Robust Distributed Compression of Symmetrically Correlated Gaussian Sources

ROBUST DISTRIBUTED COMPRESSION OF SYMMETRICALLY CORRELATED GAUSSIAN SOURCES

BY

XUAN ZHANG, B.Sc.

A THESIS

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

Master of Applied Science (2018)	McMaster University
(Electrical & Computer Engineering)	Hamilton, Ontario, Canada

TITLE:	Robust Distributed Compression of Symmetrically Correlated Gaussian Sources
AUTHOR:	Xuan Zhang B.Sc, (Communication Engineering) Northeastern University, Qinhuangdao, China
SUPERVISOR:	Dr. Jun Chen
NUMBER OF PAGES:	xi, 52

To my family and friends

Abstract

Consider a lossy compression system with ℓ distributed encoders and a centralized decoder. Each encoder compresses its observed source and forwards the compressed data to the decoder for joint reconstruction of the target signals under the mean squared error distortion constraint. It is assumed that the observed sources can be expressed as the sum of the target signals and the corruptive noises, which are generated independently from two (possibly different) symmetric multivariate Gaussian distributions. Depending on the parameters of such Gaussian distributions, the rate-distortion limit of this lossy compression system is characterized either completely or for a subset of distortions (including, but not necessarily limited to, those sufficiently close to the minimum distortion achievable when the observed sources are directly available at the decoder). The results are further extended to the robust distributed compression setting, where the outputs of a subset of encoders may also be used to produce a non-trivial reconstruction of the corresponding target signals. In particular, we obtain in the high-resolution regime a precise characterization of the minimum achievable reconstruction distortion based on the outputs of $k + 1$ or more encoders when every k out of all ℓ encoders are operated collectively in the same mode that is greedy in the sense of minimizing the distortion incurred by the reconstruction of the corresponding k target signals with respect to the average rate of these k encoders.

Key Words : Distributed compression, Gaussian source, Karush-Kuhn-Tucker conditions, mean squared error, rate-distortion.

Acknowledgements

First and foremost, I would like to express my deepest gratitude and respect to my supervisor, Dr. Jun Chen, a respectable, responsible and resourceful scholar, who has provided me with encouragement, support and valuable guidance through every stage of my master study.

I would also like to thank Dr. Jiankang Zhang and Dr. Sorina Dumitrescu for giving me valuable advice and assistance in conducting the thesis.

I would also like to thank several professors, Dr. T. Kirubarajan, Dr. J.P. Reilly and Dr. Tim Davidson for giving the assistance in fundamental theories.

I would also like to thank Mr. Li Xie and Ms. Xueping Suo for research cooperation and my colleagues Ms Yameng Chang, Ms Xiaoyu Zhao, Mr Lingshi Kong in my lab for their help.

Lastly, I forever grateful to my parents their love, support and encouragement and to my friends for their company and for bringing joy and hope in my life over the years.

Notation and abbreviations

$\mathbb{E}[\cdot]$: Expectation Operator

$(\cdot)^T$: Transpose Operator

$\mathrm{tr}(\cdot)$: Trace Operator

$\det(\cdot)$: Determinant Operator

1_j : j-Dimensional All-one Row Vector

$\mathrm{diag}^{(j)}(\kappa_1, \cdots, \kappa_j)$: $j \times j$ Diagonal Matrix with Diagonal Entries $\kappa_1, \cdots, \kappa_j$

Y^n : $(Y(1), \cdots, Y(n))$

$\mathcal{A}$: $a_1 < \cdots < a_j$

$(\omega_i)_{i \in \mathcal{A}}$: $(\omega_{a_1}, \cdots, \omega_{a_j})$

$|\mathcal{S}|$: The Cardinality of a Set $\mathcal{S}$

e : the Base of the Logarithm Function

X : Target Signals

Z : Corruptived Noises

S : Observed Sources

V : Auxiliary Random Vector

Q, U, W : l-Dimensional Zero-mean Gaussian Random Vectors

$\Theta^{(j)}$: Arbitrary Unitary Matrix

a_1, a_2, b_1, b_2, c : Nonnegative Parameters

$\hat{a_1}$, $\hat{a_2}$, $\hat{b_1}$, $\hat{b_2}$, $\hat{c}$: Nonnegative Parameters

Γ_X, Γ_Z, Γ_S, : Covariance Matrices

ρ_X, ρ_Z, ρ_S, : Correlation Coefficients

γ_X, γ_Z, γ_S, : Variances

$\mathcal{RD}_k$: A Rate Distortion Tuple $(r, d_k, \cdots, d_l)$

l, k : The Number of Encoders

λ_Q : Unique Positive Number

Contents

List of Figures

Chapter 1

Introduction and Problem Statement

1.1 Background

Since Shannon laid the foundation of the information and coding theory in 1948, in the following decades, many researchers devote themselves to develop the efficiency of the processing under wireless sensor network. The distributed source coding (DSC) was a ground-breaking theory in the information and coding theory. Nowadays, DSC theory is widely used in the fields of wireless sensor networks, video streaming and Multimedia compression. Compressing same source of data, DSC theory delivers higher compression efficiency and lower computational complexity on encoder side than traditional coding theory.

Consider a wireless sensor network where potentially noise-corrupted signals are collected and forwarded to a centralized decoder for further processing. Due to the communication constraints, it is often necessary to reduce the amount of the transmitted data by local pre-processing at each sensor. Though the multiterminal source coding theory, which aims to provide a systematic guideline for the implementation of such pre-processing, is far from being complete, significant progress has been made over the past few decades, starting from the seminal work by Slepian and Wolf on the lossless case [1] to the more recent results on the quadratic Gaussian case [2]-[17].

Arguably the greatest insight offered by this theory is that one can capitalize on the statistical dependency among the data at different sites to improve the compression efficiency even when such data need to be compressed in a purely distributed fashion. However, this performance improvement comes at a price: the compressed data from different sites might not be separably decodable, instead they need to be gathered at a central decoder for joint decompression. As a consequence, losing a portion of distributedly compressed data may render the remaining portion completely useless. Indeed, such situations are often encountered in practice.

For example, in the aforementioned wireless sensor network, it could happen that the fusion center fails to gather the complete set of compressed data needed for performing joint decompression due to unexpected sensor malfunctions or undesirable channel conditions. A natural question thus arises whether a system can harness the benefits of distributed compression without jeopardizing its functionality in adverse scenarios. Intuitively, there exists a tension between compression efficiency and system robustness. A good distributed compression system should strike a balance between these two factors.

The theory intended to characterize the fundamental tradeoff between compression efficiency and system robustness for the centralized setting is known as multiple description coding, which has been extensively studied [18]-[36].

There are several algoithms of data compression which may concern the compression efficiency. Such as dictionary algoithm, fixed bit length packing, run length encoding and Huffman encoding. One can choose the fittable alogithm for the certain senarios to improve the compression efficiency. Consider of the system robustness, it is an ability to cope the errors possibly under a certain error-tolerant rate. There are several methods to determine whether the system is robust or not. For instance, RouthHurwitz stability criterion, Nyquist stability criterion, root locus and etc.

In contrast, its distributed counterpart is far less developed, and the relevant literature is rather scarce [37]-[39].

1.2 Project Objectives

In this thesis, there are two theorems for solving the problems that discussed in the previous section. One is considered about all ℓ distributed encoders and another is considered about k encoders out of ℓ distributed encoone caders to one centralized decoder. The specific objectives are shown as follows:

1. Balance the compression efficiency and system robustness in a lossy compression system. In the present work, we consider a lossy compression system with ℓ distributed encoders and a centralized decoder. Each encoder compresses its observed source and forwards the compressed data to the decoder(see Fig. 1.1). Given the data from an arbitrary subset of encoders, the decoder is required to reconstruct the corresponding target signals within a prescribed mean squared error distortion threshold (dependent on the cardinality of that subset). It is assumed that the observed sources can be expressed as the sum of the target

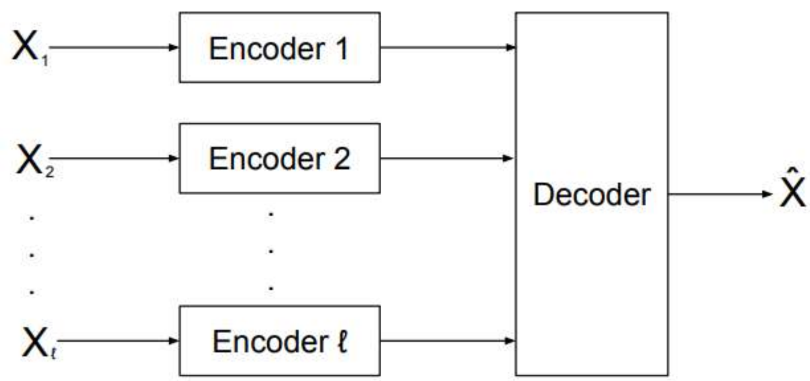

Figure 1.1: A processing from ℓ distributed encoders to a centralized decoder.

signals and the corruptive noises, which are generated independently from two (possibly different) symmetric multivariate Gaussian distributions, which this assumption is not essential for our analysis. It is adopted mainly for the purpose of making the rate-distortion expressions as explicit as possible. This setting is similar to that of the robust Gaussian CEO problem studied in [37], [38]. However, there are two major differences: the robust Gaussian CEO problem imposes the restrictions that 1) the target signal is a scalar process, and 2) the noises across different encoders are independent. Though these restrictions could be justified in certain scenarios, they were introduced largely due to the technical reliance on Oohama's bounding technique for the scalar Gaussian CEO problem [3], [6]. In this paper, we shall tackle the more difficult case where the target signals jointly form a vector process by adapting recently developed analytical methods in Gaussian multiterminal source coding theory [10], [13]-[15] to the robust compression setting.

2. Achieve the target signal to vectors by using the Gaussian multiterminal source.

We show that the theoretical difficulty caused by correlated noises can be circumvented through a coupling argument. Specifically, we introduce a fictitious signal-noise decomposition of the observed sources such that the resulting noises are independent across encoders, and couple it with the given decomposition via a Markov construction. In fact, it will become clear that this coupling argument can be useful even for analyzing those distributed compression systems with independent noises.

1.3 Thesis Structure

This thesis consists of four chapters. This first chapter provides the background and objectives for the robust distributed compression and the structure of this thesis. A problem definition and the main results of the study is presented in chapter 2. The proof of theorem 1 and the description of theorem 2 are also shown in chapter 2. The proof of theorem 2 part 1, part 2 and part 3 are all presented in chapter 3. The conclusion arising from present work and suggestions for future work are provided in chapter 4. Appendix A and Appendix B are both provided after then. A list of reference is provided at the end of the thesis.

Chapter 2

Problem Definitions and Main Results

2.1 Problem Definitions

Let the target signals $X \triangleq (X_1, \cdots, X_\ell)^T$ and the corruptive noises $Z \triangleq (Z_1, \cdots, Z_\ell)^T$ be two mutually independent ℓ-dimensional ($\ell \geq 2$) zero-mean Gaussian random vectors, and the observed sources $S \triangleq (S_1, \cdots, S_\ell)^T$ be their sum (i.e., $S = X + Z$).

Their respective covariance matrices are given by

$$\Gamma_X \triangleq \begin{pmatrix} \gamma_X & \rho_X\gamma_X & \cdots & \rho_X\gamma_X \\ \rho_X\gamma_X & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho_X\gamma_X \\ \rho_X\gamma_X & \cdots & \rho_X\gamma_X & \gamma_X \end{pmatrix},$$

$$\Gamma_Z \triangleq \begin{pmatrix} \gamma_Z & \rho_Z\gamma_Z & \cdots & \rho_Z\gamma_Z \\ \rho_Z\gamma_Z & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho_Z\gamma_Z \\ \rho_Z\gamma_Z & \cdots & \rho_Z\gamma_Z & \gamma_Z \end{pmatrix},$$

$$\Gamma_S \triangleq \begin{pmatrix} \gamma_S & \rho_S\gamma_S & \cdots & \rho_S\gamma_S \\ \rho_S\gamma_S & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho_S\gamma_S \\ \rho_S\gamma_S & \cdots & \rho_S\gamma_S & \gamma_S \end{pmatrix},$$

and satisfy $\Gamma_S = \Gamma_X + \Gamma_Z$. Moreover, we construct an i.i.d. process $\{(X(t), Z(t), S(t))\}_{t=1}^{\infty}$ such that the joint distribution of $X(t) \triangleq (X_1(t), \cdots, X_\ell(t))^T$, $Z(t) \triangleq (Z_1(t), \cdots, Z_\ell(t))^T$, and $S(t) \triangleq (S_1(t), \cdots, S_\ell(t))^T$ is the same as that of X, Z, and S for $t = 1, 2, \cdots$.

By the eigenvalue decomposition, every $j \times j$ (real) matrix

$$\Gamma^{(j)} \triangleq \begin{pmatrix} \alpha & \beta & \cdots & \beta \\ \beta & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \beta \\ \beta & \cdots & \beta & \alpha \end{pmatrix}$$

can be written as

$$\Gamma^{(j)} = \Theta^{(j)}\Lambda^{(j)}(\Theta^{(j)})^T, \tag{2.1}$$

where $\Theta^{(j)}$ is an arbitrary (real) unitary matrix with the first column being $\frac{1}{\sqrt{j}}1_j^T$, and

$$\Lambda^{(j)} \triangleq \text{diag}^{(j)}(\alpha + (j-1)\beta, \alpha - \beta, \cdots, \alpha - \beta).$$

For $j \in \{1, \cdots, \ell\}$, let $\Gamma_X^{(j)}$, $\Gamma_Z^{(j)}$, and $\Gamma_S^{(j)}$ denote the leading $j \times j$ principal submatrices of Γ_X, Γ_Z, and Γ_S, respectively; in view of (2.1), we have

$$\Gamma_X^{(j)} = \Theta^{(j)}\Lambda_X^{(j)}(\Theta^{(j)})^T,$$
$$\Gamma_Z^{(j)} = \Theta^{(j)}\Lambda_Z^{(j)}(\Theta^{(j)})^T,$$
$$\Gamma_S^{(j)} = \Theta^{(j)}\Lambda_S^{(j)}(\Theta^{(j)})^T,$$

where

$$\Lambda_X^{(j)} \triangleq \text{diag}^{(j)}(\lambda_{X,1}^{(j)}, \lambda_{X,2}, \cdots, \lambda_{X,2}),$$
$$\Lambda_Z^{(j)} \triangleq \text{diag}^{(j)}(\lambda_{Z,1}^{(j)}, \lambda_{Z,2}, \cdots, \lambda_{Z,2}),$$
$$\Lambda_S^{(j)} \triangleq \text{diag}^{(j)}(\lambda_{S,1}^{(j)}, \lambda_{S,2}, \cdots, \lambda_{S,2})$$

with

$$\lambda_{X,1}^{(j)} \triangleq (1+(j-1)\rho_X)\gamma_X,$$
$$\lambda_{X,2} \triangleq (1-\rho_X)\gamma_X,$$
$$\lambda_{Z,1}^{(j)} \triangleq (1+(j-1)\rho_Z)\gamma_Z,$$
$$\lambda_{Z,2} \triangleq (1-\rho_Z)\gamma_Z,$$
$$\lambda_{S,1}^{(j)} \triangleq (1+(j-1)\rho_S)\gamma_S,$$
$$\lambda_{S,2} \triangleq (1-\rho_S)\gamma_S.$$

Note that Γ_X, Γ_Z, and Γ_S are positive semidefinite (and consequently are well-defined covariance matrices) if and only if $\lambda_{X,1}^{(\ell)} \geq 0$, $\lambda_{X,2} \geq 0$, $\lambda_{Z,1}^{(\ell)} \geq 0$, $\lambda_{Z,2} \geq 0$, $\lambda_{S,1}^{(\ell)} \geq 0$, and $\lambda_{S,2} \geq 0$. Furthermore, we assume that $\gamma_X > 0$ since otherwise the target signals are not random. It follows by this assumption that $\gamma_S > 0$, $\lambda_{X,1}^{(\ell)}+\lambda_{X,2} > 0$, and $\lambda_{S,1}^{(\ell)} + \lambda_{S,2} > 0$.

2.1.1 The Case of All ℓ Encoders

Definition 1: Given $k \in \{1, \cdots, \ell\}$, a rate-distortion tuple $(r, d_k, \cdots, d_\ell)$ is said to be achievable if, for any $\epsilon > 0$, there exist encoding functions $\phi_i^{(n)} : \mathbb{R}^n \to \mathcal{C}_i^{(n)}$,

$i = 1, \cdots, \ell$, such that

$$\frac{1}{kn}\sum_{i\in\mathcal{A}}\log|\mathcal{C}_i^{(n)}| \leq r+\epsilon,$$
$$\mathcal{A}\subseteq\{1,\cdots,\ell\} \text{ with } |\mathcal{A}|=k, \tag{2.2}$$
$$\frac{1}{|\mathcal{A}|n}\sum_{i\in\mathcal{A}}\sum_{t=1}^{n}\mathbb{E}[(X_i(t)-\hat{X}_{i,\mathcal{A}}(t))^2] \leq d_{|\mathcal{A}|}+\epsilon,$$
$$\mathcal{A}\subseteq\{1,\cdots,\ell\} \text{ with } |\mathcal{A}|\geq k, \tag{2.3}$$

where $\hat{X}_{i,\mathcal{A}}(t) \triangleq \mathbb{E}[X_i(t)|(\phi_{i'}^{(n)}(S_{i'}^n))_{i'\in\mathcal{A}}]$. The set of all such achievable $(r, d_k, \cdots, d_\ell)$ is denoted by $\mathcal{RD}_k$.

Remark 1: Due to the symmetry of the underlying distributions, it can be shown via a timesharing argument that $\mathcal{RD}_k$ is not affected if we replace (2.2) with either of the following constraints

$$\frac{1}{n}\log|\mathcal{C}_i^{(n)}| \leq r+\epsilon, \quad i=1,\cdots,\ell,$$
$$\frac{1}{\ell n}\sum_{i=1}^{\ell}\log|\mathcal{C}_i^{(n)}| \leq r+\epsilon,$$

and/or replace (2.3) with either of the following constraints

$$\frac{1}{n}\sum_{t=1}^{n}\mathbb{E}[(X_i(t)-\hat{X}_{i,\mathcal{A}}(t))^2] \leq d_{|\mathcal{A}|}+\epsilon,$$
$$\mathcal{A}\subseteq\{1,\cdots,\ell\} \text{ with } |\mathcal{A}|\geq k,$$
$$\frac{1}{\binom{n}{j}jn}\sum_{\mathcal{A}\subseteq\{1,\cdots,\ell\}:|\mathcal{A}|=j}\sum_{i\in\mathcal{A}}\sum_{t=1}^{n}\mathbb{E}[(X_i(t)-\hat{X}_{i,\mathcal{A}}(t))^2]$$
$$\leq d_j+\epsilon, \quad j=k,\cdots,\ell.$$

Remark 2: We show in Appendix A that, for $j = k, \cdots, \ell$,

$$\begin{aligned} d_{\min}^{(j)} &\triangleq \frac{1}{j} \sum_{i=1}^{j} \mathbb{E}[(X_i - \mathbb{E}[X_i | S_1, \cdots, S_j])^2] \\ &= \frac{1}{j} d_{\min,1}^{(j)} + \frac{j-1}{j} d_{\min,2}, \end{aligned}$$

where

$$d_{\min,1}^{(j)} \triangleq \begin{cases} 0, & \lambda_{S,1}^{(j)} = 0, \\ \frac{\lambda_{X,1}^{(j)} \lambda_{Z,1}^{(j)}}{\lambda_{S,1}^{(j)}}, & \text{otherwise}, \end{cases}$$

$$d_{\min,2} \triangleq \begin{cases} 0, & \lambda_{S,2} = 0, \\ \frac{\lambda_{X,2} \lambda_{Z,2}}{\lambda_{S,2}}, & \text{otherwise}. \end{cases}$$

It is clear that $d_j > d_{\min}^{(j)}$, $j = k, \cdots, \ell$, for any $(r, d_k, \cdots, d_\ell) \in \mathcal{RD}_k$. Moreover, if $d_j \geq \gamma_X$ for some $j \in \{k, \cdots, \ell\}$, then the corresponding distortion constraint is redundant. Henceforth we shall focus on the case $d_j \in (d_{\min}^{(j)}, \gamma_X)$, $j = k, \cdots, \ell$.

2.1.2 The Case of k Out of All ℓ Encoders

Definition 2: For $d_\ell \in (d_{\min}^{(\ell)}, \gamma_X)$, let

$$r^{(\ell)}(d_\ell) \triangleq \min\{r : (r, d_\ell) \in \mathcal{RD}_\ell\}.$$

In order to state our main results, we introduce the following quantities. For any

$k \in \{1, \cdots, \ell\}$ and $d_k \in (d_{\min}^{(k)}, \gamma_X)$, let

$$\begin{aligned}
\overline{r}^{(k)}(d_k) &\triangleq \frac{1}{2k} \log \frac{(\lambda_{S,1}^{(k)} + \lambda_Q^{(k)})(\lambda_{S,2} + \lambda_Q^{(k)})^{k-1}}{(\lambda_Q^{(k)})^k}, \\
d_j^{(k)}(d_k) &\triangleq \frac{\lambda_{X,1}^{(j)}(\lambda_{Z,1}^{(j)} + \lambda_Q^{(k)})}{j(\lambda_{S,1}^{(j)} + \lambda_Q^{(k)})} \\
&\quad + \frac{(j-1)\lambda_{X,2}(\lambda_{Z,2} + \lambda_Q^{(k)})}{j(\lambda_{S,2} + \lambda_Q^{(k)})}, \quad j = k, \cdots, \ell,
\end{aligned}$$

where $\lambda_Q^{(k)}$ is the unique positive number satisfying

$$\frac{\lambda_{X,1}^{(k)}(\lambda_{Z,1}^{(k)} + \lambda_Q^{(k)})}{k(\lambda_{S,1}^{(k)} + \lambda_Q^{(k)})} + \frac{(k-1)\lambda_{X,2}(\lambda_{Z,2} + \lambda_Q^{(k)})}{k(\lambda_{S,2} + \lambda_Q^{(k)})} = d_k. \tag{2.4}$$

Our first result is a partial characterization of $r^{(\ell)}(d_\ell)$.

2.2 The Comprehension of Two Theorems

2.2.1 The definition of Theorem 1

Theorem 1: For $d_\ell \in (d_{\min}^{(\ell)}, \gamma_X)$,

$$r^{(\ell)}(d_\ell) = \overline{r}^{(\ell)}(d_\ell)$$

if either of the following conditions is satisfied:

1. $\rho_S \geq 0$ and

$$(\ell-1)\lambda_{X,2}^2(\lambda_{S,1}^{(\ell)})^2\mu^{(\ell)}(\mu^{(\ell)}-1) + \ell(\lambda_{X,1}^{(\ell)})^2\lambda_{S,2}^2 \geq 0, \tag{2.5}$$

where

$$\mu^{(\ell)} \triangleq \frac{\lambda_{S,2}-\lambda_{S,2}(\lambda_{S,2}+\lambda_Q^{(\ell)})^{-1}\lambda_{S,2}}{\lambda_{S,1}^{(\ell)}-\lambda_{S,1}^{(\ell)}(\lambda_{S,1}^{(\ell)}+\lambda_Q^{(\ell)})^{-1}\lambda_{S,1}^{(\ell)}}. \tag{2.6}$$

2. $\rho_S \leq 0$ and

$$(\lambda_{X,1}^{(\ell)})^2\lambda_{S,2}^2\nu^{(\ell)}(\nu^{(\ell)}-1) + \ell\lambda_{X,2}^2(\lambda_{S,1}^{(\ell)})^2 \geq 0, \tag{2.7}$$

where

$$\nu^{(\ell)} \triangleq \frac{\lambda_{S,1}^{(\ell)}-\lambda_{S,1}^{(\ell)}(\lambda_{S,1}^{(\ell)}+\lambda_Q^{(\ell)})^{-1}\lambda_{S,1}^{(\ell)}}{\lambda_{S,2}-\lambda_{S,2}(\lambda_{S,2}+\lambda_Q^{(\ell)})^{-1}\lambda_{S,2}}. \tag{2.8}$$

Remark 3:

1. Consider the case $\rho_S \geq 0$. When $(\ell-1)\lambda_{X,2}^2(\lambda_{S,1}^{(\ell)})^2 \leq 4\ell(\lambda_{X,1}^{(\ell)})^2\lambda_{S,2}^2$, the inequality (2.5) always holds, and $r^{(\ell)}(d_\ell)$ is characterized for all $d_\ell \in (d_{\min}^{(\ell)}, \gamma_X)$. When $(\ell-1)\lambda_{X,2}^2(\lambda_{S,1}^{(\ell)})^2 > 4\ell(\lambda_{X,1}^{(\ell)})^2\lambda_{S,2}^2$, the equation $(\ell-1)\lambda_{X,2}^2(\lambda_{S,1}^{(\ell)})^2\mu^{(\ell)}(\mu^{(\ell)}-1)+$

$\ell(\lambda_{X,1}^{(\ell)})^2\lambda_{S,2}^2 = 0$ has two real roots in the interval $[0,1]$:

$$\mu_1^{(\ell)} \triangleq \frac{1}{2} - \frac{1}{2}\sqrt{1 - \frac{4\ell(\lambda_{X,1}^{(\ell)})^2\lambda_{S,2}^2}{(\ell-1)\lambda_{X,2}^2(\lambda_{S,1}^{(\ell)})^2}},$$
$$\mu_2^{(\ell)} \triangleq \frac{1}{2} + \frac{1}{2}\sqrt{1 - \frac{4\ell(\lambda_{X,1}^{(\ell)})^2\lambda_{S,2}^2}{(\ell-1)\lambda_{X,2}^2(\lambda_{S,1}^{(\ell)})^2}}.$$

Therefore, the inequality (2.5) holds if

$$\mu^{(\ell)} \leq \mu_1^{(\ell)} \text{ or } \mu^{(\ell)} \geq \mu_2^{(\ell)}. \tag{2.9}$$

It is easy to verify that (2.9) is satisfied when $\lambda_{S,1}^{(\ell)} > \lambda_{S,2} = 0$ (which implies $\mu^{(\ell)} = 0$) or $\lambda_{S,1}^{(\ell)} = \lambda_{S,2} > 0$ (which implies $\mu^{(\ell)} = 1$). When $\lambda_{S,1}^{(\ell)} > \lambda_{S,2} > 0$, $\mu^{(\ell)}$ is a strictly decreasing function of d_ℓ, converging to 1 as $d_\ell \to d_{\min}^{(\ell)}$ and to $\frac{\lambda_{S,2}}{\lambda_{S,1}^{(\ell)}}$ as $d_\ell \to \gamma_X$; hence, it suffices to analyze the following four scenarios.

(a) $\mu_2^{(\ell)} \leq \frac{\lambda_{S,2}}{\lambda_{S,1}^{(\ell)}}$: $\mu^{(\ell)} \geq \mu_2^{(\ell)}$ is satisfied for all $d_\ell \in (d_{\min}^{(\ell)}, \gamma_X)$.

(b) $\mu_1^{(\ell)} \leq \frac{\lambda_{S,2}}{\lambda_{S,1}^{(\ell)}}$ and $\frac{\lambda_{S,2}}{\lambda_{S,1}^{(\ell)}} < \mu_2^{(\ell)} < 1$: $\mu^{(\ell)} \geq \mu_2^{(\ell)}$ is satisfied for all d_ℓ sufficiently close to $d_{\min}^{(\ell)}$.

(c) $\mu_1^{(\ell)} > \frac{\lambda_{S,2}}{\lambda_{S,1}^{(\ell)}}$ and $\mu_2^{(\ell)} < 1$: $\mu^{(\ell)} \leq \mu_1^{(\ell)}$ is satisfied for all d_ℓ sufficiently close to γ_X while $\mu^{(\ell)} \geq \mu_2^{(\ell)}$ is satisfied for all d_ℓ sufficiently close to $d_{\min}^{(\ell)}$.

(d) $\mu_1^{(\ell)} = 0$ and $\mu_2^{(\ell)} = 1$: This can happen only when $\lambda_{X,1}^{(\ell)} = 0$.

In view of the above discussion, under the condition $\rho_S \geq 0$, $r^{(\ell)}(d_\ell)$ is characterized at least for all d_ℓ sufficiently close to $d_{\min}^{(\ell)}$ unless $\lambda_{X,1}^{(\ell)} = 0$ and $\lambda_{S,1}^{(\ell)} > \lambda_{S,2}$ (note that $\lambda_{X,1}^{(\ell)} = 0$ implies $\lambda_{S,2} > 0$).

2. Consider the case $\rho_S \leq 0$. When $(\lambda_{X,1}^{(\ell)})^2\lambda_{S,2}^2 \leq 4\ell\lambda_{X,2}^2(\lambda_{S,1}^{(\ell)})^2$, the inequality (2.7) always holds, and $r^{(\ell)}(d_\ell)$ is characterized for all $d_\ell \in (d_{\min}^{(\ell)}, \gamma_X)$. When $(\lambda_{X,1}^{(\ell)})^2\lambda_{S,2}^2 > 4\ell\lambda_{X,2}^2(\lambda_{S,1}^{(\ell)})^2$, the equation $(\lambda_{X,1}^{(\ell)})^2\lambda_{S,2}^2\nu^{(\ell)}(\nu^{(\ell)}-1)+\ell\lambda_{X,2}^2(\lambda_{S,1}^{(\ell)})^2 = 0$ has two real roots in the interval $[0,1]$:

$$\nu_1^{(\ell)} \triangleq \frac{1}{2} - \frac{1}{2}\sqrt{1-\frac{4\ell\lambda_{X,2}^2(\lambda_{S,1}^{(\ell)})^2}{(\lambda_{X,1}^{(\ell)})^2\lambda_{S,2}^2}},$$

$$\nu_2^{(\ell)} \triangleq \frac{1}{2} + \frac{1}{2}\sqrt{1-\frac{4\ell\lambda_{X,2}^2(\lambda_{S,1}^{(\ell)})^2}{(\lambda_{X,1}^{(\ell)})^2\lambda_{S,2}^2}}.$$

Therefore, the inequality (2.7) holds if

$$\nu^{(\ell)} \leq \nu_1^{(\ell)} \text{ or } \nu^{(\ell)} \geq \nu_2^{(\ell)}. \tag{2.10}$$

It is easy to verify that (2.10) is satisfied when $\lambda_{S,2} > \lambda_{S,1}^{(\ell)} = 0$ (which implies $\nu^{(\ell)} = 0$) or $\lambda_{S,1}^{(\ell)} = \lambda_{S,2} > 0$ (which implies $\nu^{(\ell)} = 1$). When $\lambda_{S,2} > \lambda_{S,1}^{(\ell)} > 0$, $\nu^{(\ell)}$ is a strictly decreasing function of d_ℓ, converging to 1 as $d_\ell \to d_{\min}^{(\ell)}$ and to $\frac{\lambda_{S,1}^{(\ell)}}{\lambda_{S,2}}$ as $d_\ell \to \gamma_X$; hence, it suffices to analyze the following four scenarios.

(a) $\nu_2^{(\ell)} \leq \frac{\lambda_{S,1}^{(\ell)}}{\lambda_{S,2}}$: $\nu^{(\ell)} \geq \nu_2^{(\ell)}$ is satisfied for all $d_\ell \in (d_{\min}^{(\ell)}, \gamma_X)$.

(b) $\nu_1^{(\ell)} \leq \frac{\lambda_{S,1}^{(\ell)}}{\lambda_{S,2}}$ and $\frac{\lambda_{S,1}^{(\ell)}}{\lambda_{S,2}} < \nu_2^{(\ell)} < 1$: $\nu^{(\ell)} \geq \nu_2^{(\ell)}$ is satisfied for all d_ℓ sufficiently close to $d_{\min}^{(\ell)}$.

(c) $\nu_1^{(\ell)} > \frac{\lambda_{S,1}^{(\ell)}}{\lambda_{S,2}}$ and $\nu_2^{(\ell)} < 1$: $\nu^{(\ell)} \leq \nu_1^{(\ell)}$ is satisfied for all d_ℓ sufficiently close to γ_X while $\nu^{(\ell)} \geq \nu_2^{(\ell)}$ is satisfied for all d_ℓ sufficiently close to $d_{\min}^{(\ell)}$.

(d) $\nu_1^{(\ell)} = 0$ and $\nu_2^{(\ell)} = 1$: This can happen only when $\lambda_{X,2} = 0$.

In view of the above discussion, under the condition $\rho_S \leq 0$, $r^{(\ell)}(d_\ell)$ is characterized at least for all d_ℓ sufficiently close to $d_{\min}^{(\ell)}$ unless $\lambda_{X,2} = 0$ and $\lambda_{S,2} > \lambda_{S,1}^{(\ell)}$ (note that $\lambda_{X,2} = 0$ implies $\lambda_{S,1}^{(\ell)} > 0$).

Theorem 1 a special case of the following more general result.

2.2.2 The definition of Theorem 2

Theorem 2:

1. For $d_k \in (d_{\min}^{(k)}, \gamma_X)$,

$$(\overline{r}^{(k)}(d_k), d_k^{(k)}(d_k), \cdots, d_\ell^{(k)}(d_k)) \in \mathcal{RD}_k.$$

2. For $(r, d_k, \cdots, d_\ell) \in \mathcal{RD}_k$ with $d_k \in (d_{\min}^{(k)}, \gamma_X)$,

$$r \geq \overline{r}^{(k)}(d_k)$$

if either of the following conditions is satisfied:

i) $\rho_S \geq 0$ and

$$\begin{aligned}&(k-1)\lambda_{X,2}^2(\lambda_{S,1}^{(k)})^2\mu^{(k)}(\mu^{(k)}-1)\\&+k(\lambda_{X,1}^{(k)})^2\lambda_{S,2}^2 \geq 0,\end{aligned} \tag{2.11}$$

where $\mu^{(k)}$ is defined in (2.6) with ℓ replaced by k.

ii) $\rho_S \leq 0$ and

$$
\begin{aligned}
&(\lambda_{X,1}^{(k)})^2 \lambda_{S,2}^2 \nu^{(k)} (\nu^{(k)} - 1) + k \lambda_{X,2}^2 (\lambda_{S,1}^{(k)})^2 \\
&\geq 0, \qquad (2.12)
\end{aligned}
$$

where $\nu^{(k)}$ is defined in (2.8) with ℓ replaced by k.

(a) For $j \in \{k, \cdots, \ell\}$ and $(r, d_k, \cdots, d_\ell) \in \mathcal{RD}_k$ with $d_k \in (d_{\min}^{(k)}, \gamma_X)$ and $r = \overline{r}^{(k)}(d_k)$, we have

$$
d_j \geq d_j^{(k)}(d_k)
$$

if either of the following conditions is satisfied:

i. Condition i).

ii. $\rho_S \leq 0$, $\lambda_{S,1}^{(j)} > 0$, and

$$
\begin{aligned}
&(\nu^{(k,j)} + (k-1))(\lambda_{X,1}^{(k)})^2 \lambda_{S,2}^2 (\nu^{(k)})^2 \\
&+ (k-1)(\nu^{(k,j)} - \nu^{(k)}) \lambda_{X,2}^2 (\lambda_{S,1}^{(k)})^2 \geq 0, \qquad (2.13) \\
&(\nu^{(k,j)} - 1)(\lambda_{X,1}^{(k)})^2 \lambda_{S,2}^2 (\nu^{(k)})^2 \\
&+ ((k-1)\nu^{(k,j)} + \nu^{(k)}) \lambda_{X,2}^2 (\lambda_{S,1}^{(k)})^2 \geq 0, \qquad (2.14)
\end{aligned}
$$

where

$$
\nu^{(k,j)} = \frac{\lambda_{S,1}^{(j)} - \lambda_{S,1}^{(j)} (\lambda_{S,1}^{(j)} + \lambda_Q^{(k)})^{-1} \lambda_{S,1}^{(j)}}{\lambda_{S,2} - \lambda_{S,2} (\lambda_{S,2} + \lambda_Q^{(k)})^{-1} \lambda_{S,2}}.
$$

Proof: See Chapter 3.

Remark 4:

1. The argument in Remark 3 can be leveraged to prove that, for the case $\rho_S \geq 0$, the inequality (2.11) holds at least for all d_k sufficiently close to $d_{\min}^{(k)}$ unless $\lambda_{X,1}^{(k)} = 0$ (which can happen only when $k = \ell$) and $\lambda_{S,1}^{(k)} > \lambda_{S,2}$ (note that $\lambda_{X,1}^{(k)} = 0$ implies $\lambda_{S,2} > 0$); similarly, for the case $\rho_S \leq 0$, the inequality (2.12) holds at least for all d_k sufficiently close to $d_{\min}^{(k)}$ unless $\lambda_{X,2} = 0$ and $\lambda_{S,2} > \lambda_{S,1}^{(k)}$ (note that $\lambda_{X,2} = 0$ implies $\lambda_{S,1}^{(k)} > 0$).

2. For the case $\rho_S \leq 0$, the condition $\lambda_{S,1}^{(j)} > 0$ can be potentially violated (i.e., $\lambda_{S,1}^{(j)} = 0$) only when $j = \ell$.

3. Consider the case $\rho_S \leq 0$ and $\lambda_{S,1}^{(j)} > 0$. If $\lambda_{X,1}^{(k)} > 0$, then the inequality (2.13) holds at least for d_k sufficiently close to $d_{\min}^{(k)}$; if $\lambda_{X,1}^{(k)} = 0$, which implies $k = j = \ell$, then the inequality (2.13) always holds. The inequality (2.14) holds at least for d_k sufficiently close to $d_{\min}^{(k)}$ unless $\lambda_{X,2} = 0$ and $\lambda_{S,2} > \lambda_{S,1}^{(j)}$.

2.3 Main Results

Our main results are summarized below.

1. For the case where the decoder is only required to reconstruct the target signals based on the outputs of all ℓ encoders, the rate-distortion limit is characterized either completely or partially, depending on the parameters of signal and noise distributions.

2. For the case where the outputs of a subset of encoders may also be used to produce a non-trivial reconstruction of the corresponding target signals, the

minimum achievable reconstruction distortion based on the outputs of $k+1$ or more encoders is characterized either completely or partially, depending on the parameters of signal and noise distributions, when every k out of all ℓ encoders are operated collectively in the same mode that is greedy in the sense of minimizing the distortion incurred by the reconstruction of the corresponding k target signals with respect to the average rate of these k encoders.

Chapter 3

Proof of Theorem 2

3.1 Proof of Theorem 2 : Part 1

The following lemma can be obtained by adapting the classical result by Berger [40] and Tung [41] to the current setting.

Lemma 1: For any auxiliary random vector $V \triangleq (V_1, \cdots, V_\ell)^T$ jointly distributed with (X, Z, S) such that $\{X, Z, (S_{i'})_{i' \in \{1,\cdots,\ell\}\setminus\{i\}}, (V_{i'})_{i' \in \{1,\cdots,\ell\}\setminus\{i\}}\} \leftrightarrow S_i \leftrightarrow V_i$ form a Markov chain, $i = 1, \cdots, \ell$, and any $(r, d_k \cdots, d_\ell)$ such that

$$r1_k \in \mathcal{R}(\mathcal{A}), \quad \mathcal{A} \subseteq \{1, \cdots, \ell\} \text{ with } |\mathcal{A}| = k,$$
$$d_{|\mathcal{A}|} \geq \frac{1}{|\mathcal{A}|} \sum_{i \in \mathcal{A}} \mathbb{E}[(X_i - \mathbb{E}[X_i|(V_{i'})_{i' \in \mathcal{A}}])^2],$$
$$\mathcal{A} \subseteq \{1, \cdots, \ell\} \text{ with } |\mathcal{A}| \geq k,$$

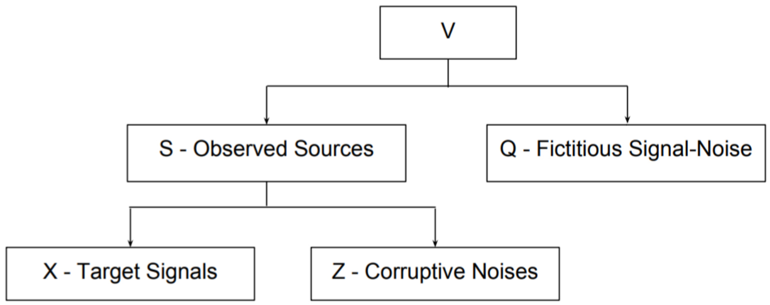

Figure 3.1: The decomposition of vector V

where $\mathcal{R}(\mathcal{A})$ denotes the set of $(r_i)_{i\in\mathcal{A}}$ satisfying

$$\sum_{i\in\mathcal{B}} r_i \geq I((S_i)_{i\in\mathcal{B}};(V_i)_{i\in\mathcal{B}}|(V_i)_{i\in\mathcal{A}\setminus\mathcal{B}}), \quad \emptyset \subset \mathcal{B} \subseteq \mathcal{A},$$

we have

$$(r, d_k \cdots, d_\ell) \in \mathcal{RD}_k.$$

Equipped with this lemma, we are in a position to prove Part 1 of Theorem 2. Let $Q \triangleq (Q_1, \cdots, Q_\ell)^T$ be an ℓ-dimensional zero-mean Gaussian random vector with covariance matrix

$$\Lambda_Q \triangleq \text{diag}^{(\ell)}(\lambda_Q, \cdots, \lambda_Q) \succ 0.$$

Moreover, we assume that Q is independent of (X, Z, S), and let

$$V_i \triangleq S_i + Q_i, \quad i = 1, \cdots, \ell.$$

The relation between all the vectors are shown in Fig. 3.1.

Clearly, $V \triangleq (V_1, \cdots, V_\ell)^T$ satisfies the condition specified in Lemma 1. Let

$$r \triangleq \frac{1}{k} I(S_1, \cdots, S_k; V_1, \cdots, V_k),$$
$$d_j \triangleq \frac{1}{j} \sum_{i=1}^{j} \mathbb{E}[(X_i - \mathbb{E}[X_i|V_1, \cdots, V_j])^2],$$
$$j = k, \cdots, \ell.$$

It is easy to show that $r1_k \in \mathcal{R}(\mathcal{A})$ for all $\mathcal{A} \subseteq \{1, \cdots, \ell\}$ with $|\mathcal{A}| = k$ by leveraging the contra-polymatroid structure [42] of $\mathcal{R}(\mathcal{A})$ and the symmetry of the underlying distributions. Let $\Lambda_Q^{(j)}$ denote the leading $j \times j$ principal submatrix of Λ_Q, $j = k, \cdots, \ell$. We have

$$\begin{aligned}
r &= \frac{1}{k}(h(V_1, \cdots, V_k) - h(V_1, \cdots, V_k|S_1, \cdots, S_k)) \\
&= \frac{1}{k}(h(S_1 + Q_1, \cdots, S_k + Q_k) - h(Q_1, \cdots, Q_k)) \\
&= \frac{1}{2k} \log \frac{\det(\Gamma_S^{(k)} + \Lambda_Q^{(k)})}{\det(\Lambda_Q^{(k)})} \\
&= \frac{1}{2k} \log \frac{\det(\Lambda_S^{(k)} + \Lambda_Q^{(k)})}{\det(\Lambda_Q^{(k)})} \\
&= \frac{1}{2k} \log \frac{(\lambda_{S,1}^{(k)} + \lambda_Q)(\lambda_{S,2} + \lambda_Q)^{k-1}}{\lambda_Q^k}.
\end{aligned}$$

Moreover, for $j = k, \cdots, \ell$,

$$\begin{aligned} d_j &= \frac{1}{j}\mathrm{tr}(\Gamma_X^{(j)} - \Gamma_X^{(j)}(\Gamma_S^{(j)} + \Lambda_Q^{(j)})^{-1}\Gamma_X^{(j)}) \\ &= \frac{1}{j}\mathrm{tr}(\Lambda_X^{(j)} - \Lambda_X^{(j)}(\Lambda_S^{(j)} + \Lambda_Q^{(j)})^{-1}\Lambda_X^{(j)}) \\ &= \frac{\lambda_{X,1}^{(j)}(\lambda_{Z,1}^{(j)} + \lambda_Q)}{j(\lambda_{S,1}^{(j)} + \lambda_Q)} + \frac{(j-1)\lambda_{X,2}(\lambda_{Z,2} + \lambda_Q)}{j(\lambda_{S,2} + \lambda_Q)}, \end{aligned}$$

which is a strictly increasing function of λ_Q, converging to $d_{\min}^{(j)}$ as $\lambda_Q \to 0$ and to γ_X as $\lambda_Q \to \infty$. One can readily complete the proof of Part 1 of Theorem 2 by invoking Lemma 1.

3.2 Proof of Theorem 2 : Part 2 and Part 3

Now we proceed to prove Part 2 and Part 3 of Theorem 2. Fix k and j with $1 \le k \le j \le \ell$. First consider the case $\Gamma_S^{(j)} \succ 0$ (i.e., $\lambda_{S,1}^{(j)} > 0$ and $\lambda_{S,2} > 0$). Let $(S_1, \cdots, S_j)^T = (U_1, \cdots, U_j)^T + (W_1, \cdots, W_j)^T$ be a fictitious signal-noise decomposition of $(S_1, \cdots, S_j)^T$, where $(U_1, \cdots, U_j)^T$ and $(W_1, \cdots, W_j)^T$ are two mutually independent j-dimensional zero-mean Gaussian vectors with covariance matrices

$$\begin{aligned} &\Gamma_U^{(j)} \succ 0, \\ &\Lambda_W^{(j)} \triangleq \mathrm{diag}^{(j)}(\lambda_W, \cdots, \lambda_W) \succ 0, \end{aligned}$$

respectively. Moreover, we couple this fictitious decomposition with the given decomposition $(S_1, \cdots, S_j)^T = (X_1, \cdots, X_j)^T + (Z_1, \cdots, Z_j)^T$ by assuming that $(\{(U_1, \cdots, U_j)^T, (W_1, \cdots, W_j)^T\} \leftrightarrow (S_1, \cdots, S_j)^T \leftrightarrow \{(X_1, \cdots, X_j)^T, (Z_1, \cdots, Z_j)^T\}$ form a Markov chain, and construct the auxiliary random processes $\{(U_1(t), \cdots, U_j(t))^T\}_{t=1}^{\infty}$ and $\{(W_1(t), \cdots, W_\ell(t))^T\}_{t=1}^{\infty}$ accordingly.

It is worth mentioning that the idea of augmenting the probability space via the introduction of auxiliary random processes is inspired by [8], [10], [13]-[15], [18], [24], [26], [28]. Our construction (without the symmetry constraint) can be viewed as a generalization of that in [10], which is restricted to the special case where the corruptive noises are absent. It should also be contrasted with the conventional approach where $(U_1, \cdots, U_j)^T$ and $(W_1, \cdots, W_j)^T$ are set respectively to be $(X_1, \cdots, X_j)^T$ and $(Z_1, \cdots, Z_j)^T$ (with the components of $(Z_1, \cdots, Z_j)^T$ assumed to be mutually independent); in general, the Markov coupling allows more flexible constructions and yields stronger results. Another novelty in our construction is that the fictitious decomposition is specified for $(S_1, \cdots, S_j)^T$ instead of $(S_1, \cdots, S_\ell)^T$. As a consequence, we can choose λ_W from $(0, \min\{\lambda_{S,1}^{(j)}, \lambda_{S,2}\})$, which may offer more freedom than $(0, \min\{\lambda_{S,1}^{(\ell)}, \lambda_{S,2}\})$ since $\min\{\lambda_{S,1}^{(j)}, \lambda_{S,2}\} \geq \min\{\lambda_{S,1}^{(\ell)}, \lambda_{S,2}\}$ and the inequality is strict when $\rho_S < 0$ and $j < \ell$.

In view of Definition 1, for any $(r, d_k \cdots, d_\ell) \in \mathcal{RD}_k$ and $\epsilon > 0$, there exist encoding functions $\phi_i^{(n)} : \mathbb{R}^n \to \mathcal{C}_i^{(n)}$, $i = 1, \cdots, j$, such that

$$\frac{1}{kn} \sum_{i \in \mathcal{A}} \log |\mathcal{C}_i^{(n)}| \le r + \epsilon, \qquad \mathcal{A} \subseteq \{1, \cdots, j\} \text{ with } |\mathcal{A}| = k, \tag{3.1}$$

$$\frac{1}{kn} \sum_{i \in \mathcal{A}} \sum_{t=1}^{n} \mathbb{E}[(X_i(t) - \hat{X}_{i,\mathcal{A}}(t))^2] \le d_k + \epsilon, \qquad \mathcal{A} \subseteq \{1, \cdots, j\} \text{ with } |\mathcal{A}| = k, \tag{3.2}$$

$$\frac{1}{jn} \sum_{i=1}^{j} \sum_{t=1}^{n} \mathbb{E}[(X_i(t) - \hat{X}_{i,\{1,\cdots,j\}}(t))^2] \le d_j + \epsilon.$$

We have

$$\begin{aligned}
&\sum_{i\in\mathcal{A}} \log |\mathcal{C}_i^{(n)}| \\
&\geq H((\phi_i^{(n)}(S_i^n))_{i\in\mathcal{A}}) \\
&= I((U_i^n)_{i\in\mathcal{A}}; (\phi_i^{(n)}(S_i^n))_{i\in\mathcal{A}}) \\
&\quad + H((\phi_i^{(n)}(S_i^n))_{i\in\mathcal{A}}|(U_i^n)_{i\in\mathcal{A}}) \\
&= I((U_i^n)_{i\in\mathcal{A}}; (\phi_i^{(n)}(S_i^n))_{i\in\mathcal{A}}) \\
&\quad + I((S_i^n)_{i\in\mathcal{A}}; (\phi_i^{(n)}(S_i^n))_{i\in\mathcal{A}}|(U_i^n)_{i\in\mathcal{A}}) \\
&= h((U_i^n)_{i\in\mathcal{A}}) + h((W_i^n)_{i\in\mathcal{A}}) \\
&\quad - h((U_i^n)_{i\in\mathcal{A}}|(\phi_i^{(n)}(S_i^n))_{i\in\mathcal{A}}) \\
&\quad - h((S_i^n)_{i\in\mathcal{A}}|(U_i^n)_{i\in\mathcal{A}}, (\phi_i^{(n)}(S_i^n))_{i\in\mathcal{A}}) \\
&= \frac{n}{2}\log((2\pi e)^k \det(\Gamma_U^{(k)})) + \frac{n}{2}\log((2\pi e)^k \det(\Lambda_W^{(k)})) \\
&\quad - h((U_i^n)_{i\in\mathcal{A}}|(\phi_i^{(n)}(S_i^n))_{i\in\mathcal{A}}) \\
&\quad - h((S_i^n)_{i\in\mathcal{A}}|(U_i^n)_{i\in\mathcal{A}}, (\phi_i^{(n)}(S_i^n))_{i\in\mathcal{A}}), \qquad (3.3)
\end{aligned}$$

where $\Gamma_U^{(k)}$ and $\Lambda_W^{(k)}$ denote the leading $k \times k$ principal submatrices of $\Gamma_U^{(j)}$ and $\Lambda_W^{(j)}$, respectively. For $t = 1, \cdots, n$, let

$$\begin{aligned}
\Sigma_{\mathcal{A}}(t) &\triangleq \mathbb{E}[(U_i(t) - \hat{U}_{i,\mathcal{A}}(t))_{i\in\mathcal{A}}^T (U_i(t) - \hat{U}_{i,\mathcal{A}}(t))_{i\in\mathcal{A}}], \\
\Delta_{\mathcal{A}}(t) &\triangleq \mathbb{E}[(S_i(t) - \tilde{S}_{i,\mathcal{A}}(t))_{i\in\mathcal{A}}^T (S_i(t) - \tilde{S}_{i,\mathcal{A}}(t))_{i\in\mathcal{A}}],
\end{aligned}$$

where

$$\hat{U}_{i,\mathcal{A}}(t) \triangleq \mathbb{E}[U_i(t)|(\phi_{i'}^{(n)}(S_{i'}^n))_{i'\in\mathcal{A}}], \quad i \in \mathcal{A},$$
$$\tilde{S}_{i,\mathcal{A}}(t) \triangleq \mathbb{E}[S_i(t)|(U_{i'}^n)_{i'\in\mathcal{A}}, (\phi_{i'}^{(n)}(S_{i'}^n))_{i'\in\mathcal{A}}], \quad i \in \mathcal{A}.$$

Moreover, let

$$\Sigma_{\mathcal{A}} \triangleq \frac{1}{n}\sum_{t=1}^{n} \Sigma_{\mathcal{A}}(t),$$
$$\Delta_{\mathcal{A}} \triangleq \frac{1}{n}\sum_{t=1}^{n} \Delta_{\mathcal{A}}(t).$$

It can be verified that

$$\begin{aligned}
&h((U_i^n)_{i\in\mathcal{A}}|(\phi_i^{(n)}(S_i^n))_{i\in\mathcal{A}}) \\
&= \sum_{t=1}^{n} h((U_i(t))_{i\in\mathcal{A}}|(\phi_i^{(n)}(S_i^n))_{i\in\mathcal{A}}, (U_i^{t-1})_{i\in\mathcal{A}}) \\
&\leq \sum_{t=1}^{n} h((U_i(t))_{i\in\mathcal{A}}|(\phi_i^{(n)}(S_i^n))_{i\in\mathcal{A}}) \\
&= \sum_{t=1}^{n} h((U_i(t) - \hat{U}_{i,\mathcal{A}}(t))_{i\in\mathcal{A}}|(\phi_i^{(n)}(S_i^n))_{i\in\mathcal{A}}) \\
&\leq \sum_{t=1}^{n} h((U_i(t) - \hat{U}_{i,\mathcal{A}}(t))_{i\in\mathcal{A}}) \\
&\leq \sum_{t=1}^{n} \frac{1}{2}\log((2\pi e)^k \det(\Sigma_{\mathcal{A}}(t))) && (3.4) \\
&\leq \frac{n}{2}\log((2\pi e)^k \det(\Sigma_{\mathcal{A}})), && (3.5)
\end{aligned}$$

where (3.4) is due to the maximum differential entropy lemma [43, p. 21], and (3.5) is due to the concavity of the log-determinant function. Similarly, we have

$$\begin{aligned} & h((S_i^n)_{i\in\mathcal{A}}|(U_i^n)_{i\in\mathcal{A}}, (\phi_i^{(n)}(S_i^n))_{i\in\mathcal{A}}) \\ & \leq \frac{n}{2}\log((2\pi e)^k \det(\Delta_{\mathcal{A}})). \end{aligned} \tag{3.6}$$

Combining (3.1), (3.3), (3.5), and (3.6) gives

$$\frac{1}{2k}\log\frac{\det(\Gamma_U^{(k)})\det(\Lambda_W^{(k)})}{\det(\Sigma_{\mathcal{A}})\det(\Delta_{\mathcal{A}})} \leq r+\epsilon. \tag{3.7}$$

For $t = 1, \cdots, n$, let

$$D_{\mathcal{A}}(t) \triangleq \mathbb{E}[(S_i(t)-\hat{S}_{i,\mathcal{A}}(t))_{i\in\mathcal{A}}^T(S_i(t)-\hat{S}_{i,\mathcal{A}}(t))_{i\in\mathcal{A}}],$$

where

$$\hat{S}_{i,\mathcal{A}}(t) \triangleq \mathbb{E}[S_i(t)|(\phi_{i'}^{(n)}(S_{i'}^n))_{i'\in\mathcal{A}}], \quad i\in\mathcal{A}.$$

Moreover, let

$$D_{\mathcal{A}} \triangleq \frac{1}{n}\sum_{t=1}^{n} D_{\mathcal{A}}(t).$$

Clearly, we have

$$0 \prec D_{\mathcal{A}} \preceq \Gamma_S^{(k)}. \tag{3.8}$$

Furthermore, as shown in Appendix B,

$$\begin{aligned}\Sigma_{\mathcal{A}} &= \Gamma_U^{(k)}(\Gamma_S^{(k)})^{-1}D_{\mathcal{A}}(\Gamma_S^{(k)})^{-1}\Gamma_U^{(k)} + \Gamma_U^{(k)} \\ &\quad - \Gamma_U^{(k)}(\Gamma_S^{(k)})^{-1}\Gamma_U^{(k)}, \quad (3.9)\\ \Delta_{\mathcal{A}} &\preceq (D_{\mathcal{A}}^{-1} + (\Lambda_W^{(k)})^{-1} - (\Gamma_S^{(k)})^{-1})^{-1}. \quad (3.10)\end{aligned}$$

The argument for (3.9) can also be leveraged to prove

$$\begin{aligned}&\frac{1}{n}\sum_{i\in\mathcal{A}}\sum_{t=1}^{n}\mathbb{E}[(X_i(t) - \hat{X}_{i,\mathcal{A}}(t))^2] \\ &= \mathrm{tr}(\Gamma_X^{(k)}(\Gamma_S^{(k)})^{-1}D_{\mathcal{A}}(\Gamma_S^{(k)})^{-1}\Gamma_X^{(k)} + \Gamma_X^{(k)} \\ &\quad - \Gamma_X^{(k)}(\Gamma_S^{(k)})^{-1}\Gamma_X^{(k)}),\end{aligned}$$

which, together with (3.2), implies

$$\begin{aligned}&\mathrm{tr}(\Gamma_X^{(k)}(\Gamma_S^{(k)})^{-1}D_{\mathcal{A}}(\Gamma_S^{(k)})^{-1}\Gamma_X^{(k)} + \Gamma_X^{(k)} \\ &- \Gamma_X^{(k)}(\Gamma_S^{(k)})^{-1}\Gamma_X^{(k)}) \le k(d_k + \epsilon). \quad (3.11)\end{aligned}$$

For $t = 1, \cdots, n$, let

$$\begin{aligned}
&\Delta_{\{1,\cdots,j\}}(t) \\
&\triangleq \mathbb{E}[(S_1(t) - \tilde{S}_{1,\{1,\cdots,j\}}(t), \cdots, S_j(t) - \tilde{S}_{j,\{1,\cdots,j\}}(t))^T \\
&\qquad (S_1(t) - \tilde{S}_{1,\{1,\cdots,j\}}(t), \cdots, S_j(t) - \tilde{S}_{j,\{1,\cdots,j\}}(t))], \\
&D_{\{1,\cdots,j\}}(t) \\
&\triangleq \mathbb{E}[(S_1(t) - \hat{S}_{1,\{1,\cdots,j\}}(t), \cdots, S_j(t) - \hat{S}_{j,\{1,\cdots,j\}}(t))^T \\
&\qquad (S_1(t) - \hat{S}_{1,\{1,\cdots,j\}}(t), \cdots, S_j(t) - \hat{S}_{j,\{1,\cdots,j\}}(t))], \\
&\delta_i(t) \triangleq \mathbb{E}[(S_i(t) - \tilde{S}_i(t))^2], \quad i = 1, \cdots, j,
\end{aligned}$$

where

$$\begin{aligned}
&\tilde{S}_{i,\{1,\cdots,j\}}(t) \\
&\triangleq \mathbb{E}[S_i(t)|U_1^n, \cdots, U_j^n, \phi_1^{(n)}(S_1^n), \cdots, \phi_j^{(n)}(S_j^n)], \\
&\qquad\qquad\qquad i = 1, \cdots, j, \\
&\hat{S}_{i,\{1,\cdots,j\}}(t) \triangleq \mathbb{E}[S_i(t)|\phi_1^{(n)}(S_1^n), \cdots, \phi_j^{(n)}(S_j^n)], \\
&\qquad\qquad\qquad i = 1, \cdots, j, \\
&\tilde{S}_i(t) \triangleq \mathbb{E}[S_i(t)|U_i^n, \phi_i^{(n)}(S_i^n)], \quad i = 1, \cdots, j.
\end{aligned}$$

Moreover, let

$$\Delta_{\{1,\cdots,j\}} \triangleq \frac{1}{n}\sum_{t=1}^{n}\Delta_{\{1,\cdots,j\}}(t),$$
$$D_{\{1,\cdots,j\}} \triangleq \frac{1}{n}\sum_{t=1}^{n}D_{\{1,\cdots,j\}}(t),$$
$$\delta_i \triangleq \sum_{t=1}^{n}\delta_i(t), \quad i=1,\cdots,j.$$

The argument for (3.10) and (3.11) can be leveraged to show that

$$\Delta_{\{1,\cdots,j\}} \preceq (D^{-1}_{\{1,\cdots,j\}} + (\Lambda_W^{(j)})^{-1} - (\Gamma_S^{(j)})^{-1})^{-1}, \tag{3.12}$$
$$\mathrm{tr}(\Gamma_X^{(j)}(\Gamma_S^{(j)})^{-1}D_{\{1,\cdots,j\}}(\Gamma_S^{(j)})^{-1}\Gamma_X^{(j)} + \Gamma_X^{(j)} - \Gamma_X^{(j)}(\Gamma_S^{(j)})^{-1}\Gamma_X^{(j)}) \leq j(d_j+\epsilon). \tag{3.13}$$

It is also clear that

$$0 < \delta_i, \quad i=1,\cdots,\ell. \tag{3.14}$$

Furthermore, in view of the fact that $S_i^n = U_i^n + W_i^n$, $i=1,\cdots,j$, and that $(U_1^n,\cdots,U_j^n),(W_1^n,\cdots,W_j^n)$ are mutually independent, we must have

$$\Delta_{\mathcal{A}} = \mathrm{diag}^{(k)}(\delta_i)_{i\in\mathcal{A}}, \tag{3.15}$$
$$\Delta_{\{1,\cdots,j\}} = \mathrm{diag}^{(j)}(\delta_1,\cdots,\delta_j). \tag{3.16}$$

Combining (3.7)–(3.16), sending $\epsilon \to 0$, and invoking a symmetrization and convexity argument shows that there exist $D^{(k)}$, $D^{(j)}$, and δ satisfying the following set of inequalities

$$\frac{1}{2k}\log\frac{\det(\Gamma_U^{(k)})}{\det(\Sigma^{(k)})}+\frac{1}{2}\log\frac{\lambda_W}{\delta}\leq r, \tag{3.17}$$

$$0 \prec D^{(k)} \preceq \Gamma_S^{(k)}, \tag{3.18}$$

$$0 < \delta, \tag{3.19}$$

$$\begin{aligned}&\mathrm{diag}^{(k)}(\delta,\cdots,\delta)\\ &\preceq((D^{(k)})^{-1}+(\Lambda_W^{(k)})^{-1}-(\Gamma_S^{(k)})^{-1})^{-1},\end{aligned} \tag{3.20}$$

$$\begin{aligned}&\mathrm{tr}(\Gamma_X^{(k)}(\Gamma_S^{(k)})^{-1}D^{(k)}(\Gamma_S^{(k)})^{-1}\Gamma_X^{(k)}+\Gamma_X^{(k)}\\ &-\Gamma_X^{(k)}(\Gamma_S^{(k)})^{-1}\Gamma_X^{(k)})\leq kd_k,\end{aligned} \tag{3.21}$$

$$\begin{aligned}&\mathrm{diag}^{(j)}(\delta,\cdots,\delta)\\ &\preceq((D^{(j)})^{-1}+(\Lambda_W^{(j)})^{-1}-(\Gamma_S^{(j)})^{-1})^{-1},\end{aligned} \tag{3.22}$$

$$\begin{aligned}&\mathrm{tr}(\Gamma_X^{(j)}(\Gamma_S^{(j)})^{-1}D^{(j)}(\Gamma_S^{(j)})^{-1}\Gamma_X^{(j)}+\Gamma_X^{(j)}\\ &-\Gamma_X^{(j)}(\Gamma_S^{(j)})^{-1}\Gamma_X^{(j)})\leq jd_j,\end{aligned} \tag{3.23}$$

where

$$\begin{aligned}D^{(k)}&=\Theta^{(k)}\mathrm{diag}^{(k)}(d_1^{(k)},d_2^{(k)},\cdots,d_2^{(k)})(\Theta^{(k)})^T,\\ D^{(j)}&=\Theta^{(j)}\mathrm{diag}^{(j)}(d_1^{(j)},d_2^{(j)},\cdots,d_2^{(j)})(\Theta^{(j)})^T\end{aligned}$$

for some $d_1^{(k)}, d_2^{(k)}, d_1^{(j)}, d_2^{(j)}$, and

$$\begin{aligned}\Sigma^{(k)}\triangleq\Gamma_U^{(k)}(\Gamma_S^{(k)})^{-1}D^{(k)}(\Gamma_S^{(k)})^{-1}\Gamma_U^{(k)}+\Gamma_U^{(k)}\\ -\Gamma_U^{(k)}(\Gamma_S^{(k)})^{-1}\Gamma_U^{(k)}.\end{aligned}$$

Equivalently, (3.17)–(3.23) can be written as

$$\frac{1}{2k}\log\frac{(\lambda_{S,1}^{(k)})^2}{(\lambda_{S,1}^{(k)}-\lambda_W)d_1^{(k)}+\lambda_{S,1}^{(k)}\lambda_W}$$
$$+\frac{k-1}{2k}\log\frac{\lambda_{S,2}^2}{(\lambda_{S,2}-\lambda_W)d_2^{(k)}+\lambda_{S,2}\lambda_W}+\frac{1}{2}\log\frac{\lambda_W}{\delta}$$
$$\leq r, \tag{3.24}$$

$$0<d_1^{(k)}\leq\lambda_{S,1}^{(k)}, \tag{3.25}$$

$$0<d_2^{(k)}\leq\lambda_{S,2}, \tag{3.26}$$

$$0<\delta, \tag{3.27}$$

$$\delta\leq((d_1^{(k)})^{-1}+\lambda_W^{-1}-(\lambda_{S,1}^{(k)})^{-1})^{-1}, \tag{3.28}$$

$$\delta\leq((d_2^{(k)})^{-1}+\lambda_W^{-1}-\lambda_{S,2}^{-1})^{-1}, \tag{3.29}$$

$$(\lambda_{X,1}^{(k)})^2(\lambda_{S,1}^{(k)})^{-2}d_1^{(k)}+\lambda_{X,1}^{(k)}-(\lambda_{X,1}^{(k)})^2(\lambda_{S,1}^{(k)})^{-1}$$
$$+(k-1)(\lambda_{X,2}^2\lambda_{S,2}^{-2}d_2^{(k)}+\lambda_{X,2}-\lambda_{X,2}^2\lambda_{S,2}^{-1})$$
$$\leq kd_k, \tag{3.30}$$

$$\delta\leq((d_1^{(j)})^{-1}+\lambda_W^{-1}-(\lambda_{S,1}^{(j)})^{-1})^{-1}, \tag{3.31}$$

$$\delta\leq((d_2^{(j)})^{-1}+\lambda_W^{-1}-\lambda_{S,2}^{-1})^{-1}, \tag{3.32}$$

$$(\lambda_{X,1}^{(j)})^2(\lambda_{S,1}^{(j)})^{-2}d_1^{(j)}+\lambda_{X,1}^{(j)}-(\lambda_{X,1}^{(j)})^2(\lambda_{S,1}^{(j)})^{-1}$$
$$+(j-1)(\lambda_{X,2}^2\lambda_{S,2}^{-2}d_2^{(j)}+\lambda_{X,2}-\lambda_{X,2}^2\lambda_{S,2}^{-1})$$
$$\leq jd_j. \tag{3.33}$$

When $\lambda_{S,1}^{(j)}\geq\lambda_{S,2}>0$, we can send $\lambda_W\to\lambda_{S,2}$ and deduce from (3.24), (3.28), (3.29), (3.31), and (3.32) that

$$\eta(d_1^{(k)}, d_2^{(k)}, \delta) \leq r, \tag{3.34}$$

$$\delta \leq ((d_1^{(k)})^{-1} + \lambda_{S,2}^{-1} - (\lambda_{S,1}^{(k)})^{-1})^{-1}, \tag{3.35}$$

$$\delta \leq d_2^{(k)}, \tag{3.36}$$

$$\delta \leq ((d_1^{(j)})^{-1} + \lambda_{S,2}^{-1} - (\lambda_{S,1}^{(j)})^{-1})^{-1}, \tag{3.37}$$

$$\delta \leq d_2^{(j)}, \tag{3.38}$$

where

$$\begin{aligned}&\eta(d_1^{(k)}, d_2^{(k)}, \delta)\\ &\triangleq \frac{1}{2k}\log\frac{(\lambda_{S,1}^{(k)})^2}{(\lambda_{S,1}^{(k)} - \lambda_{S,2})d_1^{(k)} + \lambda_{S,1}^{(k)}\lambda_{S,2}} + \frac{1}{2}\log\frac{\lambda_{S,2}}{\delta}.\end{aligned}$$

Furthermore, combining (3.33), (3.37), and (3.38) gives

$$\begin{aligned}d_j \geq\ & \frac{1}{j}((\lambda_{X,1}^{(j)})^2(\lambda_{S,1}^{(j)})^{-2}(\delta^{-1} + (\lambda_{S,1}^{(j)})^{-1} - \lambda_{S,2}^{-1})^{-1}\\ &+ \lambda_{X,1}^{(j)} - (\lambda_{X,1}^{(j)})^2(\lambda_{S,1}^{(j)})^{-1})\\ &+ \frac{j-1}{j}(\lambda_{X,2}^2\lambda_{S,2}^{-2}\delta + \lambda_{X,2} - \lambda_{X,2}^2\lambda_{S,2}^{-1}).\end{aligned} \tag{3.39}$$

Now consider the following convex optimization problem:

$$\min_{d_1^{(k)}, d_2^{(k)}, \delta} \eta(d_1^{(k)}, d_2^{(k)}, \delta) \qquad (\mathbf{P})$$

subject to (3.25), (3.26), (3.27), (3.35), (3.36), and (3.30). According to the Karush-Kuhn-Tucker conditions, $(d_1^{(k)}, d_2^{(k)}, \delta)$ is a minimizer of the convex optimization problem (**P**) if and only if (3.25), (3.26), (3.27), (3.35), (3.36), and (3.30) are satisfied, and there exist nonnegative a_1, a_2, b_1, b_2, c such that

$$\begin{aligned}&\frac{\lambda_{S,2} - \lambda_{S,1}^{(k)}}{2k((\lambda_{S,1}^{(k)} - \lambda_{S,2}^{(k)})d_1^{(k)} + \lambda_{S,1}^{(k)}\lambda_{S,2})} + a_1 \\ &- b_1(1 + \lambda_{S,2}^{-1}d_1^{(k)} - (\lambda_{S,1}^{(k)})^{-1}d_1^{(k)})^{-2} \\ &+ c(\lambda_{X,1}^{(k)})^2(\lambda_{S,1}^{(k)})^{-2} = 0, \end{aligned} \tag{3.40}$$

$$a_2 - b_2 + c(k-1)\lambda_{X,2}^2\lambda_{S,2}^{-2} = 0, \tag{3.41}$$

$$-\frac{1}{2\delta} + b_1 + b_2 = 0, \tag{3.42}$$

$$a_1(d_1^{(k)} - \lambda_{S,1}^{(k)}) = 0, \tag{3.43}$$

$$a_2(d_2^{(k)} - \lambda_{S,2}) = 0,, \tag{3.44}$$

$$b_1(\delta - ((d_1^{(k)})^{-1} + \lambda_{S,2}^{-1} - (\lambda_{S,1}^{(k)})^{-1})^{-1}) = 0, \tag{3.45}$$

$$b_2(\delta - d_2^{(k)}) = 0, \tag{3.46}$$

$$\begin{aligned}&c((\lambda_{X,1}^{(k)})^2(\lambda_{S,1}^{(k)})^{-2}d_1^{(k)} + \lambda_{X,1}^{(k)} - (\lambda_{X,1}^{(k)})^2(\lambda_{S,1}^{(k)})^{-1} \\ &+ (k-1)(\lambda_{X,2}^2\lambda_{S,2}^{-2}d_2^{(k)} + \lambda_{X,2} - \lambda_{X,2}^2\lambda_{S,2}^{-1}) - kd_k) \\ &= 0. \end{aligned} \tag{3.47}$$

Assume $d_k \in (d_{\min}^{(k)}, \gamma_X)$. It can be verified via algebraic manipulations that $\eta(d_1^{(k)}, d_2^{(k)}, \delta) =$

$\overline{r}(d_k)$ for

$$\begin{aligned}
d_1^{(k)} &\triangleq ((\lambda_{S,1}^{(k)})^{-1} + (\lambda_Q^{(k)})^{-1})^{-1}, \\
d_2^{(k)} &\triangleq (\lambda_{S,2}^{-1} + (\lambda_Q^{(k)})^{-1})^{-1}, \\
\delta &\triangleq (\lambda_{S,2}^{-1} + (\lambda_Q^{(k)})^{-1})^{-1},
\end{aligned} \tag{3.48}$$

where $\lambda_Q^{(k)}$ is given by (2.4). We shall identify the condition under which this specific $(d_1^{(k)}, d_2^{(k)}, \delta)$ is a minimizer of $(\mathbf{P})$. Clearly, (3.45)–(3.47) are satisfied. Moreover, in view of (3.43), (3.44) as well as the fact that $d_1^{(k)} < \lambda_{S,1}^{(k)}$ and $d_2^{(k)} < \lambda_{S,2}$, we must have

$$a_m = 0, \quad m = 1, 2,$$

which, together with (3.40)–(3.42), implies

$$\begin{aligned}
b_1 &= \frac{d_2^{(k)} - d_1^{(k)} + 2kc(\lambda_{X,1}^{(k)})^2(\lambda_{S,1}^{(k)})^{-2}(d_1^{(k)})^2}{2k(d_2^{(k)})^2}, \\
b_2 &= (k-1)c\lambda_{X,2}^2\lambda_{S,2}^{-2}, \\
c &= \frac{d_1^{(k)} + (k-1)d_2^{(k)}}{(\lambda_{X,1}^{(k)})^2(\lambda_{S,1}^{(k)})^{-2}(d_1^{(k)})^2 + (k-1)\lambda_{X,2}^2\lambda_{S,2}^{-2}(d_2^{(k)})^2} \\
&\quad \times \frac{1}{2k}.
\end{aligned}$$

It is obvious that b_2 and c are nonnegative. Therefore, it suffices to have $b_1 \geq 0$, which is equivalent to condition (2.11). Moreover, under this condition, every minimizer $(d_1^{(k)}, d_2^{(k)}, \delta)$ of $(\mathbf{P})$ must satisfy (3.48) due to the fact that $\frac{1}{2}\log\frac{\lambda_{S,2}}{\delta}$ is a strictly convex function of δ (in other words, (3.34), (3.25), (3.26), (3.27), (3.35), (3.36), and (3.30) imply that δ is uniquely determined and is given by (3.48) when $r = \overline{r}(d_k)$).

Hence, under condition (2.11), when $r = \overline{r}(d_k)$, we can deduce $d_j \geq d_j^{(k)}(d_k)$ by substituting (3.48) into (3.39).

When $\lambda_{S,2} \geq \lambda_{S,1}^{(j)} > 0$, we can send $\lambda_W \to \lambda_{S,1}^{(j)}$ and deduce from (3.24), (3.28), (3.29), (3.31), and (3.32) that

$$\hat{\eta}(d_1^{(k)}, d_2^{(k)}, \delta) \leq r, \tag{3.49}$$

$$\delta \leq ((d_1^{(k)})^{-1} + (\lambda_{S,1}^{(j)})^{-1} - (\lambda_{S,1}^{(k)})^{-1})^{-1}, \tag{3.50}$$

$$\delta \leq ((d_2^{(k)})^{-1} + (\lambda_{S,1}^{(j)})^{-1} - \lambda_{S,2}^{-1})^{-1}, \tag{3.51}$$

$$\delta \leq d_1^{(j)}, \tag{3.52}$$

$$\delta \leq ((d_2^{(j)})^{-1} + (\lambda_{S,1}^{(j)})^{-1} - \lambda_{S,2}^{-1})^{-1}, \tag{3.53}$$

where

$$\begin{aligned}
&\hat{\eta}(d_1^{(k)}, d_2^{(k)}, \delta) \\
&\triangleq \frac{1}{2k} \log \frac{(\lambda_{S,1}^{(k)})^2}{(\lambda_{S,1}^{(k)} - \lambda_{S,1}^{(j)})d_1^{(k)} + \lambda_{S,1}^{(k)}\lambda_{S,1}^{(j)}} \\
&+ \frac{k-1}{2k} \log \frac{\lambda_{S,2}^2}{(\lambda_{S,2} - \lambda_{S,1}^{(j)})d_2^{(k)} + \lambda_{S,2}\lambda_{S,1}^{(j)}} + \frac{1}{2} \log \frac{\lambda_{S,1}^{(j)}}{\delta}.
\end{aligned}$$

Furthermore, combining (3.33), (3.52), and (3.53) gives

$$\begin{aligned}
d_j \geq &\frac{1}{j}((\lambda_{X,1}^{(j)})^2(\lambda_{S,1}^{(j)})^{-2}\delta + \lambda_{X,1}^{(j)} - (\lambda_{X,1}^{(j)})^2(\lambda_{S,1}^{(j)})^{-1}) \\
&+ \frac{j-1}{j}(\lambda_{X,2}^2\lambda_{S,2}^{-2}(\delta^{-1} + \lambda_{S,2}^{-1} - (\lambda_{S,1}^{(j)})^{-1})^{-1} \\
&\quad + \lambda_{X,2} - \lambda_{X,2}^2\lambda_{S,2}^{-1}).
\end{aligned} \tag{3.54}$$

Now consider the following convex optimization problem:

$$\min_{d_1^{(k)}, d_2^{(k)}, \delta} \hat{\eta}(d_1^{(k)}, d_2^{(k)}, \delta) \qquad (\hat{\mathbf{P}})$$

subject to (3.25), (3.26), (3.27), (3.50), (3.51), and (3.30). According to the Karush-Kuhn-Tucker conditions, $(d_1^{(k)}, d_2^{(k)}, \delta)$ is a minimizer of the convex optimization problem $(\hat{\mathbf{P}})$ if and only if (3.25), (3.26), (3.27), (3.50), (3.51), and (3.30) are satisfied, and there exist nonnegative $\hat{a}_1, \hat{a}_2, \hat{b}_1, \hat{b}_2, \hat{c}$ such that

$$\begin{aligned}
&\frac{\lambda_{S,1}^{(j)} - \lambda_{S,1}^{(k)}}{2k((\lambda_{S,1}^{(k)} - \lambda_{S,1}^{(j)})d_1^{(k)} + \lambda_{S,1}^{(k)}\lambda_{S,1}^{(j)})} + \hat{a}_1 \\
&- \hat{b}_1(1 + (\lambda_{S,1}^{(j)})^{-1}d_1^{(k)} - (\lambda_{S,1}^{(k)})^{-1}d_1^{(k)})^{-2} \\
&+ \hat{c}(\lambda_{X,1}^{(k)})^2(\lambda_{S,1}^{(k)})^{-2} = 0, && (3.55) \\
&\frac{(k-1)(\lambda_{S,1}^{(j)} - \lambda_{S,2})}{2k((\lambda_{S,2} - \lambda_{S,1}^{(j)})d_2^{(k)} + \lambda_{S,2}\lambda_{S,1}^{(j)})} + \hat{a}_2 \\
&- \hat{b}_2(1 + (\lambda_{S,1}^{(j)})^{-1}d_2^{(k)} - \lambda_{S,2}^{-1}d_2^{(k)})^{-2} \\
&+ \hat{c}(k-1)\lambda_{X,2}^2\lambda_{S,2}^{-2} = 0, && (3.56) \\
&-\frac{1}{2\delta} + \hat{b}_1 + \hat{b}_2 = 0, && (3.57) \\
&\hat{a}_1(d_1^{(k)} - \lambda_{S,1}^{(k)}) = 0, && (3.58) \\
&\hat{a}_2(d_2^{(k)} - \lambda_{S,2}) = 0, , && (3.59) \\
&\hat{b}_1(\delta - ((d_1^{(k)})^{-1} + (\lambda_{S,1}^{(j)})^{-1} - (\lambda_{S,1}^{(k)})^{-1})^{-1}) = 0, && (3.60) \\
&\hat{b}_2(\delta - ((d_2^{(k)})^{-1} + (\lambda_{S,1}^{(j)})^{-1} - \lambda_{S,2}^{-1})^{-1}) = 0, && (3.61) \\
&\hat{c}((\lambda_{X,1}^{(k)})^2(\lambda_{S,1}^{(k)})^{-2}d_1^{(k)} + \lambda_{X,1}^{(k)} - (\lambda_{X,1}^{(k)})^2(\lambda_{S,1}^{(k)})^{-1} \\
&+ (k-1)(\lambda_{X,2}^2\lambda_{S,2}^{-2}d_2^{(k)} + \lambda_{X,2} - \lambda_{X,2}^2\lambda_{S,2}^{-1}) - kd_k) \\
&= 0. && (3.62)
\end{aligned}$$

Assume $d_k \in (d_{\min}^{(k)}, \gamma_X)$. It can be verified via algebraic manipulations that $\hat{\eta}(d_1^{(k)}, d_2^{(k)}, \delta) =$

$\overline{r}(d_k)$ for

$$\begin{aligned}
d_1^{(k)} &\triangleq ((\lambda_{S,1}^{(k)})^{-1} + (\lambda_Q^{(k)})^{-1})^{-1}, \\
d_2^{(k)} &\triangleq (\lambda_{S,2}^{-1} + (\lambda_Q^{(k)})^{-1})^{-1}, \\
\delta &\triangleq ((\lambda_{S,1}^{(j)})^{-1} + (\lambda_Q^{(k)})^{-1})^{-1},
\end{aligned} \tag{3.63}$$

where $\lambda_Q^{(k)}$ is given by (2.4). We shall identify the conditions under which this specific $(d_1^{(k)}, d_2^{(k)}, \delta)$ is a minimizer of $(\hat{\mathbf{P}})$. Clearly, (3.60)–(3.62) are satisfied. Moreover, in view of (3.58), (3.59) as well as the fact that $d_1^{(k)} < \lambda_{S,1}^{(k)}$ and $d_2^{(k)} < \lambda_{S,2}$, we must have

$$\hat{a}_m = 0, \quad m = 1, 2,$$

which, together with (3.55)–(3.57), implies

$$\begin{aligned}
\hat{b}_1 &= \frac{\delta - d_1^{(k)} + 2k\hat{c}(\lambda_{X,1}^{(k)})^2(\lambda_{S,1}^{(k)})^{-2}(d_1^{(k)})^2}{2k\delta^2}, \\
\hat{b}_2 &= \frac{(k-1)(\delta - d_2^{(k)}) + 2k(k-1)\hat{c}\lambda_{X,2}^2\lambda_{S,2}^{-2}(d_2^{(k)})^2}{2k\delta^2}, \\
\hat{c} &= \frac{d_1^{(k)} + (k-1)d_2^{(k)}}{(\lambda_{X,1}^{(k)})^2(\lambda_{S,1}^{(k)})^{-2}(d_1^{(k)})^2 + (k-1)\lambda_{X,2}^2\lambda_{S,2}^{-2}(d_2^{(k)})^2} \\
&\quad \times \frac{1}{2k}.
\end{aligned}$$

It is obvious that $\hat{c}$ is nonnegative. Therefore, it suffices to have $\hat{b}_1 \geq 0$ and $\hat{b}_2 \geq 0$, which are equivalent to conditions (2.13) and (2.14), respectively (note that, when $j = k$, condition (2.13) is redundant and condition (2.14) is simplified to condition (2.12)). Moreover, under these conditions, every minimizer $(d_1^{(k)}, d_2^{(k)}, \delta)$ of $(\hat{\mathbf{P}})$ must satisfy (3.63) due to the fact that $\frac{1}{2}\log\frac{\lambda_{S,1}^{(j)}}{\delta}$ is a strictly convex function of δ (in other words, (3.49), (3.25), (3.26), (3.27), (3.50), (3.51), and (3.30) imply that δ is uniquely determined and is given by (3.63) when $r = \overline{r}(d_k)$). Hence, under conditions (2.13) and (2.14), when $r = \overline{r}(d_k)$, we can deduce $d_j \geq d_j^{(k)}(d_k)$ by substituting (3.63) into (3.54).

For the degenerate case $\lambda_{S,1}^{(j)} > \lambda_{S,2} = 0$, we have

$$\overline{r}^{(k)}(d_k) = \frac{1}{2k} \log \frac{\gamma_X^2}{\gamma_S d_k - \gamma_X \gamma_Z},$$
$$d_j^{(k)}(d_k) = \frac{(j-k)\gamma_X^2 \gamma_Z + (k\gamma_S - j\gamma_Z)\gamma_X d_k}{(j\gamma_S - k\gamma_Z)\gamma_X - (j-k)\gamma_S d_k}.$$

The desired conclusion that $r \geq \overline{r}^{(k)}(d_k)$ and that $d_j \geq d_j^{(k)}(d_k)$ when $r = \overline{r}^{(k)}(d_k)$ follows from the corresponding result for the quadratic Gaussian multiple description problem [26], [35]. Note that $(k-1)\lambda_{X,2}^2(\lambda_{S,1}^{(k)})\mu^{(k)}(\mu^{(k)} - 1) + k(\lambda_{X,1}^{(k)})^2\lambda_{S,2}^2 = 0$ (consequently, condition (2.11) is satisfied) when $\lambda_{S,1}^{(j)} > \lambda_{S,2} = 0$. Finally, consider the degenerate case $\lambda_{S,2} > \lambda_{S,1}^{(\ell)} = 0$. It can be verified that

$$\overline{r}^{(\ell)}(d_\ell) = \frac{\ell - 1}{2\ell} \log \frac{(\ell-1)\lambda_{X,2}^2}{\ell\lambda_{S,2}d_\ell - (\ell-1)\lambda_{X,2}\lambda_{Z,2}},$$

which coincides with the rate-distortion function (normalized by ℓ) of the corresponding centralized remote source coding problem. Therefore, we must have $r \geq \overline{r}^{(\ell)}(d_\ell)$. Also, note that $(\lambda_{X,1}^{(\ell)})^2\lambda_{S,2}^2\nu^{(\ell)}(\nu^{(\ell)} - 1) + \ell\lambda_{X,2}^2(\lambda_{S,1}^{(\ell)})^2 = 0$ (consequently, condition (2.12) is satisfied for $k = \ell$) when $\lambda_{S,2} > \lambda_{S,1}^{(\ell)} = 0$. This completes the proof of Theorem 2.

Chapter 4

Conclusion And Future Work

We have studied the problem of robust distributed compression of correlated Gaussian sources in a symmetric setting and obtained a characterization of certain extremal points of the rate-distortion region. The following conclusions of the thesis were drawn from the studies.

1. For the case of all ℓ encoders, when the correlation coefficient ρ_s of the convariance matrix is a non-negative real number, $r^{(\ell)}(d_\ell)$ is characterized at least for all d_ℓ sufficiently close to $d_{\min}^{(\ell)}$ unless $\lambda_{X,1}^{(\ell)} = 0$ and $\lambda_{S,1}^{(\ell)} > \lambda_{S,2}$.

2. For the case of all ℓ encoders, when ρ_s is a non-positive real number, $r^{(\ell)}(d_\ell)$ is characterized at least for all d_ℓ sufficiently close to $d_{\min}^{(\ell)}$ unless $\lambda_{X,2}^{(\ell)} = 0$ and $\lambda_{S,2} > \lambda_{S,1}^{(\ell)}$.

3. For the case of k encoders out of ℓ encoders, when ρ_s is a non-negative real number, the condition (2.11) in chapter 2 holds at least for all d_k sufficiently close to $d_{\min}^{(k)}$ unless $\lambda_{X,1}^{(k)} = 0$ and $\lambda_{S,1}^{(k)} > \lambda_{S,2}$.

4. For the case of k encoders out of ℓ encoders, when ρ_s is a non-positive real number, the condition (2.12) holds at least for all d_k sufficiently close to $d_{\min}^{(k)}$ unless $\lambda_{X,2} = 0$ and $\lambda_{S,2} > \lambda_{S,1}^{(k)}$.

5. For the case of k encoders out of ℓ encoders, when ρ_s is a non-positive real number, the condition $\lambda_{S,1}^{(j)} > 0$ can be potentially violated ($\lambda_{S,1}^{(j)} = 0$) only when $j = \ell$.

6. For the case of k encoders out of ℓ encoders, when ρ_s is a non-positive real number and $\lambda_{S,1}^{(j)} > 0$, if $\lambda_{X,1}^{(k)} > 0$, then the condition (2.13) holds at least for d_k sufficiently close to $d_{\min}^{(k)}$; if $\lambda_{X,1}^{(k)} = 0$, which implies $k = j = \ell$, then the condition (2.13) always holds. The condition (2.14) holds at least for d_k sufficiently close to $d_{\min}^{(k)}$ unless $\lambda_{X,2} = 0$ and $\lambda_{S,2} > \lambda_{S,1}^{(j)}$.

For the future work of the related study, it is expected that one can make further progress by integrating our techniques with those developed for the quadratic Gaussian multiple description problem.

Appendix A

Calculation of $d_{\min}^{(j)}$

Assuming $\Gamma_S^{(j)} \succ 0$ (i.e., $\lambda_{S,1}^{(j)} > 0$ and $\lambda_{S,2} > 0$), we have

$$\begin{aligned}
&\sum_{i=1}^{j} \mathbb{E}[(X_i - \mathbb{E}[X_i|S_1, \cdots, S_j])^2]] \\
&= \mathrm{tr}(\Gamma_X^{(j)} - \Gamma_X^{(j)}(\Gamma_S^{(j)})^{-1}\Gamma_X^{(j)}) \\
&= \mathrm{tr}(\Lambda_X^{(j)} - \Lambda_X^{(j)}(\Lambda_S^{(j)})^{-1}\Lambda_X^{(j)}) \\
&= \frac{\lambda_{X,1}^{(j)}\lambda_{Z,1}^{(j)}}{\lambda_{S,1}^{(j)}} + (j-1)\frac{\lambda_{X,2}\lambda_{Z,2}}{\lambda_{S,2}},
\end{aligned}$$

from which the desired result follows immediately. The degenerate case $\lambda_{S,1}^{(j)} = 0$ or $\lambda_{S,2} = 0$ can be handled by performing the above analysis in a suitable subspace.

Appendix B

Proof of (3.9) and (3.10)

For $t = 1, \cdots, n$,

$$
\begin{aligned}
(G_{i,\mathcal{A}}(t))^T_{i\in\mathcal{A}} &\triangleq (U_i(t))^T_{i\in\mathcal{A}} - \mathbb{E}[(U_i(t))^T_{i\in\mathcal{A}} | (S_i(t))^T_{i\in\mathcal{A}}] \\
&= (U_i(t))^T_{i\in\mathcal{A}} - \Gamma_U^{(k)} (\Gamma_S^{(k)})^{-1} (S_i(t))^T_{i\in\mathcal{A}},
\end{aligned}
$$

which is an k-dimensional zero-mean Gaussian random vector with covariance $\Gamma_U^{(k)} - \Gamma_U^{(k)} (\Gamma_S^{(k)})^{-1} \Gamma_U^{(k)}$ and is independent of $(S_i^n)^T_{i\in\mathcal{A}}$. As a consequence,

$$
(\hat{U}_{i,\mathcal{A}}(t))^T_{i\in\mathcal{A}} = \Gamma_U^{(k)} (\Gamma_S^{(k)})^{-1} (\hat{S}_{i,\mathcal{A}}(t))^T_{i\in\mathcal{A}},
$$

$$
t = 1, \cdots, n.
$$

Now it can be readily verified that

$$\begin{aligned}\Sigma_{\mathcal{A}}(t) &= \Gamma_U^{(k)}(\Gamma_S^{(k)})^{-1}D_{\mathcal{A}}(t)(\Gamma_S^{(k)})^{-1}\Gamma_U^{(k)} \\ &\quad + \mathbb{E}[(G_{i,\mathcal{A}}(t))_{i\in\mathcal{A}}^T(G_{i,\mathcal{A}}(t))_{i\in\mathcal{A}}] \\ &= \Gamma_U^{(k)}(\Gamma_S^{(k)})^{-1}D_{\mathcal{A}}(t)(\Gamma_S^{(k)})^{-1}\Gamma_U^{(k)} + \Gamma_U^{(k)} \\ &\quad - \Gamma_U^{(k)}(\Gamma_S^{(k)})^{-1}\Gamma_U^{(k)}, \quad t = 1, \cdots, n,\end{aligned}$$

from which (3.9) follows immediately.

For $t = 1, \cdots, n$, we have

$$\begin{aligned}\Delta_{\mathcal{A}}(t) &\preceq \mathbb{E}[(S_i(t) - \tilde{S}'_{i,\mathcal{A}}(t))_{i\in\mathcal{A}}^T(S_i(t) - \tilde{S}'_{i,\mathcal{A}}(t))_{i\in\mathcal{A}}] \\ &= ((D_{\mathcal{A}}(t))^{-1} + (\Lambda_W^{(k)})^{-1} - (\Gamma_S^{(k)})^{-1})^{-1}, \end{aligned} \tag{B.1}$$

where $(\tilde{S}'_{i,\mathcal{A}}(t))_{i\in\mathcal{A}}^T$ denotes the linear MMSE estimator of $(S_i(t))_{i\in\mathcal{A}}^T$ based on $(\hat{S}_{i,\mathcal{A}}(t))_{i\in\mathcal{A}}^T$ and $(U_i(t))_{i\in\mathcal{A}}^T$. Since $(A^{-1} + B^{-1})^{-1}$ is matrix concave in A for $A \succ 0$ and $B \succ 0$, it follows that

$$\begin{aligned}&\frac{1}{n}\sum_{t=1}^{n}((D_{\mathcal{A}}(t))^{-1} + (\Lambda_W^{(k)})^{-1} - (\Gamma_S^{(k)})^{-1})^{-1} \\ &\preceq (D_{\mathcal{A}}^{-1} + (\Lambda_W^{(k)})^{-1} - (\Gamma_S^{(k)})^{-1})^{-1}.\end{aligned} \tag{B.2}$$

Combing (B.1) and (B.2) proves (3.10).

Bibliography

[1] D. Slepian and J. K. Wolf. "noiseless coding of correlated information sources,". *IEEE Trans. Inf. Theory*, vol. IT-19, no. 4, pp. 471-480, Jul. 1973.

[2] Y. Oohama. "Gaussian multiterminal source coding,". *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1912-1923, Nov. 1997.

[3] Y. Oohama. "The rate-distortion function for the quadratic Gaussian CEO problem,". *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 1057-1070, May. 1998.

[4] D. Tse V. Prabhakaran and K. Ramchandran. "Rate region of the quadratic Gaussian CEO problem,". *IEEE Int. Symp. Inform. Theory (ISIT)*, p. 117, Jun/Jul. 2004.

[5] T. Berger J. Chen, X. Zhang 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. Sel. Areas Commun*, vol. 22, no. 6, pp. 977-987, Aug. 2004.

[6] Y. Oohama. "Rate-distortion theory for Gaussian multiterminal source coding systems with several side informations at the decoder,". *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2577-2593, Jul. 2005.

[7] J. Chen and T. Berger. “Successive Wyner-Ziv coding scheme and its application to the quadratic Gaussian CEO problem,”. *IEEE Trans. Inf. Theory*, vol. 54, no. 4, pp. 1586-1603, Apr. 2008.

[8] S. Tavildar A. B. Wagner and P. Viswanath. “Rate region of the quadratic Gaussian two-encoder source-coding problem,”. *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 1938-1961, May. 2008.

[9] P. Viswanath S. Tavildar and A. B. Wagner. “The Gaussian many-help-one distributed source coding problem,”. *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 564-581, Jan. 2010.

[10] J. Chen J. Wang and X. Wu. “On the sum rate of Gaussian multiterminal source coding: New proofs and results,”. *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3946-3960, Aug. 2010.

[11] Y. Yang and Z. Xiong. “On the generalized Gaussian CEO problem,”. *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 3350-3372, Jun. 2012.

[12] Y. Zhang Y. Yang and Z. Xiong. “A new sufficient condition for sum-rate tightness in quadratic Gaussian multiterminal source coding,”. *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 408-423, Jan. 2013.

[13] J. Wang and J. Chen. “Vector Gaussian two-terminal source coding,”. *IEEE Trans. Inf. Theory*, vol. 59, no. 6, pp. 3693–3708, Jun. 2013.

[14] J. Wang and J. Chen. “Vector Gaussian multiterminal source coding,”. *IEEE Trans. Inf. Theory*, vol. 60, no. 9, pp. 5533–5552, Sep. 2014.

[15] Y. Oohama. "Indirect and direct Gaussian distributed source coding problems,". *IEEE Trans. Inf. Theory*, vol. 60, no. 12, pp. 7506–7539, Dec. 2014.

[16] F. Etezadi J. Chen and A. Khisti. "Generalized Gaussian multiterminal source coding and probabilistic graphical models,". *IEEE Int. Symp. Inform. Theory (ISIT)*, pp. 719–723, Jun. 25 - 30, 2017.

[17] Y. Chang J. Wang J. Chen, L. Xie and Y. Wang. "Generalized Gaussian multiterminal source coding: The symmetric case,". *IEEE Int. Symp. Inform. Theory (ISIT)*, arXiv:1710.04750.

[18] L. Ozarow. "On a source-coding problem with two channels and three receivers,". *Bell Syst. Tech. J*, vol. 59, no. 10, pp. 1909–1921, Dec. 1980.

[19] A. A. El Gamal and T. M. Cover. "Achievable rates for multiple descriptions,". *IEEE Trans. Inf. Theory*, vol. 28, no. 6, pp. 851–857, Nov. 1982.

[20] G. Kramer R. Venkataramani and V. K. Goyal. "Multiple description coding with many channels,". *IEEE Trans. Inf. Theory*, vol. 49, no. 9, pp. 2106–2114, Sep. 2003.

[21] R. Puri S. S. Pradhan and K. Ramchandran. "n-channel symmetric multiple descriptions—Part I: (n, k) source-channel erasure codes,". *IEEE Trans. Inf. Theory*, vol. 50, no. 1, pp. 47–61, Jan. 2004.

[22] S. S. Pradhan R. Puri and K. Ramchandran. "n-channel symmetric multiple descriptions—Part II: An achievable rate-distortion region,". *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1377–1392, Apr. 2005.

[23] T. Berger J. Chen, C. Tian and S. S. Hemami. "Multiple description quantization via Gram-Schmidt orthogonalization,". *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5197–5217, Dec. 2006.

[24] H. Wang and P. Viswanath. "Vector Gaussian multiple description with individual and central receivers,". *IEEE Trans. Inf. Theory*, vol. 53, no. 6, pp. 2133–2153, Jun. 2007.

[25] J. Chen C. Tian and S. Diggavi. "Multiuser successive refinement and multiple description coding,". *IEEE Trans. Inf. Theory*, vol. 54, no. 2, pp. 921–931, Feb. 2008.

[26] Wang and P. Viswanath. "Vector Gaussian multiple description with two levels of receivers,". *IEEE Trans. Inf. Theory*, vol. 55, no. 1, pp. 401–410, Jan. 2009.

[27] C. Tian J. Chen and S. Diggavi. "Multiple description coding for stationary Gaussian sources,". *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2868–2881, Jun. 2009.

[28] J. Chen. "Rate region of Gaussian multiple description coding with individual and central distortion constraints,". *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 3991–4005, Sep. 2009.

[29] C. Tian and J. Chen. "New coding schemes for the symmetric K-description problem,". *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 5344–5365, Oct. 2010.

[30] Y. Zhang J. Chen, S. Dumitrescu and J. Wang. "Robust multiresolution coding,". *IEEE Trans. Commun*, vol. 58, no. 11, pp. 3186–3195, Nov. 2010.

[31] L. Zhao P. Cuff J. Wang, J. Chen and H. Permuter. "On the role of the refinement layer in multiple description coding and scalable coding,". *IEEE Trans. Inf. Theory*, vol. 57, no. 3, pp. 1443–1456, Mar. 2011.

[32] Y. Zhang J. Chen and S. Dumitrescu. "Gaussian multiple description coding with low-density generator matrix codes,". *IEEE Trans. Commun*, vol. 60, no. 3, pp. 676–687, Mar. 2012.

[33] J. Chen Y. Zhang, S. Dumitrescu and Z. Sun. "LDGM-based multiple description coding for finite alphabet sources,". *IEEE Trans. Commun*, vol. 60, no. 12, pp. 3671–3682, Dec. 2012.

[34] J. Sun Y. Fan, J. Wang and J. Chen. "On the generalization of natural type selection to multiple description coding,". *IEEE Trans. Commun*, vol. 61, pp. 1361–1373, Apr. 2013.

[35] S. Shao L. Song and J. Chen. "A lower bound on the sum rate of multiple description coding with symmetric distortion constraints,". *IEEE Trans. Inf. Theory*, vol. 60, no. 12, pp. 7547–7567, Dec. 2014.

[36] J. Chen Y. Xu and Q. Wang. "The sum rate of vector Gaussian multiple description coding with tree-structured covariance distortion constraints,". *IEEE Trans. Inf. Theory*, vol. 63, no. 10, pp. 6547–6560, Oct. 2017.

[37] K. Ramchandran P. Ishwar, R. Puri and S. S. Pradhan. "On rate-constrained distributed estimation in unreliable sensor networks,". *IEEE J. Sel. Areas Commun*, vol. 23, no. 4, pp. 765–775, Apr. 2005.

[38] J. Chen and T. Berger. "Robust distributed source coding,". *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3385–3398, Aug. 2008.

[39] J. Chen and A. B. Wagner. "A semicontinuity theorem and its application to network source coding,". *IEEE Int. Symp. Inform. Theory (ISIT)y*, pp. 429–433, Jul. 6 - 11, 2008.

[40] T. Berger. "Multiterminal source coding,". *The Information Theory Approach to Communications*, (CISM International Centre for Mechanical Sciences). vol. 229, G. Longo, Ed. New York, NY, USA: Springer-Verlag, pp. 171–231, 1978.

[41] S.-Y. Tung. "Multiterminal source coding,". Ph.D. dissertation, School Electr. Eng., Cornell Univ., Ithaca, NY, USA, 1978.

[42] S. B. Wicker X. Zhang, J. Chen and T. Berger. "Successive coding in multiuser information theory,". *IEEE Trans. Inf. Theory*, vol. 53, no. 6, pp. 2246–2254, Jun. 2007.

[43] A. El Gamal and Y.-H. Kim. *Network Information Theory*, Cambridge, U.K.: Cambridge University Press, 2011.