

The Capacity of a Class of Multi-Way Relay Channels

Lawrence Ong, Sarah J. Johnson, and Christopher M. Kellett

School of Electrical Engineering and Computer Science, The University of Newcastle

Email: lawrence.ong@cantab.net; {sarah.johnson, chris.kellett}@newcastle.edu.au

Abstract—The capacity of a class of multi-way relay channels, where L users communicate via a relay (at possibly different rates), is derived for the case where the channel outputs are modular sums of the channel inputs and the receiver noise. The cut-set upper bound to the capacity is shown to be achievable. More specifically, the capacity is achieved using (i) rate splitting, (ii) functional-decode-forward, and (iii) joint source-channel coding. We note that while separate source-channel coding can achieve the common-rate capacity, joint source-channel coding is used to achieve the capacity for the general case where the users are transmitting at different rates.

I. INTRODUCTION

We consider the multi-way relay channel (MWRC), where L users ($L \geq 2$) exchange data via a relay, and where there is no direct link between the users. Common applications of this model are conference calls in the cellular network and satellite communications.

The MWRC is an extension of the two-way relay channel (TWRC) where two users ($L = 2$) exchange data via a relay (e.g., see [1], [2], [3]). The Gaussian MWRC, where the channels between the nodes are additive white Gaussian noise channels, was first investigated by Gündüz *et al.* [4]. An upper bound and a few achievable rate regions, based on the coding strategies for the relay channel, were derived using: (i) *complete-decode-forward* (CDF) where the relay completely decodes the users' messages and broadcasts a function of the messages back to the users, (ii) *compress-forward* where the relay quantizes its received signals, re-encodes and broadcasts them to the users, and (iii) *amplify-forward* where the relay simply scales and forwards what it receives. These coding strategies, however, fail to achieve the capacity of the MWRC.

Recently, *functional-decode-forward* (FDF) has been proposed for the TWRC, where the relay decodes a function of the users' messages and broadcasts the function back to the users. FDF has been shown to achieve within $\frac{1}{2}$ bit of the capacity of the Gaussian TWRC [5]. We later proposed FDF for the multi-way relay channel (MWRC), and showed that FDF achieves the *common-rate* (where all users exchange information at the same rate) capacity of the binary MWRC [6], where the channels are binary symmetric. Applying insights from the binary MWRC has allowed us to obtain the common-rate capacity of the the Gaussian MWRC with three or more users where all nodes transmit at the same power [7]. The "general"

This work is supported by the Australian Research Council under grants DP0877258 and DP1093114.

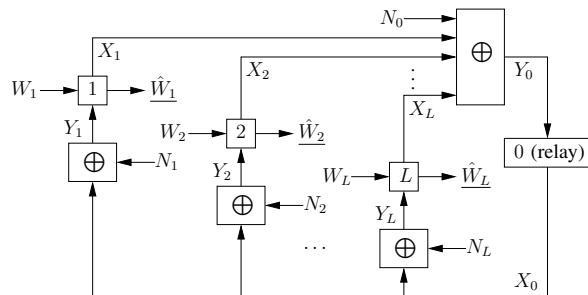


Fig. 1. The finite field adder MWRC

capacity (i.e., where users can transmit at possibly different rates) of the MWRC is not yet known.

In this paper, we work toward this goal by deriving the "general" capacity of the *finite field adder* MWRC, where the channel outputs are the summation (in finite field arithmetic) of the channel inputs and the receiver noise. We show that the capacity can be achieved by combining the ideas of (i) rate splitting, (ii) our proposed FDF [6], and (iii) the joint source-channel coding for broadcast channels by Tuncel [8]. This, to the best of our knowledge, is the first example of the MWRC where the capacity is found for all noise distributions/levels.

The rest of the paper is organized as follows. We define the channel model of the finite field adder MWRC in Sec. II, and find a capacity upper bound in Sec. III. In Sec. IV, we construct a linear code that is optimal for the point-to-point finite field adder channel. Using this linear code, we propose a coding strategy using the ideas of rate splitting, FDF, and joint source-channel coding to obtain the capacity of the finite field adder MWRC in Sec. V. Lastly, in Sec. VI, we compare the capacity with two other coding strategies, namely FDF with rate splitting and separate source-channel coding and CDF, and discuss why these two strategies fall short of the capacity.

II. CHANNEL MODEL

Fig. 1 depicts the MWRC considered in this paper, where there is no direct user-to-user link. Nodes 1, 2, ..., L are the users, and node 0 is the relay. We consider full data exchange where each user is to decode the messages from all other users. We denote by X_i node i 's input to the channel, Y_i the channel output received by node i , W_i node i 's message, and \hat{W}_i node i 's estimate of all other users' messages.

The L -user finite field adder MWRC over the finite field \mathcal{F} consists of the following:

• Uplink: $Y_0 = \left(\bigoplus_{1 \leq i \leq L} X_i \right) \oplus N_0 \triangleq X_1 \oplus X_2 \oplus \dots \oplus X_L \oplus N_0$,

• Downlink: $Y_i = X_0 \oplus N_i$, for each $i = 1, 2, \dots, L$,

where $X_i, Y_i, N_i \in \mathcal{F}$, $\forall i$, for some finite field \mathcal{F} , \oplus is the addition operation associated with \mathcal{F} , N_i are statistically independent for all i and all channel uses. Let $W_i \in \{1, 2, \dots, 2^{nR_i}\}$ be an (nR_i) -bit message, where R_i is a rational number for every $1 \leq i \leq L$, and consider n simultaneous uplink and downlink channel uses. User i 's transmit message at time t , $X_i[t]$, can only depend on its own message and its past received signals, i.e., $X_i[t] = f_{i,t}(W_i, Y_i[1], Y_i[2], \dots, Y_i[t-1])$, for $1 \leq t \leq n$. The relay's transmitted signal at any time can only depend on its past received signals, i.e., $X_0[t] = f_{0,t}(Y_0[1], Y_0[2], \dots, Y_0[t-1])$, for $1 \leq t \leq n$. After n channel uses, user i estimates the messages of all other nodes from its received signals and its own message, i.e., $\hat{W}_i = g_i(\mathbf{Y}_i, W_i)$, where $\mathbf{Y}_i = (Y_i[1], Y_i[2], \dots, Y_i[n])$. Assume that the users' messages are independent and each W_i is uniformly distributed over $\{1, 2, \dots, 2^{nR_i}\}$. We say that the rate tuple (R_1, R_2, \dots, R_L) is *achievable* if there exists some $(n, \{f_{i,t}\}_{0 \leq i \leq L, 1 \leq t \leq n}, \{g_i\}_{1 \leq i \leq L})$ such that all users can *reliably* decode the messages of all other users. We say that a user can decode a message reliably if the probability that it wrongly decodes the message can be made arbitrarily small. The *capacity* is defined as the closure of all achievable rate tuples.

III. A CAPACITY UPPER BOUND

In this section, we derive an upper bound to the capacity of the finite field adder MWRC using cut-set arguments. A cut-set upper bound to the capacity of a network is the maximum rate that information can be transferred across a *cut* separating two disjoint sets of nodes, assuming that all nodes on each side of the cut can fully cooperate. We define $R_{\min} = \min_{1 \leq j \leq L} R_j$,

$R_i^c = \sum_{j=1, j \neq i}^L R_j$, and $R_{\min}^c = \left(\sum_{j=1}^L R_j \right) - R_{\min}$. The cut-set upper bound to the capacity of the finite field adder MWRC is given in the following theorem.

Theorem 1: Consider an L -user finite field adder MWRC over \mathcal{F} . If the rate tuple (R_1, R_2, \dots, R_L) is achievable, then

$$R_{\min}^c \leq \log_2 |\mathcal{F}| - H(N_0) \quad (1)$$

$$R_i^c \leq \log_2 |\mathcal{F}| - H(N_i), \quad (2)$$

for all $1 \leq i \leq L$.

Here, $H(X) = -\sum_{x \in \mathcal{X}} p(x) \log_2 p(x)$ is the entropy.

Proof of Theorem 1: Consider a network of m nodes, in which node i sends information at the rate $R_{i,j}$ (in bits/channel use) to node j . If the set of rates $\{R_{i,j}\}$ are achievable, there exists some joint probability distribution $p(x_1, x_2, \dots, x_m)$ such that $\sum_{i \in \mathcal{S}, j \in \mathcal{S}^c} R_{i,j} \leq I(X_{\mathcal{S}}; Y_{\mathcal{S}^c} | X_{\mathcal{S}^c})$, for all $\mathcal{S} \subset \{1, 2, \dots, m\}$ [9, p. 589]. Here $X_{\mathcal{S}} = \{X_i : i \in \mathcal{S}\}$, and $\mathcal{S}^c = \{1, 2, \dots, m\} \setminus \mathcal{S}$.

First, consider the cut separating $\mathcal{S} = \{1, 2, \dots, i-1, i+2, \dots, L\}$ and $\mathcal{S}^c = \{0, i\}$ in the MWRC, for some $1 \leq i \leq L$. An upper bound to the rate R_i^c (of messages

$(W_1, W_2, \dots, W_{i-1}, W_{i+1}, \dots, W_L)$) across the cut from \mathcal{S} to \mathcal{S}^c is therefore

$$\sum_{j=1, j \neq i}^L R_j = R_i^c \leq I(X_{[1,L] \setminus \{i\}}; Y_{\{0,i\}} | X_{\{0,i\}}). \quad (3a)$$

$$= H(Y_0, Y_i | X_0, X_i) - H(Y_0, Y_i | X_{[0,L]}) \quad (3b)$$

$$= H(X_1 \oplus \dots \oplus X_{i-1} \oplus X_{i+1} \oplus \dots \oplus X_L \oplus N_0, N_i) - H(N_0, N_i) \quad (3c)$$

$$= H \left(\left(\bigoplus_{j \in [1,L] \setminus \{i\}} X_j \right) \oplus N_0 \right) - H(N_0), \quad (3d)$$

where (3d) is because $\left(\bigoplus_{j \in [1,L] \setminus \{i\}} X_j \right) \oplus N_0$ and N_i are statistically independent, so are N_0 and N_i .

Next, consider the cut separating $\mathcal{S} = \{0, 1, 2, \dots, i-1, i+2, \dots, L\}$ and $\mathcal{S}^c = \{i\}$, for some $1 \leq i \leq L$. We have the following rate constraint

$$R_i^c \leq I(X_{[0,L] \setminus \{i\}}; Y_i | X_i) \quad (4a)$$

$$= H(X_0 \oplus N_i) - H(N_i). \quad (4b)$$

The rate constraints (3d) and (4b) must be satisfied for all $1 \leq i \leq L$ for some $p(x_0, x_1, \dots, x_L)$. Note that choosing the independent and uniform distribution for each X_i , for $0 \leq i \leq L$, simultaneously maximizes all the mutual information terms in the constraints. So, combining the above rate constraints, we have Theorem 1. Note that (1) implies (3d) for all $1 \leq i \leq L$, since $R_{\min}^c = \max_{1 \leq j \leq L} R_j^c$. ■

IV. AN OPTIMAL LINEAR CODE FOR THE POINT-TO-POINT FINITE FIELD ADDER CHANNEL

Now, we consider the following linear code that maps a length- k (row vector) message $\mathbf{s} \in \mathcal{F}^k$ to a length- n (row vector) codeword $\mathbf{x} \in \mathcal{F}^n$:

$$\mathbf{x} = (\mathbf{s} \odot \mathbb{G}) \oplus \mathbf{q} = \left(\mathbf{s} \odot \begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \\ \vdots \\ \mathbf{g}_k \end{bmatrix} \right) \oplus \mathbf{q}, \quad (5)$$

where \odot is the multiplication associated with \mathcal{F} , \mathbb{G} is a fixed $k \times n$ matrix, with each element independently and uniformly chosen over \mathcal{F} , the i -th row in \mathbb{G} , \mathbf{g}_i , is a row vector of length n , and \mathbf{q} is a fixed row vector of length n , with each element independently and uniformly chosen over \mathcal{F} .

We extend the results for binary linear codes [10, p. 206–207] to finite field linear codes in the following two lemmas.

Lemma 1: Consider the linear codes defined in (5). Over the ensemble of codes, the probability that a message \mathbf{s}_1 is mapped to a given codeword \mathbf{x}_1 is $p(\mathbf{x}_1) = |\mathcal{F}|^{-n}$.

Proof of Lemma 1: There are $|\mathcal{F}|^{n(k+1)}$ ways of selecting \mathbb{G} and \mathbf{q} . As the elements are arbitrarily chosen, each unique (\mathbb{G}, \mathbf{q}) has a probability of $|\mathcal{F}|^{-n(k+1)}$ of being selected. For any \mathbb{G} , there is only one \mathbf{q} that results in the given \mathbf{x}_1 . So, there are only $|\mathcal{F}|^{nk}$ different (\mathbb{G}, \mathbf{q}) that map \mathbf{s}_1 to \mathbf{x}_1 . Hence, $p(\mathbf{x}_1) = |\mathcal{F}|^{nk} |\mathcal{F}|^{-n(k+1)} = |\mathcal{F}|^{-n}$. ■

Lemma 2: Consider the linear codes defined in (5). Let s_1 and s_2 be two different messages. The corresponding codewords $\mathbf{x}_1 = (s_1 \odot \mathbb{G}) \oplus \mathbf{q}$ and $\mathbf{x}_2 = (s_2 \odot \mathbb{G}) \oplus \mathbf{q}$ are independent.

Proof of Lemma 2: To show independence, we need to find the probabilities $p(\mathbf{x}_1)$ and $p(\mathbf{x}_2|\mathbf{x}_1)$. Equivalently, we find the probabilities $p(\mathbf{x}_1 \oplus -\mathbf{x}_2)$ and $p(\mathbf{x}_1|\mathbf{x}_1 \oplus -\mathbf{x}_2)$, where $-\mathbf{x}_2$ is the *additive inverse* of \mathbf{x}_2 in \mathcal{F} . Let s_1 and s_2 differ in the j -th position (they may differ, additionally, in other positions). So, $\mathbf{x}_1 \oplus -\mathbf{x}_2 = (s_1 \oplus -s_2) \odot \mathbb{G}$. For any $(\mathbf{g}_1, \dots, \mathbf{g}_{j-1}, \mathbf{g}_{j+1}, \dots, \mathbf{g}_k)$, there is only one \mathbf{g}_j that results in the given $(\mathbf{x}_1 \oplus -\mathbf{x}_2)$. Hence, there are only $|\mathcal{F}|^{n(k-1)}$ different \mathbb{G} 's that give $(\mathbf{x}_1 \oplus -\mathbf{x}_2)$. In addition, for any chosen \mathbb{G} , there is only one \mathbf{q} that results in the given \mathbf{x}_1 . So, there are only $|\mathcal{F}|^{n(k-1)}$ unique (\mathbb{G}, \mathbf{q}) 's that give the desired \mathbf{x}_1 and \mathbf{x}_2 . So, the probability $p(\mathbf{x}_1, \mathbf{x}_2) = |\mathcal{F}|^{n(k-1)}|\mathcal{F}|^{-n(k+1)} = |\mathcal{F}|^{-2n} = p(\mathbf{x}_1)p(\mathbf{x}_2)$. ■

With the above lemmas, we have the following theorem:

Theorem 2: Consider the finite field adder channel

$$Y = X \oplus N, \quad (6)$$

where $Y, X, N \in \mathcal{F}$, where X is the channel input, Y is the channel output, N is independent and identically distributed (i.i.d.) noise for each channel use. A transmitter sends a message $s \in \mathcal{F}^k$ over n uses of the channel (6) using the linear code in (5). The receiver can reliably decode the message from the n received signals \mathbf{Y} if n is sufficiently large and if

$$(k \log_2 |\mathcal{F}|)/n < \log_2 |\mathcal{F}| - H(N). \quad (7)$$

Sketch of proof for Theorem 2: From Lemma 1 we know that for the code defined in (5), for any codeword, each *codeletter* is uniform and i.i.d.. From Lemma 2, we know that any pair of codewords are independent of each other. Using these two facts, we can repeat the analysis of the probability of error in the proof of the channel coding theorem [9, p. 201–204] to show that the receiver can decode the message s from the n received signals \mathbf{Y} with an arbitrarily small error probability if n is sufficiently large and if $\frac{k \log_2 |\mathcal{F}|}{n} < I(X; Y)$, where X is uniformly distributed. ■

V. FUNCTIONAL-DECODE-FORWARD WITH RATE SPLITTING AND JOINT SOURCE-CHANNEL CODING

In this section we derive an achievable rate region using the linear code derived in the previous section. Consider each user i , for $1 \leq i \leq L$, sending T messages (of nR_i bits each), denoted by $(W_i[1], W_i[2], \dots, W_i[T])$. Consider a total of $(T+1)n$ channel uses. Since we consider full data exchange, user i needs to decode the messages sent by all the other users, i.e., $\{W_j[t] : \forall j \in [1, L] \setminus \{i\}, \forall t \in [1, T]\}$. Define each n channel uses as a block. In the t -th block, for $1 \leq t \leq T$, each user i sends $\mathbf{X}_i(W_i[t])$ on the uplink. In the $(t+1)$ -th block, for $1 \leq t \leq T$, the relay transmits \mathbf{X}_0 , a function of its received signals in the t -th block, on the downlink. At the end of the $(t+1)$ -th block, each user i then decodes the t -th message of all other users, i.e., $(W_1[t], \dots, W_{i-1}[t], W_{i+1}[t], \dots, W_L[t])$. So, for each pair of

the t -th block on the uplink and the $(t+1)$ -th block on the downlink, if each user can reliably decode the t -th message of all other users, then repeating the same coding scheme for all $1 \leq t \leq T$, all users can reliably decode the messages from all other users in all blocks. This means that the rate tuple $(\frac{TnR_1}{(T+1)n}, \frac{TnR_2}{(T+1)n}, \dots, \frac{TnR_L}{(T+1)n})$ is achievable. For any n, R_1, R_2, \dots, R_L , we can choose a sufficiently large T such that the achievable rate tuple is arbitrarily close to (R_1, R_2, \dots, R_L) . In this section, we derive constraints on R_1, R_2, \dots, R_L such that the rate tuple is achievable.

Since the encoding and decoding functions for all nodes are repeated in each block, we focus on the first block on the uplink and the second block on the downlink. For simplicity, we denote $W_i[1]$ by W_i in the rest of this section.

A. Uplink

Recall that $R_i^c = \sum_{j=1, j \neq i}^L R_j$, $R_{\min} = \min_{1 \leq j \leq L} R_j$, and $R_{\min}^c = (\sum_{j=1}^L R_j) - R_{\min}$. For the uplink of the MWRC, we use the idea of FDF in [6] and rate splitting. Let $R_i = R_{\min} + R'_i$. So, each message W_i can be split into $W_i = (A_i, B_i)$, where A_i is nR_{\min} bits long and B_i is nR'_i bits long. Let D , $0 \leq D < L$, be the number of users whose message is strictly more than nR_{\min} bits long. Let these users be $\{d_1, d_2, \dots, d_D\} \triangleq \mathcal{D} = \{j : R'_j > 0\}$. So, for all users $i \notin \mathcal{D}$, $W_i = A_i$ and $R'_i = 0$.

The n uplink channel uses are further split into $(L+D-1)$ sub-blocks. Each of the t -th sub-blocks for $1 \leq t \leq L-1$ consists of nR_{\min}/R_{\min}^c channel uses[†]. The t -th block for $L \leq t \leq L+D-1$ consists of $nR'_{d_{t-L+1}}/R_{\min}^c$ channel uses[†]. Note that if we the sum the number of channel uses in all sub-blocks, we get $(L-1)nR_{\min}/R_{\min}^c + n \sum_{d \in \mathcal{D}} R'_d/R_{\min}^c = n[\sum_{j=1}^L (R_{\min} + R'_j) - R_{\min}]/R_{\min}^c = n$.

In the t -th sub-block for $1 \leq t \leq L-1$, only two users transmit, using the linear code defined in (5):

$$\mathbf{X}_i = \begin{cases} (s(A_i) \odot \mathbb{G}_A) \oplus \mathbf{q}_{A,i}, & \text{if } i = t \text{ or } t+1 \\ \mathbf{o}, & \text{otherwise,} \end{cases} \quad (8)$$

where each $s(A_i)$ is a row vector of length k_A , \mathbb{G}_A is a fixed $k_A \times nR_{\min}/R_{\min}^c$ matrix, each \mathbf{X}_i and $\mathbf{q}_{A,i}$ is a row vector of length nR_{\min}/R_{\min}^c , and \mathbf{o} is the all-zero row vector (where “zero”, $\mathbf{o} \in \mathcal{F}$, is the additive identity of the field \mathcal{F}). If we say that a user i *does not transmit*, it sends $X_i = \mathbf{o}$. k_A is chosen such that

$$(k_A \log_2 |\mathcal{F}|)/n \geq R_{\min}, \quad (9)$$

so that we can define an injective (one-to-one) function that maps each A_i (of nR_{\min} bits) to a unique $s(A_i) \in \mathcal{F}^{k_A}$.

In the t -th sub-block for $L \leq t \leq L+D-1$, only one user, $d_{t-L+1} \in \mathcal{D}$, transmits using the linear code defined in (5):

$$\mathbf{X}_i = \begin{cases} (s(B_i) \odot \mathbb{G}_{B,i}) \oplus \mathbf{q}_{B,i}, & \text{if } i = d_{t-L+1} \\ \mathbf{o}, & \text{otherwise,} \end{cases} \quad (10)$$

[†]Since R_{\min} , R_{\min}^c , and $R'_{d_{t-L+1}}$ are rational numbers, there exists a (possibly large) n such that nR_{\min}/R_{\min}^c and $nR'_{d_{t-L+1}}/R_{\min}^c$ are integers.

where $\mathbf{s}(B_{d_{t-L+1}})$ is a row vector of length $k_{B,d_{t-L+1}}$, $\mathbb{G}_{B,d_{t-L+1}}$ is a fixed $k_{B,d_{t-L+1}} \times nR'_{d_{t-L+1}}/R_{\min}^c$ matrix, and each $\mathbf{X}_{d_{t-L+1}}$ and $\mathbf{q}_{B,d_{t-L+1}}$ is a fixed row vector of length $nR'_{d_{t-L+1}}/R_{\min}^c$. Similarly, $k_{B,d_{t-L+1}}$ is chosen such that

$$(k_{B,d_{t-L+1}} \log_2 |\mathcal{F}|)/n \geq R'_{d_{t-L+1}}, \quad (11)$$

so we can define an injective function that maps each $B_{d_{t-L+1}}$ (of $nR'_{d_{t-L+1}}$ bits) to a unique $\mathbf{s}(B_{d_{t-L+1}}) \in \mathcal{F}^{k_{B,d_{t-L+1}}}$.

Each element in \mathbb{G}_A , $\mathbb{G}_{B,d_{t-L+1}}$, $\mathbf{q}_{A,i}$, and $\mathbf{q}_{B,d_{t-L+1}}$ is independently and uniformly chosen over \mathcal{F} , and is fixed for all transmissions.

In the t -th sub-block for $1 \leq t \leq L-1$, the relay receives $\mathbf{Y}_0 = \mathbf{X}' \oplus \mathbf{N}_0$, where

$$\mathbf{X}' = \left([\mathbf{s}(A_t) \oplus \mathbf{s}(A_{t+1})] \odot \mathbb{G}_A \right) \oplus (\mathbf{q}_{A,t} \oplus \mathbf{q}_{A,t+1}), \quad (12)$$

which is also a linear codeword of the form (5). From Theorem 2, if nR_{\min}/R_{\min}^c is large enough and if

$$\frac{k_A \log_2 |\mathcal{F}|}{nR_{\min}/R_{\min}^c} < \log_2 |\mathcal{F}| - H(N_0), \quad (13)$$

then the relay can reliably decode the ‘‘message’’ $\mathbf{s}(A_t) \oplus \mathbf{s}(A_{t+1}) \triangleq \mathbf{s}(A_{t,t+1})$.

In the t -th sub-block for $L \leq t \leq L+D-1$, since only one user transmits, we directly apply Theorem 2. So, if

$$\frac{k_{B,d_{t-L+1}} \log_2 |\mathcal{F}|}{nR'_{d_{t-L+1}}/R_{\min}^c} < \log_2 |\mathcal{F}| - H(N_0), \quad (14)$$

then the relay can reliably decode $\mathbf{s}(B_{d_{t-L+1}})$.

Define $U \triangleq (\mathbf{s}(A_{1,2}), \mathbf{s}(A_{2,3}), \dots, \mathbf{s}(A_{L-1,L}), \mathbf{s}(B_{d_1}), \mathbf{s}(B_{d_2}), \dots, \mathbf{s}(B_{d_D}))$. On the uplink, if

$$R_{\min}^c < \log_2 |\mathcal{F}| - H(N_0), \quad (15)$$

we can always find sufficiently large n , k_A , and $\{k_{B,d}\}_{d \in \mathcal{D}}$, so that (9), (13) and (11), (14) can be satisfied in their respective sub-blocks. Hence, the relay can reliably decode U .

B. Downlink

Assume that the relay has correctly decoded U . Using the strategy of joint source-channel decoding over broadcast channels [8], the relay re-encodes U and sends it on n downlink channel uses. Each user i , for $i \in \mathcal{D}$, uses its *side information* $\mathbf{s}(B_i)$ to decode U (hence joint source-channel decoding). The users do not need to use their respective A_i in the decoding, as each A_i conveys little information about U . All users can reliably decode U if [8, Theorem 6]

$$H(U|\mathbf{s}(B_i)) < nI(X_0; Y_i), \quad \forall i \in \mathcal{D} \quad (16)$$

$$H(U) < nI(X_0; Y_i), \quad \forall i \notin \mathcal{D}, \quad (17)$$

for some $p(x_0)$. Choosing the uniform distribution for X_0 , $I(X_0; Y_i) = \log_2 |\mathcal{F}| - H(N_i)$.

Since the mapping from B_i (a random nR'_i -bit message) to $\mathbf{s}(B_i)$ is injective, $H(\mathbf{s}(B_i)) = H(B_i) = nR'_i$. Since $\mathbf{s}(A_{i,i+1}) \in \mathcal{F}^{k_A}$, $H(\mathbf{s}(A_{i,i+1})) \leq k_A \log_2 |\mathcal{F}|$, with equality iff $\mathbf{s}(A_{i,i+1})$ is uniformly distributed in \mathcal{F}^{k_A} . From Sec. V-A, $(k_A \log_2 |\mathcal{F}|)/n$ can be chosen arbitrarily close to R_{\min} . This gives $\frac{1}{n}H(U) \leq \frac{1}{n}[\sum_{i=1}^{L-1} H(\mathbf{s}(A_{i,i+1})) + \sum_{d \in \mathcal{D}} \mathbf{s}(B_d)] \leq$

$(L-1)R_{\min} + \sum_{d \in \mathcal{D}} R'_d = R_{\min}^c$, and $\frac{1}{n}H(U|\mathbf{s}(B_i)) \leq R_{\min}^c - R'_i = ([\sum_{j=1}^L R_j] - R_{\min} - R'_i) = R_i^c$. Note that for all $i \notin \mathcal{D}$, $R'_i = 0$ and hence $R_i^c = R_{\min}^c$. So, if

$$R_i^c < \log_2 |\mathcal{F}| - H(N_i), \quad \text{for all } 1 \leq i \leq L, \quad (18)$$

then (16) and (17) can both be satisfied. Note that on the downlink, linear codes are not required.

C. The Capacity of the Binary MWRC

If the rate constraints (15) and (18) are satisfied, all users are able to decode U reliably. Each user i then performs:

$$\begin{aligned} \mathbf{s}(A_{i+1}) &= \mathbf{s}(A_{i,i+1}) \oplus -\mathbf{s}(A_i), \\ \mathbf{s}(A_{i+2}) &= \mathbf{s}(A_{i+1,i+2}) \oplus -\mathbf{s}(A_{i+1}), \quad \dots, \\ \mathbf{s}(A_L) &= \mathbf{s}(A_{L-1,L}) \oplus -\mathbf{s}(A_{L-1}), \\ \mathbf{s}(A_{i-1}) &= \mathbf{s}(A_{i-1,i}) \oplus -\mathbf{s}(A_i), \\ \mathbf{s}(A_{i-2}) &= \mathbf{s}(A_{i-2,i-1}) \oplus -\mathbf{s}(A_{i-1}), \quad \dots, \\ \mathbf{s}(A_1) &= \mathbf{s}(A_{1,2}) \oplus -\mathbf{s}(A_2), \end{aligned} \quad (19)$$

and obtains $(A_1, A_2, \dots, A_{i-1}, A_{i+1}, \dots, A_L)$. Combining these with $(B_{d_1}, B_{d_2}, \dots, B_{d_D})$, each user i can reliably recover the messages of all other users, i.e., $(W_1, W_2, \dots, W_{i-1}, W_{i+1}, \dots, W_L)$.

So, all rate tuples (R_1, R_2, \dots, R_L) satisfying (15) and (18) are achievable. Since the closure of this region coincides with the capacity upper bound given in Theorem 1, we have:

Theorem 3: Consider an L -user finite field adder MWRC over \mathcal{F} . The capacity is given by all rate tuples (R_1, R_2, \dots, R_L) that satisfy (1) and (2) for all $1 \leq i \leq L$.

Remark 1: The capacity-achieving FDF does not utilize the users' received signals in their transmission. Hence, *feedback* does not increase the capacity of the finite field adder MWRC.

D. A Note on the Common-Rate Capacity

If we consider only the common rate, $R = R_i, \forall i$, we have $W_i = A_i$ and $B_i = \emptyset, \forall i$. In this case, rate splitting is not required on the uplink to get (15). Furthermore, on the downlink, since $U = (\mathbf{s}(A_{1,2}), \mathbf{s}(A_{2,3}), \dots, \mathbf{s}(A_{L-1,L}))$ has no correlation with any W_i , utilizing W_i does not help the user in decoding U . On the downlink, the relay encodes U , of $n(L-1)R$ bits, and transmits it in n channel uses. Treating the downlink from the relay to each user i as a point-to-point channel [9, p. 200], if $n(L-1)R < nI(X_0; Y_i)$, then user i can reliably decode U from its received signals without needing to use its own message (separate source-channel decoding). Hence, we get (18). Of course, after decoding U , each user needs to use its message to obtain the other users' messages using the steps in (19). But as far as channel decoding is concerned, the source messages need not be used. So, if we are only interested in the common rate case, FDF without rate splitting and separate source-channel coding is optimal (capacity-achieving) for the finite field adder MWRC.

VI. COMPARISON OF CODING STRATEGIES

Now, we compare three coding strategies for the special case when $L = 2$ and $\mathcal{F} = \{0, 1\}$, i.e., the binary TWRC. For binary N_i , we denote $\Pr\{N_i = 1\} = \rho_i$ and $H(\alpha) = -\alpha \log_2 \alpha - (1-\alpha) \log_2 (1-\alpha)$.

A. FDF with joint source-channel coding

From Theorem 3, FDF with rate splitting and joint source-channel coding achieves the capacity given by $\{(R_1, R_2) : R_1, R_2 \leq 1 - H(\rho_0), R_1 \leq 1 - H(\rho_2), R_2 \leq 1 - H(\rho_1)\}$. The capacity of the binary TWRC was reported in [11], [5].

B. FDF with separate source-channel coding

Now, we find the achievable rate region using FDF with rate splitting but with *separate* source-channel coding. The coding on the uplink is the same as that in Sec. V-A. Assuming $R_2 \geq R_1$, we have $W_1 = A_1$ and $W_2 = (A_2, B_2)$. So, on the uplink, if $R_2 < 1 - H(\rho_0)$, then the relay can reliably decode $(s(A_{1,2}), s(B_2))$. Instead of using the joint source-channel coding for the downlink described in Sec. V-B, we re-cast the downlink as a *broadcast channel with degraded message sets*, where the relay broadcasts a common message $s(A_{1,2})$ to both the users, and a private message $s(B_2)$ to user 1, and the users do not use their own messages for decoding $s(A_{1,2})$ and $s(B_2)$ (hence separate source-channel decoding). From [12], if $R_1 < 1 - H(\beta(1 - \rho_2) + (1 - \beta)\rho_2)$, $R'_2 < H(\beta(1 - \rho_1) + (1 - \beta)\rho_1) - H(\rho_1)$, and $R_1 + R'_2 < 1 - H(\rho_1)$, for some $0 \leq \beta \leq \frac{1}{2}$, then both users can reliably decode $s(A_{1,2})$ and user 1 can reliably decode $s(B_2)$ purely from their respective Y_i . The users then follow the steps in (19) to obtain the other user's message. Repeating this for the case $R_1 \geq R_2$, the achievable rate region is the convex hull of:

- \mathcal{R}_1 : all rate pairs $(R_1, R_1 + R'_2)$ satisfying
$$R_1 < 1 - H(\beta(1 - \rho_2) + (1 - \beta)\rho_2) \quad (20)$$

$$R'_2 < H(\beta(1 - \rho_1) + (1 - \beta)\rho_1) - H(\rho_1) \quad (21)$$

$$R_1 + R'_2 < 1 - \max\{H(\rho_0), H(\rho_1)\}, \quad (22)$$

for some $0 \leq \beta \leq \frac{1}{2}$, and

- \mathcal{R}_2 : all rate pairs $(R_2 + R'_1, R_2)$ satisfying
$$R_2 < 1 - H(\alpha(1 - \rho_1) + (1 - \alpha)\rho_1) \quad (23)$$

$$R'_1 < H(\alpha(1 - \rho_2) + (1 - \alpha)\rho_2) - H(\rho_2) \quad (24)$$

$$R_2 + R'_1 < 1 - \max\{H(\rho_0), H(\rho_2)\}, \quad (25)$$

for some $0 \leq \alpha \leq \frac{1}{2}$.

C. Complete-Decode-Forward

Using CDF, the relay fully decodes both W_1 and W_2 on the uplink, which is a multiple-access channel. So, if $R_1 < 1 - H(\rho_0)$, $R_2 < 1 - H(\rho_0)$, $R_1 + R_2 < 1 - H(\rho_0)$, then the relay can reliably decode W_1 and W_2 [13], [14]. Note that the last inequality implies the first two. Assuming that the relay has successfully decoded W_1 and W_2 , it broadcasts (W_1, W_2) on the downlink. Using a joint source-channel decoding, each user i , $i = 1, 2$, can reliably decode the other user's message from their respective received signals Y_i and their own messages W_i if $R_1 < 1 - H(\rho_2)$ and $R_2 < 1 - H(\rho_1)$ [15], [16]. Combining the uplink and the downlink constraints, the achievable rate region using CDF is all (R_1, R_2) satisfying:

$$R_1 < 1 - H(\rho_2), \quad R_2 \leq 1 - H(\rho_1), \quad (26)$$

$$R_1 + R_2 < 1 - H(\rho_0). \quad (27)$$

D. Discussion

Using CDF, the relay needs to fully decode the users' messages on the uplink, and this restricts the sum rate to be constrained by the uplink bandwidth, c.f. (27). So, CDF is not *uplink optimized*. On the other hand, using FDF with rate splitting and separate source-channel coding, the users' *a priori* knowledge about their own messages is not utilized during the channel decoding on the downlink – their own messages are used only *after* channel decoding. So, FDF with separate source-channel coding is not *downlink optimized*. These two coding strategies do not achieve the capacity of the finite field adder MWRC in general. FDF with rate splitting and joint source-channel coding overcomes these two shortcomings by having the relay decode only functions of the source messages on the uplink and having the users utilize their own messages in channel decoding on the downlink. This strategy indeed achieves the capacity of the finite field adder MWRC. This work suggests that for the general MWRC, functional decoding and joint source-channel coding should be utilized.

REFERENCES

- [1] R. Knopp, "Two-way radio networks with a star topology," in *Proc. Int. Zurich Seminar on Commun. (IZS)*, Zurich, Switzerland, Feb. 22-24 2006, pp. 154–157.
- [2] B. Rankov and A. Wittneben, "Achievable rate regions for the two-way relay channel," in *Proc. IEEE Int. Symposium on Inf. Theory (ISIT)*, Seattle, USA, Jul. 9-14 2006, pp. 1668–1672.
- [3] —, "Spectral efficient protocols for half-duplex fading relay channels," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 379–389, Feb. 2007.
- [4] D. Gündüz, A. Yener, A. Goldsmith, and H. V. Poor, "The multi-way relay channel," in *Proc. IEEE Int. Symposium on Inf. Theory (ISIT)*, Seoul, Korea, Jun. 28-Jul. 3 2009, pp. 339–343.
- [5] W. Nam, S. Chung, and Y. H. Lee, "Capacity bounds for two-way relay channels," in *Proc. Int. Zurich Seminar on Commun. (IZS)*, Zurich, Switzerland, Mar. 12-14 2008, pp. 144–147.
- [6] L. Ong, S. J. Johnson, and C. M. Kellett, "An optimal coding strategy for the binary multi-way relay channel," *IEEE Commun. Lett.*, vol. 14, no. 4, pp. 330–332, Apr. 2010.
- [7] L. Ong, C. M. Kellett, and S. J. Johnson, "Capacity theorems for the AWGN multi-way relay channel," in *Proc. IEEE Int. Symposium on Inf. Theory (ISIT)*, Austin, USA, Jun. 13-18 2010, pp. 664–668.
- [8] E. Tuncel, "Slepian-Wolf coding over broadcast channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1469–1482, Apr. 2006.
- [9] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Wiley-Interscience, 2006.
- [10] R. G. Gallager, *Information Theory and Reliable Communication*. Wiley, 1968.
- [11] R. Knopp, "Two-way wireless communication via a relay station," in *GDR-ISIS Meeting*, Paris, France, Mar. 29 2007.
- [12] J. Körner and K. Marton, "General broadcast channels with degraded message sets," *IEEE Trans. Inf. Theory*, vol. IT-23, no. 1, pp. 60–64, Jan. 1977.
- [13] R. Ahlswede, "Multi-way communication channels," in *Proc. IEEE Int. Symposium on Inf. Theory (ISIT)*, Tsahkadsor Armenia, USSR, Sep. 2-8 1971, pp. 23–52.
- [14] H. Liao, "A coding theorem for multiple access communication," in *Proc. IEEE Int. Symposium on Inf. Theory (ISIT)*, Asilomar, USA, Jan. 1972.
- [15] G. Kramer and S. Shamai, "Capacity for classes of broadcast channels with receiver side information," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Lake Tahoe, USA, Sep. 2-6 2007, pp. 313–318.
- [16] T. J. Oechtering, C. Schnurr, and H. Boche, "Broadcast capacity region of two-phase bidirectional relaying," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 454–458, Jan. 2008.