WYNER-ZIV CODING FOR THE HALF-DUPLEX RELAY CHANNEL

Zhixin Liu, Vladimir Stanković, and Zixiang Xiong

Department of Electrical Engineering, Texas A&M University, College Station, Texas 77843
Tel: (979) 862-8683; Fax: (979) 862-4630; E-mail: {liuzx,stankovi,zx}@ee.tamu.edu

ABSTRACT

Cover and El Gamal derived the tightest bounds on the capacity of the relay channel using random coding and proposed two coding strategies, namely decode-and-forward (DF) and compress-and-forward (CF), to provide the best known lower bounds of the achievable rate region. Depending on transmission parameters, either DF or CF could be superior. Several practical code designs based on DF have appeared recently. We present the first practical CF design for the half-duplex Gaussian relay channel based on Wyner-Ziv coding of the received source signal at the relay. Assuming ideal source and channel coding, our design achieves the lower bound of CF. It thus realizes the performance gain of CF over DF promised by the theory when the relay is close to the destination. Our practical implementation based on LDPC codes for error protection at the source and nested scalar quantization and IRA codes for Wyner-Ziv coding at the relay comes as close as 0.76 dB to the theoretical limit of CF.

1. INTRODUCTION

The relay channel is a three-terminal channel where the source communicates messages to the destination with the help of an intermediate relay node (see Fig. 1). The source broadcasts encoded messages to the relay and destination; the relay processes the received information and forwards the resulting signal to the destination. The destination collects signals from both the source and relay before attempting to recover the source information. The relay's task is thus to facilitate decoding at the destination by means of spatial/temporal diversity.

Cover and El Gamal [1] derived the tightest upper and lower bounds on the capacity of the relay channel. These two bounds in general do not coincide; one exception is when the relay channel is degraded. However, real-world wireless relay channels are not degraded, and except in few special cases [2, 3], the two bounds do not coincide. Several random coding schemes [1, 4, 2, 5, 6] have been proposed to obtain the lower bounds on the achievable rate region. These schemes can be grouped into two classes: decode-and-forward (DF) and observe-and-forward [1].

In the DF scheme, the relay decodes the received source message, re-encodes it, and forwards the resulting signal to the destination. Note that, since the relay must perfectly decode the source message, the achievable rates are bounded by the capacity of the channel between the source and relay. Consequently, if the channel between the source and destination is better than that between the source and relay, the relay cannot improve the transmission, that is, DF performs worse than direct transmission.

To alleviate this problem, a class of observe-and-forward schemes has been proposed, where the relay does not attempt to decode the signal from the source, but only observes it. For example, in the widely exploited amplify-and-forward (AF) scheme [4], which falls into this class, the relay just amplifies the received source signal. Though the AF scheme is of very low complexity, it has never been shown that it can outperform DF.

Compress-and-forward (CF), rooted in the original work of Cover and El Gamal [1], is another observe-and-forward scheme. As the name CF suggests, the relay compresses the signal it has received from the source within certain distortion. This can be achieved, for example, with a simple quantizer. Several research groups [5, 2, 3] have recently pointed out the important role of Wyner-Ziv (WZ) coding [7] (or lossy source coding with side information at the decoder) in achieving the lower performance bounds of CF. But until now, these observations have only remained on a theoretical level.

CF has higher coding complexity than AF, but it gives many rate points that are not achievable with any other coding strategy. Intuitively, DF performs well when the channel between the source and relay is clean (e.g., the relay is close to the source), whereas CF is desirable when the channel between the relay and destination is good. However, the main attraction of CF is that, in contrast to DF, it always outperforms direct transmission; thus, even if the link between the source and the relay is very poor, the relay can still help.

In this paper, we consider a half-duplex relay channel (also referred to as the 'cheap' relay channel in [5]), where the relay cannot simultaneously receive and transmit. Half-duplex relaying [6, 3, 5, 8] is more practical because constructing a relay that can operate in the full-duplex mode [1] is very expensive (due to, e.g., the large difference in the transmitting and receiving signal power levels). Wireless and Gaussian half-duplex relay channels are studied in [3], where the upper bound and the lower bounds for both DF and CF are given.

Practical code designs for wireless relay channels and cooperative networks are considered in [8, 9, 10, 11]. But these designs only exploit AF or DF, and thus, they can at their best approach the lower bound of DF, which is away from the CF limit in many cases [3]. Up to date, there is no practical code design based on CF. This is because WZ code design has not been well understood until recently [12]. A WZ coder typically consists of a quantizer followed by syndrome-based Slepian-Wolf coding [13] for compression (implemented via channel coding). Furthermore, when the channel is noisy, additional channel coding is needed to protect the syndromes in what is called distributed joint source-channel (DJSC) coding [14].

Motivated by the large coding gains offered by CF, we focus on the Gaussian half-duplex relay channel and design a practical CF-based scheme. The main idea is to split the message at the source into two parts, protect them independently with two different channel codes, and transmit the coded source in two separate fractions of a time slot; the relay compresses the source received during the relay-receive period using WZ coding, adds error protection, and sends the resulting bitstream to the destination during the relay-transmit period. Thus, the scheme consists of two classic channel coding components (for transmissions to the relay and destination) at the source and one WZ coding component (for transmission to the destination) at the relay. The latter is implemented via quantization followed by a channel coder for joint Slepian-Wolf compression [13] and error protection. We show that our scheme can achieve the lower bound of CF [3] assuming ideal source and channel coding; thus, it can be employed to realize all the performance gain of CF over DF in the wide range of SNRs as promised by the theory.

In our implementation, WZ coding and error protection at the relay are performed jointly via DJSC coding [14] with nested quantization and systematic irregular repeat-accumulate (IRA) codes [15]. We derive the inherent loss of our practical scheme from the theoretical CF bound due to practical source coding (assuming ideal channel coding); since our analysis shows a small performance loss by employing a low-dimensional source coder, for the sake of low system complexity, we resort to the simplest nested scalar quantizer. Our design with LDPC codes for error protection at the source outperforms the lower bound of DF when the relay is close to the destination. Moreover, in certain cases even when DF is theoretically better than CF, due to its high coding efficiency, our practical CF design outperforms the best practical DF implementation of [8] based on distributed turbo codes.

2. THEORETICAL BOUNDS

The relay channel is shown in Fig. 1, where c_{sr}, c_{sd}, and c_{rd} denote channel coefficients, which are assumed to be constant. We consider the half-duplex mode (achieved e.g., using time division). During each time slot, let α be the fraction of time when the relay operates in the receive mode; then, $1-\alpha$ is the fraction of time when the relay can only transmit. Let P_s and P_r be the average source and relay power constraints, respectively.

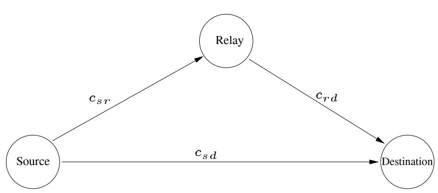

Fig. 1. The Gaussian relay channel.

The source splits a message $m \in \{1, \dots, M\}$ into two parts, m_1 and m_2, and independently encodes them into an $n\alpha$-length codeword x_{s1} and an $n(1-\alpha)$-length codeword x_{s2}. During the relay-receive period, x_{s1} is modulated to X_{s1} and broadcasted with power $\frac{kP_s}{\alpha}$, $k \leq 1$. The received signals at the relay and destination are:

$$Y_r = c_{sr}X_{s1} + Z_{sr} \tag{1}$$

$$\text{and} \quad Y_{d1} = c_{sd}X_{s1} + Z_{sd}, \tag{2}$$

respectively, where Z_{sr} and Z_{sd} are independent additive white Gaussian noises with unit power. The relay processes the received signal Y_r and obtains an $n(1-\alpha)$-length codeword x_r, which is then BPSK modulated, resulting in signal X_r, which is forwarded to the destination with power $\frac{P_r}{1-\alpha}$ during the relay-transmit period. In the meanwhile, the source modulates x_{s2} and sends the resulting signal X_{s2} with power $\frac{(1-k)P_s}{1-\alpha}$. The received signal at the destination is then

$$Y_{d2} = c_{sd}X_{s2} + c_{rd}X_r + Z, \tag{3}$$

where Z is an independent additive white Gaussian noise with unit power.

The upper bound on the capacity of the Gaussian half-duplex relay channel is derived in [3, 5] from the max-flow-min-cut theorem and is given by:

$$C_{ub} = \max_{0\leq\rho\leq 1, 0\leq\alpha\leq 1, 0\leq k\leq 1} \min\{C_{ub1}, C_{ub2}\}, \tag{4}$$

where C_{ub1} and C_{ub2} are:

$$\begin{aligned} C_{ub1} &= \alpha\frac{1}{2}\log(1+(c_{sr}^2+c_{sd}^2)\frac{kP_s}{\alpha}) + \\ &\quad (1-\alpha)\frac{1}{2}\log(1+(1-\rho^2)c_{sd}^2\frac{(1-k)P_s}{1-\alpha}), \\ C_{ub2} &= \alpha\frac{1}{2}\log(1+c_{sd}^2\frac{kP_s}{\alpha}) + (1-\alpha)\frac{1}{2}\log(1+c_{sd}^2 \\ &\quad \frac{(1-k)P_s}{1-\alpha} + c_{rd}^2 P_r + \frac{2\sqrt{\rho^2 c_{sd}^2 c_{rd}^2 (1-k)P_s P_r}}{1-\alpha}), \end{aligned}$$

respectively. Note that the parameter ρ reflects the correlation between the source and relay signals, and it can be written in closed form [3, 5].

The achievable rate for DF is [3, 5]:

$$R_{DF} \leq \max_{0\leq\rho\leq 1, 0\leq\alpha\leq 1, 0\leq k\leq 1} \min\{R_{DF1}, R_{DF2}\}, \tag{5}$$

$$\begin{aligned} \text{where} \quad R_{DF1} &= \alpha\frac{1}{2}\log(1+c_{sd}^2\frac{kP_s}{\alpha}) + \\ &\quad (1-\alpha)\frac{1}{2}\log(1+(1-\rho^2)c_{sd}^2\frac{(1-k)P_s}{1-\alpha}) \end{aligned}$$

and $R_{DF2} = C_{ub2}$. The achievable rate for CF is given by [3]:

$$R_{CF} \quad \leq \quad \max_{0\leq\alpha\leq 1, 0\leq k\leq 1} \{R_r(\alpha,k) + R_d(\alpha,k)\}, \tag{6}$$

$$\text{where} \quad R_r(\alpha,k) = \alpha\frac{1}{2}\log(1+c_{sd}^2\frac{kP_s}{\alpha} + \frac{c_{sr}^2 kP_s}{\alpha(1+\sigma_\omega^2)})$$

$$\text{and} \quad R_d(\alpha,k) = (1-\alpha)\frac{1}{2}\log(1+c_{sd}^2\frac{(1-k)P_s}{1-\alpha}),$$

with σ_ω^2 being the WZ compression noise [7] given by

$$\sigma_\omega^2 = \frac{1+(c_{sr}^2+c_{sd}^2)kP_s/\alpha}{((1+\frac{c_{rd}^2 P_r/(1-\alpha)}{1+c_{sd}^2(1-k)P_s/(1-\alpha)})^{\frac{1-\alpha}{\alpha}}-1)(1+c_{sd}^2\frac{kP_s}{\alpha})}.$$

DF and CF give the best known lower bounds of the achievable rates. Depending on transmission parameters, either DF or CF is superior [3] (see Fig. 4). Indeed, DF outperforms CF when the link between the source and relay is better than that between the relay and destination (e.g., when the relay is located close to the source); on the other hand, CF provides higher achievable rates when the link between the relay and destination is clean (e.g., when the relay is close to the destination).

3. SYSTEM DESCRIPTION

Our proposed CF-based practical scheme for half-duplex wireless relaying is depicted in Fig. 2. The source node consists of two classic channel encoders and a BPSK modulator. The relay performs DJSC encoding; that is, it *jointly* performs WZ compression of the received signal Y_r (assuming the side information Y_{d1} at the decoder) and error protection against noise in the link between the relay and destination. The receiver contains two classic channel decoders, a DJSC decoder, and an estimator.

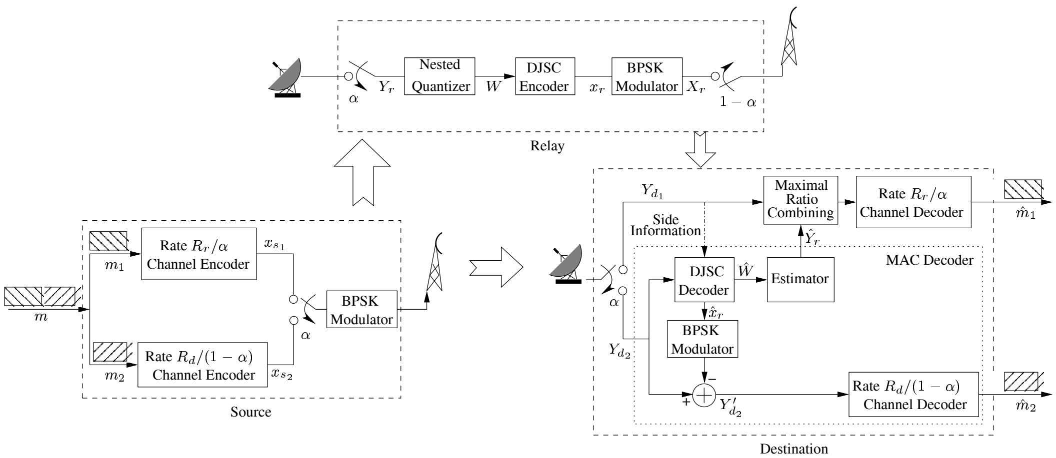

Fig. 2. Block diagram of our proposed CF-based system.

We closely follow the CF coding steps of [3]. Given the average source and relay power constraints P_s and P_r, respectively, from (6), we compute the achievable rate $R_{CF} = R_r(\alpha, k) + R_d(\alpha, k)$ by optimizing over k and α. The M-length ($R_{CF} = \log M/n$) source message m is split into two independent parts m_1 and m_2; the nR_r-length part m_1 is first encoded by the channel encoder of rate $\frac{R_r}{\alpha}$; the resulting $n\alpha$-length codeword x_{s1} is BPSK modulated to X_{s1} and sent with power $\frac{kP_s}{\alpha}$. At the relay, the received signal Y_r, given by (1), is first quantized with a nested quantizer with nesting ratio N; then, the obtained $n\alpha$ indices, W, are encoded bitplane-by-bitplane with $\log N$ systematic DJSC encoders, which jointly perform Slepian-Wolf compression and error protection [14]; their coding rates are determined according to the WZ bound [7]. Only the resulting parity-check symbols are BPSK modulated to X_r and sent during the relay-transmit period. Using an $\frac{R_d}{1-\alpha}$-rate error protection code, the source encodes m_2, modulates the obtained $n(1-\alpha)$-length codeword x_{s2} into X_{s2}, and forwards it to the destination with power $\frac{(1-k)P_s}{1-\alpha}$.

During the relay-transmit period the channel is a multiple access channel (MAC) given by (3). The receiver starts by recovering m_2 using successive cancellation decoding. That is, first X_r is reconstructed using DJSC decoding (see below); then X_r is subtracted from Y_{d2}, giving $Y'_{d2} = Y_{d2} - c_{rd}X_r = c_{sd}X_{s2} + Z$; from Y'_{d2}, m_2 is now recovered with the $\frac{R_d}{1-\alpha}$-rate error protection decoder.

DJSC decoding is done bitplane-by-bitplane similarly as in [14]. The main idea is to view the system as transmitting the bits over two channels; the first one is the actual MAC with noise $Z + X_{s2}$ which describes the distortion experienced by the parity-check bits; the second channel is the "virtual" correlation channel between Y_r and the side information Y_{d1}. Thus, to compute the log-likelihood ratio's (LLR's) needed for joint iterative decoding, for the systematic and parity-check parts, the conditional pdf of W given Y_{d1} and the conditional pdf of X_r given Y_{d2} are exploited, respectively. In this way the decoder reconstructs the quantization indices $\hat{W}$ and estimates $\hat{Y}_r$. Finally, m_1 is recovered using maximum ratio combining (MRC) of $\hat{Y}_r$ and $Y_{d1} = c_{sd}X_{s1} + Z_{sd}$ (recall that $Y_r = c_{sr}X_{s1} + Z_{sr}$) and decoding with the $\frac{R_r}{\alpha}$-rate error protection decoder.

The next proposition gives the loss of our scheme from the lower bound of CF due to practical source coding (quantization) assuming ideal channel coding.

Proposition 1 *Assuming ideal channel coding, the rate loss from the CF limit due to practical source coding is*

$$R' - R_{CF} = \frac{\alpha}{2}\log\frac{1 + c_{sd}^2\frac{kP_s}{\alpha} + \frac{c_{sr}^2 kP_s/\alpha}{1+\sigma_{p\omega}^2}}{1 + c_{sd}^2\frac{kP_s}{\alpha} + \frac{c_{sr}^2 kP_s/\alpha}{1+\sigma_{\omega}^2}}, \qquad (7)$$

where $\sigma_{p\omega}^2$ is the WZ compression noise of the source coder.

Note that $\frac{\sigma_{p\omega}^2}{\sigma_\omega^2} > 1$ quantifies the distortion loss of the practical coder from the WZ bound [7]. For example, for scalar quantization at high SNRs $\sigma_{p\omega}^2 = \frac{2\pi e}{12}\sigma_\omega^2$ [12]. In the next section, applying Proposition 1, we show that our scheme does not suffer large performance loss by employing a nested scalar quantizer.

4. SIMULATION RESULTS

In this section we compare the performance of our proposed CF scheme and the DF design in [8] (with our own implementation) for the Gaussian half-duplex relay channel (with fixed channel coefficients) described in Section 2. The experimental setup, shown in Fig. 3, is the same as in [8]. It corresponds to the practical transmission setting with frequency 2.4 GHz, path loss coefficient 3, and free-space reference distance 1 m. The relay is located along a straight line between the source and destination, which are $r = 10$ m apart. Its distance to the source is d m (with its distance to the destination being $r - d$ m).

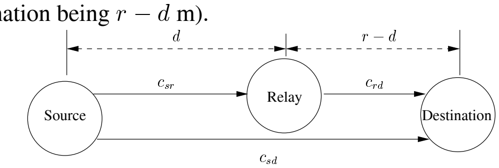

Fig. 3. Experimental setup. The relay is located along a straight line from the source to destination. The distance between the source and destination is r =10 m.

We target at the transmission rate R at 2 bits per channel use and compute the optimal (P_s, k, α) for each d ($0 < d < 10$) while fixing P_r at 70 dB. This allows us to plot the required minimum P_s according to the upper bound in (4), the lower bounds of DF in (5) and CF in (6), and the limiting performance in (7) of our practical CF scheme under scalar quantization vs. d in Fig. 4. It is seen that DF is more efficient than CF when $d < 8.1$ m; whereas CF theoretically outperforms DF when $d > 8.1$ m (i.e., when the relay is close to the destination). We thus set d to 5, 7.5 and 9 m in our CF simulations. When $d = 5$ m, the channel coefficients are set to $c_{sr}^2 = -60.97$ dB, $c_{sd}^2 = -70$ dB, and $c_{rd}^2 = -60.97$ dB; when $d = 7.5$ m, the coefficients are $c_{sr}^2 = -67.23$ dB, $c_{sd}^2 = -70$ dB, and $c_{rd}^2 = -48.45$ dB; and when $d = 9$ m, they are $c_{sr}^2 = -68.67$ dB, $c_{sd}^2 = -70.05$ dB, and $c_{rd}^2 = -40.05$ dB.

In our simulations, we assume that the distribution of source messages is uniform and that the channels' statistics are known at all three nodes. Prior to communication, the pdf's needed for decoding are stored as look-up tables at the destination. Then, we increase the power levels until the overall message bit error rate is decreased below 10^{-5}. For coding m_1 and m_2, we employ two different LDPC codes designed using density evolution. For practical DJSC coding we resort to nested scalar quantization and systematic IRA codes designed via density evolution. The maximum codeword length is $n = 100,000$ bits.

Simulation results with CF are marked by stars in Fig. 4. When $d = 9$ m, our practical result is better than the DF limit and is 0.76 dB away from the CF limit (and 1.8 dB away from the upper bound of half-duplex relaying). Note that the loss mainly comes from channel coding since the gap to the CF bound (assuming ideal channel coding) due to scalar quantization is only 0.12 dB in this case. When $d = 7.5$ and 5 m, our practical CF design performs 0.81 dB and 1.54 dB away, respectively, from the corresponding theoretical CF limit.

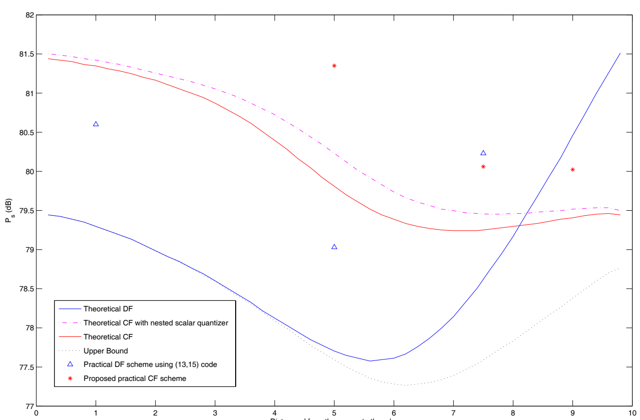

Fig. 4. Theoretical bounds (in terms of P_s) and simulation results for different d's.

Since DF outperforms CF for $d < 8.1$ m in theory, we re-implement the practical DF scheme referred to as "distributed rate 1/4 parallel strong concatenated convolutional coding" in [8] for $d = 1$, 5 and 7.5 m. The scheme exploits a $(13, 15)_8$ recursive systematic convolutional code at both the source and the relay; the source and relay jointly construct a distributed turbo coder, which besides spatial diversity gain of DF, achieves extra coding gain due to the interleaving gain and the turbo processing gain. The results are marked by triangles in Fig. 4. For $d = 1$ and 5 m, our DF implementation loses 1.3 dB and 1.67 dB, respectively, compared to the corresponding theoretical DF limit; but these results are still better than those with practical CF. When $d = 7.5$ m, however, even though DF is theoretically better than CF by 0.69 dB, our practical CF design outperforms our DF implementation of [8] by 0.17 dB in terms of P_s.

5. CONCLUSIONS AND FUTURE WORK

We have proposed the first practical CF-based scheme for the half-duplex relay channel. By employing strong channel coding at the source and 1-D nested quantization for DJSC coding, the performance of our design comes very close to the CF theoretical limits. Overall, it is better to use CF for large d (when the relay is close to the destination), and DF for small d (when the relay is close to the source). In between, there is a small range of d, for which it is better to use CF in practice even though DF is superior in theory. Our work represents a major step towards providing practical coding designs to pure multiterminal communication problems by the means of distributed source coding.

Since the loss of our design to the theoretical limits of CF is already very small, the only way of achieving sizable improvement is to reduce the gap to the upper bound. The main loss of CF comes from MAC decoding during the relay-transmit period, where the signals from the source and relay are made independent. However, we believe that further gains can be obtained by exploiting the knowledge of m_1 (and thus of the statistics of the relay signal) in encoding m_2. Then, joint decoding similar to that in [16] instead on MRC can be employed.

6. REFERENCES

[1] T. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inform. Theory*, vol. 25, pp. 572–584, Sept. 1979.

[2] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," submitted to *IEEE Trans. Inform. Theory*, Feb. 2004.

[3] A. Høst-Madsen and J. Zhang, "Capacity bounds and power allocation for the wireless relay channel," submitted to *IEEE Trans. Inform. Theory*, Sept. 2003.

[4] J.N. Laneman, D.N.C. Tse, and G.W. Wornell, "Cooperative diversity in wireless networks: efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, vol. 50, pp. 3062– 3080, Dec. 2004.

[5] M. Khojastepour, A. Sabharwal, and B. Aazhang, "On capacity of Gaussian 'cheap' relay channel," *Proc. Globecom-2003*, San Francisco, CA, Dec. 2003.

[6] K. Azarian, H. El Gamal, and P. Schniter, "Achievable diversity -vs-multiplexing tradeoffs in half-duplex cooperative channels," *Proc. ITW-2004,* San Antonio, TX, Oct. 2004.

[7] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inform. Theory*, vol. 22, pp. 1–10, Jan. 1976.

[8] M. Valenti and B. Zhao, "Distributed turbo codes: towards the capacity of the relay channel," *Proc. VTC-2003*, Orlando, FL, Oct. 2003.

[9] M. Janani, A. Hedayat, T. Hunter, and A. Nosratinia, "Coded cooperative in wireless communications: space-time transmission and iterative decoding," *IEEE Trans. Signal Processing,* vol. 52, pp. 362–371, Feb. 2004.

[10] A. Stefanov and E. Erkip, "Cooperative coding for wireless networks," *IEEE Trans. Communications,* vol. 52, pp. 1603–1612, Sept. 2004.

[11] Z. Zhang and T.M. Duman, "Capacity approaching turbo coding and iterative decoding for relay channels," submitted, 2004.

[12] Z. Xiong, A. Liveris, and S. Cheng, "Distributed source coding for sensor networks," *IEEE Signal Processing Magazine*, vol. 21, pp. 80-94, Sept. 2004.

[13] D. Slepian and J.K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory,* vol. 19, pp. 471–480, July 1973.

[14] A. D. Liveris, Z. Xiong, and C. N. Georghiades, "Joint source-channel coding of binary sources with side information at the decoder using IRA codes," *Proc. MMSP-2002*, St. Thomas, US Virgin Islands, Dec. 2002.

[15] H. Jin, A. Khandekar, and R McEliece, "Irregular repeat-accumulate codes," *Proc. 2nd Int. Symp. Turbo codes and related topics*, pp. 1–8, Sept. 2000.

[16] W. Zhong and J. Garcia-Frias, "Joint source-channel coding of correlated senders over multiple access channels," *Proc. of Allerton'04*, Monticello, IL, Oct. 2004.