Convergence Analysis of Iterative Decoding for Binary CEO Problem

Article *in* IEEE Transactions on Wireless Communications · May 2014
DOI: 10.1109/TWC.2014.041014.130921

2 authors, including:

Abolfazl Razi
Northern Arizona University
45 PUBLICATIONS **199** CITATIONS

Convergence Analysis of Iterative Decoding for Binary CEO Problem

Abolfazl Razi, *Member, IEEE,* and Ali Abedi, *Senior Member, IEEE*

***Abstract*—Estimation of a binary source using multiple observers, a variant of the so called Chief Executive Officer (CEO) problem, is considered. A low-complexity Distributed Joint Source Channel Coding (D-JSCC) based on the Parallel Concatenated Convolutional Codes (PCCC) is implemented in a cluster of sensors in a distributed fashion. Convergence of the iterative decoder is analyzed by utilizing EXtrinsic Information Transfer (EXIT) chart technique to determine the convergence region in terms of the sensors observation accuracy and channel SNR, where the iterative decoder outperforms the non-iterative one.**

This leads to design of a bi-modal decoder that adaptively switches between the iterative and non-iterative modes in order to avoid inefficient iterative information exchange without compromising the resulting Bit Error Rate (BER). This adaptive decoding algorithm saves the computational power and decoding time by a factor of about 10 by avoiding unnecessary iterations.

***Index Terms*—Algorithm design, binary CEO problem, convergence analysis, iterative decoding, distributed turbo codes.**

I. Introduction

In indirect Multi Terminal (MT) coding, when the correlation among transmitters data streams is due to the multiple observations of a common source, the compression of correlated sources reduces to the special case of Chief Executive Officer (CEO) Problem [1].

This problem has been thoroughly studied from information-theoretic perspective. The first inner and outer bounds for the rate-distortion region is derived in [2] based on quantization and binning concept. Recently, new advances are reported that tighten these performance bounds and determine the rate-distortion region for more generalized cases [3], [4], [5].

The demand for practical MT code design is recently emphasized, since it applies to an important class of Wireless Sensor Networks (WSN), where no sensor can be deployed at the exact data source location due to harsh physical or environmental conditions [6].

Manuscript received May 21, 2013; revised September 15 and November 28, 2013, February 11, 2014; accepted February 12, 2014. The associate editor coordinating the review of this paper and approving it for publication was Z. Wang.

This work is financially sponsored by National Aeronautics and Space Administration (NASA), under grant No. EP11055404916. This work is presented in part at the 46th Annual Conference on Information Sciences and Systems (CISS'12).

A. Razi is with the Department of Electrical and Computer Engineering, Duke University, Durham, NC, 27708, USA (e-mail: abolfazl.razi@duke.edu).

A. Abedi is with the Wireless Sensor Networks Laboratory (WiSe-Net), Department of Electrical and Computer Engineering, University of Maine, Orono, ME, 04469 USA (e-mail: ali.abedi@maine.edu).

Digital Object Identifier 10.1109/TWC.2014.041014.130921

The first practical implementation of Distributed Source Codes (DSC) is Distributed Source Coding Using Syndromes (DISCUS) with the idea of successive decoding at the destination [7]. One shortcoming of this approach is that it only works in the corner points of the rate region corresponding to a full compression for one node and less for the others. Therefore, sensors have different coding rates, which is not desired in homogeneous WSN design.

Low Density Parity Check (LDPC) codes are employed based on the syndrome concept to eliminate this drawback and enable flexible coding for the whole Slepian-Wolf rate region using time sharing methods. In LDPC based codes, compression is achieved by using parity check matrix at the encoder to generate syndromes [8]. However, this scheme suffers from the error propagation phenomenon in noisy wireless channels, since successive decoding is very sensitive to syndrome errors.

A different scheme using punctured parity bits of LDPC codes is proposed in [9] to combat the error propagation problem. However, this approach performs a little away from the theoretical limits and does not fully extract the correlation among sensors. This is due to the fact that the correlation model is only used in the initialization of the Belief Propagation (BP) decoding algorithm and thereafter, each constituent decoder performs individually, which is not appropriate for the common source reconstruction problem.

One important advantage of using channel codes for DSC is that they can be combined with the channel coding stage to implement Distributed Joint Source Channel Codes (D-JSCC), which further simplifies sensor structures. Recently, Irregular Repeat Accumulate (IRA) and Turbo codes are proposed to combine decoding with data fusion functionality as a practical solution for the CEO problem [10], [11], [12]. In distributed turbo codes, compression is achieved through heavy puncturing after coding [13], [14], puncturing before coding [15] or using high rate convolutional encoders [16], [17]. The main drawback in using both distributed turbo codes and LDPC codes is their limited scalability to a large number of encoders due to using joint graph for many complex decoders cooperating with one another. This limits their applicability in WSNs with poor observations (e.g. remote sensing), where a large number of sensors is required to secure a desired reliability.

In [18], a convolutional based D-JSCC with a single iterative decoder is proposed for two correlated sensors, where one sensor has complete observation accuracy and the other provides side information. The authors proposed a similar coding scheme in [19] for an arbitrary number of sensors, where each sensor has an incomplete observation. A modified

parallel-structure Multiple Turbo Decoder (MTD) is designed considering the inaccuracy of sensors observations and a new technique is proposed to estimate the observation accuracy of sensors from the received data. A power optimized Distributed Space Time Block Code (D-STBC) assisted multiple relaying method is proposed in [20] to extend this solution to a more general class of clustered networks.

The common spirit of most D-JSCC solutions for the CEO problem is to employ iterative decoding to eliminate transmission errors followed by a majority vote to exploit the correlation among various observation sequences. There is an essential trade-off for the amount of correlation sent to the receiver. The larger the correlation among the codewords, the more enhanced the source coding efficiency. However, this larger correlation deteriorates the channel coding efficiency by lowering the minimum distance. This internal trade-off was considered by the former researchers (such as authors of [21], [22]) and is addressed through balancing the two contradictory needs. For instance, employing interleaver block after encoder or using Low Density Generator Matix (LDGM) codes are proposed to increase the correlation among the codewords, which negatively affects the error recovery in low SNR regime [21], [22].

In this work, we study this problem from a different perspective and change the dilemma of "How much correlation to be sent?" to "When and how much correlation to use in the decoder?". In other words, we reevaluate the usefulness of iterative exchange of soft information among constituent decoders, when a turbo-based decoder is used to estimate a common binary source from multiple coded observations. In fact, decoding improvement through iterations is shown to be highly dependent on the system quality factors. It is noteworthy that the inefficiency of iterative decoding in some situations was observed by former researchers [23], [15], but to the best of our knowledge no comprehensive study has been conducted to identify and prevent these situations.

We develop a new convergence analysis based on EXtrinsic Information Transfer (EXIT) charts to define convergence region in terms of the sensors observation accuracies and SNR level, in which the consecutive iterations are useful. This approach leads to design a bi-modal decoder that adaptively switches between two iterative and non-iterative decoding modes based on the current channel conditions. When switching to the non-iterative mode, the decoding complexity and decoding delay is significantly reduced by avoiding unnecessary iterations.

The rest of this paper is organized as follows. In Section II, the system model is defined. Section III elaborates on the coding scheme that is applied to the system. In Section IV, EXIT chart analysis is utilized to analyze the convergence of the iterative decoder in terms of the channel quality and observation model. Simulation results are provided in Section V to verify the proposed mode selection criterion followed by concluding remarks in Section VI.

II. System Model

In the system model depicted in Fig. 1, a single source is indirectly observed by a cluster of sensors. The sensors collectively transmit their observations through orthogonal real valued AWGN channels to the fusion center. Most practical systems use orthogonal transmissions such as TDMA and FDMA for complexity considerations. Therefore, it is a widely accepted assumption in majority of practical coding designs for sensors with correlated data, including works in [9], [10], [11], [12], [18]. However, orthogonal transmission is not optimal in general [24]. The proposed coding scheme is not dependent on the Multiple Access (MA) scheme and any MA scheme including CDMA can be used after coding stage at the cost of higher complexity. Before elaborating the details of the system, we define the following notations.

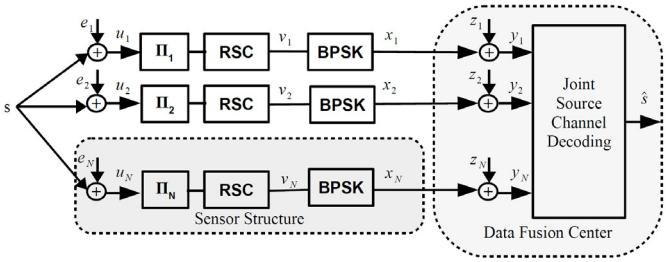

Fig. 1. System model: D-PCCC scheme is used to estimate a binary data source observed by multiple sensors.

Capital letters are used for Random Variables (RV), lower case letters for the realization of RVs, and bold-face letters for vectors and matrices. Operation $\bar{x} = 1 \oplus x$ is used to show bit flipping of a binary value x, where $\oplus$ is modulo-2 addition. $H(X)$ and $h(Y)$ are used to denote the entropy and differential entropy of discrete RV (X) and continuous valued RV (Y). The standard notation of $I(X;Y)$ is used for the Bitwise Mutual Information (BMI), which is the mutual information between a Log Likelihood Ratio (LLR) and the corresponding source bit averaged over the whole frame [25], [26]. The logarithms are e-based that provides the resulting information quantities in terms of *nat* unless noted otherwise.

The source data is assumed to be an independent identically distributed (i.i.d) equiprobable binary sequence $\{S(n)\}_{n=1}^{\infty}$. The observations of sensor i denoted as U_i is modeled as the source data passed through a virtual BSC channel with crossover probability $\beta_i \leq 0.5$. Thus, $U_i(n) = S(n) \oplus Z_i(n)$, where $Z_i(n)$ is i.i.d Bernoulli distributed observation error signal with

$$p\big(Z_i(n)=1\big) = 1 - p\big(Z_i(n)=0\big) = \beta_i. \tag{1}$$

Parameter β_i and $\bar{\beta}_i = 1 - \beta_i$ are called observation error and accuracy, respectively. The observation errors of different sensors are independent of the source and are mutually independent as well. This method is commonly used to model the observation accuracy of binary source observers.[1]

The observation of two sensors are conditionally independent given source data value, hence $\{U_i(n) \rightarrow S(n) \rightarrow$

[1]In fact, in most applications with an arbitrary set of source distribution, quantization method, and observation error model, the observed data is ultimately digitized prior to transmission. Therefore, the observation error can be translated into bit flipping of a binary data. If one could find an upper bound on the probability of bit flipping, it can be used as an asymptotic model to describe the observation inaccuracy. Hence, the virtual BSC channel becomes a general model to present the observation inaccuracy of sensors, at least as an asymptotic analysis [10].

$U_j(n)\}$ forms a Markov chain for $i \neq j$. Also, we assume equal observation error $\beta_i = \beta$ for notation simplicity, since the extension to the unequal observation error is straightforward. The pairwise crossover probability between the observations of sensors i and j is defined as $\beta_{ij} = p(U_i(n) \neq U_j(n))$, which is simply obtained as

$$\beta_{ij} = \beta_i + \beta_j - 2\beta_i\beta_j = 2\beta - \beta^2. \tag{2}$$

III. Coding Scheme

In this section, the coding scheme along with the proposed decoder is described.

A. PCCC Based Distributed Coding

The proposed D-JSCC is based on the parallel concatenation of convolutional codes (PCCC). The main difference between the proposed approach with the commonly used Distributed Turbo Codes (DTC) is that in DTC, a turbo encoder is implemented at each sensor [13]; while, we embed one Recursive Systematic Convolutional (RSC) encoder at each sensor and these encoders altogether realize a single turbo encoder. Hence, only one improved MTD based decoder is employed at the receiver, which makes the decoder less complex.

The system is composed of N sensors. At each sensor, the M-bit observed data sequence $\mathbf{u}_i = \{u_i(n)\}_{n=1}^{M}$ is encoded using an RSC encoder. The resulting systematic and parity bits are punctured to yield the output frames $\mathbf{v}_i = \{v_i(n)\}_{n=1}^{M/R_i}$, where R_i is the desired coding rate per sensor. The punctured output frame $\mathbf{v}_i$ is mapped to the BPSK symbols $\mathbf{x}_i$, using $x_i(n) = 2v_i(n) - 1$, hence $x_i(n) \in \mathcal{X} = \{-1, +1\}$.

In the first sensor, odd systematic bits and even parity bits are sent by puncturing half of the bits. We switch this pattern for the next sensor and continue this alternatively.

It was noticed that sending more parity bits slightly improves the BER performance. Therefore, we modified puncturing method for $N > 4$, such that $1/4$ and $3/4$ of transmit bits at any sensor are allocated to systematic and parity bits, respectively.

The data is interleaved using a pseudo random block interleaver Π_i prior to encoding to increase the minimum distance of the resulting multi-frame. The combination of the output frames forms a turbo-like multi-frame $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N]$. The received multi-frame at the decoder is $\mathbf{y} = [\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_N]$, where $\mathbf{y}_i = \{y_i(n)\}_{n=1}^{M/R_i}$.

We denote the transmit symbol of an arbitrary sensor by RV $X \in \mathcal{X} = \{-1, 1\}$ by omitting the time index n and position index i. If $Z \sim \mathcal{N}(0, \sigma_N^2)$ is the equivalent noise term at the receiver, then the received symbol $Y = X + Z$ has obviously the following Gaussian conditional probability density function (pdf):

$$p_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi}\sigma_N} \exp(-\frac{(y-x)^2}{2\sigma_N^2}). \tag{3}$$

B. MTD Based Decoder Structure

Inspired by the encoder structure, a MTD based decoder is employed at the receiver to decode the received multi-frame $\mathbf{y}$. As in Fig. 2, the decoder consists of Soft-Input Soft-Output (SISO) constituent decoders, each corresponding to an RSC encoder. In this work, the frames corresponding to different RSC encoders might differ by a few bits and word error is unavoidable, hence symbol based BCJR-MAP decoding algorithm is preferred over Soft Output Viterbi Algorithm (SOVA) [27]. The LLRs for the deinterleaved systematic bits are calculated as $L_{Y_i}(n) = \log(\frac{p(X_i(n)=+1|Y_i(n))}{p(X_i(n)=-1|Y_i(n)}) = 2(E_s/N_0)Y_i(n)$ for real valued AWGN channel with SNR= E_s/N_0 [28]. This initialization of input LLRs are used in one of the following decoding modes. The mode selection approach is presented in section IV.

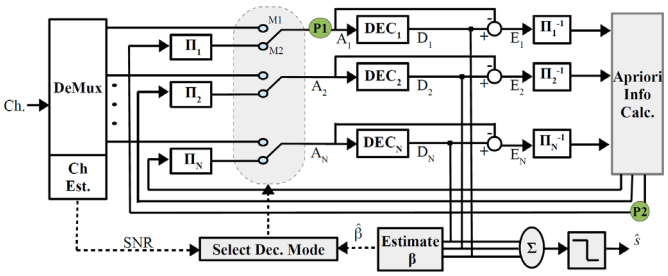

Fig. 2. Proposed bi-modal MTD decoder with observation accuracy estimation block. In non-iterative mode, the switch is in position M1.

1) Non-Iterative Mode: In this mode, shown by M1 in Fig. 2, the SISO decoders initialize the *a priori* LLRs denoted by A with the LLRs of the received systematic bits $A_i(n) = L_{Y_i}(n)$. Then, the decoders decode the received parity bits to provide output LLRs $D_i(n) = \log(\frac{p(U_i(n)=1|\mathbf{Y}_i)}{p(U_i(n)=0|\mathbf{Y}_i)})$. Then, applying the same correlation model between the source and observation bits defined in (1), the *a posteriori* likelihood of the source bits defined as $D_{s,i}(n) = \log(\frac{p(S(n)=1|\mathbf{Y}_i)}{p(S(n)=0|\mathbf{Y}_i)})$ are provided by each constituent decoder as follows:

$$\begin{aligned} D_{s,i}(n) &= \log\big(\frac{p(S(n)=1|\mathbf{Y}_i)}{p(S(n)=0|\mathbf{Y}_i)}\big) \\ &= \log\big(\frac{\beta p(U_i(n)=0|\mathbf{Y}_i) + (1-\beta)p(U_i(n)=1|\mathbf{Y}_i)}{(1-\beta)p(U_i(n)=0|\mathbf{Y}_i) + \beta p(U_i(n)=1|\mathbf{Y}_i)}\big) \\ &= \log\big(\frac{\beta + (1-\beta)\frac{p(U_i(n)=1|\mathbf{Y}_i)}{p(U_i(n)=0|\mathbf{Y}_i)}}{(1-\beta) + \beta\frac{p(U_i(n)=1|\mathbf{Y}_i)}{p(U_i(n)=0|\mathbf{Y}_i)}}\big) \\ &= \log\big(\frac{\beta + (1-\beta)e^{D_i(n)}}{(1-\beta) + \beta e^{D_i(n)}}\big) = g_\beta\big(D_i(n)\big), \end{aligned} \tag{4}$$

where we call $g_\beta(x) = \log\frac{\beta+(1-\beta)e^x}{(1-\beta)+\beta e^x}$ as the scaling function and use it frequently throughout this paper. It is easy to check that $g_\beta(x)$ is a monotonically increasing function of x for $\beta < 0.5$. The second equality in (4) follows from Markov chain property as

$$\begin{aligned} p(S(n)=1|\mathbf{Y}_i) &= \sum_{u=0,1} p(S(n)=1, U_i(n)=u|\mathbf{Y}_i) \\ &= \sum_{u=0,1} p(U_i(n)=u|\mathbf{Y}_i)p(S(n)=1|U_i(n)=u, \mathbf{Y}_i) \\ &= \sum_{u=0,1} p(U_i(n)=u|\mathbf{Y}_i)p(S(n)=1|U_i(n)=u) \\ &= \beta p(U_i(n)=0|\mathbf{Y}_i) + (1-\beta)p(U_i(n)=1|\mathbf{Y}_i). \end{aligned} \tag{5}$$

Finally, Maximum Likelihood Detection (MLD) is applied to provide an estimate of the source data, $\hat{s}(n)$. The MLD, for

the case of equal sensor observation error is reduced to the sum operation followed by a hard limiter as follows:

$$\hat{s}(n) = 0.5\big(1 + \text{sgn}\big(\sum_{i=1}^{N} D_{s,i}(n)\big)\big). \tag{6}$$

2) Iterative Mode: In this mode, the decoding is performed in consecutive iterations. The first iteration is the same as of the non-iterative decoder. The main difference is that the constituent decoders exchange soft information with one another at the end of each iteration. There are different LLR exchange structures including serial, master-slave and parallel. We choose the parallel structure due to its superior BER performance and fast convergence properties [29].

The extrinsic LLRs of decoder j at iteration r, $(E_j^{(r)})$ is scaled down with the observation error parameter β to provide new LLRs of the source bits $(E_{s,j}^{(r)})$. Then, the summation of these modified extrinsic LLRs for all constituent decoders except decoder i is summed to generate *a priori* LLRs for decoder i at the next iteration $r+1$, $(A_{s,i}^{(r+1)})$. To make the input LLRs consistent with the received symbols, we further scale down these input LLRs to obtain $(A_i^{(r+1)})$ as follows:

$$\begin{aligned} E_{s,j}^{(r)} &= g_\beta(E_j^{(r)}), \\ A_{s,i}^{(r+1)} &= E_i^{c(r)} = \sum_{\substack{j=1 \\ j\neq i}}^{N} E_{s,j}^{(r)}, \\ A_i^{(r+1)} &= g_\beta(A_{s,i}^{(r+1)}) = g_\beta\big(\sum_{\substack{j=1 \\ j\neq i}}^{N} g_\beta(E_j^{(r)})\big), \end{aligned} \tag{7}$$

where $A_i^{(r)}$ and $E_i^{(r)}$ are *a priori* and extrinsic LLRs of decoder i at iteration r, respectively. Subscript s declares that the likelihood is with respect to the source bits after the relevant scaling. $E_{s,i}^{c(r)}$ denotes the effective extrinsic LLR which is applied as *a priori* LLR to decoder i after proper scaling in the next iteration. The sequence of LLR change is depicted in Fig. 3. Note that for special case of two sensors $(N = 2)$, (7) reduces to:

$$A_i^{(r+1)} = g_\beta\big(g_\beta(E_j^{(r)})\big), \qquad i, j \in \{1, 2\}, i \neq j. \tag{8}$$

It is easy to see that $f_\beta(f_\beta(.)) = f_{2\beta(1-\beta)}(.) = f_{\beta_{ij}}(.)$. Therefore the LLR scaling in (8) from one constituent decoder to another is equivalent to the pairwise correlation between the two sensors, which verifies (7).

If β is known, it can be directly used in (4) and (7). Otherwise, we use the following estimation at the end of the first iteration. The pairwise correlation parameter between every two consecutive sensors is calculated as in (9). Then, noting $\beta_{i,j} = 2\beta - \beta^2$, the observation error parameter is found according to (10).

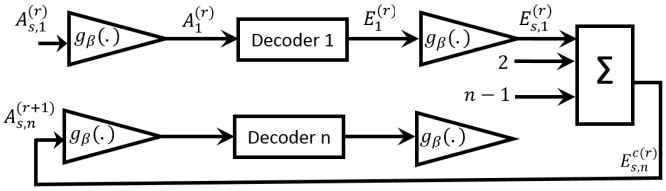

Fig. 3. LLR exchange sequence in the decoder.

$$\begin{aligned} \hat{\beta}_{i,i+1} &= \frac{1}{2M}\sum_{n=1}^{M} |\text{sgn}\big(D_i(n)\big) - \text{sgn}\big(D_{i+1}(n)\big)|, \\ & \qquad i = 1, 2, .., N-1. \end{aligned} \tag{9}$$

$$\begin{aligned} \hat{\beta}_{ij} &= \frac{1}{N-1}\sum_{i=1}^{N-1} \hat{\beta}_{i,i+1} \\ \Rightarrow \hat{\beta} &= \sqrt{1+\hat{\beta}_{ij}} - 1 \approx \frac{1}{2(N-1)}\sum_{i=1}^{N-1} \hat{\beta}_{i,i+1}. \end{aligned} \tag{10}$$

It is shown in [30], that the variance of $\hat{\beta}$ is inversely proportional to the number of symbols in a frame, M. Therefore, it provides an accurate estimate for fairly large frame sizes.

The iterative exchange of information among constituent decoders is repeated for a predetermined number of iterations, which is typically between 10 to 20. At the end of the last iteration, the source bits are estimated according to (6).

IV. EXIT Chart Analysis for MTD

EXIT charts are widely used to examine turbo decoder convergence, that shows how mutual information of LLRs with respect to the source bits are evolved in consecutive iterations [31], [26]. In this work, we apply EXIT chart analysis to the proposed decoder with some modifications based on the system model.

A procedure to derive EXIT charts for a parallel structure classical MTD is presented in [32]. This does not apply to our case, where the outputs of constituent decoders naturally tend to converge to the corresponding observation bit sequence rather than converging to the common data source sequence. Consequently, the iterative exchange of information among constituent decoders is useful only if the *a priori* information gets closer to the common data source through constituent decoder operation. The two comparison points are marked $P1$ and $P2$ in Fig. 2. Considering this fact, a modified EXIT chart based on BMI is proposed in this section.

A. EXIT Chart Definition for Distributed Coding

For BCJR Log-MAP decoding algorithm, we can focus on the symbol wise analysis by omitting the time index based on commonly accepted assumption of similar behavior for all bits regardless of their position in the sequence [25], [26]. Moreover, considering the pseudorandom interleavers and independence of the channels, the input LLRs of the constituent decoders experience independent distortion.

Based on (3) and following some mathematical manipulations as in [33], it can be shown that the obtained LLRs,

$L_Y = \log \frac{p(X=+1|Y)}{p(X=-1|Y)}$ are Gaussian RVs with mean $\mu_Y X$ and variance σ_Y^2 as:

$$L_Y = \mu_Y X + n_Y,$$
$$n_Y \sim \mathcal{N}(0, \sigma_Y^2), \quad \sigma_Y^2 = 2\mu_Y = 4/\sigma_N^2. \tag{11}$$

It is noteworthy that higher values of σ_Y provide higher certainty. For a Gaussian RV with mean μ and variance σ^2, the error probability is calculated as $Q(\mu/\sigma)$, where $\frac{\mu_Y}{\sigma_Y} = \frac{1}{2}\sigma_Y$ is called LLR SNR. Since $Q(x) = \frac{1}{\sqrt{2\pi}} \int_{t=x}^{\infty} e^{-t^2} dt$ is a monotonically decreasing function of x, higher μ/σ values yield lower error probability and hence higher certainty.

It has been shown that if both channel observations and input LLRs follow Gaussian distribution, in a MAP-family decoder with fairly large frame lengths, the extrinsic LLRs also tend to Gaussian distribution [34]. The intuitive justification is based on applying the weak law of large numbers to the summations over the random like decoder trellis structure. Moreover, extensive simulations confirm that relation (11) holds for the extrinsic LLRs as well [33]. Consequently, both input and extrinsic LLRs, A and E, can be written in the following format:

$$A = \mu_A X + n_A \quad , \quad n_A \sim \mathcal{N}(0, \sigma_A^2) \quad , \quad \mu_A = \sigma_A^2/2. \tag{12}$$
$$E = \mu_E X + n_E \quad , \quad n_E \sim \mathcal{N}(0, \sigma_E^2) \quad , \quad \mu_E = \sigma_E^2/2. \tag{13}$$

If V and X_i are BPSK versions of the source bit S and observation bit U_i, we have

$$\left.\begin{aligned} V &= 2S - 1 \\ X_i &= 2U_i - 1 \end{aligned}\right\}, (X, V \in \{-1, +1\}) \Rightarrow$$
$$p(X = -V) = 1 - p(X = V) = \beta. \tag{14}$$

The conditional pdf of input LLR, A is

$$p_A(\zeta|x) = \frac{1}{\sqrt{2\pi}\sigma_A} \exp\big(-\frac{(\zeta - \mu_A x)^2}{2\sigma_A^2}\big). \tag{15}$$

RVs A and V are conditionally independent given X; hence $\{V \to X \to A\}$ forms a Markov chain and we have

$$p_A(\zeta|v) = \sum_{x=-1,1} p_A(\zeta|x) p(x|v)$$
$$= \frac{1}{\sqrt{2\pi}\sigma_A} \big[\bar{\beta} e^{-\frac{(\zeta - \mu_A v)^2}{2\sigma_A^2}} + \beta e^{-\frac{(\zeta + \mu_A v)^2}{2\sigma_A^2}}\big]. \tag{16}$$

This conditional pdf is corresponding to a Gaussian Mixture Model (GMM), whose coefficients are derived from a Bernoulli distribution (first order binomial distribution). Throughout this paper, we call ζ a binomial-Gaussian RV with parameters $(m, \beta, \mu, \sigma^2)$, if:

$$p(\zeta; m, \mu, \sigma^2) = \sum_{i=1}^{m} \binom{m}{i} \beta^i (1-\beta)^{m-i} \mathcal{N}\big((m-2i)\mu, m\sigma^2\big). \tag{17}$$

It is known that for $\beta = 0$, the probability of input LLR error approaches zero for large variances. Accordingly, the BMI between the source data and input LLR approaches 1. We now present the following theorems for the incomplete observation accuracy case.

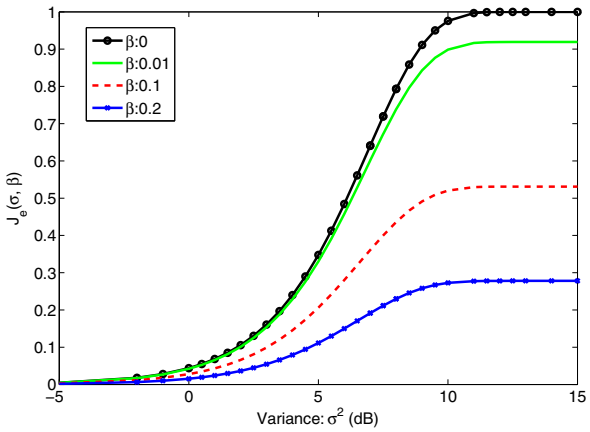

Fig. 4. Mutual information (in terms of *bit*) between the channel observation LLRs and the source data as a function of variance σ^2 for different observation error parameter β.

Theorem 4.1: The BMI between the source data and the $2nd$ order Binomial-Gaussian distributed LLR with parameter set $(m = 2, \beta, \mu_A = \sigma_A^2/2, \sigma_A^2)$ is

$$I(A;S) = J_e(\mu_A, \sigma_A, \beta)$$
$$= 1 - \frac{1}{\sqrt{2\pi}\sigma_A} \int_{-\infty}^{\infty} [\bar{\beta} e^{-\frac{(\zeta - \mu_A)^2}{2\sigma_A^2}} +$$
$$\beta e^{-\frac{(\zeta + \mu_A)^2}{2\sigma_A^2}}] \log\Big(\frac{1 + e^{-\frac{2\mu_A}{\sigma_A^2}\zeta}}{\bar{\beta} + \beta e^{-\frac{2\mu_A}{\sigma_A^2}\zeta}}\Big) d\zeta. \tag{18}$$

Proof: See Appendix A. ∎

Consequently, the mutual information $I(A;S)$ is a function of the LLR variance σ_A^2 and the observation error parameter β. If the observation accuracy is 100% (i.e. $\beta = 0$) and relation (12) holds, (18) is reduced to the well-known equation (19) for a classical MTD as in [33].

$$I(A;S) = J(\sigma_A)$$
$$= 1 - \frac{1}{\sqrt{2\pi}\sigma_A} \int_{-\infty}^{\infty} e^{-\frac{(\zeta - \mu_A)^2}{2\sigma_A^2}} \log(1 + e^{-\frac{2\mu_A}{\sigma_A^2}\zeta}) d\zeta. \tag{19}$$

In Fig. 4, the function $J_e(.)$ is depicted as a function of standard deviation σ_A and observation error parameter β. It is seen that $J_e(.)$ is a monotonically increasing function of variance for a fixed β, hence it is invertible. Also, it is observed that the function $J_e(\sigma, \beta)$ does not approach 1 even for extremely large variance σ^2. It rather approaches $1 - H(\beta)$. We note that the same relations hold for extrinsic LLRs:

$$I(E;S) = J_e(\mu_E, \sigma_E, \beta). \tag{20}$$

Note that for a one-to-one function $g_\beta(.)$, we have $I(A_i;S) = I(g_\beta(A_{s,i});S) = I(A_{s,i};S)$. So is true for E_i and $E_{s,i}$.

B. EXIT Chart Derivation for the Proposed Decoder

In this section, the equations derived in the previous section are used to develop an algorithm to derive the EXIT chart

curves. The relation between σ_A and σ_E or accordingly between $I(A;S)$ and $I(E;S)$ depends on the decoding algorithm parameters. This is generally obtained by empirical histogram methods, since no closed form is known, even for the case of complete observations [26], [31]. To develop a similar method for the proposed system, we first review the following remarks:

- **Remark 1**: Extensive simulations show that in a constituent decoder, when the input frames (including systematic and parity bits) and *a priori* LLRs (generated by the other constituent decoders) are corresponding to different versions of the same source data sequence, then the output LLRs D_i will approach Gaussian distribution with means biased either towards the corresponding bit in the input frame or the corresponding input LLR (based on the LLRs absolute values), if they are inconsistent. This is due to the large number of operations on the input frame and the LLRs on the trellis structure, which is well confirmed by Kolmogorov-Smirnov test for goodness of fit. However, after appropriate scaling of input LLRs to be consistent with the received bits, the resulting extrinsic LLRs, E_i always bias towards the observation bits.
- **Remark 2**: It is notable that the potential inefficiency of iterative information exchange among constituent decoders applied for correlated sensors was noticed in prior works. For instance, the authors in [21] have used the interleaver block after encoding to avoid harmful information exchange by generating consistent/coherent parity bits. This unusual use of interleaver is not optimal and decreases the minimum distance of the resulting codewords, especially for short frame lengths.

To implement the EXIT chart analysis, we need to calculate the following BMI between the source data and i) the input, $I(A_i^{(r)};S)$ and ii) the effective extrinsic LLRs $I(E_i^{c(r)};S)$. Noting (7), we have:

$$I(E_i^{c(r)};V) = I(A_{s,i}^{(r+1)};V) = I(A_i^{(r+1)};V) \qquad (21)$$

which quantifies the improvement of the mutual information between the common source data and *a priori* information at two consecutive iterations.

It is worth noting that in the case of complete observation accuracy ($\beta = 0$), we have $g_\beta(x) = x$ and the effective extrinsic LLR defined in (7) becomes $E_i^c = \sum_{\substack{j=1 \\ j\neq i}}^{N} E_j$, which is a Gaussian RV with the following mean and variances

$$\mu_{E_i^c} = \sum_{\substack{j=1 \\ j\neq i}}^{N} \mu_{E_j}, \qquad \sigma^2_{E_i^c} = \sum_{\substack{j=1 \\ j\neq i}}^{N} \sigma^2_{E_j} = (N-1)\sigma_E^2. \qquad (22)$$

It immediately follows that $\mu_{E_i^c} = (N-1)\mu_E = (N-1)\sigma_E^2/2 = \sigma^2_{E_i^c}/2$. Consequently, (13) holds for E_i^c and the mutual information between the source data and the effective extrinsic LLR can be easily calculated using $J(\sigma_{E_i^c})$ in (19). In the case of incomplete observation accuracy, $\beta \neq 0$, the following theorem holds for the effective extrinsic LLRs distribution, since relation (13) does not hold for E_i^c anymore.

Theorem 4.2: The extrinsic LLR, E_i^c has the following distribution with parameters $(m, \beta, \mu_E, \sigma_E^2)$:

$$\begin{aligned} p_{E_i^c}(\zeta|v) = \sum_{k=0}^{m} \binom{m}{k} \beta^k (1-\beta)^{m-k} \\ \underbrace{\big(f_Y(\zeta;-v\mu_E,\sigma_E^2) * \ldots * f_Y(\zeta;-v\mu_E,\sigma_E^2)\big)}_{k \text{ times}} \\ * \underbrace{\big(f_Y(\zeta;+v\mu_E,\sigma_E^2) * \ldots * f_Y(\zeta;+v\mu_E,\sigma_E^2)\big)}_{m-k \text{ times}}, \\ f_Y(y;\mu,\sigma^2) = \Big|\frac{e^y(1-2\beta)}{[e^y(1-\beta)-\beta][1-\beta-e^y\beta]}\Big| \\ .f_X\Big(\log\big(\frac{e^y(1-\beta)-\beta}{1-\beta-e^y\beta}\big)\Big), \quad X \sim \mathcal{N}(\mu,\sigma^2). \end{aligned} \qquad (23)$$

Proof: See Appendix B. ∎

Corollary 4.3: For the special case of complete observation accuracy ($\beta = 0$), it is easy to see that $f_Y(.) = f_X(.)$. Moreover, all terms except for $k = 0$ go away in (23) and it reduces to:

$$\begin{aligned} p_{E_i^c}(\zeta|v) &= \underbrace{f_Y(\zeta;+v\mu_E,\sigma_E^2) * \ldots * f_Y(\zeta;+v\mu_E,\sigma_E^2)}_{m \text{ times}} \\ &= \mathcal{N}(\zeta;+mv\mu_E, m\sigma_E^2). \end{aligned} \qquad (24)$$

This is consistent with (22) noting the fact that for $\beta = 0$, we have $E_i^c = \sum_{j=1}^{m} E_j$. For the case of two sensors ($m = 1$), (24) becomes $p_{E_i^c}(\zeta|v) = \mathcal{N}(\zeta;+v\mu_E,\sigma_E^2)$, which is corresponding to the classical turbo decoder.

Ultimately, the BMI between the source data and the extrinsic LLR can be calculated as

$$\begin{aligned} I(E^c;S) &= \sum_{v=\pm 1} p(v) \int_{-\infty}^{\infty} p_{E_i^c}(\zeta|v) \log\big[\frac{p_{E_i^c}(\zeta|v)}{p_{E_i^c}(\zeta)}\big] d\zeta, \\ &= 1 - \frac{1}{2} \sum_{v=\pm 1} \int_{-\infty}^{\infty} p_{E_i^c}(\zeta|v) \log\Big[\\ &\quad \frac{p_{E_i^c}(\zeta|v=-1) + p_{E_i^c}(\zeta|v=+1)}{p_{E_i^c}(\zeta|v)}\Big] d\zeta, \end{aligned} \qquad (25)$$

where we have used $p(V=1) = p(V=-1) = \frac{1}{2}$ and $\int_{-\infty}^{\infty} p_{E_i^c}(\zeta|v)d\zeta = 1$.

Plotting $I(E^c;S)$ vs $I(A;S)$ and $I(A;S)$ vs $I(E^c;S)$ curves according to Algorithm 1 represents the EXIT chart curves which are denoted by direct and reverse curves, respectively.

In Fig. 5, the modified EXIT chart is depicted versus observation accuracies and SNR values. The EXIT charts demonstrate the following properties:

1- The modified EXIT charts depend on both the channel SNR and observation accuracy. Higher observation errors suggest less advantage for iterative decoding. Likewise Lower channel SNRs present more advantage for iterative decoding.
2- Despite regular EXIT charts that are developed for classical point-to-point turbo codes, the modified EXIT chart

Algorithm 1: EXIT chart for MTD in a system with arbitrary number of sensors.

1) A sequence of binary source data bits, $\{s(n)\}_{n=1}^{M}$, is generated.
2) A sample observation sequence $\{u_i(n)\}_{n=1}^{M}$ is generated by passing the source data bits through a virtual BSC channel with crossover probability β according to (1).
3) The observation sequence is interleaved, RSC encoded, and punctured to form the output codeword $\{x_i(n)\}_{n=1}^{M/R}$ where R is the effective coding rate of each sensor. Considering the symmetry of the whole encoder/decoder structures, one constituent decoder is arbitrarily chosen.
4) The codeword is passed through an AWGN channel to form the channel observation $\{y_i(n)\}_{n=1}^{M/R}$.
5) The received symbols are unpunctured into the systematic and parity bits.
6) The input LLRs $\{A_{s,i}(n)\}_{n=1}^{M}$ are generated from the indirect observations of source data bits, by sampling pdf in (16), then are scaled down with $f_\beta(.)$ to form $\{A_i(n)\}_{n=1}^{M}$ and are fed into the constituent decoder to yield the output LLRs $\{D_i(n)\}_{n=1}^{M}$ as well as the extrinsic LLRs $\{E_i(n)\}_{n=1}^{M}$. Then the effective extrinsic LLRs are calculated using (7).
7) The observation error parameter $\hat{\beta}$ is estimated using (9) and (10). For the two decoder case, the second step in equation (10) is not required.
8) Ultimately, the mutual information between the source data bits and i) the *a priori* LLRs and ii) the effective extrinsic LLRs, $I(A_i;S)$ and $I(E_i^c;S)$ are calculated using (18) and (25).

curves do not necessarily approach full convergence point, even for very large SNR values. This results in relatively higher error floors. The lower bound on error floor can be calculated considering the error free channels from the sensors to the decoder. Error occurs if at least half of the observation bits are in error. Hence, we have the following error floor:

$$p_{\text{error}}^{(\min)} = \begin{cases} \displaystyle\sum_{k=\frac{N+1}{2}}^{N} \binom{N}{k} \beta^k (1-\beta)^{N-k}, & \text{N odd,} \\ \displaystyle\frac{1}{2}\binom{N}{N/2}\beta^{N/2}(1-\beta)^{N/2} + \\ \displaystyle\sum_{k=\frac{N}{2}+1}^{N} \binom{N}{k} \beta^k (1-\beta)^{N-k}, & \text{N even.} \end{cases} \tag{26}$$

For the two-sensor case, (26) reduces to $p_{\text{error}}^{(\min)} = \beta$. The theoretical SNR to achieve the error floor is obtained in Appendix D.

3- The initial slope of EXIT chart direct curve defines the convergence property. Near-zero initial slope declares a poor convergence property, where no significant BER improvement is obtained using iterative decoding.

4- If the number of sensors are too high, the EXIT chart direct curve approaches the horizontal line of $I(E;S) = 1$ *bit*, as stated in the following theorem. This means that almost complete certainty is obtained at the first iteration, and no considerable gain is obtainable by further iterations. Hence, non-iterative mode is always preferred.

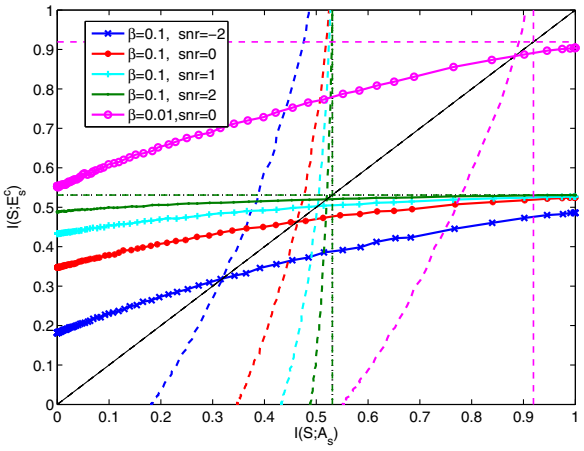

Fig. 5. EXIT charts for different observation accuracies (Number of sensors: 2). Solid and dashed lines are, respectively, corresponding to the direct and reverse curves. Dotted lines are corresponding to the maximum achievable BMI (in terms of *bit*), $1 - H(p_{\text{error}}^{(\min)})$, for error floor defined in (26).

Theorem 4.4: The BMI between the source data and the effective extrinsic LLRs for extremely large number of sensors and small observation error parameter ($\beta \to 0$) approaches 1 *bit* for any positive channel SNR.

Proof: See appendix C. ■

V. Decoding Mode Selection and Experimental Results

We define the ϵ-convergence region as the system conditions at which the iterative decoder outperforms the non-iterative one with a margin of at least ϵ, where $0 < \epsilon \ll 1$ is a small number. Specifically, we have $p_e^{\text{iter}} < (1-\epsilon)p_e^{\text{no-iter}}$, where p_e^{iter} and $p_e^{\text{no-iter}}$ are the BER of the iterative and non-iterative decoders, respectively. A corresponding relation can be found for $I(A_i;S)$ and $I(E_i^c;S)$ based on the equations derived in section IV. The non-convergence region is corresponding to the near zero initial slope of the EXIT chart direct curve, where no considerable gain is obtained by adding to the input LLRs. This region is obtained for $N = 2$ and $\epsilon = 0.1\&0.02$ by running the proposed algorithms for various pairs of (SNR, β) and analyzing the resulting EXIT charts. The ϵ-convergence region is shown with dark color in Fig. 6.

The convergence region 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 after running one decoding iteration. If the (SNR, β) pair falls in the convergence region, the iterative decoding mode is chosen. Failure in estimation of the channel SNR and

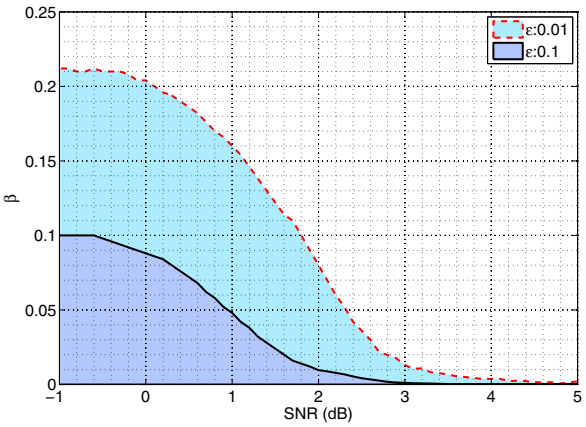

Fig. 6. Convergence region of iterative decoding algorithm in terms of the channel SNR and β for $\epsilon = 0.1$ and $\epsilon = 0.02$; outside the region, non-iterative algorithm is selected.

observation accuracy, which may cause an undesired mode selection, can be simply avoided by choosing fairly large frame lengths that ensures a high estimation accuracy. However, in the case of very rare event of the wrong mode selection, the BER performance and complexity of the proposed bimodal decoder are equivalent to the currently used single-mode decoders.

To confirm the performance improvement obtained by the mode selection criterion, we performed extensive simulations for the proposed decoder with the following parameters ($M = 2048$ bits, $N = 2$). The feedforward and feedback polynomials of the RSC encoders are arbitrarily set to $f(D) = 1+D+D^2$ and $g(D) = 1 + D^2$, respectively.

Table. VI represents the resulting BER for various (SNR, β) pairs. Iteration=1 corresponds to the non-iterative mode and iteration=10 is used to calculate the BER of the iterative mode. The last column shows the BER improvement obtained by iterative decoding. It is shown that for large SNR values, the BER of the iterative decoder is almost equal to that of the non-iterative decoder (i.e. within ϵ range). Hence, no-iterative mode is desired for complexity reduction. This SNR limit depends on the observation accuracy of sensors. The more accurate are the sensors, the higher is the SNR limit and hence the iterative mode is chosen more frequently. This is due to the fact that in this case, the observation bits of sensors and hence the extrinsic LLRs are more coherent. The results in this table are consistent with the iterative region obtained in Fig. 6. For instance, for the choice of (SNR=+4 dB, $\beta = 0.05$), there is no improvement for the iterative decoding, which falls out of the convergence region. In contrast, (SNR=0 dB, $\beta = 0.01$) presents a significant improvement for the iterative decoding and falls inside the convergence region in Fig. 6.

TABLE I
BER RESULTS FOR VARIOUS (β, SNR) PAIRS. THE SYSTEM PARAMETERS ARE $N = 2$, $M = 2048$ AND BER IS OBTAINED BY AVERAGING OVER 100 FRAMES. THE LAST COLUMN SHOWS THE PERCENTAGE OF THE BER IMPROVEMENT WITH ITERATIVE DECODING, (I.E. $100 \times (p_e^{\text{NO-ITER}} - p_e^{\text{ITER}})/p_e^{\text{NO-ITER}}$).

β	SNR (dB)	Iter:1	Iter:2	Iter: 3	Iter:10	Imp (%)
0.15	0	1.708e-01	1.649e-01	1.648e-01	1.650e-01	3.4
	+2	1.513e-01	1.507e-01	1.508e-01	1.501e-01	0.3
	+4	1.503e-01	1.506e-01	1.506e-01	1.506e-01	-0.2
0.05	0	7.524e-02	5.882e-02	5.708e-02	5.655e-02	24.8
	+2	5.091e-02	4.981e-02	4.972e-02	4.977e-02	2.2
	+4	5.007e-02	5.029e-02	5.025e-02	5.025e-02	-0.4
0.01	0	3.718e-02	1.310e-02	1.066e-02	1.033e-02	72.2
	+2	1.121e-02	9.907e-03	1.001e-02	1.001e-02	10.6
	+4	9.659e-03	9.595e-03	9.595e-03	9.595e-03	0.7

VI. Conclusions

The use of scalable and low-complexity D-PCCC with a customized MTD-based decoder to estimate a binary data source surrounded by a cluster of partially accurate observers is studied. Convergence of the iterative decoding is analyzed by introducing a modified EXIT chart technique. It is shown that the usefulness of iterative information exchange among constituent decoders depends on the observation accuracy of sensors and the channel quality. This finding is used to determine the superiority region of the iterative decoding. As a general rule, it is concluded that the iterative operation is less useful, when the channel SNR is very high or the observation accuracy of sensors are too low.

This region is derived once using extensive EXIT chart analysis and revealed to the decoder at design time. This approach reduces the complexity of decoder by avoiding useless iterations. This new approach of conditioning cooperation among constituent decoders to the current system quality factor can be used to improve the performance and reduce the complexity of the similar iterative joint decoding schemes applicable to the CEO problem.

Appendix A
Proof of Theorem 4.1

The mutual information between *a priori* information of a constituent decoder and the source data is calculated as:

$$\begin{aligned} I(A;S) = I(A;V) &= \sum_{v=-1,1} \int_{\zeta=-\infty}^{\infty} p_A(v,\zeta) \log \frac{p_A(v,\zeta)}{p_A(v)p(\zeta)} d\zeta \\ &= \frac{1}{2} \sum_{v=-1,1} \int_{-\infty}^{\infty} p_A(\zeta|v) \log \frac{p_A(\zeta|v)}{p_A(\zeta)} d\zeta \\ &= \frac{1}{2} \sum_{v=-1,1} \int_{-\infty}^{\infty} p_A(\zeta|v) \log(\\ &\quad \frac{2p_A(\zeta|v)}{p_A(\zeta|v=-1) + p_A(\zeta|v=+1)}) d\zeta. \end{aligned} \tag{27}$$

From (16) we recall that

$$\begin{aligned} p_A(\zeta|v) &= \frac{1}{\sqrt{2\pi}\sigma_A} \big[\bar{\beta}. \exp\big(-\frac{(\zeta - \mu_A v)^2}{2\sigma_A^2}\big) \\ &\quad + \beta. \exp\big(-\frac{(\zeta + \mu_A v)^2}{2\sigma_A^2}\big)\big] \end{aligned} \tag{28}$$

$$\begin{aligned} &\Rightarrow \frac{2p_A(\zeta|v=\pm 1)}{p_A(\zeta|v=-1) + p_A(\zeta|v=+1)} \\ &= \frac{\bar{\beta} + \beta. \exp(\pm\frac{2\mu_A\zeta}{\sigma_A^2})}{1 + \exp(\pm\frac{2\mu_A\zeta}{\sigma_A^2})}. \end{aligned} \tag{29}$$

Substituting (29) in (27) results in

$$I(A;S)=\frac{1}{2}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_A}\big[\bar{\beta}e^{-\frac{(\zeta+\mu_A)^2}{2\sigma_A^2}}+\beta e^{-\frac{(\zeta-\mu_A)^2}{2\sigma_A^2}}\big]\log\big(\frac{\bar{\beta}+\beta.\exp(-\frac{2\mu_A\zeta}{\sigma_A^2})}{1+\exp(-\frac{2\mu_A\zeta}{\sigma_A^2})}\big)d\zeta+\frac{1}{2}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_A}\big[\bar{\beta}e^{-\frac{(\zeta-\mu_A)^2}{2\sigma_A^2}}+\beta e^{-\frac{(\zeta+\mu_A)^2}{2\sigma_A^2}}\big]\log\big(\frac{\bar{\beta}+\beta.\exp(+\frac{2\mu_A\zeta}{\sigma_A^2})}{1+\exp(+\frac{2\mu_A\zeta}{\sigma_A^2})}\big)d\zeta. \quad (30)$$

Both integrals have the same value. This can be easily verified by change of variables ($u=-\zeta$). Hence, (30) can be rewritten as

$$I(A;S)=\frac{1}{\sqrt{2\pi}\sigma_A}\int_{-\infty}^{\infty}\big[\bar{\beta}e^{-\frac{(\zeta+\mu_A)^2}{2\sigma_A^2}}+\beta e^{-\frac{(\zeta-\mu_A)^2}{2\sigma_A^2}}\big].\big[1+\log\big(\frac{\bar{\beta}+\beta.\exp(-\frac{2\mu_A\zeta}{\sigma_A^2})}{1+\exp(-\frac{2\mu_A\zeta}{\sigma_A^2})}\big)\big]d\zeta. \quad (31)$$

Noting that $\int_{-\infty}^{\infty}p_A(\zeta|v=1)d\zeta=1$, (31) reduces to

$$I(A;S)=1-\frac{1}{\sqrt{2\pi}\sigma_A}\int_{-\infty}^{\infty}\big[\bar{\beta}e^{-\frac{(\zeta+\mu_A)^2}{2\sigma_A^2}}+\beta e^{-\frac{(\zeta-\mu_A)^2}{2\sigma_A^2}}\big].\log\big(\frac{1+\exp(-\frac{2\mu_A\zeta}{\sigma_A^2})}{\bar{\beta}+\beta.\exp(-\frac{2\mu_A\zeta}{\sigma_A^2})}\big)d\zeta. \quad (32)$$

This completes the proof.

Appendix B
Proof of Theorem 4.2

If S is the source data bit with BPSK modulated version $V=2S-1$. We present the observation bit set of the first m sensors as $\mathbf{U}^m=\{U_1,U_2,...,U_m\}$ with support set $\mathcal{U}^m$. Also we define $\mathcal{U}^{m,k}$ as a subset of $\mathcal{U}^m$ such that the observation bit is in error for k out of m sensors. It is obvious that $\mathcal{U}^{m,k}$s are disjoint sets and $\mathcal{U}^{m,0}\cup\mathcal{U}^{m,1}...\cup\mathcal{U}^{m,m}=\mathcal{U}^m$ making $\mathcal{U}^{m,i}$ a partition of $\mathcal{U}^m$. Noting that $p(U_i=S)=1-p(U_i=\bar{S})=\beta$, the probability of any randomly chosen observation set in $\mathcal{U}^m$ belongs to $\mathcal{U}^{m,k}$ follows the binomial distribution with parameters (m,β) regardless of V; i.e.,

$$p(\mathbf{U}^m\in\mathcal{U}^{m,k}|v)=p(\mathbf{U}^m\in\mathcal{U}^{m,k})=\binom{m}{k}\beta^k(1-\beta)^{m-k}. \quad (33)$$

The observation set determines the extrinsic LLR's mean value as

$$E_i\sim\begin{cases}\mathcal{N}(+V\mu_E,\sigma_E^2), & \text{if } U_i=S,\\ \mathcal{N}(-V\mu_E,\sigma_E^2), & \text{if } U_i=\bar{S},\end{cases} \quad (34)$$

which can be rewritten as

$$E_i\sim\mathcal{N}\big((2U_i-1)\mu_E,\sigma_E^2\big). \quad (35)$$

Since E_i^c is the summation of all constituent decoders extrinsic LLRs except one, the conditional pdf of E_i^c can be calculated as

$$\begin{aligned}p_{E_i^c}(\zeta|v)&=\sum_{\forall\mathbf{u}^m\in\mathcal{U}^m}p(E_i^c=\zeta|\mathbf{u}^m)p(\mathbf{u}^m|v)\\&=\sum_{k=0}^{m}p(E_i^c=\zeta|\mathbf{u}^m\in\mathcal{U}^{k,m})p(\mathbf{u}^m\in\mathcal{U}^{m,k}|v)\\&=\sum_{k=0}^{m}\binom{m}{k}\beta^k(1-\beta)^{m-k}p(E_i^c=\zeta|\mathbf{u}^m\in\mathcal{U}^{m,k}) \quad (36)\end{aligned}$$

where E_i^c is defined in (7). For notation convenience, we use new variables $x\triangleq E_i=\mathcal{N}(\mu,\sigma^2)$ and $y\triangleq g_\beta(x)=\log\big(\frac{\beta+(1-\beta)e^x}{(1-\beta)+\beta e^x}\big)$. For $\beta<0.5$, y is a monotonic increasing function of x. Using the change of variable methods in [35], we have the following pdf for y:

$$\begin{aligned}&f_X(x;\mu,\sigma^2)=\mathcal{N}(\mu,\sigma^2)\\&\Rightarrow f_Y(y;\mu,\sigma^2)=|\frac{d}{dy}(g_\beta^{-1}(y))\mid f_X(g_\beta^{-1}(y))\\&=|\frac{e^y(1-2\beta)}{[e^y(1-\beta)-\beta][1-\beta-e^y\beta]}|f_X(\log(\frac{e^y(1-\beta)-\beta}{1-\beta-e^y\beta})) \quad (37)\end{aligned}$$

It is easy to see that for $\beta=0$, we have $f_Y(y)=f_X(x)$, which verifies the correctness of (37). We also recall that the pdf of $Z=Y_1+Y_2+...Y_n$ for independent Y_i is in the form of $f_Z(\zeta)=f_{Y_1}(\zeta)*f_{Y_2}(\zeta)*...*f_{Y_m}(\zeta)$ [35]. Noting that E_i^c for $\mathbf{U}^m\in\mathcal{U}^{m,k}$ includes k extrinsic LLRs with mean $-V\mu_E$ and $(m-k)$ with mean $+V\mu_E$, all with variance σ_E^2, we have the following distribution for E_i^c after some straightforward manipulations:

$$p_{E_i^c}(\zeta|v)=\sum_{k=0}^{m}\binom{m}{k}\beta^k(1-\beta)^{m-k}.\underbrace{\big(f_Y(\zeta;-v\mu_E,\sigma_E^2)*...*f_Y(\zeta;-v\mu_E,\sigma_E^2)\big)}_{k\text{ times}}*\underbrace{\big(f_Y(\zeta;+v\mu_E,\sigma_E^2)*...*f_Y(\zeta;+v\mu_E,\sigma_E^2)\big)}_{m-k\text{ times}}, \quad (38)$$

where $f_Y(.)$ is defined in (37). This completes the proof.

Appendix C
Proof of Theorem 4.4

To prove this theorem, first we calculate $p_{E_i^c}(\zeta<0|v=1)$. From (24) we have

$$\begin{aligned}p_e&=p_{E_i^c}(\zeta<0|v=1)\\&=\int_{\zeta=-\infty}^{0}\frac{1}{\sqrt{2\pi m\sigma_E^2}}\sum_{k=0}^{m}\binom{m}{k}\beta^k(1-\beta)^{m-k}.\exp\big[-\frac{(\zeta-(2k-m)\mu_E)^2}{2m\sigma_E^2}\big]d\zeta\\&=\sum_{k=0}^{m}\binom{m}{k}\beta^k(1-\beta)^{m-k}Q\big[\frac{(2k-m)\mu_E}{\sqrt{m}\sigma_E}\big]. \quad (39)\end{aligned}$$

We note that for the case of no signal reception at the decoder (SNR=0), the mean of extrinsic LLRs are independent

of the corresponding source bits and hence $\mu_E = 0$. In this case, we have $p_e = \sum_{k=0}^{m} \binom{m}{k}\beta^k(1-\beta)^{m-k}Q(0) = \frac{1}{2}$ as expected. For a positive SNR, using normal approximation for the binomial distribution for large m, (39) converts to

$$p_e \approx \frac{1}{\sqrt{2\pi m\beta(1-\beta)}} \sum_{k=0}^{m} \Big[\exp\big[-\frac{(k-m\beta)^2}{2m\beta(1-\beta)}\big] .Q\big[\frac{(2k-m)\mu_E}{\sqrt{m}\sigma_E}\big]\Big]. \quad (40)$$

Using approximation $Q(x) \leq \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ for Q function as in [36], [37] and noting $\mu_E = \sigma_E^2/2$, results in

$$p_e \approx \frac{1}{2\pi\sqrt{m\beta(1-\beta)}} \sum_{k=0}^{m} \Big[\underbrace{\exp\big[-\frac{(k-m\beta)^2}{2m\beta(1-\beta)}\big]}_{\alpha_1} . \underbrace{\exp\big[-\frac{(m-2k)^2\sigma_E^2}{4m}\big]}_{\alpha_2}\Big]. \quad (41)$$

We know that $0 \leq \alpha_1, \alpha_2 \leq 1$. Also, $\alpha_1 = 1$ in proximity of $k = m\beta$ and decays exponentially elsewhere. So does α_2 in proximity of $k = m/2$. Therefore for $\beta \neq \frac{1}{2}$ and large m these two points are far apart; hence, $\alpha_1.\alpha_2$ is very small everywhere. Therefore, the summation approaches zero and we have $\lim_{m\to\infty} p_e = 0$.

An alternative proof is presented using jointly typical sets concept. If $v = 1$, the realization of observation set $\mathbf{u}^m = \{u_1, u_2, ...u_m\}$ most likely includes $m(1-\beta)$ correct and $m\beta$ false bits. Hence, E_i^c is a summation of $m(1-\beta)$ Gaussian RV with mean $+\mu_E$ and $m\beta$ Gaussian RV with mean $-\mu_E$, all with variance σ_E^2. Consequently, E_i^c is a Gaussian RV with mean $(1-2\beta)\mu_E$ and variance $m\sigma_E^2$. If we define $\hat{E} = \text{sgn}(E_i^c)$, the probability of LLR error, $p_e = p(\hat{E} \neq v)$ can be calculated as follows

$$\begin{aligned} p_e &= p_{E_i^c}(\zeta < 0 | v = 1) = Q\big(\frac{m(1-2\beta)\mu_E}{\sqrt{m}\sigma_E}\big) \\ &= Q\big(\sqrt{m}(1-2\beta)\mu_E/\sigma_E\big) \underset{m\to\infty}{\longrightarrow} Q(\infty) = 0. \end{aligned} \quad (42)$$

The last equality holds only for nonzero SNR ($\mu_E/\sigma_E > 0$) as mentioned in the previous proof. Consequently, we have $I(\hat{E}; S) = H(p_e)$. Applying data processing inequality to Markov chain $V \to E_i^c \to \hat{E}$ results in

$$I(E_i^c; S) \geq I(\hat{E}; S) = H(p_e) \underset{m\to\infty}{\longrightarrow} 1 \ bit\ . \quad (43)$$

This completes the proof.

APPENDIX D
SNR LIMIT TO REACH ERROR FLOOR

To calculate the minimum SNR required to reach error floor, we note that due to the decoder structure, error floor is reached if the lossless transmission of the sensors correlated data is secured. For a Gaussian MAC channel, it was shown that the maximum sum-rate is achievable by time sharing as in TDMA [24]. The maximum rate of each sensor for equal noise variance N_0, is $R_i = \frac{1}{2}\log(1 + \frac{E_S}{N_0})$, where E_s is the energy per transmit symbol.

On the other hand, recall that S is the source bit and $\mathbf{X}^N$ is the observation bits of N sensors with realization $\mathbf{x}^N$ and support set $\mathcal{X}^N, \mathcal{X} = \{0, 1\}$. Due to the equal and independent observation error of sensors, the information of each observation bit is $\frac{1}{M}H(\mathbf{X}^N)$. If the coding rate per sensor is R_i, therefore we need to have $\frac{R_i}{M}H(\mathbf{X}^N) \leq \frac{1}{2}\log(1 + \frac{E_S}{N_0})$ for lossless communication from sensors to the destination. It provides the SNR limit of $E_s/N_0(\text{min}) = e^{2R_i\frac{H(\mathbf{X}^N)}{N}} - 1$ or equivalently $E_b/N_0(\text{min}) = \frac{N}{R_i}(e^{2R_i\frac{H(\mathbf{X}^N)}{N}} - 1)$. The required joint entropy of observation bits $H(\mathbf{X}^N)$ can be easily calculated as:

$$\begin{aligned} H(\mathbf{X}^N) &= -\sum_{\forall \mathbf{x}^N \in \mathbf{X}^N} p(\mathbf{x}^N)\log(p(\mathbf{x}^N)) \\ &= -\sum_{k=0}^{N}\binom{N}{k} p(\mathbf{X}^N = \mathbf{x}^{N,k})\log(p(\mathbf{X}^N = \mathbf{x}^{N,k})) \\ &= -\sum_{k=0}^{N}\binom{N}{k}\big(\sum_{s=0}^{1} p(\mathbf{X}^N = \mathbf{x}^{N,k}|S=s)p(S=s)\big) \\ &\qquad .\log\big(\sum_{s=0}^{1} p(\mathbf{X}^N = \mathbf{x}^{N,k}|S=s)p(S=s)\big) \\ &= \sum_{k=0}^{N}\binom{N}{k}\frac{\bar{\beta}^k\beta^{N-k} + \beta^k\bar{\beta}^{N-k}}{2} \\ &\qquad .\log\big(\frac{\bar{\beta}^k\beta^{N-k} + \beta^k\bar{\beta}^{N-k}}{2}\big) \\ &= 1 - \sum_{k=0}^{N}\binom{N}{k}\frac{\bar{\beta}^k\beta^{N-k} + \beta^k\bar{\beta}^{N-k}}{2} \\ &\qquad .\log\big(\bar{\beta}^k\beta^{N-k} + \beta^k\bar{\beta}^{N-k}\big), \end{aligned} \quad (44)$$

where $\mathbf{x}^{N,k}$ is an arbitrary realization of $\mathbf{x}^N$ with k number of 0 bits.

REFERENCES

[1] T. Berger, Z. Zhang, and H. Viswanathan, "The CEO problem [multiterminal source coding]," *IEEE Trans. Inf. Theory*, vol. 42, no. 3, May 1996.

[2] 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. Areas Commun.*, vol. 22, no. 6, pp. 977–987, Aug. 2004.

[3] A. Wagner and V. Anantharam, "An improved outer bound for multiterminal source coding," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 1919–1937, May 2008.

[4] Y. Yang and Z. Xiong, "On the generalized Gaussian CEO problem," *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 3350–3372, June 2012.

[5] R. Soundararajan, A. Wagner, and S. Vishwanath, "Sum rate of the vacationing CEO problem," *IEEE Trans. Inf. Theory*, no. 99, p. 1, 2012.

[6] A. Razi and A. Abedi, "Interference reduction in wireless passive sensor networks using directional antennas," in *Proc. 2011 Caneus Fly by Wireless Workshop*, pp. 1–4.

[7] S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (discus): design and construction," in *Proc. 1999 Data Compression Conf.*, pp. 158–167.

[8] A. Liveris, Z. Xiong, and C. Georghiades, "Compression of binary sources with side information at the decoder using LDPC codes," *IEEE Commun. Lett.*, vol. 6, no. 10, pp. 440–442, Oct. 2002.

[9] M. Sartipi and F. Fekri, "Source and channel coding in wireless sensor networks using LDPC codes," in *Proc. 2004 IEEE Conf. Sensor Ad Hoc Commun. Netw.*, pp. 309–316.

[10] W. Zhong and J. Garcia-Frias, "Combining data fusion with joint source-channel coding of correlated sensors," in *Proc. 2004 IEEE Inf. Theory Workshop*, pp. 315–317.

[11] ——, "Combining data fusion with joint source-channel coding of correlated sensors using IRA codes," in *Proc. 2005 Conf. Inf. Science Syst.*
[12] J. Haghighat, H. Behroozi, and D. Plant, "Joint decoding and data fusion in wireless sensor networks using turbo codes," in *Proc. 2008 IEEE Int. Symp. Personal, Indoor Mobile Radio Commun.*, pp. 1–5.
[13] J. Garcia-Frias and Y. Zhao, "Compression of binary memoryless sources using punctured turbo codes," *IEEE Commun. Lett.*, vol. 6, no. 9, pp. 394–396, Sept. 2002.
[14] A. Aaron and B. Girod, "Compression with side information using turbo codes," in *Proc. 2002 Data Compression Conf.*, pp. 252–261.
[15] A. Liveris, Z. Xiong, and C. Georghiades, "A distributed source coding technique for correlated images using turbo-codes," *IEEE Commun. Lett.*, vol. 6, no. 9, pp. 379–381, Sept. 2002.
[16] ——, "Distributed compression of binary sources using conventional parallel and serial concatenated convolutional codes," in *Proc. 2003 Data Compression Conf.*, pp. 193–202.
[17] J. Bajcsy and P. Mitran, "Coding for the Slepian-Wolf problem with turbo codes," in *Proc. 2001 IEEE Global Telecommun. Conf.*, vol. 2, pp. 1400–1404.
[18] F. Daneshgaran, M. Laddomada, and M. Mondin, "Iterative joint channel decoding of correlated sources," *IEEE Trans. Wireless Commun.*, vol. 5, no. 10, pp. 2659–2663, Oct. 2006.
[19] A. Razi, K. Yasami, and A. Abedi, "On minimum number of wireless sensors required for reliable binary source estimation," in *Proc. 2011 IEEE Wireless Commun. Netw. Conf.*, pp. 1852–1857.
[20] A. Razi, F. Afghah, and A. Abedi, "Power optimized DSTBC assisted DMF relaying in wireless sensor networks with redundant super nodes," *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 636–645, Feb. 2013.
[21] J. Garcia-Frias, Y. Zhao, and W. Zhong, "Turbo-like codes for transmission of correlated sources over noisy channels," *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 58–66, Sept. 2007.
[22] M. Hernaez, P. Crespo, J. Del Ser, and J. Garcia-Frias, "Serially-concatenated LDGM codes for correlated sources over Gaussian broadcast channels," *IEEE Commun. Lett.*, vol. 13, no. 10, pp. 788–790, Oct. 2009.
[23] A. Liveris, Z. Xiong, and C. Georghiades, "Distributed compression of binary sources using conventional parallel and serial concatenated convolutional codes," in *Proc. 2003 Data Compression Conf.*, pp. 193–202.
[24] T. M. Cover and J. Thomas, *Elements of Information Theory*. John Wiley, 1991.
[25] R. W. Hamming, *Coding and Information Theory*. Prentice-Hall, 1986.
[26] S. ten Brink, "Convergence of iterative decoding," *Electron. Lett.*, vol. 35, no. 13, pp. 1117–1119, June 1999.
[27] P. Jung, "Comparison of turbo-code decoders applied to short frame transmission systems," *IEEE J. Sel. Areas Commun.*, vol. 14, no. 3, pp. 530–537, Apr. 1996.
[28] S. Lin and D. Costello, *Error Control Coding: Fundamentals and Applications*. Prentice Hall, 1983.
[29] D. Divsalar and E. Pollara, "Multiple turbo codes for deep-space communications," *JPL TDA Prog. Rep.*, pp. 42–121, May 1995.
[30] A. Razi and A. Abedi, "Distributed coding of sources with unknown correlation parameter," in *Proc. 2010 Int. Conf. Wireless Netw.*
[31] H. El Gamal and J. Hammons, "Analyzing the turbo decoder using the Gaussian approximation," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 671–686, Feb. 2001.
[32] S. ten Brink, "Convergence of multidimensional iterative decoding schemes," in *Proc. 2001 Asilomar Conf. Signals, Syst., Comput.*, vol. 1, pp. 270–274.
[33] ——, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1727–1737, Oct. 2001.
[34] N. Wiberg, H.-A. Loeliger, and R. Kotter, "Codes and iterative decoding on general graphs," in *Proc. 1995 IEEE Int. Symp. Inf. Theory*, p. 468.
[35] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. MacGraw Hill, 1991.
[36] G. Karagiannidis and A. Lioumpas, "An improved approximation for the Gaussian Q-function," *IEEE Commun. Lett.*, vol. 11, no. 8, pp. 644–646, Aug. 2007.
[37] J. Dyer and S. Dyer, "Corrections to, and comments on, 'An improved approximation for the Gaussian Q-function'," *IEEE Commun. Lett.*, vol. 12, no. 4, p. 231, Apr. 2008.

Abolfazl Razi is a postdoctoral researcher in the Electrical and Computer Engineering department at Duke University. He received his B.S. and M.S. in Electrical Engineering from Sharif University and Amirkabir University, respectively. He obtained his Ph.D in Electrical and Computer Engineering from University of Maine in 2013. He was a visiting graduate student at the University of Maryland in Summer and Fall 2012. He served as PACE chair of IEEE Maine Section, in 2010-2012. Prior to joining the Ph.D. program, he served as R&D expert, software engineer and project manager in several communications companies including MCI, PTK, and FARDA. He is the recipient of several academic awards including the best graduate research assistant of year from College of Engineering school, University of Maine, 2011 and the best paper award from IEEE/CANEUS fly by wireless workshop, 2011. His research interests include distributed coding, information theory, multiuser communications, Bayesian inference and non-linear estimation.

Ali Abedi received his B.Sc and M.Sc degrees in Electrical Engineering from Sharif University of Technology, and his Ph.D. in Electrical and Computer Engineering from University of Waterloo in 1996, 1998, and 2004, respectively. He joined University of Maine in 2005, where he is currently Associate Professor of Electrical and Computer Engineering and Director of WiSe-Net Lab. Prior to joining the University of Maine, he was a lecturer at Queen's University, and University of Waterloo. Dr. Abedi held Visiting Associate Professor position at the University of Maryland, College Park in 2012 and affiliate faculty position at Maine Institute for Human Genetics and Health from 2010-11.

Dr. Abedi has authored two books and over 60 research articles in Wireless Communications field. Highlights of his research include analytical performance evaluation of block codes, new methods for performance and convergence analysis of Turbo-codes, and applications of error correction codes in wireless sensor networks for structural monitoring, space explorations and biomedical engineering. His research on wireless sensing of lunar habitat was featured on NSF Science360 website in 2012. Dr. Abedi is a senior member of IEEE and candidate for IEEE Board of Directors in 2013 elections.