Large Scale Distributed Decision Making Technique and Wireless Chief Executive Officer Problem

By Xin He

A thesis submitted to
School of Information Science,
Japan Advanced Institute of Science and Technology,
in partial fulfillment of the requirements
for the degree of
Master of Information Science
Graduate Program in Information Science

Written under the direction of
Professor Tad Matsumoto

March, 2013

Large Scale Distributed Decision Making Technique and Wireless Chief Executive Officer Problem

By Xin He (1110017)

A thesis submitted to
School of Information Science,
Japan Advanced Institute of Science and Technology,
in partial fulfillment of the requirements
for the degree of
Master of Information Science
Graduate Program in Information Science

Written under the direction of
Professor Tad Matsumoto

and approved by
Professor Tad Matsumoto
Professor Mineo Kaneko
Professor Kiyofumi Tanaka

February, 2013 (Submitted)

I certify that I have prepared this Master's Thesis by myself without any inadmissible outside help.

Xin He
JAIST, Feb, 06, 2013

Author : _____

Date : _____

Supervisor : _____

Acknowledgments

First of all, I would like to show my deepest gratitude to my supervisor, Prof. Tad Matsumoto for his guidance, advice, assistance and support during my Master study. Furthermore, I wish to thank him for sharing lots of experience on how to do research and introducing me many fantastic things, extremely tasty food and a variety of beers.

I gratefully thank all the members in Matsumoto Laboratory for their kindly helps to me not only in the research but also in the daily life.

Finally, I would like to thank my parents, grandparents, relatives and Mr. Fengqing Xu for encouraging and supporting me during my whole study in Japan, and my wife Han Ge for her accompanying, supporting, and understanding.

Abstract

The primary goal of this thesis is to re-formulate various wireless networks having complex structures from the viewpoint of the Chief Executive Officer (CEO) problem. Particularly, we focus on the two major problems, wireless sensor networks (WSNs) and wireless mesh networks (WMNs). The CEO problem is described as follows: a firm's CEO aims to estimate a source data sequence $\mathbf{u}$ which cannot be observed directly; the CEO deploys a team of M agents to independently observe corrupted versions of the source data sequence $\mathbf{u}$; the agents then transmit their observations to the CEO for further processing under sum-rate constrained channels. The CEO tries to form $\mathbf{u}$ based on the received M noisy observations while keeping the distortion lower than an acceptable level. We first consider a parallel WSN modeled by the CEO problem. We apply a very simple coding scheme to the sensors (agents) and propose an efficient decoding technique with an algorithm estimating the observation error probabilities. Based on the convergence and bit-error rate (BER) performances, the proposed technique can achieve significant gains over the scheme without utilizing the correlation knowledge among the sensors. Furthermore, the BER and frame-error rate (FER) performances of the proposed technique using estimated observation error probabilities is very close to that of the proposed technique using exact observation error probabilities at the fusion center (FC).

We then focus on an issue of WMNs as a CEO problem, and provide a practical solution to a simple case of the problem. A joint decoding technique at the final destination (FD) is proposed by using the correlation knowledge between the originator-forwarding node *intra-links*. The BER performances show that the originator's information can be reconstructed at the FD even by using a very simple coding scheme. Moreover, we provide BER performance comparison between joint decoding and separate decoding strategies. The optimization of coding rate by using irregular convolutional code at originator side is also discussed in this thesis. The simulation results show that excellent performance can be achieved by the proposed system. Furthermore, extrinsic information transfer (EXIT) chart analysis is performed to investigate convergence property of the proposed technique.

Keywords: Error Probability, Estimation, CEO problem, Slepian-Wolf, wireless sensor networks, wireless mesh networks, LLR updating function

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Almost two decades have passed since the turbo coding technique was discovered by Berrou at al [1], the information and coding theory community has experienced many significant achievements. As a matter of fact, a lot of near capacity-achieving coding techniques have been proposed for point-to-point communication systems since the discovery of the Turbo codes.

Nowadays, the major focus of the research topics in information and coding theory community has shifted from point-to-point to network communications, such as cooperative communications using joint source, channel and network coding. Recently, a lot of significant theorems/theories as well as practical algorithms that can asymptotically achieve the performance bounds supported by the theory have been found in this category; the new results have been utilized by many information theory communities, including our research group, with the aim of creating a beyond-the-state-of-the-art technologies in communications systems. The most significant achievement of our research group is that the establishment of technological basis for distributed cooperative communications [2, 3] where even though *intra-links* (source-to-relay and relay-to-relay links) suffer from errors, still the transmitted information can be reconstructed at the destination by utilizing the correlation knowledge between the information part sent from the originator and relayed information.

In fact, this is a straightforward application of Slepian-Wolf theorem [4]. However, it is based on a strong assumption that the information at the source itself does not contain any errors. The goal of this research is to eliminate this assumption, and hence all the links are unreliable. This problem is called Chief Executive Officer (CEO) problem [5]

in networks information theory. The CEO problem can be applied to many forms of applications in wireless networks. This motivates us to work on the CEO problem and its applications to wireless communications. Our aim is to make a paradigm shift from the Slepian-Wolf lossless-based wireless network design to lossy link-based design, based on the CEO problem frame work.

1.2 The CEO Problem

Many categories of wireless networks, especially wireless sensor networks (WSNs), can be modeled as a CEO problem. This problem belongs to distributed lossy compression problem in networks information theory and closely related to multiple description coding problem. The terminology CEO problem originates from the situation where a firm's CEO is interested in a source data sequence $\mathbf{u}$ which cannot be observed directly. The CEO deploys a team of M agents to observe the source data sequence $\mathbf{u}$. The agents, referred to as forwarding nodes/sensors, observe independently corrupted versions of the source data sequence and then transmit their observations to a single fusion center (FC) for further processing. The CEO (FC) aims to form $\mathbf{u}$ based on the received M noisy observations while keeping the distortion lower than an acceptable level. Fig. 1.1 depicts the abstract scenario of the CEO problem system. The encoders, which are not allowed to exchange any message to utilize the correlation knowledge, are distributed in a geographical area. Since the power and the resource are limited, M agent are required to communicate with the FC under a finite sum rate constraint R. Therefore, it is crucial to introduce some coding techniques requiring very low power consumption when making such data gathering/monitoring systems.

In general, the target of the CEO problem is to determinate the minimum achievable distortion under a given sum rate constraint R. The sum rate means that the combined data rate at which the M agents may communicate information about their observations to the CEO. Berger, Zhang, and Viswanathan determine the asymptotic behavior of the minimal error occurrence probability in the limit as the number of agents M and the sum rate constraint R tend to infinity in [5] for the case where the source and observations are discrete and memoryless. The special case, where zero-mean Gaussian source is considered and the observations are corrupted by identical independent memoryless Gaussian noise with a minimum mean squared-error distortion, is called quadratic Gaussian CEO problem [6]. The rate region for this special case is studied in [7, 8, 9]. However, the rate region is not yet found for many other cases because there are a lot of mathematical difficulties

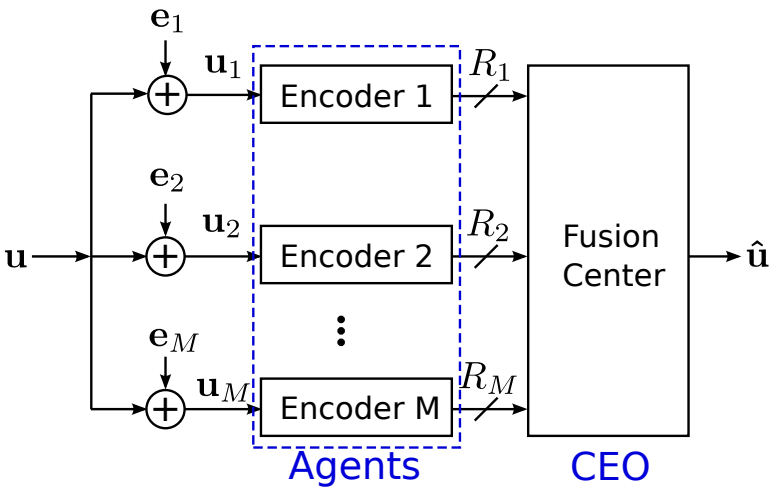

Figure 1.1: An abstract model of the CEO problem.

in analyzing the limit. Nowadays, the CEO problem attracts a lot of attention with the recognition of not only an open problem of information theory, but also significant importance on its applications to many forms of wireless communication systems, such as distributed sensor networks.

In this thesis, we propose several practical coding and decoding strategies to a parallel WSN and a WMN where none of the *last* forwarding nodes has error-free information part. In this case, a WMN can also formulated from the viewpoint of the CEO problem. Even though the theoretical bound of these application is not yet analyzed in this thesis, it is found that the our proposed technique can achieve better bit-error-rate (BER) performance than some other techniques proposed by other research groups.

1.3 Summary of Achievements and Publications

The results of a simplified wireless mesh network (WMN) which reformulated by the CEO problem has been accepted in 2012 International Symposium on Information Theory and its Applications (ISITA2012), Hawaii [10]. In this paper, we analyzed the BER performances and provide a short discussion on rate optimization.

Then, we applied the coding-decoding strategy in WSNs. The results have been accepted in IEICE Transaction on communication [11]. This paper is a joint work with a PH.D candidate Mr. Xiaobo Zhou. The part of that multiple sensors observe a binary Markov source is contributed by him.

Furthermore, a paper on how to estimate observation error probabilities in a WSN

is submitted to IEEE Communications Letters. In this paper, we propose an iterative algorithm for the purpose of estimating observation error probabilities.

The publications are listed below:

Conference Paper:

- X. He, X. Zhou, K. Anwar and T. Matsumoto, "Wireless Mesh Networks Allowing Intra -Link Errors: CEO Problem Viewpoint", 2012 International Symposium on Information Theory and its Applications (ISITA), Hawaii, October, 2012

Journal Papers:

- X. Zhou, X. He, K. Anwar, and T. Matsumoto, "GREAT-CEO: larGe scale distRibuted dEcision mAking Technique for wireless Chief Executive Officer problems", IEICE Transaction on Communications, vol. E95–B, no. 12, pp. 3654–3662, December 2012.

- X. He, X. Zhou, K. Anwar and T. Matsumoto, "Estimation of Observation Error Probability in Wireless Sensor Networks", (Submitted for publication-under review), IEEE Communications Letters

1.4 Organization of the Thesis

This thesis is organized as follows. In **Chapter 2**, we discuss the network topologies from lossless to lossy for introducing our footprint from Slepian-Wolf relaying system to the wireless communication system modeled by the CEO problem. **Chapter 3** focuses on a parallel WSN from the viewpoint of the CEO problem. **Chapter 4** introduces the coding-decoding strategies as well as the performance evaluation of a simple WMN. The conclusion and future works are summarized in **Chapter 5**.

Chapter 2

Network Topologies: Lossless to Lossy

2.1 Lossless: Slepian-Wolf Relaying System

According to the Slepian-Wolf's correlated source coding theorem [4], if the source information is compressed at a coding rate less than the entropy of each source but larger than the conditional entropy, conditioned by the other source, and if the destination knows the correlation property, still the source information can be recovered without any distortion at the destination; furthermore, this region specifies the "result" of coding, which means that the two decoders do not have to negotiate each other. The authors realized [2, 12, 13, 14] that this theorem can well be utilized in the relaying systems, because, for example in one-way relay system, the data is originated from the one single source, and hence the signal directly transmitted and the signal forwarded by the relay are correlated, even though the originated data is corrupted by errors in the source-relay link (*intra-link*). Assuming that the error occurrence property of the intra-link can be represented by a the bit-flipping model [15], which is equivalent to the binary symmetry channel (BSC) model, and if the flipping probabilities are know to the destination, iterative decoding process can well utilize the correlation property in the form of log-likelihood ratio (LLR) updating function [12, 13].

With the Slepian-Wolf theorem based relaying system, the relay (or forwarding node) does not have to perfectly correct the *intra-link* errors; it can simply extract information part from the received frame, interleave, re-encode, and then forward it to the destination. The extracted data may contain errors, but by iterative processing with the LLR updating function, the originators data can be fully recovered. This scheme is referred to

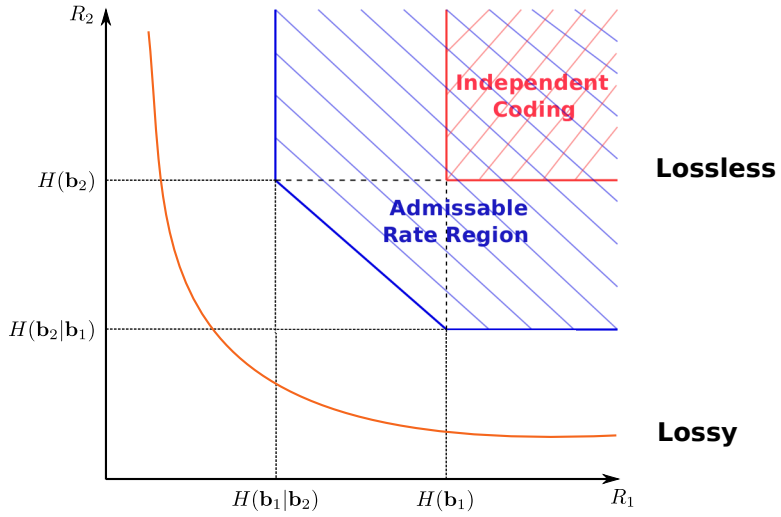

Figure 2.1: The admissible rate region of Slepian-Wolf theorem and the rate region bound for lossy coding cases. $\mathbf{b}_1$ and $\mathbf{b}_2$ represent two correlated sources.

as extract-and-forward (ErF) technique in this thesis. Ref. [2] applies the ErF concept to relay systems with the aim of exploiting source-relay correlation. Reference [14] combines the ErF concept with bit-interleaved coded modulation with iterative detection (BICM-ID). Reference [15] proposes a very simple serially concatenated coding scheme that can achieve near Slepian-Wolf/Shannon limit [4] performance, even though the structure is very simple. Reference [16] evaluates the outage probability of the ErF relay system in Rayleigh fading channels. The impacts of not only the source correlation but also the correlation of the link variation are theoretically analyzed in [16] .

As shown in Fig. 2.1, the admissible rate region is constituted as an unbounded polygon based on the Slepian-Wolf theorem. The source information can be recovered only when the compressed rate pair falls into this area. For instance, we consider the case that two binary information sequence $\mathbf{b}_1$ and $\mathbf{b}_2$ are separately encoded and jointly decoded. If $\mathbf{b}_1$ is compressed at the rate R_1 which is equal to its entropy $H(\mathbf{b}_1)$, then $\mathbf{b}_2$ can be compressed at the rate R_2 which is less than its entropy $H(\mathbf{b}_2)$, but must be greater than its conditional entropy $H(\mathbf{b}_2|\mathbf{b}_1)$, or vice versa. The admissible rate region of the Slepian-Wolf compression is given by the following three inequalities [4]:

$$R_1 \geq H(\mathbf{b}_1|\mathbf{b}_2), \tag{2.1}$$

$$R_2 \geq H(\mathbf{b}_2|\mathbf{b}_1), \tag{2.2}$$

$$R_1 + R_2 \geq H(\mathbf{b}_1, \mathbf{b}_2). \tag{2.3}$$

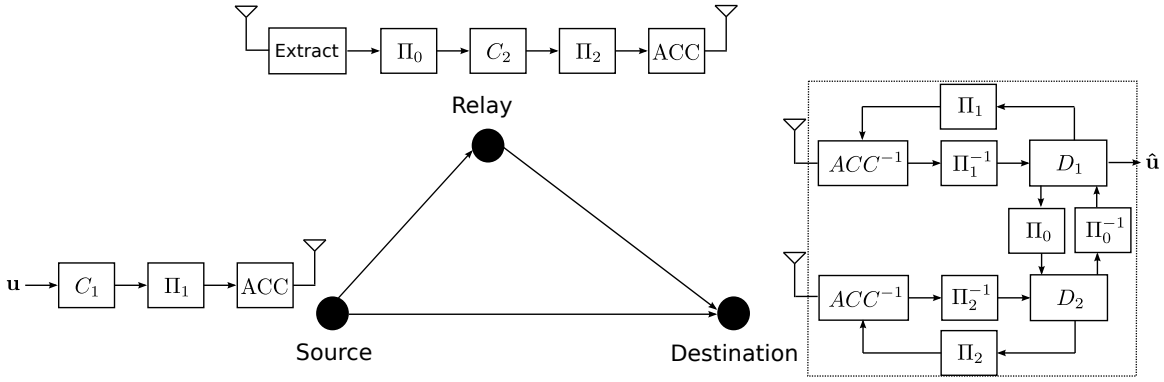

Figure 2.2: Extract-and-Forward Relay System. D_1 and D_2 denote the decoders of the channel codes C_1 and C_2 used by the source and the relay, respectively. ACC and ACC^{-1} are the accumulator and de-accumulator, respectively.

By exploiting the correlation knowledge of the data streams at the destination, the distributed source coding can achieve the same compression rate as the optimum single encoder which compresses the sources jointly.

If the correlation model of the sources can be expressed as the bit-flipping mode [17], i.e., $\mathbf{b}_2 = \mathbf{b}_1 \oplus \mathbf{e}$ and $\Pr(\mathbf{e} = 1) = p_e$, where p_e is the bit-flipping probability and $\oplus$ denotes modulo-2 addition, the Slepian-Wolf theorem can be utilized in relaying system. Assume that the appearance probabilities of the source information is equiprobable. Then, $H(\mathbf{b}_1) = H(\mathbf{b}_2) = 1$, $H(\mathbf{b}_1|\mathbf{b}_2) = H(\mathbf{b}_2|\mathbf{b}_1) = H(p_e)$, $H(\mathbf{b}_1, \mathbf{b}_2) = 1 + H(p_e)$ with $H(p_e) = -p_e \log_2(p_e) - (1-p_e)\log_2(1-p_e)$. Now, let us consider a one-way relaying system, where the relay does not aim to perfectly recover the original information transmitted by the source, but it only "extract"[1] the source information, even though the relay knows that the extracted sequence may contain some errors. As shown in Fig. 2.2, the extracted sequence representing an estimate of the original information sequence, which is then interleaved and transmitted to the common destination. Obviously, the original and extracted sequences are correlated. If we assume block fading and no heavy decoding of the channel code which quite likely causes long burst errors is performed at the relay, it is reasonable that the *intra-link* can be expressed by the bit-flipping model within the block. Based on [2, 12, 13, 14], excellent performance can be achieved through the LLR exchange in the vertical iterations between the two decoders, where to take into account the correlation, represented by the error probability of the *intra-link*.

[1]Full iterative decoding is not performed at the relay with the aim of reducing the computational burden at the relay. Instead, the relay performs Viterbi algorithm or only one iteration to get a tentative estimate of the source information sequence. This process is referred as "extract" according to [2].

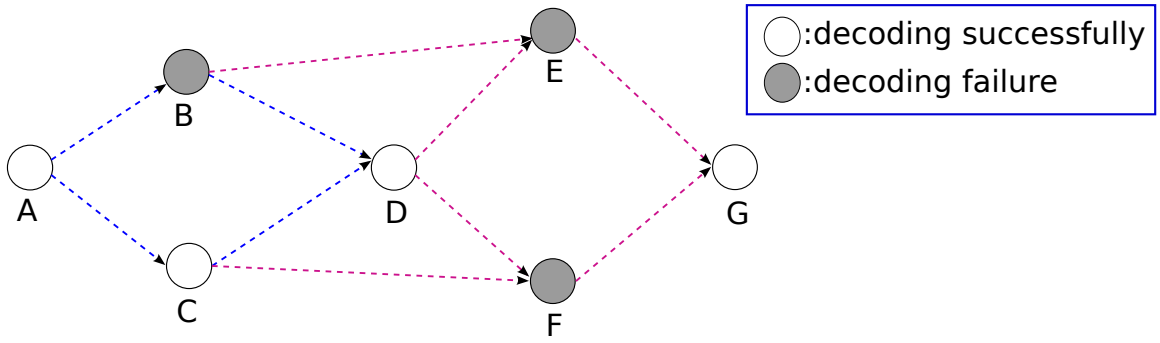

Figure 2.3: An topology of a WMN combined with Slepian-Wolf and the CEO problem.

2.2 Lossy: the CEO Problem

As mentioned in the motivation part of **Chapter 1**, the Slepian-Wolf relaying system can achieve such excellent performance is because that the correlation is exploited at the destination. We assumed that the information sequence does not contain any errors before being encoded at the source node for the Slepian-Wolf relaying system. It hence belongs to lossless case. However, if we apply the solutions of Slepian-Wolf relaying system to more complicated networks, such as wireless mesh networks/wireless sensor networks, it is quite likely that none of the forwarding nodes/sensors has error-free information part, nevertheless, they performs re-encoding of the error-corrupted versions of the information part, and forward it to the final destination (FD)/FC.

Fig. 2.3 illustrates a simple topology of a WMN, where the transmission chain $A \rightarrow \{B, C\} \rightarrow D$ belongs to Slepian-Wolf problem and $D \rightarrow \{E, F\} \rightarrow G$ typically belongs to the CEO problem. In reality, it is very common that the Slepian-Wolf and the CEO problem simultaneously exist in a large scale distributed wireless networks. Hence, researching on the CEO problem can provide appropriate frameworks for the applications with complexity-and/or-latency constraints, such as sensor networks and video/multimedia compression, since the CEO problem can shift the computational complexity from encoder side to the joint decoder side.

The rate region bound of lossy case where satisfying the rate-distortion pairs (R_1, D_1) and (R_2, D_2) are the major target is also included in Fig. 2.1. The inner bound and outer bound of the admissible rate-distortion region is derived in [9]. The sum-rate distortion function for the quadratic Gaussian CEO problem with infinite agents and the identical SNRs is provided by Oohama in [7]. In fact, the rate-distortion region for many other cases are still not yet known. It could be indeed great contributions with significant importance that if we can find precise rate-distortion region completely for the CEO problem, since

the CEO problem is related to many other distributed source coding problems.

It should be noticed that the theoretical limit of the only a specific case is known, and hence, for the systems investigated in this thesis, deriving the theoretical limit is still an open question. Therefore, in many of the performance figures, we do not include the limit lines.

Chapter 3

Wireless Sensor Networks

WSNs which composed of a large number of sensors, deployed in a geographic area to perform distributed tasks, have been recognized as an important technology, since WSNs have significant impacts on the society. In WSNs, each sensor node is required to work under very low power consumption restrictions when it performs specified tasks. Therefore, the sensor nodes have to transmit their data without requiring high transmission power, and the WSN itself has to be highly energy efficient. A technique which is effective in reducing the power consumption is to design the system so that the FC can well exploit the correlation knowledge among the observed data, as supported by the Slepian-Wolf theorem. In fact, the Slepian-Wolf theorem can be used to compress the data or to reduce the transmission power, assuming the source-channel separation [15]. Therefore, in this chapter, no specific source encoding technique is assumed.

Distributed source coding schemes for sensor networks based on the Slepian-Wolf theorem are investigated in a tutorial article [18]. In practice, most of WSNs aim to observe the same sensing object, e.g., the sensors monitor the same physical phenomenon. In network information theory, the problem of estimating the correct data emitted from the target over noisy observation link can be modeled by the CEO problem.

In this chapter, proposed coding and decoding strategies, and corresponding performance for a parallel WSN are discussed.

3.1 System Model

Figure 3.1 depicts the model of the parallel WSN system, investigated in this chapter, from the viewpoint of the CEO problem. The sensing object $\mathbf{u}$ is an independent binary sequence with $\Pr(\mathbf{u}(t) = 0) = \Pr(\mathbf{u}(t) = 1) = 1/2$. A set of sensors $\mathbf{S}$ produces the error-

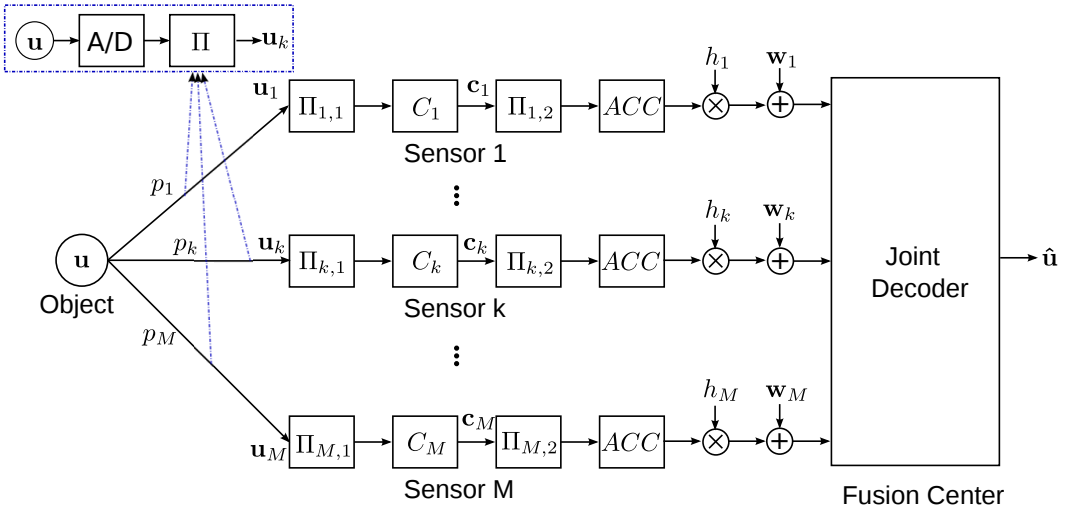

Figure 3.1: The structure of the proposed system where M sensors independently observe the same sensing object.

corrupted versions of the binary sequence $\mathbf{u}_k$, $k = 1, \cdots, M$, obtained after the interleaver Π following the analog-to-digital (A/D) convertor, corresponding to the observed samples generated by the sensing object. The observation made by each sensor is modeled by a bit-flipping model with flipping probabilities p_k for the sensor k, $k = 1, \cdots, M$, therefore, the bit sequence $\mathbf{u}_k$ can be expressed as

$$\mathbf{u}_k(t) = \mathbf{u}(t) \oplus \mathbf{e}_k(t), \tag{3.1}$$

where $\mathbf{e}_k$ is the binary error sequence with probability of $\Pr(\mathbf{e}_k(t) = 1) = p_k$. Let $\mathbf{P} = [p_1, p_2, \cdots, p_M]^{\mathrm{T}}$ denote the vector sorting the observation error probabilities. p_k denotes the correlation parameter between sensing object and observation made by the k-th sensor, where $0 \leq p_k \leq 1/2$.

The sensor k interleaves the observed bit sequence $\mathbf{u}_k$ first, using interleaver $\Pi_{k,1}$, and then encode the interleaved bit sequence $\Pi_{k,1}(\mathbf{u}_k)$ with a channel encoder C_k. The encoded bit sequence $\mathbf{c}_k$ is again interleaved by the interleaver $\Pi_{k,2}$ and doped-accumulated by doped accumulator (ACC) [2] with doping ratio P_d for $k = 1, \cdots, M$. It should be noted that the lengths of $\Pi_{k,1}$ should not necessarily be the same as that of Π. Finally, the doped-accumulated bit sequence is modulated by binary-phased-shift-keying (BPSK) and transmitted to the FC over independent additive white Gaussian noise (AWGN) channels or Rayleigh fading channels. As shown in Fig. 3.1, the signal received from the k-th sensor

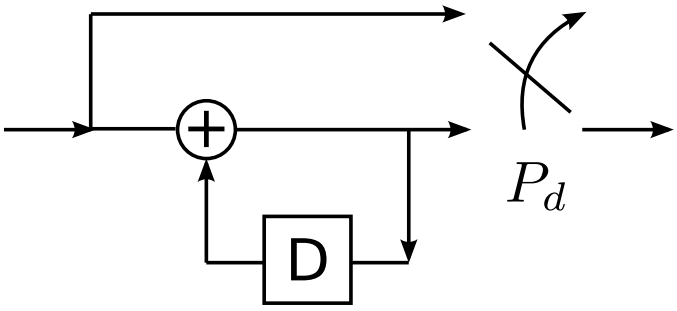

Figure 3.2: The structure of the doped accumulator (ACC), P_d denotes the doping ratio.

can be written as

$$\mathbf{y}_k = h_k \cdot \mathbf{s}_k + \mathbf{w}_k, \tag{3.2}$$

where, h_k represents the channel coefficient. The channel is static channel if h_k is a constant, while it is a block fading channel if h_k is variable over blocks-by-blocks. In the latter one, $h_k \in \mathbb{C}$ follows complex Gaussian distribution. The BPSK modulated symbol sequence at the sensor k is denoted by $\mathbf{s}_k$. $\mathbf{w}_k$ is the Gaussian noise sequence which follows $\mathcal{N}(0, \sigma^2)$, where σ^2 is the noise variance per dimension.

3.1.1 Bit-flipping Model

The bit-flipping model is widely used to exploit the correlation in cooperative communications [2, 3, 17, 14]. In a parallel WSN, the k-th sensor's observation results are A/D converted with m-bit resolution, interleaved by interleaver Π, and then transmitted to the FC by using BPSK. It should be emphasized that the same Π is commonly used by the M sensors, and the size is equivalent to the $m \times K$, where K is the number the samples. It plays a crucial role in making the length mK bit sequence random so that the observation error can well be represented by the random bit-flipping model after the interleaver Π.

The implementation of bit-flipping model can directly use modulo-2 addition operation. We first randomly generate a bit sequence $\mathbf{e}$ with the property that $\Pr(\mathbf{e}(t) = 1) = p_e$, where p_e is the error probability, then modulo-2 add with the bit sequence to be bit-flipped, which is shown in (3.1).

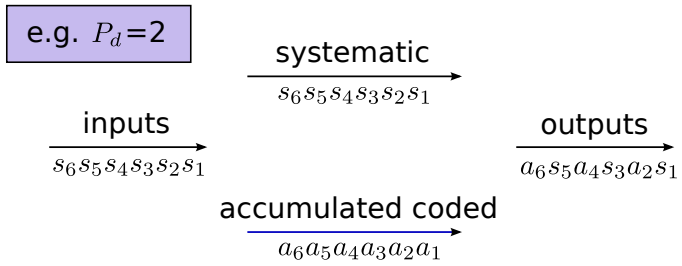

Figure 3.3: An example of ACC coding scheme with the doping ratio$P_d = 2$.

3.1.2 Doped Accumulator (ACC)

The ACC uses a memory-1 systematic recursive convolutional code structure, every P_d-th of the systematic bit is superseded by the accumulated-coded bit which is shown in Fig. 3.2. An example of how ACC works is given in Fig. 3.3.

The ACC does not change the coding rate of the system. The purpose of using the ACC is to push up the EXIT curve of inner code decoder to reach a point close enough to (1.0, 1.0) mutual information (MI) point, therefore, BER performance can be improved through the iteration process.

3.2 Decoding Algorithm

3.2.1 Basics

The LLR of a binary sequence $\mathbf{u}$ is defined as follows:

$$\mathbf{L_u} = \ln(\frac{\Pr(\mathbf{u} = +1)}{\Pr(\mathbf{u} = -1)}), \tag{3.3}$$

where, $\Pr(\mathbf{u} = +1)$ and $\Pr(\mathbf{u} = -1)$ represent the probabilities of the corresponding information bit being 0 and 1, respectively. In the decoding process, the LLR is also called soft bit, since the hard decision can be made based on the sign of the LLR value and the reliability of this hard decision largely depends on the magnitude of the LLR value. The probability of $\mathbf{u} = +1$ and $\mathbf{u} = -1$ can be straightforward calculated from the given LLR

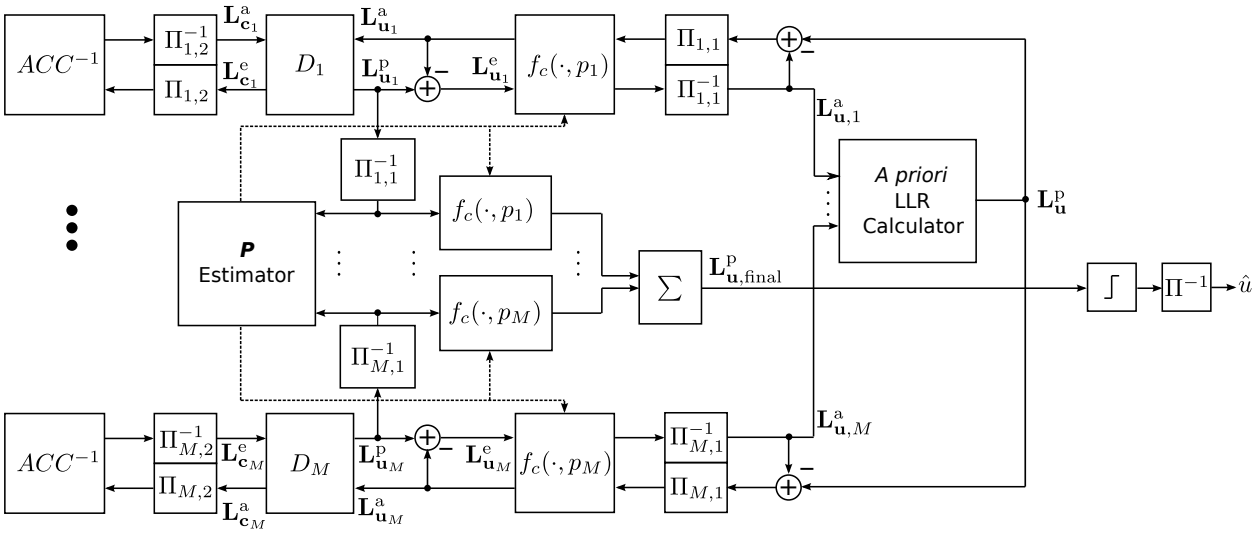

Figure 3.4: Proposed decoding strategy for a parallel sensor network.

value by:

$$\Pr(\mathbf{u} = +1) = \frac{\exp(\mathbf{L}_u)}{1 + \exp(\mathbf{L}_u)}$$
$$\Pr(\mathbf{u} = -1) = \frac{1}{1 + \exp(\mathbf{L}_u)}. \tag{3.4}$$

The *a priori* LLR $\mathbf{L}_{\mathbf{u}}^{\mathrm{a}}$ of bit sequence $\mathbf{u}$ expresses certain amount of knowledge on the bit sequence $\mathbf{u}$ which has already been known before decoding. The *a priori* LLR is usually provided by other decoders.

3.2.2 Algorithm Implementation

A block diagram of the proposed decoding algorithm is shown in Fig. 3.4. It includes local iteration (LI), which performs the extrinsic LLR exchange between the decoder ACC^{-1} of the doped-accumulator and the channel code decoders D_k, and the global iteration (GI), which performs the extrinsic LLR exchange among decoders D_k. The aim of performing GI is to utilize the correlation knowledge among the sensors through the LLR updating function f_c [15, 17], as shown in Fig. 3.4. As we can see the BER performance curves presented in **Section 3.4 "Performance Evaluation"**, the effect of performing GI is significant, hence to achieve such large gain through GI, we need to estimate the observation error probabilities. The Maximum A Posteriori (MAP) algorithm is used in ACC^{-1} and D_k.

The *LIs* are first performed for each sensor's corresponding D_k, and then *a posteriori* LLR $\mathbf{L}^{\mathrm{p}}_{\mathbf{u}_k}$ of systematic bits (information part corrupted by the observation error) output from D_k are fed into the $\mathbf{P}$ estimator to obtain the observation error probabilities p_k which can be utilized in $f_c[\cdot, p_k]$. The algorithm used in the $\mathbf{P}$ estimator is detailed in the next section. Then, the extrinsic LLR $\mathbf{L}^{\mathrm{e}}_{\mathbf{u},k}$ are input into *a priori* LLR calculator after performing f_c, p_k and de-interleaved. The *a priori* LLR calculator obtains $\mathbf{L}^{\mathrm{p}}_{\mathbf{u}}$ as:

$$\mathbf{L}^{\mathrm{p}}_{\mathbf{u}} = \sum_{k=1}^{M} \mathbf{L}^{\mathrm{a}}_{\mathbf{u},k}. \tag{3.5}$$

As indicated in Fig. 3.4, the extrinsic LLR $f_c[\mathbf{L}^{\mathrm{e}}_{\mathbf{u}_k}, p_k]$ is equivalent to *a priori* LLR $\mathbf{L}^{\mathrm{a}}_{\mathbf{u},k}$ of the *a priori* LLR calculator. Hence, $\mathbf{L}^{\mathrm{a}}_{\mathbf{u},k}$ has to be subtracted from $\mathbf{L}^{\mathrm{p}}_{\mathbf{u}}$. The interleaved version of $\mathbf{L}^{\mathrm{p}}_{\mathbf{u}} - \mathbf{L}^{\mathrm{a}}_{\mathbf{u},k}$ is input to f_c, p_k as $f_c[\Pi_{k,1}(\mathbf{L}^{\mathrm{p}}_{\mathbf{u}} - \mathbf{L}^{\mathrm{a}}_{\mathbf{u},k}), p_k]$, and then its output is fed back to D_k as the *a priori* LLR, $\mathbf{L}^{\mathrm{a}}_{\mathbf{u}_k} = f_c[\Pi_{k,1}(\mathbf{L}^{\mathrm{p}}_{\mathbf{u}} - \mathbf{L}^{\mathrm{a}}_{\mathbf{u},k}), p_k]$.

The *LI* and *GI* are performed until no more relevant gain can be achieved in *a posteriori* LLR $\mathbf{L}^{\mathrm{p}}_{\mathbf{u},\mathrm{final}}$. Then hard decisions are made based on $\mathbf{L}^{\mathrm{p}}_{\mathbf{u},\mathrm{final}}$ given by:

$$\mathbf{L}^{\mathrm{p}}_{\mathbf{u},\mathrm{final}} = \sum_{k=1}^{M} f_c[\Pi^{-1}_{k,1}(\mathbf{L}^{\mathrm{p}}_{\mathbf{u}_k}), p_k]. \tag{3.6}$$

It should be noticed that the proposed decoding technique is equivalent to performing the LLR updating by f_c-function between arbitrary M pairs of the sensors, however, the computational complexity for decoding is reduced from a combinatorial order $\binom{M}{2}$ to a linear order M. The scalability of the proposed system model is hence guaranteed.

3.2.3 LLR Updating Function

As described above, the observed binary sequence at each sensor node is the corrupted version of $\mathbf{u}$, hence it is straightforward according to [17] that we can obtain the following equation:

$$\left.\begin{aligned} \Pr(\mathbf{u}_k = 0) &= (1 - p_k)\Pr(\mathbf{u} = 0) + p_k \Pr(\mathbf{u} = 1) \\ \Pr(\mathbf{u}_k = 1) &= (1 - p_k)\Pr(\mathbf{u} = 1) + p_k \Pr(\mathbf{u} = 0) \end{aligned}\right\}. \tag{3.7}$$

Equation (3.7) is equivalent to the LLR updating function $f_c[\cdot, p_k]$ shown in (3.8), based

on the value p_k and the extrinsic LLRs of the uncoded bits obtained as the results of D_k.

$$\begin{aligned}\mathbf{L}^{\mathrm{a}}_{\mathbf{u},1} &= \Pi^{-1}_{k,1}[f_c(\mathbf{L}^{\mathrm{e}}_{\mathbf{u}_1}, p_k)] \\ &= \Pi^{-1}_{k,1}\left[\ln \frac{(1-p_k)\cdot \exp(\mathbf{L}^{\mathrm{e}}_{\mathbf{u}_1}) + p_k}{(1-p_k) + p_k \cdot \exp(\mathbf{L}^{\mathrm{e}}_{\mathbf{u}_1})}\right].\end{aligned} \tag{3.8}$$

The LLR updating function for other sensors can be derived in the similar way as shown above.

3.2.4 Correlation between Pairs of Sensors

Since the $\mathbf{P}$ estimator should know the correlation between each pair of the sensors, we can get the following equations for an example of sensor 1 and sensor 2 according to the fact that they are observing the same sensing object but suffered from errors at different positions.

$$\left.\begin{aligned}\Pr(\mathbf{u}_1 = 0) &= (1-p_1)\Pr(\mathbf{u} = 0) + p_1 \Pr(\mathbf{u} = 1) \\ \Pr(\mathbf{u}_1 = 1) &= (1-p_1)\Pr(\mathbf{u} = 1) + p_1 \Pr(\mathbf{u} = 0)\end{aligned}\right\} \tag{3.9}$$

$$\left.\begin{aligned}\Pr(\mathbf{u} = 0) &= (1-p_2)\Pr(\mathbf{u}_2 = 0) + p_2 \Pr(\mathbf{u}_2 = 1) \\ \Pr(\mathbf{u} = 1) &= (1-p_2)\Pr(\mathbf{u}_2 = 1) + p_2 \Pr(\mathbf{u}_2 = 0)\end{aligned}\right\} \tag{3.10}$$

Substituting (3.10) into (3.9), we can derive the correlation equation between sensor 1 and sensor 2, as:

$$\left.\begin{aligned}\Pr(\mathbf{u}_1 = 0) &= (1-q_{12})\Pr(\mathbf{u}_2 = 0) + q_{12} \Pr(\mathbf{u}_2 = 1) \\ \Pr(\mathbf{u}_1 = 1) &= (1-q_{12})\Pr(\mathbf{u}_2 = 1) + q_{12} \Pr(\mathbf{u}_2 = 0)\end{aligned}\right\}, \tag{3.11}$$

where $q_{12} = p_1 + p_2 - 2p_1p_2$.

The correlation equation between arbitrary pairs of the sensors can be straightforward derived as (3.11) of sensor 1 and sensor 2. Based on these correlation equations between pairs of the sensors, $\mathbf{P}$ can be easily estimated, which will be presented in the next section.

3.3 Error Probability Estimation Algorithm

Because of the fact that all the observations made by the sensors are correlated, the following pair-wise equations[1] hold according to (3.11):

$$\begin{aligned} & \hat{p}_i + \hat{p}_j - 2 \cdot \hat{p}_i \cdot \hat{p}_j = \hat{q}_{ij}, \\ \text{where,} \quad & i = 1 \cdots M, \\ & j = i + 1 \quad \text{if} \quad i = 1 \cdots M - 1, \\ & j = 1 \quad \text{if} \quad i = M, \end{aligned} \tag{3.12}$$

where following [15],

$$\hat{q}_{ij} = \frac{1}{N} \sum_{g=1}^{N} \frac{\exp(L^{\mathrm{p}}_{\mathbf{u}_i,g}) + \exp(L^{\mathrm{p}}_{\mathbf{u}_j,g})}{[1 + \exp(L^{\mathrm{p}}_{\mathbf{u}_i,g})] \cdot [1 + \exp(L^{\mathrm{p}}_{\mathbf{u}_j,g})]}, \tag{3.13}$$

with $L^{\mathrm{p}}_{\mathbf{u}_i,g} \in \mathbf{L}^{\mathrm{p}}_{\mathbf{u}_i}$, N represents the number of the *a posteriori* LLR pairs from the two decoders with their absolute values larger than a given threshold T. Since the reliability of $\hat{q}_{ij}$ is influenced by N, it is very important to choose an appropriate T value. However, how to determine the optimal T is out of the scope of this thesis.

We can reformulate (3.12) by introducing the identity matrix $\mathbf{I}$ of size M and a matrix $\mathbf{J}$ defined by (3.15), into the following form:

$$[(\mathbf{I} + \mathbf{J}) - 2 \cdot \mathrm{diag}(\hat{\mathbf{P}}) \cdot \mathbf{J}] \cdot \hat{\mathbf{P}} = \hat{\mathbf{q}}, \tag{3.14}$$

where, $\hat{\mathbf{P}} = [\hat{p}_1, \hat{p}_2, \cdots, \hat{p}_M]^T$, and $\hat{\mathbf{q}} = [\hat{q}_{12}, \hat{q}_{23}, \cdots, \hat{q}_{M1}]^T$. The $\mathrm{diag}(\cdot)$ is the operator that forms a diagonal matrix from its argument vector, and $\mathbf{J}$ is denoted as follows:

$$\mathbf{J} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots 0 \\ 0 & 0 & 1 & 0 & \cdots 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & 0 & \cdots 0 \end{bmatrix}. \tag{3.15}$$

Now, our objective is to find a nonnegative vector $\hat{\mathbf{P}}$ that minimizes $\|\mathbf{A}\hat{\mathbf{P}} - \hat{\mathbf{q}}\|^2$, which

[1] $\hat{q}_{ij}$ can be understood as the bit error probability of the $j(i)$-th sensor's link, assuming that the $i(j)$-th sensor's link is error free. Furthermore, j should not necessarily be $i + 1$. In this case, according to the selected pairs, the form the matrix $\mathbf{J}$ changes.

is formulated as follows:

$$\begin{aligned} \min \quad & \|\mathbf{A}\hat{\mathbf{P}} - \hat{\mathbf{q}}\|^2 \\ \text{s.t} \quad & \hat{\mathbf{P}} \geq 0, \end{aligned} \tag{3.16}$$

where, $\mathbf{A} = [(\mathbf{I} + \mathbf{J}) - 2 \cdot \text{diag}(\hat{\mathbf{P}}) \cdot \mathbf{J}]$.

To solve (3.16), we propose an iterative algorithm summarized in **Algorithm** 1. In this algorithm, we use the standard Nonnegative Least Squares (*lsqnonneg*) described in [19], proposed by Lawson and Hanson. The algorithm of *lsqnonneg* is detailed in **Appendix A**.

3.4 Performance Evaluation

3.4.1 Convergence Property of the Estimation

Fig. 3.5 shows the mean square estimation error of the observation error probability vector $\mathbf{P}$ versus the *GI* times where one *GI* was followed by one *LI*. The code parameters are the same as shown in Table 3.1. The results are plotted for sensor number 12 with per-link SNR and LLR threshold T as parameters.

As shown in Fig. 3.5, the mean square estimation error $|\hat{\mathbf{P}} - \mathbf{P}|^2$ decreases as iteration times increased, indicating that the more *GIs* performed, the more accurate the estimate of $\mathbf{P}$. Furthermore, the rate of convergence depends on the values of per-link SNR and T, because the values of per-link SNR and T affect the reliability of $\hat{q}_{ij}$ given by (3.13). It is interesting that with relatively low per-link SNR, the algorithm converges quickly with small T, e.g., $T = 1.5$, however, with relatively high per-link SNR, large T, e.g., $T = 2$ can lead to fast convergence.

3.4.2 BER Performance Comparison of identical P

The gain of performing *GI* to utilize the correlation knowledge over the case where *GI* is not involved in the proposed technique is evaluated, where the simulation results are shown in Fig. 3.6. It can be achieved roughly $5-8$ dB that the gain of performing *GI* depending on the number of sensors. Hence, it is necessary to estimate the error probability $\mathbf{P}$. The simulation parameters are shown in Table 3.1. Furthermore, our proposed technique can achieve better performances than the parallel concatenated convolutional codes scheme proposed in [20].

Algorithm 1: P Estimator

Input: $\mathbf{b}$, ϵ, Pre-defined maximum iterations IT_m
Output: $\hat{\mathbf{P}} \geq 0$ such that $\hat{\mathbf{P}}$=arg min $\|\mathbf{A}\hat{\mathbf{P}} - \hat{\mathbf{q}}\|^2$
Initialization: $\hat{\mathbf{P}}^{(0)} = \mathbf{0}$, Calculate $\mathbf{A}$ and $\Delta(0) = \|\mathbf{A}\hat{\mathbf{P}}^{(0)} - \hat{\mathbf{q}}\|^2$
for $l = 1$ **to** IT_m **do**
  Calculate $\hat{\mathbf{P}}^{(l)}$ by using *lsqnonneg* algorithm;
  Update $\mathbf{A} = [(\mathbf{I} + \mathbf{J}) - 2 \cdot \mathrm{diag}(\hat{\mathbf{P}}^{(l)}) \cdot \mathbf{J}]$;
  $\Delta(l) = \|\mathbf{A}\hat{\mathbf{P}}^{(l)} - \hat{\mathbf{q}}\|^2$;
  if $\Delta(l) \geq \Delta(l-1)$ **then**
    Stop;
  end
  if $\|\hat{\mathbf{P}}^{(l)} - \hat{\mathbf{P}}^{(l-1)}\|^2 \leq \epsilon$ **then**
    Stop;
  end
  $\hat{\mathbf{P}} = \hat{\mathbf{P}}^{(l)}$;
end

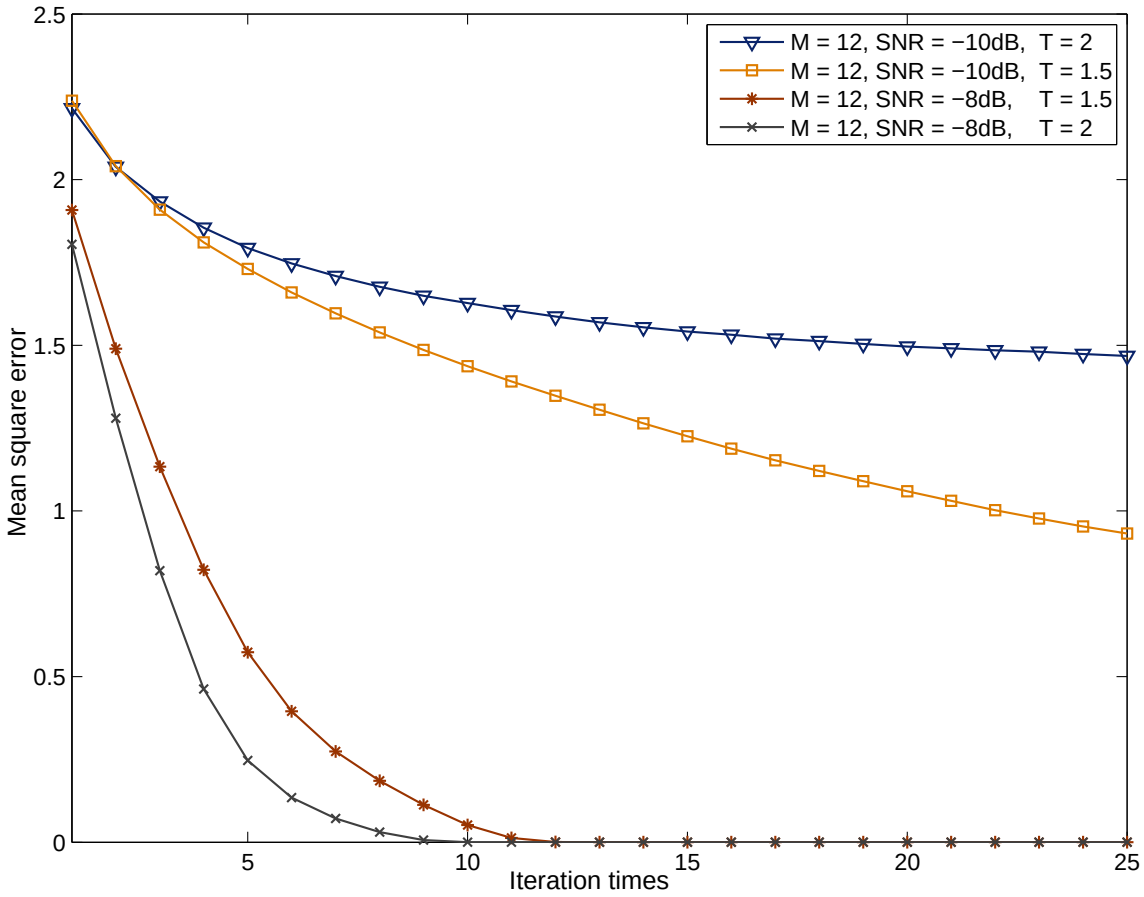

Figure 3.5: Mean square errors of estimation versus decoding iteration times.

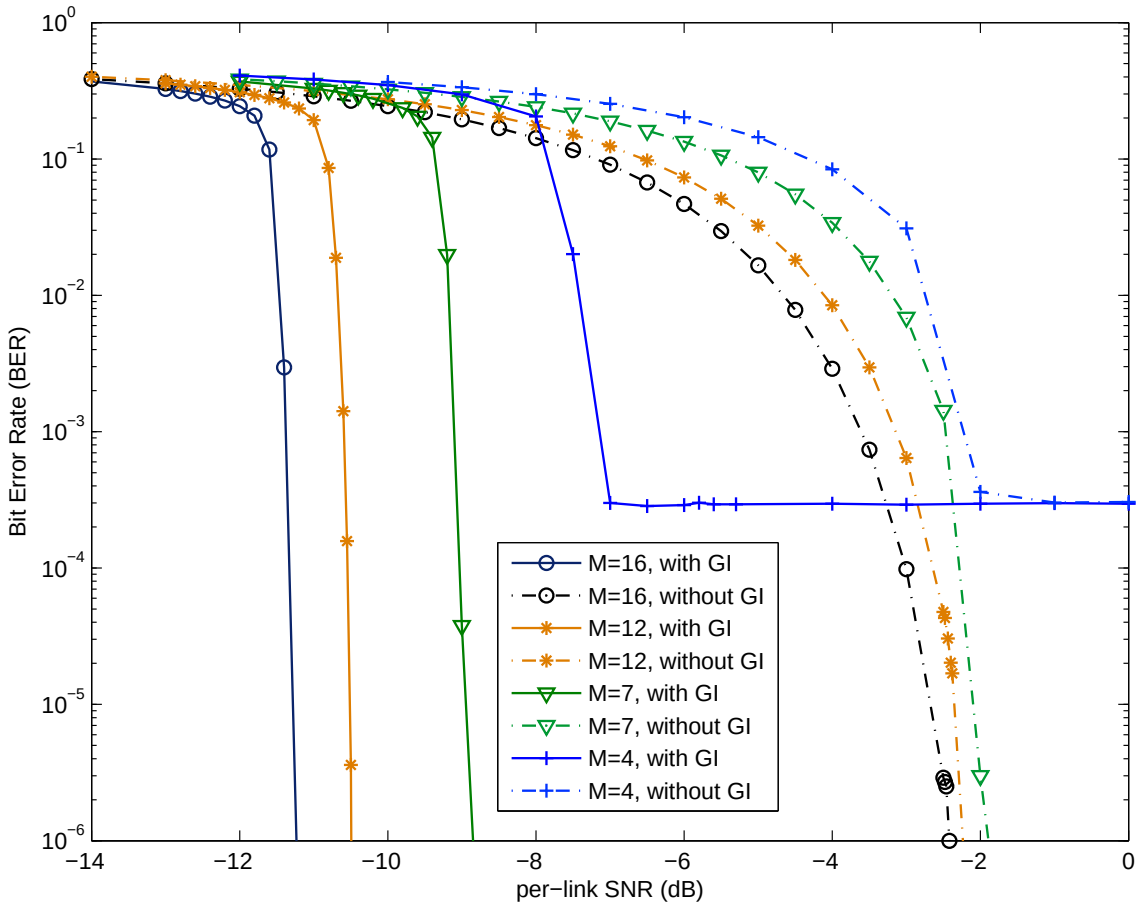

Figure 3.6: The gain of performing GI in BER performances over the case where GI is not performed with M as a parameter: solid line is the performance of utilizing the correlation knowledge, while the dash line is that of not utilizing the correlation knowledge.

Item	Setting
$\mathbf{u}$	$\Pr(\mathbf{u}=0)=\Pr(\mathbf{u}=1)=1/2$
C_k	Rate 1/2, $G=[3,2]_8$, memory-1 nonrecursive systematic convolutional code
Block length N	10^4 bits
Block	1000
Decoding Algorithm	BCJR
GI	25
P_d	1
T	2
ϵ	10^{-6}
IT_m	20

Table 3.1: The settings of simulation parameters.

The BER performances are evaluated for the cases where $\hat{\mathbf{P}}$ and $\mathbf{P}$ are used at the FC, with the numbers M of sensors as a parameter. The simulation results are shown in Fig. 3.7 . The other transmission parameters are the same as for Fig. 3.6. In the observation error probability estimator, T, ϵ and the maximum iteration times, IT_m were set as in the Table 3.1.

It is found from the figure that the BER performance can be improved by increasing the number of sensors M. With $M=4$, the error floor can not be reduced to less than 10^{-4} even by increasing per-link SNR, however it can be reduced to less than 10^{-6} with $M \geq 7$. Nevertheless, we believe that it is impossible to totally eliminate the error floor, even though it may happen at a very small BER region. The reason is because we can not completely eliminate the distortion due to the observation error, which is common to the CEO problems.

In the case all the elements of $\mathbf{P}$ have identical value p, the error floor ϖ can be calculated by (3.17):

If M is odd:

$$\varpi = \binom{M}{M}p^M + \binom{M}{M-1}p^{M-1}(1-p) + \cdots + \binom{M}{\frac{M+1}{2}}p^{\frac{M+1}{2}}(1-p)^{\frac{M-1}{2}}$$

If M is even:

$$\varpi = \binom{M}{M}p^M + \binom{M}{M-1}p^{M-1}(1-p) + \cdots + \binom{M}{\frac{M}{2}}p^{\frac{M}{2}}(1-p)^{\frac{M}{2}}/2. \tag{3.17}$$

Furthermore, (3.17) can be approximated by (3.18) if $\mathbf{P}$ is small enough, e.g., 0.01. The

M	ϖ
2	1.00×10^{-2}
3	2.97×10^{-4}
4	2.97×10^{-4}
5	9.80×10^{-6}
6	9.80×10^{-6}
7	3.39×10^{-7}
8	3.39×10^{-7}
9	1.21×10^{-8}
16	4.20×10^{-12}

Table 3.2: The error floor ϖ of different number of sensors.

error floor ϖ taking the number of sensors M as the parameter is given in Table 3.2.

$$\begin{aligned} &\text{If } M \text{ is odd:} \\ &\varpi \approx \binom{M}{\frac{M+1}{2}} p^{\frac{M+1}{2}} (1-p)^{\frac{M-1}{2}} \\ &\text{If } M \text{ is even:} \\ &\varpi \approx \binom{M}{\frac{M}{2}} p^{\frac{M}{2}} (1-p)^{\frac{M}{2}} / 2. \end{aligned} \tag{3.18}$$

Compared with the case where $\mathbf{P}$ is known, only $0.3 - 0.5$ dB loss in per-link SNR is observed when using estimate $\hat{\mathbf{P}}$ of $\mathbf{P}$, and the loss depends on the number of M of the sensors.

3.4.3 Impact of P Variation

The results shown above are assuming that all the p_k are equal to 0.01, we now consider the cases where p_k are also random variables. Fig. 3.8 shows the BER performances in the case of 7 good observations and 1 bad observation, as well as the converse case. In the simulations, the other parameters are kept the same as in Table 3.1. We use $p_k = 0.001$ for the good observation and 0.1 for the bad observation. Even in these two extreme cases, we can still estimate $\mathbf{P}$ to get close performance to the case of using exact $\mathbf{P}$ at the FC.

The BER performance of the proposed technique with the error probability vector $\mathbf{P}$ being uniformly distributed in the open interval $(0, 0.1)$ is shown in Fig. 3.9. Based on the obtained result, our proposed estimation technique can achieve very close performance

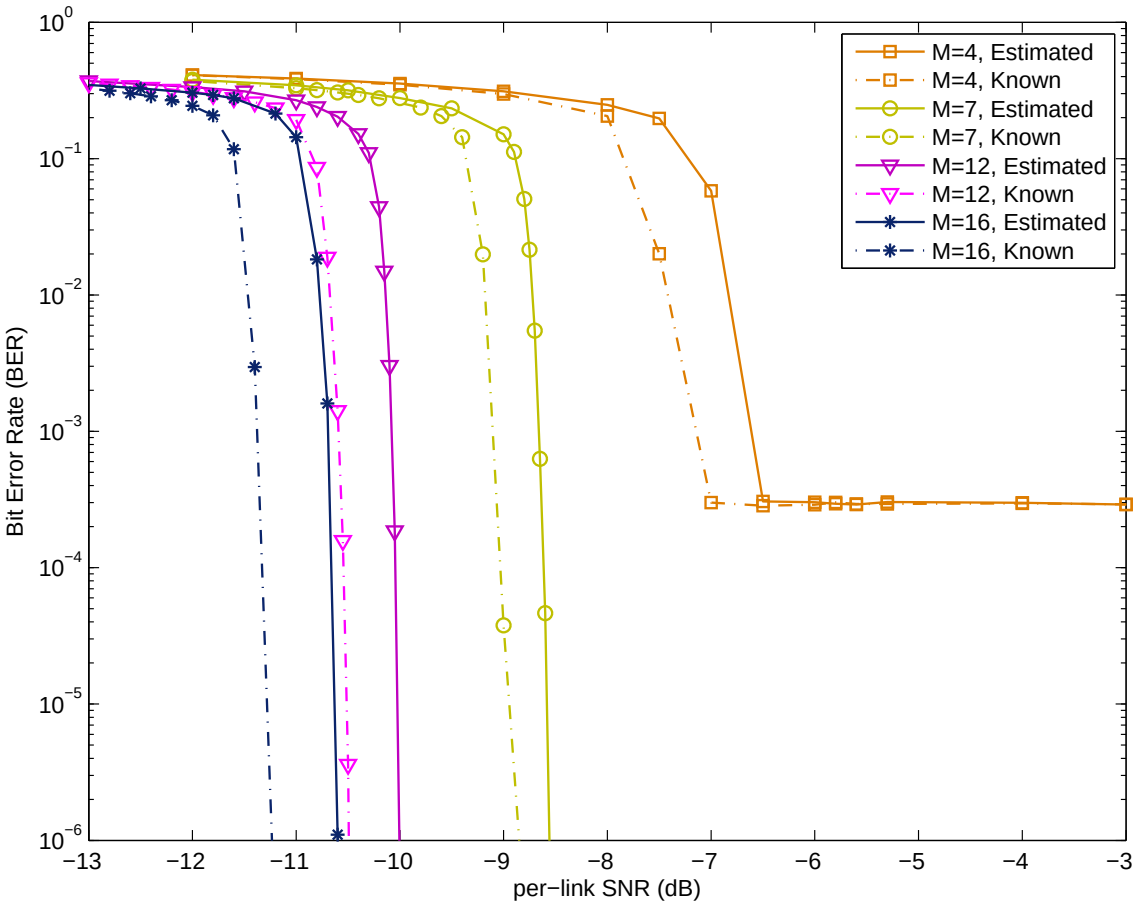

Figure 3.7: BER performance comparison for different numbers of sensors with $p_k = 0.01$, $k = 1, \cdots, M$.

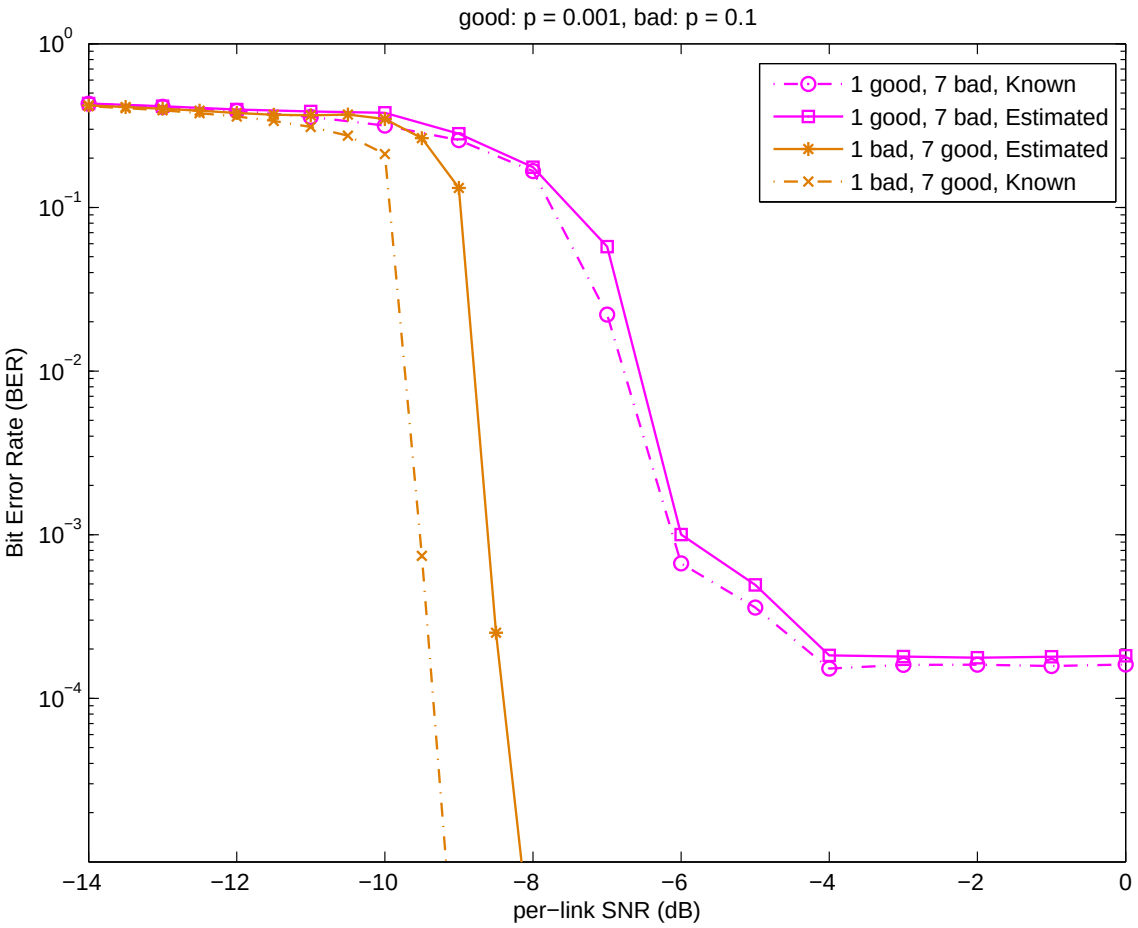

Figure 3.8: BER performance comparison for various $\mathbf{P}$.

compared to the case $\mathbf{P}$ is known to the FC.

As the results shown above, we can apply the proposed estimation technique to the practical environment, where the observation error probability varies. The loss in BER performance is spread slightly according to the value of each p_k, where $k = 1, \cdots, M$.

3.4.4 FER Performance in Block Fading Sensor-FC Links

The BER and FER performance for the case where sensors-FC links suffer from Rayleigh fading are shown in Fig. 3.10 and Fig. 3.11, respectively. In this simulation, we assume fast block fading, where fading is static within a block, but changes block-by-block. Based on the simulation results, we can still achieve around 4 dB performance gain in per-link average SNR by using the proposed technique over the case of not utilizing the correlation knowledge. Furthermore, the performance of using estimated $\hat{\mathbf{P}}$ can achieve a close performance in both BER and FER.

In Rayleigh fading channels, we improve the estimation algorithm by utilizing the adaptive threshold T algorithm when estimating $\hat{q}_{ij}$ based on the *a posteriori* LLRs from decoder D_i and D_j. The adaptive threshold T algorithm[2] is summarized as follows:

Algorithm 2: Adaptive Threshold T

Input: Initial T, Block Length: N, *a posteriori* LLR: $\mathbf{L}^{\mathrm{P}}$, Step: $\triangle T$
Output: Updated T
while length($|\mathbf{L}^{\mathrm{P}}| \geq T$) $< N \cdot \alpha\%$ **do**
  $T = T - \triangle T$;
end

The average SNR shown in the FER and BER performance figures is defined as:

$$\gamma_k = |h_k|^2 \cdot \Gamma_k, \tag{3.19}$$

with the constraint:

$$E\{|h_k|^2\} = 1, \tag{3.20}$$

where γ_k and Γ_k represent instantaneous SNR and the average SNR of the channel between the k-th sensor and FC, respectively. The channel coefficient of each channel is defined

[2] $\alpha = 75$ was used in the simulations, and the value was found empirically.

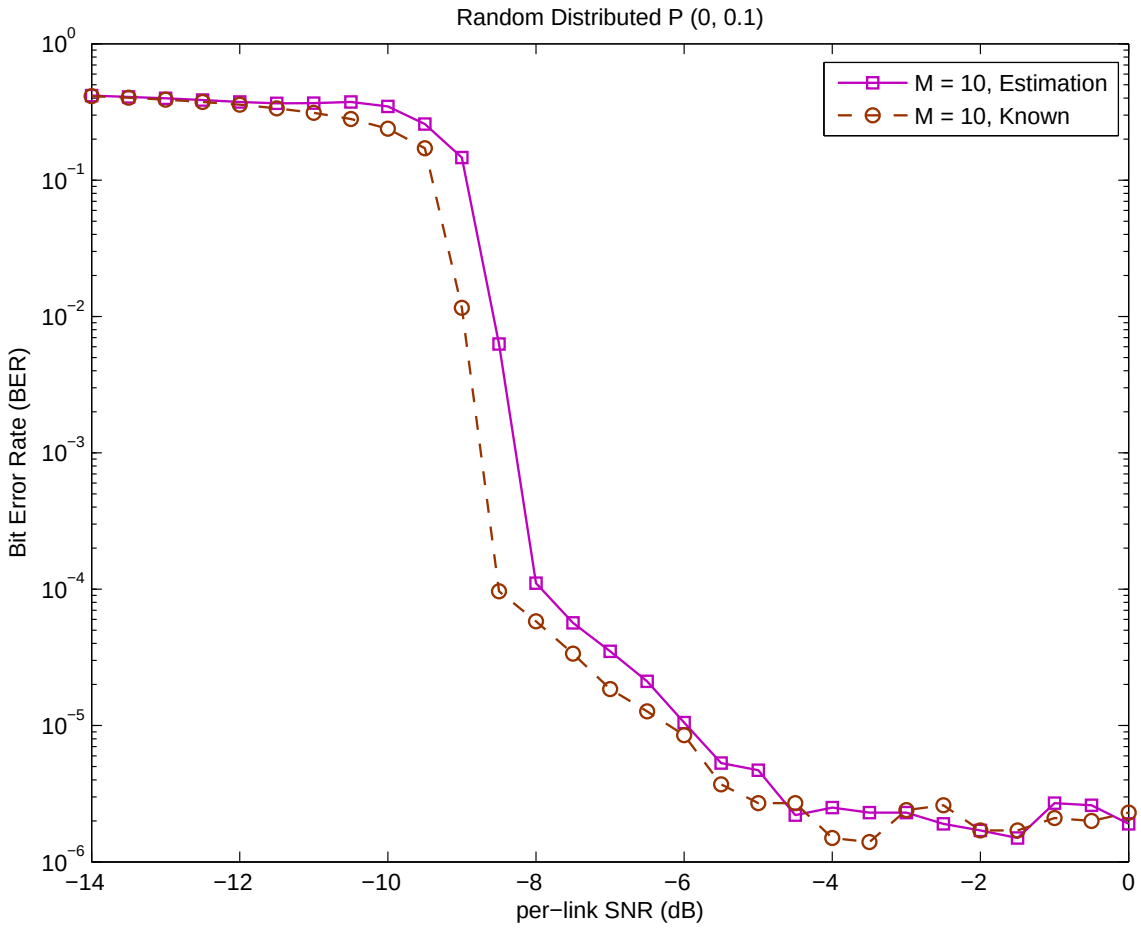

Figure 3.9: BER performance comparison for $\mathbf{P}$ following uniform distribution.

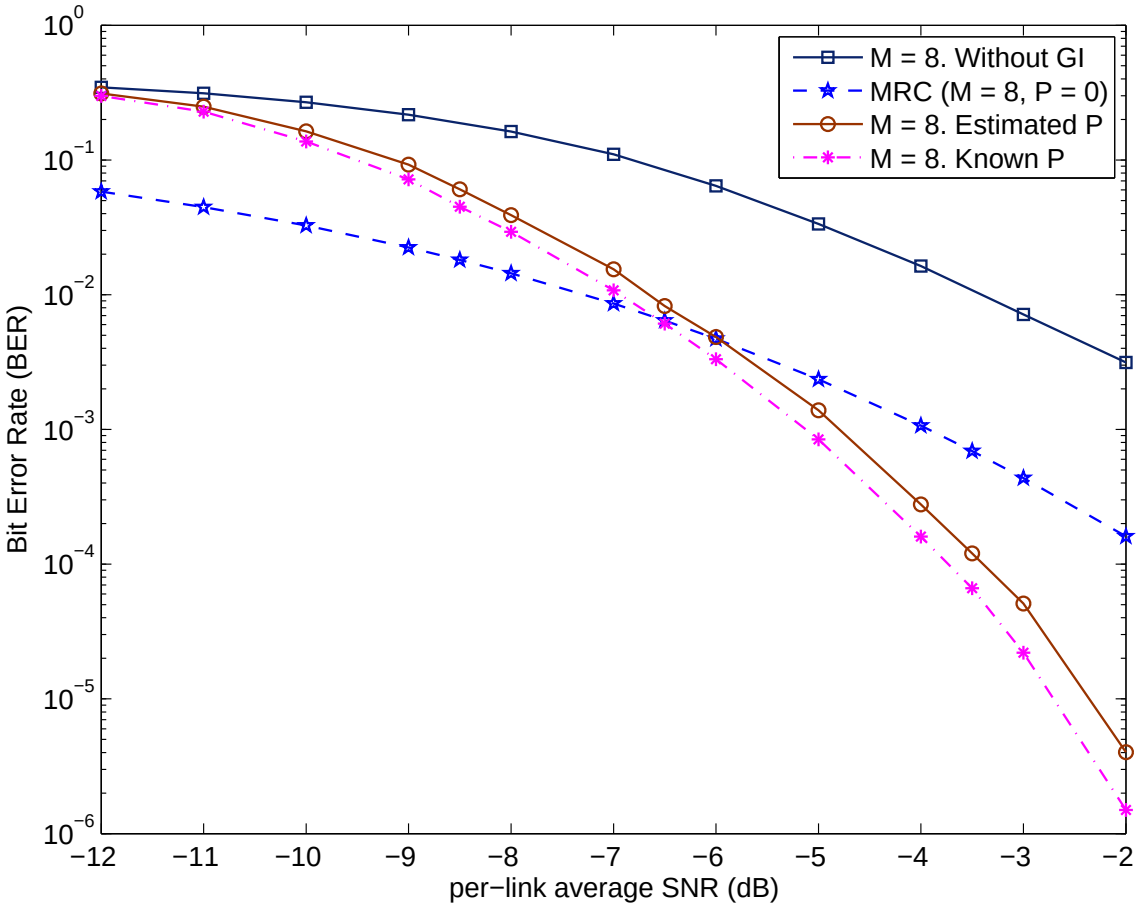

Figure 3.10: BER performance comparison for assuming that the instantaneous SNR changes link-by-link according to equation (3.19). Links are assumed to be Rayleigh fading channels. Number of sensor is 8. The observation error probabilities are set to be 0.01.

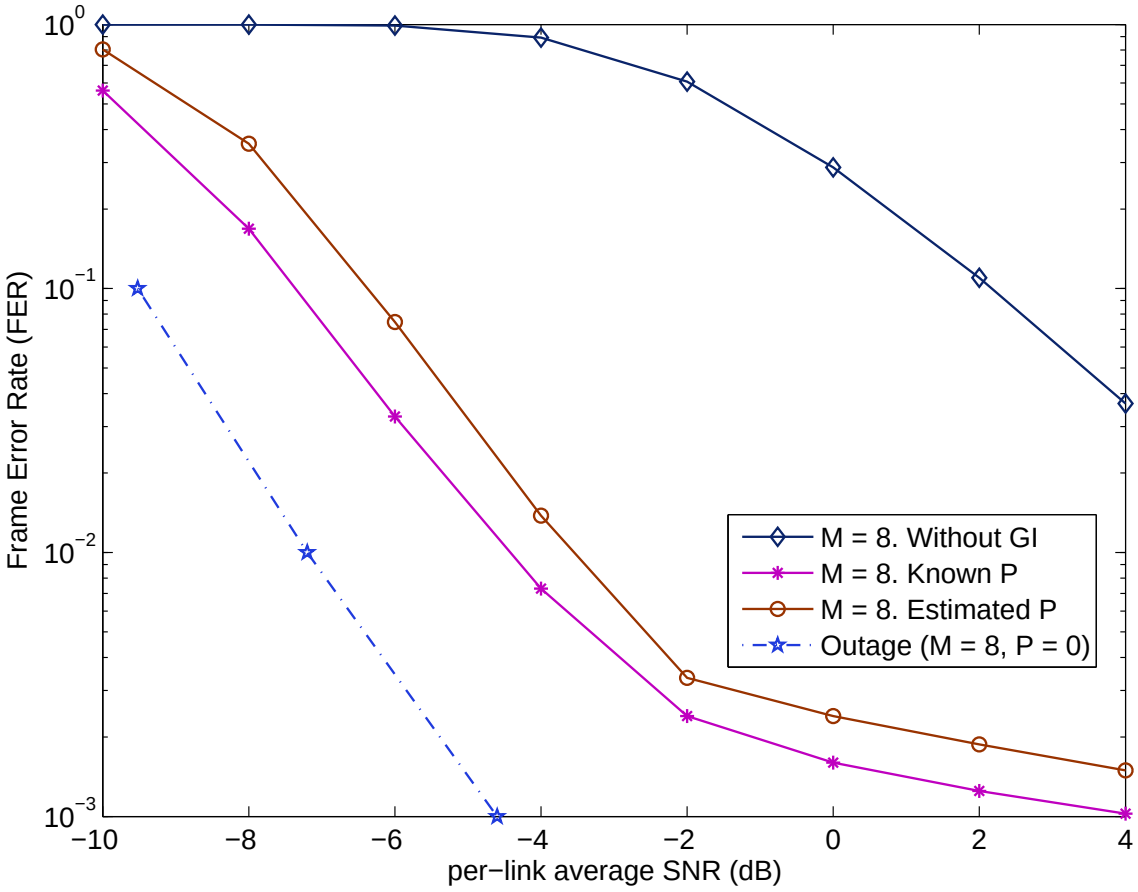

Figure 3.11: FER performance on the case of assuming the channels between sensors and FC are Rayleigh fading channels. Number of sensor is 8. The observation error probabilities are set to be 0.01.

as h_k. The probability density function (pdf) of γ_k is given by:

$$\Pr(\gamma_k) = \frac{1}{\Gamma_k} \exp(-\frac{\gamma_k}{\Gamma_k}). \tag{3.21}$$

3.5 Summary

In this chapter, we have investigated the transmission techniques of data gathered by multiple sensors to the fusion center.

We first proposed a strategy that combines the local and the global iterations without requiring heavy computational complexity, instead of exploiting the correlation knowledge over all the possible combinations of the sensor pairs. We also proposed a nonnegative constrained iterative algorithm to estimate the observation error probabilities.

It has been shown through simulations that the algorithm converges only after several iterations. In addition, the results of simulations conducted to evaluate the BER performance of the proposed techniques with the estimation algorithm can achieve only roughly $0.3 - 0.5$ dB loss in per-link SNR compared to the case of known $\mathbf{P}$ at the FC. Finally, the FER performance of the proposed techniques, where the links between the sensors and the FC are suffering from Rayleigh fading, has been evaluated.

Chapter 4

Wireless Mesh Networks

WMNs are a form of cooperative communications system, in which multiple nodes are organized in a mesh topology for relaying messages to the destinations. High data throughput, power and spectral efficiencies, as well as better resource utilization are expected. WMN systems usually consist of a group of fixed or mobile devices and hence can be deployed smoothly and flexibly in complicated environments such as in devastated and emergency situations, tunnels, oil rigs, and/or for battlefield surveillance.

In this chapter, we investigate a simple case of WMN, which is two forwarding nodes assist the originator to transmit information to FD. The schematic diagram of this simple WMN is depicted in Fig. 4.1. It should be noticed that the forwarding nodes is equivalent to the sensors in **Chapter 3**, and the originator is equivalent to the sensing object. However, since the originator can be a mobile device in a WMN, it is possible to add another outer code to remove the errors occurring in the originator-forwarding node *link*, referred to as *intra-link*.

4.1 System Model

Figure 4.2 shows a simple model of WMN, describing the viewpoint of the CEO problem, with only two forwarding nodes, assumed in this thesis. This strategy can be further extended to the case of involving more forwarding nodes, where the decoding algorithm of a parallel WSN can be used.

At the originator node, the original information bit sequence $\mathbf{x}$ to be transmitted is first encoded by C_0. The encoded bit sequence is then interleaved by random interleaver Π_0 and doped-accumulated by ACC with doping ratio P_{ori}. The output of ACC, $\mathbf{u}$, is broadcasted to the two forwarding nodes over independent BSC with crossover probabilities p_1 and

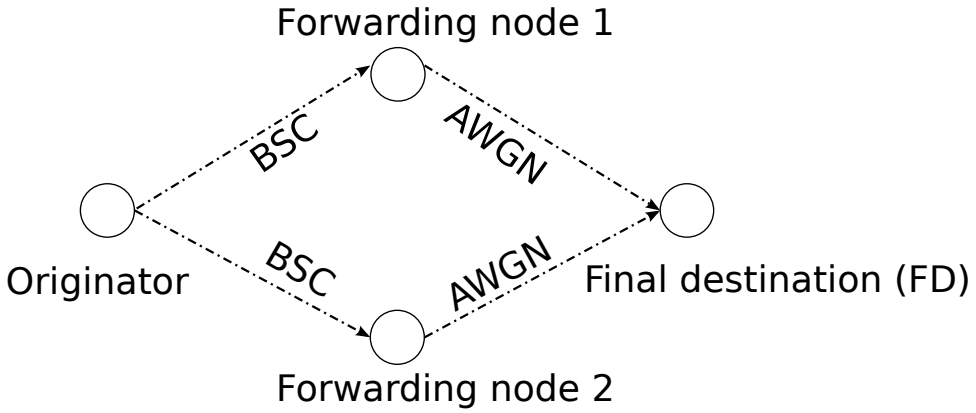

Figure 4.1: The schematic diagram of a very simple WMN, where two forwarding nodes assist the originator to communicate with the final destination.

p_2, respectively, which can be modeled by the bit-flipping model.

Aiming at perfect decoding at the forwarding nodes is out of the scope of this thesis, because it needs very strong *link-level* codes. Instead, each forwarding node makes only tentative decision on the received bit sequences, of which results $\mathbf{u}_1$ and $\mathbf{u}_2$ are first permutated by the interleavers Π_1 and Π_2 and further encoded by C_1 and C_2, respectively. The channel between the transmitted information sequence $\mathbf{u}$ and the one $\mathbf{u}_k$ obtained as the result of the tentative decoding can be seen as also BSCs with the crossover probabilities p_1 and p_2, which can be modeled by the bit-flipping model, where $k = 1, 2$ is the forwarding node index.

Since perfect recovery of information sequence transmitted from the originator node is not aimed at in the forwarding nodes, the complexity of forwarding nodes is very light. Then, the encoded sequences are again interleaved by Π_3 and Π_4, and doped-accumulated by ACC with doping ratio P_{for}. The doped-accumulated bit sequences are modulated by BPSK, i.e., $0 \rightarrow -1$ and $1 \rightarrow +1$, and then forwarded to the FD at different time slots over AWGN channels. We assume the SNRs are the same in the two channels between forwarding nodes and the FD. The received signal sequence from forwarding node k can be expressed as:

$$\mathbf{y}_k = \mathbf{s}_k + \mathbf{n}_k, \tag{4.1}$$

where the modulated signal sequence is denoted by $\mathbf{s}_k$. $\mathbf{n}_k$ represents noise that are zero mean i.i.d complex value with variance σ^2 per dimension.

After receiving the signal $\mathbf{y}$, the channel log-likelihood ratios (LLRs) are first calculated

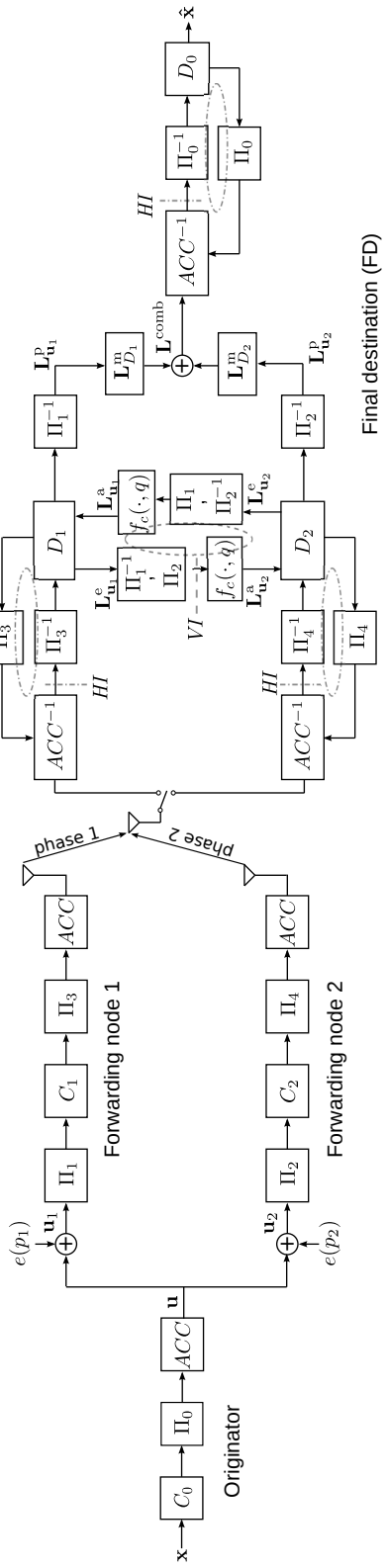

Figure 4.2: Structure of the proposed system for a simplified WMN. Only two forwarding nodes are considered.

by:

$$\mathbf{L}^{\text{ch}} = \ln \frac{\Pr(\mathbf{y}|\mathbf{s}=+1)}{\Pr(\mathbf{y}|\mathbf{s}=-1)} = \frac{2}{\sigma^2}\mathbf{y}. \tag{4.2}$$

At the FD, joint decoding is performed by exchanging the extrinsic LLRs which is detailed in the next section.

4.2 Joint Decoding Strategy

Joint decoding process is divided into two iteration processes as depicted in Fig. 4.2. We refer these two processes to Horizontal Iteration (*HI*) and Vertical Iteration (*VI*). It is necessary to mention that *HI* and *VI* is equivalent to *LI* and *GI* in a WSN, respectively. To perform the channel decoding for convolutional codes C_0, C_1 and C_2 as well as for ACC, we still perform MAP decoding using the BCJR algorithm proposed by Bahl, Cocke, Jelinek and Raviv [21] as in **Chapter 3**.

In the *HI*, the extrinsic LLRs are exchanged through the corresponding interleaver/de-interleaver between the soft-in-soft-out (SISO) decoder ACC^{-1} and SISO decoder D_1 or D_2 used by the forwarding node 1 and 2, respectively. The extrinsic information exchange is performed via *HI* until no more significant mutual information improvement can be achieved. However, activation times on the two *HI* loops are design parameters and hence optimization of the activation ordering is out of the scope of this paper. After each *HI* step, we activate *VI* loop between D_1 and D_2 by exchanging the output extrinsic LLRs of uncoded (systematic) bits output from the two decoders D_1 and D_2 via an LLR updating function $f_c(\mathbf{L}^{\text{e}}_{\mathbf{u}_k}, q)$ as:

$$f_c(\mathbf{L}^{\text{e}}_{\mathbf{u}_k}, q) = \ln \frac{(1-q)\cdot \exp(\mathbf{L}^{\text{e}}_{\mathbf{u}_k}) + q}{(1-q) + q\cdot \exp(\mathbf{L}^{\text{e}}_{\mathbf{u}_k})}, \tag{4.3}$$

with $q = p_1 + p_2 - 2p_1p_2$. The purpose of *VI* is to help two decoders D_1 and D_2 cooperate each other to reconstruct information. This is because since the uncoded bit sequences, are originated from the common originator node and forwarded by the two forwarding nodes, they are correlated. Hence, the aim is to fully exploit the knowledge about the correlation at the FD node.

After performing LLR exchange several times via the *HI*-*VI* loops, the *a posteriori* LLRs output from D_1 and D_2 are combined. Before combing, however, the LLRs are

further modified by (4.4) [22, 23].

$$\mathbf{L}^{\text{comb}} = \sum_{k=1}^{2} \mathbf{L}^{\text{m}}_{D_k} = \sum_{k=1}^{2} \ln \frac{1-p_k}{p_k} \text{sign}(\mathbf{L}^{\text{p}}_{\mathbf{u}_k}). \tag{4.4}$$

p_k represents the error probabilities of the *intra-link*. $\mathbf{L}^{\text{p}}_{\mathbf{u}_k}$ is the *a posteriori* LLR from D_k, where $k = 1, 2$ in this simple case. The function sign$(\cdot)$ takes the sign (positive or negative) of its argument.

The modified LLRs $\mathbf{L}^{\text{comb}}$ are forwarded to another horizontal iteration loop to finally obtain the originator's source information bit sequence $\hat{\mathbf{x}}$. This process is the same as the *HI* described above. Finally, hard decisions are made on the *a posteriori* LLRs obtained by the decoder D_0.

4.3 EXIT Chart Analysis

4.3.1 Basics

Extrinsic information transfer (EXIT) chart was introduced by Stephan ten Brink in [24] as a novel tool to aid the construction, analysis and optimization of concatenated error-correcting coding schemes with iterative decoding process. As a result, the convergence behavior of iterative decoding schemes can be easily understand in the EXIT chart.

As mentioned in [25], the BER performance of iterative decoding schemes contains three regions: 1) in the region of low SNR, BER is very high with negligible gain over iterations; 2) the turbo cliff happens as the SNR increased, where the error rate reduces very fast by performing enough iterations. This region is also called waterfall region; 3) the error floor appears in the region of high SNR where a quite low BER can be achieved by performing only several iterations. The error floor can be eliminated by designing the iterative systems based on EXIT chart analysis. However, as described in **Chapter 3**, the distortion caused in the CEO system results the error floor; the reason for the error floor is the CEO problem is different from the floor of the turbo system; the distortion is already included before encoding in the case of the CEO problem, and hence, it can not be predicted by the EXIT chart analysis.

In order to analyze the convergence behavior of iterative decoding at the turbo cliff region in the SNR range, a density evolution algorithm has been proposed to calculate convergence threshold for low-density parity-check (LDPC) codes over the AWGN channel and to construct LDPC capacity-approaching codes [26, 27]. The main idea of density

evolution is to track the pdfs of the exchanged information message in the iterative decoding process. The pdfs of LLRs can be assumed as being Gaussian distributed to simplify the analysis [28]. 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. Based on the EXIT chart analysis, the convergence property of the iterative systems can be visualized, and with the aid of the visualization technique, the code optimization falls into the problem of the EXIT curve matching. In [29], the authors proposed an EXIT-constrained binary switching algorithm to optimize the coding parameters of single parity check code and irregular repetition code and the mapping rule of extend mapping to achieve a very close Shannon limit performance.

Entropy and Mutual Information

In the information theory, the entropy is originally defined by Claude E. Shannon in [30], which is the measure of the uncertainty of information. It is a general concept to evaluate the amount of information in a transmitted message. The definition of the entropy is given, for a discrete random variable having probability mass function $\Pr(x)$, as:

$$H(X) = -\sum_{x \in X} \Pr(x) \log(\Pr(x)), \tag{4.5}$$

where, $H(X)$ represents the entropy of a discrete random variable X. $H(x)$ does not take negative values, with the definition $0 \log 0 = 0$.

The joint entropy $H(X, Y)$ of discrete random variables X and Y can be defined as:

$$H(X, Y) = -\sum_{x \in X} \sum_{y \in Y} \Pr(x, y) \log(\Pr(x, y)), \tag{4.6}$$

where x and y are particular values of X and Y, respectively. $\Pr(x, y)$ is the joint probability that the event $X = x$ and $Y = y$ happens. The conditional entropy of Y by given

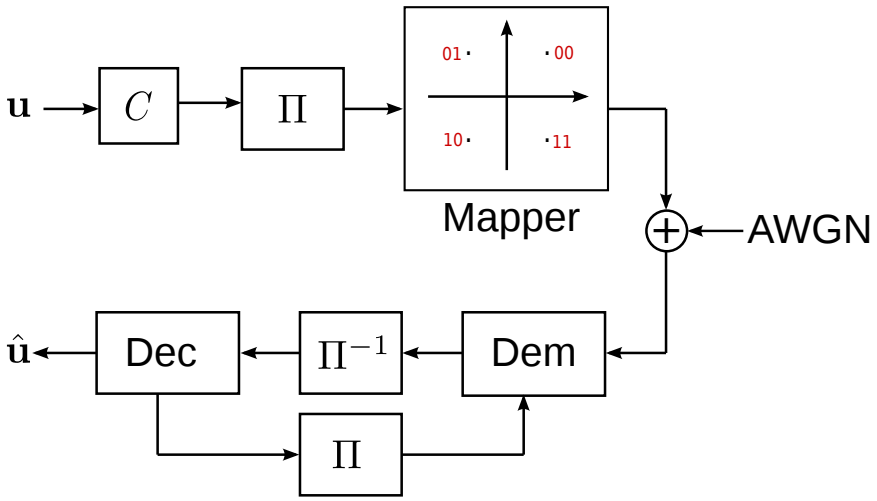

Figure 4.3: An example of a simple communication system for learning the concept of the EXIT chart. Dec: channel decoder, Dem: demapper.

X can be further defined as:

$$\begin{aligned} H(Y|X) &= \sum_{x\in X} \Pr(x) H(Y|X=x) \\ &= \sum_{x\in X} \Pr(x) \sum_{y\in Y} \Pr(y|x) \log\left(\frac{1}{\Pr(y|x)}\right) \\ &= -\sum_{x\in X}\sum_{y\in Y} \Pr(x,y) \log(\Pr(y|x)) \\ &= -\sum_{x\in X, y\in Y} \Pr(x,y) \log(\Pr(y|x)) \\ &= \sum_{x\in X, y\in Y} \Pr(x,y) \log\left(\frac{\Pr(x)}{\Pr(x,y)}\right), \end{aligned} \tag{4.7}$$

where $H(Y|X) = 0$ if and only if the value of Y is completely determined by the value of X. Conversely, $H(Y|X) = H(Y)$ if and only if Y and X are independent random variables.

The mutual information between X and Y defined in (4.8) is the relative entropy between the joint distribution and the product distribution $\Pr(x)\Pr(y)$. The mutual information indicates a quantity that measures the mutual dependence of the two random

variables.

$$I(X;Y) = \sum_{x \in X} \sum_{y \in Y} \Pr(x,y) \log \left(\frac{\Pr(x,y)}{\Pr(x) \Pr(y)} \right), \tag{4.8}$$

where $\Pr(x,y)$ is the joint distribution function of X and Y. $\Pr(x)$ and $\Pr(y)$ are the marginal probability distribution functions of X and Y, respectively, i.e., $\Pr(x) = \sum_y \Pr(x,y)$ and $\Pr(y) = \sum_x \Pr(x,y)$

The mutual information $I(X;Y)$ is the reduction in the uncertainty of X by knowing Y, which is expressed as follows:

$$I(X;Y) = H(X) - H(X|Y). \tag{4.9}$$

EXIT Chart

We consider a simple communication system which shown in Fig. 4.3. The input of the demapper is the received signal from the channel and the extrinsic information from of the channel decoder. The output LLRs of the demapper can be calculated based on the received signal and the *a priori* LLRs provided by the channel decoder (equivalent to the extrinsic information output from the channel decoder), while the output LLRs of the channel decoder is calculated using the extrinsic LLRs from the demapper as an input. It should be emphasized, according to [25], that the random interleaver Π makes the output LLRs of the demapper uncorrelated with the corresponding output LLRs of the channel decoder. In addition, the pdf of the extrinsic LLRs approach the Gaussian-like distribution as the number of iterations increased. Therefore, the LLRs L can be modeled as output of the equivalent Gaussian channel, referred to as side-channel [25], where the transmitted known information bit x is suffering from zero mean independent Gaussian noise with variance σ_N^2. The LLRs can be expressed as:

$$L = \eta_L x + n_L, \tag{4.10}$$

with a mean of LLRs $\eta_L = 2/\sigma_N^2$ and $n_L \sim \mathcal{N}(0, \sigma_L^2)$, where $\sigma_L^2 = 4/\sigma_N^2$. Since L is an LLR value based on Gaussian distribution, the mean value η_L is related to variance as:

$$\eta_L = \frac{\sigma_L^2}{2}. \tag{4.11}$$

Base on (4.11), the conditional pdf of the LLRs L is:

$$\Pr(l|B=b) = \frac{\exp(-((l-(\sigma_L^2/2)b)^2)/2\sigma_L^2)}{\sqrt{2\pi}\sigma_L}. \tag{4.12}$$

To measure the mutual information between transmitted coded bits B and the LLRs L, we use the general definition of mutual information in (4.8). Hence the mutual information $I(B;L)$ is calculated as follows:

$$I(B;L) = \frac{1}{2}\sum_{b=-1,1}\int_{-\infty}^{\infty}\Pr(l|B=b)\cdot\log_2\frac{2\Pr(l|B=b)}{\Pr(l|B=-1)+\Pr(l|B=1)}dl, \tag{4.13}$$

with the following quite general assumptions:

$$\Pr(l,b) = \Pr(l|B=b)\Pr(b) \tag{4.14}$$

$$\Pr(l) = \frac{1}{2}(\Pr(l|B=-1)+\Pr(l|B=1)) \tag{4.15}$$

$$\Pr(l|B=-1) = \Pr(l|B=1) = \frac{1}{2}, \tag{4.16}$$

where b and l are the realizations of the random variables B of the transmitted coded bits and its corresponding LLRs, respectively.

For the computational convenience, Monte Carlo simulations followed by histogram measurement can be used to calculate the MI expressed in (4.13). We can further apply the following additional assumptions on $\Pr(l|B=b)$ to avoid integration in (4.13); we assume the symmetry and consistency properties of $\Pr(l|B=b)$, as follows:

$$\text{Symmetry}: \quad \Pr(l|B=-1) = \Pr(-l|B=1) \tag{4.17}$$

$$\text{Consistency}: \quad \Pr(l|B=-1) = \Pr(-l|B=1)\exp(l), \tag{4.18}$$

and therefore,

$$\Pr(l|B=-1) = \Pr(l|B=1)\exp(l). \tag{4.19}$$

By invoking the property that LLR is ergodic, the computation of the MI can be replaced by time averaging[22]. In practice, the consistency assumption is not always satisfied. Nevertheless, we can still assume the consistency to obtain the accurate and reliable MI

[22]. Equation (4.13) hence be approximated as:

$$I(B;L) = 1 - \int_{-\infty}^{\infty} \Pr(l|B=-1) \cdot \log_2(1+\exp(-l))dl \tag{4.20}$$

$$= 1 - E[\log_2(1+\exp(-l))|B=-1] \tag{4.21}$$

$$\approx 1 - \frac{1}{N}\sum_{i=1}^{N} \log_2(1+\exp(-b_i l_i)). \tag{4.22}$$

From [22], the MI can be expressed by the so-called J-function using the Gaussian approximation. The J-function provides the relationship between the MI and the variance of LLRs σ, which is shown as:

$$J(\sigma) = I(\sigma_L = \sigma), \tag{4.23}$$

with

$$\lim_{\sigma\to 0} J(\sigma) = 0, \qquad \lim_{\sigma\to\infty} J(\sigma) = 1, \qquad \sigma > 0. \tag{4.24}$$

It is well known that the J-function can be accurately approximated by the following convenient form:

$$J(\sigma) = \left(1 - 2^{-H_1\sigma^{2H_2}}\right)^{H_3} \tag{4.25}$$

and

$$\sigma = J^{-1}(I) = \left(-\frac{1}{H_1} log_2(1 - I^{\frac{1}{H_3}})\right)^{\frac{1}{2H_2}}. \tag{4.26}$$

Fig. 4.4 shows an example of the EXIT chart of the system depicted in Fig. 4.3. The demapper curve indicates the extrinsic MI output from the demapper, while the decoder curve shows the extrinsic MI of decoder. The real MI exchange between the demapper and the channel decoder is visualized as the decoding trajectory.

4.3.2 Convergence Property

The result of EXIT chart analysis indicating the convergency properties of the proposed system is provided in this subsection. The EXIT chart analysis covers the iterations $ACC^{-1} \rightleftharpoons D_k$ and $ACC^{-1} \rightleftharpoons D_0$, where notation "$\rightleftharpoons$" denotes LLR exchange. We use three-

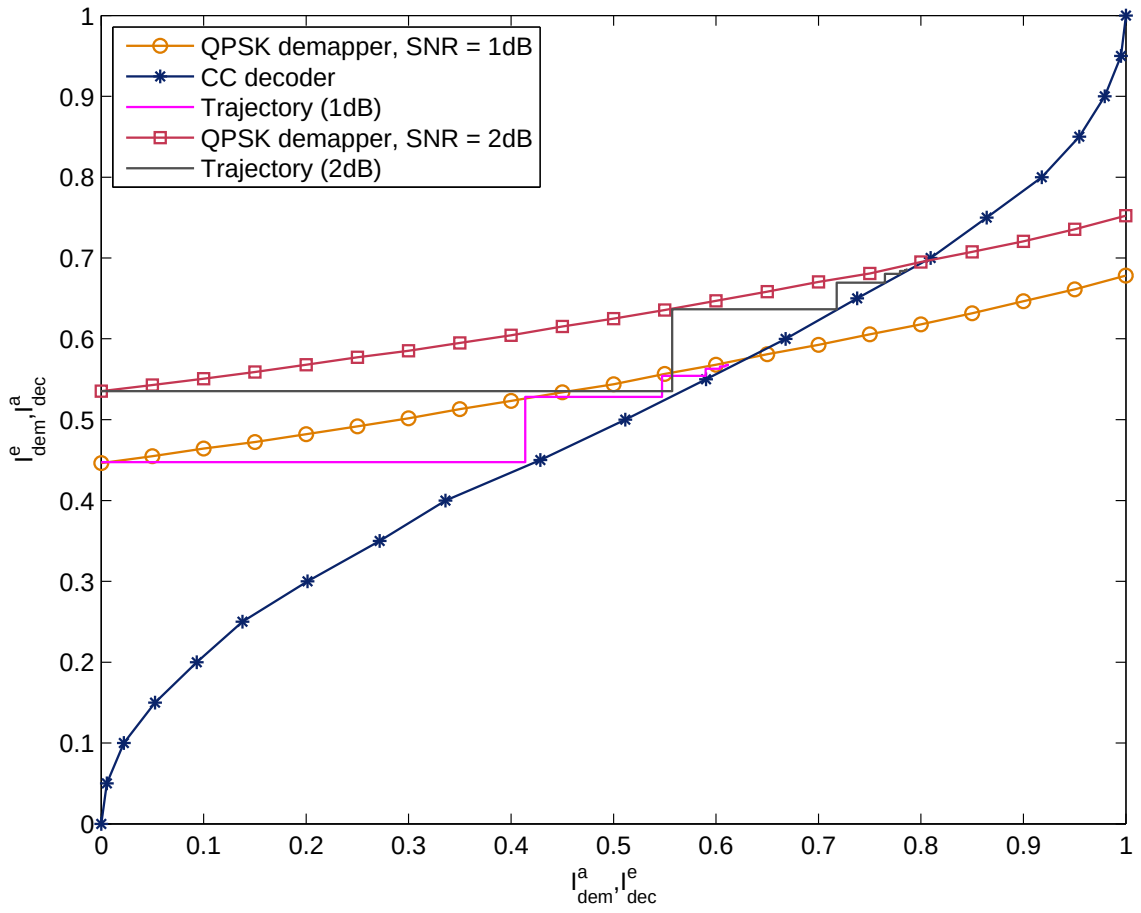

Figure 4.4: An example of the EXIT chart. Channel code: memory-1 convolutional code. Mapper: quadrature phase shift keying (QPSK) with natural mapping pattern.

dimensional (3D) EXIT chart to visualize the extrinsic information exchange between ACC^{-1} and D_1 as well as D_1 and D_2. As shown in the Fig. 4.2, two *a priori* LLR sequences are fed to D_1. Therefore, we evaluate the transfer function for D_1 in the form of:

$$I^{\mathrm{e}}_{\mathbf{c}_1} = T_1(\mathbf{L}^{\mathrm{a}}_{\mathbf{c}_1}, f_c(\mathbf{L}^{\mathrm{e}}_{\mathbf{u}_2}, q)), \tag{4.27}$$

where $I^e_{c_1}$ represents the mutual information between the LLRs of the coded bits obtained byD_1, and their corresponding transmitted bits . $\mathbf{L}^a_{\mathbf{c}_1}$ denotes the *a priori* LLRs provided into D_1, which is equivalent to the extrinsic LLRs of coded bits output from ACC^{-1}. The extrinsic LLRs of the uncoded bits output from D_2, $\mathbf{L}^e_{\mathbf{u}_2}$, is updated by the f_c function and then fed to D_1. The transfer function for D_2 can be obtained in the same way as (4.27) was derived. The q value can be estimated from the decoder output LLR, as:

$$\hat{q} = \frac{1}{N}\sum_{g=1}^{N} \frac{\exp(L^{\mathrm{p}}_{\mathbf{u}_1,g}) + \exp(L^{\mathrm{p}}_{\mathbf{u}_2,g})}{[1+\exp(L^{\mathrm{p}}_{\mathbf{u}_1,g})]\cdot[1+\exp(L^{\mathrm{p}}_{\mathbf{u}_2,g})]}, \tag{4.28}$$

with $L^{\mathrm{p}}_{\mathbf{u}_k,g} \in \mathbf{L}^{\mathrm{p}}_{\mathbf{u}_k}$, N represents the number of the *a posteriori* LLR pairs from the two decoders with their absolute values larger than a given threshold T.

The 3D EXIT chart illustrated in Fig. 4.5 is with parameters BSC crossover probabilities $p_1 = 0.05$, $p_2 = 0.06$ and doping ratio $P_{for} = 2$ at SNR $= -3.6$ dB for both the channels between the forwarding nodes and the FD. Since the D_1 and D_2 are symmetric, we only show the 3D EXIT chart of the *HI* loop for the transmission chain of the forwarding node 1 and trajectory which is obtained by evaluating the mutual information between extrinsic LLRs and the corresponding information bit sequence, by independent (non chained) simulation. Under large enough SNR and sufficient times of iterations, the trajectory can finally reach a point very close to the (1.0,1.0,1.0) MI point, which indicates that the original message, relayed by the forwarding nodes, can be reconstructed perfectly.

Figure 4.6 shows the EXIT curves and trajectory of the decoders ACC^{-1} and D_0 where $p_1 = 0.05$, $p_2 = 0.06$ and $P_{ori} = 2$. After several iterations, the trajectory achieve (1.0,1.0) point and the originator's information be recovered completely. It should be emphasized here that Fig. 4.6 indicates the case where two *HI's* perform as many iterations as no more gain in mutual information can be achieved. However, even without full iterations of two *HI* loops, which results in even smaller value of mutual information after the *a posteriori* LLRs combining, the EXIT curves of ACC^{-1} and D_0 do not intersect until a point very close to the (1.0, 1.0) MI point. According to our simulations, mutual information of 0.73

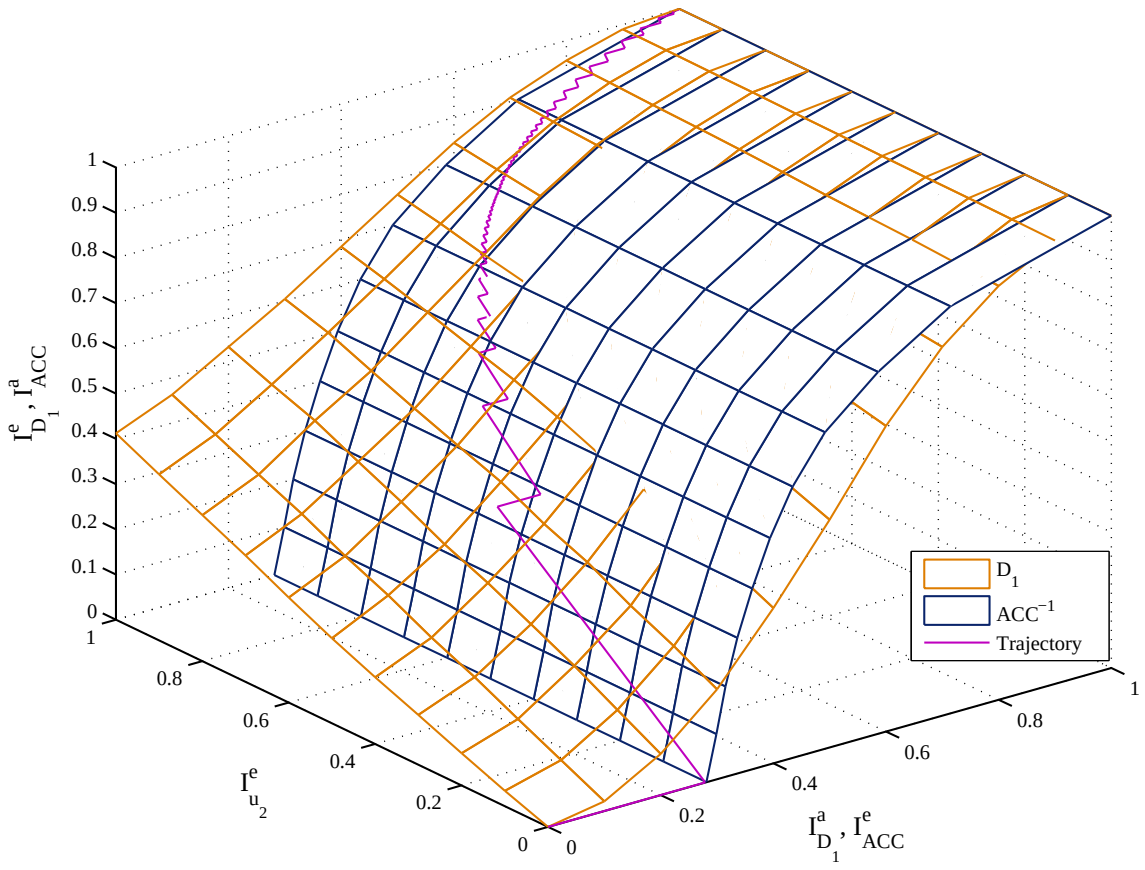

Figure 4.5: The 3D EXIT chart of the HI loop for the transmission chain of the forwarding node 1. The doping ratio for ACC is $P_{for} = 2$ and the crossover probabilities for the BSC channels are $p_1 = 0.05$ and $p_2 = 0.06$. For the both links between the forwarding nodes and the FD, SNR $= -3.6$ dB.

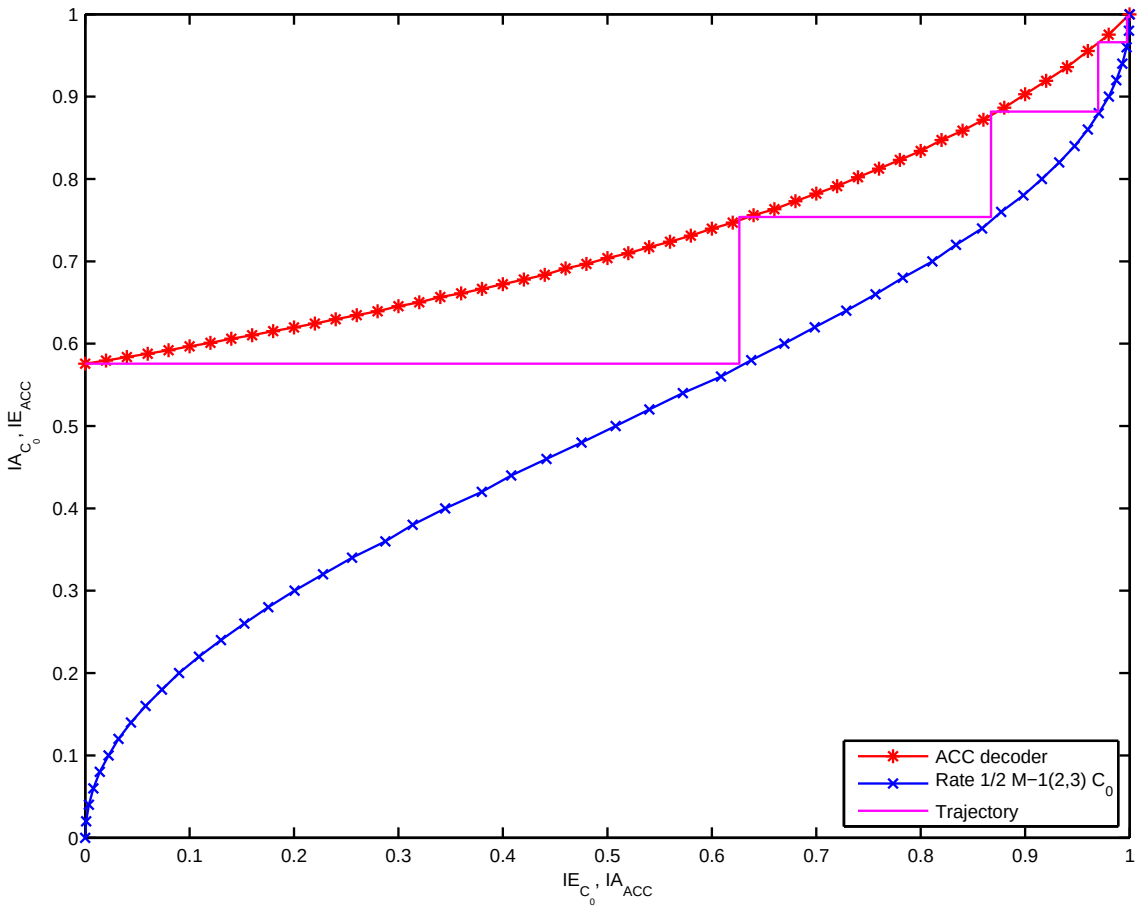

Figure 4.6: EXIT curves of ACC^{-1} with doping ratio $P_{ori} = 2$ and D_0 with rate 1/2. BSC crossover probabilities $p_1 = 0.05$ and $p_2 = 0.06$.

after the *a posteriori* LLRs combining which is the case of non-full iterations over two *HI's*, can still keep the tunnel open.

4.4 Simulation Results

BER performances of the proposed system with three representative value pairs of p_1 and p_2 are shown in Fig. 4.7, Fig. 4.8 and Fig. 4.9, respectively. In our simulations, we set the frame length at 100000 information bits. 30 *HI's* and 5 *VI's* are performed in the joint decoding part at the FD. After LLRs combing base on (4.4), we performed 10 *HI's* to get the estimate of the original information.

The BER performance of the proposed system for two relatively small p_1 and p_2 value pairs, which are described in the figure caption, are illustrated in Figs. 4.7 and 4.8, respectively. It is found from the BER simulation results that, we can achieve clear turbo cliff over the AWGN channel. Also, we can see that, a performance gain of about $2-3$ dB can be achieved by performing *VI*. The larger the gain, the smaller the $\hat{p}$ value, in which case the two forwarding nodes are highly correlated.[1]

Fig. 4.9 shows the BER performance for a relatively large p_1 and p_2 value pair ($p_1 = 0.1$, $p_2 = 0.2$). Since with such high *intra-link* error probabilities, the area of the admissible rate region supported by the Slepian-Wolf theorem is relatively small, compared with the independent coding case. Hence, the benefit of the proposed structure decreases. As shown in Fig. 4.9, turbo cliff can be achieved even without *VI* and performance gain of about 0.7 dB can be achieved with *VI* over without *VI*. Fig. 4.9 also shows the BER curve without the LLR modification before combining them, performed to take into account the *intra-link* error probabilities by (4.4). It can be seen that without the LLR modification, the error floor appears at very high BER range ($\approx 10^{-1}$), compared with that with LLR modification.

4.5 Rate Optimization

Since the gap between the EXIT curves of ACC^{-1} and D_0 shown in Fig. 4.6 is very large, we can increase the rate of C_0 by, for example using punctured convolutional codes to achieve better matching of the two EXIT curves. For this purpose, we propose an algorithm to maximize the rate of C_0 by using an irregular convolutional code [31].

[1]The CEO problem belongs to distributed lossy compression problem in Network Information Theory, and the limits for this problem, are not yet known, except for some special cases.

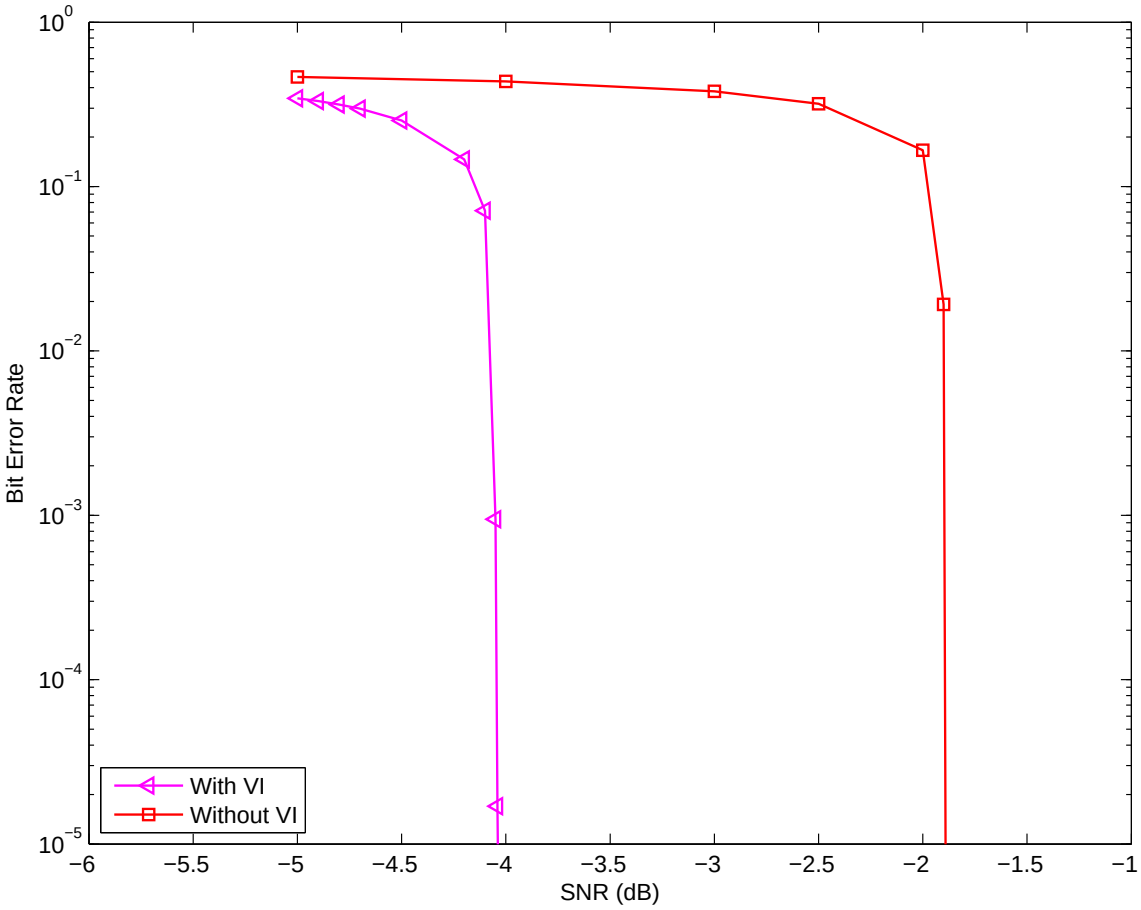

Figure 4.7: BER performance of the proposed system. The crossover probabilities for the BSC channels are $p_1 = 0.05$ and $p_2 = 0.06$.

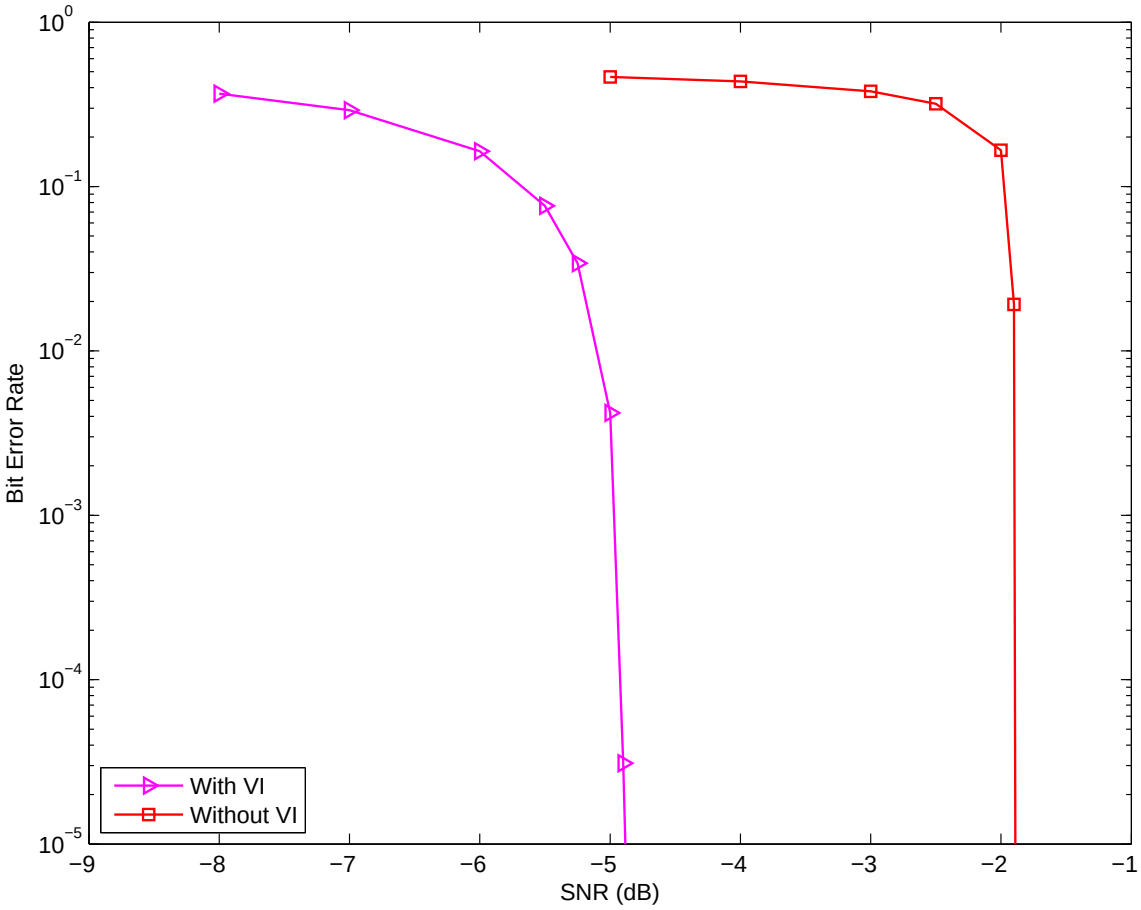

Figure 4.8: BER performances of the proposed system. The crossover probabilities for the BSC channels are $p_1 = 0.01$ and $p_2 = 0.02$.

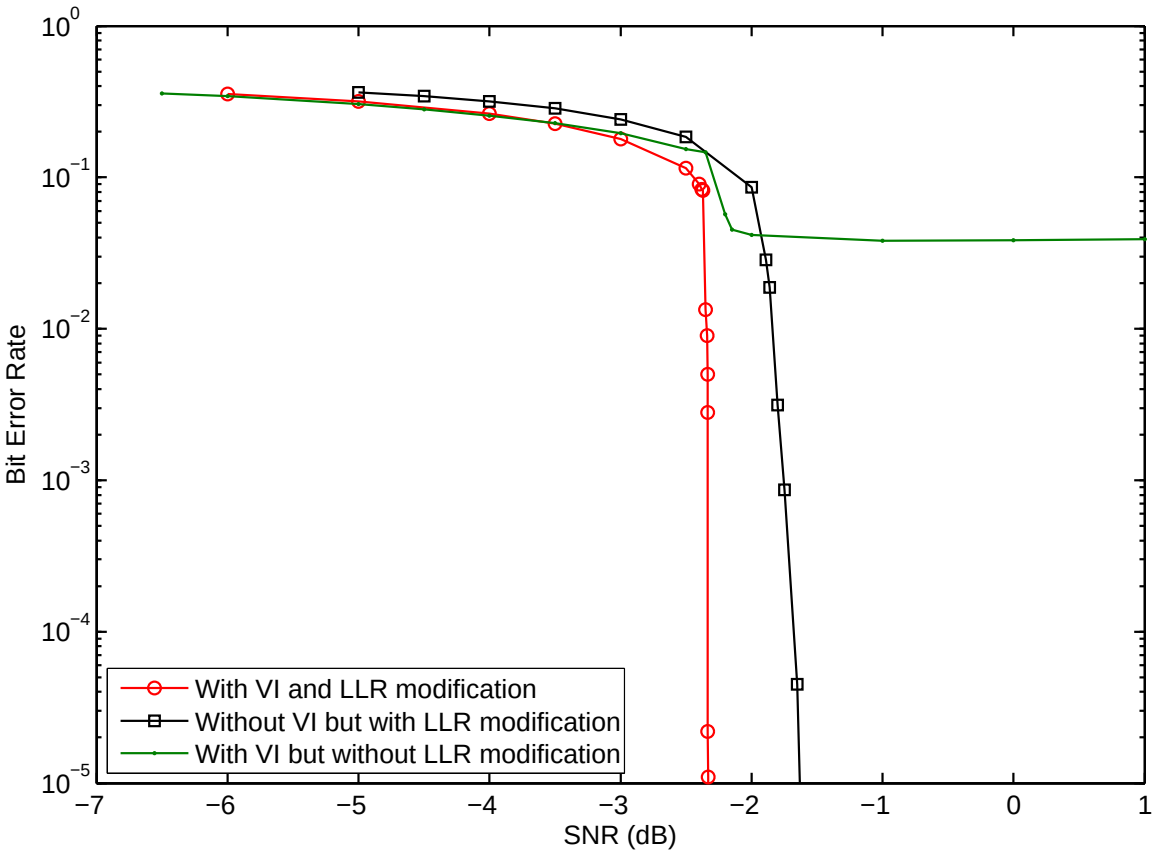

Figure 4.9: BER performances of the proposed system. The crossover probabilities for the BSC channels are $p_1 = 0.1$ and $p_2 = 0.2$.

We suppose that the C_0 is constructed from a group of subcodes to form an irregular convolutional code. In this group of subcodes, each component has different rate R_m, where R_m takes the values shown in the left column of Table 4.1. The EXIT curve of each component is shown in Fig. 4.10. An irregular convolutional code encodes the sequence of data, fractions-by-fractions, using different component in the group of subcodes, where the length of each fraction is corresponding to the ratio α_m. Fig. 4.11 provides an example of an irregular convolutional code.

Since the EXIT curve of the inner code decoder ACC^{-1} is fixed, the target of this optimization is to construct an irregular convolutional code of the outer code C_0 to maximize the rate of C_0 while keeping the convergence tunnel open. For this purpose, we propose an algorithm to minimize the horizontal gaps between the EXIT curves of ACC^{-1} and D_0. This is equivalent to maximize the rate of C_0 since the area below the EXIT curve of D_0 is determined by the rate.

Hence, the problem can be formulated into the following form:

$$\begin{aligned} &\min \sum_m -\alpha_m \cdot R_m \\ &\text{s.t.} \sum_m \alpha_m = 1 \\ &\text{and} \sum_m -\alpha_m \cdot \mathrm{F}^{\mathrm{e}}(\mathbf{IE}_{\mathrm{acc}}) < -(\mathbf{IA} + \epsilon), \end{aligned} \tag{4.29}$$

where, α_m is, as mentioned before, the ratio of using the corresponding component in the group of subcodes with rate R_m. The F^{e}-function which is an approximated function given by (4.30) that calculates the extrinsic information for different rate R_m using different parameters H_1, H_2, H_3 which are given in Table 4.1 [32]. The values of H_1, H_2 and H_3 were obtained by the least square curve matching technique. IE_{acc} is the extrinsic MI which is obtained from simulation. IA is the *a priori* MI which take value in [0∼1]. ϵ is pre-defined horizontal gaps between the two EXIT curves:

$$\mathrm{F}^{\mathrm{e}}(\mathbf{IE}_{\mathrm{acc}}) \approx (1 - 2^{-H_1 \cdot \mathbf{IE}_{\mathrm{acc}}^{H_2}})^{H_3}. \tag{4.30}$$

In this problem, α_m is the optimization variable to be determined. Obviously, this problem can be solved by linear programming (LP) technique. Fig. 4.12 shows the D_0's EXIT curve obtained by using the result of the code optimization, where the EXIT curve of ACC^{-1} is also presented. The ϵ setting for the obtained EXIT curve is shown in Table 4.2. The α_m values obtained as the result of the optimization are shown in the box in

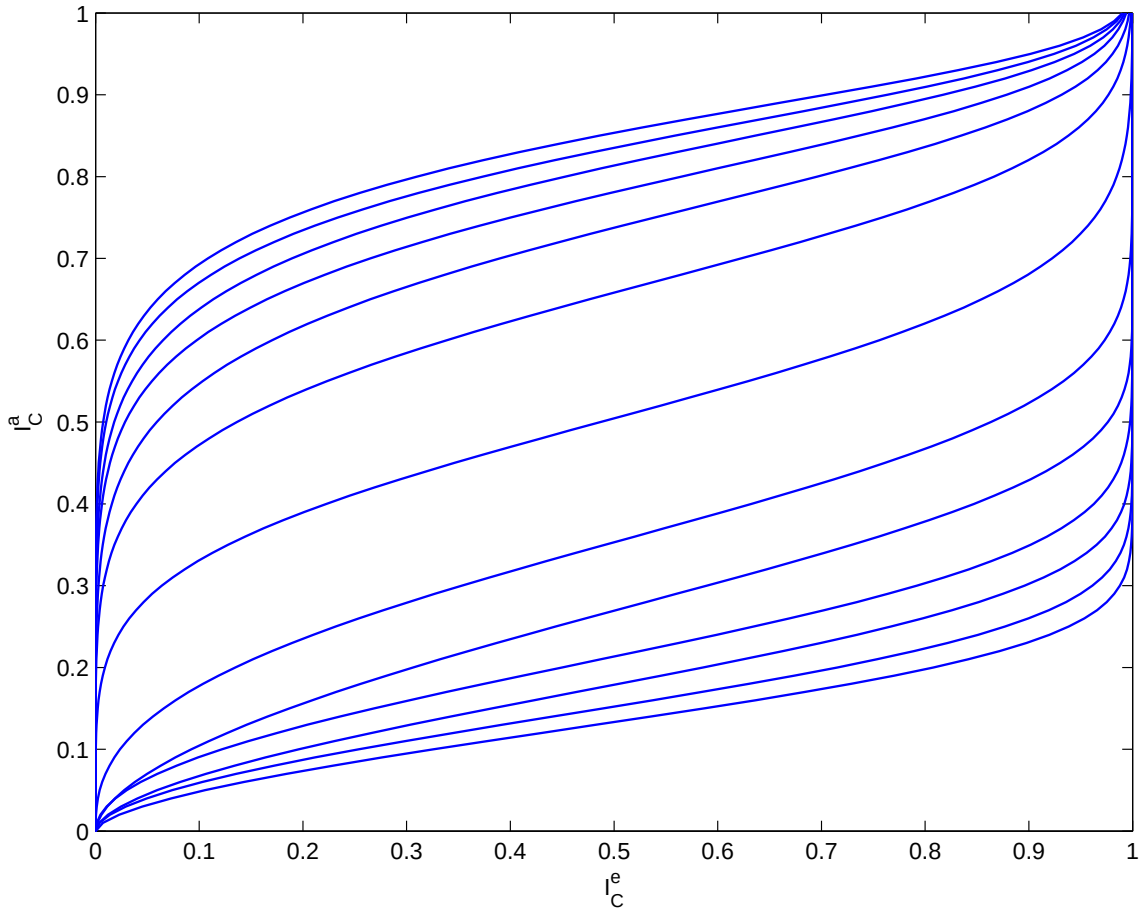

Figure 4.10: EXIT curves of each component of a group of subcodes.

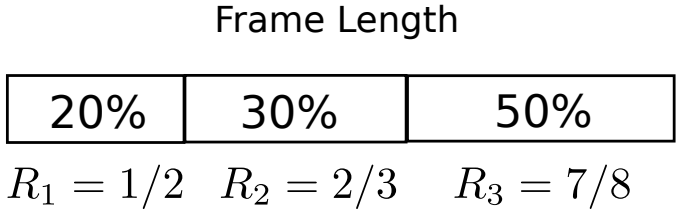

Figure 4.11: An example of irregular convolutional code which consists of three subcodes with rate are 1/2, 2/3 and 7/8, respectively.

Rate	H_1	H_2	H_3
1/8	195.400	2.9579	0.5670
1/7	118.0075	2.7906	0.6567
1/6	109.7092	3.1510	0.5521
1/5	63.5397	2.9603	0.6883
1/4	81.4071	4.3086	0.4041
1/3	31.4326	3.7342	0.6785
1/2	13.8206	3.2731	1.5506
2/3	9.6326	5.0674	1.1642
3/4	7.9191	7.6973	0.7759
4/5	6.9441	9.1556	0.7442
5/6	6.3174	12.0614	0.5828
6/7	5.9538	13.8573	0.5577
7/8	5.6122	17.1775	0.4657

Table 4.1: The coefficient of approximated function.

Fig. 4.12. It is found that the two curves are closely matched, while the convergence tunnel is open until a point very close to the (1.0, 1.0) MI point.

In fact, we do not provide in-depth information theoretic considerations on the relationship between the code rates and Hamming distortion [33] in this thesis. However, when we seek for the optimal rate allocations to the codes used by the originator and forwarding nodes, we have to first identify the relationship in the information theoretic framework of the CEO problem. Especially, the rate of C_0 should be determined by the rate-distortion function, which takes into account the Hamming distortion and allocated rate to the nodes in the network, in general.

It is expected that the more forwarding nodes involved, still error-free communication is possible with a high rate of C_0; we may be able to eliminate C_0 [20, 34] while keeping the Hamming distortion lower than a specified low enough value. In **Chapter 3**, the BER performance is improved when the number of forwarding nodes (sensors) is increased. From the analysis of the error floor, it is possible to guarantee reliable communication[2] by involving more forwarding nodes.

[2]"reliable communication" is defined from the lossy distributed coding viewpoint such that the Hamming distortion left after decoding is lower than a specified value.

MI point	1	2	3	4	5	$\cdots$	20	21
ϵ	0	0.001	0.001	0.001	0.001	$\cdots$	0.001	0

Table 4.2: The pre-defined horizontal gaps between two EXIT curves: ϵ setting.

4.6 Summary

In this chapter, we have examined coding and decoding strategies on the issue of WMNs from the CEO problem viewpoint.

In WMNs, the energy and spectrum efficiencies should be optimized as the whole network rather than an assembly of many point-to-point (P2P) connections. Each forwarding node is small transceiver with low power consumption, where energy is very scarce. We thereby proposed a very simple strategy at the forwarding nodes and a *joint* decoding scheme by exploiting the correlation knowledge among *intra-links* at the FD. Even though errors are detected in the signals received by the forwarding nodes, they are correlated because of coming from the same originating node. Therefore, utilization of the Slepian-Wolf theorem allowing distorted source recovery should be the theoretical basis of the WMNs transmission chain design.

The simulation results show that we can achieve roughly $2 - 3$ dB gain compared with separately decoding scheme. By optimizing the code parameters, close-limit BER performance can be expected, which belongs to the issue of the standard EXIT matching problem. More fundamental question is that how the relationship between the code rates and Hamming distortion can be formulated, and how it can be solved with the aim of their applications to WMNs. This is left as future study.

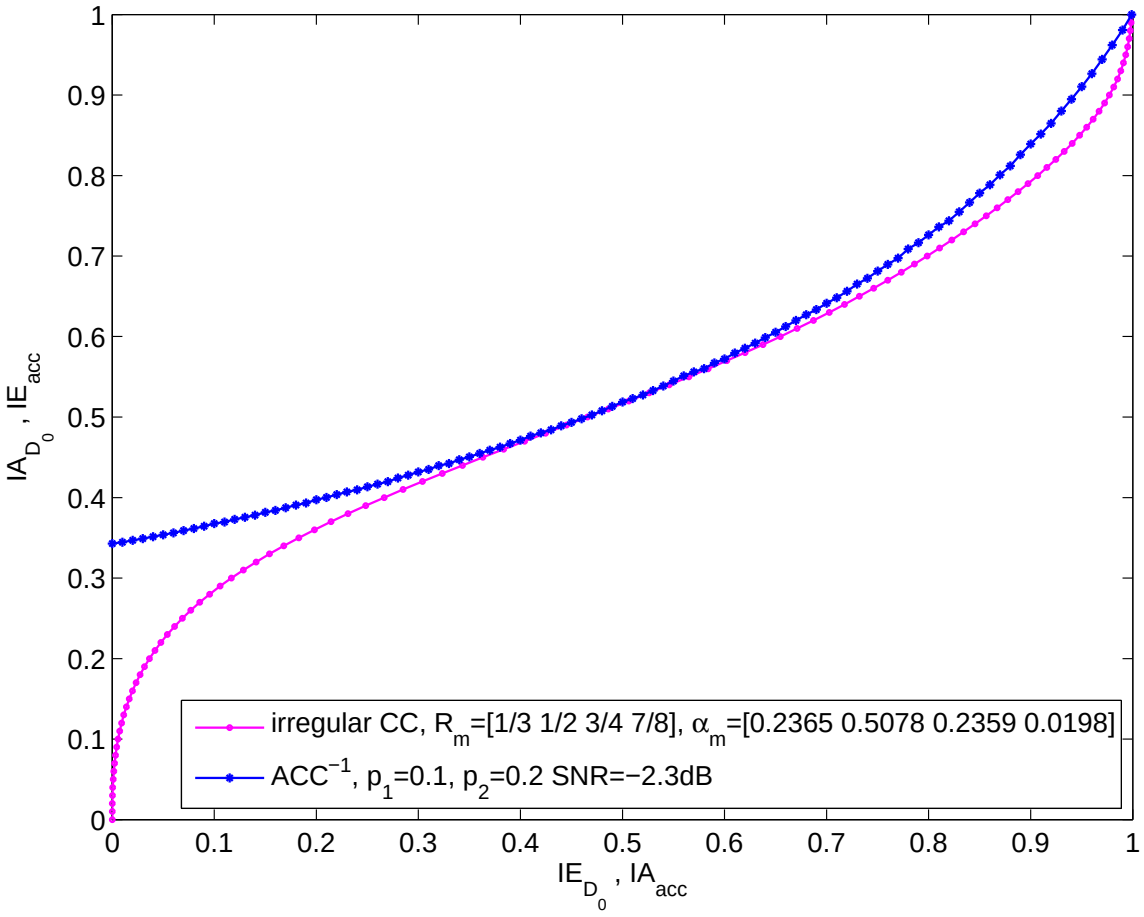

Figure 4.12: EXIT curves of ACC^{-1} with doping ratio $P_{ori} = 2$ and D_0 with different coding rates. BSC crossover probabilities $p_1 = 0.1$ and $p_2 = 0.2$. CC: convolutional code

Chapter 5

Conclusion and Future Work

5.1 Conclusion

In this thesis, we have investigated the coding-decoding strategies for a parallel WSN and a very simple WMN from the viewpoint of the CEO problem.

At the agents (sensors/forwarding nodes), heavy decoding process was not performed. Instead, ErF strategy was used for reducing the complexity of the agents, and hence also reducing the energy consumption. We then proposed the decoding technique which can well exploiting the correlation knowledge among those agents by using the LLR updating function at the CEO node (FD/FC). Excellent performances can be achieved with our proposed technique.

In **Chapter 2**, we first introduce the Slepian-Wolf relaying system which our research group has already solved with achieving excellent performances in BER/FER. Then, we simply explain the reason why we shift from the lossless cases to the lossy cases.

In **Chapter 3**, the case where multiple sensors gathering data from the sensing object aim to transmit the observed data to the FC via parallel links has been considered. We modeled this parallel WSN from the viewpoint of the CEO problem. A simple coding-decoding strategy was proposed to exploit the correlation among those sensors. Further, we proposed an iterative estimation algorithm for estimating observation error probabilities. From the simulation results, it can be found that the estimation algorithm can be applied in many situations. Thereby, reality and scalability of our proposed system were proven.

In **Chapter 4**, we further apply the coding-decoding strategy to a simple WMN where none of the forwarding nodes has error-free information part from the originator. The EXIT chart analysis and BER performance were evaluated for the proposed system model.

In addition, we have shortly discussed the rate optimization for the channel encoders in the network by using linear programming technique based on the EXIT chart analysis.

5.2 Future Work

In general, there are several tasks left for future work, which are list as follow:

1. Multiplexing transmission: Multiple access channel (MAC) and/or Orthogonal channel from sensors to FC;

2. Adaptive resource allocation: Rate optimization and power allocation optimization for the whole system model;

3. Based on the Network information theory, the rate region for given distortion pair and the bounds should be derived for more general cases;

4. Establish techniques that can evaluate convergence property of iterative techniques used to achieve the rate-distortion region;

5. Optimize rate/distortion by using the techniques which can evaluate convergence property based on the obtained rate region;

6. Apply the coding and decoding strategies to more complicated WMN/WSN with multi-hop transmission [35];

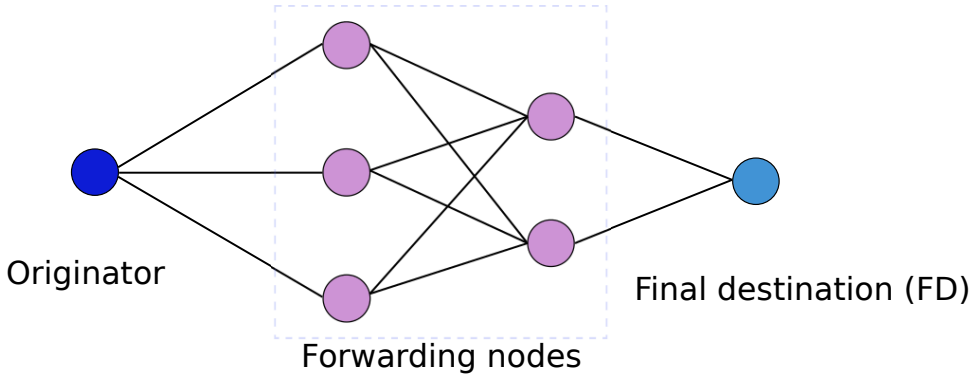

Figure 5.1: An example of a complicated WMN for further study.

7. Establish outage calculation method for the system model we assumed in this thesis from the viewpoint of the CEO problem;

8. Calculate the theoretical bound for the very simple WMN with the *intra-link* probabilities as the parameters.

Appendix A: *lsqnonneg* Algorithm

The *lsqnonneg* algorithm is the modified version of the Standard algorithm for the nonnegative least squares (NNLS) problem proposed by Lawson and Hanson, which is an active set method [19].

The NNLS problem can be formulated in the following form, which is: given a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ and the set of observed values given by $\mathbf{b} \in \mathbb{R}^m$, find a nonnegative vector $\mathbf{x} \in \mathbb{R}^n$ to minimize the functional $f(\mathbf{x}) = \frac{1}{2}\|\mathbf{Ax} - \mathbf{b}\|^2$, as:

$$\begin{aligned} \min \quad & f(\mathbf{x}) = \frac{1}{2}\|\mathbf{Ax} - \mathbf{b}\|^2 \\ s.t. \quad & \mathbf{x} \geq 0 \end{aligned} \tag{5.1}$$

$\nabla f(\mathbf{x}) = \mathbf{A}^T(\mathbf{Ax} - \mathbf{b})$ is defined to be the gradient of $f(\mathbf{x})$. The Kraush-Kuhn-Tucker (KKT) optimality conditions for the NNLS problem (5.1) are expressed as follows:

$$\begin{aligned} \mathbf{x} &\geq 0 \\ \nabla f(\mathbf{x}) &\geq 0 \\ \nabla f(\mathbf{x})^T \mathbf{x} &= 0 \end{aligned} \tag{5.2}$$

The NNLS problem quite often appears in linear algebra. Various methods have been proposed to solve the NNLS problem. Among those methods, the algorithm proposed by Lawson and Hanson is supposed to be the first method, which belongs to active set method. Mathworks further modified the standard algorithm which is summarize in **Algorithm** 3.[1] Lawson and Hanson also prove that this algorithm requires finite iterations. The algorithm can converge at a point where the KKT conditions are satisfied.

[1]The matrix $\mathbf{A}^Q$ is a matrix associated with only the variables currently in the passive set Q.

Algorithm 3: *lsqnonneg*

Input: $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$
Output: $\mathbf{x}^* \geq 0$ such that $\mathbf{x}^*$=arg min $\|\mathbf{Ax} - \mathbf{b}\|^2$
Initialization: $Q = \emptyset$, $S = 1, 2, \cdots, n$, $\mathbf{x} = \mathbf{0}$, $\mathbf{v} = \mathbf{A}^T(\mathbf{b} - \mathbf{Ax})$
while $S \neq \emptyset$ *and* $\{\max_{i \in S}(v_i) > tolerance\}$ **do**
    $j = \arg\max_{i \in S}(v_i)$;
    Include the index j in Q and remove it from S;
    $\mathbf{z}^Q = [(\mathbf{A}^Q)^T\mathbf{A}^Q]^{-1}(\mathbf{A}^Q)^T\mathbf{b}$;
    while $\min(\mathbf{z}^Q) \leq 0$ **do**
        $\beta = -\min_{i \in Q}[x_i/(x_i - z_i)]$;
        $\mathbf{x} = \mathbf{x} + \beta(\mathbf{z} - \mathbf{x})$;
        Update S and Q;
        $\mathbf{z}^Q = [(\mathbf{A}^Q)^T\mathbf{A}^Q]^{-1}(\mathbf{A}^Q)^T\mathbf{b}$;
        $\mathbf{z}^S = \mathbf{0}$;
    end
    $\mathbf{x} = \mathbf{z}$;
    $\mathbf{v} = \mathbf{A}^T(\mathbf{b} - \mathbf{Ax})$;
end

Abbreviations and Notations

ACC	doped accumulator
A/D	analog-to-digital
AWGN	additive white Gaussian noise
BCJR	MAP algorithm proposed by Bahl, Cocke, Jelinek, Raviv
BER	bit error rate
BICM-ID	bit-interleaved coded modulation with iterative detection
BPSK	binary phase-shift keying
BSC	binary symmetric channel
CEO	chief executive officer
ErF	Extract-and-Forward
EXIT	extrinsic information transfer
FC	fusion center
FD	final destination
FER	frame error rate
GI	global iteration
HI	horizontal iteration
KKT	Kraush-Kuhn-Tucker
LDPC	low-density parity-check
LI	local iteration
LLR	log-likelihood ratio
MAC	multiple access channel
MAP	maximum a posteriori probability
MI	mutual information
MSE	mean square error
NNLS	nonnegative least squares
pdf	probability density function
P2P	point-to-point
QPSK	quadrature phase-shift keying
SISO	soft-in-soft-out
VI	vertical iteration
WSN	wireless sensor network
WMN	wireless mesh network

$(\hat{\cdot})$	estimation of the argument
$(\cdot)^{-1}$	inverse of the argument
$(\cdot)^T$	transpose of a vector or matrix
$\lvert\cdot\rvert$	absolute value of the argument
$\lVert\cdot\rVert$	norm of a vector or matrix
$\mathrm{diag}(\cdot)$	diagonal matrix with vector elements on its diagonal components
$\exp(\cdot)$	exponential calculation of the argument
$E[\cdot]$	expectation of a random variable
$F^{\mathrm{e}}(\cdot)$	the function to calculate extrinsic mutual information for convolutional code
$J(\cdot)$	J-function
$\ln(\cdot)$	natural logarithm to base e
$\log_2(\cdot)$	natural logarithm to base 2
$\log(\cdot)$	natural logarithm to any bases
max	maximum value
min	minimum value
$\mathrm{sign}(\cdot)$	the sign of the argument
$T_k(\cdot,\cdot)$	transfer function for D_k

ACC^{-1}	doped accumulator decoder
B	coded bits
$\mathbf{c}_k$	coded sequence after C_k at k-th sensor
$\mathbb{C}$	the set of all complex numbers
C, C_k	channel encoder
D	distortion
D_k	decoder for encoder C_k
$\mathbf{e}_k$	the error sequence used at k-th sensor
h_k	channel coefficient for the channel between the k-th sensor and the FC
$\mathbf{I}$	the identity matrix
$I(\cdot,\cdot)$	mutual information between argument 1 and 2
I^{a}	*a priori* information
I^{e}	extrinsic information
$\mathbf{L}^{\mathrm{a}}$	*a priori* LLR sequence
$\mathbf{L}^{\mathrm{e}}$	extrinsic LLR sequence
$\mathbf{L}^{\mathrm{p}}$	*a posteriori* LLR sequence

$\mathbf{J}$	circular matrix used in the estimation algorithm
M	the number of agents
$\mathcal{N}(\cdot,\cdot)$	Gaussian distribution, argument 1: mean, and argument 2: variance
$\mathbf{n}_k$	zero mean complex AWGN sequence
n_L	zero mean complex AWGN
$\mathbf{P}$	vector of the observation error probabilities
p_k	observation error probability or *intra-link* error probability
$\Pr(\cdot)$	probability of the variable
P_d, P_{ori}, P_{for}	doping ratio of the ACC
$\mathbb{R}$	the set of all real numbers
R, R_k	coding rate
$\mathbf{s}, \mathbf{s}_k$	modulated symbol sequence
T	threshold of LLR value
$\mathbf{u}$	the binary information sequence of the sensing object or originator
$\mathbf{w}_k$	zero mean white Gaussian noise sequence
Γ	average SNR
γ	instantaneous SNR
ϖ	error floor in BER
σ^2, σ_N^2	Gaussian noise variance
σ_L^2	variance of LLR
η_L	mean of LLR

Bibliography

[1] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near shannon limit error-correcting coding and decoding: Turbo-codes. 1," in *Proc. IEEE Int Communications ICC 93. Geneva. Technical Program, Conf. Record Conf*, vol. 2, pp. 1064–1070, 1993.

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

[3] M. Cheng, X. Zhou, K. Anwar, and T. Matsumoto, "Simple relay systems with BICM-ID allowing intra-link errors," *IEICE Trans. on Comm., Special Section on Coding and Coding Theory-Based Signal Processing for Wireless Communications*, vol. E95-B, pp. 3671–3678, December 2012.

[4] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Transactions on Information Theory*, vol. 19, pp. 471– 480, July 1973.

[5] T. Berger, Z. Zhang, and H. Viswanathan, "The CEO problem," *IEEE Transactions on Information Theory*, vol. 42, pp. 887–902, May 1996.

[6] H. Viswanathan and T. Berger, "The quadratic Gaussian CEO problem," *IEEE Transactions on Information Theory*, vol. 43, pp. 1549 –1559, September 1997.

[7] Y. Oohama, "The rate-distortion function for the quadratic Gaussian CEO problem," *IEEE Transactions on Information Theory*, vol. 44, pp. 1057–1070, May 1998.

[8] V. Prabhakaran, D. Tse, and K. Ramachandran, "Rate region of the quadratic Gaussian CEO problem," in *IEEE International Symposium on Information Theory 2004 (ISIT2004)*, p. 119, 27 June–2 July 2004.

[9] J. Chen, X. Zhang, T. Berger, and S. Wicker, "An upper bound on the sum-rate distortion function and its corresponding rate allocation schemes for the CEO prob-

lem," *IEEE Journal on Selected Areas in Communications*, vol. 22, pp. 977 – 987, August 2004.

[10] X. He, X. Zhou, K. Anwar, and T. Matsumoto, "Wireless mesh networks allowing intra-link errors: CEO problem viewpoint," in *2012 International Symposium on Information Theory and its Applications (ISITA)*, (Hawaii), pp. 61–65, October 2012.

[11] X. Zhou, X. He, K. Anwar, and T. Matsumoto, "GREAT-CEO: larGe scale distRibuted dEcision mAking Technique for wireless Chief Executive Officer problems," *IEICE Trans. on Comm., Special Section on Coding and Coding Theory-Based Signal Processing for Wireless Communications*, vol. E95-B, pp. 3654–3662, December 2012.

[12] R. Thobaben, "On distributed codes with noisy relays," in *42nd Asilomar Conference on Signals, Systems and Computers*, (Pacific Grove, CA), October 2008.

[13] K. Anwar and T. Matsumoto, "Spatially concatenated codes with turbo equalization for correlated sources," *IEEE Transactions on Signal Processing*, vol. 60, pp. 5572–5577, October 2012.

[14] M. Cheng, A. Irawan, K. Anwar, and T. Matsumoto, "BICM-ID for relay system allowing intra-link errors and a similarity constellation to ARQ schemes," in *Progress in Electromagnetics Research Symposium (PIERS)*, (Kuala Lumpur, Malaysia), pp. 281–286, March 2012.

[15] J. Garcia-Frias and Y. Zhao, "Near-shannon/slepian-wolf performance for unknown correlated sources over AWGN channels," *IEEE Transactions on Communications*, vol. 53, pp. 555–559, April 2005.

[16] M. Cheng, K. Anwar, and T. Matsumoto, "Outage-Analysis of Correlated Source Transmission in Block Rayleigh Fading Channels," in *2012 IEEE Vehicular Technology Conference (VTC2012-Fall)*, (Quebec, Canada), September 2012.

[17] P.-S. Lu, V. Tervo, K. Anwar, and T. Matsumoto, "Low-complexity strategies for multiple access relaying," in *IEEE 73rd Vehicular Technology Conference (VTC Spring)*, (Budapest, Hungary), pp. 1–6, May 2011.

[18] Z. Xiong, A. Liveris, and S. Cheng, "Distributed source coding for sensor networks," *IEEE Signal Processing Magazine*, vol. 21, pp. 80–94, September 2004.

[19] D. Chen and R. J. Plemmons, "Nonnegativity constraints in numerical analysis," in *Symposium on the Birth of Numerical Analysis*, (Leuven Belgium), October 2007.

[20] A. Razi, K. Yasami, and A. Abedi, "On minimum number of wireless sensors required for reliable binary source estimation," in *Proc. IEEE Wireless Communications and Networking Conf. (WCNC)*, (Quintana-Roo, Mexico), pp. 1852–1857, Mar. 2011.

[21] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate (corresp.)," *IEEE Transactions on Information Theory*, vol. 20, pp. 284 – 287, March 1974.

[22] J. Hagenauer, "The EXIT chart - introduction to extrinsic information transfer in iterative processing," in *12th Europ. Signal Proc. Conf (EUSIPCO)*, (Vienna, Austria), pp. 1541–1548, September 2004.

[23] S. J. Johnson, *Iterative Error Correction: Turbo, Low-Density Parity-Check and Repeat-Accumulate Codes*. Cambridge University Press, 2009.

[24] S. ten Brink, "Convergence of iterative decoding," *Electronics Letters*, vol. 35, pp. 806 –808, May 1999.

[25] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Transactions on Communications*, vol. 49, pp. 1727–1737, October 2001.

[26] T. Richardson and R. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Transactions on Information Theory*, vol. 47, pp. 599 –618, February 2001.

[27] T. Richardson, M. Shokrollahi, and R. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Transactions on Information Theory*, vol. 47, pp. 619 –637, February 2001.

[28] S.-Y. Chung, T. Richardson, and R. Urbanke, "Analysis of sum-product decoding of low-density parity-check codes using a Gaussian approximation," *IEEE Transactions on Information Theory*, vol. 47, pp. 657–670, February 2001.

[29] K. Fukawa, S. Ormsub, A. Tlli, K. Anwar, and T. Matsumoto, "Exit-constrained bicm-id design using extended mapping," *EURASIP Journal on Wireless Communications and Networking*, vol. 2012, pp. 1–17, 2012.

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

[31] M. Tuchler and J. Hagenauer, "EXIT charts of irregular codes," in *2002 36th Conference on Information Sciences and Systems (CISS2002)*, (Princeton University, NJ, USA), pp. 748–753, March 2002.

[32] S. Ibi, T. Matsumoto, S. Sampei, and N. Morinaga, "EXIT chart-aided adaptive coding for MMSE turbo equalization with multilevel BICM," *IEEE Communications Letters*, vol. 10, pp. 486 –488, June 2006.

[33] A. E. Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2011.

[34] J. Haghighat, H. Behroozi, and D. V. Plant, "Joint decoding and data fusion in wireless sensor networks using turbo codes," in *IEEE 19th Int. Symp. Personal, Indoor and Mobile Radio Communications PIMRC*, (Cannes, France), pp. 1–5, September 2008.

[35] X. Zhou, A. O. Lim, K. Anwar, and T. Matsumoto, "Distributed Joint Source-Channel-Network Coding Exploiting Source Correlation for Multiple Access Relay Channel," *European Wireless Conference (EW 2013)*, 2013. Submitted for publication.