

Functional-Decode-Forward for the General Discrete Memoryless Two-Way Relay Channel

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

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

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

Abstract—We consider the general discrete memoryless two-way relay channel, where two users exchange messages via a relay, and propose two functional-decode-forward coding strategies for this channel. Functional-decode-forward involves the relay decoding a function of the users' messages rather than the individual messages themselves. This function is then broadcast back to the users, which can be used in conjunction with the user's own message to decode the other user's message. Via a numerical example, we show that functional-decode-forward with linear codes is capable of achieving strictly larger sum rates than those achievable by other strategies.

I. INTRODUCTION

We obtain two new achievable rate regions for the general discrete memoryless two-way relay channel (TWRC), in which two users exchange messages through a relay. We consider TWRCs with no direct link between the users (see Fig. 1). The new rate regions are obtained using the idea of *functional-decode-forward* (FDF), where the relay only decodes a function of the users' messages or codewords without needing to decode the messages or codewords themselves (hence saving the *uplink* bandwidth from the users to the relay). The relay then broadcasts the function to both users. The function must be defined such that knowing its own message, each user is able to decode the message sent by the other user.

We first illustrate the concept of FDF using the *noiseless* binary adder TWRC as an example, where nodes 1 and 2 (the users) exchange data through node 3 (the relay). Let $X_i \in \{0, 1\}$ be node i 's transmitted signal and $Y_i \in \{0, 1\}$ be node i 's received signal. The noiseless binary adder TWRC is defined as follows: (i) the uplink is $Y_3 = X_1 + X_2 \pmod 2$, and (ii) the downlink is $Y_1 = X_3$ and $Y_2 = X_3$. Assume that the source messages are in bits, i.e. $W_1, W_2 \in \{0, 1\}$. The well-known optimal (rate-maximizing) coding strategy is for the users to transmit uncoded information bits, i.e., $X_i = W_i$, for $i \in \{1, 2\}$, and for the relay to forward its received bits, i.e., $Y_3 = X_3$. Having received Y_1 which is $W_1 + W_2 \pmod 2$, and knowing its own message W_1 , node 1 can recover W_2 perfectly. Node 2 can recover W_1 similarly. Here, the capacity of 1 bit/channel use is achievable using this strategy.

While the bit-wise modulo-two addition of the users' messages seems to be a good function for the relay to transmit, the main challenge of FDF on a *noisy* TWRC lies in:

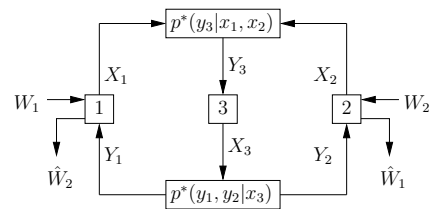


Fig. 1. The general discrete memoryless TWRC considered in this paper

- selecting a *good function* of the users' messages/codewords which the relay should decode, and
- constructing *good codes* for the users such that the relay can efficiently decode this function without needing to decode the individual users' messages/codewords.

In the case of *adder channels*, e.g., $Y_3 = X_1 + X_2 + N$, where N is the channel noise, linear codes can be used (see [1] for the case of binary adder channels, [2] for finite field adder channels, and [3], [4] for AWGN channels). Let \mathbf{X}_i be user i 's length- n linear codeword¹, for $i \in \{1, 2\}$. The structure of linear codes guarantees that $\mathbf{U} \triangleq (\mathbf{X}_1 + \mathbf{X}_2)$ is a codeword from the same code. The relay effectively receives $\mathbf{Y}_3 = \mathbf{U} + \mathbf{N}$, which is a noisy version of \mathbf{U} . Capacity-achieving linear codes have been shown to exist for this type of additive noise channel. This means if the users transmit using these linear codes, then the relay is able to efficiently decode \mathbf{U} (which is a function of the users' codewords) without having to decode the users' codewords individually. The relay then broadcasts \mathbf{U} to the users, and each user can obtain the other user's message from \mathbf{U} and its own message/codeword.

For the above adder channels, the channels actually perform the desired function by adding the users' codewords. For FDF on *general* discrete memoryless TWRCs in which the channels do not "help", it is not immediately obvious what function the relay should decode, and how the relay can decode the function without first decoding the individual messages.

In this paper, we use random linear codes for FDF on the general discrete memoryless TWRC following the idea in [5] for the multiple-access channel, i.e., the users transmit randomly generated linear codewords on the uplink. Although the uplink output Y_3 cannot be written as a (noisy) function of $X_1 + X_2$, by invoking the Markov Lemma we will prove that the relay is still able to *reliably* (i.e., with arbitrarily

This work is supported by the Australian Research Council under grants DP087725 and DP1093114.

¹Bold letters are used to denote a block of n channel uses, e.g. $\mathbf{X}_i = (X_i[1], X_i[2], \dots, X_i[n])$, where $X_i[t]$ is X_i on the t -th channel use.

small error probability) decode $\mathbf{X}_1 + \mathbf{X}_2$ without needing to decode the individual messages/codewords. The relay then broadcasts $\mathbf{X}_1 + \mathbf{X}_2$ to the users for each of them to obtain the other user's message. We call this strategy functional-decode-forward with linear codes (FDF-L).

Another method for the relay to decode a function of the users' messages in the general discrete memoryless TWRC is by using *systematic computation codes* [6] on the uplink. On the uplink, the users first send uncoded data, followed by linear-coded signals. After the relay decodes a function of the users' messages, the downlink transmission is the same as that in FDF-L. We call this strategy functional-decode-forward with systematic computation codes (FDF-S).

We will first derive two achievable rate regions for the general discrete memoryless TWRC, using FDF-L and FDF-S. We will then show, using an example, that FDF-L can achieve higher sum rates than those achievable by FDF-S and by existing coding strategies for the TWRC, including (i) the *complete-decode-forward* (CDF) coding strategy², where the relay fully decodes the messages from both users, re-encodes and broadcasts a function of the messages back to the users [7], [8], and (ii) the *compress-forward* (CF) coding strategy, where the relay quantizes its received signals, re-encodes and broadcasts the quantized signals to users [9].

II. CHANNEL MODEL

Fig. 1 depicts the general discrete memoryless TWRC considered in the paper, where users 1 and 2 exchange data through the relay (node 3). We denote by $X_i \in \mathcal{X}_i$ the channel input from node i , $Y_i \in \mathcal{Y}_i$ the channel output received by node i , and $W_i \in \mathcal{W}_i$ user i 's message. The TWRC can be completely defined by (i) the uplink channel $p^*(y_3|x_1, x_2)$, and (ii) the downlink channel $p^*(y_1, y_2|x_3)$.

Let $W_i \in \{1, 2, \dots, 2^{nR_i}\}$ be an (nR_i) -bit message, for $i \in \{1, 2\}$. Consider on each uplink and downlink, n channel uses. User i transmits $\mathbf{X}_i(W_i) = f_i(W_i)$, for $i \in \{1, 2\}$. At any time, the relay transmits a function of its previously received signals, i.e., $X_3[t] = f_{3,t}(Y_3[1], Y_3[2], \dots, Y_3[t-1])$, for $t \in \{1, 2, \dots, n\}$. After n channel uses, each user estimates the message of the other user from its received signals and its own message, i.e., $\hat{W}_2 = g_1(\mathbf{Y}_1, W_1)$ and $\hat{W}_1 = g_2(\mathbf{Y}_2, W_2)$ for users 1 and 2 respectively. \hat{W}_2 is node 1's estimate of W_2 , and vice versa. Assuming that the message pair (W_1, W_2) is uniformly distributed in $\mathcal{W}_1 \times \mathcal{W}_2$, a rate pair (R_1, R_2) is said to be *achievable* if each user can reliably decode the messages of the other. We say that a user can reliably decode a message if the probability that it wrongly decodes the message can be made arbitrarily small.

III. FIELDS AND LINEAR CODES

Now, we will present a construction of *random* linear codes with elements from finite fields. Random linear codes will be used for the users to transmit their respective messages to the

²The strategy is commonly known as decode-and-forward or decode-forward. We modified the name of this strategy here to reflect that the relay completely decodes both the users' messages before forwarding them.

relay. Using random linear codes, any two codewords are statistically pair-wise independent. This property is important for proving reliable communications. Furthermore, as mentioned in Sec. I, the structure of linear codes enables the relay to decode the desired function of the user's codewords without needing to decode the individual codewords or messages.

Let \mathcal{F} be a finite field with associated operations of addition \oplus and multiplication \odot . Consider the following codeword generating function that maps a message $\mathbf{s} \in \mathcal{F}^k$ to a codeword $\mathbf{x} \in \mathcal{F}^n$:

$$\mathbf{x} = (\mathbf{s} \odot \mathbb{G}) \oplus \mathbf{q}, \quad (1)$$

where \mathbf{x} is a row vector of length n , \mathbf{s} is a row vector of length k , \mathbb{G} is a fixed k -by- n matrix, with each element independently and uniformly chosen over \mathcal{F} , and \mathbf{q} is a fixed row vector of length n , with each element independently and uniformly chosen over \mathcal{F} . We extend Gallager's results for binary linear codes [10, p. 207] to finite-field linear codes in the following two lemmas (see [2] for the proofs):

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

Lemma 2: Consider the linear codes defined in (1). Let \mathbf{s}_1 and \mathbf{s}_2 be any two different messages. The corresponding codewords $\mathbf{x}_1 = (\mathbf{s}_1 \odot \mathbb{G}) \oplus \mathbf{q}$ and $\mathbf{x}_2 = (\mathbf{s}_2 \odot \mathbb{G}) \oplus \mathbf{q}$ are independent.

Besides the above extensions of Gallager's results for binary linear codes, we have the following additional result.

Lemma 3: Consider two linear codes: $\mathbf{x} = \mathbf{s}_1 \odot \mathbb{G} \oplus \mathbf{q}_1$, and $\mathbf{v} = \mathbf{s}_2 \odot \mathbb{G} \oplus \mathbf{q}_2$, where \mathbb{G} , \mathbf{q}_1 , and \mathbf{q}_2 are independently generated according to the uniform distribution. Any two codewords, one from each code, are independent.

Proof: From Lemma 1, we know that $p(\mathbf{x}) = |\mathcal{F}|^{-n}$ and $p(\mathbf{v}) = |\mathcal{F}|^{-n}$. We have to show that $p(\mathbf{x}, \mathbf{v}) = |\mathcal{F}|^{-2n}$. Elements in \mathbb{G} , \mathbf{q}_1 , and \mathbf{q}_2 are independent and uniformly distributed, and so each $(\mathbb{G}, \mathbf{q}_1, \mathbf{q}_2)$ has a probability of $|\mathcal{F}|^{-n(k+2)}$ of being selected. For any given \mathbb{G} , there is only one \mathbf{q}_1 and one \mathbf{q}_2 that results in the given \mathbf{x} and \mathbf{v} . So, there are only $|\mathcal{F}|^{-nk}$ different $(\mathbb{G}, \mathbf{q}_1, \mathbf{q}_2)$ that map \mathbf{s}_1 to \mathbf{x} , and \mathbf{s}_2 to \mathbf{v} . So, $p(\mathbf{x}, \mathbf{v}) = |\mathcal{F}|^{-nk} |\mathcal{F}|^{-n(k+2)} = |\mathcal{F}|^{-2n}$. ■

Remark 1: The use of dither \mathbf{q} in the above-defined linear codes of the form (1) is essential in proving Lemmas 1 and 3.

IV. TWO NEW ACHIEVABLE RATE REGIONS

A. Functional-Decode-Forward with Linear Codes (FDF-L)

We first prove the following achievable rate region for the discrete memoryless TWRC using FDF-L:

Theorem 1: Consider a TWRC where $|\mathcal{X}_1| = |\mathcal{X}_2| = |\mathcal{F}|$, for some finite field \mathcal{F} . Rename³ the elements in \mathcal{X}_1 and \mathcal{X}_2 so that $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{F}$. The rate pair (R_1, R_2) is achievable if

$$R_1, R_2 \leq \min\{I(U; Y_3), I(X_3; Y_1), I(X_3; Y_2)\}, \quad (2)$$

for

$$p(u, y_3) = \frac{1}{|\mathcal{F}|^2} \sum_{\substack{\mathbf{x}_1, \mathbf{x}_2 \\ \text{s.t. } \mathbf{x}_1 \oplus \mathbf{x}_2 = u}} p^*(y_3|x_1, x_2), \quad (3)$$

³We choose a renaming scheme that maximizes $I(U; Y_3)$.

where $U \in \mathcal{F}$, and for $p(x_3, y_1, y_2) = p(x_3)p^*(y_1, y_2|x_3)$.

Remark 2: Recall that \mathcal{F} is a finite field iff $|\mathcal{F}| = \ell^z$ for some $\ell \in \mathbb{Z}_p$ (prime numbers) and some $z \in \mathbb{Z}_+$ (positive integers). For TWRCs where $|\mathcal{X}_1| \neq |\mathcal{X}_2|$ or $|\mathcal{X}_1| = |\mathcal{X}_2| \neq \ell^z$, $\forall \ell \in \mathbb{Z}_p, \forall z \in \mathbb{Z}_+$, we select subsets $\mathcal{X}'_1 \subseteq \mathcal{X}_1$ and $\mathcal{X}'_2 \subseteq \mathcal{X}_2$ such that $|\mathcal{X}'_1| = |\mathcal{X}'_2| = \ell^z$, for some $\ell \in \mathbb{Z}_p, z \in \mathbb{Z}_+$. The result in Theorem 1 holds for any discrete memoryless TWRC with X_i replaced by $X'_i \in \mathcal{X}'_i$, for $i \in \{1, 2\}$.

1) *The Auxiliary Random Variable U :* The auxiliary random variable $U \in \mathcal{F}$ is the information that the relay recovers from its received signal Y_3 , and broadcasts to both users 1 and 2. Before proceeding to the proof of Theorem 1, we derive $p(u, y_3)$ in (3). Define

$$U \triangleq X_1 \oplus X_2. \quad (4)$$

This means $p(u|x_1, x_2) = \mathbf{1}(u = x_1 \oplus x_2)$, where $\mathbf{1}(E)$ is 1 if the event E is true, and is 0 otherwise. We can write

$$p(u, x_1, x_2, y_3) = p(x_1, x_2)p^*(y_3|x_1, x_2)\mathbf{1}(u = x_1 \oplus x_2),$$

meaning $U - (X_1, X_2) - Y_3$ forms a Markov chain.

Remark 3: The function that the relay should decode, U , is not unique. Other functions are possible as long as each user can obtain the other user's message from the function and its own message.

In FDF-L, we use the linear code structure in (1) with the same \mathbb{G} , and independent \mathbf{q}_1 and \mathbf{q}_2 for users 1 and 2 respectively (see (6) for code construction). So, for any pair of source messages, from Lemma 3, we have $p(x_1, x_2) = \frac{1}{|\mathcal{F}|^2}$, i.e., X_1 and X_2 are independent of each other and are uniformly distributed. Hence,

$$p(u, y_3) = \sum_{x_1, x_2} p(u, x_1, x_2, y_3) \quad (5a)$$

$$= \sum_{x_1, x_2} \frac{1}{|\mathcal{F}|^2} \mathbf{1}(u = x_1 \oplus x_2) p^*(y_3|x_1, x_2) \quad (5b)$$

$$= \frac{1}{|\mathcal{F}|^2} \sum_{\substack{x_1, x_2 \\ \text{s.t. } x_1 \oplus x_2 = u} p^*(y_3|x_1, x_2). \quad (5c)$$

2) *Proof of Theorem 1:* Fig. 2 depicts the relationship among the random variables for FDF-L used to achieve the rates in Theorem 1.

Uplink:

Assuming that $|\mathcal{W}_1| = |\mathcal{W}_2| = 2^{nR}$, on the uplink, for a sufficiently large n , we choose k such that $|\mathcal{F}|^k = 2^{nR}$, and define a bijective mapping from $W_i \in \{1, 2, \dots, 2^{nR}\}$ to $\mathbf{S}_i \in \mathcal{F}^k$, for $i \in \{1, 2\}$. The users transmit using linear codes of the form (1), i.e.,

$$\mathbf{X}_i(W_i) = (\mathbf{S}_i \odot \mathbb{G}) \oplus \mathbf{q}_i, \quad (6)$$

for $i \in \{1, 2\}$, where $\mathbf{S}_i \in \mathcal{F}^k$, $\mathbb{G} \in \mathcal{F}^{k \times n}$, and $\mathbf{X}_i, \mathbf{q}_i \in \mathcal{F}^n$. All elements in \mathbb{G} , \mathbf{q}_1 , and \mathbf{q}_2 are uniformly and independently chosen over \mathcal{F} , and are fixed for all channel uses.

From definition (4) we have $U \triangleq \mathbf{X}_1 \oplus \mathbf{X}_2 = (\mathbf{S}_3 \odot \mathbb{G}) \oplus \mathbf{q}_3$, which is also a codeword from a linear code of the form (1), where $\mathbf{S}_3 = \mathbf{S}_1 \oplus \mathbf{S}_2$ is the "message" and $\mathbf{q}_3 = \mathbf{q}_1 \oplus \mathbf{q}_2$. The relay estimates the codeword $U(V_3)$ from its received

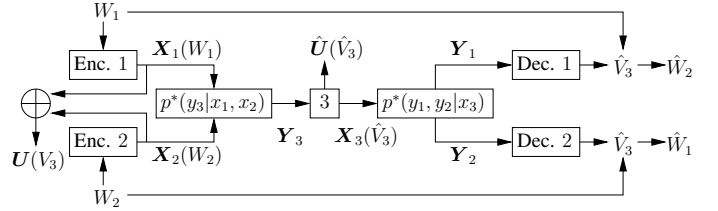


Fig. 2. Relationship among the random variables in FDF-L

signals \mathbf{Y}_3 , where $V_3 \in \{1, \dots, 2^{nR}\}$ is mapped (bijectively) to $\mathbf{S}_3 \in \mathcal{F}^k$ the same way as W_i is to \mathbf{S}_i , for $i \in \{1, 2\}$.

Definition 1: The jointly strongly δ -typical set $T_{[XY]\delta}^n$ with respect to a distribution $p(x, y)$ on $\mathcal{X} \times \mathcal{Y}$ is the set of sequences $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$ such that $N(a, b; \mathbf{x}, \mathbf{y}) = 0$ for $p(a, b) = 0$, and $\sum_{a \in \mathcal{X}} \sum_{b \in \mathcal{Y}} \frac{1}{n} N(a, b; \mathbf{x}, \mathbf{y}) - p(x, y) \leq \delta$, where $N(a, b; \mathbf{x}, \mathbf{y})$ is the number of occurrences of the pair of symbols (a, b) in the pair of sequences (\mathbf{x}, \mathbf{y}) , δ is an arbitrarily small positive real number, and the sequences in $T_{[XY]\delta}^n$ are called strongly jointly δ -typical sequences.

The relay's estimate \hat{V}_3 is the *unique* sequence $U(\hat{V}_3)$ that is jointly strongly δ -typical with its received sequence \mathbf{Y}_3 , i.e., the relay finds \hat{v}_3 such that $(U(\hat{v}_3), \mathbf{Y}_3) \in T_{[UY_3]\delta}^n$, and such that there is no $v' \neq \hat{v}_3$ where $(U(v'), \mathbf{Y}_3) \in T_{[UY_3]\delta}^n$.

Now, we bound the probability that the relay wrongly decodes V_3 . Let $W_1 = a_1$ and $W_2 = a_2$ be the transmitted messages, and $V_3 = b$ the corresponding index for U . Let

$$E_0 = \left\{ (U(b), \mathbf{Y}_3) \notin T_{[UY_3]\delta}^n \right\} \quad (7)$$

$$E_1 = \left\{ \exists v'_3 \neq b : (U(v'_3), \mathbf{Y}_3) \in T_{[UY_3]\delta}^n \right\} \quad (8)$$

$$E_2 = E_0 \cup E_1. \quad (9)$$

E_2 is the event that node 3 wrongly decodes V_3 . In addition, we define the following event:

$$E_3 = \left\{ (\mathbf{X}_1(a_1), \mathbf{X}_2(a_2), \mathbf{Y}_3) \in T_{[X_1 X_2 Y_3]\delta}^n \right\}. \quad (10)$$

E_3 is the event that the received signal \mathbf{Y}_3 is jointly strongly δ -typical with the users' codewords $\mathbf{X}_1(a_1)$ and $\mathbf{X}_2(a_2)$.

Since the users' codewords \mathbf{X}_1 and \mathbf{X}_2 share the same \mathbb{G} but have independently generated \mathbf{q}_1 and \mathbf{q}_2 , from Lemma 3, the codewords are independent. Furthermore, from Lemma 1, the codeletters $\{X_i[t] : i \in \{1, 2\}, t \in \{1, 2, \dots, n\}\}$ are also independent. From [11, Theorem 6.9], for a sufficiently large n , we have $\Pr\{E_3\} > 1 - \delta$, meaning that $(\mathbf{X}_1(a_1), \mathbf{X}_2(a_2), \mathbf{Y}_3)$ is jointly strongly δ -typical with probability tending to one, by choosing a sufficiently small $\delta > 0$.

So, for a sufficiently large n , $\Pr\{E_0^c\}$ equals

$$\begin{aligned} & \Pr \left\{ (U(b), \mathbf{Y}_3) \in T_{[UY_3]\delta}^n \right\} \\ &= \Pr \left\{ E_3^c \right\} \Pr \left\{ (U(b), \mathbf{Y}_3) \in T_{[UY_3]\delta}^n \mid E_3^c \right\} \\ & \quad + \Pr \left\{ E_3 \right\} \Pr \left\{ (U(b), \mathbf{Y}_3) \in T_{[UY_3]\delta}^n \mid E_3 \right\} \end{aligned} \quad (11a)$$

$$> \alpha + (1 - \delta) \Pr \left\{ (U(b), \mathbf{Y}_3) \in T_{[UY_3]\delta}^n \mid E_3 \right\} \quad (11b)$$

$$> \alpha + (1 - \delta)(1 - \epsilon), \quad (11c)$$

for some arbitrarily small $\epsilon > 0$, where $\alpha \triangleq \Pr\{E_3^c\} \Pr\{U(b), \mathbf{Y}_3 \in T_{[UY_3|\mathcal{F}]\delta}^n[E_3^c]\} \leq \Pr\{E_3^c\} < \delta$. Eqn. (11c) follows from the Markov Lemma [12, page 202 (Lemma 4.1)] because $U - (X_1, X_2) - Y_3$ forms a Markov chain.

Remark 4: Note that $(\mathbf{X}_1, \mathbf{X}_2, \mathbf{Y}_3)$ being jointly strongly δ -typical does not imply that (U, \mathbf{Y}_3) is jointly strongly δ -typical. However, since $U - (X_1, X_2) - Y_3$ forms a Markov chain, invoking the Markov lemma yields that (U, \mathbf{Y}_3) is jointly strongly δ -typical with probability tending to one.

It follows that

$$\Pr\{E_0\} = 1 - \{E_0^c\} < \delta + \epsilon - \delta\epsilon - \alpha < \epsilon_0, \quad (12)$$

for some arbitrarily small $\epsilon_0 > 0$, by choosing a sufficiently small δ .

Now, from Lemma 2, for any $v'_3 \neq b$, $U(v'_3)$ and $U(b)$ are independent, and hence $U(v'_3)$ and \mathbf{Y}_3 are also independent. So, we have $\Pr\{E_1\}$ equals

$$\Pr\left\{\exists v'_3 \neq b : (U(v'_3), \mathbf{Y}_3) \in T_{[UY_3|\mathcal{F}]\delta}^n\right\} \leq \sum_{v'_3 \in \{1, 2, \dots, 2^{nR}\} \setminus \{b\}} \Pr\left\{(U(v'_3), \mathbf{Y}_3) \in T_{[UY_3|\mathcal{F}]\delta}^n\right\} \quad (13a)$$

$$= (2^{nR} - 1) \Pr\left\{(U(v'_3), \mathbf{Y}_3) \in T_{[UY_3|\mathcal{F}]\delta}^n\right\} \quad (13b)$$

$$\leq (2^{nR} - 1) 2^{-n[I(U; Y_3) - \tau]} \quad (13c)$$

$$< 2^{-n[I(U; Y_3) - \tau - R]} \leq \epsilon_1, \quad (13d)$$

for some arbitrarily small $\epsilon_1 > 0$ if n is sufficiently large and if $R < I(U; Y_3) - \tau$, where $\tau \rightarrow 0$ as $|\mathcal{F}|\delta \rightarrow 0$. Here (13a) is by the union bound, and (13c) follows from [11, Lemma 7.17] as $U(v'_3)$ and \mathbf{Y}_3 are independent.

Hence, if

$$R < I(U; Y_3), \quad (14)$$

for $p(u, y_3)$ defined in (5c), then $\Pr\{E_2\} = \Pr\{E_0 \cup E_1\} \leq \Pr\{E_0\} + \Pr\{E_1\} < \epsilon_0 + \epsilon_1$, where $\epsilon_0 + \epsilon_1$ can be made arbitrarily small, i.e., the relay can *reliably* decode $\hat{V}_3 = b$.

Downlink:

Assuming that the relay has correctly decoded $U(V_3)$, it re-encodes and broadcasts the index $V_3 \in \{1, 2, \dots, 2^{nR}\}$ to the users in n downlink channel uses. For n sufficiently large, the users can reliably decode V_3 if [13, p. 567 (Theorem 15.6.3)]

$$R < I(X_3; Y_1) \quad \text{and} \quad R < I(X_3; Y_2), \quad (15)$$

for some $p(x_3)p^*(y_1, y_2|x_3)$. Note that linear codes are not required on the downlink.

Assuming node 1 correctly decodes the relay's message V_3 , knowing its own message W_1 , it can perform $U(V_3) \oplus (-\mathbf{X}_1(W_1))$ to get W_2 , where $(-\mathbf{X}_1)$ is the element-wise additive inverse of \mathbf{X}_1 . Node 2 decodes W_1 using a similar method. Combining (14) and (15), we have Theorem 1. ■

B. Functional-Decode-Forward with Systematic Computation Codes (FDF-S)

An achievable rate region for the discrete memoryless TWRC using FDF can also be obtained by using systematic computation codes [6] (instead of linear codes) on the

uplink. Similar to FDF-L, the relay computes a function of the users' codewords (the function $\mathbf{X}_1(W_1) \oplus \mathbf{X}_2(W_2)$ can again be chosen) and broadcasts this function back to the users. However, on the uplink, using systematic computation codes, the users first send uncoded transmissions to the relay, followed by a refinement stage in which the users send linear-coded transmissions. We can show that the rate region in the following theorem is achievable for the TWRC.

Theorem 2: Consider a TWRC where $|\mathcal{X}_1| = |\mathcal{X}_2| = |\mathcal{F}|$, for some finite field \mathcal{F} . Rename the elements in \mathcal{X}_1 and \mathcal{X}_2 so that $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{F}$. The rate pair (R_1, R_2) is achievable if

$$R_1 \leq \left[\frac{C_{\text{MAC}} H(W_1)}{C_{\text{MAC}} + 2H(X_1 \oplus X_2|Y_3)}, I(X_3; Y_1), I(X_3; Y_2) \right]$$

$$R_2 \leq \left[\frac{C_{\text{MAC}} H(W_2)}{C_{\text{MAC}} + 2H(X_1 \oplus X_2|Y_3)}, I(X_3; Y_1), I(X_3; Y_2) \right],$$

for some joint distributions of the form $p^*(y_3|x_1, x_2)p(x_1|w_1)p(x_2|w_2)p(w_1, w_2)$ and $p(x_3)p^*(y_1, y_2|x_3)$. Here, C_{MAC} is the maximum sum-rate of the multiple-access channel $p^*(y_3|x_1, x_2)$.

Remark 5: The above result is also valid even for TWRCs where \mathcal{X}_1 and \mathcal{X}_2 are not finite fields. See Remark 2.

The above rate region is obtained using the results in [6, Theorem 2] (by setting $V = X_1 \oplus X_2$) and [13, p. 567 (Theorem 15.6.3)]. The additional factors $H(W_1)$ and $H(W_2)$ in the above equations compared to [6, Eqn. (23)] convert computation rates to rates in bits/channel use considered in this paper. The proof is omitted because of space constraints.

V. COMPARISON OF CODING STRATEGIES

In this section, we show that the maximum sum rate obtained by FDF-L can be simultaneously higher than those achievable by FDF-S, and by two existing coding strategies: CDF and CF.

A. Existing Coding Strategies

1) *Complete-Decode-Forward (CDF):* Using CDF, the relay completely decodes the messages W_1 and W_2 sent by users 1 and 2 respectively. It then encodes and broadcasts a function of the messages to the users such that each user can recover the message sent by the other user. The overall achievable rate region is thus limited by two sets of constraints, i.e., the multiple-access constraints [14], [15] on the uplink and the broadcast constraints [16, Theorem 2.5] on the downlink, and is given in the following theorem.

Theorem 3: [see [4], [7], [8]] Consider a TWRC. The rate pair (R_1, R_2) is achievable using CDF if

$$R_1 \leq \min \left\{ I(X_1; Y_3|X_2), I(X_3; Y_2) \right\} \quad (16)$$

$$R_2 \leq \min \left\{ I(X_2; Y_3|X_1), I(X_3; Y_1) \right\} \quad (17)$$

$$R_1 + R_2 \leq I(X_1, X_2; Y_3), \quad (18)$$

for some joint distributions of the form $p(x_1)p(x_2)p^*(y_3|x_1, x_2)$ and $p(x_3)p^*(y_1, y_2|x_3)$.

2) *Compress-Forward (CF)*: Using this strategy the relay quantizes its received signal Y_3 to \hat{Y}_3 , encodes and broadcasts \hat{Y}_3 to the users. Assuming that both users can correctly decode \hat{Y}_3 , a virtual channel $X_1 \rightarrow \hat{Y}_3$ is created from user 1 to user 2 via the relay. Similarly, a virtual channel $X_2 \rightarrow \hat{Y}_3$ is created from user 2 to user 1 via the relay. The achievable rate region using CF on the TWRC is given in the following theorem.

Theorem 4: [see [9]] Consider a TWRC. The rate pair (R_1, R_2) is achievable using CF if $R_1 \leq I(X_1; \hat{Y}_3|X_2, T)$ and $R_2 \leq I(X_2; \hat{Y}_3|X_1, T)$, under the constraints $I(Y_3; \hat{Y}_3|X_1, T) < I(X_3; Y_1)$ and $I(Y_3; \hat{Y}_3|X_2, T) < I(X_3; Y_2)$, for some joint distributions of the form $p(t)p(x_1|t)p(x_2|t)p^*(y_3|x_1, x_2)p(\hat{y}_3|y_3)$ and $p(x_3)p^*(y_1, y_2|x_3)$, where $T \in \mathcal{T}$, $|\mathcal{T}| \leq 4$, $\hat{\mathcal{Y}}_3 \in \hat{\mathcal{Y}}_3$, and $|\hat{\mathcal{Y}}_3| \leq |\mathcal{Y}_3| + 3$.

B. Numerical Calculations

Now, we compare these four coding strategies on a TWRC.

1) *Channel*: We consider the following TWRC:

- $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{X}_3 = \mathcal{Y}_1 = \mathcal{Y}_2 = \{0, 1\}$, $\mathcal{Y}_3 = \{a, b, c, d\}$.
- $p^*(y_3|x_1, x_2)$ is given by the following transition matrix:

$p^*(y_3 x_1, x_2)$	$y_3 = a$	$y_3 = b$	$y_3 = c$	$y_3 = d$
$x_1 = 0, x_2 = 0$	0.1	0.2	0.3	0.4
$x_1 = 0, x_2 = 1$	0.3	0.4	0.1	0.2
$x_1 = 1, x_2 = 0$	0.4	0.3	0.2	0.1
$x_1 = 1, x_2 = 1$	0.2	0.1	0.4	0.3

Each entry in the lower right matrix denotes the conditional probability $p^*(y_3|x_1, x_2)$ that y_3 is received when (x_1, x_2) are sent. Note that $p^*(y_3|x_1, x_2)$ cannot be written as a noisy function of $x_1 \oplus x_2$.

- $p^*(y_1, y_2|x_3) = \text{BSC}_{Y_1|X_3}(0.3)\text{BSC}_{Y_2|X_3}(0.3)$, where

$$\text{BSC}_{Y|X}(\rho) \triangleq p(y|x) = \begin{cases} 1 - \rho, & \text{if } y = x \\ \rho, & \text{otherwise} \end{cases} \quad (19)$$

The downlink from the relay to each user is a binary-symmetric channel with cross-over probability $\rho = 0.3$.

2) *Achievable Sum Rates*: The maximum sum rates (i.e., $R_1 + R_2$) achievable by the different coding strategies are

- FDF-L: $R_1 + R_2 = 0.2374$.
- FDF-S: $R_1 + R_2 = 0.1602$.
- CDF: $R_1 + R_2 = 0.1536$.
- CF (an upper bound on the maximum sum rate): $R_1 + R_2 \leq 0.2374 - \zeta$, for some $\zeta > 0$.

Clearly, FDF-L outperforms the other coding strategies on this TWRC.

VI. CONCLUSION

We have proposed a functional-decode-forward coding strategy with linear codes (FDF-L) for the general discrete memoryless two-way relay channel (TWRC) and obtained a new achievable rate region. We showed that using random linear codes for the users, the relay can reliably decode a function of the users' codewords even when the channel does not perform the desired function. The function, when broadcast back to the users, allows each user to decode the other user's message.

Noting that functional decoding on the uplink of the discrete memoryless TWRC is also possible using systematic

computation codes, we obtained another achievable region for the TWRC using functional-decode-forward with systematic computation codes (FDF-S).

With an example, we numerically showed that FDF-L is capable of achieving strictly higher sum rates compared to FDF-S and two existing coding strategies, namely, complete-decode-forward and compress-forward.

However, using FDF-L or FDF-S, if the cardinalities of the user's input alphabets $|\mathcal{X}_i|$ are both not equal to that of any finite field, only subsets of \mathcal{X}_i are utilized for transmission. Furthermore, since linear codes are used for FDF-L, the distributions of the users' transmitted signals $p(x_i)$ are constrained to be uniform, which is not always optimal for the channel.

This paper nonetheless provides coding schemes for the relay to decode a function of the users' messages without having to decode the messages individually on the general discrete memoryless TWRC (which may not be additive). This strategy can be useful in multiterminal networks where different destination nodes have knowledge of some source messages and want to decode the messages of other sources.

REFERENCES

- [1] 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.
- [2] —, "The capacity of a class of multi-way relay channels," in *Proc. IEEE Int. Conf. on Commun. Syst. (ICCS)*, Singapore, Nov. 17–20 2010.
- [3] K. Narayanan, M. P. Wilson, and A. Sprintson, "Joint physical layer coding and network coding for bi-directional relaying," in *Proc. 45th Allerton Conf. on Commun., Control, and Comput.*, Monticello, USA, Sep. 26–28 2007, pp. 254–259.
- [4] 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.
- [5] B. Nazer and M. Gastpar, "The case for structured random codes: Beyond linear models," in *Proc. 46th Allerton Conf. on Commun., Control, and Comput.*, Monticello, USA, Sep. 23–26 2008, pp. 1422–1425.
- [6] —, "Computation over multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3498–3516, Oct. 2007.
- [7] 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.
- [8] S. J. Kim, P. Mitran, and V. Tarokh, "Performance bounds for bidirectional coded cooperation protocols," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 5235–5241, Nov. 2008.
- [9] C. Schnurr, T. J. Oechtering, and S. Stanczak, "Achievable rates for the restricted half-duplex two-way relay channel," in *Proc. 41st Asilomar Conf. on Signals, Syst. and Comput.*, Pacific Grove, USA, Nov. 4–7 2007, pp. 1468–1472.
- [10] R. G. Gallager, *Information Theory and Reliable Communication*. Wiley, 1968.
- [11] R. W. Yeung, *Information Theory and Network Coding*. Springer, 2008.
- [12] T. Berger, "Multiterminal source coding," in *Lecture notes presented at the 1977 CISM Summer School*, Udine, Italy, Jul. 18–20 1977, pp. 171–231.
- [13] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Wiley-Interscience, 2006.
- [14] 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.
- [15] H. Liao, "A coding theorem for multiple access communication," in *Proc. IEEE Int. Symposium on Inf. Theory (ISIT)*, Asilomar, USA, Jan. 1972.
- [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.