

Research Article

Iterative Fusion of Distributed Decisions over the Gaussian Multiple-Access Channel Using Concatenated BCH-LDGM Codes

Javier Del Ser,¹ Diana Manjarres,¹ Pedro M. Crespo,² Sergio Gil-Lopez,¹
and Javier Garcia-Frias³

¹TECNALIA-TELECOM, P. Tecnológico, Ed. 202, 48170 Zamudio, Spain

²CEIT and TECNUN (University of Navarra), 20009 Donostia-San Sebastian, Spain

³Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19716, USA

Correspondence should be addressed to Javier Del Ser, javier.delser@tecnalia.com

Received 30 November 2010; Accepted 20 January 2011

Academic Editor: Claudio Sacchi

Copyright © 2011 Javier Del Ser et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper focuses on the data fusion scenario where N nodes sense and transmit the data generated by a source S to a common destination, which estimates the original information from S more accurately than in the case of a single sensor. This work joins the upsurge of research interest in this topic by addressing the setup where the sensed information is transmitted over a Gaussian Multiple-Access Channel (MAC). We use Low Density Generator Matrix (LDGM) codes in order to keep the correlation between the transmitted codewords, which leads to an improved received Signal-to-Noise Ratio (SNR) thanks to the constructive signal addition at the receiver front-end. At reception, we propose a joint decoder and estimator that exchanges soft information between the N LDGM decoders and a data fusion stage. An error-correcting Bose, Ray-Chaudhuri, Hocquenghem (BCH) code is further applied to suppress the error floor derived from the ambiguity of the MAC channel when dealing with correlated sources. Simulation results are presented for several values of N and diverse LDGM and BCH codes, based on which we conclude that the proposed scheme outperforms significantly (by up to 6.3 dB) the suboptimum limit assuming separation between Slepian-Wolf source coding and capacity-achieving channel coding.

1. Introduction

During the last years, the scientific community has experienced an ever-growing research interest in Sensor Networks (SN) as means to efficiently monitor physical or environmental conditions without necessitating expensive deployment and/or operational costs. Generally speaking, these communication networks consist of a large number of nodes deployed over a certain geographical area and with a high degree of autonomy. Such an increased autonomy is usually attained by means of advanced battery designs, an efficient exploitation of the available radio resources, and/or cooperative communication schemes and protocols. In fact, cooperation between nearby sensors permits the network to operate as a global entity and execute actions in a computationally cheap *albeit* reliable fashion. Unfortunately, the capacity of SNs to achieve a high energy efficiency is highly determined by the scalability of these sensor

meshes. In this context, a large number of challenging paradigms have been tackled with the aim of minimizing the power consumption and improving the battery lifetime of densely populated networks. As such, it is worth to mention distributed compression [1, 2], transmission and/or cluster scheduling [3, 4], data aggregation [5–7], multihop cooperative processing [8, 9], in-network data storage [10], and power harvesting [11, 12]. This work gravitates on one of such paradigms: the centralized data fusion scenario (see Figure 1), where N nodes monitor a given information source \mathcal{S} (representing, for instance, temperature, pressure, or any other physical phenomena) and transmit their sensed data to a common receiver. This receiver will combine the data from the sensors so as to obtain a reliable estimation of the information from the original source \mathcal{S} . When the monitoring procedure at each node is subject to a non-zero probability of sensing error, intuitively one can infer that the more sensors added to this setup, the higher the accuracy

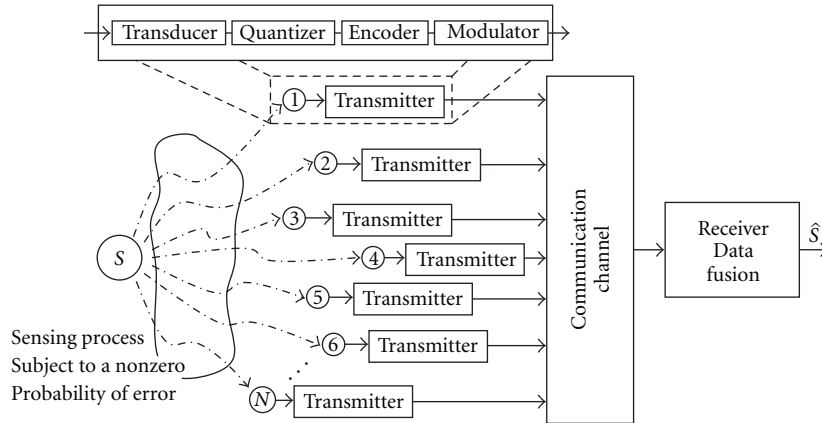


FIGURE 1: Generic data fusion scenario where N nodes sense a certain physical parameter \mathcal{S} , and transmit the sensed information to a joint receiver.

of the estimation will be with respect to the case of a single sensor. Therefore, the challenging paradigm in this specific scenario lies on how to optimally fuse the information from all sources while taking into account the aforementioned probability of sensing error, specially when dealing with practical communication channels.

One of the first contributions in this area was done by Lauer et al. in [13], who extended classical results from decision theory to the case of distributed correlated signals. Subsequently, Ekchian and Tenney [14] formulated the distributed detection problem for several network topologies. Later, in [15] Chair and Varshney derived an optimum data fusion rule which combines individually performed decisions on the data sensed at every sensor. This data fusion rule was shown to minimize the end-to-end probability of error of the overall system. More recently, several contributions have tackled the data fusion problem in diverse uncoded communication scenarios, for example, multihop networks subject to fading [16–18] and delays [19], parallel channels subject to fading [20–22], and asynchronous multiple-access channels [23, 24], among others.

On the other hand, when dealing with coded scenarios over noisy channels, it is important to point out that the data fusion problem can be regarded as a particular case of the so-called *distributed joint source-channel coding* of correlated sources, since the nonzero probability of sensing error imposes a spatial correlation among the data registered by the sensors. In the last decade, intense research effort has been conducted towards the design of practical iteratively-decodable (i.e., Turbo-like) joint source-channel coding schemes for the transmission of spatially and temporally correlated sources over diverse communication channels, for example, see [25–31] and references therein. However, these contributions address the reliable transmission of the information generated by a set of correlated sensors, whereas the encoded data fusion paradigm focuses on the reliable communication of an information source \mathcal{S} read by a set of N sensors subject to a nonzero probability of sensing error; based on this, a certain error tolerance can be permitted when detecting the data registered by a given sensor. In

this encoded data fusion setup, different Turbo-like codes have been proposed for iterative decoding and data fusion of multiple-sensor scenarios for the simplistic case of parallel AWGN channels, for example, Low Density Generator Matrix (LDGM) [32], Irregular Repeat-Accumulate (IRA) [33], and concatenated Zigzag [34] codes. In such references, it was shown that an iterative joint decoding and data fusion strategy performs better than a sequential scheme where decoding and data fusion are separately executed.

Following this research trend, this paper considers the data fusion scenario where the data sensed by N nodes is transmitted to a common receiver over a Gaussian Multiple-Access Channel (MAC). In this scenario, it is well known that the spatial correlation between the data registered by the sensors should be preserved between the transmitted signals so as to maximize the effective signal-to-noise ratio (SNR) at the receiver. On this purpose, correlation-preserving LDGM codes have been extensively studied for the problem of joint source-channel coding of correlated sensors over the MAC [35–38]. In these references, it was shown that concatenated LDGM schemes permit to drastically reduce the error floor inherent to LDGM codes. Inspired by this previous work, in this paper we take a step further by analyzing the performance of concatenated BCH-LDGM codes for encoded data fusion over the Gaussian MAC. Specifically, our contribution is twofold: on one hand, we design an iterative receiver that jointly performs LDGM decoding and data fusion based on factor graphs and the Sum Product Algorithm. On the other hand, we show that for the particular data fusion scenario under consideration, the error statistics in the decoded information from the sensors allow for the concatenation of BCH codes [39, 40] in order to decrease the aforementioned intrinsic error floor of single LDGM codes. Extensive Monte Carlo simulations will verify that the proposed concatenated BCH-LDGM codes not only outperform vastly the suboptimum limit assuming separation between distributed source and channel coding, but also reaches the theoretical residual error bound derived by assuming errorless detection and decoding of the sensor data.

The rest of the paper is organized as follows: Section 2 delves into the system model of the considered encoded data fusion scenario, whereas Section 3 elaborates on the design of the iterative decoding and data fusion procedure. Next, Section 4 discusses Monte Carlo simulation results and finally, Section 5 ends the paper by drawing some concluding remarks.

2. System Model

Figure 2 depicts the system model considered in this work. The information corresponding to a source \mathcal{S} (e.g., representing a physical parameter such as temperature) is modeled as a sequence of K i.i.d binary random variables $\{x_k^{\mathcal{S}}\}_{k=1}^K$, with $P_{x_k^{\mathcal{S}}}(0) = P_{x_k^{\mathcal{S}}}(1) = 0.5$ for all k . A set of N sensors $\{S_n\}_{n=1}^N$ registers blocks of length K $\{x_k^n\}_{k=1}^K$ ($n = 1, \dots, N$) from \mathcal{S} , subject to a probability of sensing error $p_n = \Pr\{x_k^n \neq x_k^{\mathcal{S}}\}$ for all $k \in \{1, \dots, K\}$, with $0 < p_n < 0.5$ for all $n \in \{1, \dots, N\}$. The sensed sequence at each sensor is then encoded through an outer systematic BCH code (L_{out}, K, t) , where L_{out} and t denote the output sequence length and error correction capability of the code, respectively (We hereafter adopt this nomenclature, which differs from the standard notation (L_{out}, K, d) , with d denoting the minimum distance of the BCH code.). The encoded sequence at the output of the BCH encoder is next processed through an inner LDGM code, that is, a linear code with low density generator matrix $\mathbf{G} = [\mathbf{I} \ \mathbf{P}]$. The parity check matrix of LDGM codes is expressed as $\mathbf{H} = [\mathbf{P}^T \ \mathbf{I}]$, where \mathbf{I} denotes the identity matrix, and \mathbf{P} is a $L_{\text{out}} \times (L - L_{\text{out}})$ sparse binary matrix. Variable and check degree distributions (In other words, the parity matrix \mathbf{P} of a (d_v, d_c) LDGM code has exactly d_v nonzero entries per row and d_c nonzero entries per column.) are denoted as $[d_v \ d_c]$; the overall coding rate is thus given by $R_c = R_{\text{out}}L_{\text{out}}/L = R_{\text{out}}d_c/(d_c + d_v)$, where R_{out} is the rate of the outer BCH code. Notice that due to the low density nature of LDGM matrices, correlation is preserved not only in the systematic bits but also in the coded bits. Therefore, in order to exploit this correlation, the generator matrices are set exactly the same for all sensors. The output sequence of the concatenated encoder at every sensor, $\{c_l^n\}_{l=1}^L$, is composed by a first set of K bits corresponding to the systematic bits

$\{x_l^n\}_{l=1}^K$, followed by a set of $K - L_{\text{out}}$ BCH parity bits $\{p_l^n\}_{l=K+1}^{L_{\text{out}}}$ and a final set of $L - L_{\text{out}}$ LDGM parity bits $\{p_l^n\}_{l=L_{\text{out}}+1}^L$. These encoded sequences are then BPSK (*Binary Phase Shift Keying*) modulated and transmitted to a common receiver over a Gaussian Multiple-Access Channel.

The signal at the receiver is expressed as

$$y_l = \sum_{n=1}^N h_l^n \phi(c_l^n) + n_l = b_l + n_l, \quad (1)$$

where $\phi : \{0, 1\} \rightarrow \{-\sqrt{E_c}, +\sqrt{E_c}\}$ stands for the BPSK modulation mapping, and E_c represents the average energy per channel symbol and sensor. The Gaussian MAC considered in this work assumes $h_l^n = 1$ for all $l \in \{1, \dots, L\}$ and for all $n \in \{1, \dots, N\}$, whereas $\{n_l\}_{l=1}^L$ are i.i.d. circularly symmetric complex Gaussian random variables with zero mean and variance per dimension σ^2 . Nevertheless, explanations hereafter will make no assumptions on the value of the MAC coefficients. The joint receiver must estimate the original information $\{x_k^{\mathcal{S}}\}_{k=1}^K$ generated by \mathcal{S} as $\{\hat{x}_k^{\mathcal{S}}\}_{k=1}^K$ based on the received sequence $\{y_l\}_{l=1}^L$. This will be done by applying the message-passing Sum-Product Algorithm (SPA, see [41] and references therein) over the whole factor graph describing the statistical dependence between $\{y_l\}_{l=1}^L$ and $\{x_k^{\mathcal{S}}\}_{k=1}^K$, as will be explained in next section.

3. Iterative Joint Decoding and Data Fusion

In order to estimate the aforementioned original information sequence $\{\hat{x}_k^{\mathcal{S}}\}_{k=1}^K$, the optimum joint receiver would symbolwise apply the Maximum A Posteriori (MAP) decision criterium, that is,

$$\hat{x}_k^{\mathcal{S}} = \arg \max_{x_k^{\mathcal{S}} \in \{0, 1\}} P(x_k^{\mathcal{S}} | \{y_l\}_{l=1}^L), \quad k = 1, \dots, K, \quad (2)$$

where $P(\cdot | \cdot)$ denotes conditional probability. To efficiently perform the above decision criterion, a suboptimum practical scheme would first compute the conditional probabilities of the encoded symbol c_l^n given the received sequence, which is given, for $l \in \{1, \dots, L\}$ and $n \in \{1, \dots, N\}$, as

$$\begin{aligned} P(c_l^n | y_l) &= \sum_{\sim c_l^n} P(c_l^1, \dots, c_l^N | y_l) \\ &\propto \sum_{\sim c_l^n} \exp\left(-\frac{|y_l - \phi(c_l^1)h_l^1 - \dots - \phi(c_l^{n-1})h_l^{n-1} - \phi(c_l^{n+1})h_l^{n+1} - \dots - \phi(c_l^N)h_l^N|^2}{2\sigma^2}\right), \end{aligned} \quad (3)$$

where the proportionality stands for $P(0 | y_l) + P(1 | y_l) = 1$ for all $l \in \{1, \dots, L\}$, and $\sim c_l^n$ denotes that all binary variables are included in the sum except c_l^n , that is, the sum is evaluated for all the 2^{N-1} possible

combinations of the set $\{c_l^1, \dots, c_l^{n-1}, c_l^{n+1}, \dots, c_l^N\}$. Once the L conditional probabilities for the n th sensor codeword $\{c_l^n\}_{l=1}^L$ are computed, an estimation $\{\hat{x}_k^{\mathcal{S}}\}_{k=1}^K$ of the original sensor sequence $\{x_k^{\mathcal{S}}\}_{k=1}^K$ would be obtained by performing

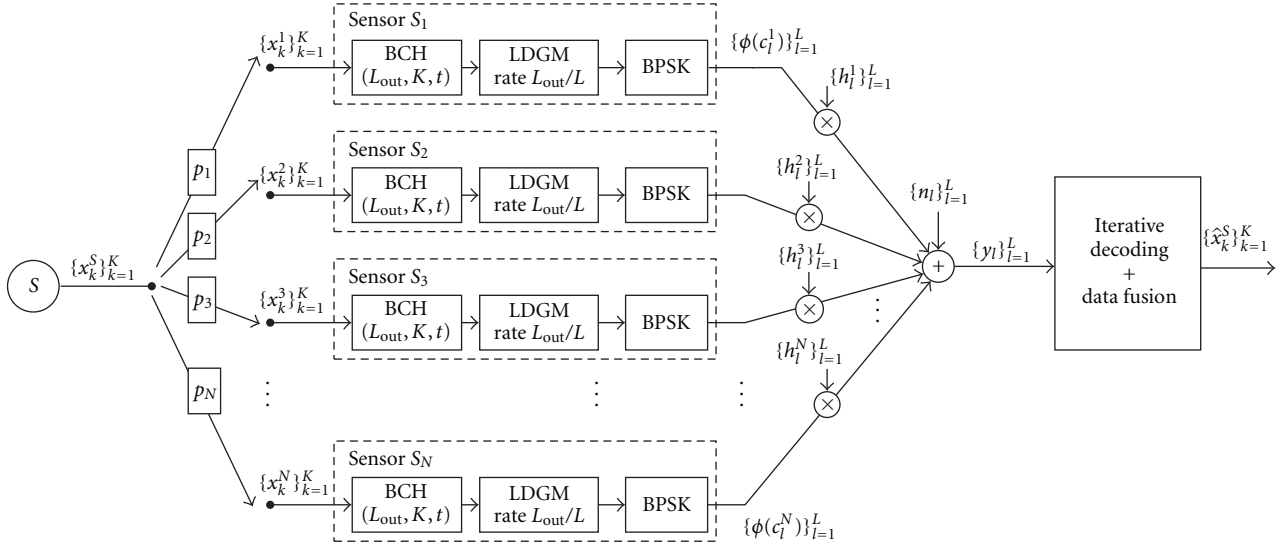


FIGURE 2: Block diagram of the considered scenario.

(1) iterative LDGM decoding based on $\{P(c_l^n | y_l)\}_{l=1}^L$ in an independent fashion with respect to the LDGM decoding procedures of the other $N - 1$ sensors and (2) an outer BCH decoding based on the hard-decoded sequence at the output of the LDGM decoder. Finally, the N recovered sensor sequences $\{\hat{x}_k^n\}_{k=1}^K$ ($n \in \{1, \dots, N\}$) would be fused to render the estimation $\{\hat{x}_k^s\}_{k=1}^K$ as

$$\hat{x}_k^s = \begin{cases} 1 & \text{if } \sum_{n=1}^N \hat{x}_k^n \geq \left\lceil \frac{N}{2} \right\rceil, \\ 0 & \text{if } \sum_{n=1}^N \hat{x}_k^n < \left\lceil \frac{N}{2} \right\rceil, \end{cases} \quad (4)$$

that is, by symbolwise majority voting over the estimated N sensor sequences. Notice that this practical scheme performs sequentially channel detection, LDGM decoding, BCH decoding, and fusion of the decoded data.

However, the performance of the above separate approach can be easily outperformed if one notices that, since we assume $0 < p_n < 0.5$ for all $n \in \{1, \dots, N\}$ (see Section 2), the sensor sequences $\{x_k^n\}_{k=1}^K$ are symbolwise spatially correlated, that is

$$\Pr\{x_k^m = x_k^n\} = p_m p_n + (1 - p_m)(1 - p_n) > 0.5, \quad (5)$$

for $n \neq m$. As widely evidenced in the literature related to the transmission of correlated information sources (see references in Section 1), this correlation should be exploited at the receiver in order to enhance the reliability of the fused sequence $\{\hat{x}_k^s\}_{k=1}^K$. In other words, the considered scenario should take advantage of this correlation, not only by means of an enhanced effective SNR at the receiver thanks to the correlation-preserving properties of LDGM codes, but also through the exploitation of the statistical relation between sequences $\{x_k^n\}_{k=1}^K$ corresponding to different sensors $n \in$

$\{1, \dots, N\}$. The latter dependence between $\{x_k^n\}_{n=1}^N$ and x_k^s can be efficiently capitalized by (1) describing the joint probability distribution of all the variables involved in the system by means of factor graphs and (2) marginalizing for \hat{x}_k^s via the message-passing Sum-Product Algorithm (SPA). This methodology allows decreasing the computational complexity with respect to a direct marginalization based on exhaustive evaluation of the entire joint probability distribution. Particularly, the statistical relation between sensor sequences is exploited in one of the compounding factor subgraphs of the receiver, as will be later detailed.

This factor graph is exemplified in Figure 3(a), where the graph structure of the joint detector, decoder, and data fusion scheme is depicted for $N = 4$ sensors. As shown in this plot, this graph is built by interconnecting different subgraphs: the graph modeling the statistical dependence between x_k^s and $\{x_k^n\}_{n=1}^N$ for all $k \in \{1, \dots, K\}$ (labeled as SENSING), the factor graph that relates sensor sequence $\{x_k^n\}_{k=1}^K$ to codeword $\{c_l^n\}_{l=1}^L$ through the LDGM parity check matrix \mathbf{H} and the BCH code (to be later detailed), and the relationship between the received sequence $\{y_l\}_{l=1}^L$ and the N codewords $\{c_l^n\}_{l=1}^L$, with $n \in \{1, \dots, N\}$ (labeled as MAC). Observe that the interconnection between subgraphs is done via variable nodes corresponding to c_l^n and x_k^n . In this context, since the concatenation of the LDGM and BCH code is systematic, variable nodes $\{c_l^n\}_{l=1}^K$ and $\{x_k^n\}_{k=1}^K$ collapse into a single node for all $n \in \{1, \dots, N\}$, which has not been shown in the plots for the sake of clarity. Before delving into each subgraph, it is also important to note that this interconnected set of subgraphs embodies an overall cyclic factor graph over which the SPA algorithm iterates—for a fixed number of iterations ℓ —in the order $\text{MAC} \rightarrow \text{LDGM}_1 \rightarrow \text{BCH}_1 \rightarrow \text{LDGM}_2 \rightarrow \dots \rightarrow \text{LDGM}_N \rightarrow \text{BCH}_N \rightarrow \text{SENSING}$.

Let us start by analyzing the MAC subgraph, which is represented in Figure 3(b). Variable nodes $\{c_l^n\}_{n=1}^N$ are linked to the received symbol y_l through the auxiliary variable

In regard to Figure 3(b), observe that a set of switches controlled by binary variables μ_1 and μ_2 drive the connection/disconnection of systematic ($l \in \{1, \dots, K\}$) and parity ($l \in \{K+1, \dots, L\}$) variable nodes from the MAC subgraph. The reason being that, as later detailed in Section 4, the degradation of the iterative SPA due to short-length cycles in the underlying factor graph can be minimized by properly setting these switches.

The analysis follows by considering Figure 3(c), where the block integrating the BCH decoder is depicted in detail. At this point it is worth mentioning that the rationale behind concatenating the BCH code with the LDGM code lies on the statistics of the errors per simulated block, as the simulation results in Section 4 will clearly show. Based on these statistics, it is concluded that such an error floor is due to most of the simulated blocks having a low number of symbols in error, rather than few blocks with errors in most of their constituent symbols. Consequently, a BCH code capable of correcting up to t errors can be applied to detect and correct such few errors per block at a small loss in performance. Having said this, the integration of the BCH decoder in the proposed iterative receiver requires some preliminary definitions.

- (i) $\delta_{k,j}^n(x)$: *a posteriori* soft information for the value $x \in \{0, 1\}$ of the node x_k^n , which is computed, at iteration j and $k \in \{1, \dots, K\}$, as the product of the *a posteriori* soft information rendered by the SPA when applied to MAC and LDGM subgraphs.
- (ii) $\delta_{l,j}^n(c)$: similar to the previously defined $\delta_{k,j}^n(x)$, this notation refers to the *a posteriori* information for the value $c \in \{0, 1\}$ of node c_l^n , which is calculated, at iteration j and $l \in \{K+1, \dots, L_{\text{out}}\}$, as the product of the corresponding *a posteriori* information produced at both MAC and LDGM subgraphs.
- (iii) $\xi_{k,j}^n(x)$: *extrinsic* soft information for $x_k^n = x \in \{0, 1\}$ built upon the information provided by the rest of sensors at iteration j and time tick $k \in \{1, \dots, K\}$.
- (iv) $\widetilde{\delta}_{k,j}^n(x)$: refined *a posteriori* soft information of node x_k^n for the value $x \in \{0, 1\}$, which is produced as a consequence of the processing stage in Figure 3(c).

Under the above definitions, the processing scheme depicted in Figure 3(c) aims at refining the input soft information coming from the MAC and LDGM subgraphs by first performing a *hard decision* (HD) on the BCH encoded sequence based on $\{\delta_{k,j}^n(x)\}_{k=1}^K$, $\{\delta_{l,j}^n(c)\}_{l=K+1}^{L_{\text{out}}}$, and the information output from the SENSING subgraph in the previous iteration, that is, $\{\xi_{k,j-1}^n(x)\}_{k=1}^K$. This is done for all $n \in \{1, \dots, N\}$ within the current iteration j . Once the binary estimated sequence $\{\widehat{c}_{l,j}^n\}_{l=1}^{L_{\text{out}}}$ corresponding to the BCH encoded block at the n th sensor is obtained and decoded, the binary output $\{\widehat{x}_{k,j}^n\}_{k=1}^K$ is utilized for adaptively

refining the *a posteriori* soft information $\{\delta_{k,j}^n(x)\}_{k=1}^K$ as $\{\widetilde{\delta}_{k,j}^n(x)\}_{k=1}^K$ under the flipping rule

$$\widetilde{\delta}_{k,j}^n(x) = \begin{cases} \max\{\delta_{k,j}^n(0), \delta_{k,j}^n(1)\} & \text{if } \widehat{x}_{k,j}^n = x, \\ \min\{\delta_{k,j}^n(0), \delta_{k,j}^n(1)\} & \text{if } \widehat{x}_{k,j}^n \neq x, \end{cases} \quad (8)$$

which is performed for $k \in \{1, \dots, K\}$. It is interesting to observe that in this expression, all those indices in error detected by the BCH decoder will consequently drive a flip in the soft information fed to the SENSING subgraph.

Finally we consider Figure 3(c) corresponding to the SENSING subgraph, where the refined soft information from all sensors is fused to provide an estimation of x_k^s as \widehat{x}_k^s . Let $\chi_{k,j}^n(x)$ denote the soft information on x_k^s (for the value $x \in \{0, 1\}$ and computed for $k \in \{1, \dots, K\}$) contributed by sensor S_n at iteration j . The SPA applied to this subgraph renders (see [41, equations (5) and (6)])

$$\chi_{k,j}^n(x) = \Gamma_{k,j}^n \left[(1 - p_n) \widetilde{\delta}_{k,j}^n(x) + p_n \widetilde{\delta}_{k,j}^n(1 - x) \right], \quad (9)$$

where p_n denotes the sensing error probability which in turn establishes the amount of correlation between sensors. Factors $\Gamma_{k,j}^n$ account for the normalization of each pair of messages, that is, $\xi_{k,j}^n(0) + \xi_{k,j}^n(1) = 1$ for all k, n, j . The estimation $\widehat{x}_k^s(j)$ of x_k^s at iteration j is then given by

$$\widehat{x}_k^s(j) = \arg \max_{x \in \{0,1\}} \prod_{n=1}^N \chi_{k,j}^n(x), \quad (10)$$

that is, by the product of all messages arriving to variable node x_k^s at iteration j . The iteration ends by computing the soft information fed back from the SENSING subgraph directly to the corresponding LDGM decoder, namely,

$$\xi_{k,j}^n(x) = \Upsilon_{k,j}^n \left[\sum_{\sim x_k^n} (1 - p_n) \prod_{m \neq n} \chi_{k,j}^m(x) + p_n \prod_{m \neq n} \chi_{k,j}^m(1 - x) \right], \quad (11)$$

where as before, $\Upsilon_{k,j}^n$ represents a normalization factor for each message pair.

4. Simulation Results

To verify the performance of the proposed system, extensive Monte Carlo simulations have been performed for $N \in \{2, 4, 6\}$ sensors and a sensing error probability set, without loss of generality, to $p_n = p = 5 \cdot 10^{-3}$ for all sensors. The experiments have been divided in two different sets so as to shed light on the aforementioned statistics of the number of errors per iterations. Accordingly, the first set does not consider any outer BCH coding, and only identical LDGM codes of rate 1/3 (input symbols per coded

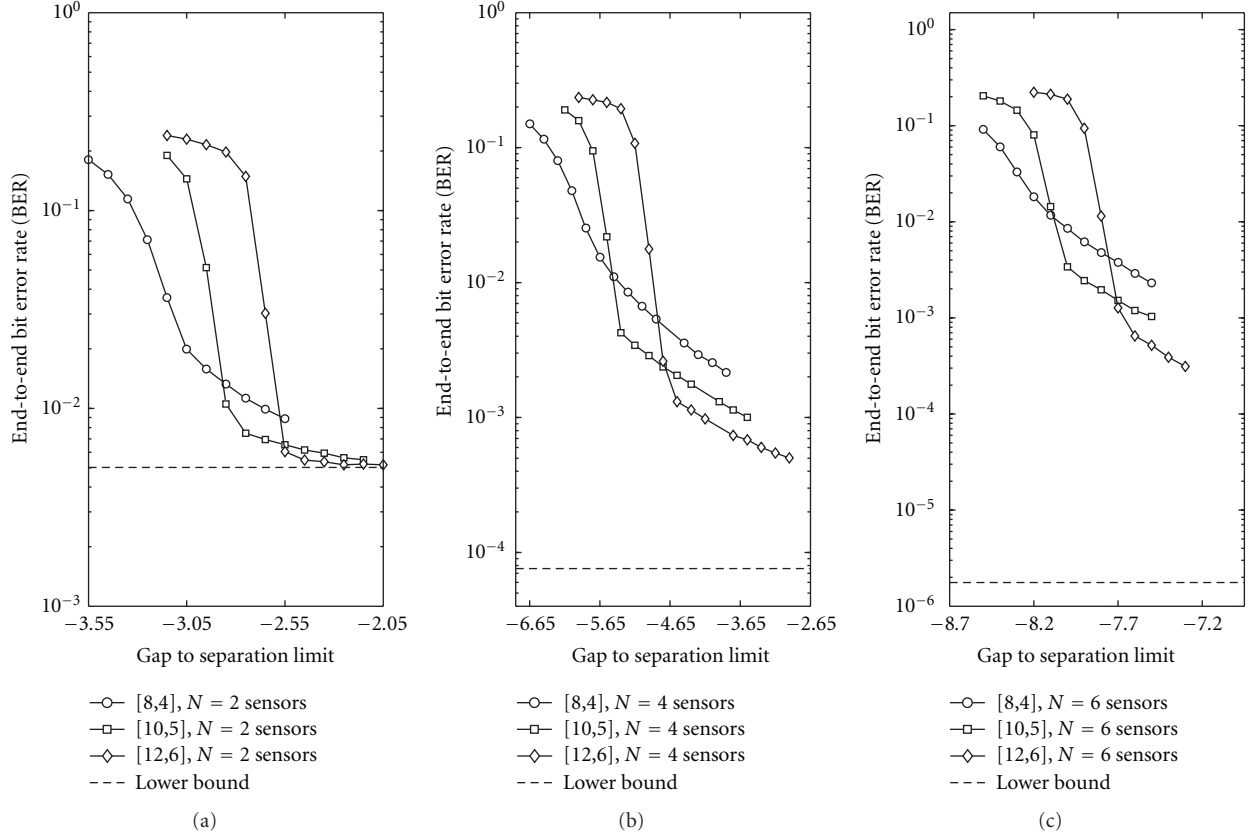


FIGURE 4: End-to-End BER versus gap to separation limit $E_b/N_0 - E_b^*/N_0$ for the Gaussian MAC with (a) $N = 2$ sensors; (b) $N = 4$ sensors; (c) $N = 6$ sensors.

symbol), variable and check degree distributions $[d_v, d_c] \in \{[8, 4], [10, 5], [12, 6]\}$, and input blocklength $K = 10000$ are utilized at every sensor. The number of iterations for the proposed iterative receiver has been set equal to $\mathcal{L} = 50$. The metric adopted for the performance evaluation is the End-to-End Bit Error Rate (BER) between x_k^s and \hat{x}_k^s , which is averaged over 2000 different information sequences per simulated point and plotted versus the E_b/N_0 ratio per sensor (energy per bit to noise power spectral density ratio). Gaussian MAC is considered in all simulations by imposing $h_l^n = 1$ for all l, n .

Before presenting the obtained simulation results, two different performance limits can be derived for each simulated case. On one hand, it can be easily shown that the aforementioned BER metric can be lower bounded by the probability of erroneously detecting x_k^s provided that all sensor symbols $\{x_k^n\}_{n=1}^N$ are perfectly recovered, which can be computed, for even N , as

$$\text{BER} \geq 0.5 \binom{N}{N/2} p^{N/2} (1-p)^{N/2} + \sum_{n=N/2+1}^N \binom{N}{n} p^n (1-p)^{N-n}, \quad (12)$$

that is, as the probability of having more than $N/2$ sensors in error. On the other hand, the minimum E_b/N_0 per sensor required for reliable transmission of all sensors can be computed by combining the Slepian-Wolf [42] Theorem for distributed compression of correlated sources and Shannon's Separation Theorem. It can be theoretically proven that this Separation Theorem does not hold for the MAC under consideration. However, this limit may serve as a theoretical reference to compare the obtained performance results. This suboptimum limit E_b^*/N_0 is computed as

$$\frac{E_b^*}{N_0} = 10 \log_{10} \left(\frac{2^{2R_c R_{\text{out}} H(S_1, \dots, S_N)} - 1}{2R_c R_{\text{out}} H(S_1, \dots, S_N)} \right) \quad (\text{dB}), \quad (13)$$

where $R_c = R_{\text{out}} d_c / (d_c + d_v)$ and the joint binary entropy of the sensors $H(S_1, \dots, S_N)$ is given by

$$H(S_1, \dots, S_N) = - \sum_{n=1}^N \binom{N}{n} \Pr\{n \text{ 0's}\} \log_2 \Pr\{n \text{ 0's}\}, \quad (14)$$

with $\Pr\{n \text{ 0's}\} = 0.5(p^n(1-p)^{N-n} + (1-p)^n p^{N-n})$ denoting the probability of having a sequence with exactly n zero symbols. In this first simulation set, no outer BCH code is used, hence $R_c = d_c / (d_c + d_v) = 1/3$.

Figure 4 summarizes the obtained results for this first set of experiments by plotting End-to-End BER versus

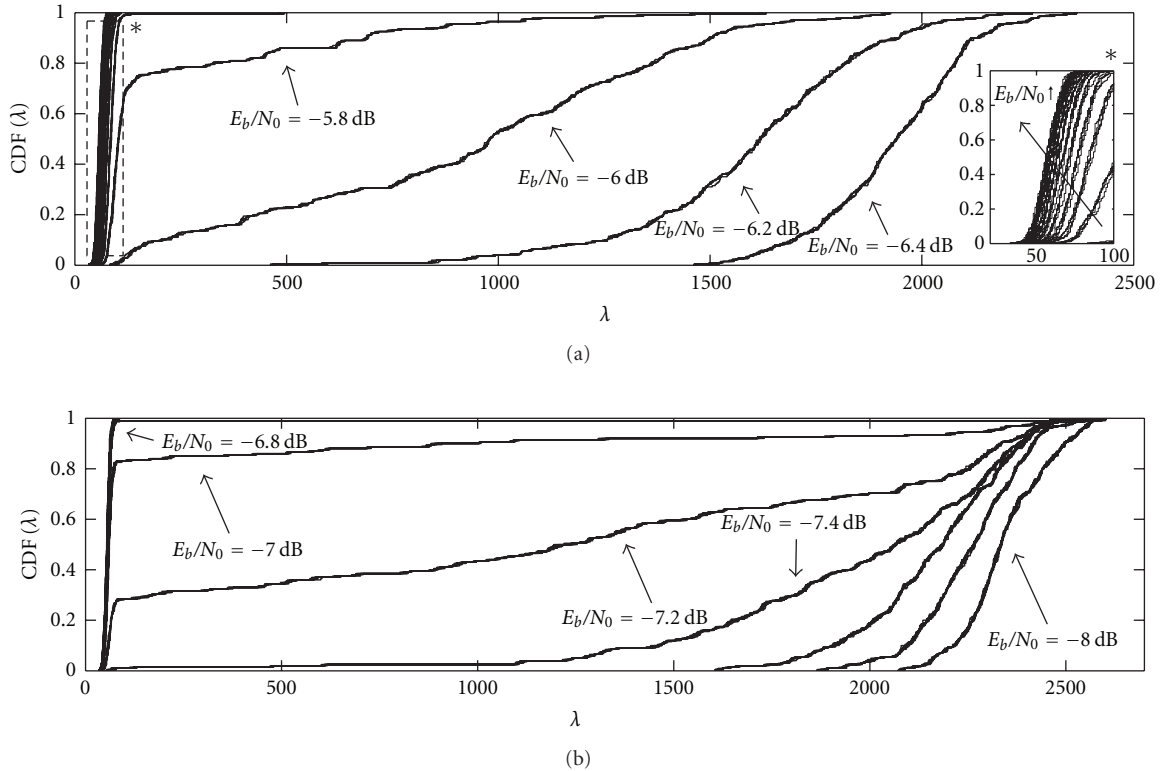


FIGURE 5: Cumulative Density Function $CDF(\lambda)$ versus number of errors per LDGM-decoded block λ for (a) $N = 4$ sensors and $[d_c d_v] = [10 5]$; (b) $N = 6$ sensors and $[d_c d_v] = [12 6]$.

the difference between the simulated E_b/N_0 and the corresponding E_b^*/N_0 limit from expression (13). Also are depicted horizontal limits corresponding to the BER lower bound from expression (12). First observe that since the aforementioned difference value is negative, the simulated E_b/N_0 is lower than E_b^*/N_0 , which verify in practice the suboptimality of the computed separation-based bound. On the other hand, notice that the set of all BER curves for $N = 2$ coincide with the lower bound in expression (12) (horizontal dashed lines), while the waterfall region of such curves degrades as $[d_v d_c]$ increases. However, for $N \in \{4, 6\}$, the error floor (due to the MAC ambiguity of the received sequence about which transmitted symbol corresponds to each sender) is higher than the lower BER bound. By increasing $[d_v d_c]$ an error floor diminishes at the cost of degrading the BER waterfall performance.

It is also important to remark that the results plotted in Figure 4 have been obtained by setting the variables controlling the switches from Figure 3(b) to $\mu_1 = \mu_2 = 1$ during the first iteration, while for the remaining $\mathcal{L} - 1$ iterations $\mu_1 = \mu_2 = 0$ (i.e., the MAC subgraph is disconnected and does not participate in the message passing procedure). The rationale behind this setup lies on the length-4 loop connecting variable nodes x_k^n , x_k^m ($m \neq n$), x_k^s and b_k for $k \in \{1, \dots, K\}$, which degrades significantly the performance of the message-passing SPA. Further simulations have been carried out to assess this degradation, which are omitted for the sake of clarity in

the present discussion. Based on this result, all simulations henceforth will utilize the same switch schedule as the one used for this first set of simulations.

To better understand the error behavior of the proposed scheme in the error floor region, it is useful to analyze the distribution of the number of errors per block at the output of the LDGM decoders. To this end, let $CDF(\lambda)$ denote the Cumulative Density Function of the number of errors per LDGM-decoded block λ at iteration \mathcal{L} , which can be empirically estimated based on the results obtained for the first set of simulations. This function $CDF(\lambda)$ is depicted for $N = 4$ and $[d_c d_v] = [10, 5]$ (Figure 5(a)) and for $N = 6$ and $[d_c d_v] = [12, 6]$ (Figure 5(b)). In this plot, such density function is depicted for every simulated E_b/N_0 point and for every compounding LDGM decoder. Observe that in all the considered E_b/N_0 range, the behavior of the CDF function results in being similar to all sensors. Furthermore, when E_b/N_0 increases (i.e., when the system operates in the error floor region), the resulting $CDF(\lambda)$ indicates that most of the decoded blocks contain a relatively small amount of errors with respect to the used blocksize $K = 10^4$. This conclusion also holds for either Figure 5(b) and the other cases addressed in the first set of simulations.

This statistical behavior of the number of errors per decoded block λ motivates the inclusion of an outer systematic BCH code whose error correction capability t is adjusted so as to correct the residual errors obtained in the error floor region. However, note that the application of

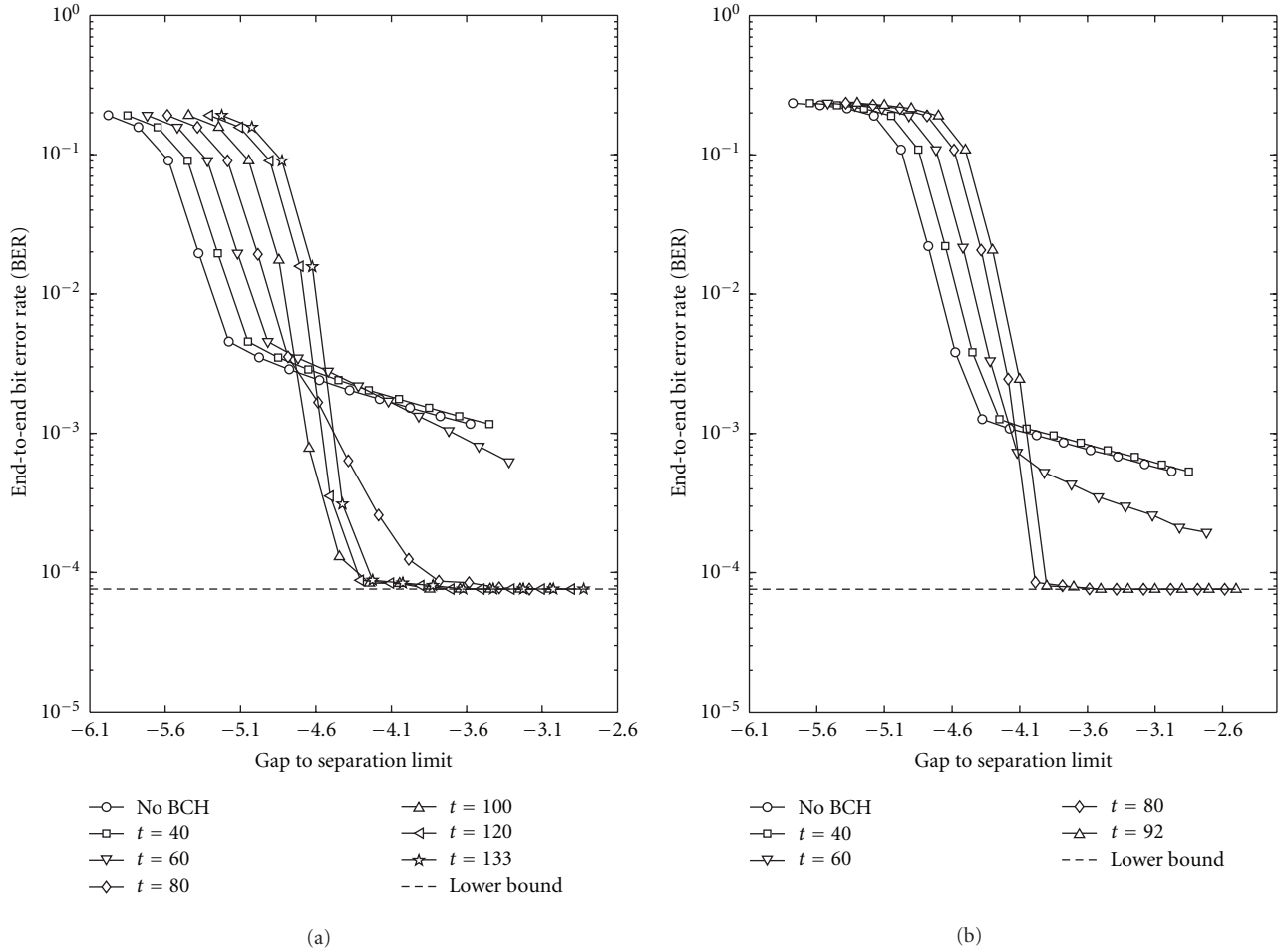


FIGURE 6: End-to-End BER versus gap to separation limit $E_b/N_0 - E_b^*/N_0$ for $N = 4$ sensors, different BCH codes and (a) $[d_c \ d_v] = [10 \ 5]$; (b) $[d_c \ d_v] = [12 \ 6]$.

an outer code involves a penalty in energy. Specifically, the E_b/N_0 ratio is increased by an amount $10 \log_{10}(1/R_{\text{out}})$ dB, where R_{out} decreases as the error capability t of the BCH code increases. Consequently, a tradeoff between t and its associated rate loss must be met. In this context, Figures 6 and 7 represent the End-to-End BER versus the gap to the separation limit $E_b/N_0 - E_b^*/N_0$ for $N = 4$ (Figures 7(a) and 7(b)), $N = 6$ (Figures 7(a) and 7(b)), and a number of BCH codes with distinct values of the error-correcting parameter t . Observe that in all cases the error floor has been suppressed by virtue of the error correcting capability of the outer BCH code, and consequently the lower bound for the BER metric in expression (12) is reached. At the same time, due to the relatively small value of t with respect to K , the energy increase incurred by concatenating an outer BCH code is less than 0.5 dB. Summarizing, the proposed iterative scheme can be regarded as an efficient and practical approach for encoded data fusion over MAC, which is shown to outperform the suboptimum separation-based limit while reaching, at the same time, the lower bound for the End-to-End BER.

5. Concluding Remarks

In this paper, we have investigated the performance of concatenated BCH-LDGM codes for iterative data fusion of distributed decisions over the Gaussian MAC. The use of LDGM codes permits to efficiently exploit the intrinsic spatial correlation between the information registered by the sensors, whereas BCH codes are selected to lower the error floor due to the MAC ambiguity about the transmitted symbols. Specifically, we have designed an iterative receiver comprising channel detection, BCH-LDGM decoding, and data fusion, which have been thoroughly detailed by means of factor graphs and the Sum-Product Algorithm. Furthermore, a specially tailored soft information flipping technique based on the output of the BCH decoding stage has also been included in the proposed iterative receiver. Extensive computer simulations results obtained for varying number of sensors, LDGM, and BCH codes have revealed that (1) our scheme outperforms significantly the suboptimum limit assuming separation between distributed source and capacity-achieving channel coding and (2) the

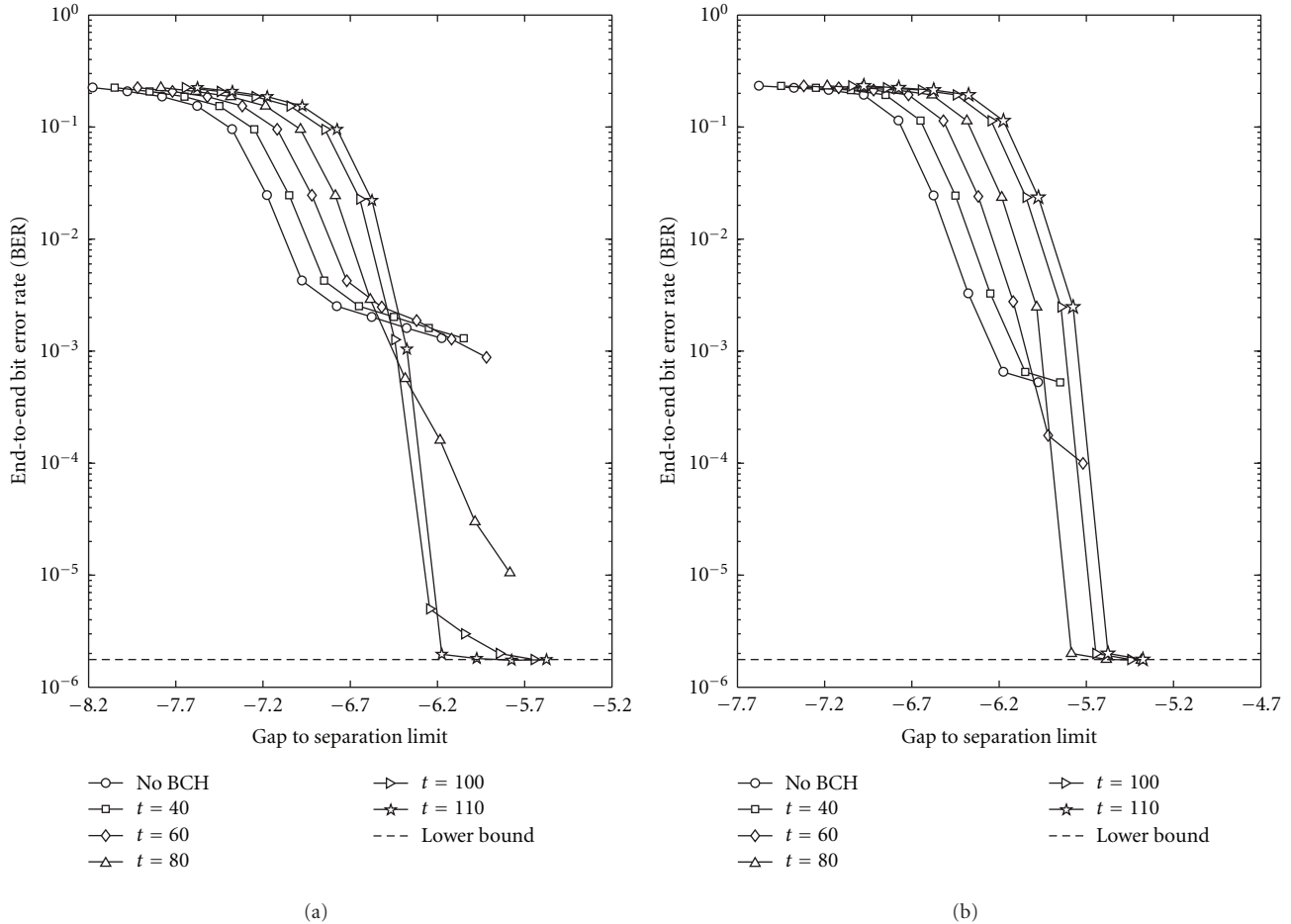


FIGURE 7: End-to-End BER versus gap to separation limit $E_b/N_0 - E_b^*/N_0$ for $N = 6$ sensors, different BCH codes and (a) $[d_c \ d_v] = [10 \ 5]$; (b) $[d_c \ d_v] = [12 \ 6]$.

obtained end-to-end error rate performance attains the theoretical lower bound assuming perfect recovery of the sensor sequences.

Acknowledgments

This work was supported in part by the Spanish Ministry of Science and Innovation through the CONSOLIDER-INGENIO (CSD200800010) and the Torres-Quevedo (PTQ-09-01-00740) funding programs and by the Basque Government through the ETORTEK programme (*Future Internet* EI08-227 project).

References

- [1] S. S. Pradhan, J. Kusuma, and K. Ramchandran, "Distributed compression in a dense microsensor network," *IEEE Signal Processing Magazine*, vol. 19, no. 2, pp. 51–60, 2002.
- [2] Z. Xiong, A. D. Liveris, and S. Cheng, "Distributed source coding for sensor networks," *IEEE Signal Processing Magazine*, vol. 21, no. 5, pp. 80–94, 2004.
- [3] Y. Yao and G. B. Giannakis, "Energy-efficient scheduling for wireless sensor networks," *IEEE Transactions on Communications*, vol. 53, no. 8, pp. 1333–1342, 2005.
- [4] M. L. Sichitiu, "Cross-layer scheduling for power efficiency in wireless sensor networks," in *Proceedings of the 23rd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '04)*, vol. 3, pp. 1740–1750, March 2004.
- [5] B. Krishnamachari, D. Estrin, and S. Wicker, "The impact of data aggregation in wireless sensor networks," in *Proceedings of the 22nd International Conference on Distributed Computing Systems*, pp. 575–578, 2002.
- [6] N. Shrivastava, C. Buragohain, D. Agrawal, and S. Suri, "Medians and beyond: new aggregation techniques for sensor networks," in *Proceedings of the 2nd International Conference on Embedded Networked Sensor Systems*, pp. 239–249, November 2004.
- [7] X. Tang and J. Xu, "Optimizing lifetime for continuous data aggregation with precision guarantees in wireless sensor networks," *IEEE/ACM Transactions on Networking*, vol. 16, no. 4, pp. 904–917, 2008.
- [8] A. Aksu and O. Ercetin, "Multi-hop cooperative transmissions in wireless sensor networks," in *Proceedings of the 2nd IEEE*

- Workshop on Wireless Mesh Networks (WiMESH '06)*, pp. 132–134, September 2006.
- [9] Y. Yuan, M. Chen, and T. Kwon, "A novel cluster-based cooperative MIMO scheme for multi-hop wireless sensor networks," *Eurasip Journal on Wireless Communications and Networking*, vol. 2006, Article ID 72493, 9 pages, 2006.
 - [10] J. Xu, X. Tang, and W. C. Lee, "EASE: an energy-efficient in-network storage scheme for object tracking in sensor networks," in *Proceedings of the 2nd Annual IEEE Communications Society Conference on Sensor and AdHoc Communications and Networks (SECON '05)*, pp. 396–405, September 2005.
 - [11] P. De Mil, B. Jooris, L. Tytgat et al., "Design and implementation of a generic energy-harvesting framework applied to the evaluation of a large-scale electronic shelf-labeling wireless sensor network," *Eurasip Journal on Wireless Communications and Networking*, vol. 2010, Article ID 343690, 12 pages, 2010.
 - [12] W. K. G. Seah, A. E. Zhi, and H. P. Tan, "Wireless sensor networks powered by ambient energy harvesting (WSN-HEAP)—survey and challenges," in *Proceedings of the 1st International Conference on Wireless Communication, Vehicular Technology, Information Theory and Aerospace and Electronic Systems Technology, Wireless (VITAE '09)*, pp. 1–5, May 2009.
 - [13] G. Lauer, N. R. Sandell Jr. et al., "Distributed detection of known signals in correlated noise," Tech. Rep. 160, Alphatech, Burlington, Mass, USA, 1982.
 - [14] L. K. Ekchian and R. R. Tenney, "Detection networks," in *Proceedings of the 21st IEEE Conference on Decision and Control*, pp. 686–691.
 - [15] Z. Chair and P. K. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 22, no. 1, pp. 98–101, 1986.
 - [16] H. Chen, P. K. Varshney, and B. Chen, "Cooperative relay for decentralized detection," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '08)*, pp. 2293–2296, April 2008.
 - [17] J. Del Ser, I. Olabarrieta, S. Gil-Lopez, and P. M. Crespo, "On the design of frequency-switching patterns for distributed data fusion over relay networks," in *Proceedings of the International ITG Workshop on Smart Antennas (WSA '10)*, pp. 275–279, February 2010.
 - [18] Y. Lin, B. Chen, and P. K. Varshney, "Decision fusion rules in multi-hop wireless sensor networks," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 41, no. 2, pp. 475–488, 2005.
 - [19] S. C. A. Thomopoulos and L. Zhang, "Distributed decision fusion in the presence of networking delays and channel errors," *Information Sciences*, vol. 66, no. 1-2, pp. 91–118, 1992.
 - [20] B. Chen, R. Jiang, T. Kasetkasem, and P. K. Varshney, "Fusion of decisions transmitted over fading channels in wireless sensor networks," in *Proceedings of the Conference Record of the Asilomar Conference on Signals, Systems and Computers*, vol. 2, pp. 1184–1188, 2002.
 - [21] R. Niu, B. Chen, and P. K. Varshney, "Decision fusion rules in wireless sensor networks using fading channel statistics," in *Proceedings of the Conference on Information Sciences and Systems*, March 2003.
 - [22] B. Chen, R. Jiang, T. Kasetkasem, and P. K. Varshney, "Channel aware decision fusion in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 52, no. 12, pp. 3454–3458, 2004.
 - [23] Y. Lin, B. Chen, and L. Tong, "Distributed detection over multiple access channels," in *2007 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '07)*, pp. 541–544, April 2007.
 - [24] W. Li and H. Dai, "Distributed detection in wireless sensor networks using a multiple access channel," *IEEE Transactions on Signal Processing*, vol. 55, no. 3, pp. 822–833, 2007.
 - [25] Y. Zhao and J. Garcia-Frias, "Turbo compression/joint source-channel coding of correlated binary sources with hidden Markov correlation," *Signal Processing*, vol. 86, no. 11, pp. 3115–3122, 2006.
 - [26] K. Kobayashi, T. Yamazato, H. Okada, and M. Katayama, "Iterative joint channel-decoding scheme using the correlation of transmitted information sequences," in *Proceedings of the International Symposium on Information Theory and Its Applications*, pp. 808–813, 2006.
 - [27] J. Garcia-Frias and Y. Zhao, "Near-Shannon/Slepian-Wolf performance for unknown correlated sources over AWGN channels," *IEEE Transactions on Communications*, vol. 53, no. 4, pp. 555–559, 2005.
 - [28] W. Zhong and J. Garcia-Frias, "LDGM codes for channel coding and joint source-channel coding of correlated sources," *Eurasip Journal on Applied Signal Processing*, vol. 2005, no. 6, pp. 942–953, 2005.
 - [29] J. Del Ser, P. M. Crespo, and O. Galdos, "Asymmetric joint source-channel coding for correlated sources with blind HMM estimation at the receiver," *Eurasip Journal on Wireless Communications and Networking*, vol. 2005, Article ID 357402, 10 pages, 2005.
 - [30] J. Garcia-Frias and J. D. Villasenor, "Joint turbo decoding and estimation of hidden Markov sources," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 9, pp. 1671–1679, 2001.
 - [31] J. Garcia-Frias, "Joint source-channel decoding of correlated sources over noisy channels," in *Proceedings of the Data Compression Conference*, pp. 283–292, March 2001.
 - [32] W. Zhong and J. Garcia-Frias, "Combining data fusion with joint source-channel coding of correlated sensors," in *Proceedings of the IEEE Information Theory Workshop (ITW '04)*, pp. 315–317, 2004.
 - [33] W. Zhong and J. Garcia-Frias, "Combining data fusion with joint source-channel coding of correlated sensors using IRA codes," in *Proceedings of the Conference on Information Sciences and Systems*, 2005.
 - [34] J. Del Ser, J. Garcia-Frias, and P. M. Crespo, "Iterative concatenated zigzag decoding and blind data fusion of correlated sensors," in *Proceedings of the International Conference on Ultra Modern Telecommunications and Workshops*, October 2009.
 - [35] J. Garcia-Frias, Y. Zhao, and W. Zhong, "Turbo-like codes for transmission of correlated sources over noisy channels," *IEEE Signal Processing Magazine*, vol. 24, no. 5, pp. 58–66, 2007.
 - [36] Y. Zhao, W. Zhong, and J. Garcia-Frias, "Transmission of correlated senders over a Rayleigh fading multiple access channel," *Signal Processing*, vol. 86, no. 11, pp. 3150–3159, 2006.
 - [37] W. Zhong, H. Chai, and J. Garcia-Frias, "LDGM codes for transmission of correlated senders over MAC," in *Proceedings of the Allerton Conference on Communication, Control, and Computing*, 2005.
 - [38] W. Zhong and J. Garcia-Fria, "Joint source-channel coding of correlated senders over multiple access channels," in *Proceedings of the Allerton Conference on Communication, Control, and Computing*, 2004.
 - [39] A. Hocquenghem, "Codes correcteurs d'erreurs," *Chiffres*, vol. 2, pp. 147–156, 1959.
 - [40] R. C. Bose and D. K. Ray-Chaudhuri, "On a class of error correcting binary group codes," *Information and Control*, vol. 3, pp. 68–79, 1960.

- [41] F. R. Kschischang, B. J. Frey, and H. A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 498–519, 2001.
- [42] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Transactions on Information Theory*, vol. 19, no. 4, pp. 471–480, 1973.