Joint Unitary Triangularization for MIMO Networks

Anatoly Khina
Dept. of EE-Systems,
Tel-Aviv University,
Tel Aviv , Israel.
Email: anatolyk@eng.tau.ac.il

Yuval Kochman*
EECS Dept., MIT
Cambridge, MA
Email: yuvalko@mit.edu

Uri Erez
Dept. of EE-Systems,
Tel-Aviv University,
Tel Aviv , Israel.
Email: uri@eng.tau.ac.il

Abstract—**This work considers communication networks where individual links can be described as MIMO channels. Unlike orthogonal modulation methods (such as the singular-value decomposition), we allow interference between sub-channels, which can be removed by the receivers via successive cancellation. The degrees of freedom earned by this relaxation are used for obtaining a basis which is simultaneously good for more than one link. Specifically, we derive necessary and sufficient conditions for shaping the ratio vector of sub-channel gains of two broadcast-channel receivers. We then apply this to two scenarios: First, in digital multicasting we present a practical capacity-achieving scheme which only uses scalar codes and linear processing. Then, we consider the joint source-channel problem of transmitting a Gaussian source over a two-user MIMO channel, where we show the existence of non-trivial cases, where the optimal distortion pair (which for high signal-to-noise ratios equals the point-to-point distortions of the individual users) may be achieved by employing a hybrid digital-analog scheme over the induced equivalent channel. Since in this approach the choice of modulation basis depends upon multiple links in the network, we coin it "network modulation".**

Index Terms—**Broadcast channel, MIMO, multicasting, generalized triangular decomposition, GSVD, GDFE, multiplicative majorization, joint source-channel coding.**

I. INTRODUCTION

The choice of modulation domain plays a major role in communication, both in deriving performance limits and in the design of practical schemes which decouple the signal processing task of channel equalization from coding. Thus, choosing the "right" basis is of central importance. For example, the capacity of the Gaussian inter-symbol interference (ISI) channel is given by the water-filling solution, applied in the frequency domain; the same transformation also allows to use popular schemes such as Orthogonal Frequency-Division Multiplexing (OFDM) which employs the discrete Fourier transform. The singular value decomposition (SVD) plays a similar role for multiple-input multiple-output (MIMO) channels. Common to both cases is *diagonalization*: They yield parallel independent equivalent channels. But do we really need such orthogonality? Capacity can be achieved for both the ISI and MIMO channels using *non-orthogonal* equivalent channels, by a receiver which performs *triangularization* of the channel[1] (rather than diagonalization) and then decision-feedback equalization or successive interference cancellation (SIC). This is done without performing any transformation at the transmitter. It is therefore natural to ask, what can be achieved by allowing *both* a transmitter transformation (in addition to the receiver one) and SIC.

One such direction, pursued by Jiang, Hager and Li [1], is the *generalized triangular decomposition* (GTD): A matrix A is decomposed as

$$A = UTV^\dagger, \qquad (1)$$

where U and V are unitary matrices, $V^\dagger$ denotes the complex conjugate of V and T is upper triangular. It is shown in [2], [3] that the transforming matrices U and V exist if and only if the diagonal elements of T obey Weyl's multiplicative majorization relation with the singular values of A (see also [4]). Since the product of these diagonal elements equals the product of the singular values of A, the decomposition performs *diagonal shaping*: it distributes the total gain between the diagonal elements in a desired way. An important special case is where it is desired to have balanced gains, i.e., the diagonal elements of T should all be equal. In that case, named the *geometric mean decomposition* (GMD) [5], the majorization condition holds for any A. When applied to MIMO communication, GMD has an advantage over SVD, that all subchannels enjoy the same gain, and thus may support codebooks of the same rate, avoiding the need for a bit-loading mechanism. This comes at the price of performing SIC at the receiver. The GMD has received considerable attention; see, e.g., [6]–[8] for some of its applications.

We take a different path, in which we wish to jointly shape the diagonals of two matrices, for the purpose of multi-terminal communication. Since with this approach the choice of basis depends upon more than one communication link, we call it *network modulation*. We jointly triangularize two matrices A_1 and A_2 as

$$A_i = U_i T_i V^\dagger, \quad i = 1, 2, \qquad (2)$$

where U_1, U_2 and V are unitary and T_1 and T_2 are upper triangular. Having the same matrix V on one of the sides of

[1]Outside the high signal-to-noise ratio regime, "near triangularization" is performed as an optimal balance between residual interference and noise.

the decomposition corresponds to applying the same transformation, and is thus suitable to two links originating (or terminating) at the same node. It turns out that in different network applications, it is important to shape the vector of *ratios* between the diagonals. We show that the sufficient and necessary condition for achievability of a ratio vector is a multiplicative majorization relation with the generalized singular values [9] of the pair (A_1, A_2).

In Section II we present known results, recalling how to achieve the point-to-point MIMO capacity using unitary triangularization. In Section III we prove the necessary and sufficient conditions for joint unitary triangularization of two matrices. In the rest of the paper we apply this result in two different scenarios, where in one we present an optimal practical scheme for a problem for which the capacity is known, and in the second we derive the (hitherto unknown) optimal performance.

Namely, for multicasting digital data over two MIMO channels, we present in Section IV a scheme which employs linear processing of scalar codebooks, much like what can be achieved in point-to-point MIMO communications using schemes such as V-BLAST [10]. This can be achieved using a uniform ratios vector, for which the majorization condition is satisfied for any channels pair. In Section V, we address the problem of transmission of an *analog* source over two MIMO links, where we show that a ratios vector of all-ones except for one element creates an equivalent channel over which a hybrid digital-analog scheme can achieve the optimal tradeoff between user distortions; thus we derive the optimal performance whenever the channels are such that this ratios vector is feasible, and argue that this is the case for a wide class of channels with two transmit antennas.

We note that the decomposition may equally be applied to cases where two transmitters communicate with a joint receiver via MIMO links (a MIMO-MAC channel). In this case the roles of the U and V matrices in (2) are interchanged. An application of the decomposition in such a setting is a MIMO extension of the "physical network coding" approach to bi-directional relays [11]. This application is beyond the scope of this paper, and appears in [12].

II. BACKGROUND: UNITARY TRIANGULARIZATION FOR MIMO CHANNELS

In this section we recall how the single-user Gaussian MIMO capacity may be achieved using multiple codebooks (each designed for a scalar AWGN channel) with SIC over an equivalent channel obtained by unitary triangularization of the form (1). To that end, we must first formalize that decomposition.

A. Unitary Triangularization

Throughout the work, we will only need to decompose matrices which belong to the following class.

Definition 1 (Proper dimensions): An $m \times n$ matrix A is said to have proper dimensions if it is full-rank and $m \geq n$.

We will need the following notation:

Definition 2 (Square part): Let A be a matrix of proper dimensions $m \times n$. The square part of A, denoted $[A]$, consists of the first n rows of A.

For decomposing non-square matrices, we need to refine the definition of triangularity.

Definition 3 (Generalized triangular matrix): Let T be a matrix of proper dimensions. We call T a generalized triangular matrix, if it has the block structure

$$T = \left(\begin{array}{c} [T] \\ 0 \end{array} \right)$$

where the square part $[T]$ is upper-triangular.

Definition 4 (Unitary triangularization): Let A be a matrix of proper dimensions. A decomposition:

$$A = UTV^\dagger$$

is called a unitary triangularization if U and V are unitary matrices, and T is generalized triangular matrix of the same dimensions as A.

Remark 1: Throughout the paper, we will assume without loss of generality that all the diagonal elements $T_{j,j}$ are real. This is similar to the definition of the SVD; any phase can be absorbed in U and V.

Note that for any unitary triangularization of a matrix A of proper dimensions $m \times n$,

$$\det\left(A^\dagger A\right) = \det\left(T^\dagger T\right) = (\det[T])^2 = \prod_{j=1}^{n} (T_{j,j})^2. \quad (3)$$

B. SIC for MIMO Channels

The exposition below follows that of the universal matrix decomposition (UCD) [13], which is in turn based upon the derivation of the MMSE version of V-BLAST, see, e.g., [14]. Later in the paper we take the triangularization to be one which is simultaneously good for two users. This is suppressed for now. We assume throughout the paper perfect channel knowledge everywhere.

We consider a point-to-point (complex) MIMO channel:

$$\boldsymbol{y} = H\boldsymbol{x} + \boldsymbol{z}, \quad (4)$$

where $\boldsymbol{x}$ is the channel input of dimensions $N_t \times 1$ subject to an average power constraint P;[2] $\boldsymbol{y}$ is the channel output vector of dimensions $N_r \times 1$; H is the channel matrix of dimensions $N_r \times N_t$ and $\boldsymbol{z}$ is an additive circularly-symmetric Gaussian noise vector of dimensions $N_r \times 1$. Without loss of generality, we assume that the noise elements are mutually-independent, identically-distributed with unit variance.

---

[2]Alternatively, one can consider an input covariance constraint $C_{\boldsymbol{x}} \triangleq E\left[\boldsymbol{x}\boldsymbol{x}^\dagger\right] \preceq C$, where by $C_1 \preceq C_2$ we mean that the matrix $(C_2 - C_1)$ is positive semi-definite.
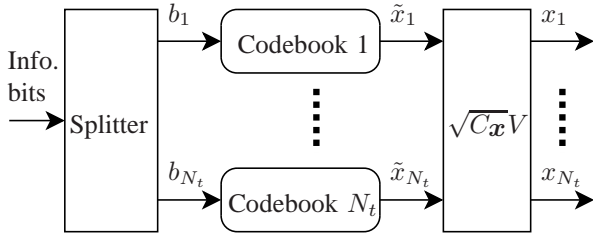
Fig. 1. Multiple-codebook transmitter with linear precoding

The capacity of this channel is given by

$$C(H, P) = \max_{C_{\boldsymbol{x}}} I(H, C_{\boldsymbol{x}}), \tag{5}$$

where the maximization is over all channel input covariance matrices $C_{\boldsymbol{x}} \geq 0$, subject to the power constraint $\text{trace}(C_{\boldsymbol{x}}) \leq P$, and

$$I(H, C_{\boldsymbol{x}}) \triangleq \log \det \left(I + H C_{\boldsymbol{x}} H^{\dagger} \right). \tag{6}$$

We may interpret $I(H, C_{\boldsymbol{x}})$ as the maximal mutual information that can be attained using an input covariance matrix $C_{\boldsymbol{x}}$, which is achievable by a Gaussian input $\boldsymbol{x}$.

In order to achieve a rate approaching this mutual information, optimal codes of long block length are needed. However, as pointed out in the introduction, we take an approach which decouples the signal-processing aspects from these of coding. We thus omit the time index throughout the paper; for example, when referring to an input vector $\boldsymbol{x}$, we mean the input at any time instant within the coding block. In a practical setting using encoder/decpder pairs of some given quality, one may easily bound the error probability of the scheme using the parameters of the codes.

When coding over a domain different than the input domain (e.g., time or space), one may start with a virtual input vector $\tilde{\boldsymbol{x}}$, related to the physical input by the linear transformation:

$$\boldsymbol{x} = \sqrt{C_{\boldsymbol{x}}} V \tilde{\boldsymbol{x}}. \tag{7}$$

We form the vector $\tilde{\boldsymbol{x}}$ in turn by taking one symbol from each of N_t parallel codebooks, of equal power $1/N_t$. The matrix V is a unitary linear precoder. See Figure 1.

Recalling the GTD (1), one may suggest to choose V by applying a unitary triangularization to

$$F \triangleq H \sqrt{C_{\boldsymbol{x}}}. \tag{8}$$

After the receiver applies the transformation $U^{\dagger}$, it is left with an equivalent triangular channel T, over which it may decode the codebooks using SIC. Unfortunately, while this "conserves" the determinant of $H C_{\boldsymbol{x}} H^{\dagger}$, it fails to do so when the identity matrix is added as in the mutual information $I(H, C_{\boldsymbol{x}})$ (6). Thus, this is optimal in the high SNR limit only, and an MMSE variation is needed in general, as next described.

We start by applying a unitary triangularization (as in definition 4) to an augmented matrix:

$$\begin{pmatrix} F \\ I \end{pmatrix} \triangleq G = U T V^{\dagger}, \tag{9}$$
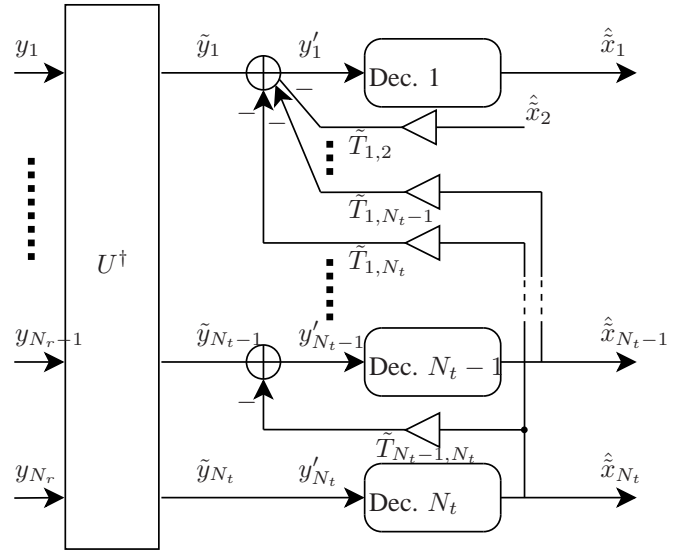

Fig. 2. SIC-based receiver.

where the identity matrix I has dimensions $N_t \times N_t$. Note that, by construction, G is of proper dimensions, regardless of the dimension and rank of the channel matrix H. That is, it has dimensions $(N_t + N_r) \times N_t$ and *full rank*. The square matrices U and V have dimensions $N_t + N_r$ and N_t, respectively. This allows to decompose the total rate into terms associated with the diagonal values of the matrix T, as follows:

$$I(H, C_{\boldsymbol{x}}) = \log \det \left(I_{N_r} + F F^{\dagger} \right) \tag{10}$$

$$= \log \det \left(I_{N_t} + F^{\dagger} F \right) \tag{11}$$

$$= \log \det \left(G^{\dagger} G \right) \tag{12}$$

$$= \sum_{j=1}^{N_t} \log(T_{j,j})^2 \triangleq \sum_{j=1}^{N_t} R_j, \tag{13}$$

where (10) follows by the definitions (6) and (8), (11) is justified by Sylvester's determinant Theorem (see e.g. [15]), (12) is a direct application of the definition (9), and the equality (13) is due to (3). Using the matrices obtained by this decomposition, the following scheme communicates scalar codebooks of rates $\{R_j\}$.

The transmitted signal is formed using (7). At the receiver, we use a matrix W, consisting of the upper-left $N_r \times N_t$ block of U: $\tilde{\boldsymbol{y}} = W^{\dagger} \boldsymbol{y}$. This results in an equivalent channel:

$$\tilde{\boldsymbol{y}} = W^{\dagger}(F V \tilde{\boldsymbol{x}} + \boldsymbol{z}) = W^{\dagger} F V \tilde{\boldsymbol{x}} + W^{\dagger} \boldsymbol{z} \triangleq \tilde{T} \tilde{\boldsymbol{x}} + \tilde{\boldsymbol{z}}. \tag{14}$$

Note that since W is not unitary, the statistics of $\tilde{\boldsymbol{z}}$ differ from those of $\boldsymbol{z}$. We denote the covariance matrix of the equivalent noise by $C_{\tilde{\boldsymbol{z}}} = W W^{\dagger}$. Finally, SIC is performed, i.e., the codebooks are decoded from last to first, where each codebook is decoded from:

$$y_j' = \tilde{y}_j - \sum_{l=j+1}^{N_t} \tilde{T}_{j,l} \hat{\tilde{x}}_l, \tag{15}$$

where $\hat{\tilde{x}}_l$ is the decoded symbol from the l-th codebook; see Figure 2. Assuming correct decoding of "past" symbols, i.e.

$\hat{\tilde{x}}_l = \tilde{x}_l$ for all $l > j$, the scalar channel for decoding of the j-th codebook is given by:

$$y'_j = \tilde{T}_{j,j}\tilde{x}_j + \sum_{l=1}^{j-1} \tilde{T}_{j,l}\tilde{x}_l + \tilde{z}_j \,. \tag{16}$$

Since $\tilde{T}$ is not triangular, the second term in this scalar channel (resulting from elements below the diagonal of $\tilde{T}$) acts as interference. The signal-to-intereference-and-noise ratio (SINR) is given by:

$$S_j = \mathrm{Var}\left(\tilde{\mathrm{x}}_j\big|\tilde{\mathbf{y}}, \tilde{\mathrm{x}}_{j+1}^{N_t}\right) = \frac{(\tilde{T}_{j,j})^2}{C_{\tilde{\mathbf{z}};j,j} + \sum_{l=1}^{j-1}(\tilde{T}_{j,l})^2} \,, \tag{17}$$

where $C_{\tilde{\mathbf{z}};i,j}$ denotes the (i,j) entry of $C_{\tilde{\mathbf{z}}}$.

The following, which is equivalent to Lemma III.3 in [13], shows optimality of the scheme.

Proposition 1: For any channel H and input covariance matrix $C_{\boldsymbol{x}}$, the SINRs S_j (17) of the transmission scheme above satisfy:

$$\log(1 + S_j) = R_j \,, \qquad \forall j = 1, \ldots, N_t \,. \tag{18}$$

where the rates R_j are given by (13).

This completes the recipe for a digital transmission scheme which achieves $I(H, C_{\boldsymbol{x}})$: for a given input covariance matrix $C_{\boldsymbol{x}}$, choose the individual codebook rates to approach $\{R_j\}$, the sum of which equals the mutual information afforded by the MIMO channel (6). By (18), the successive decoding procedure will succeed with arbitrarily low probability of error for these rates (asymptotically for high-dimensional optimal scalar AWGN codes). Taking $C_{\boldsymbol{x}}$ be the covariance matrix maximizing (5), capacity can be achieved.

The above exposition proves the optimality of the "scalar coding" approach - the combination of scalar AWGN codebooks, linear processing, and SIC. This approach offers reduced complexity and easy-to-analyze performance when the channel is known at both ends ("closed loop"). Indeed, special cases of this approach have been suggested and used. In particular, using the SVD results in a *diagonal* equivalent channel matrix T, establishing parallel virtual AWGN channels (no SIC is needed), see [15]. Other schemes, such as generalized decision feedback equalization (GDFE) and Vertical Bell-Laboratories Space-Time coding (V-BLAST), see [10], [16], are based on the QR decomposition. These do not require linear precoding, i.e., $V = I$. The UCD [13] uses both a linear precoder and SIC, in order to achieve T with diagonal elements that are all equal.

All of these schemes have significant advantages over direct capacity-achieving implementation for MIMO channels. Such high-complexity schemes, e.g., using bit-interleaved coded modulation (BICM) in conjunction with sphere detection, essentially require the same resources as if working in an "open loop" mode. Thus, the complexity involved is similar to that required for approaching the isotropic mutual information of the channel, when only the receiver knows the channel.

We conclude this section by pointing out a simple extension to a unitary transformation which induces a *block-traingular* matrix rather than a strictly triangular one. That is, if the matrix R in (9) is of the block generalized upper-triangular form:

$$T = \begin{pmatrix} T_{1,1} & T_{1,2} & \cdots & T_{1,K} \\ 0 & T_{2,2} & \cdots & T_{2,K} \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & T_{K,K} \\ 0 & \cdots & 0 & 0 \end{pmatrix}, \tag{19}$$

where $T_{k,l}$ is a block of dimensions $N_k \times N_l$, such that $\sum_{k=1}^{K} N_k = N_t$ (thus the last row of blocks consists of N_r all-zero rows). In that case, we employ $K \leq N_t$ codes in parallel, each over an equivalent $N_k \times N_k$ MIMO channel, achieved by "block-SIC":

$$\boldsymbol{y}'_j = \sum_{l=1}^{j} \tilde{T}_{j,l}\tilde{\boldsymbol{x}}_l + \tilde{\boldsymbol{z}}_j \,, \qquad j = 1, \ldots, K \,, \tag{20}$$

where $\tilde{T}_{j,l}$ si of dimensions $N_j \times N_l$. Seen as Gaussian MIMO channels (i.e., seeing residual interference as noise) we achieve, as an extension to Proposition 1, a rate

$$R_j = \log\det\left(T_{j,j}(T_{j,j})^\dagger\right) \tag{21}$$

over each such block channel.

III. JOINT TRIANGULARIZATION WITH SHAPED DIAGONAL RATIO

In this section we prove the necessary and sufficient condition for the existence of the joint triangularization (2), formally defined as follows.

Definition 5 (Joint Unitary Triangularization): Let A_1 and A_2 be matrices of proper dimensions with the same number of columns. A decomposition:

$$A_i = U_i T_i V^\dagger, \, i = 1, 2$$

is called a joint unitary triangularization if it consists of a unitary triangularization (as in Definition 4) for both A_1 and A_2.

In order to state the condition, we need the following definitions.

Definition 6 (Generalized singular values [9]): For any (ordered) matrix pair (A_1, A_2), the generalized singular values (GSVs) are the positive solutions a of the equation

$$\det\left\{A_1^\dagger A_1 - a^2 A_2^\dagger A_2\right\} = 0 \,.$$

Let the GSV vector $\boldsymbol{\mu}(A_1, A_2)$ be the vector composed of all GSVs (including their algebraic multiplicity), ordered non-increasingly.[3]

---

[3] The number of GSVs is always n, even if the number of finite solutions is smaller. We define a GSV to be infinite, if the corresponding GSV of the matrices in reverse order is zero. If the number of finite and infinite solutions is smaller than n, this suggests that the column rank can be reduced without changing the problem; we shall assume the problem is in its reduced form.

Remark 2: For matrices of proper dimensions, $\boldsymbol{\mu}$ is of length n.

Remark 3: When A_1 and A_2 are square and non-singular, $\boldsymbol{\mu}(A_1, A_2)$ consists of the singular values of $A_1 A_2^{-1}$.

Definition 7 (Diagonal ratios vector): Let T_1 and T_2 be two generalized upper-triangular matrices of proper dimensions $m_1 \times n$ and $m_2 \times n$, respectively, with non-negative diagonal elements. The diagonal ratios vector $\boldsymbol{r}(T_1, T_2) = \boldsymbol{r}([T_1], [T_2])$ is the n-length vector which contains all ratios $T_{1;j,j}/T_{2;j,j}$, ordered non-increasingly, where $T_{i;j,k}$ denotes the (j, k) entry of T_i $(i = 1, 2)$.

Definition 8 (Multiplicative majorization (see [4])): Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be two n-dimensional vectors satisfying

$$\prod_{j=1}^{n} |x_j| = \prod_{j=1}^{n} |y_j| \,.$$

Then we say that $\boldsymbol{x}$ majorizes $\boldsymbol{y}$ $(\boldsymbol{x} \succeq \boldsymbol{y})$ if for any $1 \leq k < n$,

$$\prod_{j=1}^{k} |x_j| \geq \prod_{j=1}^{k} |y_j| \,.$$

We are now ready to prove the main result of this section.

Theorem 1: Let A_1 and A_2 be two matrices of proper dimensions $m_1 \times n$ and $m_2 \times n$, respectively. Then the joint unitary triangularization of Definition 5 exists if and only if

$$\boldsymbol{\mu}(A_1, A_2) \succeq \boldsymbol{r}(T_1, T_2) \,. \tag{22}$$

Proof: **Achievability part.** We present here the proof for case when the matrices are square ($m_1 = m_2 = n$). The extension to the general proper-dimension case is relegated to Appendix A.

In the square case, A_1 and A_2 must be invertible. Define the matrix $B = A_1 A_2^{-1}$. The vector composed of the singular values of B ordered non-decreasingly coincides with the GSV vector of (A_1, A_2), $\boldsymbol{\mu}(A_1, A_2)$ (see [9], [17]). Thus, it majorizes $\boldsymbol{r}(T_1, T_2)$, by assumption. Hence, according to the GTD [1], the matrix B can be decomposed as

$$B = \tilde{U}_1 R \tilde{U}_2^\dagger \,, \tag{23}$$

where $\tilde{U}_1$ and $\tilde{U}_2$ are unitary and R is upper-triangular with a diagonal which equals the absolute values of the entries of $\boldsymbol{r}(T_1, T_2)$. Now, apply RQ decompositions to $\tilde{U}_i^\dagger A_i$ $(i = 1, 2)$ to achieve

$$\tilde{U}_i^\dagger A_i = T_i V_i^\dagger \,, \tag{24}$$

where T_i are upper-triangular with positive diagonal entries and V_i are unitary. By substituting (24) into (23) we have

$$\tilde{U}_1 T_1 V_1^\dagger V_2 T_2^{-1} \tilde{U}_2^\dagger = \tilde{U}_1 R \tilde{U}_2^\dagger \,,$$

which is equivalent to

$$V_1^\dagger V_2 = T_1^{-1} R T_2 \,. \tag{25}$$

We note that the l.h.s. of (25) is unitary, whereas its r.h.s. is an upper-triangular matrix with positive diagonal entries. An equality between such matrices can hold only if both matrices are equal to the identity matrix of the appropriate dimensions ($n \times n$). Thus, we have

$$V \triangleq V_1 = V_2 \,.$$
$$T_{1;i,i} = R_{i,i} T_{2;i,i} \,, \qquad i = 1, ..., n \,.$$

Since the diagonal of R is equal to $\boldsymbol{r}(T_1, T_2)$, this establishes the desired decomposition (2).

Converse part. Assume, in contradiction, that A_1 and A_2 can be decomposed as in (2) such that $\boldsymbol{\mu}(A_1, A_2) \not\succeq \boldsymbol{r}(T_1, T_2)$. Note that $\boldsymbol{\mu}(T_1, T_2) = \boldsymbol{\mu}(A_1, A_2)$. Moreover, $[T_1]$ and $[T_2]$ are non-singular square matrices of dimensions $n \times n$ with a GSV vector that is equal to the GSV vector of (T_1, T_2), i.e.,

$$\boldsymbol{\mu}([T_1], [T_2]) = \boldsymbol{\mu}(T_1, T_2) = \boldsymbol{\mu}(A_1, A_2) \,,$$
$$\boldsymbol{r}([T_1], [T_2]) = \boldsymbol{r}(T_1, T_2) \,.$$

Thus $\boldsymbol{\mu}([T_1], [T_2]) \not\succeq \boldsymbol{r}([T_1], [T_2])$, which in turn implies that the upper-triangular matrix $B \triangleq [T_1][T_2]^{-1}$ has a diagonal $\boldsymbol{r}([T_1], [T_2])$ and a singular values vector $\boldsymbol{\mu}([T_1], [T_2])$. But according to Weyl's condition [2]:

$$\boldsymbol{\mu}(A_1, A_2) = \boldsymbol{\mu}([T_1], [T_2]) \succeq \boldsymbol{r}([T_1], [T_2]) = \boldsymbol{r}(T_1, T_2) \,,$$

in contradiction to the assumption. ∎

Remark 4: By the unitarity of U and V, the products of $\boldsymbol{\mu}$ and $\boldsymbol{r}$ are equal. Thus, the majorization relations mean that the diagonal ratios are always "less spread" than the generalized singular values.

Remark 5 (Relation to GSVD): The GSVD [9] can be stated in a triangular form (2), with diagonals ratio $\boldsymbol{r}(T_1, T_2) = \boldsymbol{\mu}(A_1, A_2)$. Thus, the GSVD is a limiting case with maximal ratio spread.

Remark 6 (Relation to GTD): Taking in the joint decomposition $H_2 = I$ yields the GTD of H_1 [1]; further, the GSV become the singular values vector of H_1. The existence condition, in turn, reduces to the Weyl condition (see e.g. [1]). In this sense, the condition in Theorem 1 may be seen as a generalized Weyl condition for joint triangularization.

Remark 7 (Relation to the generalized Schur decomposition): This decomposition, also called the QZ-decomposition [17], is a special case of the joint decomposition with $U_1 = U_2$. It can be shown that the diagonal ratio vector induced by this decomposition is unique, i.e., requiring that the unitary matrices are the same on both sides prohibits shaping of the diagonal ratio.

The joint unitary triangularization (and, as a special case, the GTD) can also be relaxed to a block form. Here we do not require the matrices T_i to be generalized upper-triangular, but merely in blocks of size n_k, $\sum_{k=1}^{K} n_k = n$ as in (19). Let the blocks of T_i be $T_{i;k,l}$, of size $n_k \times n_l$. We define the block diagonal ratios as the absolute values of the ratios between the determinants of corresponding blocks, $|\det(T_{1;k,k})/\det(T_{2;k,k})|$, where $T_{i;j,k}$ denotes the (j, k)
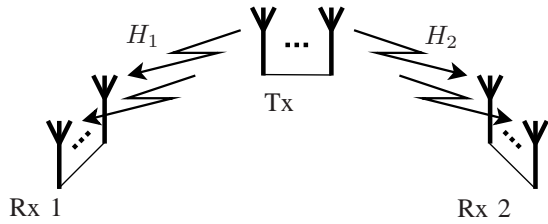
Fig. 3. Two-user MIMO BC channel. Even though both users may have similar channel quality, the actual channel matrices H_1 and H_2 differ.

block of T_i ($i = 1, 2$). Denote by $\{n_{k_l}\}_{k=1}^{K}$ the indices satisfying:

$$r_{k_1} \geq r_{k_2} \geq \cdots \geq r_{k_K},$$

where $r_k \triangleq {}^{n_k}\sqrt{|\det(T_{1;k,k})/\det(T_{2;k,k})|}$. Denote by $\boldsymbol{r}$ the K-length vector composed of $\{r_k^{n_k}\}_{k=1}^{K}$ ordered in decreasing order of r_k (i.e., $\boldsymbol{r}$ is composed of the ratios between determinants of corresponding blocks). Further denote by $\boldsymbol{\mu}$ the K-length vector composed of the products of sizes $\{k_l\}_{l=1}^{K}$ of the GSVs of (A_1, A_2) ordered non-increasingly, i.e., the first entry of $\boldsymbol{\mu}$ is the product of the largest k_1 GSVs, its second entry is the product of the next k_2 GSVs, etc. Then the desired block triangularization is possible if and only if $\boldsymbol{r} \succeq \boldsymbol{\mu}$.

IV. TRANSMISSION SCHEME FOR MULTICASTING

In this section we combine the results of Sections II and III in order to derive an optimal practical communication scheme for two-user multicasting.

The Gaussian MIMO broadcast (BC) channel has one transmit and two receive nodes, where each received signal is related to the transmitted signal through (4) (see Figure 3), i.e.:

$$\boldsymbol{y}_i = H_i \boldsymbol{x} + \boldsymbol{z}_i, \qquad i = 1, 2. \tag{26}$$

This channel has received much attention over the past decade. Unlike the single-input single-output (SISO) case, the Gaussian MIMO BC channel is not degraded. Nevertheless, capacity regions were established for some scenarios, such as private-messages only, and for a common message with a single private message, and bounds were derived for others, see [18]–[22] and references therein.

We focus our attention on the multicast (common-message) problem, the capacity of which is long known to equal the (worst-case) capacity of the compound channel [23]–[25], with the compound parameter being the channel matrix index:

$$C(H_1, H_2, P) = \max_{C_{\boldsymbol{x}}} \min_{i=1,2} I(H_i, C_{\boldsymbol{x}}), \tag{27}$$

where the maximization is under the same conditions as in (5) and $I(H, C_{\boldsymbol{x}})$ was defined in (6).

We wish to use a scalar-coding approach, as applied to the point-to-point setting in Section II. Indeed, the private-message MIMO BC capacity can be achieved by scalar-coding (in this case dirty-paper coding) techniques; see, e.g., [26]. In the presence of a common message, however, to our knowledge, no scalar capacity-approaching coding solutions are known.

QR-based schemes fail, since requiring the individual streams to be simultaneously decodable at all the receivers implies that the rate *per stream* is governed by the smallest of the corresponding diagonal elements, (potentially) inflicting an unbounded rate penalty. Adapting SVD to this scenario has an additional problem: the decomposition requires multiplying by a channel-dependent (unitary) matrix at the transmitter, which prevents from using this decomposition for more than one channel simultaneously.[4] As a result of these difficulties, other techniques were proposed, which are suboptimal in general, see, e.g., [27], [28].

In this section, we present an optimal successive-decoding (low-complexity) scheme for a two-user common-message Gaussian MIMO BC channel. Specifically, the proposed scheme is based upon SIC and good *scalar* AWGN codes, in conjunction with the following special case of the decomposition in Theorem 1:

Corollary 1: Let A_1 and A_2 be two matrices of proper dimensions, with

$$\det\left(A_1 A_1^{\dagger}\right) \geq \det\left(A_2 A_2^{\dagger}\right). \tag{28}$$

Then there exists a joint triangularization (2) where

$$T_{1;j,j} \geq T_{2;j,j} \quad \forall j = 1, \ldots, N_t.$$

Proof: An equivalent condition to (28) is that the product of the entries of $\boldsymbol{\mu} = \boldsymbol{\mu}(A_1, A_2)$ is at least one. Let $\bar{\mu} \geq 1$ be the geometrical mean of $\boldsymbol{\mu}$, and let the vector $\boldsymbol{r}$ be the same size as $\boldsymbol{\mu}$, with all the elements equal to $\bar{\mu}$. By construction, $\boldsymbol{\mu} \succeq \boldsymbol{r}$, thus by Theorem 1 there exists a joint triangularization with this ratio. Consequently, there exists a decomposition such that for all elements $T_{1;j,j} = \bar{\mu} T_{2;j,j} \geq T_{2;j,j}$. ■

Remark 8 (Admissible diagonal ratios): The proof suggests that the diagonal ratios vector be made uniform. This is always possible, but is not the only choice (unless $I(H_1, C_{\boldsymbol{x}}) = I(H_2, C_{\boldsymbol{x}})$).

For some channels H_1, H_2, let $C_{\boldsymbol{x}}$ be a capacity-achieving input covariance matrix, and assume without loss of generality that $I(H_1, C_{\boldsymbol{x}}) \geq I(H_2, C_{\boldsymbol{x}})$. Define the augmented matrices G_1 and G_2 as in (9). By Corollary 1, there exists a joint triangularization (2) such that each diagonal element of $[T_1]$ is at least equal to the corresponding element of $[T_2]$. On account of (10)-(13) we have that:

$$C(H_1, H_2) = \sum_{j=1}^{N_t} \log(T_{2;j,j})^2 \triangleq \sum_{j=1}^{N_t} R_j.$$

This rate can be approached using SIC at each receiver as in the point-to-point case of Section II. Specifically, $\tilde{\boldsymbol{x}}$ is formed

---

[4]Indeed, the GSVD allows to use a single transformation for two different channels at one of the ends, but for each virtual parallel channel it yields a different gain for each user, thus not solving the inefficiency mentioned above. In fact, using GSVD may result in *worse* performance than using a QR-based receiver without any transformation at the transmitter since the spread of the diagonal ratio is maximal, see Remark 5.

from N_t codebooks of rates $\{R_j\}$ and power $1/N_t$ each. The transmitted vector is given by the linear precoding (7) and the receiver i performs the linear transformation (14) and SIC (15) (substituting U_i and T_i for U and T, respectively). Now Proposition 1 guarantees correct decoding of all codebooks for receiver 2. Since in receiver 1 each SINR can only be greater, it will be able to decode as well.

Remark 9 (Number of codebooks): If desired, one may work with any number of codebooks above N_t, as stated in [13]. To see that, add "virtual transmit antennas" with corresponding zero channel gains. The capacity remains unchanged, and the optimal channel input covariance matrix will not allocate power to these "antennas". The number of codebooks is equal to the number of antennas, including the additional virtual ones.

Remark 10 (Private messages): If, in addition to the common message intended to both users, there are private messages (messages intended for individual users), superposition may be used. That is, part of the transmit power is dedicated to the private messages. Then, for the purpose of the common message, the transmission for the private messages is considered as noise. This approach was shown in [22] to be capacity-achieving in the presence of a single private message, and under some conditions on the rate and power - also in the presence of two private messages (even when these conditions do not hold, superposition gives the best known performance). The scheme presented in this section may be used for the common-message layer of these superposition schemes as well.

V. HDA Transmission for Source Multicasting

In this section we turn from the purely digital setting to a joint source-channel coding (JSCC) problem, where we wish to multicast an analog source to two destinations, where each destination should enjoy reconstruction quality according to the capacity afforded by its channel.

The transmission of a source over a BC channel is one of the main applications of JSCC. In this setting, JSCC may be greatly superior to transmission based upon source-channel separation. In a classical example, a white Gaussian source needs to be transmitted over a two-user AWGN BC channel, with one channel use per source sample, under mean-squared error (MSE) distortion. Analog transmission [29] achieves the optimal performance for each user as if the other user did not exist. In contrast, the separation-based scheme (concatenation of successive-refinement and broadcast codes) yields a tradeoff, where if we wish to be optimal for the user with worse signal-to-noise ratio (SNR), then both users have the same distortion, while optimality for the user with better SNR means that the distortion for the other user is trivial (equals the source variance). See, e.g., [30, App. A].

We focus on the transmission of an i.i.d. circularly-symmetric Gaussian source S to two destinations over a MIMO-BC channel (26), with one channel use per source sample. We measure the quality of the reproductions $\hat{S}_i$ using the MSE distortion measure. Thus, we wish to maximize the tradeoff between the signal-to-distortion ratios (SDRs), defined as

$$\mathrm{SDR}_i \triangleq \frac{\mathrm{Var}\,(\mathrm{S})}{\mathrm{Var}\left(\hat{\mathrm{S}}_\mathrm{i} - \mathrm{S}\right)}, \quad i = 1, 2. \tag{29}$$

The achievable SDR region $\mathcal{S}(H_1, H_2)$ is defined as the closure of all pairs which can be achieved by some encoding-decoding scheme.

This general problem of describing $\mathcal{S}(H_1, H_2)$ has not received much attention. Nevertheless, in the special cases of diagonal or Toeplitz channel matrices, it reduces to the better known problem of transmission over a colored and/or bandwidth-mismatched Gaussian BC channel, for which different schemes which outperform the separation approach have been presented, see e.g. [31]–[34]. However, even for these cases optimality claims are not abundant. In [33], Kochman and Zamir show asymptotic optimality for high SNR, where the channels have the same bandwidth as the source, and one user enjoys a better channel than the other at all frequencies. In [34], Taherzadeh and Khandani show that optimality in the slope sense (weaker than high-SNR asymptotic optimality) is possible for white channels where the bandwidth (BW) is an integer multiple of the source BW. A similar slope argument applies to the general MIMO case as well.

A simple outer bound on the achievable SDR region is given by the following.

Proposition 2: $\mathcal{S}(H_1, H_2) \subseteq \bar{\mathcal{S}}(H_1, H_2)$, where the bounding region $\bar{\mathcal{S}}(H_1, H_2)$ is given by:

$$\bigcup_{C_{\boldsymbol{x}}} \left\{(\mathrm{SDR}_1, \mathrm{SDR}_2) : \log(\mathrm{SDR}_i) \leq I(H_i, C_{\boldsymbol{x}})\right\},$$

where the union is over all matrices $C_{\boldsymbol{x}} \geq 0$ such that $\mathrm{trace}\,(C_{\boldsymbol{x}}) \leq P$, and where $I(H, C_{\boldsymbol{x}})$ was defined in (6).

The proof follows that of the classical source-channel converse [35], taking into account that both users share the same channel input.

In Section V-A we find sufficient conditions for achieving points on the boundary of this region. Then, in Section V-B we present, for the case of up to two transmit antennas, a simple sufficient condition such that all of the region $\bar{\mathcal{S}}(H_1, H_2)$ can be achieved. Unlike previous work, this proves strict, non-asymptotic optimality; it applies to some cases of color and bandwidth mismatch, although not to the white BW-expansion case.

A. Optimality by HDA Transmission

We give a constructive achievability proof, which combines a hybrid digital-analog (HDA) scheme by Mittal and Phamdo [31] with the joint triangularization approach; the optimum is achievable whenever the diagonal ratio can be shaped according to the needs of the HDA scheme. In order to understand the function of the HDA scheme, we need to consider the following related scenario. In a JSCC multicasting problem as above, the BC channel is SISO, i.e., $N_t = N_r = 1$

and the channel matrices reduce to scalars h_i. However, in addition, the transmitter node may send some digital data to the users (identical for both) over a digital channel of rate R_{digital} nats per use of the BC channel.

Proposition 3: In the setting above, the optimal performance is given by:

$$\text{SDR}_i = (1 + h_i^2 P) \cdot \exp\{R_{\text{digital}}\}.$$

Proof: We use a vector quantizer which decomposes each sample of the Gaussian source S as

$$S = \tilde{S} + Q. \tag{30}$$

The first term is the quantized source, while the second is the quantization error. By quadratic-Gaussian rate-distortion theory (see e.g. [36]), in the limit of high quantizer dimension, a quantizer of rate R_{digital} may achieve:

$$\text{SDR}_{\text{digital}} \triangleq \frac{\text{Var}(S)}{\text{Var}(Q)} = \exp\{R_{\text{digital}}\}.$$

Now the quantizer output representing $\tilde{S}$ is sent over the digital channel, thus $\tilde{S}$ can be reconstructed exactly. Given $\tilde{S}$, the reconstruction error of S becomes that of Q. That is,

$$\text{SDR}_i = \frac{\text{Var}(S)}{\text{Var}\left(\hat{Q}_i - Q\right)} = \frac{\text{Var}(Q)}{\text{Var}\left(\hat{Q}_i - Q\right)} \cdot \text{SDR}_{\text{digital}}$$

$$\triangleq \text{SDR}_{\text{analog},i} \text{SDR}_{\text{digital}},$$

where $\hat{Q}_i$ is the reconstruction of Q at receiver i using the SISO BC channel. Finally by [29], analog transmission of Q achieves $\text{SDR}_{\text{analog},i} = 1 + h_i^2 P$, yielding the desired SDRs. No scheme can achieve better performance, by considerations similar to those leading to Proposition 2. ∎

We use this HDA approach to prove the following.

Theorem 2: Denote by $\boldsymbol{\mu}$ the GSV vector of the augmented matrices (9) of the channels with some input covariance matrix $C_{\boldsymbol{x}}$. If

$$\prod_{j=1}^{N_t} \boldsymbol{\mu}_j \leq 1 \leq \prod_{j=1}^{N_t-1} \boldsymbol{\mu}_j, \tag{31}$$

then any pair $(\text{SDR}_1, \text{SDR}_2)$ such that $\log \text{SDR}_i \leq I(H_i, C_{\boldsymbol{x}})$ is achievable.

Proof: It follows by Theorem 1 that there exists a joint unitary triangularization with diagonal ratios vector which is all one except for the last element. The diagonal of T_i can thus be made to satisfy

$$T_{1;j,j} = T_{2;j,j} \triangleq t_j, \qquad \forall j = 1, \ldots, N_t - 1.$$

If we were to send digital data over the MIMO-BC channel using this particular triangularization, then by (13) we could send over these $N_t - 1$ channels a rate of:

$$R_{\text{digital}} \triangleq \sum_{j=2}^{N_t} R_j = \sum_{j=2}^{N_t} \log t_j^2.$$

This does not change if we replace, in the transmission scheme, $\tilde{x}_1$ by a different signal of the same variance P/N_t. Furthermore, regardless of the signal $\tilde{x}_1$, if the codebooks of subchannels $2, \ldots, N_t - 1$ are correctly decoded then the receiver i can obtain the equivalent channel (recall (16)):

$$\boldsymbol{y}'_{i;1} = \tilde{T}_{1;1,1} + \boldsymbol{z}_{i;1}$$

which must have by Proposition 1, a signal-to-noise ratio of

$$\text{SNR}_{\text{analog},i} = (T_{i;1,1})^2 - 1.$$

At this stage we have turned the MIMO BC channel into the combination of a digital channel of rate R_{digital} and a SISO BC channel of signal-to-noise ratios $\text{SNR}_{\text{analog},i}$. On account of Proposition 3, one can achieve

$$\log \text{SDR}_i = \log(1 + \text{SNR}_{\text{analog}}) + R_{\text{digital}}$$

$$= \sum_{j=1}^{N_t} \log(T_{i;j,j})^2$$

$$= I(H_i, C_{\boldsymbol{x}}), \qquad i = 1, 2,$$

where the last equality is on behalf of (10)-(13). ∎

Remark 11: In fact, full triagularization is not needed. It would have been sufficient to achieve a block-triangular structure, where the interference between the first $N_t - 1$ channels is arbitrary (conserving the determinant of the block in T_i). However, as indicated at the end of Section III, this does not allow to relax the condition (31). Moreover, the triangular form is advantageous from the point of view of complexity (see Section III).

Theorem 2 does not imply that $\bar{S}(H_1, H_2)$ is fully achievable, since the conditions on the GSVs should be verified separately for each optimal input covariance matrix $C_{\boldsymbol{x}}$. However, in the sequel we show that for $N_t \leq 2$, the condition can be verified directly on the channel matrices H_1 and H_2. Similarly, if the channel matrices are of (any) proper dimensions, then at the limit of high SNR (as the choice $C_{\boldsymbol{x}} = I$ becomes optimal), the GSVs of the augmented matrices approach those of (H_1, H_2), thus the condition may be applied to the channel matrices directly, verifying achievability of the whole region at once.

B. Two Transmit Antennas

In this section we consider the case where $N_t = 2$. In that case, the GSV vector $\boldsymbol{\mu}(H_1, H_2)$ has two elements. We say that the GSV vector is *mixed*, if one of the elements is at least one, and the other is at most one. The following is proven in Appendix B.

Lemma 1: Let H_1 and H_2 be two matrices of proper dimensions, with $n = 2$ columns. Let

$$G_i = \begin{pmatrix} H_i \sqrt{C} \\ I \end{pmatrix}$$

be the augmented matrices (as in (9)) for some Hermitian matrix $C \geq 0$. Then if $\boldsymbol{\mu}(H_1, H_2)$ is mixed, $\boldsymbol{\mu}(G_1, G_2)$ is mixed as well.

We use this lemma and Theorem 2 to prove the following.

Corollary 2: Let H_1, H_2 be channel matrices with $N_t = 2$. If $\boldsymbol{\mu}(H_1, H_2)$ is mixed, then the bounding region $\bar{\mathcal{S}}(H_1, H_2)$ of Proposition 2 is achievable.

Proof: For any point on the boundary of $\bar{\mathcal{S}}(H_1, H_2)$, let $\boldsymbol{\mu}$ be the GSV vector of the augmented matrices with the corresponding $C_{\boldsymbol{x}}$. By Lemma 1, $\boldsymbol{\mu}$ is mixed as well. Now if the product of $\boldsymbol{\mu}$ is at most one, we can apply Theorem 2. If it is greater than one, we switch the indices between H_1 and H_2, and then apply Theorem 2. ∎

Unfortunately, this result cannot be generalized to the case $N_t > 2$: although at any dimension it remains true that the number of GSVs smaller or greater than one is not changed by the augmentation, this property does not hold for *products* of GSVs as required for applying Theorem 2.

In order to demonstrate this result, consider the the simplest example, a diagonal two-input two-output case:[5]

$$H_i = \begin{pmatrix} \alpha_i & 0 \\ 0 & \beta_i \end{pmatrix}, \quad i = 1, 2. \tag{32}$$

The bounding SDR region $\bar{\mathcal{S}}(H_1, H_2)$ now becomes:

$$\bigcup_{0 \leq \gamma \leq 1} \left\{ (\mathrm{SDR}_1, \mathrm{SDR}_2) : \tag{33} \right.$$

$$\left. \mathrm{SDR}_i \leq \left(1 + |\alpha_i|^2 \gamma P\right)\left(1 + |\beta_i|^2 (1 - \gamma)P\right) \right\}.$$

In this expression, γ is the portion of the transmit power sent over the first band.

We point out a few special cases where points on the surface of this region are achievable by known strategies.

1) No BW expansion: analog transmission. If one of the bands has zero capacity, e.g., $\beta_1 = \beta_2 = 0$, (33) reduces to: $\mathrm{SDR}_i \leq 1 + |\alpha_i|^2 P$, which is achievable via analog transmission [29]. If for each user a different band is usable, e.g., $\alpha_1 = \beta_2 = 0$, any transmission (digital or analog) which is orthogonal between users is optimal.

2) Equal SDRs: digital transmission. A point on the boundary which satisfies $\mathrm{SDR}_1 = \mathrm{SDR}_2$ may be achieved by quantizing the source and then using a digital common-message code for the BC channel.

3) One equal band: HDA transmission. If for one of the bands the gains are equal, e.g., $|\beta_1| = |\beta_2| = \beta$, we can use that band for digital transmission with rate $R_{\mathrm{digital}} = \log(1 + \beta^2 P)$ and then apply Proposition 3 to achieve the bound (33).

Using network modulation, we can extend the HDA transmission (case 3 above), by transforming a diagonal channel where none of the gains is equal between users, to an equivalent triangular channel where for one of the bands the gain is equal. This can be done under the condition (31), which specializes to (allowing to swap roles between matrices):

$$|\alpha_1|^2 \geq |\alpha_2|^2 \quad \text{and} \quad |\beta_1|^2 \leq |\beta_2|^2 \tag{34}$$

[5]Being diagonal, this channel may be obtained from a single-input single-output Gaussian inter-symbol interference channel which has a two-step frequency response, by applying the discrete Fourier transform.
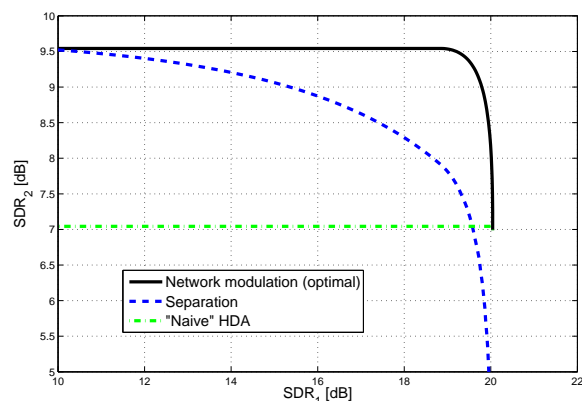

Fig. 4. Performance comparison for $\alpha_1 = 1$, $\beta_1 = 10, \alpha_2 = \beta_2 = 2, P = 1$.

or vice versa. This is an "anti-degradedness" condition: no user can have better SNR on both bands. This condition subsumes all the cases mentioned above. It is not known whether it is a necessary condition, but at least for the case where both channels are white ($\alpha_i = \beta_i$), it was shown in [37] that simultaneous optimality is *not* possible.

Figure 4 shows a numerical evaluation of performance for some gain values. It can be appreciated that the optimal performance imposes almost no tradeoff between users. Indeed the only tradeoff comes from the need to choose the same $C_{\boldsymbol{x}}$. Thus, in of high SNR, both users attain their optimal single-user performance. For comparison, we show the performance of a separation-based scheme, where a successive-refinement source code is transmitted over a digital broadcast channel code, as well as that of a "naïve" HDA scheme, where transmission is digital over one band and analog over the other.

APPENDIX A
JOINT DECOMPOSITION FOR NON-SQUARE MATRICES

In this Appendix we complete the proof of the direct part of Theorem 1, by considering the general proper-dimension case.

We start by decomposing A_i using the QR decomposition:

$$A_i = Q_i R_i, \qquad i = 1, 2,$$

where Q_i is unitary and R_i is upper-triangular with non-negative diagonal entries. Moreover, the GSV vectors of (A_1, A_2) and (R_1, R_2) are equal, since A_i and R_i are equal up to a unitary transformation on the left, i.e., $\boldsymbol{\mu}(A_1, A_2) = \boldsymbol{\mu}(R_1, R_2)$.

Since A_i is full-rank and $m_i \geq n$, the diagonal elements of R_i are all (strictly) positive and the entries on its lower $(m_i - n)$ rows are all zeros. Note that the square parts $[R_1]$ and $[R_2]$ are non-singular, with $\boldsymbol{\mu}([R_1], [R_2]) = \boldsymbol{\mu}(R_1, R_2) = \boldsymbol{\mu}(A_1, A_2)$. Thus $\boldsymbol{\mu}(\tilde{R}_1, \tilde{R}_2) \succeq \boldsymbol{r}(A_1, A_2)$. Invoking the proof for the square case in Section III, we may decompose $[R_1]$ and $[R_2]$ simultaneously as:

$$[R_1] = \tilde{U}_1 \tilde{T}_1 V^{\dagger}$$
$$[R_2] = \tilde{U}_2 \tilde{T}_2 V^{\dagger},$$

where $r(\tilde{T}_1, \tilde{T}_2) = r(A_1, A_2)$. Now, construct the augmented unitary matrices Y_i:

$$Y_i \triangleq \begin{pmatrix} \tilde{U}_i & 0 \\ 0 & I_{m_i-n} \end{pmatrix},$$

and the generalized triangular matrices T_i of dimensions $m_i \times n$:

$$T_i \triangleq \begin{pmatrix} \tilde{T}_i \\ 0 \end{pmatrix}.$$

Thus, we arrive at the desired decomposition of A_1 and A_2 (2), with $U_i \triangleq Q_i Y_i$.

APPENDIX B
PROOF OF LEMMA 1

Let $F_i = H_i \sqrt{C}$ for $i = 1, 2$. We first claim that $\boldsymbol{\mu}(F_1, F_2)$ must be mixed. This is true, since if C is non-singular then $\boldsymbol{\mu}(F_1, F_2) = \boldsymbol{\mu}(H_1, H_2)$, and if C is singular then at least one of the elements of $\boldsymbol{\mu}(F_1, F_2)$ equals one. It is left to show that if $\boldsymbol{\mu}(F_1, F_2)$ is mixed, then so is $\boldsymbol{\mu}(G_1, G_2)$. To that end, define the quadratic functions:

$$p(x) \triangleq \det\left(F_1^\dagger F_1 - x F_2^\dagger F_2\right),$$
$$q(x) \triangleq \det\left(G_1^\dagger G_1 - x G_2^\dagger G_2\right).$$

By Definition 6, the roots of $p(x)$ and $q(x)$ equal the square of the elements of $\boldsymbol{\mu}(F_1, F_2)$ and $\boldsymbol{\mu}(G_1, G_2)$, respectively. Thus it suffices to prove that if the roots of $p(x)$ are not on the same side of $x = 1$, then so are the roots of $q(x)$. By the positive semi-definitiveness of F_i and G_i, both functions are convex $\bigcup$ with $p(0), q(0), p(\infty), q(\infty) \geq 0$. By the assumption on the roots of $p(x)$, it must be that $p(1) \leq 0$. But since

$$G_1^\dagger G_1 - G_2^\dagger G_2 = F_1^\dagger F_1 - F_2^\dagger F_2$$

we have that $q(1) = p(1)$, and thus $q(1) \leq 0$. Finally, a convex $\bigcup$ continuous function which is non-negative at $x = 0$ and for $x \to \infty$ and non-positive at $x = 1$ cannot have both roots at the same side of 1.

REFERENCES

[1] Y. Jiang, W. Hager, and J. Li. The generalized triangular decompostion. *Mathematics of Computation*, vol. 77, no. 262:1037–1056, Apr. 2008.

[2] H. Weyl. Inequalities between two kinds of eigenvalues of a linear transformation. In *Proc. Nat. Acad. Sci. USA, 35*, pages 408–411, 1949.

[3] A. Horn. On the eigenvalues of a matrix with prescribed singular values. In *Proc. Amer. Math. Soc.*, pages 4–7, 1954.

[4] D. P. Palomar and Y. Jiang. MIMO transceiver design via majorization theory. *Found. Trends Commun. Inf. Theory*, 3:331–551, Nov. 2006.

[5] Y. Jiang, W. Hager, and J. Li. The geometric mean decomposition. *Linear Algebra and Its Applications*, vol. 396:373–384, Feb. 2005.

[6] C.-Y. Chen and P. P. Vaidyanathan. Precoded V-BLAST for ISI MIMO channels. In *Proceedings of the IEEE International Symposium on Circuits and Systems, ISCAS 2006.*, 2006.

[7] M. B. Shenouda and T. N. Davidson. A framework for designing MIMO systems with decision feedback equalization or Tomlinson-Harashima precoding. *IEEE Journal on Selected Areas in Communications*, 26(2):401–411, Feb. 2008.

[8] C.-C. Weng, C.-Y. Chen, and P. P. Vaidyanathan. Generalized triangular decomposition in transform coding. *IEEE Transactions on Signal Processing*, 58(2):566–574, Feb. 2010.

[9] C. F. Van Loan. Generalizing the singular value decomposition. *SIAM J. Numer.*, Anal. 13:76–83, 1976.

[10] 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 *ISSSE 1998, URSI International Symposium*, pages 295–300.

[11] M. P. Wilson, K. Narayanan, H. Pfister, and A. Sprintson. Joint physical layer coding and network coding for bi-directional relaying. *IEEE Trans. Info. Theory*, submitted.

[12] A. Khina, Y. Kochman, and U. Erez. Physical-layer MIMO relaying. In *Inter. Symp. on Inf. Theory, 2011, submitted.*

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

[14] B. Hassibi. An efficient square-root algorithm for BLAST. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 2, pages II737 –II740, 2000.

[15] E. Teletar. Capacity of the multiple antenna Gaussian channel. *Euro. Trans. Telecomm.*, 10:585–595, 1999.

[16] J. M. Cioffi and G. D. Forney Jr. Generalized Decision-Feedback Equalization for Packet Transmission with ISI and Gaussian Noise. In A. Paulraj, V. Roychowdhury, and C. Schaper, editors, *Communication, Computation, Control and Signal Processing*, pages 79–127. Kluwer, Boston, MA, 1997.

[17] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, 1983.

[18] W. Yu and J. Cioffi. Sum capacity of a Gaussian vector broadcast channel. *IEEE Trans. Info. Theory*, IT-50:1875–1892, Sept. 2004.

[19] G. Caire and S. Shamai (Shitz). On the achievable throughput of a multi-antenna Gaussian broadcast channel. *IEEE Trans. Info. Theory*, IT–49:1649–1706, July 2003.

[20] N. Jindal S. Vishwanath and A. Goldsmith. Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels. *IEEE Trans. Info. Theory*, 49:2658–2668, Oct. 2003.

[21] H. Weingarten, Y. Steinberg, and S. Shamai (Shitz). The capacity region of the Gaussian multiple-input multiple-output broadcast channel. *IEEE Trans. Info. Theory*, 52:3936–3964, Sept. 2006.

[22] H. Weingarten, Y. Steinberg, and S. Shamai (Shitz). On the capacity region of the multi-antenna broadcast channel with common messages. In *Proc. Int. Symp. Inform. Theory (ISIT), Seattle, USA*, pages 2195–2199, July 2006.

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

[24] R. L. Dobrushin. Optimal information transmission over a channel with unknown parameters (in Russian). Vol. 4, No. 12:1961–1956, Dec. 1959.

[25] J. Wolfowitz. Simultaneous channels. 4:371–386, 1960.

[26] P. Tejera, W. Utschick, G. Bauch, and J. A. Nossek. A novel decomposition technique for multiuser MIMO. In *International ITG/IEEE Workshop on Smart Antennas, Duisburg, Germany*, Apr. 2005.

[27] R. H. Gohary, T. N. Davidson, and Z.-Q Luo. An efficient design method for vector broadcast systems with common information. In *Proc. IEEE Globecom*, pages 2010–2014, Dec. 2003.

[28] M. J. Lopez. *Multiplexing, Scheduling, and Multicasting Strategies for Antenna Arrays in Wireless Networks*. PhD thesis, Massachusetts Institute of Technology, 2002.

[29] T. J. Goblick. Theoretical limitations on the transmission of data from analog sources. *IEEE Trans. Info. Theory*, IT-11:558–567, 1965.

[30] B. Chen and G. W. Wornell. Analog error-correcting codes based on chaotic dynamical systems. *IEEE Trans. Communications*, 46:881–890, July 1998.

[31] U. Mittal and N. Phamdo. Hybrid digital-analog (HDA) joint source-channel codes for broadcasting and robust communications. *IEEE Trans. Info. Theory*, IT-48:1082–1103, May 2002.

[32] V. M. Prabhakaran, R. Puri, and K. Ramchandran. A hybrid analog-digital framework for source-channel broadcast. *IEEE Trans. Info. Theory*, submitted.

[33] Y. Kochman and R. Zamir. Analog matching of colored sources to colored channels. *IEEE Trans. Info. Theory*, submitted.

[34] M. Taherzadeh and A. K. Khandani. Robust joint source-channel coding for delay-limited applications. In *ISIT-2007, Nice, France*, pages 726–730, 2007.

[35] C. E. Shannon. Coding theorems for a discrete source with a fidelity criterion. In *Institute of Radio Engineers, International Convention Record, Vol. 7*, pages 142–163, 1959.

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

[37] Z. Reznic, M. Feder, and R. Zamir. Distortion bounds for broadcasting with bandwidth expansion. *IEEE Trans. Info. Theory*, IT-52:3778–3788, Aug. 2006.