Adaptive Bi-modal Decoder for Binary Source Estimation with Two Observers

Abolfazl Razi and Ali Abedi
Wireless Sensor Networks Lab (WiSe-Net)
Department of Electrical and Computer Engineering
University of Maine, Orono, ME 04469
Email: {abolfazl.razi, ali.abedi}@maine.edu

***Abstract*—In this paper, an implementation friendly distributed version of parallel concatenated convolutional codes (D-PCCC) with a novel bi-modal iterative decoder is proposed to estimate a binary data source observed by a cluster of sensors. The convergence of the iterative decoder is analyzed by defining a modified EXIT chart. It is shown that iterative method is not always the best decoding algorithm for D-PCCC with two encoders. The system quality conditions in terms of channel SNR and sensors' observation accuracy required for usefulness of iterative algorithm is found. This provides a frame by frame mode selection criterion for decoder. The decoder chooses an appropriate iterative or non-iterative decoding mode that yields higher BER performance in both modes and reduces decoding complexity by avoiding destructive or useless iterations.**

***Index Terms*—Binary source estimation, convergence analysis, distributed source coding, EXIT chart, iterative decoding.**

I. Introduction

Practical code design for data monitoring with multiple observers has gained a lot of attention from researchers in the past decade, enabling new applications in wireless sensor networks (WSN), video coding and surveillance systems. Distributed source coding (DSC) of correlated sensors becomes more critical, when single observer systems fail to provide a reliable estimation of data due to sensors observation inaccuracy. The theoretical aspects of this problem was first studied by Slepian and Wolf and the coding rate region was defined for lossless compression [1]. Afterwards, it was generalized by finding a rate distortion function for lossy compression case [2]. One special case of this problem arises when a common source is observed by multiple sensors which is called multi-terminal coding or the chief executive officer problem (CEO) in the literature [3]. This problem has been widely investigated and the rate distortion function is derived for various cases, although there are still many unsolved problems [4], [5].

Several efficient implementations of DSC have been proposed to address this problem. Most of these schemes employ channel codes such as low-density parity-check (LDPC), irregular repeat-accumulate (IRA), low-density generator matrix (LDGM), and Turbo codes to implement an efficient DSC [6]–[8]. These approaches are easily applicable to the CEO problem when the number of sensors is rather low. The idea of using parallel concatenated convolutional codes (PCCC) with an iterative decoding algorithm is recently proposed to implement distributed joint source channel codes (D-JSCC) for correlated sources and the CEO problem as a special case in [9] and [10]. The two properties of simple structure of encoder and low complexity of decoder facilitate its usage in practical WSNs.

It is commonly agreed that in a multiple turbo decoder (MTD) used to decode a WSN cluster composed of D-PCCC encoders, iteratively exchanging information between constituent decoders enhances the system BER performance. In this paper, this presumption is reconsidered. The classical EXIT chart approach is generalized to be applied to MTD based decoders when the observation accuracies of sensors are not perfect. It is shown that superiority of iterative decoding is strongly dependent upon the observation accuracy of sensors as well as channel SNR. This analysis led to design of a bi-modal decoder that adaptively switches between two modes, iterative and non-iterative based on the system quality conditions.

The rest of this paper is organized as follows. In Section II, the system model along with the proposed encoder and decoder is defined. In Section III, a modified EXIT chart is introduced to analyze the iterative decoder's convergence property. Simulation results are evaluated in Section IV followed by conclusions in Section V.

II. System Model

Fig. 1 presents the system model used to estimate a common equiprobable binary data source, s using multiple observers. The observation of sensor i, u_i might be in error with probability β_i where without loss of generality $0 \leq \beta_i \leq \frac{1}{2}$. If N is the number of sensors, we assume equal observation error for sensors $\beta_1 = \beta_2 = ... = \beta_N = \beta$; otherwise, we choose the worst case $\beta = max\{\beta_1, \beta_2, ..., \beta_N\}$. Therefore, u_i can be written as $u_i = s \oplus e_i$, where observation error, e_i is a binary random variable (RV) independent of s with $p(e_i = 1) = 1 - p(e_i = 0) = \beta$.

Assuming independent observation error, it can be shown that the observation of any pair of sensors are conditionally independent given a source data value. Consequently $\mathbf{u}_i \rightarrow$

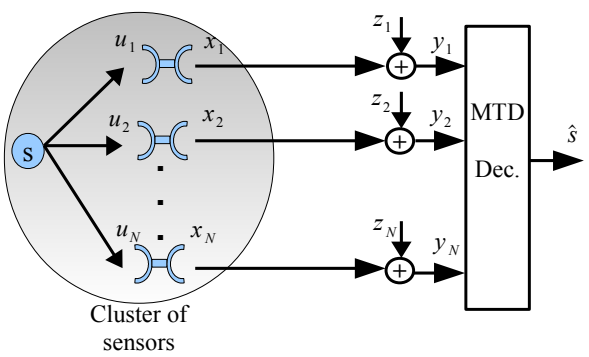

Fig. 1: System model to estimate a binary data source observed by multiple sensors.

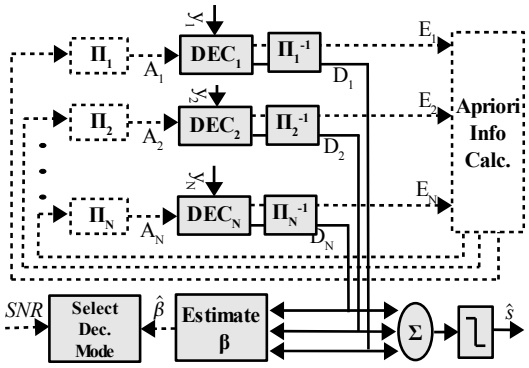

Fig. 2: Proposed bi-modal MTD decoder. In the non-iterative mode, the blocks and lines marked with dashed line go away.

$\mathbf{s} \rightarrow \mathbf{u}_j$ forms a Markov chain. The cascade of two BSC channels results in a pairwise bit flipping probability of β_{ij}, where

$$\beta_{ij} = \beta_i + \beta_j - 2\beta_i\beta_j \approx 2\beta \tag{1}$$

A. Distributed Code Design

Conventional encoding techniques are used as detailed in this section. Each sensor picks a frame of L observation bits $\{u_i(n)\}_{n=1}^{L}$, interleaves it with a pseudo-random interleaver, encodes it with a recursive systematic convolutional (RSC) encoder, punctures it with an appropriate puncturing pattern to end up with the desired coding rate per sensor denoted by R, and finally BPSK modulates it to form the output frame $\{x_i(n)\}_{n=1}^{L/R}$. The output frames are transmitted to a common data fusion center via independent AWGN channels. The received noisy frames at the destination are denoted by $\{y_i(n)\}_{n=1}^{L/R}$. In this work, the feedforward and feedback polynomials are chosen as $f(D) = 1 + D^2 + D^3$ and $g(D) = 1 + D + D^3$, which are widely used in practical systems [11].

B. Decoding Algorithm

Considering the encoding scheme, we present a novel bi-modal decoding method in this section. A multi-frame composed of output frames of the sensors altogether, forms a turbo-like encoded frame. Intuitively, a MTD with parallel structure is appropriate for decoding the received multi-frame. The structure of this decoder is depicted in Fig. 2.

The received multi-frame is divided into individual frames, each corresponding to a sensor. The decoder composed of N soft-input soft-output (SISO) constituent decoders. The received systematic and parity bits are fed into the corresponding decoder to be decoded. It is worth noting that the log-likelihood ratios (LLR) of the the received systematic bits, denoted by A_i, form the soft input of the SISO decoder i, as follows

$$A_i = 4(Es/N_0)y_i \tag{2}$$

where y_i is the received symbol and E_s/N_0 is the equivalent channel SNR at destination. Decoder i decodes the received symbols from the channel considering the a priori soft information using a maximum a posteriori (MAP) based decoding algorithm and provides the output LLRs, D_i.

Two operation modes, iterative and non-iterative are defined next. The blocks marked with dashed line in Fig. 2 contribute only to the iterative mode. If the non-iterative mode is chosen, after running the algorithm once, the output LLRs are deinterleaved appropriately and summed; the signs of the resulting values in (3) provide the first estimation of the BPSK modulated source data bits. In fact, the summation operation followed by a hard limiter is the maximum likelihood (ML) detection of the source data bits using the equal certainty output LLRs. Hence, the source is estimated as follows and the decoding algorithm terminates.

$$\hat{s}(n) = 2\ sgn\Big(\sum_{i=1}^{N} D_i(n)\Big) - 1 \tag{3}$$

where $sgn(x) = \frac{|x|}{x}$ is the sign function.

On the other hand, if the iterative mode is chosen, the soft information is exchanged between constituent decoders and the LLRs are updated in the consecutive iterations. At the end of each iteration, the extrinsic LLRs of each decoder is calculated as the additional certainty obtained at this iteration

$$E_i^{(r)} = D_i^{(r)} - A_i^{(r)} \tag{4}$$

where $E_i^{(r)}$, $D_i^{(r)}$, and $A_i^{(r)}$ are input, output, and extrinsic LLRs of decoder i at iteration r, respectively. Then, the input LLRs of each decoder for the next iteration is calculated as the summation of extrinsic LLRs of all decoders, but the current decoder, as follows

$$A_i^{(r+1)} = E_i^{(r)} = \sum_{j=1, j\neq i}^{N} E_j^{(r)} \quad , N > 2 \tag{5}$$

These new input LLRs, in turn, yield new output and extrinsic LLRs and the decoding algorithm continues for a predefined number of iterations.

In the next section, it is shown that the two iterative and non-iterative modes might outperform one another, based on the system quality conditions, which is discussed in section III. Selection parameters include i) channel SNR which is directly estimated from the received symbols by commonly used methods and ii) the observation accuracy of sensors which is estimated using a technique presented in [10].

III. EXIT Chart Analysis

One commonly used technique to investigate convergence of a MTD is EXIT chart analysis. Two important variants of EXIT chart method are based upon LLR SNR analysis and mutual information analysis. In this work, we choose the mutual information method and modify it to be applicable to the proposed iterative decoding algorithm considering the imperfect observation accuracies of sensors.

In this method, bitwise mutual information between the source data and both input and extrinsic LLRs are calculated to evaluate the input and extrinsic LLRs distance from the source data in the sense of mutual information. To explain the idea of defining a modified EXIT chart function, we note that classical EXIT chart formulation is not applied to the case of incomplete observation, since in this case the input LLRs are noisy versions of the sensors' observed bits and not the source data. Therefore, there might be a case that after running the decoding algorithm at a constituent decoder, the extrinsic LLRs get closer to the observed bits compared to the input LLRs, but at the same time, they get less close to the source data bits. In other words, in this analysis, the LLRs are generated from the observed bits, while the mutual information should be calculated with respect to the source data bits.

Considering Gaussian noise $\mathcal{N}(0,\sigma_N^2)$, if x and y represent the transmitted and the received symbols, respectively, then we have the following conditional probability density function (pdf) for y

$$p_{\mathbf{y}|\mathbf{x}}(y|x) = \frac{1}{\sqrt{2\pi}\sigma_N} exp(-\frac{(y-x)^2}{2\sigma_N^2}) \tag{6}$$

It is shown in [12] that the LLR of y, is a Gaussian RV and can be written as

$$L_y = log[\frac{p(y|x=+1)}{p(y|x=-1)}] = \mu_y x + n_y, \qquad n_y \sim \mathcal{N}(0,\sigma_y^2) \tag{7}$$

where

$$\sigma_y^2 = 2\mu_y = 4/\sigma_N^2 \tag{8}$$

It is apparent from (8) that higher channel noise power results in lower absolute value of LLR and hence lower certainty. It also has been proven that due to Gaussian distribution of input LLRs as well as channel observations and noting the random like lattice structure of MAP decoding algorithm, the output and extrinsic LLRs follow the Gaussian distribution, likewise [12].

We define the BPSK modulated version of a source bit as $v = 2s-1$. Now, omitting the sensor index, if u is an observed bit with BPSK modulated version $x = 2u-1$, recalling the observation error probability, we have the following conditional pdf for input LLR

$$p_A(\zeta|x) = \frac{1}{\sqrt{2\pi}\sigma_A} exp\big(-\frac{(\zeta-\mu_A x)^2}{2\sigma_A^2}\big) \tag{9}$$

Since observation and channel errors are independent, the triplet $\{v \to x \to A\}$ forms a Markov chain. Therefore,

$$\begin{aligned} p_A(\zeta|v) &= \sum_{x=-1,1} p_A(\zeta|x)p(x|v) \\ &= \frac{1}{\sqrt{2\pi}\sigma_A}\big[\bar{\beta}e^{-\frac{(\zeta-\mu_A v)^2}{2\sigma_A^2}} + \beta e^{-\frac{(\zeta+\mu_A v)^2}{2\sigma_A^2}}\big] \end{aligned} \tag{10}$$

This conditional pdf includes four parameters $(m = 2, \beta, \mu_A, \sigma_A^2)$. This pdf is called *binomial-Gaussian* distribution of order m, throughout this paper. To calculate the mutual information between the source data bit and input LLR, we recall the definition of mutual information between a discrete and a continuous-valued RV as follows

$$\begin{aligned} I(A;v) &= \sum_{v=-1,1} \int_{\zeta=-\infty}^{\infty} p_A(v,\zeta) log \frac{p_A(v,\zeta)}{p_A(v)p(\zeta)} d\zeta \\ &= \frac{1}{2} \sum_{v=-1,1} \int_{-\infty}^{\infty} p_A(\zeta|v) log \frac{2p_A(\zeta|v)}{p_A(\zeta|v=-1)+p_A(\zeta|v=+1)} d\zeta \end{aligned} \tag{11}$$

According to (10), it can be easily seen that

$$\frac{2p_A(\zeta|v=\pm 1)}{p_A(\zeta|v=-1)+p_A(\zeta|v=+1)} = \frac{\bar{\beta}+\beta e^{\pm\frac{2\mu_A\zeta}{\sigma_A^2}}}{1+e^{\pm\frac{2\mu_A\zeta}{\sigma_A^2}}} \tag{12}$$

Substituting (12) in (11) with some mathematical manipulations results in

$$\begin{aligned} &I(A;v) = J_e(\sigma_A,\mu_A,\beta) = \\ &1 - \frac{1}{\sqrt{2\pi}\sigma_A} \int_{-\infty}^{\infty} \big[\bar{\beta}e^{-\frac{(\zeta+\mu_A)^2}{2\sigma_A^2}} + \beta e^{-\frac{(\zeta-\mu_A)^2}{2\sigma_A^2}}\big] log\big(\frac{1+e^{-\frac{2\mu_A\zeta}{\sigma_A^2}}}{\bar{\beta}+\beta e^{-\frac{2\mu_A\zeta}{\sigma_A^2}}}\big) d\zeta \end{aligned} \tag{13}$$

Noting the relation $\mu_A = \sigma^2/2$, (13) reduces to :

$$\begin{aligned} &I(A;v) = J_e(\sigma_A,\beta) = \\ &1 - \frac{1}{\sqrt{2\pi}\sigma_A} \int_{-\infty}^{\infty} [\bar{\beta}e^{-\frac{(\zeta-\sigma_A^2/2)^2}{2\sigma_A^2}} + \beta e^{-\frac{(\zeta+\sigma_A^2/2)^2}{2\sigma_A^2}}] log(\frac{1+e^{-\zeta}}{\bar{\beta}+\beta e^{-\zeta}}) d\zeta \end{aligned} \tag{14}$$

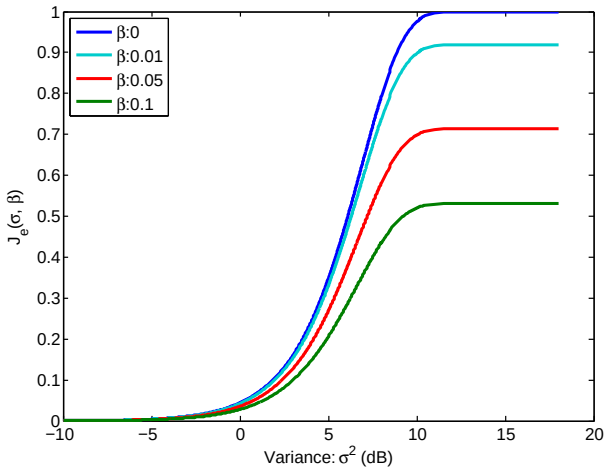

Fig. 3: Mutual information between the input LLR and source data as a function of LLR variance σ^2 and observation error parameter β.

This function is depicted in Fig. 3 which demonstrates the mutual information dependence upon channel SNR and β. It is simply verified that for the case of no observation error $\beta = 0$, (14) simplifies to the well known equation (15) derived for a classical turbo decoder in [12].

$$\begin{aligned} I_{A,v} &= J(\sigma_A) \\ &= 1 - \frac{1}{\sqrt{2\pi}\sigma_A} \int_{-\infty}^{\infty} e^{-\frac{(\zeta-\mu_A)^2}{2\sigma_A^2}} log(1 + e^{-\frac{2\mu_A}{\sigma_A^2}\zeta}) d\zeta \end{aligned} \tag{15}$$

For the sake of simplicity, in this paper we consider a system with two sensors. Similar to classical EXIT charts, we derive the modified EXIT charts by performing extensive simulations to obtain empirical LLR histograms. Then, we follow the bellow procedure to find the relation between σ_A and σ_E or correspondingly between $I_{A,v}$ and $I_{E,v}$.

To do so, a sample input sequence of length L, $\{s(n)\}_{n=1}^{L}$ is randomly generated. This is passed through two independent virtual BSC channels with parameter β to form two observed sequences $\{u_i(n)\}_{n=1}^{L},\ i = 1, 2$. The two observed sequences are encoded and appropriately punctured to form two output frames $\{x_i(n)\}_{n=1}^{L/R}$. Noting the symmetry of decoders, the first constituent decoder is chosen for simulation. Thus, the noisy version of the first encoded sequence $\{y_1(n)\}_{n=1}^{L/R}$ is used to perform decoding; while the a priori LLRs $\{A_1(n)\}_{n=1}^{L}$ are made using the second observed sequence $\{u_2(n)\}_{n=1}^{L}$ for different LLR variances σ_A^2. The binomial-Gaussian distributed input LLR sequence with variance σ_A^2 yields extrinsic LLRs with the same distribution and a different variance σ_E^2 that can be found from the extrinsic LLRs histogram analysis. Ultimately, the mutual information between the source data and both input and extrinsic LLRs, $I_{A,v}$ and $I_{E,v}$ are found using (14). Plotting direct curve, ($I_{E,v}$ vs $I_{A,v}$) and its reverse curve, ($I_{A,v}$ vs $I_{E,v}$) provides EXIT chart curves. For simplicity, we only plot the direct curve and avoid plotting the reverse curve, and the stair case-like plots that show how soft information is exchanged between two constituent decoders.

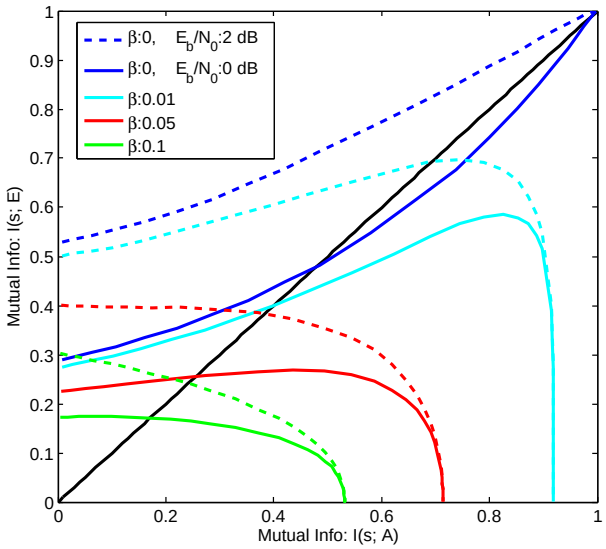

Fig. 4: Modified EXIT chart function for different values of observation error parameter, β. Solid and dashed lines represent $E_b/N_0 = 0dB$ and $E_b/N_0 = 2dB$, respectively

IV. Simulation Results

Fig. 4 presents the EXIT chart direct curves for different SNR and observation error values. It is seen that for $\beta = 0$, the EXIT chart direct curve is equivalent to the one of regular turbo decoder, and approaches the full convergence point, $I(s;E) = I(s;A) = 1$. If SNR is higher than a certain value, then direct curve lies above $I(s;A) = I(s;E)$ curve and a tunnel opens up between the two curves. This means that the decoder totally converges and BER approaches zero.

For the incomplete observation case, where $\beta \neq 0$, the initial slope of the curve is not always positive and depends on SNR and β values. Negative initial slope means that input LLRs with higher certainty result in extrinsic LLRs with lower certainty; hence iterative exchange of LLRs between decoders is not useful, as justified in Section III. Also, it is seen that extremely large LLR values in this case do not cause perfect input LLR certainty. This is because input LLRs are made from the sensors observed sequence and not the source data bits. Therefore, for very large LLR values we have

$$\lim_{\sigma_A^2 \to \infty} I(s;A) = I(s;u) = 1 - H(\beta) \tag{16}$$

where $1 - H(\beta) = 1 + \beta log_2(\beta) + \bar{\beta} log_2(\bar{\beta})$ is the information capacity of a BSC channel with crossover probability β. Based on data processing inequality, we have $I(s;A) \leq I(s;u)$. According to (8), very large LLR variances means error free channels from sensors to the decoder. Therefore, $I(s;A)$

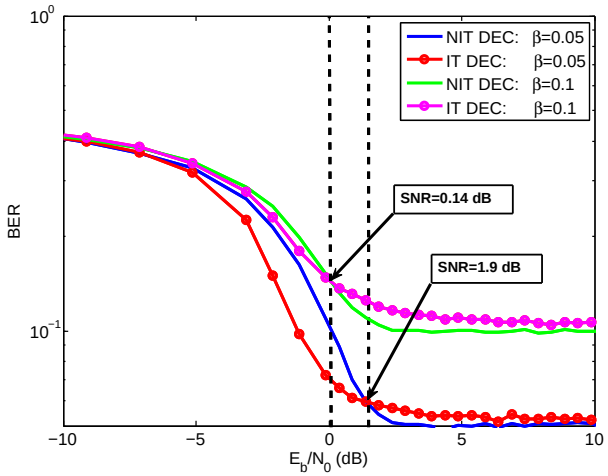

Fig. 5: BER performance comparison of the iterative and non-iterative decoder for different observation error parameters. Number of sensors is 2.

approaches $I(s;u) = 1 - H(\beta)$. At this point, we also have $I(s;E) = 0$. This is because high certainty input LLRs which are generated from the other sensor's observation bits are in conflict with the received systematic and parity bits, hence the constituent decoder fails to decode the received sequence.

To find the conditions where the iterative decoder outperforms the non-iterative one, we are interested in (SNR, β) pairs where the initial slope of the EXIT chart direct curve is positive. We can interpret from Fig. 4 that for a given SNR value, higher observation errors result in lower performance obtained by iterative information exchange. Also, for a fixed β, lower SNR shows higher performance for the iterative decoder.

For instance, for observation accuracy of $\beta = 0.1$ represented by cyan color in Fig. 4, it is shown that $E_b/N_0 = 0\ dB$ results in zero initial slope or almost the same performance for both decoders; while $E_b/N_0 = 2\ dB$ results in negative initial slope, or equivalently lower performance for the iterative decoder. This suggests $E_b/N_0 = 0\ dB$ as border line SNR for $\beta = 0.1$, where the decoder switches between iterative and non-iterative modes. Similarly, this SNR threshold for $\beta = 0.05$ can be found to be about $2\ dB$.

This fact is confirmed by BER performance comparison between the iterative and non-iterative decoders for observation error parameters $\beta = 0.05$ and $\beta = 1$, presented in Fig. 5. The number of iterations is 20 and the the frame length is set to $L = 2000$. The SNR threshold, where the superiority of iterative and non-iterative decoders to one another switches, is depicted with dashed lines. These thresholds are in accordance with the ones found by analyzing the proposed EXIT charts.

V. Conclusion

In this work, an easy-to-implement low-complexity distributed coding based on D-PCCC is proposed for a binary CEO problem that is appropriate for applications with tiny sensors. The use of a parallel structure MTD decoder for this problem is analyzed and the widely accepted hypothesis of usefulness of iterative information exchange between constituent decoders is criticized. It was shown that usefulness of iterative decoding is strongly dependent on the sensors' observation accuracy and channel SNR. As a general rule, the information exchange is more useful for lower SNRs and higher observation accuracies.

This finding led to design of a bi-modal decoder that extracts the channel SNR and observation error parameter β from the received frames and selects the more appropriate decoding mode. This new decoder not only outperforms both single mode iterative and non-iterative decoders, but also reduces decoding complexity by avoiding useless iterations. Since the region of superiority of iterative decoding in terms of (SNR, β) pairs can be found once and revealed to the decoder at design time, no additional run-time complexity is imposed on the decoder.

References

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

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

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

[4] J. Wang, J. Chen, and X. Wu, "On the sum rate of Gaussian multiterminal source coding: New proofs and results," *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3946–3960, Aug. 2010.

[5] J. Chen and T. Berger, "Successive Wyner-Ziv coding scheme and its application to the quadratic Gaussian CEO problem," *IEEE Trans. Inf. Theory*, vol. 54, no. 4, pp. 1586–1603, Apr. 2008.

[6] J. Garcia-Frias and Z. Xiong, "Distributed source and joint source-channel coding: from theory to practice," in *Proc. IEEE Acoustics, Speech, and Signal Processing Conference (ICASSP '05)*, vol. 5, Mar. 2005, pp. 1093–1096.

[7] S. Qaisar and H. Radha, "Multipath multi-stream distributed reliable video delivery in wireless sensor networks," in *Proc. the 43th Annual Conf. Inf. Sciences Sys. (CISS' 09)*, Mar. 2009, pp. 207 –212.

[8] R. Halloush and H. Radha, "Practical distributed video coding based on source rate estimation," in *Proc. the 44th Annual Conf. Inf. Sciences Sys. (CISS' 10)*, Mar. 2010, pp. 1 –6.

[9] 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.

[10] A. Razi, K. Yasami, and A. Abedi, "On minimum number of wireless sensors required for reliable binary source estimation," in *Proc. IEEE Wireless Commun. & Networking Conf. (WCNC '11)*, Mar. 2011, pp. 1852–1857.

[11] M. Martina, M. Nicola, and G. Masera, "A flexible UMTS-WiMax turbo decoder architecture," *IEEE Trans. Circuits and Systems: Express Briefs*, vol. 55, no. 4, pp. 369–373, Apr. 2008.

[12] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1727 –1737, Oct. 2001.