Decode-and-Forward for the Gaussian Relay Channel via Standard AWGN Coding and Decoding

Anatoly Khina, Or Ordentlich, Uri Erez
Dept. EE-Systems, TAU
Tel Aviv, Israel
Email: {anatolyk,ordent,uri}@eng.tau.ac.il

Yuval Kochman
School of CSE, HUJI
Jerusalem, Israel
Email: yuvalko@cs.huji.ac.il

Gregory W. Wornell
Dept. EECS, MIT
Cambridge, MA 02139, USA
Email: gww@mit.edu

***Abstract*—This work considers practical implementation of the decode-and-forward relaying protocol for the full-duplex Gaussian relay channel. Unlike previous works which developed coding techniques tailored to this protocol, it is shown that standard fixed-rate codes which are good for the Gaussian scalar channel suffice to approach the theoretical performance promised by this protocol. The proposed technique employs only linear operations and successive interference cancelation in conjunction with a fixed-rate base code, and the achievable rate is solely dictated by the performance of the base code. The same approach and results carry over to the multiple-antenna case as well.**

I. Introduction

Relaying techniques are a key element in enhancing the capacity of wireless networks. Accordingly, a great amount of research effort has been devoted in recent years to studying both the information-theoretic limits of networks incorporating relays, and to developing suitable coding techniques.

While the capacity of even the most basic relay channel, namely the scalar Gaussian single-relay channel [1] remains an open problem, achievable rates regions as well as outer bounds have been established for basic models, following the pioneering work of Cover and El Gamal [2]. The achievable rates are largely based on a few key relaying approaches, and can roughly be categorized around the Amplify-and-Forward (AF), Compress-and-Forward (CF), and Decode-and-Forward (DF) protocols; see e.g. [2], [3].

In this work we consider the Gaussian full-duplex[1] relay channel in which the relay may receive and transmit simultaneously. The channel model is depicted in Figure 1.

The source transmits a signal that is received at both the destination and relay; the relay transmits an additional signal to the destination. Thus, the channel input–output relationship is

$$\begin{aligned} y^{\text{rel}} &= h_{\text{src,rel}} x^{\text{src}} + z^{\text{rel}} , \\ y^{\text{dst}} &= h_{\text{src,dst}} x^{\text{src}} + h_{\text{rel,dst}} x^{\text{rel}} + z^{\text{dst}} , \end{aligned} \tag{1}$$

where x, y and z denote the channel input, output and noises, respectively, using superscripts and subscripts 'dst', 'src' and 'rel' to indicate 'destination', 'source' and 'relay', respectively. Without loss of generality, all channel inputs are subject to the same average power constraint P and all gains are real and non-negative.[2] We assume full knowledge of all channel gains at the source, destination and relay nodes ("closed loop").

We further focus on the DF protocol which we review in Section II-A. In the DF protocol, a sequence of messages is transmitted. The key feature of the DF protocol is encoding each message twice using two independent codebooks and transmitting each message over two consecutive time frames. In the first time frame the message is encoded at the source using the first codebook. The relay, which is assumed to have a better channel than the destination, decodes the message. The destination is unable to decode at this stage. Rather, it forms a list of possible candidates for the transmitted message. In the second time frame, the relay and source both encode the message using the second codebook and transmit it coherently to the destination, where the source transmits also via superposition a new message to be decoded by the relay. Now the destination can create another list of candidates and intersect it with the list from the previous time frame. If the transmission rate is chosen correctly, only the transmitted message would fall in the intersection of both lists.

The question of how to implement the list decoders above is non-trivial and has motivated numerous works that proposed coding techniques tailored to DF relaying; see [5]–[11] and references therein. In contrast, the purpose of the present work is to provide a framework for constructing coding schemes, allowing to approach the theoretical performance of the DF protocol using only "off-the-shelf" scalar codes (and decoders) designed for the additive white Gaussian noise (AWGN) channel.

Our approach is based upon the observation that in the Gaussian case, the DF scenario can be mathematically formulated as an equivalent multiple-input multiple-output (MIMO) multicast (or common message broadcast) scenario as developed in Section III. This equivalent scenario, consists of a transmit node equipped with two "antennas" communicating with two receive nodes (users). The first receive node is the relay which has one receive antenna, corresponding to the single time frame it is allowed for decoding the message. The second receive node is the destination which is equipped with two receive antennas corresponding to the two time frames it utilizes to decode the message. The fact that the relay coherently transmits with the source during the second time frame is also taken into account in the derived equivalent MIMO multicast model.

[1] For a practical scheme for the half-duplex setting see [4].
[2] This follows since any phase can be absorbed in the transmit signals.

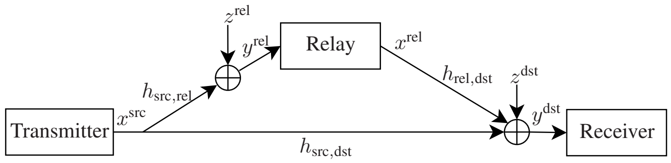

Fig. 1. Gaussian MIMO relay channel.

In [12] a signal processing architecture was proposed for the Gaussian multiple-input multiple-output (MIMO) multicast scenario, that is based on a novel joint unitary decomposition. It was demonstrated that one can approach the multicast capacity using scalar point-to-point encoders and decoders with successive interference cancelation (SIC). Applying this architecture to the equivalent multicast model we derive in this work, yields a low complexity optimal scheme for the DF protocol. In our scheme, the coding task is reduced to that of coding over an AWGN channel, whereas the relay network topology is accounted for by the linear processing.

The approach we develop readily carries over to MIMO DF relaying and can be generalized to more relays, as described in Section IV.

II. Background

We now recall the decode-and-forward protocol for the Gaussian SISO relay channel, as well as the practical MIMO multicast scheme of [12]. In the sequel, we show how the latter may be combined with the DF protocol, to arrive at a practical DF scheme.

A. Full-Duplex Decode-and-Forward Protocol

The DF scheme has a sequential nature: the data is partitioned into a sequence of messages $\{w_i\}$. At each time instance, the source and relay transmit functions of a "sliding window" of messages, and the relay and destination decode messages sequentially. Thus, both the relay and the destination employ successive interference cancellation (SIC). Specifically, the DF scheme (for the Gaussian case) consists of the following.

Codebook construction: Generate two different (*independent*) good AWGN codebooks of the same length n and power P and rates to be determined in the sequel. We denote the codebooks by $\mathcal{C}^{\mathrm{C}}$ and $\mathcal{C}^{\mathrm{D}}$, where the superscript C stands for 'coherent', and D — for 'direct', the operational meanings of which will become apparent in the sequel.

We now describe transmission of block/message i. Let $\boldsymbol{x}_i^{\mathrm{C}}$ and $\boldsymbol{x}_i^{\mathrm{D}}$ be the codewords corresponding to a message w_i in $\mathcal{C}^{\mathrm{C}}$ and $\mathcal{C}^{\mathrm{D}}$, respectively.[3]

Source: The i-th block signal of length n, $\boldsymbol{x}_i^{\mathrm{src}}$, sent by the source, is equal to the sum of $\boldsymbol{x}_i^{\mathrm{C}}$ and $\boldsymbol{x}_{i+1}^{\mathrm{D}}$, where a portion ρ^2 ($0 \le \rho^2 \le 1$) of the available power P is allocated to $\boldsymbol{x}_i^{\mathrm{C}}$ and the rest — to $\boldsymbol{x}_{i+1}^{\mathrm{D}}$. Thus, the signal in block (or "time frame") i sent by the source, $\boldsymbol{x}_i^{\mathrm{src}}$, is equal to

$$\boldsymbol{x}_i^{\mathrm{src}} = \rho \boldsymbol{x}_i^{\mathrm{C}} + \sqrt{1-\rho^2}\boldsymbol{x}_{i+1}^{\mathrm{D}} . \tag{2}$$

[3] In this subsection, all vectors are taken to be row vectors.

Relay: At each time frame $i-1$, the relay recovers w_i from a single output block $\boldsymbol{y}_{i-1}^{\mathrm{rel}}$ by decoding $\boldsymbol{x}_i^{\mathrm{D}}$, where w_{i-1} is assumed to be known (assuming correct decoding of previous codewords, and that w_1 is predetermined). Since w_{i-1}, and hence also $\boldsymbol{x}_{i-1}^{\mathrm{C}}$ are assumed to be known, the contribution of the latter can be subtracted from $\boldsymbol{y}_{i-1}^{\mathrm{rel}}$, resulting (assuming correct decoding of previous codewords):

$$\begin{aligned} \tilde{\boldsymbol{y}}_{i-1}^{\mathrm{rel}} &= \boldsymbol{y}_{i-1}^{\mathrm{rel}} - h_{\mathrm{src,rel}}\rho \boldsymbol{x}_{i-1}^{\mathrm{C}} \\ &= \sqrt{1-\rho^2} h_{\mathrm{src,rel}} \boldsymbol{x}_i^{\mathrm{D}} + \boldsymbol{z}_{i-1}^{\mathrm{rel}} . \end{aligned} \tag{3}$$

At time frame i (and again assuming correct decoding), the relay knows w_i (and hence also $\boldsymbol{x}_i^{\mathrm{C}}$) and sends

$$\boldsymbol{x}_i^{\mathrm{rel}} = \boldsymbol{x}_i^{\mathrm{C}} . \tag{4}$$

Destination: At each time frame i, the destination recovers w_i based on two consecutive output blocks $\boldsymbol{y}_{i-1}^{\mathrm{dst}}$ and $\boldsymbol{y}_i^{\mathrm{dst}}$, assuming w_{i-1} is known (was decoded correctly from previous outputs, except for w_1 which is predetermined). In essence, the destination recovers w_i from two observations of the encoded message: the first being a noisy version of $\boldsymbol{x}_i^{\mathrm{D}}$, the other being a noisy version of $\boldsymbol{x}_i^{\mathrm{C}}$.

Specifically, we subtract $h_{\mathrm{src,dst}}\rho \boldsymbol{x}_{i-1}^{\mathrm{C}}$ from $\boldsymbol{y}_{i-1}^{\mathrm{dst}}$ to arrive at

$$\begin{aligned} \tilde{\boldsymbol{y}}_{i-1}^{\mathrm{dst}} &= \boldsymbol{y}_{i-1}^{\mathrm{dst}} - h_{\mathrm{src,dst}}\rho \boldsymbol{x}_{i-1}^{\mathrm{C}} \\ &= \sqrt{1-\rho^2} h_{\mathrm{src,dst}} \boldsymbol{x}_i^{\mathrm{D}} + \boldsymbol{z}_{i-1}^{\mathrm{dst}} . \end{aligned} \tag{5}$$

This serves as the first noisy observation.

The second noisy observation is that of $\boldsymbol{x}_i^{\mathrm{C}}$, which is obtained from $\boldsymbol{y}_i^{\mathrm{dst}}$ as follows. the component of $\boldsymbol{x}_i^{\mathrm{C}}$ in the source signal $\boldsymbol{x}_i^{\mathrm{src}}$ and the signal transmitted by the relay $\boldsymbol{x}_i^{\mathrm{rel}}$ sum coherently, whereas $\boldsymbol{x}_{i+1}^{\mathrm{D}}$ plays the role of an AWGN, namely,

$$\boldsymbol{y}_i^{\mathrm{dst}} = \left(\rho h_{\mathrm{src,dst}} + h_{\mathrm{rel,dst}}\right) \boldsymbol{x}_i^{\mathrm{C}} + \boldsymbol{z}_i^{\mathrm{equiv}} , \tag{6}$$

where

$$\boldsymbol{z}_i^{\mathrm{equiv}} = \sqrt{1-\rho^2} h_{\mathrm{src,dst}} \boldsymbol{x}_i^{\mathrm{D}} + \boldsymbol{z}_i^{\mathrm{dst}} \tag{7}$$

is of power $P_{\mathrm{equiv}} = \left(1-\rho^2\right) h_{\mathrm{src,dst}}^2 P + 1$. Normalizing the power of the noise $\boldsymbol{z}_i^{\mathrm{equiv}}$, i.e., dividing $\boldsymbol{y}_i^{\mathrm{dst}}$ by $\sqrt{P_{\mathrm{equiv}}}$, we arrive at

$$\begin{aligned} \overline{\boldsymbol{y}}_i^{\mathrm{dst}} &= \frac{1}{\sqrt{P_{\mathrm{equiv}}}} \boldsymbol{y}_i^{\mathrm{dst}} \\ &= \frac{\rho h_{\mathrm{src,dst}} + h_{\mathrm{rel,dst}}}{\sqrt{\left(1-\rho^2\right) h_{\mathrm{src,dst}}^2 P + 1}} \boldsymbol{x}_i^{\mathrm{C}} + \overline{\boldsymbol{z}}_i^{\mathrm{dst}} , \end{aligned} \tag{8}$$

where $\overline{\boldsymbol{z}}_i^{\mathrm{dst}}$ is of unit power.

Note that the relay recovers w_i based only on a single observation (3), whereas the destination uses two observations (5) and (8) to recover the same information. Thus, the achievable rate of the DF protocol is limited by the minimum of the mutual informations of the two [2]:[4]

[4] Assuming many blocks, the loss due to predetermining w_1 is negligible.

$$
\begin{aligned}
R_{\text{DF}} = \min \Bigg\{ & \log\left(1 + \left(1-\rho^2\right) h^2_{\text{src,rel}} P\right), \\
& \log\left(1 + \left(1-\rho^2\right) h^2_{\text{src,dst}} P\right) \\
& \quad + \log\left(1 + \frac{\left(\rho h_{\text{src,dst}} + h_{\text{rel,dst}}\right)^2 P}{\left(1-\rho^2\right) h^2_{\text{src,dst}} P + 1}\right) \Bigg\} \qquad (9) \\
= \min \Bigg\{ & \log\left(1 + \left(1-\rho^2\right) S_{\text{src,rel}}\right), \\
& \log\left(1 + S_{\text{src,dst}} + S_{\text{rel,dst}} + 2\rho\sqrt{S_{\text{src,dst}} S_{\text{rel,dst}}}\right) \Bigg\},
\end{aligned}
$$

where we define the signal-to-noise ratios (SNRs) as $S \triangleq h^2 P$ where h may correspond to $h_{\text{rel,dst}}$, $h_{\text{src,dst}}$ or $h_{\text{src,rel}}$.

Remark 1: The technique above is applicable for any value of ρ between 0 and 1. An optimization of this parameter and its optimal explicit value along with the corresponding optimal rate in (9), can be evaluated; see, e.g., [13].

Nevertheless, it is not clear how to combine the information carried by the two codewords $\boldsymbol{x}^{\text{C}}$ and $\boldsymbol{x}^{\text{D}}$ and at the same time recover it at the relay, using a *practical* scheme. Different approaches, e.g., list decoding, were proposed, but these are still hard to implement in practice; see [3] for a detailed survey of these schemes. In the sequel, we show how to overcome this hurdle, i.e., design a practical scheme that approaches (9). This is done by showing that this problem may be recast as an effective MIMO multicast one, and then using the multicast scheme presented next in Section II-B.

B. Codes for Common-Message MIMO Multicast

We now describe the main tool which is used in this work. We follow [12] where a practical coding scheme is introduced for a related problem, where data needs to be multicasted to two users over a MIMO broadcast channel. The channel model is described by

$$\boldsymbol{y}_k = H_k \boldsymbol{x} + \boldsymbol{z}_k \,, \qquad (10)$$

where $\boldsymbol{y}_k$ is the received $M_{r_k} \times 1$ vector of user k ($k = 1, 2$), $\boldsymbol{x}$ denotes the $N_t \times 1$ complex-valued input vector which is limited to an average power P per symbol,[5] H_k is the $M_{r_k} \times N_t$ complex channel matrix to user k, and $\boldsymbol{z}_k$ is assumed to be a circularly-symmetric Gaussian vector with identity covariance matrix.

In this subsection we suppress time indices for simplicity of exposition.

For a single user, the rate achievable for an $M_r \times N_t$ channel matrix H and an input covariance matrix $K \triangleq E[\boldsymbol{x}\boldsymbol{x}^\dagger]$ is equal to the Gaussian mutual information (MI) between the input and output vectors

$$R(H, K) \triangleq \log\left|I_{N_r} + HKH^\dagger\right| \,, \qquad (11)$$

where $|\cdot|$ denotes the determinant and I_{M_r} is an identity matrix of dimension M_r. The common-message ("multicast") capacity of this channel is given by the compound-channel capacity [14]

$$C_{\text{common}} = \max_K \min_{i=1,2} R(H_k, K) \,.$$

We start with N_t codebooks, each one of them good for a SISO Gaussian channel of rate R_j to be determined and unit power. At each time instance we form a vector $\tilde{\boldsymbol{x}}$ using one sample from each codebook. The transmitted vector is thus

$$\boldsymbol{x} = BV\tilde{\boldsymbol{x}} \,, \qquad (12)$$

where $E\left[\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}^\dagger\right] = I_{N_t}$, V is unitary and B is a precoding matrix satisfying

$$BB^\dagger = K \,, \qquad (13)$$

meaning that the input covariance matrix is indeed equal to K. Receiver k ($k = 1, 2$) computes[6]

$$\tilde{\boldsymbol{y}}_k = U_k^\dagger \boldsymbol{y}_k \,, \qquad (14)$$

and then decodes the N_t codes using SIC, starting from $\tilde{x}_{N_t}$, where $\tilde{x}_k$ denotes the k-th entry of $\tilde{\boldsymbol{x}}$. The following theorem, due to [12], shows that this strategy is optimal.

Theorem 1: For any two channel matrices H_1 and H_2, and input covariance matrix K such that $R(H_1, K) = R(H_2, K) = R$, there exist U_1, U_2 and a unitary V such that

$$
\begin{aligned}
S_{1,j} &= S_{2,j} \\
R_j &= \log\left(1 + S_{1,j}\right) \,, \qquad \forall j = 1, \ldots, N_t \\
R &= \sum_{j=1}^{N_t} R_j \,,
\end{aligned}
$$

where $S_{k,j}$ is the signal-to-interference-and-noise ratio of the j-th symbol at receiver k, given symbols $\tilde{x}_{j+1}^{N_t}$.

By the first two equations, codebooks of rate

$$R_j = \log(1 + S_{1,j})$$

can be decoded by both receivers (using standard AWGN decoding); by the third equation, the sum of these rates equals the optimum.

Remark 2: Using an input covariance matrix K over the channels described by the channel matrices H_1 and H_2, is mathematically equivalent to working with a unit covariance matrix I_{N_t} over equivalent channel matrices $F_1 \triangleq H_1 B$ and $F_2 \triangleq H_2 B$, respectively.

III. Main Result: Full-Duplex Decode-and-Forward Scheme

We now show how the decode-and-forward protocol for the SISO Gaussian relay channel may be formulated as an equivalent MIMO multicast scenario. Applying the scheme of Section II-B to this scenario reduces the coding task to that of Gaussian SISO (fixed-rate) coding, for which any standard coding technique can be used.

[5] Alternatively, one can consider an input covariance constraint.

[6] The receiver matrices U_k are not necessarily unitary.

Recall that the relay uses only a *single* observation as reflected in (3), to recover w_i. The destination, on the other hand, makes use of *two* consecutive observation blocks, as given in (5) and (8), to recover the same message w_i. This can be reformulated in an equivalent matrix notation as

$$\begin{aligned} \mathcal{Y}_i^{\text{rel}} &= \mathcal{H}_{\text{rel}} \mathcal{X}_i + \mathcal{Z}_i^{\text{rel}} , \\ \mathcal{Y}_i^{\text{dst}} &= \mathcal{H}_{\text{dst}} \mathcal{X}_i + \mathcal{Z}_i^{\text{dst}} , \end{aligned} \tag{15}$$

where

$$\begin{aligned} \mathcal{H}_{\text{rel}} &= \sqrt{2} \begin{pmatrix} \sqrt{1-\rho^2} h_{\text{src,rel}} & 0 \end{pmatrix} \\ \mathcal{H}_{\text{dst}} &= \sqrt{2} \begin{pmatrix} \sqrt{1-\rho^2} h_{\text{src,dst}} & 0 \\ 0 & \frac{\rho h_{\text{src,dst}} + h_{\text{rel,dst}}}{\sqrt{((1-\rho^2) h_{\text{src,dst}}^2 P + 1}} \end{pmatrix} \end{aligned} \tag{16}$$

are the effective channel matrices,

$$\begin{aligned} \mathcal{Z}_i^{\text{rel}} &\triangleq \boldsymbol{z}_{i-1}^{\text{rel}} \\ \mathcal{Z}_i^{\text{dst}} &\triangleq \begin{pmatrix} \boldsymbol{z}_{i-1}^{\text{dst}} \\ \overline{\boldsymbol{z}}_i^{\text{dst}} \end{pmatrix} \end{aligned}$$

are additive white noises with identity covariance matrices,

$$\mathcal{X}_i \triangleq \frac{1}{\sqrt{2}} \begin{pmatrix} \boldsymbol{x}_i^{\text{D}} \\ \boldsymbol{x}_i^{\text{C}} \end{pmatrix} \tag{17}$$

is the channel input vector subject to a power constraint P, and

$$\begin{aligned} \mathcal{Y}_i^{\text{rel}} &\triangleq \tilde{\boldsymbol{y}}_{i-1}^{\text{rel}} \\ \mathcal{Y}_i^{\text{dst}} &\triangleq \begin{pmatrix} \tilde{\boldsymbol{y}}_{i-1}^{\text{dst}} \\ \overline{\boldsymbol{y}}_i^{\text{dst}} \end{pmatrix} \end{aligned}$$

are the effective output vector at the relay and the destination, respectively.

The input dimension of the matrices $\mathcal{H}_{\text{rel}}$ and $\mathcal{H}_{\text{dst}}$ is equal to two, since the input signal $\mathcal{X}_i$ consists of two independent codewords. The output of these matrices, reflect the number of observation blocks utilized, by each of the receive nodes, to recover these codewords.

Thus, the problem of implementing the decode-and-forward protocol over the Gaussian relay channel is equivalent to a MIMO multicast problem, for which the scheme of Section II-B can be applied, resulting in a practical scheme that achieves the decode-and-forward rate (9). Note that in this case B of (12) is equal the identity matrix I_2.[7]

Remark 3: In order to approach the rate given in (9), $\boldsymbol{x}_i^{\text{C}}$ and $\boldsymbol{x}_i^{\text{D}}$ need to be independent. In the proposed scheme this is indeed the case, since we start with independent codebooks of the same power (the entries of $\tilde{\boldsymbol{x}}$ in (12) which in this case are the entries of $\mathcal{X}$ in (17)) and send two orthonormal combinations of the two (materialized by the multiplication by a unitary matrix V in (12)). If the codebooks (entries of $\tilde{\boldsymbol{x}}$ in (12) or $\mathcal{X}$ in (17)) are good (SISO) AWGN codes, this implies that they have a Gaussian distribution, which, together with the orthogonality property, implies in turn the independence of $\boldsymbol{x}_i^{\text{C}}$ and $\boldsymbol{x}_i^{\text{D}}$.

[7]Note that the power allocation of $\boldsymbol{x}^{\text{C}}$ and $\boldsymbol{x}^{\text{D}}$ is absorbed in the effective channel matrices $\mathcal{H}^{\text{rel}}$ and $\mathcal{H}^{\text{dst}}$.

IV. Extensions

A. MIMO Relay Channel

The DF technique and scheme used for the SISO case, can be readily extended to the MIMO case, where each of the nodes (source, relay and destination) is equipped with multiple antennas. In this case the scalar channel coefficients h are replaced by channel matrices H whose dimensions are determined by the number of antennas at the corresponding transmit and receive ends.

Direct extension of the technique of Section II-A for the MIMO case calls for replacing the scalar codebooks $\boldsymbol{x}^{\text{D}}$ and $\boldsymbol{x}^{\text{C}}$ with vector codebooks whose entries are vectors of the same dimension as the transmitted signal $\boldsymbol{x}^{\text{src}}$.

In this section, with some abuse of notation, boldface is used to denote spatial vectors rather than time vectors, as was the case in Sections II-A and III. Furthermore, vectors are taken to be column vectors where the time index is suppressed, and the codeword is represented by a single symbol.

Without loss of generality, we assume that the number of the transmit antennas at the source and the relay is equal, i.e., the column dimensions of $H_{\text{src,dst}}$ and $H_{\text{rel,dst}}$ are equal, since otherwise, we may pad the matrix with the lower column dimension with all-zero columns. Thus, $\boldsymbol{x}^{\text{src}}$ and $\boldsymbol{x}^{\text{rel}}$, and hence also $\boldsymbol{x}^{\text{D}}$ and $\boldsymbol{x}^{\text{C}}$, are all of the same dimension.

We will see shortly that standard scalar codes suffice for this case as well.

Note also that in this case, the covariance matrices corresponding to $\boldsymbol{x}^{\text{D}}$ and $\boldsymbol{x}^{\text{C}}$ at the source and to $\boldsymbol{x}^{\text{C}}$ — at the relay, can be shaped to improve the achievable rate, such that each of these covariance matrices satisfy the power constraints [15]. More generally, we take $\boldsymbol{x}^{\text{D}}$ and $\boldsymbol{x}^{\text{C}}$ to be independent and white of total average unit power and multiply them by suitable precoding matrices, as in (12)–(13), subject to an average power constraint P. Hence the signals sent by the source (2) and the relay (4) should be replaced, in the MIMO case, by

$$\begin{aligned} \boldsymbol{x}_i^{\text{src}} &= \rho B_{\text{src}}^{\text{C}} \boldsymbol{x}_i^{\text{C}} + \sqrt{1-\rho^2} B_{\text{src}}^{\text{D}} \boldsymbol{x}_{i+1}^{\text{D}} , \\ \boldsymbol{x}_i^{\text{rel}} &= B_{\text{rel}}^{\text{C}} \boldsymbol{x}_i^{\text{C}} , \end{aligned}$$

respectively, where $B_{\text{src}}^{\text{C}}$, $B_{\text{src}}^{\text{D}}$ and $B_{\text{rel}}^{\text{C}}$ are the precoding matrices satisfying the power constraints:

$$\text{trace}\left\{ B_{\text{src}}^{\text{C}} {B_{\text{src}}^{\text{C}}}^{\dagger} \right\}, \text{trace}\left\{ B_{\text{src}}^{\text{D}} {B_{\text{src}}^{\text{D}}}^{\dagger} \right\}, \text{trace}\left\{ B_{\text{rel}}^{\text{C}} {B_{\text{rel}}^{\text{C}}}^{\dagger} \right\} \leq P .$$

Remark 4: In the SISO case, the signal sent by the relay (4) and the corresponding component $\boldsymbol{x}^{\text{C}}$ in the signal of (2) need to be multiplied by an appropriate phase, such that they sum coherently (this was absorbed in the channel coefficients h in the exposition of the SISO channel in Section II-A, which were assumed to be real and non-negative); in the MIMO case, this generalizes to multiplying by appropriate unitary matrices prior to the covariance shaping, which together constitute the precoding matrices.

The channel output at the relay, after subtracting the component corresponding to $\boldsymbol{x}^{\text{C}}$ (MIMO equivalent of (3), assuming correct decoding), is equal to

$$R_{\mathrm{DF}} = \min\Bigg\{ \log\left|I + \left(1-\rho^2\right) H_{\mathrm{src,rel}} B_{\mathrm{src}}^{\mathrm{D}} \left(H_{\mathrm{src,rel}} B_{\mathrm{src}}^{\mathrm{D}}\right)^\dagger\right|, \tag{18}$$

$$\log\left|I + \left(1-\rho^2\right) H_{\mathrm{src,dst}} B_{\mathrm{src}}^{\mathrm{D}} \left(H_{\mathrm{src,dst}} B_{\mathrm{src}}^{\mathrm{D}}\right)^\dagger\right| + \log\left|I + K_{\mathrm{equiv}}^{-1}\left(\rho H_{\mathrm{src,dst}} B_{\mathrm{src}}^{\mathrm{C}} + H_{\mathrm{rel,dst}} B_{\mathrm{rel}}^{\mathrm{C}}\right)\left(\rho H_{\mathrm{src,dst}} B_{\mathrm{src}}^{\mathrm{C}} + H_{\mathrm{rel,dst}} B_{\mathrm{rel}}^{\mathrm{C}}\right)^\dagger\right| \Bigg\}$$

$$\begin{aligned}\tilde{\boldsymbol{y}}_{i-1}^{\mathrm{rel}} &= \boldsymbol{y}_{i-1}^{\mathrm{rel}} - \rho H_{\mathrm{src,rel}} B_{\mathrm{src}}^{\mathrm{C}} \boldsymbol{x}_{i-1}^{\mathrm{C}} \\ &= \sqrt{1-\rho^2} H_{\mathrm{src,rel}} B_{\mathrm{src}}^{\mathrm{D}} \boldsymbol{x}_i^{\mathrm{D}} + \boldsymbol{z}_{i-1}^{\mathrm{rel}} .\end{aligned}$$

At the destination, (5) is replaced by

$$\boldsymbol{y}_{i-1}^{\mathrm{dst}} = \sqrt{1-\rho^2} H_{\mathrm{src,dst}} B_{\mathrm{src}}^{\mathrm{D}} \boldsymbol{x}_i^{\mathrm{D}} + \boldsymbol{z}_{i-1}^{\mathrm{dst}} ,$$

whereas (6) and (7) are replaced by

$$\begin{aligned}\boldsymbol{y}_i^{\mathrm{dst}} &= \left(\rho H_{\mathrm{src,dst}} B_{\mathrm{src}}^{\mathrm{C}} + H_{\mathrm{rel,dst}} B_{\mathrm{rel}}^{\mathrm{C}}\right) \boldsymbol{x}_i^{\mathrm{C}} + \boldsymbol{z}_i^{\mathrm{equiv}} , \\ \boldsymbol{z}_i^{\mathrm{equiv}} &= \sqrt{1-\rho^2} H_{\mathrm{src,dst}} B_{\mathrm{src}}^{\mathrm{D}} \boldsymbol{x}_i^{\mathrm{D}} + \boldsymbol{z}_i^{\mathrm{dst}} ,\end{aligned}$$

respectively. The covariance matrix of the Gaussian noise is

$$K_{\mathrm{equiv}} \triangleq (1-\rho^2) H_{\mathrm{src,dst}} B_{\mathrm{src}}^{\mathrm{D}} \left(H_{\mathrm{src,dst}} B_{\mathrm{src}}^{\mathrm{D}}\right)^\dagger + I .$$

This matrix is positive-definite and therefore can be decomposed according to the Cholesky Decomposition as

$$K_{\mathrm{equiv}} = L_{\mathrm{equiv}} L_{\mathrm{equiv}}^\dagger , \tag{19}$$

where L is invertible. Hence, by applying L_{equiv}^{-1} to $\boldsymbol{y}_i^{\mathrm{dst}}$ on the right, we arrive at

$$\overline{\boldsymbol{y}}_i^{\mathrm{dst}} = L_{\mathrm{equiv}}^{-1}\left(\rho H_{\mathrm{src,dst}} B_{\mathrm{src}}^{\mathrm{C}} + H_{\mathrm{rel,dst}} B_{\mathrm{rel}}^{\mathrm{C}}\right) \boldsymbol{x}_i^{\mathrm{C}} + \overline{\boldsymbol{z}}_i^{\mathrm{dst}} ,$$

where $\overline{\boldsymbol{z}}_i^{\mathrm{dst}} = L_{\mathrm{equiv}}^{-1} \boldsymbol{z}_i^{\mathrm{equiv}}$ has an identity covariance matrix.

Thus, the achievable rate of this scheme in the MIMO case, is equal to (18).

The effective matrices (16) in the scheme of Section III, need to be replaced, in the MIMO case, by the following channel matrices:

$$\mathcal{H}_{\mathrm{rel}} = \sqrt{2}\begin{pmatrix}\sqrt{1-\rho^2} H_{\mathrm{src,rel}} B_{\mathrm{src}}^{\mathrm{D}} & 0\end{pmatrix}$$

$$\mathcal{H}_{\mathrm{dst}} = \sqrt{2}\begin{pmatrix}\sqrt{1-\rho^2} H_{\mathrm{src,dst}} B_{\mathrm{src}}^{\mathrm{D}} & 0 \\ 0 & L_{\mathrm{equiv}}^{-1}\Big(\rho H_{\mathrm{src,dst}} B_{\mathrm{src}}^{\mathrm{C}} \\ & \quad + H_{\mathrm{rel,dst}} B_{\mathrm{rel}}^{\mathrm{C}}\Big)\end{pmatrix}$$

over which any good "off-the-shelf" codebooks for the SISO AWGN channel can be used.

Remark 5: The above scheme works for any admissible choices of the precoding matrices (viz. matrices satisfying the power constraints) and $0 \le \rho \le 1$. Thus, optimization over these parameters can be performed to maximize the achievable rate (18).

B. More Relays and Equal-Rate Codebooks

In the scheme of Section III (and Section IV-A), the SISO codes used (constituting $\mathcal{X}$) are of different rates. Constructing a scheme in which all the SISO codes are of equal rates is more appealing, since this allows using the same codebook over all the SISO sub-channels as well as avoiding the need of a bid-loading mechanism; see [16], for a detailed explanation.

A Seemingly different problem is extending the scheme of Section III to the case of more relays (possibly without a direct link). For this, the multicast scheme of Section II-B needs to be extended to the case of more than two users.

Both of these problems can be resolved simultaneously by incorporating a space–time coding structure, but come at the expense of a greater latency at the output, as explained in [17].

References

[1] E. C. Van der Meulen. Three-terminal communication channels. *Adv. Appl. Prob.*, 3:120–54, 1971.

[2] T. M. Cover and A. A. El Gamal. Capacity theorems for the relay channel. *IEEE Trans. Info. Theory*, IT-25:572–584, Sept. 1979.

[3] G. Kramer, M. Gastpar, and P. Gupta. Cooperative strategies and capacity theorems for relay networks. *IEEE Trans. Info. Theory*, IT-51:3037–3063, Sept. 2005.

[4] A. Khina, Y. Kochman, U. Erez, and G. W. Wornell. Incremental coding over MIMO channels. In *Proceedings of IEEE Info. Theory Workshop (ITW), Paraty, Brazil*, pages 400–404, Oct. 2011.

[5] P. Razaghi and W. Yu. Bilayer low-density parity-check codes for decode-and-forward in relay channels. *IEEE Trans. Info. Theory*, 53:3723–3739, Oct. 2007.

[6] M. H. Azmi and J. Yuan. Performance of bilayer-lengthened LDPC codes under joint decoding. In *Proceedings of IEEE Info. Theory Workshop (ITW), Taormina, Sicily, Italy*, pages 163–167, Oct. 2009.

[7] M. H. Azmi, J. Yuan, G. Lechner, and L. K. Rasmussen. Design of multi-edge-type bilayer-expurgated LDPC codes for decode-and-forward in relay channels. 59:2993–3006, Nov. 2011.

[8] M. A. Khojastepour, N. Ahmed, and B. Aazhang. Code design for the relay channel and factor graph decoding. In *The Proceedings of Thirty-Eighth Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA*, pages 2000–2004, Nov. 2004.

[9] Z. Zhang and T. M. Duman. Capacity-approaching turbo coding and iterative decoding for relay channels. 53:1895–1905, Nov. 2005.

[10] J. Ezri and M. Gastpar. On the performance of independently designed LDPC codes for the relay channel. In *Proc. Int. Symp. Inform. Theory (ISIT), Seattle, WA, USA*, pages 977–981, July 2006.

[11] Y. Song and N. Devroye. Lattice codes for the Gaussian relay channel: decode-and-forward and compress-and-forward. *IEEE Trans. Info. Theory*, submitted, Oct. 2011.

[12] A. Khina, Y. Kochman, and U. Erez. Joint unitary triangularization for MIMO networks. *IEEE Trans. Signal Proc.*, 60:326–336, Jan. 2012.

[13] A. El Gamal and T.-H. Kim. *Network Information Theory*. Cambridge University Press, 2011.

[14] D. Blackwell, L. Breiman, and A. J. Thomasian. The capacity of a class of channels. *The Annals of Math. Stat.*, 30:1229–1241, Dec. 1959.

[15] B. Wang, J. Zhang, and A. Høst-Madsen. On the capacity of MIMO relay channels. *IEEE Trans. Info. Theory*, 51:29–43, Jan. 2005.

[16] Y. Jiang, W. Hager, and J. Li. Uniform channel decomposition for MIMO communications. *IEEE Trans. Signal Proc.*, vol. 53:4283–4294, Nov. 2005.

[17] I. Livni, A. Khina, A. Hitron, and U. Erez. Space-time MIMO multicasting. submitted to International Symposium on Information Theory (ISIT2012), Cambridge, MA, USA. Available at arXiv:1204.1677.