Decode-and-Forward Relaying via Standard AWGN Coding and Decoding

Anatoly Khina, Yuval Kochman, *Member, IEEE*, Uri Erez, *Member, IEEE*, and Gregory W. Wornell, *Fellow, IEEE*

***Abstract*—A framework is developed for decode-and-forward based relaying using standard coding and decoding that are good for the single-input single-output (SISO) additive white Gaussian noise channel. The framework is applicable to various scenarios and demonstrated for several important cases. Each of these scenarios is transformed into an equivalent Gaussian multiple-input multiple-output (MIMO) common-message broadcast problem, which proves useful even when all links are SISO ones. Over the effective MIMO broadcast channel, a recently developed Gaussian MIMO common-message broadcast scheme is applied. This scheme transforms the MIMO links into a set of parallel SISO channels with no loss of mutual information, using linear pre- and post-processing combined with successive decoding. Over these resulting SISO channels, "off-the-shelf" scalar codes may be used.**

***Index Terms*—Relay channel, rateless coding, MIMO channels, successive interference cancellation.**

I. Introduction

Relaying techniques are a key element in enhancing the performance of wireless networks. Accordingly, a great amount of research effort has been devoted 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, introduced in [1], [2], remains unknown, achievable rate regions, as well as outer bounds, have been established for basic models, following the pioneering work of [3]. The achievable regions 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) paradigms. We refer the reader to [4]–[6] for a review of these and other approaches.

Another key ingredient in enhancing the performance of wireless networks is the use of multiple-input multiple-output (MIMO) links [7], [8]. MIMO transmission can be used in conjunction with any relaying technique to further boost performance over that of single-input single-output (SISO) transmission.

The present work provides a unified framework for constructing coding schemes that approach the optimal performances of DF protocols over different Gaussian channel topologies, using only "off-the-shelf" scalar codes designed for the additive white Gaussian noise (AWGN) channel, for both SISO and MIMO links, where perfect channel knowledge is available at all transmission ends ("*closed loop*"). Specifically, we treat the important special cases of half-duplex (HD) and full-duplex (FD) transmission over the ("*triangular*") relay channel [1]–[3] (both SISO and MIMO). The approach proposed in this work is demonstrated for these basic settings, laying the foundations for treating more complex Gaussian networks in a unified manner.

The question of how to implement the full- and half-duplex DF protocols over the relay channel is non-trivial even for the SISO case, and has motivated numerous works that have proposed coding techniques tailored to DF relaying; see [9]–[15] and references therein for suggested coding techniques for full-duplex relaying. For half-duplex relaying, Mitran et al. [5] proposed a practical scheme, that employs rateless coding; transmission is divided into two phases: During the first ("*listening*" or "*broadcast*") phase, the source broadcasts to the relay and the destination, at the end of which the relay is able to recover the message conveyed by the source. During the second ("*collaboration*" or "*multiple-access*") phase, both the relay and the source transmit *coherently*, until the destination is able to decode the message as well.[1] We note that the practicality of the scheme assumes the availability of good rateless codes for the AWGN channel. Recent works addressing rateless coding for the AWGN channel include [16]–[18]. Since in such a transmission scheme, the relay decodes the message from only the first part of the codeword — while the destination from its full blocklength — DF over the HD relay channel is sometimes referred to as *rateless relaying* [6], [19];

The material in this paper was presenetd in part at the 2011 IEEE Information Theory Workshop (ITW), Paraty, Brazil, and at the 2012 IEEE ITW, Lausanne, Switzerland. The work of A. Khina was supported by the Feder Family Award, by the Yitzhak and Chaya Weinstein Research Institute for Signal Processing, and by the Clore Israel Foundation. The work of Y. Kochman was supported, in part, by the Israel Science Foundation (ISF) under grant No. 956/12. The work of U. Erez was supported, in part, by the ISF under Grant No. 1956/15. The work of G. W. Wornell was supported, in part, by the Air Force Office of Scientific Research (AFOSR) under Grant No. FA9550-11-1-0183, and the National Science Foundation (NSF) under Grant No. CCF-1319828.

A. Khina was with the Department of Electrical Engineering—Systems, Tel Aviv University, Tel Aviv 69978, Israel. He is now with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125, USA (e-mail: khina@caltech.edu).

Y. Kochman is with the School of Computer Science and Engineering, Hebrew University of Jerusalem, Jerusalem 91905, Israel (e-mail: yuvalko@cs.huji.ac.il).

U. Erez is with the Department of Electrical Engineering—Systems, Tel Aviv University, Tel Aviv 69978, Israel (e-mail: uri@eng.tau.ac.il).

G. W. Wornell is with the Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge 02139, USA (e-mail: gww@mit.edu).

[1]This scheme assumes a degraded relay channel, namely that the relay is able to recover the message prior to the destination; otherwise, the destination is able to decode the message before the relay, and this scheme reduces to that of direct transmission ("silent" relay) throughout the whole transmission process.

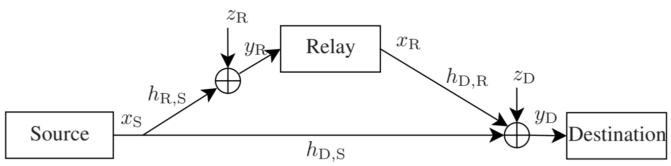


Fig. 1: SISO relay channel.

see [20]–[23] and references therein for works developing various code designs for HD DF.

In the framework proposed in this paper, there is no need for explicit design of rateless codes for a relay network. Rather, rateless coding naturally emerges as an integral part in the designed transmission schemes. In particular, the rateless transmission scheme for the scalar AWGN channel, developed previously in [16], is shown to be a special case of the design problem addressed in this paper, as well as its MIMO counterpart, and hence can be treated in a similar manner. The resulting rateless coding scheme may also serve as a building block in MIMO networks with several nodes, where one node wishes to transmit the same information to several users, where the latter nodes may wish to forward the decoded message further to other destinations.

We show that it is possible to reduce the coding task for the scenarios above to standard AWGN scalar coding, where encoding and decoding may be carried out using a *modulation–code separation* (MCS) approach. This is done by employing judiciously designed linear processing, in conjunction with successive interference cancellation (SIC). That is, encoder and decoder modules designed for SISO AWGN channels can be used in a straightforward manner while the specific relay architecture will influence only the linear processing employed.

The system architecture developed in this paper for the relaying problem builds on the work in [24], [25], where a novel joint unitary decomposition is derived and is tightly related to common-message broadcast (BC) (or *physical-layer multicast*) over Gaussian MIMO channels.

The rest of the paper is organized as follows. We begin in Section II by describing the half- and full-duplex SISO relay channel models and the corresponding decode-and-forward relaying variants. The coding schemes sought in this paper are formally defined in Section II-C, and a capacity-achieving scheme that uses such codes for the problem of broadcasting a common message over MIMO channels is presented in Section III. Building on this MIMO common-message BC scheme, we construct decode-and-forward schemes for the half- and full-duplex relay channel settings in Section IV, and extend them in Section V to the MIMO case. These results are then utilized for the derivation of a MIMO rateless coding scheme in Section VI. Finally, generalizations of the architecture to other input constraints and more relays, with or without "line-of-sight", are discussed in Section VII.

II. Decode-and-Forward Relaying

In the triangular relay channel model, a source sends a signal to a destination and a relay, where the relay transmits an additional signal to the destination. We consider the SISO relay channel setting, depicted in Fig. 1, in which all signals pass through SISO channels:

$$y_{\rm R} = h_{\rm R,S} x_{\rm S} + z_{\rm R} \,, \tag{1a}$$

$$y_{\rm D} = h_{\rm D,S} x_{\rm S} + h_{\rm D,R} x_{\rm R} + z_{\rm D} \,, \tag{1b}$$

where x, y and z denote channel input, output and noise vectors, respectively, using subscripts 'S', 'R' and 'D' to indicate '*source*', '*relay*' and '*destination*', respectively. The channel gains are denoted by h with two subscripts, where the first subscript indicates the channel output ("*destination node*"), and the second indicates the channel input ("*originating node*"). We assume full knowledge of all channel gains at all communication nodes (closed loop). We further assume that all noise processes are mutually independent and circularly-symmetric Gaussian. Without loss of generality we assume that these noise processes are white and of unit power, that all gains are real and non-negative,[2] and that all channel inputs are subject to the same average power constraint P.

Different settings have been considered for this channel, the two most prominent being half- and full-duplex relaying.[3] Since the capacity of this channel is not known in general (for half- or full-duplex), different transmission protocols have been proposed.

In the rest of this section, we recall the decode-and-forward half- and full-duplex relaying protocols, in Sections II-A and II-B, respectively. We formally define, in Section II-C, the class of "practical schemes" we are after, that materialize these protocols.

A. Half-Duplex Relaying

In the half-duplex setting, the relay may either receive or transmit at each time instant, but not both simultaneously. Thus, the system may be either in BC mode:

$$y_{\rm R} = h_{\rm R,S} x_{\rm S} + z_{\rm R} \,, \tag{2a}$$

$$y_{\rm D} = h_{\rm D,S} x_{\rm S} + z_{\rm D} \,, \tag{2b}$$

or in a multiple-access (MAC) mode:

$$y_{\rm D} = h_{\rm D,S} x_{\rm S} + h_{\rm D,R} x_{\rm R} + z_{\rm D} \,. \tag{3}$$

We consider the construction of a two-phase DF protocol for the case of $h_{\rm R,S} > h_{\rm D,S}$,[4] following the rateless relaying scheme proposed by Mitran et al. [5]. During the first (BC) phase, consisting of n_1 uses of the channel (2), both the destination and the relay receive information from the source, until the relay is able to decode; during the second (MAC) phase, consisting of $(n_2 - n_1)$ uses of the channel (3), the relay and the destination transmit coherently until the destination is able to decode the transmitted message as well. Assume that the ratio used between the total duration and the MAC duration

[2]Since any phase can be absorbed in the transmit signals.

[3]Other settings, such as their "non-coherent" variants (see, e.g., [26]), can be treated in a similar manner to the one proposed in this work.

[4]Otherwise, the relay channel is reversely degraded and the decode-and-forward scheme reduces to that of direct transmission, that is, point-to-point communication where the relay stays silent. This is indeed optimal in the HD case; see, e.g., [26] for further details.

is rational,[5] i.e., $n_1 = m_1 n$ and $n_2 = m_2 n$, for some positive integers m_1, m_2 and n. Denote by P_1 and P_2 the average power per channel use used by the source during the first and second phases, respectively; these power allocations need to satisfy the total power constraint

$$m_1 P_1 + (m_2 - m_1) P_2 \leq m_2 P.$$

The optimal allocation should be chosen to maximize the achievable rate, but its exact value is not material for the formulation of the protocol.

Denote further the average power used during the second phase by the relay by P_{R}, which satisfies, in turn, the constraint

$$(m_2 - m_1) P_{\mathrm{R}} \leq m_2 P.$$

Protocol (Half-duplex decode-and-forward relaying).

Offline:

- Choose admissible power allocations P_1, P_2 and P_{R}.
- Construct a good code of blocklength n_2 that can be recovered (w.h.p.) from either
  - n_1 channel uses over an AWGN channel with signal-to-noise ratio (SNR) equal to $|h_{\mathrm{R,S}}|^2 P_1$.
  - n_2 channel uses with the first n_1 channel uses having an SNR of $|h_{\mathrm{D,S}}|^2 P_1$, and the remaining $(n_2 - n_1)$ channel uses having an SNR of $\left(h_{\mathrm{D,S}}\sqrt{P_2} + h_{\mathrm{D,R}}\sqrt{P_{\mathrm{R}}}\right)^2$.

Source during BC phase: Transmits the first n_1 entries of the codeword with average power P_1.

Source during MAC phase: Transmits the remaining $(n_2 - n_1)$ entries of the codeword with average power P_2.

Relay during BC phase: Recovers the codeword from the first n_1 channel uses.

Relay during MAC phase: Transmits the last $(n_2 - n_1)$ entries of the codeword with average power P_{R}.

Destination: Receives the first n_1 (corrupted) codeword entries sent only by the source, and then the remaining $(n_2 - n_1)$ entries transmitted by both the source and the relay, and recovers the codeword.

The optimal achievable rate using this scheme is equal to (see, e.g., [5], [19])

$$\begin{aligned} R = \frac{1}{m_2} \min \Bigg\{ & m_1 \log\left(1 + |h_{\mathrm{R,S}}|^2 P_1\right), \\ & m_1 \log\left(1 + |h_{\mathrm{D,S}}|^2 P_1\right) \\ & + (m_2 - m_1) \log\left(1 + \left[h_{\mathrm{D,S}}\sqrt{P_2} + h_{\mathrm{D,R}}\sqrt{P_{\mathrm{R}}}\right]^2\right) \Bigg\}, \end{aligned} \tag{4}$$

where the multiplication by a factor of $1/m_2$ normalizes the rate per physical channel use.

[5]Since the rational numbers form a dense subset of the reals, any real ratio can be approached arbitrarily closely.

Note that the message is recovered after a different number of time instants at the relay and the destination. Moreover, the destination observes a "varying-SNR" channel, i.e., different SNRs during the two transmission phases of the protocol. This, in turn, complicates the implementation of this protocol.

B. Full-Duplex Relaying

In the full-duplex setting [3], the relay may receive and transmit simultaneously. Hence the DF protocol, in this case, has a sequential nature: The data is partitioned into a sequence of messages $\{m[i]\}$. At each time instant, the source and relay transmit functions of a "sliding window" of messages, and the relay and destination decode messages sequentially.

In this subsection we consider explicitly transmission blocks, represented by row vectors and denoted by $x^n \triangleq \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}$.

Detailed Protocol (Full-duplex decode-and-forward relaying).

Offline:

- Generate two different (*independent*) good AWGN codebooks of the same length n, power P, and rate to be specified in the sequel. We denote the codebooks by $\mathcal{C}_{\mathrm{C}}$ and $\mathcal{C}_{\mathrm{I}}$, where the subscript 'C' stands for 'coherent', and 'I' — for 'incoherent', the operational meanings of which will become apparent shortly.
- Denote by ρ^2 the portion of the available power P that is allocated to to codebook $\mathcal{C}_{\mathrm{C}}$, where $0 \leq \rho \leq 1$; the remaining power $(1 - \rho^2)P$ is allocated to $\mathcal{C}_{\mathrm{I}}$.

Source: For each block i, the information message $m[i]$ is encoded into two codewords, $x_{\mathrm{C}}^n[i] \in \mathcal{C}_{\mathrm{C}}$ and $x_{\mathrm{I}}^n[i] \in \mathcal{C}_{\mathrm{I}}$. The i-th block signal of length n, $x_{\mathrm{S}}^n[i]$, sent by the source, is equal to the sum of $x_{\mathrm{C}}^n[i]$ and $x_{\mathrm{I}}^n[i+1]$. Thus, the signal in block (or "time frame") i sent by the source, $x_{\mathrm{S}}^n[i]$, is equal to

$$x_{\mathrm{S}}^n[i] = \rho x_{\mathrm{C}}^n[i] + \sqrt{1 - \rho^2}\, x_{\mathrm{I}}^n[i+1]\,. \tag{5}$$

Relay as Receiver: At time frame $(i-1)$, $m[i-1]$ is assumed to be known to the relay (assuming correct decoding of previous codewords, and predetermined $m[1]$), and the output at the relay is

$$\begin{aligned} y_{\mathrm{R}}^n[i-1] &= x_{\mathrm{S}}^n[i-1] + z_{\mathrm{R}}^n[i-1] \\ &= \rho h_{\mathrm{R,S}} x_{\mathrm{C}}^n[i-1] + \sqrt{1-\rho^2}\, h_{\mathrm{R,S}} x_{\mathrm{I}}^n[i] + z_{\mathrm{R}}^n[i-1]\,. \end{aligned}$$

From this *single* output block, $m[i]$ is recovered by decoding $x_{\mathrm{I}}^n[i]$, where the contribution of $x_{\mathrm{C}}^n[i-1]$ (which is known, since $m[i-1]$ is known) is subtracted from $y_{\mathrm{R}}^n[i-1]$, resulting in:

$$\tilde{y}_{\mathrm{R}}^n[i-1] = y_{\mathrm{R}}^n[i-1] - h_{\mathrm{R,S}} \rho x_{\mathrm{C}}^n[i-1] \tag{6a}$$

$$= \sqrt{1-\rho^2}\, h_{\mathrm{R,S}} x_{\mathrm{I}}^n[i] + z_{\mathrm{R}}^n[i-1]\,. \tag{6b}$$

Relay as Transmitter: At time frame i, the relay knows $m[i]$ (and hence also $x_{\mathrm{C}}^n[i]$) and sends

$$x_{\mathrm{R}}^n[i] = x_{\mathrm{C}}^n[i]\,. \tag{7}$$

Destination: At each time frame i, the destination recovers $m[i]$ based on *two* consecutive output blocks $y_{\mathrm{D}}^n[i-1]$ and

$y_{\mathrm{D}}^n[i]$, assuming $m[i-1]$ is known (was decoded correctly from previous outputs, except for $m[1]$ which is predetermined). In essence, the destination recovers $m[i]$ from two observations of the encoded message: the first being a noisy version of $x_{\mathrm{I}}^n[i]$, whereas the other being a noisy version of $x_{\mathrm{C}}^n[i]$.

Specifically, we subtract $h_{\mathrm{D,S}}\rho x_{\mathrm{C}}^n[i-1]$ from $y_{\mathrm{D}}^n[i-1]$ to arrive at

$$\begin{aligned}\tilde{y}_{\mathrm{D}}^n[i-1] &= y_{\mathrm{D}}^n[i-1] - h_{\mathrm{D,S}}\rho x_{\mathrm{C}}^n[i-1] &\text{(8a)}\\ &= \sqrt{1-\rho^2}h_{\mathrm{D,S}}x_{\mathrm{I}}^n[i] + z_{\mathrm{D}}^n[i-1]\,. &\text{(8b)}\end{aligned}$$

This serves as the first noisy observation.

The second noisy observation is that of $x_{\mathrm{C}}^n[i]$, which is obtained from $y_{\mathrm{D}}^n[i]$ as follows. The components of $x_{\mathrm{C}}^n[i]$ in the source signal $x_{\mathrm{S}}^n[i]$ and the signal transmitted by the relay $x_{\mathrm{R}}^n[i]$ sum up coherently, whereas $x_{\mathrm{I}}^n[i+1]$ plays the role of noise, namely,[6]

$$y_{\mathrm{D}}^n[i] = (\rho h_{\mathrm{D,S}} + h_{\mathrm{D,R}})\, x_{\mathrm{C}}^n[i] + z_{\text{equiv}}^n[i]\,, \tag{9}$$

where

$$z_{\text{equiv}}^n[i] = \sqrt{1-\rho^2}h_{\mathrm{D,S}}x_{\mathrm{I}}^n[i] + z_{\mathrm{D}}^n[i] \tag{10}$$

is of power $P_{\text{equiv}} = \left(1-\rho^2\right)h_{\mathrm{D,S}}^2 P + 1$. Normalizing the power of the noise $\boldsymbol{z}_i^{\text{equiv}}$, i.e., dividing $y_{\mathrm{D}}^n[i]$ by $\sqrt{P_{\text{equiv}}}$, we arrive at

$$\begin{aligned}\bar{y}_{\mathrm{D}}^n[i] &= \frac{1}{\sqrt{P_{\text{equiv}}}} y_{\mathrm{D}}^n[i] &\text{(11a)}\\ &= \frac{\rho h_{\mathrm{D,S}} + h_{\mathrm{D,R}}}{\sqrt{\left(1-\rho^2\right)h_{\mathrm{D,S}}^2 P + 1}} x_{\mathrm{C}}^n[i] + \bar{z}_{\mathrm{D}}^n[i]\,, &\text{(11b)}\end{aligned}$$

where $\bar{z}_{\mathrm{D}}^n[i]$ is of unit power.

Note that the relay recovers $m[i]$ based on only a single observation (6), whereas the destination uses two observations (8) and (11) to recover the same information. Thus, assuming independent Gaussian codebooks, the achievable rate of the DF protocol is limited by the minimum of the mutual informations of the two [3] (see also [26]):[7]

$$\begin{aligned}R_{\mathrm{DF}} = \min\Bigg\{ &\log\left(1 + \left(1-\rho^2\right)h_{\mathrm{R,S}}^2 P\right), \\ &\log\left(1 + \left(1-\rho^2\right)h_{\mathrm{D,S}}^2 P\right) \\ &+ \log\left(1 + \frac{\left(\rho h_{\mathrm{D,S}} + h_{\mathrm{D,R}}\right)^2 P}{\left(1-\rho^2\right)h_{\mathrm{D,S}}^2 P + 1}\right)\Bigg\} &\text{(12a)}\\ = \min\Bigg\{ &\log\left(1 + \left(1-\rho^2\right)S_{\mathrm{R,S}}\right), \\ &\log\left(1 + S_{\mathrm{D,S}} + S_{\mathrm{D,R}} + 2\rho\sqrt{S_{\mathrm{D,S}}S_{\mathrm{D,R}}}\right)\Bigg\}, &\text{(12b)}\end{aligned}$$

where we define the SNRs as $S \triangleq h^2 P$ where h may correspond to $h_{\mathrm{D,R}}$, $h_{\mathrm{D,S}}$ or $h_{\mathrm{R,S}}$.

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 (12), can be evaluated; see, e.g., [26].

[6] Again, we assume, w.l.o.g., that $h_{\mathrm{D,S}}$ and $h_{\mathrm{D,R}}$ are real and positive.

[7] Assuming many blocks, the loss due to predetermining $m[1]$ is negligible.

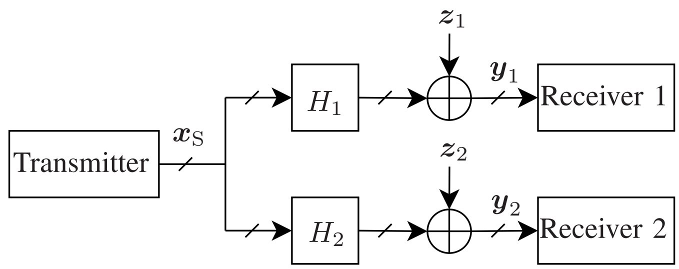


Fig. 2: Two-user MIMO broadcast channel.

Even though the rate of the FD DF protocol is fully determined, it is not clear how to achieve it using a *practical scheme*. The main difficulty is in how to combine the information carried by the two codewords x_{C}^n and x_{I}^n (drawn from two independent codebooks, but carrying the same information) and at the same time recover it at the relay. Different approaches, e.g., list decoding, were proposed, but these are still hard to implement in practice; see [4] 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 (12), as defined next.

Remark 2. Using the same codebook for both x_{C}^n and x_{I}^n, i.e., $\mathcal{C}_{\mathrm{C}} \equiv \mathcal{C}_{\mathrm{I}}$, and maximum-ratio combining (MRC), is suboptimal, since in order to maximize mutual information, the channel inputs in the two channel blocks need to be statistically independent.

C. Black-Box Coding Schemes

In the present work we follow a “black box” approach to coding: We only allow the use of standard codes that are good for SISO point-to-point AWGN channels, along with linear pre- and post-processing and SIC. This way, the linear processing and the SIC procedures allow to decouple the coding task from the modulation, where only the latter is tailored to the specific channel topology. We further want any performance loss of the whole transmission scheme to be governed solely by the loss of the SISO point-to-point AWGN codebooks. We refer to such schemes as MCS schemes.[8]

In the sequel we shall show that both of the DF variants, namely half- and full-duplex, can be transformed into equivalent MIMO common-message BC problems. We therefore describe next how to construct an MCS scheme for MIMO common-message BC.

III. MCS Schemes for MIMO Common-Message Broadcast via Matrix Decomposition

We now describe the main tool used in this work. We review the results of [24] where an MCS scheme is introduced for

[8] Of course, since SIC is employed, error propagation needs to be taken into account, and can bounded in a simple manner using the union bound. Due to these issues, we make no claims for optimality in terms of error probability, but only in terms of achievable rates.

MIMO common-message BC. The channel model, depicted in Fig. 2, is described by

$$\boldsymbol{y}_i = H_i \boldsymbol{x}_{\mathrm{S}} + \boldsymbol{z}_i \,, \tag{13}$$

where $\boldsymbol{y}_i$ is the received $M_i \times 1$ vector of user i ($i = 1, 2$), $\boldsymbol{x}_{\mathrm{S}}$ denotes the $N_{\mathrm{S}} \times 1$ complex-valued input vector which is limited to an average power P, H_i is the $M_i \times N_{\mathrm{S}}$ complex channel matrix to user i, and $\boldsymbol{z}_i$ is assumed to be a circularly-symmetric Gaussian vector of length M_i with identity covariance matrix.

For communication to a single user, the rate achievable for an $M \times N_{\mathrm{S}}$ 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

$$C(H, K) \triangleq \log \left| I + HKH^\dagger \right| \,, \tag{14}$$

where $|\cdot|$ denotes the determinant and I is an identity matrix of dimension M. The common-message BC capacity of this channel is given by the compound-channel capacity [27]–[29]

$$C_{\text{common}} = \max_K \min_{i=1,2} C(H_i, K) \,. \tag{15}$$

The scheme of [24], which achieves the capacity (15), relies upon the following joint matrix decomposition, also developed in [24].

Theorem 1 (Joint equi-diagonal triangularization [JET]). *Let A_1 and A_2 be complex-valued full-rank matrices of dimensions $M_1 \times N$ and $M_2 \times N$, respectively, such that $M_1 \geq N$ and $M_2 \geq N$ (implying that A_i are of rank N). If $\det(A_1^\dagger A_1) = \det(A_2^\dagger A_2)$, then A_1 and A_2 can be jointly decomposed as*

$$\begin{aligned} A_1 &= U_1 T_1 V^\dagger \\ A_2 &= U_2 T_2 V^\dagger \,, \end{aligned}$$

where U_1, U_2 and V are unitary matrices of dimensions $M_1 \times M_1$, $M_2 \times M_2$ and $N \times N$, respectively; T_1 and T_2 are generalized upper-triangular matrices of dimensions $M_1 \times N$ and $M_2 \times N$, respectively, with positive equal *diagonal elements, viz.*

$$\begin{aligned} T_{1;ij} &= T_{2;ij} = 0, \quad i > j \\ T_{1;ii} &= T_{2;ii} \triangleq t_i \,, \end{aligned}$$

where $T_{k;ij}$ denotes the (i, j) entry of T_k.

The following definition will prove useful in applying Theorem 1 to common-message BC, as well as all other communication settings considered in the sequel.

Definition 1 (Effective MMSE matrix). Let H be a channel matrix of dimensions $M \times N_{\mathrm{S}}$ and let K be the $N_{\mathrm{S}} \times N_{\mathrm{S}}$ input covariance matrix used over this channel. Then, the corresponding *effective minimum mean square error (MMSE) matrix* is the $(M + N_{\mathrm{S}}) \times N_{\mathrm{S}}$ matrix G:

$$G \triangleq \begin{pmatrix} HB \\ I \end{pmatrix} \,,$$

where I is the identity matrix of dimension N_{S} and B is any matrix satisfying $BB^\dagger = K$.[9]

This definition serves as the MIMO channel analogue of the MMSE variant of decision feedback equalization for linear time-invariant systems [30].

By applying the decomposition of Theorem 1 to the effective MMSE matrices of (13), the following scheme, reminiscent of V-BLAST [30]–[32], may be devised, which transforms the problem to that of transmission over parallel SISO channels.

Remark 3. We assume that the channel matrices H_1 and H_2, and the input covariance matrix K used, satisfy $C(H_1, K) = C(H_2, K)$. This incurs no loss of generality, since the common-message capacity is limited to the minimum between the two (15). This, in turn, means that the corresponding effective MMSE matrices, G_1 and G_2, satisfy $|G_1^\dagger G_1| = |G_2^\dagger G_2|$. In practice, the user having larger capacity will enjoy, in each scalar sub-channel in the scheme to follow, excess SNR that will not be utilized, due to the bottleneck being the other user. For a more thorough account, see [24], [25, Example 3].

Scheme (MIMO common-message broadcast).

Offline:

- Select an admissible $N_{\mathrm{S}} \times N_{\mathrm{S}}$ input covariance matrix K that satisfies the input power constraint and an $N_{\mathrm{S}} \times N_{\mathrm{S}}$ matrix B satisfying $BB^\dagger = K$.
- Construct the effective MMSE matrices G_1 and G_2, of dimensions $(M_1 + N_{\mathrm{S}}) \times N_{\mathrm{S}}$ and $(M_2 + N_{\mathrm{S}}) \times N_{\mathrm{S}}$, respectively, corresponding to the channel matrices H_1 and H_2 of (13) and K.
- Apply the JET of Theorem 1 to G_1 and G_2, to obtain the unitary matrices U_1, U_2, and V, of dimensions $(M_1 + N_{\mathrm{S}})$, $(M_2 + N_{\mathrm{S}})$, and N_{S}, respectively, and the generalized upper-triangular matrices T_1 and T_2 of dimensions $(M_1 + N_{\mathrm{S}}) \times N_{\mathrm{S}}$ and $(M_2 + N_{\mathrm{S}}) \times N_{\mathrm{S}}$, respectively.
- Denote the N_{S} diagonal elements of T_1 and T_2 by $\{t_j\}$ (which are equal for both matrices).
- Denote by $\tilde{U}_k$ ($k = 1, 2$) the $M_k \times N_{\mathrm{S}}$ sub-matrix of U_k composed of its first M_k rows.
- Construct good codes for scalar AWGN channels of SNRs $\{t_j^2 - 1\}$ and blocklength n.

Transmitter:

- Constructs n vectors $\tilde{\boldsymbol{x}}_{\mathrm{S}}$ of length N_{S} each, by taking one sample from each codebook.
- Combines all these codewords by multiplying each $\tilde{\boldsymbol{x}}_{\mathrm{S}}$ by the unitary matrix V and by B:

$$\boldsymbol{x}_{\mathrm{S}} = BV\tilde{\boldsymbol{x}}_{\mathrm{S}} \,.$$

- Transmits the n vectors $\boldsymbol{x}_{\mathrm{S}}$.

Receiver $\boldsymbol{k}$ ($\boldsymbol{k = 1, 2}$):

- Receives the n output vectors $\boldsymbol{y}_k$.

[9] Such a B can always be constructed, e.g., using the Cholesky decomposition or unitary diagonalization.

- For each $\tilde{\boldsymbol{y}}_k$, calculates
$$\tilde{\boldsymbol{y}}_k = \tilde{U}_k^\dagger \boldsymbol{y}_k \,.$$
- Decodes the codewords using successive interference cancellation, starting from the N_S-th codeword and ending with the first one: The N_S-th codeword is decoded first, using the N_S-th element of $\tilde{\boldsymbol{y}}_k$, treating the other codewords as AWGN. The effect of the N_S-th element of $\tilde{\boldsymbol{x}}_\mathrm{S}$ is then subtracted out from the remaining elements of $\tilde{\boldsymbol{y}}_k$. Next, the $(N_\mathrm{S}-1)$ codeword is decoded, using the $(N_\mathrm{S}-1)$ element of $\tilde{\boldsymbol{y}}_k$ — and so forth.

Using capacity-achieving codes of SNRs $\{t_j^2-1\}$, i.e., codes of rates close to
$$\left\{R_j \middle| R_j = \log\left(t_j^2\right), j \in \{1,\dots,N_\mathrm{S}\}\right\}, \tag{16}$$
the whole scheme achieves the common-message BC capacity (15). This is formally proved in [24], and is, in turn, a simple extension of the optimality proof of V-BLAST for the MIMO single-user ("point-to-point") channel [33] (see also [34, Appendix A] and [35]).

Remark 4. The resulting diagonal values $\{t_j\}$ are greater or equal to 1, due to the presence of I in the construction of the effective MMSE matrices G_1 and G_2 (recall Definition 1). Thus, the rates $\{R_j\}$ in (16) are all non-negative.

Remark 5. Using an input covariance matrix K over the channels described by the matrices H_1 and H_2, is mathematically equivalent to working with a unit covariance matrix I ("white input") over equivalent channel matrices $F_1 \triangleq H_1 B$ and $F_2 \triangleq H_2 B$, respectively.

We next show how this MIMO common-message BC scheme can be leveraged for the construction of practical DF schemes for the half- and full-duplex relay channel settings.

IV. MCS Schemes for Decode-and-Forward Relaying

In this section we construct MCS schemes for half- and full-duplex decode-and-forward relaying. We do this by translating the two problems to equivalent MIMO common-message BC ones, which allows us, in turn, to apply the scheme of Section III.

A. Half-Duplex Relaying

Assume, as in Section II-A, that the ratio used between the total duration and the MAC duration is rational, i.e., $n_1 = m_1 n$ and $n_2 = m_2 n$, for some positive integers m_1, m_2 and n. We view the n_i $(i = 1,2)$ channel uses as n channel uses of an "*augmented channel*" with m_2 "transmit antennas" (corresponding to m_2 physical channel uses batched together). The number of "receive antennas" of the augmented channel corresponds to the number of channel uses needed by the relay and the destination for the recovery of the message. Denote by I_m the identity matrix of dimension m, and by $0_{k\times\ell}$ the all-zero matrix of dimensions $k\times\ell$. Hence, the $m_1 \times m_2$ augmented matrix from the source to the relay is
$$\underline{H}_\mathrm{R,S} = \begin{pmatrix} h_\mathrm{R,S} I_{m_1} & 0_{m_1\times m_2-m_1}\end{pmatrix}, \tag{17}$$
whereas the $m_2 \times m_2$ augmented matrices from the source to the destination and the relay have the following forms
$$\underline{H}_\mathrm{D,S} = h_\mathrm{D,S} I_{m_2}\,,$$
$$\underline{H}_\mathrm{D,R} = h_\mathrm{D,R} I_{m_2}\,.$$

Note that, for notational convenience, we assume that the relay can receive and transmit at all times, but during the BC and MAC phases it has zero transmit and receive gains, respectively. The latter is manifested in the zero columns in (17), whereas the former is represented by the following $m_2 \times m_2$ power-allocation matrices
$$\underline{B}_\mathrm{S} = \begin{pmatrix} \sqrt{P_1} I_{m_1} & 0_{m_1\times m_2-m_1} \\ 0_{m_2-m_1\times m_1} & \sqrt{P_2} I_{m_2-m_1}\end{pmatrix},$$
$$\underline{B}_\mathrm{R} = \begin{pmatrix} \multicolumn{2}{c}{0_{m_1\times m_2}} \\ 0_{m_2-m_1\times m_1} & \sqrt{P_\mathrm{R}} I_{m_2-m_1}\end{pmatrix},$$
where the input power constraints are translated into
$$\mathrm{trace}\left\{\underline{B}_\mathrm{S}^2\right\} \le m_2 P\,,$$
$$\mathrm{trace}\left\{\underline{B}_\mathrm{R}^2\right\} \le m_2 P\,.$$
Moreover, since during the MAC stage the relay has full knowledge of the message, we may view the source and the relay as a single node that may coherently transmit over an effective point-to-point channel to the destination.

The transmission task, therefore, reduces to that of two-user MIMO common-message BC over
$$\underline{\boldsymbol{y}}_\mathrm{R} = \underline{H}_1 \underline{\boldsymbol{x}} + \underline{\boldsymbol{z}}_\mathrm{R}\,, \tag{18a}$$
$$\underline{\boldsymbol{y}}_\mathrm{D} = \underline{H}_2 \underline{\boldsymbol{x}} + \underline{\boldsymbol{z}}_\mathrm{D}\,, \tag{18b}$$
where $\underline{H}_1$ and $\underline{H}_2$ are the total augmented channel matrices of dimensions $m_1 \times m_2$ and $m_2 \times m_2$, to the relay and the destination:
$$\underline{H}_1 = \underline{H}_\mathrm{R,S} \underline{B}_\mathrm{S}\,, \tag{19a}$$
$$\underline{H}_2 = \underline{H}_\mathrm{D,S} \underline{B}_\mathrm{S} + \underline{H}_\mathrm{D,R} \underline{B}_\mathrm{R}\,. \tag{19b}$$
$\underline{\boldsymbol{x}}$, $\underline{\boldsymbol{y}}$ and $\underline{\boldsymbol{z}}$ represent the corresponding *augmented* input, output and noise vectors, respectively.

Note that the m_2-length input vector $\underline{\boldsymbol{x}}$ has entries of unit power, as the power coefficients were absorbed in the total augmented channel matrices (19).

Over these effective channel matrices, the MCS scheme for MIMO common-message BC of Section III may be readily applied, achieving a rate of
$$R = \min_{i=1,2} \frac{1}{m_2} C\left(\underline{H}_i, I\right)\,, \tag{20}$$
where again we multiply by a factor $1/m_2$ to normalize the rate per physical channel use. This rate is equal, in turn, to the desired rate of (4).

Note that, in order to materialize a single input vector $\underline{\boldsymbol{x}}$ of length m_2, m_2 physical channel inputs x_S are required. Since only the first m_1 columns of $\underline{H}_\mathrm{R,S}$ are non-zero, sending $\underline{\boldsymbol{x}}$ or its first m_1 entries followed by zeros is equivalent over this channel. Thus, by transmitting, during each of the first n_1 (physical) channel uses, only the first m_1 entries of the appropriate augmented input $\underline{\boldsymbol{x}}$, we are able to effectively transmit the whole $\underline{\boldsymbol{x}}$ vector over the channel $\underline{H}_{\mathrm{R,S}}$. To accommodate for the second user (the columns of which are all non-zero, in general), the source then sends over the remaining $(n_2 - n_1)$ *physical* channel uses x_{S} — the remaining parts of the n_1 *augmented* vectors $\underline{\boldsymbol{x}}$, in (any) systematic manner; see Remark 7 for further discussion of the order of transmission and Fig. 3 — for two such options. During the MAC phase, the relay joins the source in transmitting a linear combination of the codewords. That is, it transmits $\underline{\boldsymbol{x}}$ multiplied by the unitary matrix V and by $\sqrt{P_{\mathrm{R}}}$.

For completeness of presentation, we next describe the entire MCS scheme for HD-DF relaying (with the MIMO common-message BC scheme of Section III encapsulated within the half-duplex protocol of Section II-A).

Scheme (Half-duplex decode-and-forward relaying).

Offline:

- Choose the power allocations P_1, P_2 and P_{R}.
- Construct the effective ("total augmented") channel matrices (19).
- Construct the effective MMSE matrices $\underline{G}_1$ and $\underline{G}_2$, of dimensions $(m_1{+}m_2){\times}m_2$ and $2m_2{\times}m_2$, corresponding to the effective channel matrices $\underline{H}_1$ and $\underline{H}_2$, respectively.
- Apply the JET of Theorem 1 to $\underline{G}_1$ and $\underline{G}_2$, to obtain the unitary matrices U_1, U_2, and V, of dimensions $(m_1 + m_2)$, $2m_2$, and m_2, respectively, and the generalized upper-triangular matrices T_1 and T_2 of dimensions $(m_1 + m_2) \times m_2$ and $2m_2 \times m_2$, respectively.
- Denote the m_2 diagonal elements of T_1 and T_2 by $\{t_j\}$ (which are equal for both matrices).[10]
- Denote by $\tilde{U}_k$ $(k = 1, 2)$ the $m_k \times m_2$ sub-matrix composed of the first m_k rows of U_k.
- Construct good codes for scalar AWGN channels of SNRs $\{t_j^2 - 1\}$ and blocklength n.

Source:

- Constructs n vectors $\underline{\tilde{\boldsymbol{x}}}$ of length m_2 each, by taking one sample from each codebook.
- Combines all these codewords by multiplying each $\underline{\tilde{\boldsymbol{x}}}$ by the unitary matrix V and by $\underline{B}_{\mathrm{S}}$, both of dimensions $m_2 \times m_2$:

$$\underline{\boldsymbol{x}}_{\mathrm{S}} = \underline{B}_{\mathrm{S}} V \underline{\tilde{\boldsymbol{x}}} \, .$$

Source during BC phase: Transmits the first m_1 entries of each of the n vectors $\underline{\boldsymbol{x}}_{\mathrm{S}}$.

Source during MAC phase: Transmits the rest of the $(m_2 - m_1)$ entries of each of the n vectors $\underline{\boldsymbol{x}}_{\mathrm{S}}$.[11]

Relay during BC phase:

- After receiving the first $n_1 = m_1 n$ output vectors $\boldsymbol{y}_{\mathrm{R}}$, constructs the n vectors $\underline{\boldsymbol{y}}_{\mathrm{R}}$.
- Multiples each vector $\underline{\boldsymbol{y}}_{\mathrm{R}}$ by $\tilde{U}_1^\dagger$.
- Recovers the codewords using SIC, resulting in SINR values of $\{t_j^2 - 1\}$.

Relay during MAC phase: Transmits the last $(m_2 - m_1)$ entries of $\underline{\boldsymbol{x}}_{\mathrm{R}}$, which is equivalent to sending $\underline{\tilde{\boldsymbol{x}}}$ multiplied by the unitary matrix V and by $\underline{B}_{\mathrm{R}}$:

$$\underline{\boldsymbol{x}}_{\mathrm{R}} = \underline{B}_{\mathrm{R}} V \underline{\tilde{\boldsymbol{x}}},$$

since the first m_1 rows of $\underline{B}_{\mathrm{R}}$ are all zero.

Destination:

- Receives the first n_1 vectors sent only by the source, and then the remaining $(n_2 - n_1)$ vectors transmitted by both, the source and the relay.
- Constructs the n output vectors $\underline{\boldsymbol{y}}_{\mathrm{D}}$ out of the $n_2 = m_2 n$ received vectors $\boldsymbol{y}_{\mathrm{D}}$.
- Multiplies each vector $\underline{\boldsymbol{y}}_{\mathrm{D}}$ by $\tilde{U}_2^\dagger$.
- Recovers the codebooks using SIC, resulting in SINR values of $\{t_j^2 - 1\}$.

Using capacity-achieving codes designed for AWGN channels, i.e., codes of rates close to

$$\left\{ R_j \middle| R_j = \log\left(t_j^2\right) \, , j \in \{1, \dots, m_2\} \right\} ,$$

the desired rate of (4) is achieved.

Thus, we were able to construct an MCS scheme that achieves a total transmission rate approaching the optimum (for this protocol).

Remark 6. The scheme uses m_2 parallel codes. In practice, a large number of codes may have a negative impact (due to considerations such as error propagation and channel coherence time). Thus, when implementing such a scheme, one may choose to use an approximate fraction m_1/m_2 with smaller denominator, striking a balance between the different losses.

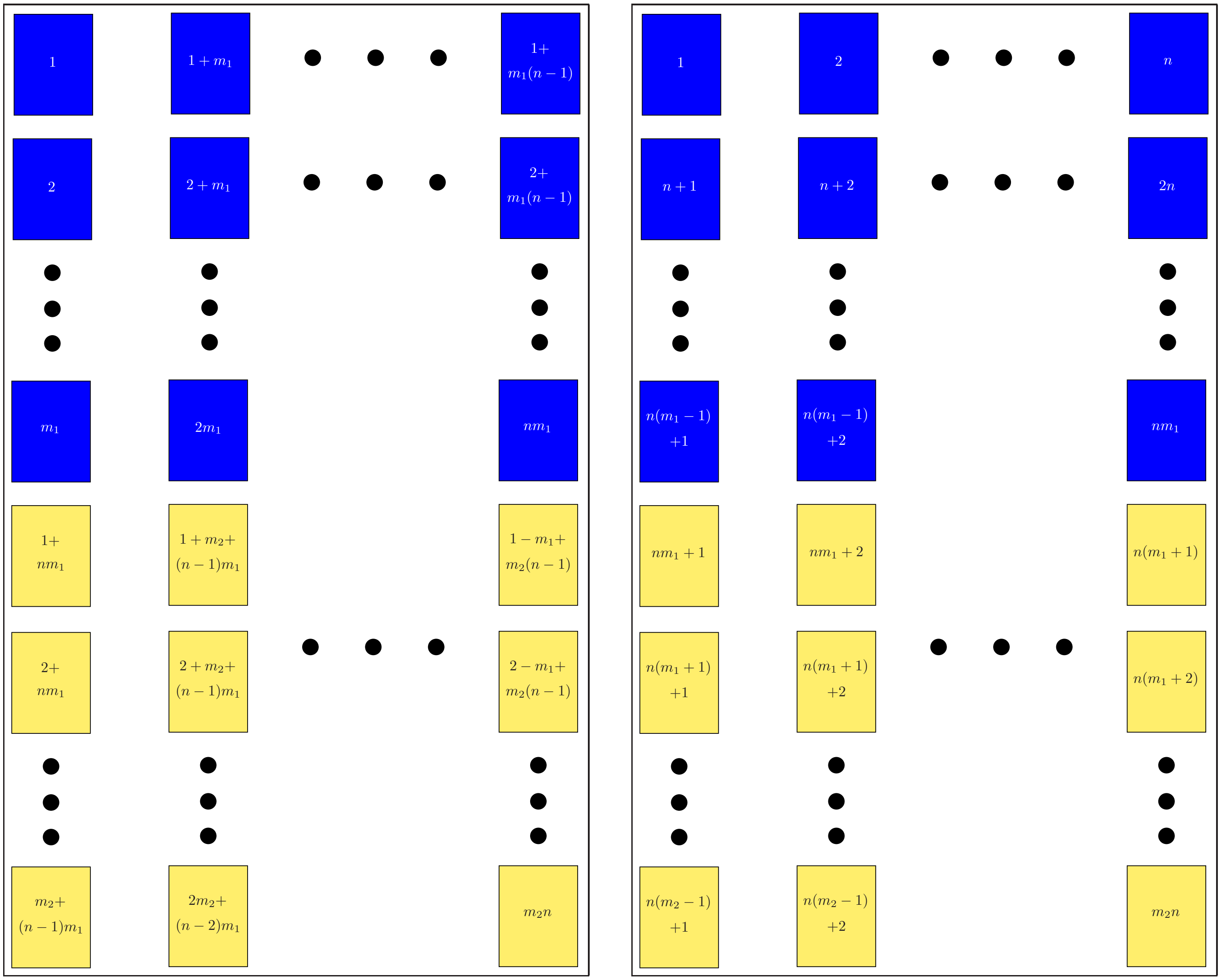
Remark 7. Different orderings ("interleaving") of the entries of $\boldsymbol{x}_{\mathrm{S}}$ (and $\boldsymbol{x}_{\mathrm{R}}$) are possible. Any ordering/interleaving of the first m_1 entries of each of the n_1 vectors $\boldsymbol{x}_{\mathrm{S}}$, which constitute the first entries of the n_1 vectors $\underline{\tilde{\boldsymbol{x}}}$, followed by any ordering of the rest, is applicable [achieves (20)]. Two orderings are of particular interest. The first, illustrated in Fig. 3a, is the one that enables to construct the vectors $\underline{\boldsymbol{y}}_{\mathrm{R}}$ and $\underline{\boldsymbol{y}}_{\mathrm{D}}$ in the fastest manner possible; this, in turn, enables to process these vectors (applying the unitary transformations $\tilde{U}_1$ and $\tilde{U}_2$) as early as possible. The second ordering of interest, depicted in Fig. 3b, is the systematic ordering according to which first the first entry ("first layer") of each of the vectors $\boldsymbol{x}_{\mathrm{S}}$ is sent, followed by the second entry ("second layer") of each of the vectors $\boldsymbol{x}_{\mathrm{S}}$, etc.; this ordering can be implemented more easily, being "more systematic", and is suitable for cases of more relays, where each relay starts to transmit after a different number of channel uses. In this case, the first "layer" is the least common multiple (LCM) of $\{n_i\}$ — the number of uses needed by each of the relays and the destination to recover reliably the transmitted codewords.

B. Full-Duplex Relaying

We now show how to materialize the FD-DF relaying protocol via an MCS scheme. To that end, we show how this

[10] Similarly to Remark 3, we assume, that $|G_1^\dagger G_1| = |G_2^\dagger G_2|$. Otherwise, the achievable rate is limited to the minimum of the two determinants and can be consequently improved by varying the transmission duration time of the relay.

[11] See Remark 7 for discussion of the transmission order.



(a) Transmission ordering that enables efficient processing (processing can start earlier than in other orderings) of the received signals.

(b) Systematic transmission ordering suitable for schemes with more than two users.

Fig. 3: Examples of interleaving/ordering of transmitted signals. Each rectangle represents a *physical* input signal x_{S}; each column represents an augmented input vector $\underline{\boldsymbol{x}}_{\mathrm{S}}$ of length m_2; the indices inside the rectangle represent the transmission ordering; rectangles with dark fill are received by both the relay and the destination, whereas the lighter-filled rectangles are designated to the end destination only.

protocol may be formulated as an equivalent MIMO common-message BC one. This allows, in turn, to apply the scheme of Section III.

We extend the notation x^n as follows. $\boldsymbol{x}^n \triangleq \begin{pmatrix} \boldsymbol{x}_1 & \cdots & \boldsymbol{x}_n \end{pmatrix}$ is a matrix of dimensions $m \times n$ of m-length column vectors $\boldsymbol{x}$.

Recall that the relay uses only a *single* observation, as reflected in (6), to recover $m[i]$. The destination, on the other hand, makes use of *two* consecutive observation blocks, as given in (8) and (11), to recover the same message $m[i]$. This can be reformulated in an equivalent matrix notation as

$$
\begin{aligned}
\underline{\boldsymbol{y}}_{\mathrm{R}}^n[i] &= \underline{H}_{\mathrm{R}} \underline{\boldsymbol{x}}_{\mathrm{S}}^n[i] + \underline{\boldsymbol{z}}_{\mathrm{R}}^n[i] \,, \\
\underline{\boldsymbol{y}}_{\mathrm{D}}^n[i] &= \underline{H}_{\mathrm{D}} \underline{\boldsymbol{x}}_{\mathrm{S}}^n[i] + \underline{\boldsymbol{z}}_{\mathrm{D}}^n[i] \,,
\end{aligned}
$$

where

$$
\underline{H}_{\mathrm{R}} = \sqrt{2} \begin{pmatrix} \sqrt{1-\rho^2} h_{\mathrm{R,S}} & 0 \end{pmatrix} \tag{21a}
$$

$$
\underline{H}_{\mathrm{D}} = \sqrt{2} \begin{pmatrix} \sqrt{1-\rho^2} h_{\mathrm{D,S}} & 0 \\ 0 & \frac{\rho h_{\mathrm{D,S}} + h_{\mathrm{D,R}}}{\sqrt{((1-\rho^2) h_{\mathrm{D,S}}^2 P + 1}} \end{pmatrix} \tag{21b}
$$

are the effective channel matrices,

$$
\begin{aligned}
\underline{\boldsymbol{z}}_{\mathrm{R}}^n[i] &\triangleq z_{\mathrm{R}}^n[i-1] \\
\underline{\boldsymbol{z}}_{\mathrm{D}}^n[i] &\triangleq \begin{pmatrix} z_{\mathrm{D}}^n[i-1] \\ \tilde{z}_{\mathrm{D}}^n[i] \end{pmatrix}
\end{aligned}
$$

are additive white noise vectors with identity covariance matrices,

$$
\underline{\boldsymbol{x}}_{\mathrm{S}}^n[i] \triangleq \frac{1}{\sqrt{2}} \begin{pmatrix} x_{\mathrm{I}}^n[i] \\ x_{\mathrm{C}}^n[i] \end{pmatrix}
$$

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

$$\underline{\boldsymbol{y}}_{\mathrm{R}}^n[i] \triangleq \tilde{y}_{\mathrm{R}}^n[i-1]$$
$$\underline{\boldsymbol{y}}_{\mathrm{D}}^n[i] \triangleq \begin{pmatrix} \tilde{y}_{\mathrm{D}}^n[i-1] \\ \bar{y}_{\mathrm{D}}^n[i] \end{pmatrix}$$

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

The input dimension of the matrices $\underline{H}_{\mathrm{R}}$ and $\underline{H}_{\mathrm{D}}$ is equal to two, since the input signal $\underline{\boldsymbol{x}}_{\mathrm{S}}[i]$ consists of two independent codewords. The output dimensions of these matrices reflect the number of observation blocks utilized by each of the receive nodes to recover these codewords.

Scheme (Full-duplex decode-and-forward relaying).

Offline:

- Choose the power allocation coefficient $\rho \in [0, 1]$.
- Construct the effective (augmented) matrices $\underline{H}_{\mathrm{R}}$ and $\underline{H}_{\mathrm{D}}$ of (21).
- Construct the effective MMSE matrices $\underline{G}_{\mathrm{R}}$ and $\underline{G}_{\mathrm{D}}$, of dimensions 3×2 and 4×2, respectively, corresponding to the effective channel matrices $\underline{H}_{\mathrm{R}}$ and $\underline{H}_{\mathrm{D}}$.
- Apply the JET of Theorem 1 to $\underline{G}_{\mathrm{R}}$ and $\underline{G}_{\mathrm{D}}$, to obtain the unitary matrices U_1, U_2, and V, of dimensions 3, 4, and 2, respectively, and the generalized upper-triangular matrices T_1 and T_2 of dimensions 3×2 and 4×2, respectively.
- Since the diagonals of T_1 and T_2 are equal up to a scaling factor, we denote the smaller diagonal pair by $\{t_1, t_2\}$.
- Denote by $\tilde{U}_1$ the 1×2 sub-matrix composed of the first row of U_1, and by $\tilde{U}_2$ — the 2×2 sub-matrix composed of the first two rows of U_2.
- Construct two good codes for scalar AWGN channels of SNRs $\{t_j^2 - 1\}$ and blocklength n.

Source:

- Splits sub-message i, $m[i]$, into two (independent) parts and encodes them into two codewords, each from a different codebook.
- Constructs n vectors $\tilde{\underline{\boldsymbol{x}}}$ of length two each, by taking one sample from each codebook.
- Combines these codewords by multiplying each $\tilde{\underline{\boldsymbol{x}}}$ by the 2×2 unitary matrix V and $\sqrt{P/2}$:
$$\frac{1}{\sqrt{2}} \begin{pmatrix} x_{\mathrm{I}} \\ x_{\mathrm{C}} \end{pmatrix} \triangleq \underline{\boldsymbol{x}}_{\mathrm{S}} = \frac{1}{\sqrt{2}} \sqrt{P} V \tilde{\underline{\boldsymbol{x}}} \,.$$
- At time block i, transmits
$$x_{\mathrm{S}}^n[i] = \rho x_{\mathrm{C}}^n[i] + \sqrt{1-\rho^2} x_{\mathrm{I}}^n[i+1].$$

Relay as Receiver: At time frame $(i-1)$:

- Computes[12]
$$\begin{aligned}\underline{\boldsymbol{y}}_{\mathrm{R}}^n[i] &\triangleq y_{\mathrm{R}}^n[i-1] - h_{\mathrm{R,S}} \rho x_{\mathrm{C}}^n[i-1] \\ &= \sqrt{1-\rho^2} h_{\mathrm{R,S}} x_{\mathrm{I}}^n[i] + z_{\mathrm{R}}^n[i-1] \\ &= \underline{H}_{\mathrm{R}} \underline{\boldsymbol{x}}_{\mathrm{S}}^n[i] + \underline{\boldsymbol{z}}_{\mathrm{R}}^n[i] \,.\end{aligned}$$
- Multiplies $\underline{\boldsymbol{y}}_{\mathrm{R}}^n[i]$ by $\tilde{U}_1^\dagger$.
- Recovers the two codewords using SIC, resulting in signal-to-interference-and-noise ratio (SINR) values of (at least) $\{t_1^2 - 1, t_2^2 - 1\}$.

Relay as Transmitter: At time frame i, sends[13]

$$x_{\mathrm{R}}^n[i] = x_{\mathrm{C}}^n[i] \,.$$

Destination:

- For the recovery of $m[i]$, calculates
$$\underline{\boldsymbol{y}}_{\mathrm{D}}^n[i] \triangleq \begin{pmatrix} y_{\mathrm{D}}^n[i-1] - h_{\mathrm{D,S}} \rho x_{\mathrm{C}}^n[i-1] \\ \frac{1}{\sqrt{P_{\text{equiv}}}} y_{\mathrm{D}}^n[i] \end{pmatrix} \tag{22a}$$
$$= \begin{pmatrix} \sqrt{1-\rho^2} h_{\mathrm{D,S}} x_{\mathrm{I}}^n[i] + z_{\mathrm{D}}^n[i-1] \\ \frac{\rho h_{\mathrm{D,S}} + h_{\mathrm{D,R}}}{\sqrt{(1-\rho^2) h_{\mathrm{D,S}}^2 P + 1}} x_{\mathrm{C}}^n[i] + \bar{z}_{\mathrm{D}}^n[i] \end{pmatrix} \tag{22b}$$
$$= \underline{H}_{\mathrm{D}} \underline{\boldsymbol{x}}_{\mathrm{S}}^n[i] + \underline{\boldsymbol{z}}_{\mathrm{D}}^n[i] \,. \tag{22c}$$
- Multiplies $\underline{\boldsymbol{y}}_{\mathrm{R}}^n[i]$ by $\tilde{U}_1^\dagger$.
- Recovers the two codewords using SIC, resulting in SINR values of (at least) $\{t_1^2 - 1, t_2^2 - 1\}$.

Remark 8. In order to approach the rate R_{DF} of (12) using random Gaussian codes $x_{\mathrm{C}}^n[i]$ and $x_{\mathrm{I}}^n[i]$, one needs these codes to be independent. In the proposed scheme, this will be the case if one takes the entries of $\tilde{\underline{\boldsymbol{x}}}$ from equal-power independent random Gaussian codebooks, since an orthogonal transformation conserves independence in the Gaussian case.

V. Extension to MIMO Relaying

In this section we extend the results of Section IV to the MIMO case.

In the MIMO variant of the relay channel, the SISO links, denoted by 'h' in (1), are replaced by MIMO ones, denoted by 'H'. Thus, (1) is replaced with

$$\begin{aligned}\boldsymbol{y}_{\mathrm{R}} &= H_{\mathrm{D,S}} \boldsymbol{x}_{\mathrm{S}} + \boldsymbol{z}_{\mathrm{R}} \,, \\ \boldsymbol{y}_{\mathrm{D}} &= H_{\mathrm{D,S}} \boldsymbol{x}_{\mathrm{S}} + H_{\mathrm{D,R}} \boldsymbol{x}_{\mathrm{R}} + \boldsymbol{z}_{\mathrm{D}} \,,\end{aligned}$$

where channel input and output dimensions are denoted by N and M, respectively. The Gaussian noise vectors are mutually independent and circularly-symmetric with identity covariance matrices. As in the SISO case, all channel inputs are subject to the same power constraint P and knowledge of all channel matrices at all communication nodes is assumed ("closed loop").

A. Half-Duplex Relaying

In this case, the BC (2) and MAC (3) modes of the SISO half-duplex case are replaced with

$$\begin{aligned}\boldsymbol{y}_{\mathrm{R}} &= H_{\mathrm{R,S}} \boldsymbol{x}_{\mathrm{S}} + \boldsymbol{z}_{\mathrm{R}} \,, \\ \boldsymbol{y}_{\mathrm{D}} &= H_{\mathrm{D,S}} \boldsymbol{x}_{\mathrm{S}} + \boldsymbol{z}_{\mathrm{D}} \,,\end{aligned}$$

and

$$\boldsymbol{y}_{\mathrm{D}} = H_{\mathrm{D,S}} \boldsymbol{x}_{\mathrm{S}} + H_{\mathrm{D,R}} \boldsymbol{x}_{\mathrm{R}} + \boldsymbol{z}_{\mathrm{D}} \,,$$

[12] As for the general FD-DF protocol of Section II-B, we assume correct decoding of previous messages, including $m[i-1]$ (and $x_{\mathrm{C}}^n[i-1]$), as well as known predetermined $m[1]$.

[13] Again, as in the protocol of Section II-B, assuming correct decoding, $m[i]$, and hence also $x_{\mathrm{C}}^n[i]$, are known at time frame i.

respectively.

As in the treatment for the SISO case in Section IV-A, we construct an MCS scheme that approaches capacity via the MCS scheme for MIMO common-message BC of Section III. To that end, we use a MIMO version of the representation (18), repeated here for convenience:

$$\underline{\boldsymbol{y}}_{\mathrm{R}} = \underline{H}_1 \underline{\boldsymbol{x}} + \underline{\boldsymbol{z}}_{\mathrm{R}} \,, \tag{23a}$$
$$\underline{\boldsymbol{y}}_{\mathrm{D}} = \underline{H}_2 \underline{\boldsymbol{x}} + \underline{\boldsymbol{z}}_{\mathrm{D}} \,. \tag{23b}$$

The power allocation coefficients of the SISO case ($\sqrt{P_1}$, $\sqrt{P_2}$ and $\sqrt{P_{\mathrm{R}}}$), are replaced by beamforming matrices (B_1, B_2 and B_{R}, respectively). In contrast to the SISO case, alignment of the transmission signals during the MAC phase, carried in order to achieve full coherence, is non-trivial, and one has to optimize over the beamforming matrices.[14]

We assume again, as in the SISO case, that for the chosen beamforming matrices B_1, B_2 and B_{R}: $n_1 = m_1 n$ and $n_2 = m_2 n$, for some positive integers m_1, m_2 and n, where n_1 and n_2 are the blocklengths of the two transmission phases as defined in Section II-A.

Thus, the equivalent augmented channel matrices of (23), in the MIMO case, take the form

$$\begin{aligned}\underline{H}_1 &= \underline{H}_{\mathrm{R,S}} \underline{B}_{\mathrm{S}} \\ \underline{H}_2 &= \underline{H}_{\mathrm{D,S}} \underline{B}_{\mathrm{S}} + \underline{H}_{\mathrm{D,R}} \underline{B}_{\mathrm{R}} \,,\end{aligned}$$

where

$$\begin{aligned}\underline{H}_{\mathrm{R,S}} &= \begin{pmatrix} I_{m_1} \otimes H_{\mathrm{R,S}} & 0_{M_{\mathrm{R}} m_1 \times N_{\mathrm{S}}(m_2 - m_1)} \end{pmatrix} \,, \\ \underline{H}_{\mathrm{D,S}} &= I_{m_2} \otimes H_{\mathrm{D,S}} \,, \\ \underline{H}_{\mathrm{D,R}} &= I_{m_2} \otimes H_{\mathrm{D,R}} \,;\end{aligned}$$

and

$$\underline{B}_{\mathrm{S}} = \begin{pmatrix} I_{m_1} \otimes B_1 & 0_{N_{\mathrm{S}} m_1 \times N_{\mathrm{S}}(m_2 - m_1)} \\ 0_{N_{\mathrm{S}}(m_2 - m_1) \times N_{\mathrm{S}} m_1} & I_{(m_2 - m_1)} \otimes B_2 \end{pmatrix} \,, \tag{24a}$$
$$\underline{B}_{\mathrm{R}} = \begin{pmatrix} & 0_{N_{\mathrm{S}} m_1 \times N_{\mathrm{S}} m_2} \\ 0_{N_{\mathrm{S}}(m_2 - m_1) \times N_{\mathrm{S}} m_1} & I_{(m_2 - m_1)} \otimes B_{\mathrm{R}} \end{pmatrix} \,, \tag{24b}$$

where $\otimes$ denotes the Kronecker product operation (see, e.g., [36, Ch. 4]), i.e., $I_m \otimes H$ is a block diagonal matrix with m blocks that are all equal to H.

For these matrices, according to Theorem 1, an MCS scheme achieving a rate of

$$R = \min_{i=1,2} \log \left| I + \underline{H}_i \underline{H}_i^{\dagger} \right|$$

can be devised, by a simple adaptation of the scheme for the SISO case of Section IV-B.

B. Full-Duplex Relaying

As in the HD setting, the DF protocol and scheme used for the FD DF variant for the SISO case in Sections II-B and IV-B, can be readily extended to the MIMO case. This calls for replacing the scalar codebooks x_{I}^n and x_{C}^n with vector codebooks whose entries are vectors of the same dimension as the transmitted signal $\boldsymbol{x}_{\mathrm{S}}$.

[14]The optimization of B_2 and B_{R} can be simplified by noting that multiplying both matrices by the same unitary matrix on the right, results in no change in performance.

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

Note also that in this case the beamforming matrices, corresponding to $\boldsymbol{x}_{\mathrm{I}}$ and $\boldsymbol{x}_{\mathrm{C}}$ at the source and to $\boldsymbol{x}_{\mathrm{C}}$ — at the relay, can be shaped to improve the achievable rate, such that each of these covariance matrices satisfies the power constraints. More generally, we take $\boldsymbol{x}_{\mathrm{I}}$ and $\boldsymbol{x}_{\mathrm{C}}$ to be independent and white of total average unit power and multiply them by suitable precoding matrices, subject to an average power constraint P. Hence, the signals sent by the source (5) and the relay (7) should be replaced, in the MIMO case, with

$$\begin{aligned}\boldsymbol{x}_{\mathrm{S}}^n[i] &= \rho B_{\mathrm{S}}^{(\mathrm{C})} \boldsymbol{x}_{\mathrm{C}}^n[i] + \sqrt{1-\rho^2} B_{\mathrm{S}}^{(\mathrm{I})} \boldsymbol{x}_{\mathrm{I}}[i+1] \,, \\ \boldsymbol{x}_{\mathrm{R}}^n[i] &= B_{\mathrm{R}}^{(\mathrm{C})} \boldsymbol{x}_{\mathrm{C}}^n[i] \,,\end{aligned}$$

respectively, where $B_{\mathrm{S}}^{(\mathrm{C})}$, $B_{\mathrm{S}}^{(\mathrm{I})}$ and $B_{\mathrm{R}}^{(\mathrm{C})}$ are the precoding matrices satisfying the power constraints:

$$\begin{aligned}\operatorname{trace}\left\{ B_{\mathrm{S}}^{(\mathrm{C})} B_{\mathrm{S}}^{(\mathrm{C})\dagger} \right\} &\le P \,, \\ \operatorname{trace}\left\{ B_{\mathrm{S}}^{(\mathrm{I})} B_{\mathrm{S}}^{(\mathrm{I})\dagger} \right\} &\le P \,, \\ \operatorname{trace}\left\{ B_{\mathrm{R}}^{(\mathrm{C})} B_{\mathrm{R}}^{(\mathrm{C})\dagger} \right\} &\le P \,.\end{aligned}$$

Again, these precoding matrices should be chosen to maximize the achievable rate, but their exact choice is not material for the construction of the protocol and the corresponding MCS scheme.

Remark 9. In the SISO case, the signal sent by the relay (7) and the corresponding component x_{C}^n in the signal of (5) 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 Sections II and IV, which were assumed to be real and positive); 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 x_{C}^n (the MIMO equivalent of (6), assuming correct decoding), is equal to

$$\begin{aligned}\tilde{\boldsymbol{y}}_{\mathrm{R}}^n[i-1] &= \boldsymbol{y}_{\mathrm{R}}^n[i-1] - \rho H_{\mathrm{R,S}} B_{\mathrm{S}}^{(\mathrm{C})} \boldsymbol{x}_{\mathrm{C}}^n[i-1] \\ &= \sqrt{1-\rho^2} H_{\mathrm{R,S}} B_{\mathrm{S}}^{(\mathrm{I})} \boldsymbol{x}_{\mathrm{I}}^n[i] + \boldsymbol{z}_{\mathrm{R}}^n[i-1] \,.\end{aligned}$$

At the destination, (22) is replaced by

$$\boldsymbol{y}_{\mathrm{D}}^n[i-1] = \sqrt{1-\rho^2} H_{\mathrm{D,S}} B_{\mathrm{S}}^{(\mathrm{I})} \boldsymbol{x}_{\mathrm{I}}^n[i] + \boldsymbol{z}_{\mathrm{D}}^n[i-1] \,,$$

whereas (9) and (10) are replaced by

$$\begin{aligned}\boldsymbol{y}_{\mathrm{D}}^n[i] &= \left(\rho H_{\mathrm{D,S}} B_{\mathrm{S}}^{(\mathrm{C})} + H_{\mathrm{D,R}} B_{\mathrm{R}}^{(\mathrm{C})} \right) \boldsymbol{x}_{\mathrm{C}}^n[i] + \boldsymbol{z}_{\mathrm{equiv}}^n[i] \,, \\ \boldsymbol{z}_{\mathrm{equiv}}^n[i] &= \sqrt{1-\rho^2} H_{\mathrm{D,S}} B_{\mathrm{S}}^{(\mathrm{I})} \boldsymbol{x}_{\mathrm{I}}^n[i] + \boldsymbol{z}_{\mathrm{D}}^n[i] \,,\end{aligned}$$

respectively. The covariance matrix of the Gaussian noise is

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

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

$$K_{\text{equiv}} = L_{\text{equiv}} L_{\text{equiv}}^{\dagger} \, ,$$

where L is invertible. Hence, by applying L_{equiv}^{-1} to $\boldsymbol{y}_{\text{D}}^{n}[i]$ on the right, we arrive at

$$\bar{\boldsymbol{y}}_{\text{D}}^{n}[i] = L_{\text{equiv}}^{-1} \left(\rho H_{\text{D,S}} B_{\text{S}}^{(\text{C})} + H_{\text{D,R}} B_{\text{R}}^{(\text{C})} \right) \boldsymbol{x}_{\text{C}}^{n}[i] + \bar{\boldsymbol{z}}_{\text{D}}^{n}[i] \, ,$$

where $\bar{\boldsymbol{z}}_{\text{D}}^{n}[i] = L_{\text{equiv}}^{-1} \boldsymbol{z}_{\text{equiv}}^{n}[i]$ has an identity covariance matrix.

The effective matrices (21) of the scheme of Section IV-A, need to be replaced, in the MIMO case, by the following channel matrices:

$$\underline{H}_{\text{R}} = \sqrt{2} \begin{pmatrix} \sqrt{1-\rho^2} H_{\text{R,S}} B_{\text{S}}^{(\text{I})} & 0 \end{pmatrix}$$

$$\underline{H}_{\text{D}} = \sqrt{2} \begin{pmatrix} \sqrt{1-\rho^2} H_{\text{D,S}} B_{\text{S}}^{(\text{I})} & 0 \\ 0 & L_{\text{equiv}}^{-1} \left[\rho H_{\text{D,S}} B_{\text{S}}^{(\text{C})} + H_{\text{D,R}} B_{\text{R}}^{(\text{C})} \right] \end{pmatrix} .$$

$$\begin{aligned} R_{\text{DF}} = \min \Big\{ & C\left(H_{\text{R,S}} B_{\text{S}}^{(\text{I})}, \left[1-\rho^2\right] I \right) , \\ & C\left(H_{\text{D,S}} B_{\text{S}}^{(\text{I})}, \left[1-\rho^2\right] I \right) \\ & + C\left(L_{\text{equiv}}^{-1} \left[\rho H_{\text{D,S}} B_{\text{S}}^{(\text{C})} + H_{\text{D,R}} B_{\text{R}}^{(\text{C})} \right], I \right) \Big\} , \end{aligned}$$

where $C(\cdot,\cdot)$ was defined in (14).

VI. Application to MIMO Broadcast Rateless Coding

In this section, we treat the problem of designing practical incremental redundancy codes over the AWGN channel, both for the SISO and MIMO cases, by constructing an MCS scheme. As discussed in Sections I and II-A, this problem is closely related to the half-duplex relaying problem. In fact, careful scrutiny reveals that it is actually a special case of the half-duplex relay problem and therefore the MCS schemes for the SISO and MIMO variants of Sections IV-A and V-A, respectively, may be readily applied.

We start by providing a formal definition of the problem, following Shulman [37].

A. Problem Setting

Consider the MIMO broadcast channel of Section III, and denote the point-to-point capacity of user i by

$$C_i \triangleq C(H_i, K_i) \, ,$$

where K_i is the covariance matrix maximizing (14) for $H = H_i$ under a power constraint trace$\{K_i\} \leq P$.

In a two-user rateless setting, the transmitter needs to send the same k bits to both receivers. Each receiver "listens" to the transmission from time instant 1 until it is able to reliably decode all bits; then it may tune out. The *online time* of user i, denoted by n_i, is the number of channel uses that is required for reliable recovery of the information, and the resulting effective rates are defined as

$$R_i \triangleq \left\lfloor \frac{k}{n_i} \right\rfloor , \quad i = 1, 2 \, , \tag{25}$$

where $\lfloor \cdot \rfloor$ denotes the "floor" operation. The following, due to Shulman [37], states the optimal rates.

Proposition 1. *The effective rate pair (R_1, R_2) is achievable under power constraint P if and only if there exists a covariance matrix K with trace$\{K\} \leq P$, such that*

$$\frac{C(H_i, K)}{R_i} + C_i \left[\frac{1}{R_i} - \frac{1}{R_{\bar{i}}} \right]^{+} \geq 1 \, , \qquad i = 1, 2 \, ,$$

where $[a]^{+} \triangleq \max\{a, 0\}$ and

$$\bar{i} = \begin{cases} 2 & i = 1 \\ 1 & i = 2 \end{cases} .$$

This result can be understood as follows. Without loss of generality, assume that $C(H_1, K) \geq C(H_2, K)$. For the first $n_1 = k / C(H_1, K)$ channel uses, the transmitter uses the covariance matrix K, at the end of which the first user obtains enough mutual information to decode the message, and may tune out. Once only the second user is online, the transmitter switches to the best-matched covariance matrix for its channel, K_2. With this,

$$n_2 - n_1 = \frac{k - n_1 C(H_2, K)}{C_2}$$

additional uses are needed until the second user obtains enough mutual information, as well. We see, then, that the only compromise is in the choice of covariance matrix for the first period; given this choice, each receiver is able to use all the mutual information provided by the channel, as if it were a point-to-point scenario. Unfortunately, this information-theoretic result does not tell us how to achieve these rates using *practical* codes; in the sequel, we construct MCS schemes for this problem, by reducing it to that of a ("classical") MIMO common-message BC one (as in Section III).

Remark 10. Proposition 1 generalizes to more than two users in a straightforward manner; see [37].

Remark 11. A scheme for the case of equal channel matrices up to a scalar constant, which is equivalent to the scenario of a single channel matrix but unknown SNR, can be obtained relying on the geometric mean decomposition (GMD). For further details see [38].

B. Two-User Rateless MCS Scheme via Half-Duplex Relaying

In order to design an MCS scheme for the two-user case, we observe that (for a specific choice of K in Proposition 1) this problem can be regarded as a special case of the half-duplex (or "*rateless relay*") problem of Sections II-A and IV-A, where the relay "tunes in" until it is able to decode, but does not transmit any signal (corresponds to assigning $B_2 = 0$ in (24b)).

For completeness, we next present the MIMO incremental redundancy coding scheme, which is first reformulated as a "classical" (non incremental-redundancy) common-message BC problem, using a matrix representation, as in the half-duplex scheme of Section IV-A (being a special case of the latter).

Again, we restrict attention to rational ratios between the effective rates, i.e.,

$$n_1 = m_1 n \tag{26a}$$

$$n_2 = m_2 n \tag{26b}$$

for some positive integers m_1, m_2 and n. We view the n_i $(i = 1, 2)$ channel uses as n uses of the equivalent channels,

$$\underline{\boldsymbol{y}}_1 = \underline{H}_1 \underline{\boldsymbol{x}}_{\mathrm{S}} + \underline{\boldsymbol{z}}_1 \,,$$
$$\underline{\boldsymbol{y}}_2 = \underline{H}_2 \underline{\boldsymbol{x}}_{\mathrm{S}} + \underline{\boldsymbol{z}}_2 \,,$$

represented by the block-diagonal matrices

$$\underline{H}_1 = \underline{H}_1' \underline{B} \,, \tag{27a}$$

$$\underline{H}_2 = \underline{H}_2' \underline{B} \,, \tag{27b}$$

of dimensions $m_1 M_1 \times m_2 N_{\mathrm{S}}$ and $m_2 M_2 \times m_2 N_{\mathrm{S}}$, respectively, where

$$\underline{H}_1' = \begin{pmatrix} I_{m_1} \otimes H_1 & 0_{m_1 M_1 \times (m_2 - m_1) N_{\mathrm{S}}} \end{pmatrix}$$
$$\underline{H}_2' = I_{m_2} \otimes H_2 \,,$$

and[15]

$$\underline{B} = \begin{pmatrix} I_{m_1} \otimes B & 0_{m_1 N_{\mathrm{S}} \times (m_2 - m_1) N_{\mathrm{S}}} \\ 0_{(m_2 - m_1) N_{\mathrm{S}} \otimes m_1 N_{\mathrm{S}}} & I_{m_2 - m_1} \otimes B_2 \end{pmatrix},$$

where B and B_2 are (any) beamforming matrices satisfying $BB^\dagger = K$ and $B_2 B_2^\dagger = K_2$.[16] Define $\underline{K} = \underline{B}\underline{B}^\dagger$.

Note that using (25) and (26), the effective rates of these equivalent channels are *equal*, i.e.,

$$R \triangleq C\left(\underline{H}_1', \underline{K}\right) = C\left(\underline{H}_2', \underline{K}\right) \,, \tag{28}$$

or alternatively (see Remark 5),

$$R \triangleq C\left(\underline{H}_1, I\right) = C\left(\underline{H}_2, I\right) \,.$$

Consequently, we can apply the MIMO common-message BC scheme of Theorem 1 to the matrices $\underline{H}_1'$ and $\underline{H}_2'$ and input covariance matrix $\underline{K}$, or alternatively to the equivalent channel matrices $\underline{H}_1 \triangleq \underline{H}_1' \underline{B}$ and $\underline{H}_2 \triangleq \underline{H}_2' \underline{B}$ with unit covariance matrix, such that the optimum rates, as given in Proposition 1, are achieved.

Scheme (Two-user rateless coding).

Offline:

- Choose admissible beamforming matrices B and B_2.
- Construct the effective (augmented) channel matrices $\underline{H}_1$ and $\underline{H}_2$ (27).
- Construct the effective MMSE matrices $\underline{G}_1$ and $\underline{G}_2$, of dimensions $(m_1 M_1 + m_2 N_{\mathrm{S}}) \times m_2 N_{\mathrm{S}}$ and $m_2 (M_2 + N_{\mathrm{S}}) \times m_2 N_{\mathrm{S}}$, respectively, corresponding to the effective channel matrices $\underline{H}_1$ and $\underline{H}_2$.
- Apply the JET of Theorem 1 to $\underline{G}_1$ and $\underline{G}_2$, to obtain the unitary matrices U_1, U_2, and V, of dimensions $(m_1 M_1 + m_2 N_{\mathrm{S}})$, $m_2 (M_2 + N_{\mathrm{S}})$, and $m_2 N_{\mathrm{S}}$, respectively, and the generalized upper-triangular matrices T_1 and T_2 of dimensions $(m_1 M_1 + m_2 N_{\mathrm{S}}) \times m_2 N_{\mathrm{S}}$ and $m_2 (M_2 + N_{\mathrm{S}}) \times m_2 N_{\mathrm{S}}$, respectively.
- Denote the $m_2 N_{\mathrm{S}}$ diagonal elements of T_1 and T_2 by $\{t_j\}$ (which are equal for both matrices).
- Denote by $\tilde{U}_k$ $(k = 1, 2)$ the $m_k M_k \times m_2 N_{\mathrm{S}}$ sub-matrix composed of the first $m_k M_k$ rows of U_k.
- Construct good codes for scalar AWGN channels of SNRs $\{t_j^2 - 1\}$ and blocklength n.

Transmitter:

- Constructs n vectors $\underline{\boldsymbol{x}}$ of length $m_2 N_{\mathrm{S}}$ using one sample from each codebook.
- Combines all these codewords by multiplying each $\tilde{\underline{\boldsymbol{x}}}$ by the unitary matrix V and by $\underline{B}$, both of dimensions $m_2 N_{\mathrm{S}} \times m_2 N_{\mathrm{S}}$:

$$\underline{\boldsymbol{x}}_{\mathrm{S}} = \underline{B} V \tilde{\underline{\boldsymbol{x}}} \,.$$

- Transmits the first $m_1 M_1$ entries of each of the n vectors $\underline{\boldsymbol{x}}_{\mathrm{S}}$ over m_1 *physical* channel inputs $\boldsymbol{x}_{\mathrm{S}}$.
- Transmits the rest of the entries of each of the n vectors $\underline{\boldsymbol{x}}_{\mathrm{S}}$ over $(m_2 - m_1)$ physical channel inputs $\boldsymbol{x}_{\mathrm{S}}$.

Receiver 1:

- After receiving the first $n_1 = m_1 n$ output vectors $\boldsymbol{y}_{\mathrm{D}}$, constructs the n output vectors $\underline{\boldsymbol{y}}_1$.[17]
- Multiplies each vector $\underline{\boldsymbol{y}}_1$ by $\tilde{U}_1^\dagger$.
- Recovers the codewords using SIC, resulting in SINR values of $\{t_j^2 - 1\}$.

Receiver 2:

- Constructs the n output vectors $\underline{\boldsymbol{y}}_2$ out of the $n_2 = m_2 n$ received vectors $\boldsymbol{y}_{\mathrm{D}}$.
- Multiplies each vector $\underline{\boldsymbol{y}}_2$ by $\tilde{U}_2^\dagger$.
- Recovers the codewords using SIC, resulting in SINR values of $\{t_j^2 - 1\}$.

Using capacity-achieving scalar AWGN codes, i.e., codes of rates close to

$$\left\{ R_j \middle| R_j = \log\left(t_j^2\right) \,, j \in \{1, \ldots, m_2 N_{\mathrm{S}}\} \right\} \,,$$

the desired rate of (28) is achieved.

Remark 12. A similar scheme can be constructed for varying (*known*) blocks on the diagonal of $\underline{H}_2'$. In fact, such a variation was already used in the scheme of Section IV-A for the half-duplex relay problem. In the case of different blocks on the diagonal, however, the blocks on the diagonal of $\underline{B}$ will vary as well. Yet, even in this more general case, fixed-rate scalar AWGN codebooks suffice for achieving the optimum.

[15] Mind the different meaning of the indices of the beamforming matrices, from that of (24), in this case.

[16] In contrast to the half-duplex transmission scheme, in this case, beamforming matrices which give rise to the same covariance matrix will result in the same performance.

[17] As explained in Section IV-A, the other columns of $U_1^\dagger$ are multiplied by zeros because of the zero columns in $\underline{H}_1$. Hence, the first $m_1 n$ physical output vectors $\boldsymbol{y}_{\mathrm{D}}$ suffice to reconstruct the "interesting" part of $\underline{\boldsymbol{y}}_1$ at receiver 1.

C. SISO Rateless Coding

We now show that the scheme proposed in this section may be helpful even outside the MIMO setting. Consider the case of a SISO channel, i.e., the channel matrices in (13) reduce to scalars. This problem was considered in [16] and is described by the channel

$$y = \alpha x + z\,,$$

where α takes one of the values $\{\alpha_1, \alpha_2, \ldots\}$, where

$$i \log(1 + |\alpha_i|^2) = R.$$

That is, the values $\{\alpha_i | \, i = 1, \ldots, M\}$ are chosen such that the ratios between the transmission durations lengths $\{n_i | \, i = 1, \ldots, M\}$ required to recover the same message, satisfy $n_i = in$, for $i = 1, \ldots, M$ and (any) integer n, where the latter will serve as the basic blocklength of the schemes to be described next.

A practical scheme based on linear pre- and post-processing and SIC was proposed in [16] for any number M of possible α_i values, and it was shown that a perfect solution (i.e. capacity-achieving) exists for $M = 2$ for any R, or for $M = 3$ up to some critical value of R.[18]

The case $M = 2$ falls under the category of channels addressed in Section VI-B, with the augmented matrices being

$$\underline{H}_1 = \begin{pmatrix} \alpha_1 & 0 \end{pmatrix},$$
$$\underline{H}_2 = \begin{pmatrix} \alpha_2 & 0 \\ 0 & \alpha_2 \end{pmatrix}.$$

Interestingly, the explicit computation of the matrices U_1, U_2 and V of that scheme exactly coincides with the matrices derived in [16]. However, our approach provides a straightforward extension to other cases with either two channel uses or more, as long as the channel materialization is known to be one of two options. Specifically, let $\alpha_{1,1}, \ldots, \alpha_{1,M}$ and $\alpha_{2,1}, \ldots, \alpha_{2,M}$ be the sequences of SNRs according to both options, such that

$$\sum_{i=1}^{M} \log(1 + \alpha_{1,i}) = \sum_{i=1}^{M} \log(1 + \alpha_{2,i}) = R.$$

Then, our scheme achieves exactly the optimum rate R. This holds also in the specific case where the trailing SNRs of one of the sequences are zero, and hence applies to the two-option rateless coding problem. For the case of $M = 3$ channel options (α values), the critical rate can be derived via necessary and sufficient conditions for the possibility to attain JET for three matrices [40]. For a discussion for more than three channel options ($M > 3$), see Section VII.

VII. Concluding Remarks

The schemes derived in this work are applicable for any choice of input covariance matrix K, achieving the corresponding mutual information. Thus, the average power input constraints assumed in this work can be replaced with other input constraints, e.g., covariance constraints or individual power constraints. Of course, this requires to carry out the optimization problem of finding the optimal covariance matrix K for the chosen constraints. The resulting covariance matrix K can then be used in all of the proposed schemes.

Note that in the proposed schemes, the SISO codes used are of different rates, which may be somewhat undesirable. A seemingly different problem is extending the schemes of Section IV to the case of more relays (possibly without a direct link). For this to be possible, the MIMO common-message BC scheme of Section III 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, at the expense of greater latency at the output, as explained in [25].

We finally note that networks containing more than one source node as well as various secrecy scenarios can be treated in a similar fashion as well, as demonstrated in [41], [42].

[18] A numerical derivation of the precoding matrix V in the case of a rateless code (even for parameters for which a perfect decomposition is not possible) is available in [39].

Acknowledgments

The authors thank Or Ordentlich for helping with the analysis of the full-duplex relay channel scenario of Section IV-B.

References

[1] E. C. van der Meulen, "Transmission of information in a T-terminal discrete memoryless channel," Ph.D. dissertation, University of California, Berkeley, CA, USA, June 1968.
[2] ——, "Three-terminal communication channels," *Adv. Appl. Prob.*, vol. 3, pp. 120–54, 1971.
[3] T. M. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Info. Theory*, vol. 25, no. 5, pp. 572–584, Sep. 1979.
[4] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Info. Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.
[5] P. Mitran, H. Ochiai, and V. Tarokh, "Space–time diversity enhancements using collaborative communications," *IEEE Trans. Info. Theory*, vol. 51, no. 6, pp. 2041–2057, June 2005.
[6] J. Castura and Y. Mao, "Rateless coding and relay networks," *IEEE Sig. Proc. Magazine*, vol. 24, no. 5, pp. 27–35, Sep. 2007.
[7] E. Telatar, "Capacity of the multiple antenna Gaussian channel," *Europ. Trans. Telecommun.*, vol. 10, pp. 585–595, Nov. 1999.
[8] G. Foschini, "Layered space–time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Sys. Tech. Jour.*, vol. 1, no. 2, pp. 41–59, 1996.
[9] P. Razaghi and W. Yu, "Bilayer low-density parity-check codes for decode-and-forward in relay channels," *IEEE Trans. Info. Theory*, vol. 53, no. 10, pp. 3723–3739, Oct. 2007.
[10] M. H. Azmi and J. Yuan, "Performance of bilayer-lengthened LDPC codes under joint decoding," in *Proc. IEEE Int. Symp. on Info. Theory (ISIT)*, Taormina, Sicily, Italy, Oct. 2009, pp. 163–167.
[11] 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," *IEEE Trans. Comm.*, vol. 59, no. 11, pp. 2993–3006, Nov. 2011.
[12] M. A. Khojastepour, N. Ahmed, and B. Aazhang, "Code design for the relay channel and factor graph decoding," in *Proc. Asilomar Conf. on Sig., Sys. and Comp.*, vol. 2, Pacific Grove, CA, USA, Nov. 2004, pp. 2000–2004.
[13] Z. Zhang and T. M. Duman, "Capacity-approaching turbo coding and iterative decoding for relay channels," *IEEE Trans. Comm.*, vol. 53, no. 11, pp. 1895–1905, Nov. 2005.
[14] J. Ezri and M. Gastpar, "On the performance of independently designed LDPC codes for the relay channel," in *Proc. IEEE Int. Symp. on Info. Theory (ISIT)*, Seattle, WA, USA, July 2006, pp. 977–981.
[15] Y. Song and N. Devroye, "Lattice codes for the Gaussian relay channel: Decode-and-forward and compress-and-forward," *IEEE Trans. Info. Theory*, vol. 59, no. 8, pp. 4927–4948, Aug. 2013.

[16] U. Erez, M. D. Trott, and G. W. Wornell, “Rateless coding for Gaussian channels,” *IEEE Trans. Info. Theory*, vol. 58, no. 2, pp. 530–547, Feb. 2012.

[17] R. Palanki and J. S. Yedidia, “Rateless codes on noisy channels,” in *Proc. IEEE Int. Symp. on Info. Theory (ISIT)*, Chicago, IL, USA, June 2004, p. 37.

[18] R. Liu, P. Spasojevic, and E. Soljanin, “Incremental redundancy cooperative coding for wireless networks: Cooperative diversity, coding, and transmission energy gains,” *IEEE Trans. Info. Theory*, vol. 54, no. 3, pp. 1207–1224, Mar. 2008.

[19] J. Castura and Y. Mao, “Rateless coding for wireless relay channels,” *IEEE Trans. Wireless Comm.*, vol. 6, no. 5, pp. 1638–1642, May 2007.

[20] A. Chakrabarti, A. de Baynast, A. Sabharwal, and B. Aazhang, “Low density parity check codes for the relay channel,” *IEEE J. Selected Areas in Comm.*, vol. 25, no. 2, pp. 280–291, Feb. 2007.

[21] J. Hu and T. M. Duman, “Low density parity check codes over wireless relay channels,” *IEEE Trans. Wireless Comm.*, vol. 6, no. 9, pp. 3384–3394, Sep. 2007.

[22] J. P. Cances and V. Meghdadi, “Optimized low density parity check codes designs for half duplex relay channels,” *IEEE Trans. Wireless Comm.*, vol. 8, no. 7, pp. 3390–3395, July 2009.

[23] T. Van Nguyen, A. Nosratinia, and D. Divsalar, “Bilayer protograph codes for half-duplex relay channels,” *IEEE Trans. Wireless Comm.*, vol. 12, no. 5, pp. 1969–1977, May 2011.

[24] A. Khina, Y. Kochman, and U. Erez, “Joint unitary triangularization for MIMO networks,” *IEEE Trans. Sig. Proc.*, vol. 60, no. 1, pp. 326–336, Jan. 2012.

[25] A. Khina, I. Livni, A. Hitron, and U. Erez, “Joint unitary triangularization for Gaussian multi-user MIMO networks,” *IEEE Trans. Info. Theory*, vol. 61, no. 5, pp. 2662–2692, May 2015.

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

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

[28] R. L. Dobrushin, “Optimal information transmission over a channel with unknown parameters,” (in Russian) *Radiotekh. i Elektron.*, vol. 4, no. 12, pp. 1951–1956, Dec. 1959.

[29] J. Wolfowitz, “Simultaneous channels,” *Arch. Rational Mech. Anal.*, vol. 4, pp. 371–386, 1960.

[30] J. M. Cioffi, G. P. Dudevoir, M. V. Eyuboglu, and G. D. Forney Jr., “MMSE decision-feedback equalizers and coding — Part I: Equalization results,” *IEEE Trans. Comm.*, vol. 43, no. 10, pp. 2582–2594, Oct. 1995.

[31] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, “V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel,” in *Proc. URSI Int. Symp. Sig., Sys., Elect. (ISSSE)*, Sep./Oct. 1998, pp. 295–300.

[32] J. M. Cioffi, G. P. Dudevoir, M. V. Eyuboglu, and G. D. Forney Jr., “MMSE decision-feedback equalizers and coding — Part II: Coding results,” *IEEE Trans. Comm.*, vol. 43, no. 10, pp. 2595–2594, Oct. 1995.

[33] B. Hassibi, “An efficient square-root algorithm for BLAST,” in *Proc. IEEE Int. Conf. on Acoust. Speech and Sig. Proc. (ICASSP)*, vol. 2, Istanbul, Turkey, June 2000, pp. 737–740.

[34] H. El Gamal, G. Caire, and M. O. Damen, “Lattice coding and decoding achieve the optimal diversity–multiplexing tradeoff of MIMO channels,” *IEEE Trans. Info. Theory*, vol. 50, no. 6, pp. 968–985, June 2004.

[35] Y. Jiang, W. Hager, and J. Li, “Uniform channel decomposition for MIMO communications,” *IEEE Trans. Sig. Proc.*, vol. 53, no. 11, pp. 4283–4294, Nov. 2005.

[36] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. Cambridge: Cambridge University Press, 1991.

[37] N. Shulman, “Universal channel coding,” Ph.D. dissertation, Tel Aviv University, Tel Aviv, Israel, 2004.

[38] A. Khina, Y. Kochman, U. Erez, and G. W. Wornell, “Incremental coding over MIMO channels,” in *Proc. IEEE Info. Theory Workshop (ITW)*, Paraty, Brazil, Oct. 2011, pp. 400–404.

[39] M. D. Trott, U. Erez, and G. W. Wornell, “Rateless coding for Gaussian channels: Matlab script,” Feb. 2012. [Online]. Available: www.eng.tau.ac.il/∼uri/scripts/rateless_matlab_script.zip

[40] A. Hitron, “Linear precoding and coding for Gaussian channels,” Master’s thesis, Tel Aviv University, Sep. 2012. [Online]. Available: http://www.eng.tau.ac.il/∼uri/theses/hitron_msc.pdf

[41] A. Khina, Y. Kochman, and U. Erez, “Physical-layer MIMO relaying,” in *Proc. IEEE Int. Symp. on Info. Theory (ISIT)*, St. Petersburg, Russia, July/Aug. 2011, pp. 2437–2441.

[42] A. Khina, Y. Kochman, and A. Khisti, “Decomposing the MIMO wiretap channel,” in *Proc. IEEE Int. Symp. on Info. Theory (ISIT)*, Honolulu, HI, USA, June/July 2014, pp. 206–210.

Anatoly Khina (S’08) was born in Moscow, USSR, on September 10, 1984. He received the B.Sc. and M.Sc. degrees (both *summa cum laude*), and completed his Ph.D. studies in 2006, 2010 and 2015, respectively, all in electrical engineering at Tel Aviv University. His research interests include information theory, signal processing, digital communications and matrix analysis. He is currently a Postdoctoral Scholar in the Department of Electrical Engineering at the California Institute of Technology, Pasadena, CA, USA.

In parallel to his studies, Anatoly has been working as an engineer in various algorithms, software and hardware R&D positions. He is a recipient of the Fulbright and Rothschild Postdoctoral Fellowships, Clore Scholarship, Trotsky Award, Weinstein Prize for research in signal processing, and the first prize for outstanding research work in the field of communication technologies of the Advanced Communication Center (ACC) Feder Family Award.

Yuval Kochman (S06–M09) received his B.Sc. (*cum laude*), M.Sc. (*cum laude*) and Ph.D. degrees from Tel Aviv University in 1993, 2003 and 2010, respectively, all in electrical engineering. During 2009–2011, he was a Postdoctoral Associate at the Signals, Informtion and Algorithms Laboratory at the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA. Since 2012, he has been with the School of Computer Science and Engineering at the Hebrew University of Jerusalem.

Outside academia, he has worked in the areas of radar and digital communications. His research interests include information theory, communications and signal processing.

Uri Erez (M’09) was born in Tel-Aviv, Israel, on October 27, 1971. He received the B.Sc. degree in mathematics and physics and the M.Sc. and Ph.D. degrees in electrical engineering from Tel-Aviv University in 1996, 1999, and 2003, respectively. During 2003–2004, he was a Postdoctoral Associate at the Signals, Information and Algorithms Laboratory at the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA. Since 2005, he has been with the Department of Electrical Engineering–Systems at Tel-Aviv University.

His research interests are in the general areas of information theory and digital communications. He served in the years 2009–2011 as Associate Editor for Coding Techniques for the IEEE TRANSACTIONS ON INFORMATION THEORY.

Gregory W. Wornell (S83–M91–SM00–F04) Received the B.A.Sc. degree in electrical engineering from the University of British Columbia, Vancouver, BC, Canada, and the S.M. and Ph.D. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, in 1985, 1987, and 1991, respectively.

Since 1991, he has been on the faculty at MIT, where he is the Sumitomo Professor of Engineering in the department of Electrical Engineering and Computer Science (EECS). He leads the Signals, Information, and Algorithms Laboratory in the Research Laboratory of Electronics,an dco-chairs the EECS department graduate program. He has held visiting appointments at the former AT&T Bell Laboratories, Murray Hill, NJ, USA, the University of California, Berkeley, CA, USA, and Hewlett-Packard Laboratories, Palo Alto, CA, USA. His research interests and publications span the areas of signal processing, digital communication, and information theory, and include algorithms and architectures for wireless networks, sensing and imaging systems, digitally enhanced analog circuits and systems, multimedia applications, and aspects of computational biology and neuroscience.

Dr. Wornell has been involved in the Information Theory and Signal Processing Societies of the IEEE in a variety of capacities, and maintains a number of close industrial relationships and activities. He has won a number of awards for both his research and teaching.