PERFORMANCE ANALYSIS OF SYMBOL TIMING ESTIMATORS FOR TIME-VARYING MIMO CHANNELS

A Thesis

by

FLAVIU GABRIEL PANDURU

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

August 2004

Major Subject: Electrical Engineering

PERFORMANCE ANALYSIS OF SYMBOL TIMING ESTIMATORS FOR TIME-VARYING MIMO CHANNELS

A Thesis

by

FLAVIU GABRIEL PANDURU

Submitted to Texas A&M University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Approved as to style and content by:

Erchin Serpedin (Chair of Committee)	
Costas N. Georghiades (Member)	Aydin Karsilayan (Member)
Marina Vannucci (Member)	Chanan Singh (Head of Department)

August 2004

Major Subject: Electrical Engineering

ABSTRACT

Performance Analysis of Symbol Timing Estimators
for Time-varying MIMO Channels. (August 2004)
Flaviu Gabriel Panduru, B.S.; M.S., Politehnica University, Bucharest, Romania
Chair of Advisory Committee: Dr. Erchin Serpedin

The purpose of this thesis is to derive and analyze the theoretical limits for estimating the symbol timing delay of Multiple-Input Multiple-Output (MIMO) systems. Two main $N \times M$ system models are considered, where N represents the number of transmit antennas and M denotes the number of receive antennas, the 2×2 system used by S.-A. Yang and J. Wu and the 4×4 system used by Y.-C. Wu and E. Serpedin. The second model has been extended to take into account the symbol time-varying fading. The theoretical estimation limits are shown by several bounds: modified Cramer-Rao bound (MCRB), Cramer-Rao bound (CRB) and Barankin bound (BB). BB will be exploited to obtain accurate information regarding the necessary length of data to obtain good estimation. Two scenarios for synchronization are presented: data-aided (DA) and non-data-aided (NDA). Two models for the fading process are considered: block fading and symbol time-varying fading, respectively, the second case being assumed to be Rayleigh distributed. The asymptotic Cramer-Rao bounds for low signal-to-noise ratio (low-SNR) and for high-SNR are derived and the performance of several estimators is presented. The performance variation of bounds and estimators is studied by varying different parameters, such as the number of antennas, the length of data taken into consideration during the estimation process, the SNR, the oversampling factor, the power and the Doppler frequency shift of the fading.

To my parents

ACKNOWLEDGMENTS

Among all the people I should express my gratitude for, my adviser deserves particular attention and many thanks, not only for the insight in the thesis writing process and for the excellent academic supervision, but for actually changing my life by motivating me in doing research and writing this thesis. Dr. Serpedin helped me overcome the difficulties and frustration encountered in the research process and offered me guidance and support when I needed it the most. I am really grateful, and I consider myself fortunate to have had such an adviser.

I want to thank Dr. Georghiades for his supportive and enlightening attitude towards me and other students and for making the wireless group such a nice environment. For his generosity and integrity I consider him to be a moral example for everyone. I admire and appreciate him also for his incredible capacity of always having a smile and a good word for his students.

I also wish to express my appreciation and thank my other committee members, Dr. Karsilayan, Dr. Matis and Dr. Vannucci for their kindness and supportiveness.

I thank all my colleagues and friends, especially Yik-Chung and Pradeep, for their useful advice. Finally, many thanks to Alireza for being such a trustful friend.

TABLE OF CONTENTS

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

FIGURE Page

CHAPTER I

INTRODUCTION

A. Motivation and Problem Statement

The estimation of the timing delay is of crucial importance for communications systems because of the need to maximize the signal-to-noise ratio and to reduce the inter-symbol interference.

Due to the sampling at the receiver side, the transmit and the receive signals, as well as the noise associated with the receiver, can be expressed in vectorial form. Therefore, the propagation channel between the base station and the mobile can be modelled as a matrix. In the case of block fading, each channel associated with a pair of antennas is considered to be characterized by a complex constant for the whole period of a transmission block. The elements of the channel matrix can be assumed independent or correlated with a fixed correlation matrix. In the case of symbol-level time-varying (TV) fading, the propagation coefficients for each pair of antennas are characterized by time-varying processes, characterized by known or quasi-known statistical properties. These processes are sampled with the sampling rate $1/T_s$ and the received signal can be expressed in a new vectorial form.

The use of multiple antennas in a communications system is equivalent with adding new dimensions. The spatial diversity introduced is an effective mean to increase the system's performance. The capacity of MIMO antenna systems has been proven in the literature to be much higher than the one obtained in the case of single antenna systems. The timing estimation problem becomes more complicated, but the ability to express all signals in matrix form allows the solution to be expressed in a

---

The journal model is *IEEE Transactions on Automatic Control.*

compact form and to evidence the gain produced by the multi-antenna diversity.

The analysis of the performance of known estimators is a first step in the process of finding better estimators and optimizing the system parameters. To evaluate the accuracy of a certain estimator, theoretical bounds are needed, in order to show how far the performance of the suboptimal realizable estimator is from the optimal theoretical limit. The mean square error is a good indicator of performance because the variance of the timing error is a meaningful parameter to determine. From this point of view, the computation of bounds like the Cramer-Rao bound and the Barankin bound becomes very justified.

Finally, identifying the parameters that influence the performance of the system, as well as studying their individual and joint contributions to the improvement or degradation of the accuracy of the estimation, is an important step in the design process.

B. Outline of the Thesis

Chapter II is providing some background about the main issue of this thesis, the timing recovery. The process of timing recovery is presented in the general frame of synchronization for communications systems. After highlighting the importance of the timing delay estimation, the mean square error is presented as being a good performance criterion and the main bounds used in this thesis, the Cramer-Rao bound, the modified Cramer-Rao bound, and the Barankin bound, are introduced, briefly pointing their main individual advantages and characteristics. Chapter III describes the propagation and system models used in this thesis. The fading variations are assumed to respect the Jakes' model and can be simulated by using autoregressive methods or the eigendecomposition of the autocorrelation matrix. The main system

model is introduced in Chapter III.C and is extended in order to account for the TV fading. Chapter IV introduces the Cramer-Rao bound and the Barankin bound as the main performance indicator limits used in this thesis, presents the methods of determining them (with most of the computations detailed in Appendix A-E) and shows the obtained results. In Chapter V a detailed analysis of the system models is presented. A cost function is introduced, an optimization method is proven to provide better results in certain conditions, and the asymptotic Cramer-Rao bounds are presented as auxiliary performance indicators. The variations in the estimation accuracy are studied by modifying different channel or system parameters, the Barankin bound is exploited to provide information about the required length of the observation and the performance of several estimators is compared. Finally, Chapter VI summarizes the thesis, presenting some conclusions and suggestions for future research work.

CHAPTER II

BACKGROUND

A. Symbol Timing Recovery

For accurate recovery of the transmitted data symbols, the communication receivers require knowledge about the channel. However, in many practical situations, this information is not available from the beginning. Therefore, several channel parameters need to be estimated using the information contained in the received signal. The main parameters required are the carrier frequency, the carrier phase and the timing delay [1]. In a typical coherent receiver, the received signal is first coherently demodulated and low-pass filtered to recover the modulating message signal. The next step is the sampling of the message signal at the symbol rate and the recovery of the transmitted data symbols. Although the receiver generally knows the symbol rate, it does not know when to sample the signal for the best signal-to-noise ratio (SNR) performance. The objective of the symbol timing recovery is to find the best instants for sampling the received signal.

Generally, in communication systems, the process of synchronization follows a certain pattern. First synchronization step is the acquisition of the carrier. In coherent demodulators, the achievement of the carrier recovery consists of the generation of a reference carrier whose phase equals the one of the transmitted signal carrier. The discrepancy in frequency comes from the deviation of the transmitter and receiver oscillators and from the Doppler effect [1]. The carrier phase synchronization is generally achieved by a phase-locked-loop (PLL) circuit. Throughout this thesis, no carrier frequency error is assumed, except the frequency Doppler shift, which is considered non-zero in the symbol level time-varying fading scenario case.

The second synchronization step, which is the focus of this thesis, is the symbol synchronization, which is a requirement for all digital communication systems which transmit the information synchronously. Assuming a symbol period of T and a relative delay ε, the samples at $t = kT + \varepsilon T$ are required in order to reduce the inter-symbol interference (ISI) effects. Assuming a sampling period of T_s, only the samples at moments kT_s are available. As described in [2], for analog receivers the synchronization relies on controlling the sampling instants of the received signal, whereas for digital receivers there will always exist a small difference between T_s and T that will produce, on long term, cycle slips. Therefore, it is necessary to obtain samples of a matched filter output at the symbol rate $1/T$ from the signal samples available at the rate $1/T_s$. A second problem for synchronous sampling, due to the generally unknown propagation delay on the communication chain between the transmitter and the receiver, is that the symbol timing must be derived from the received signal. These two synchronization issues are generally referred to as low-level synchronization. The so-called high-level synchronization deals with discrete parameters like words, frames, and packets.

The errors in the estimation of ε degrade the overall performance of the system, increasing the ISI. Ultimately, these errors increase the bit error probability and decrease the performance of the communication system. The symbol timing synchronization is generally achieved using a delay-locked-loop (DLL) circuit.

In some cases, when the data is *a priori* known, this knowledge can be used to improve the estimation. There are two cases when this assumption is valid. When a known data sequence is transmitted, such as during a preamble of a data packet, the data-aided (DA) recovery techniques are used. Even after the initial training period, if the SNR is sufficiently high, the estimated symbols can be used by the same DLL circuit and the same estimation procedure can continue for the whole duration of

transmission. This method is called decision-directed (DD).

When the SNR is low, the data is not sufficiently accurate to be used in the timing recovery process, and non-data-aided (NDA) techniques have to be used. The classical approaches to NDA synchronization are presented in [3]. The first method uses the signal's cyclostationarity, usually the second-order statistics [4, 5, 6] and the second method relies on treating the data symbols as random variables and using the maximum-likelihood (ML) principle. In the second case, the SNR is generally considered to be low and most of the algorithms are derived as approximations of the ML estimator. One of the most classical algorithms that exploits the cyclostationarity is the one proposed by Oerder and Meyr (O&M) in [7], and one of the well known detectors using the ML approach is the one proposed by Gardner in [8].

B. The Cramer-Rao Lower Bound

For timing estimation schemes, the variance of the timing error is a good measure of the system performance. However, for most practical suboptimal estimators, the variance of the estimation error cannot be determined analytically. Therefore, the derivation of good bounds becomes helpful in the sense of showing how far a particular suboptimal realizable estimator is from the optimal estimator in terms of its variance. The Cramer-Rao bound (CRB) is a fundamental lower bound on the error variance of unbiased estimators and a benchmark for many practical estimators. Under some regularity assumptions and for a large number of observations, the CRB is generally achievable.

When computing the Cramer-Rao bound, one of the assumptions made is the differentiability of the likelihood function with respect to the parameter to determine. When working with digital signals, this assumption calls for a need for oversampling,

such that a good approximation of the continuous signal can be obtained using a dense set of samples.

For a particular model, the CRB depends on the particular sequence of training symbols (in the DA case), or on the statistics of the data used (in the NDA case).

For a scalar parameter, the CRB is defined as follows [9]. If the probability density function $p(\mathbf{r};\varepsilon)$ satisfies the regularity condition

$$E\left[\frac{\partial \ln p(\mathbf{r};\varepsilon)}{\partial \varepsilon}\right] = 0 \qquad \text{for all } \varepsilon \ ,$$

where the expectation is taken with respect to $p(\mathbf{r};\varepsilon)$, the variance of any unbiased estimator $\hat{\varepsilon}$ must satisfy

$$\text{var}(\hat{\varepsilon}) \geq \frac{1}{-E\left[\frac{\partial^2 \ln p(\mathbf{r};\varepsilon)}{\partial \varepsilon^2}\right]} = \frac{1}{-\int \frac{\partial^2 \ln p(\mathbf{r};\varepsilon)}{\partial \varepsilon^2} p(\mathbf{r};\varepsilon) d\mathbf{r}} \ , \tag{2.1}$$

where the derivative is evaluated at the true value of ε and the expectation is taken with respect to $p(\mathbf{r};\varepsilon)$.

For a vector parameter $\varepsilon = [\ \varepsilon_1 \quad \varepsilon_2 \quad \cdots \quad \varepsilon_p\ ]^T$, assuming the estimator $\hat{\varepsilon}$ is unbiased, the CRB places a bound on the variance of each element [9]

$$\text{var}(\hat{\varepsilon_i}) \geq [J^{-1}(\varepsilon)]_{ii} \ ,$$

where $J(\varepsilon)$ is the $p \times p$ Fisher information matrix, defined as

$$[J(\varepsilon)]_{ij} = -E\left[\frac{\partial^2 \ln p(\mathbf{r};\varepsilon)}{\partial \varepsilon_i \partial \varepsilon_j}\right] \ ,$$

for $i, j = 1, 2, \cdots, p$.

The matrix J is in general symmetrical and positive definite.

A lot of literature has been dedicated to the computation of the CRB for various communications models. The CRB for DA timing recovery has been presented in [10]. Reference [10] also reveals the connection between a particular training sequence

and its performance limit and indicates how to compute the CRB in the time and frequency domain, respectively. Reference [11] computes the CRB for linearly modulated signals with no ISI in the NDA case, while [12] takes into consideration the time-selective fading for amplitude phase-modulated signals, assuming that both the multiplicative and the additive white noise are Gaussian. For deterministic signals, the CRB is also computed in [13], assuming the presence of both multiplicative and additive noise. An example of extension of the CRB computation to MIMO communication channels is derived in [14] for frequency offset estimation, assuming a block fading model.

The CRB is a valid bound for any unbiased quadratic estimator, but in many cases the computation of the CRB becomes very difficult, because the statistics of the observation depend not only on the vector parameter to be estimated, but also on some nuisance parameters that normally are not estimated. These difficulties created a need for a simpler, yet looser bound, the modified Cramer-Rao bound (MCRB), which will be described in the next section. An unifying approach for derivation of various bounds used in constrained optimization problems has been presented in [15]. Reference [15] explains that even though the CRB can be asymptotically reached by the maximum-likelihood method for high SNR and a large number of independent snapshots, it is too optimistic under more practical estimation conditions (low SNR, small number of snapshots). In this sense, better (tighter) bounds, like the Battacharyya bound and the Barankin bound have been proposed. In the authors' view, the Barankin bound is not as popular as the CRB because of the increased computational cost and the lack of underlying assumptions in the bound derivation.

C. The Modified Cramer-Rao Bound

The modified Cramer-Rao bound (MCRB) proves to be useful in many practical situations where the true CRB is difficult to compute. Although the MCRB is always looser than the true CRB, it is generally easier to obtain, and in many cases is as tight as the true CRB at high SNR. This bound has been introduced for the scalar case in [16] and extended to a vectorial form in [17]. A comparison between the true CRB and the MCRB in the scalar case is presented in [18], evidencing that at high SNR, when the scalar parameter is coupled with the nuisance parameters, the MCRB can be quite loose compared to the true CRB. This paper further presents several cases where the MCRB practically coincides with the asymptotic Cramer-Rao bound (ACRB), defined as the analytical approximation of the true CRB for high SNR. An expression of the MCRB used in the block fading case for one of the two models analyzed in this thesis is explained in [3].

The MCRB is defined as follows [16]. Considering $\hat{\varepsilon}(\mathbf{r})$ to be an unbiased estimator of ε and denoting by $\mathbf{u}$ the random vector of the nuisance parameters, representing all the parameters that do not have to be estimated, including the data, with known probability density function $p(\mathbf{u})$, independent of ε, the joint probability density function (pdf) can be expressed as

$$p(\mathbf{r};\varepsilon) = \int_{-\infty}^{\infty} p(\mathbf{r} \mid \mathbf{u};\varepsilon)p(\mathbf{u})d\mathbf{u} = E_{\mathbf{u}}[p(\mathbf{r} \mid \mathbf{u};\varepsilon)] \ . \tag{2.2}$$

Generally, $p(\mathbf{r} \mid \mathbf{u};\varepsilon)$ is easily available. In many cases, the computation of (2.1) is untractable, either because the integration in (2.2) is difficult or because the expectation in (2.1) cannot be computed. By interchanging the order of the expectation and the logarithm operators, a new bound, easier to compute, is obtained. This bound is

named the modified Cramer-Rao bound

$$\mathrm{MCRB}(\varepsilon) = \frac{1}{E_{\mathbf{u}}\left[E_{\mathbf{r}|\mathbf{u}}\left[-\frac{\partial^2 \ln(p(\mathbf{r}|\mathbf{u};\varepsilon))}{\partial \varepsilon^2}\right]\right]} ,$$

where $E_{\mathbf{r}|\mathbf{u}}[\cdot]$ denotes the expectation with respect to the pdf $p(\mathbf{r} \mid \mathbf{u}; \varepsilon)$.

The Cramer-Rao bound (2.1) can be also expressed as

$$\mathrm{CRB}(\varepsilon) = \frac{1}{E\left[\left[\frac{\partial \ln p(\mathbf{r};\varepsilon)}{\partial \varepsilon}\right]^2\right]} = \frac{1}{\int \left[\frac{\partial \ln p(\mathbf{r};\varepsilon)}{\partial \varepsilon}\right]^2 p(\mathbf{r};\varepsilon) d\mathbf{r}} ,$$

and the corresponding expression for the MCRB is

$$\mathrm{MCRB}(\varepsilon) = \frac{1}{E_{\mathbf{u}}\left[E_{\mathbf{r}|\mathbf{u}}\left[\frac{\partial \ln(p(\mathbf{r}|\mathbf{u};\varepsilon))}{\partial \varepsilon}\right]^2\right]} .$$

In general, it can be shown that $\mathrm{MCRB}(\varepsilon) \leq \mathrm{CRB}(\varepsilon)$.

D. The Barankin Bound

As described in the previous section, in certain conditions the maximum-likelihood estimators can perform very close to the Cramer-Rao bound. Thus, the CRB is a good bound provided that the observation time is sufficiently large and the SNR large enough. When these conditions do not hold, a threshold effect can be noticed. Below a critical SNR value, the optimum achievable estimator deviates radically from the CRB. Barankin [19] proposed a class of lower bounds for unbiased estimators. The Barankin bound (BB) is a tighter bound than the CRB and, through the threshold effect, gives additional information about the required observation length. The most general version of the BB is the greatest lower bound on the variance of any unbiased estimator and can be asymptotically achieved. For most practical problems, this version of the BB is too complex to be evaluated. Therefore, most derivations in the literature use slightly simplified variants. Using the Schwarz inequality, the author in

[20] derived lower bounds in the case of scalar valued parameters. Using these bounds in the case of pulse-position-modulated (PPM) signals, it has been proven in [21] that the threshold effect characterizes all nonlinear modulation communications schemes. A geometric interpretation of the BB has been given in [22]. For the problem of time delay estimation, the BB is determined in [23] using a dense set of test points and the threshold phenomenon is analyzed in [24].

The BB is defined as follows [25]. Assume Ω a sample space of points ω and $P(\omega \mid \varepsilon)$ a family of probability measures on Ω indexed by a parameter ε, taking values in the index set π and having a density function with respect to a measure μ. For all measurable sets E

$$P(E \mid \varepsilon) = \int_E p(\omega \mid \varepsilon) d\mu(\omega) \ .$$

For $g(\cdot)$ a function defined on π and for $\hat{g}(\varepsilon)$ an unbiased estimator of $g(\varepsilon)$

$$\int \hat{g}(\omega) p(\omega \mid \varepsilon) d\mu(\omega) = g(\varepsilon) \ , \quad \forall \varepsilon \in \pi \ .$$

The BB takes the form

$$\mathrm{var}(\hat{g}) \geq \sup_{\varepsilon_i, a_i} \frac{\left\{\sum_{i=1}^n a_i \left[g(\varepsilon_i) - g(\varepsilon)\right]\right\}^2}{\int \left[\sum_{i=1}^n a_i L(\omega; \varepsilon_i, \varepsilon)\right]^2 p(\omega \mid \varepsilon) d\mu(\omega)} \ ,$$

where $L(\omega; \varepsilon_i, \varepsilon) = p(\omega \mid \varepsilon_i)/p(\omega \mid \varepsilon)$ and the supremum is to be taken over all finite families of $\varepsilon_i \in \pi$ and real valued $a_i's$.

A more general bound for a hybrid parameter vector (including some deterministic and some random entries) has been derived in [26]. The BB is a particular case of the above mentioned bound for deterministic parameter estimation and the Bobrovsky-Zakai bound is a particular case for random parameter estimation. References [27, 26] show that the inclusion of additional test points does not reduce the BB. Also, [26] shows that, although there is no formal procedure for choosing the

test points, usually these points are the ones that locally maximize the correlation function and that in most practical examples the insertion of additional points does not change the bound significantly, so that in many cases even one properly chosen test point is good enough.

A simplified version of the BB, considering a single test point, is the one proposed by Chapman and Robbins in [28]. This version is broadly used in the literature [27, 29] and this is the version used in this thesis. For the estimation of a scalar real parameter ε, the Chapman-Robbins version of the BB can be presented as follows [29]. Denoting by $p(\mathbf{r};\varepsilon)$ the pdf of a vector $\mathbf{r}$, for a given ε, and considering η to be a real number independent of $\mathbf{r}$ such that $\varepsilon+\eta$ ranges over all possible values of ε, any unbiased estimator $\hat{\varepsilon}$ satisfies $\mathrm{var}(\hat{\varepsilon}) \geq \mathrm{BB}$, where BB is given by

$$\mathrm{BB} = \sup_{\eta} \frac{\eta^2}{\int \frac{p(\mathbf{r};\varepsilon+\eta)^2}{p(\mathbf{r};\varepsilon)} d\mathbf{r} - 1} \geq \mathrm{CRB} \ , \tag{2.3}$$

and the CRB is given by

$$\mathrm{CRB} = \lim_{\eta \to 0} \frac{\eta^2}{\int \frac{p(\mathbf{r};\varepsilon+\eta)^2}{p(\mathbf{r};\varepsilon)} d\mathbf{r} - 1} = \frac{1}{\mathrm{var}\left[\frac{\partial \ln p(\mathbf{r};\varepsilon)}{\partial \varepsilon}\right]} \ .$$

The following ratio measures the deviation of the BB from the CRB

$$\frac{\mathrm{BB}}{\mathrm{CRB}} = \sup_{\eta} \frac{\eta^2 \mathrm{var}\left[\frac{\partial \ln p(\mathbf{r};\varepsilon)}{\partial \varepsilon}\right]}{\int \frac{p(\mathbf{r};\varepsilon+\eta)^2}{p(\mathbf{r};\varepsilon)} d\mathbf{r} - 1} \ .$$

When this ratio equals 1, the supremum is reached for $\eta = 0$ and the two bounds coincide. When the ratio is larger than 1, there is a threshold effect and the supremum is reached for $\eta \neq 0$.

CHAPTER III

MODELS

A. Modelling the Time-Varying Fading

For an $N \times M$ MIMO system there exists a number of NM propagation channels. The assumption considered throughout this thesis is that these channels are independent and follow the same statistics. For each channel, the fading is modelled as a complex Gaussian process and the statistic properties of the channel coefficients depend on the Doppler frequency $f_d = v f_0/c$, where v is the speed of the mobile, c is the speed of light and f_0 is the carrier frequency of the communication system. The time-variance of the system is evidenced by the autocorrelation function of the fading components.

Under some conditions (see, for example, [30]), the Jakes' model can be assumed and the power spectral density (PSD) associated with each fading component is given by

$$S(f) = \begin{cases} \frac{\sigma_\alpha^2}{\pi f_d \sqrt{1-\left(\frac{f}{f_d}\right)^2}} & \mid f \mid \leq f_d \\ 0 & \text{elsewhere} \end{cases} . \tag{3.1}$$

The corresponding continuous time autocorrelation function is

$$R(\tau) = \sigma_\alpha^2 J_0(2\pi f_d \tau) \ , \tag{3.2}$$

where $J_0(\cdot)$ denotes the Bessel function of the first kind and zero order and σ_α^2 denotes the fading's variance. The discrete version of the autocorrelation function is

$$R[n] = \sigma_\alpha^2 J_0(2\pi f_d \mid n \mid T_s) \ , \tag{3.3}$$

which is equivalent with the normalization of the Doppler frequency by the sampling rate of $1/T_s$ [30]. This is the theoretical autocorrelation model used throughout this

thesis.

For simulation purposes, methods to generate a fading whose PSD approximates the one given by (3.1) are needed. Some of the most effective methods include the use of autoregressive (AR) models. As a general rule, the higher the order of the model, the better the approximation and the bigger the complexity. Several AR models have been presented in [31]. A powerful method is described in [30, 32]. Reference [32] also shows how to take into account the multiple cross-correlations of the fading. For simulation purposes, this thesis follows the method described in [32] to generate independent realizations of the fading, considering an AR model of order 5. The method, as presented in the above-mentioned paper, consists of generating correlated complex Gaussian processes $\mathbf{h}[n]$, according to the p^{th} order autoregressive model

$$\mathbf{h}[n] = -\sum_{k=1}^{p} \mathbf{A}[k]\mathbf{h}[n-k] + \mathbf{w}[n] \ ,$$

where $\mathbf{A}$[k] are $V \times V$ matrices containing the AR coefficients and

$$\mathbf{w}[n] = \left[\begin{array}{cccc} w_1[n] & w_2[n] & \cdots & w_V[n] \end{array} \right]^T$$

is the driving noise, considered to be complex white Gaussian with zero mean and covariance matrix $\mathbf{Q} = E\{\mathbf{w}[n]\mathbf{w}[n]^H\}$. For an observation matrix

$$\underline{\mathbf{H}}[n] = \left[\begin{array}{cccc} \mathbf{h}[n-1]^T & \mathbf{h}[n-2]^T & \cdots & \mathbf{h}[n-p]^T \end{array} \right]^T \ ,$$

the model covariance matrix is defined as $\mathbf{R_{hh}} = E\{\underline{\mathbf{H}}[n]\underline{\mathbf{H}}[n]^H\}$.

Using the Levinson-Wiggins-Robinson (LWR) algorithm, the coefficients $\mathbf{A}[k]$ can be determined by solving the following multichannel Yule-Walker system of equa-

tions

$$\begin{pmatrix} \mathbf{R_{hh}}[0] & \mathbf{R_{hh}}[-1] & \cdots & \mathbf{R_{hh}}[-p+1] \\ \mathbf{R_{hh}}[1] & \mathbf{R_{hh}}[0] & \cdots & \mathbf{R_{hh}}[-p+2] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R_{hh}}[p-1] & \mathbf{R_{hh}}[p-2] & \cdots & \mathbf{R_{hh}}[0] \end{pmatrix} \begin{pmatrix} \mathbf{A}^H[1] \\ \mathbf{A}^H[2] \\ \vdots \\ \mathbf{A}^H[p] \end{pmatrix} = - \begin{pmatrix} \mathbf{R_{hh}}[1] \\ \mathbf{R_{hh}}[2] \\ \vdots \\ \mathbf{R_{hh}}[p] \end{pmatrix} .$$

As indicated in [32], the $V \times V$ covariance matrix of the driving noise vector process can be computed as

$$\mathbf{Q} = \mathbf{R_{hh}}[0] + \sum_{k=1}^{p} \mathbf{R_{hh}}[-k]\mathbf{A}^H[k] \ .$$

The next step is to perform the Cholesky factorization $\mathbf{Q} = \mathbf{G}\mathbf{G}^H$. Finally, the driving process is generated as $\mathbf{w}[n] = \mathbf{G}\mathbf{z}$, where $\mathbf{z}$ is a $V \times 1$ vector of independent zero mean complex Gaussian elements with unit variance.

B. Eigendecomposition-Based Model

This section introduces briefly the first model used in this thesis. The approach taken in this model, described in detail in [33], has the advantage of accommodating an ML estimator of the time delay based on the eigendecomposition of the autocorrelation matrix. The time-varying channels' gains are considered multiplicative distortions modelled as linear combinations of the eigenfunctions. Although it can be easily extended to an arbitrary number of antennas, this model, as presented in [33], assumes a MIMO system with $N = 2$ transmitting antennas and $M = 2$ receiving antennas.

The continuous time transmitted signals are given by

$$d_i(t) = \sum_{k=1}^{L_0} d_{ik} g(t - kT) \qquad i = 1, 2 \ ,$$

where L_0 denotes the length of the training sequence, i is the index of the correspond-

ing antenna, T denotes the symbol time duration, d_{ik} stands for the training sequence corresponding to the transmit antenna i, and $g(t)$ is the square-root raised cosine pulse with roll-off factor β. Throughout all simulations, a factor $\beta = 0.3$ has been considered. The data is considered to be generated by the orthogonal sequences of length $L_0 = 12$ [33]: $d_{1k} = [+----+--+++-]$ and $d_{2k} = [++-++--+---+-++]$.

Let $h_{ij}(t)$ and $n_j(t)$ denote the multiplicative fading between the transmit antenna i and receive antenna j and the cyclic symmetric complex Gaussian noise associated with receiving antenna j, respectively $(i, j = 1, 2)$. The signal $r_j(t)$ received at antenna j can be expressed as

$$r_j(t) = \sum_{i=1}^{2} d_i(t - \varepsilon)h_{ij}(t) + n_j(t) \ . \tag{3.4}$$

For Rayleigh fading, $h_{ij}(t)$ is modelled as a zero-mean Gaussian random process with the autocorrelation function

$$R_\alpha(t, u) = E[h_{ij}(t)h_{ij}^*(u)] = \sigma_\alpha^2 J_0(2\pi f_d(t - u)) \ ,$$

where $*$ represents the complex conjugate. This formula is similar to (3.2). The next step is to find the eigenfunctions $f_k(t)$ and the eigenvalues λ_k that satisfy

$$\int_{T_i}^{T_f} R_\alpha(t, u)f_k(u)du = \lambda_k f_k(t) \ , \tag{3.5}$$

where T_i and T_f are the limits of the observation time. Using a good resolution, (3.5) can be approximated with its discrete version.

Next, the fading is expressed as

$$h_{ij}(t) = \lim_{S\to\infty} h_{ijS}(t) \ ,$$

where

$$h_{ijS}(t) = \sum_{k=1}^{S} c_{ijk} f_k(t) \ , \tag{3.6}$$

each c_{ijk} is a complex Gaussian random variable with variance λ_k, and S stands for the number of eigenvalues taken into consideration during the eigen-decomposition.

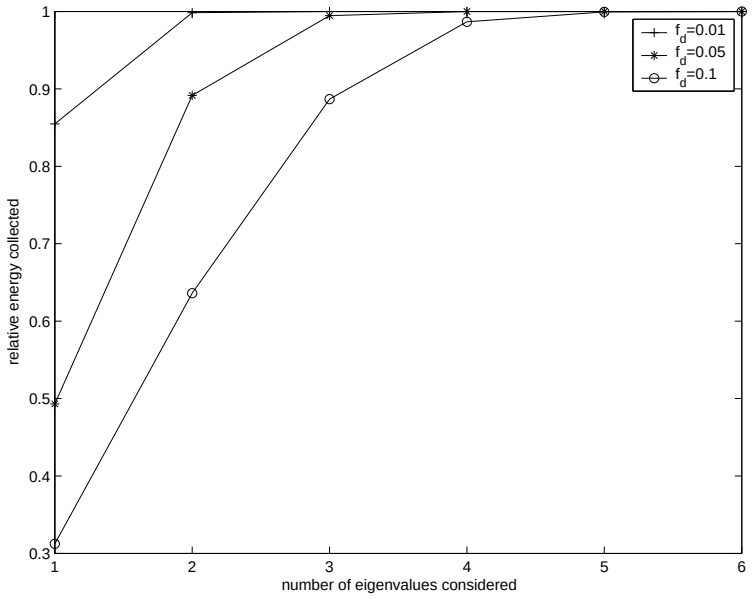


Fig. 1. Relative energy function of the number of eigenvalues considered

For simulation purposes, it is useful to decide how many eigenvalues S are needed for each value of the Doppler shift f_d in order to collect most of the fading energy. The result in represented in Fig. 1 where, for good approximation, an oversampling factor $Q = 16$ has been used in formula (3.5). Fig. 1 shows that even for very fast fading ($f_d = 0.1$), $S = 5$ assures an excellent approximation. For slow fading ($f_d = 0.01$), even $S = 2$ is enough. In future simulations $S = 4$ will be considered in (3.6) and (5.1).

C. Main Model

1. Block Fading

This section presents the main model used throughout this thesis as it has been described in [34] and [3], and the next section extends it to accommodate the time-varying fading. Let N denote the number of transmit antennas and M the number of receive antennas. Let h_{ij} be the complex channel coefficient between the i^{th} transmit antenna and the j^{th} receive antenna. The channel transfer function can be compactly described by the matrix

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{21} & \cdots & h_{N1} \\ h_{12} & h_{22} & \cdots & h_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ h_{1M} & h_{2M} & \cdots & h_{NM} \end{bmatrix} . \tag{3.7}$$

The complex envelope of the received signal at the j^{th} receive antenna is

$$r_j(t) = \sqrt{\frac{E_s}{NT}} \sum_{i=1}^{N} h_{ij} \sum_{m} d_i(m) g(t - mT - \varepsilon T) + n_j(t) \ ,$$

which, after sampling at a rate $1/T_s$, can be expressed as

$$r_j(kT_s) = \sqrt{\frac{E_s}{NT}} \sum_{i=1}^{N} h_{ij} \sum_{m} d_i(m) g(kT_s - mT - \varepsilon T) + n_j(kT_s) \ ,$$

where $k = 0, 1, \cdots, L_0 Q - 1$. In these formulas, E_s/N is the normalized symbol energy, T is the symbol duration, $d_i(m)$ represents the zero-mean complex-valued symbols transmitted from antenna i and $g(t)$ is the square-root raised cosine pulse with roll-off factor β. Throughout the simulations it has been considered a roll-off factor $\beta = 0.3$. The parameter ε represents the timing offset to estimate, considered to take values in the interval $[0, 1]$. The term $n_j(t)$ is the noise associated with the

j^{th} receive antenna and it is assumed to be a complex valued, circularly-distributed white Gaussian process with power spectral density N_0. For an oversampling factor $Q \geq 2$, the sampling rate is $f_s = 1/T_s = Q/T$. With L_0 representing the number of symbols observed and L_g the number of symbols affected by ISI on each side of $g(t)$, the number of consecutive samples becomes $L_0 Q$ and the sampled form of r_j can be expressed as a product of matrices

$$\mathbf{r}_j = \xi \mathbf{A}_\varepsilon \mathbf{Z} \mathbf{H}_{j,:}^T + \mathbf{n}_j \ ,$$

where $\xi = \sqrt{E_s/NT}$, $\quad \mathbf{r}_j = [r_j(0) \quad r_j(T_s) \quad \cdots \quad r_j((L_0Q-1)T_s)]^T$,

$\mathbf{A}_\varepsilon = [\mathbf{a}_{-L_g}(\varepsilon) \quad \mathbf{a}_{-L_g+1}(\varepsilon) \quad \cdots \quad \mathbf{a}_{L_0+L_g-1}(\varepsilon)]$,
$\mathbf{a}_i(\varepsilon) = [g(-iT - \varepsilon T) \quad g(T_s - iT - \varepsilon T) \quad \cdots \quad g((L_0Q-1)T_s - iT - \varepsilon T)]$,
$\mathbf{Z} = [\mathbf{d}_1 \quad \mathbf{d}_2 \quad \cdots \quad \mathbf{d}_N]$, $\qquad \mathbf{d}_i = [d_i(-L_g) \quad d_i(-L_g+1) \quad \cdots \quad d_i(L_0+L_g-1)]^T$,
$\mathbf{n}_j = [n_j(0) \quad n_j(1) \quad \cdots \quad n_j(L_0Q-1)]^T$.

For all simulations, unless otherwise stated, it has been considered that $L_0 = 32$ and $L_g = 4$.

After stacking all the received vectors $\mathbf{r}_j$ one under another, the following general model is obtained

$$\mathbf{r} = \xi(\mathbf{I}_M \otimes \mathbf{A}_\varepsilon)\mathrm{vec}(\mathbf{Z}\mathbf{H}^T) + \mathbf{n} \ , \tag{3.8}$$

where $\mathbf{r} = \begin{bmatrix} \mathbf{r}_1^T & \mathbf{r}_2^T & \cdots & \mathbf{r}_M^T \end{bmatrix}^T$ and $\mathbf{n} = \begin{bmatrix} \mathbf{n}_1^T & \mathbf{n}_2^T & \cdots & \mathbf{n}_M^T \end{bmatrix}^T$. Notation $\otimes$ denotes the Kronecker product and $\mathrm{vec}(\cdot)$ is the operator that stacks the columns of a matrix one under another.

After some straightforward processing, formula (3.8) can be adapted to a particular model, corresponding to the data-aided (DA) case

$$\mathbf{r} = \xi(\mathbf{I}_M \otimes \mathbf{A}_\varepsilon \mathbf{Z})\mathrm{vec}(\mathbf{H}^T) + \mathbf{n} \ . \tag{3.9}$$

2. Symbol Time-Varying Fading

The model introduced in this section will be used for all simulations in this thesis, unless otherwise stated. Since there are L_0Q observation samples, in the symbol time-varying case the matrix $\mathbf{H}$ needs to be redefined, extending its number of rows by L_0Q. Considering between each pair of antennas different values of the fading at different instants of time, correlated according to (3.3), the channel transfer function can be described in this case by the $N \times ML_0Q$ size matrix

$$\mathbf{H}^T = \begin{bmatrix} h_{11}(0) & \cdots & h_{11}((L_0Q-1)T_s) & h_{12}(0) & \cdots & h_{1M}((L_0Q-1)T) \\ h_{21}(0) & \ddots & \vdots & h_{22}(0) & \ddots & \vdots \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ h_{N1}(0) & \cdots & h_{N1}((L_0Q-1)T_s) & h_{N2}(0) & \cdots & h_{NM}((L_0Q-1)T) \end{bmatrix} . \tag{3.10}$$

By creating a repetition of $\mathbf{A}_\varepsilon$

$$\mathbf{T}_\varepsilon = \mathbf{1}_{M\times 1} \otimes \mathbf{A}_\varepsilon \ ,$$

where $\mathbf{1}_{M\times 1}$ represents a matrix of size $M \times 1$ whose elements are all equal to 1, a new form for equation (3.8) is obtained

$$\mathbf{r} = \mathrm{diag}\{\xi \mathbf{T}_\varepsilon \mathbf{Z}\mathbf{H}^T\} + \mathbf{n} \ ,$$

where $\mathrm{diag}\{\cdot\}$ represents the operator that retains only the diagonal elements of a matrix. Introducing a new operator named $\mathrm{sample}_\varsigma\{\cdot\}$ that keeps only the lines of index $k\varsigma^{th}+1$ of a matrix, $\mathbf{r}$ can also be expressed as

$$\mathbf{r} = \mathrm{sample}_{(ML_0Q+1)} \left\{ \mathrm{vec}(\xi \mathbf{T}_\varepsilon \mathbf{Z}\mathbf{H}^T) \right\} + \mathbf{n} \ .$$

Using twice the fact that $\text{vec}(\mathbf{AYB}) = (\mathbf{B}^T \otimes \mathbf{A})\text{vec}(\mathbf{Y})$ on $\text{vec}(\mathbf{T}_\varepsilon \mathbf{Z}\mathbf{H}^T)$, the above equation becomes successively

$$\mathbf{r} = \xi \text{sample}_{(ML_0Q+1)} \left\{ (\mathbf{I}_{ML_0Q} \otimes \mathbf{T}_\varepsilon)\text{vec}(\mathbf{Z}\mathbf{H}^T) \right\} + \mathbf{n} \ , \tag{3.11}$$

$$\mathbf{r} = \xi \text{sample}_{(ML_0Q+1)} \left\{ (\mathbf{I}_{ML_0Q} \otimes \mathbf{T}_\varepsilon)(\mathbf{I}_{ML_0Q} \otimes \mathbf{Z})\text{vec}(\mathbf{H}^T) \right\} + \mathbf{n} \ .$$

Using the fact that $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC}) \otimes (\mathbf{BD})$, the previous equation becomes

$$\mathbf{r} = \xi \text{sample}_{(ML_0Q+1)} \left\{ (\mathbf{I}_{ML_0Q} \otimes \mathbf{T}_\varepsilon \mathbf{Z})\text{vec}(\mathbf{H}^T) \right\} + \mathbf{n} \ .$$

Property

For any matrices $\mathbf{A}$ and $\mathbf{B}$ of size $P \times MN$ and $M \times N$ respectively, the following property of the $\text{sample}_\varsigma\{\cdot\}$ operator is true

$$\text{sample}_\varsigma\{\mathbf{A}\text{vec}(\mathbf{B})\} = \text{sample}_\varsigma\{\mathbf{A}\}\text{vec}(\mathbf{B}) \ .$$

Using this property, $\mathbf{r}$ can be expressed as

$$\mathbf{r} = \xi \text{sample}_{(ML_0Q+1)} \left\{ (\mathbf{I}_{ML_0Q} \otimes \mathbf{T}_\varepsilon \mathbf{Z}) \right\} \text{vec}(\mathbf{H}^T) + \mathbf{n} \ .$$

Because the "sampling" rate of the $\text{sample}_\varsigma\{\cdot\}$ operator is $ML_0Q + 1$ and the size of the identity matrix is $(ML_0Q) \times (ML_0Q)$, the following model for the DA time-varying channel case is obtained

$$\mathbf{r} = \xi \mathbf{\Psi}_\varepsilon \text{vec}(\mathbf{H}^T) + \mathbf{n} \ , \tag{3.12}$$

where $\mathbf{\Psi}_\varepsilon$ is a matrix having the following block diagonal structure

$$\mathbf{\Psi}_\varepsilon = \begin{bmatrix} [\mathbf{T}_\varepsilon \mathbf{Z}]_{1,:} & 0 & \cdots & 0 \\ 0 & [\mathbf{T}_\varepsilon \mathbf{Z}]_{2,:} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & [\mathbf{T}_\varepsilon \mathbf{Z}]_{ML_0Q,:} \end{bmatrix} . \tag{3.13}$$

A similar form can be obtained for the NDA case, following the same steps between equations (3.11) and (3.12), with the only difference that the matrix $\mathbf{Z}$ is left inside the vec operator in equation (3.11). Thus, the following model for the NDA time-varying case is obtained

$$\mathbf{r} = \xi \mathbf{\Psi}_\varepsilon \text{vec}(\mathbf{Z}\mathbf{H}^T) + \mathbf{n} \ , \tag{3.14}$$

where

$$\mathbf{\Psi}_\varepsilon = \begin{bmatrix} [\mathbf{T}_\varepsilon]_{1,:} & 0 & \cdots & 0 \\ 0 & [\mathbf{T}_\varepsilon]_{2,:} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & [\mathbf{T}_\varepsilon]_{ML_0Q,:} \end{bmatrix} . \tag{3.15}$$

Expressions (3.13) and (3.15) can be further simplified noticing the structure of the matrix $\mathbf{T}_\varepsilon$. Because $\mathbf{T}_\varepsilon = \mathbf{1}_{M\times 1} \otimes \mathbf{A}_\varepsilon$, the final forms of (3.13) and (3.15) are as follow

- for DA case

$$\mathbf{\Psi}_\varepsilon = \mathbf{I}_M \otimes \begin{bmatrix} [\mathbf{A}_\varepsilon \mathbf{Z}]_{1,:} & 0 & \cdots & 0 \\ 0 & [\mathbf{A}_\varepsilon \mathbf{Z}]_{2,:} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & [\mathbf{A}_\varepsilon \mathbf{Z}]_{L_0Q,:} \end{bmatrix} . \tag{3.16}$$

- for NDA case

$$\mathbf{\Psi}_{\varepsilon} = \mathbf{I}_M \otimes \begin{bmatrix} [\mathbf{A}_{\varepsilon}]_{1,:} & 0 & \cdots & 0 \\ 0 & [\mathbf{A}_{\varepsilon}]_{2,:} & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & [\mathbf{A}_{\varepsilon}]_{L_0 Q,:} \end{bmatrix} . \tag{3.17}$$

Equations (3.12) and (3.14) can be written in the following general form

$$\mathbf{r} = \xi \mathbf{\Psi}_{\varepsilon} \mathbf{h} + \mathbf{n} \ , \tag{3.18}$$

where $\mathbf{\Psi}_{\varepsilon}$ has been defined in (3.16) and (3.17) and

$$\mathbf{h} = \begin{cases} \text{vec}(\mathbf{H}^T) & \text{in the DA case} \\ \text{vec}(\mathbf{Z}\mathbf{H}^T) & \text{in the NDA case} \end{cases} . \tag{3.19}$$

CHAPTER IV

COMPUTATION OF BOUNDS

A. Block Fading Channel Model

1. Modified Cramer-Rao Bound

a. Data-Aided Scenario

In the DA block fading case, the MCRB has been found to be [34]

$$\text{MCRB} = \frac{NQ}{2M} \cdot \left(\frac{E_s}{N_0}\right)^{-1} \cdot \frac{1}{\text{tr}(\mathbf{Z}^H \mathbf{D}_\varepsilon^H \mathbf{D}_\varepsilon \mathbf{Z})} \,, \tag{4.1}$$

where $\text{tr}(\cdot)$ denotes the trace of a matrix, $\mathbf{D}_\varepsilon$ represents the derivative of the matrix $\mathbf{A}_\varepsilon$ with respect to the time delay ε, $\mathbf{D}_\varepsilon = d\mathbf{A}_\varepsilon/d\varepsilon$ and all the other parameters have been defined in Chapter III.C.

An equivalent form for the extended model can be derived by observing the similarity between the equations (3.12) and (3.9). Thus

$$\text{MCRB} = \frac{NQ}{2M} \cdot \left(\frac{E_s}{N_0}\right)^{-1} \cdot \frac{1}{\text{tr}(\boldsymbol{\Upsilon}_\varepsilon^H \boldsymbol{\Upsilon}_\varepsilon)} \,,$$

where, similarly to (3.16)

$$\boldsymbol{\Upsilon}_\varepsilon = \mathbf{I}_M \otimes \begin{bmatrix} [\mathbf{D}_\varepsilon \mathbf{Z}]_{1,:} & 0 & \cdots & 0 \\ 0 & [\mathbf{D}_\varepsilon \mathbf{Z}]_{2,:} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & [\mathbf{D}_\varepsilon \mathbf{Z}]_{L_0 Q,:} \end{bmatrix} . \tag{4.2}$$

Due to the $\text{tr}(\cdot)$ operator, both forms of MCRB produce the same result.

The data is generated as in [35]. The first step is the construction of a Chu

sequence of length L_0

$$\mathbf{s} = \left[\begin{array}{cccc} s(0) & s(1) & ... & s(L_0-1) \end{array} \right].$$

Using the sequence $\mathbf{s}$, a second sequence is constructed as

$$\mathbf{s}' = \left[\begin{array}{cccccccc} s(0) & s(1) & ... & s(L_0-1) & s(0) & s(1) & ... & s(2NL_g-1) \end{array} \right].$$

Finally, the training sequences are given by

$$\mathbf{d}_i = \left[\begin{array}{ccc} s'((2i-1)L_g) & ... & s'((2i-1)L_g+L_0-1) \end{array} \right].$$

b. Non-Data-Aided Scenario

In the NDA block fading case, for a particular half-rate orthogonal space-time block code, described by the matrix

$$\left[\begin{array}{cccc} d_1 & d_2 & d_3 & d_4 \\ -d_2 & d_1 & -d_4 & d_3 \\ -d_3 & d_4 & d_1 & -d_2 \\ -d_4 & -d_3 & d_2 & d_1 \\ d_1^* & d_2^* & d_3^* & d_4^* \\ -d_2^* & d_1^* & -d_4^* & d_3^* \\ -d_3^* & d_4^* & d_1^* & -d_2^* \\ -d_4^* & -d_3^* & d_2^* & d_1^* \end{array} \right], \tag{4.3}$$

the MCRB has been found to be [34]

$$\text{MCRB} = \frac{NQ}{2M} \cdot \left(\frac{E_s}{N_0} \right)^{-1} \cdot \frac{1}{\text{tr}(\mathbf{D}_\varepsilon^H \mathbf{D}_\varepsilon)} \ .$$

As in the DA case, an equivalent form for the extended model can be derived by observing the similarity between the equations (3.14) and (3.8). Thus

$$\text{MCRB} = \frac{NQ}{2M} \cdot \left(\frac{E_s}{N_0}\right)^{-1} \cdot \frac{1}{\text{tr}(\boldsymbol{\Upsilon}_{\varepsilon}^H \boldsymbol{\Upsilon}_{\varepsilon})} \ ,$$

where, similarly to (3.17)

$$\boldsymbol{\Upsilon}_{\varepsilon} = \mathbf{I}_M \otimes \begin{bmatrix} [\mathbf{D}_{\varepsilon}]_{1,:} & 0 & \cdots & 0 \\ 0 & [\mathbf{D}_{\varepsilon}]_{2,:} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & [\mathbf{D}_{\varepsilon}]_{L_0 Q,:} \end{bmatrix} . \tag{4.4}$$

2. Modified Barankin Bound

For its relative ease of computation, a modified version of the BB is introduced in this section. This version is obtained by replacing the pdf-s by their fading-conditioned versions using the same procedure that distinguishes the MCRB from the true CRB. Therefore, the name of this bound will be the Modified Barankin Bound (MBB) throughout this thesis.

In its Chapman-Robbins form (see (2.3)), the MBB is

$$\text{MBB} = \sup_{\eta} \frac{\eta^2}{\int \left[\int \frac{p(\mathbf{r}|\mathbf{h};\varepsilon+\eta)^2}{p(\mathbf{r}|\mathbf{h};\varepsilon)} p(\mathbf{h}) d\mathbf{h} \right] d\mathbf{r} - 1} \geq \text{MCRB} \ . \tag{4.5}$$

Using the Cauchy-Schwarz inequality it can be proven (see Appendix A) that this bound is always looser than the true BB. Also, this bound is always larger than or equal to the MCRB, and thus it is a valid bound for timing estimation. The derivation of the MBB is presented in Appendix A. The general formula (A.2) can be particularized by taking into account two small observations. First, a change of

variables is needed

$$\sigma^2 \to \frac{N_0 Q}{T}$$

$$\xi^2 \to \frac{E_s}{NT}$$

and thus, $2\xi^2/\sigma^2 \to 2E_s/N_0 QN$.

The second observation regards the particular structure taken by the matrix $\mathbf{\Psi}_\varepsilon$ in (A.2). In the NDA case, $\mathbf{\Psi}_\varepsilon = \mathbf{I}_M \otimes \mathbf{A}_\varepsilon$ (see (3.8)) and in the DA case $\mathbf{\Psi}_\varepsilon = \mathbf{I}_M \otimes \mathbf{A}_\varepsilon \mathbf{Z}$ (see (3.9)). Therefore, it is possible to generally express $\mathbf{\Psi}_\varepsilon$ as $\mathbf{\Psi}_\varepsilon = \mathbf{I}_M \otimes \mathbf{\Xi}_\varepsilon$, and the products $\mathbf{\Psi}_\alpha^H \mathbf{\Psi}_\beta$ take a simpler form

$$\mathrm{tr}(\mathbf{\Psi}_\alpha^H \mathbf{\Psi}_\beta) = \mathrm{tr}((\mathbf{I}_M \otimes \mathbf{\Xi}_\alpha^H)(\mathbf{I}_M \otimes \mathbf{\Xi}_\beta)) = \mathrm{tr}(\mathbf{I}_M \otimes (\mathbf{\Xi}_\alpha^H \mathbf{\Xi}_\beta) = M\mathrm{tr}(\mathbf{\Xi}_\alpha^H \mathbf{\Xi}_\beta) \ .$$

In the block fading DA case the MBB takes the form

$$\mathrm{MBB} = \sup_\eta \frac{\eta^2}{\exp\left\{\frac{E_s}{N_0} \cdot \frac{2M}{QN} \cdot \mathrm{tr}(\mathbf{Z}^H (\mathbf{A}_{\varepsilon+\eta} - \mathbf{A}_\varepsilon)^H (\mathbf{A}_{\varepsilon+\eta} - \mathbf{A}_\varepsilon)\mathbf{Z})\right\} - 1} \ .$$

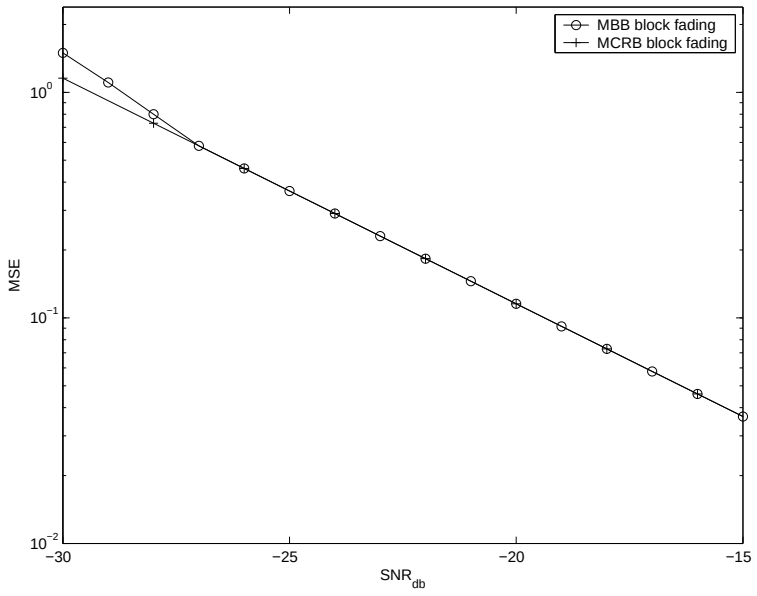


Fig. 2. MBB and MCRB for block fading

Fig. 2 compares the MBB and the MCRB in the DA case. The threshold effect is visible, but at very low SNR (-27 dB). The same value is obtained in the NDA case. For the MCRB-optimized data sequences that will be introduced in Chapter V.B, the threshold appears at -29 dB. The conclusion is that, although it is a valid bound, the MBB is too close to the MCRB to be of any practical use as a tighter bound. The threshold can give some information regarding the required length of the training data, but more accurate results can be obtained using the true BB.

For the block fading NDA case the MBB takes the form

$$\mathrm{MBB} = \sup_{\eta} \frac{\eta^2}{\exp\left\{\frac{E_s}{N_0} \cdot \frac{2M}{QN} \cdot \mathrm{tr}((\mathbf{A}_{\varepsilon+\eta} - \mathbf{A}_{\varepsilon})^H (\mathbf{A}_{\varepsilon+\eta} - \mathbf{A}_{\varepsilon}))\right\} - 1} \ .$$

3. Cramer-Rao Bound

In the block fading case the Fisher Information Matrix (FIM) has the following form

$$\left[\begin{array}{c|c} \mathbf{A} & \mathbf{B}^T \\ \hline \mathbf{B} & \mathbf{C} \end{array}\right] \ ,$$

where

$$\mathbf{A} = \begin{bmatrix} J_{\sigma_\alpha^2 \sigma_\alpha^2} & J_{\sigma_\alpha^2 N_0} \\ J_{\sigma_\alpha^2 N_0} & J_{N_0 N_0} \end{bmatrix} \ ,$$

$$\mathbf{B} = \begin{bmatrix} J_{\sigma_\alpha^2 \varepsilon} & J_{N_0 \varepsilon} \end{bmatrix} \ ,$$

and

$$\mathbf{C} = \begin{bmatrix} J_{\varepsilon\varepsilon} \end{bmatrix} \ ,$$

where (see (B.13))

$$J_{\alpha\beta} = E\left[-\frac{\partial^2 \mathbf{\Lambda}}{\partial\alpha\partial\beta}\right] = \mathrm{tr}\left(\tilde{\mathbf{R}}^{-1} \frac{\partial \tilde{\mathbf{R}}}{\partial\alpha} \tilde{\mathbf{R}}^{-1} \frac{\partial \tilde{\mathbf{R}}}{\partial\beta}\right) \ .$$

After computing the elements of the matrices $\mathbf{A}, \mathbf{B}$ and $\mathbf{C}$, the CRB is found using the formula

$$\text{CRB} = \frac{1}{\mathbf{C} - \mathbf{B}\mathbf{A}^{-1}\mathbf{B}^T} \geq \frac{1}{\mathbf{C}} = J_{\varepsilon\varepsilon}^{-1} \tag{4.6}$$

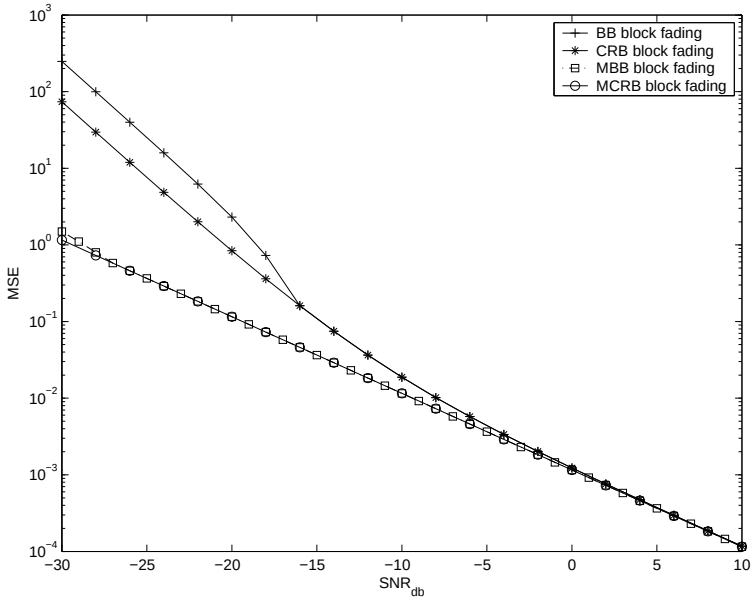
The details of the computation are given in Appendix B. In the above inequality, the left hand side represents the CRB computed in the presence of nuisance parameters, and the right hand side denotes the CRB computed when the timing delay is the only unknown. The significance of this inequality is that the presence of additional unknown parameters increases the bound. Introducing a cost function, Chapter V.A presents the relative difference between the CRB computed in the two scenarios and its dependence on f_d.

4. Barankin Bound

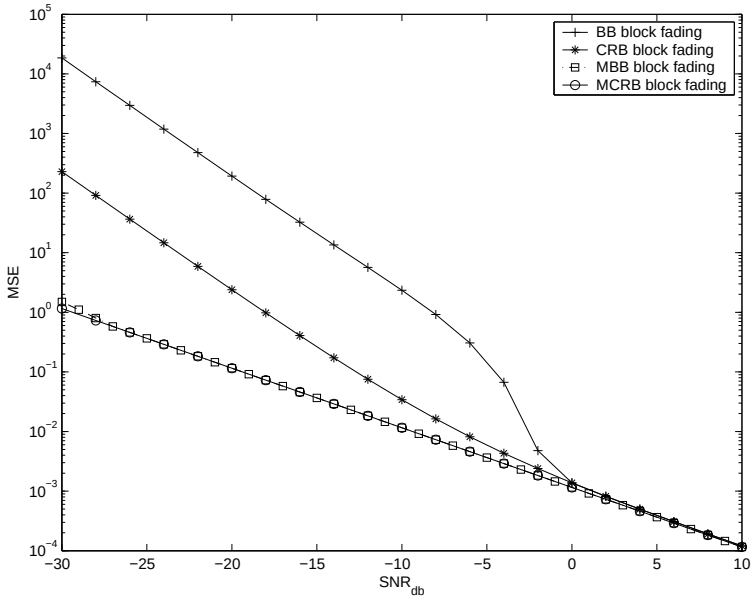
For the eigendecomposition-based model introduced in Chapter III.B the covariance matrix $\tilde{\mathbf{R}}$ of the received signal has been determined in (B.2) and the associated pdf $p_{\bar{\mathbf{r}}}(\bar{\mathbf{r}})$ in (B.4). The matrix $\tilde{\mathbf{R}}$ depends on ε through the combination of terms involving $\bar{\mathbf{d}}_1$ and $\bar{\mathbf{d}}_2$. To highlight this dependence, $\tilde{\mathbf{R}}_\varepsilon$ will denote the matrix $\tilde{\mathbf{R}}$ corresponding to the time delay ε and $\tilde{\mathbf{R}}_{\varepsilon+\eta}$ the matrix $\tilde{\mathbf{R}}$ corresponding to the time delay $\varepsilon + \eta$.

Similarly, for the main model, the general form of the covariance matrix $\tilde{\mathbf{R}}$ of the received signal has been presented in (B.9). The particular forms taken in the DA and the NDA case have been presented in (B.10) and (B.11) and the associated pdf $p_{\mathbf{r}}(\mathbf{r})$ in (B.12). The matrix $\tilde{\mathbf{R}}$ depends on ε through $\mathbf{A}_\varepsilon$. In order to stress out this dependence the notations $\tilde{\mathbf{R}}_\varepsilon$ and $\tilde{\mathbf{R}}_{\varepsilon+\eta}$ will be used. The details of computing the BB are presented in Appendix C.

Fig. 3 presents all the bounds introduced for the block fading case.



a) DA case



b) NDA case

Fig. 3. Bounds for block fading

It is important to notice that the MCRB has a slope of -10 dB/decade because this slope is approximately the slope of the true CRB at medium SNR values even in the symbol-level time-varying fading case, as described in the next section.

B. Cramer-Rao Bound for TV Fading

The Fisher Information Matrix (FIM) has the following form

$$\left[\begin{array}{c|c} \mathbf{A} & \mathbf{B}^T \\ \hline \mathbf{B} & \mathbf{C} \end{array}\right] ,$$

where

$$\mathbf{A} = \begin{bmatrix} J_{\sigma_\alpha^2\sigma_\alpha^2} & J_{\sigma_\alpha^2 f_d} & J_{\sigma_\alpha^2 N_0} \\ J_{\sigma_\alpha^2 f_d} & J_{f_d f_d} & J_{f_d N_0} \\ J_{\sigma_\alpha^2 N_0} & J_{f_d N_0} & J_{N_0 N_0} \end{bmatrix} , \tag{4.7}$$

$$\mathbf{B} = \begin{bmatrix} J_{\sigma_\alpha^2 \varepsilon} & J_{f_d \varepsilon} & J_{N_0 \varepsilon} \end{bmatrix} , \tag{4.8}$$

and

$$\mathbf{C} = \begin{bmatrix} J_{\varepsilon\varepsilon} \end{bmatrix} , \tag{4.9}$$

where (see (D.6))

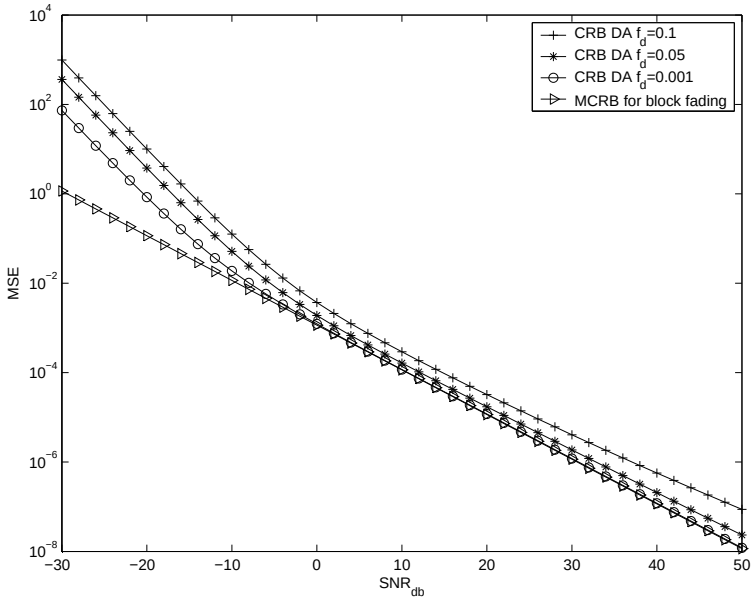
$$J_{\alpha\beta} = E\left[-\frac{\partial^2 \mathbf{\Lambda}}{\partial\alpha\partial\beta}\right] = \mathrm{tr}\left(\tilde{\mathbf{R}}^{-1}\frac{\partial\tilde{\mathbf{R}}}{\partial\alpha}\tilde{\mathbf{R}}^{-1}\frac{\partial\tilde{\mathbf{R}}}{\partial\beta}\right) .$$

After computing the elements of the matrices $\mathbf{A}, \mathbf{B}$ and $\mathbf{C}$, the CRB is found using the formula

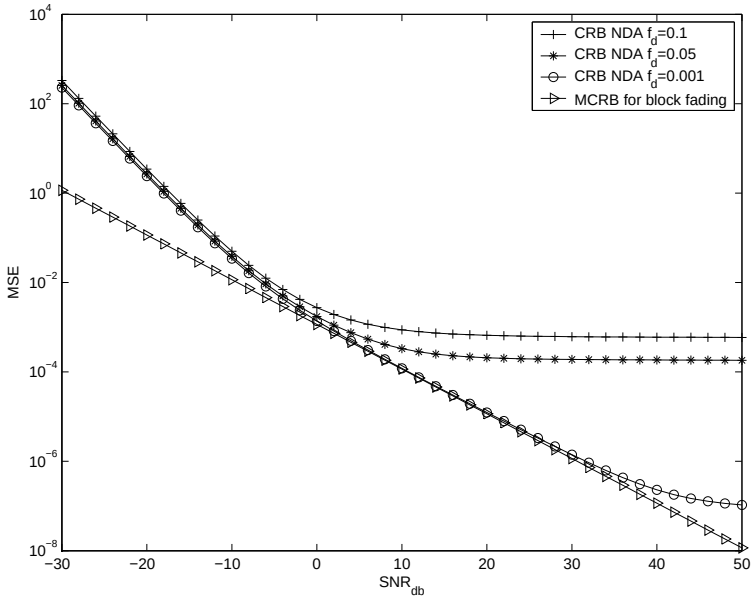
$$\mathrm{CRB} = \frac{1}{\mathbf{C} - \mathbf{B}\mathbf{A}^{-1}\mathbf{B}^T} \geq \frac{1}{\mathbf{C}} = J_{\varepsilon\varepsilon}^{-1} . \tag{4.10}$$

The same comments made for (4.6) apply for the above inequality.

Fig. 4 presents the variations with the Doppler shift of the true CRB, in both the DA and NDA scenarios. For reference, the MCRB in the block fading case



a) DA case



b) NDA case

Fig. 4. CRB for TV fading

is also included in this figure. For the DA case, the training sequence based on Chu sequences, introduced in Section A, has been chosen in a suboptimal way, in the sense that the MCRB in the DA case almost coincides with the MCRB for the NDA case. This suboptimal training sequence will be used in order to highlight the differences between the DA and the NDA case. However, Chapter V.B will show how to optimize the training sequence in order to minimize the MCRB. The advantages of DA estimation will become even more obvious at that time but for the moment all simulations are restricted to the suboptimal training sequences. Because the larger the f_d the larger the bound, as seen in Fig. 4, the best estimators perform worse when the fading is fast varying and this rule remains valid for the BB, as it will be presented in Section C. Also, Fig. 4 shows that the true CRB generally consists of three parts. The first part is the region of low SNR, characterized by a slope of -20 dB/decade. In this region the performance of any unbiased estimator decreases drastically with the noise power. At very high SNR the curves flatten and a floor effect appears. This effect appears in Fig. 4 *b*. The floor effect also appears in the DA case, but at higher SNR. The intermediate region from -20 dB/decade slope to 0 dB/decade slope is the region of medium SNR. Since most systems operate at such values of SNR, in this region the MCRB, which has a slope of -10 dB/decade, is generally a good approximation for the true CRB. For the considered model, the intermediate region extends approximately from 0 dB to 30 dB for block fading. It is visible in Fig. 4 that at very slow fading ($f_d = 0.001$) the CRB for both DA and NDA case almost coincide in the intermediate SNR region with the MCRB obtained in the block fading case.

A comparison of the DA and NDA bounds evidences that, except for the intermediate region, at slow fading the DA CRB (Fig. 4 *a*) outperforms the NDA CRB (Fig. 4 *b*). When the fading becomes very fast varying ($f_d = 0.1$), at low SNR the DA algorithms perform worse but, due to the floor effect, at medium and high SNR they

perform much better than their NDA equivalents. In Fig. 4 the NDA CRB displays less sensitivity to the Doppler shift at very low SNR. Therefore, the conclusion is that in this region there should exist robust estimators.

C. Barankin Bound for TV Fading

For the eigendecomposition-based model introduced in Chapter III.B the covariance matrix $\tilde{\mathbf{R}}$ of the received signal has been determined in (D.2) and the associated pdf $p_{\bar{\mathbf{r}}}(\bar{\mathbf{r}})$ in (B.4). The matrix $\tilde{\mathbf{R}}$ depends on ε through $\bar{\mathbf{d}}_1$ and $\bar{\mathbf{d}}_2$. To highlight this dependence, $\tilde{\mathbf{R}}_\varepsilon$ will denote the matrix $\tilde{\mathbf{R}}$ corresponding to the time delay ε and $\tilde{\mathbf{R}}_{\varepsilon+\eta}$ the matrix $\tilde{\mathbf{R}}$ corresponding to the time delay $\varepsilon+\eta$.

Similarly, for the main model, the general form of the covariance matrix $\tilde{\mathbf{R}}$ of the received signal has been presented in (D.4) and the associated pdf $p_{\mathbf{r}}(\mathbf{r})$ in (D.5). The matrix $\tilde{\mathbf{R}}$ depends on ε through $\mathbf{\Psi}_\varepsilon$, which is a function of $\mathbf{A}_\varepsilon$ through the particular forms (3.16) for the DA case and (3.17) for the NDA case. In order to highlight this dependence the notations $\tilde{\mathbf{R}}_\varepsilon$ and $\tilde{\mathbf{R}}_{\varepsilon+\eta}$ will be used. The details of computing the BB are presented in Appendix C.

For the main model, the BB has been found to have the form

$$\mathrm{BB} = \sup_{\eta} \frac{\eta^2 |\, \tilde{\mathbf{R}}_{\varepsilon+\eta}\tilde{\mathbf{R}}_{\varepsilon}^{-1} \,||\, 2\mathbf{I} - \tilde{\mathbf{R}}_{\varepsilon+\eta}\tilde{\mathbf{R}}_{\varepsilon}^{-1} \,|}{1 - |\, \tilde{\mathbf{R}}_{\varepsilon+\eta}\tilde{\mathbf{R}}_{\varepsilon}^{-1} \,||\, 2\mathbf{I} - \tilde{\mathbf{R}}_{\varepsilon+\eta}\tilde{\mathbf{R}}_{\varepsilon}^{-1} \,|} \ . \tag{4.11}$$

Expressions (C.5) and (C.8) are basically the same formula, since in (C.8) each covariance matrix can be expressed as the Kronecker product between $\mathbf{I}_M$ and a more fundamental matrix. This is valid for DA and NDA case, for block and for TV fading. Using the standard determinant properties $|\, \mathbf{I}_M \otimes \mathbf{\Delta} \,| = |\, \mathbf{\Delta} \,|^M$, and in the particular case $M = 2$, (C.8) takes a form similar to (C.5).

Fig. 5 presents the variations with the Doppler shift of the BB in the DA case.

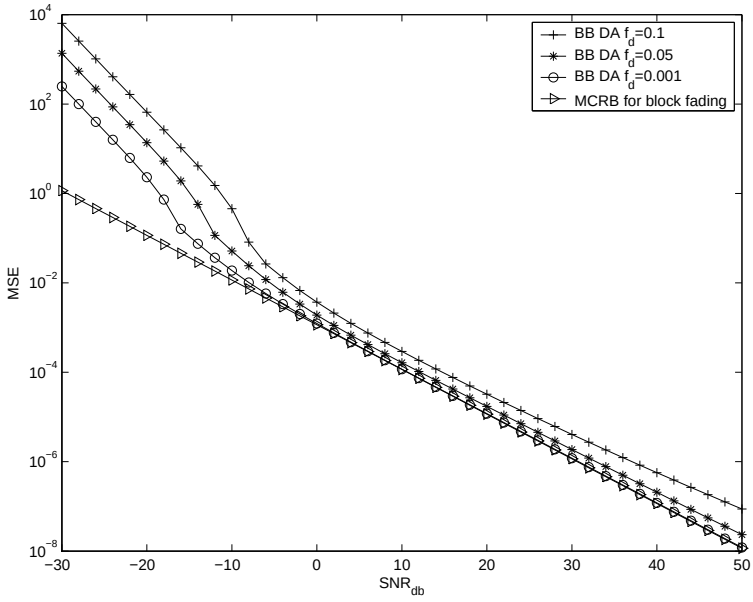
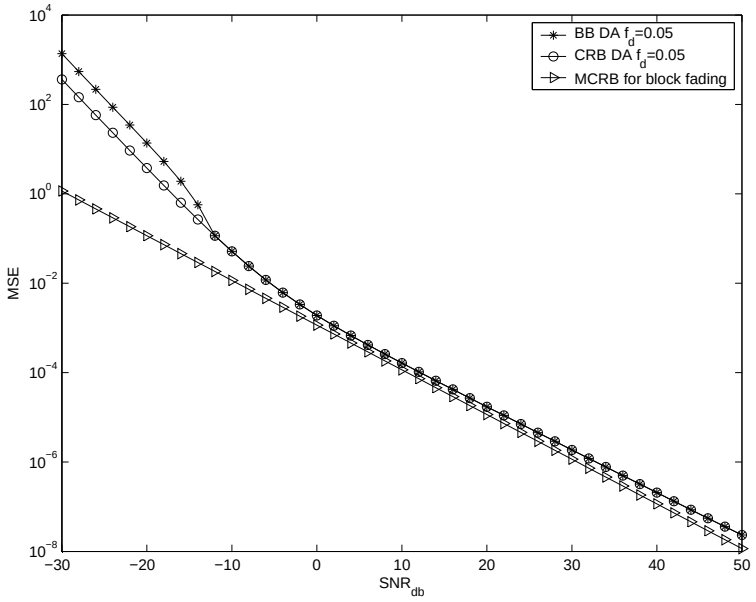


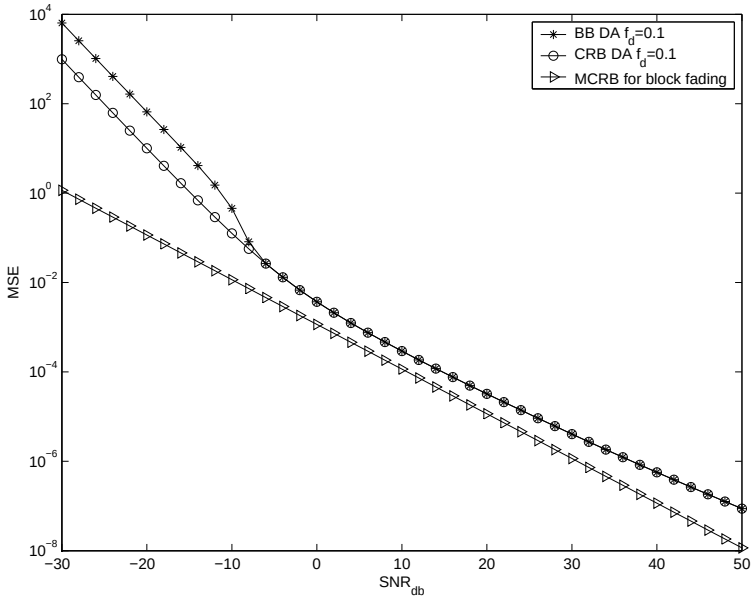
Fig. 5. BB for TV fading

It can be noticed that the remark about the variation of the CRB with the Doppler shift remains valid for the BB and the bound increases with f_d. For clarity, the corresponding CRBs have not been plotted, but the threshold points are obvious. The SNR associated with the threshold is an increasing function of f_d.

Fig. 6 *a* and *b* compare the BB and the CRB obtained for a certain Doppler shift. In these comparative plots, the threshold behavior is clearly depicted. This Doppler-dependent threshold will be analyzed further in Chapter V.E in order to get an in-depth view over the required length of training data.



a) moderate fading



b) very fast fading

Fig. 6. BB variation with the Doppler shift

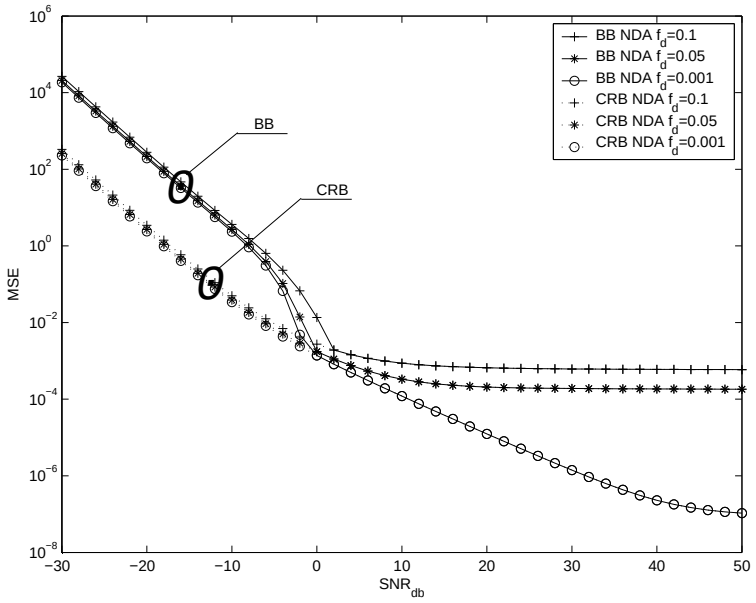


Fig. 7. BB and CRB in the NDA case

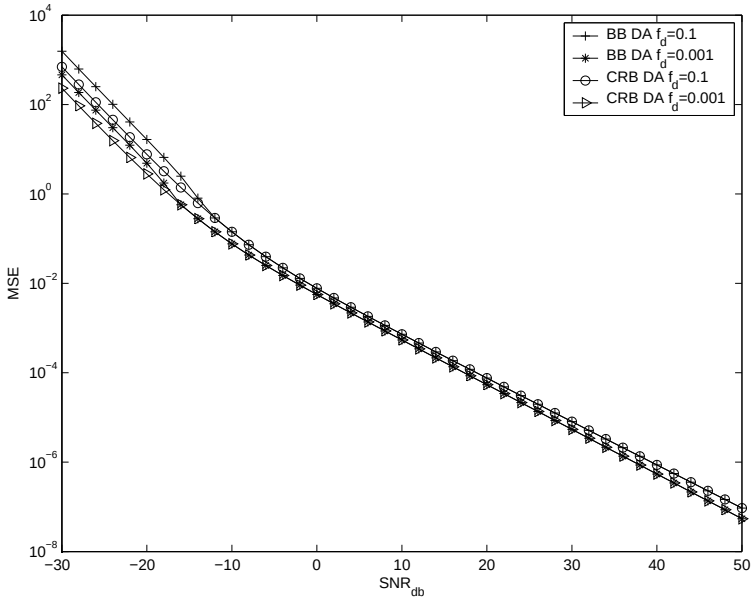


Fig. 8. BB and CRB for the eigendecomposition-based model

For comparison purposes, Fig. 7 presents the BB and the CRB for the NDA case, and Fig. 8 presents the BB and the CRB associated with the eigendecomposition-based model introduced in Chapter III.B. From these figures it can be inferred that the SNR threshold and the bounds are increasing functions of f_d. Not only the CRB, but also the BB in the NDA case is very little dependent on f_d at low SNR. Also, by comparing Fig. 5 and Fig. 7, it appears that the threshold phenomenon occurs at higher SNR in the NDA case, meaning that, in general, a higher length of data is needed in the NDA case. Also, although by comparing the CRB in the DA and NDA case it can be concluded that at very low SNR the NDA algorithms are superior, the BB proves that even in that SNR region the DA estimators are preferable.

CHAPTER V

ANALYSIS OF BOUNDS AND ESTIMATORS

A. Cost Function

In the timing estimation process it is useful to define a cost function that determines the price paid for not knowing the nuisance parameters $(\sigma_\alpha^2, f_d, N_0)$. The absolute cost is the difference between the CRBs computed assuming unknown and known nuisance parameters, respectively. Referring to (4.10), the absolute cost is found to

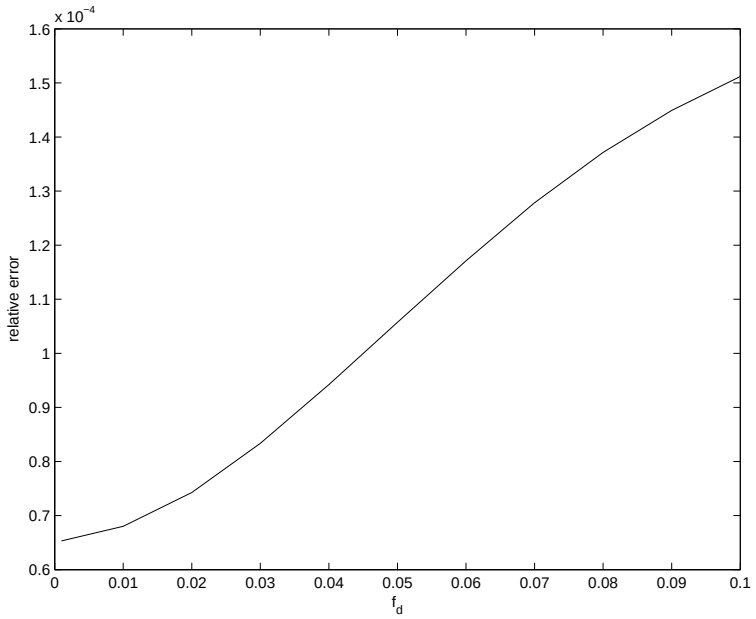


Fig. 9. Relative cost of not knowing the nuisance parameters

be the difference between the left hand side and the right hand side terms. The relative cost is defined as the ratio between this difference and the CRB determined

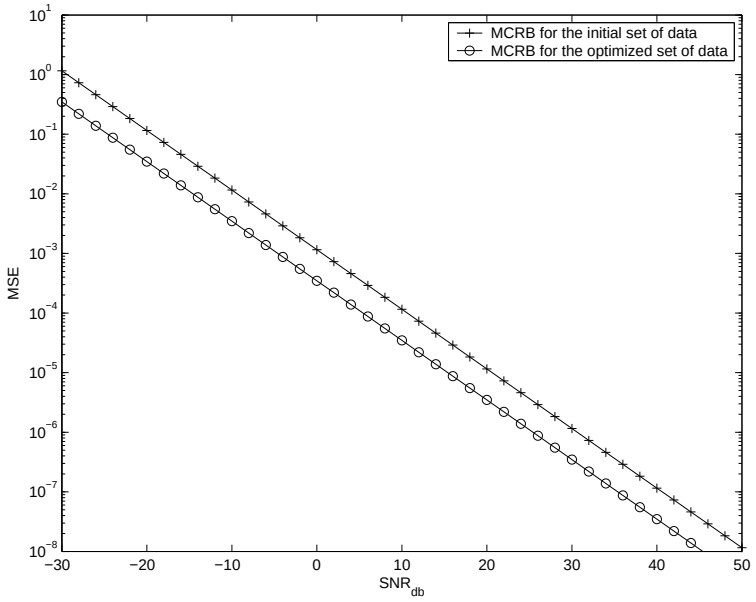
assuming unknown nuisance parameters and is mathematically expressed as

$$\text{Cost} = \frac{\mathbf{B}\mathbf{A}^{-1}\mathbf{B}^{T}}{\mathbf{C}} \ .$$

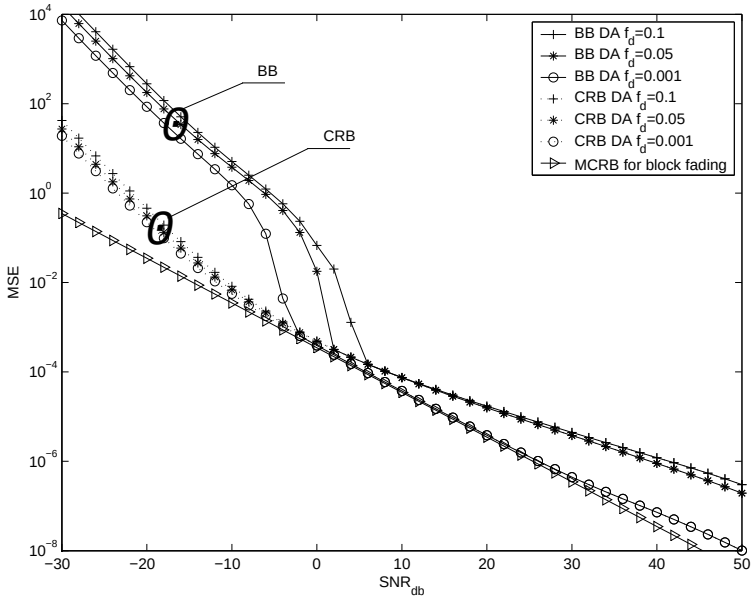
Fig. 9 presents the dependence of the average relative cost, defined as the relative cost averaged over all values of SNR in the range of interest $(-30, 50)$ dB, with respect to f_d, in the NDA case. The relative cost is quite small, being of the order of 10^{-4} and 10^{-5} in the NDA and DA case, respectively. Therefore, the knowledge upon the nuisance parameters $(\sigma_\alpha^2, f_d, N_0)$ can not improve substantially the performance of the timing delay estimation. From Fig. 9 it turns out that the cost is an increasing function with respect to f_d.

B. Optimization over the Training Data

It has been found that the data matrix $\mathbf{Z}$ that minimizes the MCRB is formed by the N scaled eigenvectors corresponding to the N largest eigenvalues of the matrix $\mathbf{D}_\varepsilon^H \mathbf{D}_\varepsilon$, where $\mathbf{D}_\varepsilon$ has been described in (4.1). Fig. 10 *a* shows the improvement in the MCRB obtained by using this optimized data matrix. The gain turns out to be approximately 5 dB. Fig. 10 *b* shows the BB and the CRB for the MCRB-optimized training sequence. At low SNR, the CRB and the BB are less dependent on f_d than their equivalents from the non-optimized case. The BB threshold effect occurs at higher SNR in the optimized case. From these regards, the CRB and the BB for the MCRB-optimized data behave similarly with their equivalents from the NDA case.

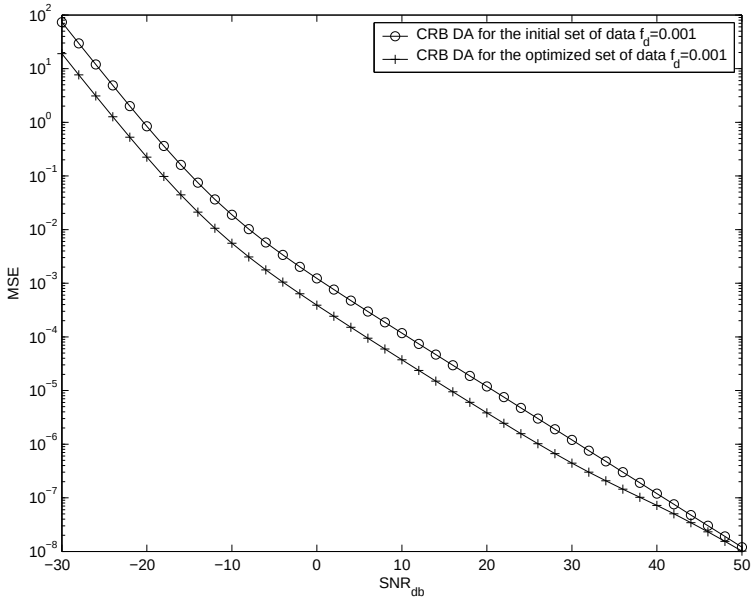


a) MCRB gain

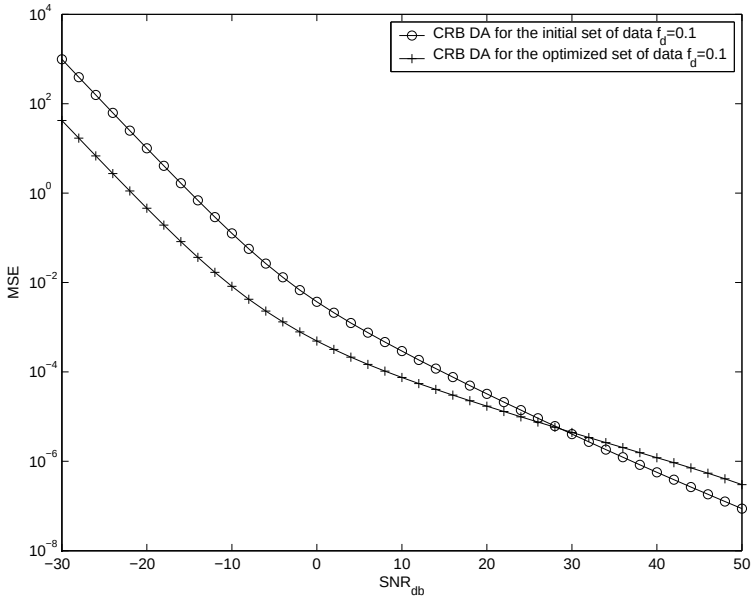


b) BB and CRB

Fig. 10. Bounds for MCRB-optimized data



a) very slow fading



b) very fast fading

Fig. 11. Comparison of CRB for different sets of data

Fig. 11 compares the CRB for the initial set of data with the one for the optimized data set, assuming different Doppler shifts. It turns out that, although for slow fading the MCRB-optimized CRB is below the initial CRB for the whole region of interest, for very fast fading it is below only for SNR < 28 dB. A comparative analysis of the BB between Fig. 10 *b* and Fig. 5 reveals that for low SNR the BB is lower for the initial set of data. Therefore, it can be concluded that the optimization described is useful for any f_d only for the average SNR region, which extends approximately from 0 dB to 30 dB.

C. Asymptotic Cramer-Rao Bounds (ACRB)

This section presents the asymptotic CRB (ACRB) for low and high SNR, respectively, assuming both the DA and NDA scenarios. The details of computation are presented in Appendix E. The asymptotic bounds for low SNR, as well as the ones for high SNR, are characterized by a Doppler dependent multiplicative constant. The low-SNR ACRB has a slope of 20 dB/decade (varies inversely proportional with the second power of the SNR). With increasing SNR each CRB reduces up to its asymptotic value for high SNR. This high-SNR asymptotic CRB has a slope of 0 dB/decade (does not depend on the SNR) and therefore is characterized by a constant. The floor is due to the multiplicative noise, because at high SNR the additive noise can be neglected.

For different Doppler shifts, Table I presents the constants associated with the low-SNR ACRB (LACRB) and the high-SNR ACRB (HACRB) in three scenarios: data-aided considering the Chu training sequence (DA1), data-aided considering the MCRB-optimized sequences (DA2) and non-data-aided (NDA). The LACRB and the HACRB turn out to be increasing functions with respect to f_d in all three scenarios.

This fact was expected, because the CRB and the BB are also increasing functions with respect to f_d.

Table I. Constants associated with the ACRB

f_d	DA1		DA2		NDA	
	LACRB	HACRB	LACRB	HACRB	LACRB	HACRB
0.1	9.78E-4	2.9E-12	4.18E-5	5.5E-11	3.28E-4	5.45E-4
0.09	8.61E-4	1.1E-12	3.88E-5	5.24E-11	3.11E-4	4.43E-4
0.08	6.87E-4	2.41E-13	3.57E-5	4E-11	2.94E-4	3.85E-4
0.07	5.24E-4	9.2E-14	3.26E-5	2.37E-11	2.79E-4	2.58E-4
0.06	4.51E-4	2E-14	2.95E-5	8.55E-12	2.66E-4	2.35E-4
0.05	3.61E-4	6.67E-15	2.67E-5	1.66E-12	2.54E-4	1.62E-4
0.04	2.64E-4	3E-15	2.44E-5	4E-13	2.45E-4	9.96E-5
0.03	2.05E-4	3E-15	2.24E-5	4E-13	2.37E-4	5.85E-5
0.02	1.53E-4	3E-15	2.07E-5	4E-13	2.32E-4	2.2E-5
0.01	9.73E-5	3E-15	1.93E-5	4E-13	2.28E-4	6.6E-6
0.001	7.25E-5	3E-15	1.86E-5	4E-13	2.27E-4	7.62E-8

For the two DA scenarios, Fig. 12 and Fig. 13 present the LACRB and HACRB for several $f_d's$ together with the true CRB for the corresponding f_d, and Fig. 14 depicts the same curves in the NDA scenario.

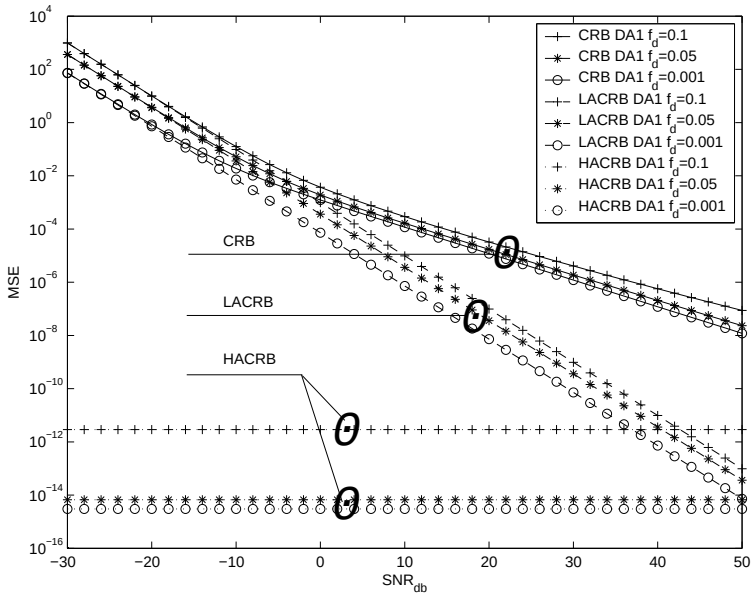


Fig. 12. ACRB for the DA case with Chu training sequences

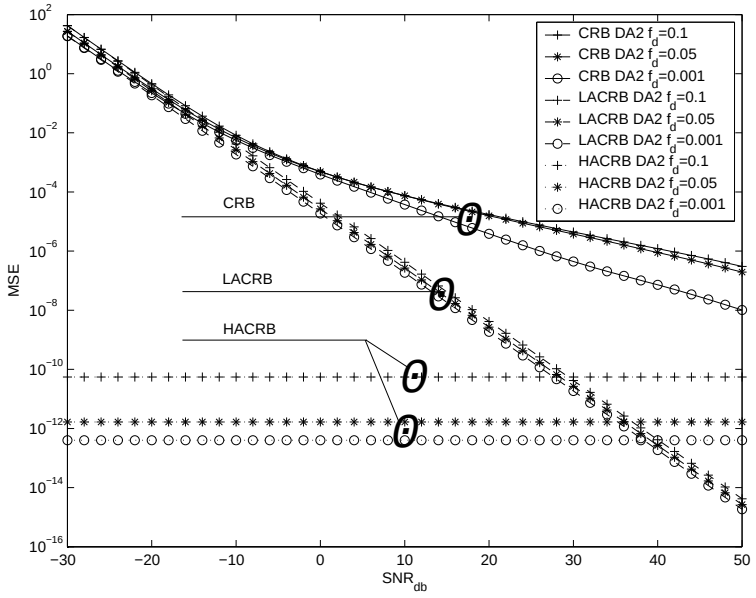


Fig. 13. ACRB for the DA case with optimized training sequences

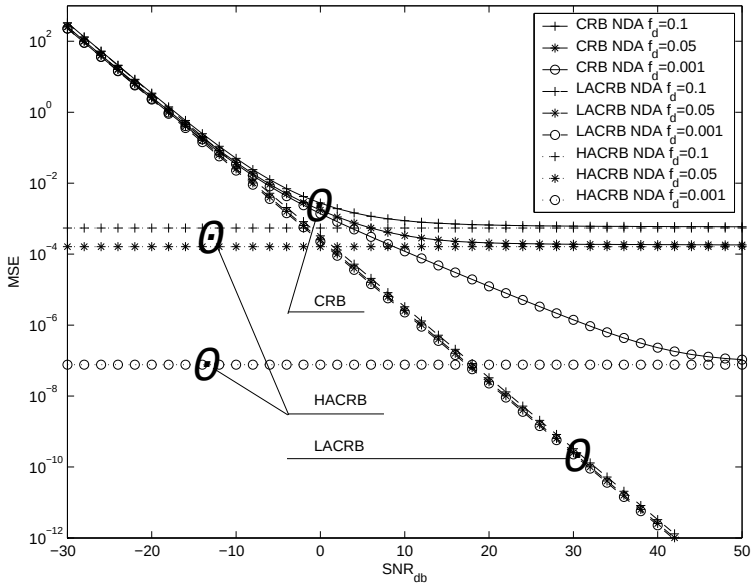
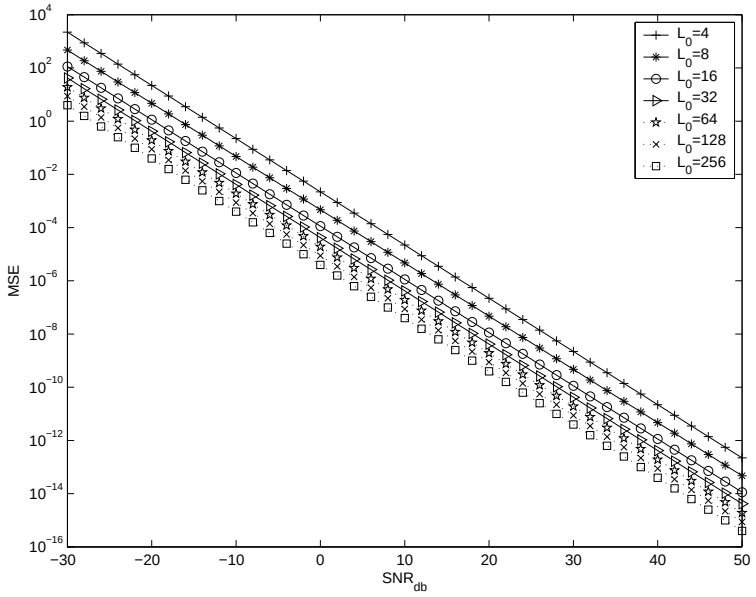


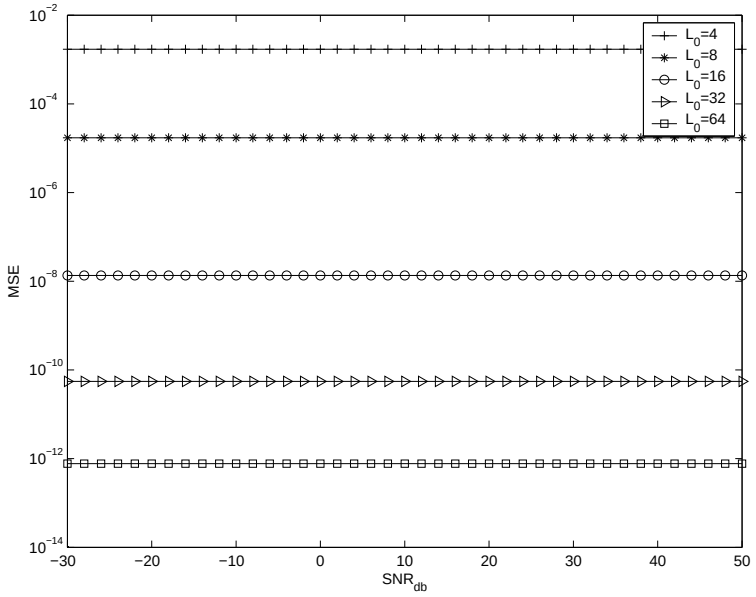
Fig. 14. ACRB for the NDA case

D. The Influence of the Length of Data

This section presents the variation of the ACRB at low and high SNR, respectively, with respect to the length of training data L_0. For the MCRB-optimized training sequences case, Fig. 15 and Fig. 16 are showing the evolution of these bounds and the ACRB associated constants, respectively, with respect to the length of data L_0, assuming very fast fading conditions ($f_d = 0.1$). The ACRB associated constants have been defined in Section C and the MCRB-optimized training sequences have been introduced in Section B. Intuitively, the LACRB and the HACRB are expected to decrease with L_0, because an increase in the length of training data should improve the accuracy of the estimation. Fig. 16 shows that these improvements have a tendency to saturate. Therefore, there is no reason to increase the length of the training data beyond a certain Doppler shift dependent limit.

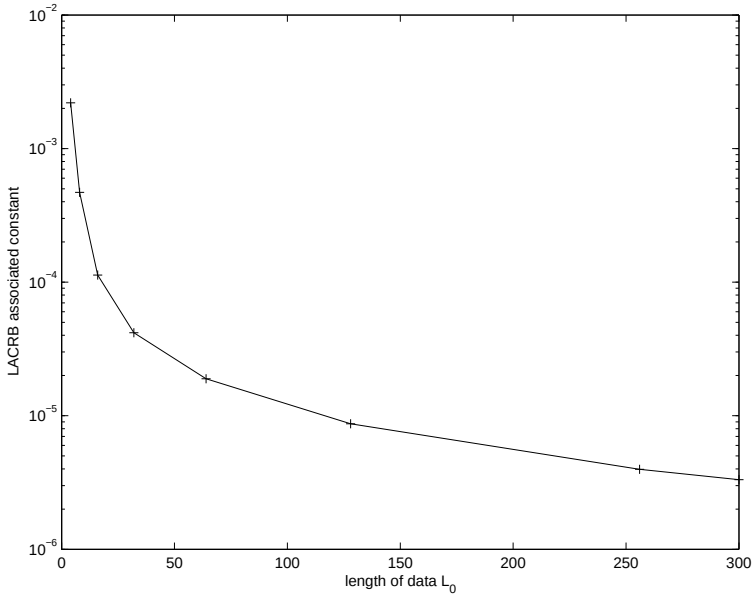


a) LACRB variation

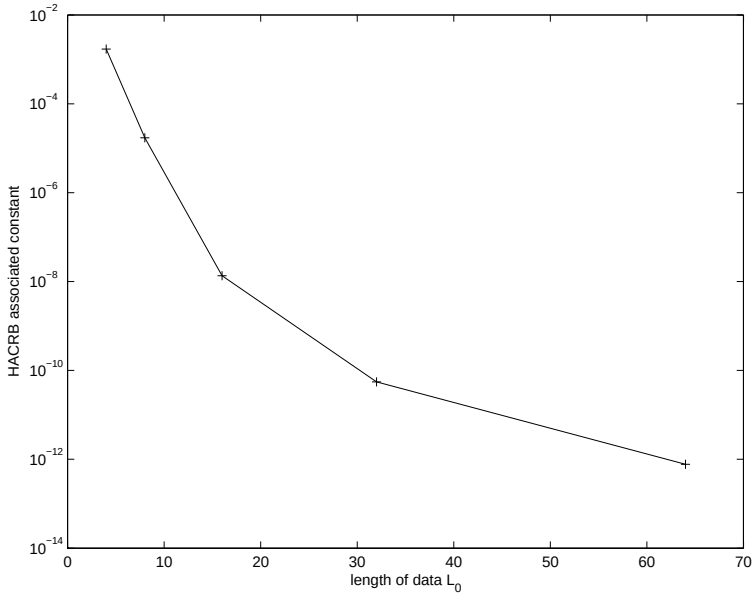


b) HACRB variation

Fig. 15. ACRB variation with respect to the length of data



a) $LACRB$



b) $HACRB$

Fig. 16. Variation of the ACRB associated constants with respect to the length of data

E. Barankin Bound Threshold Analysis

For all DA communications systems it is important to minimize the length of the training sequences in order to maximize the transmission rate and to avoid the waste of energy, because the training symbols do not contain any information. MIMO systems generally need more training than the single antenna systems and therefore the issue of minimizing the length of the training data is even more critical. Also, in TV fading channels very long training sequences are unacceptable because the propagation parameters change in time and the estimation efficiency degrades. On the other hand, the training data should be long enough to assure a good quality of the transmission, described by the bit error rate (BER). For all the considerations above mentioned, it is important to establish the appropriate length that meets all the requirements.

The BB can give a more accurate estimate of the needed length than the CRB because it is a tighter bound. Due to the sudden depart of the BB from the CRB, the quality of the estimation is sharply decreasing. Therefore, the threshold SNR, defined as the SNR where this departure occurs, should be at most the lower limit of the range in which a particular communications system is designed to function. This threshold SNR can be exploited in the design process as follows. For a certain SNR, denoted in this thesis by "target SNR", a certain MSE is established as the highest acceptable value and is denoted by "target MSE". The threshold SNR of a properly designed communications system should be lower than the target SNR. Also, the MSE corresponding to the threshold SNR should be at most equal to the target MSE.

The length of the data should be chosen such that the SNR and the MSE associated to the threshold are, within the required limits, as close to the target values

as possible. Fig. 17 shows an example of design, assuming Chu training sequences and very fast fading ($f_d = 0.1$). The variation of the threshold are shown for training sequences having a length L_0 between 4 and 128 symbols. Both the MSE and the SNR thresholds are decreasing functions with respect to L_0. Therefore, for any L_0 beyond a certain value, the design requirements are met. For example, considering a target MSE of 10^{-1} and a target SNR of -12 dB, any $L_0 \geq 32$ satisfies the requirements. Although sequences with $L_0 = 64$ or $L_0 = 128$ are acceptable, they are overdesigned. Taking into consideration the arguments from the beginning of this section, such systems could prove to be suboptimal.

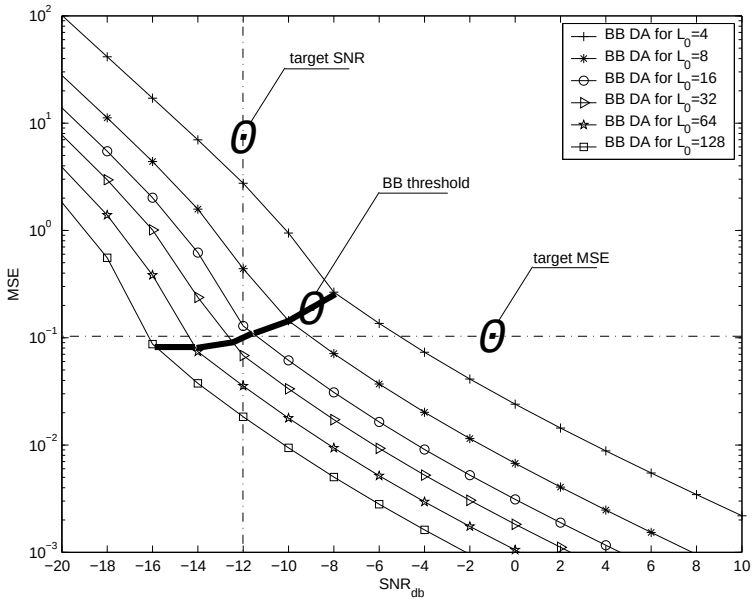


Fig. 17. Threshold variation for Chu training sequences

Fig. 18 shows that similar threshold variations are found in the case of MCRB-optimized sequences, for slow fading and low L_0.

For long sequences in slow fading and for fast fading, due to the design based on eigendecomposition and due to the increasing size of the matrix involved, these

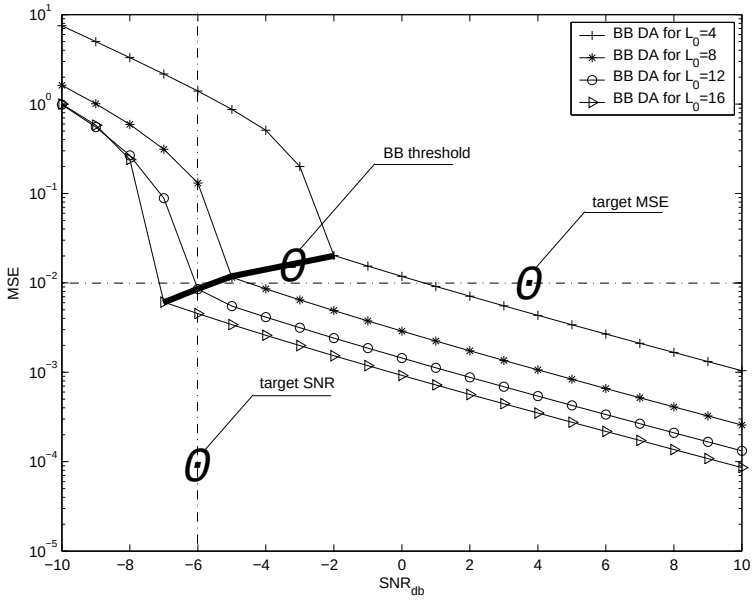


Fig. 18. Threshold variation for short optimized sequences and very slow fading

sequences exhibit a different behavior, presented in Fig. 19 and Fig. 20. To exemplify, in the case of fast fading, for a target MSE of 10^{-2} and a target SNR of 0 dB, the acceptable training length has to be between 8 and 16. For $L_0 < 8$ the CRB at the target SNR becomes larger than the target MSE, and for $L_0 > 16$, although the CRB is still inside the design limits, due to the increasing threshold SNR and to the increasing distance between the BB and the CRB, the BB becomes larger than the target MSE.

The design of a training sequence exhibiting a different behavior for short and long L_0, respectively, as it is the case depicted in Fig. 18 and Fig. 19, is more restrictive than the case presented in Fig. 17 and generally less prohibitive than the case depicted in Fig. 20. Starting from low values of L_0, if the increase of L_0 for the first region does not provide the required results, the search continues in the second region similarly to the case presented in Fig. 20.

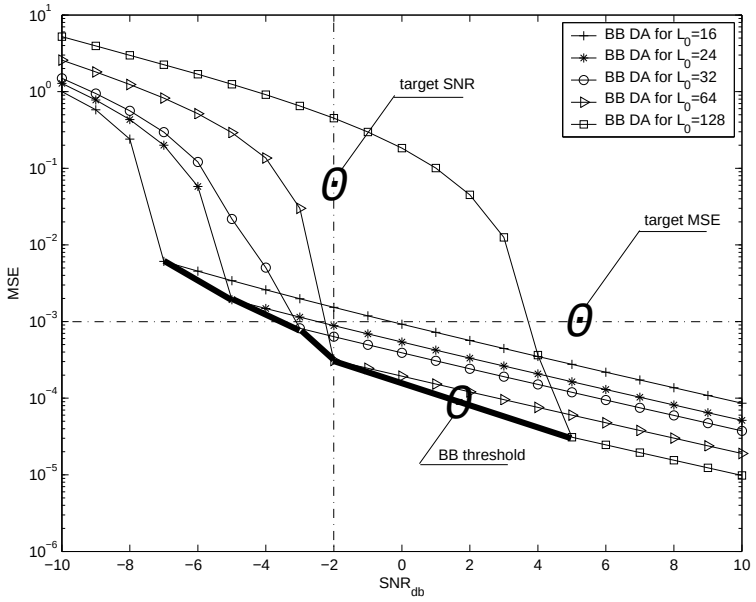


Fig. 19. Threshold variation for long optimized sequences and very slow fading

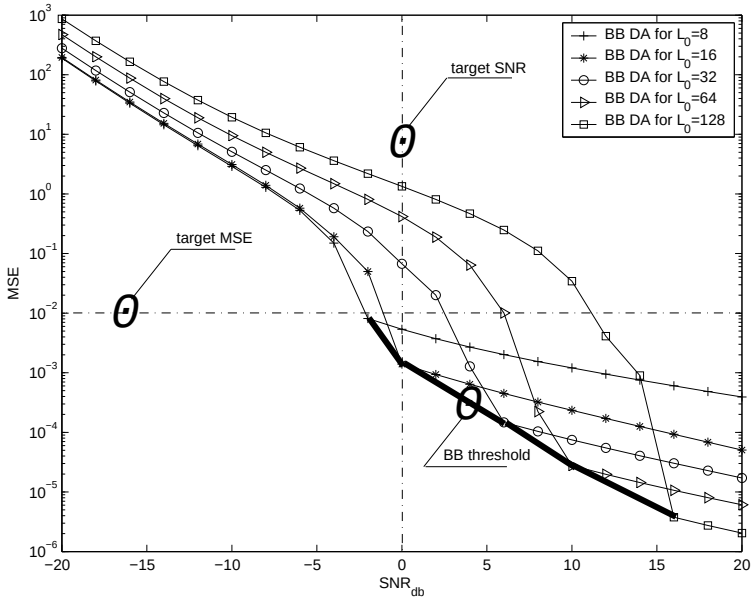


Fig. 20. Threshold variation for optimized sequences and very fast fading

F. The Influence of the Oversampling Factor

Fig. 21 represents the CRB computed under slow fading conditions for the eigendecomposition-based model introduced in Chapter III.B at different values of the oversampling factor Q. The improvement due to the increasing resolution becomes smaller and smaller. From $Q = 2$ to $Q = 4$ there is a gain of approximately 2 dB. The additional gain becomes approximately 0.4 dB when the oversampling factor increases from $Q = 4$ to $Q = 8$ and becomes almost insignificant when Q changes from 8 to 16. Under very fast fading conditions the gains observed are 1.6 dB, respectively 0.3 dB, when Q changes from 2 to 4 and 4 to 8, respectively.

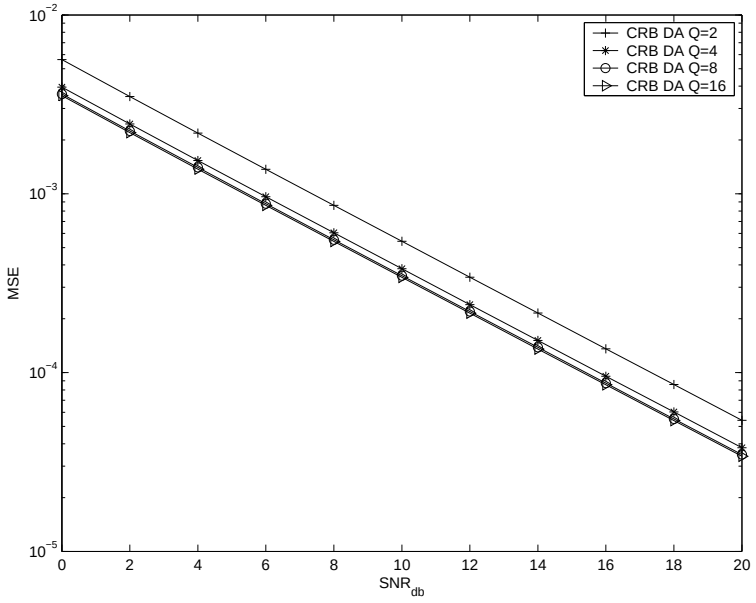


Fig. 21. Influence of the oversampling factor Q

This dependence on Q is valid only for the model above mentioned and is due to the limited precision of the eigendecomposition process. In other words, for low oversampling factors Q, the resolution is not good enough and (3.5) can not be ap-

proximated by its discrete version. The performance of the main model introduced in Chapter III.C does not depend on Q.

G. The Influence of the Fading's Variance

As expected, considering the forms taken by the covariance matrix (see (D.1), (D.2), (D.3) and (D.4)) for the models introduced in Chapter III.B and C, in the Rayleigh case any increase in the variance of the fading is equivalent with an increase in the signal's energy. Therefore, changing σ_α^2 is the same as changing the system's SNR and an equivalent SNR can be defined as

$$\mathrm{SNR}_{eq} = \mathrm{SNR}\sigma_\alpha^2$$

Fig. 22 shows the variation of the CRB with respect to the fading's variance for the

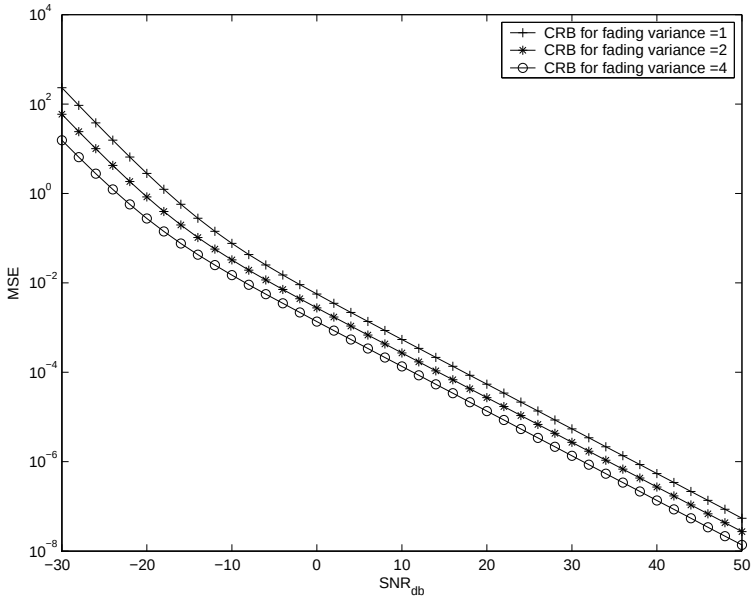


Fig. 22. Influence of the fading's variance

eigendecomposition-based model introduced in Chapter III.B, assuming the fading is very slow varying ($f_d = 0.001$). All the plots in this thesis, except Fig. 22, consider the equivalent SNR or, in other words, assume $\sigma_\alpha^2 = 1$.

H. Eigendecomposition Estimator

For the eigendecomposition-based model introduced in Chapter III.B, the authors in [33] found an ML estimate of the timing delay ε based on the eigenvalue decomposition of the autocorrelation matrix. The expression is proved to be

$$\varepsilon_{ML} = \arg\max_{\varepsilon} \left[\lim_{S\to\infty} \sum_{j=1}^{2} \mathbf{V}_{jS}^H(\varepsilon)(\mathbf{R} + N_0\mathbf{\Lambda}_{S2}^{-1})^{-1}\mathbf{V}_{jS}(\varepsilon) \right] , \tag{5.1}$$

where N_0 is the PSD of the noise, $\mathbf{R}$ is a $2S \times 2S$ matrix with elements r_{mn} given by

$$r_{mn} = \int_{T_i}^{T_f} d_{i1}(t) f_{k1}(t) d_i^*(t) f_k^*(t) dt \ ,$$

with $m = (i-1)S + k, \quad n = (i_1 - 1)S + k_1, \quad i, i_1 = 1, 2, \quad k, k_1 = 1, 2, \cdots, S,$ $\mathbf{V}_{jS}(\varepsilon) = \begin{bmatrix} v_{j11} & v_{j12} & \cdots & v_{j1S} & v_{j21} & v_{j22} & \cdots & v_{j2S} \end{bmatrix}^T$, $v_{jik} = \int_{T_i}^{T_f} r_j(t) d_i^*(t) f_k^*(t) dt$ $\mathbf{\Lambda}_{S2} = \text{diag}[\mathbf{\Lambda}_S \quad \mathbf{\Lambda}_S]$ and $\mathbf{\Lambda}_S = \text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_S)$. As explained in Chapter III.B, S is the number of eigenvalues considered in the eigendecomposition process, $d_i(t)$ is the training data transmitted by antenna i, and $f_k(t)$ and λ_k stand for the eigenfunctions and the eigenvalues of the autocorrelation matrix, respectively.

This estimator's performance is presented in Fig. 23 together with the CRB, both computed for different values of f_d.

Although its performance turns out to be quite far from the theoretical CRB, this estimator has the merit of being robust, in the sense that the errors in the estimation of f_d do not influence much the estimation accuracy of ε. For considerations explained in Section F, an oversampling factor $Q = 16$ has been used for simulations.

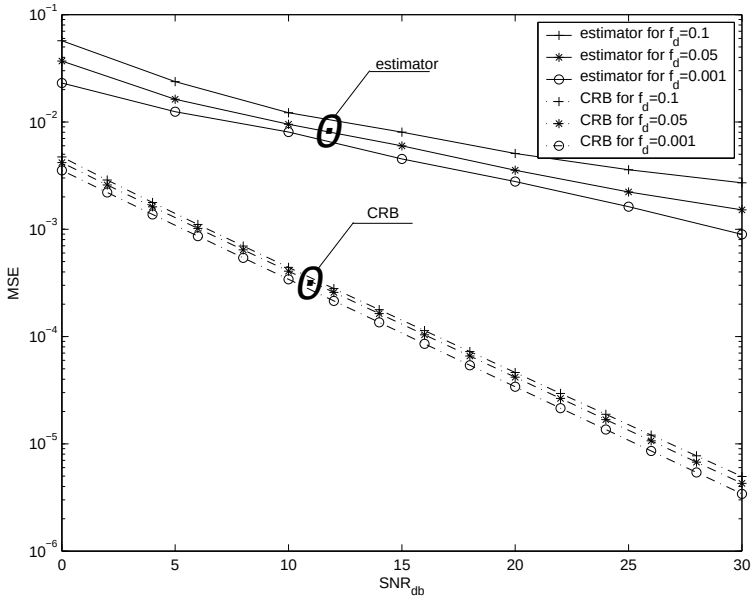


Fig. 23. Eigendecomposition estimator

I. Block Fading Estimator and Influence of the Number of Antennas

In the block fading scenario, the following estimator has been found [34] for the main model introduced in Chapter III.C

$$\hat{\varepsilon} = \arg\max_{\varepsilon}(\mathbf{\Lambda}(\varepsilon)) \tag{5.2}$$

where

- for DA case

$$\mathbf{\Lambda}(\varepsilon) = \sum_{j=1}^{M} \mathbf{r}_j^H \mathbf{A}_\varepsilon \mathbf{Z}(\mathbf{Z}^H \mathbf{A}_\varepsilon^H \mathbf{A}_\varepsilon \mathbf{Z})^{-1} \mathbf{Z}^H \mathbf{A}_\varepsilon^H \mathbf{r}_j$$

- for NDA case

$$\mathbf{\Lambda}(\varepsilon) = \sum_{j=1}^{M} \mathbf{r}_j^H \mathbf{A}_\varepsilon (\mathbf{A}_\varepsilon^H \mathbf{A}_\varepsilon)^{-1} \mathbf{A}_\varepsilon^H \mathbf{r}_j$$

The performance of this estimator is presented in Fig. 24 and Fig. 25 for the DA case and for the number of receive antennas varying between 1 and 4. The corresponding MCRBs are also depicted in these figures. Fig. 24 shows that the MCRB becomes smaller when the number M of receive antennas increases. Also, the performance of the estimator is improving with M. Fig. 25 shows also that the performance of the estimator becomes closer to the theoretical limit for larger M. It has been found that the number of transmit antennas does not influence the performance of the system. Referring to Fig. 3, it is noticed that for the considered range ((0 - 30) dB), the CRB and the BB coincide with the MCRB. The estimator performs very well, but it is very complex, as $\mathbf{\Lambda}(\varepsilon)$ should theoretically be computed for all possible values of (ε).

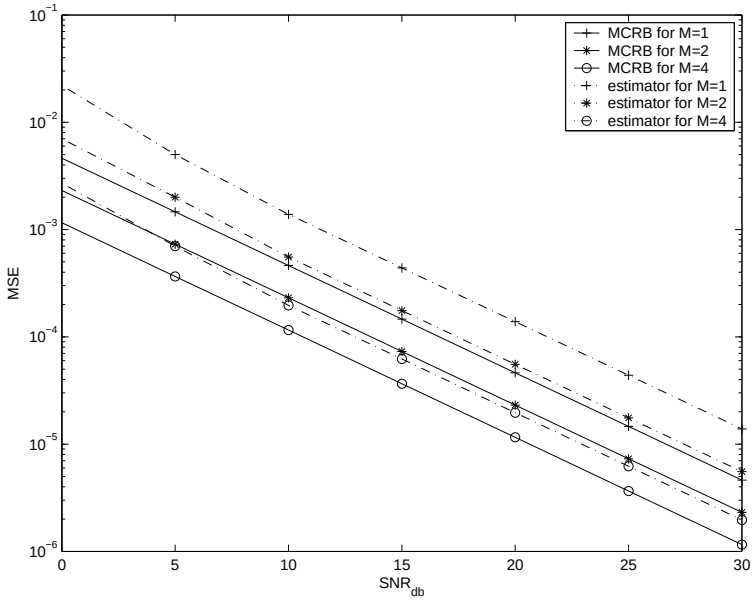


Fig. 24. Estimator for M = 1, 2 and 4

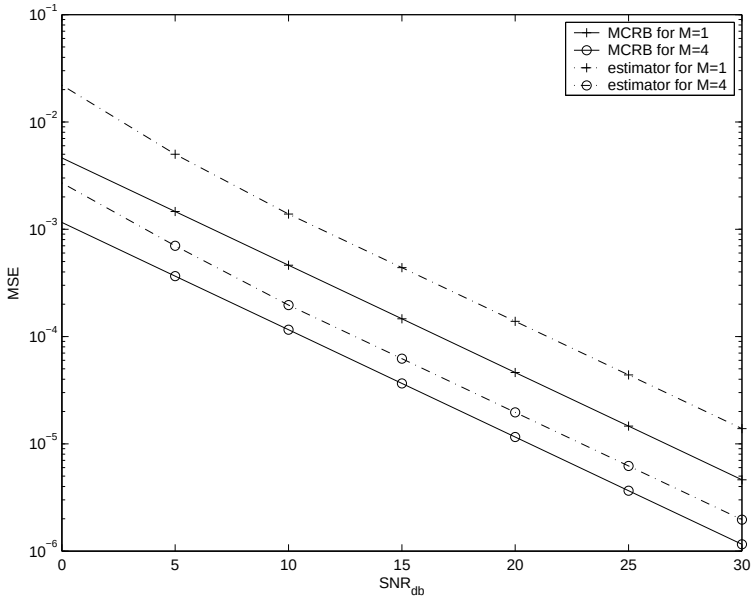


Fig. 25. Estimator for M = 1 and 4

J. Non-Linearity Based Non-Data-Aided Estimators

In practice, the complexity of the estimators should be as low as possible. Therefore, estimators based on exhaustive searching algorithms, like the one presented in Section I, are not used in practice. In the NDA scenario, there exists a family of estimators that extract the information about the timing delay from the phase of a spectral line, obtained after passing the received signal, filtered and sampled, through a non-linear device. In the continuous case, the useful spectral component at $1/T_s$ could be extracted by a PLL and a narrow-band filter. In the discrete case, the same component is obtained by computing the complex Fourier coefficient at the symbol rate. The general form of such an estimator, with the notations used in this thesis, is

$$\hat{\varepsilon} = -\frac{1}{2\pi}\arg\left\{\sum_{k=0}^{L_0Q-1} F(r_j(kT_s))\exp\frac{-j2\pi k}{Q}\right\} \tag{5.3}$$

where $F(\cdot)$ stands for the non-linearity. The most used non-linearities are the square-law (SL), the absolute value (AV), the fourth-law (FL) and the logarithmic non-linearities (LOG). The most popular in the literature is the square-law non-linearity and the estimator is also known as the Oerder&Meyr estimator [7]. The performance

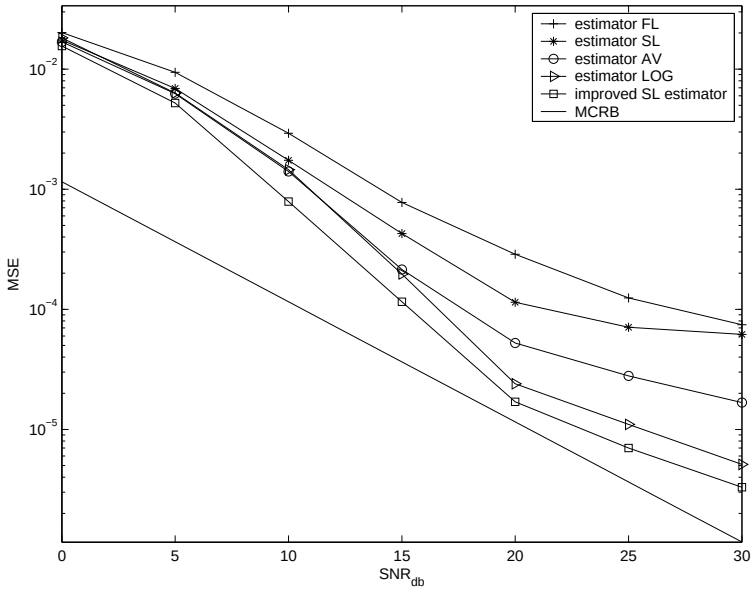


Fig. 26. Non-linearity based estimators

of the NDA estimators is depicted in Fig. 26, for all the non-linearities mentioned. The forms taken by the function $F(\cdot)$ are as follows

$$F(r_j(kT_s)) =$$

- $\mid r_j(kT_s) \mid^2$ for SL

- $\mid r_j(kT_s) \mid$ for AV

- $\mid r_j(kT_s) \mid^4$ for FL

- $\log(1 + SNR^2 \mid r_j(kT_s) \mid^2)$ for LOG

At the output of the non-linearity the bandwidth is increased. For example, considering an excess bandwidth factor β and the SL non-linearity, in order to avoid the aliasing effect for frequencies below $1/T_s$, the sampling rate needs to be at least $(2+\beta)/T_s$. In practical applications, the oversampling factor is chosen in general to be $Q = 4$, and this is the factor used for the simulations depicted in Fig. 26. Besides the estimators above mentioned and the MCRB, presented as a reference, there appears one more curve on Fig. 26, denoted by the name "the improved SL estimator". This is a generalization of the Oerder&Meyr estimator, is easier to compute, does not require an oversampling factor larger than 2 and is defined as follows [36].

$$\hat{\varepsilon} = -\frac{1}{2\pi}\arg\left\{\sum_{k=0}^{K-1} \mathbf{\Lambda}(k)\exp\frac{-j2\pi k}{K}\right\} , \tag{5.4}$$

where $\mathbf{\Lambda}(\varepsilon) = \sum_{j=1}^{M} \mathbf{r}_j^H \mathbf{A}_\varepsilon (\mathbf{A}_\varepsilon^H \mathbf{A}_\varepsilon)^{-1} \mathbf{A}_\varepsilon^H \mathbf{r}_j$ is computed over a set of K uniformly spaced values from 0 to $(K-1)/K$. Besides the better performance compared to the classical SL estimator, visible on Fig. 26, the improved SL estimator is easier to implement, because the factors $\mathbf{A}_\varepsilon (\mathbf{A}_\varepsilon^H \mathbf{A}_\varepsilon)^{-1} \mathbf{A}_\varepsilon^H$ can be pre-computed and $K = 4$ is enough to obtain good performance. In all the simulations above, the overall estimate $\hat{\varepsilon}$ has been computed as the average between the individual estimates of the M receive antennas.

CHAPTER VI

CONCLUSIONS

A. Summary of the Thesis

To summarize, the work in this thesis focused on the issue of timing delay synchronization. The mean square error has been presented as a good indicator of the performance of estimators. Bounds like the Cramer-Rao bound and the Barankin bound have been used to analyze the quality of estimation. An extended model for time varying MIMO systems has been found to provide compact expressions for these bounds. The fading has been assumed to be time-varying, independent from one pair of antennas to another, and Rayleigh distributed, and to respect the Jakes' model. Autoregressive methods can be used to generate the fading time-varying coefficients. A detailed analysis of the system models has been presented, pointing out the influence of different parameters, like the length of the training data, the oversampling factor and the fading's variance, upon the bounds. A cost function has been introduced in order to highlight the price paid for not knowing the nuisance parameters, the threshold effect in the Barankin bound has been exploited in order to provide information about the required length of the observation and the performance of several estimators has been compared.

B. Suggestions for Future Work

In Fig. 16 it has been shown that the coefficients of the low- and high-SNR ACRB, respectively, have a tendency to reach a saturation floor when the length of observation increases over a certain limit. Due to the limited precision of the simulation process, this floor cannot be determined precisely, especially at high SNR. Therefore,

developing an analytical expression for the evolution of these coefficients with respect to the observation length could be useful.

The model introduced in Chapter III.C assumes independent realizations of the fading from one pair of antennas to another. Expressing a model for the symbol-level time-varying correlated fading case and studying the influence that the correlation has upon the bounds and estimators presented could be of great importance. In the block fading case, a model that accounts for correlated coefficients between pairs of antennas has been presented in [34], where the transmit and the receive correlation matrix, respectively, are based on measurements performed by Nokia.

Some additional interesting problems to explore would be to optimize the bandwidth and to consider a spectral approach in expressing the asymptotic bounds. Also, the performance of the NDA estimators depicted in Fig. 26 depends on the roll-off factor. In this thesis, it was assumed that $\beta = 0.3$, but a study could be done by varying it from 0 to 1.

A detailed analysis for all bounds and estimators presented in this thesis could be done for other types of fading. In-depth analysis of the Rice fading could be done in a manner similar to the one used in this thesis.

REFERENCES

[1] T. F. Wong and T. M. Lok, notes for "Theory of Digital Communications", Chapter 3: Synchronization, Accessed May 2004, Available online at http://www.wireless.ece.ufl.edu/twong/Notes/Comm/ch3.pdf.

[2] H. Meyr, M. Moeneclaey, and S. A. Fechtel, *Digital Communication Receivers: Synchronization, Channel Estimation, and Signal Processing*, New York, N.Y., John Wiley & Sons, 1998.

[3] G. Vazquez, J. Riba, G. B. Giannakis, Y. Hua, P. Stoica, and L. Tong, *Signal Processing Advanced Wireless Mobile Communications*, vol. 2, ch. 9., pp. 357-402, Englewood Cliffs, N.J., Prentice-Hall, Jul. 2000, Accessed May 2004, Available online at http://gpstsc.upc.es/comm/jriba/publications_data.html.

[4] W. A. Gardner, "The Role of Spectral Correlation in Design and Performance Analysis of Synchronizers", *IEEE Transactions on Communications*, vol. 34, pp. 1089-1095, Nov. 1986.

[5] F. Gini and G. B. Giannakis, "Frequency Offset and Symbol Timing Recovery in Flat Fading Channels: a Unifying Cyclostationary Approach", *IEEE Transactions on Communications*, vol. 46, pp. 400-411, Mar. 1998.

[6] J. Riba and G. Vazquez, "Bayesian Recursive Estimation of Frequency and Timing Exploiting the Cyclostationary Property", *Signal Processing (EURASIP)*, vol. 40, pp. 21-37, Oct. 1994.

[7] M. Oerder and H. Meyr, "Digital Filter and Square Timing Recovery", *IEEE Transactions on Communications*, vol. 36, pp. 605-612, May 1988.

[8] F. M. Gardner, "A BPSK/QPSK Timing-Error Detector for Sampled Receivers", *IEEE Transactions on Communications*, vol. 34, pp. 423-429, May 1986.

[9] S. M. Kay, *Fundamentals of Statistical Signal Processing, vol. 1: Estimation Theory*, Upper Saddle River, N.J., Prentice Hall, 1993.

[10] Y. Jiang, F.-W. Sun, and J. S. Baras, "On the Performance Limits of Data-Aided Synchronization", *IEEE Transactions on Information Theory*, vol. 49, no. 1, pp. 191-203, Jan. 2003.

[11] I. Bergel and A. J. Weiss, "Cramer-Rao Bound on Timing Recovery of Linearly Modulated Signals with No ISI", *IEEE Transactions on Communications*, vol. 51, no. 4, pp. 634-640, Apr. 2003.

[12] M. Ghogho, A. K. Nandi, and A. Swami, "Cramer-Rao Bounds and Maximum Likelihood Estimation for Random Amplitude Phase-Modulated Signals", *IEEE Transactions on Signal Processing*, vol. 47, no. 11, pp. 2905-2916, Nov. 1999.

[13] A. Swami, "Cramer-Rao Bounds for Deterministic Signals in Additive and Multiplicative Noise", *Signal Processing*, vol. 53, pp. 231-244, 1996.

[14] O. Besson and P. Stoica, "On Parameter Estimation of MIMO Flat-Fading Channels with Frequency Offsets", *IEEE Transactions on Signal Processing*, vol. 51, no. 3, pp. 602-613, Mar. 2003.

[15] P. Forster and P. Larzabal, "On Lower Bounds for Deterministic Parameter Estimation", *International Conference on Acoustics, Speech, and Signal Processing, (ICASSP)*, vol. 2, pp. 1137-1140, 2002.

[16] A. D'Andrea, U. Mengali, and R. Reggiannini, "The Modified Cramer-Rao Bound and Its Application to Synchronization Problems", *IEEE Transactions*

on Communications, vol. 42, no. 2/3/4, pp. 1391-1399, Feb./Mar./Apr. 1994.

[17] F. Gini, R. Reggiannini, and U. Mengali, "The Modified Cramer-Rao Bound in Vector Parameter Estimation", *IEEE Transactions on Communications*, vol. 46, no. 1, pp. 52-60, Jan. 1998.

[18] M. Moeneclaey, "On the True and the Modified Cramer-Rao Bounds for the Estimation of a Scalar Parameter in the Presence of Nuisance Parameters", *IEEE Transactions on Communications*, vol. 46, no. 11, pp. 1536-1544, Nov. 1998.

[19] E. W. Barankin, "Locally Best Estimates", *Annals of Mathematical Statistics*, vol. 20, pp. 477-501, 1949.

[20] J. Kiefer, "On Minimum Variance Estimation", *Annals of Mathematical Statistics*, vol. 23, pp. 627-629, 1952.

[21] R. J. McAulay and L. Seidman, "A Useful Form of the Barankin Lower Bound and Its Application to PPM Threshold Analysis", *IEEE Transactions on Information Theory*, vol. IT-15, no. 2, pp. 273-279, Mar. 1969.

[22] J. Albuquerque, "The Barankin Bound: A Geometric Interpretation", *IEEE Transactions on Information Theory*, vol. 19, no. 4, pp. 559-561, Jul. 1973.

[23] A. Zeira and P. M. Schultheiss, "Realizable Lower Bounds for Time Delay Estimation", *IEEE Transactions on Signal Processing*, vol. 41, no. 11, pp. 3102-3113, Nov. 1993.

[24] A. Zeira and P. M. Schultheiss, "Realizable Lower Bounds for Time Delay Estimation: Part 2-Threshold Phenomena", *IEEE Transactions on Signal Processing*, vol. 42, no. 5, pp. 1001-1007, May 1994.

[25] R. J. McAulay and E. M. Hofstetter, "Barankin Bounds on Parameter Estimation", *IEEE Transactions on Information Theory*, vol. IT-17, no. 6, pp. 669-676, Nov. 1971.

[26] I. Reuven and H. Messer, "A Barankin-Type Lower Bound on the Estimation Error of a Hybrid Parameter Vector", *IEEE Transactions on Information Theory*, vol. 43, no. 3, pp. 1084-1093, 1997.

[27] L.N. Atallah, J.-P. Barbot, and P. Larzabal, "From Chapman-Robbins Bound Towards Barankin Bound in Threshold Behaviour Prediction", *Electronics Letters*, vol. 40, no. 4, pp. 279-280, Feb. 2004.

[28] D. G. Chapman and H. Robbins, "Minimum Variance Estimation Without Regularity Assumption", *Annals of Mathematical Statistics*, vol. 22, pp. 581-586, 1951.

[29] L. Knockaert, "The Barankin Bound and Threshold Behavior in Frequency Estimation", *IEEE Transactions on Signal Processing*, vol. 45, no. 9, pp. 2398-2401, Sep. 1997.

[30] K. E. Baddour and N. C. Beaulieu, "Autoregressive Models for Fading Channel Simmulation", *IEEE Global Telecommunications Conference, 2001 (GLOBECOM '01)*, vol. 2 , pp. 1187-1192, Nov. 2001.

[31] L. Lindbom, A. Ahlen, M. Sternad and M. Falkenstrom, "Tracking of Time-Varying Mobile Radio Channels-Part II: A Case Study", *IEEE Transactions on Communications*, vol. 50, issue 1, pp. 156-167, Jan. 2002.

[32] K. E. Baddour and N. C. Beaulieu, "Accurate Simulation of Multiple Cross-Correlated Fading Channels", *IEEE International Conference on Communica-*

tions 2002 (ICC 2002), vol. 1, pp. 267-271, Apr.-May 2002.

[33] S.-A. Yang and J. Wu, "A Robust Timing Synchronization Scheme in Multiple Antenna Systems with Doppler Frequency Shifts", *IEEE Communications Letters*, vol. 7, no. 3, pp. 118-120, Mar. 2003.

[34] Y.-C. Wu and E. Serpedin, "Maximum Likelihood Symbol Timing Estimation in MIMO Correlated Fading Channels", *Wireless Communications and Mobile Computing Journal (WCMC), Special Issue on "Multiple-Input Multiple-Output (MIMO) Communications" (submitted)*, Nov. 2003.

[35] Y.-C. Wu S.-C. Chan, and E. Serpedin, "Symbol-Timing Estimation in Space-Time Coding Systems Based on Orthogonal Training Sequences", *IEEE Transaction on Wireless Communications, (accepted for publication)*, Mar. 2003.

[36] Y.-C. Wu and E. Serpedin, "Low-Complexity Feedforward Symbol Timing Estimator Using Conditional Maximum-Likelihood Principle", *IEEE Communications Letters*, vol. 8, no. 3, pp. 168-170, Mar. 2004.

[37] F. Gini, M. Luise, and R. Reggiannini, "Cramer-Rao Bounds in the Parametric Estimation of Fading Radiotransmission Channels", *IEEE Transactions on Communications*, vol. 46, no. 10, pp. 1390-1398, Oct. 1998.

[38] R. Dabora, J. Goldberg, and H. Messer, "Cramer Rao Bound Analysis for Data Aided Time Synchronization of MSK over a Fast Fading Channel", *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2000 (ICASSP '00)*, vol. 5, pp. 2521-2524, 5-9 Jun. 2000.

APPENDIX A

DERIVATION OF THE MBB

The inequality MBB $\leq$ BB can be proven using the Cauchy-Schwarz inequality

$$\left(\int_a^b f(h)g(h)dh\right)^2 \leq \int_a^b f^2(h)dh \int_a^b g^2(h)dh \ .$$

With $f^2(\mathbf{h}) = \frac{p^2(\mathbf{r},\mathbf{h},\varepsilon+\eta)}{p(\mathbf{r},\mathbf{h},\varepsilon)}$ and $g^2(\mathbf{h}) = p(\mathbf{r},\mathbf{h},\varepsilon)$, the Cauchy-Schwarz inequality resumes to

$$\frac{\left(\int p(\mathbf{r},\mathbf{h},\varepsilon+\eta)d\mathbf{h}\right)^2}{\int p(\mathbf{r},\mathbf{h},\varepsilon)d\mathbf{h}} \leq \int \frac{p^2(\mathbf{r},\mathbf{h},\varepsilon+\eta)}{p(\mathbf{r},\mathbf{h},\varepsilon)}d\mathbf{h} \ ,$$

$$\frac{p^2(\mathbf{r};\varepsilon+\eta)}{p(\mathbf{r};\varepsilon)} \leq \int \frac{p^2(\mathbf{r} \mid \mathbf{h};\varepsilon+\eta)}{p(\mathbf{r} \mid \mathbf{h};\varepsilon)}p(\mathbf{h})d\mathbf{h} \ ,$$

from which it is obvious that

$$\sup_{\eta}\frac{\eta^2}{\int \frac{p^2(\mathbf{r};\varepsilon+\eta)}{p(\mathbf{r};\varepsilon)}d\mathbf{r} - 1} \geq \sup_{\eta}\frac{\eta^2}{\int \left[\int \frac{p^2(\mathbf{r}|\mathbf{h};\varepsilon+\eta)}{p(\mathbf{r}|\mathbf{h};\varepsilon)}p(\mathbf{h})d\mathbf{h}\right] d\mathbf{r} - 1} \ .$$

Thus, it has been proven that the MBB given by expression (4.5) is always looser than the BB given by expression (2.3). To determine the forms of the MBB for the model used in this thesis, the first step is to note that, in general, for circular white Gaussian noise

$$\mathbf{r} = \xi\mathbf{\Psi}_{\varepsilon}\mathbf{h} + \mathbf{n} \ , \tag{A.1}$$

and

$$p(\mathbf{r} \mid \mathbf{h};\varepsilon) = \frac{1}{(\pi\sigma^2)^{L_0 Q}}\exp\left[-\frac{(\mathbf{r}-\xi\mathbf{\Psi}_{\varepsilon}\mathbf{h})^H(\mathbf{r}-\xi\mathbf{\Psi}_{\varepsilon}\mathbf{h})}{\sigma^2}\right] \ .$$

Therefore

$$\int E_{\mathbf{h}}\left[\frac{p^2(\mathbf{r}|\mathbf{h};\varepsilon+\eta)}{p(\mathbf{r}|\mathbf{h};\varepsilon)}\right]d\mathbf{r} = \int E_{\mathbf{h}}\left[\exp\left\{\frac{2\xi^2((\mathbf{\Psi}_{\varepsilon+\eta}-\mathbf{\Psi}_{\varepsilon})\mathbf{h})^H((\mathbf{\Psi}_{\varepsilon+\eta}-\mathbf{\Psi}_{\varepsilon})\mathbf{h})}{\sigma^2}\right\}\right.$$

$$\cdot \frac{1}{(\pi\sigma^2)^{L_0 Q}} \exp\left\{\frac{(\mathbf{r}-\xi(2\mathbf{\Psi}_{\varepsilon+\eta}-\mathbf{\Psi}_{\varepsilon})\mathbf{h})^H(\mathbf{r}-\xi(2\mathbf{\Psi}_{\varepsilon+\eta}-\mathbf{\Psi}_{\varepsilon})\mathbf{h})}{\sigma^2}\right\} d\mathbf{r}$$

$$= E_{\mathbf{h}}\left[\exp\left\{\frac{2\xi^2((\mathbf{\Psi}_{\varepsilon+\eta}-\mathbf{\Psi}_{\varepsilon})\mathbf{h})^H((\mathbf{\Psi}_{\varepsilon+\eta}-\mathbf{\Psi}_{\varepsilon})\mathbf{h})}{\sigma^2}\right\}\right]$$

$$\cdot \int E_{\mathbf{h}}\left[\frac{1}{(\pi\sigma^2)^{L_0 Q}} \exp\left\{\frac{(\mathbf{r}-\xi(2\mathbf{\Psi}_{\varepsilon+\eta}-\mathbf{\Psi}_{\varepsilon})\mathbf{h})^H(\mathbf{r}-\xi(2\mathbf{\Psi}_{\varepsilon+\eta}-\mathbf{\Psi}_{\varepsilon})\mathbf{h})}{\sigma^2}\right\}\right] d\mathbf{r}$$

$$= E_{\mathbf{h}}\left[\exp\left\{\frac{2\xi^2((\mathbf{\Psi}_{\varepsilon+\eta}-\mathbf{\Psi}_{\varepsilon})\mathbf{h})^H((\mathbf{\Psi}_{\varepsilon+\eta}-\mathbf{\Psi}_{\varepsilon})\mathbf{h})}{\sigma^2}\right\}\right] \ ,$$

because the integral part in the last equality corresponds to a Gaussian distribution. Because $\mathbf{h}$ is a column vector, a $\mathrm{tr}(\cdot)$ operator can be introduced. Thus,

$$\int E_{\mathbf{h}}\left[\frac{p^2(\mathbf{r}|\mathbf{h};\varepsilon+\eta)}{p(\mathbf{r}|\mathbf{h};\varepsilon)}\right] d\mathbf{r}$$

$$= E_{\mathbf{h}}\left[\exp\left\{\frac{2\xi^2}{\sigma^2}\mathrm{tr}\left(\mathbf{h}^H(\mathbf{\Psi}_{\varepsilon+\eta}-\mathbf{\Psi}_{\varepsilon})^H(\mathbf{\Psi}_{\varepsilon+\eta}-\mathbf{\Psi}_{\varepsilon})\mathbf{h}\right)\right\}\right]$$

$$= E_{\mathbf{h}}\left[\exp\left\{\frac{2\xi^2}{\sigma^2}\mathrm{tr}\left((\mathbf{\Psi}_{\varepsilon+\eta}-\mathbf{\Psi}_{\varepsilon})^H(\mathbf{\Psi}_{\varepsilon+\eta}-\mathbf{\Psi}_{\varepsilon})\mathbf{h}\mathbf{h}^H\right)\right\}\right]$$

$$= \exp\left\{\frac{2\xi^2}{\sigma^2}\mathrm{tr}\left((\mathbf{\Psi}_{\varepsilon+\eta}-\mathbf{\Psi}_{\varepsilon})^H(\mathbf{\Psi}_{\varepsilon+\eta}-\mathbf{\Psi}_{\varepsilon})E_{\mathbf{h}}\left[\mathbf{h}\mathbf{h}^H\right]\right)\right\} \ .$$

In the DA case $\mathbf{h} = \mathrm{vec}(\mathbf{H}^T)$ (see (3.9)), where $\mathbf{H}$ has been defined in (3.7). Assuming independent fading coefficients, $E_{\mathbf{h}}\left[\mathbf{h}\mathbf{h}^H\right]$ is an identity matrix multiplied by σ_α^2. In the NDA case, $\mathbf{h} = \mathrm{vec}(\mathbf{Z}\mathbf{H}^T)$ (see (3.8)) and for the code described by the matrix (4.3), it has been shown [34] that $E_{\mathbf{h}}\left[\mathbf{h}\mathbf{h}^H\right]$ is also an identity matrix multiplied by σ_α^2. Therefore, assuming $\sigma_\alpha^2 = 1$

$$\int E_{\mathbf{h}}\left[\frac{p^2(\mathbf{r} \mid \mathbf{h};\varepsilon+\eta)}{p(\mathbf{r} \mid \mathbf{h};\varepsilon)}\right] d\mathbf{r} = \exp\left\{\frac{2\xi^2}{\sigma^2}\mathrm{tr}\left((\mathbf{\Psi}_{\varepsilon+\eta}-\mathbf{\Psi}_{\varepsilon})^H(\mathbf{\Psi}_{\varepsilon+\eta}-\mathbf{\Psi}_{\varepsilon})\right)\right\} \ ,$$

where $\mathbf{\Psi}_{\varepsilon}$ has been defined in (3.16) and (3.17). Plugging this into (4.5)

$$\mathrm{MBB} = \sup_{\eta} \frac{\eta^2}{\exp\left\{\frac{2\xi^2}{\sigma^2}\mathrm{tr}\left((\mathbf{\Psi}_{\varepsilon+\eta}-\mathbf{\Psi}_{\varepsilon})^H(\mathbf{\Psi}_{\varepsilon+\eta}-\mathbf{\Psi}_{\varepsilon})\right)\right\} - 1} \ . \tag{A.2}$$

APPENDIX B

DERIVATION OF THE CRB FOR BLOCK FADING

For the eigendecomposition-based model introduced in Chapter III.B the signal received by antenna j ($j = 1, 2$) can be expressed in vectorial form. Because the oversampling factor is Q and the number of symbols considered for processing is L_0, all vectors have length L_0Q. The fading coefficients h_{ij} are assumed constant for a block interval and independent from one pair of antennas to another, and are assumed to be identically and circularly distributed complex Gaussian variables, with zero-mean and variance σ_α^2. In vectorial form, expression (3.4) can be written as

$$\bar{\mathbf{r}}_j = \bar{\mathbf{d}}_1 h_{1j} + \bar{\mathbf{d}}_2 h_{2j} + \bar{\mathbf{n}}_j \ , \tag{B.1}$$

where $\bar{\mathbf{d}}_i = \left[\begin{array}{cccc} d_i(0) & d_i(T_s) & \cdots & d_i((L_0Q-1)T_s) \end{array} \right]$, $(i = 1, 2)$, $h_{ij} \sim N(0, \sigma_\alpha^2)$ and $\bar{\mathbf{n}}_j \sim N(0, N_0\mathbf{I}_{L_0Q})$. Therefore, $\bar{\mathbf{r}}_j$ is also a Gaussian vector with zero-mean and covariance matrix

$$E\left[\bar{\mathbf{r}}_j^H \bar{\mathbf{r}}_j\right] = E\left[(\bar{\mathbf{d}}_1 h_{1j} + \bar{\mathbf{d}}_2 h_{2j} + \bar{\mathbf{n}}_j)^H (\bar{\mathbf{d}}_1 h_{1j} + \bar{\mathbf{d}}_2 h_{2j} + \bar{\mathbf{n}}_j)\right] \ .$$

Because $E\left[h_{ij} h_{kl}\right] = 0$ for $i \neq k$ or $j \neq l$ the above equation becomes

$$\tilde{\mathbf{R}} = E\left[\bar{\mathbf{r}}_j^H \bar{\mathbf{r}}_j\right] = \sigma_\alpha^2 \left[\bar{\mathbf{d}}_1^H \bar{\mathbf{d}}_1 + \bar{\mathbf{d}}_2^H \bar{\mathbf{d}}_2\right] + N_0 \mathbf{I}_{L_0Q} \ . \tag{B.2}$$

Therefore

$$p_{\bar{\mathbf{r}}_j}(\bar{\mathbf{r}}_j) = \frac{1}{\pi^{L_0Q} \mid \tilde{\mathbf{R}} \mid} \exp\left[-\bar{\mathbf{r}}_j \tilde{\mathbf{R}}^{-1} \bar{\mathbf{r}}_j^H\right] \ . \tag{B.3}$$

Using the assumption of independence between the signals received by the two antennas

$$p_{\bar{\mathbf{r}}}(\bar{\mathbf{r}}) = \prod_{j=1}^{2} p_{\bar{\mathbf{r}}_j}(\bar{\mathbf{r}}_j) \ , \tag{B.4}$$

the log-likelihood function becomes

$$\mathbf{\Lambda} = \ln p_{\bar{\mathbf{r}}}(\bar{\mathbf{r}}) = -2L_0 Q \ln(\pi) - 2\ln \mid \tilde{\mathbf{R}} \mid - \sum_{j=1}^{2} \bar{\mathbf{r}}_j \tilde{\mathbf{R}}^{-1} \bar{\mathbf{r}}_j^H \ .$$

The second term's partial derivative w.r.t. β is

$$\frac{\partial \ln \mid \tilde{\mathbf{R}} \mid}{\partial \beta} = \frac{\partial \mathrm{tr}(\ln \tilde{\mathbf{R}})}{\partial \beta} = \mathrm{tr}\left(\frac{\partial \ln \tilde{\mathbf{R}}}{\partial \beta}\right) = \mathrm{tr}\left(\tilde{\mathbf{R}}^{-1} \frac{\partial \tilde{\mathbf{R}}}{\partial \beta}\right) \ ,$$

where $\beta = \sigma_\alpha^2, N_0, \varepsilon$.

The third term's partial derivative w.r.t. β is

$$\frac{\partial (\bar{\mathbf{r}}_j \tilde{\mathbf{R}}^{-1} \bar{\mathbf{r}}_j^H)}{\partial \beta} = \bar{\mathbf{r}}_j \frac{\partial \tilde{\mathbf{R}}^{-1}}{\partial \beta} \bar{\mathbf{r}}_j^H = -\bar{\mathbf{r}}_j \tilde{\mathbf{R}}^{-1} \frac{\partial \tilde{\mathbf{R}}}{\partial \beta} \tilde{\mathbf{R}}^{-1} \bar{\mathbf{r}}_j^H \ .$$

Adding all terms, the general form of the derivative of the log-likelihood function w.r.t. β becomes

$$\frac{\partial \mathbf{\Lambda}}{\partial \beta} = -2\mathrm{tr}\left(\tilde{\mathbf{R}}^{-1} \frac{\partial \tilde{\mathbf{R}}}{\partial \beta}\right) + \sum_{j=1}^{2} \bar{\mathbf{r}}_j \tilde{\mathbf{R}}^{-1} \frac{\partial \tilde{\mathbf{R}}}{\partial \beta} \tilde{\mathbf{R}}^{-1} \bar{\mathbf{r}}_j^H \ . \tag{B.5}$$

The partial derivatives of $\tilde{\mathbf{R}}$ w.r.t. β are computed as follows

$$\frac{\partial \tilde{\mathbf{R}}}{\partial \sigma_\alpha^2} = \left[\bar{\mathbf{d}}_1^H \bar{\mathbf{d}}_1 + \bar{\mathbf{d}}_2^H \bar{\mathbf{d}}_2\right] \ ,$$

$$\frac{\partial \tilde{\mathbf{R}}}{\partial N_0} = \mathbf{I}_{L_0 Q} \ ,$$

$$\frac{\partial \tilde{\mathbf{R}}}{\partial \varepsilon} = \sigma_\alpha^2 \left[\frac{\partial}{\partial \varepsilon}\left[\bar{\mathbf{d}}_1^H \bar{\mathbf{d}}_1 + \bar{\mathbf{d}}_2^H \bar{\mathbf{d}}_2\right]\right] \ .$$

For $\beta = \sigma_\alpha^2$, (B.5) becomes

$$\frac{\partial \mathbf{\Lambda}}{\partial \sigma_\alpha^2} = 2\mathrm{tr}\left(\tilde{\mathbf{R}}^{-1}\left[\bar{\mathbf{d}}_1^H \bar{\mathbf{d}}_1 + \bar{\mathbf{d}}_2^H \bar{\mathbf{d}}_2\right]\right) - \sum_{j=1}^{2} \bar{\mathbf{r}}_j \tilde{\mathbf{R}}^{-1}\left[\bar{\mathbf{d}}_1^H \bar{\mathbf{d}}_1 + \bar{\mathbf{d}}_2^H \bar{\mathbf{d}}_2\right] \tilde{\mathbf{R}}^{-1} \bar{\mathbf{r}}_j^H \ . \tag{B.6}$$

To prove that the regularity condition is respected for this partial derivative, first the

expected value of the second term in (B.6) is computed as follows

$$E\left[\sum_{j=1}^{2}\bar{\mathbf{r}}_j\tilde{\mathbf{R}}^{-1}\left[\bar{\mathbf{d}}_1^H\bar{\mathbf{d}}_1+\bar{\mathbf{d}}_2^H\bar{\mathbf{d}}_2\right]\tilde{\mathbf{R}}^{-1}\bar{\mathbf{r}}_j^H\right]=E\left[\sum_{j=1}^{2}\mathrm{tr}\left(\tilde{\mathbf{R}}^{-1}\left[\bar{\mathbf{d}}_1^H\bar{\mathbf{d}}_1+\bar{\mathbf{d}}_2^H\bar{\mathbf{d}}_2\right]\tilde{\mathbf{R}}^{-1}\bar{\mathbf{r}}_j^H\bar{\mathbf{r}}_j\right)\right]$$

$$=\sum_{j=1}^{2}\mathrm{tr}\left(\tilde{\mathbf{R}}^{-1}\left[\bar{\mathbf{d}}_1^H\bar{\mathbf{d}}_1+\bar{\mathbf{d}}_2^H\bar{\mathbf{d}}_2\right]\tilde{\mathbf{R}}^{-1}E\left[\bar{\mathbf{r}}_j^H\bar{\mathbf{r}}_j\right]\right)=\sum_{j=1}^{2}\mathrm{tr}\left(\tilde{\mathbf{R}}^{-1}\left[\bar{\mathbf{d}}_1^H\bar{\mathbf{d}}_1+\bar{\mathbf{d}}_2^H\bar{\mathbf{d}}_2\right]\tilde{\mathbf{R}}^{-1}\tilde{\mathbf{R}}\right)$$

$$=2\mathrm{tr}\left(\tilde{\mathbf{R}}^{-1}\left[\bar{\mathbf{d}}_1^H\bar{\mathbf{d}}_1+\bar{\mathbf{d}}_2^H\bar{\mathbf{d}}_2\right]\right)\ .$$

Introducing this in (B.6) proves that $E\left[\partial\mathbf{\Lambda}/\partial\sigma_\alpha^2\right]=0$. The same procedure could be used to prove that the regularity condition is respected for all first order partial derivatives. For brevity, these derivations will be skipped. Starting from (B.5), the second order mix derivatives are computed in the general form

$$E\left[-\frac{\partial^2\mathbf{\Lambda}}{\partial\alpha\partial\beta}\right]=2\frac{\partial}{\partial\alpha}\mathrm{tr}\left(\tilde{\mathbf{R}}^{-1}\frac{\partial\tilde{\mathbf{R}}}{\partial\beta}\right)-\sum_{j=1}^{2}\mathrm{tr}\left(\frac{\partial\tilde{\mathbf{R}}^{-1}}{\partial\alpha}\frac{\partial\tilde{\mathbf{R}}}{\partial\beta}\tilde{\mathbf{R}}^{-1}E\left[\bar{\mathbf{r}}_j^H\bar{\mathbf{r}}_j\right]\right)$$

$$-\sum_{j=1}^{2}\mathrm{tr}\left(\tilde{\mathbf{R}}^{-1}\frac{\partial^2\tilde{\mathbf{R}}}{\partial\alpha\partial\beta}\tilde{\mathbf{R}}^{-1}E\left[\bar{\mathbf{r}}_j^H\bar{\mathbf{r}}_j\right]\right)-\sum_{j=1}^{2}\mathrm{tr}\left(\tilde{\mathbf{R}}^{-1}\frac{\partial\tilde{\mathbf{R}}}{\partial\beta}\frac{\partial\tilde{\mathbf{R}}^{-1}}{\partial\alpha}E\left[\bar{\mathbf{r}}_j^H\bar{\mathbf{r}}_j\right]\right)$$

$$=2\mathrm{tr}\left(\frac{\partial\tilde{\mathbf{R}}^{-1}}{\partial\alpha}\frac{\partial\tilde{\mathbf{R}}}{\partial\beta}+\tilde{\mathbf{R}}^{-1}\frac{\partial^2\tilde{\mathbf{R}}}{\partial\alpha\partial\beta}-\frac{\partial\tilde{\mathbf{R}}^{-1}}{\partial\alpha}\frac{\partial\tilde{\mathbf{R}}}{\partial\beta}-\tilde{\mathbf{R}}^{-1}\frac{\partial^2\tilde{\mathbf{R}}}{\partial\alpha\partial\beta}-\frac{\partial\tilde{\mathbf{R}}}{\partial\beta}\frac{\partial\tilde{\mathbf{R}}^{-1}}{\partial\alpha}\right)$$

Because the first four terms cancel each other

$$E\left[-\frac{\partial^2\mathbf{\Lambda}}{\partial\alpha\partial\beta}\right]=-2\mathrm{tr}\left(\frac{\partial\tilde{\mathbf{R}}}{\partial\beta}\frac{\partial\tilde{\mathbf{R}}^{-1}}{\partial\alpha}\right)=2\mathrm{tr}\left(\frac{\partial\tilde{\mathbf{R}}}{\partial\beta}\tilde{\mathbf{R}}^{-1}\frac{\partial\tilde{\mathbf{R}}}{\partial\alpha}\tilde{\mathbf{R}}^{-1}\right)\ .$$

Thus, the following general form is obtained

$$J_{\alpha\beta}=E\left[-\frac{\partial^2\mathbf{\Lambda}}{\partial\alpha\partial\beta}\right]=2\mathrm{tr}\left(\tilde{\mathbf{R}}^{-1}\frac{\partial\tilde{\mathbf{R}}}{\partial\alpha}\tilde{\mathbf{R}}^{-1}\frac{\partial\tilde{\mathbf{R}}}{\partial\beta}\right)\ . \tag{B.7}$$

The derivation of the block fading CRB for the main model introduced in Chapter III.C starts with the general model (A.1), where $\mathbf{h}$ is defined as in (3.19) for an associated matrix $\mathbf{H}$ defined by (3.7). By comparing (3.8), (3.9) and (A.1), it is

obtained that

$$\mathbf{\Psi}_{\varepsilon} = \begin{cases} \mathbf{I}_M \otimes \mathbf{A}_{\varepsilon}\mathbf{Z} & \text{DA case} \\ \mathbf{I}_M \otimes \mathbf{A}_{\varepsilon} & \text{NDA case} \end{cases} . \tag{B.8}$$

As explained in Appendix A, $E_{\mathbf{h}}[\mathbf{h}\mathbf{h}^H]$ is the identity matrix multiplied by σ_α^2 in the DA and NDA cases considered. Under the same Gaussian assumptions as for the previous model, $\mathbf{r}$ is a Gaussian vector with zero-mean and covariance matrix

$$E[\mathbf{r}\mathbf{r}^H] = E[(\xi\mathbf{\Psi}_{\varepsilon}\mathbf{h} + \mathbf{n})(\xi\mathbf{h}^H\mathbf{\Psi}_{\varepsilon}^H + \mathbf{n}^H)] = \xi^2 E[\mathbf{\Psi}_{\varepsilon}\mathbf{h}\mathbf{h}^H\mathbf{\Psi}_{\varepsilon}^H] + E[\mathbf{n}\mathbf{n}^H] ,$$

$$\tilde{\mathbf{R}} = E[\mathbf{r}\mathbf{r}^H] = \xi^2\mathbf{\Psi}_{\varepsilon}E[\mathbf{h}\mathbf{h}^H]\mathbf{\Psi}_{\varepsilon}^H + N_0\mathbf{I}_{ML_0Q} = \xi^2\sigma_\alpha^2\mathbf{\Psi}_{\varepsilon}\mathbf{\Psi}_{\varepsilon}^H + N_0\mathbf{I}_{ML_0Q} . \tag{B.9}$$

Taking (B.8) into account, the particular versions of $\tilde{\mathbf{R}}$ become

$$\tilde{\mathbf{R}} = \mathbf{I}_M \otimes \left(\xi^2\sigma_\alpha^2\mathbf{A}_{\varepsilon}\mathbf{Z}\mathbf{Z}^H\mathbf{A}_{\varepsilon}^H + N_0\mathbf{I}_{L_0Q}\right) \quad \text{DA case} , \tag{B.10}$$

$$\tilde{\mathbf{R}} = \mathbf{I}_M \otimes \left(\xi^2\sigma_\alpha^2\mathbf{A}_{\varepsilon}\mathbf{A}_{\varepsilon}^H + N_0\mathbf{I}_{L_0Q}\right) \quad \text{NDA case} . \tag{B.11}$$

The pdf of the Gaussian vector $\mathbf{r}$, having mean zero and a covariance matrix given by (B.9), has the form

$$p_{\mathbf{r}}(\mathbf{r}) = \frac{1}{\pi^{ML_0Q} \mid \tilde{\mathbf{R}} \mid} \exp\left[-\mathbf{r}^H\tilde{\mathbf{R}}^{-1}\mathbf{r}\right] , \tag{B.12}$$

and the log-likelihood function becomes

$$\mathbf{\Lambda} = \ln p_{\mathbf{r}}(\mathbf{r}) = -ML_0Q\ln(\pi) - \ln \mid \tilde{\mathbf{R}} \mid -\mathbf{r}^H\tilde{\mathbf{R}}^{-1}\mathbf{r} .$$

Similarly to the computations for the previous model, it is obtained that

$$J_{\alpha\beta} = E\left[-\frac{\partial^2\mathbf{\Lambda}}{\partial\alpha\partial\beta}\right] = \mathrm{tr}\left(\tilde{\mathbf{R}}^{-1}\frac{\partial\tilde{\mathbf{R}}}{\partial\alpha}\tilde{\mathbf{R}}^{-1}\frac{\partial\tilde{\mathbf{R}}}{\partial\beta}\right) . \tag{B.13}$$

The partial derivatives of $\tilde{\mathbf{R}}$ w.r.t. β are computed as follows

$$\frac{\partial\tilde{\mathbf{R}}}{\partial\sigma_\alpha^2} = \xi^2\mathbf{\Psi}_{\varepsilon}\mathbf{\Psi}_{\varepsilon}^H ,$$

$$\frac{\partial \tilde{\mathbf{R}}}{\partial N_0} = \mathbf{I}_{ML_0Q} \ ,$$

$$\frac{\partial \tilde{\mathbf{R}}}{\partial \varepsilon} = \xi^2 \sigma_\alpha^2 \left[\boldsymbol{\Upsilon}_\varepsilon \boldsymbol{\Psi}_\varepsilon^H + \boldsymbol{\Psi}_\varepsilon \boldsymbol{\Upsilon}_\varepsilon^H \right] \ ,$$

where

$$\boldsymbol{\Upsilon}_\varepsilon = \frac{\partial \boldsymbol{\Psi}_\varepsilon}{\partial \varepsilon} = \begin{cases} \mathbf{I}_M \otimes \mathbf{D}_\varepsilon \mathbf{Z} & \text{DA case} \\ \mathbf{I}_M \otimes \mathbf{D}_\varepsilon & \text{NDA case} \end{cases} .$$

APPENDIX C

DERIVATION OF THE BB

In this Appendix, a general form for the BB, valid for both the block and the TV fading will be derived. The BB has been defined in (2.3) as

$$\mathrm{BB} = \sup_{\eta} \frac{\eta^2}{\int \frac{p(\mathbf{r};\varepsilon+\eta)^2}{p(\mathbf{r};\varepsilon)} d\mathbf{r} - 1} \ . \tag{C.1}$$

For the eigendecomposition-based model introduced in Chapter III.B, (B.3) and (B.4) lead to the following pdf

$$p_{\bar{\mathbf{r}}}(\bar{\mathbf{r}}) = \frac{1}{\pi^{2L_0Q} \mid \tilde{\mathbf{R}} \mid^2} \exp\left[-\bar{\mathbf{r}}_1 \tilde{\mathbf{R}}^{-1} \bar{\mathbf{r}}_1^H - \bar{\mathbf{r}}_2 \tilde{\mathbf{R}}^{-1} \bar{\mathbf{r}}_2^H \right] \ .$$

The dependence on the time delay is stressed out as

$$p(\bar{\mathbf{r}};\varepsilon) = \frac{1}{\pi^{2L_0Q} \mid \tilde{\mathbf{R}}_{\varepsilon} \mid^2} \exp\left[-\bar{\mathbf{r}}_1 \tilde{\mathbf{R}}_{\varepsilon}^{-1} \bar{\mathbf{r}}_1^H - \bar{\mathbf{r}}_2 \tilde{\mathbf{R}}_{\varepsilon}^{-1} \bar{\mathbf{r}}_2^H \right] \ , \tag{C.2}$$

$$p(\bar{\mathbf{r}};\varepsilon+\eta) = \frac{1}{\pi^{2L_0Q} \mid \tilde{\mathbf{R}}_{\varepsilon+\eta} \mid^2} \exp\left[-\bar{\mathbf{r}}_1 \tilde{\mathbf{R}}_{\varepsilon+\eta}^{-1} \bar{\mathbf{r}}_1^H - \bar{\mathbf{r}}_2 \tilde{\mathbf{R}}_{\varepsilon+\eta}^{-1} \bar{\mathbf{r}}_2^H \right] \ . \tag{C.3}$$

Squaring (C.3) leads to

$$p^2(\bar{\mathbf{r}};\varepsilon+\eta) = \frac{1}{\pi^{4L_0Q} \mid \tilde{\mathbf{R}}_{\varepsilon+\eta} \mid^4} \exp\left[-2\bar{\mathbf{r}}_1 \tilde{\mathbf{R}}_{\varepsilon+\eta}^{-1} \bar{\mathbf{r}}_1^H - 2\bar{\mathbf{r}}_2 \tilde{\mathbf{R}}_{\varepsilon+\eta}^{-1} \bar{\mathbf{r}}_2^H \right] \ . \tag{C.4}$$

The integral part in (C.1) can be computed by plugging (C.2) and (C.4)

$$\int \frac{p^2(\mathbf{r};\varepsilon+\eta)}{p(\mathbf{r};\varepsilon)} d\mathbf{r}$$

$$= \int \frac{|\tilde{\mathbf{R}}_{\varepsilon}|^2}{|\tilde{\mathbf{R}}_{\varepsilon+\eta}|^4} \frac{1}{\pi^{2L_0Q}} \exp\left[-\bar{\mathbf{r}}_1 (2\tilde{\mathbf{R}}_{\varepsilon+\eta}^{-1} - \tilde{\mathbf{R}}_{\varepsilon}^{-1}) \bar{\mathbf{r}}_1^H - \bar{\mathbf{r}}_2 (2\tilde{\mathbf{R}}_{\varepsilon+\eta}^{-1} - \tilde{\mathbf{R}}_{\varepsilon}^{-1}) \bar{\mathbf{r}}_2^H \right] d\mathbf{r}$$

$$= \int\int \frac{|\tilde{\mathbf{R}}_{\varepsilon}|^2}{|\tilde{\mathbf{R}}_{\varepsilon+\eta}|^4} \frac{1}{\pi^{2L_0Q}} \exp\left[-\bar{\mathbf{r}}_1 (2\tilde{\mathbf{R}}_{\varepsilon+\eta}^{-1} - \tilde{\mathbf{R}}_{\varepsilon}^{-1}) \bar{\mathbf{r}}_1^H - \bar{\mathbf{r}}_2 (2\tilde{\mathbf{R}}_{\varepsilon+\eta}^{-1} - \tilde{\mathbf{R}}_{\varepsilon}^{-1}) \bar{\mathbf{r}}_2^H \right] d\mathbf{r}_1 d\mathbf{r}_2$$

$$= \frac{|\tilde{\mathbf{R}}_{\varepsilon}|^2}{|\tilde{\mathbf{R}}_{\varepsilon+\eta}|^4 |2\tilde{\mathbf{R}}_{\varepsilon+\eta}^{-1} - \tilde{\mathbf{R}}_{\varepsilon}^{-1}|^2} = \frac{1}{|\tilde{\mathbf{R}}_{\varepsilon+\eta}\tilde{\mathbf{R}}_{\varepsilon}^{-1}|^2 |2\mathbf{I} - \tilde{\mathbf{R}}_{\varepsilon+\eta}\tilde{\mathbf{R}}_{\varepsilon}^{-1}|^2} ,$$

where the last equality has been obtained using standard determinant properties. Plugging this integral into (C.1), the form of the BB for the eigen-decomposition model becomes

$$\mathrm{BB} = \sup_{\eta} \frac{\eta^2}{\frac{1}{|\tilde{\mathbf{R}}_{\varepsilon+\eta}\tilde{\mathbf{R}}_{\varepsilon}^{-1}|^2 |2\mathbf{I} - \tilde{\mathbf{R}}_{\varepsilon+\eta}\tilde{\mathbf{R}}_{\varepsilon}^{-1}|^2} - 1} = \sup_{\eta} \frac{\eta^2 \mid \tilde{\mathbf{R}}_{\varepsilon+\eta}\tilde{\mathbf{R}}_{\varepsilon}^{-1} \mid^2 \mid 2\mathbf{I} - \tilde{\mathbf{R}}_{\varepsilon+\eta}\tilde{\mathbf{R}}_{\varepsilon}^{-1} \mid^2}{1 - \mid \tilde{\mathbf{R}}_{\varepsilon+\eta}\tilde{\mathbf{R}}_{\varepsilon}^{-1} \mid^2 \mid 2\mathbf{I} - \tilde{\mathbf{R}}_{\varepsilon+\eta}\tilde{\mathbf{R}}_{\varepsilon}^{-1} \mid^2} \tag{C.5}$$

In order to be able to apply this formula directly, the matrix $\tilde{\mathbf{R}}_{\varepsilon}$ has to be invertible. The nondetermination obtained in the case $\eta \to 0$ can be solved with the l'Hopital rule. In limit towards zero, the BB becomes the CRB. The same method is used to determine the BB for the main model defined in Chapter III.C. The pdf $p_{\mathbf{r}}(\mathbf{r})$ is given by (D.5)

$$p_{\mathbf{r}}(\mathbf{r}) = \frac{1}{\pi^{ML_0Q} \mid \tilde{\mathbf{R}} \mid} \exp\left[-\mathbf{r}^H \tilde{\mathbf{R}}^{-1} \mathbf{r}\right] .$$

Expressing the dependence on the time delay

$$p(\mathbf{r}; \varepsilon) = \frac{1}{\pi^{ML_0Q} \mid \tilde{\mathbf{R}}_{\varepsilon} \mid} \exp\left[-\mathbf{r}^H \tilde{\mathbf{R}}_{\varepsilon}^{-1} \mathbf{r}\right] , \tag{C.6}$$

$$p(\mathbf{r}; \varepsilon + \eta) = \frac{1}{\pi^{ML_0Q} \mid \tilde{\mathbf{R}}_{\varepsilon+\eta} \mid} \exp\left[-\mathbf{r}^H \tilde{\mathbf{R}}_{\varepsilon+\eta}^{-1} \mathbf{r}\right] . \tag{C.7}$$

After squaring (C.7) and introducing it, together with (C.6), in the integral from (C.1), following the same procedure as for the previous model, it is obtained

$$\int \frac{p^2(\mathbf{r}; \varepsilon + \eta)}{p(\mathbf{r}; \varepsilon)} d\mathbf{r} = \frac{\mid \tilde{\mathbf{R}}_{\varepsilon} \mid}{\mid \tilde{\mathbf{R}}_{\varepsilon+\eta} \mid^2 \mid 2\tilde{\mathbf{R}}_{\varepsilon+\eta}^{-1} - \tilde{\mathbf{R}}_{\varepsilon}^{-1} \mid} .$$

Then, the BB for the main model becomes

$$\mathrm{BB} = \sup_{\eta} \frac{\eta^2 \mid \tilde{\mathbf{R}}_{\varepsilon+\eta}\tilde{\mathbf{R}}_{\varepsilon}^{-1} \mid\mid 2\mathbf{I} - \tilde{\mathbf{R}}_{\varepsilon+\eta}\tilde{\mathbf{R}}_{\varepsilon}^{-1} \mid}{1 - \mid \tilde{\mathbf{R}}_{\varepsilon+\eta}\tilde{\mathbf{R}}_{\varepsilon}^{-1} \mid\mid 2\mathbf{I} - \tilde{\mathbf{R}}_{\varepsilon+\eta}\tilde{\mathbf{R}}_{\varepsilon}^{-1} \mid} . \tag{C.8}$$

APPENDIX D

DERIVATION OF THE CRB FOR TV FADING

With a few changes, the procedure to compute the CRB for TV fading follows the same guidelines presented in Appendix B. For the eigendecomposition-based model introduced in Chapter III.B, the fading coefficients h_{ij} can be expressed as L_0Q-length vectors, independent from one pair of antennas to another, following a circularly complex Gaussian distribution, with zero-mean and a covariance matrix $\mathbf{R}_\alpha$, whose elements are given by (see (3.3))

$$(\mathbf{R}_\alpha)_{ij} = \sigma_\alpha^2 J_0(2\pi f_d \mid i-j \mid T_s) \ . \tag{D.1}$$

An expression similar to (B.1) can be derived

$$\bar{\mathbf{r}}_j = \bar{\mathbf{d}}_1 * \bar{\mathbf{h}}_{1j} + \bar{\mathbf{d}}_2 * \bar{\mathbf{h}}_{2j} + \bar{\mathbf{n}}_j \ ,$$

where $\bar{\mathbf{d}}_i = \left[\begin{array}{cccc} d_i(0) & d_i(T_s) & \cdots & d_i((L_0Q-1)T_s) \end{array} \right]$, $\bar{\mathbf{n}}_j \sim N(0, N_0\mathbf{I}_{L_0Q})$, $(i,j=1,2)$, $\bar{\mathbf{h}}_{ij} \sim N(0,\mathbf{R}_\alpha)$, and $*$ denotes the point-to-point vectors multiplication. Therefore, $\bar{\mathbf{r}}_j$ is also a Gaussian vector with zero-mean and covariance

$$\tilde{\mathbf{R}} = E\left[\bar{\mathbf{r}}_j^H \bar{\mathbf{r}}_j\right] = \mathbf{R}_\alpha \left[\bar{\mathbf{d}}_1^H \bar{\mathbf{d}}_1 + \bar{\mathbf{d}}_2^H \bar{\mathbf{d}}_2\right] + N_0\mathbf{I}_{L_0Q} \ . \tag{D.2}$$

The first term of the sumation in (D.2) has been obtained assuming independence between the vectors $\bar{\mathbf{d}}_i$ and $\bar{\mathbf{h}}_{ij}$. The derivations from (B.3) to (B.5) remain valid. The partial derivatives $\partial\tilde{\mathbf{R}}/\partial\beta$, where $\beta = \sigma_\alpha^2, f_d, N_0, \varepsilon$, can be determined as

$$\frac{\partial\tilde{\mathbf{R}}}{\partial\sigma_\alpha^2} = \frac{\partial\mathbf{R}_\alpha}{\partial\sigma_\alpha^2}\left[\bar{\mathbf{d}}_1^H \bar{\mathbf{d}}_1 + \bar{\mathbf{d}}_2^H \bar{\mathbf{d}}_2\right] \ ,$$

$$\frac{\partial\tilde{\mathbf{R}}}{\partial f_d} = \frac{\partial\mathbf{R}_\alpha}{\partial f_d}\left[\bar{\mathbf{d}}_1^H \bar{\mathbf{d}}_1 + \bar{\mathbf{d}}_2^H \bar{\mathbf{d}}_2\right] \ ,$$

$$\frac{\partial \tilde{\mathbf{R}}}{\partial N_0} = \mathbf{I}_{L_0 Q} \ ,$$

$$\frac{\partial \tilde{\mathbf{R}}}{\partial \varepsilon} = \mathbf{R}_\alpha \left[\frac{\partial}{\partial \varepsilon} \left[\bar{\mathbf{d}}_1^H \bar{\mathbf{d}}_1 + \bar{\mathbf{d}}_2^H \bar{\mathbf{d}}_2 \right] \right] \ .$$

Introducing the partial derivatives of $\tilde{\mathbf{R}}$ in (B.5), it is obtained that (B.7) is still formally valid in the TV case.

The derivation of the CRB for the main model introduced in Chapter III.C starts with the general model (3.18), where $\mathbf{h}$ is defined as in (3.19) for an associated matrix $\mathbf{H}$ defined by (3.10). Matrix $\mathbf{\Psi}_\varepsilon$ has been defined by (3.16) in the DA case and by (3.17) in the NDA case.

Vector $\mathbf{h}$ is a concatenation of Gaussian vectors in both cases. After studying the structure of the extended matrix $\mathbf{H}$ defined by (3.10), the covariance matrix of $\mathbf{h}$ turns out to be

$$\mathbf{\Gamma_h} = E[\mathbf{h}\mathbf{h}^H] = \begin{cases} \mathbf{I}_M \otimes \mathbf{R}_\alpha \otimes \mathbf{I}_N & \text{DA case} \\ \mathbf{I}_M \otimes \mathbf{R}_\alpha \otimes \mathbf{I}_{L_0+2L_g} & \text{NDA case} \end{cases} , \qquad \text{(D.3)}$$

where $\mathbf{R}_\alpha$ is the $L_0 Q \times L_0 Q$ autocorrelation matrix of the signal samples, defined according to (D.1). Under Gaussian assumptions, $\mathbf{r}$ is also a Gaussian vector with zero-mean and covariance matrix

$$E[\mathbf{r}\mathbf{r}^H] = E[(\xi \mathbf{\Psi}_\varepsilon \mathbf{h} + \mathbf{n})(\xi \mathbf{h}^H \mathbf{\Psi}_\varepsilon^H + \mathbf{n}^H)] = \xi^2 E[\mathbf{\Psi}_\varepsilon \mathbf{h}\mathbf{h}^H \mathbf{\Psi}_\varepsilon^H] + E[\mathbf{n}\mathbf{n}^H] \ ,$$

$$\tilde{\mathbf{R}} = E[\mathbf{r}\mathbf{r}^H] = \xi^2 \mathbf{\Psi}_\varepsilon \mathbf{\Gamma_h} \mathbf{\Psi}_\varepsilon^H + N_0 \mathbf{I}_{M L_0 Q} \ . \qquad \text{(D.4)}$$

Similarly with the block fading case, it is determined

$$p_{\mathbf{r}}(\mathbf{r}) = \frac{1}{\pi^{M L_0 Q} \mid \tilde{\mathbf{R}} \mid} \exp\left[-\mathbf{r}^H \tilde{\mathbf{R}}^{-1} \mathbf{r} \right] \ , \qquad \text{(D.5)}$$

$$\mathbf{\Lambda} = \ln p_{\mathbf{r}}(\mathbf{r}) = -M L_0 Q \ln(\pi) - \ln \mid \tilde{\mathbf{R}} \mid - \mathbf{r}^H \tilde{\mathbf{R}}^{-1} \mathbf{r} \ ,$$

$$J_{\alpha\beta} = E\left[-\frac{\partial^2 \mathbf{\Lambda}}{\partial\alpha\partial\beta}\right] = \mathrm{tr}\left(\tilde{\mathbf{R}}^{-1}\frac{\partial\tilde{\mathbf{R}}}{\partial\alpha}\tilde{\mathbf{R}}^{-1}\frac{\partial\tilde{\mathbf{R}}}{\partial\beta}\right) \ . \tag{D.6}$$

The partial derivatives $\partial\tilde{\mathbf{R}}/\partial\beta$ can be determined as

$$\frac{\partial\tilde{\mathbf{R}}}{\partial\sigma_\alpha^2} = \xi^2 \mathbf{\Psi}_\varepsilon \frac{\partial \mathbf{\Gamma_h}}{\partial\sigma_\alpha^2}\mathbf{\Psi}_\varepsilon^H \ , \tag{D.7}$$

$$\frac{\partial\tilde{\mathbf{R}}}{\partial f_d} = \xi^2 \mathbf{\Psi}_\varepsilon \frac{\partial \mathbf{\Gamma_h}}{\partial f_d}\mathbf{\Psi}_\varepsilon^H \ , \tag{D.8}$$

$$\frac{\partial\tilde{\mathbf{R}}}{\partial N_0} = \mathbf{I}_{ML_0Q} \ , \tag{D.9}$$

$$\frac{\partial\tilde{\mathbf{R}}}{\partial\varepsilon} = \xi^2\left[\mathbf{\Upsilon}_\varepsilon \mathbf{\Gamma_h}\mathbf{\Psi}_\varepsilon^H + \mathbf{\Psi}_\varepsilon\mathbf{\Gamma_h}\mathbf{\Upsilon}_\varepsilon^H\right] \ , \tag{D.10}$$

where $\mathbf{\Upsilon}_\varepsilon$ has been defined by (4.2) in the DA case and by (4.4) in the NDA case.

Using (D.3), the partial derivatives of $\mathbf{\Gamma_h}$ are determined to be

$$\frac{\partial \mathbf{\Gamma_h}}{\partial\sigma_\alpha^2} = \begin{cases} \mathbf{I}_M \otimes \frac{\partial \mathbf{R}_\alpha}{\partial\sigma_\alpha^2} \otimes \mathbf{I}_N & \text{DA case} \\ \mathbf{I}_M \otimes \frac{\partial \mathbf{R}_\alpha}{\partial\sigma_\alpha^2} \otimes \mathbf{I}_{L_0+2L_g} & \text{NDA case} \end{cases} \ , \tag{D.11}$$

$$\frac{\partial \mathbf{\Gamma_h}}{\partial f_d} = \begin{cases} \mathbf{I}_M \otimes \frac{\partial \mathbf{R}_\alpha}{\partial f_d} \otimes \mathbf{I}_N & \text{DA case} \\ \mathbf{I}_M \otimes \frac{\partial \mathbf{R}_\alpha}{\partial f_d} \otimes \mathbf{I}_{L_0+2L_g} & \text{NDA case} \end{cases} \ . \tag{D.12}$$

APPENDIX E

DERIVATION OF THE ACRB FOR TV FADING

The derivation of the ACRB for low-SNR starts with the general autocorrelation matrix introduced as (D.4) and with the general formula for the FIM elements, introduced as (D.6)

$$\tilde{\mathbf{R}} = E[\mathbf{r}\mathbf{r}^H] = \xi^2 \mathbf{\Psi}_\varepsilon \mathbf{\Gamma}_\mathbf{h} \mathbf{\Psi}_\varepsilon^H + N_0 \mathbf{I}_{ML_0Q} \ , \tag{E.1}$$

$$J_{\alpha\beta} = E\left[-\frac{\partial^2 \mathbf{\Lambda}}{\partial\alpha\partial\beta}\right] = \mathrm{tr}\left(\tilde{\mathbf{R}}^{-1}\frac{\partial\tilde{\mathbf{R}}}{\partial\alpha}\tilde{\mathbf{R}}^{-1}\frac{\partial\tilde{\mathbf{R}}}{\partial\beta}\right) \ . \tag{E.2}$$

The partial first order derivatives have been computed in (D.7)-(D.10) and they do not depend on N_0.

Formula (E.1) can be re-expressed as

$$\tilde{\mathbf{R}} = N_0\left(\mathbf{I}_{ML_0Q} + \frac{\xi^2 \mathbf{\Psi}_\varepsilon \mathbf{\Gamma}_\mathbf{h} \mathbf{\Psi}_\varepsilon^H}{N_0}\right) \ .$$

For a high constant ρ and an arbitrary matrix $\mathbf{K}$ the following approximation can be used [37]

$$\left(\mathbf{I} + \frac{\mathbf{K}}{\rho}\right)^{-1} \simeq \mathbf{I} - \frac{\mathbf{K}}{\rho} \ .$$

Therefore, at low-SNR

$$\tilde{\mathbf{R}}^{-1} \simeq (N_0)^{-1}\left(\mathbf{I}_{ML_0Q} - \frac{\xi^2 \mathbf{\Psi}_\varepsilon \mathbf{\Gamma}_\mathbf{h} \mathbf{\Psi}_\varepsilon^H}{N_0}\right) \ . \tag{E.3}$$

Because the partial derivatives do not depend on N_0, the following general form is obtained by introducing (E.3) in (E.2)

$$J_{\alpha\beta} = \mathrm{tr}\left((N_0)^{-2} \left(\mathbf{I}_{ML_0Q} - \frac{\xi^2 \boldsymbol{\Psi}_\varepsilon \boldsymbol{\Gamma}_\mathbf{h} \boldsymbol{\Psi}_\varepsilon^H}{N_0} \right) \frac{\partial \tilde{\mathbf{R}}}{\partial \alpha} \left(\mathbf{I}_{ML_0Q} - \frac{\xi^2 \boldsymbol{\Psi}_\varepsilon \boldsymbol{\Gamma}_\mathbf{h} \boldsymbol{\Psi}_\varepsilon^H}{N_0} \right) \frac{\partial \tilde{\mathbf{R}}}{\partial \beta} \right) . \tag{E.4}$$

When N_0 is large, (E.4) can be approximated with

$$J_{\alpha\beta} = \mathrm{tr}\left((N_0)^{-2} \frac{\partial \tilde{\mathbf{R}}}{\partial \alpha} \frac{\partial \tilde{\mathbf{R}}}{\partial \beta} \right) = (N_0)^{-2} \mathrm{tr}\left(\frac{\partial \tilde{\mathbf{R}}}{\partial \alpha} \frac{\partial \tilde{\mathbf{R}}}{\partial \beta} \right) .$$

All the elements of the FIM are obtained as the product between a term independent on N_0 and a term dependent on the second power of N_0. Therefore the low-SNR ACRB can be also expressed as the product between a term independent on SNR and N_0^2.

The independent part of the elements of the FIM is computed as follows.

The elements of the sub-matrix $\mathbf{A}$ in the FIM (see (4.7)) are

$$J_{\sigma_\alpha^2 \sigma_\alpha^2} = \xi^4 \mathrm{tr}\left(\boldsymbol{\Psi}_\varepsilon \frac{\partial \boldsymbol{\Gamma}_\mathbf{h}}{\partial \sigma_\alpha^2} \boldsymbol{\Psi}_\varepsilon^H \boldsymbol{\Psi}_\varepsilon \frac{\partial \boldsymbol{\Gamma}_\mathbf{h}}{\partial \sigma_\alpha^2} \boldsymbol{\Psi}_\varepsilon^H \right) ,$$

$$J_{\sigma_\alpha^2 f_d} = \xi^4 \mathrm{tr}\left(\boldsymbol{\Psi}_\varepsilon \frac{\partial \boldsymbol{\Gamma}_\mathbf{h}}{\partial \sigma_\alpha^2} \boldsymbol{\Psi}_\varepsilon^H \boldsymbol{\Psi}_\varepsilon \frac{\partial \boldsymbol{\Gamma}_\mathbf{h}}{\partial f_d} \boldsymbol{\Psi}_\varepsilon^H \right) ,$$

$$J_{\sigma_\alpha^2 N_0} = \xi^2 \mathrm{tr}\left(\boldsymbol{\Psi}_\varepsilon \frac{\partial \boldsymbol{\Gamma}_\mathbf{h}}{\partial \sigma_\alpha^2} \boldsymbol{\Psi}_\varepsilon^H \right) ,$$

$$J_{f_d f_d} = \xi^4 \mathrm{tr}\left(\boldsymbol{\Psi}_\varepsilon \frac{\partial \boldsymbol{\Gamma}_\mathbf{h}}{\partial f_d} \boldsymbol{\Psi}_\varepsilon^H \boldsymbol{\Psi}_\varepsilon \frac{\partial \boldsymbol{\Gamma}_\mathbf{h}}{\partial f_d} \boldsymbol{\Psi}_\varepsilon^H \right) ,$$

$$J_{f_d N_0} = \xi^2 \mathrm{tr}\left(\boldsymbol{\Psi}_\varepsilon \frac{\partial \boldsymbol{\Gamma}_\mathbf{h}}{\partial f_d} \boldsymbol{\Psi}_\varepsilon^H \right) ,$$

$$J_{N_0 N_0} = ML_0Q .$$

The elements of the sub-matrix $\mathbf{B}$ (see (4.8)) are

$$J_{\sigma_\alpha^2 \varepsilon} = \xi^4 \mathrm{tr}\left(\left(\boldsymbol{\Psi}_\varepsilon \frac{\partial \boldsymbol{\Gamma}_\mathbf{h}}{\partial \sigma_\alpha^2} \boldsymbol{\Psi}_\varepsilon^H \right) \left[\boldsymbol{\Upsilon}_\varepsilon \boldsymbol{\Gamma}_\mathbf{h} \boldsymbol{\Psi}_\varepsilon^H + \boldsymbol{\Psi}_\varepsilon \boldsymbol{\Gamma}_\mathbf{h} \boldsymbol{\Upsilon}_\varepsilon^H \right] \right) ,$$

$$J_{f_d \varepsilon} = \xi^4 \mathrm{tr}\left(\left(\boldsymbol{\Psi}_\varepsilon \frac{\partial \boldsymbol{\Gamma}_\mathbf{h}}{\partial f_d} \boldsymbol{\Psi}_\varepsilon^H \right) \left[\boldsymbol{\Upsilon}_\varepsilon \boldsymbol{\Gamma}_\mathbf{h} \boldsymbol{\Psi}_\varepsilon^H + \boldsymbol{\Psi}_\varepsilon \boldsymbol{\Gamma}_\mathbf{h} \boldsymbol{\Upsilon}_\varepsilon^H \right] \right) ,$$

$$J_{N_0\varepsilon} = \xi^2 \mathrm{tr}\left(\left[\boldsymbol{\Upsilon}_\varepsilon \boldsymbol{\Gamma}_{\mathbf{h}} \boldsymbol{\Psi}_\varepsilon^H + \boldsymbol{\Psi}_\varepsilon \boldsymbol{\Gamma}_{\mathbf{h}} \boldsymbol{\Upsilon}_\varepsilon^H\right]\right) \ .$$

Finally, the element of the sub-matrix $\mathbf{C}$ (see (4.9)) is

$$J_{\varepsilon\varepsilon} = \xi^4 \mathrm{tr}\left(\left[\boldsymbol{\Upsilon}_\varepsilon \boldsymbol{\Gamma}_{\mathbf{h}} \boldsymbol{\Psi}_\varepsilon^H + \boldsymbol{\Psi}_\varepsilon \boldsymbol{\Gamma}_{\mathbf{h}} \boldsymbol{\Upsilon}_\varepsilon^H\right]\left[\boldsymbol{\Upsilon}_\varepsilon \boldsymbol{\Gamma}_{\mathbf{h}} \boldsymbol{\Psi}_\varepsilon^H + \boldsymbol{\Psi}_\varepsilon \boldsymbol{\Gamma}_{\mathbf{h}} \boldsymbol{\Upsilon}_\varepsilon^H\right]\right) \ .$$

The derivation of the ACRB for high-SNR starts with the same general auto-correlation matrix introduced as (D.4) and with the general formula for the FIM elements, introduced as (D.6).

$$\tilde{\mathbf{R}} = E[\mathbf{r}\mathbf{r}^H] = \xi^2 \boldsymbol{\Psi}_\varepsilon \boldsymbol{\Gamma}_{\mathbf{h}} \boldsymbol{\Psi}_\varepsilon^H + N_0 \mathbf{I}_{ML_0Q} \ ,$$

$$J_{\alpha\beta} = E\left[-\frac{\partial^2 \boldsymbol{\Lambda}}{\partial\alpha\partial\beta}\right] = \mathrm{tr}\left(\tilde{\mathbf{R}}^{-1}\frac{\partial\tilde{\mathbf{R}}}{\partial\alpha}\tilde{\mathbf{R}}^{-1}\frac{\partial\tilde{\mathbf{R}}}{\partial\beta}\right) \ .$$

The partial first order derivatives have been computed in (D.7)-(D.10). At high SNR the inverse of the matrix $\tilde{\mathbf{R}}$ can be computed by ignoring the noise [38]

$$\tilde{\mathbf{R}}^{-1} = \frac{1}{\xi^2}\left(\boldsymbol{\Psi}_\varepsilon \boldsymbol{\Gamma}_{\mathbf{h}} \boldsymbol{\Psi}_\varepsilon^H\right)^{-1} \ .$$

The elements of the sub-matrix $\mathbf{A}$ in the FIM (see (4.7)) are

$$J_{\sigma_\alpha^2\sigma_\alpha^2} = \tfrac{1}{(\sigma_\alpha^2)^2} ML_0Q \ ,$$

$$J_{\sigma_\alpha^2 f_d} = \tfrac{1}{\sigma_\alpha^2}\mathrm{tr}\left(\left(\boldsymbol{\Psi}_\varepsilon \boldsymbol{\Gamma}_{\mathbf{h}} \boldsymbol{\Psi}_\varepsilon^H\right)^{-1}\left(\boldsymbol{\Psi}_\varepsilon \tfrac{\partial \boldsymbol{\Gamma}_{\mathbf{h}}}{\partial f_d} \boldsymbol{\Psi}_\varepsilon^H\right)\right) \ ,$$

$$J_{\sigma_\alpha^2 N_0} = \tfrac{1}{\sigma_\alpha^2}\tfrac{1}{\xi^2}\mathrm{tr}\left(\left(\boldsymbol{\Psi}_\varepsilon \boldsymbol{\Gamma}_{\mathbf{h}} \boldsymbol{\Psi}_\varepsilon^H\right)^{-1}\right) \ ,$$

$$J_{f_d f_d} = \mathrm{tr}\left(\left(\boldsymbol{\Psi}_\varepsilon \boldsymbol{\Gamma}_{\mathbf{h}} \boldsymbol{\Psi}_\varepsilon^H\right)^{-1}\left(\boldsymbol{\Psi}_\varepsilon \tfrac{\partial \boldsymbol{\Gamma}_{\mathbf{h}}}{\partial f_d} \boldsymbol{\Psi}_\varepsilon^H\right)\left(\boldsymbol{\Psi}_\varepsilon \boldsymbol{\Gamma}_{\mathbf{h}} \boldsymbol{\Psi}_\varepsilon^H\right)^{-1}\left(\boldsymbol{\Psi}_\varepsilon \tfrac{\partial \boldsymbol{\Gamma}_{\mathbf{h}}}{\partial f_d} \boldsymbol{\Psi}_\varepsilon^H\right)\right) \ ,$$

$$J_{f_d N_0} = \tfrac{1}{\xi^2}\mathrm{tr}\left(\left(\boldsymbol{\Psi}_\varepsilon \boldsymbol{\Gamma}_{\mathbf{h}} \boldsymbol{\Psi}_\varepsilon^H\right)^{-1}\left(\boldsymbol{\Psi}_\varepsilon \tfrac{\partial \boldsymbol{\Gamma}_{\mathbf{h}}}{\partial f_d} \boldsymbol{\Psi}_\varepsilon^H\right)\left(\boldsymbol{\Psi}_\varepsilon \boldsymbol{\Gamma}_{\mathbf{h}} \boldsymbol{\Psi}_\varepsilon^H\right)^{-1}\right) \ ,$$

$$J_{N_0 N_0} = \tfrac{1}{\xi^4}\mathrm{tr}\left(\left(\boldsymbol{\Psi}_\varepsilon \boldsymbol{\Gamma}_{\mathbf{h}} \boldsymbol{\Psi}_\varepsilon^H\right)^{-1}\left(\boldsymbol{\Psi}_\varepsilon \boldsymbol{\Gamma}_{\mathbf{h}} \boldsymbol{\Psi}_\varepsilon^H\right)^{-1}\right) \ .$$

The elements of the sub-matrix $\mathbf{B}$ (see (4.8)) are

$$J_{\sigma_\alpha^2\varepsilon} = \tfrac{1}{\sigma_\alpha^2}\mathrm{tr}\left(\left(\boldsymbol{\Psi}_\varepsilon\boldsymbol{\Gamma}_{\mathbf{h}}\boldsymbol{\Psi}_\varepsilon^H\right)^{-1}\left(\boldsymbol{\Upsilon}_\varepsilon\boldsymbol{\Gamma}_{\mathbf{h}}\boldsymbol{\Psi}_\varepsilon^H + \boldsymbol{\Psi}_\varepsilon\boldsymbol{\Gamma}_{\mathbf{h}}\boldsymbol{\Upsilon}_\varepsilon^H\right)\right) ,$$

$$J_{f_d\varepsilon} = \mathrm{tr}\left(\left(\boldsymbol{\Psi}_\varepsilon\boldsymbol{\Gamma}_{\mathbf{h}}\boldsymbol{\Psi}_\varepsilon^H\right)^{-1}\left(\boldsymbol{\Psi}_\varepsilon\frac{\partial\boldsymbol{\Gamma}_{\mathbf{h}}}{\partial f_d}\boldsymbol{\Psi}_\varepsilon^H\right)\left(\boldsymbol{\Psi}_\varepsilon\boldsymbol{\Gamma}_{\mathbf{h}}\boldsymbol{\Psi}_\varepsilon^H\right)^{-1}\left(\boldsymbol{\Upsilon}_\varepsilon\boldsymbol{\Gamma}_{\mathbf{h}}\boldsymbol{\Psi}_\varepsilon^H + \boldsymbol{\Psi}_\varepsilon\boldsymbol{\Gamma}_{\mathbf{h}}\boldsymbol{\Upsilon}_\varepsilon^H\right)\right) ,$$

$$J_{N_0\varepsilon} = \tfrac{1}{\xi^2}\mathrm{tr}\left(\left(\boldsymbol{\Psi}_\varepsilon\boldsymbol{\Gamma}_{\mathbf{h}}\boldsymbol{\Psi}_\varepsilon^H\right)^{-1}\left(\boldsymbol{\Psi}_\varepsilon\boldsymbol{\Gamma}_{\mathbf{h}}\boldsymbol{\Psi}_\varepsilon^H\right)^{-1}\left(\boldsymbol{\Upsilon}_\varepsilon\boldsymbol{\Gamma}_{\mathbf{h}}\boldsymbol{\Psi}_\varepsilon^H + \boldsymbol{\Psi}_\varepsilon\boldsymbol{\Gamma}_{\mathbf{h}}\boldsymbol{\Upsilon}_\varepsilon^H\right)\right) .$$

Finally, the element of the sub-matrix $\mathbf{C}$ (see (4.9)) is

$$J_{\varepsilon\varepsilon} =$$
$$\mathrm{tr}\left(\left(\boldsymbol{\Psi}_\varepsilon\boldsymbol{\Gamma}_{\mathbf{h}}\boldsymbol{\Psi}_\varepsilon^H\right)^{-1}\left(\boldsymbol{\Upsilon}_\varepsilon\boldsymbol{\Gamma}_{\mathbf{h}}\boldsymbol{\Psi}_\varepsilon^H + \boldsymbol{\Psi}_\varepsilon\boldsymbol{\Gamma}_{\mathbf{h}}\boldsymbol{\Upsilon}_\varepsilon^H\right)\left(\boldsymbol{\Psi}_\varepsilon\boldsymbol{\Gamma}_{\mathbf{h}}\boldsymbol{\Psi}_\varepsilon^H\right)^{-1}\left(\boldsymbol{\Upsilon}_\varepsilon\boldsymbol{\Gamma}_{\mathbf{h}}\boldsymbol{\Psi}_\varepsilon^H + \boldsymbol{\Psi}_\varepsilon\boldsymbol{\Gamma}_{\mathbf{h}}\boldsymbol{\Upsilon}_\varepsilon^H\right)\right) .$$

None of these elements depend on N_0 and therefore the ACRB is independent of SNR.

VITA

Flaviu Gabriel Panduru was born in Craiova, Romania, on November 8, 1972. After obtaining a master's degree in Radiocommunications from the Politehnica University, Bucharest, Romania, and working for 4 years in Bucharest as a design and maintenance engineer for a national wireless communications company, he came to Texas A&M University to pursue a master's degree in Electrical Engineering. His email address is fpanduru@wclf.tamu.edu.

The typist for this thesis was Flaviu Gabriel Panduru.