MULTITERMINAL VIDEO CODING:
FROM THEORY TO APPLICATION

A Dissertation

by

YIFU ZHANG

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

DOCTOR OF PHILOSOPHY

August 2012

Major Subject: Electrical Engineering

MULTITERMINAL VIDEO CODING

FROM THEORY TO APPLICATION

A Dissertation

by

YIFU ZHANG

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

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Zixiang Xiong
Committee Members,	Tie Liu
	Alex Sprintson
	Anxiao Jiang
Head of Department,	Costas Georghiades

August 2012

Major Subject: Electrical Engineering

ABSTRACT

Multiterminal Video Coding

From Theory to Application. (August 2012)

Yifu Zhang, M.S., Tsinghua University;

B.S., Tsinghua University

Chair of Advisory Committee: Zixiang Xiong

Multiterminal (MT) video coding is a practical application of the MT source coding theory. For MT source coding theory, two problems associated with achievable rate regions are well investigated into in this thesis: a new sufficient condition for BT sum-rate tightness, and the sum-rate loss for quadratic Gaussian MT source coding. Practical code design for ideal Gaussian sources with quadratic distortion measure is also achieved for cases more than two sources with minor rate loss compared to theoretical limits. However, when the theory is applied to practical applications, the performance of MT video coding has been unsatisfactory due to the difficulty to explore the correlation between different camera views. In this dissertation, we present an MT video coding scheme under the H.264/AVC framework. In this scheme, depth camera information can be optionally sent to the decoder separately as another source sequence. With the help of depth information at the decoder end, inter-view correlation can be largely improved and thus so is the compression performance. With the depth information, joint estimation from decoded frames and side information at the decoder also becomes available to improve the quality of reconstructed video frames. Experimental result shows that compared to separate encoding, up to 9.53% of the bit rate can be saved by the proposed MT scheme using decoder depth information, while up to 5.65% can be saved by the scheme without depth camera information. Comparisons to joint video coding schemes are also provided.

To my Parents

ACKNOWLEDGMENTS

I would like to give my greatest gratitude to Dr. Zixiang Xiong for his enlightening guidance. He has been a source of encouragement throughout my graduate studies. I would like to thank Dr. Yang Yang for his generous help without which I could not be where I am today. I would also like to thank Dr. Tie Liu, Dr. Alex Sprintson, and Dr. Anxiao Jiang for serving on my committee. I want to thank my colleagues in the Multimedia Laboratory for sharing their insightful knowledge with me. I also want to express my hearty appreciation to my parents, who gave me not only life, but all their love.

TABLE OF CONTENTS

CHAPTER Page

CHAPTER Page

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

FIGURE Page

FIGURE Page

FIGURE Page

CHAPTER I

INTRODUCTION

Multiterminal (MT) video coding refers to the problem of separate encoding and joint decoding of multiple correlated video sequences. These sequences are usually captured by closely positioned, synchronized cameras. In this procedure, different encoders (or camera views) are not allowed to communicate with each other, while bit streams for different camera views are decoded jointly. MT video coding is underpinned by the MT source coding problem [1], which deals with separate encoding and joint decoding of multiple correlated sources under distortion constraints. MT source coding is the lossy version of the distributed source coding problem first studied by Slepian and Wolf [2], who showed separate *lossless* encoding of two correlated sources (with joint decoding) suffers no rate loss when compared to joint encoding (and decoding). Later, Wyner and Ziv [3] extended a special case of Slepian-Wolf (SW) coding to lossy source coding with side information at the decoder, and showed that there is in general a rate loss with Wyner-Ziv (WZ) coding when compared to source coding with side information at both the encoder and decoder. One special case of WZ coding (with no rate loss) is when the source and side information are jointly Gaussian and the distortion measure is the mean square error (MSE).

Generally, two classes of MT source coding problems, namely *direct* MT source coding [1, 4, 5] and *indirect* MT source coding [6, 7], have been studied. The latter is often referred to as the CEO problem, where different terminals observe and separately encode noisy versions of a *single* remote source, which is to be reconstructed at the decoder. Recently, the CEO problem has been generalized to the setup with

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

multiple remote sources under a sum-distortion constraint [8, 9, 10].

By connecting the quadratic Gaussian MT source coding problem to the quadratic Gaussian CEO problem [11], Wagner *et al.* [12] showed sum-rate tightness of the Berger-Tung (BT) rate region for the two-terminal and positive-symmetric cases. Wang *et al.* [13] then provided an alternative proof based on an estimation-theoretic result, which also leads to a sufficient condition for BT sum-rate tightness. Yang and Xiong [14] started with a generalized quadratic Gaussian CEO problem and proved sum-rate tightness in the bi-eigen equal-variance with equal distortion (BEEV-ED) case. Although the BEEV-ED case satisfies the sufficient condition given in [13], the proof technique for the converse theorem is different and examples more explicit.

This thesis work starts from theoretical problems of MT source coding. First, a new and more inclusive sufficient condition than Wang *et al.*'s [13] for BT sum-rate tightness is provided. The main novelty is to consider a larger set of remote sources, such that the observation noises between the MT and remote sources have a *block-diagonal* covariance matrix, instead of a diagonal matrix as assumed in [13]. By restricting the noise covariance matrix to have K 2×2 diagonal blocks and $(L-2K)$ 1×1 diagonal blocks, we build a connection between the L-terminal problem and K two-terminal problems with *matrix-distortion* constraint.

Another problem in MT source coding theory is the sum-rate loss of quadratic Gaussian multiterminal source coding, i.e., the difference between the minimum sum-rates of distributed encoding and joint encoding (both with joint decoding) of correlated Gaussian sources subject to MSE distortion constraints on individual sources. With the minimum sum-rate given for the above-mentioned special cases of quadratic Gaussian MT source coding, it is interesting to investigate the sum-rate loss of distributed encoding as compared to joint encoding (and decoding) of Gaussian sources. However, since the minimum sum-rate for MT coding is not known in general, it is impossible to compute the sum-rate loss for all quadratic Gaussian L-terminal source coding problems. In addition, due to the individual distortion constraints, characterizing the minimum sum-rate of joint encoding becomes more difficult as the number of sources L increases.

Fortunately, with the Berger-Tung (BT) inner rate region available, we have an upper bound on the minimum sum-rate of distributed encoding. On the other hand, by relaxing the individual distortion constraints in the joint encoding problem to a sum-distortion constraint (that equals the sum of the individual target distortions), the joint encoding minimum sum-rate can be easily lower-bounded by the solution to a reverse water-filling problem [15]. By taking the difference between the upper bound for distributed encoding (with larger minimum sum-rate) and the lower bound for joint encoding (with smaller minimum sum-rate), we obtain an upper bound on the sum-rate loss for the general distributed encoding problem.

An important step in this work is devoted to proving that under the *non-degraded* assumption, that is, all target distortions are simultaneously achievable by a Gaussian BT scheme, this upper bound approaches its supremum in the BEEV-ED case, where the supremum sum-rate loss is proved to increase almost linearly with L, with an asymptotic slope of 0.1083 b/s per source as L goes to infinity. The non-degraded assumption is made such that the upper bound on the distributed encoding sum-rate can be expressed simply in terms of the eigenvalues of the source covariance matrix after proper normalization. Then because both the upper bound on the minimum sum-rate of distributed encoding and the lower bound on the minimum sum-rate of joint encoding are achieved with equality in the BEEV-ED case, we conclude that under the same assumption, the supremum sum-rate loss of quadratic Gaussian L-terminal source coding also increases almost linearly with L. It is worth noting that this result is obtained even though we currently do not have the full knowledge of the minimum sum-rate of the quadratic Gaussian MT source coding problem.

Following the theoretical work on MT source coding, practical code design are examined for quadratic Gaussian source coding with more than two terminals in both the indirect and direct setups. The focus is on cases when the BT sum-rate bounds are known to be tight. Previous research on MT source code design has mostly focused on the two-terminal case. Pradhan and Ramchandran provided a code design based on generalized coset codes for the two-terminal quadratic Gaussian CEO problem [16]. Yang *et al.* [17] proposed an SW coded quantization (SWCQ) framework for both direct and indirect quadratic Gaussian two-terminal source coding; SWCQ utilizes trellis-coded quantization (TCQ) [18] followed by low-density parity-check (LDPC) codes for SW compression. Since TCQ and LDPC codes are limit-approaching techniques for quantization and SW compression, respectively, results in [17] show only a 0.139-0.194 bit per sample (b/s) loss from the sum-rate bound of quadratic Gaussian two-terminal source coding.

Our practical designs follow the same principle of SWCQ based on TCQ and LDPC codes as in [17]. In addition to TCQ, we also employ trellis-coded vector quantization (TCVQ) [19, 20] to improve the coding efficiency in the low-rate regime (e.g., when the rate is less than one b/s for some terminals). Assuming ideal TCQ and limit-approaching LDPC coding, we show that by varying the encoding and corresponding decoding order of different terminals/observations, SWCQ can approach all corner points of the rate region of generalized Gaussian CEO coding as well as the sum-rate bound of quadratic Gaussian MT source coding. Simulations using 8192-state TCQ/TCVQ and length-10^6 LDPC codes show a sum-rate loss of only 0.106-0.162 b/s with three and four terminals.

In going from code design for two terminals to that for more than two terminals, the main issue we have to deal with in this paper is increased complexity. For the two-terminal case, LDPC profiles in the SWCQ scheme of [17] are individually designed for every SW coded bit plane of every WZ coded terminals. However, when the number of terminals/sources increases, the brute-force design method of [17] for LDPC codes becomes impractical. Therefore, the analysis of the bit-plane-wise correlation channel between the quantized source and its decoder side information becomes important. Based on the analysis, we provide approximate distributions of these channels that match well with the true distributions obtained from the real data. Our simulations show that LDPC codes designed for the approximate channel distributions suffer no loss when compared to LDPC codes designed for the true channel distributions. It provides a bridge between the theory of MT source coding and the practice of multiview/MT video coding.

On the application part, MT video coding for camera arrays and distributed video sensor networks has become a very active area of research in recent years. For example, [21] uses turbo codes to outperform the JPEG2000 standard separate encoding scheme (or simulcast scheme). In [22] and [23], nested lattice codes for DCT and wavelet transform coefficients are studied. [24] employs a six-parameter affine transform model for inter-view prediction to outperform the simulcast scheme using H.263 standard. The geometry constraints for multiple view images are analyzed in [25] and [26], and bit savings are achieved compared to JPEG2000. However, the latest H.264/AVC video coding standard [27] proves to be much more efficient in rate-distortion (R-D) performance for simulcast schemes and is difficult to outperform by MT video coding schemes. In our earlier work on two-terminal video coding [28], a bit rate saving of about 1% is achieved compared to H.264/AVC simulcast scheme by coding one sequence by H.264/AVC and the other by Wyner-Ziv video coding. We also treated three-terminal video coding [29].

The most important step in MT video coding is side information generation at the decoder, including generation of a side information frame for a WZ coded frame, and subsequently the side information for all the WZ coded components of the WZ coded frame. The quality of side information frame determines the transmission rate of WZ coded camera view sequences. To acquire high-quality side information frames from decoded frames in other camera view sequences, the configuration of camera setup and depth information (the distance value map between objects in the scene and the camera) are needed to find pixel-to-pixel correspondence between different view frames.

In all current MT video coding schemes, the depth information are acquired at the decoder by processing the decoded texture sequences. This restricts the depth accuracy and consequently, the R-D performance. Therefore, in this work, we are looking into an alternative, i.e., collecting depth information at the encoder and sending it to the decoder separately. This still conforms with the MT setup. With depth information included in the scheme, side information for MT video coding can also be made much more accurate and better R-D performance can be expected.

Joint estimation [17] is an important step in WZ coding, which reconstructs the final signal using decoded information and side information jointly. If we also consider joint estimation in MT video coding, depth information becomes more favorable since that by providing accurate geometrical information at the decoder, side information for pixel values can be acquired from previously decoded simultaneous frames from other camera views, and thus joint estimation from side information and decoded frames of the current decoded sequence is available, which can also improve the quality of the reconstructed frames. Moreover, joint estimation can also be used for multiview video coding (MVC) scheme if depth information is provided. This reemphasizes the importance of depth information in 3-D video applications.

Additionally, using depth information at the decoder also reduces decoder complexity, since a MT scheme without depth information at the decoder usually needs to employ complicated stereo matching algorithms to find pixel mapping between different views for better R-D performance [28], while MT scheme with depth information at the decoder can get such pixel mapping from depth information by simple affine transform.

On the other hand, for hardware implementation, different types of depth cameras have been provided for research and even commercial use. For example, the SwissRanger series range camera [30] can directly capture the depth map of a scene in real time; the successful launch of Microsoft Kinect makes deployment of cheap commodity depth cameras a step closer to reality. Although constrained by its relatively lower resolution and high geometrical distortion compared to traditional video cameras (or texture cameras), depth cameras can provide more accurate depth information for the background as well as objects that cannot be easily discerned by existing stereo matching algorithms, especially when the number of stereo views is limited (e.g., no more than three). Thus, using depth information in MT video coding can be expected to be more popular in practice as such devices become more advanced.

Therefore, in the application part of this thesis, we provide schemes and experiment results on MT video coding with/without separate depth information sent to the decoder and used as side information, aiming at enhancing the R-D performance of the the current MT coding schemes, compared to the simulcast video coding scheme. We implemented our proposed scheme to both standard MVC test sequences and a sequence with actual synchronized low-resolution depth sequence. For MVC test sequences, since *a priori* depth information is not available for such sequences, we first generate low-resolution (thus low transmission rate) depth information by processing simultaneous original frames from different views. Such depth information is encoded and transmitted to the decoder separately to satisfy the MT coding constraint, and then used at the joint decoder to help construct better side information for MT coding. The proposed MT video coding scheme is implemented and experimented under the H.264/AVC standard framework. Therefore, by comparing with the simulcast video coding scheme using H.264/AVC joint model (JM) reference software [31], we found that by transmitting additional depth information to the joint decoder, up to 9.53% of the bit rate can be saved compared to the simulcast video coding scheme, while 5.65% can be saved in the case no depth information transmitted to the decoder. We also compared our result to the MVC scheme using H.264/AVC joint multiview video model (JMVM) reference software [32, 33], and it shows that the MT scheme with depth information at the decoder still suffers an average sum rate loss up to 8.54% compared to the MVC scheme.

The remainder of the thesis is organized as follows: Chapter II provides a new sufficient condition for sum-rate tightness in MT source coding after a brief summary of the background of Gaussian quadratic MT source coding theory; Chapter III gives some new results on sum-rate loss of Gaussian quadratic MT source coding; Chapter IV deals with practical code design problems for MT source coding with know tight sum-rate bounds; Chapter V focuses on the application of MT source coding, MT video coding; and finally Chapter VI concludes the dissertation.

CHAPTER II

A NEW SUFFICIENT CONDITION FOR SUM-RATE TIGHTNESS IN QUADRATIC GAUSSIAN MT SOURCE CODING

In this chapter, we first provide a brief review of the setup of quadratic Gaussian MT source coding problem in Section A and existing results in this setup in Section 2. Section B studies the two-terminal source coding problem with matrix-distortion constraint, and provides an improved lower bound on the sum-rate. Section C states our main results on a new sufficient condition for sum-rate tightness, and presents a degraded example belonging to the block-degraded case that satisfies our new condition. Section D gives a simplified sufficient condition for the sum-rate tightness in the non-degraded cases, followed by two additional examples satisfying the simplified condition.

A. The quadratic Gaussian MT source coding problem

1. Quadratic Gaussian direct MT coding

For any integer L, denote $\mathcal{L} = \{1, 2, ..., L\}$. Let $Y_{\mathcal{L}} = (Y_1, Y_2, ..., Y_L)^{\mathrm{T}}$ be a length-L vector Gaussian source with mean $\mathbf{0}$ and covariance matrix $\Sigma_{Y_{\mathcal{L}}}$. Also denote $Y_{\mathcal{S}_k}$ as the length-$|\mathcal{S}_k|$ subvector of $Y_{\mathcal{L}}$ indexed by $\mathcal{S}_k$. For an integer n, let $\boldsymbol{Y}_{\mathcal{L}} = (Y_{\mathcal{L},1}, Y_{\mathcal{L},2}, ..., Y_{\mathcal{L},n})$ be an $L \times n$ matrix with $Y_{\mathcal{L},i}$, $i = 1, 2, ..., n$ being n independent drawings of $Y_{\mathcal{L}}$. Also denote $\boldsymbol{Y}_{\mathcal{S}_k}$ as the $|\mathcal{S}_k| \times n$ submatrix of $\boldsymbol{Y}_{\mathcal{L}}$ with column indices $\mathcal{S}_k$. For any $L \times n$ random matrix $\boldsymbol{Y}_{\mathcal{L}}$ and any random object ω, define the conditional covariance matrix of $\boldsymbol{Y}_{\mathcal{L}}$ given ω as

$$\mathrm{cov}(\boldsymbol{Y}_{\mathcal{L}}|\omega) \triangleq \frac{1}{n}\mathrm{E}\left[\left(\boldsymbol{Y}_{\mathcal{L}} - \mathrm{E}(\boldsymbol{Y}_{\mathcal{L}}|\omega)\right)\left(\boldsymbol{Y}_{\mathcal{L}} - \mathrm{E}(\boldsymbol{Y}_{\mathcal{L}}|\omega)\right)^{\mathrm{T}}\right]. \tag{2.1}$$

Consider the task of separately compressing a length-n block of sources $\boldsymbol{Y}_{\mathcal{L}}$ at L encoders and jointly reconstructing $\boldsymbol{Y}_{\mathcal{L}}$ as $\hat{\boldsymbol{Y}}_{\mathcal{L}}$ at a central decoder subject to individual distortion constraints $\boldsymbol{D}_{\mathcal{L}} = \{D_1, D_2, ..., D_L\}$. For compact notation, subscript $\mathcal{L}$ will be dropped in the rest of the thesis if no ambiguity is incurred. This problem is known as the *quadratic Gaussian MT source coding problem*, whose block diagram is depicted in Fig 1.

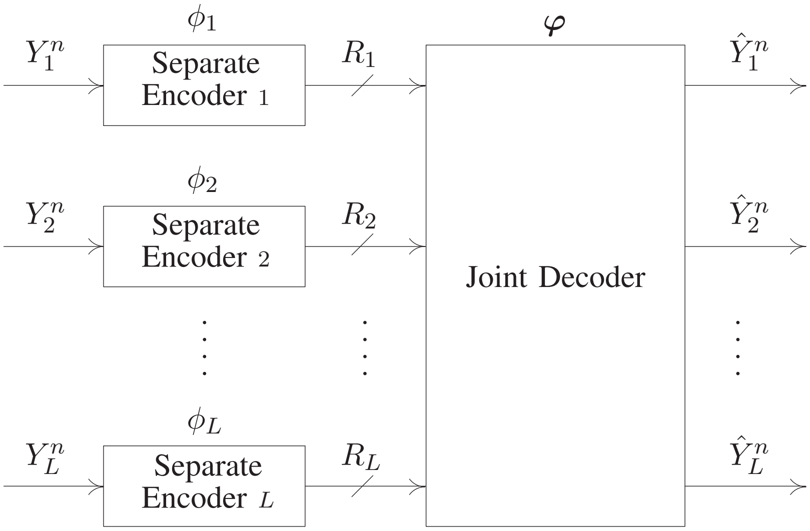


Fig. 1. The quadratic Gaussian MT source coding problem.

Let

$$\phi_j^{(n)} : \mathbb{R}^n \mapsto \left\{1, 2, ..., 2^{R_j^{(n)}} - 1\right\}, \quad j \in \mathcal{L} \tag{2.2}$$

be the j-th encoder function and

$$\psi_j^{(n)} : \left\{1, \dots, 2^{R_1^{(n)}} - 1\right\} \times \left\{1, \dots, 2^{R_2^{(n)}} - 1\right\} \times \cdots \times \left\{1, \dots, 2^{R_L^{(n)}} - 1\right\} \mapsto \mathbb{R}^n \tag{2.3}$$

be the reconstruction function for $\boldsymbol{Y}_j$. Denote W_j as the transmitted symbol at the j-th encoder, and $R_{\boldsymbol{\Sigma_Y}}^{\mathrm{MT}}\left(\phi_{\mathcal{L}}^{(n)}, \psi_{\mathcal{L}}^{(n)}\right) = \sum_{j\in\mathcal{L}} R_j^{(n)}$ as the sum-rate of the MT coding scheme $\left(\phi_{\mathcal{L}}^{(n)}, \psi_{\mathcal{L}}^{(n)}\right)$. We say a rate tuple $(R_1, ..., R_L)^{\mathrm{T}}$ is $(\boldsymbol{\Sigma_Y}, \boldsymbol{D})$-achievable if there exists a sequence of schemes $\left\{\left(\phi_{\mathcal{L}}^{(n)}, \psi_{\mathcal{L}}^{(n)}\right) : n \in \mathbb{N}^{+}\right\}$ such that

$$\begin{aligned} \limsup_{n\to\infty} R_j^{(n)} &\leq R_j, \quad \text{for any } j \in \mathcal{L}, && (2.4)\\ \limsup_{n\to\infty} \frac{1}{n}\mathrm{E}\left[(Y_{j,i} - \hat{Y}_{j,i})^2\right] &\leq D_j, \quad \text{for any } j \in \mathcal{L}. && (2.5) \end{aligned}$$

Define the $(R_1, ..., R_L)^{\mathrm{T}}$ is $(\boldsymbol{\Sigma_Y}, \boldsymbol{D})$-achievable rate region $\mathcal{R}_{\boldsymbol{\Sigma_Y}}^{\mathrm{MT}}(\boldsymbol{D})$ as the convex closure of all $(R_1, ..., R_L)^{\mathrm{T}}$-achievable rate tuples, i.e.,

$$\mathcal{R}_{\boldsymbol{\Sigma_Y}}^{\mathrm{MT}}(\boldsymbol{D}) = \mathrm{cl}\left\{(R_1, R_2, \ldots, R_L)^{\mathrm{T}} : (R_1, R_2, \ldots, R_L)^{\mathrm{T}} \text{ is } (\boldsymbol{\Sigma_Y}, \boldsymbol{D}) \text{ achievable.}\right\} \quad (2.6)$$

The *minimum sum-rate* with respect to $(\boldsymbol{\Sigma_Y}, \boldsymbol{D})$ is then defined as

$$R_{\boldsymbol{\Sigma_Y}}^{\mathrm{MT}}(\boldsymbol{D}) = \inf\left\{\sum_{i=1}^{L} R_i : (R_1, R_2, \ldots, R_L)^{\mathrm{T}} \in \mathcal{R}_{\boldsymbol{\Sigma_Y}}^{\mathrm{MT}}(\boldsymbol{D})\right\}. \quad (2.7)$$

In order to study the sum-rate loss, we also consider the problem of joint encoding (and joint decoding) of the same length-L Gaussian vector source $\boldsymbol{Y}$. Let $\left(\phi_{\mathrm{Joint}}^{(n)}, \varphi_{\mathrm{Joint}}^{(n)}\right)$ be a pair of joint encoding/decoding functions defined as

$$\begin{aligned} \phi_{\mathrm{Joint}}^{(n)} : \quad & \underbrace{\mathbb{R}^n \times \ldots \times \mathbb{R}^n}_{L} \rightarrow \left\{1, 2, ..., M_{\mathrm{Joint}}^{(n)}\right\}, \\ \varphi_{\mathrm{Joint}}^{(n)} : \quad & \left\{1, 2, ..., M_{\mathrm{Joint}}^{(n)}\right\} \rightarrow \underbrace{\mathbb{R}^n \times \ldots \times \mathbb{R}^n}_{L}. \end{aligned}$$

A non-negative rate R is $(\boldsymbol{\Sigma_Y}, \boldsymbol{D})$-*jointly-achievable* if there exists a sequence of schemes $\left\{\left(\phi_{\mathrm{Joint}}^{(n)}, \varphi_{\mathrm{Joint}}^{(n)}\right) : n \in \mathbb{N}^{+}\right\}$ such that

$$\begin{aligned} \limsup_{n\to\infty} \frac{1}{n} \log_2 M_{\mathrm{Joint}}^{(n)} &\leq R, \\ \limsup_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} E\left[(Y_{i,j} - \hat{Y}_{i,j})^2\right] &\leq D_i, \forall i \in \mathcal{L}. \end{aligned}$$

are satisfied. The *joint encoding minimum sum-rate* with respect to $(\boldsymbol{\Sigma_Y}, \boldsymbol{D})$ is similarly defined as

$$R^{\text{Joint}}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}) = \min\{R : R \text{ is } (\boldsymbol{\Sigma_Y}, \boldsymbol{D}) -\text{jointly}-\text{achievable}\}\,.$$

Then the *sum-rate loss* of distributed over joint encoding is defined as

$$R^{\Delta}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}) = R^{\text{MT}}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}) - R^{\text{Joint}}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}).$$

Berger and Tung [1, 4] provide an *inner rate region* inside which all rate tuples are $(\boldsymbol{\Sigma_Y}, \boldsymbol{D})$-achievable. In this paper, we restrict ourselves to a subset of the Berger-Tung inner rate region inside which all points can be achieved by parallel Gaussian test channels. This subset is referred to as the *Berger-Tung (BT)* inner rate region in the sequel. Let $\boldsymbol{U}_{\mathcal{L}} = (U_1, U_2, \ldots, U_L)^{\text{T}}$ be a length-L auxiliary random vector such that

- $U_i = Y_i + Q_i, i = 1, 2, \ldots, L$, where $Q_i \sim \mathcal{N}(0, \sigma^2_{Q_i})$, and all Q_i's are independent of each other and of all Y_i's,
- $\boldsymbol{U}_{\mathcal{L}}$ satisfies $\text{E}\left\{(Y_i - E(Y_i|\boldsymbol{U}_{\mathcal{L}}))^2\right\} \leq D_i$ for all $i = 1, 2, \ldots, L$,

and define $\mathcal{U}(\boldsymbol{\Sigma_Y}, \boldsymbol{D})$ as the set of all auxiliary random vectors $\boldsymbol{U}$ that satisfy the above conditions. Then the following lemma gives the BT inner rate region, the proof can be found in [1, 4].

Lemma 1. *Define*

$$\mathcal{R}^{\text{BT}}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}) = \bigcup_{U_{\mathcal{L}} \in \mathcal{U}(\Sigma_{Y_{\mathcal{L}}}, D_{\mathcal{L}})} \left\{ (R_1, R_2, \ldots, R_L)^{\text{T}} : \sum_{i \in \mathcal{A}} R_i \geq I(Y_{\mathcal{A}}; \boldsymbol{U}_{\mathcal{A}} | \boldsymbol{U}_{\mathcal{L}-\mathcal{A}}) \right\}, \quad (2.8)$$

then

$$\mathcal{R}^{\text{BT}}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}) \subseteq \mathcal{R}^{\text{MT}}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}). \quad (2.9)$$

In particular, the BT minimum sum-rate

$$R^{\mathrm{BT}}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}) = \inf\left\{\ \sum_{i=1}^{L} R_i : (R_1, R_2, \ldots, R_L)^{\mathrm{T}} \in \mathcal{R}^{\mathrm{BT}}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D})\right\}$$

$$= \inf_{\boldsymbol{\Sigma_Q} \in {}_{L} : \left[\left(\boldsymbol{\Sigma_Y}^{-1} + \boldsymbol{\Sigma_Q}^{-1}\right)^{-1}\right]_{j,j} \leq D_j,\ \forall j \in \mathcal{L}} \frac{1}{2}\log_2\left[\frac{|\boldsymbol{\Sigma_Y}|}{|(\boldsymbol{\Sigma_Y}^{-1} + \boldsymbol{\Sigma_Q}^{-1})^{-1}|}\right] \quad (2.10)$$

satisfies

$$R^{\mathrm{MT}}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}) \leq R^{\mathrm{BT}}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}), \quad (2.11)$$

where $_L$ *denotes the set of all* $L \times L$ *positive definite (p.d.) diagonal matrices.*

For example, the BT rate region for the quadratic Gaussian two-terminal source coding problem with $\boldsymbol{\Sigma_Y} = \begin{bmatrix} \sigma^2_{Y_1} & \rho\sigma_{Y_1}\sigma_{Y_2} \\ \rho\sigma_{Y_1}\sigma_{Y_2} & \sigma^2_{Y_2} \end{bmatrix}$ is given by

$$\mathcal{R}^{\mathrm{BT}}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}) = \hat{\mathcal{R}}^{\mathrm{BT}}_1(D_1, D_2) \cap \hat{\mathcal{R}}^{\mathrm{BT}}_2(D_1, D_2) \cap \hat{\mathcal{R}}^{\mathrm{BT}}_{12}(D_1, D_2), \quad (2.12)$$

where

$$\hat{\mathcal{R}}^{\mathrm{BT}}_i(D_1, D_2) = \left\{(R_1, R_2) : R_i \geq \frac{1}{2}\log^+\left[(1 - \rho^2 + \rho^2 2^{-2R_j})\frac{\sigma^2_{Y_i}}{D_i}\right]\right\}, i, j = 1, 2, i \neq j, \quad (2.13)$$

$$\hat{\mathcal{R}}^{\mathrm{BT}}_{12}(D_1, D_2) = \left\{(R_1, R_2) : R_1 + R_2 \geq \frac{1}{2}\log^+\left[(1 - \rho^2)\frac{\beta_{max}\sigma^2_{Y_1}\sigma^2_{Y_2}}{2D_1D_2}\right]\right\}, \quad (2.14)$$

with $\beta_{max} = 1 + \sqrt{1 + \frac{4\rho^2 D_1 D_2}{(1-\rho^2)^2\sigma^2_{Y_1}\sigma^2_{Y_2}}}$, and $\log^+ x = \max\{\log x, 0\}$. The BT rate region with $\sigma^2_{Y_1} = \sigma^2_{Y_2} = 1$, $\rho = 0.9$, and $D_{\mathcal{L}} = (0.1, 0.1)^{\mathrm{T}}$ is shown in Fig. 2, where $\partial\hat{\mathcal{R}}^{\mathrm{BT}}_i(D_1, D_2)$ and $\partial\hat{\mathcal{R}}^{\mathrm{BT}}_{12}(D_1, D_2)$ are the boundaries of $\hat{\mathcal{R}}^{\mathrm{BT}}_i(D_1, D_2)$ and $\hat{\mathcal{R}}^{\mathrm{BT}}_{12}(D_1, D_2)$, respectively.

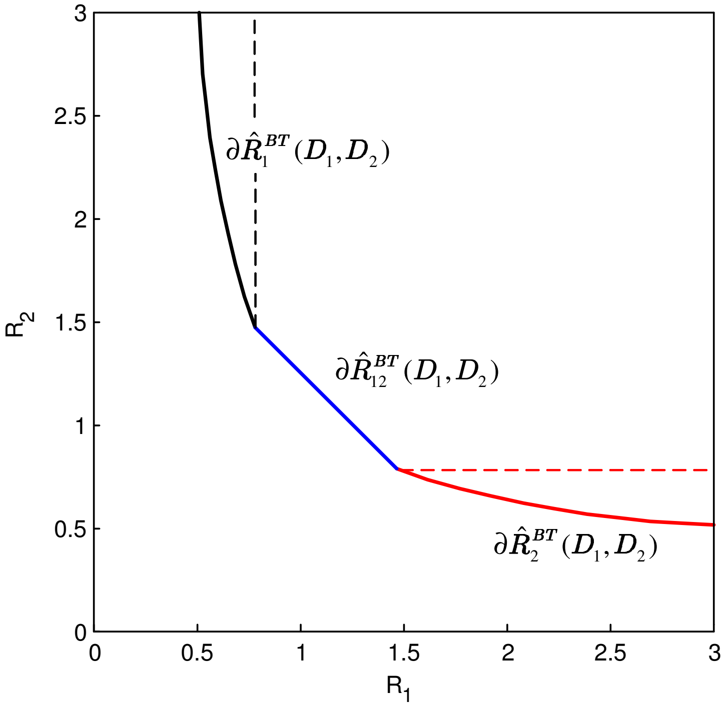


Fig. 2. An example of the BT rate region for the quadratic Gaussian two-terminal source coding problem.

2. Existing results on sum-rate tightness

Wagner *et al.* [12] proved that for the two-terminal case (with $L = 2$), the BT minimum sum-rate is equal to the MT minimum sum-rate, as stated in the following lemma,

Lemma 2 ([12]). *For any positive-definite symmetric $\boldsymbol{\Sigma_Y} \in \mathbb{R}^{2\times 2}$ and any positive real vector $\boldsymbol{D} = (D_1, D_2)^{\mathrm{T}}$, it holds that*

$$R^{\mathrm{MT}}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}) = R^{\mathrm{BT}}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}),$$

where $R^{\mathrm{MT}}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D})$ and $R^{\mathrm{BT}}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D})$ are the MT and BT sum-rate for $\boldsymbol{\Sigma_Y}$ with distortion constraint $\boldsymbol{D}$ respectively.

They also showed tightness of the BT sum-rate bound for the positive symmetric case, i.e., (2.16) holds for any $L \times L$ *positive-symmetric* matrices of the form

$$\Sigma_{Y_{\mathcal{L}}} = \mathcal{S}_L(a,b) \triangleq \begin{bmatrix} a & b & b & ... & b \\ b & a & b & ... & b \\ ... & ... & ... & ... & ... \\ b & b & b & ... & a \end{bmatrix}, \tag{2.15}$$

for some $a > b > 0$ and any $\boldsymbol{D} = (D, D, ..., D)^{\mathrm{T}}$ for some $D > 0$. Fig. 3 depicts the QB rate region of quadratic Gaussian three-terminal source coding with $\rho = 0.8$ and $D = 0.05$, which is a 3-D extension of the rate region for quadratic Gaussian two-terminal source coding [34, 35]. The sum-rate bound is the hexagonal portion of the hyperplane defined by $R_1 + R_2 + R_3 = R_{\boldsymbol{Y}}(3, 0.8, 0.05) = 4.865$ b/s. The six corner points of the hexagon corresponds to different encoding orders for the three sources.

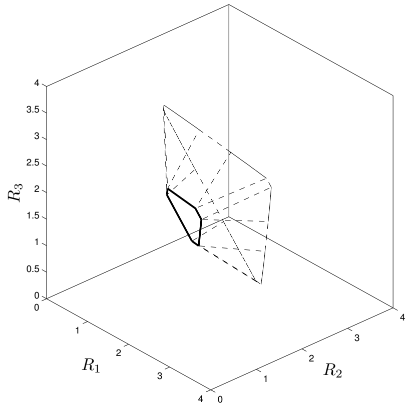


Fig. 3. The BT rate region of quadratic Gaussian three-terminal source coding in the positive symmetric case with $\rho = 0.8$ and $D = 0.05$.

It is recently proved in [14] that tightness of the BT minimum sum-rate also holds for a more general class called BEEV-ED, where the source covariance matrix $\boldsymbol{\Sigma_Y}$ is bi-eigen equal-variance (BEEV), i.e., $\boldsymbol{\Sigma_Y}$ has equal diagonal element and two distinct eigenvalues, and the target distortions are equal for all sources. We summarize these results in the following three lemmas.

Lemma 3 ([14])**.** *For any positive-definite symmetric $\mathbf{\Sigma}_{\boldsymbol{Y}}$ with equal diagonal element $a > 0$ and two distinct eigenvalues, and any positive real number $D \in (0, a]$, it holds that*

$$R^{\mathrm{MT}}_{\mathbf{\Sigma}_{\boldsymbol{Y}}}(D\mathbf{1}) = R^{\mathrm{BT}}_{\mathbf{\Sigma}_{\boldsymbol{Y}}}(D\mathbf{1}).$$

Moreover, the optimal sequence of schemes that approaches the minimum sum-rate $R^{\mathrm{MT}}_{\mathbf{\Sigma}_{\boldsymbol{Y}}}(D\mathbf{1})$ must also approach the target distortion vector $D\mathbf{1}$.

The most general cases of quadratic Gaussian MT source coding problem with tight sum-rate are provided by Wang *et al.* [13]. Their proof contains four major steps.

- First, the L MT sources $\boldsymbol{Y}$ are connected to L remote sources $\boldsymbol{X}$ such that

  $$\boldsymbol{Y} = \boldsymbol{X} + \boldsymbol{N} \tag{2.16}$$

  with $\boldsymbol{N}$ being a zero-mean Gaussian vector independent of $\boldsymbol{X}$ with a *diagonal* covariance matrix

  $$\mathbf{\Sigma}_{\boldsymbol{N}} = \begin{bmatrix} \sigma^2_{N_1} & 0 & ... & 0 \\ 0 & \sigma^2_{N_2} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \sigma^2_{N_L} \end{bmatrix}. \tag{2.17}$$

  Then they use the Markov chain $\boldsymbol{X} \to \boldsymbol{Y} \to W$ to obtain an *estimation-theoretic result* that $\mathrm{cov}(\boldsymbol{Y}|\boldsymbol{X}, W)$ must also be diagonal.

- Exploit the *semidefinite partial order* of the distortion matrices, which is due to the fact that a linear *minimum mean squared error* (MMSE) estimator cannot outperform its optimal MMSE counterpart, to show that

$$\mathrm{cov}(\boldsymbol{Y}|\boldsymbol{X},W) \preceq \left((\mathrm{cov}(\boldsymbol{Y}|W))^{-1} + \boldsymbol{\Sigma}_{\boldsymbol{N}}^{-1} - \boldsymbol{\Sigma}_{\boldsymbol{Y}}^{-1}\right)^{-1}.$$

- A lower bound on the MT minimum sum-rate $R_{\boldsymbol{\Sigma_Y}}^{\mathrm{MT}}(\boldsymbol{D})$ is derived by exploiting the diagonal structure of $\mathrm{cov}(\boldsymbol{Y}|\boldsymbol{X},W)$.
- Form a convex optimization problem that minimizes the above lower bound over $\boldsymbol{D} \triangleq \mathrm{cov}(\boldsymbol{Y}|W)$ and $\gamma \triangleq \mathrm{diag}(\mathrm{cov}(\boldsymbol{Y}|\boldsymbol{X},W))$, and establish a sufficient condition for the $\boldsymbol{D}$ and γ that correspond to the optimal BT scheme to satisfy the the KKT condition of the optimization problem.

Specifically, let $\mathscr{P}_L^{\succeq}$ be the set of $L\times L$ p.s.d. matrices and $\mathbb{d}$ be the set of diagonal matrices. Define $\mathscr{D}(\boldsymbol{D},\boldsymbol{\Sigma_Y})$ as the set of all BT-achievable distortion matrices that satisfy the distortion constraints, and $\mathscr{N}(\boldsymbol{\Sigma_Y})$ as the set of all possible diagonal covariance matrices $\Sigma_{\boldsymbol{N}}$, i.e.,

$$\mathscr{D}(\boldsymbol{D},\boldsymbol{\Sigma_Y}) \triangleq \left\{\boldsymbol{D} \in \mathbb{R}^{L\times L} : [\boldsymbol{D}]_{j,j} = D_j, \forall j \in \mathcal{L}, \text{ and } \boldsymbol{D}^{-1} - \boldsymbol{\Sigma}_{\boldsymbol{Y}}^{-1} \in \mathscr{P}^{\succeq} \cap \mathbb{d}\right\}, \tag{2.18}$$

$$\mathscr{N}(\boldsymbol{\Sigma}_Y) \triangleq \left\{\boldsymbol{\Sigma} \in \mathscr{P}^{\succeq} \cap \mathbb{d} : \boldsymbol{\Sigma} \succeq \boldsymbol{\Sigma_Y}\right\}. \tag{2.19}$$

Wang *et al.*'s result [13] is summarized in the following theorem.

Theorem 1 ([13])**.** *If for some* $\boldsymbol{D} \in \mathscr{D}(\boldsymbol{D},\boldsymbol{\Sigma_Y})$ *and* $\boldsymbol{\Sigma_N} \in \mathscr{N}(\boldsymbol{\Sigma}_Y)$*, there exists a diagonal matrix* $\boldsymbol{\Pi} = \mathrm{diag}(\pi_1, ..., \pi_L)$ *such that*

$$\boldsymbol{D}\left(\boldsymbol{\Pi} - \boldsymbol{D}^{-1} + \boldsymbol{D}^{-1}\left(\boldsymbol{D}^{-1} + \boldsymbol{\Sigma}_{\boldsymbol{N}}^{-1} - \boldsymbol{\Sigma}_{\boldsymbol{Y}}^{-1}\right)^{-1}\boldsymbol{D}^{-1}\right)\boldsymbol{D} \tag{2.20}$$

is a p.s.d. matrix with the same diagonal elements as those of $\left(\boldsymbol{D}^{-1} + \boldsymbol{\Sigma}_{\boldsymbol{N}}^{-1} - \boldsymbol{\Sigma}_{\boldsymbol{Y}}^{-1}\right)^{-1}$, *then the BT sum-rate bound is tight, i.e.,*

$$\mathcal{R}_{\mathbf{\Sigma}_{\boldsymbol{Y}}}^{\mathrm{BT}}(\boldsymbol{D}) = \mathcal{R}_{\mathbf{\Sigma}_{\boldsymbol{Y}}}^{\mathrm{MT}}(\boldsymbol{D}). \tag{2.21}$$

Using a different technique, sum-rate tightness for a special *bi-eigen equal-variance with equal distortion* class of MT problems was proved by Yang and Xiong [14]. That is, (2.16) holds for any $\mathbf{\Sigma}_{\boldsymbol{Y}} \in \mathcal{B}$ and $\boldsymbol{D} = (D, D, ..., D)^{\mathrm{T}}$ for some $D > 0$, where $\mathcal{B}$ denotes the set of all $L \times L$ p.s.d. matrices with two distinct eigenvalues and equal diagonal elements.

3. Quadratic Gaussian indirect MT coding

a. The Gaussian CEO problems

Let X be a Gaussian remote source with zero mean and variance σ_X^2 and $\boldsymbol{Y}_{\mathcal{L}} = (X, X, \cdots, X)^{\mathrm{T}} + \boldsymbol{N}_{\mathcal{L}}$ be the observations, where $\mathcal{L} = \{1, 2, \ldots, L\}$, $\boldsymbol{Y}_{\mathcal{L}} = (Y_1, Y_2, \ldots, Y_L)^{\mathrm{T}}$ and $\boldsymbol{N}_{\mathcal{L}} = (N_1, N_2, \ldots, N_L)^{\mathrm{T}}$ is a length-L Gaussian vector noise independent of X with zero mean and covariance matrix $\mathbf{\Sigma}_{\boldsymbol{N}_{\mathcal{L}}} = \mathrm{diag}\left(\sigma_{N_1}^2, \sigma_{N_2}^2, \ldots, \sigma_{N_L}^2\right)$.

Each of the L encoders observes exactly one component of $\boldsymbol{Y}_{\mathcal{L}}$, and separately encodes a length-n block of its own observation $Y_\ell^n = (Y_{\ell,1}, Y_{\ell,2}, \ldots, Y_{\ell,n})$ into $W_\ell \in \left\{1, 2, \ldots, 2^{R_\ell}\right\}$, using function

$$\phi_\ell : \ \mathbb{R}^n \mapsto \left\{1, 2, \ldots, 2^{R_\ell}\right\}, \ \ell \in \mathcal{L}. \tag{2.22}$$

The joint decoder receives W_ℓ for all $\ell \in \mathcal{L}$ before reconstructing $X^n = (X_1, X_2, \ldots, X_n)$ as $\hat{X}^n = (\hat{X}_1, \hat{X}_2, \ldots, \hat{X}_n)$ using function

$$\psi : \left\{1, 2, \ldots, 2^{R_1}\right\} \times \cdots \times \left\{1, 2, \ldots, 2^{R_L}\right\} \mapsto \mathbb{R}^n. \tag{2.23}$$

The block diagram for the Gaussian CEO problem is depicted in Fig. 4.

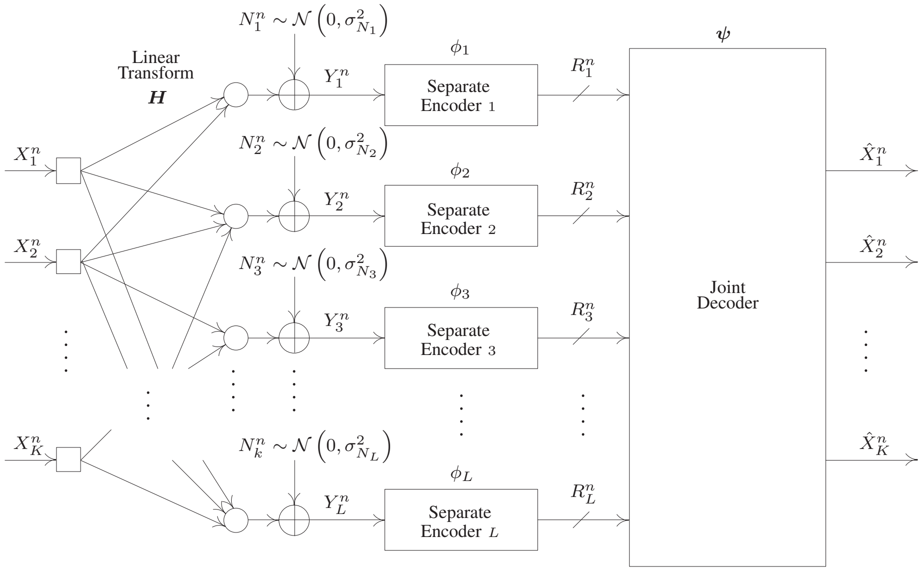


Fig. 4. The Gaussian CEO problem.

We say a rate vector $(R_1, R_2, \ldots, R_L)^{\mathrm{T}}$ is $(\sigma_X^2, \boldsymbol{\Sigma}_{\boldsymbol{N}_{\mathcal{L}}}, D)$-*achievable* for distortion measure D if there exist L encoder functions ϕ_l, $l = 1, 2, \ldots, L$ and a decoder function ψ such that the distortion constraint

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mathrm{E}\left[\left(X_i - \hat{X}_i\right)^2\right] \leq D, \tag{2.24}$$

is satisfied. The achievable rate region $\boldsymbol{\mathcal{R}}_X(\sigma_X^2, \boldsymbol{\Sigma}_{\boldsymbol{N}_{\mathcal{L}}}, D)$ is defined as the convex hull of the set of all achievable rate vectors, i.e.,

$$\boldsymbol{\mathcal{R}}_X\left(\sigma_X^2, \boldsymbol{\Sigma}_{\boldsymbol{N}_{\mathcal{L}}}, D\right) = \mathrm{cl}\left\{(R_1, \ldots, R_L)^{\mathrm{T}} : (R_1, \ldots, R_L)^{\mathrm{T}} \text{ is } (\sigma_X^2, \boldsymbol{\Sigma}_{\boldsymbol{N}_{\mathcal{L}}}, D)\text{-achievable}\right\}. \tag{2.25}$$

Similar to that of direct MT coding, the BT inner rate region in the indirect case can be defined as

$$\boldsymbol{\mathcal{R}}_X^{\mathrm{BT}}\left(\sigma_X^2, \boldsymbol{\Sigma}_{\boldsymbol{N}_{\mathcal{L}}}, D\right) = \mathrm{cl}\left\{ \bigcup_{U_{\mathcal{L}} \in \mathcal{U}\left(\sigma_X^2, \boldsymbol{\Sigma}_{\boldsymbol{N}_{\mathcal{L}}}, D\right)} \left\{ \left(R_1, \ldots, R_L\right)^{\mathrm{T}} \middle| \sum_{i \in \mathcal{A}} R_i \geq I\left(Y_{\mathcal{A}}; U_{\mathcal{A}} | U_{\mathcal{L}-\mathcal{A}}\right) \right\}, \right\} \tag{2.26}$$

where $\mathcal{A} \subseteq \mathcal{L}$, and $\mathcal{U}\left(\sigma_X^2, \boldsymbol{\Sigma}_{\boldsymbol{N}_{\mathcal{L}}}, D\right)$ is the set of all length-L auxiliary random vectors $U_{\mathcal{L}} = \left(U_1, \ldots, U_L\right)^{\mathrm{T}}$ such that $U_i = Y_i + Q_i$, for $i = 1, \ldots, L$, and $\mathrm{E}\left\{\left(X - \mathrm{E}\left(X | U_{\mathcal{L}}\right)\right)^2\right\} \leq D$, where $Q_i \sim \mathcal{N}\left(0, \sigma_{Q_i}^2\right)$ and Q_i's are independent of each other as well as of all Y_i's.

Oohama [11] proved that the BT rate region is tight for the Gaussian CEO problem with any L, i.e.,

$$\boldsymbol{\mathcal{R}}_X\left(\sigma_X^2, \boldsymbol{\Sigma}_{\boldsymbol{N}_{\mathcal{L}}}, D\right) = \boldsymbol{\mathcal{R}}_X^{\mathrm{BT}}\left(\sigma_X^2, \boldsymbol{\Sigma}_{\boldsymbol{N}_{\mathcal{L}}}, D\right), \tag{2.27}$$

which implies that the BT sum-rate bound is also tight. Specifically, if we assume $\sigma_{N_1}^2 = \cdots = \sigma_{N_L}^2 = \sigma_N^2$, the BT sum-rate is [11, 36]

$$R_X\left(\sigma_X^2, \sigma_N^2 \cdot \boldsymbol{I}, D\right) = R_X^{\mathrm{BT}}\left(\sigma_X^2, \sigma_N^2 \cdot \boldsymbol{I}, D\right) = \frac{L}{2} \log^+ \frac{L D^{1-1/L}\left(\sigma_X^2\right)^{1+1/L}}{L D \sigma_X^2 - \sigma_N^2\left(\sigma_X^2 - D\right)^+}. \tag{2.28}$$

b. The generalized Gaussian CEO problem

Let K and L be two positive integers. Denote $\mathcal{K} = \{1, 2, \ldots, K\}$. Let $\boldsymbol{X}_{\mathcal{K}} = \left(X_1, X_2, \ldots, X_K\right)^{\mathrm{T}}$ be a length-K Gaussian source vector with zero mean and covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{X}}$, and $\boldsymbol{H}$ be an $L \times K$ matrix with full column rank. Define the observations $\boldsymbol{Y}_{\mathcal{L}}$ as $\boldsymbol{Y}_{\mathcal{L}} = \boldsymbol{H}\boldsymbol{X}_{\mathcal{K}} + \boldsymbol{N}_{\mathcal{L}}$.

Encoding and decoding of $\boldsymbol{Y}_{\mathcal{L}}$ are similar to those in the original Gaussian CEO problem in Section a. The only difference is that the joint decoder aims to reconstruct

$\boldsymbol{X}_{\mathcal{K}} = (\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_K)^{\mathrm{T}}$ as $\hat{\boldsymbol{X}}_{\mathcal{K}} = (\hat{\boldsymbol{X}}_1, \hat{\boldsymbol{X}}_2, \ldots, \hat{\boldsymbol{X}}_K)^{\mathrm{T}}$, using function

$$\boldsymbol{\psi} : \left\{1, 2, \ldots, 2^{R_1}\right\} \times \cdots \times \left\{1, 2, \ldots, 2^{R_L}\right\} \mapsto \mathbb{R}^{K \times n}. \tag{2.29}$$

Obviously, the Gaussian CEO problem corresponds to the special case with $K = 1$ and $\boldsymbol{H} = \mathbf{1}$.

Denote $\mathbb{T} = (\boldsymbol{\Sigma}_{\boldsymbol{X}}, \boldsymbol{H}, \boldsymbol{\Sigma}_{\boldsymbol{N}_{\mathcal{L}}})$, we say a rate vector $(R_1, R_2, \ldots, R_L)^{\mathrm{T}}$ is $(\mathbb{T}, D)$-*achievable* if there exist L encoder functions ϕ_l, $l = 1, 2, \ldots, L$ and a decoder function ψ such that the sum-distortion constraint

$$\limsup_{n \to \infty} \frac{1}{n} \sum_{k=1}^{K} \sum_{i=1}^{n} \mathrm{E}\left[\left(X_{k,i} - \hat{X}_{k,i}\right)^2\right] \leq D, \tag{2.30}$$

is satisfied. The achievable rate region is

$$\boldsymbol{\mathcal{R}}_{\boldsymbol{X}_{\mathcal{K}}}(\mathbb{T}, D) = \mathrm{cl}\left\{(R_1, \ldots, R_L)^{\mathrm{T}} : (R_1, \ldots, R_L)^{\mathrm{T}} \text{ is } (\mathbb{T}, D)\text{-achievable}\right\}. \tag{2.31}$$

Similarly, the BT rate region for this case is defined as

$$\boldsymbol{\mathcal{R}}^{\mathrm{BT}}_{\boldsymbol{X}_{\mathcal{K}}}(\mathbb{T}, D) = \mathrm{cl}\left\{\bigcup_{U_{\mathcal{L}} \in \mathcal{U}(\mathbb{T}, D)} \left\{(R_1, \ldots, R_L)^{\mathrm{T}} \middle| \sum_{i \in \mathcal{A}} R_i \geq I\left(Y_{\mathcal{A}}; U_{\mathcal{A}} | U_{\mathcal{L}-\mathcal{A}}\right)\right\}\right\}, \tag{2.32}$$

where $\mathcal{A} \subseteq \mathcal{L}$, and $\mathcal{U}(\mathbb{T}, D)$ is the set of all length-L auxiliary random vectors $U_{\mathcal{L}} = (U_1, \ldots, U_L)^{\mathrm{T}}$ such that $U_i = Y_i + Q_i$, for $i = 1, 2, \ldots, L$, and $\sum_{i=1}^{K} \mathrm{E}\left\{(X_i - \mathrm{E}(X_i | U_{\mathcal{L}}))^2\right\} \leq D$, where $Q_i \sim \mathcal{N}\left(0, \sigma^2_{Q_i}\right)$ and Q_i's are independent to each other as well as to all Y_i's.

Oohama [9, 10] and Yang *et al.* [8] provided sufficient conditions for rate region tightness. For example, assume two independent remote sources[1]

$$(X_1, X_2) \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\right), \tag{2.33}$$

with target distortion $D = 0.88$, transform matrix

$$\boldsymbol{H} = \begin{pmatrix} -\frac{\sqrt{3}}{3} & 0 \\ -\frac{\sqrt{3}}{3} & -\frac{\sqrt{2}}{2} \\ -\frac{\sqrt{3}}{3} & \frac{\sqrt{2}}{2} \end{pmatrix}, \tag{2.34}$$

and observation noise covariance matrix

$$\boldsymbol{\Sigma}_{\boldsymbol{N}_{\{1,2,3\}}} = \mathrm{diag}(0.5, 0.6, 0.7). \tag{2.35}$$

According to [8, 9, 10], the BT rate region shown in Fig. 5 for this case is tight. Consequently, the BT minimum sum-rate is tight as well and can be calculated by (2.10) as

$$R_{\boldsymbol{X}_{\mathcal{K}}}(\mathbb{T}, D) = R^{\mathrm{BT}}_{\boldsymbol{X}_{\mathcal{K}}}(\mathbb{T}, D) = 8.948\,\mathrm{b/s}. \tag{2.36}$$

B. The two-terminal source coding problem with a matrix-distortion constraint

In order to go beyond Wang *et al.*'s sufficient condition [13], which assumes independent observation noises as seen in (2.17) and is derived using classical Gaussian rate-distortion function, in this paper we allow 2×2 block-correlation among the observation noises. Consequently, the derivation of the new lower bound requires us to consider a variant of the two-terminal source coding problem where the two individual

[1] Correlation between remote sources can always be absorbed into the transformation matrix $\boldsymbol{H}$.

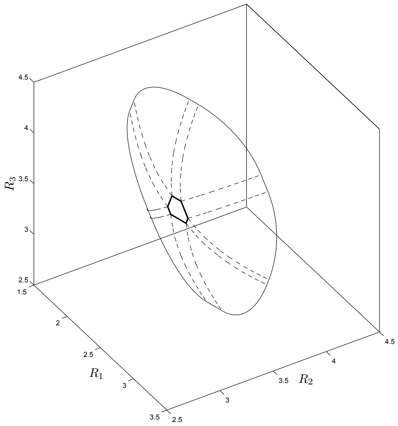


Fig. 5. The BT rate region of generalized quadratic Gaussian CEO problem with $K = 2$ remote sources and $L = 3$ observations, as defined by (2.33), (2.34) and (2.35).

distortion constraints are replaced by a 2×2 matrix-distortion constraint. Although the original quadratic Gaussian two-terminal source coding problem is completely solved [12, 34], due to the different distortion constraints, the exact achievable rate region for the matrix-distortion constrained two-terminal problem is still unknown. In this section, we derive a lower bound on the sum-rate of the matrix-distortion constrained two-terminal problem, which serves as the key to our main results given in the next section.

Assume that length-n blocks of Gaussian sources $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$ are separated compressed at the two encoders, while the decoder tries to reconstruct $\boldsymbol{Y}_{\mathcal{L}}$ such that

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}\left[(Y_{\mathcal{L},i} - \hat{Y}_{\mathcal{L},i})(Y_{\mathcal{L},i} - \hat{Y}_{\mathcal{L},i})^{\mathrm{T}}\right] \preceq \boldsymbol{D}_2 = \begin{bmatrix} D_1 & \theta\sqrt{D_1 D_2} \\ \theta\sqrt{D_1 D_2} & D_2 \end{bmatrix}, \tag{2.37}$$

where $\boldsymbol{A} \preceq \boldsymbol{B}$ means $\boldsymbol{B} - \boldsymbol{A}$ is a p.s.d. matrix, and denote the minimum sum-rate of such a problem as $R^{\mathrm{MT}}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}_2)$. Compared to the original quadratic Gaussian two-terminal source coding problem with individual distortion constraints, we have

$$R^{\mathrm{MT}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}\left((D_1, D_2)^{\mathrm{T}}\right) = \inf_{\theta\in[-1,1]} R^{\mathrm{MT}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}\left(\left[\begin{array}{cc} D_1 & \theta\sqrt{D_1 D_2} \\ \theta\sqrt{D_1 D_2} & D_2 \end{array}\right]\right). \tag{2.38}$$

Although Wagner *et al.*'s paper [12] focused on the original quadratic Gaussian two-terminal source coding problem, their converse proof has already explored the relationship in (2.38) to some extent, and provided a composite lower bound on the sum-rate of the two-terminal source coding problem with matrix-distortion constraint, namely,

$$R^{\mathrm{MT}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}_2) \geq \max\left\{R^{\mathrm{coop}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}_2), R^{\mu}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}_2)\right\}, \tag{2.39}$$

where

$$R^{\mathrm{coop}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}_2) = \frac{1}{2}\log\frac{|\boldsymbol{\Sigma}_{\boldsymbol{Y}}|}{|\boldsymbol{D}_2|},\ \ R^{\mu}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}_2) = R_{\boldsymbol{\Sigma}_{\boldsymbol{Y}},\mu}(\tilde{\mu}^{\mathrm{T}}\boldsymbol{D}_2\tilde{\mu}),$$

and $\tilde{\mu} = \left(\sqrt{D_2}, \sqrt{D_1}\right)^{\mathrm{T}}$, and $R_{\boldsymbol{\Sigma}_{\boldsymbol{Y}},\mu}(d)$ denotes the minimum sum-rate of the μ-sum problem with target distortion d.

We now give the exact form of a new lower bound that is inspired by Wang *et al.*'s work [13] and partially tighter than Wagner *et al.*'s bound in (2.39). Note that there is no loss in assuming that the correlation coefficient ρ between Y_1 and Y_2 is non-negative.

Lemma 4. *For any pair of* 2×2 *matrices*

$$\boldsymbol{\Sigma}_{\boldsymbol{Y}} = \left[\begin{array}{cc} \sigma^2_{Y_1} & \rho\sigma_{Y_1}\sigma_{Y_2} \\ \rho\sigma_{Y_1}\sigma_{Y_2} & \sigma^2_{Y_2} \end{array}\right], \tag{2.40}$$

$$\boldsymbol{D}_2 = \left[\begin{array}{cc} D_1 & \theta\sqrt{D_1 D_2} \\ \theta\sqrt{D_1 D_2} & D_2 \end{array}\right] \tag{2.41}$$

such that

$$\rho \geq 0, \text{ and } \boldsymbol{D}_2 \preceq \boldsymbol{\Sigma_Y}, \tag{2.42}$$

it holds that

$$\begin{aligned} R^{\mathrm{MT}}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}_2) &\geq \underline{R}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}_2) \\ &\triangleq \max\left\{R_{\mathrm{lb}}(\boldsymbol{\Sigma_Y}, \boldsymbol{D}_2), R_{\mu}(\boldsymbol{\Sigma_Y}, \boldsymbol{D}_2)\right\} \\ &= \begin{cases} R_{\mu}(\boldsymbol{\Sigma_Y}, \boldsymbol{D}_2) & \theta \leq \tilde{\theta} \\ R_{\mathrm{lb}}(\boldsymbol{\Sigma_Y}, \boldsymbol{D}_2) & \theta > \tilde{\theta} \end{cases}, \end{aligned} \tag{2.43}$$

where

$$\begin{aligned} R_{\mu}(\boldsymbol{\Sigma_Y}, \boldsymbol{D}_2) &= \frac{1}{2}\log\frac{v_1v_2(v_1v_2(1-\rho^2)+2\rho(1+\theta))}{(1+\theta)^2} \\ R_{\mathrm{lb}}(\boldsymbol{\Sigma_Y}, \boldsymbol{D}_2) &= \frac{1}{2}\log\frac{v_1^3v_2^3(1-\rho^2)^2}{(1-\theta)^2(v_1v_2(1-\rho^2)+2\rho(1+\theta))}, \end{aligned} \tag{2.44}$$

with $v_1 = \frac{\sigma_{Y_1}}{\sqrt{D_1}}$, $v_2 = \frac{\sigma_{Y_2}}{\sqrt{D_2}}$, *and*

$$\tilde{\theta} \triangleq \frac{\sqrt{v_1^2v_2^2(1-\rho^2)^2+4\rho^2}-v_1v_2(1-\rho^2)}{2\rho}. \tag{2.45}$$

Particularly, if $\theta \leq \tilde{\theta}$, *the lower bound is tight, i.e.,* $R^{\mathrm{MT}}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}_2) = \underline{R}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}_2)$.

Proof. Before proving Lemma 4, we define an equivalent two-terminal problem, with

$$\boldsymbol{\Sigma_Y} = \begin{bmatrix} v_1^2 & \rho v_1 v_2 \\ \rho v_1 v_2 & v_2^2 \end{bmatrix}, \quad \text{and } \boldsymbol{D}_2 = \begin{bmatrix} 1 & \theta \\ \theta & 1 \end{bmatrix}. \tag{2.46}$$

Then we need to prove $R^{\mathrm{MT}}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}_2) \geq R_{\mu}(\boldsymbol{\Sigma_Y}, \boldsymbol{D}_2)$ and $R^{\mathrm{MT}}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}_2) \geq R_{\mathrm{lb}}(\boldsymbol{\Sigma_Y}, \boldsymbol{D}_2)$.

To prove Lemma 4, let

$$X = Y_1 + Y_2 + Z, \tag{2.47}$$

where $Z \sim \mathcal{N}(0, \sigma_Z^2)$ with $\sigma_Z^2 = \frac{v_1 v_2 (1-\rho^2)}{\rho}$. Then the variance of X can be computed as $\sigma_X^2 = \frac{(v_1+v_2\rho)(v_2+v_1\rho)}{\rho}$, and it can be easily verified that

$$Y_{\mathcal{L}} \;=\; [\alpha_1, \alpha_2]^{\mathrm{T}} \cdot X + \left[\tilde{N}_1, \tilde{N}_2\right]^{\mathrm{T}}, \tag{2.48}$$

with $\alpha_1 = \frac{v_1\rho}{v_2+v_1\rho}$, $\alpha_2 = \frac{v_2\rho}{v_1+v_2\rho}$, $\left[\tilde{N}_1, \tilde{N}_2\right]^{\mathrm{T}} \sim \mathcal{N}(\mathbf{0}, \mathrm{diag}(\sigma^2_{\tilde{N}_1}, \sigma^2_{\tilde{N}_2}))$, and $\sigma^2_{\tilde{N}_1} = \frac{v_1^2 v_2(1-\rho^2)}{v_2+v_1\rho}$, $\sigma^2_{\tilde{N}_2} = \frac{v_2^2 v_1(1-\rho^2)}{v_1+v_2\rho}$. Moreover, any scheme that achieves a distortion matrix $\boldsymbol{D}_2$ on $Y_{\mathcal{L}}$ must be able to achieve a distortion of $[1\ 1] \cdot \boldsymbol{D}_2 \cdot [1\ 1]^{\mathrm{T}} + \sigma_Z^2$ on X.

Hence

$$\begin{aligned}
H(W_{\mathcal{L}}) &= I(\boldsymbol{Y}_{\mathcal{L}}, \boldsymbol{X}; W_{\mathcal{L}}) \\
&= I(\boldsymbol{X}; W_{\mathcal{L}}) + \sum_{i=1}^{2} I(\boldsymbol{Y}_i; W_i | \boldsymbol{X}) \qquad (2.49) \\
&= h(\boldsymbol{X}) - h(\boldsymbol{X}|W_{\mathcal{L}}) + \sum_{i=1}^{2} h(\boldsymbol{Y}_i|\boldsymbol{X}) - h(\boldsymbol{Y}_i; W_i|\boldsymbol{X}) \\
&\geq \frac{n}{2}\log\frac{\sigma_X^2}{[1\ 1]\cdot\boldsymbol{D}_2\cdot[1\ 1]^{\mathrm{T}} + \sigma_Z^2} + \frac{n}{2}\log\frac{\sigma^2_{\tilde{N}_1}\sigma^2_{\tilde{N}_2}}{\gamma_1\gamma_2} \qquad (2.50) \\
&\geq \frac{n}{2}\log\frac{\sigma_X^2}{2+2\theta+\frac{v_1v_2(1-\rho^2)}{\rho}} + \frac{n}{2}\log\frac{v_1^3 v_2^3(1-\rho^2)^2}{(v_2+v_1\rho)(v_1+v_2\rho)\gamma_1\gamma_2},
\end{aligned}$$

where (2.49) uses the fact that $W_i \to Y_i^n \to \boldsymbol{X} \to (Y_j^n, W_j)$ form a Markov chain for any $i, j \in \{1, 2\}$ and $i \neq j$, in (2.50) we define $\gamma_i \triangleq \frac{1}{n}\sum_{j=1}^{n} \mathrm{var}(Y_{i,j}|W_i, \boldsymbol{X})$ and use the fact that Gaussian random variables maximize entropy over those with a fixed variance.

On the other hand, due to [13, Lemma 1], we known that $\frac{1}{n}\sum_{i=1}^{n} \mathrm{cov}(\boldsymbol{Y}_{\mathcal{L},i}|X_i, W_{\mathcal{L}}) = \mathrm{diag}(\gamma_1, \gamma_2)$. Then [13, Lemma 3] implies that

$$\frac{1}{n}\sum_{i=1}^{n} \mathrm{cov}(\boldsymbol{Y}_{\mathcal{L},i}|X_i, W_{\mathcal{L}}) \preceq \left(\boldsymbol{D}_2^{-1} + \boldsymbol{\Sigma}^{-1}_{\tilde{\boldsymbol{N}}_{\mathcal{L}}} - \boldsymbol{\Sigma}^{-1}_{\boldsymbol{Y}}\right)^{-1}, \tag{2.51}$$

with $\Sigma_{\tilde{N}_{\mathcal{L}}} = \operatorname{diag}\left(\sigma^2_{\tilde{N}_1}, \sigma^2_{\tilde{N}_2}\right)$, i.e.,

$$\operatorname{diag}(\gamma_1, \gamma_2) \preceq \left(\begin{bmatrix} 1 & \theta \\ \theta & 1 \end{bmatrix}^{-1} + \frac{\rho}{v_1 v_2 (1-\rho^2)} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right)^{-1}, \tag{2.52}$$

which can be combined with (2.50) to form a semi-definite optimization problem that minimizes

$$\mathcal{F}(\gamma_1, \gamma_2) \;\triangleq\; \frac{1}{2} \log \frac{1}{\gamma_1 \gamma_2} \tag{2.53}$$

over γ_1 and γ_2 subject to

$$\mathcal{G}(\gamma_1, \gamma_2) \;\triangleq\; \begin{bmatrix} 1 & \theta \\ \theta & 1 \end{bmatrix}^{-1} + \frac{\rho}{v_1 v_2 (1-\rho^2)} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} - \operatorname{diag}(\gamma_1^{-1}, \gamma_2^{-1}) \preceq \mathbf{0}. \tag{2.54}$$

The Lagrangian is

$$\mathbb{L}(\gamma_1, \gamma_2) = \mathcal{F}(\gamma_1, \gamma_2) + \operatorname{tr}(\Lambda \mathcal{G}(\gamma_1, \gamma_2)), \tag{2.55}$$

where Λ is a p.s.d. matrix. Then the KKT condition is given by

$$\begin{aligned} \nabla_{\gamma_i} \mathbb{L}(\gamma_1, \gamma_2) &= 0, \quad i = 1, 2, && (2.56) \\ \mathcal{G}(\gamma_1, \gamma_2) &\preceq \mathbf{0}, && (2.57) \\ \Lambda \mathcal{G}(\gamma_1, \gamma_2) &= \mathbf{0}. && (2.58) \end{aligned}$$

Solving the (2.56) and (2.58), we get two sets of solutions, namely,

$$\gamma_1 = 1-\theta,\ \gamma_2 = 1-\theta,\ \Lambda = \frac{1-\theta}{2} \cdot \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \tag{2.59}$$

and

$$\gamma_1 = \frac{v_1 v_2(1-\rho^2)(1+\theta)}{v_1 v_2(1-\rho^2)+2\rho(1+\theta)},\ \gamma_2 = \frac{v_1 v_2(1-\rho^2)(1+\theta)}{v_1 v_2(1-\rho^2)+2\rho(1+\theta)},$$
$$\Lambda = \frac{v_1 v_2(1-\rho^2)(1+\theta)}{v_1 v_2(1-\rho^2)+2\rho(1+\theta)} \cdot \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}. \tag{2.60}$$

Then it is easy to verify that the first set of solution satisfies (2.57) if $\theta \geq \tilde{\theta}$, while the second set of solution satisfies (2.57) if $\theta \leq \tilde{\theta}$. Hence the optimal solutions of γ_1 and γ_2 are

$$\gamma_1 = \gamma_2 \quad = \quad \begin{cases} \frac{v_1 v_2(1-\rho^2)(1+\theta)}{v_1 v_2(1-\rho^2)+2\rho(1+\theta)} & \theta \leq \tilde{\theta} \\ 1-\theta & \theta > \tilde{\theta} \end{cases}, \tag{2.61}$$

which directly lead to (2.43).

To prove tightness of the lower bound $\underline{R}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}_2)$ when $\theta \leq \tilde{\theta}$, we construct a BT scheme with distortion matrix

$$\tilde{\boldsymbol{D}}_2 = (\boldsymbol{\Sigma_Y}^{-1} + \mathrm{diag}(q_1, q_2)^{-1})^{-1} = \begin{bmatrix} \frac{(1+\theta)(v_1 v_2(1-\rho^2)+\rho(1+\theta))}{(v_1 v_2(1-\rho^2)+2\rho(1+\theta))} & \frac{\rho(1+\theta)^2}{(v_1 v_2(1-\rho^2)+2\rho(1+\theta))} \\ \frac{\rho(1+\theta)^2}{(v_1 v_2(1-\rho^2)+2\rho(1+\theta))} & \frac{(1+\theta)(v_1 v_2(1-\rho^2)+\rho(1+\theta))}{(v_1 v_2(1-\rho^2)+2\rho(1+\theta))} \end{bmatrix},$$

and sum-rate

$$\frac{1}{2}\log\frac{|\boldsymbol{\Sigma_Y}|}{|\tilde{\boldsymbol{D}}_2|} = \frac{1}{2}\log\frac{v_1^2 v_2^2(1-\rho^2)}{\frac{v_1 v_2(1+\theta)^2(1-\rho^2)}{(v_1 v_2(1-\rho^2)+2\rho(1+\theta))}} = \underline{R}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}_2), \tag{2.62}$$

where

$$q_1 = \frac{v_1^2 v_2(1-\rho^2)(1+\theta)}{v_1^2 v_2(1-\rho^2)-(v_2-\rho v_1)(1+\theta)},\ q_2 = \frac{v_1 v_2^2(1-\rho^2)(1+\theta)}{v_1 v_2^2(1-\rho^2)-(v_1-\rho v_2)(1+\theta)}.$$

Then tightness is proved by verifying

$$\boldsymbol{D}_2 - \tilde{\boldsymbol{D}}_2 = \frac{\rho\left(1-\theta^2\right) - v_1 v_2 \theta \left(1-\rho^2\right)}{\left(v_1 v_2 \left(1-\rho^2\right) + 2\rho\left(1+\theta\right)\right)} \cdot \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \succeq \mathbf{0}, \tag{2.63}$$

where the last matrix inequality is due to the facts that $f_1(\theta) \triangleq (v_1v_2(1-\rho^2) + 2\rho(1+\theta)) > 0$, $f_2(\theta) \triangleq \rho(1-\theta^2) - v_1v_2\theta(1-\rho^2)$ is monotone decreasing in the range $\theta \in [-1, \tilde{\theta})$, $f_2(\tilde{\theta}) = 0$, and the assumption that $\theta \le \tilde{\theta}$. □

Note that unlike the original two-terminal problem, the new lower bound $\underline{R}^{\text{sum}}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}_2)$ does not always meet the BT upper bound, which is given by

$$R^{\text{BT}}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}_2) = \max\left\{R_{\text{lb}}(\boldsymbol{\Sigma_Y}, \boldsymbol{D}_2), R^{\mu}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}_2)\right\} = \begin{cases} R^{\mu}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}_2) & \theta \le \tilde{\theta} \\ R_{\text{ub}}(\boldsymbol{\Sigma_Y}, \boldsymbol{D}_2) & \theta > \tilde{\theta} \end{cases} \tag{2.64}$$

with

$$R_{\text{ub}}(\boldsymbol{\Sigma_Y}, \boldsymbol{D}_2) = \frac{1}{2}\log\frac{v_1v_2(v_1v_2(1-\rho^2) - 2\rho(1-\theta))}{(1-\theta)^2}. \tag{2.65}$$

Obviously, if $\theta > \tilde{\theta}$, the two bounds do not coincide, and we can easily compute the gap between them as

$$\begin{aligned} R^{\Delta}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}_2) &\triangleq \underline{R}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}_2) - R^{\text{BT}}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}_2) \\ &= R_{\text{ub}}(\boldsymbol{\Sigma_Y}, \boldsymbol{D}_2) - R_{\text{lb}}(\boldsymbol{\Sigma_Y}, \boldsymbol{D}_2) \\ &= \frac{1}{2}\log\frac{(v_1v_2(1-\rho^2) - 2\rho(1-\theta))(v_1v_2(1-\rho^2) + 2\rho(1+\theta))}{v_1^2v_2^2(1-\rho^2)^2}. \end{aligned} \tag{2.66}$$

To evaluate the maximum value of $R^{\Delta}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}_2)$, we compute the feasible range of θ, which is constrained by the assumption $\boldsymbol{D}_2 \preceq \boldsymbol{\Sigma_Y}$, and given by $\theta \in (\underline{\theta}, \overline{\theta})$ with

$$\underline{\theta} = \max\left\{-1, -\sqrt{(v_1^2-1)(v_2^2-1)} - \rho v_1 v_2\right\}, \ \overline{\theta} = \min\left\{1, \sqrt{(v_1^2-1)(v_2^2-1)} + \rho v_1 v_2\right\}. \tag{2.67}$$

Now due to the assumption that $\rho \geq 0$, $R^{\Delta}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}_2)$ is monotone increasing in θ in the range $(\tilde{\theta}, \overline{\theta})$. Hence

$$\sup_{\theta \in (\tilde{\theta}, \overline{\theta})} R^{\Delta}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}_2) = \lim_{\theta \to \overline{\theta}} R^{\Delta}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}_2) \leq \lim_{\theta \to 1} R^{\Delta}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}_2) = \frac{1}{2} \log \left(1 + \frac{4\rho}{v_1 v_2 (1 - \rho^2)} \right). \tag{2.68}$$

We thus conclude that although the lower bound $\underline{R}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}_2)$ is not always tight, the gap to the upper bound $R^{\mathrm{BT}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}_2)$ cannot exceed a certain threshold that depends only on v_1, v_2, and ρ.

On the other hand, if we calculate the improvement from Wagner *et al.*'s lower bound (2.39) to our new one $\underline{R}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}_2)$ with $\theta \in (\tilde{\theta}, \overline{\theta})$, we obtain

$$\underline{R}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}_2) - \max\left\{ R^{\mathrm{coop}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}_2), R^{\mu}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}_2) \right\} = \frac{1}{2} \log \frac{(v_1 v_2 (1 + \theta)(1 - \rho^2)}{(1 - \theta)(v_1 v_2 (1 - \rho^2) + 2\rho(1 + \theta))}, \tag{2.69}$$

which obviously goes to infinity as $\theta \to 1$, this means that the improvement can be infinitely large for any value of v_1, v_2, and ρ such that $\overline{\theta}$ defined in (2.67) equals to one.

A comparison among Wagner's lower bound [12], our partially improved lower bound, and the BT upper bound with $\sigma^2_{Y_1} = \sigma^2_{Y_2} = 1$, $\rho = 0.9$, $D_1 = 0.1$, $D_2 = 0.05$ is shown in Fig. 6. We can clearly observe that the gap from our new lower bound to the BT upper bound is much smaller than that to the lower bound in [12].

C. Main results

1. Definitions and preliminaries

Before stating our main results, we need to give some definitions and review the subgradient-based KKT condition.

Let $\pi = \{\pi_1, ..., \pi_L\}$ be a permutation of $\boldsymbol{\mathcal{L}}$, and   be the corresponding $L \times L$

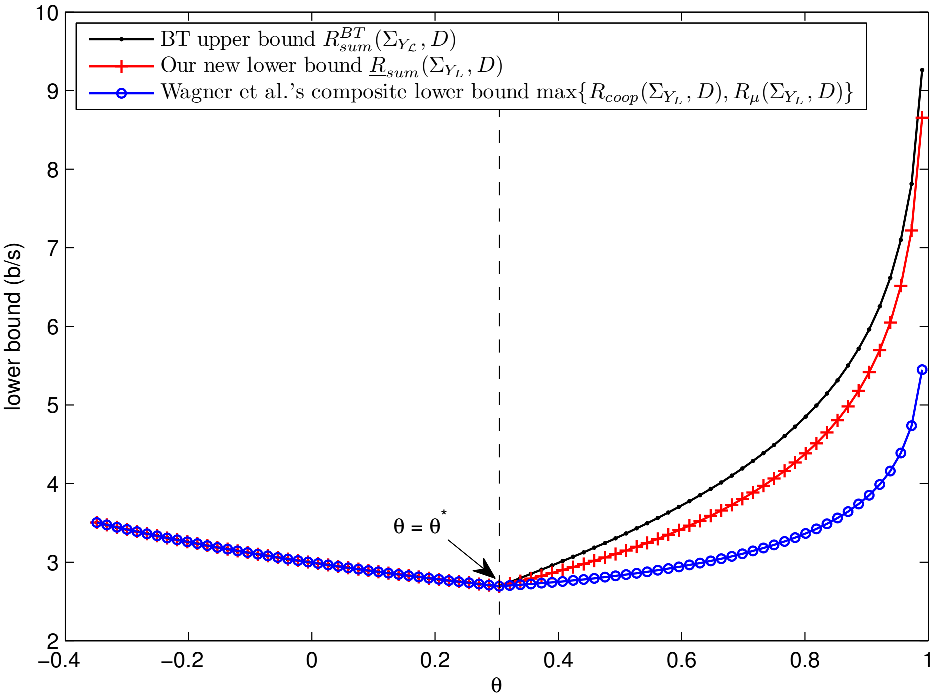


Fig. 6. Comparison among Wagner's lower bound [12], our partially improved lower bound, and the BT upper bound.

permutation matrix such that $\mathcal{L} = \pi$. We say an $L \times L$ matrix Σ is $\pi^{(K)}$ block-diagonal if it is symmetric and can be written as

$$\Sigma = \quad \cdot \begin{bmatrix} a_{1,1} & a_{1,2} & 0 & 0 & \dots & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ a_{1,2} & a_{2,2} & 0 & 0 & \dots & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & a_{3,3} & a_{3,4} & \dots & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & a_{3,4} & a_{4,4} & \dots & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & & \\ 0 & 0 & 0 & 0 & \dots & \dots & a_{2K-1,2K-1} & a_{2K-1,2K} & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & \dots & a_{2K-1,2K} & a_{2K,2K} & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & \dots & 0 & 0 & a_{2K+1} & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & \dots & 0 & 0 & 0 & \dots & a_{L-1} & 0 \\ 0 & 0 & 0 & 0 & \dots & \dots & 0 & 0 & 0 & \dots & 0 & a_L \end{bmatrix} ^{\mathrm{T}}, \quad (2.70)$$

and denote $\Upsilon_K(\pi)$ as the set of all $\pi^{(K)}$ block-diagonal matrices. Equivalently,

$\boldsymbol{\Sigma} \in \Upsilon_K(\pi)$ if and only if $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}^{\mathrm{T}}$ and

$$\boldsymbol{\Sigma}_{\pi_i,\pi_j} = 0 \text{ if } \begin{cases} i,j \in \{1,2,...,2K\} \text{ s.t. } \left\lceil \frac{i}{2} \right\rceil \neq \left\lceil \frac{j}{2} \right\rceil, \\ i,j \in \{2K+1, 2K+2, ..., L\} \text{ s.t. } i \neq j, \\ i \in \{2K+1, 2K+2, ..., L\} \text{ and } j \in \{1,2,...,2K\}, \\ i \in \{1,2,...,2K\} \text{ and } j \in \{2K+1, 2K+2, ..., L\}. \end{cases} \tag{2.71}$$

Comparing (2.70) and (2.17), it is clear that all diagonal matrices are also $\pi^{(K)}$ block-diagonal, but the converse is not true for $K \geq 1$, i.e.,

$$\mathbb{d} \subsetneq \Upsilon_K(\pi) \text{ for } 1 \leq K \leq \left\lfloor \frac{L}{2} \right\rfloor \text{ and any permutation } \pi. \tag{2.72}$$

Consequently, if we define

$$\mathscr{N}_{\pi^{(K)}}(\boldsymbol{\Sigma_Y}) \triangleq \left\{ \boldsymbol{\Sigma} \in \mathscr{P}^{\succeq} \cap \Upsilon_K(\pi) : \boldsymbol{\Sigma} \succeq \boldsymbol{\Sigma_Y} \right\}, \tag{2.73}$$

and compare with $\mathscr{N}(\boldsymbol{\Sigma_Y})$ defined in (2.19), it holds that

$$\mathscr{N}(\boldsymbol{\Sigma_Y}) \;=\; \mathscr{N}_{\mathbb{I}^{(0)}}(\boldsymbol{\Sigma_Y}) \;\subseteq\; \mathscr{N}_{\pi^{(K)}}(\boldsymbol{\Sigma_Y}) \tag{2.74}$$

for any $0 \leq K \leq \left\lfloor \frac{L}{2} \right\rfloor$ and permutation π, where $\mathbb{I}$ denotes the identity permutation that maps $\boldsymbol{\mathcal{L}}$ to itself.

For a set of L Gaussian sources $\boldsymbol{Y}$ and a $\boldsymbol{\Sigma_N} \in \Upsilon_K(\pi)$ such that $\boldsymbol{\Sigma_N} \preceq \boldsymbol{\Sigma_Y}$, let $M = \mathrm{rank}(\boldsymbol{\Sigma_Y} - \boldsymbol{\Sigma_N})$ and the singular value decomposition of $\boldsymbol{\Sigma_Y} - \boldsymbol{\Sigma_N}$ be

$$\boldsymbol{\Sigma_Y} - \boldsymbol{\Sigma_N} = \boldsymbol{T}^{\mathrm{T}} \mathrm{diag}(\sigma^2_{X_1}, \sigma^2_{X_2}, ..., \sigma^2_{X_M}, 0, ..., 0) \boldsymbol{T}. \tag{2.75}$$

Then define $\Sigma_{\boldsymbol{X}_\mathcal{M}} = \mathrm{diag}(\sigma^2_{X_1}, \sigma^2_{X_2}, ..., \sigma^2_{X_M})$, $\boldsymbol{H} = \boldsymbol{T}_{\mathcal{M},\mathcal{L}}$, and let

$$\boldsymbol{X}_\mathcal{M} \triangleq \boldsymbol{A}\boldsymbol{Y} + \boldsymbol{Z}_\mathcal{L}, \tag{2.76}$$

with $Z_{\mathcal{L}} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{B})$ independent of $\boldsymbol{Y}$, where

$$\boldsymbol{A} = \boldsymbol{\Sigma}_{\boldsymbol{X}_{\mathcal{M}}} \boldsymbol{H} \boldsymbol{\Sigma}_{\boldsymbol{Y}}{}^{-1},\ \boldsymbol{B} = \boldsymbol{\Sigma}_{\boldsymbol{X}_{\mathcal{M}}} - \boldsymbol{\Sigma}_{\boldsymbol{X}_{\mathcal{M}}} \boldsymbol{H} \boldsymbol{\Sigma}_{\boldsymbol{Y}}{}^{-1} \boldsymbol{H}^{\mathrm{T}} \boldsymbol{\Sigma}_{\boldsymbol{X}_{\mathcal{M}}}. \tag{2.77}$$

It is trivial to verify that the M Gaussian *remote sources* $\boldsymbol{X}_{\mathcal{M}} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{X}_{\mathcal{M}}})$ satisfy

$$\boldsymbol{Y} = \boldsymbol{H}^{\mathrm{T}} \boldsymbol{X}_{\mathcal{M}} + \boldsymbol{N}, \tag{2.78}$$

with the L *observation noises* $N_{\mathcal{L}} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{N}})$ independent of $X_{\mathcal{M}}$.

Next, we briefly review the subgradient-based KKT conditions for non-differentiable convex optimization problems. The original KKT condition is a necessary condition for global optimality in a convex optimization problem with differentiable objective function and equality/inequality constraints. However, when dealing with non-differentiable convex optimization problems, subgradient-based KKT condition has to be used instead. We call $\boldsymbol{g}$ a subgradient [37] of a non-differentiable scalar-valued vector function f at point $\boldsymbol{x}$, if

$$f(\boldsymbol{y}) \quad \geq \quad f(\boldsymbol{x}) + \boldsymbol{g}^{\mathrm{T}}(\boldsymbol{y} - \boldsymbol{x}) \text{ for all } \boldsymbol{y}. \tag{2.79}$$

In particular, if $f = \max\{f_1, f_2\}$ with f_1 and f_2 being convex and differentiable such that $f_1(\boldsymbol{x}_0) = f_2(\boldsymbol{x}_0)$, then the subgradients of f at $\boldsymbol{x}_0$ form a line segment between $\nabla f_1(\boldsymbol{x}_0)$ and $\nabla f_2(\boldsymbol{x}_0)$. The set of all subgradients of a function f at some point $\boldsymbol{x}$ is called the subdifferential of f at $\boldsymbol{x}$, and denoted as $\partial f(\boldsymbol{x})$. The subdifferential of $\underline{R}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{\Gamma})$ is given in the following lemma.

Lemma 5. *Assume that $\boldsymbol{\Sigma}_{\boldsymbol{Y}}$ and $\boldsymbol{D}_2$ take forms of (2.40) and (2.41), respectively, such that $\boldsymbol{D}_2 \preceq \boldsymbol{\Sigma}_{\boldsymbol{Y}}$. Then the subdifferential of $\underline{R}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}_2)$ (as a function of $\boldsymbol{D}_2$) at*

$$\boldsymbol{D}_2 = \tilde{\boldsymbol{D}}_2 \ \triangleq \ \begin{bmatrix} D_1 & \tilde{\theta}\sqrt{D_1 D_2} \\ \tilde{\theta}\sqrt{D_1 D_2} & D_2 \end{bmatrix} \tag{2.80}$$

is a line segment

$$\partial \underline{R}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}_2) \mid_{\boldsymbol{D}_2=\tilde{\boldsymbol{D}}_2} = \left\{ -\frac{1}{2} \tilde{\boldsymbol{D}}_2^{-1} \Psi \ \tilde{\boldsymbol{D}}_2^{\ -1} : \Psi \in \ {}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}) \right\},$$

where

$$ {}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}) \triangleq \left\{ \begin{bmatrix} D_1 & (\alpha + (1-\alpha)(2|\tilde{\theta}| - 1))s\sqrt{D_1 D_2} \\ (\alpha + (1-\alpha)(2|\tilde{\theta}| - 1))s\sqrt{D_1 D_2} & D_2 \end{bmatrix} : \alpha \in [0,1] \right\},$$

with $\tilde{\theta}$ *defined in (2.45) and* $s \triangleq \mathrm{sign}(\tilde{\theta})$.

Proof. First, due to the assumption that $\tilde{\boldsymbol{D}}_2^{-1} - \boldsymbol{\Sigma}_{\boldsymbol{Y}}^{-1}$ is a p.s.d. diagonal matrix, we must have

$$\theta \ = \ \begin{cases} \frac{\sqrt{1-2\rho^2+\rho^4+4\rho^2 d_1^2 d_2^2}-(1-\rho^2)}{2\rho d_1 d_2} & \rho \geq 0 \\ \frac{-\sqrt{1-2\rho^2+\rho^4+4\rho^2 d_1^2 d_2^2}-(1-\rho^2)}{2\rho d_1 d_2} & \rho < 0 \end{cases}, \tag{2.81}$$

with $d_1 = \sqrt{D_1}$ and $d_2 = \sqrt{D_2}$. Now since

$$\underline{R}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}_2) = \max\left\{ R_{\mathrm{lb}}(\boldsymbol{\Sigma_Y}, \boldsymbol{D}_2), R^{\mu}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}_2) \right\},$$

we compute

$$\nabla_{\boldsymbol{D}_2} R_{\mathrm{lb}}(\boldsymbol{\Sigma_Y}, \boldsymbol{D}_2) \mid_{\boldsymbol{D}_2=\tilde{\boldsymbol{D}}_2} = \kappa \cdot \begin{bmatrix} \frac{1}{D_1} & \frac{s(1-2|\theta|)}{\sqrt{D_1 D_2}} \\ \frac{s(1-2|\theta|)}{\sqrt{D_1 D_2}} & \frac{1}{D_2} \end{bmatrix},$$

$$\nabla_{\boldsymbol{D}_2} R^{\mu}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}_2) \mid_{\boldsymbol{D}_2=\tilde{\boldsymbol{D}}_2} = \chi \cdot \begin{bmatrix} \frac{1}{D_1} & \frac{s}{\sqrt{D_1 D_2}} \\ \frac{s}{\sqrt{D_1 D_2}} & \frac{1}{D_2} \end{bmatrix}, \tag{2.82}$$

where

$$\kappa = \frac{\rho^4 - 2d_1d_2\rho^3 - 2\rho^2 + 4\rho^2 d_1^2 d_2^2 + 2d_1d_2\rho + 1}{2(1-\rho^2)^2} - \frac{\rho^2 + 2d_1d_2\rho - 1}{2(1-\rho^2)^2}\sqrt{1 - 2\rho^2 + \rho^4 + 4\rho^2 d_1^2 d_2^2},$$
$$\chi = -\frac{\rho^4 + 2d_2\rho^3 d_1 + 4\rho^2 d_2^2 d_1^2 - 2\rho^2 - 2d_2\rho d_1 + 1}{2(1-\rho^2)^2} + \frac{2d_1d_2\rho - 1 + \rho^2}{2(1-\rho^2)^2}\sqrt{1 - 2\rho^2 + \rho^4 + 4\rho^2 d_1^2 d_2^2}. \tag{2.83}$$

Finally, it is easy to verify that

$$\begin{aligned} -\tilde{\boldsymbol{D}}_2 \cdot \nabla_{\boldsymbol{D}_2} R_{\mathrm{lb}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D}_2) \mid_{\boldsymbol{D}_2 = \tilde{\boldsymbol{D}}_2} \cdot \tilde{\boldsymbol{D}}_2 &= \begin{bmatrix} D_1 & s(1-2|\theta|)\sqrt{D_1D_2} \\ s(1-2|\theta|)\sqrt{D_1D_2} & D_2 \end{bmatrix}, \\ -\tilde{\boldsymbol{D}}_2 \cdot \nabla_{\boldsymbol{D}_2} R^{\mu}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}_2) \mid_{\boldsymbol{D}_2 = \tilde{\boldsymbol{D}}_2} \cdot \tilde{\boldsymbol{D}}_2 &= \begin{bmatrix} D_1 & s\sqrt{D_1D_2} \\ s\sqrt{D_1D_2} & D_2 \end{bmatrix}, \end{aligned}$$

and Lemma 5 readily follows. □

For a convex optimization problem with objective function f, inequality constraints $g_i \leq 0$ for $j = 1, ..., m$ and equality constraints $h_j = 0$ for $j = 1, ..., l$, the global optimal point $\boldsymbol{x} = \boldsymbol{x}^*$ must satisfy

$$\begin{aligned} \boldsymbol{0} &\in \partial f(\boldsymbol{x}^*) + \sum_{i=1}^{m} \mu_i \partial g_i(\boldsymbol{x}^*) + \sum_{j=1}^{l} \lambda_j \partial h_i(\boldsymbol{x}^*), \\ g_i(\boldsymbol{x}^*) &\leq 0, i = 1, 2, ..., m, \\ h_j(\boldsymbol{x}^*) &= 0, j = 1, 2, ..., l, \\ \mu_i &\geq 0, i = 1, 2, ..., m, \\ \mu_i g_i(\boldsymbol{x}^*) &= 0, i = 1, 2, ..., m, \end{aligned}$$

for some μ_i's and λ_j's.

2. A new sufficient condition for sum-rate tightness

Now we are ready to state our main result on a new sufficient condition for the tightness of BT minimum sum-rate. Consider an MT source coding problem defined by $\boldsymbol{\Sigma}_{\boldsymbol{Y}}$ and $\boldsymbol{D}$. Denote the BT minimum sum-rate as $R^{\mathrm{BT}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D})$, and assume that the optimal BT scheme achieves a distortion matrix $\tilde{\boldsymbol{D}}_2$. The main result of this paper is given in the following theorem.

Theorem 2. *$R^{\mathrm{BT}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}) = R^{\mathrm{MT}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D})$ if there exists a permutation π, a $\pi^{(K)}$ block diagonal p.d. matrix $\boldsymbol{\Sigma}_{\boldsymbol{N}}$ such that $\boldsymbol{\Sigma}_{\boldsymbol{N}} \preceq \boldsymbol{\Sigma}_{\boldsymbol{Y}}$, an $L \times L$ p.s.d. matrix $\boldsymbol{\Omega}$, an $L \times L$ p.s.d. diagonal matrix $\boldsymbol{\Pi}$, and a set of K 2×2 p.s.d. matrices $\boldsymbol{\Theta}_j$, $j \in \mathcal{K}$ such that the following conditions are satisfied:*

$$\tilde{\boldsymbol{D}}_2 \left(\boldsymbol{\Pi} - \tilde{\boldsymbol{D}}_2^{-1} + \tilde{\boldsymbol{D}}_2^{-1} \left(\tilde{\boldsymbol{D}}_2^{-1} + \boldsymbol{\Sigma}_{\boldsymbol{N}}^{-1} - \boldsymbol{\Sigma}_{\boldsymbol{Y}}^{-1} \right)^{-1} \tilde{\boldsymbol{D}}_2^{-1} \right) \tilde{\boldsymbol{D}}_2 = \boldsymbol{\Lambda} - \boldsymbol{\Omega}, \tag{2.84}$$

$$\langle \boldsymbol{\Lambda} \rangle_j^{\pi} + \boldsymbol{\Theta}_j - \quad \left(\langle \boldsymbol{\Sigma}_{\boldsymbol{N}} \rangle_j^{\pi}, \mathrm{diag}(\langle \tilde{\boldsymbol{\Gamma}} \rangle_j^{\pi}) \right) \ni \boldsymbol{0}, \forall j \in \mathcal{K}, \tag{2.85}$$

$$\textit{for } k = 2K+1, ..., L, \quad [\boldsymbol{\Lambda}]_{\pi_k, \pi_k} = \left[\tilde{\boldsymbol{\Gamma}} \right]_{\pi_k, \pi_k}, \tag{2.86}$$

$$\boldsymbol{\Omega} \left(\boldsymbol{\Sigma}_{\boldsymbol{Y}}^{-1} - \tilde{\boldsymbol{D}}_2^{-1} \right) = \boldsymbol{0}, \tag{2.87}$$

$$\boldsymbol{\Theta}_j \left(\langle \boldsymbol{\Sigma}_{\boldsymbol{N}} \rangle_j^{\pi} - \langle \tilde{\boldsymbol{\Gamma}} \rangle_j^{\pi} \right) = \boldsymbol{0}, \forall j \in \mathcal{K}, \tag{2.88}$$

$$[\boldsymbol{\Pi}]_{j,j} \left(\left[\tilde{\boldsymbol{D}}_2 \right]_{j,j} - D_j \right) = 0, \forall j \in \mathcal{K}, \tag{2.89}$$

where $\langle \boldsymbol{C} \rangle_j^{\pi}$ denotes the 2×2 submatrix constructed from the (π_{2j-1}, π_{2j})-th row and (π_{2j-1}, π_{2j})-th column of $\boldsymbol{C}$, and

$$\tilde{\boldsymbol{\Gamma}} \triangleq \left(\tilde{\boldsymbol{D}}_2^{-1} + \boldsymbol{\Sigma}_{\boldsymbol{N}}^{-1} - \boldsymbol{\Sigma}_{\boldsymbol{Y}}^{-1} \right)^{-1}. \tag{2.90}$$

Proof. To prove Theorem 2, we need the following two lemmas.

Lemma 6. *For any random objects $\boldsymbol{Y}_{\mathcal{L}}$ and $\boldsymbol{X}_{\mathcal{M}}$, if*

$$[\text{cov}(\boldsymbol{Y}_{\mathcal{L}}|\boldsymbol{X}_{\mathcal{M}})]_{i,j} \quad = \quad 0 \tag{2.91}$$

for some $i, j \in \mathcal{L}$, then

$$[\text{cov}(\boldsymbol{Y}_{\mathcal{L}}|\boldsymbol{X}_{\mathcal{M}}, W_{\mathcal{L}})]_{i,j} \quad = \quad 0 \tag{2.92}$$

for any L functions $W_{\mathcal{L}} \triangleq \left\{\psi_1^{(n)}(\boldsymbol{Y}_1), \psi_2^{(n)}(\boldsymbol{Y}_2), ..., \psi_L^{(n)}(\boldsymbol{Y}_L)\right\}$.

Proof. To prove Lemma 6, we need to use [13, Lemma 1], which is stated in the following proposition for the sake of completion.

Proposition 1. *For integers n, m and random variables X and ω, let $\boldsymbol{X}$ be a row vector of n independent drawings of X, and $\boldsymbol{Y}(\omega)$ be any $1 \times m$ vector of measurable functions of ω. Then it holds that*

$$\text{E}\left[(\boldsymbol{X} - \text{E}(\boldsymbol{X}|\omega))^{\text{T}}\boldsymbol{Y}(\omega)\right] = \mathbf{0}_{n\times m}. \tag{2.93}$$

Now (2.91) and the definition of $W_{\mathcal{L}}$ imply that the Markov chains $W_i \to \boldsymbol{Y}_i \to \boldsymbol{X}_{\mathcal{M}} \to (\boldsymbol{Y}_j, W_j)$ and $W_j \to \boldsymbol{Y}_j \to \boldsymbol{X}_{\mathcal{M}} \to (\boldsymbol{Y}_i, W_i)$ hold. Hence (2.92) must hold since

$$\begin{aligned}
[\text{cov}(\boldsymbol{Y}_{\mathcal{L}}|\boldsymbol{X}_{\mathcal{M}}, W_{\mathcal{L}})]_{i,j} &= \text{E}\left[(\boldsymbol{Y}_i - \text{E}(\boldsymbol{Y}_i|\boldsymbol{X}_{\mathcal{M}}, W_{\mathcal{L}}))(\boldsymbol{Y}_j - \text{E}(\boldsymbol{Y}_j|\boldsymbol{X}_{\mathcal{M}}, W_{\mathcal{L}}))^{\text{T}}\right] \\
&= \text{E}\left[(\boldsymbol{Y}_i - \text{E}(\boldsymbol{Y}_i|\boldsymbol{X}_{\mathcal{M}}, W_i))(\boldsymbol{Y}_j - \text{E}(\boldsymbol{Y}_j|\boldsymbol{X}_{\mathcal{M}}, W_j))^{\text{T}}\right] \qquad (2.94) \\
&= \text{E}\left[(\boldsymbol{Y}_i - \text{E}(\boldsymbol{Y}_i|\boldsymbol{X}_{\mathcal{M}}, \boldsymbol{Y}_j, W_i, W_j))(\boldsymbol{Y}_j - \text{E}(\boldsymbol{Y}_j|\boldsymbol{X}_{\mathcal{M}}, W_j))^{\text{T}}\right] \\
&\qquad (2.95) \\
&= 0, \qquad (2.96)
\end{aligned}$$

where (2.94) and (2.95) are due to the above two Markov chains, and (2.96) used Proposition 1 and the fact that $(\boldsymbol{Y}_j - E(\boldsymbol{Y}_j|\boldsymbol{X}_{\mathcal{M}}, W_j))$ is a function of $\omega \triangleq (\boldsymbol{X}_{\mathcal{M}}, \boldsymbol{Y}_j, W_i, W_j)$.

□

Lemma 7. *For any pair $(X_{\mathcal{M}}, Y_{\mathcal{L}})$ satisfying (2.78) and any $\boldsymbol{D}$, there exists a $\boldsymbol{D}_2 \in \mathbb{R}^{L\times L}$ and a*

$$\boldsymbol{\Gamma} = \ {}^{\mathrm{T}}\mathrm{diag}(\boldsymbol{\Gamma}_1, \ldots, \boldsymbol{\Gamma}_K, \gamma_{K+1}, \ldots, \gamma_L) \ \ \in \Upsilon_K(\pi) \tag{2.97}$$

such that

$$\mathrm{diag}(\boldsymbol{D}_2) \leq \boldsymbol{D}, \ \boldsymbol{\Gamma} \preceq \left(\boldsymbol{D}_2^{-1} + \boldsymbol{\Sigma}_{\boldsymbol{N}}^{-1} - \boldsymbol{\Sigma}_{\boldsymbol{Y}}^{-1}\right)^{-1}, \tag{2.98}$$

and the sum-rate of the quadratic Gaussian L-terminal problem satisfies

$$R^{\mathrm{MT}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}) \geq \frac{1}{2}\log\frac{|\boldsymbol{\Sigma}_{\boldsymbol{X}_{\mathcal{M}}}|}{|\boldsymbol{A}\boldsymbol{D}_2\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{B}|} + \sum_{k=1}^{K} \underline{R}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}\{\pi_{2k-1},\pi_{2k}\}}|\boldsymbol{X}_{\mathcal{M}}}(\boldsymbol{\Gamma}_k) + \frac{1}{2}\sum_{i=K+1}^{L}\log\frac{\sigma^2_{N_{\pi_i}}}{\gamma_i}, \tag{2.99}$$

where $\boldsymbol{\Sigma}_{\boldsymbol{Y}\{\pi_{2k-1},\pi_{2k}\}}|\boldsymbol{X}_{\mathcal{M}}$ denotes the conditional covariance matrix of $(Y_{\pi_{2k-1}}, Y_{\pi_{2k}})^{\mathrm{T}}$ given $\boldsymbol{X}_{\mathcal{M}}$, and $\boldsymbol{A}$ and $\boldsymbol{B}$ are defined in (2.77).

Proof. First, given $\boldsymbol{\Sigma}_{\boldsymbol{N}} \in \Upsilon_K(\pi)$ and $\boldsymbol{\Sigma}_{\boldsymbol{N}} \preceq \boldsymbol{\Sigma}_{\boldsymbol{Y}}$, we can always apply (2.75) to find an $M \times L$ matrix $\boldsymbol{H}$ and (2.76) to construct M remote sources $X_{\mathcal{M}}$ such that (2.78) holds. This implies that $\boldsymbol{\Sigma}_{\boldsymbol{N}} = \mathrm{cov}(\boldsymbol{Y}_{\mathcal{L}}|\boldsymbol{X}_{\mathcal{M}}) \in \Upsilon_K(\pi)$. Then we can apply Lemma 6, and obtain that $\mathrm{cov}(\boldsymbol{Y}_{\mathcal{L}}|\boldsymbol{X}_{\mathcal{M}}, W_{\mathcal{L}}) \in \Upsilon_K(\pi)$. Hence we can denote

$$\boldsymbol{\Gamma} \ \stackrel{\Delta}{=} \ \mathrm{cov}(\boldsymbol{Y}_{\mathcal{L}}|\boldsymbol{X}_{\mathcal{M}}, W_{\mathcal{L}}) \ \in \ \Upsilon_K(\pi), \tag{2.100}$$

which takes form of (2.90).

On the other hand, due to (2.78), we know that any scheme that achieves a distortion matrix of $\boldsymbol{D}_2$ on $Y_{\mathcal{L}}$ must be able to achieve a distortion matrix of $\boldsymbol{A}\boldsymbol{D}_2\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{B}$ on $\boldsymbol{X}_{\mathcal{M}}$.

Similar to (2.49), we write

$$\begin{aligned}
&H(W_{\mathcal{L}})\\
=&I(\boldsymbol{Y}_{\mathcal{L}}, \boldsymbol{X}; W_{\mathcal{L}})\\
=&I(\boldsymbol{X}; W_{\mathcal{L}}) + \sum_{i=1}^{K} I(\boldsymbol{Y}_{\{\pi_{2k-1},\pi_{2k}\}}; W_{\{\pi_{2k-1},\pi_{2k}\}} | \boldsymbol{X}_{\mathcal{M}}) + \sum_{i=K+1}^{L} I(\boldsymbol{Y}_{\pi_i}; W_{\pi_i} | \boldsymbol{X}_{\mathcal{M}})
\end{aligned} \tag{2.101}$$

$$\begin{aligned}
=&h(\boldsymbol{X}) - h(\boldsymbol{X}|W_{\mathcal{L}}) + \sum_{i=1}^{K} I(\boldsymbol{Y}_{\{\pi_{2k-1},\pi_{2k}\}}; W_{\{\pi_{2k-1},\pi_{2k}\}} | \boldsymbol{X}_{\mathcal{M}})\\
&+ \sum_{i=K+1}^{L} (h(\boldsymbol{Y}_{\pi_i} | \boldsymbol{X}_{\mathcal{M}}) - h(\boldsymbol{Y}_{\pi_i}; W_{\pi_i} | \boldsymbol{X}_{\mathcal{M}}))
\end{aligned} \tag{2.102}$$

$$\geq \frac{1}{2} \log \frac{|\boldsymbol{\Sigma}_{\boldsymbol{X}_{\mathcal{M}}}|}{|\boldsymbol{A}\boldsymbol{D}_2\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{B}|} + \sum_{i=1}^{K} I(\boldsymbol{Y}_{\{\pi_{2k-1},\pi_{2k}\}}; W_{\{\pi_{2k-1},\pi_{2k}\}} | \boldsymbol{X}_{\mathcal{M}}) + \frac{1}{2} \sum_{i=K+1}^{L} \log \frac{\sigma^2_{N_{\pi_i}}}{\gamma_i}, \tag{2.103}$$

where (2.103) comes from the assumption that the achieved distortion is no larger than $\boldsymbol{D}_2$ in the p.d. sense, and the definitions $\mathrm{cov}(\boldsymbol{Y}_{\{\pi_{2k-1},\pi_{2k}\}} | W_{\{\pi_{2k-1},\pi_{2k}\}}, \boldsymbol{X}_{\mathcal{M}}) = \boldsymbol{\Gamma}_k$ and $\gamma_i = \frac{1}{n}\sum_{j=1}^{n} \mathrm{var}(Y_{i,j}|W_i, \boldsymbol{X})$. Now comparing (2.99) with (2.103), we only need to show that

$$I(\boldsymbol{Y}_{\{\pi_{2k-1},\pi_{2k}\}}; W_{\{\pi_{2k-1},\pi_{2k}\}} | \boldsymbol{X}_{\mathcal{M}}) \geq nR_{\boldsymbol{\Sigma}_{\boldsymbol{Y}_{\{\pi_{2k-1},\pi_{2k}\}}|\boldsymbol{X}_{\mathcal{M}}}}(\boldsymbol{\Gamma}_k) \tag{2.104}$$

holds for any $k \in \mathcal{K}$.

Assume that (2.104) does not hold for some $k \in \mathcal{K}$, i.e., there exist encoders $\psi^{(n)}_{\pi_{2k-1}}$ and $\psi^{(n)}_{\pi_{2k}}$ such that

$$\begin{aligned}
\mathrm{cov}(\boldsymbol{Y}_{\{\pi_{2k-1},\pi_{2k}\}} | W_{\{\pi_{2k-1},\pi_{2k}\}}, \boldsymbol{X}_{\mathcal{M}}) &= \boldsymbol{\Gamma}_k,\\
I(\boldsymbol{Y}_{\{\pi_{2k-1},\pi_{2k}\}}; W_{\{\pi_{2k-1},\pi_{2k}\}} | \boldsymbol{X}_{\mathcal{M}}) &< nR_{\boldsymbol{\Sigma}_{\boldsymbol{Y}_{\{\pi_{2k-1},\pi_{2k}\}}|\boldsymbol{X}_{\mathcal{M}}}}(\boldsymbol{\Gamma}_k).
\end{aligned} \tag{2.105}$$

Then consider the matrix-distortion constrained two-terminal problem with sources

$$\tilde{Y}_{\{\pi_{2k-1},\pi_{2k}\}} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_{\boldsymbol{Y}_{\{\pi_{2k-1},\pi_{2k}\}}|\boldsymbol{X}_{\mathcal{M}}}) \tag{2.106}$$

and target distortion matrix $\mathbf{\Gamma}_k$. Now let $\boldsymbol{X}_{\mathcal{M}}$ be a length-n block of samples independently draw from $X_{\mathcal{M}} = \boldsymbol{A}\boldsymbol{Y}_{\mathcal{L}} + \boldsymbol{Z}_{\mathcal{L}}$ according to (2.76). Also assume that $\boldsymbol{X}_{\mathcal{M}}$ is independent of the sources $\tilde{\boldsymbol{Y}}_{\{\pi_{2k-1},\pi_{2k}\}}$ and available at both the encoders and the decoder. Let

$$\bar{\boldsymbol{Y}}_{\{\pi_{2k-1},\pi_{2k}\}} = \tilde{\boldsymbol{Y}}_{\{\pi_{2k-1},\pi_{2k}\}} + \boldsymbol{H}^{\mathrm{T}}_{\mathcal{M},\{\pi_{2k-1},\pi_{2k}\}}\boldsymbol{X}_{\mathcal{M}}, \tag{2.107}$$

where $\boldsymbol{H}$ is the $M \times L$ matrix satisfying (2.78). It is obvious that $\bar{\boldsymbol{Y}}_{\{\pi_{2k-1},\pi_{2k}\}}$ has a covariance matrix of $\Sigma_{\boldsymbol{Y}_{\{\pi_{2k-1},\pi_{2k}\}}}$, hence we can blindly apply the same encoders $\psi^{(n)}_{\pi_{2k-1}}$ and $\psi^{(n)}_{\pi_{2k}}$ on $\bar{\boldsymbol{Y}}_{\{\pi_{2k-1},\pi_{2k}\}}$ to generate $W_{\{\pi_{2k-1},\pi_{2k}\}}$ before using Slepian-Wolf coding with decoder side information $\boldsymbol{X}_{\mathcal{M}}$, to achieve a final rate of

$$H(W_{\{\pi_{2k-1},\pi_{2k}\}}|\boldsymbol{X}_{\mathcal{M}}) = I(\boldsymbol{Y}_{\{\pi_{2k-1},\pi_{2k}\}}; W_{\{\pi_{2k-1},\pi_{2k}\}}|\boldsymbol{X}_{\mathcal{M}}) < nR_{\mathbf{\Sigma}_{\boldsymbol{Y}_{\{\pi_{2k-1},\pi_{2k}\}}|\boldsymbol{X}_{\mathcal{M}}}}(\mathbf{\Gamma}_k), \tag{2.108}$$

and a distortion matrix of $\mathbf{\Gamma}_k = \mathrm{cov}(\boldsymbol{Y}_{\{\pi_{2k-1},\pi_{2k}\}}|W_{\{\pi_{2k-1},\pi_{2k}\}}, \boldsymbol{X}_{\mathcal{M}})$, which contradicts with the definition of $R_{\mathbf{\Sigma}_{\boldsymbol{Y}_{\{\pi_{2k-1},\pi_{2k}\}}|\boldsymbol{X}_{\mathcal{M}}}}(\mathbf{\Gamma}_k)$. Then Lemma 7 follows from (2.101), (2.104), and Lemma 4. □

Remarks:

- Lemma 6 ensures that $\mathrm{cov}(\boldsymbol{Y}_{\mathcal{L}}|\boldsymbol{X}_{\mathcal{M}}, W_{\mathcal{L}})$ in (2.92) shares the same structure with $\mathbf{\Sigma}_{\boldsymbol{N}} = \mathrm{cov}(\boldsymbol{Y}_{\mathcal{L}}|\boldsymbol{X}_{\mathcal{M}})$ in (2.91), which is assumed to be block-diagonal in this paper. Note that this property holds even for non-block-diagonal $\mathbf{\Sigma}_{\boldsymbol{N}}$'s.
- This structural similarity between $\mathbf{\Sigma}_{\boldsymbol{N}} = \mathrm{cov}(\boldsymbol{Y}_{\mathcal{L}}|\boldsymbol{X}_{\mathcal{M}})$ and $\mathrm{cov}(\boldsymbol{Y}_{\mathcal{L}}|\boldsymbol{X}_{\mathcal{M}}, W_{\mathcal{L}})$ is a key to the proof of Lemma 5, since it restricts $\mathrm{cov}(\boldsymbol{Y}_{\mathcal{L}}|\boldsymbol{X}_{\mathcal{M}}, W_{\mathcal{L}})$, which

equals to $\boldsymbol{\Gamma}$ in (2.97), to be block-diagonal, and hence makes the lower bound (2.99) much simpler.

Now we proceed to the proof of Theorem 2.

Due to Lemma 7, to find the best lower bound on $R^{\mathrm{MT}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D})$, we need to solve the following optimization problem for given $(X_{\mathcal{M}}, Y_{\mathcal{L}})$ and $\boldsymbol{D}$ satisfying (2.78),

$$\begin{array}{rl}
\text{Min.} & \frac{1}{2}\log\frac{|\boldsymbol{\Sigma}_{\boldsymbol{X}_{\mathcal{M}}}|}{|\boldsymbol{A}\boldsymbol{D}_2\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{B}|} + \sum_{k=1}^{K}\underline{R}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}_{\{\pi_{2k-1},\pi_{2k}\}}|\boldsymbol{X}_{\mathcal{M}}}}(\boldsymbol{\Gamma}_k) + \frac{1}{2}\sum_{i=K+1}^{L}\log\frac{\sigma^2_{N_{\pi_i}}}{\gamma_i} \\
\text{over} & \boldsymbol{D}_2, \boldsymbol{\Gamma}_1, \dots, \boldsymbol{\Gamma}_K, \gamma_{2K+1}, \dots, \gamma_L \\
\text{s.t.} & \boldsymbol{\Gamma} \preceq (\boldsymbol{\Sigma}_{\boldsymbol{N}}^{-1} + \boldsymbol{D}_2^{-1} - \boldsymbol{\Sigma}_{\boldsymbol{Y}}^{-1})^{-1}, \\
& \mathbf{0} \ \prec \ \boldsymbol{D}_2 \ \preceq \ \boldsymbol{\Sigma}_{\boldsymbol{Y}}, \\
& [\boldsymbol{D}_2]_{j,j} \ \leq \ D_j, \quad \text{for any } j \in \mathcal{L}, \\
& \mathbf{0} \ \prec \ \boldsymbol{\Gamma}_k \ \preceq \ \Sigma_{N_{\{\pi_{2k-1},\pi_{2k}\}}} \forall k \in \mathcal{K}, \\
& 0 < \gamma_k \leq \sigma^2_{N_{\pi_k}}, \ k = 2K+1, \dots, L,
\end{array}$$

which is clearly convex. The Lagrangian is

$$\begin{aligned}
\mathbb{L} = & -\frac{1}{2}\log|\boldsymbol{A}\boldsymbol{D}_2\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{B}| + \sum_{k=1}^{K}\underline{R}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}_{\{\pi_{2k-1},\pi_{2k}\}}|\boldsymbol{X}_{\mathcal{M}}}}(\boldsymbol{\Gamma}_k) - \frac{1}{2}\sum_{i=K+1}^{L}\log\gamma_i \\
& + \operatorname{tr}(\Lambda((\boldsymbol{\Sigma}_{\boldsymbol{N}}^{-1} + \boldsymbol{D}_2^{-1} - \boldsymbol{\Sigma}_{\boldsymbol{Y}}^{-1}) - \boldsymbol{\Gamma}^{-1})) + \operatorname{tr}(\Omega(\boldsymbol{\Sigma}_{\boldsymbol{Y}}^{-1} - \boldsymbol{D}_2^{-1})) \\
& + \sum_{i=1}^{K}\operatorname{tr}(\Theta_i(\boldsymbol{\Sigma}^{-1}_{N_{\{\pi_{2i-1},\pi_{2i}\}}} - \boldsymbol{\Gamma}_i^{-1})) + \sum_{j=1}^{L}\operatorname{tr}(\Pi_j\boldsymbol{E}_j\boldsymbol{D}_2\boldsymbol{E}_j),
\end{aligned}$$

where Λ, Ω, Θ_i, $i \in \mathcal{K}$, Π_j, $j \in \mathcal{L}$ are p.s.d. matrices, and $\boldsymbol{E}_i$ is the $L \times L$ single-entry matrix whose (i,i)-th element is one.

Assume that the optimal BT scheme achieves a distortion matrix $\tilde{\boldsymbol{D}}_2$, and $\tilde{\boldsymbol{\Gamma}}$ as defined in (2.90), then by applying Lemma 5, we obtain the subgradient based KKT conditions at $(\tilde{\boldsymbol{D}}_2, \tilde{\boldsymbol{\Gamma}})$, which are

$$\begin{aligned}
\tilde{\boldsymbol{D}}_2 \left(\boldsymbol{\Pi} - \tilde{\boldsymbol{D}}_2^{-1} + \tilde{\boldsymbol{D}}_2^{-1} \left(\tilde{\boldsymbol{D}}_2^{-1} + \boldsymbol{\Sigma}_{\boldsymbol{N}}^{-1} - \boldsymbol{\Sigma}_{\boldsymbol{Y}}^{-1}\right)^{-1} \tilde{\boldsymbol{D}}_2^{-1}\right) \tilde{\boldsymbol{D}}_2 &= \boldsymbol{\Lambda} - \boldsymbol{\Omega}, \\
\langle\boldsymbol{\Lambda}\rangle_j^{\pi} + \boldsymbol{\Theta}_j - \quad \left(\langle\boldsymbol{\Sigma}_{\boldsymbol{N}}\rangle_j^{\pi}, \mathrm{diag}(\langle\tilde{\boldsymbol{\Gamma}}\rangle_j^{\pi})\right) &\ni \boldsymbol{0}, \forall j \in \mathcal{K}, \\
\text{for } k = 2K+1, ..., L, \quad [\boldsymbol{\Lambda}]_{\pi_k,\pi_k} &= \left[\tilde{\boldsymbol{\Gamma}}\right]_{\pi_k,\pi_k}, \\
\boldsymbol{\Omega}\left(\boldsymbol{\Sigma}_{\boldsymbol{Y}}^{-1} - \tilde{\boldsymbol{D}}_2^{-1}\right) &= \boldsymbol{0}, \\
\boldsymbol{\Theta}_j(\langle\boldsymbol{\Sigma}_{\boldsymbol{N}}\rangle_j^{\pi} - \langle\tilde{\boldsymbol{\Gamma}}\rangle_j^{\pi}) &= \boldsymbol{0}, \forall j \in \mathcal{K}, \\
[\boldsymbol{\Pi}]_{j,j}\,(\left[\tilde{\boldsymbol{D}}_2\right]_{j,j} - D_j) &= 0, \forall j \in \mathcal{K},
\end{aligned}$$

where $\boldsymbol{\Pi}$, $\boldsymbol{\Lambda}$, $\boldsymbol{\Omega}$, and $\boldsymbol{\Theta}_j$'s are the p.s.d. Lagrangian multipliers. Then Theorem 2 readily follows. □

- Example 1: the block-degraded case

All known cases of quadratic Gaussian MT source coding problems with tight sum-rate bound belong to the non-degraded subclass, where all target distortions are met with equalities (i.e., all distortion constraints are active [38]) in the optimal BT scheme. In this subsection, we first study a block-degraded case, and independently show sum-rate tightness in this case (under certain condition). Then we give a numerical example to confirm that the set of block-degraded case with tight sum-rate intersects with the one defined by the sufficient condition in Theorem 2.

Consider a special case of quadratic Gaussian MT source coding, where the vector source $Y_{\mathcal{L}}$ and the target distortion vector $\boldsymbol{D}$ can be partitioned into K groups, namely, $(Y_{\mathcal{S}_1}, D_{\mathcal{S}_1})$, $(Y_{\mathcal{S}_2}, D_{\mathcal{S}_2}), \ldots, (Y_{\mathcal{S}_K}, D_{\mathcal{S}_K})$, and for any $k \in \mathcal{K}$, there exists an integer $\mathbb{i}(k) \in \mathcal{S}_k$, such that

$$Y_j \;=\; Y_{\mathbb{i}(k)} + Z_j, \quad \text{and} \quad D_j \geq D_{\mathbb{i}(k)} + \sigma_{Z_j}^2, \forall j \in \mathcal{S}_k, \tag{2.109}$$

where $Z_j \ \sim \mathcal{N}(0, \sigma^2_{Z_j})$ with $\sigma^2_{Z_j} > 0$ for $j \in \mathcal{S}_k - \{\mathbb{i}(k)\}$ and $\sigma^2_{Z_{\mathbb{i}(k)}} = 0$ is independent of $Y_{\mathbb{i}(k)}$ and Z_j's are mutually independent. Each $Y_{\mathbb{i}(k)}$, $k \in \mathcal{K}$ is called the *group leader* in $Y_{\mathcal{S}_k}$, and denote $\bar{\boldsymbol{Y}}_{\mathcal{K}} = (Y_{\mathbb{i}(1)}, Y_{\mathbb{i}(2)}, \dots, Y_{\mathbb{i}(k)})^{\mathrm{T}}$, $\bar{\boldsymbol{D}}_{\mathcal{K}} = (D_{\mathbb{i}(1)}, D_{\mathbb{i}(2)}, \dots, D_{\mathbb{i}(k)})^{\mathrm{T}}$. We say a pair $(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})$ is *block-degraded (BD)* if they satisfy the above condition. The K components of $\bar{\boldsymbol{Y}}_{\mathcal{K}}$ are referred to as *core sources* while the other $L - K$ as *redundant sources.*

Equivalently, $(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})$ is BD if there exists a partition $\mathcal{P} = \{\mathcal{S}_k : k \in \mathcal{K}\}$ of $\mathcal{L}$ and another pair $(\boldsymbol{\Sigma}_{\bar{\boldsymbol{Y}}_{\mathcal{K}}}, \bar{\boldsymbol{D}}_{\mathcal{K}})$ such that

$$\boldsymbol{\Sigma}_{\boldsymbol{Y}} = \boldsymbol{G}_{\mathcal{P}} \boldsymbol{\Sigma}_{\bar{\boldsymbol{Y}}_{\mathcal{K}}} \boldsymbol{G}_{\mathcal{P}}^{\mathrm{T}} + \Sigma_{\boldsymbol{Z}_{\mathcal{L}}}, \tag{2.110}$$

$$D_{\mathbb{i}(k)} = \bar{D}_k, \forall k \in \mathcal{K}, \tag{2.111}$$

$$D_j \geq \bar{D}_k + [\Sigma_{\boldsymbol{Z}_{\mathcal{L}}}]_{j,j}, \forall\ j \in \mathcal{S}_k - \{\mathbb{i}(k)\} \text{ and } k \in \mathcal{K}, \tag{2.112}$$

where $\boldsymbol{G}_{\mathcal{P}}$ is an $L \times K$ matrix whose $(j, \mathbb{i}(k))$-th element is one for all $j \in \mathcal{S}_k$, $k \in \mathcal{K}$ with the rest being zero, and $\Sigma_{\boldsymbol{Z}_{\mathcal{L}}}$ is a diagonal matrix whose diagonal elements are positive with exceptions that $[\Sigma_{Z_{\mathcal{L}}}]_{\mathbb{i}(k),\mathbb{i}(k)} = 0$. Then an L-terminal quadratic Gaussian MT source coding problem with a BD pair $(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})$ automatically induces a K-terminal source coding problem defined by the pair $(\boldsymbol{\Sigma}_{\bar{\boldsymbol{Y}}_{\mathcal{K}}}, \bar{\boldsymbol{D}}_{\mathcal{K}})$.

Consider a BD pair $(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})$ with partition $\mathcal{P} = \{\mathcal{S}_k : k \in \mathcal{K}\}$ and $(\boldsymbol{\Sigma}_{\bar{\boldsymbol{Y}}_{\mathcal{K}}}, \bar{\boldsymbol{D}}_{\mathcal{K}}, \boldsymbol{\Sigma}_{\boldsymbol{Z}_{\mathcal{L}}})$ satisfying (2.110)-(2.112). We say a matrix Λ is $\mathcal{P}$-block-diagonal if $[\Lambda]_{i,j} = 0$ for any $i \in \mathcal{S}_k, j \in \mathcal{S}_l$ with $k, l \in \mathcal{K}, k \neq l$, and denote $\mathbb{d}_{\mathcal{P}}$ as the set of all $\mathcal{P}$-block-diagonal matrices. For two $L \times L$ matrices A and B, we write $A \overset{\mathcal{P}}{\equiv} B$ if $[A]_{i,j} = [B]_{i,j}$ for any $i, j \in \mathcal{S}_k$ with some $k \in \mathcal{K}$.

We claim that for a BD pair $(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})$, tightness of the BT sum-rate bound in the induced K-terminal quadratic Gaussian MT source coding problem implies tightness of the same bound in the original L-terminal problem, which is stated in the following.

Lemma 8. *For any BD pair* $(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})$*, if the BT minimum sum-rate is tight for the induced* K*-terminal source coding problem, i.e.,*

$$R^{\mathrm{MT}}_{\boldsymbol{\Sigma}_{\bar{\boldsymbol{Y}}_{\mathcal{K}}}}(\bar{\boldsymbol{D}}_{\mathcal{K}}) = R^{\mathrm{BT}}_{\boldsymbol{\Sigma}_{\bar{\boldsymbol{Y}}_{\mathcal{K}}}}(\bar{\boldsymbol{D}}_{\mathcal{K}}) \tag{2.113}$$

then it must also be tight for the original MT source coding problem defined by $(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})$*, i.e.,*

$$R^{\mathrm{MT}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}_{\mathcal{L}}) = R^{\mathrm{BT}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}_{\mathcal{L}}) = R^{\mathrm{MT}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\bar{\boldsymbol{D}}_{\mathcal{K}}).$$

Proof. First, it is obvious that $R^{\mathrm{BT}}_{\boldsymbol{\Sigma}_{Y_{\mathcal{L}}}}(\boldsymbol{D}_{\mathcal{L}}) = R^{\mathrm{BT}}_{\boldsymbol{\Sigma}_{\bar{\boldsymbol{Y}}_{\mathcal{K}}}}(\bar{\boldsymbol{D}}_{\mathcal{K}})$. Then assume that there is a sequence of schemes $\left\{(\phi^{(n)}_{\mathcal{L}}, \psi^{(n)}_{\mathcal{L}}) : n \in \mathbb{N}^+\right\}$ such that

$$\limsup_{n\to\infty} \sum_{j\in\mathcal{L}} R^{(n)}_j < R^{\mathrm{BT}}_{\boldsymbol{\Sigma}_{\bar{\boldsymbol{Y}}_{\mathcal{K}}},}(\bar{\boldsymbol{D}}_{\mathcal{K}}), \tag{2.114}$$

$$\limsup_{n\to\infty} \frac{1}{n}\mathrm{E}\left[\left(Y_{j,i} - \hat{Y}_{j,i}\right)^2\right] \le D_j, \text{ for any } j \in \mathcal{L}. \tag{2.115}$$

Now consider another sequence of schemes $\left\{(\bar{\phi}^{(n)}_{\mathcal{L}}, \bar{\psi}^{(n)}_{\mathcal{L}}) : n \in \mathbb{N}^+\right\}$ such that for any $k \in \mathcal{K}$,

$$\bar{\phi}^{(n)}_{\mathbb{i}(k)}(\boldsymbol{Y}_{\mathbb{i}(k)}) = \boxtimes_{j\in\mathcal{S}_k} \bar{W}_j, \tag{2.116}$$

$$\bar{\phi}^{(n)}_j(\boldsymbol{Y}_j) \equiv 0 \text{ for any } j \in \mathcal{S}_k - \{\mathbb{i}(k)\}, \tag{2.117}$$

where

$$\bar{W}_{\mathbb{i}(k)} \triangleq W_{\mathbb{i}(k)} = \phi^{(n)}_{\mathbb{i}(k)}(\boldsymbol{Y}_{\mathbb{i}(k)}), \tag{2.118}$$

$$\bar{W}_j \triangleq \phi^{(n)}_j(\boldsymbol{Y}_{\mathbb{i}(k)} + \boldsymbol{Z}_j), \tag{2.119}$$

with $\bar{Z}_j \sim \mathcal{N}(0, \sigma^2_{Z_j})$ being independent of $Y_{\mathcal{L}}$, "$\boxtimes$" denotes Cartesian product, and

$$\bar{\psi}^{(n)}_{\mathbb{i}(k)}(W_{\mathcal{L}}) = \psi^{(n)}_{\mathbb{i}(k)}(\bar{W}_{\mathcal{L}}). \tag{2.120}$$

Then we must have

$$R^{\mathrm{MT}}_{\boldsymbol{\Sigma}_{\bar{\boldsymbol{Y}}_{\mathcal{K}}}}\left(\phi^{(n)}_{\mathcal{L}},\psi^{(n)}_{\mathcal{L}}\right)=R^{\mathrm{MT}}_{\boldsymbol{\Sigma}_{\bar{\boldsymbol{Y}}_{\mathcal{K}}}}\left(\bar{\phi}^{(n)}_{\mathcal{L}},\bar{\psi}^{(n)}_{\mathcal{L}}\right), \tag{2.121}$$

$$\Rightarrow \limsup_{n\to\infty} R^{\mathrm{MT}}_{\boldsymbol{\Sigma}_{\bar{\boldsymbol{Y}}_{\mathcal{K}}}}\left(\bar{\phi}^{(n)}_{\mathcal{L}},\bar{\psi}^{(n)}_{\mathcal{L}}\right)=\limsup_{n\to\infty} R^{\mathrm{MT}}_{\boldsymbol{\Sigma}_{\bar{\boldsymbol{Y}}_{\mathcal{K}}}}\left(\phi^{(n)}_{\mathcal{L}},\psi^{(n)}_{\mathcal{L}}\right)<R^{\mathrm{BT}}_{\boldsymbol{\Sigma}_{\bar{\boldsymbol{Y}}_{\mathcal{K}}}}\left(\bar{\boldsymbol{D}}_{\mathcal{K}}\right), \tag{2.122}$$

and

$$\limsup_{n\to\infty}\frac{1}{n}\mathrm{E}\left[\left(Y_{j,i}-\mathrm{E}(Y_{j,i}|\bar{W}_{\mathcal{L}})\right)^2\right]\leq\begin{cases} D_j, & j=\mathbb{i}(k)\text{ for some }k\in\mathcal{K}\\ D_{\mathbb{i}(k)}+\sigma^2_{Z_j}\leq D_j, & j\in\mathcal{S}_k-\{\mathbb{i}(k)\}\text{ for some }k\in\mathcal{K}\end{cases}.$$

Hence the sequence of schemes $\left\{(\bar{\phi}^{(n)}_{\mathcal{L}},\bar{\psi}^{(n)}_{\mathcal{L}}):n\in\mathbb{N}^+\right\}$ achieves the distortion vector $\boldsymbol{D}$ and a sum-rate smaller than $R^{\mathrm{BT}}_{\boldsymbol{\Sigma}_{\bar{\boldsymbol{Y}}_{\mathcal{K}}}}(\bar{\boldsymbol{D}}_{\mathcal{K}})$. On the other hand, $\left\{(\bar{\phi}^{(n)}_{\mathcal{L}},\bar{\psi}^{(n)}_{\mathcal{L}}):n\in\mathbb{N}^+\right\}$ is also an achievable sequence of schemes for the induced K-terminal problem, for which the BT sum-rate bound $R^{\mathrm{BT}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D})=R^{\mathrm{BT}}_{\boldsymbol{\Sigma}_{\bar{\boldsymbol{Y}}_{\mathcal{K}}}}(\bar{\boldsymbol{D}}_{\mathcal{K}})$ is known to be tight, leading to a contradiction. □

Remarks:

- Although Wang *et al.*'s sufficient condition [13] for sum-rate tightness does not include any degraded case, one can easily use Lemma 8 to generate a BD example with tight sum-rate bound. In fact, with slight modifications (with details omitted), Wang *et al.*'s proof [13] can also be generalized to directly show sum-rate tightness for such BD cases without explicitly using Lemma 8.

- We note that Lemma 8 only guarantees the sum-rate tightness of a *subset* of the BD subclass of quadratic Gaussian MT source coding problems. Moreover, this subset intersects with the one defined by the sufficient condtion in Theorem 2, as shown in the following numerical example.

A specific numerical example that satisfies the requirements in both Theorem 2 and Lemma 8 is as follows. Let $L = 4$,

$$\mathbf{\Sigma}_{\boldsymbol{Y}} = \begin{bmatrix} 1.0000 & 0.9000 & 0.8000 & 0.8000 \\ 0.9000 & 1.0000 & 0.7000 & 0.7000 \\ 0.8000 & 0.7000 & 1.0000 & 1.0000 \\ 0.8000 & 0.7000 & 1.0000 & 1.1000 \end{bmatrix}, \tag{2.123}$$

and

$$\boldsymbol{D} = (0.3760, 0.35, 0.3, 0.5)^{\mathrm{T}}, \tag{2.124}$$

The optimal BT distortion matrix is

$$\tilde{\boldsymbol{D}}_2 = \begin{bmatrix} 0.3760 & 0.2740 & 0.1818 & 0.1818 \\ 0.2740 & 0.3500 & 0.1231 & 0.1231 \\ 0.1818 & 0.1231 & 0.3000 & 0.3000 \\ 0.1818 & 0.1231 & 0.3000 & 0.4000 \end{bmatrix}, \tag{2.125}$$

hence this example is degraded since $D_4 = 0.5$ is not achieved with equality in the optimal BT distortion matrix $\tilde{\boldsymbol{D}}_2$.

We first verify that this example satisfies the sufficient condition in Theorem 2. Let $\pi = \{1, 2, 3, 4\}$ and

$$\mathbf{\Sigma}_{\boldsymbol{N}} = \begin{bmatrix} 0.2942 & 0.2852 & 0 & 0 \\ 0.2852 & 0.4535 & 0 & 0 \\ 0 & 0 & 0.0923 & 0 \\ 0 & 0 & 0 & 0.1923 \end{bmatrix} \tag{2.126}$$

be a $\pi^{(K)}$ p.d. block diagonal matrix with $K = 1$. Then $M = 4$,

$$\boldsymbol{\Sigma}_{\boldsymbol{X}_{\mathcal{M}}} = \begin{bmatrix} 3.1162 & 0 & 0 & 0 \\ 0 & 0.0923 & 0 & 0 \\ 0 & 0 & 0.0377 & 0 \\ 0 & 0 & 0 & 0.0061 \end{bmatrix}, \ \boldsymbol{H} = \begin{bmatrix} -0.4712 & -0.4130 & -0.5511 & -0.5511 \\ 0 & 0 & 0.7071 & -0.7071 \\ 0.5417 & 0.5619 & -0.4421 & -0.4421 \\ -0.6961 & 0.7167 & 0.0290 & 0.0290 \end{bmatrix}. \tag{2.127}$$

Now the following p.s.d. matrices

$$\boldsymbol{\Lambda} = \begin{bmatrix} 0.2248 & 0.2489 & 0.0967 & 0.0967 \\ 0.2489 & 0.2791 & 0.1075 & 0.1075 \\ 0.0967 & 0.1075 & 0.0783 & 0 \\ 0.0967 & 0.1075 & 0 & 0.1923 \end{bmatrix}, \ \boldsymbol{\Omega} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.1000 \end{bmatrix}, \ \boldsymbol{\Theta}_{\mathbf{1}} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$
$$\boldsymbol{\Pi} = \begin{bmatrix} 1.0377 & 0 & 0 & 0 \\ 0 & 1.8957 & 0 & 0 \\ 0 & 0 & 2.6331 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \ \tilde{\boldsymbol{\Gamma}} = \begin{bmatrix} 0.2248 & 0.1753 & 0 & 0 \\ 0.1753 & 0.2791 & 0 & 0 \\ 0 & 0 & 0.0783 & 0 \\ 0 & 0 & 0 & 0.1923 \end{bmatrix} \tag{2.128}$$

satisfy all the KKT conditions. Note that $\tilde{\boldsymbol{\Gamma}}$ in (2.128) has the same structure as $\boldsymbol{\Sigma}_{\boldsymbol{N}}$ in (2.126), which is consistent with Lemma 6. In addition, $(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})$ is a line segment

$$(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D}) = \left\{ \alpha \cdot \begin{bmatrix} 0.2248 & 0.2505 \\ 0.2505 & 0.2791 \end{bmatrix} + (1-\alpha) \cdot \begin{bmatrix} 0.2248 & 0.1001 \\ 0.1001 & 0.2791 \end{bmatrix} : \alpha \in [0,1] \right\}. \tag{2.129}$$

On the other hand, it is easy to verify that $(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})$ is a BD pair with

$$\mathcal{P} = \{\{1\}, \{2\}, \{3,4\}\}, \ \boldsymbol{\Sigma}_{\bar{\boldsymbol{Y}}_{\mathcal{K}}} = \begin{bmatrix} 1.0000 & 0.9000 & 0.8000 \\ 0.9000 & 1.0000 & 0.7000 \\ 0.8000 & 0.7000 & 1.0000 \end{bmatrix},$$

$$\Sigma_{\boldsymbol{Z}_{\mathcal{L}}} = \mathrm{diag}(0,0,0,0.1), \ \bar{\boldsymbol{D}}_{\mathcal{K}} = (0.3760, 0.35, 0.3)^{\mathrm{T}},$$

and the induced three-terminal quadratic Gaussian MT source coding problem defined by $(\boldsymbol{\Sigma}_{\bar{\boldsymbol{Y}}_{\mathcal{K}}}, \bar{D}_{\mathcal{L}})$ has a tight sum-rate bound due to Theorem 2. Hence we conclude that the above four-terminal numerical example of quadratic Gaussian MT source coding problem also satisfies the simple sufficient condition in Lemma 8.

D. A simplified sufficient condition

Although the sufficient condition given in Theorem 2 is more inclusive than that in [13], it is rather complicated and hard to verify. However, in the *non-degraded* case where the optimal BT scheme quantizes every source, and achieves all L target distortions with equalities, the sufficient condition in Theorem 2 can be further simplified. Note that the non-degraded case is of special interest since all the previously known quadratic Gaussian MT source coding problems with tight sum-rate bound belong to this case.

Corollary 1. *For an MT source coding problem defined by $\boldsymbol{\Sigma}_{\boldsymbol{Y}}$ and $\boldsymbol{D}$, if the optimal BT distortion matrix $\tilde{\boldsymbol{D}}_2$ satisfies $\mathrm{diag}(\tilde{\boldsymbol{D}}_2) = \boldsymbol{D}$ and $\tilde{\boldsymbol{D}}_2^{-1} - \boldsymbol{\Sigma}_{\boldsymbol{Y}}^{-1}$ is a p.d. matrix, then $R^{\mathrm{BT}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}) = R^{\mathrm{MT}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D})$ if there exists a permutation π and a $\pi^{(K)}$ block diagonal p.d. matrix $\boldsymbol{\Sigma}_{\boldsymbol{N}}$ such that $\boldsymbol{\Sigma}_{\boldsymbol{N}} \preceq \boldsymbol{\Sigma}_{\boldsymbol{Y}}$,*

$$\boldsymbol{\Lambda} \triangleq \tilde{\boldsymbol{D}}_2 \left(\boldsymbol{\Pi} - \tilde{\boldsymbol{D}}_2^{-1} + \tilde{\boldsymbol{D}}_2^{-1} \left(\tilde{\boldsymbol{D}}_2^{-1} + \boldsymbol{\Sigma}_{\boldsymbol{N}}^{-1} - \boldsymbol{\Sigma}_{\boldsymbol{Y}}^{-1} \right)^{-1} \tilde{\boldsymbol{D}}_2^{-1} \right) \tilde{\boldsymbol{D}}_2 \qquad (2.130)$$

is a p.s.d. matrix, and

$$\operatorname{sign}\left(\left[\tilde{\mathbf{\Gamma}}\right]_{\pi_{2k-1},\pi_{2k-1}}\right)\cdot[\mathbf{\Lambda}]_{\pi_{2k-1},\pi_{2k}} \geq 2|\,[\mathbf{\Gamma}]_{\pi_{2k-1},\pi_{2k}}\,| - \sqrt{[\mathbf{\Gamma}]_{\pi_{2k-1},\pi_{2k-1}}\,[\mathbf{\Gamma}]_{\pi_{2k},\pi_{2k}}} \tag{2.131}$$

is satisfied for all $k \in \mathcal{K}$, where $\tilde{\mathbf{\Gamma}}$ is defined in (2.90) and

$$\mathbf{\Pi} \triangleq \operatorname{diag}\left((\tilde{\boldsymbol{D}}_2 \odot \tilde{\boldsymbol{D}}_2)^{-1}\boldsymbol{D}\right), \tag{2.132}$$

with $\odot$ denoting Hadamard product (entrywise product).

Proof. First, due to the assumption that $\tilde{\boldsymbol{D}}_2^{-1} - \mathbf{\Sigma}_{\boldsymbol{Y}}^{-1} \succ \mathbf{0}$, (2.87) implies that $\mathbf{\Omega} = \mathbf{0}$, which, combined with (2.84), directly leads to (2.130). On the other hand, $\tilde{\boldsymbol{D}}_2^{-1} - \mathbf{\Sigma}_{\boldsymbol{Y}}^{-1} \succ \mathbf{0}$ also ensures that

$$\tilde{\mathbf{\Gamma}} = (\tilde{\boldsymbol{D}}_2^{-1} + \mathbf{\Sigma}_{\boldsymbol{N}}^{-1} - \mathbf{\Sigma}_{\boldsymbol{Y}}^{-1})^{-1} \prec \mathbf{\Sigma}_{\boldsymbol{N}}, \tag{2.133}$$

hence (2.88) is true if and only if $\mathbf{\Theta}_j = \mathbf{0}$ for all $j \in \mathcal{K}$.

Now (2.85) becomes

$$\langle \mathbf{\Lambda} \rangle_j^\pi - \quad \left(\langle \mathbf{\Sigma}_{\boldsymbol{N}} \rangle_j^\pi, \operatorname{diag}(\langle \tilde{\mathbf{\Gamma}} \rangle_j^\pi)\right) \quad \ni \quad \mathbf{0}, \forall j \in \mathcal{K}, \tag{2.134}$$

then due to the fact that all 2×2 matrices in $\left(\langle \mathbf{\Sigma}_{\boldsymbol{N}} \rangle_j^\pi, \operatorname{diag}(\langle \tilde{\mathbf{\Gamma}} \rangle_j^\pi)\right)$ have the same diagonal elements as those of $\langle \tilde{\mathbf{\Gamma}} \rangle_j^\pi$, we know that

$$[\mathbf{\Lambda}]_{\pi_k,\pi_k} \quad = \quad \left[\tilde{\mathbf{\Gamma}}\right]_{\pi_k,\pi_k}, \ \forall\ k = 1, 2, ..., 2K. \tag{2.135}$$

Hence by combining (2.86) and (2.135), we obtain

$$\begin{aligned}
&\operatorname{diag}(\boldsymbol{\Lambda}) = \operatorname{diag}\left(\tilde{\boldsymbol{\Gamma}}\right) \\
\Leftrightarrow\ &\operatorname{diag}\left(\tilde{\boldsymbol{D}}_2(\boldsymbol{\Pi} - \tilde{\boldsymbol{D}}_2^{-1} + \tilde{\boldsymbol{D}}_2^{-1}(\tilde{\boldsymbol{D}}_2^{-1} + \boldsymbol{\Sigma}_{\boldsymbol{N}}^{-1} - \boldsymbol{\Sigma}_{\boldsymbol{Y}}^{-1})^{-1}\tilde{\boldsymbol{D}}_2^{-1})\tilde{\boldsymbol{D}}_2\right) = \operatorname{diag}\left((\tilde{\boldsymbol{D}}_2^{-1} + \boldsymbol{\Sigma}_{\boldsymbol{N}}^{-1} - \boldsymbol{\Sigma}_{\boldsymbol{Y}}^{-1})^{-1}\right) \\
\Leftrightarrow\ &\operatorname{diag}(\tilde{\boldsymbol{D}}_2\boldsymbol{\Pi}\tilde{\boldsymbol{D}}_2) = \operatorname{diag}(\tilde{\boldsymbol{D}}_2) \\
\Leftrightarrow\ &\sum_{j=1}^{L}\left[\tilde{\boldsymbol{D}}_2\right]_{i,j}^2 \cdot [\boldsymbol{\Pi}]_{j,j} = \left[\tilde{\boldsymbol{D}}_2\right]_{i,i},\ \forall\ i \in \mathcal{L} \\
\Leftrightarrow\ &(\tilde{\boldsymbol{D}}_2 \odot \tilde{\boldsymbol{D}}_2)\operatorname{diag}(\boldsymbol{\Pi}) = \operatorname{diag}(\tilde{\boldsymbol{D}}_2) = \boldsymbol{D} \\
\Leftrightarrow\ &\operatorname{diag}(\boldsymbol{\Pi}) = (\tilde{\boldsymbol{D}}_2 \odot \tilde{\boldsymbol{D}}_2)^{-1}\boldsymbol{D}, \qquad (2.136)
\end{aligned}$$

and (2.132) is proved.

Finally, (2.135) holds if there exists an $\alpha \in [0,1]$ such that

$$\begin{aligned}
[\boldsymbol{\Lambda}]_{\pi_{2k-1},\pi_{2k}} = &\left(\alpha + (1-\alpha)(2|\frac{\left[\tilde{\boldsymbol{\Gamma}}\right]_{\pi_{2k-1},\pi_{2k}}}{\sqrt{\left[\tilde{\boldsymbol{\Gamma}}\right]_{\pi_{2k-1},\pi_{2k-1}}\left[\tilde{\boldsymbol{\Gamma}}\right]_{\pi_{2k},\pi_{2k}}}}| - 1)\right) \\
&\cdot \operatorname{sign}\left(\left[\tilde{\boldsymbol{\Gamma}}\right]_{\pi_{2k-1},\pi_{2k-1}}\right)\sqrt{\left[\tilde{\boldsymbol{\Gamma}}\right]_{\pi_{2k-1},\pi_{2k-1}}\left[\tilde{\boldsymbol{\Gamma}}\right]_{\pi_{2k},\pi_{2k}}}. \qquad (2.137)
\end{aligned}$$

Now (2.137) is equivalent to

$$\operatorname{sign}\left(\left[\tilde{\boldsymbol{\Gamma}}\right]_{\pi_{2k-1},\pi_{2k-1}}\right) \cdot [\boldsymbol{\Lambda}]_{\pi_{2k-1},\pi_{2k}} \leq \sqrt{\left[\tilde{\boldsymbol{\Gamma}}\right]_{\pi_{2k-1},\pi_{2k-1}}\left[\tilde{\boldsymbol{\Gamma}}\right]_{\pi_{2k},\pi_{2k}}} \qquad (2.138)$$

and

$$\operatorname{sign}\left(\left[\tilde{\boldsymbol{\Gamma}}\right]_{\pi_{2k-1},\pi_{2k-1}}\right) \cdot [\boldsymbol{\Lambda}]_{\pi_{2k-1},\pi_{2k}} \geq 2\left|\left[\tilde{\boldsymbol{\Gamma}}\right]_{\pi_{2k-1},\pi_{2k}}\right| - \sqrt{\left[\tilde{\boldsymbol{\Gamma}}\right]_{\pi_{2k-1},\pi_{2k-1}}\left[\tilde{\boldsymbol{\Gamma}}\right]_{\pi_{2k},\pi_{2k}}}, \qquad (2.139)$$

where (2.138) is automatically satisfied since

$$[\boldsymbol{\Lambda}]_{\pi_{2k-1},\pi_{2k-1}} = \left[\tilde{\boldsymbol{\Gamma}}\right]_{\pi_{2k-1},\pi_{2k-1}},\ [\boldsymbol{\Lambda}]_{\pi_{2k},\pi_{2k}} = \left[\tilde{\boldsymbol{\Gamma}}\right]_{\pi_{2k},\pi_{2k}},\ \text{and } \langle\boldsymbol{\Lambda}\rangle_j^{\pi} \succeq \boldsymbol{0}. \qquad (2.140)$$

Hence (2.131) must hold. □

- Example 2: the block-circulant case

We study a special class of quadratic Gaussian MT source coding problem named the block-circulant case.

Let $L = 2m$ be an even number, and assume that the source covariance matrix $\mathbf{\Sigma}_{\boldsymbol{Y}}$ is *block-circulant*, i.e., it is of the form

$$\mathbf{\Sigma}_{\boldsymbol{Y}} = \begin{bmatrix} B_1 & B_2 & B_3 & ... & B_m \\ B_m & B_1 & B_2 & ... & B_{m-1} \\ ... & ... & ... & ... & ... \\ B_2 & B_3 & B_4 & ... & B_1 \end{bmatrix},$$

where $B_i = B_{m+2-i}$ for $i = 2, 3, ..., m$ are p.d. symmetric 2×2 blocks of the form

$$B_i = \begin{bmatrix} b_{i,1} & b_{i,2} \\ b_{i,2} & b_{i,1} \end{bmatrix}. \tag{2.141}$$

Denote $\mathbb{C}_L$ as the set of all $L \times L$ block-circulant matrices. We state several important properties of block-circulant matrices.

- Any $\mathbf{\Sigma} \in \mathbb{C}_L$ can be diagonalized by

$$\boldsymbol{G}_L \triangleq \boldsymbol{F}_m \otimes \boldsymbol{F}_2, \tag{2.142}$$

with $\otimes$ denoting Kronecker product, and $\boldsymbol{F}_m$ being the $m \times m$ *real Fourier* *matrix* [14] (which is orthogonal with $\boldsymbol{F}_m\boldsymbol{F}_m^{\mathrm{T}} = \boldsymbol{I}_m$). For example, when $L = 6$,

$$\boldsymbol{G}_6 = \boldsymbol{F}_3 \otimes \boldsymbol{F}_2 = \begin{bmatrix} 0.4082 & 0.4082 & 0 & 0 & 0.5774 & 0.5774 \\ 0.4082 & -0.4082 & 0 & 0 & 0.5774 & -0.5774 \\ 0.4082 & 0.4082 & 0.5000 & 0.5000 & -0.2887 & -0.2887 \\ 0.4082 & -0.4082 & 0.5000 & -0.5000 & -0.2887 & 0.2887 \\ 0.4082 & 0.4082 & -0.5000 & -0.5000 & -0.2887 & -0.2887 \\ 0.4082 & -0.4082 & -0.5000 & 0.5000 & -0.2887 & 0.2887 \end{bmatrix}. \tag{2.143}$$

- $\mathbb{C}_L$ is a ring under matrix addition and multiplication. In particular, $\mathbb{C}_L$ is closed under the following operation

$$\boldsymbol{A} \star \boldsymbol{B} \triangleq \boldsymbol{A} - \boldsymbol{A}(\boldsymbol{A}+\boldsymbol{B})^{-1}\boldsymbol{A} = \boldsymbol{B} - \boldsymbol{B}(\boldsymbol{A}+\boldsymbol{B})^{-1}\boldsymbol{B} \in \mathbb{C}_L, \ \forall \ \boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}_L. \tag{2.144}$$

- For any $\boldsymbol{A} \in \mathbb{C}_L$, there are $2 \cdot \lceil \frac{L+1}{2} \rceil$ degrees of freedom in the L eigenvalues of $\boldsymbol{A}$, with $\lceil x \rceil$ denoting the smallest integer larger than x.

We say a quadratic Gaussian MT source coding problem belongs to the *block-circulant* case if the source covariance matrix is block-circulant and all the target distortions are equal, i.e., $\boldsymbol{\Sigma}_{\boldsymbol{Y}} \in \mathbb{C}_L$ and $\boldsymbol{D} = D \cdot \boldsymbol{1}$. An important fact for this special case, which follows directly from the properties of block-circulant matrices, is that the optimal BT distortion matrix can be expressed analytically with

$$\tilde{\boldsymbol{D}}_2 = \boldsymbol{\Sigma}_{\boldsymbol{Y}} \star q\boldsymbol{I}_L, \tag{2.145}$$

where q satisfies

$$\sum_{i=1}^{L} \frac{1}{\frac{1}{\lambda_i} + \frac{1}{q}} = LD, \tag{2.146}$$

with λ_i, $i \in \mathcal{L}$ being the L eigenvalues of $\boldsymbol{\Sigma_Y}$.

Now we are ready to investigate the tightness condition provided by Wang *et al.* [13] for this block-circulant case, which is given in the following lemma.

Lemma 9. *For any block-circulant quadratic Gaussian MT source coding problem, Wang* et al.*'s tightness condition [13, Lemma 4] for the sum-rate bound to be tight is equivalent to*

$$\mathrm{diag}((\tilde{\boldsymbol{D}}_2 \odot \tilde{\boldsymbol{D}}_2)^{-1} D\mathbf{1}) \succeq \tilde{\boldsymbol{D}}_2^{-1} - \tilde{\boldsymbol{D}}_2^{-1}(\tilde{\boldsymbol{D}}_2^{-1} + \lambda_{\min}^{-1}\boldsymbol{I}_L - \boldsymbol{\Sigma_Y}^{-1})^{-1}\tilde{\boldsymbol{D}}_2^{-1}, \quad (2.147)$$

with $\tilde{\boldsymbol{D}}_2$ defined in (2.145) and $\lambda_{\min}$ being the smallest eigenvalue of $\boldsymbol{\Sigma_Y}$.

Proof. We only need to show that if

$$\mathrm{diag}((\tilde{\boldsymbol{D}}_2 \odot \tilde{\boldsymbol{D}}_2)^{-1} D\mathbf{1}) \succeq \tilde{\boldsymbol{D}}_2^{-1} - \tilde{\boldsymbol{D}}_2^{-1}(\tilde{\boldsymbol{D}}_2^{-1} + \Theta^{-1} - \boldsymbol{\Sigma_Y}^{-1})^{-1}\tilde{\boldsymbol{D}}_2^{-1} \qquad (2.148)$$

holds for some p.s.d. diagonal matrix $\Theta = \mathrm{diag}(\mu_1, \mu_2, ..., \mu_L)$ such that

$$\boldsymbol{\Sigma_Y} \ \succeq \ \Theta, \qquad (2.149)$$

then (2.147) must also hold.

In fact, due to the symmetric properties of block-circulant matrices, it is easy to show that if both (2.148) and (2.149) hold for $\Theta = \mathrm{diag}(\mu_1, \mu_2, ..., \mu_L)$, then they must also hold for

$$\Theta_k^{\dagger} = \mathrm{diag}(\mu_{\varsigma(k,1)}, \mu_{\varsigma(k,2)}, \mu_{\varsigma(k+1,1)}, \mu_{\varsigma(k+1,2)}, ..., \mu_{\varsigma(k+m-1,1)}, \mu_{\varsigma(k+m-1,2)}), \qquad (2.150)$$

for any $k \in \{0, 1, ..., m-1\}$, as well as

$$\Theta_k^{\ddagger} = \mathrm{diag}(\mu_{\varsigma(k,2)}, \mu_{\varsigma(k,1)}, \mu_{\varsigma(k+1,2)}, \mu_{\varsigma(k+1,1)}, ..., \mu_{\varsigma(k+m-1,2)}, \mu_{\varsigma(k+m-1,1)}), \qquad (2.151)$$

where $\varsigma(j,i) \triangleq 2 \cdot (j \bmod m) + i$. Hence (2.147) must be true since

$$
\begin{aligned}
&\operatorname{diag}((\tilde{\boldsymbol{D}}_2 \odot \tilde{\boldsymbol{D}}_2)^{-1} D\mathbf{1}) \\
\succeq\; & \frac{1}{L}\sum_{k=1}^{m}\left[\tilde{\boldsymbol{D}}_2^{-1} - \tilde{\boldsymbol{D}}_2^{-1}(\tilde{\boldsymbol{D}}_2^{-1} + (\Theta_k^{\dagger})^{-1} - \boldsymbol{\Sigma}_{\boldsymbol{Y}}^{-1})^{-1}\tilde{\boldsymbol{D}}_2^{-1}\right] \\
&+ \frac{1}{L}\sum_{k=1}^{m}\left[\tilde{\boldsymbol{D}}_2^{-1} - \tilde{\boldsymbol{D}}_2^{-1}(\tilde{\boldsymbol{D}}_2^{-1} + (\Theta_k^{\ddagger})^{-1} - \boldsymbol{\Sigma}_{\boldsymbol{Y}}^{-1})^{-1}\tilde{\boldsymbol{D}}_2^{-1}\right] \\
\succeq\; & \tilde{\boldsymbol{D}}_2^{-1} - \left(\tilde{\boldsymbol{D}}_2\left(\tilde{\boldsymbol{D}}_2^{-1} + \left(\frac{1}{L}\sum_{k=1}^{m}\Theta_k^{\dagger} + \frac{1}{L}\sum_{k=1}^{m}\Theta_k^{\ddagger}\right)^{-1} - \boldsymbol{\Sigma}_{\boldsymbol{Y}}^{-1}\right)\tilde{\boldsymbol{D}}_2\right)^{-1} && (2.152) \\
\succeq\; & \tilde{\boldsymbol{D}}_2^{-1} - \tilde{\boldsymbol{D}}_2^{-1}\left(\tilde{\boldsymbol{D}}_2^{-1} + \lambda_{\min}^{-1}\boldsymbol{I}_L - \boldsymbol{\Sigma}_{\boldsymbol{Y}}^{-1}\right)^{-1}\tilde{\boldsymbol{D}}_2^{-1}, && (2.153)
\end{aligned}
$$

where (2.152) is due to the concavity of $\tilde{\boldsymbol{D}}_2^{-1} - \tilde{\boldsymbol{D}}_2^{-1}\left(\tilde{\boldsymbol{D}}_2^{-1} + \Theta^{-1} - \boldsymbol{\Sigma}_{\boldsymbol{Y}}^{-1}\right)^{-1}\tilde{\boldsymbol{D}}_2^{-1}$ with respect to Θ, and (2.153) uses the fact that

$$
\boldsymbol{\Sigma}_{\boldsymbol{Y}} \succeq \Theta_k^{\dagger},\quad \boldsymbol{\Sigma}_{\boldsymbol{Y}} \succeq \Theta_k^{\ddagger} \Rightarrow \boldsymbol{\Sigma}_{\boldsymbol{Y}} \succeq \frac{1}{L}\sum_{k=1}^{m}\Theta_k^{\dagger} + \frac{1}{L}\sum_{k=1}^{m}\Theta_k^{\ddagger} = \frac{1}{L}\sum_{i=1}^{L}\mu_i\boldsymbol{I}_L \Rightarrow \frac{1}{L}\sum_{i=1}^{L}\mu_i \le \lambda_{\min}. \tag{2.154}
$$

□

With Lemma 9, one can easily test whether Wang *et al.*'s tightness condition is satisfied by a block-circulant case of quadratic Gaussian MT source coding problem. For example, let $L = 4$ and

$$
\boldsymbol{\Sigma}_{\boldsymbol{Y}} = \begin{bmatrix} 1.0000 & 0.5000 & 0.9750 & 0.4800 \\ 0.5000 & 1.0000 & 0.4800 & 0.9750 \\ 0.9750 & 0.4800 & 1.0000 & 0.5000 \\ 0.4800 & 0.9750 & 0.5000 & 1.0000 \end{bmatrix} \in \mathbb{C}_4, \tag{2.155}
$$

and $\boldsymbol{D} = 0.1362 \cdot \boldsymbol{1}$. Then the optimal BT distortion matrix is

$$\tilde{\boldsymbol{D}}_2 = \begin{bmatrix} 0.1362 & 0.0189 & 0.1142 & 0.0018 \\ 0.0189 & 0.1362 & 0.0018 & 0.1142 \\ 0.1142 & 0.0018 & 0.1362 & 0.0189 \\ 0.0018 & 0.1142 & 0.0189 & 0.1362 \end{bmatrix}. \tag{2.156}$$

We first use Lemma 9 to test Wang *et al.*'s tightness condition, which is not satisfied since

$$\begin{aligned} \operatorname{diag}\left((\tilde{\boldsymbol{D}}_2 \odot \tilde{\boldsymbol{D}}_2)^{-1} D\boldsymbol{1}\right) &= 4.1631 \boldsymbol{I}_4 \\ &\not\preceq \begin{bmatrix} 7.5599 & 5.4290 & -3.6183 & -5.7492 \\ 5.4290 & 7.5599 & -5.7492 & -3.6183 \\ -3.6183 & -5.7492 & 7.5599 & 5.4290 \\ -5.7492 & -3.6183 & 5.4290 & 7.5599 \end{bmatrix} \\ &= \tilde{\boldsymbol{D}}_2^{-1} - \tilde{\boldsymbol{D}}_2^{-1}\left(\tilde{\boldsymbol{D}}_2^{-1} + \lambda_{\min}^{-1}\boldsymbol{I}_L - \boldsymbol{\Sigma}_{\boldsymbol{Y}}^{-1}\right)^{-1}\tilde{\boldsymbol{D}}_2^{-1}. \end{aligned} \tag{2.157}$$

However, it is easy to verify that this example does satisfy the condition given in Corollary 1, since when $\pi = \{1, 2, 3, 4\}$ and

$$\boldsymbol{\Sigma}_{\boldsymbol{N}} = \begin{bmatrix} 0.0250 & 0.0200 & 0 & 0 \\ 0.0200 & 0.0250 & 0 & 0 \\ 0 & 0 & 0.0250 & 0.0200 \\ 0 & 0 & 0.0200 & 0.0250 \end{bmatrix} \in \Upsilon_2(\pi), \tag{2.158}$$

$\tilde{\boldsymbol{\Gamma}}$ and $\boldsymbol{\Lambda}$ defined in (2.90) and (2.130) satisfy for $k = 1, 2$:

$$\begin{aligned}\operatorname{sign}\left(\left[\tilde{\boldsymbol{\Gamma}}\right]_{2k-1,2k}\right) \cdot [\boldsymbol{\Lambda}]_{2k-1,2k} &= 0.0219 \\ &\geq 0.0171 = 2\left[\tilde{\boldsymbol{\Gamma}}\right]_{2k-1,2k} - \sqrt{\left[\tilde{\boldsymbol{\Gamma}}\right]_{2k-1,2k-1}\left[\tilde{\boldsymbol{\Gamma}}\right]_{2k,2k}}.\end{aligned} \tag{2.159}$$

Remarks:

- Unlike the known cases with tight sum-rate bound including the two-terminal case [12], the positive-symmetric case [12], and the BEEV-ED case [14], some of the block-circulant cases might not have a tight sum-rate bound if they do not satisfy the requirements in Corollary 1.

- We pick the block-circulant case as an example mainly because of the nice properties in this case that enable us to analytically evaluate the sufficient condition in Theorem 1 without a full search over $\boldsymbol{\Sigma}_{\boldsymbol{N}} \in \mathscr{N}(\boldsymbol{\Sigma}_{\boldsymbol{Y}})$.

- Example 3: another numerical example

Now we give a general numerical example that satisfies the requirement of Corollary 1.

Let $L = 3$,

$$\boldsymbol{\Sigma}_{\boldsymbol{Y}} = \begin{bmatrix} 1.0000 & 0.9500 & 0.7000 \\ 0.9500 & 1.0000 & 0.6000 \\ 0.7000 & 0.6000 & 1.0000 \end{bmatrix}, \tag{2.160}$$

and

$$\boldsymbol{D} = (0.4, 0.45, 0.3)^{\mathrm{T}}. \tag{2.161}$$

Let $\pi = \{1, 2, 3\}$ and

$$\mathbf{\Sigma}_{\boldsymbol{N}} = \begin{bmatrix} 0.4827 & 0.5074 & 0 \\ 0.5074 & 0.6205 & 0 \\ 0 & 0 & 0.0512 \end{bmatrix} \tag{2.162}$$

be a $\pi^{(K)}$ p.d. block diagonal matrix with $K = 1$.

Then the BT minimum sum-rate bound for the MT source coding problem defined by $\mathbf{\Sigma}_{\boldsymbol{Y}}$ and $\boldsymbol{D}$ is tight, since $\tilde{\mathbf{\Gamma}}$ and $\mathbf{\Lambda}$ defined in (2.90) and (2.130) satisfy

$$\operatorname{sign}(\left[\tilde{\mathbf{\Gamma}}\right]_{1,2}) \cdot [\mathbf{\Lambda}]_{1,2} = 0.3596 \geq 0.2815 = 2\left[\tilde{\mathbf{\Gamma}}\right]_{1,2} - \sqrt{\left[\tilde{\mathbf{\Gamma}}\right]_{1,1}\left[\tilde{\mathbf{\Gamma}}\right]_{2,2}}.$$

We have shown that the sum-rate tightness in the above numerical example is ensured by Corollary 1. In addition, it can be verified numerically that it does not satisfy the tightness condition provided by Wang *et al.* [13].

CHAPTER III

RESULTS ON THE SUM-RATE LOSS OF QUADRATIC GAUSSIAN MT SOURCE CODING

In this chapter, Section A reviews related existing results. Section B states our main result on the supremum sum-rate loss over the non-degraded case of quadratic Gaussian MT source coding, followed by its achievability proof and an outline of its converse proof, which is detailed in Section C. Section D gives some discussions and comparisons.

Notation-wise, we denote $\mathbf{0}_{m\times n}$ and $\mathbf{1}_{m\times n}$ as the all-zero and all-one matrix of size $m \times n$, respectively, with the subscript dropped if it is clear from the context.

A. Existing knowledge on the sum-rate loss of Gaussian quadratic MT source coding

We say an $L \times L$ matrix $\mathbf{\Sigma}$ is *symmetric* if $\mathbf{\Sigma} = \boldsymbol{\mathcal{S}}_L(a,b)$ for some $a > 0$ and $-\frac{a}{L-1} < b < a$, with $\boldsymbol{\mathcal{S}}_L(a,b)$ denoting the $L \times L$ matrix whose diagonal elements equal to a with all off-diagonal elements being b. Note that among the L eigenvalues of $\mathbf{\Sigma}$, there are only two distinct numbers $a + (L-1)b$ and $a - b$, with the latter repeated $L-1$ times. In addition, a symmetric matrix $\mathbf{\Sigma}$ is called *positive symmetric* if $b \geq 0$ and *negative symmetric* if $b < 0$. A quadratic Gaussian MT problem is positive- or negative-symmetric if $\mathbf{\Sigma}_{\boldsymbol{Y}}$ is so and $\boldsymbol{D} = D\mathbf{1}$ for some $D > 0$.

Without knowing the exact minimum sum-rate bound $R^{\mathrm{MT}}_{\mathbf{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D})$ for quadratic Gaussian MT source coding in general, little has been done to compute the sum-rate loss $R^{\Delta}_{\mathbf{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D})$. The best known upper bound on the sum-rate loss is one b/s for the two-terminal source coding problem, which is proved by Zamir [39] for continuous source distributions and MSE distortion measure[1]. Since jointly Gaussian sources are continuous, we thus have the following lemma, which is also proved in [40].

Lemma 10 ([39, 40])**.** *For any positive-definite* $\boldsymbol{\Sigma}_{\boldsymbol{Y}} \in \mathbb{R}^{2\times 2}$ *and any positive real distortion vector* $\boldsymbol{D} = (D_1, D_2)^{\mathrm{T}}$*, it holds that*

$$R^{\Delta}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}) \leq 1 b/s.$$

It is shown in [41] that for two jointly Gaussian sources, as the target distortions D_1 and D_2 go to zero, the sum-rate loss $R^{\Delta}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D})$ also goes to zero. This result is consistent with the Slepian-Wolf theorem [2]. One can thus loosely think of MT source coding as the lossy version of Slepian-Wolf coding. For MT source coding with more than two sources, there is still no prior knowledge about the sum-rate loss.

B. Main result on the supremum sum-rate loss

Now we state our main result on the supremum sum-rate loss of quadratic Gaussian MT source coding.

Theorem 3. *For any* $L \geq 2$ *it holds that*

$$\sup_{(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D}) \in \mathscr{S}_L^{\mathrm{BT}}} R^{\Delta}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}) = L \cdot \max\left[\tau\left(\frac{\lfloor Lx^\star \rfloor}{L}\right), \tau\left(\frac{\lceil Lx^\star \rceil}{L}\right)\right], \tag{3.1}$$

where $\lfloor\cdot\rfloor$ *and* $\lceil\cdot\rceil$ *is respectively the floor and ceiling function,*

$$\mathscr{S}_L^{\mathrm{BT}} = \big\{(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D}) : \exists\, \boldsymbol{\mathcal{D}} \in \mathbb{R}^{L\times L} s.t.\ \mathrm{diag}(\boldsymbol{\mathcal{D}}) = \boldsymbol{D} \textit{ and}$$
$$\boldsymbol{\mathcal{D}} = \boldsymbol{\Sigma}_{\boldsymbol{Y}} - \boldsymbol{\Sigma}_{\boldsymbol{Y}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}} + \boldsymbol{\Lambda})^{-1}\boldsymbol{\Sigma}_{\boldsymbol{Y}} \textit{ for some p.s.d. diagonal } \boldsymbol{\Lambda}\big\}, \tag{3.2}$$

[1] In the same paper [39], Zamir also conjectured that the supremum rate loss for the Wyner-Ziv problem with MSE distortion measure is 0.1083 b/s.

$\tau : [0,1] \mapsto \mathbb{R}$ *is defined as*

$$\tau(x) \triangleq \begin{cases} \frac{x}{2}\log_2\left[\frac{1+x}{2}+\frac{1}{2}\sqrt{(1-x)(5-x)}\right]+\frac{1-x}{2}\log_2\left[\frac{(1+3x)\sqrt{1-x}}{(1+x)\sqrt{1-x}+x\sqrt{5-x}}\right], & x<1, \\ 0, & x=1, \end{cases}$$

and

$$x^\star \triangleq \arg\max_{x\in[0,1]} \tau(x) \approx 0.8151108221.$$

Theorem 3 gives the supremum sum-rate loss in the quadratic Gaussian MT problem under the *non-degraded* assumption $(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D}) \in \mathscr{S}_L^{\mathrm{BT}}$, meaning that all target distortions are simultaneously achievable by BT schemes. Fig. 7 plots this supremum sum-rate loss as a function of L. We observe that the supremum increases almost linearly in L. This observation is confirmed by the following corollary, which shows that the asymptotic slope of the supremum sum-rate loss equals to

$$l^\star \;\triangleq\; \max_{x\in[0,1]} \tau(x) \;=\; \tau(x^\star) \;=\; 0.1083256073.$$

The asymptotic function $0.1083L$ is also plotted in Fig. 7 for comparison.

Corollary 2. *It holds that*

$$\lim_{L\to\infty}\left[\frac{1}{L}\sup_{(\boldsymbol{\Sigma}_{\boldsymbol{Y}},\boldsymbol{D})\in\mathscr{S}_L^{\mathrm{BT}}} R^{\Delta}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D})\right] = l^\star.$$

Proof. We have

$$\begin{aligned}\lim_{L\to\infty}\left[\max\left(\tau(\frac{\lfloor Lx^\star\rfloor}{L}),\tau(\frac{\lceil Lx^\star\rceil}{L})\right)\right] &= \lim_{L\to\infty}\tau\left(\frac{1}{L}\cdot\arg\max_{N\in\{\lfloor Lx^\star\rfloor,\lceil Lx^\star\rceil\}}\tau\left(\frac{N}{L}\right)\right)\\ &= \tau\left(\lim_{L\to\infty}\left[\frac{1}{L}\cdot\arg\max_{N\in\{\lfloor Lx^\star\rfloor,\lceil Lx^\star\rceil\}}\tau\left(\frac{N}{L}\right)\right]\right) \qquad (3.3)\\ &= \tau\left(x^\star\right) = l^\star, \qquad (3.4)\end{aligned}$$

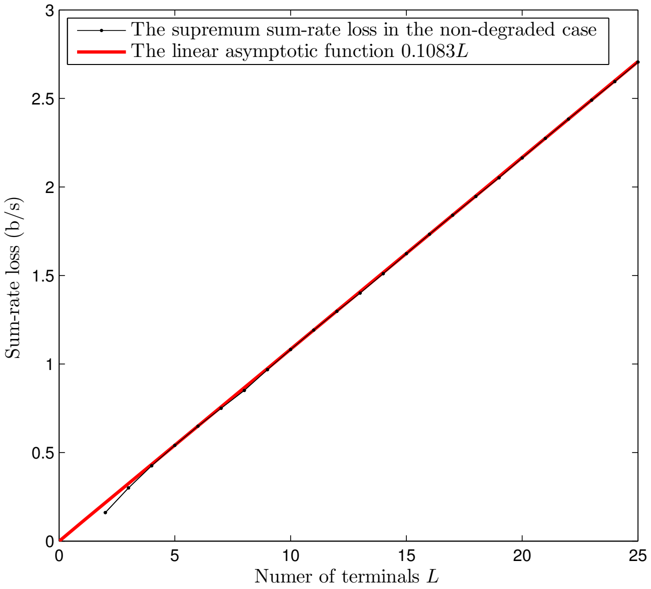


Fig. 7. The supremum sum-rate loss in the non-degraded case and its linear asymptotic function.

where (3.3) is due to the fact that continuous functions transform limits into limits, and (3.4) holds because

$$\frac{1}{L} \cdot \arg \max_{N \in \{\lfloor Lx^\star \rfloor, \lceil Lx^\star \rceil\}} \tau\left(\frac{N}{L}\right) \in \left[x^\star - \frac{1}{L},\ x^\star + \frac{1}{L}\right).$$

□

To prove Theorem 3, we need to show that the right-hand-side (r.h.s) of (3.1) is both an lower bound and upper bound on the supremum sum-rate loss $\sup_{(\boldsymbol{\Sigma_Y}, \boldsymbol{D}) \in \mathscr{S}_L^{\mathrm{BT}}} R_{\boldsymbol{\Sigma_Y}}^{\Delta}(\boldsymbol{D})$, which will be referred to as the achievability proof and the converse proof, respectively. In the next two subsections, we provide the achievability proof and an outline of the converse proof, whose detail is given in Section C.

1. The achievability proof

We need to show that

$$\sup_{(\boldsymbol{\Sigma}_{\boldsymbol{Y}},\boldsymbol{D})\,\in\,\mathscr{S}_L^{\mathrm{BT}}} R^{\Delta}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}) \geq L\cdot\max\left[\tau\left(\frac{\lfloor Lx^{\star}\rfloor}{L}\right),\tau\left(\frac{\lceil Lx^{\star}\rceil}{L}\right)\right]. \tag{3.5}$$

To do this, we provide a sequence of quadratic Gaussian MT problems that belongs to the BEEV-ED subclass for which the limit of sum-rate loss $R^{\Delta}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D})$ equals to the r.h.s. of (3.5).

For a fixed L, denote

$$N^{\star} \;\triangleq\; \arg\max_{N\in\{\lfloor Lx^{\star}\rfloor,\lceil Lx^{\star}\rceil\}} \tau\left(\frac{N}{L}\right), \tag{3.6}$$

and define a sequence of covariance matrices $\boldsymbol{\Sigma}_{\boldsymbol{Y}}(k;L)$, $k\in\mathbb{N}$ as

$$\boldsymbol{\Sigma}_{\boldsymbol{Y}}(k;L) = \boldsymbol{F}_L\cdot\mathrm{diag}\left(\left(\nu\left(\frac{N^{\star}}{L}\right),k\right)_{\mathcal{A},\mathcal{L}-\mathcal{A}}\right)\cdot\boldsymbol{F}_L^{\mathrm{T}}, \tag{3.7}$$

with $\boldsymbol{F}_L$ being the $L\times L$ *real Fourier matrix* [14], $(\lambda,\Lambda)_{\mathcal{A},\mathcal{L}-\mathcal{A}}$ as a length-L vector whose $N^{\star}$ elements indexed by $\boldsymbol{\mathcal{A}}$ are λ with all the rest being Λ, $\nu:[0,1]\mapsto\mathbb{R}$ is a continuous function defined as

$$\nu(x)\triangleq\frac{-x^2+4x-1+(1-x)\sqrt{(1-x)(5-x)}}{2},$$

where

$$\mathcal{A} = \begin{cases} \{1\} \cup \bigcup_{i \in \mathcal{B}} \{i, L+2-i\}, & L \text{ is odd and } N^\star \text{ is odd} \\ \bigcup_{i \in \mathcal{B}} \{i, L+2-i\}, & L \text{ is odd and } N^\star \text{ is even} \\ \{1\} \cup \bigcup_{i \in \mathcal{B}} \{i, L+2-i\} \text{ or } \left\{\frac{L}{2}+1\right\} \cup \bigcup_{i \in \mathcal{B}} \{i, L+2-i\}, & L \text{ is even and } N^\star \text{ is odd} \\ \left\{1, \frac{L}{2}+1\right\} \cup \bigcup_{i \in \mathcal{C}} \{i, L+2-i\} \text{ or } \bigcup_{i \in \mathcal{B}} \{i, L+2-i\}, & L \text{ is even and } N^\star \text{ is even} \end{cases} \tag{3.8}$$

where $\mathcal{B}, \mathcal{C} \subset \left\{2, 3, \ldots, \left\lfloor \frac{L}{2} \right\rfloor\right\}$ with $|\mathcal{B}| = \left\lfloor \frac{N^\star}{2} \right\rfloor$ and $|\mathcal{C}| = \left\lfloor \frac{N^\star}{2} \right\rfloor - 1$.

Then we have the following lemma which ensures that the sequence of pairs $(\mathbf{\Sigma}_{\boldsymbol{Y}}(k; L), \mathbf{1})$ can approach the r.h.s. of (3.5).

Lemma 11. *As $k \to \infty$, it holds that*

$$\lim_{k \to \infty} R^{\Delta}_{\mathbf{\Sigma}_{\boldsymbol{Y}}(k;L)}(\mathbf{1}) = L \cdot \max \left[\tau\left(\frac{\lfloor L x^\star \rfloor}{L} \right), \tau\left(\frac{\lceil L x^\star \rceil}{L} \right) \right].$$

Proof. To prove Lemma 11, we first compute a general formula for the sum-rate loss between distributed encoding and joint encoding of Gaussian sources with BEEV covariance matrix $\mathbf{\Sigma} = \mathcal{B}_{L,N}(\lambda, \Lambda, \mathcal{A}, \boldsymbol{T})$ and equal target distortion $\boldsymbol{D} = D\mathbf{1}$, where

$$\mathcal{B}_{L,N}(\lambda, \Lambda, \mathcal{A}, \boldsymbol{T}) \ \overset{\Delta}{=} \ \boldsymbol{T} \text{diag}((\lambda, \Lambda)_{\mathcal{A}, \mathcal{L}-\mathcal{A}}) \boldsymbol{T}^{\mathrm{T}},$$

for some set $\mathcal{A} \subset \mathcal{L}$ with $|\mathcal{A}| = N$, $\Lambda > \lambda > 0$, and orthogonal matrix $\boldsymbol{T}$ satisfying [14]

$$\sum_{j \in \mathcal{A}} \boldsymbol{T}_{i,j}^2 = \frac{N}{L} \quad \text{for any} \quad i \in \mathcal{L}. \tag{3.9}$$

Note that in the trivial case with $D \geq \frac{N\lambda + (L-N)\Lambda}{L}$, both $R^{\mathrm{MT}}_{\mathbf{\Sigma}_{\boldsymbol{Y}}}(D\mathbf{1})$ and $R^{\mathrm{Joint}}_{\mathbf{\Sigma}_{\boldsymbol{Y}}}(D\mathbf{1})$

are zero, hence we can assume that

$$D \;<\; \frac{N\lambda + (L-N)\Lambda}{L}.$$

It is proved in [14] that the distributed encoding minimum sum-rate in this BEEV-ED case is given by

$$R^{\mathrm{MT}}_{\boldsymbol{\Sigma_Y}}(D\mathbf{1}) + R^{\mathrm{BT}}_{\boldsymbol{\Sigma_Y}}(D\mathbf{1}) = \frac{L-N}{2}\log_2(1+\Lambda p) + \frac{N}{2}\log_2(1+\lambda p), \tag{3.10}$$

where p is the solution to

$$\frac{L-N}{\frac{1}{\Lambda}+p} + \frac{N}{\frac{1}{\lambda}+p} = LD. \tag{3.11}$$

On the other hand, the joint encoding minimum sum-rate in the BEEV-ED case is given by the reverse water-filling formula as

$$R^{\mathrm{Joint}}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}) = \frac{L-N}{2}\log_2\left(\frac{\Lambda}{\min\{\Lambda, w\}}\right) + \frac{N}{2}\log_2\left(\frac{\lambda}{\min\{\lambda, w\}}\right), \tag{3.12}$$

where w is the unique solution to

$$(L-N)\min\{\Lambda, w\} + N\min\{\lambda, w\} = LD. \tag{3.13}$$

Combining (3.10) with (3.12), we know that the sum-rate loss in the BEEV-ED case is given by

$$\begin{aligned} R^{\Delta}_{\boldsymbol{\Sigma_Y}}(D\mathbf{1}) &= R^{\mathrm{MT}}_{\boldsymbol{\Sigma_Y}}(D\mathbf{1}) - R^{\mathrm{Joint}}_{\boldsymbol{\Sigma_Y}}(D\mathbf{1}) \\ &= \frac{L-N}{2}\log_2\left[\left(\frac{1}{\Lambda}+p\right)\min\{\Lambda, w\}\right] + \frac{N}{2}\log_2\left[\left(\frac{1}{\lambda}+p\right)\min\{\lambda, w\}\right], \end{aligned} \tag{3.14}$$

where p and w are the solutions to (3.11) and (3.13), respectively.

Now we prove Lemma 11. First, it is easy to verify that $\boldsymbol{\Sigma_Y}(k;L) = \mathcal{B}_{L,N^\star}\left(\nu\left(\frac{N^\star}{L}\right), k, \boldsymbol{\mathcal{A}}, \boldsymbol{F}_L\right)$

with $\mathcal{A}$ and $\boldsymbol{F}_L$ satisfying the requirement (3.9). Hence we can use (3.14) to compute the sum-rate loss for each pair $(\boldsymbol{\Sigma}_{\boldsymbol{Y}}(k;L), \mathbf{1})$. Solving for w and p, we obtain

$$\begin{aligned} w &= \frac{L - N^\star \lambda}{L - N^\star}, \\ p &= \sqrt{\frac{1}{4} + \frac{1}{2\lambda} + \frac{1}{4\lambda^2} - \frac{N^\star}{L\lambda} - \frac{1}{2k} - \frac{1}{2k\lambda} + \frac{N^\star}{kL} + \frac{1}{4k^2}} + \frac{L\lambda - L - \lambda}{2L\lambda}, \end{aligned}$$

where $\lambda = \nu(\frac{N^\star}{L})$. Hence

$$\begin{aligned} R^{\Delta}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}(k;L)}(\mathbf{1}) = & \frac{L - N^\star}{2} \log_2 \left[\left(\frac{1}{k} + p \right) \min \{k, w\} \right] \\ & + \frac{N^\star}{2} \log_2 \left[\left(\nu^{-1} \left(\frac{N^\star}{L} \right) + p \right) \min \left\{ \nu \left(\frac{N^\star}{L} \right), w \right\} \right] \\ = & \frac{L - N^\star}{2} \log_2 \left[w \left(\frac{1}{k} + p \right) \right] + \frac{N^\star}{2} \log_2 (1 + \lambda p) . \end{aligned}$$

Let $k \to \infty$, then

$$p = \sqrt{\frac{1}{4} + \frac{1}{2\lambda} + \frac{1}{4\lambda^2} - \frac{N^\star}{L\lambda}} + \frac{L\lambda - L - \lambda}{2L\lambda} = 1 - \frac{2N^\star}{L + N^\star + \sqrt{(L - N^\star)(5L - N^\star)}},$$

and it is easy to verify that

$$R^{\Delta}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}(k;L)}(\mathbf{1}) \quad \overset{k \to \infty}{=} \quad \tau \left(\frac{N^\star}{L} \right).$$

Hence Lemma 11 is proved. □

Lemma 11 directly leads to (3.5) after verifying $(\boldsymbol{\Sigma}_{\boldsymbol{Y}}(k;L), \mathbf{1}) \in \mathscr{S}_L^{\mathrm{BT}}$. Though its detailed proof is postponed to Appendix A, we give several examples of the above defined $\boldsymbol{\Sigma}_{\boldsymbol{Y}}(k;L)$ matrices.

For $L = 5$, the supremum sum-rate loss of 0.54103 b/s can be approached from below by the sequence of pairs $(\boldsymbol{\Sigma}_{\boldsymbol{Y}}(k;5), \mathbf{1})$ defined in (3.7) with $N^\star = 4$ and $\mathcal{A} =$

$\{2, 3, 4, 5\}$, i.e.,

$$\begin{aligned}\boldsymbol{\Sigma}_{\boldsymbol{Y}}(k;5) &= \boldsymbol{F}_5 \cdot \operatorname{diag}\left(k, \nu\left(\frac{4}{5}\right), \nu\left(\frac{4}{5}\right), \nu\left(\frac{4}{5}\right), \nu\left(\frac{4}{5}\right)\right) \cdot \boldsymbol{F}_5^{\mathrm{T}} \\ &= \frac{1}{5}\boldsymbol{\mathcal{S}}_5\left(k + 4\nu\left(\frac{4}{5}\right), k - \nu\left(\frac{4}{5}\right)\right)\end{aligned}$$

with $\boldsymbol{\mathcal{S}}_L(a, b)$ denoting the $L \times L$ matrix whose diagonal elements equal to a with all off-diagonal elements being b, i.e., the supremum sum-rate loss for $L = 5$ can be approached in the positive symmetric case. Clearly, as $k \to \infty$,

$$\lim_{k\to\infty} \frac{5}{k}\boldsymbol{\Sigma}_{\boldsymbol{Y}}(k;5) \quad = \quad \mathbf{1}_{5\times 5}.$$

Remark 1: One can easily verify that for any $L \leq 7$, it is always true that $N^\star = L-1$, i.e., the supremum sum-rate loss in the non-degraded case can be achieved when $\boldsymbol{\Sigma}_{\boldsymbol{Y}}$ is BEEV with $L - 1$ small eigenvalues and one large eigenvalue, which is indeed the positive symmetric case defined in Section A. Conversely, $N^\star < L - 1$ holds for any $L > 7$. Hence the supremum sum-rate loss in the non-degraded case can be achieved in the positive symmetric case *if and only if* $L \leq 7$.

For $L = 8$, the supremum sum-rate loss of 0.85120 b/s can be approached from below by the sequence of pairs $(\boldsymbol{\Sigma}_{\boldsymbol{Y}}(k;8), \mathbf{1})$ defined in (3.7) with $N^\star = 6$ and $\mathcal{A} = \{2, 3, 4, 6, 7, 8\}$, i.e.,

$$\boldsymbol{\Sigma}_{\boldsymbol{Y}}(k;8) = \boldsymbol{F}_8 \cdot \operatorname{diag}\left(k, \nu\left(\frac{3}{4}\right), \nu\left(\frac{3}{4}\right), \nu\left(\frac{3}{4}\right), k, \nu\left(\frac{3}{4}\right), \nu\left(\frac{3}{4}\right), \nu\left(\frac{3}{4}\right)\right) \cdot \boldsymbol{F}_8^{\mathrm{T}},$$

which, after reordering (without affecting the sum-rate loss), is equivalent to a block-positive-symmetric matrix which satisfies

$$\lim_{k\to\infty} \frac{4}{k}\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{Y}}(k;8) \quad = \quad \begin{bmatrix} \mathbf{1}_{4\times 4} & \mathbf{0}_{4\times 4} \\ \mathbf{0}_{4\times 4} & \mathbf{1}_{4\times 4} \end{bmatrix}.$$

Remark 2: $L = 8$ is not the only case when the supremum sum-rate loss can be approached in the block-positive-symmetric case. In fact, if

$$N^\star < L - 1 \ \&\ L = L' \cdot (L - N^\star) \text{ with } L' \in \{4, 5, 6\}\,, \tag{3.15}$$

then the supremum sum-rate loss in the L-terminal case equals to $(L - N^\star)$ times that in the L'-terminal case, hence can be approached in the block-positive-symmetric case with block size $L' \times L'$. Conversely, using the strict concavity of $\tau(x)$ function, it is not hard to show that (3.15) is also a necessary condition for the supremum to be approachable in the block-positive-symmetric case. Furthermore, it is easy to check that (3.15) holds *if and only if*

$$L \in \{8, 10, 12, 15, 18, 20, 24, 25, 30\}\,. \tag{3.16}$$

In general, the supremum in (3.1) cannot be approached in (block-)positive-symmetric cases. For example, when $L = 11$, the supremum sum-rate loss of 1.19152 b/s is approached from below by the sequence of pairs $(\boldsymbol{\Sigma}_{\boldsymbol{Y}}(k; 11), \mathbf{1})$ defined in (3.7) with $N^\star = 9$ and $\mathcal{A} = \{1, 2, 3, 5, 6, 7, 8, 10, 11\}$, i.e.,

$$\begin{aligned}\boldsymbol{\Sigma}_{\boldsymbol{Y}}(k; 11) = \boldsymbol{F}_{11} \cdot \operatorname{diag}\Bigg(&\nu\left(\frac{9}{11}\right), \nu\left(\frac{9}{11}\right), \nu\left(\frac{9}{11}\right), k, \nu\left(\frac{9}{11}\right),\\ &\nu\left(\frac{9}{11}\right), \nu\left(\frac{9}{11}\right), \nu\left(\frac{9}{11}\right), k, \nu\left(\frac{9}{11}\right), \nu\left(\frac{9}{11}\right)\Bigg) \cdot \boldsymbol{F}_{11}^{\mathrm{T}},\end{aligned}$$

which is a circulant symmetric matrix that satisfies

$$\lim_{k\to\infty}\frac{11}{2k}\boldsymbol{\Sigma}_{\boldsymbol{Y}}(k;11)=\begin{bmatrix} 1 & a_1 & a_2 & a_3 & a_4 & a_5 & a_5 & a_4 & a_3 & a_2 & a_1 \\ a_1 & 1 & a_1 & a_2 & a_3 & a_4 & a_5 & a_5 & a_4 & a_3 & a_2 \\ a_2 & a_1 & 1 & a_1 & a_2 & a_3 & a_4 & a_5 & a_5 & a_4 & a_3 \\ a_3 & a_2 & a_1 & 1 & a_1 & a_2 & a_3 & a_4 & a_5 & a_5 & a_4 \\ a_4 & a_3 & a_2 & a_1 & 1 & a_1 & a_2 & a_3 & a_4 & a_5 & a_5 \\ a_5 & a_4 & a_3 & a_2 & a_1 & 1 & a_1 & a_2 & a_3 & a_4 & a_5 \\ a_5 & a_5 & a_4 & a_3 & a_2 & a_1 & 1 & a_1 & a_2 & a_3 & a_4 \\ a_4 & a_5 & a_5 & a_4 & a_3 & a_2 & a_1 & 1 & a_1 & a_2 & a_3 \\ a_3 & a_4 & a_5 & a_5 & a_4 & a_3 & a_2 & a_1 & 1 & a_1 & a_2 \\ a_2 & a_3 & a_4 & a_5 & a_5 & a_4 & a_3 & a_2 & a_1 & 1 & a_1 \\ a_1 & a_2 & a_3 & a_4 & a_5 & a_5 & a_4 & a_3 & a_2 & a_1 & 1 \end{bmatrix} \tag{3.17}$$

with

$$a_1=-0.142, a_2=-0.960, a_3=0.415, a_4=0.841, a_5=-0.655.$$

2. Outline of the converse proof

We need to show that

$$\sup_{(\boldsymbol{\Sigma}_{\boldsymbol{Y}},\boldsymbol{D})\in\mathscr{S}_L^{\mathrm{BT}}} R^{\Delta}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D})\leq L\cdot\max\left[\tau\left(\frac{\lfloor Lx^\star\rfloor}{L}\right),\tau\left(\frac{\lceil Lx^\star\rceil}{L}\right)\right]. \tag{3.18}$$

The direct proof is started by noting the fact that

$$\sup_{(\boldsymbol{\Sigma}_{\boldsymbol{Y}},\boldsymbol{D})\in\mathscr{S}_L^{\mathrm{BT}}} R^{\Delta}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D})\leq\sup_{(\boldsymbol{\Sigma}_{\boldsymbol{Y}},\boldsymbol{D})\in\mathscr{S}_L^{\mathrm{BT}}} R^{\mathrm{BT}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D})-R^{\mathrm{Joint}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}), \tag{3.19}$$

which is due to Lemma 1 and the definition of $R^{\Delta}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D})$.

The rest of converse proof contains three steps.

- First, we show that to compute the r.h.s of (3.19), we only need to search over a subclass of the non-degraded cases $\mathscr{S}_L^{\mathrm{BT}}$ called *regular cases*, for which the difference $R_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}^{\mathrm{BT}}(\boldsymbol{D}) - R_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}^{\mathrm{Joint}}(\boldsymbol{D})$ can be further upper-bounded by a function that only depends on the eigenvalues of $\boldsymbol{\Sigma}_{\boldsymbol{Y}}$. We then formulate the core optimization problem $\mathbb{P}_0^L$ over the eigenvalues of $\boldsymbol{\Sigma}_{\boldsymbol{Y}}$ by allowing the eigenvalues to take the value of infinity, so that supremum can be achievable and thus be replaced by maximum. However, the resulting optimization problem belongs to the class called nonconvex *mixed-integer nonlinear programming* (MINLP) problems [42, Section 1.1], which is NP-complete in general [42].

- To solve the above MINLP problem, we prove that its optimal solution must be achieved in the case when the eigenvalues of $\boldsymbol{\Sigma}_{\boldsymbol{Y}}$ take at most four distinct values. Hence $\mathbb{P}_0^L$ is simplified to an equivalent optimization problem $\mathbb{P}_1^L$. Then we formulate a set of auxiliary optimization problem $\mathbb{P}_2(x)$ with $x \in [0, 1)$ by relaxing the integer design variables in $\mathbb{P}_1^L$ to take continuous value, and show that for a given L, the largest number among the L maximum function values of the auxiliary problems $\left\{\mathbb{P}_2(x) : x \in \left\{0, \frac{1}{L}, ..., \frac{L-1}{L}\right\}\right\}$ is an upper bound on that of the core optimization problem $\mathbb{P}_0^L$. This upper bound is tight if it is achieved when the eigenvalues of $\boldsymbol{\Sigma}_{\boldsymbol{Y}}$ only takes two distinct values, one of which is a finite function of L with the other being infinity.

- Finally, we separately treat the cases when $L \notin \{2, 3, 4, 8\}$ and $L \in \{2, 3, 4, 8\}$. In the former case, it is proved using *rigorous numerical methods* (interval arithmetic to be specific) that the above upper bound is indeed achieved in the bi-eigen case. In the latter case, we directly solve $\mathbb{P}_1^L$ by exhausting all possible combinations of the integer design variables. Fortunately, the maximum function value is also achieved in the above bi-eigen case. The last step is to verify that the maximum function value in $\mathbb{P}_0^L$ equals to the r.h.s. of (3.18) for all L.

C. The converse proof of Theorem 3

In this section, we give the complete converse proof of Theorem 3. The three major steps are summarized as: formulation of the core optimization problem, reduction to the quad-eigen case and relaxation, and solution via rigorous numerical methods, with details provided in the following three subsections, respectively.

1. The core optimization problem

Define

$$\mathscr{D}^{\mathrm{BT}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}) = \left\{\boldsymbol{\mathcal{D}} : \boldsymbol{\mathcal{D}} = \boldsymbol{\Sigma}_{\boldsymbol{Y}} - \boldsymbol{\Sigma}_{\boldsymbol{Y}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}} + \boldsymbol{\Lambda})^{-1}\boldsymbol{\Sigma}_{\boldsymbol{Y}} \text{ for some p.s.d. and diagonal } \boldsymbol{\Lambda}\right\}, \tag{3.20}$$

and

$$\mathscr{D}^{\mathrm{BT}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D}) = \left\{\boldsymbol{\mathcal{D}} \in \mathscr{D}^{\mathrm{BT}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}) : \mathrm{diag}(\boldsymbol{\mathcal{D}}) \leq \boldsymbol{D}\right\},$$
$$\mathscr{D}_{=}^{\mathrm{BT}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D}) = \left\{\boldsymbol{\mathcal{D}} \in \mathscr{D}^{\mathrm{BT}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}) : \mathrm{diag}(\boldsymbol{\mathcal{D}}) = \boldsymbol{D}\right\}.$$

Also define

$$\mathscr{D}^{\mathrm{Joint}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}) = \left\{\boldsymbol{\mathcal{D}} : \boldsymbol{\mathcal{D}}^{\mathrm{T}} = \boldsymbol{\mathcal{D}} \text{ and } \boldsymbol{0} \preceq \boldsymbol{\mathcal{D}} \preceq \boldsymbol{\Sigma}_{\boldsymbol{Y}}\right\},$$

and

$$\mathscr{D}^{\mathrm{Joint}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D}) = \left\{\boldsymbol{\mathcal{D}} \in \mathscr{D}^{\mathrm{Joint}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}) : \mathrm{diag}(\boldsymbol{\mathcal{D}}) \leq \boldsymbol{D}\right\},$$
$$\mathscr{D}_{=}^{\mathrm{Joint}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D}) = \left\{\boldsymbol{\mathcal{D}} \in \mathscr{D}^{\mathrm{Joint}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}) : \mathrm{tr}(\boldsymbol{\mathcal{D}}) = \sum_{i \in \mathcal{L}} D_i\right\}.$$

In words, $\mathscr{D}^{\mathrm{BT}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}})$ and $\mathscr{D}^{\mathrm{Joint}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}})$ denote the set of distortion matrices that are *BT-achievable* and *joint-achievable* for a given source covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{Y}}$, respectively. $\mathscr{D}^{\mathrm{BT}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})$ and $\mathscr{D}^{\mathrm{Joint}}_{=}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})$ contain all BT- and joint-achievable distortion matrices that meet the distortion constraints defined by $\boldsymbol{D}$, respectively. And $\mathscr{D}^{\mathrm{BT}}_{=}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})$ further restricts to the BT-achievable matrices that meet *all* the distortion constraints *with equalities*, while $\mathscr{D}^{\mathrm{Joint}}_{=}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})$ the joint-achievable matrices that achieve a *sum-distortion* of $\sum_{i\in\mathcal{L}} D_i$. Then the following relationships are obvious,

$$\begin{aligned} \mathscr{D}^{\mathrm{BT}}_{=}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D}) \subset \mathscr{D}^{\mathrm{BT}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D}) \subset \mathscr{D}^{\mathrm{BT}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}), \\ \mathscr{D}^{\mathrm{Joint}}_{=}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D}) \subset \mathscr{D}^{\mathrm{Joint}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D}) \subset \mathscr{D}^{\mathrm{Joint}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}). \end{aligned} \tag{3.21}$$

In particular, it is proved in [13, Theorem 4] that if $(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D}) \in \mathscr{S}^{\mathrm{BT}}_{L}$, $\mathscr{D}^{\mathrm{BT}}_{=}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})$ must be a singleton set, hence we can denote the single element in $\mathscr{D}^{\mathrm{BT}}_{=}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})$ as $\boldsymbol{\mathcal{D}}^{\mathrm{BT}}_{=}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})$.

We say a pair $(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})$ is *regular* if $(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D}) \in \mathscr{S}^{\mathrm{BT}}_{L}$,

$$\boldsymbol{\mathcal{D}}^{\mathrm{BT}}_{=}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D}) = \boldsymbol{\Sigma}_{\boldsymbol{Y}} - \boldsymbol{\Sigma}_{\boldsymbol{Y}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}} + \frac{1}{p} \cdot \boldsymbol{I})^{-1}\boldsymbol{\Sigma}_{\boldsymbol{Y}} \tag{3.22}$$

for some real number $p > 0$, and $\sum_{i\in\mathcal{L}} D_i = L$. Denote $\mathscr{R}_L$ as the set of all regular $(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})$ pairs. Then we have the following lemma, which shows that any pair $(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D}) \in \mathscr{S}^{\mathrm{BT}}_{L}$ can be *regularized* without affecting the distributed encoding and joint encoding minimum sum-rates, hence the sum-rate loss between them.

Lemma 12. *For any pair $(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D}) \in \mathscr{S}^{\mathrm{BT}}_{L}$, there exists a regular pair $(\boldsymbol{\Sigma}_{\boldsymbol{Y}'}, \boldsymbol{D}') \in \mathscr{R}_L$ such that $R^{\mathrm{MT}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}) = R^{\mathrm{MT}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}'}}(\boldsymbol{D}')$ and $R^{\mathrm{Joint}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}) = R^{\mathrm{Joint}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}'}}(\boldsymbol{D}')$, hence*

$$R^{\Delta}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}) \;=\; R^{\Delta}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}'}}(\boldsymbol{D}'). \tag{3.23}$$

In addition, for any $\boldsymbol{\Sigma}_{\boldsymbol{Y}}$*, if* $\mathrm{tr}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}) \geq L$ *there exists a unique* $\boldsymbol{D}$ *such that* $(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D}) \in \mathscr{R}_L$*, otherwise no such* $\boldsymbol{D}$ *exists.*

Proof. Due to the definition of $\mathscr{D}_{=}^{\mathrm{BT}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})$, we assume that $\boldsymbol{\mathcal{D}}_{=}^{\mathrm{BT}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})$ satisfies

$$\boldsymbol{\mathcal{D}}_{=}^{\mathrm{BT}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D}) = \boldsymbol{\Sigma}_{\boldsymbol{Y}} - \boldsymbol{\Sigma}_{\boldsymbol{Y}} \left(\boldsymbol{\Sigma}_{\boldsymbol{Y}} + \mathrm{diag}\left(\frac{1}{p_1}, \frac{1}{p_2}, ..., \frac{1}{p_L} \right) \right)^{-1} \boldsymbol{\Sigma}_{\boldsymbol{Y}}$$

for some $p_i > 0$ for $i \in \boldsymbol{\mathcal{L}}$. Denote $\boldsymbol{\Lambda} = \mathrm{diag}\left(\frac{1}{p_1}, \frac{1}{p_2}, ..., \frac{1}{p_L} \right)$, and define

$$p = \frac{1}{L} \sum_{i=1}^{L} p_i D_i, \ \boldsymbol{\mathcal{E}} = \mathrm{diag}\left(\frac{p_1}{p}, \frac{p_2}{p}, \ldots, \frac{p_L}{p} \right), \ \boldsymbol{\Sigma}_{\boldsymbol{Y}'} = \boldsymbol{\mathcal{E}}^{\frac{1}{2}} \boldsymbol{\Sigma}_{\boldsymbol{Y}} \boldsymbol{\mathcal{E}}^{\frac{1}{2}}, \ \boldsymbol{D}' = \boldsymbol{\mathcal{E}} \boldsymbol{D}.$$

Then $(\boldsymbol{\Sigma}_{\boldsymbol{Y}'}, \boldsymbol{D}') \in \mathscr{R}_L$, since

$$\sum_{i \in \mathcal{L}} D_i' = \sum_{i \in \mathcal{L}} \frac{p_i D_i}{p} = L.$$

We also have

$$\begin{aligned} \boldsymbol{\Sigma}_{\boldsymbol{Y}'} - \boldsymbol{\Sigma}_{\boldsymbol{Y}'} \left(\boldsymbol{\Sigma}_{\boldsymbol{Y}'} + \frac{1}{p} \boldsymbol{I} \right)^{-1} \boldsymbol{\Sigma}_{\boldsymbol{Y}'} &= \boldsymbol{\mathcal{E}}^{\frac{1}{2}} \boldsymbol{\Sigma}_{\boldsymbol{Y}} \boldsymbol{\mathcal{E}}^{\frac{1}{2}} - \boldsymbol{\mathcal{E}}^{\frac{1}{2}} \boldsymbol{\Sigma}_{\boldsymbol{Y}} \boldsymbol{\mathcal{E}}^{\frac{1}{2}} \left(\boldsymbol{\mathcal{E}}^{\frac{1}{2}} \boldsymbol{\Sigma}_{\boldsymbol{Y}} \boldsymbol{\mathcal{E}}^{\frac{1}{2}} + \boldsymbol{\mathcal{E}}^{\frac{1}{2}} \boldsymbol{\Lambda} \boldsymbol{\mathcal{E}}^{\frac{1}{2}} \right)^{-1} \boldsymbol{\mathcal{E}}^{\frac{1}{2}} \boldsymbol{\Sigma}_{\boldsymbol{Y}} \boldsymbol{\mathcal{E}}^{\frac{1}{2}} \\ &= \boldsymbol{\mathcal{E}}^{\frac{1}{2}} \boldsymbol{\mathcal{D}}_{=}^{\mathrm{BT}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D}) \boldsymbol{\mathcal{E}}^{\frac{1}{2}}, \end{aligned}$$

and

$$\mathrm{diag}\left(\boldsymbol{\mathcal{E}} \boldsymbol{\mathcal{D}}_{=}^{\mathrm{BT}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D}) \boldsymbol{\mathcal{E}} \right) \ = \ (\boldsymbol{\mathcal{E}} \odot \boldsymbol{\mathcal{E}}) \boldsymbol{D} = \boldsymbol{D}',$$

which implies that

$$\boldsymbol{\mathcal{D}}_{=}^{\mathrm{BT}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}'}, \boldsymbol{D}') \ = \ \boldsymbol{\mathcal{E}}^{\frac{1}{2}} \boldsymbol{\mathcal{D}}_{=}^{\mathrm{BT}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D}) \boldsymbol{\mathcal{E}}^{\frac{1}{2}},$$

due to the definition of $\mathscr{D}_{=}^{\mathrm{BT}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}'}, \boldsymbol{D}')$ and its singleton nature.

Now consider a scheme $(\boldsymbol{\phi}^{(n)}, \boldsymbol{\varphi}^{(n)})$ for the distributed encoding problem of $\boldsymbol{Y}$ with covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{Y}}$ and target distortion vector $\boldsymbol{D}$. Then for the distributed encoding problem of $\boldsymbol{Y}'$ with covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{Y}'}$ and target distortion vector $\boldsymbol{D}'$, consider the scheme $(\tilde{\boldsymbol{\phi}}^{(n)}, \tilde{\boldsymbol{\varphi}}^{(n)})$ that

1. scales the i-th source block $Y_i'^n$ by a factor of $\sqrt{\frac{p}{p_i}}$, then the scaled sources $U_i^n = \sqrt{\frac{p}{p_i}} Y_i'^n$, $i \in \mathcal{L}$ must be i.i.d with covariance matrix $\boldsymbol{\mathcal{E}}^{-\frac{1}{2}} \boldsymbol{\Sigma}_{\boldsymbol{Y}'} \boldsymbol{\mathcal{E}}^{-\frac{1}{2}} = \boldsymbol{\Sigma}_{\boldsymbol{Y}}$,
2. applies $\phi_i^{(n)}$ on the i-th scaled source block U_i^n,
3. reconstructs U_i^n using $\boldsymbol{\varphi}^{(n)}$ as $\hat{U}_i^n$,
4. reconstructs $Y_i'^n$ as $\hat{Y}_i'^n = \sqrt{\frac{p_i}{p}} \hat{U}_i^n$.

Obviously, the new scheme $(\tilde{\boldsymbol{\phi}}^{(n)}, \tilde{\boldsymbol{\varphi}}^{(n)})$ must have the same sum-rate as $(\boldsymbol{\phi}^{(n)}, \boldsymbol{\varphi}^{(n)})$, and achieve a distortion

$$\frac{1}{n} \sum_{j=1}^{n} \mathrm{E}\left[d(Z_{i,j}, \hat{Z}_{i,j}) \right] = \frac{p_i}{p} \cdot \left[\frac{1}{n} \sum_{j=1}^{n} E\left[d(U_{i,j}, \hat{U}_{i,j}) \right] \right] \leq \frac{p_i}{p} D_i = D_i'$$

for $Y_i'^n$. Hence any $(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})$-achievable sum-rate must also be $(\boldsymbol{\Sigma}_{\boldsymbol{Y}'}, \boldsymbol{D}')$-achievable. The converse that $(\boldsymbol{\Sigma}_{\boldsymbol{Y}'}, \boldsymbol{D}')$-achievable implies $(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})$-achievable can be proved in the same way. Hence we must have $R^{\mathrm{MT}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}) = R^{\mathrm{MT}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}'}}(\boldsymbol{D}')$.

Using the same technique, $R^{\mathrm{Joint}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}) = R^{\mathrm{Joint}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}'}}(\boldsymbol{D}')$ comes from the equivalence between $(\boldsymbol{\Sigma}_{\boldsymbol{Y}'}, \boldsymbol{D}')$-joint-achievable and $(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})$-joint-achievable. □

A natural corollary of Lemma 12 is stated as follows.

Corollary 3. *It holds that*

$$\sup_{(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D}) \in \mathscr{S}_L^{\mathrm{BT}}} R^{\Delta}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}) = \sup_{(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D}) \in \mathscr{R}_L} R^{\Delta}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}).$$

Remark 3: We introduce the concept of regularity due to two reasons: first, Corollary 3 ensures that to compute the supremum sum-rate loss over all pairs $(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D}) \in \mathscr{S}_L^{\mathrm{BT}}$, it is sufficient to consider the regular pairs; more importantly, as will be shown below, once a pair $(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})$ is regularized, there exist simple upper/lower bounds on the

BT/joint encoding minimum sum-rates that are expressed only as a function of the eigenvalues of $\boldsymbol{\Sigma_Y}$ (note that $\boldsymbol{D}$ is uniquely determined by $\boldsymbol{\Sigma_Y}$).

The main idea of finding an equivalent (in the sense of (3.23)) regular pair $(\boldsymbol{\Sigma_{Y'}}, \boldsymbol{D'}) \in \mathscr{R}_L$ is based on the fact that simultaneously scaling the i-th source Y_i by a factor of $t_i \neq 0$ and the corresponding target distortion D_i by a factor of t_i^2 does not change the distributed encoding or joint encoding minimum sum-rate. One can also define $\mathscr{R}_L$ as, e.g., the set of $(\boldsymbol{\Sigma_Y}, \boldsymbol{D})$ pairs such that $(\boldsymbol{\Sigma_Y}, \boldsymbol{D}) \in \mathscr{S}_L^{\mathrm{BT}}$ and (3.22) holds with $p = 1$. In fact, as long as (3.22) holds and there is only one degree of freedom in the two values p and $\sum_{i \in \mathcal{L}} D_i$, Lemma 12 is always true. We choose the definition such that $\sum_{i \in \mathcal{L}} D_i = L$ and leave the one degree of freedom to p because this leads to simplifications in the sequel.

Denote the L eigenvalues of $\boldsymbol{\Sigma_Y}$ as λ_i, $i \in \mathcal{L}$ and without loss of generality assume that they are in a non-decreasing order, i.e., $\lambda_i \leq \lambda_j$ for $1 \leq i < j \leq L$. Let $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, ..., \lambda_L)^{\mathrm{T}}$. Assuming $(\boldsymbol{\Sigma_Y}, \boldsymbol{D}) \in \mathscr{R}_L$, it must be true that $\sum_{i \in \mathcal{L}} \lambda_i = \mathrm{tr}(\boldsymbol{\Sigma_Y}) \geq L$, since otherwise the target distortion vector $\boldsymbol{D}$ cannot be achieved by a BT scheme. Then we have an upper bound on the minimum BT sum-rate as well as a lower bound on the minimum joint-encoding rate, which are given in the following lemma.

Lemma 13. *For any $L \geq 2$, the following equations hold*

$$R_{\boldsymbol{\Sigma_Y}}^{\mathrm{BT}}(\boldsymbol{D}) \leq \sum_{i \in \mathcal{L}} \frac{1}{2} \log_2 (1 + \lambda_i p) \triangleq \overline{R}^{\mathrm{BT}}(\boldsymbol{\lambda}), \tag{3.24}$$

$$R_{\boldsymbol{\Sigma_Y}}^{\mathrm{Joint}}(\boldsymbol{D}) \geq \begin{cases} \sum_{i=W+1}^{L} \frac{1}{2} \log_2 \frac{\lambda_i}{w}, & W < L, \\ 0, & W = L \end{cases} \triangleq \underline{R}^{\mathrm{Joint}}(\boldsymbol{\lambda}), \tag{3.25}$$

where p is the solution to

$$\sum_{i\in\mathcal{L}} \frac{\lambda_i}{1+\lambda_i p} = \sum_{i\in\mathcal{L}} D_i = L, \tag{3.26}$$

the water level w equals to one when $\sum_{i\in\mathcal{L}} \lambda_i = L$, and otherwise equals to the unique solution to the reverse water-filling problem [15]

$$\sum_{i\in\mathcal{L}} \min(\lambda_i, w) = L \tag{3.27}$$

and

$$W \;=\; |\,\{i \in \mathcal{L} : \lambda_i < w\}\,| \tag{3.28}$$

with $|A|$ denoting the cardinality of the set A.

Proof. We first upper-bound the minimum BT sum-rate as

$$\begin{aligned}
R^{\mathrm{BT}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}) &= \min_{\boldsymbol{U}\,\in\,\mathcal{U}(\boldsymbol{\Sigma}_{\boldsymbol{Y}},\boldsymbol{D})} I(\boldsymbol{Y};\boldsymbol{U}) \qquad &(3.29)\\
&= \min_{\boldsymbol{\mathcal{D}}\,\in\,\mathscr{D}^{\mathrm{BT}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}},\boldsymbol{D})} \frac{1}{2}\log_2 \frac{\det^+(\boldsymbol{\Sigma}_{\boldsymbol{Y}})}{\det^+(\boldsymbol{\mathcal{D}})} &(3.30)\\
&\le \frac{1}{2}\log_2 \frac{\det^+(\boldsymbol{\Sigma}_{\boldsymbol{Y}})}{\det^+(\boldsymbol{\mathcal{D}}^{\mathrm{BT}}_{=}(\boldsymbol{\Sigma}_{\boldsymbol{Y}},\boldsymbol{D}))} &(3.31)\\
&= \frac{1}{2}\log_2 \frac{\prod_{i\in\mathcal{L}:\lambda_i>0}\lambda_i}{\prod_{i\in\mathcal{L}:\lambda_i>0}\frac{\lambda_i}{1+\lambda_i p}} &(3.32)\\
&= \sum_{i\in\mathcal{L}} \frac{1}{2}\log_2\left[1+\lambda_i p\right] = \overline{R}^{\mathrm{BT}}(\boldsymbol{\lambda}),
\end{aligned}$$

where (3.29) and (3.30) come from the definitions of $R^{\mathrm{BT}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D})$ and $\mathscr{D}^{\mathrm{BT}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}},\boldsymbol{D})$, respectively, with $\det^+(\boldsymbol{A})$ denoting the product of positive eigenvalues of matrix $\boldsymbol{A}$, (3.31) is due to the fact that $\boldsymbol{\mathcal{D}}^{\mathrm{BT}}_{=}(\boldsymbol{\Sigma}_{\boldsymbol{Y}},\boldsymbol{D}) \in \mathscr{D}^{\mathrm{BT}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}},\boldsymbol{D})$, and (3.32) is true since in the regular case, $\boldsymbol{\mathcal{D}}^{\mathrm{BT}}_{=}(\boldsymbol{\Sigma}_{\boldsymbol{Y}},\boldsymbol{D})$ must equal to $\boldsymbol{\Sigma}_{\boldsymbol{Y}} - \boldsymbol{\Sigma}_{\boldsymbol{Y}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}} + \frac{1}{p}\cdot\boldsymbol{I})^{-1}\boldsymbol{\Sigma}_{\boldsymbol{Y}}$ (whose eigenvalues are $\frac{\lambda_i}{1+\lambda_i p}$ for $i \in \boldsymbol{\mathcal{L}}$) with p being the solution to

$$\sum_{i\in\mathcal{L}} \frac{\lambda_i}{1+\lambda_i p} = \mathrm{tr}(\boldsymbol{\mathcal{D}}^{\mathrm{BT}}_{=}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})) \;=\; \sum_{i\in\mathcal{L}} D_i \;=\; L.$$

Similarly, we obtain a lower bound on the joint encoding minimum sum-rate,

$$\begin{aligned}
R^{\mathrm{Joint}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D}) &= \min_{\boldsymbol{V}:\, \mathrm{diag}\{E[(\boldsymbol{Y}-E(\boldsymbol{Y}|\boldsymbol{V}))(\boldsymbol{Y}-E(\boldsymbol{Y}|\boldsymbol{V}))^{\mathrm{T}}]\} \,\leq \boldsymbol{D}} I(\boldsymbol{Y};\boldsymbol{V}) && (3.33)\\
&= \min_{\boldsymbol{\mathcal{D}}\,\in\,\mathscr{D}^{\mathrm{Joint}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}},\boldsymbol{D})} \frac{1}{2}\log_2 \frac{\det^{+}(\boldsymbol{\Sigma}_{\boldsymbol{Y}})}{\det^{+}(\boldsymbol{\mathcal{D}})} && (3.34)\\
&\geq \min_{\boldsymbol{\mathcal{D}}\,\in\,\mathscr{D}^{\mathrm{Joint}}_{=}(\boldsymbol{\Sigma}_{\boldsymbol{Y}},\boldsymbol{D})} \frac{1}{2}\log_2 \frac{\det^{+}(\boldsymbol{\Sigma}_{\boldsymbol{Y}})}{\det^{+}(\boldsymbol{\mathcal{D}})} && (3.35)\\
&= \sum_{i\in\mathcal{L}} \frac{1}{2}\log_2 \left[\max(1,\frac{\lambda_i}{w})\right] && (3.36)\\
&= \begin{cases} \sum_{i=W+1}^{L} \frac{1}{2}\log_2 \frac{\lambda_i}{w} & W < L \\ 0 & W = L \end{cases} = \underline{R}^{\mathrm{Joint}}(\boldsymbol{\lambda}),
\end{aligned}$$

where (3.33) is the single-letter rate-distortion function of $\boldsymbol{Y}$ with vector distortion constraint $\boldsymbol{D}$, (3.34) is true because Gaussian distribution maximizes differential entropy for a given covariance matrix, (3.35) is due to the relation (3.21), and in (3.36) we used the fact that reverse water-filling on the eigenvalues of a multivariate Gaussian random vector can achieve its rate-sum-distortion function (see, e.g., [15, p.315]). □

Due to (2.9), (3.24) and (3.25), the sum-rate loss $R^{\Delta}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(\boldsymbol{D})$ for $(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D}) \in \mathscr{S}_L^{\mathrm{BT}}$

is upper-bounded as

$$\begin{aligned}
R^{\Delta}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}) &\leq R^{\mathrm{BT}}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}) - R^{\mathrm{Joint}}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}) \\
&\leq \overline{R}^{\mathrm{BT}}(\boldsymbol{\lambda}) - \underline{R}^{\mathrm{Joint}}(\boldsymbol{\lambda}) \\
&= \frac{1}{2}\sum_{i\in\mathcal{L}} \log_2(1+\lambda_i p) + \frac{1}{2}\sum_{i=W+1}^{L} \log_2 \frac{w}{\lambda_i} \\
&= \begin{cases} \frac{1}{2}\sum_{i\in\mathcal{L}} \log_2(1+\lambda_i p) & W = L \\ \frac{1}{2}\sum_{i\in\mathcal{L}} \log_2(\frac{1}{\lambda_i}+p) & W = 0 \\ \begin{array}{l} \frac{1}{2}\sum_{i=1}^{W} \log_2(1+\lambda_i p) \\ +\frac{1}{2}\sum_{i=W+1}^{L} \log_2(w(\frac{1}{\lambda_i}+p)) \end{array} & 0<W<L \end{cases}
\end{aligned} \tag{3.37}$$

where p and w are the solutions to (3.26) and (3.27), respectively, and W is defined in (3.28), note that we used the fact that $w = 1$ when $W = 0$.

To compute the supremum value of (3.37) over all $\boldsymbol{\lambda}$'s satisfying $\sum_{i\in\mathcal{L}} \lambda_i = \mathrm{tr}(\boldsymbol{\Sigma_Y}) \geq L$ (due to Lemma 12), we need to allow λ_i's to take the value of infinity, or equivalently, we denote $v_i = \frac{1}{\lambda_i}$ for $i > W$ such that $v_i \in [0, \frac{1}{w}]$. We formulate the core optimization problem $\mathbb{P}_0^L$ as follows

$$\begin{aligned}
\mathbb{P}_0^L: \quad \text{Max.} \quad & f_0(\boldsymbol{\xi}, w, p, W) && \triangleq \frac{1}{2}\sum_{i=1}^{W}\log_2(1+\lambda_i p) + \frac{1}{2}\sum_{i=W+1}^{L}\log_2(w(v_i+p)) \\
\text{over} \quad & && \lambda_1, ..., \lambda_W, v_{W+1}, ..., v_L \in \mathbb{R},\ w, p \in \mathbb{R}^+,\ \text{and } W \in \{0, 1, ..., L\}, \\
\text{s.t.} \quad & h_{01}(\boldsymbol{\xi}, w, p, W) && \triangleq \sum_{i=1}^{W}\frac{\lambda_i}{1+\lambda_i p} + \sum_{i=W+1}^{L}\frac{1}{v_i+p} - L \ = \ 0, && (3.38) \\
& h_{02}(\boldsymbol{\xi}, w, p, W) && \triangleq \sum_{i=1}^{W}\lambda_i + (L-W)w - L \ = \ 0, && (3.39) \\
& g_{0i}(\boldsymbol{\xi}, w, p, W) && \triangleq \lambda_i - w \leq 0, \quad i = 1, 2, ..., W, && (3.40) \\
& g_{0i}(\boldsymbol{\xi}, w, p, W) && \triangleq v_i - \frac{1}{w} \leq 0,\ i = W+1, W+2, ..., L, && (3.41) \\
& j_{0i}(\boldsymbol{\xi}, w, p, W) && \triangleq -\lambda_i \leq 0,\ i = 1, 2, ..., W, && (3.42) \\
& j_{0i}(\boldsymbol{\xi}, w, p, W) && \triangleq -v_i \leq 0,\ i = W+1, W+2, ..., L, && (3.43) \\
& k_0(\boldsymbol{\xi}, w, p, W) && \triangleq W - L \ \leq \ 0. && (3.44)
\end{aligned}$$

$\mathbb{P}_1^L$: Max. $f_1(\lambda, w, \eta, p, N, M, K)$ over $\lambda, w, \eta, p \in \mathbb{R}$ and $N, M, K \in \{0\} \cup \boldsymbol{\mathcal{L}}$

$$\text{s.t. } h_{10}(\lambda, w, \eta, p, N, M, K) \triangleq \frac{N\lambda}{1+\lambda p} + \frac{Mw}{1+wp} + \frac{(L-N-M-K)\eta}{1+\eta p} + \frac{K}{p} - L = 0, \tag{3.45}$$

$$h_{11}(\lambda, w, \eta, p, N, M, K) \triangleq w = \frac{L-N\lambda}{L-N} = 0, \tag{3.46}$$

$$0 \leq N \leq L-1,\ 1 \leq N+M \leq L-1,\ K \leq L-N-M, \tag{3.47}$$

and $0 \leq \lambda < w < \eta < \infty$.

where $\boldsymbol{\xi} = (\lambda_1, ..., \lambda_W, v_{W+1}, ..., v_L)^{\mathrm{T}}$, $f_0(\boldsymbol{\xi}, w, p, W)$ is the sum-rate loss $R^{\Delta}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D})$ defined in (3.37), (3.38) and (3.39) are equality constraints, and (3.40) - (3.44) are inequality constraints. Unfortunately, it is very hard to directly solve $\mathbb{P}_0^L$, since it is a nonconvex MINLP problem [42] with integer design variable W.

2. Reduction to the quad-eigen case and relaxation

Instead of directly solving $\mathbb{P}_0^L$, we state the following lemma, which shows that the optimal solution to $\mathbb{P}_0^L$ must be achieved when λ_i's take at most four different values.

Lemma 14. *The optimal solution to $\mathbb{P}_0^L$ must be achieved when the eigenvalues λ_i's satisfy*

1. $\lambda_i = \lambda$, *for all* $i = 1, 2, ..., N$, *with* $0 \leq \lambda < 1$ *and* $0 \leq N \leq L - K$.
2. $\lambda_i = w$, *for* $i = N+1, N+2, ..., N+M$, *where* M *is some non-negative integer such that* $1 \leq N + M = W \leq L - K$.
3. $\lambda_i = \eta \in (w, \infty)$, *for* $i = N+M+1, N+M+2, ..., L-K$, *where* K *is such that* $N + M + K \leq L$.
4. $\lambda_i = \infty$ *for* $i = L - K + 1, ..., L$ *due to assumption.*

In other words, the optimal covariance matrix $\mathbf{\Sigma_Y}$ *that achieves the supremum sum-rate loss in the regular case can only have at most four distinct eigenvalues, taking values from the set* $\{\lambda, w, \eta, \infty\}$ *with* $0 \leq \lambda < w < \eta < \infty$.

Proof. We first show that the Karush-Kuhn-Tucker (KKT) condition is a necessary condition for optimality by proving that for fixed $W \in \{0, 1, ..., L\}$, the equality constraints h_{01}, h_{02} and any possible combinations of active inequality constraints chosen from $\{g_{0i} : i \in \mathcal{L}\} \cup \{j_{0i} : i \in \mathcal{L}\}$ must satisfy the linear independence constraint qualification [43, p. 247], i.e., their gradients are linearly independent at any $(\boldsymbol{\xi}, w, p)$. In fact, if we compute the gradients of the above $2L + 2$ functions with respect to

$(\boldsymbol{\xi}, w, p)$,

$$\begin{aligned}
\nabla_{(\boldsymbol{\xi},w,p)} h_{01} &= \Big(\frac{1}{(1+\lambda_1 p)^2}, \ldots, \frac{1}{(1+\lambda_W p)^2}, \frac{-1}{(v_{W+1}+p)^2}, \ldots, \frac{-1}{(v_L+p)^2}, 0, \\
&\qquad -\sum_{i=1}^{W} \frac{\lambda_i^2}{(1+\lambda_i p)^2} - \sum_{i=W+1}^{L} \frac{1}{(v_i+p)^2}\Big)^{\mathrm{T}}, \\
\nabla_{(\boldsymbol{\xi},w,p)} &= (\underbrace{1, ..., 1}_{W}, \underbrace{0, ..., 0}_{L-W}, L-W, 0)^{\mathrm{T}}, \\
\nabla_{(\boldsymbol{\xi},w,p)} g_{0i} &= \begin{cases} (\underbrace{0, ..., 0}_{i-1}, 1, \underbrace{0, ..., 0}_{L-i}, -1, 0)^{\mathrm{T}} & i \le W \\ (\underbrace{0, ..., 0}_{i-1}, 1, \underbrace{0, ..., 0}_{L-i}, \frac{1}{w^2}, 0)^{\mathrm{T}} & W < i \le L \end{cases}, \\
\nabla_{(\boldsymbol{\xi},w,p)} j_{0i} &= (\underbrace{0, ..., 0}_{i-1}, -1, \underbrace{0, ..., 0}_{L-i}, 0, 0)^{\mathrm{T}}, \ i \in \boldsymbol{\mathcal{L}}.
\end{aligned}$$

Note that g_{0i} and j_{0i} cannot be both active. We observe that among all partial derivatives with respect to p, $\frac{\partial h_{01}}{\partial p}$ is the only non-zero one. Hence we only need to show that the following matrix is non-singular,

$$\begin{bmatrix}
1 & 1 & ... & 1 & 0 & ... & 0 & L-W \\
1 & 0 & ... & 0 & 0 & ... & 0 & b_1 \\
0 & 1 & ... & 0 & 0 & ... & 0 & b_2 \\
... & ... & ... & ... & ... & ... & ... & ... \\
0 & 0 & ... & 1 & 0 & ... & 0 & b_W \\
0 & 0 & ... & 0 & 1 & ... & 0 & b_{W+1} \\
... & ... & ... & ... & ... & ... & ... & ... \\
0 & 0 & ... & 0 & 0 & ... & 1 & b_L
\end{bmatrix},$$

where $b_i \in \{0, -1\}$ for $i \le W$ and $b_i \in \left\{0, \frac{1}{w^2}\right\}$ otherwise, which is obvious since its determinant equals to $(-1)^L \cdot (L - W - \sum_{i=1}^{W} b_i) \neq 0$.

Now we know that the optimal solution to $\mathbb{P}_0^L$ must satisfy the KKT condition for some $W \in \{0, 1, ..., L\}$. Hence we use the Lagrangian of $\mathbb{P}_0^L$

$$\begin{aligned} L(\mathbb{P}_0^L) = & -f_0(\boldsymbol{\xi}, w, p, W) + \alpha \cdot h_{01}(\boldsymbol{\xi}, w, p, W) + \beta \cdot h_{02}(\boldsymbol{\xi}, w, p, W) + \sum_{i=1}^{L} \gamma_i \cdot g_{0i}(\boldsymbol{\xi}, w, p, W) \\ & + \sum_{i=1}^{W} \zeta_i \cdot j_{0i}(\boldsymbol{\xi}, w, p, W), \end{aligned}$$

to compute the KKT condition (for a fixed W) as

$$\begin{aligned} \frac{\partial L(\mathbb{P}_0)}{\partial \lambda_i} &= -\frac{1}{2\ln 2}\frac{p}{1+\lambda_i p} + \frac{\alpha}{(1+\lambda_i p)^2} + \beta + \gamma_i - \zeta_i \;=\; 0, \\ & \quad i = 1, 2, ..., W, \qquad (3.48) \\ \frac{\partial L(\mathbb{P}_0)}{\partial v_i} &= -\frac{1}{2\ln 2}\frac{1}{v_i + p} - \frac{\alpha}{(v_i+p)^2} + \gamma_i - \zeta_i \;=\; 0, \\ & \quad i = W+1, ..., L, \qquad (3.49) \\ \frac{\partial L(\mathbb{P}_0)}{\partial p} &= -\frac{1}{2\ln 2}\left[\sum_{i=1}^{W}\frac{\lambda_i}{1+\lambda_i p} + \sum_{i=W+1}^{L}\frac{1}{v_i+p}\right] \\ & \quad -\alpha\left(\sum_{i=1}^{W}\frac{\lambda_i^2}{(1+\lambda_i p)^2} + \sum_{i=W+1}^{L}\frac{1}{(v_i+p)^2}\right) = 0, \qquad (3.50) \\ \frac{\partial L(\mathbb{P}_0)}{\partial w} &= -\frac{L-W}{2\ln 2}\cdot\frac{1}{w} + (L-W)\beta - \sum_{i=1}^{W}\gamma_i + \sum_{i=W+1}^{L}\frac{\gamma_i}{w^2} \\ &= 0, \qquad (3.51) \end{aligned}$$

and

$$\begin{aligned} h_{01}(\boldsymbol{\xi}, w, p, W) &= 0, \\ h_{02}(\boldsymbol{\xi}, w, p, W) &= 0, \\ \gamma_i \cdot g_{0i}(\boldsymbol{\xi}, p, w, W) &= 0, \quad i = 1, 2, ..., L, \qquad (3.52) \\ \zeta_i \cdot j_{0i}(\boldsymbol{\xi}, p, w, W) &= 0, \quad i = 1, 2, ..., W, \\ \gamma_i &\geq 0, \qquad (3.53) \\ \zeta_i &\geq 0. \end{aligned}$$

Note that we may assume without missing the optimal solution that the equalities in g_{0i} are not achieved for any $i \in \{W+1, ..., L\}$, since otherwise the solution must satisfy the KKT condition for some $W' > i \geq W$. On the other hand, if the equality in g_{0i} holds for some $i \leq W$, i.e., $\lambda_i = 0$, then the optimization problem reduces to $\mathbb{P}_0^{L-1}$ after replacing L by $L-1$ in h_{01} and h_{02}. Hence in the rest of the proof, we assume $\gamma_i = 0$ for all $i \in \{W+1, ..., L\}$ and $\zeta_i = 0$ for all $i \leq W$.

From (3.50) and the fact that $\sum_{i=1}^{W} \frac{\lambda_i}{1+\lambda_i p} + \sum_{i=W+1}^{L} \frac{1}{v_i+p} = L$, we get

$$\alpha = -\frac{L}{2\ln 2\left[\sum_{i=1}^{W} \frac{\lambda_i^2}{(1+\lambda_i p)^2} + \sum_{i=W+1}^{L} \frac{1}{(v_i+p)^2}\right]} < 0. \tag{3.54}$$

On the other hand, (3.51) and (3.53) imply that

$$\beta = \frac{1}{2\ln 2 \cdot w} + \frac{\sum_{i=1}^{W} \gamma_i}{L-W} \ > \ 0. \tag{3.55}$$

Now let $\mathcal{G} \subseteq \{1, 2, ..., W\}$ be the index set such that

$$\begin{cases} \gamma_i = 0 & \text{for all } i \in \mathcal{G} \\ \gamma_i > 0 & \text{for all } i \in \{1, 2, ..., W\} - \mathcal{G} \end{cases}. \tag{3.56}$$

Then for any $i \in \mathcal{G}$, (3.48) and (3.56) tell us that

$$-\frac{1}{2\ln 2}\frac{p}{1+\lambda_i p} + \frac{\alpha}{(1+\lambda_i p)^2} + \beta \ = \ 0. \tag{3.57}$$

Since for $i \in \{1, 2, ..., W\}$, $\lambda_i \leq w < \infty$, we can combine (3.54), (3.55), and (3.57) and write

$$\beta(1+\lambda_i p)^2 - \frac{p}{2\ln 2}(1+\lambda_i p) + \alpha = 0.$$

Assume there are $i, j \in \mathcal{G}$ such that $\lambda_i \neq \lambda_j$. Then λ_i and λ_j are two distinct positive roots of

$$\beta(1+\lambda p)^2 - \frac{p}{2\ln 2}(1+\lambda p) + \alpha = 0. \tag{3.58}$$

However, it is obvious that (3.58) has only one positive root (since $\beta > 0$ and $\alpha\beta < 0$), namely

$$\lambda = \frac{1}{p}\left[\frac{p+\sqrt{p^2-16\alpha\beta\ln^2 2}}{4\beta\ln 2} - 1\right],$$

and we have a contradiction. Hence for any $i, j \in \mathcal{G}$, we must have $\lambda_i = \lambda_j = \lambda \leq w$. Let N be the cardinality of $\mathcal{G}$. It is easy to prove that $\lambda \leq 1$ since otherwise

$$\sum_{i=1}^{L} \min\{\lambda_i, w\} = \sum_{i\in\mathcal{G}\cap\{1,2,...,W\}} \lambda_i + \sum_{i\in\{1,2,...,W\}-\mathcal{G}} \lambda_i + Kw \geq L\lambda > L.$$

Similarly, for any $i \in \{W+1, W+2, ..., L\}$, due to (3.49) and (3.54), it must hold that

$$\frac{1}{(v_i+p)\cdot 2\ln 2} + \frac{\alpha}{(v_i+p)^2} + \zeta_i = 0,$$

which implies that for any $i, j \in \{W+1, W+2, ..., L\}$ such that $v_i, v_j > 0$, we must have $v_i = v_j = -(2\ln 2\cdot\alpha + p)$, i.e., $\lambda_i = \lambda_j = \eta \triangleq -\frac{1}{2\ln 2\cdot\alpha+p}$. On the other hand, if $v_i = 0$, then we must have $\lambda_i = \infty$. We denote K as the number of infinite eigenvalues.

Moreover, due to (3.52) and (3.56), we know that $\lambda_i = w$ for any $i \in \{1, 2, ..., W\}-\mathcal{G}$. Hence the optimal $\boldsymbol{\xi}$ must correspond to a covariance matrix with at most four distinct eigenvalues $\{\lambda, w, \eta, \infty\}$ such that $0 \leq \lambda \leq w < \eta \leq \infty$. In addition, we can also assume without losing generality that $0 \leq \lambda < w < \eta < \infty$. Denote M as the cardinalities of the set $\{1, 2, ..., W\} - \mathcal{G}$. Then we must have $N + M = W$.

Now we show that $N + M \geq 1$. Otherwise assume that the optimal solution satisfies $W = 0$, i.e., $\lambda_i > w$ for all $i \in \mathcal{L}$, and $w = 1$ due to (3.39). Then it must hold that $\lambda_i = \eta > 1$ for all $i \leq L - K$. Then the cost function $f_0(\boldsymbol{\xi}, w, p, W)$ becomes

$$f_3(\eta, p, K) = \frac{L-K}{2}\log_2(\frac{1}{\eta} + p) + \frac{K}{2}\log_2(p),$$

and the constraints are

$$\frac{L-K}{\frac{1}{\eta}+p} + \frac{K}{p} - L = 0, \quad \text{ and } \ 1 - \eta < 0.$$

First consider the case when $K = 0$, then the cost function is $f_3(\eta, p, K) = \frac{L}{2}\log_2(\frac{1}{\eta} + p) = L\log_2(1) = 0$, which means $K = 0$ corresponds to a zero sum-rate loss. Similarly, when $K = L$, the sum-rate loss must also be zero. Then consider the case when $K \in \{1, 2, ..., L-1\}$, we have $\eta = \frac{Lp-K}{Lp(1-p)}$, and the cost function is $f_4(p, K) = \frac{K}{2}\log_2(\frac{p(L-K)}{Lp-K}) + \frac{K}{2}\log_2(p)$. Clearly, for any $K \in \{1, 2, ..., L-1\}$, $f_4(p, K)$ is a monotone decreasing function of p in the range $(\frac{K}{L}, 1)$, since

$$\frac{\partial f_4(p, K)}{\partial p} \;=\; -\frac{(1-p)LK}{p(Lp-K) \cdot 2\ln 2} \;<\; 0,$$

where the last inequality is due to the fact that $\eta = \frac{Lp-K}{Lp(1-p)} > 0$. Now since η is a monotone increasing function of p, we know that for any $K \in \{1, 2, ..., L-1\}$, $f_4(p, K)$ is maximized as $\eta = \frac{Lp-K}{Lp(1-p)} \to 1$, i.e., $p \to \sqrt{\frac{K}{L}}$. This means another solution with $N^* = 0$, $M^* = L - K$, $K^* = K$, $w^* = 1$, and $p^* = \sqrt{\frac{K}{L}}$ must achieve a larger cost function value, which contradicts with the assumption. Hence it must hold that $N + M \geq 1$.

Finally, we show that $N + M = W \leq L - 1$. Otherwise we must have $W = L$, which means $\lambda_i \leq w$ for all $i \in \mathcal{L}$, and $\sum_{i=1}^{L}\lambda_i = L$ (due to (3.39)). Then (3.38) is true if and only if $p = 0$, which implies that the cost function $f_0(\boldsymbol{\xi}, w, p, W) = \frac{1}{2}\sum_{i=1}^{L}\log_2(1 + \lambda_i \cdot 0) = 0$. Therefore, the optimal solution to $\mathbb{P}_0^L$ cannot be such that $W = L$, since the supremum sum-rate loss is obviously larger than zero. □

Due to Lemma 14, we can define

$$f_1(\lambda, w, \eta, p, N, M, K) \triangleq \frac{N}{2}\log_2(1+\lambda p) + \frac{M}{2}\log_2(1+wp) \\ + \frac{L-N-M-K}{2}\log_2\left(w\left(\frac{1}{\eta}+p\right)\right) + \frac{K}{2}\log_2(wp),$$

and restate $\mathbb{P}_0^L$ as $\mathbb{P}_1^L$ defined at the bottom of the page.

Although $\mathbb{P}_1^L$ is still a nonconvex MINLP optimization problem due to the discreteness of (N, M, K), one can always exhaust all (N, M, K) triples satisfying (3.47) and find the maximum function value $f_1(\lambda, w, \eta, p, N, M, K)$ for each triple under the constraints (3.45) and (3.46). The sub-problem of $\mathbb{P}_1^L$ corresponding to a fixed (N, M, K)-triple is denoted as $\mathbb{P}_1^L(N, M, K)$. However, as L goes to infinity, the complexity of the above method becomes intractable.

Our approach of solving $\mathbb{P}_1^L$ is to define a set of auxiliary continuous optimization problems $\mathbb{P}_2(x)$ parameterized by $x \in [0, 1)$, such that for each fixed $N \in \{0, 1, ..., L-1\}$, the maximum over the solutions to all $\mathbb{P}_1^L(N, M, K)$ problems (with M and K vary) must be upper-bounded by that to $\mathbb{P}_2(\frac{N}{L})$, with equality holds when the later is achieved in the bi-eigen case (corresponding to $M = 0$ and $K = L - N$ in $\mathbb{P}_1^L$).

First, we eliminate K by upper-bounding $f_1(\lambda, w, \eta, p, N, M, K)$. Let t be the solution to

$$\frac{\eta}{1+\eta p} = t \cdot \frac{w}{1+wp} + (1-t) \cdot \frac{1}{p}, \tag{3.59}$$

then we must have $t \in [0, 1)$. Since the function $-\log_2(\cdot)$ is convex, we have

$$\begin{aligned}
\log_2\left(w\left(\frac{1}{\eta}+p\right)\right) &= -\log_2\left(\frac{1}{w}\cdot\frac{\eta}{1+\eta p}\right) \\
&= -\log_2\left[\frac{1}{w}\cdot\left(t\cdot\frac{w}{1+wp}+(1-t)\cdot\frac{1}{p}\right)\right] \\
&= -\log_2\left[t\cdot\frac{1}{1+wp}+(1-t)\cdot\frac{1}{wp}\right] \\
& {}_2(1+wp)+(1-t)\log\log_2(wp).
\end{aligned}$$

Thus if we define

$$M' = M + (L - M - N - K)t, \tag{3.60}$$

which is a real number between 0 and $L - N$, the constraint (3.45) becomes

$$\frac{N\lambda}{1+\lambda p} + \frac{M'w}{1+wp} + \frac{(L-N-M')}{p} = L, \tag{3.61}$$

$$\tag{3.62}$$

and the objective function $f_1(\lambda, w, \eta, p, N, M, K)$ can be upper-bounded by

$$f_1(\lambda, w, \eta, p, N, M, K) \le \frac{N}{2}\log_2(1+\lambda p) + \frac{M'}{2}\log_2(1+wp) + \frac{L-N-M'}{2}\log_2(wp) \tag{3.63}$$

$$\triangleq \overline{f}_1(\lambda, w, p, N, M'). \tag{3.64}$$

Clearly, (3.63) holds with equality if $M = L - N - K = 0$, or equivalently, $M' = 0$.

Next, we relax N to be a real number in $[0, L)$ (since Lemma 14 proves that the optimal solution to $\mathbb{P}_0$ cannot be achieved when $N = L$), and denote $x = \frac{N}{L}$, and

$y = \frac{M'}{L}$. For a fixed $x \in [0,1)$, we eliminate w and y in $\overline{f}_1(\lambda, w, p, N, M')$ using

$$\begin{aligned} w &= \frac{L - N\lambda}{L - N} = \frac{1 - x\lambda}{1 - x} && (3.65) \\ y &= \frac{(L - N - Lp + Lp\lambda - Lp^2\lambda)(L - N + Lp - Lpn\lambda)}{L(1 + p\lambda)(L - N)} \\ &= \frac{(1 - x - p + p\lambda - p^2\lambda)(1 - x + p - px\lambda)}{(1 + p\lambda)(1 - x)}, && (3.66) \end{aligned}$$

and obtain the *relaxed* optimization problem $\mathbb{P}_2(x)$ for a fixed $x \in [0,1)$ (note that $p \leq 1$ due to (3.38) and (3.41)) as follows,

$$\begin{aligned} \mathbb{P}_2(x): \quad \text{Max.} \quad & f_2(\lambda, p; x) \text{ over } 0 \leq \lambda < 1, 0 < p \leq 1, \\ \text{s.t.} \quad & g_2(\lambda, p; x) \overset{\Delta}{=} 1 - x - p + p\lambda - p^2\lambda \geq 0, \qquad (3.67) \end{aligned}$$

where

$$\begin{aligned} f_2(\lambda, p; x) &\overset{\Delta}{=} \overline{f}_1(\lambda, w, p, Lx, Ly) \\ &= \frac{x}{2}\log_2(1 + \lambda p) + \frac{y}{2}\log_2(1 + wp) + \frac{1 - x - y}{2}\log_2(wp) && (3.68) \\ &= \frac{x}{2}\log_2(1 + \lambda p) + \frac{1 - x}{2}\log_2\left(\frac{1 - x\lambda}{1 - x}p\right) \\ &\quad + \frac{(1 - x - p + p\lambda - p^2\lambda)(1 - x + p - px\lambda)}{2(1 + p\lambda)(1 - x)} \cdot \log_2\left(1 + \frac{1 - x}{p(1 - x\lambda)}\right), \end{aligned}$$

and the constraint (3.67) is equivalent to $y \geq 0$ with y given in (3.66). From (3.61) - (3.66), it is clear that

$$f_1(\lambda, w, \eta, p, N, M, K) \leq \overline{f}_1(\lambda, w, p, N, M') = L \cdot f_2(\lambda, p; x) \qquad (3.69)$$

for any (w, η, M, K) satisfying (3.45) - (3.47), where $x = \frac{N}{L}$, t is the solution to (3.59), and M' is defined in (3.60). Moreover, (3.69) holds with equality if

$$\begin{aligned} & M = L - N - K = 0 \quad \Leftrightarrow \quad M' = 0 \quad \Leftrightarrow \quad y = 0 \\ & \quad \Leftrightarrow \quad g_2(\lambda, p\,; x) = 1 - x - p + p\lambda - p^2\lambda = 0. \end{aligned} \tag{3.70}$$

Denote the solutions to $\mathbb{P}_0^L$, $\mathbb{P}_1^L$, and $\mathbb{P}_2(x)$ as

$$\begin{aligned} \mathrm{sol}(\mathbb{P}_0^L) &= \left\{f^{\max}; \boldsymbol{\xi}^{\max}, w^{\max}, p^{\max}, W^{\max}\right\}, \\ \mathrm{sol}(\mathbb{P}_1^L) &= \left\{f^{\max_1}; \lambda^{\max_1}, \eta^{\max_1}, p^{\max_1}, w^{\max_1}, N^{\max_1}, M^{\max_1}, K^{\max_1}\right\}, \\ \mathrm{sol}(\mathbb{P}_2(x)) &= \left\{f^{\max_2}(x); \lambda^{\max_2}(x), p^{\max_2}(x)\right\}, \end{aligned}$$

respectively. Then the relationship among $\mathrm{sol}(\mathbb{P}_0^L)$, $\mathrm{sol}(\mathbb{P}_1^L)$, and $\mathrm{sol}(\mathbb{P}_2(x))$ is given in the following lemma.

Lemma 15. *It holds for any $L \geq 2$ that*

$$f^{\max} \;=\; f^{\max_1} \;\leq\; L \cdot \max_{N \in \{0,1,\ldots,L-1\}} f^{\max_2}\left(\frac{N}{L}\right), \tag{3.71}$$

with equality holds if the largest $f^{\max_2}\left(\frac{N}{L}\right)$ over all $N \in \{0, 1, ..., L-1\}$ is achieved on the boundary of (3.67), i.e.,

$$g_2\left(\lambda^{\max_2}\left(\frac{N^{\max_2}}{L}\right), p^{\max_2}\left(\frac{N^{\max_2}}{L}\right), \frac{N^{\max_2}}{L}\right) = 0, \tag{3.72}$$

where $N^{\max_2} = \arg\max_{N \in \{0,1,\ldots,L-1\}} f^{\max_2}\left(\frac{N}{L}\right)$. Moreover, if (3.72) holds, we must *have*

$$W^{\max} = N^{\max_1} = N^{\max_2}, \quad M^{\max_1} = 0, \tag{3.73}$$

$$K^{\max_1} = L - N^{\max_2}, \tag{3.74}$$

$$\lambda_i^{\max} = \lambda^{\max_1} = \lambda^{\max_2}\left(\frac{N^{\max_2}}{L}\right) = \nu\left(\frac{N^{\max_2}}{L}\right), i \leq N^{\max_2}, \tag{3.75}$$

$$v_i^{\max} = 0, \qquad i > N^{\max_2}, \tag{3.76}$$

$$p^{\max} = p^{\max_1} = p^{\max_2}\left(\frac{N^{\max_2}}{L}\right), \tag{3.77}$$

$$w^{\max} = w^{\max_1} = \frac{L - N^{\max_2}\lambda^{\max_2}(\frac{N^{\max_2}}{L})}{L - N^{\max_2}}, \tag{3.78}$$

i.e., the optimal solution for $\mathbb{P}_0^L$ corresponds to a covariance matrix with two distinct values: $\lambda^{\max_2}(\frac{N^{\max_2}}{L})$ and ∞.

Proof. First, due to Lemma 14, sol($\mathbb{P}_1$) is equivalent to sol($\mathbb{P}_0$) in the sense that

$$f^{\max} \quad = \quad f^{\max_1}, \tag{3.79}$$

and

$$\begin{aligned}
\lambda_i^{\max} &= \lambda^{\max_1}, i = 1, 2, ..., N^{\max_1} \\
\lambda_i^{\max} &= w^{\max_1}, i = N^{\max_1} + 1, ..., N^{\max_1} + M^{\max_1}, \\
\lambda_i^{\max} &= \eta^{\max_1}, i = N^{\max_1} + M^{\max_1}, \ldots, N^{\max_1} + M^{\max_1} + K^{\max_1}, \\
\lambda_i^{\max} &= \infty, i = N^{\max_1} + M^{\max_1} + K^{\max_1} + 1, ..., L, \\
p^{\max} &= p^{\max_1}, \\
w^{\max} &= w^{\max_1}, \\
W^{\max} &= N^{\max_1} + M^{\max_1}.
\end{aligned}$$

Then (3.71) follows directly from (3.79), (3.69) - (3.70), and the equivalence between the constraints (3.45) and (3.61).

In addition, if (3.72) holds, due to (3.70), we know that

$$f_1(\lambda, w, \eta, p, N, M, K) \;=\; L \cdot f_2\left(\lambda, p; \frac{N}{L}\right),$$

where $\lambda = \lambda^{\max_2}\left(\frac{N^{\max_2}}{L}\right)$, $w = \frac{L - N^{\max_2}\lambda^{\max_2}\left(\frac{N^{\max_2}}{L}\right)}{L - N^{\max_2}}$, $p = p^{\max_2}\left(\frac{N^{\max_2}}{L}\right)$, $N = N^{\max_2}$, $M = 0$, $K = L - N^{\max_2}$, and η be any real number larger than w. This means that the $\max_{N \in \{0,1,\ldots,L-1\}} f^{\max_2}\left(\frac{N}{L}\right)$, which is an upper bound on $f^{\max_1}$, is also achievable in $\mathbb{P}_1$. Hence (3.71) holds with equality if (3.72) is true, and (3.73) - (3.78) are trivial consequences. □

Moreover, if the solution to $\mathbb{P}_2(x)$ is achieved on the boundary $g_2(\lambda, p; x) = 0$, then it must be the solution to

$$\begin{aligned} \mathbb{P}_{2b}(x) : \text{Max.} \;\; & f_{2b}(\lambda, p; x) \text{ over } 0 \leq \lambda < 1, 0 < p \leq 1, \\ \text{s.t.} \quad & g_2(\lambda, p; x) \;=\; 0, \end{aligned}$$

where

$$f_{2b}(\lambda, p; x) \triangleq \frac{x}{2}\log_2(1 + \lambda p) + \frac{1-x}{2}\log_2\left(\frac{1 - x\lambda}{1-x}p\right).$$

The following lemma gives the exact form of the solution to $\mathbb{P}_{2b}(x)$ for any $x \in [0, 1)$.

Lemma 16. *The maximum function value for* $\mathbb{P}_{2b}(x)$ *is*

$$f^{max_{2b}} \;=\; \tau(x), \tag{3.80}$$

which is achieved when

$$\lambda^{max_{2b}} \;=\; \nu(x), \tag{3.81}$$

$$p^{max_{2b}} \;=\; \mu(x) \;\triangleq\; 1 - \frac{2x}{(1+x) + \sqrt{(1-x)(5-x)}}. \tag{3.82}$$

Proof. First, when $x = 0$, $f_{2b}(\lambda, p; x) \equiv 0$, and (3.80) - (3.82) hold since $\tau(0) = 0$,

$\nu(0) = 0$, and $\mu(0) = 1$.

Then consider the case when $x \in (0, 1)$.

1. If $\lambda = 0$, then $p = 1 - x$ due to (3.82), and the objective function becomes $f_{2b}(0, 1 - x; x) = 0$, thus the maximum cannot be achieved when $\lambda = 0$.
2. If $p = 1$, then $x = 0$ must hold due to (3.82), contradicts with the assumption that $x \in (0, 1)$.
3. Hence the optimal function value must be achieved when $\lambda, p \in (0, 1)$, which implies that (3.82) is the only active constraint, whose gradient is $(p(1-p), \lambda - 2\lambda p - 1)^{\mathrm{T}} \neq \mathbf{0}$. Therefore, the linear independence constraint qualification [43, p. 247] must be satisfied at the optimal point, and the KKT condition is necessary for global optimality.

The Lagrangian is

$$L(\mathbb{P}_{2b}(x)) \quad = \quad -f_{2b}(\lambda, p; x) + \alpha \cdot g_2(\lambda, p; x),$$

and the KKT condition is

$$\begin{aligned}
\frac{\partial L(\mathbb{P}_{2b}(x))}{\partial \lambda} &= \frac{1}{2\ln 2}\frac{x(1-x-p+\lambda p)}{(1+\lambda p)(1-\lambda x)} + \alpha p(1-p) = 0,\\
\frac{\partial L(\mathbb{P}_{2b}(x))}{\partial p} &= -\frac{1}{2\ln 2}\frac{1-x+\lambda p}{p(1+\lambda p)} - \alpha(1-\lambda+2\lambda p) = 0,\\
g_2(\lambda, p; x) &= 1 - x - p + p\lambda - p^2\lambda = 0,
\end{aligned}$$

which leads to two and only two sets of solutions, namely (the corresponding α^+ and

α^- are omitted),

$$\begin{aligned} p^{max_{2b}} = \mu(x), \quad \lambda^{max_{2b}} &= \nu(x) \\ p^- = 1 - \frac{2x}{(1+x) - \sqrt{(1-x)(5-x)}}&, \\ \lambda^- = \frac{-x^2+4x-1-(1-x)\sqrt{(1-x)(5-x)}}{2}&. \end{aligned}$$

One can verify that the first set of solution satisfies the KKT condition, while the second set $\lambda = \lambda^-$ and $p = p^-$ is not feasible since

$$p^-\lambda^- = -\frac{3(1-x)+\sqrt{(1-x)(5-x)}}{1+x+\sqrt{(1-x)(5-x)}} < 0.$$

Hence the maximum function value $f_{2b}(\lambda, p; x)$ is achieved at $(\lambda^{\max_{2b}}, p^{\max_{2b}})$, with a maximum function value of

$$f_{2b}(\lambda^{\max_{2b}}, p^{\max_{2b}}; x) = \tau(x).$$

Lemma 16 is proved. □

Remark 4: The original problem $\mathbb{P}_0^L$ involves optimization over L eigenvalues $\{\lambda_1, \lambda_2, ..., \lambda_L\}$ (which are allowed to be infinity). In Lemma 14, it is shown that the optimal solution can only take four different eigenvalues $(\lambda, w, \eta, \infty)$, leading to an equivalent optimization problem $\mathbb{P}_1^L$. Now Lemma 15 further proves that for any integer $L \geq 2$, if the maximum function value $f^{\max_2}(x)$ of all auxiliary optimization problems $\mathbb{P}_2(x)$ with $x \in \left\{0, \frac{1}{L}, ..., \frac{L-1}{L}\right\}$ is achieved on the boundary of (3.67), then $M^{\max_1} = L - N^{\max_1} - M^{\max_1} - K^{\max_1} = 0$, i.e., the maximum function value $f^{\max}$ in the original problem, which is an upper bound on the L-terminal supremum sum-rate loss (normalized by L) over $(\boldsymbol{\Sigma_Y}, \boldsymbol{D}) \in \mathscr{R}_L$, must be achieved when the source covariance matrix $\boldsymbol{\Sigma_Y}$ is bi-eigen, with eigenvalues $\lambda^{\max_1}$ and ∞ repeated $N^{\max_1}$ and $L - N^{\max_1}$ times, respectively. Moreover, if (3.72) holds, Lemma 16 gives the exact form of $f^{\max}$ as

$$f^{\max} = L \cdot \max_{N \in \{0,1,\ldots,L-1\}} \tau\left(\frac{N}{L}\right).$$

3. Solution via rigorous numerical methods

Due to Lemmas 15 and 16, $\mathbb{P}_0^L$ would be solved if (3.72) holds for all integer $L \geq 2$. Unfortunately, it can be easily verified numerically that (3.72) is not true for $L \in \{2,3,4,8\}$. To see this, we plot in Fig. 8 the numerically computed maximum function value of $f_2(\lambda, p\,; x)$ over all $p \in (0,1]$ for fixed x and λ subject to the constraint $g_2(\lambda, p; x) \geq 0$, i.e.,

$$f^*(x,\lambda) \quad \triangleq \quad \max_{p \in (0,1] : g_2(\lambda,p;x) \geq 0} f_2(\lambda, p; x).$$

For comparison, we also plot the maximum of $f_2(\lambda, p; x)$ when the constraint is forced to be satisfied with equality, i.e.,

$$f^b(x,\lambda) \triangleq \max_{p \in (0,1] : g_2(\lambda,p;x) = 0} f_2(\lambda, p; x).$$

An obvious fact is that $f^*(x,\lambda) \geq f^b(x,\lambda)$ for any (x,λ) pair. Then the maximum function value in sol($\mathbb{P}_2(\frac{N^{\max_2}}{L})$), i.e.,

$$f^{\max_2}\left(\frac{N^{\max_2}}{L}\right) = \max_{N \in \{0,1,\ldots,L\}} \left[\max_{\lambda \in [0,1)} f^*\left(\frac{N}{L}, \lambda\right)\right]$$

is numerically computed as ($N^{\max_2}$'s are numerically found)

$$f^{\max_2}\left(\frac{1}{2}\right) = 0.1015,\ f^{\max_2}\left(\frac{2}{3}\right) = 0.1043,\ f^{\max_2}\left(\frac{3}{4}\right) = 0.1066,\ f^{\max_2}\left(\frac{6}{8}\right) = 0.1066$$

for $L = 2, 3, 4, 8$, respectively. On the other hand, the maximum value on the boundary defined as

$$f^{\mathrm{max_b}}\left(\frac{N^{\mathrm{max_b}}}{L}\right) = \max_{N\in\{0,1,\ldots,L\}}\left[\max_{\lambda\in[0,1)} f^{b}\left(\frac{N}{L},\lambda\right)\right]$$

with $N^{\mathrm{max_b}} = \arg\max_{N\in\{0,1,\ldots,L\}}\left[\max_{\lambda\in[0,1)} f^{b}(\frac{N}{L},\lambda)\right]$ can be computed for $L = 2, 3, 4, 8$ respectively as

$$f^{\mathrm{max_b}}\left(\frac{1}{2}\right) = 0.0805,\ f^{\mathrm{max_b}}\left(\frac{2}{3}\right) = 0.1001,\ f^{\mathrm{max_b}}\left(\frac{3}{4}\right) = 0.1064,\ f^{\mathrm{max_b}}\left(\frac{6}{8}\right) = 0.1064.$$

We observe that $f^{\mathrm{max_b}}(\frac{N^{\mathrm{max_b}}}{L}) < f^{\mathrm{max_2}}(\frac{N^{\mathrm{max_2}}}{L})$ for $L = 2, 3, 4, 8$, which means $\mathrm{sol}(\mathbb{P}_2(\frac{1}{2}))$, $\mathrm{sol}(\mathbb{P}_2(\frac{2}{3}))$, $\mathrm{sol}(\mathbb{P}_2(\frac{3}{4}))$, and $\mathrm{sol}(\mathbb{P}_2(\frac{6}{8}))$ for $L = 2, 3, 4, 8$ are not achieved on the boundary $g_2(\lambda, p; x) = 0$. These numerical results are plotted in Fig. 8, with the optimal λ's given by

$$\begin{aligned}
\lambda^{\mathrm{max_2}}\left(\frac{N^{\mathrm{max_2}}}{L}\right) &\triangleq \arg\max_{\lambda\in[0,1)} f^{*}\left(\frac{N^{\mathrm{max_2}}}{L},\lambda\right) = 0.8328, 0.8453, 0.8548,\\
\lambda^{\mathrm{max_b}}\left(\frac{N^{\mathrm{max_2}}}{L}\right) &\triangleq \arg\max_{\lambda\in[0,1)} f^{*}\left(\frac{N^{\mathrm{max_b}}}{L},\lambda\right) = 0.7500, 0.8114, 0.8475,
\end{aligned}$$

for $L = 2, 3, 4$, respectively (the results for $L = 8$ are not plotted in Fig. 8 since $\mathrm{sol}(\mathbb{P}_2(\frac{6}{8}))$ is exactly the same as $\mathrm{sol}(\mathbb{P}_2(\frac{3}{4}))$).

Therefore, we separately treat the cases when $L = 5, 6, 7$ and $L \geq 9$, for which (3.72) can be proved; and the cases when $L = 2, 3, 4, 8$, for which $\mathbb{P}_1^L$ is directly solved. Correspondingly, we have the following two lemmas.

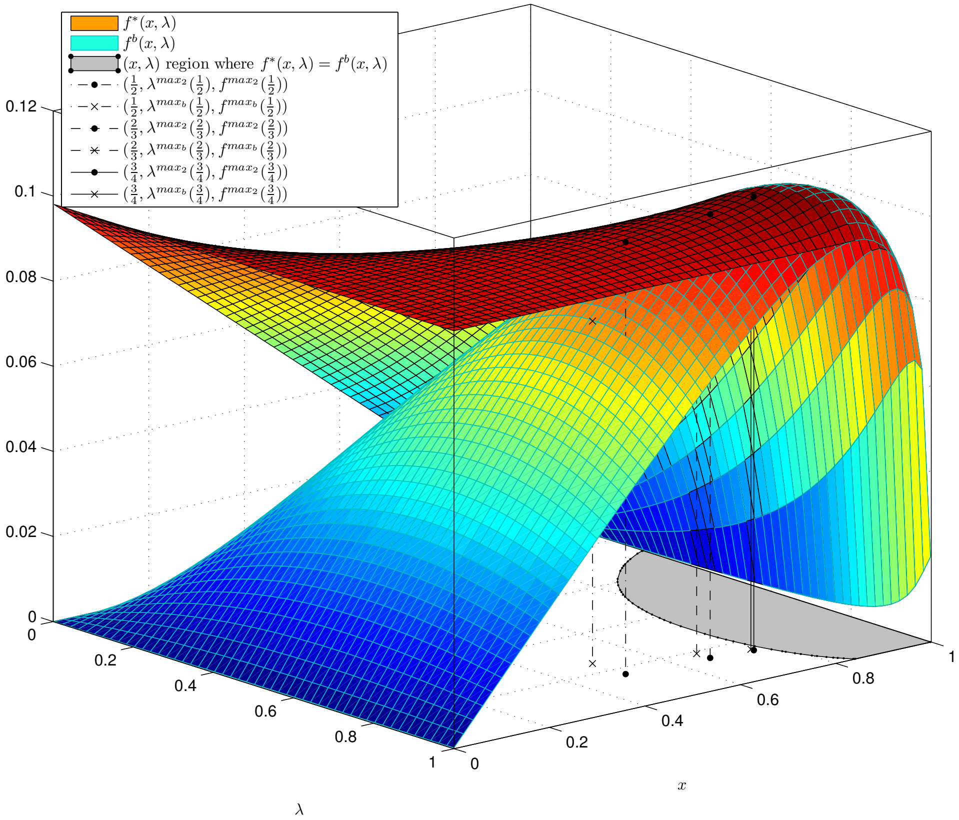


Fig. 8. Numerical comparisons between $f^*(x,\lambda)$ & $f^b(x,\lambda)$ and $f^{\max_2}(\frac{N^{\max_2}}{L})$ & $f^{\max_b}(\frac{N^{\max_b}}{L})$ for $L = 2,3,4$. The shaded region in the (x,λ) plane corresponds to all points satisfying $f^*(x,\lambda) = f^b(x,\lambda)$. If $f^{\max_2}(\frac{N^{\max_2}}{L})$ is achieved in this region, so will be $f^{\max_b}(\frac{N^{\max_b}}{L})$.

Lemma 17. *If $L \notin \{2,3,4,8\}$, then (3.72) holds, and the solution to $\mathbb{P}_0^L$ is given by*

$$\mathrm{sol}(\mathbb{P}_0^L) = \begin{cases} f^{\max} = L \cdot \tau(\frac{N^\star}{L}), \\ \boldsymbol{\xi}^{\max} = \left(\underbrace{\nu\left(\frac{N^\star}{L}\right), \ldots, \nu\left(\frac{N^\star}{L}\right)}_{N^\star}, \underbrace{0, ..., 0}_{L-N^\star} \right)^{\mathrm{T}} \\ w^{\max} = \frac{L-N^\star\nu(\frac{N^\star}{L})}{L-N}, \\ p^{\max} = \mu(\frac{N^\star}{L}), \\ W^{\max} = N^\star. \end{cases}, \tag{3.83}$$

with $N^\star$ defined in (3.6).

Proof. See Appendix A. □

Lemma 18. *If $L \in \{2, 3, 4, 8\}$, then the solution to $\mathbb{P}_1^L$ is achieved in the bi-eigen case, i.e.,*

$$M^{\max_1} \;=\; 0, \quad K^{\max_1} \;=\; L - N^{\max_1}. \tag{3.84}$$

and the solution to $\mathbb{P}_0^L$ is also given by (3.83).

Proof. We compute the lower bound (A.25) of $f^{\max_1}(N)$ (which is also the maximum function value on the boundary $g_2(\lambda^{\max_2}(x), p^{\max_2}(x); x) = 0$) for each $N \in \{0, 1, 2, ..., L-1\}$, and an upper bound of $f^{\max_1}(N)$ under the non-boundary assumption that $g_2(\lambda^{\max_2}(x), p^{\max_2}(x); x) > 0$

$$f^{\max_1}(N) \leq f^{\max_2}\left(\frac{N}{L}\right) \leq \overline{f}^{g>0}\left(\frac{N}{L}\right) \tag{3.85}$$

where the first and second inequalities are true due to Lemma 15 and (A.22), respectively. We observe in Table I that for pairs $(N, L) = (2, 3), (3, 4), (6, 8)$, the lower bound (A.25) of $f^{\max_1}(N)$ is larger than the lower and upper bounds of $f^{\max_1}(N)$ for all other N values. Hence we must have $N^{\max_1} = 2, 3, 6$ for $L = 3, 4, 8$, respectively.

Now we solve $\mathbb{P}_1^L$ separately for $L = 2, 3, 4, 8$. First consider $L = 2$, for which (3.84) must hold since there are exactly two eigenvalues and the trivial case with two equal eigenvalues leads to independent sources, and thus zero sum-rate loss.

When $L = 3$, we already know that $N^{\max_1} = 2$. Then the three optimal eigenvalues can be either (λ, λ, η) or $(\lambda, \lambda, \infty)$ (since $N^{\max_1} + M^{\max_1} \leq L - 1$), both of which correspond to the bi-eigen case, hence (3.84) must hold for $L = 3$.

Similarly, when $L = 4$, since $N^{\max_1} = 3$, the four optimal eigenvalues can be either $(\lambda, \lambda, \lambda, \eta)$ or $(\lambda, \lambda, \lambda, \infty)$, both of which correspond to the bi-eigen case, hence

Table I. Lower bounds of $f^{\max_1}(N)$ on the boundary of (3.72) and upper bounds of $f^{\max_1}(N)$ over non-boundary points for $L \in \{2, 3, 4, 8\}$ and $N \in \{0, 1, ..., L-1\}$.

L	N	lower bound (A.24)	upper bound (3.85)	$\lambda^*(L, N)$
2	0	0.0000000000	0.0981455396	0.9999975000
	1	**0.0804820237**	**0.1015973757**	0.8329675000
3	0	0.0000000000	0.0981455396	0.9999975000
	1	0.0560016357	0.0999813399	0.8250425000
	2	**0.1000689444**	**0.1043755056**	0.8454025000
4	0	0.0000000000	0.0981455396	0.9999975000
	1	0.0426767359	0.0993992764	0.8220575000
	2	0.0804820237	0.1015973757	0.8329675000
	3	**0.1064002237**	**0.1067163310**	0.8549075000
8	0	0.0000000000	0.0981455396	0.9999975000
	1	0.0217593854	0.0987030783	0.8184625000
	2	0.0426767359	0.0993992764	0.8220575000
	3	0.0624294813	0.1003194211	0.8267475000
	4	0.0804820237	0.1015973757	0.8329675000
	5	0.0958492158	0.1035061119	0.8416575000
	6	**0.1064002237**	**0.1067163310**	0.8549075000
	7	0.1058563749	0.0377093873	0.3678625000

(3.84) must hold for $L = 4$.

When $L = 8$, since $N^{\max_1} = 6$, the optimal eigenvalues can be of the forms

$$(\lambda, \lambda, \lambda, \lambda, \lambda, \lambda, \infty, \infty), (\lambda, \lambda, \lambda, \lambda, \lambda, \lambda, \eta, \eta), (\lambda, \lambda, \lambda, \lambda, \lambda, \lambda, w, \infty),$$
$$(\lambda, \lambda, \lambda, \lambda, \lambda, \lambda, \eta, \infty), (\lambda, \lambda, \lambda, \lambda, \lambda, \lambda, w, \eta), \tag{3.86}$$

while the other possible form $(\lambda, \lambda, \lambda, \lambda, \lambda, \lambda, w, w)$ cannot achieve the maximum function value since $N^{\max_1} + M^{\max_1} \leq L - 1$. To prove (3.84), we only need to show that the maximum function value $f^{\max_1}$ must not be achieved by the eigenvalues taking the last three forms in (3.86).

- The first case $(\lambda,\lambda,\lambda,\lambda,\lambda,\lambda,\infty,\infty)$ can be absorbed into the second case $(\lambda,\lambda,\lambda,\lambda,\lambda,\lambda,\eta,\eta)$ by relaxing η to take the value of infinity. Denote $\theta=\frac{1}{\eta}$, thus $\theta\geq 0$, with $\theta=0$ corresponding to the first case. Then $f^{\max_1}$ must be the solution to the optimization problem of maximizing $f_1(\lambda,4-3\lambda,\frac{1}{\theta},p,6,0,0)=3\log_2(1+\lambda p)+\log_2((4-3\lambda)(\theta+p))$ over $\lambda,\theta,p\in\mathbb{R}$ while subjecting to $h_{10}(\lambda,4-3\lambda,\frac{1}{\theta},p,6,0,0)=\frac{6\lambda}{1+\lambda p}+\frac{2}{\theta+p}-8=0$, $0\leq\lambda<1$, $0<p\leq 1$, and $\theta\geq 0$. It is easy to show that the maximum function value of $f^{\max_1}=0.8512017896$ is achieved when

$$\begin{cases} \lambda^{\max_1} &= \frac{23+\sqrt{17}}{32}=0.8475970508,\\ p^{\max_1} &= \frac{3\sqrt{17}-5}{16}=0.4605823048, \quad ,\\ \theta &= 0 \end{cases} \tag{3.87}$$

which corresponds to the first case $(\lambda,\lambda,\lambda,\lambda,\lambda,\lambda,\infty,\infty)$.

- Similarly, in the third case $(\lambda,\lambda,\lambda,\lambda,\lambda,\lambda,\eta,\infty)$, we relax η to take the value of infinity, and denote $\theta=\frac{1}{\eta}$. By solving the optimization problem of maximizing $f_1(\lambda,4-3\lambda,\frac{1}{\theta},p,6,0,1)=3\log_2(1+\lambda p)+\frac{1}{2}\log_2((4-3\lambda)(\theta+p))+\frac{1}{2}\log_2((4-3\lambda)p)$ over $\lambda,\theta,p\in\mathbb{R}$ while subjecting to $h_{10}(\lambda,4-3\lambda,\frac{1}{\theta},p,6,0,1)=\frac{6\lambda}{1+\lambda p}+\frac{1}{\theta+p}+\frac{1}{p}-8=0$, $0\leq\lambda<1$, $0<p\leq 1$, and $\theta\geq 0$, we obtain the same solution given by (3.87). This means that the supremum function value over the third case is strictly smaller than that in the first case where $\theta=0$. Hence the optimal eigenvalues cannot be of the form $(\lambda,\lambda,\lambda,\lambda,\lambda,\lambda,\eta,\infty)$.

- In the fourth and fifth cases $(\lambda,\lambda,\lambda,\lambda,\lambda,\lambda,w,\infty)$ and $(\lambda,\lambda,\lambda,\lambda,\lambda,\lambda,w,\eta)$, denote $\theta=\frac{1}{\eta}$, then the third case corresponds to $\theta=0$. $f^{\max_1}$ must be the solution to the optimization problem of maximizing $f_1(\lambda,4-3\lambda,\frac{1}{\theta},p,6,1,1=3\log_2(1+\lambda p)+\frac{1}{2}\log_2(1+(4-3\lambda)p)+\frac{1}{2}\log_2((4-3\lambda)(\theta+p))$, over $\lambda,\theta,p\in\mathbb{R}$ while subjecting to $h_{10}(\lambda,4-3\lambda,\frac{1}{\theta},p,6,1,1)=\frac{6\lambda}{1+\lambda p}+\frac{4-3\lambda}{1+(4-3\lambda)p}+\frac{1}{\theta+p}-8=0$, $0\leq\lambda<1$, $0<p\leq 1$, and $\theta\geq 0$. Note that $w=4-3\lambda$ due to (3.46). The solution is easily found

to be $\lambda = 0.8845122817$, $p = 0.3355552814$, and $\theta = 0$, then the corresponding supremum function value of $f = 0.8207035176 < 0.8512017896$ over both cases is achieved in the fourth case $(\lambda, \lambda, \lambda, \lambda, \lambda, \lambda, w, \infty)$. Hence the optimal eigenvalues cannot be of the form $(\lambda, \lambda, \lambda, \lambda, \lambda, \lambda, w, \infty)$ or $(\lambda, \lambda, \lambda, \lambda, \lambda, \lambda, w, \eta)$. Therefore, we conclude that for $L = 8$, the maximum function value of $f^{\max_1} = 0.8512017896$ is achieved by eigenvalues $(\lambda^{\max_1}, \lambda^{\max_1}, \lambda^{\max_1}, \lambda^{\max_1}, \lambda^{\max_1}, \lambda^{\max_1}, \infty, \infty)$ with $\lambda^{\max_1} = 0.8475970508$, hence (3.84) must hold for $L = 8$.

We thus have proved that (3.84) holds for $L \in \{2, 3, 4, 8\}$. Now it is easy to see that the sub-problem $\mathbb{P}_1^L(N, 0, L-N)$ (whose maximum function value is denoted as $f^{\max_1}(N, 0, L-N)$) is indeed equivalent to $\mathbb{P}_{2b}(\frac{N}{L})$ (with its objective function scaled by L), hence

$$f^{\max} = f^{\max_1} + \max_{N \in \{0,1,\ldots,L\}} f^{\max_1}(N, 0, L-N) = \max_{N \in \{0,1,\ldots,L\}} L \cdot \tau\left(\frac{N}{L}\right) = L \cdot \tau\left(\frac{N^\star}{L}\right).$$

and all other equations in (3.83) follows from (3.73) - (3.78), (3.81) and (3.82). □

The converse of Theorem 3 is proved since (3.18) follows directly from Lemmas 17 and 18.

D. Discussions

In this section, we first give an explanation why the $l^\star = 0.1083$ bit per sample per source supremum sum-rate loss coincides with the conjectured supremum Wyner-Ziv rate loss [39], then discuss an example in the two-terminal case, for which the non-degraded requirement in Theorem 3 is not needed because both the distributed encoding and joint encoding minimum sum-rates can be written in explicit forms. We also compute the supremum sum-rate loss in the positive symmetric case and compare it to that in the more general non-degraded case.

1. A coincidence with the rate loss in Wyner-Ziv coding

The asymptotic supremum sum-rate loss of $l^\star = 0.1083$ bit per sample per source echoes Zamir's conjecture on the supremum Wyner-Ziv rate loss [39]. In fact, the two numbers coincide because they are obtained through two *equivalent* optimization problems, as will be shown in this subsection.

In the Wyner-Ziv case, the 0.1083 b/s rate loss is achieved in the mixture Gaussian case with two mixture components [39]. In order to compare with our results, we consider a more general setting. Let $L \geq 2$, and $\{\sigma_i^2 : i \in \mathcal{L}\}$ be an ordered set of positive numbers such that $\sigma_i^2 \leq \sigma_j^2$ for any $1 \leq i < j \leq L$. Consider a mixture Gaussian source defined by $X = X_I$, with $X_i \sim \mathcal{N}(0, \sigma_i^2)$, and I as a discrete random variable taking value from the index set $\mathcal{L}$ and is independent of X_i's. Let $Pr(I = i) = q_i$ for $i \in \mathcal{L}$. The source X is available at the encoder, while the random variable I serves as the decoder side information. Denote D_{WZ} as the target Wyner-Ziv distortion. Then using the same arguments as in [39], it can be shown that the Wyner-Ziv rate-distortion function in this case is given by

$$R_{\text{WZ}}(D_{\text{WZ}}) = \sum_{i \in \mathcal{L}} \frac{q_i}{2} \log_2 \left(1 + \frac{\sigma_i^2}{\sigma_n^2}\right),$$

where σ_n^2 is the solution to

$$\sum_{i \in \mathcal{L}} q_i \cdot \left(\frac{1}{\sigma_i^2} + \frac{1}{\sigma_n^2}\right)^{-1} = D_{\text{WZ}},$$

and the conditional rate-distortion function is given by the reverse water-filling formula

$$R_{X|I}(D_{\text{WZ}}) \quad = \quad \sum_{i \in \mathcal{L} \,:\, \sigma_i^2 > w_{\text{WZ}}} \frac{q_i}{2} \log_2 \frac{w}{\sigma_i^2},$$

with w_{WZ} being the solution to

$$\sum_{i\in\mathcal{L}} q_i \cdot \min\left(\sigma_i^2, w_{\mathrm{WZ}}\right) = D_{\mathrm{WZ}}.$$

Then we observe that $L \cdot R_{\mathrm{WZ}}(D_{\mathrm{WZ}})$ and $L \cdot R_{X|I}(D_{\mathrm{WZ}})$ are exactly the same as $\overline{R}^{\mathrm{BT}}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D})$ and $\underline{R}^{\mathrm{Joint}}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D})$ defined in (3.24) and (3.25), respectively, after setting $D_{\mathrm{WZ}} = 1$, $q_i = \frac{1}{L}$ for $i \in \mathcal{L}$, and interchanging the following pairs,

$$\sigma_i^2 \leftrightarrow \lambda_i, \quad p \leftrightarrow \frac{1}{\sigma_n^2}, \quad w \leftrightarrow w_{\mathrm{WZ}}.$$

Hence the optimization problem of maximizing the Wyner-Ziv rate loss $R_{\mathrm{WZ}}(D_{\mathrm{WZ}}) - R_{X|I}(D_{\mathrm{WZ}})$ under the constraint that $q_i = \frac{1}{L}$ for $i \in \mathcal{L}$ is indeed equivalent to the core optimization problem $\mathbb{P}_0^L$. Then as $L \to \infty$, the constraint $q_i = \frac{1}{L}$ vanishes because σ_i^2's can take repeating values, and rational numbers are dense on the real line. Therefore, the supremum Wyner-Ziv rate loss in the above-defined mixture Gaussian case equals to the limit of (per-source) supremum sum-rate loss in the quadratic Gaussian case (under the non-degraded assumption) as L goes to infinity. Moreover, both supremums are achieved in the bimodal/bi-eigen case that are illustrated in Fig. 9.

2. The special two-terminal case

Our main result in Theorem 3 is derived only for the non-degraded case. However, if we consider the simplest case $L = 2$, the statement will still hold without making the non-degraded assumption. In this subsection, we will compute the exact MT sum-rate and joint-encoding sum-rate for the two-terminal case and show that the sum-rate loss between them is equal to the supremum value in the non-degraded case. Without loss of generality, we will assume that $\boldsymbol{\Sigma_Y} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$ for some $0 < \rho < 1$

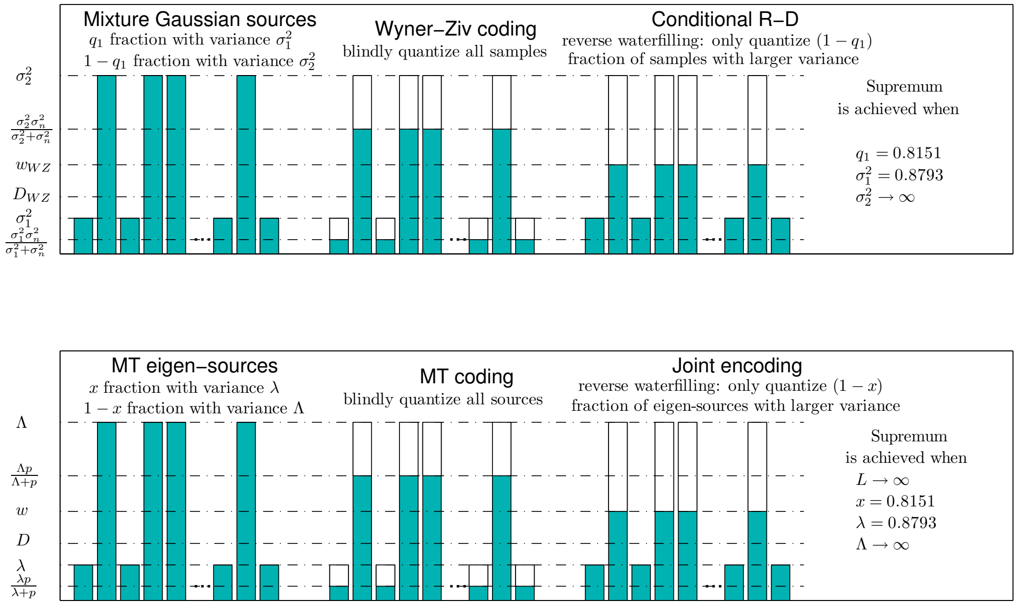

Mixture Gaussian sources
q_1 fraction with variance σ_1^2
$1-q_1$ fraction with variance σ_2^2
Wyner−Ziv coding
blindly quantize all samples
Conditional R−D
reverse waterfilling: only quantize $(1-q_1)$
fraction of samples with larger variance
σ_2^2
$\frac{\sigma_2^2\sigma_n^2}{\sigma_2^2+\sigma_n^2}$
w_{WZ}
D_{WZ}
σ_1^2
$\frac{\sigma_1^2\sigma_n^2}{\sigma_1^2+\sigma_n^2}$
Supremum
is achieved when
$q_1 = 0.8151$
$\sigma_1^2 = 0.8793$
$\sigma_2^2 \to \infty$
MT eigen−sources
x fraction with variance λ
$1-x$ fraction with variance Λ
MT coding
blindly quantize all sources
Joint encoding
reverse waterfilling: only quantize $(1-x)$
fraction of eigen-sources with larger variance
Λ
$\frac{\Lambda p}{\Lambda+p}$
w
D
λ
$\frac{\lambda p}{\lambda+p}$
Supremum
is achieved when
$L \to \infty$
$x = 0.8151$
$\lambda = 0.8793$
$\Lambda \to \infty$

Fig. 9. A comparison between the supremum rate loss in quadratic Gaussian Wyner-Ziv coding and the supremum sum-rate loss in quadratic Gaussian MT coding.

throughout this subsection (note that the cases when $\rho = 0$ and $\rho = 1$ are trivial and result in zero sum-rate loss).

We first find the degraded case for $L = 2$. In fact, it is easy to find the two solutions of $\mathbf{\Lambda}$ to $\text{diag}(\mathbf{\Sigma}_{\boldsymbol{Y}} - \mathbf{\Sigma}_{\boldsymbol{Y}}(\mathbf{\Sigma}_{\boldsymbol{Y}} + \mathbf{\Lambda})^{-1}\mathbf{\Sigma}_{\boldsymbol{Y}}) = (D_1, D_2)^{\mathrm{T}}$ as required by (3.2),

$$\text{diag}\left(\frac{(1-\rho^2)(2D_1+\rho^2-1\pm\sqrt{1+\rho^4-2\rho^2+4D_1D_2\rho^2})}{2((1-D_1)-\rho^2(1-D_2))},\right.$$
$$\left.\frac{(1-\rho^2)(2D_2+\rho^2-1\pm\sqrt{1+\rho^4-2\rho^2+4D_1D_2\rho^2})}{2((1-D_2)-\rho^2(1-D_1))}\right).$$

Hence a pair $(\mathbf{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})$ is non-degraded (which is equivalent to $\mathbf{\Lambda} \succeq 0$) if and only if

$$\max\{D_1, D_2\} \leq 1 - \rho^2(1 - \min\{D_1, D_2\}) \Leftrightarrow \rho \leq \sqrt{\frac{1-\max\{D_1, D_2\}}{1-\min\{D_1, D_2\}}}. \qquad (3.88)$$

The MT minimum sum-rate is given in [12] as

$$R^{\mathrm{MT}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}((D_1, D_2)^{\mathrm{T}}) = \begin{cases} \frac{1}{2}\log_2 \frac{1}{\min\{D_1,D_2\}} & \rho \geq \sqrt{\frac{1-\max\{D_1,D_2\}}{1-\min\{D_1,D_2\}}} \\ \frac{1}{2}\log_2 \frac{(1-\rho^2)\beta_{\max}}{2D_1D_2} & \rho < \sqrt{\frac{1-\max\{D_1,D_2\}}{1-\min\{D_1,D_2\}}} \end{cases}, \tag{3.89}$$

where $\beta_{\max} = 1 + \sqrt{1 + \frac{4\rho^2 D_1 D_2}{(1-\rho^2)^2}}$, while the joint encoding minimum sum-rate is given in the following lemma.

Lemma 19. *The joint encoding minimum sum-rate for* $\boldsymbol{\Sigma}_{\boldsymbol{Y}} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$ *and* $\boldsymbol{D} = (D_1, D_2)^{\mathrm{T}}$ *is given by*

$$R^{\mathrm{Joint}}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}\left((D_1, D_2)^{\mathrm{T}}\right) = \begin{cases} \frac{1}{2}\log_2 \frac{1}{\min\{D_1,D_2\}}, & \rho \geq \rho^{\dagger} \\ \frac{1}{2}\log_2 \frac{(1-\rho^2)}{(1-\theta^2_{\max})D_1D_2} & \rho^{\ddagger} \leq \rho < \rho^{\dagger} \\ \frac{1}{2}\log_2 \frac{(1-\rho^2)}{D_1D_2}, & \rho < \rho^{\ddagger} \end{cases}, \tag{3.90}$$

where $\rho^{\dagger} = \sqrt{\frac{1-\max\{D_1,D_2\}}{1-\min\{D_1,D_2\}}}$, $\rho^{\ddagger} = \sqrt{(1-D_1)(1-D_2)}$, *and* $\theta_{\max} = \frac{\rho - \sqrt{(1-D_1)(1-D_2)}}{\sqrt{D_1D_2}}$.

Proof. Computing the joint encoding minimum sum-rate is equivalent to solving the following convex semidefinite programming problem,

$$\begin{aligned} \text{Min.} \quad & -\log\det(\boldsymbol{\mathcal{D}}) \\ \text{s.t.} \quad & \mathbf{0} \preceq \boldsymbol{\mathcal{D}} \preceq \boldsymbol{\Sigma}_{\boldsymbol{Y}} \\ & \boldsymbol{E}_1\boldsymbol{\mathcal{D}}\boldsymbol{E}_1 \preceq D_1\boldsymbol{E}_1 \\ & \boldsymbol{E}_2\boldsymbol{\mathcal{D}}\boldsymbol{E}_2 \preceq D_2\boldsymbol{E}_2, \end{aligned}$$

where $\boldsymbol{E}_i$ is the 2×2 matrix whose (i, i)-th element is one with all others being zero. It is easy to verify that the Slater's condition [43] holds, hence KKT condition is necessary and sufficient for global optimality. Note that the optimal $\boldsymbol{\mathcal{D}}$ cannot be singular, hence the constraint $\mathbf{0} \preceq \boldsymbol{\mathcal{D}}$ must not be active.

The lagrangian is $-\log\det(\boldsymbol{\mathcal{D}})+\mathrm{tr}(\boldsymbol{\Omega}(\boldsymbol{\mathcal{D}}-\boldsymbol{\Sigma}_{\boldsymbol{Y}}))+\sum_{i=1}^{2}\mathrm{tr}(\boldsymbol{\Pi}_i(\boldsymbol{E}_i\boldsymbol{\mathcal{D}}\boldsymbol{E}_i-D_i\boldsymbol{E}_i))$ where $\boldsymbol{\Omega}$ and $\boldsymbol{\Pi}_i$'s are p.s.d. Lagrangian multipliers, and the KKT condition is

$$\begin{aligned}
-\boldsymbol{\mathcal{D}}^{-1}+\boldsymbol{\Omega}+\boldsymbol{E}_1\boldsymbol{\Pi}_1\boldsymbol{E}_1+\boldsymbol{E}_2\boldsymbol{\Pi}_2\boldsymbol{E}_2 &= \mathbf{0},\\
\boldsymbol{\Omega}(\boldsymbol{\mathcal{D}}-\boldsymbol{\Sigma}_{\boldsymbol{Y}}) &= \mathbf{0},\\
\boldsymbol{\Pi}_i(\boldsymbol{E}_i\boldsymbol{\mathcal{D}}\boldsymbol{E}_i-D_i\boldsymbol{E}_i) &= \mathbf{0}, \quad i=1,2,\\
\boldsymbol{\mathcal{D}}-\boldsymbol{\Sigma}_{\boldsymbol{Y}} &\preceq \mathbf{0},\\
\boldsymbol{E}_1\boldsymbol{\mathcal{D}}\boldsymbol{E}_1-D_1\boldsymbol{E}_1 &\preceq \mathbf{0},\\
\boldsymbol{E}_2\boldsymbol{\mathcal{D}}\boldsymbol{E}_2-D_2\boldsymbol{E}_2 &\preceq \mathbf{0}.
\end{aligned}$$

First, if the case is degraded, we assume that $D_2 > 1-\rho^2(1-D_1)$ without loss of generality. It is easy to show that

$$\boldsymbol{\mathcal{D}}^\dagger = \begin{bmatrix} D_1 & \rho D_1 \\ \rho D_1 & 1-\rho^2(1-D_1) \end{bmatrix},\ \boldsymbol{\Omega}^\dagger = \frac{1}{1-\rho^2}\cdot\begin{bmatrix} \rho^2 & -\rho \\ -\rho & 1 \end{bmatrix},\ \boldsymbol{\Pi}_1^\dagger = \begin{bmatrix} \frac{1}{D_1} & 0 \\ 0 & 0 \end{bmatrix},\ \boldsymbol{\Pi}_2^\dagger = \mathbf{0}$$

satisfy the KKT condition. Hence the minimum joint encoding sum-rate is $\frac{1}{2}\log_2\frac{\det(\boldsymbol{\Sigma}_{\boldsymbol{Y}})}{\det(\boldsymbol{\mathcal{D}}^\dagger)} = \frac{1}{2}\log_2\frac{1}{D_1}$.

Then consider the case when $\rho^\ddagger \le \rho < \rho^\dagger$. One can verify that

$$\boldsymbol{\mathcal{D}}^\dagger = \begin{bmatrix} D_1 & \rho-\rho^\ddagger \\ \rho-\rho^\ddagger & D_2 \end{bmatrix},\ \boldsymbol{\Omega}^\dagger = \frac{1}{\det(\boldsymbol{\mathcal{D}}^\dagger)}\cdot\begin{bmatrix} (1-D_2)(\frac{\rho}{\rho^\ddagger}-1) & \rho^\ddagger-\rho \\ \rho^\ddagger-\rho & (1-D_1)(\frac{\rho}{\rho^\ddagger}-1) \end{bmatrix},$$

$$\boldsymbol{\Pi}_1^\dagger = \frac{1}{\det(\boldsymbol{\mathcal{D}}^\dagger)}\cdot\begin{bmatrix} 1-\rho\sqrt{\frac{1-D_2}{1-D_1}} & 0 \\ 0 & 0 \end{bmatrix},\ \boldsymbol{\Pi}_2^\dagger = \frac{1}{\det(\boldsymbol{\mathcal{D}}^\dagger)}\cdot\begin{bmatrix} 0 & 0 \\ 0 & 1-\rho\sqrt{\frac{1-D_1}{1-D_2}} \end{bmatrix},$$

satisfy the KKT condition, resulting in a minimum joint encoding sum-rate of

$$\frac{1}{2}\log_2\frac{\det(\boldsymbol{\Sigma}_{\boldsymbol{Y}})}{\det(\boldsymbol{\mathcal{D}}^\dagger)} = \frac{1}{2}\log_2\frac{(1-\rho^2)}{(1-\theta_{\max}^2)D_1D_2}.$$

Finally, when $\rho < \sqrt{(1-D_1)(1-D_2)}$, the KKT condition holds for

$$\boldsymbol{\mathcal{D}}^{\dagger} = \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix}, \ \boldsymbol{\Omega}^{\dagger} = \mathbf{0}, \ \boldsymbol{\Pi}_1^{\dagger} = \begin{bmatrix} \frac{1}{D_1} & 0 \\ 0 & 0 \end{bmatrix}, \ \boldsymbol{\Pi}_2^{\dagger} = \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{D_2} \end{bmatrix},$$

and (3.90) readily follows. □

Comparing (3.89) with (3.90), we observe that in the degraded case defined by (3.88), the two minimum sum-rates are the same, which means the sum-rate loss is always zero. This fact and Theorem 3 lead to

$$\begin{aligned}\sup_{(\boldsymbol{\Sigma_Y},\boldsymbol{D})} R^{\Delta}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}) &= \max\left\{ \sup_{(\boldsymbol{\Sigma_Y},\boldsymbol{D})\,\in\,\mathscr{S}_L^{\mathrm{BT}}} R^{\Delta}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}), \sup_{(\boldsymbol{\Sigma_Y},\boldsymbol{D})\,\notin\,\mathscr{S}_L^{\mathrm{BT}}} R^{\Delta}_{\boldsymbol{\Sigma_Y}}(\boldsymbol{D}) \right\} \\ &= \max\left\{ 2\cdot\max\left[\tau(\frac{\lfloor 2x^{\star}\rfloor}{2}), \tau(\frac{\lceil 2x^{\star}\rceil}{2})\right], 0 \right\} \\ &= \frac{1}{2}\log_2\frac{5}{4} \ \approx \ 0.161 \text{ b/s.}\end{aligned}$$

Remark 5: The 0.161 b/s supremum sum-rate loss for $L = 2$ is much smaller than the one b/s upper bound provided by Zamir in [39], although the latter is a universal upper bound (for MSE distortion measure) that does not required the sources to be jointly Gaussian.

3. Comparison with the supremum sum-rate loss in the symmetric case

In the examples given in Section 1, we already know that the supremum sum-rate loss under the non-degraded assumption equals to that in the positive symmetric case if and only if $L \leq 7$, it is thus interesting to also compute the supremum in the later case for $L > 7$.

For the positive symmetric case [12], there is no loss of generality to assume that $\boldsymbol{\Sigma_Y} = \boldsymbol{\mathcal{S}}_L(1,\rho)$ and $\boldsymbol{D} = D\mathbf{1}$ with $D < 1$. Then the optimal joint encoding scheme is through reverse water-filling [15], with the minimum rate given by

$$R_{\boldsymbol{\Sigma_Y}}^{\text{Joint}}(D\mathbf{1}) = \begin{cases} \frac{1}{2}\log_2 \frac{\delta_L(\rho)}{D^L\delta_L(1-\frac{1-\rho}{D})} & \text{if } D > 1-\rho \\ \frac{1}{2}\log_2 \frac{\delta_L(\rho)}{D^L} & \text{if } D \le 1-\rho \end{cases},$$

where $\delta_L(x) \overset{\Delta}{=} (1-x)^{L-1}(1+(L-1)x)$ for any $-\frac{1}{L-1} \le x \le 1$.

On the other hand, for any $L \ge 2$, $\rho \in (0,1)$ and $D \in (0,1)$, the minimum sum-rate of quadratic Gaussian MT source coding is given in exact form as

$$R_{\boldsymbol{\Sigma_Y}}^{\text{MT}}(D\mathbf{1}) = R_{\boldsymbol{\Sigma_Y}}^{\star}(D\mathbf{1}) = \frac{1}{2}\log_2 \frac{\delta_L(\rho)}{D^L\delta_L(\theta^{\text{MT}})},$$

where

$$\theta^{\text{MT}} \;\; = \;\; t^{\text{MT}} + \sqrt{(t^{\text{MT}})^2 + 1/(L-1)}, \tag{3.91}$$

with $t^{\text{MT}} = \frac{L-2}{2(L-1)} - \frac{(1-\rho)(1+(L-1)\rho)}{2(L-1)D\rho}$. The proof can be found in [12, 13, 44].

Now we can compute the exact sum-rate loss in this positive symmetric case.

$$R_{\boldsymbol{\Sigma_Y}}^{\Delta}(D\mathbf{1}) = R_{\boldsymbol{\Sigma_Y}}^{\text{MT}}(D\mathbf{1}) - R_{\boldsymbol{\Sigma_Y}}^{\text{Joint}}(D\mathbf{1}) = \begin{cases} \frac{1}{2}\log_2 \frac{\delta_L(\theta^{\text{Joint}})}{\delta_L(\theta^{\text{MT}})} & D > 1-\rho \\ \frac{1}{2}\log_2 \frac{1}{\delta_L(\theta^{\text{MT}})} & D \le 1-\rho \end{cases},$$

where $\theta^{\text{Joint}} = 1 - \frac{1-\rho}{D}$ and θ^{MT} is given in (3.91).

An example of the sum-rate loss $R_{\boldsymbol{\Sigma_Y}}^{\Delta}(D\mathbf{1})$ is plotted in Fig. 10 as a function of ρ and D for $L = 2$. When $\rho = 0$, all sources are independent, hence $R_{\boldsymbol{\Sigma_Y}}^{\text{MT}}(D\mathbf{1}) = R_{\boldsymbol{\Sigma_Y}}^{\text{Joint}}(D\mathbf{1}) = \frac{L}{2}\log_2\frac{1}{D}$ and $R_{\boldsymbol{\Sigma_Y}}^{\Delta}(D\mathbf{1}) = 0$; when $\rho = 1$, all sources are statistically identical, thus coding one of them suffices, hence $R_{\boldsymbol{\Sigma_Y}}^{\text{MT}}(D\mathbf{1}) = R_{\boldsymbol{\Sigma_Y}}^{\text{Joint}}(D\mathbf{1}) = \frac{1}{2}\log_2\frac{1}{D}$ and $R_{\boldsymbol{\Sigma_Y}}^{\Delta}(D\mathbf{1}) = 0$; when $D = 0$, we have a Slepian-Wolf coding problem of L sources, hence $R_{\boldsymbol{\Sigma_Y}}^{\Delta}(D\mathbf{1}) = 0$ due to the no rate loss conclusion of the Slepian-Wolf theorem [2] and its extensions[15, 45]; finally, when $D = 1$, $R_{\boldsymbol{\Sigma_Y}}^{\text{MT}}(D\mathbf{1}) = R_{\boldsymbol{\Sigma_Y}}^{\text{Joint}}(D\mathbf{1}) = 0$ and the rate loss is also zero.

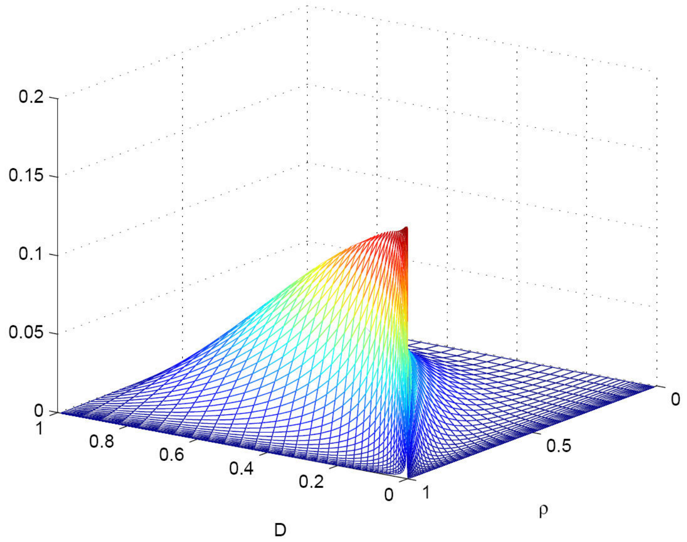


Fig. 10. The sum-rate loss $R^{\Delta}_{\mathbf{\Sigma}_{\boldsymbol{Y}}}(D\mathbf{1})$ for quadratic Gaussian MT source coding in the positive symmetric case for $L = 2$.

For any fixed $\rho \in (0, 1)$, there is a maximum sum-rate loss over all D's, and this maximum sum-rate loss (as a function of ρ) monotonically increases to a supremum value as $\rho \to 1$. Moreover, it is seen from Fig. 10 that the distortion that maximizes the sum-rate loss goes to zero as $\rho \to 1$. This implies that the supremum sum-rate loss is approached from below as both minimum sum-rates $R^{\text{Joint}}_{\mathbf{\Sigma}_{\boldsymbol{Y}}}(D\mathbf{1})$ and $R^{\text{MT}}_{\mathbf{\Sigma}_{\boldsymbol{Y}}}(D\mathbf{1})$ go to infinity, while the difference between them remains finite. And the sum-rate loss $R^{\Delta}_{\mathbf{\Sigma}_{\boldsymbol{Y}}}(D\mathbf{1})$ has a singularity point at $(\rho, D) = (1, 0)$.

The exact form of the supremum sum-rate loss in the positive symmetric case is given in the following lemma.

Lemma 20. *For a given $L \geq 2$, the supremum sum-rate loss over all possible ρ's and*

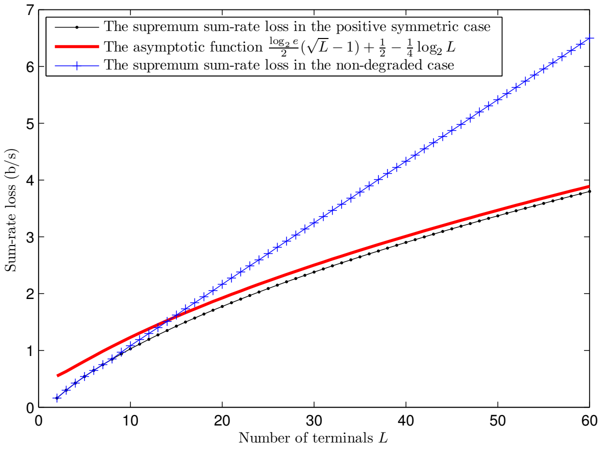


Fig. 11. The supremum sum-rate loss of quadratic Gaussian MT source coding in the positive symmetric case. The supremums are the same in the positive symmetric case and the non-degraded case when $L \leq 7$.

D's is

$$\sup_{\rho\in(0,1),D\in(0,1)} R^{\Delta}_{\boldsymbol{\Sigma_Y}}(D\mathbf{1}) = \frac{L-1}{2}\log_2 \frac{1-\frac{2L+1-\sqrt{1+4L}}{2L^2}}{1-\frac{-1+\sqrt{1+4L}}{2L}} + \frac{1}{2}\log_2 \frac{1+(L-1)\frac{2L+1-\sqrt{1+4L}}{2L^2}}{1+(L-1)\frac{-1+\sqrt{1+4L}}{2L}} \tag{3.92}$$

$$\overset{L\to\infty}{\rightsquigarrow} \frac{\sqrt{L}-1}{2}\log_2 \mathrm{e} + \frac{1}{2} - \frac{1}{4}\log_2 L = 0.7213\sqrt{L} + o(\sqrt{L}), \tag{3.93}$$

where $A \overset{L\to\infty}{\rightsquigarrow} B$ *means* $\lim_{L\to\infty}(A-B)=0$.

Proof. First, for fixed L and ρ, θ^{MT} is a monotone increasing function of $D \in (0,1)$ because

$$\frac{\partial\theta^{\mathrm{MT}}}{\partial D} = \frac{\partial\theta^{\mathrm{MT}}}{\partial t^{\mathrm{MT}}}\cdot\frac{\partial t^{\mathrm{MT}}}{\partial D} = \left(1+\frac{t^{\mathrm{MT}}}{\sqrt{(t^{\mathrm{MT}})^2+\frac{1}{L-1}}}\right)\cdot\left(\frac{(1-\rho)(1+(L-1)\rho)}{2(L-1)D^2\rho}\right) > 0.$$

Then $\delta_L(\theta^{\mathrm{MT}})$ is a monotone decreasing function of $D \in (0,1)$. Consequently, when $D \leq 1-\rho$, $R^{\Delta}_{\boldsymbol{\Sigma_Y}}(D\mathbf{1}) = -\frac{1}{2}\log_2 \delta_L(\theta^{\mathrm{MT}})$ is a monotone increasing function of $D \in (0,1)$.

Hence we have

$$\begin{aligned}\sup_{\rho\in(0,1),D\in(0,1)} R^{\Delta}_{\boldsymbol{\Sigma_Y}}(D\mathbf{1}) &= \max\left\{\sup_{\rho\in(0,1),D\in(1-\rho,1)} R^{\Delta}_{\boldsymbol{\Sigma_Y}}(D\mathbf{1}), \sup_{\rho\in(0,1),D\in(0,1-\rho]} R^{\Delta}_{\boldsymbol{\Sigma_Y}}(D\mathbf{1})\right\}\\ &= \max\left\{\sup_{\rho\in(0,1),D\in(1-\rho,1)} R^{\Delta}_{\boldsymbol{\Sigma_Y}}(D\mathbf{1}), \sup_{\rho\in(0,1)} R^{\Delta}_{\boldsymbol{\Sigma_Y}}((1-\rho)\mathbf{1})\right\}\\ &= \sup_{\rho\in(0,1),D\in[1-\rho,1)} R^{\Delta}_{\boldsymbol{\Sigma_Y}}(D\mathbf{1})\\ &= \sup_{\rho\in(0,1),D\in[1-\rho,1)} \frac{1}{2}\log_2\frac{\delta_L(\theta^{\mathrm{Joint}})}{\delta_L(\theta^{\mathrm{MT}})}.\end{aligned}$$

Now denote $\mathscr{F}_L(\rho, D) = \frac{\delta_L(\theta^{\mathrm{Joint}})}{\delta_L(\theta^{\mathrm{MT}})}$, we have

$$\begin{aligned}\frac{\partial\mathscr{F}_L(\rho,D)}{\partial D} = \frac{\partial\left[\frac{\delta_L(\theta^{\mathrm{Joint}})}{\delta_L(\theta^{\mathrm{MT}})}\right]}{\partial D} &= \frac{-L(L-1)(1-\theta^{\mathrm{Joint}})^{L-2}(1-\theta^{\mathrm{MT}})^{L-2}}{\delta_L^2(\theta^{\mathrm{MT}})}\\ &\cdot\left[\theta^{\mathrm{Joint}}(1-\theta^{\mathrm{MT}})(1+(L-1)\theta^{\mathrm{MT}})\frac{\partial\theta^{\mathrm{Joint}}}{\partial D}\right.\\ &\left.-\theta^{\mathrm{MT}}(1-\theta^{\mathrm{Joint}})(1+(L-1)\theta^{\mathrm{Joint}})\frac{\partial\theta^{\mathrm{MT}}}{\partial D}\right].\end{aligned} \tag{3.94}$$

Setting $\frac{\partial\mathscr{F}_L(\rho,D)}{\partial D}$ to zero, we have a unique solution in $[1-\rho, 1)$, namely,

$$D^{\star}_{\rho} = \begin{cases} \frac{(1+\rho)^2(1-\rho)}{1+2\rho} & L = 2\\ \frac{1-\rho}{2\rho(L-2)(2+(L-2)\rho)}\cdot[-\sqrt{1+4\rho+4\rho^2(L-1)} & \\ \quad +(2(L-1)(L-2)\rho^2+2(2L-3)\rho+1)] & L > 2. \end{cases}$$

Then we compute

$$\begin{aligned}\theta^{\mathrm{MT}}\mid_{D=D^{\star}_{\rho}} &= \frac{-1+\sqrt{1+4\rho+4\rho^2(L-1)}}{2(1+(L-1)\rho)} \triangleq \theta^{\mathrm{MT}}_{\max}(\rho),\\ \theta^{\mathrm{Joint}}\mid_{D=D^{\star}_{\rho}} &\frac{2\rho(1+(L-1)\rho)+1-\sqrt{1+4\rho+4\rho^2(L-1)}}{\rho(L-1)(2+(L-1)\rho)+1} \triangleq \theta^{\mathrm{Joint}}_{\max}(\rho).\end{aligned}$$

Hence

$$\frac{\partial \mathscr{F}_L(\rho, D^\star_\rho)}{\partial \rho} = \begin{cases} \frac{-2}{\delta_2^2(\theta^{\mathrm{MT}}_{\max}(\rho))}\left[-\frac{\rho(1+2\rho)^2}{(1+\rho)^7}\right], & L = 2 \\ [\mathscr{A} + \mathscr{B}\sqrt{1+4\rho+4\rho^2(L-1)}]\cdot & \\ \quad \frac{-L(L-1)(1-\theta^{\mathrm{Joint}}_{\max}(\rho))^{L-2}(1-\theta^{\mathrm{MT}}_{\max}(\rho))^{L-2}}{\delta_L^2(\theta^{\mathrm{MT}}_{\max}(\rho))}, & L > 2, \end{cases}$$

where $\mathscr{A}$ and $\mathscr{B}$ are rational functions of L and ρ. We observe that for $L = 2$, $\frac{\partial \mathscr{F}_L(\rho, D^\star_\rho)}{\partial \rho} > 0$ for any $\rho \in (0,1)$. Moreover, it is not hard to verify that $\mathscr{A}$ and $\mathscr{B}$ satisfy $\mathscr{B} < 0$ and the following condition,

$$\mathscr{A}^2 - \mathscr{B}^2 \times (1+4\rho+4\rho^2(L-1)) = -\frac{\rho(L-2)^2(2+(L-2)\rho)^2}{(1+(L-1)\rho)^7} < 0,$$

which implies that $\frac{\partial \mathscr{F}_L(\rho, D^\star_\rho)}{\partial \rho} > 0$ for any $L \in \mathbb{N} \cap (2, \infty)$ and $\rho \in (0,1)$, hence

$$\begin{aligned} \sup_{\rho\in(0,1),D\in(0,1)} R^{\Delta}_{\mathbf{\Sigma}_{\boldsymbol{Y}}}(D\mathbf{1}) &= \sup_{\rho\in(0,1),D\in[1-\rho,1)} \frac{1}{2}\log_2 \frac{\delta_L(\theta^{\mathrm{Joint}})}{\delta_L(\theta^{\mathrm{MT}})} && (3.95) \\ &= \lim_{\rho\to1} \frac{1}{2}\log_2 \mathscr{F}_L(\rho, D^\star_\rho) && (3.96) \\ &= \lim_{\rho\to1} \frac{1}{2}\log_2 \frac{\delta_L(\theta^{\mathrm{Joint}}_{\max}(\rho))}{\delta_L(\theta^{\mathrm{MT}}_{\max}(\rho))} && (3.97) \\ &= \frac{1}{2}\log_2 \frac{\delta_L(\lim_{\rho\to1}\theta^{\mathrm{Joint}}_{\max}(\rho))}{\delta_L(\lim_{\rho\to1}\theta^{\mathrm{MT}}_{\max}(\rho))} && (3.98) \\ &= \frac{1}{2}\log_2 \frac{\delta_L(\frac{2L+1-\sqrt{1+4L}}{2L^2})}{\delta_L(\frac{-1+\sqrt{1+4L}}{2L})} \\ &= \frac{L-1}{2}\log_2 \frac{1-\frac{2L+1-\sqrt{1+4L}}{2L^2}}{1-\frac{-1+\sqrt{1+4L}}{2L}} + \frac{1}{2}\log_2 \frac{1+(L-1)\frac{2L+1-\sqrt{1+4L}}{2L^2}}{1+(L-1)\frac{-1+\sqrt{1+4L}}{2L}} \\ &\overset{L\to\infty}{\rightsquigarrow} \frac{1}{2}\log_2(1-\frac{1}{L})^{L-1} - \frac{L-1}{2\sqrt{L}}\log_2(1-\frac{1}{\sqrt{L}})^{\sqrt{L}} - \frac{1}{2}\log_2\frac{1}{\sqrt{4L}} \\ &\overset{L\to\infty}{\rightsquigarrow} \frac{\log_2 \mathrm{e}}{2}(\sqrt{L}-1) + \frac{1}{2} - \frac{1}{4}\log_2 L = 0.7213\sqrt{L} + o(\sqrt{L}). \end{aligned}$$

□

From (3.93), we see that as L increases, the supremum sum-rate loss in the positive symmetric case increases in the order of $\sqrt{L}$, since $\lim_{L\to\infty} \frac{1/2-1/4\log_2 L}{\sqrt{L}} = 0$. Fig.

11 plots the supremum sum-rate loss $\sup_{\rho\in(0,1),D\in(0,1)} R^{\Delta}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}(D\mathbf{1})$ and its asymptotic function in (3.93), as well as the supremum sum-rate loss in the non-degraded case for comparison.

CHAPTER IV

PRACTICAL CODE DESIGN FOR MT SOURCE CODING

In this chapter, the practical coding scheme for more than two terminals is proposed in Section A. Section B presents our approximation analysis on the correlation channel for LDPC code design. Section C provides simulation results to show the small sum-rate loss of our practical design.

A. Proposed scheme for multiterminal source coding

In this section, we present our proposed code design for both quadratic Gaussian direct and indirect MT coding with more than two terminals based on SWCQ, where TCQ (TCVQ in low rate regime) is used for source quantization, and LDPC-based SW compression is employed to exploit the source correlation after quantization. Moreover, the correlation model between quantized sources is analyzed for SWCQ code design. Our aim is to approach all corner points of the sum-rate bound – other points on the sum-rate bound can be achieved by time sharing.

1. TCQ (TCVQ) quantizer design

The two components of SWCQ are quantization and SW coding. According to [35], both have to be optimal in order to approach the sum-rate bound: the quantizer needs to achieve the maximum 1.53 dB granular gain for Gaussian sources and SW coders must compress the quantized sources to their joint entropy.

TCQ [18] provides an efficient means of quantization. Given a rate R b/s and a memory size M, TCQ constructs an expanded signal set (ESS) $\mathcal{D}$ of size 2^{R+1}, i.e.,

$$\mathcal{D} = \left\{ \left(-2^R + \frac{1}{2}\right)\Delta, \left(-2^R + \frac{3}{2}\right)\Delta, \ldots, \left(2^R - \frac{1}{2}\right)\Delta \right\},$$

where Δ is the quantization step, and a rate-1/2 trellis of memory M, whose polynomials can be chosen according to [20]. Then for a source sequence, TCQ employs the Viterbi algorithm to find the sequence of codewords that is closest to the source sequence in the MSE sense.

To keep the quantization noises independent of difference sources, a dithering sequence can be generated (and then added to each source) by a simple i.i.d. uniformly distributed source, which reduces the complexity of TCQ when compared to dithered lattice quantization (this requires the dither sequence to be uniformly distributed over the basic Voronoi region) [20].

For practical TCQ design, we use the polynomial searching algorithm in [20] to find a good trellis for 8192-state (memory-13) TCQ with a granular gain of $g_{\text{TCQ}} = 1.428$ dB. The loss compared to the maximum possible granular gain $g_{\max} = 1.53$ dB is about 0.1 dB.

Since the trellis bit in TCQ has memory (whereas the codeword bits are sample-wise independent given the trellis bit), if we directly transmit the trellis bit using 1 b/s, the rate will be too high when the total rate budget for some terminal is less than 1 b/s. This scenario often arises in the low-rate regime. Hence we resort to k-D TCVQ [19] so that the rate for transmitting the trellis bit is $1/k$ b/s. TCVQ in conjunction with SW encoding forms an SWC-TCVQ [20] scheme for WZ coding, in which the trellis bit is transmitted without compression. It is difficult to analyze the asymptotical performance of SWC-TCVQ in the low-rate regime due to the complexity when computing the conditional distribution of the source given TCQ/TCVQ quantized side information.

In practical design, we use 2-D 8192-state TCVQ with a maximal granular gain of $g_{\text{TCVQ}} = 1.345$ dB, which is smaller than g_{TCQ} with the same memory size due to the relatively smaller increase of minimum Euclidean distance of subdivision of cosets

[46].

2. SW code design based on LDPC codes

SW coding is implemented via syndrome-based binning. Each bit plane of a quantized source is partitioned into bins indexed by syndromes of a channel code. The encoder computes the syndrome $\boldsymbol{s} = \boldsymbol{x}\boldsymbol{H}^{\mathrm{T}}$ and sends it to the decoder at rate $R^{\mathrm{SW}} = (n - k)/n$ b/s, where $\boldsymbol{x}$ is a length-n binary input sequence and $\boldsymbol{H}$ is the $(n-k) \times n$ parity-check matrix of the LDPC code. Based on the side information $\boldsymbol{y}$ and received syndrome $\boldsymbol{s}$, the decoder finds the recovered sequence $\hat{\boldsymbol{x}}$ in the coset $\mathcal{X}_{\boldsymbol{s}} = \left\{\boldsymbol{x} \in \{0,1\}^n : \boldsymbol{x}\boldsymbol{H}^{\mathrm{T}} = \boldsymbol{s}\right\}$, i.e.,

$$\hat{\boldsymbol{x}} = \arg\max_{\boldsymbol{x} \in \mathcal{C}_{\boldsymbol{s}}} p(\boldsymbol{x}|\boldsymbol{y}). \tag{4.1}$$

In practical SW code design, we choose LDPC codes because of their capacity-approaching performance and flexibility in code design using density evolution. First, for each SW encoder, a certain number of training blocks (e.g., ten length-10^6 blocks) of source samples and side information samples are generated to estimate the actual correlation model between each WZ coded bit plane of the quantized sources and the side information. The LDPC code degree profiles are first designed with differential evolution [47] using the estimated correlation model, parity check matrices are then randomly generated according to the corresponding node-perspective degree profiles. Finally a full-search algorithm is employed to find length-four cycles in the corresponding Tanner graph for removal. This becomes harder as the rate of the LDPC code decreases. However, at large block lengths (e.g., 10^6 bits), these short cycles will not affect the decoding performance (in terms of bit error rate) very much.

3. Proposed scheme for direct MT coding

With the SWCQ components given above, we can set up the MT coding scheme. For the direct MT coding setup that the BT sum-rate bound is tight, since the sum-rate bound, denoted as $\partial\boldsymbol{\mathcal{R}}_{\boldsymbol{Y}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})$, is an $(L-1)$-dimensional contra-polymatroid [48], a corner point $\boldsymbol{R} = \left(R_{(1)}, R_{(2)}, \ldots, R_{(L)}\right)^{\mathrm{T}}$ corresponds to a coding scheme with

$$R_{(1)} = H\left(Y^n_{(1),\mathrm{Q}}\right) \text{ and } R_{(i)} = H\left(Y^n_{(i),\mathrm{Q}}|Y^n_{(1),\mathrm{Q}}, \ldots, Y^n_{(i-1),\mathrm{Q}}\right), \tag{4.2}$$

for $i = 2, 3, \ldots, L$, where $\left\{R_{(1)}, \ldots, R_{(L)}\right\}$ and $\left\{Y^n_{(1),\mathrm{Q}}, \ldots, Y^n_{(L),\mathrm{Q}}\right\}$ are the same arbitrary permutation of $\{R_1, \ldots, R_L\}$ and quantized indices $\left\{Y^n_{1,\mathrm{Q}}, \ldots, Y^n_{L,\mathrm{Q}}\right\}$, respectively. The sum rate can be written as

$$\sum_{i=1}^{L} R_{(i)} = \frac{1}{n} H\left(Y^n_{(1)}, \ldots, Y^n_{(L)}\right) \tag{4.3}$$

according to the chain rule.

Since $Y^n_1, Y^n_2, \ldots, Y^n_L$ are symmetric, (4.2) can be rewritten as

$$R_{(i)} = H_i, \quad i = 1, 2, \ldots, L, \tag{4.4}$$

where $\{H_1, H_2, \ldots, H_L\}$ are L possible rates for corner points. Therefore, the number of corner points for the L-terminal symmetric case is $L!$. Without loss of generality, we pick the corner point $\boldsymbol{R}_1 = (R_1, R_2, \ldots, R_L)^{\mathrm{T}}$ as an example, which corresponds to the coding scheme shown in Fig. 12. In this scheme, Y^n_1 is first quantized and entropy encoded by $\mathcal{E}^{\mathrm{ENT}}$ assuming that it cannot receive any side information at the decoder end from the other $L-1$ sources. Then, the second source Y^n_2 is similarly quantized but encoded with an SW encoder $\mathcal{E}^{\mathrm{SW}}_1$ and decoded using the decoded version $\tilde{Y}^n_1$ of Y^n_1 as side information. Similarly, the other $L-2$ sources $Y^n_3, Y^n_4, \ldots, Y^n_L$ are also quantized and encoded with SW encoders $\mathcal{E}^{\mathrm{SW}}_1, \mathcal{E}^{\mathrm{SW}}_2, \ldots, \mathcal{E}^{\mathrm{SW}}_{L-1}$, respectively, assuming each source Y_i^n can use the decoded version $\tilde{Y}_1^n, \tilde{Y}_2^n, \ldots, \tilde{Y}_{i-1}^n$ of $Y_1^n, Y_2^n, \ldots, Y_{i-1}^n$ as side information at the decoder end, the corresponding side information can be written as

$$\boldsymbol{Z}_{i-1} = Z_{i-1}^n = \begin{cases} \tilde{Y}_{i-1}^n, & i = 2; \\ \mathcal{C}_{i-2}\left[\tilde{Y}_1^n, \tilde{Y}_2^n, \ldots, \tilde{Y}_{i-1}^n\right], & i = 3, \ldots, L, \end{cases} \tag{4.5}$$

where $\mathcal{C}_{i-2}$ is a linear function, which means $\boldsymbol{Z}_{i-1}$ is a linear combination of the previous dequantized sources. If we assume ideal quantization of the input jointly Gaussian sources in the sense that the quantization errors are also Gaussian and independent of the sources, then $\boldsymbol{Z}_{i-1}$ provides a sufficient statistic for decoding Y_i^n. Finally, the recovered sources $\hat{Y}_1^n, \hat{Y}_2^n, \ldots, \hat{Y}_L^n$ are generated by a linear estimator $\mathcal{C}_\mathrm{e}$ based on the decoded signals $\tilde{Y}_1^n, \tilde{Y}_2^n, \ldots, \tilde{Y}_3^n$.

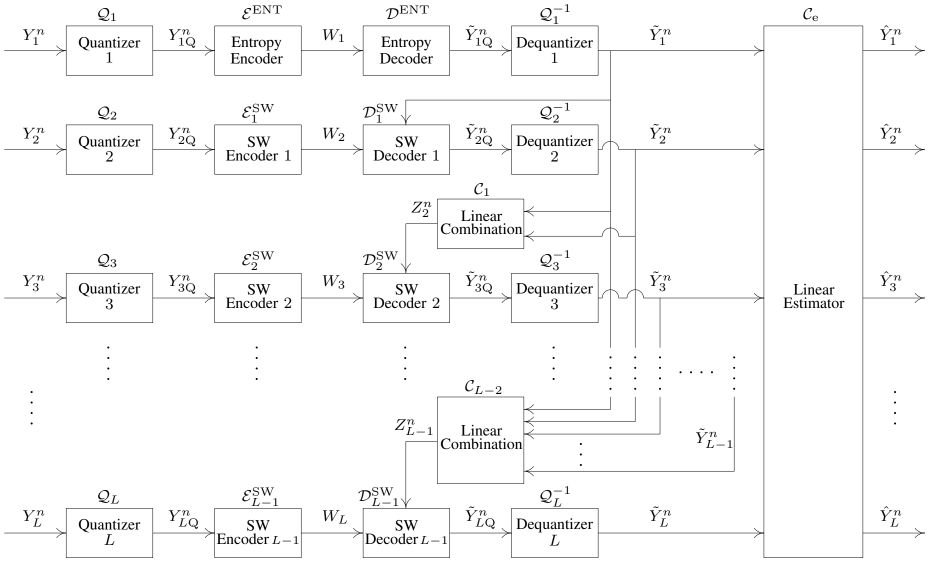


Fig. 12. Block diagram of the proposed SWCQ scheme for direct MT source coding.

Using the above method, we can achieve all $L!$ corner points in $\partial\boldsymbol{\mathcal{R}}_{\boldsymbol{Y}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})$. By changing the encoding order of the sources, we can approach all corner points of sum-rate bound as shown in Fig. 3 for the three-terminal positive symmetric case.

Moreover, since $\partial\boldsymbol{\mathcal{R}}_{\boldsymbol{Y}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})$ is a convex set in a $(L-1)$-dimensional hyperplane, all the other points in $\partial\boldsymbol{\mathcal{R}}_{\boldsymbol{Y}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})$ can be approached by time sharing.

Another method to achieve an arbitrary point $\boldsymbol{R}$ in $\partial\boldsymbol{\mathcal{R}}_{\boldsymbol{Y}}(\boldsymbol{\Sigma}_{\boldsymbol{Y}}, \boldsymbol{D})$ is source splitting, the two-terminal case of which has been exploited in [17, 49, 50]. In the three-terminal case, we can fix Y_1^n while splitting Y_2^n and Y_3^n into Y_{21}^n, Y_{22}^n, Y_{31}^n, and Y_{32}^n, respectively. The encoding and decoding order are set as $Y_{21}^n \rightarrow Y_{31}^n \rightarrow Y_1^n \rightarrow Y_{32}^n \rightarrow Y_{22}^n$. It is easy to show that every point on the sum-rate bound can be approached by source splitting. However, source splitting becomes more involved when the number of terminals increases. Thus we do not pursue source splitting in this work.

4. Coding scheme for indirect MT coding

We describe the scheme for the generalized CEO problem, which subsumes the original Gaussian CEO problem. Similar to the original CEO coding scheme in [17], we employ the same encoder and decoder as described in Section 3, except that the linear estimator uses dequantized observations $\tilde{Y}_1^n$, $\tilde{Y}_2^n$, $\ldots$, $\tilde{Y}_L^n$ to reconstruct the K remote sources directly, instead of reconstructing the L observations. The coding scheme for this case is shown in Fig. 13.

5. High rate analysis of the proposed scheme

In our proposed scheme, since quantization is followed by binning-based SW coding, the total loss can be divided into quantization loss due to source coding and binning loss due to channel coding. Similar to the high-rate performance analysis for the two-encoder case in [17], if we assume ideal binning by capacity-achieving (e.g., LDPC) channel coding and restrict ourselves to the high-rate scenario, i.e., $\max\{D_1^*, D_2^*, \ldots, D_L^*\} \rightarrow 0$, where D_i^* is the target distortion for the i-th source, $i = 1, \ldots, L$, the asymptotical performance of our TCQ-based SWCQ schemes for

MT source coding can be characterized by the following theorem.

Theorem 4. *If the BT sum-rate bound is tight for a quadratic Gaussian MT source coding problem, let* $(R_1^*, R_2^*, \ldots, R_L^*)$ *be a corner point on the BT sum-rate bound, then under ideal SW coding, the achievable sum-rate of our TCQ-based SWCQ scheme satisfies*

$$R = \sum_{i=1}^{L} R_i = \sum_{i=1}^{L} \left(R_i^* + \frac{1}{2} \log \left(2\pi \mathrm{e} G_{\mathcal{Q}_i} \right) \right) + o(1), \tag{4.6}$$

where $G_{\mathcal{Q}_1}, G_{\mathcal{Q}_2}, \ldots, G_{\mathcal{Q}_L}$ *are the equivalent normalized second moments of the Voronoi regions for the* L *trellis coded quantizers* $\mathcal{Q}_1, \mathcal{Q}_2, \ldots, \mathcal{Q}_L$*. And* $o(1) \to 0$ *as* $\max\{D_1^*, D_2^*, \ldots, D_L^*\} \to 0$ *and block length* $n \to \infty$*.*

Proof. Without loss of generality, assuming that the source vector $\boldsymbol{Y}$ is encoded in the order $Y_1, Y_2, \ldots, Y_L$, then Y_1 is first encoded with dithered TCQ quantizer $\mathcal{Q}_1$ which uses an ESS of size 2^{R+1}, with $\tilde{R} = 1$ and step size Δ_1. Thus, the ESS

$$\mathcal{D} = \left\{ -2^R + \frac{1}{2}\Delta_1, -2^R + \frac{3}{2}\Delta_1, \ldots, 2^R - \frac{1}{2}\Delta_1 \right\} \tag{4.7}$$

is partitioned into $2^{\tilde{R}+1} = 4$ cosets $\mathcal{D}_0$, $\mathcal{D}_1$, $\mathcal{D}_2$ and $\mathcal{D}_3$, each with 2^{R-1} points. Then by Proposition 1 in [17], we have

$$\mathrm{P}\left\{ \hat{Y}_{1,i} \in \mathcal{D}_c \middle| Y_{1,i} = y_{1,i} \right\} = \mathrm{P}\left\{ \hat{Y}_{1,i} \in \mathcal{D}_{(c+j) \bmod 4} \middle| Y_{1,i} = y_{1,i} + j\Delta_1 \right\}, \tag{4.8}$$

for $i = 0, 1, \ldots, n-1$, $c, j = 0, 1, 2, 3$, and $\left(-2^R + 1.5\right)\Delta_1 \le y_{1,i}, y_{1,i} + j\Delta_1 \le \left(2^R - 0.5\right)\Delta_1$. Denote the trellis bit vector of $\mathcal{Q}_1$ as $\boldsymbol{m}_1 = (m_{1,0}, m_{1,1}, \ldots, m_{1,n-1})^{\mathrm{T}}$, and the codeword vector $\boldsymbol{w}_1 = (w_{1,0}, w_{1,1}, \ldots, w_{1,n-1})^{\mathrm{T}}$. If we directly transmit the trellis bit vector $\boldsymbol{m}_1$ using 1 b/s (since $\tilde{R} = 1$) without SW coding, the practical rate will be

$$R_1 = 1 + \frac{1}{n}\sum_{i=0}^{n-1}\int_{-\frac{\Delta_1}{2}}^{\frac{\Delta_1}{2}} \frac{1}{\Delta_1} H\left(W_{1,i}\left|C_{1,i}, V_{1,i}\right.\right) \mathrm{d}v_{1,i}, \tag{4.9}$$

where $\boldsymbol{V}_1 = (V_{1,0}, V_{1,1}, \ldots, V_{1,n-1})^{\mathrm{T}}$ is a length-n vector of i.i.d. random dithers and $\boldsymbol{C}_1 = (C_{1,0}, C_{1,1}, \ldots, C_{1,n-1})^{\mathrm{T}}$ is the coset index vector.

Since the conditional distribution of $Y_{1,i}$ given $C_{1,i}$ and $V_{1,i}$ completely determines the conditional entropy $H\left(W_{1,i}\left|C_{1,i}, V_{1,i} = v_{1,i}\right.\right)$ in (4.9), we have for $i = 0, 1, \ldots, n-1$,

$$p\left(Y_{1,i} = y_{1,i}\left|C_{1,i} = c_{1,i}, V_{1,i} = v_{1,i}\right.\right) = \frac{p\left(Y_{1,i} = y_{1,i} + v_{1,i}\right) \cdot P\left(C_{1,i} = c_{1,i}\left|Y_{1,i} = y_{1,i}\right.\right)}{P\left(C_{1,i} = c_{1,i}\right)}. \tag{4.10}$$

Next we consider the WZ coding components that quantizes $Y_2^n, \ldots, Y_L^n$ and compresses the output $Y_{k,\mathrm{Q}}^n = \mathcal{Q}_k\left(Y_k^n\right)$ $(k = 2, 3, \ldots, L)$ to R_k b/s. Let the ESS step size of the employed TCQ be Δ_k, and the dequantized version

$$\left(\tilde{\boldsymbol{Y}}_1, \tilde{\boldsymbol{Y}}_2, \ldots, \tilde{\boldsymbol{Y}}_L\right) = \left(\tilde{Y}_1^n, \tilde{Y}_2^n, \ldots, \tilde{Y}_L^n\right) = \left(\boldsymbol{Y}_1 + \boldsymbol{Q}_1, \boldsymbol{Y}_2 + \boldsymbol{Q}_2, \ldots, \boldsymbol{Y}_L + \boldsymbol{Q}_L\right), \tag{4.11}$$

where

$$\boldsymbol{Q}_k = Q_k^n = \left(\boldsymbol{Y}_k + \boldsymbol{V}_k\right) - \mathcal{Q}_k^{-1}\left[\mathcal{Q}_k\left(\boldsymbol{Y}_k + \boldsymbol{V}_k\right)\right], \tag{4.12}$$

for $k = 1, 2, \ldots, L$ are zero-mean independent Gaussian random variables that are also independent of $Y_1, Y_2, \ldots, Y_L$. According to (4.5), similar to (4.9) and (4.10), for

$k = 2, 3, \ldots, L$, we have

$$\begin{aligned} R_k &= 1 + \frac{1}{n} H\left(\boldsymbol{W}_k \left| \boldsymbol{M}_k, \boldsymbol{V}_k, \boldsymbol{Z}_{k-1} \right.\right) \\ &\leq 1 + \frac{1}{n} \sum_{i=0}^{n-1} H\left(W_{k,i} \left| M_{k,i}, V_{k,i}, Z_{k-1,i} \right.\right) \\ &= 1 + \frac{1}{n} \sum_{i=0}^{n-1} \int_{-\frac{\Delta_k}{2}}^{\frac{\Delta_k}{2}} \frac{1}{\Delta_k} H\left(W_{k,i} \left| M_{k,i}, V_{k,i}, Z_{k-1,i} \right.\right) \mathrm{d}v_{k,i}, \end{aligned} \tag{4.13}$$

and

$$\begin{aligned} & p\left(Y_{k,i} = y_{k,i} \left| C_{k,i} = c_{k,i}, V_{k,i} = v_{k,i}, Z_{k-1,i} = z_{k-1,i} \right.\right) \\ &= p\left(Y_{k,i} = y_{k,i} + v_{k,i} \left| C_{k,i} = c_{k,i}, V_{1,i} = 0, Z_{k-1,i} = z_{k-1,i} \right.\right) \\ &= \frac{p\left(Y_{k,i} = y_{k,i} + v_{k,i} \left| Z_{k-1,i} = z_{k-1,i} \right.\right)}{P\left(C_{k,i} = c_{k,i} \left| Z_{k-1,i} = z_{k-1,i} \right.\right)} \cdot P\left(C_{k,i} = c_{k,i} \left| Y_{k,i} = y_{k,i} \right.\right) \end{aligned} \tag{4.14}$$

where $V_k^n = \{V_{k,1}, V_{k,2}, \ldots, V_{k,n-1}\}$ is a length-n vector of i.i.d. random dithers, and the last equation in (4.14) comes from Markov chain $Z_{k-1,i} \to Y_{k,i} \to C_{k,i}$.

In the case of high-rate transmission, we can assume that

$$\Delta_k \to 0, \quad k = 1, 2, \ldots, L. \tag{4.15}$$

Thus, we have

$$\begin{aligned} p\left(W_{1,i} = j \left| C_{1,i} = c_{1,i}, V_{1,i} = v_{1,i} \right.\right) &= p\left(Y_{1,i} + v_{1,i} \in \mathcal{W}_j \left| C_{1,i} = c_{1,i}, V_{1,i} = v_{1,i} \right.\right) \\ &\approx p\left(Y_{1,i} + v_{1,i} \in \mathcal{W}_j\right), \end{aligned} \tag{4.16}$$

where

$$\mathcal{W}_j = \left[\left(4j + c_{1,i} - 2^{R_1^Q}\right)\Delta_1, \left(4j + c_{1,i} - 2^{R_1^Q} + 1\right)\Delta_1\right]. \tag{4.17}$$

Then we have

$$\begin{aligned}
&\lim_{\Delta_1 \to 0} H\left(W_{1,i} \left| C_{1,i} = c_{1,i}, V_{1,i} = v_{1,i} \right.\right) \\
&= - \lim_{\Delta_1 \to 0} \sum_{j=0}^{2^R - 1} \left(p\left(W_{1,i} \left| C_{1,i} = c_{1,i}, V_{1,i} = v_{1,i} \right.\right) \cdot \log p\left(W_{1,i} \left| C_{1,i} = c_{1,i}, V_{1,i} = v_{1,i} \right.\right)\right) \\
&= \lim_{\Delta_1 \to 0} \left(h\left(Y_{1,i} + v_{1,i}\right) - \log\left(4\Delta_1\right)\right) = h\left(Y_{1,i}\right) - \log\left(4\Delta_1\right). \qquad (4.18)
\end{aligned}$$

Similarly, for $k = 2, 3, \ldots, L-1$,

$$\begin{aligned}
&p\left(W_{k,i} = j \left| C_{k,i} = c_{k,i}, V_{k,i} = v_{k,i}, Z_{k-1,i} = z_{k-1,i} \right.\right) \\
&= p\left(Y_{k,i} + v_{k,i} \in \mathcal{W}_j \left| C_{k,i} = c_{k,i}, V_{k,i} = v_{k,i}, Z_{k-1,i} = z_{k-1,i} \right.\right) \\
&= \int_{\mathcal{W}_j} \frac{p\left(Y_{k,i} + v_{k,i} = \tau \left| Z_{k-1,i} = z_{k-1,i} \right.\right) \cdot P\left(C_{k,i} = c_{k,i} \left| Y_{k,i} = \tau \right.\right)}{P\left(C_{k,i} = c_{k,i} \left| Z_{k-1,i} = z_{k-1,i} \right.\right)} \mathrm{d}\tau \\
&\approx p\left(Y_{k,i} + v_{k,i} \in \mathcal{W}_j \left| Z_{k-1,i} = z_{k-1,i} \right.\right), \qquad (4.19)
\end{aligned}$$

where τ^* is some value of τ in $\mathcal{W}_j$, and then

$$\begin{aligned}
&\lim_{\Delta_k \to 0} H\left(W_{k,i} \left| C_{k,i}, Z_{k-1,i}, V_{k,i} = v_{k,i} \right.\right) \\
&= - \lim_{\Delta_k \to 0} \int_{\mathbb{R}} \Big[\sum_{j=0}^{2^R - 1} p\left(Y_{k,i} + v_{k,i} \in \mathcal{W}_j \left| Z_{k-1,i} = z_{k-1,i} \right.\right) \\
&\quad \cdot \log p\left(Y_{k,i} + v_{k,i} \in \mathcal{W}_j \left| Z_{k-1,i} = z_{k-1,i} \right.\right) \Big] \mathrm{d}z_{k-1,i} \\
&= \lim_{\Delta_k \to 0} h\left(Y_{k,i} + v_{k,i} \left| Z_{k-1,i} \right.\right) - \log\left(4\Delta_k\right) \\
&= h\left(Y_{k,i} \left| Z_{k-1,i} \right.\right) - \log\left(4\Delta_k\right). \qquad (4.20)
\end{aligned}$$

If we assume ideal SW coding, the distortion can be written as

$$d_k = \frac{1}{n} \mathrm{E}\left[\left\|Q_k^n\right\|_2^2\right] = \frac{1}{n} \mathrm{E}\left[\left\|\tilde{Y}_k^n - Y_k^n\right\|_2^2\right] = \mathcal{V}_k^{\frac{2}{n}} G_{\mathcal{Q}_k} = \left(2\Delta_k\right)^2 G_{\mathcal{Q}_k}, \quad k = 1, 2, \ldots, L, \qquad (4.21)$$

where $\mathcal{V}_k$ is the volume of the Voronoi region of the current quantizer $\mathcal{Q}_k$. Therefore, we can proceed by

$$\begin{aligned}
R_1 &= \lim_{\Delta_1 \to 0} \left(1 + \frac{1}{n} \sum_{i=0}^{n-1} \int_{-\frac{\Delta_1}{2}}^{\frac{\Delta_1}{2}} \frac{1}{\Delta_1} H\left(W_{1,i} \left| C_{1,i}, V_{1,i} = v_{1,i}\right.\right) \mathrm{d}v_{1,i} \right) \\
&= \lim_{\Delta_1 \to 0} 1 + \frac{1}{n} \left(\sum_{i=0}^{n-1} h\left(Y_{1,i}\right) - \log\left(4\Delta_1\right) \right) \\
&= 1 + \frac{1}{2} \log\left(2\pi \mathrm{e} \sigma_Y^2\right) - \log\left(2\sqrt{\frac{d_1}{G_{\mathcal{Q}_1}}} \right) \\
&= R_1^* + \frac{1}{2} \log\left(2\pi \mathrm{e} G_{\mathcal{Q}_1}\right), \qquad (4.22)
\end{aligned}$$

where $\sigma_{\boldsymbol{Y}_1}^2$ is the variance of $\boldsymbol{Y}_1$, and for $k = 2, 3, \ldots, L$,

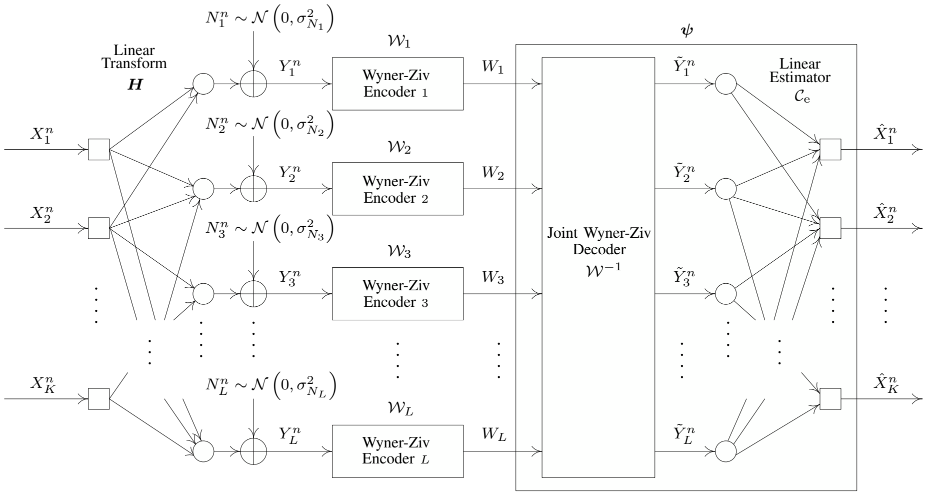
$$\begin{aligned}
R_k &= \lim_{\Delta_k \to 0} \left(1 + \frac{1}{n} \sum_{i=0}^{n-1} \int_{-\frac{\Delta_k}{2}}^{\frac{\Delta_k}{2}} \frac{1}{\Delta_k} H\left(W_{k,i} \left| C_{k,i}, Z_{k-1,i}, V_{k,i} = v_{k,i}\right.\right) \mathrm{d}v_{k,i} \right) \\
&= \frac{1}{2} \log\left(\frac{\sigma_{\boldsymbol{Y}_k | \boldsymbol{Z}_{k-1}}^2}{d_k} \right) + \frac{1}{2} \log\left(2\pi \mathrm{e} G_{\mathcal{Q}_k}\right) \\
&= R_k^* + \frac{1}{2} \log\left(2\pi \mathrm{e} G_{\mathcal{Q}_k}\right), \qquad (4.23)
\end{aligned}$$

where $\sigma_{\boldsymbol{Y}_k | \boldsymbol{Z}_{k-1}}^2$ is the variance of $\boldsymbol{Y}_k$ given $\boldsymbol{Z}_{k-1}$.

Finally, the theorem is proved by adding together (4.22) and (4.23). □

B. Correlation channel modeling

Due to the use of TCQ/TCVQ, the bit-plane-wise correlation channel between the quantized source $Y_{i,\mathrm{Q}}$ and the decoder side information Z_{i-1}^{SW} is not Gaussian. In addition, the correlation channel between any pair of quantized sources is not Gaussian either. We mathematically model the bit-plane-wise correlation channel between $Y_{i,\mathrm{Q}}$ and Z_{i-1}^{SW} to facilitate the design of LDPC profiles for SW compression. Assume we know Z_{i-1}^{SW} and the distribution $p\left(Y_i \left| Z_{i-1}^{\mathrm{SW}}\right.\right)$, and the trellis bit b_0 and the first $j-1$



$N_1^n \sim \mathcal{N}\left(0, \sigma_{N_1}^2\right)$
$\mathcal{W}_1$
$\boldsymbol{\psi}$
Linear Transform $\boldsymbol{H}$
Y_1^n
Wyner-Ziv Encoder 1
W_1
$\tilde{Y}_1^n$
Linear Estimator $\mathcal{C}_{\mathrm{e}}$
X_1^n
$\hat{X}_1^n$
$N_2^n \sim \mathcal{N}\left(0, \sigma_{N_2}^2\right)$
$\mathcal{W}_2$
Y_2^n
Wyner-Ziv Encoder 2
W_2
$\tilde{Y}_2^n$
X_2^n
$\hat{X}_2^n$
$N_3^n \sim \mathcal{N}\left(0, \sigma_{N_3}^2\right)$
Joint Wyner-Ziv Decoder $\mathcal{W}^{-1}$
$\mathcal{W}_3$
Y_3^n
Wyner-Ziv Encoder 3
W_3
$\tilde{Y}_3^n$
X_K^n
$N_L^n \sim \mathcal{N}\left(0, \sigma_{N_L}^2\right)$
$\hat{X}_K^n$
$\mathcal{W}_L$
Y_L^n
Wyner-Ziv Encoder L
W_L
$\tilde{Y}_L^n$

Fig. 13. Block diagram of the proposed coding scheme for the generalized Gaussian CEO problem. The Gaussian CEO problem corresponds to the case with $K = 1$.

bit planes $b_1, \ldots, b_{j-1}$ of $Y_{i,\mathrm{Q}}$ have been decoded. For simplicity and without loss of generality, we also assume that the quantization step $\Delta = 1$ (otherwise the sources can always be scaled up or down). Since we are using a rate-1/2 TCQ, b_0 is decoded into a coset $c_0 \in \{0, 1, 2, 3\}$, the centers of quantization cells for the j-th bit $b_j = 1$ are $\left\{m_j + 2^{j+2}u \,\middle|\, u = 0, 1, \ldots, 2^{R+1-j} - 1\right\}$, and those for $b_j = 0$ are shifted by 2^{j+1}, where m_j is the shift of $b_0, \ldots, b_{j-1}$ and c_0 and can be written as

$$m_j = -2^R + \frac{1}{2} + c_0 + \sum_{k=1}^{j-1} 2^{k+1} b_k. \tag{4.24}$$

We approximate the conditional probability distribution function (p.d.f.) $p(y|c_0, b_1, \ldots, b_R)$ as $p_{\mathrm{G}}(y - m^R)$, with

$$p_{\mathrm{G}}(y) \triangleq \begin{cases} \dfrac{\mathrm{e}^{-\frac{\pi \mathrm{e} y^2}{4}}}{\int_{-2}^{2} \mathrm{e}^{-\frac{\pi \mathrm{e} x^2}{4}} \mathrm{d}x}, & y \in [-2, 2] \\ 0, & \text{otherwise.} \end{cases}, \tag{4.25}$$

which is obtained by bounding the ideal Gaussian quantization noise of variance $\frac{2}{\pi \mathrm{e}}$

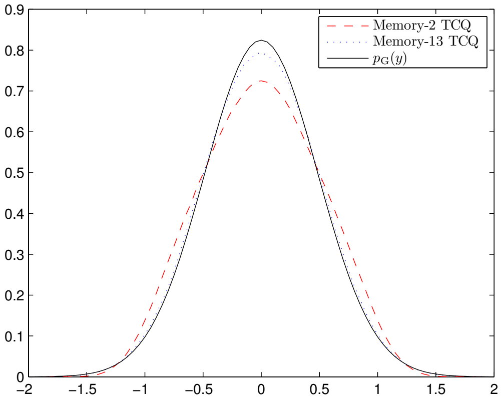


Fig. 14. Quantization error distributions of memory-2 TCQ and memory-13 TCQ as compared to the approximated distribution $p_{\mathrm{G}}(y)$.

within the range of $(-2, 2)$. This approximation becomes more accurate as the TCQ memory size increases, as shown in Fig. 14. It can also be seen from Fig. 15, where the relative entropy between the TCQ quantization error distributions and the approximated distribution decreases as the TCQ memory increases. Then, assuming the TCQ rate is high enough such that var $\left(Y_i \,\middle|\, Z_{i-1}^{\mathrm{SW}}\right) \gg \Delta = 1$, the conditional p.d.f. superposition of shifted copies of $p_{\mathrm{G}}(y)$, each centered at $\left\{-2^R + m_j + 2^{j+1} + 2^{j+2}u \,\middle|\, u = 0, 1, \ldots, 2^{R+1-j} - 1\right\}$ i.e.,

$$p(y|c_0, b_1, \ldots, b_{j-1}, b_j = 1) = \frac{\sum_{u=0}^{2^{R+1-j}-1} p_{\mathrm{G}}\left(y - m_j - 2^{j+2}u\right)}{\int_{-\infty}^{\infty} \sum_{u=0}^{2^{R+1-j}-1} p_{\mathrm{G}}\left(x - m_j - 2^{j+2}u\right) \mathrm{d}x}. \qquad (4.26)$$

Similarly, we have

$$p(y|c_0, b_1, \ldots, b_{j-1}, b_j = 0) = \frac{\sum_{u=0}^{2^{R+1-j}-1} p_{\mathrm{G}}\left(y - m_j - 2^{j+1} - 2^{j+2}u\right)}{\int_{-\infty}^{\infty} \sum_{u=0}^{2^{R+1-j}-1} p_{\mathrm{G}}\left(x - m_j - 2^{j+1} - 2^{j+2}u\right) \mathrm{d}x}. \qquad (4.27)$$

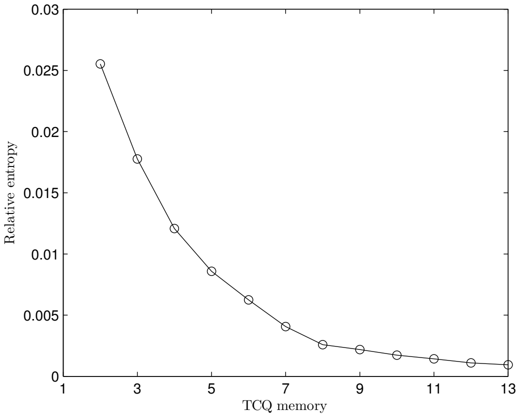


Fig. 15. The relative entropy between the TCQ quantization error distribution (of different memory sizes) and the approximated distribution $p_{\mathrm{G}}(y)$.

On the other hand,

$$\begin{aligned}\frac{\mathrm{P}\left\{b_j = 1 \left| Z_{i-1}^{\mathrm{SW}}, c_0, b_1, \ldots, b_{j-1}\right.\right\}}{\mathrm{P}\left\{b_j = 0 \left| Z_{i-1}^{\mathrm{SW}}, c_0, b_1, \ldots, b_{j-1}\right.\right\}} &= \frac{\int_{-\infty}^{\infty} p\left(y \left| Z_{i-1}^{\mathrm{SW}}, c_0, b_1, \ldots, b_{j-1}, b_j = 1\right.\right) \mathrm{d}y}{\int_{-\infty}^{\infty} p\left(y \left| Z_{i-1}^{\mathrm{SW}}, c_0, b_1, \ldots, b_{j-1}, b_j = 0\right.\right) \mathrm{d}y} \\ &= \frac{\int_{-\infty}^{\infty} p\left(y \left| Z_{i-1}^{\mathrm{SW}}\right.\right) p\left(y \left| c_0, b_1, \ldots, b_{j-1}, b_j = 1\right.\right) \mathrm{d}y}{\int_{-\infty}^{\infty} p\left(y \left| Z_{i-1}^{\mathrm{SW}}\right.\right) p\left(y \left| c_0, b_1, \ldots, b_{j-1}, b_j = 0\right.\right) \mathrm{d}y},\end{aligned} \tag{4.28}$$

with (4.28) being true when $p\left(y \left| Z_{i-1}^{\mathrm{SW}}\right.\right)$ and $p\left(y \left| c_0, b_1, \ldots, b_j\right.\right)$ are independent, and $p\left(Z_{i-1}^{\mathrm{SW}}\right)$ and $\mathrm{P}\left(c_0, b_1, \ldots, b_j\right)$ are independent as well, which holds when the rate of TCQ is high. This means that under the high-rate assumption, $p\left(y \left| Z_{i-1}^{\mathrm{SW}}, c_0, b_1, \ldots, b_j\right.\right)$ can be considered as $p(y|c_0, b_1, \ldots, b_j)$ enveloped by $p\left(y \left| Z_{i-1}^{\mathrm{SW}}\right.\right)$, whose variance is determined by $\boldsymbol{\Sigma_Y}$, $\boldsymbol{D}$, and the source encoding order. An example is plotted in Fig. 16, from which it can be seen that the statistical distribution and its approximation are almost identical, both enveloped by the conditional distribution $p\left(y \left| Z_{i-1}^{\mathrm{SW}}\right.\right)$.

Therefore, using (4.26), (4.27), and (4.28), the log-likelihood ratio (LLR) of the

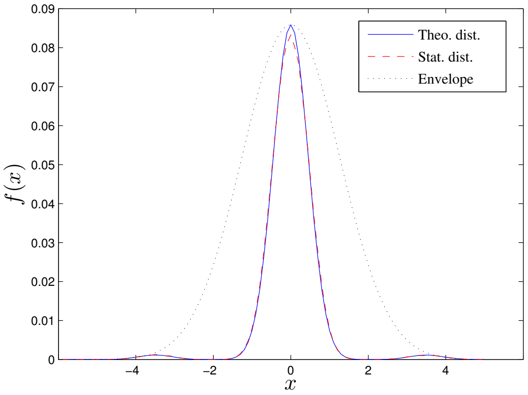


Fig. 16. The conditional distribution $p\left(y\left|Z_{i-1}^{\mathrm{SW}}, c_0, b_1, \ldots, b_{j-1}, b_j = 1\right.\right)$ and its approximation of the first WZ coder (second bit plane of the symmetric case with $\rho = 0.8$ and $D = 0.05$).

j-th bit plane when $Z_{i-1}^{\mathrm{SW}} = a$ can be calculated by

$$\mathrm{LLR}(a, c_0, \ldots, b_{j-1}, 1) = \log \frac{\mathrm{P}\left\{b_j = 1\left|Z_{i-1}^{\mathrm{SW}} = a, c_0, b_1, \ldots, b_{j-1}\right.\right\}}{\mathrm{P}\left\{b_j = 0\left|Z_{i-1}^{\mathrm{SW}} = a, c_0, b_1, \ldots, b_{j-1}\right.\right\}}, \tag{4.29}$$

and

$$\mathrm{LLR}(a, c_0, \ldots, b_{j-1}, 0) = \log \frac{\mathrm{P}\left\{b_j = 0\left|Z_{i-1}^{\mathrm{SW}} = a, c_0, b_1, \ldots, b_{j-1}\right.\right\}}{\mathrm{P}\left\{b_j = 1\left|Z_{i-1}^{\mathrm{SW}} = a, c_0, b_1, \ldots, b_{j-1}\right.\right\}}. \tag{4.30}$$

Then, by going over the range of Z_{i-1}^{SW}, c_0, and $b_1, \ldots, b_{j-1}$ to calculate different LLR values, we can get $p_{\mathrm{LLR}, Z_{i-1}^{\mathrm{SW}}, c_0, b_1, \ldots, b_j}(\cdot)$, which is the approximate p.d.f. of the joint distribution of LLR, Z_{i-1}^{SW}, c_0, and $b_1, \ldots, b_j$. Then the approximate LLR distribution $f(l)$ can be calculated by

$$f(l) = \sum_{c_0, b_1, \ldots, b_{j-1}} \int_{-\infty}^{\infty} p_{\mathrm{LLR}, Z_{i-1}^{\mathrm{SW}}, c_0, b_1, \ldots, b_j}\left(\mathrm{LLR} = l, Z_{i-1}^{\mathrm{SW}} = a, c_0, b_1, \ldots, b_{j-1}\right) da. \tag{4.31}$$

Note that this distribution is an average given $b_j = 1$ and $b_j = 0$, where the j-th bit can be 1 or 0 with equal probability. Since the bit-plane-wise correlation channel is symmetric, the conditional LLR distributions given $b_j = 1$ or $b_j = 0$, denoted as $f(l|1)$ and $f(l|0)$, respectively, satisfy $f(l|1) = \mathrm{e}^l f(l|0)$ [51], thus $f(l|1)$ and $f(l|0)$ can be written as

$$f(l|1) = \frac{\mathrm{e}^l}{1+\mathrm{e}^l} f(l), \quad f(l|0) = \frac{1}{1+\mathrm{e}^l} f(l). \tag{4.32}$$

An example of $f(l|1)$ v.s. $f(l|0)$ is shown in Fig. 17, from which we can see that the theoretical/approximate LLR distribution is almost identical to the practical one. LDPC code designs can be carried out by using the approximate LLR distribution instead of the practical one acquired from training data. Compared to practical training based design, our design based on the approximate distribution suffers no additional rate loss in our simulations.

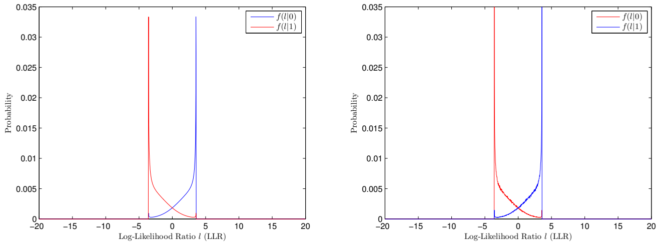


Fig. 17. The theoretical (left) and practical (right) LLR distribution of the first WZ coder (second bit plane of the symmetric case with $\rho = 0.8$ and $D = 0.05$).

Since the conditional Gaussian distribution $p\left(Y_i \middle| Z_{i-1}^{\mathrm{SW}}\right)$ can be estimated by $\boldsymbol{\Sigma_Y}$ and $\boldsymbol{D}$, the LLR distribution of each WZ coded bit plane at each source terminal can be pre-calculated if $\boldsymbol{\Sigma_Y}$ and $\boldsymbol{D}$ are given. It can be seen that the LLR distribution, and hence the LDPC code rate are only determined by the variance of the distribution $p\left(Y_i \middle| Z_{i-1}^{\mathrm{SW}}\right)$ and the bit plane position. Therefore, a library of LDPC profiles can be built up off-line and the code rate can vary from 0 to 1. Then for any quadratic Gaussian MT problem, the LDPC profile for each bit plane of the WZ coded sources can be determined by looking up the library. Fig. 18 depicts the average right profile degree and rate loss to theoretical capacity for different LDPC rates in our LDPC code library.

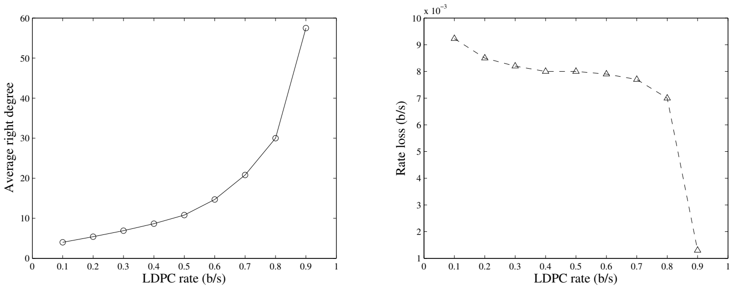


Fig. 18. The average right profile degree and rate loss to capacity of the offline designed LDPC code library.

Instead of the above approximation method, the correlation model between quantized sources can acquired by data training, i.e., encoding several training blocks of data to get an approximate correlation distribution before encoding and decoding the actual sources [20]. This approach provides no additional coding gain compared to our proposed approximation method, while in the meantime obviously causing delays to the entire coding process.

C. Experimental results

We present our experimental results in accordance with the theoretical reviews of tight BT bounds in Section 2, Chapter II. Three- and four-terminal cases are considered in our simulations. In the positive symmetric setup with different ρ and D values, coding results in both high-rate and low-rate regimes are given. The block length is fixed at 10^6 bits, and the bit error rate in SW decoding ranges from 10^{-7} to 4×10^{-6}. Cases with more than four terminals can be achieved using similar coding schemes and performance similar to the three- and four-terminal cases is expected. Detailed results are given in the following subsections.

1. Quadratic Gaussian direct MT coding

a. The positive symmetric case

In this setup, the sources are zero mean, jointly Gaussian with identical variance 1 and correlation coefficient $\rho = 0.80$. We consider three- and four-terminal cases.

High-rate scenario (with identical target distortion $D = 0.05$): In this case, the experimental result of our practical code design is given in Table II. It is seen that after meeting the target distortion D, our practical SWCQ design only suffers a small rate loss of about 0.05 b/s at each terminal.

Table II. Ideal and practical corner point coding rates (in b/s) using TCQ quantizer in high-rate scenarios.

	Three terminal			Four terminal			
	#1	#2	#3	#1	#2	#3	#4
Ideal rate	2.077	1.468	1.320	2.061	1.454	1.307	1.239
Practical rate	2.102	1.527	1.398	2.086	1.510	1.385	1.331

We employ 8192-state TCQ source encoder and irregular LDPC code to approach the MT sum-rate bound. Table III lists bit-plane level conditional entropies and the practical LDPC code profiles for the three-terminal coding that approaches a corner point − the ideal corner points can be calculated according to the method in Section 3. It is seen that the sum-rate loss due to practical coding is 0.162 b/s. The practical sum-rate region as compared to the MT sum-rate bound is depicted in Fig. 19.

Low rate scenario (with identical target distortion $D = 0.10$): For relatively lower transmission rate, e.g., 1 b/s or lower, we use SWC-TCVQ with 8192-state trellis to reduce trellis bit rate. The rest of the low-rate scheme stays the same as in the high-rate scenario. The quantizer choices and the ideal/practical rates are given

Table III. Entropies versus practical rates for MT source coding approaching a corner point using TCQ and SW coding, together with the LDPC code profiles used for SW compression. The correlation coefficient is $\rho = 0.80$ and target distortion is $D = 0.05$. Bit planes not transmitted are omitted in the table.

Component	Bit Plane #	Conditional Entropy (b/s)	Practical Rate (b/s)	Irregular LDPC Code Profile (Edge Perspective)
Y_1	All	2.077	2.102	–
Y_2	1	1.000*	1.000	–
	2	0.499	0.507	$\lambda(x) = 0.1413x + 0.2229x^2 + 0.0129x^5 + 0.0879x^6 + 0.0560x^9 + 0.0220x^{10} + 0.0300x^{11} + 0.0023x^{12} + 0.0648x^{13} + 0.0174x^{14} + 0.0168x^{17} + 0.0212x^{18} + 0.0018x^{34} + 0.0413x^{35} + 0.0134x^{46} + 0.0059x^{47} + 0.0448x^{48} + 0.0676x^{49} + 0.0567x^{98} + 0.0732x^{99}$; $\rho(x) = 0.2000x^9 + 0.8000x^{10}$.
	3	0.014	0.020	$\lambda(x) = 0.0070x + 0.3537x^2 + 0.0285x^5 + 0.0622x^6 + 0.0140x^9 + 0.2723x^{10} + 0.0077x^{26} + 0.0183x^{27} + 0.1404x^{45} + 0.0280x^{46} + 0.0177x^{62} + 0.0183x^{63} + 0.0319x^{99}$; $\rho(x) = x^{299}$.
	All	1.513	1.527	–
Y_3	1	1.000*	1.000	–
	2	0.380	0.388	$\lambda(x) = 0.1019x + 0.2589x^2 + 0.0056x^5 + 0.0321x^6 + 0010x^7 + 0.0008x^8 + 0.1053x^9 + 0.0870x^{10} + 0.1319x^{21} + 0.0190x^{22} + 0.0146x^{37} + 0.0176x^{38} + 0.0066x^{44} + 0.0317x^{45} + 0.0117x^{96} + 0.1741x^{97}$; $\rho(x) = x^{14}$.
	3	0.005	0.010	$\lambda(x) = 0.0974x + 0.2474x^2 + 0.0093x^5 + 0.2864x^7 + 0.0339x^{19} + 0.0156x^{20} + 0.0090x^{21} + 0.1339x^{22} + 0.1356x^{34} + 0.0216x^{40} + 0.0099x^{42}$; $\rho(x) = x^{549}$.
	All	1.385	1.398	–
Total	–	4.975	5.027	–

* We directly compute the conditional entropy of the trellis bit plane assuming it is memoryless given the side information.

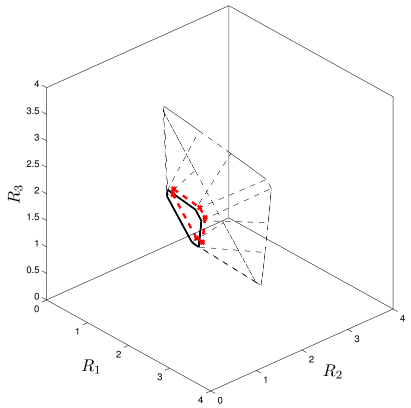


Fig. 19. The BT rate region v.s. practical sum-rate bound for the three-terminal symmetric case with $\rho = 0.80$ and $D = 0.05$. The practical sum-rate bound is enclosed by the dashed line with crosses.

in Table IV.

From Table IV we see that compared with the results in high-rate scenario in Table III, the relative rate loss for each terminal is higher than those in the high-rate scenario. This is partially due to the higher granular loss of TCVQ. However, SWC-TCVQ is much more efficient in the low-rate scenario than SWC-TCQ, since the rate loss is no higher than 0.077 b/s even for a transmission rate as low as 0.797 b/s, compared to the rate loss of 0.078 b/s for a transmission rate of 1.320 b/s when employing the SWC-TCQ scheme as shown in Table III.

Table IV. Ideal and practical corner point coding rates (in b/s) and quantizer choice in low-rate scenarios.

	Three terminal			Four terminal			
	#1	#2	#3	#1	#2	#3	#4
Ideal rate	1.506	1.019	0.884	1.471	0.993	0.860	0.797
Practical rate	1.546	1.073	0.950	1.512	1.049	0.929	0.874
Quantizer	TCQ	TCVQ	TCVQ	TCQ	TCVQ	TCVQ	TCVQ

b. The BEEV-ED case

For the BEEV-ED setup, we designed a coding scheme in the high-rate scenario for a four-terminal scenario with source covariance matrix

$$\mathbf{\Sigma}_{\boldsymbol{Y}} = \begin{pmatrix} 1.0 & 0.3 & 0.4 & 0.0 \\ 0.3 & 1.0 & 0.0 & -0.4 \\ 0.4 & 0.0 & 1.0 & 0.3 \\ 0.0 & -0.4 & 0.3 & 1.0 \end{pmatrix}, \tag{4.33}$$

and uniform target distortion $D = 0.05$. It can be verified that the two distinct eigenvalues of $\mathbf{\Sigma}_{\boldsymbol{Y}}$ are $\Lambda = 1.5$ and $\lambda = 0.5$ (each repeated twice) and the BT sum-rate bound is tight. A corner point on the theoretical BT bound is calculated with (4.2) as

$$(R_1 \;\; R_2 \;\; R_3 \;\; R_4) = (2.150 \;\; 2.088 \;\; 2.027 \;\; 1.965) \;\; \text{b/s}, \tag{4.34}$$

with a sum-rate of $R_{\boldsymbol{Y}}(\mathbf{\Sigma}_{\boldsymbol{Y}}, D) = 8.230$ b/s. The practical encoding rates in our SWC-TCQ design are

$$(R_1^* \;\; R_2^* \;\; R_3^* \;\; R_4^*) = (2.173 \;\; 2.130 \;\; 2.071 \;\; 2.010) \;\; \text{b/s} \tag{4.35}$$

when the target distortions are met.

c. A general nonsymmetric setup

According to the sufficient condition [13], we designed a coding scheme in high-rate scenario for a three-terminal scenario the following setup.

$$\mathbf{\Sigma}_{\boldsymbol{Y}} = \begin{pmatrix} 1.0 & 0.7 & 0.8 \\ 0.7 & 1.0 & 0.9 \\ 0.8 & 0.9 & 1.0 \end{pmatrix}, \tag{4.36}$$

and target distortion

$$(D_1 \;\; D_2 \;\; D_3) = (0.030 \;\; 0.025 \;\; 0.020)\,. \tag{4.37}$$

It is easy to verify that $\mathbf{\Sigma}_{\boldsymbol{Y}}$ and $\boldsymbol{D}$ satisfy the sufficient condition for tight BT sum-rate bound. A corner point on the theoretical BT sum-rate bound is calculated with (2.10) and (4.2) as

$$(R_1 \;\; R_2 \;\; R_3) = (2.492 \;\; 2.143 \;\; 1.450) \;\; \text{b/s}, \tag{4.38}$$

with a sum-rate of $R_{\boldsymbol{Y}}(\mathbf{\Sigma}_{\boldsymbol{Y}}, D) = 6.085$ b/s. The practical encoding rates in our SWC-TCQ design are

$$(R_1^* \;\; R_2^* \;\; R_3^*) = (2.513 \;\; 2.188 \;\; 1.513) \;\; \text{b/s}, \tag{4.39}$$

when the target distortions are met.

2. Quadratic Gaussian indirect MT coding

a. The Gaussian CEO case

We present results for two Gaussian CEO cases with $L = 3$ and $L = 4$, respectively. The remote source is set to be $X \sim \mathcal{N}(0, 0.80)$ with i.i.d. observation noise variance

$\sigma_N^2 = 0.20$ for both cases, with target distortion $D = 0.07808$ when $L = 3$ and $D = 0.06032$ when $L = 4$.

For $L = 3$, a corner point on the tight theoretical BT sum-rate bound is calculated with (2.10) and (4.2) as

$$(R_1 \;\; R_2 \;\; R_3) = (2.076 \;\; 1.468 \;\; 1.320) \;\; \text{b/s}, \tag{4.40}$$

with a sum-rate of $R_X(\boldsymbol{\Sigma}_{\boldsymbol{N}_{\mathcal{L}}}, D) = 4.864$ b/s. The practical encoding rates in our SWC-TCQ design are

$$(R_1^* \;\; R_2^* \;\; R_3^*) = (2.102 \;\; 1.527 \;\; 1.398) \;\; \text{b/s}, \tag{4.41}$$

when the target distortion is met.

For $L = 4$, a corner point on the tight theoretical BT sum-rate bound is calculated with (2.10) and (4.2) as

$$(R_1 \;\; R_2 \;\; R_3 \;\; R_4) = (2.061 \;\; 1.454 \;\; 1.307 \;\; 1.239) \;\; \text{b/s}, \tag{4.42}$$

with a sum-rate of $R_X\left(\sigma_X^2, \boldsymbol{\Sigma}_{\boldsymbol{N}_{\mathcal{L}}}, D\right) = 6.061$ b/s. The practical encoding rates in our SWC-TCQ design are

$$(R_1^* \;\; R_2^* \;\; R_3^* \;\; R_4^*) = (2.086 \;\; 1.510 \;\; 1.385 \;\; 1.331) \;\; \text{b/s}, \tag{4.43}$$

when the target distortion is met.

b. The generalized Gaussian CEO case

For $L = 3$, we use the example defined by (2.33), (2.34) and (2.35). In this case, the observation covariance matrix is

$$\mathbf{\Sigma}_{\boldsymbol{Y}} = \begin{pmatrix} 0.8333 & 0.3333 & 0.3333 \\ 0.3333 & 1.9333 & -0.6667 \\ 0.3333 & -0.6667 & 2.0333 \end{pmatrix}. \tag{4.44}$$

A corner point on the theoretical BT bound is calculated with (2.10) and (4.2) as

$$(R_1 \;\; R_2 \;\; R_3) = (2.279 \;\; 3.444 \;\; 3.225) \;\; \text{b/s}, \tag{4.45}$$

and the sum-rate is $R_{\boldsymbol{X}_{\mathcal{K}}}(\mathbb{T}, D) = 8.948$ b/s. The practical encoding rates in our SWC-TCQ design are

$$(R_1^* \;\; R_2^* \;\; R_3^*) = (2.302 \;\; 3.496 \;\; 3.271) \;\; \text{b/s}, \tag{4.46}$$

when the target sum-distortion is met. The practical sum-rate bound as compared to the theoretical one is shown in Fig. 20.

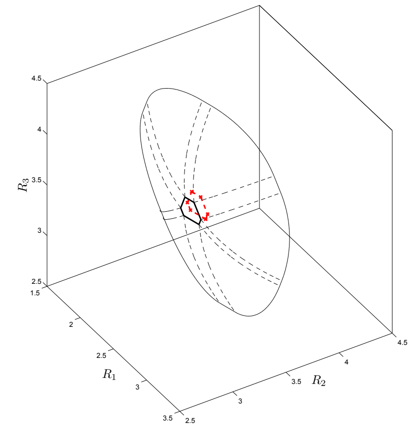


Fig. 20. The theoretical rate region v.s. practical sum-rate bound for the generalized CEO case defined by (2.33), (2.34) and (2.35). The practical sum-rate bound is enclosed by the dashed line with crosses.

For $L = 4$, we use $K = L - 1 = 3$ Gaussian i.i.d. remote sources $X_1, X_2, X_3 \sim \mathcal{N}(0, 1.0000)$ with i.i.d. observation noises $N_1, \ldots, N_4 \sim \mathcal{N}(0, 0.2500)$ and target sum-distortion $D = 0.6193$. The transform matrix is

$$\boldsymbol{H} = \begin{pmatrix} 0.0000 & 0.5000 & 0.7071 \\ 0.7071 & -0.5000 & 0.0000 \\ 0.0000 & 0.5000 & -0.7071 \\ -0.7071 & -0.5000 & 0.0000 \end{pmatrix}, \tag{4.47}$$

yielding an observation covariance matrix

$$\boldsymbol{\Sigma_Y} = \begin{pmatrix} 1.0000 & -0.2500 & -0.2500 & -0.2500 \\ -0.2500 & 1.0000 & -0.2500 & -0.2500 \\ -0.2500 & -0.2500 & 1.0000 & -0.2500 \\ -0.2500 & -0.2500 & -0.2500 & 1.0000 \end{pmatrix}. \tag{4.48}$$

A corner point on the theoretical BT sum-rate bound is calculated from (2.10) and (4.2) as

$$(R_1 \;\; R_2 \;\; R_3 \;\; R_4) = (3.318 \;\; 3.272 \;\; 3.190 \;\; 2.991) \text{ b/s}, \tag{4.49}$$

and the tight sum-rate is $R_{\boldsymbol{X}_{\mathcal{K}}}(\mathbb{T}, D) = 12.771$ b/s. The practical encoding rates in our SWC-TCQ design are

$$(R_1^* \;\; R_2^* \;\; R_3^* \;\; R_4^*) = (3.339 \;\; 3.318 \;\; 3.234 \;\; 3.034) \text{ b/s}, \tag{4.50}$$

when the target sum-distortion is met.

CHAPTER V

MT VIDEO CODING

Following the theoretical analysis and code design for ideal sources, the practical application on video compression is investigated in this chapter. Section A illustrates the detailed MT video coding scheme without depth information transmitted to decoder, including side information generation, SW code design and decoder side joint estimation, Section B describes the usage of separately transmitted depth information in MT video coding, Section C gives the experiment results on depth camera assisted MT video coding.

A. MT video coding without depth camera assistance

In this section, we provide detailed description of our MT video coding scheme without depth information at the decoder. The whole scheme is implemented under the H.264/AVC framework. Generally, we follow the MT source coding scheme for quadratic Gaussian sources in Chapter IV, which proves to approach the theoretical achievable sum rate asymptotically, and the block diagram of the scheme is shown in Fig. 21.

Given equal distortion measure $D_i = D$, $i = 1, \ldots, L$ each texture camera sequence $\mathcal{H}_i$, $i = 1 \ldots, L$, is encoded separately and transmitted to the decoder end. The first texture sequence $\mathcal{H}_1$ is encoded using the original H.264/AVC scheme (this can be considered as an entropy coding scheme), and other sequences $\mathcal{H}_i$, $i = 2, \ldots, L$ are WZ encoded under the H.264/AVC framework, assuming that side information $\mathrm{SI}_1, \ldots, \mathrm{SI}_{(i-1)}$ are generated from previous $i - 1$ sequences.

Therefore, if we assume that WZ coding approaches conditional entropy, the

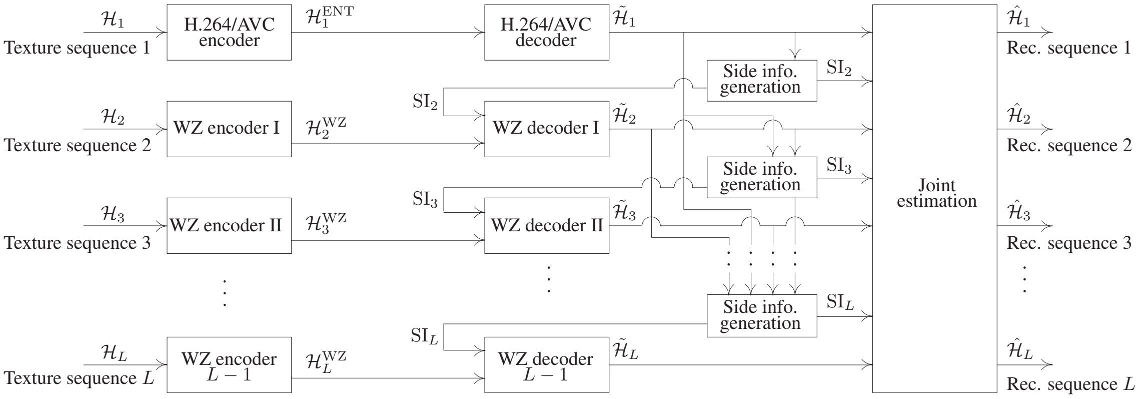


Fig. 21. The block diagram of the proposed MT video coding scheme.

total transmission rate $R_{\mathrm{T}}^{\mathrm{MT}}(D)$ of the MT coding scheme can be written as

$$R_{\mathrm{T}}^{\mathrm{MT}}(D) = \sum_{i=1}^{L} R_{\mathrm{T}_i}^{\mathrm{MT}}(\mathcal{H}_i, D) \tag{5.1}$$

$$= H\left(\hat{\mathcal{H}}_1(D)\right) + \sum_{i=2}^{L} H\left(\hat{\mathcal{H}}_i(D)\,\middle|\,\hat{\mathcal{H}}_1(D), \ldots, \hat{\mathcal{H}}_{i-1}(D)\right), \tag{5.2}$$

where $R_{\mathrm{T}_i}^{\mathrm{MT}}(\mathcal{H}_i, D)$ is the rate of texture sequence $\mathcal{H}_i$ given distortion measure D, and $\hat{\mathcal{H}}_i(D)$ is the i-th reconstructed sequence given distortion D. It can be seen that compared to joint video coding scheme, whose rate $R_{\mathrm{T}}^{\mathrm{joint}}(D)$ can be written as

$$R_{\mathrm{T}}^{\mathrm{joint}}(D) = \sum_{i=1}^{L} H\left(\hat{\mathcal{H}}_i(D)\,\middle|\,\hat{\mathcal{H}}_1(D), \ldots, \hat{\mathcal{H}}_{i-1}(D), \hat{\mathcal{H}}_{i+1}(D), \ldots, \hat{\mathcal{H}}_L(D)\right), \tag{5.3}$$

and the simulcast video coding scheme, whose rate $R_{\mathrm{T}}^{\mathrm{simul}}(D)$ can be written as

$$R_{\mathrm{T}}^{\mathrm{simul}}(D) = \sum_{i=1}^{L} H\left(\hat{\mathcal{H}}_i(D)\right), \tag{5.4}$$

we have

$$R_{\mathrm{T}}^{\mathrm{joint}}(D) \leq R_{\mathrm{T}}^{\mathrm{MT}}(D) \leq R_{\mathrm{T}}^{\mathrm{simul}}(D), \tag{5.5}$$

since condition reduces entropy.

Correspondingly at the joint decoder, the coded texture sequences $\mathcal{H}_1^{\mathrm{ENT}}$ and $\mathcal{H}_i^{\mathrm{WZ}}$, $i = 2\ldots, L$ are decoded sequentially, with each $\mathcal{H}_i^{\mathrm{WZ}}$, $i = 2\ldots, L$ using previously decoded $i-1$ sequences $\hat{\mathcal{H}}_j$, $j = 1, \ldots, i-1$, as well as the decoded depth information $\hat{\mathcal{D}}$ as side information. It should be noticed that for a sequence $\mathcal{H}_i$, $i = 2\ldots, L$, since its side information SI_i at the decoder is generated independently of $\mathcal{H}_i$, joint estimation, i.e., using the side information SI_i and decoded sequence $\tilde{\mathcal{H}}_i$, can be used in the decoder to improve the quality of reconstructed sequence, as shown in Fig. 21.

1. Side information frame generation

In detail, to perform WZ coding (SW coding of coded components), side information is necessary at the decoder. Side information generation is a major step in MT video coding, since the quality of the side information determines the rate of WZ coding. Though different video sequences are highly correlated, the correlation model is complicated if no depth information is provided, since we have no knowledge of the pixel-to-pixel correlation between simultaneous frames from different camera views. On the other hand, let $\mathcal{H}_{i,t}(x, y)$ be a pixel at position (x, y) of frame $\mathcal{H}_{i,t}$ for the i-th view at time slot t, if we have acquired depth information $\mathcal{D}_t$, we can always locate a corresponding pixel position (x', y') in another frame $\mathcal{H}_{i',t}$ from the i'-th view, meaning that $\mathcal{H}_{i,t}(x, y)$ and $\mathcal{H}_{i,t}(x', y')$ represent the same object position in the scene at time slot j. Thus, the correlation between $\mathcal{H}_{i,t}$ and $\mathcal{H}_{i',t}$ becomes pixel-wise, and therefore much easier to be utilized in side information generation for WZ coded components of $\mathcal{H}_{i,t}$.

If depth information can be provided at the decoder, side information generation can become easier and its quality can be better for anchor frames. This is also the major difference between MT video coding with and without depth camera. Therefore, here we describe the side information generation for anchor frames when depth information is not transmitted to the decoder, and the case of non-anchor frames will be further discussed in detail in Section B since it is done similarly as the case with depth camera information.

In our implementation, since temporal prediction can provide information about current depth from previous frames, we treat *anchor frames* (they are only allowed to be predicted by simultaneous frames from other camera views) and *non-anchor frames* (they are allowed to be predicted both temporally from the same view and spatially from other camera views) differently.

In this scheme, if an anchor frame $\mathcal{H}_{2,t}$ in the second view is directly entropy coded under H.264/AVC framework with given quantization parameter (QP) q, for the next anchor frame $\mathcal{H}_{2,t}$ in the second view, we can not apply accurate frame warping since there is no depth information at the decoder. Therefore, we need to follow the algorithm in [28], in which $\mathcal{H}_{2,t}$ is coded in two layers, a coarse layer and a refinement layer, and the two layers are transmitted sequentially. The decoded coarse layer $\tilde{\mathcal{H}}_{2,t}^{\mathrm{C}}$ is a low-quality reconstruction resulted by quantization using a larger QP $q' = q + 12$, such that the two quantizers $\mathbb{C}(q)$ and $\mathbb{C}(q')$ are *embedded*, which means that for quantized DCT coefficients, the zero-th quantization cells of $\mathbb{C}(q')$ contains five quantization cells of $\mathbb{C}(q)$ and other quantization cells of $\mathbb{C}(q')$ contains four. Therefore, the decoded refinement layer $\tilde{\mathcal{H}}_{2,t}^{\mathrm{R}}$ contains only indices $\mathrm{QI}_{2,t}$ for smaller quantization cells of $\mathbb{C}(q)$ in the larger quantization cells of $\mathbb{C}(q')$.

Thus, since $\tilde{\mathcal{H}}_{2,t}^{\mathrm{C}}$ are first obtained by the decoder as a coarse version of $\tilde{\mathcal{H}}_{2,t}$, it can be used jointly with $\tilde{\mathcal{H}}_{1,t}$ at the decoder for depth estimation, and thus the side information for the refinement layer $\tilde{\mathcal{H}}_{2,t}^{\mathrm{R}}$ can be obtained by warping $\tilde{\mathcal{H}}_{1,t}$ to the using the estimated depth information. Moreover, since the depth information can be further estimated using $\tilde{\mathcal{H}}_{1,t}$ and $\tilde{\mathcal{H}}_{2,t}$ and warped to the following views, this two-layer transmission scheme is not necessary for WZ compression of anchor frames of the i-th view when $i \geq 3$. The details of depth estimation frame warping are discussed in Section B.

2. SW coding of frames

Since the proposed MT scheme is implemented under the H.264/AVC framework, the side information frames $\mathrm{SI}_{i,t}^{(1)}, \ldots, \mathrm{SI}_{i,t}^{(i-1)}$, $i = 2, \ldots, L$, $t = 1 \ldots, n$, can not be used directly for WZ coding. Our approach is to encode both the sequence frame $\mathcal{H}_{i,t}$ and side information frames $\mathrm{SI}_{i,t}^{(1)}, \ldots, \mathrm{SI}_{i,t}^{(i-1)}$ by H.264/AVC encoder, and perform SW coding for each bit plane of different components of the H.264/AVC bitstream of $\mathcal{H}_{i,t}$, using the corresponding components of $\mathrm{SI}_{i,t}^{(j)}$ (or the coded bitstream of $\mathrm{SI}_{i,t}^{(j)}$), $j = 1, \ldots, i-1$, as side information. We also need to treat anchor frames and non-anchor frames differently in this step.

a. Anchor frame coding

In MT video coding scheme, anchor frame coding is similar to that of an intra-predicted frame (I-frame) in H.264/AVC, except that for $i = 2, \ldots, L$, major coding components of $\mathcal{H}_i$, such as intra-prediction modes and quantized DCT coefficients, are not entropy coded, but rather SW coded with decoder side information. The detailed coding process is shown in Fig. 22.

Consider one anchor frame $\mathcal{H}_{i,t}$ and its side information frame $\mathrm{SI}_{i,t}^{(j)}$, $j < i$, we can divide the information into three components: intra-prediction mode $\mathrm{MI}_{i,t}$, quantized DCT coefficients $\mathrm{DI}_{i,t}$, and other information $\mathrm{OI}_{i,t}$. $\mathrm{MI}_{i,t}$ and $\mathrm{DI}_{i,t}$ are SW coded at different rates. $\mathrm{OI}_{i,t}$ is entropy coded in the same way as in H.264/AVC. The decoded component $\hat{\mathrm{OI}}_{i,t}$ is later combined with SW decoded $\hat{\mathrm{MI}}_{i,t}$ and $\hat{\mathrm{DI}}_{i,t}$ to form the decoded frame $\tilde{\mathcal{H}}_{i,t}$.

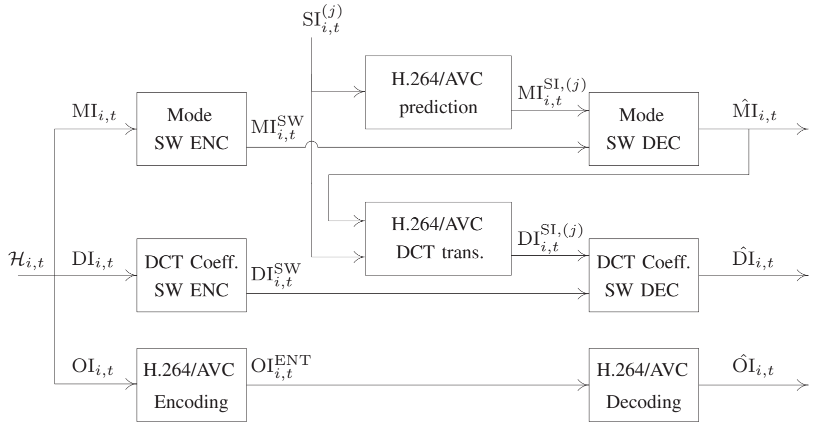

$\mathrm{SI}_{i,t}^{(j)}$
H.264/AVC prediction
$\mathrm{MI}_{i,t}^{\mathrm{SI},(j)}$
$\mathrm{MI}_{i,t}$
Mode SW ENC
$\mathrm{MI}_{i,t}^{\mathrm{SW}}$
Mode SW DEC
$\hat{\mathrm{MI}}_{i,t}$
H.264/AVC DCT trans.
$\mathrm{DI}_{i,t}^{\mathrm{SI},(j)}$
$\mathcal{H}_{i,t}$
$\mathrm{DI}_{i,t}$
DCT Coeff. SW ENC
$\mathrm{DI}_{i,t}^{\mathrm{SW}}$
DCT Coeff. SW DEC
$\hat{\mathrm{DI}}_{i,t}$
$\mathrm{OI}_{i,t}$
H.264/AVC Encoding
$\mathrm{OI}_{i,t}^{\mathrm{ENT}}$
H.264/AVC Decoding
$\hat{\mathrm{OI}}_{i,t}$

Fig. 22. The coding process of an anchor frame components.

Intra prediction modes: Since the intra-prediction mode for each block decides the residual block and thus the DCT coefficients, $\mathrm{MI}_{i,t}$ should be first encoded and decoded so that $\hat{\mathrm{MI}}_{i,t}$ can be used to help compress $\mathrm{DI}_{i,t}$. Suppose we are using warped frame $\mathrm{SI}_{i,t}^{(j)}$, we first encode it using H.264/AVC and calculate the decoder side information $\mathrm{MI}_{i,t}^{\mathrm{SI}}$ for SW coding of $\mathrm{MI}_{i,t}$.

Intra frame DCT coefficients: In the second step, the warped frame $\mathrm{SI}_{i,t}^{(j)}$ is re-encoded using H.264/AVC with $\hat{\mathrm{MI}}_{i,t}$ (instead of $\mathrm{MI}_{i,t}^{\mathrm{SI}}$) being the intra-prediction mode decisions. The resulting DCT coefficients (before quantization) $\mathrm{DI}_{i,t}^{\mathrm{SI}}$ serve as the decoder side information for SW coding of $\mathrm{DI}_{i,t}$. Finally, the above decoded components $\hat{\mathrm{MI}}_{i,t}$ and $\hat{\mathrm{DI}}_{i,t}$ are combined with other decoded components $\hat{\mathrm{OI}}_{i,t}$ to construct the decoded anchor frame $\tilde{\mathcal{H}}_{i,t}$. An example of the correlation model of quantized non-zero I-frame coefficients and their decoder side information is shown in Fig. 23.

Remark:

For anchor frames of the second view, since only the refinement layers are transmitted to the decoder, the only component that can be SW compressed is quantization cell indices $\mathrm{QI}_{2,t}$, which is part of the DCT coefficients.

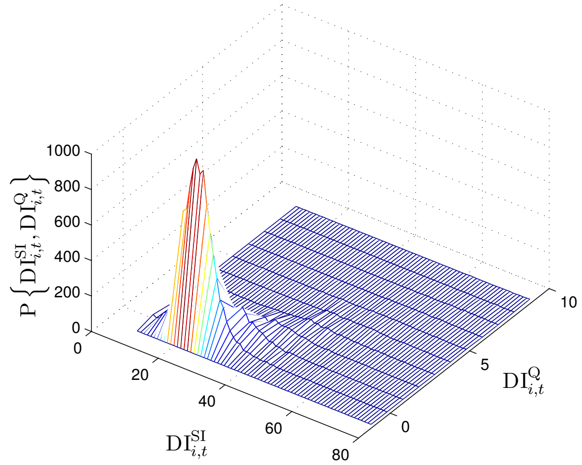


Fig. 23. An example of the correlation model of quantized non-zero I-frame coefficients and their decoder side information.

b. Non-anchor frame coding

Non-anchor frame coding in MT video coding scheme is similar to that of a predicted frame (P-frame or B-frame) in H.264/AVC. The difference is that for $\mathcal{H}_i$, $i = 2, \ldots, L$, the major coding components, such as inter-prediction mode, motion vector difference (MVD), as well as the DCT coefficients, are not entropy coded but SW coded, as shown in Fig. 24.

Inter prediction modes: As shown in the figure, consider one non-anchor $\mathcal{H}_{i,t}$ and its j-th side information frame $\mathrm{SI}^{(j)}_{i,t}$, the inter-prediction mode $\mathrm{MP}_{i,t}$ is first SW coded with side information $\mathrm{MP}^{\mathrm{SI},(j)}_{i,t}$ generated by H.264/AVC coding of $\mathrm{SI}^{(j)}_{i,t}$. In detail, inter prediction modes in H.264/AVC includes prediction block sizes, prediction directions, and reference frame indices. In SW coding of these components, the corresponding components acquired after H.264/AVC coding of $\mathrm{SI}^{(j)}_{i,t}$ are used as side information bit plane wise.

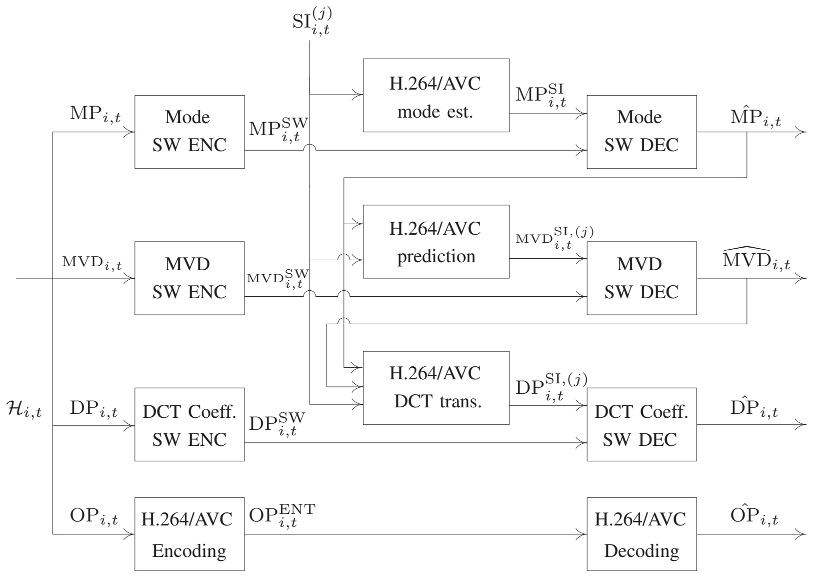

$\mathrm{SI}_{i,t}^{(j)}$
H.264/AVC
mode est.
$\mathrm{MP}_{i,t}^{\mathrm{SI}}$
$\mathrm{MP}_{i,t}$
Mode
SW ENC
$\mathrm{MP}_{i,t}^{\mathrm{SW}}$
Mode
SW DEC
$\hat{\mathrm{MP}}_{i,t}$
H.264/AVC
prediction
$\mathrm{MVD}_{i,t}^{\mathrm{SI},(j)}$
$\mathrm{MVD}_{i,t}$
MVD
SW ENC
$\mathrm{MVD}_{i,t}^{\mathrm{SW}}$
MVD
SW DEC
$\widehat{\mathrm{MVD}}_{i,t}$
H.264/AVC
DCT trans.
$\mathrm{DP}_{i,t}^{\mathrm{SI},(j)}$
$\mathcal{H}_{i,t}$
$\mathrm{DP}_{i,t}$
DCT Coeff.
SW ENC
$\mathrm{DP}_{i,t}^{\mathrm{SW}}$
DCT Coeff.
SW DEC
$\hat{\mathrm{DP}}_{i,t}$
$\mathrm{OP}_{i,t}$
H.264/AVC
Encoding
$\mathrm{OP}_{i,t}^{\mathrm{ENT}}$
H.264/AVC
Decoding
$\hat{\mathrm{OP}}_{i,t}$

Fig. 24. The code process of a non-anchor frame components.

Motion vector differences: After inter prediction modes have been correctly decoded, we then SW code MVD using side information $\mathrm{MVD}_{i,t}^{\mathrm{SI},(j)}$ generated by H.264/AVC coding of $\mathrm{SI}_{i,t}^{(j)}$ with $\hat{\mathrm{MP}}_{i,t}$ as the inter-prediction mode. MVD has horizontal and vertical components, both using the corresponding components of $\mathrm{MVD}_{i,t}^{\mathrm{SI},(j)}$ as side information bit plane wise.

DCT coefficients: After all the motion information are successfully decoded, $\mathrm{SI}_{i,t}^{(j)}$ is coded again by H.264/AVC with inter-prediction mode and MVD set to $\hat{\mathrm{MP}}$ and the decoded MVD's $\widehat{\mathrm{MVD}}_{i,t}$ respectively, and the resulting DCT coefficients $\mathrm{DP}_{i,t}^{\mathrm{SI},(j)}$ (before quantization) serve as the decoder side information for SW coding of $\mathrm{DP}_{i,t}$. Finally, the above decoded components $\hat{\mathrm{MP}}_{i,t}$, $\widehat{\mathrm{MVD}}_{i,t}$, and $\hat{\mathrm{DP}}_{i,t}$ are combined with other decoded components $\hat{\mathrm{OP}}_{i,t}$ to reconstruct the non-anchor frame $\hat{\mathcal{H}}_{i,t}$.

c. Slepian-Wolf code design

Practical SW coding is implemented via syndrome-based binning. Each bit plane of a quantized source is partitioned into bins indexed by syndromes of a channel code. The encoder computes the syndrome $\boldsymbol{s} = \boldsymbol{x}\boldsymbol{H}^{\mathrm{T}}$ and sends it to the decoder at rate $R^{\mathrm{SW}} = (n-k)/n$ b/s, where $\boldsymbol{x}$ is a length-n binary sequence and $\boldsymbol{H}$ is the $(n-k) \times n$ parity-check matrix. At the decoder end, based on the side information $\boldsymbol{y}$ and received syndrome $\boldsymbol{s}$, the decoder finds the recovered sequence $\hat{\boldsymbol{x}}$ in the coset $\mathbb{C}_{\boldsymbol{s}} = \left\{\boldsymbol{x} \in \{0,1\}^n : \boldsymbol{x}\boldsymbol{H}^{\mathrm{T}} = \boldsymbol{s}\right\}$, i.e.,

$$\hat{\boldsymbol{x}} = \arg\max_{\boldsymbol{x} \in \mathbb{C}_{\boldsymbol{s}}} p(\boldsymbol{x}|\boldsymbol{y}) \tag{5.6}$$

For practical SW coding, we choose LDPC codes because of their capacity-approaching performance and flexibility in code design. The message-passing decoding algorithm can also be applied to SW coding with a little modification. LDPC codes can be designed for different components according to the log-likelihood ratio (LLR) distribution of the correlation channel. In our case, the LLR distribution is very similar to that of a binary AWGN channel, e.g., the conditional probability distribution function (p.d.f.) of the LLR of one WZ coded component is shown in Fig. 25. Therefore, we choose LDPC codes designed for AWGN channels at different rates for different component correlation models.

d. Joint estimation at the decoder

In our scheme, when decoding the frames from the i-th view with depth information $\hat{\mathcal{D}}$ at the decoder, as shown in Fig. 21, the simultaneous frame from previously $i-1$ decoded views can be warped to the current view frame $\mathcal{H}_i$ as side information frame SI_i. If the depth information $\hat{\mathcal{D}}$ is accurate enough, SI_i can be considered as consisting

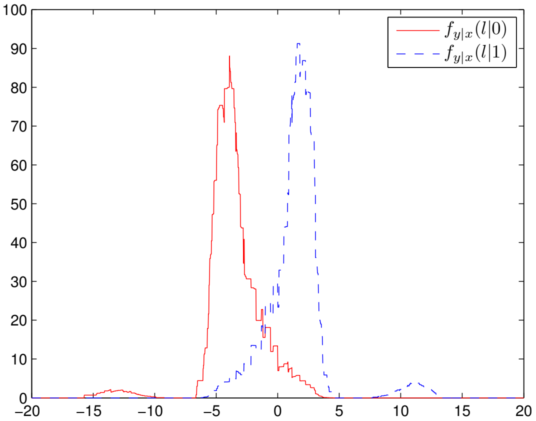


Fig. 25. The conditional p.d.f.'s of the LLR of the 1st bit plane of anchor frame DCT coefficients (non-zero DC coefficient signs), given the transmitted bit is 0 or 1.

of $i-1$ noised versions of $\mathcal{H}_i$, i.e.,

$$\mathrm{SI}_i^{(j)} = \mathcal{H}_i + \mathcal{N}_i^{\mathrm{S},j},\ j = 1, \ldots, i-1, \tag{5.7}$$

where $\mathcal{N}_i^{\mathrm{S},j}$ represents the noise between $\mathcal{H}_i$ and previously decoded frame $\hat{\mathcal{H}}_j$, which is caused by camera sensor difference, inaccurate pixel mapping, as well as random thermal noise. On the other hand, the decoded version $\tilde{\mathcal{H}}_i$ is also a noised version of $\mathcal{H}_i$, i.e.,

$$\tilde{\mathcal{H}}_i = \mathcal{H}_i + \mathcal{N}_i^{\mathrm{C}}, \tag{5.8}$$

where $\mathcal{N}_i^{\mathrm{C}}$ represents the noise caused by quantization, filtering, etc. Since $\mathcal{N}_i^{\mathrm{C}}$ and $\mathcal{N}_i^{\mathrm{S}}$ are from different sources, they can be considered independent. Therefore, if the reconstructed version $\hat{\mathcal{H}}_i$ is jointly estimated from $\tilde{\mathcal{H}}_i$ and SI_i with MSE criterion, the noise power can be reduced. If the noises are AWGN noise, we know that a linear combination of input signals weighted by noise variances is optimal. However, in our case, $\mathcal{N}_i^{\mathrm{C}}$ is nearly Laplacian and $\mathcal{N}_i^{\mathrm{S}}$ is much more complicated. Thus, the optimal estimator is difficult to determine. Therefore the optimal estimator $\hat{\mathcal{H}}_i\left(\tilde{\mathcal{H}}_i, \mathrm{SI}_i, \hat{\mathcal{D}}\right)$ should be computed using Bayes model.

To achieve an optimal estimation $\hat{\mathcal{H}}_i$ from $\tilde{\mathcal{H}}_i$ and SI_i in MSE sense, we need to find the conditional expectation, i.e.,

$$\hat{\mathcal{H}}_i = \mathrm{E}\left[\mathcal{H}_i \,\middle|\, \tilde{\mathcal{H}}_i, \mathrm{SI}_i, \hat{\mathcal{D}}\right], \tag{5.9}$$

for $i = 2, \ldots, L$. If assuming pixel-wise independence for a given frame, we have

$$\hat{\mathcal{H}}_i(x,y) = \mathrm{E}\left[\mathcal{H}_i(x,y) \,\middle|\, \tilde{\mathcal{H}}_i(x,y), \mathrm{SI}_i(x,y)\right], \tag{5.10}$$

for any pixel (x, y) in the frame $\mathcal{H}_i$. Then, for a given set of values of $\mathcal{H}_i(x, y) = h$ and $\mathrm{SI}_i(x, y) = s$, we have

$$\begin{aligned} &\mathrm{E}\left[\mathcal{H}_i(x,y) \,\middle|\, \tilde{\mathcal{H}}_i(x,y) = h, \mathrm{SI}_i(x,y) = s\right] \\ &= \int_{\mathcal{H}_i(x,y)} \mathcal{H}_i(x,y) p\left(\mathcal{H}_i(x,y) \,\middle|\, \tilde{\mathcal{H}}_i(x,y) = h, \mathrm{SI}_i(x,y) = s\right) \mathrm{d}\mathcal{H}_i(x,y), \end{aligned} \tag{5.11}$$

and by Bayes' theorem, the conditional p.d.f. can be written as

$$\begin{aligned} &p\left(\mathcal{H}_i(x,y) \,\middle|\, \tilde{\mathcal{H}}_i(x,y) = h, \mathrm{SI}_i(x,y) = s\right) \\ &= \frac{p\left(\tilde{\mathcal{H}}_i(x,y) = h, \mathrm{SI}_i(x,y) = s \,|\mathcal{H}_i(x,y)\right)}{p\left(\tilde{\mathcal{H}}_i(x,y) = h, \mathrm{SI}_i(x,y) = s\right)} \cdot p\left(\mathcal{H}_i(x,y)\right) \\ &= \frac{p\left(\tilde{\mathcal{H}}_i(x,y) = h \,|\mathcal{H}_i(x,y)\right) p\left(\mathrm{SI}_i(x,y) = s \,|\mathcal{H}_i(x,y)\right)}{p\left(\tilde{\mathcal{H}}_i(x,y) = h, \mathrm{SI}_i(x,y) = s\right)} \cdot p\left(\mathcal{H}_i(x,y)\right) \\ &= \frac{p_{\mathcal{N}_i^{\mathrm{C}}}\left(h - \mathcal{H}_i(x,y)\right) p_{\mathcal{N}_i^{\mathrm{S}}}\left(s - \mathcal{H}_i(x,y)\right)}{p\left(\tilde{\mathcal{H}}_i(x,y) = h, \mathrm{SI}_i(x,y) = s\right)} \cdot p\left(\mathcal{H}_i(x,y)\right), \end{aligned} \tag{5.12}$$

where $p_{\mathcal{N}_i^{\mathrm{C}}}(\cdot)$ and $p_{\mathcal{N}_i^{\mathrm{S}}}(\cdot)$ are the p.d.f.'s of $\mathcal{N}_i^{\mathrm{C}}$ and $\mathcal{N}_i^{\mathrm{S}}$, respectively. Since the values of $\mathcal{H}_i(x,y)$ and $\mathrm{SI}_i(x,y)$ are known, by (5.11) and (5.12), we have

$$\begin{aligned}&\mathrm{E}\left[\mathcal{H}_i(x,y)\,\middle|\,\tilde{\mathcal{H}}_i(x,y)=h,\mathrm{SI}_i(x,y)=s\right]\\ =&\int_{\mathcal{H}_i(x,y)} p\left(\mathcal{H}_i(x,y)\right)\mathcal{H}_i(x,y)p_{\mathcal{N}_i^{\mathrm{C}}}\left(h-\mathcal{H}_i(x,y)\right)p_{\mathcal{N}_i^{\mathrm{S}}}\left(s-\mathcal{H}_i(x,y)\right)\mathrm{d}\mathcal{H}_i(x,y).\end{aligned} \tag{5.13}$$

For simplicity, we assume that the statistics for $\mathcal{N}_i^{\mathrm{C}}$ and $\mathcal{N}_i^{\mathrm{S}}$ are independent of pixel position (x,y) as well as the temporal order i, and we also assume that the original pixel luma value $\mathcal{H}_i(x,y)$ appears with equal probability. Then (5.13) can be simplified as

$$\mathrm{E}\left[\mathcal{H}_i(x,y)\,\middle|\,\tilde{\mathcal{H}}_i(x,y)=h,\mathrm{SI}_i(x,y)=s\right]=\int_{h_{\min}}^{h_{\max}} xp_{\mathcal{N}_i^{\mathrm{C}}}\left(h-x\right)p_{\mathcal{N}_i^{\mathrm{S}}}\left(s-x\right)\mathrm{d}x, \tag{5.14}$$

where $h_{\min}$ and $h_{\max}$ are the minimum and maximum possible values of $\mathcal{H}_i(x,y)$. By using (5.14), the optimal estimation of $\mathcal{H}_i(x,y)$ for any given $\tilde{\mathcal{H}}_i$ and SI_i can be trained from the a few frames shot by the set of cameras with identical configuration.

From the above analysis, it should be noticed that the joint estimation process only requires accurate pixel-to-pixel correspondence. Therefore, as long as the pixel-wise depth information is acquired at the decoder, joint estimation can be performed from multiple decoded simultaneous frames from other camera views, which is not limited to MT video coding and can be applied to various multiview video applications.

B. Depth camera assisted MT video coding

In this section, we chiefly describe how to use separately encoded and transmitted depth information in MT video coding. The block diagram of MT video coding scheme with depth camera assistance is shown in Fig. 26. In this scheme, depth information $\mathcal{D}$ consists of depth images (or depth maps) of the current scene at time slots $1, 2, \dots, n$, which are denoted as $\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_n$.

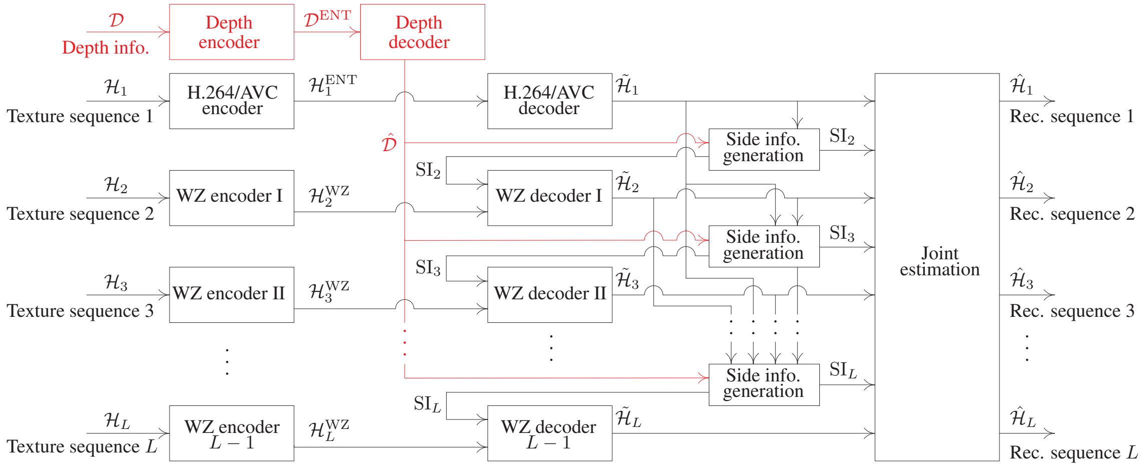


Fig. 26. The block diagram of the proposed MT video coding scheme with depth camera assistance.

Therefore, if we assume that WZ coding approaches conditional entropy, the total transmission rate $R^{\mathrm{MT}}_{\mathrm{TD}}(D)$ of the MT coding scheme with depth camera help can be written as

$$R^{\mathrm{MT}}_{\mathrm{TD}}(D) = R^{\mathrm{MT}}_{\mathrm{D}}(\mathcal{D}) + \sum_{i=1}^{L} R^{\mathrm{MT}}_{\mathrm{T}i}(\mathcal{H}_i, D, \hat{\mathcal{D}}) \tag{5.15}$$

$$= H\left(\hat{\mathcal{D}}\right) + H\left(\hat{\mathcal{H}}_1(D)\right) + \sum_{i=2}^{L} H\left(\hat{\mathcal{H}}_i(D) \middle| \hat{\mathcal{H}}_1(D), \dots, \hat{\mathcal{H}}_{i-1}(D), \hat{\mathcal{D}}\right), \tag{5.16}$$

where $R^{\mathrm{MT}}_{\mathrm{D}}(\mathcal{D})$ is the rate of the depth information $\mathcal{D}$, $R^{\mathrm{MT}}_{\mathrm{T}_i}(\mathcal{H}_i, D)$ is the rate of texture sequence $\mathcal{H}_i$ given distortion measure D and reconstructed depth information $\hat{\mathcal{D}}$. It can be seen that compared to (5.2), the MT scheme with depth camera help transmit depth information $\mathcal{D}$ with additional rate $H\left(\hat{\mathcal{D}}\right)$. However, it is shown in the Section C that by utilizing depth information at the decoder, the correlation at the decoder can be largely improved and thus the total rate is actually reduced. Moreover, a compromise should be made between the quality and transmission rate of the depth information since more accurate depth information costs more rate to transmit but provides higher correlation between different views thus lowering the texture sequence rate $H\left(\hat{\mathcal{H}}_i(D)\middle|\hat{\mathcal{H}}_1(D),\ldots,\hat{\mathcal{H}}_{i-1}(D),\hat{\mathcal{D}}\right)$, $i = 2,\ldots,L$, and *vice versa.*

As mentioned in Section A, we only discuss the side information generation process here for both MT video coding with/without depth camera, since this is the difference between the two schemes.

1. Anchor frame warping

Since an anchor frame is only allowed to be predicted from simultaneous frames from other views, we need to transmit depth information to the decoder for anchor frames, and the decoded frames from other views can be warped to the current view using depth information. The issue of depth information compression has been well discussed [52, 53, 54]. However, since the accuracy of depth value is critical in our application, we code each depth value losslessly, while downsampling the resolution of depth image to restrict the transmission rate $R_{\mathrm{D}}(\mathcal{D})$ of the depth information.

Since the depth information at time t consists of a depth map with each pixel showing the depth values, we can consider the depth information $\mathcal{D}_t$ is also a picture frame shot by an imaginary camera with known configurations and parameters and synchronized with the camera set. Thus for a pixel position (x_D, y_D) in the depth map (or the *depth coordinate*), $\mathcal{D}_t(x_D, y_D)$ is the implicit depth value (the distance between its corresponding object position in the space, or *world coordinate* and the imaginary camera). Then by multiview geometry, we have [55]:

$$\mathcal{D}_t(x_D, y_D) \cdot (x_D, y_D, 1)^{\mathrm{T}} = \boldsymbol{K}_D \boldsymbol{R}_D (X_D, Y_D, Z_D, 1)^{\mathrm{T}}, \tag{5.17}$$

where $\boldsymbol{K}_D$ and $\boldsymbol{K}_D$ are the 3×3 intrinsic matrices and $\boldsymbol{R}_D$, $\boldsymbol{R}_D$ are the 3×4 rotation matrices (or extrinsic matrices) of the imaginary depth camera, and (X_D, Y_D, Z_D) is the corresponding object position in the world coordinate. Therefore, the actual depth (Z_D) of any world coordinate position (X_D, Y_D, Z_D) at time t can be derived if it is shown in $\mathcal{D}_t$.

With the decoded depth information $\hat{\mathcal{D}}$, the warping process is as follows (shown in Fig. 27). Let (X, Y, Z) be an actual position in the world coordinate, (x_m, y_m) and (x_n, y_n) $(1 \leq m, n \leq L)$ be the corresponding points in the *camera coordinates* of the m-th and n-th camera views, respectively. We have [55]:

$$z_m \cdot (x_m, y_m, 1)^{\mathrm{T}} = \boldsymbol{K}_m \boldsymbol{R}_m (X, Y, Z, 1)^{\mathrm{T}} \tag{5.18}$$

$$z_n \cdot (x_n, y_n, 1)^{\mathrm{T}} = \boldsymbol{K}_n \boldsymbol{R}_n (X, Y, Z, 1)^{\mathrm{T}}, \tag{5.19}$$

where $\boldsymbol{K}_m$ and $\boldsymbol{K}_n$ are the intrinsic matrices and $\boldsymbol{R}_m$, $\boldsymbol{R}_n$ are the rotation matrices, and z_m, z_n are implicit depth values associated with pixel positions (x_m, y_m) and (x_n, y_n) in the two camera coordinates, respectively. These matrices can be calibrated and calculated when the camera positions and focal lengths are fixed. For each pixel position (x_m, y_m) in the m-th camera view frame, we acquire the corresponding depth Z in the world coordinate from the decoded depth information $\hat{\mathcal{D}}$ by using (5.17), thus we can solve for (z_m, X, Y) from the three equations implied by (5.18). Then using (5.19), the pixel correspondence $(x_m, y_m) \leftrightarrow (x_n, y_n)$ between the left and the right view can be computed, and warping can be performed from a left view frame to a new frame such that it looks like it were shot by the right view.

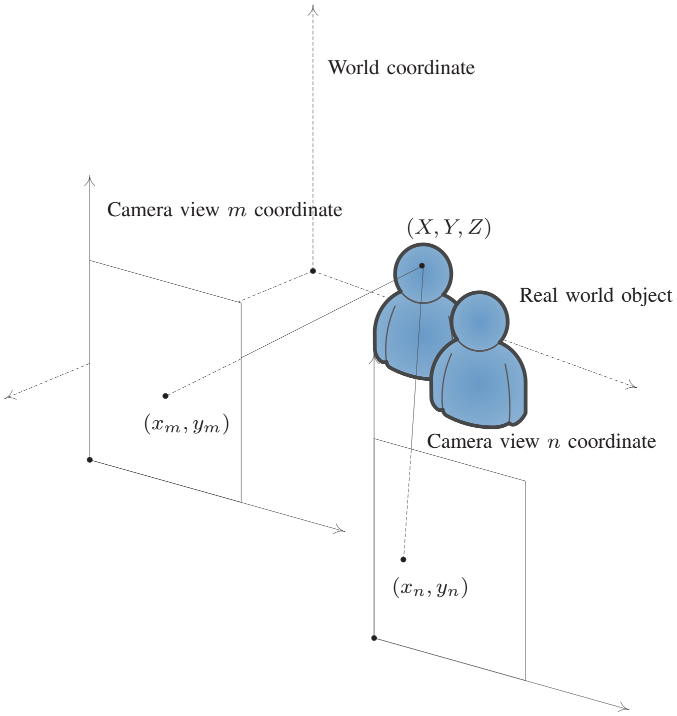


Fig. 27. Multiview geometry for frame warping from camera view m to camera view n.

It should be noticed that if we replace one of the equations from (5.18) and (5.19) with (5.17) using the same world coordinate position (X, Y, Z), then similar derivations follow, which means the the decoded depth information $\hat{\mathcal{D}}_t$ can also be warped to any camera view. The decoded depth information warped to the i-th camera view can be denoted as $\hat{\mathcal{D}}_{i,t}$. Then the pixel-to-pixel mapping between $\hat{\mathcal{D}}_{i,t}$ and $\mathcal{H}_{i,t}$ can be acquired, facilitating further discussion in this paper.

It can be seen that the warping process is time invariant if camera positions and configurations are fixed when the video sequences are shot. Therefore, we can denote the warping function from j-th view to the i-th view as $W_{j,i}(\cdot)$. With the above approach, for any given anchor frame $\mathcal{H}_{i,t}$, we can get its side information frame (denoted as $\mathrm{SI}_{i,t}^{(j)}$) from $\hat{\mathcal{H}}_{j,t}$, i.e.,

$$\mathrm{SI}_{i,t}^{(j)} = W_{j,i}\left(\hat{\mathcal{H}}_{j,t}, \hat{\mathcal{D}}_{i,t}\right),\ j = 1, 2, \ldots, i-1. \tag{5.20}$$

Since we have all previous $i-1$ frames at time t decoded, the decoder side information $\mathrm{SI}_{i,t}$ for an anchor frame $\mathcal{H}_{i,t}$ consists of $i-1$ side information frames warped from previously decoded frames:

$$\begin{aligned} \mathrm{SI}_{i,t} &= \left\{\mathrm{SI}_{i,t}^{(1)}, \mathrm{SI}_{i,t}^{(2)}, \ldots, \mathrm{SI}_{i,t}^{(i-1)}\right\} \\ &= \left\{W_{1,i}\left(\hat{\mathcal{H}}_{1,t}, \hat{\mathcal{D}}_{i,t}\right), W_{2,i}\left(\hat{\mathcal{H}}_{2,t}, \hat{\mathcal{D}}_{i,t}\right), \ldots, W_{i-1,i}\left(\hat{\mathcal{H}}_{i-1,t}, \hat{\mathcal{D}}_{i,t}\right)\right\}. \end{aligned} \tag{5.21}$$

If the depth information $\mathcal{D}_t$ is accurate enough, the only noise in $\mathrm{SI}_{i,t}$ comes from camera sensor difference and thermal noise, which can be considered as spatial and temporal independent. Therefore, we can generate decoder side information for WZ coded components of $\mathcal{H}_{i,t}$ with high precision.

Remarks:

Occlusion will occur in both depth map warping and texture frame warping with known depth, which means that scenes shot by one camera view might be occluded thus not appear in another camera view at the same time, thus no depth or texture values can be assigned to the occluded regions the during warping. Occlusion is caused by the geometry of the scene and the configuration of camera set, etc. To minimize the affect from occlusion in frame warping, first we can utilize warping from multiple views to the current view, which can provide more complete information of the scene, since a region in the world coordinate occluded in one view could be captured in another view at the same time. For those regions still occluded after the first step, we need to search for the nearest available neighbor in the frame for depth or texture values to construct a reasonable side information frame.

2. Non-anchor frame warping

For a non-anchor frame $\mathcal{H}_{i,t}$, since temporal prediction can be performed, the depth information for such frames can be predicted by the frame motion information of previously decoded frames from other camera views, i.e., the motion between $\hat{\mathcal{H}}_{j,t}$ and $\hat{\mathcal{H}}_{j,t'}$, $j = 1, \ldots, i-1$, together with the depth information $\mathcal{D}_{t'}$ from previously decoded frames at time t', thus saving transmission rate for depth information while maintaining acceptable depth information quality. If $\mathcal{H}_{1,t'}, \ldots, \mathcal{H}_{L,t'}$ are coded as anchor frames, the depth information $\mathcal{D}_{t'}$ can be used directly for depth information estimation at time t since $\mathcal{D}_{t'}$ is transmitted to the decoder losslessly. On the other hand, if $\mathcal{H}_{1,t'}, \ldots, \mathcal{H}_{L,t'}$ are coded as non-anchor frames, we do not have instant depth information at the decoder. However, since we already have the reconstructed frames $\hat{\mathcal{H}}_{1,t'}, \ldots, \hat{\mathcal{H}}_{L,t'}$, the decoder side depth information can be estimated from any two reconstructed frames, and then used to further estimation of depth information at time t can be performed.

In detail, let $\mathcal{H}_{i,t}$, $\mathcal{H}_{i,t'}$ be two right view frames, where frames at time t' have all been reconstructed at the decoder thus we have $\hat{\mathcal{H}}_{1,t'} \ldots, \hat{\mathcal{H}}_{L,t'}$, and now we are trying to estimate the depth information for frame warping at time t: $\tilde{\mathcal{D}}_t$. First, consider the case that $\mathcal{H}_{1,t'}, \ldots, \mathcal{H}_{L,t'}$ are coded as anchor frames. The scheme in this case is shown in Fig. 28.

In this case, let $\mathcal{H}_{j,t}$, $j < i$, be the j-th view frame at time t, and since it is coded before $\mathcal{H}_{i,t}$, we already have its reconstruction $\hat{\mathcal{H}}_{j,t}$. Thus, we can estimate the motion $M_j^{(t',t)}$ in the j-th view between time t' and t from $\hat{\mathcal{H}}_{j,t'}$ to $\hat{\mathcal{H}}_{j,t}$, i.e., finding $M_j^{(t',t)}$ such that for every pixel (x, y) in the j-th view frame, we have

$$\hat{\mathcal{H}}_{j,t'}\left(x + M_{j,h}^{(t',t)}(x,y), y + M_{j,v}^{(t',t)}(x,y)\right) = \hat{\mathcal{H}}_{j,t}(x,y). \tag{5.22}$$

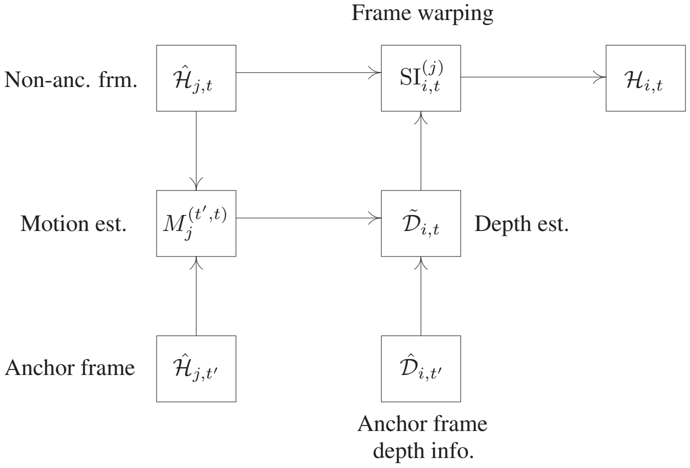

Frame warping
Non-anc. frm.
$\hat{\mathcal{H}}_{j,t}$
$\mathrm{SI}_{i,t}^{(j)}$
$\mathcal{H}_{i,t}$
Motion est.
$M_j^{(t',t)}$
$\tilde{\mathcal{D}}_{i,t}$
Depth est.
Anchor frame
$\hat{\mathcal{H}}_{j,t'}$
$\hat{\mathcal{D}}_{i,t'}$
Anchor frame
depth info.

Fig. 28. The non-anchor frame warping process with neighboring anchor frames.

where $M_{j,h}^{(t',t)}(x,y)$ and $M_{j,v}^{(t',t)}(x,y)$ are the horizontal and vertical components of motion vector at pixel position (x,y) of $M_j^{(t',t)}$. From Section 1, since the depth information $\hat{\mathcal{D}}_{t'}$ is known, then the pixel mapping between the i-th and j-th views at time t' can be acquired, i.e., for a given pixel position (x',y') in $\mathcal{H}_{i,t'}$, we can find its corresponding pixel position (x,y) in $\mathcal{H}_{j,t'}$ by using $\hat{\mathcal{D}}_{t'}$. Moreover, since these two positions maps to one point in the world coordinate, and the motion vector in the j-th camera view coordinate is $M_j^{(t',t)}(x,y)$, from the geometry shown in Fig. 29, we can find that the best estimation of the motion vector in the i-th camera view coordinate is

$$\tilde{M}_i^{(t',t)}(x',y') = \frac{d_i(x',y')}{d_j(x,y)} \cdot M_j^{(t',t)}(x,y) \cdot \cos\theta_{ij}, \tag{5.23}$$

where $d_i(x,y)$, $d_j(x,y)$ are the distance from the world coordinate pixel to camera i and j, respectively, and θ_{ij} is the angle formed by the normal rays of the two cameras. These parameters can be derived from the extrinsic and intrinsic matrices of camera i and j.

Thus, since $\hat{\mathcal{D}}_{t'}$ can be warped to the i-th view to $\hat{\mathcal{D}}_{i,t'}$, which has identical

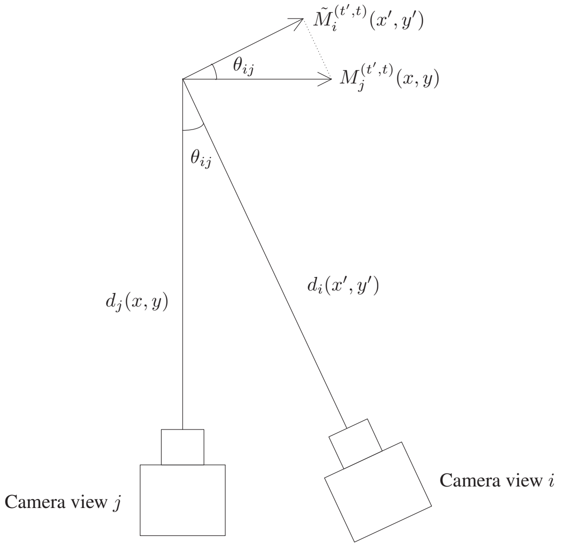


Fig. 29. The geometry of motion vector estimation between different camera views.

camera configurations and parameters as $\mathcal{H}_{i,t}$, from (5.22) and (5.23) we can estimate the depth information for the i-th view at time t using

$$\begin{aligned}\tilde{\mathcal{D}}_{i,t}(x',y') &= \hat{\mathcal{D}}_{i,t'}\left(x' + \tilde{M}_{i,h}^{(t',t)}(x',y'), y' + \tilde{M}_{i,v}^{(t',t)}(x',y')\right)\\ &= \hat{\mathcal{D}}_{i,t'}\left(x' + \frac{d_i(x',y')}{d_j(x,y)} \cdot M_{j,h}^{(t',t)}(x,y)\cdot\cos\theta_{ij}, y' + \frac{d_i(x',y')}{d_j(x,y)}\cdot M_{j,v}^{(t',t)}(x,y)\cdot\cos\theta_{ij}\right),\end{aligned} \tag{5.24}$$

where $\tilde{M}_{i,h}^{(t',t)}(x',y')$ and $\tilde{M}_{i,v}^{(t',t)}(x',y')$ are the horizontal and vertical components of the estimated motion vector $\tilde{M}_i^{(t',t)}(x',y')$.

In the above process, similar to the warping process in Section 1, occlusion will occur because of the difference of depth distribution in the two views and the inaccuracy of depth estimation, which means the estimation from $\hat{\mathcal{D}}_{i,t'}$ to $\tilde{\mathcal{D}}_{i,t}$ not one-one mapping. To eliminate such occlusion with minimal loss, the position in $\tilde{\mathcal{D}}_{i,t}$ that can not find a proper estimation from $\hat{\mathcal{D}}_{i,t'}$ will share the depth value from the neighboring pixel with smallest depth.

Thus, after it is acquired, $\tilde{\mathcal{D}}_{i,t}$ can be further warped to the j-th view as the estimation of depth $\tilde{\mathcal{D}}_{j,t}$ by using (5.17), and then $\hat{\mathcal{H}}_{j,t}$ can be warped to the i-th view using $\tilde{\mathcal{D}}_{j,t}$ as a side information frame $\mathrm{SI}_{i,t}^{(j)}$, and the warping process is complete.

In the case that $\mathcal{H}_{1,t'}, \ldots, \mathcal{H}_{L,t'}$ are coded as non-anchor frames, the depth estimation scheme is shown in Fig. 30. In this case, the difference is that there is no instantly decoded depth information for the reference frames. Therefore, instead of using $\hat{\mathcal{D}}_t$ to estimate $\tilde{\mathcal{D}}_t$ directly, we need to first estimate the reference frame depth, denoted as $\tilde{\mathcal{D}}^*_{t'}$. To estimate depth from multiple views, various stereo matching algorithms can be applied [56], which should consider matching cost of pixel texture values differences and texture shape differences and thus can produce well shaped depth (or disparity) images. In this application, we are considering R-D performance instead of depth map integrity, and a simplified stereo matching algorithm for depth estimation is illustrated as follows.

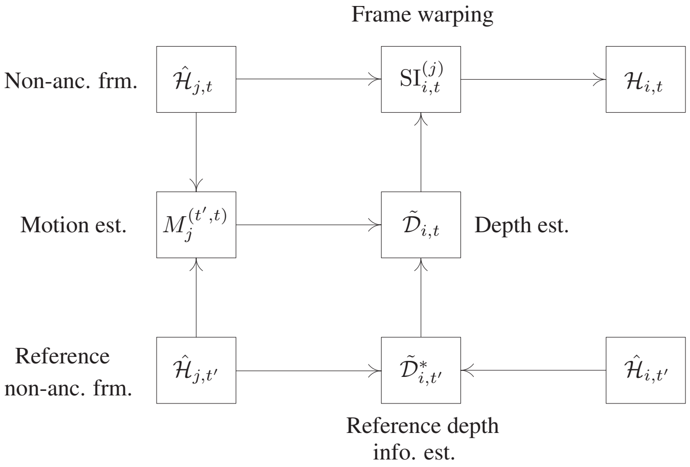

Frame warping
Non-anc. frm.
$\hat{\mathcal{H}}_{j,t}$
$\mathrm{SI}_{i,t}^{(j)}$
$\mathcal{H}_{i,t}$
Motion est.
$M_j^{(t',t)}$
$\tilde{\mathcal{D}}_{i,t}$
Depth est.
Reference non-anc. frm.
$\hat{\mathcal{H}}_{j,t'}$
$\tilde{\mathcal{D}}^*_{i,t'}$
$\hat{\mathcal{H}}_{i,t'}$
Reference depth info. est.

Fig. 30. The non-anchor frame warping process with neighboring non-anchor frames.

For a reasonable estimation, notice that we have all decoded reference frames $\mathcal{H}_{1,t'}, \ldots, \mathcal{H}_{L,t'}$ available at the decoder, thus for each pair, say from j-th to i-th view, of frame warping, we can estimate the reference depth information only $\tilde{\mathcal{D}}^*_{i,t'}$ from the two reference frames $\hat{\mathcal{H}}_{j,t'}$ and $\hat{\mathcal{H}}_{i,t'}$ by searching for the disparity between them. This can be written as an optimization problem as

$$\tilde{\mathcal{D}}^*_{i,t'} = \arg \min_{\tilde{\mathcal{D}}_{i,t'}} \left\| W_{j,i} \left(\hat{\mathcal{H}}_{j,t'}, \tilde{\mathcal{D}}_{i,t'} \right) - \hat{\mathcal{H}}_{i,t'} \right\|_2^2 . \tag{5.25}$$

This problem can be solved similarly to the process of motion search in video coding without R-D optimization using MSE criterion if we treat the spatial difference as temporal, and the complexity is therefore of the same magnitude.

From the above description, we can see that if depth information is available at the decoder, frame warping for both anchor frames and non-anchor frames can be implemented by limited calculations, which can greatly reduce the decode side complexity while keeping the frame warping accuracy for acceptable R-D performance, since complicated stereo matching algorithms are not necessary for finding the pixel mapping between frames from different views.

Remarks:

1. In Section 1 and 2, we described the side information generation process in MT coding with depth camera assistance. For non-anchor frames in the MT video coding scheme without depth information transmitted, the process in 2 can be exactly followed, since depth information can always be estimated from decoded reference frames.

2. The warping process in Section 1 and the depth estimation process in Section 2 are also used in anchor frame warping in the MT scheme without depth camera, where the depth information $\tilde{\mathcal{D}}_t$ is estimated by the reconstructed anchor frame $\tilde{\mathcal{H}}_{1,t}$ from the first view and the decoded coarse layer $\tilde{\mathcal{H}}^{\mathrm{C}}_{2,t}$ of the anchor frame from the second view.

C. Experimental results

We implemented our MT video coding scheme (both with and without depth camera assistance) for two scenarios: multiview video sequence set *argon* with synchronized depth sequence, and standard MVC test sequences *akko&kayo* and *rena*, for which we generated corresponding depth sequence by ourselves. The results for these two scenarios are shown in Subsection 1 and 2, respectively.

1. Experiment on sequence *argon*

In order to get correlated video sequences, we employ $L = 4$ HD cameras (Pilot GigE vision cameras by Basler Vision Technologies) which output colored video frames, and one depth camera (SwissRanger SR4000), which outputs the depth value for every points in the current scene, in grey-scale form. The four HD cameras are fixed closely and horizontally to a cage, while the depth camera is placed closely above one of the HD cameras. This compact setup is implemented for higher correlation between different views and thus better coding performance, as shown in Fig. 31, and an example of the original depth camera frame is shown in Fig. 32. To ensure synchronization between different cameras, all cameras are triggered by a single series rectangular wave at 20 Hz[1]. This system can be easily extended to the scenario with $L > 4$ cameras.

The cameras are calibrated before shooting video sequences. While calibration for cameras with identical resolution has been well investigated [57] and thus can be accurately done, calibration between an HD-resolution texture camera and a QCIF-resolution depth camera can have significant calibration error. Therefore, after apply-

[1] The hardware setup was implemented at AT&T Labs-Research, Florham Park, NJ 07932.

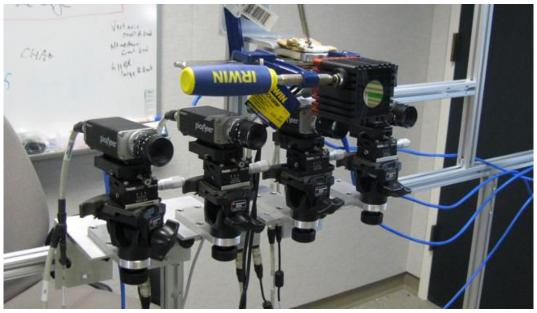


Fig. 31. The camera system setup for the collection of four correlated video sequences.

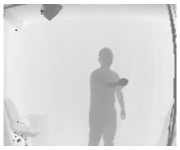


Fig. 32. An original depth frame with QCIF resolution.

ing the calibration method in [57], to reduce the calibration error and the inaccuracy brought by upsampling of the depth map to HD resolution, the depth sequence needs to be compressed before transmission, and then further refined after decoding before using it to assist pixel warping between different views.

- **Compression of depth sequence:** Several approaches have been proposed to compress the depth sequence (see e.g., [52, 53, 54]). In our setup, we use H.264/AVC to compress the depth sequence (to R_d bytes with distortion D_d) for simplicity. The distortion D_d measures the inaccuracy between the true depth map and decoded depth map. Therefore, given decoded depth sequence with distortion, the accuracy of decoder end pixel mapping between two camera views depends on D_d and thus R_d. A high quality depth sequence costs more rate to encode, but gives better decoder side information and thus lowering the rate R_t for the texture sequence, and *vice versa*. Therefore, a tradeoff has to be made between R_t and R_d. Given a fixed texture sequence quantization parameter (QP) q_t, R_t and R_d are both functions of depth sequence QP q_d. Thus, we need to solve the optimization problem

$$R^* = \min_{q_d} R_t(q_d) + R_d(q_d), \tag{5.26}$$

to achieve the best compression performance.

In practice, since QP's are integers, the problem can be solved by searching over q_d. For example, given q_d, compute $R(q_d) = R_t(q_d) + R_t(q_d)$ and $R(q_d + \Delta q_1) = R_t(q_d + \Delta q_1) + R_t(q_d + \Delta q_1)$ using MT video coding scheme. If $R(q_d) > R(q_d + \Delta q_1)$, then compute $R(q_d + \Delta q_1 + \Delta q_2)$; otherwise compute $R(q_d - \Delta q_1)$. This process is performed until a minimum rate $R\left(q_d^\star\right)$ is achieved (both $R\left(q_d^\star - \Delta q_n\right)$ and $R\left(q_d^\star + \Delta q_n\right)$ are larger than $R\left(q_d^\star\right)$ for small Δq_n).

In our experiment with *argon*, we search over different depth sequence QP's at an average texture sequence PSNR of 47.0 dB for the first GOP. As shown in Fig. 33, the optimal rate is achieved at $q_d^\star = 28$, with $R_d(28) = 656$ bytes (for the first GOP).

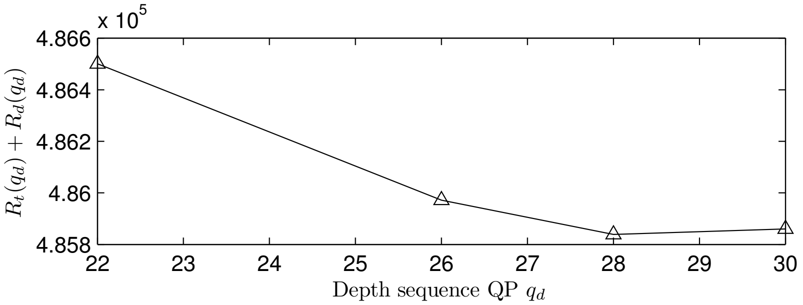


Fig. 33. MT video coding at rate $R_t(q_d)+R_d(q_d)$ with depth camera vs. depth sequence QP q_d for the first GOP.

- **Refinement of depth sequence:** We refine the low-resolution depth sequence frames from the depth camera to fit their corresponding HD frames. A decoded depth map typically has both calibration and upsampling error, which means the pixel position error in the upsampled depth map can be quite large. It is thus not practical to employ existing refinement algorithms directly. We thus devise a successive bilateral filtering refinement algorithm, based on the layer-wise algorithm of [58].

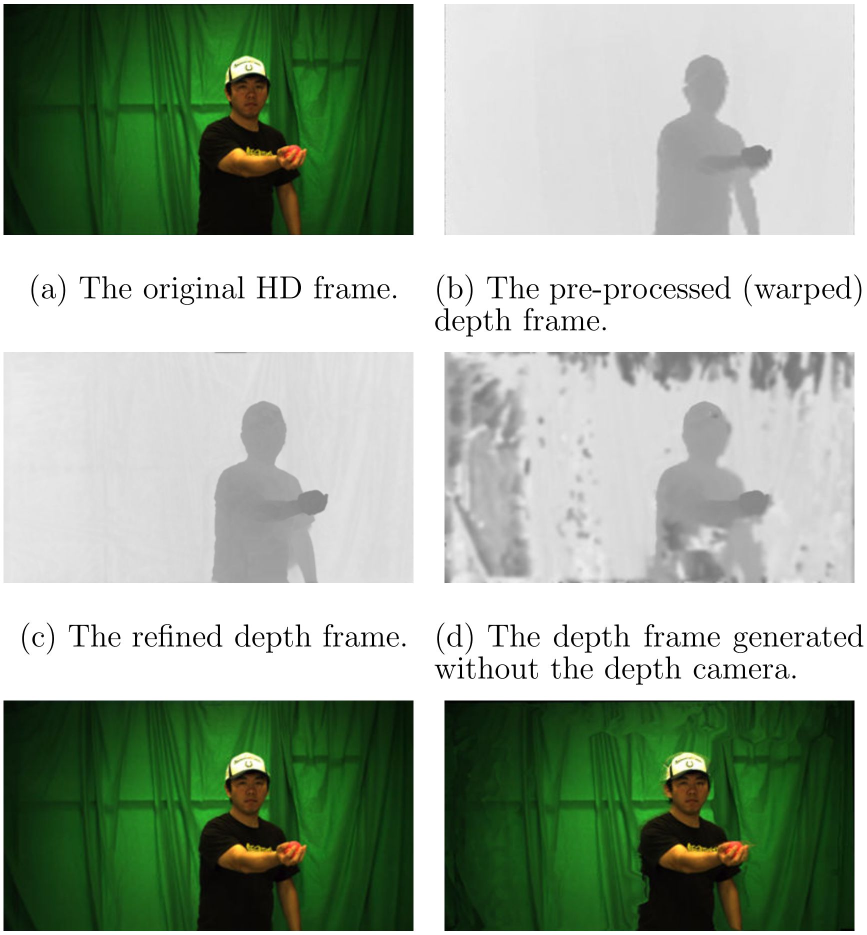
Our successive algorithm can be viewed as a k-step bilateral filtering refinement method when the HD frame has m^k times resolution (in width or height) of the depth frame. In each refinement step, the depth map is upsampled only m times before filtering to ensure the calibration and upsampling error is limited and thus corrected by the fixed-window-size bilateral filter.

It should be noted that the assistance of decoded depth maps is that it provides a relatively accurate initialization for the iterative refinement algorithm. Therefore, even highly-quantized depth maps will not deteriorate the refined depth map much, since the quantization error can be compensated by stereo matching.

Our algorithm can also be used when the depth camera is turned off. In this case it becomes one of the stereo matching algorithms for multiple view vision. This facilitates comparing depth camera assisted MT video coding with the case when there is no depth camera.

For successive depth refinement, since the scene range of the depth camera is larger than that of the HD cameras and the depth camera has a resolution between 10 to 20 times lower than the HD cameras (this can be seen by comparing Fig. 32, Fig. 34(a), and Fig. 34(b)), we choose $k = 2$ and $m = 4$, which means we first refine the depth map to 1/4 of the HD size of the texture views, then upsample it to full-HD size to make a successive refinement. We also utilize left and right HD frames (when available) that can be warped to the current camera view using the calibration parameters, in order to help refinement of the current depth map.

An example of the successive depth refinement result is shown in Fig. 34. The effectiveness of depth camera assistance can be easily seen by comparing Figs. 34(c) and 34(d). We can also see from Figs. 34(c) and 34(b) that the refined depth map is much closer to the true depth distribution than the original one.



(a) The original HD frame.

(b) The pre-processed (warped) depth frame.

(c) The refined depth frame.

(d) The depth frame generated without the depth camera.

(e) Side information with depth camera help.

(f) Side information without depth camera help.

Fig. 34. An example of depth map refinement.

For MT video coding of the four HD texture sequences (partially shown in Figs. 34), each sequence has 100 frames, which are divided into 10 GOPs, each with 10 frames using an “IPPP...” structure. We set QP = 22 and follow the H.264/AVC scheme, the only difference is that Wyner-Ziv coding compresses to the SW rate (or conditional entropy), whereas H.264/AVC compresses to the self entropy of the quantization indices. Since all the cameras are fixed, we assume that the camera parameters are known to the decoder before the decoding process.

With the help of depth sequence (coded with QP=28), the average PSNR (over 100 frames) of the side information for Wyner-Ziv coding of $\mathcal{H}_3$ (the last center view) is 38.6 dB. In contrast, the corresponding average PSNR is 31.7 dB without depth, i.e., the side information is solely generated from the previously decoded texture frames. In Fig. 34, compared with the original HD texture frame (Fig. 34(a)), we can see that the quality of a side information frame with depth camera assistance (Fig. 35(d)) is much higher in both background and foreground than that without depth camera assistance (Fig. 34(f)).

We compare MT video coding (with and without depth sequence) with both simulcast and JMVM coding, because the former gives an upper sum-rate limit of MT video coding, while the latter provides a loose lower bound on the sum-rate. The sum-rate comparisons and percentage of rate savings over H.264/AVC based simulcast are given in Table V, where MT I denotes the MT coding scheme without the help of depth camera at the decoder end, and MT II denotes the MT coding scheme with this help. And for MT II, a rate of 6933 bytes for the depth sequence is included when counting the total bit rate. In each case, the same average PSNR’s of 46.82 dB, 47.23 dB, 47.14 dB, and 47.21 dB for the four texture sequences respectively are achieved.

From Table V, we see that:

- MT video coding with depth sequence only gains 1.43% in sum-rate over that without depth sequence, even though the quality of the decoder side information with the help of a depth camera significantly improves that without the depth camera (as seen in Fig. 34). This means that much needs to be done to improve the performance of our MT video coding scheme − in terms of turning better side information quality into larger sum-rate savings.

- MT video coding with depth sequence saves 2.59% in sum-rate over simulcast, whereas JMVM does 6.98% better. This underlines the difficulty of significantly outperforming simulcast with both distributed MT video coding and joint JMVM coding, largely due to the inaccuracy of the depth information.

- The R-D performance improvements (for both MT II over MT I, and MT over simulcast) are achieved at the cost of higher complexity at the decoder, the encoding complexity of MT I and MT II stays roughly the same, which complies with the application requirement of MT video coding.

2. Experiment on standard MVC test sequences *akko&kayo* and *rena*

We also implemented our scheme for the standard MVC test sequence sets *akko&kayo* (5 views) and *rena* (8 views) with resolution 640 × 480 and sequence length 300. The two sets of sequences are provided with intrinsic and extrinsic matrices for each view, and thus frame warping is available. Since we compare the performance of our MT scheme with JM reference software for simulcast encoding and JMVM reference software for joint encoding, and only bi-directional temporal prediction is allowed in the prediction structure of the current JMVM, we need to make the temporal prediction structure identical for the three schemes for fair comparison. Therefore we follow the standard MVC prediction temporal structure (one anchor frame for

Table V. Sum-rates comparison of different schemes.

GOP	**Sum-rate (in bytes)**			
number	**Simulcast**	**MT I**	**MT II**	**JMVM**
1	496436	489893	482118	454844
2	623899	616986	610627	583718
3	627720	620919	611322	585217
4	528388	521036	511645	484062
5	559221	552213	541295	517049
6	650392	643889	636377	611772
7	665571	658950	651919	628064
8	599016	591632	582869	554862
9	564457	556905	546755	519087
10	664972	658280	650261	623988
Total	5980072	5910703	5825188	5562663
Savings	–	**1.16%**	**2.59%**	**6.98%**

every 8 frames and hierarchical bi-directional prediction is used for the 7 non-anchor frames between two nearest anchor frames) and “IPPP” mode is used for inter-view prediction. Other H.264/AVC settings include anchor frame QP 22, high profile with FRExt off, luma only mode, and CAVLC for entropy coding.

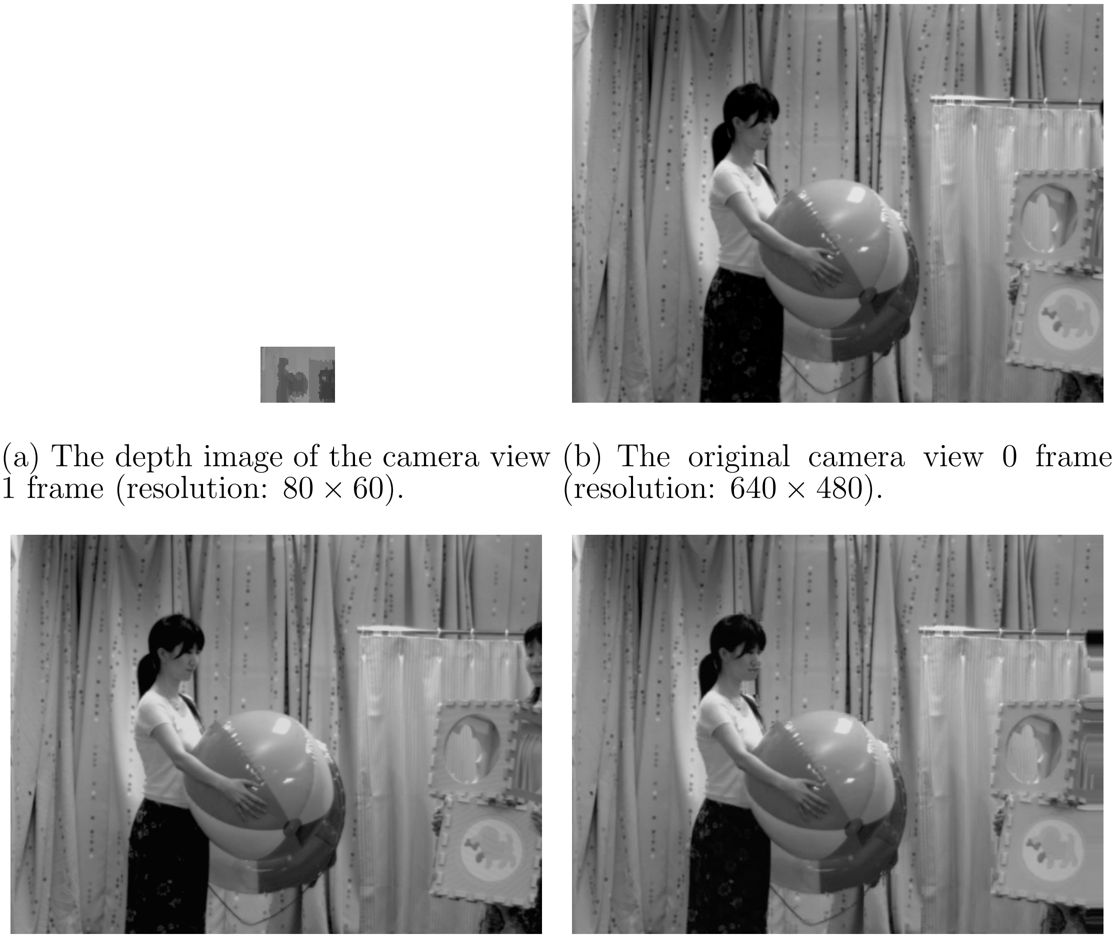
Since the actual depth data is not available for these two sets for sequences, we semi-manually generated downsampled depth images (with resolution 80 × 60) for each anchor frame that is WZ coded. The depth images are generated using segmentation and block pixel matching (since the camera parameters are given, the distances between matched pixels can be transferred to depth values). Since the camera position configuration is simple and the correlation between different views is relatively high, even the depth generated is coarse, the accuracy of frame warping is still acceptable. Thus, we expect that if more advanced depth cameras can be provided, the performance of our scheme can be largely improved.

An example of depth image assisted frame warping can be seen in Fig. 35 and Fig. 36. We can see that although the depth image is rather inaccurate in shape (mainly because the downsampling process), the quality of side information frame is still acceptable. For comparison, we also tried the MPEG 3DV depth estimation software DERS [59] for depth generation, and the resulting depth and side information frames are shown in Fig. 37 (we only provided results for *akko&kayo*, since DERS software is not designed for the camera configuration in *rena*, which is not configured with parallel but angled cameras). We can see that though the depth information generated by DERS is more detailed, the visual quality of side information frame does not differ much from that of our approach. And the actual PSNR of side information frame by DERS is nearly 5 dB lower than that of our approach, since whereas the shape of depth image is close to the truth, the depth values are not as accurate as those of our method. In addition, since it is more detailed, transmitting the depth image generated by DERS costs much more bits (59,928 bits by H.264/AVC lossless encoding for the frame in Fig. 35(a), compared to 2,030 bits for the frame in Fig. 35(a)).

The comparison of average side information frame quality (measured in PSNR) with that of decoded frames is shown in Table VI. It can be seen that the PSNR of side information frames is about 10 dB lower than that of decoded frames. This also indicates that in the joint estimation step, the weight of decoded frames should be much larger than that of side information frames. For simplicity, in the joint estimation process, we use linear combination of a decoded frame $\tilde{\mathcal{H}}_{i,t}$ and a side information frame $\mathrm{SI}_{i,t}^{(i-1)}$ from the nearest view, i.e.,

$$\hat{\mathcal{H}}_{i,t} = c_{\mathrm{c}}\tilde{\mathcal{H}}_{i,t} + c_{\mathrm{s}}\mathrm{SI}_{i,t}^{(i-1)}, \tag{5.27}$$

and the coefficients acquired by training, and the PSNR gain for using joint estimation is 0.12 dB for *akko&kayo* and 0.21 dB for *rena* with depth camera assistance, which corresponds to bit rate savings of 2.11% and 3.47% of total bit rate, respectively.



(a) The depth image of the camera view 1 frame (resolution: 80 × 60).

(b) The original camera view 0 frame (resolution: 640 × 480).

(c) The camera view 1 frame to be WZ coded.

(d) The side information frame for 35(c) warped from camera view 0 frame (PSNR: 32.24dB).

Fig. 35. An example of depth image assisted frame warping, from sequence set *akko&kayo*, camera view 0 and 1, first frame.

Table VI. Comparison of average frame quality (in PSNR, dB) between decoded frames and corresponding side information frames (with depth camera assistance). Anchor frame QP is set at 22.

	akko&kayo (5 views)		*rena* (5 views)	
	Anchor	Non-anchor	Anchor	Non-anchor
Dec.	44.15	42.56	46.89	45.05
SI	32.24	30.75	37.67	36.96

The bit rate saving comparisons can be see in Table VII and VIII, in which MT I and MT II correspond to scheme without and with depth camera assistance respectively. In this table, it is shown that with depth camera assistance, the sum rate saving is improved by 4% on average, even at the cost of extra bit rate for depth information. We can also see that the sum rate saving achieved by the proposed MT scheme is about half of that achieved by the joint scheme. This means that the sum rate loss of the MT scheme over joint scheme is 8.53% of the total rate on average of the two sets of sequences. The sum rate loss is more significant than that in quadratic Gaussian MT source stated in Chapter IV, and the main reason is that both the sources themselves and correlation between different sources are much more complicated (with non-stationary distribution, spatial and temporal memory, etc.) than the model in theoretical analysis. Another cause is the inaccuracy of depth information, which is relatively coarse and thus the mapping error exists, especially at object edges.

It also can be seen that while the MT anchor frame saving is smaller than that of joint scheme (Since we only transmit depth information for anchor frames, the bit rate cost of depth image is counted in anchor frame bit rates in the Table VII and VIII), MT scheme achieves better savings for non-anchor frames compared to joint scheme (spatial prediction for non-anchor frames gain little). This is mainly due to the fact that joint estimation can be used for non-anchor frames with the presence of depth information, thus the information from other views can be better exploited.

Table VII. Bit rates (in bytes) of different schemes and their rate savings compared to the simulcast scheme at the same target average PSNR (sequence *akko&kayo*, 5 views).

	Anchor	Non-anc.	Total	Saving
Simul.	4610772	3077424	7688196	—
MT I	4385191	2877692	7262883	5.53%
MT II	4097542	2877692	6975234	9.27%
Joint	3178183	3071071	6249254	18.72%

Table VIII. Bit rates (in bytes) of different schemes and their rate savings compared to the simulcast scheme at the same target average PSNR (sequence *rena*, 8 views).

	Anchor	Non-anc.	Total	Saving
Simul.	4171724	3592631	7764355	—
MT I	4100289	3216446	7316735	5.77%
MT II	3780787	3216446	6997233	9.88%
Joint	2796082	3609400	6405482	17.50%

Detailed MT saving results with depth camera assistance is shown in Table IX and X, in which bit rate savings for different components in the bitstream are provided. From the table, we can see that by providing depth information at the decoder, we transmit 1.29% and 0.92% more bit rate for the two sequence sets respectively. However, this rate loss can be compensated by much more rate savings acquired from higher side information quality, e.g., the joint estimation is not applicable without depth information at the decoder, and the rate savings from joint estimation only are already more than the bit rate for depth information for each sequence set.

Table IX. Rate savings (in bytes) achieved by different components in the bit stream (sequence *akko&kayo*). Average mutual information is provided for WZ coded components.

Component	Mutual info.	Bytes Saved	% saved
Intra mode	0.23	34314	0.44%
Anchor DCT coeff.	0.40	395481	5.09%
Depth info.	—	-71729	-0.92%
Inter mode	0.09	7902	0.10%
MVD	0.15	24604	0.32%
Non-anc. DCT coeff.	0.24	107006	1.38%
Joint Estimation	—	269544	3.47%

Additionally, we fix the average transmission rate (*akko&kayo* at 1.2 Mbps per view, *rena* at 0.6 Mbps per view, at frame rate of 30 fps) and compare the PSNR performance vs. different frames. The result is shown in Fig. 38. and 39

Remarks:

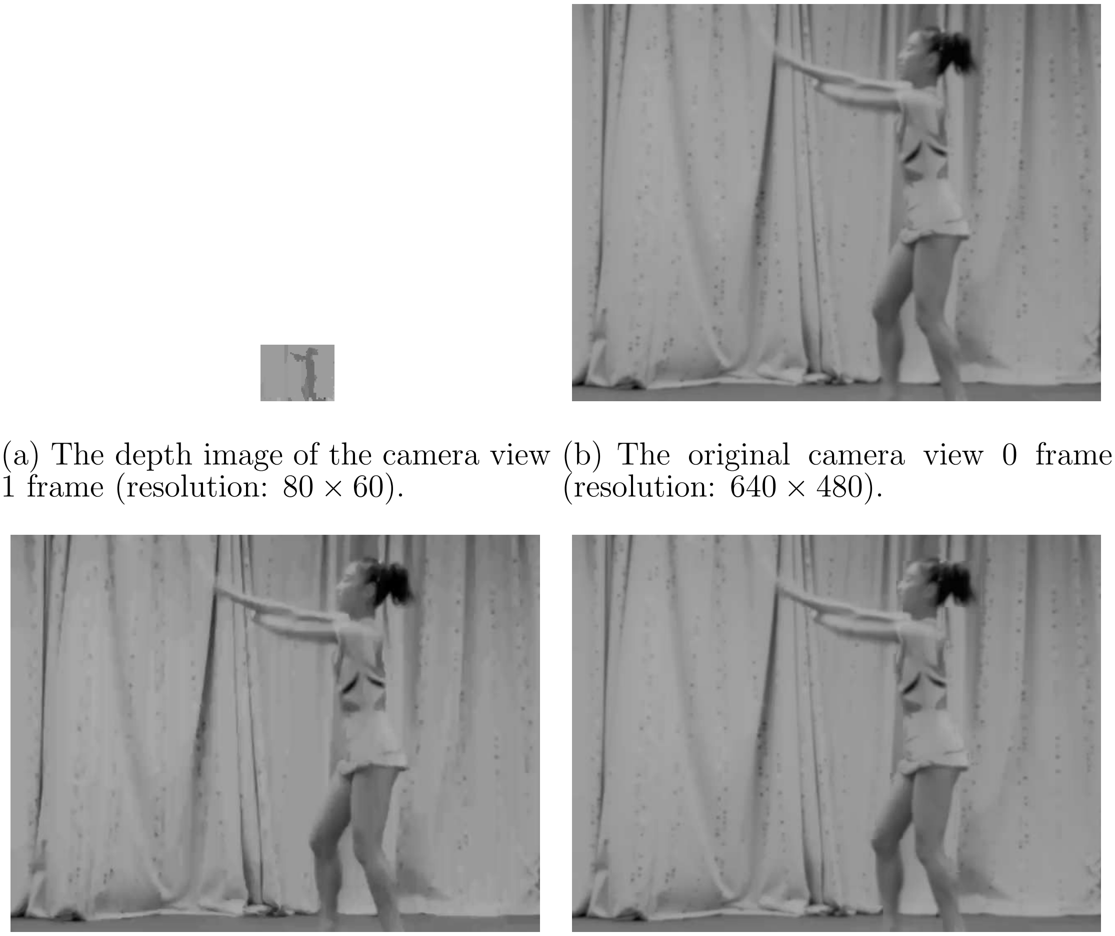
- The current depth information is generated by processing video frames from different views jointly. However, since we encode and transmit it separately, this scheme is still a MT scheme if we consider the depth information as another video source, which contains the geometrical relation between different texture sequences. Particularly, if the depth information can be automatically collected

Table X. Rate savings (in bytes) achieved by different components in the bit stream (sequence *rena*). Average mutual information is provided for WZ coded components.

Component	Mutual info.	Bytes Saved	% saved
Intra mode	0.31	74243	0.97%
Anchor DCT coeff.	0.58	507657	6.59%
Depth info.	—	-99238	-1.29%
Inter mode	0.08	5953	0.08%
MVD	0.13	24313	0.32%
Non-anc. DCT coeff.	0.09	38026	0.49%
Joint Estimation	—	162009	2.11%

and synchronized with the texture sequences (new devices, such as depth cameras, are available now, but well calibrated and synchronized sequences are still rare), the application of the proposed scheme will becomes straightforward.

- By transmitting additional depth information to the decoder, the R-D performance of MT video coding can be improved. This indicates that the correlation between different camera view sequences can be more thoroughly exploited by providing their geometrical relations, which help to acquire pixel-to-pixel correspondence between simultaneous frames from different camera views. If such correspondence is accurate, the correlation model between different views will become much simpler since only pixel value noise that is independent between different views needs to be considered. Therefore, the advantage of using depth information should not only benefit MT video coding, but also become a more efficient means of representation for other applications with multiple correlated video sequences.



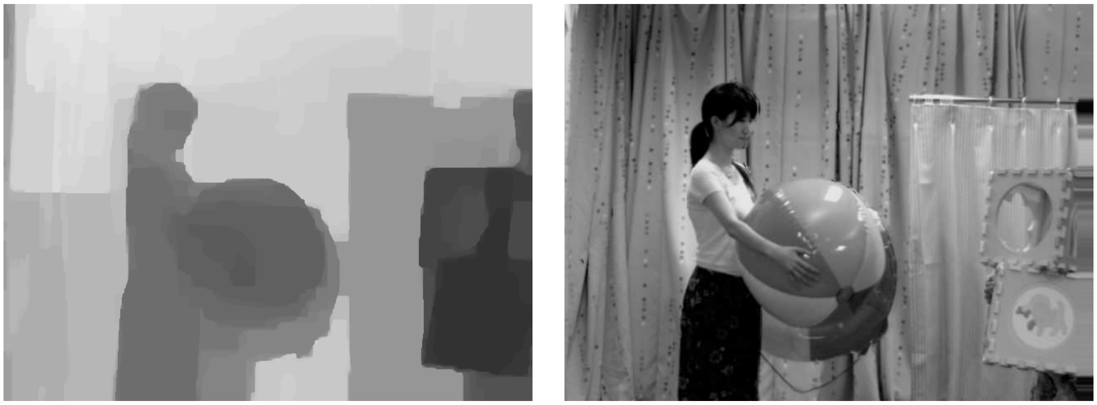
(a) The depth image of the camera view 1 frame (resolution: 80 × 60).

(b) The original camera view 0 frame (resolution: 640 × 480).

(c) The camera view 1 frame to be WZ coded.

(d) The side information frame for 36(c) warped from camera view 0 frame (PSNR: 37.67dB).

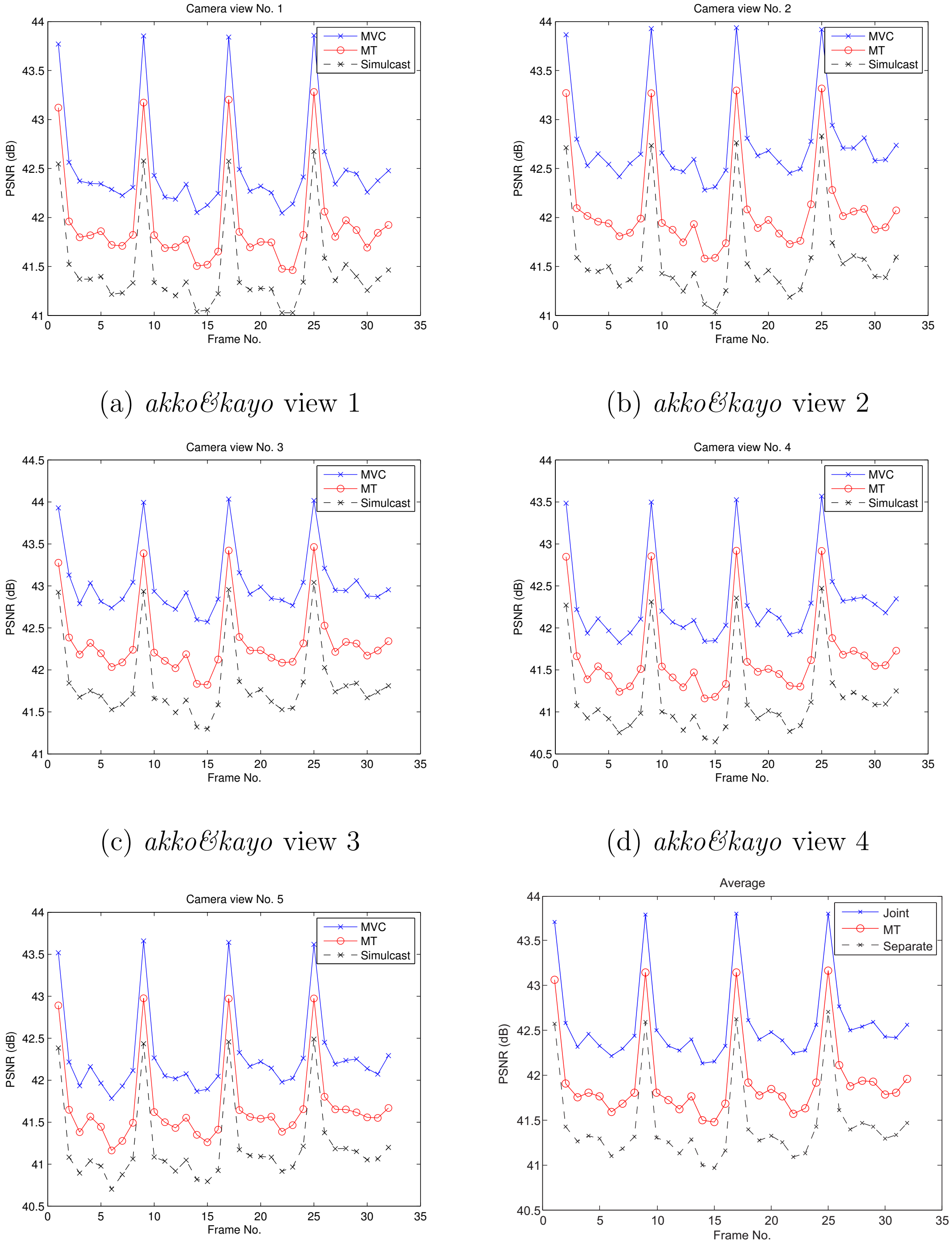
Fig. 36. An example of depth image assisted frame warping, from sequence set *rena*, camera view 0 and 1, first frame.



(a) The depth image of camera view 1 frame generated by DERS.

(b) The side information frame for 35(c) warped from camera view 0 frame by DERS (PSNR 27.38dB).

Fig. 37. Depth and side information frame using DERS for depth estimation.



(a) *akko&kayo* view 1

(b) *akko&kayo* view 2

(c) *akko&kayo* view 3

(d) *akko&kayo* view 4

(e) *akko&kayo* view 5

(f) *akko&kayo* average

Fig. 38. Comparison of PSNR (in dB) vs. frame number for simulcast, MT and joint schemes. First GOP, 1 anchor frame followed by 7 hierarchically bi-predicted non-anchor frames. Sequence *akko&kayo*.

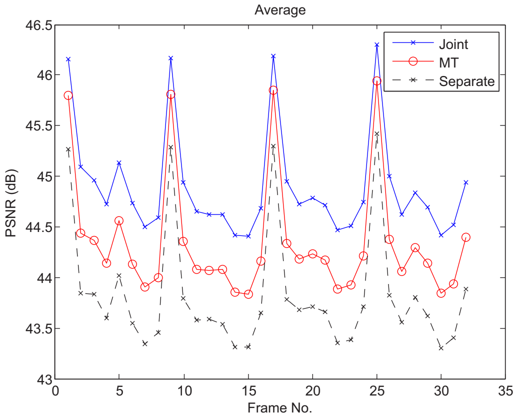


Fig. 39. Comparison of PSNR (in dB) vs. frame number for simulcast, MT and joint schemes. First GOP, 1 anchor frame followed by 7 hierarchically bi-predicted non-anchor frames. Sequence *rena*.

CHAPTER VI

CONCLUSIONS

In this dissertation, the theory and application of MT video coding is analyzed in detail. First, two theoretical results in quadratic Gaussian MT source coding are shown. A new sufficient condition is provided for BT sum-rate tightness. And the behavior of sum-rate loss in quadratic Gaussian MT source coding is also analyzed.

Following the theoretical results and extending the code designs in [17] for quadratic Gaussian two-terminal source coding, this dissertation also proposed the first code design for quadratic Gaussian MT direct and indirect source coding problems that have recently been shown to have tight sum-rate bound. TCQ/TCVQ and LDPC codes are employed to approach corner points of the rate region. A model-based approximation to the LLR distribution is provided to simplify the LDPC code design with no additional rate loss. Simulation results show that in the three- and four-terminal cases in the high-rate scenario, the sum-rate loss due to practical coding can be achieved as low as 0.106 b/s for a transmission rate of 9.095 b/s, while in the low-rate scenario the sum-rate loss is 0.146 b/s for a sum rate of 4.131 b/s. The rate loss in the low-rate scenario is relatively higher due to the smaller granular gain of TCVQ.

In accordance with the advancement of research on RD analysis in MT source coding theory, and the development of distributed video sensor network and its 3-D applications, we provided detailed analysis and experiment results for MT video coding with depth information separately transmitted to the decoder under the H.264/AVC framework in this dissertation. By utilizing the depth information to acquire better decoder side information, we are able to achieve an average sum rate saving of about 9.58% over simulcast scheme implemented by JM reference software, about 4% better than that without depth camera. Comparison to MVC scheme implemented by JMVM reference software shows that MT scheme suffers a sum rate loss of 8.53%, which conforms with the result in MT source coding theory. Moreover, given depth information, joint estimation can also be performed to improve the quality of reconstructed sequence frames for both MT and MVC schemes, which indicates the importance of using additional depth information in 3D video applications. Since the depth information we use is still simple and relatively coarse, we expect that if more advanced depth information collecting devices are equipped in multiview video sequence acquisition, the performance of MT video coding could be further improved, and even joint video coding scheme other applications, e.g., 3D-TV, free view-point TV, etc., would also benefit from this.

REFERENCES

[1] T. Berger, “Multiterminal source coding,” in *The Information Theory Approach to Communications.* Springer-Verlag, G. Longo, Ed., New York, NY, 1977.

[2] D. Slepian and J. Wolf, “Noiseless coding of correlated information sources,” *IEEE Trans. Inform. Theory*, vol. 19, pp. 471-480, Jul. 1973.

[3] A. Wyner, “The rate-distortion function for source coding with side information at the decoder,” *IEEE Trans. Inform. Theory*, vol. 20, pp. 1-10, Jan. 1976.

[4] S. Tung, “Multiterminal rate-distortion theory,” Ph.D. dissertation, School of Electrical Engineering, Cornell University, Ithaca, NY, 1978.

[5] T. Berger, K. B. Housewright, J. K. Omura S. Tung and J. Wolfowitz, “An upper bound on the rate distortion function for source coding with partial side information at the decoder,”, *IEEE Trans. Inform. Theory*, vol. 25, pp. 664-666, 1979.

[6] H. Yamamoto and K. Itoh, “Source coding theory for multiterminal communication system with a remote source,” *Trans. IECE Japan*, vol. E63, pp. 700-706, Oct. 1980.

[7] T. Flynn and R. Gray, “Encoding of correlated observations,” *IEEE Trans. Inform. Theory*, vol. 33, pp. 773-787, Nov. 1987.

[8] Y. Yang and Z. Xiong, “On the generalized Gaussian CEO problem,” *IEEE Trans. Inform. Theory*, vol. 58, June 2012.

[9] Y. Oohama, "Distributed source coding for correlated memoryless Gaussian sources," submitted to *IEEE Trans. Inform. Theory.* Available at http://arxiv.org/PS_cache/arxiv/pdf/0908/0908.3982v4.pdf.

[10] Y. Oohama, "Distributed source coding of correlated Gaussian sources", submitted to *IEEE Trans. Inform. Theory.* Available at http://arxiv.org/PS_cache/arxiv/pdf/1007/1007.4418v2.pdf.

[11] Y. Oohama, "Rate-distortion theory for Gaussian multiterminal source coding systems with several side informations at the decoder," *IEEE Trans. Inform. Theory*, vol. 51, pp. 2577-2593, Jul. 2005.

[12] A. Wagner, S. Tavildar, and P. Viswanath, "The rate region of the quadratic Gaussian two-terminal source-coding problem," *IEEE Trans. Inform. Theory*, vol. 54, pp. 1938-1961, May 2008.

[13] J. Wang, J. Chen, and X. Wu, "On the sum rate of Gaussian multiterminal source coding: New proofs and results," *IEEE Trans. Inform. Theory*, vol. 56, pp. 3946-3960, Aug. 2010.

[14] Y. Yang and Z. Xiong, "The sum-rate bound for a new class of quadratic Gaussian multiterminal source coding problems," *IEEE Trans. Inform. Theory*, vol. 58, February 2012.

[15] T. Cover and J. Thomas, *Element of Information Theory*, Wiley, John & Sons, Inc., 1991.

[16] S. Pradhan and K. Ramchandran, "Generalized coset codes for distributed binning," *IEEE Trans. Inform. Theory*, vol. 51, pp. 3457-3474, Oct. 2005.

[17] Y. Yang, V. Stanković, Z. Xiong, and W. Zhao, "On multiterminal source code design," *IEEE Trans. Inform. Theory*, vol. 54, pp. 2278-2302, May 2008.

[18] M. Marcellin and T. Fischer, "Trellis coded quantization of memoryless and Gaussian-Markov sources," *IEEE Trans. Communications*, vol. 38, pp. 82-93, Jan. 1990.

[19] T. Fischer, M. Marcellin, and M. Wang, "Trellis-coded vector quantization," *IEEE Trans. Inform. Theory*, vol. 37, pp. 1551-1566, Nov. 1991.

[20] Y. Yang, S. Cheng, Z. Xiong, and W. Zhao, "Wyner-Ziv coding based on TCQ and LDPC codes," *IEEE Trans. Communications*, vol. 54, pp. 2278-2302, May 2008.

[21] X. Zhu, A. Aaron and B. Girod, "Distributed compression for large camera arrays," *Proc. IEEE Workshop on Statistical Signal Proc.*, St. Louis, MO, Sept. 2003.

[22] M. Flierl and B. Girod, "Coding of multi-view image sequences with video sensors," *Proc. of ICIP'06*, pp. 609-612, Atlanta, GA, Oct. 2006.

[23] M. Flierl and P. Vandergheynst, "Distributed coding of highly correlated image sequences with motion-compensated temporal wavelets," *EURASIP J. Applied Signal Proc.*, Article ID 46747, 2006.

[24] X. Guo, Y. Lu, F. Wu, W. Gao, and S. Li, "Distributed multi-view video coding," *Proc. SPIE VCIP*, San Jose, CA, Jan. 2006.

[25] I. Tosic and P. Frossard, "Geometry-based distributed scene representation with omnidirectional vision sensors," *IEEE Trans. Image Processing*, vol. 17, pp. 1033-1046, Jul. 2008.

[26] N. Gehrig and P. Dragotti, "Geometry-driven distributed compression of the plenoptic function: Performance bounds and constructive algorithms," *IEEE Trans. Image Processing*, vol. 18, pp. 457-470, Mar. 2009.

[27] T. Wiegand, G. Sullivan, G. Bjøtegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 560–576, Jul. 2003.

[28] Y. Yang, V. Stanković, Z. Xiong, and W. Zhao, "Two-terminal video coding," *IEEE Trans. Image Processing*, vol. 18, pp. 534-551, Mar. 2009.

[29] Y. Zhang, Y. Yang, and Z. Xiong, "Three-terminal video coding," in *Proc. IEEE Multimedia Signal Processing Workshop*, Rio de Janeiro, Brazil, October 2009.

[30] MESA Imaging, http://www.mesa-imaging.ch/.

[31] H.264/AVC reference software JM 18.0, available at http://iphome.hhi.de/suehring/tml/download/.

[32] "Joint multiview video model (JMVM) 8.0," JVT-AA207, Geneva, Switzerland, Apr. 2008.

[33] The JMVM (joint multiview video model) software, Jun. 2008. http://iphome.hhi.de/suehring/tml/download/.

[34] Y. Oohama, "Gaussian multiterminal source coding," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1912-1923, Nov. 1997.

[35] A. Wagner and V. Anantharam, "An improved outer bound for multiterminal source coding," *IEEE Trans. Inform. Theory*, vol. 54, pp. 1919-1937, May 2008.

[36] J. Chen, X. Zhang, T. Berger, and S. B. Wicker, "An upper bound on the sum-rate distortion function and its corresponding rate allocation schemes for the CEO problem," *IEEE J. Select. Areas in Comm.*, vol. 22, pp. 977-987, Aug. 2004.

[37] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, MA: Athena Scientific, 1999.

[38] S. Boyd, L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge 2004.

[39] R. Zamir, "The rate loss in the Wyner-Ziv problem," *IEEE Trans. Inform. Theory*, vol. 42, pp. 2073-2084, Nov. 1996.

[40] H. Feng, "On the rate loss of multiterminal source codes," *Proc. ISIT-2006*, Seattle, WA, July 2006.

[41] R. Zamir and T. Berger, "Multiterminal source coding with high resolution," *IEEE Trans. Inform. Theory*, pp. 106-117, Jan. 1999.

[42] I. Nowak, *Relaxation and decomposition methods for mixed integer nonlinear programming*, Birkhäuser, Basel, 2005.

[43] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*, Wiley, John & Sons, Inc., 2006, 3rd Edition.

[44] Y. Yang and Z. Xiong, "The supremum sum-rate loss of quadratic Gaussian direct multiterminal source coding," *Proc. UCSD Workshop on Information Theory and its Applications*, San Diego, CA, Jan. 2008.

[45] J. Wolf, "Data reduction for multiple correlated sources," *Proc. 5th Colloquium Microwave Communication*, pp. 287-295, June 1973.

[46] Y. Sun, Y. Yang, A. Liveris, V. Stanković, and Z. Xiong, "Near-capacity dirty-paper code design: A source-channel coding approach," *IEEE Trans. Inform. Theory*, vol. 55, pp. 3013-3031, Jul. 2009.

[47] K. Price, and R. Storn, "Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces," *J. Global Optimiz.*, vol. 11, pp. 341-359, 1997.

[48] D. N. C. Tse and S. V. Hanly, "Multiaccess fading channels-part I: polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2796-2815, Nov. 1998.

[49] R. Zamir, S. Shamai, and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1250-1276, Jun. 2002.

[50] J. Chen, and T. Berger, "Successive Wyner-Ziv coding scheme and its application to the quadratic Gaussian CEO problem,", *IEEE Trans. Inform. Theory*, vol. 54, pp. 1586-1603, Apr. 2008.

[51] T. Richardson, M. Amin Shokrollahi, and R. Urbanke, "Design of capacity-approaching irregular low-density parity check codes,", *IEEE Trans. Inform. Theory*, vol. 47, pp. 619-637, Feb. 2001.

[52] H. Oh and Y. Ho, "H.264-based depth map sequence coding using motion information of corresponding texture video," *LNCS*, pp. 898-907, Dec. 2006.

[53] M. Kang, C. Lee, J. Lee and Y. Ho, "Adaptive geometry-based intra prediction for depth video coding," *Proc. ICME*, pp. 1230-1235, Jul. 2010.

[54] H. Yuan, Y. Chang, J. Huo, F. Yang, and Z. Lu, “Model-based joint bit allocation between texture videos and depth maps for 3-D video coding,” *IEEE Trans. Circuits and Systems for Video Technology*, vol. 21, Apr. 2011.

[55] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 1st ed., Cambridge University Press, 2000.

[56] D. Scharstein and R. Szeliski, “A Taxonomy and evaluation of dense two-frame stereo correspondence algorithms”, *International Journal of Computer Vision*, pp.7-42, 2002.

[57] http://www.vision.caltech.edu/bouguetj/calib_doc/.

[58] Q. Yang, R. Yang, J. Davis, and D. Nistér, “Spatial-depth super resolution for range images,” *Proc. CVPR’07*, Minneapolis, MN, Jun. 2007.

[59] “3DV depth estimation and view synthesis software package”, ISO/IEC JTC1/SC29/WG11, MPEG2011/N12188, available at http://mpeg.chiariglione.org/working_documents/explorations/3dav/DERS.zip.

APPENDIX A

PROOF OF LEMMA 12

Proof. To prove Lemma 17, we use a rigorous numerical method that is based on the following two propositions.

Proposition 2. *Denote the optimal* p *in* $\mathbb{P}_2(x)$ *for fixed* $(x, \lambda) \in [0,1)^2$ *as*

$$p^*(x,\lambda) \ \triangleq \ \arg \max_{p \in (0,1] : g_2(\lambda, p\,; x) \geq 0} f_2(\lambda, p\,; x), \tag{A.1}$$

then $p^*(x, \lambda)$ *must satisfy*

$$p^*(x,\lambda) \ = \ \min(p^f(x,\lambda), p^g(x,\lambda)), \tag{A.2}$$

where $p^f(x,\lambda)$*,* $p^g(x,\lambda)$ *are the solutions to (A.3) and (A.4) as follows.*

$$\frac{p(2-4x\lambda+4p\lambda+2x\lambda^2-5xp\lambda^2+xp\lambda^3-2xp^2\lambda^3+2p^2\lambda^2)}{(1+p\lambda)^2(1-x)} \cdot \ln\left(\frac{(1-x\lambda)p}{1-x+(1-x\lambda)p}\right) = 1, \tag{A.3}$$

$$1 - x - p + p\lambda - p^2\lambda = 0. \tag{A.4}$$

Proof. In this proof, we denote $f = f_2(\lambda, p\,; x)$, $g = g_2(\lambda, p\,; x)$. First, we show that for any fixed x and λ, f is a concave function of p. In fact, (A.5) - (A.7) hold as follows,

$$\begin{aligned}
2\ln 2 \frac{\partial^2 f}{\partial p^2} &= \frac{2-4x\lambda+4p\lambda+2x\lambda^2-5xp\lambda^2+xp\lambda^3-2xp^2\lambda^3+2p^2\lambda^2}{(1+p\lambda)^2(1-x+p-px\lambda)} + \ln(\frac{(1-x\lambda)p}{1-x+(1-x\lambda)p})\cdot \\
&\quad \frac{2(1-3xp^2\lambda^3-3xp\lambda^2-2x\lambda+3p\lambda+x\lambda^2+3p^2\lambda^2-xp^3\lambda^4+p^3\lambda^3)}{(1-x)(1+p\lambda)^3} && \text{(A.5)} \\
&\leq \frac{2-4x\lambda+4p\lambda+2x\lambda^2-5xp\lambda^2+xp\lambda^3-2xp^2\lambda^3+2p^2\lambda^2}{(1+p\lambda)^2(1-x+p-px\lambda)} + \Big(-\frac{1-x}{1-x+(1-x\lambda)p}\Big)\cdot \\
&\quad \frac{2(1-3xp^2\lambda^3-3xp\lambda^2-2x\lambda+3p\lambda+x\lambda^2+3p^2\lambda^2-xp^3\lambda^4+p^3\lambda^3)}{(1-x)(1+p\lambda)^3} && \text{(A.6)} \\
&= -\frac{xp\lambda^2(1-\lambda)(3+p\lambda)}{(1+p\lambda)^3(1-x+p-px\lambda)} \ \leq \ 0. && \text{(A.7)}
\end{aligned}$$

where (A.6) is due to the fact that

$$\begin{aligned}&(1 - 3xp^2\lambda^3 - 3xp\lambda^2 - 2x\lambda + 3p\lambda + x\lambda^2 + 3p^2\lambda^2 - xp^3\lambda^4 + p^3\lambda^3) \qquad \text{(A.8)}\\ =&(1 - 2x\lambda + x\lambda^2) + (3p^2\lambda^2 - 3xp^2\lambda^3) + (3p\lambda - 3xp\lambda^2) + (p^3\lambda^3 - xp^3\lambda^4) \geq 0,\end{aligned}$$

and $\ln(1 - x) \leq -x$, with $x = \frac{1-x}{1-x+(1-x\lambda)p} \in (0, 1)$. Hence the (scaled) first order derivative

$$\begin{aligned}2\ln 2\frac{\partial f}{\partial p} &= 1 + p(2 - 4x\lambda + 4p\lambda + 2x\lambda^2 - 5xp\lambda^2\\ &\quad + xp\lambda^3 - 2xp^2\lambda^3 + 2p^2\lambda^2)(1 + p\lambda)^{-2}(1 - x)^{-1}\\ &\quad \cdot \ln\left(\frac{(1 - x\lambda)p}{1 - x + (1 - x\lambda)p}\right)\\ &\triangleq 1 + \mathscr{C}(\lambda, p, x) \cdot \ln[\mathscr{D}(\lambda, p, x)] \qquad \text{(A.9)}\end{aligned}$$

is monotonically decreasing in p. Moreover, since $\frac{\partial f}{\partial p}$ satisfies

$$\lim_{p\to 0}\left[2\ln 2\frac{\partial f}{\partial p}\right] = 1 > 0,$$

and

$$\begin{aligned}2\ln 2\frac{\partial f}{\partial p}\bigg|_{p=1} &= 1 + \frac{(2 - 4x\lambda + 4\lambda + 2x\lambda^2 - 5x\lambda^2 + x\lambda^3 - 2x\lambda^3 + 2\lambda^2)}{(1 + \lambda)^2(1 - x)}\\ &\quad \cdot \ln\left(\frac{1 - x\lambda}{1 - x + (1 - x\lambda)}\right)\\ &\leq 1 + \frac{(2 - 4x\lambda + 4\lambda + 2x\lambda^2 - 5x\lambda^2 + x\lambda^3 - 2x\lambda^3 + 2\lambda^2)}{(1 + \lambda)^2(1 - x)} \cdot \left(-\frac{1 - x}{2 - x - x\lambda}\right)\\ &= -\frac{x(1 - \lambda)}{(1 + \lambda)^2(2 - x - x\lambda)} < 0.\end{aligned}$$

we know that for any (x, λ) pair, there must be a solution to (A.3) in the range

$p \in (0,1)$, which means $p^f(x,\lambda)$ is well defined and must satisfy

$$2\ln 2 \frac{\partial f}{\partial p} \begin{cases} > \; 0 & p \in (0, p^f(x,\lambda)) \\ = \; 0 & p = p^f(x,\lambda) \\ < \; 0 & p \in (p^f(x,\lambda), 1] \end{cases} .$$

Then since $p^g(x,\lambda)$ is the solution to $g = 1-x-p+p\lambda-p^2\lambda = 0$, and g is monotonically decreasing in p, we know that $p = p^f(x,\lambda)$ must satisfy $g \geq 0$ if $p^f(x,\lambda) \leq p^g(x,\lambda)$; and f must be monotone increasing in $p \in (0, p^g(x,\lambda))$ if $p^g(x,\lambda) < p^f(x,\lambda)$. Hence (A.2) must hold for any fixed (x,λ). □

Proposition 3. *For any rectangular region* $\Omega = \left\{(x,\lambda) : \underline{x} \leq x < \overline{x}, \underline{\lambda} \leq \lambda < \overline{\lambda}\right\}$ *with* $\underline{x}, \overline{x}, \underline{\lambda}, \overline{\lambda} \in [0,1)$*, define*

$$\underline{p} = \min\left(\max(\frac{\kappa(1-\overline{x})}{1-\overline{x}\underline{\lambda}}, \underline{p}^f(\Omega)), \underline{p}^g(\Omega)\right),\ \overline{p} = \min(\overline{p}^f(\Omega), \overline{p}^g(\Omega)), \overline{w} = \frac{1-\overline{x}\underline{\lambda}}{1-\overline{x}},$$
$$\overline{y} = \frac{(1+\underline{p}\overline{\lambda}-\underline{x}-\underline{p}-\underline{p}^2\overline{\lambda})(1-\underline{x}+\overline{p}-\overline{p}\underline{\lambda x})}{(1+\underline{p}\underline{\lambda})(1-\overline{x})},$$

where

$$\kappa = -\frac{1}{2\mathscr{W}_{-1}(-\frac{1}{2\sqrt{e}})+1} \quad = \quad 0.39795255$$

is a constant satisfying

$$1 + 2\kappa \cdot \ln\left[\frac{\kappa}{1+\kappa}\right] = 0, \tag{A.10}$$

with $\mathscr{W}_{-1}(x)$ *being the Lambert W function in the branch* $[-\frac{1}{e}, 0)$*, and* $\underline{p}^f(\Omega)$*,* $\overline{p}^f(\Omega)$*,*

$\underline{p}^g(\Omega)$, $\overline{p}^g(\Omega)$ *are the solutions of* $p \in (0, 1]$ *to (A.11) - (A.14)as follows, respectively.*

$$1 + \frac{p(2 - 4\underline{x\lambda} + 4p\overline{\lambda} + 2\overline{x}\overline{\lambda}^2 - 5\underline{xp}\underline{\lambda}^2 + \overline{x}p\overline{\lambda}^3 - 2\underline{xp}^2\underline{\lambda}^3 + 2p^2\overline{\lambda}^2)}{(1 + p\underline{\lambda})^2(1 - \overline{x})} \ln\left(\frac{(1 - \underline{x}\overline{\lambda})p}{1 - \underline{x} + (1 - \underline{x}\overline{\lambda})p}\right) = 0, \tag{A.11}$$

$$1 + \frac{p(2 - 4\overline{x}\overline{\lambda} + 4p\underline{\lambda} + 2\overline{x}\underline{\lambda}^2 - 5\overline{x}p\overline{\lambda}^2 + \underline{xp}\underline{\lambda}^3 - 2\overline{x}p^2\overline{\lambda}^3 + 2p^2\underline{\lambda}^2)}{(1 + p\overline{\lambda})^2(1 - \underline{x})} \ln\left(\frac{(1 - \overline{x}\underline{\lambda})p}{1 - \overline{x} + (1 - \overline{x}\underline{\lambda})p}\right) = 0, \tag{A.12}$$

$$1 - \overline{x} - p + p\underline{\lambda} - p^2\underline{\lambda} = 0, \tag{A.13}$$

$$1 - \underline{x} - p + p\overline{\lambda} - p^2\overline{\lambda} = 0. \tag{A.14}$$

Then for any $(x, \lambda) \in \Omega$, *the optimal* $p^*(x, \lambda)$ *defined in (A.1) must satisfy*

$$\underline{p} \ \leq \ p^*(x, \lambda) \ \leq \ \overline{p}, \tag{A.15}$$

with the corresponding maximum function value upper-bounded by

$$\max_{(x,\lambda)\in\Omega} f^*(x, \lambda) \leq \overline{f}(\Omega) \triangleq \frac{\overline{x}}{2}\log_2(1 + \overline{\lambda}\overline{p}) + \frac{\overline{y}}{2}\log_2(1 + \overline{wp}) + \frac{1 - \underline{x} - \overline{y}}{2}\log_2(\overline{wp}). \tag{A.16}$$

Proof. In this proof, we again denote $f = f_2(\lambda, p; x)$, $g = g_2(\lambda, p; x)$.

In the rectangular region $(x, \lambda) \in \Omega$ where $0 \leq \underline{x} \leq x < \overline{x} < 1$ and $0 \leq \underline{\lambda} \leq \lambda < \overline{\lambda} \leq 1$, we can lower- and upper-bound $\mathscr{C}(\lambda, p, x)$ and $\mathscr{D}(\lambda, p, x)$ (defined in (A.9)) as

$$\mathscr{C}(\lambda, p, x) \mid_{(x,\lambda)\in\Omega} \geq \underline{\mathscr{C}}(\Omega, p) \triangleq \frac{p(2 - 4\overline{x}\overline{\lambda} + 4p\underline{\lambda} + 2\overline{x}\underline{\lambda}^2 - 5\overline{x}p\overline{\lambda}^2 + \underline{xp}\underline{\lambda}^3 - 2\overline{x}p^2\overline{\lambda}^3 + 2p^2\underline{\lambda}^2)}{(1 + p\overline{\lambda})^2(1 - \underline{x})},$$

$$\mathscr{C}(\lambda, p, x) \mid_{(x,\lambda)\in\Omega} \leq \overline{\mathscr{C}}(\Omega, p) \triangleq \frac{p(2 - 4\underline{x\lambda} + 4p\overline{\lambda} + 2\overline{x}\overline{\lambda}^2 - 5\underline{xp}\underline{\lambda}^2 + \overline{x}p\overline{\lambda}^3 - 2\underline{xp}^2\underline{\lambda}^3 + 2p^2\overline{\lambda}^2)}{(1 + p\underline{\lambda})^2(1 - \overline{x})},$$

$$\mathscr{D}(\lambda, p, x) \mid_{(x,\lambda)\in\Omega} \geq \underline{\mathscr{D}}(\Omega, p) \triangleq \frac{(1 - \underline{x}\overline{\lambda})p}{1 - \underline{x} + (1 - \underline{x}\overline{\lambda})p},$$

$$\mathscr{D}(\lambda, p, x) \mid_{(x,\lambda)\in\Omega} \leq \overline{\mathscr{D}}(\Omega, p) \triangleq \frac{(1 - \overline{x}\underline{\lambda})p}{1 - \overline{x} + (1 - \overline{x}\underline{\lambda})p},$$

where the last two inequalities hold because $\mathscr{D}(\lambda, p, x)$ is monotonically increasing in x and monotonically decreasing in λ. Clearly,

$$\ln[\underline{\mathscr{D}}(\Omega, p)] < 0, \ \ln[\overline{\mathscr{D}}(\Omega, p)] < 0, \ \underline{\mathscr{C}}(\Omega, p) > 0, \ \overline{\mathscr{C}}(\Omega, p) > 0, \tag{A.17}$$

where (A.17) is true because

$$\begin{aligned} &2 - 4x\lambda + 4p\lambda + 2x\lambda^2 - 5xp\lambda^2 + xp\lambda^3 - 2xp^2\lambda^3 + 2p^2\lambda^2 \\ &= 2p^2\lambda^2(1 - x\lambda) + p\lambda(1 - x\lambda)(4 - x\lambda) + 2(1 - x\lambda)^2 + x\lambda^2(1 - x)(2 + p\lambda) > 0, \end{aligned}$$

for any $(x, \lambda) \in [0, 1)^2$ (and thus $(\overline{x}, \underline{\lambda})$ and $(\underline{x}, \overline{\lambda})$). Hence

$$\begin{aligned} 2\ln 2\frac{\partial f}{\partial p}\;|_{(x,\lambda)\in\Omega} &\geq 1 + \overline{\mathscr{C}}(\Omega, p)\cdot\ln[\underline{\mathscr{D}}(\Omega, p)] \triangleq \dot{f}^-(\Omega, p), \text{ for any } p \in (0, 1] \quad \text{(A.18)} \\ 2\ln 2\frac{\partial f}{\partial p}\;|_{(x,\lambda)\in\Omega} &\leq 1 + \underline{\mathscr{C}}(\Omega, p)\cdot\ln[\overline{\mathscr{D}}(\Omega, p)] \triangleq \dot{f}^+(\Omega, p), \text{ for any } p \in (0, 1]. \end{aligned}$$

Now $\underline{p}^f(\Omega)$, $p^f(x, \lambda)$, and $\overline{p}^f(\Omega)$ are the solutions to $\dot{f}^-(\Omega, p) = 0$, $\frac{\partial f}{\partial p} = 0$, and $\dot{f}^+(\Omega, p) = 0$, respectively, and we claim that

$$\underline{p}^f(\Omega) \;\leq\; p^f(x, \lambda) \;\leq\; \overline{p}^f(\Omega) \quad \text{for any } (x, \lambda) \in \Omega, \qquad \text{(A.19)}$$

since otherwise assume that, e.g., $\underline{p}^f(\Omega) > p^f(x, \lambda)$ for some $(x, \lambda) \in \Omega$, leading to a contradiction

$$0 = \frac{\partial f}{\partial p}\;|_{p=p^f(x,\lambda)} > \frac{\partial f}{\partial p}\;|_{p=\underline{p}^f(\Omega)} \geq \frac{1}{2\ln 2}\dot{f}^-\left(\Omega, \underline{p}^f(\Omega)\right) = 0,$$

where the first inequality is true because $\frac{\partial f}{\partial p}$ is monotonically decreasing in p, and the second inequality is due to (A.18). On the other hand, from (A.10) and the facts that

$$\begin{aligned} 2\ln 2\frac{\partial f}{\partial p} &= 1 + \left[\frac{2p(1 - x\lambda)}{1 - x} - \frac{xp\lambda(1 - \lambda)(2 + p\lambda)}{(1 + p\lambda)^2(1 - x)}\right]\cdot\ln[\mathscr{D}(\Omega, p)] \\ &\geq 1 + \frac{2p(1 - x\lambda)}{1 - x}\cdot\ln[\mathscr{D}(\Omega, p)] = 1 + 2\left[\frac{p(1 - x\lambda)}{1 - x}\right]\cdot\ln\left[\frac{\frac{p(1-x\lambda)}{1-x}}{1 + \frac{p(1-x\lambda)}{1-x}}\right], \end{aligned}$$

$$g \geq 1 - \overline{x} - p + p\underline{\lambda} - p^2\underline{\lambda},\; g \leq 1 - \underline{x} - p + p\overline{\lambda} - p^2\overline{\lambda},$$

we can use similar argument as in the proof of (A.19) to show that for any $(x, \lambda) \in \Omega$,

$$\frac{p^f(x,\lambda)(1-x\lambda)}{1-x} \geq \kappa \quad \Rightarrow \quad p^f(x,\lambda) \; \geq \; \frac{\kappa(1-x)}{1-x\lambda} \geq \frac{\kappa(1-\overline{x})}{1-\overline{x}\underline{\lambda}}. \tag{A.20}$$

and

$$\underline{p}^g(\Omega) \;\leq\; p^g(x,\lambda) \;\leq\; \overline{p}^f(\Omega) \quad \text{for any } (x,\lambda) \in \Omega. \tag{A.21}$$

Thus (A.15) follows from (A.2), (A.19), (A.20), (A.21), and the definitions of $\underline{p}$ and $\overline{p}$.

Finally, to prove (A.16), we use the equivalent definition of f given in (3.68) with w and y defined in (3.65) and (3.66), respectively. Then (A.16) is due to the facts that $w \leq \overline{w}$ and $y \leq \overline{y}$ for any $(x, \lambda) \in \Omega$ and $p = p^*(x, \lambda)$. □

Now we split the rectangle $0 \leq x < 1, 0 \leq \lambda < 1$ into N_c small rectangular cells denoted as

$$\Omega_k \;\;=\;\; \left\{(x,\lambda) : \underline{x}_k \leq x < \overline{x}_k, \underline{\lambda}_k \leq \lambda < \overline{\lambda}_k\right\},$$

for $k = 1, 2, ..., N_c$, and compute the four values $\underline{p}^f(\Omega)$, $\overline{p}^f(\Omega)$, $\underline{p}^g(\Omega)$, $\overline{p}^g(\Omega)$ for each cell. Then we can define

$$\mathbf{\Omega}^{g=0} \;\;=\;\; \bigcup_{k \in \{1,2,...,N_c\} \,:\, \underline{p}^f(\Omega_k) \geq \overline{p}^g(\Omega_k)} \Omega_k.$$

Obviously, $\mathbf{\Omega}^{g=0}$ is an (x, λ) region inside which the maximum $f^*(x, \lambda)$ must be achieved on the boundary $g_2(\lambda, p^*(x, \lambda); x) = 0$, since for any $(x, \lambda) \in \Omega_k$ such that

$\underline{p}^f(\Omega_k) \geq \overline{p}^g(\Omega_k)$, it must be true that

$$\begin{aligned} & p^g(x,\lambda) \;\leq\; \overline{p}^g(\Omega_k) \;\leq\; \underline{p}^f(\Omega_k) \;\leq\; p^f(x,\lambda) \\ \Rightarrow\;\; & p^*(x,\lambda) = \min(p^f(x,\lambda), p^g(x,\lambda)) = p^g(x,\lambda) \\ \Rightarrow\;\; & p^*(x,\lambda) \text{ satisfies } g_2(\lambda, p^*(x,\lambda)\,; x) = 0. \end{aligned}$$

Conversely, for a pair (x,λ), if $f^*(x,\lambda)$ is not achieved on the boundary $g_2(\lambda, p^*(x,\lambda)\,;x) = 0$, then we must have $(x,\lambda) \notin \mathbf{\Omega}^{g=0}$.

Then if $g_2(\lambda^{\max_2}(x), p^{\max_2}(x)\,;x) > 0$ for some $x \in [0,1)$, i.e., if $f^{\max_2}(x)$ is not achieved on the boundary, it must hold that $(x, \lambda^{\max_2}(x)) \notin \mathbf{\Omega}^{g=0}$, which is equivalent to $(x, \lambda^{\max_2}(x)) \in \Omega_k$ for some $\Omega_k \not\subset \mathbf{\Omega}^{g=0}$, due to the definition of $\mathbf{\Omega}^{g=0}$. Hence

$$f^{\max_2}(x) \leq \max_{k\in\{1,2,\ldots,N_c\}:\Omega_k\not\subset\mathbf{\Omega}^{g=0},\underline{x}_k\leq x<\overline{x}_k} \overline{f}(\Omega_k) \overset{\Delta}{=} \overline{f}^{g>0}(x), \tag{A.22}$$

where $\overline{f}^{g>0}(x)$ can be computed numerically up to arbitrary precision for any given $x \in [0,1)$.

Now we proceed to prove Lemma 17. To do this, we compute

$$\max_{x\in[0,1)} \overline{f}^{g>0}(x) = \max_{k\in\{1,2,\ldots,N_c\}:\Omega_k\not\subset\mathbf{\Omega}^{g=0}} \overline{f}(\Omega_k) = 0.1076069180 \overset{\Delta}{=} f^o_{\max},$$

where the maximum of 0.1076069180 is achieved in the rectangular region Ω_k centered at $\frac{\underline{x}_k+\overline{x}_k}{2} = 0.7760825000, \frac{\underline{\lambda}_k+\overline{\lambda}_k}{2} = 0.8730725000$. Then for any $x \in [0,1)$, if the maximum function value $f^{\max_2}(x)$ is achieved at a non-boundary point, i.e., $g_2(\lambda^{\max_2}(x), p^{\max_2}(x)\,;x) > 0$, then due to (A.22), it holds that

$$f^{\max_2}(x) \leq \overline{f}^{g>0}(x) \leq f^o_{\max} = 0.1076069180. \tag{A.23}$$

On the other hand, since the solution to $\mathbb{P}_{2b}(x)$ is always a lower bound on that to $\mathbb{P}_2(x)$, we apply Lemma 16 and obtain

$$f^{\max_2}\left(\frac{N}{L}\right) \geq \tau\left(\frac{N}{L}\right), \tag{A.24}$$

and it is easy to verify that

$$f^{\max_1}(N) \geq f_1(\nu(x), 1, 1, \mu(x), N, 0, L-N) = \tau\left(\frac{N}{L}\right), \tag{A.25}$$

where $f^{\max_1}(N)$ denotes the maximum function value in $\mathbb{P}_1^L$ when $N \in \{0, 1, ..., L-1\}$ is fixed. In addition, due to the fact that $\frac{\partial\tau(x)}{\partial x} > 0$ when $x < x^\star$, and $\frac{\partial\tau(x)}{\partial x} < 0$ when $x > x^\star$, we know that for any integer $L \geq 2$, if there exists an integer N such that $0.777 \leq \frac{N}{L} \leq 0.849$, it must hold that

$$\tau\left(\frac{N}{L}\right) \geq \min(\tau(0.777), \tau\,(0.849)) = 0.1076149432 > f^o_{\max}.$$

Now assume that (3.72) does not hold for some $L \in \{5, 6, 9, 10, ...\}$. Then we have a contradiction

$$f^o_{\max} \geq f^{\max_2}\left(\frac{N^{\max_2}}{L}\right) \geq f^{\max_2}\left(\frac{N}{L}\right) \geq \tau\left(\frac{N}{L}\right) > f^o_{\max},$$

where the first inequality is due to the assumption and (A.23), the second inequality comes from the definition of $N^{\max_2}$, the third inequality is proved in (A.24), while the last inequality is due to the fact that for any $L \in \{5, 6, 9, 10, ...\}$, there always exists an $N \in \{1, 2, ..., L-1\}$ such that $0.777 \leq \frac{N}{L} \leq 0.849$. To verify the latter fact, we write

$$0.777 \leq \frac{4}{5}, \frac{5}{6}, \frac{7}{9}, \frac{8}{10}, \frac{9}{11}, \frac{10}{12}, \frac{11}{13}, \frac{11}{14} \leq 0.849,$$

and note that $0.849 - 0.777 = 0.072 > \frac{1}{L}$ for any integer $L > 14$.

The last case is when $L = 7$, for which no integer $N \in \{0, 1, 2, ..., 6\}$ satisfies $0.777 \leq \frac{N}{L} \leq 0.849$. For this case, we prove (3.72) by computing $\overline{f}^{g>0}\left(\frac{N}{L}\right)$ for all

$N \in \{0, 1, 2, ..., 6\}$, which gives

$$\max_{N\in\{0,1,2,...,6\}} \overline{f}^{g>0}\left(\frac{N}{L}\right) = \max\{0.0983224677, 0.0987674177, 0.0996073967, 0.1007807547, 0.1025463428, 0.1055466980, 0.0443653176\} = 0.1055466980. \tag{A.26}$$

Then (3.72) must be true since otherwise we have a contradiction

$$0.1055466980 \geq \overline{f}^{g>0}\left(\frac{N^{\max_2}}{L}\right) \geq f^{\max_2}\left(\frac{N^{\max_2}}{7}\right) \geq f^{\max_2}\left(\frac{N}{7}\right) \geq \tau\left(\frac{6}{7}\right) = 0.1071970579,$$

where the four inequalities are due to (A.26), (A.22), definition of $N^{\max_2}$, and (A.24), respectively.

Now we have proved that (3.72) holds for any $L \notin \{2, 3, 4, 8\}$. To verify the first equation in (3.83), we need to show that for $L \geq 2$,

$$L \cdot \max_{N\in\{1,2,...,L-1\}} \tau(\frac{N}{L}) = L \cdot \max\left[\tau\left(\frac{\lfloor Lx^\star \rfloor}{L}\right), \tau\left(\frac{\lceil Lx^\star \rceil}{L}\right)\right],$$

which follows from the facts that $\frac{\partial \tau(x)}{\partial x} > 0$ when $x < \frac{\lfloor Lx^\star \rfloor}{L} < x^\star$ and $\frac{\partial \tau(x)}{\partial x} < 0$ when $x > \frac{\lceil Lx^\star \rceil}{L} > x^\star$. All other equations in (3.83) are trivial consequences of (3.73) - (3.78), (3.81) and (3.82). □

VITA

Name Yifu Zhang

Education Doctor of Philosophy (August 2008 – August 2012)
Major: Electrical Engineering
Texas A&M University, College Station, TX 77840

Master of Science (September 2005 – July 2008)
Major: Electronic Engineering
Tsinghua University, Beijing, P.R. China, 100084

Bachelor of Science (September 2001 – July 2005)
Major: Automation
Tsinghua University, Beijing, P. R. China, 100084

Permanent address Department of Electrical & Computer Engineering,
Texas A&M University, College Station, TX 77843

The typist for this dissertation was Yifu Zhang.