HELPER-ASSISTED LOSSY MULTITERMINAL SOURCE CODING: RATE-DISTORTION ANALYSIS AND CODING DESIGN

WENSHENG LIN

Japan Advanced Institute of Science and Technology

Doctoral Dissertation

HELPER-ASSISTED LOSSY MULTITERMINAL SOURCE CODING: RATE-DISTORTION ANALYSIS AND CODING DESIGN

WENSHENG LIN

Supervisor: Professor Tad Matsumoto

Graduate School of Advanced Science and Technology
Japan Advanced Institute of Science and Technology

Information Science

September 2019

External reviewers:
Professor Pierluigi Salvo Rossi
Professor Gerhard Kramer
Postdoctoral Researcher Germán Bassi

Internal reviewers:
Associate Professor Brian Michael Kurkoski
Professor Mineo Kaneko

Abstract

This dissertation investigates several topics belonging to the category of helper-assisted lossy multiterminal source coding, including multiterminal source coding with a helper, binary chief executive officer (CEO) problem with a helper, lossy source coding with helpers, lossy communications with lossy-forwarding (LF), and practical coding design.

Initially, for multiterminal source coding with a helper, we derive an inner bound on the achievable rate-distortion region, which is then utilized to evaluate the upper bound of the outage probability over block Rayleigh fading channels. The numerical results demonstrate the performance improvement and the diversity gain by introducing a helper. Interestingly, the system with a helper has higher energy efficiency while also reducing the outage probability.

Subsequently, we solve the binary CEO problem with a helper by decomposing it into two steps as multiterminal source coding and final decision. We derive an outer bound on the achievable rate-distortion region, and formulate a convex optimization problem to minimize the distortions at the first step of multiterminal source coding with a helper. For the step of final decision, we investigate the distortion propagating from the joint decoding results to the final decision.

Moreover, we present an inner bound on the achievable rate-distortion region for lossy source coding with helpers by the proof of achievability. The theoretical inner bound is verified to be a generalization of the Wyner-Ziv theorem.

For lossy communications with an LF relay, we determine an inner bound on the achievable rate-distortion region of lossy source coding with a helper for the first step. Then, we calculate the upper bound of the outage probability over block Rayleigh fading channels. We also conduct a series of simulations to compare the outage performance of LF with that of amplify-and-forward (AF) and decode-and-forward (DF).

Finally, we develop the hybrid majority voting (HMV) code for practical lossy compression. We theoretically analyze the rate-distortion performance of the HMV code, and prove that it has superior performance in spite of low complexity. In addition, we find the bit flipping (BF) code as the complement code of HMV code for successive refinement. By this means, the distortion in the standalone link can be conspicuously reduced while the refinement link can keep almost the same performance as before. Furthermore, we conclude the methodology of hybrid codes design for lossy source coding. We develop the hybrid code based on the Hamming codes as an example, and also find its syndrome as the complement code for successive refinement.

Keywords: Multiterminal source coding, lossy compression, side information, rate-distortion, outage probability.

Acknowledgment

The research works presented in this dissertation were carried out in part at the Graduate School of Advanced Science and Technology, Japan Advanced Institute of Science and Technology (JAIST), Japan, and in part at the Centre for Wireless Communications (CWC), University of Oulu, Finland.

First and foremost, I am extremely grateful to the Government of China and the China Scholarship Council (CSC) for financial support during my doctoral life. Meanwhile, I am also indebted to JAIST for exempting my entrance fee and tuition fee of doctoral program.

I would like to express my strong love to this university, JAIST. Without the knowledge learned here, so many ideas of research work cannot come out. The most conspicuous advantage of JAIST is its excellent facilities, such as the supercomputer which basically allows common students to run a program up to 144 threads. I benefit a lot from the supercomputer when evaluating the system performance through simulations. It helps me obtain results faster, so that I can have enough time to address reviewers' comments and extra time to do new researches.

On my way to pursuing the Ph.D. degree, the person who provided most instructions and help to me is my supervisor, Professor Tad Matsumoto. During my most arduous time when I was a research student, he not only gave me the advice on researches and daily life, but also recommended me to be a research assistant for financial support. Therefore, I would like to express my sincere sense of gratitude to him.

Moreover, I also received much guidance from my second supervisor, Associate Professor Brian Michael Kurkoski, internal advisor for minor research, Professor Mineo Kaneko, and external advisor for minor research, Professor Markku Juntti from University of Oulu. I would like to express my thanks to them for their valuable advice on my research.

I am also grateful to the external reviewers and examiners, Professor Pierluigi Salvo Rossi from Norwegian University of Science and Technology, Professor Gerhard Kramer from Technical University of Munich, and Postdoctoral Researcher Germán Bassi from KTH Royal Institute of Technology. They gave me a lot of significant comments to refine this dissertation.

Every time I recall the starting of my relations with JAIST, my mind goes back to the autumn in 2014, when Professor Jianwu Dang introduced JAIST to me for the first time. I would like to specially thank him also for recommending me to Professor Tad Matsumoto.

Since I started researches and taking courses in JAIST, there were many friends providing me help and support. I am grateful to Associate Professor Xin He from Anhui Normal University, Dr. Meng Cheng, Dr. Shen Qian, Dr. Ken Qin and other friends in JAIST. When I went to University of Oulu for my minor research, I obtained a lot of help from Dr. Jiguang He. Thus, I also would like to express my gratitude to him. Besides the academic life, my daily life is also enriched by the Chinese Student and Scholar Association (CSSA) of JAIST, which often organizes interesting activities. I would like to thank the committee members in the CSSA of JAIST, including Qisen Wang, Jingyu Guo, Zhiyong Qiu, Zhongsheng Wang, Zi Wang, Minghong Fang, Dazhao Xie, Yuhao Xiao, Hanyang Ge, Youming Fan, Pujun Chang, Zhizhou He, and Zhaolun Huang.

I will never forget the happiness when I received my first acceptance notification of journal paper. Before that, my other papers were rejected for six times in total. Every time the editors and reviewers pointed out my mistakes, I grew up and could finally polish my paper to have a superior quality. Therefore, I would like to thank all the editors and reviewers whether they criticize or appreciate my paper.

During my spare time, anime and music help me escape from heavy works and nervous life.

Therefore, I want to thank Yoshino Nanjō, Kana Hanazawa, Ayane Sakura, Aoi Yūki, Maaya Sakamoto, Maaya Uchida, Inori Minase, Yui Ogura, Rina Hidaka, Ayana Taketatsu, Rumi Ōkubo, Ai Kayano, Kana Asumi, Saori Ōnishi, Sakura Tange, Rie Takahashi, Miyuki Sawashiro, Nao Tōyama, Shiori Izawa, Yui Horie, Yukari Tamura, Saori Hayami, Hiroshi Kamiya, Natsuki Hanae, Yoshitsugu Matsuoka, and other voice actors/actresses for their supreme performing arts.

Finally, I owe my family, especially my father Lixing Lin, my mother Qiongxian Hu, my uncles Liming Lin and Qiaomao Hu, a great debt of gratitude for their encouragement when I got stuck in research.

List of Abbreviations

ACC	Accumulator
ADF	Adaptive decode-and-forward
AF	Amplify-and-forward
AWGN	Additive white Gaussian noise
BCJR	Bahl-Cocke-Jelinek-Raviv
BER	Bit error rate
BF	Bit flipping
BPSK	Binary phase shift keying
BSC	Binary symmetric channel
CC	Convolutional code
CEO	Chief executive officer
CF	Compress-and-forward
CSI	Channel state information
DF	Decode-and-forward
DMS	Discrete memoryless source
DPF	Distortion propagating function
DSBS	Doubly symmetric binary source
FER	Frame error rate
HMV	Hybrid majority voting
i.i.d.	Independent and identically distributed
IoT	Internet of Things
JPEG	Joint photographic experts group
LDPC	Low-density parity-check
LF	Lossy-forwarding
LLR	Log-likelihood ratio
MAC	Medium access control
MARC	Multiple access relay channel
MDC	Multiple description coding
MP3	MPEG audio layer 3
MPEG	Moving picture experts group
MV	Majority voting
PCM	Pulse-code modulation
pdf	Probability density function
pmf	Probability mass function
QoS	Quality of Service
R-D	Relay-destination
S-D	Source-destination
S-R	Source-relay
SMV	Single-compression with MV code
SNR	Signal-to-noise ratio
WSN	Wireless sensor network

List of Symbols

ACC^{-1}	decoder of ACC
$\mathcal{B}$	equal-size bin used in joint typicality coding
$\mathbf{b}$	a vector
$\mathbf{b}^{\text{T}}$	the transpose of the vector $\mathbf{b}$
$\|\mathbf{b}\|_2$	l_2-norm of $\mathbf{b}$
$\text{Bern}(\rho)$	the Bernoulli distribution which takes the value 1 with probability ρ
$C(\cdot)$	the Shannon capacity using Gaussian codebook
CC^{-1}	decoder of CC
D	distortion requirement
d	distortion measure
$\tilde{d}$	a dummy variable for a specific value of distortion
$\text{DSBS}(\rho)$	doubly symmetric binary source with crossover probability ρ
E	transmitting symbol energy
$\mathcal{E}$	an event
$\text{E}(\cdot)$	the expected value
$\exp(\cdot)$	natural exponential function
$f(\cdot)$	the *pdf* of a random variable
$f_b(\cdot)$	joint binary entropy function for correlated sources
$f_c(\cdot)$	LLR updating function for correlated sources
$\text{F}_{\text{DP}}(\cdot)$	distortion propagating function
G	geometric gain
$g(\cdot)$	the algorithm for generating the helper information
h	complex channel gain
$\|h\|$	the modulus of h
$H(\cdot)$	the entropy of a random variable
$H_b(\cdot)$	binary entropy function
$H_b^{-1}(\cdot)$	the inverse of binary entropy function
i	link index
$I(X;Y)$	the mutual information between X and Y
L	the number of multiple links
$\mathcal{L}$	a set of $\{1, 2, \cdots, L\}$
LLR^a	*a priori* LLR
LLR^e	extrinsic LLR
LLR^p	*a posteriori* LLR
$\log(\cdot)$	logarithmic function
M	the random variable representing the encoded codeword
$p(\cdot)$	the *pmf* of a random variable
P_{out}	outage probability
$\text{PB}(\cdot)$	Poisson binomial distortion function
$\text{pow}(a, b)$	a to the power of b, i.e., a^b
$\Pr\{\cdot\}$	the probability of an event
Q	the auxiliary variable resulting from time-sharing scheme

R	link rate
$[R]^-$	$\min\{1, R\}$
$\mathcal{R}(\cdot)$	achievable rate-distortion region
r	channel coding rate
$\mathcal{S}$	a subset of $\mathcal{L}$
$\mathcal{S}^c$	the complementary set of $\mathcal{S}$
S_j	the j-th element of the set $\mathcal{S}$
S_j^k	a set of $\{S_j, S_{j+1}, \cdots, S_{k-1}, S_k\}$
t	time index
$\mathcal{T}_\epsilon^{(n)}$	the set of jointly ϵ-typical n-sequences
U	the auxiliary variable representing the compressed information of X
V	the auxiliary variable representing the compressed information of Y
X	the random variable denoting the source
$X_{\mathcal{L}}$	a set of $\{X_i \| i \in \mathcal{L}\}$
X^n	the source sequence with n bits
$\mathcal{X}$	the finite alphabet of the random variable X
$\|\mathcal{X}\|$	the cardinality of $\mathcal{X}$
x	the realization of the random variable X
$\hat{x}$	the estimate of x
Y	the random variable denoting the side information
Z	the random variable denoting noise
$\delta(\epsilon)$	a function of ϵ that tends to zero as $\epsilon \to 0$
ϵ	an arbitrary small positive number
γ	instantaneous SNR of the channel
$\overline{\gamma}$	average SNR of the channel
φ	the mapping for encoding
ψ	the mapping for decoding
Π	interleaver
Π^{-1}	deinterleaver
ρ	the crossover probability between two random variables
$\Theta(\cdot)$	the constraint on source coding rate
$a * b$	binary convolution process, i.e., $a * b = a(1-b) + b(1-a)$
$\binom{n}{k}$	binomial coefficient, i.e., n choose k
$X \oplus Y$	modulo-2 sum

Contents

CHAPTER 1

Introduction

Nowadays, Internet of Things (IoT) becomes the technical basis of smart society [1], where numerous sensors and/or robots collect data and monitor objects instead of human. In general, the facilities communicate with each other through wireless channels for mobility and extendibility, and therefore wireless sensor networks (WSNs) are widely implemented to support IoT [2–5]. Essentially, the fundamental framework of WSNs and IoT is multiterminal source coding, in which the correlated sources are separately encoded in distributed encoders, while the received codewords are jointly decoded in a common decoder.

Traditionally, lossless recovery of the information is needed in various communications scenarios which require high fidelity and reliability. There are already some research achievements related to lossless communications in WSNs. Zou *et al.* [6] proposed a data coding and transmission method, which can losslessly recover the original data despite the data loss occurred during transmissions, for structural health monitoring by wireless smart sensor network. In [7], Long and Xiang developed a lossless data compression algorithm based on run-length encoding and Huffman coding for energy saving in WSNs. Dedeoglu *et al.* [8] presented a distributed optimization algorithm for power allocation in lossless data gathering WSNs.

However, in IoT systems, the major task is to make some judgements other than losslessly reconstruct the source information. Thus, the system is still able to make correct judgements, as long as the distortions of the source estimates are within a specified degree. Especially in big data era, large quantities of data packets transmitted through networks result in the significant power consumption and the bandwidth shortage. If the estimates of the source information are not necessarily lossless, as exemplified in IoT systems, we can save power and bandwidth by reducing the transmission rates. Consequently, there is an interesting trade-off between link rates and final distortions in lossy multiterminal source coding.

To date, the concept of helper has been introduced into diverse communication systems to make transmissions more robust and reliable [9–12]. Inspired by these research works, we are interested in the performance improvement by introducing helper(s) into the communication system. Obviously, it can easily be expected that the system can satisfy lower distortion requirements by adding helper(s). Nevertheless, there might be some problems regarding the resource efficiency, e.g., how much performance gain we can obtain from the helper, or whether the performance gain can increase linearly by adding more helpers. To answer these questions, we have to specifically calculate the performance gains. Therefore, this dissertation aims at the performance analysis and practical coding design for helper-assisted lossy multiterminal source coding.

1.1 Preliminaries

1.1.1 Multiterminal Source Coding for Direct Transmissions

The general model of multiterminal source coding for direct transmissions is depicted in Fig. 1.1, where two correlated source sequences X_1^n and X_2^n are separately encoded into two codewords M_1

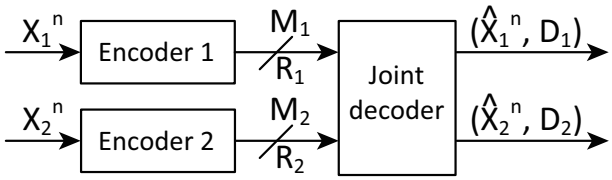


Fig. 1.1. The general model of multiterminal source coding for direct transmissions.

and M_2 to satisfy the link rates R_1 and R_2, respectively. Then, a joint decoder utilizes both M_1 and M_2 to construct the estimates $\hat{X}_1^n$ and $\hat{X}_2^n$, which may deviate from the source sequences X_1^n and X_2^n within the distortion requirements D_1 and D_2.

For the system shown in Fig. 1.1 with $D_1 = D_2 = 0$, i.e., lossless multiterminal source coding, Slepian and Wolf [13] determined the achievable rate region with two discrete memoryless source (DMS) for the first time. Surprisingly, even though the distributed encoders do not communicate with each other, the achievable rate region is still the same as that of joint encoding. Then, Cover [14] generalized the Slepian-Wolf theorem to the case with arbitrary number of sources. However, the exact achievable rate-distortion region is still an open question for the system without necessary requirements of the full source recoveries. The most classical results of lossy multiterminal source coding problem are the inner and outer bounds on the achievable rate-distortion region derived by Berger [15] and Tung [16].

Regarding the system with Gaussian sources, Oohama [17] devoted efforts to the inner and outer bounds on the rate-distortion region for Gaussian multiterminal source coding under squared distortion measures. Subsequently, Wagner *et al.* [18] determined the rate-distortion region of the quadratic Gaussian source coding problem with two sources, and provided the proofs of achievability and the converse.

1.1.2 The CEO Problem

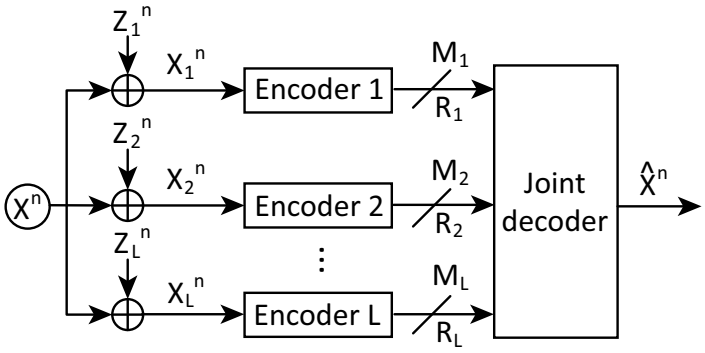


Fig. 1.2. The CEO problem.

Fig. 1.2 illustrates another interesting problem in the category of multiterminal source coding, i.e., the chief executive officer (CEO) problem [19], where a CEO (joint decoder) is interested in a hidden source X. However, due to severe communication environment in real world, such as long distance and shadow, direct transmission from the source to the CEO is not available. Therefore, the CEO has to only rely on some agents (encoders) which can observe the source X, although the observations $X_1^n, X_2^n, \cdots, X_L^n$ may also suffer from noises $Z_1^n, Z_2^n, \cdots, Z_L^n$. We spontaneously

want to know how much fidelity $\hat{X}^n$ can achieve if the strength of noise and the link rates are specified.

Chen and Berger focused on a CEO system with two agents in [20], where they developed a robust distributed coding scheme and proved the optimality in various special cases. In [21], Oohama presented classical results of the rate-distortion function for the CEO problem with Gaussian sources and squared distortion measure. The CEO problem with binary sources was solved by He *et al.* [22], who presented a lower bound of Hamming distortion for the binary CEO problem with two sources. Then, the result was further extended to solve the binary CEO problem with arbitrary number of sources in [23].

1.1.3 Multiterminal Source Coding with Side Information

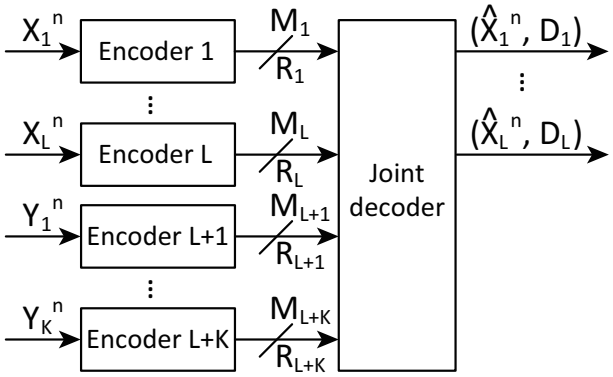


Fig. 1.3. The general model of multiterminal source coding with side information.

Actually, not all of the sequences sent from encoders need to be reconstructed in practical communication systems, where some of the transmitters only act as helpers to provide compressed side information as illustrated in Fig. 1.3. For some case that there is no rate constraint on the helper link, the joint decoder can directly receive the side information without compression.

There are already a lot of research achievements with respect to multiterminal source coding with only one source to be recovered. In [24], Ahlswede and Korner determined the rate region of the lossless source coding problem with a helper. For lossy communication systems, Wyner and Ziv [25] characterized the rate-distortion function of lossy compression with noncausal side information only available at the decoder. Sechelea *et al.* [26] analyzed the lossy compression of a binary source with correlated side information available at both encoder and decoder in depth. In [27], Rahman and Wagner showed interest in using a helper to provide side information for the problem of vector Gaussian source coding, and they identified the corresponding achievable rate region. Sgarro [28] characterized the achievable rate region for the system where one source needs to be recovered at different joint decoders with different side information. Timo *et al.* [29] derived an upper bound on the rate-distortion function for lossy source coding with various side information utilized in many decoders.

Another special case is that the helper can also directly observe the source, and hence the helper can provide side information more efficiently. This concept is referred to as successive refinement [30], which is widely implemented to satisfying different Quality of Service (QoS) with diverse users, especially for streaming media. Fig. 1.4 depicts the simplest system mode of successive refinement. A sequence X^n is encoded into two codewords M_1 and M_2 at rates R_1 and

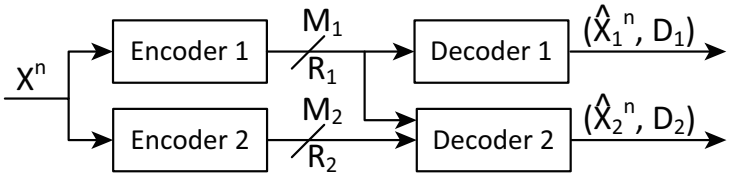


Fig. 1.4. The general model of successive refinement.

R_2, respectively. The first link is a standalone link, i.e., the decoder 1 generates a lossy recovery $\hat{X}_1^n$ to satisfy a distortion requirement D_1 only by exploiting M_1. In contrast, the second link is a refinement link where the decoder 2 can jointly utilize M_1 and M_2 to reconstruct $\hat{X}_2^n$ with a lower distortion D_2. The coding technique for successive refinement has a more generic name, i.e., multiple description coding (MDC) [31]. Wolf *et al.* [32] characterized a necessary condition on the achievable rate-distortion region of MDC for the first time, and the necessary condition was further investigated for a binary source by Witsenhausen and Wyner in [33]. Then, El Gamal and Cover [34] derived an inner bound on the achievable rate-distortion region for MDC.

If the system contains only one link of source to be reconstructed and more than one link of helpers, it is classified into many-help-one problem. For the many-help-one problem with Gaussian sources, Oohama [35] obtained the rate-distortion region in the case that the helper information are conditionally independent if the target source is given. Wolf *et al.* [36] proposed an inner bound on the rate region of binary many-help-one problem, in which the source has to be recovered losslessly.

For the case with more than one source to be recovered, Han and Kobayashi studied a multiterminal source coding problem for losslessly reconstructing many sources with many helpers in [37], where an inner bound is derived by utilizing a coding scheme based on the joint typical sequence. In [38], Wyner determined the rate region for the lossless problem with two source links and one helper link, under the condition that each source link can only separately utilize its own data and the helper data. Rey Vega *et al.* in [39] focused on a lossy source coding problem for three terminals, containing both an encoder and decoder interactively performing encoding and decoding.

1.1.4 Lossy-Forwarding

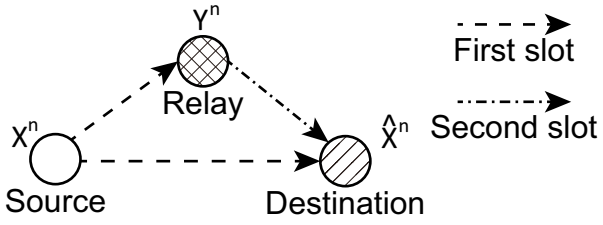


Fig. 1.5. The simplest system model of a lossy relaying system.

Relaying system is an implementation of multiterminal source codding. Recently, lossy-forwarding (LF) [40] has attracted significant attention of academia and industry, owing to its great potential in harsh communication environment. As shown in Fig. 1.5, a source broadcasts the sequence X^n to a destination and a relay at the first time slot. Then, the destination aims to

recover the source sequence after receiving the assisted information from the relay at the second time slot. If the capacity constraint on the source-relay (S-R) link is relatively strict, the relay cannot forward the message correctly. Once errors are detected in the decoded data sequence, the traditional decode-and-forward (DF) scheme discards the data sequence without forwarding to the destination. However, from the viewpoint of multiterminal source coding, the relay sequence containing intra-link errors has correlation with the source sequence as well. By the LF strategy, the relay still continues to send the error-corrupted sequence Y^n to the destination, and hence the final estimate can be refined with the side information provided from the relay despite the link rate of relay channel.

So far, a number of scholars have made efforts to investigate LF. Base on the Slepian-Wolf theorem, Hu and Li [41] proposed the novel LF relaying strategy for the first time, to help the destination recover data losslessly. In [42], Cheng *et al.* derived the outage probability for an LF relaying system with three nodes communicating through block Rayleigh fading channels. Zhou *et al.* [43] evaluated the exact outage probability over independent block Rayleigh fading channels for LF relaying system. As for the practical techniques related to LF, researchers in [44–46] provided diverse coding schemes based on the turbo code [47]. Brulatout *et al.* [48] presented a medium access control (MAC) layer protocol which cooperates with LF techniques in physical layer. In [49], Wolf *et al.* designed an optimal power allocation scheme among a source and two LF relays by taking into account outage probability.

1.1.5 Rayleigh Fading Channel

Rayleigh fading channel is a widely implemented channel model to reflect the effect of radio signal propagation. The complex channel gain h of a Rayleigh fading channel follows the two-dimensional Gaussian distribution. For a modulated symbol $x(t)$ sent at the t-th time index, the received signal is expressed as

$$x'(t) = h\sqrt{G} \cdot x(t) + z(t), \tag{1.1}$$

where z represents the zero-mean additive white Gaussian noise (AWGN), and G is the geometric gain due to transmission distance. Let $E = \mathrm{E}[|x(t)|^2]$ be the transmitting symbol energy, and the variance of all z be equal to $N_0/2$ per dimension. The average signal-to-noise ratio (SNR) is given by

$$\overline{\gamma} = G \cdot \mathrm{E}[|h|^2] \cdot \frac{E}{N_0}. \tag{1.2}$$

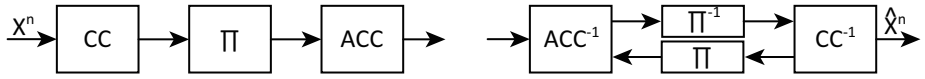
Then, the instantaneous SNR can be calculated by

$$\gamma = |h|^2 \cdot \overline{\gamma}. \tag{1.3}$$

We can finally obtain the probability density function (*pdf*) of instantaneous SNR as

$$f(\gamma) = \frac{1}{\overline{\gamma}} \exp(-\frac{\gamma}{\overline{\gamma}}). \tag{1.4}$$

1.1.6 Basic Channel Coding Scheme Throughout the Dissertation



(a) Encoder. (b) Decoder.

Fig. 1.6. The basic channel coding scheme throughout the dissertation.

Without explicit specification, we implement the coding scheme illustrated in Fig. 1.6 as the basic channel coding scheme throughout this dissertation.

To start with, the sequence X^n is encoded with a convolutional code (CC) as the outer code. Then, the output of CC is interleaved by Π for the purpose of exploiting the principle of turbo codes. Finally, an accumulator (ACC) [46] encodes the interleaved sequence as the inner code.

The structure of corresponding decoder is depicted in Fig. 1.6(b), where ACC^{-1} and CC^{-1} represent the decoder of ACC and CC, respectively. In decoding process, ACC^{-1} decodes the received symbols and output log-likelihood ratio (LLR) for the first step. After deinterleaving by Π^{-1}, the LLR of outer code is then decoded by CC^{-1}. Moreover, CC^{-1} also yields the extrinsic LLR to be utilized as the *a priori* information for refining the decoding result of ACC^{-1}. By several rounds of iterative decoding described above, we can further eliminate the negative impact of the low channel SNR.

1.2 Motivation

1.2.1 The State-of-the-Art

1.2.1.1 Theoretical Framework

Currently, the theoretical framework of multiterminal source coding is highly matured. Numerous researchers unified and generalized the classical theorems, such as the Slepian-Wolf theorem, the Berger-Tung bounds, and the Wyner-Ziv theorem. For instance, Wagner and Anantharam [50] studied a multiterminal source coding problem with one link of uncompressed side information available. In [51], Jana and Blahut derived the bounds for lossless and lossy multiterminal source coding systems where lossless and lossy links are mixed.

Regarding the CEO problem, the case with binary sources was solved in [22, 23] as mentioned above. When solving the binary CEO problem, He *et al.* [22] divided the communication into a successive encoding/decoding process, i.e., encoding/decoding the multiple sources and then combining the joint decoding results. In the first step of multiterminal source coding, they derived an outer bound of the rate-distortion region for binary sources. Then, the outer bound was extended to the case with arbitrary number of binary sources in [23]. For the final decision of binary CEO problem, the bit error probability of binary data gathering by soft combining was analyzed in [52], where many correlated sources have diverse bit-flipping probabilities. In terms of decoding algorithms for binary CEO problem, Razi and Abedi [53] developed a method to analyze the convergence of iterative decoding for binary CEO problem. An iterative joint decoding algorithm was implemented into the WSNs with binary sources according to the model of binary CEO problem by Haghighat *et al.* in [54]. He *et al.* developed a joint decoding algorithm for binary

CEO problem in [55], where the joint decoder recursively performs soft decoding and updates LLR by exchanging the mutual information among the data sequences.

Based on the classical rate region of multiterminal source, many scholars also investigated the outage probability for communications suffering from channel fadings. Laneman *et al.* characterized the outage probabilities of amplify-and-forward (AF) and DF relaying strategies for Rayleigh fading channels in [56], where the relay and source messages can be regarded as correlated information. Zhou *et al.* [57] derived the outage probability for the system with two correlated binary sources communicating through orthogonal multiple access relay channel (MARC) over block Rayleigh fadings. In [58], Lu *et al.* analyzed the outage probability of the MARC system where two correlated sources suffer from block Rayleigh fadings, and the estimate of source in the relay may contain errors. The popular LF relaying strategy has also been analyzed in depth with respect to the outage probability. Qian *et al.* [59] made a comparison of outage probability under spatially and temporally correlated fading among LF, DF and adaptive decode-and-forward (ADF). In [60], Qian *et al.* analyzed the theoretical performance of an LF system with three nodes suffering from independent block Nakagami-m fading.

1.2.1.2 Practical Coding Techniques

In coding theory, data compression is a classical topic including two fundamental categories, i.e., lossless and lossy. The lossless compression has been well studied during the last several decades, e.g., Shannon coding [61], Huffman coding [62] and Lempel-Ziv coding [63, 64]. Regarding lossy compression for continuous source and multimedia data, there are many technologies, such as pulse-code modulation (PCM) for continuous source, MPEG audio layer 3 (MP3) for audio [65], joint photographic experts group (JPEG) [66] for image, and MPEG-4 [67] for video. Even though the multimedia data is in a digital format, they are basically continuous sources with correlations between information bits.

Nevertheless, the lossy compression for DMS is not easy, because the distance between the codewords and the original sequences is considered to be more crucial than the correlations between symbols. Although the optimal performance can be achieved for sufficiently long sequence according to Shannon's lossy source coding theorem [68], we need significantly huge memory to store the codebook for joint typicality coding. For lossy multiterminal source coding, although the coding schemes in the Wyner-Ziv theorem and the Berger-Tung inner bound have superior performance regarding coding rates versus distortions, they are too complex to be implemented into practical systems. Not only is it difficult to find a theoretically optimal codebook with respect to a specified distortion, but we also have to find diverse codebooks for different rate or distortion requirements.

Moreover, another problem for practical lossy source coding is how to refine the estimate of source if extra information is available. For continuous sources, there are a number of researches which focus on practical MDC algorithms, e.g., image [69], audio [70] and video [71]. However, up to our best knowledge, no research aims at establishing the practical coding scheme for successive refinement with DMS, although the theoretical investigations have reached already highly matured level.

1.2.2 Beyond the State-of-the-Art

Despite a large number of theoretical achievements in multiterminal source coding, the outage probability is still unknown for lossy end-to-end multiterminal communications with a helper.

Obviously, the distortions of recovered observations are determined by the link rates, which are derived from the instantaneous channel capacities. According to Shannon's lossy source-channel separation theorem [61, 68], the distortion occurring in block fading channels can be equivalently evaluated for the case, where the information sequence is compressed into a codeword with lower rate by lossy source coding such that the codeword can be losslessly transmitted through fading channels. In order to conduct the outage probability analysis, it is necessary to determine the achievable rate-distortion region for multiterminal source coding with a helper. Then, based on Shannon's lossy source-channel separation theorem for multiterminal communications [72], the results of rate-distortion analysis can be further utilized in the derivation of outage probability in block fading channels.

With regard to the binary CEO problem, we are interested in the performance improvement provided by a helper. Based on the previous achievements, we establish the framework of the binary CEO problem with a helper as a successive process with two steps, i.e., multiterminal source coding with a helper and final decision.

For more than one helper, the achievable rate-distortion region has not been determined yet for lossy many-help-one problem. We make our contribution to deriving an inner bound on the achievable rate-distortion region for lossy source coding with multiple helpers.

Subsequently, we consider the implementation of multiterminal source coding in relaying systems. As stated above, there are already a lot of works related to outage probability analysis of lossless relaying systems. Nonetheless, the performance analysis has not been finished yet for the LF relaying systems with lossy reconstructions allowed at the destination, which is concisely named as lossy LF relaying.

Finally, notice that it is hard to implement the joint typicality coding scheme used in theoretical analysis to practical systems. Thus, we develop a practical lossy source coding algorithm so-called hybrid majority voting (HMV) code, which requires relatively low complexity and exhibits good performance. In addition, we further apply the HMV code to successive refinement, by finding a complement coding scheme that contains the information of lost part caused in lossy source coding.

1.3 Notations and Definitions

For the purpose of conciseness in derivations and distinction between similar terminologies, this section introduces the common definitions used throughout this dissertation.

1.3.1 Random Variables and Sets

The random variables and their realizations are denoted by uppercase and lowercase letters, respectively. In particular, we use i to denote the link index and t to denote the time index. Generally, X, Y and M stand for source information, helper information, and encoded codeword, respectively. U and V represent the compressed information of X and Y, respectively. Calligraphic letters $\mathcal{X}$, $\mathcal{Y}$, $\cdots$ denote the finite alphabets of a random variable. The superscript of a random vector and its realization represent the length of the vector.

In particular, we define $\mathcal{L} = \{1, 2, \cdots, L\}$, and $\mathcal{S}$ is a subset of $\mathcal{L}$. Furthermore, $\mathcal{S}^c$ represents the complementary set of $\mathcal{S}$. We define S_j as the j-th element of the set $\mathcal{S}$, and $S_j^k = \{S_j, S_{j+1}, \cdots, S_{k-1}, S_k\}$. The random variable with a finite alphabet as subscript stands

for a set of all random variables with index in the finite alphabet, such as $X_{\mathcal{L}} = \{X_i | i \in \mathcal{L}\}$. The cardinality of a set is denoted by $|\cdot|$.

1.3.2 Functions and Operations

For a function $F(\cdot)$, $F^{-1}(\cdot)$ stands for the corresponding inverse function. The common functions used throughout this dissertation are defined in the following.

Definition 1.1: The entropy of a random variable X with probability mass function (*pmf*) $p(x)$ is defined as

$$H(X) = -\sum_{x \in \mathcal{X}} p(x) \log p(x). \tag{1.5}$$

In particular, $H_b(\cdot)$ denotes the binary entropy function.

Definition 1.2: The mutual information between two random variables X and Y is defined as

$$I(X;Y) = \sum_{(x,y) \in \mathcal{X} \times \mathcal{Y}} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}. \tag{1.6}$$

Definition 1.3: *Joint binary entropy function for correlated sources.* According to [23], given a set of crossover probabilities $\{\mathcal{P}\}$ with a common binary source $X \sim \mathrm{Bern}(0.5)$, the joint entropy $f_b(\cdot)$ of the outputs from independent binary symmetric channels (BSCs) is calculated as

$$f_b(\{\mathcal{P}\}) = -\sum_{j=1}^{2^{|\mathcal{P}|}} q_j \log_2(q_j), \tag{1.7}$$

where

$$q_j = 0.5 \left(\prod_{k \in \mathcal{A}_j} p_k \prod_{k' \in \mathcal{A}_j^c} \bar{p}_{k'} + \prod_{k \in \mathcal{A}_j} \bar{p}_k \prod_{k' \in \mathcal{A}_j^c} p_{k'} \right), \tag{1.8}$$

with $\bar{p} = 1 - p$ and $\mathcal{A}_j$ traversing all the subsets of $\{1, 2, \cdots, |\mathcal{P}|\}$.

In addition, we define the following functions and operations for the convenience in derivation. We define $\mathrm{pow}(a,b) = a^b$, and $[R]^- = \min\{1, R\}$. The operation $*$ denotes the binary convolution process, i.e., $a * b = a(1-b) + b(1-a)$.

1.3.3 Distortion Measure

The distortion measure $d : \mathcal{X} \times \mathcal{X} \mapsto [0, \infty)$ is defined to describe the distortion level between $x(t)$ and $\hat{x}(t)$ at t-th time index. Particularly, if the source is binary, the distortion level is described by the Hamming distortion measure as

$$d(x(t), \hat{x}(t)) = \begin{cases} 1, & \text{if } x(t) \neq \hat{x}(t), \\ 0, & \text{if } x(t) = \hat{x}(t). \end{cases} \tag{1.9}$$

For the entire sequence, the average distortion between x_i^n and $\hat{x}_i^n$ is defined as

$$d(x^n, \hat{x}^n) = \frac{1}{n} \sum_{t=1}^{n} d\left(x(t), \hat{x}(t)\right). \tag{1.10}$$

1.3.4 Clarification of Terminologies

Table 1.1. THE ATTRIBUTES OF SIMILAR TERMINOLOGIES

Terminology	The way of collecting information	Information needed to be recovered	Able to generate its own information
Sensor	actively detect	yes	no
Agent	actively detect	not necessarily	no
Helper	actively detect or passively receive	no	yes
Relay	passively receive	no	yes

To avoid confusion, Table 1.1 summarizes the major attributes of the similar terminologies used in this dissertation. Hence, we can distinguish similar terminologies by their different attributes.

1.4 Outline of the Dissertation

The objective of this dissertation is to present the theoretical analyses and coding design for several typical problems of helper-assisted lossy multiterminal source coding.

Chapter 2 starts from the simplest case of helper-assisted lossy multiterminal source coding, i.e., only two sources and one helper in the system. To investigate the performance improvement provided by a helper, initially, we determine an inner bound on the achievable rate-distortion region for binary sources. Then, we evaluate the system performance through the achievable rate-distortion region with diverse correlation levels of sources and distortion requirements. Based on Shannon's lossy source-channel separation theorem, we further derive the upper bound of the outage probability of the system over block Rayleigh fading channels. The diversity gain provided a helper is verified in the numerical results.

Chapter 3 focuses on the case without direct transmission link from the source to the destination, i.e., the binary CEO problem with a helper. To begin with, we use a successive decoding scheme to decompose the binary CEO problem with a helper into the multiterminal source coding and final decision problems. Then, we present an outer bound on the rate-distortion region for multiterminal source coding with binary sources and a helper. After solving a convex optimization problem formulated from the derived outer bound, we obtain the final distortion by substituting the minimized distortions of observation into the distortion propagating function (DPF), which is derived to bridge the relationship between the joint decoding results and final decision. Finally, we analyze the trade-off of rate-distortion through theoretical calculation and simulations. We also have an in-depth discussion on the differences of system performance improvement between locating a helper and including an additional agent.

Chapter 4 analyzes the performance gain by adding more than one helper for lossy source coding with one target source. First of all, we perform the theoretical analysis to derive an inner bound on the achievable rate-distortion region for lossy source coding with helpers. The numerical results precisely match the Wyner-Ziv theorem when there is only one helper link and no rate constraint on the helper link. Moreover, a series of simulations are conducted for the performance

evaluation provided that the link rates are constrained by channel capacities. Although there is an obvious gap between the theoretical and simulation results, the performance curves show similar tendencies in terms of the SNR versus bit error rate (BER).

Chapter 5 establishes the theoretical framework towards the utilization of helper in practical systems, i.e., lossy communications with the aid of LF relaying. For in-depth performance analysis, the problem is decomposed into two parts as follows: a point-to-point coding problem in the S-R link, and a lossy source coding problem with a helper in the source-destination (S-D) and relay-destination (R-D) links. To begin with, we derive an inner bound on the achievable rate-distortion region of the lossy source coding with a helper. Then, we focus on the analysis of outage probability over block Rayleigh fading channels. Finally, a practical encoding/decoding scheme is proposed for the evaluation of system performance by computer simulations. Due to the suboptimal channel coding and incomplete utilization of joint typicality, the theoretical performance cannot be achieved in the simulation; however, the tendency of curves in simulations matches that in theoretical calculation.

Chapter 6 performs practical coding design for lossy compression and successive refinement with DMS. Inspired by the coding scheme used in the classic rate-distortion theorem, we find a series of basic majority voting (MV) codes and analyze their rate-distortion performance. We then present an algorithm to find two component MV codes and apply them to lossy compression, group by group, to construct the HMV codes. Moreover, we further implement the HMV code to successive refinement by the means of developing the bit flipping (BF) code as the complement code. We also evaluate the performance of the HMV code through simulations, the results of which indicate that the HMV code makes it possible to easily control efficiency and complexity. By utilizing the HMV code and the BF code for successive refinement, the standalone link can satisfy lower distortion than puncturing; meanwhile, the refinement link has almost the same performance as puncturing for relatively large R_1. Based on the duality between source coding and channel coding, we propose the methodology for hybrid codes design, where the Hamming codes are utilized as an example for designing hybrid codes. We also find that the syndrome of the Hamming codes can be further used as the complement code for successive refinement.

Chapter 7 summarizes the main results and concludes this dissertation. We also present the perspective of helper-assisted lossy multiterminal source coding and several directions of future studies.

1.5 Summary of Contributions

This dissertation is written as a monograph based on four journal papers [73–76] and one conference paper [77]. The author has taken the main responsibility for deriving theoretical equations, designing practical coding schemes, developing simulation programs, and writing all the papers. The co-authors provided guidance, helps, comments and criticism during the research and writing processing.

The main contributions of this dissertation are summarized as follows:

- We derive an inner bound on the achievable rate-distortion region of lossy multiterminal source coding problem with two binary sources and a helper. Base on the derived inner bound and Shannon's lossy source-channel separation theorem, we further calculate the upper bound of the outage probabilities over block Rayleigh fading channels. By utilizing the derived mathematical results, we conduct an in-depth investigation of performance

improvement by introducing a helper. It is remarkable that a helper can not only enlarge the achievable rate-distortion region, but also provide diversity gains and reduce the outage probability.

- For the binary CEO problem with a helper, we derive an outer bound on the rate-distortion region of multiterminal source coding problem with many agents and a helper for binary sources. Then, the outer bound is utilized to formulate a convex optimization problem for minimizing the distortions when reconstructing observations. Moreover, we analyze the distortion propagating from the estimate of agent sequences to the final decision. By substituting the solution of the convex optimization problem for minimizing the distortions in recovered observations, we investigate the trade-off of rate-distortion for the binary CEO problem with a helper. Besides, we make a comparison of performance improvement between the system with a helper and that with an additional agent through simulations.
- We present an inner bound on the achievable rate-distortion region of lossy source coding with helpers for general sources. For the helper information being independent with each other if the source is given, we further calculate the rate-distortion function for doubly symmetric binary source (DSBS), and extend the results to joint source-channel coding. The theoretical results are consistent to the Wyner-Ziv theorem as the special case in the sense that there is only one full-rate helper in the system.
- For lossy communications with an LF relay, we determine an inner bound on the achievable rate-distortion region. Based on the derived inner bound on the achievable rate-distortion region, we investigate the upper bound of the outage probability assuming block Rayleigh fading channels in the relaying system; knowing the upper bound of the outage probability allows the system designers to build the communications systems based on the safer side of specification. The numerical results demonstrate the relationship of outage probability to average SNR, expected distortion and relay location. Moreover, we make a comparison of the outage probability among AF, DF and LF through simulations.
- Finally, we develop a practical lossy compression scheme, i.e., the HMV code, and perform the theoretical rate-distortion analysis for it. We also find the BF code as the complement code of HMV code for successive refinement. To evaluate the performance of the HMV and BF codes, we conduct a series of simulations, where the results demonstrate the better performance of the HMV code than puncturing. Moreover, we propose the methodology of hybrid codes design for lossy source coding. The hybrid code based on the Hamming codes is presented as an example; meanwhile, we also find the corresponding complement code for successive refinement by calculating the syndrome of the Hamming codes.

CHAPTER 2

Multiterminal Source Coding with a Helper

This chapter starts the performance analysis of helper-assisted lossy multiterminal source coding from the simplest case, i.e., only two sources and one helper in the system. The main goal is to analyze the rate-distortion performance of multiterminal source coding with a helper for binary sources and the outage probability over block Rayleigh fading channels. Notice that the exact achievable rate-distortion region for lossy multiterminal source coding is still an open problem. Thus, for the achievable rate-distortion region used in the derivation of outage probability, we only consider the inner bound, i.e., the lossy recoveries must satisfy the distortion requirements if the link rates are larger than the inner bound.

2.1 Problem Statement

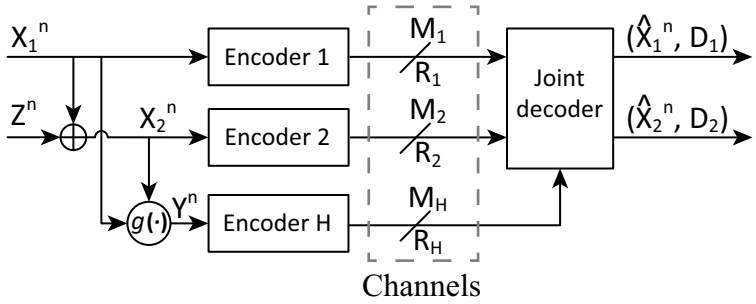


Fig. 2.1. The model of multiterminal source coding with two binary sources and one helper.

We consider the simplest case of multiterminal source coding with a helper, i.e., there are only two binary sources in the system as illustrated in Fig. 2.1. There are two independent and identically distributed (i.i.d.) sequences $x_1^n = \{x_1(t)\}_{t=1}^n$ and $x_2^n = \{x_2(t)\}_{t=1}^n$, generated by two correlated DMSs X_1 and X_2, respectively. At t-th time slot, $x_i(t)$ takes values from the binary alphabet $\mathcal{X}_i = \{0, 1\}$ for $i \in \{1, 2\}$. Thus, X_2 is equivalent to the output of a BSC with input X_1 and crossover probability ρ and vice versa, i.e., $X_2 = X_1 \oplus Z$ with $Z \sim \mathrm{Bern}(\rho)$. In this chapter, we consider the sources $X_i \sim \mathrm{Bern}(0.5)$ for $i \in \{1, 2\}$. Since the helper information $y^n = \{y(t)\}_{t=1}^n$ highly depends on the helper structure, we assume without loss of generality that Y is a function $g(\cdot)$ of X_1 and X_2.

To begin with, three sequences x_1^n, x_2^n and y^n are independently encoded by encoder 1, encoder 2 and encoder H at coding rates R_1, R_2 and R_{H}, respectively. The encoding process can be performed by assigning an index M to each sequence according to the following mapping rules:

$$\varphi_i : \mathcal{X}_i^n \mapsto \mathcal{M}_i = \{1, 2, \cdots, 2^{nR_i}\} \text{ for } i \in \{1, 2\}, \tag{2.1}$$

$$\varphi_{\mathrm{H}} : \mathcal{Y}^n \mapsto \mathcal{M}_{\mathrm{H}} = \{1, 2, \cdots, 2^{nR_{\mathrm{H}}}\}. \tag{2.2}$$

Subsequently, the encoding outputs $\varphi_1(x_1^n)$, $\varphi_2(x_2^n)$ and $\varphi_{\mathrm{H}}(y^n)$ are transmitted to a common receiver. In contrast to distributed compressions in the encoders, the decoder can jointly construct

the estimates $\hat{x}_1^n$ and $\hat{x}_2^n$ from indices $\varphi_1(x_1^n)$ and $\varphi_2(x_2^n)$ by utilizing the compressed side information $\varphi_{\mathrm{H}}(y^n)$. The reconstruction process can be implemented by the mapping as follows:

$$\psi : \mathcal{M}_1 \times \mathcal{M}_2 \times \mathcal{M}_{\mathrm{H}} \mapsto \mathcal{X}_1^n \times \mathcal{X}_2^n. \tag{2.3}$$

Since the estimate $\hat{x}_i^n$ may occasionally deviate from the observation x_i^n if the rates are not large enough, the Hamming distortion measure $d_i : \mathcal{X}_i \times \mathcal{X}_i \mapsto \{0, 1\}$ is applied to describe the distortion level between x_i and $\hat{x}_i$. For given distortion requirements (D_1, D_2), the rate-distortion region $\mathcal{R}(D_1, D_2)$, consisting of all achievable rate triplets $(R_1, R_2, R_{\mathrm{H}})$, is defined as

$$\begin{aligned}\mathcal{R}(D_1, D_2) = \big\{&(R_1, R_2, R_{\mathrm{H}}) : (R_1, R_2, R_{\mathrm{H}}) \text{ is admissible such that} \\ &\lim_{n\to\infty} \mathrm{E}(d_i(x_i^n, \hat{x}_i^n)) \le D_i + \epsilon, \text{ for } i \in \{1, 2\}, \text{ and any } \epsilon > 0\big\}.\end{aligned} \tag{2.4}$$

In the literature, Berger [15] and Tung [16] determined the inner and outer bounds on the achievable rate-distortion region for the system with two sources only. Wagner and Anantharam [50] derived an outer bound on the achievable rate-distortion region for the case with many sources and one link for uncompressed side information utilization. The main theoretical results in this chapter are an inner bound on the achievable rate-distortion region, and an upper bound of the outage probability over block Rayleigh fading channels.

2.2 Rate-Distortion Analysis

2.2.1 Achievable Rate-Distortion Region

First, from [77], the inner bound on the achievable rate-distortion region with general sources is

$$R_1 > I(X_1; U_1|U_2, V, Q), \tag{2.5}$$
$$R_2 > I(X_2; U_2|U_1, V, Q), \tag{2.6}$$
$$R_1 + R_2 > I(X_1, X_2; U_1, U_2|V, Q), \tag{2.7}$$
$$R_{\mathrm{H}} > I(Y; V), \tag{2.8}$$

for some conditional *pmf* $p(q)p(u_1|x_1, q)p(u_2|x_2, q)p(v|y)$, and functions $\hat{x}_i(u_1, u_2, v, q)$ such that $\mathrm{E}(d_i(X_i, \hat{X}_i)) \le D_i$ for $i \in \{1, 2\}$.

U_i and V are auxiliary variables containing the compressed information in M_i and M_{H} for X_i and Y, respectively; Q is an auxiliary variable resulting from time-sharing between the cases that one of the coding rates is large enough to independently satisfy the corresponding distortion requirement. Since Q is an auxiliary variable of time-sharing, we calculate the inner bound with binary sources for $|\mathcal{Q}| = 1$ for the first step. Then, we equivalently implement the time-sharing scheme by using a dummy variable.

First, consider

$$\begin{aligned}R_1 &> I(X_1; U_1|U_2, V) \\ &= H(U_1|U_2, V) - H(U_1|X_1, U_2, V) \\ &= H(U_1|U_2, V) - H(U_1|X_1, X_2, U_2, V) - I(U_1; X_2|X_1, U_2, V) \\ &= H(U_1|U_2, V) - H(U_1|X_1, X_2, U_2) - I(U_1; X_2|X_1, U_2, V) \qquad (2.9) \\ &= H(U_1|U_2, V) - H(U_1|X_1) - I(U_1; X_2|X_1, U_2, V) \qquad (2.10)\end{aligned}$$

$$
\begin{aligned}
&= H(U_1|U_2,V) - H(U_1|X_1) - H(X_2|X_1,U_2,V) + H(X_2|X_1,U_1,U_2,V) \\
&= H(U_1|U_2,V) - H(U_1|X_1) - H(X_2|X_1,U_2,V) + H(X_2|X_1,U_2,V) \qquad (2.11)\\
&= H(U_1|U_2,V) - H(U_1|X_1) \\
&= H(U_1|U_2) - I(U_1;V|U_2) - H(U_1|X_1) \\
&= H_b(D_1 * \rho * D_2) - I(U_1;V|U_2) - H_b(D_1), \qquad (2.12)
\end{aligned}
$$

where (2.9-2.11) follow since $V \to Y \to (X_1, X_2) \to U_1$, $U_2 \to X_2 \to X_1 \to U_1$ and $U_1 \to X_1 \to X_2$ form three Markov chains, respectively, with the first Markov chain resulting from the fact that Y is a function of X_1 and X_2. Symmetrically, we have

$$R_2 > H_b(D_1 * \rho * D_2) - H_b(D_2) - I(U_2;V|U_1). \qquad (2.13)$$

Then, consider

$$
\begin{aligned}
R_1 + R_2 &> I(X_1, X_2; U_1, U_2|V) \\
&= H(U_1, U_2|V) - H(U_1, U_2|X_1, X_2, V) \\
&= H(U_1, U_2|V) - H(U_1, U_2|X_1, X_2) \qquad (2.14)\\
&= H(U_1, U_2) - I(U_1, U_2; V) - H(U_1, U_2|X_1, X_2) \\
&= H(U_1) + H(U_2|U_1) - I(U_1, U_2; V) - H(U_1|X_1, X_2) - H(U_2|X_1, X_2, U_1) \\
&= H(U_1) + H(U_2|U_1) - I(U_1, U_2; V) - H(U_1|X_1) - H(U_2|X_2) \qquad (2.15)\\
&= 1 + H_b(D_1 * \rho * D_2) - I(U_1, U_2; V) - H_b(D_1) - H_b(D_2), \qquad (2.16)
\end{aligned}
$$

where (2.14) follows since $V \to Y \to (X_1, X_2) \to (U_1, U_2)$ form a Markov chain, and (2.15) follows since $X_2 \to X_1 \to U_1$ and $U_1 \to X_1 \to X_2 \to U_2$ form two Markov chains.

Notice that in (2.12), (2.13) and (2.16), it is hard to calculate $I(U_1;V|U_2)$, $I(U_2;V|U_1)$ and $I(U_1, U_2; V)$ without a specific helper structure. We consider the theoretical optimal case in which the helper sequence contains the mutual information of sources as much as possible. From the Markov chain $U_1 \to (X_1, X_2) \to Y \to V$, we have[1]

$$
\begin{aligned}
I(U_1;V|U_2) &\le I(Y;V) \\
&\le [R_{\mathrm{H}}]^-, \qquad (2.17)
\end{aligned}
$$

and hence

$$R_1 > H_b(D_1 * \rho * D_2) - H_b(D_1) - [R_{\mathrm{H}}]^-. \qquad (2.18)$$

Likewise, we have

$$R_2 > H_b(D_1 * \rho * D_2) - H_b(D_2) - [R_{\mathrm{H}}]^-. \qquad (2.19)$$

Form the Markov chain $(U_1, U_2) \to (X_1, X_2) \to Y \to V$, we have

$$
\begin{aligned}
I(U_1, U_2; V) &\le I(Y;V) \\
&\le [R_{\mathrm{H}}]^-, \qquad (2.20)
\end{aligned}
$$

and

$$R_1 + R_2 > 1 + H_b(D_1 * \rho * D_2) - H_b(D_1) - H_b(D_2) - [R_{\mathrm{H}}]^-. \qquad (2.21)$$

[1]It should be noticed that (2.17) might not always hold for *arbitrary* helper structure.

To visually present the inner bound with the time-sharing scheme, the rate-distortion region is divided into three parts, as follows:
(a) for some $0 \leq \tilde{d} \leq D_2$,

$$\begin{cases} R_1 > H_b(D_1 * \rho * \tilde{d}) - H_b(D_1) - [R_{\mathrm{H}}]^-, \\ R_2 > 1 - H_b(\tilde{d}), \end{cases} \tag{2.22}$$

(b) for some $0 \leq \tilde{d} \leq D_1$,

$$\begin{cases} R_1 > 1 - H_b(\tilde{d}), \\ R_2 > H_b(\tilde{d} * \rho * D_2) - H_b(D_2) - [R_{\mathrm{H}}]^-, \end{cases} \tag{2.23}$$

(c) common case,

$$R_1 + R_2 > 1 + H_b(D_1 * \rho * D_2) - H_b(D_1) - H_b(D_2) - [R_{\mathrm{H}}]^-, \tag{2.24}$$

where $\tilde{d}$ is a dummy variable. We calculate the rates R_1, R_2 and $R_1 + R_2$ with given D_1 and D_2, respectively. Then, we plot the rate-distortion region by combining the three parts shown above, which is equivalent to the time-sharing concept. It should be noted that the inner bound in (2.22-2.24) can be derived only if the equality of (2.17) and (2.20) holds, and hence it is not an inner bound in general.

2.2.2 Performance Evaluation

This section evaluates the system performance with respect to rate-distortion, based on the derived inner bound in (2.22-2.24). Therefore, the inner bound plotted in the following is not an inner bound in general, and it is only for the case that (2.17) and (2.20) hold with equality.

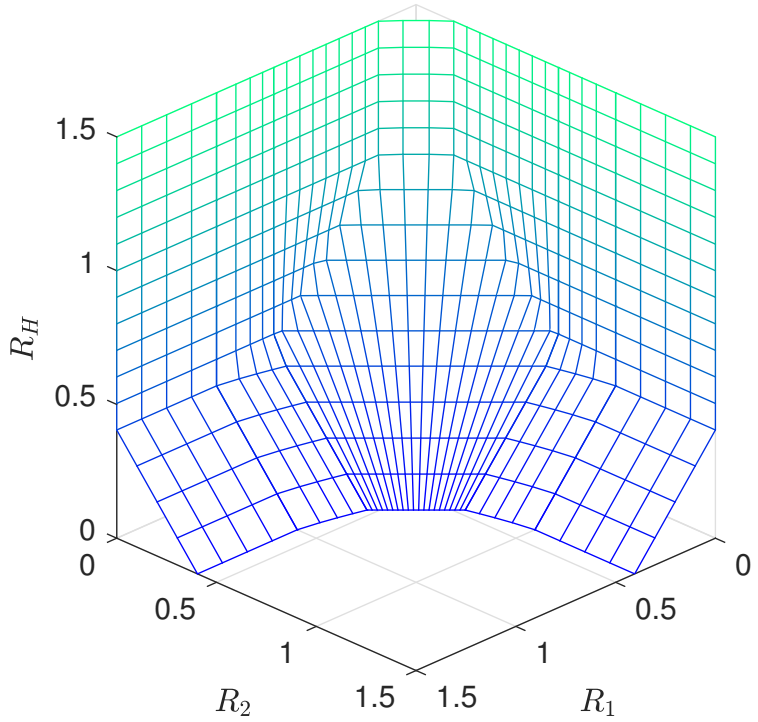
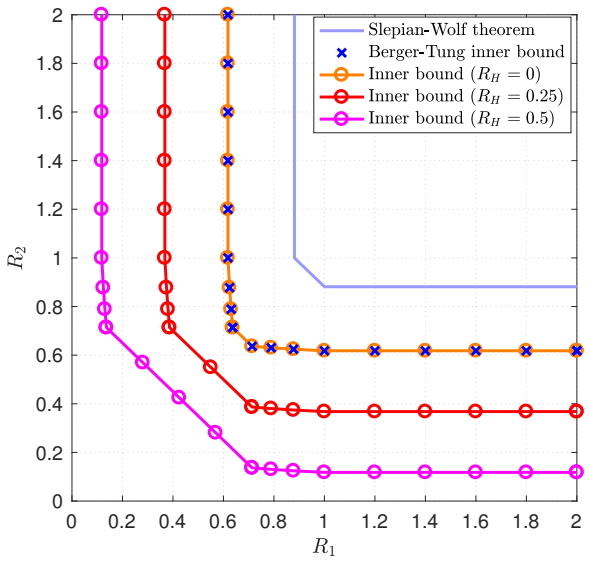
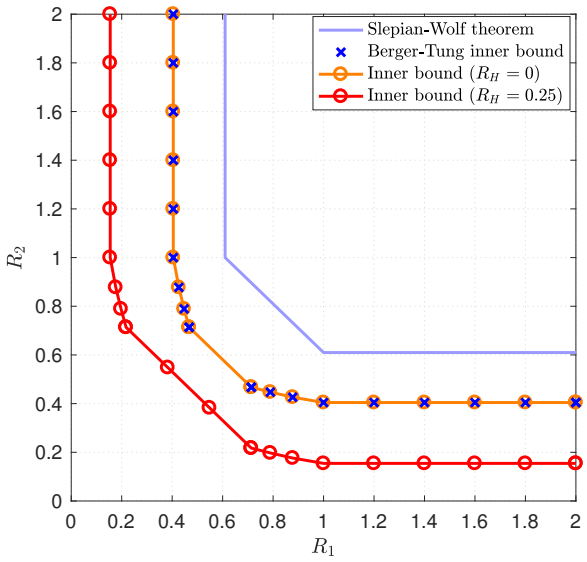


Fig. 2.2. An overall view of the inner bound on the achievable rate-distortion region, with $\rho = 0.15$ and $D_1 = D_2 = 0.05$.

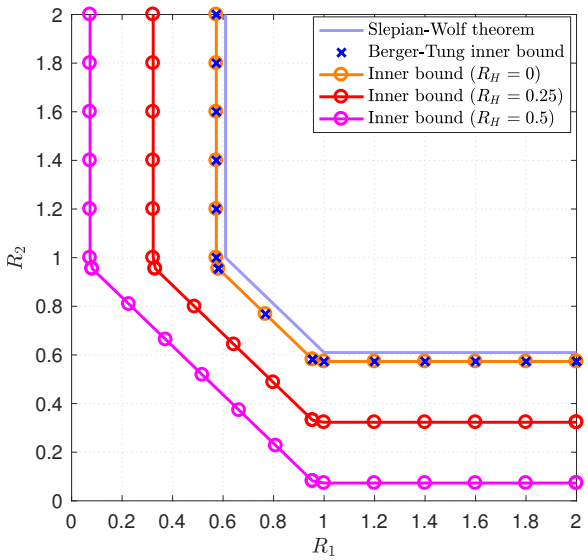
An overall view of the inner bound on the achievable rate-distortion region is provided in Fig. 2.2, where the crossover probability ρ between sources is set at 0.15 and the desired distortion pair (D_1, D_2) is set at $(0.05, 0.05)$. Clearly, the helper can enhance the robustness of transmissions by expanding the achievable rate-distortion region as R_{H} increases. However, the above part of the region, i.e., $R_{\mathrm{H}} \geq 1$, does not change even if the helper rate increases.



(a) $\rho = 0.3, D_1 = D_2 = 0.05$.



(b) $\rho = 0.15, D_1 = D_2 = 0.05$.



(c) $\rho = 0.15, D_1 = D_2 = 0.005$.

Fig. 2.3. The inner bound on the achievable rate-distortion region for given R_{H}.

In order to verify the results by classical theorems, i.e., the Slepian-Wolf theorem and Berger-Tung inner bound, we plot the inner bound on the achievable rate-distortion region for given values of R_{H} in Fig. 2.3. Note that the Slepian-Wolf theorem is for lossless multiterminal source coding, i.e., $D_1 = D_2 = 0$. Hence, the inner bound is closer to the Slepian-Wolf region when the required

distortions are smaller, as shown in Fig. 2.3(c); conversely, the constraints on the coding rates become less strict if more distortions are acceptable. Besides, the inner bound derived in this chapter perfectly coincides with the Berger-Tung inner bound when $R_{\mathrm{H}} = 0$, in spite of the distortion requirements and the correlation level between sources. This phenomenon results from the fact that the system model shown in Fig. 2.1 reduces to the model used in the Berger-Tung inner bound, if the helper link is removed, i.e., equivalently $R_{\mathrm{H}} = 0$. Consequently, the Berger-Tung inner bound can be utilized as a baseline for comparison when analyzing the performance improvement provided by the helper. It is noticeable that the achievable rate-distortion region is obviously enlarged as the helper rate increases, i.e., the communications become more reliable after introducing a helper.

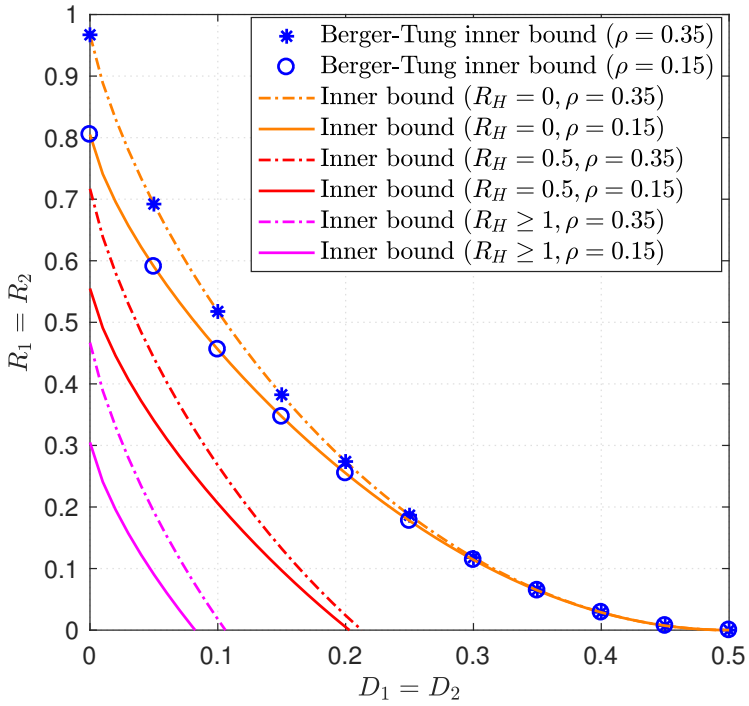


Fig. 2.4. The inner bound on the rate-distortion function for symmetric source links.

We can investigate the trade-off between the coding rates and the desired distortions in depth from the curve of the rate-distortion, by setting two source links as symmetric links, i.e., $R_1 = R_2$ and $D_1 = D_2$. As depicted in Fig. 2.4, the derived inner bound with $R_{\mathrm{H}} = 0$ also precisely matches the Berger-Tung inner bound in terms of the rate-distortion function. For $R_{\mathrm{H}} > 0$, we can observe significant performance gain provided by a helper. Another interesting observation is that there are still some distortions when $R_1 = R_2 = 0$, even if $R_{\mathrm{H}} \geq 1$. The reason for the unavoidable distortion is that $H(X_1^n, X_2^n) = H(X_1^n) + H(X_2^n|X_1^n) = n + nH_b(\rho) > n \geq H(Y^n)$ for all $\rho > 0$. Hence, the decoder can losslessly reconstruct X_1 and X_2 only by the helper information Y, if and only if the sources are completely correlated. There is also no doubt that more correlated sources, i.e., ρ is smaller, require lower coding rates for satisfying the same distortion requirements. However, the gap between the curves with different ρ becomes narrower as the distortion requirements go larger. Therefore, with the increment of the desired distortions, the correlations between sources have less effect on the coding rates.

2.3 Outage Probability Analysis

Now, we start to analyze the outage probability for the system illustrated in Fig. 2.1 suffering from independent block Rayleigh fadings. The outage event occurs, if at least one of the recovered sequences cannot satisfy the specified distortion requirement of the source links. According to Shannon's lossy source-channel separation theorem, the channel condition can be taken into account by constraining the source coding rate with channel capacity. It should be emphasized here that the source and helper sequences are not actually encoded into the codeword M for lossy compression, although we utilize the framework of lossy distributed multiterminal source coding problem to equivalently analyze the distortions determined by link rates.

2.3.1 Derivation of Outage Probability

Due to the independent block Rayleigh fading assumption on each link, the channel gains h_1, h_2 and h_{H} in the link of X_1, X_2 and Y independently follow the two-dimensional Gaussian distribution. Hence, the *pdf*s of instantaneous SNRs in the three links are

$$f(\gamma_i) = \frac{1}{\overline{\gamma}_i} \exp(-\frac{\gamma_i}{\overline{\gamma}_i}), \text{ for } i \in \{1, 2\}, \tag{2.25}$$

$$f(\gamma_{\mathrm{H}}) = \frac{1}{\overline{\gamma}_{\mathrm{H}}} \exp(-\frac{\gamma_{\mathrm{H}}}{\overline{\gamma}_{\mathrm{H}}}). \tag{2.26}$$

For the purpose of simplicity, we assume that the channel state information (CSI) is only available at the receiver sides, and the effect of shadowing is not taken into account.

According to Shannon's lossy source-channel separation theorem, the distortion requirements (D_1, D_2) can be satisfied if the following inequalities hold:

$$R_i(D_i) \leq \Theta_i(\gamma_i) = \frac{C(\gamma_i)}{r_i}, \text{ for } i \in \{1, 2\}, \tag{2.27}$$

$$R_{\mathrm{H}} \leq \Theta_{\mathrm{H}}(\gamma_{\mathrm{H}}) = \frac{C(\gamma_{\mathrm{H}})}{r_{\mathrm{H}}}, \tag{2.28}$$

where $C(\cdot)$ is the Shannon capacity using Gaussian codebook; r_i and r_{H} stand for the end-to-end coding rates. Therefore, if the link rates $(R_1, R_2, R_{\mathrm{H}})$ supported by $(\Theta_1(\gamma_1), \Theta_2(\gamma_2), \Theta_{\mathrm{H}}(\gamma_{\mathrm{H}}))$ fall outside the achievable rate-distortion region, the distortion requirements (D_1, D_2) cannot be satisfied, i.e., outage event occurs.

For a given value of R_{H} and specified distortion requirements (D_1, D_2), the shape of achievable rate-distortion region can be illustrated as Fig. 2.5. In order to calculate the outage probability, we divide the area of outage into six subareas with corresponding probabilities $P_1, P_2, \cdots, P_6$. Hence, the outage probability is calculated as

$$P_{\text{out}} = \sum_{k=1}^{6} P_k. \tag{2.29}$$

Noticed that we use the inner bound on the achievable rate-distortion region here; therefore, the calculation result is the upper bound of the outage probability.

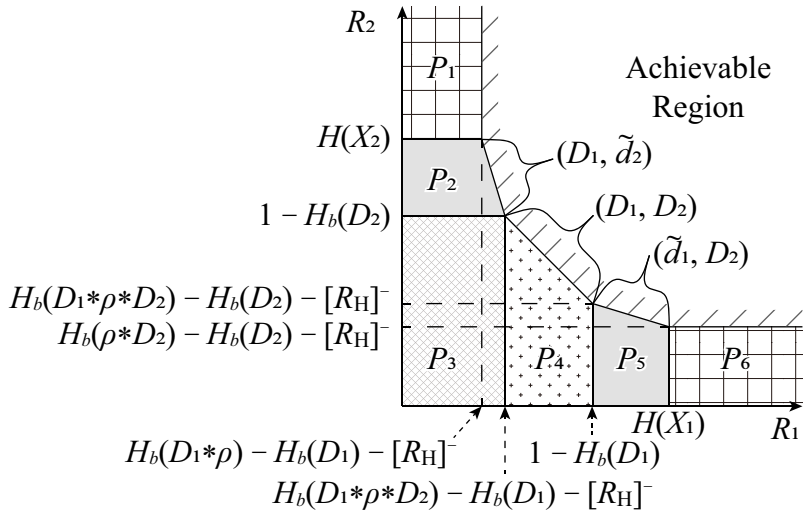


Fig. 2.5. Achievable rate-distortion region for calculating outage probability.

First, consider

$$\begin{aligned}
P_1 &= \Pr\{0 \le R_1 \le H_b(D_1 * \rho) - H_b(D_1) - [R_{\mathrm{H}}]^-, H(X_2) \le R_2, 0 \le R_{\mathrm{H}}\} \\
&= \Pr\{0 \le \Theta_1(\gamma_1) \le \lambda_1(0), H(X_2) \le \Theta_2(\gamma_2), 0 \le \Theta_{\mathrm{H}}(\gamma_{\mathrm{H}})\} \\
&= \Pr\{\Theta_1^{-1}(0) \le \gamma_1 \le \Theta_1^{-1}[\lambda_1(0)], \Theta_2^{-1}[H(X_2)] \le \gamma_2, \Theta_{\mathrm{H}}^{-1}(0) \le \gamma_{\mathrm{H}}\} \\
&= \int_{\Theta_{\mathrm{H}}^{-1}(0)}^{\infty} d\gamma_{\mathrm{H}} \int_{\Theta_1^{-1}(0)}^{\Theta_1^{-1}[\lambda_1(0)]} d\gamma_1 \cdot \int_{\Theta_2^{-1}[H(X_2)]}^{\infty} f(\gamma_2) f(\gamma_1) f(\gamma_{\mathrm{H}}) d\gamma_2, \qquad (2.30)
\end{aligned}$$

where $\lambda_i(\tilde{d}) = \max\{0, H_b(D_i * \rho * \tilde{d}) - H_b(D_i) - [\Theta_{\mathrm{H}}(\gamma_{\mathrm{H}})]^-\}$ for $i \in \{1, 2\}$. Notice that $H_b(D_i * \rho * \tilde{d}) \le 1$ and $H_b(D_i) \ge 0$; hence, $\lambda_i(\tilde{d}) = 0$ when $\Theta_{\mathrm{H}}(\gamma_{\mathrm{H}}) = R_{\mathrm{H}} > 1$. Therefore, P_1 can be further calculated as

$$\begin{aligned}
P_1 &= \int_{\Theta_{\mathrm{H}}^{-1}(0)}^{\Theta_{\mathrm{H}}^{-1}(1)} d\gamma_{\mathrm{H}} \int_{\Theta_1^{-1}(0)}^{\Theta_1^{-1}[\lambda_1(0)]} d\gamma_1 \cdot \int_{\Theta_2^{-1}[H(X_2)]}^{\infty} f(\gamma_2) f(\gamma_1) f(\gamma_{\mathrm{H}}) d\gamma_2 \\
&\quad + \int_{\Theta_{\mathrm{H}}^{-1}(1)}^{\infty} d\gamma_{\mathrm{H}} \int_{\Theta_1^{-1}(0)}^{\Theta_1^{-1}[0]} d\gamma_1 \cdot \int_{\Theta_2^{-1}[H(X_2)]}^{\infty} f(\gamma_2) f(\gamma_1) f(\gamma_{\mathrm{H}}) d\gamma_2 \\
&= \int_{\Theta_{\mathrm{H}}^{-1}(0)}^{\Theta_{\mathrm{H}}^{-1}(1)} d\gamma_{\mathrm{H}} \int_{\Theta_1^{-1}(0)}^{\Theta_1^{-1}[\lambda_1(0)]} d\gamma_1 \cdot \int_{\Theta_2^{-1}[H(X_2)]}^{\infty} f(\gamma_2) f(\gamma_1) f(\gamma_{\mathrm{H}}) d\gamma_2 + 0 \\
&= \frac{1}{\overline{\gamma}_{\mathrm{H}}} \cdot \exp\left(-\frac{\Theta_2^{-1}[H(X_2)]}{\overline{\gamma}_2}\right) \cdot \int_{\Theta_{\mathrm{H}}^{-1}(0)}^{\Theta_{\mathrm{H}}^{-1}(1)} \exp\left(-\frac{\gamma_{\mathrm{H}}}{\overline{\gamma}_{\mathrm{H}}}\right) \\
&\quad \cdot \left[1 - \exp\left(-\frac{\Theta_1^{-1}[\lambda_1(0)]}{\overline{\gamma}_1}\right)\right] d\gamma_{\mathrm{H}}. \qquad (2.31)
\end{aligned}$$

Next, consider

$$\begin{aligned}
P_2 &= \Pr\{0 \le R_1 \le H_b(D_1 * \rho * \tilde{d}_2) - H_b(D_1) - [R_{\mathrm{H}}]^-, 1 - H_b(D_2) \le R_2 < H(X_2), 0 \le R_{\mathrm{H}}\} \\
&= \Pr\{\Theta_1^{-1}(0) \le \gamma_1 \le \Theta_1^{-1}[\lambda_1(\tilde{d}_2)], \Theta_2^{-1}[1 - H_b(D_2)] \le \gamma_2 < \Theta_2^{-1}[H(X_2)], \Theta_{\mathrm{H}}^{-1}(0) \le \gamma_{\mathrm{H}}\}
\end{aligned}$$

$$
\begin{aligned}
&= \int_{\Theta_{\mathrm{H}}^{-1}(0)}^{\infty} d\gamma_{\mathrm{H}} \int_{\Theta_2^{-1}[1-H_b(D_2)]}^{\Theta_2^{-1}[H(X_2)]} d\gamma_2 \cdot \int_{\Theta_1^{-1}(0)}^{\Theta_1^{-1}[\lambda_1(\tilde{d}_2)]} f(\gamma_1) f(\gamma_2) f(\gamma_{\mathrm{H}}) d\gamma_1 \\
&= \int_{\Theta_{\mathrm{H}}^{-1}(0)}^{\Theta_{\mathrm{H}}^{-1}(1)} d\gamma_{\mathrm{H}} \int_{\Theta_2^{-1}[1-H_b(D_2)]}^{\Theta_2^{-1}[H(X_2)]} d\gamma_2 \cdot \int_{\Theta_1^{-1}(0)}^{\Theta_1^{-1}[\lambda_1(\tilde{d}_2)]} f(\gamma_1) f(\gamma_2) f(\gamma_{\mathrm{H}}) d\gamma_1 \\
&\quad + \int_{\Theta_{\mathrm{H}}^{-1}(1)}^{\infty} d\gamma_{\mathrm{H}} \int_{\Theta_2^{-1}[1-H_b(D_2)]}^{\Theta_2^{-1}[H(X_2)]} d\gamma_2 \cdot \int_{\Theta_1^{-1}(0)}^{\Theta_1^{-1}[0]} f(\gamma_1) f(\gamma_2) f(\gamma_{\mathrm{H}}) d\gamma_1 \\
&= \int_{\Theta_{\mathrm{H}}^{-1}(0)}^{\Theta_{\mathrm{H}}^{-1}(1)} d\gamma_{\mathrm{H}} \int_{\Theta_2^{-1}[1-H_b(D_2)]}^{\Theta_2^{-1}[H(X_2)]} d\gamma_2 \cdot \int_{\Theta_1^{-1}(0)}^{\Theta_1^{-1}[\lambda_1(\tilde{d}_2)]} f(\gamma_1) f(\gamma_2) f(\gamma_{\mathrm{H}}) d\gamma_1 + 0 \\
&= \frac{1}{\overline{\gamma}_2 \overline{\gamma}_{\mathrm{H}}} \cdot \int_{\Theta_{\mathrm{H}}^{-1}(0)}^{\Theta_{\mathrm{H}}^{-1}(1)} d\gamma_{\mathrm{H}} \\
&\quad \cdot \int_{\Theta_2^{-1}[1-H_b(D_2)]}^{\Theta_2^{-1}[H(X_2)]} \exp\left(-\frac{\gamma_2}{\overline{\gamma}_2} - \frac{\gamma_{\mathrm{H}}}{\overline{\gamma}_{\mathrm{H}}}\right) \cdot \left[1 - \exp\left(-\frac{\Theta_1^{-1}[\lambda_1(\tilde{d}_2)]}{\overline{\gamma}_1}\right)\right] d\gamma_2, \qquad (2.32)
\end{aligned}
$$

where $\tilde{d}_i = H_b^{-1}(1 - [R_i]^-)$ for $i \in \{1, 2\}$.

For P_3, consider

$$
\begin{aligned}
P_3 &= \Pr\{0 \le R_1 < H_b(D_1 * \rho * D_2) - H_b(D_1) - [R_{\mathrm{H}}]^-, 0 \le R_2 < 1 - H_b(D_2), 0 \le R_{\mathrm{H}}\} \\
&= \Pr\{\Theta_1^{-1}(0) \le \gamma_1 < \Theta_1^{-1}[\lambda_1(D_2)], \Theta_2^{-1}(0) \le \gamma_2 < \Theta_2^{-1}[1 - H_b(D_2)], \Theta_{\mathrm{H}}^{-1}(0) \le \gamma_{\mathrm{H}}\} \\
&= \int_{\Theta_{\mathrm{H}}^{-1}(0)}^{\infty} d\gamma_{\mathrm{H}} \int_{\Theta_2^{-1}(0)}^{\Theta_2^{-1}[1-H_b(D_2)]} d\gamma_2 \int_{\Theta_1^{-1}(0)}^{\Theta_1^{-1}[\lambda_1(D_2)]} f(\gamma_1) f(\gamma_2) f(\gamma_{\mathrm{H}}) d\gamma_1 \\
&= \int_{\Theta_{\mathrm{H}}^{-1}(0)}^{\Theta_{\mathrm{H}}^{-1}(1)} d\gamma_{\mathrm{H}} \int_{\Theta_2^{-1}(0)}^{\Theta_2^{-1}[1-H_b(D_2)]} d\gamma_2 \cdot \int_{\Theta_1^{-1}(0)}^{\Theta_1^{-1}[\lambda_1(D_2)]} f(\gamma_1) f(\gamma_2) f(\gamma_{\mathrm{H}}) d\gamma_1 \\
&\quad + \int_{\Theta_{\mathrm{H}}^{-1}(1)}^{\infty} d\gamma_{\mathrm{H}} \int_{\Theta_2^{-1}(0)}^{\Theta_2^{-1}[1-H_b(D_2)]} d\gamma_2 \cdot \int_{\Theta_1^{-1}(0)}^{\Theta_1^{-1}[0]} f(\gamma_1) f(\gamma_2) f(\gamma_{\mathrm{H}}) d\gamma_1 \\
&= \int_{\Theta_{\mathrm{H}}^{-1}(0)}^{\Theta_{\mathrm{H}}^{-1}(1)} d\gamma_{\mathrm{H}} \int_{\Theta_2^{-1}(0)}^{\Theta_2^{-1}[1-H_b(D_2)]} d\gamma_2 \cdot \int_{\Theta_1^{-1}(0)}^{\Theta_1^{-1}[\lambda_1(D_2)]} f(\gamma_1) f(\gamma_2) f(\gamma_{\mathrm{H}}) d\gamma_1 + 0 \\
&= \frac{1}{\overline{\gamma}_{\mathrm{H}}} \cdot \left[1 - \exp\left(-\frac{\Theta_2^{-1}[1 - H_b(D_2)]}{\overline{\gamma}_2}\right)\right] \\
&\quad \cdot \int_{\Theta_{\mathrm{H}}^{-1}(0)}^{\Theta_{\mathrm{H}}^{-1}(1)} \exp\left(-\frac{\gamma_{\mathrm{H}}}{\overline{\gamma}_{\mathrm{H}}}\right) \cdot \left[1 - \exp\left(-\frac{\Theta_1^{-1}[\lambda_1(D_2)]}{\overline{\gamma}_1}\right)\right] d\gamma_{\mathrm{H}}. \qquad (2.33)
\end{aligned}
$$

To calculate P_4, consider

$$
\begin{aligned}
P_4 &= \Pr\{H_b(D_1 * \rho * D_2) - H_b(D_1) - [R_{\mathrm{H}}]^- \le R_1 < 1 - H_b(D_1), \\
&\qquad 0 \le R_2 \le 1 + H_b(D_1 * \rho * D_2) - H_b(D_1) - H_b(D_2) - [R_{\mathrm{H}}]^- - [R_1]^-, 0 \le R_{\mathrm{H}}\} \\
&= \Pr\{\Theta_1^{-1}[\lambda_1(D_2)] \le \gamma_1 \le \Theta_1^{-1}[1 - H_b(D_1)], \Theta_2^{-1}(0) \le \gamma_2 \le \Theta_2^{-1}(\mu_1), \Theta_{\mathrm{H}}^{-1}(0) \le \gamma_{\mathrm{H}}\} \\
&= \int_{\Theta_{\mathrm{H}}^{-1}(0)}^{\infty} d\gamma_{\mathrm{H}} \int_{\Theta_1^{-1}[\lambda_1(D_2)]}^{\Theta_1^{-1}[1-H_b(D_1)]} d\gamma_1 \cdot \int_{\Theta_2^{-1}(0)}^{\Theta_2^{-1}(\mu_1)} f(\gamma_2) f(\gamma_1) f(\gamma_{\mathrm{H}}) d\gamma_2 \\
&= \int_{\Theta_{\mathrm{H}}^{-1}(0)}^{\Theta_{\mathrm{H}}^{-1}(1)} d\gamma_{\mathrm{H}} \int_{\Theta_1^{-1}[\lambda_1(D_2)]}^{\Theta_1^{-1}[1-H_b(D_1)]} d\gamma_1 \cdot \int_{\Theta_2^{-1}(0)}^{\Theta_2^{-1}(\mu_1)} f(\gamma_2) f(\gamma_1) f(\gamma_{\mathrm{H}}) d\gamma_2
\end{aligned}
$$

$$
\begin{aligned}
&+\int_{\Theta_{\mathrm{H}}^{-1}(1)}^{\infty} d\gamma_{\mathrm{H}} \int_{\Theta_1^{-1}[0]}^{\Theta_1^{-1}[1-H_b(D_1)]} d\gamma_1 \cdot \int_{\Theta_2^{-1}(0)}^{\Theta_2^{-1}(\mu_1)} f(\gamma_2) f(\gamma_1) f(\gamma_{\mathrm{H}}) d\gamma_2 \\
&= \frac{1}{\overline{\gamma}_1 \overline{\gamma}_{\mathrm{H}}} \cdot \int_{\Theta_{\mathrm{H}}^{-1}(0)}^{\Theta_{\mathrm{H}}^{-1}(1)} d\gamma_{\mathrm{H}} \cdot \int_{\Theta_1^{-1}[\lambda_1(D_2)]}^{\Theta_1^{-1}[1-H_b(D_1)]} \exp\left(-\frac{\gamma_1}{\overline{\gamma}_1} - \frac{\gamma_{\mathrm{H}}}{\overline{\gamma}_{\mathrm{H}}}\right) \cdot \left[1 - \exp\left(-\frac{\Theta_2^{-1}(\mu_1)}{\overline{\gamma}_2}\right)\right] d\gamma_1 \\
&\quad + \frac{1}{\overline{\gamma}_1} \cdot \exp\left(-\frac{\Theta_{\mathrm{H}}^{-1}(1)}{\overline{\gamma}_{\mathrm{H}}}\right) \\
&\quad \cdot \int_{\Theta_1^{-1}[0]}^{\Theta_1^{-1}[1-H_b(D_1)]} \exp\left(-\frac{\gamma_1}{\overline{\gamma}_1}\right) \cdot \left[1 - \exp\left(-\frac{\Theta_2^{-1}(\mu_1')}{\overline{\gamma}_2}\right)\right] d\gamma_1, \qquad (2.34)
\end{aligned}
$$

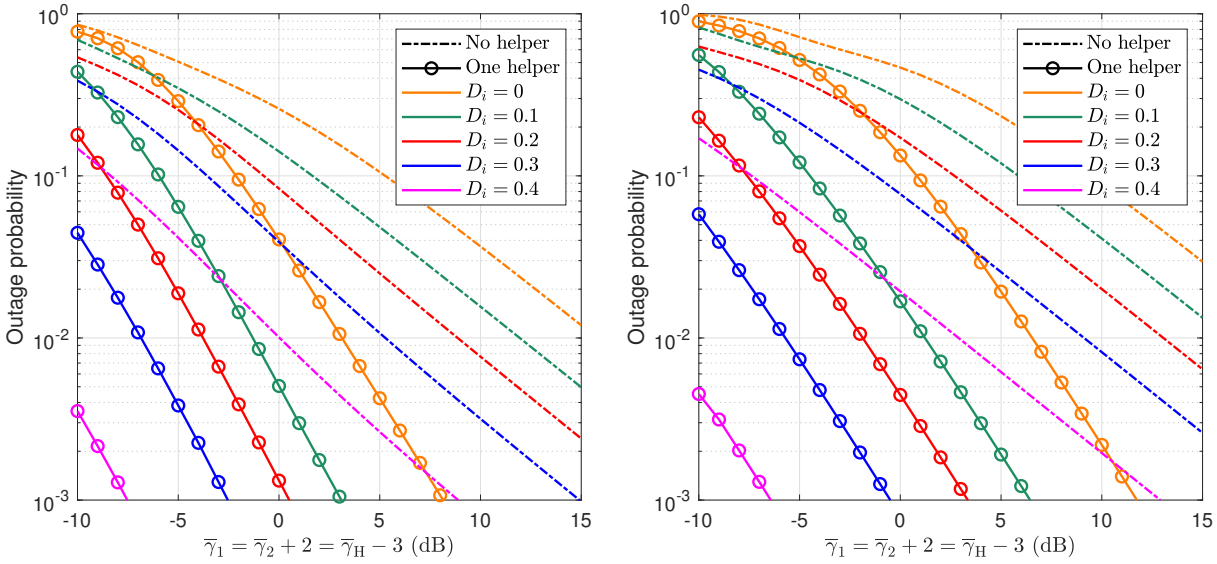
where $\mu_i = \max\{0, 1 + H_b(D_1 * \rho * D_2) - H_b(D_1) - H_b(D_2) - [\Theta_{\mathrm{H}}(\gamma_{\mathrm{H}})]^- - [\Theta_i(\gamma_i)]^-\}$ and $\mu_i' = \max\{0, H_b(D_1 * \rho * D_2) - H_b(D_1) - H_b(D_2) - [\Theta_i(\gamma_i)]^-\}$ for $i \in \{1, 2\}$.

Similar to the calculation of P_2 and P_1, we have

$$
\begin{aligned}
P_5 &= \frac{1}{\overline{\gamma}_1 \overline{\gamma}_{\mathrm{H}}} \cdot \int_{\Theta_{\mathrm{H}}^{-1}(0)}^{\Theta_{\mathrm{H}}^{-1}(1)} d\gamma_{\mathrm{H}} \\
&\quad \cdot \int_{\Theta_1^{-1}[1-H_b(D_1)]}^{\Theta_1^{-1}[H(X_1)]} \exp\left(-\frac{\gamma_1}{\overline{\gamma}_1} - \frac{\gamma_{\mathrm{H}}}{\overline{\gamma}_{\mathrm{H}}}\right) \cdot \left[1 - \exp\left(-\frac{\Theta_2^{-1}[\lambda_2(\tilde{d}_1)]}{\overline{\gamma}_2}\right)\right] d\gamma_1, \qquad (2.35) \\
P_6 &= \frac{1}{\overline{\gamma}_{\mathrm{H}}} \cdot \exp\left(-\frac{\Theta_1^{-1}[H(X_1)]}{\overline{\gamma}_1}\right) \\
&\quad \cdot \int_{\Theta_{\mathrm{H}}^{-1}(0)}^{\Theta_{\mathrm{H}}^{-1}(1)} \exp\left(-\frac{\gamma_{\mathrm{H}}}{\overline{\gamma}_{\mathrm{H}}}\right) \cdot \left[1 - \exp\left(-\frac{\Theta_2^{-1}[\lambda_2(0)]}{\overline{\gamma}_2}\right)\right] d\gamma_{\mathrm{H}}. \qquad (2.36)
\end{aligned}
$$

2.3.2 Numerical Results

Fig. 2.6 depicts the upper bound of the outage probability for lossy communications with two sources and one helper. For the purpose of plotting the curves in a 2D plane, the average SNRs are set at different values but change at the same speed for each link, i.e., $\overline{\gamma}_1 = \overline{\gamma}_2 + 2 = \overline{\gamma}_{\mathrm{H}} - 3$. It should be explained here that the curves for the case without a helper only show first order diversity, due to the definition of outage event, i.e., the recoveries of the both sources cannot satisfy the distortion requirements. We can clearly observe from the decay of the curves that with a helper, second order diversity can be achieved. Therefore, the case with a helper providing the side information achieves higher diversity order than the case without a helper. Interestingly, the energy efficiency of the whole system can be improved by introducing a helper. For example, for two sources without a helper, first order diversity means that the outage probability decreases from 10^{-1} to 10^{-2} requires 10 dB increment of per-link average SNR (20 dB in total). However, for two sources with a helper, second order diversity means that it only requires 5 dB increment of per-link average SNR (15 dB in total). Consequently, the system with a helper consumes lower sum power for the same decrement of outage probability. From Fig. 2.6(a) and Fig. 2.6(b), it is found that the smaller the correlation between the sources, i.e., ρ is larger, the higher the outage probability, which is consistent to our expected results. In addition, it is noticeable that the system will have lower outage probability, if the distortion requirements are less strict. Hence, identifying how to specify the distortion requirements is a valuable and interesting work in the design of practical systems.



(a) $\rho = 0.1$. (b) $\rho = 0.4$.

Fig. 2.6. The upper bound of the outage probability for the sources with different correlation levels.

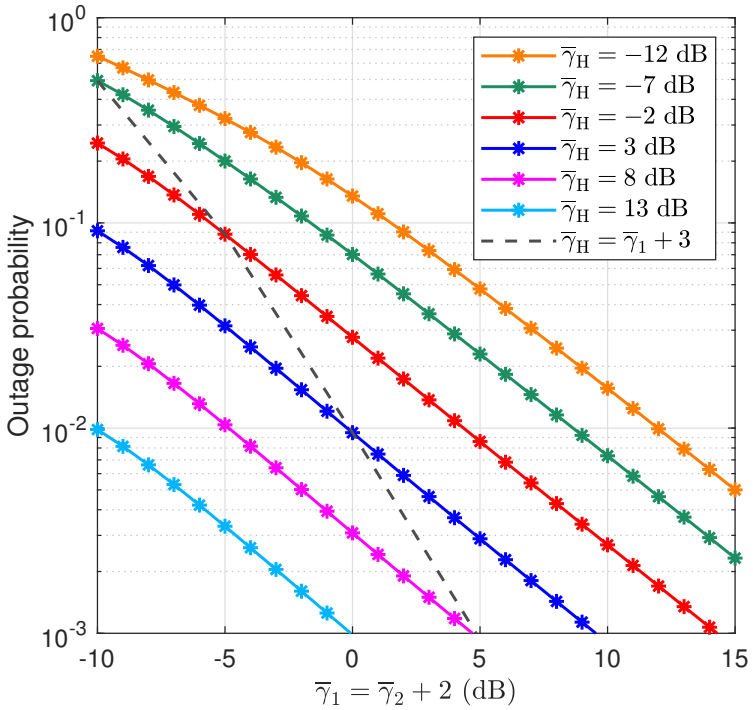


Fig. 2.7. The upper bound of the outage probability for diverse average SNRs in the helper link, with $\rho = 0.2$ and $D_1 = D_2 = 0.1$.

Fig. 2.7 illustrates the upper bound of the outage probability for fixed average SNR in the helper link, i.e., only $\overline{\gamma}_1 = \overline{\gamma}_2 + 2$. It is obvious that the outage curves are shifted to left by increasing the average SNR in the helper link. However, the gap between the outage curves becomes narrower when $\overline{\gamma}_H$ is relatively small. Because the channel capacity of the helper link cannot support lossless

transmission of the helper information if $\gamma_{\mathrm{H}} < 0$ dB, while there is no loss of the helper information for $\gamma_{\mathrm{H}} \geq 0$ dB. Besides, notice that the slope of curve in Fig. 2.7 is less steep than that of the case with one helper in Fig. 2.6. The reason for this observation is because the value of $\overline{\gamma}_{\mathrm{H}}$ is fixed (does not change along with $\overline{\gamma}_1$ and $\overline{\gamma}_2$) in Fig. 2.7, and the effect of performance improvement provided by a helper appears in the form of parallel shifting of outage curves. If we use a dashed line to connect the points where $\overline{\gamma}_1 = \overline{\gamma}_2 + 2 = \overline{\gamma}_{\mathrm{H}} - 3$ like in Fig. 2.6, we can observe second order diversity.

2.4 Summary

In this chapter, we have analyzed the performance of lossy multiterminal source coding with a helper in terms of rate-distortion and outage probability over independent block Rayleigh fading channels. Based on Shannon's lossy source-channel separation theorem, we start the performance analysis from multiterminal source coding with a helper, and then take the channel conditions into consideration in the derivation of outage probability. For the rate-distortion analysis, we derive an inner bound on the achievable rate-distortion region for multiterminal source coding with two binary sources and one helper. The theoretical results demonstrate that a helper can obviously enlarge the achievable rate-distortion region; moreover, the inner bound derived in this dissertation coincides with the Berger-Tung inner bound when the helper rate decreases to 0, i.e., the helper link is equivalently removed. Then, we apply the derived inner bound into the outage probability analysis. In order to briefly calculate the outage probability, the area outside the achievable rate-distortion region is divided into several subareas. Finally, we investigate the performance improvement by introducing a helper with respect to the outage probability. The curves of outage probability indicate that a helper can make the system achieve higher order of diversity, and improve the energy efficiency for the whole system.

CHAPTER 3

Binary CEO Problem with a Helper

Chapter 2 validates the performance gain provided by a helper for direct transmissions from the sources to the destination. In this chapter, we are interested in the performance improvement by introducing a helper into the system without direct transmission from the source to the destination, i.e., the binary CEO problem. Inspired by the solutions of the binary CEO problem in [22, 23], we can start the performance analysis of the binary CEO problem with a helper from a multiterminal source coding problem and then investigate the distortion of final decision. Moreover, we conduct a series of simulations to compare the performance improvement between locating a helper and introducing an additional agent.

3.1 Problem Statement

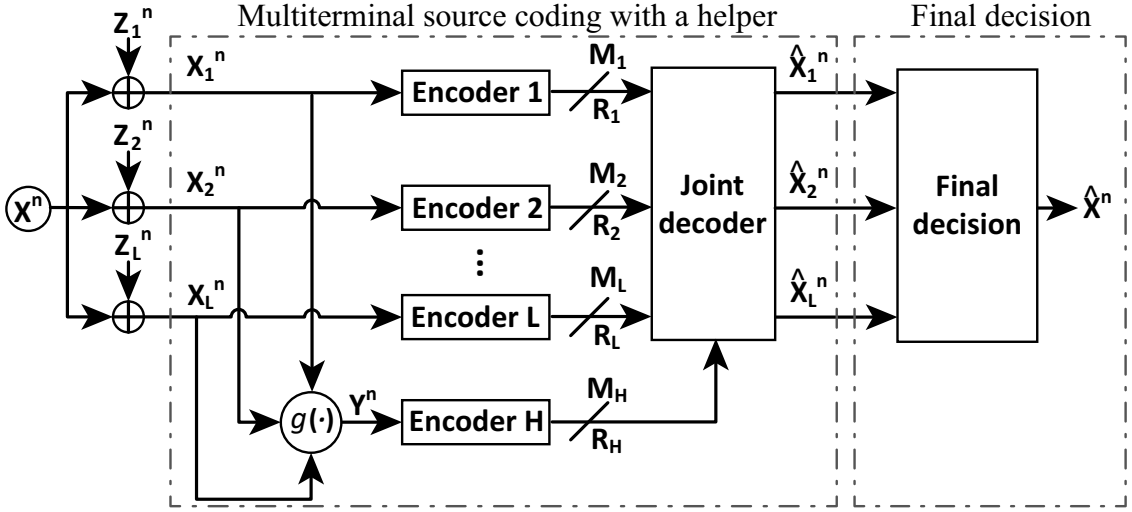


Fig. 3.1. The binary CEO problem with a helper.

As illustrated in Fig. 3.1[1], there is a binary source X acts as a common source in the binary CEO problem with a helper. The DMS X generates i.i.d. sequence $x^n = \{x(t)\}_{t=1}^n$ by taking values from the binary alphabet $\mathcal{X} = \{0, 1\}$ for each time slot. The source X is observed by L agents at the same time. Due to the influence of noise, the observation $x_i^n = \{x_i(t)\}_{t=1}^n$ may contain errors $z_i^n = \{z_i(t)\}_{t=1}^n$ for $i \in \mathcal{L}$. Hence, the error probability $\Pr\{x_i(t) \neq x(t)\} = \rho_i$ for $Z_i \sim \mathrm{Bern}(\rho_i)$. Simultaneously, a helper generates the helper sequence $y^n = \{y(t)\}_{t=1}^n$ from the agent sequences bit by bit, and then transmits the helper sequence to the destination after compressing it. Therefore, all the X_i and Y can be also regarded as DMS. The sequences x_i^n and y^n are encoded at rates R_i

[1]It should be noted that this dissertation solve the binary CEO problem with a helper by a successive decoding scheme, i.e., first reconstructing $\hat{X}_{\mathcal{L}}^n$ and then making the final decision $\hat{X}^n$. The successive decoding scheme may not achieve the optimal performance as directly reconstructing $\hat{X}^n$ from $\{M_1, M_2, \cdots, M_L, M_{\mathrm{H}}\}$.

and R_{H} by encoder i and encoder H, respectively. Encoder i and encoder H assign an index to each sequence according to the following mapping rules:

$$\varphi_i : \mathcal{X}_i^n \mapsto \mathcal{M}_i = \{1, 2, \cdots, 2^{nR_i}\}, \tag{3.1}$$

$$\varphi_{\mathrm{H}} : \mathcal{Y}^n \mapsto \mathcal{M}_{\mathrm{H}} = \{1, 2, \cdots, 2^{nR_{\mathrm{H}}}\}. \tag{3.2}$$

Then, the encoder outputs $M_{\mathcal{L}}$ and M_{H} are transmitted to a joint decoder. The joint decoder constructs the estimates $\hat{x}_{\mathcal{L}}^n$ from indices $M_{\mathcal{L}}$ and M_{H} by utilizing the mapping rule, as:

$$\psi : \mathcal{M}_1 \times \mathcal{M}_2 \times \cdots \times \mathcal{M}_L \times \mathcal{M}_{\mathrm{H}} \mapsto \mathcal{X}_1^n \times \mathcal{X}_2^n \times \cdots \times \mathcal{X}_L^n. \tag{3.3}$$

Since the estimate $\hat{x}_i^n$ may occasionally deviate from the observation x_i^n, the Hamming distortion measure is implemented to describe the distortion level between $x_i(t)$ and $\hat{x}_i(t)$. For given distortion values $D_{\mathcal{L}}$, the rate-distortion region $\mathcal{R}(D_{\mathcal{L}})$, consisting of all achievable rate tuple $(R_{\mathcal{L}}, R_{\mathrm{H}})$, is defined as

$$\begin{aligned}\mathcal{R}(D_{\mathcal{L}}) = \{(R_{\mathcal{L}}, R_{\mathrm{H}}) : (R_{\mathcal{L}}, R_{\mathrm{H}}) \text{ is admissible such that} \\ \lim_{n\to\infty} \mathrm{E}(d_i(x_i^n, \hat{x}_i^n)) \leq D_i + \epsilon, \text{ for } i \in \mathcal{L}, \text{ and any } \epsilon > 0\}.\end{aligned} \tag{3.4}$$

Finally, the destination reconstructs the estimate of x^n from $\hat{x}_{\mathcal{L}}^n$. Obviously, the distortion between x_i^n and $\hat{x}_i^n$ will determine the final estimate $\hat{x}^n$. Hence, the final distortion

$$\mathrm{E}\left[\frac{1}{n}\sum_{t=1}^{n} d(x(t), \hat{x}(t))\right] \leq D + \epsilon, \tag{3.5}$$

can be formulated as a function $\mathrm{F}_{\mathrm{DP}}(\cdot)$ of $D_{\mathcal{L}}$, where $\mathrm{F}_{\mathrm{DP}}(\cdot)$ is referred to as DPF. The DPF is defined as $D = \mathrm{F}_{\mathrm{DP}}(D_{\mathcal{L}})$, which highly depends on the decision rule. It should be emphasized here that the DPF limits the decoding scheme to successive decoding.

In the literature, He *et al.* [22] derived an outer bound on the achievable rate-distortion region for the binary CEO problem with two sources only, and then the outer bound was extended to the case with arbitrary number of sources in [23]. The major theoretical result in this chapter is the outer bound on the achievable rate-distortion region for the binary CEO problem with many sources and one helper.

3.2 The Step of Multiterminal Source Coding

Since distortions will propagate from the decoding results of multiterminal source coding to the final decision, we should minimize the distortion from the first step of multiterminal source coding. Therefore, for distortion minimization by convex optimization, we need to derive an outer bound on the achievable rate-distortion region of multiterminal source coding with a helper.

3.2.1 Outer Bound on the Rate-Distortion Region

From the extended Berger-Tung outer bound [50] with multiple sources and one link of side information, we can obtain the outer bound for multiterminal source coding with a helper as presented in the following proposition.

Proposition 3.1: Let $(X_{\mathcal{L}}, Y)$ be a $(L+1)$-component DMS and $d_i(x_i, \hat{x}_i)$ for $i \in \mathcal{L}$ be distortion measures. If a rate tuple $(R_{\mathcal{L}}, R_{\mathrm{H}})$ is achievable with distortion tuple $D_{\mathcal{L}}$ for distributed

lossy source coding with a helper observing Y, then it must satisfy the inequalities

$$\sum_{i\in\mathcal{S}} R_i \geq I(X_{\mathcal{L}}; U_{\mathcal{S}}|U_{\mathcal{S}^c}, V), \tag{3.6}$$

$$R_{\mathrm{H}} \geq I(Y; V), \tag{3.7}$$

for some conditional *pmf* $p(u_{\mathcal{L}}, v|\ x_{\mathcal{L}}, y)$, and functions $\hat{x}_i(u_{\mathcal{L}}, v)$ such that $U_i \to X_i \to X_j$, $X_i \to X_j \to U_j$, and $X_{\mathcal{S}} \to Y \to V$ form Markov chains and $\mathrm{E}(d_i(X_i, \hat{X}_i)) \leq D_i$, where $i, j \in \mathcal{L}$ and $i \neq j$.

It is easy to understand that the constraint of (3.7) results from the rate limit on the helper instead of uncompressed side information. Now, we start to derive the outer bound for binary sources. Consider

$$\begin{aligned}
\sum_{i\in\mathcal{S}} R_i &\geq I(X_{\mathcal{L}}; U_{\mathcal{S}}|U_{\mathcal{S}^c}, V) \\
&= H(X_{\mathcal{L}}|U_{\mathcal{S}^c}, V) - H(X_{\mathcal{L}}|U_{\mathcal{S}}, U_{\mathcal{S}^c}, V) \\
&= H(X_{\mathcal{L}}|U_{\mathcal{S}^c}) - I(X_{\mathcal{L}}; V|U_{\mathcal{S}^c}) - H(X_{\mathcal{L}}|U_{\mathcal{S}}, U_{\mathcal{S}^c}) + I(X_{\mathcal{L}}; V|U_{\mathcal{S}}, U_{\mathcal{S}^c}) \\
&= I(X_{\mathcal{L}}; U_{\mathcal{S}}|U_{\mathcal{S}^c}) - I(X_{\mathcal{L}}; V|U_{\mathcal{S}^c}) + I(X_{\mathcal{L}}; V|U_{\mathcal{S}}, U_{\mathcal{S}^c}) \\
&= I(X_{\mathcal{L}}; U_{\mathcal{S}}|U_{\mathcal{S}^c}) + I(X_{\mathcal{S}}; V|U_{\mathcal{L}}) - I(X_{\mathcal{S}}; V|U_{\mathcal{L}}) - I(X_{\mathcal{L}}; V|U_{\mathcal{S}^c}) \\
&\quad + I(X_{\mathcal{L}}; V|U_{\mathcal{S}}, U_{\mathcal{S}^c}).
\end{aligned} \tag{3.8}$$

Then, we calculate $I(X_{\mathcal{L}}; U_{\mathcal{S}}|U_{\mathcal{S}^c}) + I(X_{\mathcal{S}}; V|U_{\mathcal{L}})$ and $-I(X_{\mathcal{S}}; V|U_{\mathcal{L}}) - I(X_{\mathcal{L}}; V|U_{\mathcal{S}^c}) + I(X_{\mathcal{L}}; V|U_{\mathcal{S}}, U_{\mathcal{S}^c})$ separately. Consider

$$\begin{aligned}
&I(X_{\mathcal{L}}; U_{\mathcal{S}}|U_{\mathcal{S}^c}) + I(X_{\mathcal{S}}; V|U_{\mathcal{L}}) \\
&\quad= I(X_{\mathcal{S}}; U_{\mathcal{S}}|U_{\mathcal{S}^c}) + I(X_{\mathcal{S}^c}; U_{\mathcal{S}}|U_{\mathcal{S}^c}, X_{\mathcal{S}}) + I(X_{\mathcal{S}}; V|U_{\mathcal{L}}) \\
&\quad= I(X_{\mathcal{S}}; U_{\mathcal{S}}|U_{\mathcal{S}^c}) + H(X_{\mathcal{S}^c}|U_{\mathcal{S}^c}, X_{\mathcal{S}}) - H(X_{\mathcal{S}^c}|U_{\mathcal{S}}, U_{\mathcal{S}^c}, X_{\mathcal{S}}) + I(X_{\mathcal{S}}; V|U_{\mathcal{L}}) \\
&\quad= I(X_{\mathcal{S}}; U_{\mathcal{S}}|U_{\mathcal{S}^c}) + H(X_{\mathcal{S}^c}|U_{\mathcal{S}^c}, X_{\mathcal{S}}) - H(X_{\mathcal{S}^c}|U_{\mathcal{S}^c}, X_{\mathcal{S}}) + I(X_{\mathcal{S}}; V|U_{\mathcal{L}}) \qquad (3.9) \\
&\quad= I(X_{\mathcal{S}}; U_{\mathcal{S}}|U_{\mathcal{S}^c}) + I(X_{\mathcal{S}}; V|U_{\mathcal{L}}) \\
&\quad= I(X_{\mathcal{S}}; U_{\mathcal{S}}, U_{\mathcal{S}^c}) - I(X_{\mathcal{S}}; U_{\mathcal{S}^c}) + I(X_{\mathcal{S}}; V|U_{\mathcal{S}}, U_{\mathcal{S}^c}) \\
&\quad= I(X_{\mathcal{S}}; U_{\mathcal{L}}, V) - I(X_{\mathcal{S}}; U_{\mathcal{S}^c}) \\
&\quad\geq I(X_{\mathcal{S}}; \hat{X}_{\mathcal{S}}) - I(X_{\mathcal{S}}; U_{\mathcal{S}^c}), \qquad (3.10)
\end{aligned}$$

where (3.9) follows the fact that $U_i \to X_i \to X_j$ form a Markov chain for $i \in \mathcal{S}$ and $j \in \mathcal{S}^c$; (3.10) follows data processing inequality when estimating $\hat{X}_{\mathcal{S}}$ from $(U_{\mathcal{L}}, V)$. Notice that $I(X_{\mathcal{S}}; \hat{X}_{\mathcal{S}})$ is the mutual information required for recovering $\hat{X}_{\mathcal{S}}$, and $I(X_{\mathcal{S}}; U_{\mathcal{S}^c})$ is the mutual information provided by the remaining agent links except the helper link. Hence, $I(X_{\mathcal{S}}; \hat{X}_{\mathcal{S}}) - I(X_{\mathcal{S}}; U_{\mathcal{S}^c})$ represents the link rates required for multiterminal source coding without a helper, and we can further bound $I(X_{\mathcal{L}}; U_{\mathcal{S}}|U_{\mathcal{S}^c}) + I(X_{\mathcal{S}}; V|U_{\mathcal{L}})$ by applying the result in [23] into (3.10), as

$$I(X_{\mathcal{L}}; U_{\mathcal{S}}|U_{\mathcal{S}^c}) + I(X_{\mathcal{S}}; V|U_{\mathcal{L}}) \geq f_b(\{\rho_{\mathcal{S}}, \alpha_{\mathcal{S}^c}\}) - f_b(\{\alpha_{\mathcal{S}^c}\}) - \sum_{i\in\mathcal{S}} H_b(D_i), \tag{3.11}$$

where $\alpha_i = \rho_i * H_b^{-1}(1 - [R_i]^-)$, and $f_b(\cdot)$ is the joint binary entropy function for correlated sources defined in **Definition 1.3**.

Next, consider

$$\begin{aligned}
&- I(X_{\mathcal{S}}; V|U_{\mathcal{L}}) - I(X_{\mathcal{L}}; V|U_{\mathcal{S}^c}) + I(X_{\mathcal{L}}; V|U_{\mathcal{S}}, U_{\mathcal{S}^c}) \\
&\quad = -I(X_{\mathcal{S}}; V|U_{\mathcal{L}}) - H(V|U_{\mathcal{S}^c}) + H(V|X_{\mathcal{L}}, U_{\mathcal{S}^c}) + H(V|U_{\mathcal{S}}, U_{\mathcal{S}^c}) - H(V|X_{\mathcal{L}}, U_{\mathcal{S}}, U_{\mathcal{S}^c}) \\
&\quad = -I(X_{\mathcal{S}}; V|U_{\mathcal{L}}) - H(V|U_{\mathcal{S}^c}) + H(V|X_{\mathcal{L}}) + H(V|U_{\mathcal{S}}, U_{\mathcal{S}^c}) - H(V|X_{\mathcal{L}}) \qquad (3.12) \\
&\quad = H(V|U_{\mathcal{L}}, X_{\mathcal{S}}) - H(V|U_{\mathcal{L}}) - H(V|U_{\mathcal{S}^c}) + H(V|U_{\mathcal{S}}, U_{\mathcal{S}^c}) \\
&\quad = -I(V; U_{\mathcal{S}}, X_{\mathcal{S}}|U_{\mathcal{S}^c}) \\
&\quad = -I(V; X_{\mathcal{S}}|U_{\mathcal{S}^c}), \qquad (3.13)
\end{aligned}$$

where (3.12) and (3.13) follow since $U_i \to X_i \to V$ form a Markov chain for $i \in \mathcal{L}$. To further calculate (3.13), consider

$$\begin{aligned}
I(V; X_{\mathcal{S}}|U_{\mathcal{S}^c}) &\leq I(V; Y) \qquad (3.14) \\
&\leq [R_{\mathrm{H}}]^-, \qquad (3.15)
\end{aligned}$$

where (3.14) follows since $X_{\mathcal{S}} \to Y \to V$ form a Markov chain. Notice that the equality of (3.14) holds when the helper only utilize $X_{\mathcal{S}}$ to generate the helper information Y. Moreover, the equality of (3.15) holds when the helper rate is completely exploited for compression. Therefore, (3.15) might not hold for all cases regardless the specified helper structure. Consequently, there is no waste of the helper rate if the conditions for the equality of (3.14) and (3.15) are satisfied. In this case, it is obvious that the structure of the helper is optimal. By assuming that we have an optimal helper and substituting (3.11), (3.13) and (3.15) into (3.8), we can finally obtain

$$\sum_{i\in\mathcal{S}} R_i \geq f_b(\{\rho_{\mathcal{S}}, \alpha_{\mathcal{S}^c}\}) - f_b(\{\alpha_{\mathcal{S}^c}\}) - \sum_{i\in\mathcal{S}} H_b(D_i) - [R_{\mathrm{H}}]^-. \qquad (3.16)$$

Remark 1: Since we assume that the structure of the helper is optimal, the constraint on the helper link, i.e., the inequality (3.7), is satisfied by the helper encoder, which finds a proper codeword M_{H} to make $I(Y;V)$ as large as possible.

Remark 2: If we set $R_{\mathrm{H}} = 0$, i.e., the helper link is equivalently cut off, (3.16) reduces to the outer bound without a helper in [23].

3.2.2 Distortion Minimization by Convex Optimization

Since the final distortion D is a function of $D_{\mathcal{L}}$ by the successive decoding for given $R_{\mathcal{L}}$ and R_{H}, we can first minimize the l_2-norm of the vector $[D_1, D_2, \cdots, D_L]$ by solving a convex optimization problem [22, 23], which is formulated from the outer bound. Then, we calculate the minimum distortion D^* by substituting the solution of the convex optimization problem into DPF. Notice that for a practical communication system, the channel capacity should also be taken into consideration. According to Shannon's lossy source-channel separation theorem [61, 68], the distortion tuple $D_{\mathcal{L}}$ is achievable if the following inequalities hold:

$$R_i(D_i) \leq \Theta_i(\gamma_i) = \frac{C(\gamma_i)}{r_i}, \text{ for } i \in \mathcal{L}, \qquad (3.17)$$

$$R_{\mathrm{H}} \leq \Theta_{\mathrm{H}}(\gamma_{\mathrm{H}}) = \frac{C(\gamma_{\mathrm{H}})}{r_{\mathrm{H}}}. \qquad (3.18)$$

By applying the outer bound derived above, we can formulate the convex optimization problem for the system with an optimal helper as

$$\begin{aligned}
&\min_{D_1,D_2,\cdots,D_L} \|[D_1, D_2, \cdots, D_L]\|_2 \\
&s.t. \\
&\quad \sum_{i\in\mathcal{S}} H_b(D_i) \geq f_b(\{\rho_{\mathcal{S}}, \alpha_{\mathcal{S}^c}\}) - f_b(\{\alpha_{\mathcal{S}^c}\}) - \sum_{i\in\mathcal{S}}[\Theta_i(\gamma_i)]^- - [\Theta_{\mathrm{H}}(\gamma_{\mathrm{H}})]^-, \\
&\quad 0 \leq D_i \leq 0.5, i \in \mathcal{L}.
\end{aligned} \tag{3.19}$$

After solving the convex optimization problem, we can obtain the minimum value of distortion D_i^* for $i \in \mathcal{L}$. Then, we use the estimates $\hat{X}_{\mathcal{L}}^n$ with minimum distortions $D_{\mathcal{L}}^*$ to make final decision.

3.3 Final Distortion Analysis

As stated before, the final distortion highly depends on the decision rule, and hence we have to specify the decision rule for final distortion analysis. In the following, we consider the theoretically optimal case and a practical case, i.e., MV decision, respectively.

3.3.1 Optimal Decision

To facilitate the derivation, we define $\beta_i = \rho_i * D_i$. For the optimal decision rule, consider

$$\begin{aligned}
H(X) - H_b(\tilde{d}) & \\
&\leq I(X;\hat{X}) \\
&\leq I(X;\hat{X}_{\mathcal{L}}) && (3.20) \\
&= H(X) + H(\hat{X}_{\mathcal{L}}) - H(X,\hat{X}_{\mathcal{L}}) \\
&\leq H(X) + f_b(\{\beta_{\mathcal{L}}\}) - f_b(\{0,\beta_{\mathcal{L}}\}), && (3.21)
\end{aligned}$$

where the steps are justified as:
(3.20) the probable information loss in the final decision,
(3.21) X can be regarded as the output of a BSC with itself as the input and the crossover probability equal to 0.

Consequently, we have

$$\tilde{d} \geq H_b^{-1}[f_b(\{0,\beta_{\mathcal{L}}\}) - f_b(\{\beta_{\mathcal{L}}\})]. \tag{3.22}$$

Obviously, the minimum final distortion D, i.e., the distortion by optimal decision, is given by

$$D = H_b^{-1}[f_b(\{0,\beta_{\mathcal{L}}\}) - f_b(\{\beta_{\mathcal{L}}\})]. \tag{3.23}$$

3.3.2 Majority Voting Decision

Since the optimal decision specifies a universal lower bound, here, we consider another practical decision scheme, i.e., MV. The distortion between X^n and $\hat{X}^n$ by MV is the sum probability of

several events in a Poisson binomial process [52]. We introduce a function to evaluate the distortion in a Poisson binomial process as follows:

Definition 3.1: Poisson binomial distortion function [52]. The distortion between X^n and $\hat{X}^n$, which is estimated from $\hat{X}^n_{\mathcal{L}}$ by MV, is calculated by $D = \mathrm{PB}(\beta_{\mathcal{L}})$ as

$$\mathrm{PB}(\beta_{\mathcal{L}}) = \begin{cases} \sum\limits_{j=\frac{L+1}{2}}^{L} \Pr(J=j), & \text{if } L \text{ is odd,} \\ \frac{1}{2}\Pr(J=\frac{L}{2}) + \sum\limits_{j=\frac{L}{2}+1}^{L} \Pr(J=j), & \text{if } L \text{ is even,} \end{cases} \tag{3.24}$$

where

$$\Pr(J=j) = \begin{cases} \prod\limits_{i=1}^{L}(1-\beta_i), & j=0, \\ \frac{1}{j}\sum\limits_{i=1}^{j}(-1)^{(i-1)}\Pr(J=j-i)\eta(i), & j>0, \end{cases} \tag{3.25}$$

with $\eta(i) = \sum_{k=1}^{L}(\frac{\beta_k}{1-\beta_k})^i$ for $0 \le j \le L$.

By utilizing the Poisson binomial distortion function, we can calculate the distortion D by MV among $\hat{X}_{\mathcal{L}}$ as

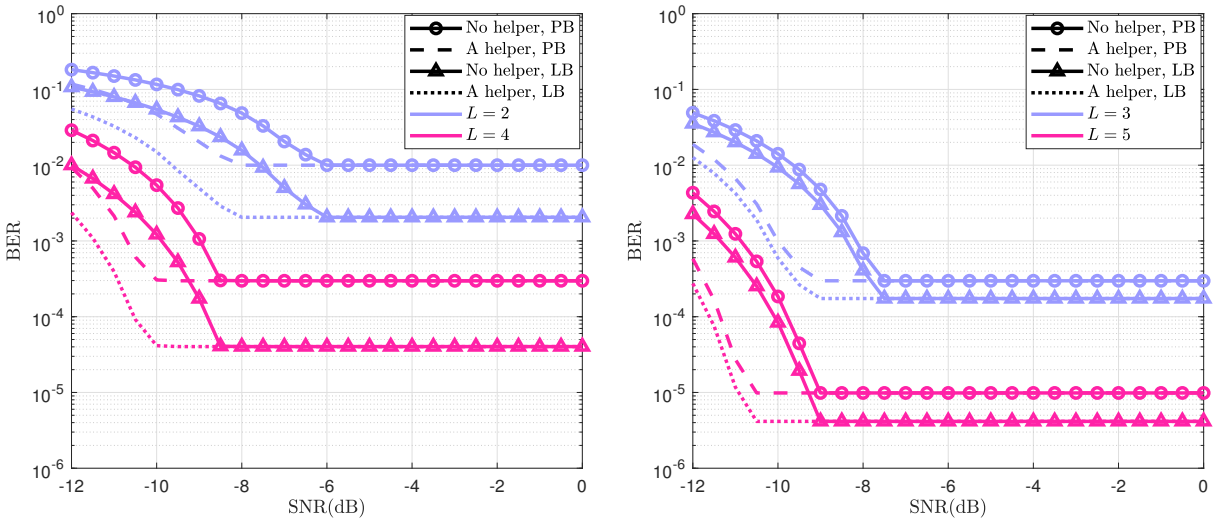
$$D = \mathrm{PB}(\beta_{\mathcal{L}}). \tag{3.26}$$

Remark: If ρ_i are various among all links, weighted MV should be implemented to generate more accurate estimate of X^n. The error probability by weighted MV is presented in Appendix A. In this chapter, we focus on the system with homogeneous agents for simplicity, and the results can be easily extended to the case with heterogeneous agents according to Appendix A.

In summary, the DPFs for optimal decision and MV are (3.23) and (3.26), respectively. Therefore, we can obtain the minimum final distortion between X^n and $\hat{X}^n$ by substituting the solution $D^*_{\mathcal{L}}$ of the convex optimization problem into DPF as $D^* = \mathrm{F}_{\mathrm{DP}}(\beta^*_{\mathcal{L}})$, where $\beta^*_{\mathcal{L}} = \{\rho_i * D^*_i | i \in \mathcal{L}\}$.

3.3.3 Numerical Results

Now, we start investigations on the trade-off of rate-distortion for the binary CEO problem with a helper through numerical results. A memoryless source $X \sim \mathrm{Bern}(0.5)$ is used in the following. Initially, we compare the BER performance between MV and optimal decision. By utilizing the DPF after solving a corresponding convex optimization problem, we can depict the curve of SNR for each link versus BER as shown in Fig. 3.2. All the crossover probabilities between X and X_i are set at the same value of 0.01. Moreover, the end-to-end coding rate is set at $\frac{1}{2}$, and the SNR is set at the same level for all of the agent and helper links. Notice from the results that whether there is a helper or not, a gap obviously appears between the Poisson binomial (PB) process, i.e., MV, and the theoretical lower bound (LB), i.e., optimal decision. The reason for the performance gap is that it is extremely difficult to completely utilize the mutual information between X^n and $X^n_{\mathcal{L}}$. For instance, assuming that there are $2K$ agents with $X_t = X_{1,t} = \cdots = X_{K,t} = 0$ and $X_{K+1,t} = \cdots = X_{2K,t} = 1$ at some time index t, it is obvious that there is some mutual information between X_t and $\{X_{1,t}, \cdots, X_{K,t}\}$. However, decision error will still occur in the Poisson binomial



(a) The number of agents is even.

(b) The number of agents is odd.

Fig. 3.2. Comparison of system performance between MV and optimal decision, where $\rho_i = 0.01$.

process, resulting in the loss of mutual information between X_t and $\hat{X}_t$. It should be also noticed that the gap is sensitive to the number of agents, i.e., the gap is smaller when there are odd number of agents. Because there could be equal number of "0" and "1" at the same bit of the agent sequences, if the number of agents is even. In this draw case, the bit of final decision by MV is randomly selected, resulting in more performance loss. Hence, the MV decision can achieve better performance if odd number of agents are deployed. For the effect of a helper, we find that the helper can reduce BER at small SNR value range, but keep the same BER floor as the case without a helper. Because for sufficiently large SNR, there is already no distortion between X_i^n and $\hat{X}_i^n$, and the side information Y^n generated from $X_{\mathcal{L}}^n$ becomes redundant.

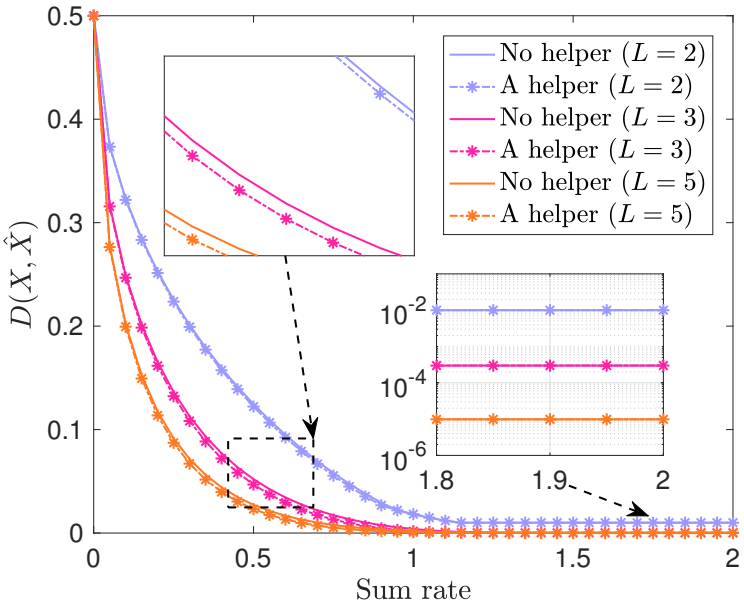
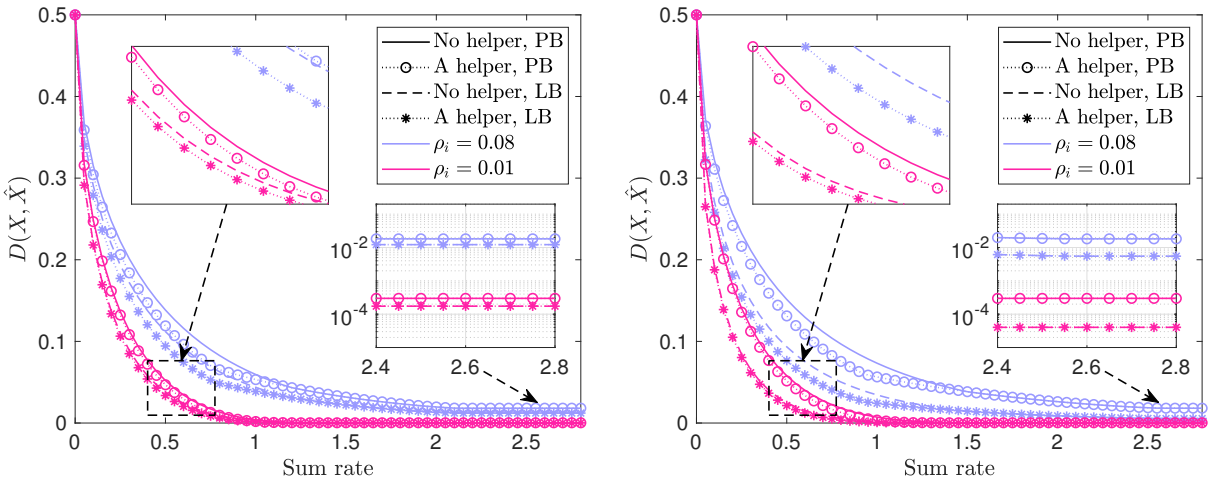


Fig. 3.3. The trade-off of rate-distortion with a MV helper, where $\rho_i = 0.01$ for all agents.

We further investigate the effect on rate-distortion by introducing a helper with MV decision in Fig. 3.3. Since the rate allocation scheme is out of the scope in this dissertation, the rate is evenly allocated to all nodes including agents and helper. Surprisingly, the helper can reduce the final distortion D before achieving the distortion floor, even though the agent rate decreases by sharing the sum rate with a helper. This phenomenon indicates that it is possible to improve the system performance for multiple access channels with the same sum rate by introducing a helper. However, the curve with a helper will finally converge with the no-helper case at the same distortion floor. Notice the fact that the curve with a helper is still valuable, although the improvement is very small.



(a) $L = 3$. (b) $L = 4$.

Fig. 3.4. The trade-off of rate-distortion with diverse number of agents.

Finally, we make a comparison of rate-distortion between MV and optimal decision with diverse ρ_i. Fig. 3.4 shows an inclination that the more correlated observations are, i.e., ρ_i is smaller, the faster the distortion floor is achieved. Because it is easier to minimize the distortion between X_i^n and $\hat{X}_i^n$ for more correlated observations, owing to more mutual information among the observations. Moreover, as the correlation among the agent observations increases, the distortion floor decreases for both MV and optimal decision. We can also find the same phenomenon as Fig. 3.2 that the gap between MV and optimal decision is smaller for odd number of agents, even if the number of agents in Fig. 3.4(a) is less than that in Fig. 3.4(b). In addition, since in the situation where the mutual information is lost less frequently with smaller ρ_i, the gap between MV and optimal decision is smaller with more correlated observations.

3.4 Practical Performance Evaluation

3.4.1 Simulation Design

In this section, we evaluate the practical performance of binary CEO problem with a helper through simulations. As depicted in Fig. 3.5, there are L encoders separately encode the observations X_i^n, which is detected from a common source sequence X^n and mixed with the error Z_i^n. Simultaneously, the encoder H encodes the side information Y^n generated from the agent observations.

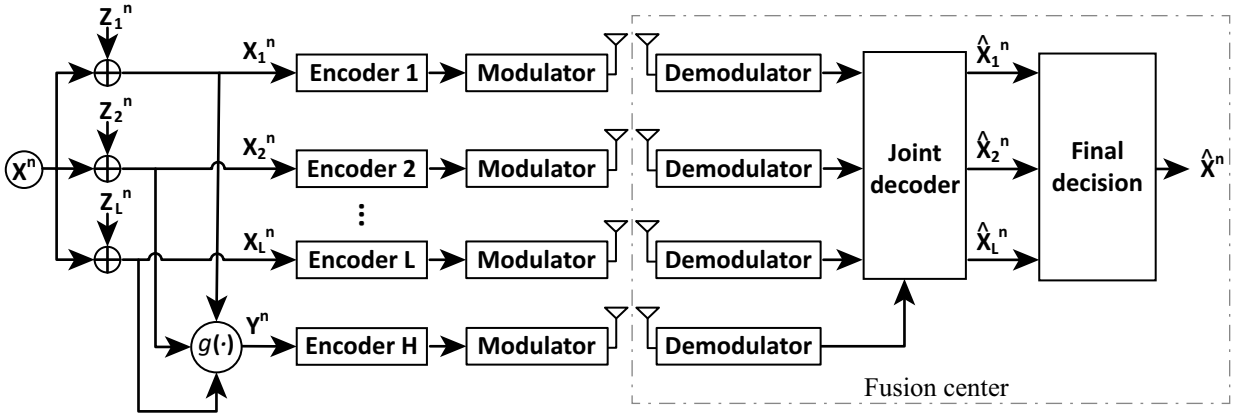


Fig. 3.5. Simulation system.

Next, the encoded sequences are sent to a fusion center through AWGN channels after modulation. The fusion center first demodulates the received signals and then jointly decodes them. Finally, the estimate of all sequences in the last round of iteration is used to make final decision.

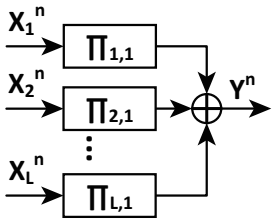
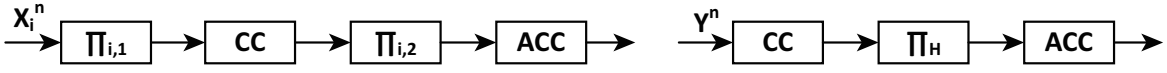


Fig. 3.6. Generation of helper information by modulo-2 sum.

Regarding the helper sequence Y^n, since its optimal structure is still an open problem, we select two frequently implemented structures in practice as $g(\cdot)$, i.e., the helper information generated by modulo-2 sum or MV. The $g(\cdot)$ by modulo-2 sum is illustrated in Fig. 3.6, where X_i^n is interleaved by $\Pi_{i,1}$ for the first step; subsequently, Y^n is produced by the modulo-2 sum of interleaved $X_{\mathcal{L}}^n$ bit by bit. By the interleaver $\Pi_{i,1}$, the distribution of Y is approximate to $\mathrm{Bern}(0.5)$, i.e., Y^n can contain side information as much as possible.



(a) Agent encoder i. (b) Helper encoder H.

Fig. 3.7. The structure of encoders.

Fig. 3.7 shows the structure of encoders. In order to better exploit the correlation among $X_{\mathcal{L}}^n$, the interleaver $\Pi_{i,1}$ is used to disperse noises into different bits of X_i^n. For the helper, only one interleaver Π_{H} is needed between CC and ACC.

The general structure of joint decoder is depicted in Fig. 3.8. In local iteration, the extrinsic information is exchanged between ACC^{-1} and CC^{-1} via an interleaver Π and its corresponding deinterleaver Π^{-1}. After several rounds of local iteration, the CC^{-1} outputs the *a posteriori* LLR (LLR^p) of information bits. In global iteration, an extrinsic information exchanger updates the *a*

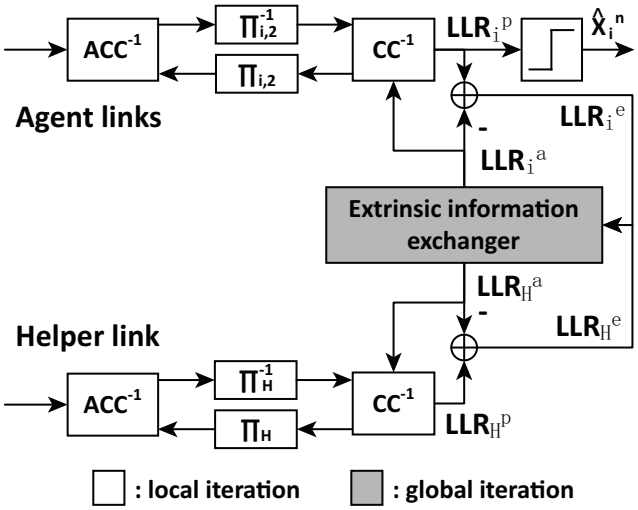
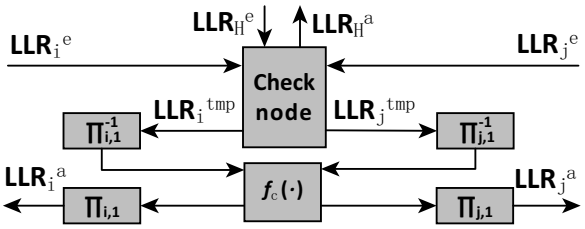
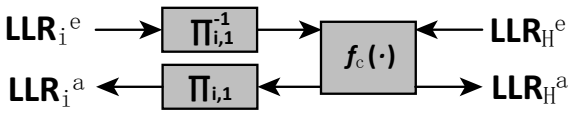


Fig. 3.8. The structure of joint decoder.

priori LLR (LLR^a) with the extrinsic LLR (LLR^e), which is calculated by $(\mathrm{LLR}^p - \mathrm{LLR}^a)$. The joint decoder alternately executes local iteration and global iteration, until the mutual information calculated from LLR_i^p is large enough or the maximum iteration time is exceeded. Finally, the estimate of X_i^n is produced by the hard decision of LLR_i^p in the last round of local iteration.



(a) The helper information is generated by modulo-2 sum, where $i, j \in \mathcal{L}$ and $i \neq j$.



(b) The helper information is generated by MV.

Fig. 3.9. The structure of extrinsic information exchanger.

The extrinsic information exchanger updates the LLR^a according to the rules shown in Fig. 3.9. For the case with helper information generated by modulo-2 sum, the LLR^e of all agents and the helper is first exchanged based on the same principle for the check node of low-density parity-check (LDPC) codes [78], as

$$\mathrm{LLR}_i^{\mathrm{tmp}} = \mathrm{LLR}_i^e - 2 \cdot \operatorname{arctanh}\left(\prod_{j \in \mathcal{W}} \tanh \frac{-\mathrm{LLR}_j^e}{2} \cdot \tanh \frac{-\mathrm{LLR}_{\mathrm{H}}^e}{2} \right), \tag{3.27}$$

$$\mathrm{LLR}_{\mathrm{H}}^a = \mathrm{LLR}_{\mathrm{H}}^e - 2 \cdot \operatorname{arctanh}\left(\prod_{i \in \mathcal{L}} \tanh \frac{-\mathrm{LLR}_i^e}{2} \right), \tag{3.28}$$

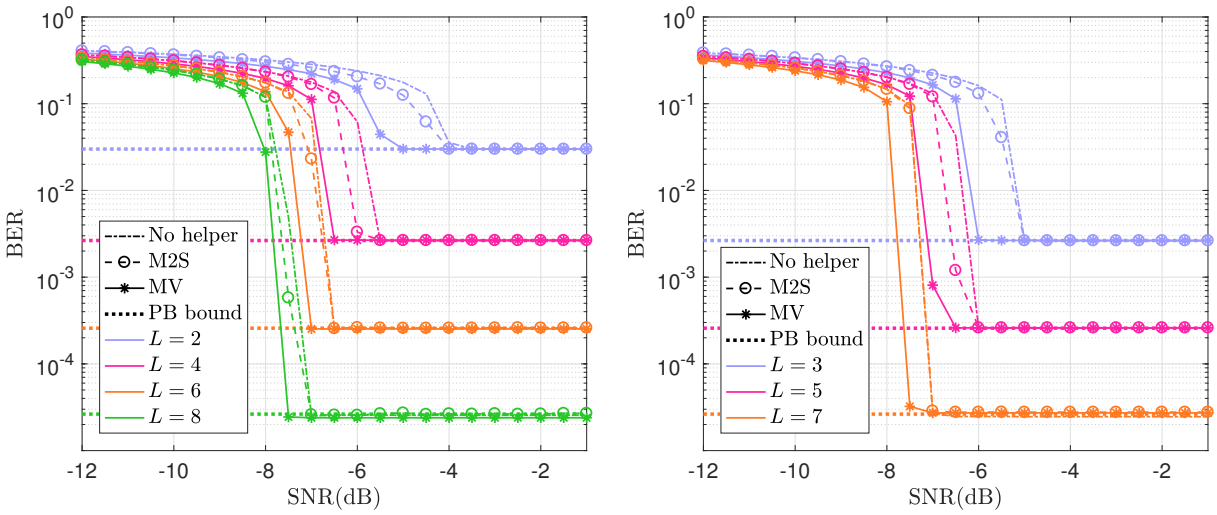
where $\mathcal{W} = \mathcal{L} \backslash i$, and $\mathrm{LLR}_i^{\mathrm{tmp}}$ is the temporary result for agents. Then, according to the correlation model [79], the $\mathrm{LLR}_i^{\mathrm{tmp}}$ is deinterleaved by $\Pi_{i,1}^{-1}$ and further calculated by the LLR updating function $f_c(\cdot)$ for correlated sources [80]. Finally, $\Pi_{i,1}$ interleaves the output of $f_c(\cdot)$ again to provide LLR_i^a. The structure of extrinsic information exchanger is much simpler for $g(\cdot)$ being

MV. As illustrated in Fig. 3.9(b), $f_c(\cdot)$ directly updates all of the LLR_i^e after deinterleaving with the LLR_{H}^e, and then the outputs from $f_c(\cdot)$ for agents are interleaved into LLR_i^a.

3.4.2 Simulation Results

Table 3.1. BASIC SETTINGS OF SIMULATION PARAMETERS

Parameter	Value	Parameter	Value
Number of blocks	1000	Block length	10000 bits
Distribution of X	$\text{Bern}(0.5)$	Maximum iteration time	30
Generator polynomial of CC	$([3,2]3)_8$	Rate of CC	$1/2$
Type of interleaver	random interleaver	Modulation method	BPSK

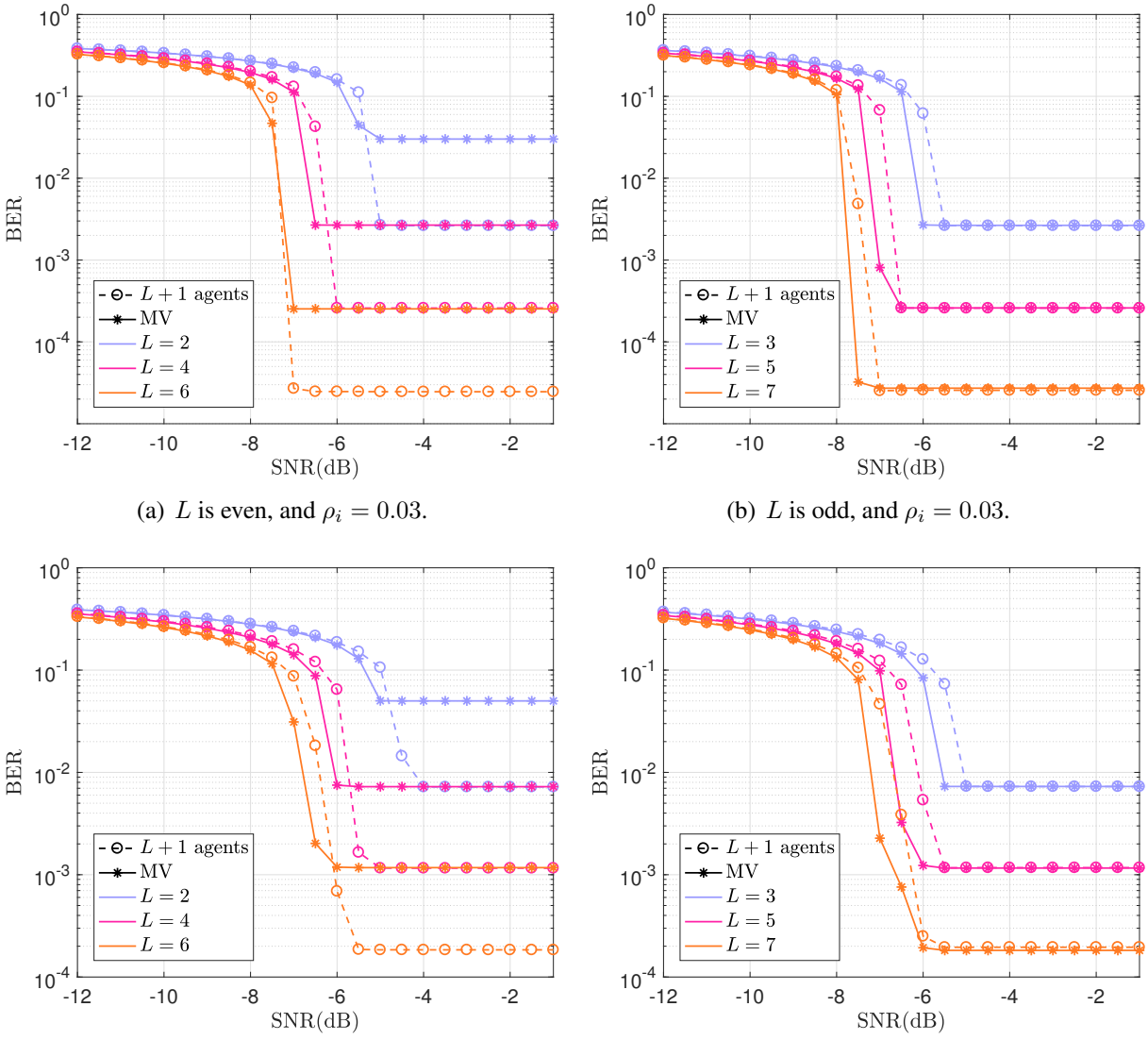


(a) L is even. (b) L is odd.

Fig. 3.10. Comparison of simulation results between different structures of helper, where $\rho_i = 0.03$.

Fig. 3.10 compares the simulation results between $g(\cdot)$ being modulo-2 sum (M2S) and MV, where the basic parameter settings are listed in Table 3.1. Since the MV decision rule is implemented in simulations, we use the limit derived from Poisson binomial as the theoretical bound. Clearly, the trade-off of rate-distortion in simulations perfectly matches that in theoretical analysis, i.e., a helper can shift the turbo cliff to left but cannot reduce the BER floor. It should be highlighted that shifting SNR to left also means eliminating distortions for low SNR level. Moreover, the helper with $g(\cdot)$ being MV can obviously reduce the SNR threshold more than the helper which generates its information by modulo-2 sum. The reason for the difference of performance is that the distortion is included in the helper sequence with modulo-2 sum operation if there is only one check node. If one of the LLR_i is with the opposite sign, all of the other $(L-1)$ LLRs will get negative helper information. However, such negative helper information cannot be reversed again as in the LDPC codes, because no other check nodes exist in the system with only

one helper. Hence, if the SNR is in an extremely low level, the helper with $g(\cdot)$ being modulo-2 sum will lose its effect due to the large distortion of X_i. This problem also results in the reduction of helper efficiency for large L, i.e., it is difficult to shift the turbo cliff by further increasing L. Therefore, the system with only one helper for $g(\cdot)$ being modulo-2 sum cannot obtain enough gains as the LDPC codes with a lot of check nodes. Nevertheless, the helper with its information generated by MV still can obviously reduce the SNR threshold with larger L. Because not only can the helper information generated by MV preserve large enough mutual information among $X_{\mathcal{L}}$ for lower error-corrupted probability ρ_i, but the distortions occurring at the small part of nodes are also not dominant when exchanging extrinsic information. Besides, the simulation results for both structures of helper can achieve the bound derived from Poisson binomial process. Hence, we can draw a conclusion that it is convincible and effective to predict the trade-off of rate-distortion for a practical system by applying the theoretical results.



(a) L is even, and $\rho_i = 0.03$.

(b) L is odd, and $\rho_i = 0.03$.

(c) L is even, and $\rho_i = 0.05$.

(d) L is odd, and $\rho_i = 0.05$.

Fig. 3.11. Comparison between one additional agent and a helper.

Finally, we make a comparison between $(L+1)$ agents and a helper for diverse ρ_i as illustrated

in Fig. 3.11. Since MV shows a better performance than modulo-2 sum as the helper structure, we only plot the curves of MV for comparison. It is remarkable in Fig. 3.11(a) and Fig. 3.11(c) that the system with $(L+1)$ agents can achieve the BER floor of the case with $(L+2)$ agents, when L is an even number. For example, when $L = 2$, by adding one more agent, the system can achieve the same BER floor as $L = 4$. Because one additional agent provides extra information of X, and this avoids the draw case that equal number of "0" and "1" appear in the same bit of agent sequences when L is even. However, the system with a helper has a lower SNR threshold for arbitrary L, due to the larger mutual information between Y^n and $X_{\mathcal{L}}^n$ by MV. Especially when ρ_i becomes larger or L is odd, the gap of turbo cliff between $(L + 1)$ agents and a helper is very conspicuous, e.g., Fig. 3.11(c) and Fig. 3.11(b) have more obvious gap than the case of Fig. 3.11(a). We can also find that the system with $(L + 1)$ agents keeps the same BER floor as one-helper system when L is odd as shown in Fig. 3.11(b) and Fig. 3.11(d), even though one more agent provides extra mutual information about X. Consequently, except the additional implementation cost needed, it is better to add a helper than one more agent for a system with odd number of agents. If the wireless channels are good enough, one more agent can make the estimate of source more accurate for the system already having even number of agents. However, in an extremely noisy environment, i.e., before achieving BER floor and/or ρ_i is relatively large, the system performance can benefit more from a helper than adding an additional agent. For instance, in Fig. 3.11(c) with $L = 6$ and SNR less than -6 dB, the curve with a helper has lower BER than the curve with $(L + 1)$ agents. In addition, there is another noticeable phenomenon that the BER floor is not immediately achieved for relatively larger ρ_i and L after turbo cliff appears, e.g., $L = 7$ with a helper in Fig. 3.11(d). Because for the bits at the same time index, they need to obtain enough extrinsic information from each other, so as to decode correctly when SNR is extremely low. Hence, only the bits almost without corrupting errors can accumulate their extrinsic information and be correctly decoded. The SNR threshold decreases as the number of agents increases, while larger SNR is required for large ρ_i to correctly decode all bits at the same time index. Once the SNR threshold decreases to the level in which the bits with some corrupting errors decode fail, the BER is not able to achieve the BER floor as soon as the turbo cliff appears.

3.5 Summary

We have analyzed and evaluated the performance of the binary CEO problem with a helper. To begin with, we decompose the binary CEO problem with a helper into two sub problems as multiterminal source coding with a helper and final decision. Subsequently, we derive an outer bound on the achievable rate-distortion region for the multiterminal source coding step and the DPF for the final decision step. Based on the derived outer bound, a convex optimization problem is formulated to minimize the distortion of observations with given agent and helper link rates. By substituting the minimized distortions into DPF, we investigate the relationship between link rates and final distortion. Although there is an obvious gap between MV and optimal decision, they show the same tendency on the trade-off of rate-distortion. Finally, we have risen an encoding/decoding scheme and design a simulation for practical performance evaluation, so as to compare with the theoretical results and analyze the trade-off of rate-distortion for the binary CEO problem with a helper. Both the theoretical and simulation results indicate that a helper can reduce the SNR threshold, while the BER floor does not change. Moreover, a helper with its structure as MV has a better performance than an additional agent for the system with odd number of agents or in extremely noisy communication environment. These significant observations are extremely useful

for the design of practical systems.

CHAPTER 4

Lossy Source Coding with Multiple Helpers

Chapter 2 and Chapter 3 have verified that a helper can enhance the system performance. Therefore, it spontaneously arouses our curiosity about how much the performance can be promoted by adding more helpers. The major objective of this chapter is to investigate the rate-distortion performance for lossy source coding with many helpers. We derive an inner bound on the achievable rate-distortion region for general sources, and then calculate it for binary sources. The tendency of theoretical results is further confirmed in the performance evaluation by simulations.

4.1 Problem Statement

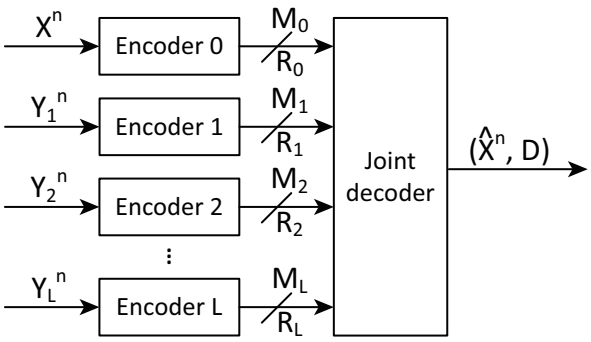


Fig. 4.1. The system model of lossy source coding with helpers.

As depicted in Fig. 4.1, there is a target source X and many helpers Y_i in the system. Although the number of helpers may be limited by space and channel resources in practical scenario, we do not constrain it in theoretical analysis. In total, there are $(L+1)$ DMS $(X, Y_1, \cdots, Y_L)$, with X and Y_i taking values from corresponding finite alphabets $\mathcal{X}$ and $\mathcal{Y}_i$ at each time slot, respectively. $x^n = \{x(t)\}_{t=1}^n$ and $y_i^n = \{y_i(t)\}_{t=1}^n$ indicate the i.i.d. sequences generated from the sources X and Y_i, respectively. Then, the sequences x^n and y_i^n are observed and transmitted to a common decoder, after compression by encoder 0 and encoder i, respectively. Due to some restrictions in practice, e.g., the deployment of encoders is distributed and located at different places, the observed sequences x^n and y_i^n have to be encoded into codewords separately. The encoders compress the sequences x^n and y_i^n at rates R_0 and R_i, respectively, by assigning an index to each sequence according to the following mapping rules:

$$\varphi_0 : \mathcal{X}^n \mapsto \mathcal{M}_0 = \{1, 2, \cdots, 2^{nR_0}\}, \tag{4.1}$$

$$\varphi_i : \mathcal{Y}_i^n \mapsto \mathcal{M}_i = \{1, 2, \cdots, 2^{nR_i}\}, \text{ for } i \in \mathcal{L}. \tag{4.2}$$

Without loss of generality, we assume that 2^{nR_0} and 2^{nR_i} are integer numbers.

Once all of the encoding outputs $\varphi_0(x^n), \varphi_1(y_1^n), \cdots, \varphi_L(y_L^n)$ are received, the joint decoder constructs the estimate $\hat{x}^n$ of x^n by jointly utilizing the received codewords in contrast to the

distributed compression at the encoders. The reconstruction process is expressed by the mapping as:

$$\psi : \mathcal{M}_0 \times \mathcal{M}_1 \times \cdots \times \mathcal{M}_L \mapsto \mathcal{X}^n. \tag{4.3}$$

Generally, distortion happens when the estimate $\hat{x}^n$ does not fully contain the information of x^n. The distortion measure $d : \mathcal{X} \times \mathcal{X} \mapsto [0, \infty)$ is defined to describe the degree of distortion between x and $\hat{x}$. For given distortion requirement D, the achievable rate-distortion region $\mathcal{R}(D)$, consisting of all achievable rate tuple of $(R_0, R_{\mathcal{L}})$, is defined as

$$\begin{aligned}\mathcal{R}(D) = \{(R_0, R_{\mathcal{L}}) : (R_0, R_{\mathcal{L}}) \text{ is admissible such that} \\ \lim_{n\to\infty} \mathrm{E}\left(d(x^n, \hat{x}^n)\right) \leq D + \epsilon, \text{ for any } \epsilon > 0\}.\end{aligned} \tag{4.4}$$

In the literature, Oohama [35] determined the achievable rate-distortion region for Gaussian sources. Wolf *et al.* [36] derived an inner bound on the rate region for lossless case with binary sources. The main theoretical results in this chapter are the inner bounds on the achievable rate-distortion region for general sources and binary sources.

4.2 Rate-Distortion Analysis

4.2.1 Inner Bound for General Sources

Initially, we derive an inner bound on the achievable rate-distortion region for lossy source coding with helpers.

Proposition 4.1: Let $(X, Y_{\mathcal{L}})$ be an $(L+1)$-component DMS and $d(x, \hat{x})$ be distortion measure. A rate tuple $(R_0, R_{\mathcal{L}})$ is achievable with distortion requirement D for distributed lossy source coding with more than one helper if

$$R_0 > I(X; U|V_{\mathcal{L}}), \tag{4.5}$$

$$\sum_{i\in\mathcal{S}} R_i > I(Y_{\mathcal{S}}; V_{\mathcal{S}}|V_{\mathcal{S}^c}), \tag{4.6}$$

for some conditional *pmf* $p(u|x) \cdot \prod_{i=1}^{L} p(v_i|y_i)$ and function $\hat{x}(u, v_{\mathcal{L}})$ such that $\mathrm{E}(d(X, \hat{X})) \leq D$, with $U \to X \to Y_i \to V_i$ and $V_i \to Y_i \to X \to Y_j \to V_j$ forming Markov chains for $i, j \in \mathcal{L}$ and $i \neq j$.

Proof of Proposition 4.1: We use an $(L + 1)$-dimension distributed compress-bin scheme for lossy source coding, and analyze the expected distortion of this scheme with respect to rate constraints. In the following, we assume that $\epsilon_1 < \epsilon_2 < \epsilon_3 < \epsilon$.

Codebook generation. Fix a conditional *pmf* $p(u|x) \cdot \prod_{i=1}^{L} p(v_i|y_i)$ and a function $\hat{x}(u, v_{\mathcal{L}})$ such that $\mathrm{E}(d(X, \hat{X})) \leq D/(1 + \epsilon)$. Let $\tilde{R}_0 \geq R_0$ and $\tilde{R}_i \geq R_i$ for $i \in \mathcal{L}$. Randomly and independently generate $2^{n\tilde{R}_0}$ sequences $u^n(k_0) \sim \prod_{t=1}^{n} p_U(u), k_0 \in \mathcal{K}_0 = \{1, 2, \cdots, 2^{n\tilde{R}_0}\}$. Similarly, for $i \in \mathcal{L}$, randomly and independently generate $2^{n\tilde{R}_i}$ sequences $v_i^n(k_i) \sim \prod_{t=1}^{n} p_{V_i}(v_i), k_i \in \mathcal{K}_i = \{1, 2, \cdots, 2^{n\tilde{R}_i}\}$. Then, partition the set of indices $k_0 \in \mathcal{K}_0$ into equal-size bins $\mathcal{B}_0(m_0) = \{(m_0 - 1) \cdot 2^{n(\tilde{R}_0 - R_0)} + 1, \cdots, m_0 \cdot 2^{n(\tilde{R}_0 - R_0)}\}$ for $m_0 \in \mathcal{M}_0$, and also partition the set of indices $k_i \in \mathcal{K}_i$ into equal-size bins $\mathcal{B}_i(m_i) = \{(m_i - 1) \cdot 2^{n(\tilde{R}_i - R_i)} + 1, \cdots, m_i \cdot 2^{n(\tilde{R}_i - R_i)}\}$ for $m_i \in \mathcal{M}_i, i \in \mathcal{L}$. This codebook structure is utilized in the encoders and the decoder.

Encoding. Upon observing x^n, encoder 0 finds an index $k_0 \in \mathcal{K}_0$ such that $(u^n(k_0), x^n) \in \mathcal{T}_{\epsilon_1}^{(n)}$. If there is more than one such index k_0, encoder 0 selects one of them uniformly at random. If there is no such index k_0, encoder 0 selects an index from $\mathcal{K}_0$ uniformly at random. Similarly, for $i \in \mathcal{L}$, encoder i finds an index $k_i \in \mathcal{K}_i$ such that $(v_i^n(k_i), y_i^n) \in \mathcal{T}_{\epsilon_1}^{(n)}$. If there is more than one such index k_i, encoder i selects one of them uniformly at random. If there is no such index k_i, encoder i selects an index from $\mathcal{K}_i$ uniformly at random. Then, encoder 0 and encoder i send the indices m_0 and m_i such that $k_0 \in \mathcal{B}_0(m_0)$ and $k_i \in \mathcal{B}_i(m_i)$, respectively.

Decoding. The decoder finds the unique index tuple $(\hat{k}_0, \hat{k}_{\mathcal{L}}) \in \mathcal{B}_0(m_0) \times \mathcal{B}_1(m_1) \times \cdots \times \mathcal{B}_L(m_L)$ such that $(u^n(\hat{k}_0), v_1^n(\hat{k}_1), \cdots, v_L^n(\hat{k}_L)) \in \mathcal{T}_\epsilon^{(n)}$. If there is such a unique index tuple $(\hat{k}_0, \hat{k}_{\mathcal{L}})$, the reconstruction is computed bit by bit as $\hat{x}_t(u_t(\hat{k}_0), v_{1,t}(\hat{k}_1), \cdots, v_{L,t}(\hat{k}_L))$; otherwise, $\hat{x}^n$ is set to arbitrary sequence in $\mathcal{X}^n$.

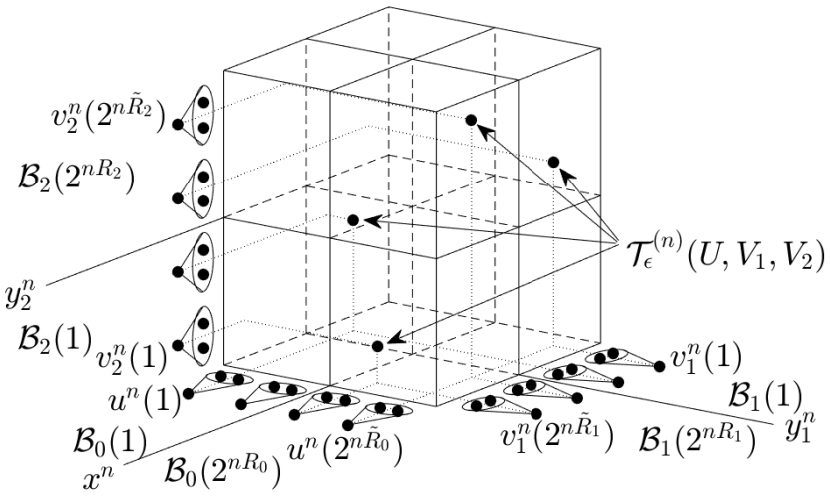


Fig. 4.2. An example of the distributed compress-bin scheme with $L = 2$.

An example of the distributed compress-bin scheme with $L = 2$ is depicted in Fig. 4.2. Now, we analyze the expected distortion of the distributed compress-bin scheme. Let $(K_0, K_{\mathcal{L}})$ denote the index tuple for the chosen $(U^n, V_{\mathcal{L}}^n)$ tuple, $(M_0, M_{\mathcal{L}})$ be the tuple of corresponding bin indices, and $(\hat{K}_0, \hat{K}_{\mathcal{L}})$ be the tuple of decoded indices. Define the "error" event

$$\mathcal{E} = \{(U^n(\hat{K}_0), V_1^n(\hat{K}_1), \cdots, V_L^n(\hat{K}_L), X^n, Y_1^n, \cdots, Y_L^n) \notin \mathcal{T}_\epsilon^{(n)}\}, \tag{4.7}$$

and consider the following sub-events:

$$\mathcal{E}_1 = \{(U^n(k_0), X^n) \notin \mathcal{T}_{\epsilon_1}^{(n)} \text{ for all } k_0 \in \mathcal{K}_0\}, \tag{4.8}$$

$$\mathcal{E}_2 = \{(V_i^n(k_i), Y_i^n) \notin \mathcal{T}_{\epsilon_1}^{(n)} \text{ for all } k_i \in \mathcal{K}_i, i \in \mathcal{L}\}, \tag{4.9}$$

$$\mathcal{E}_3 = \{(U^n(K_0), X^n, Y_1^n) \notin \mathcal{T}_{\epsilon_2}^{(n)}\}, \tag{4.10}$$

$$\mathcal{E}_4 = \{(U^n(K_0), X^n, V_1^n(K_1), Y_1^n) \notin \mathcal{T}_{\epsilon_3}^{(n)}\}, \tag{4.11}$$

$$\mathcal{E}_5 = \{(U^n(K_0), X^n, V_1^n(K_1), Y_1^n, \cdots, V_L^n(K_L), Y_L^n) \notin \mathcal{T}_\epsilon^{(n)}\}, \tag{4.12}$$

$$\mathcal{E}_6 = \{(V_1^n(\tilde{k}_1), \cdots, V_L^n(\tilde{k}_L)) \in \mathcal{T}_\epsilon^{(n)} \text{ for some } \tilde{k}_{\mathcal{L}} \in \mathcal{B}_1(M_1) \times \cdots \times \mathcal{B}_L(M_L), \tilde{k}_{\mathcal{L}} \neq K_{\mathcal{L}}\}, \tag{4.13}$$

$$\mathcal{E}_7 = \{(U^n(\tilde{k}_0), V_1^n(K_1), \cdots, V_L^n(K_L)) \in \mathcal{T}_\epsilon^{(n)} \text{ for some } \tilde{k}_0 \in \mathcal{B}_0(M_0), \tilde{k}_0 \neq K_0\}. \tag{4.14}$$

$\mathcal{E}_1$ and $\mathcal{E}_2$ represent encoding error events in encoder 0 and encoder i for $i \in \mathcal{L}$, respectively. $\mathcal{E}_5$ occurs if joint typicality decoding fails, with $\mathcal{E}_3$ and $\mathcal{E}_4$ being its sub-events. $\mathcal{E}_6$ and $\mathcal{E}_7$ mean that there are more than one decoding result, and hence a decoding error event occurs. Notice that the "error" event occurs only if $(U^n(K_0), V_1^n(K_1), \cdots, V_L^n(K_L), X^n, Y_1^n, \cdots, Y_L^n) \notin \mathcal{T}_\epsilon^{(n)}$ or $(\tilde{k}_0, \tilde{k}_\mathcal{L}) \neq (K_0, K_\mathcal{L})$. By the union of the events bound, we have

$$\begin{aligned}\Pr(\mathcal{E}) \leq \Pr(\mathcal{E}_1) + \Pr(\mathcal{E}_2) + \Pr(\mathcal{E}_1^c \cap \mathcal{E}_3) + \Pr(\mathcal{E}_3^c \cap \mathcal{E}_4) + \Pr(\mathcal{E}_4^c \cap \mathcal{E}_5) \\ + \Pr(\mathcal{E}_6) + \Pr(\mathcal{E}_7).\end{aligned} \tag{4.15}$$

We bound each term as follows. First, by the covering lemma [81], $\Pr(\mathcal{E}_1)$ tends to zero as $n \to \infty$ if

$$\tilde{R}_0 > I(X;U) + \delta(\epsilon_1), \tag{4.16}$$

and $\Pr(\mathcal{E}_2)$ tends to zero as $n \to \infty$ if

$$\tilde{R}_i > I(Y_i;V_i) + \delta(\epsilon_1). \tag{4.17}$$

Since $\mathcal{E}_1^c = \{(U^n(K_0), X^n) \in \mathcal{T}_{\epsilon_1}^{(n)}\}$, $Y_1^n | \{U^n(K_0) = u^n, X^n = x^n\} \sim \prod_{t=1}^n p_{Y_1|X}(y_{1,t}|x_t)$. By the conditional typicality lemma [81], $\Pr(\mathcal{E}_1^c \cap \mathcal{E}_3)$ approaches zero as $n \to \infty$.

To bound $\Pr(\mathcal{E}_3^c \cap \mathcal{E}_4)$, let $(u^n, x^n, y_1^n) \in \mathcal{T}_{\epsilon_2}^{(n)}(U, X, Y_1)$, and consider

$$\begin{aligned}&\Pr\{V_1^n(K_1) = v_1^n | U^n(K_0) = u^n, X^n = x^n, Y_1^n = y_1^n\} \\ &\quad = \Pr\{V_1^n(K_1) = v_1^n | Y_1^n = y_1^n\} \\ &\quad = p(v_1^n | y_1^n).\end{aligned} \tag{4.18}$$

First, notice that by the covering lemma, $\Pr\{V_1^n(K_1) \in \mathcal{T}_{\epsilon_2}^{(n)}(V_1|y_1^n) | Y_1^n = y_1^n\}$ converges to 1 as $n \to \infty$, i.e., $p(v_1^n|y_1^n)$ satisfies the first condition of the Markov lemma [81]. Then, similar to the proof of the Berger-Tung inner bound, shown in *Lemma 12.3* in [81], $p(v_1^n|y_1^n)$ also satisfies the second condition of the Markov lemma. Hence, according to the Markov lemma, we have

$$\lim_{n\to\infty} \Pr\{(u^n, x^n, y_1^n, V_1^n(K_1)) \in \mathcal{T}_{\epsilon_3}^{(n)} | U^n(K_0) = u^n, X^n = x^n, Y_1^n = y_1^n\} = 1, \tag{4.19}$$

if $(u^n, x^n, y_1^n) \in \mathcal{T}_{\epsilon_2}^{(n)}(U, X, Y_1)$ and $\epsilon_2 < \epsilon_3$ is sufficiently small. Therefore, $\Pr(\mathcal{E}_3^c \cap \mathcal{E}_4)$ tends to zero as $n \to \infty$. By recursively utilizing the similar derivation for bounding $\Pr(\mathcal{E}_3)$ and $\Pr(\mathcal{E}_3^c \cap \mathcal{E}_4)$, we can obtain that $\Pr(\mathcal{E}_4^c \cap \mathcal{E}_5)$ tends to zero as $n \to \infty$.

To bound $\Pr(\mathcal{E}_6)$, we introduce the following two lemmas:

Lemma 4.1: *Joint typicality lemma for multiple random variables.* Let $(V_\mathcal{S}, V_{\mathcal{S}^c}) \sim p(v_\mathcal{S}, v_{\mathcal{S}^c})$. If $\tilde{v}_i^n \sim \prod_{t=1}^n p_{V_i}(v_{i,t})$ for $i \in \mathcal{S}$, and $\tilde{v}_i^n$ is an arbitrary random sequence for $i \in \mathcal{S}^c$, then

$$\begin{aligned}&\Pr\{(\tilde{V}_\mathcal{S}^n, \tilde{v}_{\mathcal{S}^c}^n) \in \mathcal{T}_\epsilon^{(n)}(V_\mathcal{S}, V_{\mathcal{S}^c})\} \\ &\quad \leq \mathrm{pow}\left(2, -n\Big[\sum_{j=2}^{|\mathcal{S}|} I(V_{S_1^{j-1}}; V_{S_j}) + I(V_\mathcal{S}; V_{\mathcal{S}^c}) - \delta(\epsilon)\Big]\right).\end{aligned} \tag{4.20}$$

Lemma 4.2: *Mutual packing lemma for multiple random variables.* Let $(V_\mathcal{S}, V_{\mathcal{S}^c}) \sim p(v_\mathcal{S}, v_{\mathcal{S}^c})$. For $i \in \mathcal{S}$, let $V_i^n(k_i) \sim \prod_{t=1}^n p_{V_i}(v_{i,t})$, $k_i \in \mathcal{K}_i = \{1, 2, \cdots, 2^{nr_i}\}$. For $i \in \mathcal{S}^c$,

let $\tilde{V}_i^n$ be an arbitrarily distributed random sequence. Assume that $(V_i^n(k_i) : i \in \mathcal{S}, k_i \in \mathcal{K}_i)$ and $(\tilde{V}_i^n : i \in \mathcal{S}^c)$ are independent of each other. Then, $\delta(\epsilon)$ exists that tends to zero as $\epsilon \to 0$ such that

$$\lim_{n\to\infty} \Pr\{(V_{S_1}^n(k_{S_1}), \cdots, V_{S_{|\mathcal{S}|}}^n(k_{S_{|\mathcal{S}|}}), \tilde{V}_{\mathcal{S}^c}^n) \in \mathcal{T}_\epsilon^{(n)} \text{ for some } k_i \in \mathcal{K}_i, i \in \mathcal{S}\} = 0, \tag{4.21}$$

if

$$\sum_{i\in\mathcal{S}} r_i < \sum_{j=2}^{|\mathcal{S}|} I(V_{S_1^{j-1}}; V_{S_j}) + I(V_{\mathcal{S}}; V_{\mathcal{S}^c}) - \delta(\epsilon). \tag{4.22}$$

The proofs of **Lemma 4.1** and **Lemma 4.2** are provided in Appendix B and Appendix C, respectively.

If $\tilde{R}_i = R_i$ for $i \in \mathcal{S}^c$, notice that (4.17) becomes

$$R_i > I(Y_i; V_i) + \delta(\epsilon_1), \tag{4.23}$$

and hence R_i is already large enough for link i. Moreover, since $\tilde{R}_i - R_i = 0$, there is only one index in $\mathcal{B}_i$ for $i \in \mathcal{S}^c$. Hence, $\tilde{k}_i = K_i$ for $\tilde{k}_i \in \mathcal{B}_i$. Then, $\mathcal{E}_6$ can be simplified as

$$\begin{aligned}\mathcal{E}_6 = \{&(V_{S_1}^n(\tilde{k}_{S_1}), \cdots, V_{S_{|\mathcal{S}|}}^n(\tilde{k}_{S_{|\mathcal{S}|}}), V_{S_1^c}^n(K_{S_1^c}), \cdots, V_{S_{|\mathcal{S}^c|}^c}^n(K_{S_{|\mathcal{S}^c|}^c})) \in \mathcal{T}_\epsilon^{(n)} \\ &\text{for some } \tilde{k}_{\mathcal{S}} \neq K_{\mathcal{S}}, \tilde{k}_{\mathcal{S}} \in \mathcal{B}_{S_1}(M_{S_1}) \times \cdots \times \mathcal{B}_{S_{|\mathcal{S}|}}(M_{S_{|\mathcal{S}|}})\}.\end{aligned} \tag{4.24}$$

Following a similar argument as *Lemma 11.1* in [81] in the proof of the Wyner-Ziv theorem, we have

$$\begin{aligned}\Pr(\mathcal{E}_6) \le \Pr\{&(V_{S_1}^n(\tilde{k}_{S_1}), \cdots, V_{S_{|\mathcal{S}|}}^n(\tilde{k}_{S_{|\mathcal{S}|}}), V_{S_1^c}^n(K_{S_1^c}), \cdots, V_{S_{|\mathcal{S}^c|}^c}^n(K_{S_{|\mathcal{S}^c|}^c})) \in \mathcal{T}_\epsilon^{(n)} \\ &\text{for some } \tilde{k}_{\mathcal{S}} \in \mathcal{B}_{S_1}(1) \times \cdots \times \mathcal{B}_{S_{|\mathcal{S}|}}(1)\},\end{aligned} \tag{4.25}$$

$$\Pr(\mathcal{E}_7) \le \Pr\{(U^n(\tilde{k}_0), V_1^n(K_1), \cdots, V_L^n(K_L)) \in \mathcal{T}_\epsilon^{(n)} \text{ for some } \tilde{k}_0 \in \mathcal{B}_0(1)\}. \tag{4.26}$$

According to **Lemma 4.2** and the packing lemma [81], $\Pr(\mathcal{E}_6)$ and $\Pr(\mathcal{E}_7)$ tend to zero as $n \to \infty$, respectively, if

$$\sum_{i\in\mathcal{S}} (\tilde{R}_i - R_i) < \sum_{j=2}^{|\mathcal{S}|} I(V_{S_1^{j-1}}; V_{S_j}) + I(V_{\mathcal{S}}; V_{\mathcal{S}^c}) - \delta(\epsilon), \tag{4.27}$$

$$\tilde{R}_0 - R_0 < I(U; V_{\mathcal{L}}) - \delta(\epsilon). \tag{4.28}$$

By combining (4.16), (4.17), (4.27) and (4.28), we have shown that $\Pr(\mathcal{E})$ tends to zero as $n \to \infty$ if

$$R_0 > I(X; U) + \delta(\epsilon_1) - I(U; V_{\mathcal{L}}) + \delta(\epsilon), \tag{4.29}$$

$$\sum_{i\in\mathcal{S}} R_i > \sum_{i\in\mathcal{S}} [I(Y_i; V_i) + \delta(\epsilon_1)] - \sum_{j=2}^{|\mathcal{S}|} I(V_{S_1^{j-1}}; V_{S_j}) - I(V_{\mathcal{S}}; V_{\mathcal{S}^c}) + \delta(\epsilon). \tag{4.30}$$

We can further calculate (4.29) as

$$\begin{aligned} R_0 &> I(X; U) + \delta(\epsilon_1) - I(U; V_{\mathcal{L}}) + \delta(\epsilon), \\ &= I(X, V_{\mathcal{L}}; U) - I(U; V_{\mathcal{L}}) + \delta'(\epsilon) \qquad (4.31) \\ &= I(X; U|V_{\mathcal{L}}) + \delta'(\epsilon), \qquad (4.32)\end{aligned}$$

where (4.31) follows since $V_{\mathcal{L}} \to Y_{\mathcal{L}} \to X \to U$ forms a Markov chain, and $\delta'(\epsilon) = \delta(\epsilon_1) + \delta(\epsilon)$.

Then, (4.30) can further be reduced to:

$$\begin{aligned}\sum_{i\in\mathcal{S}} R_i &> \sum_{i\in\mathcal{S}} [I(Y_i;V_i) + \delta(\epsilon_1)] - \sum_{j=2}^{|\mathcal{S}|} I(V_{S_1^{j-1}};V_{S_j}) - I(V_{\mathcal{S}};V_{\mathcal{S}^c}) + \delta(\epsilon) \\ &= I(Y_{S_1};V_{S_1}) + I(Y_{S_2};V_{S_2}) - I(V_{S_1};V_{S_2}) + \sum_{j=3}^{|\mathcal{S}|}\Big[I(Y_{S_j};V_{S_j}) - I(V_{S_1^{j-1}};V_{S_j})\Big] \\ &\quad - I(V_{\mathcal{S}};V_{\mathcal{S}^c}) + \delta'(\epsilon), \end{aligned} \tag{4.33}$$

where $\delta'(\epsilon) = |\mathcal{S}| \cdot \delta(\epsilon_1) + \delta(\epsilon)$. Consider

$$\begin{aligned} & I(Y_{S_1^{j-1}};V_{S_1^{j-1}}) + I(Y_{S_j};V_{S_j}) - I(V_{S_1^{j-1}};V_{S_j}) \\ &= I(Y_{S_1^{j-1}};V_{S_1^{j-1}}) + I(Y_{S_j};V_{S_j}) - I(V_{S_1^{j-1}};V_{S_j}) + H(Y_{S_1^{j-1}}|V_{S_1^{j-1}},Y_{S_j}) \\ &\quad - H(Y_{S_1^{j-1}}|V_{S_1^{j-1}},Y_{S_j}) \\ &= I(Y_{S_1^{j-1}};V_{S_1^{j-1}}) + I(Y_{S_j};V_{S_j}) - I(V_{S_1^{j-1}};V_{S_j}) + H(Y_{S_1^{j-1}}|V_{S_1^{j-1}},Y_{S_j}) \\ &\quad - H(Y_{S_1^{j-1}}|V_{S_1^{j-1}},Y_{S_j},V_{S_j}) \qquad (4.34) \\ &= I(Y_{S_1^{j-1}};V_{S_1^{j-1}}) + I(Y_{S_j};V_{S_j}) - I(V_{S_1^{j-1}};V_{S_j}) + I(Y_{S_1^{j-1}};V_{S_j}|V_{S_1^{j-1}},Y_{S_j}) \\ &= I(Y_{S_1^{j-1}},Y_{S_j};V_{S_1^{j-1}}) + I(Y_{S_j},V_{S_1^{j-1}};V_{S_j}) - I(V_{S_1^{j-1}};V_{S_j}) \\ &\quad + I(Y_{S_1^{j-1}};V_{S_j}|V_{S_1^{j-1}},Y_{S_j}) \qquad (4.35) \\ &= I(Y_{S_1^{j}};V_{S_1^{j-1}}) + I(Y_{S_j};V_{S_j}|V_{S_1^{j-1}}) + I(Y_{S_1^{j-1}};V_{S_j}|V_{S_1^{j-1}},Y_{S_j}) \\ &= I(Y_{S_1^{j}};V_{S_1^{j-1}}) + I(Y_{S_1^{j-1}},Y_{S_j};V_{S_j}|V_{S_1^{j-1}}) \\ &= I(Y_{S_1^{j}};V_{S_1^{j-1}}) + I(Y_{S_1^{j}};V_{S_j}|V_{S_1^{j-1}}) \\ &= I(Y_{S_1^{j}};V_{S_1^{j-1}},V_{S_j}) \\ &= I(Y_{S_1^{j}};V_{S_1^{j}}), \qquad (4.36) \end{aligned}$$

where (4.34) follows the fact that V_{S_j} is a function of Y_{S_j}, and (4.35) follows that $V_{S_1^{j-1}} \to Y_{S_1^{j-1}} \to Y_{S_j} \to V_{S_j}$ forms a Markov chain. By substituting (4.36) into (4.33) for $j = 2$, and then recursively for $j = 3, \cdots, |\mathcal{S}|$, we have

$$\begin{aligned}\sum_{i\in\mathcal{S}} R_i &> I(Y_{S_1^2};V_{S_1^2}) + \sum_{j=3}^{|\mathcal{S}|}\Big[I(Y_{S_j};V_{S_j}) - I(V_{S_1^{j-1}};V_{S_j})\Big] - I(V_{\mathcal{S}};V_{\mathcal{S}^c}) + \delta'(\epsilon) \\ &= I(Y_{S_1^{|\mathcal{S}|}};V_{S_1^{|\mathcal{S}|}}) - I(V_{\mathcal{S}};V_{\mathcal{S}^c}) + \delta'(\epsilon) \\ &= I(Y_{\mathcal{S}};V_{\mathcal{S}}) - I(V_{\mathcal{S}};V_{\mathcal{S}^c}) + \delta'(\epsilon) \\ &= I(Y_{\mathcal{S}},V_{\mathcal{S}^c};V_{\mathcal{S}}) - I(V_{\mathcal{S}};V_{\mathcal{S}^c}) + \delta'(\epsilon) \qquad (4.37) \\ &= I(Y_{\mathcal{S}};V_{\mathcal{S}}|V_{\mathcal{S}^c}) + \delta'(\epsilon), \qquad (4.38) \end{aligned}$$

where (4.37) follows the fact that $V_{\mathcal{S}^c} \to Y_{\mathcal{S}^c} \to Y_{\mathcal{S}} \to V_{\mathcal{S}}$ forms a Markov chain.

Notice that $(U^n(K_0), V_1^n(K_1), \cdots, V_L^n(K_L), X^n, Y_1^n, \cdots, Y_L^n) \in \mathcal{T}_\epsilon^{(n)}$, when there is no "error". Therefore, by the law of total expectation and the typical average lemma [81], the

asymptotic distortion, averaged over the random codebook and encoding, is upper bounded as

$$\begin{aligned}\lim_{n\to\infty}\sup \mathrm{E}(d(X^n,\hat{X}^n)) &\le \lim_{n\to\infty}\sup\left[d_{\max}\cdot\Pr(\mathcal{E})+(1+\epsilon)\cdot\mathrm{E}(d(X,\hat{X}))\cdot\Pr(\mathcal{E}^c)\right]\\ &\le D,\end{aligned} \tag{4.39}$$

if the inequalities in (4.32) and (4.38) are satisfied. Finally, from the continuity of mutual information and taking $\epsilon \to 0$, we complete the proof of **Proposition 4.1**.

4.2.2 Inner Bound for Binary Sources

Here, we analyze the achievable rate-distortion region for binary sources. Consider a DSBS(ρ_i) (X, Y_i) with $X \sim \mathrm{Bern}(0.5)$ and $Y_i \sim \mathrm{Bern}(0.5)$ for $i \in \mathcal{L}$, where $\rho_i = \Pr\{x \neq y_i\}$, $\rho_i \in [0, \frac{1}{2}]$. The distortion measure is set as the Hamming distortion measure for binary sources.

Now, we calculate the constraints of the achievable rate-distortion region for DSBS. First, consider

$$\begin{aligned}R_0(D) &> I(X;U|V_{\mathcal{L}})\\ &= H(U|V_{\mathcal{L}}) - H(U|V_{\mathcal{L}},X)\\ &= H(U|V_{\mathcal{L}}) - H(U|X) &(4.40)\\ &= H(U,V_{\mathcal{L}}) - H(V_{\mathcal{L}}) - H_b(D), &(4.41)\end{aligned}$$

where (4.40) follows since $V_{\mathcal{L}} \to X \to U$ forms a Markov chain.

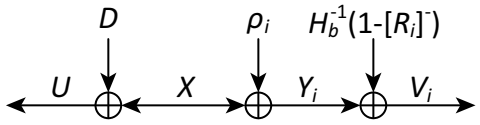


Fig. 4.3. The test channels for binary sources.

Since $V_i \to Y_i \to X \to Y_j \to V_j$ forms a Markov chain for $i \neq j$, i.e., Y_i are independent to each other if X is given, we can obtain the test channel shown in Fig. 4.3. Then, by **Definition 1.3**, we can calculate (4.41) as

$$R_0(D) > f_b(\{D, \alpha_{\mathcal{L}}\}) - f_b(\{\alpha_{\mathcal{L}}\}) - H_b(D), \tag{4.42}$$

where $\alpha_i = \rho_i * H_b^{-1}(1 - [R_i]^-)$. Notice that $D = 0.5$ if $R_0 = 0$ according to (4.42). However, it is obvious that by decoding only with the compressed side information $V_{\mathcal{L}}$, $\hat{X}$ still can achieve the distortion

$$\begin{aligned}D' &= H_b^{-1}[H(X|V_{\mathcal{L}})]\\ &= H_b^{-1}[H(X,V_{\mathcal{L}}) - H(V_{\mathcal{L}})]\\ &= H_b^{-1}[f_b(\{0,\alpha_{\mathcal{L}}\}) - f_b(\{\alpha_{\mathcal{L}}\})], &(4.43)\end{aligned}$$

where (4.43) holds since X can be regarded as the output of a BSC with itself as input and the crossover probability $\rho_0 = 0$. Therefore, the optimal performance can be achieved by time-sharing

between rate-distortion coding and zero-rate decoding only with the compressed side information. Consequently, we can obtain the rate-distortion function for DSBS, as

$$R_0(D) = \begin{cases} \omega(D), & \text{for } 0 \leq D \leq D_c, \\ (D - D') \cdot \omega'(D_c), & \text{for } D_c < D \leq D', \\ 0, & \text{for } D' < D, \end{cases} \tag{4.44}$$

where $\omega(D) = f_b(\{D, \alpha_{\mathcal{L}}\}) - f_b(\{\alpha_{\mathcal{L}}\}) - H_b(D)$ with $\omega'(D)$ being the derivative of $\omega(D)$, and D_c is the solution to the equation $\omega(D_c) = (D_c - D') \cdot \omega'(D_c)$.

Finally, we extend the above results of lossy source coding with helpers into joint source-channel coding based on Shannon's lossy source-channel separation theorem. Assuming that the channels are orthogonal, the link rates are constrained by:

$$R_0(D) \leq \Theta_0(\gamma_0) = \frac{C(\gamma_0)}{r_0}, \tag{4.45}$$

$$R_i \leq \Theta_i(\gamma_i) = \frac{C(\gamma_i)}{r_i}, \text{ for } i \in \mathcal{L}. \tag{4.46}$$

By combining (4.44) with (4.45) and (4.46), we have

$$\Theta_0(\gamma_0) \geq \begin{cases} \omega(D), & \text{for } 0 \leq D \leq D_c, \\ (D - D') \cdot \omega'(D_c), & \text{for } D_c < D \leq D', \\ 0, & \text{for } D' < D, \end{cases} \tag{4.47}$$

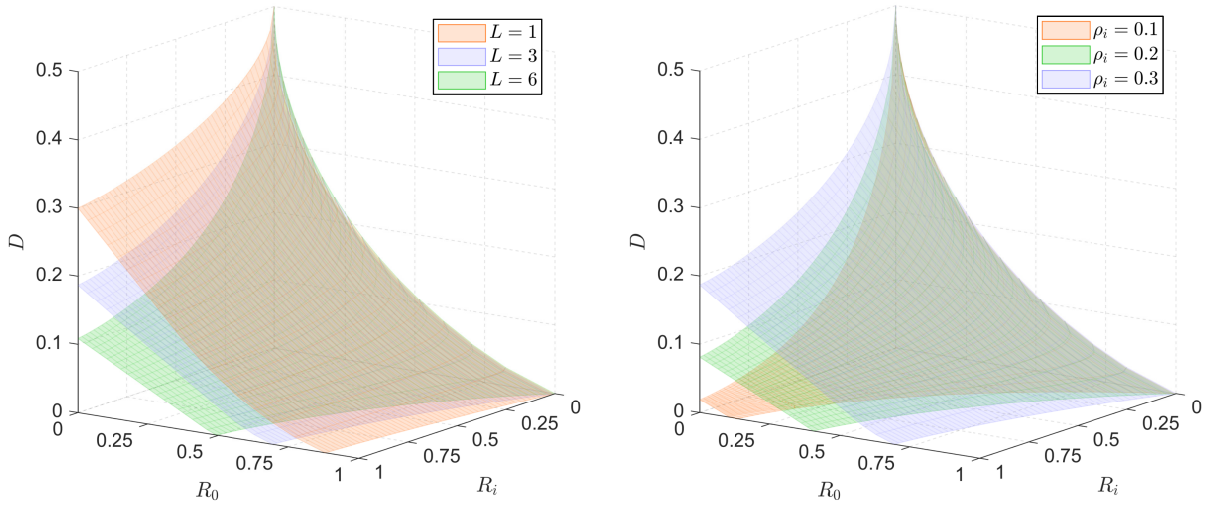
with $\alpha_i = \rho_i * H_b^{-1}(1 - [\Theta_i(\gamma_i)]^-)$ for calculating $\omega(D)$ and D'.

Remark: If a distortion requirement is given, we can evaluate whether the SNR of all links can satisfy the distortion requirement by (4.47). Conversely, if the SNR values of all links are given, we can utilize (4.47) to calculate the final distortion.

4.2.3 Numerical Results

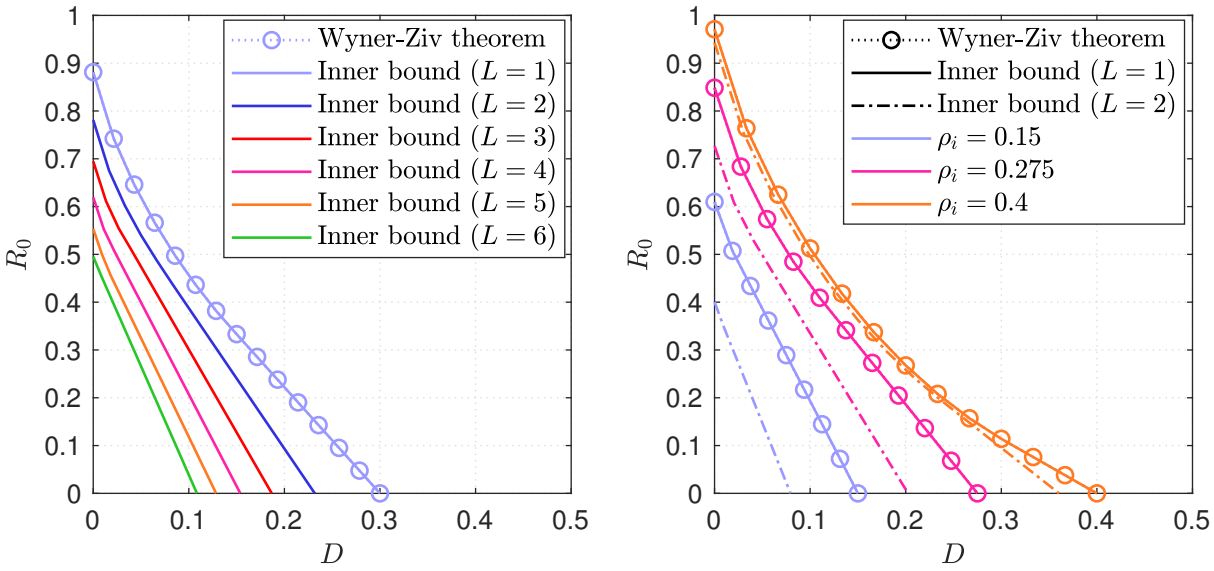
The relationship between the link rates and the final distortion is illustrated in Fig. 4.4, where we set all R_i at the same value, i.e., homogeneous helper links, so that the achievable rate-distortion region is able to be plotted within three dimensions. From the whole view, we can see that the distortion of X drops from 0.5 as R_0 and R_i gradually increase from 0. Moreover, the distortion decreases faster for larger L and smaller ρ_i in Fig. 4.4(a) and Fig. 4.4(b), respectively. It is also remarkable that all surfaces of the rate-distortion function intersect at one same curve in the R_0-D coordinate plane, i.e., $R_i = 0$. Obviously, the system model is equivalent to independent lossy source coding if $R_i = 0$, and hence $R_0(D)$ reduces to the classical rate-distortion function, which is not affected by the number of helpers and the correlations between sources. Another important phenomenon is that the distortion cannot be entirely eliminated to zero in the R_i-D coordinate plane. Therefore, the estimate $\hat{X}$ must be a lossy version of X when there is no information of X directly available for $\rho_i > 0$.

For given R_i, we can obtain the curves shown in Fig. 4.5 by projecting the surfaces of the rate-distortion function onto the R_0-D coordinate plane. Interestingly, the curves of the inner bound based on **Proposition 4.1** perfectly coincide with the curves of the Wyner-Ziv theorem for arbitrary ρ_i if there is only one helper without rate constraint. This phenomenon results from the fact that



(a) $\rho_i = 0.3$. (b) $L = 3$.

Fig. 4.4. The inner bound on the achievable rate-distortion region for homogeneous helper links.



(a) $\rho_i = 0.3$. (b) Diverse ρ_i.

Fig. 4.5. The inner bound on the rate-distortion function for given $R_i = 1$.

the theoretical model of the lossy source coding with helpers reduces to the Wyner-Ziv problem when $L = 1$ and $R_i = 1$. In addition, Fig. 4.5(a) demonstrates that the distortion can be reduced by introducing extra helpers; however, the gap between L and $(L + 1)$ becomes narrower along with the increment of helpers. Consequently, it is harder to obtain more gains when the number of helpers is already large enough. In Fig. 4.5(b), we can clearly observe that the curve shift to the left for small ρ_i, i.e., the distortion is smaller for more correlated sources. Meanwhile, the gap between L and $(L+1)$ is wider for the sources with high correlations, and hence it is more efficient to introduce extra helpers for more correlated sources.

4.3 Performance Evaluation

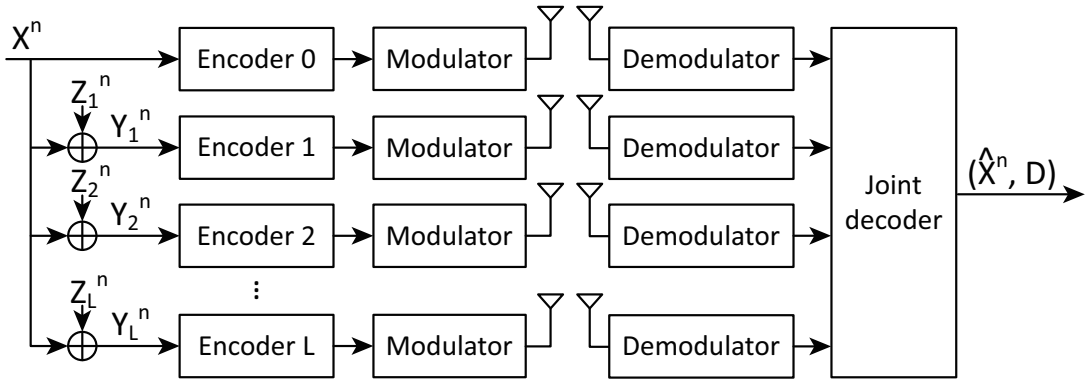
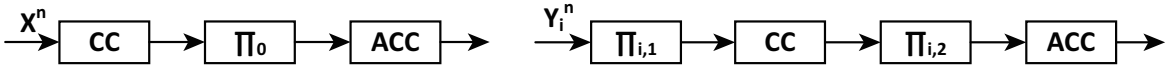


Fig. 4.6. A practical instance of lossy communications with helpers.

In this section, we start to evaluate the practical system performance for lossy communications with helpers as depicted in Fig. 4.6. There is one target sequence X^n and L helper sequences Y_i^n corrupted by Z_i^n with $Z_i \sim \mathrm{Bern}(\rho_i)$. To begin with, encoder 0 and encoder i encode their own sequence, respectively, and send the codeword through AWGN channels after modulation. The objective of this simulation is to compare the practical performance with the theoretical bound. Therefore, in order to make the final distortion as small as possible, the joint decoder starts to decode and produce estimate $\hat{X}^n$ after receiving and demodulating the signals in all the links. If the aim of a system is to satisfy a specified distortion requirement, the joint decoder may decrease the latency and complexity by decoding with fewer helper sequences. In other words, after receiving signals from some links, the joint decoder can first evaluate the SNR of received signals and the crossover probabilities between X and Y_i by the error probability estimation algorithm proposed in [55]. Then, it calculates the final distortion with already received signals by (4.47). If the expected final distortion is not larger than the given distortion requirement, the joint decoder starts decoding process; otherwise, it continues receiving the signals from the remaining links until the expected final distortion is small enough.



(a) The encoder of the target source. (b) The encoder of i-th helper.

Fig. 4.7. The structure of encoders.

Since the distributed compress-bin scheme utilized in the theoretical proof requires extremely huge memory to store the codebook, we design a practical coding scheme for simulation as shown in Fig. 4.7. An additional interleaver $\Pi_{i,1}$ is used to disperse noises into different bits before CC in the helper link.

As depicted in Fig. 4.8, ACC^{-1} decodes the inner code, and then CC^{-1} decodes the outer code after deinterleaving in Π^{-1}. Next, the extrinsic information is interleaved and subsequently exchanged to ACC^{-1} as the *a priori* information in local iteration. In the global iteration, LLR^p output from CC^{-1} is updated via an extrinsic information exchanger, which inputs LLR^e and

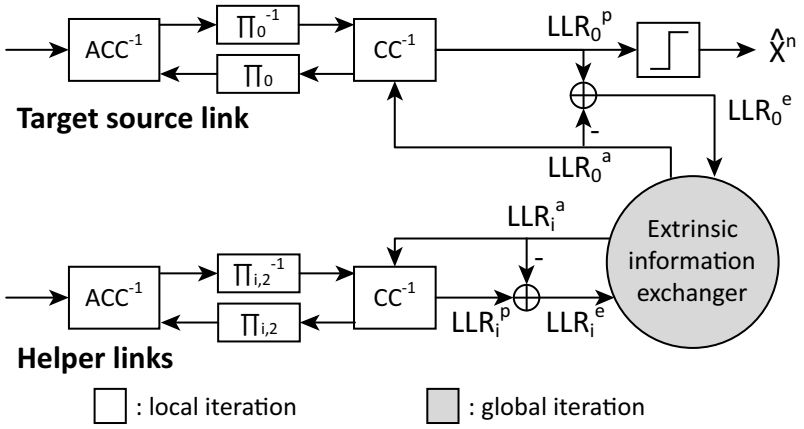
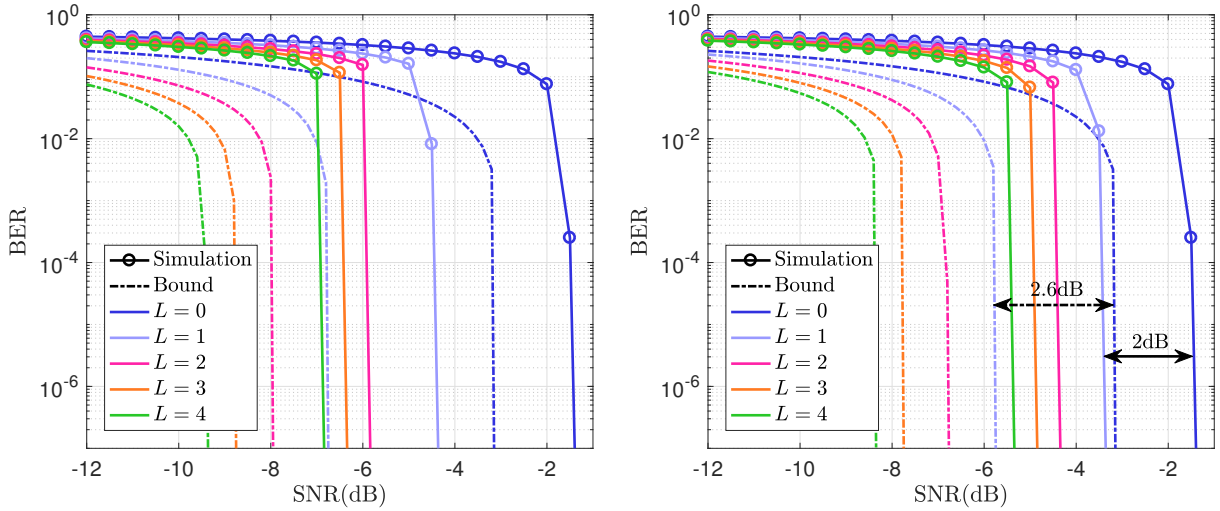


Fig. 4.8. The structure of the joint decoder.

outputs LLR^a. The extrinsic information exchanger calculates LLR^e by the LLR updating function $f_c(\cdot)$ for correlated sources [80].

Table 4.1. BASIC SETTINGS OF SIMULATION PARAMETERS

Parameter	Value	Parameter	Value
Number of blocks	1000	Block length	10000 bits
Generator polynomial of CC	$([3,2]3)_8$	Type of interleaver	random interleaver
Modulation method	BPSK	Maximum iteration time	30



(a) $\rho_i = 0.01$. (b) $\rho_i = 0.1$.

Fig. 4.9. Simulation results, where SNR is set at the same value for all links.

With the basic parameter settings listed in Table 4.1, the simulation results in Fig. 4.9 show the similar tendency as the curves of the theoretical bound. Clearly, the SNR threshold becomes lower as the number of helpers increases; however, the turbo cliff shifts to the left less rapidly

for the system with more helper. By the comparison between Fig. 4.9(a) and Fig. 4.9(b), we can find that the more independent the sources are, the higher SNR threshold is required. The performance gap between the theoretical and simulation results is due to the following two factors, i.e., the suboptimal channel coding scheme and incomplete utilization of joint typicality in the simulation. First, notice that there is an obvious gap between the theoretical and simulation results even for the case without any helper, because it is hard to achieve the Shannon limit by the relatively simple channel coding scheme used in the simulation. Besides the loss of performance due to channel coding, another key factor for the gap between the theoretical and simulation results is the incomplete utilization of joint typicality in the simulation. For instance, as shown in Fig. 4.9(b), there is a 2.6 dB gain for the theoretical results between no helper and one helper; however, only 2 dB gain can be achieved in the simulation for the same condition. This observation implies that the joint typicality is not completely utilized for joint decoding in the simulation as the distributed compress-bin scheme used in the theoretical analysis.

4.4 Summary

We have analyzed the performance of lossy communications with helpers. Initially, we start from the theoretical analysis of lossy source coding with helpers. After deriving an inner bound on the achievable rate-distortion region, we further calculate the rate-distortion function for binary sources. Subsequently, the results of lossy source coding with a helper is extended to joint source-channel coding based on Shannon's lossy source-channel separation theorem. The theoretical results perfectly match the Wyner-Ziv theorem, if there is only one helper and no rate limit on it. Finally, we present a distributed encoding and joint decoding scheme to evaluate the practical performance for an instance of lossy communications with helpers via a series of simulations. The comparison between the theoretical and simulation results inspires us that the system performance can be further improved, if there is a better coding scheme which can more efficiently utilizes the joint typicality of the coded sequences. Moreover, both the theoretical and simulation results indicate that the additional helper provides even smaller gains as the number of helpers becomes large.

CHAPTER 5

Lossy LF Relaying

From Chapter 2 to Chapter 4, we have established the theoretical framework for the performance analysis of helper-assisted lossy multiterminal source coding. This chapter applies the theoretical framework to a practical communication system, i.e., lossy communications with LF relaying, for the analyses of rate-distortion and outage probability. In traditional relaying systems, the destination requires high fidelity of the recovered information. However, when the channel condition is not good enough, the destination is probably not able to losslessly reconstruct the source sequence. Actually, lossy reconstructions within a distortion level are also acceptable as exemplified in IoT systems. With a specified acceptable distortion requirement, we can reduce the power consumption or transmission bandwidth by lossy compression than lossless communication. Thus, the trade-off between the link rates and the expected distortion degree is a very interesting topic in the big data era, especially for numerous electronic devices, of which power is supplied by small battery.

5.1 Problem Statement

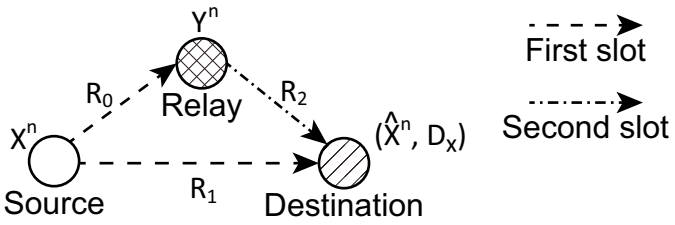


Fig. 5.1. The system model of lossy LF relaying.

The system model of lossy LF relaying is depicted in Fig. 5.1[1]. To make equations more concise, we denote variables for the S-R, S-D and R-D links with subscripts 0, 1 and 2, respectively. Due to the condition of wireless channels in practical systems, the S-R, S-D and R-D links have to satisfy the rates R_0, R_1 and R_2, respectively. From the aspect of whole system, the reconstruction of the source sequence will contain distortion, if the rate triplet (R_0, R_1, R_2), supported by the channel conditions in the S-R, S-D, and R-D links, respectively, does not satisfy the achievable rate region of lossless communications.

Notice that it is difficult to directly derive the explicit expression of the distortions resulting from the rate constraints on wireless channels. Shannon provided a way to equivalently determine the distortion corresponding to the channel capacity, i.e., compressing the data sequence by lossy

[1]It is noticed that the S-D link serves as a parallel alternative link at the second time slot. The system reliability can be further enhanced, if both the source and the relay transmit information by multiple access at the second time slot. However, for simplicity, this dissertation does not consider the case that the source transmits information twice, and it is left as the future work.

source coding to satisfy the channel capacity which is achievable by lossless channel coding[2]. Since the source coding and channel coding are separately performed, this idea is referred to as Shannon's lossy source-channel separation theorem.

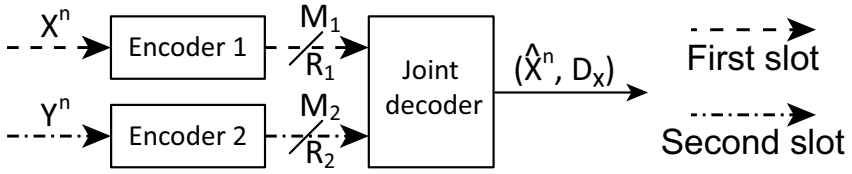


Fig. 5.2. The multiterminal source coding problem composed of the S-D and R-D links.

Likewise, to analyze the performance of lossy LF relaying, we can start from the fundamentals of multiterminal source coding. Since the relay receives data only from the source, the analysis of the S-R link can be easily handled by Shannon's lossy source-channel separation theorem for point-to-point communication. Regarding the remaining S-D and R-D links, we can first consider a multiterminal source coding problem illustrated in Fig. 5.2, where the encoder of X and the encoder of Y have to compress the sequences X^n and Y^n into codewords M_1 and M_2 at the rates R_1 and R_2, respectively.

By taking values from a finite alphabet $\mathcal{X}$ for each time index t, a common DMS X generates i.i.d. sequence $x^n = \{x(t)\}_{t=1}^n$. The encoder of X encodes the sequence x^n by mapping it into an index as:

$$\varphi_1 : \mathcal{X}^n \mapsto \mathcal{M}_1 = \{1, 2, \cdots, 2^{nR_1}\}. \tag{5.1}$$

Since the relay sequence $y^n = \{y(t)\}_{t=1}^n$ is an error-corrupted version of x^n, y^n is also an i.i.d. sequence with each bits belonging to a finite alphabet $\mathcal{Y}$. Similar to the encoder of X, the encoder of Y encodes the sequence y^n by assigning an index according to the mapping rule:

$$\varphi_2 : \mathcal{Y}^n \mapsto \mathcal{M}_2 = \{1, 2, \cdots, 2^{nR_2}\}. \tag{5.2}$$

The joint decoder in the destination node starts decoding after receiving the encoder outputs $\varphi_1(x^n)$ and $\varphi_2(y^n)$. Unlike the distributed compression in encoders, the joint decoder constructs the estimate $\hat{x}^n$ from the index $\varphi_1(x^n)$ with the assistance of the compressed side information $\varphi_2(y^n)$. The recovering progress is implemented by the following mapping as:

$$\psi : \mathcal{M}_1 \times \mathcal{M}_2 \mapsto \mathcal{X}^n. \tag{5.3}$$

Due to the possible deviation of x from $\hat{x}$, the distortion measure $d_X : \mathcal{X} \times \mathcal{X} \mapsto [0, d_{X,\max}]$ is defined to describe the distortion level between x and its estimate $\hat{x}$. With a desired distortion value D_X, the rate region $\mathcal{R}(D_X)$, consisting of all achievable rate pairs (R_1, R_2), is defined as

$$\begin{aligned}\mathcal{R}(D_X) = \{&(R_1, R_2) : (R_1, R_2) \text{ is admissible such that} \\ &\lim_{n\to\infty} \mathrm{E}[d_X(x^n, \hat{x}^n)] \le D_X + \epsilon, \text{ for any } \epsilon > 0\}.\end{aligned} \tag{5.4}$$

[2]By utilizing the duality between source coding and channel coding, the information loss due to channel conditions can be equivalently analyzed by lossy source coding, followed by lossless transmission through wireless channels. Eventually, compression is not performed by the encoder, but fading variation may reduce the rate supported by the channel. However, in theoretical distortion analysis, we can formulate the problem in this way for simplicity and without loss of generality.

For the point-to-point communication in the S-R link, the distortion between X and Y depends on R_0. For the cooperative communications in the S-D and R-D links, the distortion D_X depends on R_1, R_2 and the correlations between X and Y. Therefore, the final distortion D_X is eventually determined by R_0, R_1 and R_2. After determining the achievable rate-distortion region $\mathcal{R}(D_X)$ and the correlations between X and Y, we can obtain the relationship between the final distortion D_X and channel capacities of all three links by utilizing Shannon's lossy source-channel separation theorem.

Given a set of channel capacities for three links, we can calculate the expected minimum distortion based on the derived achievable rate-distortion region. If the expected distortion is larger than a specified distortion requirement, the communications are not reliable and outage event occurs. The channel capacities are random variables in fading channels, and hence the outage event randomly occurs with a probability, which is referred to as the outage probability. With a specified channel model, we can obtain the distributions of channel capacities and further calculate the outage probability, i.e., the probability that the instantaneous channel capacities cannot satisfy the distortion requirement.

In this chapter, the S-R, S-D and R-D links are assumed to suffer from independent block Rayleigh fading, with the channel gains as h_0, h_1 and h_2, respectively. Therefore, we can obtain the *pdf* of instantaneous SNR γ_i as

$$f(\gamma_i) = \frac{1}{\overline{\gamma}_i} \exp(-\frac{\gamma_i}{\overline{\gamma}_i}), \text{ for } i \in \{0, 1, 2\}. \tag{5.5}$$

For the purpose of simplicity, we assume that the CSI is only available at the receiver sides, and the effect of shadowing is not taken into account.

In the literature, Ahlswede and Korner [24] determined the achievable rate region of the lossless source coding with a helper. Zhou *et al.* [43] presented both the exact and approximate calculations of outage probability for lossless LF relaying system over independent block Rayleigh fading channels. The major results in this chapter are the inner bound on the achievable rate-distortion region and the upper bound of the outage probability for lossy LF relaying.

5.2 Rate-Distortion Analysis

To begin with, we determine an inner bound on the achievable rate-distortion region for lossy source coding with a helper as provided in the following proposition.

Proposition 5.1: Let (X, Y) be a 2-component DMS and $d_X(x, \hat{x})$ be a distortion measure. A rate pair (R_1, R_2) is achievable with distortion requirement D_X for lossy source coding of X with a helper Y if

$$R_1 > I(X; U|V), \tag{5.6}$$
$$R_2 > I(Y; V), \tag{5.7}$$

for some conditional *pmf* $p(u|x)p(v|y)$ and function $\hat{x}(u, v)$ such that $\mathrm{E}[d_X(X, \hat{X})] \leq D_X$, with $U \to X \to Y \to V$ forming a Markov chain.

The achievability for **Proposition 5.1** can be easily proved by applying **Proposition 4.1** with $L = 1$.

In order to draw a precise shape of the inner bound on the achievable rate-distortion region, we need a specified distribution of source. Since digital signals are quite often assumed in LF,

we start to use binary source as an instance in the following. Consider a binary source $X \sim \mathrm{Bern}(0.5)$, it is easy to find that Y, U and V also follow the $\mathrm{Bern}(0.5)$ distribution separately. For the purpose of deriving the relationship between R_0 and the distortion occurring in the S-R link, we can equivalently calculate the correlations between X and Y based on Shannon's lossy source-channel separation theorem. To satisfy the channel capacity by lossy source coding, we have

$$\begin{aligned} R_0 &> I(X;Y) \\ &= H(X) - H(X|Y) \\ &= 1 - H_b(\rho), \end{aligned} \tag{5.8}$$

where ρ represents the crossover probability between X and Y. Likewise, for the R-D link with the crossover probability ρ' between Y and V, we have

$$\begin{aligned} R_2 &> I(Y;V) \\ &= 1 - H_b(\rho'). \end{aligned} \tag{5.9}$$

From (5.6), for the S-D link, we have

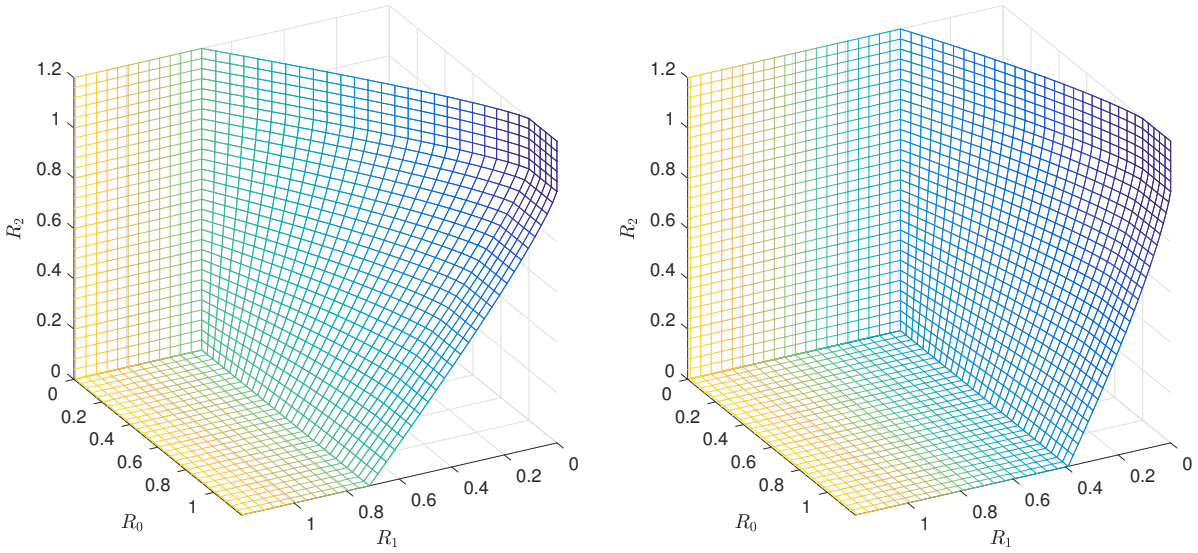
$$\begin{aligned} R_1 &> I(X;U|V) \\ &= H(U|V) - H(U|X,V) \\ &= H(U|V) - H(U|X) && (5.10) \\ &= H_b(\rho' * \rho * D_X) - H_b(D_X), && (5.11) \end{aligned}$$

where (5.10) and (5.11) follows since $V \to Y \to X \to U$ forms a Markov chain with the crossover probabilities ρ', ρ and D_X, respectively.

Consequently, we can obtain the inner bound on the achievable rate-distortion region with given distortion requirement as

$$\begin{cases} R_0 > 1 - H_b(\rho), \\ R_1 > H_b(\rho' * \rho * D_X) - H_b(D_X), \\ R_2 > 1 - H_b(\rho'). \end{cases} \tag{5.12}$$

If the desired distortion is given, we can illustrate the inner bound on the achievable rate-distortion region as in Fig. 5.3. It is remarkable that arbitrary R_0 and R_2 are admissible if R_1 is not less than $1 - H_b(D_X)$. Obviously, the compressed side information provided by the relay becomes redundant when R_1 is large enough for independent decoding. Hence, the distortion D_X can be easily satisfied by independent decoding for $R_1 \geq 1 - H_b(D_X)$ according to the lossy source coding theorem for point-to-point communication. Fig. 5.3(a) and Fig. 5.3(b) also demonstrate that the achievable rate-distortion region extends when the desired distortion becomes relatively large. Moreover, the part of surface is not flat for R_0, R_1 and R_2 all being less than 1. Because R_0 or R_2 needs more increment to compensate the decrease of R_1, due to the distortion propagating from the S-R link to the R-D link. Another interesting observation is that the achievable rate-distortion region is symmetric with respect to the plane of $R_0 = R_2$. Therefore, the S-R and R-D links have the same importance for system design, such as in determining power allocation and/or relay location.



(a) $D_X = 0.05, 1 - H_b(D_X) \approx 0.7136$. (b) $D_X = 0.15, 1 - H_b(D_X) \approx 0.3902$.

Fig. 5.3. The inner bound on the achievable rate-distortion region for specified D_X.

5.3 Outage Probability Analysis

In this section, we derive the upper bound of the outage probability for the lossy LF relaying, based on the inner bound of the achievable rate-distortion region derived in (5.12).

5.3.1 Outage Event of Lossy LF Relaying

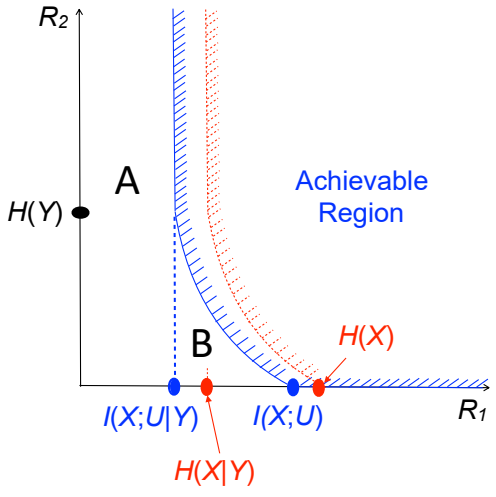


Fig. 5.4. The achievable rate-distortion region for X and Y, where the blue solid line and the red dashed line stand for lossy and lossless cases, respectively.

Here, we focus on the transmissions of the S-D and R-D links, which directly determine the occurrence of outage event, i.e., the destination cannot guarantee the reconstruction of X with the distortion smaller than D_X. For the influence of the S-R link, we treat the crossover probability ρ

between X and Y as a parameter determined by R_0. By this means, we can obtain achievable rate-distortion region for given R_0 as illustrated in Fig. 5.4, where the rate pair (R_1, R_2) is achievable if (5.6) and (5.7) are satisfied. To facilitate the outage calculation provided later in this chapter, the unachievable rate region is divided to two sub-regions, A and B, as indicated by

$$\begin{cases} \mathrm{A} \triangleq \{0 \le R_1 \le I(X;U|Y), 0 \le R_2\}, \\ \mathrm{B} \triangleq \{I(X;U|Y) \le R_1 \le I(X;U|V), 0 \le R_2 \le H(Y)\}. \end{cases} \tag{5.13}$$

To conveniently calculate $I(X;U|Y)$, we can utilize the result in (5.11) by letting $V = Y$ and $\rho' = 0$. Consequently, we have

$$\begin{cases} \mathrm{A} \triangleq \{0 \le R_1 \le H_b(\rho * D_X) - H_b(D_X), 0 \le R_2\}, \\ \mathrm{B} \triangleq \{H_b(\rho * D_X) - H_b(D_X) \le R_1 \le H_b(\rho' * \rho * D_X) - H_b(D_X), 0 \le R_2 \le 1\}. \end{cases} \tag{5.14}$$

Intuitively, the rate constraint in (5.11) indicates that:

- For $H(Y) \le R_2$, Y can successfully decoded with $Y = V$, i.e., $\rho' = 0$. The transmission with distortion D_X can be supported as long as $R_1 > H_b(0 * \rho * D_X) - H_b(D_X) = H_b(\rho * D_X) - H_b(D_X)$, which reduces to the Wyner-Ziv theorem.
- Even with $0 < R_2 < H(Y)$, Y can be partially recovered at the destination as V. V containing errors serves as the compressed side information for recovering X as long as $R_1 > H_b(\rho' * \rho * D_X) - H_b(D_X)$.
- In the case $R_2 = 0$ ($\rho' = 0.5$), i.e., the R-D link is broken down, the conditions in (5.11) become to $R_1 > H_b(0.5 * \rho * D_X) - H_b(D_X) = 1 - H_b(D_X)$, which reduces to the classical rate-distortion function.

Based on the discussion above, (5.11) can be rewritten explicitly as

$$R_1 > \begin{cases} H_b(\rho * D_X) - H_b(D_X), & \text{for } H(Y) \le R_2, \\ H_b(\rho' * \rho * D_X) - H_b(D_X), & \text{for } 0 < R_2 < H(Y), \\ 1 - H_b(D_X), & \text{for } R_2 = 0. \end{cases} \tag{5.15}$$

With the help of the compressed side information V, the outage event occurs when the rate pair (R_1, R_2) falls inside the unachievable regions, i.e., region A and B in Fig. 5.4. The outage probability P_{out} can be defined by taking average over all the transmissions, which results in

$$\begin{aligned} P_{\text{out}} =& \Pr\{(R_1, R_2) \in \mathrm{A} \cup \mathrm{B}\} \\ =& \Pr\{\rho = 0, (R_1, R_2) \in \mathrm{A} \cup \mathrm{B}\} + \Pr\{\rho \in (0, 0.5], (R_1, R_2) \in \mathrm{A} \cup \mathrm{B}\} \\ =& \Pr\{\rho = 0, (R_1, R_2) \in \mathrm{A}\} + \Pr\{\rho = 0, (R_1, R_2) \in \mathrm{B}\} \\ &+ \Pr\{\rho \in (0, 0.5], (R_1, R_2) \in \mathrm{A}\} + \Pr\{\rho \in (0, 0.5], (R_1, R_2) \in \mathrm{B}\} \\ =& \Pr\{\rho = 0, 0 \le R_1 \le H_b(\rho * D_X) - H_b(D_X), 0 \le R_2\} \\ &+ \Pr\{\rho = 0, H_b(\rho * D_X) - H_b(D_X) \le R_1 \le H_b(\rho' * \rho * D_X) - H_b(D_X), 0 \le R_2 \le 1\} \\ &+ \Pr\{0 < \rho \le 0.5, 0 \le R_1 \le H_b(\rho * D_X) - H_b(D_X), 0 \le R_2\} \\ &+ \Pr\{0 < \rho \le 0.5, H_b(\rho * D_X) - H_b(D_X) \le R_1 \le H_b(\rho' * \rho * D_X) - H_b(D_X), \\ &\qquad 0 \le R_2 \le 1\} \\ =& \Pr\{\rho = 0, 0 \le R_1 \le 0, 0 \le R_2\} \end{aligned}$$

$$
\begin{aligned}
&+\Pr\{\rho=0, 0\le R_1\le H_b(\rho'*D_X)-H_b(D_X), 0\le R_2\le 1\}\\
&+\Pr\{0<\rho\le 0.5, 0\le R_1\le H_b(\rho*D_X)-H_b(D_X), 0\le R_2\}\\
&+\Pr\{0<\rho\le 0.5, H_b(\rho*D_X)-H_b(D_X)\le R_1\le H_b(\rho'*\rho*D_X)-H_b(D_X),\\
&\qquad 0\le R_2\le 1\}\\
=&P_{1,\mathrm{A}}+P_{1,\mathrm{B}}+P_{2,\mathrm{A}}+P_{2,\mathrm{B}},
\end{aligned}
\tag{5.16}
$$

where $P_{1,\mathrm{A}}$, $P_{1,\mathrm{B}}$, $P_{2,\mathrm{A}}$ and $P_{2,\mathrm{B}}$ are defined for conciseness. The first subscript 1 and 2 represent the events $\rho=0$ and $\rho\in(0,0.5]$, while the second subscript A and B represent that the rate pair (R_1,R_2) falls inside the region A and B, respectively.

5.3.2 Outage Derivation

For calculating the outage probability, first we establish the relationship between γ_i and R_i for $i\in\{0,1,2\}$. Since orthogonal transmissions are assumed in the system, from the Shannon's lossy source-channel separation theorem, the relationship between the instantaneous channel SNR γ_i and its corresponding rate constraint R_i are given by

$$
R_i=\Theta_i(\gamma_i)=\begin{cases}\dfrac{C(\gamma_0)}{r_X}=\dfrac{E^n}{2r_X}\log_2\left(1+\dfrac{2\gamma_0}{E^n}\right), i=0,\\ \dfrac{C(\gamma_1)}{r_X}=\dfrac{E^n}{2r_X}\log_2\left(1+\dfrac{2\gamma_1}{E^n}\right), i=1,\\ \dfrac{C(\gamma_2)}{r_Y}=\dfrac{E^n}{2r_Y}\log_2\left(1+\dfrac{2\gamma_2}{E^n}\right), i=2,\end{cases}
\tag{5.17}
$$

where r_X and r_Y represent the channel coding rates for X^n and Y^n, respectively; E^n is the signaling dimensionality. By combining the results with (5.8), the crossover probability ρ between X and Y can be expressed with the function of γ_0 as

$$
\rho=H_b^{-1}\left[1-\Theta_0(\gamma_0)\right].
\tag{5.18}
$$

With the assumption that each link suffers from statistically independent block Rayleigh fading, each term of the outage probability expression in (5.16) can be further expressed as

$$
\begin{aligned}
P_{1,\mathrm{A}}=&\Pr\{\rho=0, 0\le R_1\le 0, 0\le R_2\}\\
=&\Pr\left\{\Theta_0^{-1}(1)\le\gamma_0, \Theta_1^{-1}(0)\le\gamma_1\le\Theta_1^{-1}(0), \Theta_2^{-1}(0)\le\gamma_2\right\}\\
=&\int_{\Theta_2^{-1}(0)}^{\infty}d\gamma_2\int_{\Theta_1^{-1}(0)}^{\Theta_1^{-1}(0)}d\gamma_1\int_{\Theta_0^{-1}(1)}^{\infty}f(\gamma_0)f(\gamma_1)f(\gamma_2)d\gamma_0,\\
=&0,
\end{aligned}
\tag{5.19}
$$

$$
\begin{aligned}
P_{1,\mathrm{B}}=&\Pr\{\rho=0, 0\le R_1\le H_b(\rho'*D_X)-H_b(D_X), 0\le R_2\le 1\}\\
=&\Pr\left\{\Theta_0^{-1}(1)\le\gamma_0, \Theta_1^{-1}(0)\le\gamma_1\le\Theta_1^{-1}[H_b(\xi(\gamma_2,D_X))-H_b(D_X)],\right.\\
&\left.\Theta_2^{-1}(0)\le\gamma_2\le\Theta_2^{-1}(1)\right\}\\
=&\int_{\Theta_2^{-1}(0)}^{\Theta_2^{-1}(1)}d\gamma_2\int_{\Theta_1^{-1}(0)}^{\Theta_1^{-1}[H_b(\xi(\gamma_2,D_X))-H_b(D_X)]}d\gamma_1\int_{\Theta_0^{-1}(1)}^{\infty}f(\gamma_0)f(\gamma_1)f(\gamma_2)d\gamma_0
\end{aligned}
$$

$$
\begin{aligned}
=&\frac{1}{\overline{\gamma}_2}\exp\left(-\frac{\Theta_0^{-1}(1)}{\overline{\gamma}_0}\right)\\
&\cdot\int_{\Theta_2^{-1}(0)}^{\Theta_2^{-1}(1)}\exp\left(-\frac{\gamma_2}{\overline{\gamma}_2}\right)\cdot\left[1-\exp\left(-\frac{\Theta_1^{-1}\left\{H_b[\xi(\gamma_2,D_X)]-H_b(D_X)\right\}}{\overline{\gamma}_1}\right)\right]d\gamma_2, \quad (5.20)
\end{aligned}
$$

$$
\begin{aligned}
P_{2,\mathrm{A}} =&\Pr\left\{0<\rho\le 0.5, 0\le R_1\le H_b(\rho * D_X)-H_b(D_X), 0\le R_2\right\}\\
=&\Pr\left\{\Theta_0^{-1}(0)\le\gamma_0<\Theta_0^{-1}(1), \Theta_1^{-1}(0)\le\gamma_1\le\Theta_1^{-1}\{H_b\left[\xi(\gamma_0,D_X)\right]-H_b(D_X)\},\right.\\
&\left.\Theta_2^{-1}(0)\le\gamma_2\right\}\\
=&\int_{\Theta_0^{-1}(0)}^{\Theta_0^{-1}(1)}d\gamma_0\int_{\Theta_1^{-1}(0)}^{\Theta_1^{-1}\{H_b[\xi(\gamma_0,D_X)]-H_b(D_X)\}}d\gamma_1\int_{\Theta_2^{-1}(0)}^{\infty}f(\gamma_2)f(\gamma_1)f(\gamma_0)d\gamma_2\\
=&\frac{1}{\overline{\gamma}_0}\exp\left(-\frac{\Theta_2^{-1}(0)}{\overline{\gamma}_2}\right)\\
&\cdot\int_{\Theta_0^{-1}(0)}^{\Theta_0^{-1}(1)}\exp\left(-\frac{\gamma_0}{\overline{\gamma}_0}\right)\cdot\left[1-\exp\left(-\frac{\Theta_1^{-1}\left\{H_b\left[\xi(\gamma_0,D_X)\right]-H_b(D_X)\right\}}{\overline{\gamma}_1}\right)\right]d\gamma_0, \quad (5.21)
\end{aligned}
$$

and

$$
\begin{aligned}
P_{2,\mathrm{B}} =&\Pr\left\{0<\rho\le 0.5, H_b(\rho * D_X)-H_b(D_X)\le R_1\le H_b(\rho' * \rho * D_X)-H_b(D_X),\right.\\
&\left.0\le R_2\le 1\right\}\\
=&\Pr\left\{\Theta_0^{-1}(0)\le\gamma_0<\Theta_0^{-1}(1),\right.\\
&\Theta_1^{-1}\{H_b\left[\xi(\gamma_0,D_X)\right]-H_b(D_X)\}\le\gamma_1\le\Theta_1^{-1}\{H_b\left[\zeta(\gamma_2,\gamma_0)*D_X\right]-H_b(D_X)\},\\
&\left.\Theta_2^{-1}(0)\le\gamma_2\le\Theta_2^{-1}(1)\right\}\\
=&\int_{\Theta_0^{-1}(0)}^{\Theta_0^{-1}(1)}d\gamma_0\int_{\Theta_2^{-1}(0)}^{\Theta_2^{-1}(1)}d\gamma_2\int_{\Theta_1^{-1}\{H_b[\xi(\gamma_0,D_X)]-H_b(D_X)\}}^{\Theta_1^{-1}\{H_b[\zeta(\gamma_2,\gamma_0)*D_X]-H_b(D_X)\}}f(\gamma_1)f(\gamma_2)f(\gamma_0)d\gamma_1\\
=&\frac{1}{\overline{\gamma}_0\overline{\gamma}_2}\int_{\Theta_0^{-1}(0)}^{\Theta_0^{-1}(1)}d\gamma_0\int_{\Theta_2^{-1}(0)}^{\Theta_2^{-1}(1)}\exp\left(-\frac{\gamma_0}{\overline{\gamma}_0}-\frac{\gamma_2}{\overline{\gamma}_2}\right)\\
&\cdot\left[\exp\left(-\frac{\Theta_1^{-1}\{H_b\left[\xi(\gamma_0,D_X)\right]-H_b(D_X)\}}{\overline{\gamma}_1}\right)\right.\\
&\left.-\exp\left(-\frac{\Theta_1^{-1}\{H_b\left[\zeta(\gamma_2,\gamma_0)*D_X\right]-H_b(D_X)\}}{\overline{\gamma}_1}\right)\right]d\gamma_2, \quad (5.22)
\end{aligned}
$$

where $\xi(\gamma_i,\tilde{\rho})=H_b^{-1}[1-\Theta_i(\gamma_i)]*\tilde{\rho}$ and $\zeta(\gamma_i,\gamma_j)=H_b^{-1}[1-\Theta_i(\gamma_i)]*H_b^{-1}[1-\Theta_j(\gamma_j)]$. Since there is not an explicit expression for the inverse of binary entropy function, it is hard to further calculate the integral and obtain a precise closed form. Instead, we utilize computer to calculate the numerical results for analyzing the outage probability.

5.3.3 Numerical Results

The upper bound of the outage probabilities for specified distortion requirement D_X is presented in Fig. 5.5, where average SNR is set at the same value for all three links. Clearly, the lossy LF relaying achieves lower outage probability with larger accepted distortion D_X. It should be noticed

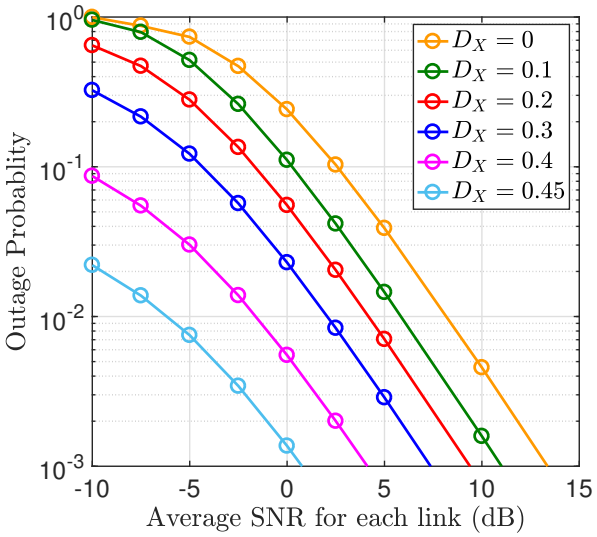


Fig. 5.5. The upper bound of the outage probability with different distortion level.

that the outage probability equals to zero when $D_X = 0.5$. This is because that $D_X = 0.5$ indicates any distortion can be accepted at the destination, and therefore, there will be no more outage in this case.

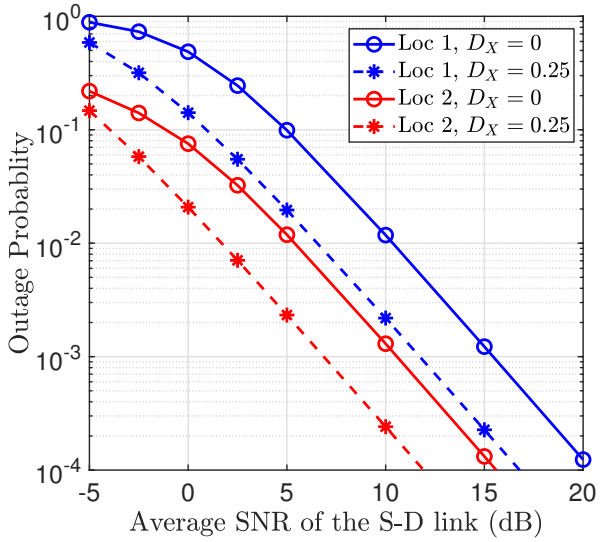


Fig. 5.6. The upper bound of the outage probability for different relay locations.

The outage curves of the lossy LF relaying are shown in Fig. 5.6 for two different relay location scenarios. With w_i for $i \in \{0, 1, 2\}$ denoting the distance of its corresponding link, we can empirically calculate $G_i = (w_1/w_i)^{3.52}$ by normalizing G_1 to the unity [82], where the exponent 3.52 represents the path loss factor which corresponds to the density of the obstacles of the area. We set $w_0 = w_1 = w_2$ in location scenario 1 (Loc 1), while $w_0 = 0.25w_1$ and $w_2 = 0.75w_1$ in location scenario 2 (Loc 2). In either the Loc 1 or Loc 2, lower outage probability can be achieved by allowing distortion at the destination. Moreover, since the distances of the S-R and R-D links in Loc 2 are both smaller than that in Loc 1, outage events occur with lower probability in Loc 2 than in Loc 1.

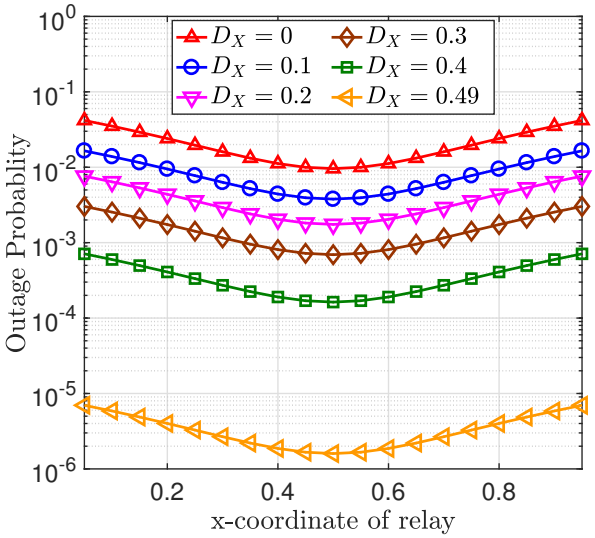


Fig. 5.7. The optimal relay positions of the lossy LF relaying, where $\overline{\gamma}_1 = 5$ dB.

Fig. 5.7 shows the impact of the relay location on the outage probability, with $\overline{\gamma}_1 = 5$ dB. The relay is located on the line between the source and the destination. It is found that the lowest outage probability can be achieved when the relay is located at the midpoint regardless the distortion level. It is also observed that the outage curves are symmetric with respect to the midpoint of the S-D link. This is because in the lossy LF relaying, the errors due to the S-R link can be corrected at the destination, and therefore, the midpoint ($w_0 = w_2$) is the optimal point where the contributions of the S-R and R-D links are balanced. This phenomenon indicates that the S-R and R-D links are of the same significance for system design, which perfectly matches with the finding in Fig. 5.3.

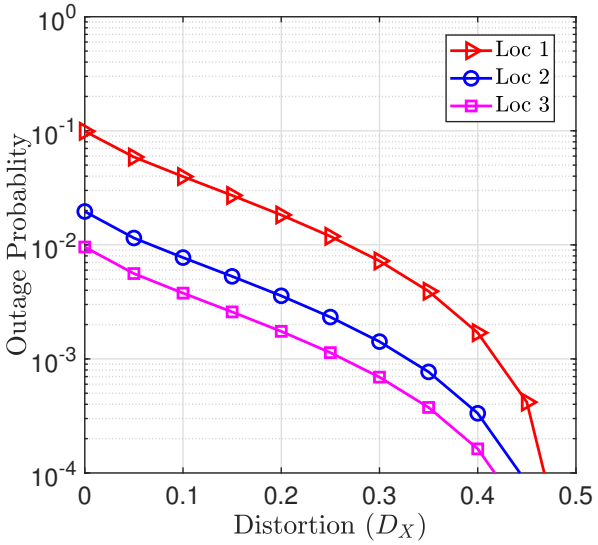


Fig. 5.8. The upper bound of the outage probability versus distortion level where $\overline{\gamma}_1 = 5$ (dB).

Fig. 5.8 shows the upper bound of the outage probability versus the distortion level D_X, with different relay location scenarios are considered. We set $w_0 = w_1 = w_2$ in Loc 1, $w_0 = 0.25w_1$ and $w_2 = 0.75w_1$ in Loc 2, and $w_0 = w_2 = 0.5w_1$ in Loc 3. It is observed that when the relay at the same location, outage probability decreases as the distortion level increases. It can also be

seen from the figure that, the outage performance in Loc 2 is superior than that obtained in Loc 1. This is because the quality of the S-R link in Loc 2 is better than that in Loc 1, resulting in lower probability of the S-R link transmission failure. From intuitive discussion for Fig. 5.7, we can understand the fact that the lossy LF relaying shows the best outage performance in Loc 3, since the relay is at the midpoint. Another interesting finding is that, the outage probability decreases almost linearly with D_X when the value of distortion is small (roughly less than 0.3); however, the outage probability decreases significantly when D_X is larger than 0.3. This observation can explain the reason why the gap between the curve with $D_X = 0.4$ and that with $D_X = 0.49$ suddenly becomes large in Fig. 5.7.

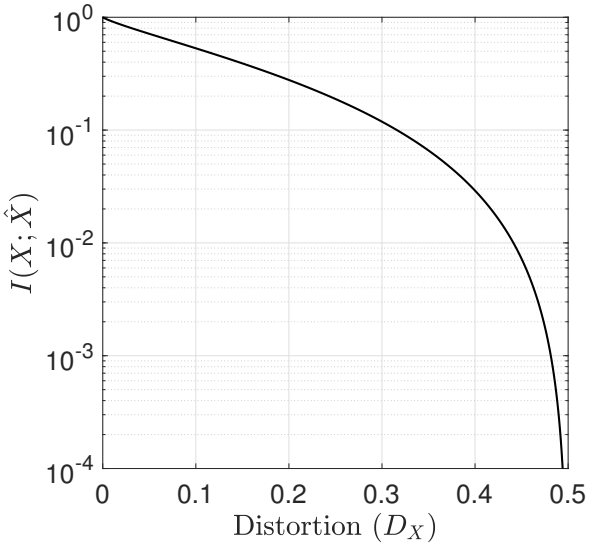


Fig. 5.9. The required mutual information between X and $\hat{X}$ for given distortion requirement.

The approximately linear tendency of the outage probability for $D_X \leq 0.3$ eventually results from the the required mutual information between X and $\hat{X}$. As illustrated in Fig. 5.9, $I(X;\hat{X})$ is almost linear in logarithmic scale as the distortion requirement D_X changes from 0 to 0.3. Since $I(X;\hat{X})$ is supported by the channel capacities, the channel capacities required for satisfying specified D_X decrease faster when $D_X > 0.3$. Moreover, the outage probability is also supported by the channel capacities, and hence its decay becomes sharper due to the faster reduction of required channel capacities for $D_X > 0.3$.

5.4 Performance Evaluation

5.4.1 Simulation Design

Here, we start to evaluate the system performance for a practical wireless communications network. As illustrated in Fig. 5.10, there are three nodes in the system containing source, relay and destination. In the first slot, the source node encodes sequence X^n by encoder 1 and broadcasts the modulated signal through Rayleigh channels. Then, the relay node decodes the received signal by a decoder after demodulation and makes hard decision into Y^n, while the destination node just stores the received signal. In the second slot, the relay node encodes Y^n by encoder 2 and subsequently sends the modulated signal to the destination node. As soon as the destination node

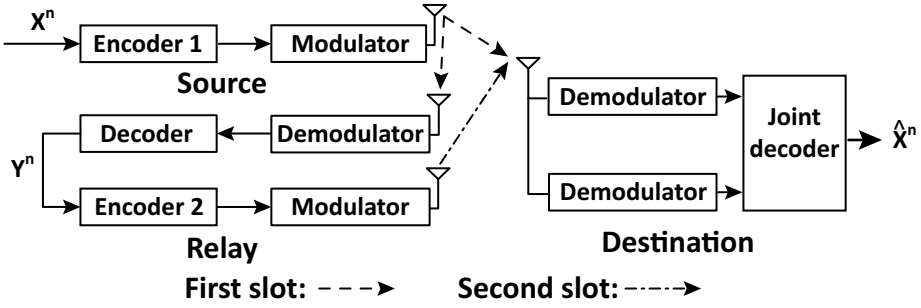
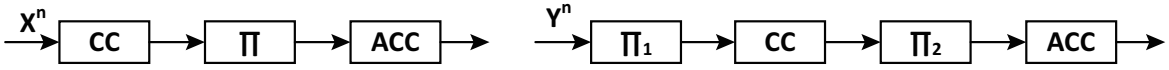


Fig. 5.10. The system model for simulation.

receives the signal from the relay node, it starts to jointly decodes the received signals and finally outputs the estimate $\hat{X}^n$.



(a) Encoder 1. (b) Encoder 2.

Fig. 5.11. The structure of encoders.

The structure of encoders is shown in Fig. 5.11. For the sake of iteration gains between X^n and Y^n in joint decoding, Y^n is interleaved by Π_1 at the beginning of encoding.

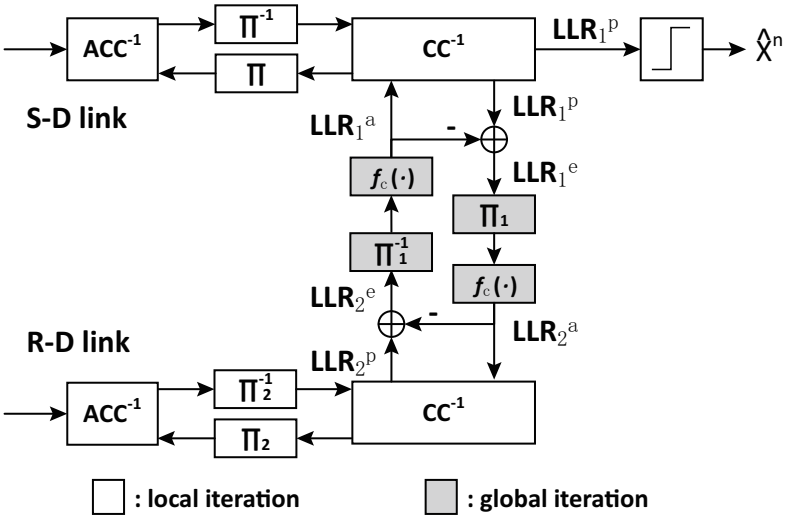


Fig. 5.12. The structure of joint decoder.

Fig. 5.12 depicts the structure of the joint decoder in the destination node. To begin with, the demodulated signal in each link is separately decoded by ACC^{-1} and CC^{-1}. In the local iteration, the extrinsic information is exchanged between ACC^{-1} and CC^{-1} via an interleaver Π and a deinterleaver Π^{-1}. After CC^{-1} outputs LLR^p at the end of local iteration, the joint decoder calculates LLR^e by subtracting LLR^a from LLR^p. When exchanging LLR between X^n and Y^n, the error probability of Y^n is first estimated by the algorithm proposed in [55], and then the *a priori* LLR is updated by the LLR updating function $f_c(\cdot)$ [80] with the extrinsic LLR as input. By this means, the relay information provides less extrinsic information if more errors exist in Y^n,

and hence X^n is insulated from the errors in Y^n. Due to the interleaving process on Y^n before CC, LLR_1^e should be interleaved by Π_1 and LLR_2^e should be deinterleaved by Π_1^{-1} when exchanging the extrinsic information in the global iteration. Finally, the estimate $\hat{X}^n$ is made by hard decision from LLR_1^p, if the maximum iteration time is exceeded or no more gains of the mutual information on LLR_1^p can be obtained in iterations.

5.4.2 Simulation Results

Table 5.1. BASIC PARAMETER SETTINGS

Parameter	Value	Parameter	Value
Frame length	10^4 bits	Type of interleaver	random interleaver
Number of frames	10^6	Generator polynomial of CC	$([3, 2]3)_8$
Rate of CC	$1/2$	Decoding algorithm for CC	BCJR algorithm [83]
Modulation method	BPSK	Maximum iteration time	30

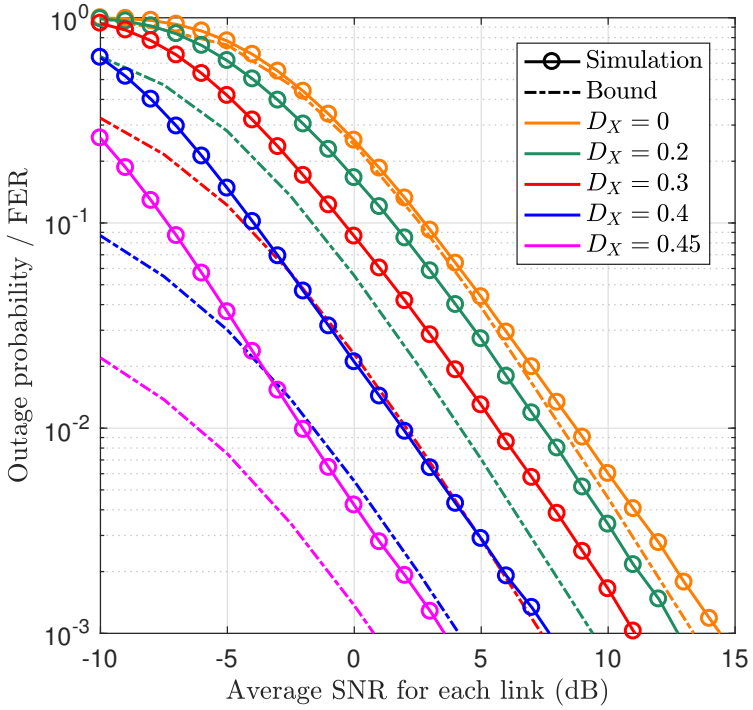


Fig. 5.13. Comparison between theoretical bound and simulation results.

The simulation result with parameter settings listed in Table 5.1 is shown in Fig. 5.13, which compares the theoretical outage probability and frame error rate (FER) in simulation. For simplicity, average SNR is set at the same value for the S-D, S-R and R-D links. In Fig. 5.13, it is clear that the simulation result has the same tendency and similar slope as the theoretical bound, even though there is an obvious gap between them. Moreover, the gap between the

simulation and theoretical results becomes larger as D_X increases. This phenomenon indicates that the practical scheme used in simulation is more efficient when the distortion requirement is more strict. There are two major factors which result in the loss of system performance. First, notice that with relatively simple channel coding scheme, it is hard to achieve Shannon limit, and hence there is also a gap between the FER in simulation and the theoretical outage probability of the network, as a whole. Another significant factor is that, the practical coding scheme used in simulation exchanges the extrinsic information in bit-wise; however, the distributed compress-bin scheme in the achievability proof jointly decodes in sequence-wise, to completely exploit the joint typicality. Therefore, the practical coding scheme cannot utilize the joint typicality as efficiently as the distributed compress-bin scheme, resulting in the loss of system performance.

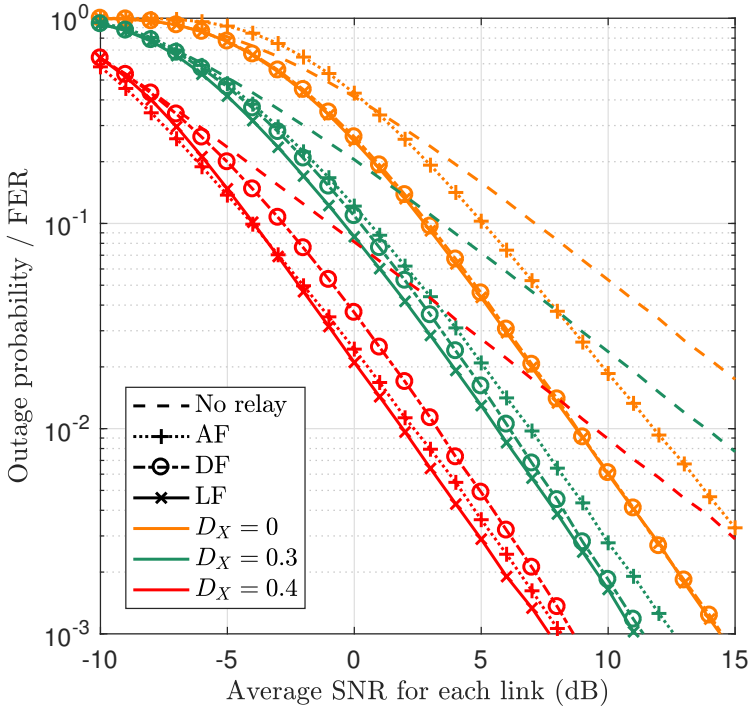


Fig. 5.14. Comparison among different forwarding schemes.

Fig. 5.14 compares the practical performance for diverse relaying schemes, including AF, DF, LF, and the case without relay[3]. Obviously, the curves with a relay have the same decay of the performance curve independently of the relaying scheme, while the slope of the curve without a relay is less steep compared to the curves with a relay. This observation demonstrates the diversity gains achieved by introducing a relay, although there are some gaps between different relaying schemes. It is found that AF has a worse performance than DF and LF for relatively small distortion requirement, because DF and LF can eliminate the errors from the S-R link, while AF amplifies the signals along with the noise. In addition, by encoding again at the relay, the data received from the S-D and R-D links are equivalent to distributed turbo codes, and hence DF and LF have coding gains while AF cannot. However, when the required distortion becomes very large, e.g., $D_X = 0.4$, AF has a better performance than DF. The reason is that large D_X requires even lower

[3]Although there are also many other forwarding schemes, such as ADF [84] and compress-and-forward (CF) [85], we only compare LF with the most fundamental AF and DF schemes for simplicity. In the case of lossless transmission, their performances are compared in [86]

SNR, which makes more errors exist in the decoding result at the relay; therefore, the DF relay discards the data sequences and stop forwarding more frequently. However, the system with an LF relay still has a lower outage probability than that with an AF relay, due to the utilization of correlations in the error-corrupted sequences.

5.5 Summary

We have analyzed the performance of lossy LF relaying, where distortion is allowed in the destination with the assistance of an LF relay. To begin with, we divided the system into two sub problems, i.e., lossy point-to-point communication for the S-R link, and the multiterminal source coding problem for the S-D and R-D links. The sub problem for the S-R link can be easily solved by Shannon's lossy source coding theorem and lossy source-channel separation theorem. Then, for the multiterminal source coding problem in the S-D and R-D links, we derive an inner bound on the achievable rate-distortion region. We further determine the relationship between final distortion and the rate constraints due to the channel condition, by applying Shannon's lossy source-channel separation theorem to the achievable rate-distortion region. Moreover, we analyze the outage probability for specified distortion requirements over block Rayleigh fading channels. Finally, we design a simulation system to evaluate the practical performance of FER. Comparing to the theoretical outage probability, we find that the tendency of simulation result matches with theoretical analysis. Especially for the case with strict distortion requirement, the FER in simulation is very close to the theoretical outage probability.

CHAPTER 6

Practical Coding Design for Lossy Compression

This chapter focus on coding design of lossy compression for practical use. Since the joint typicality coding scheme utilized in the theoretical achievability proof requires extremely huge memory for storing codebook, it is hard to be implemented in practical scenario. Instead, puncturing is frequently used in the design of various source and/or channel coding scheme due to its simplicity. Notice that puncturing can be also regarded as a lossy source coding scheme. If there is a coding scheme that has a better performance than puncturing, we can improve the performance by replacing puncturing with some better coding scheme.

6.1 Performance Analysis of Puncturing

Consider an n-bits binary sequence X^n with $X \sim \mathrm{Bern}(0.5)$, if X^n is punctured with rate R, the punctured $n(1-R)$ bits has 0.5 probability of error. Consequently, the expected distortion with puncturing is given by

$$D_{\mathrm{P}} = \frac{0.5 \cdot n(1-R)}{n} = \frac{1-R}{2}, \tag{6.1}$$

which is obviously linear to the rate R. Intuitively, lossy source coding can only achieve linear performance with linear algebra method, i.e., multiplying a generation matrix. Puncturing is equivalent to multiplying a diagonal matrix with the diagonal elements for the punctured bit being 0 and other diagonal elements being 1, and hence its performance is linear.

However, according to Shannon's lossy source coding theorem, the binary rate-distortion function for X^n is

$$R = 1 - H_b(D). \tag{6.2}$$

In the achievability proof of Shannon's lossy source coding theorem, 2^n sequences x^n are mapped to 2^{nR} sequences $\hat{x}^n$ such that $(x^n, \hat{x}^n)$ is jointly typical. This joint typicality coding requires extremely huge memory for storing codebook; however, the codebook with non-linear mapping can achieve optimal performance as $n \to \infty$. Therefore, a coding scheme containing non-linear process is the key to make performance closer to the theoretical limit. Inspired by the joint typicality coding scheme, we find a rate adaptive lossy source coding scheme, named HMV code, which can achieve the performance shown in Fig. 6.1. The HMV code can make a trade-off between efficiency and complexity.

6.2 Majority Voting Code

Consider a special case of joint typicality coding, i.e., $R = \frac{1}{n}$, and we have to map 2^n sequences to 2 sequence. It is not difficult to show that two sequences with all zeros and all ones are optimal, because they have maximum Hamming distance n with each other. Since the distance

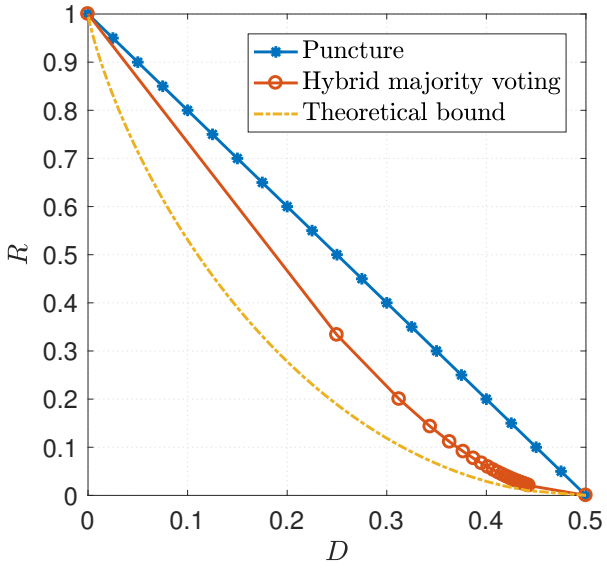


Fig. 6.1. The performance of hybrid majority voting code.

between all zeros/ones and other sequences are smaller than n, the distortion is smaller compared to using other codeword as the estimate of sequences. It is also obvious that in order to decrease the distortion, a sequence should be mapped to all ones if the sequence contains more 1 than 0; otherwise, the sequence should be mapped to all zeros. Then, we use only 1-bit codewords "0" and "1" to represent two sequences, i.e., all zeros and all ones, respectively, we have successfully compressed an n-bits sequence into 1-bit codeword. In summary, we compress an n-bits sequence into the codeword "1" if the number of ones is more than $\frac{n}{2}$; otherwise, we compress it into "0". By majority voting, we can simply perform the lossy compression described above, of which the performance becomes closer to the theoretical bound as n goes large.

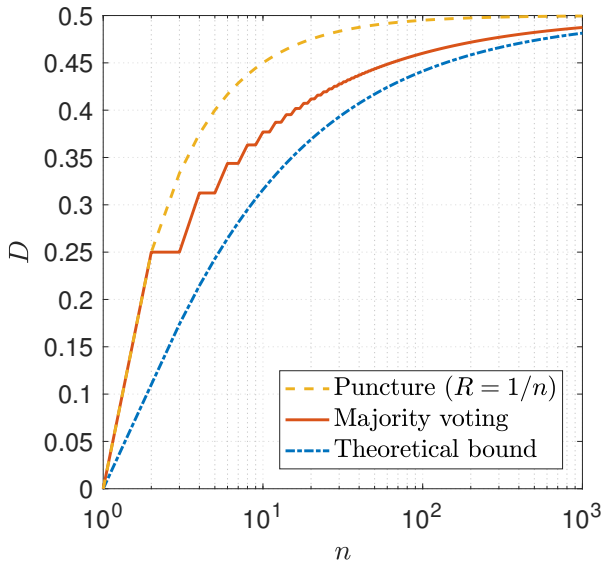


Fig. 6.2. The performance of majority voting code with diverse sequence length.

Now, we analyze the expected distortion of the MV code. Without loss of generality, we

assuming that we receive the 1-bit sequence "0". For an integer $l < \frac{n}{2}$, the number of sequence with l bits errors is $\binom{n}{l}$. For $l = \frac{n}{2}$ with even n, only half of the sequences are mapped to all zeros, and hence the number of sequence with l bits errors is $\frac{1}{2}\binom{n}{l}$. Since there are $2^{(n-1)}$ sequences encoded to "0" and all sequences are generated with the same probability, the expected distortion of the MV code is calculated as

$$D_{\text{MV}}(n) = \begin{cases} \dfrac{1}{2^{n-1}} \displaystyle\sum_{l=0}^{\frac{n-1}{2}} \frac{l}{n}\binom{n}{l}, & \text{for } n \text{ is odd,} \\ \dfrac{1}{2^{n-1}} \left(\displaystyle\sum_{l=0}^{\frac{n}{2}-1} \frac{l}{n}\binom{n}{l} + \frac{1}{4}\binom{n}{\frac{n}{2}} \right), & \text{for } n \text{ is even.} \end{cases} \tag{6.3}$$

The performance of the MV code is depicted in Fig. 6.2, which demonstrates the expected tendency that the distortion of MV code becomes close to the theoretical bound for sufficiently long sequence.

Then, we start lossy compression with arbitrary R based on MV code. Notice that there are nR bits in encoded sequence. The most simple way is compressing $k = n(1-R)+1$ bits into 1 bit and keeping the remaining $nR-1$ bits the same as origin. By this single-compression with MV code (SMV), we can satisfy the distortion

$$D_{\text{SMV}}(n,k) = \begin{cases} \dfrac{1}{2^{k-1}} \displaystyle\sum_{l=0}^{\frac{k-1}{2}} \frac{l}{n}\binom{k}{l}, & \text{for } k \text{ is odd,} \\ \dfrac{1}{2^{k-1}} \left(\displaystyle\sum_{l=0}^{\frac{k}{2}-1} \frac{l}{n}\binom{k}{l} + \frac{k}{4n}\binom{k}{\frac{k}{2}} \right), & \text{for } k \text{ is even.} \end{cases} \tag{6.4}$$

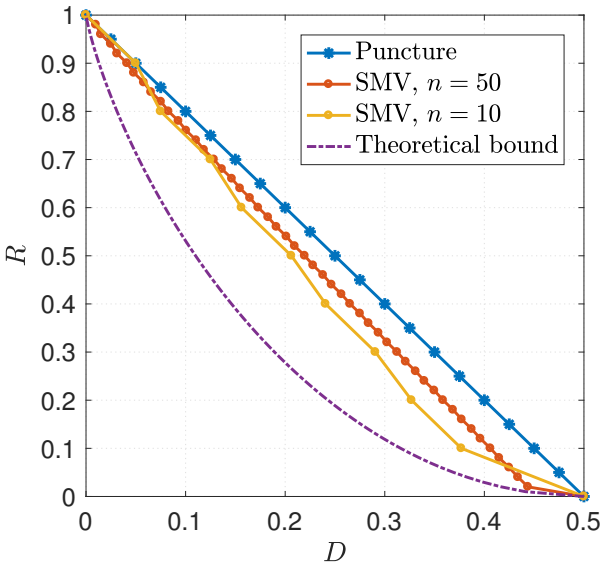


Fig. 6.3. The performance of single-compression with MV code.

The performance of SMV is shown in Fig. 6.3 for various values of n. It is sensible that the performance becomes even worse as n increases except the case with relatively small R. The

reason for this observation is that the remaining $nR - 1$ bits have not been exploited in lossy compression, and hence more bits are wasted when n becomes larger.

6.3 Hybrid Majority Voting Code

6.3.1 Encoding

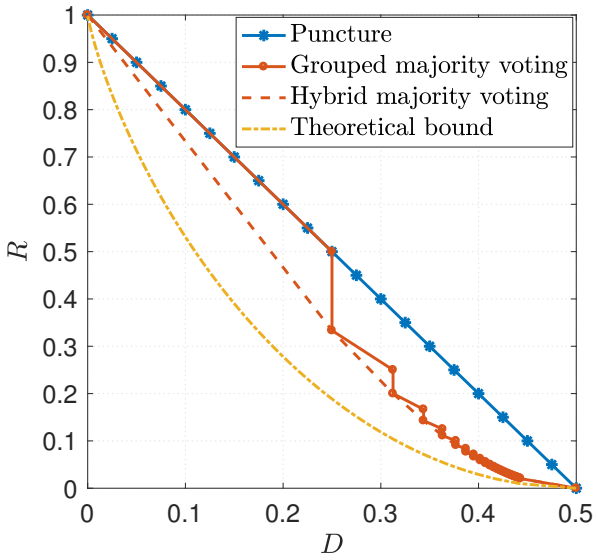


Fig. 6.4. The performance of grouped MV code and HMV code.

For the purpose of utilizing all remaining bits for lossy compression, a reasonable way is to divide the sequence into nR groups and then compress group by group. Assuming that the sequence can be divided into nR groups with equivalent size $s = \frac{1}{R}$, we can obtain the performance curve illustrated in Fig. 6.4. It should be noticed that the distortion for $R = \frac{1}{s}$ with even s is the same as that for $R = \frac{1}{s+1}$, i.e., the distortion can be satisfied with lower rate. Therefore, the group with odd size is better than that with even size.

Now, we extend this idea to general cases with R being arbitrary number from 0 to 1. By mixing two MV codes with different odd sizes s and $s + 2$, the average rate between $\frac{1}{s}$ and $\frac{1}{s+2}$ can be satisfied. If we use i groups of the rate $\frac{1}{s}$ MV code and j groups of the rate $\frac{1}{s+2}$ MV code as component codes, the average rate and the average distortion is a linear combination of the rate and distortion of two component codes, respectively. Since each group of the MV code yields 1 bit compressed coding result, $i + j = nR$. From the total bits in the MV code groups, we have $i \cdot s + j \cdot (s + 2) = n$. Consequently, we can obtain $i = n[(s + 2)R - 1]/2$ and $j = n(1 - sR)/2$.

The algorithm of the HMV code, consisting of two component MV codes, is summarized in Algorithm 6.1, of which the gist is to first find appropriate values of s, i and j given sequence length n and compression rate R. Then, encode group by group.

6.3.2 Decoding

There are two types of decoding algorithm, i.e., hard decoding and soft decoding. The hard decoding algorithm for the HMV code is also very simple, i.e., repeating the coded bit s times for

Algorithm 6.1. HYBRID MAJORITY VOTING CODE

Input: sequence X^n, R
Output: coded sequence Y^{nR}
set $s_{\text{mid}} = \text{floor}(1/R)$;
if s_{mid} is odd **then**
  set $s = s_{\text{mid}}$;
else
  set $s = s_{\text{mid}} - 1$;
end if
set $i = n[(s+2)R - 1]/2$ and $j = n(1 - sR)/2$;
for $t = 1$ **to** i **do**
  encode t-th group in X^n to $Y(t)$ by the rate $\frac{1}{s}$ MV code;
end for
for $t = i + 1$ **to** $i + j$ **do**
  encode t-th group in X^n to $Y(t)$ by the rate $\frac{1}{s+2}$ MV code;
end for

the first i-th group and $(s+2)$ times for the last j-th group. Consequently, the expected distortion of the HMV code can be expressed as

$$D_{\text{HMV}} = \frac{i \cdot s \cdot D_{\text{MV}}(s) + j \cdot (s+2) \cdot D_{\text{MV}}(s+2)}{n}. \tag{6.5}$$

The soft decoding yields LLR which is useful for exchanging the mutual information in iterative decoding of codes having multiple constituency components [87]. We first transform the received signal into LLR, and then take the expected distortion into account based on the correlation model in [79]. With the *a priori* LLR of received signal being LLR^a, the *a posteriori* LLR of the coded bits can be calculated as

$$\text{LLR}^p = \log \frac{(1-d) \cdot \exp(\text{LLR}^a) + d}{(1-d) + d \cdot \exp(\text{LLR}^a)}, \tag{6.6}$$

where d is set at $D_{\text{MV}}(s)$ and $D_{\text{MV}}(s+2)$ for the group coded by the rate $\frac{1}{s}$ and $\frac{1}{s+2}$ MV codes, respectively. After calculating the LLR^p of coded $(i+j)$-bits, the LLR^p_X of the sequence X^n is reproduced by performing repetition of the corresponding bit in LLR^p for s and $(s+2)$ times in the groups of the rate $\frac{1}{s}$ and $\frac{1}{s+2}$ MV codes, respectively. In this way, the LLR^p_X can be jointly exploited in further iteration of decoding, depending on specified algorithm for exchanging mutual information.

6.4 Implementation in Successive Refinement

Now, we start to implement the HMV code to successive refinement with DMS. The challenges of practical coding design for successive refinement with DMS include the following two aspects. First, for the standalone link, we have to find a lossy source coding scheme that minimizes the average distance between the codeword and information sequences. Meanwhile, for the refinement link, there should also be a complement coding scheme, which contains the information of lost part caused by the lossy source coding in the standalone link. The simplest way for lossy source coding

is puncturing, and the information of lost part is straightforwardly obtained in the form of the punctured bits. Consequently, we can decompose the information sequence into two codewords by puncturing for successive refinement. For convenience, this process that encoding a sequence into two codewords for successive refinement with DMS is simply referred to as codeword decomposition.

6.4.1 Problem of Codewords Overlapping

In the following, we focus on a DMS $X \sim \mathrm{Bern}(0.5)$. According to the rate-distortion theorem for successive refinement [81], the rates are achievable if

$$R_1 \geq I(X; \hat{X}_1), \tag{6.7}$$

$$R_1 + R_2 \geq I(X; \hat{X}_1, \hat{X}_2). \tag{6.8}$$

Obviously, $I(X; \hat{X}_2)$ can be equal to 1, i.e., $\hat{X}_2^n$ is a lossless recovery of X^n, if $R_1 + R_2 \geq 1$. For instance, the simplest method is to generate M_1 by puncturing X^n at rate R_1, and keep the punctured bits as M_2. Notice that the standalone link is equivalent to utilizing lossy source coding, and as illustrated in Fig. 6.1, the performance has a big gap between puncturing and the theoretical limit for the standalone link. The reason for the performance loss is that puncturing is performed bit by bit and does not efficiently utilize the minimum distance between the whole sequences.

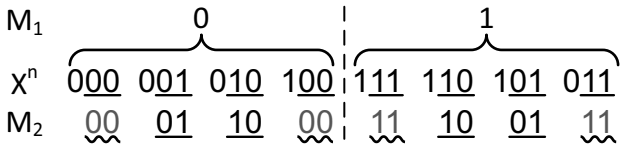


Fig. 6.5. Overlap of codewords happens if M_1 and M_2 are generated by the HMV code and puncturing, respectively.

For the purpose of improving the performance for the standalone link, we can use the HMV code instead of puncturing. However, for the refinement link, the performance with M_1 generated by the HMV code may be worse than that by puncturing. For example, we consider a special case $n = 3$, $R_1 = \frac{1}{3}$ and $R_2 = \frac{2}{3}$. It is obvious that D_2 can be easily reduced to 0 by puncturing X^n to generate M_1 and keeping the punctured bits as M_2. Then, we discuss the performance of the refinement link with M_1 and M_2 generated by the HMV code and puncturing, respectively. Since M_1 is generated by MV of 3 bits, each bit has the same weight of information. Therefore, we can puncture the first bit of X^n to generate M_2 without loss of generality. Notice from Fig. 6.5 that there are some cases that X^n cannot be losslessly reconstructed by M_1 and M_2, because M_1 and M_2 cannot determine a unique sequence of X^n due to the presence of the overlapped part in the codewords. For instance, X^n could be "000" or "100" if $M_1 =$ "0" and $M_2 =$ "00", resulting in the distortion when reconstructing X^n.

Consequently, codeword decomposition with the HMV code and puncturing may degrade the performance of the whole system, although the distortion in the standalone link is smaller. It is obvious that if some different sequences of X^n have the same M_1 and M_2, the reconstruction is not unique, resulting in a distortion. Hence, the necessary condition for lossless successive refinement is that all possible combinations of the decomposed codewords M_1 and M_2 are unique. If the standalone link utilizes the HMV code, it is hard to find a relatively simple coding scheme that contains all the information of lost part caused by the HMV code.

6.4.2 Codeword Decomposition with HMV Code

Notice that the principle of the HMV code is to divide the sequence X^n into many groups of bits with the group size being an odd number s or $(s+2)$, and then perform MV group by group. Therefore, in order to find the complement code of the HMV code, we can first find the complement code of MV code in a group with s bits, i.e., the rate $\frac{1}{s}$ MV code. For a sequence with s bits, the codeword of rate $\frac{1}{s}$ MV code is only 1 bit, and hence its complement code has $(s-1)$ bits. Since s is an odd number, all possible sequences with s bits contain two cases, i.e., the sequences with more "0" than "1", and its opposite. If we flip all bits of a sequence with more "0", the flipped sequence must have more "1" and vice versa. Hence, the number of the sequences with more "0" is equal to that of the sequences with more "1", i.e., 2^{s-1}. Notice that 2^{s-1} is also the number of all possible codewords with $(s-1)$ bits. Therefore, we can find a bijection that maps a s-bits sequence with more "0" to a $(s-1)$-bits codeword, and the bijection can be also utilized for the s-bits sequence with more "1" by flipping all bits of the sequence with more "0".

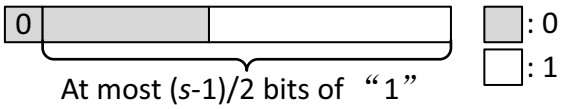


Fig. 6.6. The s-bits sequences with more "0" and the first bit being "0".

As depicted in Fig. 6.6, for the sequences with more "0" and the first bit being "0", the remaining $(s-1)$ bits contain at most $(s-1)/2$ bits of "1". Therefore, the number of the sequences with more "0" and the first bit being "0" is

$$\sum_{l=0}^{(s-1)/2} \binom{s-1}{l}, \tag{6.9}$$

which is also equal to the number of the $(s-1)$-bits codewords with "1" not more than "0". Thus, for the sequences with more "0" and the first bit being "0", we can directly use the last $(s-1)$ bits of the sequence as the codeword. Then, the remaining part of the bijection is mapping the sequences with "1" as the first bit to the codewords with more "1".

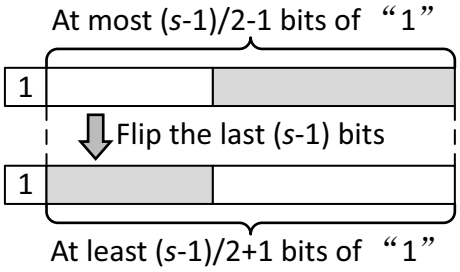


Fig. 6.7. The s-bits sequences with more "0" and the first bit being "1".

As illustrated in Fig. 6.7, if the first bit is "1" for a s-bits sequence with more "0", the remaining $(s-1)$ bits contain at most $[(s-1)/2-1]$ bits of "1". Consequently, the number of all possible sequences is

$$\sum_{l=0}^{\frac{s-1}{2}-1} \binom{s-1}{l} = \sum_{l=\frac{s-1}{2}+1}^{s-1} \binom{s-1}{l}. \tag{6.10}$$

Notice from (6.10) that the right side is equal to the number of the $(s-1)$-bits codewords with more "1". In addition, if we flip the remaining $(s-1)$ bits of the sequence with more "0" and the first bit being "1", the flipped $(s-1)$ bits must contain more "1". Hence, the mapping for the remaining part of the bijection can be simply performed by removing the first bit of the sequence and flipping the remaining bits to generate the codeword. Likewise, for a sequence with more "1", we first flipping all bits of the sequence to make the flipped sequence containing more "0", and then we can perform the same bijection stated above. Since the major process for this simple mapping rule is to flip bits, the bijection is referred to as BF code.

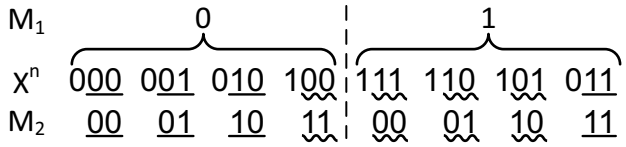


Fig. 6.8. Codewords decomposition by the HMV code and the BF code.

Fig. 6.8 shows a simple example of the codewords decomposition with the HMV code and the BF code for $n = 3$. It is obvious that $X^n \to (M_1, M_2)$ is a bijection, and hence X^n can be losslessly recovered from M_1 and M_2. Consequently, the BF code is the complement code of the HMV code, and a sequence can be decomposed into two codewords by the HMV code and the BF code for lossless successive refinement.

Algorithm 6.2. BIT FLIPPING CODE

Input: A group of bits B^s in the sequence X^n
Output: BF codeword W^{s-1}
if $B(1)$ is 1 **then**
  set $W(1:s-1) = \text{flip}(B(2:s))$;
else
  set $W(1:s-1) = B(2:s)$;
end if

Algorithm 6.2 summarizes the encoding algorithm of the BF code for a group of bits in the sequence X^n. To decompose a sequence into two codewords by the HMV and BF codes, we should generate the BF codewords group by group corresponding to the HMV code. For a group of bits in the sequence, if the first bit is "1", the codeword is generated by removing the first bit and flipping the remaining bits; otherwise, the codeword is the same as the remaining bits.

For the purpose of reconstructing the original sequence, we can perform Algorithm 6.3 group by group. To begin with, we recover the first bit by flipping the HMV codeword in the same group of the BF codeword, if more than half of the bits in the BF codeword are "1"; otherwise, the first bit is the same as the HMV codeword. If the recovered first bit is "1", the remaining bits in the group are generated by flipping the BF codeword; otherwise, the remaining bits keep the same as the group of the BF codeword.

Remark: Notice that $1 \geq R_1 + R_2 \geq R_1$ in the general case where the refinement link still is lossy. Since the HMV code and the BF code encode group by group, the codeword decomposition can be easily extended from lossless case to lossy case by mixing a certain ratio of lossless codeword decomposition and lossy source coding with the HMV code. More specifically, we only generate the BF code for some of the groups in X^n to satisfy the rate R_2, and hence the

Algorithm 6.3. DECODING OF CODEWORD DECOMPOSITION

Input: HMV codeword Y in a group, BF codeword W^{s-1} in the same group
Output: A group of bits B^s in the sequence X^n
set n_1 as the number of "1" in W^{s-1};
if $n_1 > \frac{s-1}{2}$ **then**
  set $B(1) = \text{flip}(Y)$;
else
  set $B(1) = Y$;
end if
if $B(1)$ is 1 **then**
  set $B(2:s) = \text{flip}(W(1:s-1))$;
else
  set $B(2:s) = W(1:s-1)$;
end if

groups with the BF code can be losslessly recovered while the groups without the BF code are still lossy. With this technique, we can trade off (R_1, R_2) and (D_1, D_2). Finally, from the view of the whole sequence, the refinement link is lossy but still has a better performance than utilizing puncturing.

6.5 Performance Evaluation

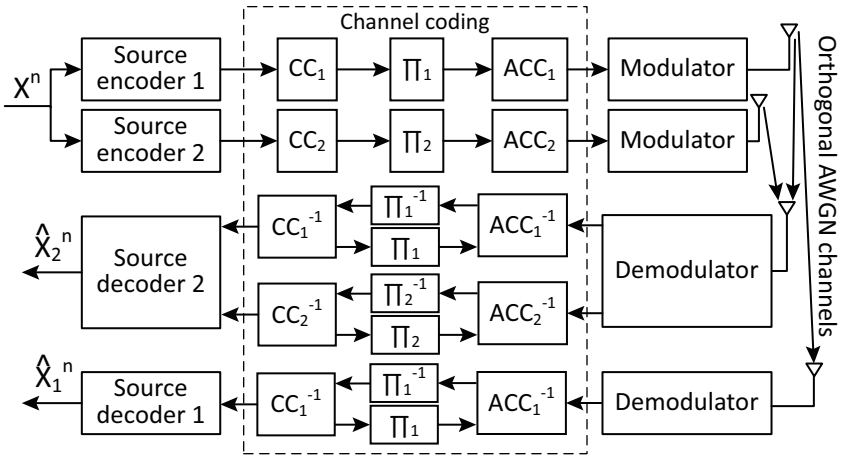


Fig. 6.9. The system model for simulations.

We design a simulation system illustrated in Fig. 6.9 to evaluate the performance of codeword decomposition. The sequence X^n is decomposed into two codewords by two source encoders for the first step. In each link, the codeword is outer-encoded by CC, and an ACC is deployed for inner-encoding after an random interleaver Π. After modulation, the outer-coded bits are transmitted to receivers via orthogonal AWGN channels. Then, the receivers demodulate the received signals, and further perform iterative decoding by CC^{-1} and ACC^{-1}. The interleaver Π and deinterleaver Π^{-1} are utilized to exchange extrinsic information between CC^{-1} and ACC^{-1}. Finally, the recoveries are reconstructed by source decoders with the hard decisions of channel coding.

Table 6.1. BASIC SETTINGS OF SIMULATION PARAMETERS

Parameter	Value	Parameter	Value
Sequence length	10^4 bits	The number of sequence	10^5
Generator polynomial of CC	$([3, 2]3)_8$	Type of interleaver	random interleaver
Modulation method	BPSK	Maximum iteration time	20

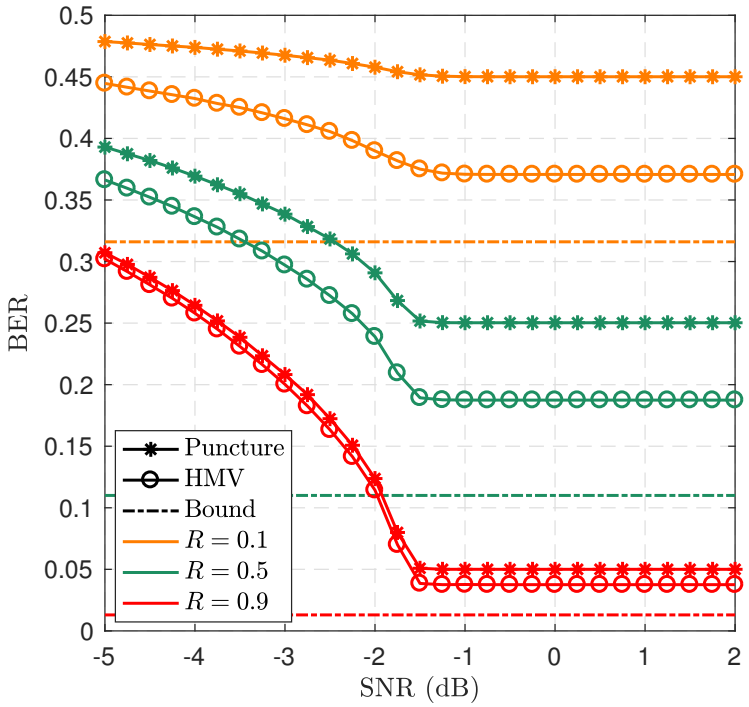
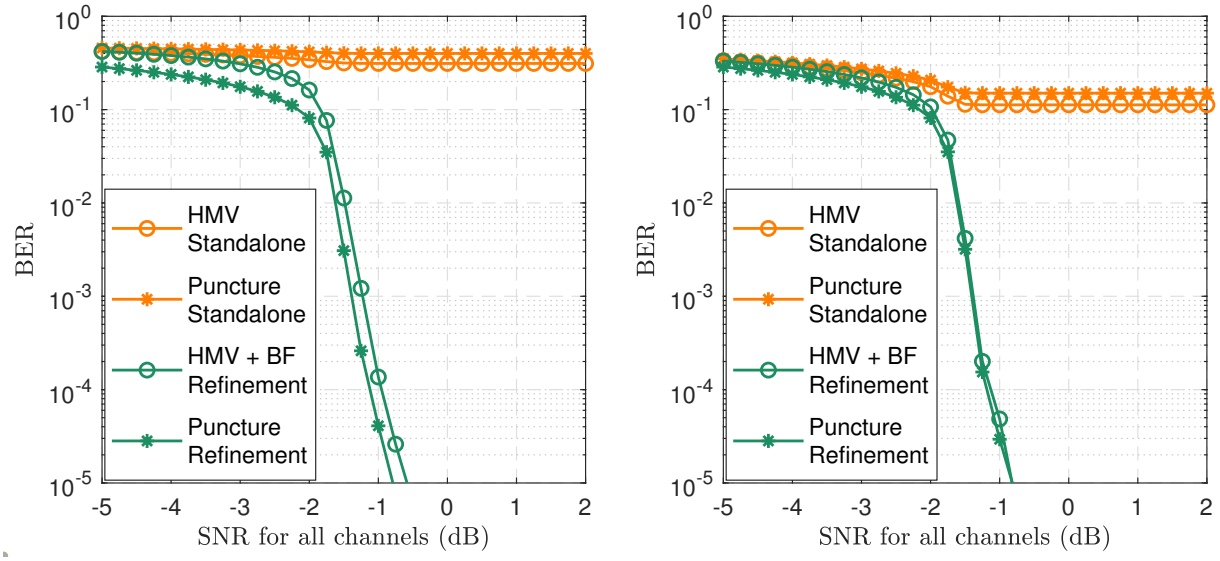


Fig. 6.10. The performance of joint source-channel coding in the standalone link.

The basic setting of simulation parameters is listed in Table 6.1. First of all, we investigate the performance in the standalone link as depicted in Fig. 6.10. It is clear that the joint source-channel coding scheme based on the HMV code has lower BER than puncturing, with the same SNR value. Furthermore, the SNR floor for the HMV code is also closer to the theoretical bound derived from the rate-distortion function. This observation confirms that a coding scheme can get better performance by replacing the puncturing component with the HMV code.

Fig. 6.11 shows the performance of codeword decomposition for the whole system. Clearly, the HMV code has lower D_1 than puncturing in the standalone link, because it builds correlations among the bits of DMS by dispersing information within a group. As the complement code of the HMV code, the BF code also generates correlated bits in a group, resulting in the error of a bit propagating to other bits in the same group. The group length is larger for smaller R_1, and hence the problem of error propagation becomes more severe due to more influenced bits in a group. Consequently, the codeword decomposition by the HMV code and the BF code has more performance loss in the refinement link only when SNR is not large enough. However, if noise can be completely eliminated by channel coding, there is no error floor for the codeword decomposition by the HMV code. Moreover, if R_1 is relatively large, D_2 for codeword decomposition by the HMV



(a) $R_1 = 0.2$. (b) $R_1 = 0.7$.

Fig. 6.11. The BER performance for the whole system.

code in the refinement link is almost the same as puncturing. Even for a very small $R_1 = 0.2$, the performance loss of the refinement link is less than 0.25 dB. Anyway, in spite of R_1, the standalone link always has obvious performance gain by the HMV code than puncturing.

6.6 Methodology for Hybrid Codes Design

As presented above, given a specified codebook of a rate-distortion code, we can calculate its source coding rate and expected distortion, which can be plotted as one point of the rate-distortion function. If we have a series of rate-distortion codes, we can utilize the linear combination of them to satisfy arbitrary rate and distortion between two component codes. Now, we introduce the general method for finding component codes, which can be used to construct new hybrid codes with lower distortion than the HMV code.

6.6.1 Hybrid Codes Based on the Duality of Channel Coding

From the duality between source coding and channel coding, we can notice that the MV code is dual to the repeat code. For encoding, the MV code maps n bits to 1 bit, while the repeat code maps 1 bit to n bits, and the decoding process is converse. This finding inspires us to exploit the optimal codes to compose hybrid codes. In the following, we use the Hamming codes [88] as an example to explain the method of hybrid codes design based on the duality between source coding and channel coding.

For a Hamming code with n total bits and k data bits, it has $(n-k)$ parity bits which can detect 1-bit error at $(2^{n-k}-1)$ different positions. Therefore, in decoding process, each possible codeword can be mapped from 2^{n-k} instances of received sequence, including 1 sequence without error and $(2^{n-k}-1)$ sequences with one bit error. Notice that 1 data sequence with k bits is mapped to 1 codeword with n bits, and hence the number of possible codewords is 2^k. Consequently, 2^k

possible codewords can be mapped form $2^k \times 2^{n-k} = 2^n$ received sequences in total, i.e., the mapping of 2^k possible codewords can completely cover the sequence space with n bits.

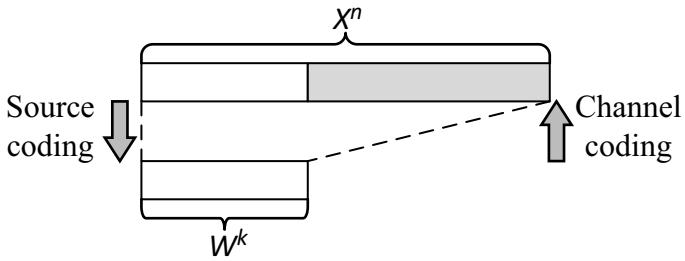


Fig. 6.12. Duality between source coding and channel coding.

By making reverse of the encoding/decoding process of the Hamming codes as Fig. 6.12, we can map n-bits sequence to k-bits sequence, i.e., equivalent to lossy source coding with $R = \frac{k}{n}$. More specifically, for encoding a source sequence X^n, we first use the parity-check matrix to calculate its syndrome. Then, if any bit of the syndrome is not equal to "0", we flip 1 bit in X^n according to the syndrome. Finally, the systematic bits of the output are selected as the k-bits codeword W^k of lossy source coding. For decoding, the estimate $\hat{X}^n$ can be easily reconstructed by utilizing the generator matrix of the Hamming codes with W^k as the input.

Obviously, there are $(2^{n-k}-1)$ cases where the distance between $\hat{X}^n$ and X^n is 1, i.e., $(2^{n-k}-1)$ bits of error in 2^{n-k} sequences with length being n. Moreover, we have $(2^{n-k}-1) = n$ for the Hamming codes. Hence, the expected distortion of lossy source coding based on the Hamming codes can be calculated as

$$D_{\text{Hamming}} = \frac{2^{n-k}-1}{2^{n-k} \times n} = \frac{1}{2^{n-k}}. \tag{6.11}$$

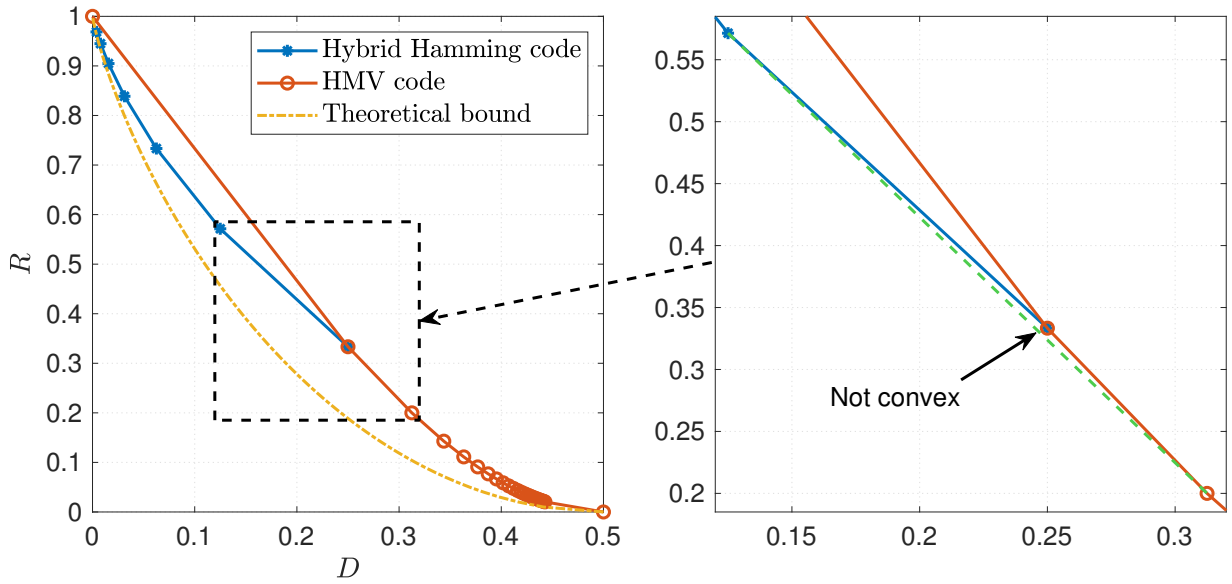


Fig. 6.13. The expected distortion of hybrid codes based on the Hamming codes.

Fig. 6.13 shows the rate-distortion function for the hybrid code composed of the Hamming codes with diverse rates. Interestingly, the curves of the hybrid Hamming code and the HMV code

cross at $(\frac{1}{4}, \frac{1}{3})$. Because the Hamming$(3, 1)$ code is essentially the same as the repeat code (MV code) with 3 bits. For $R > \frac{1}{3}$, the hybrid Hamming code has lower distortion than the HMV code. Therefore, we can use both the Hamming code and the MV code as component codes to generate a new hybrid codes with better performance. Notice that the curve combined with the Hamming code and the MV code is not convex between $(\frac{1}{8}, \frac{4}{7})$ and $(\frac{5}{16}, \frac{1}{5})$, which is found in the magnified part of Fig. 6.13 with the dashed line connecting the two points below the rate-distortion functions of the Hamming code and the MV code. Hence, to optimize the performance for $\frac{1}{5} < R < \frac{4}{7}$, we can utilize the Hamming$(7, 4)$ code and the rate $\frac{1}{5}$ MV code as component codes.

6.6.2 Codeword Decomposition with the Hamming Codes

Since the Hamming(n, k) code has $(n-k)$-bits syndrome, we can identify a unique n-bits sequence by k data bits and $(n-k)$-bits syndrome. Therefore, the syndrome of n-bits sequence completely contains the lost part of information in lossy source coding by the Hamming codes. With the $(n-k)$-bits syndrome as the complement code, we are able to losslessly recover the n-bits sequence by joint decoding for successive refinement.

For a sequence X^n, we can follow the algorithm described above to generate the codeword W^k for lossy source coding. To generate the corresponding complement codeword W_C^{n-k}, we only need to calculate the syndrome by the parity-check matrix with X^n as the input. In the decoding for successive refinement, we can calculate the n-bits sequence $\hat{X}^n$ by the generator matrix of the Hamming codes with W^k as the input for the first step. Then, we flip 1 bit in $\hat{X}^n$ at the position specified by the syndrome W_C^{n-k}, if W_C^{n-k} is not all "0" bits. By this means, we can easily utilize the Hamming codes and its syndrome for lossy source coding and successive refinement.

6.7 Summary

We have developed a lossy compression scheme, i.e., the HMV code, with relatively high efficiency and low complexity for practical use. Based on a special and simplest case of the joint typicality coding scheme, we have found the basic MV code and analyze its rate-distortion performance. Then, we exploited the two MV component codes to construct the HMV code, for the purpose of adapting arbitrary compression rate. The encoding and hard/soft decoding algorithms for the HMV code were also presented in detail. Moreover, we also implement the HMV code to successive refinement with DMS. We utilize the HMV code in the standalone link to achieve a better performance than puncturing, and then develop the BF code to contain all information of lost part caused by the HMV code. In addition, we conducted a series of simulations to compare the performance difference between the HMV code and puncturing. The simulation results verify that the HMV code can achieve better performance than puncturing. Although the simulation results demonstrate that there is a small trade-off of performance between two links, the performance of the whole system is obviously better for relatively large R_1 by codeword decomposition with the HMV code and the BF code. It should be emphasized that the proposed technique can easily perform the trade-off between (R_1, R_2) and (D_1, D_2), in the case $1 \geq R_1 + R_2 \geq R_1$. Finally, we conclude the methodology for hybrid codes design by utilizing the duality between source coding and channel coding. One of the optimal codes, i.e., the Hamming codes, is exemplified for applying the channel coding scheme into lossy source coding. We also find the syndrome of the Hamming codes as the complement code for successive refinement.

CHAPTER 7

Conclusion and Outlook

This dissertation have investigated several interesting topics belonging to the category of helper-assisted lossy multiterminal source coding.

Initially, we analyze the performance improvement by introducing a helper into the system of multiterminal source coding with two correlated sources. An inner bound on the achievable rate-distortion region is derived, and then utilized to calculate the upper bound of the outage probability over block Rayleigh fading channels. In the numerical results, the derived inner bound accurately match the Berger-Tung inner bound, when the helper link is equivalently removed. The results also demonstrate that a helper can obviously extend the achievable rate-distortion region and decrease the outage probability.

Then, we concentrate on the binary CEO problem with a helper. In order to investigate the performance limit, we divide the binary CEO problem with a helper into two sub problems as multiterminal source coding with a helper and final decision. By this means, we derive an outer bound on the achievable rate-distortion region, and formulate a convex optimization problem to minimize the distortions in the step of multiterminal source coding. We further analyze the distortion propagating from the decoding results of multiterminal source coding to the final decision for theoretically optimal decision and MV decision. Through simulations, we also compare the performance gain between introducing a helper and locating an extra agent.

Subsequently, we make efforts to evaluate the trade-off between link rates and final distortion for lossy source coding with helpers. Based on joint typicality coding, we prove the achievability of an inner bound on the rate-distortion region for general sources. Then, we further calculate the inner bound for the case with binary sources. If there is only one helper and no rate constraint on it, we find that the derived inner bound precisely coincides with the Wyner-Ziv theorem for arbitrary correlation level between sources.

Moreover, we investigate an implementation of multiterminal source coding in wireless cooperative communication networks, i.e., lossy communications with an LF relay. For the first step, we determine an inner bound on the achievable rate-distortion region of lossy source coding with a helper. Then, we calculate the upper bound of the outage probability for lossy LF relaying over block Rayleigh fading channels. The theoretical results indicate that outage event occurs less frequently, if lossless recovery is not necessarily needed in the destination. In addition, we make a comparison of outage probability among AF, DF and LF for the relaying system allowing lossy communications through simulations.

Finally, we develop the HMV code, i.e., a practical lossy source coding scheme with obviously higher efficiency than puncturing, in contrast to its simple coding and decoding algorithms. We also theoretically analyze the rate-distortion performance of the HMV code. After finding the BF code as the corresponding complement code, the HMV code is implemented to successive refinement with DMS. The simulations results confirm that, compared to puncturing, the HMV code can significantly reduce the distortion in the standalone link, while the performance of the refinement link keeps almost the same. To conclude the methodology of hybrid codes design, we present an example based on the Hamming codes, and develop the corresponding complement code for successive refinement by calculating the syndrome of the Hamming codes.

In the future, we can extend the related studies in the following directions:

- The binary CEO problem with a helper is solved in this dissertation by decomposing it into a successive process, i.e., multiterminal source coding with a helper, and then final decision. However, the optimality of this successive decoding scheme is an open problem. Therefore, it is a meaningful work to evaluate the optimality of successive decoding scheme, or find another better decoding scheme for the binary CEO problem with a helper.
- The theoretical outage probability of the binary CEO problem with a helper and lossy communications with helpers requires extremely complicated multiple integral; nevertheless, it is still a very interesting implementation to wireless cooperative communications over fading channels.
- For simplicity in theoretical analysis, we assume the channels to be orthogonal for the binary CEO problem and lossy communications with helpers. Multiple access techniques and multiuser detection schemes are the key to applying the theoretical results to practical systems.
- Only the inner bound, i.e., sufficient condition, on the achievable rate-distortion region is derived for lossy source coding with helpers in this dissertation. For the necessary condition of lossy source coding with helpers, to determine the outer bound also make significant sense.
- Regarding lossy communications in relaying systems, we only derive the theoretical outage probability for LF relaying. The theoretical outage probability for lossy communications with AF or DF will provide a more distinct view of performance limit.
- The LF strategy analyzed in this dissertation always forwards the relay sequence to the destination regardless of the intra-link error. Although the performance of distortion will not be worse by the adaptive LLR updating function, the performance gain becomes very small if the relay sequence contains too many errors. However, the relay still consumes the same power to forward the sequence and the destination needs to perform relatively complicated algorithm for joint decoding. Thus, the trade-off between power consumption and final distortion is of great importance for practical system design. One reasonable method is to evaluate the error probability in the relay sequence, and compare with some threshold before forwarding the sequence to the destination.
- Relatively simple channel coding schemes are applied in simulations for verifying the tendency of theoretical results in this dissertation. Since the encoded sequences in different links can be regarded as distributed codes, it is possible to obtain more distributed coding gains by replacing the relatively simple channel coding scheme with more complicated codes, such as turbo codes with long memories, LDPC codes and polar codes [89].
- Within the framework of grouped lossy source coding for practical use, the way to further improve the coding efficiency is to find some simple component code in rate $\frac{u}{s}$, i.e., compress an s-bits sequence group into a u-bits codeword with better performance and then construct a new type of hybrid code satisfying lower distortion. However, it might not be easy to find the corresponding complement code for successive refinement.
- Another difficult but valuable work is the practical coding design, which can have high performance while keeps relatively low complexity, for multiterminal source coding with more than one source to be reconstructed.

APPENDIX A

Error Probability by Weighted MV

Regarding the error probability by weighted MV, the final decision $\hat{x}$ follows [90]:

$$\hat{x} = \begin{cases} 1, & \text{if } \mathbf{w}\mathbf{b}^{\mathrm{T}} > 0, \\ 0, & \text{otherwise}, \end{cases} \tag{A.1}$$

where $\mathbf{w} = [\log \frac{1-p_1}{p_1}, \cdots, \log \frac{1-p_L}{p_L}]$ and $\mathbf{b} = 2 \cdot [\hat{x}_1, \cdots, \hat{x}_L] - 1$. Similarly to the Poisson binomial process, the error probability for the estimate of x is given by

$$p_e = \Pr\left\{ \sum_{k \in \mathcal{B}_+} w_k > \sum_{j \in \mathcal{B}_-} w_j \right\} + \frac{1}{2} \Pr\left\{ \sum_{k \in \mathcal{B}_+} w_k = \sum_{j \in \mathcal{B}_-} w_j \right\}, \tag{A.2}$$

where $\mathcal{B}_+ = \{i | b_i = +1\}$ and $\mathcal{B}_- = \{i | b_i = -1\}$. Note that in order to calculate (A.2), it needs to carry out the search over all the possible combinations of w_i.

APPENDIX B

Proof of Lemma 4.1

First, consider

$$
\begin{aligned}
&\Pr\{(\tilde{V}_{\mathcal{S}}^n, \tilde{v}_{\mathcal{S}^c}^n) \in \mathcal{T}_\epsilon^{(n)}(V_{\mathcal{S}}, V_{\mathcal{S}^c})\} \\
&\quad= \sum_{\tilde{v}_{\mathcal{S}}^n \in \mathcal{T}_\epsilon^{(n)}(V_{\mathcal{S}}|\tilde{v}_{\mathcal{S}^c}^n)} \prod_{j\in\mathcal{S}} p(\tilde{v}_j^n) \\
&\quad\le \sum_{\tilde{v}_{\mathcal{S}}^n \in \mathcal{T}_\epsilon^{(n)}(V_{\mathcal{S}}|\tilde{v}_{\mathcal{S}^c}^n)} \mathrm{pow}\left(2, -n\sum_{j\in\mathcal{S}}(1-\epsilon)H(V_j)\right) \\
&\quad= \sum_{\tilde{v}_{\mathcal{S}}^n \in \mathcal{T}_\epsilon^{(n)}(V_{\mathcal{S}}|\tilde{v}_{\mathcal{S}^c}^n)} \mathrm{pow}\left(2, -n(1-\epsilon)\left[H(V_{\mathcal{S}}) + \sum_{j=2}^{|\mathcal{S}|} I(V_{S_1^{j-1}}; V_{S_j})\right]\right) \\
&\quad= \left|\mathcal{T}_\epsilon^{(n)}(V_{\mathcal{S}}|\tilde{v}_{\mathcal{S}^c}^n)\right| \cdot \mathrm{pow}\left(2, -n(1-\epsilon)\left[H(V_{\mathcal{S}}) + \sum_{j=2}^{|\mathcal{S}|} I(V_{S_1^{j-1}}; V_{S_j})\right]\right) \\
&\quad\le \mathrm{pow}\left(2, n(1+\epsilon)H(V_{\mathcal{S}}|V_{\mathcal{S}^c}) - n(1-\epsilon)\left[H(V_{\mathcal{S}}) + \sum_{j=2}^{|\mathcal{S}|} I(V_{S_1^{j-1}}; V_{S_j})\right]\right) \\
&\quad= \mathrm{pow}\left(2, -n\left[-(1+\epsilon)H(V_{\mathcal{S}}|V_{\mathcal{S}^c}) + (1-\epsilon)H(V_{\mathcal{S}}) + (1-\epsilon)\sum_{j=2}^{|\mathcal{S}|} I(V_{S_1^{j-1}}; V_{S_j})\right]\right) \\
&\quad= \mathrm{pow}\left(2, -n\left[I(V_{\mathcal{S}}; V_{\mathcal{S}^c}) + \sum_{j=2}^{|\mathcal{S}|} I(V_{S_1^{j-1}}; V_{S_j}) - \delta(\epsilon)\right]\right), \qquad \text{(B.1)}
\end{aligned}
$$

where $\delta(\epsilon) = \epsilon\left[H(V_{\mathcal{S}}|V_{\mathcal{S}^c}) + H(V_{\mathcal{S}}) + \sum_{j=2}^{|\mathcal{S}|} I(V_{S_1^{j-1}}; V_{S_j})\right]$. This completes the proof of **Lemma 4.1**.

APPENDIX C

Proof of Lemma 4.2

Define the events

$$\tilde{\mathcal{E}}_k = \{(V_{S_1}^n(k_{S_1}), \cdots, V_{S_{|\mathcal{S}|}}^n(k_{S_{|\mathcal{S}|}}), \tilde{V}_{\mathcal{S}^c}^n) \in \mathcal{T}_\epsilon^{(n)}\} \text{ for } k_i \in \mathcal{K}_i, i \in \mathcal{S}. \tag{C.1}$$

By the union of events bound, the probability of the event of interest can be bounded as

$$\begin{aligned}
\Pr\left(\bigcup_{k_i \in \mathcal{K}_i, i \in \mathcal{S}} \tilde{\mathcal{E}}_k\right) &\leq \sum_{k_i \in \mathcal{K}_i, i \in \mathcal{S}} \Pr(\tilde{\mathcal{E}}_k) \\
&= \prod_{i \in \mathcal{S}} 2^{nr_i} \cdot \Pr(\tilde{\mathcal{E}}_k) \\
&\leq \prod_{i \in \mathcal{S}} 2^{nr_i} \cdot \mathrm{pow}\left(2, -n\left[\sum_{j=2}^{|\mathcal{S}|} I(V_{S_1^{j-1}}; V_{S_j}) + I(V_{\mathcal{S}}; V_{\mathcal{S}^c}) - \delta(\epsilon)\right]\right) \qquad \text{(C.2)} \\
&= \mathrm{pow}\left(2, n\sum_{i \in \mathcal{S}} r_i - n\left[\sum_{j=2}^{|\mathcal{S}|} I(V_{S_1^{j-1}}; V_{S_j}) + I(V_{\mathcal{S}}; V_{\mathcal{S}^c}) - \delta(\epsilon)\right]\right), \qquad \text{(C.3)}
\end{aligned}$$

where (C.2) follows according to **Lemma 4.1**. Notice that (C.3) tends to zero as $n \to \infty$ if

$$\sum_{i \in \mathcal{S}} r_i < \sum_{j=2}^{|\mathcal{S}|} I(V_{S_1^{j-1}}; V_{S_j}) + I(V_{\mathcal{S}}; V_{\mathcal{S}^c}) - \delta(\epsilon). \tag{C.4}$$

This completes the proof of **Lemma 4.2**.

References

[1] B. Ahlgren, M. Hidell, and E. C.-H. Ngai, "Internet of Things for smart cities: Interoperability and open data," *IEEE Internet Computing*, vol. 20, no. 6, pp. 52–56, Dec. 2016.

[2] A. Alaiad and L. Zhou, "Patients' adoption of WSN-based smart home healthcare systems: An integrated model of facilitators and barriers," *IEEE Transactions on Professional Communication*, vol. 60, no. 1, pp. 4–23, Mar. 2017.

[3] Q. Chi, H. Yan, C. Zhang, Z. Pang, and L. Da Xu, "A reconfigurable smart sensor interface for industrial WSN in IoT environment," *IEEE transactions on industrial informatics*, vol. 10, no. 2, pp. 1417–1425, May 2014.

[4] S. R. Nimbargi, S. Hadawale, and G. Ghodke, "Tsunami alert & detection system using IoT: A survey," in *International Conference on Big Data, IoT and Data Science*. Pune, India: IEEE, Dec. 2017, pp. 182–184.

[5] B. Siregar, A. B. A. Nasution, and F. Fahmi, "Integrated pollution monitoring system for smart city," in *International Conference on ICT For Smart Society (ICISS)*. Surabaya, Indonesia: IEEE, Jul. 2016, pp. 49–52.

[6] Z. Zou, Y. Bao, F. Deng, and H. Li, "An approach of reliable data transmission with random redundancy for wireless sensors in structural health monitoring," *IEEE Sensors Journal*, vol. 15, no. 2, pp. 809–818, Feb. 2015.

[7] S. Long and P. Xiang, "Lossless data compression for wireless sensor networks based on modified bit-level RLE," in *8th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM)*. Shanghai, China: IEEE, Sep. 2012, pp. 1–4.

[8] V. Dedeoglu, S. Perreau, and A. Grant, "Distributed energy consumption minimization for lossless data gathering wireless sensor networks," in *Australian Communications Theory Workshop (AusCTW)*. Wellington, New Zealand: IEEE, Feb. 2012, pp. 145–149.

[9] A. Irawan and T. Matsumoto, "Feedback-assisted correlated packet transmission with a helper," *IEEE Transactions on Vehicular Technology*, May 2018.

[10] R. Tandon, S. Ulukus, and K. Ramchandran, "Secure source coding with a helper," *IEEE Transactions on Information Theory*, vol. 59, no. 4, pp. 2178–2187, Oct. 2013.

[11] P. Ostovari, J. Wu, and A. Khreishah, "Cooperative internet access using helper nodes and opportunistic scheduling," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 7, pp. 6439–6448, Jul. 2017.

[12] S. Gherekhloo and A. Sezgin, "Latency-limited broadcast channel with cache-equipped helpers," *IEEE Transactions on Wireless Communications*, Jul. 2017.

[13] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Transactions on Information Theory*, vol. 19, no. 4, pp. 471–480, Jul. 1973.

[14] T. Cover, "A proof of the data compression theorem of Slepian and Wolf for ergodic sources (corresp.)," *IEEE Transactions on Information Theory*, vol. 21, no. 2, pp. 226–228, Mar. 1975.

[15] T. Berger, "Multiterminal source coding," in *The Information Theory Approach to Communications*, G. Longo, Ed. New York: Springer-Verlag, 1978, pp. 171–231.

[16] S. Y. Tung, "Multiterminal source coding," Ph.D. dissertation, School of Electrical Engineering, Cornell University, Ithaca, New York, 1978.

[17] Y. Oohama, "Gaussian multiterminal source coding," *IEEE Transactions on Information Theory*, vol. 43, no. 6, pp. 1912–1923, Nov. 1997.

[18] A. B. Wagner, S. Tavildar, and P. Viswanath, "Rate region of the quadratic Gaussian two-encoder source-coding problem," *IEEE Transactions on Information Theory*, vol. 54, no. 5, pp. 1938–1961, May 2008.

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

[20] J. Chen and T. Berger, "Robust distributed source coding," *IEEE Transactions on Information Theory*, vol. 54, no. 8, pp. 3385–3398, Aug. 2008.

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

[22] X. He, X. Zhou, P. Komulainen, M. Juntti, and T. Matsumoto, "A lower bound analysis of Hamming distortion for a binary CEO problem with joint source-channel coding," *IEEE Transactions on Communications*, vol. 64, no. 1, pp. 343–353, Jan. 2016.

[23] X. He, X. Zhou, M. Juntti, and T. Matsumoto, "A rate-distortion region analysis for a binary CEO problem," in *IEEE 83rd Vehicular Technology Conference (VTC Spring)*, Nanjing, China, May 2016, pp. 1–5.

[24] R. Ahlswede and J. Korner, "Source coding with side information and a converse for degraded broadcast channels," *IEEE Transactions on Information Theory*, vol. 21, no. 6, pp. 629–637, Nov. 1975.

[25] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Transactions on information Theory*, vol. 22, no. 1, pp. 1–10, Jan. 1976.

[26] A. Sechelea, A. Munteanu, S. Cheng, and N. Deligiannis, "On the rate-distortion function for binary source coding with side information," *IEEE Transactions on Communications*, vol. 64, no. 12, pp. 5203–5216, Dec. 2016.

[27] M. S. Rahman and A. B. Wagner, "Rate region of the vector Gaussian one-helper source-coding problem," *IEEE Transactions on Information Theory*, vol. 61, no. 5, pp. 2708–2728, May 2015.

[28] A. Sgarro, "Source coding with side information at several decoders," *IEEE Transactions on Information Theory*, vol. 23, no. 2, pp. 179–182, Mar. 1977.

[29] R. Timo, T. Chan, and A. Grant, "Rate distortion with side-information at many decoders," *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 5240–5257, Aug. 2011.

[30] W. H. R. Equitz and T. M. Cover, "Successive refinement of information," *IEEE Transactions on Information Theory*, vol. 37, no. 2, pp. 269–275, 1991.

[31] V. K. Goyal, "Multiple description coding: Compression meets the network," *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 74–93, 2001.

[32] J. K. Wolf, A. D. Wyner, and J. Ziv, "Source coding for multiple descriptions," *The Bell System Technical Journal*, vol. 59, no. 8, pp. 1417–1426, 1980.

[33] H. S. Witsenhausen and A. D. Wyner, "Source coding for multiple descriptions II: A binary source," *Bell System Technical Journal*, vol. 60, no. 10, pp. 2281–2292, 1981.

[34] A. El Gamal and T. Cover, "Achievable rates for multiple descriptions," *IEEE Transactions on Information Theory*, vol. 28, no. 6, pp. 851–857, 1982.

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

[36] A. Wolf, D. C. González, M. Dörpinghaus, J. C. S. Santos Filho, and G. Fettweis, "On the binary lossless many-help-one problem with independently degraded helpers," in *56th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. Monticello, IL, USA: IEEE, Oct. 2018, pp. 538–542.

[37] T. Han and K. Kobayashi, "A unified achievable rate region for a general class of multiterminal source coding systems," *IEEE Transactions on Information Theory*, vol. 26, no. 3, pp. 277–288, May 1980.

[38] A. Wyner, "On source coding with side information at the decoder," *IEEE Transactions on Information Theory*, vol. 21, no. 3, pp. 294–300, May 1975.

[39] L. Rey Vega, P. Piantanida, and A. O. Hero, "The three-terminal interactive lossy source coding problem," *IEEE Transactions on Information Theory*, vol. 63, no. 1, pp. 532–562, Jan. 2017.

[40] J. He, V. Tervo, X. Zhou, X. He, S. Qian, M. Cheng, M. Juntti, and T. Matsumoto, "A tutorial on lossy forwarding cooperative relaying," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 66–87, Firstquarter 2018.

[41] R. Hu and J. Li, "Exploiting Slepian-Wolf codes in wireless user cooperation," in *IEEE 6th Workshop on Signal Processing Advances in Wireless Communications*, New York, USA, Jun. 2005, pp. 275–279.

[42] M. Cheng, K. Anwar, and T. Matsumoto, "Outage probability of a relay strategy allowing intra-link errors utilizing Slepian-Wolf theorem," *EURASIP Journal on Advances in Signal Processing*, vol. 2013, no. 1, p. 34, Dec. 2013.

[43] X. Zhou, M. Cheng, X. He, and T. Matsumoto, "Exact and approximated outage probability analyses for decode-and-forward relaying system allowing intra-link errors," *IEEE Transactions on Wireless Communications*, vol. 13, no. 12, pp. 7062–7071, Dec. 2014.

[44] B. Zhao and M. C. Valenti, "Distributed turbo coded diversity for relay channel," *Electronics Letters*, vol. 39, no. 10, pp. 786–787, May 2003.

[45] H. H. Sneessens, J. Louveaux, and L. Vandendorpe, "Turbo-coded decode-and-forward strategy resilient to relay errors," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Las Vegas, NV, USA, Apr. 2008, pp. 3213–3216.

[46] K. Anwar and T. Matsumoto, "Accumulator-assisted distributed turbo codes for relay systems exploiting source-relay correlation," *IEEE Communications Letters*, vol. 16, no. 7, pp. 1114–1117, Jul. 2012.

[47] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: Turbo-codes," *IEEE Transactions on Communications*, vol. 44, no. 10, pp. 1261–1271, Oct. 1996.

[48] M. Brulatout, H. Khalife, V. Conan, S. Szott, M. Natkaniec, K. Kosek-Szott, and L. Prasnal, "A cooperative MAC protocol for lossy forwarding networks," in *Wireless Days (WD)*. Toulouse, France: IEEE, Mar. 2016, pp. 1–3.

[49] A. Wolf, M. Matthé, A. Festag, and G. Fettweis, "Outage based power allocation for a lossy forwarding two-relaying system," in *IEEE 20th International Workshop on Computer Aided Modelling and Design of Communication Links and Networks (CAMAD)*, Guildford, UK, Sep. 2015, pp. 267–272.

[50] A. B. Wagner and V. Anantharam, "An improved outer bound for multiterminal source coding," *IEEE Transactions on Information Theory*, vol. 54, no. 5, pp. 1919–1937, May 2008.

[51] S. Jana and R. Blahut, "Canonical description for multiterminal source coding," in *IEEE International Symposium on Information Theory (ISIT)*, Toronto, ON, Canada, Jul. 2008, pp. 697–701.

[52] X. He, X. Zhou, M. Juntti, and T. Matsumoto, "Data and error rate bounds for binary data gathering wireless sensor networks," in *IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Stockholm, Sweden, Jun. 2015, pp. 505–509.

[53] A. Razi and A. Abedi, "Convergence analysis of iterative decoding for binary CEO problem," *IEEE Transactions on Wireless Communications*, vol. 13, no. 5, pp. 2944–2954, May 2014.

[54] J. Haghighat, H. Behroozi, and D. V. Plant, "Iterative joint decoding for sensor networks with binary CEO model," in *IEEE 9th Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Recife, Brazil, Jul. 2008, pp. 41–45.

[55] X. He, X. Zhou, K. Anwar, and T. Matsumoto, "Estimation of observation error probability in wireless sensor networks," *IEEE Communications Letters*, vol. 17, no. 6, pp. 1073–1076, Jun. 2013.

[56] J. N. Laneman, D. N. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3062–3080, 2004.

[57] X. Zhou, P.-S. Lu, K. Anwar, and T. Matsumoto, "Correlated sources transmission in orthogonal multiple access relay channel: Theoretical analysis and performance evaluation," *IEEE Transactions on Wireless Communications*, vol. 13, no. 3, pp. 1424–1435, Jan. 2014.

[58] P.-S. Lu, X. Zhou, and T. Matsumoto, "Outage probabilities of orthogonal multiple-access relaying techniques with imperfect source-relay links," *IEEE Transactions on Wireless Communications*, vol. 14, no. 4, pp. 2269–2280, Dec. 2015.

[59] S. Qian, J. He, M. Juntti, and T. Matsumoto, "Fading correlations for wireless cooperative communications: Diversity and coding gains," *IEEE Access*, Apr. 2017.

[60] S. Qian, X. Zhou, X. He, J. He, M. Juntti, and T. Matsumoto, "Performance analysis for lossy-forward relaying over Nakagami-m fading channels," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 11, pp. 10 035–10 043, Nov. 2017.

[61] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, Jul. 1948.

[62] D. A. Huffman, "A method for the construction of minimum-redundancy codes," *Proceedings of the IRE*, vol. 40, no. 9, pp. 1098–1101, Sep. 1952.

[63] J. Ziv and A. Lempel, "A universal algorithm for sequential data compression," *IEEE Transactions on Information Theory*, vol. 23, no. 3, pp. 337–343, May 1977.

[64] ——, "Compression of individual sequences via variable-rate coding," *IEEE Transactions on Information Theory*, vol. 24, no. 5, pp. 530–536, Sep. 1978.

[65] ISO/IEC International Standard 11172-3, "Coding of moving pictures and associated audio for digital storage media at up to about 15 Mbit/s - Part3: Audio," 1993.

[66] ISO/IEC International Standard 10918-1, "Information technology – Digital compression and coding of continuous-tone still images: Requirements and guidelines," 1992.

[67] ISO/IEC International Standard 14496-10, "Information technology – Coding of audio-visual objects – Part 10: Advanced video coding," 2003.

[68] C. E. Shannon, "Coding theorems for a discrete source with a fidelity criterion," *IRE Nat. Conv. Rec*, vol. 4, no. 142-163, p. 1, Mar. 1959.

[69] S. D. Servetto, K. Ramchandran, V. A. Vaishampayan, and K. Nahrstedt, "Multiple description wavelet based image coding," *IEEE Transactions on Image Processing*, vol. 9, no. 5, pp. 813–826, 2000.

[70] W. Jiang and A. Ortega, "Multiple description speech coding for robust communication over lossy packet networks," in *IEEE International Conference on Multimedia and Expo*, vol. 1. New York, NY, USA: IEEE, Jul. 2000, pp. 444–447.

[71] Y. Wang, A. R. Reibman, and S. Lin, "Multiple description coding for video delivery," *Proceedings of the IEEE*, vol. 93, no. 1, pp. 57–70, 2005.

[72] J.-J. Xiao and Z.-Q. Luo, "Multiterminal source–channel communication over an orthogonal multiple-access channel," *IEEE Transactions on Information Theory*, vol. 53, no. 9, pp. 3255–3264, Sep. 2007.

[73] W. Lin, X. He, M. Juntti, and T. Matsumoto, "Binary data gathering with a helper in Internet of Things: Distortion analysis and performance evaluation," *IEEE Access*, vol. 7, pp. 12 855–12 867, Jan. 2019.

[74] W. Lin and T. Matsumoto, "Performance analysis of distortion-acceptable cooperative communications in wireless sensor networks for Internet of Things," *IEEE Sensors Journal*, vol. 19, no. 5, pp. 1979–1989, Mar. 2019.

[75] W. Lin, S. Qian, and T. Matsumoto, "Lossy-forward relaying for lossy communications: Rate-distortion and outage probability analyses," *IEEE Transactions on Wireless Communications*, vol. 18, no. 8, pp. 3974–3986, Aug. 2019.

[76] W. Lin and T. Matsumoto, "An efficiency-complexity controllable rate adaptive lossy source coding—Hybrid majority voting code," *IEEE Communications Letters*, vol. 22, no. 12, pp. 2419–2422, Dec. 2018.

[77] ——, "An analysis of performance improvement by a helper for wireless sensor networks," in *IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Bologna, Italy, Sep. 2018.

[78] R. Gallager, "Low-density parity-check codes," *IRE Transactions on Information Theory*, vol. 8, no. 1, pp. 21–28, Jan. 1962.

[79] J. Garcia-Frias and Y. Zhao, "Near-Shannon/Slepian-Wolf performance for unknown correlated sources over AWGN channels," *IEEE Transactions on Communications*, vol. 53, no. 4, pp. 555–559, Apr. 2005.

[80] X. Zhou, X. He, K. Anwar, and T. Matsumoto, "GREAT-CEO: larGe scale distRibuted dEcision mAking Technique for wireless Chief Executive Officer problems," *IEICE Transactions on Communications*, vol. 95, no. 12, pp. 3654–3662, Dec. 2012.

[81] A. El Gamal and Y.-H. Kim, *Network information theory*. Cambridge, UK: Cambridge University Press, 2011.

[82] R. Youssef and A. G. i. Amat, "Distributed serially concatenated codes for multi-source cooperative relay networks," *IEEE Transactions on Wireless Communications*, vol. 10, no. 1, pp. 253–263, Jan. 2011.

[83] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate (corresp.)," *IEEE Transactions on Information Theory*, vol. 20, no. 2, pp. 284–287, Mar. 1974.

[84] M. R. Souryal and B. R. Vojcic, "Performance of amplify-and-forward and decode-and-forward relaying in rayleigh fading with turbo codes," in *IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, vol. 4, Toulouse, France, May 2006, pp. 681–684.

[85] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Transactions on Information Theory*, vol. 51, no. 9, pp. 3037–3063, Aug. 2005.

[86] S. Qian, V. Tervo, J. He, M. Juntti, and T. Matsumoto, "A comparative study of different relaying strategies over one-way relay networks," in *22th European Wireless Conference*, Oulu, Finland, May 2016, pp. 1–6.

[87] J. Hagenauer, "The EXIT chart – introduction to extrinsic information transfer in iterative processing," in *12th European Signal Processing Conference*. Vienna, Austria: IEEE, Sep. 2004, pp. 1541–1548.

[88] R. W. Hamming, "Error detecting and error correcting codes," *The Bell System Technical Journal*, vol. 29, no. 2, pp. 147–160, Apr. 1950.

[89] E. Arikan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Transactions on Information Theory*, vol. 55, no. 7, pp. 3051–3073, Jul. 2009.

[90] X. He, *Binary information sensing and multiterminal source coding: Rate-distortion analysis and transmission design*. Nomi, Japan: JAIST Press, Jun. 2016.

Publications

[1] W. Lin, S. Qian, and T. Matsumoto, "Lossy-forward relaying for lossy communications: Rate-distortion and outage probability analyses," *IEEE Transactions on Wireless Communications*, vol. 18, no. 8, pp. 3974–3986, Aug. 2019.

[2] W. Lin and T. Matsumoto, "Performance analysis of distortion-acceptable cooperative communications in wireless sensor networks for Internet of Things," *IEEE Sensors Journal*, vol. 19, no. 5, pp. 1979–1989, Mar. 2019.

[3] W. Lin, X. He, M. Juntti, and T. Matsumoto, "Binary data gathering with a helper in Internet of Things: Distortion analysis and performance evaluation," *IEEE Access*, vol. 7, pp. 12 855–12 867, Jan. 2019.

[4] W. Lin and T. Matsumoto, "An efficiency-complexity controllable rate adaptive lossy source coding—Hybrid majority voting code," *IEEE Communications Letters*, vol. 22, no. 12, pp. 2419–2422, Dec. 2018.

[5] W. Lin and T. Matsumoto, "An analysis of performance improvement by a helper for wireless sensor networks," in *IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Bologna, Italy, Sep. 2018.