Joint Decoding and Data Fusion in Wireless Sensor Networks Using Turbo Codes

Javad Haghighat
Department of Electrical and Computer Engineering
McGill University
Montreal, Quebec, Canada, H3A 2A7
Email: javad.haghighat@mcgill.ca

Hamid Behroozi
Department of Mathematics and Statistics
Queen's University
Kingston, Ontario, Canada, K7L 3N6
Email: behroozi@queensu.ca

David V. Plant
Department of Electrical and Computer Engineering
McGill University
Montreal, Quebec, Canada, H3A 2A7
Email: david.plant@mcgill.ca

***Abstract*—We consider the problem of joint decoding and data-fusion in data gathering sensor networks modeled by the Chief Executive Officer (CEO) problem. Correlation between sensors' data is known at the fusion center and is employed to update extrinsic information received from soft-in soft-out (SISO) decoders. It is shown in the literature that this scheme has a lower bit error rate compared with the schemes that separately decode data received from each sensor and then estimate the value of the source. Previous works consider correlated Gaussian sources and apply a single SISO decoder. We consider the binary CEO problem, where all sensors observe the same binary source corrupted by independent binary noises, and apply turbo codes to encode and transmit them to the fusion center. We show how extrinsic information is passed between SISO decoders and the vertical-decoding unit that updates extrinsic information using channel correlations. We illustrate the performance of the joint decoder for different correlations and rates. Simulation results show promising improvements compared with the separate decoding scheme. We also compare the bit error rates achieved by turbo codes with the ones achieved by convolutional codes and discuss the results.**

Index Terms— CEO problem, sensor networks, iterative decoding, turbo codes.

I. Introduction

The increasing attention given to new applications of wireless sensor networks (WSNs) is a reason for new interests in evaluating source-channel communications in multi-terminal systems. Such applications include environmental and structural monitoring, rescue operations and disaster recovery, health care and medical applications, film-making and media production, to name a few [1], [2]. A WSN consists of a collection of small, low-power sensor nodes spread across a geographical area for performing distributed sensing tasks and measuring physical phenomena. It is a rapidly-deployable network that does not require any fixed infrastructure. In this paper, we consider a data gathering WSN. For instance, consider WSNs at sites of accidents such as collapse of a building (Fig. 1) to detect and locate trapped survivors, or to track natural gas and toxic substances [2]. This type of WSN can be modeled by the CEO problem [3], which is an abstract model for remote monitoring in wireless networks.

In the CEO problem, a CEO is interested in a source that cannot be observed directly. N agents (sensors) observe independent noisy versions of the source, separately encode their observations, and then transmit through rate-constrained channels to a single fusion center (FC) for further processing. The scenario is shown in Fig. 2. The FC intends to form an optimal estimate of the source based on information received from the agents. The CEO problem is considered for the discrete case in [3] and for the quadratic Gaussian case in [4], [5], [6], [7], [8].

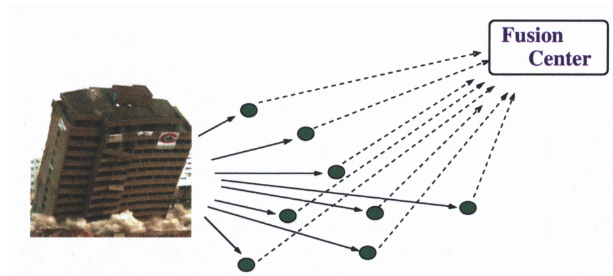

Fig. 1. A data-gathering wireless sensor network. The target data is observed by N sensors (the figure shows $N = 8$ sensors). The sensors encode and transmit their observations to a decoder or fusion center. The decoder aims to obtain an estimate of the source.

We assume that the coded sequences from sensors are transmitted to the FC through independent binary symmetric channels (BSCs). When the channels between sensors and the FC are ideal (i.e. with cross over probability of zero), FC may estimate the source by separately decoding the data received from each sensor and voting among the binary outcomes. However, for noisy channels this separate decoding generally leads to a suboptimal performance. In [9] an iterative joint decoding algorithm is proposed for data-gathering WSNs. After decoding the data received from each sensor, the soft values, i.e., log-likelihood ratios (LLRs), are passed to a separate unit that updates them by taking the statistical knowledge, i.e. correlation between transmitted data, into account. Updated information is returned to the decoders and the process continues for a definite number of iterations. This is similar to the problem of turbo equalization [10], [11] where the knowledge of inter-symbol interference is employed to update LLRs for the next round of decoding. An approach similar to [9] appears in a pioneer work of [12] where low-density generator matrix (LDGM) codes are applied for binary CEO model with independent additive white Gaussian noise (AWGN) channels between sensors and the FC. However, that work considers iterative decoding over the graph of the whole system (instead of iterating extrinsic information between distinct modules). This will reduce the bit error probability at the expense of increasing decoding complexity.

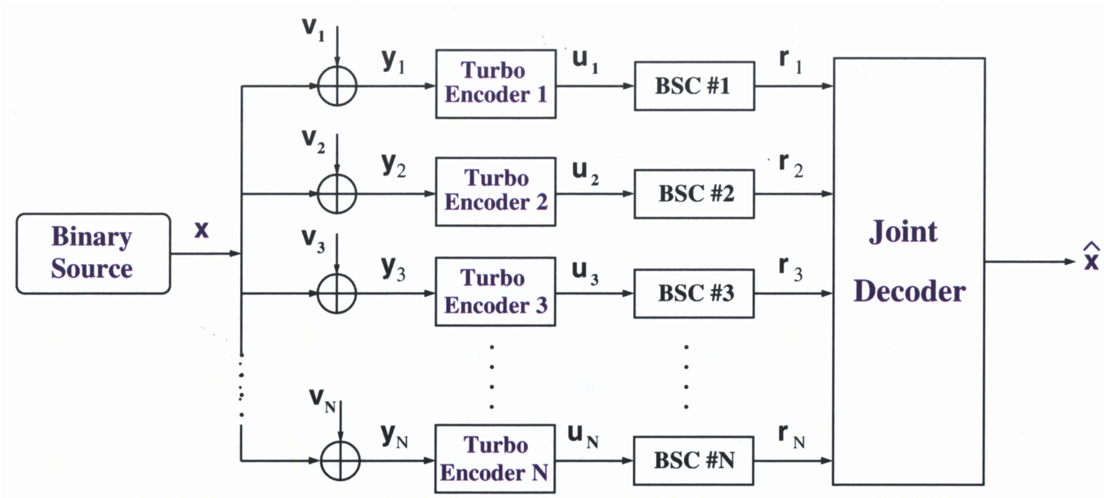

Fig. 2. A sensor network with binary CEO model. $\mathbf{x}$ is a sequence of i.i.d. Bernoulli random variables with $\Pr(x(k) = 0) = 0.5$. The vectors $\mathbf{y}_n$'s are observations of $\mathbf{x}$ through independent binary-symmetric channels (BSCs) with cross-over probabilities p_s. The sensors encode and transmit their observations as $\mathbf{u}'_n s$ with a rate R_n to a joint decoder through independent BSCs with cross-over probabilities p_c . The decoder fuses these N sequences together and provides an estimate of the source sequence $\mathbf{x}$ as $\hat{\mathbf{x}}$.

We apply turbo codes for data-gathering WSNs modeled by the binary CEO problem. A binary source corrupted by independent binary noises is observed by different sensors. Each sensor encodes its observation by two parallel concatenated convolutional codes, punctures parities to achieve the target rate, and transmits through BSC to the FC. The FC passes the extrinsic information between the SISO decoders related to each sensor, and meanwhile delivers extrinsic information received from SISO decoders of all sensors to two vertical decoder (VD) units that update extrinsic information by taking correlation into account. We compare the bit error rate of this scheme with the bit error rate achieved by the separate decoding scheme. We also compare the bit error rate achieved by applying turbo codes with the one achieved by applying convolutional codes and discuss the results. We should mention that a similar work on turbo encoding and decoding of correlated sources appreas in [13], where a correlation decoder is applied between the turbo decoders to provide updated probabilities to the turbo decoders. However, this correlation decoder is not the same as the decoder we propose here, as it addresses a different source model.

The rest of this paper is organized as follows. Section II presents the system model and definitions. In Section III we formulate the update rules for L-values in the iterative joint decoding scheme. Simulation results and discussions are presented in Section IV. Section V concludes the paper.

II. System Model

Figure 2 shows the binary CEO model for a data-gathering WSN. The source is an i.i.d. binary sequence $\mathbf{x} = \{x(k)\}_{k=1}^{K}$, $\Pr(x(k) = 0) = \Pr(x(k) = 1) = 0.5$ where K is the block length of the source sequence and is called the message block length. This symmetric source is monitored by N sensors. Sensor number n receives a noisy observation of the source sequence as $\mathbf{y}_n = \{y_n(k)\}_{k=1}^{K}$ after passing through a BSC with crossover probability p_s. In other words for each $1 \leq k \leq K$, $y_n(k) = x_n(k) \oplus v_n(k)$ where $\oplus$ denotes the modula 2 addition and $v_n(k)$ is a binary i.i.d. random variable that takes "1" with probability p_s and "0" with probability $1 - p_s$. Each sensor encodes its data using two parallel concatenated convolutional codes (turbo code). Turbo encoders of different sensors are not able to communicate with each other to directly exploit the correlation between their inputs. The coded sequence, $\mathbf{u}_n = \{u_n(k)\}_{k=1}^{K+M}$, consists of information bits, $\{u_n(k)\}_{k=1}^{K} = \{y_n(k)\}_{k=1}^{K}$, and parity bits $\{u_n(k)\}_{k=K+1}^{K+M}$ that are generated by the two convolutional codes and are punctured to achieve a desired rate. The positions of punctured parities are selected randomly, but are fixed and known at both encoder and decoder. The FC receives the data of all sensors as a matrix $\{r_n(k)\}_{k=1}^{K+M}$, $1 \leq n \leq N$ and estimates the source as a sequence $\{\hat{x}(k)\}_{k=1}^{K}$.

For ideal BSCs where $p_c = 0$, $\{y_n(k)\}_{k=1}^{K}$, $1 \leq n \leq N$ are available at the FC. It is straightforward to show that in this case maximum likelihood estimator is a simple detector that determines the value of the source by *voting*, as follows:

$$\hat{x}(k) = \begin{cases} 0 & \sum_{n=1}^{N} y_n(k) \leq \frac{N}{2} \\ 1 & \text{otherwise} \end{cases} \tag{1}$$

Note that when N is an even number, $\sum_{n=1}^{N} y_n(k) = \frac{N}{2}$ gives no information about $x(k)$, and the bit error probability remains 0.5. Therefore, in this case we let the estimator in (1) always reproduce a zero. Estimating $x(k)$ using (1) gives the bit error probability expressed by (2) (the probability that majority number of sensors vote for a value that is different from the actual value of $x(k)$). This is the minimum bit error probability that the system could achieve using any coding scheme. The irreducible minimum bit error probability given by (2) is due to the noisy source observations. For ideal channels, a simple uncoded transmission and voting will achieve this minimum. For non-ideal channels, this minimum can be achieved if there exists a channel code with rate $\frac{K}{K+M}$ at each sensor node that asymptotically achieves zero error probability for the BSC with crossover probability p_c (the channel capacity theorem requires $\frac{K}{K+M}$ to be greater than or equal to $1 + p_c \log p_c + (1 - p_c) \log(1 - p_c)$, the capacity of the BSC).

$$\Pr(\hat{x}(k) \neq x(k)) = \begin{cases} \frac{1}{2}\binom{N}{\frac{N}{2}} p_s^{\frac{N}{2}}(1-p_s)^{\frac{N}{2}} + \sum\limits_{n=\frac{N}{2}+1}^{N} \binom{N}{n} p_s^n (1-p_s)^{N-n} & N \quad \text{even} \\ \sum\limits_{n=\frac{N+1}{2}}^{N} \binom{N}{n} p_s^n (1-p_s)^{N-n} & N \quad \text{odd} \end{cases} \tag{2}$$

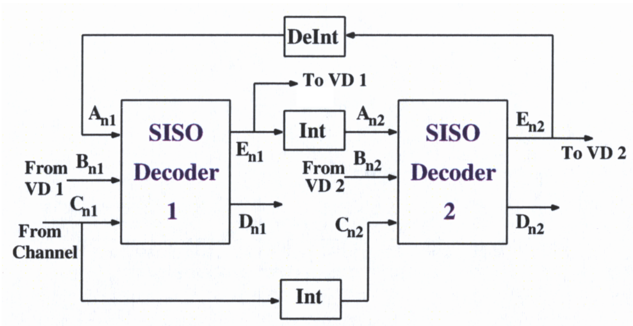

Fig. 3. Block diagram of horizontal decoder for sensor number n. VD denotes vertical decoder.

If such channel code exists, the FC can separately decode the data received from each sensor and then vote to determine the value of the source. However, in practice, when codes are of limited block lengths, a joint decoding approach is preferred and performs substantially better than separate decoding method (see Section IV). Such joint decoder applies its knowledge of the correlation between sensors' data, to update the extrinsic information. In Section III we show how to implement such joint decoding scheme for sensor networks equipped by turbo codes.

III. Implementation of Joint Decoding for Turbo Codes

The joint decoding scheme presented here is a connection-development between coding and data fusion steps. The joint decoding is performed in two steps. Following the notation of [14], we call these steps the *Horizontal* and the *Vertical* decoding. In the horizontal decoding, each two SISO decoders associated with a sensor node exchange their extrinsic information. Meanwhile, two separate vertical decoding units collect extrinsic information from the same SISO decoders of all sensors (i.e. vertical decoder 1 from all first SISO decoders, and vertical decoder 2 from all second SISO decoders) and update them for the next iteration.

Horizontal Decoding: Figure 3 shows the block diagram of horizontal decoder for sensor number n. For sensor n, SISO decoder number i, $i = 1, 2$, receives three vectors of L-values, $\mathbf{C}_{ni}$ received from the channel, $\mathbf{B}_{ni}$ received from the vertical decoder i, and $\mathbf{A}_{ni}$ *a priori* information received from the SISO decoder connected to it. For initialization, $\mathbf{A}_{ni}$ and $\mathbf{B}_{ni}$ are set to zero for $i = 1, 2$. The SISO decoder runs BCJR algorithm [15] to calculate the *a posteriori* information $\mathbf{D}_{ni}$, and expresses the extrinsic information, $\mathbf{E}_{ni}$, by subtracting *a priori* L-values from *a posteriori* L-values. The extrinsic information is interleaved (deinterleaved) to act as *a priori* L-values for the other SISO decoder. $\mathbf{E}_{ni}$ is also delivered to the vertical decoder to extract $\mathbf{B}_{ni}$ for the next decoding iteration.

Vertical Decoding: The ith vertical decoder ($i = 1, 2$) collects a posteriori L-values $\mathbf{D}_{ni}$ for all sensors $n = 1, \cdots, N$. Fix k and let n change from 1 to N. The vertical decoder starts by calculating initial probabilities:

$$\Pr(y_n(k) = 0) = \frac{\exp(D_{ni}(k))}{1 + \exp(D_{ni}(k))},$$

and

$$\Pr(y_n(k) = 1) = \frac{1}{1 + \exp(D_{ni}(k))}.$$

Therefore, if $f(y_n(k))$ represents the probability mass function of $y_n(k)$, we have:

$$f(y_n(k)) = \begin{cases} \frac{\exp(D_{ni}(k))}{1+\exp(D_{ni}(k))} & y_n(k) = 0 \\ \frac{1}{1+\exp(D_{ni}(k))} & y_n(k) = 1 \end{cases} \tag{3}$$

Then, it estimates the joint probability mass function $f(x(k), y_1(k), ..., y_N(k))$ as:

$$f(x(k), y_1(k), ..., y_N(k)) = \frac{1}{2} \alpha p_s^w (1-p_s)^{N-w} \prod_{n=1}^{N} f(y_n(k)), \tag{4}$$

where

$$w = \sum_{n=1}^{N} \left((x(k) \oplus y_n(k) \right).$$

In (4), the coefficient $\frac{1}{2}$ represents the probability of $x(k)$ and α is a coefficient that normalizes the following summation to one:

$$\sum_{x(k)=0}^{1} \sum_{y(1,k)=0}^{1} \ldots \sum_{y(N,k)=0}^{1} f(x(k), y_1(k), ..., y_N(k)) = 1.$$

Now, for each n the vertical decoder calculates the marginal probability mass function:

$$\begin{aligned} \tilde{f}(y_n(k)) = \beta \times & \sum_{x(k)=0}^{1} \sum_{y_1(k)=0}^{1} \ldots \sum_{y_{n-1}(k)=0}^{1} \\ & \sum_{y_{n+1}(k)=0}^{1} \ldots \sum_{y_N(k)=0}^{1} f(x(k), y_1(k), ..., y_N(k)), \end{aligned} \tag{5}$$

for $y_n(k) = 0, 1$, respectively. Again, coefficient β is used for normalization. We note that the marginal probability mass function contains the information about $y_n(k)$ provided by all other nodes (excluding $y_n(k)$). Therefore, the *extrinsic* L-value is expressed by:

$$B_{ni}(k) = \log \frac{\tilde{f}(0)}{\tilde{f}(1)}. \tag{6}$$

The vertical decoder repeats the same procedure for all $1 \leq k \leq K$ to calculate the vector $\{B_{ni}(k)\}_{k=1}^{K}$, and feeds it back to the SISO decoders.

After running a definite number of horizontal-vertical iterations, each decoder ($n = 1, \cdots, N$) makes a binary decision on $\hat{x}(k)$ by looking at $D_{n2}(k)$, and the final value of $\hat{x}(k)$ is defined as the majority of these decisions. Notice that although similar to separate decoding, $\hat{x}(k)$ is found by voting, several rounds of vertical decoding have provided additional information that increases the reliability of the decision. In Section IV we show that the joint design of decoding and data fusion leads to substantial performance gain compared with decoupled designs.

IV. Simulation Results

In all simulations we consider systematic recursive convolutional codes with generator polynomials $G(D) = (1, \frac{1+D^2}{1+D+D^2})$. For separate and joint decoding of turbo codes, as well as joint decoding of convolutional codes, the number of iterations is fixed to 4. The source block length is fixed to $K = 1000$ bits and number of sensors is fixed to $N = 4$. For the case of convolutional codes, only one SISO decoder represents each sensor node at the decoder and all these SISO decoders are connected and exchange their information via a single vertical decoder. We note that uncoded transmission over the BSC results in a total crossover probability of $p_s(1-p_c)+p_c(1-p_s)$ between $x(k)$ and each received bit $r_n(k)$. Thus, bit error probability of this transmission could be found after replacing p_s in (2) by $p_s(1-p_c)+p_c(1-p_s)$. These bit error rate curves are provided in all figures for comparison. As discussed before, bit error rate achieved by uncoded transmission for ideal channels ($p_c = 0$) indicates the minimum achievable bit error probability of the system (Equation (2)).

Figure 4 shows the bit error rate achieved by separate and joint decoding of turbo codes, respectively. The number of sensors is set to $N = 4$ and the crossover probability between source and sensors is $p_s = 0.01$. Separate encoding reaches the minimum achievable bit error probability (3.0×10^{-4}) for channels with cross over probability $p_c \leq 0.05$, while joint decoding achieves this minimum for all channels with $p_c \leq 0.14$. As an example to compare the performance of joint and separate decoding, assume that a bit error rate of 10^{-3} is desired. Joint decoding can attain this criterion for channels with crossover probability of $p_c = 0.16$ while separate decoding requires a crossover probability less than 0.08. This means that joint decoding can tolerate twice more channel errors compared with the separate decoding.

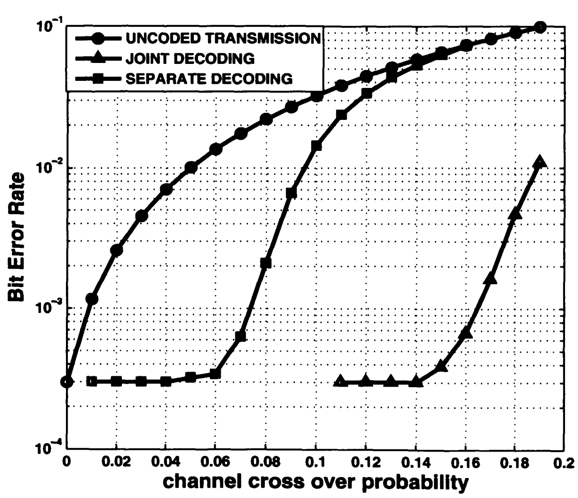

Fig. 4. Bit error rate achieved by $(2000, 1000)$ turbo codes for $N = 4$ sensors and $p_s = 0.01$.

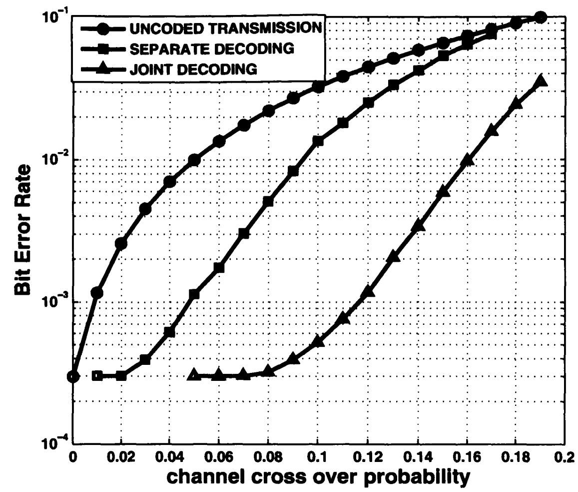

Fig. 5. Bit error rate achieved by $(2000, 1000)$ convolutional codes for $N = 4$ sensors and $p_s = 0.01$.

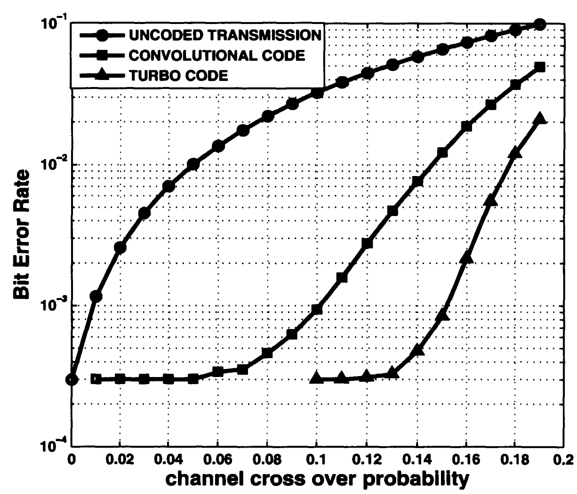

Fig. 6. Bit error rate achieved by joint decoding for $(1800, 1000)$ turbo and convolutional codes. $N = 4$ and $p_s = 0.01$.

Figure 5 shows similar results when turbo codes are replaced by convolutional codes. By comparing the two curves for bit error rate of 10^{-3} we observe that joint decoding tolerates 2.4 times more channel errors ($p_c = 0.05$ for separate decoding and $p_c = 0.12$ for joint decoding). Also, comparing the curves for joint decoding in Fig. 4 and Fig. 5 shows that turbo codes can keep the BER below 10^{-3} for all $p_c \leq 0.16$, while convolutional codes require $p_c \leq 0.12$. This means that turbo codes can correct 33% more channel errors. This percentage will increase

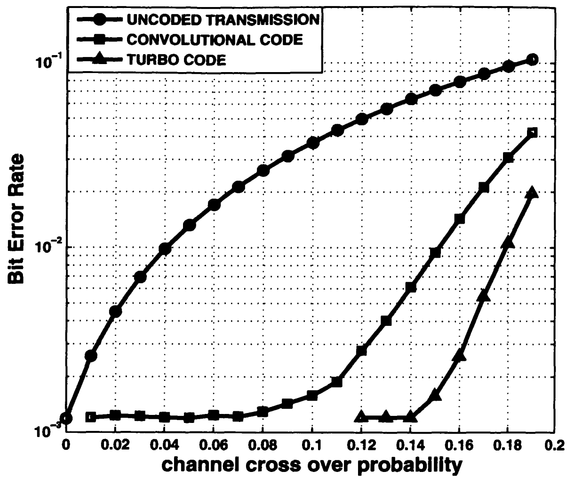

Fig. 7. Bit error rate achieved by joint decoding for $(2000, 1000)$ turbo and convolutional codes. $N = 4$ and $p_s = 0.02$.

by increasing the code rate. This is illustrated in Fig. 6, where $(1800, 1000)$, i.e. rate $\frac{5}{9}$ codes are applied. Turbo codes and convolutional codes reach bit error rate of 10^{-3} at $p_c = 0.15$, and 0.10 respectively. In other words, turbo codes tolerate 50% more channel errors (they could correct 1.5 times more channel errors), while maintaining the same bit error rate.

Finally, Fig. 7 compares the performance of $(2000, 1000)$ turbo and convolutional codes for $N = 4$ sensors and $p_s = 0.02$. The minimum achievable bit error probability is 1.2×10^{-3}. Turbo codes reach this minimum at $p_c = 0.14$ whereas convolutional codes achieve it for $p_c = 0.07$. We observe that by reducing the code rate, both convolutional and turbo codes gain better error correction properties. However, this gain is more significant in the case of turbo codes; and therefore the gap between the two bit error rate curves increases by decreasing the rate.

V. Conclusion

We considered joint design of channel coding and data fusion steps in data-gathering sensor networks using turbo codes. The design is based on exploiting the knowledge of correlation between sensors' data at the fusion center. Two vertical decoders located at the fusion center, employ this knowledge to update extrinsic information received from SISO decoders of all sensors. We showed how extrinsic information is passed between $2 \times N$ SISO decoders and the vertical decoders. We evaluated the performance of joint decoding algorithm for systems modeled by binary CEO problem, where N sensors observe the same binary source corrupted by independent binary noises, encode and transmit their information to the FC through independent BSCs. Our simulation results show that the joint design of decoding and data fusion steps allows for substantial performance gains over separate decoding and data fusion designs. We also provided comparisons between the bit error rates achieved by convolutionally coded and turbo coded systems for different correlations and rates. The results show that turbo codes are able to correct up to 1.5 times more channel errors while maintaining the same bit error rate.

Acknowledgment

The authors would like to thank the Natural Science and Engineering Council of Canada for its support.

References

[1] I. F. Akyildiz, W. Su, and E. C. Y. Sankarasubramaniam, "A survey on sensor networks," *IEEE Commun. Mag.*, vol. 40, no. 8, pp. 102–114, Aug. 2002.

[2] A. J. Goldsmith and S. B. Wicker, "Design challenges for energy-constrained ad hoc wireless networks," *IEEE Wireless Commun. Mag.*, vol. 9, no. 4, pp. 8–27, Aug. 2002.

[3] T. Berger, Z. Zhang, and H. Viswanathan, "The CEO problem," *IEEE Trans. Inf. Theory*, vol. 42, no. 3, pp. 887–902, May 1996.

[4] H. Viswanathan and T. Berger, "The quadratic Gaussian CEO problem," *IEEE Trans. Inf. Theory*, vol. 43, no. 5, pp. 1549–1559, Sep. 1997.

[5] Y. Oohama, "Multiterminal source coding for correlated memoryless Gaussian sources with several side information at the decoder," in *Proc. IEEE Inf. Theory and Comm. Workshop*, Jun. 1999, p. 100.

[6] J. Chen, X. Zhang, T. Berger, and S. B. 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.

[7] V. Prabhakaran, D. Tse, and K. Ramchandran, "Rate region of the quadratic Gaussian CEO problem," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Chicago, IL, USA, Jun. 2004, p. 119.

[8] Y. Oohama, "Rate-distortion theory for Gaussian multiterminal source coding systems with several side informations at the decoder," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2577–2593, Jul. 2005.

[9] S. Howard and P. Flikkema, "Integrated source-channel decoding for correlated data-gathering sensor networks," in *Proc. of IEEE Wireless Communications and Networking Conference*, Las Vegas, NV, Mar. 2008.

[10] C. Douillard, M. Jezequel, C. Berrou, A. Picart, P. Didier, and A. Glavieux, "Iterative correction of intersymbol interference: Turbo equalization," *European Trans. on Telecomm.*, vol. 6, pp. 507–511, Sep. 1995.

[11] R. Koetter, A. Singer, and M. Tuchler, "Turbo equalization," *IEEE Signal Process. Mag.*, vol. 21, no. 1, pp. 67–80, Jan. 2004.

[12] W. Zhong and J. Garcia-Frias, "Combining data fusion with joint source-channel coding of correlated sensors," in *Proc. IEEE Inf. Theory Workshop (ITW)*, San Antonio, Texas (invited paper), Oct. 2004.

[13] Y. Zhao and J. Garcia-Frias, "Joint estimation and compression of nonbinary sources using punctured turbo codes," *IEEE Trans. Commun.*, vol. 53, no. 3, pp. 385–390, Mar. 2005.

[14] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Trans. Inf. Theory*, vol. 42, no. 2, pp. 429–445, Mar. 1996.

[15] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inf. Theory*, vol. 20, no. 2, pp. 284–287, Mar. 1974.