Estimation of Observation Error Probability in Wireless Sensor Networks

Xin He, Xiaobo Zhou, *Student Member, IEEE,* Khoirul Anwar, *Member, IEEE,* and Tad Matsumoto, *Fellow, IEEE*

***Abstract*—In this letter, we first of all propose for a parallel wireless sensor network (WSN) a decoding technique that well exploits the correlation knowledge of the sensing data to be transmitted from each sensor to the fusion center (FC). This letter then derives an algorithm to estimate the observation error probabilities, representing the correlation, of the links between the sensing object and sensors. The convergence of the algorithm is also evaluated. Furthermore, the comparison of bit-error-rate (BER) performance between two cases, one uses estimated observation error probabilities, the other assumes the full knowledge of the observation error probabilities, is made. The simulation results show that the difference is only around $0.3 - 0.5$ dB in per-link signal-to-noise power ratio (SNR), depending on the number of sensors.**

***Index Terms*—Error probability, estimation, CEO problem, wireless sensor networks, LLR updating function.**

I. Introduction

WIRELESS sensor networks (WSNs) composed of a large number of sensors, deployed in a geographic area to perform distributed tasks, have been recognized as an important technology, since WSNs have significant impacts on the society. In WSNs, each sensor node is required to work under very low power consumption restrictions when it performs specified tasks. Therefore, the sensor nodes have to transmit their data without requiring high transmission power, and the WSN itself has to be highly energy efficient. A technique which is effective in reducing the power consumption is to design the system so that the fusion center (FC) can well exploit the correlation knowledge among the observed data, as supported by the Slepian-Wolf theorem [1]. In fact, the Slepian-Wolf theorem can be used to compress the data or to reduce the transmission power, assuming the source-channel separation [2]. Therefore, in this letter, no specific source encoding technique is assumed.

Distributed source coding schemes for sensor networks based on the Slepian-Wolf theorem are investigated in a tutorial article [3]. In practice, most of WSNs aim to observe the same sensing object, e.g., the sensors monitor the same physical phenomenon. In network information theory, the problem of estimating the correct data emitted from the target over noisy observation link can be modeled by the chief executive officer (CEO) problem [4], as defined in [5]. The quadratic Gaussian CEO problem is studied in [6], where it is assumed that the observed data samples are correlated Gaussian random variables. Rate-distortion function for a parallel Gaussian CEO model is investigated in [7], where also successive coding/decoding strategy is briefly introduced.

Manuscript received February 15, 2013. The associate editor coordinating the review of this letter and approving it for publication was I.-R. Chen.

This research is supported in part by a Japan Society for the Promotion of Science (JSPS) Grant under Scientific Research KAKENHI KIBAN (C) No. 22560367 and (B) No. 23360170, and in part, the SWOCNET Project funded by Academy of Finland.

X. He, X. Zhou, and K. Anwar are with the School of Information Science, Japan Advanced Institute of Science and Technology (JAIST), 1-1 Asahidai, Nomi, Ishikawa, Japan 923-1292 (e-mail: {hexin, xiaobo, anwar-k}@jaist.ac.jp).

T. Matsumoto is with the School of Information Science, JAIST, 1-1 Asahidai, Nomi, Ishikawa, Japan 923-1292 (e-mail: matumoto@jaist.ac.jp). He is also with the Centre for Wireless Communication, University of Oulu, FI-90014 Finland (e-mail: tadashi.matsumoto@ee.oulu.fi).

Digital Object Identifier 10.1109/LCOMM.2013.042313.130361

In [8], a source type identification technique at the FC using dumb sensors is investigated from the viewpoint of hypothesis testing. Instead, this letter focus on the coding/decoding techniques of analog-to-digital (A/D)-converted binary sequence; the k-th sensor's observation results are A/D-converted with m-bit resolution, interleaved by interleaver Π, and then transmitted to the FC by using binary-phased-shift-keying (BPSK). It should be emphasized that the same Π is commonly used by the M sensors, and the size is equivalent to the $m \times K$, where K is number of the samples. It plays a crucial role in making the length mK bit sequence random so that the observation error can well be represented by the random bit-flipping model after the interleaver Π. Both static Additive White Gaussian Noise (AWGN) and Rayleigh fading channels are assumed for the channels from sensors to the FC.

A technique that utilizes the knowledge of the observation error probabilities at the FC is proposed in [9]; each sensor uses parallel concatenated convolutional codes (PCCC) for the data transmission to the FC, and the extrinsic log-likelihood ratio (LLR) from the each decoder is combined; and after they are combined, the extrinsic information of the combiner is fed back to the each PCCC decoder, weighted by the observation error probabilities p_k, $k = 1, \cdots, M$. In [10], we propose a technique to more efficiently utilize the knowledge about p_k than in [9], where the combined output extrinsic LLR is updated by using so-called f_c-function [10]; it is assumed that the p_k values are known to the FC.

The major purpose of this letter is to propose a nonnegative constrained iterative algorithm to estimate the observation error probabilities for a WSN having an arbitrary number of sensors. The estimated observation error probabilities are used in the LLR exchange between the decoders, referred to as global iteration (*GI*), utilized in this letter to exploit the knowledge of p_k and thereby to improve the system performance. Furthermore, the decoding-complexity is reduced to a linear order by introducing local iteration (*LI*) and *GI*.

The rest of this letter is organized as follows. In Section II, the system model of the parallel WSN is described. Section III describes the decoding strategy which well utilizes the correlation knowledge. The proposed error probability estimation algorithm is detailed in Section IV. The convergence property of the proposed observation error estimation algorithm and

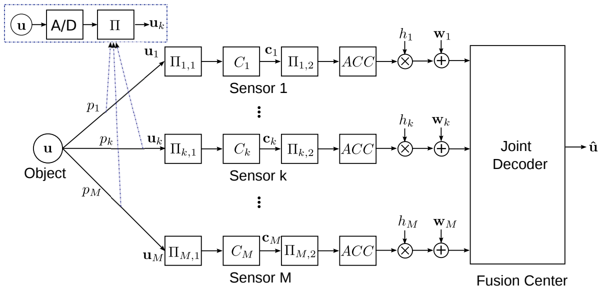

Fig. 1. Structure of the proposed system model.

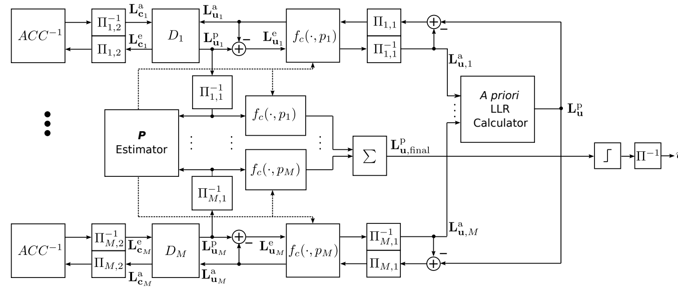

Fig. 2. Proposed decoding strategy for a parallel sensor network.

bit-error-rate (BER) performances are evaluated in Section V. We conclude this letter in Section VI with some concluding statements.

II. System Model

Fig. 1 depicts the model of the parallel WSN system, investigated in this letter. A set of sensors $\mathbf{S}$ produces the error-corrupted versions of the binary sequence $\mathbf{u}_k$, $k = 1, \cdots, M$, obtained after the interleaver Π following the A/D convertor, corresponding to the observed samples generated by the sensing object. The observations made by the sensors are correlated, which, as noted above, are modeled by bit-flipping models with flipping probabilities p_k for the sensor k, $k = 1, \cdots, M$. Let $\mathbf{P} = [p_1, p_2, \cdots, p_M]^{\mathrm{T}}$ denote the vector sorting the observation error probabilities. The sensor k interleaves the observed bit sequence $\mathbf{u}_k$ first, using interleaver $\Pi_{k,1}$, and then encode the interleaved bit sequence $\Pi_{k,1}(\mathbf{u}_k)$ with a channel encoder C_k.

The encoded bit sequence $\mathbf{c}_k$ is further interleaved by the interleaver $\Pi_{k,2}$ and doped-accumulated by ACC [11] with doping ratio P_d for $k = 1, \cdots, M$. It should be noted that the lengths of $\Pi_{k,1}$ should not necessarily be the same as that of Π. Finally, the doped-accumulated bit sequence is modulated by BPSK and transmitted to the FC over independent static AWGN channels or Rayleigh fading channels. As shown in Fig. 1, the signal received from the k-th sensor can be written as:

$$\mathbf{y}_k = h_k \cdot \mathbf{s}_k + \mathbf{w}_k, \tag{1}$$

where, h_k represents the channel coefficient. The BPSK modulated symbol sequence at the sensor k is denoted by $\mathbf{s}_k$. $\mathbf{w}_k$ is a zero mean Gaussian noise sequence with variance σ^2 per dimension.

III. Decoding Algorithm

A block diagram of the proposed decoding algorithm is shown in Fig. 2. It includes *LI*, which performs the extrinsic LLR exchange between the decoder ACC^{-1} of the doped-accumulator and the channel code decoders D_k, and the *GI*, which performs the extrinsic LLR exchange among decoders D_k. The aim of performing *GI* is to utilize the correlation knowledge among the sensors through the LLR updating function f_c [2], as shown in Fig. 2. As we can see in the BER performance curves presented in Section V, the effect of performing *GI* is significant, hence to achieve such large gain through *GI*, we need to estimate the observation error probabilities.

After the *LI*, *a posteriori* LLR $\mathbf{L}^{\mathrm{p}}_{\mathbf{u}_k}$ of systematic bits (information part corrupted by the observation error) output from D_k are fed into the $\mathbf{P}$ estimator to obtain the observation error probabilities p_k which can be utilized in $f_c(\cdot, p_k)$. The algorithm used in the $\mathbf{P}$ estimator is detailed in the next section. Then, the extrinsic LLR $\mathbf{L}^{\mathrm{e}}_{\mathbf{u},k}$ are input into *a priori* LLR calculator after performing $f_c(\cdot, p_k)$ and de-interleaved. The *a priori* LLR calculator obtains $\mathbf{L}^{\mathrm{p}}_{\mathbf{u}}$ as:

$$\mathbf{L}^{\mathrm{p}}_{\mathbf{u}} = \sum_{k=1}^{M} \mathbf{L}^{\mathrm{a}}_{\mathbf{u},k} = \sum_{k=1}^{M} f_c[\Pi^{-1}_{k,1}(\mathbf{L}^{\mathrm{e}}_{\mathbf{u}_k}), p_k]. \tag{2}$$

As indicated in Fig. 2, the extrinsic LLR $f_c[\mathbf{L}^{\mathrm{e}}_{\mathbf{u}_k}, p_k]$ is equivalent to *a priori* LLR $\mathbf{L}^{\mathrm{a}}_{\mathbf{u},k}$ of the *a priori* LLR calculator. Hence, $\mathbf{L}^{\mathrm{a}}_{\mathbf{u},k}$ has to be subtracted from $\mathbf{L}^{\mathrm{p}}_{\mathbf{u}}$. The interleaved version of $\mathbf{L}^{\mathrm{p}}_{\mathbf{u}} - \mathbf{L}^{\mathrm{a}}_{\mathbf{u},k}$ is input to $f_c(\cdot, p_k)$ as $f_c[\Pi_{k,1}(\mathbf{L}^{\mathrm{p}}_{\mathbf{u}} - \mathbf{L}^{\mathrm{a}}_{\mathbf{u},k}), p_k]$, and then its output is fed back to D_k as the *a priori* LLR, $\mathbf{L}^{\mathrm{a}}_{\mathbf{u}_k} = f_c[\Pi_{k,1}(\mathbf{L}^{\mathrm{p}}_{\mathbf{u}} - \mathbf{L}^{\mathrm{a}}_{\mathbf{u},k}), p_k]$.

The *LI* and *GI* are performed until no more relevant gain can be achieved in *a posteriori* LLR $\mathbf{L}^{\mathrm{p}}_{\mathbf{u},\mathrm{final}}$. Then hard decisions are made based on $\mathbf{L}^{\mathrm{p}}_{\mathbf{u},\mathrm{final}}$ given by:

$$\mathbf{L}^{\mathrm{p}}_{\mathbf{u},\mathrm{final}} = \sum_{k=1}^{M} f_c[\Pi^{-1}_{k,1}(\mathbf{L}^{\mathrm{p}}_{\mathbf{u}_k}), p_k]. \tag{3}$$

It should be noticed that the proposed decoding technique is equivalent to performing the LLR updating by f_c-function between arbitrary M pairs of the sensors, however, the computational complexity for decoding is reduced from a combinatorial order $\binom{M}{2}$ to a linear order M.

IV. Error Probability Estimation Algorithm

Because of the fact that all the observations made by the sensors are correlated, the following pair-wise equations[1] hold:

$$\hat{p}_i + \hat{p}_j - 2 \cdot \hat{p}_i \cdot \hat{p}_j = \hat{q}_{ij}, \tag{4}$$

where $i = 1 \cdots M$. $j = i + 1$ if $i = 1 \cdots M - 1$ and $j = 1$ if $i = M$. By following [2],

$$\hat{q}_{ij} = \frac{1}{N} \sum_{1}^{N} \frac{\exp(\mathbf{L}^{\mathrm{p}}_{\mathbf{u}_i}) + \exp(\mathbf{L}^{\mathrm{p}}_{\mathbf{u}_j})}{[1 + \exp(\mathbf{L}^{\mathrm{p}}_{\mathbf{u}_i})] \cdot [1 + \exp(\mathbf{L}^{\mathrm{p}}_{\mathbf{u}_j})]}. \tag{5}$$

N represents the number of the *a posteriori* LLR pairs from the two decoders with their absolute values larger than a given

[1] q_{ij} can be understood as the bit error probability of the $j(i)$-th sensor's link, assuming that the $i(j)$-th sensor's link is error free. Furthermore, j should not necessarily be $i + 1$. In this case, according to the selected pairs, the form the matrix $\mathbf{J}$ changes.

threshold T. Since the reliability of $\hat{q}_{ij}$ is influenced by N, it is very important to choose an appropriate T value. However, how to determine the optimal T is out of the scope of this letter.

We can reformulate (4) by introducing the identity matrix $\mathbf{I}$ of size M and a matrix $\mathbf{J}$ defined by (7), into the following form:

$$[(\mathbf{I}+\mathbf{J}) - 2 \cdot \mathrm{diag}(\hat{\mathbf{P}}) \cdot \mathbf{J}] \cdot \hat{\mathbf{P}} = \hat{\mathbf{q}}, \tag{6}$$

where, $\hat{\mathbf{P}} = [\hat{p}_1, \hat{p}_2, \cdots, \hat{p}_M]^T$, and $\hat{\mathbf{q}} = [\hat{q}_{12}, \hat{q}_{23}, \cdots, \hat{q}_{M1}]^T$. The $\mathrm{diag}(\cdot)$ is the operator that forms a diagonal matrix from its argument vector, and $\mathbf{J}$ is denoted as follows:

$$\mathbf{J} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots 0 \\ 0 & 0 & 1 & 0 & \cdots 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & 0 & \cdots 0 \end{bmatrix}. \tag{7}$$

Now, our objective is to find a nonnegative vector $\hat{\mathbf{P}}$ that minimizes $\|\mathbf{A}\hat{\mathbf{P}} - \hat{\mathbf{q}}\|^2$, which is formulated as follows:

$$\begin{aligned} \min \quad & \|\mathbf{A}\hat{\mathbf{P}} - \hat{\mathbf{q}}\|^2 \\ \text{s.t} \quad & \hat{\mathbf{P}} \succeq \mathbf{0}, \end{aligned} \tag{8}$$

where, $\mathbf{A} = [(\mathbf{I}+\mathbf{J}) - 2 \cdot \mathrm{diag}(\hat{\mathbf{P}}) \cdot \mathbf{J}]$.

To solve (8), this letter proposes an iterative algorithm summarized in Algorithm 1. In this algorithm, we use the standard Nonnegative Least Squares (*lsqnonneg*) described in [12], proposed by Lawson and Hanson.

Algorithm 1: $\mathbf{P}$ Estimator

Input: $\hat{\mathbf{q}}$, ϵ, Pre-defined maximum iterations IT_m
Output: $\hat{\mathbf{P}} \succeq \mathbf{0}$ such that $\|\mathbf{A}\hat{\mathbf{P}} - \hat{\mathbf{q}}\|^2$ is minimized.
Initialization: $\hat{\mathbf{P}}^{(0)} = \mathbf{0}$, Calculate $\mathbf{A}$ and $\Delta(0) = \|\mathbf{A}\hat{\mathbf{P}}^{(0)} - \hat{\mathbf{q}}\|^2$
for $l = 1$ **to** IT_m **do**
  Calculate $\hat{\mathbf{P}}^{(l)}$ by using *lsqnonneg* algorithm;
  Update $\mathbf{A} = [(\mathbf{I}+\mathbf{J}) - 2 \cdot \mathrm{diag}(\hat{\mathbf{P}}^{(l)}) \cdot \mathbf{J}]$;
  $\Delta(l) = \|\mathbf{A}\hat{\mathbf{P}}^{(l)} - \hat{\mathbf{q}}\|^2$;
  if $\Delta(l) \geq \Delta(l-1)$ **then**
    Exit For;
  end
  if $\|\hat{\mathbf{P}}^{(l)} - \hat{\mathbf{P}}^{(l-1)}\|^2 \leq \epsilon$ **then**
    Exit For;
  end
end
$\hat{\mathbf{P}} = \hat{\mathbf{P}}^{(l)}$;

V. Performance Evaluation

A. Convergence Property

Fig. 3 shows the mean square estimation error of the observation error probability vector $\mathbf{P}$ versus the *GI* times where one *LI* was followed by one *GI*. The code parameters are the same as in the case of Fig. 4 described in the next subsection. The results are plotted for sensor number 12 with per-link SNR and LLR threshold T as parameters.

As shown in Fig. 3, the mean square estimation error $|\hat{\mathbf{P}} - \mathbf{P}|^2$ decreases as iteration times increased, indicating that the more *GIs* performed, the more accurate the estimate of $\mathbf{P}$. Furthermore, the rate of convergence depends on the values of per-link SNR and T, because the values of per-link SNR and T affect the reliability of $\hat{q}_{ij}$ given by (5).

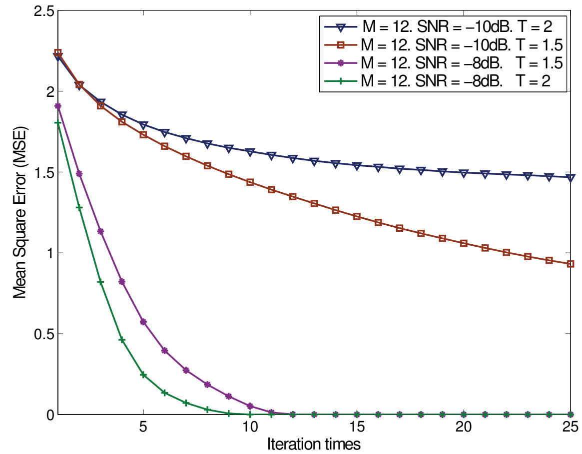

Fig. 3. Mean square errors of estimation versus decoding iteration times.

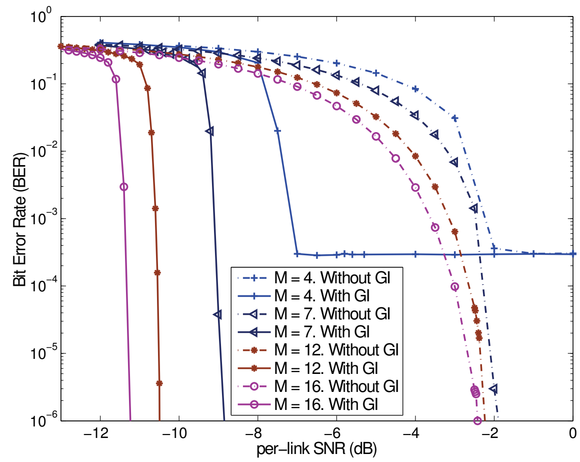

Fig. 4. BER performances with and without *GI* in AWGN channels.

B. BER Comparison

Fig. 4 illustrates the difference on BER performances between the case of involving *GIs* and that of not involving *GI*. In this simulation, a half rate memory-1 nonrecursive systematic convolutional code with the polynomial $G = [3, 2]_8$ was used for C_k. The doping ratio P_d was set to 1. The observation error probabilities p_k were all set to 0.01, and block length was 10^3 bits. We performed 25 *GIs* for $\mathbf{P}$ estimation. The channels between sensors and the FC are independent AWGN channels, where the channel coefficient $h_k = 1$. Fig. 5 shows the BER performances for the cases where $\hat{\mathbf{P}}$ and $\mathbf{P}$ are used at the FC, with the numbers M of sensors as a parameter. The other transmission parameters are the same as for Fig. 4. In $\mathbf{P}$ estimator, T and ϵ were set to 2 and 10^{-6}, respectively, and the maximum iteration times, IT_m, were set at 20. The channels are assumed to be the same as in Fig. 4.

The BER performances in the case sensors-FC links suffer from block Rayleigh fading are shown in Fig. 6. We also evaluate the gain with *GI* over without *GI* in Rayleigh fading

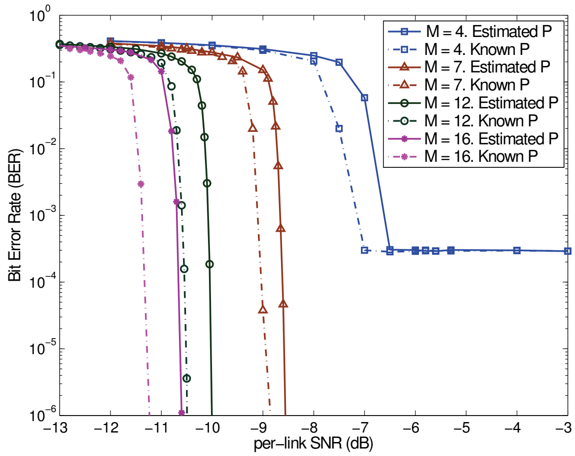

Fig. 5. BER performance comparison for different number of sensors.

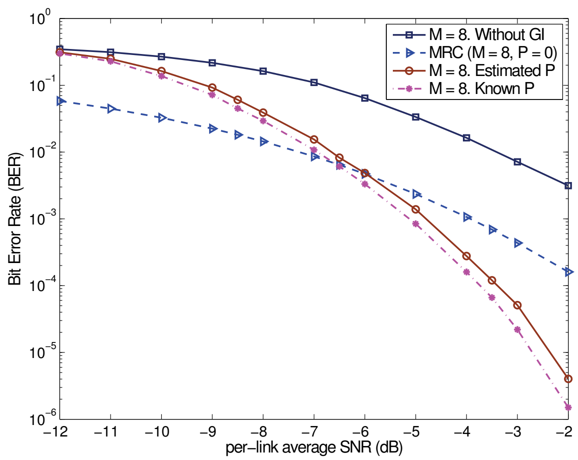

Fig. 6. BER performance in block Rayleigh fading channels.

channel. We can still achieve around 4 dB gain when $M = 8$ at 10^{-4} BER range. Furthermore, it is found that the BER curves obtained by using actual $\mathbf{P}$ and its estimate $\hat{\mathbf{P}}$ are very close (around 0.2 dB loss) with each other.

It is found from the figure that the BER performance can be improved by increasing the number of sensors M. With $M = 4$, the error floor can not be reduced to less than 10^{-4} by increasing per-link SNR, however it can be reduced to less than 10^{-6} with $M \geq 7$. Nevertheless, we believe that it is impossible to totally eliminate the error floor, even though it may happen at a very small BER region. The reason is because we can not completely eliminate the distortion due to the observation error, which is common to the CEO problems. Compared with the case where $\mathbf{P}$ is known, only $0.3-0.5$ dB loss in per-link SNR is observed when using estimated $\hat{\mathbf{P}}$, where the loss depends on M. Furthermore, our proposed technique has around $1-3$ dB improvement in per-link SNR compared with the scheme proposed in [9].

VI. Conclusion

In this letter, we have investigated for a parallel wireless sensor network the transmission techniques of data gathered by multiple sensors to the fusion center. We first proposed a new algorithm that, instead of exploiting the correlation knowledge over all the possible combinations of the sensor pairs, combines the local and the global iterations to avoid the necessity of heavy computational complexity. We also proposed a nonnegative constrained iterative algorithm to estimate the observation error probabilities. It has been shown through simulations that the algorithm converges only after several iterations. In addition, the results of BER performance show that the proposed technique using $\hat{\mathbf{P}}$ can achieve only roughly $0.3-0.5$ dB loss in per-link SNR compared to the case actual $\mathbf{P}$ is used in static AWGN channel, and 0.2 dB loss in Rayleigh fading channel.

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] J. Garcia-Frias and Y. Zhao, "Near-Shannon/Slepian-Wolf performance for unknown correlated sources over AWGN channels," *IEEE Trans. Commun.*, vol. 53, no. 4, pp. 555–559, Apr. 2005.

[3] Z. Xiong, A. D. Liveris, and S. Cheng, "Distributed source coding for sensor networks," *IEEE Signal Process. Mag.*, vol. 21, no. 5, pp. 80–94, Sept. 2004.

[4] A. Gamal and Y. Kim, *Network Information Theory*. Cambridge University Press, 2011.

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

[6] Y. Oohama, "The rate-distortion function for the quadratic Gaussian CEO problem," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 1057–1070, May 1998.

[7] H. Behroozi and M. Soleymani, "Performance of the successive coding strategy in the CEO problem," in *Proc. 2005 IEEE Global Telecomm. Conf.*, vol. 3, pp. 1347–1352.

[8] K. Liu and A. Sayeed, "Type-based decentralized detection in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 1899–1910, May 2007.

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

[10] X. Zhou, X. He, K. Anwar, and T. Matsumoto, "GREAT-CEO: larGe scale distRibuted dEcision mAking Technique for wireless Chief Executive Officer problems," *IEICE Trans. Commun.*, vol. E95-B, no. 12, pp. 3654–3662, Dec. 2012.

[11] K. Anwar and T. Matsumoto, "Accumulator-assisted distributed turbo codes for relay system exploiting source-relay correlations," *IEEE Commun. Lett.*, vol. 16, no. 7, pp. 1114–1117, 2012.

[12] D. Chen and R. J. Plemmons, "Nonnegativity constraints in numerical analysis," in *2007 Symposium on the Birth of Numerical Analysis*.