7-1-2010

Distributed Coding of Sources with Unknown Correlation Parameter

Ali Abedi
University of Maine - Main, abedi@eece.maine.edui

Abolfazl Razi
University of Maine - Main, arazi@eece.main.edu

Distributed Coding of Sources with Unknown Correlation Parameter

A. Razi, and A. Abedi
Electrical and Computer Engineering Department, University of Maine, Orono, ME, USA
Email: {arazi,abedi}@eece.maine.edu

Abstract— *Binary source estimation using a cluster of Wireless Sensor Network (WSN) is considered. The correlation between binary data source and received data at each sensor is modeled as an independent Binary Symmetric (BSC) Channel. A novel distributed joint source channel coding (DJSCC) scheme is proposed to estimate source data using the intrinsic correlation in received data from sensors. Encoder is based on distributed structure of turbo codes and decoder is equipped with a self correlation estimation block to track data correlation adaptively.*

Low complexity encoding, fast decoding convergence, and self correlation parameter estimation capability makes the proposed system an appropriate choice for distributed source estimation in wireless sensor networks with unknown, unpredictable or even time-varying correlation parameter.

Keywords: Wireless sensor networks, Joint source channel codes, Distrbuted coding, turbo codes

1. Introduction

In majority of wireless sensor networks, a sheer numbers of sensors collect data in an area and transmit it to a central base station. Dividing sensors into clusters with correlated data and employing distributed coding optimizes the overall system data flow efficiency. In this approach a group of tiny sensors inside a cluster encode the data in a distributed manner without communicating with each other and collectively transmit it to a sink node, where data is processed or relayed to a central base station using various relay mechanism such as Amplify and Forward (AAF) and Decode and Forward (DAF) [1].

The idea behind that is a theorem introduced by Slepian and Wolf in 1973 [2]. They proved that it is possible to design source codes for a group of transmitters such that they can compress and transmit data at rate below joint entropy $H(X_1, X_2, ..., X_M)$ which is much less than independent data compression $H(X_1) + ... + H(X_M)$. Optimum coding rate is achieved without need for inter transmitter communication. Wyner and Ziv later extended this result to continuous valued Gaussian distributed sources and defined rate distortion function for lossy coding [3].

The first realization of distributed source coding algorithm based on Slepian-Wolf theorem is introduced by Ramchandran, et. al. employing syndrom concept that is called DISCUSS [4]. More powerful source codes are developed based on known coding schemes such as Low Density Parity Check Codes (LDPC) [5], Turbo Codes [6], [7], Irregular Repeat Accumulate (IRA) [8] , and Low Density Generator Matrix Codes (LDGM) [9]. Recent research works combine source and channel coding to propose distributed joint source-channel coding to improve system overall performane and reduce system complexity [10].

One special case is a scenario that a number of sensors around a common data source, transmit noisy version of data to a single receiver which is an example for Chief Executive Officer (CEO) problem [11]. The receiver aims to estimate the source data using data fusion and joint decoding of correlated sources [12]. In this paper, we address this scenario with a new method based on distributed turbo encoder structure at sensor side and a modified turbo decoder at the receiver side.

In the proposed system, the intrinsic correlation between sensors is used as redundancy information to implement joint source channel coding with low complexity structure. Despite the low complexity structure of encoders, since they form a distributed turbo encoder, the proposed system performs very close to well known powerful channel codes in terms of BER performance.

Moreover, a new method is introduced to estimate the correlation among sensor's data at at the receiver to make the proposed system applicable in the case of unknown and time varying correlation parameters. Scalability is another benefit of the proposed system model, since one may add as many as required sensors to each cluster employing multi-branch turbo codes.

The rest of this paper is organized as follows. In section 2, CEO problem of a single WSN cluster with BSC based correlation model is investigated. In section 3, a new distributed coding scheme based on turbo encoder is proposed. Section 4 describes structure of employed decoder equipped with a self correlation estimation block. Section 5 includes simulation results followed by conclusion in section 6.

2. Correlation modeling

A cluster of sensors in proximity of a source is considered. The scenario is depicted in fig. 1, where binary source data $S = \{s^1, s^2, ..., s^N\}$ is an independent identically distributed

$(i.i.d)$ Bernoulli sequence with length N and $P(s^k = 0) = P(s^k = 1) = 1/2$. Each sensor measures a noisy version of data denoted by X_i. The channel error between source data, S and data sensed by i^{th} senor is modeled as a BSC channel with crossover probability of P_i, therefore the k^{th} bit in X_i is

$$x_i^k = s^k \oplus e_i^k \tag{1}$$

where e_i^k is the error of bit k in the i^{th} sensor with probability of $P(e_i^k = 1) = P_i,\ P(e_i^k = 0) = 1 - P_i$. Hereafter, without loss of generality it is assumed that $0 \leq P_1 \leq P_2 \leq ... \leq P_M \ll 1/2$, where P_i is the correlation parameter between source data and observation of the i^{th} sensor.

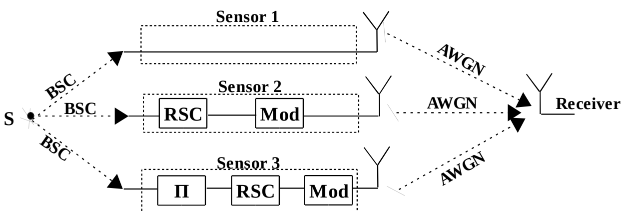

Fig. 1: A cluster of the three sensors observe a single binary data source, and form a distributed turbo encoder structure

The channels are assumed to be memoryless, hence the sequence X_i is also an $i.i.d$ sequence, i.e.

$$P(X_1, X_2, ..., X_M|S) = \prod_{k=1}^{N} P(x_1^k, x_2^k, ..., x_M^k|s^k) \tag{2}$$

Since the BSC channels between source and sensors are assumed to be independent, the correlation parameter between each set of two sources can be modeled as follows:

$$\begin{aligned}
P(x_i^k = 0|x_j^k = 0) &= \frac{P(x_i^k = 0, x_j^k = 0)}{P(x_j^k = 0)} \\
&= \frac{\sum_{m=0}^{1} P(x_i^k = 0, x_j^k = 0|s^k = m)P(s^k = m)}{\sum_{n=0}^{1} P(x_j^k = 0|s^k = n)P(s^k = n)} \\
&= \frac{\sum_{m=0}^{1} P(x_i^k = 0|s^k = m)P(x_j^k = 0|s^k = m)P(s^k = m)}{\sum_{n=0}^{1} P(x_j^k = 0|s^k = n)P(s^k = n)} \\
&= \frac{(1/2)[(1 - P_i)(1 - P_j) + P_iP_j]}{1/2} = 1 + 2P_iP_j - P_i - P_j
\end{aligned} \tag{3}$$

Similarly it can be shown that:

$$\begin{aligned}
P(x_i^k = 1|x_j^k = 1) &= 1 + 2P_iP_j - P_i - P_j \\
P(x_i^k = 1|x_j^k = 0) &= P_i + P_j - 2P_iP_j \\
P(x_i^k = 0|x_j^k = 1) &= P_i + P_j - 2P_iP_j
\end{aligned} \tag{4}$$

for small values of P_i and P_j, the term P_iP_j is negligible and equation set (4) can be approximated by

$$\begin{aligned}
P(x_i^k = 0|x_j^k = 0) = P(x_i^k = 1|x_j^k = 1) &= 1 - P_i - P_j \\
P(x_i^k = 1|x_j^k = 0) = P(x_i^k = 0|x_j^k = 1) &= P_i + P_j
\end{aligned} \tag{5}$$

This means that, sensors are pairwise correlated and correlation between the i^{th} and the j^{th} sensors can be modeled as a BSC channel with parameter $P_{ij} = P_i + P_j - 2P_iP_j$. This result can also be demonstrated with another intuitive approach. A BSC channel with equiprobable inputs is bidirectional and we have $P(X = x|Y = y) = P(Y = y|X = x)$, therefore the channel between two sensors can be modeled as a cascade of two consecutive BSC channels with parameters P_i and P_j that results in a BSC channel with parameter $(1-P_i)P_j + P_i(1-P_j) = P_i + P_j - 2P_iP_j \approx P_i + P_j$. These results are used in decoder implementation of the proposed method to estimate the correlation parameter at the receiver.

Data at the i^{th} sensor is coded to form $C_i = c_i^1, c_i^2, ..., c_i^{N/R},\ i = 0, 1, ...M$, where R is coding rate at each sensor. The coded data is modulated with Binary Phase Shift Keying Scheme $(BPSK)$ to form output symbol frame $Y_i = y_i^1, y_i^2, ..., y_i^{N/R}$.

$$y_i^k = 2c_i^k - 1 \tag{6}$$

The output frames of sensors are transmitted through noisy orthogonal channels to the destination. The received signal is the combination of all $R_i = Y_i + N_i$, where N_i is Gaussian channel noise between the i^{th} sensor and the receiver. The k^{th} bit of R_i is

$$r_i^k = y_i^k + n_i^k \tag{7}$$

The goal is to estimate the common data which is shown by $\hat{S}$ based on the received signals, hence the problem is to minimize $E[(S - \hat{S})^2|R_1, R_2, ..., R_M]$.

3. Distributed coding for three sensors with correlated data

In this section, a coding method for the simple case of three sensors $M = 3$ with equal correlation parameters $P_1 = P_2 = P_3 = P$ is introduced. In the proposed approach as depicted in fig. 1, the first sensor transmits its data without coding so that the output stream of the first sensor contains only the systematic bits $C_1 = \{c_1^1, c_1^2, ..., c_1^N\} =$

$\{i_1^1, i_1^2, ..., i_1^N\}$, the second sensor encodes data using an RSC encoder with polynomial $f(D) = 1 + D^2 + D^3$ and $g(D) = 1 + D + D^3$, where $f(.)$ and $g(.)$ are forward and feedback polynomials, respectively. The resulting parity bits $C_2 = \{c_2^1, c_2^2, ..., c_2^N\} = \{p_1^1, p_1^2, ..., p_1^N\}$ are transmitted to the receiver. Finally the third sensor encodes its data with the same RSC encoder after random interleaving to form the output sequence $C_3 = \{c_3^1, c_3^2, ..., c_3^N\} = \{p_2^1, p_2^2, ..., p_2^N\}$. These three sensors together form a distributed structure of a generic turbo encoder. Data in three branches of encoder are not the same, while they are correlated versions of a single data source and consequently may differ slightly. This property which makes the encoder structure different from ordinary turbo codes, will be considered in design of decoder structure that is described in the next section.

4. Decoding structure

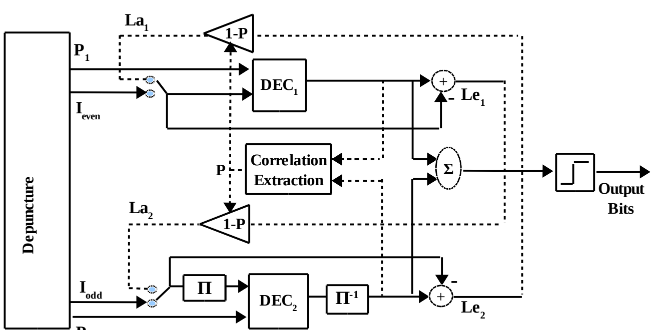

Fig. 2: Decoder structure based on turbo decoder with self correlation estimation block

All the frames transmitted by sensors are gathered at the receiver to form one systematic bit set and two parity sets. Inspired by this fact, a turbo-based decoder with some modifications is implemented at the receiver. The Max-Log-MAP algorithm is used in RSC decoders to decrease the complexity utilizing $log(e^x + e^y) \approx max(x, y)$ approximation [13].

As shown in fig. 2, in the designed turbo decoder, the two parity sets correspond to correlated versions of the systematic bits. Therefore applying the systematic bits to one of the decoders as apriori information as is done in a standard turbo decoder, biases the output of this particular decoder to systematic bits. This effect degrades the performance, especially for large values of correlation crossover probability. To remove this undesired effect, half of the systematic bits (odd bits) is applied as apriori information to the first decoder and the second part (even bits) is applied to the second decoder. This modification, makes a balance between these two decoders to increases the BER performance.

The second modification is based on considering the relation between [the] inputs of three encoders according to (5) at transmitter side. The similar relation, $L_{a_2} = (1-2\hat{P}) \times L_{e_1}$ is applied to [the] extrinsic Log likelihood Ratio (LLR) of each decoder before using them as apriori information for the other decoder, where $\hat{P}$ is the estimated value of correlation crossover probability using a new method presented in next section.

The last modification is that the average of output LLRs of both decoders $L_{av}^{out} = \frac{L_1^{out} + L_2^{out}}{2}$ is used instead of using output LLRs of one decoder in a standard turbo decoder, to estimate the data in [the] last iteration. The BER improvement due to these modifications are presented in section 5.

4.1 Correlation extraction by post processing of estimated data

In this section, the correlation estimation scheme is demonstrated. $Y_i = \{y_i^1, y_i^2, ..., y_i^N\}$ is the transmitted BPSK symbols by the i^{th} sensor. If $\hat{y}_i^k$ be the estimation of these symbols at the i^{th} RSC decoder, after each iteration of decoding algorithm, it intends to get more and more close to transmitted systematic symbols and cancels the noise effect, Therefore we have

$$\hat{y}_i^k \approx y_i^k = 2x_i^k - 1 \tag{8}$$

Considering the correlation model in (1), it converts to

$$\hat{y}_i^k = 2(s^k \oplus e_i^k) - 1 \tag{9}$$

To estimate the BSC crossover probability between sensors as correlation parameter sensors, we define a new random variable ρ_{ij}^k as

$$\rho_{ij}^k = \frac{| \hat{y}_i^k - \hat{y}_j^k |}{2} \tag{10}$$

substituting (9) in (10) results in

$$\rho_{ij}^k = | (s^k \oplus e_i^k) - (s^k \oplus e_j^k) | \tag{11}$$

There is two possible cases, first we consider a case when the error occurs in both sensors or in none of the sensors, $e_i^k = e_j^k$. The probability of this event is $(1 - P_i)(1 - P_j) + P_i P_j = 1 + 2P_i P_j - P_i - P_j$ and in this case, the two terms cancel each other,

$$\begin{aligned} P(\rho_{ij}^k = 0) &= P(e_i^k = e_j^k) \\ &= 1 + 2P_i P_j - P_i - P_j \end{aligned} \tag{12}$$

The other case is when the error occurs just in one sensor. It occurs with probability $P_i(1 - P_j) + (1 - P_i)P_j = P_i + P_j - 2P_i P_j$, without loss of generality, bit flipping due to error is assumed to be at sensor i, $e_i^k = \bar{e}_j^k = 1 - e_j^k = 1$, it follows,

$$\rho_{ij}^k = \mid (s^k \oplus 1) - (s^k \oplus 0) \mid$$
$$= \mid \bar{s}^k - s^k \mid = 1 \quad if\ e_i^k \neq e_j^k \tag{13}$$

we used the fact that, s^k is either 0 or 1, therefore $\bar{s}^k - s^k = \pm 1$. Therefore we have

$$P(\rho_{ij}^k = 1) = P(e_i^k \neq e_j^k) = P_i + P_j - 2P_iP_j \tag{14}$$

Expected value of ρ_{ij}^k is derived using (12) and (14)

$$E(\rho_{ij}^k) = 0.P(\rho_{ij}^k = 0) + 1.P(\rho_{ij}^k = 1)$$
$$= (1 - P_i)P_j + (1 - P_j)P_i = P_{ij} \tag{15}$$

Since ρ_{ij}^k is either 0 or 1, we have $(\rho_{ij}^k)^2 = \rho_{ij}^k$, and consequently

$$E[(\rho_{ij}^k)^2] = E(\rho_{ij}^k) = P_{ij} \tag{16}$$

ρ_{ij}^k and ρ_{ij}^l are independent RVs, and their joint expected value can be calculated as

$$E(\rho_{ij}^k \rho_{ij}^l) = E(\rho_{ij}^k)E(\rho_{ij}^l) = P_{ij}^2 \tag{17}$$

Taking the average of ρ_{ij}^k over all bits of the frame, a new variable $\hat{P}$ is defined as

$$\hat{P} = \frac{\sum_{k=1}^{N} \rho_{ij}^k}{N} \tag{18}$$

Using the above calculation it concludes that

$$E(\hat{P}) = \frac{\sum_{k=1}^{N} E(\rho_{ij}^k)}{N} = E(\rho_{ij}^k) = P_{ij} \tag{19}$$

$$E(\hat{P}^2) = \frac{1}{N^2} E(\sum_{k=1}^{N} \sum_{l=1}^{N} \rho_{ij}^k \rho_{ij}^l)$$
$$= \frac{1}{N^2} \sum_{k=1}^{N} E[(\rho_{ij}^k)^2] + \frac{1}{N^2} \sum_{k=1}^{N} \sum_{l=1, l \neq k}^{N} E(\rho_{ij}^k \rho_{ij}^l)$$
$$= \frac{1}{N}(P_{ij}) + \frac{N(N-1)}{N^2}(P_{ij})^2 \tag{20}$$
$$\sigma_{\hat{P}}^2 = E(\hat{P}^2) - E^2(\hat{P}) = \frac{1}{N}(P_{ij} - P_{ij}^2) \tag{21}$$

Therefore $\hat{P}$ is a RV with mean P_{ij} and variance $\sigma_{\hat{P}}^2$ which approaches zero if the frame length N is selected large enough. In the other words, $\hat{P}$ is a good approximation of pairwise BSC crossover probability P_{ij} that models the correlation among sensors. The above justification is true for arbitrary number of sensors and for especial case of three sensors with equal crossover probability we have $P_1 = P_2 = P_3 = P$, and consequently $\hat{P} \approx P_{ij} \approx P_i + P_j = 2P$.

5. Simulation Results

In this section the simulation results for proposed systems is presented. In all simulations unless otherwise explicitly specified, data frames with 256 bits length, BPSK modulation, i.i.d equiprobable Bernoulli bit stream, orthogonal AWGN channels and BSC correlation model with parameter P=0.01 are used.

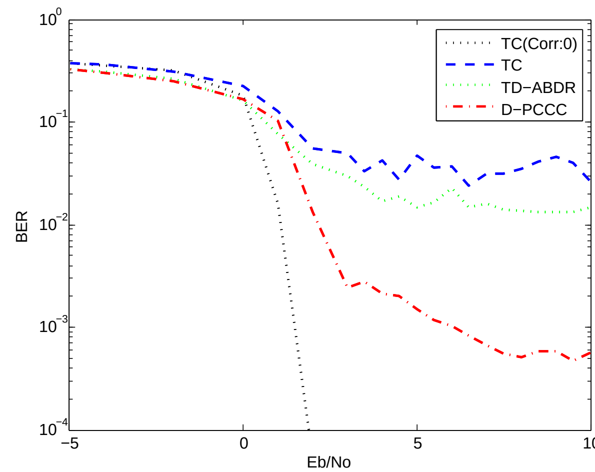

Fig. 3: Comparison of different decoding schemes for 3 sensors with correlated data, correlation parameter: 0.01, TC: Turbo Decoder, TC-ABDR: Modified Turbo Decoder with Average Based Decision Rule, D-PCCC Distributed Parallel Concatenated Convolutional Coders with Modified Multiple Turbo Decoder

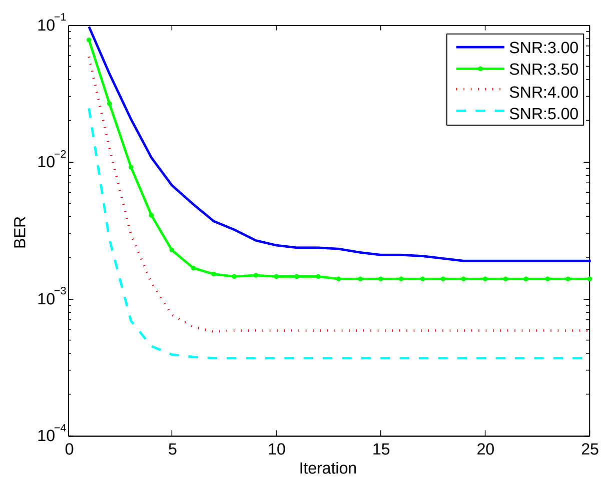

Fig. 4: BER performance of modified Multiple Turbo Decoder for 4 sensors versus different decoder iterations, Correlation parameter is 0.01

Fig. 3 shows the system BER performance [for/of] the proposed system in which, three sensors performing as a distributed turbo encoder. TC represents basic turbo decoder applied at the receiver. While Modified-TC represents the

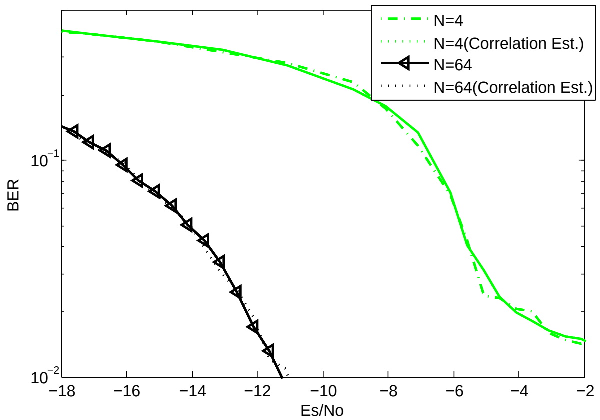

Fig. 5: comparison of BER performance of modified Multiple Turbo Decoder with known correlation and self estimation of correlation at the receiver, correlation parameter: 0.05

modified average based decision rule turbo decoder. The simulation results show 50% improvement in BER floor, because BER floor of proposed scheme is about 0.02, which is approximately half of the BER floor of a generic turbo decoder in the same SNR value.

Convergence of the proposed system is analyzed in fig. 4. The decoder almost reaches its final performance for iteration values between 5 and 10 for different SNR values, while this value is normally more than 10 for a turbo decoder. This is based on the fact that the output LLR's of RSC decoders represent the correlated versions of the source data. Therefore the system performance can not be improved by more iterations and system saturates very fast.

Fig. 5 demonstrates the effect of the correlation parameter estimation in the proposed method. In the first method, the correlation parameter is assumed known at the receiver, while in the second method, it is estimated by post processing of the received symbols. Simulation results prove the accuracy of the developed method since the system BER performance for two cases are almost the same and there is no considerable BER performance degradation due to estimation of correlation parameter at the receiver.

6. Conclusion

In this article, a distributed joint source-channel coding technique based on turbo codes for a CEO problem in WSNs is proposed. The simulation results show about 50% improvement in BER floor for the proposed system [in comparison with/compared to] the same scenario and using a standard turbo decoder.

The BSC crossover probability is considered as correlation parameter of system model and is used in proposed decoder structure to improve performance and convergence of decoding scheme. A new method is developed to estimate the correlation parameter at the decoder. It makes the proposed system an appropriate option to employ in applications with unknown, unpredictable or time-varying correlation factor without considerable loss in [the] system BER performance.

The system converges even faster than an standard turbo decoder. Data rate of sensors and consequently the system data rate can be managed using different puncturing methods.

Another benefit of the proposed system is scalability of turbo encoder employing multi branch structure. Therefore arbitrary number of sensors in each cluster can be selected based on [the] average correlation parameter, sensors' power constraint, noise power, coding rate and available bandwith. The system can be extended for other channel models such as fading and interference channels.

References

[1] Y. Li, "Distributed coding for cooperative wireless networks: An overview and recent advances," *IEEE Communications Magazine*, vol. 47, no. 8, pp. 71 - 77, Aug. 2009.

[2] D. Slepian, and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Transactions on Information Theory*, vol. 19, no. 4, pp. 471-480, Jul. 1983.

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

[4] S. S. Pradhan, and K. Ramchandran, "Distributed source coding using syndromes (DISCUS): design and construction", *IEEE Transactions on Information Theory*, vol. 49, no. 3, pp. 626-643, Mar. 2003.

[5] A. D. Liveris, Z. Xiong, and C. N.Georghiades, "Compression of binary sources with side information at the decoder using LDPC codes," *IEEE Communications Letters*, vol. 6, pp. 440-442, Oct. 2002.

[6] J. Bajcsy, and P. Mitran, "Coding for the Slepian-Wolf problem with turbo codes,"*Global Telecommunications Conference (GLOBECOM) 2001*, vol. 2, pp. 1400-1404, Nov. 2001.

[7] A. Aaron, and B. Girod, "Compression with side information using turbo codes," *IEEE Data Compression Conference (DCC) 2002*, pp. 252-261, Apr. 2002.

[8] V. Stankovic, A.D. Liveris, X. Zixiang Xiong, and C.N. Georghiades, "On code design for the Slepian-Wolf problem and lossless multiterminal networks," *IEEE Transactions on Information Theory*, vol. 52, pp. 1495 - 1507, 2006.

[9] C. Stefanovic, D. Vukobratovic, and V. Stankovic, "*IEEE Information Theory Workshop, ITW 2009*, pp. 208 - 212, 2009.

[10] D. Gunduz, E. Erkip, A. Goldsmith, and H.V. Poor, "Source and Channel Coding for Correlated Sources Over Multiuser Channels," *IEEE Transactions on Information Theory* vol. 55, pp. 3927 - 3944, Sep. 2009.

[11] T. Berger, Zhen Zhang, and H. Viswanathan, "The CEO Problem," *IEEE Transactions on Information Theory* pp. 887-902, 1996.

[12] W. Zhong, and J. Garcia-Frias, "Combining data fusion with joint source-channel coding of correlated sensors," *IEEE Information Theory Workshop*, pp. 315-317, Oct. 2004.

[13] R. Ghaffar, and R. Knopp, "Analysis of Low Complexity Max Log MAP Detector and MMSE Detector for Interference Suppression in Correlated Fading," *Global Telecommunications Conference (GLOBECOM) 2009*, pp. 1-6, 2009.