B symbols in time leads to sufficient results (see Figure 10b).

### 3.4.3. Transmit diversity schemes for OFDM

Considering transmit diversity schemes for IEEE 802.11a, one can distinguish between schemes which are compatible with the standard and those which are not. STBC belong to the latter ones. In contrast, delay diversity schemes need no modification of the receiver at all, thus they are fully standard conform.

#### Delay Diversity and Cyclic Delay Diversity

Delay diversity means transmitting the same OFDM symbol in time domain, including the CP, with a certain delay for each antenna. Due to synchronization constraints, the maximum delay is restricted to the remaining length of the CP, which is the total length of the CP minus the channel impulse response length. A better solution especially for OFDM systems is *cyclic* delay diversity (CDD), for example, known from [16, 17]. Using CDD, cyclically time-shifted OFDM symbols are *simultaneously* transmitted by each antenna. It is important to note that the signal is shifted *before* inserting the CP. Compared to noncyclic delay diversity, there is no strong restriction for the length of the delay. The allowed maximum length equals the FFT length.

The principle of all delay diversity schemes is to increase the length of the channel impulse response seen at the receiver, that is, the channel transfer function becomes more frequency selective. The added diversity is only exploited by the channel decoder [17]; in contrast to other transmit diversity schemes, there is no SNR enhancement. Because the superposition of the transmitted signals from each antenna at the receiver using delay diversity or CDD is equivalent to a single transmit antenna system with extended channel impulse length, no changes are necessary at the receiver.

Figure 11 shows the increased frequency selectivity due to CDD by means of two measured channel transfer functions at 2.4 GHz. In Figure 11b, a CDD system with 2 transmit antennas and a delay of 6 samples ( $\equiv 0.3 \mu\text{s}$ ) was used. For comparison, the single antenna case is presented in Figure 11a. The magnitudes of the channel coefficients for each subcarrier in the OFDM system were obtained by the estimation based on the IEEE 802.11a preamble.

Although there is no restriction for the delay length using CDD in theory, problems may occur if a noise reduction (NR) of the estimated channel transfer function by windowing in time domain [15] is carried out. In this case, the *increased* channel impulse length due to CDD has to be considered when fixing the NR window length. If the window length is too short, the channel impulse response will be falsified, which significantly reduces the performance of the NR.

#### STBC: Alamouti Scheme

The transmit diversity scheme proposed by Alamouti [12] is based on non-frequency-selective or flat-fading channel assumptions. Therefore, in case of OFDM, the coding of the

<sup>7</sup>Root mean square:  $\sqrt{1/N \sum_{k=1}^N |s(k)|^2}$ .

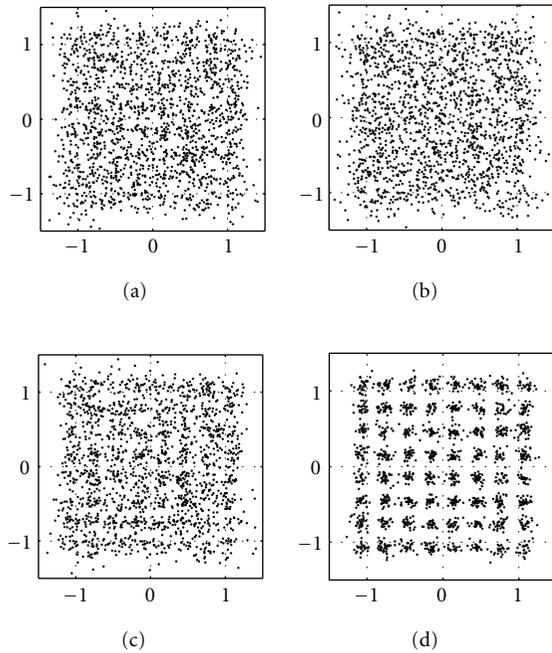


FIGURE 12: Signal constellations: (a) receive signal 1, (b) receive signal 2, (c) MRC of receive signals 1 and 2, (d) MRC of receive signals 1 to 4.

transmit symbols according to Alamouti has to be done in *frequency* domain, that is, before processing the IFFT at the transmitter, and decoding after applying the FFT at the receiver. So, in contrast to delay diversity schemes, two IFFT processing units are needed. Because the OFDM demodulation (FFT) is the inverse operation of the modulation (IFFT), the equations describing the Alamouti coding (8) remain unchanged for OFDM transmission, except for the transmit symbols  $s_i$  as well as the receive symbols  $r_i$  becoming OFDM symbols in frequency domain, that is, they consist of 52 (number of subcarriers) PSK or QAM symbols each.

In contrast to all delay diversity schemes, using the Alamouti transmit diversity scheme is *not* compatible to IEEE 802.11a. In addition to the modifications in the receiver according to the Alamouti decoding, a modification of the channel estimation and synchronization algorithms as well as a new preamble structure is necessary.

#### 3.4.4. Receive diversity scheme for OFDM

The application of receive diversity to a transmission system based on IEEE 802.11a can be realized without any changes to the transmitter, that is, absolutely standard conform. One possible method is maximum ratio combining (MRC), on which we will focus in the following.

##### Maximum Ratio Combining

In order to combine the received symbols according to the MRC principle, in frequency domain, the not-yet-equalized

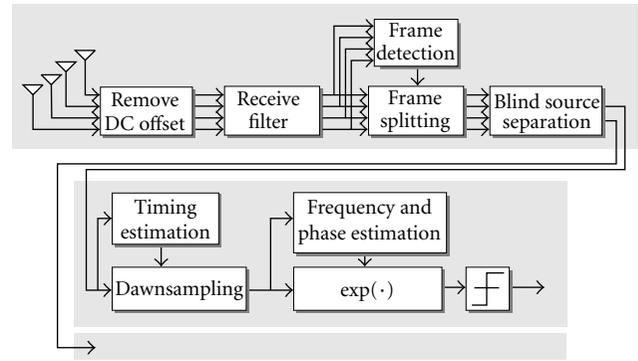


FIGURE 13: BSS setup.

values on each subcarrier are multiplied with the corresponding conjugate complex channel coefficient. The resulting values of all receive antennas are then added up separately for each subcarrier and afterwards, in case of QAM symbols, normalized to the sum of power of the channel coefficients resulting from the different antennas. After a soft-decision demapping, the weighted bits are multiplied with the inverse of the normalization factors used before.

A measurement example obtained with MASI can be seen in Figure 12. In that case, the BER for the signal received on the first and second receive antennas after channel decoding is 0.11 and 0.5, respectively. MRC of the two received signals (see Figure 12c) results in a reduction of the BER to  $7.97 \cdot 10^{-3}$ , whereas the combining of four receive signals (see Figure 12d), obtained with two additional receive antennas, leads to an error-free reception.

#### 3.5. Blind source separation

BSS algorithms are able to separate different signals from a multisensor setup. The only knowledge used to achieve this goal is that the signals should be statistically independent.

We choose the BSS setup in favor of classical pilot-based spatial multiplexing schemes like VBLAST, because this enables us to rely on well-known algorithms for frequency and timing, estimation. In the BSS setup frequency and timing, estimation can be done on every separated data stream independently and therefore these setups are applicable even in multiuser scenarios.

To apply source separation techniques in communications, we are using the setup depicted in Figure 13. First of all, the DC offset caused by the direct conversion frontend is removed. After root-raised cosine filtering, a frame synchronization according to Section 3.1 is carried out. To separate the independent components, we can apply a BSS algorithm directly to the oversampled signal. For this step, we choose the JADE [18] algorithm<sup>8</sup> as a spatial-only separation approach.

<sup>8</sup>We also successfully used other approaches like fastICA [19] and SSARS [20].

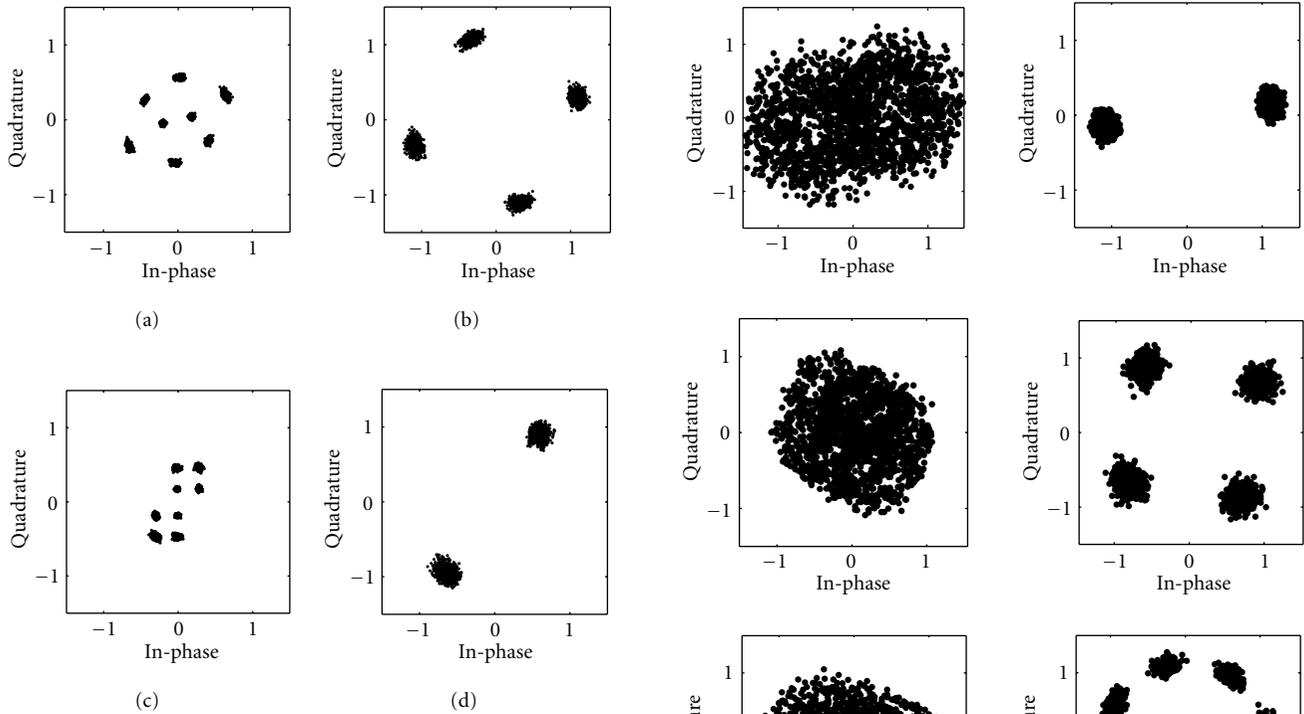


FIGURE 14:  $2 \times 2$  signal constellations before ((a) and (c)) and after ((b) and (d)) BSS.

The separation leads to data streams which are processed in the classical way like in single antenna systems. We synchronize to the symbol timing using the method presented in [10]. In order to determine the carrier frequency offset, we apply a nonlinearity and a frequency estimation.

Measurements were done with a sampling frequency of  $f_s = 10$  MHz in order to get an approximately flat channel. In order to visualize the successful separation, we simultaneously transmit signals with different modulation schemes.

Figure 14 depicts the separation of a BPSK and a QPSK signal sent in parallel and received by two antennas. The signal constellation before separation is obtained by using the timing information estimated after separation. As one can see in Figure 14, the signal streams are properly separated. Figures 14a and 14c show that in this particular measurement, the signal of the BPSK signal was dominant.

The separation procedure can be easily extended to a system with four transmit and receive antennas. The results are depicted in Figure 15. It can be seen that even in this situation, a proper blind separation is possible.

Based on our experiences, we can state that it is practically possible to apply separation algorithms for separation of communication signals in MIMO setups, even if the properties of the modulation schemes are not taken into account. This makes our setup interesting for interference scenarios. If a knowledge of the symbol alphabet of a signal is additionally exploited, the BSS can be used as a frontend to spatial interference cancellation algorithms like VBLAST [21].

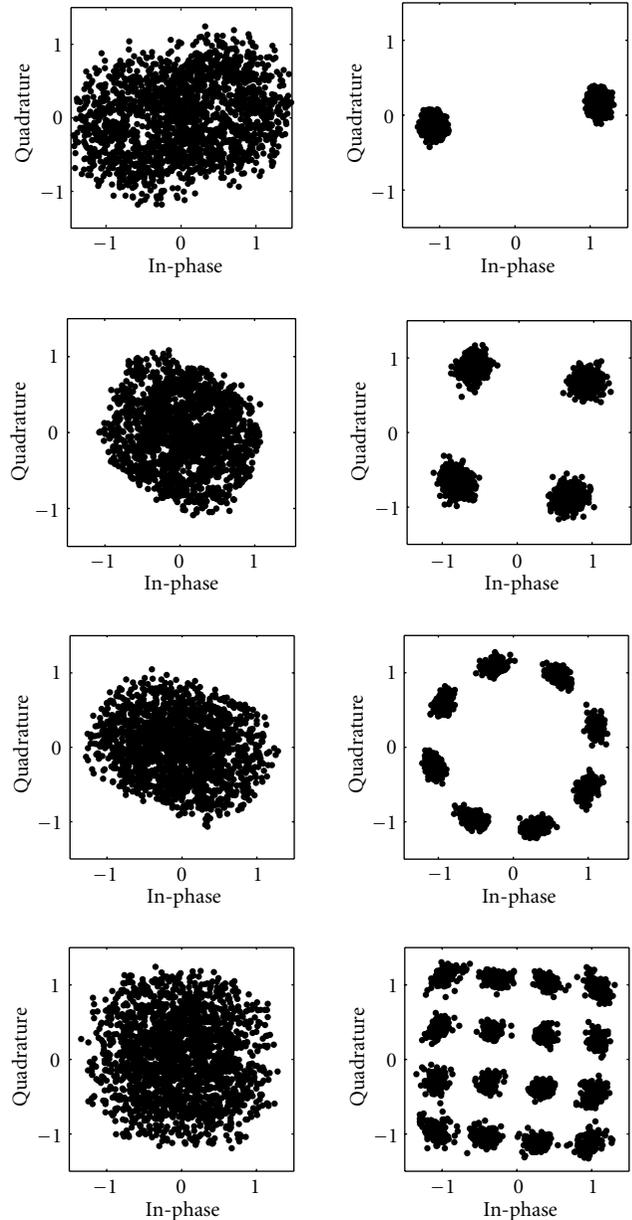


FIGURE 15:  $4 \times 4$  signal constellations before (left) and after (right) BSS.

#### 4. CONCLUSIONS

In this paper, we introduced a very flexible low-cost measurement system which allows the testing of nearly all MIMO communications setups currently under discussion. Arbitrary signals can be generated and transmitted in real time. However, the offline processing concept significantly reduces the complexity of the demonstrator. In contrast to a real-time simulator, this has enabled us to freely investigate optimal and suboptimal algorithms. Moreover, we are not limited to a special simulation software. A wide range of applications was presented. In order to show the nature of the MIMO

channel, we accomplished some indoor measurements of frequency responses. Furthermore, receive and transmit diversity schemes to gain performance from the spatial channel were considered. In theory, receive and transmit diversity are interchangeable. However, in practice, we observed that orthogonal STBCs are more sensitive to estimation errors. As an example for OFDM, we evaluated a system according to IEEE 802.11a to which we successfully applied several transmit and receive diversity schemes. The feasibility of BSS for communications systems under realistic conditions was studied. During our indoor measurements, we could hardly produce scenarios that prevent the BSS from working. Consequently, BSS algorithms, which can be directly applied to the oversampled received signal without timing and carrier offset synchronization, are suitable for robust frontend processing.

## REFERENCES

- [1] T. Haase, "Aufbau einer digitalen Funkübertragungsstrecke bei 2.4 GHz für Anwendungen innerhalb von Gebäuden," Diplomarbeit, University of Bremen, Bremen, Germany, 1999.
- [2] R. Gozali, R. Mostafa, R. C. Palat, et al., "Virginia-tech space-time advanced radio (VT-STAR)," in *IEEE Radio and Wireless Conference*, pp. 227–231, Waltham, Mass, USA, 2001.
- [3] L. Brühl, C. M. Walke, W. Keusgen, and C. Degen, *SABA - Ein Echtzeit-Demonstrator für MIMO- und Multi-User-Systeme mit adaptiven Gruppenantennen*, DFG Kolloquium, Kaiserslautern, October 2001.
- [4] T. Horseman, J. Webber, A. Nix, and M. Beach, "MIMO test bed design and performance assessment of candidate algorithms," Tech. Rep. D542 Part 3, IST-1999-10322 SATURN, 2003.
- [5] P. Murphy, F. Lou, and P. Frantz, "A hardware testbed for the implementation and evaluation of MIMO algorithms," in *IEEE International Conference on Mobile and Wireless Communications Networks*, Singapore, October 2003.
- [6] R. Thomä, A. Richter, U. Trautwein, D. Hampicke, and G. Sommerkorn, "Superresolution measurement and simulation of vector radio channels," in *2000 International Symposium on Antennas and Propagation*, pp. 249–252, Fukuoka, Japan, August 2000.
- [7] K. D. Kammeyer and V. Kühn, *Matlab in der Nachrichtentechnik*, Informations- und Kommunikationstechnik. J. Schlembach-Verlag, Weil der Stadt, Germany, 1st edition, 2001.
- [8] T. H. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*, Cambridge University Press, Cambridge, 1998.
- [9] R. H. Walden, "Analog-to-digital converter survey and analysis," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 4, pp. 539–550, 1999.
- [10] S. J. Lee, "A new non-data-aided feedforward symbol timing estimator using two samples per symbol," *IEEE Communications Letters*, vol. 6, no. 5, pp. 205–207, 2002.
- [11] R. Matzner, "An SNR estimation algorithm for complex baseband signals using higher order statistics," *Facta Universitatis: Electronics and Energetics*, vol. 6, no. 1, pp. 41–52, 1993.
- [12] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, 1998.
- [13] M. Stege, M. Bronzel, and G. Fettweis, "On the performance of space-time-blockcodes," in *Proc. IEEE 53rd Vehicular Technology Conference*, vol. 3, pp. 2282–2286, Rhodes, Greece, May 2001.
- [14] IEEE, *High-speed Physical Layer in the 5 GHz Band*, IEEE Std 802.11a, 1999.
- [15] H. Schmidt, *OFDM für die drahtlose Datenübertragung innerhalb von Gebäuden*, Ph.D. thesis, University of Bremen, Bremen, Germany, 2001.
- [16] A. Dammann and S. Kaiser, "Transmit/receive antenna diversity techniques for OFDM systems," *European Transactions on Telecommunications*, vol. 13, no. 5, pp. 531–538, 2002.
- [17] M. Bossert, A. Huebner, F. Schuehlein, E. Costa, and H. Haas, "On cyclic delay diversity in OFDM based transmission schemes," in *Proc. 7th International OFDM-Workshop*, Hamburg, Germany, September 2002.
- [18] J.-F. Cardoso and A. Souloumiac, "Blind beamforming for non-Gaussian signals," *IEE Proceedings Part F: Radar and Signal Processing*, vol. 140, no. 6, pp. 362–370, 1993.
- [19] E. Bingham and A. Hyvärinen, "A fast fixed-point algorithm for independent component analysis of complex-valued signals," *International Journal of Neural Systems*, vol. 10, no. 1, pp. 1–8, 2000.
- [20] M. Feng and K. D. Kammeyer, "Blind source separation for communication signals using antenna arrays," in *Proc. IEEE Int. Conference on Personal Wireless Communications*, Florence, Italy, October 1998.
- [21] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. Int. Symp. Signals, Systems, Electronics*, pp. 295–300, Pisa, Italy, September 1998.

**J. Rinas** studied electrical engineering at the University of Bremen, where he finished his Diplomarbeit (equivalent to M.S.) on RAKE receiver structures for the UMTS in 2000. In the same year, he joined the Department of Telecommunications at the University of Bremen as a Ph.D. student. His main research interests are blind source separation in MIMO communication systems and their practical realizations.



**R. Seeger** studied electrical engineering at the University of Bremen, where he finished his Diplomarbeit (equivalent to M.S.) on the design and implementation of parametric filters on a real-time platform in February 1999. In the same year, he joined the Department of Telecommunications at the University of Bremen as a Ph.D. student. His main research interests are space-time processing for the UMTS downlink and practical realization aspects of communication systems.



**L. Brötje** was born in Bremen, Germany, in 1973. He studied communications at the University of Bremen and finished his Diplomarbeit (equivalent to M.S.) in 2000. Currently, he is working on his Ph.D., focused on WLAN-systems (IEEE 802.11a/g). His main research topics are nonlinearities, for example, I/Q imbalances, DC offsets, and synchronizations aspects.



**S. Vogeler** studied electrical engineering at the University of Bremen, where he finished his Diplomarbeit (equivalent to M.S.) on finite alphabet-based blind channel estimation for OFDM systems in June 2001. In the same year, he joined the Department of Telecommunications at the University of Bremen as a Ph.D. student. His main research interests comprise the coexistence problems of different wireless LAN standards as well as the application of OFDM transmission techniques in case of strong Doppler influence.



**T. Haase** studied electrical engineering at the University of Bremen, where he finished his Diplomarbeit (equivalent to M.S.) on the hardware design of a 2.4 GHz wireless transmission system for indoor applications in December 1999. From January 2000 to April 2003, he worked at the Department of Telecommunications, the University of Bremen as a Technician. His main research interest is the design of electronic devices for communications. Since May 2004, he has been working at the ZARM Technik GmbH where he develops electronic devices for space applications.



**K.-D. Kammeyer** received the Diplom degree in electrical engineering (equivalent to M.S.) from Berlin University of Technology, Germany, in 1972, and the Ph.D. degree from Erlangen University, Germany, in 1977. From 1972 to 1979, he worked in the field of data transmission, digital signal processing, and digital filters at the Universities of Berlin, Saarbrücken, and Erlangen, all in Germany. From 1979 to 1984, he was with Paderborn University, Germany, where he was engaged in the development of digital broadcasting systems. During the following decade, he was Professor for digital signal processing in communications at Hamburg University of Technology, Germany. In 1995, he was appointed Professor for telecommunications at the University of Bremen, Germany. His research interests are digital (adaptive) systems and signal processing in mobile communication systems (GSM, UMTS, and multicarrier systems). Since 1989, he is active in the field of higher-order statistics. Professor Kammeyer holds 14 patent families. He has published three course books as well as 75 technical papers.



# Exploiting Phase Diversity for CDMA2000 1X Smart Antenna Base Stations

**Seongdo Kim**

*Advanced R&D Team, Ace Technology, Seoul 137-130, Korea  
Email: ksd544@acetech.co.kr*

**Seunghoon Hyeon**

*School of Electrical and Computer Engineering, Hanyang University, Seoul 133-791, Korea  
Email: hsheon@dsplab.hanyang.ac.kr*

**Seungwon Choi**

*School of Electrical and Computer Engineering, Hanyang University, Seoul 133-791, Korea  
Email: choi@ieee.org*

*Received 25 June 2003; Revised 4 March 2004*

A performance analysis of an access channel decoder is presented which exploits a diversity gain due to the independent magnitude of received signals energy at each of the antenna elements of a smart-antenna base-station transceiver subsystem (BTS) operating in CDMA2000 1X signal environment. The objective is to enhance the data retrieval at cellsite during the access period, for which the optimal weight vector of the smart antenna BTS is not available. It is shown in this paper that the access channel decoder proposed in this paper outperforms the conventional one, which is based on a single antenna channel in terms of detection probability of access probe, access channel failure probability, and Walsh-code demodulation performance.

**Keywords and phrases:** phase diversity, access channel, searching, Walsh-code demodulation, CDMA2000 1X.

## 1. INTRODUCTION

As the demand of mobile communications increases rapidly, the 3G mobile communication system must provide various contents in as high as possible data rate compared to the conventional 2G systems [1]. In order to provide the various contents to the increased number of users, it is necessary to secure an extremely accurate detection of synchronization information between base-station transceiver subsystem (BTS) and each of the subscribers together with a good demodulation technique [2]. For improving the quality of communication services capable of providing the various contents, the mobile communication network based on a single antenna BTS must increase the number of base stations. However, to increase the number of base stations is very costly and it also causes the cell planning to be very complicated because of the frequent handoff [3].

Smart antenna technology [4, 5] has been considered as being a solution to increase the communication capacity and improve the communication quality [6, 7] as well without too much investment required for increasing the number of base stations. In order for the smart antenna system to work as desired, however, the weight vector should be provided in

such a way that a nice-shaped beam pattern is generated in accordance with the directions of desired and/or undesired signal sources [8].

It is noteworthy that the data format of the access channel in CDMA2000 1X system is the same as that employed in the current IS-95 CDMA system. It particularly means that access channel consists of the preamble, data, and cyclic redundancy code (CRC) with the data rate being 4.8 kbps. As the access channel is not accompanied with pilot channel, the contents of the access channel should be decoded through a noncoherent detection [9, 10, 11]. For that reason, during the access state, which exists before a traffic channel is set up between BTS and a given subscriber such that an optimal weight vector has not yet been computed according to the received data at the BTS, the advantages of smart antenna, which is available due to the nice-shaped beam pattern, cannot be provided. In this paper, we apply the principle of phase diversity technology [12] to the access channel decoder in order to enhance the data retrieval during the access state. The phase diversity can be obtained in any antenna array system which uses the envelope detection procedure regardless of antenna spacing due to the fact that the

energy of received signal at each antenna element consists of many signal components transmitted from a large number of users with all statistically independent arrival angles. Consequently, the energy at each antenna element becomes statistically independent as the number of users increases in a given CDMA signal environment. The gain of phase diversity can be exploited in the access channel decoder, of which the function is basically to compute the correlation energy of received signal with each of the Walsh codes by simply summing up the correlation energies calculated at each of the antenna elements.

The superiority of the proposed access channel decoder exploiting the phase diversity has been confirmed in terms of searching capability, Walsh demodulation, access failure probability, and so forth. In this paper, the access failure has been found by checking the CRC that is contained in the access probe.

This paper is composed as follows. Section 2 summarizes the concept of the phase diversity. Section 3 introduces the structure of access channel decoder for a hardware implementation. Section 4 presents the performance analysis of the proposed access channel decoder in comparison to the conventional one. Section 5 presents the concluding remarks.

## 2. PHASE DIVERSITY

Phase diversity technology is based on a fact that the magnitude of received signal energy at each antenna element of a smart antenna system is independent of each other because the phase of every component, which is determined by the arrival angle and carrier phase delay associated with the corresponding mobile terminal, is statistically independent of each other. It may sound contradictory to another fact that the received signal at each antenna element is coherent to each other, that is, the magnitude of received signal is the same at every antenna element and only the phase varies when the angle spread is not too wide and the antenna spacing is not far greater than a half wavelength. What has to be carefully considered in the discussion of the phase diversity technology is that the received signal at each antenna element is composed of plural signal components, each of which is transmitted from a corresponding mobile terminal. Each of the received signal components at each antenna element is fully coherent to each other such that the magnitude is exactly the same at every antenna element and the phase difference between adjacent elements is  $\pi \sin(\theta)$  when the antenna spacing is a half wavelength where  $\theta$  is the arrival angle measured from the broad side of the array. However, due to the large number of transmitting mobile terminals, the magnitude of the received signal which consists of the large number of the signal components is independent at every antenna element. The independency of the received signal energy can be clarified through the equations shown in this section.

The phase difference between adjacent antenna elements due to a single signal component, say, a signal transmitted from  $l$ th subscriber, is a function of incident angle  $\theta_l$  and

antenna spacing  $d$ , that is,  $\varphi(\theta_l, d)$ . For simplicity but without loss of generality, we assume that the first antenna element is the reference antenna. Then, holding back the terms related to the multipath and angle spread until we quote Section 4, after the frequency down-conversion, the in-phase and quadrature component of the received signal at the  $n$ th element can be respectively written as follows:

$$\begin{aligned}\hat{I}_n(t) &= \sum_{m=1}^M \hat{S}_m(t) \cos [\phi_m + (n-1)\varphi(\theta_m, d)], \\ \hat{Q}_n(t) &= \sum_{m=1}^M \hat{S}_m(t) \sin [\phi_m + (n-1)\varphi(\theta_m, d)],\end{aligned}\quad (1)$$

where the subscripts  $n$  and  $m$  are the indices for denoting the antenna element and signal source, that is, the transmitting subscriber, respectively,  $\hat{S}_m(t)$  is the magnitude of the received signal,  $M$  is the total number of signal components impinging upon the antenna element, and  $\phi_m$  is the carrier phase delay. The noise term has been deleted for ease of explanation. Assuming that the signal transmitted from the  $l$ th subscriber is the desired one, after the despreading procedure with the PN code assigned to the  $l$ th subscriber, say,  $p_l(t)$ , the in-phase and Quadrature component of the received signal at the  $n$ th element can be written as follows:

$$\begin{aligned}I_n(t) &= S_l(t) \cos [\phi_l + (n-1)\varphi(\theta_l, d)] \\ &+ \sum_{m=1, m \neq l}^M S_m(t) \cos [\phi_m + (n-1)\varphi(\theta_m, d)], \\ Q_n(t) &= S_l(t) \sin [\phi_l + (n-1)\varphi(\theta_l, d)] \\ &+ \sum_{m=1, m \neq l}^M S_m(t) \sin [\phi_m + (n-1)\varphi(\theta_m, d)],\end{aligned}\quad (2)$$

where  $I_n(t) = \int_T \hat{I}_n(t) p_l(t) dt$ ,  $Q_n(t) = \int_T \hat{Q}_n(t) p_l(t) dt$ , and  $S_m(t) = \int_T \hat{S}_m(t) p_l(t) dt$  with the integral period  $T$  being determined by the processing gain. Note that the first terms in the right-hand side of (2) are the desired ones while the last terms are the interfering ones. The key part of the phase diversity is that the signal at each antenna element is independent of each other because the signal at each antenna element consists of a large number of signals transmitted from randomly located large number of mobile terminals. Once again, this is because the interfering terms consist of  $M-1$  terms of which the incident angles are all independent such that the magnitude of the received signal at each element, which is determined by a vector sum of  $M$  signal components, must be determined in a random fashion.

In Walsh-code demodulator in access channel decoder, as there are 64 Walsh codes in CDMA2000 1X system, each symbol shown in (2) should be correlated with each of the 64 Walsh codes at each antenna element to produce the decision variables of Walsh demodulator. Then, the decision variables for  $k = 1, 2, \dots, 64$  at the  $n$ th antenna element, which is obtained by correlating the symbol with each of the 64 Walsh

codes, can be written as follows:

$$\begin{aligned} I_{n,k}(t) &= S_{l,k}(t) \cos [\phi_l + (n-1)\varphi(\theta_l, d)] \\ &+ \sum_{m=1, m \neq l}^M S_{m,k}(t) \cos [\phi_m + (n-1)\varphi(\theta_m, d)], \\ Q_{n,k}(t) &= S_{l,k}(t) \sin [\phi_l + (n-1)\varphi(\theta_l, d)] \\ &+ \sum_{m=1, m \neq l}^M S_{m,k}(t) \sin [\phi_m + (n-1)\varphi(\theta_m, d)], \end{aligned} \quad (3)$$

where the subscript  $k$  is the Walsh index and the  $k$ th decision variable is obtained as  $I_{n,k}(t) = \int_{T_W} I_n(t) W_k(t) dt$ ,  $Q_{n,k}(t) = \int_{T_W} Q_n(t) W_k(t) dt$ , and  $S_{m,k}(t) = \int_{T_W} S_m(t) W_k(t) dt$  for  $k = 1, 2, \dots, 64$  with the integral period  $T_W$  being determined by the length of the Walsh code. Since the interfering terms in (3) can be approximated to Gaussian as the number of signal components is sufficiently large, (3) can be rewritten as follows:

$$\begin{aligned} I_{n,k}(t) &= G[S_{l,k}(t) \cos(\Theta_{n,l}), \sigma^2], \\ Q_{n,k}(t) &= G[S_{l,k}(t) \sin(\Theta_{n,l}), \sigma^2], \end{aligned} \quad (4)$$

where  $G[\mu, \sigma^2]$  denotes a Gaussian random variable with mean  $\mu$  and variance  $\sigma^2$ , where the variance is determined by the sum of the interferers' power measured at the receiver, and  $\Theta_{n,l} = \phi_l + (n-1)\varphi(\theta_l, d)$ .

What is claimed in the phase diversity technology is that the decision variable should be computed as a sum of all the values obtained at every antenna channel. Thus, the decision variable  $Z_k(t)$  for estimating the Walsh number in the Walsh demodulator is

$$Z_k(t) = \sum_{n=1}^N [I_{n,k}^2(t) + Q_{n,k}^2(t)] \quad \text{for } k = 1, 2, \dots, 64. \quad (5)$$

As  $I_{n,k}(t)$  and  $Q_{n,k}(t)$  are Gaussian random variables as discussed above, the decision variable  $Z_k(t)$  is a noncentric chi-squared random variable with the degree of freedom being  $2N$ . The probability density function of  $Z_k(t)$  can be written as follows [13]:

$$\begin{aligned} P_Z(\alpha) &= \frac{(\alpha/\sigma^2 b)^{(N-1)/2}}{2\sigma^2} e^{-(1/2)(b+\alpha/\sigma^2)} I_{N-1} \left( \sqrt{\frac{b\alpha}{\sigma^2}} \right) \quad \text{for } \alpha \geq 0, \\ &= 0 \quad \text{for } \alpha < 0, \end{aligned} \quad (6)$$

where  $I_{N-1}(\bullet)$  is a modified Bessel function of the first kind with order  $N-1$  and the noncentric parameter  $b$  is  $N(S_{l,k}(t)^2/\sigma^2)$ . Note that when the magnitude of the desired signal is zero, the probability density function of  $Z_k(t)$  becomes a centric chi-squared random variable. In this case, the probability density function of  $Z_k(t)$  can be written as follows:

$$p_Z(\alpha) = \frac{\alpha^{(N-1)}}{(\sigma^2)^N 2^N \Gamma(N)} e^{-\alpha/2\sigma^2}, \quad (7)$$

where  $\Gamma(\bullet)$  is the Gamma function.

The average and variance of the centric chi-squared random variable are  $2N\sigma^2$  and  $4N\sigma^4$ , respectively, and those for the noncentric chi-squared random variable are  $N(2\sigma^2 + S_{l,k}(t)^2)$  and  $4N\sigma^2(\sigma^2 + S_{l,k}(t)^2)$ , respectively. Note that both the mean and variance increase linearly as the number of antenna elements  $N$  increases. This suggests that the performance of the envelope detection, that is,  $Z_k(t) = \sum_{n=1}^N [I_{n,k}^2(t) + Q_{n,k}^2(t)]$ , improves linearly as the number of antenna elements increases. Note that there would be no gain at all if the variance increases in proportion to  $N^2$  as the mean increases in proportion to  $N$ . Consequently, the phase diversity technique, with the detection variables being computed as suggested in (5), increases the signal-to-interference ratio (SIR) by nearly  $N$  times where  $N$  is the number of antenna elements in the array system. Note that there is no weight computation involved in the phase diversity technique.

### 3. ACCESS CHANNEL DECODER

Figure 1 illustrates a block diagram of the access channel modulator operating in CDMA2000 1X mobile communication system. As the terminal may access the BTS at any moment, the BTS in CDMA2000 1X system receives the access information based on a noncoherent detection. Note that, as shown in Figure 1, the information in the access channel is transmitted in a low data rate, that is, 4.8 kbps. It can also be observed in Figure 1 that the access channel modulator includes the 64-ary Walsh modulation and the offset quadrature phase shift keying (OQPSK) modulation [14]. It is also noteworthy that the OQPSK modulation in the access channel modulator can provide a diversity gain because it adopts two BPSK modulations for the  $I$ -channel and  $Q$ -channel, respectively, of a given signal.

The access channel decoder implemented in our smart antenna BTS consists of a searcher, OQPSK demodulator, Walsh demodulator, Viterbi decoder, CRC checker, and so forth, in such a way that it can demodulate the access channel data modulated through the procedure shown in Figure 1. Among the blocks in the access channel decoder, the searcher and Walsh demodulator employs the proposed phase diversity technology of which the detailed application method and hardware structures are introduced in Sections 3.1 and 3.2.

#### 3.1. Searcher

The searcher in the access channel decoder performs PN code acquisition for retrieving the access channel information at cellsite, using the preambles given at the beginning part of access probe [15]. As described earlier, in order to exploit the phase diversity gain, the correlation energies obtained at each of the antenna channels are summed up each time to produce the detection variable, with which the searcher detects the peak correlation energy to estimate the propagation delay of the target subscriber. Figure 2 illustrates the hardware structure of the access searcher implemented in our smart antenna system operating in CDMA2000 1X system. Note that the correlation energies computed at each of the antenna channels are summed up to form the detection variable for determining whether or not the current time lag corresponds

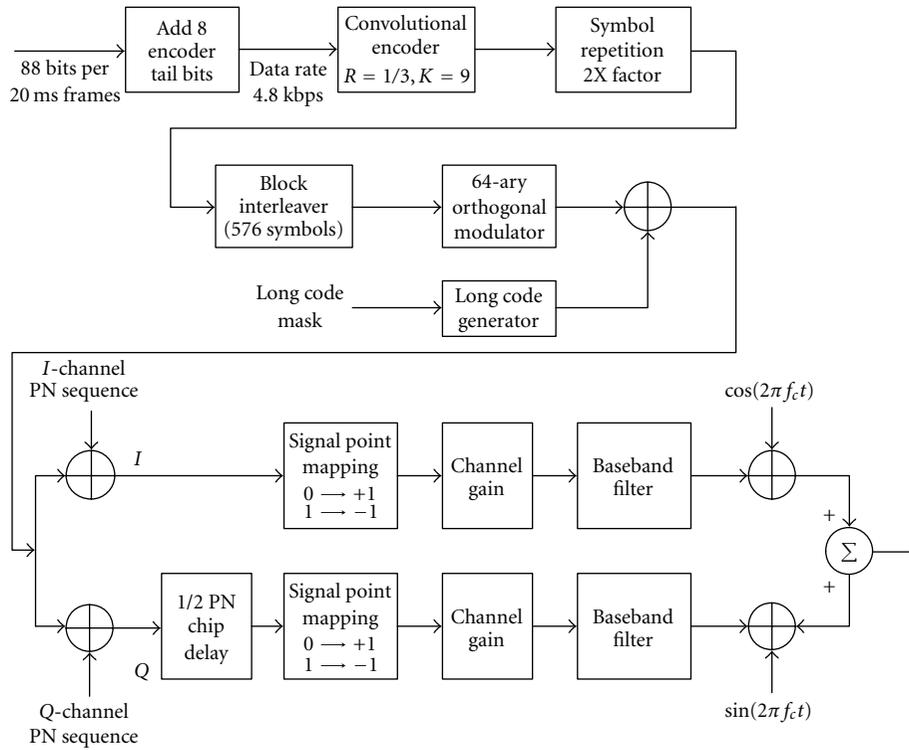


FIGURE 1: Structure of access channel modulator of CDMA2000 1X system.

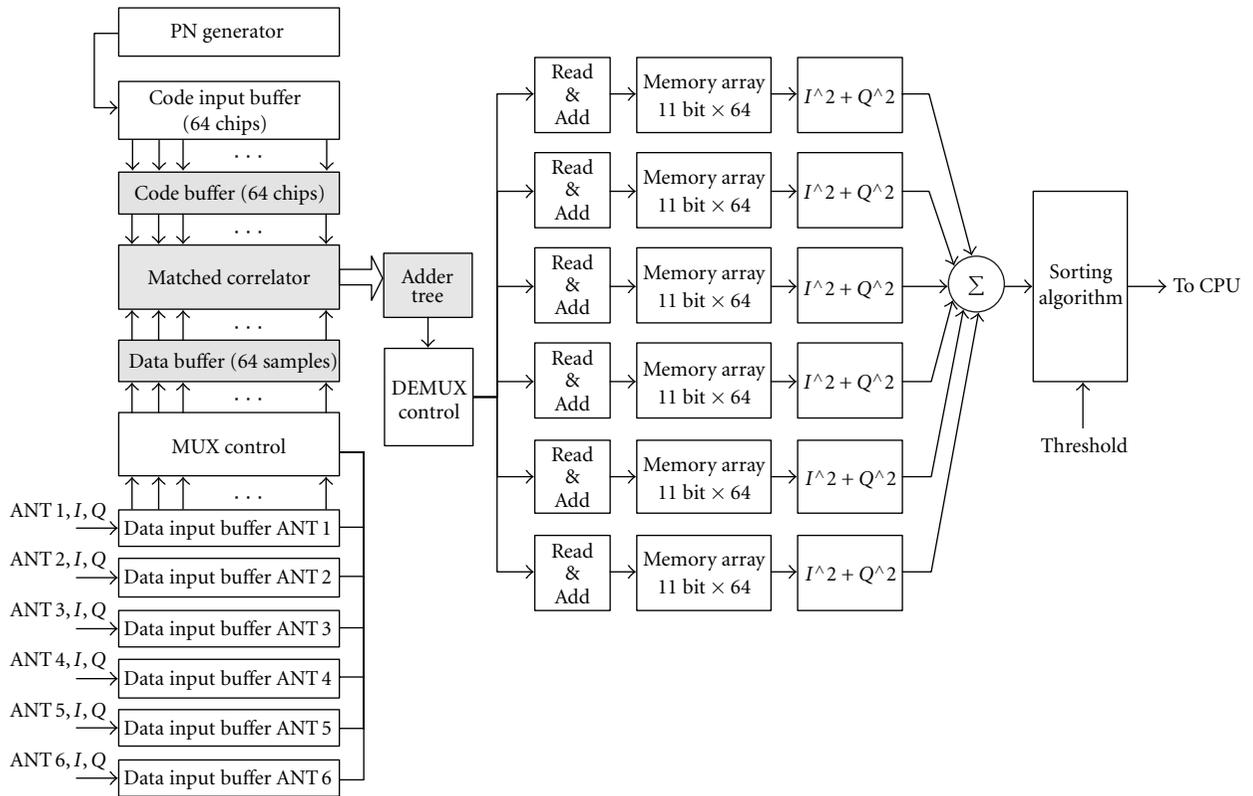


FIGURE 2: Structure of the proposed searcher.

to the peak correlation value. In our smart antenna BTS, 6 antenna channels are activated for reverse link, that is,  $N = 6$ .

As shown in Figure 2, the  $I$  and  $Q$  components of the received signal obtained after the frequency-down and analog-to-digital conversion at each antenna channel, that is,  $\{\text{ANT}_i I, Q \text{ for } i = 1, 2, \dots, N\}$ , respectively, are fed to the corresponding data input buffer.

The searching function, that is, PN code acquisition, is performed after analog-to-digital conversion of frequency down-converted baseband signal received at each of the antenna channels. The detailed procedure of computing the correlation energy in the hardware implementation of our access channel searcher is as follows. First, the correlation between the received data and PN code of the desired subscriber is performed for the period of 1/4 PCG (power control group), which is of 384-chip duration, with the time lag of correlation being shifted by 1/2 chip at a time. At each antenna channel, the correlation energy is obtained by averaging the correlation results for the period of 7 PCG. Consequently, the total period for obtaining the correlation energy at each antenna channel is for  $7 \times 4 \times 384$  chips. As for the window size, that is, the length of the time interval for which the peak correlation energy is searched, it has been set to 30-chip duration, which means the correlation energy is computed 60 times because the time lag for computing the correlation is jumping by 1/2 chip each time the correlation is computed as mentioned above. Summarizing the above, the searching energy is computed for each time lag  $i$  by

$$Z_{\text{search}}\left(\frac{i}{2}\right) = \sum_{j=1}^N \left\{ \begin{aligned} & \left[ \sum_{k=0}^{28} \sum_{n=1}^{384} x_j^I[n + 384 \times k] \right. \\ & \quad \left. \cdot \text{PN}^I \left[ n + 384 \times k - \frac{i}{2} \right] \right]^2 \\ & + \left[ \sum_{k=0}^{28} \sum_{n=1}^{384} x_j^Q[n + 384 \times k] \right. \\ & \quad \left. \cdot \text{PN}^Q \left[ \left( n - \frac{1}{2} \right) + 384 \times k - \frac{i}{2} \right] \right]^2 \end{aligned} \right\} \quad (8)$$

for  $i = 0, 1, 2, \dots, 59$ ,

where  $Z_{\text{search}}(t)$  is the searching energy to be computed,  $x_j^I(t)$  and  $x_j^Q(t)$  are, respectively, the in-phase and quadrature component of the received data at  $j$ th antenna element, and  $\text{PN}^I(t)$  and  $\text{PN}^Q(t)$  denote the PN code of the desired subscriber assigned to In-phase and Quadrature component, respectively, in CDMA2000 1X system. Note that the time index includes  $(i/2)$  term because the time lag for computing the correlation energy is shifted by 1/2-chip duration each time as described previously. It should also be observed that the received data  $x_j^I(t)$  and  $x_j^Q(t)$  consist of many signal components transmitted from all the subscribers operating in a given cell through the traffic as well as access channels. In the access searcher operating in accordance with (8), the phase

diversity can be exploited because the correlation energy is obtained through the summation of all the correlation results at each of the antenna channels. The objective of the searcher in the access channel decoder is to find the time index  $(i/2)$  for which the peak value of the correlation energy is given as a result of computing (8) for every  $i$ . The problem of setting the threshold value to determine whether or not each of the correlation values corresponds to a peak is not included in the scope of this paper.

### 3.2. Walsh demodulator

The objective of the Walsh demodulator is to find the Walsh number that corresponds to the information of 6-bit word transmitted from the desired subscriber. So, the question to be answered in the Walsh demodulator is “Which one of  $\{W_0, W_1, W_2, \dots, W_{63}\}$  has been transmitted from the target subscriber?” Figure 3 illustrates the Walsh demodulator implemented in our smart antenna BTS [16] operating in the reverse link of CDMA2000 1X signal environment that can fully exploit the gain of phase diversity [17]. At each antenna channel, the received signal is first mixed with  $\cos \omega_1 t$  and  $\sin \omega_1 t$  to produce the baseband in-phase and quadrature component  $r_{I,n}(t)$  and  $r_{Q,n}(t)$ , respectively, for  $n = 1, 2, \dots, N$ , where  $\omega_1$  is the carrier frequency with  $N$  being the number of antenna elements in the smart antenna system. For ease of explanation, A-to-D converter is omitted in Figure 3. Then, the received data are descrambled with the long and short PN code, that is,  $c(t)$ , and  $p_I(t)$  and  $p_Q(t)$  for in-phase and quadrature, respectively, in Figure 3, assigned to the desired subscriber. Note that the descrambling procedure is performed using the timing information, that is,  $(i/2)$  as shown in (8), given from the access searcher. Then, the descrambled received data are fed to the input ports of the Walsh demodulator at each antenna channel to be correlated with each of the 64 Walsh words. In Figure 3, the correlation of the received data with each of the Walsh words is denoted as  $\langle \bullet, W_k(t) \rangle$ , which is performed by  $\int_{T_W} \bullet \cdot W_k(t) dt$  with the integration period  $T_W$  being the Walsh word length as mentioned in Section 2. Note that in order to provide the phase diversity gain mentioned earlier, the decision variable  $Z_k$  for  $k = 1, 2, \dots, 64$  is obtained through the summation procedure of correlation values computed at each of the antenna channels in Figure 3 [3]. As the objective of the Walsh demodulator is irrelevant to the time index, the time index in the decision variable  $Z_k(t)$  as shown in (5) is omitted in this section.

Figure 4 is the photograph of a channel card implemented in our smart antenna BTS [17]. Each channel card installed in a given smart antenna BTS includes all the necessary modules for modulating or demodulating the signal to be transmitted or received to or from the corresponding subscriber. The access channel searcher and the Walsh demodulator discussed in this section are included in the demodulation part of the channel card. As shown in Figure 4, the demodulator has been implemented with 5 of 1 million-gate FPGAs (field programmable gate array)—Altera’s APEX EP20K1000EBC652. In fact, the demodulation part of the channel card consists of the access channel searcher, Walsh

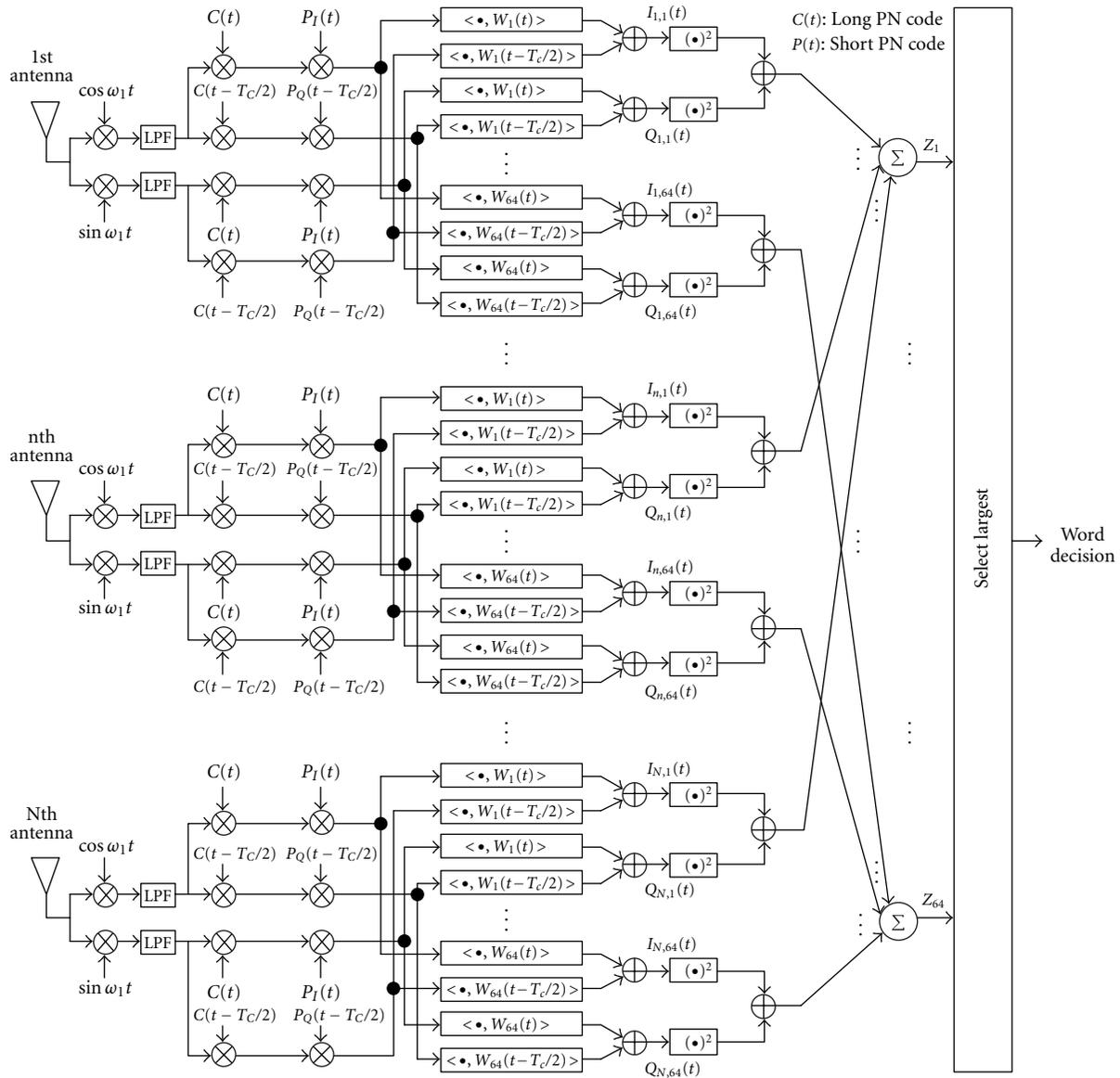


FIGURE 3: Walsh demodulator providing the phase diversity in smart antenna base station.

demodulator, both of which are discussed in this section as main topics of this paper, traffic channel demodulator for spreading the received traffic data, demodulator controller that has been implemented with a digital signal processing (DSP)—TMS320C6203, and channel card controller that has been implemented with MPC860. All these modules for the demodulation have been implemented in the 5 FPGAs shown in Figure 4. The demodulator shown in Figure 4 provides 4 fingers to each user for the RAKE reception [18]. The received data through the access channel allocated to each of the 4 fingers by the access searcher are demodulated in the Walsh demodulator implemented with the FPGAs. The demodulator controller demodulates the access probe from the retrieved 6-bit words which are obtained as a result of the Walsh demodulation. The state of the desired subscriber can

also be monitored from these 6-bit words information because the CRC of the desired subscriber is included in the access probe. The CRC information are transferred to the channel card controller. It is also the channel card controller that performs the interface between physical layer and higher layer as well as the call processing using the message retrieved from the access probe.

## 4. PERFORMANCE ANALYSIS

### 4.1. Signal modeling and experimental environment

In this subsection, the performance analysis of the proposed access channel decoder is presented through various computer simulations and experimental results. The received baseband signal at the  $n$ th antenna element, assuming the



FIGURE 4: Channel card for smart antenna system.

$n_0$ th element is reference antenna element, can be written as follows [19, 20]:

$$x_n(t) = \sum_{m=1}^M \sum_{k=1}^{K_m} \left( \sum_{q=1}^{L_k} s_{m,k}(t) e^{j2\pi(f_d \cos \varphi_{m,k,q} t - f_c \tau_{m,k,q})} \right) \times e^{-j(n-n_0)\pi \sin \theta_{m,k,q}} + n(t), \quad (9)$$

where the indices  $m$ ,  $k$ , and  $q$  are used to denote the subscriber, propagation path, and scattered components, respectively,  $s_{m,k}(t)$  is the received signal from the  $m$ th subscriber through the  $k$ th path,  $f_d$  is Doppler shift,  $\varphi$  is the moving direction of the desired subscriber measured from the broad side of the array antenna,  $f_c$  is the carrier frequency,  $\tau$  is the propagation delay of the signal,  $\theta$  is the arrival angle of the received signal, and  $n(t)$  is a zero-mean Gaussian random quantity determined by the noise. Note that the signal model shown in (9) includes the angular spread. Note that  $M$ ,  $K_m$ , and  $L_k$  are the number of subscribers, multipaths, and scattered components, respectively.

It has been assumed in our simulations that the propagation delay in a cluster are all the same such that  $\tau_{m,k,q} \cong \tau_{m,k}$  [19]. The magnitude of each of the multipaths has been determined in accordance with the 6-finger model given in 3GPP2 recommendation as follows: 0.6369 : 0.5742 : 0.3623 : 0.2536 : 0.2595 : 0.0407, which results in the power ratio of 0.4056 : 0.3297 : 0.1313 : 0.0643 : 0.0673 : 0.0017. Note that for the retrieval of access channel data, which is the main issue of this paper, the largest 4 fingers out of the 6, of which the instantaneous magnitude varies at every sampling time due to the fading, are taken because the number of fingers in our receiver is 4 as mentioned earlier.

#### 4.2. Performance of the searcher

In the access state, during which the pilot signals are not available, the searching, that is, PN code acquisition, is performed using the preamble given at the access probe. In

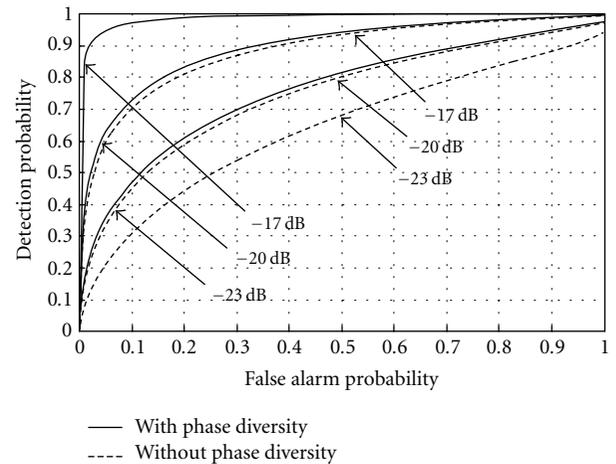


FIGURE 5: False alarm probability versus detection probability in AWGN environment.

our smart antenna BTS, the access probe is composed of 5-frame preamble and 4-frame message. The integration period for the searcher has been set to 384-chip duration ( $312.5 \mu\text{second}$ ), which is 1/4 PCG. As the phase shift of the received signal is about  $9^\circ$  for  $f_d = 80 \text{ Hz}$  for the integration period of  $312.5 \mu\text{second}$ , the correlation energy in the searcher is obtained by averaging the correlation values for the period of 7 PCG for higher accuracy as stated in Section 3.

Figures 5 and 6 illustrate the performance of the access searcher in terms of false alarm probability and detection probability in additive white Gaussian noise (AWGN) and fading channel, respectively. The numbers given inside the parenthesis in Figures 5 and 6 denote the signal-to-interference-plus-noise ratio (SINR), being  $-17 \text{ dB}$ ,  $-20 \text{ dB}$ , and  $-23 \text{ dB}$ . It has been found in our extensive computer simulations that the searcher designed by the proposed

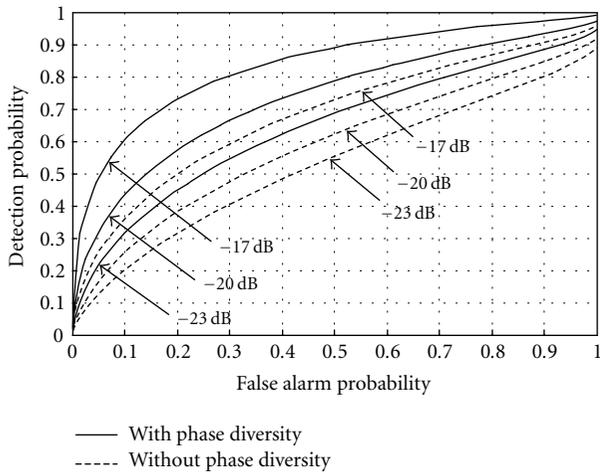


FIGURE 6: False alarm probability versus detection probability in the fading environment.

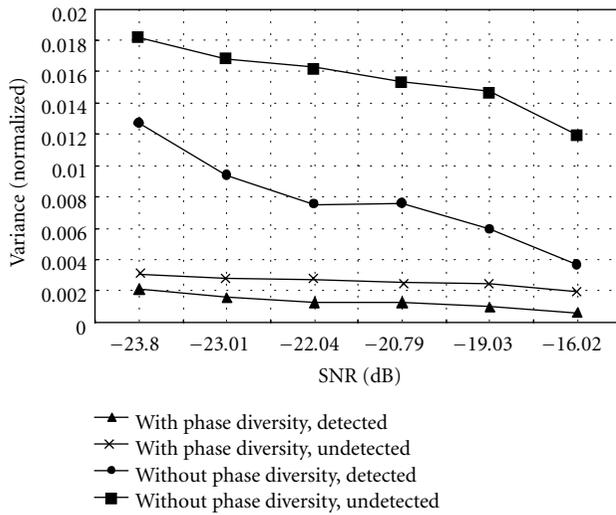


FIGURE 7: Variance of searching energy.

technique that exploits the phase diversity provides for about 3.3 dB gain in the SINR for AWGN channel. The gain in the fading environments has been found to be about 5.5 dB.

The performance improvement provided by the proposed searcher can also be found in the distribution of the variance of the detection variable, that is,  $E[\{Z(i/2) - E[Z(i/2)]\}^2]$ , where  $E[\bullet]$  denotes the expectation of  $\bullet$  and  $Z(i/2)$  is the detection variable, that is, correlation energy, as defined in (8). Figure 7 illustrates the variance of the normalized detection variable. Note that the labeling “detected” or “undetected” means that the time lag of the detection variable, that is,  $(i/2)$ , is matched or mismatched with the actual propagation delay of the signal, respectively. In computing the variance, the correlation energy itself has been normalized with the average correlation energy, that is,  $E[Z(i_0/2)]$ , obtained in the detected case. It should be ob-

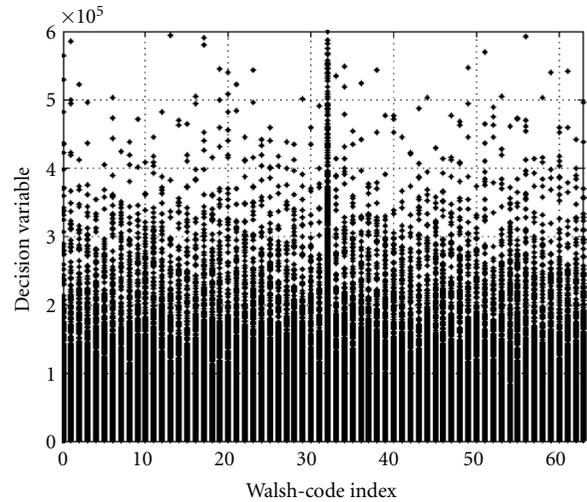


FIGURE 8: Decision variable of the Walsh demodulator for single antenna system.

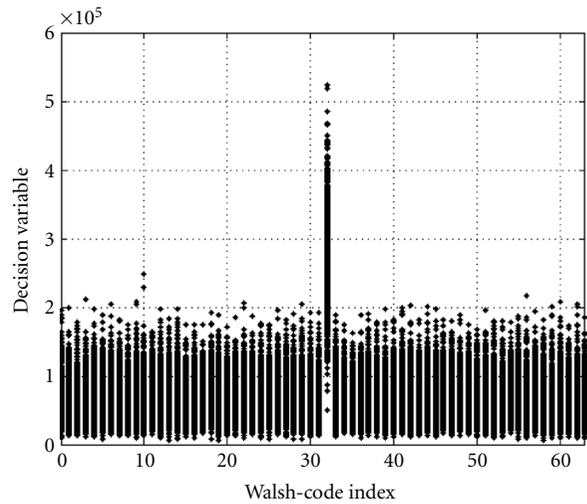


FIGURE 9: Decision variable of the Walsh demodulator for array antenna system.

served in Figure 7 that the variance in both the detected and undetected cases is significantly reduced, which means the detection capability is enhanced and the false alarm is suppressed as much as the variance is reduced. Recall that the variance of the correlation energy is actually an uncertainty in the detection procedure of the searcher. It has been found in our extensive simulations that the variance is reduced nearly by  $1/N$ , that is,  $1/6$  in our smart antenna BTS. This result could be predicted from the discussions of Section 2.

### 4.3. Performance of the Walsh demodulator

The performance of the Walsh demodulator is presented in terms of the distribution of the detection variable  $Z_k$  and the improvement in  $E_b/N_0$ . Figures 8 and 9 illustrate the distribution of the decision variable  $Z_k$  in an AWGN channel of SNR = -24 dB. Each of the 64 Walsh words is correlated with

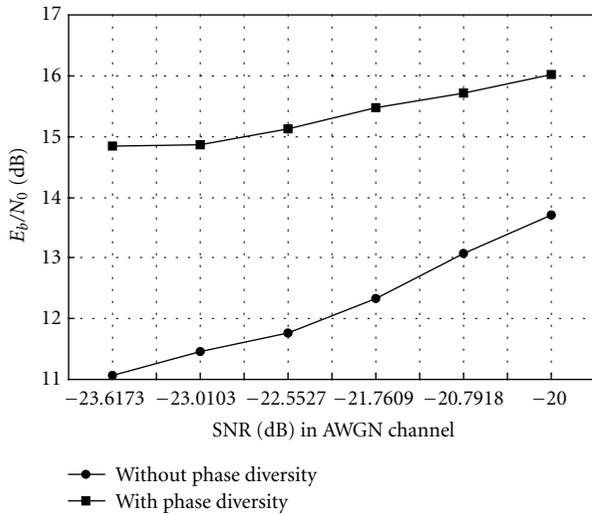
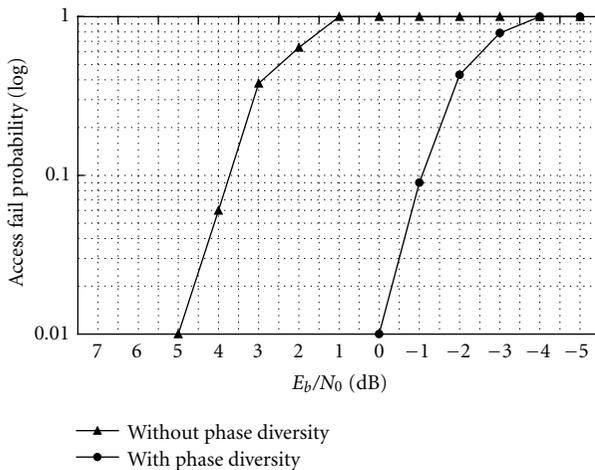
FIGURE 10:  $E_b/N_0$  enhancement in the Walsh demodulator.

FIGURE 11: Access fail probability in AWGN environment.

received signal of 64-chip length when a test signal corresponding to  $W_{32}$  is assumed to be transmitted from a mobile terminal. From Figures 8 and 9, it can be observed that the decision variable  $Z_{32}$  provided from the proposed Walsh demodulator employing the phase diversity technology is significantly distinguished from the other correlation results, that is,  $Z_{i, i \neq 32}$ , because the variance in computing each of the correlation results in the proposed system is reduced almost proportionally to the number of antenna elements, that is, 6 in our smart antenna BTS. Consequently, the performance of the Walsh demodulator is improved by adopting the array system with phase diversity.

Figure 10 illustrates an improvement of the proposed Walsh demodulator in terms of  $E_b/N_0$ . The value for  $E_b/N_0$  has been obtained from the ratio between  $E[Z_{32}]$  and the standard deviation in computing the average value of the decision variable. As shown in Figure 10,  $E_b/N_0$  in the Walsh demodulator is increased by about 2.5 ~ 4 dB.

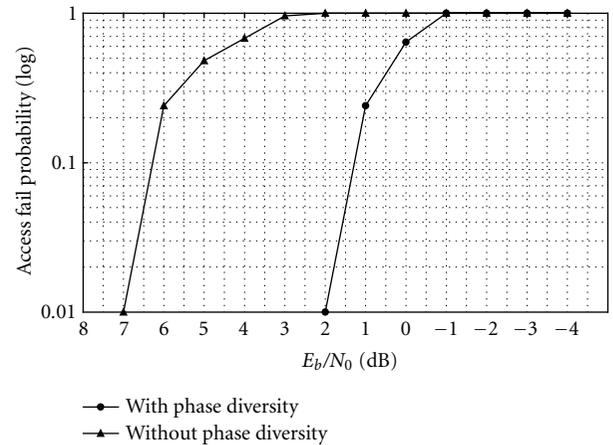


FIGURE 12: Access fail probability in the fading environment.

TABLE 1: Measured  $E_b/N_0$  at access channel decoder output.

Channel $E_b/N_0$ (dB)	$E_b/N_0$ at decoder output (dB)	
	Conventional	Phase diversity
5	11.2	16.1
6	11.8	16.7
7	12.5	17.3
8	13	17.9

#### 4.4. Performance of entire access channel decoder

According to the requirement regarding the access probe test, specified by IS-97D [21], the access failure probability should not exceed 1% when  $E_b/N_0$  is 6.5 dB. In this section, the access failure probability provided by the proposed access channel decoder is presented. Note that the access failure is detected from the failure of CRC contained in the received access probe.

Figures 11 and 12 illustrate the access failure probability in AWGN and fading circumstances, respectively. Note that the required  $E_b/N_0$  for the failure probability to be 1% is enhanced by about 5 dB. It particularly means that the smart antenna BTS employing the phase diversity can increase the cell size, compared to a conventional BTS consisting of a single antenna channel, so much that the access signal arrives at the BTS with the power of about 5 dB lower. Consequently, communication performance can be enhanced even during the access state for which the pilot data are not available such that the optimal parameters for nice beam pattern are not yet obtained.

Table 1 represents the measurements of  $E_b/N_0$  at the output port of the proposed access channel decoder of our smart antenna system prepared together with the noise generator and fading emulator as shown in Figure 13. As shown in the table, the proposed access channel decoder enhances the  $E_b/N_0$  by about 5 dB. Note that the measurements shown in Table 1 very much coincides with the simulation results shown in Figures 11 and 12.

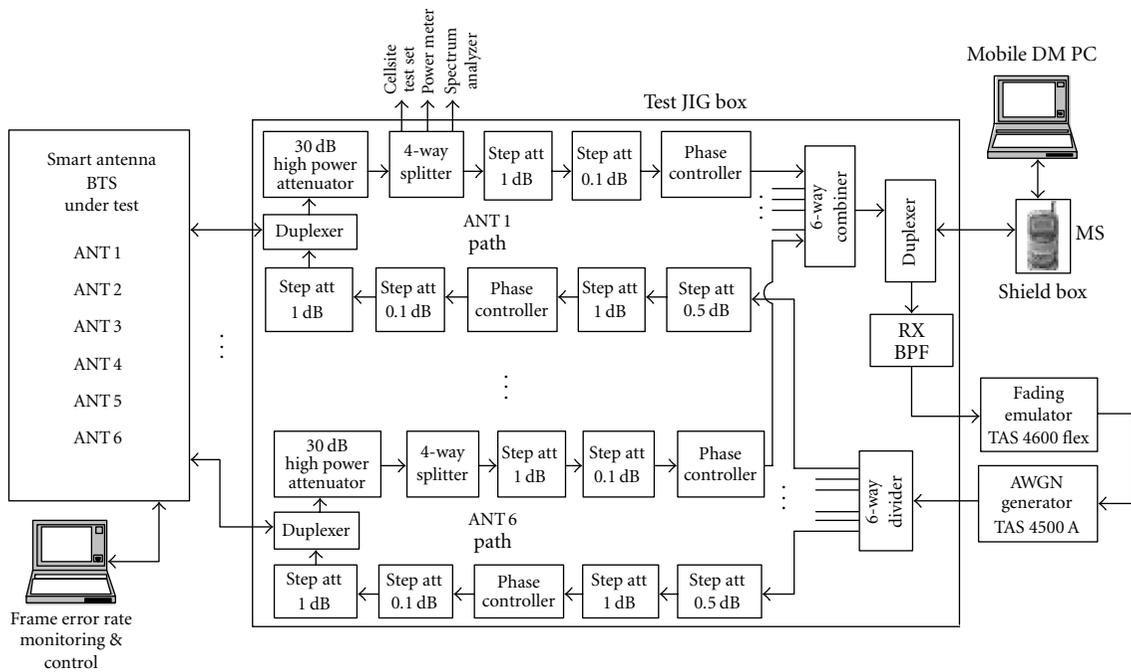


FIGURE 13: Experimental environment for access channel decoder of smart antenna BTS.

## 5. CONCLUSIONS

In this paper, we present an access channel decoder for the smart antenna BTS. The proposed decoder consists of a searcher, OQPSK demodulator, Walsh decoder, Viterbi decoder, and CRC checker. The proposed smart antenna system implemented in this paper exploits the gain of phase diversity with 6 antenna elements. The performance of the proposed access channel decoder has been demonstrated for the access searcher and Walsh demodulator in terms of enhanced detection/false alarm probability and improved  $E_b/N_0$ , respectively. The overall performance of the proposed access channel decoder has been presented in terms of access failure probability. As for the access probe test required by IS-97D [21], the proposed access channel decoder provides about 5 dB improvement in the minimum  $E_b/N_0$  required for the failure rate to be less than 1% both in AWGN and the fading circumstances. More specifically speaking, for a cell with a radius of 2 km, for example, the proposed technique can increase the distance of reverse search coverage by about 76% if the path loss in a given mobile environment is proportional to the fourth-power of the path distance [22]. The phase diversity technology proposed in this paper can be used in various systems of next generation mobile communications such as WCDMA systems as well as CDMA2000 1X systems.

## ACKNOWLEDGMENT

This work was supported by HY-SDR Research Center at Hanyang University, Seoul, Korea under the ITRC program of MIC, Korea.

## REFERENCES

- [1] F. Adachi, M. Sawahashi, and H. Suda, "Wideband DS-CDMA for next-generation mobile communications systems," *IEEE Communications Magazine*, vol. 36, no. 9, pp. 56–69, 1998.
- [2] A. J. Paulraj and B. C. Ng, "Space-time modems for wireless personal communications," *IEEE Personal Communications*, vol. 5, no. 1, pp. 36–48, 1998.
- [3] J. S. Lee and L. E. Miller, *CDMA Systems Engineering Handbook*, Artech House, Norwood, Mass, USA, 1998.
- [4] J. C. Liberti Jr., and T. S. Rappaport, *Smart Antennas for Wireless Communications: IS-95 and Third Generation CDMA Applications*, Prentice-Hall, Upper Saddle River, NJ, USA, 1999.
- [5] B. Widrow, P. E. Mantey, L. J. Griffiths, and B. B. Goode, "Adaptive antenna systems," *Proceedings of the IEEE*, vol. 55, no. 12, pp. 2143–2159, 1967.
- [6] A. J. Paulraj and C. B. Papadias, "Space-time processing for wireless communications," *IEEE Signal Processing Magazine*, vol. 14, no. 6, pp. 49–83, 1997.
- [7] B. Suard, A. F. Naguib, G. Xu, and A. J. Paulraj, "Performance analysis of CDMA mobile communication systems using antenna array," in *Proc. International Conference on Acoustics, Speech and Signal Processing*, vol. IV, pp. 153–156, Minneapolis, Minn, USA, April 1993.
- [8] J. S. Thompson, P. M. Grant, and B. Mulgrew, "Smart antenna arrays for CDMA systems," *IEEE Personal Communications*, vol. 3, no. 5, pp. 16–25, 1996.
- [9] T. Cover, R. McEliece, and E. Posner, "Asynchronous multiple-access channel capacity," *IEEE Trans. Inform. Theory*, vol. 27, no. 4, pp. 409–413, 1981.
- [10] F. Khaleghi and M. Kavehrad, "A new correlator receiver architecture for noncoherent optical CDMA networks with bipolar capacity," *IEEE Trans. Communications*, vol. 44, no. 10, pp. 1335–1339, 1996.

- [11] M. K. Varanasi and A. Russ, "Noncoherent decorrelative detection for nonorthogonal multipulse modulation over the multiuser Gaussian channel," *IEEE Trans. Communications*, vol. 46, no. 12, pp. 1675–1684, 1998.
- [12] H. Hwang, J. Park, Y. Kim, S. Ahn, and S. Choi, "Performance improvement of smart antenna system in terms of searching and tracking capability due to phase diversity technique in IS2000 1X signal environment," in *IEEE Antennas and Propagation Society International Symposium*, vol. 1, pp. 614–617, San Antonio, Tex, USA, June 2002.
- [13] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York, NY, USA, 1965.
- [14] 3GPP2 C.S0002-C, "Physical layer standard for cdma2000 spread spectrum systems," Release C, v1.0, pp. 3-23–3-26, May 2002.
- [15] H. Park, B. Kang, S. Bang, and Y. Han, "A maximum-likelihood technique for preamble search in a CDMA reverse link," in *1997 IEEE 47th Vehicular Technology Conference*, vol. 3, pp. 1614–1618, Phoenix, Ariz, USA, May 1997.
- [16] S. Choi, W. Lee, S. Hyeon, H. Bahk, C. Lee, and J. Kim, "Performance of smart antenna base station implemented for CDMA2000 1X," in *MPRG Annual Symposium on Wireless Personal Communications*, Blacksburg, Va, USA, June 2003.
- [17] S. Choi, H. M. Son, and T. K. Sarkar, "Implementation of a smart antenna system on a general-purpose digital signal processor utilizing a linearized CGM," *Digital Signal Processing*, vol. 7, no. 2, pp. 105–119, 1997.
- [18] Y. Kim, H. Im, J. Park, H. Bahk, J. Kim, and S. Choi, "Implementation of smart antenna base station with a novel searcher and tracker for CDMA2000 1X," in *IEEE 8th International Conference on Communication Systems*, vol. 1, pp. 394–398, Singapore, November 2002.
- [19] A. F. Naguib, *Adaptive antennas for CDMA wireless networks*, Ph.D. Dissertation, Department of Electrical Engineering, Stanford University, Stanford, Calif, USA, August 1996.
- [20] S. Choi, J. Choi, H. Im, and B. Choi, "A novel adaptive beamforming algorithm for antenna array CDMA systems with strong interferers," *IEEE Trans. Vehicular Technology*, vol. 51, no. 5, pp. 808–816, 2002.
- [21] 3GPP2 C.S0010-B, "Recommended minimum performance standards for cdma2000 spread spectrum base stations," Release B, v1.0, pp. 3-23–3-26, December 2002.
- [22] J. S. Lee, *Lectures on IS-95 CDMA Cellular System Design and the Principles of Cell Planning*, J. S. Lee Associates, Rockville, Md, USA, 1996.

**Seunghyeon Hyeon** received his B.S. and M.S. degrees in electronic communication engineering from Hanyang University, Seoul, Korea, in 2002 and 2004, respectively. Since 2002, he has been with the Communication Signal Processing Laboratory, Hanyang University, Seoul, Korea, where he had developed the smart antenna beamforming module and a DSP algorithm for real-time applications. His current research focuses on implementation of a smart antenna system for third-generation mobile communication systems.



**Seungwon Choi** received his B.S. degree from Hanyang University, Seoul, Korea, and his M.S. degree from Seoul National University, Korea, in 1980 and 1982, respectively, both in electronics engineering, his M.S. degree (computer engineering) in 1985, and his Ph.D. degree (electrical engineering) in 1988, both from Syracuse University, Syracuse, NY. From 1988 to 1989, he was with the Department of Electrical and Computer Engineering, Syracuse University, Syracuse, NY, as an Assistant Professor. In 1989, he joined the Electronics and Telecommunications Research Institute, Daejeon, Korea. From 1990 to 1992, he was with the Communications Research Laboratory, Tokyo, Japan, as a Science and Technology Agency Fellow, developing the adaptive antenna array systems and adaptive equalizing filters. He joined Hanyang University, Seoul, Korea, in 1992, as an assistant Professor. He is a Professor at the School of Electrical and Computer Engineering, Hanyang University. Since 2003, Dr. Choi has been serving as the representative of the ITU Region 3 for SDR (Software Defined Radio) Forum. His research interests include digital communications and adaptive signal processing with a recent focus on the implementation of the smart antenna systems for both mobile communication systems and wireless data systems.



**Seongdo Kim** was born in Daegu, Korea, on March 19, 1965. He received his B.S. and M.S. degrees in electronic communication engineering from Hanyang University, Seoul, Korea, in 1988 and 1990, respectively. He is pursuing the Ph.D. degree in the Department of Electronic Communication Engineering at Hanyang University, Seoul, Korea. From 1990 to 1999, he worked as a Senior Researcher at the Agency for Defence Development in Daejeon, Korea, engaged in the design of military communication system. From 2000 to 2002, he worked as a Senior Researcher at Hyundai electronic in Icheon, Korea, engaged in the design of wideband digital IF and RF subsystem of 3G BTS. In 2003, he joined Ace Technology, Seoul, Korea, where he is presently the Director of Media Communication Institute. His research interests are smart antenna algorithm, design of CDMA BTS system, and application of digital predistortion (DPD) technology.

