Chapter Author: Dr. Abolfazl Razi (Northern Arizona University)

Chapter 8

Information-flow in hard to reach areas: Source estimation using correlated streams

In a typical Wireless Sensor Network (WSN), a number of spatially distributed sensors measure a set of desired environmental and physical parameters across a field and pass the collected data wirelessly to a central station for further processing. Various aspects of WSN systems including efficient data acquisition, optimal sampling rate and distributed compression, error recovery, information fusion, localization and tracking, routing and networking are studied in the past decade. These studies have led to design of novel algorithms to solve different technical challenges and alleviate implementation barriers. This book chapter considers a specific variant of WSNs, where data sources are located in hard to reach areas and therefore additional challenges and issues arise. An important scenario is field monitoring under extreme environmental conditions. Some real-world examples include jet engine temperature monitoring [1] [2], structural health monitoring [3] [4], underwater passive aquatic listeners [4], and space exploration using composite sensors [5]. In some other applications such as navigation systems using battery-free sensors [6] placing sensors at exact data source locations is not feasible or economically viable since the source locations are not stationary or not known in advance. Figure 8-1 provides an illustrative example, where a set of autonomous Unmanned Aerial Vehicles (UAVs) track an animal of interest, collect intermittent and partially accurate imagery data and send it to a ground station for further processing [7].

In these scenarios, an important challenge is the low fidelity of collected information not only due to communication errors, but also due to inaccuracies in the measurements of remotely deployed sensors [8] [9]. To address this issue, distributed coding emerged as an efficient technique to collect information more elegantly by individual sensors such that a reliable

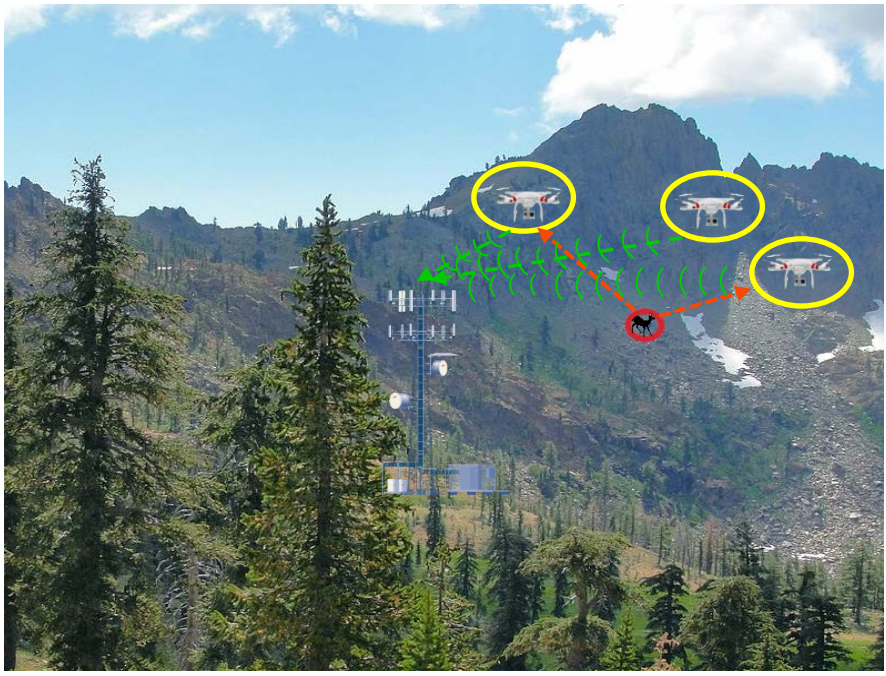


Figure 8-1 Habitat monitoring: a set of autonomous UAVs monitor the object of interest in an inaccessible area and collectively report their observations to a central ground station for further processing.

information is extractable by fusing partially accurate observations [10]. This book chapter is devoted to exploring the state of the art in distributed coding techniques. In particular, a comprehensive review on the fundamental theoretical aspects of this problem is provided with a summary of recent developments. In this context, a novel and implementation-friendly solution is proposed to address this problem and answer a set of more specific technical questions including: i) *how many sensors are optimal in a given conditions?* ii) w*hat is the fidelity of collected information?* and iii) *what is the highest achievable information rate?* We finally provide insights about the integration of the proposed method with modern networking approaches in order to extend it to large-scale multi-hop communication infrastructure.

8.1 INTRODUCTION

A contemporary WSN consists of a multiple sensors mainly deployed in autonomous mobile ground or aerial vehicles with wireless transmission capabilities in order to collectively monitor environmental and physical

parameters and transmit the collected data to a central data fusion unit for further processing. Wireless sensing is an integral part of the emerging Internet of Things (IoT) platforms with a broad range of applications including smart homes, smart grids, transportation and traffic control, aviation, wildlife monitoring, fire control, robotics and industrial automation [12] [13]. Due to the Ad-hoc nature of WSNs with dynamic and vibrant typologies, utilizing a central controller is costly, unreliable and inefficient. Therefore, distributed algorithm design and implementation has gained a lot of attention in research community to realize agile, flexible and scalable data acquisition, compression and transmission in WSNs [14].

The majority of distributed algorithms aim at realizing a smooth and reliable data-flow from sensors to a central processing station by accommodating optimized compression, networking and error recovery techniques based on the assumption that sensor measurements accurately represent the actual monitored parameters. However, there are scenarios where placing sensors in exact data locations is not technically feasible due to a number of reasons including extreme environmental conditions (e.g. jet engine), lack of prior knowledge about source locations (e.g. fire control systems in forests), and maintenance difficulty (e.g. space exploration). In such scenarios, an effective solution is to deploy a cluster of sensors in proximity of a putative data source and process the collected data. In this regard, Distributed Source Coding (DSC) schemes are designed to exploit the intrinsic spatial correlation among sensors observations to compensate for sensors observation errors. This solution brings forward a new set of questions such as i) how many sensor are sufficient to achieve a certain level of fidelity, ii) how to compensate sensing inaccuracies, iii) how often to collect sensor readings and iv) how to exploit the spatial correlation among sensors.

A comprehensive study of ongoing research in designing DSC algorithms during the past years reveals that most recent studies have been focused on developing theoretically near-optimal algorithms. However, the high complexity of these algorithms have prevented their commercial implementations in tiny sensors with limited capabilities [11]. Furthermore, most algorithms are designed and optimized based on two assumptions of data sources being stationary and observation models being fully known. Bothe assumptions are far from reality. Therefore, the majority of currently available System on Chip (SOC) sensor platforms such as Telos B Motes, Stargate, Mica2, Tmote Sky and IBM cricket stick with the conventional point-to-point coding techniques in their physical layer implementations following IEEE 802.15 standard series and do not utilize recently proposed DSC schemes despite their proven superior performances [12] [13] [14] [15].

8 Information-flow in hard to reach areas

This book chapter is devoted to study and design of practically feasible and self-tuning distributed algorithms for robust and adaptive data compression, error recovery and networking for WSNs with partially accurate sensors under dynamic environmental conditions. The main emphasis will be placed on a special case of remote sensing of a common data source with correlated sensors, the so called the Chief Execute Officer (CEO) problem. A novel implementation of distributed coding is discussed in detail, which is based on Parallel Concatenation of Convolution Codes (PCCC) with a novel bi-modal decoder at the receiver.

8.2 THE CHIEF EXECUTIVE OFFICER (CEO) PROBLEM

The mathematical model of the CEO problem (Figure 8-2) is first proposed by Toby Berger in his seminal paper to capture the spatial correlation among the readings of a cluster of sensors that collectively monitor a common data source [16]. This model is inspired by a similar problem in business studies, where a CEO of a company interviews his untrustworthy employees individually and process the interrogation results to extract reliable information about his plant.

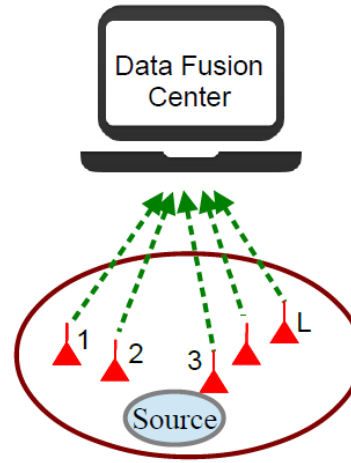


Figure 8-2 CEO Problem: A cluster of L sensors collectively monitor a common source and send their observations to a data fusion center.

The CEO problem is a special case of a more general problem formulation, where L sensors collectively monitor K sources with correlated data, as depicted in Figure 8-3. In this figure, S_i is the i[th] data source and U_j, E_j, X_j, N_j, Y_j and R_j, respectively, denote the observation, observation error, coded sample, channel noise, received sample and transmission rate of the j[th] sensor. The main goal is to design L distributed encoders with minimal rates R_i and one joint decoder in order to estimate K sources with a predefined distortion limit $d(S_i, \hat{S}_i) < D$, where $d(.,.)$ is an arbitrary distance measure. Here, we assume additive observation error model $U_i = \sum_{S_j \in \chi_i} S_j + E_i$, where χ_i represents a set of sources observed by sensor i. For the sake of simplicity, we disregard the interference effect and assume parallel channel model with additive noise $Y_i = X_i + W_i$, where X_i is the coded sample and W_i is the white zero-mean Gaussian noise with variance N_i.

This general model reduces for some important special cases if certain simplifying assumptions hold. For instance, we can set observation error to zero, i.e. $E_i = 0, \forall i = 1,2, \dots, L$, which converts the indirect observation scenario to a simpler case of direct observation [17]. Another special case arises if any of K utilized sensors monitors only one specific source (i.e. $\chi_i = \{S_i\}$), and the correlation among sensor readings are restricted to the intrinsic correlations among the sources [18] [11]. If full information about some of the sensors are considered available to the receiver and the objective is to recover the other sensors' readings, the problem reduces to the coding with side information [19] [20]. Considering only one source $K = 1$, boils the problem down to the CEO problem, where the objective is to estimate the common

source [16]. In all cases, the sensors are not allowed to convene and exchange information with one another; hence, distributed coding is utilized [21].

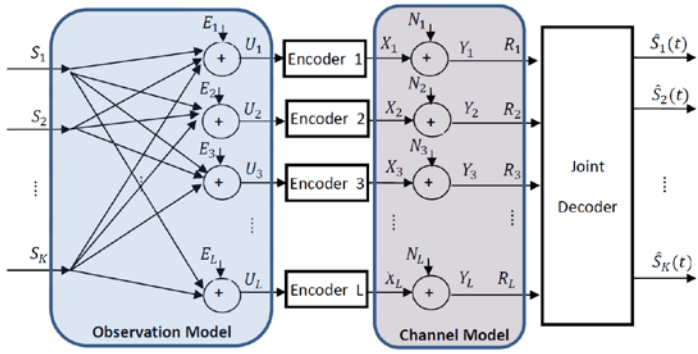


Figure 8-3 Multi-terminal source coding for sensors with direct and indirect observations.

8.2.1 Binary Correlation Model

The most commonly adopted model for correlation among sensor readings is joint Gaussian distribution, where the common source and the observations of sensors are jointly distributed according to a multivariate Gaussian distribution. This model is flawed in reality, since different data sources such light, temperature, humidity, stress, pressure, and etc. do not follow normal distribution in general [22] [23] [24] [25] [26] [27]. Furthermore, observation noise is not always normally distributed. The third drawback of this theoretical model is the ignorance of real-world sensor implementations. In practice, sensors read continuous-valued variables and convert them into binary streams after different stages of quantization (e.g. linear uniform PCM, Lloyd-Max quantizer or vector quantization [28] [29]), companding (e.g. A-Law and μ-Law for speech signals [30] [31]), and sample to bit mapping (e.g. Gray coding [32] [33]).

The recent developments of practical source codes are typically applied to binary bitstreams after a digitization step. To fill this gap between the theoretical framework and the practical implementation, we assume that the source generates binary symbols, which is observed by sensor i through a binary symmetric channel (BSC) with crossover probability β_i. Namely, we have $U_i = S + E_i, S \sim Bernoulli(0.5), E_i \sim Bernoulli(\beta_i)$, where $\beta_i = \boldsymbol{E}[Pr(S \neq U_i)]$ is the expected bit flipping probability from the common source to sensor i for a given source type and sensor implementation. This

model is very general and mimics the impact of noise and other non-linearity sources at different stages of the path from source to the representative binary bitstream. Therefore, this model has been widely adopted in various studies [34] [19] [35] [36] [37] [38]. With this correlation model, the aforementioned CEO model converts to a special cased called binary CEO (bCEO).

8.2.2 Information-theoretic review

In the CEO problem, the goal is to find the rate region $\mathcal{R}_*(D) = \cup_{(R_1,R_2,\dots R_L)} R_D(R_1, R_2, \dots, R_L)$, which includes all combination of rate L-tuples. Here, R_i is the coding rate of sensor i and $R_D(R_1, R_2, \dots, R_L)$ represents all rate tuples for which there exist a joint-decoder capable of estimating the common source within a predefined distortion limit $D^{(n)} = \frac{1}{n}\boldsymbol{E}[\, d(S^n, \hat{S}^n)] < D$, for a given distance measure $d(.,.)$, when the codeword block length n is chosen large enough. One may also be interested in solving a simpler problem of minimizing the sum rate, which is defined as $R = \sum_{i=1}^{L} R_i$, for a given target distortion threshold D.

This problem is intensively investigated from the information-theoretic perspective, yet the exact rate-distortion function for the CEO problem is not known in general [39]. The first information-theoretic result, reported by Toby Berger in [40], states that if the number of agents (sensors) approach infinity ($L \to \infty$) and the agents are allowed to share their observations, then they are able to provide an accurate estimate of the common source symbol S with a limited distortion based on a distortion rate function $D(R)$. As a special case, if the sum-rate R exceeds the entropy of the common source $H(S)$, the decoder fully recovers the source symbol with an arbitrary low distortion, i.e. $R > H(S) \Rightarrow D(R) = 0$.

However, if the agents are not allowed to convene, which is the case in the CEO problem, there does not exist a finite value of the total rate R, for which even infinitely many agents can make the distortion D arbitrarily small. It was shown that for an infinitely large number of agents, the distortion decays exponentially as sum-rate approaches infinity. This fact implies that in practical implementation of distributed coding with a limited number of sensors, we need to settle with an approximate source estimation and not seek an error-free source recovery. This phenomenon is due to the inherent uncertainty of indirect observations. In contrast, for direct observation, correlated sources can be fully recovered within a rate region defined by the famous Slepian-Wolf theorem [18]. A concrete characterization of rate-distortion region, known as Berger-Tung rate region proposed in [41] based

on performing Gaussian quantization method on the observation bits and then applying Slepian-Wolf coding to the quantized the bits.

An important variant of CEO problem is called Quadratic Gaussian CEO (QGCEO) problem, where the source and observation errors are identically and independently distributed (i.i.d.) Gaussian variables. The problem of finding distortion-rate for the QGCEO problem is studied from different perspectives including sum rate distortion function derivation [42] [43], computationally tractable sum-rate calculation [44], sum-rate loss due to lack of inter-sensor communications [45], multiple source estimation [46] [47] [48] and robust coding to recover both the common source and the sensor readings with a certain fidelity [49]. Extensions of these results to vector Gaussian CEO problem can be found in [50] [51] [52] [53].

The CEO problem can be generalized in the sense that the CEO is interested in estimating both common source data and sensors observations with arbitrary fidelity. If the maximum allowable distortions of observations reconstruction approach infinity meaning that the CEO does not care about the observation itself, this problem reduces to the classical CEO problem. On the other hand, if the maximum allowable distortions of the common source observations reconstruction approaches infinity, the problem is reduced to two-terminal direct observation case. The rate region for this general case is characterized in [49].

Another important variant with practical advantages is the binary CEO problem, as discussed earlier. Theoretical limits on the rate-distortion region for the bCEO problem is studied in [39] [54]. However, these results do not provide a closed-form expression for the rate distortion function and rather provide complex expressions for the upper and lower bounds on the rate distortion function. These expressions are complicated to quantify even for the special case of two sensors $L = 2$ and equal observation error $\beta_1 = \beta_2 = \beta$, and consequently are not much useful in designing practical systems.

In this book chapter, we will show how our simplified calculations for the end-to-end coding rate proposed in [55] can be used to configure a cluster of sensors for real-world applications to achieve an acceptable level of accuracy. This method is based on finding a polymatroid region for the rates that fully recovers the correlated sensor readings instead of finding the rates required for recovering the common data source [56] [57].

8.3 PRACTICAL CODING DESIGN FOR THE BINARY CEO PROBLEM

8.3.1 Distributed source coding vs joint coding

In order to elucidate the concept of distributed source coding, here we present a simple example to show its contrast with joint coding. The main distinction between the two methods is the exchange of information among sensors. Consider two sensors with correlated binary inputs. Each sensor reading $\boldsymbol{U}_i$ is a sequence of three bits $\boldsymbol{U}_i = [U_{i1}U_{i2}U_{i3}]$, therefore it takes $2^3 = 8$ options $\boldsymbol{U}_i \in \mathcal{U} = \{S_0, \dots, S_7\}$ as depicted in Figure 8-4. Based on the observation model, U_1 and U_2 can differ at most in one bit (i.e. $d_{Hamming}(U_1, U_2) \leq 1$). The direct implication of this correlation model is that for each realization of $\boldsymbol{U}_1$, there are only 4 possibilities for $\boldsymbol{U}_2$, which include $\boldsymbol{U}_2 = \boldsymbol{U}_1 \oplus \boldsymbol{E}$, where $\boldsymbol{E} = \{[000], [001], [010], [001]\}$. Consequently, the entropy of $\boldsymbol{U}_1$ is $H(\boldsymbol{U}_1) = 3$ bits and the conditional entropy of $\boldsymbol{U}_2$ given $\boldsymbol{U}_1$ is $H(\boldsymbol{U}_2|\boldsymbol{U}_1) = H(\boldsymbol{U}_1 \oplus \boldsymbol{E}|\boldsymbol{U}_1) = H(\boldsymbol{E}) = 2$ bits. The practical implementations of joint coding is straightforward. Sensor 1 sends the plain reading $\boldsymbol{U}_1$ with 3 bits to the destination. Sensor 2 receives $\boldsymbol{U}_1$ and calculates $\boldsymbol{E} = \boldsymbol{U}_1 \oplus \boldsymbol{U}_2$ and then maps it into two bits $[E_1E_2]$ and sends the result to the receiver. The receiver performs sequential decoding (also called onion pealing) by determining $\boldsymbol{U}_1$ first and then calculating $\boldsymbol{E}$ based on received bits $[E_1E_2]$. Subsequently, it calculates $\boldsymbol{U}_2 = \boldsymbol{U}_1 \oplus \boldsymbol{E}$ (Figure 8-4a).

Lossless distributed coding relies on the famous Slepian-Wolf theorem, which suggest that we should be able to implement distributed coding of rates $(R_1, R_2) = \big(H(\boldsymbol{U}_1), H(\boldsymbol{U}_2|\boldsymbol{U}_1)\big) = (3,2)$ without inter-sensor

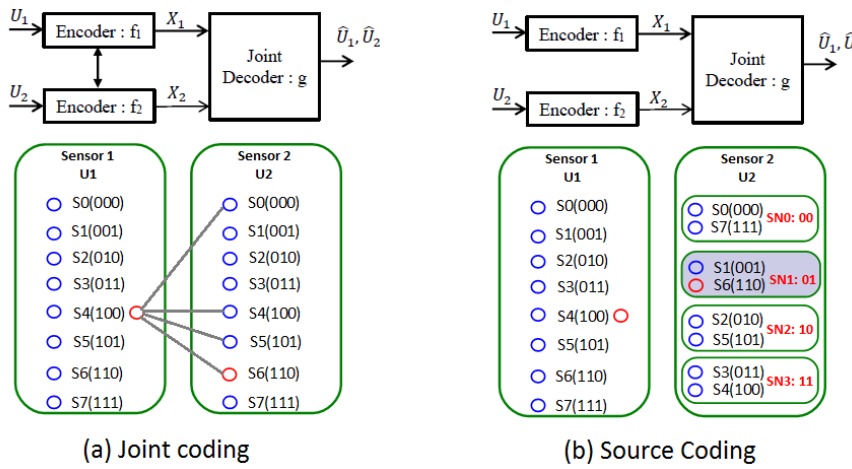


Figure 8-4 Joint coding versus distributed coding for correlated binary sources.

communication such that the receiver can perfectly recover $\boldsymbol{U}_1$ and $\boldsymbol{U}_2$. Consider the following implementation, where sensor 1 sends its plain reading U_1 using three bits as before. Sensor 2, however, divides the space of 8 possible outcomes into 4 disjoint partitions called cosets. Each coset SN_i encompasses two symbols with Hamming distance of 3. Each coset is mapped to a 2-bit sequence $C = [C_1, C_2]$ called syndrome as shown in Figure 8-4b. The receiver first identifies $\boldsymbol{U}_1$ from the received bits $[U_{11}U_{12}U_{13}]$, and them identifies the coset SN_i based on the received syndrome bits $C = [C_1, C_2]$. Each coset includes two symbols and the one with a lower Hamming distance to $\boldsymbol{U}_1$ is mapped to $\boldsymbol{U}_2$. This simple technique is a variant of a more general technique called Syndrome based distributed coding [58].

8.3.2 Review on Practical DSC Methods

Various practical implementations of DSC are proposed in the literature [11] [59] [60]. The first realization of DSC was introduced by Ramchandran, et. al. employing the concept of Syndrom based coding [58]. Afterwards, different implementations of DSCs were developed based on the more powerful channel codes including Low Density Parity Check Codes (LDPC) [19], turbo codes [35] [61] [62] [63] [64], turbo codes with pre-puncturing [65], turbo codes with post puncturing [37], Irregular Repeat Accumulate (IRA) codes [66] , and Low Density Generator Matrix Codes (LDGM) [67] [68].

Similar to point to point communications, combining the DSC stage with the subsequent channel coding stage to develop a single encoder called Distributed Joint Source Channel Coding (D-JSCC) is more efficient in terms of complexity and implementation cost. The most reported realizations of D-JSCCs are based on LDPC and turbo like codes [69] [70] [71] [63] [72]. However, these methods have not yet widely utilized in practice mainly due to implementation issues including the need for a prior knowledge of the observation model, the requirement of time-invariant observation noise, decoding complexity, sensitivity to a sensor failure and complication of scalability to a large number of sensors [11] [63] [36] [57] [71] [69].

In this chapter, we propose a D-JSCC method, which provides a superior error recovery performance compared to the state of the art D-JSCC methods and enjoys a simple encoder and low-complexity decoder structure. The proposed scheme is robust to sensor failures and continues to operate if a sufficient number of sensors are still operational. Due to the linear complexity of the proposed decoding algorithm in the number of sensors, this method is easily scalable to a large number of sensors. The bimodal operation eliminates

the unnecessary decoding iterations and reduces the complexity of the decoder by a factor of 10 to 20 in some cases. Finally, by employing a novel implementation of observation parameter extraction, no prior knowledge is required about the observation accuracy of sensors and the proposed system works even in time-varying observation models.

The following sections elaborate on the details of the proposed method and offer solutions for finding the optimal number of sensors, analyzing the convergence of the decoding algorithm and quantifying the performance of the system, when integrated with relaying techniques.

8.4 D-PCCC WITH BI-MODAL DECODER

The proposed scheme is based on utilizing a distributed version of parallel concatenation of convolutional codes (D-PCCC) in the sensors and an adaptive bimodal iterative decoder in the receiver as follows.

8.4.1 Encoder Structure

Each sensor is equipped with a pseudo-random interleaver, a RSC encoder and a puncturing module, which collectively form a distributed version of PCCC encoder as depicted in Figure 8-5. The rate of sensor i after puncturing the output codeword is R_i, where we set $R_i = R, \forall\, i \in \{1,2,..,L\}$ for homogeneous sensors unless otherwise specified. Here, the coding rate per sensor can be less than or greater than 1 in contrast to a typical individual encoder. An important distinction with the conventional PCCC scheme is that the input bits of each constituent encoder correspond to the noisy versions of common source bits, and hence may be slightly different.

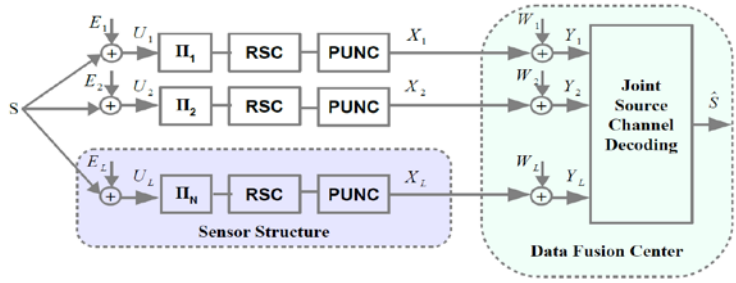


Figure 8-5 Coding Scheme: Each sensor includes an interleaver, a RSC encoder and a puncturing module, which collectively form D-PCCC scheme. Figure is from

We follow the BSC observation model, as described in section 8.2.1. We note that the independence of observation errors for different sensors implies that the pairwise correlation between the observations of sensors i and j is modeled as a BSC channel with cross over probability $\beta_{ij} = 1 + 2\beta_i\beta_j - \beta_i - \beta_j$, which simplifies to $1 + 2\beta^2 - 2\beta$ for equal observation error $\beta_1 = \beta_2 = \cdots = \beta_L = \beta$. For small values of β, we can use the $\beta_{ij} \approx 1 - 2\beta$ approximation [73] [74].

8.4.2 Decoder Structure

Inspired by the encoder structure, we develop a Multiple-branch Turbo Decoder (MTD) at the receiver with some modifications and additional features to accommodate the distributed nature of the encoder, as depicted in Figure 8-6.

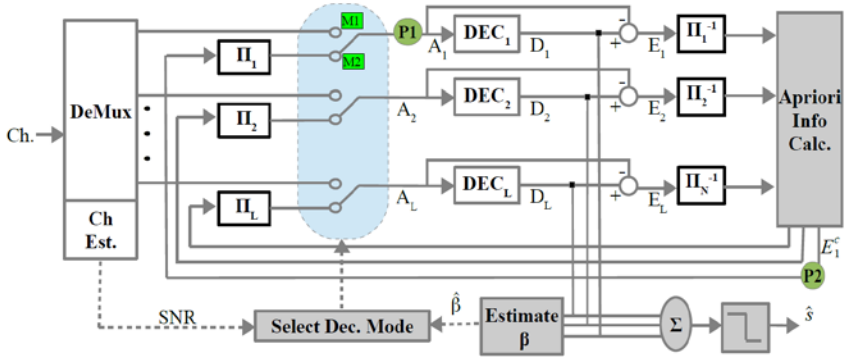


Figure 8-6 Proposed bi-modal parallel-structure MTD decoder equipped with an observation parameter extraction module. Figure is from *[83]*.

The decoder includes L constituent Soft-Input Soft-Output (SISO) decoders operate in parallel. Each constituent decoder executes the BCJR Max-Log-MAP algorithm using $log(e^x + e^y) \approx max(x, y)$ approximation for reduced complexity [75] that iteratively decodes the received bits in order to recover the relevant sensors observation bits [74].

Each constituent decoder receives a-priori LLRs denoted by $A_i^{(t)}(k)$ for $k = 1,2, \dots, N$ at iteration t and calculates the a-posteriori LLRs denoted by $D_i^{(t)}(k)$. Here, N is the length of dataword, i.e. the number of information bits. The extrinsic LLRs are defined as $E_i^{(t)}(k) = D_i^{(t)}(k) - A_i^{(t)}(k)$. In contrast to the standard MTD, here the information bits of each constituent decoder represent the corresponding sensor's observation bits, therefore the extrinsic LLRs converge to the observation bits rather than the common source bits, which

may cause convergence issues. The following is a list of modifications made to address this issue:

- **Decoder initialization:** In the first iteration, all SISO decoders are initialized with the LLRs that correspond to the observation bits of the respective sensor, which is in contrast with the classic MTD decoder, where only one RSC decoder is initialized with the LLRs of the common information bits.
- **LLR scaling:** In a classical MTD, the input LLRs for decoder i, $A_i^{(t)}(k)$ are calculated as the average of the extrinsic LLRs of all other RSC decoders (i.e. $A_i^{(t)}(k) = \sum_{j=1, j\neq i}^{N} E_j^{(t-1)}(k)$). In the proposed scenario, the extrinsic LLRs represent the corresponding sensor's observation bits. Therefore, we define a new set of extrinsic LLRs to represent the likelihood of the common source bits as follows:

$$E_{i,S}^{(t)}(n)\& = log_2\left(\frac{Pr(S(n)=1|\boldsymbol{U}_i)}{Pr(S(n)=0|\boldsymbol{U}_i)}\right) = log_2\left(\frac{\hat{\beta}+(1-\hat{\beta})\,2^{E_i^{(t)}(n)}}{(1-\hat{\beta})+\hat{\beta}2^{E_i^{(t)}(n)}}\right), \qquad (1)$$

which is the non-linearly scaled versions of the original extrinsic LLRs, $E_i^{(t)}(n)$. Here, $\boldsymbol{U}_i = [U_i(1), \dots, U_i(N)]$ is the observation vector of sensor i and $\hat{\beta}$ is the estimate of β. The a-priori LLRs of decoder i is then calculated as follows:

$$A_i^{(t)} = \sum_{\substack{j=1 \\ j\neq i}}^{N} E_{i,S}^{(t-1)} = \sum_{\substack{j=1 \\ j\neq i}}^{N} log_2\left(\frac{\hat{\beta}\ +\ (1-\hat{\beta})\,2^{E_i^{(t-1)}}}{(1-\hat{\beta})+\hat{\beta}2^{E_i^{(t-1)}}}\right)$$

$$= log\left[\prod_{\substack{j=1 \\ j\neq i}}^{N} \frac{\hat{\beta}+(1-\hat{\beta})\,2^{E_i^{(t-1)}}}{(1-\hat{\beta})+\hat{\beta}2^{E_i^{(t-1)}}}\right], \qquad (2)$$

The details of derivations are provided in [76].

- **Observation accuracy estimation**: In order to accommodate time-variant and unknown observation error parameter β, an observation error estimation module is developed to estimate β by the receiver. We first estimate the transmitted information symbol $\hat{x}_i(k)$ by hard-thresholding the output LLRs at the end of each iteration (i.e. $\hat{x}_i(k) = sign[D_i(k)]$). Then, we use the following estimator:

$$\hat{\beta} \approx \frac{1}{2LN(N-1)} \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j\neq i}}^{N-1} \sum_{k=1}^{L} \rho_{i,j}(k)\,,\ \rho_{i,j}(k) = \frac{|\hat{x}_i(k)-\hat{x}_j(k)|}{2}, \qquad (3)$$

which is based on estimating the pairwise correlation followed by averaging over all pairs of sensors. It is easy to show that $\hat{\beta}$ is normally distributed with mean β and variance $\frac{2\beta(1-\beta)(1-2\beta+2\beta^2)}{LN(N-1)} \approx \frac{2\beta}{LN(N-1)}$. Hence, it approaches the observation error parameter β for large enough N and L values.

- **Decision phase:** In the last iteration, hard decision is performed based on the average of a-posteriori LLRs, $D_{(av)} = \frac{1}{N}\sum_{i=1}^{N} D_i$ instead of the output LLRs of a particular SISO decoder D_i, for the similar reason of avoiding bias to a particular encoder/decoder pair.

- **Bimodal operation:** The last modification is using bimodal operation (iterative and non-iterative modes) as detailed in the following section.

8.4.3 Bimodal Decoding Operation

In a majority of the previously reported D-JSCC schemes for correlated sensors, iterative decoding is deployed in order to exploit the correlation among sensors' observations~ [11] [77] [71] [78] [63] [69], [79] [57] [80].

In this section, we revisit the presumption of superiority of iterative decoding for the proposed system by answering the following question, "*Under what conditions does the iterative decoding improve the estimation accuracy?*"

An intuitive conjecture is that iterative decoding is beneficial when the observation accuracy of sensors and consequently their pairwise correlations are high enough. In fact, the uselessness of iterative decoding, when the correlation among sensor observations is relatively low, is noticed by former researchers in several studies including [65] [37] [64]. To the best of our knowledge, however, no study has been conducted to characterize this phenomenon. In this section, we propose a method to obtain the convergence region of the decoder, where the iterative decoding is beneficial.

Consider the block diagram of the decoder depicted in Figure 8-6. The idea is to verify whether the estimates of the common source bits provided by a constituent decoder improve during one full iteration (from point P_1 to P_2 in Figure 8-6) or not. The technique we use here is a modified version of Extrinsic Information Transfer (EXIT) charts.

EXIT chart is a powerful technique to quantify the improvement of LLRs' relevance to the information bits during one decoding iteration by a

constituent decoder [81]. A typical EXIT chart for a standard turbo decoder ($\beta = 0$) is depicted in Figure 8-7. However, its conventional form is not applicable to our system model noting the fact that the extrinsic LLRs of each decoder converge to the corresponding sensor's observation, and not to the common source bits.

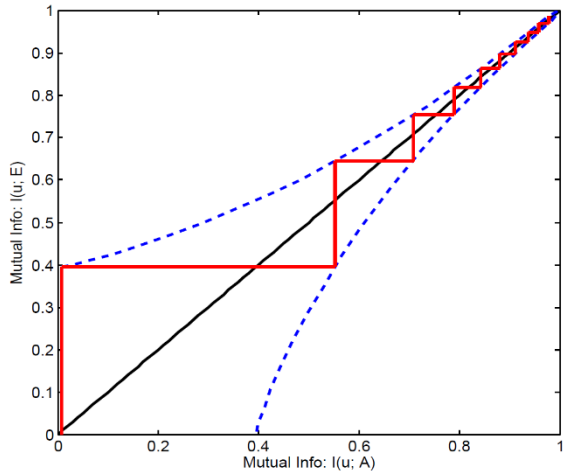


Figure 8-7 Conventional EXIT chart for the extreme case of complete observation accuracy ($\boldsymbol{\beta = 0}$, $\boldsymbol{E_b/N_0 = 1}$ dB).

Here, we provide the sketch of the modified EXIT chart technique that is used to find the convergence region. Interested users are referred to [82] and [83] for more details. To develop an EXIT chart, we note that the input LLRs received from the channel are Normally distributed, hence we have:

$$A^{(1)} = log\frac{Pr(X=+1|Y)}{Pr(X=-1|Y)} = \mu_Y\, Y + n_Y,$$
$$n_Y \sim \mathcal{N}(0,\sigma_Y^2), \quad \sigma_Y^2 = 2\mu_Y = 4/\sigma_N^2, \qquad (4)$$

where X and Y denote the BPSK modulated transmitted and received bits. It has been shown that if both channel observations and input LLRs follow Gaussian distribution, a MAP-family decoder with a fairly large frame length generates a-posteriori LLRs, which tend to follow a Gaussian distribution [84]. The intuitive justification is based on applying the weak law of large numbers to the summands over the random like decoder trellis structure. Moreover, extensive simulations confirm that relation (4) holds for the extrinsic LLRs as well [81]. Consequently, both a-priori and extrinsic LLRs, A and E, can be written in the following format:

$$E = \mu_E\, Y + n_E, \quad n_E \sim \mathcal{N}(0,\sigma_E^2), \quad \mu_E = \frac{\sigma_E^2}{2}, \qquad (5)$$

Now, if V denotes the BPSK modulated version of the source bit, the input LLR (A) exhibits the following bimodal distribution:

$$P_A(\zeta|v) = \sum_{x=-1,+1} P_A(\zeta|x)P(x|v) = \frac{1}{\sqrt{2\pi}\sigma_A}\left[\bar{\beta}e^{-\frac{(\zeta-\mu_A v)^2}{2\sigma_A^2}} + \beta e^{-\frac{(\zeta+\mu_A v)^2}{2\sigma_A^2}}\right], \quad \bar{\beta} = 1-\beta, \tag{6}$$

which is called the m^{th} order Binomial-Gaussian distribution with parameter set $(m = 2, \beta, \mu, \sigma^2)$ in this writing. This distribution for $\beta = 0$ boils down to a Gaussian distribution as expected. Consequently, the Bitwise-Mutual Information (BMI), between the a-priori LLRs and the source bits are calculated as:

$$I(A;\, V) = 1 - \frac{1}{\sqrt{2\pi}\sigma_A}\int_{-\infty}^{\infty}\left[\bar{\beta}e^{-\frac{(\zeta-\mu_A v)^2}{2\sigma_A^2}} + \beta e^{-\frac{(\zeta+\mu_A v)^2}{2\sigma_A^2}}\right] log\left(\frac{1+e^{-2\frac{\mu_A}{\sigma_A^2}\zeta}}{\bar{\beta}+\beta e^{-2\frac{\mu_A}{\sigma_A^2}\zeta}}\right)d\zeta. \tag{7}$$

For more than two constituent decoders ($m > 2$), an additional layer of averaging is required after calculating extrinsic LLRs. The resulting equations are more complex for $m > 2$ and are omitted here for the sake of brevity [83]. Note that $I(A;\, V)$ is a function of β and channel SNR. A typical $I(A;\, V)$ is depicted in Figure 8-8.

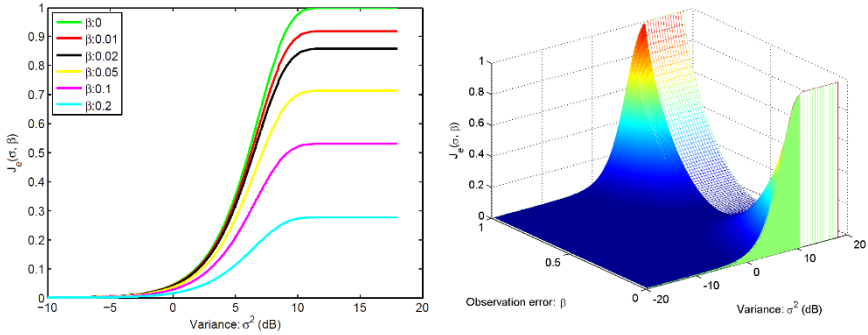


Figure 8-8 Bitwise mutual information between the a-priori LLRs and the source data as a function of noise variance σ^2 and observation error β. The left and right hand side figures present the 2D and 3D views. Figure is from *[76]*.

Similar equations hold for extrinsic LLRs. Note that in contrast to the conventional turbo decoder, here $I(A;V)$ and $I(E;V)$ do not approach 1, even for extremely large LLR values $\sigma_A \to \infty$, rather they approach $1 - H(\beta)$ [83].

In order to derive modified EXIT chart curves, we generate random source bits, and then we pass the source bits through a BSC channel with crossover probability β to obtain the simulated observation bits. Subsequently, we generate a-priori LLRs for the first iteration according to (4) for a given channel SNR. Then, we execute the decoding algorithm and obtain extrinsic LLRs (E) after proper scaling and averaging as mentioned in section 8.4.2. The parameters of extrinsic LLRs (μ_E, σ_E^2) are estimated afterwards by fitting a Gaussian distribution. Finally, we calculate $I(A;V)$ and $I(E;V)$ using (7).

Plotting $I(E;V)$ vs $I(A;V)$ as well as the reverse curves, i.e. $I(A;V)$ vs $I(E;V)$ presents the modified EXIT charts as depicted in Figure 8-9. If the direct curve falls above the unit-slope line and presents a positive initial slope, then iterative decoding is beneficial, since the BMI between the LLRs and the source bits exhibit improvement for a decoding cycle. The point at which the direct and reverse curves meet presents the final achievable error recovery performance and defines the number of useful iterations. The closer this point to the upper right corner $[I(A;V), I(E;V)] = [1,1]$, the lower the error floor. The general trend obtained from this figure is that lower channel SNRs and higher observation accuracies are suggestive of more advantage for iterative decoding.

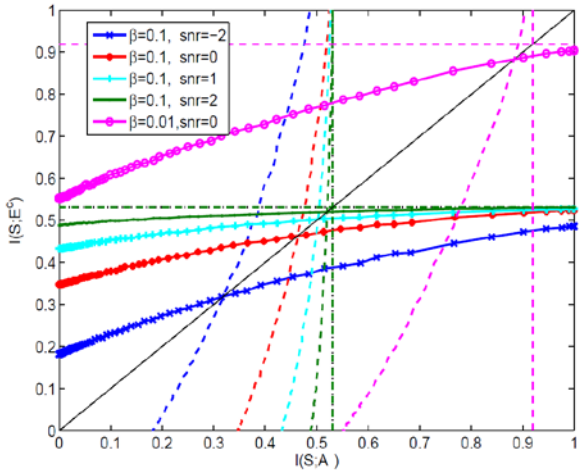


Figure 8-9 Modified EXIT charts for different observation accuracies (Number of sensors is $L = 2$). Figure is from *[83]*.

In order to find the convergence regions, we develop (β, SNR) grids and generate EXIT charts for all combinations of (β, SNR). Then, we identify the conditions that yield a positive initial slope of ϵ for the EXIT chart direct curve. A sample convergence region is shown in Figure 8-10.

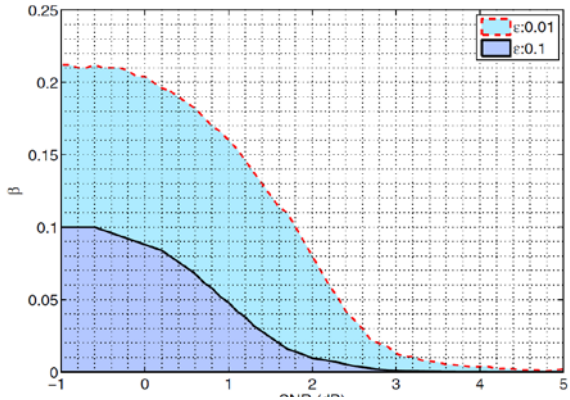


Figure 8-10 Convergence region of iterative decoding algorithm in terms of channel SNR and β for ε = 0.1 and ε = 0.01. Here ε is the level of minimum improvement due to iterative decoding. Outside the convergence region, non-iterative algorithm is selected. Figure is from [83]**.**

We note that developing convergence region for the (β, SNR) grid is a time consuming task, but it is performed only once per system and is revealed to the decoder. Once the decoder receives a new multi-frame from the sensors, it estimates the channel SNR and the observation accuracy parameter β. If the estimated (SNR, β) pair falls within the convergence region, the iterative mode is chosen, otherwise LLR exchange is bypassed to avoid unnecessary iterations.

8.5 PERFORMANCE ANALYSIS

To confirm the performance improvement obtained by the mode selection mechanism, we performed extensive Monte Carlo simulations for the proposed scheme (D-PCCC) with the following parameters ($M = 2000{\sim}bits, N = 8$ and $\beta = 0.1$). We compare the results against similar representative methods including distributed turbo coding in [71] and D-LDGM codes in [57] under equivalent conditions.

The error probability of AWGN channel $Pr(error) = Q(\sqrt{2SNR})$ is used to approximate the BSC channel used in [71] with an equivalent AWGN channel. The energy per source bit is calculated as $\frac{E_b}{N_0} = \frac{N}{R_i} E_s/N_0$. Equivalent coding rate, $R_{eq} = \frac{N}{R_i}$ is used to make a fair comparison choices between different schemes, where N is the number of sensors.

The comparison results show that in the low observation accuracy scenario, $\beta = 0.1$, the proposed scheme performs close to the more complex LDGM codes, with performance degradation of less than 1dB. The superiority of the proposed coding scheme over the classical DTC is shown as well. This performance improvement is achieved with lower decoding complexity, when the decoder switches to the non-iterative mode. The order of decoding complexity reduction with respect to classical DTC can be as high as the typical number of decoding iterations, which is 10 to 20.

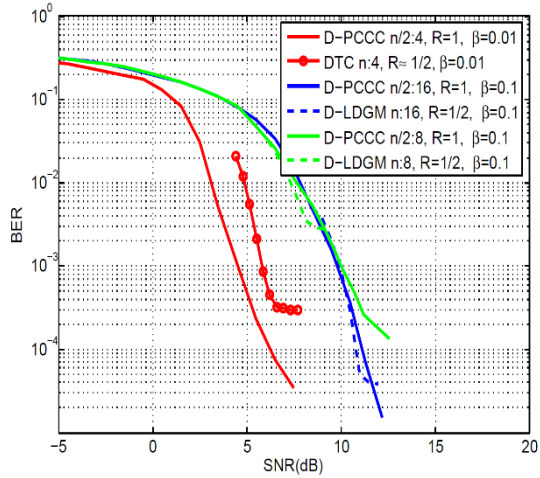


Figure 8-11 BER performance comparison of the proposed scheme with the state of the art D-JSCC methods.

8.5.1 Optimal number of sensors

Apparently, using a higher number of sensors at each cluster reduces the estimation error, but on the other hand, it increases the bandwidth usage and the decoding complexity. A key question is *"What is the minimum number of sensors that yields a desired performance?"* In order to answer this question, we provide an approximate information-theoretic analysis. We note that the rate region of the bCEO problem even for the simplest case of two sensors is not yet characterized in terms of a closed-form expression [39] [54]. Therefore, the proposed simple method is valuable in practical system design.

In this regard, we first note that the virtual channel from the common source to the final destination is the cascade of a broadcast BSC channel and a set of parallel AWGN channels. It can be shown that for a cluster of N sensors

with transmit power P, the observation error β and the noise variance σ_N^2, the capacity of this virtual channel is [55]:

$$C = \frac{1}{2(2\,\pi\,\sigma_N^2)^{\frac{N}{2}}} \int_{y^n} \left[(\gamma_0 + \gamma_1)\, log\left(\frac{\gamma_0 + \gamma_1}{2}\right) - \gamma_0\, log\, \gamma_0 - \gamma_1\, log\, \gamma_1 \right] d_{y^N}, \quad (8)$$

where we have:

$$\gamma_\alpha = \sum_{k=0}^{N} \binom{N}{k} \beta^k (1 - \beta)^{N-k}\, exp\left(- \frac{\sum_{i=0}^{k} \left(y_i + (2\alpha - 1)\sqrt{P}\right)^2 + \sum_{i=k+1}^{N} \left(y_i - (2\alpha - 1)\sqrt{P}\right)^2}{2\sigma_N^2} \right) \quad (9)$$

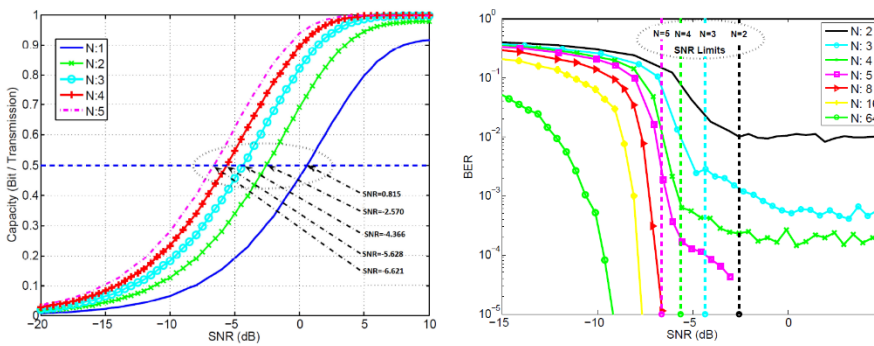
Therefore, the channel capacity is $C = f_c\,(N, \frac{P}{\sigma_N^2}, \beta)$, which is a closed-form function of the number of sensors (N), the channel SNR ($\frac{P}{\sigma_N^2}$) and the observation error parameter (β). Now, note the fact that the BER floor is fully defined by the number of sensors in an error-free environment as follows:

$$P_{error}^{(min)} = \begin{cases} \sum_{k=\frac{N+1}{2}}^{N} \binom{N}{k} \beta^k + (1-\beta)^{N-k} & N\ odd \\ \frac{1}{2}\binom{N}{\frac{N}{2}} \beta^{\frac{N}{2}} (1-\beta)^{\frac{N}{2}} + \sum_{k=\frac{N}{2}+1}^{N} \binom{N}{k} \beta^k + (1-\beta)^{N-k} & N\ even \end{cases} \quad (10)$$

Based on the channel-coding theorem, there exist an encoder with rate $R \leq C$ that can be employed by the source for an error-free transmission of information bits to the common destination. Now, we conjecture that if we employ a proper coding technique of rate $R \leq C$ at the sensors, the final recovery error approaches the error floor, which is the best achievable error probability. Based on this argument, for a given set of system conditions $\beta, \frac{P}{\sigma_N^2}$, R, we can obtain the optimal number of sensors N by solving $R = f_c(N, \frac{P}{\sigma_N^2}, \beta)$. This approach is demonstrated in Figure 8-12.

In Figure 8-12a, the channel capacity is plotted for a fixed observation error parameter $\beta = 0.01$. The capacity is an increasing function of the channel SNR. If a certain value of capacity is desired, the minimum required SNR level is less for a higher number of utilized sensors. For instance, to achieve a capacity of $1/2$ bits/transmission, the required SNR values for *N=2,3,4,5*

sensors are -2.570 dB, -4.366 dB, -5.628 dB, and -6.621 dB, respectively. In another interpretation, for power constrained sensors (limited SNRs), this graph can be used to determine the minimum number of sensors to achieve a certain level of capacity. For instance, at least $N=4$ sensors are required to achieve a capacity of $1/2$ bits/transmission at SNR=-5 dB. This is verified by extensive Monte Carlo simulations for ($\beta = 0.01, N = 256$ *bits and* $R_i = \frac{1}{2}$ for different number of sensors as depicted in Figure 8-12b, where the decoder reaches its error floor at SNR=-5 dB if at least 4 sensors are utilized in the system.



(a) Channel Capacity vs SNR, $\beta = 0.01$. (b) Decoder BER Performance

Figure 8-12 a) Information capacity of system vs observation accuracy (BSC crossover probability) and channel quality (SNR) for different number of sensors. b) The BER performance of the proposed coding scheme with different number of sensors. In both figures, $\beta = 0.01$. Figures are from *[55]***.**

8.6 EXTENSION TO MULTI-HOP NETWORKS

So far, a collocated network scenario is considered, where all sensors directly communicate to a data fusion center. This scenario does not cover many applications, where the use of multi-hop communication is unavoidable due to a long distance between the source and data fusion center. The question here is how to extend the obtained results to such cases.

In order to address this requirement, we propose to use the idea of inner channels by replacing the multi-hop channel from each sensor to the fusion center as a BSC channel with an equivalent crossover probability.

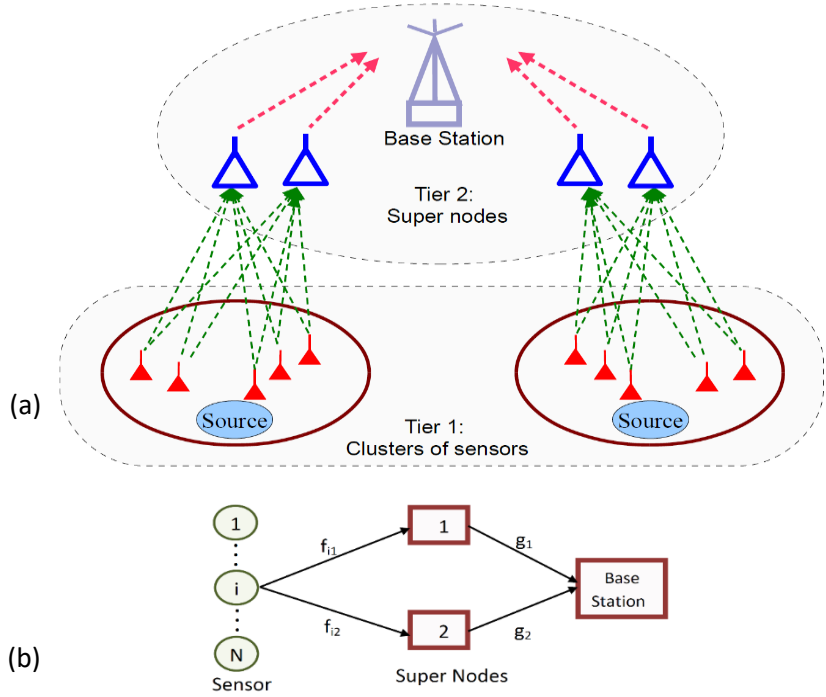


Figure 8-13 a) System model for two-tiered double-sink wireless sensor network. b) Channel coefficients for the communication links from sensors to the base station via two super nodes.. Figure is from *[85]***.**

For instance, in [85], a two-tiered network, where the sensors at each cluster communicate to the base station through two intermediate relay nodes is considered (Figure 8-13a). Utilizing two super nodes (also called cluster heads) as relay nodes has an obvious advantage of robustness to cluster head failure in practice. A Demodulate and Forward (DMF) relaying method assisted by distributed version of Space Time Block Codes (D-STBC) is utilized in order to improve the BER performance and overcome the potential fading effects.

Suppose that the total transmission power P is divided between the two tiers (sensors and super nodes) with a ratio of α. As such, each sensor consumes $\alpha \frac{P}{N}$ and each relay node consumes $(1-\alpha)\frac{P}{2}$. The inner channel error probability can be calculated as follows [85]:

$$p_e^{(in)} = \frac{1}{4(\alpha\gamma+1)} + \frac{1}{2(\eta\gamma+2)^2} - \frac{1}{4(\alpha\gamma+1)(\eta\gamma+2)^2}, \qquad (11)$$

where we have:

$$\gamma = \frac{P}{N}, \quad \gamma_1 = \alpha\,\gamma, \qquad \gamma_2 = \eta\,\gamma, \qquad \eta = \frac{3\alpha\bar{\alpha}}{2\,\alpha + 1}. \tag{12}$$

Figure 8-14 presents the impact of power allocation on inner channel error probability. Despite STBC assisted Amplify and Forward (AF) relaying, where the optimality achieved if the total power is equally divided between the two tiers (sensors and relays), the performance curve of DMF relaying mode is not symmetric. Dashed and solid lines in this figure present analytical and simulation results for the resulting error probability as a function of power allocation parameter α. The optimal power allocation scenario can be chosen based on the average SNR value of the system, $\frac{P}{N_1}$.

We complete this section by comparing the end-to-end error probability for the basis scenario with one and two relay nodes (cluster heads) with and without STBC coding in presence of channel fading effect in Figure 8-15.

It is shown that using two super-nodes increases the BER performance by about 2∼3 dB due to the space diversity gain. Also, about 1 dB additional gain is obtained using D-STBC due to the space-time diversity. All schemes ultimately reach the error floor for extremely large SNR values, which is corresponding to the error-free communication as calculated in (10).
We finally note that the concept of inner channel is general and similar performance analysis and power allocation policies can be applied to other network configurations.

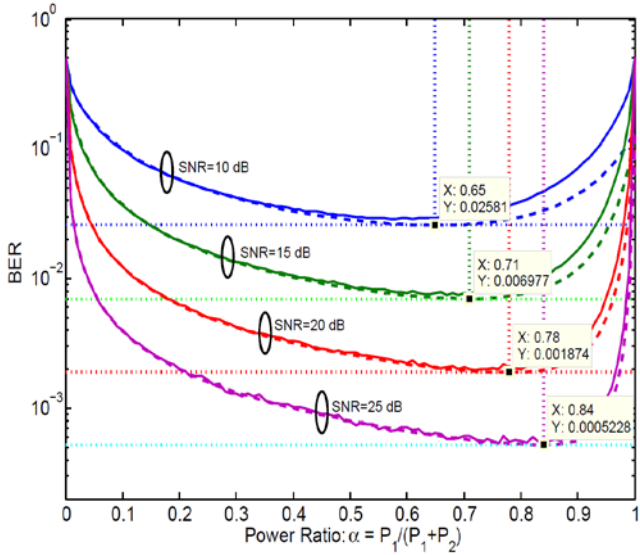


Figure 8-14 Inner channel error probability vs power allocation parameter α. Figure is from [85].

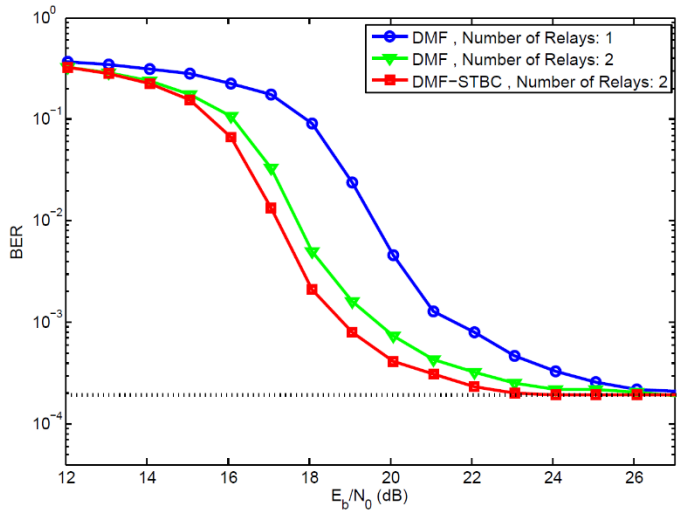


Figure 8-15 Comparison of system performance for different number of super-nodes, with and without STBC coding at super-nodes ($N = 4$). Figure is from *[85]*.

8.7 DISCUSSIONS AND FUTURE DIRECTIONS

This chapter explores various algorithms designed by the coding research community to realize compression and error recovery in WSNs with partially accurate and remotely deployed sensors. Despite their promising theoretical performances, most of these algorithms are not well suited for practical applications due to a number of drawbacks. In this regard, a practical and implementation-friendly solution is reviewed for this problem.

Previously reported algorithms, as discussed in section 8.3, are not easily scalable to a large number of sensors. However, employing a large number of sensors at each cluster is unavoidable when the observation accuracy of sensors are low. Fusing multiple sensor readings compensate for the observation errors that occur prior to encoding. Therefore, there is an essential need for an easily scalable distributed joint source-channel coding algorithm. This solution should be flexible enough to accommodate dynamic clustering with varying number of moving sensors at each cluster. Further, the developed algorithm should be capable of accommodating time varying and unknown observation models and should be easily integrated with the current sensor structures and contemporary networking protocols. To address these requirements, in section 8.4 an adaptive algorithm is presented for efficient data collection across a data field, where source locations are inaccessible due to either unknown source locations or harsh environmental conditions [86].

The commonly accepted presumption of superiority of iterative decoding over non-iterative decoding is revisited in section 8.4.3. Indeed, the proposed convergence analysis, developed based on EXIT chart technique, suggests that in some system conditions, the soft information exchange among constituent decoders at the destination do not improve the overall end-to-end error rate. Therefore, avoiding these unnecessary information exchange cycles reduces the average decoding complexity [82] [83]. Further, in section 8.5.1, an easy-to-evaluate information-theoretic function is derived to determine the minimum number of sensors to reach the highest achievable error recovery performance [55].

To generalize the proposed scheme to large-scale networks, a two-tiered clustered system model is considered in section 8.6, where each cluster includes a single data source, which is remotely monitored by surrounding sensors. The sensors within a cluster collectively compress and transmit their observations to a data fusion center via two super-nodes. This two-tiered network model is robust to super-node failure in contrast with the traditionally designed single super-node systems [73]. The set of two-hop communication channels from the sensors to the data fusion center via two

super-nodes is modeled as inner channels. This enabled us to analyze the end-to-end performance of the whole system by evaluating the distributed coding operation over inner channels. The available total power is optimally split to the existing sensor and super nodes such that the overall data throughput is maximized [85]. This system demonstrates the easy integration of our proposed coding scheme with desired relaying methods.
This book chapter suggests several directions as continuation of the current studies as follows:

- Designing and optimizing the coding parameters for the more general case of unequal observation parameters $\beta_i \neq \beta$ and unequal channel qualities $SNR_i \neq SNR$.
- The proposed scheme is robust to sensors failure and stays operational as long as one or some of the sensors are active. However, each sensor failure degrades the overall end-to-end BER performance. This performance degradation can be compensated by changing the coding rate of other active sensor. This method can be studied as another possible extension of this work, which involves quantification of the BER performance degradation due to a sensor failure and finding the new coding rate to compensate for this loss.
- Developing more elegant learning algorithms (such as reinforcement learning) in order to tune the required coding rates without the need for estimating the sensors measurement accuracies and channel qualities can simplify the implementation of sensors.
- A preliminary study is conducted in [86] to interrogate multiple SAW based passive sensors using a single interrogator. If all the passive sensors observe a common data source, similar techniques can be developed for passive sensors to harness the correlation among sensor observations to enhance the final estimation accuracy.

REFERENCES

[1] R. Schmidt, J. R. Dickman and D. R. Kiracofe, "Engine Health Monitoring Using Acoustic Sensors". USA Patent 15/063,582, 14 September 2017.

[2] E. Dudzik, A. Abedi, D. Hummels and M. d. Cunha, "Orthogonal code design for passive wireless sensors," in IEEE Biennial Symposium on Communications, 2008.

[3] J. Zhang, G. Y. Tian, A. M. Marindra, A. I. Sunny and A. B. Zhao, "A review of passive RFID tag antenna-based sensors and systems for structural health monitoring applications," Sensors, vol. 17, no. 2, p. 265, 2017.

[4] Z. Michael, E. Anagnostou, J. Nystuen and M. Anagnostou, "UPAL: Underwater passive aquatic listener," in MTS/IEEE OCEANS'15, Washington, 2015.

[5] K. Krishen, "Space applications for ionic polymer-metal composite sensors, actuators, and artificial muscles," Acta Astronautica, vol. 64, no. 11, pp. 1160-1166, 2009.

[6] E. Berkcan and Y. Lee, "Passive wireless sensors for turbomachines and method of operating the same". Patent 9,909,443, 6 March 2018.

[7] J. Hodgson, R. Mott, S. M. Baylis, T. T. Pham, S. Wotherspoon, A. D. Kilpatrick, R. R. Segaran, I. Reid, A. Terauds and L. P. Koh, "Drones count wildlife more accurately and precisely than humans," Methods in Ecology and Evolution, 2017.

[8] T. Li, J. M. Corchado, S. Sun and J. Bajo, "Clustering for filtering: Multi-object detection and estimation using multiple/massive sensors," Information Sciences, vol. 388, pp. 172-190, 2017.

[9] R. Gravina, P. Alinia, H. Ghasemzadeh and G. Fortino, "Multi-sensor fusion in body sensor networks: State-of-the-art and research challenges," Information Fusion, vol. 35, pp. 68-80, 2017.

[10] Z. Xiong, A. D. Liveris and S. Cheng, "Distributed source coding for sensor networks," IEEE signal processing magazine , vol. 21, no. 5, pp. 80-94, 2004.

[11] J. Garcia-Frias and Z. Xiong, "Distributed source and joint source-channel coding: from theory to practice," in Proceedings IEEE Acoustics, Speech, and Signal Processing Conference (ICASSP), 2005.

[12] J. Hill, M. Horton, R. Kling and L. Krishnamurthy, "The platforms enabling wireless sensor networks," ACM Communications, vol. 47, no. 6, pp. 41-46, Jun 2004.

[13] Cricket v2 User Manual, Cambridge, Lab, MIT Computer Science \& Artificial Intelligence, 2005.

[14] A. Rowe, R. Mangharam and R. Rajkumar, "FireFly: A Time Synchronized Real-Time Sensor Networking Platform," in Wireless Ad Hoc Networking: Personal-Area, Local-Area, and the Sensory-Area Networks, CRC Press Book Chapter, 2006.

[15] R. Mangharam, A. Rowe and R. Rajkumar, "FireFly: A Cross-Layer Platform for Wireless Sensor Networks," Real Time Systems Journal, Special Issue on Real-Time Wireless Sensor Networks, vol. 37, no. 3, pp. 181-231, Dec. 2007.

[16] T. Berger and Z. Zhang, "On the CEO problem," in IEEE International Symposium Information Theory (ISIT), 1994.

[17] Y. Oohama, "Indirect and Direct Gaussian Distributed Source Coding Problems," IEEE Transactions on Information Theory, vol. 60, no. 12, pp. 7506-7539, December 2014.

[18] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," IEEE Transactions on Information Theory, vol. 19, no. 4, pp. 471-408, July 1973.

[19] A. Liveris, Z. Xiong and C. Georghiades, "Compression of binary sources with side information at the decoder using LDPC codes," IEEE Communication Letters, vol. 6, no. 10, pp. 440-442, October 2002.

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

[21] Y. Li, "Distributed coding for cooperative wireless networks: An overview and recent advances," IEEE Communication Magazine, vol. 47, no. 8, pp. 71-77, August 2009.

[22] M. Breton and B. Kovatchev, "Analysis, Modeling, and Simulation of the Accuracy of Continuous Glucose Sensors," Journal of Diabetes Science Technology, vol. 2, no. 5, pp. 835-862, September 2008.

[23] K. Petrosyants and N. I. Rjabov, "Temperature sensors modeling for smart power ICs," in the 27th Annual IEEE Semiconductor Thermal Measurement and Management Symposium (SEMI-THERM), 2011.

[24] P. Robins, V. Rapley and P. Thomas, "A review of wireless SAW sensors," vol. 47, no. 2, pp. 317-332, 2000.

[25] C. Mendis, A. Skvortsov, A. Gunatilaka and S. Karunasekera, "Performance of Wireless Chemical Sensor Network With Dynamic Collaboration," IEEE Sensors Journal, vol. 12, no. 8, pp. 2630-2637, 2012.

[26] O. Tigli and M. Zaghloul, "Surface acoustic wave SAW biosensors," in the 53rd IEEE International Midwest Symposium on Circuits and Systems (MWSCAS), 2010.

[27] Y. Hovakeemian, K. Naik and A. Nayak, "A survey on dependability in Body Area Networks," in the 5th International Symposium on Medical Information Communication Technology (ISMICT), 2011.

[28] P. Bylanski, "PCM A-Law decoder using a circulating method," Electronics Letters, vol. 12, no. 2, pp. 57-58, 1976.

[29] R. Gray, "Vector quantization," IEEE ASSP Magazine, vol. 1, no. 2, pp. 4-29, 1984.

[30] P. Bylanski, "Signal-processing operations on A-Law companded speech," Electronics Letters, vol. 15, no. 12, pp. 697-698, 1979.

[31] P. Dvorsky, J. Londak, O. Labaj and P. Podhradsky, "Comparison of codecs for videoconferencing service in NGN," in ELMAR, 2012.

[32] Y. Linde, A. Buzo and R. Gray, "An Algorithm for Vector Quantizer Design," IEEE Transactions on Communications, vol. 28, no. 1, pp. 84-95, 1980.

[33] R. Gray and D. Neuhoff, "Quantization," IEEE Transactions on Information Theory, vol. 44, no. 6, pp. 2325-2383, 1998.

[34] Z. Krusevac, P. Rapajic and R. Kennedy, "Concept of time varying binary symmetric model-channel uncertainty modeling," in International Conference on Communications Systems (ICCS), 2004.

[35] A. Aaron and B. Girod, "Compression with side information using turbo codes," in Data Compression Conference (DCC), 2002.

[36] J. Garcia-Frias and Y. Zhao, "Compression of binary memoryless sources using punctured turbo codes," IEEE Communication Letters, vol. 6, no. 9, pp. 394-396, September 2002.

[37] A. Liveris, Z. Xiong and C. Georghiades, "Distributed compression of binary sources using conventional parallel and serial concatenated convolutional codes," in Data Compression Conference (DCC), 2003.

[38] Z. Tu, J. Li and R. Blum, "Compression of a binary source with side information using parallelly concatenated convolutional codes," in IEEE Global Telecommunications Conference (GLOBECOM), 2004.

[39] X. He, X. Zhou, M. Juntti and T. Matsumoto, "A Rate-Distortion Region Analysis for a Binary CEO Problem," in IEEE 83rd Vehicular Technology Conference (VTC Spring), Nanjing, 2016.

[40] T. Berger, Z. Zhang and H. Viswanathan, "The CEO problem [multiterminal source coding]," IEEE Transactions on Information Theory, vol. 42, no. 3, pp. 887-902, May 1996.

[41] S. Y. Tung, "Multiterminal source coding," Ph.D. dissertation, School of Electrical Engineering, Cornell University, Ithaca, NY, May 1998.

[42] H. Viswanathan and T. Berger, "The quadratic Gaussian CEO problem," IEEE Transactions on Information Theory, vol. 43, no. 5, pp. 1549-1559, September 1997.

[43] 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 problem," IEEE J. Sel. Area Communication, vol. 22, no. 6, pp. 977-987, August 2004.

[44] T. A. Courtade, "Outer Bounds for Multiterminal Source Coding based on Maximal Correlation," CoRR, 2013.

[45] Y. Yang, Y. Zhang and Z. Xiong, "On the Sum-Rate Loss of Quadratic Gaussian Multiterminal Source Coding," IEEE Transactions on Information Theory, vol. 57, no. 9, pp. 5588-5614, 2011.

[46] Y. Yang, Y. Zhang and Z. Xiong, "The generalized quadratic Gaussian CEO problem: New cases with tight rate region and applications," in IEEE International Symposium on Information Theory Proceedings (ISIT), 2010.

[47] Y. Yang and Z. Xiong, "The Sum-Rate Bound for a New Class of Quadratic Gaussian Multiterminal Source Coding Problems," IEEE Transactions on Information Theory, vol. 58, no. 2, pp. 693-707, 2012.

[48] Y. Yang and Z. Xiong, "On the Generalized Gaussian CEO Problem," IEEE Transactions on Information Theory, vol. 58, no. 6, pp. 3350-3372, jun 2012.

[49] J. Chen and T. Berger, "Robust Distributed Source Coding," IEEE Transactions on Information Theory, vol. 54, no. 8, pp. 3385-3398, 2008.

[50] S. avildar and P. Viswanath, "On the Sum-rate of the Vector Gaussian CEO Problem," in Thirty-Ninth Asilomar Conference on Signals, Systems and Computers, 2005.

[51] G. Zhang and W. Kleijn, "Bounding the Rate Region of the Two-Terminal Vector Gaussian CEO Problem," in Data Compression Conference (DCC), 2011.

[52] S.-C. Lin and H.-J. Su, "Vector Wyner-Ziv coding for vector Gaussian CEO problem," in the 41st Annual Conference on Information Sciences and Systems (CISS), 2007.

[53] J. Chen and J. Wang, "On the vector Gaussian CEO problem," in IEEE International Symposium Information Theory (ISIT), 2011.

[54] X. He, X. Zhou, P. Komulainen, M. Juntti and T. Matsumoto, "A Lower Bound Analysis of Hamming Distortion for a Binary CEO Problem With Joint Source-Channel Coding," IEEE Transactions on Communications, vol. 64, no. 1, pp. 343-353, January 2016.

[55] A. Razi, K. Yasami and A. Abedi, "On Minimum Number of Wireless Sensors Required for Reliable Binary Source Estimation," in IEEE

Wireless Communication and Networking Conference (WCNC), 2011.

[56] J. Del Ser, J. Garcia-Frias and P. Crespo, "Iterative concatenated Zigzag decoding and blind data fusion of correlated sensors," in International Conference on Ultra Modern Telecommunications Workshops (ICUMT), 2009.

[57] W. Zhong and J. Garcia-Frias, "Combining data fusion with joint source-channel coding of correlated sensors," in IEEE Information Theory Workshop (ITW), 2004.

[58] S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (DISCUS): design and construction," in Data Compression Conference (DCC), 1999.

[59] S. Qaisar and H. Radha, "Multipath multi-stream distributed reliable video delivery in Wireless Sensor Networks," in the 43th Annual Conference Information Sciences Systems (CISS), 2009.

[60] I. Esnaola and J. Garcia-Frias, "Analysis and optimization of distributed linear coding of Gaussian sources," in the 43th Annual Conference Information Sciences Systems (CISS), 2009.

[61] J. Bajcsy and P. Mitran, "Coding for the Slepian-Wolf problem with turbo codes," in IEEE Global Telecommunications Conference (GLOBECOM), 2001.

[62] M. Valenti and B. Zhao, "Distributed turbo codes: towards the capacity of the relay channel," in the 58th IEEE Vehicular Technology Conference (VTC), 2003.

[63] F. Daneshgaran, M. Laddomada and M. Mondin, "Iterative joint channel decoding of correlated sources," IEEE Transactions on Wireless Communication, vol. 5, no. 10, pp. 2659-2663, October 2006.

[64] J. Garcia-Frias, Y. Zhao and W. Zhong, "Turbo-Like Codes for Transmission of Correlated Sources over Noisy Channels," IEEE Signal Processing Magazine, vol. 24, no. 5, pp. 58-66, September 2007.

[65] A. Liveris, Z. Xiong and C. Georghiades, "A distributed source coding technique for correlated images using turbo-codes," IEEE Communications Letters, vol. 6, no. 9, pp. 379-381, September 2002.

[66] V. Stankovic, A. Liveris, Z. Xiong and C. Georghiades, "On code design for the Slepian-Wolf problem and lossless multiterminal networks," IEEE Transactions on Information Theory, vol. 52, no. 4, pp. 1495-1507, April 2006.

[67] M. Hernaez, P. Crespo, J. Del Ser and J. Garcia-Frias, "Serially-concatenated LDGM codes for correlated sources over gaussian broadcast channels," IEEE Communication Letters, vol. 13, no. 10, pp. 788-790, October 2009.

[68] C. Stefanovic, D. Vukobratovic and V. Stankovic, "On distributed LDGM and LDPC code design for networked systems," in IEEE Information Theory Workshop (ITW), 2009.

[69] M. Sartipi and F. Fekri, "Source and channel coding in wireless sensor networks using LDPC codes," in IEEE Conference Sensor and Ad Hoc Communication and Networks (SECON), 2004.

[70] A. Yedla, H. Pfister and K. Narayanan, "Can iterative decoding for erasure correlated sources be universal?," in 47th Annual Allerton Conference Communication, Control, and Computing, 2009.

[71] J. Haghighat, H. Behroozi and D. Plant, "Joint decoding and data fusion in wireless sensor networks using turbo codes," in the 19th IEEE International Symposium Personal, Indoor and Mobile Radio Communication (PIMRC, 2008.

[72] J. Garcia-Frias, "Joint source-channel decoding of correlated sources over noisy channels," in Data Compression Conference (DCC), 2001.

[73] A. Razi, F. Afghah and A. Abedi, "Binary source estimation using two-tiered sensor network," IEEE Communication Letter, vol. 5, no. 4, pp. 449-451, April 2011.

[74] A. Razi and A. Abedi, "Distributed Coding of Sources with Unknown Correlation Parameter," in Proceedings International Conference Wireless Networking (ICWN), Las Vegas, 2010.

[75] R. Ghaffar and R. Knopp, "Analysis of Low Complexity Max Log MAP Detector and MMSE Detector for Interference Suppression in Correlated Fading," in IEEE Global Telecommunications Conference, (GLOBECOM), 2009.

[76] A. Razi, "Distributed Adaptive Algorithm Design For Joint Data Compression And Coding In Dynamic Wireless Sensor Networks," PhD Thesis, University of Maine, May 2013.

[77] A. Liveris, Z. Xiong and C. Georghiades, "Joint source-channel coding of binary sources with side information at the decoder using IRA codes," in IEEE Workshop Multimedia Signal Processing, 2002.

[78] J. Haghighat, H. Behroozi and D. Plant, "Iterative joint decoding for sensor networks with binary CEO model," in the 9th IEEE Workshop Signal Processing Advances in Wireless Communication (SPAWC) , 2008.

[79] D. Gunduz, E. Erkip, A. Goldsmith and H. Poor, "Source and Channel Coding for Correlated Sources Over Multiuser Channels," IEEE Transactions on Information Theory, vol. 55, no. 9, pp. 3924-3944, sEPTEMBER 2009.

[80] R. Maunder, J. Wang, S. Ng, L.-L. Yang and L. Hanzo, "On the Performance and Complexity of Irregular Variable Length Codes for Near-Capacity Joint Source and Channel Coding," IEEE Transactions

on Wireless Communication, vol. 7, no. 4, pp. 1338-1347, aPRIL 2008.

[81] S. Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," IEEE Transactions on Communication, vol. 49, no. 10, pp. 1727 -1737, October 2001.

[82] A. Razi and A. Abedi, "Adaptive Bi-modal Decoder for Binary Source Estimation with Two Observers," in the 46th Annual Conference Information Science Systems (CISS), Princeton, 2012.

[83] A. Razi and A. Abedi, "Convergence Analysis of Iterative Decoding for Binary CEO Problem," IEEE Transactions on Wireless Communications, vol. 13, no. 5, pp. 2944-2954, May 2014.

[84] N. Wiberg, H.-A. Loeliger and R. Kotter, "Codes and iterative decoding on general graphs," in IEEE International Symposium Information Theory (ISIT), 1995.

[85] A. Razi, F. Afghah and A. Abedi, "Power Optimized DSTBC Assisted DMF Relaying in Wireless Sensor Networks with Redundant Super nodes," IEEE transactions on Wireless Communications, vol. 12, no. 2, pp. 636-645, February 2013.

[86] A. Razi and A. Abedi, "Interference reduction in Wireless Passive Sensor Networks using directional antennas," in 4th Annual IEEE/Caneus Fly by Wireless Workshop (FBW), Montreal, CA, 2011.