Offset Encoding for Multiaccess Relay Channels

Lalitha Sankaranarayanan, *Student Member, IEEE*, Gerhard Kramer, *Member, IEEE*, and Narayan B. Mandayam, *Senior Member, IEEE*

***Abstract*— An offset encoding technique is presented that improves sliding-window decoding with decode-and-forward for K-user multiaccess relay channels. The technique offsets user transmissions by one block per user and achieves the corner points of the backward decoding rate region but with a smaller delay.**

***Index Terms*— Multiaccess communication, relaying, encoding, cooperative systems.**

I. INTRODUCTION

The multiaccess relay channel (MARC) is a network where several users communicate with a single destination in the presence of a relay [1]. Several coding strategies for the relay channel [2], [3] extend readily to the MARC [4], [5]. For example, the strategy of [3, Theorem 1], now often called *decode-and-forward* (DF), has a relay that decodes user messages before forwarding them to the destination [4], [5]. Similarly, the strategy in [3, Theorem 6], now often called *compress-and-forward* (CF), has the relay quantize its output symbols and transmit the resulting quantization bits to the destination [5].

For the classic relay channel, several block-Markov encoding and decoding techniques achieve the DF rate in [3, Theorem 1] (see [4, Sec. I]):

- *irregular* encoding (different size codebooks at the source and relay) and *successive* decoding [3, Theorem 1],
- *regular* encoding (same size codebooks at the source and relay) and *sliding-window* decoding [6],
- regular encoding and *backward decoding* [7].

One can, in fact, use irregular encoding with any of the above decoding methods. The above techniques have all been generalized to multiple relay networks [4], [8]–[11]. For the MARC, however, the different DF decoding methods do not always yield the same rate region. For example, we show that backward decoding can give larger rates than sliding-window decoding (see also [12], [13]). On the other hand, sliding-window decoding decodes blocks of message bits at regular intervals before all channel-symbol blocks are transmitted. This is useful: if the sliding window length is much smaller than the backward decoding delay, then sliding-window decoding is preferable for *streaming* data applications.

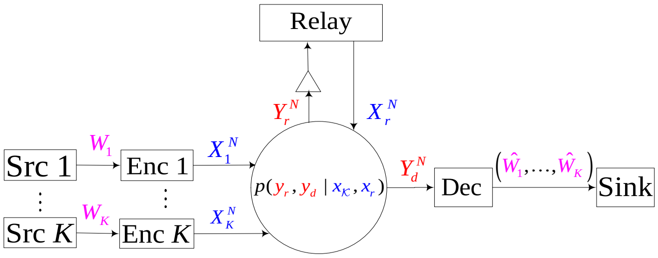


Fig. 1. A K-user multiaccess relay channel.

To compare the methods, suppose we use backward decoding for B message blocks transmitted in $B+1$ channel-symbol blocks. The decoding delay is then $B+1$ channel-symbol blocks for the first message block. Our main contribution is an *offset encoding* technique for sliding-window decoding that recovers the corner points of the backward decoding rate region with a delay of $K+1$ channel-symbol blocks for every message block. The total number of channel-symbol blocks required is $B+K$. Note that K can be much smaller than B, e.g., if the relay serves only a small number of users at a time. For the non-corner boundary points of the backward decoding rate region, we would require time-sharing that increases decoding delay. However, rate-splitting methods can potentially avoid the time-sharing delays [14].

This correspondence is organized as follows. In Section II we present the MARC model and summarize the DF random code construction of [4, Appendix A]. In Section III, we review the backward decoding rate region and compute the sliding-window decoding rate region. The latter region is in general smaller than the former. In Section IV, we describe offset encoding and develop its rate region when combined with sliding-window decoding. Section V concludes the paper.

II. PRELIMINARIES

A. Model and Notation

The K-user MARC has K sources, one relay, and one destination (see Fig. 1). The sources emit the messages W_k, $k = 1, 2, \ldots, K$, that are statistically independent and take on values uniformly in the sets $\{1, 2, \ldots, M_k\}$. The channel is used N times so that the rate of W_k is $R_{W_k} = B_{W_k}/N$ bits per channel use where $B_{W_k} = \log_2 M_k$ bits. The channel input $X_{k,i}$ from source k at time i, $i = 1, 2, \ldots, N$, is a function of W_k, while the relay's channel input $X_{r,i}$ is a causal function of its received signals $Y_r^{i-1} = (Y_{r,1}, Y_{r,2}, \ldots, Y_{r,i-1})$. The destination uses the N channel outputs Y_d^N to decode the K mes-

---

The work of L. Sankaranarayanan and N. B. Mandayam was supported in part by the National Science Foundation under Grant No. ITR-0205362. The work of G. Kramer was partially supported by the Board of Trustees of the University of Illinois Subaward No. 04-217 under NSF Grant No. CCR-0325673. The material in this correspondence was presented in part at the 42^{nd} Annual Allerton Conference on Communications, Control, and Computing, Monticello, IL, Sep. 2004.

L. Sankaranarayanan and N. B. Mandayam are with the WINLAB, Rutgers University.

G. Kramer is with Bell Labs, Alcatel-Lucent.

	Block 1	Block 2	Block 3		Block B	Block B+1
User 1	$\underline{x}_1(w_{1,1},1)$	$\underline{x}_1(w_{1,2},w_{1,1})$	$\underline{x}_1(w_{1,3},w_{1,2})$	…	$\underline{x}_1(w_{1,B},w_{1,B\text{-}1})$	$\underline{x}_1(1,w_{1,B})$
	$\underline{v}_1(1)$	$\underline{v}_1(w_{1,1})$	$\underline{v}_1(w_{1,2})$		$\underline{v}_1(w_{1,B\text{-}1})$	$\underline{v}_1(w_{1,B})$
User 2	$\underline{x}_2(w_{2,1},1)$	$\underline{x}_2(w_{2,2},w_{2,1})$	$\underline{x}_2(w_{2,3},w_{2,2})$	…	$\underline{x}_2(w_{2,B},w_{2,B\text{-}1})$	$\underline{x}_2(1,w_{2,B})$
	$\underline{v}_2(1)$	$\underline{v}_2(w_{2,1})$	$\underline{v}_2(w_{2,2})$		$\underline{v}_2(w_{2,B\text{-}1})$	$\underline{v}_2(w_{2,B})$
Relay	$\underline{x}_r(1,1)$	$\underline{x}_r(w_{1,1},w_{2,1})$	$\underline{x}_r(w_{1,2},w_{2,2})$	…	$\underline{x}_r(w_{1,B\text{-}1},w_{2,B\text{-}1})$	$\underline{x}_r(w_{1,B},w_{2,B})$

Fig. 2. Regular encoding for a two-user MARC assuming the relay decodes correctly.

sages as $(\hat{W}_1, \hat{W}_2, \dots, \hat{W}_K)$. We write $\mathcal{K} = \{1, 2, \dots, K\}$, $X_{\mathcal{S}} = \{X_k : k \in \mathcal{S}\}$ for all $\mathcal{S} \subseteq \mathcal{K}$, $\mathcal{S}^c$ to denote the complement of $\mathcal{S}$ in $\mathcal{K}$, and $|\mathcal{S}|$ for the cardinality of $\mathcal{S}$. The channel is time-invariant and memoryless with the conditional probability distribution

$$p(y_r, y_d | x_{\mathcal{K}}, x_r). \tag{1}$$

The capacity region $\mathcal{C}_{MARC}$ is the closure of the set of rate tuples $(R_{W_1}, R_{W_2}, \dots, R_{W_K})$ for which the destination can, for sufficiently large N, decode the K source messages with an arbitrarily small positive error probability.

As further notation, we write $R_{\mathcal{S}} = \sum_{k \in \mathcal{S}} R_k$, $[m, n] = \{m, m+1, \dots, n\}$, and we use the vector notation $\underline{x}_k$ for length-n codewords of user k. We use the usual notation for entropy and mutual information [15], [16] and take all logarithms to the base 2 so that our rate units are bits. We write random variables (e.g. W_k) with uppercase letters and their realizations (e.g. w_k) with the corresponding lowercase letters.

B. Random Code Construction

A DF code construction is presented in [4, Appendix A] and we review it below. This construction is common to all the decoding methods considered below and it uses independent random variables V_k, $k = 1, 2, \dots, K$, to help the sources cooperate with the relay.

Random Code Construction:

Consider the probability distribution

$$\left(\prod_{k=1}^{K} p(v_k) p(x_k | v_k) \right) \cdot p(x_r | v_{\mathcal{K}}). \tag{2}$$

We use regular encoding. For each k, generate 2^{nR_k} codewords $\underline{v}_k(s_k)$, $s_k = 1, 2, \dots, 2^{nR_k}$, by choosing the letters $v_{k,i}(s_k)$, $i = 1, 2, \dots, n$, independently with distribution $p(v_k)$. Similarly, for every $\underline{v}_k(s_k)$ generate 2^{nR_k} codewords $\underline{x}_k(w_k, s_k)$, $w_k = 1, 2, \dots, 2^{nR_k}$, by choosing the letters $x_{k,i}(w_k, s_k)$ independently with distribution $p_{X_k|V_k}(\cdot | v_{k,i}(s_k))$ for all i. Finally, generate one length-n relay codeword $\underline{x}_r(s_1, s_2, \dots, s_K)$ for each tuple $(s_1, s_2, \dots, s_K)$ by choosing $x_{r,i}(s_1, s_2, \dots, s_K)$ independently with distribution $p_{X_r|V_1,V_2,\dots,V_k}(\cdot | v_{1,i}(s_1), \dots, v_{K,i}(s_K))$ for all i.

The above procedure is repeated $B+1$ times, once for each block, and the b^{th} codebook is used in block b, $b = 1, 2, \dots, B+1$. The encoding procedure of [4, Appendix A] proceeds as follows. We change this procedure in Sec. IV.

Regular Block Markov Encoding:

Encoder k parses w_k into B blocks $w_{k,1}, w_{k,2}, \dots, w_{k,B}$, each having nR_k bits, and transmits these messages over $B+1$ channel-symbol blocks as shown in Fig. 2. The relay sends the codeword $\underline{x}_r(s_{1,b}, s_{2,b}, \dots, s_{K,b})$ in block b where $s_{k,b}$ is the relay's estimate of $w_{k,b-1}$ from block $b-1$. We set $s_{k,1} = 1$ and $w_{k,B+1} = 1$ for all k. We thus have $N = n(B+1)$ and $B_{W_k} = nR_kB$ so the overall rate of user k is $R_{W_k} = R_k \cdot B / (B+1)$ which approaches R_k for large B.

III. DECODE-AND-FORWARD

A. Backward Decoding

Consider a 2-user MARC where the sources and the relay use the block-Markov encoding method described above. The relay decodes the messages reliably if (see Appendix I)

$$R_1 \le I(X_1; Y_r | X_2 V_1 V_2 X_r) \tag{3}$$
$$R_2 \le I(X_2; Y_r | X_1 V_1 V_2 X_r) \tag{4}$$
$$R_1 + R_2 \le I(X_1 X_2; Y_r | V_1 V_2 X_r). \tag{5}$$

The destination decodes the message blocks in reverse order using its channel-symbol blocks $\underline{y}_{d,B+1}, \underline{y}_{d,B}, \dots, \underline{y}_{d,2}$. The resulting destination rate bounds are (see Appendix I)

$$R_1 \le I(X_1 X_r; Y_d | X_2 V_2) \tag{6}$$
$$R_2 \le I(X_2 X_r; Y_d | X_1 V_1) \tag{7}$$
$$R_1 + R_2 \le I(X_1 X_2 X_r; Y_d). \tag{8}$$

Fig. 3 shows the rate region defined by (3)-(8). For a K-user MARC, these bounds generalize as follows.

Theorem 1: The capacity region of a K-user MARC includes the union of the set of rate tuples $(R_1, R_2, \dots, R_K)$ that satisfy, for all $\mathcal{S} \subseteq \mathcal{K}$,

$$R_{\mathcal{S}} \le \min \begin{pmatrix} I(X_{\mathcal{S}}; Y_r | X_{\mathcal{S}^c} V_{\mathcal{K}} X_r U), \\ I(X_{\mathcal{S}} X_r; Y_d | X_{\mathcal{S}^c} V_{\mathcal{S}^c} U) \end{pmatrix} \tag{9}$$

where the union is over all distributions that factor as

$$p(u) \cdot \left(\prod_{k=1}^{K} p(x_k, v_k | u) \right) \cdot p(x_r | v_{\mathcal{K}}, u) \cdot p(y_r, y_d | x_{\mathcal{K}}, x_r). \tag{10}$$

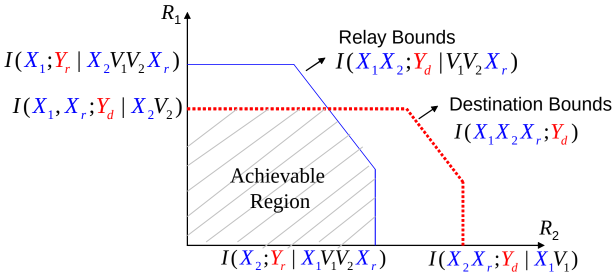


Fig. 3. Rate region achieved by DF for a two-user MARC.

Proof: See Appendix I. ■

Remark 1: The *time-sharing* random variable U ensures that the region of Theorem 1 is convex. For simplicity, we will develop the theory below for a constant U only.

Remark 2: The destination decodes the message blocks $w_{k,B}, w_{k,B-1}, \ldots, w_{k,1}$ with delays of $2, 3, \ldots, B+1$ channel-symbol blocks, respectively. Note that B must be large to ensure that the rate-loss factor $B/(B+1)$ due to block Markov encoding is close to 1.

B. Sliding-Window Decoding

Suppose the destination uses sliding-window decoding, i.e., the destination decodes the message pair $(w_{1,b}, w_{2,b})$ transmitted in block b by using $\underline{y}_{d,b}$ and $\underline{y}_{d,b+1}$. For example, in Fig. 2, the destination decodes $(w_{1,2}, w_{2,2})$ by using $\underline{y}_{d,2}$ and $\underline{y}_{d,3}$. Observe that $(w_{1,b+1}, w_{2,b+1})$ is not known while decoding $(w_{1,b}, w_{2,b})$. One can check that the bounds in (6)-(8) are replaced by

$$\begin{aligned} R_1 &\le I(X_1;Y_d|X_2V_1V_2X_r) + I(V_1X_r;Y_d|V_2) &(11)\\ R_2 &\le I(X_2;Y_d|X_1V_1V_2X_r) + I(V_2X_r;Y_d|V_1) &(12)\\ R_1 + R_2 &\le I(X_1X_2X_r;Y_d). &(13)\end{aligned}$$

The analysis used to obtain (11)-(13) is similar to that presented in Appendix II and is hence omitted. In brief, the term $I(X_1;Y_d|X_2V_1V_2X_r)$ in (11) results from $\underline{y}_{d,b}$ while the term $I(V_1X_r;Y_d|V_2)$ is due to $\underline{y}_{d,b+1}$. In fact, the same bounds result if one increases the sliding window length to decode messages from many past blocks, unless this window includes block $B+1$. The bounds (12) and (13) are obtained similarly.

We next compare (6)-(8) and (11)-(13). Obviously, the bounds (8) and (13) are the same. But consider the right-hand side of (6) that expands as

$$\begin{aligned} I(X_1X_r;Y_d|X_2V_2) &= I(X_1V_1X_r;Y_d|X_2V_2) &(14)\\ &= I(X_1;Y_d|X_2V_1V_2X_r)\\ &\quad + I(V_1X_r;Y_d|X_2V_2). &(15)\end{aligned}$$

where (14) follows from the Markov chain $(V_1,V_2) - (X_1,X_2,X_r) - Y_d$ and (15) from the chain rule for mutual information. We further have

$$\begin{aligned} I(V_1X_r;Y_d|X_2V_2) &= I(V_1X_r;X_2Y_d|V_2) &(16)\\ &\ge I(V_1X_r;Y_d|V_2) &(17)\end{aligned}$$

where (16) follows from the Markov chain $X_2 - V_2 - (V_1, X_r)$.

Note that (17) holds with equality if and only if

$$I(V_1X_r;X_2|V_2Y_d) = 0. \tag{18}$$

Comparing (15) and (17) with (11), we see that backward decoding is at least as good as sliding-window decoding.

We show by example that backward decoding can be strictly better than sliding-window decoding. Consider a MARC with $\{0,1\}$ inputs X_1, X_2, and X_r. Suppose we have

$$\begin{aligned} Y_r &= X_1 + X_2 &(19)\\ Y_d &= X_1 + X_r &(20)\end{aligned}$$

where we use integer addition. Any DF rate region must be in the capacity region of the user-to-relay multiaccess channel. This capacity region in bits per channel use is given by (see [16, p. 392])

$$R_1 \le 1, \quad R_2 \le 1, \quad R_1 + R_2 \le 3/2. \tag{21}$$

One can check that backward decoding achieves this region with independent and coin-tossing V_1, V_2, X_1, X_2, and X_r. However, for sliding-window decoding the bounds (3)-(5) and (11)-(13) are

$$\begin{aligned} R_1 &\le H(X_1|V_1) &(22)\\ R_2 &\le \min\left(H(X_2|V_2), H(X_1+X_r|V_1) - H(X_1|V_1)\right) &(23)\\ R_1 + R_2 &\le \min\left(H(X_1+X_2|V_1V_2), H(X_1+X_r\right). &(24)\end{aligned}$$

Suppose we desire $R_2 = 1$ so that (23) implies that X_2 is coin-tossing and independent of V_2. For such V_2 and X_2 the bound (24) implies

$$\begin{aligned} R_1 + R_2 &\le H(X_1+X_2|V_1V_2)\\ &= 1 + H(X_1|V_1)/2. &(25)\end{aligned}$$

We further have from (23) that

$$\begin{aligned} R_2 &\le H(X_1+X_r|V_1) - H(X_1|V_1)\\ &\le \log_2 3 - H(X_1|V_1). &(26)\end{aligned}$$

The combination of $R_2 = 1$, (25) and (26) gives

$$R_1 \le H(X_1|V_1)/2 \le (\log_2(3) - 1)/2 \approx 0.292. \tag{27}$$

The same bound results if we add a time-sharing random variable U to all the entropies in (22)-(24). Sliding-window decoding cannot therefore achieve the backward decoding corner point $(R_1, R_2) = (1/2, 1)$.

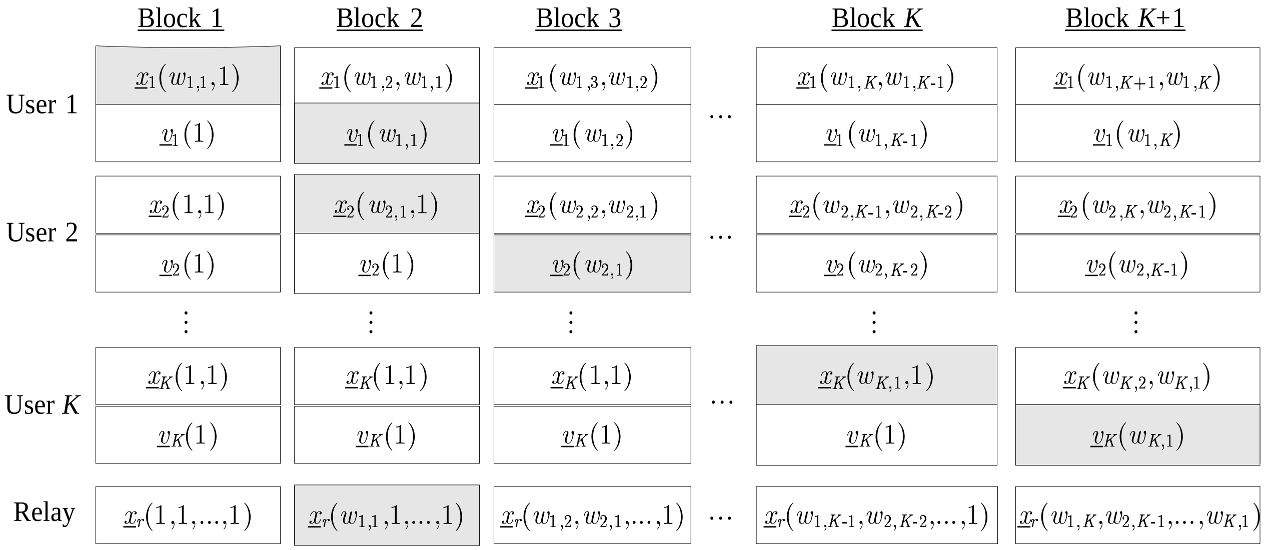
For $K > 2$, the bounds (11)-(13) generalize to

$$R_{\mathcal{S}} \le I(X_{\mathcal{S}};Y_d|X_{\mathcal{S}^c}V_{\mathcal{S}}X_r) + I(V_{\mathcal{S}}X_r;Y_d|V_{\mathcal{S}^c}) \tag{28}$$

for all $\mathcal{S} \subseteq \mathcal{K}$. One can show that the bounds in (28) are in general more restrictive than the corresponding destination bounds in (9) for all $\mathcal{S} \subset \mathcal{K}$.

IV. OFFSET ENCODING

To improve sliding-window decoding, we offset the message blocks from the K sources by one block per source. Let π denote a permutation (order) of the source indices. We let user $\pi(i)$ start transmitting in block i, i.e., we set $w_{\pi(i),b} = 1$ for $b < i$ and $b \ge B + i$. The resulting message-to-codeword



	Block 1	Block 2	Block 3		Block K	Block K+1
User 1	$\underline{x}_1(w_{1,1},1)$	$\underline{x}_1(w_{1,2},w_{1,1})$	$\underline{x}_1(w_{1,3},w_{1,2})$	$\cdots$	$\underline{x}_1(w_{1,K},w_{1,K\text{-}1})$	$\underline{x}_1(w_{1,K+1},w_{1,K})$
	$\underline{v}_1(1)$	$\underline{v}_1(w_{1,1})$	$\underline{v}_1(w_{1,2})$		$\underline{v}_1(w_{1,K\text{-}1})$	$\underline{v}_1(w_{1,K})$
User 2	$\underline{x}_2(1,1)$	$\underline{x}_2(w_{2,1},1)$	$\underline{x}_2(w_{2,2},w_{2,1})$	$\cdots$	$\underline{x}_2(w_{2,K\text{-}1},w_{2,K\text{-}2})$	$\underline{x}_2(w_{2,K},w_{2,K\text{-}1})$
	$\underline{v}_2(1)$	$\underline{v}_2(1)$	$\underline{v}_2(w_{2,1})$		$\underline{v}_2(w_{2,K\text{-}2})$	$\underline{v}_2(w_{2,K\text{-}1})$
	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
User K	$\underline{x}_K(1,1)$	$\underline{x}_K(1,1)$	$\underline{x}_K(1,1)$	$\cdots$	$\underline{x}_K(w_{K,1},1)$	$\underline{x}_K(w_{K,2},w_{K,1})$
	$\underline{v}_K(1)$	$\underline{v}_K(1)$	$\underline{v}_K(1)$		$\underline{v}_K(1)$	$\underline{v}_K(w_{K,1})$
Relay	$\underline{x}_r(1,1,\ldots,1)$	$\underline{x}_r(w_{1,1},1,\ldots,1)$	$\underline{x}_r(w_{1,2},w_{2,1},\ldots,1)$	$\cdots$	$\underline{x}_r(w_{1,K\text{-}1},w_{2,K\text{-}2},\ldots,1)$	$\underline{x}_r(w_{1,K},w_{2,K\text{-}1},\ldots,w_{K,1})$

Fig. 4. Offset encoding for a K-user MARC assuming the relay decodes correctly.

mappings with offset order $\pi = (1, 2, \ldots, K)$ are shown in Fig. 4. Observe that offset encoding uses $B + K$ channel-symbol blocks so the overall rate-loss factor is $B/(B + K)$.

The relay decodes at the end of each block as before. We thus require

$$R_{\mathcal{S}} \leq I(X_{\mathcal{S}}; Y_r | X_{\mathcal{S}^c} V_{\mathcal{K}} X_r) \tag{29}$$

for all $\mathcal{S} \subseteq \mathcal{K}$ as in (9). In block b, the relay sends the codeword $\underline{x}_r(s_{\mathcal{K},b})$, where $s_{\mathcal{K},b} = \{s_{k,b} : k \in \mathcal{K}\}$ and $s_{k,b}$ is the relay's estimate of $w_{k,b-k}$.

The destination uses a sliding window of length $K + 1$ to decode the message blocks with the same index b. Hence, the combined encoding and decoding delay for every message block is $K + 1$ channel-symbol blocks. We summarize the resulting rate bounds below and give the performance analysis in Appendices II and III.

A. Two Users with Joint Decoding

Consider $K = 2$ and suppose the offset order is $\pi = (1, 2)$. Suppose we decode $(w_{1,b}, w_{2,b})$ jointly by using $\underline{y}_{d,b}$, $\underline{y}_{d,b+1}$, and $\underline{y}_{d,b+2}$. The analysis in Appendix II shows that we can achieve (R_1, R_2) satisfying

$$R_1 \leq I(X_1 X_r; Y_d | X_2 V_2) \tag{30}$$

$$R_2 \leq I(X_2; Y_d | V_1 V_2 X_r) + I(V_2; Y_d) \tag{31}$$

$$R_1 + R_2 \leq I(X_1 X_2 X_r; Y_d). \tag{32}$$

Note that (30) is the same as (6) but (31) is different from (7). The difference arises because the destination does not know $w_{1,b+1}$ or $w_{1,b+2}$ when decoding $w_{2,b}$. We can show that (7) is in general larger than (31) by expanding (7) as

$$\begin{aligned} I(X_2 X_r; Y_d | X_1 V_1) &= I(X_2 V_2 X_r; Y_d | X_1 V_1) && (33)\\ &= I(X_2; Y_d | X_1 V_1 V_2 X_r) \\ &\quad + I(V_2 X_r; Y_d | X_1 V_1) && (34) \end{aligned}$$

where (33) follows from the Markov chain $(V_1, V_2) - (X_1, X_2, X_r) - Y_d$. But the first mutual information term in (34) satisfies

$$\begin{aligned} I(X_2; Y_d | X_1 V_1 V_2 X_r) &= I(X_2; X_1 Y_d | V_1 V_2 X_r) && (35)\\ &\geq I(X_2; Y_d | V_1 V_2 X_r) && (36) \end{aligned}$$

where (35) follows from the Markov chain $X_1 - (V_1, V_2, X_r) - X_2$. Similarly, the second mutual information term in (34) satisfies

$$\begin{aligned} I(V_2 X_r; Y_d | X_1 V_1) &\geq I(V_2; Y_d | X_1 V_1) && (37)\\ &= I(V_2; X_1 V_1 Y_d) && (38)\\ &\geq I(V_2; Y_d) && (39) \end{aligned}$$

where (38) follows from the independence of (X_1, V_1) and V_2. It thus seems that we do not achieve all of the backward decoding region. However, we next show that we can obtain the corner points of the backward decoding region.

Note that there are several types of corner points depending on whether the polytopes defined by the relay bounds (3)-(5) and the destination bounds (6)-(8) intersect. However, we need to consider only the latter bounds because the relay's rate bounds are the same for both non-offset and offset encoding. We therefore focus on the destination decoder, and when we write "achieve" below we are ignoring the relay bounds (3)-(5).

Consider the corner point labeled "$\pi = (1, 2)$" in Fig. 5. We can achieve this point (ignoring the relay bounds (3)-(5)) provided that the sum of (30) and (31) is less restrictive than (32). But (32) expands as

$$\begin{aligned} R_1 + R_2 &\leq I(X_1 X_2 X_r; Y_d) && (40)\\ &= I(X_1 X_2 V_2 X_r; Y_d) && (41)\\ &= I(X_1 X_r; Y_d | X_2 V_2) + I\left(X_2 V_2; Y_d\right). && (42) \end{aligned}$$

where (41) follows from the Markov chain V_2 −

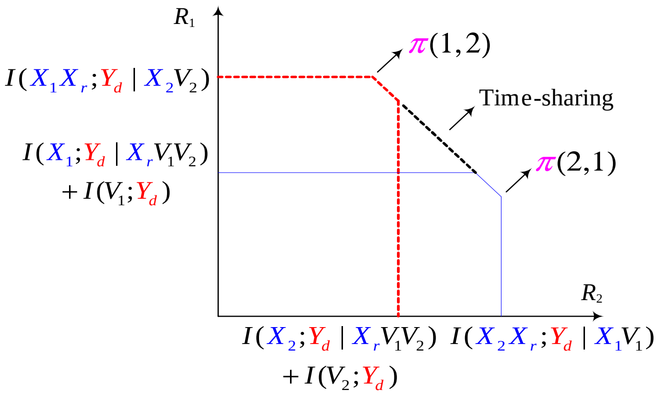


Fig. 5. Rate region with sliding-window decoding and offset encoding.

$(X_1, X_2, X_r) - Y_d$. We further have

$$\begin{aligned} I(X_2V_2;Y_d) &= I(X_2;Y_d|V_2) + I(V_2;Y_d) && (43)\\ &\le I(X_2;V_1X_rY_d|V_2) + I(V_2;Y_d) && (44)\\ &= I(X_2;Y_d|V_1V_2X_r) + I(V_2;Y_d) && (45) \end{aligned}$$

where (45) follows from the Markov chain $X_2 - V_2 - (V_1, X_r)$. Thus, we achieve the corner point under consideration. For the offset order $\pi = (2,1)$, we similarly obtain the corner point labeled "$\pi = (2,1)$" in Fig. 5. The non-corner points of the backward decoding region are achieved by time-sharing.

B. K Users with Successive Decoding

We wish to show that offset encoding recovers the backward decoding corner points for $K > 2$. However, the generalization of (11)-(13) is unwieldy and gives limited insight. Instead, we use successive decoding inside the sliding window to obtain the backward decoding corner points.

Suppose the offset order is $\pi = (1, 2, \ldots, K)$ as in Fig. 4. Consider the window with the channel-symbol blocks $\underline{y}_{d,1}, \underline{y}_{d,2}, \cdots, \underline{y}_{d,K+1}$. In this window, the destination successively decodes $w_{K,1}, w_{K-1,1}, \ldots, w_{1,1}$ by assuming that its past decoding steps were successful. In Appendix III, we show that one can approach the rate point $\underline{R} = (R_1, R_2, \ldots, R_K)$ with

$$R_k = \begin{cases} I(X_1X_r;Y_d|X_{[2,K]}V_{[2,K]}) & k = 1\\ I(X_kV_k;Y_d|X_{[k+1,K]}V_{[k+1,K]}) & k > 1. \end{cases} \quad (46)$$

The codewords contributing to these rates are shown as shaded blocks in Fig. 4. Let $\mathcal{S}(L) = \{1, 2, \ldots, L\}$ where $1 \le L \le K$. One can check that $\underline{R}$ satisfies

$$R_{\mathcal{S}(L)} = I(X_{[1,L]}X_r;Y_d|X_{[L+1,K]}V_{[L+1,K]}) \quad (47)$$

which means that $R_{\mathcal{S}(L)}$ satisfies the destination bound

$$R_{\mathcal{S}} \le I(X_{\mathcal{S}}X_r;Y_d|X_{\mathcal{S}^c}V_{\mathcal{S}^c}) \quad (48)$$

in (9) with equality for $\mathcal{S} = \mathcal{S}(L)$.

Now consider any $\underline{R}'$, $\underline{R}''$, and α for which $0 < \alpha < 1$ and $\underline{R} = \alpha\underline{R}' + (1-\alpha)\underline{R}''$. If $\underline{R}'$ and $\underline{R}''$ are points that satisfy (48) for all $\mathcal{S} \subseteq \mathcal{K}$ then (47) suffices to establish that $\underline{R} = \underline{R}' = \underline{R}''$. The point $\underline{R}$ is thus a corner point of the polytope defined by the destination bounds in (48) [17, p. 121]. We can achieve the other corner points by changing the offset order. Finally, we achieve the non-corner points by time-sharing.

V. CONCLUSIONS

We presented an offset encoding technique for DF that improves the rate region of sliding-window decoding. The technique achieves the corner points of the backward decoding rate region but avoids the excessive delay associated with backward decoding. Offset encoding will clearly apply to other multi-terminal problems [13], [18]–[20].

APPENDIX I
BACKWARD DECODING ANALYSIS

We derive the DF rate bounds for discrete memoryless MARCs, $K = 2$, and backward decoding. The random code construction and the encoding are described in Section II-B and we use typical sequence decoders. Define the set of typical sequences of length n with respect to ϵ and $P_{X,Y}(\cdot)$ as

$$T_\epsilon^{(n)}(X,Y) = \left\{ (\underline{x},\underline{y}) : \left| \frac{n(a,b|\underline{x},\underline{y})}{n} - P_{X,Y}(a,b) \right| \le \frac{\epsilon}{|\mathcal{X}|\cdot|\mathcal{Y}|} \right\} \quad (49)$$

where $n(a,b|\underline{x},\underline{y})$ is the number of times the pair (a,b) occurs in the sequence of pairs $(x_1,y_1), (x_2,y_2), \ldots (x_n,y_n)$. $\mathcal{X}$ and $\mathcal{Y}$ are the alphabets of X and Y with cardinalities $|\mathcal{X}|$ and $|\mathcal{Y}|$, respectively.

Decoding:

1) *At the relay*: The relay decodes $(w_{1,b}, w_{2,b})$ in block b, $b = 1, 2, \ldots, B$, by using $\underline{y}_{r,b}$ and by assuming that its message estimates in the previous blocks are correct. More precisely, the relay decodes by finding a $(\tilde{w}_{1,b}, \tilde{w}_{2,b})$ such that

$$\begin{aligned} &(\underline{x}_1(\tilde{w}_{1,b}, w_{1,b-1}), \underline{x}_2(\tilde{w}_{2,b}, w_{2,b-1}), \underline{v}_1(w_{1,b-1}),\\ &\quad \underline{v}_2(w_{2,b-1}), \underline{x}_r(w_{1,b-1}, w_{2,b-2}), \underline{y}_{r,b}) \in\\ &\qquad T_\epsilon^{(n)}(X_1, X_2, V_1, V_2, X_r, Y_r). \quad (50) \end{aligned}$$

We assume that the correct codewords are identified as being typical since this is a high probability event for large n. With this assumption the relay makes an error only if it identifies a $(\tilde{w}_{1,b}, \tilde{w}_{2,b}) \ne (w_{1,b}, w_{2,b})$ that satisfies (50). This error event can be further split into three disjoint error events. The first error event has a $\tilde{w}_{1,b} \ne w_{1,b}$ and $\tilde{w}_{2,b} = w_{2,b}$ satisfying (50). Using [4, Lemma 1] and the union bound, the probability of this event is at most

$$2^{n(R_1 - I(X_1;Y_r|X_2V_1V_2X_r)+6\epsilon)}. \quad (51)$$

Thus, for reliable decoding we set

$$R_1 < I(X_1;Y_r|X_2V_1V_2X_r). \quad (52)$$

The second error event has $\tilde{w}_{1,b} = w_{1,b}$ and a $\tilde{w}_{2,b} \neq w_{2,b}$ satisfying (50). By symmetry to (52), we set

$$R_2 < I(X_2; Y_r | X_1 V_1 V_2 X_r). \tag{53}$$

The third error event has a $\tilde{w}_{1,b} \neq w_{1,b}$ and a $\tilde{w}_{2,b} \neq w_{2,b}$ satisfying (50). We again use [4, Lemma 1] to bound the probability of this event by

$$2^{n(R_1+R_2-I(X_1X_2;Y_r|V_1V_2X_r)+6\epsilon)}. \tag{54}$$

Reliable decoding thus requires

$$R_1 + R_2 < I(X_1 X_2; Y_r | V_1 V_2 X_r). \tag{55}$$

2) *At the destination*: The destination collects all of its $B+1$ output blocks. Starting from the last block, the destination decodes $(w_{1,b}, w_{2,b})$, $b = B, B-1, \ldots, 1$ by using $\underline{y}_{d,b+1}$ and by assuming that its previously decoded message estimates are correct. More precisely, the destination decodes by finding a $(\tilde{w}_{1,b}, \tilde{w}_{2,b})$ such that

$$\begin{aligned}&(\underline{x}_1(w_{1,b+1}, \tilde{w}_{1,b}), \underline{x}_2(w_{2,b+1}, \tilde{w}_{2,b}), \underline{v}_1(\tilde{w}_{1,b}),\\ &\quad \underline{v}_2(\tilde{w}_{2,b}), \underline{x}_r(\tilde{w}_{1,b}, \tilde{w}_{2,b}), \underline{y}_{d,b+1})\\ &\quad\quad \in T_\epsilon^{(n)}(X_1, X_2, V_1, V_2, X_r, Y_d).\end{aligned} \tag{56}$$

As before we assume that the correct codewords are identified as being typical. Again, three kinds of error events can occur in decoding $(w_{1,b}, w_{2,b})$. Using [4, Lemma 1] and the union bound, we follow the same decoding steps as for the relay decoder to show that

$$R_1 < I(X_1 X_r; Y_d | X_2 V_2) \tag{57}$$

$$R_2 < I(X_2 X_r; Y_d | X_1 V_1) \tag{58}$$

$$R_1 + R_2 < I(X_1 X_r X_2; Y_d) \tag{59}$$

ensures reliable communications.

Combining (52), (53), (55), and (57)-(59), we have the bounds (3)-(8). The analysis generalizes in a straightforward way for entropy-typical sequences [16, p. 51], the addition of a time-sharing random variable U [16, p. 396], and $K > 2$.

APPENDIX II
SLIDING-WINDOW JOINT DECODING ANALYSIS

We derive the DF rate bounds for $K = 2$, offset encoding, and sliding-window decoding. Without loss of generality, we consider the offset order $\pi = (1, 2)$. Section II-B describes the random code construction.

Encoding: Consider block b.

1) Source 1 transmits $\underline{x}_k(w_{1,b}, w_{1,b-1})$ while source 2 transmits $\underline{x}_2(w_{2,b-1}, w_{2,b-2})$ where $w_{2,-1}$, $w_{2,0}$, $w_{2,1}$, $w_{1,0}$, $w_{1,B+1}$, $w_{1,B+2}$, and $w_{2,B+2}$ are set to 1.
2) The relay transmits $\underline{x}_r(s_{1,b}, s_{2,b})$ where $(s_{1,b}, s_{2,b})$ is the message pair decoded at the relay in block $(b-1)$.

Decoding:

1) *At the relay*: The relay decoder error analysis is the same as that described in Appendix I up to changes in the message indices. We therefore have the same rate bounds (52), (53), and (55).
2) *At the destination*: The destination decodes $(w_{1,b}, w_{2,b})$ by using $\underline{y}_{d,b}$, $\underline{y}_{d,b+1}$, and $\underline{y}_{d,b+2}$ and by assuming that no errors were made up to block b. More precisely, the destination decodes by finding a $(\tilde{w}_{1,b}, \tilde{w}_{2,b})$ such that three events occur:

$$\begin{aligned}\mathcal{E}_1 : \quad &(\underline{v}_1(w_{1,b-1}), \underline{v}_2(w_{2,b-2}), \underline{x}_1(\tilde{w}_{1,b}, w_{1,b-1}),\\ &\underline{x}_2(w_{2,b-1}, w_{2,b-2}), \underline{x}_r(w_{1,b-1}, w_{2,b-2}), \underline{y}_{d,b})\\ &\in T_\epsilon^{(n)}(V_1, V_2, X_1, X_2, X_r, Y_d)\end{aligned} \tag{60}$$

$$\begin{aligned}\mathcal{E}_2 : \quad &(\underline{v}_1(\tilde{w}_{1,b}), \underline{v}_2(w_{2,b-1}), \underline{x}_2(\tilde{w}_{2,b}, w_{2,b-1}),\\ &\underline{x}_r(\tilde{w}_{1,b}, w_{2,b-1}), \underline{y}_{d,b+1})\\ &\in T_\epsilon^{(n)}(V_1, V_2, X_2, X_r, Y_d)\end{aligned} \tag{61}$$

$$\mathcal{E}_3 : \quad (\underline{v}_2(\tilde{w}_{2,b}), \underline{y}_{d,b+2}) \in T_\epsilon^{(n)}(V_2, Y_d). \tag{62}$$

Note that the codebooks in different blocks are generated independently, so the above three error events are independent. As before, we consider three disjoint error events that can occur in decoding $(w_{1,b}, w_{2,b})$. The first event has a $\tilde{w}_{1,b} \neq w_{1,b}$ and $\tilde{w}_{2,b} = w_{2,b}$ satisfying (60)-(62). We upper bound the probability of this error event using [4, Lemma 1] and the union bound as

$$\begin{aligned}&\sum_{\tilde{w}_{1,b} \neq w_{1,b}} \Pr(\mathcal{E}_1 \cap \mathcal{E}_2 \cap \mathcal{E}_3)\\ &= \sum_{\tilde{w}_{1,b} \neq w_{1,b}} \Pr(\mathcal{E}_1) \cdot \Pr(\mathcal{E}_2) \cdot \Pr(\mathcal{E}_3) &(63)\\ &\leq 2^{n(R_1 - I(X_1;Y_d|X_2V_1V_2X_r) - I(V_1X_r;Y_d|X_2V_2) + 12\epsilon)} &(64)\end{aligned}$$

where we used $\Pr(\mathcal{E}_3) \leq 1$. Thus, we set

$$R_1 < I(X_1 X_r; Y_d | X_2 V_2). \tag{65}$$

Consider next the case where $\tilde{w}_{1,b} = w_{1,b}$ but $\tilde{w}_{2,b} \neq w_{2,b}$. The expression (63) with the summation over $\tilde{w}_{2,b} \neq w_{2,b}$ instead of $\tilde{w}_{1,b} \neq w_{1,b}$ is upper bounded as

$$2^{n(R_2 - I(X_2;Y_d|V_1V_2X_r) - I(V_2;Y_d) + 12\epsilon)} \tag{66}$$

where we used $\Pr(\mathcal{E}_1) \leq 1$. We thus require

$$R_2 < I(X_2; Y_d | V_1 V_2 X_r) + I(V_2; Y_d). \tag{67}$$

Finally, consider the case $\tilde{w}_{1,b} \neq w_{1,b}$ and $\tilde{w}_{2,b} \neq w_{2,b}$. The expression (63) except with the summation over both $\tilde{w}_{1,b} \neq w_{1,b}$ and $\tilde{w}_{2,b} \neq w_{2,b}$ is upper bounded as

$$2^{n(R_1+R_2-I(X_1X_2X_r;Y_d)+18\epsilon)}. \tag{68}$$

For reliable decoding, we thus require

$$R_1 + R_2 < I(X_1 X_2 X_r; Y_d). \tag{69}$$

Combining (65), (67), and (69), we obtain (30)-(32). Again, the analysis generalizes in a straightforward way for entropy-typical sequences, the addition of a time-sharing random variable U, and $K > 2$.

APPENDIX III

SLIDING-WINDOW SUCCESSIVE DECODING ANALYSIS

We derive DF rate bounds for $K \geq 2$, offset encoding, and sliding-window decoding. However, the destination decoder now performs successive rather than joint decoding. Without loss of generality, we consider the offset order $\pi = (1, 2, \dots, K)$. Section II-B describes the random code construction, and the encoding and relay decoding are the same as in Appendix II.

Decoding at the destination: Consider the window with the channel-symbol blocks $\underline{y}_{d,1}, \underline{y}_{d,2}, \dots, \underline{y}_{d,K+1}$. As explained in Section IV-B, the destination successively decodes in the reverse order $w_{K,1}, w_{K-1,1}, \dots, w_{1,1}$ (see the shaded blocks in Fig. 4). The destination further assumes that its past decoding steps were successful, and we perform our analysis with the same assumption. For $k = K, K-1, \dots, 2$, the destination finds a $\tilde{w}_{k,1}$ such that two events occur:

$$\begin{aligned} \mathcal{E}_1: \quad & (\underline{v}_k(\tilde{w}_{k,1}), \underline{v}_{[k+1,K]}(1), \underline{x}_{[k+1,K]}(1,1), \underline{y}_{d,k+1}) \\ & \in T_\epsilon^{(n)}(V_k, V_{[k+1,K]}, X_{[k+1,K]}, Y_d) \end{aligned} \tag{70}$$

$$\begin{aligned} \mathcal{E}_2: \quad & (\underline{x}_k(\tilde{w}_{k,1}), \underline{v}_{[k,K]}(1), \underline{x}_{[k+1,K]}(1,1), \underline{y}_{d,k}) \\ & \in T_\epsilon^{(n)}(X_k, V_{[k,K]}, X_{[k+1,K]}, Y_d) \end{aligned} \tag{71}$$

where $\underline{v}_{[i,j]}(1) = \{v_i(1), v_{i+1}(1), \dots, v_j(1)\}$ and similarly for $\underline{x}_{[i,j]}(1)$ and $\underline{v}_{\mathcal{K}}(1)$ below. As before, the above events are independent and we assume that the correct codewords are identified as being typical. The destination thus makes an error only if it identifies a $\tilde{w}_{k,1} \neq w_{k,1}$ that satisfies both (70) and (71). We upper bound the probability of this event using [4, Lemma 1] as

$$\begin{aligned} \sum_{\tilde{w}_{k,1} \neq w_{k,1}} & \Pr(\mathcal{E}_1) \cdot \Pr(\mathcal{E}_2) \leq \\ & 2^{n\left(R_k - I(X_k V_k; Y_d | X_{[k+1,K]} V_{[k+1,K]}) + 12\epsilon\right)}. \end{aligned} \tag{72}$$

For $k > 1$, we therefore require

$$R_k < I(X_k V_k; Y_d | X_{[k+1,K]} V_{[k+1,K]}). \tag{73}$$

For $k = 1$, we add $\underline{x}_r(\cdot)$ to (70) and (71) as follows:

$$\begin{aligned} \mathcal{E}_1: \quad & (\underline{v}_1(\tilde{w}_{1,1}), \underline{v}_{[2,K]}(1), \underline{x}_{[2,K]}(1,1), \\ & \underline{x}_r(\tilde{w}_{1,1}, 1, \dots, 1), \underline{y}_{d,2}) \\ & \in T_\epsilon^{(n)}(V_1, V_{[2,K]}, X_{[2,K]}, X_r, Y_d) \end{aligned} \tag{74}$$

$$\begin{aligned} \mathcal{E}_2: \quad & (\underline{x}_1(\tilde{w}_{1,1}, 1), \underline{v}_{\mathcal{K}}(1), \underline{x}_{[2,K]}(1,1), \underline{x}_r(1,1,\dots,1), \underline{y}_{d,1}) \\ & \in T_\epsilon^{(n)}(X_1, V_{\mathcal{K}}, X_{[2,K]}, X_r, Y_d). \end{aligned} \tag{75}$$

The resulting bound is

$$R_1 < I(X_1 X_r; Y_d | X_{[2,K]} V_{[2,K]}). \tag{76}$$

One can check that the above analysis generalizes to $b > 1$.

ACKNOWLEDGMENTS

The authors would like to thank Prof. Liang-Liang Xie of the University of Waterloo and the anonymous reviewer for their useful comments that helped improve the paper.

REFERENCES

[1] G. Kramer and A. J. van Wijngaarden, “On the white Gaussian multiple-acess relay channel,” in *Proc. 2000 IEEE Int. Symp. Inform. Theory*, Sorrento, Italy, June 2000, p. 40.

[2] E. C. van der Meulen, “Three-terminal communication channels,” *Adv. Applied Probability*, vol. 3, pp. 120–154, 1971.

[3] T. Cover and A. El Gamal, “Capacity theorems for the relay channel,” *IEEE Trans. Inform. Theory*, vol. 25, no. 5, pp. 572–584, Sept. 1979.

[4] G. Kramer, M. Gastpar, and P. Gupta, “Cooperative strategies and capacity theorems for relay networks,” *IEEE Trans. Inform. Theory*, vol. 51, no. 9, pp. 3027–3063, Sept. 2005.

[5] L. Sankaranarayanan, G. Kramer, and N. B. Mandayam, “Capacity theorems for the multiple-access relay channel,” in *Proc. 42nd Annual Allerton Conf. on Commun., Control, and Computing*, Monticello, IL, Sept. 2004.

[6] A. B. Carleial, “Multiple-access channels with different generalized feedback signals,” *IEEE Trans. Inform. Theory*, vol. 28, no. 6, pp. 841–850, Nov. 1982.

[7] F. M. J. Willems, “Informationtheoretical Results for the Discrete Memoryless Multiple Access Channel,” Ph.D. dissertation, Doctor in de Wetenschappen Proefschrift, Katholieke Universiteit Leuven, Leuven, Belgium, Oct. 1982.

[8] M. R. Aref, “Information flow in relay networks,” Ph.D. dissertation, Stanford University, Stanford, CA, Oct. 1980.

[9] P. Gupta and P. R. Kumar, “Towards an information theory of large networks: An achievable rate region,” *IEEE Trans. Inform. Theory*, vol. 49, no. 8, pp. 1877–1894, Aug. 2003.

[10] L.-L. Xie and P. R. Kumar, “An achievable rate for the multiple-level relay channel,” *IEEE Trans. Inform. Theory*, vol. 51, no. 4, pp. 1348–1358, Apr. 2005.

[11] M. Sikora, J. N. Laneman, M. Haenggi, D. J. Costello, and T. E. Fuja, “Bandwidth- and power-efficient routing in linear wireless networks,” *IEEE Trans. Inform. Theory*, vol. 52, no. 6, pp. 2624–2633, June 2006.

[12] F. K. C.-M. Zeng and A. Buzo, “Achievability proofs of some multiuser channel coding theorems using backward decoding,” *IEEE Trans. Inform. Theory*, vol. 35, no. 6, pp. 1160–1165, Nov. 1989.

[13] J. N. Laneman and G. Kramer, “Window decoding for the multiaccess channel with generalized feedback,” in *Proc. 2004 IEEE Int. Symp. Inform. Theory*, Chicago, IL, June 27 - July 2, 2004.

[14] B. Rimoldi and R. Urbanke, “A rate-splitting approach to the gaussian multiple access channel,” *IEEE Trans. Inform. Theory*, vol. 42, no. 2, pp. 348–363, Mar. 1996.

[15] R. G. Gallager, *Information Theory and Reliable Communication*. John Wiley, 1968.

[16] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[17] M. Padberg, *Linear Optimizations and Extensions*. Springer, 1991.

[18] L. Sankaranarayanan, G. Kramer, and N. B. Mandayam, “Cooperation vs. hierarchy: An information-theoretic comparison,” in *Proc. Int. Symp. Inf. Theory*, Adelaide, Australia, Sept. 2005.

[19] L.-L. Xie and P. R. Kumar, “A multi-relay scheme for the multi-source multi-cast network,” in *Proc. 2006 IEEE Intl. Symp. Inform. Theory*, Seattle, WA, July 2006.

[20] ——, “Multi-source, multi-destination, multi-relay wireless networks,” Aug. 2006, submitted to the *IEEE Trans. Inform. Theory*.