ON THE RATE-DISTORTION FUNCTION OF SYMMETRIC REMOTE GAUSSIAN MULTITERMINAL SOURCE CODING: BOUNDS AND ASYMPTOTICS

ON THE RATE-DISTORTION FUNCTION OF SYMMETRIC REMOTE GAUSSIAN MULTITERMINAL SOURCE CODING: BOUNDS AND ASYMPTOTICS

By SIYAO ZHOU, M.S.

A THESIS

SUBMITTED TO THE ELECTRICAL AND COMPUTER ENGINEERING

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Doctor of Philosophy(2022)	McMaster University
(Electrical and Computer Engineering)	Hamilton, Ontario, Canada

TITLE:	On the Rate-Distortion Function of Symmetric Remote Gaussian Multiterminal Source Coding: Bounds and Asymptotics
AUTHOR:	Siyao Zhou M.S. (Transportation and Communication Engineering), Nanjing University of Aeronautics and Astronautics, Nanjing, China
SUPERVISOR:	Dr. Jun Chen
NUMBER OF PAGES:	xii, 104

Lay Abstract

A remote multiterminal source coding network consists of L encoders and a decoder is considered. Each encoder observes a source and sends a compressed version to the decoder. The decoder produces a joint reconstruction of target signals with the distortion below a given threshold. The minimum compression rate of this network versus the distortion threshold is referred to as the rate-distortion function. This thesis focuses the symmetric quadratic Gaussian case of the remote multiterminal source coding problem, where the observed sources can be expressed as the sum of target signals and corruptive noises which are independently generated from two symmetric multivariate Gaussian distributions. For this special case, an explicit lower bound on the rate-distortion function is established and is shown to partially coincide with the well-known Berger-Tung upper bound. Moreover, it is proved that the aforementioned lower bound is tighter than the centralized-encoding lower bound. The asymptotic behaviors of these upper and lower bounds are analyzed in the large L limit.

Abstract

Due to the development of the Internet of Things (IoT), one frequently encounters the scenarios where data collected at different sites need to be compressed and then forwarded to a fusion center for joint processing. As such data are typically correlated from one site to another, it is desirable to capitalize on the statistical dependencies to improve the compression efficiency. The multiterminal source coding problem and its variants aim to characterize the performance limits of this type of distributed compression systems.

This thesis is divided into two major parts. The first part deals with so-called remote multiterminal source coding, where L encoders compress their respective observations and send the compressed data to a central decoder for the joint reconstruction of target signals. The fundamental limit of remote multiterminal source coding is characterized by the rate-distortion function, which delineates the optimal tradeoff between the compression rate and the reconstruction distortion. For simplicity, it is assumed that the observed sources can be expressed as the sum of target signals and corruptive noises which are independently generated from two symmetric multivariate Gaussian distributions. For this special case, an explicit lower bound on the rate-distortion function is established and is shown to match the well-known Berger-Tung upper bound in some distortion regimes. The asymptotic gap between the upper and lower bounds is computed in the large L limit.

The second part considers the centralized encoding setting where the L sources are jointly observed and compressed by a single encoder. The rate-distortion function for this setting is completely characterized and is leveraged as a rate-distortion lower bound for the symmetric remote Gaussian multiterminal source coding problem in view of the fact that centralized encoding is more powerful than distributed encoding. It is shown that this centralized-encoding lower bound is not as tight as the lower bound established in the first part. The asymptotic analysis of this centralized-encoding lower bound is also provided.

To my wife Jingjing Qian

Acknowledgements

I would like to express my deepest gratitude to Dr. Jun Chen for his enlightening guidance. He would always be there when I needed help. None of these works would be done without his instructions, suggestions and help. I would always be inspired by his immense knowledge and the attitude toward work he has demonstrated.

I would like to thank my committee members Dr. Dongmei Zhao, Dr. Sorina Dumitrescu, Dr. Tim Davidson, and Dr. Jian-kang Zhang for their guidance and support over the years. May Dr. Zhang rest in peace.

I would also want to express my gratitude to have the fortune to work with Huan Liu and Botao Xiao, they were always ready to share their insights and knowledge with me. I would also express my thankfulness to Yangyi Liu, Kangdi Shi, Xiaohong Liu, Zhihao Shi and Peiyao Chen for their support both in my studies and life.

I would like to express my gratefulness to the China Scholarship Council for providing the funding that allowed me to pursue my graduate studies abroad.

Last but not least, I would like to thank my parents for their emotional support throughout my Ph.D.. I would also like to thank my wife and colleague, Jingjing Qian. To them, I dedicate this thesis. This unique experience would be the most valuable memory in my life.

Table of Contents

List of Figures

Abbreviations and Notation

Abbreviations

IoT	Internet of Things
FL	Federated Learning
MT	Multi-terminal
CEO	Chief Executive Officer
SW	Slepian-Wolf
WZ	Wyner-Ziv
BT	Berger-Tung
UB	Upper Bound
LB	Lower Bound
RD	Rate-distortion
ZMGV	Zero-mean Gaussian Vector

SMG Symmetric Multivariate Gaussian

RV Random Variable

ML Machine Learning

MMSE Minimum Mean Square Error

Notation

$\mathbb{E}(.)$ Expectation operator

$(.)^T$ Transpose operator

$\mathbf{tr}(.)$ Trace operator

$\mathbf{det}(.)$ Determinant operator

$\mathbf{diag}^{(L)}(\kappa_1, \ldots, \kappa_L)$

$L \times L$ diagonal matrix with diagonal entries $\kappa_1, \ldots, \kappa_L$

$\mathbf{1}_L$ L-dimensional all-one row vector

X^n Abbreviation of $(X_1, \ldots, X_n)$

$(\omega_{a_\ell})_{\ell \in \mathcal{A}}$ $(\omega_{a_1}, \ldots, \omega_{a_L})$ for a set $\mathcal{A}$ with elements $(1, \cdots, L)$,

$|\mathcal{X}|$ Cardinality of a set $\mathcal{X}$

$\{X_\ell^{(n)}\}$ L dimension random sequences with lenth n

Σ Covariance matrix of multivariate random variables

I Mutual information

Declaration of Academic Achievement

1. Wang, Y., Xie, L., Zhou, S., Wang, M. and Chen, J., 2019. Asymptotic rate-distortion analysis of symmetric remote Gaussian source coding: Centralized encoding vs. distributed encoding. Entropy, 21(2), p.213.

2. Xie, L., Tu, X., Zhou, S. and Chen, J., 2020. Generalized Gaussian Multiterminal Source Coding in the High-Resolution Regime. IEEE Transactions on Communications, 68(6), pp.3782-3791.

3. Zhou, S., Salehkalaibar, S., Qian, J., Chen, J., Shi, W., Ge, Y. and Tong, W., 2022. On Distributed Lossy Coding of Symmetrically Correlated Gaussian Sources. IEEE Transactions on Communications, 70(12), pp. 7957-7972.

4. Zhou, S., Qian, J., Chen, J., Symmetric Remote Gaussian Source Coding with a Centralized Encoder, submitted for publication.

Chapter 1

Introduction

1.1 Multiterminal Source Coding

The research on multiterminal source coding can be traced back to the seminal paper by Slepian and Wolf [26]. They proved a counterintuitive result that distributed compression of two correlated sources can be as efficient as joint compression in terms of the minimum achievable sum rate. The problem of lossless source coding with a helper was studied by Wyner [36] and Ahlswede and Korner [1] independently. Kobayashi and Han [15] found a unified description of the rate region of a more general class of multiterminal source coding systems, which subsumes the results in [37, 1, 36, 24, 17] as special cases. Witsenhausen and Wyner [35] investigated the property of a certain function arising from [36], which was later known as the information bottleneck function. Their work together with [1] inspired the landmark paper by Wyner and Ziv [37], which initiated the study of lossy compression in the context of multiterminal source coding. Specifically, in [37], Wyner and Ziv considered the problem of lossy source coding with side information only available at the decoder, which can be viewed as a generalization of the classical rate-distortion theory [25] where side information is not available at all and should be distinguished from the case [2] where side information is available at both the encoder and decoder. They provided a complete characterization of the rate-distortion function for this problem, which involves an auxiliary random variable with a certain Markov property. A concise proof of this result can be found in the standard textbook on network information theory by El Gamal and Kim [13]. The achievability part of the proof relies on the idea of random binning originated in [26] and dcrystallized by Cover [10]. A direct generalization of the Wyner-Ziv result to the two-terminal lossy source coding setting was given by Berger [3] and Tung [27]. However, for that setting, they only established an inner bound and an outer bound of the rate-distortion region. Characterizing the exact rate-distortion region is a longstanding open problem. Berger et al. [5] established an upper bound on the rate-distortion function for source coding with partial side information at the decoder. Later Kapsi and Berger [16] considered a general multiterminal source coding problem where the availability of partial side information is optional. Berger and Yeung [4] tackled the problem of two-terminal source coding with one of the sources required to be reconstructed perfectly. Using the entropy power inequality, Oohama [19] solved the problem of quadratic Gaussian source coding with a helper, which in turn provides a partial characterization of the rate region of the quadratic Gaussian two-terminal source coding problem. This result generalizes an intriguing observation made by Wyner and Ziv [38] that for the quadratic Gaussian case, the rate-distortion function remains the same even if the encoder has access to the decoder side information.

1.2 The CEO Problem

Soon after the celebrated works of Slepian and Wolf [26] and Wyner and Ziv [37], Berger [3] and Tung [27] launched the investigation of multiterminal lossy source coding. They considered the case where the target sources are directly available at the encoders. A variant of the Berger-Tung problem was studied by Flynn and Gray [14], where the encoders can only observe noise corrupted versions of target sources. We shall distinguish the above two classes of problems as direct and indirect (or remote) multiterminal source coding problems, respectively. Note that the latter problem is often referred to as the CEO problem [6, 28] when the noisy observations at the encoders are conditionally independent given the target sources. The CEO problem was first formulated in [6]. Its quadratic Gaussian version, first studied by Viswanathan and Berger [28], has received particular attention. A major progress on the quadratic Gaussian CEO problem was made by Oohama [20], who derived an explicit expression of the rate-distortion function in the asymptotic regime where the number of encoders tends to infinity. The rate-distortion region for this problem was completely characterized by Prabhakaran et al. [23] and Oohama [21]. Their success can be largely attributed to the use of Shannon's entropy power inequality as a bounding technique for establishing the converse coding theorem.

1.3 Remote Multiterminal Source Coding

As mentioned before, the direct multiterminal source coding problem was first studied by Berger [3] and Tung [27], who established the best known achievable rate-distortion region, commonly known as the Berger-Tung inner bound. Later, the indirect multiterminal source coding problem was investigated by Yamamoto and Itoh [40] and Flynn and Gray [14]. Oohama [19] determined the rate region of a relaxed version of the quadratic Gaussian two-terminal source coding problem, where the distortion constraint is only imposed on one of the two sources. This result, together with the Berger-Tung inner bound, yielded a partial characterization of the boundary of the rate region of the original Gaussian two-terminal source coding problem. It remained to show that the minimum sum rate of the Berger-Tung inner bound is tight for the purpose of characterizing the whole rate region. This was accomplished by Wagner et al. [29] via the construction of a composite lower bound. Specifically, they showed that coupling the cooperative lower bound with the rate-distortion function of a suitably defined Gaussian CEO problem yielded the desired converse. Wang et al. [32] proposed a new method for determining the minimum sum rate of Gaussian two-terminal source coding by exploiting the relationship between the semidefinite partial order of the distortion covariance matrices associated with the MMSE estimation and the optimal linear estimation. With this method, they further derived a general lower bound on the sum rate of Gaussian multiterminal source coding and established a set of sufficient conditions under which the lower bound is tight.

One key insight from the works by Wagner et al. [29] and Wang et al. [32] is that one may create a conditional independence structure and exploit it in the converse argument even if such a structure is not explicit in the problem formulation. This suggests that it might be possible to go beyond the CEO problem to deal with more general remote multiterminal source coding problems. In this thesis, we shall show that this is indeed the case. In particular, we circumvent the technical difficulty (namely, a lack of conditional independence structures) caused by correlated noises through a fictitious signal-noise decomposition and obtain some conclusive results regarding the rate-distortion function of symmetric remote Gaussian multiterminal source coding.

Chapter 2

Symmetric Remote Gaussian Multiterminal Source Coding

2.1 Abstract

A distributed lossy compression network with L encoders and a decoder is considered. Each encoder observes a source and sends a compressed version to the decoder. The decoder produces a joint reconstruction of target signals with the mean squared error distortion below a given threshold. It is assumed that the observed sources can be expressed as the sum of target signals and corruptive noises which are independently generated from two symmetric multivariate Gaussian distributions. The minimum compression rate of this network versus the distortion threshold is referred to as the rate-distortion function, for which an explicit lower bound is established by solving a convex program induced by a fictitious signal-noise decomposition. Our lower bound matches the well-known Berger-Tung upper bound for some values of the distortion threshold. The asymptotic gap between the upper and lower bounds is characterized

in the large L limit.

2.2 Introduction

Recently, there has been an increase in the deployment of sensor applications in wireless networks as parts of the future Internet of Things (IoT), thanks to the decreasing cost of sensors. One of the theoretical challenges that arises in these systems is to reduce the amount of data that is transmitted in the network by processing it locally at each sensor. A possible solution to this problem is to exploit the statistical dependency among the data at different sensors to get an improved compression efficiency. The multi-terminal source coding theory aims to develop suitable schemes for that purpose and characterize the corresponding performance limits. There have been significant amount of works over the past few decades in this area, e.g., Slepian-Wolf source coding [26] for lossless compression, more recent works on Gaussian multi-terminal source coding and its variants [19, 20, 23, 33, 8, 29, 32, 30, 31, 22, 9, 21]. An interesting regime that has received particular attention (see, e.g., [20]) is when the number of encoders in the network approaches infinity. This asymptotic regime reflects the typical scenarios in sensor fusion and is also relevant to some emerging machine learning applications (esp., federated learning) that leverage distributed compression to reduce the communication cost between the central server and a massive number of edge devices for training a global model.

In the present chapter, we study a remote multiterminal source coding system with L distributed encoders and a central decoder. Each encoder compresses its observed source sequence and forwards the compressed version to the decoder. The decoder is required to reconstruct the target signals with the mean squared error distortion below a given threshold. It is assumed that the observed sources can be expressed as the sum of target signals and corruptive noises which are generated independently according to two symmetric multivariate Gaussian distributions. We are interested in characterizing the minimum required compression rate as a function of the distortion threshold, which is known as the rate-distortion function. Our setup is different from the Gaussian CEO problem [6] in two aspects. Firstly, the target signals are assumed to form a vector process. Secondly, the noises across different encoders are allowed to be correlated with each other. Notice that these two relaxations do not exist in the original Gaussian CEO problem where the target signal is a scalar process and the noises across different encoders are independent.

As a main contribution of this chapter, we establish an explicit lower bound on the rate-distortion function of symmetric remote Gaussian multiterminal source coding by solving a convex program induced by a fictitious signal-noise decomposition and make a systematic comparison with the well-known Berger-Tung upper bound [13, Thm 12.1]. It should be mentioned that the symmetry assumption adopted in our setup is not critical for our analysis. It only helps us to present the rate-distortion expressions in explicit forms. We also provide an asymptotic analysis of the upper and lower bounds in the large L limit, extending Oohama's celebrated result [20] for the Gaussian CEO problem.

The rest of this chapter is organized as follows. The system model and some preliminaries are presented in Chapter 2.3. The main results are stated in Chapter 2.4 while their proofs are given in Chapters 2.5.1, 2.5.2 and 2.5.3. Chapter 2.6 contains some concluding remarks.

2.3 System Model

Consider a multi-terminal source coding problem with L distributed encoders and a centralized decoder. There are L sources $(X_1, \ldots, X_L) \in \mathbb{R}^L$, which form a zero-mean Gaussian vector. The encoders observe the noisy versions of these sources, denoted by $(Y_1, \ldots, Y_L) \in \mathbb{R}^L$, which can be expressed as

$$Y_\ell = X_\ell + Z_\ell, \qquad \ell \in \{1, \ldots, L\}, \tag{2.3.1}$$

where $(Z_1, \ldots, Z_L)$ is a zero-mean Gaussian random vector independent of $(X_1, \ldots, X_L)$. We define $\mathbf{X} := (X_1, \ldots, X_L)^T$, $\mathbf{Y} := (Y_1, \ldots, Y_L)^T$, and $\mathbf{Z} := (Z_1, \ldots, Z_L)^T$. The distributions of $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$ are determined by their covariance matrices Σ_X, Σ_Y and Σ_Z, respectively.

The source vector $\mathbf{X}$ together with the noise vector $\mathbf{Z}$ and the corrupted version $\mathbf{Y}$ generates an i.i.d. process $\{(\mathbf{X}_i, \mathbf{Y}_i, \mathbf{Z}_i)\}$. Each encoder $\ell \in \{1, \ldots, L\}$ assigns a message $M_\ell \in \mathcal{M}_\ell$ to its observed sequence Y_ℓ^n using an encoding function $\phi_\ell^{(n)} \colon \mathbb{R}^n \to \mathcal{M}_\ell$ such that $M_\ell := \phi_\ell^{(n)}(Y_\ell^n)$. Given $(M_1, \ldots, M_L)$, the decoder produces a reconstruction $(\hat{X}_1^n, \ldots, \hat{X}_L^n) := g^{(n)}(M_1, \ldots, M_L)$ using a decoding function $g^{(n)} \colon \mathcal{M}_1 \times \ldots \mathcal{M}_L \to \mathbb{R}^{L \times n}$.

Definition 1. *A rate-distortion pair (R, D) is said to be achievable if for any $\epsilon > 0$, there exist encoding functions $\phi_\ell^{(n)}$, $\ell \in \{1, \cdots, L\}$, and a decoding function $g^{(n)}$ such that*

$$\frac{1}{n} \sum_{\ell=1}^{L} \log |\mathcal{M}_\ell| \leq R + \epsilon, \tag{2.3.2}$$

and

$$\frac{1}{nL}\sum_{\ell=1}^{L}\sum_{i=1}^{n}\mathbb{E}[(X_{\ell,i}-\hat{X}_{\ell,i})^2] \leq D+\epsilon. \tag{2.3.3}$$

For every D, let $\mathcal{R}(D)$ denote the infimum of R such that (R, D) is achievable. We shall refer to $\mathcal{R}(D)$ as the rate-distortion function.

2.3.1 Preliminaries

For a given $L \times L$ matrix

$$\Gamma := \begin{pmatrix} \alpha & \beta & \dots & \beta \\ \beta & \alpha & \dots & \beta \\ \vdots & \vdots & \dots & \vdots \\ \beta & \beta & \dots & \alpha \end{pmatrix}, \tag{2.3.4}$$

it follows by the eigenvalue decomposition that we can write

$$\Gamma = \Theta\Lambda\Theta^T, \tag{2.3.5}$$

where Θ is an arbitrary unitary matrix with the first column being $\frac{1}{\sqrt{L}}\mathbf{1}_L^T$ and

$$\Lambda := \text{diag}^{(L)}(\alpha+(L-1)\beta, \alpha-\beta, \dots, \alpha-\beta).$$

In this section, we assume that the covariance matrix Σ_*, $* \in \{X, Y, Z\}$, can be written as

$$\Sigma_* := \begin{pmatrix} \sigma_*^2 & \rho_*\sigma_*^2 & \dots & \rho_*\sigma_*^2 \\ \rho_*\sigma_*^2 & \sigma_*^2 & \dots & \rho_*\sigma_*^2 \\ \vdots & \vdots & \dots & \vdots \\ \rho_*\sigma_*^2 & \rho_*\sigma_*^2 & \dots & \sigma_*^2 \end{pmatrix}, \tag{2.3.6}$$

for some σ_* and ρ_*. Therefore, we can write

$$\Sigma_* = \Theta\Lambda_*\Theta^T, \tag{2.3.7}$$

where

$$\Lambda_* := \mathrm{diag}^{(L)}(\lambda_*, \gamma_*, \dots, \gamma_*) \tag{2.3.8}$$

with

$$\begin{aligned} \lambda_* &:= (1 + (L-1)\rho_*)\sigma_*^2, \\ \gamma_* &:= (1-\rho_*)\sigma_*^2. \end{aligned} \tag{2.3.9}$$

Note that it suffices to specify Σ_X and Σ_Y since $\Sigma_Y = \Sigma_X + \Sigma_Z$ (i.e., $\sigma_Y^2 = \sigma_X^2 + \sigma_Z^2$ and $\rho_Y\sigma_Y^2 = \rho_X\sigma_X^2 + \rho_Z\sigma_Z^2$). It is also clear that $\lambda_Y = \lambda_X + \lambda_Z$ and $\gamma_Y = \gamma_X + \gamma_Z$. To ensure that the covariance matrices are positive semi-definite and the source vector $\mathbf{X}$ is not deterministic, we assume $\sigma_X^2 > 0$, $\sigma_Z^2 \geq 0$, $\rho_X \in [-\frac{1}{L-1}, 1]$ and $\rho_Z \in [-\frac{1}{L-1}, 1]$; we further assume Σ_Y is positive definite, i.e., $\min(\lambda_Y, \gamma_Y) > 0$.

2.4 Main Results

First, we review some results of [33]. The following theorem gives an upper bound on the rate-distortion function $\mathcal{R}(D)$. Let

$$d_{\min} := \frac{\lambda_X \lambda_Z}{L\lambda_Y} + \frac{(L-1)\gamma_X\gamma_Z}{L\gamma_Y}, \tag{2.4.1}$$

and

$$\overline{\mathcal{R}}(D) := \frac{1}{2}\log\left(1+\frac{\lambda_Y}{\lambda_Q}\right) + \frac{L-1}{2}\log\left(1+\frac{\gamma_Y}{\lambda_Q}\right), \tag{2.4.2}$$

where λ_Q is a positive number satisfying

$$\lambda_X\left(1-\frac{\lambda_X}{\lambda_Y+\lambda_Q}\right) + (L-1)\gamma_X\left(1-\frac{\gamma_X}{\gamma_Y+\lambda_Q}\right) = LD. \tag{2.4.3}$$

Theorem 1 (Upper bound of Thm 2 in [33])**.** *For $D \in (d_{\min}, \sigma_X^2)$, we have*

$$\mathcal{R}(D) \leq \overline{\mathcal{R}}(D). \tag{2.4.4}$$

Proof of Theorem 1: See Appendix A.1.

Remark 1. *It can be observed that $\overline{\mathcal{R}}(D)$, given in* (2.4.2)*, is expressed as the sum of two terms. These two terms correspond to the compression rates required for the larger eigenvalue λ_Y and the smaller eigenvalue γ_Y, respectively. The second term has the coefficient $L-1$, which is consistent with the fact that the eigenvalue γ_Y appears $L-1$ times in the diagonal matrix Λ_Y. A similar observation can be made for the distortion expression as given in* (2.4.3).

Next, we review a result of [33], which provides a lower bound on the rate-distortion function $\mathcal{R}(D)$ in the form of a minimization program. Define

$$\Omega(\alpha,\beta,\delta) := \frac{1}{2}\log\frac{\lambda_Y^2}{(\lambda_Y-\lambda_W)\alpha+\lambda_Y\lambda_W} + \frac{L-1}{2}\log\frac{\gamma_Y^2}{(\gamma_Y-\lambda_W)\beta+\gamma_Y\lambda_W} + \frac{L}{2}\log\frac{\lambda_W}{\delta}, \tag{2.4.5}$$

where $\lambda_W = \min(\lambda_Y, \gamma_Y)$. Let $\underline{\mathcal{R}}(D)$ be the solution of the following optimization problem:

$$\underline{\mathcal{R}}(D) := \min_{\alpha,\beta,\delta} \ \Omega(\alpha,\beta,\delta), \tag{2.4.6a}$$

$$s.t. \ 0 < \alpha \leq \lambda_Y, \tag{2.4.6b}$$

$$0 < \beta \leq \gamma_Y, \tag{2.4.6c}$$

$$0 < \delta, \tag{2.4.6d}$$

$$\delta \leq (\alpha^{-1} + \lambda_W^{-1} - \lambda_Y^{-1})^{-1}, \tag{2.4.6e}$$

$$\delta \leq (\beta^{-1} + \lambda_W^{-1} - \gamma_Y^{-1})^{-1}, \tag{2.4.6f}$$

$$\lambda_X^2\lambda_Y^{-2}\alpha + \lambda_X - \lambda_X^2\lambda_Y^{-1} + (L-1)(\gamma_X^2\gamma_Y^{-2}\beta + \gamma_X - \gamma_X^2\gamma_Y^{-1}) \leq LD. \tag{2.4.6g}$$

Theorem 2 (Lower bound of Thm 2 in [33])**.** *For $D \in (d_{\min}, \sigma_X^2)$, we have*

$$\mathcal{R}(D) \geq \underline{\mathcal{R}}(D). \tag{2.4.7}$$

Proof of Theorem 2: See Appendix A.2.

In the following, we derive the explicit solution of the above program. Define the following rate-distortion expressions:

$$\underline{\mathcal{R}}^c(D) = \begin{cases} \underline{\mathcal{R}}_1^c(D) & \text{if} \quad D \le \mathsf{D}_{\text{th}}^c, \\ \underline{\mathcal{R}}_2^c(D) & \text{if} \quad D > \mathsf{D}_{\text{th}}^c, \end{cases} \tag{2.4.8}$$

$$\underline{\hat{\mathcal{R}}}^c(D) := \begin{cases} \underline{\hat{\mathcal{R}}}_1^c(D) & \text{if} \quad D \le \hat{\mathsf{D}}_{\text{th}}^c, \\ \underline{\hat{\mathcal{R}}}_2^c(D) & \text{if} \quad D > \hat{\mathsf{D}}_{\text{th}}^c, \end{cases} \tag{2.4.9}$$

where

$$\begin{aligned} \underline{\mathcal{R}}_1^c(D) :=& \frac{L+1}{2} \log(L+1)\gamma_X^2\gamma_Y^{-1}\Big(LD - \lambda_X - (L-1)(\gamma_X - \gamma_X^2\gamma_Y^{-1}) \\ &+ \lambda_X^2\lambda_Y^{-2}(\lambda_Y + (\gamma_Y^{-1} - \lambda_Y^{-1})^{-1})\Big)^{-1} \\ &+ \frac{1}{2}\log \lambda_X^2\gamma_X^{-2}(\lambda_Y\gamma_Y^{-1} - 1)^{-1} + \frac{L}{2}\log\frac{L-1}{L}, \end{aligned} \tag{2.4.10}$$

$$\begin{aligned} \underline{\hat{\mathcal{R}}}_1^c(D) :=& \frac{2L-1}{2} \log(2L-1)\lambda_X^2\lambda_Y^{-1}\Big(LD - \lambda_X - (L-1)(\gamma_X - \gamma_X^2\gamma_Y^{-1}) \\ &+ \lambda_X^2\lambda_Y^{-1} + (L-1)\gamma_X^2\gamma_Y^{-2}(\lambda_Y^{-1} - \gamma_Y^{-1})^{-1}\Big)^{-1} \\ &+ \frac{L-1}{2}\log \gamma_X^2\lambda_X^{-2}(\gamma_Y\lambda_Y^{-1} - 1)^{-1} + \frac{L}{2}\log\frac{1}{L}, \end{aligned} \tag{2.4.11}$$

$$\underline{\mathcal{R}}_2^c(D) := \frac{L}{2}\log\frac{(L-1)\gamma_X^2\gamma_Y^{-1}}{LD - \lambda_X - (L-1)(\gamma_X - \gamma_X^2\gamma_Y^{-1})}, \tag{2.4.12}$$

$$\underline{\hat{\mathcal{R}}}_2^c(D) := \frac{L}{2}\log\frac{\lambda_X^2\lambda_Y^{-1}}{LD - \lambda_X - (L-1)\gamma_X + \lambda_X^2\lambda_Y^{-1}}, \tag{2.4.13}$$

and

$$\mathsf{D}^{c}_{\mathrm{th}} := \lambda_X^2(\lambda_Y - \gamma_Y)^{-1} + \frac{1}{L}((L-1)\gamma_X^2(\gamma_X^{-1} - \gamma_Y^{-1}) + \lambda_X), \tag{2.4.14}$$

$$\hat{\mathsf{D}}^{c}_{\mathrm{th}} := \gamma_X^2(\gamma_Y - \lambda_Y)^{-1} + \frac{1}{L}((L-1)\gamma_X + \lambda_X^2(\lambda_X^{-1} - \lambda_Y^{-1})). \tag{2.4.15}$$

Moreover, define the following parameters:

$$\mu_1 := \frac{1}{2} - \frac{1}{2}\sqrt{1 - \frac{4L}{L-1}\lambda_X^2\lambda_Y^{-2}\gamma_X^{-2}\gamma_Y^2}, \tag{2.4.16}$$

$$\mu_2 := \frac{1}{2} + \frac{1}{2}\sqrt{1 - \frac{4L}{L-1}\lambda_X^2\lambda_Y^{-2}\gamma_X^{-2}\gamma_Y^2}, \tag{2.4.17}$$

$$\nu_1 := \frac{1}{2} - \frac{1}{2}\sqrt{1 - 4L\gamma_X^2\lambda_Y^2\lambda_X^{-2}\gamma_Y^{-2}}, \tag{2.4.18}$$

$$\nu_2 := \frac{1}{2} + \frac{1}{2}\sqrt{1 - 4L\gamma_X^2\lambda_Y^2\lambda_X^{-2}\gamma_Y^{-2}}, \tag{2.4.19}$$

and

$$\begin{aligned}\mathsf{D}_{\mathrm{th},1} :=& \frac{1}{L}\Big(\lambda_X + (L-1)\gamma_X - \lambda_X^2(\lambda_Y - \gamma_Y)^{-1} + (L-1)\gamma_X^2(\lambda_Y - \gamma_Y)^{-1} \\ &- \mu_2(L-1)\gamma_X^2\gamma_Y^{-1}(1 - \gamma_Y\lambda_Y^{-1})^{-1} + \frac{1}{\mu_2}\lambda_X^2\lambda_Y^{-1}(\lambda_Y\gamma_Y^{-1} - 1)^{-1}\Big),\end{aligned} \tag{2.4.20}$$

$$\begin{aligned}\mathsf{D}_{\mathrm{th},2} :=& \frac{1}{L}\Big(\lambda_X + (L-1)\gamma_X - \lambda_X^2(\lambda_Y - \gamma_Y)^{-1} + (L-1)\gamma_X^2(\lambda_Y - \gamma_Y)^{-1} \\ &- \mu_1(L-1)\gamma_X^2\gamma_Y^{-1}(1 - \gamma_Y\lambda_Y^{-1})^{-1} + \frac{1}{\mu_1}\lambda_X^2\lambda_Y^{-1}(\lambda_Y\gamma_Y^{-1} - 1)^{-1}\Big),\end{aligned} \tag{2.4.21}$$

$$
\begin{aligned}
\hat{\mathsf{D}}_{\text{th},1} :=& \frac{1}{L}\Big(\lambda_X + (L-1)\gamma_X - \lambda_X^2(\lambda_Y - \gamma_Y)^{-1} + (L-1)\gamma_X^2(\lambda_Y - \gamma_Y)^{-1} \\
&- \frac{1}{\nu_2}(L-1)\gamma_X^2\gamma_Y^{-1}(1-\gamma_Y\lambda_Y^{-1})^{-1} + \nu_2\lambda_X^2\lambda_Y^{-1}(\lambda_Y\gamma_Y^{-1}-1)^{-1}\Big),
\end{aligned} \tag{2.4.22}
$$

$$
\begin{aligned}
\hat{\mathsf{D}}_{\text{th},2} :=& \frac{1}{L}\Big(\lambda_X + (L-1)\gamma_X - \lambda_X^2(\lambda_Y - \gamma_Y)^{-1} + (L-1)\gamma_X^2(\lambda_Y - \gamma_Y)^{-1} \\
&- \frac{1}{\nu_1}(L-1)\gamma_X^2\gamma_Y^{-1}(1-\gamma_Y\lambda_Y^{-1})^{-1} + \nu_1\lambda_X^2\lambda_Y^{-1}(\lambda_Y\gamma_Y^{-1}-1)^{-1}\Big).
\end{aligned} \tag{2.4.23}
$$

Theorem 3 (Lower bound)**.** *The lower bound $\underline{\mathcal{R}}(D)$ is completely characterized as follows.*

- $\lambda_Y \geq \gamma_Y$*:*

  1. *If $\lambda_X^2\gamma_Y^2 \geq \frac{L-1}{4L}\gamma_X^2\lambda_Y^2$ or if $\lambda_X^2\gamma_Y^2 < \frac{L-1}{4L}\gamma_X^2\lambda_Y^2$ and $\mu_2 \leq \frac{\gamma_Y}{\lambda_Y}$, then for $D \in (d_{\min}, \sigma_X^2)$, we have*

$$
\underline{\mathcal{R}}(D) = \overline{\mathcal{R}}(D). \tag{2.4.24}
$$

  2. *If $\lambda_X^2\gamma_Y^2 < \frac{L-1}{4L}\gamma_X^2\lambda_Y^2$, $\mu_1 \leq \frac{\gamma_Y}{\lambda_Y}$ and $\frac{\gamma_Y}{\lambda_Y} < \mu_2 < 1$, then for $D \in (d_{\min}, \sigma_X^2)$, we have*

$$
\underline{\mathcal{R}}(D) = \begin{cases} \overline{\mathcal{R}}(D), & D \leq \mathsf{D}_{th,1}, \\ \underline{\mathcal{R}}^c(D), & D > \mathsf{D}_{th,1}. \end{cases} \tag{2.4.25}
$$

  3. *If $\lambda_X^2\gamma_Y^2 < \frac{L-1}{4L}\gamma_X^2\lambda_Y^2$, $\mu_1 > \frac{\gamma_Y}{\lambda_Y}$ and $\mu_2 < 1$, then for $D \in (d_{\min}, \sigma_X^2)$, we*

have

$$\underline{\mathcal{R}}(D) = \begin{cases} \overline{\mathcal{R}}(D), & D \leq \mathsf{D}_{th,1}, \\ \underline{\mathcal{R}}^c(D), & \mathsf{D}_{th,1} < D < \mathsf{D}_{th,2}, \\ \overline{\mathcal{R}}(D), & D \geq \mathsf{D}_{th,2}. \end{cases} \tag{2.4.26}$$

4. *If* $\lambda_X^2\gamma_Y^2 < \frac{L-1}{4L}\gamma_X^2\lambda_Y^2$, $\mu_1 = 0$ *and* $\mu_2 = 1$ *(or equivalently, if* $\lambda_X = 0$*), then for* $D \in (d_{\min}, \sigma_X^2)$*, we have*

$$\underline{\mathcal{R}}(D) = \underline{\mathcal{R}}^c(D). \tag{2.4.27}$$

- $\gamma_Y \geq \lambda_Y$*:*

1. *If* $\gamma_X^2\lambda_Y^2 \geq \frac{1}{4L}\lambda_X^2\gamma_Y^2$ *or if* $\gamma_X^2\lambda_Y^2 < \frac{1}{4L}\lambda_X^2\gamma_Y^2$ *and* $\nu_2 \leq \frac{\lambda_Y}{\gamma_Y}$*, then for* $D \in (d_{\min}, \sigma_X^2)$*, we have*

$$\underline{\mathcal{R}}(D) = \overline{\mathcal{R}}(D). \tag{2.4.28}$$

2. *If* $\gamma_X^2\lambda_Y^2 < \frac{1}{4L}\lambda_X^2\gamma_Y^2$, $\nu_1 \leq \frac{\lambda_Y}{\gamma_Y}$ *and* $\frac{\lambda_Y}{\gamma_Y} < \nu_2 < 1$*, then for* $D \in (d_{\min}, \sigma_X^2)$*, we have*

$$\underline{\mathcal{R}}(D) = \begin{cases} \overline{\mathcal{R}}(D), & D \leq \hat{\mathsf{D}}_{th,1}, \\ \hat{\underline{\mathcal{R}}}^c(D), & D > \hat{\mathsf{D}}_{th,1}. \end{cases} \tag{2.4.29}$$

3. *If* $\gamma_X^2\lambda_Y^2 < \frac{1}{4L}\lambda_X^2\gamma_Y^2$, $\nu_1 > \frac{\lambda_Y}{\gamma_Y}$ *and* $\nu_2 < 1$*, then for* $D \in (d_{\min}, \sigma_X^2)$*, we have*

$$\underline{\mathcal{R}}(D) = \begin{cases} \overline{\mathcal{R}}(D), & D \leq \hat{\mathsf{D}}_{th,1}, \\ \hat{\underline{\mathcal{R}}}^c(D), & \hat{\mathsf{D}}_{th,1} < D < \hat{\mathsf{D}}_{th,2}, \\ \overline{\mathcal{R}}(D), & D \geq \hat{\mathsf{D}}_{th,2}. \end{cases} \tag{2.4.30}$$

4. *If* $\gamma_X^2\lambda_Y^2 < \frac{1}{4L}\lambda_X^2\gamma_Y^2$, $\nu_1 = 0$ *and* $\nu_2 = 1$ *(or equivalently, if* $\gamma_X = 0$*), then*

for $D \in (d_{\min}, \sigma_X^2)$*, we have*

$$\underline{\mathcal{R}}(D) = \hat{\underline{\mathcal{R}}}^c(D). \tag{2.4.31}$$

Proof. See Chapter 2.5.1. □

According to Theorem 3, under some conditions, the lower bound $\underline{\mathcal{R}}(D)$ matches the upper bound $\overline{\mathcal{R}}(D)$. The gap between the lower and upper bounds will be investigated in the following example for some values of the parameters.

Example 1: In this example, we compare the upper bound $\overline{\mathcal{R}}(D)$ with the lower bound $\underline{\mathcal{R}}(D)$. We set $L = 10$. In Fig. 2.1a and Fig. 2.1b, we plot $\overline{\mathcal{R}}(D)$ and $\underline{\mathcal{R}}(D)$ with $D \in (d_{\min}, \sigma_X^2)$ for the following three cases.

- Case 1: $\lambda_X = 0.8$, $\gamma_X = 1$, $\lambda_Y = 5$, and $\gamma_Y = 4$. In this case, we have $d_{\min} = 0.7422$ and $\sigma_X^2 = 0.98$. As can be seen from the figure, $\underline{\mathcal{R}}(D)$ coincides with $\overline{\mathcal{R}}(D)$ for all $D \in (d_{\min}, \sigma_X^2)$, so $\mathcal{R}(D)$ is completely determined.

- Case 2: $\lambda_X = 0.5$, $\gamma_X = 1$, $\lambda_Y = 6$, and $\gamma_Y = 3$. In this case, we have $d_{\min} \approx 0.646$, $\mathsf{D}_{\text{th},1} \approx 0.691$, $\mathsf{D}^c_{\text{th}} \approx 0.733$ and $\sigma_X^2 = 0.95$. As can be observed from both figures, $\underline{\mathcal{R}}(D)$ coincides with $\overline{\mathcal{R}}(D)$ for $D \in (d_{\min}, \mathsf{D}_{\text{th},1}]$ and consequently $\mathcal{R}(D)$ is determined over this interval (see the diamond-line portion of Fig. 2.1b). For $D \in (\mathsf{D}_{\text{th},1}, \mathsf{D}^c_{\text{th}}]$, $\underline{\mathcal{R}}(D)$ is characterized by $\underline{\mathcal{R}}^c_1(D)$ (see the plus-line portion of Fig. 2.1b). For $D \in (\mathsf{D}^c_{\text{th}}, \sigma_X^2)$, $\underline{\mathcal{R}}(D)$ is characterized by $\underline{\mathcal{R}}^c_2(D)$ (see the cross-line portion of Fig. 2.1b).

- Case 3: $\lambda_X = 1$, $\gamma_X = 0.45$, $\lambda_Y = 12$, and $\gamma_Y = 2.4$. In this case, we have $d_{\min} = 0.4207$, $\mathsf{D}_{\text{th},1} \approx 0.453$, $\mathsf{D}_{\text{th},2} \approx 0.489$ and $\sigma_X^2 = 0.505$. For $D \in (d_{\min}, \mathsf{D}_{\text{th},1}]$

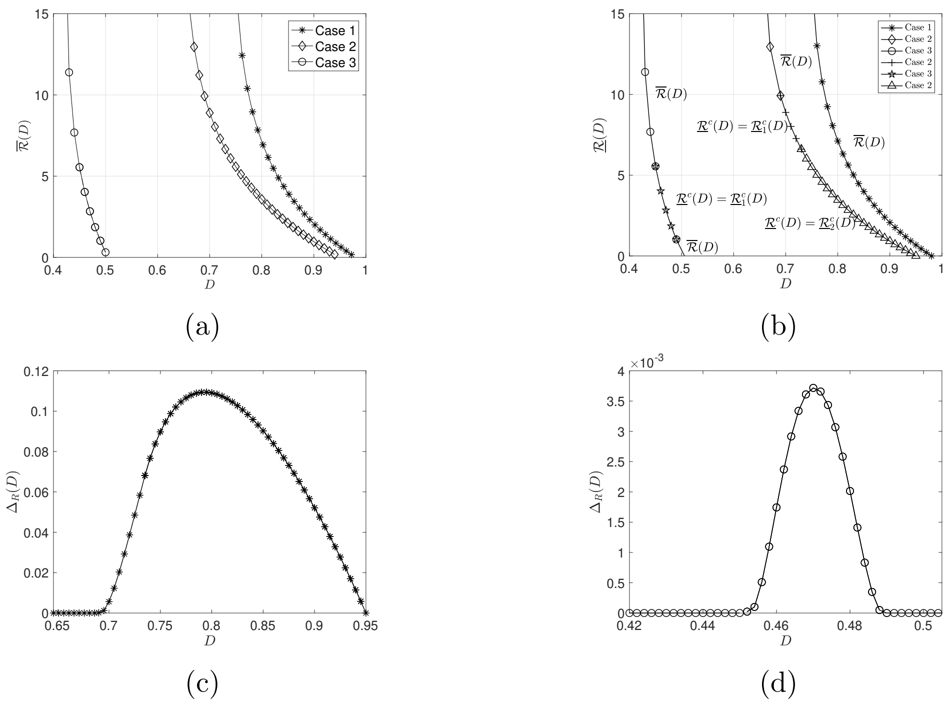

Figure 2.1: (a) Upper bound $\overline{\mathcal{R}}(D)$ with $D \in (d_{\min}, \sigma_X^2)$ for the three cases. (b) Lower bound $\underline{\mathcal{R}}(D)$ with $D \in (d_{\min}, \sigma_X^2)$ for the three cases. (c) $\Delta_R(D)$ with $D \in (d_{\min}, \sigma_X^2)$ for Case 2. (d) $\Delta_R(D)$ with $D \in (d_{\min}, \sigma_X^2)$ for Case 3.

and $D \in [\mathsf{D}_{\text{th},2}, \sigma_X^2)$, $\underline{\mathcal{R}}(D)$ coincides with $\overline{\mathcal{R}}(D)$ and consequently $\mathcal{R}(D)$ is determined over these two intervals (see the cie-line portion of Fig. 2.1b). For $D \in (\mathsf{D}_{\text{th},1}, \mathsf{D}_{\text{th},2})$, $\underline{\mathcal{R}}(D)$ is characterized by $\underline{\mathcal{R}}_1^c(D)$ (see the pentagonal-line portion of Fig. 2.1b).

As can be observed from Fig. 2.1a and Fig. 2.1b, there exists a gap between $\overline{\mathcal{R}}(D)$ and $\underline{\mathcal{R}}(D)$ in Cases 2 and 3. We plot this gap, denoted by $\Delta_R(D)$, with $D \in (d_{\min}, \sigma_X^2)$ for these two cases in Fig. 2.1c and Fig. 2.1d, respectively.

Now, we proceed to study the asymptotic behavior of the rate-distortion bounds $\overline{\mathcal{R}}(D)$ and $\underline{\mathcal{R}}(D)$ when L tends to infinity. In the discussion below, it is necessary to assume that $\rho_X, \rho_Z \in [0, 1]$. First, we perform the asymptotic analysis for $\overline{\mathcal{R}}(D)$.

Define

$$d_{\min}^{\infty} := \begin{cases} \frac{\sigma_X^2 \sigma_Z^2}{\sigma_X^2 + \sigma_Z^2}, & \rho_X \sigma_X^2 + \rho_Z \sigma_Z^2 = 0, \\ \frac{\rho_X \rho_Z \sigma_X^2 \sigma_Z^2}{\rho_X \sigma_X^2 + \rho_Z \sigma_Z^2} + \gamma_X \gamma_Z \gamma_Y^{-1}, & \rho_X \sigma_X^2 + \rho_Z \sigma_Z^2 > 0, \end{cases} \tag{2.4.32}$$

$$\mathsf{D}_{\text{th},0}^{\infty} := \frac{\rho_X \rho_Z \sigma_X^2 \sigma_Z^2}{\rho_X \sigma_X^2 + \rho_Z \sigma_Z^2} + \gamma_X, \tag{2.4.33}$$

$$\xi := \left(\frac{\rho_X}{1 - \rho_X} \right) \cdot \left(\frac{1 - \rho_Y}{\rho_Y} \right), \tag{2.4.34}$$

and

$$\overline{\mathcal{R}}^{\infty}(D) := \frac{L}{2} \log \frac{\sigma_X^4}{(\sigma_X^2 + \sigma_Z^2) D - \sigma_X^2 \sigma_Z^2}, \tag{2.4.35}$$

$$\begin{aligned} \overline{\mathcal{R}}_1^{\infty}(D) :=& \frac{L}{2} \log \frac{\gamma_X^2 \gamma_Y^{-1}}{D - d_{\min}^{\infty}} + \frac{1}{2} \log L + \frac{1}{2} \log \frac{(\rho_X \sigma_X^2 + \rho_Z \sigma_Z^2)(\mathsf{D}_{\text{th},0}^{\infty} - D)}{\gamma_X^2} \\ &+ \frac{(\mathsf{D}_{\text{th},0}^{\infty} - \xi \gamma_X^2 \gamma_Y^{-1} - D)^2}{2(\mathsf{D}_{\text{th},0}^{\infty} - D)(D - d_{\min}^{\infty})} + O\left(\frac{1}{L} \right), \end{aligned} \tag{2.4.36}$$

$$\begin{aligned} \overline{\mathcal{R}}_2^{\infty}(D) :=& \frac{1}{2} \xi \sqrt{L} + \frac{1}{4} \log L + \frac{1}{2} \log\left(\frac{\rho_X}{1 - \rho_X}\right) \\ &- \frac{\xi(\rho_X \gamma_X - \gamma_Z + (1 - \rho_X^2)\sigma_Z^2)}{4\gamma_X(\rho_X^2 \sigma_X^2 + \rho_X \rho_Z \sigma_Z^2)} + O\left(\frac{1}{\sqrt{L}} \right), \end{aligned} \tag{2.4.37}$$

$$\overline{\mathcal{R}}_3^{\infty}(D) := \frac{1}{2} \log \frac{\rho_X^2 \sigma_X^4}{(\rho_X \sigma_X^2 + \rho_Z \sigma_Z^2)(D - \mathsf{D}_{\text{th},0}^{\infty})} + \frac{(1 - \rho_Y)(\sigma_X^2 - D)}{2\rho_Y (D - \mathsf{D}_{\text{th},0}^{\infty})} + O\left(\frac{1}{L} \right). \tag{2.4.38}$$

Theorem 4 (Asymptotic Expression of Upper Bound). *1. If $\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2 = 0$, then for $D \in (d_{\min}^{\infty}, \sigma_X^2)$, we have*

$$\overline{\mathcal{R}}(D) = \overline{\mathcal{R}}^{\infty}(D). \tag{2.4.39}$$

2. If $\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2 > 0$, then for $D \in (d_{\min}^{\infty}, \sigma_X^2)$, we have

$$\overline{\mathcal{R}}(D) = \begin{cases} \overline{\mathcal{R}}_1^{\infty}(D), & D < \mathsf{D}_{th,0}^{\infty}, \\ \overline{\mathcal{R}}_2^{\infty}(D), & D = \mathsf{D}_{th,0}^{\infty}, \\ \overline{\mathcal{R}}_3^{\infty}(D), & D > \mathsf{D}_{th,0}^{\infty}. \end{cases} \tag{2.4.40}$$

Proof. See Chapter 2.5.2. □

Next, we perform the asymptotic analysis for $\underline{\mathcal{R}}(D)$. Define

$$\mathsf{D}_{\mathrm{th},1}^{\infty} := \frac{\rho_X\rho_Z\sigma_X^2\sigma_Z^2}{\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2} + \gamma_X - \frac{1+\sqrt{1-4\xi^2}}{2}\gamma_X^2\gamma_Y^{-1}, \tag{2.4.41}$$

$$\mathsf{D}_{\mathrm{th},2}^{\infty} := \frac{\rho_X\rho_Z\sigma_X^2\sigma_Z^2}{\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2} + \gamma_X - \frac{1-\sqrt{1-4\xi^2}}{2}\gamma_X^2\gamma_Y^{-1}, \tag{2.4.42}$$

and

$$\begin{aligned} \underline{\mathcal{R}}_1^{\infty}(D) :=& \frac{L+1}{2}\log\frac{\gamma_X^2\gamma_Y^{-1}}{D-d_{\min}^{\infty}} + \frac{1}{2}\log L + \frac{1}{2}\frac{(1-2\xi)\gamma_X^2\gamma_Y^{-1}}{D-d_{\min}^{\infty}} \\ &+ \frac{1}{2}\log\frac{\rho_X^2(1-\rho_Y)}{(1-\rho_X)^2\rho_Y} + O\left(\frac{1}{L}\right), \end{aligned} \tag{2.4.43}$$

$$\underline{\mathcal{R}}_2^{\infty}(D) := \frac{L}{2}\log\frac{\sigma_X^4}{\gamma_Y D - \sigma_X^2\gamma_Z} - \frac{1}{2}\frac{D-\sigma_X^2}{D-\sigma_X^2+\sigma_X^4\gamma_Y^{-1}} + O\left(\frac{1}{L}\right). \tag{2.4.44}$$

Theorem 5 (Asymptotic Expression of Lower Bound)**.** *1. If $\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2 = 0$, then for $D \in (d_{\min}^{\infty}, \sigma_X^2)$, we have*

$$\underline{\mathcal{R}}(D) = \overline{\mathcal{R}}^{\infty}(D). \tag{2.4.45}$$

2. If $\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2 > 0$, $\rho_X > 0$ and $\xi \geq \frac{1}{2}$, then for $D \in (d_{\min}^{\infty}, \sigma_X^2)$, we have

$$\underline{\mathcal{R}}(D) = \begin{cases} \overline{\mathcal{R}}_1^{\infty}(D), & D < \mathsf{D}_{th,0}^{\infty}, \\ \overline{\mathcal{R}}_2^{\infty}(D), & D = \mathsf{D}_{th,0}^{\infty}, \\ \overline{\mathcal{R}}_3^{\infty}(D), & D > \mathsf{D}_{th,0}^{\infty}. \end{cases} \tag{2.4.46}$$

3. If $\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2 > 0$, $\rho_X > 0$ and $\xi < \frac{1}{2}$, then for $D \in (d_{\min}^{\infty}, \sigma_X^2)$, we have

$$\underline{\mathcal{R}}(D) = \begin{cases} \overline{\mathcal{R}}_1^{\infty}(D), & D \leq \mathsf{D}_{th,1}^{\infty}, \\ \underline{\mathcal{R}}_1^{\infty}(D), & \mathsf{D}_{th,1}^{\infty} < D < \mathsf{D}_{th,2}^{\infty}, \\ \overline{\mathcal{R}}_1^{\infty}(D), & \mathsf{D}_{th,2}^{\infty} \leq D < \mathsf{D}_{th,0}^{\infty}, \\ \overline{\mathcal{R}}_2^{\infty}(D), & D = \mathsf{D}_{th,0}^{\infty}, \\ \overline{\mathcal{R}}_3^{\infty}(D), & D > \mathsf{D}_{th,0}^{\infty}. \end{cases} \tag{2.4.47}$$

4. If $\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2 > 0$ and $\rho_X = 0$, then for $D \in (d_{\min}^{\infty}, \sigma_X^2)$, we have

$$\underline{\mathcal{R}}(D) = \underline{\mathcal{R}}_2^{\infty}(D). \tag{2.4.48}$$

Proof. See Chapter 2.5.3. □

The following corollary provides the (asymptotic) gap between $\overline{\mathcal{R}}(D)$ and $\underline{\mathcal{R}}(D)$.

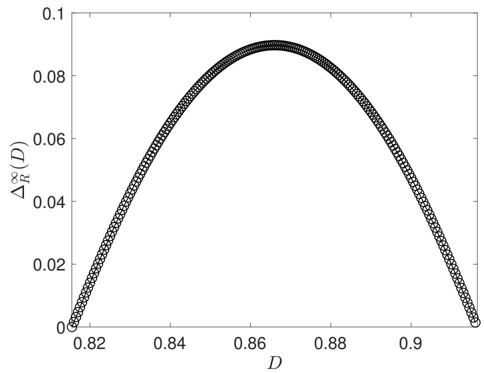

Figure 2.2: $\Delta_R^\infty(D)$ with $\rho_X = 0.3$, $\sigma_X^2 = 1$, $\rho_Y = 0.5$, $\sigma_Y^2 = 5$ and $D \in (\mathsf{D}_{\text{th},1}^\infty, \mathsf{D}_{\text{th},2}^\infty)$, where $\mathsf{D}_{\text{th},1}^\infty \approx 0.816$ and $\mathsf{D}_{\text{th},2}^\infty \approx 0.917$.

Define

$$\Delta_R^{(\infty)}(D) := \frac{(\mathsf{D}_{\text{th},1}^\infty - D)(\mathsf{D}_{\text{th},2}^\infty - D)}{2\left(\mathsf{D}_{\text{th},0}^\infty - D\right)(D - d_{\min}^\infty)} + \frac{1}{2}\log\frac{\gamma_Y^2}{\xi^2\gamma_X^4}\left(\mathsf{D}_{\text{th},0}^\infty - D\right)(D - d_{\min}^\infty). \quad (2.4.49)$$

Corollary 1. *The gap between $\overline{\mathcal{R}}(D)$ and $\underline{\mathcal{R}}(D)$ is given as follows.*

1. *If $\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2 = 0$, then for $D \in (d_{\min}^\infty, \sigma_X^2)$, we have*

$$\overline{\mathcal{R}}(D) - \underline{\mathcal{R}}(D) = 0. \quad (2.4.50)$$

2. *If $\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2 > 0$, $\rho_X > 0$ and $\xi \geq \frac{1}{2}$, then for $D \in (d_{\min}^\infty, \sigma_X^2)$, we have*

$$\overline{\mathcal{R}}(D) - \underline{\mathcal{R}}(D) = 0. \quad (2.4.51)$$

3. *If $\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2 > 0$, $\rho_X > 0$ and $\xi < \frac{1}{2}$, then for $D \in (d_{\min}^\infty, \sigma_X^2)$, we have*

$$\lim_{L\to\infty} \overline{\mathcal{R}}(D) - \underline{\mathcal{R}}(D) = \begin{cases} 0, & D \leq \mathsf{D}_{th,1}^\infty \quad or \quad D \geq \mathsf{D}_{th,2}^\infty, \\ \Delta_R^\infty(D), & \mathsf{D}_{th,1}^\infty < D < \mathsf{D}_{th,2}^\infty. \end{cases} \quad (2.4.52)$$

4. If $\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2 > 0$ and $\rho_X = 0$, then for $D \in (d_{\min}^{\infty}, \sigma_X^2)$, we have

$$\overline{\mathcal{R}}(D) - \underline{\mathcal{R}}(D) = O(\log L). \tag{2.4.53}$$

As can be seen from the above corollary, under the third condition, the lower and upper bounds asymptotically match for all D except when $\mathsf{D}_{\mathrm{th},1}^{\infty} < D < \mathsf{D}_{\mathrm{th},2}^{\infty}$. Fig. 2.2 plots the function $\Delta_R^{(\infty)}(D)$, which characterizes the asymptotic gap between $\overline{\mathcal{R}}(D)$ and $\underline{\mathcal{R}}(D)$ (as L tends to infinity) in the interval $\mathsf{D}_{\mathrm{th},1}^{\infty} < D < \mathsf{D}_{\mathrm{th},2}^{\infty}$, for some values of parameters.

2.5 Proof of Results

2.5.1 Proof of Theorem 3

Before starting the proof, we introduce another representation of $\overline{\mathcal{R}}(D)$ (defined in (2.4.2)–(2.4.3)) which will be repeatedly used in the sequel. Define

$$\lambda_I^{-1} := \lambda_Y^{-1} + \lambda_Q^{-1}, \tag{2.5.1}$$

$$\gamma_I^{-1} := \gamma_Y^{-1} + \lambda_Q^{-1}. \tag{2.5.2}$$

Corollary 2. *$\overline{\mathcal{R}}(D)$ can be alternatively expressed as*

$$\overline{\mathcal{R}}(D) = \frac{1}{2}\log(\lambda_I^{-1}\lambda_Y) + \frac{L-1}{2}\log\left(1 + \gamma_Y(\lambda_I^{-1} - \lambda_Y^{-1})\right), \tag{2.5.3}$$

where

$$\lambda_X^2\lambda_Y^{-2}\lambda_I + \lambda_X - \lambda_X^2\lambda_Y^{-1} + (L-1)(\gamma_X^2\gamma_Y^{-2}(\lambda_I^{-1} + \gamma_Y^{-1} - \lambda_Y^{-1})^{-1} + \gamma_X - \gamma_X^2\gamma_Y^{-1}) = LD,$$

or in the following form

$$\overline{\mathcal{R}}(D) = \frac{1}{2}\log\left(1 + \lambda_Y(\gamma_I^{-1} - \gamma_Y^{-1})\right) + \frac{L-1}{2}\log(\gamma_I^{-1}\gamma_Y), \tag{2.5.4}$$

where

$$\lambda_X^2\lambda_Y^{-2}(\gamma_I^{-1} + \lambda_Y^{-1} - \gamma_Y^{-1})^{-1} + \lambda_X - \lambda_X^2\lambda_Y^{-1} + (L-1)(\gamma_X^2\gamma_Y^{-2}\gamma_I + \gamma_X - \gamma_X^2\gamma_Y^{-1}) = LD. \tag{2.5.5}$$

Now, consider the optimization problem in Theorem 2 as follows:

$$\min_{\alpha,\beta,\delta} \quad \Omega(\alpha, \beta, \delta), \tag{2.5.6a}$$

$$\text{s.t.} \quad \text{constraints } (2.4.6\text{b}) - (2.4.6\text{g}). \tag{2.5.6b}$$

Based on the fact that $\lambda_Y \geq \gamma_Y$ or $\gamma_Y \geq \lambda_Y$, we get two different cases.

First, consider the case $\lambda_Y \geq \gamma_Y > 0$, where we have $\lambda_W = \gamma_Y$. Thus, the objective function reduces to

$$f(\alpha, \delta) := \frac{1}{2}\log\frac{\lambda_Y^2}{(\lambda_Y - \gamma_Y)\alpha + \lambda_Y\gamma_Y} + \frac{L}{2}\log\frac{\gamma_Y}{\delta}, \tag{2.5.7}$$

and the constraints (2.4.6b)-(2.4.6g) are simplified as follows:

$$0 < \alpha \leq \lambda_Y, \tag{2.5.8a}$$

$$0 < \beta \leq \gamma_Y, \tag{2.5.8b}$$

$$0 < \delta, \tag{2.5.8c}$$

$$\delta \leq (\alpha^{-1} + \gamma_Y^{-1} - \lambda_Y^{-1})^{-1}, \tag{2.5.8d}$$

$$\delta \leq \beta, \tag{2.5.8e}$$

$$\lambda_X^2 \lambda_Y^{-2} \alpha + \lambda_X - \lambda_X^2 \lambda_Y^{-1} + (L-1)(\gamma_X^2 \gamma_Y^{-2} \beta + \gamma_X - \gamma_X^2 \gamma_Y^{-1}) \leq LD. \tag{2.5.8f}$$

Since the objective function does not depend on parameter β, we can eliminate β from the constraints (2.5.8b), (2.5.8e) and (2.5.8f). Thus, we get the following new constraints:

$$0 < \alpha \leq \lambda_Y, \tag{2.5.9a}$$

$$0 < \delta, \tag{2.5.9b}$$

$$\delta \leq (\alpha^{-1} + \gamma_Y^{-1} - \lambda_Y^{-1})^{-1}, \tag{2.5.9c}$$

$$\delta \leq \gamma_Y, \tag{2.5.9d}$$

$$\lambda_X^2 \lambda_Y^{-2} \alpha + \lambda_X - \lambda_X^2 \lambda_Y^{-1} + (L-1)(\gamma_X^2 \gamma_Y^{-2} \delta + \gamma_X - \gamma_X^2 \gamma_Y^{-1}) \leq LD. \tag{2.5.9e}$$

Given constraint (2.5.9a), the inequality (2.5.9c) is more restricting compared to

(2.5.9d), so the above constraints reduce to

$$0 < \alpha \leq \lambda_Y, \tag{2.5.10a}$$

$$0 < \delta \leq (\alpha^{-1} + \gamma_Y^{-1} - \lambda_Y^{-1})^{-1}, \tag{2.5.10b}$$

$$\lambda_X^2 \lambda_Y^{-2} \alpha + \lambda_X - \lambda_X^2 \lambda_Y^{-1} + (L-1)(\gamma_X^2 \gamma_Y^{-2} \delta + \gamma_X - \gamma_X^2 \gamma_Y^{-1}) \leq LD. \tag{2.5.10c}$$

Then, the goal is to minimize $f(\alpha, \delta)$ subject to the constraints (2.5.10), which is a convex program. According to the KKT optimality conditions, there exist non-negative Lagrange multipliers $\{\omega_1, \omega_2, \omega_3\}$ and optimal solutions (α^*, δ^*) such that

$$\frac{\gamma_Y - \lambda_Y}{2((\lambda_Y - \gamma_Y)\alpha^* + \lambda_Y \gamma_Y)} + \omega_1 - \omega_2(1 + (\gamma_Y^{-1} - \lambda_Y^{-1})\alpha^*)^{-2} + \omega_3 \lambda_X^2 \lambda_Y^{-2} = 0, \tag{2.5.11a}$$

$$-\frac{L}{2\delta^*} + \omega_2 + (L-1)\omega_3 \gamma_X^2 \gamma_Y^{-2} = 0, \tag{2.5.11b}$$

$$\omega_1(\alpha^* - \lambda_Y) = 0, \tag{2.5.11c}$$

$$\omega_2(\delta^* - ((\alpha^*)^{-1} + \gamma_Y^{-1} - \lambda_Y^{-1})^{-1}) = 0, \tag{2.5.11d}$$

$$\omega_3(\lambda_X^2 \lambda_Y^{-2} \alpha^* + \lambda_X - \lambda_X^2 \lambda_Y^{-1} + (L-1)(\gamma_X^2 \gamma_Y^{-2} \delta^* + \gamma_X - \gamma_X^2 \gamma_Y^{-1}) - LD) = 0. \tag{2.5.11e}$$

In the following, we consider two different cases for the Lagrange multipliers.

Case 1 ($\omega_2 > 0$):

In this case, the KKT conditions in (2.5.11) reduce to

$$\frac{\gamma_Y - \lambda_Y}{2((\lambda_Y - \gamma_Y)\alpha^* + \lambda_Y\gamma_Y)} + \omega_1 - \omega_2(1 + (\gamma_Y^{-1} - \lambda_Y^{-1})\alpha^*)^{-2} + \omega_3\lambda_X^2\lambda_Y^{-2} = 0, \quad (2.5.12a)$$

$$-\frac{L}{2\delta^*} + \omega_2 + (L-1)\omega_3\gamma_X^2\gamma_Y^{-2} = 0, \quad (2.5.12b)$$

$$\omega_1(\alpha^* - \lambda_Y) = 0, \quad (2.5.12c)$$

$$\delta^* - ((\alpha^*)^{-1} + \gamma_Y^{-1} - \lambda_Y^{-1})^{-1} = 0, \quad (2.5.12d)$$

$$\omega_3(\lambda_X^2\lambda_Y^{-2}\alpha^* + \lambda_X - \lambda_X^2\lambda_Y^{-1} + (L-1)(\gamma_X^2\gamma_Y^{-2}\delta^* + \gamma_X - \gamma_X^2\gamma_Y^{-1}) - LD) = 0. \quad (2.5.12e)$$

Assume that α^* and δ^* satisfy

$$\lambda_X^2\lambda_Y^{-2}\alpha^* + \lambda_X - \lambda_X^2\lambda_Y^{-1} + (L-1)(\gamma_X^2\gamma_Y^{-2}\delta^* + \gamma_X - \gamma_X^2\gamma_Y^{-1}) = LD. \quad (2.5.13)$$

Solving the set of equations in (2.5.12) yields

$$\omega_1 = 0, \quad (2.5.14a)$$

$$\omega_2 = \frac{L}{2\delta^*} - (L-1)\omega_3\gamma_X^2\gamma_Y^{-2}, \quad (2.5.14b)$$

$$\omega_3 = \frac{\frac{L}{2\delta^*}(1 + (\gamma_Y^{-1} - \lambda_Y^{-1})\alpha^*)^{-2} + \frac{1}{2}(\gamma_Y^{-1} - \lambda_Y^{-1})(1 + (\gamma_Y^{-1} - \lambda_Y^{-1})\alpha^*)^{-1}}{\lambda_X^2\lambda_Y^{-2} + (L-1)\gamma_X^2\gamma_Y^{-2}(1 + (\gamma_Y^{-1} - \lambda_Y^{-1})\alpha^*)^{-2}}. \quad (2.5.14c)$$

Notice that $\omega_3 \geq 0$ since $\lambda_Y \geq \gamma_Y$. We should make sure that $\omega_2 \geq 0$. This gives the following inequality:

$$\frac{1}{2L}(\gamma_Y^{-1} - \lambda_Y^{-1})(1 + (\gamma_Y^{-1} - \lambda_Y^{-1})\alpha^*)^{-1} \leq \frac{1}{2(L-1)\delta^*}\lambda_X^2\gamma_X^{-2}\lambda_Y^{-2}\gamma_Y^2, \quad (2.5.15)$$

which can be equivalently written as

$$\delta^* \leq \frac{L}{L-1}\lambda_X^2\gamma_X^{-2}\lambda_Y^{-2}\gamma_Y^2\left(\left(\gamma_Y^{-1}-\lambda_Y^{-1}\right)^{-1}+\alpha^*\right), \tag{2.5.16}$$

combining the above inequality with (2.5.13), we can write

$$LD \leq L\lambda_X^2(\lambda_Y-\gamma_Y)^{-1}+(L-1)\gamma_X^2(\gamma_X^{-1}-\gamma_Y^{-1})+\lambda_X-(L+1)\lambda_X^2\lambda_Y^{-1}+(L+1)\alpha^*\lambda_X^2\lambda_Y^{-2}. \tag{2.5.17}$$

Define

$$\lambda_I := \alpha^*. \tag{2.5.18}$$

Considering (2.5.16) with (2.5.12d) and re-arranging the terms yields the following constraint:

$$(L-1)\gamma_X^2\gamma_Y^{-2}(\lambda_I^{-1}+\gamma_Y^{-1}-\lambda_Y^{-1})^{-1}-L\lambda_I\lambda_X^2\lambda_Y^{-2} \leq L\lambda_X^2\lambda_Y^{-1}(\lambda_Y\gamma_Y^{-1}-1)^{-1}. \tag{2.5.19}$$

Re-arranging the terms in (2.5.17) and (2.5.13), we have

$$LD \leq L\lambda_X^2\lambda_Y^{-1}(\lambda_Y\gamma_Y^{-1}-1)^{-1}+(L-1)\gamma_X^2(\gamma_X^{-1}-\gamma_Y^{-1})+\lambda_X-\lambda_X^2\lambda_Y^{-1}+(L+1)\lambda_I\lambda_X^2\lambda_Y^{-2}, \tag{2.5.20}$$

$$LD = \lambda_X-\lambda_X^2\lambda_Y^{-1}+(L-1)(\gamma_X^2\gamma_Y^{-2}(\lambda_I^{-1}+\gamma_Y^{-1}-\lambda_Y^{-1})^{-1}+\gamma_X-\gamma_X^2\gamma_Y^{-1})+\lambda_X^2\lambda_Y^{-2}\lambda_I. \tag{2.5.21}$$

Thus, we define the following set as the admissible distortion set:

$$\begin{aligned}\mathcal{D}_1(\lambda_I) :=& \{D \in (d_{\min}, \sigma_X^2): \\ & LD \leq L\lambda_X^2\lambda_Y^{-1}(\lambda_Y\gamma_Y^{-1}-1)^{-1} + (L-1)\gamma_X^2(\gamma_X^{-1}-\gamma_Y^{-1}) \\ & + \lambda_X - \lambda_X^2\lambda_Y^{-1} + (L+1)\lambda_I\lambda_X^2\lambda_Y^{-2}, \\ & LD = \lambda_X - \lambda_X^2\lambda_Y^{-1} + (L-1)(\gamma_X^2\gamma_Y^{-2}(\lambda_I^{-1}+\gamma_Y^{-1}-\lambda_Y^{-1})^{-1} \\ & + \gamma_X - \gamma_X^2\gamma_Y^{-1}) + \lambda_X^2\lambda_Y^{-2}\lambda_I\}.\end{aligned} \tag{2.5.22}$$

Plugging (2.5.12d) into (2.5.7) and considering (2.5.13) yields the rate-distortion expression $\overline{\mathcal{R}}(D)$ defined in (2.5.3) subject to constraint (2.5.4).

Case 2 ($\omega_2 = 0$):

In this case, the KKT conditions in (2.5.11) reduce to

$$\omega_1 = \frac{\lambda_Y - \gamma_Y}{2((\lambda_Y - \gamma_Y)\alpha^* + \lambda_Y\gamma_Y)} - \omega_3\lambda_X^2\lambda_Y^{-2}, \tag{2.5.23a}$$

$$\omega_3 = \frac{L}{2\delta^*(L-1)}\gamma_Y^2\gamma_X^{-2}, \tag{2.5.23b}$$

$$\omega_1(\alpha^* - \lambda_Y) = 0, \tag{2.5.23c}$$

$$\lambda_X^2\lambda_Y^{-2}\alpha^* + \lambda_X - \lambda_X^2\lambda_Y^{-1} + (L-1)(\gamma_X^2\gamma_Y^{-2}\delta^* + \gamma_X - \gamma_X^2\gamma_Y^{-1}) - LD = 0. \tag{2.5.23d}$$

To solve the above set of equations, we consider two different subcases: $\omega_1 > 0$ and $\omega_1 = 0$.

Subcase a ($\omega_1 = 0$):

Solving the set of equations in (2.5.23) with $\omega_1 = 0$ yields

$$\alpha^* = \frac{1}{2\omega_3}\lambda_X^{-2}\lambda_Y^2 - (\gamma_Y^{-1} - \lambda_Y^{-1})^{-1}, \tag{2.5.24a}$$

$$\delta^* = \frac{L}{2(L-1)\omega_3}\gamma_Y^2\gamma_X^{-2}, \tag{2.5.24b}$$

$$\omega_3 = \frac{L+1}{2} \cdot \frac{1}{LD - \lambda_X - (L-1)(\gamma_X - \gamma_X^2\gamma_Y^{-1}) + \lambda_X^2\lambda_Y^{-2}(\lambda_Y + (\gamma_Y^{-1} - \lambda_Y^{-1})^{-1})}. \tag{2.5.24c}$$

Recalling the definition of λ_I in (2.5.18), considering (2.5.24a) with (2.5.24c) and re-arranging the terms, we get the following equation:

$$LD = L\lambda_X^2\lambda_Y^{-1}(\lambda_Y\gamma_Y^{-1} - 1)^{-1} + (L-1)\gamma_X^2(\gamma_X^{-1} - \gamma_Y^{-1}) + \lambda_X - \lambda_X^2\lambda_Y^{-1} + (L+1)\lambda_I\lambda_X^2\lambda_Y^{-2}. \tag{2.5.25}$$

Notice that (2.5.11d) with $\omega_2 = 0$ implies that

$$\delta^* < ((\alpha^*)^{-1} + \gamma_Y^{-1} - \lambda_Y^{-1})^{-1}. \tag{2.5.26}$$

Moreover, (2.5.24a) with the fact that $\alpha^* < \lambda_Y$ gives

$$\omega_3 > \frac{1}{2}\lambda_X^{-2}(\lambda_Y - \gamma_Y), \tag{2.5.27}$$

which together with (2.5.24c) yields the following constraint on D:

$$LD < L\lambda_X^2(\lambda_Y - \gamma_Y)^{-1} + (L-1)\gamma_X^2(\gamma_X^{-1} - \gamma_Y^{-1}) + \lambda_X. \tag{2.5.28}$$

Plugging (2.5.24a) and (2.5.24b) into (2.5.26) and re-arranging the terms give the following condition:

$$L\lambda_X^2\lambda_Y^{-1}(\lambda_Y\gamma_Y^{-1}-1)^{-1} < (L-1)\gamma_X^2\gamma_Y^{-2}(\lambda_I^{-1}+\gamma_Y^{-1}-\lambda_Y^{-1})^{-1} - L\lambda_I\lambda_X^2\lambda_Y^{-2}, \quad (2.5.29)$$

combining (2.5.29) with (2.5.25) yields

$$LD < \lambda_X - \lambda_X^2\lambda_Y^{-1} + (L-1)(\gamma_X^2\gamma_Y^{-2}(\lambda_I^{-1}+\gamma_Y^{-1}-\lambda_Y^{-1})^{-1} + \gamma_X - \gamma_X^2\gamma_Y^{-1}) + \lambda_X^2\lambda_Y^{-2}\lambda_I. \quad (2.5.30)$$

The conditions (2.5.25) and (2.5.30) define the following distortion set:

$$\begin{aligned} \mathcal{D}_1^c(\lambda_I) := \{D \in (d_{\min}, \sigma_X^2) : \\ & LD = L\lambda_X^2\lambda_Y^{-1}(\lambda_Y\gamma_Y^{-1}-1)^{-1} + (L-1)\gamma_X^2(\gamma_X^{-1}-\gamma_Y^{-1}) \\ & + \lambda_X - \lambda_X^2\lambda_Y^{-1} + (L+1)\lambda_I\lambda_X^2\lambda_Y^{-2}, \\ & LD < \lambda_X - \lambda_X^2\lambda_Y^{-1} + (L-1)(\gamma_X^2\gamma_Y^{-2}(\lambda_I^{-1}+\gamma_Y^{-1}-\lambda_Y^{-1})^{-1} \\ & + \gamma_X - \gamma_X^2\gamma_Y^{-1}) + \lambda_X^2\lambda_Y^{-2}\lambda_I\}. \end{aligned} \quad (2.5.31)$$

In summary, for this subcase, $D \in \mathcal{D}_1^c(\lambda_I)$ while the constraint (2.5.28) holds. Plugging (2.5.24a)–(2.5.24c) into (2.5.7) gives the rate-distortion expression $\underline{\mathcal{R}}_1^c(D)$ defined in (2.4.10).

Subcase b ($\omega_1 > 0$):

Here, we get the following solution to (2.5.23):

$$\alpha^* = \lambda_Y, \tag{2.5.32a}$$

$$\delta^* = \frac{L}{2(L-1)\omega_3}\gamma_Y^2\gamma_X^{-2}, \tag{2.5.32b}$$

$$\omega_1 = \frac{\lambda_Y^{-1} - \gamma_Y\lambda_Y^{-2}}{2} - \omega_3\lambda_X^2\lambda_Y^{-2}, \tag{2.5.32c}$$

$$\omega_3 = \frac{L}{2(LD - \lambda_X - (L-1)(\gamma_X - \gamma_X^2\gamma_Y^{-1}))}. \tag{2.5.32d}$$

Considering the fact that $\omega_1 \geq 0$ yields the following constraint:

$$\omega_3 \leq \frac{1}{2}\lambda_X^{-2}(\lambda_Y - \gamma_Y). \tag{2.5.33}$$

Combining the above inequality with (2.5.32d), we get

$$LD \geq L\lambda_X^2(\lambda_Y - \gamma_Y)^{-1} + (L-1)\gamma_X^2(\gamma_X^{-1} - \gamma_Y^{-1}) + \lambda_X. \tag{2.5.34}$$

With a similar reason to the previous subcase (by considering distortion constraints (2.5.25) and (2.5.30)), we also know that $D \in \mathcal{D}_1^c(\lambda_I)$. In summary, for this subcase, the distortion set is restricted to $\mathcal{D}_1^c(\lambda_I)$ while constraint (2.5.34) holds. Plugging (2.5.32a) and (2.5.32b) into (2.5.7) while considering (2.5.32d) gives the rate-distortion expression $\underline{\mathcal{R}}_2^c(D)$ defined in (2.4.12).

To sum up all of the above cases, we have

$$\underline{\mathcal{R}}(D) = \begin{cases} \overline{\mathcal{R}}(D), & D \in \mathcal{D}_1(\lambda_I), \\ \underline{\mathcal{R}}^c(D), & D \in \mathcal{D}_1^c(\lambda_I), \end{cases} \tag{2.5.35}$$

where $\underline{\mathcal{R}}^c(D)$ is defined in (2.4.8).

Next, consider the case $\gamma_Y \geq \lambda_Y > 0$, where we have $\lambda_W = \lambda_Y$. Thus, the objective function (2.4.5) reduces to

$$f(\beta, \delta) := \frac{L-1}{2} \log \frac{\gamma_Y^2}{(\gamma_Y - \lambda_Y)\beta + \lambda_Y \gamma_Y} + \frac{L}{2} \log \frac{\lambda_Y}{\delta}, \tag{2.5.36}$$

subject to the following constraints:

$$0 < \beta \leq \gamma_Y, \tag{2.5.37a}$$

$$0 < \delta \leq (\beta^{-1} + \lambda_Y^{-1} - \gamma_Y^{-1})^{-1}, \tag{2.5.37b}$$

$$\lambda_X^2 \lambda_Y^{-2} \delta + \lambda_X - \lambda_X^2 \lambda_Y^{-1} + (L-1)(\gamma_X^2 \gamma_Y^{-2} \beta + \gamma_X - \gamma_X^2 \gamma_Y^{-1}) \leq LD. \tag{2.5.37c}$$

Then, the goal is to minimize $f(\beta, \delta)$ subject to the constraints (2.5.36).

According to the KKT optimality conditions, there exist nonnegative Lagrange multipliers $\{\hat{\omega_1}, \hat{\omega_2}, \hat{\omega_3}\}$ and optimal solutions (β^*, δ^*) such that

$$\frac{(L-1)(\lambda_Y - \gamma_Y)}{2((\gamma_Y - \lambda_Y)\beta^* + \lambda_Y \gamma_Y)} + \hat{\omega_1} - \hat{\omega_2}(1 + (\lambda_Y^{-1} - \gamma_Y^{-1})\beta^*)^{-2} + \hat{\omega_3}(L-1)\gamma_X^2 \gamma_Y^{-2} = 0, \tag{2.5.38a}$$

$$-\frac{L}{2\delta^*} + \hat{\omega_2} + \hat{\omega_3} \lambda_X^2 \lambda_Y^{-2} = 0, \tag{2.5.38b}$$

$$\hat{\omega_1}(\beta^* - \gamma_Y) = 0, \tag{2.5.38c}$$

$$\hat{\omega_2}(\delta^* - ((\beta^*)^{-1} + \lambda_Y^{-1} - \gamma_Y^{-1})^{-1}) = 0, \tag{2.5.38d}$$

$$\hat{\omega_3}(\lambda_X^2 \lambda_Y^{-2} \delta^* + \lambda_X - \lambda_X^2 \lambda_Y^{-1} + (L-1)(\gamma_X^2 \gamma_Y^{-2} \beta^* + \gamma_X - \gamma_X^2 \gamma_Y^{-1}) - LD) = 0. \tag{2.5.38e}$$

In the following, we also consider two different cases for the Lagrange multipliers.

Case 1 ($\hat{\omega_2} > 0$):

In this case, the KKT conditions in (2.5.38) reduce to the following:

$$\frac{(L-1)(\lambda_Y - \gamma_Y)}{2((\gamma_Y - \lambda_Y)\beta^* + \lambda_Y\gamma_Y)} + \hat{\omega_1} - \hat{\omega_2}(1 + (\lambda_Y^{-1} - \gamma_Y^{-1})\beta^*)^{-2} + \hat{\omega_3}(L-1)\gamma_X^2\gamma_Y^{-2} = 0, \tag{2.5.39a}$$

$$-\frac{L}{2\delta^*} + \hat{\omega_2} + \hat{\omega_3}\lambda_X^2\lambda_Y^{-2} = 0, \tag{2.5.39b}$$

$$\hat{\omega_1}(\beta^* - \gamma_Y) = 0, \tag{2.5.39c}$$

$$\delta^* - ((\beta^*)^{-1} + \lambda_Y^{-1} - \gamma_Y^{-1})^{-1} = 0, \tag{2.5.39d}$$

$$\hat{\omega_3}(\lambda_X^2\lambda_Y^{-2}\delta^* + \lambda_X - \lambda_X^2\lambda_Y^{-1} + (L-1)(\gamma_X^2\gamma_Y^{-2}\beta^* + \gamma_X - \gamma_X^2\gamma_Y^{-1}) - LD) = 0. \tag{2.5.39e}$$

Consider the case where β^* and δ^* satisfy the following:

$$\lambda_X^2\lambda_Y^{-2}\delta^* + \lambda_X - \lambda_X^2\lambda_Y^{-1} + (L-1)(\gamma_X^2\gamma_Y^{-2}\beta^* + \gamma_X - \gamma_X^2\gamma_Y^{-1}) = LD. \tag{2.5.40}$$

Solving the set of equations in (2.5.39) yields the following.

$$\hat{\omega_1} = 0, \tag{2.5.41a}$$

$$\hat{\omega_2} = \frac{L}{2\delta^*} - \hat{\omega_3}\lambda_X^2\lambda_Y^{-2}, \tag{2.5.41b}$$

$$\hat{\omega_3} = \frac{\frac{L}{2\delta^*}(1 + (\lambda_Y^{-1} - \gamma_Y^{-1})\beta^*)^{-2} + \frac{L-1}{2}(\lambda_Y^{-1} - \gamma_Y^{-1})(1 + (\lambda_Y^{-1} - \gamma_Y^{-1})\beta^*)^{-1}}{(L-1)\gamma_X^2\gamma_Y^{-2} + \lambda_X^2\lambda_Y^{-2}(1 + (\lambda_Y^{-1} - \gamma_Y^{-1})\beta^*)^{-2}}. \tag{2.5.41c}$$

Notice that $\hat{\omega_3} \geq 0$ since $\lambda_Y \leq \gamma_Y$. We should make sure that $\hat{\omega_2} \geq 0$. This gives the following inequality:

$$\frac{1}{2}(\lambda_Y^{-1} - \gamma_Y^{-1})(1 + (\lambda_Y^{-1} - \gamma_Y^{-1})\beta^*)^{-1} \leq \frac{L}{2\delta^*}\lambda_X^{-2}\gamma_X^2\lambda_Y^2\gamma_Y^{-2}, \tag{2.5.42}$$

which can be equivalently written as

$$\delta^* \leq L\lambda_X^{-2}\gamma_X^2\lambda_Y^2\gamma_Y^{-2}\left(\left(\lambda_Y^{-1} - \gamma_Y^{-1}\right)^{-1} + \beta^*\right) \tag{2.5.43}$$

Combining the above inequality with (2.5.40), we can write:

$$LD \leq L\gamma_X^2(\gamma_Y - \lambda_Y)^{-1} + (L-1)\gamma_X^2(\gamma_X^{-1} - \gamma_Y^{-1}) + \lambda_X - \lambda_X^2\lambda_Y^{-1} - L\gamma_X^2\gamma_Y^{-1} + (2L-1)\beta^*\gamma_X^2\gamma_Y^{-2}.$$

Case 2 ($\hat{\omega}_2 = 0$):

In this case, the KKT conditions in (2.5.38) reduce to the following:

$$\hat{\omega}_1 = \frac{(L-1)(\gamma_Y - \lambda_Y)}{2((\gamma_Y - \lambda_Y)\beta^* + \lambda_Y\gamma_Y)} - \hat{\omega}_3(L-1)\gamma_X^2\gamma_Y^{-2}, \tag{2.5.44a}$$

$$\hat{\omega}_3 = \frac{L}{2\delta^*}\lambda_Y^2\lambda_X^{-2}, \tag{2.5.44b}$$

$$\hat{\omega}_1(\beta^* - \gamma_Y) = 0, \tag{2.5.44c}$$

$$\lambda_X^2\lambda_Y^{-2}\delta^* + \lambda_X - \lambda_X^2\lambda_Y^{-1} + (L-1)(\gamma_X^2\gamma_Y^{-2}\beta^* + \gamma_X - \gamma_X^2\gamma_Y^{-1}) - LD = 0. \tag{2.5.44d}$$

To solve the above set of equations, we consider two different subcases $\hat{\omega}_1 > 0$ and $\hat{\omega}_1 = 0$ in the following.

Subcase a ($\hat{\omega}_1 > 0$):

Here, we get the following solution to (2.5.44):

$$\beta^* = \gamma_Y, \tag{2.5.45a}$$

$$\delta^* = \frac{L}{2\hat{\omega_3}}\lambda_Y^2\lambda_X^{-2}, \tag{2.5.45b}$$

$$\hat{\omega_1} = \frac{(L-1)(\gamma_Y^{-1} - \lambda_Y\gamma_Y^{-2})}{2} - \hat{\omega_3}(L-1)\gamma_X^2\gamma_Y^{-2}, \tag{2.5.45c}$$

$$\hat{\omega_3} = \frac{L}{2(LD - \lambda_X + \lambda_X^2\lambda_Y^{-1} - (L-1)\gamma_X)}. \tag{2.5.45d}$$

Consider the fact that $\hat{\omega_1} \geq 0$ which yields the following constraint

$$\hat{\omega_3} \leq \frac{1}{2}\gamma_X^{-2}(\gamma_Y - \lambda_Y). \tag{2.5.46}$$

Plugging (2.5.45a) and (2.5.45b) into (2.5.7) and considering inequality (2.5.46) yields the RHS of equation (2.4.29). Equality (2.5.45d) together with (2.5.46) yields the following:

$$L\gamma_X^2(\gamma_Y - \lambda_Y)^{-1} + (L-1)\gamma_X - \lambda_X^2\lambda_Y^{-1} + \lambda_X \leq LD. \tag{2.5.47}$$

Subcase b ($\hat{\omega_1} = 0$):

Solving the set of equations in (2.5.44) with $\hat{\omega_1} = 0$ yields the following:

$$\beta^* = \frac{1}{2\hat{\omega_3}}\gamma_X^{-2}\gamma_Y^2 - (\lambda_Y^{-1} - \gamma_Y^{-1})^{-1}, \tag{2.5.48a}$$

$$\delta^* = \frac{L}{2\hat{\omega_3}}\lambda_Y^2\lambda_X^{-2}, \tag{2.5.48b}$$

$$\hat{\omega_3} = \frac{2L-1}{2} \tag{2.5.48c}$$

$$\cdot \frac{1}{LD - \lambda_X - (L-1)(\gamma_X - \gamma_X^2\gamma_Y^{-1}) + \lambda_X^2\lambda_Y^{-1} + (L-1)\gamma_X^2\gamma_Y^{-2}(\lambda_Y^{-1} - \gamma_Y^{-1})^{-1}}. \tag{2.5.48d}$$

Considering the fact that $\gamma_Y \geq \beta^* > 0$, we get the following constraint:

$$\frac{1}{2}\gamma_X^{-2}\gamma_Y^2(\lambda_Y^{-1} - \gamma_Y^{-1}) \geq \hat{\omega}_3 \geq \frac{1}{2}\gamma_X^{-2}\gamma_Y^2((\lambda_Y^{-1} - \gamma_Y^{-1})^{-1} + \gamma_Y)^{-1}. \tag{2.5.49}$$

Plugging (2.5.48a)–(2.5.48d) into (2.5.36) yields the rate-distortion function in (2.4.31). The equality (2.5.48d) together with the condition (2.5.49) leads to the following constraint on distortion:

$$\begin{aligned} &L\gamma_X^2(\gamma_Y - \lambda_Y)^{-1} + (L-1)\gamma_X - \lambda_X^2\lambda_Y^{-1} + \lambda_X - (2L-1)\gamma_X^2\gamma_Y^{-1} \leq LD \leq \\ &L\gamma_X^2(\gamma_Y - \lambda_Y)^{-1} + (L-1)\gamma_X^2(\gamma_X^{-1} - \gamma_Y^{-1}) + \lambda_X - \lambda_X^2\lambda_Y^{-1} + (L-1)\gamma_X^2\gamma_Y^{-1}. \end{aligned} \tag{2.5.50}$$

the expression can be simplified as

$$\lambda_X^2\lambda_Y^{-2}(\gamma_I^{-1} + \lambda_Y^{-1} - \gamma_Y^{-1})^{-1} - L\gamma_I\gamma_X^2\gamma_Y^{-2} \leq L\gamma_X^2\gamma_Y^{-1}(\gamma_Y\lambda_Y^{-1} - 1)^{-1}, \tag{2.5.51}$$

the admissible distortion set is given by

$$\begin{aligned} \mathcal{D}_2(\gamma_I) := \{D \in (d_{\min}, \sigma_X^2)\colon& \\ & LD \leq L\gamma_X^2\gamma_Y^{-1}(\gamma_Y\lambda_Y^{-1} - 1)^{-1} + (L-1)\gamma_X^2(\gamma_X^{-1} - \gamma_Y^{-1}) \\ & + \lambda_X - \lambda_X^2\lambda_Y^{-1} + (2L-1)\gamma_I\gamma_X^2\gamma_Y^{-2}, \\ & LD = \lambda_X - \lambda_X^2\lambda_Y^{-1} + (L-1)(\gamma_X^2\gamma_Y^{-2}\gamma_I + \gamma_X - \gamma_X^2\gamma_Y^{-1}) \\ & + \lambda_X^2\lambda_Y^{-2}(\gamma_I^{-1} + \lambda_Y^{-1} - \gamma_Y^{-1})^{-1}\}, \end{aligned} \tag{2.5.52}$$

where we have $\underline{\mathcal{R}}(D) = \overline{\mathcal{R}}(D)$. Moreover, under the condition

$$\lambda_X^2\lambda_Y^{-2}(\gamma_I^{-1} + \lambda_Y^{-1} - \gamma_Y^{-1})^{-1} - L\gamma_I\gamma_X^2\gamma_Y^{-2} > L\gamma_X^2\gamma_Y^{-1}(\gamma_Y\lambda_Y^{-1} - 1)^{-1}, \tag{2.5.53}$$

the admissible distortion set is given by

$$\begin{aligned}\mathcal{D}_2^c(\gamma_I) := \{D \in (d_{\min}, \sigma_X^2)\colon \\ & LD = L\gamma_X^2\gamma_Y^{-1}(\gamma_Y\lambda_Y^{-1} - 1)^{-1} + (L-1)\gamma_X^2(\gamma_X^{-1} - \gamma_Y^{-1}) \\ & + \lambda_X - \lambda_X^2\lambda_Y^{-1} + (2L-1)\gamma_I\gamma_X^2\gamma_Y^{-2}, \\ & LD < \lambda_X - \lambda_X^2\lambda_Y^{-1} + (L-1)(\gamma_X^2\gamma_Y^{-2}\gamma_I + \gamma_X - \gamma_X^2\gamma_Y^{-1}) \\ & + \lambda_X^2\lambda_Y^{-2}(\gamma_I^{-1} + \lambda_Y^{-1} - \gamma_Y^{-1})^{-1}\},\end{aligned} \tag{2.5.54}$$

where the lower bound takes the expression $\underline{\mathcal{R}}^c(D)$ defined in (2.4.9). Thus, the case of $\gamma_Y \geq \lambda_Y$ can be summarized as follows:

$$\underline{\mathcal{R}}(D) = \begin{cases} \overline{\mathcal{R}}(D), & D \in \mathcal{D}_2(\gamma_I), \\ \hat{\underline{\mathcal{R}}}^c(D), & D \in \mathcal{D}_2^c(\gamma_I). \end{cases} \tag{2.5.55}$$

After characterizing the lower bound under two complement sets for each of the cases $\lambda_Y \geq \gamma_Y$ and $\gamma_Y \geq \lambda_Y$, it just remains to explicitly determine the sets $\mathcal{D}_1(\lambda_I)$ and $\mathcal{D}_2(\gamma_I)$. According to the definition of the set $\mathcal{D}_1(\lambda_I)$ in (2.5.22), the two conditions inside the set can be recast as the non-negativity condition for a certain expression. So, the proof is continued by investigating the sign of this expression over different intervals. Similar analyses can be done for $\mathcal{D}_2(\gamma_I)$ as well. Such steps have been taken in [33, Remark 3] and the result is summarized in the following.

$$\mathcal{D}_1(\lambda_I) = \begin{cases} \{D \in (d_{\min}, \sigma_X^2)\} & \text{if } \lambda_X^2\gamma_X^{-2}\gamma_Y^2\lambda_Y^{-2} \geq \frac{L-1}{4L}, \\ \{D \in (d_{\min}, \sigma_X^2)\} & \text{if } \lambda_X^2\gamma_X^{-2}\gamma_Y^2\lambda_Y^{-2} < \frac{L-1}{4L} \text{ and } \mu_2 \leq \frac{\gamma_Y}{\lambda_Y}, \\ \{D \in (d_{\min}, \mathsf{D}_{\text{th},1})\} & \text{if } \lambda_X^2\gamma_X^{-2}\gamma_Y^2\lambda_Y^{-2} < \frac{L-1}{4L}, \mu_1 \leq \frac{\gamma_Y}{\lambda_Y} \text{ and } \frac{\gamma_Y}{\lambda_Y} < \mu_2 < 1, \\ \{D \in (d_{\min}, \mathsf{D}_{\text{th},1}) \cup (\mathsf{D}_{\text{th},2}, \sigma_X^2)\} & \text{if } \lambda_X^2\gamma_X^{-2}\gamma_Y^2\lambda_Y^{-2} < \frac{L-1}{4L},\ \mu_1 > \frac{\gamma_Y}{\lambda_Y} \text{ and } \mu_2 < 1, \\ \emptyset & \text{if } \lambda_X^2\gamma_X^{-2}\gamma_Y^2\lambda_Y^{-2} < \frac{L-1}{4L},\ \mu_1 = 0 \text{ and } \mu_2 = 1, \end{cases} \tag{2.5.56}$$

and

$$\mathcal{D}_2(\gamma_I) = \begin{cases} \{D \in (d_{\min}, \sigma_X^2)\} & \text{if } \gamma_X^2\lambda_X^{-2}\lambda_Y^2\gamma_Y^{-2} \geq \frac{1}{4L}, \\ \{D \in (d_{\min}, \sigma_X^2)\} & \text{if } \gamma_X^2\lambda_X^{-2}\lambda_Y^2\gamma_Y^{-2} < \frac{1}{4L} \text{ and } \nu_2 \leq \frac{\lambda_Y}{\gamma_Y}, \\ \{D \in (d_{\min}, \hat{\mathsf{D}}_{\text{th},1})\} & \text{if } \gamma_X^2\lambda_X^{-2}\lambda_Y^2\gamma_Y^{-2} < \frac{1}{4L}, \nu_1 \leq \frac{\lambda_Y}{\gamma_Y} \text{ and } \frac{\lambda_Y}{\gamma_Y} < \nu_2 < 1, \\ \{D \in (d_{\min}, \hat{\mathsf{D}}_{\text{th},1}) \cup (\hat{\mathsf{D}}_{\text{th},2}, \sigma_X^2)\} & \text{if } \gamma_X^2\lambda_X^{-2}\lambda_Y^2\gamma_Y^{-2} < \frac{1}{4L}, \nu_1 > \frac{\lambda_Y}{\gamma_Y} \text{ and } \nu_2 < 1, \\ \emptyset & \text{if } \gamma_X^2\lambda_X^{-2}\lambda_Y^2\gamma_Y^{-2} < \frac{1}{4L}, \nu_1 = 0 \text{ and } \nu_2 = 1. \end{cases} \tag{2.5.57}$$

This completes the proof.

2.5.2 Proof of Theorem 4

First, notice that the distortion constraint in (2.4.3) can be written as

$$(\lambda_X + (L-1)\gamma_X - LD)\lambda_Q^2 + (\phi_1\gamma_Y + (L-1)\phi_2\lambda_Y - \phi_3(\gamma_Y + \lambda_Y))\lambda_Q - \phi_3\lambda_Y\gamma_Y = 0, \tag{2.5.58}$$

where $\phi_1 := \lambda_X^2 \lambda_Y^{-1}$, $\phi_2 := \gamma_X^2 \gamma_Y^{-1}$ and $\phi_3 := LD + \phi_1 + (L-1)\phi_2 - (\lambda_X + (L-1)\gamma_X)$. The equation in (2.5.58) can be equivalently written as

$$a\lambda_Q^2 + b\lambda_Q + c = 0, \tag{2.5.59}$$

where $a := (\sigma_X^2 - D)L$, $b := g_1 L^2 + g_2 L$ and $c := h_1 L^2 + h_2 L$ and

$$g_1 := \rho_X \rho_Z \sigma_X^2 \sigma_Z^2 + (\rho_X \sigma_X^2 + \rho_Z \sigma_Z^2)(\gamma_X - D), \tag{2.5.60a}$$

$$g_2 := \sigma_X^2(\gamma_Z + \gamma_Y) - \rho_X \sigma_X^2 \gamma_X - 2\gamma_Y D, \tag{2.5.60b}$$

$$\begin{aligned} h_1 :=& \rho_X \rho_Z \sigma_X^2 \sigma_Z^2 \gamma_Y + (\rho_X \sigma_X^2 + \rho_Z \sigma_Z^2)(\gamma_X \gamma_Z - \gamma_Y D) \\ =& \gamma_Y (\rho_X \sigma_X^2 + \rho_Z \sigma_Z^2)(d_{\min}^{\infty} - D), \end{aligned} \tag{2.5.60c}$$

$$h_2 := \rho_X \sigma_X^2 \gamma_Z^2 + \rho_Z \sigma_Z^2 \gamma_X^2 + \gamma_X \gamma_Z \gamma_Y - \gamma_Y^2 D. \tag{2.5.60d}$$

We consider three different cases based on the value of g_1.

Case1 ($g_1 > 0$): In this case, we have

$$\lambda_Q = \frac{-b+\sqrt{b^2-4ac}}{2a} \tag{2.5.61a}$$

$$= \frac{-b+b\sqrt{1-\frac{4ac}{b^2}}}{2a} \tag{2.5.61b}$$

$$= \frac{-b+b(1-\frac{2ac}{b^2}-\frac{2a^2c^2}{b^4}+O(\frac{1}{L^3}))}{2a} \tag{2.5.61c}$$

$$= -\frac{c}{b}-\frac{ac^2}{b^3}+O\left(\frac{1}{L^2}\right) \tag{2.5.61d}$$

$$= -\frac{h_1L+h_2}{g_1L+g_2}-\frac{(\sigma_X^2-D)(h_1L+h_2)^2}{(g_1L+g_2)^3}+O\left(\frac{1}{L^2}\right) \tag{2.5.61e}$$

$$= -\frac{h_1L+h_2}{g_1L}\left(1-\frac{g_2}{g_1L}+O(\frac{1}{L^2})\right)-\frac{(\sigma_X^2-D)h_1^2}{g_1^3L}+O\left(\frac{1}{L^2}\right) \tag{2.5.61f}$$

$$= -\frac{h_1}{g_1}-\left(\frac{h_2}{g_1}-\frac{g_2h_1}{g_1^2}+\frac{(\sigma_X^2-D)h_1^2}{g_1^3}\right)\frac{1}{L}+O\left(\frac{1}{L^2}\right) \tag{2.5.61g}$$

$$= \eta_1+\frac{\eta_2}{L}+O\left(\frac{1}{L^2}\right), \tag{2.5.61h}$$

where (2.5.61c) follows because $\sqrt{1+x} = 1+\frac{1}{2}x-\frac{1}{8}x^2+O(x^3)$ and (2.5.61f) follows because $\frac{1}{1+x} = 1 - x + O(x^2)$. Now, plugging the above into (2.4.2) yields

$$
\begin{aligned}
&\frac{1}{2}\log\frac{\lambda_Y+\lambda_Q}{\lambda_Q}+\frac{L-1}{2}\log\frac{\gamma_Y+\lambda_Q}{\lambda_Q} && \text{(2.5.62a)}\\
=&\frac{1}{2}\log\frac{\lambda_Y+\eta_1+\frac{\eta_2}{L}+O(\frac{1}{L^2})}{\eta_1+\frac{\eta_2}{L}+O(\frac{1}{L^2})}\\
&+\frac{L-1}{2}\log\frac{\gamma_Y+\eta_1+\frac{\eta_2}{L}+O(\frac{1}{L^2})}{\eta_1+\frac{\eta_2}{L}+O(\frac{1}{L^2})} && \text{(2.5.62b)}\\
=&\frac{1}{2}\log\frac{(1+(L-1)\rho_Y)\sigma_Y^2+\eta_1+\frac{\eta_2}{L}+O(\frac{1}{L^2})}{\eta_1+\frac{\eta_2}{L}+O(\frac{1}{L^2})}\\
&+\frac{L-1}{2}\log\frac{\gamma_Y+\eta_1+\frac{\eta_2}{L}+O(\frac{1}{L^2})}{\eta_1+\frac{\eta_2}{L}+O(\frac{1}{L^2})} && \text{(2.5.62c)}\\
=&\frac{1}{2}\log\left(\frac{L\rho_Y\sigma_Y^2}{\eta_1}+O(1)\right)\\
&+\frac{L-1}{2}\log\left(\left(\frac{\gamma_Y+\eta_1}{\eta_1}+\frac{\eta_2}{L\eta_1}+O\left(\frac{1}{L^2}\right)\right)\left(1-\frac{\eta_2}{L\eta_1}+O\left(\frac{1}{L^2}\right)\right)\right) && \text{(2.5.62d)}\\
=&\frac{1}{2}\log\left(\frac{L\rho_Y\sigma_Y^2}{\eta_1}+O(1)\right)+\frac{L-1}{2}\log\left(\frac{\gamma_Y+\eta_1}{\eta_1}-\frac{\eta_2\gamma_Y}{L\eta_1^2}+O\left(\frac{1}{L^2}\right)\right) && \text{(2.5.62e)}\\
=&\frac{1}{2}\log L+\frac{1}{2}\log\frac{\rho_Y\sigma_Y^2}{\eta_1+\gamma_Y}+\frac{L}{2}\log\frac{\eta_1+\gamma_Y}{\eta_1}-\frac{\eta_2\gamma_Y}{2\eta_1(\eta_1+\gamma_Y)}+O\left(\frac{1}{L}\right) && \text{(2.5.62f)}\\
=&\frac{1}{2}\log L+\frac{1}{2}\log\frac{\rho_X\sigma_X^2+\rho_Z\sigma_Z^2}{\eta_1+\gamma_Y}+\frac{L}{2}\log\frac{\eta_1+\gamma_Y}{\eta_1}-\frac{\eta_2\gamma_Y}{2\eta_1(\eta_1+\gamma_Y)}+O\left(\frac{1}{L}\right), && \text{(2.5.62g)}
\end{aligned}
$$

where (2.5.62d) follows because $\frac{1}{1+x} = 1 - x + O(x^2)$ and (2.5.62f) follows because $\log(1 + x) = x + O(x^2)$. With some straightforward calculations, we can show that each term of the above expression can be written as follows:

$$\frac{1}{2}\log\frac{\rho_X\sigma_X^2+\rho_Z\sigma_Z^2}{\eta_1+\gamma_Y}=\frac{1}{2}\log\frac{\rho_X\rho_Z\sigma_X^2\sigma_Z^2+(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)(\gamma_X-D)}{\gamma_X^2}, \tag{2.5.63a}$$

$$\frac{L}{2}\log\frac{\eta_1+\gamma_Y}{\eta_1}=\frac{L}{2}\log\frac{(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)\gamma_X^2}{(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)(\gamma_Y D-\gamma_X\gamma_Z)-\rho_X\rho_Z\gamma_Y\sigma_X^2\sigma_Z^2}, \tag{2.5.63b}$$

$$\begin{aligned}-\frac{\gamma_Y\eta_2}{2\eta_1(\eta_1+\gamma_Y)}&=\frac{\gamma_Y(\sigma_X^2\rho_Z\sigma_Z^2-(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)D)^2}{2(\rho_X\rho_Z\sigma_X^2\sigma_Z^2+(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)(\gamma_X-D))(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)\gamma_Y(D-d_{\min}^\infty)}\\ &=\frac{\gamma_Y(\sigma_X^2\rho_Z\sigma_Z^2-(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)D)^2}{2(\rho_X\rho_Z\sigma_X^2\sigma_Z^2+(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)(\gamma_X-D))((\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)(\gamma_Y D-\gamma_X\gamma_Z)-\rho_X\rho_Z\gamma_Y\sigma_X^2\sigma_Z^2)}.\end{aligned} \tag{2.5.63c}$$

Moreover, notice that $g_1 > 0$ and $D > d_{\min}^\infty$ implies $\rho_X\sigma_X^2+\rho_Z\sigma_Z^2 > 0$ and $\eta_1 > 0$ from (2.5.60b) and (2.5.60d). Considering these conditions, (2.5.62g)–(2.5.63) and simplifying the terms, we get the first clause of (2.4.40).

Case 2 ($g_1 = 0$): We consider two different subcases.

Subcase 1 ($\rho_X\sigma_X^2+\rho_Z\sigma_Z^2=0$): The distortion constraint in (2.4.3) simplifies to

$$L\sigma_X^2-\frac{L\sigma_X^4}{\sigma_X^2+\sigma_Z^2+\lambda_Q}=LD, \tag{2.5.64}$$

or equivalently,

$$\lambda_Q=\frac{\sigma_X^4}{\sigma_X^2-D}-\sigma_X^2-\sigma_Z^2. \tag{2.5.65}$$

Plugging the above solution in (2.4.2), we get the rate-distortion expression in (2.4.39).

Subcase 2 ($\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2 > 0$): In this case, we have

$$\lambda_Q = \frac{-g_2L + g_2L\sqrt{1 - 4\frac{(\sigma_X^2-D)(h_1L+h_2)}{g_2^2}}}{2L(\sigma_X^2 - D)} \tag{2.5.66a}$$

$$= \frac{-g_2L + L^{\frac{3}{2}}\sqrt{-4(\sigma_X^2 - D)h_1}\sqrt{1 + \frac{g_2^2-4(\sigma_X^2-D)h_2}{-4(\sigma_X^2-D)h_1L}}}{2L(\sigma_X^2 - D)} \tag{2.5.66b}$$

$$= \frac{-g_2L + L^{\frac{3}{2}}\sqrt{-4(\sigma_X^2 - D)h_1}(1 - \frac{g_2^2-4(\sigma_X^2-D)h_2}{8(\sigma_X^2-D)h_1L} + O(\frac{1}{L^2}))}{2L(\sigma_X^2 - D)} \tag{2.5.66c}$$

$$= \sqrt{-\frac{h_1L}{\sigma_X^2 - D}} - \frac{g_2}{2(\sigma_X^2 - D)} + \frac{g_2^2 - 4(\sigma_X^2 - D)h_2}{8\sqrt{-(\sigma_X^2 - D)^3h_1L}} + O(\frac{1}{L^{\frac{3}{2}}}) \tag{2.5.66d}$$

$$= \alpha_1\sqrt{L} + \alpha_2 + O(\frac{1}{\sqrt{L}}). \tag{2.5.66e}$$

Moreover, the condition $\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2 > 0$ together with $\sigma_X^2 > D > d_{\min}^{\infty}$ and $g_1 = 0$ implies $\rho_X > 0$, $\gamma_X > 0$ and $\alpha_1 > 0$. Then, we get the following:

$$D = \frac{\rho_X\rho_Z\sigma_X^2\sigma_Z^2}{\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2} + \gamma_Y, \tag{2.5.67a}$$

$$\sigma_X^2 - D = \frac{\rho_X^2\sigma_X^4}{\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2}, \tag{2.5.67b}$$

$$h_1 = -(\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2)\gamma_X^2, \tag{2.5.67c}$$

$$g_2 = \frac{-\rho_Z\gamma_X\sigma_X^2\sigma_Z^2 - \rho_X\gamma_X\gamma_Y\sigma_X^2 + \rho_X^2\sigma_X^4\gamma_Z + \rho_X\gamma_Z\sigma_X^4 - \rho_X\rho_Z\gamma_X\sigma_X^2\sigma_Z^2}{\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2}, \tag{2.5.67d}$$

$$\alpha_1 = \frac{(\rho_X\sigma_X^2 + \rho_X\sigma_Z^2)\gamma_X}{\rho_X\sigma_X^2}, , \tag{2.5.67e}$$

$$\alpha_2 = \frac{\rho_Z\gamma_X\sigma_X^2\sigma_Z^2 + \rho_X\gamma_X\gamma_Y\sigma_X^2 - \rho_X^2\sigma_X^4\gamma_Z - \rho_X\gamma_Z\sigma_X^4 + \rho_X\rho_Z\gamma_X\sigma_X^2\sigma_Z^2}{2\rho_X^2\sigma_X^4}. \tag{2.5.67f}$$

Now, we simplify each term of the rate in (2.4.2). Consider the first term of (2.4.2) as follows:

$$\begin{aligned}
\frac{1}{2}\log\frac{\lambda_Y+\lambda_Q}{\lambda_Q} &= \frac{1}{2}\log\frac{\lambda_Y}{\lambda_Q}+\frac{1}{2}\log\frac{\lambda_Y+\lambda_Q}{\lambda_Y} &\text{(2.5.68a)}\\
&= \frac{1}{2}\log\frac{L(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)+\gamma_Y}{\alpha_1\sqrt{L}+O(1)}+\frac{1}{2}\log\frac{\lambda_Y+\lambda_Q}{\lambda_Y} &\text{(2.5.68b)}\\
&= \frac{1}{2}\log\frac{L(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)+\gamma_Y}{\alpha_1\sqrt{L}+O(1)}+O\left(\frac{1}{\sqrt{L}}\right) &\text{(2.5.68c)}\\
&= \frac{1}{4}\log L+\frac{1}{2}\log\frac{\rho_X\sigma_X^2+\rho_Z\sigma_Z^2}{\alpha_1}+O\left(\frac{1}{\sqrt{L}}\right) &\text{(2.5.68d)}\\
&= \frac{1}{4}\log L+\frac{1}{2}\log\rho_X\gamma_X^{-1}\sigma_X^2+O\left(\frac{1}{\sqrt{L}}\right), &\text{(2.5.68e)}
\end{aligned}$$

where (2.5.68b) follows from the definition of λ_Y in (2.3.9) and the definition of λ_Q in (2.5.66e), (2.5.68c) follows because $\frac{\lambda_Q}{\lambda_Y}=O(\frac{1}{\sqrt{L}})$ and $\log(1+x)=O(x)$, (2.5.68e) follows from the definition of α_1 in (2.5.67e).

The second term of (2.4.2) can be simplified as follows:

$$
\begin{aligned}
&\frac{L-1}{2}\log\frac{\gamma_Y+\lambda_Q}{\lambda_Q} && \text{(2.5.69a)}\\
=&\frac{L-1}{2}\log(1+\frac{\gamma_Y}{\lambda_Q}) && \text{(2.5.69b)}\\
=&\frac{L-1}{2}\left(\frac{\gamma_Y}{\lambda_Q}-\frac{\gamma_Y^2}{2\lambda_Q^2}+O\left(\frac{1}{L^{\frac{3}{2}}}\right)\right) && \text{(2.5.69c)}\\
=&\frac{L-1}{2}\left(\frac{\gamma_Y}{\alpha_1\sqrt{L}+\alpha_2+O(\frac{1}{\sqrt{L}})}-\frac{\gamma_Y^2}{2(\alpha_1\sqrt{L}+O(1))^2}+O\left(\frac{1}{L^{\frac{3}{2}}}\right)\right) && \text{(2.5.69d)}\\
=&\frac{L-1}{2}\Big(\frac{\gamma_Y}{\alpha_1\sqrt{L}}\left(1-\frac{\alpha_2}{\alpha_1\sqrt{L}}+O\left(\frac{1}{L}\right)\right)-\frac{\gamma_Y^2}{2\alpha_1^2L}\left(1+O\left(\frac{1}{\sqrt{L}}\right)\right)+O\left(\frac{1}{L^{\frac{3}{2}}}\right)\Big) && \text{(2.5.69e)}\\
=&\frac{\gamma_Y\sqrt{L}}{2\alpha_1}-\frac{\gamma_Y(\gamma_Y+2\alpha_2)}{4\alpha_1^2}+O\left(\frac{1}{\sqrt{L}}\right) && \text{(2.5.69f)}\\
=&\frac{\rho_X\gamma_Y\sigma_X^2\sqrt{L}}{2\gamma_X(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)}-\frac{\gamma_Y(\rho_X\sigma_X^4(\gamma_X-\rho_X\gamma_Z)+(1+\rho_X)\rho_Z\sigma_X^2\sigma_Z^2\gamma_X)}{4(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)^2\gamma_X^2}\\
&+O\left(\frac{1}{\sqrt{L}}\right), && \text{(2.5.69g)}
\end{aligned}
$$

where (2.5.69c) follows because $\frac{\gamma_Y}{\lambda_Q}=O\left(\frac{1}{\sqrt{L}}\right)$ and $\log(1+x)=x-\frac{1}{2}x^2+O(x^3)$, (2.5.69e) follows because $\frac{1}{1+x}=1-x+O(x^2)$. Considering the fact that $g_1=0$, using approximations (2.5.68e) and (2.5.69g) and simplifying the terms, we get the second clause of (2.4.40).

Case 3 ($g_1 < 0$): Here, we have

$$\lambda_Q = \frac{-g_1L^2 - g_2L + \sqrt{(g_1L^2 + g_2L)^2 - 4L(\sigma_X^2 - D)(h_1L^2 + h_2L)}}{2L(\sigma_X^2 - D)} \tag{2.5.70a}$$

$$= \frac{-g_1L^2 - g_2L + \sqrt{g_1^2L^4 + (2g_1g_2 - 4(\sigma_X^2 - D)h_1)L^3 + (g_2^2 - 4(\sigma_X^2 - D)h_2)L^2}}{2L(\sigma_X^2 - D)} \tag{2.5.70b}$$

$$= \frac{-g_1L^2 - g_2L - g_1L^2(1 + (g_1g_2 - 2(\sigma_X^2 - D)h_1)\frac{1}{g_1^2L} + O(\frac{1}{L^2}))}{2L(\sigma_X^2 - D)} \tag{2.5.70c}$$

$$= -\frac{g_1}{\sigma_X^2 - D}L - \frac{g_2 - (\sigma_X^2 - D)h_1}{\sigma_X^2 - D} + O\left(\frac{1}{L}\right) \tag{2.5.70d}$$

$$= \frac{(\gamma_X - D)(\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2) - \rho_X\rho_Z\sigma_X^2\sigma_Z^2}{\sigma_X^2 - D} + O(1) \tag{2.5.70e}$$

$$= \beta_1 L + O(1), \tag{2.5.70f}$$

where (2.5.70c) follows because $\sqrt{1+x} = 1 + \frac{1}{2}x + O(x^2)$. We then use the above approximation to calculate each term of the rate in (2.4.2) as follows:

$$\frac{1}{2}\log\frac{\lambda_Y + \lambda_Q}{\lambda_Q} = \frac{1}{2}\log\frac{\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2 + \beta_1}{\beta_1} + O\left(\frac{1}{L}\right) \tag{2.5.71a}$$

$$= \frac{1}{2}\log\frac{\rho_X^2\sigma_X^4}{(\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2)(D - \gamma_X) - \rho_X\rho_Z\sigma_X^2\sigma_Z^2} + O\left(\frac{1}{L}\right), \tag{2.5.71b}$$

and

$$\frac{L-1}{2}\log\frac{\gamma_Y+\lambda_Q}{\lambda_Q} = \frac{L-1}{2}\left(\frac{\gamma_Y}{\lambda_Q}+O(\frac{1}{L^2})\right) \tag{2.5.72a}$$

$$= \frac{\gamma_Y}{2\beta_1}+O\left(\frac{1}{L}\right) \tag{2.5.72b}$$

$$= \frac{\gamma_Y(\sigma_X^2-D)}{2(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)(D-\gamma_X)-2\rho_X\rho_Z\sigma_X^2\sigma_Z^2}+O\left(\frac{1}{L}\right). \tag{2.5.72c}$$

Considering the fact that $g_1 < 0$, using approximations (2.5.71b) and (2.5.72c) and simplifying the terms, we get the third clause of (2.4.40). This concludes the proof.

2.5.3 Proof of Theorem 5

First, notice that $\rho_X, \rho_Z \in [0, 1]$ implies $\lambda_Y \geq \gamma_Y$. We consider four different cases.

Case 1 ($\rho_X\sigma_X^2+\rho_Z\sigma_Z^2 = 0$): In this case, the condition $\lambda_X^2\gamma_Y^2 \geq \frac{L-1}{4L}\gamma_X^2\lambda_Y^2$ is satisfied trivially for all L. So, we are under the first condition of Theorem 3, and consequently

$$\underline{\mathcal{R}}(D) = \overline{\mathcal{R}}(D) = \overline{\mathcal{R}}^{\infty}(D). \tag{2.5.73}$$

This yields the first condition of Theorem (5), where the rate-distortion expression is given by (2.4.45).

Case 2 ($\rho_X\sigma_X^2+\rho_Z\sigma_Z^2 > 0$**,** $\rho_X > 0$**,** $\xi \geq \frac{1}{2}$): In this case, we are under the first condition of Theorem 3. This can be readily verified when $\gamma_X = 0$. When $\gamma_X > 0$, we have

$$\lambda_X^2\lambda_Y^{-2}\gamma_X^{-2}\gamma_Y^2 = \frac{(1+(L-1)\rho_X)^2(1-\rho_Y)^2}{(1+(L-1)\rho_Y)^2(1-\rho_X)^2} \tag{2.5.74a}$$

$$= \xi^2 + \frac{2\xi^2(\rho_Y-\rho_X)}{\rho_X\rho_Y L} + O\left(\frac{1}{L^2}\right) \tag{2.5.74b}$$

$$\geq \frac{1}{4} \quad \text{for all sufficiently large } L \tag{2.5.74c}$$

$$\geq \frac{L-1}{4L}, \tag{2.5.74d}$$

where (2.5.74c) can be verified by considering $\xi = \frac{1}{2}$ (which implies $\rho_Y > \rho_X$) and $\xi > \frac{1}{2}$ separately. In summary, the analysis of this case yields (2.4.46).

Case 3 ($\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2 > 0$, $\rho_X > 0$, $\xi < \frac{1}{2}$): In this case, we are under the third condition of Theorem 3. This is because of the fact that $\mu_2 < 1$,

$$\lambda_X^2\lambda_Y^{-2}\gamma_X^{-2}\gamma_Y^2 = \xi^2 + O\left(\frac{1}{L}\right) \tag{2.5.75a}$$

$$< \frac{L-1}{4L} \quad \text{for all sufficiently large } L, \tag{2.5.75b}$$

and

$$\mu_1 = \frac{1}{2} - \frac{1}{2}\sqrt{1 - \frac{4L}{L-1}\lambda_X^2\lambda_Y^{-2}\gamma_X^{-2}\gamma_Y^2} \tag{2.5.76a}$$

$$= \frac{1}{2} - \frac{1}{2}\sqrt{1-4\xi^2} + O\left(\frac{1}{L}\right) \tag{2.5.76b}$$

$$> \frac{\gamma_Y}{\lambda_Y} \quad \text{for all sufficiently large } L, \tag{2.5.76c}$$

where the last inequality follows because $\frac{\gamma_Y}{\lambda_Y} = O\left(\frac{1}{L}\right)$. Thus, we continue with approximating $\mathsf{D}_{\text{th},1}$, $\mathsf{D}_{\text{th},2}$ and the rate-distortion expressions. We approximate $\mathsf{D}_{\text{th},1}$

and $\mathsf{D}_{\text{th},2}$ for large L as follows:

$$\mathsf{D}_{\text{th},1} = \frac{\rho_X \rho_Z \sigma_X^2 \sigma_Z^2}{\rho_X \sigma_X^2 + \rho_Z \sigma_Z^2} + \gamma_X - \frac{1 + \sqrt{1 - 4\xi^2}}{2} \gamma_X^2 \gamma_Y^{-1} + O\left(\frac{1}{L}\right) \tag{2.5.77a}$$

$$= \mathsf{D}_{\text{th},1}^{\infty} + O\left(\frac{1}{L}\right), \tag{2.5.77b}$$

and

$$\mathsf{D}_{\text{th},2} = \frac{\rho_X \rho_Z \sigma_X^2 \sigma_Z^2}{\rho_X \sigma_X^2 + \rho_Z \sigma_Z^2} + \gamma_X - \frac{1 - \sqrt{1 - 4\xi^2}}{2} \gamma_X^2 \gamma_Y^{-1} + O\left(\frac{1}{L}\right) \tag{2.5.78a}$$

$$= \mathsf{D}_{\text{th},2}^{\infty} + O\left(\frac{1}{L}\right). \tag{2.5.78b}$$

Now, it remains to approximate the rate-distortion expressions. In the intervals $D < \mathsf{D}_{\text{th},1}^{\infty}$ and $D > \mathsf{D}_{\text{th},2}^{\infty}$, $\overline{\mathcal{R}}(D)$ can be approximated as in Theorem 4, which leads to the expression in (2.4.40). In the interval $\mathsf{D}_{\text{th},1}^{\infty} < D < \mathsf{D}_{\text{th},2}^{\infty}$, we need to approximate $\underline{\mathcal{R}}^c(D)$. For the rate-distortion expression $\underline{\mathcal{R}}^c(D)$, notice that the second clause of (2.4.8) is not active for large L since

$$L\lambda_X^2(\lambda_Y - \gamma_Y)^{-1} + (L-1)\gamma_X^2(\gamma_X^{-1} - \gamma_Y^{-1}) + \lambda_X = L^2 \rho_X^2 \rho_Y^{-1} \sigma_X^4 \sigma_Y^{-2} + O(L) > LD. \tag{2.5.79}$$

Thus, we need to approximate $\underline{\mathcal{R}}_1^c(D)$ defined in (2.4.10) for large L. Consider the following term in the first logarithm. We have

$$
\begin{aligned}
&LD - \lambda_X - (L-1)(\gamma_X - \gamma_X^2\gamma_Y^{-1}) + \lambda_X^2\lambda_Y^{-2}(\lambda_Y + (\gamma_Y^{-1} - \lambda_Y^{-1})^{-1}) \\
=&L(D - \rho_X\sigma_X^2 - (\gamma_X - \gamma_X^2\gamma_Y^{-1}) + \rho_X^2\rho_Y^{-1}\sigma_X^4\sigma_Y^{-2}) + \\
&(2\rho_X\rho_Y^{-1}(1-\rho_X)\sigma_X^4\sigma_Y^{-2} - \gamma_X^2\gamma_Y^{-1}) + O\left(\frac{1}{L}\right) \\
=&LA + B + O\left(\frac{1}{L}\right). \qquad (2.5.80a)
\end{aligned}
$$

Thus, plugging the above into $\underline{\mathcal{R}}_1^c(D)$ in (2.4.10), we can approximate the first logarithm as follows:

$$
\begin{aligned}
&\frac{L+1}{2}\log\frac{(L+1)\gamma_Y^{-1}\gamma_X^2}{LA + B + O\left(\frac{1}{L}\right)} \\
&=\frac{L+1}{2}\log\frac{\gamma_Y^{-1}\gamma_X^2}{A + \frac{1}{L+1}(-A+B)} \qquad (2.5.81a) \\
&=\frac{L+1}{2}\log\frac{\gamma_Y^{-1}\gamma_X^2}{A} + \frac{A-B}{2A} + O\left(\frac{1}{L}\right) \qquad (2.5.81b) \\
&=\frac{L+1}{2}\log\frac{\gamma_Y^{-1}\gamma_X^2}{D - \rho_X\sigma_X^2 - (\gamma_X - \gamma_X^2\gamma_Y^{-1}) + \rho_X^2\rho_Y^{-1}\sigma_X^4\sigma_Y^{-2}} \\
&\quad + \frac{1}{2}\frac{D + 2\gamma_X^2\gamma_Y^{-1} - \sigma_X^2 + \rho_X(3\rho_X - 2)\rho_Y^{-1}\sigma_X^4\sigma_Y^{-2}}{D - \rho_X\sigma_X^2 - (\gamma_X - \gamma_X^2\gamma_Y^{-1}) + \rho_X^2\rho_Y^{-1}\sigma_X^4\sigma_Y^{-2}} + O\left(\frac{1}{L}\right) \qquad (2.5.81c) \\
&=\frac{L+1}{2}\log\frac{\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2\gamma_Y^{-1}\gamma_X^2}{(\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2)(D - (\gamma_X - \gamma_X^2\gamma_Y^{-1})) - \rho_X\rho_Z\sigma_X^2\sigma_Z^2} \\
&\quad + \frac{1}{2}\frac{(\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2)(D + 2(1-\xi)\gamma_X^2\gamma_Y^{-1} - \gamma_X) - \rho_X\rho_Z\sigma_X^2\sigma_Z^2}{(\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2)(D - (\gamma_X - \gamma_X^2\gamma_Y^{-1})) - \rho_X\rho_Z\sigma_X^2\sigma_Z^2} + O\left(\frac{1}{L}\right). \\
&\qquad (2.5.81d)
\end{aligned}
$$

The second logarithm of (2.4.10) can be approximated as follows:

$$\frac{1}{2}\log \lambda_X^2 \gamma_X^{-2}(\lambda_Y \gamma_Y^{-1} - 1)^{-1} = \frac{1}{2}\log L + \frac{1}{2}\log\left(\left(\frac{\rho_X}{1-\rho_X}\right)^2 \left(\frac{1-\rho_Y}{\rho_Y}\right)\right) + O\left(\frac{1}{L}\right). \tag{2.5.82}$$

The third logarithm of (2.4.10) can also be approximated as follows:

$$\frac{L}{2}\log\left(1 - \frac{1}{L}\right) = -\frac{1}{2} + O\left(\frac{1}{L^2}\right). \tag{2.5.83}$$

Plugging (2.5.81d) and (2.5.82) into (2.4.10) yields

$$\begin{aligned}\underline{\mathcal{R}}_1^c(D) =& \frac{L+1}{2}\log\frac{\gamma_Y^{-1}\gamma_X^2}{(\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2)(D - (\gamma_X - \gamma_X^2\gamma_Y^{-1})) - \rho_X\rho_Z\sigma_X^2\sigma_Z^2} + \frac{1}{2}\log L \\ &+ \frac{1}{2}\frac{(1-2\xi)\gamma_X^2\gamma_Y^{-1}}{(\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2)(D - (\gamma_X - \gamma_X^2\gamma_Y^{-1})) - \rho_X\rho_Z\sigma_X^2\sigma_Z^2} \\ &+ \frac{1}{2}\log\left(\left(\frac{\rho_X}{1-\rho_X}\right)^2\left(\frac{1-\rho_Y}{\rho_Y}\right)\right) + O\left(\frac{1}{L}\right) \\ =& \underline{\mathcal{R}}_1^\infty(D). \end{aligned} \tag{2.5.84a}$$

The above expression can be further simplified to (2.4.43). Moreover, the two boundary points $D = \mathsf{D}_{\text{th},1}^\infty$ and $D = \mathsf{D}_{\text{th},2}^\infty$ can be easily handled by considering the fact that $\overline{\mathcal{R}}_1^\infty(\mathsf{D}_{\text{th},1}^\infty) = \underline{\mathcal{R}}_1^\infty(\mathsf{D}_{\text{th},1}^\infty)$ and $\overline{\mathcal{R}}_1^\infty(\mathsf{D}_{\text{th},2}^\infty) = \underline{\mathcal{R}}_1^\infty(\mathsf{D}_{\text{th},2}^\infty)$. In summary, the analysis of this case yields (2.4.47).

Case 4 ($\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2 > 0$ and $\rho_X = 0$): In this case, we are under the second condition of Theorem 3 since

$$\mu_1 = \frac{1}{2} - \frac{1}{2}\sqrt{1 - \frac{4L}{L-1}\lambda_Y^{-2}\gamma_Y^2} \tag{2.5.85a}$$

$$= \frac{1}{2} - \frac{1}{2}\sqrt{1 - \frac{4L}{L-1}\left(\frac{(1-\rho_Y)^2}{L^2\rho_Y^2} + O\left(\frac{1}{L^3}\right)\right)} \tag{2.5.85b}$$

$$= \frac{(1-\rho_Y)^2}{L^2\rho_Y^2} + O\left(\frac{1}{L^3}\right) \tag{2.5.85c}$$

$$\leq \frac{\gamma_Y}{\lambda_Y} \quad \text{for all sufficiently large } L, \tag{2.5.85d}$$

and

$$1 > \mu_2 \tag{2.5.86a}$$

$$= 1 - \frac{(1-\rho_Y)^2}{L^2\rho_Y^2} + O\left(\frac{1}{L^3}\right) \tag{2.5.86b}$$

$$> \frac{\gamma_Y}{\lambda_Y} \quad \text{for all sufficiently large } L, \tag{2.5.86c}$$

and

$$\lambda_X^2\lambda_Y^{-2}\gamma_X^{-2}\gamma_Y^2 = O(\frac{1}{L^2}) \tag{2.5.87a}$$

$$< \frac{L-1}{4L} \quad \text{for all sufficiently large } L, \tag{2.5.87b}$$

where (2.5.85d) and (2.5.86c) are due to $\frac{\gamma_Y}{\lambda_Y} = \frac{1-\rho_Y}{L\rho_Y} + O(\frac{1}{L^2})$. Here, $\mathsf{D}_{\text{th},1}$ simplifies as follows:

$$\mathsf{D}_{\text{th},1} = \sigma_X^2 - \sigma_X^4\gamma_Y^{-1} + O\left(\frac{1}{L}\right) \tag{2.5.88a}$$

$$= d_{\min}^{\infty} + O\left(\frac{1}{L}\right). \tag{2.5.88b}$$

So, for all $D \in (d_{\min}^{\infty}, \sigma_X^2)$, the lower bound is given by $\underline{\mathcal{R}}^c(D)$ when L is large enough. It just remains to approximate $\underline{\mathcal{R}}^c(D)$. Notice that the second clause of (2.4.8) is active since

$$L\lambda_X^2(\lambda_Y - \gamma_Y)^{-1} + (L-1)\gamma_X^2(\gamma_X^{-1} - \gamma_Y^{-1}) + \lambda_X = L(\sigma_X^2 - \sigma_X^4\gamma_Y^{-1}) + O(1) \quad (2.5.89a)$$

$$= Ld_{\min}^{\infty} + O(1) < LD. \quad (2.5.89b)$$

The rate-distortion expression $\underline{\mathcal{R}}_2^c(D)$ can be approximated as follows:

$$\underline{\mathcal{R}}_2^c(D) = \frac{L}{2}\log\frac{(L-1)\gamma_X^2\gamma_Y^{-1}}{LD - \lambda_X - (L-1)(\gamma_X - \gamma_X^2\gamma_Y^{-1})} \quad (2.5.90a)$$

$$= \frac{L}{2}\log\frac{\sigma_X^4}{\gamma_Y D - \sigma_X^2\gamma_Z} - \frac{1}{2}\frac{D - \sigma_X^2}{D - \sigma_X^2 + \sigma_X^4\gamma_Y^{-1}} + O\left(\frac{1}{L}\right) \quad (2.5.90b)$$

$$= \underline{\mathcal{R}}_2^{\infty}(D). \quad (2.5.90c)$$

In summary, the analysis of this case yields (2.4.48). This concludes the proof.

2.6 Conclusion

We have studied the problem of distributed compression of symmetrically correlated Gaussian sources. An explicit lower bound on the rate-distortion function is established and is shown to partially coincide with the Berger-Tung upper bound. The asymptotic expressions for the upper and lower bounds are derived in the large L limit. It is of considerable theoretical interest to develop new bounding techniques to close the gap between the two bounds.

Chapter 3

Symmetric Remote Gaussian Source Coding with a Centralized Encoder

3.1 Abstract

The problem of symmetric remote Gaussian source coding with a centralized encoder is considered in this chapter. The rate-distortion function for this problem is completely characterized and is leveraged as a rate-distortion lower bound for the symmetric remote Gaussian multiterminal source coding problem. It is shown that this centralized-encoding lower bound is not as tight as the lower bound established via the fictitious signal-noise decomposition approach. The asymptotic analysis of this centralized-encoding lower bound is also provided.

3.2 Introduction

The Internet of Things(IoT) has witnessed growing popularity due to its diverse applications, one of the main challenges that arises when deploying the wireless sensors network is to transmit and reconstruct the noise-corrupted data with the communication cost constraints. One possible way of balancing the trade-off between reducing communication costs and promoting the reconstruction quality is to exploit the statistical dependency among the data at different sensors. Multiterminal source coding theory provides a systematic guideline for the implementation of such pre-processing, there has been a significant amount of works over the past decades in this area.

In their celebrated paper, Slepian and Wolf [26] showed that it is possible to encode correlated information sources separately and decode them jointly with a vanishing error, without paying more in rate relative to jointly encoding the sources. That is, even though the encoders operate independently of one another, they can reduce the sum rate to the joint entropy of the dependent sources. Then Wyner and Ziv [37] extend rate-distortion theory to the case in which side information is present at the decoder. Berger [3] and Tung [27] extend the Wyner-Ziv coding and generalize the Slepian-Wolf problem by considering general distortion criteria on the source reconstruction.

An interesting regime that has received particular attention is when the number of encoders in the network tends to infinity [33]. This asymptotic regime reflects the typical scenarios in various emerging machine learning applications where the number of clients in the centralized sensor network is large, e.x. federated learning [18].

Another special attention has been paid to the setting known as generalized quadratic Gaussian multiterminal source coding, where the unobserved source and additive noise are jointly Gaussian and the distortion measure is deployed as the mean square error. For the centralized coding scheme, the rate-distortion function is given by the celebrated reverse water-filling result [11]. However, for the distributed coding case, the exact characterization of the rate-distortion function remains an open problem, Wagner et al. [29] gave a complete solution for the case when $L = 2$, beyond that the understanding is rather limited. Moreover, there is strong evidence that for most generalized Gaussian multiterminal source coding systems, the rate-distortion function might not be expressed in a closed form solution [39]. Indeed, the existing conclusive results for the distributed coding case are typically given in the form of convex programming [34, 33, 41]. Therefore, even if one manages to solve the generalized Gaussian multiterminal source coding problem completely, extracting useful insights from such a solution can still be non-trivial.

In this part we consider an indirect lossy source coding system with centralized encoder and decoder while L correlated sources and L correlated noises comprise the corrupted observed source, the sources and noises are independent. This setting enables the encoder to jointly compress the noise corrupted source and forwards the compressed data to the decoder, the decoder is required to reconstruct the corresponding target source within a prescribed mean squared error distortion threshold. It is assumed that the observed sources can be expressed as the sum of the target sources and the corruptive noises, which are generated independently from two symmetric multivariate Gaussian distributions.

Our problem settings share some similarities with the afore-mentioned reverse water-filling problem and Gaussian CEO problem [7, 20]. However, the major difference between our setting and the reverse water-filling problem is that in our setting we consider a more general scenario where additive noise could be correlated. Moreover, previous work has shown that this type of theoretic difficulty caused by correlated noises can be circumvented through a fictitious signal-noise decomposition of the observed sources, such idea can be found in [29, 33]. The major difference between CEO problem and our work is that the target signal for CEO problem is a scalar process, while in our settings it is a vector process.

In Wang et al. [33] work, they considered a generalized version multiterminal source coding problem with distributed encoding scheme and gave a closed form solution of the upper bound of the rate-distortion function. Later in Zhou et al. [41] work, they solved the lower bound of the rate-distortion function explicitly, the noticeable find on the asymptotic gap between the upper bound and lower bound makes people more curious about the characteristics of the rate-distortion function, this discovery leads directly to the motivation to our work. Inspired by Wagner et al. [29] and Wang et al. [29], where the formal work determines the rate region of the quadratic Gaussian two-encoder source-coding problem by comparing two types of lower bound on the sum rate, while the latter work also determined the rate region of that problem, by bounding the rate-distortion function with the upper bound and the lower bound, our intuition is to find a lower bound of the rate-distortion function that would be better than that described in [41] settings.

Intuitively, the optimal rate-distortion performance of any generalized multiterminal source coding system must be no superior to that of its centralized counterpart and no inferior to that of its distributed counterpart, then in this work, our questions boil down to whether the asymptotic rate-distortion function of centralized setting is greater than or equal to the lower bound of the rate-distortion function for distributed setting.

The rest of this chapter is organized as follows. The problem definitions and the main results are presented in Section 3.3. The detailed proofs are provided in Sections 3.5. The chapter is concluded in Section 3.6.

3.3 System Model

Consider a L terminal indirect distributed source coding system with centralized encoding scheme and distributed encoding scheme. For centralized case, encoder encodes $\ell \in \{1, \cdots, L\}$ observed source sequences $\{X_\ell^{(n)}\}$ with a encoding function $\phi^{(n)} : \mathbb{R}^{L\times n} \to \mathcal{M}^{L\times n}$, producing a set of indexes $\{M_\ell^{(n)}\}$, then the decoder reconstructs target source sequences by implementing a mapping $g^{(n)} : \mathcal{M}^{L\times n} \to \mathbb{R}^{L\times n}$, producing a set of estimation of source sequences $\hat{X}_\ell^{(n)}$. Similarly setting for distributed case, where,encoding function $\phi_\ell^{(n)} : \mathbb{R}^n \to \mathcal{M}^n$ and a decoding function $g_\ell^{(n)} : \mathcal{M}^n \to \mathbb{R}^n$.

Now we are in the position of giving the achievable definition of the rate-distortion pair.

Definition 2 (Centralized encoding)**.** *A rate-distortion pair (R, D) is said to be achievable with centralized encoding if, for any $\epsilon > 0$, there exists an encoding function $\phi^{(n)}$ such that*

$$\begin{aligned} &\frac{1}{n}\log\left|M^{(n)}\right| \le R + \epsilon, \\ &\frac{1}{Ln}\sum_{\ell=1}^{L}\sum_{i=1}^{n}\mathbb{E}\left[\left(X_\ell(i) - \hat{X}_\ell(i)\right)^2\right] \le D + \epsilon, \end{aligned} \tag{3.3.1}$$

where $\hat{X}_\ell(i) \triangleq \mathbb{E}\left[X_\ell(t) \mid \left(\phi^{(n)}\left(Y^n\right)\right)\right]$. For a given D, the minimum R such that

(R, D) is achievable with centralized encoding is denoted by $\widehat{\mathcal{R}}(D)$.

Definition 3 (Distributed encoding)**.** *A rate-distortion pair (R, D) is said to be achievable with distributed encoding if, for any $\epsilon > 0$, there exist encoding functions $\phi_\ell^{(n)}, \ell \in \{1, \cdots, L\}$, such that*

$$\begin{aligned} &\frac{1}{n}\sum_{\ell=1}^{L}\log\left|M_\ell^{(n)}\right| \leq R + \epsilon, \\ &\frac{1}{Ln}\sum_{\ell=1}^{L}\sum_{i=1}^{n}\mathbb{E}\left[\left(X_\ell(i) - \hat{X}_\ell(i)\right)^2\right] \leq D + \epsilon, \end{aligned} \tag{3.3.2}$$

where $\hat{X}_\ell(i) \triangleq \mathbb{E}\left[X_\ell(i) \mid \left(\phi_1^{(n)}\left(Y_1^n\right), \cdots, \phi_L^{(n)}\left(Y_L^n\right)\right)\right]$. For a given D, the minimum R such that (R, D) is achievable with distributed encoding is denoted by $\widetilde{\mathcal{R}}(D)$.

We have already defined $\widehat{\mathcal{R}}(D)$ as the rate-distortion function of symmetrically correlated Gaussian source coding with centralized encoding and $\widetilde{\mathcal{R}}(D)$ as the rate-distortion function of symmetrically correlated Gaussian source coding with distributed encoding. One can know from [41] that an explicit lower bound for $\widetilde{\mathcal{R}}(D)$ is established and it matches the well-known Berger-Tung upper bound for some values of the distortion threshold. We will define $\underline{\mathcal{R}}(D)$ as the lower bound of the rate-distortion function with distributed encoding.

3.3.1 Preliminary

Let $Y \triangleq (Y_1, \cdots, Y_L)^T$ be the sum of two mutually independent L-dimensional ($L \geq 2$) zero-mean Gaussian random vectors, source $X \triangleq (X_1, \cdots, X_L)^T$ and noise $Z \triangleq (Z_1, \cdots, Z_L)^T$, which can be expressed as

$$Y_\ell = X_\ell + Z_\ell, \qquad \ell \in \{1, \ldots, L\}. \tag{3.3.3}$$

The distributions of X, Y and Z are determined by their covariance matrices Σ_*, $* \in \{X, Y, Z\}$, written as

$$\Sigma_* \triangleq \begin{pmatrix} \sigma_*^2 & \rho_*\sigma_*^2 & \dots & \rho_*\sigma_*^2 \\ \rho_*\sigma_*^2 & \sigma_*^2 & \dots & \rho_*\sigma_*^2 \\ \vdots & \vdots & \dots & \vdots \\ \rho_*\sigma_*^2 & \rho_*\sigma_*^2 & \dots & \sigma_*^2 \end{pmatrix}, \tag{3.3.4}$$

which satisfy $\Sigma_Y = \Sigma_X + \Sigma_Z$ (i.e., $\sigma_Y^2 = \sigma_X^2 + \sigma_Z^2$ and $\rho_Y\sigma_Y^2 = \rho_X\sigma_X^2 + \rho_Z\sigma_Z^2$). To ensure that the covariance matrices are positive semi-definite and the source vector X is not deterministic, we assume $\sigma_X^2 > 0$, $\sigma_Z^2 \geq 0$, $\rho_X \in [-\frac{1}{L-1}, 1]$ and $\rho_Z \in [-\frac{1}{L-1}, 1]$. Moreover the source vector X together with the noise vector Z and the corrupted version Y generates an i.i.d. process $\{(X(t), Y(t), Z(t))\}_{t=1}^{\infty}$.

By the eigenvalue decomposition, a given $L \times L$ matrix

$$\Gamma \triangleq \begin{pmatrix} \alpha & \beta & \dots & \beta \\ \beta & \alpha & \dots & \beta \\ \vdots & \vdots & \dots & \vdots \\ \beta & \beta & \dots & \alpha \end{pmatrix}, \tag{3.3.5}$$

can be written as

$$\Gamma = \Theta\Lambda\Theta^T, \tag{3.3.6}$$

where Θ is an arbitrary unitary matrix with the first column being $\frac{1}{\sqrt{L}}\mathbf{1}_L^T$ and

$$\Lambda \triangleq \mathrm{diag}^{(L)}\left(\alpha + (L-1)\beta, \alpha - \beta, \dots, \alpha - \beta\right). \tag{3.3.7}$$

Based on this, we can write

$$\Sigma_* = \Theta\Lambda_*\Theta^T, \quad * \in \{X, Y, Z\}, \tag{3.3.8}$$

where

$$\Lambda_* \triangleq \mathrm{diag}^{(L)}(\lambda_*\gamma_*, \ldots, \gamma_*), \tag{3.3.9}$$

with

$$\lambda_* \triangleq (1 + (L-1)\rho_*)\sigma_*^2, \tag{3.3.10}$$

$$\gamma_* \triangleq (1 - \rho_*)\sigma_*^2. \tag{3.3.11}$$

Note that $\lambda_Y = \lambda_X + \lambda_Z$ and $\gamma_Y = \gamma_X + \gamma_Z$.

3.3.2 Remark

Let $d_{\min}$ be the minimum achievable distortion when $\{Y(t)\}_{t=1}^{\infty}$ is directly available at the decoder ($\widehat{\mathcal{R}}(D) = \widetilde{\mathcal{R}}(D) = \infty$ for $D \leq d_{\min}$), where (detailed derivation is provided in Chapter 3.5)

$$d_{\min} \triangleq \begin{cases} \frac{(L-1)\sigma_X^2\gamma_Z}{L\sigma_X^2+(L-1)\gamma_Z}, & \rho_X = -\frac{1}{L-1}, \\ \sigma_X^2 - \frac{\lambda_X^2}{L\lambda_Y} - \frac{(L-1)\gamma_X^2}{L\gamma_Y}, & \rho_X \in \left(-\frac{1}{\ell-1}, 1\right), \\ \frac{\sigma_X^2\lambda_Z}{L\sigma_X^2+\lambda_Z}, & \rho_X = 1, \end{cases} \tag{3.3.12}$$

Moreover, it is easy to show that $\widehat{\mathcal{R}}(D) = \widetilde{\mathcal{R}}(D) = 0$ for $D \geq \sigma_X^2$ (since the distortion constraint is trivially satisfied with the reconstruction set to be zero). Henceforth we shall focus on the case $D \in (d_{\min}, \sigma_X^2)$.

3.4 Main Results

Theorem 6. *For* $D \in (d_{\min}, \sigma_X^2)$,

$$\underline{\mathcal{R}}(D) \geq \widehat{\mathcal{R}}(D). \tag{3.4.1}$$

Proof. See Chapter 3.5.1. □

Theorem 7 (Centralized encoding)**.** *For* $D \in (d_{\min}, \sigma_X^2)$,

$$\widehat{\mathcal{R}}(D) = \begin{cases} \frac{L-1}{2} \log \frac{L\sigma_X^4}{\left(L\sigma_X^2+(L-1)\gamma_Z\right)D-(L-1)\sigma_X^2\gamma_Z}, & \rho_X = -\frac{1}{L-1}, \\ \frac{1}{2}\log^+ \frac{\lambda_X^2}{\lambda_Y\delta} + \frac{L-1}{2}\log^+ \frac{\gamma_X^2}{\gamma_Y\delta}, & \rho_X \in \left(-\frac{1}{\ell-1}, 1\right), \\ \frac{1}{2}\log \frac{L\sigma_X^4}{(L\sigma_X^2+\lambda_Z)D-\sigma_X^2\lambda_Z}, & \rho_X = 1, \end{cases} \tag{3.4.2}$$

where

$$\delta \triangleq \begin{cases} D - d_{\min}, & D \leq \min\left\{\frac{\lambda_X^2}{\lambda_Y}, \frac{\gamma_X^2}{\gamma_Y}\right\} + d_{\min}, \\ \frac{L(D-d_{\min})}{L-1} - \frac{\lambda_X^2}{(L-1)\lambda_Y}, & D > \frac{\lambda_X^2}{\lambda_Y} + d_{\min} \\ L\left(D - d_{\min}\right) - \frac{(L-1)\gamma_X^2}{\gamma_Y}, & D > \frac{\gamma_X^2}{\gamma_Y} + d_{\min}. \end{cases} \tag{3.4.3}$$

Proof. See Chapter 3.5.2. □

Example 1: In this example, we compare the lower bound of rate-distortion function $\underline{\mathcal{R}}(D)$ under distributed encoding with the rate-distortion function $\widehat{\mathcal{R}}(D)$ under centralized encoding. In Fig. 3.1, we plot the difference between $\underline{\mathcal{R}}(D)$ and $\widehat{\mathcal{R}}(D)$,

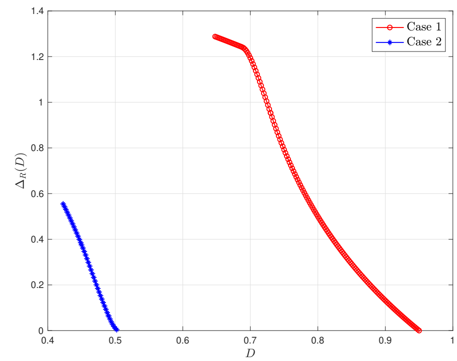

Figure 3.1: $\Delta_R(D)$ with $D \in (d_{\min}, \sigma_X^2)$ for Case 1 and Case 2.

denoted by $\Delta_R(D)$, with $D \in (d_{\min}, \sigma_X^2)$ for the following two cases. According to [41], in these two cases, there exists a gap between the lower bound $\underline{\mathcal{R}}(D)$ and the Berger-Tung upper bound for some values of the distortion threshold. We set $L = 10$.

- Case 1: $\lambda_X = 0.5$, $\gamma_X = 1$, $\lambda_Y = 6$, and $\gamma_Y = 3$. In this case, we have $d_{\min} \approx 0.646$ and $\sigma_X^2 = 0.95$.
- Case 2: $\lambda_X = 1$, $\gamma_X = 0.45$, $\lambda_Y = 12$, and $\gamma_Y = 2.4$. In this case, we have $d_{\min} = 0.4207$ and $\sigma_X^2 = 0.505$.

As can be observed from the figure, the gap $\Delta_R(D)$ is nonnegative, which verifies $\widehat{\mathcal{R}}(D)$ will not be greater than $\underline{\mathcal{R}}(D)$.

In the following, we proceed to study the asymptotic behavior of the two rate-distortion bounds when $L \to \infty$. In the discussion below, it is necessary to restrict attention to the case $\rho_X, \rho_Z \in [0, 1]$; moreover, without loss of generality, we assume

$D \in \left(d_{\min}^{(\infty)}, \sigma_X^2\right)$, where

$$d_{\min}^{(\infty)} \triangleq \lim_{\ell\to\infty} d_{\min} = \begin{cases} \frac{\rho_X\sigma_X^2\rho_Z\sigma_Z^2}{\rho_X\sigma_X^2+\rho_Z\sigma_Z^2} + \frac{\gamma_X\gamma_Z}{\gamma_X+\gamma_Z}, & \rho_X \in (0,1), \\ \frac{\sigma_X^2\gamma_Z}{\sigma_X^2+\gamma_Z}, & \rho_X = 0, \\ \frac{\sigma_X^2\rho_Z\sigma_Z^2}{\sigma_X^2+\rho_Z\sigma_Z^2}, & \rho_X = 1. \end{cases} \tag{3.4.4}$$

Theorem 8 (Centralized encoding for asymptotic regime)**.** *For* $D \in \left(d_{\min}^{(\infty)}, \sigma_X^2\right)$,

1. $\rho_X = 0$, $\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2 > 0$ ($\rho_Z > 0$, $\sigma_Z^2 \neq 0$),

$$\widehat{\mathcal{R}}^{\infty}(D) = \frac{L-1}{2}\log\frac{\sigma_X^4}{(\sigma_X^2+\gamma_Z)D - \sigma_X^2\gamma_Z} + \frac{1}{2}\frac{\sigma_X^4}{(\sigma_X^2+\gamma_Z)D - \sigma_X^2\gamma_Z} - \frac{1}{2} + O\left(\frac{1}{L}\right). \tag{3.4.5}$$

2. $\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2 = 0$ ($\rho_X = 0$, $\rho_Z = 0$ *or* $\sigma_Z^2 = 0$),

$$\widehat{\mathcal{R}}^{\infty}(D) = \frac{L}{2}\log\frac{\sigma_X^4}{(\sigma_X^2+\sigma_Z^2)D - \sigma_X^2\sigma_Z^2}. \tag{3.4.6}$$

3. $\rho_X > 0$, $\rho_X\sigma_X^2 + \rho_Z\sigma_Z^2 > 0$,

$$\widehat{\mathcal{R}}^{\infty}(D) = \begin{cases} \frac{L}{2}\log\frac{\gamma_X^2(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)}{(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)((\gamma_X+\gamma_Z)D-\gamma_X\gamma_Z)-\rho_X\sigma_X^2\rho_Z\sigma_Z^2(\gamma_X+\gamma_Z)} & \\ +\frac{1}{2}\log L + \hat{\alpha} + O\left(\frac{1}{L}\right), & D < \mathsf{D}_{th,0}^{\infty}, \\ \frac{1}{2}\log L + \frac{1}{2}\log\frac{\rho_X^2\sigma_X^4(\gamma_X+\gamma_Z)}{(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)\gamma_X^2} + \frac{1}{2}\frac{(\gamma_Z\rho_X\sigma_X^2-\gamma_X\rho_Z\sigma_Z^2)^2}{\gamma_X^2(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)^2} + O\left(\frac{1}{L}\right), & D = \mathsf{D}_{th,0}^{\infty}, \\ \frac{1}{2}\log\frac{\rho_X^2\sigma_X^4}{(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)(D-\gamma_X)-\rho_X\sigma_X^2\rho_Z\sigma_Z^2} + O\left(\frac{1}{L}\right), & D > \mathsf{D}_{th,0}^{\infty}, \end{cases} \tag{3.4.7}$$

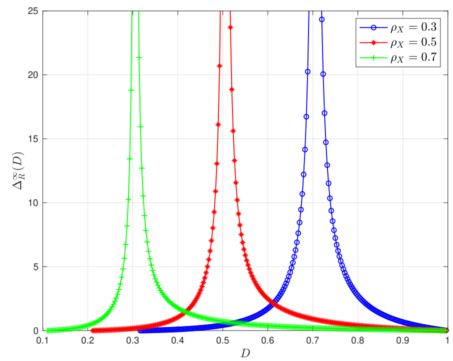

Figure 3.2: $\Delta_R^\infty(D)$ with $\sigma_X^2 = 1$, $\rho_Z = 0.05$ and $\sigma_Z^2 = 0.1$ for different ρ_X.

where

$$\hat{\alpha} \triangleq \frac{1}{2}\log\frac{\rho_X^2\sigma_X^4(\gamma_X+\gamma_Z)}{(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)\gamma_X^2} + \frac{1}{2}\frac{(\gamma_Z\rho_X\sigma_X^2-\gamma_X\rho_Z\sigma_Z^2)^2}{(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)^2((\gamma_X+\gamma_Z)D-\gamma_X\gamma_Z)-\rho_X\sigma_X^2\rho_Z\sigma_Z^2(\gamma_X+\gamma_Z)(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)} \tag{3.4.8}$$

and

$$\mathsf{D}_{th,0}^\infty \triangleq \frac{\rho_X\rho_Z\sigma_X^2\sigma_Z^2}{\rho_X\sigma_X^2+\rho_Z\sigma_Z^2}+\gamma_X. \tag{3.4.9}$$

Proof. See Section 3.5.3. □

Example 2: In this example, we plot the function $\Delta_R^\infty(D)$, which characterized the asymptotic difference between $\underline{\mathcal{R}}^\infty(D)$ and $\widehat{\mathcal{R}}^\infty(D)$ (as L tends to infinity). See Fig. 3.2 and 3.3 for some graphical illustrations of $\Delta_R^\infty(D)$ with different parameter settings.

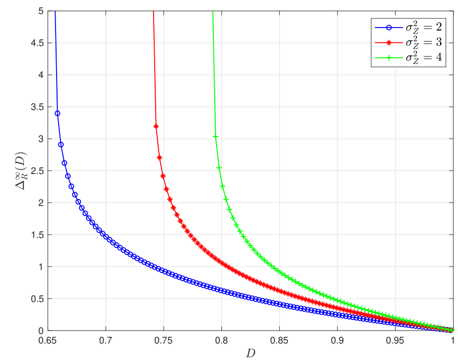

Figure 3.3: $\Delta_R^\infty(D)$ with $\rho_X = 0$, $\sigma_X^2 = 1$ and $\rho_Z = 0.05$ for different σ_Z^2.

3.5 Proof of Results

3.5.1 Proof of Theorem 6

From [34] [12], we know that $\widehat{\mathcal{R}}(D)$ is given by the solution of the following optimization problem:

$$\begin{aligned} \widehat{\mathcal{R}}(D) = \min_{p_{\hat{X}|Y}} \quad & I(Y;\hat{X}) \\ \textbf{s.t.} \quad & \mathbb{E}\left[(X-\hat{X})^T(X-\hat{X})\right] \le LD, \end{aligned} \tag{3.5.1}$$

where

$$I(Y;\hat{X}) = \frac{1}{2}\log\frac{\det(\Sigma_Y)}{\det(\mathcal{D})} = \frac{1}{2}\log\frac{\lambda_Y}{\alpha} + \frac{L-1}{2}\log\frac{\gamma_Y}{\beta} \triangleq \Theta(\alpha,\beta). \tag{3.5.2}$$

It is known that a result of [33][41] provides a lower bound on the rate-distortion function with distributed encoding by solving a minimization program, which defines

$$\Omega(\alpha,\beta,\delta) = \frac{1}{2}\log\frac{\lambda_Y^2}{(\lambda_Y-\lambda_W)\alpha+\lambda_Y\lambda_W} + \frac{L-1}{2}\log\frac{\gamma_Y^2}{(\gamma_Y-\lambda_W)\beta+\gamma_Y\lambda_W} + \frac{L}{2}\log\frac{\lambda_W}{\delta},$$

where $\lambda_W = \min(\lambda_Y, \gamma_Y)$. Let $\underline{\mathcal{R}}(D)$ be the solution of the following optimization problem:

$$
\begin{aligned}
\underline{\mathcal{R}}(D) = \min_{\alpha,\beta,\delta} \ & \Omega(\alpha,\beta,\delta), \\
\textbf{s.t.} \ & 0 < \alpha \le \lambda_Y, \\
& 0 < \beta \le \gamma_Y, \\
& 0 < \delta, \\
& \delta \le (\alpha^{-1} + \lambda_W^{-1} - \lambda_Y^{-1})^{-1}, \\
& \delta \le (\beta^{-1} + \lambda_W^{-1} - \gamma_Y^{-1})^{-1}, \\
& \lambda_X^2 \lambda_Y^{-2} \alpha + \lambda_X - \lambda_X^2 \lambda_Y^{-1} + (L-1)(\gamma_X^2 \gamma_Y^{-2} \beta + \gamma_X - \gamma_X^2 \gamma_Y^{-1}) \le LD.
\end{aligned} \tag{3.5.3}
$$

One can readily prove the constraints under (α, β) for distributed case are same as the centralized case. In the following, we compare $\Theta(\alpha,\beta)$ and $\Omega(\alpha,\beta,\delta)$ under the restriction of the given constraints (3.5.3).

Subcase 1: $\lambda_Y \ge \gamma_Y > 0$. We can send $\lambda_W = \gamma_Y$, then

$$
\begin{aligned}
& \delta \le (\alpha^{-1} + \gamma_Y^{-1} - \lambda_Y^{-1})^{-1}, \\
& \delta \le \beta.
\end{aligned} \tag{3.5.4}
$$

1. If $(\alpha^{-1} + \gamma_Y^{-1} - \lambda_Y^{-1})^{-1} \ge \beta$, we have

$$
\Omega(\alpha,\beta,\delta) \ge \frac{1}{2} \log \frac{\lambda_Y^2}{(\lambda_Y - \gamma_Y)\alpha + \lambda_Y \gamma_Y} + \frac{L}{2} \log \frac{\gamma_Y}{\beta} \tag{3.5.5}
$$

and

$$
\Theta(\alpha,\beta) = \frac{1}{2} \log \frac{\lambda_Y}{\alpha} + \frac{L-1}{2} \log \frac{\gamma_Y}{\beta}. \tag{3.5.6}
$$

Note that

$$\frac{1}{2}\log\frac{\lambda_Y}{\alpha} - \frac{1}{2}\log\frac{\lambda_Y^2}{(\lambda_Y-\gamma_Y)\alpha+\lambda_Y\gamma_Y} = \frac{1}{2}\log\frac{\alpha+\gamma_Y(1-\frac{\alpha}{\lambda_Y})}{\alpha} \geq 0 \tag{3.5.7}$$

due to $\alpha \leq \lambda_Y$ and

$$\frac{L}{2}\log\frac{\gamma_Y}{\beta} - \frac{L-1}{2}\log\frac{\gamma_Y}{\beta} = \frac{1}{2}\log\frac{\gamma_Y}{\beta} \geq 0, \tag{3.5.8}$$

so that

$$\frac{1}{2}\log\frac{\gamma_Y}{\beta} - \frac{1}{2}\log\frac{\alpha+\gamma_Y(1-\frac{\alpha}{\lambda_Y})}{\alpha} = \frac{1}{2}\log\frac{(\alpha^{-1}+\gamma_Y^{-1}-\lambda_Y^{-1})^{-1}}{\beta} \tag{3.5.9}$$

is always non-negative since the assumption $(\alpha^{-1}+\gamma_Y^{-1}-\lambda_Y^{-1})^{-1} \geq \beta$, which means $\Omega(\alpha,\beta,\delta) \geq \Theta(\alpha,\beta)$ under this case.

2. If $(\alpha^{-1}+\gamma_Y^{-1}-\lambda_Y^{-1})^{-1} < \beta$, we have

$$\Omega(\alpha,\beta,\delta) \geq \frac{1}{2}\log\frac{\lambda_Y^2}{(\lambda_Y-\gamma_Y)\alpha+\lambda_Y\gamma_Y} + \frac{L}{2}\log\frac{\gamma_Y}{(\alpha^{-1}+\gamma_Y^{-1}-\lambda_Y^{-1})^{-1}} \tag{3.5.10}$$

and

$$\Theta(\alpha,\beta) = \frac{1}{2}\log\frac{\lambda_Y}{\alpha} + \frac{L-1}{2}\log\frac{\gamma_Y}{\beta}. \tag{3.5.11}$$

Note that

$$\frac{1}{2}\log\frac{\lambda_Y}{\alpha} - \frac{1}{2}\log\frac{\lambda_Y^2}{(\lambda_Y-\gamma_Y)\alpha+\lambda_Y\gamma_Y} = \frac{1}{2}\log\frac{\alpha+\gamma_Y(1-\frac{\alpha}{\lambda_Y})}{\alpha} \geq 0 \tag{3.5.12}$$

due to $\alpha \le \lambda_Y$. Since $(\alpha^{-1} + \gamma_Y^{-1} - \lambda_Y^{-1})^{-1} < \beta$, we have

$$\frac{L}{2} \log \frac{\gamma_Y}{(\alpha^{-1} + \gamma_Y^{-1} - \lambda_Y^{-1})^{-1}} - \frac{L-1}{2} \log \frac{\gamma_Y}{\beta} > \frac{1}{2} \frac{\gamma_Y}{(\alpha^{-1} + \gamma_Y^{-1} - \lambda_Y^{-1})^{-1}} \quad (3.5.13)$$

and it can be easily verified that

$$\frac{1}{2} \frac{\gamma_Y}{(\alpha^{-1} + \gamma_Y^{-1} - \lambda_Y^{-1})^{-1}} = \frac{1}{2} \log \frac{\alpha + \gamma_Y(1 - \frac{\alpha}{\lambda_Y})}{\alpha}. \quad (3.5.14)$$

Therefore, $\Omega(\alpha, \beta, \delta) \ge \Theta(\alpha, \beta)$ under this case.

Subcase 2: $\gamma_Y \ge \lambda_Y > 0$. We can send $\lambda_W = \lambda_Y$, then

$$\begin{aligned} &\delta \le \alpha, \\ &\delta \le (\beta^{-1} + \lambda_Y^{-1} - \gamma_Y^{-1})^{-1}. \end{aligned} \quad (3.5.15)$$

1. If $(\beta^{-1} + \lambda_Y^{-1} - \gamma_Y^{-1})^{-1} \ge \alpha$, we have

$$\Omega(\alpha, \beta, \delta) \ge \frac{L-1}{2} \log \frac{\gamma_Y^2}{(\gamma_Y - \lambda_Y)\beta + \gamma_Y \lambda_Y} + \frac{L}{2} \log \frac{\lambda_Y}{\alpha} \quad (3.5.16)$$

and

$$\Theta(\alpha, \beta) = \frac{L-1}{2} \log \frac{\gamma_Y}{\beta} + \frac{1}{2} \log \frac{\lambda_Y}{\alpha}. \quad (3.5.17)$$

Note that

$$\frac{L-1}{2} \log \frac{\gamma_Y}{\beta} - \frac{L-1}{2} \log \frac{\gamma_Y^2}{(\gamma_Y - \lambda_Y)\beta + \lambda_Y \gamma_Y} = \frac{L-1}{2} \log \frac{\beta + \lambda_Y(1 - \frac{\beta}{\gamma_Y})}{\beta} \ge 0 \quad (3.5.18)$$

due to $\beta \leq \gamma_Y$ and

$$\frac{L}{2}\log\frac{\lambda_Y}{\alpha} - \frac{1}{2}\log\frac{\lambda_Y}{\alpha} = \frac{L-1}{2}\log\frac{\lambda_Y}{\alpha} \geq 0, \tag{3.5.19}$$

so that

$$\frac{L-1}{2}\log\frac{\lambda_Y}{\alpha} - \frac{L-1}{2}\log\frac{\beta + \lambda_Y(1-\frac{\beta}{\gamma_Y})}{\beta} = \frac{L-1}{2}\log\frac{(\beta^{-1}+\lambda_Y^{-1}-\gamma_Y^{-1})^{-1}}{\alpha} \tag{3.5.20}$$

is always non-negative since the assumption $(\beta^{-1}+\lambda_Y^{-1}-\gamma_Y^{-1})^{-1} \geq \alpha$, which means $\Omega(\alpha,\beta,\delta) \geq \Theta(\alpha,\beta)$ under this case.

2. If $(\beta^{-1}+\lambda_Y^{-1}-\gamma_Y^{-1})^{-1} < \alpha$, we have

$$\Omega(\alpha,\beta,\delta) \geq \frac{L-1}{2}\log\frac{\gamma_Y^2}{(\gamma_Y-\lambda_Y)\beta+\gamma_Y\lambda_Y} + \frac{L}{2}\log\frac{\lambda_Y}{(\beta^{-1}+\lambda_Y^{-1}-\gamma_Y^{-1})^{-1}} \tag{3.5.21}$$

and

$$\Theta(\alpha,\beta) = \frac{L-1}{2}\log\frac{\gamma_Y}{\beta} + \frac{1}{2}\log\frac{\lambda_Y}{\alpha}. \tag{3.5.22}$$

Note that

$$\frac{L-1}{2}\log\frac{\gamma_Y}{\beta} - \frac{L-1}{2}\log\frac{\gamma_Y^2}{(\gamma_Y-\lambda_Y)\beta+\lambda_Y\gamma_Y} = \frac{L-1}{2}\log\frac{\beta+\lambda_Y(1-\frac{\beta}{\gamma_Y})}{\beta} \geq 0 \tag{3.5.23}$$

due to $\beta \leq \gamma_Y$. Since $(\beta^{-1}+\lambda_Y^{-1}-\gamma_Y^{-1})^{-1} < \alpha$, we have

$$\frac{L}{2}\log\frac{\lambda_Y}{(\beta^{-1}+\lambda_Y^{-1}-\gamma_Y^{-1})^{-1}} - \frac{1}{2}\log\frac{\lambda_Y}{\alpha} > \frac{L-1}{2}\frac{\lambda_Y}{(\beta^{-1}+\lambda_Y^{-1}-\gamma_Y^{-1})^{-1}}, \tag{3.5.24}$$

and it can be easily verified that

$$\frac{L-1}{2}\frac{\lambda_Y}{(\beta^{-1}+\lambda_Y^{-1}-\gamma_Y^{-1})^{-1}} = \frac{L-1}{2}\log\frac{\beta+\lambda_Y(1-\frac{\beta}{\gamma_Y})}{\beta}, \tag{3.5.25}$$

which results $\Omega(\alpha, \beta, \delta) \geq \Theta(\alpha, \beta)$ under this case.

In summary, it is always true that $\Omega(\alpha, \beta, \delta) \geq \Theta(\alpha, \beta)$ under all conditions. Together with the same constraints, it can be concluded that $\underline{\mathcal{R}}(D) \geq \widehat{\mathcal{R}}(D)$. This completes the proof of Theorem 6.

3.5.2 Proof of Theorem 7

$\widehat{\mathcal{R}}(D)$ is given by the solution of the following optimization problem:

$$\begin{aligned} &\min_{p_{\hat{X}|Y}} \; I(Y;\hat{X}) \\ &\textbf{s.t. } \mathbb{E}\left[(X-\hat{X})^T(X-\hat{X})\right] \leq LD, \\ &X \leftrightarrow Y \leftrightarrow \hat{X} \text{ form a Markov chain.} \end{aligned} \tag{3.5.26}$$

It can be transferred equivalently to

$$\begin{aligned} &\min_{p_{\hat{X}|\bar{Y}}} \; I(\bar{Y};\hat{X}) \\ &\textbf{s.t. } \mathbb{E}\left[(\bar{X}-\hat{X})^T(\bar{X}-\hat{X})\right] \leq LD, \\ &\bar{X} \leftrightarrow \bar{Y} \leftrightarrow \hat{X} \text{ form a Markov chain.} \end{aligned} \tag{3.5.27}$$

where $\bar{X} \triangleq \Theta^T X, \bar{Z} \triangleq \Theta^T Z$, and $\bar{Y} \triangleq \Theta^T Y$. As defined in Chapter 3.3, Θ^T is an arbitrary unitary matrix with the first row being $\frac{1}{\sqrt{L}} 1_L$. Moreover, we can have

$$Ld_{\min} = \mathbb{E}\left[(\bar{X} - \hat{Y})^T (\bar{X} - \hat{Y})\right] = \mathbb{E}\left[(\bar{X} - \mathbb{E}[\bar{X} \mid \bar{Y}])^T (\bar{X} - \mathbb{E}[\bar{X} \mid \bar{Y}])\right]. \quad (3.5.28)$$

where $\hat{Y} \triangleq \mathbb{E}[\bar{X} \mid \bar{Y}]$.

It can be verified that $\bar{X}$ and $\bar{Z}$ are two independent zero-mean Gaussian random vectors with covariance matrices

$$\Lambda_* \triangleq \operatorname{diag}^{(L)}(\lambda_* \gamma_*, \ldots, \gamma_*), \quad * \in \{X, Z\}. \quad (3.5.29)$$

Denote the ℓ-th components of $\bar{X}, \bar{Z}$ and $\bar{Y}$ by $\bar{X}_\ell, \bar{Z}_\ell$ and $\bar{Y}_\ell$, respectively, $\ell = 1, \cdots, L$. We have

$$\begin{aligned}
\mathbb{E}\left[\left(\bar{X}_1\right)^2\right] &= \lambda_X = L\rho_X \sigma_X^2 + \gamma_X, \\
\mathbb{E}\left[\left(\bar{X}_\ell\right)^2\right] &= \gamma_X, \quad \ell = 2, \cdots, L, \\
\mathbb{E}\left[\left(\bar{Z}_1\right)^2\right] &= \lambda_Z = L\rho_Z \sigma_Z^2 + \gamma_Z, \\
\mathbb{E}\left[\left(\bar{Z}_\ell\right)^2\right] &= \gamma_Z, \quad \ell = 2, \cdots, L.
\end{aligned} \quad (3.5.30)$$

Since $\bar{Y}_\ell = \bar{X}_\ell + \bar{Z}_\ell$,

$$\begin{aligned}
\mathbb{E}\left[\left(\bar{Y}_1\right)^2\right] &= \lambda_Y = L\rho_Y \sigma_Y^2 + \gamma_Y = L(\rho_X \sigma_X^2 + \rho_Z \sigma_Z^2) + \gamma_X + \gamma_Z, \\
\mathbb{E}\left[\left(\bar{Y}_\ell\right)^2\right] &= \gamma_Y = \gamma_X + \gamma_Z, \quad \ell = 2, \cdots, L.
\end{aligned} \quad (3.5.31)$$

Now denote the ℓ-th component of $\hat{Y}$ by $\hat{Y}_\ell \triangleq \mathbb{E}[\hat{X}_\ell \mid \hat{Y}_\ell], \ell = 1, \cdots, L$. According to linear estimation, we have

$$\begin{aligned}
&\mathbb{E}\left[\left(\hat{Y}_1\right)^2\right] = \begin{cases} 0, & \rho_X = -\frac{1}{L-1}, \\ \frac{\left(L\rho_X\sigma_X^2+\gamma_X\right)^2}{L(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)+\gamma_X+\gamma_Z}, & \rho_X \in \left(-\frac{1}{L-1}, 1\right], \end{cases} \\
&\mathbb{E}\left[\left(\hat{Y}_\ell\right)^2\right] = \begin{cases} \frac{\gamma_X^2}{\gamma_X+\gamma_Z}, & \rho_X \in \left[-\frac{1}{L-1}, 1\right), \\ 0, & \rho_X = 1, \end{cases} \quad \ell = 2, \cdots, L.
\end{aligned} \tag{3.5.32}$$

Note that

$$\mathbb{E}\left[(\bar{X}-\hat{Y})^T(\bar{X}-\hat{Y})\right] = \sum_{\ell=1}^{L}\mathbb{E}\left[\left(\bar{X}_\ell\right)^2\right] - \sum_{\ell=1}^{L}\mathbb{E}\left[\left(\hat{Y}_\ell\right)^2\right], \tag{3.5.33}$$

this holds because $\hat{Y} \triangleq \mathbb{E}[\bar{X} \mid \bar{Y}]$. Therefore, together with equations (1)-(4), we derive

$$\begin{aligned}
d_{\min} &= \frac{1}{L}\mathbb{E}\left[(\bar{X}-\hat{Y})^T(\bar{X}-\hat{Y})\right] = \sum_{\ell=1}^{L}\mathbb{E}\left[\left(\bar{X}_\ell\right)^2\right] - \sum_{\ell=1}^{L}\mathbb{E}\left[\left(\hat{Y}_\ell\right)^2\right] \\
&= \begin{cases} \frac{(L-1)\sigma_X^2\gamma_Z}{L\sigma_X^2+(L-1)\gamma_Z}, & \rho_X = -\frac{1}{L-1}, \\ \sigma_X^2 - \frac{\lambda_X^2}{L\lambda_Y} - \frac{(L-1)\gamma_X^2}{L\gamma_Y}, & \rho_X \in \left(-\frac{1}{\ell-1}, 1\right), \\ \frac{\sigma_X^2\lambda_Z}{L\sigma_X^2+\lambda_Z}, & \rho_X = 1, \end{cases}
\end{aligned} \tag{3.5.34}$$

Since it is obvious that $\hat{Y}$ is determined by $\bar{Y}$ and from (3.5.27) we know $\bar{X} \leftrightarrow \bar{Y} \leftrightarrow \hat{X}$ form a Markov chain, the distortion constraint can be written as

$$\begin{aligned}
\mathbb{E}\left[(\bar{X}-\hat{X})^T(\bar{X}-\hat{X})\right] &= \mathbb{E}\left[(\bar{X}-\hat{Y})^T(\bar{X}-\hat{Y})\right] + \mathbb{E}\left[(\hat{Y}-\hat{X})^T(\hat{Y}-\hat{X})\right] \\
&= Ld_{\min} + \mathbb{E}\left[(\hat{Y}-\hat{X})^T(\hat{Y}-\hat{X})\right]
\end{aligned} \tag{3.5.35}$$

Therefore, (3.5.27) is equivalent to

$$\begin{aligned} \min_{p_{\hat{X}|\hat{Y}}} \quad & I(\hat{Y};\hat{X}) \\ \textbf{s.t.} \quad & \mathbb{E}\left[(\hat{Y}-\hat{X})^T(\hat{Y}-\hat{X})\right] \le L(D-d_{\min}), \end{aligned} \tag{3.5.36}$$

which is the well-known reverse water-filling problem. Consider the following three subcases separately.

1. When $\rho_X = -\frac{1}{L-1}$, we have

$$\widehat{\mathcal{R}}(D) = \frac{L-1}{2}\log\frac{\gamma_X^2}{(\gamma_X+\gamma_Z)\delta_1}, \tag{3.5.37}$$

where $\delta_1 = \frac{L}{L-1}(D-d_{\min})$, which is derived by

$$L(D-d_{\min}) = (L-1)\delta_1. \tag{3.5.38}$$

In this case, $\gamma_X = \frac{L}{L-1}\sigma_X^2$ and $d_{\min} = \frac{(L-1)\sigma_X^2\gamma_Z}{L\sigma_X^2+(L-1)\gamma_Z}$, then we get

$$\widehat{\mathcal{R}}(D) = \frac{L-1}{2}\log\frac{L\sigma_X^4}{(L\sigma_X^2+(L-1)\gamma_Z)\,D-(L-1)\sigma_X^2\gamma_Z}. \tag{3.5.39}$$

2. When $\rho_X = 1$, we have

$$\widehat{\mathcal{R}}(D) = \frac{1}{2}\log\frac{(L\rho_X\sigma_X^2+\gamma_X)^2}{(L(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)+\gamma_X+\gamma_Z)\,\delta_2}, \tag{3.5.40}$$

where $\delta_2 = L(D-d_{\min})$. In this case, $\gamma_X = 0$ and $d_{\min} = \frac{\sigma_X^2\lambda_Z}{L\sigma_X^2+\lambda_Z}$, thus

$$\widehat{\mathcal{R}}(D) = \frac{1}{2}\log\frac{L\sigma_X^4}{(L\sigma_X^2+\lambda_Z)D-\sigma_X^2\lambda_Z}. \tag{3.5.41}$$

3. When $\rho_X \in (-\frac{1}{L-1}, 1)$, according to [11] Theorem 10.3.3, one can readily get the reverse water-filling solution as follows:

$$\widehat{\mathcal{R}}(D) = \frac{1}{2}\log^{+}\frac{\lambda_X^2}{\lambda_Y\delta} + \frac{L-1}{2}\log^{+}\frac{\gamma_X^2}{\gamma_Y\delta}. \tag{3.5.42}$$

If $D - d_{\min} \leq \min\{\frac{\lambda_X^2}{\lambda_Y}, \frac{\gamma_X^2}{\gamma_Y}\}$, then

$$\delta = D - d_{\min}. \tag{3.5.43}$$

If $D - d_{\min} > \frac{\lambda_X^2}{\lambda_Y}$, then

$$\begin{aligned} L(D - d_{\min}) &= \frac{\lambda_X^2}{\lambda_Y} + (L-1)\delta, \\ \delta &= \frac{L\,(D - d_{\min})}{L-1} - \frac{\lambda_X^2}{(L-1)\lambda_Y}. \end{aligned} \tag{3.5.44}$$

If $D - d_{\min} > \frac{\gamma_X^2}{\gamma_Y}$, then

$$\begin{aligned} L(D - d_{\min}) &= \delta + (L-1)\frac{\gamma_X^2}{\gamma_Y}, \\ \delta &= L(D - d_{\min}) - \frac{(L-1)\gamma_X^2}{\gamma_Y}. \end{aligned} \tag{3.5.45}$$

In summary, for $D \in (d_{\min}, \sigma_X^2)$,

$$\widehat{\mathcal{R}}(D) = \begin{cases} \frac{L-1}{2}\log\frac{L\sigma_X^4}{\left(L\sigma_X^2+(L-1)\gamma_Z\right)D-(L-1)\sigma_X^2\gamma_Z}, & \rho_X = -\frac{1}{L-1}, \\ \frac{1}{2}\log^{+}\frac{\lambda_X^2}{\lambda_Y\delta} + \frac{L-1}{2}\log^{+}\frac{\gamma_X^2}{\gamma_Y\delta}, & \rho_X \in \left(-\frac{1}{\ell-1}, 1\right), \\ \frac{1}{2}\log\frac{L\sigma_X^4}{(L\sigma_X^2+\lambda_Z)D-\sigma_X^2\lambda_Z}, & \rho_X = 1, \end{cases} \tag{3.5.46}$$

where

$$\delta \triangleq \begin{cases} D - d_{\min}, & D \leq \min\left\{\frac{\lambda_X^2}{\lambda_Y}, \frac{\gamma_X^2}{\gamma_Y}\right\} + d_{\min}, \\ \frac{L(D-d_{\min})}{L-1} - \frac{\lambda_X^2}{(L-1)\lambda_Y}, & D > \frac{\lambda_X^2}{\lambda_Y} + d_{\min} \\ L\left(D - d_{\min}\right) - \frac{(L-1)\gamma_X^2}{\gamma_Y}, & D > \frac{\gamma_X^2}{\gamma_Y} + d_{\min}. \end{cases} \tag{3.5.47}$$

In particular, when $\rho_X = 0$ and $\rho_Z \in (0, 1]$, we have

$$\widehat{\mathcal{R}}(D) = \frac{1}{2}\log^+ \frac{\sigma_X^4}{(\sigma_X^2 + \lambda_Z)\delta} + \frac{L-1}{2}\log^+ \frac{\sigma_X^4}{(\sigma_X^2 + \gamma_Z)\delta}, \tag{3.5.48}$$

where

$$\delta \triangleq \begin{cases} D - d_{\min}, & D \leq \frac{\sigma_X^4}{\sigma_X^2+\lambda_Z} + d_{\min}, \\ \frac{L(D-d_{\min})}{L-1} - \frac{\sigma_X^4}{(L-1)(\sigma_X^2+\lambda_Z)}, & D > \frac{\sigma_X^4}{\sigma_X^2+\lambda_Z} + d_{\min}. \end{cases} \tag{3.5.49}$$

This completes the proof of Theorem 7.

3.5.3 Proof of Theorem 8

1. Setting $\rho_X = 0$ and $\rho_Z \in (0, 1]$, first we have

$$\begin{aligned}
d_{\min} &= \sigma_X^2 - \frac{\lambda_X^2}{L\lambda_Y} - \frac{(L-1)\gamma_X^2}{L\gamma_Y} && \text{(3.5.50a)}\\
&= \sigma_X^2 - \frac{1}{L}\frac{\sigma_X^4}{\sigma_X^2+\lambda_Z} - \frac{L-1}{L}\frac{\sigma_X^4}{\sigma_X^2+\gamma_Z} && \text{(3.5.50b)}\\
&= \frac{\sigma_X^2\gamma_Z}{\sigma_X^2+\gamma_Z} + \frac{1}{L}\frac{\sigma_X^4(\lambda_Z-\gamma_Z)}{(\sigma_X^2+\gamma_Z)(\sigma_X^2+\lambda_Z)} && \text{(3.5.50c)}\\
&= \frac{\sigma_X^2\gamma_Z}{\sigma_X^2+\gamma_Z} + \frac{\sigma_X^4\rho_Z\sigma_Z^2}{(\sigma_X^2+\gamma_Z)\rho_Z\sigma_Z^2 L + (\sigma_X^2+\gamma_Z)^2} && \text{(3.5.50d)}\\
&= \frac{\sigma_X^2\gamma_Z}{\sigma_X^2+\gamma_Z} + \frac{\sigma_X^4}{\sigma_X^2+\gamma_Z}\frac{1}{L}\left(\frac{1}{1+\frac{\sigma_X^2+\gamma_Z}{\rho_Z\sigma_Z^2}\frac{1}{L}}\right) && \text{(3.5.50e)}\\
&= \frac{\sigma_X^2\gamma_Z}{\sigma_X^2+\gamma_Z} + \frac{\sigma_X^4}{\sigma_X^2+\gamma_Z}\frac{1}{L}\left(1-\frac{\sigma_X^2+\gamma_Z}{\rho_Z\sigma_Z^2}\frac{1}{L}+O\left(\frac{1}{L^2}\right)\right) && \text{(3.5.50f)}\\
&= \frac{\sigma_X^2\gamma_Z}{\sigma_X^2+\gamma_Z} + \frac{\sigma_X^4}{\sigma_X^2+\gamma_Z}\frac{1}{L} - \frac{\sigma_X^4}{\rho_Z\sigma_Z^2}\frac{1}{L^2}+O\left(\frac{1}{L^3}\right) && \text{(3.5.50g)}\\
&\triangleq d_{\min}^{(\infty)} + \frac{\sigma_X^4}{\sigma_X^2+\gamma_Z}\frac{1}{L} - \frac{\sigma_X^4}{\rho_Z\sigma_Z^2}\frac{1}{L^2}+O\left(\frac{1}{L^3}\right). && \text{(3.5.50h)}
\end{aligned}$$

where (3.5.50f) holds because $\frac{1}{1+x} = 1 - x + O(x^2)$. Moreover here and later we assume $\sigma_Z^2 \neq 0$.

For the special case where $\sigma_Z^2 = 0$, we have $d_{\min} = 0$ and

$$\widehat{\mathcal{R}}(D) = \frac{L}{2}\log\frac{\sigma_X^2}{D}, \quad \text{for } D \in (0, \sigma_X^2), \tag{3.5.51}$$

which can be generalized into the second case of Theorem 8.

For this case, it can be deduced from Theorem 7 that

$$\widehat{\mathcal{R}}(D) = \begin{cases} \frac{1}{2}\log\frac{\sigma_X^4}{(\sigma_X^2+\lambda_Z)(D-d_{\min})} + \frac{L-1}{2}\log\frac{\sigma_X^4}{(\sigma_X^2+\gamma_Z)(D-d_{\min})}, & D \in (d_{\min}, \frac{\sigma_X^4}{\sigma_X^2+\lambda_Z} + d_{\min}], \\ \frac{L-1}{2}\log\frac{(L-1)\sigma_X^4(\sigma_X^2+\lambda_Z)}{(\sigma_X^2+\gamma_Z)\left(L(D-d_{\min})(\sigma_X^2+\lambda_Z)-\sigma_X^4\right)}, & D \in (\frac{\sigma_X^4}{\sigma_X^2+\lambda_Z} + d_{\min}, \sigma_X^2), \end{cases} \tag{3.5.52}$$

We can also know when $L \to \infty$, only the second condition will be active, this is because

$$\begin{aligned} \frac{\sigma_X^4}{\sigma_X^2+\lambda_Z} + d_{\min} &= \frac{\sigma_X^4}{L\rho_Z\sigma_Z^2+\gamma_Z+\sigma_X^2} + d_{\min} \\ &= \frac{\sigma_X^2\gamma_Z}{\sigma_X^2+\gamma_Z} + O\left(\frac{1}{L}\right) \\ &= d_{\min}^{(\infty)} + O\left(\frac{1}{L}\right). \end{aligned} \tag{3.5.53}$$

Therefore it follows that $d \in (d_{\min}^{(\infty)}, \sigma_X^2)$ and consequently

$$\widehat{\mathcal{R}}(D) = \frac{L-1}{2}\log\frac{(L-1)\sigma_X^4(\sigma_X^2+\lambda_Z)}{(\sigma_X^2+\gamma_Z)\left(L(D-d_{\min})(\sigma_X^2+\lambda_Z)-\sigma_X^4\right)} \tag{3.5.54}$$

when L comes to sufficiently large. Note that

$$\begin{aligned} & L(D-d_{\min})(\sigma_X^2+\lambda_Z) \\ =& L\left(D - \frac{\sigma_X^2\gamma_Z}{\sigma_X^2+\gamma_Z} - \frac{\sigma_X^4}{\sigma_X^2+\gamma_Z}\frac{1}{L} + \frac{\sigma_X^4}{\rho_Z\sigma_Z^2}\frac{1}{L^2} - O\left(\frac{1}{L^3}\right)\right)(L\rho_Z\sigma_Z^2+\gamma_Z+\sigma_X^2) \\ =& \left(D - \frac{\sigma_X^2\gamma_Z}{\sigma_X^2+\gamma_Z}\right)\rho_Z\sigma_Z^2L^2 + \left((\sigma_X^2+\gamma_Z)D - \sigma_X^2\gamma_Z - \frac{\sigma_X^4\rho_Z\sigma_Z^2}{\sigma_X^2+\gamma_Z}\right)L + O\left(\frac{1}{L}\right) \end{aligned} \tag{3.5.55}$$

where the first equality holds for (3.5.50g) and

$$\begin{aligned}(L-1)\sigma_X^4(\sigma_X^2+\lambda_Z) &= (L-1)\sigma_X^4(L\rho_Z\sigma_Z^2+\gamma_Z+\sigma_X^2)\\ &= \sigma_X^4\rho_Z\sigma_Z^2L^2+\sigma_X^4(\sigma_X^2+\gamma_Z-\rho_Z\sigma_Z^2)L-\sigma_X^4(\sigma_X^2+\gamma_Z).\end{aligned} \tag{3.5.56}$$

Plugging (3.5.55) and (3.5.56) into (3.5.54) gives

$$\begin{aligned}&\widehat{\mathcal{R}}(D)\\ =&\frac{L-1}{2}\log\frac{\sigma_X^4\rho_Z\sigma_Z^2L^2+\sigma_X^4(\sigma_X^2+\gamma_Z-\rho_Z\sigma_Z^2)L-\sigma_X^4(\sigma_X^2+\gamma_Z)}{((\sigma_X^2+\gamma_Z)D-\sigma_X^2\gamma_Z)\rho_Z\sigma_Z^2L^2+\left((\sigma_X^2+\gamma_Z)((\sigma_X^2+\gamma_Z)D-\sigma_X^2\gamma_Z)-\sigma_X^4\rho_Z\sigma_Z^2\right)L-\sigma_X^4(\sigma_X^2+\gamma_Z)+O\left(\frac{1}{L}\right)}\\ =&\frac{L-1}{2}\log\frac{\sigma_X^4\rho_Z\sigma_Z^2L+\sigma_X^4(\sigma_X^2+\gamma_Z-\rho_Z\sigma_Z^2)}{((\sigma_X^2+\gamma_Z)D-\sigma_X^2\gamma_Z)\rho_Z\sigma_Z^2L+\left((\sigma_X^2+\gamma_Z)((\sigma_X^2+\gamma_Z)D-\sigma_X^2\gamma_Z)-\sigma_X^4\rho_Z\sigma_Z^2\right)}+O\left(\frac{1}{L}\right)\\ =&\frac{L-1}{2}\log\frac{\sigma_X^4}{(\sigma_X^2+\gamma_Z)D-\sigma_X^2\gamma_Z}+\frac{L-1}{2}\log\frac{\rho_Z\sigma_Z^2L+(\sigma_X^2+\gamma_Z-\rho_Z\sigma_Z^2)}{\rho_Z\sigma_Z^2L+(\sigma_X^2+\gamma_Z)-\frac{\sigma_X^4\rho_Z\sigma_Z^2}{(\sigma_X^2+\gamma_Z)D-\sigma_X^2\gamma_Z}}+O\left(\frac{1}{L}\right)\\ =&\frac{L-1}{2}\log\frac{\sigma_X^4}{(\sigma_X^2+\gamma_Z)D-\sigma_X^2\gamma_Z}+\frac{L-1}{2}\log\left(1+\frac{\frac{\sigma_X^4\rho_Z\sigma_Z^2}{(\sigma_X^2+\gamma_Z)D-\sigma_X^2\gamma_Z}-\rho_Z\sigma_Z^2}{\rho_Z\sigma_Z^2L+(\sigma_X^2+\gamma_Z)-\frac{\sigma_X^4\rho_Z\sigma_Z^2}{(\sigma_X^2+\gamma_Z)D-\sigma_X^2\gamma_Z}}\right)+O\left(\frac{1}{L}\right)\\ =&\frac{L-1}{2}\log\frac{\sigma_X^4}{(\sigma_X^2+\gamma_Z)D-\sigma_X^2\gamma_Z}+\frac{L-1}{2}\left(\frac{\frac{\sigma_X^4\rho_Z\sigma_Z^2}{(\sigma_X^2+\gamma_Z)D-\sigma_X^2\gamma_Z}-\rho_Z\sigma_Z^2}{\rho_Z\sigma_Z^2L+(\sigma_X^2+\gamma_Z)-\frac{\sigma_X^4\rho_Z\sigma_Z^2}{(\sigma_X^2+\gamma_Z)D-\sigma_X^2\gamma_Z}}+O\left(\frac{1}{L^2}\right)\right)+O\left(\frac{1}{L}\right)\\ =&\frac{L-1}{2}\log\frac{\sigma_X^4}{(\sigma_X^2+\gamma_Z)\,D-\sigma_X^2\gamma_Z}+\frac{1}{2}\frac{\sigma_X^4}{(\sigma_X^2+\gamma_Z)\,D-\sigma_X^2\gamma_Z}-\frac{1}{2}+O\left(\frac{1}{L}\right)\end{aligned} \tag{3.5.57}$$

where (3.5.57) is due to $\log(1+x)=x+O(x^2)$. In particular, if $\rho_Z=1$ ($\gamma_Z=0$), we have $d_{\min}^{(\infty)}=0$ and

$$\widehat{\mathcal{R}}(D)=\frac{L-1}{2}\log\frac{\sigma_X^2}{D}+\frac{1}{2}\frac{\sigma_X^2}{D}-\frac{1}{2}+O\left(\frac{1}{L}\right). \tag{3.5.58}$$

2. Setting $\rho_X = 0$ and $\rho_Z = 0$, we have

$$\begin{aligned} d_{\min} &= \sigma_X^2 - \frac{\sigma_X^4}{(\sigma_X^2 + \sigma_Z^2)L} - \frac{L-1}{L}\frac{\sigma_X^4}{\sigma_X^2 + \sigma_Z^2} \\ &= \sigma_X^2 - \frac{\sigma_X^4}{\sigma_X^2 + \sigma_Z^2} \\ &= \frac{\sigma_X^2\sigma_Z^2}{\sigma_X^2 + \sigma_Z^2} \end{aligned} \tag{3.5.59}$$

and

$$\begin{aligned} \widehat{\mathcal{R}}(D) &= \frac{1}{2}\log\frac{\sigma_X^4}{(\sigma_X^2 + \sigma_Z^2)\delta} + \frac{L-1}{2}\log\frac{\sigma_X^4}{(\sigma_X^2 + \sigma_Z^2)\delta} \\ &= \frac{L}{2}\log\frac{\sigma_X^4}{(\sigma_X^2 + \sigma_Z^2)\delta} \end{aligned} \tag{3.5.60}$$

where $\delta = D - d_{\min}$ since $D - d_{\min} = D - \sigma_X^2 + \frac{\sigma_X^4}{\sigma_X^2+\sigma_Z^2} < \frac{\sigma_X^4}{\sigma_X^2+\sigma_Z^2}$, so

$$\widehat{\mathcal{R}}(D) = \frac{L}{2}\log\frac{\sigma_X^4}{\left(\sigma_X^2 + \sigma_Z^2\right)D - \sigma_X^2\sigma_Z^2} \tag{3.5.61}$$

for $D \in (\frac{\sigma_X^2\sigma_Z^2}{\sigma_X^2+\sigma_Z^2}, \sigma_X^2)$.

3. Setting $\rho_X = 1$ and $\rho_Z \in [0, 1]$, we have

$$\begin{aligned} d_{\min} &= \frac{\sigma_X^2\lambda_Z}{L\sigma_X^2 + \lambda_Z} \\ &= \frac{L\sigma_X^2\rho_Z\sigma_Z^2 + \sigma_X^2\gamma_Z}{L(\sigma_X^2 + \rho_Z\sigma_Z^2) + \gamma_Z} \\ &= \frac{\sigma_X^2\rho_Z\sigma_Z^2}{\sigma_X^2 + \rho_Z\sigma_Z^2} + O\left(\frac{1}{L}\right) \\ &\triangleq d_{\min}^{(\infty)} + O\left(\frac{1}{L}\right) \end{aligned} \tag{3.5.62}$$

and

$$\begin{aligned}\widehat{\mathcal{R}}(D) &= \frac{1}{2}\log\frac{L\sigma_X^4}{(L\sigma_X^2+\lambda_Z)D-\sigma_X^2\lambda_Z}\\ &= \frac{1}{2}\log\frac{L\sigma_X^4}{(L\sigma_X^2+L\rho_Z\sigma_Z^2+\gamma_Z)D-\sigma_X^2(L\rho_Z\sigma_Z^2+\gamma_Z)}\\ &= \frac{1}{2}\log\frac{\sigma_X^4}{(\sigma_X^2+\rho_Z\sigma_Z^2)\,D-\sigma_X^2\rho_Z\sigma_Z^2}+O\left(\frac{1}{L}\right)\end{aligned} \tag{3.5.63}$$

for $D \in (\frac{\sigma_X^2\rho_Z\sigma_Z^2}{\sigma_X^2+\rho_Z\sigma_Z^2}, \sigma_X^2)$.

4. Setting $\rho_X \in (0,1)$ and $\rho_Z \in [0,1]$, we have

$$\begin{aligned}d_{\min} &= \sigma_X^2-\frac{\lambda_X^2}{L\lambda_Y}-\frac{(L-1)\gamma_X^2}{L\gamma_Y}\\ &= \sigma_X^2-\frac{(L\rho_X\sigma_X^2+\gamma_X)^2}{L^2(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)+(\gamma_X+\gamma_Z)L}-\frac{L-1}{L}\frac{\gamma_X^2}{\gamma_X+\gamma_Z}\\ &= \sigma_X^2-\frac{\gamma_X^2}{\gamma_X+\gamma_Z}+\frac{\gamma_X^2}{L(\gamma_X+\gamma_Z)}-\frac{(L\rho_X\sigma_X^2+\gamma_X)^2}{L^2(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)+(\gamma_X+\gamma_Z)L}\\ &= \rho_X\sigma_X^2+\frac{\gamma_X\gamma_Z}{\gamma_X+\gamma_Z}+\frac{-(\gamma_X+\gamma_Z)\rho_X^2\sigma_X^4L+\gamma_X(\gamma_X\rho_Z\sigma_Z^2-(\gamma_X+2\gamma_Z)\rho_X\sigma_X^2)}{(\gamma_X+\gamma_Z)(L(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)+\gamma_X+\gamma_Z)}\\ &= \rho_X\sigma_X^2-\frac{\rho_X^2\sigma_X^4}{\rho_X\sigma_X^2+\rho_Z\sigma_Z^2}+\frac{\gamma_X\gamma_Z}{\gamma_X+\gamma_Z}+\frac{(\gamma_Z\rho_X\sigma_X^2-\gamma_X\rho_Z\sigma_Z^2)^2}{(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)^2(\gamma_X+\gamma_Z)L}+O\left(\frac{1}{L^2}\right)\\ &= \frac{\rho_X\sigma_X^2\rho_Z\sigma_Z^2}{\rho_X\sigma_X^2+\rho_Z\sigma_Z^2}+\frac{\gamma_X\gamma_Z}{\gamma_X+\gamma_Z}+\frac{(\gamma_Z\rho_X\sigma_X^2-\gamma_X\rho_Z\sigma_Z^2)^2}{(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)^2(\gamma_X+\gamma_Z)L}+O\left(\frac{1}{L^2}\right)\\ &\triangleq d_{\min}^{(\infty)}+\frac{(\gamma_Z\rho_X\sigma_X^2-\gamma_X\rho_Z\sigma_Z^2)^2}{(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)^2(\gamma_X+\gamma_Z)L}+O\left(\frac{1}{L^2}\right).\end{aligned} \tag{3.5.64}$$

Moreover, it always holds that $\frac{\lambda_X^2}{\lambda_Y} > \frac{\gamma_X^2}{\gamma_Y}$ when $L \to \infty$, so it can be deduced from Theorem 7 that

$$\widehat{\mathcal{R}}(D) = \begin{cases} \frac{1}{2}\log\frac{(L\rho_X\sigma_X^2+\gamma_X)^2(\gamma_X+\gamma_Z)}{\gamma_X^2(L(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)+\gamma_X+\gamma_Z)} + \frac{L}{2}\log\frac{\gamma_X^2}{(\gamma_X+\gamma_Z)(D-d_{\min})}, & D\in(d_{\min}, \frac{\gamma_X^2}{\gamma_X+\gamma_Z}+d_{\min}], \\ \frac{1}{2}\log\frac{(L\rho_X\sigma_X^2+\gamma_X)^2(\gamma_X+\gamma_Z)}{(L(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)+\gamma_X+\gamma_Z)(L(\gamma_X+\gamma_Z)(D-d_{\min})-(L-1)\gamma_X^2)}, & D\in(\frac{\gamma_X^2}{\gamma_X+\gamma_Z}+d_{\min}, \sigma_X^2), \end{cases} \tag{3.5.65}$$

where $\frac{\gamma_X^2}{\gamma_X+\gamma_Z}+d_{\min}$ converges to $\frac{\rho_X\sigma_X^2\rho_Z\sigma_Z^2}{\rho_X\sigma_X^2+\rho_Z\sigma_Z^2}+\gamma_X$ as $L\to\infty$ due to (3.5.64). Now consider the following two subcases separately.

- $D\in(\frac{\rho_X\sigma_X^2\rho_Z\sigma_Z^2}{\rho_X\sigma_X^2+\rho_Z\sigma_Z^2}+\frac{\gamma_X\gamma_Z}{\gamma_X+\gamma_Z}, \frac{\rho_X\sigma_X^2\rho_Z\sigma_Z^2}{\rho_X\sigma_X^2+\rho_Z\sigma_Z^2}+\gamma_X]$

  In this case, we have $D\in(d_{\min}, \frac{\gamma_X^2}{\gamma_X+\gamma_Z}+d_{\min}]$ and consequently

$$\begin{aligned}\widehat{\mathcal{R}}(D) &= \frac{1}{2}\log\frac{(L\rho_X\sigma_X^2+\gamma_X)^2(\gamma_X+\gamma_Z)}{\gamma_X^2(L(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)+\gamma_X+\gamma_Z)} + \frac{L}{2}\log\frac{\gamma_X^2}{(\gamma_X+\gamma_Z)(D-d_{\min})} \\ &= \frac{L}{2}\log\frac{\gamma_X^2}{\gamma_X+\gamma_Z} - \frac{L}{2}\log(D-d_{\min}) + \frac{1}{2}\log\frac{\gamma_X+\gamma_Z}{\gamma_X^2} \\ &+ \frac{1}{2}\log\frac{(L\rho_X\sigma_X^2+\gamma_X)^2}{L(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)+\gamma_X+\gamma_Z}.\end{aligned} \tag{3.5.66}$$

  Note that

$$\frac{1}{2}\log\frac{(L\rho_X\sigma_X^2+\gamma_X)^2}{L(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)+\gamma_X+\gamma_Z} = \frac{1}{2}\log L + \frac{1}{2}\log\frac{\rho_X^2\sigma_X^4}{\rho_X\sigma_X^2+\rho_Z\sigma_Z^2} + O\left(\frac{1}{L}\right) \tag{3.5.67}$$

and

$$\frac{1}{2}\log(D-d_{\min}) \tag{3.5.68a}$$
$$=\frac{1}{2}\log\left(D-\frac{\rho_X\sigma_X^2\rho_Z\sigma_Z^2}{\rho_X\sigma_X^2+\rho_Z\sigma_Z^2}-\frac{\gamma_X\gamma_Z}{\gamma_X+\gamma_Z}-\frac{(\gamma_Z\rho_X\sigma_X^2-\gamma_X\rho_Z\sigma_Z^2)^2}{(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)^2(\gamma_X+\gamma_Z)L}-O\left(\frac{1}{L^2}\right)\right) \tag{3.5.68b}$$
$$=\frac{1}{2}\log\left(D-\frac{\rho_X\sigma_X^2\rho_Z\sigma_Z^2}{\rho_X\sigma_X^2+\rho_Z\sigma_Z^2}-\frac{\gamma_X\gamma_Z}{\gamma_X+\gamma_Z}\right) \tag{3.5.68c}$$
$$-\frac{1}{2}\frac{(\gamma_Z\rho_X\sigma_X^2-\gamma_X\rho_Z\sigma_Z^2)^2}{(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)^2((\gamma_X+\gamma_Z)D-\gamma_X\gamma_Z)-\rho_X\sigma_X^2\rho_Z\sigma_Z^2(\gamma_X+\gamma_Z)(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)}\frac{1}{L} \tag{3.5.68d}$$
$$+O\left(\frac{1}{L^2}\right) \tag{3.5.68e}$$

where (3.5.68b) is due to (3.5.64) and (3.5.68e) holds for $\log(1+x)=x+O(x^2)$. Substituting (3.5.67) and (3.5.68e) into (3.5.66) and re-arranging the terms yield

$$\begin{aligned}
&\widehat{\mathcal{R}}(D)\\
=&\frac{L}{2}\log\frac{\gamma_X^2(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)}{(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)\left((\gamma_X+\gamma_Z)D-\gamma_X\gamma_Z\right)-\rho_X\sigma_X^2\rho_Z\sigma_Z^2(\gamma_X+\gamma_Z)}\\
&+\frac{1}{2}\log L+\frac{1}{2}\log\frac{\rho_X^2\sigma_X^4\left(\gamma_X+\gamma_Z\right)}{(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)\gamma_X^2}\\
&+\frac{1}{2}\frac{(\gamma_Z\rho_X\sigma_X^2-\gamma_X\rho_Z\sigma_Z^2)^2}{(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)^2\left(\left(\gamma_X+\gamma_Z\right)D-\gamma_X\gamma_Z\right)-\rho_X\sigma_X^2\rho_Z\sigma_Z^2(\gamma_X+\gamma_Z)(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)}\\
&+O\left(\frac{1}{L}\right).
\end{aligned} \tag{3.5.69}$$

In particular, when $D = \frac{\rho_X\sigma_X^2\rho_Z\sigma_Z^2}{\rho_X\sigma_X^2+\rho_Z\sigma_Z^2} + \gamma_X$, we have

$$\widehat{\mathcal{R}}(D) = \frac{1}{2}\log L + \frac{1}{2}\log\frac{\rho_X^2\sigma_X^4\left(\gamma_X+\gamma_Z\right)}{(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)\gamma_X^2} + \frac{1}{2}\frac{(\gamma_Z\rho_X\sigma_X^2-\gamma_X\rho_Z\sigma_Z^2)^2}{\gamma_X^2(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)^2} + O\left(\frac{1}{L}\right). \tag{3.5.70}$$

- $D \in (\frac{\rho_X\sigma_X^2\rho_Z\sigma_Z^2}{\rho_X\sigma_X^2+\rho_Z\sigma_Z^2} + \gamma_X, \sigma_X^2)$

  This case falls into where $D \in (\frac{\gamma_X^2}{\gamma_X+\gamma_Z} + d_{\min}, \sigma_X^2)$, so we have

$$\begin{aligned}\widehat{\mathcal{R}}(D) &= \frac{1}{2}\log\frac{(L\rho_X\sigma_X^2+\gamma_X)^2(\gamma_X+\gamma_Z)}{(L(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)+\gamma_X+\gamma_Z)(L(\gamma_X+\gamma_Z)(D-d_{\min})-(L-1)\gamma_X^2)}\\ &= \frac{1}{2}\log\frac{\rho_X^2\sigma_X^4(\gamma_X+\gamma_Z)}{(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)((\gamma_X+\gamma_Z)(D-d_{\min})-\gamma_X^2)} + O\left(\frac{1}{L}\right),\end{aligned} \tag{3.5.71}$$

  where

$$\begin{aligned}&(\gamma_X+\gamma_Z)(D-d_{\min})-\gamma_X^2\\ =&(\gamma_X+\gamma_Z)(D-\frac{\rho_X\sigma_X^2\rho_Z\sigma_Z^2}{\rho_X\sigma_X^2+\rho_Z\sigma_Z^2}-\frac{\gamma_X\gamma_Z}{\gamma_X+\gamma_Z})-\gamma_X^2\\ =&\frac{(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)(\gamma_X+\gamma_Z)(D-\gamma_X)-\rho_X\sigma_X^2\rho_Z\sigma_Z^2(\gamma_X+\gamma_Z)}{\rho_X\sigma_X^2+\rho_Z\sigma_Z^2}.\end{aligned} \tag{3.5.72}$$

  Substituting (3.5.72) into (3.5.71) yields

$$\widehat{\mathcal{R}}(D) = \frac{1}{2}\log\frac{\rho_X^2\sigma_X^4}{(\rho_X\sigma_X^2+\rho_Z\sigma_Z^2)(D-\gamma_X)-\rho_X\sigma_X^2\rho_Z\sigma_Z^2} + O\left(\frac{1}{L}\right) \tag{3.5.73}$$

  which can cover the case 3 where $\rho_X = 1$ and $\rho_Z \in [0, 1]$.

In summary, we have

$$d_{\min}^{(\infty)} \triangleq \lim_{\ell \to \infty} d_{\min} = \begin{cases} \frac{\rho_X \sigma_X^2 \rho_Z \sigma_Z^2}{\rho_X \sigma_X^2 + \rho_Z \sigma_Z^2} + \frac{\gamma_X \gamma_Z}{\gamma_X + \gamma_Z}, & \rho_X \in (0,1), \\ \frac{\sigma_X^2 \gamma_Z}{\sigma_X^2 + \gamma_Z}, & \rho_X = 0, \\ \frac{\sigma_X^2 \rho_Z \sigma_Z^2}{\sigma_X^2 + \rho_Z \sigma_Z^2}, & \rho_X = 1, \end{cases} \tag{3.5.74}$$

and for $D \in \left(d_{\min}^{(\infty)}, \sigma_X^2\right)$,

1. $\rho_X = 0$, $\rho_Z \in (0,1]$, $\sigma_Z^2 \neq 0$,

$$\widehat{\mathcal{R}}^{\infty}(D) = \frac{L-1}{2} \log \frac{\sigma_X^4}{(\sigma_X^2 + \gamma_Z) D - \sigma_X^2 \gamma_Z} + \frac{1}{2} \frac{\sigma_X^4}{(\sigma_X^2 + \gamma_Z) D - \sigma_X^2 \gamma_Z} - \frac{1}{2} + O\left(\frac{1}{L}\right). \tag{3.5.75}$$

2. $\rho_X = 0$, $\rho_Z = 0$ or $\sigma_Z^2 = 0$,

$$\widehat{\mathcal{R}}^{\infty}(D) = \frac{L}{2} \log \frac{\sigma_X^4}{(\sigma_X^2 + \sigma_Z^2) D - \sigma_X^2 \sigma_Z^2}. \tag{3.5.76}$$

3. $\rho_X \in (0,1]$, $\rho_Z \in [0,1]$,

$$\begin{aligned} &\widehat{\mathcal{R}}^{\infty}(D) \\ &= \begin{cases} \frac{L}{2} \log \frac{\gamma_X^2 (\rho_X \sigma_X^2 + \rho_Z \sigma_Z^2)}{(\rho_X \sigma_X^2 + \rho_Z \sigma_Z^2)((\gamma_X + \gamma_Z) D - \gamma_X \gamma_Z) - \rho_X \sigma_X^2 \rho_Z \sigma_Z^2 (\gamma_X + \gamma_Z)} & \\ + \frac{1}{2} \log L + \hat{\alpha} + O\left(\frac{1}{L}\right), & D < \mathsf{D}_{\text{th},0}^{\infty}, \\ \frac{1}{2} \log L + \frac{1}{2} \log \frac{\rho_X^2 \sigma_X^4 (\gamma_X + \gamma_Z)}{(\rho_X \sigma_X^2 + \rho_Z \sigma_Z^2) \gamma_X^2} + \frac{1}{2} \frac{(\gamma_Z \rho_X \sigma_X^2 - \gamma_X \rho_Z \sigma_Z^2)^2}{\gamma_X^2 (\rho_X \sigma_X^2 + \rho_Z \sigma_Z^2)^2} + O\left(\frac{1}{L}\right), & D = \mathsf{D}_{\text{th},0}^{\infty}, \\ \frac{1}{2} \log \frac{\rho_X^2 \sigma_X^4}{(\rho_X \sigma_X^2 + \rho_Z \sigma_Z^2)(D - \gamma_X) - \rho_X \sigma_X^2 \rho_Z \sigma_Z^2} + O\left(\frac{1}{L}\right), & D > \mathsf{D}_{\text{th},0}^{\infty}, \end{cases} \end{aligned} \tag{3.5.77}$$

where

$$\hat{\alpha} \triangleq \frac{1}{2} \log \frac{\rho_X^2 \sigma_X^4 (\gamma_X + \gamma_Z)}{(\rho_X \sigma_X^2 + \rho_Z \sigma_Z^2)\gamma_X^2} + \frac{1}{2} \frac{(\gamma_Z \rho_X \sigma_X^2 - \gamma_X \rho_Z \sigma_Z^2)^2}{(\rho_X \sigma_X^2 + \rho_Z \sigma_Z^2)^2 ((\gamma_X + \gamma_Z) D - \gamma_X \gamma_Z) - \rho_X \sigma_X^2 \rho_Z \sigma_Z^2 (\gamma_X + \gamma_Z)(\rho_X \sigma_X^2 + \rho_Z \sigma_Z^2)}, \tag{3.5.78}$$

and

$$\mathsf{D}_{\mathrm{th},0}^{\infty} \triangleq \frac{\rho_X \rho_Z \sigma_X^2 \sigma_Z^2}{\rho_X \sigma_X^2 + \rho_Z \sigma_Z^2} + \gamma_X. \tag{3.5.79}$$

This completes the proof of Theorem 8.

3.6 Conclusions

We considered an indirect distributed lossy source coding system with a centralized encoder and a centralized decoder. We provide the explicit rate-distortion function for this setting under both non-asymptotic and asymptotic regimes. Furthermore, we make the comparison between the asymptotic function for the centralized case and the asymptotic distributed lower bound, results show that the asymptotic rate-distortion function for the centralized case is smaller than the asymptotic distributed lower bound, which gave us the hint that the former does not make a better lower bound for the distributed case.

Chapter 4

Conclusion

The problem of symmetric remote Gaussian multiterminal source coding is studied. An explicit lower bound on the rate-distortion function for this problem is derived by solving a convex program induced a fictitious signal-noise decomposition. It is further shown that this lower bound partially coincides with the Berger-Tung upper bound, yielding a complete characterization of the rate-distortion function in certain regimes. The asymptotic expressions of the derived lower bound and the Berger-Tung upper bound are computed when the number of encoders tends to infinity.

The rate-distortion function of symmetric remote Gaussian source coding with a centralized encoder is completely characterized. It then serves as a second rate-distortion lower bound for symmetric remote Gaussian multiterminal source coding. It is shown that this lower bound is not as tight as the first one, demonstrating the advantage of the fictitious signal-noise decomposition approach. The asymptotic behavior of the second lower bound is also analyzed.

It is of considerable interest to fully characterize the rate-distortion function of symmetric remote Gaussian multiterminal source coding. However, the existing bounding techniques appear to be inadequate for this purpose. For future work, we attempt to first identify the looseness of our (first) lower bound and the Berger-Tung upper bound, especially in the regime where the gap between the two diverges asymptotically.

Appendix A

Your Appendix

A.1 Proof of Theorem 1

The proof is built upon the so-called Berger-Tung upper bound [13, Thm 12.1] as summarized in the following lemma.

Lemma 1. *Let* $\boldsymbol{V} := (V_1, \ldots, V_L)^T$ *be an auxiliary random vector jointly distributed with* $(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{Y})$ *such that* $(\boldsymbol{X}, \boldsymbol{Z}, \{Y_\ell\}_{\ell \in \mathcal{A} \setminus \ell}, \{V_\ell\}_{\ell \in \mathcal{A} \setminus \ell}) \to Y_\ell \to V_\ell$ *form a Markov chain for* $\ell \in \mathcal{A}$, $\mathcal{A} = \{1, \ldots, L\}$, *and any* (r, d) *such that*

$$r \in \mathcal{R}(\mathcal{A}), \tag{A.1.1}$$

and

$$d \geq \frac{1}{\ell} \sum_{\ell \in \mathcal{A}} \mathbb{E}[(X_\ell - \mathbb{E}[X_\ell | (V_\ell)_{\ell \in \mathcal{A}}])^2], \tag{A.1.2}$$

where $\mathcal{R}$ denoted the set of $(r_\ell)_{\ell\in\mathcal{A}}$ satisfying

$$\sum_{\ell\in\mathcal{B}} r_\ell \geq I((Y_\ell)_{\ell\in\mathcal{B}};(V_\ell)_{\ell\in\mathcal{B}}|(V_\ell)_{\ell\in\mathcal{A}\setminus\ell\in\mathcal{B}}) \tag{A.1.3}$$

with $\phi\subset\mathcal{B}\subseteq\mathcal{A}$, we have

$$(r,d)\in\mathcal{RD}(\mathcal{A}). \tag{A.1.4}$$

Let $\mathbf{Q}:=(Q_1,\ldots,Q_L)^T$ be an L-dimensional zero-mean Gaussian random vector with covariance matrix

$$\Lambda_Q=\mathrm{diag}^{(L)}(\lambda_Q,\ldots,\lambda_Q)\succ 0, \tag{A.1.5}$$

where $\lambda_Q>0$. We assume $\mathbf{Q}$ is independent of $(\mathbf{X},\mathbf{Z},\mathbf{Y})$. Define the following auxiliary random variables:

$$V_\ell=Y_\ell+Q_\ell,\qquad \ell\in\{1,\ldots,L\}. \tag{A.1.6}$$

Note that the resulting $\mathbf{V}$ satisfies the Markov chain constraints specified in Lemma 1. Let

$$r=\frac{1}{L}I(Y_1,\cdots,Y_L;V_1,\cdots,V_L), \tag{A.1.7}$$

$$d=\frac{1}{L}\sum_{\ell\in\mathcal{A}}\mathbb{E}[(X_\ell-\mathbb{E}[X_\ell|V_1,\cdots,V_L])^2], \tag{A.1.8}$$

we have

$$\begin{aligned} r &= \frac{1}{L}\left(h\left(V_1,\cdots,V_L\right) - h\left(V_1,\cdots,V_L|Y_1,\cdots,Y_L\right)\right) \\ &= \frac{1}{L}\left(h\left(Y_1+Q_1,\cdots,Y_L+Q_L\right) - h\left(Q_1,\cdots,Q_L\right)\right) \\ &= \frac{1}{2L}\log\frac{\det(\Gamma_Y+\Lambda_Q)}{\det(\Lambda_Q)} \\ &= \frac{1}{2L}\log\frac{\det(\Lambda_Y+\Lambda_Q)}{\det(\Lambda_Q)} \\ &= \frac{1}{2L}\log\left(1+\frac{\lambda_Y}{\lambda_Q}\right) + \frac{L-1}{2L}\log\left(1+\frac{\gamma_Y}{\lambda_Q}\right), \end{aligned} \tag{A.1.9}$$

and

$$\begin{aligned} d &= \frac{1}{L}tr\left(\Gamma_X - \Gamma_X\left(\Gamma_Y+\Lambda_Q\right)^{-1}\Gamma_X\right) \\ &= \frac{1}{L}tr\left(\Lambda_X - \Lambda_X\left(\Lambda_Y+\Lambda_Q\right)^{-1}\Lambda_X\right) \\ &= \frac{1}{L}\lambda_X\left(1-\frac{\lambda_X}{\lambda_Y+\lambda_Q}\right) + \frac{L-1}{L}\gamma_X\left(1-\frac{\gamma_X}{\gamma_Y+\lambda_Q}\right), \end{aligned} \tag{A.1.10}$$

which is a strictly increasing function of λ_Q, converging to d_{min} as $\lambda_Q \to 0$ and to γ_X as $\lambda_Q \to \infty$. One can readily complete the proof of Theorem 1 by invoking Lemma 1 and the standard proof of Berger-Tung upper bound.

A.2 Proof of Theorem 2

Let

$$(Y_1,\ldots,Y_L)^T = (U_1,\ldots,U_L)^T + (W_1,\ldots,W_L)^T, \tag{A.2.1}$$

where $(U_1,\ldots,U_L)^T$ and $(W_1,\ldots,W_L)^T$ are two mutually independent L-dimensional zero-mean Gaussian vectors with covariance matrices $\Sigma_U \succ 0$ and

$$\Lambda_W = \mathrm{diag}^{(L)}(\lambda_W,\ldots,\lambda_W) \succ 0. \tag{A.2.2}$$

Then, two auxiliary random processes $\{(U_{1,t}, \ldots, U_{L,t})^T\}_{t=1}^{\infty}$ and $\{(W_{1,t}, \ldots, W_{L,t})^T\}_{t=1}^{\infty}$ are constructed in an i.i.d. manner.

According to Definition 1, for any $R \geq \mathcal{R}(D)$ and $\epsilon > 0$, there exist encoding and decoding functions such that

$$\frac{1}{n}\sum_{\ell=1}^{L} \log|\mathcal{M}_\ell| \leq R + \epsilon, \tag{A.2.3}$$

and

$$\frac{1}{nL}\sum_{\ell=1}^{L}\sum_{t=1}^{n} \mathbb{E}[(X_{\ell,t} - \hat{X}_{\ell,t})^2] \leq D + \epsilon. \tag{A.2.4}$$

The proof is divided into several steps as follows.

Simplifying the Rate Constraint: Lower bounding $\sum_\ell log(|C_\ell^{(n)}|)$ we have

$$\begin{aligned}
LR &\geq \sum_\ell log(|C_\ell^{(n)}|) \\
&\geq H(M_L^{(n)}) \\
&= I(U_L^n; M_L^{(n)}) + H(M_L^{(n)}|U_L^n) \\
&= I(U_L^n; M_L^{(n)}) + I(Y_L^n; M_L^{(n)}|U_L^n) \\
&= h(U_L^n) - h(U_L^n|M_L^{(n)}) + h(Y_L^n|U_L^n) - h(Y_L^n|M_L^{(n)}, U_L^n) \\
&= \frac{n}{2}\log((2\pi e)^L \det(\Gamma_U^{(n)})) + \frac{n}{2}\log((2\pi e)^L \det(\Lambda_W^{(n)})) - h(U_L^n|M_L^{(n)}) - h(Y_L^n|M_L^{(n)}, U_L^n),
\end{aligned} \tag{A.2.5}$$

we let

$$\Delta_{U|M} := \frac{1}{n}\sum_t \mathbb{E}\left[(U_L(t) - \hat{U}_L(t))^T(U_L(t) - \hat{U}_L(t))\right], \tag{A.2.6}$$

$$\Delta_{Y|U,M} := \frac{1}{n}\sum_t \mathbb{E}\left[(Y_L(t) - \hat{Y}_L(t)^T(Y_L(t) - \hat{Y}_L(t)))\right], \tag{A.2.7}$$

where

$$\hat{U}_L(t) := \mathbb{E}\left[U_L(t)|M_L(t)\right], \tag{A.2.8}$$

$$\hat{Y}_L(t) := \mathbb{E}\left[Y_L(t)|U_L(t), M_L(t)\right]. \tag{A.2.9}$$

It then can be verified that

$$\begin{aligned} h(U_L^n|M_L^{(n)}) &= \sum_{t=1}^{n} h(U_L(t)|M_L^{(n)}, U_L^{t-1}) \\ &\leq \sum_{t=1}^{n} h(U_L(t)|M_L^{(n)}) \\ &= \sum_{t=1}^{n} h(U_L(t) - \hat{U}_L(t)|M_L^{(n)}) \\ &\leq \sum_{t=1}^{n} h(U_L(t) - \hat{U}_L(t)) \\ &\leq \sum_{t=1}^{n} \frac{1}{2}\log((2\pi e)^L \det(\Delta_{U|M})) \\ &= \frac{n}{2}\log((2\pi e)^L \det(\Delta_{U|M})), \end{aligned} \tag{A.2.10}$$

which can be achieved by leveraging the maximum differential entropy lemma, and the concavity of the log-determinant function. Similarly, we have

$$\begin{aligned}
&h(Y_L^n|M_L^{(n)}, U_L^n) \\
=&\sum_{t=1}^{n} h(Y_L(t)|M_L^{(n)}, U_L(t), Y_L^{t-1}) \\
\leq&\sum_{t=1}^{n} h(Y_L(t)|M_L^{(n)}, U_L(t)) \\
=&\sum_{t=1}^{n} h(Y_L(t) - \hat{Y}_L(t)|M_L^{(n)}, U_L(t)) \\
\leq&\sum_{t=1}^{n} h(Y_L(t) - \hat{Y}_L(t)) \\
\leq&\sum_{t=1}^{n} \frac{1}{2}\log((2\pi e)^L \det(\Delta_{Y|U,M})) \\
=&\frac{n}{2}\log((2\pi e)^L \det(\Delta_{Y|U,M})).
\end{aligned} \tag{A.2.11}$$

Combining the above results, we get

$$\frac{1}{2L}\log\frac{\det(\Sigma_U)\det(\Lambda_W)}{\det(\Delta_{U|M})\det(\Delta_{Y|U,M})} \leq R + \epsilon. \tag{A.2.12}$$

Simplifying the Distortion Constraint: We define

$$\Delta_{Y|M} := \frac{1}{n}\sum_{t=1}^{n} \mathbb{E}\left[(Y_L(t) - \bar{Y}_L(t))^T(Y_L(t) - \bar{Y}_L(t))\right], \tag{A.2.13}$$

where

$$\bar{Y}_L(t) := \mathbb{E}\left[Y_L(t)|M_L^{(n)}\right], \tag{A.2.14}$$

clearly, we have

$$0 \prec \Delta_{Y|M} \preceq \Sigma_Y. \tag{A.2.15}$$

With some matrix calculations as in [33, Appendix B], one can show that

$$\Delta_{U|M} = \Sigma_U \Sigma_Y^{-1} \Delta_{Y|M} \Sigma_Y^{-1} \Sigma_U + \Sigma_U - \Sigma_U \Sigma_Y^{-1} \Sigma_U, \tag{A.2.16}$$

$$\Delta_{Y|U,M} \preceq (\Delta_{Y|M}^{-1} + \Lambda_W^{-1} - \Sigma_Y^{-1})^{-1}. \tag{A.2.17}$$

Similar to $\Delta_{U|M}$ as in (A.2.16), one can show that

$$\frac{1}{n} \sum_{\ell=1}^{L} \sum_{t=1}^{n} \mathbb{E}[(X_{\ell,t} - \hat{X}_{\ell,t})^2] = \mathrm{tr}(\Sigma_X \Sigma_Y^{-1} \Delta_{Y|M} \Sigma_Y^{-1} \Sigma_X + \Sigma_X - \Sigma_X \Sigma_Y^{-1} \Sigma_X). \tag{A.2.18}$$

Combining (A.2.18) and (A.2.4), we get

$$\mathrm{tr}(\Sigma_X \Sigma_Y^{-1} \Delta_{Y|M} \Sigma_Y^{-1} \Sigma_X + \Sigma_X - \Sigma_X \Sigma_Y^{-1} \Sigma_X) \leq L(D + \epsilon). \tag{A.2.19}$$

Moreover, let

$$\delta_\ell := \sum_{t=1}^{n} \mathbb{E}\left[(Y_\ell(t) - \tilde{Y}_\ell(t))^2 \right], \tag{A.2.20}$$

where

$$\tilde{Y}_\ell(t) := \mathbb{E}\left[Y_\ell(t) | U_\ell^n, M_\ell \right], \tag{A.2.21}$$

it is clear that

$$\delta_\ell > 0, \qquad \ell \in [1, L]. \tag{A.2.22}$$

Furthermore, since $Y_\ell^n = U_\ell^n + W_\ell^n$, $\ell \in [1, L]$, and $(U_1^n, \ldots, U_L^n)$ and $(W_1^n, \ldots, W_L^n)$ are mutually independent, we have

$$\Delta_{Y|U,W} = \mathrm{diag}^{(L)}(\delta_1, \ldots, \delta_L). \tag{A.2.23}$$

Formulating the Optimization Problem: Considering (A.2.12), (A.2.22), (A.2.23), (A.2.15), (A.2.17), (A.2.19) and letting $\epsilon \to 0$, one can show using symmetrization and convexity arguments that there exist Δ with identical diagonal entries as well as identical off-diagonal entries and δ such that

$$\frac{1}{2}\log\frac{\det(\Sigma_U)}{\det(\Delta_{U|M})}+\frac{L}{2}\log\frac{\lambda_W}{\delta}\leq R, \tag{A.2.24a}$$

$$0 \prec \Delta \preceq \Sigma_Y, \tag{A.2.24b}$$

$$0 < \delta, \tag{A.2.24c}$$

$$\mathrm{diag}^{(L)}(\delta,\ldots,\delta) \preceq (\Delta^{-1}+\Lambda_W^{-1}-\Sigma_Y^{-1})^{-1}, \tag{A.2.24d}$$

$$\mathrm{tr}(\Sigma_X\Sigma_Y^{-1}\Delta\Sigma_Y^{-1}\Sigma_X+\Sigma_X-\Sigma_X\Sigma_Y^{-1}\Sigma_X)\leq LD. \tag{A.2.24e}$$

Using the eigenvalue decomposition, we have $\Delta = \Theta\ \mathrm{diag}(\alpha,\beta,\ldots,\beta)\ \Theta^T$ for some positive α and β. So, inequality (A.2.24a) can be equivalently written as

$$\frac{1}{2}\log\frac{\lambda_Y^2}{(\lambda_Y-\lambda_W)\alpha+\lambda_Y\lambda_W}+\frac{L-1}{2}\log\frac{\gamma_Y^2}{(\gamma_Y-\lambda_W)\beta+\gamma_Y\lambda_W}+\frac{L}{2}\log\frac{\lambda_W}{\delta}\leq R, \tag{A.2.25}$$

and (A.2.24b)–(A.2.24e) reduce to the constraints (2.4.6b)-(2.4.6g). Thus, minimizing the left-hand side of (A.2.25) over (α,β,δ) subject to the constraints (2.4.6b)-(2.4.6g) and sending λ_W to $\min(\lambda_Y,\gamma_Y)$ yields the desired lower bound.

Bibliography

[1] R. Ahlswede and J. Korner. Source coding with side information and a converse for degraded broadcast channels. *IEEE Transactions on Information Theory*, 21 (6):629–637, 1975.

[2] T. Berger. Rate distortion theory: A mathematical basis for data compression. *IEEE Transactions on Communications*, 20(6):1202–1202, 1972.

[3] T. Berger. Multiterminal source coding. *The information theory approach to communications*, 1978.

[4] T. Berger and R. W. Yeung. Multiterminal source encoding with one distortion criterion. *IEEE Transactions on Information Theory*, 35(2):228–236, 1989.

[5] T. Berger, K. Housewright, J. Omura, S. Yung, and J. Wolfowitz. An upper bound on the rate distortion function for source coding with partial side information at the decoder. *IEEE Transactions on Information Theory*, 25(6): 664–666, 1979.

[6] T. Berger, Z. Zhang, and H. Viswanathan. The ceo problem [multiterminal source coding]. *IEEE Transactions on Information Theory*, 42(3):887–902, 1996.

[7] J. Chen and T. Berger. Robust distributed source coding. *IEEE Transactions on Information Theory*, 54(8):3385–3398, 2008.

[8] J. Chen and T. Berger. Successive wyner–ziv coding scheme and its application to the quadratic gaussian ceo problem. *IEEE Transactions on Information Theory*, 54(4):1586–1603, 2008.

[9] J. Chen, F. Etezadi, and A. Khisti. Generalized gaussian multiterminal source coding and probabilistic graphical models. In *2017 IEEE International Symposium on Information Theory (ISIT)*, pages 719–723. IEEE, 2017.

[10] T. Cover. A proof of the data compression theorem of slepian and wolf for ergodic sources (corresp.). *IEEE Transactions on Information Theory*, 21(2): 226–228, 1975.

[11] T. M. Cover and J. A. Thomas. *Elements of Information Theory, 2nd Ed.* Wiley, 2006.

[12] R. Dobrushin and B. Tsybakov. Information transmission with additional noise. *IRE Transactions on Information Theory*, 8(5):293–304, 1962.

[13] A. El Gamal and Y.-H. Kim. *Network information theory*. Cambridge university press, 2011.

[14] T. Flynn and R. Gray. Encoding of correlated observations. *IEEE Transactions on Information Theory*, 33(6):773–787, 1987.

[15] T. Han and K. Kobayashi. A unified achievable rate region for a general class of multiterminal source coding systems. *IEEE Transactions on Information Theory*, 26(3):277–288, 1980.

[16] A. Kaspi and T. Berger. Rate-distortion for correlated sources with partially separated encoders. *IEEE Transactions on Information Theory*, 28(6):828–840, 1982.

[17] J. Korner and K. Marton. Images of a set via two channels and their role in multi-user communication. *IEEE Transactions on Information Theory*, 23(6): 751–761, 1977.

[18] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas. Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics*, pages 1273–1282. PMLR, 2017.

[19] Y. Oohama. Gaussian multiterminal source coding. *IEEE Transactions on Information Theory*, 43(6):1912–1923, 1997.

[20] Y. Oohama. The rate-distortion function for the quadratic gaussian ceo problem. *IEEE Transactions on Information Theory*, 44(3):1057–1070, 1998.

[21] Y. Oohama. Rate-distortion theory for gaussian multiterminal source coding systems with several side informations at the decoder. *IEEE Transactions on Information Theory*, 51(7):2577–2593, 2005.

[22] Y. Oohama. Indirect and direct gaussian distributed source coding problems. *IEEE Transactions on Information Theory*, 60(12):7506–7539, 2014.

[23] V. Prabhakaran, D. Tse, and K. Ramachandran. Rate region of the quadratic gaussian ceo problem. In *International Symposium onInformation Theory, 2004. ISIT 2004. Proceedings.*, page 119. IEEE, 2004.

[24] A. Sgarro. Source coding with side information at several decoders. *IEEE Transactions on Information Theory*, 23(2):179–182, 1977.

[25] C. E. Shannon. A mathematical theory of communication. *The Bell system technical journal*, 27(3):379–423, 1948.

[26] D. Slepian and J. Wolf. Noiseless coding of correlated information sources. *IEEE Transactions on information Theory*, 19(4):471–480, 1973.

[27] S. Tung. Multiterminal source coding. *Ph. D. thesis, School of Electrical Engineering, Cornell University*, 1978.

[28] H. Viswanathan and T. Berger. The quadratic gaussian ceo problem. *IEEE Transactions on Information Theory*, 43(5):1549–1559, 1997.

[29] A. B. Wagner, S. Tavildar, and P. Viswanath. Rate region of the quadratic gaussian two-encoder source-coding problem. *IEEE Transactions on Information Theory*, 54(5):1938–1961, 2008.

[30] J. Wang and J. Chen. Vector gaussian two-terminal source coding. *IEEE transactions on information theory*, 59(6):3693–3708, 2013.

[31] J. Wang and J. Chen. Vector gaussian multiterminal source coding. *IEEE transactions on information theory*, 60(9):5533–5552, 2014.

[32] J. Wang, J. Chen, and X. Wu. On the sum rate of gaussian multiterminal source coding: New proofs and results. *IEEE Transactions on Information Theory*, 56 (8):3946–3960, 2010.

[33] Y. Wang, L. Xie, X. Zhang, and J. Chen. Robust distributed compression of symmetrically correlated gaussian sources. *IEEE Transactions on Communications*, 67(3):2343–2354, 2018.

[34] Y. Wang, L. Xie, S. Zhou, M. Wang, and J. Chen. Asymptotic rate-distortion analysis of symmetric remote gaussian source coding: Centralized encoding vs. distributed encoding. *Entropy*, 21(2):213, 2019.

[35] H. Witsenhausen and A. Wyner. A conditional entropy bound for a pair of discrete random variables. *IEEE Transactions on Information Theory*, 21(5): 493–501, 1975.

[36] A. Wyner. On source coding with side information at the decoder. *IEEE Transactions on Information Theory*, 21(3):294–300, 1975. doi: 10.1109/TIT.1975.1055374.

[37] A. Wyner and J. Ziv. The rate-distortion function for source coding with side information at the decoder. *IEEE Transactions on information Theory*, 22(1): 1–10, 1976.

[38] A. D. Wyner. The rate-distortion function for source coding with side information at the decoder\3-ii: General sources. *Information and control*, 38(1):60–80, 1978.

[39] L. Xie, X. Tu, S. Zhou, and J. Chen. Generalized gaussian multiterminal source coding in the high-resolution regime. *IEEE Transactions on Communications*, 68(6):3782–3791, 2020.

[40] H. Yamamoto and K. Itoh. Source coding theory for multiterminal communication systems with a remote source. *IEICE TRANSACTIONS (1976-1990)*, 63 (10):700–706, 1980.

[41] S. Zhou, S. Salehkalaibar, J. Qian, J. Chen, W. Shi, Y. Ge, and W. Tong. On distributed lossy coding of symmetrically correlated gaussian sources. *IEEE Transactions on Communications*, 70(12):7957–7972, 2022. doi: 10.1109/TCOMM.2022.3218822.