Rate-Distortion Regions for Successively Structured Multiterminal Source Coding Schemes

Hamid Behroozi

A Thesis
in
The Department
of
Electrical and Computer Engineering

Presented in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy at
Concordia University
Montréal, Québec, Canada

June 2007

Your file Votre référence
ISBN: 978-0-494-31141-7
Our file Notre référence
ISBN: 978-0-494-31141-7

ABSTRACT

Rate-Distortion Regions for Successively Structured
Multiterminal Source Coding Schemes

Hamid Behroozi, PhD.
Concordia University, 2007

Multiterminal source coding refers to separate encoding and joint decoding of multiple correlated sources. Joint decoding requires all the messages to be decoded simultaneously which is exponentially more complex than a sequence of single-message decodings. Inspired by previous work on successive coding strategy, which is based on successive decoding structure, we apply the successive Wyner-Ziv coding to different schemes of multiterminal source coding problem. We address the problem from an information theoretic perspective and determine the rate region for three different multiterminal coding schemes: Gaussian CEO problem, *1-helper* problem, and *2-terminal* source coding problem. We prove that the optimal sum-rate distortion performance for the CEO problem is achievable using the successive coding strategy which is essentially a low complexity approach for obtaining a prescribed distortion. We show that if the sum-rate tends to infinity for a finite number of agents (sensors), the optimal rate allocation strategy assigns equal rates to all agents. The same result is obtained when the number of agents tends to infinity while the sum-rate is finite. Then, we consider *1-helper* source coding scheme where one source provides partial side information to the decoder to help the reconstruction of the other source. Our results show that the successive coding strategy is an optimal strategy in this scheme in the sense of achieving the rate-distortion function. For the *2-terminal* source coding problem, we develop connections between source encoding and data fusion steps and prove that the whole rate-distortion region is achievable using the successive coding strategy. Comparing the performance of the sequential coding with the performance of the successive coding, we show that there is no sum-rate loss when the side information is not available at the encoder. This result is of special interest in some applications such as video coding where there are processing and storage constraints at the encoder. Based on the successive coding strategy, we provide an achievable rate-distortion region for the *m-terminal* source coding.

We also consider a distributed network, modeled by CEO problem with Gaussian multiple access channel (MAC), where L noisy observations of a memoryless Gaussian source are transmitted through an additive white Gaussian MAC to a decoder. The decoder wishes to reconstruct the main source with an average distortion D at the smallest possible power consumption in the communication link. Our goal is to characterize the power-distortion region achievable by any coding strategy regardless of delay and complexity. We obtain a necessary condition for achievability of all power-distortion tuples $(P_1, P_2, ..., P_L, D)$. Also, analyzing the uncoded transmission scheme provides a sufficient condition for achievability of $(P_1, P_2, ..., P_L, D)$. Then, we consider a symmetric case of the problem where the observations of agents have the same noise level and the transmitting signals are subject to the same average power constraint. We show that in this case the necessary and sufficient conditions coincide and give the optimal power-distortion region. Therefore, in the symmetric case of Gaussian CEO problem uncoded transmission over a Gaussian MAC performs optimally for any finite number of agents.

Dedicated to

my beloved wife, *Sepideh*

my mother, *Fatemeh*

my brother, *Vahid*,

my sisters, *Neda, Negar, Negin*

and in loving memory of my father, *Ali*

ACKNOWLEDGEMENTS

Before all, I must thank God for all the opportunities that He has put in front of me. I wish to thank my supervisor, Professor M. Reza Soleymani, for giving me the opportunity to work in his group, for his support and encouragement throughout my research and preparation of this thesis. He entrusted me with the creative freedom early on to explore research ideas, for which I am grateful. Without his guidance, I would not have been able to reach this milestone.

I would like to thank all the members of my Ph.D. defense committee: Prof. Amir K. Khandani, Prof. Yousef R. Shayan, Prof. Walaa Hamouda, and Prof. Sudhir P. Mudur and the committee chair, Prof. Thomas G. Fevens. I appreciate their supportive efforts, valuable comments, and all the time they spent to read this thesis. My special thanks goes to the external examiner, Prof. Amir K. Khandani, of University of Waterloo, for his technical insight and the suggestions he made for the future research. I would like to thank Prof. Eugene I. Plotkin for his constructive advices that improved the quality of my work.

I also want to thank all current and previous members of the Wireless and Satellite Communications Laboratory for their help and comments and for creating a collaborative environment for research. Specially I would like to thank my lifelong friend, Dr. Javad Haghighat, for many fruitful discussions we had during this work; he has taught me so much and have been the greatest friend one could have.

I would like to thank FQRNT (le Fonds Québécois de la Recherche sur la Nature et les Technologies), the Québec Government, Concordia University, Hydro Quebec, and NSERC (Natural Sciences and Engineering Research Council of Canada) for their financial support, scholarships and awards.

Many thanks also go to my beloved mother, Fatemeh, my late father, Ali, my brother, Vahid, and my dear sisters, Neda, Negar and Negin. Their tremendous support and encouragement throughout the years from my remote hometown inspire me to overcome every difficult time. I would not be where I am today without their love, help, and unlimited sacrifices. I only hope I have lived up to their dreams and expectations and made all their trials and supports worthwhile. I hope this thesis would fulfill part of my father's wish for me: to be one of the best researchers in my field of interest.

Last but certainly not the least, I would like to express my sincere thanks to my wonderful wife, Sepideh, the love of my life. Her unconditional love, support and understanding helped me through this long journey. I wish this thesis could be my precious gift to her.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SYMBOLS

$R(D)$	Rate-distortion function
$C(P)$	Channel capacity
$h(X)$	Differential entropy of continuous random variable X
$H(X)$	Entropy of random variable X
$H(X\|Y)$	Conditional entropy of X given Y
$H(X,Y)$	Joint entropy of random variables X and Y
$I(X;Y)$	Mutual information between random variables X and Y
$\overline{R}$	Total rate (sum-rate)
R_i	Communication rate at ith agent (sensor node)
L	Total number of agents (sensors)
D	Expected value of distortion
D_i	Expected value of distortion after ith decoder
D_{min}	Minimum achievable distortion
D_0^*	A lower bound on the minimum achievable distortion
σ_X^2	Variance of the source
$\|A\|$	Determinant of the matrix A
K_S	Covariance matrix of the random variable S
$\mathbf{X}$	Random vector
X	Random variable
$\mathbf{x}$	Sample vector
x	Sample value
N_i	Noise variance in the observation of the ith agent
P_i	Average power constraint for the transmitted signal of the ith agent

LIST OF ABBREVIATIONS

CEO	Chief Executive (Estimation) Officer
FC	Fusion Center
SNR	Signal-to-Noise Ratio
SQNR	Signal-to-Quantization Noise Ratio
DSC	Distributed Source Coding
SWC	Slepian-Wolf Coding
AWGN	Additive White Gaussian Noise
i.i.d.	Independent and Identically Distributted
WSN	Wireless Sensor Network
pdf	Probability Density Function
MSE	Mean-Squared Error
MMSE	Minimum Mean-Squared Error
CRT	Communication Rate Throughput
OPTA	Optimum Performance Theoretically Attainable

Chapter 1

Introduction

Multiterminal source coding refers to separate encoding and joint decoding of multiple correlated sources. Joint decoding requires all messages to be decoded simultaneously which is exponentially more complex than a sequence of single-message decodings. We consider successive coding as a low complexity coding scheme for multiterminal source coding. It allows us to analyze the multiterminal schemes with finite number of sources. In this thesis, we analyze the rate-distortion function of three different multiterminal schemes: multiterminal noisy (remote) source coding scheme called the CEO problem, *1-helper* source coding scheme and *2-terminal* source coding problem.

In this chapter, we first give a background on classical source coding. Then different multiterminal source coding schemes will be described. The organization of the thesis concludes this chapter.

1.1 Classical Source Coding

Shannon initiated the studies on source coding in his original paper [1]. He stated the source coding theorem for lossless data compression, which investigates the question of how short a description is possible if we want perfect reconstruction of a given amount of information. He also introduced the idea of lossy data compression which considers the shortest possible description to represent information within some prescribed distortion. Since the source does not need to be reconstructed perfectly in this case, such a source coding scheme is called a "lossy" scheme. Shannon investigated lossy data compression, which is also called source coding with a fidelity criterion, more precisely in [2]. Fidelity is measured by a fidelity criterion such as average distortion. The objective of this problem is to determine the minimum description rate required to reconstruct a source with respect to a target distortion. This is called the rate-distortion function. Since his work, many research works have been done on the subject of rate-distortion theory for point-to-point communication systems (e.g., [3, 4, 5, 6]), much of which has been reviewed in the survey paper of Berger and Gibson [7].

But new issues have been addressed beyond the classical source coding problem.

1.2 Multiterminal Source Coding

1.2.1 Source Coding with Side Information

In the point-to-point communication there is only a single source of information at encoder and decoder, i.e., the message at the encoder, and the received signal at the decoder. If any type of useful information is added to encoder or decoder's information, it would be called "side" information [8]. Side information is often available to improve the rate-distortion performance of the system. One simple example is when an observer records a signal X to be conveyed to a receiver who has an old version of the signal, Y which is correlated with X. Wyner and Ziv determine the rate-distortion function of the source coding with uncoded side information at the decoder [9]. Wyner extends the result for the Gaussian case with mean squared-error distortion measure [10]. Today, **Wyner-Ziv coding** refers to lossy source coding with decoder side information.

1.2.2 Noisy (Remote) Source Coding

Noisy source coding refers to coding of noisy (or imperfect) observations. For instance, consider a distributed monitoring system or a distributed video sensor network. In such a system, the encoder may not have direct access to the source of interest. Instead, only a corrupted or noisy version of the source is available at the encoder. This scheme is also termed as the **remote source coding** problem and is studied in [11]. Wolf and Ziv consider a version of this scheme in which the source is corrupted by an additive noise [12]. Recent extensions to the non-Gaussian sources appear in [13].

1.2.3 Noisy Source Coding with Side Information

This scheme is the generalization of two previous schemes. The generalized Wyner-Ziv source coding for noisy encoder observations, which is also called the remote source coding with side information, is appeared in [14]. This problem is also known as the noisy Wyner-Ziv coding, i.e., lossy coding of noisy observations with side information available at the decoder and not at the encoder. Similar rate-distortion analysis is presented in [15], and recently in [8].

1.2.4 Source Coding for Multiterminal Scenarios

Information sources in many networks such as wireless sensor networks are distributed. Since the sources are correlated, the correlation should be exploited to avoid redundant transmission. This requires the theory of distributed coding of correlated sources. In other words, many wireless networks are working under some power constraints at their nodes. On the other hand, the bit rate directly impacts transmission power consumption at a node. Therefore, determining minimum rates at which these correlated, physically separated, sources can be compressed, given that they are going to be reconstructed at a joint decoder, is critical. Moreover, in some real situations such as distributed monitoring system, multiple encoders only have partial access to the sources of interest. In fact, there are many schemes or scenarios in practice whose rate-distortion regions are not completely characterized.

In general, Multiterminal Source Coding [16] refers to the compression of multiple correlated sources that cannot communicate with each other (distributed coding). The outputs of these sources will be sent to a fusion center (FC) or a Chief Executive (Estimation) Officer (CEO) (e.g. the base station) for joint decoding. In this thesis, we focus on lossy source coding and investigate the rate-distortion regions of multiterminal source coding schemes.

Slepian and Wolf establish information-theoretic bounds for distributed lossless coding [17]. Wyner and Ziv extend the results for lossy source coding with decoder side information [9]. Over the last 30 years, significant effort has been made on finding a complete characterization of multiterminal rate-distortion region. But even concrete examples of the problem are hard to analyze. For instance, the complete characterization of the rate region for *2-terminal* source coding for the Gaussian case has been found recently [18]. Nevertheless, today multiterminal source coding is still of special interest; not only because it is an open problem of information theory [19], but also because of its application in wireless communication systems. In fact, the increasing attention given to new applications such as wireless video networks or distributed sensor networks is a reason for new interests in multiterminal source coding schemes. A wireless sensor network (WSN) is a particular type of Ad Hoc network where the nodes have severe energy constraints. It consists of a large number of sensors spread across a geographical area for information gathering. Each sensor may have capabilities of wireless communications, signal processing of its observation or measurement and networking the data, thus, it must be self-configuring. These networks are widely used in many applications and have great capabilities for consumer, military, and civic applications. We can consider sensor networks that forecast the weather, control the traffic, or provide security in shopping malls or other places. They also can be deployed rapidly and be used efficiently at the sites of accidents such as collapse of a building to detect and locate trapped survivors, or to track natural gas and toxic substances [20]. The benefit of these distributed networks is their ease of deployment since they do not need any infrastructure which is required in wireless communications using base stations. Distributed sensor networks are typically operated under constraints on system resources such as *power* and *bandwidth.* Since the bit rate directly impacts transmission power consumption at a node, an efficient high ratio compression is the main requirement of distributed wireless sensor networks.

In practice, there are many other problems that are closely related to multiterminal source coding problem. The biological and machine pattern recognition problem from life sciences, recently considered by Westover and O'Sullivan [21, 22], is closely related to multiterminal source coding. Motivated by the problem of communication over relay networks, Gastpar extends Wyner-Ziv to the case of multiple sources [23]. Another reason that the multiterminal source coding has captured a lot of attention from the research community is the duality relationship among many of these problems such as duality between source coding and information embedding or data hiding, duality between rate-distortion and channel capacity, etc [24, 25, 26, 27, 28].

Although the theory of multiterminal source coding started more than 30 years ago, a complete characterization of its rate-distortion region still remains unknown. Even concrete examples of this problem are hard to analyze. For instance, the whole rate-distortion region of the *2-terminal* source coding scheme for Gaussian sources with mean-squared error (MSE) distortion has been recently characterized. We consider successive coding as a low complexity coding scheme that allows us to analyze the multiterminal schemes with finite number of sources.

1.3 Successive Coding Strategy

Previous researches on coding of multiterminal schemes are based on joint decoding of all messages. Joint decoding requires all the messages to be decoded simultaneously which is exponentially more complex than a sequence of single-message decodings. Inspired by previous work on successive coding strategy, we apply the successive Wyner-Ziv coding to the multiterminal source coding schemes.

This is a decentralized strategy because at each stage, only the knowledge sharing between each encoder and decoder is needed. When an encoder encodes a message, it considers two things. First, its observations and second, its statistical knowledge about the messages that the decoder has already received from other nodes in the network. The latter is known as the "decoder side information" in the sense of Wyner and Ziv [9]. At the decoder, messages from sources are decoded sequentially in order to increase the fidelity of estimation at each decoding step.

The successive Wyner-Ziv coding strategy allows us to derive an achievable rate region for multiterminal schemes. Its successive structuring provides flexibility to deal with distributed signal processing. From the perspective of robustness, this scheme performs well, i.e., if we timeshare among different successive decoding schemes, no matter which node fails in the network, the decoder can obtain a nontrivial estimate of the source. Finally, by applying the successive coding the available practical Wyner-Ziv coding techniques are applicable to more general distributed and multiterminal source coding problems. We address the problem from an information theoretic perspective and determine the rate region for three different multiterminal coding schemes: Gaussian CEO problem, *1-helper* source coding scheme and *2-terminal* source coding problem.

First, we consider multiterminal noisy source coding problem, called the CEO problem, introduced in [29]. In this problem, the CEO is interested in an underlying source. L agents (sensors) observe independently noisy versions of the source signal X. Agents separately communicate information about their observations to the FC through rate-constrained noiseless channels without collaborating. The FC desires to form an optimal estimate of X based on information received from the agents. The objective of the CEO problem is to determine the minimum achievable distortion under a sum-rate constraint. By sum-rate, we mean the total rate at which the agents may communicate information about their observations to the FC. The special case of Gaussian source and noise statistics with MSE distortion is called the quadratic Gaussian CEO problem, first introduced in [30]. For this case, the rate region is known [31, 32, 33, 34]. We apply the successive coding strategy and show that it is an optimal strategy in the Gaussian CEO problem in the sense of achieving the sum-rate distortion function of the problem.

Then, we evaluate the performance of the successive coding strategy for the problem of multiterminal lossy coding of correlated Gaussian sources. We first consider the *1-helper* source coding problem where one source provides partial side information to the decoder to help reconstruction of the other source signal. Our results show that the successive coding strategy is an optimal strategy in sense of achieving the rate-distortion function of the *1-helper* problem. In the *2-terminal* source coding scheme, we are interested in reconstructing both sources at the decoder. We prove that successive coding can achieve the whole rate-region of the problem. Finally, we provide an achievable rate-distortion region for the *m-helper* problem in which multiple correlated sources transmit their information to a FC for further processing. One of these sources is the source of interest, which is called the *primary* source, but other sources act as helpers by sending correlated information (which is called side information) to help reproduction of the primary source signal [35]. We also provide an inner bound for the rate-region of the *m-terminal* source coding problem, where all of the m sources are to be reconstructed at the FC.

In the last part of this research, we consider Gaussian CEO problem with Gaussian multiple access channel (MAC) where L noisy observations of a memoryless Gaussian source are transmitted through an additive white Gaussian MAC to a single FC. The encoders are distributed and cannot cooperate to exploit their correlation. Each encoder is subject to a transmission cost constraint. This constraint comes from the restrictions on the resources such as power and bandwidth that are available at each agent (sensor node). Since the final goal is to reconstruct the main source to within some prescribed distortion level at the smallest cost in the communication link, finding a suitable coding strategy to achieve this goal is critical. Our interest lies in deriving the achievable power-distortion region, while the fidelity of estimation at the FC is measured by the MSE distortion. We first obtain a necessary condition for achievability of all power-distortion tuples $(P_1, P_2, ..., P_L, D)$. Our proof is based on using the data processing inequality and analyzing the remote source coding scenario, where the agents observations are given to one common encoder. On the other hand, it is shown [36, 37] that for a point-to-point transmission of a single Gaussian source through an additive white Gaussian noise (AWGN) channel, if the channel bandwidth is equal to the source bandwidth, a simple uncoded transmission achieves the optimal power-distortion tradeoff. In [38] the Gaussian CEO problem in a symmetric environment is considered, where the agents observations have the same noise level and the transmitting terminals are subject to the same average power constraint. The authors show that as the number of agents tends to infinity, uncoded transmission achieves the smallest possible distortion. In the recent work of Lapidoth *et al.* [39] sending a memoryless Bi-variate Gaussian source over an interfering MAC is evaluated. They have shown that in the symmetric case, uncoded transmission is optimal below a threshold signal-to-noise ratio (SNR). Motivated by this result, we analyze the performance of the uncoded transmission scheme in the Gaussian CEO problem. Analyzing the uncoded transmission scheme provides a sufficient condition for achievability of $(P_1, P_2, ..., P_L, D)$. We show that, in the symmetric case, these necessary and sufficient conditions coincide and give the optimal power-distortion tradeoff. This is the same result as the recent work of Gastpar [40] which is obtained independently. It shows that in the symmetric case of Gaussian CEO problem uncoded transmission over a Gaussian MAC performs optimally for any finite number of agents.

1.4 Thesis Outline

The rest of the thesis is organized as follows:

Chapter 2, Background, is devoted to the literature review and background of source coding, distributed source coding and rate-distortion theory. We focus on Gaussian rate-distortion and present different multiterminal coding schemes that we are going to characterize their rate-distortion regions based on successive coding strategy. We also present a comprehensive background on the concept of successive coding in this chapter.

In **Chapter 3, Successively Structured Gaussian CEO Problem,** we apply the successive coding strategy to the Gaussian CEO problem and derive the optimal rate allocation scheme to achieve the minimum distortion under a sum-rate constraint. One of the main contributions of this work is to show that the sum-rate distortion function of the Gaussian CEO problem can be achieved by a sequence of successively structured Wyner-Ziv codes. Based on these results, we characterize the whole rate region of the Gaussian CEO problem. We also demonstrate that if the sum-rate, $\overline{R}$, grows to infinity with a finite number of agents or if the number of agents, L, tends to infinity under a sum-rate constraint, a sequence of successively structured Wyner-Ziv codes with equal communication rates at agents $\left(R_1 = R_2 = \cdots = R_L = \overline{R}/L\right)$ converges to the rate-distortion function. Hence, we can simplify rate allocation problem in a general parallel network with L agents by assigning equal rates to agents, provided that the average rate per agent is either very large or is very small. The similar result for the large number of agents is also presented in [8]. In the last part of Chapter 3, a solution for the communication throughput of a Gaussian relay network is presented.

In **Chapter 4, Successively Structured Gaussian Multiterminal Source Coding Schemes,** we evaluate the performance of the successive coding strategy for the multiterminal lossy coding of correlated Gaussian sources. We consider the *m-helper* problem for the special case of $m = 1$ where one source provides partial side information to the decoder to help reconstruction of the other source. Our results show that the successive coding strategy, which is inherently a low complexity coding scheme of obtaining a prescribed distortion, is an optimal strategy in the sense of achieving the rate-distortion function of the *1-helper* problem. By developing connections between source encoding and data fusion step, it is shown that the whole rate distortion region for the *2-terminal* source coding problem is achievable using the successive coding strategy. Comparing the performance of the sequential coding with the performance of the successive coding, we show that there is no sum-rate loss when the side information is not available at the encoder. We provide an achievable rate-distortion region for the *m-helper* problem and also derive an inner bound for the rate-region of the *m-terminal* source coding scheme.

In **Chapter 5, Gaussian CEO Problem with Gaussian Multiple Access Channel,** we consider a distributed network, modeled by CEO problem with Gaussian MAC, where L distributed agents transmit noisy observations of a source through a Gaussian MAC to a common destination. The goal is to characterize the optimal tradeoff between the transmission cost, i.e., the power vector $\boldsymbol{P} = (P_1, P_2, ..., P_L)$, and the average estimation distortion, D. We present necessary and sufficient conditions for achievability of $(L+1)-$tuples $(P_1, P_2, ..., P_L, D)$. In the symmetric case these conditions agree and provide the optimal power-distortion tradeoff. We show that in the symmetric case for any finite L, uncoded transmission performs optimally and achieves the smallest possible distortion.

Chapter 6, Conclusion, summarizes the contributions of this thesis and gives suggestions for future work.

Chapter 2

Background

In this chapter we review the rate-distortion theory for lossy source coding. We consider a general model for multiterminal (distributed) source coding and focus on Wyner-Ziv coding and its generalized version to noisy encoder observations, called remote source coding with side information. Then we present three multiterminal schemes that we are going to obtain their rate-distortion functions based on the successive coding strategy. We present the literature review for all three schemes and provide the existing rate-distortion functions. Then, we review the successive coding strategy as an enabling technique that allows us to analyze the multiterminal coding schemes. A summary of the chapter is presented at the end.

2.1 Rate-Distortion Theory

Rate-distortion theory was introduced by Claude Shannon in his original work [1]. It gives the theoretical bounds for how much compression can be achieved using lossy data compression methods. Results of the rate-distortion theory are obtained without consideration of a specific coding method. It addresses the problem of determining the minimum amount of information rate R that should be communicated over a channel, so that the source can be reconstructed at the receiver without exceeding a given distortion D. In the rate-distortion theory, the rate is usually understood as the number of bits per source symbol to be transmitted, and the distortion is usually defined as the variance of the difference between the input and the output signal, i.e., the MSE distortion.

2.1.1 Rate-Distortion Function

Let X be a source that produces a sequence of symbols $X^n = X_1, X_2, ..., X_n$ in an independent and identically distributed (*i.i.d.*) fashion according to the distribution $p(x)$ for $x \in \mathcal{X}$. The source encoder represents each n-symbol sequence $X^n \in \mathcal{X}^n$ from the source with an index $\varphi(X^n) \in \{1, 2, ..., 2^{nR}\}$. The source decoder represents X^n by an estimate $\hat{X}^n \in \hat{\mathcal{X}}^n$, which is also called the reproduced (or reconstructed) sequence. The source coding scheme is illustrated in Figure 2.1. Since we have a compression, the rate always holds $R < 1$.

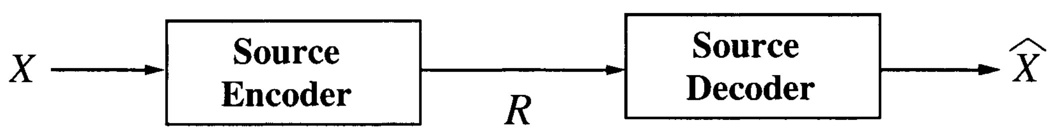

Figure 2.1: Source coding scheme.

A **distortion measure** (or function) is a mapping from pairs of source symbols and reproduced symbols to the positive real line (the set of non-negative real numbers). In fact, the *single-letter* distortion is a measure of the cost of representing the symbol X by the symbol $\hat{X}$. Thus, if X is the source symbol and $\hat{X}$ is its representation, the distortion $d(X, \hat{X})$ is a positive real number, i.e., $d : \mathcal{X} \times \hat{\mathcal{X}} \longrightarrow [0, \infty)$. In most cases, the reproduction alphabet $\hat{\mathcal{X}}$ is the same as the source alphabet $\mathcal{X}$.

The distortion between sequences is defined by the average distortion per source symbol, i.e., $d(X^n, \hat{X}^n) = \frac{1}{n}\sum_{i=1}^{n} d(X_i, \widehat{X}_i)$. The fidelity criterion is the **total average distortion** which is defined by $E\left[d(X^n, \hat{X}^n)\right] = E\left[\frac{1}{n}\sum_{i=1}^{n} d(X_i, \widehat{X}_i)\right]$ where the expectation is with respect to the probability distribution on $\mathcal{X}$.

A $(2^{nR}, n)$ rate-distortion code for a source X under distortion measure d is defined by encoder and decode functions (φ, ψ) such that

$$
\begin{aligned}
\varphi &: \mathcal{X}^n \longrightarrow \{1, 2, ..., 2^{nR}\} \\
\psi &: \{1, 2, ..., 2^{nR}\} \longrightarrow \hat{\mathcal{X}}^n \\
\triangle &= \tfrac{1}{n} \textstyle\sum_{i=1}^{n} Ed(X_i, \widehat{X}_i)
\end{aligned} \tag{2.1}
$$

where $\widehat{X}_i$ is the ith component of $\widehat{X}^n = \psi(\varphi(X^n))$. A rate-distortion pair (R, D) is achievable if a sequence of $(2^{nR}, n)$ rate-distortion codes (φ, ψ) exists such that we have $\lim_{n \longrightarrow \infty} Ed(X^n, \psi(\varphi(X^n))) \leq D$, i.e., $\lim_{n \longrightarrow \infty} \triangle \leq D$. The **rate-distortion region** $\mathcal{R}$ for a source X is the closure of the set of achievable pairs (R, D), where closure of a set S is the union of S with the set of limit points of S. Now we can define a function to describe the boundary of the rate-distortion region: The **rate-distortion function** $R(D)$ is the infimum of rates R such that the pair (R, D) is in the rate-distortion region $\mathcal{R}$ of the source for a given average distortion D, i.e.,

$$
R(D) = \inf_{D} \{R : (R, D) \in \mathcal{R}\}. \tag{2.2}
$$

Shannon [1] gives an information-theoretic characterization of the rate-distortion function:

The information rate-distortion function $R^{(I)}(D)$ for a source X under distortion measure $d(X, \hat{X})$ is:

$$
R^{(I)}(D) = \inf_{\hat{X} \in \mathcal{M}_X(D)} I(X; \hat{X}), \tag{2.3}
$$

where $\mathcal{M}_X(D)$ is the closure of the set of all random variables $\hat{X}$ described by a test channel $p(\hat{x} \mid x)$ such that $Ed(X, \hat{X}) \leq D$. In other words,

$$
R^{(I)}(D) = \inf_{p(\hat{x}|x): \sum_{x,\hat{x}} p(x) p(\hat{x}|x) d(x,\hat{x}) \leq D} I(X; \hat{X}). \tag{2.4}
$$

It is shown that the rate-distortion function of an *i.i.d.* source X with distribution $p(x)$ and bounded distortion function $d(X, \hat{X})$ equals the corresponding information rate-distortion function. In this work, we have considered Gaussian sources with MSE distortion measure. The rate-distortion function for a Gaussian source $\mathcal{N}(0, \sigma^2)$ under a MSE distortion function is:

$$R(D) = \begin{cases} \frac{1}{2}\log\left(\frac{\sigma_X^2}{D}\right) & 0 \le D \le \sigma^2 \\ 0 & D > \sigma^2 \end{cases} \quad (2.5)$$

When one moves beyond the classical source coding problem, new issues arise: source coding with side information, noisy (remote) source coding, Multiterminal (or distributed) source coding, etc.

2.2 Multiterminal (Distributed) Source Coding

Multiterminal source coding or distributed data compression refers to separate lossy encoding and joint decoding of multiple correlated sources. In this more general scenario, L sources are observed at L separate encoders. The goal it to jointly estimate all observations, rather than some underlying source. Rate-distortion theory for multiterminal coding problem was first studied by Wyner and Ziv. The analysis of this scenario, ranging from one source with coded side information to L sources has been studied during years [9, 41, 42, 43, 31, 34, 32, 33].

Consider a communication system with two correlated sources, X and Y (see Figure 2.2). Assume that X and Y are not co-located or cannot cooperate to directly exploit the correlation. Therefore, the sources are encoded independently or are "distributed". On the other hand, the receiver can see both encoded sequences and can perform joint decoding.

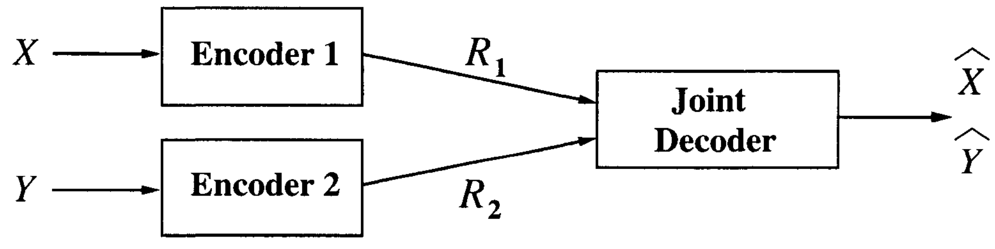

Figure 2.2: Distributed compression of two correlated sources, X and Y . The decoder jointly decodes X and Y.

2.2.1 Lossless Distributed Source Coding

In lossless source coding, we want to represent the information perfectly. Assume that $\{(x_i, y_i)\}_{i=1}^{\infty}$ be a sequence of *i.i.d.* drawings of a pair of correlated discrete random variables X and Y. Shannon's source coding theory states that if we want to encode X and Y together, a rate given by joint entropy of $H(X, Y)$ is sufficient to have lossless compression, i.e., after decompression we obtain $\hat{X} = X$ and $\hat{Y} = Y$. To reach this efficient compression rate, we can compress Y into $H(Y)$ bits per sample at the encoder and transmit it to the decoder. Based on this knowledge (i.e., complete knowledge of Y available at the encoder and decoder) we can compress X into $H(X \mid Y)$ bits per sample. But what should we do if we have a distributed nature like in the sensor network? What is the best achievable rate if X and Y must be separately encoded?

Slepian and Wolf [17] showed that the joint entropy $R = H(X, Y)$ is still achievable as long as the individual rate for each source is at least its conditional entropy given the other source. Therefore, based on Slepian-Wolf Coding (SWC), the encoder of X by just knowing the joint distribution of X and Y can perform as well as the encoder with complete knowledge of Y. There is no loss of coding efficiency with separate encoding compared with joint encoding as long as joint decoding is performed.

2.2.2 Lossy Source Coding with Side Information

In the lossy compression, the decoder produces the source estimate to an acceptable average distortion. Wyner-Ziv coding [9] refers to lossy compression with decoder side information. The Wyner-Ziv coding scheme is shown in Fig. 2.3. We need to encode X under the constraint that the average distortion between X and the estimated version $\widehat{X}$ is less than or equal D, assuming that the side information Y is available at the decoder but not at the encoder. The important fact about Wyner-Ziv coding is that it usually suffers rate loss when compared to lossy coding of X when the side information Y is available at both the encoder and the decoder. One exception is when X and Y are jointly Gaussian with MSE distortion measure. There is no rate loss with Wyner-Ziv coding in this case, which is of special interest in practice such as video sensor networks since many image and video sources (after mean abstraction) can be modeled as jointly Gaussian [44].

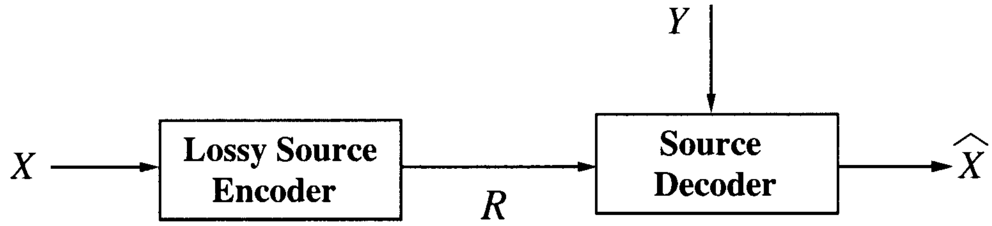

Figure 2.3: Wyner-Ziv coding scheme.

2.2.2.1 Wyner-Ziv coding for Quadratic Gaussian case

Assume X and Y are zero mean and stationary jointly Gaussian memoryless sources and the distortion metric is MSE, i.e., $d(x,\widehat{x}) = (x-\widehat{x})^2$. Let the covariance matrix of X and Y be $\Lambda = \begin{bmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix}$ with correlation coefficient ρ such that $|\rho| < 1$. If we represent the Wyner-Ziv rate-distortion function by $R_{wz}^*(D)$ and conditional rate-distortion function, i.e., rate-distortion function when both encoder and decoder have access to the side information by $R_{X|Y}(D)$, where, D is the upper limit of average distortion, then

$$R_{wz}^*(D) = R_{X|Y}(D) = \frac{1}{2}\log^+\left[\frac{\sigma_{X|Y}^2}{D}\right] = \frac{1}{2}\log^+\left[\frac{\sigma_X^2(1-\rho^2)}{D}\right], \tag{2.6}$$

where $\log^+ x = \max\{\log_2 x, 0\}$, $\rho = \frac{E[XY]}{\sigma_X\sigma_Y}$ and $\sigma_{X|Y}^2 = E\left[(X - E[X \mid Y])^2\right] = E[X^2] - \frac{E^2[XY]}{E[Y^2]}$. So there is no rate loss with Wyner-Ziv coding in this quadratic Gaussian case [10].

2.2.3 Noisy Source Coding with Side Information

Fig. 2.4 shows the noisy (remote) source coding scheme with decoder side information which is also called the generalized Wyner-Ziv source coding for noisy encoder observations [45]. The length-n *i.i.d.* source vector $\mathbf{X}$ is observed via two memoryless channel as $\mathbf{Y}_e$ at the encoder, and as $\mathbf{Y}_d$ at the decoder. Based on its observation, $\mathbf{Y}_e$, the encoder transmits a message m to the decoder over a rate-constrained noiseless channel. The decoder produces source estimate $\widehat{\mathbf{X}}$ as a function of m and its side information $\mathbf{Y}_d$. The goal is to minimize the transmission rate needed to guarantee that the source decoder can approximate the source to within average distortion D, i.e., $E\{d(\mathbf{X}, \widehat{\mathbf{X}})\} \leq D$ provided that side information is available at the decoder. Because of imperfect encoder observations, $\mathbf{X}$ is not uniquely determined from the encoder observation $\mathbf{Y}_e$. Thus the problem is different from Wyner-Ziv source coding.

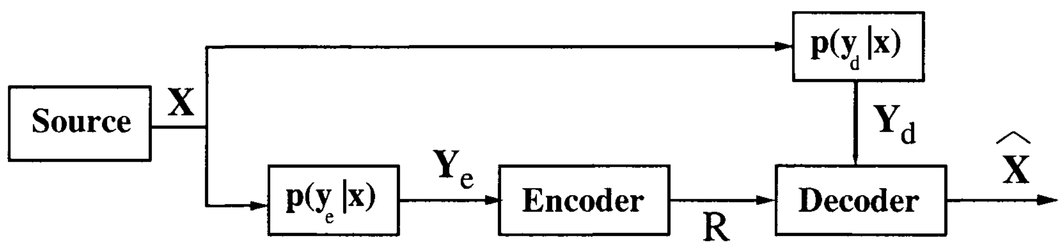

Figure 2.4: Wyner-Ziv source coding with noisy encoder observation.

2.2.3.1 Noisy Wyner-Ziv Coding for Quadratic Gaussian case

Assume $\mathbf{Y}_e$ and $\mathbf{Y}_d$ are observations of the *i.i.d.* Gaussian source $\mathbf{X}$ through AWGN channels, i.e., $Y_{e,i} = X_i + V_{e,i}$ and $Y_{d,i} = X_i + V_{d,i}$, where $V_{e,i} \sim N(0, N_e)$ and $V_{d,i} \sim N(0, N_d)$. We assume that two noises are independent of each other and of the source. It is shown that [14, 15, 45] the rate-distortion function for this problem is as follows:

$$R(D) = \frac{1}{2} \log \left[\frac{\sigma^2_{X|Y_d} - \sigma^2_{X|Y_e,Y_d}}{D - \sigma^2_{X|Y_e,Y_d}} \right] \tag{2.7}$$

where $\sigma^2_{X|Y_e,Y_d} \leq D \leq \sigma^2_{X|Y_d}$. The conditional variance $\sigma^2_{X|Y_d}$ represents the minimum mean-squared error (MMSE) in $\mathbf{X}$ given the decoder observation $\mathbf{Y}_d$, and $\sigma^2_{X|Y_e,Y_d}$ is similarly defined while both encoder and decoder observations, i.e., $\mathbf{Y}_e$ and $\mathbf{Y}_d$ are available for the estimation of $\mathbf{X}$.

But there are many different scenarios for distributed source coding. In those cases, a general expression for the rate-distortion function or rate-region is unknown. We consider successive coding strategy as an enabling technique that allows us to analyze the multiterminal schemes. Its successive structure is suitable for networks because the nature of data in networks is distributed. Although this simple strategy may have suboptimal performance in general, we prove that in three multiterminal schemes it is optimal in the sense of achieving the rate-distortion function.

We consider three extensions of the source coding problem to the distributed setting, called the CEO problem, *1-helper* problem and the *2-terminal* source coding problem. All these problems can be considered as examples of the general multiterminal source coding problem, illustrated in Fig. 2.5.

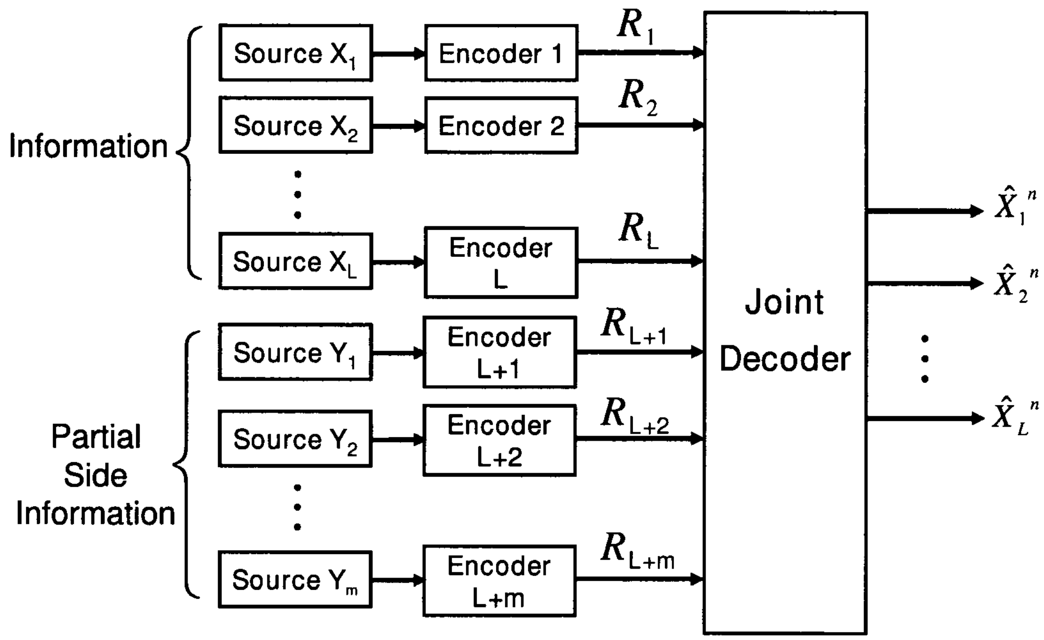

Figure 2.5: General model for multiterminal source coding problem. Each encoder separately encodes its message while the joint decoder, based on all received messages, obtains estimates of L primary sources, $X_1, X_2, ..., X_L$. Auxiliary sources $Y_1, Y_2, ..., Y_m$, which are also called helpers, provide partial side information to help the decoder to reconstruct primary sources.

$L+m$ encoders must encode $L+m$ correlated sources separately, each one subject to a rate constraint, so that a joint decoder (FC) with access to codes from all encoders can obtain estimate of L sources, where m other sources play the role of partial side information to help the decoder to reproduce the transmitted sequence of the main L sources within the prescribed average distortions.

2.2.4 Extensions of Noisy Wyner-Ziv Source Coding

2.2.4.1 CEO problem

The CEO problem, which is in fact the multiterminal noisy source coding problem, is an abstract model for remote monitoring (or sensing) and distributed compression in wireless networks. In this problem, a chief executive officer (CEO) is interested in an underlying source. L agents observe independently corrupted versions of the source. Each agent has a noiseless, rate-constrained channel to the CEO. Without collaborating, the agents must transmit messages across these channels to the CEO so that the CEO can reconstruct an estimate of the source to within some degree of fidelity. The scenario of the CEO problem is shown in Fig. 2.6.

The special case of Gaussian source and noise statistics with MSE distortion is called the quadratic Gaussian CEO problem and is introduced in [30]. Since each observation is the source corrupted by an additive white Gaussian noise (AWGN), it is also called AWGN CEO problem [13]. For this version of the CEO problem the rate-region is known [31, 32, 33, 34]. The objective of the CEO problem has been to determine the minimum achievable distortion under a sum-rate constraint. It is shown [34, 32, 33] that the sum-rate distortion function of the Gaussian CEO problem can be expressed as

$$\overline{R}(D) = \frac{1}{2}\log^{+}\left\{\frac{\sigma_X^2}{D}\frac{1}{\prod_{i=1}^{M} N_i}\left(\frac{M}{\frac{1}{\sigma_X^2}+\sum_{i=1}^{M}\frac{1}{N_i}-\frac{1}{D}}\right)^{M}\right\}, \tag{2.8}$$

where M is the largest integer between 1 and L that satisfies $\frac{M}{N_M} \geq \left(\frac{1}{\sigma_X^2}+\sum_{i=1}^{M}\frac{1}{N_i}-\frac{1}{D}\right)$.

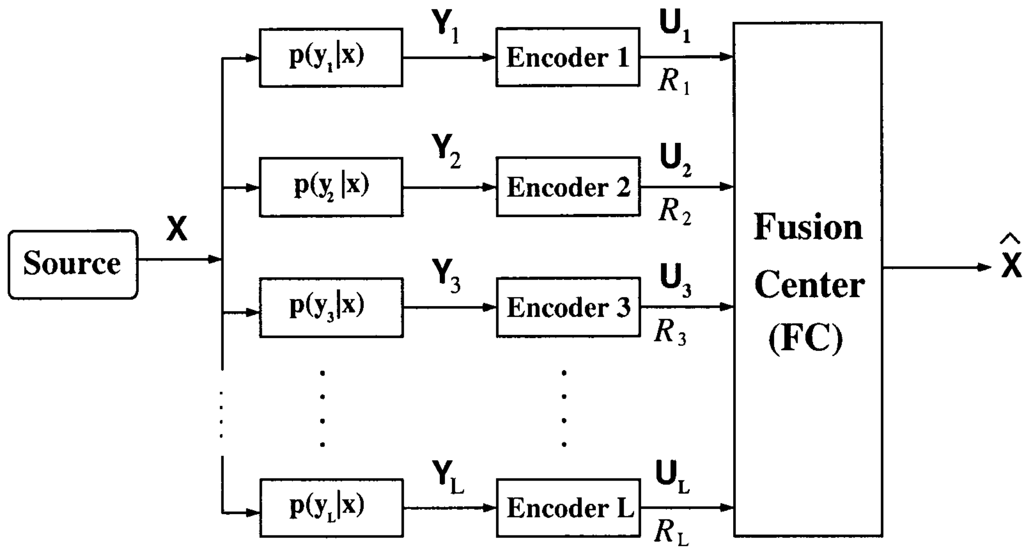

Figure 2.6: The CEO model. The target data X is observed by L wireless agents (sensors) as Y_is. Agents encode and transmit their observations through rate constrained noiseless channels to a FC. The FC obtains an estimate of the source X within an acceptable degree of fidelity.

In this work, we prove that the successive coding strategy can achieve this sum-rate distortion for any finite number of agents and therefore it is optimal.

2.2.4.2 *1-helper* Source Coding

As a special case, consider the case of 2 correlated sources, where one source (called *auxiliary source*) plays the role of partial side information to help the decoder reproduce the transmitted sequence of the other source (which is called the *primary source*). For this problem which is called the *1-helper* problem the admissible rate-region, i.e., the set of all transmission rates for which the primary source can be decoded with an arbitrary small error probability, is determined [46, 47]. The scenario is illustrated in Fig. 2.7. The result is extended to the many-help-one problem, where there are several auxiliary sources that act as the side information at the decoder [48, 49]. They consider a special case where the auxiliary sources are conditionally independent if the primary source is given.

The *1-helper* problem for the correlated memoryless Gaussian sources and squared distortion measures is investigated in [42]. Oohama shows that his outer bound for the

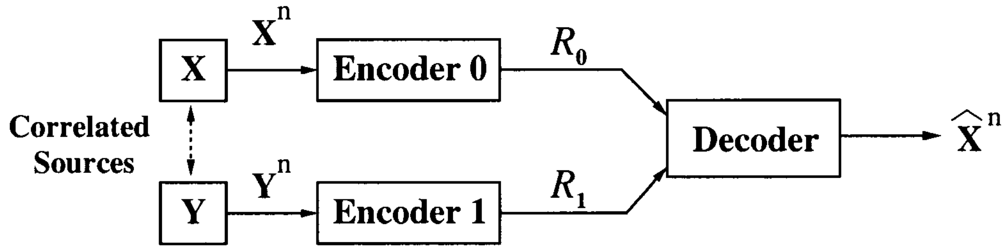

Figure 2.7: *1-helper* source coding scheme. Source X, which is called the *primary source*, is to be reconstructed at the decoder. Source Y, which is called the *auxiliary source* or the *helper*, provides partial side information to the decoder to help reconstruction of the primary source.

2-terminal source coding when combined with the inner bound of Berger [16] and Tung [50] determines the rate-distortion function of the *1-helper* problem.

Theorem 1 *([42]) The rate-distortion function of the Gaussian 1-helper problem can be expressed as*

$$R_0(D_0) = \frac{1}{2}\log^+\left[\frac{\sigma_X^2}{D_0}\left(1-\rho^2+\rho^2 2^{-2R_1}\right)\right], \tag{2.9}$$

where R_0 represent the rate of the primary source and R_1 is the helper's rate.

Oohama extends his results to more than two sources for a certain class of $m+1$ correlated sources, where m source signals are independent noisy versions of the primary source, i.e., $X_i = X_0 + N_i$, $i \in \{1, 2, ..., m\}$. In other words, sources X_1 to X_m are conditionally independent given the source X_0 [31, 34]. The result is extended to the case where $m+1$ sources satisfy a kind of tree structure on their correlation [51]. This condition contains the conditionally independent condition as a special case. In [35], the general correlation structure for the many-help-one problem, which is also called *m-helper problem,* is considered. For this case, the authors of [35] derive a lower bound on the rate-distortion function. However, in [51] it is claimed that the bounding method of [35] does not provide a tight result and the problem still remains open.

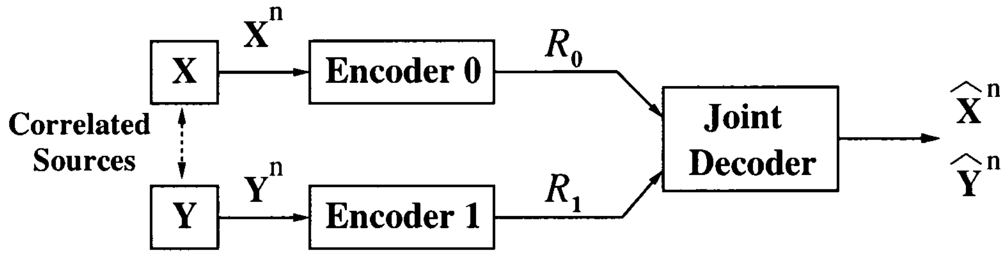

Figure 2.8: *2-terminal* source coding scheme. X and Y are two correlated sources with correlation coefficient ρ. Both sources are to be reconstructed a the decoder.

2.2.4.3 *2*-Terminal Source Coding

In this scheme, two correlated sources are separately encoded and sent to a single decoder, where the decoder reconstruct both sources. The corresponding coding scheme is shown in Fig. 2.8. Berger and Yeung [41] solve the rate-distortion problem for the situation in which the reconstruction of X must be perfect, while that of Y is subject to a distortion criterion. Oohama [42] gives the solution of the Wyner-Ziv problem with coded side information for the case of two sources. He derives an outer region for the rate-distortion region of the *2-terminal* source coding problem and demonstrates that the inner region obtained by Berger [16] and Tung [50] is partially tight [42]. Wagner *et al.* complete the characterization of the rate region for the *2-terminal* source coding problem by showing that the inner bound of Berger and Tung is also tight in the sum-rate [18].

Based on the results of [42] and [18], the whole rate-distortion region of the *2-terminal* source coding system can be characterized by

$$\mathcal{R}(D_0, D_1) = \mathcal{R}_0^*(D_0) \cap \mathcal{R}_1^*(D_1) \cap \widetilde{R}_{01}(D_0, D_1), \tag{2.10}$$

where

$$\mathcal{R}_0^*(D_0) = \left\{(R_0, R_1) : R_0 \geq \frac{1}{2}\log^+\left[\frac{\sigma_X^2}{D_0}\left(1 - \rho^2 + \rho^2 2^{-2R_1}\right)\right]\right\}, \tag{2.11}$$

$$\mathcal{R}_1^*(D_1) = \left\{(R_0, R_1) : R_1 \geq \frac{1}{2}\log^+\left[\frac{\sigma_Y^2}{D_1}\left(1 - \rho^2 + \rho^2 2^{-2R_0}\right)\right]\right\}, \tag{2.12}$$

$$\widetilde{R}_{01}(D_0, D_1) = \left\{(R_0, R_1) : R_0 + R_1 \geq \frac{1}{2}\log^+\left[\frac{\sigma_X^2 \sigma_Y^2}{D_0 D_1}\frac{\beta_{max}}{2}\left(1 - \rho^2\right)\right]\right\}, \tag{2.13}$$

and

$$\beta_{max} = 1 + \sqrt{1 + \frac{4\rho^2}{(1-\rho^2)^2} \times \frac{D_0 D_1}{\sigma_X^2 \sigma_Y^2}}. \tag{2.14}$$

In Chapter 3 and Chapter 4, we apply the successive Wyner-Ziv coding strategy to these multiterminal coding schemes and obtain corresponding rate-distortion functions. We also characterize the rate-region for those coding schemes.

2.3 Successive Wyner-Ziv Coding

The successive coding approach is well-known in the research community. One of the first applications of the successive Wyner-Ziv coding is presented in the successive refinement scenario. In the successive refinement source coding [52, 53, 54] at first, the source will be described by a few bits of information. The description of the source can be improved when more bits of information are available. The scenario is depicted in Fig. 2.9. A source is successively refinable if encoding in multiple stages incurs no rate loss as compared with optimal rate-distortion encoding. It is shown that Gaussian sources with the squared error distortion measure are successively refinable [53]. Steinberg and Merhav [55] consider the problem of successive refinement in the presence of side information. They answer the question whether such a progressive encoding causes rate loss as compared with a single stage Wyner- Ziv coding. It is shown that the jointly Gaussian sources (with the squared error distortion measure) are successively refinable in the Wyner-Ziv setting. It is also shown that there is no rate loss when the difference between the source and the side information is Gaussian and independent of the side information [56]. The characterization of the rate-distortion region for successive refinability for more than two-stage systems is presented in [57]. Application of successive coding strategy in the CEO problem is presented in [45, 58]. In the recent work of [58] encoder i splits R_i to m_i pieces, then it uses Wyner-Ziv codes successively to convey m_i pieces to the decoder. The authors use the idea of source splitting for lossless source coding [59] and its adopted version for the distributed lossy source coding [60], which is called quantization splitting,

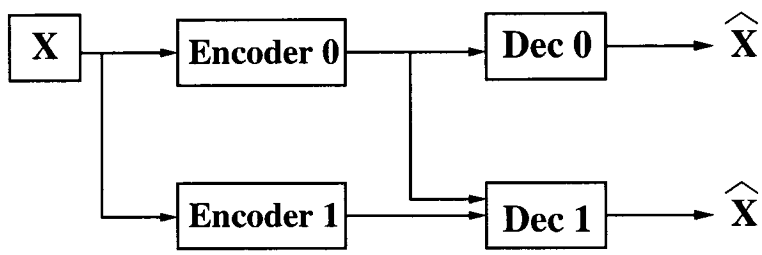

Figure 2.9: Successive refinement coding.

to show that any rate tuple in the rate region of Gaussian CEO problem can be achieved by a low complexity successive Wyner-Ziv coding scheme with at most $2L - 1$ steps. More precisely, this scheme requires $\sum_{i=1}^{L} m_i$ Wyner-Ziv coding steps.

In the successive Wyner-Ziv coding scheme [45], the whole process of quantization, communication and estimation is designed to make use of the statistical knowledge of encoder about decoder's data as the decoder side information in the Wyner-Ziv sense. This is a decentralized strategy because at each stage, only the knowledge sharing between each encoder and decoder is needed. This model couples the application-layer estimation problem with the physical-layer communication problem [8]. So, when an encoder encodes a message, it considers two things. First, its observations and second, its statistical knowledge about the messages that the decoder has already received from other nodes in the network. The latter is known as the "decoder side information" in the sense of Wyner and Ziv [9]. At the decoder, instead of joint one-shot decoding, messages from sources are decoded sequentially in order to increase the fidelity of estimation at each decoding step. By accumulating more data at each step, the decoder has a better side information to use in the next decoding step. Thus, this strategy is also known as the side information aware coding strategy. It gives a low complexity way to attain the prescribed distortion. The joint decoding approach requires all messages decoded simultaneously which is exponentially more complex than a sequence of Wyner-Ziv stages. In fact, joint decoding is very difficult to implement in practice, because random codes have a decoding complexity of the order of $2^{nI(\mathbf{Y};\mathbf{W})}$ where n is the block length, $\mathbf{Y}$ and $\mathbf{W}$ are vectors of input and output of encoders, respectively [58]. The successive coding structure allows us to analyze the finite-L region. Its successive structuring provides flexibility to deal with distributed signal processing and is suitable for networks because the nature of data in networks is distributed [8]. From the perspective of robustness, this scheme performs well [61], i.e., no matter which node fails in the network, the decoder can obtain a non-trivial estimate of the source. But in the joint decoding of all messages, any corruption in the transmitted codewords may cause a complete failure in the decoding [58]. Finally, by applying the successive coding the available practical Wyner-Ziv coding techniques are applicable to more general distributed and multiterminal source coding problems.

Using this strategy, the CEO problem can be decomposed into a sequence of L noisy Wyner-Ziv coding schemes [45]. Although, in general, this simple strategy has suboptimal performances in networks, we evaluate the performance of successive coding strategy in the CEO problem in Chapter 3 and show that it is an optimal strategy in the sense of achieving the sum-rate distortion function of the Gaussian CEO problem. In Chapter 4, it is shown that the successive coding strategy is also optimal in the *1-helper* problem and in the *2- terminal* source coding problem and can achieve the whole rate regions.

2.4 Chapter Summary

In this chapter we reviewed the lossy source coding with side information, which is termed as Wyner-Ziv coding. Then we considered its generalization to noisy observations, which is called remote source coding with decoder side information. Then we moved beyond the Wyner-Ziv scheme to multiterminal coding scheme and provided existing rate-distortion functions for three scenarios in multiterminal source coding, called the CEO problem, *1-helper* problem and *2-terminal* source coding problem. We presented the successive coding strategy as a low complexity coding technique that allows us to analyze multiterminal source coding schemes with finite number of sources. By applying the successive coding strategy, any multiterminal source coding scheme can be decomposed to a sequence of successively structured Wyner-Ziv coding schemes. We will show that this low complexity coding scheme can achieve the optimal performance in the three multiterminal coding schemes presented.

Chapter 3

Successively Structured Gaussian CEO Problem

3.1 Introduction

3.1.1 The CEO Problem

Consider the data gathering application in wireless sensor networks (WSNs): a large number of sensors are deployed in a field to measure a physical phenomenon such as temperature or pressure. They communicate information about their measurements at a limited rate to a single fusion center (FC) for further processing. The scenario is depicted in Fig. 3.1. The key challenge in such a data gathering application is conserving energy of distributed wireless sensor nodes and maximizing their lifetime. Since the sensor measurements are correlated, the correlation should be exploited to avoid redundant transmission. In fact, since the bit-rate directly impacts power consumption at a sensor node, by eliminating the data redundancy and reducing the communication load we can manage the energy resources carefully. The behavior of wireless networks such as WSNs can be modeled by the CEO problem [62, 14, 63, 15, 29, 30, 43, 64, 32, 33]. It is an abstract model for remote monitoring (or sensing) and distributed compression in wireless

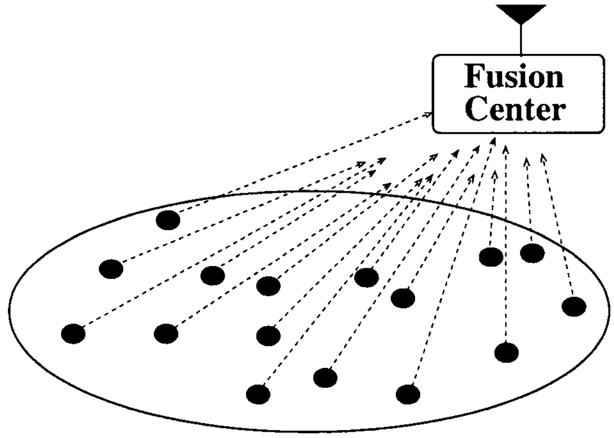

Figure 3.1: Data gathering application in wireless sensor networks. A large number of sensor nodes are deployed in a field to measure a physical phenomenon and then transmit their information to a FC. Since the sensors need to operate over a long period of time, a careful management of the energy resources is the main concern in WSNs.

networks.

In the CEO problem, L agents (sensors) observe independent noisy versions of the source signal X. Agents communicate information about their observations separately to the FC at rate $\{R_i\}_{i=1}^{L}$ without conferring with each other. The FC desires to form an optimal estimate of X based on information received from the agents. The CEO model is shown in Fig. 2.6. The objective of the CEO problem is to determine the minimum achievable distortion under a sum rate constraint [30]. By sum rate, we mean the total rate at which the agents may communicate information about their observations to the FC.

3.1.2 Previous Work

The tradeoff between the sum-rate and the distortion is considered for the discrete case in [29] and for the quadratic Gaussian case in [30] where the problem is considered in the limit of a large number of high rate agents with the same quality of observations. In [43] Oohama gives a complete characterization of the sum-rate distortion function for the Gaussian CEO problem when the number of agents tends to infinity. Oohama presents a solution to the AWGN CEO problem with finite number of agents in [31, 34]. Inner and outer bounds on the rate-distortion region for the general problem have been established in [65, 32]. A complete characterization of the rate region for the Gaussian CEO problem is investigated by Prabhakaran *et al.* [33] where they considered a finite number of agents with the differing quality of observations. In fact, they showed that the achievable rate region established by Chen *et al.* in [32] is indeed the rate-distortion region. A variation of the Gaussian CEO problem is investigated by Pandya *et al.* [35], where each agent observes a Gaussian source of dimension L. They obtain a lower bound for the sum-rate distortion function of the problem. In general, the vector Gaussian CEO problem which is a natural generalization of the scalar Gaussian CEO problem is studied in [66] where inner and outer bounds of the rate distortion region are derived. In [67] Oohama obtains tighter bounds than those of [66] on the rate region of the vector Gaussian CEO problem. As other works in this area we can mention [68] and [69].

The coding/decoding strategy of [43, 31, 65, 32, 33, 34] is based on the joint decoding of all messages. As a result, the decoding process cannot begin before all messages have been received at the joint decoder. Also, in practice, joint decoding of all agent transmissions at the FC is very difficult to implement because random codes have a decoding complexity of the order of $2^{nI(Y_1,Y_2,\ldots,Y_L;U_1,U_2,\ldots,U_L)}$ where n is the code block length. In other words, joint decoding requires all the messages to be decoded simultaneously which is exponentially more complex than single-message decoding. Therefore, as a low complexity way, we use the successive Wyner-Ziv coding [9] in the CEO problem and evaluate its rate-distortion performance compared with the rate-distortion function of the CEO problem. Although in general this simple strategy has a suboptimal performance, we will show that for the Gaussian CEO problem consisting of L agents, successively structured Wyner-Ziv codes can achieve the sum-rate distortion function of the problem.

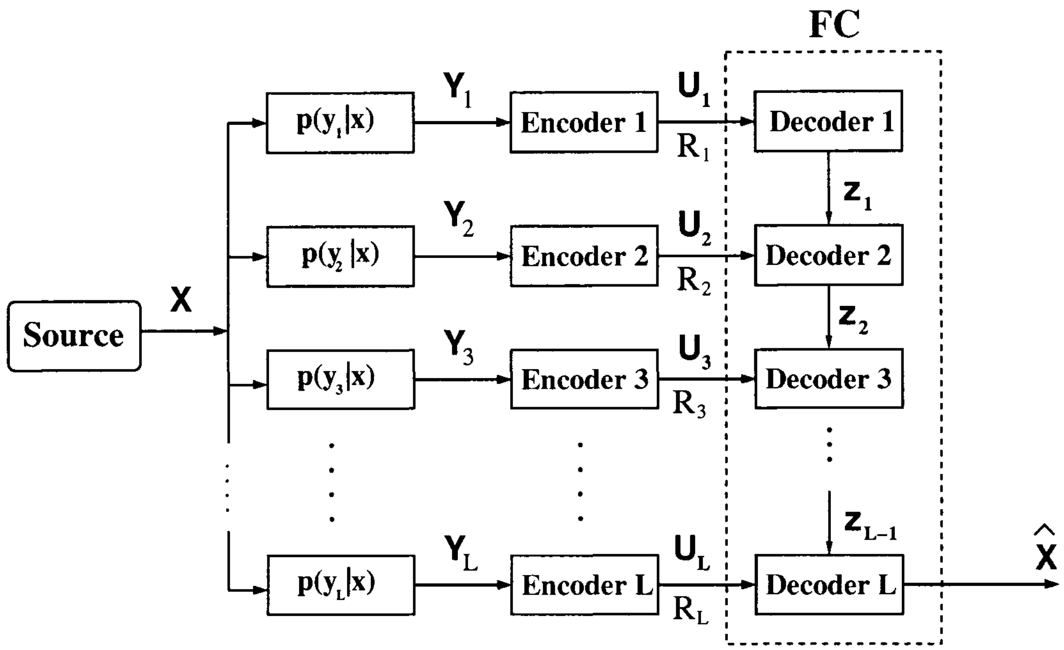

Figure 3.2: The CEO Model based on the successive Wyner-Ziv coding strategy. Each encoder transmits the data based on the side information available at the decoder. The decoder works in a sequential manner to increase the fidelity of estimate at each decoding step.

3.1.3 Successive Coding for CEO Problem

In the successively structured CEO problem, the CEO problem is decomposed into a sequence of data fusion encoding and decoding blocks based on the noisy Wyner-Ziv results [70, 45]. The scenario is shown in Fig. 3.2. Based on this strategy, the distortion sum-rate tradeoff for the CEO problem consisting of two agents with the same SNR is obtained in [45] and it is shown that this is the distortion sum-rate function using the distortion lower bound of Oohama [43]. In this chapter, we consider the CEO problem consisting of L agents ($L \geq 2$) with differing SNRs.

3.1.4 Main Contribution

Our main contribution in this chapter is to show that the sum-rate distortion function of the Gaussian CEO problem can be achieved by a sequence of successively structured

Wyner-Ziv codes. Therefore, the high complexity optimal source code can be decomposed into a sequence of low complexity Wyner-Ziv codes. We apply the successive coding strategy to the Gaussian CEO problem and derive the optimal rate allocation scheme to achieve the minimum distortion under a sum-rate constraint. Then, we determine the optimal sum-rate distortion tradeoff. Comparing with the result of [33], we show that our result is the optimum sum-rate distortion function for the Gaussian CEO problem. We demonstrate that if the number of agents, L, tends to infinity under a sum-rate constraint or if the sum-rate, $\overline{R}$, grows to infinity with a finite number of agents, a sequence of successively structured Wyner-Ziv codes with identical agent communication rates $\left(R_1 = R_2 = \cdots = R_L = \overline{R}/L\right)$ converges to the rate-distortion function. Hence, we can simplify rate allocation problem in a general parallel network with L agents by assigning equal rates to agents, provided that the average rate per agent is either very large or is very small. Finally, based on our results and the results of [8], we present a solution for deriving the communication throughput of a relay network consisting of L relays.

The rest of this chapter is organized as follows: In Section 3.2, we present the system model and problem formulation. In Section 3.3, we use the successive coding strategy and obtain the optimal rate allocation scheme to achieve the minimum distortion under a sum-rate constraint. The rate region for the CEO problem, some extreme cases in the CEO problem and the rate loss with respect to the remote source coding are presented in Section 3.4. The application of our result to the computation of the communication throughput of the relay networks is presented in Section 3.5. Finally, we conclude the chapter in Section 3.6 with a summary of our results.

3.2 Problem Formulation

The distributed network model studied in this chapter is shown in Fig. 3.3. In this model, a firm's CEO is interested in the data sequence $\{X(t)\}_{t=1}^{\infty}$. The target data, $\{X(t)\}_{t=1}^{\infty}$, cannot be observed directly. The CEO deploys a team of L agents to observe the source data sequence. The agents observe independent noisy versions of this sequence, represented by the set $\{Y_i(t)\}_{t=1}^{\infty}$, $i = 1, ..., L$. Agents cannot cooperate to exploit their correlation. They communicate information about their observed data to the FC through noiseless channels at communication rate R_i ($i = 1, ..., L$). This rate constraint comes from the restrictions on the resources such as bandwidth and power that are available at agents (for example, available at sensor nodes).

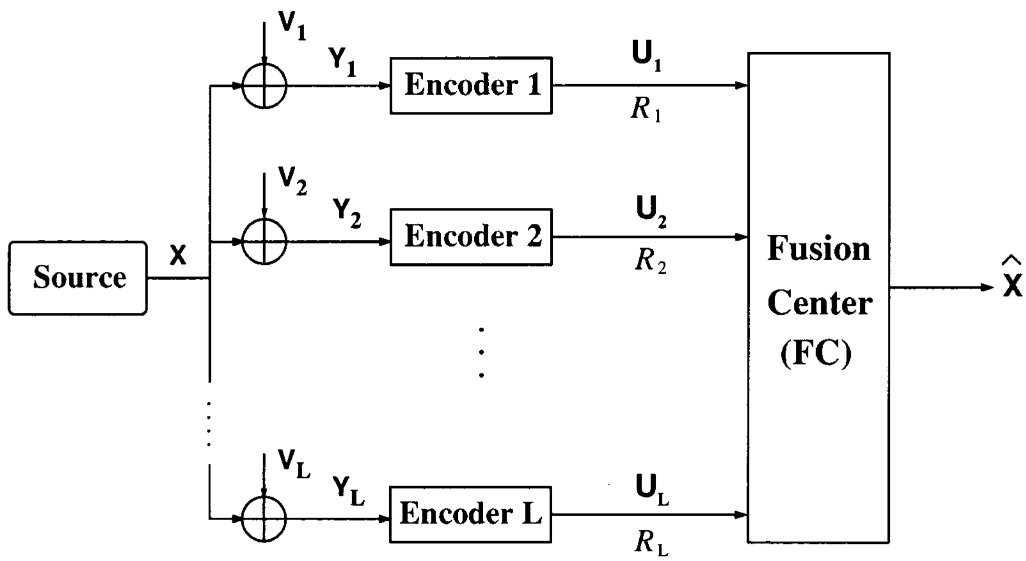

Figure 3.3: Gaussian CEO problem (AWGN CEO problem). The target data X is observed by L agents as Y_is. Agents encode and transmit their observations through rate constrained noiseless channels to a FC. The FC desires to obtain an estimate of the source X within an acceptable degree of fidelity.

Symbolically, for each observation time $t = 1, 2, 3, ..., n, ...$

$$Y_i(t) = X(t) + V_i(t), \qquad i = 1, ..., L \tag{3.1}$$

where $X(t) \sim \mathcal{N}(0, \sigma_X^2)$ and $V_i(t) \sim \mathcal{N}(0, N_i)$ which is *i.i.d.* over i and t. Random variables $Y_i(t)$ for ($i = 1, ..., L$) are conditionally independent given the source $X(t)$. The common alphabets of random variables $X(t)$ and $Y_i(t)$ for $t = 1, 2, ...$ are denoted by $\mathcal{X}$ and $\mathcal{Y}_i$, respectively. We represent n independent instances of $\{X(t)\}_{t=1}^{\infty}$ and $\{Y_i(t)\}_{t=1}^{\infty}$ by data sequences $X^n = \{X(1), X(2), ..., X(n)\}$ and $Y_i^n = \{Y_i(1), Y_i(2), ..., Y_i(n)\}$, respectively. Agent i encodes its observation data sequence Y_i^n into a source code sequence

$\varphi_i\left(Y_i^n\right)$ separately where the encoder functions are defined as

$$\varphi_i : \mathcal{Y}_i^n \rightarrow \mathcal{C}_i = \{1, 2, ..., |\mathcal{C}_i|\} \tag{3.2}$$

for $i = 1, 2, ..., L$. The coded sequences from L agents are sent to the FC with the rate constraints

$$\frac{1}{n}\log_2|\mathcal{C}_i| \leq R_i + \delta \ , \ i = 1, 2, ..., L \tag{3.3}$$

where δ is an arbitrarily small positive number. The task of FC is to reconstruct the source message to an acceptable degree of fidelity. Since we have considered noiseless channels between encoders and the decoder, the decoder observes L-tuple $\varphi^L = \left(\varphi_1\left(Y_1^n\right), ..., \varphi_L\left(Y_L^n\right)\right)$, and makes an estimate of the source X^n as $\widehat{X}^n$. The decoder function is given by

$$\psi : \mathcal{C}_1 \times \cdots \times \mathcal{C}_L \rightarrow \mathcal{X}^n. \tag{3.4}$$

Therefore, the reconstructed signal can be represented by $\widehat{X}^n = \psi\left(\varphi_1\left(Y_1^n\right), ..., \varphi_L\left(Y_L^n\right)\right)$. The fidelity measure is the average distortion, defined as $\Delta = E\left[\frac{1}{n}\sum_{t=1}^n d(X(t), \widehat{X}(t))\right]$ where $d : \mathcal{X}^2 \rightarrow [0, \infty)$ is the MSE distortion measure, i.e., $d(X, \widehat{X}) = (X - \widehat{X})^2$. Let $\mathcal{F}_\delta^{(n)}\left(R_1, ..., R_L\right)$ denote all $(L+1)-$tuple encoder and decoder functions $(\varphi_1, ..., \varphi_L, \psi)$ that satisfy (3.2)-(3.4). The goal is to determine the tradeoff between the total rate of encoders, $\sum_{i=1}^L R_i = \overline{R}$, and the minimum achievable distortion. A sum-rate distortion pair $\left(\overline{R}, D\right)$ is admissible if for any $\delta > 0$ and any $n \geq n_0(\delta)$ there exists a pair $\left(\varphi^L, \psi\right) \in \mathcal{F}_\delta^{(n)}\left(R_1, ..., R_L\right)$ such that $\Delta \leq D + \delta$.

To analyze the CEO problem, we apply the successive Wyner Ziv coding strategy. This strategy is a joint design of source coding, communication and data fusion steps. When an agent encodes a message, previously decoded messages that are available at the decoder act as the decoder side information. At the decoder, instead of joint decoding, messages from agents are decoded sequentially with the objective of increasing the fidelity of estimation at each decoding step. Hence, this strategy simplifies the analysis of the CEO problem by decomposing the CEO problem into L successive Wyner-Ziv coding cases [8]. In Section 3.3 we evaluate the sum-rate distortion performance of the successive coding strategy in the Gaussian CEO problem and prove that this strategy can achieve the sum-rate distortion function of the CEO problem.

3.3 Optimal Rate Allocation Strategy

Assume that the memoryless Gaussian source X is observed at each agent in independent AWGN. If we represent the length-n source vector by $\mathbf{X}$ and the observation of the i^{th} agent by $\mathbf{Y_i}$, then $\mathbf{Y_i} = \mathbf{X} + \mathbf{V_i}$ where $\mathbf{X} \sim \mathcal{N}(0, \sigma_X^2 \mathbf{I})$ and $\mathbf{V_i} \sim \mathcal{N}(0, N_i \mathbf{I})$. Observations are conditionally independent given the source. The FC produces the source estimate $\widehat{\mathbf{X}}$ to an acceptable degree of fidelity. The measure of the fidelity is the average distortion criterion, i.e., $\frac{1}{n} E\left[\sum_{i=1}^{n} d(x_i, \widehat{x}_i)\right]$ where d is the MSE distortion measure. The distortion-rate performance of the successive coding strategy for the quadratic Gaussian CEO problem is investigated in [70]. It is shown that the following distortion-rate tradeoff can be achieved:

$$D_i = \frac{D_{i-1} N_i}{D_{i-1} + N_i} + \left(\frac{D_{i-1}^2}{D_{i-1} + N_i} \right) 2^{-2R_i}, \qquad i = 1, ..., L \tag{3.5}$$

where $D_0 = \sigma_X^2$, D_i is the distortion achieved by the i^{th} decoder and R_i is the communication rate between the i^{th} agent and the FC in terms of bits per observation sample. First we consider the 2-agent CEO problem and obtain optimal rate allocations that minimize the final estimation distortion under a sum-rate constraint. Then, in Section 3.3.2, we use this result as the base case of an induction to obtain the optimal rate allocation strategy in the $L-$agent CEO problem.

3.3.1 CEO Problem with 2 Agents

Here, $\begin{pmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{pmatrix} = \begin{pmatrix} I \\ I \end{pmatrix} \mathbf{X} + \begin{pmatrix} \mathbf{V}_1 \\ \mathbf{V}_2 \end{pmatrix}$ where $\mathbf{V}_1 \sim \mathcal{N}(0, N_1 I)$ and $\mathbf{V}_2 \sim \mathcal{N}(0, N_2 I)$.

Theorem 2 *The optimal sum-rate distortion tradeoff for the 2-agent Gaussian CEO problem based on the successive coding strategy can be expressed as*

$$\overline{R}(D) = \begin{cases} \frac{1}{2}\log^+\left(\frac{4\sigma_X^2}{DN_1N_2\left(\frac{1}{D^*}-\frac{1}{D}\right)^2}\right) & D^* < D < min\left(D_{11}, D_{22}\right) \\ \frac{1}{2}\log^+\left(\frac{\sigma_X^4}{D\left(\sigma_X^2+N_1\right)-\sigma_X^2 N_1}\right) & D_{11} < D < \sigma_X^2 \quad \& \quad N_1 < N_2 \\ \frac{1}{2}\log^+\left(\frac{\sigma_X^4}{D\left(\sigma_X^2+N_2\right)-\sigma_X^2 N_2}\right) & D_{22} < D < \sigma_X^2 \quad \& \quad N_2 < N_1 \end{cases} \tag{3.6}$$

where $D^* = \left(\frac{1}{\sigma_X^2}+\frac{1}{N_2}+\frac{1}{N_1}\right)^{-1}$, $D_{11} = \left(\frac{1}{\sigma_X^2}+\frac{1}{N_1}-\frac{1}{N_2}\right)^{-1}$, *and* $D_{22} = \left(\frac{1}{\sigma_X^2}+\frac{1}{N_2}-\frac{1}{N_1}\right)^{-1}$.

Proof By using (3.5) for $i = 1, 2$ we obtain

$$D_2 = \frac{\sigma_X^2\left(1+\frac{\sigma_X^2}{N_1}2^{-2R_1}\right)}{1+\frac{\sigma_X^2}{N_1}} \times \frac{\frac{N_2}{N_1}\left(1+\frac{\sigma_X^2}{N_1}\right)+\frac{\sigma_X^2}{N_1}\left(2^{-2R_2}+\frac{\sigma_X^2}{N_1}2^{-2(R_1+R_2)}\right)}{\frac{N_2}{N_1}\left(1+\frac{\sigma_X^2}{N_1}\right)+\frac{\sigma_X^2}{N_1}\left(1+\frac{\sigma_X^2}{N_1}2^{-2R_1}\right)}. \tag{3.7}$$

It can easily be seen from Equation (3.7) that under a fixed R_1, by increasing R_2, the achievable distortion decreases. Also, under a fixed R_2, the distortion will be reduced by increasing R_1. These results are somehow obvious since we expect to have less distortion by increasing the rate. We optimize the distortion expression D_2 over a sum-rate constraint $\overline{R}$ to minimize the achievable distortion using the successive coding strategy. The optimization is over the fraction of the total rate allocated to each agent: $R_1 = \lambda\overline{R}$ and $R_2 = (1-\lambda)\overline{R}$, given the total fixed communication rate $R_1 + R_2 = \overline{R}$. Therefore, considering the rates in form of nats (instead of bits), D_2 is as follows:

$$D_2 = \frac{\sigma_X^2\left(1+\frac{\sigma_X^2}{N_1}e^{-2\lambda\overline{R}}\right)}{1+\frac{\sigma_X^2}{N_1}} \times \frac{\frac{N_2}{N_1}\left(1+\frac{\sigma_X^2}{N_1}\right)+\frac{\sigma_X^2}{N_1}\left(e^{-2(1-\lambda\overline{R})}+\frac{\sigma_X^2}{N_1}e^{-2\overline{R}}\right)}{\frac{N_2}{N_1}\left(1+\frac{\sigma_X^2}{N_1}\right)+\frac{\sigma_X^2}{N_1}\left(1+\frac{\sigma_X^2}{N_1}e^{-2\lambda\overline{R}}\right)}. \tag{3.8}$$

To obtain the optimum fractional rate (λ), we take the derivation of D_2 with respect to λ and set the result equal to zero,

$$\begin{aligned} \frac{d}{d\lambda}D_2 \quad = \quad & \left[e^{-2\overline{R}}\left(\frac{N_2}{N_1}+\frac{\sigma_X^2}{N_1}\left(1+\frac{N_2}{N_1}\right)\right)\right]e^{4\lambda\overline{R}}+\left[2\left(\frac{\sigma_X^2}{N_1}\right)^2 e^{-2\overline{R}}\right]e^{2\lambda\overline{R}} \\ & -\left[e^{-2\overline{R}}\left(\frac{N_2}{N_1}\left(\frac{\sigma_X^2}{N_1}\right)^2+\left(\frac{N_2}{N_1}-1\right)\left(\frac{\sigma_X^2}{N_1}\right)^3\right)+\left(\frac{N_2}{N_1}\right)^2\left(1+\frac{\sigma_X^2}{N_1}\right)^2\right] = 0. \end{aligned} \tag{3.9}$$

This is a quadratic equation with respect to $e^{2\lambda\overline{R}}$. The acceptable root of this equation is as follows:

$$e^{2\lambda\overline{R}} = \left[\frac{-\left(\frac{\sigma_X^2}{N_1}\right)^2 + \frac{N_2}{N_1}\left(1+\frac{\sigma_X^2}{N_1}\right)\sqrt{\left(\frac{\sigma_X^2}{N_1}\right)^2 + \left(\frac{N_2}{N_1}+\frac{N_2}{N_1}\frac{\sigma_X^2}{N_1}+\frac{\sigma_X^2}{N_1}\right)e^{2\overline{R}}}}{\left(\frac{N_2}{N_1}+\frac{N_2}{N_1}\frac{\sigma_X^2}{N_1}+\frac{\sigma_X^2}{N_1}\right)}\right]. \quad (3.10)$$

Therefore, the optimal fractional rate allocation factor can be obtained as

$$\lambda_{opt} = \frac{1}{2\overline{R}}\ln\left[\frac{-\left(\frac{\sigma_X^2}{N_1}\right)^2 + \frac{N_2}{N_1}\left(1+\frac{\sigma_X^2}{N_1}\right)\gamma}{\left(\frac{N_2}{N_1}+\frac{N_2}{N_1}\frac{\sigma_X^2}{N_1}+\frac{\sigma_X^2}{N_1}\right)}\right], \quad (3.11)$$

where $\gamma = \sqrt{\left(\frac{\sigma_X^2}{N_1}\right)^2 + \left(\frac{N_2}{N_1}+\frac{N_2}{N_1}\frac{\sigma_X^2}{N_1}+\frac{\sigma_X^2}{N_1}\right)e^{2\overline{R}}}$. Since the second derivative of D_2 with respect to λ is positive, λ_{opt} is the value of λ that gives the *minimum* of the distortion-rate tradeoff. We substitute this optimal λ_{opt} in equation (3.8) to obtain the minimum achievable distortion,

$$D_{min} = \frac{\sigma_X^2\frac{N_2}{N_1}\left(\gamma+\frac{\sigma_X^2}{N_1}\right)}{\gamma\left(\frac{N_2}{N_1}+\frac{N_2}{N_1}\frac{\sigma_X^2}{N_1}+\frac{\sigma_X^2}{N_1}\right)^2}\left[\frac{N_2}{N_1}+\frac{N_2}{N_1}\frac{\sigma_X^2}{N_1}+\frac{\sigma_X^2}{N_1}+\frac{\sigma_X^2}{N_1}e^{-2\overline{R}}\left(\gamma+\frac{\sigma_X^2}{N_1}\right)\right]. \quad (3.12)$$

To simplify the optimal distortion expression, multiply Equation (3.12) by $\frac{\left(\gamma-\frac{\sigma_X^2}{N_1}\right)^2}{\left(\gamma-\frac{\sigma_X^2}{N_1}\right)^2}$, to get

$$\begin{aligned} D_{min} &= \left[\left(\frac{N_2}{N_1}+\frac{N_2}{N_1}\frac{\sigma_X^2}{N_1}+\frac{\sigma_X^2}{N_1}\right)\left(\gamma-\frac{\sigma_X^2}{N_1}\right)+\frac{\sigma_X^2}{N_1}e^{-2\overline{R}}\left(\gamma^2-\left(\frac{\sigma_X^2}{N_1}\right)^2\right)\right] \\ &\quad \times\frac{\sigma_X^2\frac{N_2}{N_1}\left(\gamma^2-\left(\frac{\sigma_X^2}{N_1}\right)^2\right)}{\gamma\left(\frac{N_2}{N_1}+\frac{N_2}{N_1}\frac{\sigma_X^2}{N_1}+\frac{\sigma_X^2}{N_1}\right)^2\left(\gamma-\frac{\sigma_X^2}{N_1}\right)^2}. \end{aligned} \quad (3.13)$$

Also, from the definition of γ , we can substitute the result of equation below in (3.13),

$$\gamma^2-\left(\frac{\sigma_X^2}{N_1}\right)^2 = \left(\frac{N_2}{N_1}+\frac{N_2}{N_1}\frac{\sigma_X^2}{N_1}+\frac{\sigma_X^2}{N_1}\right)e^{2\overline{R}}. \quad (3.14)$$

Therefore, the minimum achievable distortion is as follows:

$$D_{min} = \frac{\left(\frac{N_2}{N_1}\right)\sigma_X^2 e^{2\overline{R}}}{\left(\gamma-\frac{\sigma_X^2}{N_1}\right)^2} = \frac{\left(\frac{N_2}{N_1}\right)\sigma_X^2 e^{2\overline{R}}}{\left(\sqrt{\left(\frac{\sigma_X^2}{N_1}\right)^2+\left(\frac{N_2}{N_1}+\frac{N_2}{N_1}\frac{\sigma_X^2}{N_1}+\frac{\sigma_X^2}{N_1}\right)e^{2\overline{R}}}-\frac{\sigma_X^2}{N_1}\right)^2}. \quad (3.15)$$

Since the λ_{opt} of (3.11) is not always between 0 and 1, we should obtain the best value for λ to obtain minimum distortion. When the λ_{opt} is greater than one, we should choose λ to be equal one in (3.8) to obtain the minimum value of D_2. This is the same as assigning the whole rate to the first agent. Also, when λ_{opt} is less than zero, we should assign the whole rate to the second agent, i.e., $\lambda = 0$. As a result,

$$D_{min} = \begin{cases} \sigma_X^2 e^{2\overline{R}} \left(\frac{N_2}{N_1}\right) \left(\gamma - \frac{\sigma_X^2}{N_1}\right)^{-2} & 0 < \lambda_{opt} < 1 \\ \sigma_X^2 \left(1 + \frac{\sigma_X^2}{N_1} e^{-2\overline{R}}\right) \left(1 + \frac{\sigma_X^2}{N_1}\right)^{-1} & \lambda_{opt} > 1 \\ \sigma_X^2 \left(1 + \frac{\sigma_X^2}{N_2} e^{-2\overline{R}}\right) \left(1 + \frac{\sigma_X^2}{N_2}\right)^{-1} & \lambda_{opt} < 0 \end{cases} \tag{3.16}$$

Solving (3.16) in terms of $\overline{R}$ will result in (3.6). This completes the Proof.

Comparing our results with the sum-rate distortion function of the 2−agent CEO problem, presented in [65, 32, 33], reveals that the achievable sum-rate distortion tradeoff in (3.6) is indeed the rate-distortion function of the CEO problem consisting of 2 agents with different *SNR*s. In other words, the successive Wyner-Ziv coding is an optimal strategy in the CEO problem in the sense of achieving the sum-rate distortion function. By doing some manipulations, the optimal rate allocation scheme can be obtained as follows:

Corollary 1 *The optimal rate allocation strategy that can achieve the sum-rate distortion function of the 2-agent Gaussian CEO problem can be expressed as follows:*

$$\begin{cases} R_1(D) = \frac{1}{2}\log\left(\frac{2\left(\frac{1}{\sigma_X^2}+\frac{1}{N_1}\right)-\left(\frac{1}{D^*}-\frac{1}{D}\right)}{\frac{N_1}{\sigma_X^2}\left(\frac{1}{D^*}-\frac{1}{D}\right)}\right), \\ R_2(D) = \frac{1}{2}\log\left(\frac{4}{DN_2\left(\frac{1}{D^*}-\frac{1}{D}\right)\left(2\left(\frac{1}{\sigma_X^2}+\frac{1}{N_1}\right)-\left(\frac{1}{D^*}-\frac{1}{D}\right)\right)}\right). \end{cases} \tag{3.17}$$

These are the optimal values of compression rates for the successively structured CEO problem that minimize the final distortion.

3.3.2 CEO Problem with L Agents ($L > 2$)

We aim to obtain the optimal rate allocation strategy that can achieve minimum distortion under a sum-rate constraint in a Gaussian CEO problem consisting of L agents ($L > 2$).

To minimize the final distortion D_L under the sum-rate constraint $\sum_{i=1}^{L} R_i = \overline{R}$, we introduce the Lagrange multiplier μ and define

$$J(R_1, R_2, ..., R_L) = D_L + \mu \left(\sum_{i=1}^{L} R_i \right). \tag{3.18}$$

Taking derivative of J with respect to R_i and setting the derivative to 0 leads to

$$\frac{\partial D_L}{\partial R_1} = \frac{\partial D_L}{\partial R_2} = ... = \frac{\partial D_L}{\partial R_L} = -\mu. \tag{3.19}$$

If we obtain the derivative of D_L with respect to R_i, the optimum values of $(R_1, ..., R_L)$ which minimize the final distortion achievable D_L can be obtained by solving $(L-1)$ equations of (3.19) and the sum-rate constraint of $\sum_{i=1}^{L} R_i = \overline{R}$. The following lemma determines the derivative of D_L with respect to R_k.

Lemma 1 *The derivative of D_n with respect to R_k for $k \leq n \leq L$ can be expressed as:*

$$\frac{\partial D_n}{\partial R_k} = \frac{(-2\log 2)\, D_{k-1}^2 2^{-2R_k}}{D_{k-1} + N_k} \prod_{i=k}^{n-1} \left[\frac{D_i^2 + 2N_{i+1}D_i + N_{i+1}^2 2^{2R_{i+1}}}{(D_i + N_{i+1})^2\, 2^{2R_{i+1}}} \right], \tag{3.20}$$

where $D_0 = \sigma_X^2$.

Proof See Appendix I at the end of this chapter.

Thus, from (3.19) we have

$$\frac{\partial D_L}{\partial R_k} = \frac{\partial D_L}{\partial R_{k+1}}. \tag{3.21}$$

Substituting (3.20) in (3.21) and doing some manipulations gives the following result.

Corollary 2 *The optimal rates in the quadratic Gaussian CEO problem can be found using the following iterative expression:*

$$R_{k+1} = \frac{1}{2}\log_2 \left[\frac{D_k^2 2^{2R_k} (D_{k-1} + N_k)(D_k + N_{k+1})}{D_{k-1}^2 N_{k+1}^2} - \frac{D_k^2 + 2N_{k+1}D_k}{N_{k+1}^2} \right]. \tag{3.22}$$

Using (3.22) we can obtain the optimum values of R_i's that minimize the final achievable distortion D_L under the constraint that $\sum_{i=1}^{L} R_i = \overline{R}$. Without loss of generality, we assume that $N_1 \leq N_2 \leq \cdots \leq N_L$. The values of $(R_1, R_2, ..., R_L)$ are the optimal rates for agents in the sense of permitting the source decoder to reconstruct X within the fidelity of D_L which is the minimum distortion attainable with the successive coding strategy.

Lemma 2 *The optimal rate allocation for the quadratic Gaussian CEO problem based on the successive coding strategy can be expressed as*

$$R_j(D_M) = \frac{1}{2}\log_2\left(\frac{M}{N_j}\frac{S_j}{S_{j-1}}\frac{1}{\left(\frac{1}{D^*(M)} - \frac{1}{D_M}\right)}\right), \tag{3.23}$$

for $j = 1, 2, ..., M$ *where,*

$$D^*(M) = \left(\frac{1}{\sigma_X^2} + \sum_{i=1}^{M}\frac{1}{N_i}\right)^{-1} \tag{3.24}$$

is the infinite rate lower bound on the distortion, which can be achieved when all the agents are given infinite rate. The number of active agents is M where M is the largest integer value between 1 and L that satisfies the following inequality:

$$\frac{M}{N_M} \geq \left(\frac{1}{D^*(M)} - \frac{1}{D_M}\right), \tag{3.25}$$

and

$$S_j = M\left(\frac{1}{\sigma_X^2} + \sum_{i=1}^{j}\frac{1}{N_i}\right) - j\left(\frac{1}{D^*(M)} - \frac{1}{D_M}\right). \tag{3.26}$$

The average distortion after each step of decoding can be expressed as

$$D_j = \frac{M}{S_j}, \qquad j = 1, 2, \ldots, M. \tag{3.27}$$

Proof The proof is presented in Appendix II at the end of this chapter.

As the final goal for the CEO problem, we are interested in the tradeoff between the estimation distortion and the sum-rate. The following theorem gives the sum-rate distortion performance of the successive coding strategy in the CEO problem.

Theorem 3 *The sum-rate distortion tradeoff based on the successive Wyner-Ziv coding strategy for the quadratic Gaussian CEO problem consisting of L agents is*

$$\overline{R}(D) = \frac{1}{2}\log_2\left\{\frac{\sigma_X^2}{D}\frac{1}{\prod_{i=1}^{M} N_i}\left(\frac{M}{\frac{1}{D^*(M)} - \frac{1}{D}}\right)^M\right\} \tag{3.28}$$

for $D^*(M) \leq D \leq \sigma_X^2$ where M is the number of active agents, which is the largest integer value between 1 and L such that $\frac{M}{N_M} \geq \left(\frac{1}{D^*(M)} - \frac{1}{D}\right)$.

Proof By considering optimal rate allocation of (3.23), adding R_1 to R_M will result in the sum-rate distortion tradeoff in (3.28).

Comparing the result of Theorem 3 with the results of [32] and [33] reveals that the achievable sum-rate distortion tradeoff for the Gaussian CEO problem based on the successive coding strategy is indeed the sum-rate distortion function of this problem. Therefore, the successive coding strategy achieves the sum-rate distortion function. This strategy has low complexity for obtaining a prescribed distortion and thus is simple to implement for wireless networks.

Using (3.23), the difference between the rates of two sequential stages in the successive coding strategy can be expressed as

$$R_{l+1} - R_l = \frac{1}{2}\log_2\left(\frac{N_l}{N_{l+1}}\frac{S_{l+1}S_{l-1}}{S_l^2}\right) \tag{3.29}$$

It is easy to show that $0 < \frac{N_l}{N_{l+1}}\frac{S_{l+1}S_{l-1}}{S_l^2} < 1$. Therefore, R_l decreases in l. In other words, each encoder encodes for a decoder with better side information, therefore, can obtain a better compression rate than its preceding encoders [8].

From Lemma 2, it is clear that the number of active agents depends on the degree of fidelity (distortion) at which the underlying source can be estimated by the FC. For the optimal rate allocation, we always give preference to agents with small noise variances. The agent with the smallest noise variance will receive the full rate allocation, i.e., $R_1 = \overline{R}$ if $\frac{1}{\sigma_X^2} \leq \frac{1}{D} \leq \frac{1}{\sigma_X^2} + \frac{1}{N_1} - \frac{1}{N_2}$. In general, $R_j = R_{j+1} = ...R_L = 0$ if

$$\frac{1}{\sigma_X^2} + \sum_{i=1}^{j-1}\frac{1}{N_i} - \frac{j-1}{N_{j-1}} \leq \frac{1}{D} \leq \frac{1}{\sigma_X^2} + \sum_{i=1}^{j}\frac{1}{N_i} - \frac{j}{N_j}$$

In fact, if we want to obtain an estimate of the source with an average distortion of $\left(\frac{1}{\sigma_X^2} + \sum_{i=1}^{L} \frac{1}{N_i} - \frac{L}{N_L}\right)^{-1}$ or better, all of the agents should transmit their observations. Therefore, the optimal rate allocation based on the successive coding strategy has the behavior of the Generalized Water-filling approach [32]. As an example, the sum-rate distortion function for the case of $L = 3$ agents is shown in Fig. 3.4. The dotted line shows the case when all the agents transmit their observations. The solid line demonstrates the case when the two agents with the higher quality of observations transmit while the dashed line corresponds to the case when only the agent with the best quality of observation transmits its observation. Therefore, the number of active agents depends on the desired degree of fidelity. As the value of acceptable distortion decreases, the number of active agents increases. An interesting result is that even when the SNRs are identical at different agents, our rate allocation scheme assigns different rates to different agents to minimize the final distortion under a sum-rate constraint, i.e., if $N_1 = N_2 = \ldots = N_L = N$, then

$$R_j = \frac{1}{2}\log_2\left[\frac{\frac{L}{N}\left(\frac{L}{\sigma_X^2} + j\left(\frac{1}{D_L} - \frac{1}{\sigma_X^2}\right)\right)}{\left(\frac{1}{\sigma_X^2} + \frac{L}{N} - \frac{1}{D_L}\right)\left(\frac{L}{\sigma_X^2} + (j-1)\left(\frac{1}{D_L} - \frac{1}{\sigma_X^2}\right)\right)}\right], \qquad j = 1, 2, \ldots, L\ . \tag{3.30}$$

Therefore, we have the asymmetric rate allocation. In addition, all of the agents have non-zero rates and none of them will receive the total sum-rate.

To understand how the distortion, D_j, changes with j, consider the following statement: The ratio of D_j to D_{j-1} can be expressed as:

$$\frac{D_j}{D_{j-1}} = 1 - \frac{\frac{M}{N_j} - \left(\frac{1}{D^*(M)} - \frac{1}{D_M}\right)}{S_j}. \tag{3.31}$$

Since M is chosen such that $\frac{M}{N_M} \geq \left(\frac{1}{D^*(M)} - \frac{1}{D_M}\right)$ and we assume that $N_j \leq N_M$, so $D_j \leq D_{j-1}$. Hence, the average distortion decreases after each decoding step. Intuitively, since the FC obtains higher amount of data after each step, it has better side information to use in the decoding process. Note that the rate of improvement in the distortion decreases after each stage of the successive coding. In other words, we can show that

$$\frac{D_{j-1}}{D_j} > \frac{D_j}{D_{j+1}} \geq 1. \tag{3.32}$$

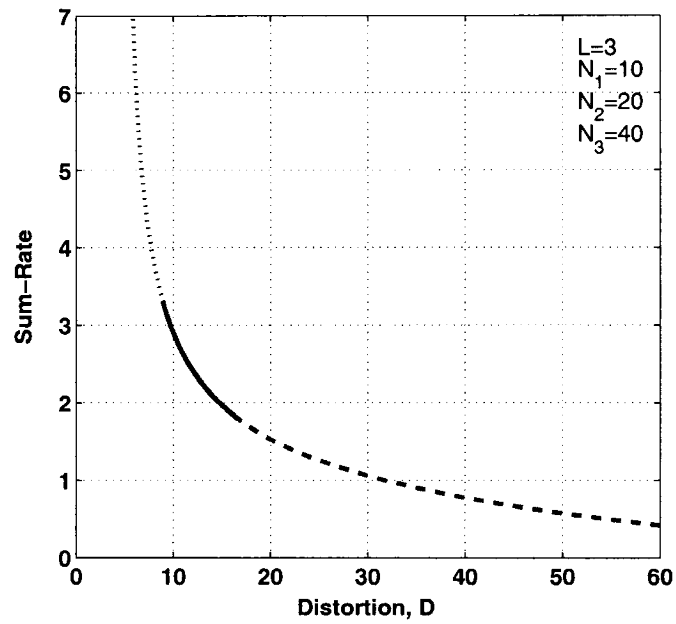

Figure 3.4: Sum-rate distortion function for the 3-node CEO problem. Dotted line: all the agents transmit their observations. Solid line: two agents with the higher quality of observations transmit. Dashed line: only the agent with the best quality of observation, i.e., agent 1 transmits its observation.

3.3.3 A Lower bound on the Minimum Achievable Distortion

In this part, a lower bound on the minimum achievable distortion for the Gaussian CEO problem with source variance σ_X^2 and observation noise variances N_i is presented. We show that the minimum achievable distortion satisfies, $D_{min} \geq D^*$ where

$$D^* = \left(\frac{1}{\sigma_X^2} + \sum_{i=1}^{L} \frac{1}{N_i} \right)^{-1}. \tag{3.33}$$

This lower bound can be found when the decoder has direct access to the observations of the agents. The system model for this idealization is shown in Fig. 3.5. Thus, if agents are given infinite rate, they can simply forward their observations to the CEO at full resolution. Then the CEO uses all the observations to make the minimum mean-square estimation $E\left[X \mid Y_1, Y_2, ..., Y_L\right]$ resulting in estimation error $\sigma^2_{X|Y_1,Y_2,...,Y_L}$, which is equal to D^*. If $\overline{R} \rightarrow \infty$, then the minimum achievable distortion of the successive

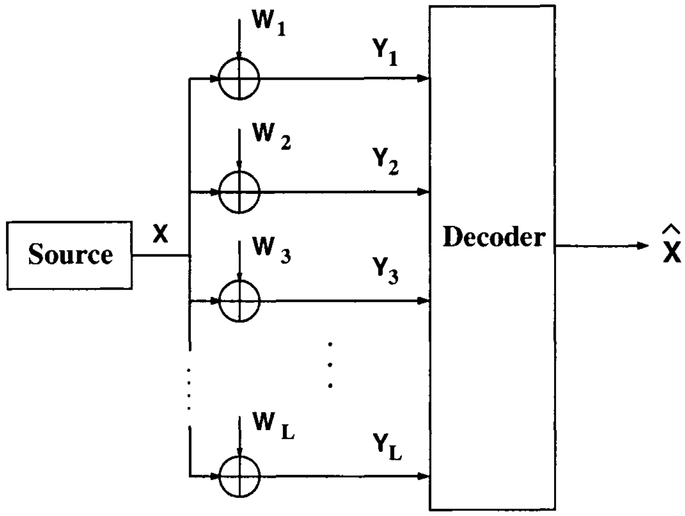

Figure 3.5: The network model for the ideal case when the decoder has direct access to the observations of the agents.

coding strategy would be the same as this lower bound.

Now, we want to show that this lower bound can be obtained using Shannon's source-channel coding theorem (also known as Shannon's separation theorem) [71]. The network of Fig. 3.5 can be considered as the system in which the data of a Gaussian source is distributed across the multi-antenna channel with one transmit antenna (the source) and L receive antennas [38]. The system would be an ergodic point-to-point communication system [38]. Therefore, the Shannon's separation theorem can be applied.

For this channel, we have

$$\begin{bmatrix} Y_1 \\ \vdots \\ Y_L \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} X + \begin{bmatrix} V_1 \\ \vdots \\ V_L \end{bmatrix} \tag{3.34}$$

which can be shown as the vector equation

$$\mathbf{Y} = \mathbf{S} + \mathbf{V}. \tag{3.35}$$

The capacity of this channel can be computed as [5]

$$C = \frac{1}{2} \log_2 \left(\frac{|K_S + K_V|}{|K_V|} \right) \tag{3.36}$$

where $|A|$ represents the determinant of matrix A, K_S and K_V are the covariance matrices of random vectors S and V, respectively. Since K_s is an $L \times L$ matrix with identical entries σ_X^2, and $K_V = diag(N_1, N_2, ..., N_L)$, it can be shown that [32]

$$|K_S + K_V| = \left[1 + \sigma_X^2 \left(\sum_{i=1}^{L} \frac{1}{N_i}\right)\right] \prod_{i=1}^{L} N_i \tag{3.37}$$

Thus, the capacity formula can be simplified as

$$C = \frac{1}{2} \log_2 \left[1 + \sigma_X^2 \sum_{i=1}^{L} \left(\frac{1}{N_i}\right)\right] \tag{3.38}$$

Also, the rate-distortion function of the Gaussian source with variance σ_X^2 is $R(D) = \frac{1}{2} \log_2 \frac{\sigma_X^2}{D}$. The minimum achievable distortion for the Gaussian source across this multi-antenna channel can be obtained using separation theorem. We put $R \leq C$ and get $D \geq \left(\frac{1}{\sigma_X^2} + \sum_{i=1}^{L} \frac{1}{N_i}\right)^{-1} = D^*$.

3.3.4 Numerical Results

We assume that the target data is an *i.i.d.* zero mean Gaussian process with variance $\sigma_X^2 = 100$. Fig. 3.6 shows the decrease in the mean-squared estimation distortion versus the total number of agents where agents have equal $SNR = \frac{\sigma_X^2}{N}$, i.e., $V_1 \sim V_2 \sim ... \sim V_L \sim \mathcal{N}(0, N)$. The results are plotted for $\overline{R} = 0.25L, 0.5L, L$ when the noise variance of the observations is 10. The data points that correspond to each agent's estimation error are connected. As the total number of agents increases, the achievable distortion decreases since the CEO accumulates more data and can obtain a better source estimate. In addition, increasing the sum-rate results in less estimation distortion. However, the decrease in the distortion becomes negligible as the number of agents gets large.

In Fig. 3.7, for a network with 8 agents, the minimum achievable distortion as a function of sum-rate $\overline{R}$ for different noise levels is illustrated. In each case all agents have the same observation SNR. For all sum-rates the distortion decreases with decreasing the

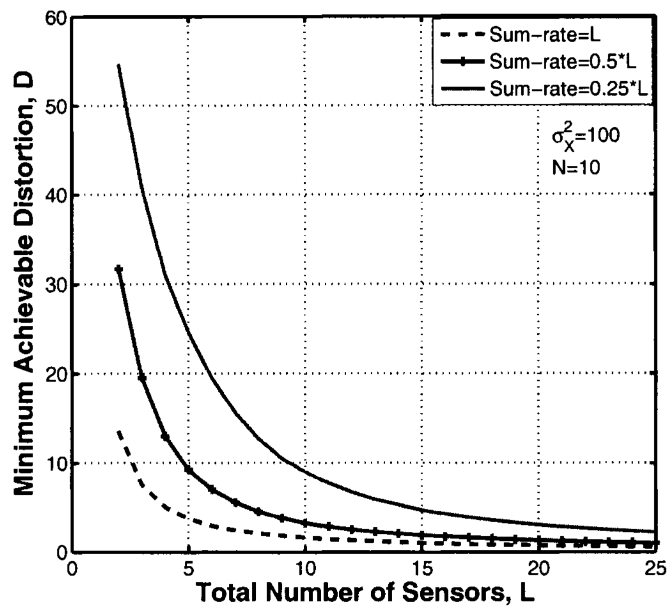

Figure 3.6: CEO estimation performance. Estimation distortion D as a function of total number of agents.

noise level of observations. The distortion tends to its minimum value, when the sum-rate passes a threshold. This minimum distortion is the limit of distortion when $\overline{R}$ tends to infinity. Therefore, having the sum-rate of *30* for our example will give the same distortion as the network with infinite rate agents.

Final estimation distortion versus sum-rate for different values of σ_X^2 is plotted in Fig. 3.8. The distortion decreases when σ_X^2 takes smaller values. We see that the distortions for all cases become the same for large sum-rates. This is because the lower bound of the minimum distortion is proportional to $\left(\frac{1}{\sigma_X^2} + \frac{L}{N}\right)^{-1}$ (see Equation (3.24)). Since for our example, $\frac{1}{\sigma_X^2} \ll \frac{L}{N}$, the achievable distortion for large value of sum-rate is the same for different source variances and is equal to $\left(\frac{L}{N}\right)^{-1} = 0.8^{-1} = 1.25$.

In Fig. 3.9, the optimal percentage of total rate assigned to each agent for the case of $L = 4$ agents all with the same $SNR = 5$ is illustrated. These percentages are approximately $100/L$ if the sum-rate is either very large or very small. This is the same as the results of [8] for the case of two agents. Therefore, in a general wireless network

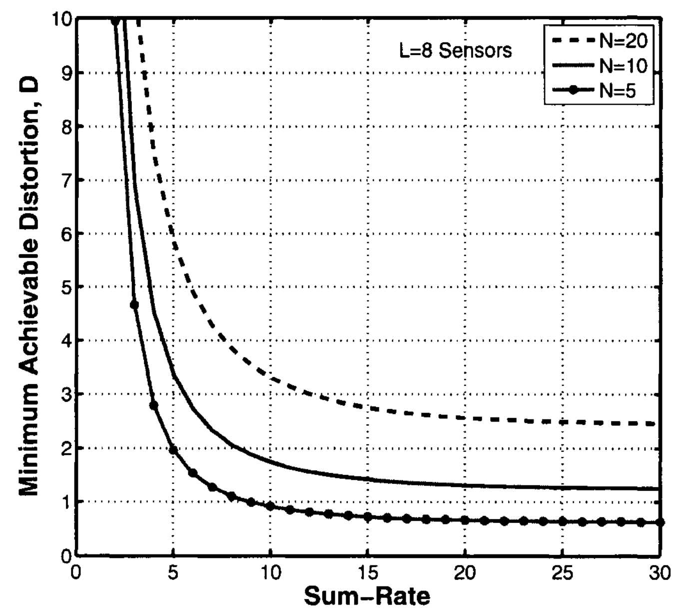

Figure 3.7: Estimation distortion versus the sum-rate for different noise levels.

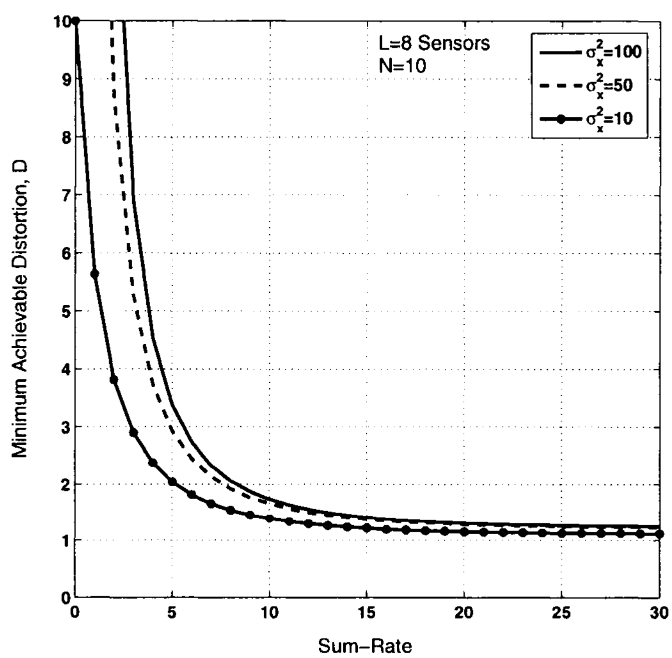

Figure 3.8: Estimation distortion versus the sum-rate for different source variances.

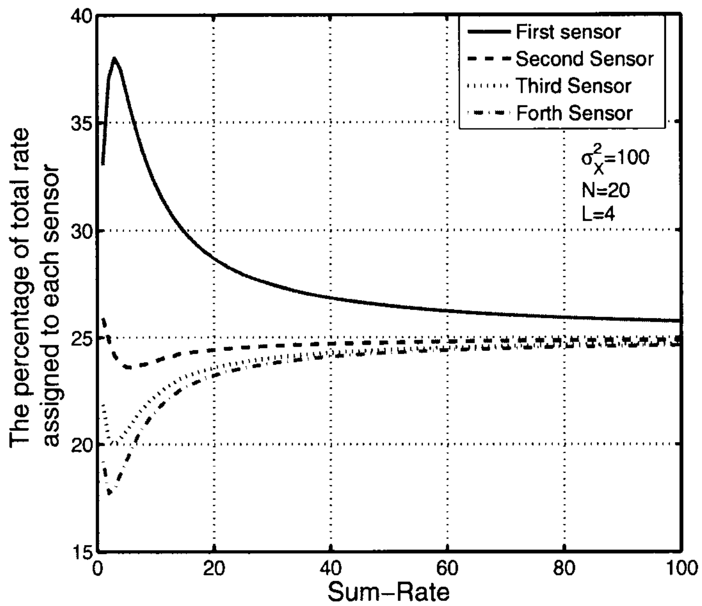

Figure 3.9: The percentage of total rate assigned to each agent for the network of $L = 4$ agents. All agents have the same noise level of 10. The source variance is fixed to 100.

of L agents, if the average rate per agent gets small or if the sum-rate is very large for a fixed L, the rate-allocation under a sum-rate constraint can be simplified by assigning equal-rate to agents. We will prove this result in Section 3.4.2.

Fig. 3.10 shows the case of $L = 3$ when the noise levels of agents are different. As $\overline{R}$ gets large, equal-rate allocation per agent yields the same distortion as the optimal case.

3.4 Rate Region, Rate Loss and Asymptotic Cases

3.4.1 Achievable Rate Region and Optimal Rate Allocation Region

Let $\mathcal{R}(D)$ denote the achievable rate region with respect to distortion D, i.e., the set of all achievable $(L+1)$-tuples $(R_1, R_2, ..., R_L, D)$. Assume $\pi = (\pi_1, \pi_2, ..., \pi_L)$ is a permutation of the set $\mathcal{I}_L = \{1, 2, ..., L\}$. Using the successive coding strategy, there

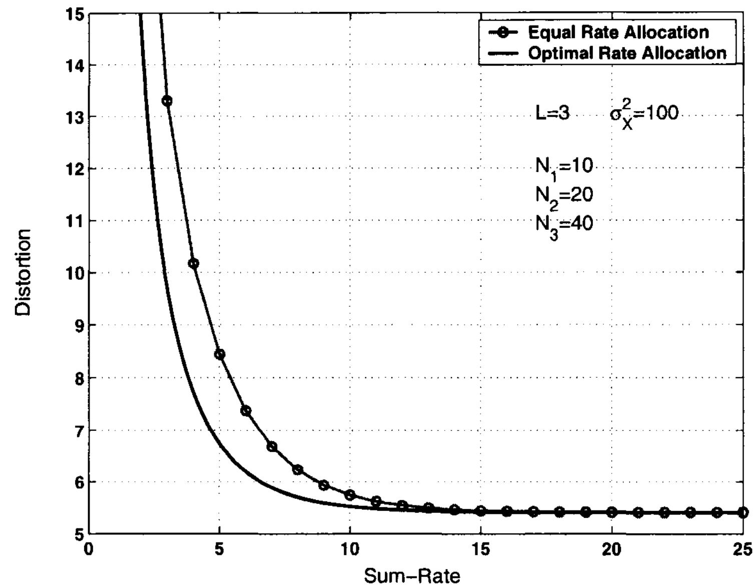

Figure 3.10: Distortion sum-rate tradeoff for the network with $L = 3$ agents. The noise levels of agents are different and the variance of the source is 100.

are $L!$ possible orderings for the coding/decoding process in the CEO problem. A given permutation π has the following rate region:

$$\mathcal{R}_\pi(D) = \left\{ \left(R_{\pi(1)}, R_{\pi(2)}, ..., R_{\pi(L)}\right) : R_{\pi(L)} \geq \frac{1}{2}\log_2 \left[\frac{D_{L-1}^2}{D\left(D_{L-1} + N_{\pi(L)}\right) - D_{L-1}N_{\pi(L)}}\right]\right\}, \tag{3.39}$$

where

$$D_i = \frac{D_{i-1}N_{\pi(i)}}{D_{i-1} + N_{\pi(i)}} + \left(\frac{D_{i-1}^2}{D_{i-1} + N_{\pi(i)}}\right) 2^{-2R_{\pi(i)}}, \; i = 1, ..., L-1 \tag{3.40}$$

and $D_0 = \sigma_X^2$. The lower convex hull of these $L!$ regions specifies the achievable rate region of the CEO problem, $\mathcal{R}(D)$. Any $(L+1)$-tuple $(R_1, R_2, ..., R_L, D)$ in this rate region is achievable based on the time-sharing among these $L!$ successive coding schemes.

Let $\mathcal{R}_{opt}(D)$ denote the set of all rate vectors $(R_1, R_2, ..., R_L)$ in $\mathcal{R}(D)$ that attain the sum-rate distortion function of $\overline{R}(D)$. This is indeed the optimal rate allocation region in the achievable rate region $\mathcal{R}(D)$. Each successive coding scheme has a set of optimal compression rates for the agents in the CEO problem in the sense that they attain the minimum distortion for a given sum-rate $\overline{R}$. Since the number of active agents depends on the noise variances, each of these $L!$ encoding/decoding schemes has a different number of active agents, M_π. M_π is the largest integer value between 1 and L that satisfies the equation $\frac{M_\pi}{\max\{N_{\pi(1)},\ldots,N_{\pi(M_\pi)}\}} \geq \left(\frac{1}{D^*_\pi(M_\pi)} - \frac{1}{D}\right)$ where $D^*_\pi(M_\pi) = \left(\frac{1}{\sigma_X^2} + \sum_{i=1}^{M_\pi} \frac{1}{N_{\pi(i)}}\right)^{-1}$. By assigning zero rate to non-active agents, there are $L!$ set of optimal rates $\left(R_{\pi(1)}, R_{\pi(2)}, \ldots, R_{\pi(L)}\right)$. The optimal rate allocation region $\mathcal{R}_{opt}(D)$ is the convex hull of these $L!$ vertices $\left(R_{\pi(1)}, R_{\pi(2)}, \ldots, R_{\pi(L)}\right)$. The coordinates of these vertices can be obtained as

$$\begin{cases} R_{\pi(j)} = \frac{1}{2}\log_2\left(\frac{M_\pi}{N_{\pi(j)}} \frac{S_{\pi(j)}}{S_{\pi(j-1)}} \frac{1}{\left(\frac{1}{D^*_\pi(M_\pi)} - \frac{1}{D}\right)}\right) & j = 1, 2, \ldots, M_\pi \\ R_{\pi(j)} = 0 & j = M_\pi + 1, \ldots, L \end{cases} \tag{3.41}$$

where

$$S_{\pi(j)} = M_\pi\left(\frac{1}{\sigma_X^2} + \sum_{i=1}^{j} \frac{1}{N_{\pi(i)}}\right) - j\left(\frac{1}{D^*_\pi(M_\pi)} - \frac{1}{D}\right). \tag{3.42}$$

Example: Rate Region of 2-Agent CEO Problem

To apply the successive coding strategy in the 2-agent CEO problem, there are two possible orderings for the encoding/decoding process: (a) the message from node 1 is designed to be decoded first, and (b) the message from node 2 is designed to be decoded first. Ordering (a) minimizes the compression rate R_2 and ordering (b) minimizes the compression rate R_1. The rate-region of the 2-agent CEO problem is shown in Fig. 3.11. The solid line AB is where the sum-rate is minimized. The region above the dash line is the achievable rate region where R_2 is minimized (ordering(a)) and the region above the dash-dot line is the achievable rate region where R_1 is minimized (ordering (b)). In fact, the lower convex hull for both rate-regions of ordering (a) and (b) will identify the rate region. Note that the rate allocations of (3.17) which is the optimal rate allocation under a sum-rate constraint for ordering (a) corresponds to the point A in Fig. 3.11. Therefore, the optimal rate allocation scheme that achieve distortion D using minimum sum-rate $\overline{R}$

corresponds to the points on line AB where

$$\begin{cases} R_{1B}(D) = \frac{1}{2}\log\left(\frac{4}{DN_1\left(\frac{1}{D^*}-\frac{1}{D}\right)\left(2\left(\frac{1}{\sigma_X^2}+\frac{1}{N_2}\right)-\left(\frac{1}{D^*}-\frac{1}{D}\right)\right)}\right) \\ R_{2B}(D) = \frac{1}{2}\log\left(\frac{2\left(\frac{1}{\sigma_X^2}+\frac{1}{N_2}\right)-\left(\frac{1}{D^*}-\frac{1}{D}\right)}{\frac{N_2}{\sigma_X^2}\left(\frac{1}{D^*}-\frac{1}{D}\right)}\right) \end{cases} \tag{3.43}$$

Allowing time-sharing between ordering (a) and (b) achieves all the intermediate points of line AB.

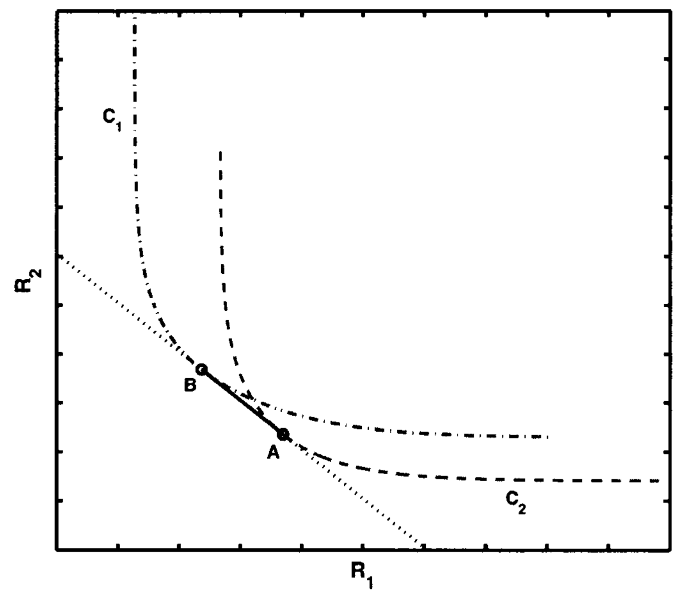

Figure 3.11: The rate-region under a fixed distortion D for the 2-agent CEO problem. The dash-dot line shows the equation (3.44) where R_1 is minimized (ordering (b)). The dash line shows the equation (3.45) where R_2 is minimized (ordering (a)). Line AB is the boundary of achievable rate region where the sum-rate $R_1 + R_2$ is minimized (equation (3.46)).

The whole rate region of the 2-agent CEO problem can be completely characterized as the lower convex envelope of the following regions:

$$R_1 \geq \frac{1}{2}\log^+\left(\frac{\frac{\sigma_X^2}{DN_1}\left(1+\frac{\sigma_X^2}{N_2}\right)^{-1}\left(1+\frac{\sigma_X^2}{N_2}2^{-2R_2}\right)^2}{\left(\frac{1}{D^*}-\frac{1}{D}\right)+\frac{\sigma_X^2}{N_2}2^{-2R_2}\left(\frac{1}{N_1}-\frac{1}{D}\right)}\right), \tag{3.44}$$

$$R_2 \geq \frac{1}{2}\log^+\left(\frac{\frac{\sigma_X^2}{DN_2}\left(1+\frac{\sigma_X^2}{N_1}\right)^{-1}\left(1+\frac{\sigma_X^2}{N_1}2^{-2R_1}\right)^2}{\left(\frac{1}{D^*}-\frac{1}{D}\right)+\frac{\sigma_X^2}{N_1}2^{-2R_1}\left(\frac{1}{N_2}-\frac{1}{D}\right)}\right), \tag{3.45}$$

$$R_1 + R_2 \geq \frac{1}{2}\log^+\left(\frac{4\sigma_X^2}{DN_1N_2\left(\frac{1}{D^*}-\frac{1}{D}\right)^2}\right). \tag{3.46}$$

From the perspective of robustness, this scheme performs better than the joint decoding scheme [61], i.e., no matter which node fails, the CEO can obtain a nontrivial estimate of the source.

3.4.2 Asymptotic Cases

We consider an equal *SNR* network, $N_1 = N_2 = ..., N_L = N$. Two asymptotic cases are considered: $L \to \infty$ and $\overline{R} \to \infty$. The first case is also investigated by Draper [8]. We show that if the average rate per agent is small or if the sum-rate $\overline{R}$ is very large for a fixed L, then the performance of equal rate allocation for the successive coding strategy converges to the rate-distortion function.

3.4.2.1 Large Number of Agents (Nodes)

For the network with equal SNR nodes, the sum-rate distortion function can be expressed as

$$\overline{R}(D) = \frac{1}{2}\log_2\left\{\frac{\sigma_X^2}{D}\right\} + \frac{L}{2}\log_2\left\{\frac{\frac{L}{N}}{\frac{L}{N}+\frac{1}{\sigma_X^2}-\frac{1}{D}}\right\}. \tag{3.47}$$

Rewriting Equation (3.30) in terms of (3.47), we can show that

$$R_j = \frac{\overline{R}}{L} - \frac{1}{2L}\log_2\left\{\frac{\sigma_X^2}{D}\right\} + \frac{1}{2}\log_2\left\{\frac{\frac{L}{\sigma_X^2}+j\left(\frac{1}{D}-\frac{1}{\sigma_X^2}\right)}{\frac{L}{\sigma_X^2}+(j-1)\left(\frac{1}{D}-\frac{1}{\sigma_X^2}\right)}\right\}. \tag{3.48}$$

As the number of nodes tends to infinity, $L \to \infty$, the sum-rate distortion function of (3.47) can be expressed as

$$\lim_{L\to\infty} \overline{R}(D) = \frac{1}{2}\log_2\left\{\frac{\sigma_X^2}{D}\right\} + \frac{N}{2}\left(\frac{1}{D} - \frac{1}{\sigma_X^2}\right)(\log_2 e)\,. \tag{3.49}$$

Thus, using (3.48) we can show that $\lim_{L\to\infty, j\neq k} \frac{R_j}{R_k} = 1$. As a result, the optimal rate allocation for the successive coding strategy allocates an equal rate of $\frac{\overline{R}}{L}$ to each agent. In other words, when the average rate per node approaches zero, the equal rate allocation scheme for the successive coding strategy does not cause any extra distortion compared to the optimal rate allocation scheme. The same result is presented in [8] but the approach is different. Since they did not have the expression for the optimal rates, they showed that the equal rate allocation of $R_j = \frac{\overline{R}}{L}$ combining with the iterative expression of D_j will result in the sum-rate of (3.49) as L grows to infinity.

3.4.2.2 Large Sum-Rate

From (3.28), we know that allowing sum-rate to grow to infinity is equivalent to assigning the final distortion D_L to the lower bound of (3.24). Therefore, using (3.48), the optimal rate allocation for the large values of sum-rate can be written as

$$R_j = \frac{\overline{R}}{L} - \frac{1}{2L}\log_2\left\{1 + \frac{L\sigma_X^2}{N}\right\} + \frac{1}{2}\log_2\left\{\frac{\frac{1}{\sigma_X^2} + j\frac{1}{N}}{\frac{1}{\sigma_X^2} + (j-1)\frac{1}{N}}\right\}. \tag{3.50}$$

Thus, $\lim_{\overline{R}\to\infty, j\neq k} \frac{R_j}{R_k} = 1$. Hence, as the sum-rate grows to infinity, the optimal rate allocation for the successive coding allocates an equal rate of $\frac{\overline{R}}{L}$ to each agent.

3.4.2.3 Numerical Results

To compare the behavior of equal rate allocation for intermediate number of agents (sensors) and intermediate values of sum-rate, we provide some numerical results for a wireless sensor network modeled by the Gaussian CEO problem. The similar plot for intermediate number of sensors is presented in [8]. We assume that the target data is an *i.i.d.* zero mean Gaussian process with variance $\sigma_X^2 = 100$. If the distortion achieved by equal-rate allocation is represented by D_{eql} and the minimum achievable distortion of the optimal

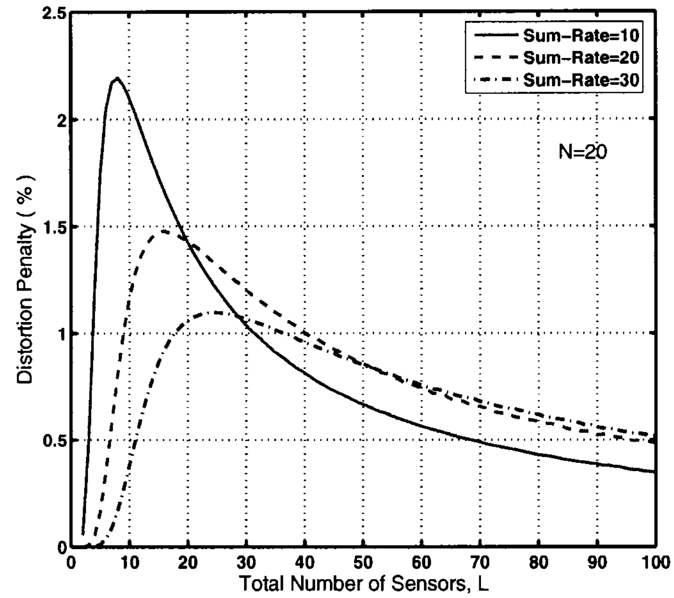

Figure 3.12: The distortion penalty caused by assigning equal rates to sensors versus total number of sensors, parametrized by the sum-rate, $\overline{R}$. All the sensors have the same noise variance, $N = 20$.

rate allocation is termed by D_{min}, the percentage of extra distortion caused by assigning equal rates to sensors can be calculated as

$$Distortion \quad Penalty(\%) = \frac{D_{eql} - D_{min}}{D_{min}} \times 100. \tag{3.51}$$

The percentage of distortion penalty versus total number of sensors in the network is shown in Fig. 3.12. The noise variance is fixed at $N = 20$. We plot the results for three different sum-rates, $\overline{R} = 10, 20$, and 30. It is observable that this equal-rate allocation scheme may not cause a large extra distortion compared to the minimum achievable distortion. The penalty approaches zero as the number of sensors grows to infinity.

We plot the distortion penalty versus the sum-rate, parametrized by the number of sensors in Fig. 3.13. We see that for all three values of L, the distortion penalty vanishes as the sum-rate grows to large values.

Hence, the fraction of the total rate allocated to each sensor is approximately $1/L$ if

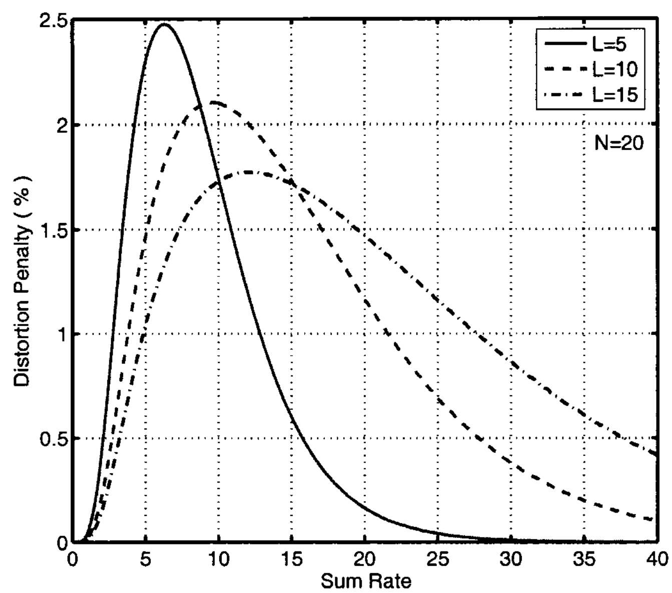

Figure 3.13: The percentage of excess distortion caused by assigning equal rates to sensors (agents) versus the sum-rate, parametrized by the total number of sensors. All the sensors have the same noise variance, $N = 20$.

the sum-rate $\overline{R}$ is very large for a fixed L or if L is very large for a fixed sum-rate. Thus, we can simplify the rate allocation problem in a general parallel network with L agents by assigning equal rates to agents. This scheme may not cause a large extra distortion compared to the minimum achievable distortion.

3.4.3 Rate Loss in the Gaussian CEO problem

In this part, we investigate the effect of "distributed" source coding compared with the joint encoding of correlated data sequences. We show that if the agents in the CEO problem are allowed to collaborate, then we require a lower rate for achieving the same distortion. The latter scheme which is called the full cooperative CEO problem or the remote source coding problem [3] is shown in Fig. 3.14. The sum-rate distortion function of this scheme can be obtained as follows:

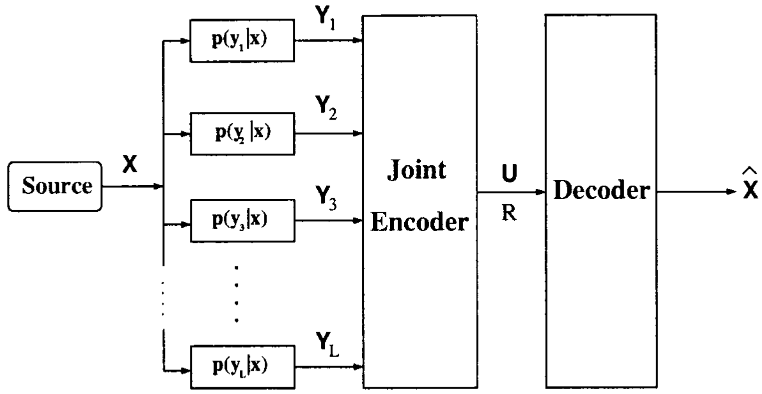

Figure 3.14: Remote source coding problem. The joint encoder has access to all observations.

Lemma 3 The sum-rate distortion function for the Gaussian remote source coding problem, depicted in Fig. 3.14, can be expressed as

$$R_X^{remote}(D) = \frac{1}{2}\log_2\left(\frac{\sigma_X^2}{D}\frac{\left(\sum_{i=1}^{L}\frac{1}{N_i}\right)}{\left(\frac{1}{\sigma_X^2}+\sum_{i=1}^{L}\frac{1}{N_i}-\frac{1}{D}\right)}\right) \qquad \left(\frac{1}{\sigma_X^2}+\sum_{i=1}^{L}\frac{1}{N_i}\right)^{-1} < D < \sigma_X^2 \tag{3.52}$$

Proof The rate-distortion function for the problem of Fig. 3.14 can be defined as

$$R_X^{remote}(D) = \min_{U\in U_X(D)} I(Y_1, Y_2, ..., Y_L; U) \tag{3.53}$$

where $U_X(D) = \{U : X \rightarrow Y_i \rightarrow U; E[d(X,U) \leq D\}$. Based on the Wyner-Ziv source coding for noisy observations [8] we can calculate (3.53) and obtain the following rate-distortion tradeoff:

$$R_X^{remote}(D) = \frac{1}{2}\log_2\left(\frac{\sigma_X^2-\sigma_{X|Y_1,Y_2,...,Y_L}^2}{D-\sigma_{X|Y_1,Y_2,...,Y_L}^2}\right) = \frac{1}{2}\log_2\left(\frac{\sigma_X^2-\left(\frac{1}{\sigma_X^2}+\sum_{i=1}^{L}\frac{1}{N_i}\right)^{-1}}{D-\left(\frac{1}{\sigma_X^2}+\sum_{i=1}^{L}\frac{1}{N_i}\right)^{-1}}\right), \tag{3.54}$$

which results in (3.52).

Comparison of (3.28) with (3.52) shows that there is a rate loss when the agents are not allowed to collaborate. The rate loss can be expressed as [72]

$$R_X^{Loss}(D) = \overline{R}(D) - R_X^{remote}(D). \tag{3.55}$$

By subtracting (3.52) from (3.28), we obtain the rate loss of

$$R_X^{Loss}(D) = \frac{1}{2}\log_2\left\{\frac{\frac{1}{\sigma_X^2}+\sum_{i=1}^{L}\frac{1}{N_i}-\frac{1}{D}}{\left(\frac{1}{\sigma_X^2}+\sum_{i=1}^{M}\frac{1}{N_i}-\frac{1}{D}\right)^M}\right\} + \frac{1}{2}\log_2\left\{\frac{M^M}{\left(\prod_{i=1}^{M}N_i\right)\left(\sum_{i=1}^{L}\frac{1}{N_i}\right)}\right\}. \tag{3.56}$$

When we have a homogeneous network, i.e., all the agents have the same $SNR(=\sigma_X^2/N)$, then all L agents are active and $M = L$. For this case the rate loss is obtained as

$$R_X^{Loss}(D) = \frac{L-1}{2}\log_2\left\{\frac{\frac{L}{N}}{\frac{1}{\sigma_X^2}+\frac{L}{N}-\frac{1}{D}}\right\}. \tag{3.57}$$

We observe that when the number of agents tends to infinity, the rate loss of the CEO problem is

$$\lim_{L\to\infty} R_X^{Loss}(D) = \frac{N}{2}\left(\frac{1}{D}-\frac{1}{\sigma_X^2}\right)(\log_2 e). \tag{3.58}$$

In fact, in this limit, the remote source coding achieves the standard quadratic Gaussian rate distortion function, $R_X^{remote}(D) = \frac{1}{2}\log_2\left(\frac{\sigma_X^2}{D}\right)$, while the CEO problem achieves the sum-rate distortion of (3.49) which incurs a rate loss of (3.58). We also observe that as the average distortion D decreases the rate loss of the CEO problem increases.

3.5 Application

In this section, we consider the application of our results for computing the communication rate throughput of relay channels. The relay channel first has been introduced by Van der Meulen [73]. Capacity of a single-relay channel has been found by Cover and El Gamal [74].

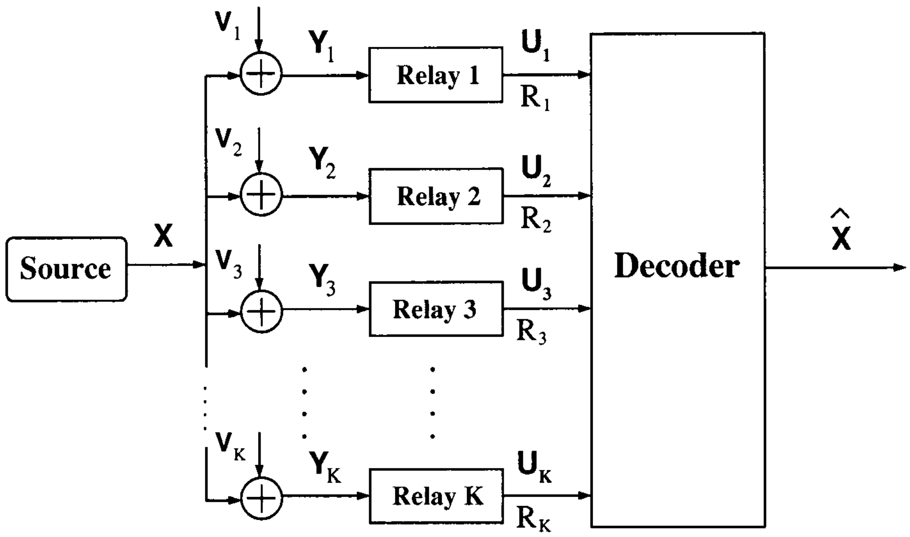

Figure 3.15: Parallel Gaussian $K-$relay network.

We consider the parallel Gaussian $K-$relay network of Fig. 3.15 that consists of a single source-destination pair and K relays [45, 75, 76], where the source has the power P, the K relays observe the source corrupted by AWGN of variance N_i and the central decoder (the FC) receives the messages from the relays through rate-constrained noiseless channels. We assume that the relays transmit rate-constrained messages to the CEO under a sum-rate constraint $\overline{R}$.

To detect the transmitted message at the decoder, there are two different schemes [45]: first scheme is to use broadcast codes for communication to the relays. Let R_{BC} denote the maximum reliable sum-rate to the relays. Then, the communication rate throughput (CRT) of this strategy would be $min\{R_{BC}, \overline{R}\}$. In the second scheme which is called "Estimate and Detect" strategy [45], codewords are generated in an *i.i.d.* Gaussian manner. Then, the relays based on their observations transmit bit streams to the FC. The FC makes an estimate of the codeword and then detects the message based on its estimate. Since the output of the decoder $\widehat{X}$ and codeword are jointly typical, standard typicality decoding will work [8]. If we denote this strategy by EstDet, the CRT of this strategy is

$$\begin{aligned} R_{EstDet} &= I(X;U_1,...,U_K) = h(X) - h(X \mid U_1,...,U_K) \\ &= \tfrac{1}{2}\log\left[2\pi e\sigma_X^2\right] - \tfrac{1}{2}\log\left[2\pi e\sigma^2_{X|U_1,...,U_K}\right] = \tfrac{1}{2}\log\left[\tfrac{P}{D}\right] \end{aligned} \quad (3.59)$$

where $U_1, ..., U_K$ are the auxiliary random variables of the K relays and D is the final distortion of (3.28) with $\sigma_X^2 = P$. Therefore, for each value of D, we can obtain R_{EstDet} and $\overline{R}$ using (3.59) and (3.28), respectively. As a result, CRT of this strategy versus $\overline{R}$ can be obtained. Note that in the CEO problem, the minimum sum-rate is the minimum of total information exchanged between encoders and the decoder, i.e., $\overline{R} = \min I(Y_1, ..., Y_K; U_1, ..., U_K)$.

To compare the two strategies, we consider the relay network in which the relays have the same noise variances, i.e., $N_1 = N_2 = ..., N_L = N$. Thus, $R_{BC} = \frac{1}{2}\log\left[1 + \frac{P}{N}\right]$. Fig. 3.16 shows the CRT of the two strategies versus $\overline{R}$. The value of SNR$(= \frac{P}{N})$ is fixed to 40. As the number of relays increases the CRT of EstDet strategy increases while the CRT of broadcast coding does not change. Also, for large values of $\overline{R}$, the EstDet strategy outperforms broadcast coding since the EstDet can exploit the diversity of the relay observations. Our result is an extension of the result of [45], where the CRT versus the sum-rate for a parallel two-relay network is derived.

3.6 Chapter Summary

We considered a distributed network, modeled by the CEO problem, in which L agents communicate their observations of the target data sequence to the FC using limited transmission rate. We used the successive Wyner-Ziv coding strategy and obtained the optimal sum-rate distortion tradeoff and optimal rate allocation scheme for the Gaussian CEO problem. Our result indicates that the sum-rate distortion function of the CEO problem is attained using the successive coding strategy which is a less complex way than the joint decoding method. We demonstrated that if the average rate per agent is small or if the sum-rate is very large for a fixed L, then the performance of equal rate allocation scheme for the successive coding strategy converges to the rate-distortion function.

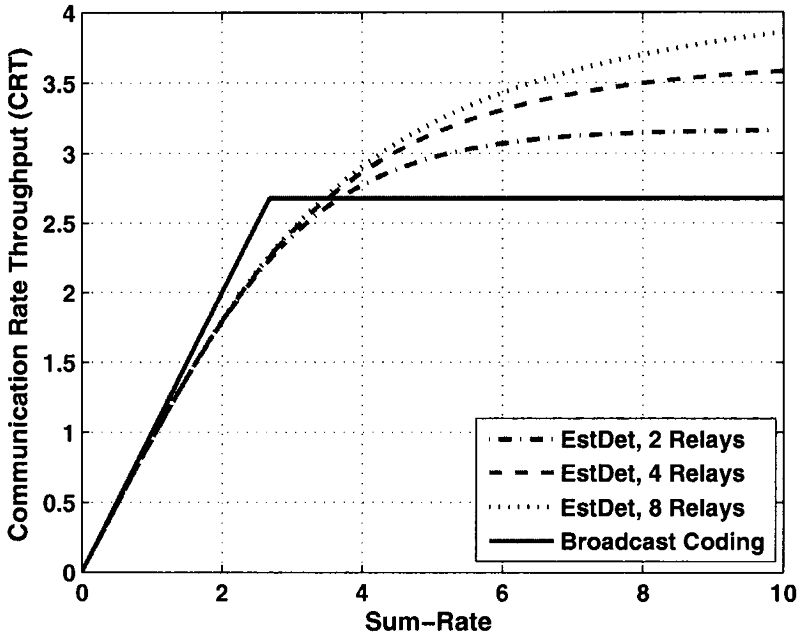

Figure 3.16: CRT of a Gaussian relay network achieved by two approaches: Broadcast Coding and Estimate and Detect (EstDet) strategy. SNR is fixed to 40.

Appendix I- Proof of Lemma 1

We use the induction method. As the base case, it can be shown that the result is true for $n = 2$ by taking derivative of D_n using Equation (3.5). Now, as the induction hypothesis, we assume that this result is true for $n \geq 2$ and we need to show that it is true for $n + 1$. From (3.5) D_{n+1} can be expressed as:

$$D_{n+1} = \frac{D_n N_{n+1} + D_n^2 2^{-2R_{n+1}}}{D_n + N_{n+1}} \tag{3.60}$$

For $k \leq n$, D_n is a function of R_k. Therefore,

$$\frac{\partial D_{n+1}}{\partial R_k} = \frac{\left(N_{n+1}\frac{\partial D_n}{\partial R_k} + 2D_n\frac{\partial D_n}{\partial R_k}2^{-2R_{n+1}}\right)}{(D_n + N_{n+1})} - \frac{\left(\frac{\partial D_n}{\partial R_k}\right)\left(D_n N_{n+1} + D_n^2 2^{-2R_{n+1}}\right)}{(D_n + N_{n+1})^2} \tag{3.61}$$

This expression can be simplified as follows:

$$\frac{\partial D_{n+1}}{\partial R_k} = \frac{\frac{\partial D_n}{\partial R_k}\left(D_n^2 + 2D_nN_{n+1} + N_{n+1}^2 2^{2R_{n+1}}\right)}{\left(D_n + N_{n+1}\right)^2 2^{2R_{n+1}}} \tag{3.62}$$

Substituting (3.20) in (3.62) leads to

$$\frac{\partial D_{n+1}}{\partial R_k} = \frac{(-2\log 2)\, D_{k-1}^2 2^{-2R_k}}{D_{k-1} + N_k} \prod_{i=k}^{n} \frac{D_i^2 + 2N_{i+1}D_i + N_{i+1}^2 2^{2R_{i+1}}}{\left(D_i + N_{i+1}\right)^2 2^{2R_{i+1}}} \tag{3.63}$$

for $k \leq n$. To complete the proof, we just need to show that the result is also true for $k = n + 1$. Since D_n is not a function of R_{n+1}, by using (3.60), we obtain

$$\frac{\partial D_{n+1}}{\partial R_{n+1}} = \frac{D_n^2\,(-2\log 2)\, 2^{-2R_{n+1}}}{D_n + N_{n+1}} \tag{3.64}$$

This completes the proof.

Appendix II- Proof of Lemma 2

First, we prove that our result holds for R_1 and D_1. We use the induction method. As the base case, we can show that the result is true for $M = 2$ by solving equations $\frac{\partial D_2}{\partial R_1} = \frac{\partial D_2}{\partial R_2}$ and $\sum_{i=1}^{2} R_i = \overline{R}$ where D_2 can be obtained from (3.5). The proof of this base case is presented in Section 3.3.1. As the induction hypothesis, we assume that R_1 is correct for the case of M agents. Now, if we consider the case of $M + 1$ agents with optimal rates of $R_1^{'}, R_2^{'}, ..., R_{M+1}^{'}$, the formula of $R_2^{'}$ is the same as R_1 where σ_X^2 is replaced by D_1, $(N_1, ..., N_M)$ are replaced by $(N_2, N_3, ..., N_{M+1})$ and D_M is replaced by D_{M+1}. Therefore,

$$R_2^{'} = \frac{1}{2}\log_2\left(\frac{M \times \left(\frac{1}{D_1} + \frac{1}{N_2}\right) - B_1}{N_2 \times B_1/D_1}\right) \tag{3.65}$$

where $B_1 = \left(\frac{1}{D_1} + \sum_{i=2}^{M+1}\frac{1}{N_i} - \frac{1}{D_{M+1}}\right)$. Now we use Corollary 2 to obtain rate $R_1^{'}$ in terms of $R_2^{'}$, D_1 and $D_0 = \sigma_X^2$, i.e.,

$$2^{2R_1'} = \frac{\sigma_X^4 \left(D_1^2 + 2N_2D_1 + N_2^2 2^{2R_2'}\right)}{D_1^2 (\sigma_X^2 + N_1)(D_1 + N_2)} \tag{3.66}$$

By substituting D_1 of (3.5) and R_2' of (3.65) in (3.66) and doing some manipulations we obtain R_1' as

$$R_1' = \frac{1}{2}\log_2 \left(\frac{(M+1)\left(\frac{1}{\sigma_X^2} + \frac{1}{N_1}\right) - B_2}{N_1 \times B_2/\sigma_X^2} \right) \tag{3.67}$$

where $B_2 = \left(\frac{1}{\sigma_X^2} + \sum_{i=1}^{M+1} \frac{1}{N_i} - \frac{1}{D_{M+1}}\right)$. This completes the proof for R_1. Using (3.5) and the expression for R_1, D_1 can be obtained as $D_1 = \frac{M}{S_1}$. This result shows that the expression for D_j is correct for the case of $j = 1$. We need to prove that (3.23) and (3.27) are valid for all values of $j > 1$. We consider (D_{j-1}, R_j) as F_j. We assume that we know F_j and we want to obtain F_{j+1}. Using (3.5), the value of D_j can be obtained. We need to obtain R_{j+1} and D_{j+1} from (D_{j-1}, R_j, D_j). We use Corollary 2 to rewrite R_{j+1} in terms of R_j, D_j and D_{j-1}:

$$R_{j+1} = \frac{1}{2}\log_2 \left[\frac{\left(\frac{M}{S_j}\right)^2 \frac{MS_j}{N_j A S_{j-1}} \left(\frac{M}{S_{j-1}} + N_j\right)\left(\frac{M}{S_j} + N_{j+1}\right) - \frac{M^4}{S_j^2 S_{j-1}^2} - \frac{2M^3 N_{j+1}}{S_j S_{j-1}^2}}{N_{j+1}^2 \frac{M^2}{S_{j-1}^2}} \right] \tag{3.68}$$

where $A = \left(\frac{1}{\sigma_X^2} + \sum_{i=1}^{M} \frac{1}{N_i} - \frac{1}{D_M}\right)$. Since $S_j - S_{j-1} = \frac{M}{N_j} - A$, by doing some manipulations we obtain

$$R_{j+1} = \frac{1}{2}\log_2 \frac{M^2 + MN_{j+1}S_j - MAN_{j+1}}{AS_jN_{j+1}^2} = \frac{1}{2}\log_2 \frac{MS_{j+1}}{AS_jN_{j+1}} \tag{3.69}$$

Comparing (3.69) with (3.23) we see that our result is correct. Now we obtain D_{j+1} from the iterative expression of (3.5). We substitute $R_{j+1} = \frac{1}{2}\log_2 \frac{MS_{j+1}}{AS_jN_{j+1}}$ and $D_j = \frac{M}{S_j} = \frac{M}{S_{j+1} + A - \frac{M}{N_{j+1}}}$ in (3.5) for $i = j+1$. After some manipulations we obtain $D_{j+1} = \frac{M}{S_{j+1}}$. This completes the proof.

Chapter 4

Successively Structured Gaussian Multiterminal Source Coding Schemes

4.1 Introduction

Multiterminal source coding or distributed data compression refers to separate lossy encoding and joint decoding of multiple correlated sources [16]. Consider the following scenario for the signal detection: a sensor is used to measure a physical phenomenon and transmit limited rate information to a FC for further processing. Also, suppose another sensor is able to measure the phenomenon but it is farther from the phenomenon than the first sensor and hence it is not worth to reproduce its signal at the FC. The scenario is shown in Fig. 4.1. Since both sensors measure the same phenomenon, their information is correlated. Therefore, the second sensor can provide partial side information to reduce the rate required at the first sensor and hence save transmission energy for a given distortion. In fact, the information of the second sensor can be used at the FC to obtain higher reliability and lower probability of detection error. This motivating application requires the theory of separate coding of correlated sources.

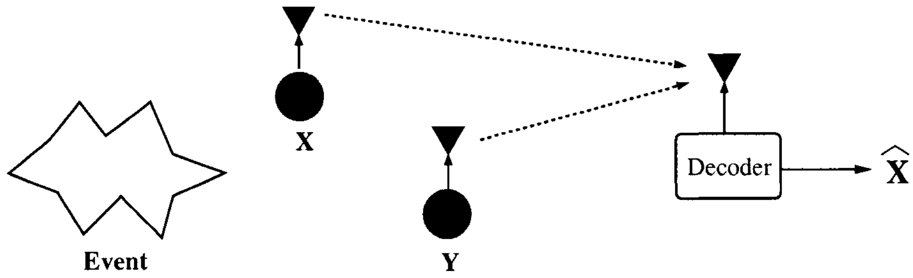

Figure 4.1: *1-helper* scenario. Two sensors are deployed to measure a physical phenomenon (an event). At the decoder, we are only interested in the signal reconstruction of the sensor X which is nearer to the event than other sensor, Y.

4.1.1 Previous Work

The rate-distortion theory for the multiterminal source coding problem was first studied by Wyner and Ziv [9]. They obtained the rate-distortion function of a single source when the decoder can observe full resolution side information about the source. The important fact about the Wyner-Ziv coding is that it usually suffers rate loss compared to the lossy coding of X when the side information Y is available at both the encoder and the decoder. One exception is when X and Y are jointly Gaussian with MSE distortion measure. There is no rate loss with the Wyner-Ziv coding in this case, which is of special interest in some applications such as video sensor networks since many image and video sources (after mean abstraction) can be modeled as jointly Gaussian [44]. This result is the dual of Costa's dirty paper theorem for the channel coding with side information at the transmitter only [77]. In this work, we also focus on the Gaussian case. The generalized Wyner-Ziv source coding for noisy encoder observations is appeared in [14]. This problem is also known as the noisy Wyner-Ziv coding, i.e., lossy coding of noisy observations with side information available at the decoder and not at the encoder. Similar rate-distortion analysis is presented in [15], and recently in [8]. Systematic lossy source/channel coding with uncoded side information at the decoder is presented in [78]. Berger and Yeung [41] have solved the rate-distortion problem for two sources when both of them are reconstructed. They consider the situation in which the reconstruction of X must be perfect, while that of Y is subject to a distortion criterion. Oohama [42] gives the solution of the Wyner-Ziv problem with coded side information for the case of two sources. He derives an outer region for the rate-distortion region of the *2-terminal* source coding problem, depicted in Fig. 4.2, and demonstrates that the inner region obtained by Berger [16] and Tung [50] is partially tight [42]. Wagner *et al.* complete the characterization of the rate region for the *2-terminal* source coding problem by showing that the inner bound of Berger and Tung is also tight in the sum-rate [18]. Oohama extends his results to more than two sources for a certain class of $m+1$ correlated sources, where m source signals are independent noisy versions of the main source, i.e., $X_i = X_0 + N_i$, $i \in \{1, 2, ..., m\}$. In other words, sources X_1 to X_m are conditionally independent given the source X_0 [79].

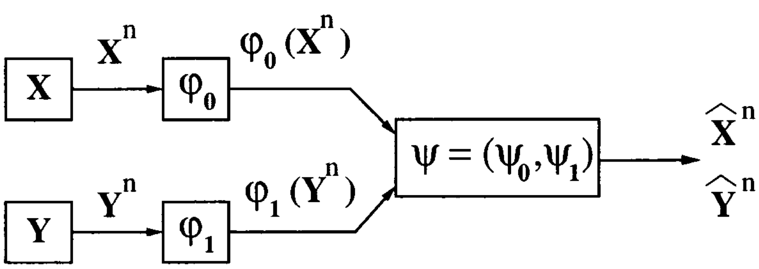

Figure 4.2: *2-terminal* source coding. X and Y are two correlated sources.

Consider the case of two correlated sources, where one source plays the role of partial side information to help the decoder to reconstruct the transmitted sequence of the other source within a prescribed average distortion. This problem, which is also called *1-helper* problem, for the correlated memoryless Gaussian sources and squared distortion measures was investigated in [42]. Oohama shows that his outer bound for the *2-terminal* source coding when combined with the inner bound of Berger and Tung determines the rate-distortion function of the *1-helper* problem. Oohama extends his results to more than two sources for a certain class of $m+1$ correlated sources, where sources X_1 to X_m are conditionally independent given the source X_0 [31, 34]. The result is extended to the case where $m+1$ sources satisfy a kind of tree structure on their correlation [51]. This condition contains the conditionally independent condition of [31, 34] as a special case. In [35], the general correlation structure for the *m-helper problem* is considered. For this case, they derive a lower bound on the rate-distortion function.

In this work, we consider the successive coding strategy for the *1-helper* and *2-terminal* source coding problems. The successively structured *2-terminal* source coding problem is similar to the sequential coding of two correlated sources [80], where for sources X and Y, the encoding of Y depends on X but the encoding of X does not depend on Y (illustrated in Fig. 4.3).

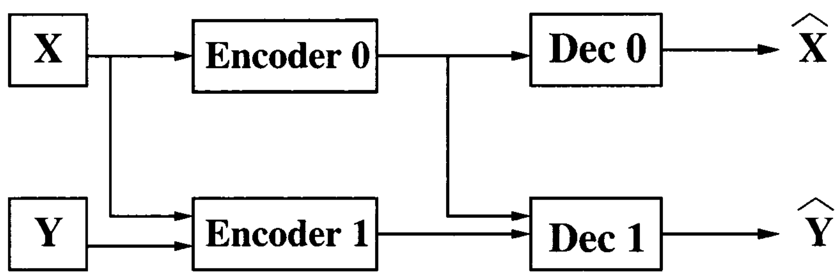

Figure 4.3: Sequential coding of correlated sources.

4.1.2 Main Contribution

Our main contributions can be summarized as follows: We show that for both the Gaussian *1-helper* problem and the Gaussian *2-terminal* source coding problem, successively structured Wyner-Ziv codes can achieve the rate-distortion bounds. In fact, by developing connections between source encoding and data fusion steps, it is shown that the whole rate-distortion region for the *2-terminal* source coding problem is achievable using the successive coding strategy. Therefore, the high complexity optimal source code can be decomposed into a sequence of low complexity Wyner-Ziv codes. By comparing the results of the successive coding strategy for the *2-terminal* source coding and the sequential coding of correlated Gaussian sources, we demonstrate that there is no sum-rate loss when the output of the first encoder is not available at the second encoder. The result is of special interest in some applications such as video coding where there are processing and storage constraints at the encoder. This successive coding approach leads us to derive an inner bound for the rate region of the *m-terminal* source coding where all the sources are reconstructed at the fusion center (FC) with specified distortions.

The rest of this chapter is organized as follows: In Section 4.2, we present the system model and problem formulation. In Section 4.3, we use the successive coding strategy in order to obtain the rate-distortion regions of the *1-helper* and *2-terminal* source coding schemes. We generalize the achievable rate-distortion region of the successive coding for the *m-terminal* problem in Section 4.4. Summary of the chapter is given in Section 4.5.

4.2 Problem Formulation

Let X and Y be correlated Gaussian random variables such that $\{(X_t, Y_t)\}_{t=0}^{\infty}$ are jointly stationary Gaussian memoryless sources. For each observation time $t = 1, 2, 3, ...,$ the random pair (X_t, Y_t) takes a value in real space $\mathcal{X} \times \mathcal{Y}$ and has a probability density function (pdf) $p_{X,Y}(x, y)$ of $\mathcal{N} \sim (0, \Lambda)$ where the covariance matrix Λ is given by

$$\Lambda = \begin{bmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix} \quad , \ -1 < \rho < 1 \tag{4.1}$$

We represent n independent instances of $\{X_t\}_{t=1}^{\infty}$ and $\{Y_t\}_{t=1}^{\infty}$ by data sequences $X^n = \{X_1, X_2, ..., X_n\}$ and $Y^n = \{Y_1, Y_2, ..., Y_n\}$, respectively.

The *2-terminal* coding system is shown in Fig. 4.2. Correlated sources, X and Y, are not co-located or cannot cooperate to directly exploit their correlation. Data sequences X^n and Y^n are separately encoded to $\varphi_0(X^n)$ and $\varphi_1(Y^n)$. The encoder functions are defined by

$$\begin{aligned} \varphi_0 &: \mathcal{X}^n \to \mathcal{C}_1 = \{1, 2, ..., C_1\}, \\ \varphi_1 &: \mathcal{Y}^n \to \mathcal{C}_2 = \{1, 2, ..., C_2\}. \end{aligned} \tag{4.2}$$

The coded sequences are sent to the FC or the hub node with the rate constraints

$$\frac{1}{n} \log C_i \leq R_i + \delta \ , \ i = 0, 1 \tag{4.3}$$

where δ is an arbitrary prescribed positive number. The decoder observes $(\varphi_0(X^n), \varphi_1(Y^n))$, decodes all messages, and makes estimates of all sources. The decoder function is given by $\psi = (\psi_0, \psi_1)$ where its components are defined by

$$\begin{aligned} \psi_0 &: \mathcal{C}_0 \times \mathcal{C}_1 \rightarrow \mathcal{X}^n, \\ \psi_1 &: \mathcal{C}_0 \times \mathcal{C}_1 \rightarrow \mathcal{Y}^n. \end{aligned} \tag{4.4}$$

Let $d_0 : \mathcal{X}^2 \rightarrow [0, \infty)$ and $d_1 : \mathcal{Y}^2 \rightarrow [0, \infty)$ be the squared distortion measures, i.e., $d_0(X, \widehat{X}) = (X - \widehat{X})^2$ and $d_1(Y, \widehat{Y}) = (Y - \widehat{Y})^2$. For the reconstructed signals $\widehat{X}^n = \psi_0(\varphi_0(X^n), \varphi_1(Y^n))$ and $\widehat{Y}^n = \psi_1(\varphi_0(X^n), \varphi_1(Y^n))$, the average distortions Δ_0, Δ_1 can be defined by

$$\begin{aligned} \Delta_0 &= E\left[\frac{1}{n}\sum_{t=1}^{n} d_1(X_t, \widehat{X}_t)\right], \\ \Delta_1 &= E\left[\frac{1}{n}\sum_{t=1}^{n} d_2(Y_t, \widehat{Y}_t)\right]. \end{aligned} \tag{4.5}$$

For given distortion levels D_0 and D_1, a rate pair (R_0, R_1) is admissible if for any $\delta > 0$ and any $n \geq n_0(\delta)$ there exists a triple $(\varphi_0, \varphi_1, \psi)$ satisfying (4.2)-(4.4) such that $\Delta_i \leq D_i + \delta$ for $i = 0, 1$. The rate-distortion region $\mathcal{R}(D_0, D_1)$ can be defined as

$$\mathcal{R}(D_0, D_1) = \left\{(R_0, R_1) : (R_0, R_1) \ is \ admissible\right\}. \tag{4.6}$$

The whole rate-distortion region of the *2-terminal* coding scheme is presented in Chapter 2, (2.11)-(2.14). Our goal is to show that this rate-distortion region can be achieved using the successive coding strategy. There are two steps to reach this goal: 1) Obtaining two curved parts of the $\mathcal{R}(D_0, D_1)$, i.e., (2.11) and (2.12); 2) Obtaining the straight-line segment of $\mathcal{R}(D_0, D_1)$, i.e., the sum-rate limit of (2.13).

4.3 *2-Terminal* Source Coding

4.3.1 *1-helper* Source Coding Scheme

The two-curved portions of the rate region for the *2-terminal* coding system are in fact the rate-distortion regions of the *1-helper* coding schemes, where one source provides partial side information to the decoder to help reconstruction of the other source signal.

Assume that distortion level D_1 is sufficiently large, i.e., the goal is only to reconstruct X, and the other source, Y, is used as a helper. In other words, the helper will not be reconstructed, and it is just used as auxiliary information to reconstruct the main source. This problem is called the *1-helper* problem. The associated coding system is depicted in Fig. 4.4. The rate region does not depend on the distortion level D_1, and it can be represented by $R_0(D_0)$, i.e., $R_0(D_0) = \Big\{(R_0, R_1) : (R_0, R_1) \in \mathcal{R}(D_0, D_1) \ for \ some \ D_1 > 0 \Big\}$. It is clear that $R_0(D_0)$ is an outer region of $\mathcal{R}(D_0, D_1)$.

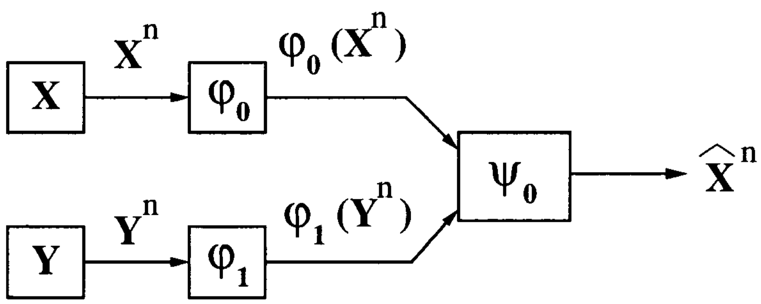

Figure 4.4: *1-helper* coding scheme.

We determine the rate-distortion performance of the successive coding strategy in the *1-helper* problem. By applying the successive coding/decoding strategy in the *1-helper* problem, the problem can be decomposed into *two* successive coding stages. The scenario is presented in Fig. 4.5. Each source encodes its message while previously decoded message that is available at the decoder acts as the decoder side information. At the FC, instead of joint decoding, messages from encoders are decoded sequentially in order to increase the fidelity of estimation at each decoding step.

Theorem 4 *Successively structured Wyner-Ziv codes can achieve the rate-distortion function of the 1-helper coding scheme:*

$$R_0(D_0) = \frac{1}{2}\log^+\left[\frac{\sigma_X^2}{D_0}\left(1-\rho^2+\rho^2 2^{-2R_1}\right)\right], \tag{4.7}$$

In other words, for every $D_0 > 0$, the achievable rate region of the successive coding strategy for the 1-helper coding scheme can be represented by

$$\mathcal{R}_0(D_0) = \quad \Big\{(R_0, R_1) : R_0 \geq \tfrac{1}{2}\log^+\left[\tfrac{\sigma_X^2}{D_0}\left(1-\rho^2+\rho^2 2^{-2R_1}\right)\right]\Big\}. \tag{4.8}$$

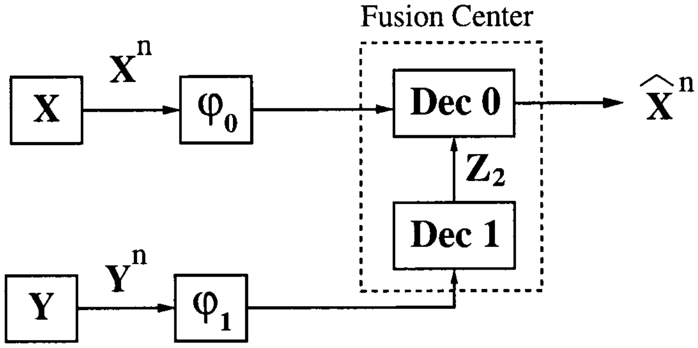

Figure 4.5: *1-helper* coding system with the successive coding strategy.

Proof Let Z_2 denote the output signal generated by the helper at the decoder. Then, X can be encoded at the Wyner-Ziv rate

$$R_0(D_0) = \frac{1}{2}\log\left(\frac{\sigma^2_{X|Z_2}}{D_0}\right). \tag{4.9}$$

The helper forms the MMSE estimate of X from Y, given by

$$X_2 = E[X \mid Y] = \rho\frac{\sigma_X}{\sigma_Y}Y. \tag{4.10}$$

Then, the helper encodes X_2 at rate R_1, thus the quantization error (distortion) can be expressed as

$$D_1 = var(X_2)2^{-2R_1} = \rho^2\sigma_X^2 2^{-2R_1}. \tag{4.11}$$

Finally, Z_2 can be written as $X_2 + E_2$, where E_2 is the quantization error with variance D_1 given above, and E_2 is independent of X_2 in the limit of large block length. The result is that the conditional variance of X given Z_2 is equal to the estimation MMSE between X and X_2 plus the quantization error variance D_1. We have

$$\sigma^2_{X|Z_2} = E\left[|X - X_2|^2\right] + D_1 = \sigma_X^2\left(1-\rho^2\right) + \sigma_X^2\rho^2 2^{-2R_1}. \tag{4.12}$$

Replacing this into (4.9), we obtain the result of (4.7). This is the achievable rate-distortion region by the successive coding strategy. Comparing our result with the results of [42] and [18] in (2.11) shows that by applying the successive coding strategy, the rate-distortion function for the *1-helper* coding scheme is achievable. This completes the proof.

Remark 1 The Wyner-Ziv result [9] can be considered as the special case for the *1-helper* source coding scheme when a high resolution encoding is used to encode the side information. In other words, by letting $R_1 \to \infty$, the rate-distortion function derived by Wyner and Ziv can be obtained.

Remark 2 By rewriting equation (4.7) as $D_0 = \sigma_X^2 2^{-2R_0}\left(1 - \rho^2 + \rho^2 2^{-2R_1}\right)$ it can be seen that each additional bit of the primary source description reduces the average final distortion by a factor of 4; however, each additional bit of the helper description can reduce the average distortion by at most a factor of 4. This maximum reduction occurs when the correlation coefficient is one. Also, under a constant rate of the helper, D_0 is a decreasing function of the correlation coefficient.

Remark 3 We can also obtain the result of Theorem 4 when decoder 1 decodes the associated source signal Y instead of X. Then, this decoded signal is used as the side information for decoder 0. This could be shown as follows. The encoder of Y encodes and transmits its signal with the rate

$$R_1(D_1) = \frac{1}{2}\log\left(\frac{\sigma_Y^2}{D_1}\right), \qquad 0 \le D_1 \le \sigma_Y^2 \tag{4.13}$$

where D_1 is the average distortion in estimating Y at decoder 1. The encoder of X can encode the source X by considering its statistical knowledge about the decoder's data, i.e., $Z_2 = \widetilde{Y}$, with the Wyner-Ziv rate

$$R_0(D_0) = \frac{1}{2}\log\left(\frac{\sigma^2_{X|\widetilde{Y}}}{D_0}\right) = \frac{1}{2}\log\left(\frac{\sigma_X^2\left(1-\rho^2_{X\widetilde{Y}}\right)}{D_0}\right) \tag{4.14}$$

where $Z_2 = \widetilde{Y} = Y + V$ with $V \sim \mathcal{N}(0, D_1\sigma_Y^2/\left(\sigma_Y^2 - D_1\right))$. In fact, we use an innovation form [5] to rewrite the relationship between the estimate of the source Y at decoder 1, i.e., $\widehat{Y}$, and the source Y as $\widehat{Y} = (Y + V)\alpha$ where $V \sim \mathcal{N}(0, D_1\sigma_Y^2/\left(\sigma_Y^2 - D_1\right))$ and $\alpha = (1 - D_1/\sigma_Y^2)$. This relationship gives us the distortion of $E\left[(Y - \widehat{Y})^2\right] = D_1$. The message Z_2 can be considered as $\widehat{Y}/\alpha = Y + V$, i.e., Z_2 can be considered as the source

Y in an AWGN, V. Therefore, $\rho_{X\tilde{Y}} = \rho\sqrt{\sigma_Y^2 - D_1}/\sigma_Y$. By doing some manipulations, we will see that the rate-distortion tradeoff of (4.14) will be the same as (4.7).

4.3.1.1 Geometric Sphere-Packing argument

In this subsection, we show how to obtain the rate-distortion function of the *1-helper* problem from geometric sphere-packing arguments [26, 25]. These arguments for the noisy Wyner-Ziv source coding are presented in [8].

At the first step, we obtain the encoding rate of the side information Y. Since our final goal is to reconstruct X at the decoder, we must consider how much the encoded message W_1, once decoded, will help to reconstruct X. It can be shown that the estimating function at decoder 1 is equal to $\widehat{X} = \frac{\sigma_X^2}{\sigma_X^2 - \sigma_{X|Y}^2} W_1$, i.e., this function can achieve the distortion level of D_{X1}. The decoder 1 only knows the transmitted vector lies in an uncertainty sphere of radius $r_{u_1} = \sqrt{n\left(\sigma_X^2 + \varepsilon_{11}\right)}$ which is centered at the source estimate $E\left[X\right]$. The scenario is depicted in Fig. 4.6. The encoder of the auxiliary source maps Y to the label of the quantization region in which it lies. The radius of each spherical quantization region is equal to $r_{q_1} = \sqrt{n\frac{\sigma_X^2}{\sigma_X^2 - \sigma_{X|Y}^2}\left(D_{X1} - \sigma_{X|Y}^2 - \varepsilon_{22}\right)}$. If no two quantization regions share the same label in each large sphere of radius r_{u_1}, the quantization region in which Y is located can be determined without error. We determine the minimum number of quantization spheres required to cover the uncertainty sphere. This number is lower bounded by the ratio of volumes of spheres: $M_1 \geq \frac{k(n)\left(r_{u_1}\right)^n}{k(n)\left(r_{q_1}\right)^n}$ where $k(n)$ is a coefficient that is a function of the dimension. Thus, the lower bound on the rate-distortion function $R_1(D_{X1})$ can be obtained as

$$R_1 = \frac{1}{n}\log_2 M_1 \geq \frac{1}{2}\log\left(\frac{\sigma_X^2 + \varepsilon_{11}}{\frac{\sigma_X^2}{\sigma_X^2 - \sigma_{X|Y}^2}\left(D_{X1} - \sigma_{X|Y}^2 - \varepsilon_{22}\right)}\right) > \frac{1}{2}\log\left(\frac{\sigma_X^2 - \sigma_{X|Y}^2}{D_{X1} - \sigma_{X|Y}^2}\right) \tag{4.15}$$

In the next step, based solely on the available side information at the decoder 0, Z_2, the decoder can have a MMSE equal to $\sigma_{X|Z_2}^2 = D_{X1}$. Therefore, the transmitted vector

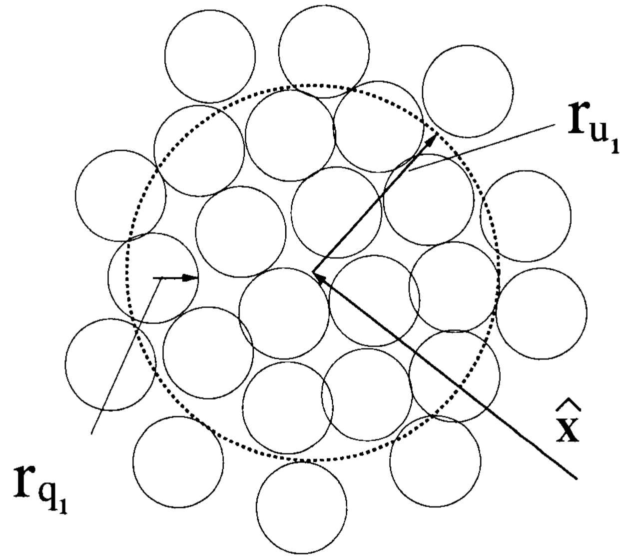

Figure 4.6: Sphere-packing argument to determine the rate-distortion function for the *1-helper* problem. The source estimate lies within the dotted circle with the radius of r_{u_1}. The solid small circles with the radius r_{q_1} correspond to quantization regions at the encoder of the auxiliary source.

X falls in an uncertainty sphere of radius $r_u = \sqrt{n\left(\sigma^2_{X|Z_2} + \varepsilon_1\right)} = \sqrt{n\left(D_{X1} + \varepsilon_1\right)}$ which is centered at the source estimate $E\left[X \mid Z_2\right]$. This sphere is shown by the dotted circle in Fig. 4.7. The encoder 0 maps X to the label of the quantization region in which it lies, where the radius of each quantization sphere equals to $r_q = \sqrt{n\left(D_0 - \varepsilon_2\right)}$. Thus, the minimum number of quantization spheres of radius r_q required to cover the larger sphere of uncertainty is lower bounded by

$$M \geq \frac{k(n)\left(\sqrt{n\left(D_{X1} + \varepsilon_1\right)}\right)^n}{k(n)\left(\sqrt{n\left(D_0 - \varepsilon_2\right)}\right)^n}$$

As a result, the lower bound on the rate-distortion function of the *1-helper* problem can be derived as

$$R_0 = \frac{1}{n}\log_2 M \geq \frac{1}{2}\log\left(\frac{D_{X1} + \varepsilon_1}{D_0 - \varepsilon_2}\right) > \frac{1}{2}\log\left(\frac{D_{X1}}{D_0}\right). \tag{4.16}$$

By substituting D_{X1} from (4.15) into (4.16), the same result as (4.7) will be obtained.

To apply the successive coding strategy in the *2-terminal* source coding scheme,

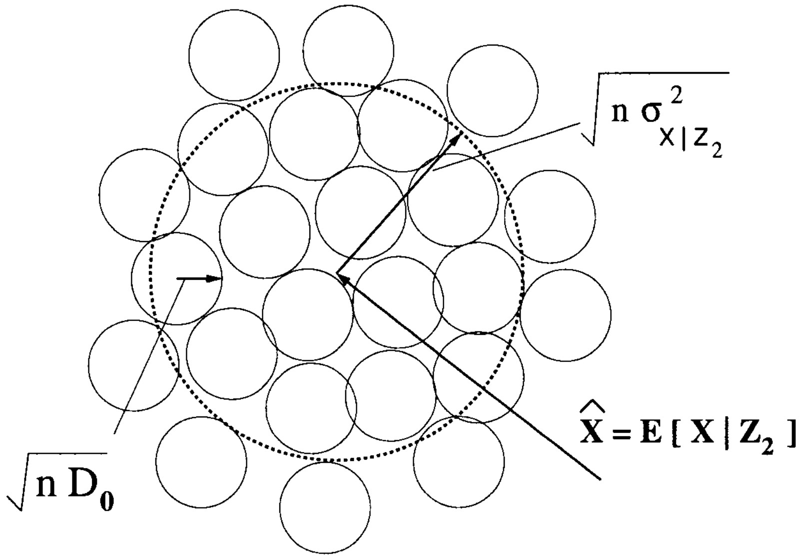

Figure 4.7: Sphere covering argument to determine the rate-distortion function for the *1-helper* problem. The source estimate given the side information lies within the dotted circle. The solid small circles with the radius $\sqrt{nD_0}$ correspond to quantization regions at the encoder of the primary source.

there are two possible orderings for the coding/decoding process: (a) the message from source Y is encoded and transmitted first, and (b) the message from source X is encoded and transmitted first. Ordering (a) determines the rate-distortion function for the *1-helper* problem when Y is the helper and therefore the distortion for reconstructing X is minimized. Ordering (b) determines the rate-distortion function for the *1-helper* problem when X is the helper and therefore the distortion for reconstructing Y is minimized. These two rate-distortion functions determine two curved portions of the rate region $\mathcal{R}(D_0, D_1)$ for *2-terminal* source coding problem.

4.3.2 Sum-rate limit for the *2-terminal* rate-region

Now we want to obtain the last part of the $\mathcal{R}(D_0, D_1)$ which is defined by the boundary of the sum-rate. Using (4.13) and (4.14), ordering (a) gives the minimum sum-rate of

$$R_0 + R_1 = \frac{1}{2} \log \left[\frac{\sigma_X^2 \sigma_Y^2}{D_0 D_1} \left(1 - \rho^2\right) + \frac{\sigma_X^2 \rho^2}{D_0} \right]. \tag{4.17}$$

By symmetry, ordering (b) gives the minimum sum-rate of

$$R_0 + R_1 = \frac{1}{2}\log\left[\frac{\sigma_X^2\sigma_Y^2}{D_0D_1}\left(1-\rho^2\right) + \frac{\sigma_Y^2\rho^2}{D_1}\right]. \tag{4.18}$$

By allowing time sharing between ordering (a) and (b), the achievable sum-rate can be expressed as

$$R_0 + R_1 = \frac{1}{2}\log\left[\frac{\sigma_X^2\sigma_Y^2}{D_0D_1}\left(1-\rho^2\right) + \frac{\sigma_X^2\rho^2}{D_0}\right] + \frac{1}{2}\log\left[\frac{\sigma_X^2\left(1-\rho^2\right)+D_0\rho^2\,\sigma_Y^2}{\sigma_Y^2\left(1-\rho^2\right)+D_1\rho^2\,\sigma_X^2}\right]^{\alpha}, \tag{4.19}$$

where $0 \leq \alpha \leq 1$. Since the sum-rate of (4.19) is achievable, it is an upper bound for the sum-rate part in the rate region of *2-terminal* source coding problem. By subtracting (4.19) from (2.13) and doing some manipulations, we can show that the sum-rate distortion function of the *2-terminal* source coding is achievable if

$$\rho^2 = \frac{max\left(D_0/\sigma_X^2, D_1/\sigma_Y^2\right)-1}{min\left(D_0/\sigma_X^2, D_1/\sigma_Y^2\right)-1}. \tag{4.20}$$

In this case, the successive coding strategy degrades to the no-helper problem where only the source with the minimum average distortion D_i $(i = 0, 1)$ should be encoded.

In our analysis, so far, we have assumed that each decoder desires to reconstruct the corresponding source within the given fidelity, D_i. But we can develop connections between data fusion and source coding with side information. In this scenario, the encoder quantizes and encodes the source to a degree of fidelity less than D_i to reduce the transmission rate. But this reduction does not affect the quality of the reconstructed signals at the FC. In fact, the remaining correlation among the decoded signals enables the FC to reconstruct the sources into desired degrees of fidelity. By using a linear estimator after the source decoder, both source signals can be reproduced at the desired degrees of fidelity, D_i's. The joint design of the source encoding and the data fusion will yield substantial performance gain over decoupled designs. Based on this scenario, we demonstrate that the successive coding strategy can achieve all the rate-distortion tuples (R_0, R_1, D_0, D_1) belonging to the inner bound of Berger and Tung [16, 50]. In fact, the

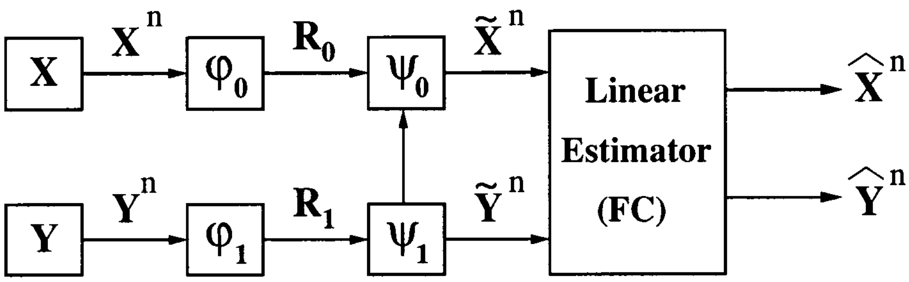

Figure 4.8: Block diagram of the successive coding strategy in the *2-terminal* source coding problem. The linear estimator fuses both received signals softly to produce the estimates of both sources.

whole rate-distortion region of the *2-terminal* source coding problem can be characterized by applying the successive coding strategy.

Theorem 5 *Successive coding strategy can achieve the sum-rate distortion function of the 2-terminal source coding scheme.*

Proof Our proof is similar to the proof of source-splitting method presented in [81, 82]. Consider the ordering (a) for the *2-terminal* source coding, depicted in Fig. 4.8. The source encoder φ_1, quantizes Y^n and then compresses the quantized signal. The output message W_1 is transmitted at rate R_1. At the decoder side, W_1 is decompressed and it is used to reconstruct Y^n as $\widetilde{Y}^n$. By exploring the remained correlation between $\widetilde{Y}^n$ and X^n, the encoder φ_0 compresses the quantized version of X^n at rate R_0. Using $\widetilde{Y}^n$ as the side information, the decoder ψ_0 decodes the received signal to $\widetilde{X}^n$. Given $\widetilde{X}^n$ and $\widetilde{Y}^n$, the linear estimator reproduces X^n and Y^n using linear combination of the inputs:

$$\begin{aligned} \widehat{X}^n &= \alpha_0 \widetilde{Y}^n + \beta_0 \widetilde{X}^n, \\ \widehat{Y}^n &= \alpha_1 \widetilde{Y}^n + \beta_1 \widetilde{X}^n. \end{aligned} \tag{4.21}$$

Define the quantization errors as $E_0^n = \widetilde{X}^n - X^n$ and $E_1^n = \widetilde{Y}^n - Y^n$. We first derive the estimator coefficients. The average quantization distortions Δ_0, Δ_1 can be defined by

$$\begin{aligned} \Delta_0 = E\left[\tfrac{1}{n}\textstyle\sum_{t=1}^n (X_t - \widetilde{X}_t)^2\right] = E\left[\tfrac{1}{n}\textstyle\sum_{t=1}^n E_{0,t}^2\right], \\ \Delta_1 = E\left[\tfrac{1}{n}\textstyle\sum_{t=1}^n (Y_t - \widetilde{Y}_t)^2\right] = E\left[\tfrac{1}{n}\textstyle\sum_{t=1}^n E_{1,t}^2\right]. \end{aligned} \tag{4.22}$$

Due to the orthogonal properties for optimal estimations, applying the projection theorem results in:

$$E\left\{(X-\widehat{X})\widetilde{X}\right\}=0 \quad \& \quad E\left\{(X-\widehat{X})\widetilde{Y}\right\}=0,$$
$$E\left\{(Y-\widehat{Y})\widetilde{X}\right\}=0 \quad \& \quad E\left\{(Y-\widehat{Y})\widetilde{Y}\right\}=0. \tag{4.23}$$

As a result, by doing some manipulation, the estimator coefficients that minimize the average distortions $E\left\{(X-\widehat{X})^2\right\}$ and $E\left\{(Y-\widehat{Y})^2\right\}$ can be obtained as

$$\begin{array}{ll} \alpha_0 = \frac{\rho\sigma_X\sigma_Y\Delta_0}{\Delta^*} & \beta_0 = \frac{(\sigma_Y^2+\Delta_1)\sigma_X^2-\rho^2\sigma_X^2\sigma_Y^2}{\Delta^*} \\ \alpha_1 = \frac{(\sigma_X^2+\Delta_0)\sigma_Y^2-\rho^2\sigma_X^2\sigma_Y^2}{\Delta^*} & \beta_1 = \frac{\rho\sigma_X\sigma_Y\Delta_1}{\Delta^*}, \end{array} \tag{4.24}$$

where $\Delta^* = (\sigma_X^2+\Delta_0)(\sigma_Y^2+\Delta_1)-\rho^2\sigma_X^2\sigma_Y^2$. Based on the rate-distortion theory and Wyner-Ziv results, the transmission rates of this scheme can be expressed as

$$nR_0 \geq I(X^n;\widetilde{X}^n) - I(\widetilde{X}^n;\widetilde{Y}^n), \tag{4.25}$$

$$nR_1 \geq I(Y^n;\widetilde{Y}^n). \tag{4.26}$$

For jointly Gaussian random variables X and Y, $I(X;Y) = -\frac{1}{2}\log(1-\rho^2)$. We just need to obtain the correlation coefficients of $\rho_{X\widetilde{X}}$, $\rho_{\widetilde{X}\widetilde{Y}}$, and $\rho_{Y\widetilde{Y}}$. We can show

$$\rho_{X\widetilde{X}}^2 = \frac{\sigma_X^2}{\sigma_X^2+\Delta_0}, \quad \rho_{Y\widetilde{Y}}^2 = \frac{\sigma_Y^2}{\sigma_Y^2+\Delta_1}, \quad \rho_{\widetilde{X}\widetilde{Y}}^2 = \frac{\rho^2\sigma_X^2\sigma_Y^2}{(\sigma_X^2+\Delta_0)(\sigma_Y^2+\Delta_1)}. \tag{4.27}$$

Thus, (4.25) and (4.26) can be computed as

$$R_0 \geq \frac{1}{2}\log\left(\frac{(\sigma_X^2+\Delta_0)(\sigma_Y^2+\Delta_1)-\rho^2\sigma_X^2\sigma_Y^2}{\Delta_0(\sigma_Y^2+\Delta_1)}\right), \tag{4.28}$$

$$R_1 \geq \frac{1}{2}\log\left(\frac{\sigma_Y^2+\Delta_1}{\Delta_1}\right). \tag{4.29}$$

The overall average distortion can be expressed as

$$\begin{aligned} D_0^* &= E\left\{(X-\widehat{X})^2\right\} = E\left\{(X-\alpha_0\widetilde{Y}-\beta_0\widetilde{X})^2\right\} \\ &= \sigma_X^2+\alpha_0^2(\sigma_Y^2+\Delta_1)+\beta_0^2(\sigma_X^2+\Delta_0)-2\alpha_0\rho\sigma_X\sigma_Y-2\beta_0\sigma_X^2+2\alpha_0\beta_0\rho\sigma_X\sigma_Y, \end{aligned} \tag{4.30}$$

$$\begin{aligned} D_1^* &= E\left\{(Y-\widehat{Y})^2\right\} = E\left\{(Y-\alpha_1\widetilde{Y}-\beta_1\widetilde{X})^2\right\} \\ &= \sigma_Y^2+\alpha_1^2(\sigma_Y^2+\Delta_1)+\beta_1^2(\sigma_X^2+\Delta_0)-2\alpha_1\sigma_Y^2-2\beta_1\rho\sigma_X\sigma_Y+2\alpha_1\beta_1\rho\sigma_X\sigma_Y. \end{aligned} \tag{4.31}$$

Using (4.24), average distortions (4.30) and (4.31) can be simplified as

$$D_0^* = \frac{\left(\sigma_Y^2\left(1-\rho^2\right)+\Delta_1\right)\sigma_X^2\Delta_0}{\Delta^*}, \tag{4.32}$$

$$D_1^* = \frac{\left(\sigma_X^2\left(1-\rho^2\right)+\Delta_0\right)\sigma_Y^2\Delta_1}{\Delta^*}. \tag{4.33}$$

We show that one of the corner points on the sum-rate bound of (2.13) is achievable with the Wyner-Ziv scheme of ordering (a). We prove this result by construction. Let

$$\Delta_0 = \frac{2D_0\sigma_X^2\sigma_Y^2\left(1-\rho^2\right)}{\beta_{max}\sigma_X^2\sigma_Y^2\left(1-\rho^2\right)-2D_0\sigma_Y^2}, \tag{4.34}$$

$$\Delta_1 = \frac{2D_1\sigma_X^2\sigma_Y^2\left(1-\rho^2\right)}{\beta_{max}\sigma_X^2\sigma_Y^2\left(1-\rho^2\right)-2D_1\sigma_X^2}. \tag{4.35}$$

Substituting (4.34) and (4.35) in (4.28) and (4.29) reveals that

$$R_{0A} = \frac{\beta_{max}\sigma_X^2\sigma_Y^2\left(1-\rho^2\right)^2}{D_0\left(\beta_{max}\sigma_Y^2\left(1-\rho^2\right)-2D_1\rho^2\right)}, \tag{4.36}$$

$$R_{1A} = \frac{\beta_{max}\sigma_Y^2\left(1-\rho^2\right)-2D_1\rho^2}{2D_1\left(1-\rho^2\right)}. \tag{4.37}$$

R_{0A} and R_{1A} satisfy the sum-rate distortion function of (2.13). The achievable average distortions (4.32) and (4.33) can be expressed as $D_0^* = D_0$ and $D_1^* = D_1$. Therefore, the sum-rate distortion function of (2.13) is achievable. Ordering (b) gives similar results which can be obtained by interchanging subscripts. Therefore, the optimal rate allocation scheme that achieve any distortion pair (D_0, D_1) using minimum sum-rate $\overline{R}$ of (2.13) corresponds to the points on line AB where

$$R_{0B} = \frac{1}{2}\log\left(\frac{\sigma_X^2+\Delta_0}{\Delta_0}\right) = \frac{\beta_{max}\sigma_X^2\left(1-\rho^2\right)-2D_0\rho^2}{2D_0\left(1-\rho^2\right)}, \tag{4.38}$$

$$R_{1B} = \frac{1}{2}\log\left(\frac{\left(\sigma_X^2+\Delta_0\right)\left(\sigma_Y^2+\Delta_1\right)-\rho^2\sigma_X^2\sigma_Y^2}{\Delta_1\left(\sigma_X^2+\Delta_0\right)}\right) = \frac{\beta_{max}\sigma_X^2\sigma_Y^2\left(1-\rho^2\right)^2}{D_1\left(\beta_{max}\sigma_X^2\left(1-\rho^2\right)-2D_0\rho^2\right)}. \tag{4.39}$$

Allowing time-sharing between ordering (a) and (b) achieves all the intermediate points of line AB.

4.3.3 Comparison with Sequential coding of correlated sources

In this part, we make a comparison between the minimum sum-rate of the sequential coding [80] and the sum-rate of the successive coding strategy for coding of correlated Gaussian sources. In fact, the scenario of successive coding in the *2-terminal* source coding problem is similar to the scenario of sequential coding of 2 correlated sources. The sequential coding scheme is illustrated in Fig. 4.3. In this coding system, the encoding is performed sequentially, i.e., the encoding of Y depends on the source X but not vice versa. In [80], it is shown that the minimal sum-rate for the sequential coding of correlated Gaussian sources is given by

$$R_0 + R_1 = \frac{1}{2}\log\left(\frac{\sigma_X^2\sigma_Y^2\left(1-\rho^2\left(1-D_0/\sigma_X^2\right)\right)}{D_0 D_1}\right). \tag{4.40}$$

Now, consider the ordering (b) of the successive coding strategy. The minimum sum-rate of the successive coding strategy for this model is given in (4.18). By rewriting (4.18), we observe that this is the same as the minimum sum-rate of sequential coding in (4.40). Therefore, there is no sum-rate loss with the successive coding compared with the sequential coding of correlated Gaussian sources. This means that the availability of side information at the encoder does not improve the sum-rate distortion function. This observation is of special interest in some applications. For instance, in most practical Video Codecs there are storage constraints. In particular, the encoder for encoding a frame can retain a copy of the previous frame but not earlier frames [80]. This suggests that we should consider the successive coding as a promising technique to achieve the minimum sum-rate with no storage constraint at the encoder.

4.4 *M-Terminal* Source Coding

In the multiterminal source coding, the FC decodes messages of all sources. The multiterminal coding system with *m* terminals is shown in Fig. 4.9.

Data sequences Y_i^n are separately encoded to $\varphi_i\left(Y_i^n\right)$ where the encoder functions

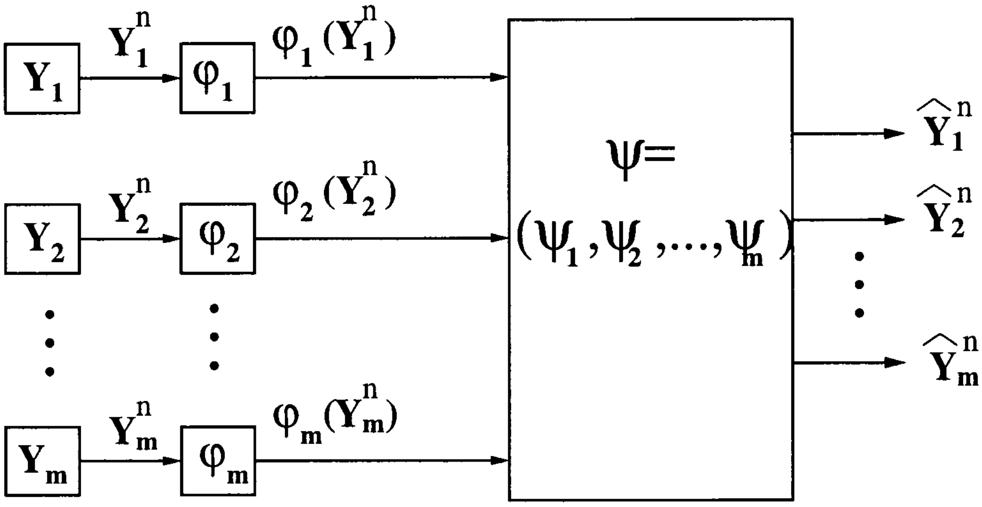

Figure 4.9: Multiterminal source coding scheme. Y_i's for $i = 1, ..., m$ are m correlated sources that are separately encoded. The joint decoder wants to obtain estimates of all the sources.

are defined by

$$\varphi_i : \mathcal{Y}_i^n \to \mathcal{C}_i = \{1, 2, ..., C_i\}. \tag{4.41}$$

The coded sequences are sent to the FC with the rate constraints

$$\frac{1}{n} \log C_i \leq R_i + \delta \ , \ i = 1, 2, ..., m \tag{4.42}$$

where δ is an arbitrary prescribed positive number. The decoder observes m transmitted sequences $(\varphi_1 (Y_1^n), ..., \varphi_m (Y_m^n))$, decodes all the messages, and makes estimates of all the sources. The decoder function is given by $\psi = (\psi_1, \psi_2, ..., \psi_m)$ where its components are defined by

$$\psi_i : \mathcal{C}_1 \times \mathcal{C}_2 \times \cdots \times \mathcal{C}_m \to \mathcal{Y}_i^n. \tag{4.43}$$

Let $d_i : \mathcal{Y}_i^2 \to [0, \infty)$ be the squared distortion measures, i.e., $d_i(Y_i, \widehat{Y}_i) = (Y_i - \widehat{Y}_i)^2$. For the reconstructed signals $\widehat{Y}_i^n = \psi_i (\varphi_1 (Y_1^n), ..., \varphi_m (Y_m^n))$, the average distortions $\Delta_1, \Delta_2, ..., \Delta_m$ can be defined by

$$\Delta_i = E\left[\frac{1}{n}\sum_{t=1}^{n} d_i(Y_{it}, \widehat{Y}_{it})\right]. \tag{4.44}$$

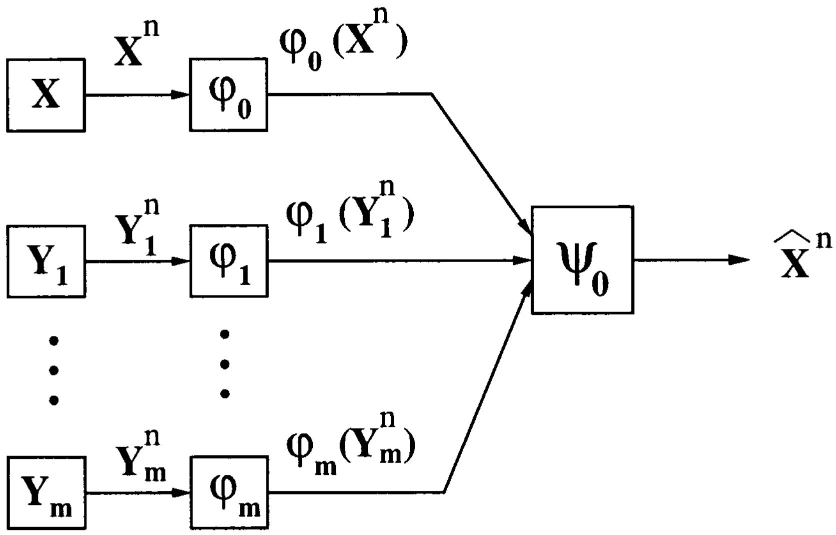

Figure 4.10: *m-helper* source coding scheme.

For given distortion levels $(D_1, D_2, ..., D_m)$, an m-tuple set of rates $(R_1, R_2, ..., R_m)$ is admissible if for any $\delta > 0$ and any $n \geq n_0(\delta)$ there exists a $(m+1)-$tuple $(\varphi_1, \varphi_2, ..., \varphi_m, \psi)$ satisfying (4.41)-(4.43) such that $\Delta_i \leq D_i + \delta$ for $i = 1, 2, ..., m$. The rate-distortion region $\mathcal{R}(D_1, D_2, ..., D_m)$ can be defined as all the m-tuple sets of rates $(R_1, R_2, ..., R_m)$ that are admissible.

Similar to the previous section, we first consider the special case of *m-terminal* coding system, where the goal is to estimate one of these sources while other sources provide partial side information to the decoder to help reconstruction of the primary source. Based on the result of the following section, we can obtain an inner region for the *m-terminal* source coding scheme.

4.4.1 Achievable Rate-Region of a special case: *m-helper* problem

The m*-helper* system and the system based on the successive coding strategy are shown in Figures 4.10 and 4.11, respectively.

By generalizing the results of Section 4.3 and considering the fact that each decoder decodes its corresponding source, the rates of encoders can be computed as follows:

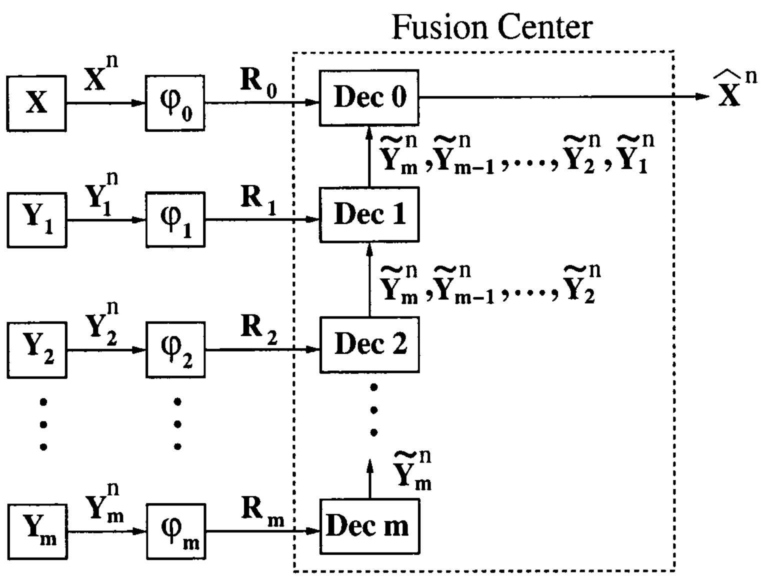

Figure 4.11: Successively structured *m-helper* source coding scheme.

$$\left\{\begin{array}{lcc} R_m & = & \frac{1}{2}\log\frac{\sigma^2_{Y_m}}{D_m} \\ R_{m-1} & = & \frac{1}{2}\log\frac{\sigma^2_{Y_{m-1}|Z_m}}{D_{m-1}} \\ R_{m-2} & = & \frac{1}{2}\log\frac{\sigma^2_{Y_{m-2}|Z_m,Z_{m-1}}}{D_{m-2}} \\ \vdots & & \vdots \\ R_1 & = & \frac{1}{2}\log\frac{\sigma^2_{Y_1|Z_m,Z_{m-1},\ldots,Z_2}}{D_1} \\ R_0 & = & \frac{1}{2}\log\frac{\sigma^2_{X|Z_m,Z_{m-1},\ldots,Z_2,Z_1}}{D_0} \end{array}\right. \tag{4.45}$$

where $Z_i = \widetilde{Y}_i = Y_i + \widetilde{V}_i$, $\widetilde{V}_i \sim \mathcal{N}(0, D_i/\alpha_i)$, $\alpha_i = \left(1 - D_i/\sigma^2_{Y_i}\right)$ and D_i $(i = 1, ..., m)$ is the average distortion in estimating source Y_i at decoder i. To obtain the final achievable rate-distortion region $R_0(D_0)$, we should obtain the conditional variances in (4.45) in terms of the source variances and the correlation coefficients between sources.

If X_i is a zero mean Gaussian random variable, $X_i \sim \mathcal{N}(0, \sigma^2_X)$, and X_i's are jointly Gaussian, then the conditional probability density function of $f(X_{n+1} \mid X_1, X_2, ..., X_n)$ can be obtained as follows:

$$f(X_{n+1} \mid X_1, X_2, ..., X_n) = \frac{1}{\sqrt{2\pi P}} exp\left[-\frac{(X_{n+1} - a_1X_1 - a_2X_2 - \cdots - a_nX_n)^2}{2P}\right], \tag{4.46}$$

where

$$P = \sigma^2_{X_{n+1}|X_1,X_2,\dots,X_n} = R_{n+1,n+1} - a_1 R_{1,n+1} - a_2 R_{2,n+1} - \cdots - a_n R_{n,n+1} \tag{4.47}$$

and $R_{i,j} = E[X_i X_j] = \rho_{X_i X_j}\sigma_{X_i}\sigma_{X_j}$. Coefficients $a_1, a_2, \dots, a_n$ can be computed from the following equation:

$$E[(X_{n+1} - a_1 X_1 - a_2 X_2 - \cdots - a_n X_n) X_i] = 0, \tag{4.48}$$

for $i = 1, 2, \dots, n$. This equation can be represented as the series of equations as follows:

$$\begin{cases} a_1 R_{11} + a_2 R_{21} + \cdots + a_n R_{n,1} &= R_{n+1,1} \\ a_1 R_{12} + a_2 R_{22} + \cdots + a_n R_{n,2} &= R_{n+1,2} \\ \vdots & \vdots \\ a_1 R_{1,n} + a_2 R_{2,n} + \cdots + a_n R_{n,n} &= R_{n+1,n} \end{cases} \tag{4.49}$$

To solve the system, consider it as the matrix equation of $\mathbf{A}\mathbf{R}_{XX} = \mathbf{R}_{X_{n+1}X}$ where $\mathbf{A} = [\ a_1 \quad a_2 \quad \dots \quad a_n\ ]$, $\mathbf{X} = [\ X_1 \quad X_2 \quad \dots \quad X_n\ ]$, $\mathbf{R}_{XX} = E[\mathbf{X}^t\mathbf{X}]$, and $\mathbf{R}_{X_{n+1}X} = [\ R_{n+1,1} \quad R_{n+1,2} \quad \dots \quad R_{n+1,n}\ ]$. Therefore, the coefficient vector of $\mathbf{A}$ can be obtained by

$$\mathbf{A} = \mathbf{R}_{X_{n+1}X}\mathbf{R}^{-1}_{XX}. \tag{4.50}$$

As a result, the conditional variances required in (4.45) can be obtained versus the correlation functions. Since the final goal is to obtain $R_0(D_0)$ in terms of source variances and correlation coefficients between sources, we just need to derive the relation between correlation coefficients $(\rho_{Y_i Z_j}, \rho_{Z_i Z_j}, \rho_{X Z_i})$ and correlation coefficients ρ_{XY_i} and $\rho_{Y_i Y_j}$. These coefficients can be obtained from the following equations:

$$\begin{aligned} \rho_{Y_i Z_j} &= \rho_{Y_i Y_j}\sqrt{\sigma^2_{Y_j} - D_j}/\sigma_{Y_j}, \\ \rho_{X Z_i} &= \rho_{X Y_i}\sqrt{\sigma^2_{Y_i} - D_i}/\sigma_{Y_i}, \\ \rho_{Z_i Z_j} &= \rho_{Y_i Y_j}\sqrt{\sigma^2_{Y_i} - D_i}\sqrt{\sigma^2_{Y_j} - D_j}/\sigma_{Y_i}\sigma_{Y_j}. \end{aligned} \tag{4.51}$$

4.4.2 Special Case: *2-helper* Source Coding Scheme

From (4.45), we know that

$$\begin{aligned} R_2(D_2) &= \tfrac{1}{2}\log\left(\frac{\sigma^2_{Y_2}}{D_2}\right), \\ R_1(D_1) &= \tfrac{1}{2}\log\left(\frac{\sigma^2_{Y_1|Z_2}}{D_1}\right), \\ R_0(D_0) &= \tfrac{1}{2}\log\left(\frac{\sigma^2_{X|Z_1,Z_2}}{D_0}\right), \end{aligned} \tag{4.52}$$

where

$$\begin{aligned} \sigma^2_{Y_1|Z_2} &= \sigma^2_{Y_1}\left(1-\rho^2_{Y_1Z_2}\right), \\ \sigma^2_{X|Z_1,Z_2} &= \sigma^2_X\left(1-\rho^2_1\right)/\left(1-\rho^2_{Z_1Z_2}\right), \\ \rho^2_1 &= \rho^2_{XZ_1}+\rho^2_{XZ_2}+\rho^2_{Z_1Z_2}-2\rho_{XZ_1}\rho_{XZ_2}\rho_{Z_1Z_2}. \end{aligned} \tag{4.53}$$

Combining these equations with (4.51) results in

$$R_0(D_0) = \frac{1}{2}\log\left[\frac{\sigma^2_X\left(1-\rho^2_1\right)}{D_0\left(1-\rho^2_{Z_1Z_2}\right)}\right], \tag{4.54}$$

where ρ_1^2 and $\rho^2_{Z_1Z_2}$ in terms of correlation coefficients among X, Y_1 and Y_2 are as follows:

$$\begin{aligned} \rho_1^2 = &\quad \rho^2_{XY_1}+\rho^2_{XY_2}+\rho^2_{Y_1Y_2}-2\rho_{XY_1}\rho_{XY_2}\rho_{Y_1Y_2} \\ &+ \left(2\rho_{XY_1}\rho_{XY_2}\rho_{Y_1Y_2}-\rho^2_{XY_1}-\rho^2_{Y_1Y_2}\right)\times\left(1-\rho^2_{Y_1Y_2}+\rho^2_{Y_1Y_2}2^{-2R_2}\right)2^{-2R_1} \\ &+ \left(2\rho_{XY_1}\rho_{XY_2}\rho_{Y_1Y_2}-\rho^2_{XY_2}-\rho^2_{Y_1Y_2}\right)2^{-2R_2} \\ &+ \left(\rho^2_{Y_1Y_2}-2\rho_{XY_1}\rho_{XY_2}\rho_{Y_1Y_2}\right)\times\left(1-\rho^2_{Y_1Y_2}+\rho^2_{Y_1Y_2}2^{-2R_2}\right)2^{-2R_1}2^{-2R_2}, \end{aligned} \tag{4.55}$$

and

$$\rho^2_{Z_1Z_2} = \rho^2_{Y_1Y_2}\left(1-2^{-2R_2}\right)\times\left(1-2^{-2R_1}\left(1-\rho^2_{Y_1Y_2}+\rho^2_{Y_1Y_2}2^{-2R_2}\right)\right). \tag{4.56}$$

Therefore, (4.54) defines an achievable rate-distortion region for the *2-helper* problem based on the successive coding strategy.

Remark 4 If $R_1 \to \infty$ and $R_2 \to \infty$, i.e., full resolution side information, the rate-distortion tradeoff for the *2-helper* coding scheme, (4.54), can be simplified to

$$R_0(D_0) = \frac{1}{2}\log\left[\frac{\sigma_X^2\left(1-\rho_{XY_1}^2-\rho_{XY_2}^2-\rho_{Y_1Y_2}^2+2\rho_{XY_1}\rho_{XY_2}\rho_{Y_1Y_2}\right)}{D_0\left(1-\rho_{Y_1Y_2}^2\right)}\right], \tag{4.57}$$

which is equal to the conditional rate-distortion function, $R_{X|Y_1,Y_2}$, when both encoder and decoder have access to Y_1 and Y_2. Therefore, there is no rate loss with the successive coding when compared to lossy coding of X with the side information (Y_1 and Y_2) available at both the encoder and the decoder.

Example 1 Consider the *2-helper* problem with $R_2 = 0$. Since the rate of the encoder 2 is zero, there is no help from Y_2. Therefore, we expect to obtain the rate-distortion for the *1-helper* problem. By substituting $R_2 = 0$ in (4.54), we obtain

$$R_0(D_0) = \frac{1}{2}\log\left[\frac{\sigma_X^2}{D_0}\left(1-\rho_{XY_1}^2+\rho_{XY_1}^2 2^{-2R_1}\right)\right], \tag{4.58}$$

which is the rate-distortion function of the *1-helper* source coding scheme.

4.4.3 An Inner Region for *m-Terminal* Coding Scheme

Assume $\Pi_i = (\pi_{i1}, \pi_{i2}, ..., \pi_{im})$ is a permutation of the set $\mathcal{I}_m = \{1, 2, ..., m\}$. Using the successive coding strategy, there are $m!$ possible orderings for the coding/decoding process in the *m-terminal* source coding scheme. For a given permutation Π, the following rate region is achievable:

$$\mathcal{R}_\Pi\left(D_{\Pi(1)}, D_{\Pi(2)}, ..., D_{\Pi(m)}\right) = \mathcal{R}^*_{\Pi(m)}\left(D_{\Pi(m)}\right) \cap \widehat{\mathcal{R}}_\Pi\left(D_{\Pi(1)}, D_{\Pi(2)}, ..., D_{\Pi(m)}\right) \tag{4.59}$$

where

$$\mathcal{R}^*_{\Pi(j)}\left(D_{\Pi(j)}\right) = \left\{\left(R_{\Pi(1)}, R_{\Pi(2)}, ..., R_{\Pi(m)}\right) :\ R_{\Pi(j)} \geq \tfrac{1}{2}\log_2\left[\frac{\sigma^2_{Y_{\Pi(j)}|Z_{\Pi(j-1)},...,Z_{\Pi(1)}}}{D_{\Pi(j)}}\right]\right\}, \tag{4.60}$$

$$\begin{aligned}\widehat{\mathcal{R}}_\Pi\left(D_{\Pi(1)}, D_{\Pi(2)}, ..., D_{\Pi(m)}\right) = \Big\{&\left(R_{\Pi(1)}, R_{\Pi(2)}, ..., R_{\Pi(m)}\right) :\\ &\textstyle\sum_{j=1}^m R_{\Pi(j)} \geq \tfrac{1}{2}\log_2\left[\frac{\prod_{j=1}^m \sigma^2_{Y_{\Pi(j)}|Z_{\Pi(j-1)},...,Z_{\Pi(1)}}}{\prod_{j=1}^m D_{\Pi(j)}}\right]\Big\},\end{aligned} \tag{4.61}$$

and $Z_{\Pi(j)} = Y_{\Pi(j)} + \widetilde{V}_{\Pi(j)}$, $\widetilde{V}_{\Pi(j)} \sim \mathcal{N}(0, D_{\Pi(j)}/\alpha_{\Pi(j)})$, $\alpha_{\Pi(j)} = \left(1 - D_{\Pi(j)}/\sigma^2_{Y_{\Pi(j)}}\right)$. The inner region for the rate region of the multiterminal source coding can be represented as

$$\mathcal{R}_{in}\left(D_1, D_2, ..., D_m\right) = \left\{ \left\{ \cap_{i=1}^{m!} \mathcal{R}^*_{\Pi_i(m)}\left(D_{\Pi_i(m)}\right) \right\} \cap \widehat{\mathcal{R}}_{min}\left(D_1, D_2, ..., D_m\right) \right\}, \tag{4.62}$$

where

$$\widehat{\mathcal{R}}_{min}\left(D_1, D_2, ..., D_m\right) = \min_{1 \leq i \leq m} \widehat{\mathcal{R}}_{\Pi_i}\left(D_{\Pi_i(1)}, D_{\Pi_i(2)}, ..., D_{\Pi_i(m)}\right). \tag{4.63}$$

$\mathcal{R}^*_{\Pi_i(m)}\left(D_{\Pi_i(m)}\right)$ is in fact the rate-distortion tradeoff of the $(m-1)$-*helper* coding scheme for the ith ordering of coding/decoding (out of $m!$ possible orderings).

4.5 Chapter Summary

In this chapter, we have determined the rate-distortion function of the *1-helper* problem based on the successive coding strategy. Compared with the sequential coding for the Gaussian correlated sources, there is no loss of sum-rate for the successive coding where the side information is not available at the encoder. We have demonstrated that the whole rate distortion region for the *2-terminal* source coding problem is achievable using the successive coding strategy. Finally, we derived an achievable rate-distortion region for the *m-helper* problem and also provided an inner bound for the rate region of the *m-terminal* source coding scheme based on the successive coding strategy.

Chapter 5

Gaussian CEO Problem with Gaussian Multiple Access Channel

5.1 Introduction

We consider a distributed network, modeled by Gaussian CEO problem [29], [30], [43], where L noisy observations of a memoryless Gaussian source are transmitted through an additive white Gaussian multiple access channel (MAC) to a single fusion center. The encoders are distributed and cannot cooperate to exploit their correlation. Each encoder is subject to a transmission cost constraint. This constraint comes from the restrictions on the resources such as bandwidth and power that are available at each agent. The scenario is illustrated in Fig. 5.1 (See Section 5.3 for details). The FC wishes to reconstruct the main source with an average distortion D at the smallest cost in the communication link. The goal in this problem, which is also called *multiterminal source-channel communication problem* [83] is to characterize all cost-distortion pairs achievable by any coding strategy in an information-theoretic sense regardless of delay and complexity. Our interest lies in determining the power-distortion region, while the fidelity of estimation at the FC is measured by the MSE distortion.

For the considered Gaussian network, the optimal power-distortion tradeoff and its corresponding optimal coding strategy still remain unknown. However, it is well known [36], [37] that for a point-to-point transmission of a single Gaussian source through an AWGN channel, a simple uncoded transmission is optimal if the source bandwidth is equal to the channel bandwidth. In the recent work of Lapidoth *et al.* [39], the authors consider sending a memoryless Bi-variate Gaussian source over an interfering MAC. They have shown that in the symmetric case, where the source components are of the same variance and the transmitted signals are subject to the same average power constraint, uncoded transmission is optimal below a threshold SNR. In [38] Gastpar and Vetterli consider the Gaussian sensor network modeled by the CEO problem in a symmetric environment, where the sensors observations have the same noise level and the transmitting terminals are subject to the same average power constraint. They show that as the number of sensors tends to infinity, uncoded transmission achieves the smallest possible distortion. Their proof is based on analyzing the idealized system in which the sensors are ideally linked to the destination. They show that as the number of sensors $L \to \infty$, the distortion of the uncoded transmission and the distortion of the idealized system coincide. However, two things remain unknown: (i) What is the optimal power-distortion tradeoff in a Gaussian CEO problem with a finite number of agents? (ii) Does the uncoded transmission perform optimally, in the sense of achieving the optimal power-distortion tradeoff, in a Gaussian CEO problem with finite L?

Our main contribution in this chapter is to obtain a necessary condition for achievability of all transmission cost-distortion tuples $(P_1, P_2, ..., P_L, D)$. Our proof is based on analyzing the remote source coding scenario, where the agents observations are given to one common encoder, and using the data processing inequality. Analyzing the uncoded transmission scheme in considered Gaussian network provides a sufficient condition for achievability of $(P_1, P_2, ..., P_L, D)$. We show that, in the symmetric case, these necessary and sufficient conditions coincide and give the optimal power-distortion tradeoff. Our analysis also shows that in the symmetric case of Gaussian CEO problem with a Gaussian MAC, uncoded transmission performs optimally for any finite number of agents.

The remainder of this chapter is organized as follows. In Section 5.2, we present the system model and problem formulation. Section 5.3 provides necessary and sufficient conditions for the achievability of power-distortion tuples $(P_1, P_2, ..., P_L, D)$. The optimal power-distortion tradeoff for the symmetric case is also presented. Section 5.4 concludes the chapter.

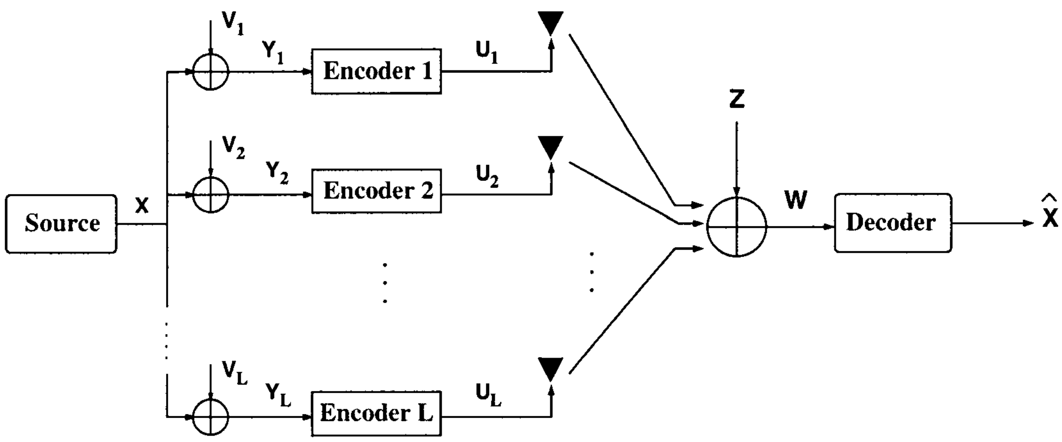

Figure 5.1: Gaussian CEO problem with a Gaussian multiple access channel (MAC). Z represents the additive white Gaussian noise (AWGN) of the channel.

5.2 Problem Formulation

The distributed network studied in this chapter is modeled by a Gaussian CEO problem [29, 30] with Gaussian MAC, which is shown in Fig. 5.1. In this model [30], for each observation time $t = 1, 2, 3, ...$

$$Y_i(t) = X(t) + V_i(t) \qquad i = 1, ..., L \tag{5.1}$$

where $X(t) \sim \mathcal{N}(0, \sigma_X^2)$ and $V_i(t) \sim \mathcal{N}(0, N_i)$ which is *i.i.d.* over i and t. Random variables $Y_i(t)$ for $(i = 1, ..., L)$ are conditionally independent given the source $X(t)$. We represent n independent instances of $\{X(t)\}_{t=1}^{\infty}$ and $\{Y_i(t)\}_{t=1}^{\infty}$ by data sequences $X^n = \{X(1), X(2), ..., X(n)\}$ and $Y_i^n = \{Y_i(1), Y_i(2), ..., Y_i(n)\}$, respectively. The correlated sources cannot cooperate to directly exploit their correlation. Data sequences

Y_i^n are separately encoded to $\varphi_i\left(Y_i^n\right) = U_i^n$ where the encoder functions are defined as

$$\varphi_i : \mathcal{Y}_i^n \to \mathcal{C}_i^{(n)}, \tag{5.2}$$

for $i = 1, 2, ..., L$ and $\mathcal{C}_i^{(n)} = \left\{1, 2, ..., \left|\mathcal{C}_i^{(n)}\right|\right\}$ denotes the codebook of agent i. The agents communicate the coded sequences to the FC through an additive white Gaussian MAC. The transmitted sequences U_i^n are average-power limited to P_i, i.e.,

$$\frac{1}{n}\sum_{t=1}^{n} E\left[|U_i(t)|^2\right] \leq P_i + \delta \qquad i = 1, 2, ..., L \tag{5.3}$$

where δ is an arbitrary prescribed positive number. In other words, the coding function be chosen to ensure that $\lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} E\left[|U_i(t)|^2\right] \leq P_i$. The time-$t$ output of the Gaussian MAC is given by

$$W(t) = \sum_{j=1}^{L} U_j(t) + Z(t), \tag{5.4}$$

where the terms $\{Z(t)\}_{t=1}^{\infty}$ are *i.i.d.* zero-mean variance-σ_Z^2 Gaussian random variables that are independent of the source sequence. Based on the channel output W^n, the FC makes an estimate of the main source X^n as $\widehat{X}^n$. It produces the source estimate $\widehat{X}^n$ to an acceptable degree of fidelity D. The measure of the fidelity is the average distortion criterion, i.e., $\Delta = \frac{1}{n}E\left[\sum_{j=1}^{n} d(X(j), \widehat{X}(j))\right]$ where $d(X(j), \widehat{X}(j))$ is the MSE distortion measure. The reconstructed signal can be described by $\widehat{X}^n = \psi\left(W^n\right)$, where the decoder function is described by

$$\psi : \mathcal{W}^n \to \mathcal{X}^n, \tag{5.5}$$

and $\mathcal{W}$ is the common alphabet of the random variables $W(t)$ for $t = 1, 2, \cdots$.

Let $\boldsymbol{P} = (P_1, P_2, ..., P_L)$ and $\mathcal{F}_\delta^{(n)}\left(P_1, P_2, ..., P_L\right)$ denote all $(L+1)-$tuple encoder and decoder functions $(\varphi_1, ..., \varphi_L, \psi)$ that satisfy (5.2)-(5.5). For a particular coding scheme $(\varphi_1, ..., \varphi_L, \psi)$, the performance is determined by the required cost vector $\boldsymbol{P}$ and the incurred distortion D. For any target distortion $D \geq 0$, the power-distortion region is defined in [83], [38] as

$$\mathcal{P}(D) = \left\{(P_1, P_2, ..., P_L) \mid (\boldsymbol{P}, D) \ is \ admissible\right\}.$$

A power-distortion pair $(\boldsymbol{P}, D)$ is admissible if for any $\delta > 0$ and any $n \geq n_0(\delta)$ there exists an $(L+1)-$tuple set $(\varphi_1, ..., \varphi_L, \psi) \in \mathcal{F}_\delta^{(n)}(P_1, P_2, ..., P_L)$ such that $\Delta \leq D+\delta$. In other words, $(\boldsymbol{P}, D)$ is admissible if there is a coding scheme that can achieve a distortion close to D while satisfying the transmission cost constraints.

5.3 Power-Distortion Tradeoff

We present necessary and sufficient conditions for achievability of $(P_1, P_2, ..., P_L, D)$.

5.3.1 Necessary Condition

Theorem 6 *A necessary condition for the achievability of* $(P_1, P_2, ..., P_L, D)$ *is that*

$$R_X^{rem}(D) \leq \frac{1}{2}\log_2\left(1 + \frac{\overline{P}}{\sigma_Z^2}\right), \tag{5.6}$$

where

$$R_X^{rem}(D) = \frac{1}{2}\log_2\left(\frac{\sigma_X^2 \sum_{i=1}^L \frac{1}{N_i}}{D} \frac{1}{\frac{1}{D_0^*} - \frac{1}{D}}\right) \qquad D_0^* < D < \sigma_X^2 \tag{5.7}$$

$$D_0^* = \left(\frac{1}{\sigma_X^2} + \sum_{i=1}^L \frac{1}{N_i}\right)^{-1}, \tag{5.8}$$

and

$$\overline{P} = \sum_{i=1}^L P_i + 2\sigma_X^2 \sum_{i=1}^L \sum_{j=2, j\neq i}^L \sqrt{\frac{P_i P_j}{(\sigma_X^2 + N_i)(\sigma_X^2 + N_j)}}. \tag{5.9}$$

Proof This necessary condition follows from two different concepts: data processing inequality and the remote source coding. Consider a block of n source symbols X^n and the corresponding observations of the L agents $\boldsymbol{Y}^n = \{Y_i^n\}_{i=1}^L$. If we denote the channel inputs produced by an arbitrary code by $\boldsymbol{U}^n = \{U_i^n\}_{i=1}^L$, the corresponding block of the channel output can be represented by W^n. By using the data processing inequality, the mutual information between the observation vector and the vector of the source estimate,

$I(\boldsymbol{Y}^n; \widehat{X}^n)$, can be upper bounded by the mutual information between transmitted waveforms and received waveform, $I(U_1^n, U_2^n, ..., U_L^n; W^n) = I(\boldsymbol{U}^n; W^n)$. This latter is upper bounded by the capacity of an additive white Gaussian noise channel with the power constraint $\sum_{i=1}^{L} P_i + 2\sigma_X^2 \sum_{i=1}^{L} \sum_{j=2, j\neq i}^{L} \sqrt{\frac{P_i P_j}{(\sigma_X^2 + N_i)(\sigma_X^2 + N_j)}}$. This can be shown as follows:

$$\begin{aligned} \tfrac{1}{n} I(\boldsymbol{U}^n; W^n) &= \tfrac{1}{n} h(W^n) - \tfrac{1}{n} h(W^n \mid \boldsymbol{U}^n) \\ &\overset{(a)}{=} \tfrac{1}{n} h(W^n) - \tfrac{1}{n} h(Z^n \mid \boldsymbol{U}^n) \\ &\overset{(b)}{=} \tfrac{1}{n} h(W^n) - \tfrac{1}{n} h(Z^n) \end{aligned} \tag{5.10}$$

where (a) follows from the equation $W^n = \sum_{i=1}^{L} U_i^n + Z^n$ and from the definition of conditional entropy, and (b) follows from the fact that U_i^n's are independent of Z^n. To calculate $\frac{1}{n} h(W^n)$ in (5.10) we need to obtain

$$\begin{aligned} &Var\,(W(k)) = Var\,(U_1(k) + \cdots + U_L(k) + Z(k)) = \\ &\left\{ \sum_{i=1}^{L} \sigma_{i,k}^2 + 2 \sum_{i=1}^{L} \sum_{j=2, j\neq i}^{L} \rho_{U_i(k) U_j(k)} \sqrt{\sigma_{i,k}^2 \sigma_{j,k}^2} + \sigma_Z^2 \right\} \overset{(c)}{\leq} \\ &\left\{ \sum_{i=1}^{L} \sigma_{i,k}^2 + 2 \sum_{i=1}^{L} \sum_{j=2, j\neq i}^{L} \rho_{Y_i(k) Y_j(k)} \sqrt{\sigma_{i,k}^2 \sigma_{j,k}^2} + \sigma_Z^2 \right\} \end{aligned} \tag{5.11}$$

where $U_i(k)$ is the k-th component of $U_i^n = \varphi_i(Y_i^n)$ and $\sigma_{i,k}^2 = Var\,(U_i(k)) = E\left[|U_i(k)|^2\right]$. Inequality of (c) follows from the fact that the maximum correlation coefficient between any two finite-variance functions of two jointly Gaussian random variables is equal to the correlation coefficient between original random variables. More precisely, if $Corr(X, Y)$ is the classical (Pearson) correlation between X and Y, the maximum correlation coefficient is defined in [84] as

$$\begin{aligned} \rho_{sup}(X, Y) \;=\; & sup\,\{Corr\,(\varphi_1(X), \varphi_2(Y)) : \\ & 0 < E\,|\varphi_1(X)|^2 < \infty, \\ & 0 < E\,|\varphi_2(Y)|^2 < \infty \quad \} \end{aligned} \tag{5.12}$$

It is shown that [85] for a bivariate Gaussian vector (X, Y), $\rho_{sup}(X, Y) = |Corr(X, Y)|$. It completes the justification of (c). From (5.1), we have $\rho_{Y_i Y_j} = \frac{\sigma_X^2}{\sqrt{(\sigma_X^2 + N_i)(\sigma_X^2 + N_j)}}$.

Substituting this correlation coefficient in (5.11) and using the power constraints of (5.3) for the transmitted coded sequences U_i^n's, and also using the Cauchy-Schwarz inequality [86] we will obtain

$$\frac{1}{n}\sum_{k=1}^{n} Var\left(W(k)\right) \leq$$

$$\sum_{i=1}^{L} P_i + 2\sigma_X^2 \sum_{i=1}^{L} \sum_{j=2, j\neq i}^{L} \sqrt{\frac{P_i P_j}{\left(\sigma_X^2 + N_i\right)\left(\sigma_X^2 + N_j\right)}} + \sigma_Z^2 \tag{5.13}$$

Since *i.i.d.* random variables with normal distribution maximize the differential entropy subject to sum of the variances [5], we obtain $\frac{1}{n}h(W^n) \leq \frac{1}{2}\log\left\{2\pi e\left(\overline{P} + \sigma_Z^2\right)\right\}$ where $\left(\overline{P} + \sigma_Z^2\right)$ is the right hand side of (5.13). Thus, (5.10) can be simplified as

$$\frac{1}{n} I(\boldsymbol{U}^n; W^n) \leq \frac{1}{2}\log_2\left(1 + \frac{\overline{P}}{\sigma_Z^2}\right) \tag{5.14}$$

Therefore, using the data processing inequality, $\frac{1}{n}I(\boldsymbol{Y}^n; \widehat{X}^n) \leq \frac{1}{2}\log_2\left(1 + \frac{\overline{P}}{\sigma_Z^2}\right)$. On the other hand, a lower bound on $I(Y_1, ..., Y_L; \widehat{X})$ can be obtained by considering the remote source coding problem introduced in [3], where the agents in the CEO problem are allowed to collaborate, i.e., the L-tuple $\underline{Y} = (Y_1, ..., Y_L)$ is observed by one common encoder. The scenario is also presented in Chapter 2 and Chapter 3 and it is illustrated in Fig. 3.14. The rate-distortion function for the remote source coding problem of Fig. 3.14 is obtained in Chapter 3 which is expressed in (5.7). Now we put this lower bound of $I(\boldsymbol{Y}^n; \widehat{X}^n)$ less than or equal to the upper bound of (5.14) (which is the capacity of the AWGN channel with the power constraint of (5.13)),

$$R_X^{rem}(D) \leq \frac{1}{n} I(\boldsymbol{Y}^n; \widehat{X}^n) \leq \frac{1}{n} I(\boldsymbol{U}^n; W^n) \leq \frac{1}{2}\log_2\left(1 + \frac{\overline{P}}{\sigma_Z^2}\right) \tag{5.15}$$

Thus, the necessary condition in (5.6) will be obtained. This completes the proof.

By substituting (3.52) in (5.6) for the symmetric case, the following result will be obtained:

Corollary 3 *In the symmetric case of Gaussian CEO problem where $N_1 = N_2 = \cdots = N_L = N$ and $P_1 = P_2 = \cdots = P_L = P$, the necessary condition for the achievability of (P, D_l) can be expressed as*

$$D_l \geq \sigma_X^2 \frac{\sigma_Z^2 (\sigma_X^2 + N) + NLP}{\sigma_Z^2 (\sigma_X^2 + N) + LP (L\sigma_X^2 + N)}. \tag{5.16}$$

5.3.2 Sufficient Condition

Based on analyzing the uncoded transmission in the CEO problem, we present a sufficient condition for achievability of $(P_1, P_2, ..., P_L, D)$. In this approach which is also called "analog forwarding" [83] or "amplify-and-forward" [87] approach, each agent transmits the scaled version of its observation, scaled to its power constraint, i.e.,

$$U_i(t) = \alpha_i Y_i(t) \ \ where \quad \alpha_i = \sqrt{\frac{P_i}{\sigma_X^2 + N_i}} \tag{5.17}$$

According to Fig. 5.2, the received signal at the FC is

$$\begin{aligned} W(t) &= \textstyle\sum_{i=1}^L U_i(t) + Z(t) \\ &= \textstyle\sum_{i=1}^L \sqrt{\frac{P_i}{\sigma_X^2+N_i}} X(t) + \sqrt{\frac{P_i}{\sigma_X^2+N_i}} V_i(t) + Z(t). \end{aligned} \tag{5.18}$$

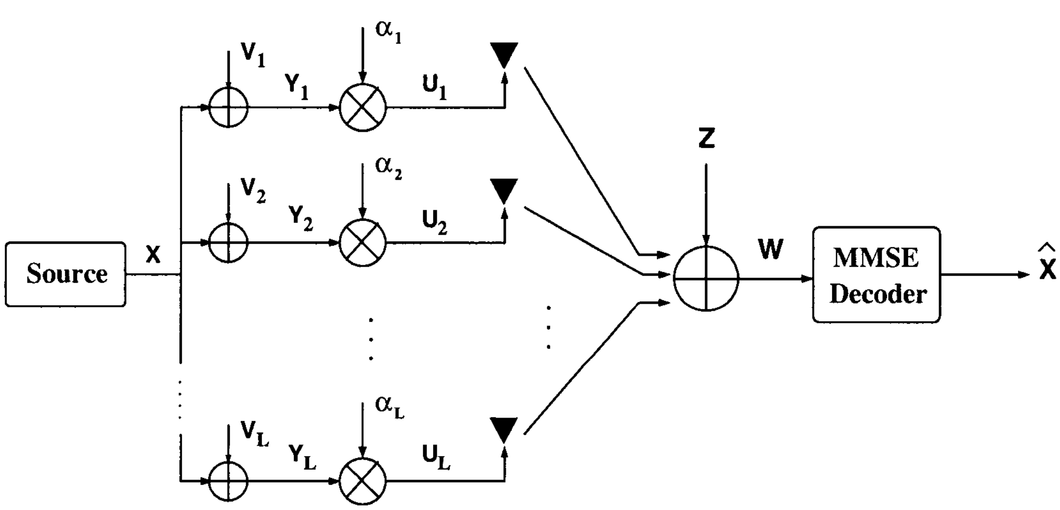

Figure 5.2: Uncoded transmission in a Gaussian CEO problem with Gaussian MAC.

Since the encoding is memoryless, the optimum estimator is the MMSE estimator of $X(t)$ from the received signals $\{W(t) : 1 \leq t < \infty\})$ [38], which can be obtained by $\widehat{X}(t) = E[X(t) \mid W(t)]$. The average cost of MMSE estimator, which is the MSE distortion D_u, satisfies

$$\frac{1}{D_u} = \frac{1}{\sigma_X^2} + \left(\sum_{i=1}^{L} \sqrt{\frac{P_i}{\sigma_X^2 + N_i}}\right)^2 \left(\sum_{i=1}^{L} \frac{P_i N_i}{\sigma_X^2 + N_i} + \sigma_Z^2\right)^{-1}. \tag{5.19}$$

The power-distortion region achieved by the uncoded transmission approach can be considered as an inner region for the whole power-distortion region. Therefore, we will have the following sufficient condition for the achievability of $(P_1, P_2, ..., P_L, D)$:

Theorem 7 *For the $(L+1)$-tuple $(P_1, P_2, ..., P_L, D)$ to be achievable it suffices that the following condition holds:*

$$\frac{1}{D} \leq \frac{1}{\sigma_X^2} + \left(\sum_{i=1}^{L} \sqrt{\frac{P_i}{\sigma_X^2 + N_i}}\right)^2 \left(\sum_{i=1}^{L} \frac{P_i N_i}{\sigma_X^2 + N_i} + \sigma_Z^2\right)^{-1}. \tag{5.20}$$

In the symmetric case, the sufficient condition can be represented by

$$\frac{1}{D} \leq \frac{1}{\sigma_X^2} + \left(\frac{L^2 P}{\sigma_X^2 + N}\right) \left(\frac{LPN}{\sigma_X^2 + N} + \sigma_Z^2\right)^{-1}. \tag{5.21}$$

Comparing the result of Corollary 3 and Equation (5.21) reveals that the necessary condition and the sufficient condition agree in the symmetric case. Hence, we derive the "optimum performance theoretically attainable" (OPTA) [3], for the symmetric Gaussian CEO problem with additive white Gaussian MAC. This is the same result as the recent work of [40] which is obtained independently.

Corollary 4 *For the symmetric Gaussian CEO problem with additive white Gaussian MAC,*

(i) The optimal distortion-power tradeoff can be represented by

$$D = \sigma_X^2 \frac{\sigma_Z^2 (\sigma_X^2 + N) + NLP}{\sigma_Z^2 (\sigma_X^2 + N) + LP (L\sigma_X^2 + N)} \tag{5.22}$$

(ii) Uncoded transmission achieves the optimal distortion-power tradeoff in (5.22) and in fact is the optimal transmission strategy for symmetric Gaussian CEO problem with Gaussian MAC.

Remark 5 (Quasi-Static Flat Rayleigh Fading Channel) The same analysis can be done for transmission of a Gaussian source over a quasi-static (slow varying) flat Rayleigh fading channel. We expect that, when the amplitudes of fading coefficients are the same, uncoded transmission performs optimally and can achieve the smallest possible distortion.

Remark 6 (Communication between the agents) In general, we expect that the communication between agents enhance the performance. However, in [38], based on analyzing the idealized system, it is shown that as the number of agents grows to infinity, communication among agents does not improve the performance of symmetric CEO problem. Our analysis shows that the same result holds for any finite number of agents in the symmetric case. More specifically, in the derivation of the necessary condition in (5.6) we assume that the noisy observations of the source X are given to one common encoder. It means that the necessary condition in (5.6) is obtained based on full cooperation among agents. Since (5.6) is the optimal power-distortion tradeoff in the symmetric case of Gaussian CEO problem, hence, there is no penalty for the fact that the agents are distributed in the symmetric CEO problem. In other words, we do not get a better performance if we allow communication between the agents.

5.4 Chapter Summary

It is well known that for a point-to-point transmission of a Gaussian source across an AWGN channel, the uncoded transmission achieves the optimal power-distortion tradeoff. Also, it is shown by Gastpar and Vetterli that for the considered Gaussian CEO problem under Gaussian multiple access, as the number of agents grows to infinity, the uncoded transmission achieves the smallest possible distortion. However, it was unknown what the optimal power-distortion tradeoff is in a Gaussian CEO problem with a finite number of agents and whether the uncoded transmission achieves the optimal power-distortion tradeoff. In this chapter, first we provided necessary and sufficient conditions for achievability of $(L+1)-$tuples $(P_1, P_2, ..., P_L, D)$. Then we obtained the optimal power-distortion tradeoff for the symmetric case and proved that for any finite number of agents, uncoded transmission performs optimally and achieves the smallest possible distortion.

Chapter 6

Conclusion

In multi-user information theory, multiterminal source coding is of special interest. It considers separate coding of multiple correlated sources that are not allowed to collaborate to exploit their correlation and reduce transmission rate (and save transmission power).

Although the theory of multiterminal or distributed source coding started more than 30 years ago, finding a complete characterization for its rate-distortion region still remains open. It is one of the long-standing open problems in information theory and even concrete examples of this problem are hard to analyze. For instance, the whole rate-distortion region of the *2-terminal* source coding scheme for Gaussian sources with MSE distortion is characterized in 2005.

Previous works on coding of multiterminal schemes are based on joint decoding of all messages. It requires that all messages be decoded simultaneously which is exponentially more complex than a sequence of single-message decodings. Inspired by previous work on successive coding strategy, we consider successive Wyner-Ziv coding as an enabling technique with low complexity that allows us to analyze the multiterminal coding schemes with finite number of sources. We focus on Gaussian sources and determine the rate-region for three different multiterminal coding schemes based on the successive coding strategy: CEO problem (multiterminal remote source coding scheme), *1-helper* coding scheme (a special case of many-help-one problem with one helper), and

2-terminal source coding scheme.

First, we apply the successive coding strategy to the Gaussian CEO problem and derive the optimal rate allocation scheme to achieve the minimum distortion under a sum-rate constraint. We prove that the sum-rate distortion function of the Gaussian CEO problem can be achieved by a sequence of successively structured Wyner-Ziv codes. Therefore, the high complexity optimal source code can be decomposed into a sequence of low complexity Wyner-Ziv codes. We show that if the sum-rate tends to infinity for a finite number of agents, the optimal rate allocation strategy assigns equal rates to all sensors. The same result is obtained when the number of agents tends to infinity while the sum-rate is finite.

Then, we consider *1-helper* coding scheme where one source (auxiliary source which is also called the helper) provides partial side information to the decoder to help the reconstruction of the other source (primary source). Our results show that the successive coding strategy is an optimal strategy in this problem in the sense of achieving the rate-distortion function. For the *2-terminal* source coding scheme, we develop connections between source encoding and data fusion steps and prove that the whole rate-distortion region is achievable using the successive coding strategy. Comparing the performance of the sequential coding with the performance of the successive coding, we show that there is no sum-rate loss when the side information is not available at the encoder. This result is of special interest in some applications such as video coding where there are processing and storage constraints at the encoder. Based on the successive coding strategy, we provide an achievable rate-distortion region for the *m-terminal* source coding scheme.

We also consider transmission of noisy versions of a Gaussian source through a Gaussian multiple access channel to a single FC. It can be modeled as the Gaussian CEO problem with the Gaussian MAC. The decoder wishes to reconstruct the main source with an average distortion D at the smallest possible power consumption in the communication link. Our goal is to characterize the power-distortion region achievable by any coding strategy regardless of delay and complexity. We obtain a necessary condition for achievability of all power-distortion tuples $(P_1, P_2, ..., P_L, D)$. Also, analyzing the uncoded transmission scheme provides a sufficient condition for achievability of $(P_1, P_2, ..., P_L, D)$. Then, we consider a symmetric case of the problem where the observations of agents have the same noise level and the transmitting signals are subject to the same average power constraint. We show that in this case the necessary and sufficient conditions coincide and give the optimal power-distortion region. Therefore, in the symmetric case of Gaussian CEO problem with additive white Gaussian MAC uncoded transmission performs optimally for any finite number of agents. This is the same result as the recent work of Gastpar [40] which is obtained independently.

6.1 Suggestions of future research

In this work we considered *2-terminal* source coding where the encoders are not allowed to cooperate before transmission to the joint decoder. One can consider *2-terminal* source coding scheme with partially cooperating encoders, where the encoders are connected by communication links with finite capacities. Hence, before encoding and transmitting their data to the decoder, they exchange information to increase the reliability of their information. Since the successive coding allows determining the achievable region for multiterminal schemes, by applying the successive coding strategy, an achievable rate-distortion tradeoff of the cooperative *2-terminal* source coding scheme can be obtained. Then one can verify that if this achievable tradeoff is the rate-distortion function of the scheme. As a special case, consider cooperative *1-helper* problem, where only one link from the primary source to the auxiliary source is added to the *1-helper* scheme introduced before. Analyzing this scheme is useful for some practical applications such as relay networks.

In this thesis we also obtained an achievable rate region for *2-helper* and in general *m-helper* problem based on successive coding strategy. One can verify whether the successive coding is optimal in those coding schemes. We expect that in the Gaussian / MSE case, successive coding can achieve the optimal performance in multiterminal source coding schemes.

If one node drops out of a wireless network, due to hardware failure or any other reason, a joint decoder may not be able to decode and nothing can be recovered. By using timesharing of successive coding schemes in a multiterminal scenario, we may obtain a nontrivial estimate of the source. However, depending on the order of decodings at the FC, some schemes will perform better than others. An analysis of a 4-node Gaussian CEO problem is presented in [61], where a tradeoff between robustness and efficiency for different successive schemes is provided by studying the combinatorial properties of the associated directed graphs. General analysis of this tradeoff for the Gaussian CEO problem with L agents and for *m-terminal* source coding scheme is an open area of research.

To code or not to code? This has been an essential question which has not been answered for many multiterminal schemes. Based on analyzing the remote source coding scheme for the CEO problem and using data processing inequality, we proved that the uncoded transmission is optimal in the symmetric Gaussian CEO problem. One can use similar approach for multiterminal coding schemes such as asymmetric *2-terminal* coding scheme, asymmetric CEO problem, or *m-terminal* coding scheme $(m > 2)$ and try to answer following questions for each scheme:

1. Does uncoded transmission perform optimally?
2. What is the optimal power-distortion region for the scheme?
3. What kind of coding strategy can achieve the optimal power-distortion tradeoff of the scheme?

Bibliography

[1] C. E. Shannon, "A mathematical theory of communication," *Bell Systems Technical Journal*, vol. 27, pp. 379–423, pp. 623–656, 1948.

[2] ——, "Coding theorems for a discrete source with a fidelity criterion," *IRE Convention Rec.*, vol. 7, p. 142, 1959.

[3] T. Berger, *Rate-distortion theory: A mathematical basis for data compression.* Englewood Cliffs, NJ: Prentice-Hall, 1971.

[4] R. J. McEliece, "The theory of information and coding," *Encylcopedia of Mathematics and its Applications*, vol. 3, Addison-Wesley, Reading, Mass., 1977.

[5] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* New York: Wiley, 1991.

[6] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression.* Norwell, MA: Kluwer Academic Publishers, 1992.

[7] T. Berger and J. D. Gibson, "Lossy source coding," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2693–2723, Oct. 1998.

[8] S. C. Draper, "Successive structuring of source coding algorithms for data fusion, buffering, and distribution in networks," Ph.D. dissertation, Massachusetts Institute of Technology (MIT), Cambridge, MA, Jun. 2002.

[9] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inf. Theory*, vol. 22, pp. 1–10, Jan. 1976.

[10] A. D. Wyner, "The rate-distortion function for source coding with side information at the decoder ii: General sources," *Information and Control*, vol. 38, pp. 60–80, Jul. 1978.

[11] R. Dobrushin and B. Tsybakov, "Information transmission with additional noise," *IEEE Trans. Inf. Theory*, vol. 8, no. 6, pp. 293–304, Sep. 1962.

[12] J. K. Wolf and J. Ziv, "Transmission of noisy information to a noisy receiver with minimum distortion," *IEEE Trans. Inf. Theory*, vol. 16, no. 4, pp. 406–411, Jul. 1970.

[13] K. Eswaran, "Remote source coding and AWGN CEO problems," Master's thesis, EECS Department, University of California, Berkeley, Jan. 2006.

[14] H. Yamamoto and K. Itoh, "Source coding theory for multiterminal communication systems with a remote source," *Trans. of the IECE of Japan*, vol. E63, no. 10, pp. 700–706, Oct. 1980.

[15] T. J. Flynn and R. M. Gray, "Encoding of correlated observations," *IEEE Trans. Inf. Theory*, vol. 33, no. 6, pp. 773–787, Nov. 1987.

[16] T. Berger, "Multiterminal source coding," *The Information Theory Approach to Communications (CISM Courses and Lectures)*, pp. 171–231, 1978.

[17] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. 19, pp. 471–480, Jul. 1973.

[18] A. B. Wagner, S. Tavildar, and P. Viswanath, "The rate region of the Quadratic Gaussian two-terminal source-coding problem," *IEEE Trans. Inf. Theory*, submitted for publication, Feb. 2006. [Online]. Available: http://www.arxiv.org/PS_cache/cs/pdf/0510/0510095.pdf

[19] S. D. Servetto, "The sum-rate of multiterminal source coding with two encoders," *IEEE Trans. Inf. Theory*, submitted for publication, Apr. 2006 (original title: The Region of Achievable Rates for Multiterminal Source Coding, Part I).

[20] A. J. Goldsmith and S. B. Wicker, "Design challenges for energy-constrained ad hoc wireless networks," *IEEE Wireless Commun. Mag.*, vol. 9, no. 4, pp. 8–27, Aug. 2002.

[21] M. B. Westover, "Image representation and pattern recognition in brains and machines," Ph.D. dissertation, Washington University, St. Louis, MO, Sep. 2006.

[22] M. B. Westover and J. A. O'Sullivan, "Achievable rates for pattern recognition," *IEEE Trans. Inf. Theory*, submitted for publication, Sep. 2005. [Online]. Available: http://arxiv.org/abs/cs.IT/0509022/

[23] M. Gastpar, "The Wyner-Ziv problem with multiple sources," *IEEE Trans. Inf. Theory*, vol. 50, no. 11, pp. 2762–2768, Nov. 2004.

[24] T. M. Cover and M. Chiang, "Duality between channel capacity and rate distortion with two-sided state information," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1629–1638, Jun. 2002.

[25] S. S. Pradhan, J. Chou, and K. Ramchandran, "Duality between source coding and channel coding and its extension to the side information case," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1181–1203, May 2003.

[26] R. J. Barron, B. Chen, and G. W. Wornell, "The duality between information embedding and source coding with side information and some applications," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1159–1180, May 2003.

[27] W. Yu, "Duality and the value of cooperation in distributive source and channel coding problems," in *Proc. 41st Allerton Conf. on Communication, Control and Computing*, Urbana, IL, 2003.

[28] H. Wang and P. Viswanath, "Fixed binning schemes for channel and source coding problems: An operational duality," in *Proc. 38th Annual Conference on Information Sciences and Systems (CISS)*, Princeton, NJ, Mar. 2004.

[29] T. Berger, Z. Zhang, and H. Viswanathan, "The CEO problem," *IEEE Trans. Inf. Theory*, vol. 42, no. 3, pp. 887–902, May 1996.

[30] H. Viswanathan and T. Berger, "The Quadratic Gaussian CEO problem," *IEEE Trans. Inf. Theory*, vol. 43, no. 5, pp. 1549–1559, Sep. 1997.

[31] Y. Oohama, "Multiterminal source coding for correlated memoryless Gaussian sources with several side information at the decoder," in *Proc. IEEE Inf. Theory and Comm. Workshop*, Jun. 1999, p. 100.

[32] J. Chen, X. Zhang, T. Berger, and S. B. Wicker, "An upper bound on the sum-rate distortion function and its corresponding rate allocation schemes for the CEO problem," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 977–987, Aug. 2004.

[33] V. Prabhakaran, D. Tse, and K. Ramchandran, "Rate region of the quadratic Gaussian CEO problem," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Chicago, IL, USA, Jun. 2004, p. 119.

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

[35] A. Pandya, A. Kansal, G. Pottie, and M. Srivastava, "Lossy source coding of multiple Gaussian sources: *m-helper* problem," in *Proc. IEEE Inf. Theory Workshop (ITW)*, San Antonio, Texas, Oct. 2004.

[36] T. J. Goblick, "Theoretical limitations on the transmission of data from analog sources," *IEEE Trans. Inf. Theory*, vol. 11, pp. 558–567, Oct. 1965.

[37] M. Gastpar, B. Rimoldi, and M. Vetterli, "To code, or not to code: Lossy source channel communication revisited," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1147–1158, May 2003.

[38] M. Gastpar and M. Vetterli, "Source-channel communication in sensor networks," in *Proc. 2nd Int. Workshop Information Processing in Sensor Networks*, Palo Alto, CA, Apr. 2003, pp. 162–177.

[39] A. Lapidoth and S. Tinguely, "Sending a Bi-variate Gaussian source over a Gaussian MAC," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Seattle, Washington,, Jul. 2006, pp. 2124–2128.

[40] M. Gastpar, "Uncoded transmission is exactly optimal for a simple Gaussian sensor network," in *Proc. 2007 Information Theory and Applications Workshop*, San Diego, CA, USA, Jan. 2007.

[41] T. Berger and R. Yeung, "Multiterminal source encoding with one distortion criterion," *IEEE Trans. Inf. Theory*, vol. 35, no. 2, pp. 228–236, Mar. 1989.

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

[43] ——, "The rate-distortion function for the Quadratic Gaussian CEO problem," *IEEE Trans. Inf. Theory*, vol. 4, no. 3, pp. 1057–1070, May 1998.

[44] Z. Xiong, V. Stankovic, S. Cheng, A. Liveris, and Y. Sun, "Source-channel coding for algebraic multiterminal binning," in *Proc. IEEE Inf. Theory Workshop (ITW)*, San Antonio, Texas, Oct. 2004.

[45] S. C. Draper and G. W. Wornell, "Side information aware coding strategies for sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 966–976, Aug. 2004.

[46] A. D. Wyner, "On source coding with side information at the decoder," *IEEE Trans. Inf. Theory*, vol. 21, pp. 294–300, May 1975.

[47] R. F. Ahlswede and J. K¨orner, "Source coding with side information and a converse for degraded broadcast channels," *IEEE Trans. Inf. Theory*, vol. 21, pp. 629–637, Nov. 1975.

[48] J. K¨orner and K. Marton, "How to encode the module-two sum of binary sources," *IEEE Trans. Inf. Theory*, vol. 25, pp. 219–221, Mar. 1975.

[49] S. I. Gelfand and M. S. Pinsker, "Source coding with incomplete side information," *(in Russian) Probl. Pered. Inform.*, vol. 15, no. 2, pp. 45–57, 1979.

[50] S. Y. Tung, "Multiterminal source coding," Ph.D. dissertation, School of Electrical Engineering, Cornell University, Ithaca, NY, May 1978.

[51] Y. Oohama, "Gaussian multiterminal source coding with several side informations at the decoder," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Seattle, Washington, Jul. 2006.

[52] V. N. Koshelev, "Hierarchical coding of discrete sources," *Probl. Pered. Inform.*, vol. 16, no. 3, pp. 31–49, 1980.

[53] W. H. R. Equitz and T. M. Cover, "Successive refinement of information," *IEEE Trans. Inf. Theory*, vol. 37, no. 2, pp. 269–275, Mar. 1991.

[54] B. Rimoldi, "Successive refinement of information: Characterization of achievable rates," *IEEE Trans. Inf. Theory*, vol. 40, no. 1, pp. 253–259, Jan. 1994.

[55] Y. Steinberg and N. Merhav, "On successive refinement for the Wyner-Ziv problem," *IEEE Trans. Inf. Theory*, vol. 50, no. 8, pp. 1636–1654, Aug. 2004.

[56] S. Cheng and Z. Xiong, "Successive refinement for the Wyner-Ziv problem and layered code design," *IEEE Trans. Signal Process.*, vol. 53, no. 2, pp. 3269–3281, Aug. 2005.

[57] C. Tian and S. Diggavi, "On multistage successive refinement for Wyner-Ziv source coding with degraded side information," *IEEE Trans. Inf. Theory*, submitted for publication, Apr. 2006. [Online]. Available: http://licos.epfl.ch/Papers/MultiStageWZ.pdf

[58] J. Chen and T. Berger, "Successive Wyner-Ziv coding scheme and its application to the quadratic Gaussian CEO problem," *IEEE Trans. Inf. Theory*, submitted for publication, Apr. 2006. [Online]. Available: http://arxiv.org/PS_cache/cs/pdf/0604/0604077.pdf

[59] B. Rimoldi and R. Urbanke, "Asynchronous Slepian-Wolf coding via source-splitting," in *Proc. IEEE Int. Symp. Information Theory*, Ulm, Germany, Jun. 1997, p. 271.

[60] R. Zamir, S. Shamai, and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1250–1276, Jun. 2002.

[61] X. Zhang and S. B. Wicker, "Robustness vs. Efficiency for the Quadratic Gaussian CEO problem," in *Proc. Allerton Conference*, Sep. 2004.

[62] S. I. Gelfand and M. S. Pinsker, "Coding of sources on the basis of observations with incomplete information," *Problems of Information Transmission*, vol. 15, no. 2, pp. 115–125, 1979.

[63] H. Yamamoto, "Wyner-Ziv theory for a general function of the correlated sources," *IEEE Trans. Inf. Theory*, vol. 28, no. 5, pp. 803–807, Sep. 1982.

[64] R. Zamir and T. Berger, "Multiterminal source coding with high resolution," *IEEE Trans. Inf. Theory*, vol. 45, no. 1, pp. 106–117, Jan. 1999.

[65] J. Chen, X. Zhang, T. Berger, and S. B. Wicker, "Rate allocation in distributed sensor network," in *Proc. Allerton Conference*, 2003.

[66] X. Zhang and S. B. Wicker, "On the rate region of the vector Gaussian CEO problem," in *Proc 39th Annual Conference on Information Sciences and Systems (CISS)*, The Johns Hopkins University, Baltimore, MD, USA, Mar. 2005.

[67] Y. Oohama, "Rate distortion region for separate coding of correlated Gaussian remote observations," in *Proc. 43rd Allerton Conference*, Sep. 2005.

[68] J. J. Xiao and Z. Q. Luo, "Optimal rate allocation for the vector Gaussian CEO problem," in *Proc. 1st IEEE Int. Workshop on Computational advances in Multi-Sensor Adaptive Processing*, Dec. 2005, pp. 56–59.

[69] S. Tavildar and P. Viswanath, "On the sum-rate of the vector Gaussian CEO problem," in *Proc. Asilomar conference on signals and systems*, 2005.

[70] S. C. Draper and G. W. Wornell, "Successively structured CEO problems," in *Proc. IEEE Int. Symp. Information Theory*, Lausanne, Switzerland, Jun. 2002, p. 65.

[71] C. Shannon, *The Mathematical Theory of Communication.* Urbana: University of Illinois Press, 1949.

[72] K. Eswaran and M. Gastpar, "Rate loss in the CEO problem," in *Proc 39th Annual Conference on Information Sciences and Systems (CISS)*, The Johns Hopkins University, Baltimore, MD, USA, Mar. 2005.

[73] E. C. van der Meulen, "Transmission of information in a T-terminal discrete memoryless channel," Ph.D. dissertation, Dept. Statistics, Univ. California, Berkeley, CA, 1968.

[74] T. M. Cover and A. A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sep. 1979.

[75] B. Schein and R. Gallager, "The Gaussian parallel relay network," in *Proc. IEEE Int. Symp. Information Theory*, Sorrento, Italy, Jun. 2000, p. 22.

[76] B. E. Schein, "Distributed coordination in network information theory," Ph.D. dissertation, Massachusetts Institute of Technology (MIT), Cambridge, MA, Sep. 2001.

[77] M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[78] S. Shamai, S. Verdu, and R. Zamir, "Systematic lossy source/channel coding," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 564–579, Mar. 1998.

[79] Y. Oohama, "Separate source coding of correlated gaussian remote sources," in *Proc. 2006 Information Theory and Applications Workshop*, San Diego, CA, USA, Jan. 2006.

[80] H. Viswanathan and T. Berger, "Sequential coding of correlated sources," *IEEE Trans. Inf. Theory*, vol. 46, no. 1, pp. 236–246, Jan. 2000.

[81] Y. Yang, V. Stankovic, Z. Xiong, and W. Zhao, "On multiterminal source code design," in *Proc. IEEE Data Compression Conference (DCC), J. A. Storer, M. Cohn (eds.), IEEE Computer Society Press*, Snowbird, Utah, Mar. 2005.

[82] ——, "On multiterminal source code design," *IEEE Trans. Inf. Theory*, submitted for publication.

[83] J.-J. Xiao and Z.-Q. Luo, "Multiterminal source-channel communication under orthogonal multiple access," *IEEE Trans. Inf. Theory*, submitted for publication. [Online]. Available: http://www.ece.umn.edu/users/luozq/recent_work.html

[84] H. O. Hirschfeld, "A connection between correlation and contingency," *Proc. Cambridge Philos. Soc.*, vol. 31, pp. 520–524, 1935.

[85] H. O. Lancaster, "Some properties of the bivariate normal distribution considered in the form of a contingency table," *Biometrika*, vol. 44, pp. 289–292, 1957.

[86] R. V. Kadison, "A generalized schwarz inequality and algebraic invariants for operator algebras," *Annals of Mathematics*, vol. 56, pp. 494–503, Nov. 1952.

[87] J. N. Laneman, "Cooperative diversity in wireless networks: Algorithms and architectures," Ph.D. dissertation, Massachusetts Institute of Technology (MIT), Cambridge, MA, Aug. 2002.