Since the two rate regions are now symmetric (see Fig. 4), the second constraint becomes equality. This expresses λ as a function of a, and plugging into (24), we get the desired result. □

REFERENCES

[1] M. J. Lee, J. Zheng, Y.-B. Ko, and D. M. Shrestha, "Emerging standards for wireless mesh technology," *IEEE Wireless Commun.*, vol. 13, no. 2, pp. 56–63, Apr. 2006.

[2] IEEE 802.16's Relay Task Group [Online]. Available: http://www.ieee802.org/16/relay

[3] E. van der Meulen, "Three-terminal communication channels," *Adv. Appl. Probab.*, vol. 3, pp. 120–154, 1971.

[4] T. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sep. 1979.

[5] A. El Gamal and S. Zahedi, "Minimum energy communication over a relay channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Ykohama, Japan, 2003, p. 344.

[6] S. Zahedi, M. Mohseni, and A. El Gamal, "On the capacity of AWGN relay channels with linear relaying functions," in *Proc. IEEE Int. Symp. Inf. Theory*, Chicago, 2004, p. 399.

[7] Y. Liang and V. V. Veeravalli, "Gaussian orthogonal relay channels: Optimal resource allocation and capacity," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3284–3289, Sep. 2005.

[8] A. Høst-Madsen and J. Zhang, "Capacity bounds and power allocation for wireless relay channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2020–2040, Jun. 2005.

[9] M. A. Khojastepour, A. Sabharwal, and B. Aazhang, "Bounds on achievable rates for general multi-terminal networks with practical constraints," in *Proc. Inf. Process. Sens. Netw.: Second Int. Work.*, Palo Alto, CA, Apr. 2003.

[10] A. Høst-Madsen, "Capacity bounds for cooperative diversity," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1522–1544, Apr. 2006.

[11] M. Katz and S. Shamai (Shitz), "Relaying protocols for two colocated users," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2329–2344, Jun. 2006.

[12] A. Steiner and S. Shamai (Shitz), "Single-user broadcasting protocols over a two-hop relay fading channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 4821–4838, Nov. 2006.

[13] J. N. Laneman, D. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[14] K. Azarian, H. El Gamal, and P. Schniter, "On the achievable diversity – Multiplexing tradeoff in half-duplex cooperative channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4152–4172, Dec. 2005.

[15] B. Schein and R. Gallager, "The Gaussian parallel relay network," in *Proc. IEEE Int. Symp. Inf. Theory*, 2000.

[16] B. Schein, "Distributed coordination in network information theory," Ph.D., MIT, Cambridge, MA, 2001.

[17] M. Gastpar, G. Kramer, and P. Gupta, "The multiple-relay channel: Coding and antenna-clustering capacity," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT'02)*, Lausanne, Switzerland, Jun. 2002, p. 137.

[18] E. M. Yeh and R. A. Berry, "Throughput optimal control of cooperative relay networks," in *Proc. 2005 Int. Symp. Inf. Theory*, Adelaide, Australia, Sep. 4–9, 2005, pp. 1206–1210.

[19] L. Xie and P. Kumar, "A network information theory for wireless communication: Scaling laws and optimal operation," *IEEE Trans. Inf. Theory*, vol. 50, no. 5, pp. 748–767, 2004.

[20] A. Reznik, S. Kulkarni, and S. Verdu, "Degraded Gaussian multirelay channel: Capacity and optimal power allocation," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3037–3046, 2004.

[21] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.

[22] M. Gastpar and M. Vetterli, "On the capacity of large Gaussian relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 765–779, Mar. 2005.

[23] A. Sanderovich, S. Shamai (Shitz), Y. Steinberg, and G. Kramer, "Communication via decentralized processing," in *Proc. Int. Symp. Inf. Theory (ISIT)*, Adelaide, Australia, Sep. 4–9, 2005, pp. 1201–1205.

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

[25] D. Slepian and J. K. Wolf, "A coding theorem for multiple access channels with correlated sources," *Bell Syst. Tech. J.*, vol. 52, no. 7, pp. 1037–1076, 1973.

[26] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Telecommun.*, vol. 10, pp. 585–595, Nov./Dec. 1999.

[27] V. Prabhakaran, D. Tse, and K. Ramchandran, "Rate region of the quadratic Gaussian CEO problem," in *Proc. ISIT 2004*, Jun.–Jul. 27–2, 2004, p. 117.

[28] 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.

Offset Encoding for Multiple-Access Relay Channels

Lalitha Sankar, *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 multiple-access relay channels. The technique offsets user transmissions by one block per user and achieves the corner points of the destination's backward decoding rate regions but with a smaller delay. As a result, one achieves boundary points of the best known decode-and-forward rate regions with a smaller delay than with backward decoding.**

***Index Terms*—Cooperative systems, encoding, multiple-access communication, relaying.**

I. INTRODUCTION

The multiple-access 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];

Manuscript received August 14, 2006; revised April 13, 2007. The work of L. Sankar (previously Sankaranarayanan) and N. B. Mandayam was supported in part by the National Science Foundation under Grant ITR-0205362. The work of G. Kramer was supported in part by the Board of Trustees of the University of Illinois Subaward 04-217 under the National Science Foundation Grant CCR-0325673. The material in this correspondence was presented in part at the 42nd Annual Allerton Conference on Communications, Control, and Computing, Monticello, IL, September 2004.

L. Sankar is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: lalitha@princeton.edu).

N. B. Mandayam is with the Wireless Information Network Laboratory (WINLAB), Rutgers University, North Brunswick, NJ 08902 USA (e-mail: lalitha@winlab.rutgers.edu; narayan@winlab.rutgers.edu).

G. Kramer is with Bell Laboratories, Alcatel-Lucent, Murray Hill, NJ 07974-0636 USA (e-mail: gkr@bell-labs.com).

Communicated by A. El Gamal, Guest Editor for the Special Issue on Relaying and Cooperation.

Color versions of Figures 1, 3, and 5 in this correspondence are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2007.904791

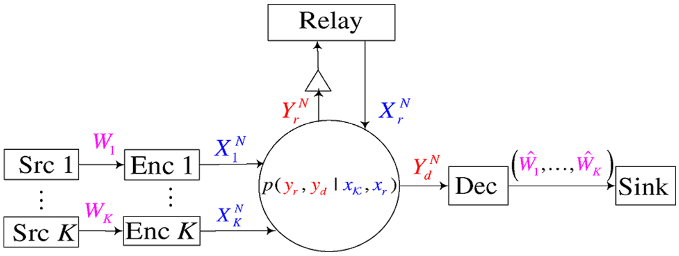

Fig. 1. A K-user MARC.

- *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* applications.

To compare the methods, suppose the destination uses 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, where we measure delay from the start of the block to the time the block is decoded. Our main contribution is an *offset encoding* technique for sliding-window decoding that recovers the corner points of the destination's backward decoding rate regions 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 regions, we use a combination of offset encoding, no-offset encoding, and/or time sharing between different offset encoding methods. Note, however, that time sharing increases decoding delay; rate-splitting methods might perhaps avoid this delay [14], [15].

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 correspondence.

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 messages as $(\hat{W}_1, \hat{W}_2, \ldots, \hat{W}_K)$. We write $\mathcal{K} = \{1,2,\ldots,K\}$, $X_{\mathcal{S}} = \{X_k : k \in \mathcal{S}\}$ 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

$$p_{Y_r,Y_d|X_{\mathcal{K}},X_r}(y_r, y_d | x_{\mathcal{K}}, x_r). \tag{1}$$

The capacity region $\mathcal{C}_{\text{MARC}}$ is the closure of the set of rate tuples $(R_{W_1}, R_{W_2}, \ldots, R_{W_K})$ 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 $R_{\mathcal{S}} = \sum_{k\in\mathcal{S}} R_k$, $[m,n] = \{m, m+1, \ldots, n\}$, 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 [16], [17] and take all logarithms to the base 2 so that our rate units are bits. We write random variables (e.g., W_k) with upper case letters and their realizations (e.g., w_k) with the corresponding lower case letters. We drop subscripts on probability distributions if the arguments are lower case versions of the random variables, e.g., we write (1) as $p(y_r, y_d | x_{\mathcal{K}}, x_r)$.

We assume familiarity of the reader with basic notions of backward decoding and joint decoding as described in [4], [6], [7], [10], [12].

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, $k = 1,2,\ldots,K$, to help the sources cooperate with the relay. The V_k have finite alphabets.

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 $\underline{v}_k(s_k)$, $s_k = 1,2,\ldots,2^{nR_k}$, by choosing the letters $v_{k,i}(s_k)$, $i = 1,2,\ldots,n$, independently with distribution $p(v_k)$. Similarly, for every $v_k(s_k)$ generate 2^{nR_k} codewords $x_k(w_k, s_k)$, $w_k = 1,2,\ldots,2^{nR_k}$, 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 $\underline{x}_r(s_1, s_2, \ldots, s_K)$ for each tuple $(s_1, s_2, \ldots, s_K)$ by choosing $x_{r,i}(s_1, s_2, \ldots, s_K)$ independently with distribution $p_{X_r|V_1,V_2,\ldots,V_k}(\cdot|v_{1,i}(s_1), \ldots, v_{K,i}(s_K))$ for all i.

The above code construction procedure is repeated $B+1$ times, once for each block, and the bth codebook is used in block b, $b = 1,2,\ldots,B+1$. Note that the codebooks are independent across blocks; this fact simplifies the error analysis [6], [10]. The encoding procedure of [4, Appendix A] proceeds as follows. We change this procedure in Section IV.

Regular Block Markov Encoding:

Encoder k parses w_k into B blocks $w_{k,1}, w_{k,2}, \ldots, w_{k,B}$, each having nR_k bits, and transmits these messages over $B+1$ channel-symbol blocks as shown in Fig. 2 for $K = 2$. More generally, user k transmits $\underline{x}_k(w_{k,b}, w_{k,b-1})$ in block b where $w_{k,0} = w_{k,B+1} = 1$ for all k. The relay sends the codeword $x_r(s_{1,b}, s_{2,b}, \ldots, s_{K,b})$ in block b where $s_{k,b}$ is the relay's estimate of $w_{k,b-1}$ from block $b-1$. We set $s_{k,1} = 1$ for all k. We thus have

	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.

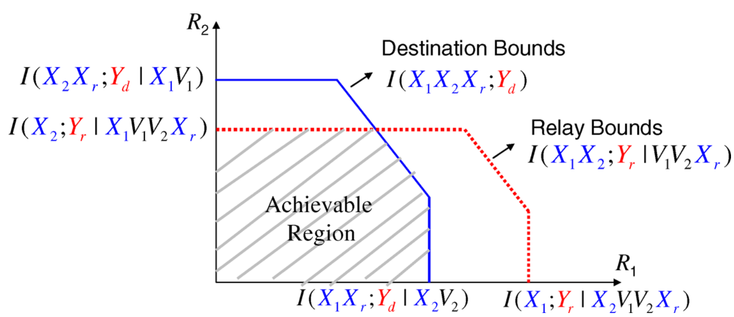

Fig. 3. Example of a rate region achieved by DF and backward decoding for a two-user MARC.

$N = n(B+1)$ 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 two-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)

$$R_1 \le I(X_1;Y_r|X_2V_1V_2X_r) \quad (3)$$
$$R_2 \le I(X_2;Y_r|X_1V_1V_2X_r) \quad (4)$$
$$R_1+R_2 \le I(X_1X_2;Y_r|V_1V_2X_r). \quad (5)$$

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)

$$R_1 \le I(X_1X_r;Y_d|X_2V_2) \quad (6)$$
$$R_2 \le I(X_2X_r;Y_d|X_1V_1) \quad (7)$$
$$R_1+R_2 \le I(X_1X_2X_r;Y_d). \quad (8)$$

Fig. 3 shows an example of 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 $\mathcal{S} \subseteq \mathcal{K}$

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

where the union is over all distributions that factor as

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

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 $w_{k,B}, w_{k,B-1}, \dots, w_{k,1}$ with delays of $2, 3, \dots, B+1$ 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 $y_{d,b}$ and $y_{d,b+1}$. 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

$$R_1 \le I(X_1;Y_d|X_2V_1V_2X_r) + I(V_1X_r;Y_d|V_2) \quad (11)$$
$$R_2 \le I(X_2;Y_d|X_1V_1V_2X_r) + I(V_2X_r;Y_d|V_1) \quad (12)$$
$$R_1+R_2 \le I(X_1X_2X_r;Y_d). \quad (13)$$

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 $y_{d,b}$ while the term $I(V_1X_r;Y_d|V_2)$ is due to $y_{d,b+1}$. 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

$$I(X_1X_r;Y_d|X_2V_2) = I(X_1V_1X_r;Y_d|X_2V_2) \quad (14)$$
$$= I(X_1;Y_d|X_2V_1V_2X_r) + I(V_1X_r;Y_d|X_2V_2). \quad (15)$$

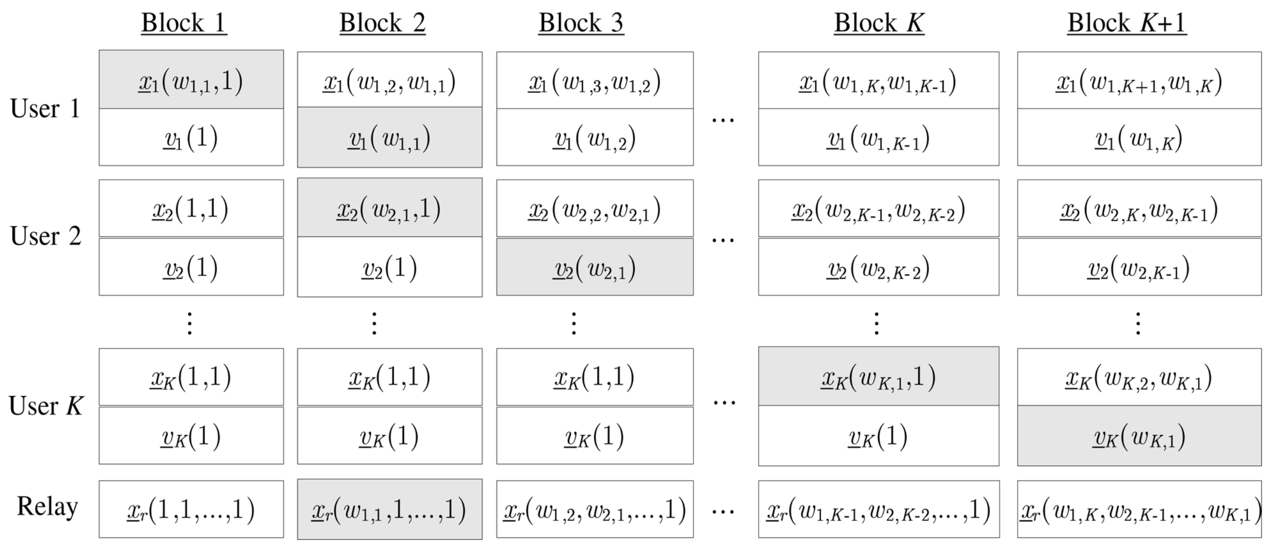

Fig. 4. Offset encoding for a K-user MARC assuming the relay decodes correctly.

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

$$I(V_1X_r; Y_d|X_2V_2) = I(V_1X_r; X_2Y_d|V_2) \tag{16}$$
$$\geq I(V_1X_r; Y_d|V_2) \tag{17}$$

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 the right-hand side of (6) is at least the right-hand side of (11). By symmetry, the right-hand side of (7) is at least the right-hand side of (12). Hence, 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

$$Y_r = X_1 + X_2 \tag{19}$$
$$Y_d = X_1 + X_r \tag{20}$$

where we use integer addition. Any DF rate region must be in the capacity region of the user-to-relay multiple-access channel. This capacity region in bits per channel use is given by (see [17, p. 392])

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

One can check that backward decoding achieves this largest possible DF region for the MARC 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), (11)–(13) are

$$R_1 \leq H(X_1|V_1) \tag{22}$$
$$R_2 \leq \min\left(H(X_2|V_2),\, I(V_2X_r; Y_d|V_1) \right) \tag{23}$$
$$R_1 + R_2 \leq \min\left(H(X_1 + X_2|V_1V_2),\, H(X_1 + X_r) \right). \tag{24}$$

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

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

We further have from (23) that

$$\begin{aligned} R_2 &\leq I(V_2X_r; Y_d|V_1) \\ &= H(X_1 + X_r|V_1) - H(X_1|V_1) \\ &\leq \log_2 3 - H(X_1|V_1). \end{aligned} \tag{26}$$

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

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

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

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

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, i.e., $\pi = (\pi(1), \pi(2), \ldots, \pi(K))$ where $\pi(i) \in \mathcal{K}$ for all i and $\{\pi(i) : i = 1, 2, \ldots, K\} = \mathcal{K}$. We let user $\pi(i)$ start transmitting in block i, i.e., we set $w_{\pi(i),b} = 1$ for $b < i$ and $b \geq B + i$. The resulting message-to-codeword 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, except that $s_{\pi(i),b}$ is now the relay's estimate of $w_{\pi(i),b-i}$. 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 $x_r(s_{\mathcal{K},b})$ where $s_{\mathcal{K},b} = \{s_{k,b} : k \in \mathcal{K}\}$.

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 the destination decodes $(w_{1,b}, w_{2,b})$ 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

$$R_1 \leq I(X_1X_r;Y_d|X_2V_2) \tag{30}$$
$$R_2 \leq I(X_2;Y_d|V_1V_2X_r) + I(V_2;Y_d) \tag{31}$$
$$R_1 + R_2 \leq I(X_1X_2X_r;Y_d). \tag{32}$$

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 $w_{2,b}$, in contrast with the situation of no offset discussed in Section III-B. We can show that (7) is in general larger than (31) by expanding (7) as (see (14) and (15))

$$\begin{aligned} I(X_2X_r;Y_d|X_1V_1) &= I(X_2V_2X_r;Y_d|X_1V_1) && (33)\\ &= I(X_2;Y_d|X_1V_1V_2X_r) \\ &\quad + I(V_2X_r;Y_d|X_1V_1) && (34)\end{aligned}$$

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

$$\begin{aligned} I(X_2;Y_d|X_1V_1V_2X_r) &= I(X_2;X_1Y_d|V_1V_2X_r) && (35)\\ &\geq I(X_2;Y_d|V_1V_2X_r) && (36)\end{aligned}$$

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

$$\begin{aligned} I(V_2X_r;Y_d|X_1V_1) &\geq I(V_2;Y_d|X_1V_1) && (37)\\ &= I(V_2;X_1V_1Y_d) && (38)\\ &\geq I(V_2;Y_d) && (39)\end{aligned}$$

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

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. We focus on the destination bounds because the relay bounds are the same for both no-offset and offset encoding. Note, however, that if the polytopes intersect as in Fig. 3, then one of the corner points of the shaded region is not a corner point of the destination's backward decoding region. To achieve such points, it turns out that we can use either no-offset or offset encoding, as shown below. Alternatively, we could time-share between different offset orders, but this increases the decoding delay.

Consider the corner point

$$(R_1,R_2) = (I(X_1X_r;Y_d|X_2V_2),\, I(X_2V_2;Y_d)) \tag{40}$$

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

$$\begin{aligned} R_1+R_2 &\leq I(X_1X_2X_r;Y_d) && (41)\\ &= I(X_1X_2V_2X_r;Y_d) && (42)\\ &= I(X_1X_r;Y_d|X_2V_2) + I(X_2V_2;Y_d). && (43)\end{aligned}$$

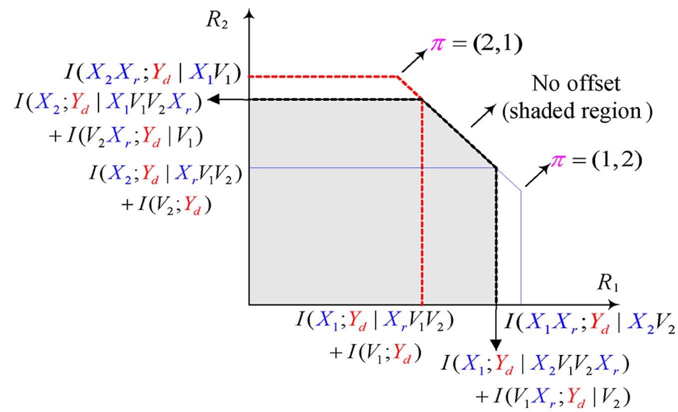

Fig. 5. Rate region with sliding-window decoding and offset encoding.

where (42) follows from the Markov chain $V_2 - (X_1,X_2,X_r) - Y_d$. We further have

$$\begin{aligned} I(X_2V_2;Y_d) &= I(X_2;Y_d|V_2) + I(V_2;Y_d) && (44)\\ &\leq I(X_2;V_1X_rY_d|V_2) + I(V_2;Y_d) && (45)\\ &= I(X_2;Y_d|V_1V_2X_r) + I(V_2;Y_d) && (46)\end{aligned}$$

where (46) 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 shaded region in Fig. 5 shows the points achieved by no-offset encoding that are defined by (11)–(13). Interestingly, the union of rate pairs achieved by the three methods (no-offset encoding, offset encoding with $\pi = (1,2)$, offset encoding with $\pi = (2,1)$) is precisely the backward decoding rate region. Time sharing between offset orders is therefore not needed.

Finally, we remark that the above shows that offset encoding improves sliding-window decoding, since one now achieves the corner point of the example in Section III-B.

B. K Users With Successive Decoding

We wish to show that offset encoding recovers the destination's backward decoding corner points for $K > 2$. However, the generalization of (30)–(32) is unwieldy and gives limited insight. Instead, we use successive decoding inside the sliding window to obtain the backward decoding corner points.

We begin by considering the set function (see (9))

$$f(\mathcal{S}) = \begin{cases} I(X_{\mathcal{S}}X_r;Y_d|X_{\mathcal{S}^C}V_{\mathcal{S}^C}), & \mathcal{S} \subseteq \mathcal{K}, \mathcal{S} \neq \emptyset \\ 0, & \mathcal{S} = \emptyset \end{cases} \tag{47}$$

for some distribution satisfying (10) with U a constant. We claim that $f(\cdot)$ is submodular [18, Ch. 44]. To see this, consider k_1 and k_2 in $\mathcal{K}$ with $k_1 \neq k_2$, $k_1 \notin \mathcal{S}$, $k_2 \notin \mathcal{S}$, and expand

$$\begin{aligned} &f(\mathcal{S}\cup\{k_1\}) + f(\mathcal{S}\cup\{k_2\}) \\ &= I(X_{\mathcal{S}}X_{k_1}V_{k_1}X_r;Y_d|X_{(\mathcal{S}\cup\{k_1\})^C}V_{(\mathcal{S}\cup\{k_1\})^C}) \\ &\quad + I(X_{\mathcal{S}}X_{k_2}V_{k_2}X_r;Y_d|X_{(\mathcal{S}\cup\{k_2\})^C}V_{(\mathcal{S}\cup\{k_2\})^C}) && (48)\\ &= I(X_{k_1}V_{k_1};Y_d|X_{(\mathcal{S}\cup\{k_1\})^C}V_{(\mathcal{S}\cup\{k_1\})^C}) \\ &\quad + I(X_{\mathcal{S}}X_r;Y_d|X_{\mathcal{S}^C}V_{\mathcal{S}^C}) \\ &\quad + I(X_{\mathcal{S}}X_{k_2}V_{k_2}X_r;Y_d|X_{(\mathcal{S}\cup\{k_2\})^C}V_{(\mathcal{S}\cup\{k_2\})^C}) && (49)\end{aligned}$$

where (48) follows from the Markov chain $V_{\mathcal{K}} - (X_{\mathcal{K}}, X_r) - Y_d$ and (49) from the chain rule for mutual information. We lower-bound the first term in (49) as

$$\begin{aligned} &H(X_{k_1}V_{k_1}|X_{(\mathcal{S}\cup\{k_1\})^C}V_{(\mathcal{S}\cup\{k_1\})^C}) \\ &\quad - H(X_{k_1}V_{k_1}|X_{(\mathcal{S}\cup\{k_1\})^C}V_{(\mathcal{S}\cup\{k_1\})^C}Y_d) \\ &= H(X_{k_1}V_{k_1}|X_{(\mathcal{S}\cup\{k_1,k_2\})^C}V_{(\mathcal{S}\cup\{k_1,k_2\})^C}) \\ &\quad - H(X_{k_1}V_{k_1}|X_{(\mathcal{S}\cup\{k_1\})^C}V_{(\mathcal{S}\cup\{k_1\})^C}Y_d) \qquad (50) \\ &\geq I(X_{k_1}V_{k_1};Y_d|X_{(\mathcal{S}\cup\{k_1,k_2\})^C}V_{(\mathcal{S}\cup\{k_1,k_2\})^C}) \qquad (51) \end{aligned}$$

where (50) follows from the independence of the (X_k, V_k) and (51) because conditioning cannot increase entropy. The expression (51) added to the final term in (49) is

$$I(X_{\mathcal{S}\cup\{k_1,k_2\}}X_r;Y_d|X_{(\mathcal{S}\cup\{k_1,k_2\})^C}V_{(\mathcal{S}\cup\{k_1,k_2\})^C}). \qquad (52)$$

Inserting (51) into (49), we have

$$f(\mathcal{S}\cup\{k_1\}) + f(\mathcal{S}\cup\{k_2\}) \geq f(\mathcal{S}) + f(\mathcal{S}\cup\{k_1,k_2\}) \qquad (53)$$

for all $\mathcal{S} \subseteq \mathcal{K}$. The set function $f(\cdot)$ is therefore submodular by [18, Theorem 44.1, p. 767].

The preceding shows that the rate region defined by the destination bounds (see (9))

$$R_{\mathcal{S}} \leq I(X_{\mathcal{S}}X_r;Y_d|X_{\mathcal{S}^c}V_{\mathcal{S}^c}), \qquad \mathcal{S} \subseteq \mathcal{K} \qquad (54)$$

is a polymatroid associated with $f(\cdot)$ (see [18, p. 767]). But the nonzero corner points $\underline{R} = (R_1, R_2, \ldots, R_K)$ of this polymatroid are known to be given by (see [18, p. 777])

$$R_{\pi(k)} = \begin{cases} f(\{\pi(1),\ldots,\pi(k)\}) \\ -f(\{\pi(1),\ldots,\pi(k-1)\}), & k \leq \ell \\ 0, & k > \ell \end{cases} \qquad (55)$$

where π is a permutation of the source indices, $k = 1, 2, \ldots, K$, and $\ell = 1, 2, \ldots, K$. For example, consider $\pi = (1, 2, \ldots, K)$ for which (55) evaluates to

$$R_k = \begin{cases} I(X_1X_r;Y_d|X_{[2,K]}V_{[2,K]}), & k = 1 \\ I(X_kV_k;Y_d|X_{[k+1,K]}V_{[k+1,K]}), & 2 \leq k \leq \ell \\ 0, & k > \ell \end{cases} \qquad (56)$$

where $X_{[K+1,K]}$ and $V_{[K+1,K]}$ are considered to be constants.

We are mainly interested in the corner points of the *base polytope* defined by $\ell = K$ in (55) (see [18, p. 767]) because the other corner points are achieved by discarding message bits. The expression (55) shows that there are up to $K!$ base polytope corner points, namely, one point for each π.

Suppose the offset order is $\pi = (1, 2, \ldots, K)$ as in Fig. 4. Consider the window with the channel-symbol blocks $\underline{y}_{d,1}, \underline{y}_{d,2}, \ldots, \underline{y}_{d,K+1}$. 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

$$R_k = \begin{cases} I(X_1X_r;Y_d|X_{[2,K]}V_{[2,K]}), & k = 1 \\ I(X_kV_k;Y_d|X_{[k+1,K]}V_{[k+1,K]}), & 2 \leq k \leq K \end{cases} \qquad (57)$$

where $X_{[K+1,K]}$ and $V_{[K+1,K]}$ are considered to be constants. The codewords contributing to these rates are shown as shaded blocks in Fig. 4. But the rates (57) are precisely the rates in (56) for $\ell = K$. Hence, we achieve the desired corner point. We can achieve the other corner points by changing the offset order π. Finally, we can achieve the non-corner points by time-sharing between offset orders. An interesting open problem is whether the union of rate points achieved by using all combinations of offset orderings and no-offsets gives the backward decoding rate region (see Fig. 5). If so, then as for $K = 2$ there is no need to time-share between offset orders.

V. Conclusion

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 destination's backward decoding rate region but avoids the excessive delay associated with backward decoding. Offset encoding will clearly apply to other multiterminal problems [13], [19]–[21].

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 (strongly) typical sequence decoders. Let $n(a, b|\underline{x}, \underline{y})$ be the number of times the pair (a, b) occurs in the sequence $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$, and let $\mathcal{X}$ and $\mathcal{Y}$ be the alphabets of X and Y with cardinalities $|\mathcal{X}|$ and $|\mathcal{Y}|$, respectively. Define the set of typical sequences of length n with respect to ϵ and $P_{X,Y}(\cdot)$ as

$$T_\epsilon^{(n)}(X,Y) = \left\{ (\underline{x},\underline{y}) : \left| \frac{n(a,b|\underline{x},\underline{y})}{n} - P_{X,Y}(a,b) \right| \leq \frac{\epsilon}{|\mathcal{X}|\cdot|\mathcal{Y}|} \text{ for all } (a,b) \text{ and } n(a,b|\underline{x},\underline{y}) = 0 \text{ if } P_{X,Y}(a,b) = 0 \right\}. \qquad (58)$$

We refer to [22, Ch. 2] for properties of such sequences.

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 (see [3]). More precisely, the relay decodes by finding a $(\tilde{w}_{1,b}, \tilde{w}_{2,b})$ such that

$$\begin{aligned} &(\underline{x}_1(\tilde{w}_{1,b}, w_{1,b-1}), \underline{x}_2(\tilde{w}_{2,b}, w_{2,b-1}), \underline{v}_1(w_{1,b-1}), \\ &\quad \underline{v}_2(w_{2,b-1}), \underline{x}_r(w_{1,b-1}, w_{2,b-1}), \underline{y}_{r,b}) \\ &\quad \in T_\epsilon^{(n)}(X_1, X_2, V_1, V_2, X_r, Y_r). \qquad (59) \end{aligned}$$

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 $(\tilde{w}_{1,b}, \tilde{w}_{2,b}) \neq (w_{1,b}, w_{2,b})$ that satisfies (59). This error event can be further split into three disjoint error events. The first error event has a $\tilde{w}_{1,b} \neq w_{1,b}$ and $\tilde{w}_{2,b} = w_{2,b}$ satisfying (59). Using [4, Lemma 1] and the union bound, the probability of this event is at most

$$2^{n(R_1 - I(X_1;Y_r|X_2V_1V_2X_r)+6\epsilon)}. \qquad (60)$$

Thus, for reliable decoding we set

$$R_1 < I(X_1;Y_r|X_2V_1V_2X_r). \qquad (61)$$

The second error event has $\tilde{w}_{1,b} = w_{1,b}$ and a $\tilde{w}_{2,b} \neq w_{2,b}$ satisfying (59). By symmetry to (61), we set

$$R_2 < I(X_2;Y_r|X_1V_1V_2X_r). \qquad (62)$$

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 (59). We again use [4, Lemma 1] to bound the probability of this event by

$$2^{n(R_1+R_2-I(X_1X_2;Y_r|V_1V_2X_r)+6\epsilon)}. \tag{63}$$

Reliable decoding thus requires

$$R_1 + R_2 < I(X_1X_2;Y_r|V_1V_2X_r). \tag{64}$$

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 (see [3]). More precisely, the destination decodes by finding a $(\tilde{w}_{1,b}, \tilde{w}_{2,b})$ such that

$$\begin{aligned}(\underline{x}_1(w_{1,b+1}, \tilde{w}_{1,b}), \underline{x}_2(w_{2,b+1}, \tilde{w}_{2,b}), \underline{v}_1(\tilde{w}_{1,b}), \\ \underline{v}_2(\tilde{w}_{2,b}), \underline{x}_r(\tilde{w}_{1,b}, \tilde{w}_{2,b}), \underline{y}_{d,b+1}) \\ \in T_\epsilon^{(n)}(X_1, X_2, V_1, V_2, X_r, Y_d).\end{aligned} \tag{65}$$

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

$$R_1 < I(X_1X_r;Y_d|X_2V_2) \tag{66}$$
$$R_2 < I(X_2X_r;Y_d|X_1V_1) \tag{67}$$
$$R_1 + R_2 < I(X_1X_rX_2;Y_d) \tag{68}$$

ensures reliable communications.

Combining (61), (62), (64), and (66)–(68), we have the bounds (3)–(8). The analysis carries over in a straightforward way to weakly typical (or entropy-typical) sequences [17, p. 51], the addition of a time-sharing random variable U [17, 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 $\underline{x}_1(w_{1,b}, w_{1,b-1})$ while source 2 transmits $\underline{x}_2(w_{2,b-1}, w_{2,b-2})$ where $w_{2,-1}$, $w_{2,0}$, $w_{1,0}$, $w_{1,B+1}$, $w_{1,B+2}$, and $w_{2,B+1}$ 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 (61), (62), and (64).
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 following three events occur:

$$\begin{aligned}\mathcal{E}_1 : (\underline{v}_1(w_{1,b-1}), \underline{v}_2(w_{2,b-2}), \underline{x}_1(\tilde{w}_{1,b}, w_{1,b-1}), \\ \underline{x}_2(w_{2,b-1}, w_{2,b-2}), \underline{x}_r(w_{1,b-1}, w_{2,b-2}), \underline{y}_{d,b}) \\ \in T_\epsilon^{(n)}(V_1, V_2, X_1, X_2, X_r, Y_d)\end{aligned} \tag{69}$$

$$\begin{aligned}\mathcal{E}_2 : (\underline{v}_1(\tilde{w}_{1,b}), \underline{v}_2(w_{2,b-1}), \underline{x}_2(\tilde{w}_{2,b}, w_{2,b-1}), \\ \underline{x}_r(\tilde{w}_{1,b}, w_{2,b-1}), \underline{y}_{d,b+1}) \\ \in T_\epsilon^{(n)}(V_1, V_2, X_2, X_r, Y_d)\end{aligned} \tag{70}$$

$$\mathcal{E}_3 : (\underline{v}_2(\tilde{w}_{2,b}), \underline{y}_{d,b+2}) \in T_\epsilon^{(n)}(V_2, Y_d). \tag{71}$$

Note that the codebooks in different blocks are generated independently (see Section II-B) so the above three events are independent (see [6], [10]). 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 (69)–(71). We upper-bound the probability of this error event using [4, Lemma 1] and the union bound as

$$\begin{aligned}&\textstyle\sum_{\tilde{w}_{1,b}\neq w_{1,b}} \Pr(\mathcal{E}_1 \cap \mathcal{E}_2 \cap \mathcal{E}_3) \\ &= \sum_{\tilde{w}_{1,b}\neq w_{1,b}} \Pr(\mathcal{E}_1)\cdot\Pr(\mathcal{E}_2)\cdot\Pr(\mathcal{E}_3) &(72)\\ &\le 2^{n(R_1 - I(X_1;Y_d|X_2V_1V_2X_r) - I(V_1X_r;Y_d|X_2V_2)+12\epsilon)} &(73)\\ &= 2^{n(R_1 - I(X_1X_r;Y_d|X_2V_2)+12\epsilon)} &(74)\end{aligned}$$

where we used $\Pr(\mathcal{E}_3) \le 1$ for (73) and (14)–(15) for (74). Thus, we set

$$R_1 < I(X_1X_r;Y_d|X_2V_2). \tag{75}$$

Consider next the case where $\tilde{w}_{1,b} = w_{1,b}$ but $\tilde{w}_{2,b} \neq w_{2,b}$. The expression (72) 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

$$2^{n(R_2 - I(X_2;Y_d|V_1V_2X_r) - I(V_2;Y_d)+12\epsilon)} \tag{76}$$

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

$$R_2 < I(X_2;Y_d|V_1V_2X_r) + I(V_2;Y_d). \tag{77}$$

Finally, consider the case $\tilde{w}_{1,b} \neq w_{1,b}$ and $\tilde{w}_{2,b} \neq w_{2,b}$. The expression (72) with the summation now over both $\tilde{w}_{1,b} \neq w_{1,b}$ and $\tilde{w}_{2,b} \neq w_{2,b}$ is upper-bounded as

$$\begin{aligned}&2^{n(R_1+R_2)}\cdot 2^{-nI(X_1;Y_d|X_2V_1V_2X_r)+n6\epsilon} \\ &\quad\cdot 2^{-nI(X_2V_1X_r;Y_d|V_2)+n6\epsilon}\cdot 2^{-nI(V_2;Y_d)+n6\epsilon} &(78)\\ &= 2^{n(R_1+R_2-I(X_1X_2X_r;Y_d)+18\epsilon)} &(79)\end{aligned}$$

where we have used the chain rule for mutual information and the Markov chain $(V_1, V_2) - (X_1, X_2, X_r) - Y_d$. For reliable decoding, we thus require

$$R_1 + R_2 < I(X_1X_2X_r;Y_d). \tag{80}$$

Combining (75), (77), and (80), we obtain (30)–(32). Again, the analysis carries over in a straightforward way to weakly 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 $K \ge 2$, offset encoding, and sliding-window decoding. We further focus on the message blocks $w_{k,b}$ with $b = 1$. However, the destination now performs successive rather than joint decoding. Without loss of generality, we consider the offset order $\pi = (1,2,\ldots,K)$. 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 $\underline{y}_{d,1}, \underline{y}_{d,2}, \ldots, \underline{y}_{d,K+1}$. As explained in Section IV-B, the destination successively decodes in the reverse order

$w_{K,1}, w_{K-1,1}, \ldots, w_{1,1}$ (see the shaded blocks in Fig. 4 for the $w_{(k,l)}$ with $k = K$, $k = 2$, and $k = 1$). The destination further assumes that its past decoding steps were successful, and we perform our analysis with the same assumption. For $k = K, K-1, \ldots, 2$, the destination finds a $\tilde{w}_{k,1}$ such that the following two events occur:

$$\begin{aligned}\mathcal{E}_{1,k} : (&\underline{v}_k(\tilde{w}_{k,1}), \underline{v}_{[k+1,K]}(1), \underline{x}_{[k+1,K]}(1,1), \underline{y}_{d,k+1}) \\ &\in T_\epsilon^{(n)}(V_k, V_{[k+1,K]}, X_{[k+1,K]}, Y_d)\end{aligned} \tag{81}$$

$$\begin{aligned}\mathcal{E}_{2,k} : (&\underline{x}_k(\tilde{w}_{k,1}, 1), \underline{v}_{[k,K]}(1), \underline{x}_{[k+1,K]}(1,1), \underline{y}_{d,k}) \\ &\in T_\epsilon^{(n)}(X_k, V_{[k,K]}, X_{[k+1,K]}, Y_d)\end{aligned} \tag{82}$$

where $\underline{v}_{[i,j]}(1) = \{v_i(1), v_{i+1}(1), \ldots, v_j(1)\}$ and similarly for $\underline{x}_{[i,j]}(1,1)$ and $\underline{v}_{\mathcal{K}}(1)$ below. As before, we assume that variables with vacuous index sets are appropriate constants, e.g., we assume that all the entries of $\underline{v}_{[K+1,K]}$ are the same constant $V_{[K+1,K]}$.

The events $\mathcal{E}_{1,k}$ and $\mathcal{E}_{2,k}$ 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 (81) and (82). We upper-bound the probability of this event using [4, Lemma 1] as

$$\begin{aligned}&\sum_{\tilde{w}_{k,1} \neq w_{k,1}} \Pr(\mathcal{E}_{1,k}) \cdot \Pr(\mathcal{E}_{2,k}) \\ &\qquad \le 2^{n\left(R_k - I(X_k V_k; Y_d | X_{[k+1,K]} V_{[k+1,K]}) + 12\epsilon\right)}.\end{aligned} \tag{83}$$

For $2 \le k \le K$, we therefore require

$$R_k < I(X_k V_k; Y_d | X_{[k+1,K]} V_{[k+1,K]}). \tag{84}$$

For $k = 1$, we add $\underline{x}_r(\cdot)$ to (81) and (82) as follows:

$$\begin{aligned}\mathcal{E}_{1,1} : (&\underline{v}_1(\tilde{w}_{1,1}), \underline{v}_{[2,K]}(1), \underline{x}_{[2,K]}(1,1), \underline{x}_r(\tilde{w}_{1,1}, 1, \ldots, 1), \underline{y}_{d,2}) \\ &\in T_\epsilon^{(n)}(V_1, V_{[2,K]}, X_{[2,K]}, X_r, Y_d)\end{aligned} \tag{85}$$

$$\begin{aligned}\mathcal{E}_{2,1} : (&\underline{x}_1(\tilde{w}_{1,1}, 1), \underline{v}_{\mathcal{K}}(1), \underline{x}_{[2,K]}(1,1), \underline{x}_r(1, 1, \ldots, 1), \underline{y}_{d,1}) \\ &\in T_\epsilon^{(n)}(X_1, V_{\mathcal{K}}, X_{[2,K]}, X_r, Y_d).\end{aligned} \tag{86}$$

The resulting bound is

$$R_1 < I(X_1 X_r; Y_d | X_{[2,K]} V_{[2,K]}). \tag{87}$$

For example, for $K = 2$, the two rate bounds are

$$R_2 < I(X_2 V_2; Y_d) \tag{88}$$

$$R_1 < I(X_1 X_r; Y_d | X_2 V_2) \tag{89}$$

and one can approach the corner point (40). One can check that the above analysis generalizes to $b > 1$.

Acknowledgment

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

References

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

[2] E. C. van der Meulen, "Three-terminal communication channels," *Adv. Appl. Probab.*, vol. 3, pp. 120–154, 1971.

[3] T. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. IT-25, pp. 572–584, Sep. 1979.

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

[5] L. Sankaranarayanan, G. Kramer, and N. B. Mandayam, "Capacity theorems for the multiple-access relay channel," in *Proc. 42nd Annu. Allerton Conf. Communications, Control, and Computing*, Monticello, IL, Sep. 2004, pp. 1782–1791.

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

[7] F. M. J. Willems, "Information-theoretical Results for the Discrete Memoryless Multiple Access Channel," Doctor in de Wetenschappen Proefschrift (Ph.D.) dissertation, Katholieke Univ. Leuven, Leuven, Belgium, 1982.

[8] M. R. Aref, "Information Flow in Relay Networks," Ph.D. dissertation, Stanford Univ., Stanford, CA, 1980.

[9] P. Gupta and P. R. Kumar, "Towards an information theory of large networks: An achievable rate region," *IEEE Trans. Inf. 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. Inf. 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. Inf. Theory*, vol. 52, no. 6, pp. 2624–2633, Jun. 2006.

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

[13] J. N. Laneman and G. Kramer, "Window decoding for the multiple-access channel with generalized feedback," in *Proc. 2004 IEEE Int. Symp. Information Theory*, Chicago, IL, Jun./Jul. 2004, p. 281.

[14] B. Rimoldi and R. Urbanke, "A rate-splitting approach to the Gaussian multiple access channel," *IEEE Trans. Inf. Theory*, vol. 42, no. 2, pp. 364–375, Mar. 1996.

[15] A. Grant, B. Rimoldi, R. Urbanke, and P. Whiting, "Rate-splitting multiple-access for discrete memoryless channels," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 873–890, Mar. 2001.

[16] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

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

[18] A. Schrijver, *Combinatorial Optimization: Polyhedra and Efficiency*. New York: Springer-Verlag, 2003.

[19] L. Sankaranarayanan, G. Kramer, and N. B. Mandayam, "Cooperation vs. hierarchy: An information-theoretic comparison," in *Proc. Int. Symp. Information Theory*, Adelaide, Australia, Sep. 2005, pp. 411–415.

[20] L.-L. Xie and P. R. Kumar, "A multi-relay scheme for the multi-source multi-cast network," in *Proc. IEEE Int. Symp. Information Theory*, Seattle, WA, Jul. 2006, pp. 2383–2387.

[21] L.-L. Xie and P. R. Kumar, "Multisource, multidestination, multirelay wireless networks," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3586–3595, Oct. 2007.

[22] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Budapest, Hungary: Académiai Kiadó, 1981.