XIN HE

BINARY INFORMATION SENSING AND MULTITERMINAL SOURCE CODING: RATE-DISTORTION ANALYSIS AND TRANSMISSION DESIGN

Academic dissertation to be presented with the assent of the Doctoral Training Committee of Technology and Natural Sciences of the University of Oulu for public defence in Collaboration room 6, Asahidai, on 10 June 2016, at 3 p.m.

UNIVERSITY OF OULU
and
JAPAN ADVANCED INSTITUTE OF SCIENCE AND TECHNOLOGY
2016

Supervised by
Professor Markku Juntti
Professor Tad Matsumoto

Reviewed by
Professor Gerhard Kramer
Professor Antoine Berthet
Professor Ping Li
Associate Professor Brian Kurkoski
Professor Mineo Kaneko

ISBN 978-4-903092-45-4 (Paperback)

He, Xin, Binary Information Sensing and Multiterminal Source Coding: Rate-Distortion Analysis and Transmission Design.
University of Oulu Graduate School; University of Oulu, Faculty of Information Technology and Electrical Engineering; Centre for Wireless Communications-Radio Technologies (CWC-RT); Infotech Oulu; Japan Advanced Institute of Science and Technology, School of Information Science, Information Theory and Signal Processing Laboratory

Abstract

A binary chief executive officer (CEO) problem is investigated with the application to an information sensing network, where noise-corrupted versions of a binary sequence are forwarded by a group of sensors to a single destination over orthogonal multiple access channels. The primary goal of this thesis is to provide theoretical analysis of the system performance and to design practical transmission techniques by applying recent results of multiterminal rate-distortion theory.

Concatenated convolutional codes and a joint decoding scheme are proposed for the binary information sensing network. The performance limits for threshold signal-to-noise ratio (SNR) and the bit error probability (BEP) floor are analytically derived using the Slepian-Wolf theorem and the Poisson binomial process, respectively. Furthermore, a BEP floor lower bound is obtained from the classical rate-distortion theory. We further derive an outer bound on the rate-distortion region for the binary CEO problem through the multiterminal source coding. Derived outer bound is applied to the problem of acquiring the lower bound on the Hamming distortion in the framework of convex optimization for arbitrary value of SNR. Computer simulations are carried out to verify the theoretical results.

Furthermore, an optimal power allocation scheme is proposed for the binary information sensing network from rate-distortion perspective. Based on the simulation results, our proposed power allocation scheme outperforms the uniform power allocation method.

Keywords: CEO problem, multiterminal source coding, rate-distortion, achievable rate, Hamming distortion, sensor network

To my family

Preface

This research work has been carried out under a curriculum that is organized by the Collaborative Education Program of Centre for Wireless Communications (CWC), University of Oulu, and Japan Advanced Institute of Science and Technology (JAIST), Japan from 2013-2016. At JAIST, the research was carried out at Information Theory and Signal Processing Laboratory, School of Information Science. At University of Oulu, the work was conducted at Department of Communications Engineering (DCE) and CWC.

First of all, I would like to express my sincere gratitude to my principal advisors Professor Tad Matsumoto and Professor Markku Juntti for offering me the opportunities to pursue the Ph.D. degree at two universities. Their guidance, advice, and support for how to conduct research work as well as write scientific papers always impress me. Without their valuable discussions and comments, I would not have completed this research work smoothly. I look forward to further collaboration in the future.

I am grateful to the reviewers and examiners, Professor Gerhard Kramer from Technical University of Munich, Germany, Professor Antoine Berthet from CentraleSupélec, France, Professor Ping Li from City University of Hongkong, and Professor Brian Kurkoski, Professor Mineo Kaneko from JAIST, for their valuable comments and suggestions to improve the quality of the thesis.

I wish to express my special thanks to Dr. Zongliang Zhuansun. Without his guidance and help, I could not have the chance to study abroad. He shared with me a large amout of knowledge in computer science when I was an undergraduate student. His enthusiasm and preciseness in doing academic research encouraged me and I decided to proceed with my postgraduate study in Japan. I want to thank Professor Yasutada Oohama from University of Electro-Communications (UEC), Japan for his valuable discussions on the rate-distortion theory when I visited him at UEC.

I also would like to thank my current and previous colleagues at JAIST and CWC for their helps not only in the research but also the daily life. My special thank goes to Dr. Xiaobo Zhou, Dr. Meng Cheng and Dr. Penshun Lu, who helped me a lot during my studies. We worked together and made fruitful joint publications. The special thank also extends to Mr. Weiwei Jiang, Mr. Shen Qian, Mr. Kun Wu, Mr. Sekiya Ryouta, Dr. Reza Asvadi, Dr. Valtteri Tervo, Dr. Petri Komulainen, Dr. Antti Tolli, Dr. Marian Codreanu, Dr. Iqbal Hussain and Dr. Khoirul Anwar for their helpful discussions.

I would like to thank the administrative staff Kirsi Ojutkangas, Aya Inoue, Tomoko Taniguchi and Uiko Hashimoto for their help.

Many thanks to my friends, Xin Wang, Kai Wang, Yini Wang, Yiran Wang, Bo Li, Yingyue Xu, Jiawei Lin, Yuping Feng, Xin Fang, Xiaodong Liu, Zeyun Zhu, Guangzhe Huang, Penshun Lu at University of Oulu, and Xiaobo Zhou, Meng Cheng, Weiwei Jiang, Jiajie Xue, Erick Gar, Fan Zhou, Tong Huang, Xiaoguang Li, Xuyang Zhao, Mihray Ahmet, Xinzhu Wang, Debin Zhang and Jiuyang Zhang at JAIST. It was very impressive and joyful to talk and gather with them.

This research work has been funded by several projects and programs, including JAIST doctoral researcher fellowship (DRF) program, Links-on-the-fly Technology for Robust, Efficient and Smart Communication in Unpredictable Environments (RESUCE) project funded by European Commission (EC) under 7th Framework Program (FP7), Network compression based wireless cooperative communication systems (NETCOBRA), Slepian-Wolf Cooperation Wireless Networks (SWOCNET), and Sensing, Compression, Communications and Data Fusion in Wireless Sensor Networks (SeCoFu) projects funded by the Academy Finland as well as by a personal scholarship from Nokia Foundation. I want to take this opportunity to thank their financial support which enabled me to devote my time to research.

I am grateful to my grandparents Jiegao He and Gentao Jin, and my father Yanliang He for their endless love and support to my daily life. Also, I want to show my special thank to Qing He, Hongliang Sun, Yayu Sun, Caixia He, Jiegen He, Yongkang He, Chunxia He, Weiguo He, Tao He, Guixia He, Wenxia He, Changzhou Wang, Hui Wang, Jian Wang, Songe Wang, Guohua He, Ainan He, Yin Han, Qian Ge, Fengqing Xu for their support and sharing happiness. For sure, I miss you, my mother Jiee Jin.

Finally, I would like to express my deepest gratitude to my wife, Han Ge for her love, understanding, support and encouragement.

Without my family members, I could not have chance to continue my study and research, hence, this thesis is dedicated to them.

Nomi, May 13th, 2016 Xin He

List of abbreviations

ACC	Doped-accumulator
ACC^{-1}	Decoder of doped-accumulator
AWGN	Additive white Gaussian noise
BCJR	Bahl-Cocke-Jelinek-Raviv
BEP	Bit error probability
BER	Bit error rate
BICM-ID	Bit-interleaved coded-modulation with iterative decoding
BPSK	Binary phase-shift keying
BSC	Binary symmetric channel
CEO	Chief executive officer
CF	Compress-and-forward
CI	Conditionally independent
DEC	Decoder
DeM	Demodulator
DF	Decode-and-forward
DPI	Data processing inequality
DSC	*Distributed source coding*
ENC	Encoder
EPI	Entropy power inequality
EXIT	Extrinsic information transfer
FER	Frame error rate
i.i.d.	Independently and identically distributed
JSC	Joint source-channel
KKT	Karush-Kuhn-Tucker
LDPC	Low-density parity-check
LF	Lossy forwarding
LLR	Log-likelihood ratio
MAC	Multiple access channel
MGL	Mrs. Gerber's lemma
MSP	Modified set partition
NOMA	Non-orthogonal multiple access

P2P	Point-to-point
PA	Power allocation
pdf	Probability density function
pmf	Probability mass function
QAM	Quadrature amplitude modulation
QoS	Quality of service
QPSK	Quadrature phase-shift keying
RV	Random variable
SNR	Signal-to-noise ratio
TDMA	Time division multiple access
VQ	Vector quantization
WMN	Wireless mesh network
WSN	Wireless sensor network
XOR	Exclusive-OR
3D	Three-dimensional

List of Symbols

$\mathscr{A}^{\mathrm{C}}$	complementary set of $\mathscr{A}$
B	random variable that generate binary noise
$\mathbf{b}$	sequence of the binary noise
C	channel capacity
CC_i	channel encoder at sensor i
DCC_i	decoder of CC_i
D	distortions
i	sensor index
$I^{\mathrm{a}}_{\mathbf{c}_i}$	mutual information by evaluating a priori LLR sequence of $\mathbf{c}_i$
$I^{\mathrm{e}}_{\mathbf{x}_i}$	mutual information by evaluating extrinsic LLR sequence of $\mathbf{x}_i$
$I^{\mathrm{e}}_{\mathbf{c}_i}$	mutual information by evaluating extrinsic LLR sequence of $\mathbf{c}_i$
L	the number of terminals (sensors)
$\mathscr{L}$	set $\{1,2,\cdots,L\}$
$\mathscr{L}\setminus i$	remove i from set $\mathscr{L}$
$\mathbf{l}^{\mathrm{c}}_i$	channel LLR sequence of link i
$\mathbf{l}^{\mathrm{a}}_{\mathbf{x},i}$	a priori LLR sequence on $\mathbf{x}$ to link i
$\mathbf{l}^{\mathrm{a}}_{\mathbf{x}_i}$	extrinsic LLR sequence on $\mathbf{x}_i$
$\mathbf{l}^{\mathrm{e}}_{\mathbf{x},i}$	extrinsic LLR sequence on $\mathbf{x}$ of link i
$\mathbf{l}^{\mathrm{e}}_{\mathbf{c}_i}$	extrinsic LLR sequence on $\mathbf{c}_i$
$\mathbf{l}^{\mathrm{e}}_{\mathbf{x}_i}$	extrinsic LLR sequence on $\mathbf{c}_i$
$\mathbf{l}^{\mathrm{e}}_{\mathbf{x},i}$	extrinsic LLR sequence on $\mathbf{x}$ of link i
$\mathbf{l}^{\mathrm{p}}_{\mathbf{x}_i}$	a posteriori LLR sequence on $\mathbf{x}_i$
$\mathbf{l}^{\mathrm{p}}_{\mathbf{x}}$	a posteriori LLR sequence on $\mathbf{x}$
P_d	doping ratio
p_i	observation error probability of sensor i
$p_X(x)$	probability mass function of X
$p_{XY}(x,y)$	joint probability mass function of X and Y
$p_{X\|Y}(x\|y)$	conditional probability mass function of X given Y
p_{lb}	lower bound on bit error floor
R	source coding rate
R^{c}	channel coding rate

$\mathscr{R}^{\mathrm{o}}$	outer bound on rate region
$\mathscr{R}^{\mathrm{i}}$	inner bound on rate region
$\mathscr{R}^{\mathrm{SW}}$	Slepian-Wolf rate region
$\mathscr{R}_0^{\mathrm{WZ}}(D)$	Wyner-Ziv rate-distortion region
$\mathscr{R}^{\mathrm{YI}}$	Oohama's rate-distortion bound of Gaussian multiterminal source coding
r	end-to-end coding rate
$\mathscr{S}$	subset of $\mathscr{L}$
$\mathbf{s}$	modulated signal sequence
t	timing index
U_i, W	auxiliary variables in source coding problem
X	random variable
X^n	vector of random variables $[X(1), X(2), \cdots, X(n)]$
$X_{\mathscr{L}}$	a set of random variables $\{X_i \mid i \in \mathscr{L}\}$
$\mathscr{X}$	alphabet set of random variable X
$\mathbf{x}$	binary information sequence $[x(1), x(2), \cdots, x(n)]$ generated by the sensing object
$x(t)$	t-th component of $\mathbf{x}$
$\mathbf{y}$	received signal sequence
Z	Gaussian random variable (noise)
$\mathbf{z}$	AWGN sequence
φ	encoder fucntion
ψ	decoder function
Π	interleaver
ε	arbitrary small positive number
σ^2	variance of Gaussian random variable
Σ	covariance matrix of correlated Gaussian source
ρ	correlation parameter
δ	convergence control threshold
$\hat{\cdot}$	estimation
η_i	parametrization variable
λ_i	Lagrange multiplier for inequality constraint
μ	Lagrange multiplier for equality constraint

$Bern(p)$	Bernoulli distribution which takes value 1 with probability p
$CN(\mu, \sigma^2)$	complex Gaussian distribution with mean μ and variance σ^2 per dimension
$C(\cdot)$	Shannon channel capacity

$d(a,b)$	distortion measure function between a and b
$E[\cdot]$	the expectation of random variable
exp	natural exponential function
$f_c(\cdot)$	LLR updating function
$g(\cdot)$	rate-distortion function used in binary Wyner-Ziv bound
$h(\cdot)$	joint entropy function of correlated binary sources
$H(X)$	entropy of random variable X
$H(X,Y)$	joint entropy of random variables X,Y
$H(X\|Y)$	conditional entropy of X given Y
$H_2(\cdot)$	binary entropy function
$H_2^{-1}(\cdot)$	inverse of binary entropy function $[0,1]\mapsto[0,0.5]$
$I(X;Y)$	mutual information between X and Y
inf$[\cdot]$	the infimum of a subset
$\ln(\cdot)$	natural logarithm
$\log(\cdot)$	logarithm function with base 2
$\lg(\cdot)$	logarithm function with base 10
$Log(\cdot)$	logarithm distribution
min	minimization
max	maximization
PB$(\cdot)$	function of calculating bit error floor using Poisson binomial process
$R(D)$	rate-distortion function
$R^{\mathrm{CEO}}(D)$	rate-distortion function of quadratic CEO problem
$\Re(\cdot)$	real part of a complex value
sup$[\cdot]$	the supremum of a subset
$T(\cdot,\cdot)$	extrinsic information transfer chart transform function
$\|\mathscr{X}\|$	cardinality of a set
$\|a\|$	absolute value of a
$a*b$	binary convolution, i.e., $a*b=a(1-b)+b(1-a)$
$[a]^+$	$\max\{0,a\}$
$\log^+(a)$	$\max\{0,\log(a)\}$
$[a]^-$	$\min\{1,a\}$

Contents

1 Introduction

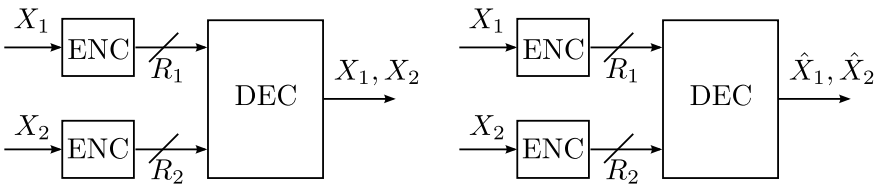
Wireless sensor networking (WSN) has attracted significant attention by the sensing and wireless communication research communities, and their practical deployment is also gaining momentum. The applications of WSNs are numerous, and they can be deployed in unpredictable environments to perform various distributed sensing tasks, for which it is extraordinarily important to design good transmission and scheduling techniques in order to make the sensor network highly energy-efficient. The information gathered by different sensors are often correlated. Therefore, *distributed source coding* (DSC) techniques inspired, e.g., by the Slepian-Wolf theorem, can provide an effective framework to design the networks efficiently. This can be utilized to decrease the transmission rate or to reduce the transmit power. Xiong *et al*. discussed the DSC schemes for WSNs in the framework of the Slepian-Wolf theorem in a tutorial paper [1]. This thesis deals with the design of transmission techniques for a particular WSN, which is the binary information sensing network, and the theoretical analysis of this network based on multiterminal information theory in the category of DSC.

This chapter begins with a historical review on classical problems of DSC and recent results. Afterwards, the motivation of this research work is briefly introduced together with a literature review of relevant work. Finally, the outline of this thesis is provided, followed by the author's contributions.

1.1 Background and History

A number of emerging applications, such as WSN, wireless mesh network (WMN), wireless cooperative relaying networks and wireless video, involve multiple correlated sources, which are required to be independently compressed at distributed terminals. In such applications, distributed terminals usually cannot communicate with each other due to the constraints on power and complexity. *Distributed source coding*, which handles the problem of compressing multiple correlated sources, has thus gained a lot of attention during the last four decades.

The fundamental question of DSC is to determine the tradeoff between the encoding rates and the fidelities of the recovered correlated sources. It is well known that this tradeoff has been mathematically investigated in the framework of rate-distortion analysis in information theory. The study of rate-distortion theory becomes extremely



(a) Lossless case. **(b) Lossy case.**

Fig. 1. Distributed source coding problem with two terminals.

important since it provides us with the design criteria for the design of the transmission techniques of communication networks.

As shown in Fig. 1, there are two categories of DSC problem from the level of distortion perspective; one is *lossless* DSC and the other *lossy* DSC. The level of distortion in the lossless DSC is required to be arbitrarily small while that in the lossy DSC is allowed at a certain value with a specified distortion measure function. The problem shown in Fig. 1 is also referred to as multiterminal source coding, which has a rich history.

There is a further significant distinction in DSC from the viewpoint of whether or not the sources are directly available at the encoders, which is *direct* versus *remote* DSC. In the direct DSC, the encoder directly accesses the source which is of interest by the decoder, while in the latter one, the encoder only has the noisy version of the source through an observation process. There is a famous problem in the category of remote DSC, which is so-called chief executive officer (CEO) problem. This dissertation focuses on the CEO problem and its application to binary information sensing network.

1.1.1 Lossless Distributed Source Coding

In 1973, Slepian and Wolf started the pioneering work for establishing the fundamental lossless DSC theorem with two correlated sources [2]. The admissible rate region of the compression rate pair is determined for which two sources can be reconstructed with an arbitrarily small error probability (which is referred as losslessly in usual Shannon sense). Surprisingly, the Slepian-Wolf theorem showed that two correlated sources X_1 and X_2 can be losslessly recovered with joint decoding, as long as the compression rate of each source and the sum rate are larger than their conditional entropy and their joint entropy, respectively.

Based on the setup of the Slepian-Wolf theorem, various extensions have been done by researchers. The Slepian-Wolf theorem was generalized by Cover [3] to arbitrary number of correlated sources. Ahlwede and Körner focused on a specific case, where the decoder is only interested in recovering one of the sources, let's say, X_2, and the other source X_1 acts as the side information. This problem is referred as source coding with side information. The rate region was then derived in [4] through the direct coding and converse proof. Wyner independently studied the source coding with side information problem in [5], where two decoders reconstruct two sources with the help of the common side information. A specific of the source coding with side information problem where a primary source is the exclusive-OR (XOR) version of two helper information was studied by Körner and Marton [6]. On the other hand, the practical design of the Slepian-Wolf coding is also an active research topic, such as using low-density parity-check (LDPC) codes [7], turbo codes [8] and rateless codes [9].

Gel'fand and Pinsker [10] examined the rate region for perfectly reproducing an underlying source via corrupted observations of the source based on an important assumption, that the observations are conditionally independent (CI) given the underlying source. This problem is also called lossless CEO problem, which has great impact on solving the CEO problem in lossy DSC.

1.1.2 Lossy Distributed Source Coding

As shown in Fig. 1(b), we review another particular case of the DSC problem where the decoder only needs to recover the sources at certain distortions with a distortion measure function, for example, Hamming distortion.

Wyner and Ziv studied the source coding problem with side information available at the decoder, in which the main source is required to be recovered within a fidelity criterion [11]. The rate-distortion function was derived and it showed that the required transmission rate with Wyner-Ziv source coding is larger than that of the side information being available both at the encoder and decoder. This finding, however, is in contrast to the Slepian-Wolf theorem, where knowledge of side information at the encoder does not result in a reduction of the transmission rate. Subsequently, Wyner generalized the Wyner-Ziv coding problem to nondiscrete sources [12]. Recently, the Wyner-Ziv coding was extended in various directions. Gastpar considered the Wyner-Ziv coding with multiple correlated sources [13], where an inner bound of achievable rate region and a matching outer bound are derived. Draper derived the rate-distortion

function of the remote Wyner-Ziv coding problem, where the encoder accesses to noisy observations of the main source [14]. The Wyner-Ziv coding was intensively studied with the application to wireless video networks, such as [15–18], and to the wireless relaying network as an optimal compress-and-forward (CF) strategy [19–21].

Berger and Tung characterized the inner and outer bounds on the rate-distortion region for a multiterminal source coding problem [22, 23], where correlated sources are independently encoded and jointly decoded, as shown in Fig. 1(b). Oohama studied the rate-distortion theory for multiterminal source coding with correlated memoryless Gaussian sources and squared distortion measures [24]. He derived an explicit form of the outer bound on the rate-distortion region, and showed that Berger-Tung inner bound is partially tight. Furthermore, he gave a rigorous proof of direct coding theorem using random coding arguments originated by Berger [22], Han and Kobayshi [25]. Yeung considered the multiterminal source coding problem with only one distortion criterion and derived the rate-distortion bound [26]. Oohama [27] and Pandya *et al*. [28] independently examined the rate-distortion bound of a Guassian multiterminal source coding problem, where multiple correlated Gaussian sources are encoded in a distributed sense, and one of them is the source of interest by the decoder (many-help-one problem). Wagner recently derived an improved outer bound on the rate-distortion region for the multiterminal source coding problem [29].

The CEO problem, which is meaningful in the category of remote multiterminal source coding, has intensively studied. Berger *et al*. originated the terminology CEO problem from the viewpoint of distributed communication/estimation[30]. The CEO problem is described as follows. A CEO is interested in a underlying source that cannot be observed directly. A group of L agents (terminals) is employed by the CEO to independently observe the underlying source and generate the corrupted versions of the source. Each agent then independently encodes the noisy observations under a sum rate constraint R and transmit the encoded information to a data processing center. Finally, the CEO reproduces the estimates of the underlying source as accurate as possible. The major objective of studying the CEO problem is to determine the tradeoff between the sum rate and distortion when L goes to infinity.

The asymptotic behavior of the minimum distortion function with respect to the rate was determined in the limit as L and R tend to infinity, for discrete memoryless sources [30]. Viswanathan and Berger examined the rate-distortion behavior for the quadratic Gaussian CEO problem, where the source and observations are assumed to jointly Gaussian distributed [31]. Oohama further derived an explicit form of the rate-distortion

Table 3. The historical problems of the distributed source coding.

Lossless DSC	Lossy DSC
	Wyner-Ziv coding [11]
	generalized Wyner-Ziv coding [12]
	Berger-Tung inner and outer bounds [22, 23]
Slepian-Wolf coding [2]	source coding with side Information [26]
generalized Slepian-Wolf coding [3]	Improved outer bound for multiterminal
source coding with helper [4, 5]	source coding [29]
lossless CEO problem [10]	Gaussian multiterminal source coding [24]
specified helper problem (XOR) [6]	Many-help-one problem [28]
	CEO problem [30]
	Gaussian CEO problem [36]
	vector Gaussian CEO problem [33, 34]

function for the quadratic Gaussian CEO problem using the CI property. Ekrem *et al*. and Wang *et al*. independently provided an outer bound on the rate-distortion region for the vector Gaussian CEO problem, where a vector Gaussian source is observed by agents [32–34]. A binary erasure CEO problem, where a binary source is observed through multiple independent erasure channels by the agents, was studied by Wagner and Anantharam [35].

In summary, Table 3 shows the important problems in the category of DSC.

1.2 Motivation

The DSC problem has strong connections with the design of wireless networks, as well as distributed estimation tasks. We consider the information sensing network and its connection to the multiterminal source coding problem with the aim of providing the analytical assessment to the system performance and designing optimal power allocation schemes and/or scheduling protocols.

1.2.1 The state-of-the-art

Information sensing is straightforwardly modeled by the CEO problem in network information theory. As shown in Fig. 2, an example of information sensing network is deployed to monitor the environment. For such networks, power is usually limited, and

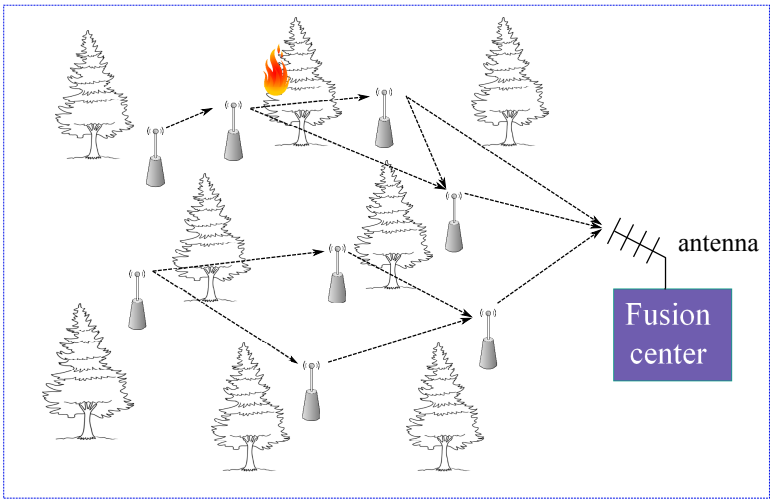


Fig. 2. An example of information sensing network.

hence, designing transmission techniques, which exploit correlation knowledge from the view of DSC, become important.

Gaussian case

Correlated data gathering problem in the context of sensor networks was considered in [37], and the Slepian-Wolf theorem was applied to result in an optimal coding strategy.

Side information aware coding strategies based on a generalization of Wyner-Ziv coding were proposed for a tree-structured sensor network [38], where serial and parallel networks are emphasized. Behroozi *et al*. obtained the optimal distortion and sum-rate tradeoff using successive coding strategy for the quadratic Gaussian CEO problem [39, 40], and proposed an optimal rate allocation scheme [41]. Successive Wyner-Ziv coding was applied to the quadratic Gaussian CEO problem to achieve the every point in the rate region [42], based on the nature that the rate-distortion region is contra-polymatroid [43, 44].

In [45], the authors investigated the joint source-channel decoding for the Gaussian sensor network, where the correlation among the observation data is regarded as a global code. They explored the method of how to design the decoder of this global code,

and then the decoder is concatenated with the decoder for the applied error correcting codes at sensor nodes to form the joint source-channel decoding process.

Joint source-channel coding strategy for a wireless sensor network under delay, data buffer size and energy constraints was further considered in [46]. The compressed samples generated by a time varying Gaussian source are forwarded to a destination over a fading channel in several time slots. For such scenario, an optimal transmission policy was formulated in the framework of convex optimization from the rate-distortion viewpoint. A joint source-channel coding using lattices was further considered for the Gaussian sensor network with a large number of sensors [47]. By adopting lattices, higher dimensions of the channel space were utilized to asymptotically achieve the Wyner-Ziv bound.

The problem of combining the quadratic Gaussian CEO problem and the multiple access channel (MAC) was investigated in [48], with the application to Gaussian relaying network.

Binary case

For binary data gathering sensor nodes, a coding scheme using convolutional code and an iterative joint decoding algorithm, was proposed in [49]. The joint decoding was divided into two steps. One is horizontal iteration which performs the Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm, while the other one is vertical iteration that exploits the correlation by evaluating the joint probability mass function (pmf) of the common binary source and the observations.

A coding scheme based on the parallel concatenated convolutional codes was proposed in [50], where the extrinsic log-likelihood ratio (LLR) sequence is weighted by the observation error probabilities at the decoder. In addition, the capacity of the equivalent parallel channel was derived to verify the bit error rate (BER) performance taking into account the error probability of the observed data sequence. However, this calculation restricted to the case that the observation error probability is identical.

In [51], an adaptive bi-modal decoder for a binary source estimation involving two sensors was proposed based on the modified extrinsic information transfer (EXIT) chart analysis. Convergence property analysis of the iterative decoding algorithm was presented based on the modified EXIT chart analysis for binary information sensing network in [52]. It shows that the iterative process is less useful if the channel quality is very good or the observation accuracy is very low.

In [53], the authors proposed an encoding/decoding technique which can significantly improve the BER performance by exploiting the correlation knowledge through the LLR updating function [54], for both a binary independently and identically distributed (i.i.d.) source and a binary Markov source. In [55], the authors proposed a non-negative constrained iterative algorithm for estimating the observation error probabilities in a WSN having an arbitrary number of sensors.

The rate-distortion behavior was investigated for a simple case of the binary CEO problem in [56], where the rate is equally allocated to the terminals. The distortion function with respect to the sum rate was derived in the limit as $L \to \infty$.

1.2.2 Beyond the state-of-the-art

As stated above, the major focus of this thesis is the binary CEO problem with the application to the binary information sensing network from the theoretical analysis to the practical scheme design. For the binary CEO problem, most of the previous work only concentrate on the design of practical encoding/decoding algorithms. This motivates us to analyze the binary CEO problem from the rate-distortion perspective using multiterminal source coding theory. The binary CEO problem is related to an important toy scenario in "Links-on-the-fly Technology for Robust, Efficient, and Smart Communication in Unpredictable Environments (RESCUE)" project. In this toy scenario, we consider the situation where a group of relays is deployed in the disaster area to collect information, however, none of relays has correct information. This is a basic toy scenario in RESCUE project and can be extended to more complicated cases, such as one-way relaying network and multi-way relaying network.

Therefore, we produce the following tasks toward solving the binary CEO problem.

- The rate-distortion region analysis of the binary CEO problem needs to be performed using the multiterminal source coding theory.
- The encoding/decoding algorithm should be designed to exploit the benefit of the network structure and the correlation knowledge within the structure.
- The problem of minimizing the transmission power allocated to all the sensors in the network as a whole, while keeping the quality of service (QoS) requirement, needs to be investigated. We can implement the results of the rate-distortion region analysis to this problem, where the transmission power corresponds to the rate and the QoS is regarded as the distortion level.

It should be emphasized here that the binary CEO problem can be also seen as the theoretical model of parallel relaying network that utilizes the lossy forwarding (LF) concept[1]. Hence, this work can be extended in the context of various wireless networks for designing efficient protocols and justifying system performance theoretically, particularly, the toy scenarios 1, 2 and 3 in RESCUE project [57].

1.3 Outline of the Dissertation

The goal of this dissertation is to provide theoretical analysis of the binary CEO problem and its application to binary information sensing network.

In Chapter 2, the background knowledge required in the theoretical analysis is summarized. We first review the basic concept of entropy and mutual information. Several important inequalities including Fano's inequality, Mrs. Gerber's lemma (MGL) and data processing inequality (DPI) are provided for the converse coding proof. Then, the classical results in multiterminal source coding, such as the Slepian-Wolf theorem, the Wyner-Ziv bound are included. After that, we briefly discuss the channel coding theorem and source-channel separation theorem. Finally, we provide the practical channel coding and extrinsic information transfer chart analysis, followed by the decision rules of binary information sensing.

In Chapter 3, the achievable rate and the bit error probability (BEP) floor of the binary information sensing network are investigated. The encoding and decoding algorithms are proposed for the binary information sensing network to exploit the correlation knowledge of sensing data. The main theoretical results are: 1) the signal-to-noise ratio (SNR) limit is converted from the achievable rate based on the Slepian-Wolf theorem and source-channel separation theorem in orthogonal additive white Gaussian noise (AWGN) channels; 2) the BEP floor, which is a common phenomenon in the binary information sensing network caused by the observation error, is analytically calculated using the Poisson binomial process and binary rate-distortion function. Then, a series of computer simulations is conducted to verify these limits, including the SNR limit and the BEP floor. Finally, a three-dimensional (3D) EXIT chart analysis is performed to confirm the simulation results.

In Chapter 4, we derive a theoretical lower bound on the Hamming distortion from rate-distortion viewpoint for the binary information sensing network with joint source-

[1]LF protocol is originated from the decode-and-forward (DF) relaying strategy. The relay node always keeps activated to forward the decoded message regardless of the error being introduced.

channel (JSC) setup. The objective of this network is to estimate a single binary source over orthogonal AWGN channels by performing independent coding at each sensor node and joint decoding at the fusion center. For the simplicity of analysis, we focus on orthogonal transmissions from L sensors/terminals to the fusion center, and we separate the stages of JSC decoding and the final decision on the common source [52, 58, 59]. Hence, deriving the theoretical lower bound on the Hamming distortion is equivalent to minimizing a distortion function subject to a series of inequalities obtained based on the source-channel separation theorem for lossy source coding.

First, a specified example of which the number of terminals $L = 2$ is discussed in detail. The binary information sensing network is modeled as a binary CEO problem. In order to address the two-terminal binary CEO problem, we consider a more general problem, which is a binary multiterminal source coding. We then derive an outer bound on the rate-distortion region for the binary multiterminal source coding through the converse coding proof. The derived outer bound is briefly compared to the classical regions including the Slepian-Wolf theorem, the Wyner-Ziv coding and Berger-Tung inner bound. The relationship between the binary CEO problem and the binary multiterminal source coding is connected through a distortion function. The theoretical lower bound on the Hamming distortion is formulated as a convex optimization problem that obtains the minimum values of distortions of the binary multiterminal source coding and map them back to the distortion of the binary CEO problem. Finally, computer simulations using our proposed encoding/decoding algorithm are performed to check the theoretical lower bound on the Hamming distortion.

After that, we extend the solutions of the case $L = 2$ to arbitrary number of L. We establish an outer bound on the rate-distortion region through the converse coding proof. However, the outer bound is an approximation since the proof of using test channel to bound the mutual information is lacking. Nevertheless, the lower bound on the Hamming distortion using the approximated outer bound can still act as an useful reference in verifying the performance of designed encoding/decoding algorithms for binary information sensing network. Furthermore, the superiority of our proposed joint decoding algorithm is examined analytically with the comparison to the separate decoding scheme based on the rate-distortion region analysis.

In Chapter 5, an optimal power allocation scheme is proposed to minimize the Hamming distortion subject to a fixed total power in the context of a cluster binary information sensing network. We first formulate the minimization problem from the rate-distortion perspective. To avoid deriving the analytical form of the inverse of the

binary entropy function, we then reformulate the optimization problem in the framework of convex optimization by maximizing the weighted channel capacity. From the computer simulations, the proposed power allocation scheme gains around $1.5 \sim 2$ dB from the case of uniformly distributing the power.

In Chapter 6, we conclude this thesis and summarize the main results. We further provide some future studies and insight discussion of how to derive the rate-distortion function of the binary CEO problem directly.

1.4 Summary of Contributions

This thesis is written as a monograph based on two journal papers [60, 61], and two conference papers [62, 63]. The first journal paper [60] has already been published and the other one [61] is now under review. The author has had the main responsibility for performing the analysis, conducting simulations, and writing all the papers [60, 62, 63]. For the paper [61], the author has worked as the corresponding author who has provided the main idea, written Section I, and verified the simulation results. Other authors provided helps, comments and criticism during the writing processing.

Besides these publications, the author published another conference paper [64] and co-authored several papers in the relevant area [65–70] during his doctoral study. The author made contributions to Section IV and the revision of [65], Section III of [66], Section IV of [67], Section I of [68], and the outage probability of CF in [69]. The author performed the black-box simulations for [70]. Furthermore, the author has been involved in writing technical reports[2] [57, 71, 72] under the RESCUE project. The author contributed to Section 4.1 of [57], Chapter 2, Sections 3.1.1, 4.1, 4.2 and 5.1 of [71], Sections 2.2.1, 2.2.2 and 2.3.6 of [72], and provided the help in editing [57].

In summary, the main contributions of this thesis are summarized as follows.

– A simple encoding scheme and a joint decoding algorithm which exploit the correlation knowledge are proposed for the binary information sensing network. The Slepian-Wolf theorem is applied to analyze the theoretical limit on the threshold SNR. The BEP floor is also derived based on the Poisson binomial process and the binary rate-distortion function.
– An outer bound on the rate-distortion region for the binary multiterminal source coding is derived through the converse coding proof. An explicit form of the outer bound

[2]The technical report is not official publication.

is provided for the case $L = 2$ of the binary multiterminal source coding. Furthermore, an approximated outer bound for an arbitrary number of L is also proved.

– The binary CEO problem is connected with the binary multiterminal source coding problem in terms of the distortion level through established distortion functions.
– We obtain the theoretical lower bound on the Hamming distortion for the binary information sensing network with JSC setup, based on the derived outer bound and the source-channel separation theorem.
– An optimal power allocation scheme is proposed for the binary information sensing network by maximizing the weighted channel capacity. Based on computer simulations, it is found that our proposed power allocation scheme outperforms the uniform power allocation case.
– An insightful discussion on how to derive the rate-distortion region for the two-terminal binary CEO problem is provided.

2 Preliminaries

In this chapter, the necessary background knowledge for this research is reviewed. We first introduce the basic concepts of entropy and mutual information in Section 2.1. After that, in Section 2.2, several important inequalities, such as Fano's inequality and MGL, used in the converse proof of rate-distortion analysis are included. We then review several classical results in DSC problem in Section 2.3, such as Slepian-Wolf theorem, Berger-Tung inner and outer bounds. The Shannon's channel coding theorem and the source-channel separation theorem are briefly discussed in Section 2.4 and Section 2.5, respectively. The superiority of the proposed encoding scheme is analyzed using extrinsic information transfer chart in Section 2.6. Finally, the decision rules of majority voting and soft combining are described in Section 2.7.

2.1 Entropy and Mutual Information

In information theory, *entropy* was originally described by Shannon in [73] for communication system, which is the measure of the uncertainty of information. Consider a random variable (RV) X taking i.i.d. values from a finite alphabet $\mathscr{X}$ with a probability mass function (pmf) $p_X(x) = \Pr\{X = x\}$, The entropy of RV X is given as

$$H(X) = -\sum_{x\in\mathscr{X}} p_X(x)\log(p_X(x)). \tag{1}$$

$H(X)$ does not take negative values with the definition $0\log 0 = 0$. The base of logarithm determines the unit of entropy, e.g., if the base is 2, entropy is expressed in bits.

The joint entropy $H(X,Y)$ of discrete random variables X and Y with a joint pmf $p_{XY}(x,y) = \Pr\{X = x, Y = y\}$ is defined as

$$H(X,Y) = -\sum_{x\in\mathscr{X}}\sum_{y\in\mathscr{Y}} p_{XY}(x,y)\log(p_{XY}(x,y)), \tag{2}$$

The conditional entropy of Y by given X can be further defined as

$$\begin{aligned} H(Y|X) &= \sum_{x\in\mathscr{X}} p_X(x)H(Y|X=x) \\ &= -\sum_{x\in\mathscr{X}} p_X(x)\sum_{y\in\mathscr{Y}} p_{Y|X}(y|x)\log(p_{Y|X}(y|x)) \\ &= -\sum_{x\in\mathscr{X}}\sum_{y\in\mathscr{Y}} p_{XY}(x,y)\log(p_{Y|X}(y|x)), \end{aligned} \tag{3}$$

where $H(Y|X) = 0$ if and only if the exact state of Y can be completely determined by X. Conversely, $H(Y|X) = H(Y)$ if and only if Y and X are independent.

The above definitions can be easily extended to the case of continuous RV's by taking integral instead of taking summation.

The chain rule of entropy [74] is simply obtained as

$$H(Y|X) = H(X,Y) - H(X). \tag{4}$$

The proof of (4) is a very basic training in information theory and hence we omitted.

The mutual information measures a quantity of the mutual dependence of two RV's, i.e., how much the information of a RV can be obtained from the other one. The mutual information between discrete RV's X and Y defined in (5) is the relative entropy between the joint distribution $p_{XY}(x,y)$ and the product distribution $p_X(x)p_Y(y)$.

$$I(X;Y) = \sum_{x\in\mathscr{X}} \sum_{y\in\mathscr{Y}} p_{XY}(x,y) \log\left(\frac{p_{XY}(x,y)}{p_X(x)p_Y(y)}\right), \tag{5}$$

where $p_X(x)$ and $p_Y(y)$ are the marginal probability distribution functions of X and Y. The mutual information $I(X;Y)$ indicates the amount of reduction in the uncertainty of X by knowing Y (or Y by knowing X), which is expressed as follows.

$$\begin{aligned} I(X;Y) &= H(X) - H(X|Y) \\ &= H(Y) - H(Y|X). \end{aligned} \tag{6}$$

2.2 Useful Inequalities

This section provides several important inequalities used in the converse proof of the outer bound on the rate-distortion region.

Fano's Inequality

Fano's inequality is used in the converse proof of many coding theorems. Fano's inequality, which is summarized in Lemma 1, provides the average information loss in terms of the error probability p_e of transmitting a source over a noisy channel.

Theorem 1 (Fano's inequality [74])**.** *Let a pair of RV's (X, Y) follow a joint pmf $p_{XY}(x,y)$ and $p_e = \Pr\{X \neq g(Y)\}$, where $g(\cdot)$ is a function that generate the estimation of X from the output Y. Then we have*

$$H(X|Y) \leq H_2(p_e) + p_e \log(|\mathscr{X}| - 1) \leq 1 + p_e \log(|\mathscr{X}| - 1), \tag{7}$$

with $|\mathscr{X}|$ representing the cardinality of the set $\mathscr{X}$ and $H_2(\cdot)$ being the binary entropy function, i.e., $H_2(a) = -a\log(a) - (1-a)\log(1-a)$.

The reason of the term $|\mathscr{X}| - 1$ is that if we observe an error event and $\hat{X}$, then X should be different to $\hat{X}$ and hence can only take at most $|\mathscr{X}| - 1$ values. It should be mentioned here that if the deterministic function $g(\cdot)$ does not take value from the same alphabet set $\mathscr{X}$, we should slightly modify (7) to

$$H(X|Y) \leq H_2(p_e) + p_e \log(|\mathscr{X}|) \leq 1 + p_e \log(|\mathscr{X}|). \tag{8}$$

Data Processing Inequality

The DPI states that the content of the information cannot increase through a physical process, which is a very important theoretical concept in information theory.

Theorem 2 (Data processing inequality [74])**.** *Suppose that RV's X, Y and Z forms a Markov chain as $X \to Y \to Z$, i.e., $p_{XYZ}(x,y,z) = p_{Z|Y}(z|y)p_{Y|X}(y|x)p_X(x)$, then the following inequalities hold.*

$$I(X;Y) \geq I(X;Z) \quad \textit{and} \quad I(Y;Z) \geq I(X;Z). \tag{9}$$

Equality holds if and only if $I(X;Y|Z) = 0$.

The DPI can be easily proved using the chain rule of mutual information and Markov property, i.e., X and Z are conditionally independent if Y is given. Intuitively, the DPI states that information will be lost if further process is taken.

Mrs. Gerber's Lemma

The MGL is usually used in the converse proof of coding theorems for the binary case, such as the capacity of the binary broadcast channel.

Lemma 3 (Mrs. Gerber's lemma [75])**.** *Let X and W be two RV's and X be binary. If a RV Y following a Bernoulli distribution with probability p, which is denoted as $Bern(p)$, is independent of (X,W) then*

$$H(Z|W) \geq H_2[H_2^{-1}(H(X|W)) * p], \tag{10}$$

where $Z = X \oplus Y$ is the XORed version of X and Y, and $H_2^{-1}(\cdot) : [0,1] \mapsto [0,0.5]$ represents the inverse function of $H_2(\cdot)$. The operation $*$ denotes the binary convolution process, i.e., $a * b = a(1-b) + b(1-a)$.

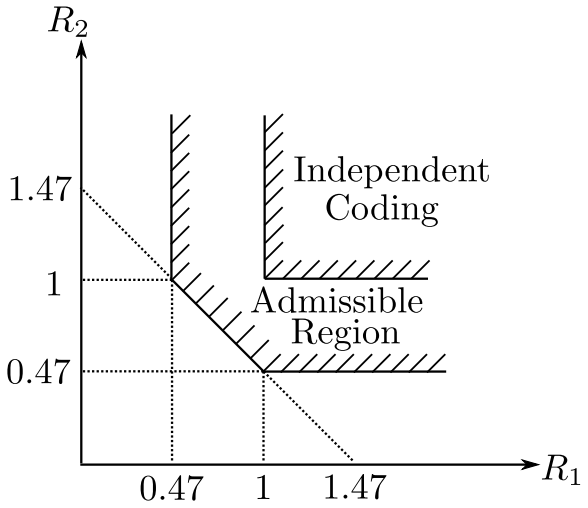


Fig. 3. Slepian-Wolf admissible rate region. X_1 follows $Bern(0.5)$, and X_2 is the output of a binary symmetric channel with crossover probability 0.1 when X_1 is the input.

The MGL is related to the entropy power inequality (EPI) introduced by Shannon [73]. However, the generalization of the MGL to multiple dependent RV's not only considering one W is difficult, according to our limited knowledge.

2.3 Theorems in Distributed Source Coding

The theorems, which were established in the previous work and are necessary in the analysis for our work, are summarized in this section.

2.3.1 Slepian-Wolf Theorem

Consider a source coding problem (as shown in Fig. 1(a)) where the decoder aims at perfectly reproducing two corrected sources, which are independently compressed at two terminals. Two correlated sources X_1 and X_2 make n independent drawing from a joint pmf $p_{X_1X_2}(x_1,x_2)$ to form two correlated information sequences X_1^n and X_2^n. Each encoder separately assigns an index from the set $\{1,2,\cdots,2^{nR_i}\}$ for the sequence X_i^n and sends the index to the decoder, $i=1,2$. After received the indices from the encoder, the joint decoder assigns an estimate $(\hat{X}_1^n,\hat{X}_2^n)\in\mathscr{X}_1^n\times\mathscr{X}_2^n$ or reports an error. The

probability of error is defined as

$$p_e^{(n)} = \Pr\{(\hat{X}_1^n, \hat{X}_2^n) \neq (X_1^n, X_2^n)\}. \tag{11}$$

A rate pair (R_1, R_2) is said to be admissible if there exists codes $(n, 2^{nR_1}, 2^{nR_2})$ such that $p_e^{(n)} \to 0$ with $n \to \infty$. Slepian and Wolf characterized the admissible rate region which is the closure of the set of admissible rate pairs of such problem.

Theorem 4 (Slepian-Wolf theorem [2])**.** *For two discrete memoryless sources X_1 and X_2 which draws i.i.d. from a joint pmf $p_{X_1X_2}(x_1, x_2)$, the admissible rate region $\mathscr{R}$ is given by*

$$\mathscr{R}^{SW} = \Big\{(R_1, R_2) : \begin{aligned} R_1 &\geq H(X_1|X_2) \\ R_2 &\geq H(X_2|X_1) \\ R_1 + R_2 &\geq H(X_1, X_2) \end{aligned} \Big\}.$$

The Slepian-Wolf theorem is proved based on *random binning*, which is a key concept in DSC for partitioning the outcomes of the random source. Some specific linear codes were constructed to achieve the Slepian-Wolf bound [76, 77].

Figure 3 shows the admissible rate region of compressing two binary sources X_1 and X_2 by applying the Slepian-Wolf theorem. As a reference, the rate region of two independent sources is also included in Fig. 3. From this example, the total minimal compression rate is significantly decreased compared to the independent case.

The Slepian-Wolf theorem was extended to an arbitrary number of correlated sources, as shown in the following theorem.

Theorem 5 (Generalized Slepian-Wolf theorem[78])**.** *In order to achieve lossless compression of L correlated sources $\{X_1, X_2, \cdots, X_L\}$, the source coding rate R_i should satisfy the following conditions*

$$\sum_{i \in \mathscr{S}} R_i \geq H(X_{\mathscr{S}}|X_{\mathscr{S}^{\mathrm{C}}}) \quad \textit{for all } \mathscr{S} \subseteq \{1, 2, \cdots, L\}, \tag{12}$$

where $\mathscr{S}^{\mathrm{C}} = \{1, 2, \cdots, L\} \setminus \mathscr{S}$ represents the complementary set of $\mathscr{S}$ and $X_{\mathscr{S}} = \{X_i | i \in \mathscr{S}\}$.

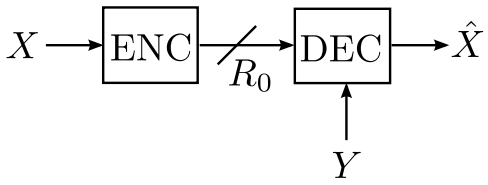


Fig. 4. The Wyner-Ziv source coding problem.

For example, the admissible rate region for the case $L = 3$ is

$$R_1 \geq H(X_1|X_2,X_3), \tag{13}$$
$$R_2 \geq H(X_2|X_1,X_3), \tag{14}$$
$$R_3 \geq H(X_3|X_1,X_2), \tag{15}$$
$$R_1+R_2 \geq H(X_1,X_2|X_3), \tag{16}$$
$$R_1+R_3 \geq H(X_1,X_3|X_2), \tag{17}$$
$$R_2+R_3 \geq H(X_2,X_3|X_1), \tag{18}$$
$$R_1+R_2+R_3 \geq H(X_1,X_2,X_3). \tag{19}$$

2.3.2 Wyner-Ziv Bound

The source coding problem which Wyner and Ziv examined in [5] is shown in Fig. 4. In this problem, a pair of dependent RV's (X,Y) produce i.i.d. sequences; the encoder assigns an index from the set $\{1,2,\cdots,2^{nR_0}\}$ for the source sequence X^n; the decoder reconstructs the estimates $\hat{X}^n$ using the side information generated by Y and the received index from the encoder. Wyner-Ziv derived the rate-distortion function $R_0^{\mathrm{WZ}}(D)$, i.e., the codes $(n,2^{nR_0})$ with rate R_0 achieves $\limsup_{n\to\infty} E[d(X^n,\hat{X}^n)] \leq D$, and provided the converse and direct coding proofs.

Theorem 6 (Wyner-Ziv bound [5, 79])**.** *Given a pair of discrete memoryless sources (X,Y) generates i.i.d. realizations, and there exists a RV Z such that X,Y and Z forms a Markov chain $Z \to X \to Y$ in this order. The decoder reproduce the estimates $\hat{X}$ based on Z and Y. Then a rate R is achievable if $R \geq R_0^{WZ}(D)$, where $R_0^{WZ}(D)$ is given as*

$$R_0^{WZ}(D) = \inf[I(X;Z|Y)], \tag{20}$$

with a distortion measure $D = E[d(X,\hat{X})]$ and $\inf[\cdot]$ representing the infimum of a subset.

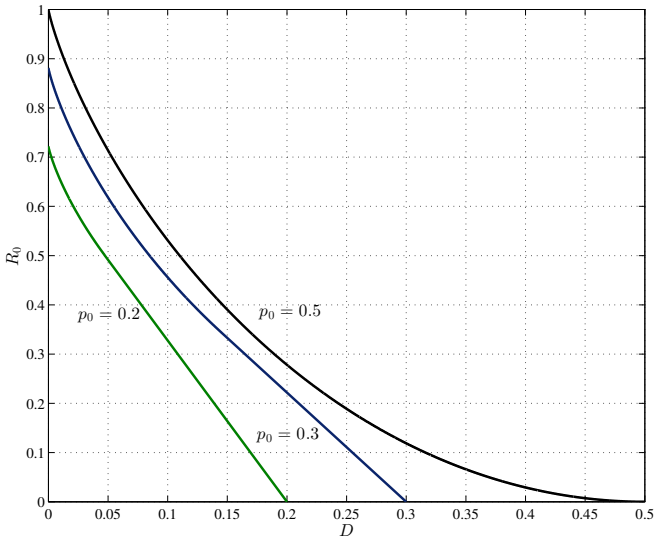


Fig. 5. Binary Wyner-Ziv bound on rate-distortion.

We now consider a specific example of Wyner-Ziv source coding where X follows $Bern(0.5)$, and Y is the output of a binary symmetric channel (BSC) with crossover probability being p_0 and the input X. The distortion measure is specified to Hamming distortion. As a consequence of applying Theorem 6, the following corollary is obtained.

Corollary 7 (Binary-Hamming Wyner-Ziv bound [5])**.**

$$R_0^{WZ}(D) = \inf[\alpha g(\beta_1) + (1-\alpha) g(\beta_2)], \tag{21}$$

with

$$g(D) = \begin{cases} H_2(p_0 * D) - H_2(D), & 0 \le D \le p_0 \\ 0, & D = p_0, \end{cases} \tag{22}$$

and the infimum being with respect to all $\alpha \in [0,1]$ *and* $\beta_1, \beta_2 \in [0, p_0]$.

Figure 5 shows the rate-distortion bounds using Corollary 7. As it is found from the figure, the binary Wyner-Ziv bound includes two parts: (i) the graph of function $g(D)$ for $D \le d_c$; and (ii) the straight line which is tangent to the graph of $g(D)$ for $d_c \le D \le p_0$, since distortion level p_0 is achievable by $R_0 = 0$. Here, the point d_c is the

solution of

$$\frac{g(d_c)}{d_c - p_0} = \frac{\mathrm{d}g(D)}{\mathrm{d}D}|_{D=d_c}, \tag{23}$$

where $\frac{\mathrm{d}g(D)}{\mathrm{d}D}$ represents the derivative of $g(D)$ with respect to D. Furthermore, it reduces to the conventional rate-distortion function for a binary source if $p_0 = 0.5$, i.e., the helper is independent with the main source.

2.3.3 Multiterminal Source coding

Considering a multiterminal source coding problem shown in Fig. 1(b), two correlated sources X_1 and X_2 form two n independent copies X_1^n and X_2^n from a joint pmf $p_{X_1X_2}(x_1,x_2)$, as in Slepian-Wolf coding. Each encoder independently sends an index taking from the set $\{1,2,\cdots,2^{nR_i}\}$ for the sequence X_i^n to the decoder, $i = 1,2$. After that, the joint decoder assigns an estimate $(\hat{X}_1^n, \hat{X}_2^n)$ for each received index pair. A rate pair (R_1,R_2) is said to be achievable for distortion pair (D_1,D_2) if there exists codes $(n, 2^{nR_1}, 2^{nR_2})$ such that

$$\lim_{n\to\infty} \sup E\{\frac{1}{n}\sum_{t=1}^{n} d(x_i(t),\hat{x}_i(t))\} \leq D_i, i = 1,2, \tag{24}$$

where $d(\cdot)$ is the distortion measure function and $\sup[\cdot]$ represents the supremum of a subset.

The exact bound on the rate-distortion region of such problem is still an open question. Alternatively, Berger [22] and Tung [23] derived the inner and outer bounds of the rate-distortion region.

Berger-Tung Inner Bound

Theorem 8 (Berger-Tung Inner Bound [22, 78])**.** *For a sequence $\{X_1(t),X_2(t)\}_{t=1}^{\infty}$ of a discrete RV pair (X_1,X_2) drawing i.i.d. from a joint pmf $p_{X_1X_2}(x_1,x_2)$, where $X_1(t) \in \mathscr{X}_1$ and $X_2(t) \in \mathscr{X}_2$. Then for any rate pair that satisfy*

$$\begin{cases} R_1 & \geq & I(X_1;U_1|U_2,Q), \\ R_2 & \geq & I(X_2;U_2|U_1,Q), \\ R_1+R_2 & \geq & I(X_1,X_2;U_1,U_2|Q), \end{cases} \tag{25}$$

there exists an integer n and mappings

$$\begin{array}{rrcl} f_1: & \mathscr{X}_1^n & \to & \mathscr{M}_1 = \{1,2,\cdots,2^{nR_1}\}, \\ f_2: & \mathscr{X}_2^n & \to & \mathscr{M}_2 = \{1,2,\cdots,2^{nR_2}\}, \\ g: & \mathscr{M}_1 \times \mathscr{M}_2 & \to & \mathscr{X}_1^n \times \mathscr{X}_2^n, \end{array} \tag{26}$$

such that

$$E[d_1(X_1,\hat{X}_1)] \le D_1 \text{ and } E[d_2(X_2,\hat{X}_2)] \le D_2, \tag{27}$$

for some conditional pmf $p_Q(q)p_{U_1|X_1Q}(u_1|x_1,q)p_{U_2|X_2Q}(u_2|x_2,q)$ *with* $|\mathscr{U}_i| \le |\mathscr{X}_i| + 4$, $i = 1,2$, *i.e.,* U_i *and* X_i *form a Markov chain by given* Q *as* $U_1 \to X_1 \to X_2 \to U_2|Q$.

It is easy to see that the Slepian-Wolf theorem is a special case of the Berger-Tung inner bound by setting $D_1 = D_2 = 0$.

Berger-Tung Outer Bound

As stated above, the rate-distortion region of lossy source coding problem is still not known in general. The lower convex envelope of the Berger-Tung inner bound determines an upper bound of the rate-distortion region, while the lower convex envelope of the Berger-Tung outer bound determines a lower bound of the rate-distortion region. Here, the Berger-Tung outer bound is reviewed.

Theorem 9 (Berger-Tung Outer Bound [78])**.** *For distributed lossy source coding problem, a rate pair* (R_1,R_2) *is achievable with given a distortion pair* (D_1,D_2) *only if there exists two auxiliary RV's* U_1 *and* U_2 *which satisfy the inequalities*

$$\begin{cases} R_1 & \ge & I(X_1,X_2;U_1|U_2), \\ R_2 & \ge & I(X_1,X_2;U_2|U_1) \\ R_1+R_2 & \ge & I(X_1,X_2;U_1,U_2), \end{cases} \tag{28}$$

for some conditional pmf $p_{U_1U_2|X_1X_2}(u_1,u_2|x_1,x_2)$ *such that* $U_1 \to X_1 \to X_2$ *and* $X_1 \to X_2 \to U_2$ *form two independent Markov chains and* $E[d_i(X_i,\hat{X}_i)] \le D_i$ *with* $\hat{X}_i = f_i(U_1,U_2)$, $i = 1,2$.

The expression of the outer bound is very similar to the Berger-Tung inner bound except requiring the time sharing variable Q to ensure the convexity of the envelope of the inner bound. Furthermore, the Markov condition is weaker than that in Berger-Tung inner bound. Berger-Tung outer bound determines the necessary requirement of the admissible rate pair (R_1,R_2), i.e., in order to achieve the average distortions (D_1,D_2), the rate pair *must* satisfy the inequalities in Berger-Tung outer bound.

Oohama determined the rate-distortion region of a particular case of the multiterminal source coding problem [24], which is so-called Gaussian multiterminal source coding. Two correlated memoryless Gaussian sources X_1 and X_2 having a joint probability density function (pdf) $p_{X_1X_2}(x_1,x_2)$ are independently encoded and jointly decoded within squared distortions D_1, D_2, where the pdf $p_{X_1X_2}(x_1,x_2)$ is given by

$$p_{X_1X_2}(x_1,x_2) = \frac{1}{2\pi|\Sigma|^{1/2}} \exp\{-\frac{1}{2}\mathbf{x}\Sigma^{-1T}\mathbf{x}\} \tag{29}$$

with $\mathbf{x} = (x_1,x_2) \in \mathscr{X}_1 \times \mathscr{X}_2$ and Σ being a covariance matrix as

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}, \quad -1 < \rho < 1. \tag{30}$$

An outer bound of the rate-distortion region was derived in [24], where a rigorous proof of the direct coding theorem is also given. Let

$$R_1(D_1) = \left\{ (R_1,R_2) : R_1 \geq \frac{1}{2}\log^+[\frac{\sigma_1^2}{D_1}(1-\rho^2+\rho^2 2^{-2R_2})] \right\}, \tag{31}$$

$$R_2(D_2) = \left\{ (R_1,R_2) : R_2 \geq \frac{1}{2}\log^+[\frac{\sigma_2^2}{D_2}(1-\rho^2+\rho^2 2^{-2R_1})] \right\}, \tag{32}$$

$$R_{12}(D_1,D_2) = \left\{ (R_1,R_2) : R_1+R_2 \geq \frac{1}{2}\log^+[(1-\rho^2)\frac{\sigma_1^2\sigma_2^2}{D_1D_2}] \right\}, \tag{33}$$

where $\log^+(a) = \max\{\log(a),0\}$. Then the following theorem holds.

Theorem 10 (Oohama's outer bound [24])**.** *For all $D_1, D_2 > 0$, $\mathscr{R}(D_1,D_2) \subseteq \mathscr{R}^{YI}(D_1,D_2)$, where $\mathscr{R}^{YI}(D_1,D_2) = R_1(D_1) \cap R_2(D_2) \cap R_{12}(D_1,D_2)$, and $\mathscr{R}(D_1,D_2) = \{(R_1,R_2) :$ (R_1,R_2) is admissible$\}$.*

The converse proof of deriving Theorem 10 is used as a very important reference in our work for deriving the outer bound on the rate-distortion region for binary multiterminal source coding.

2.3.4 The CEO Problem

The CEO problem of which an abstract model is shown in Fig. 6 provides a theoretical framework of designing the transmission techniques and scheduling protocols for distributed WSNs. A $(n, 2^{nR_1}, \cdots, 2^{nR_L}, \sum_{i=1}^{L} R_i \leq R)$ code for the CEO problem involves

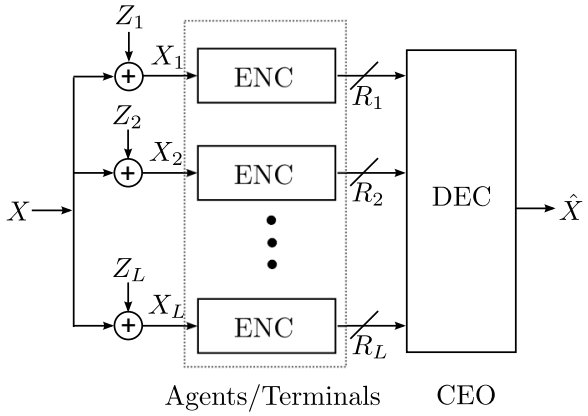


Fig. 6. The abstract model of the CEO problem.

– L encoders send indices taking from the set $\{1,2,\cdots,2^{nR_i}\}$ for each sequence X_i^n, $i = 1,\cdots,L$.
– the decoder generate an estimate $\hat{X}^n$ based on the received L indices.

A rate-distortion pair (R,D) is said to be achievable if

$$\lim_{n\to\infty} \sup E[\frac{1}{n}\sum_{t=1}^{n} d_i(x(t),\hat{x}(t))] \leq D \tag{34}$$

for some codes $(n, 2^{nR_1},\cdots,2^{nR_L}, \sum_{i=1}^{L} R_i \leq R)$.

In general, to determine the achievable rate-distortion region for the CEO problem is an open question except the quadratic Gaussian case. We review an established rate-distortion function derived by Oohama [36] for the quadratic Gaussian CEO problem.

Quadratic Gaussian CEO

The quadratic Gaussian CEO problem assumes that the source X and the multiple observations $X_1,\cdots,X_L$ are jointly Gaussian distributed. Assume that X follows Gaussian distribution with mean 0 and variance σ_X^2, which draws an i.i.d. data sequence $\{X(t)\}_{t=1}^{\infty}$ by taking values from $\mathscr{X}$. Let $\{X_i(t)_{t=1}^{\infty}\}$, $i = 1,\cdots,L$, represent noisy versions of $\{X(t)\}_{t=1}^{\infty}$ corrupted by independent AWGN, i.e.,

$$X_i(t) = X(t) + Z_i(t), \tag{35}$$

where $Z_i(t)$ are independent Gaussian RV's following Gaussian distribution with mean 0 and variance σ_Z^2. The quadratic Gaussian CEO problem is described as follows. The

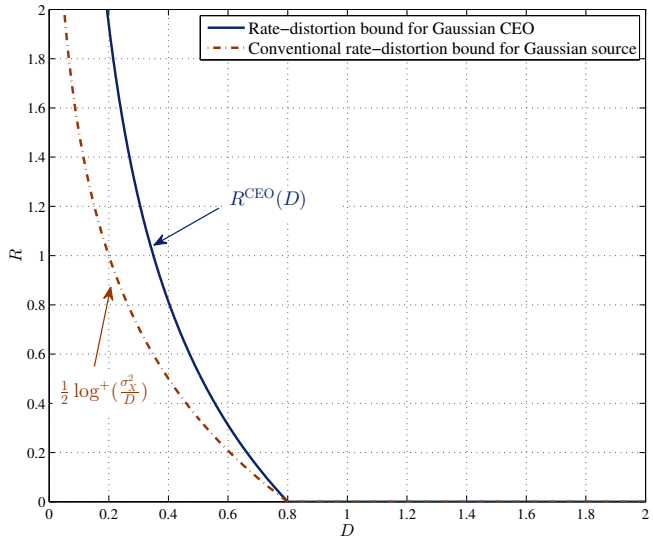


Fig. 7. An example of $R(D)$ with $\sigma_X^2 = 0.8$.

CEO aims at producing the estimates of $\{X(t)\}_{t=1}^{\infty}$ by deploying a team of L agents. Each agent i independently encodes the corrupted version $\{X_i(t)_{t=1}^{\infty}\}$ by the encoder function $\varphi_i(X_i)$ with a sum rate constraint R. The CEO processes information $\varphi_i(X_i)$ to generate the estimation $\hat{X}$ through the decoder function $\psi(\varphi_1(X_1),\cdots,\varphi_L(X_L))$. Mathematically, φ_i is defined as

$$\varphi_i : \mathscr{X}^n \to \mathscr{M}_i = \{1,2,\cdots,2^{nR_i}\}, \tag{36}$$

with the sum rate constraint

$$\sum_{i=1}^{L} R_i \leq R. \tag{37}$$

On the other hand, the decoder function ψ is expressed by

$$\psi : \mathscr{M}_1 \times \mathscr{M}_2 \times \cdots \times \mathscr{M}_L \to \mathscr{X}^n. \tag{38}$$

The explicit formula of the rate-distortion function $R^{\mathrm{CEO}}(D)$ is provided in the following theorem, which is derived by Oohama using CI property.

Theorem 11 (Rate-distortion function of Gaussian CEO [36]). *For every $D > 0$*

$$R^{CEO}(D) = \frac{\sigma_Z^2}{2\sigma_X^2}\left[\frac{\sigma_X^2}{D} - 1\right]^+ + \frac{1}{2}\log^+(\frac{\sigma_X^2}{D}). \tag{39}$$

where $[a]^+ = \max\{0, a\}$.

Figure 7 shows an example of $R^{\mathrm{CEO}}(D)$ function of Theorem 11. The classical rate-distortion function for a Gaussian source is also depicted as a reference. It can be found from the results, the difference between $R^{\mathrm{CEO}}(D)$ and classical rate-distortion function becomes small when D goes large, i.e., the rate loss is large for relatively small D.

2.4 Channel Coding Theorem

We now consider a dual setting that transmits the compressed message over a noisy channel. In information theory and communications, there are various models to describe the statistical properties of noisy channels, such as, BSC, AWGN channel. Since the noisy channel introduces errors into the message, there arises a question what is the condition on the transmission rate R in order to correct all the errors. Shannon established the channel coding theorem to describe the requirement of R for reliable communications. Before going to the detail of the channel coding theorem, we first review an essential property of the noisy channel. Let

$$C = \max_{p_X(x)} I(X;Y) \tag{40}$$

be the capacity of a discrete memoryless channel $p_{Y|X}(y|x)$.

The channel coding theorem states the condition on transmission rate R with respect to C as follows.

Theorem 12 (Channel coding theorem [73, 74]). *For a discrete memoryless channel, if $R \leq C$, reliable communication can be achieved. Conversely, in order to guarantee reliable communication, the transmission rate R* must *satisfy the condition $R \leq C$.*

The channel coding theorem is also called Shannon's second coding theorem [73], which proves the existence of codes, however, how to construct such codes is not provided.

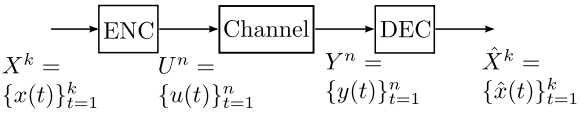


Fig. 8. Joint source-channel coding setup.

2.5 Source-Channel Separation Theorem

There exists an assumption that the channels between the encoders and the decoder are perfect behind the discussions of Section 2.3. However, in many situations, the compressed sources must be transmitted to the decoder over noisy channels, as in Section 2.4. In this section, we review a more general joint source-channel setup, which is illustrated in Fig. 8. The encoder sends a codeword $U^n \in \mathscr{U}^n$ for the source samples $X^k \in \mathscr{X}^k$, and the decoder generates an estimate $\hat{X}^k$ based on the received sequence Y^n. A prevalent way, in this case, is to perform source and channel encoding and decoding separately. For the point-to-point (P2P) communication with stationary memoryless sources and channels, Shannon proved that such strategy is asymptotically optimal [73], which is called Shannon's source-channel separation theorem.

Theorem 13 (Source-channel separation theorem)**.** *Given a discrete memoryless source X and an average distortion measure $d(x,\hat{x})$ with rate-distortion function $R(D)$ and a discrete memoryless channel with capacity C, the following statement hold*

$$\begin{aligned} &\textit{If } kR(D) < nC, \quad \textit{then there exists a sequence of JSC codes such that} \\ &\qquad \lim_{n\to\infty} \sup E[d(X^k, \hat{X}^k)] \leq D. \end{aligned} \tag{41}$$

where k is source samples and n is the channel symbols. $R(D)$ is expressed in bits per sample and the capacity C is in bits per channel symbol.

The crucial point, however, is whether source-channel separation theorem holds for complicated setups, such as sending multiple sources over MAC. Unfortunately, the separation theorem does not hold in general, where a series of contradictory examples can be found in [80–82]. In this dissertation, we only focus on orthogonal MAC, hence, Theorem 13 can be appropriately applied in analyzing the system performance, i.e., it is an optimal way that performs source and channel encoding sequentially in this case.

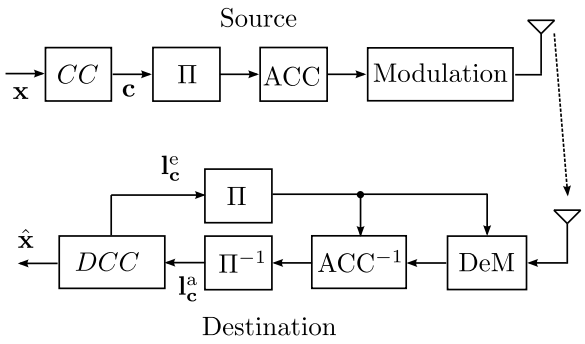


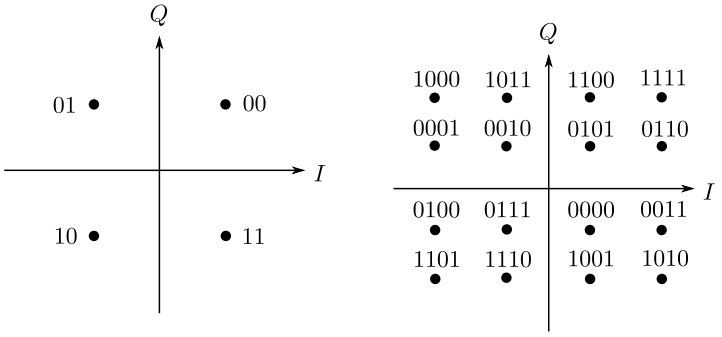
Fig. 9. The proposed encoding scheme with iterative decoding. DeM: demapper.

2.6 Channel Coding, Modulation and EXIT Chart Analysis

2.6.1 Channel Coding and Modulation

To protect the message to be transmitted over noisy channels, channel coding which adds redundancy to the message is performed after the source coding. A variety of channel coding schemes were designed during the last several decades, such as convolutional codes, LDPC codes and fountain codes.

In this subsection, we provide the advantages of our proposed encoding scheme used in binary information sensing networks. As shown in Fig. 9, the information sequence $\mathbf{x}$ is first encoded by the convolution encoder *CC* with generate polynomial $G = [11, 10]$ at coding rate $\frac{1}{2}$. The encoded sequence is then interleaved by a random interleaver Π, of which the role is to enable the iterative decoding according to turbo principle. After that, the interleaved sequence is doped-accumulated by an accumulator ACC and then mapped to the symbol based on the constellation patterns shown in Fig. 10. Note that ACC is a systematic recursive convolutional code with $G = [11, 10]$, and the systematic bit is regularly superseded by the coded bit based on the doping ratio for achieving coding rate 1. Furthermore, the modulation scheme, particularly, higher order modulations used here is based on bit-interleaved coded-modulation with iterative decoding (BICM-ID) technique, where the demapper/demodulation process is involved in the iterative decoding to improve the performance. Hence, the constellation of quadrature phase-shift keying (QPSK) and 16-quadrature amplitude modulation (QAM) are not Gray mapping. Note that for binary phase-shift keying (BPSK) iterative demapper is not needed.



(a) Natural mapping of QPSK.

(b) MSP of 16QAM.

Fig. 10. Constellation of QPSK and 16QAM.

Iterative decoding is performed for the proposed encoding scheme. The decoders ACC^{-1} and *DCC* use the logarithmic versions of maximum a posteriori algorithm, such as BCJR to generate the LLR sequences. Moreover, the function of demapper is to generate the extrinsic LLR of k-th bit of symbol $s(t)$ by

$$\begin{aligned} l^{\mathrm{e}}(s(t)^k) &= \ln \frac{\Pr(y(t)|s(t)^k=1)}{\Pr(y(t)|s(t)^k=0)} \\ &= \ln \frac{\sum\limits_{s(t)\in S_1^k} p(y(t)|s(t)) \prod_{j=1,j\neq k}^{m} \exp(s(t)^j \cdot l^{\mathrm{a}}(s(t)^j))}{\sum\limits_{s(t)\in S_0^k} p(y(t)|s(t)) \prod_{j=1,j\neq k}^{m} \exp(s(t)^j \cdot l^{\mathrm{a}}(s(t)^j))}, \end{aligned} \tag{42}$$

where $\ln(\cdot)$ is the natural logarithm function. $s(t)^k$ represents k-th bit of symbol $s(t)$ and S_b^k is the subsets of the symbols of which the k-th bit is b, $b = 0, 1$. The a priori LLR $l^{\mathrm{a}}(s(t)^j))$ is the feedback from the decoder *DCC*. The bit per symbol in the constellation is denoted by m, i.e., $m = 2$ for QPSK and $m = 4$ for 16QAM. The extrinsic LLR $l^{\mathrm{e}}(s(t)^k)$, $t = 1, \cdots, 2n$, $k = 1, \cdots, m$ forms the LLR sequence $\mathbf{l}_{\mathbf{v}}^{\mathrm{e}}$ and fed into ACC^{-1}.

The proposed encoding scheme is very simple, since the memory size of the convolutional code is 1, however, it achieves superior BER performance against channel noises, which is verified through EXIT chart analysis.

2.6.2 EXIT Chart Analysis

Basics

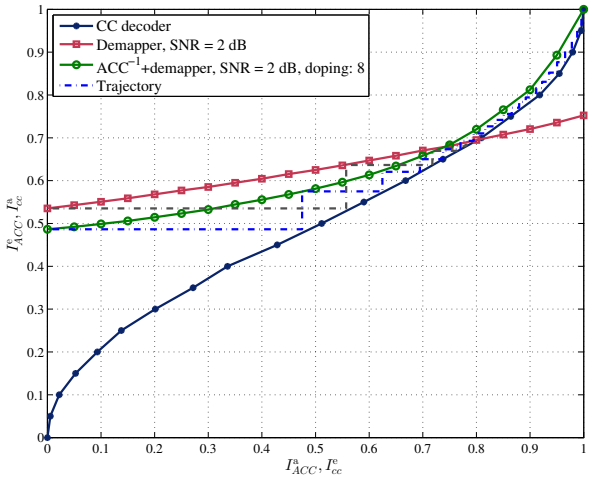
EXIT chart is a novel tool to analyze the convergency property of iterative decoding of concatenated soft-in-soft-out decoders [83, 84].

As mentioned in [84], the BER performance of iterative decoding usually contains three regions: 1) in the region of low SNR, BER is very high with negligible gain over iterations; 2) the turbo cliff region where the BER is significantly reduced by increasing SNR by a small amount when enough iterations are performed; 3) the error floor appears in the region of high SNR where a quite low BER can be achieved by performing only several iterations. However, it is well known that with the help of doped-accumulator (ACC), the error floor can be eliminated, which is supported by the EXIT chart analysis [85].

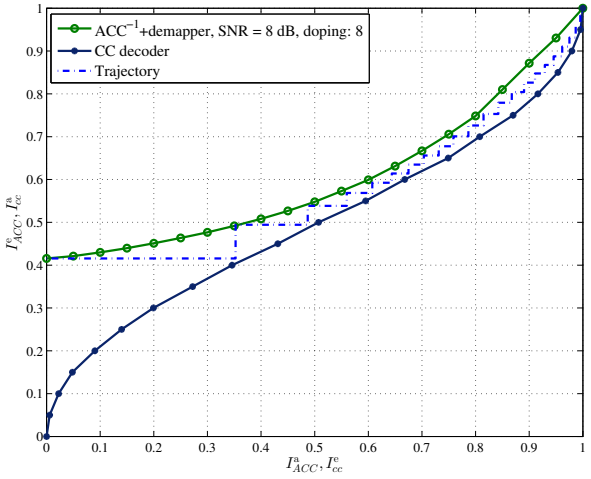
In order to analyze the convergence behavior of iterative decoding, a density evolution algorithm has been proposed to calculate the convergence threshold for LDPC codes over the AWGN channel and to construct LDPC capacity-approaching codes [86, 87]. The main idea of density evolution is to track the pdf of the exchanged information message in the iterative decoding process. The pdf of LLRs can be assumed to be Gaussian distributed. EXIT chart visualizes the density evolution of extrinsic LLR over the iteration using the mutual information between the coded bits at the transmitter and the corresponding LLRs at the receiver. It can be visually understand the exchange of extrinsic LLR through the trajectory in the EXIT chart analysis, and hence, the turbo cliff region can be predicted. Furthermore, with the aid of the EXIT chart, the code optimization falls into the problem of the EXIT curve matching [88].

EXIT Chart

We consider the analysis of the proposed encoding scheme with iterative decoding, as shown in Fig. 9. The output LLR sequence $\mathbf{l}_{\mathbf{v}}^{\mathrm{e}}$ of the demapper together with ACC^{-1} is generated from the received signals $\mathbf{y}$ and the a priori LLR $\mathbf{l}_{\mathbf{v}}^{\mathrm{a}}$ fed back from the decoder *DCC*. However, the extrinsic LLR sequence $\mathbf{l}_{\mathbf{c}}^{\mathrm{e}}$ of *DCC* is generated by BCJR algorithm with the input LLR $\mathbf{l}_{\mathbf{c}}^{\mathrm{a}}$ which is artificially generated. The mutual information $I_{ACC}^{\mathrm{e}} = I(\mathbf{l}_{\mathbf{v}}^{\mathrm{e}}; \mathbf{v})$ and $I_{CC}^{\mathrm{e}} = I(\mathbf{l}_{\mathbf{c}}^{\mathrm{e}}; \mathbf{c})$ need to be computed.



(a) QPSK with natural mapping.



(b) 16QAM with MSP mapping.

Fig. 11. EXIT chart analyses of using proposed encoding and iterative decoding.

The mutual information $I(B;L_b^{\mathrm{e}})$ between the RV L_b^{e} and B, where L_b^{e} and B represent the RV's that generate the realizations of the extrinsic LLR l and the bit $b \in \{+1,-1\}$, respectively. As we know, the pdf of RV L_b^{e} approaches Gaussian distribution as the number of iterations is enough, thus, L_b^{e} is modelled by

$$L_b^{\mathrm{e}} = \mu_l b + n_l \tag{43}$$

with n_l is the Gaussian RV with mean 0 and variance σ_l^2. The mean of L_b^{e} is represented by $\mu_l = \frac{\sigma_l^2}{2}$. According the definition of mutual information, $I(B;L_b^{\mathrm{e}})$ is calculated by

$$I(B;L_b^{\mathrm{e}}) = \sum_{b=+1,-1} \int_{-\infty}^{+\infty} p(l|b)p(b) \log \frac{p(l|b)}{p(l)} dl. \tag{44}$$

Using the assumption that $p(b) = \frac{1}{2}$ for both $b = +1,-1$ and $p(l) = \frac{1}{2}[p(l|b=-1) + p(l|b=+1)]$, we have

$$I(B;L_b^{\mathrm{e}}) = \frac{1}{2} \sum_{b=-1,1} \int_{-\infty}^{+\infty} p(l|b) \cdot \log \frac{2p(l|b)}{p(l|b=-1) + p(l|b=+1)} dl, \tag{45}$$

with

$$p(l|b) = \frac{1}{\sqrt{2\pi\sigma_l^2}} \exp\left(-\frac{(l - \frac{\sigma_l^2}{2} b)^2}{2\sigma_l^2} \right). \tag{46}$$

Applying (45) to our system, we can evaluate the mutual information between the output LLRs of DCC, the demapper together with ACC^{-1} and the corresponding bits. The values of mutual information are plotted in Fig. 11.

From the EXIT chart analysis, it is found that the curve of the demapper together with ACC^{-1} approached the $(1.0, 1.0)$ mutual information point, however, the curve of only using demapper (without ACC) cannot achieve that point. In other words, the BER performance can be significantly improved with ACC, in particular, the error floor region is removed. Furthermore, through EXIT chart analyses, the doping ratios of ACC for BPSK, QPSK and 16-QAM modulations are obtained.

2.7 Decision of Binary Information Sensing

2.7.1 Binary Information Sensing

The binary information sensing is the major application to be considered. As shown in Fig. 12, multiple sensors observe a common binary source which produces binary i.i.d.

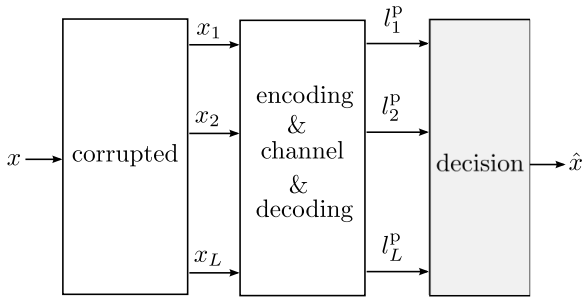


Fig. 12. An abstract model of binary information sensing.

sequence $\mathbf{x} = [x(1), \cdots, x(n)]$ from the Bernoulli distribution. The noisy observations $\mathbf{x}_i = [x_i(1), \cdots, x_i(n)]$ are encoded independently by each sensor and sent to the fusion center over noisy channel. The observation error probability of i-th link is p_i. At the fusion center, the joint decoding process the a posteriori LLR $\mathbf{l}_i^{\mathrm{p}}$ based on the received signal from the channel. The detail of the encoding, property of channels and decoding is omitted since we focus on the decision rule. Also, in the figure, only one realization of the source is presented.

2.7.2 Majority Voting

A simple decision rule for generating estimates $\hat{x}$ from multiple independent LLR values $l_1^{\mathrm{p}}, \cdots, l_L^{\mathrm{p}}$ is majority voting, i.e., the LLR values are first converted to hard bits based on the sign of LLRs, $+ \mapsto 1$ and $- \mapsto 0$, and then comparing the number of 1's and 0's of the hard bits. The majority voting decision is expressed as

$$\hat{x} = \begin{cases} 1, & \sum_{i=1}^{L} \hat{x}_i > \frac{L}{2} \\ 0, & \text{otherwise} \end{cases}, \tag{47}$$

where $\hat{x}_i$ is the generated bits from l_i^{p} based on hard decision. The majority voting decision rule deals with hard bit, however, the information is lost during the hard decision process of the LLR values. The other drawback of majority voting is that the observation error p_i is not utilized in the decision. Intuitively, if p_i of each link is not the same, the LLR value of the link with smaller p_i should assign larger weight. Thus, majority voting works inadequately for the case p_i of each link are various.

2.7.3 Soft Combining

To avoid disadvantages of the majority voting, we proposed a soft combining technique to generate the estimate $\hat{x}$ as

$$\hat{x} = \begin{cases} 1, & \sum_{i=1}^{L} f_c(l_i^{\mathrm{p}}, p_i) > 0 \\ 0, & \text{otherwise} \end{cases}, \tag{48}$$

where function f_c is the LLR updating function [54], the definition of which is

$$f_c(l_i^{\mathrm{p}}, p_i) = \ln \frac{(1-p_i)\exp(l_i^{\mathrm{p}}) + p_i}{(1-p_i) + p_i \exp(l_i^{\mathrm{p}})}, \tag{49}$$

Soft combining outperforms the majority voting since the LLR is weighted by p_i through f_c function. However, it is difficult to analyze the decision error probability for the soft combining.

2.8 Summary

This chapter provided the background concepts and theorems used throughout this research. First of all, we reviewed the definition of entropy and mutual information. Several necessary inequalities, such as MGL, for conversely proving the rate-distortion region were summarized. After that, the well-known theorems in the category of DSC were presented. We briefly discussed the channel coding theorem and the source-channel separation theorem, followed by the practical channel coding techniques, EXIT chart analysis and decision rules used in binary information sensing.

3 Analyses of Asymptotic Sum Rate Limit and Bit Error Rate Floor

We consider a binary information sensing network, in which multiple sensors observe a single binary source to produce erroneous observations. They are independently encoded by each sensor and transmitted to a fusion center over orthogonal AWGN channels. The fusion center reconstructs the estimates of the common binary source from the received signals by a joint decoder. As stated in Chapter 1, the problem of estimating a single source via multiple unreliable sensors (agents) is modeled by the binary CEO problem. However, due to the difficulty in deriving the rate-distortion function for the binary CEO problem, we analyze the binary information sensing network using the Slepian-Wolf theorem in this chapter.

Numerous encoding/decoding algorithms have been proposed for binary information sensing networks [50, 51, 53, 55]. In [50], authors proposed an encoding scheme using parallel concatenated convolutional codes and a joint decoding method using weighted extrinsic LLR to utilize the correlation knowledge. An adaptive bi-modal decoder for the binary information sensing network having two sensors was proposed in [51]. We studied the binary information sensing network in [53] and [55]. A very simple encoding and joint decoding algorithm using LLR updating function was proposed, the BER performance of which is significantly improved [53]. We further proposed an algorithm for estimating the correlation parameters in [55].

In AWGN channels, the BER curve of the proposed techniques for the binary information sensing network exhibits a sharp turbo cliff at a certain threshold SNR like the standard turbo codes do. However, the BER curve exhibits a flat error floor, which keeps at a certain level even we increase the SNR. It is clearly understood that the floor is due to the fact that the observations suffer from errors before being channel-encoded for transmission. Then, there arises a question: what the reference values of the theoretical threshold SNR and the bit error floor are to be compared with?

To answer this question, we use the Slepian-Wolf theorem, channel capacity theorem and separability to derive the theoretical SNR limit converted from the achievable sum rate. Furthermore, the BEP floor is analytically calculated on the basis of Poisson binomial process, and its lower bound is obtained from the binary rate-distortion function.

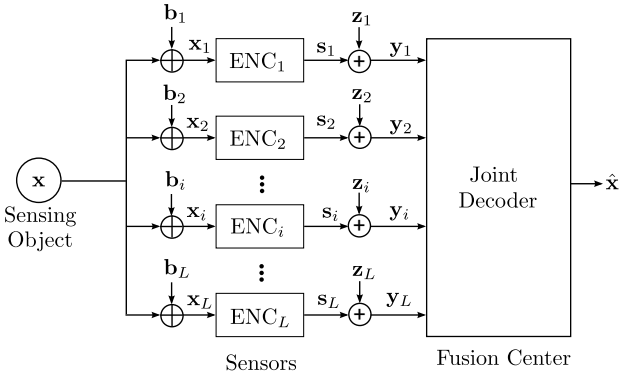


Fig. 13. A schematic diagram of a binary information sensing network.

The rest of this chapter is organized as follows. Section 3.1 states the problem of binary information sensing. The approximation of the limit on sum rate using the Slepian-Wolf theorem is derived in Section 3.2. The BEP floor and its lower bound are analyzed in Section 3.3. The encoding/decoding algorithm is presented in Section 3.4. Numerical results obtained through computer simulations and EXIT chart analysis are shown in Section 3.5. Finally, Section 3.6 concludes this chapter.

3.1 Problem Statement

The block diagram of the binary information sensing network to be analyzed is illustrated in Fig. 13. A sensing object, producing binary i.i.d. data $\mathbf{x} = [x(t)]_{t=1}^{n}$ from a source $X \sim Bern(0.5)$, is observed by L sensors, where t represents the time index with n being the block length. The observation $\mathbf{x}_i$ made by the i-th sensor is a corrupted version of $\mathbf{x}$ by the binary error sequence $\mathbf{b}_i$ with $\Pr(b_i(t) = 1) = p_i$, where $i = 1, 2, \cdots, L$ indicates the index of sensors. The observation $\mathbf{x}_i$ can be seen as the output of a BSC with associated crossover probability p_i when $\mathbf{x}$ is the input, i.e., $\mathbf{x}_i = \mathbf{x} \oplus \mathbf{b}_i$. Consequently, the observations are correlated with each other since they have the common input $\mathbf{x}$ of the independent BSCs. In other words, X and X_i follows a joint pmf $p_{XX_1\cdots X_L}(x, x_1, \cdots, x_L)$, where X_i is the random variable that generates observation sequences $\mathbf{x}_i$.

Each sensor encodes its observation $\mathbf{x}_i$ by an joint source-channel encoder ENC_i and transmits the corresponding coded data to the fusion center over independent AWGN

channels. The fusion center performs joint decoding to form estimates $\hat{\mathbf{x}}$ of the original source.

It should be emphasized here that the difference between our approach and Berger-Tung coding strategy. In the latter one, a vector quantization (VQ) is used to generate codewords for source samples, followed by a binning step. However, we focus on the case where we observe binary sequences. Moreover, our purpose is to determine how much transmission power (dual to the compression rate) can be saved by exploiting the source correlation at the joint decoding stage. As a result, encoding process by the VQ is not explicitly considered here.

3.2 Achievable Sum Rate and SNR limit in AWGN Channels

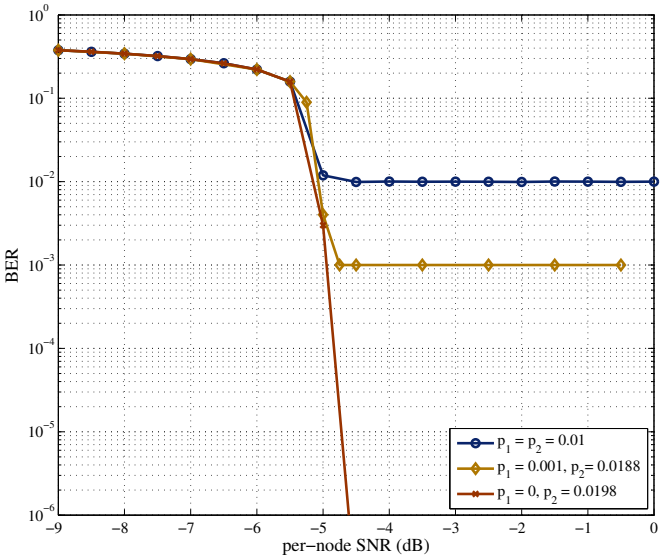
In order to analyze the theoretical SNR limit for the binary information sensing network, the achievable rate region of the source coding should be clarified based on source-channel separation theorem. However, directly deriving the achievable rate region using the source coding model, the binary CEO problem, is not easy. We then modify the problem into a relatively simple one in order to approximately analyze the achievable rate region. It is found through simulation results, as shown in Fig. 14, that the turbo cliff appears at the SNR value that the bit error floor is reached. Hence, for deriving the theoretical threshold SNR, we assume that observations are losslessly recovered, which corresponds to the situation that the bit error floor is achieved. As a reference, in Fig. 14, the BER performance of the case that one of p_i is 0 is also presented. This case is equivalent to one-way relaying system which has been intensively studied by using the Slepian-Wolf theorem in [65, 89]. It is found that the turbo cliff almost appears at the same SNR point if the correlation level is the same.

Therefore, the Slepian-Wolf theorem is applied to obtain the achievable sum rate for the binary information sensing network.

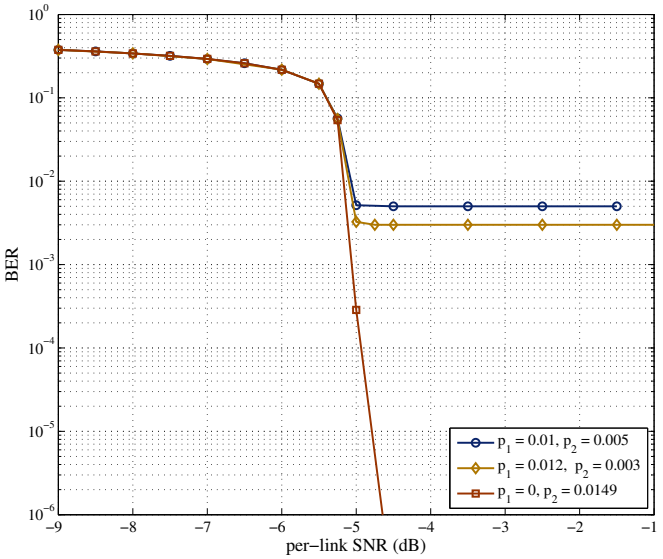
Since the channels between the sensors and the fusion center are assumed to be *orthogonal* with each other, source-channel separation holds in this case [54]. Therefore, based on *separability*, the SNR limit converted from the achievable sum rate is shown as below.

Proposition 14.

$$SNR_{lim} = \begin{cases} 10\lg\left(\frac{2^{2R_i^c H(X_{\mathscr{L}})/L}-1}{2}\right), & \textit{one-dimensional signal} \\ 10\lg\left(2^{R_i^c H(X_{\mathscr{L}})/L}-1\right), & \textit{two-dimensional signal} \end{cases}, \tag{50}$$



(a) Correlation level: $p_1 * p_2 = p_1(1-p_2)+p_2(1-p_1) = 0.0198$**.**



(b) Correlation level: $p_1 * p_2 = p_1(1-p_2)+p_2(1-p_1) = 0.0149$**.**

Fig. 14. BER comparison between the lossless and lossy cases.

where R_i^{c} represents the end-to-end rate of the i-th sensor node and $\lg(\cdot)$ *is the logarithm function with base* 10. *$H(X_{\mathscr{L}}) = H(X_1, \cdots, X_L)$ is the joint entropy of sources X_i, $i = 1, \cdots, L$. The SNR limit SNR_{lim} is expressed in dB.*

Proof. Starting from source-channel separation theorem, we have

$$\sum_{i=1}^{L} R_i(D_i) R_i^{\mathrm{c}} \leq \sum_{i=1}^{L} C(\gamma_i), \tag{51}$$

where $R_i(D_i)$ is the required rates in bits per sample that achieve Hamming distortions D_i between $\mathbf{x}_i$ and $\hat{\mathbf{x}}_i$, and R_i^{c} is the end-to-end rate in samples per channel use. $C(\gamma_i)$ is the Shannon's channel capacity function in bits per channel use and γ_i the SNR. By letting $D_i \to 0$, we can rewrite (51) to

$$R_i^{\mathrm{c}} \sum_{i=1}^{L} R_i \leq LC(\gamma_i). \tag{52}$$

We need to compute the sum rate $\sum_{i=1}^{L} R_i$ to obtain the threshold γ_i. As we know, the source coding of Fig. 13 belongs to Slepian-Wolf coding problem when D_i *goes to* 0. Using Slepian-Wolf theorem summarized in Theorem 5, the sum rate $\sum_{i=1}^{L} R_i = H(X_{\mathscr{L}})$.

Therefore, substituting $\sum_{i=1}^{L} R_i = H(X_{\mathscr{L}})$ into (52), we have

$$\gamma_{\mathrm{lim}} = \begin{cases} \frac{2^{2R_i^{\mathrm{c}} H(X_{\mathscr{L}})/L} - 1}{2}, & \text{one-dimensional signal} \\ 2^{R_i^{\mathrm{c}} H(X_{\mathscr{L}})/L} - 1, & \text{two-dimensional signal} \end{cases}, \tag{53}$$

based on an assumption that *capacity-achieving code* is used at each sensor node. Then by changing the unit of γ_{lim} to a logarithmic unit decibel, we obtain (50). □

It should be emphasized here that the derived SNR limit SNR_{lim} assumes that the minimum Hamming distortion (BEP floor) is reached in the binary sensing network, since we set D_i goes to 0. The minimum Hamming distortion is equivalent to the bit error floor in the BER performance, which is analyzed in the next section. Now we need to calculate $R_{\mathrm{sum}} = \sum_{i=1}^{L} R_i = H(X_{\mathscr{L}})$.

Given the fact that x_i, $i = 1, \cdots, L$, is the result of passing x through a BSC with crossover probability p_i, where x_i and x represent the i.i.d. realizations of X_i and X, respectively, the joint probability $\Pr(x_1, x_2, \cdots, x_L)$ is formulated as

$$\Pr(x_1, x_2, \cdots, x_L) = \frac{1}{2} \prod_{k \in \mathscr{A}} (1 - p_k) \prod_{j \in \mathscr{A}^{\mathrm{C}}} p_j + \frac{1}{2} \prod_{k \in \mathscr{A}} p_k \prod_{j \in \mathscr{A}^{\mathrm{C}}} (1 - p_j), \tag{54}$$

where $\mathscr{A} = \{i \in \mathscr{L} | x_i = 0\}$ and $\mathscr{A}^{\mathrm{C}} = \mathscr{L} \setminus \mathscr{A}$ is the complement set of $\mathscr{A}$ with $\mathscr{L} = \{1, \cdots, L\}$. For example, setting $L = 3$ with $x_1 = 0$, $x_2 = 1$ and $x_3 = 0$, the set $\mathscr{A}$ is equal to $\{1,3\}$ and $\mathscr{A}^{\mathrm{C}} = \{2\}$.

Therefore, the joint entropy $H(X_{\mathscr{L}})$ with $X_{\mathscr{L}} = \{X_i | i \in \mathscr{L}\}$, which is equivalent to R_{sum}, is calculated as

$$H(X_{\mathscr{L}}) = -\sum_{x_i \in \{0,1\}} \Pr(x_1, x_2, \cdots, x_L) \log(\Pr(x_1, x_2, \cdots, x_L)). \tag{55}$$

However, the computational complexity increases very fast if L goes large. We further simplify the calculation in (55) by considering the symmetric property of $\Pr(x_1, x_2, \cdots, x_L)$, and function $h(\cdot)$ is defined to calculate the joint entropy $H(X_{\mathscr{L}})$ as follows.

Definition 1. A function $h(p_{\mathscr{L}})$ expresses the joint entropy of the outputs of independent BSCs with a set of crossover probabilities $p_{\mathscr{L}}$ and the input X. Based on this fact, function $h(p_{\mathscr{L}})$ is

$$h(p_{\mathscr{L}}) = \begin{cases} 0, & \text{if } |\mathscr{L}| = 0, \\ 1, & \text{if } |\mathscr{L}| = 1, \\ -2 \cdot \sum\limits_{k=1}^{2^{(|\mathscr{L}|-1)}} q_k \log(q_k), & \text{otherwise}, \end{cases} \tag{56}$$

with $q_k = 0.5 \cdot (\prod_{j \in \mathscr{A}_k} p_j \prod_{l \in \mathscr{A}_k^c} (1 - p_l) + \prod_{j \in \mathscr{A}_k} (1 - p_j) \prod_{l \in \mathscr{A}_k^c} p_l)$ and $\mathscr{A}_k \subseteq \{1, 2, \cdots, |\mathscr{L}| - 1\}$.

3.3 BEP Floor Analysis

To analyze the bit error floor that appears in the BER performance curve when performing the encoding/decoding technique proposed in [53] and [55], we assume the channels between the sensors and the fusion center are noiseless, since the bit error floor appears in the high SNR regime and it is only determined by the observation error probabilities. It should be emphasized here that the bit error floor is not caused by the coding technique, while it is caused by the system structure. The observation errors are inserted in the source sequence before the encoding process. In this section, we statistically evaluate the value of the bit error floor, which is referred to as BEP floor, and its theoretical lower bound.

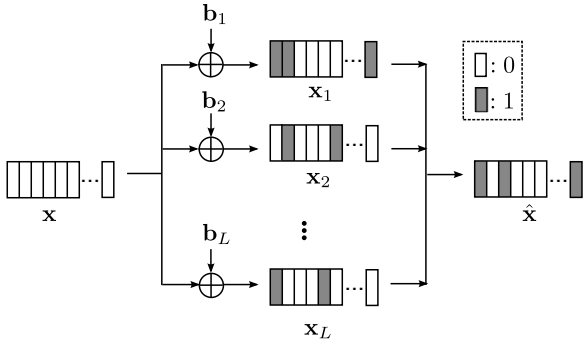


Fig. 15. The model of analyzing the BEP floor.

3.3.1 Poisson-Binomial Approximation

Without loss of generality, we assume the source $\mathbf{x}$ is an all zero sequence with n bits, as shown in Fig. 15. The majority decision rule to generate $\hat{x}(t)$ is

$$\hat{x}(t) = \begin{cases} 1, & \text{if } 1(\mathbf{X}(t)) > 0(\mathbf{X}(t)), \\ 0, & \text{otherwise}, \end{cases} \tag{57}$$

where $1(\mathbf{X}(t))$ and $0(\mathbf{X}(t))$ count the number of 1's and 0's in the t-th column of $\mathbf{X}$, respectively, with $\mathbf{X}$ being a $L \times n$ matrix that stores $\mathbf{x}_i$ row-by-row. The decision error occurs when $\hat{x}(t)$ is decided to be 1. Hence, the BEP floor is analyzed by determining the probability of the number of 1's from L independent Bernoulli sequences having different probabilities of "1". In the statistics, the Poisson-binomial distribution is the probability distribution of the sum of independent Bernoulli trials that are not necessarily identically distributed. The probability of J-times occurrence of the error in L-times repeated binary trials with different crossover probabilities is [90]

$$\Pr(J = j) = \begin{cases} \prod_{i=1}^{L} (1 - p_i), & j = 0, \\ \frac{1}{j} \sum_{k=1}^{j} (-1)^{(k-1)} \Pr(J = j-k) U(k), & j > 0, \end{cases} \tag{58}$$

where $U(k) = \sum_{i=1}^{L} \left(\frac{p_i}{1-p_i} \right)^k$ and $0 \leq j \leq L$.

As a result, the BEP floor with different observation noise probabilities is calculated by

$$\mathrm{PB}(p_{\mathscr{L}}) = \Pr(\hat{x}(t) \neq x(t)) = \begin{cases} \sum\limits_{j=\frac{L+1}{2}}^{L} \Pr(J=j), & \text{if } L \text{ is odd,} \\ \frac{1}{2}\Pr(J=\frac{L}{2}) + \sum\limits_{j=\frac{L}{2}+1}^{L} \Pr(J=j), & \text{if } L \text{ is even.} \end{cases} \tag{59}$$

Obviously, it is reduced to a binomial process if p_i is identical.

3.3.2 Theoretical Lower Bound on the BEP Floor

The BEP floor analysis provided in the previous subsection is based only on the generalized majority logic (Poisson-binomial) analysis. However, it does not take into account the impact of soft-combining of the LLRs. The exact BEP floor analysis taking soft-combining into account is studied in [61, 91]. Instead, we further provide a theoretical lower bound on the BEP floor p_{lb}.

According to the rate-distortion theory for a binary source [74], the theoretical lower bound of the BEP floor is given in the following proposition.

Proposition 15 (Lower bound on BEP floor)**.** *Assume that a random variable $X \sim Bern(0.5)$, and X_i is the output variable of a BSC with crossover probability p_i, where $0 \leq p_i \leq \frac{1}{2}$. The minimum error probability of estimating X from X_i is given by*

$$p_{lb} = H_2^{-1}[1 + \sum_{i=1}^{L} H_2(p_i) - H(X_{\mathscr{L}})]. \tag{60}$$

Proof. According to the rate-distortion function for binary source [74], we have (61)

$$\begin{aligned} 1 - H_2(\tilde{d}) &\leq I(X;\hat{X}) && (61)\\ &\leq I(X;X_{\mathscr{L}}) && (62)\\ &= H(X) - H(X|X_{\mathscr{L}}) && (63)\\ &= 1 - H(X,X_{\mathscr{L}}) + H(X_{\mathscr{L}}) && (64)\\ &= 1 - \{H(X) + H(X_1|X) + \cdots + H(X_L|X)\} + H(X_{\mathscr{L}}) && (65)\\ &= 1 - \{1 + H_2(p_1) + \cdots + H_2(p_L)\} + H(X_{\mathscr{L}}), && (66) \end{aligned}$$

where $\tilde{d}$ is a dummy variable, and the steps are justified as:
(62) there exists information loss in the process of obtaining $\hat{X}$ from $X_{\mathscr{L}}$,

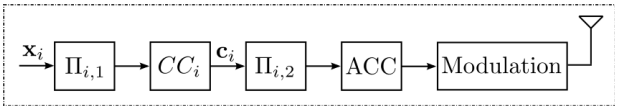


Fig. 16. Proposed encoding scheme at each sensor node.

(65) assume $X_i \to X \to (X_1, \cdots, X_{i-1}, X_{i+1}, \cdots, X_L)$ forms Markov chains, i.e., given X, X_i are independent to each other [10]. A detailed explanation of the CI is given in Appendix 1.

Thus, it is obvious from (66) that $\tilde{d} \geq H_2^{-1}[1 + H_2(p_1) + \cdots + H_2(p_L) - H(X_{\mathscr{L}})]$. Therefore, the lower bound on the BEP floor p_{lb} in (60) is obtained by setting to the minimal value of $\tilde{d}$. □

The lower bound p_{lb} is fully based on theoretical analysis regardless of any specific decision rules. As it is seen from (61) and (62), we assume that there is *no information loss* when producing $\hat{X}$ from $X_{\mathscr{L}}$. However, in practice, this is difficult to achieve. It is left as a future study to find better decision schemes that the soft combining.

3.4 Encoding and Joint Decoding Algorithms

3.4.1 Encoding Scheme

There are a variety of encoding techniques that can be used for the binary information sensing network. However, as stated above, the primary purpose of this chapter is to study the asymptotic limit of the sum rate and the BEP floor, and hence, we simply follow the transmission technique proposed in [53, 55].

As shown in Fig. 16, observation $\mathbf{x}_i$ is first interleaved by an interleaver $\Pi_{i,1}$ whose role is to enable the iterative decoding process to utilize the correlation knowledge according to the turbo principle. The interleaved sequence is then encoded by an encoder CC_i resulting in the coded sequence $\mathbf{c}_i$. It is further interleaved by another interleaver $\Pi_{i,2}$, the length of which depends on the coding rate of the encoder CC_i. After that, the interleaved version of $\mathbf{c}_i$ is doped-and-accumulated by a so-called ACC with a doping ratio P_d. Finally, the modulated symbol sequence $\mathbf{s}_i$ generated by a specified modulation scheme for the doped-accumulated binary sequence, is transmitted over AWGN channels to the fusion center. For simplicity, the modulation schemes used in this chapter are BPSK and QPSK, and we assume that the channels from the sensors to the fusion center are orthogonal to each other, indicating that the transmission from a sensor does

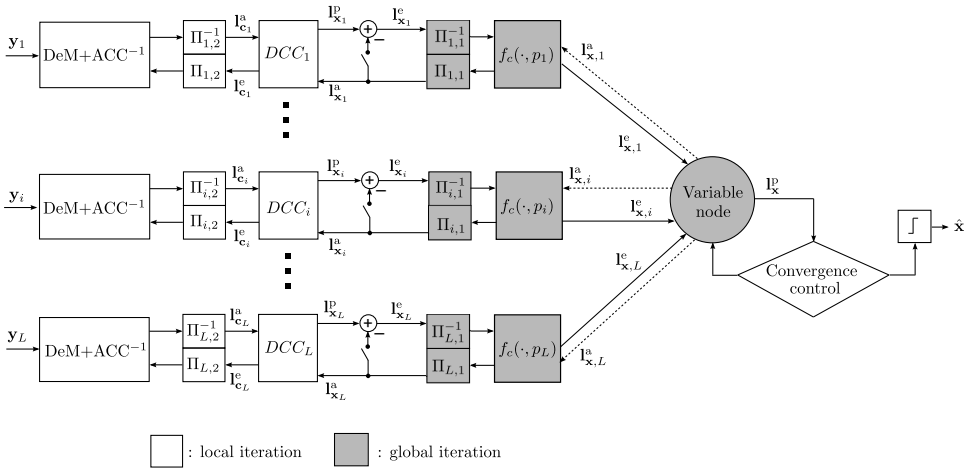


Fig. 17. The block diagram of the joint decoder that utilizes the correlation knowledge through global iteration.

not suffer interference from the other sensors. Therefore, the received signal $\mathbf{y}_i$ in the fusion center from the i-th sensor is simply expressed as

$$y_i(t) = s_i(t) + z_i(t), \tag{67}$$

where $z_i(t)$ generating from $CN(0, \sigma^2)$ is the complex AWGN sequence with the variance per dimension σ^2.

3.4.2 Joint Decoding Algorithm

The block diagram of the proposed joint decoding technique is shown in Fig. 17. Because of the fact that the received data sequences from the sensors are highly correlated, global iteration is introduced in the iterative decoding process with the view to utilizing the correlation knowledge, i.e., the extrinsic LLR is exchanged between global and L local iterations to enhance the BER performance and thereby to reduce the power consumption of each sensor.

The decoding process is divided into following steps.

(a) Initialization: The channel state information is fully known to the fusion center. Thereby, after receiving the signal $\mathbf{y}_i$ from the i-th sensor, the channel LLR $\mathbf{l}^{\mathrm{c}}_i =$

$[l_i^{\mathrm{c}}(t)]_{t=1}^{n}$ is calculated by[3]

$$l_i^{\mathrm{c}}(t) = \frac{2}{\sigma^2}\Re(y_i(t)), \tag{68}$$

where $\Re(\cdot)$ takes the real part of a complex value in its argument. In addition, the extrinsic LLR $\mathbf{l}^{\mathrm{e}}_{\mathbf{c}_i}$ of the coded bits and the *a priori* LLR $\mathbf{l}^{\mathrm{a}}_{\mathbf{x}_i}$ of the systematic bits, which are corresponding to the *a priori* LLR fed back to ACC^{-1} after interleaving, and the *a priori* LLR to DCC_i, respectively, are set to 0.

(b) Local iteration: Local iterations including the decoding processes of decoders ACC^{-1} and DCC_i are concurrently performed once.

(c) Global iteration: The extrinsic LLR $\mathbf{l}^{\mathrm{e}}_{\mathbf{x}_i}$, which is obtained by subtracting the *a priori* LLR $\mathbf{l}^{\mathrm{a}}_{\mathbf{x}_i}$ from the *a posteriori* LLR $\mathbf{l}^{\mathrm{p}}_{\mathbf{x}_i}$, is deinterleaved by $\Pi_{i,1}^{-1}$ and then updated through function[4] f_c [54] as

$$\mathbf{l}^{\mathrm{e}}_{\mathbf{x},i} = f_c\{\Pi_{i,1}^{-1}(\mathbf{l}^{\mathrm{e}}_{\mathbf{x}_i}), p_i\}. \tag{69}$$

Function f_c can be obtained by first converting the LLR into the probabilities of the systematic bits being "0" and "1", and then modifying the values by taking into account the knowledge of p_i. The purpose of function f_c is to avoid the error propagation while exchanging likelihood information between local and global iterations.

The interleaver $\Pi_{i,1}$ makes the observation errors randomly distributed in the observation, and hence, global iteration can be involved in the decoding process to achieve significant improvement on the BER performance.

Since the scenario of multiple sensors observing the common sensing object can be seen equivalent to performing the repetition coding, a degree L variable node combines $L-1$ extrinsic LLRs $\mathbf{l}^{\mathrm{e}}_{\mathbf{x},j}$, $j \neq i$, obtained as the results of local iterations, to calculate the message fed back to the decoder DCC_i of the k-th local iteration. Thus, the *a priori* LLR $\mathbf{l}^{\mathrm{a}}_{\mathbf{x},i}$ is then obtained as

$$\mathbf{l}^{\mathrm{a}}_{\mathbf{x},i} = \sum_{j \in \mathscr{L}\backslash i} \mathbf{l}^{\mathrm{e}}_{\mathbf{x},j}, \tag{70}$$

with $\mathscr{L} = \{1, 2, \cdots, L\}$ being the set of the indices of the sensors, and $\mathscr{L}\backslash i$ means removing i from the set $\mathscr{L}$.

[3]This calculation is applied for BPSK demodulation, however, for other modulation schemes, the channel LLR is generated by demodulator (DeM) based on BICM-ID method [92].

[4]In this work, the observation error probability p_i is assumed to be known to the fusion center. We can use the proposed algorithm in [55] to accurately estimate them only in the fusion center.

The output of $f_c(\mathbf{l}^{\mathrm{a}}_{\mathbf{x},i}, p_i)$ is interleaved by $\Pi_{i,1}$ to form the *a priori* LLR $\mathbf{l}^{\mathrm{a}}_{\mathbf{x}_i}$ which is fed back to DCC_i in the corresponding i-th local iteration.

(d) Convergence control: For saving the computational power of the destination node, we add the convergence control after the global iteration to remove unnecessary iterations. The mutual information on the sum LLR is estimated using the averaging method after the second global iteration as [93]

$$MI = I(\mathbf{l}^{\mathrm{p}}_{\mathbf{x}}; \mathbf{x}) = 1 - \frac{1}{n}\sum_{t=1}^{n} H_2\Big(\frac{\exp(|l^{\mathrm{p}}_{\mathbf{x}}(t)|/2)}{\exp(|l^{\mathrm{p}}_{\mathbf{x}}(t)|/2) + \exp(-|l^{\mathrm{p}}_{\mathbf{x}}(t)|/2)}\Big). \quad (71)$$

The difference on the mutual information is calculated by subtracting the value of the last iteration from that of the current iteration. If the difference is smaller than a given threshold δ , e.g., 10^{-4}, or the number of total iterations reaches the pre-setting value, the iteration process is terminated.

(e) The decoding process is performed in an iterative manner where steps (b) to (d) are repeated with the switches shown in Fig. 17 being closed.

(f) Hard decision: A hard decision on $\mathbf{x}$ is made to make a final estimate $\hat{\mathbf{x}}$ of $\mathbf{x}$, based on the *a posteriori* LLR $\mathbf{l}^{\mathrm{p}}_{\mathbf{x}}$ originated by summing up all the deinterleaved and f_c-updated versions of the *a posteriori* LLR $\mathbf{l}^{\mathrm{p}}_{\mathbf{x}_i}$.

3.5 Simulations for Verification

A series of simulations have been conducted to verify the performance limits on threshold SNRs and BEP floors. This section provides the results of the simulations.

3.5.1 Parameters in Encoding/Decoding Algorithm

The common parameters assumed in the simulations are summarized in Table 4.

3.5.2 Identical Observation Error Probability p_i

First, we consider the case each sensor has the same observation error probability, let say, $p_i = 0.01$. The analytical results of SNR limits, approximated BEP floors using Poisson binomial and their lower bounds for $L = 2, 4, 6$ and 8 are summarized in Table 5.

Figure 18 illustrates the BER performance using the proposed encoding/joint decoding algorithm. The modulation scheme used here is BPSK. The results of the theoretical

Table 4. The settings of simulation parameters.

Parameter	Value
Block length n	10000 bits
Block	1000
Interleavers	random
Encoder CC_i	Rate $1/2$, $G = (3,2)_8$, memory-1 nonrecursive systematic convolutional code
Doping ratio P_d	1 for BPSK and 8 for QPSK
Modulation	BPSK and QPSK with natural mapping
Decoding Algorithm	log-maximum *a posteriori*
The number of iterations:	50
δ	10^{-4}

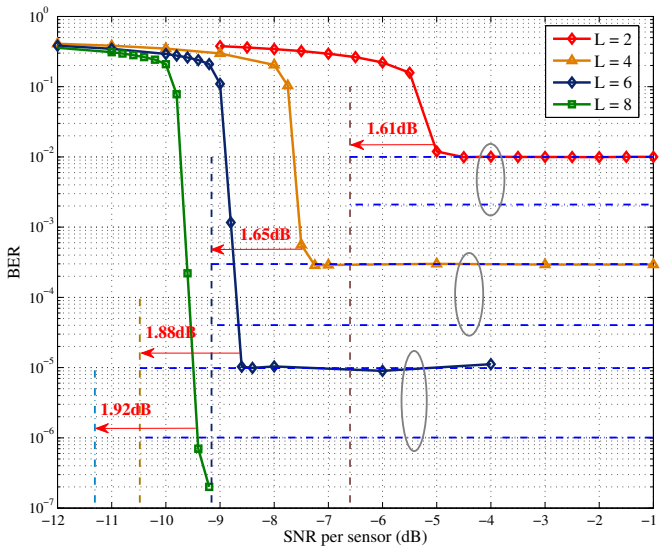


Fig. 18. BER performance of our proposed technique with the number of sensors $L = \{2,4,6,8\}$. p_i are set at 0.01. The corresponding SNR_{lim} are plotted in vertical dash-dot lines. Approximated BEP floors and the lower bounds p_{lb} on the BEP floors are presented in horizontal dashed and dash-dot lines, respectively.

Table 5. The analytical results of the threshold SNR and the BEP floor. (identical p_i)

L	SNR_{lim} (dB)	BEP floor	p_{lb}
2	-6.6	1×10^{-2}	2.1×10^{-3}
4	-9.157	2.98×10^{-4}	4.03×10^{-5}
6	-10.48	9.85×10^{-6}	1.01×10^{-6}
8	-11.31	3.42×10^{-7}	2.84×10^{-8}

analysis presented in Table 5 are also shown as references. It can be clearly seen that in all the cases, the BER curves exhibit a sharp drop in the values at their corresponding threshold SNR values, like turbo-cliff of the turbo-codes. The difference between the threshold SNR obtained by the simulations and the analytical SNR limits shown in Table 5 are $\{1.61, 1.65, 1.88, 1.92\}$ dB for $L = \{2, 4, 6, 8\}$. The gap is because theoretical SNR limits assume that the capacity-achieving code is used at each sensor node, however, the encoding scheme in simulations is not optimized.

Furthermore, such sharp a decrease in BER suddenly plateaus at a value, however, the appearance of the bit error floor is different from that with the turbo codes; the bit error floor is flat with our proposed techniques, while that with the turbo codes it still has a decay which is due to the property of the consistent codes. The approximated BEP floor using Poisson-binomial function $\text{PB}(\cdot)$ is very accurate in these cases. However, the lower bound on the BEP floor is difficult to achieve by the proposed joint decoding algorithm. As it is found from the results, for each parameter case, a gap between the simulations and lower bounds can be clearly observed. The reason is that the lower bound on the BEP floor is derived by assuming there is *no information loss* during the decision process from the observations.

3.5.3 Diverse Observation Error Probability p_i

The impact of variation on p_i to theoretical analyses is considered. The values of p_i are set based on: (i) predefined values, $\{0.002, 0.025, 0.075\}$ for $L = 3$, $\{0.005, 0.0145, 0.015, 0.025, 0.03\}$ for $L = 5$ and $\{0.0003, 0.005, 0.01, 0.015, 0.02, 0.02, 0.05\}$ for $L = 7$; (ii) the pmf of logarithmic distribution $Log(\xi)$, as an example of the case the p_i follows a specific distribution due to some practical reason, as

$$p_i = \frac{-1}{\ln(1-\xi)} \frac{\xi^i}{i}, \tag{72}$$

Table 6. The analytical results of the threshold SNR and the BEP floor. (various p_i)

cases	L	ξ	SNR_{lim} (dB)	BEP floor	p_{lb}
	3		-7.02	2.1×10^{-3}	4.44×10^{-4}
predefined p_i	5	–	-9.223	4.6×10^{-5}	1.58×10^{-5}
	7		-10.28	1.33×10^{-6}	8.37×10^{-8}
	6	0.9	-2.297	0.0148	0.002
$Log(\xi)$	7	0.99	-3.729	6.8×10^{-4}	1.741×10^{-4}
	9	0.97	-4.087	1.488×10^{-4}	2.323×10^{-5}

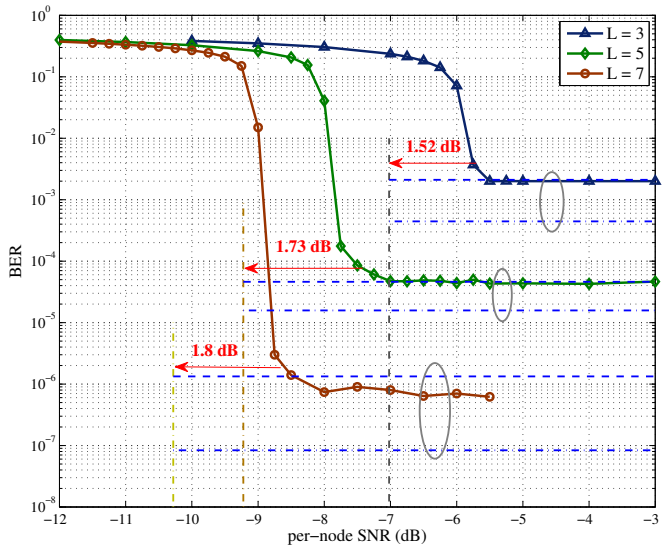


Fig. 19. BER performance comparison with the number of sensors $L = \{3,5,7\}$. The theoretical results including SNR limits and BEP floors are plotted according to the values shown in Table 6.

where $0 < \xi < 1$ is the parameter of $Log(\xi)$ and $\ln(\cdot)$ is the natural logarithm function. If the obtained p_i is larger than 0.5, we simply adjust $p_i = 1 - p_i$.

Figure 19 shows the BER performance with $L = \{3,5,7\}$, where the observation error probability p_i of each sensor was set as the predefined values. Even in these cases, the BER performance gap to the theoretical limits are about $\{1.52, 1.73, 1.8\}$ dB.

Figure 20 depicts the BER performance using logarithmic distributed p_i and QPSK modulation for $L = \{6,7,9\}$. SNR limits and the lower bounds on the BEP floor are also presented in Fig. 20 according to the values of Table 6. The gap between the

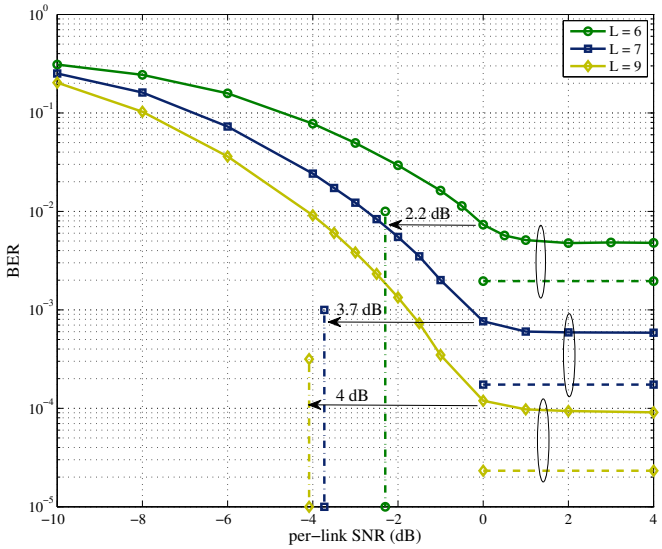


Fig. 20. BER performance comparison with the number of sensors $L = \{6,7,9\}$. The observation error probabilities follows the logarithmic distribution with the parameter ξ shown in shown in Table 6. The modulation of each link is QPSK with natural mapping that results in $R_i^{\mathbf{c}} = 1$ sample per channel use.

SNR limits and the SNR points of turbo cliff are around $\{2.2, 3.7, 4\}$ dB for different number of sensors, which are larger than the previous cases. The reason is because the obtained p_i are relatively large from the logarithmic distribution using the setting ξ and the modulation is QPSK, and thus the BER performance cannot achieve significant gains through global iterations.

It is also found that the bit error floors shown in Figs. 19 and 20 are placed between the results of the Poisson-binomial approximation and the rate-distortion lower bound. In other words, the predicted BEP floors based on Poisson-binomial process are not accurate for the case having various p_i values. As a result, to more accurately predict the system performance with respect to BEP floor, the soft combining need to be taken into account.

We thus derive the BEP floor using different weights based on p_i [61], the detail of which is provided in Appendix 2. We use this method to calculate the BEP floor for the case that the Poisson-binomial is not accurate and list the results in Table 7. Through the

Table 7. The actual bit error floor and predicted BEP floor based on soft combining.

L	p_i	Actual error floor	predicted BEP floor
7	predefined	6.2×10^{-7}	5.44×10^{-7}
6	$p_i \sim Log(0.9)$	0.0048	0.0048
7	$p_i \sim Log(0.99)$	5.9×10^{-4}	5.914×10^{-4}
9	$p_i \sim Log(0.97)$	9.38×10^{-5}	9.628×10^{-5}

comparisons, the predicted BEP floor using this method exactly matches the simulation results. The negligible difference is due to the accuracy of simulations.

Moreover, in the practical encoding/decoding design, source compression is not carried out. The rate loss is large for two cases: (i) source entropy is less than 1; (ii) the channel is good, i.e., the SNR is relatively large. In these two cases, compression should be performed to reduce the redundant information.

3.5.4 Verification by EXIT Analysis

The convergence behavior of the i-th local iteration is evaluated by the EXIT chart analysis [84, 93]. The 3D EXIT chart analysis is used to verify the convergence of local iteration, taking into the account of the helper information of global iteration.

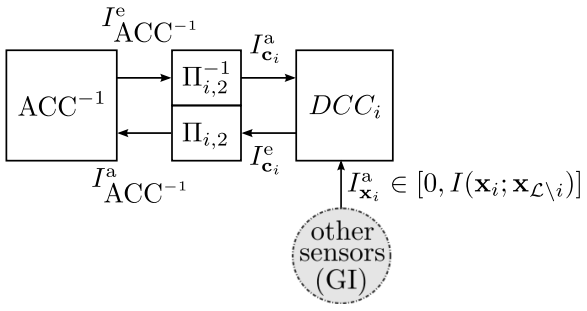


Fig. 21. The abstract model for EXIT analysis. GI: global iteration.

The abstract model indicating the relationships of the mutual information for EXIT chart analysis is shown in Fig. 21. For obtaining the 3D EXIT chart, the *a priori* input LLR fed back to ACC^{-1} was first artificially generated for different values of $I^{\mathrm{a}}_{\mathrm{ACC}^{-1}}$, $0 \leq I^{\mathrm{a}}_{\mathrm{ACC}^{-1}} \leq 1$. The mutual information $I^{\mathrm{e}}_{\mathrm{ACC}^{-1}}$ was then calculated by evaluating the histogram of the output extrinsic LLR output from ACC^{-1}. Furthermore, the EXIT

function for DCC_i is determined by evaluating the histogram of LLR $\mathbf{l}^{\mathrm{e}}_{\mathbf{c}_i}$ in the form of

$$I^{\mathrm{e}}_{\mathbf{c}_i} = T_i(\mathbf{l}^{\mathrm{a}}_{\mathbf{c}_i}, \mathbf{l}^{\mathrm{a}}_{\mathbf{x}_i}), \tag{73}$$

where $I^{\mathrm{e}}_{\mathbf{c}_i}$ is the mutual information between the extrinsic LLR $\mathbf{l}^{\mathrm{e}}_{\mathbf{c}_i}$ and the coded bits $\mathbf{c}_i$, indicating that $I^{\mathrm{e}}_{\mathbf{c}_i} = I(\mathbf{l}^{\mathrm{e}}_{\mathbf{c}_i}; \mathbf{c}_i)$. Since $I^{\mathrm{e}}_{\mathbf{c}_i}$ is the output of a function with two inputs, $\mathbf{l}^{\mathrm{a}}_{\mathbf{c}_i}$ and $\mathbf{l}^{\mathrm{a}}_{\mathbf{x}_i}$, they were artificially generated for different values of $I^{\mathrm{a}}_{\mathbf{c}_i}$, $0 \leq I^{\mathrm{a}}_{\mathbf{c}_i} \leq 1$, and $I^{\mathrm{a}}_{\mathbf{x}_i} = I(\mathbf{l}^{\mathrm{a}}_{\mathbf{x}_i}; \mathbf{x}_i)$, $0 \leq I^{\mathrm{a}}_{\mathbf{x}_i} \leq I(\mathbf{x}_{\mathscr{L}\setminus i}; \mathbf{x}_i)$, respectively.

The mutual information $I^{\mathrm{e}}_{\mathbf{c}_i}$ was then calculated by evaluating the histogram of $\mathbf{l}^{\mathrm{e}}_{\mathbf{c}_i}$ output from DCC_i. Finally, the 3D EXIT chart is plotted based on the values of the mutual information.

The 3D EXIT chart shown in Fig. 22 was obtained by setting $p_i = 0.01$, $i = 1, \cdots, 4$, $SNR = -7.5$ dB, and Fig. 23 obtained by setting p_i as the predefined random values, $SNR - 7$ dB. From Figs. 22 and 23, it is found that the effectiveness of global iteration is significant. The two EXIT planes are stuck at the initial stage, however, the tunnel between two EXIT planes opens as $I^{\mathrm{a}}_{\mathbf{x}_i}$ becomes large. The 3D trajectories obtained by evaluating $I^{\mathrm{a}}_{\mathbf{x}_i}$, $I^{\mathrm{e}}_{\mathrm{ACC}^{-1}}$ and $I^{\mathrm{e}}_{\mathbf{c}_i}$ in the real simulations are also presented in Figs. 22 and 23. As we can see, the trajectories asymptotically reach a point very close to $(1.0, I(\mathbf{x}_i; \mathbf{x}_{\mathscr{L}\setminus i}), 1.0)$ mutual information point, indicating that the information $\mathbf{x}_i$ transmitted from the i-th sensor node can be recovered with an arbitrary low error probability.

3.6 Summary

We investigated the transmission techniques for the binary information sensing network, where multiple sensors observe a common binary source and transmit their erroneous observations to the fusion center. The theoretical limit on the sum rate was analyzed based on the framework of the Slepian-Wolf theorem in lossless distributed source coding problem. We then converted the sum rate to the SNR limit of capacity based on separability theorem. Numerical results shown that the SNR values where turbo cliffs happened in the BER curves are only around $1.5 \sim 2$ dB to the corresponding SNR limit. Furthermore, the EXIT chart analysis clearly verified the practical threshold SNR values in the BER performance. The BEP floor, which is common to the CEO problem, was also approximated by evaluating the Poisson-binomial distribution. The simulation results were very close to the approximated BEP floor. In addition, the lower bound on the BEP floor was derived from the rate-distortion function for binary sources.

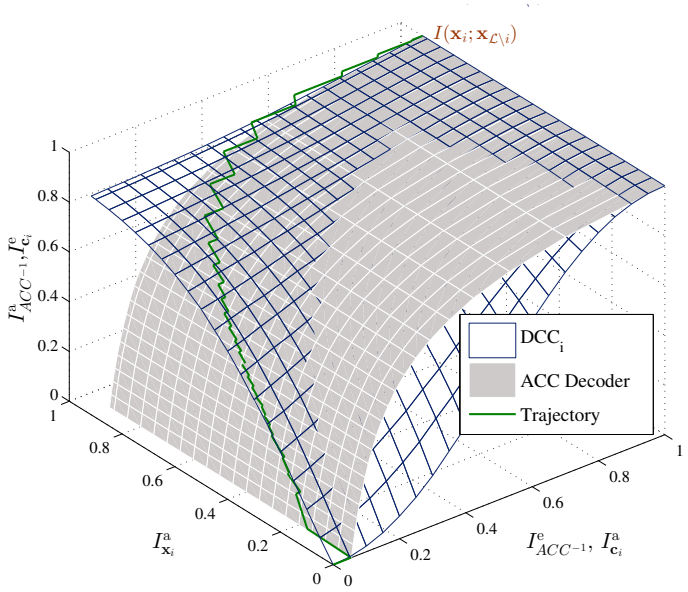


Fig. 22. 3D EXIT curves of ACC^{-1} and DCC_i at $SNR = -7.5$ dB and $L = 4$.

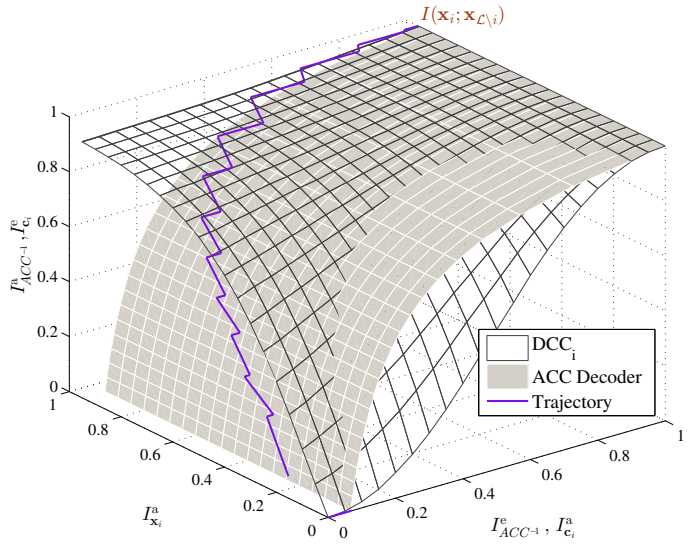


Fig. 23. 3D EXIT curves of ACC^{-1} and DCC_i at $SNR = -7$ dB and $L = 5$.

4 Hamming Distortion Bounds of Binary Information Sensing

We analyzed the theoretical limits in AWGN channels using the Slepian-Wolf theorem for the binary information sensing network in **Chapter 3**. However, the limits are obtained with a strong assumption that the observations are recovered losslessly in usual Shannon sense, which is not always true. Instead, in this chapter, we analyze a problem that how small a Hamming distortion level the fusion center can achieve from the rate-distortion perspective by taking into account distortions of reconstructed observations. To this end, we formulate a minimization problem to obtain the lower bounds on Hamming distortion using the JSC setup with orthogonal MAC components.

In order to solve the minimization problem, we first model the source coding of the binary information sensing network by the binary CEO problem. We then reduce the binary CEO problem to a binary multiterminal source coding problem, which plays the core role in solving the main problem. An outer bound for the rate-distortion region of the binary multiterminal source coding problem is then derived by providing the converse coding proof. We establish the connection with respect to the Hamming distortion level between the binary CEO problem and the binary multiterminal source coding problem. Finally, the minimization problem is formulated in the framework of convex optimization. It should be emphasized here that our purpose is not intended to derive a tight rate-distortion bound for the binary CEO problem. Instead, we focus on the derivation of a lower bound on the Hamming distortion that can be used as a reference of the BER performance curves of the encoding/decoding algorithms, including the technique proposed in our contributions [53, 55].

This chapter is organized as follows. In Section 4.1, the system model and the problem to be solved are described. The derivation of the outer bound and its proof for the binary multiterminal source coding problem with two terminals is detailed in Section 4.2. The problem of how to obtain the lower bound on the Hamming distortion is formulated in Section 4.3. Section 4.4 provides the numerical results of simulations as well as their corresponding lower bounds. We extend the analysis to the binary information sensing network with multiple sensors in Section 4.5. Finally, we conclude this chapter in Section 4.6.

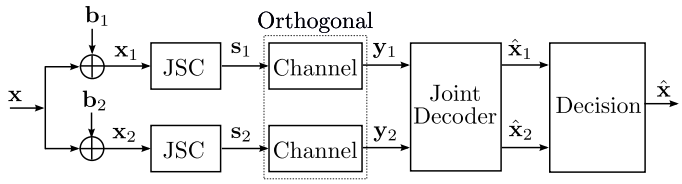


Fig. 24. The abstract system model of estimating a single source through two independent nodes with joint source-channel coding.

4.1 Problem Statement

The system model of estimating a single source through two sensors/terminals is depicted in Fig. 24. A common i.i.d. source X produces a sequence $\mathbf{x} = [x(t)]_{t=1}^{n}$ by taking values from a binary set $\mathscr{X} = \{0,1\}$ with equal probability. Source X is observed by two nodes and forwarded to a single destination. Due to the inaccuracy of the estimation and/or limited received signal power at nodes, such as in WSN and WMN, the sequences received by the nodes may contain errors[5], and the nodes still forward the erroneous sequences to the destination, which is referred as LF [65, 95]. The error probabilities $\Pr(x_1(t) \neq x(t))$ and $\Pr(x_2(t) \neq x(t))$ are denoted as p_1 and p_2, respectively, i.e., $\Pr(b_i(t) = 1) = p_i$ for the binary noise sequence $\mathbf{b}_i = [b_i(t)]_{t=1}^{n}$, $i = 1,2$. At the nodes, the noisy versions $\mathbf{x}_1 = [x_1(t)]_{t=1}^{n}$ and $\mathbf{x}_2 = [x_2(t)]_{t=1}^{n}$ of $\mathbf{x}^n$ are independently encoded by two JSC encoders to generate symbol sequences $\mathbf{s}_1 = [s_1(t)]_{t=1}^{k_1}$ and $\mathbf{s}_2 = [s_2(t)]_{t=1}^{k_2}$ with coding rates $r_i = n/k_i$, $i = 1,2$. The symbol sequences $\mathbf{s}_1$ and $\mathbf{s}_2$ are then transmitted to the destination over two orthogonal AWGN channels, as

$$\mathbf{y}_i = h_i \cdot \mathbf{s}_i + \mathbf{z}_i, \quad i = 1,2, \tag{74}$$

where h_i and $\mathbf{z}_i = [z(t)]_{t=1}^{k_i}$ represent the channel gain and the AWGN sequence at the destination, respectively. The orthogonality can be achieved by any scheduled multiple access scheme, like time division multiple access (TDMA), i.e., $\mathbf{s}_1$ and $\mathbf{s}_2$ can be transmitted at different time intervals. The destination performs JSC decoding to form estimates $\hat{\mathbf{x}}_i$ of the sequences $\mathbf{x}_i$, $i = 1,2$. We define the expected Hamming distortion measures $E[\frac{1}{n}\sum_{t=1}^{n} d(x_i(t), \hat{x}_i(t))] \leq D_i + \varepsilon$ to evaluate the error probability $\Pr(x_i(t) \neq \hat{x}_i(t))$

[5]In WMN applications, the nodes correspond to the transceivers in the multiple routes. In a WMN, a source communicates with a destination through multiple intermediate nodes if they are not within the communication coverage. If errors are allowed in the messages forwarded by the intermediate nodes, the WMN can be also modeled as the model shown in Fig. 24 [94].

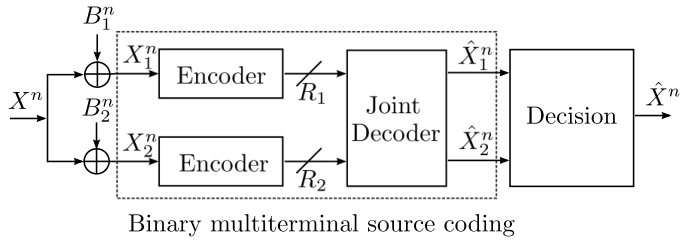


Fig. 25. The abstract model of the binary CEO problem with two independent nodes.

with

$$d(x_i(t),\hat{x}_i(t)) = \begin{cases} 1, & \text{if} \quad x_i(t) \neq \hat{x}_i(t), \\ 0, & \text{if} \quad x_i(t) = \hat{x}_i(t), \end{cases} \tag{75}$$

and ε representing an arbitrarily small positive number.

Finally, the destination reconstructs the source information $\mathbf{x}^n$ of which the estimate is denoted as $\hat{\mathbf{x}}^n$ based on a decision rule from $\hat{\mathbf{x}}_1^n$ and $\hat{\mathbf{x}}_2^n$. Therefore, the distortion measure $E[\frac{1}{n}\sum_{t=1}^n d(x(t),\hat{x}(t))] \leq D+\varepsilon$ can be formulated as a function of D_i, $i = 1,2$, as $D = F_d(D_1,D_2)$, where function $F_d(\cdot)$ is detailed in Section 4.3. It should be emphasized here that function $D = F_d(D_1,D_2)$ limits the decoding scheme to which first reconstructs $\mathbf{x}_1^n$ and $\mathbf{x}_2^n$ and then makes the decision on $\mathbf{x}^n$ from those reconstructions (it is referred to as sequential decoding), as shown in Fig. 24. The optimality of such a decoding scheme is an open problem, but it is definitely of interest for practical systems. Furthermore, $F_d(D_1,D_2)$ largely depends on the decision rule, i.e., there exists different function $F_d(D_1,D_2)$ for different decision rules[6].

According to the source-channel separation theorem for lossy source coding [96], distortion D_1 and D_2 can be achieved if the following inequalities hold:

$$\begin{cases} R_1(D_1)\cdot r_1 & \leq C(\gamma_1), \\ R_2(D_2)\cdot r_2 & \leq C(\gamma_2), \end{cases} \tag{76}$$

where $R_i(D_i)$ is the rate-distortion function for the source coding and $C(\gamma)$ is the Shannon capacity using Gaussian codebook[7] with the argument γ denoting the signal-to-

[6]It has been assumed in this setup that 1) each encoder uses joint typicality encoding and binning based on random coding arguments, and the decoder performs joint typicality decoding with a sufficiently large n to achieve the average distortion D_i as in the Berger-Tung source coding problem [78]; 2) the errors occurring in each sequence $\mathbf{x}_i^n$ are i.i.d. In the practical system, we use random interleavers to asymptotically make this assumption practical. As shown in Section 4.4, the simulation results are consistent with the lower bound calculation based on soft combining decision.

[7]For one-dimensional signal, $C(\gamma) = \frac{1}{2}\log_2(1+2\gamma)$, and for two-dimensional signal, $C(\gamma) = \log_2(1+\gamma)$ [74].

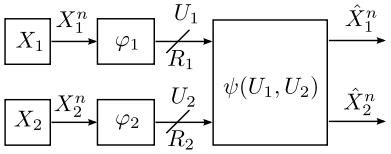


Fig. 26. The binary multiterminal source coding problem for two correlated binary sources.

noise ratio (SNR) of the channel. As stated above, r_1 and r_2 represent end-to-end coding rates of two links. Our goal is to derive the theoretical lower bound on the Hamming distortion for the system shown in Fig. 24. It is equivalent to minimizing the expected Hamming distortion D through a function $F_d(D_1,D_2)$ under constraints shown in (76), as

$$\min_{D_1,D_2} \quad D = F_d(D_1,D_2) \tag{77}$$

$$\text{s.t.} \tag{78}$$

$$R_1(D_1)\cdot r_1 \le C(\gamma_1),$$

$$R_2(D_2)\cdot r_2 \le C(\gamma_2),$$

The minimization being performed in (77) is for a specific system which maps the average distortions D_1 and D_2 to D, since function $F_d(D_1,D_2)$ is defined for designated decision rules. To achieve this goal by solving (77), we turn to derive the rate-distortion function $R_i(D_i)$ for the problem shown in Fig. 24 and to establish the function $D = F_d(D_1,D_2)$ for the decision rule used at the destination.

4.2 Rate-Distortion Region Analysis

4.2.1 Outer Bound on the Rate-Distortion Region

Source Coding

In network information theory, the source coding of the communication system shown in Fig. 24 is modeled by the binary CEO problem. The abstract model of the binary CEO problem is illustrated in Fig. 25. In order to derive the rate-distortion function $R_i(D_i)$, we first reduce the binary CEO problem to a binary multiterminal source coding problem. An outer bound for the rate-distortion region which is determined by the rate-distortion function $R_i(D_i)$ is then derived for the binary multiterminal source coding problem through the converse proof, as in the Gaussian case [24].

The binary multiterminal source coding problem which we consider is depicted in Fig. 26. Since random sources X_1^n and X_2^n originate from the common source X^n, the random variable pair (X_1, X_2) follow a joint pmf $p_{X_1X_2}(x_1,x_2) = \Pr\{X_1 = x_1, X_2 = x_2\}$ given by

$$p_{X_1X_2}(x_1,x_2) = \begin{cases} \frac{1}{2}\rho, & \text{if } x_1 \neq x_2, \\ \frac{1}{2}(1-\rho), & \text{otherwise}, \end{cases} \tag{79}$$

where $\rho = \Pr(x_1 \neq x_2)$ is the correlation parameter between the sources X_1 and X_2, i.e., X_2 can be seen as the output of a BSC with the crossover probability ρ where X_1 is the input. Two encoders independently encode X_1^n and X_2^n at rates R_1 and R_2 as

$$\varphi_1 : \mathscr{X}^n \to \mathscr{M}_1 = \{1, 2, \cdots, 2^{nR_1}\},$$
$$\varphi_2 : \mathscr{X}^n \to \mathscr{M}_2 = \{1, 2, \cdots, 2^{nR_2}\}.$$

The encoder output sequences $U_1 = \varphi_1(X_1^n)$ and $U_2 = \varphi_2(X_2^n)$ are transmitted to a common receiver. It jointly decodes the received samples to construct the estimates $(\hat{X}_1^n, \hat{X}_2^n)$ of the source pair (X_1^n, X_2^n) denoted as $(\hat{X}_1^n, \hat{X}_2^n) = \psi(\varphi_1(X_1^n), \varphi_2(X_2^n))$.

For given distortion values $D_1 \in [0, \frac{1}{2}]$ and $D_2 \in [0, \frac{1}{2}]$, the rate-distortion region $\mathscr{R}(D_1, D_2)$ is defined as

$$\mathscr{R}(D_1,D_2) = \Big\{(R_1,R_2) : (R_1,R_2) \text{ is admissible such that } E\frac{1}{n}\sum_{t=1}^{n} d(x_i(t), \hat{x}_i(t)) \leq D_i + \varepsilon, i = 1,2\Big\}.$$

It should be emphasized here that the admissible rate-distortion region may *not* applied to the binary CEO problem directly, since the strategy at the CEO is specified to two-step decoding. The admissible rate-distortion region defined above limits the problem which has the specific setup.

Main Results

We provide a bound $\mathscr{R}^{\mathrm{o}}(D_1, D_2)$ of the rate-distortion region $\mathscr{R}(D_1, D_2)$.

Definition 2. Let

$$\mathscr{R}_1^{\mathrm{o}}(D_1) = \Big\{(R_1,R_2) : R_1 \geq H_2[\rho * H_2^{-1}(1-[R_2]^-)] - H_2(D_1)\Big\}, \tag{80}$$

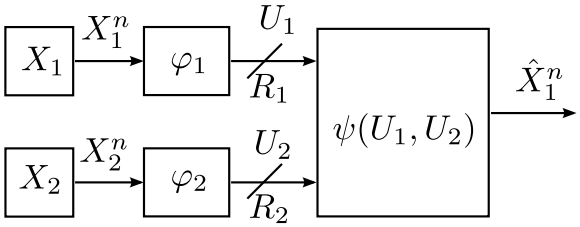


Fig. 27. Case 1: X_2 **acts as the side information.**

$$\mathscr{R}_2^{\mathrm{o}}(D_2) = \Big\{(R_1,R_2): \\ R_2 \geq H_2[p * H_2^{-1}(1-[R_1]^-)] - H_2(D_2)\Big\}, \tag{81}$$

and

$$\mathscr{R}_{12}^{\mathrm{o}}(D_1,D_2) = \Big\{(R_1,R_2): \\ R_1+R_2 \geq 1+H_2(p)-H_2(D_1)-H_2(D_2)\Big\}, \tag{82}$$

with $[R_i]^- = \min\{1,R_i\}$.

For every $D_1 \in [0,\frac{1}{2}]$ and $D_2 \in [0,\frac{1}{2}]$,

$$\mathscr{R}^{\mathrm{o}}(D_1,D_2) = \mathscr{R}_1^{\mathrm{o}}(D_1) \bigcap \mathscr{R}_2^{\mathrm{o}}(D_2) \bigcap \mathscr{R}_{12}^{\mathrm{o}}(D_1,D_2). \tag{83}$$

In what follows, we prove that $\mathscr{R}_1^{\mathrm{o}}(D_1)$, $\mathscr{R}_2^{\mathrm{o}}(D_2)$ and $\mathscr{R}_{12}^{\mathrm{o}}(D_1,D_2)$ are the supersets of the regions of $\mathscr{R}(D_1,D_2)$. It means that the following theorem holds.

Theorem 16 (Outer bound on rate-distortion region [60])**.** *$\mathscr{R}^o(D_1,D_2)$ is an outer bound for the rate-distortion region $\mathscr{R}(D_1,D_2)$; i.e., $\mathscr{R}(D_1,D_2) \subseteq \mathscr{R}^o(D_1,D_2)$.*

Proofs

Proof of Theorem 16 (Converse). To prove Theorem 16, the following three different cases of the binary multiterminal source coding problem are considered.

Case 1. In order to prove that $\mathscr{R}(D_1,D_2) \subseteq \mathscr{R}_1^{\mathrm{o}}(D_1)$, we assume that the rate pair $(R_1,R_2) \in \mathscr{R}(D_1,D_2)$ and show that this implies that $(R_1,R_2) \in \mathscr{R}_1^{\mathrm{o}}(D_1)$. In the proof, X_2^n is first reconstructed without constraint on D_2 which results in (89), and then $\hat{X}_2^n$ is

regarded as the side information to recover X_1^n, as shown in Fig. 27. Assume that a rate pair (R_1, R_2) achieves distortion D_1, then

$$\begin{aligned} n\cdot(R_1+\varepsilon) &\geq H(U_1) \\ &\geq H(U_1|U_2) && (84)\\ &= I(X_1^n;U_1|U_2) && (85)\\ &= I(X_1^n;U_1,U_2)-I(X_1^n;U_2) && (86)\\ &\geq I(X_1^n;\hat{X}_1^n)-I(X_1^n;U_2), && (87)\end{aligned}$$

where the steps are justified, because
(84) conditioning reduces entropy,
(85) U_1 is a function of X_1^n,
(86) the chain rule of mutual information,
(87) $X_1^n \to (U_1,U_2) \to \hat{X}_1^n$ forms a Markov chain.

Now our aim is to lower bound $I(X_1^n;\hat{X}_1^n)$ and upper bound $I(X_1^n;U_2)$. Since $I(X_1^n;\hat{X}_1^n) = H(X_1^n)-H(X_1^n|\hat{X}_1^n) = n-H(X_1^n|\hat{X}_1^n)$, to lower bound $I(X_1^n;\hat{X}_1^n)$ is equivalent to upper bound $H(X_1^n|\hat{X}_1^n)$. According to the Fano's inequality, we have

$$\begin{aligned} H(X_1^n|\hat{X}_1^n) &\leq n\cdot H_2(D_1)+n\cdot D_1\cdot\log(|\mathscr{X}|-1) \\ &= n\cdot H_2(D_1). && (88)\end{aligned}$$

On the other hand, since $I(X_1^n;U_2) = H(X_1^n)-H(X_1^n|U_2) = n-H(X_1^n|U_2)$, an upper bound on $I(X_1^n;U_2)$ corresponds to the lower bound on $H(X_1^n|U_2)$. Observing that $X_1^n \to X_2^n \to U_2$ forms a Markov chain, it can be shown that $H(X_1^n|U_2) \geq nH_2(\rho * \beta)$ by MGL, where $\beta = \frac{1}{n}H_2^{-1}[H_2(X_2^n|U_2)]$. Since the binary convolution $*$ is monotonically increasing with respect to β if ρ is fixed, we need to find the minimizing value of β to lower bound $H_2(\rho*\beta)$[8]. We also have the rate constraint on R_2 as

$$\begin{aligned} n\cdot(R_2+\varepsilon) &\geq H(U_2) \\ &= I(X_2^n;U_2). && (89)\end{aligned}$$

Letting $\varepsilon \to 0$, we have $n\cdot R_2 \geq I(X_2^n;U_2) = H(X_2^n)-H(X_2^n|U_2) = n-nH_2(\beta)$, and hence $\beta \geq H_2^{-1}(1-R_2)$. Therefore, the lower bound on $H(X_1^n|U_2)$ is given by

$$H(X_1^n|U_2) \geq n\cdot H_2[\rho * H_2^{-1}(1-R_2)]. \quad (90)$$

[8]The binary entropy function is a monotonically increasing function in the interval $[0,\frac{1}{2}]$.

From (87), (88) and (90), we can obtain

$$\begin{aligned} & n \cdot (R_1 + \varepsilon) \\ & \quad \geq n - n \cdot H_2(D_1) - n + n \cdot H_2[\rho * H_2^{-1}(1 - R_2)] \\ & \quad = n \cdot H_2[\rho * H_2^{-1}(1 - R_2)] - n \cdot H_2(D_1). \end{aligned} \tag{91}$$

The rate-distortion region $\mathscr{R}_1^{\mathrm{o}}(D_1)$ shown in (80) is obtained by letting $\varepsilon \to 0$ in (91). Thus, the rate pair (R_1, R_2) satisfies condition (80); i.e., $(R_1, R_2) \in \mathscr{R}_1^{\mathrm{o}}(D_1)$.

Case 2. The source X_1^n acts as a helper to reconstruct X_2^n under the required distortion level D_2. This is the case symmetric with *Case 1*. The rate-distortion region shown in (81) can be proved in the same way as in *Case 1*.

Case 3. Here, we prove that $(R_1, R_2) \in \mathscr{R}(D_1, D_2)$ implies $(R_1, R_2) \in \mathscr{R}_{12}^{\mathrm{o}}(D_1, D_2)$. To this end, assume that the rate pair (R_1, R_2) achieves distortion D_1 for X_1 and D_2 for X_2. In the following proof, decoder ψ jointly reconstructs the sources X_1^n and X_2^n under required distortions D_1 and D_2. The following inequalities are obtained:

$$\begin{aligned} n \cdot (R_1 + R_2 + \varepsilon) & \geq H(U_1) + H(U_2) \\ & \geq H(U_1, U_2) \\ & = I(X_1^n, X_2^n; U_1, U_2) \qquad (92) \\ & = I(X_1^n; U_1, U_2) + I(X_2^n; U_1, U_2 | X_1^n) \\ & = I(X_1^n; U_1, U_2) + I(X_2^n; X_1^n, U_1, U_2) - n \cdot I(X_1; X_2) \\ & \geq I(X_1^n; \hat{X}_1^n) + I(X_2^n; \hat{X}_2^n) - n \cdot I(X_1; X_2), \qquad (93) \end{aligned}$$

where (92) holds since U_i is a function of X_i^n, $i = 1, 2$. Similarly, by utilizing Fano's inequality to lower bound $I(X_1^n; \hat{X}_1^n)$ and $I(X_2^n; \hat{X}_2^n)$, we have

$$n \cdot (R_1 + R_2 + \varepsilon) \geq n + n \cdot H_2(\rho) - n \cdot H_2(D_1) - n \cdot H_2(D_2). \tag{94}$$

Letting $\varepsilon \to 0$ in the above inequality, we conclude that (82) holds. That is, $(R_1, R_2) \in \mathscr{R}_{12}^{\mathrm{o}}(D_1, D_2)$.

Through these three cases, it can be concluded that the admissible rate pair $(R_1 + \varepsilon, R_2 + \varepsilon) \in \mathscr{R}^{\mathrm{o}}(D_1, D_2)$. Furthermore, the monotonicity of the outer bound $\mathscr{R}^{\mathrm{o}}(D_1, D_2)$ [23] implies that $\mathscr{R}^{\mathrm{o}}(D_1, D_2) \subseteq \mathscr{R}^{\mathrm{o}}(D_1 + \varepsilon, D_2 + \varepsilon)$. Since (R_1, R_2) is admissible, we conclude that $\mathscr{R}(D_1, D_2) \subseteq \mathscr{R}^{\mathrm{o}}(D_1, D_2)$ by letting $\varepsilon \to 0$. □

In summary, the outer bound on the rate-distortion function $R_i(D_i)$ is given by

$$\begin{cases} R_1(D_1) \geq H_2[\rho * H_2^{-1}(1-R_2(D_2))] - H_2(D_1), \\ R_2(D_2) \geq H_2[\rho * H_2^{-1}(1-R_1(D_1))] - H_2(D_2), \\ \sum_{i=1}^{2} R_i(D_i) \geq 1 + H_2(\rho) - \sum_{i=1}^{2} H_2(D_i). \end{cases} \tag{95}$$

4.2.2 Inner Bound

As it is known that the exact rate-distortion bound for lossy multiterminal source coding problem lies between the Berger-Tung inner and outer bounds [78]. We also derived the rate-distortion region $\mathscr{R}^{\mathrm{i}}(D_1,D_2)$ based on the Berger-Tung inner bound [57] as

$$\mathscr{R}^{\mathrm{i}}(D_1,D_2) = \mathscr{R}_1^{\mathrm{i}}(D_1) \bigcap \mathscr{R}_2^{\mathrm{i}}(D_2) \bigcap \mathscr{R}_{12}^{\mathrm{i}}(D_1,D_2) \tag{96}$$

with

$$\begin{cases} \mathscr{R}_1^{\mathrm{i}}(D_1) = \{(R_1,R_2) | R_1 \geq H_2(\rho * D_1 * D_2) - H_2(D_1)\}, \\ \mathscr{R}_2^{\mathrm{i}}(D_2) = \{(R_1,R_2) | R_2 \geq H_2(\rho * D_1 * D_2) - H_2(D_2)\}, \\ \mathscr{R}_{12}^{\mathrm{i}}(D_1,D_2) = \{(R_1,R_2) | \\ \qquad R_1 + R_2 \geq 1 + H_2(\rho * D_1 * D_2) - \sum_{i=1}^{2} H_2(D_i)\}, \end{cases}$$

for every $0 \leq D_1, D_2 \leq \frac{1}{2}$. The detailed proof is given in Appendix 3.

4.2.3 Remarks

We now show that the derived outer and inner bounds on the rate-distortion region is connected to the classical results mentioned in Chapter 2.

Remark 1. If either $R_1 = 0$ or $R_2 = 0$, i.e., one of two encoders is breakdown in the network, $\mathscr{R}^{\mathrm{o}}(D_1,D_2)$ is then consistent with the classical rate-distortion function $1 - H_2(D_i)$ for the binary source.

Remark 2. If the distortions D_1 and D_2 are required to be arbitrarily small, then $\mathscr{R}^{\mathrm{o}}(D_1,D_2)$ reduces to the Slepian-Wolf rate region [2] for correlated binary sources if we set $D_1 \to 0$ and $D_2 \to 0$. The Slepian-Wolf rate region and $\mathscr{R}^{\mathrm{o}}(D_1,D_2)$ are shown in Fig. 29. Obviously, it is found that by allowing nonzero distortion, the sources can be further compressed compared to the Slepian-Wolf lossless case.

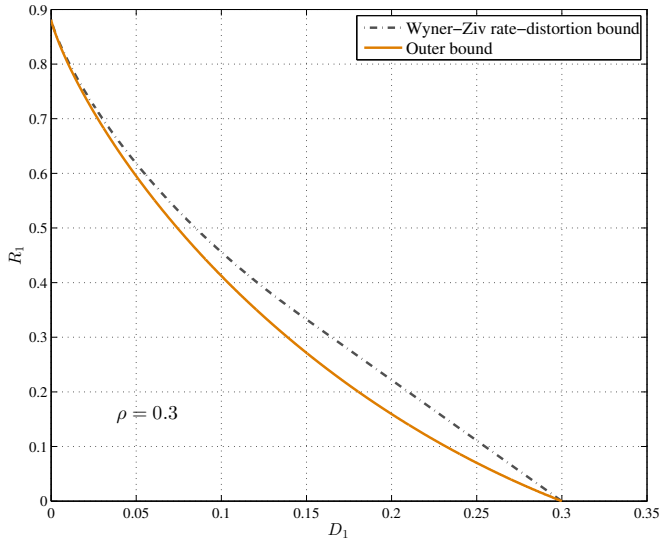
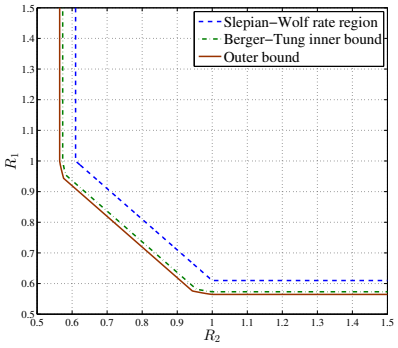


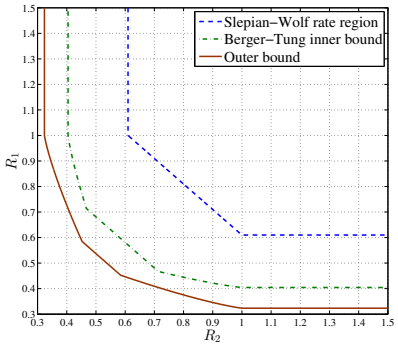
Fig. 28. The comparison of Wyner-Ziv rate-distortion bound and derived outer bound. The correlation ρ between two sources is set at 0.3.

Remark 3. If we are interested in reconstructing only one of the two sources, say X_1, and there is no rate limit on describing X_2^n, i.e., $R_2 \geq \frac{1}{n}H(X_2^n)$, then it is equivalent to the Wyner-Ziv compression problem [78]. Fig. 28 plots the rate-distortion bound $R_1^{\mathrm{WZ}}(D_1)$ of the Wyner-Ziv source coding [11] and our derived outer bound. In this case, $\mathscr{R}_1^{\mathrm{o}}(D_1)$ is not tight, since it can be found from Fig. 28 that the rate-distortion region of the Wyner-Ziv problem lies inside of $\mathscr{R}_1^{\mathrm{o}}(D_1)$.

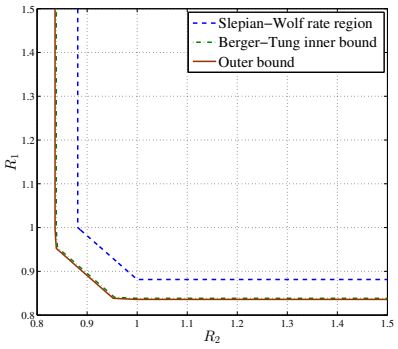
Remark 4. In Fig. 29, the Berger-Tung inner bound for binary case $\mathscr{R}^{\mathrm{i}}(D_1, D_2)$ is also presented as a reference to verify how close the bounds $\mathscr{R}^{\mathrm{o}}(D_1, D_2)$ and $\mathscr{R}^{\mathrm{i}}(D_1, D_2)$ are. It can be seen from the figure that they are very close to each other for small values of D_1 and D_2, i.e., the outer bound can be considered as a useful reference in the evaluation of the BER performance, even though there exists a small gap. The gap between the Berger-Tung inner bound and the derived outer bound is sensitive to both ρ and D_i. If ρ goes large and/or D_i small, the gap becomes relatively small. However, to resolve this gap, further insightful discussions are still needed as in Gaussian multiterminal source coding [24].



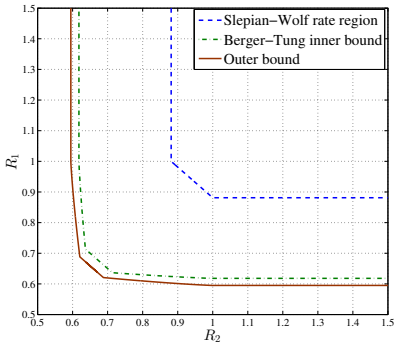
(a) $\rho = 0.15$, $D_1 = D_2 = 0.005$.



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



(c) $\rho = 0.3$, $D_1 = D_2 = 0.005$.



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

Fig. 29. The comparison of $\mathscr{R}^{\mathbf{o}}(D_1, D_2)$, Berger-Tung inner bound and Slepian-Wolf admissible rate region.

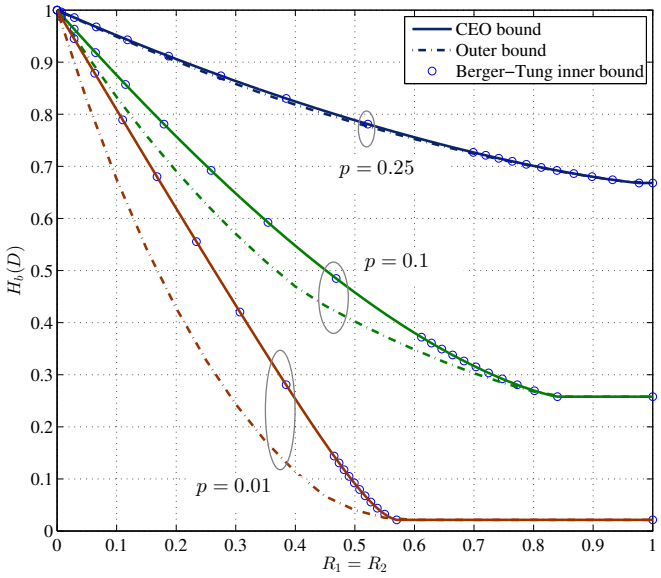


Fig. 30. The comparison of two-step outer bound and Berger-Tung inner bound with optimal decision, and the direct outer bound of the binary CEO problem. We assume that $p = p_1 = p_2$ and $R_1 = R_2$.

Remark 5. In Fig. 30, we compare three rate-distortion bounds for the binary CEO problem. As a reference, we directly derived the outer bound on the rate-distortion region for the two-node binary CEO problem, which is summarized in Appendix 4. This outer bound is referred as direct outer bound. Our derived outer bound with optimal decision, which is stated in the next section, is not tight for the binary CEO problem. However, if the observation accuracy is low, let say, $p_1 = p_2 = 0.25$, the gap between the derived two-step outer bound and the direct outer bound is negligible. Furthermore, for large rate R_i, two bounds exactly match with each other. Hence, the derived outer bound is tight for relatively large p and/or R_i based on the results shown in Fig. 30, which is consistent with the above discussion in Remark 4. Also, it is interesting to find that the Berger-Tung inner bound with optimal decision coincides with the direct outer bound for any p_i. As a result, it is concluded that the binary CEO problem with two nodes is solved.

4.3 Problem Formulation: Hamming Distortion Lower Bounds

4.3.1 Distortion Function

As stated in Section 4.1, distortion D is a function of distortions D_i, $i = 1,2$. Function $F_d(D_1,D_2)$ is obtained by evaluating the relationship between the binary CEO and the binary multiterminal source coding problems in terms of distortions, where the model of the relationship is shown in Fig. 31. The estimate $\hat{X}$ is obtained based on the decision rule from the outputs of two parallel BSCs with crossover probabilities $p_1 * D_1$, $p_2 * D_2$ and input X. The distortion D largely depends on the decision rule used by the destination. Here we only consider two decision rules. One is the soft combining decision and the other the optimal decision.

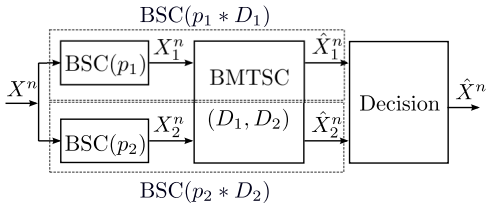


Fig. 31. The relationship between the binary CEO and multiterminal source coding problems. BMTSC: binary multiterminal source coding.

Soft combining decision

Distortion D is obtained by evaluating the probability of an error event. Let $\theta_1 = p_1 * D_1$ and $\theta_2 = p_2 * D_2$. Without loss of generality, we assume that $\theta_1 \leq \theta_2$. Hence, the error event is composed of two independent events: node 1 makes a wrong decision and node 2 makes correct decision or both node 1 and node 2 make erroneous decisions. Therefore, the distortion D in this case is approximated by $D \cong \theta_1(1-\theta_2) + \theta_1\theta_2 = \theta_1$. It can be found that the corner point θ_1 or θ_2 in the rate-distortion region is achieved. Hence, the soft combining decision rule can be seen as being equivalent to that derived from the time sharing method.

Optimal decision

Since the block length is assumed to be infinite and the code is random, an optimal lower bound on the distortion D is determined by applying Proposition 15, as

$$D = H_2^{-1}[H_2(\theta_1) + H_2(\theta_2) - H_2(\theta_1 * \theta_2)]. \tag{97}$$

It should be emphasized here that the optimal decision acts as a universal lower bound on the Hamming distortion for specific schemes which assume sequential decoding. However, in the design of practical encoding/decoding algorithms, we do not consider this decision rule.

In summary, the distortion level D of the two decision rules described above is given as

$$D = F_d(D_1, D_2) = \begin{cases} \min\{\theta_1, \theta_2\}, & \text{soft combining,} \\ H_2^{-1}[H_2(\theta_1) + H_2(\theta_2) - H_2(\theta_1 * \theta_2)], & \text{optimal.} \end{cases} \tag{98}$$

4.3.2 Convex Optimization: Minimizing Distortion

By substituting the rate-distortion function (95) and (98) into the minimization problem (77), we have

$$\begin{aligned} \min_{D_1, D_2} \quad & D \\ \text{s.t.} \quad & \\ & H_2[\rho * H_2^{-1}(1 - \tfrac{C(\gamma_2)}{r_2})] - H_2(D_1) \leq \frac{C(\gamma_1)}{r_1}, \\ & H_2[\rho * H_2^{-1}(1 - \tfrac{C(\gamma_1)}{r_1})] - H_2(D_2) \leq \frac{C(\gamma_2)}{r_2}, \\ & 1 + H_2(\rho) - H_2(D_1) - H_2(D_2) \leq \frac{C(\gamma_1)}{r_1} + \frac{C(\gamma_2)}{r_2}, \\ & D_i \leq \frac{1}{2}, \quad i = 1, 2, \\ & D_i \geq 0, \quad i = 1, 2. \end{aligned} \tag{99}$$

The reason of using the derived outer bound, not the Berger-Tung inner bound is that, the outer bound can be easily formulated as a convex optimization. The Berger-Tung inner bound includes term $D_1 * D_2$ in the binary entropy function which cannot be easily handled in the minimization. It is found that distortion $D = F_d(D_1, D_2)$ is monotonically increasing function on the intervals $D_i \in [0, \frac{1}{2}]$, $i = 1, 2$ for both the soft

combining decision and optimal decision rules, and the proof is detailed in Appendix 5. Furthermore, since the sequential decoding (first reconstructs $\mathbf{x}_1$ and $\mathbf{x}_2$, then makes decision on $\mathbf{x}$) is applied, we first minimize the ℓ_2-norm of a vector $[D_1, D_2]$ instead of directly minimizing D, as

$$\begin{aligned}
\min_{D_1,D_2} \quad & \|[D_1, D_2]\|_2 \\
\text{s.t.} \quad & \\
-H_2(D_1) - H_2(D_2) &\leq \frac{C(\gamma_1)}{r_1} + \frac{C(\gamma_2)}{r_2} - 1 - H_2(\rho), \\
-H_2(D_1) &\leq \frac{C(\gamma_1)}{r_1} - H_2[\rho * H_2^{-1}(1 - \frac{C(\gamma_2)}{r_2})], \\
-H_2(D_2) &\leq \frac{C(\gamma_2)}{r_2} - H_2[\rho * H_2^{-1}(1 - \frac{C(\gamma_1)}{r_1})], \\
D_i &\leq \frac{1}{2}, \quad i = 1, 2, \\
-D_i &\leq -0, \quad i = 1, 2,
\end{aligned} \tag{100}$$

to obtain the minimal values of D_1 and D_2, and then map them to D by using function $F_d(D_1, D_2)$.

It is easily found that the problem (100) is convex since the objective function is convex and function $-H_2(\cdot)$ is also convex. Therefore, it can be efficiently solved using convex optimization tools. Assume that the minimum values of D_1 and D_2 obtained through the convex optimization are denoted as $D_1^\star$ and $D_2^\star$, respectively. Substituting $D_1^\star$ and $D_2^\star$ into (98), the minimum distortion value $D^\star$ is then obtained through

$$D^\star = \begin{cases} \min\{\theta_1^\star, \theta_2^\star\}, & \text{soft combining,} \\ H_2^{-1}[H_2(\theta_1^\star) + H_2(\theta_2^\star) - H_2(\theta_1^\star * \theta_2^\star)], & \text{optimal,} \end{cases} \tag{101}$$

where $\theta_1^\star$ and $\theta_2^\star$ are $p_1 * D_1^\star$ and $p_2 * D_2^\star$, respectively. It should be emphasized here that the distortion D_1 or D_2 should be set to 0 in the optimization problem (99) if $\frac{C(\gamma_1)}{r_1}$ or $\frac{C(\gamma_2)}{r_2}$ is larger than or equal to 1, which is the binary entropy of the source X_1 and X_2. The reason is that a source can be reconstructed under an arbitrary small error probability if the source coding rate is larger than its entropy even in the case the helper does not exist [74].

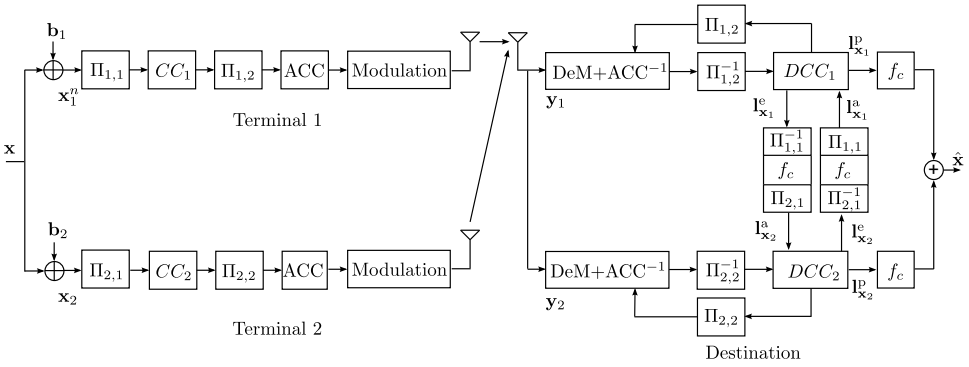


Fig. 32. Block diagram of the encoding/decoding algorithm.

4.4 Verification of Hamming Distortion Lower Bounds

4.4.1 Simulation Settings

We briefly explain the practical encoding/decoding algorithm [53, 55] which is illustrated in Fig. 32. This algorithm is used to verify the theoretical Hamming distortion lower bounds. As illustrated in Fig. 32, each node encodes its erroneous sequence by using a serially concatenated memory-1 convolutional code and ACC. The encoder output sequences are then modulated and transmitted to the destination over statistically independent AWGN and block Rayleigh fading channels, where the channel gain h_i is static within each block but varies independently block-by-block. At the destination, iterative decoding process is carried out between the decoders of the convolutional code and the ACC, as well as between the two decoders of the convolutional codes through the LLR updating function f_c to modify the extrinsic LLR, according to the error probabilities p_1 and p_2.

The lower bounds[9] on the Hamming distortion for different SNR values γ_1, γ_2 are obtained through solving the convex optimization problem which we presented in Section 4.3. The results are shown in Figs. 33–36 for AWGN channels and Fig. 37 for block Rayleigh fading channels. The common parameters used in conducting the simulations are shown in Table 8.

[9]The terminology "lower bound" used here is due to the Hamming distortion is calculated based on the derived outer bound, even though the approximation of the objective functions is used.

Table 8. The settings of simulation parameters.

Parameter	Value
Block length n	10000 bits for AWGN and 2048 bits for fading
Block	1000 for AWGN and 10000 for fading
Interleavers	random
Encoder CC_i	Rate $1/2$, $G = (3,2)_8$, memory-1 nonrecursive systematic convolutional code
Doping ratio P_d	1
Modulation	BPSK and QPSK with natural mapping
Decoding Algorithm	log-maximum *a posteriori*
The number of iterations:	30

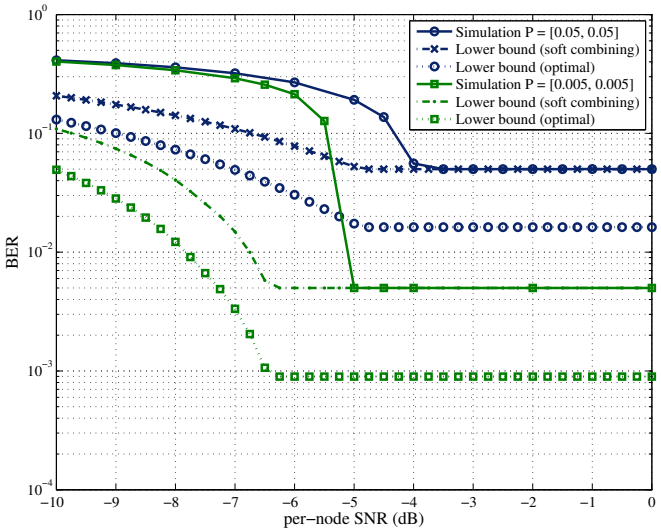


Fig. 33. Symmetric P and SNR. BPSK is used for both nodes.

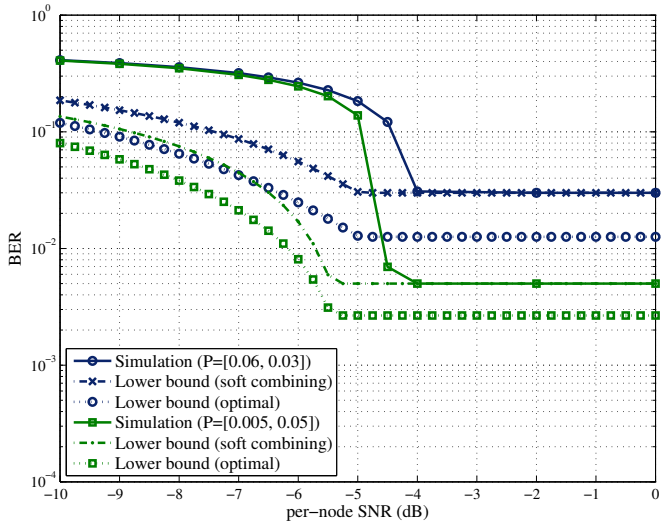


Fig. 34. Asymmetric P and symmetric SNR. BPSK is used for both nodes.

4.4.2 Numerical Results

Figure 33 shows the error probability lower bounds and the BER versus SNR when p_1, p_2 and SNRs of the two nodes are set identically; this is referred as the symmetric case. It can be found that, the BER curves obtained by simulations and the theoretical lower bounds on the Hamming distortion exhibit a similar tendency. The gap between the simulated BER and theoretical lower bound on Hamming distortion is caused by: (i) the derived outer bound is not tight, and thus smaller Hamming distortion is obtained for fixed rates; (ii) the Hamming distortion lower bound is obtained by assuming the optimal source coding rate is adopted based on separability, however, fixed coding rate is used in simulations.

Furthermore, it is clearly found that the error floor of the BER obtained by the simulation and the lower bound on the Hamming distortion based on soft combining match exactly. The reason is that if the SNRs of two nodes are large enough, the distortion levels D_1 and D_2 are almost 0, which results in the error floor being determined completely by the error probabilities p_1 and p_2. A gap clearly appears between the Hamming distortion lower bounds using the soft combining and optimal decision rules. The reason is twofold: 1) the optimality of the soft combining cannot be guaranteed; 2) optimal decision is derived based on the assumption of the binary rate-distortion function with-

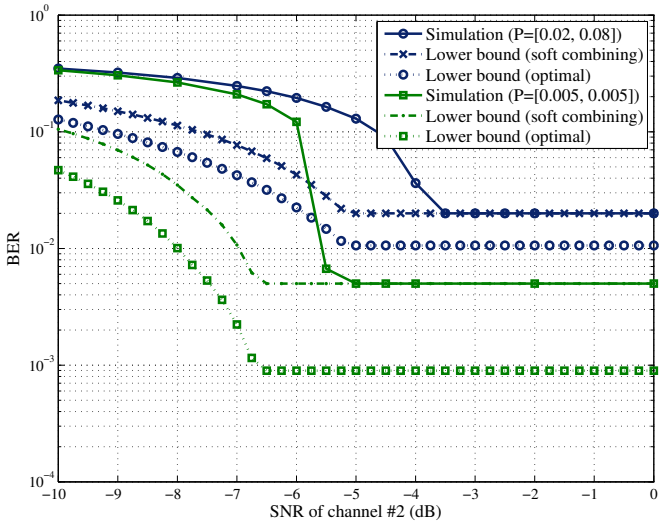


Fig. 35. Asymmetric r_1 and r_2. The coding rates r_1 and r_2 are set at $\frac{1}{4}$ and $\frac{1}{2}$, respectively. The transmit power of two nodes is the same. BPSK is used for both nodes.

out any loss during processing the information. To find a better decision rule than soft combining rule is left as a future study. However, it is clear that the Hamming distortion lower bound deriving from the optimal decision cannot be exceeded.

The impact of the variation of the error probabilities p_1, p_2 and the coding rates r_i are evaluated in AWGN channels. Fig. 34 shows the results for asymmetric p_1 and p_2 but symmetric SNRs. When the coding rates[10] r_1 and r_2 are set as $\frac{1}{4}$ and $\frac{1}{2}$, respectively, the BER performance shown in Fig. 35 is obtained. We further consider using different modulation schemes for the nodes to achieve different rates of the channel code in Fig. 36, where QPSK is used for node 1 and BPSK for node 2. Even in these asymmetric cases, the theoretical lower bounds on the Hamming distortion can still provide us with a useful reference when we evaluate the BER performance of practical systems. Furthermore, the theoretical lower bounds on the Hamming distortion obtained based on our derived outer bound exhibit similar behaviors to those of the BER curves found by simulations.

[10]We simply transmit the output of ACC without doping to achieve rate $\frac{1}{4}$. No optimized design of the channel code is considered.

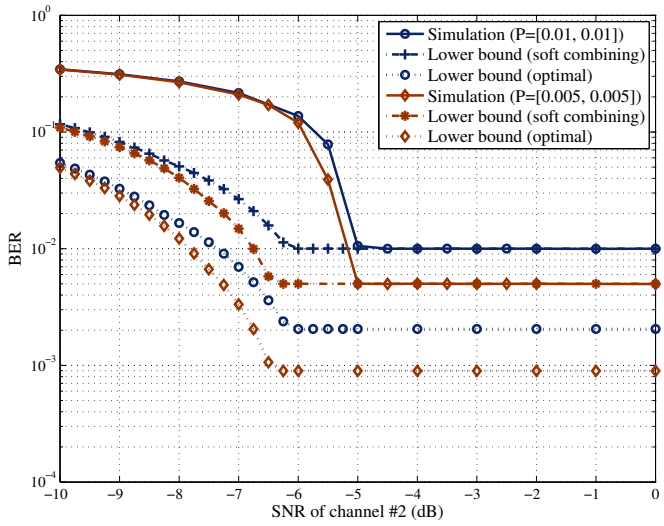


Fig. 36. Asymmetric r_1 and r_2. The coding rates r_1 and r_2 are set at 1 and $\frac{1}{2}$, respectively. The transmit power of two nodes is the same. QPSK is used for node 1 and BPSK for node 2.

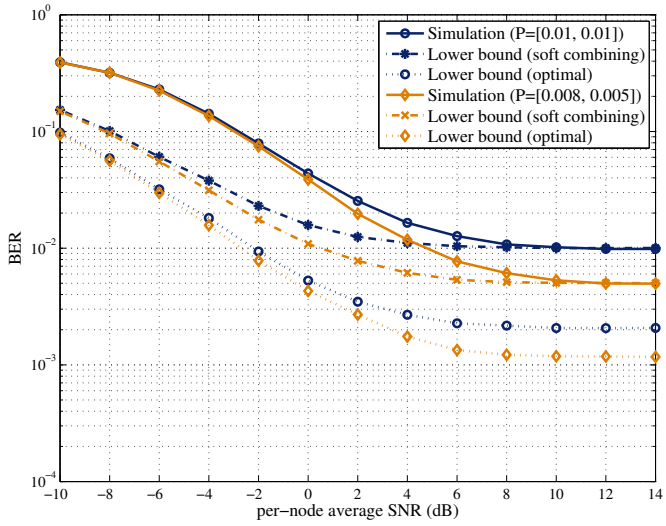


Fig. 37. BER performances over Rayleigh fading channels. Both nodes use BPSK modulation.

In both the symmetric and asymmetric cases, the threshold SNR value at which turbo cliff in the BER obtained by the simulation is around 1.5 dB larger than that observed in the theoretical lower bounds in static AWGN channels. In addition, since the lower bounds on the Hamming distortion plateaus at a certain level even if the power is increased at high SNR regime, increasing the number of nodes is a proper way to improve performance in the practical deployment.

In Fig. 37, the channels between two nodes and the destination experience independent block Rayleigh fading. Therefore, the instantaneous SNRs of two nodes are different while the average SNRs of the two channels are the same. The lower bounds on the Hamming distortion shown in Fig. 37 are calculated as

$$D^{\star}_{\text{fading}} = \int_0^{+\infty} \int_0^{+\infty} D^{\star}(\gamma_1, \gamma_2) \cdot \Pr(\gamma_1) \cdot \Pr(\gamma_2) d\gamma_1 d\gamma_2, \tag{102}$$

where $D^{\star}(\gamma_1, \gamma_2)$ is the result of (101), obtained for static AWGN channels. $\Pr(\gamma_i)$ is the probability density function of the SNR γ_i, which follows the Rayleigh distribution. We use Monte Carlo method to obtain the lower bounds on the average Hamming distortion $D^{\star}_{\text{fading}}$ instead of theoretically calculating (102). In the Rayleigh fading case, the shape of the BER curves and the lower bounds on the Hamming distortion are almost the same. Two points need to be emphasized here. The analytical solution of (102) is difficult to find, because $D^{\star}(\gamma_1, \gamma_2)$ is obtained by solving the formulated convex optimization using *cvx* tool. The other point is that, the outage probability approaches 1 using the definition that the outage event happens when the package cannot *losslessly* recovered. Hence, the definition of outage should be changed in this case. We follow the method of using Slepian-Wolf theorem and separability to calculate the outage probability p_{out} for the situation that bit error floor is reached [57, Section 4.2], where the definition of outage event is

$$\begin{cases} \text{Outage}, & D > \min\{p_1, p_2\} \\ \text{Success}, & \text{otherwise} \end{cases} \tag{103}$$

The detail of deriving p_{out} for two-node case is shown in Appendix 6. We compare the theoretical outage probability p_{out} and the frame error rate (FER)[11], where the results is shown in Fig. 38. The FER performance obtained by practical encoding/decoding algorithm are around $1 \sim 2$ dB in average SNR to the theoretical outage p_{out}.

[11] The frame is error if and only if $D > \min\{p_1, p_2\}$.

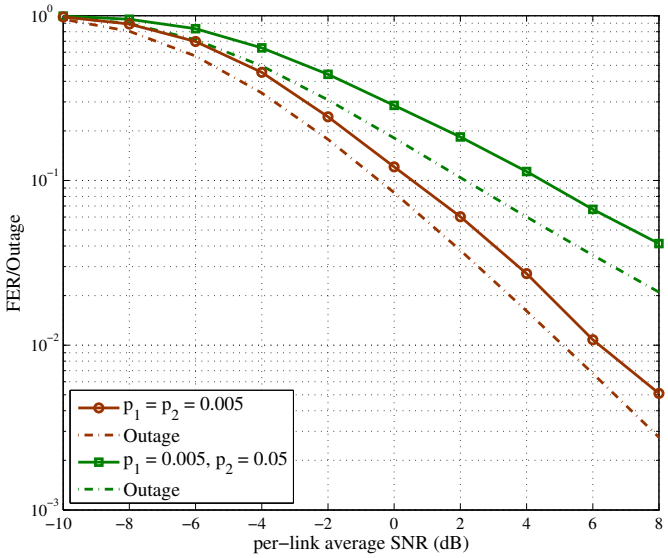


Fig. 38. Comparison between FER and theoretical p_{out}.

4.5 Extension to Multiple Terminals

The outer bound on the rate-distortion region for $L = 2$ terminals is proved and discussed in detail. However, it is worth to show the possibility of extending the proposed solutions to the general binary CEO problem with an arbitrary number of terminals. In this section, the same approach where the rate-distortion region of the binary CEO problem is solved through establishing the relationship with the binary multiterminal source coding is applied to the general binary CEO problem.

4.5.1 Problem Statement

The binary multiterminal source coding which we consider is depicted in Fig. 39. Let the information sequence of source $X_i^n = \{x_i(t)\}_{t=1}^n$, $i = 1,2,\cdots,L$, be binary i.i.d. It should emphasized here that X_i^n is generated from X^n through a BSC with crossover probability p_i in order to make connection with the binary CEO problem. Each encoder independently encodes the data sequences X_i^n with the coding rate R_i as

$$\varphi_i : \mathscr{X}^n \to \mathscr{M}_i = \{1,2,\cdots,2^{n\cdot R_i}\}, \tag{104}$$

where φ_i is the i-th encoder function. The encoder outputs $\varphi_i(X_i^n)$ are transmitted to a common receiver over noiseless channels. The common receiver jointly produces estimates $\hat{X}_i^n$ of the sources based on the received sequences from the agents as $\hat{X}_i^n = \psi[\varphi_1(X_1^n), \cdots, \varphi_L(X_L^n)]$, where ψ is the decoder function.

Let $d_i(x_i(t), \hat{x}_i(t))$ be the average Hamming distortion measure. For given positive numbers $D_i \in [0, 1/2]$, we define the rate-distortion region $\mathscr{R}(D_{\mathscr{L}})$ as

$$\mathscr{R}(D_{\mathscr{L}}) = \{(R_{\mathscr{L}}) : (R_{\mathscr{L}}) \text{ is admissible such that } E\frac{1}{n}\sum_{t=1}^{n} d_i(x_i(t), \hat{X}_i(t)) \leq D_i + \varepsilon\},$$

where $\mathscr{L} = \{1, \cdots, L\}$, $R_{\mathscr{L}} = \{R_i | i \in \mathscr{L}\}$ and $D_{\mathscr{L}} = \{D_i | i \in \mathscr{L}\}$. We provide an outer bound $\mathscr{R}^{\text{o}}(D_{\mathscr{L}})$ for the rate-distortion region $\mathscr{R}(D_{\mathscr{L}})$ with the converse proof in the next section.

4.5.2 Rate-Distortion Region Analysis

Definition 3. $\mathscr{R}^{\text{o}}(D_{\mathscr{L}}) = \bigcap_{\mathscr{S}} \{\mathscr{R}^{\text{o}}_{\mathscr{S}}(D_{\mathscr{S}})\}$, $\forall \mathscr{S} \subseteq \mathscr{L}$ and $\mathscr{S} \neq \emptyset$, with

$$\mathscr{R}^{\text{o}}_{\mathscr{S}}(D_{\mathscr{S}}) = \Bigg\{ (R_{\mathscr{S}}) : \forall R_j, j \in \mathscr{S}^c = \mathscr{L} \setminus \mathscr{S}, \sum_{i \in \mathscr{S}} R_i \geq h(\{p_{\mathscr{S}}, \alpha_{\mathscr{S}^c}\}) - h(\{\alpha_{\mathscr{S}^c}\}) - \sum_{i \in \mathscr{S}} H_2(D_i) \Bigg\}, \tag{105}$$

where

$$p_{\mathscr{S}} = \{p_i | i \in \mathscr{S}\}, \tag{106}$$

$$\alpha_{\mathscr{S}^c} = \{\alpha_j | j \in \mathscr{S}^c\}, \tag{107}$$

$$\alpha_j = p_j * H_2^{-1}(1 - [R_j]^-), \tag{108}$$

set $\mathscr{S}^c$ is the complementary set of $\mathscr{S}$ and $[a]^- = \min\{1, a\}$. Function $h(\cdot)$ is defined in Definition 1.

Theorem 17. $\mathscr{R}(D_{\mathscr{L}}) \subseteq \mathscr{R}^o(D_{\mathscr{L}})$.

Proof

Converse proof of Theorem 17. In order to easily present the proof of the outer bound, we take $L = 3$ as a basic example. Let $U_i = \varphi(X_i^n)$.

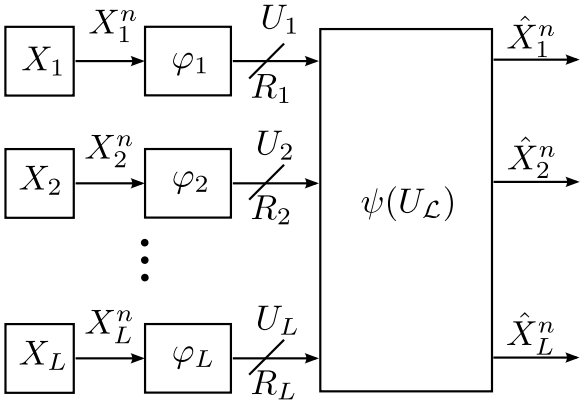


Fig. 39. The binary multiterminal source coding problem for L correlated binary sources.

Case 1. $\mathscr{S} = \{1\}$ and $\mathscr{S}^c = \{2,3\}$. In this case, sources X_2^n and X_3^n operate as helpers for recovering X_1^n. The case belongs to the category of many-help-one problems in the network information theory. A specific two-help-one problem where the primary source, which the decoder wants to reproduce is the XOR version of two helpers was studied by Körner and Marton [6]. Furthermore, the many-help-one problem for correlated Gaussian sources was studied by Oohama [27] and Pandya *et al*. [28], respectively. Assume that D_1 is achieved by a rate triple (R_1, R_2, R_3), then the following equations hold:

$$\begin{aligned} n(R_1+\varepsilon) &\geq H(U_1) \\ &\geq H(U_1|U_2,U_3) && (109)\\ &= I(X_1^n;U_1|U_2,U_3) && (110)\\ &= I(X_1^n;U_1,U_2,U_3) - I(X_1^n;U_2,U_3) && (111)\\ &\geq I(X_1^n;\hat{X}_1^n) - I(X_1^n;U_2,U_3) && (112) \end{aligned}$$

where the steps are justified since
(109) conditioning reduces the entropy,
(110) U_1 is a function of X_1^n,
(111) the chain rule of mutual information,
(112) data processing inequality,

$$n(R_2+\varepsilon) \geq H(U_2) = I(X_2^n;U_2), \tag{113}$$

$$n(R_3+\varepsilon) \geq H(U_3) = I(X_3^n;U_3). \tag{114}$$

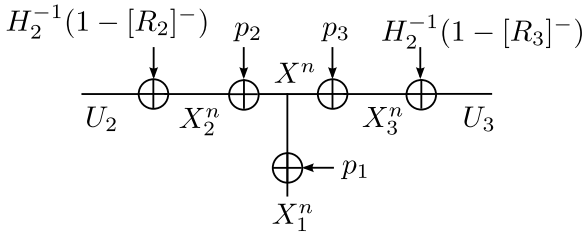


Fig. 40. The test BSC model for proving the outer bound.

Now, we need to find the lower bound on the term $I(X_1^n;\hat{X}_1^n)$ and the upper bound on $I(X_1^n;U_2,U_3)$. As same as in $L=2$ case, the lower bound on $I(X_1^n;\hat{X}_1^n)$ is obtained by applying Fano's inequality

$$I(X_1^n;\hat{X}_1^n) \geq n - nH_2(D_1). \tag{115}$$

The upper bound on $I(X_1^n;U_2,U_3)$ is derived based on the test BSC model shown in Fig. 40 and inequalities (113), (114), as[12]

$$\begin{aligned} I(X_1^n;U_2,U_3) &= nH(X_1)+H(U_2,U_3)-H(X_1^n,U_2,U_3) \\ &\leq n+nh(\{\alpha_2,\alpha_3\})-nh(\{p_1,\alpha_2,\alpha_3\}), \end{aligned} \tag{116}$$

where (116) is obtained based on the fact that X_1^n, U_2 and U_3 are the outputs from a BSC or a cascade BSC channels when X^n is the input. By substituting (115) and (116) into (112) and letting ε go to 0, we conclude that

$$R_1 \geq h(\{p_1,\alpha_2,\alpha_3\})-h(\{\alpha_2,\alpha_3\})-H_2(D_1). \tag{117}$$

For the cases $\mathscr{S}=\{2\}$ and $\mathscr{S}=\{3\}$, the bounds can be obtained in the same way.

Case 2. $\mathscr{S}=\{1,2\}$ and $\mathscr{S}^c=\{3\}$. The source X_3^n acts as the helper to recover X_1^n and X_2^n. Gastpar derived the inner and outer bounds for the rate-distortion region of independently compressing two or more correlated sources with side information available at the decoder [13]. In this case, except that (114) holds, the following inequalities

[12]Inspired by the MGL, we establish the test BSC model to bound the mutual information. However, this bound may exist a gap to the global optimal bound. The validation of the global optimality is left as a future study. Indeed, the bound derived from the current setup can still serve as a useful reference in the power allocation and scheduling of WSNs. Some further discussions on this issue are provided in subsection 4.5.4.

also hold.

$$
\begin{aligned}
n(R_1 + R_2 + \varepsilon) &\geq H(U_1) + H(U_2) \\
&\geq H(U_1, U_2 | U_3) \\
&= I(X_1^n, X_2^n; U_1, U_2 | U_3) \\
&= I(X_1^n, X_2^n; U_1, U_2, U_3) - I(X_1^n, X_2^n; U_3) \\
&= I(X_1^n; U_1, U_2, U_3) + I(X_2^n; U_1, U_2, U_3 | X_1^n) - I(X_1^n, X_2^n; U_3) \\
&\geq I(X_1^n; \hat{X}_1^n) + I(X_2^n; U_1, U_2, U_3, X_1^n) \\
&- I(X_1^n; X_2^n) - I(X_1^n; U_3) - I(X_2^n; U_3 | X_1^n) \qquad (118) \\
&\geq I(X_1^n; \hat{X}_1^n) + I(X_2^n; \hat{X}_2^n) \\
&- I(X_1^n; X_2^n) - I(X_1^n; U_3) - I(X_2^n; U_3, X_1^n) + I(X_1^n; X_2^n) \\
&\geq H(X_1^n, X_2^n, U_3) - H(U_3) - n[H_2(D_1) + H_2(D_2)] \qquad (119) \\
&\geq n[h(\{p_1, p_2, \alpha_3\}) - h(\{\alpha_3\}) - H_2(D_1) - H_2(D_2)] \qquad (120)
\end{aligned}
$$

where (118) holds because the chain rule of mutual information and data processing inequality, (119) follows from several steps of elementary calculation, and (120) is obtained based on the same test BSC model which is shown in Fig. 40.

By letting ε go to 0, it is concluded that $R_1 + R_2 \geq h(\{p_1, p_2, \alpha_3\}) - h(\{\alpha_3\}) - H_2(D_1) - H_2(D_2)$. The other two similar cases with $\mathscr{S} = \{1,3\}$ and $\mathscr{S} = \{2,3\}$ can be followed the same derivation which is shown above.

Case 3. $\mathscr{S} = \{1,2,3\}$ and $\mathscr{S}^c = \emptyset$. Assume that a rate triple (R_1, R_2, R_3) achieves the required distortions D_1, D_2 and D_3, we have the following inequalities.

$$
\begin{aligned}
n(R_1 + R_2 + R_3 + \varepsilon) &\geq H(U_1) + H(U_2) + H(U_3) \\
&\geq H(U_1, U_2, U_3) \\
&= I(X_1^n, X_2^n, X_3^n; U_1, U_2, U_3) \\
&= I(X_1^n; U_1, U_2, U_3) + I(X_2^n, X_3^n; U_1, U_2, U_3 | X_1^n) \\
&\geq I(X_1^n; \hat{X}_1^n) + I(X_2^n, X_3^n; U_1, U_2, U_3, X_1^n) - I(X_2^n, X_3^n; X_1^n) \\
&\geq I(X_1^n; \hat{X}_1^n) + I(X_2^n; U_1, U_2, U_3, X_1^n) + I(X_3^n; U_1, U_2, U_3, X_1^n | X_2^n) - I(X_2^n, X_3^n; X_1^n) \\
&\geq I(X_1^n; \hat{X}_1^n) + I(X_2^n; \hat{X}_2^n) + I(X_3^n; \hat{X}_3^n) - I(X_3^n; X_2^n) - I(X_2^n, X_3^n; X_1^n) \\
&= nh(\{p_1, p_2, p_3\}) - n \sum_{i=1}^{3} H_2(D_i). \qquad (121)
\end{aligned}
$$

The derivation of above inequalities can be straightforwardly extended to the general case having an arbitrary number of sources. In summary, we conclude that

$$\sum_{i\in\mathscr{S}} R_i \geq h(\{p_{\mathscr{S}}, \alpha_{\mathscr{S}^c}\}) - h(\{\alpha_{\mathscr{S}^c}\}) - \sum_{i\in\mathscr{S}} H_2(D_i). \tag{122}$$

Hence, we can conclude that $\mathscr{R}(D_{\mathscr{L}}) \subseteq \mathscr{R}^{\mathrm{o}}(D_{\mathscr{L}})$. □

Remark 6. The outer bound $\mathscr{R}^{\mathrm{o}}(D_{\mathscr{L}})$ for the rate-distortion region is a convex hull of a set of rate tuples.

Remark 7. If all the distortion levels $D_{\mathscr{L}}$ approach to 0, the outer bound $\mathscr{R}^{\mathrm{o}}(D_{\mathscr{L}})$ coincides with the Slepian-Wolf theorem with multiple correlated sources [3, 78].

Remark 8. Consider the case two agents observe the same source X, i.e., $L = 2$. In this case, since $h(\cdot)$ has a very simple form, we can obtain the following rate-distortion inequalities by substituting p_i and D_i into the outer bound expression

$$\begin{cases} R_1(D_1) \geq H_2(p_1 * \alpha_2) - H_2(D_1), \\ R_2(D_2) \geq H_2(p_2 * \alpha_1) - H_2(D_2), \\ \sum\limits_{i=1}^{2} R_i(D_i) \geq 1 + H_2(p_1 * p_2) - \sum\limits_{i=1}^{2} H_2(D_i), \end{cases} \tag{123}$$

which is consistent with the results shown in $L = 2$ case.

4.5.3 Sum Rate versus Distortion

The distortion level D is examined with respect to the sum rate R, giving fixed values of L and the observation error probabilities p_i. The results of D versus R are shown in Fig. 41 with different L. The distortion D is the result of

$$D = \mathrm{PB}(p_{\mathscr{L}} * D^{\star}_{\mathscr{L}}), \tag{124}$$

where $p_{\mathscr{L}} * D^{\star}_{\mathscr{L}} = \{p_1 * D^{\star}_1, \cdots, p_L * D^{\star}_L\}$ and $D^{\star}_{\mathscr{L}} = \{D^{\star}_i | i \in \mathscr{L}\}$ is given by

$$\begin{aligned} D^{\star}_i \quad &= \arg\min \quad \|[D_1, \cdots, D_L]\|_2 \\ &\text{s.t.} \quad \begin{cases} \sum\limits_{i\in\mathscr{S}} H_2(D_i) \geq h(\{p_{\mathscr{S}}, \alpha_{\mathscr{S}^c}\}) - h(\{\alpha_{\mathscr{S}^c}\}) - \sum\limits_{i\in\mathscr{S}} \frac{R}{L}, \\ 0 \leq D_i \leq 0.5, \qquad i \in \mathscr{L} \end{cases} \end{aligned} \tag{125}$$

and $\mathrm{PB}(\cdot)$ calculated the error probability based on the Poisson binomial process using (59).

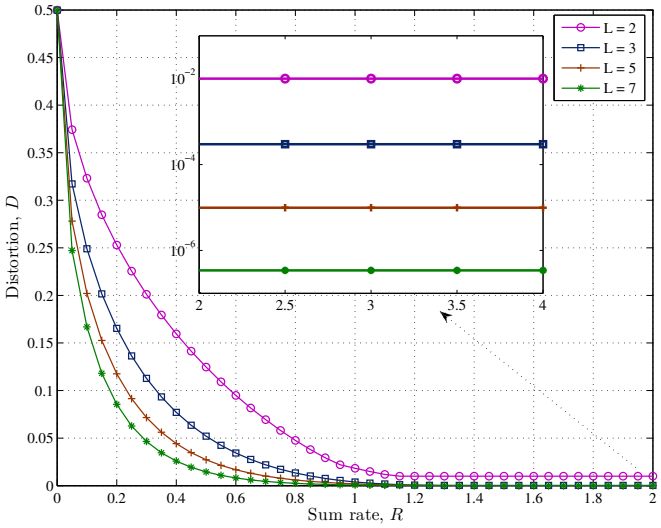


Fig. 41. Sum rate versus distortion. The rate is equally allocated to each link. The observation error probabilities are set to 0.01**.**

It should be emphasized here that the sum rate R is equally allocated to each link without considering any optimal rate allocation scheme. It can be seen from the figure that, D becomes small, if the number of links L increased and/or the sum rate R is large. However, D converges to a certain level (not equal to 0) even we increase R which is very clear in the enlarged view. The certain level is given by $\mathrm{PB}(p_{\mathscr{L}})$ by assuming the distortions D_i asymptotically approach 0. Furthermore, this level also decreased when the number of agents increases.

4.5.4 Brief Discussions of using test BSC

In the proof of the outer bound, there is an important step of bounding the mutual information term $I(X_1^n;U_2,\cdots,U_L)$ using the test BSCs. In other words, the cardinality bound on $\mathscr{U}_i$ is assumed to be 2. In [97], Soumya gave a proof of reducing the cardinality bound of auxiliary RV U_i in multiterminal source coding problem. It is found that the cardinality bound can be reduced as $|\mathscr{U}_i| \leq |\mathscr{X}_i|$. Applying the result in our specified binary Hamming case, it is reasonable to set the cardinality bound $|\mathscr{U}_i|$ to 2.

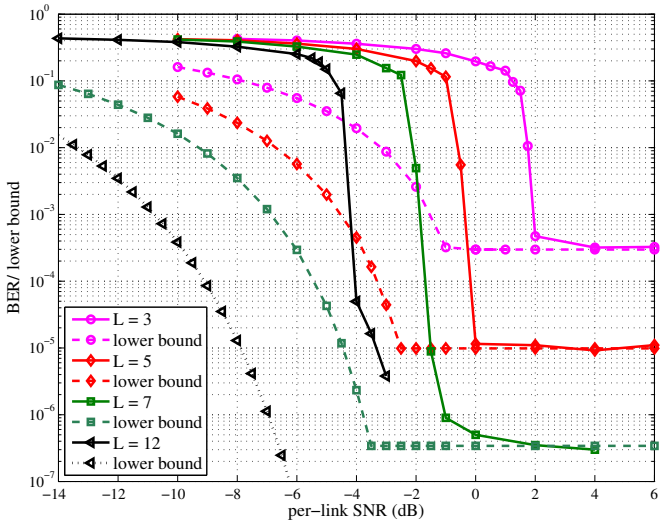


Fig. 42. Simulation results verify the lower bounds on Hamming distortion using 16QAM and identical p_i.

Besides this, the further discussion is needed for whether it is optimal to use BSC to lower bound the conditional entropy terms, such as $H(X_1^n|U_2,U_3)$. Inspired by the proof of MGL, we use the BSC as the test channel in the derivation. However, the extension of MGL to this general setup still needs some efforts. Hence, the outer bound on the rate-distortion region is only an approximation.

4.5.5 Numerical Results

A series of simulations are performed to verify the Hamming distortion lower bound that obtained by solving the following convex optimization problem

$$\begin{aligned} \min \quad & \|[D_1,\cdots,D_L]\|_2 \\ \text{s.t.} \quad & \begin{cases} -\sum\limits_{i\in\mathscr{S}} H_2(D_i) \le \sum\limits_{i\in\mathscr{S}} \frac{C(\gamma_i)}{r_i} - h(\{p_{\mathscr{S}},\alpha_{\mathscr{S}^c}\}) + h(\{\alpha_{\mathscr{S}^c}\}), \\ 0 \le D_i \le 0.5, \qquad i\in\mathscr{L} \end{cases} \end{aligned} \tag{126}$$

and mapping the minimal values to D using Poisson binomial process.

To obtain the BER performance, the practical encoding/decoding algorithm is used as the same as in Fig. 16 and Fig. 17 (Section 3.4 of Chapter 3), where the modulation

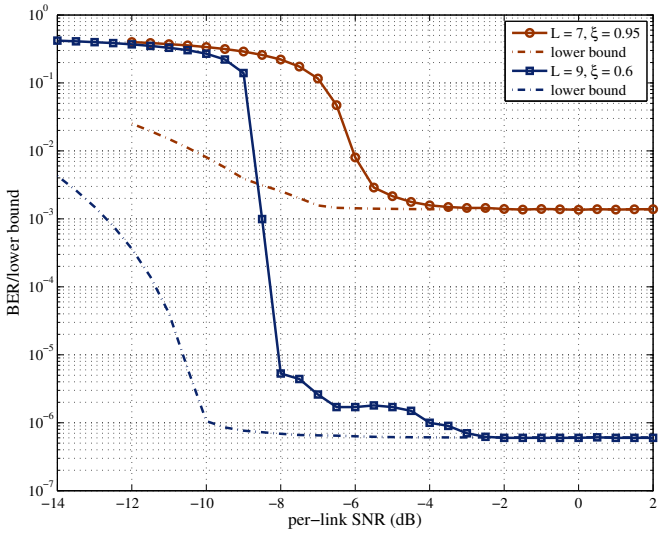
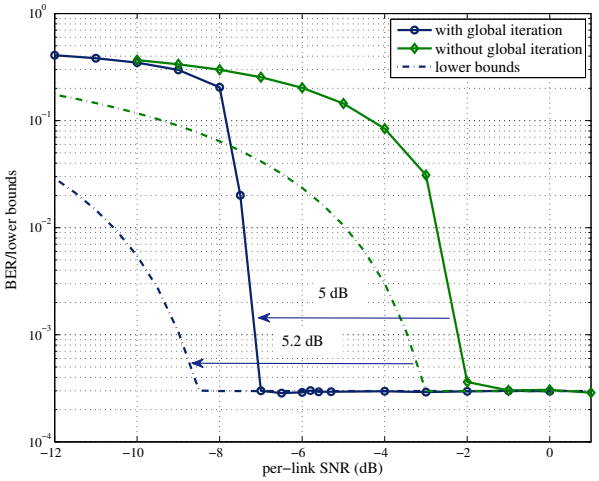


Fig. 43. Simulation results verify the lower bounds on Hamming distortion using BPSK and logarithmic distributed p_i.

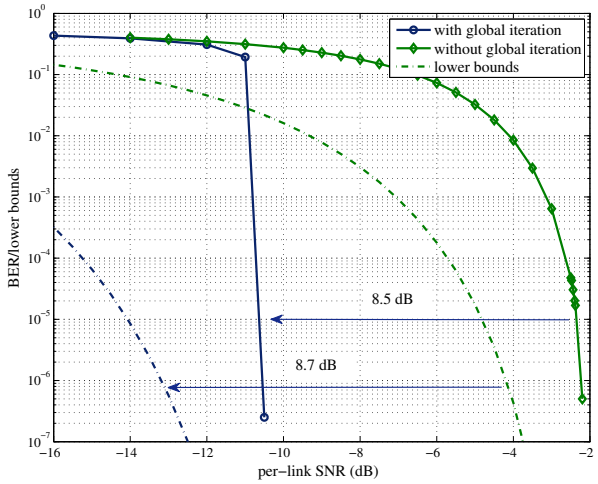
scheme is assumed to be 16-QAM with modified set partition (MSP) mapping [98] and BPSK, respectively.

Figure 42 shows the simulation results and their corresponding Hamming distortion lower bounds for identical p_i. As we can see from the figure that, the Hamming distortion lower bounds and the simulation results have very similar tendency. However, if L goes large, the difference between the simulation results and the theoretical bounds also becomes significant. The reason is that the theoretical bounds are obtained by assuming the capacity-achieving code is used at each sensor node. To further analyze the impact of p_i variation, Fig. 43 shows the BER performance using the obtained p_i from logarithmic distribution. Their Hamming distortion lower bounds also presents as references. Note that in these cases, the Hamming distortion bound on D is obtained using soft combining (Appendix 2) after the minimal distortions D_i of each link are available. From the simulations results, it concludes that the impact of p_i variation to the bound analysis is not significant.

Furthermore, the superiority of performing global iteration in decoding process is proved through the theoretical analysis. It is found that from Fig. 44, the performance gain in AWGN channels is around 5 dB for $L = 4$ and 8.5 dB for $L = 12$, respectively.



(a) $L = 4$, $p_i = 0.01$.



(b) $L = 12$, $p_i = 0.01$.

Fig. 44. Comparison on theoretical Hamming distortion lower bounds by assuming whether correlation is utilized through *global iteration*.

Simulation results using the proposed encoding/decoding algorithm also confirm the gain. The Hamming distortion lower bound for the case global iteration is not performed is given by

$$D^{\text{noGI}} = \text{PB}(\{p_1 * H_2^{-1}(1-[C(\gamma_1)]^-), \cdots, p_L * H_2^{-1}(1-[C(\gamma_L)]^-)\}). \tag{127}$$

In other words, D^{noGI} is the result of solving the following convex optimization problem.

$$D^{\text{noGI}} = \text{PB}(p_{\mathscr{L}} * D_{\mathscr{L}}^{noGI\star}), \tag{128}$$

where $p_{\mathscr{L}} * D_{\mathscr{L}}^{noGI\star} = \{p_1 * D_1^{noGI\star}, \cdots, p_L * D_L^{noGI\star}\}$ and $D_{\mathscr{L}}^{noGI\star} = \{D_i^{noGI\star} | i \in \mathscr{L}\}$ is given by

$$\begin{aligned} \min \quad & \|[D_1, \cdots, D_L]\|_2 \\ \text{s.t.} \quad & \begin{cases} -\sum\limits_{i\in\mathscr{S}} H_2(D_i) \le \sum\limits_{i\in\mathscr{S}} \frac{C(\gamma_i)}{r_i} - |\mathscr{S}|, \\ 0 \le D_i \le 0.5, \qquad i \in \mathscr{L} \end{cases} \end{aligned} \tag{129}$$

Comparing the constraints of (126) and (129), it is obviously found that $|\mathscr{S}| \ge h(\{p_{\mathscr{S}}, \alpha_{\mathscr{S}^c}\}) - h(\{\alpha_{\mathscr{S}^c}\})$ with equality holding if and only if $p_i = 0.5$. Hence, the Hamming distortion lower bound D^{noGI} is greater than D.

4.6 Summary

In this chapter, we examined theoretically the lower bound on the Hamming distortion for the binary information sensing network modelled by the binary CEO problem, where several independent terminals forward the erroneous versions of a common binary source to the destination over static AWGN and block Rayleigh fading channels.

We first considered a simple case that the number of terminals is 2. The binary CEO problem was first formulated as the binary multiterminal source coding problem, which is the core part of the binary CEO problem. The outer bound on the rate-distortion region for the binary multiterminal source coding problem was then derived based on the converse proof of the bound. The relationship between the binary CEO problem and the binary multiterminal source coding problem in terms of the distortion function has been established. According to the lossy source-channel separation theorem, the lower bound on the Hamming distortion was formulated by minimizing the distortion function subject to the inequalities between the derived outer bound and the channel capacities. The problem of obtaining the lower bound on the Hamming distortion was solved in the

framework of convex optimization, and the results of Hamming distortion lower bounds only apply to schemes which use sequential decoding. Through a series of simulations, it has been shown that the BER curves obtained with a practical encoding/decoding algorithm is consistent with the result of the theoretical lower bounds on the Hamming distortion.

We further extended discussions to the binary CEO problem having arbitrary number of terminals. An approximated outer bound was derived through the converse coding proof. The outer bound on the rate-distortion region was used to obtain the theoretical lower bound on the Hamming distortion for the general binary information sensing network as the case of two terminals. Finally, we simply discussed the superiority of our proposed decoding algorithm from the rate-distortion perspective.

5 Power Allocation of Binary Information Sensing Networks

In this chapter, a power allocation (PA) scheme is proposed to minimize the distortion with a fixed total transmission power in the context of binary information sensing network.

In sensor networks, each sensor node is equipped with a battery and the power is usually scarce owing to the limit battery size. Hence, the power allocation and scheduling in sensor networks become extremely important. In [99, 100], an optimization framework for joint source coding, routing and resource allocation was presented in sensor networks. The distortion and power were weighted by two vectors in the optimization problem to achieve the goal of balancing the tradeoff between them. The optimization problem was solved efficiently in the dual domain. Optimal power allocation for Gaussian sensor network with distortion constraints was considered in [101], where both TDMA and non-orthogonal multiple access (NOMA) schemes are assumed in transmission phases of sensors.

We consider the power allocation in AWGN channel for binary information sensing network from rate-distortion perspective in order to achieve optimum distortion under total power constraints. The problem is formulated in the convex optimization framework and is solved using Karush-Kuhn-Tucker (KKT) conditions. Computer simulations are performed to show the advantage of our proposed PA scheme.

This chapter is organized as follows. Section 5.1 describes the problem, followed by the proposed PA algorithms in Section 5.2. The numerical results are presented in Section 5.3. We conclude this chapter in Section 5.4.

5.1 Problem Statement

In the binary sensing network which is illustrated in Fig. 45, a group of sensors is distributed in a specific area to observe a binary sensing object. There exists a leader, which is often called cluster head, in this group to collect the position of each sensor. The leader is in charge of calculating the observation accuracy of each sensor based on their position and determining the power ratio of each sensor. The leader then broad-

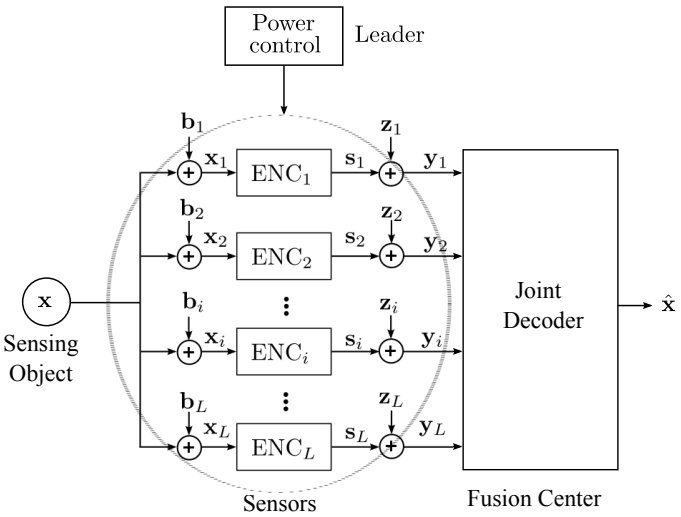


Fig. 45. System model of binary information sensing network with power allocation.

casts the power ratio to each sensor. Meanwhile, we simply suppose that the power ratio is correctly received by each sensor.

After this initialization stage, each sensor encodes its observation $\mathbf{x}_i$ using the concatenated convolutional codes. The coded sequences are then modulated by BPSK and transmitted to the fusion center over an AWGN channel with the corresponding power ratio. For simplicity, the channel noise at the receiver side is normalized to 1. The problem is to determine the optimal power ratio at the leader node with fixed total power E_T to achieve minimal BER.

5.2 Proposed Power Allocation Schemes

As stated in Chapter 4, the binary sensing network is modeled by the binary CEO problem, and we formulated the minimization problem of obtaining the lower bound on Hamming distortion. Similarly, the problem of optimal power allocation is formulated

as follows.

$$\min_{\beta_i} \quad \mathrm{PB}(p_1 * d_1, \cdots, p_L * d_L) \tag{130}$$

s.t.

$$-\sum_{i\in\mathscr{S}} H_2(d_i) \leq \sum_{i\in\mathscr{S}} \frac{C(\beta_i E_T)}{r_i} - h(\{p_{\mathscr{S}}, \alpha_{\mathscr{S}^c}\}) + h(\{\alpha_{\mathscr{S}^c}\}), \tag{131}$$

$$\sum_{i=1}^{L} \beta_i = 1, \tag{132}$$

$$\beta_i \geq 0. \tag{133}$$

The formulated problem is non-convex and not easy to solve. The reason is that the analytical form of $h(\cdot)$ is complicated and the parameters $\alpha_{\mathscr{S}^c}$ depend on the allocated power. Moreover, the objective function $\mathrm{PB}(\cdot)$ is not convex. We thus propose a brute force search-based heuristic method to obtain the optimal power ratio $\beta_i^\star$, which is summarized in Algorithm 1.

Algorithm 1: Heuristic method for power allocation

Input: p_i, E_T
Output: $\beta_i^\star$ such that d is minimized
Initialization: ascending sort p_i, $d_- = \infty$
for $\beta_1 = 0$ **to** 1 **do**
  for $\beta_j = 0$ **to** $1 - \sum_{k=1}^{j-1} \beta_k$, $(j = 2, \cdots, L-1)$ **do**
    $\beta_L = 1 - \sum_{k=1}^{L-1} \beta_k$;
    if $\beta_k > \beta_l$, $(\forall k > l)$ **then**
      Continue;
    end
    Obtain minimal d_i for each link based on the outer bound;
    Caculate d using $\mathrm{PB}(p_1 * d_1, \cdots, p_L * d_L)$;
    if $d < d_-$ **then**
      $\beta_i^\star = \beta_i$, $i = 1, \cdots, L$;
      $d_- = d$;
    end
  end
end

However, the computational complexity of Algorithm 1 is very high, since the number of combinations on β_i is very large. Even we add a condition that $\beta_i < \beta_j, \forall i < j$

to reduce the number of combinations, the heuristic method is not tractable in practice, particularly for large number of sensors, e.g., $L = 7$.

Hence, we need to formulate this problem instead of using our derived outer bound. To simplify this problem, we use the summation of $p_i * d_i$, which is the mean value of the Poisson binomial distribution as the objective function. Furthermore, we suppose that the observations are independent of each other. As a consequence, the problem (130) is formulated to

$$\min_{\beta_i} \quad \sum_{i=1}^{L}(1-2p_i)d_i + p_i \tag{134}$$

s.t.

$$1 - H_2(d_i) \leq C(\beta_i E_T), \tag{135}$$

$$\sum_{i=1}^{L}\beta_i = 1, \tag{136}$$

$$\beta_i \quad \geq 0. \tag{137}$$

It should be emphasized here that the solution of this problem only provides a suboptimal power ratio of the main problem, since the effectiveness of the correlation is not taken into account. The complexity of the problem, however, is significantly reduced. In the minimization problem (134), it is easily found that d_i works as a dummy variable. Due to the fact that d_i is proportional to $-C(\beta_i E_T)$, we reformulate the problem (134) into the following maximization problem

$$\max_{\beta_i} \quad \sum_{i=1}^{L}(1-2p_i)C(\beta_i E_T) \tag{138}$$

s.t.

$$\sum_{i=1}^{L}\beta_i = 1, \tag{139}$$

$$\beta_i \geq 0, \tag{140}$$

to avoid analytically deriving the inverse function of $H_2^{-1}(\cdot)$. Obviously, the problem (138) is convex [102], since the capacity function $C(\cdot)$ is concave and the constraints are linear. The KKT conditions of (138) are summarized in Appendix 7. From the KKT conditions, we have the analytical solutions for optimum power ratio $\beta_i^\star$ in (138), as

$$\beta_i^\star = \begin{cases} \frac{1-2p_i}{\mu^\star} - \frac{1}{E_T}, & \mu^\star < (1-2p_i)E_T, \\ 0, & \mu^\star \geq (1-2p_i)E_T, \end{cases} \tag{141}$$

Table 9. Optimal power ratio of each sensor obtained by proposed PA scheme.

p_i	E_T (dB)	$\beta^\star$
$\{0.005, 0.05, 0.15\}$	-5	$\{0.6745, 0.3255, 0\}$
	0	$\{0.529, 0.39, 0.081\}$
	10	$\{0.397, 0.3517, 0.2513\}$
$\{0.005, 0.01, 0.02, 0.1, 0.3\}$	-5	$\{0.3811, 0.3453, 0.2736, 0, 0\}$
	0	$\{0.3271, 0.3137, 0.2869, 0.0723, 0\}$
	10	$\{0.2595, 0.256, 0.2487, 0.1905, 0.0453\}$
$\{0.002, 0.008, 0.01, 0.05, 0.07, 0.1, 0.15\}$	-5	$\{0.3596, 0.3174, 0.3033, 0.0197, 0, 0\}$
	0	$\{0.2631, 0.2479, 0.2428, 0.1414, 0.0904, 0.0144, 0\}$
	10	$\{0.1722, 0.169, 0.1679, 0.146, 0.135, 0.1186, 0.0913\}$

with

$$\sum_{i=1}^{L} \max\{0, \frac{1-2p_i}{\mu^\star} - \frac{1}{E_T}\} = 1, \tag{142}$$

where μ is the Lagrange multiplier for the equality constraint (140). We apply the proposed optimal PA scheme to binary information sensing network and make comparison with the uniform PA method.

5.3 Numerical Results

By using (141), the optimal power ratios are obtained for some fixed E_T and sets of observation error probabilities, the results of which are showed in Table 9. Based on these results, it is found that $\beta_i^\star$ is equal to 0 in some cases, e.g., $L = 3$ with $E_T = -5$ dB. Hence, our proposed scheme can be perceived as a scheduling method in binary information sensor network. In this case, the leader (cluster header) sends a control message to those sensors of which the power ratio is 0 to force them into sleeping mode. After a constant time interval, those sensors are activated and report their positions to the cluster header. Our proposed joint decoding algorithm that utilizes the correlation knowledge is adopted in order to obtain the BER performance.

Figures 46–48 demonstrate the BER performance versus total SNR using our proposed power allocation scheme and uniform power allocation. For $L = 3$ and $L = 5$, the BER performance using the power ratio obtained from the heuristic method are also presented. From these results, it exhibits that our proposed PA scheme gains around $1.5 \sim 2$ dB in terms of the total power compared to the uniform PA case for the specific

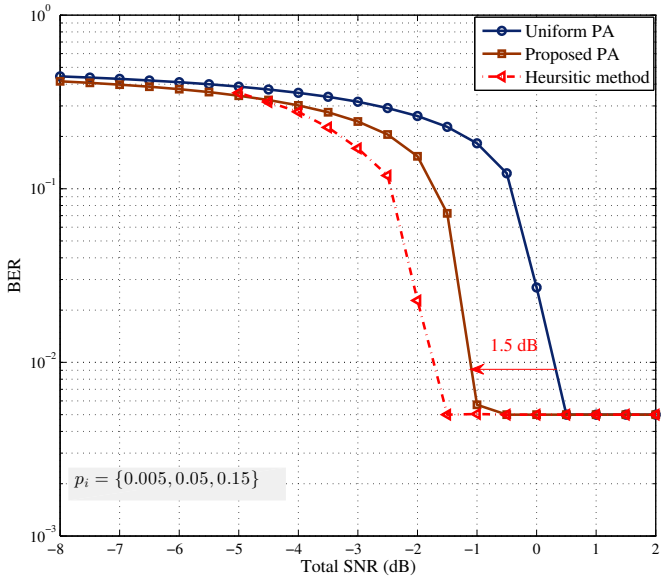


Fig. 46. BER performance of using the proposed PA, uniform PA and heuristic method for $L = 3$.

observation error probabilities. Compared to the heuristic method, the loss of our proposed PA is very small, 0.5 dB for $L = 3$, while almost no loss for $L = 5$. The loss is relatively large if the BER performance is dominated by a small number of sensors, e.g., the sensor with $p = 0.005$ in 3-node case dominates the performance. Based on these comparisons, our proposed power allocation scheme can be seen as a good approximation of the main problem. Moreover, the benefit of our proposed power allocation is the low complexity. In heuristic method, 39 and 106 combinations, respectively, were searched for $L = 3$ and $L = 5$ at each SNR point. As a consequence, the proposed PA can be easily applied in practical situations for the binary information sensing compared to the heuristic method.

At the high SNR regime, the BER performance obtained by all the schemes converges together. The reason is twofold: 1) at high SNR regime, the proposed power allocation scheme approaches to the uniform allocation, while in heuristic method, power is uniformly allocated to the dominated sensors; 2) the BER performance is only determined by the error probability p_i when the power is large enough.

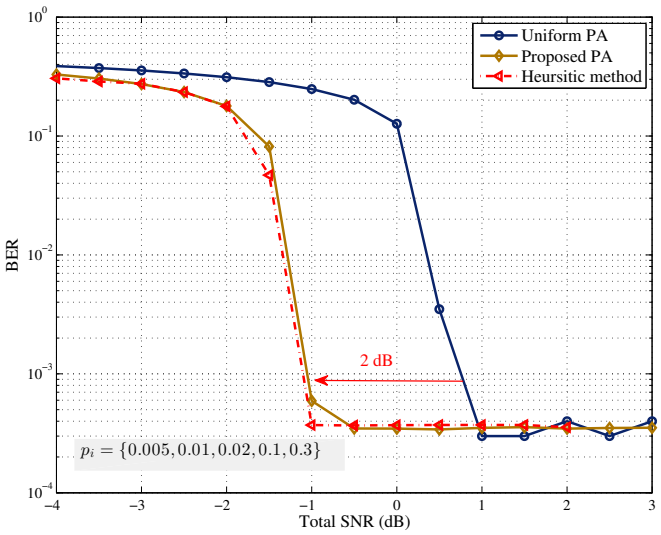


Fig. 47. BER performance of using the proposed PA, uniform PA and heuristic method for $L = 5$.

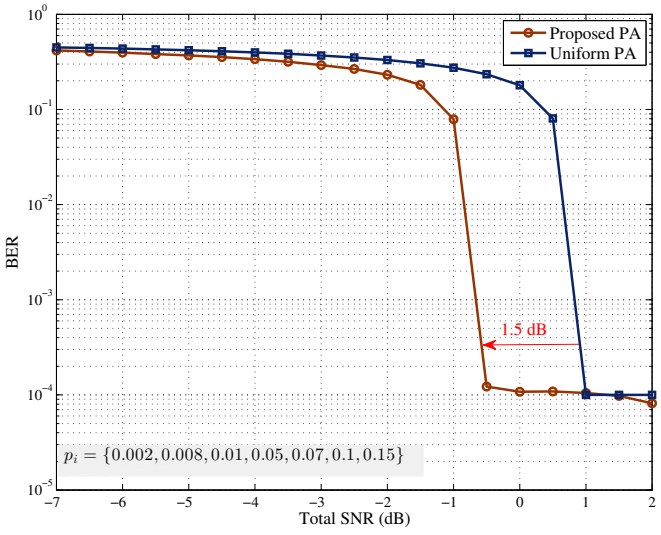


Fig. 48. BER performance of using the proposed PA and uniform PA for $L = 7$.

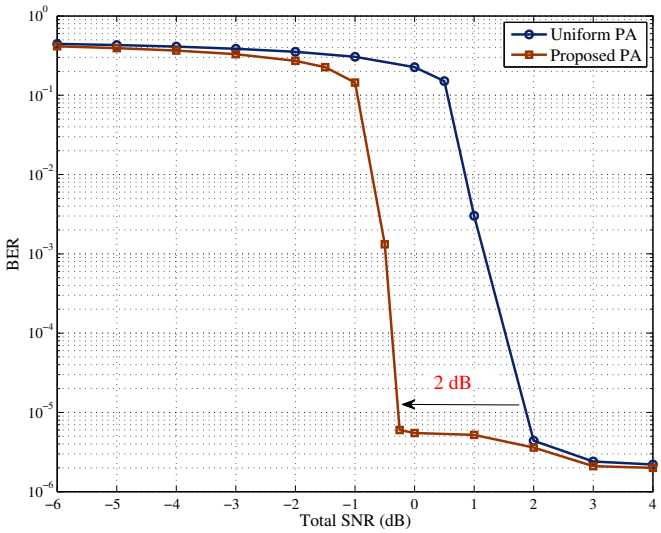


Fig. 49. BER performance of using the proposed PA and uniform PA for $L = 9$. p_i follows logarithmic distribution with $\xi = 0.6$.

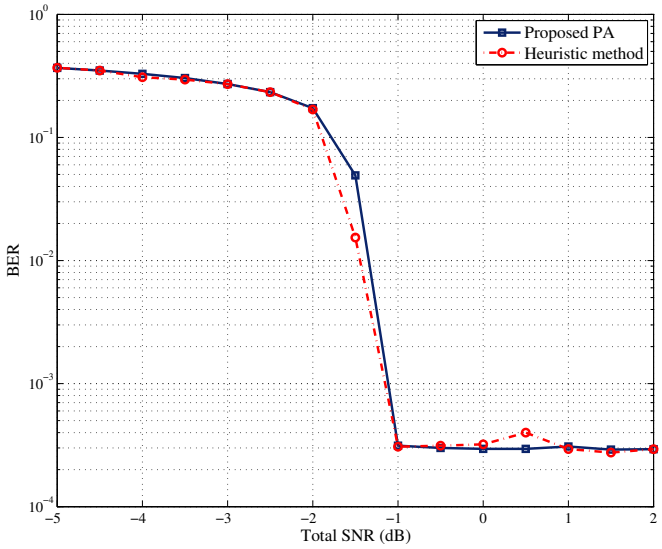


Fig. 50. BER performance of using proposed PA and heuristic method.

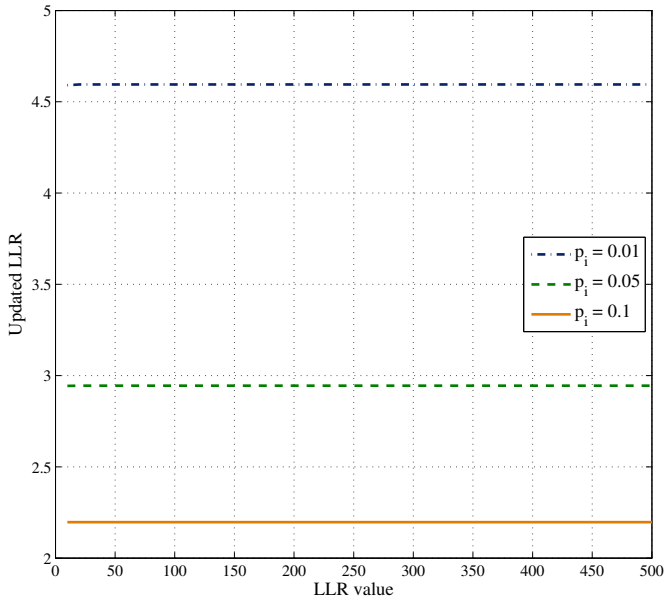


Fig. 51. Updated LLR values through LLR updating function f_c.

The observation error probability p_i is set at predefined values in above analyses, however, Fig. 49 illustrates the BER performance of the proposed power allocation scheme when p_i follows logarithmic distribution. It is found that the proposed power allocation scheme can achieve 2 dB gain in total power compared to the uniform power allocation.

We also consider the power allocation for the case p_i are equally distributed, where the BER performances are illustrated in Fig. 50 . The proposed convex optimization method (138) results in same power ratio for each link since the capacity of each link is weighted by the same value. However, the heuristic method obtains different power ratio at some SNR points, for example, at $-4, -3.5$ and -1.5 dB, the optimal power ratio $\beta_i^\star$ are $\{0.5, 0.45, 0.05\}$, $\{0.45, 0.4, 0.15\}$ and $\{0.7, 0.15, 0.15\}$, respectively. However, the difference in terms of the BER performance between the proposed power allocation and the heuristic methods is not significant. The reason is that, as shown in Fig. 51, the value of the updated LLR is significantly affected by the observation error probability p_i. However, the input LLR that depends on the channel SNR does not change the up-

dated LLR much. Hence, if p_i is the same, we can simply allocate equal power to each sensor.

In general, our proposed scheme works appropriately, however, the proposed PA is centralized, in which the cluster head needs to collect the information from the sensors. The complexity of the problem may increase if the number of sensors is large in a group. In the future, we need to find a solution to make the PA distributed, and/or to seek for a good partition method to group sensors.

5.4 Summary

We proposed an optimal power allocation scheme for binary information sensing network from the viewpoint of rate-distortion. For the original formulated problem, we proposed a brute-force search-based heuristic method to get the optimal power ratio. However, it is necessary to check a great number of combinations for the power ratio in the heuristic method. In order to reduce the complexity, we reformulated the optimization problem to maximize the sum of channel capacities weighted by the observation accuracy $(1-2p_i)$. Through computer simulations, it showed that our proposed power allocation scheme outperforms the uniform allocation method and is easy to deploy in practice compared to the heuristic method. Even though our proposed scheme is not distributed, it still be applicable to small-scale sensor network or parallel relaying network. In the future, we need to find a distributed power allocation to reduce the overhead when collecting the information from each sensor node by the cluster head.

6 Conclusion and Outlook

6.1 Conclusion

In this work, the theoretical analysis of the binary CEO problem was carried out from the view of rate-distortion behavior. We *solved the rate-distortion bound for the binary CEO problem with two-terminal* by establishing a direct outer bound and a two-step Berger-Tung inner bound with optimal decision on the rate-distortion region. It was found that the inner bound and the outer bound exactly match with each other. Furthermore, for the binary CEO problem with multiple terminals, we established a *loose outer bound* on the rate-distortion region using the specified two-step setup. Meanwhile, the practical encoding/decoding algorithm was proposed for the binary information sensing network to utilize the correlation knowledge of the observations. A power allocation scheme based on the derived loose outer bound was proposed to achieve better BER performance under total power constraints. In summary, the significance of this work is listed below.

– The two-terminal binary CEO problem was solved.
– The proposed joint decoding algorithm has wide applications, such as parallel relaying networks.
– Proposed power allocation scheme for binary information sensing networks achieved significant improvements in terms of BER performance in AWGN channels.

The contributions of each chapter are summarized below.

In Chapter 3, the limit on the achievable sum rate and the BEP floor of the binary information sensing network, including the threshold SNR converted from the sum rate requirement and the BEP floor lower bound, was investigated. The main results are as follows.

– We proposed the encoding and decoding algorithms that exploit the correlation knowledge of sensing data, and introduced the convergence control using mutual information measurement to improve the efficiency of the joint decoding algorithm. The proposed joint decoding is *scalable* and is with *linear complexity* in terms of the number of sensors. Furthermore, the proposed joint decoding algorithm is also applicable to the relaying system using lossy forwarding and the automatic repeat request (ARQ) system [103].

– The *achievable sum rate* was derived using the Slepian-Wolf theorem, and was converted to corresponding SNR limit based on the source-channel separation theorem for transmitting correlated data over orthogonal AWGN channels. It was found that there exists an around 2 dB gap between the SNR limits and the simulation results. The reason is that the SNR limit was obtained based on using the capacity-achieving code.
– The BEP floor was analytically calculated based on Poisson binomial process and its lower bound was obtained based on the binary rate-distortion function and the conditional independent property.
– Finally, a 3D EXIT chart analysis was performed to confirm the simulation results.

Chapter 4 derived a theoretical lower bound on the Hamming distortion from rate-distortion perspective for the binary information sensing network withJSC setup, in which a single binary source is observed by multiple sensors and transmitted the coded data to the fusion center over orthogonal AWGN channels. The analysis, however, limited the decoding schemes to which separate the stages of JSC decoding and the final decision on the common source.

– We explored the specified case where the number of terminals $L = 2$ in details. The binary information sensing network was first modeled as a binary CEO problem. For solving the rate-distortion region for the two-terminal binary CEO problem, a more general problem, which is a binary multiterminal source coding, was considered. We then derived an outer bound on the rate-distortion region for the binary multiterminal source coding through the converse coding proof. The derived outer bound has a very simple form and can be applied to practical wireless sensor/relaying network to analyze the system performance.

  Moreover, comparisons between the outer bound and the classical results, including Slepian-Wolf rate region, Wyner-Ziv bound and Berger-Tung inner bound were made by drawing the rate-distortion region. Also, we derived a direct outer bound on the rate-distortion region for the binary CEO problem. Through the comparisons, our two-step outer bound [13] is not tight in general. However, if the observation accuracy is low and/or the distortions are small, the two-step outer bound has negligible gaps to the corresponding Berger-Tung inner bounds.

[13]The two-step outer bound is the established outer bound for the binary multiterminal source coding problem, and then map to the binary CEO problem based on decision rules.

After that, we established the relationship between the binary CEO problem and the binary multiterminal source coding using a distortion function for several specific decision rules. The theoretical lower bound on the Hamming distortion was then formulated in the framework of convex optimization. Finally, computer simulations using our proposed encoding/decoding algorithm were performed to verify the theoretical lower bound on the Hamming distortion.

– We further established an outer bound on the rate-distortion region through the converse coding proof for the binary CEO problem with arbitrary number of terminals. The derived outer bound, however, is only an approximation since it is lack of rigorous proof of using test channel to bound the mutual information. The lower bound on the Hamming distortion using the approximated outer bound still acted as a useful reference of the performance obtained using the encoding/decoding algorithms in the binary information sensing network. Furthermore, the superiority of our proposed joint decoding algorithm was investigated analytically with the comparison to the independent decoding scheme.

In Chapter 5, an optimal power allocation scheme was proposed to minimize the Hamming distortion subject to a fixed total power in the context of a cluster binary information sensing network.

– We formulated the power-distortion optimization problem based on the derived two-step outer bound in Chapter 4 and the source-channel separation theorem. A heuristic method was proposed to solve the formulated optimization problem. However, the complexity of the heuristic method is very high.
– To reduce the complexity, we reformulated the optimization problem by relaxing several conditions, such as, correlation was not taken into account. After that, to avoid the inverse of the binary entropy function, we modified the optimization problem using the property that distortion is proportional to the negative channel capacity. An analytical solution was subsequently obtained from the KKT conditions of the problem of maximizing weighted channel capacities.
– From the computer simulations, the proposed power allocation scheme achieved $1.5 \sim 2$ dB gains compared to the uniform power allocation case. In addition, the difference between the proposed power allocation scheme and the heuristic method was not significant.

6.2 Future Studies

In this section, we give several directions of extending our work.

Directly derive the rate-distortion bound for the binary CEO problem

As an important future study, the outer bound on the rate-distortion region for the binary CEO problem with multiple terminals should be directly analyzed and compare with the derived two-step outer bound with optimal decision, as the two-terminal binary CEO problem case. Also, the Berger-Tung inner bound should be extended to the case having multiple terminals. After that, comparisons between the Berger-Tung inner bound and the direct outer bound can be made to show the tightness of the bounds.

Rate-distortion Analysis with energy harvesting sensors

The energy harvesting sensor network has attracted a lot of attention recently in order to improve the lifetime of the sensor network [46, 104–109]. Based on the derived rate-distortion bound for the binary CEO problem, it is worth to study the design of the power allocation schemes and scheduling protocols for the binary information sensing network with energy harvesting.

Outage analysis of the parallel relaying network

The outage analysis of one-way relaying network using LF protocol was intensively studied in [65, 89, 110, 111], however, the problem becomes complicated when the number of relays increased. As an extension of our work, theoretical analysis of combining the rate-distortion analysis and the impact of fading variation is a good direction, including outage probability derivation and the design of power optimization scheme.

Power allocation for sensor network

In this work, we proposed a power allocation scheme for the binary information sensing network in AWGN channel. However, the power allocation works in centralized manner, where the channel state information and observation accuracy are needed in the cluster head. In the future, it is important to design distributed power allocation

scheme to avoid the signaling overhead of collecting information from each sensor by the cluster head.

Decision rule

So far, we considered the majority voting and soft combining to generate the hidden source information from the observations. Through simulations, the soft combining outperformed the majority voting, however, it has a clear gap to the theoretical lower bound. In the future, we need to find some other decision rules to improve the performance.

Source compression

In this work, source compression was not carried out in the the practical encoding design. However, it caused rate loss for the cases: (i) the source entropy is less than 1, and (ii) the channel is good. Combining source compression, using such as Huffman coding, and channel coding is left as a future study.

Nonorthogonal multiple access channel (NOMA)

We mainly focused on the orthogonal MAC, and thus source coding and channel coding can be separately analyzed according to the separability on source-channel coding. However, if the channels are nonorthogonal, the proposed method of analyzing the rate-distortion bound is suboptimal. As a future study, the analysis of rate-distortion behavior for binary information sensing should be extended to NOMA channels, since the throughput of the network is significantly improved by NOMA.

References

1. Xiong Z, Liveris AD & Cheng S (2004) Distributed source coding for sensor networks. IEEE Signal Processing Magazine 21(5): 80–94.
2. Slepian D & Wolf J (1973) Noiseless coding of correlated information sources. IEEE Transactions on Information Theory 19(4): 471–480.
3. Cover T (1975) A proof of the data compression theorem of Slepian and Wolf for ergodic sources (corresp.). IEEE Transactions on Information Theory 21(2): 226–228.
4. Ahlswede R & Körner J (1975) Source coding with side information and a converse for degraded broadcast channels. IEEE Transactions on Information Theory 21(6): 629–637.
5. Wyner A (1975) On source coding with side information at the decoder. IEEE Transactions on Information Theory 21(3): 294–300.
6. Körner J & Marton K (1979) How to encode the modulo-two sum of binary sources (corresp.). IEEE Transactions on Information Theory 25(2): 219–221.
7. Fang Y, Jeon G & Jeong J (2009) Rate-adaptive compression of LDPC syndromes for Slepian-Wolf coding. In: Proceedings of the Picture Coding Symposium (PCS), pp. 1–4.
8. Li J, Tu Z & Blum R (2004) Slepian-Wolf coding for nonuniform sources using turbo codes. In: Proceedings of the Data Compression Conference (DCC), pp. 312–321.
9. Eckford A & Yu W (2005) Rateless Slepian-Wolf codes. In: Proceedings of the Thirty-Ninth Asilomar Conference on Signals, Systems and Computers, pp. 1757–1761.
10. Gel'fand S & Pinsker M (1979) Coding of sources on the basis of observations with incomplete information (in Russian). Problemy Peredachi Informatsii (Problem of Infromation Transmission) 25(2): 219–221.
11. Wyner A & Ziv J (1976) The rate-distortion function for source coding with side information at the decoder. IEEE Transactions on Information Theory 22(1): 1–10.
12. Wyner A (1978) The rate-distortion function for source coding with side information at the decoder-II: General sources. Information and Control 38: 60–80.
13. Gastpar M (2004) The Wyner-Ziv problem with multiple sources. IEEE Transactions on Information Theory 50(11): 2762–2768.
14. Draper SC (2002) Successive structuring of source coding algorithms for data fusion, buffering, and distribution in networks. Ph.D. thesis, Electrical Engineering and Computer Science, Massachusetts Institute of Technology.
15. Aaron A, Rane S, Zhang R & Girod B (2003) Wyner-Ziv coding for video: applications to compression and error resilience. In: Proceedings of the Data Compression Conference (DCC), pp. 93–102.
16. Yaacoub C, Farah J & Pesquet-Popescu B (2007) Joint source-channel Wyner-Ziv coding in wireless video sensor networks. In: Proceedings of the IEEE International Symposium on Signal Processing and Information Technology, pp. 225–228.
17. Areia J, Ascenso J, Brites C & Pereira F (2007) Wyner-Ziv stereo video coding using a side information fusion approach. In: Proceedings of the IEEE 9th Workshop on Multimedia Signal Processing (MMSP), pp. 453–456.
18. Pereira F, Brites C, Ascenso J & Tagliasacchi M (2008) Wyner-Ziv video coding: A review of the early architectures and further developments. In: Proceedings of the IEEE International Conference on Multimedia and Expo, pp. 625–628.

19. Liu Z, Stanković V & Xiong Z (2005) Wyner-Ziv coding for the half-duplex relay channel. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '05), volume 5, pp. 1113–1116.
20. Uppal M, Liu Z, Stankovic V & Xiong Z (2009) Compress-forward coding with BPSK modulation for the half-duplex Gaussian relay channel. IEEE Transactions on Signal Processing 57(11): 4467–4481.
21. Sneessens H, Vandendorpe L & Laneman J (2009) Adaptive compress-and-forward relaying in fading environments with or without Wyner-Ziv coding. In: Proceedings of the IEEE International Conference on Communications (ICC), pp. 1–5.
22. Berger T (Ed. New York, Springer-Verlag, 1978) Multiterminal source coding. In: The Information Theory Approach to Communications, G. Longo, volume 229, pp. 171–231.
23. Tung SY (1978) Multiterminal Source Coding. Ph.D. thesis, Cornell University.
24. Oohama Y (1997) Gaussian multiterminal source coding. IEEE Transactions on Information Theory 43(6): 1912–1923.
25. Han T & Kobayashi K (1980) A unified achievable rate region for a general class of multiterminal source coding systems. IEEE Transactions on Information Theory 26(3): 277–288.
26. Yeung RWH (1988) Some results on multiterminal source coding. Ph.D. thesis, Cornell University.
27. Oohama Y (2005) Rate-distortion theory for Gaussian multiterminal source coding systems with several side informations at the decoder. IEEE Transactions on Information Theory 51(7): 2577–2593.
28. Pandya A, Kansal A, Pottie G & Srivastava M (2004) Lossy source coding of multiple Gaussian sources: m-helper problem. In: Proceedings of the IEEE Information Theory Workshop (ITW), pp. 34–38.
29. Wagner A & Anantharam V (2005) An improved outer bound for the multiterminal source-coding problem. In: Proceedings of the IEEE International Symposium on Information Theory (ISIT), pp. 1406–1410.
30. Berger T, Zhang Z & Viswanathan H (1996) The CEO problem. IEEE Transactions on Information Theory 42(3): 887–902.
31. Viswanathan H & Berger T (1997) The quadratic Gaussian CEO problem. IEEE Transactions on Information Theory 43(5): 1549–1559.
32. Ekrem E & Ulukus S (2012) An outer bound for the vector Gaussian CEO problem. In: Proceedings of the IEEE International Symposium on Information Theory Proceedings (ISIT), pp. 576–580.
33. Ekrem E & Ulukus S (2014) An outer bound for the vector Gaussian CEO problem. IEEE Transactions on Information Theory 60(11): 6870–6887.
34. Wang J & Chen J (2012) On the vector gaussian L-terminal CEO problem. In: Proceedings of the IEEE International Symposium on Information Theory Proceedings (ISIT), pp. 571–575.
35. Wagner A & Anantharam V (2008) An improved outer bound for multiterminal source coding. IEEE Transactions on Information Theory 54(5): 1919–1937.
36. Oohama Y (1998) The rate-distortion function for the quadratic Gaussian CEO problem. IEEE Transactions on Information Theory 44(3): 1057–1070.
37. Cristescu R, Beferull-Lozano B & Vetterli M (2004) On network correlated data gathering. In: Proceedings of the IEEE International Conference on Computer Communications (INFOCOM), volume 4, pp. 2571–2582.

38. Draper S & Wornell GW (2004) Side information aware coding strategies for sensor networks. IEEE Journal on Selected Areas in Communications 22(6): 966–976.
39. Behroozi H & Soleymani M (2005) Performance of the successive coding strategy in the CEO problem. In: Proceedings of the IEEE Global Communications Conference, volume 3, pp. 1347–1352.
40. Behroozi H & Soleymani M (2007) Distortion sum-rate performance of successive coding strategy in quadratic gaussian CEO problem. IEEE Transactions on Wireless Communications 6(12): 4361–4365.
41. Behroozi H & Soleymani M (2009) Optimal rate allocation in successively structured Gaussian CEO problem. IEEE Transactions on Wireless Communications 8(2): 627–632.
42. Chen J & Berger T (2008) Successive Wyner-Ziv coding scheme and its application to the quadratic Gaussian CEO problem. IEEE Transactions on Information Theory 54(4): 1586–1603.
43. Chen J, Zhang X, Berger T & Wicker S (2004) An upper bound on the sum-rate distortion function and its corresponding rate allocation schemes for the CEO problem. IEEE Journal on Selected Areas in Communications 22(6): 977–987.
44. Viswanath P (2004) Sum rate of Multiterminal Gaussian source coding, volume 229. Providence, RI: American Mathemtical Society.
45. Howard S & Flikkema P (2008) Integrated source-channel decoding for correlated data-gathering sensor networks. In: Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC), pp. 1261–1266.
46. Orhan O, Günduz D & Erkip E (2015) Source-channel coding under energy, delay, and buffer constraints. IEEE Transactions on Wireless Communications 14(7): 3836–3849.
47. Karlsson J & Skoglund M (2011) Lattice-based source-channel coding in wireless sensor networks. In: Proceedings of the IEEE International Conference on Communications (ICC), pp. 1–5.
48. Kochman Y & Zamir R (2008) The joint gaussian CEO/MAC problem and relaying. In: Proceedings of the 46th Annual Allerton Conference on Communication, Control, and Computing, pp. 1554–1561.
49. Haghighat J, Behroozi H & Plant D (2008) Iterative joint decoding for sensor networks with binary CEO model. In: Proceedings of the IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), pp. 41–45.
50. Razi A, Yasami K & Abedi A (2011) On minimum number of wireless sensors required for reliable binary source estimation. In: Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC), pp. 1852–1857. Mexico.
51. Razi A & Abedi A (2012) Adaptive bi-modal decoder for binary source estimation with two observers. In: Proceedings of the Annual Conference on Information Systems and Sciences (CISS), pp. 1–5. Princeton, NJ, USA.
52. Razi A & Abedi A (2014) Convergence analysis of iterative decoding for binary CEO problem. IEEE Transactions on Wireless Communications 13(5): 2944–2954.
53. Zhou X, He X, Anwar K & Matsumoto T (2012) GREAT-CEO: larGe scale distRibuted dEcision mAking Technique for wireless Chief Executive Officer problems. IEICE Transactions on Communications E95-B(12): 3654–3662.
54. Garcia-Frias J & Zhao Y (2005) Near-Shannon/Slepian-Wolf performance for unknown correlated sources over AWGN channels. IEEE Transactions on Communications 53(4): 555–559.

55. He X, Zhou X, Anwar K & Matsumoto T (2013) Estimation of Observation Error Probability in Wireless Sensor Networks. IEEE Communications Letters 17(6): 1073–1076.
56. Oohama Y (2009) Information sensing and multiterminal source coding (in Japanese). Proceedings of the Institute of Statistical Mathematics 57(2): 233–251.
57. ICT-619555 RESCUE Deliverable (2015) D1.2.1-Assessment on Feasibility, Achievability, and Limits. Technical report.
58. Zhong W & Garcia-Frias J (2004) Combining data fusion with joint source-channel coding of correlated sensors. In: Proceedings of the IEEE Information Theory Workshop (ITW), pp. 315–317.
59. Sartipi M & Fekri F (2004) Source and channel coding in wireless sensor networks using LDPC codes. In: Proceedings of the IEEE Conference on Sensor and Ad Hoc Communications and Networks, pp. 309–316.
60. He X, Zhou X, Komulainen P, Juntti M & Matsumoto T (2016) A lower bound analysis of Hamming distortion for a binary CEO problem with joint source-channel coding. IEEE Transactions on Communications 64(1): 343–353.
61. Jiang W, He X & Matsumoto T (2016) Power allocation in an asymmetric wireless sensor network. IEEE Communications Letters , submitted (April 2016).
62. He X, Zhou X, Juntti M & Matsumoto T (2015) Data and error rate bounds for binary data gathering wireless sensor networks. In: Proceedings of the IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), pp. 505–509.
63. He X, Zhou X, Juntti M & Matsumoto T (2016) A rate-distortion region analysis for a binary CEO problem. In: Proceedings of the IEEE 83rd Vehicular Technology Conference. Nanjing, China.
64. He X, Juntti M, Zhou X, Komulainen P & Matsumoto T (2015) CEO problem based analysis of D2D cooperative user pairing. In: Proceedings of the IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), pp. 685–689.
65. Zhou X, Cheng M, He X & Matsumoto T (2014) Exact and approximated outage probability analyses for decode-and-forward relaying system allowing intra-link errors. IEEE Transactions on Wireless Communications 13(12): 7062–7071.
66. Zhou X, Cheng M, He X, Anwar K & Matsumoto T (2014) Outage analysis of decode-and-forward relaying system allowing intra-link errors. In: Proceedings of the European Wireless Conference, pp. 1–6.
67. Zhou X, He X, Juntti M & Matsumoto T (2015) Outage probability of correlated binary source transmission over fading multiple access channels. In: Proceedings of the IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), pp. 96–100.
68. Jiang W, He X, Qian S, Juntti M & Matsumoto T (2015) Finite-SNR diversity-multiplexing tradeoff for decode-and-forward relaying system allowing intra-link errors. In: Proceedings of the IEEE 10th International Conference on Information, Communications and Signal Processing (ICICS).
69. Qian S, Zhou X, He X, Juntti M & Matsumoto T (2015) Outage analysis for lossy-forward relaying: Impact of line-of-sight component. IEEE Transactions on Vehicular Technology Submitted.
70. Tervo V, He X, Zhou X, Komulainen P, Kuehlmorgen S, Wolf A & Festag A (2016) An error rate model of relay communications with lossy forwarding and joint decoding. In:

Proceedings of the IEEE International Conference on Communications Workshops. Kuala Lumpur, Malaysia.
71. ICT-619555 RESCUE Deliverable (2014) D2.1.1-Interim Report of detailed coding/decoding algorithms and correlation estimation techniques. Technical report.
72. ICT-619555 RESCUE Deliverable (2015) D2.3-Feasibility Assessment in model based environments. Technical report.
73. Shannon C (1948) A mathematical theory of communication. Bell System Technical Journal 27: 379–423, 623–656.
74. Cover T & Thomas J (2006) Elements of Information Theory. USA: John Wiley & Sons, Inc., 2nd edition.
75. Wyner AD & Ziv J (1973) A theorem on the entropy of certain binary sequence and applications: Part I. IEEE Transactions on Information Theory 19(6): 769–772.
76. Liveris AD, Xiong Z & Georghiades CN (2002) Compression of binary sources with side information at the decoder using LDPC codes. IEEE Communications Letters 6(10): 440–442.
77. Stanković V, Liveris AD, Xiong Z & Georghiades CN (2006) On code design for the Slepian-Wolf problem and lossless multiterminal networks. IEEE Transactions on Information Theory 52(4): 1495–1507.
78. Gamal A & Kim Y (2011) Network Information Theory. Cambridge University Press.
79. Dragotti PL & Gastpar M (2009) Distributed Source Coding: Theory, Algorithms and Applications. Academic Press.
80. Cover T, Gamal A & Salehi M (1980) Multiple access channels with arbitrarily correlated sources. IEEE Transactions on Information Theory 26(6): 648–657.
81. Gastpar M & Vetterli M (2002) On the capacity of wireless networks: the relay case. In: Proceedings of the IEEE International Conference on Computer Communications, volume 3, pp. 1577–1586 vol.3.
82. Gastpar MC & Vetterli M (2003) Source-Channel Communication in Sensor Networks. Lecture Notes in Computer Science 2634: 162–177.
83. ten Brink S (1999) Convergence of iterative decoding. Electronics Letters 35(10): 806–808.
84. ten Brink S (2001) Convergence behavior of iteratively decoded parallel concatenated codes. IEEE Transactions on Communications 49(10): 1727 –1737.
85. Anwar K & Matsumoto T (2012) Accumulator-assisted distributed turbo codes for relay system exploiting source-relay correlations. IEEE Communications Letters 16(7): 1114–1117.
86. Richardson T & Urbanke R (2001) The capacity of low-density parity-check codes under message-passing decoding. IEEE Transactions on Information Theory 47(2): 599 –618.
87. Richardson T, Shokrollahi M & Urbanke R (2001) Design of capacity-approaching irregular low-density parity-check codes. IEEE Transactions on Information Theory 47(2): 619 – 637.
88. Fukawa K, Ormsub S, Tölli A, Anwar K & Matsumoto T (2012) Exit-constrained BICM-ID design using extended mapping. EURASIP Journal on Wireless Communications and Networking 2012: 1–17. URI: http://dx.doi.org/10.1186/1687-1499-2012-40.
89. Cheng M, Anwar K & Matsumoto T (2013) Outage probability of a relay strategy allowing intra-link errors utilizing Slepian-Wolf theorem. EURASIP Journal on Advances in Signal Processing 2013(1): 1–12.

90. Wang YH (1993) On the number of successes in independent trials. Statistica Sinica 3(2): 295–312.
91. Wolf A, Matthe M & Fettweis G (2015) Improved source correlation estimation in wireless sensor networks. In: Proceedings of the IEEE International Conference on Communication Workshop (ICCW), pp. 2121–2126.
92. Li X & Ritcey J (1999) Trellis-coded modulation with bit interleaving and iterative decoding. IEEE Journal on Selected Areas in Communications 17(4): 715–724.
93. Hagenauer J (2004) The EXIT Chart - Introduction to Extrinsic Information Transfer in iterative processing. In: Proceedings of the European Signal Processing Conference, pp. 1541–1548.
94. He X, Zhou X, Anwar K & Matsumoto T (2012) Wireless mesh networks allowing intra-link errors: CEO problem viewpoint. In: Proceedings of the International Symposium on Information Theory and Its Applications (ISITA), pp. 61–65. Hawaii, USA.
95. Lu PS, Zhou X, Anwar K & Matsumoto T (2014) Joint adaptive network-channel coding for energy-efficient multiple-access relaying. IEEE Transactions on Vehicular Technology 63(5): 2298–2305.
96. Xiao JJ & Luo ZQ (2007) Multiterminal source-channel communication over an orthogonal multiple-access channel. IEEE Transactions on Information Theory 53(9): 3255–3264.
97. Jana S (2009) Alphabet sizes of auxiliary random variables in canonical inner bounds. In: Proceedings of the 43rd Annual Conference on Information Sciences and Systems (CISS), pp. 67–71.
98. Chindapol A & Ritcey J (2001) Design, analysis, and performance evaluation for BICM-ID with square QAM constellations in rayleigh fading channels. IEEE Journal on Selected Areas in Communications 19(5): 944–957.
99. Yu W & Yuan J (2005) Joint source coding, routing and resource allocation for wireless sensor networks. In: Proceedings of the IEEE International Conference on Communications (ICC), volume 2, pp. 737–741 Vol. 2.
100. Yuan J & Yu W (2008) Joint source coding, routing and power allocation in wireless sensor networks. IEEE Transactions on Communications 56(6): 886–896.
101. Li J, Dey S & Evans J (2007) Maximal lifetime rate and power allocation for sensor networks with data distortion constraints. In: Proceedings of the IEEE International Conference on Communications (ICC), pp. 3678–3685.
102. Boyd S & Vandenberghe L (2004) Convex Optimization. Cambridge University Press, New York, NY, USA.
103. Irawan A, Anwar K & Matsumoto T (2015) Lossy forwarding technique for parallel multihop-multirelay systems. In: Proceedings of the IEEE 82nd Vehicular Technology Conference (VTC Fall),, pp. 1–5.
104. Brunelli D, Benini L, Moser C & Thiele L (2008) An efficient solar energy harvester for wireless sensor nodes. In: Proceedings of the Design, Automation and Test in Europe (DATE), pp. 104–109.
105. Ha D (2011) Small scale energy harvesting - principles, practices and future trends. In: Proceedings of the IEEE 14th International Symposium on Design and Diagnostics of Electronic Circuits Systems (DDECS), pp. 9–9.
106. Ren X & Liang W (2012) Delay-tolerant data gathering in energy harvesting sensor networks with a mobile sink. In: Proceedings of the IEEE Global Communications Conference (GLOBECOM), pp. 93–99.

107. Nayyar A, Basar T, Teneketzis D & Veeravalli V (2012) Communication scheduling and remote estimation with energy harvesting sensor. In: Proceedings of the IEEE 51st Annual Conference on Decision and Control (CDC), pp. 843–848.
108. Ding Z, Perlaza S, Esnaola I & Poor H (2014) Power allocation strategies in energy harvesting wireless cooperative networks. IEEE Transactions on Wireless Communications 13(2): 846–860.
109. Ku M, Li W, Chen Y & Liu K (2015) Advances in energy harvesting communications: Past, present, and future challenges. IEEE Communications Surveys & Tutorials PP(99): 1–1.
110. Cheng M, Anwar K & Matsumoto T (2012) Outage analysis of correlated source transmission in block rayleigh fading channels. In: Proceedings of the IEEE Vehicular Technology Conference, pp. 1–5.
111. Qian S, Cheng M, Anwar K & Matsumoto T (2013) Outage probability analysis for correlated sources transmission over Rician fading channels. In: Proceedings of the IEEE 24th International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC), pp. 1087–1091.
112. Brannstrom F, Rasmussen L & Grant A (2005) Convergence analysis and optimal scheduling for multiple concatenated codes. IEEE Transactions on Information Theory 51(9): 3354–3364.
113. Chen D & Plemmons RJ (2007) Nonnegativity constraints in numerical analysis. In: Proceedings of the Symposium on the Birth of Numerical Analysis. Leuven Belgium.
114. Courtade TA (2012) Two problems in multiterminal information theory. Ph.D. thesis, Department of Electrical Engineering, University of California, Los Angeles.
115. Kostina V & Verdu S (2012) Fixed-length lossy compression in the finite blocklength regime. IEEE Transactions on Information Theory 58(6): 3309–3338.
116. Lin S & Costello DJ (2004) Error Control Coding. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2nd edition.
117. Liu K & Sayeed A (2007) Type-based decentralized detection in wireless sensor networks. IEEE Transactions on Signal Processing 55(5): 1899–1910.
118. Proakis J & Salehi M (2007) Digital Communications. McGraw-Hill Education, 5th edition.
119. Shannon CE (1959) Coding theorems for a discrete source with a fidelity criterion. IRE National Convention Record 4(1): 142–163.
120. Tüchler M (2002) Convergence prediction for iterative decoding of threefold concatenated systems. In: Proceedings of the IEEE Global Communications Conference, volume 2, pp. 1358–1362.
121. Vazquez-Araujo F, Fresnedo O, Castedo L & Garcia-Frias J (2014) Analog joint source-channel coding over MIMO channels. EURASIP Journal on Wireless Communications and Networking 2014(1): 25.
122. Viswanathan R & Varshney P (1997) Distributed detection with multiple sensors I. Fundamentals. Proceedings of the IEEE 85(1): 54–63.
123. Yuan J & Yu W (2008) Joint source coding, routing and power allocation in wireless sensor networks. IEEE Transactions on Communications 56(6): 886–896.
124. Zhang H & Ma H (2010) Delay-efficient rate control for Wyner-Ziv video coding in wireless video sensor networks using network coding. In: Proceedings of the IEEE International Conference on Multimedia and Expo (ICME), pp. 243–248.

Appendix 1 Conditional Independence of Observations

The CI property is used in deriving the rate-distortion function as well as the theoretical BEP floor lower bound. Therefore, we give an intuitive verification of the CI for the correlated binary discrete memoryless sources. Assume that X is a binary source generated from $Bern(p)$, and X_i is the noisy observations of X through independent BSCs. According to the basic probability theory, we have the following Table 10 to show how to calculate the joint probability $p_{XX_1X_2}(x,x_1,x_2)$.

Table 10. Joint probability of correlated binary sources.

x	x_1	x_2	$p_X(x)$	$p_{XX_1X_2}(x,x_1,x_2)$
0	0	0	p	$p\bar{p}_1\bar{p}_2$
0	0	1	p	$p\bar{p}_1p_2$
0	1	0	p	$pp_1\bar{p}_2$
0	1	1	p	pp_1p_2
1	0	0	$\bar{p}$	$\bar{p}p_1p_2$
1	0	1	$\bar{p}$	$\bar{p}p_1\bar{p}_2$
1	1	0	$\bar{p}$	$\bar{p}\bar{p}_1p_2$
1	1	1	$\bar{p}$	$\bar{p}\bar{p}_1\bar{p}_2$

In the table, the notation $\bar{p} = 1-p$. It is easily found that $p_{XX_1X_2}(x,x_1,x_2) = p_X(x)p_{X_1|X}(x_1|x)p_{X_2|X}(x_2|x)$. And hence, the observations are conditional independent by given the underlying source. The calculation can be straightforward extended to the case of $p_{XX_1\cdots X_L}(x,x_1,\cdots,x_L)$.

Appendix 2 BEP Floor of Soft Combining Decision

For the case p_i are various over links, the BEP floor obtained by Poisson-binomial process is not accurate. Thus, the BEP floor needs to be analyzed by taking soft combining into account. In soft combining decision, the LLR sequence is weighted by p_i using function f_c. It is equivalent to a weighted majority voting, for which the hard decision of $\hat{x}$ follows

$$\hat{x} = \begin{cases} 1, & \mathbf{w}\mathbf{v}^{\mathrm{T}} > 0 \\ 0, & \text{otherwise} \end{cases} \tag{143}$$

where $\mathbf{w} = [\log\frac{1-p_1}{p_1}, \cdots, \log\frac{1-p_L}{p_L}]$ and $\mathbf{v} = \mathrm{sign}([l_1^{\mathrm{p}}, \cdots, l_L^{\mathrm{p}}])$, with $\mathrm{sign}(\cdot)$ taking the sign of its argument, i.e., 1 for positive numbers, -1 for negative numbers. Similar to the Poisson binomial process by assuming that 0 is transmitted, the BEP floor is given by

$$p_e = \Pr\{\sum_{k\in\mathscr{V}_+} w_k > \sum_{j\in\mathscr{V}_-} w_j\} + \frac{1}{2}\Pr\{\sum_{k\in\mathscr{V}_+} w_k = \sum_{j\in\mathscr{V}_-} w_j\}, \tag{144}$$

where $\mathscr{V}_+ = \{i|v_i = +1\}$ and $\mathscr{V}_- = \{i|v_i = -1\}$. Note that the difference between the BEP floor using Poisson binomial process and (144) is that, the error is determined by the number of 1's and 0's in Poisson binomial, and the weights of positive and negative signs in soft combining, respectively. To compute (144), it needs to carry out the search of the possible combinations of w_i over the power set of $\{1, \cdots, L\}$.

Appendix 3 Calculation of the Inner Bound

The Berger-Tung inner bound on the rate-distortion region for the binary case with Hamming distortion is calculated based on the test BSCs. By using the Markov property and the chain rules of entropy and mutual information, $\mathscr{R}^{\mathrm{i}}(D_1, D_2)$ shown in (97) is obtained in the following way

$$\begin{aligned}
R_1 &\geq I(X_1; V_1|V_2) \\
&= H(V_1|V_2) - H(V_1|X_1, V_2) && (145) \\
&= H(V_1|V_2) - H(V_1|X_1) && (146) \\
&= H_2(D_1 * p_1 * p_2 * D_2) - H_2(D_1) \\
R_2 &\geq H_2(D_1 * p_1 * p_2 * D_2) - H_2(D_2) && (147)
\end{aligned}$$

$$\begin{aligned}
R_1 + R_2 &\geq I(X_1, X_2; V_1, V_2) \\
&= H(V_1, V_2) - H(V_1, V_2|X_1, X_2) \\
&= H(V_1) + H(V_1|V_2) - H(V_1|X_1, X_2) - H(V_2|X_1, X_2, V_1) && (148) \\
&= 1 + H_2(D_1 * p_1 * p_2 * D_2) - H(V_1|X_1) - H(V_2|X_1, X_2) && (149) \\
&= 1 + H_2(D_1 * p_1 * p_2 * D_2) - H(V_1|X_1) - H(V_2|X_2) && (150) \\
&= 1 + H_2(D_1 * p_1 * p_2 * D_2) - H_2(D_1) - H_2(D_2)
\end{aligned}$$

where the steps are justified, with:
(145) the chain rule for mutual information,
(146) given X_1, V_1 and V_2 are conditionally independent,
(147) symmetric to the calculation of R_1,
(148) the chain rule for entropy,
(149) given X_1, X_2 and V_1 are conditionally independent, also, V_1 and V_2 are conditionally independent given X_1 and X_2,
(150) given X_2, X_1 and V_2 are conditionally independent.

It should be emphasized here that the timing sharing variable Q is not involved in the above calculation, while the equations are based on [22]. In order to visually present the Berger-Tung inner bound, the rate-distortion region is divided into three parts, as

(a) for some $0 \leq \tilde{d} \leq D_2$

$$\begin{cases} R_1 & \geq & H_2(D_1 * p_1 * p_2 * \tilde{d}) - H_2(D_1), \\ R_2 & \geq & 1 - H_2(\tilde{d}), \end{cases} \tag{151}$$

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

$$\begin{cases} R_2 & \geq & H_2(D_2 * p_1 * p_2 * \tilde{d}) - H_2(D_2) \\ R_1 & \geq & 1 - H_2(\tilde{d}); \end{cases} \tag{152}$$

(c)

$$R_1 + R_2 \geq 1 + H_2(D_1 * p_1 * p_2 * D_2) - H_2(D_1) - H_2(D_2), \tag{153}$$

where $\tilde{d}$ is a dummy variable. We calculate the rates R_1, R_2 as well as $R_1 + R_2$ with given D_1 and D_2, respectively, and then plot the rate-distortion region by combining the three parts shown above, which is similar to the time sharing concept.

Appendix 4 Direct Outer Bound for the Two-node Binary CEO Problem

Let

$$\mathscr{R}_L(D)=\Big\{(R_1,\cdots,R_L,D):\quad \text{there exists } \varphi_1,\cdots,\varphi_L,\psi)\quad \text{such that}$$

$$\frac{1}{n}\log|\varphi_i|\le R_i \tag{154}$$

$$\frac{1}{n}Ed(X^n,\hat{X}^n)=\frac{1}{n}\sum_{t=1}^{n}Ed(X(t),\hat{X}(t))\le D\Big\}. \tag{155}$$

be the rate-distortion region of the binary CEO problem with $d(,)$ being the Hamming distortion measure and $|\varphi_i|=2^{nR_i}$ denoting the range of cardinality of φ_i. Our aim is to derive a good outer bound on $\mathscr{R}_L(D)$. Assume that $(R_1,R_2,D)\in\mathscr{R}_L(D)$ and define

$$\begin{cases} U_i=\varphi_i(X_i^n)\\ \tilde{X}_t=[X(1),\cdots,X(t-1),X(t+1),\cdots,X(n)]\\ \varepsilon_t=\Pr\{x_i\neq\hat{x}_i\}. \end{cases} \tag{156}$$

Then we can obtain the inequality as

$$\frac{1}{n}H(X^n|\hat{X}^n)\le H_2(D). \tag{157}$$

A proof of (157) is shown as follows.

$$\begin{aligned} \frac{1}{n}H(X^n|\hat{X}^n) &= \frac{1}{n}\sum_{t=1}^{n}H(X(t)|X^{t-1}\hat{X}^n) && (158)\\ &\le \frac{1}{n}\sum_{t=1}^{n}H(X(t)|\hat{X}(t)) && (159)\\ &\le \frac{1}{n}\sum_{t=1}^{n}(\varepsilon_t(\log(|\mathscr{X}|-1))+H_2(\varepsilon_t)) && (160)\\ &= \frac{1}{n}\sum_{t=1}^{n}H_2(\varepsilon_t) && (161)\\ &\le H_2(\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t) && (162)\\ &\le H_2(D) && (163) \end{aligned}$$

with the steps being justified by
(158) chain rule of entropy,
(159) conditioning reduces entropy,
(160) Fano's inequality,
(162) Jensen's inequality and H_2 is concave,
(163) definition.

Based on the assumption $(R_1, R_2, D) \in \mathscr{R}_L(D)$ and several steps of basic calculation, we have

$$\begin{cases} R_i & \geq \frac{1}{n}\sum_{t=1}^{n} I(U_i; X_i(t)|U_{\mathscr{L}\setminus i}, \tilde{X}_t) \\ R_{\text{sum}} & \geq \frac{1}{n}\sum_{t=1}^{n} I(U_{\mathscr{L}}; X_{\mathscr{L}}(t)|\tilde{X}_t) \\ H_2(D) & \geq \frac{1}{n} H(X|U_{\mathscr{L}}) \end{cases}, \tag{164}$$

and the outer bound on $\mathscr{R}_L(D)$ as

$$\begin{aligned} \mathscr{R}_L^{\text{o}}(D) = \Big\{ (R_{\mathscr{L}}) : & R_i \geq I(U_i; X_i|U_{\mathscr{L}\setminus i}) \\ & R_{\text{sum}} \geq I(U_{\mathscr{L}}; X_{\mathscr{L}}) \\ & H_2(D) \geq H(X|U_{\mathscr{L}}) \\ & \text{for some } U_i \text{ with independent Markov chains} \\ & U_i \to X_i \to X \to X_{\mathscr{L}\setminus i} \to U_{\mathscr{L}\setminus i} \\ & X \to (X_1, \cdots, X_L) \to (U_1, \cdots, U_L) \\ & |\mathscr{U}_i| \leq |\mathscr{X}_i| + 7 \Big\}. \end{aligned} \tag{165}$$

The proof of $\mathscr{R}_L(D) \subseteq \mathscr{R}_L^{\text{o}}(D)$ is based on the lossless CEO problem originated by Gel'fand and Pinsker [10] and is omitted here.

The outer bound $\mathscr{R}_L^{\text{o}}(D)$ needs to be computed using a good parametrization method. Similar to the quadratic Gaussian CEO problem, we introduce following terms

$$\eta_i = I(X_i; U_i|X) \text{ for all } i \in \mathscr{L} \tag{166}$$

to parameterize the outer bound. Then $\mathscr{R}_L^{\text{o}}(D)$ is represented by

$$\begin{cases} R_i & \geq \eta_i + H(X|U_{\mathscr{L}\setminus i}) - H_2(D) \\ R_{\text{sum}} & \geq 1 - H_2(D) + \sum_{i=1}^{L} \eta_i \end{cases} \tag{167}$$

The term η_i can be easily obtained by MGL. Then the questions remain for future study are a series of minimization problems, denoted as $\min H(X|U_{\mathscr{S}})$ with $\mathscr{S} \subseteq \mathscr{L}$. So far,

in this dissertation, we use test BSC to obtain the minimal value on $H(X|U_1,U_2)$ for the two-node case. It is found that the results of using test BSC consistent with that of using a brute-force search over a fine mesh of conditional distributions $p_{U_i|X_i}(u_i|x_i)$ [114].

Appendix 5 Monotonicity of Distortion D

Majority decision. $D = \min\{\theta_1, \theta_2\}$. Since θ_i, $i = 1,2$ is the result of the binary convolution of p_i and D_i, θ_i is obviously increasing as D_i is increasing, when p_i is fixed.

Optimal decision. $D = H_2^{-1}[H_2(\theta_1) + H_2(\theta_2) - H_2(\theta_1 * \theta_2)]$.

In this case, D is a composite function of $H_2^{-1}(\cdot)$ and $H_2(\theta_1) + H_2(\theta_2) - H_2(\theta_1 * \theta_2)$. Since the function $H_2^{-1}(\cdot)$ is monotonically increasing, we only need to prove that $g(\theta_1, \theta_2) = H_2(\theta_1) + H_2(\theta_2) - H_2(\theta_1 * \theta_2)$ is also an increasing function of θ_1 and θ_2.

Assume θ_2 is fixed. The partial derivative $\frac{\partial g(\theta_1,\theta_2)}{\partial \theta_1}$ on θ_1 is

$$\frac{\partial g(\theta_1, \theta_2)}{\partial \theta_1} = \log \frac{1-\theta_1}{\theta_1} - (1-2\theta_2) \cdot \log \frac{1-\theta_1 * \theta_2}{\theta_1 * \theta_2}. \tag{168}$$

In order to prove that (168) is nonnegative, we should prove

$$\frac{1-\theta_1}{\theta_1} \geq \left(\frac{1-\theta_1 * \theta_2}{\theta_1 * \theta_2}\right)^{(1-2\theta_2)}. \tag{169}$$

The above always holds according to the monotonically increasing property of function $\log(\cdot)$. As $0 \leq \theta_i \leq \frac{1}{2}$, $i = 1,2$ and $0 \leq \theta_1 * \theta_2 \leq \frac{1}{2}$ is assumed, the following inequalities are obtained after several steps of elementary calculation

$$\frac{1-\theta_1}{\theta_1} \geq \left(\frac{1-\theta_1}{\theta_1}\right)^{(1-2\theta_2)} \geq \left(\frac{1-\theta_1 * \theta_2}{\theta_1 * \theta_2}\right)^{(1-2\theta_2)}. \tag{170}$$

Therefore, it is found that (168) can not take negative values according to (170). Symmetrically, we can assume θ_1 is fixed, and show that the partial derivative $\frac{\partial g(\theta_1,\theta_2)}{\partial \theta_2}$ on θ_2 is also nonnegative. Hence, $g(\theta_1, \theta_2)$ is increasing in the dimension of θ_1 and θ_2, respectively. Based on the above two cases, it is concluded that the distortion D is increasing with respect to D_1 and D_2.

Appendix 6 Outage Probability Derivation

Following the method of deriving the outage probability for one-way relaying network [110], the outage probability p_{out} of binary information sensing over Rayleigh fading channels is obtained. Based on the Slepian-Wolf theorem, if the rate pair (R_1, R_2) falls into parts 1 and 2, as shown in Fig. 52, both the correlated sources can be recovers in arbitrarily small error probability. Furthermore, if (R_1, R_2) is in part 3 or 4, source 2 or 1 can be losslessly recovered. Define P_i, $i = 1,2,3,4$ as the probabilities that (R_1, R_2) falls into part i, respectively, then based on the definition of the outage event, p_{out} is expresses as

$$p_{\text{out}} = \begin{cases} 1-(P_1+P_2+P_4), & p_1 < p_2 \\ 1-(P_1+P_2+P_3), & p_1 > p_2 \\ 1-(P_1+P_2+P_3+P_4), & p_1 = p_2 \end{cases} . \tag{171}$$

Note that outage happens if and only if the final distortion D is larger than $\min\{p_1, p_2\}$, and thus, the sensor with smaller p_i or both sensors if $p_1 = p_2$ dominate the performance. In the calculation of outage probability, P_3 and/or P_4 should be subtracted accordingly. Now, we need to compute P_1, P_2, P_3 and P_4.

Assume the instantaneous SNR γ_i, $i = 1,2$ follows Rayleigh fading, as

$$p_{\Gamma_i}(\gamma_i) = \frac{1}{\Gamma_i} \exp(-\frac{\gamma_i}{\Gamma_i}). \tag{172}$$

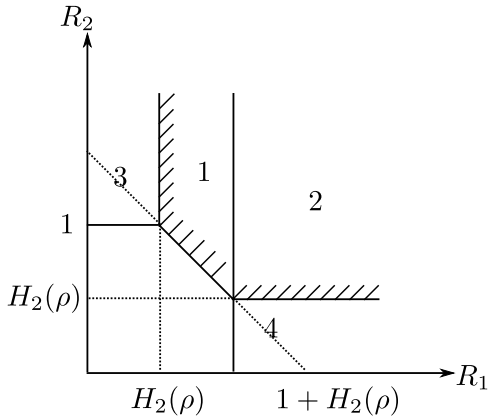


Fig. 52. Slepian-Wolf rate region for analyzing outage probability.

Based on the Slepain-Wolf theorem and separability of source and channel, we have

$$\begin{aligned}
P_1 &= \Pr\{H_2(\rho) < R_1 < 1, R_1 + R_2 > 1 + H_2(\rho)\} \\
&= \Pr\{2^{r_1 H_2(\rho)} - 1 < \gamma_1 < 2^{r_1} - 1, 2^{[r_2(1+H_2(\rho)) - \frac{r_2}{r_1}\log(1+\gamma_1)]} - 1 < \gamma_2\} \\
&= \int_{2^{r_1 H_2(\rho)}-1}^{2^{r_1}-1} \int_{2^{[r_2(1+H_2(\rho)) - \frac{r_2}{r_1}\log(1+\gamma_1)]}-1}^{+\infty} p_{\Gamma_1}(\gamma_1) p_{\Gamma_2}(\gamma_2) d\gamma_1 d\gamma_2 \\
&= \frac{1}{\Gamma_1} \int_{2^{r_1 H_2(\rho)}-1}^{2^{r_1}-1} \exp(-\frac{\gamma_1}{\Gamma_1}) \left[\exp(-\frac{\gamma_1}{\Gamma_1})\right]_{2^{[r_2(1+H_2(\rho)) - \frac{r_2}{r_1}\log(1+\gamma_1)]}-1}^{+\infty} d\gamma_1 \\
&= \frac{1}{\Gamma_1} \int_{2^{r_1 H_2(\rho)}-1}^{2^{r_1}-1} \exp(\frac{1}{\Gamma_2} - \frac{\gamma_1}{\Gamma_1} - \frac{2^{r_2(1+H_2(\rho))}}{\Gamma_2 (1+\gamma_1)^{\frac{r_2}{r_1}}}) d\gamma_1, \qquad (173)
\end{aligned}$$

$$\begin{aligned}
P_2 &= \Pr\{R_1 > 1, R_2 > H_2(\rho)\} \\
&= \Pr\{2^{r_1} - 1 < \gamma_1 < \infty, 2^{r_2 H_2(\rho)} - 1 < \gamma_2 < \infty\} \\
&= \int_{2^{r_1}-1}^{+\infty} \int_{2^{r_2 H_2(\rho)}-1}^{+\infty} p_{\Gamma_1}(\gamma_1) p_{\Gamma_2}(\gamma_2) d\gamma_1 d\gamma_2 \\
&= \int_{2^{r_1}-1}^{+\infty} p_{\Gamma_1}(\gamma_1) d\gamma_1 \int_{2^{r_2 H_2(\rho)}-1}^{+\infty} p_{\Gamma_2}(\gamma_2) d\gamma_2 \\
&= \exp[-(\frac{2^{r_1}-1}{\Gamma_1} + \frac{2^{r_2 H_2(\rho)}-1}{\Gamma_2})], \qquad (174)
\end{aligned}$$

$$\begin{aligned}
P_3 &= \Pr\{0 < R_1 < H_2(\rho), R_2 > 1\} \\
&= \Pr\{0 < \gamma_1 < 2^{r_1 H_2(\rho)} - 1, 2^{r_2} - 1 < \gamma_2 < \infty\} \\
&= \int_{0}^{2^{r_1 H_2(\rho)}-1} \int_{2^{r_2}-1}^{+\infty} p_{\Gamma_1}(\gamma_1) p_{\Gamma_2}(\gamma_2) d\gamma_1 d\gamma_2 \\
&= \int_{0}^{2^{r_1 H_2(\rho)}-1} p_{\Gamma_1}(\gamma_1) d\gamma_1 \int_{2^{r_2}-1}^{+\infty} p_{\Gamma_2}(\gamma_2) d\gamma_2 \\
&= [1 - \exp(-\frac{2^{r_1 H_2(\rho)}-1}{\Gamma_1})] \exp(-\frac{2^{r_2}-1}{\Gamma_2}). \qquad (175)
\end{aligned}$$

Similarly,

$$P_4 = [1 - \exp(-\frac{2^{r_2 H_2(\rho)}-1}{\Gamma_2})] \exp(-\frac{2^{r_1}-1}{\Gamma_1}). \qquad (176)$$

The outage is then obtained by (171) depending on the values of p_1 and p_2. Note that the derivation is using the capacity function with two-dimensional signal. The outage probability can be similarly calculated for the one-dimensional signal.

Appendix 7 The KKT Conditions of (138)

The KKT conditions for the optimization problem (138) are summarized as follows.

$$\begin{aligned}
\beta_i^\star &\geq 0, && (177)\\
\sum_i \beta_i^\star &= 1, && (178)\\
\lambda_i^\star &\geq 0, && (179)\\
\lambda_i^\star \beta_i^\star &= 0, && (180)\\
-\frac{(1-2p_i)E_T}{1+\beta_i^\star E_T} - \lambda_i^\star + \mu^\star &= 0, && (181)
\end{aligned}$$

where $i = 1, \cdots, L$, and λ_i, μ are the introduced Lagrange multipliers for the constraints. It is found that this problem is very similar to the water-filling algorithm [118] in wireless communications, and we can easily get the analytical solutions for $\beta_i^\star$ from the KKT conditions.

List of original publications

1. 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, June 2013.
2. 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 Transaction on Communications, vol.64, no.1, pp. 343–353, January 2016.
3. W. Jiang, X. He and T. Matsumoto, "Power Allocation in an Asymmetric Wireless Sensor Network", IEEE Communications Letters, under review. (corresponding author)
4. 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 Transaction on Wireless Communication, vol. 13, no.12, pp. 7062–7071 December 2014.
5. S. Qian, X. Zhou, X. He, M. Juntti and T. Matsumoto, "Outage Analysis for Lossy-Forward Relaying: Impact of Line-of-Sight Component", IEEE Transaction on Vehicular Technology, under review.
6. X. He, M. Juntti, X. Zhou, P. Komulainen and T. Matsumoto, "CEO Problem based Analysis of D2D Cooperative User Pairing", IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Stockholm, Sweden, June 28–July 1, 2015.
7. X. He, X. Zhou, M. Juntti and T. Matsumoto, "Data and Error Rate Bounds for Binary Data Gathering Wireless Sensor Networks", IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Stockholm, Sweden, June 28–July 1, 2015.
8. X. He, X. Zhou, M. Juntti and T. Matsumoto, "A Rate-Distortion Region Analysis for a Binary CEO Problem", IEEE 83rd Vehicular Technology Conference (VTC) 2016 Spring, Nanjing, May 15–18, 2016.
9. X. Zhou, X. He, M. Juntti and T. Matsumoto, "Outage Probability of Correlated Binary Source Transmission over Fading Multiple Access Channels", IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Stockholm, Sweden, June 28–July 1, 2015.
10. W. Jiang, X. He, S. Qian, M. Juntti and T. Matsumoto, "Finite-SNR Diversity-Multiplexing Tradeoff for Decode-and-Forward Relaying System Allowing Intra-link Errors", IEEE 10th International Conference on Information, Communications and Signal Processing (ICICS), Singapore, December 2–4, 2015.
11. W. Jiang, X. He, S. Qian and T. Matsumoto, "Opportunistic Relay Selection for Lossy-forwarding", IEEE PIMRC, 2016, under review.
12. V. Tervo, X. He, X. Zhou, P. Komulainen, S. Kuehlmorgen, A. Wolf and A. Festag, "An error rate model of relay communications with lossy forwarding and joint decoding", IEEE International Conference on Communications Workshop, Kuala Lumpur, Malaysia, May 23–27, 2016.
13. X. Zhou, M. Cheng, X. He, K. Anwar and T. Matsumoto, "Outage Analysis of Decode-and-Forward Relaying System Allowing Intra-link Errors", European Wireless 2014, Barcelona, May 14–16, 2014.