Modulation for MIMO Networks with Several Users

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

Ayal Hitron
Dept. of EE-Systems, TAU
Tel Aviv, Israel
Email: ayal@eng.tau.ac.il

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

***Abstract*—In a recent work, a capacity-achieving scheme for the common-message two-user MIMO broadcast channel, based on single-stream coding and decoding, was described. This was obtained via a novel joint unitary triangularization which is applied to the corresponding channel matrices. In this work, the triangularization is generalized, to *any* (finite) number of matrices, allowing multi-user applications. To that end, multiple channel uses are jointly treated, in a manner reminiscent of space-time coding. As opposed to the two-user case, in the general case there does not always exist a perfect (capacity-achieving) solution. However, a nearly optimal scheme (with vanishing loss in the limit of large blocks) always exists. Common-message broadcasting is but one example of communication networks with MIMO links which can be solved using an approach coined "Network Modulation"; the extension beyond two links carries over to these problems.**

I. Introduction

Multiple-input multiple-output (MIMO) Gaussian channels are a basic building block of many communication networks, due to their potential to enhance the throughput of communication systems, and have been extensively studied both in terms of the theoretical limits (see, e.g., [1]) as well as in terms of modulation and coding schemes that allow to approach these limits. In different communication scenarios, different assumptions on the channel behavior and of the availability of channel state information are appropriate (see [2], [3] and references therein).

A recent approach, coined "Network Modulation" [4], tackles the problem of conveying information over different multiple-antenna multi-terminal networks where full channel state information is available at all terminals (i.e., a fully closed-loop scenario). The approach is based on jointly triangularizing several matrices using the same unitary matrix on one side (joint encoder or decoder) and different unitary matrices on the other side (separate decoders or encoders), such that the diagonals of the resulting triangular matrices satisfy desirable properties, e.g., that they are equal. This decomposition, along with successive interference cancellation (SIC) or dirty paper coding (DPC) [5], transforms the channels into parallel scalar additive white Gaussian noise channels (AWGN). Thus, employing this scheme along with (any) scalar codes which are good for the AWGN channel, provides "practical" capacity-achieving schemes, for scenarios in which the capacity is known. Furthermore, somewhat surprisingly, it has been demonstrated that the approach allows to obtain new achievable rate regions to several information-theoretic problems, such as the two-way MIMO relay problem [6] and the problem of joint source-channel coding of a source over a MIMO broadcast channel [4].

A scenario of significant importance is that of sending a common message over a MIMO Gaussian broadcast (BC) channel, henceforth the *multicasting* scenario. The channel is given by

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

where $\mathbf{x}$ is the complex-valued channel input vector of length n subject to a power constraint, $\mathbf{y}_i$ $(i = 1, 2)$ is the output vector of user i of length m_i, H_i is the $m_i \times n$ complex channel matrix to user i, and $\mathbf{z}_i$ is an additive circularly-symmetric complex Gaussian noise vector of length m_i. Without loss of generality, we assume that both the noise elements and the input signal have unit power, i.e., $z_i \sim \mathcal{CN}(0, I_{m_i})$ and $\mathbb{E}\left[\mathbf{x}^\dagger \mathbf{x}\right] \leq 1$, where $\dagger$ denotes the conjugate transpose operator. It was shown in [4] that it is possible to jointly triangularize the channel matrices H_1 and H_2 using unitary matrices, such that the ratio between the resulting diagonals is constant. This in turn allows to achieve the common-message capacity using single-stream encoding and decoding of standard AWGN codes along with SIC (much like in V-BLAST transmission for a single user [7]). As we recall in the sequel, the problem of multicasting over several MIMO channels is tightly connected to the problems of universal coding over parallel channels, as well as rateless coding for Gaussian channels. Thus, the results we derive are relevant also to the latter problems.

The joint triangularization of [4] was limited to only two matrices, and hence, only two-user multicasting (or "perfect two-rate" in the rateless problem [8]) could be treated. The aim of the current work is to generalize the network modulation approach to more than two users. This is done by utilizing multiple uses of the channel, reminiscent of space-time coding techniques [9], [10].

II. Background: Network Modulation

In this section we recall the joint triangularization of two matrices [4], and its application to the two-user multicasting problem. We then demonstrate the relevance of the scheme to the special case of a two-rate scalar Gaussian rateless problem.

* This work was supported in part by the U.S. - Israel Binational Science Foundation under grant 2008/455.

A. Unitary Matrix Triangularization

The network modulation approach is based on several forms of matrix decompositions, one of which is the *geometric mean decomposition* (GMD) [11]. For simplicity, we will only consider the decomposition of *square* matrices throughout this work. As we show in the sequel, this does not pose any restriction on the communication problems addressed. The GMD [11] of an $n \times n$ matrix A is given by:

$$A = UTV^{\dagger}\,, \tag{2}$$

where U and V are $n \times n$ unitary matrices, and T is an upper-triangular $n \times n$ matrix such that all its diagonal values equal λ, where λ is a real-valued non-negative number.

Building on the GMD, the following decomposition, which will be referred to as Joint Equi-diagonal Triangularization (JET), was introduced in [4]. Let A_1 and A_2 be two complex matrices of dimensions $n \times n$ such that $\det(A) = \det(B)$. Then, the joint triangularization of A_1 and A_2 is given by:

$$\begin{aligned} A_1 &= U_1 R_1 V^{\dagger} \\ A_2 &= U_2 R_2 V^{\dagger}\,, \end{aligned} \tag{3}$$

where U_1, U_2, V are $n \times n$ unitary matrices, and R_1, R_2 are upper-triangular $n \times n$ matrices with *the same* real-valued, non-negative diagonal values, namely,

$$[R_1]_{ii} = [R_2]_{ii} \quad \forall i = 1, \ldots, n\,.$$

B. MIMO Multicast Scheme

We now recall how the JET decomposition can be used to obtain a practical scheme for transmitting a common message over a MIMO Gaussian BC with two receivers, as described by (1). Define the mutual information between a Gaussian input vector $\mathbf{x}$, having a covariance matrix $C_{\mathbf{x}} \triangleq \mathbb{E}\left[\mathbf{x}\mathbf{x}^{\dagger}\right]$, and the channel output $\mathbf{y}_i$, by

$$I(H_i, C_{\mathbf{x}}) \triangleq \log\det\left(I + H_i C_{\mathbf{x}} H_i^{\dagger}\right)\,. \tag{4}$$

The *common-message capacity* is given by the (worst-case) compound-channel capacity expression (see, e.g., [12]):

$$C = \max_{C_{\mathbf{x}}:\,\mathrm{tr}(C_{\mathbf{x}})\leq 1}\ \min_{i=1,2} I(H_i, C_{\mathbf{x}})\,. \tag{5}$$

Let $C_{\mathbf{x}}$ be an admissible covariance matrix, and assume for simplicity that $I(H_1, C_{\mathbf{x}}) = I(H_2, C_{\mathbf{x}}) = R$. The following scheme [4] achieves the rate R.

Define the following *augmented matrices*:

$$\tilde{G}_i \triangleq \begin{bmatrix} F_i \\ I_n \end{bmatrix}\,,$$

where $F_i \triangleq H_i\sqrt{C_{\mathbf{x}}}$ and I_n is the $n \times n$ identity matrix.

Next, the matrices $\tilde{G}_i$ are transformed into square matrices, by means of the QR decomposition:

$$\tilde{G}_i = Q_i G_i\,, \tag{6}$$

where Q_i is an $(m_i + n) \times n$ matrix with orthonormal columns and G_i is an $n \times n$ upper-triangular matrix with real-valued positive diagonal elements. Now, assuming that $I(H_1, C_{\mathbf{x}}) = I(H_2, C_{\mathbf{x}}) = R$, this implies [4, Proposition 1]:

$$\det(G_1) = \det(G_2) = 2^{\frac{R}{2}}\,.$$

Therefore, G_1, G_2 can be jointly triangularized using the JET:

$$G_i = U_i R_i V^{\dagger}\,, \qquad i = 1, 2\,, \tag{7}$$

where R_1 and R_2 are upper-triangular, having the same diagonal elements. The transmission scheme is as follows:

1) Construct n optimal codes for scalar AWGN channels. The k-th codebook is designed for a SISO AWGN channel with a rate $2\log r_k$, where r_k is the k-th diagonal element of R_1 (and also of R_2).
2) In each channel use, an n-length vector $\tilde{\mathbf{x}}$ is formed using one sample from each codebook. The transmitted vector $\mathbf{x}$ is then obtained using the following precoder:
$$\mathbf{x} = \sqrt{C_{\mathbf{x}}}V\tilde{\mathbf{x}}\,. \tag{8}$$
3) At the receiving ends, the i-th user calculates
$$\tilde{\mathbf{y}}_i = U_i^{\dagger}\tilde{Q}_i^{\dagger}\mathbf{y}_i\,, \tag{9}$$
where $\tilde{Q}_i$ consists of the first n rows of Q_i.
4) Finally, the codebooks are decoded using SIC, starting from the n-th codeword and ending with the first one: The n-th codeword is decoded first, using the n-th element of $\tilde{\mathbf{y}}_i$, treating the other codewords as AWGN. The effect of the n-th element of $\tilde{\mathbf{x}}$ is then subtracted out from the remaining elements of $\tilde{\mathbf{y}}$. Next, the $(n-1)$-th codeword is decoded, using the $(n-1)$-th element of $\tilde{\mathbf{y}}_i$ - and so forth.

The optimality of this scheme was proved in [4, Sec. IV].

Example 1 (Application to the two-rate rateless problem): Consider the scalar Gaussian rateless problem defined in [13]:

$$y_m = \alpha x_m + z_m\,, \qquad m = 1, 2, \ldots, M\,.$$

The gain α is known only to the receiver, and can take one of M possible values, such that a gain of α_m implies that the message should be decodable using m received blocks: [1]

$$R = m\log(1 + |\alpha_m|^2)\,, \qquad m = 1, 2, \ldots, M\,.$$

Specializing the problem to the case of one (possible) incremental redundancy block ($M = 2$), the perfect two-rate rateless problem can be viewed as a 2-user MIMO-BC channel with channel matrices

$$H_1 = \begin{pmatrix} \alpha_1 & 0 \\ 0 & 0 \end{pmatrix}, \quad H_2 = \begin{pmatrix} \alpha_2 & 0 \\ 0 & \alpha_2 \end{pmatrix}.$$

Applying the scheme of Section II-B yields the following precoding matrix [8]:

$$V = \sqrt{\frac{1}{2^{R/2}+1}}\begin{pmatrix} 1 & 2^{R/4} \\ 2^{R/4} & -1 \end{pmatrix}\,,$$

which coincides with the result in [13, Section III].

[1] Alternatively, this can be viewed as a scheme that works for every value of α, but designed to be optimal only for M specific values.

Erez, Trott and Wornell [13] also treated the case of $M = L = 3$, and found a condition for which a "perfect" scheme exists. In the sequel we will shed light on this condition.

III. Joint Triangularization of Many Matrices

In this section we extend the network modulation technique to a any finite number of users, using a *recursion principle*. Specifically, given K matrices $G_1, \ldots, G_K$, we wish to find K matrices with orthonormal columns $U_1, \ldots, U_k$, and another such matrix V, such that the matrices $R_i \triangleq U_i^\dagger G_i V$ are upper-triangular, having *equal* diagonals. We shall refer to this as K-matrix JET, or simply K-JET.

The proof of the existence of a JET decomposition of two matrices G_1 and G_2 [4] is based upon applying the GMD (2) to the single matrix $G_1 G_2^{-1}$. Similarly, we show in the following lemma that $(K+1)$-JET is equivalent to simultaneous GMD of K matrices, which will be referred to as K-GMD.

Lemma 1: Let $G_1, \ldots, G_{K+1}$ be $n \times n$ complex valued matrices with equal determinants, and define the K matrices:

$$A_i = G_i G_{K+1}^{-1}\,, \qquad i = 1, \ldots, K\,. \tag{10}$$

Then, there exist $K+1$ matrices with orthonormal columns $U_1, \ldots, U_{K+1}$, of dimensions $n \times m$, such that

$$U_i^\dagger A_i U_{K+1} = T_i\,, \qquad i = 1, \ldots, K\,, \tag{11}$$

where $\{T_i\}$ are upper-triangular with all diagonal entries equal to 1, *if and only if* there exists an $n \times m$ matrix V with orthonormal columns, such that

$$U_i^\dagger G_i V = R_i\,, \qquad i = 1, \ldots, K+1\,,$$

where $\{R_i\}$ are upper-triangular with equal diagonals.

Proof: See a constructive proof in Appendix A. ∎

Remark 1: Constructing matrices with constant diagonals could be advantageous in practice, as this corresponds to equal gains of all the resulting sub-channels, and hence enables to use the *same* (single) codebook over all of them.

We are thus left with the task of performing K-GMD to K matrices. In Section IV we state sufficient and necessary conditions for the existence of the above decomposition for the special case of two real-valued 2×2 matrices. We will then, in Section V, present a different approach, involving joint triangularization of block-diagonal matrices, which enables a nearly-optimal network-modulation scheme, even when exact triangularization is not possible.

IV. Exact Triangularization with Constant Diagonals of Two Real-Valued 2×2 Matrices

We now provide a necessary and sufficient condition for the existence of 2-GMD for *real-valued* 2×2 matrices.

Theorem 1 (2-GMD for 2×2 real-valued matrices): Let A_1 and A_2 be *real-valued* 2×2 matrices with determinants equal to 1. Apply (any) JET decomposition to them: [2]

$$A_i = U_i^{\mathrm{JET}} R_i \left(V^{\mathrm{JET}}\right)^\dagger, \qquad i = 1, 2\,, \tag{12}$$

[2]The JET decomposition is, in general, not unique.

where:

$$R_i = \begin{pmatrix} r_1 & x_i \\ 0 & r_2 \end{pmatrix}.$$

Then, there exist three complex-valued 2×2 unitary matrices $U_1^{\mathrm{GMD}}, U_2^{\mathrm{GMD}}, V^{\mathrm{GMD}}$ such that:

$$\left(U_i^{\mathrm{GMD}}\right)^\dagger A_i V^{\mathrm{GMD}} = \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}$$

if and only if the following inequality is satisfied:

$$r_2 \left(\frac{x_1 + x_2}{2}\right)^2 \leq r_2 + \frac{x_1 x_2}{r_1 - r_2}. \tag{13}$$

Without loss of generality, we can assume that the solution is of the form:

$$V^{\mathrm{GMD}} = \begin{pmatrix} s_1 & s_2 \\ s_2^* & -s_1^* \end{pmatrix}. \tag{14}$$

Proof: The proof is straightforward, and is given in Appendix C. ∎

Remark 2: Although this theorem is valid only for *real-valued* matrices A_1 and A_2, the resulting unitary matrices U_1, U_2, and V are, in general, *complex-valued*. In Section V-A we bring a restatement of the theorem, which involves only real-valued orthogonal transformations.

Remark 3: This theorem can be applied to the three-rate rateless problem defined in Section II-B. This yields a condition for the existence of a perfect scheme, namely, $R \leq 6 \log\left(\frac{3+\sqrt{5}}{2}\right) \approx 8.331$, as in [13]. The details are given in Appendix D.

V. Space-Time Triangularization

As indicated by Theorem 1, joint triangularization with constant diagonal values is not always possible. However, even when the condition for joint triangularization does not hold, we can still perform nearly-optimal network modulation, by utilizing multiple uses of the same channel realization. The idea of mixing the same symbols between multiple channel uses has much in common with Space-Time Codes [9], [10].

A. Restatement of Theorem 1

In order to introduce the space-time like structure, we start by a restatement of Theorem 1.

Recall the two-user common-message broadcast MIMO channel (1) with two transmit antennas ($n = 2$), and a general number of antennas m_i at each receiver. We now utilize transmission in two consecutive time instances (as in [9]). This is equivalent to sending extended symbols over the following *extended channel*:

$$\mathbf{Y}_i = \mathcal{H}_i \mathbf{X} + \mathbf{Z}_i\,, \qquad i = 1, 2\,.$$

The extended vectors $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ are composed of two "physical" input, output, and noise vectors, respectively, and $\mathcal{H}_i$ is the $(2m_i) \times 4$ *extended channel matrix* defined as ($i = 1, 2$)

$$\mathcal{H}_i = \lceil H_i \rfloor_{\otimes 2} \tag{15}$$

where $\lceil A \rfloor_{\otimes N}$ denotes the Kronecker product $I_N \otimes A$, viz. a block-diagonal matrix with N blocks of A on its diagonal:

$$\lceil A \rfloor_{\otimes N} \triangleq \begin{pmatrix} A & 0 & \cdots & 0 \\ 0 & A & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A \end{pmatrix} .$$

The power constraint now becomes $\mathbb{E}\left[\mathbf{X}^T \mathbf{X}\right] \leq 2$.

Let $C_{\mathbf{x}}$ be a covariance matrix satisfying $\operatorname{trace}(C_{\mathbf{x}}) \leq 1$, and define the augmented matrices G_i as in (6). Following Lemma 1, we define the two 2×2 matrices:

$$A_1 \triangleq G_1 G_3^{-1} \,,\; A_2 \triangleq G_2 G_3^{-1} \,.$$

Also define the following 4×4 extended matrices ($i = 1, 2$):

$$\mathcal{G}_i \triangleq \lceil G_i \rfloor_{\otimes 2} \,,\; \mathcal{A}_i \triangleq \lceil A_i \rfloor_{\otimes 2} \,. \tag{16}$$

Since the matrices A_1 and A_2 are real-valued matrices, we can obtain 2-GMD of the matrices $\mathcal{A}_1$ and $\mathcal{A}_2$ under the same conditions as in Theorem 1, such that all the involved unitary transformations become *real-valued*. Following Lemma 1, this yields a 3-JET of the three matrices $\mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_3$:

$$\mathcal{G}_i = \mathcal{U}_i \mathcal{R}_i \mathcal{V}^\dagger \,,$$

where $\mathcal{R}_i$ are upper triangular with equals diagonals.

In particular, the complex precoding matrix V^{GMD} given by (14) implies the following (real) orthogonal space-time block code structure of $\mathcal{V}^{\mathrm{GMD}}$ [14]:

$$\mathcal{V}^{\mathrm{GMD}} = \begin{pmatrix} a & -c & b & d \\ b & d & -a & c \\ c & a & d & -b \\ -d & b & c & a \end{pmatrix} .$$

The same scheme as in Section II-B can now be employed, such that the two channel uses are effectively transformed into four scalar AWGN channels, having the same capacities for all three users. Note that the matrix $\tilde{Q}_i$ of (9) is replaced with its extended version, $\left\lceil \tilde{Q}_i \right\rfloor_{\otimes 2}$.

B. *Nearly-Optimal* 2-*GMD*

We now show how to utilize a space-time structure in order to obtain nearly-optimal joint triangularization of two matrices, such that the resulting triangular matrices have a constant diagonal. This method will later be generalized to any number of matrices, using Lemma 1. The resulting scheme becomes asymptotically optimal for large values of N, where N is the number of channel uses assembled together for the purpose of joint decomposition. Note that the proposed scheme is nearly optimal for *any* two complex-valued channels H_i (and not restricted to real-valued matrices, in contrast to the perfect construction of Theorem 1).

Theorem 2 (Nearly-Optimal 2*-GMD):* Let A_1 and A_2 be two complex-valued $n \times n$ matrices, and define the following $nN \times nN$ extended matrices:

$$\mathcal{A}_i = \lceil A_i \rfloor_{\otimes N} \,, \quad i = 1, 2, \,. \tag{17}$$

Then there exist three $nN \times n\big(N-(n-1)\big)$ matrices $\mathcal{U}_1, \mathcal{U}_2, \mathcal{V}$ with orthonormal columns, such that:

$$\mathcal{U}_i^\dagger \mathcal{A}_i \mathcal{V} = \begin{pmatrix} 1 & * & \cdots & * & * \\ 0 & 1 & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & * \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix} , \qquad i = 1, 2 \,.$$

By using this decomposition together with Lemma 1, the same scheme as in Section II-B can be employed, such that the N channel uses are effectively transformed into $n(N-n+1)$ scalar AWGN channels. The sum of the capacities of these channels tends to the capacity of the original channel for large values of N, where the only loss comes from edge effects (truncating the extreme $n(n-1)$ elements).

The full proof of the theorem is given in Appendix E. The main idea of the proof is demonstrated by the proof for the 2×2 case, presented next.

Proof of Theorem 2 for $n = 2$: We start by jointly triangularizing the matrices A_1 and A_2:

$$\left(U_i^{\mathrm{JET}}\right)^\dagger A_i V^{\mathrm{JET}} = \begin{pmatrix} r_1 & x_i \\ 0 & r_2 \end{pmatrix} \tag{18}$$

where $r_1 r_2 = 1$. We now apply the decomposition (18) to each block separately, using:

$$\mathcal{U}_i^{\mathrm{JET}} = \lceil U_i^{\mathrm{JET}} \rfloor_{\otimes N} \,,\; \mathcal{V}^{\mathrm{JET}} = \lceil V^{\mathrm{JET}} \rfloor_{\otimes N} \,,$$

which yields the matrices

$$\left(\mathcal{U}_i^{\mathrm{JET}}\right)^\dagger \mathcal{A}_i \mathcal{V}^{\mathrm{JET}} = \begin{pmatrix} r_1 & x_i & 0 & 0 & \cdots & 0 & 0 \\ 0 & r_2 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & r_1 & x_i & \cdots & 0 & 0 \\ 0 & 0 & 0 & r_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & r_1 & x_i \\ 0 & 0 & 0 & 0 & \cdots & 0 & r_2 \end{pmatrix} . \tag{19}$$

Note that the sub-matrix

$$\Lambda = \begin{pmatrix} r_2 & 0 \\ 0 & r_1 \end{pmatrix}$$

does not depend on i, and therefore it can be decomposed using the GMD (2), $\Lambda = U^{\mathrm{GMD}} T \big(V^{\mathrm{GMD}}\big)^\dagger$, where T is upper-triangular with only 1s on the diagonal. We use this decomposition to construct a second transformation – only this time it is not be applied on each block separately, but rather "mixes" pairs of consecutive blocks, using:

$$\mathcal{U}^{\mathrm{GMD}} = \begin{pmatrix} 0 \;\; 0 \;\; \cdots \;\; 0 \\ \lceil U^{\mathrm{GMD}} \rfloor_{\otimes (N-1)} \\ 0 \;\; 0 \;\; \cdots \;\; 0 \end{pmatrix} , \mathcal{V}^{\mathrm{GMD}} = \begin{pmatrix} 0 \;\; 0 \;\; \cdots \;\; 0 \\ \lceil V^{\mathrm{GMD}} \rfloor_{\otimes (N-1)} \\ 0 \;\; 0 \;\; \cdots \;\; 0 \end{pmatrix} .$$

Applying this transformation to (19) yields the following $(2N-2)\times(2N-2)$ upper-triangular matrix:

$$\mathcal{U}_i^\dagger \mathcal{A}_i \mathcal{V} = \begin{pmatrix} 1 & * & \cdots & * & * \\ 0 & 1 & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & * \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix},$$

where $\mathcal{U}_i \triangleq \mathcal{U}_i^{\mathrm{JET}}\mathcal{U}^{\mathrm{GMD}}$ and $\mathcal{V} \triangleq \mathcal{V}^{\mathrm{JET}}\mathcal{V}^{\mathrm{GMD}}$. ■

C. Nearly-Optimal K-GMD

By using Lemma 1, we can generalize Theorem 2 to any number of users, as follows:

Theorem 3 (Nearly-Optimal K-GMD): Let $A_1,\dots,A_K$ be K complex-valued $n\times n$ matrices with determinants equal to 1, and define $\mathcal{A}_1,\dots,\mathcal{A}_K$ as in (17). Then there exist $K+1$ matrices $\mathcal{U}_1,\dots,\mathcal{U}_K,\mathcal{V}$, with orthonormal columns, such that:

$$\mathcal{U}_i^\dagger \mathcal{A}_i \mathcal{V} = \begin{pmatrix} 1 & * & \cdots & * & * \\ 0 & 1 & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & * \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}, \qquad i=1,\dots,K\,.$$

Proof: A sketch of the proof is given in Appendix B. ■

VI. Discussion

Theorem 1 provides sufficient and necessary conditions for joint GMD of two *real-valued* 2×2 matrices. This naturally raises the question of how this condition can be carried over to the complex-valued case, and to general dimensions $n\times n$.

Furthermore, we demonstrated that (exact) K-GMD, not using any space-time structure, is not always possible. Nevertheless, *nearly*-optimal communication schemes can always be constructed, which become optimal in the limit of large N. It remains an open question whether an exact triangularization can be obtained using only a finite number of channel uses.

Acknowledgements

The authors would like to thank Yuval Kochman for constant help throughout this work.

Appendix A
Proof of Lemma 1

Proof of Lemma 1: The direct part holds trivially. We are therefore left with the task of proving the converse part. We start with the QR decomposition $G_{K+1}^{-1}U_{K+1}=VS$, where V is of dimensions $n\times m$ with orthonormal columns, and S is an $m\times m$ upper-triangular matrix. Thus, using (10) and (11), we obtain the following equalities:

$$\begin{aligned} U_i^\dagger G_i V S &= T_i\,, \qquad i=1,\dots,K \\ U_{K+1}^\dagger G_{K+1} V S &= I\,. \end{aligned}$$

Multiplying by S^{-1} on the right yields:

$$\begin{aligned} U_i^\dagger G_i V &= T_i S^{-1}\,, \qquad i=1,\dots,K \\ U_{K+1}^\dagger G_{K+1} V &= S^{-1}\,. \end{aligned}$$

Since T_i are upper-triangular with only 1s on the diagonal, the matrices $R_i \triangleq T_i S^{-1}$ $(i=1,\dots,K)$ and $R_{K+1}\triangleq S^{-1}$ have equal diagonals, which completes the proof. ■

Appendix B
Sketch of Proof of Theorem 3

Proof Idea: The theorem has already been proved for the special case of $K=2$. For larger values of K we prove by induction, applying repeatedly Lemma 1 and of Theorem 2:

1) According to Lemma 1, performing K-GMD is equivalent to $(K+1)$-JET. We can thus transform K upper-triangular matrices with *constant* diagonal values into $K+1$ upper-triangular matrices of the same size, $R_1,\dots,R_{K+1}$ with *equal* diagonals.
2) Given the matrices $R_1,\dots,R_{K+1}$, construct the block-diagonal *extended* matrices $\mathcal{R}_i$, as in (17). Using the technique of Theorem 2, we construct matrices with orthonormal columns, $\mathcal{U}_1^{(K+1)},\dots,\mathcal{U}_{K+1}^{(K+1)},\mathcal{V}^{(K+1)}$, such that the matrices $\left(\mathcal{U}_i^{(K+1)}\right)^\dagger \mathcal{R}_i\mathcal{V}^{(K+1)}$ are upper-triangular, with constant diagonals. Finally, the loss in rate could be made arbitrarily small by taking N to be sufficiently large.

Appendix C
Condition for 2-GMD of Real-Valued 2×2 Matrices

We now prove the necessary and sufficient condition for the existence of joint-triangularization of two 2×2 real-valued matrices.

Proof of Theorem 1: Let A_1 and A_2 be real-valued 2×2 matrices with determinants equal to 1. Apply the JET decomposition to these matrices, to obtain

$$A_i = U_i^{\mathrm{JET}} R_i \left(V^{\mathrm{JET}}\right)^\dagger, \qquad i=1,2, \tag{20}$$

where:

$$R_i = \begin{pmatrix} r_1 & x_i \\ 0 & r_2 \end{pmatrix}$$

such that $r_1r_2=1$. The matrices $U_i^{\mathrm{JET}},V^{\mathrm{JET}}$ are real-valued unitary matrices, and we can assume without loss of generality that $\det\left(V^{\mathrm{JET}}\right)=1$.

If there exist three complex-valued 2×2 unitary matrices U_1,U_2,V such that:

$$U_i^\dagger A_i V = \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}, \tag{21}$$

then according to (20),

$$\left(U_i^{\mathrm{GMD}}\right)^\dagger R_i V^{\mathrm{GMD}} = \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix},$$

where

$$\begin{aligned} U_i^{\mathrm{GMD}} &= \left(U_i^{\mathrm{JET}}\right)^\dagger U_i \\ V^{\mathrm{GMD}} &= \left(V^{\mathrm{JET}}\right)^\dagger V\,. \end{aligned}$$

Denote the entries of the first column of V^{GMD} by s_1 and s_2^*, i.e.,

$$V^{\mathrm{GMD}} = \begin{pmatrix} s_1 & * \\ s_2^* & * \end{pmatrix} ,$$

where $|s_1|^2 + |s_2|^2 = 1$. The first column of $R_1 V^{\mathrm{GMD}}$ and of $R_2 V^{\mathrm{GMD}}$ is therefore:

$$R_i V^{\mathrm{GMD}} = \begin{pmatrix} r_1 s_1 + x_i s_2^* & * \\ r_2 s_2^* & * \end{pmatrix} , \qquad i = 1, 2 ,$$

where x_1 and x_2 denote the off-diagonal elements of R_1 and R_2 respectively. These two columns must have a norm of 1, namely:

$$|r_1 s_1 + x_i s_2^*|^2 + |r_2 s_2^*|^2 = 1 , \qquad i = 1, 2 .$$

Since r_1, r_2, x are real-valued, s_1 and s_2 must satisfy the following three equations:

$$\begin{aligned} |s_1|^2 + |s_2|^2 &= 1 \\ r_1^2|s_1|^2 + (x_1^2 + r_2^2)|s_2|^2 + 2r_1x_1\mathrm{Re}(s_1 s_2) &= 1 \\ r_1^2|s_1|^2 + (x_2^2 + r_2^2)|s_2|^2 + 2r_1x_2\mathrm{Re}(s_1 s_2) &= 1 . \end{aligned}$$

Denoting $\alpha \equiv \frac{s_1}{s_2^*}$, and substituting α in these equations, results in:

$$1 + |\alpha|^2 = \frac{1}{|s_2|^2} \qquad (22)$$

$$r_1^2|\alpha|^2 + (x_1^2 + r_2^2) + 2r_1x_1\mathrm{Re}(\alpha) = \frac{1}{|s_2|^2} \qquad (23)$$

$$r_1^2|\alpha|^2 + (x_2^2 + r_2^2) + 2r_1x_2\mathrm{Re}(\alpha) = \frac{1}{|s_2|^2} . \qquad (24)$$

Subtracting (24) from (23) yields:

$$(x_1 + x_2) + 2r_1\mathrm{Re}(\alpha) = 0 ,$$

So we have:

$$|\alpha|^2 = \frac{1}{|s_2|^2} - 1 \qquad (25)$$

$$\mathrm{Re}(\alpha) = -\left(\frac{x_1 + x_2}{2}\right) r_2 \qquad (26)$$

$$(\mathrm{Re}(\alpha))^2 = \left(\frac{x_1 + x_2}{2}\right)^2 r_2^2 \qquad (27)$$

$$(\mathrm{Im}(\alpha))^2 = \frac{1}{|s_2|^2} - 1 - \left(\frac{x_1 + x_2}{2}\right)^2 r_2^2 . \qquad (28)$$

Thus, equation (23) and (24) become:

$$|s_2|^2 = \frac{r_1^2 - 1}{r_1^2 - r_2^2 + x_1x_2} ,$$

and therefore equation (27) becomes

$$(\mathrm{Im}(\alpha))^2 = \frac{r_1^2 - r_2^2 + x_1x_2}{r_1^2 - 1} - 1 - \left(\frac{x_1 + x_2}{2}\right)^2 r_2^2 .$$

Thus, the following conditions are necessary and sufficient for the existence of a solution:

$$\begin{aligned} \frac{r_1^2 - 1}{r_1^2 - r_2^2 + x_1x_2} &\geq 0 \\ \frac{r_1^2 - r_2^2 + x_1x_2}{r_1^2 - 1} - 1 - \left(\frac{x_1 + x_2}{2}\right)^2 r_2^2 &\geq 0 , \end{aligned}$$

which are equivalent to

$$r_2 \left(\frac{x_1 + x_2}{2}\right)^2 \leq r_2 + \frac{x_1x_2}{r_1 - r_2} . \qquad (29)$$

This proves that (29) is a *necessary* condition for the existence of the decomposition (21).

Now, assume that this condition holds, and define the matrix:

$$V^{\mathrm{GMD}} = \begin{pmatrix} s_1 & s_2 \\ s_2^* & -s_1^* \end{pmatrix} .$$

We now apply the QR decomposition to the matrices $R_1 V^{\mathrm{GMD}}$ and $R_2 V^{\mathrm{GMD}}$:

$$\left(U_i^{\mathrm{GMD}}\right)^\dagger R_i V^{\mathrm{GMD}} = \begin{pmatrix} a_i & b_i \\ 0 & c_i \end{pmatrix} . \qquad (30)$$

The first columns of both $R_1 V^{\mathrm{GMD}}$ and $R_2 V^{\mathrm{GMD}}$ have norms equal to 1. Therefore, from the construction of the QR decomposition, it follows that $a_1 = a_2 = 1$. Consequently, since both matrices have a unit determinant, $c_1 = c_2 = 1$ must hold as well. Thus, (30) becomes:

$$\left(U_i^{\mathrm{GMD}}\right)^\dagger R_i V^{\mathrm{GMD}} = \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} ,$$

and therefore,

$$A_i = U_i \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} V^\dagger ,$$

where

$$\begin{aligned} U_i &= U_i^{\mathrm{JET}} U_i^{\mathrm{GMD}} \\ V &= V^{\mathrm{JET}} V^{\mathrm{GMD}} . \end{aligned}$$

Furthermore, since the matrix V^{GMD} is of the form

$$V^{\mathrm{GMD}} = \begin{pmatrix} s_1 & s_2 \\ s_2^* & -s_1^* \end{pmatrix} ,$$

and V^{JET} is a real-valued unitary matrix with unit determinant, it is easy to see that the matrix V is also of the form

$$V = \begin{pmatrix} s_1 & s_2 \\ s_2^* & -s_1^* \end{pmatrix} ,$$

which completes the proof of the theorem.

■

Appendix D
Three-Rate Rateless

We now consider the three-rate "rateless" problem, as defined in Section II-B, with $M = L = 3$ and a given rate R:

$$\begin{aligned} \mathcal{H}_1 &= \begin{pmatrix} \alpha_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ \mathcal{H}_2 &= \begin{pmatrix} \alpha_2 & 0 & 0 \\ 0 & \alpha_2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ \mathcal{H}_3 &= \begin{pmatrix} \alpha_3 & 0 & 0 \\ 0 & \alpha_3 & 0 \\ 0 & 0 & \alpha_3 \end{pmatrix} , \end{aligned}$$

where $\alpha_1, \alpha_2, \alpha_3$ are the positive values satisfying $\log(1+\alpha_1^2) = 2\log(1+\alpha_2^2) = 3\log(1+\alpha_3^2) = R$. As in the 2-rate case, the covariance matrix in this problem is the identity matrix, $C_{\mathbf{x}} = I$. Since $\mathcal{H}_3$ is a s scaled identity matrix, we can ignore it and concentrate on the remaining two matrices.

The augmented matrices, as defined in (6), are:

$$\begin{aligned} \mathcal{G}_1 &= \begin{pmatrix} 2^{\frac{R}{2}} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ \mathcal{G}_2 &= \begin{pmatrix} 2^{\frac{R}{4}} & 0 & 0 \\ 0 & 2^{\frac{R}{4}} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

The decomposition (12) becomes:

$$\begin{aligned} \mathcal{R}_1 &= 2^{\frac{R}{2}} \cdot \begin{pmatrix} 1 & z & w \\ 0 & 2^{-\frac{R}{12}} & x \\ 0 & 0 & 2^{\frac{R}{12}} \end{pmatrix} \\ \mathcal{R}_2 &= 2^{\frac{R}{2}} \cdot \begin{pmatrix} 1 & z & 0 \\ 0 & 2^{-\frac{R}{12}} & 0 \\ 0 & 0 & 2^{\frac{R}{12}} \end{pmatrix}, \end{aligned}$$

where

$$x = -\left(1 - 2^{-\frac{R}{6}}\right)\sqrt{1 + 2^{\frac{R}{6}} + 2^{\frac{R}{3}}}.$$

It then follows from Theorem 1 that there exists a perfect solution over the complex field if and only if:

$$x^2 - 4 \leq 0,$$

or explicitly:

$$2^{-\frac{R}{3}}\left(1 + 2^{\frac{R}{6}}\right)^2\left(1 - 3 \cdot 2^{\frac{R}{6}} + 2^{\frac{R}{3}}\right) \leq 0.$$

This condition is satisfied if and only if:

$$R \leq 6\log\left(\frac{3+\sqrt{5}}{2}\right) \approx 8.331,$$

which coincides with the result that was obtained in [13].

For rates higher than this threshold, a perfect capacity-achieving solution does not exist. However, as explained earlier, multiple channel usages can be utilized in order to approach capacity asymptotically.

Appendix E
Nearly Optimal 2-GMD for $n \geq 2$

We now bring the proof of Theorem 2 for the general case $n \geq 2$.

Proof of Theorem 2:

Let A_1 and A_2 be the two $n \times n$ complex valued matrices, with determinants equal to 1. As in (16), we define the extended matrices,

$$\mathcal{A}_i = \begin{pmatrix} A_i & 0 & \cdots & 0 \\ 0 & A_i & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_i \end{pmatrix}.$$

We are looking for three $nN \times n(N-n+1)$ matrices $\mathcal{U}_1, \mathcal{U}_2, \mathcal{V}$ with orthonormal columns, such that

$$\mathcal{U}_i^\dagger \mathcal{A}_i \mathcal{V}$$

are upper-triangular, with only 1s on the diagonal.

We accomplish that using three steps:

a) Joint Triangularization: As in the $n = 2$ proof, we start by jointly triangularizing the matrices A_1 and A_2:

$$\left(U_i^{\mathrm{JET}}\right)^\dagger A_i V^{\mathrm{JET}} = \begin{pmatrix} r_1 & * & \cdots & * \\ 0 & r_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_n \end{pmatrix}.$$

We now apply this transformation to each block separately:

$$\mathcal{A}_i^{\mathrm{JET}} = \left(\mathcal{U}_i^{\mathrm{JET}}\right)^\dagger \mathcal{A}_i \mathcal{V}^{\mathrm{JET}}, \tag{31}$$

where

$$\begin{aligned} \mathcal{U}_i^{\mathrm{JET}} &= \begin{pmatrix} U_i^{\mathrm{JET}} & 0 & \cdots & 0 \\ 0 & U_i^{\mathrm{JET}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & U_i^{\mathrm{JET}} \end{pmatrix} \\ \mathcal{V}^{\mathrm{JET}} &= \begin{pmatrix} V^{\mathrm{JET}} & 0 & \cdots & 0 \\ 0 & V^{\mathrm{JET}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & V^{\mathrm{JET}} \end{pmatrix}. \end{aligned}$$

b) Reordering: It will now be convenient to re-order the columns of $\mathcal{A}_i^{\mathrm{JET}}$ such that the following columns:

$$kn, kn + (n-1), kn + 2(n-1), \cdots, kn + (n-1)^2$$

will become "grouped together" for every k.[3] Formally, We do so by introducing the following $nN \times n(N-n+1)$ re-ordering matrix $\mathcal{O}$:

$$\mathcal{O}_{ij} = \begin{cases} 1 & i = \pi_j \\ 0 & \text{Otherwise,} \end{cases} \tag{32}$$

where the function π is defined as follows:

- For $1 \leq j \leq n(N-n+1)$,

$$\pi_j = (n-1) \cdot [(j-1)\%n] + n \cdot \left\lfloor \frac{j-1}{n} \right\rfloor + n.$$

As a result of this re-ordering, we obtain an upper-triangular $(N-n+1)n \times (N-n+1)n$ matrix, which has a block-triangular structure:

$$\mathcal{O}^T \mathcal{A}_i^{\mathrm{JET}} \mathcal{O} = \left(\begin{array}{c:c:c:c:c} \Lambda & * & \cdots & * & * \\ \hdashline 0 & \Lambda & \cdots & * & * \\ \hdashline \vdots & \vdots & \vdots & \ddots & \vdots \\ \hdashline 0 & 0 & \cdots & \Lambda & * \\ \hdashline 0 & 0 & \cdots & 0 & \Lambda \end{array}\right),$$

[3]Note that this set includes exactly one symbol from each of n consecutive channel uses.

where

$$\Lambda = \begin{pmatrix} r_n & 0 & \cdots & 0 \\ 0 & r_{n-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_1 \end{pmatrix}.$$

c) GMD: Since the matrix Λ does not depend on i, we can decompose it using GMD:

$$\Lambda = U^{\mathrm{GMD}} T \big(V^{\mathrm{GMD}}\big)^{\dagger},$$

where T is upper-triangular with only 1s on its diagonal.

We now apply the GMD to all the blocks of $\mathcal{O}^T \mathcal{A}_i^{\mathrm{JET}} \mathcal{O}$:

$$\begin{aligned} \mathcal{U}^{\mathrm{GMD}} &= \begin{pmatrix} U^{\mathrm{GMD}} & 0 & \cdots & 0 \\ 0 & U^{\mathrm{GMD}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & U^{\mathrm{GMD}} \end{pmatrix} \\ \mathcal{V}^{\mathrm{GMD}} &= \begin{pmatrix} V^{\mathrm{GMD}} & 0 & \cdots & 0 \\ 0 & V^{\mathrm{GMD}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & V^{\mathrm{GMD}} \end{pmatrix}, \end{aligned}$$

to obtain:

$$\big(\mathcal{U}_i^{\mathrm{GMD}}\big)^{\dagger} \mathcal{O}^T \mathcal{G}_i^{\mathrm{JET}} \mathcal{O} \mathcal{V}^{\mathrm{GMD}} = \left(\begin{array}{c|c|c|c|c} T_i & * & \cdots & * & * \\ \hline 0 & T_i & \cdots & * & * \\ \hline \vdots & \vdots & \vdots & \ddots & \vdots \\ \hline 0 & 0 & \cdots & T_i & * \\ \hline 0 & 0 & \cdots & 0 & T_i \end{array}\right), \tag{33}$$

where T_i are upper-triangular with only 1s on the diagonal.

We conclude by combining (33) with (31) to obtain:

$$\mathcal{U}_i^{\dagger} \mathcal{A}_i \mathcal{V} = \left(\begin{array}{c|c|c|c|c} T_i & * & \cdots & * & * \\ \hline 0 & T_i & \cdots & * & * \\ \hline \vdots & \vdots & \vdots & \ddots & \vdots \\ \hline 0 & 0 & \cdots & T_i & * \\ \hline 0 & 0 & \cdots & 0 & T_i \end{array}\right),$$

where

$$\mathcal{U}_i = \mathcal{U}_i^{\mathrm{JET}} \mathcal{O} \mathcal{U}^{\mathrm{GMD}} \tag{34}$$

$$\mathcal{V} = \mathcal{V}^{\mathrm{JET}} \mathcal{O} \mathcal{V}^{\mathrm{GMD}}, \tag{35}$$

which completes the proof of the theorem.

■

REFERENCES

[1] E. Telatar, "Capacity of the multiple antenna Gaussian channel," *Europ. Trans. Telecommun.*, vol. 10, pp. 585–595, Nov. 1999.

[2] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bolcseki, "An overview of MIMO communications - a key to gigabit wireless," *Proceedings of the IEEE*, vol. 92, no. 2, pp. 198 – 218, Feb. 2004.

[3] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications*. Cambridge University Press, 2005.

[4] A. Khina, Y. Kochman, and U. Erez, "Joint unitary triangularization for MIMO networks," *IEEE Trans. Signal Processing*, Submitted, Dec. 2010. [Online]. Available: http://arxiv.org/abs/1012.4715

[5] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. Information Theory*, vol. IT-29, pp. 439–441, May 1983.

[6] A. Khina, Y. Kochman, and U. Erez, "Physical-layer MIMO relaying," in *Proceedings of ISIT-2011, St. Petersburg, Russia*, July 2011.

[7] 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*, pp. 295–300.

[8] A. Khina, Y. Kochman, U. Erez, and G. W. Wornell, "Incremental coding over MIMO channels," submitted to Information Theory Workshop (ITW2011), Parati, Brazil.

[9] S. Alamouti, "A simple transmit diversity technique for wireless communications," *Selected Areas in Communications, IEEE Journal on*, vol. 16, no. 8, pp. 1451 –1458, Oct. 1998.

[10] V. Tarokh, N. Sheshadri, and A. R. Calderbank, "Space-time block codes form orthogonal designs," *IEEE Trans. Information Theory*, vol. IT-45, pp. 1456–1467, July 1999.

[11] Y. Jiang, W. Hager, and J. Li, "The geometric mean decompostion," *Lin. Algebra and Its Applications*, vol. vol. 396, pp. 373–384, Feb. 2005.

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

[13] U. Erez, M. D. Trott, and G. W. Wornell, "Rateless coding for Gaussian channels," 2007. [Online]. Available: http://arxiv.org/abs/0708.2575

[14] A. V. Geramita and J. Seberry, "Orthogonal designs, quadratic forms and Hadamard matrices," *Lec. Notes in Pure and App. Mathematics*, 1979.