Joint Unitary Triangularization for Gaussian Multi-User MIMO Networks

Anatoly Khina, Idan Livni, Ayal Hitron, and Uri Erez

Technical Report, Dept. of EE-Systems, Tel Aviv University, March 27, 2015

***Abstract*—The problem of transmitting a common message to multiple users over the Gaussian multiple-input multiple-output broadcast channel is considered, where each user is equipped with an arbitrary number of antennas. A closed-loop scenario is assumed, for which a practical capacity-approaching scheme is developed. By applying judiciously chosen unitary operations at the transmit and receive nodes, the channel matrices are triangularized so that the resulting matrices have equal diagonals, up to a possible multiplicative scalar factor. This, along with the utilization of successive interference cancellation, reduces the coding and decoding tasks to those of coding and decoding over the single-antenna additive white Gaussian noise channel. Over the resulting effective channel, any off-the-shelf code may be used. For the two-user case, it was recently shown that such joint unitary triangularization is always possible. In this paper, it is shown that for more than two users, it is necessary to carry out the unitary linear processing jointly over multiple channel uses, i.e., space–time processing is employed. It is further shown that exact triangularization, where all resulting diagonals are equal, is still not always possible, and appropriate conditions for the existence of such are established for certain cases. When exact triangularization is not possible, an asymptotic construction is proposed, that achieves the desired property of equal diagonals up to edge effects that can be made arbitrarily small, at the price of processing a sufficiently large number of channel uses together.**

***Index Terms*—Matrix decompositions, space–time modulation, common-message broadcast, physical-layer multicast, Gaussian MIMO, successive interference cancellation.**

I. Introduction

A RECURRING theme in digital communications is the use of a standard "off-the-shelf" coding module in combination with appropriate linear pre/post processing which is tailored to the specific channel model. Such methods are appealing due to their low complexity of implementation as well as conceptually, since the tasks of coding and modulation are effectively decoupled.

The simplest example of the decoupling approach is provided by the singular-value decomposition (SVD) in communication for single-user (SU) Gaussian multiple-input multiple-output (MIMO) channels. In this case, the MIMO channel is transformed into diagonal form, corresponding to parallel scalar channels. If one allows pre- or post-interference cancellation, a much broader class of decompositions may be employed. For SU MIMO communication, this includes the widely used schemes based on the QR decomposition, namely, V-BLAST/GDFE [1], [2]. Further applicable decompositions that allow to approach capacity via decoupling, include the geometric-mean decomposition (GMD) [3]–[5] for the SU case, and its generalization — block diagonal GMD [6] — for private-message broadcast (BC).

In the present work, we aim to extend the decoupling approach beyond the single-user Gaussian MIMO channel, to the more general problem of common-message BC. That is, we consider a scenario in which a transmitter, equipped with multiple antennas, wishes to send simultaneously the same ("common") message to multiple users, each of which equipped with (any number of) multiple antennas.

The capacity of this scenario, referred to as common-message BC (or "physical-layer multicast"), is well known, and is given by the compound channel capacity [7]–[9]. Unfortunately, whereas for the problem of transmitting private messages over the Gaussian MIMO BC channel, capacity can be achieved via decoupling (in conjunction with dirty-paper coding; see, e.g., [6], [10]), practical schemes that attain an analogous result for the common-message counterpart are not hitherto known.

Beyond being important in its own right, common-message BC serves as the basis for various communication settings, since many communication scenarios can be transformed into an equivalent MIMO common-message BC setting. This is the case for rateless coding over SISO and MIMO Gaussian channels [11] (see also Section V-E), permuted channels [12] (see also Section V-F), half- and full-duplex SISO and MIMO relaying [11], [13], two-way MIMO relaying [14], [15] and many others.

Extension of the decoupling approach, which is at the heart of single-user scalar systems, to the multiple-user MIMO common-message BC problem requires, however, overcoming a major hurdle: Not only is simultaneous diagonalization impossible, even the existence of appropriate joint triangularization for two users was not known to be possible until recently [16].

Hence, different practical approaches have been proposed over the years for the problem of conveying a common message over Gaussian MIMO broadcast channels. However, none of these approaches is capacity achieving in general, even for simple cases. To illustrate this, we consider a simple three-

---

The authors are with the Department of Electrical Engineering, Tel Aviv university, Israel, email: {anatolyk,idanlivn,ayal,uri}@eng.tau.ac.il Parts of this work were presented at the International Symposium on Information Theory (ISIT) 2011 in Saint Petersburg, Russia, and at the International Symposium on Information Theory (ISIT) 2012 in Cambridge, MA, USA.

The work of Anatoly Khina was supported in part by the Feder family award, the David and Paulina Trotsky Foundation, and by the Clore Israel Foundation. The work of Anatoly Khina and Ayal Hitron was supported in part by the Yitzhak and Chaya Weinstein Research Institute for Signal Processing. The work of Uri Erez was supported in part by the Israel Science Foundation under Grant No. 1557/12.

user example.

Example 1 (Degrees-of-freedom mismatch): Consider the following three-user channel:[1]

$$\boldsymbol{y}_k = H_k \boldsymbol{x} + \boldsymbol{z}_k \,, \qquad k = 1, 2, 3 \,,$$

where $\boldsymbol{z}_k$ is an additive white Gaussian noise (AWGN), specifically we assume to be circularly-symmetric Gaussian noise with unit power for each element $\mathcal{CN}(0, I)$, $\boldsymbol{x}$ is the channel vector subject to an average power constraint P, H_k are the complex-valued channel matrices

$$H_1 = \begin{pmatrix} \alpha_1 & 0 \end{pmatrix}, \, H_2 = \begin{pmatrix} 0 & \alpha_1 \end{pmatrix}, \, H_3 = \begin{pmatrix} \alpha_2 & 0 \\ 0 & \alpha_2 \end{pmatrix},$$

and α_1 and α_2 are chosen such that the WI capacities of all three channels are equal, viz.

$$C^{\text{common}} = C_{\text{WI}} \triangleq \log(1 + |\alpha_1|^2 P/2) = 2 \log(1 + |\alpha_2|^2 P/2) \,.$$

This example models a three-user "degrees-of-freedom-mismatch" scenario, in which the first two users are equipped with a single antenna each (i.e., they have only one degree of freedom), whereas the third user is equipped with two antennas (i.e., has two degrees of freedom).

Of course, from a purely information-theoretic viewpoint, a random i.i.d. Gaussian codebook over time and space is simultaneously good (i.e., capacity achieving) for all three users in the example. However when considering *practical* codes, the situation is very different.

To the best of our knowledge, known *practical* schemes are limited to the smallest number of degrees of freedom ("multiplexing gain") of the different users, or incorporate time- or frequency-sharing, which again lose degrees of freedom. Thus, these schemes achieve only a fraction of the available degrees of freedom. Alternatively, maximal degrees-of-freedom open-loop techniques may be used (e.g., in the case of two transmit antennas as in the example, golden code modulation [17]–[20]). However, these are far from capacity-achieving at low to moderate transmission rates.

By using single-stream communication, in the high SNR regime, the third user is able to achieve only half of its individual capacity. On the other hand, transmitting two streams across the two transmit antennas, results in a loss of half of the capacity of users 1 and 2. Another approach considered in the literature for this problem is that of using a "pure open-loop" approach, namely Alamouti modulation [21] — for the two-transmit antenna case, and orthogonal space–time block coding (OSTBC) [22] — for more. The performance of these schemes does not depend on the number of receivers. However, this universality comes at the price of a substantial rate loss for MIMO channels having several receive antennas, as these schemes use only a single stream, thus failing to achieve the multiplexing gain offered by the MIMO channel of user 3 in the example.[2] Also note that time/frequency sharing incur a great loss in performance (up to half of the capacity in this case). Other techniques that can be applied for this scenario [23]–[25] are also suboptimal in general.

The aim of the present work is to develop a practical capacity-achieving scheme for the Gaussian MIMO common-message broadcast MIMO setting via decoupling, allowing to utilize a "black box" approach to coding. Namely, this approach allows constructing a capacity-achieving scheme that utilizes only "off-the-shelf" encoders and decoders designed for scalar AWGN channels, together with simple signal processing tools.

We construct a capacity-approaching scheme that applies judiciously chosen unitary operations to the time-extended channel matrices at the transmitter and the receivers in conjunction with successive interference cancellation. In contrast to the open-loop OSTBC structures, that strive for an "orthogonal design" structure, i.e., to diagonalize the channel matrices (see, e.g., [22]), the space–time structure presented in this work results in triangular matrices, similar to those of V-BLAST/GDFE, but having *equal diagonals*. This gives rise to effective parallel scalar additive white Gaussian noise (AWGN) channels, over which standard codes can be used to approach capacity. Thus, the proposed scheme can be thought of as an "interpolation" between the open-loop OSTBC and the closed-loop SU V-BLAST/SVD ones.

The results of this paper generalize those of [16], in which the case of only two users was considered, for which it suffices to apply unitary transformations directly to the channel matrices. For more users, on the other hand, we show that jointly processing multiple channel uses is necessary. That is, the unitary transformations are applied to time-extended channel matrices.

The rest of the paper is organized as follows. In Section II we present the notations that are used throughout the paper. In Section III we define the Gaussian MIMO common-message BC channel model. In Section IV we recall known schemes for the single-user case, relying on various forms of unitary matrix decompositions. In Section V we suggest a generalization of the SU schemes to the multi-user scenario, based on newly developed matrix decompositions and derive necessary and sufficient conditions for the existence of such decompositions in some scenarios. Then, in Section VI, we generalize the multi-user scheme by employing space–time coding and discuss the existence of "perfect" decompositions needed for such a construction. In Section VII we utilize the space–time structure in order to develop a practical scheme, which is nearly optimal and asymptotically achieves the capacity for any number of users, even when "perfect decompositions" are not possible. Finally, in Section VIII we present some extensions of the results and conclude in Section IX.

II. NOTATION

The following notation will be used throughout the paper:

- Channel matrix of dimension $n_r \times n_t$: H, where n_r and n_t stand for the number of antennas at the receiver and at the transmitter, respectively.
- Channel gain: α.
- Augmented channel matrix: $\tilde{H}$, see Definition 2 in Section IV-C.

---

[1] Throughout this paper, vectors are denoted by boldface lower case letters, and matrices are denoted by upper case letters. Logarithms are taken to base 2 and rates are given in bits.

[2] Moreover, for more than two transmit antennas, the OSTBC of [22] attain strictly less than one degree of freedom.

- Channel canonical matrix: G, see Definition 3 in Section IV-C.
- General square complex matrix of dimensions $n \times n$: A.
- Hermitian square matrix: S.
- Upper triangular matrix with diagonal $\boldsymbol{r}$: R.
- Upper triangular matrix with a constant diagonal: T.
- Real-valued diagonal matrix: D.
- Complex-valued matrices whose columns are orthonormal (which are unitary, in case these matrices are square): U,V,Q.
- The Identity matrix: I.
- Capital script letters denote time-extended matrices: $\mathcal{H}, \mathcal{A}, \mathcal{S}, \mathcal{R}, \mathcal{T}, \mathcal{U}, \mathcal{V}, \mathcal{Q}, \mathcal{G}$, see Section VI-A.
- Number of users: K.
- Number of time extensions: N.
- Vectors are denoted by boldface lower case letters. For example, $\boldsymbol{x}$ denotes the transmitted vector, $\boldsymbol{y}$ — the received vector, and $\boldsymbol{z}$ — the noise vector.
- Time-extended vectors are denoted by script lower case letters. For example, $\mathscr{x}, \mathscr{y}$ and $\mathscr{z}$ denote extended transmit, received and noise vectors, respectively.
- Indices: j, k, l, m, p, q.
- Channel capacity: C.
- All logarithms are taken to base 2. All rates are given in bits per two dimensions (complex channel use).
- Average power constraint: P.
- Covariance matrix of the vector $\boldsymbol{x}$: $C_{\boldsymbol{x}}$.
- Singular values and generalized singular values: $\boldsymbol{\sigma}$, $\boldsymbol{\mu}$.
- Real and imaginary parts of a complex number: $\mathrm{Re}\{\cdot\}$, $\mathrm{Im}\{\cdot\}$.
- Expected value of a random variable: $\mathbb{E}(\cdot)$.
- Vector ℓ_2 norm: $\|\cdot\|$.
- Determinant of a matrix: $\det(\cdot)$.
- Trace of a matrix: $\mathrm{tr}\,(\cdot)$.
- Adjugate (the transpose of the cofactor) matrix: $\mathrm{adj}(\cdot)$.

III. COMMON-MESSAGE BROADCAST CHANNEL MODEL

The K-user Gaussian MIMO broadcast channel consists of one transmit and K receive nodes, where each received signal is related to the transmitted signal through a MIMO link:[3]

$$\boldsymbol{y}_k = H_k \boldsymbol{x} + \boldsymbol{z}_k \,, \qquad k = 1, \ldots, K \,, \tag{1}$$

where $\boldsymbol{x}$ is the channel input of dimensions $n_t \times 1$, and is subject to an average power constraint P;[4] $\boldsymbol{y}_k$ is the channel output vector of receiver k ($k = 1, \ldots, K$) of dimensions $n_r^{(k)} \times 1$; H_k is the channel matrix to user k of dimensions $n_r^{(k)} \times n_t$; and $\boldsymbol{z}_k$ is an additive circularly-symmetric Gaussian noise vector of dimensions $n_r^{(k)} \times 1$, where, without loss of generality, we assume that the noise elements are mutually independent and identically distributed with unit power.

The aim of the transmitter is to send the same (common) message to all the receivers. The capacity of this scenario is well known to equal the (worst-case) capacity of the compound channel [7]–[9], with the compound parameter being the channel matrix index:

$$C\left(\{H_k\}_{k=1}^K, P\right) = \max_{C_{\boldsymbol{x}}} \min_{k=1,\ldots,K} I(H_k, C_{\boldsymbol{x}}) \,, \tag{2}$$

where $I(H_k, C_{\boldsymbol{x}})$ is the mutual information between the channel input $\boldsymbol{x}$ and the channel output $\boldsymbol{y}_i$, obtained by taking $\boldsymbol{x}$ to be Gaussian with covariance matrix $C_{\boldsymbol{x}}$:

$$I(H, C_{\boldsymbol{x}}) \triangleq \log\det\left(I + H C_{\boldsymbol{x}} H^\dagger\right) \,, \tag{3}$$

and the maximization is carried over all admissible input covariance matrices $C_{\boldsymbol{x}}$, satisfying the power constraint $\mathrm{tr}\,(C_{\boldsymbol{x}}) \leq P$.

For $K = 1$ (SU), the capacity (2) can be achieved via the decoupling approach in several ways, each corresponding to a different matrix decomposition.

IV. SINGLE-USER SCHEME VIA MATRIX TRIANGULARIZATION: KNOWN RESULTS

In this section we briefly recall some important matrix decompositions, and the associated SU communication schemes. In Section IV-A we recall the generalized triangular decomposition (GTD), and some of its important special cases which include the SVD, QR, and GMD. A geometrical interpretation of these decompositions is provided in Section IV-B. In Section IV-C, we describe how the GTD can be used in order to construct a practical capacity-achieving communication scheme for the SU Gaussian MIMO communication problem.

A. Generalized Triangular Decomposition

We only consider the decomposition of *square* invertible matrices throughout this work. As we show in the sequel, this does not impose any restriction on the communication problems addressed.

The next theorem uses the following definition:

Definition 1 (Multiplicative Majorization (See [26])): Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be two n-dimensional vectors of positive elements. Denote by $\tilde{\boldsymbol{x}}$ and $\tilde{\boldsymbol{y}}$ the vectors composed of the entries of $\boldsymbol{x}$ and $\boldsymbol{y}$, respectively, ordered non-increasingly. We say that $\boldsymbol{x}$ majorizes $\boldsymbol{y}$ ($\boldsymbol{x} \succeq \boldsymbol{y}$) if they have equal products:

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

and their (ordered) elements satisfy, for any $1 \leq l < n$,

$$\prod_{j=1}^l \tilde{x}_j \geq \prod_{j=1}^l \tilde{y}_j \,.$$

Theorem 1 (Generalized Triangular Decomposition): Let A be an invertible matrix of dimensions $n \times n$ and $\boldsymbol{r}$ be an n-dimensional vector of positive elements. A GTD of the matrix A is given by:

$$A = URV^\dagger \,, \tag{4}$$

where U, V are unitary matrices, and R is an upper triangular matrix with a prescribed set of diagonal values $\boldsymbol{r}$, where $r_j =$

---

[3] For ease of notation, in the case $K = 1$ we denote the single channel matrix H_1 by H.

[4] Alternatively, one can consider any other input covariance constraint, e.g., individual power constraints, and covariance matrix constraints. Given any covariance matrix, the approach described in the sequel may be applied to approach (3).

R_{jj}. This decomposition exists if and only if the vector $\boldsymbol{r}$ is majorized by the singular-values vector of A:

$$\boldsymbol{\sigma}(A) \succeq \boldsymbol{r} \,. \tag{5}$$

In other words, the singular values are an extremal case for the diagonal of all possible unitary triangularizations.

The necessity of the majorization condition was proven by Weyl [27], and the sufficiency of this condition — by Horn [28]. Explicit constructions of the decomposition were introduced in [29] and [30].

We now recall three important special cases of the GTD.

1) SVD (See, .e.g., [31]): An important special case of the GTD is the SVD, in which the resulting matrix R in (4) is a *diagonal* matrix, such that the diagonal elements of R are equal to the singular values of the original matrix A.

2) QR Decomposition (See, .e.g., [31]): Another important special case of the GTD is the QR decomposition, in which the matrix V in (4) equals to the identity matrix and hence does not depend on the matrix A. This decomposition can be constructed by performing Gram-Schmidt orthonormalization on the (ordered) columns of the matrix A.

3) GMD (See [3]–[5]): A GMD of a square complex invertible matrix A is given by:

$$A = UTV^\dagger \,, \tag{6}$$

where U, V are unitary matrices, and T is an upper triangular matrix such that all its diagonal values equal to the geometric mean of the singular values of A, which is real and positive.

Note that this decomposition always exists if A is invertible (since the vector of singular values of A necessarily majorizes the vector of diagonal elements of T), but is not unique.

B. Geometric Interpretation of the GTD

We give a geometric interpretation of the GTD of Theorem 1, for the special case of 2×2 real matrices. A similar geometric interpretation can be devised for the general case.

In the real case, unitary matrices reduce to (real) orthogonal ones. In the 2×2 case, these orthogonal matrices are merely rotation matrices.[5] Thus, the matrices U and V of Theorem 1 are rotation matrices, namely,

$$\begin{aligned} V &= \begin{pmatrix} \cos\theta_r & -\sin\theta_r \\ \sin\theta_r & \cos\theta_r \end{pmatrix} \\ U &= \begin{pmatrix} \cos\theta_\ell & -\sin\theta_\ell \\ \sin\theta_\ell & \cos\theta_\ell \end{pmatrix} , \end{aligned} \tag{7}$$

where θ_r and θ_ℓ are the rotation angles.

Denote the columns of the matrix to be decomposed, A, by $\boldsymbol{a}$ and $\boldsymbol{b}$:

$$A \triangleq \begin{pmatrix} \boldsymbol{a} & \boldsymbol{b} \end{pmatrix} \triangleq \begin{pmatrix} a_x & b_x \\ a_y & b_y \end{pmatrix}$$

and assume, without loss of generality, $\det(A) = 1$.

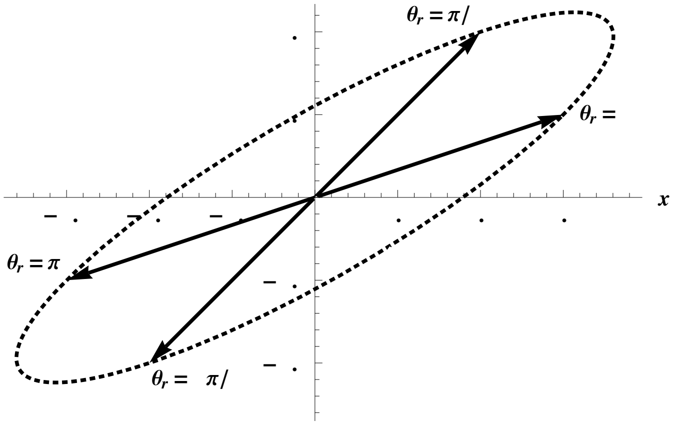


Fig. 1: All possible column vectors of AV, where V is a rotation matrix, and $\boldsymbol{a} = (3/2, 1/2)^\dagger$ and $\boldsymbol{b} = (1, 1)^\dagger$. The arrows correspond to $\tilde{\boldsymbol{a}}$ at different angles.

By multiplying A by V on the right, we obtain

$$\begin{aligned} AV &= \begin{bmatrix} a_x \cos\theta_r + b_x \sin\theta_r & -a_x \sin\theta_r + b_x \cos\theta_r \\ a_y \cos\theta_r + b_y \sin\theta_r & -a_y \sin\theta_r + b_y \cos\theta_r \end{bmatrix} \\ &= \begin{bmatrix} a_x \cos\theta_r + b_x \sin\theta_r & a_x \cos(\theta_r + \frac{\pi}{2}) + b_x \sin(\theta_r + \frac{\pi}{2}) \\ a_y \cos\theta_r + b_y \sin\theta_r & a_y \cos(\theta_r + \frac{\pi}{2}) + b_y \sin(\theta_r + \frac{\pi}{2}) \end{bmatrix} \end{aligned} \tag{8}$$

By varying the rotation angle θ_r, it is readily verified that the resulting column vectors in (8), move along an ellipse, centered at the origin. This is illustrated in Figure 1, for a specific choice of A, where we define $\tilde{A} \triangleq AV$ and its columns — by $\tilde{\boldsymbol{a}}$ and $\tilde{\boldsymbol{b}}$.

After applying V on the right, we multiply the resulting matrix $\tilde{A}$ by a rotation matrix $U^\dagger$ on the left. The latter operation rotates the column vectors $\tilde{\boldsymbol{a}}$ and $\tilde{\boldsymbol{b}}$, by an angle $(-\theta_\ell)$ (the minus is due to the transposition of U prior to multiplication). The angle θ_ℓ is chosen such that $U^\dagger \tilde{\boldsymbol{a}}$ is aligned with the x-axis. This is illustrated for a specific choice of $\tilde{\boldsymbol{a}}$ and $\tilde{\boldsymbol{b}}$ in Figure 2.

Remark 1: Since the orthogonal matrix V is applied on the right, the norms of the *rows* of A are not affected. Nevertheless, the *columns* of AV have different norms, in general, from those of the columns of A, as can be seen from (8). The multiplication on the left by $U^\dagger$, on the other hand, does not change the norms of the columns. As for the angle between the column vectors — multiplication by a unitary matrix V on the right changes the relative angle between the two vectors, unlike a unitary operation applied on the left, which only rotates the two vectors together, but does not change the relative angle between the two.

Since the norms of the columns are not affected by unitary operations applied on the left, the possible values on the diagonal of the resulting triangular matrix in the GTD, are fully determined by the norms ("lengths") of the column vectors resulting after applying V on the right, which in turn, vary together on an ellipse.

We next interpret geometrically the special cases of SVD, QR and GMD (for the real 2×2 case).

---

[5] In general, reflection matrices need to be considered in conjunction with the rotation matrices. However, reflection matrices are not needed for the construction of GTD, as will become clear in the sequel.

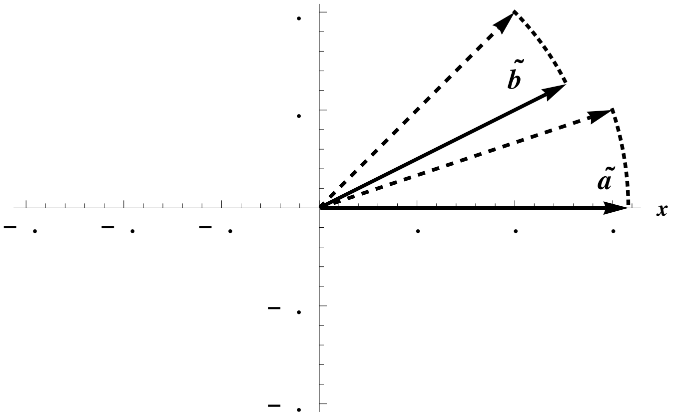
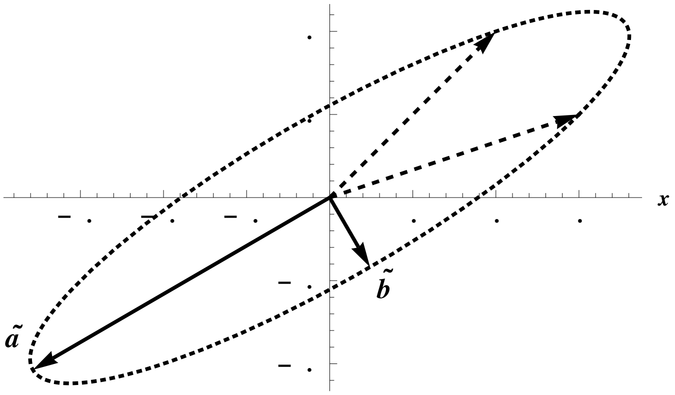
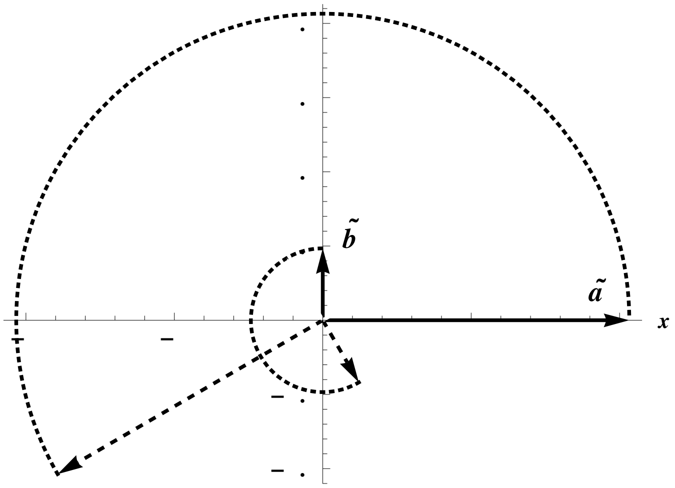


Fig. 2: Rotation by $U^\dagger$ of $\tilde{\boldsymbol{a}}$ and $\tilde{\boldsymbol{b}}$, resulting from $U^\dagger \tilde{A}$ (or alternatively of $\boldsymbol{a}$ and $\boldsymbol{b}$ in the QR decomposition case), until $U^\dagger \tilde{\boldsymbol{a}}$ is aligned with the x-axis, for $\tilde{\boldsymbol{a}} = (3/2, 1/2)^\dagger$ and $\tilde{\boldsymbol{b}} = (1,1)^\dagger$.

1) SVD: In this decomposition, the resulting columns, at the end of the process, must be orthogonal. This is established by choosing θ_r such that the relative angle between the resulting vectors, after the multiplication by V, is $\pi/2$. As we show below, this is always possible. Afterwards, the two vectors are rotated together via the left-multiplication by $U^\dagger$, until they lie parallel to the axes. This process is demonstrated in Figure 3. Moreover, the resulting orthogonal vectors correspond also to the longest and shortest ("extreme") possible diagonal values achievable via the GTD. This can also be seen in Figure 3 and is formally stated in the following lemma. Note that this is a special (2×2) case of the majorization property (5) of the GTD. Here, we provide a geometric proof.



(a) Right rotation by $\theta_r = 3.865$, for which the vectors are orthogonal.



(b) Left rot. by $\theta_\ell = 3.667$, for which the vectors are aligned with the axes.

Fig. 3: SVD for $\boldsymbol{a} = (3/2, 1/2)^\dagger$ and $\boldsymbol{b} = (1,1)^\dagger$.

Proof: The norm of $\tilde{\boldsymbol{a}}$ after applying a rotation matrix V on the right is

$$\begin{aligned}\left\|\tilde{\boldsymbol{a}}\right\|^2 &= \left(a_x \cos\theta_r + b_x \sin\theta_r\right)^2 + \left(a_y \cos\theta_r + b_y \sin\theta_r\right)^2 \\ &= \frac{1}{2}\left(a_x^2 + a_y^2 + b_x^2 + b_y^2\right) \\ &\quad + \frac{1}{2}\left(a_x^2 + a_y^2 - b_x^2 - b_y^2\right)\cos 2\theta_r \\ &\quad + \left(a_x b_x + a_y b_y\right)\sin 2\theta_r \,.\end{aligned}$$

Similarly, the norm of $\tilde{\boldsymbol{b}}$ is given by

$$\begin{aligned}\left\|\tilde{\boldsymbol{b}}\right\|^2 &= \frac{1}{2}\left(a_x^2 + a_y^2 + b_x^2 + b_y^2\right) \\ &\quad - \frac{1}{2}\left(a_x^2 + a_y^2 - b_x^2 - b_y^2\right)\cos 2\theta_r \\ &\quad - \left(a_x b_x + a_y b_y\right)\sin 2\theta_r \,.\end{aligned}$$

The extreme values of $\left\|\tilde{\boldsymbol{a}}\right\|^2$ and $\left\|\tilde{\boldsymbol{b}}\right\|^2$ are achieved at θ_r satisfying:

$$-\frac{\mathrm{d}\left(\left\|\tilde{\boldsymbol{b}}\right\|^2\right)}{\mathrm{d}\theta_r} = \frac{\mathrm{d}\left(\left\|\tilde{\boldsymbol{a}}\right\|^2\right)}{\mathrm{d}\theta_r} \tag{9a}$$

$$= -\left(a_x^2 + a_y^2 - b_x^2 - b_y^2\right)\sin 2\theta_r \tag{9b}$$

$$+ 2\left(a_x b_x + a_y b_y\right)\cos 2\theta_r \tag{9c}$$

$$= 0 \,. \tag{9d}$$

On the other hand, the vectors $\tilde{\boldsymbol{a}}$ and $\tilde{\boldsymbol{b}}$ are orthogonal for θ_r values satisfying

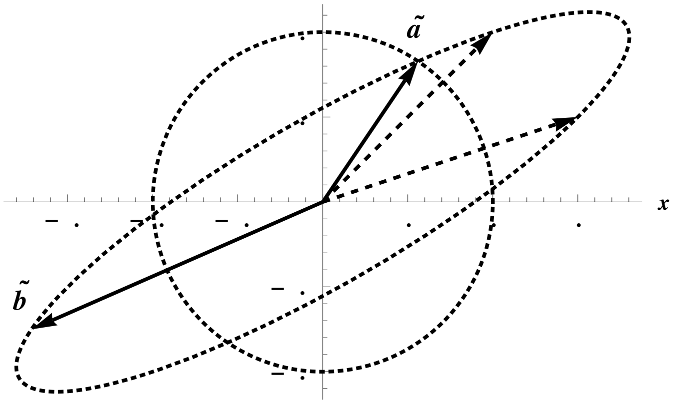
$$\langle \boldsymbol{a}, \boldsymbol{b} \rangle = -\frac{1}{2}\left(a_x^2 + a_y^2 - b_x^2 - b_y^2\right)\sin 2\theta \tag{10a}$$

$$+ \left(a_x b_x + a_y b_y\right)\cos 2\theta \tag{10b}$$

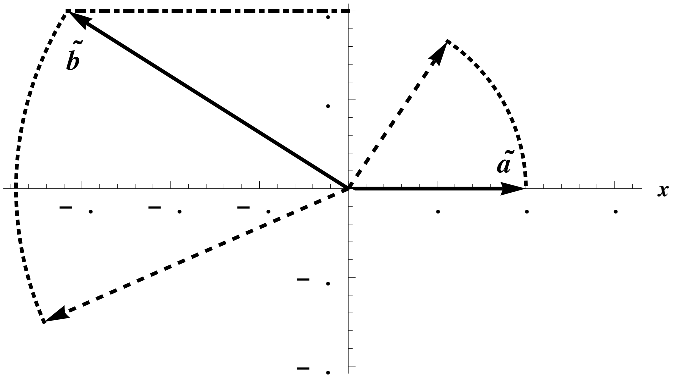
$$= 0 \,. \tag{10c}$$

Observing that the requirements of (9) and (10) are the same, and that the second derivatives of $\left\|\tilde{\boldsymbol{a}}\right\|^2$ and $\left\|\tilde{\boldsymbol{b}}\right\|^2$ are opposite, we conclude the desired result. ■

2) QR Decomposition: In this decomposition no right rotation V is applied, i.e., $V = I$ or equivalently $\theta_r = 0$. Thus, a left rotation is applied to the columns of A, until the first column vector is aligned with the x-axis. This suggests that the first diagonal element is equal to the norm of the first column of A (prior to rotation); the second diagonal element can be computed from the determinant and the first diagonal vector,



(a) Right rotation by $\theta_r = 1.843$, for which the first vector has unit norm.



(b) Left rotation by $\theta_\ell = 0.977$, for which the first vector is aligned with the $\boldsymbol{x}$ axis.

Fig. 4: GMD for $\boldsymbol{a} = (3/2, 1/2)^\dagger$ and $\boldsymbol{b} = (1, 1)^\dagger$.

or alternatively by computing the norm of the orthogonal component of the second column vector to the first one. See Figure 2.

3) GMD: In this decomposition the angle θ_r is chosen such that the length (norm) of $\tilde{\boldsymbol{a}}$ is equal to 1, or equivalently we seek for an angle θ_r for which the ellipse intersects with the unit circle. Since both the ellipse and the unit circle (which corresponds to the 2×2 identity matrix) have determinants equal to 1 (i.e., have the same area) and both are centered at the origin, they must intersect at exactly 4 points, unless the ellipse is itself the unit circle (in which case there is an infinite number of intersection points). The operation on the left rotates the two vectors until the first is aligned with the x-axis. Moreover, since unitary operations preserve volume (absolute value of the determinant), the second diagonal element must be 1 as well. That is, the projection of the second vector on the y-axis is equal to 1. The remaining element may be found, e.g., via the Frobenius norm, which is again invariant under rotations on both sides, and its sign may be easily determined as well. This is demonstrated in Figure 4.

C. SU MIMO transmission via Matrix Triangularization

We now review the capacity-approaching communication schemes that utilize the above matrix decompositions. For the SU case (i.e., $K = 1$ in (1)), a practical communication scheme can be obtained by applying the SVD to the channel matrix H:

$$H = UDV^\dagger \,.$$

By applying the pre-processing matrix V at the transmitter and the post-processing matrix $U^\dagger$ at the receiver, the resulting effective channel matrix becomes *diagonal*, and therefore the capacity can be achieved using off-the-shelf codes, designed for *scalar* SU AWGN channels. The rates of those codes are determined by the SNRs of the independent scalar sub-channels, namely, by the diagonal elements of the diagonal matrix D (after allocating power to the resulting sub-channels, via water-pouring).

We now review a more general scheme, applicable to any GTD rather than the special case of SVD. This scheme is based upon the derivation of the MMSE variant of Vertical Bell-Laboratories Space–Time coding (V-BLAST), see, e.g., [4], [32], [33].

Definition 2 (Augmented Matrix): Define the following *augmented matrix*:[6]

$$\tilde{H} \triangleq \begin{pmatrix} HC_{\boldsymbol{x}}^{1/2} \\ I_{n_t} \end{pmatrix} \,, \tag{11}$$

where I_{n_t} is the $n_t \times n_t$ identity matrix. Next, the matrix $\tilde{H}$ is transformed into a square matrix, by means of the QR decomposition.

Definition 3 (Channel Canonical Matrix): Let $\tilde{H}$ be the augmented matrix (11), and let

$$\tilde{H} = QG \,,$$

where Q is an $(n_r + n_t) \times n_t$ matrix with orthonormal columns and G is an $n_t \times n_t$ upper triangular matrix with real-valued positive diagonal elements. The matrix G will be referred to as a *channel canonical matrix*, reminiscent of the system canonical response defined in [34] for LTI scalar systems.

Now the matrix G is decomposed according to the GTD:

$$G = URV^\dagger \,, \tag{12}$$

where R is upper triangular whose diagonal values are equal to the prescribed diagonal elements $r_1, \ldots, r_{n_t}$ (which satisfy the multiplicative majorization condition of Definition 1), and $r_j^2 - 1$ are the effective signal-to-noise ratios of the scalar sub-channels.

Remark 2: Due to the presence of the identity matrix I_{n_t} in (11), it follows that the the diagonal elements of G and R are necessarily greater or equal to 1, and their determinants are greater than 1.[7]

The transmission scheme is as follows:

1) Construct n_t codewords, each from a codebook matched to a scalar AWGN channel of signal-to-noise ratio (SNR) $r_j^2 - 1$. That is, up to a rate of $\log r_j^2$.

---

[6] $C_{\boldsymbol{x}}^{1/2}$ is any matrix B satisfying $BB^\dagger = C_{\boldsymbol{x}}$, and can be found, e.g., via the Cholesky decomposition.

[7] Assuming a "canonical QR decomposition" is used, i.e., the one that results in positive diagonal entries in the triangular matrix.

2) In each channel use, an n_t-length vector $\tilde{\boldsymbol{x}}$ is formed using one sample from each codebook. The transmitted vector $\boldsymbol{x}$ is then obtained using the following linear precoder:

$$\boldsymbol{x} = C_{\boldsymbol{x}}^{1/2} V \tilde{\boldsymbol{x}} \,.$$

3) The receiver calculates

$$\tilde{\boldsymbol{y}} = U^{\dagger} \tilde{Q}^{\dagger} \boldsymbol{y} \,, \tag{13}$$

where $\tilde{Q}$ consists of the first n_t rows of Q.

4) Finally, the codebooks are decoded using successive interference cancellation, starting from the n_t-th codeword and ending with the first one: The n_t-th codeword is decoded first, using the n_t-th element of $\tilde{\boldsymbol{y}}$, treating the other codewords as AWGN. The effect of the n_t-th element of $\tilde{\boldsymbol{x}}$ is then subtracted out from the remaining elements of $\tilde{\boldsymbol{y}}$. Next, the $(n_t - 1)$-th codeword is decoded, using the $(n_t - 1)$-th element of $\tilde{\boldsymbol{y}}$ — and so forth.

The proof of optimality of this scheme, i.e., that it is capacity achieving, appears in [33, Lemma III.3].

Note that each element of $\tilde{\boldsymbol{x}}$ should be understood to correspond to a symbol of a codebook of length L. Thus, the index time is suppressed. Similarly, the successive interference cancellation process of recovering the codebooks from $\tilde{\boldsymbol{y}}$ should be understood, again, to correspond to a symbol of a codebook of length L. Our analysis is not affected by the exact value of L, but rather only by the gap to capacity of the base code. Hence, in order to approach capacity, L needs to be large. Throughout this paper, we assume capacity-achieving scalar (base) codes; any loss in these codes, would translate in a straightforward manner to a loss in the overall scheme.

Remark 3: If we take $V = I$ in (12), namely use the QR decomposition, we obtain a transmission scheme that requires no precoding at the transmitter. Since the QR decomposition is unique, we have no freedom in choosing the diagonal values r_j. Alternatively, the matrices U and V can be chosen according to the SVD. In this case, the resulting matrix R in (12) is *diagonal*, and therefore the channel is transformed into parallel independent scalar sub-channels and there is no need to perform successive interference cancellation. As in the case of the QR decomposition, the SVD is unique, and there is no freedom in choosing the diagonal values r_j (which, in this case, are the singular values of the matrix G). Finally, If the matrices U and V are chosen according to the GMD (6), then all the values r_j are equal, meaning that all the codebooks in the scheme have the same rate. Moreover, in this case the *same* scalar codebook can be used over all the sub-channels.[8] This special case is known as the uniform channel decomposition (UCD) [33].

Remark 4 (Decoding Order): In step 4 of the scheme, one could decode the codebooks in a different order. This corresponds to replacing the QR decomposition (12) with Gram-Schmidt orthonormalization in a different order, e.g., QL decomposition. Alternatively, this could be represented in the notations of this section by retaining the QR decomposition, but performing it on a column-permuted matrix $G\Pi$, where Π is some permutation matrix. This, in general, would alter the rate allocation between the different sub-streams.

V. MULTI-USER SCHEME VIA MATRIX TRIANGULARIZATION

The goal of this section is to generalize the point-to-point communication scheme, presented in Section IV-C, to the K-user BC channel defined in Section III. This is a generalization of the two-user case ($K = 2$) that was considered in [16].

We start in Section V-A by defining some forms of joint decomposition of K matrices. Namely, we define the K-user geometric mean decomposition (K-GMD) and the K-user joint equi-diagonal triangularization (K-JET). A communication scheme for the K-user common-message BC setting, based on these decompositions, is described in Section V-C. Unfortunately, these decompositions do not always exist; In Section V-D we provide necessary and sufficient conditions for the existence of these decompositions, for a certain special case.

A. K-JET and K-GMD

We now present the definitions of K-GMD and K-JET — decompositions of K square matrices of the same dimensions and having the same determinant.

Definition 4 (K-JET): Let $A_1, \ldots, A_K$ be K invertible complex matrices of dimensions $n \times n$, with equal determinants. A K-JET of these matrices is a decomposition

$$A_k = U_k R_k V^{\dagger} \,, \qquad k = 1, \ldots, K \,, \tag{14}$$

where $U_1, \ldots, U_K, V$ are $n \times n$ unitary matrices, and $R_1, \ldots, R_K$ are upper triangular $n \times n$ matrices with *the same* real, positive diagonal values, namely,

$$[R_1]_{jj} = \cdots = [R_K]_{jj} \,, \qquad j = 1, \ldots, n \,. \tag{15}$$

Remark 5: For $K = 2$, 2-JET will be simply referred to as JET. JET of two matrices was introduced in [16], where it was proved to always exist (for any two matrices A_1 and A_2 with equal determinants).

Remark 6: The K-JET of Definition 4 easily extends to matrices with non-equal determinants as follows. Define the normalized matrices

$$\tilde{A}_k \triangleq \left| \det(A_k) \right|^{-1/n} A_k \,.$$

These scaled matrices have unit determinants.[9] Applying K-JET to the scaled matrices $\{\tilde{A}_k\}$, results in triangular matrices $\{\tilde{R}_k\}$ with equal diagonals, and a set of unitary matrices $\{U_k\}$ and V. This, in turn, suggests the following joint decomposition of the matrices $\{A_k\}$:

$$A_k = U_k R_k V^{\dagger} \,, \qquad k = 1, \ldots, K \,,$$

where

$$R_k \triangleq \left| \det(A_k) \right|^{1/n} \tilde{R}_k \,.$$

---

[8] In practice, the codebooks should not be identical, though they can, for example, be derived from a common base codebook via scrambling.

[9] Up to a scalar phase which can be absorbed in the left-unitary matrices $\{U_k\}$.

Thus, K-JET applied to matrices having non-equal determinants, gives rise to triangular matrices having *proportional* diagonals (instead of the equal diagonals, in the equal-determinant case). This is illustrated in the following example.

Example 2: Consider the following two matrices having *non-equal* determinants:

$$A_1 = \begin{pmatrix} 2 & 1 \\ 0 & 8 \end{pmatrix} = 4 \underbrace{\begin{pmatrix} 0.5 & 0.25 \\ 0 & 2 \end{pmatrix}}_{\tilde{A}_1}, \qquad \det(A_1) = 16 ,$$

$$A_2 = \begin{pmatrix} 5 & -2 \\ 0 & 5 \end{pmatrix} = 5 \underbrace{\begin{pmatrix} 1 & -0.4 \\ 0 & 1 \end{pmatrix}}_{\tilde{A}_2}, \qquad \det(A_2) = 25 .$$

By applying JET to $\tilde{A}_1$ and $\tilde{A}_2$, we obtain the following triangular matrices:

$$\tilde{R}_1 \approx \begin{pmatrix} 1.20 & -1.48 \\ 0 & 0.84 \end{pmatrix} \Rightarrow R_1 = 4 \begin{pmatrix} 1.20 & -1.48 \\ 0 & 0.84 \end{pmatrix}$$
$$\tilde{R}_2 \approx \begin{pmatrix} 1.20 & -0.17 \\ 0 & 0.84 \end{pmatrix} \Rightarrow R_2 = 5 \begin{pmatrix} 1.20 & -0.17 \\ 0 & 0.84 \end{pmatrix} ,$$

by applying the unitary matrices

$$U_1 \approx \begin{pmatrix} -0.22 & -0.98 \\ 0.98 & -0.22 \end{pmatrix}, \quad U_2 \approx \begin{pmatrix} -0.87 & -0.49 \\ 0.49 & -0.87 \end{pmatrix}$$
$$V \approx \begin{pmatrix} -0.81 & -0.58 \\ 0.58 & -0.81 \end{pmatrix} .$$

Hence, the original matrices A_1 and A_2 can be simultaneously triangularized as follows

$$A_1 \approx 4 U_1 \begin{pmatrix} 1.20 & -1.48 \\ 0 & 0.84 \end{pmatrix} V^\dagger \approx U_1 \begin{pmatrix} 4.79 & -5.91 \\ 0 & 3.34 \end{pmatrix} V^\dagger ,$$

$$A_2 \approx 5 U_2 \begin{pmatrix} 1.20 & -0.17 \\ 0 & 0.84 \end{pmatrix} V^\dagger \approx U_2 \begin{pmatrix} 5.99 & -0.85 \\ 0 & 4.18 \end{pmatrix} V^\dagger .$$

Definition 5 (K-GMD): The K-GMD is a special case of the K-JET where the entries on the diagonal are constant, namely

$$[R_k]_{jj} = \sqrt[n]{\det A_k} , \qquad \begin{array}{l} k = 1, \ldots, K \\ j = 1, \ldots, n . \end{array}$$

In this case the resulting upper triangular matrices will be denoted by T_k (instead of R_k for the general K-JET):

$$A_k = U_k T_k V^\dagger , \quad k = 1, \ldots, K . \tag{16}$$

Remark 7: For $K = 1$, 1-GMD reduces to the GMD of (6). The proof of the existence of a JET of two matrices A_1 and A_2 [16] is based upon applying the GMD (6) to the (single) matrix $A_1 A_2^{-1}$. This technique is generalized for more matrices in the next lemma.

Lemma 1 (Equivalence of Square K-GMD and (K+1)-JET): Let $A_1, \ldots, A_{K+1}$ be $n \times n$ full-rank complex-valued matrices with equal determinants, and define the K matrices:

$$B_k = A_k A_{K+1}^{-1} , \qquad k = 1, \ldots, K . \tag{17}$$

Then the following two statements are equivalent:

1) There exist $K+1$ unitary matrices $U_1, \ldots, U_K, U_{K+1}$, of dimensions $n \times n$, such that
$$U_k^\dagger B_k U_{K+1} = T_k , \qquad k = 1, \ldots, K , \tag{18}$$
where $\{T_k\}$ are $n \times n$ upper triangular with all diagonal entries equal to 1.
2) There exist $K+2$ unitary matrices $U_1, \ldots, U_{K+1}, V$, of dimensions $n \times n$, such that
$$U_k^\dagger A_k V = R_k , \qquad k = 1, \ldots, K+1 ,$$
where $\{R_k\}$ are $n \times n$ upper triangular with equal diagonals, as in (15).

Proof: First, assume that statement 2 holds. Thus, there exist $K+2$ unitary matrices $U_1, \ldots, U_{K+1}, V$, of dimensions $n \times n$, such that

$$U_k^\dagger A_k V = R_k , \qquad k = 1, \ldots, K+1 ,$$

where $\{R_k\}$ are $n \times n$ upper triangular with equal diagonals. This implies that

$$\begin{aligned} U_k^\dagger B_k U_{K+1} &= U_k^\dagger A_k A_{K+1}^{-1} U_{K+1} \\ &= U_k^\dagger A_k V V^\dagger A_{K+1}^{-1} U_{K+1} \\ &= R_k R_{K+1}^{-1} \\ &= T_k , \end{aligned}$$

where T_k is upper triangular with all the diagonal elements equal to 1, which results in statement 1.

Now, assume that statement 1 holds. Perform the QR decomposition on the matrix $A_{K+1}^{-1} U_{K+1}$:

$$A_{K+1}^{-1} U_{K+1} = V R ,$$

where V is a unitary matrix of dimensions $n \times n$, and R is an $n \times n$ upper triangular matrix. Thus, substituting (17), we obtain the following equalities:

$$\begin{aligned} U_k^\dagger A_k V R &= U_k^\dagger A_k A_{K+1}^{-1} U_{K+1} \\ &= U_k^\dagger B_k U_{K+1} , \qquad k = 1, \ldots, K , \end{aligned}$$

which, according to (18), is equal to

$$U_k^\dagger A_k V R = T_k , \qquad k = 1, \ldots, K . \tag{19}$$

On the other hand, we have

$$U_{K+1}^\dagger A_{K+1} V R = U_{K+1}^\dagger A_{K+1} A_{K+1}^{-1} U_{K+1} \tag{20a}$$
$$= U_{K+1}^\dagger U_{K+1} = I . \tag{20b}$$

Multiplying (19) and (20) by R^{-1} on the right yields:

$$\begin{aligned} U_k^\dagger A_k V &= T_k R^{-1} , \qquad k = 1, \ldots, K \\ U_{K+1}^\dagger A_{K+1} V &= R^{-1} . \end{aligned}$$

Since T_k are upper triangular with only 1s on the diagonal, the matrices $R_k \triangleq T_k R^{-1}$ ($k = 1, \ldots, K$) and $R_{K+1} \triangleq R^{-1}$ have equal diagonals, which completes the proof. ■

Remark 8: As a consequence of Lemma 1, if it is possible to perform K-GMD on *any* K full rank square matrices having the same determinant, then it is also possible to perform $(K+1)$-JET on *any* $K+1$ full rank square matrices of the same dimensions and the same determinant, and vice versa. In particular, since 1-GMD is always possible, it is also always possible to perform 2-JET on *any* two full rank square matrices of the same dimensions and equal determinants.

Remark 9: The condition of equal determinants in Definitions 4 and 5 may be replaced with a slightly weaker condition of equal absolute values of the determinants, i.e.,

$$|\det(A_1)| = |\det(A_2)| = \cdots = |\det(A_K)| \ .$$

This is easily achieved by multiplying by additional diagonal phase matrices on the left in (14) and (16).

B. Geometric Interpretation of the JET

Following the geometric interpretation of the GTD in Section IV-B, we give a geometric interpretation of the JET for the special case of 2×2 matrices:

$$A_1 \triangleq \begin{pmatrix} \boldsymbol{a}^{(1)} & \boldsymbol{b}^{(1)} \end{pmatrix} = \begin{pmatrix} a_x^{(1)} & b_x^{(1)} \\ a_y^{(1)} & b_y^{(1)} \end{pmatrix}$$

$$A_2 \triangleq \begin{pmatrix} \boldsymbol{a}^{(2)} & \boldsymbol{b}^{(2)} \end{pmatrix} = \begin{pmatrix} a_x^{(2)} & b_x^{(2)} \\ a_y^{(2)} & b_y^{(2)} \end{pmatrix} ,$$

where $\boldsymbol{a}^{(i)}$ and $\boldsymbol{b}^{(i)}$ are the first and second columns of A_i $(i = 1, 2)$, respectively. The interpretation for the general case is a simple extension of the 2×2 case. As in Section IV-B, we assume, without loss of generality, that $\det(A_1) = \det(A_2) = 1$.

By multiplying both matrices A_1 and A_2 on the right by the same rotation matrix V (7), we obtain $(i = 1, 2)$

$$A_i V \triangleq \begin{pmatrix} \tilde{\boldsymbol{a}}^{(i)} & \tilde{\boldsymbol{b}}^{(i)} \end{pmatrix} = \begin{bmatrix} a_x^{(i)} \cos\theta_r + b_x \sin\theta_r & a_x^{(i)} \cos(\theta_r + \frac{\pi}{2}) + b_x^{(i)} \sin(\theta_r + \frac{\pi}{2}) \\ a_y^{(i)} \cos\theta_r + b_y \sin\theta_r & a_y^{(i)} \cos(\theta_r + \frac{\pi}{2}) + b_y^{(i)} \sin(\theta_r + \frac{\pi}{2}) \end{bmatrix} . \tag{21}$$

That is, we obtain two ellipses of equal area (absolute value of determinant), centered at the origin (see Figure 5a). The norms of the first column vectors in (21), $\tilde{\boldsymbol{a}}^{(1)}$ and $\tilde{\boldsymbol{a}}^{(2)}$, are 2π-cyclic continuous functions of θ_r. Thus, using the intermediate value theorem, there exists an angle θ_r (and in fact, four such angles per cycle) for which the norms of $\tilde{\boldsymbol{a}}^{(1)}$ and $\tilde{\boldsymbol{a}}^{(2)}$ are equal, as illustrated in Figure 5b.

Multiplying each of the resulting matrices, $A_i V$, on the left, by an appropriate rotation matrix $U_i^\dagger$, where

$$U_i = \begin{pmatrix} \cos\theta_\ell^{(i)} & -\sin\theta_\ell^{(i)} \\ \sin\theta_\ell^{(i)} & \cos\theta_\ell^{(i)} \end{pmatrix} ,$$

rotates both column vectors of $A_i V$ by the same angle, $\left(-\theta_\ell^{(i)}\right)$, without altering their norms. Thus, by choosing $\theta_\ell^{(i)}$, such that $U_i^\dagger \tilde{\boldsymbol{a}}^{(i)}$ are aligned with the x-axis, for both $i = 1, 2$, we achieve the desired decomposition, as depicted in Figures 5c and 5d.

Remark 10: JET of more than two matrices is not possible, in general. This may be seen in the 2×2 case, that while every two ellipses must intersect for some value of θ_r, due to the intermediate value theorem, there is no hope for simultaneous intersection of more trajectories.

C. MIMO Common-Message Broadcast Scheme via Matrix Decomposition

The scheme of Section IV-C can be generalized for the K-user BC channel (1) in a straightforward manner, by replacing the GTD (4) with the K-JET (14).

Let $C_{\boldsymbol{x}}$ be an admissible covariance matrix. As will be explained in Remark 12, we can assume without loss of generality that $I(H_1, C_{\boldsymbol{x}}) = \cdots = I(H_K, C_{\boldsymbol{x}})$. The following scheme achieves the rate $I(H_i, C_{\boldsymbol{x}})$. Therefore, the common-message BC capacity (2) can be achieved by an appropriate choice of the matrix $C_{\boldsymbol{x}}$.

Applying Definitions 2 and 3 we define

$$\tilde{H}_k \triangleq \begin{pmatrix} H_k C_{\boldsymbol{x}}^{1/2} \\ I_{n_t} \end{pmatrix} , \tag{22}$$

$$\tilde{H}_k = Q_k G_k , \qquad k = 1, \ldots, K , \tag{23}$$

where I_{n_t} is the $n_t \times n_t$ identity matrix, $C_{\boldsymbol{x}}$ is any admissible covariance matrix, the matrices $\tilde{H}_k$ are the augmented channel matrices, Q_k are $(n_r^{(k)} + n_t) \times n_t$ matrices with orthonormal columns, and G_k are the canonical channel matrices of dimensions $n_t \times n_t$ and are upper triangular with real positive diagonal elements.

Now, assume that there exists a K-JET of the matrices G_k:

$$G_k = U_k R_k V^\dagger , \qquad k = 1, \ldots, K ,$$

where R_k are upper triangular matrices whose diagonal values are equal to $r_1, \ldots, r_{n_t}$. Then, the same transmission scheme as in Section IV-C may be employed, where in step 3 the k-th receiver uses the matrices Q_k and U_k in (13).

Remark 11: As in Remark 3, if the K-JET in the above scheme is also a K-GMD (16), then the capacity (2) can be achieved using *the same* scalar codebook over all scalar sub-channels.

Remark 12: Consider the case where, for the optimal input covariance matrix $C_{\boldsymbol{x}}$, the mutual informations to the different users, $\{I(H_k, C_{\boldsymbol{x}})\}$, are not all equal. In this case, the common-message BC capacity (2) is limited to the minimum of these mutual informations:

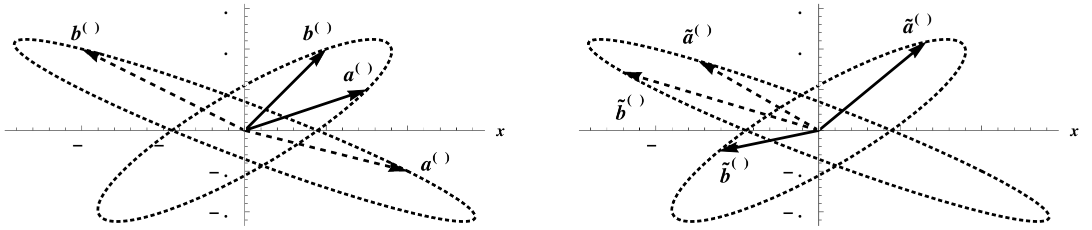
$$C\left(\{H_k\}_{k=1}^K, P\right) = \max_{C_{\boldsymbol{x}}} \min_{k=1,\ldots,K} I(H_k, C_{\boldsymbol{x}}) .$$

Rewriting these mutual informations in terms of the channel canonical matrices $\{G_k\}$:

$$\begin{aligned} I(H_k, C_{\boldsymbol{x}}) &= \log\det\left(I + H_k C_{\boldsymbol{x}} H_k^\dagger\right) \\ &= \log\det\left(I + H_k^\dagger C_{\boldsymbol{x}}^{1/2\dagger} C_{\boldsymbol{x}}^{1/2} H_k\right) \\ &= \log\det\left(\tilde{H}_k^\dagger \tilde{H}_k\right) \\ &= \log\det\left((Q_k G_k)^\dagger Q_k G_k\right) \\ &= \log\det\left(G_k^\dagger Q_k^\dagger Q_k G_k\right) \\ &= \log\det\left(G_k^\dagger G_k\right) \\ &= 2\log|\det(G_k)| , \end{aligned}$$

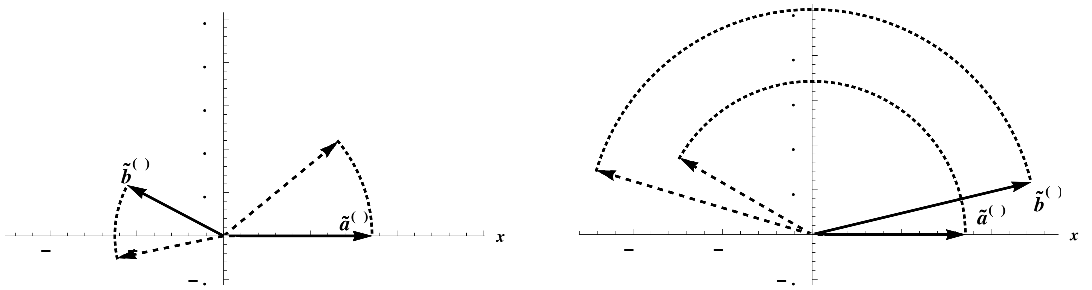
we have

$$C\left(\{H_k\}_{k=1}^K, P\right) = 2\log \min_{k=1,\ldots,K} \det(G_k) ,$$



(a) All possible column vectors of $A_i V$, where V is a rotation matrix; the original column vectors (for $V = I$) are depicted explicitly.

(b) Right rotation by $\theta_r \approx 1.34$, for which the resulting vectors $\tilde{\boldsymbol{a}}^{(1)}$ and $\tilde{\boldsymbol{a}}^{(2)}$ have equal norms.



(c) Left rotation by $\theta_\ell^{(1)} \approx 0.69$ of the first matrix, for which the first vector is aligned with the x-axis.

(d) Left rotation by $\theta_\ell^{(2)} \approx 2.62$ of the first matrix, for which the first vector is aligned with the x-axis.

Fig. 5: JET for $\boldsymbol{a}^{(1)} = (3/2, 1/2)^\dagger, \boldsymbol{b}^{(1)} = (1,1)^\dagger, \boldsymbol{a}^{(2)} = (2, -0.5)^\dagger, \boldsymbol{b}^{(2)} = (-2, 1)^\dagger$

where the absolute value operation may be dropped as explained in Remark 2.

Thus, the common-message BC capacity is dictated by the user having the minimal $\det(G_k)$.

Applying K-JET to the matrices $\{G_k\}$, results in *proportional* diagonal elements (in contrast to the equal diagonals resulting when all mutual informations are equal; see Remark 6). Since these effective diagonal entries correspond to the effective SNRs of the effective scalar sub-channels observed by each user, this implies, in turn, that the users having larger mutual information have larger effective SNRs. However, since the common-message BC capacity is limited to the minimum of the mutual informations, the excess SNRs of the users with larger mutual informations (and $\det(G_k)$) has no effect on achievable rate.

This "bottleneck phenomenon" is illustrated in the following example.

Example 3 (Example 2 Continued): Consider the two channel canonical matrices G_1 and G_2 (replacing A_1 and A_2 in Example 2).

$$G_1 = \begin{pmatrix} 2 & 1 \\ 0 & 8 \end{pmatrix} = 4 \begin{pmatrix} 0.5 & 0.25 \\ 0 & 2 \end{pmatrix}, \qquad \det(G_1) = 16,$$
$$G_2 = \begin{pmatrix} 5 & -2 \\ 0 & 5 \end{pmatrix} = 5 \begin{pmatrix} 1 & -0.4 \\ 0 & 1 \end{pmatrix}, \qquad \det(G_2) = 25.$$

By applying JET to G_1 and G_2 we obtain

$$R_1 = 4 \begin{pmatrix} 1.20 & -1.48 \\ 0 & 0.84 \end{pmatrix}$$
$$R_2 = 5 \begin{pmatrix} 1.20 & -0.17 \\ 0 & 0.84 \end{pmatrix}$$

The corresponding common-message BC capacity is, therefore,

$$\begin{aligned} C &= 2 \log \min \left\{ \det(G_1), \det(G_2) \right\} \\ &\approx \underbrace{2 \log (4 \times 1.20)}_{\text{Rate of stream 1}} + \underbrace{2 \log (4 \times 0.84)}_{\text{Rate of stream 2}} \approx 8. \end{aligned}$$

Thus, the rates of the two streams are dictated by user 1, whereas user 2 has excess effective SNR in each of the streams.

Remark 13 (Decoding Order): Recall that in the single-user case, there is no loss (in terms of achievable rates) in restricting attention to upper triangular decomposition at the receiver, since any ordering can be represented as a permutation of the matrix R in (12), namely,

$$G = U \Pi R \Pi^\dagger V^\dagger, \tag{24}$$

where Π is a permutation matrix. Since permutation matrices are unitary, (24) falls under the framework (12) without permutations. In the multi-user case, on the other hand, each receiver can choose a different decoding order, which implies that the different permutation matrices cannot be absorbed in the (single) matrix V. Hence, there is a loss of generality in the proposed scheme. This restriction is removed in Section VIII-B.

D. Perfect 2-GMD for 2×2 Matrices

In this section we provide necessary and sufficient conditions for the existence of 2-GMD for 2×2 matrices. The conditions are stated in the following theorem. As explained in Remark 12, we can assume without loss of generality that both matrices have determinants equal to 1. According to Lemma 1, this also provides a necessary and sufficient condition for the existence of a 3-JET for 2×2 matrices.

Theorem 2 (2-GMD for 2×2 Matrices): Let A_1 and A_2 be *complex-valued* 2×2 matrices with determinants equal to 1. Then, there exist complex-valued 2×2 unitary matrices U_1, U_2, V such that:

$$U_k^\dagger A_k V = \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}, \quad k = 1, 2, \tag{25}$$

if and only if the following inequality is satisfied:

$$F_1\left(A_1^\dagger A_1 - I, A_2^\dagger A_2 - I\right) \geq 0, \tag{26}$$

where

$$F_1(S_1, S_2) \triangleq \det\left(S_1 \operatorname{adj}(S_2) - S_2 \operatorname{adj}(S_1)\right), \tag{27}$$

and $*$ represents an arbitrary value value (which may differ between the two matrices).

Remark 14: Even in the case where the matrices A_1 and A_2 are real-valued, the resulting unitary matrices U_1, U_2, and V are, in general, *complex-valued*. In fact, if A_1, A_2 are real valued, then it can be easily shown that the matrices U_1, U_2 and V are real-valued if and only if (26) holds *with equality*. In Section VI-C we show how to obtain a communication scheme that involves only real-valued orthogonal transformations, under the same condition (26), using a space–time structure.

Remark 15: Using this theorem, a sufficient and necessary condition for the existence of a 2-GMD for two 3×3 *diagonal* matrices can be derived. The method of this derivation is demonstrated via an example of the "rateless" problem with three rates in Section V-E2.

The following lemma, the proof of which is given in Appendix A, will be used in the proof of the theorem.

Lemma 2: Let S_1 and S_2 be complex-valued Hermitian 2×2 matrices. Then, there exists a complex-valued vector $\boldsymbol{v} \in \mathbb{C}^2$, such that

$$\begin{aligned} \boldsymbol{v}^\dagger S_1 \boldsymbol{v} &= 0 \\ \boldsymbol{v}^\dagger S_2 \boldsymbol{v} &= 0 \\ \boldsymbol{v} &\neq 0, \end{aligned}$$

if and only if the following conditions hold:

$$\det(S_1) \leq 0 \tag{28a}$$

$$\det(S_2) \leq 0 \tag{28b}$$

$$F_1(S_1, S_2) \geq 0, \tag{28c}$$

where F_1 is defined as in (27).

Proof of Theorem 2: Let V be a 2×2 unitary matrix, and denote by $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ the first and second columns of V, respectively. Note that

$$A_k V = (A_k \boldsymbol{v}_1 | A_k \boldsymbol{v}_2), \quad k = 1, 2.$$

We now perform the QR decomposition on the above matrices:

$$A_1 V = U_1 T_1 \tag{29a}$$

$$A_2 V = U_2 T_2, \tag{29b}$$

where U_1, U_2 are unitary and T_1, T_2 are upper triangular. Since we have

$$T_k = U_k^\dagger A_k V = \left(U_k^\dagger A_k v_1 \;\middle|\; U_k^\dagger A_k v_2 \right),$$

and the norm of $A_k v_1$ equals that of $U_k^\dagger A_k v_1$, the upper-left element of T_1 and T_2 is equal to 1,

$$T_k = \begin{pmatrix} 1 & * \\ 0 & * \end{pmatrix}, \quad k = 1, 2, \tag{30}$$

if and only if:

$$\begin{aligned} \|A_1 \boldsymbol{v}_1\| &= 1 \\ \|A_2 \boldsymbol{v}_1\| &= 1. \end{aligned}$$

Also, since V is required to be unitary, the norm of $\boldsymbol{v}_1$ must equal 1:

$$\|\boldsymbol{v}_1\| = 1.$$

Note that for every $\boldsymbol{v}_1$, we can choose a unit-norm vector $\boldsymbol{v}_2$ that spans the subspace orthogonal to $\boldsymbol{v}_1$, thus constructing a unitary matrix V. Also, since V is unitary, $\det(A_1 V) = \det(A_2 V) = 1$, and therefore from (30) it follows that the bottom-right element also equals 1.

Combining the above observations, it follows that there exists a 2×2 unitary matrix V such that the decomposition (29) is possible, where T_1, T_2 have only 1s on their diagonals, if and only if the first column of V, denoted by $\boldsymbol{v}_1$, satisfies the following three equations:

$$\begin{aligned} \boldsymbol{v}_1^\dagger A_1^\dagger A_1 \boldsymbol{v}_1 &= 1 \\ \boldsymbol{v}_1^\dagger A_2^\dagger A_2 \boldsymbol{v}_1 &= 1 \\ \boldsymbol{v}_1^\dagger \boldsymbol{v}_1 &= 1, \end{aligned}$$

or equivalently,

$$\begin{aligned} \boldsymbol{v}_1^\dagger (A_1^\dagger A_1 - I) \boldsymbol{v}_1 &= 0 \\ \boldsymbol{v}_1^\dagger (A_2^\dagger A_2 - I) \boldsymbol{v}_1 &= 0 \\ \boldsymbol{v}_1^\dagger \boldsymbol{v}_1 &= 1. \end{aligned}$$

Note that since $\det(A_1) = \det(A_2) = 1$, we have

$$\begin{aligned} \det\left(A_1^\dagger A_1 - I\right) &\leq 0 \\ \det\left(A_2^\dagger A_2 - I\right) &\leq 0. \end{aligned}$$

Using this result along with the result of Lemma 2 with $S_k = A_k^\dagger A_k - I$, proves the theorem. ■

Corollary 1: Theorem 2 can easily be generalized as follows: for any $r > 0$, there exist three complex-valued 2×2 unitary matrices U_1, U_2, V such that:

$$U_k^\dagger A_k V = \begin{pmatrix} r & * \\ 0 & 1/r \end{pmatrix}, \quad k = 1, 2,$$

if and only if the following conditions are satisfied:

$$\begin{aligned} \det\left(A_1^\dagger A_1 - r^2 I\right) &\leq 0 \\ \det\left(A_2^\dagger A_2 - r^2 I\right) &\leq 0 \\ F_1\left(A_1^\dagger A_1 - r^2 I, A_2^\dagger A_2 - r^2 I\right) &\geq 0. \end{aligned}$$

The proof of the corollary follows along the same line as that of Theorem 2 with obvious modifications.

E. Example: “Rateless” Codes over the AWGN Channel

We now consider the problem of constructing scalar Gaussian rateless codes, treated in [35].[10] The constructed codes are designed for a complex AWGN channel,

$$\boldsymbol{y}_l = \alpha \boldsymbol{x}_l + \boldsymbol{z}_l \,, \qquad l = 1, 2, \dots \,, \tag{31}$$

where α is a channel gain that varies from receiver to receiver, $\boldsymbol{x}_l$ is the channel input vector of M symbols, $\boldsymbol{z}_l$ is a noise vector of M i.i.d. complex Gaussian random variables, each of variance 1, and $\boldsymbol{y}_l$ is the vector of M channel output symbols. The channel input is average-power limited, without loss of generality, to power 1.

We assume that α can take one of K possible values, such that a gain of α_k implies that the message should be decodable using only the first k received blocks.[11] The gains are such that, for any value of k, the total capacity is the same:

$$C = k \log(1 + |\alpha_k|^2) \,, \qquad k = 1, 2, \dots, K \,.$$

This implies that the compound capacity is achieved by a white input distribution.

The scheme proposed in [35] consists of dividing the information message into L sub-messages (“layers”), encoding each sub-message using a (fixed-block) codebook, designed for a scalar AWGN channel, and sending in each block some linear combination of those codewords. In the sequel we will consider only the case where $K = L$, i.e., the number of codewords used by the scheme is equal to the highest possible number of blocks received by the receiver.

Alternatively, this problem can be viewed as a K-user MIMO common-message BC problem, as follows: the K transmission blocks (31) can be considered as a single transmission over a Gaussian MIMO channel, with channel matrix

$$H = \begin{pmatrix} \alpha_k & 0 & \cdots & 0 \\ 0 & \alpha_k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_k \end{pmatrix} .$$

Since the k-th user is allowed to use only the first k blocks, this is equivalent to removing the last $K - k$ rows from the corresponding channel matrix, namely, the channel matrix of the k-th user becomes:

$$H_k = \left(\begin{array}{ccccccc} \multicolumn{4}{c}{\overbrace{\hphantom{\alpha_k \quad 0 \quad \cdots \quad 0}}^{k}} & \multicolumn{3}{c}{\overbrace{\hphantom{0 \quad \cdots \quad 0}}^{K-k}} \\ \alpha_k & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \alpha_k & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \alpha_k & 0 & \cdots & 0 \end{array} \right) . \tag{32}$$

Since the capacity-achieving distribution in this problem is white, this translates to an input covariance matrix which is a scaled identity matrix. Namely, $C_{\boldsymbol{x}} = I$.

Alternatively, the channel matrix of the k-th user can be viewed as a square $K \times K$ diagonal matrix, where the last $K - k$ diagonal elements are forced to be zeros:

$$H_k = \left(\begin{array}{ccccccc} \multicolumn{4}{c}{\overbrace{\hphantom{\alpha_k \quad 0 \quad \cdots \quad 0}}^{k}} & \multicolumn{3}{c}{\overbrace{\hphantom{0 \quad \cdots \quad 0}}^{K-k}} \\ \alpha_k & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \alpha_k & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \alpha_k & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{array} \right) .$$

This alternative representation yields the same results as the representation (32).

We now recover the results of [35], giving explicit constructions for $K = 2, 3$.

1) Two Rates ($K = 2$): Specializing the problem to the case of one (possible) incremental redundancy block ($K = 2$), the two channel matrices are (same as H_1 and H_3 in Example 1)

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

where α_1, α_2 are values satisfying

$$\log(1 + |\alpha_1|^2) = 2 \log(1 + |\alpha_2|^2) = C \,.$$

Applying the scheme of Section V-C yields the following precoding matrix [11]:

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

which coincides with the result in [35, Sec. III].

2) Three Rates: The case of $K = 3$ was also treated in [35], where a condition for which a “perfect” scheme exists was derived. We will now shed light on this condition.

Again, representing the problem as a MIMO common-message BC one, the three possible channel matrices are:

$$\begin{aligned} H_1 &= \begin{pmatrix} \alpha_1 & 0 & 0 \end{pmatrix} , \\ H_2 &= \begin{pmatrix} \alpha_2 & 0 & 0 \\ 0 & \alpha_2 & 0 \end{pmatrix} , \\ H_3 &= \begin{pmatrix} \alpha_3 & 0 & 0 \\ 0 & \alpha_3 & 0 \\ 0 & 0 & \alpha_3 \end{pmatrix} , \end{aligned}$$

where $\alpha_1, \alpha_2, \alpha_3$ are values satisfying

$$\log(1 + |\alpha_1|^2) = 2 \log(1 + |\alpha_2|^2) = 3 \log(1 + |\alpha_3|^2) = C \,.$$

The channel canonical matrices, as defined in (23), are:

$$\begin{aligned} G_1 &= \begin{pmatrix} 2^{\frac{C}{2}} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} , \\ G_2 &= \begin{pmatrix} 2^{\frac{C}{4}} & 0 & 0 \\ 0 & 2^{\frac{C}{4}} & 0 \\ 0 & 0 & 1 \end{pmatrix} , \end{aligned}$$

---

[10] A numerical derivation of the precoding matrix V in the case of a rateless code (even for parameters for which a perfect decomposition is not possible) is available in [36].

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

$$G_3 = \begin{pmatrix} 2^{\frac{C}{3}} & 0 & 0 \\ 0 & 2^{\frac{C}{3}} & 0 \\ 0 & 0 & 2^{\frac{C}{3}} \end{pmatrix} .$$

Since G_3 is a s scaled identity matrix, we are in fact seeking a 2-GMD of the remaining two matrices. Thus, denoting $b = 2^{\frac{C}{12}}$, we need to perform a 2-GMD on the following two 3×3 matrices,

$$G_1 = \begin{pmatrix} b^6 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} , \quad G_2 = \begin{pmatrix} b^3 & 0 & 0 \\ 0 & b^3 & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

Equivalently, dividing both matrices by b^2, we are seeking a 2-GMD of the following two 3×3 matrices, both having a determinant equal to 1:

$$A_1 = \begin{pmatrix} b^4 & 0 & 0 \\ 0 & b^{-2} & 0 \\ 0 & 0 & b^{-2} \end{pmatrix}$$

$$A_2 = \begin{pmatrix} b & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & b^{-2} \end{pmatrix} .$$

As shown in Appendix C, this reduces to performing 2-GMD on the following two 2×2 matrices:

$$\tilde{A}_1 = \begin{pmatrix} \frac{\sqrt{1-b^2+b^8}}{b^2} & \frac{b^6-1}{b\sqrt{(1-b^2+b^8)(1+b^2+b^4)}} \\ 0 & \frac{b^2}{\sqrt{1-b^2+b^8}} \end{pmatrix}$$

$$\tilde{A}_2 = \begin{pmatrix} b & 0 \\ 0 & b^{-1} \end{pmatrix} .$$

We have:

$$F_1(\tilde{A}_1^\dagger \tilde{A}_1 - I, \tilde{A}_2^\dagger \tilde{A}_2 - I) = \\ -\frac{(b^2-1)^4(b^2+1)^2(1+b^2+b^4)(1-3b^2+b^4)}{b^{12}} ,$$

where F_1 is defined in (27). According to Theorem 2, there exists a solution if and only if this value is non-negative, namely,

$$1 - 3 \cdot 2^{\frac{C}{6}} + 2^{\frac{C}{3}} \leq 0 .$$

This condition is satisfied if and only if:

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

which coincides with the result that was obtained in [35], where arduous algebraic manipulations were used to obtain this condition.

Finally, we note that there exists a similar result for four rates ($K = 4$). In this case, it is shown in [37] that there exists a perfect solution if and only if the rate C does not exceed a critical rate, which equals approximately 10.55.

F. Example: Arbitrarily Permuted Parallel Channels

The problem of transmitting information over arbitrarily permuted parallel channels was studied by Willems and Gorokhov [38] and by Hof et al. [39]. In this point-to-point scenario, the transmitter is connected to the receiver via M parallel memoryless channels, sharing the same input alphabet, the transition matrices of which are known at the transmitter but not their order. Namely, at each time instant, the transmitter generates M input symbols to be sent over the M parallel channels, and these symbols are then *permuted* by a one-to-one-mapping (permutation) $\pi \in S_M$ from $\{1, \ldots, M\}$ onto itself.

The permutation π is arbitrary, yet constant throughout the transmission block, and is known to the receiver but not to the transmitter. The aim of the receiver is to recover the transmitted message with arbitrarily small error probability. This channel model is of relevance in scenarios where the gains of the channels are generated according to an i.i.d. distribution, and one may choose the "design gains" so as to minimize the outage probability; for details see [38, Sec. VII].

In this section we describe a practical capacity-achieving scheme for the *Gaussian* case, described by

$$y_m = \alpha_m x_m + z_m , \quad m = 1, 2, \ldots, M , \tag{33}$$

where x_m is the input to the m-th channel and is subject to a power constraint[12]

$$\mathbb{E}\left(|x_m|^2\right) \leq 1 , \tag{34}$$

y_m is the output of the m-th channel, and $\{z_m\}$ are i.i.d. circularly-symmetric Gaussian variables with unit variance, independent of $\{x_m\}$. The gains $\{\alpha_m\}$ are known to the receiver, whereas the transmitter knows the gains up to an unknown permutation. Namely, the transmitter knows the gains but not their *order*.

The M parallel channels (33) may be regarded as a single MIMO channel,

$$\boldsymbol{y} = H\boldsymbol{x} + \boldsymbol{z} ,$$

where $\boldsymbol{x}$ is the channel input vector of length M, and $\boldsymbol{z}$ is a circularly-symmetric white Gaussian random vector of length M and identity covariance matrix. The channel matrix H is an $M \times M$ diagonal matrix, which is known at the receiver:

$$H = \begin{pmatrix} \alpha_1 & 0 & \cdots & 0 \\ 0 & \alpha_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_M \end{pmatrix} .$$

The transmitter knows the matrix H, up to the unknown order of the diagonal elements.

The latter is, in turn, equivalent to broadcasting the same (common) message to $K = M!$ receivers simultaneously, where the channel matrix to user k is

$$H_k \triangleq \begin{pmatrix} \alpha_{\pi_k(1)} & 0 & \cdots & 0 \\ 0 & \alpha_{\pi_k(2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_{\pi_k(M)} \end{pmatrix} ,$$

and $\pi_k \in S_K$ is a permutation which is different for each user. As a consequence, this transmission problem may be

---

[12] Alternatively, the individual power constraints can be replaced by a sum-power constraint. However, both cases reduce to the same result.

regarded as a special case of the common-message Gaussian MIMO broadcast one. Under the power constraint (34), the capacity of this common-message BC scenario is obtained by taking $C_{\boldsymbol{x}} = I$ in (2), namely,

$$C = \sum_{m=1}^{M} \log\left(1 + |\alpha_m|^2\right) .$$

We now show how the same transmission schemes as described in the previous sections can be used in this scenario for $M = 2 \to K = 2$ and $M = 3 \to K = 6$. We give here only the results without proofs. The full details are given in [12].

For the case of $M = 2$, the channel can be in one of two “states”:

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

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

where $\alpha_1, \alpha_2 \geq 0$ are known.

Since there are only two options for the channel matrix H, the capacity in this case can be achieved using JET, as described in Section V-A. Specifically, capacity is achieved by choosing the precoding matrix to be the (scaled) Hadamard matrix (which coincides with the 2×2 DFT matrix):

$$V = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} .$$

Similarly, in the case of three parallel channels ($M = 3$), we have:

$$H = \begin{pmatrix} \alpha_1 & 0 & 0 \\ 0 & \alpha_2 & 0 \\ 0 & 0 & \alpha_3 \end{pmatrix} ,$$

where $\alpha_1, \alpha_2, \alpha_3 \geq 0$ are known, up to an unknown permutation. In this case, capacity is achieved by the following precoding matrix, which is the 3×3 DFT matrix:

$$V = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 1 & 1 \\ 1 & e & e^{-1} \\ 1 & e^{-1} & e \end{pmatrix} ,$$

where $e \triangleq e^{2\pi i/3}$.

For $M \geq 4$, capacity is no longer achieved using a DFT precoding matrix. Nevertheless, extension of the above scheme to $4 \leq M \leq 6$ is possible [12] by utilizing algebras of higher dimensions, such as the quaternion algebra. These algebras can be materialized using a space–time structure over the complex or real fields. Moreover, the complex field may be represented over the reals by incorporating time extensions, as is explained in the sequel — in Section VI-C.

In the next section we describe the space–time structure that is used for the construction of joint triangularization of more than two matrices.

VI. SPACE–TIME TRIANGULARIZATION

A. Introduction

As indicated by Theorem 2, joint (unitary) triangularization with constant diagonal values (K-GMD) is not always possible. However, even when the condition for joint triangularization does not hold, it is possible to gain more mathematical degrees of freedom 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 OSTBC [21], [22]. However, whereas space–time processing has traditionally been applied to an open-loop communication scenario, in the present work it will be applied to the closed-loop common-message BC problem.

We first recall the idea of linear space–time codes, also known as linear dispersion codes (see, e.g., [40]), which will be used as a building block for the proposed communication scheme. For this, we consider the point-to-point MIMO Gaussian channel, with an $n_r \times n_t$ channel matrix H,

$$\boldsymbol{y} = H\boldsymbol{x} + \boldsymbol{z} .$$

We now utilize transmission over N consecutive blocks, assuming that the channel matrix H does not change between these blocks. This is equivalent to sending time-extended symbols over the following *time-extended channel*:

$$\mathpzc{y} = \mathcal{H}\mathpzc{x} + \mathpzc{z} . \tag{35}$$

The time-extended vectors $\mathpzc{x}, \mathpzc{y}, \mathpzc{z}$ are composed of N “physical” (concatenated) input, output, and noise vectors, respectively, and $\mathcal{H}$ is the $(Nn_r) \times (Nn_t)$ *time-extended channel matrix* defined as

$$\mathcal{H} = \lceil H \rfloor_{\otimes N} , \tag{36}$$

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:

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

In linear space–time modulation (also known as “space–time coding”) the extended input vector $\mathpzc{x}$ is obtained by linearly combining independent streams of data symbols.[13] Of special interest are modulations that possess a certain structure with the aim of facilitating decoding. Such a family includes OSTBCs, and in particular Alamouti modulation [21]. When using an OSTBC, the transmitter applies a unitary transformation, *which does not depend on the channel matrix H*, to the data symbols, and the receiver applies another orthogonal transformation to the channel output, such that the effective channel matrix is transformed into a *diagonal* form, over which communication is possible using off-the-shelf codes designed for *scalar* AWGN channels. Thus, *simultaneous* diagonalization of all possible channel matrices, is attained.

---

[13] The transformation may, more generally, be taken to be linear over the reals. Nevertheless, for the purposes of this paper it suffices to consider only linear transformations over the complex numbers.

Unfortunately, OSTBCs that universally achieve the white-input capacity of every channel, as is the case for Alamouti modulation, do not exist for MISO channels with more than 2 transmit antennas, let alone for MIMO channels [22], [41].

In this work, we use the idea of space–time modulation, but instead of diagonalizing the channel matrices, we are content with *triangularization*. This, in turn, requires the employment of another ingredient to the communication scheme, namely, successive interference cancellation at the receivers. Further, in contrast to OSTBC, where the same transformation is applied to a continuum of channels, the proposed approach is applicable to only a finite number of channel matrices.

B. Space–Time Common-Message BC Scheme

We now introduce the space-time common-message BC scheme. Recall the common-message broadcast MIMO channel (1) with K users and n_t transmit antennas. We now utilize transmission over N consecutive blocks, assuming that the channel matrices do not change between these blocks. This is equivalent to sending extended symbols over the following *time-extended channels*:

$$\mathscr{y}_k = \mathcal{H}_k \mathscr{x} + \mathscr{z}_k \,, \qquad k = 1, \ldots, K \,,$$

where the time-extended vectors $\mathscr{x}, \mathscr{y}_k, \mathscr{z}_k$ and time-extended matrices $\mathcal{H}_k$ are defined as in (35) and (36).[14] The power constraint now becomes $\mathbb{E}\left[\mathscr{x}^\dagger \mathscr{x}\right] \leq NP$.

Let $C_{\boldsymbol{x}}$ be an $n_t \times n_t$ covariance matrix satisfying $\operatorname{tr}(C_{\boldsymbol{x}}) \leq P$. As explained in Remark 12, we can assume without loss of generality that

$$I(H_1, C_{\boldsymbol{x}}) = \cdots = I(H_K, C_{\boldsymbol{x}}) = C \,.$$

Define the matrices $\tilde{H}_k$, Q_k, and G_k as in (22) and (23). Further define the following time-extended channel canonical matrices:

$$\mathcal{G}_k \triangleq \lceil G_k \rfloor_{\otimes N} \,, \quad k = 1, \ldots, K \,.$$

Now, assume that there exists a K-JET of the matrices $\mathcal{G}_k$:

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

where $\mathcal{R}_k$ are upper triangular matrices whose diagonal values are equal to $r_1, \ldots, r_{n_t N}$. Then, the same transmission scheme as in Section IV-C can be employed, with the following replacements:

- The transmitted vector $\boldsymbol{x}$ is replaced by the time-extended vector $\mathscr{x}$
- The received vector $\boldsymbol{y}$ is replaced by the time-extended vector $\mathscr{y}_k$
- In step 3, the k-th user uses the matrix $\mathcal{U}_k$ instead of U in (13), and the matrix $\tilde{Q}$ is replaced with its time-extended version, $\left\lceil \tilde{Q}_k \right\rfloor_{\otimes N}$, where $\tilde{Q}_k$ consists of the first n_t rows of Q_k.

C. Space–Time 2-GMD for 2×2 Matrices

We now consider the special case where the transmitter is equipped with 2 antennas, and we are interested in performing 2-GMD, or alternatively, 3-JET, on the extended matrices.

As we saw in Section V-D, 2-GMD of 2×2 matrices is not always possible. This raises the question whether we can exploit the space–time structure to perform 2-GMD on the extended matrices, even in cases where 2-GMD of the original (not time-extended) matrices is not possible.

For a general number of antennas n_t, we know that space–time structures can sometimes enable GMD in cases where it is not possible without time extensions (see, e.g., [12]). However, in some cases, space–time structures cannot help. Such is the case for $n_t = 2$, as implied by the following theorem which is proved in Appendix D.

Theorem 3: Let A_1 and A_2 be *complex-valued* 2×2 matrices with determinants equal to 1, such that condition (26) does not hold (namely, there does not exist a 2-GMD of the matrices A_1 and A_2). Let $N \in \mathbb{N}$, and define the following extended matrices:

$$\mathcal{A}_k \triangleq \lceil A_k \rfloor_{\otimes N} \,, \quad k = 1, 2 \,.$$

Then, there also does not exist 2-GMD of the matrices $\mathcal{A}_1, \mathcal{A}_2$, for any value of $N \in \mathbb{N}$.

Consider now the case where the channel matrices are real-valued, and we allow the use of only orthogonal real-valued matrices U_k, V in the communication scheme. Then, if condition (26) holds, a space–time structure with $N = 2$ enables 2-GMD. This is explained in the following corollary.

Corollary 2: If condition (26) holds, then according to Theorem 2 we can perform 2-GMD on A_1, A_2 (25) with complex-valued unitary matrices U_1, U_2, V. In particular, we can assume that the three matrices U_1, U_2, V are of the following form:

$$\begin{pmatrix} a+bi & c+di \\ c-di & -a+bi \end{pmatrix} . \tag{37}$$

This implies that there exists a 2-GMD of the extended matrices with $N = 2$, $\mathcal{A}_1$ and $\mathcal{A}_2$, where the corresponding real-valued orthogonal matrices $\mathcal{U}_1, \mathcal{U}_2, \mathcal{V}$ are derived from U_1, U_2, V (37) as follows:

$$\begin{pmatrix} a & -b & c & -d \\ c & d & -a & -b \\ b & a & d & c \\ -d & c & b & -a \end{pmatrix} .$$

However, more extensions, i.e., $N \geq 3$, cannot help to construct (perfect) 2-GMD, due to Theorem 3.

---

[14] This technique can be extended to the case where the channel matrices are time-varying. In this case, the time-extended channel matrices of (36) are replaced by the block-diagonal matrices

$$\mathcal{H}_k = \begin{pmatrix} H_k^{(1)} & 0 & \cdots & 0 \\ 0 & H_k^{(2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & H_k^{(N)} \end{pmatrix} .$$

VII. NEARLY-OPTIMAL K-GMD

As indicated by Theorem 2, joint triangularization with constant diagonal values (K-GMD) is not always possible even if we consider time-extended channel matrices, as in Theorem 3.

The question is whether we may use the transmission scheme, presented in Section V-C, for the general multi-user problem. We now demonstrate that although perfect decomposition is not possible in general, we can still perform *nearly-optimal* triangularization, by utilizing multiple uses of the same channel realization.

There are many ways to define "nearly optimal". Commonly, this term refers to a problem with some optimization criterion, or some error criterion, where the optimization solution or the error are bounded, based on some statistical assumptions. Here, we refer to a different meaning. We strive for an explicit lower bound on the communication rate (without any statistical assumption on the generation processes of the channel matrices), which is asymptotically optimal, in the number of time extensions utilized. These two goals are achieved by defining "nearly optimal K-GMD", in which the resulting matrices are as in "perfect K-GMD" form — upper triangular matrices with equal and constant diagonal elements — up to a small number of diagonal elements, which becomes negligible as the number of time extensions grows. This is defined formally as follows.

Definition 6 (Nearly-Optimal K-GMD): Let $A_1, \dots, A_K$ be complex-valued $n \times n$ matrices with determinants equal to 1. Consider a sequence of decompositions (for each N) of the following form. For each N, define the following $nN \times nN$ extended matrices:

$$\mathcal{A}_k \triangleq \lceil A_k \rfloor_{\otimes N} \,, \quad k = 1, \dots, K \,,$$

and the $(K+1)$ matrices $\mathcal{U}_1, \dots, \mathcal{U}_K, \mathcal{V}$ of dimensions $nN \times \tilde{n}$, with orthonormal columns, such that:

$$\mathcal{U}_k^\dagger \mathcal{A}_k \mathcal{V} = \begin{pmatrix} 1 & * & \cdots & * & * \\ 0 & 1 & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & * \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \triangleq \mathcal{T}_k \,, \quad k = 1, \dots, K \,, \tag{38}$$

where $*$ represents some value (which may differ within each matrix as well as between different ones).

We say that the sequence of decompositions is *nearly-optimal K-GMD*, if

$$\lim_{N \to \infty} \frac{\tilde{n}}{nN} = 1 \,.$$

Theorem 4 (Existence of nearly-optimal K-GMD): For any K complex-valued $n \times n$ matrices $A_1, \dots, A_K$ with determinants equal to 1, there exists a sequence of *nearly-optimal K-GMD* with $\tilde{n} = n\big(N - (n^{K-1} - 1)\big)$, where $N \geq n^{K-1}$.

Note that again, as was explained in Remark 12, we assume, w.l.o.g., that all matrices have determinants equal to 1.

The proof of the theorem is given in the form of a constructive algorithm. The algorithm for the general case is presented in Appendix I. Also, implementations of the algorithm in Matlab and Python are available in [42] and [43], respectively. In order to simplify the understanding of the algorithm, we demonstrate the algorithm for some special cases, each of which illustrates a different aspect of the general case. In Section VII-B we present the algorithm for the simplest case of 2-GMD of extended 2×2 matrices, with any number of time extensions. In Appendix F we present the algorithm for the case of 3-GMD of extended 2×2 matrices with only $N = 4$ extensions. In Appendix G we generalize this for general K-GMD of extended 2×2 matrices. Finally, in Appendix H, we present the algorithm for 2-GMD of extended $n \times n$ matrices.

We note that, similarly to the case of perfect triangularization, nearly optimal K-GMD is equivalent to nearly optimal $(K+1)$-JET. This is formally stated in the following lemma, which is a generalization of Lemma 1 to the non square-matrix case, and is proved in Appendix E.

Lemma 3 (Equivalence of K-GMD and (K+1)-JET): Let $A_1, \dots, A_{K+1}$ be $n \times n$ full-rank complex-valued matrices with equal determinants, and define the K matrices:

$$B_k = A_k A_{K+1}^{-1} \,, \qquad k = 1, \dots, K \,. \tag{39}$$

Then, the following two statements are equivalent:

1) There exist $K+1$ matrices with orthonormal columns $U_1, \dots, U_K, U_{K+1}$, of dimensions $n \times \tilde{n}$, such that
$$U_k^\dagger B_k U_{K+1} = T_k \,, \qquad k = 1, \dots, K \,, \tag{40}$$
where $\{T_k\}$ are $\tilde{n} \times \tilde{n}$ upper triangular with all diagonal entries equal to 1.
2) There exist $K+2$ matrices with orthonormal columns $U_1, \dots, U_{K+1}, V$, of dimensions $n \times \tilde{n}$, such that
$$U_k^\dagger A_k V = R_k \,, \qquad k = 1, \dots, K+1 \,,$$
where $\{R_k\}$ are $\tilde{n} \times \tilde{n}$ upper triangular with equal diagonals, as in (15).

Nearly optimal K-GMD is readily applied for K-user common-message BC: Transmission is carried over the equal sub-channel gains whereas the non-equal ones are discarded.

Corollary 3 (Achievable Rates via Nearly-Optimal K-GMD): Let $H_1, \dots, H_K$ be complex-valued channel matrices of dimensions $n_r^{(1)} \times n_t, \dots, n_r^{(K)} \times n_t$, respectively, and $C_{\boldsymbol{x}}$ be an $n_t \times n_t$ covariance matrix satisfying the power constraint $\mathrm{tr}\,(C_{\boldsymbol{x}}) \leq P$. Define $\{\mathcal{H}_k\}$, $\{G_k\}$, and $\{\mathcal{G}_k\}$ as in Section VI-B with $N \geq n_t^{K-1}$ time extensions. Without loss of generality, assume that

$$I(H_1, C_{\boldsymbol{x}}) = \cdots = I(H_K, C_{\boldsymbol{x}}) = C \triangleq n_t \log(1 + \mathrm{SNR}_{\mathrm{eff}}) \,.$$

Then, the following common-message BC rate is achieved:

$$R = \left[1 - \frac{n_t^{K-1} - 1}{N}\right] n_t \log\left(1 + \frac{N}{N - (n_t^{K-1} - 1)} \mathrm{SNR}_{\mathrm{eff}}\right) \tag{41a}$$

$$\geq \left[1 - \frac{n_t^{K-1} - 1}{N}\right] C \,, \tag{41b}$$

using equal-rate capacity-achieving scalar AWGN codes. By taking $N \to \infty$, the achievable rate R achieves capacity.

Proof of Corollary 3: Apply Theorem 4 to $\{\mathcal{G}_k\}$ to obtain the square upper triangular matrices $\{\mathcal{T}_k\}$ of dimensions $n_t\big(N - (n_t^{K-1} - 1)\big)$ with constant diagonals. By using the transmission scheme of Section VI-B over $\{\mathcal{T}_k\}$ a rate of (41b) is achieved. By allocating power and rate only to the $n_t\big(N - (n_t^{K-1} - 1)\big)$ non-discarded streams corresponding to the (constant) diagonal values in $\{\mathcal{T}_k\}$ in (38), the improved rate of (41a) is achieved. ∎

Remark 16: Any nearly optimal K-GMD sequence (not necessarily the one specified in Theorem 4) allows to approach capacity in the limit of $N \to \infty$.

We now demonstrate Corollary 3 for two special cases.

Example 4 (Example 1 Revisited): We reexamine the three-user degrees-of-freedom mismatch setting that was introduced in Example 1 in Section I, which we reproduce here for convenience. We have three users with the following channel matrices:

$$H_1 = \begin{pmatrix} \alpha_1 & 0 \end{pmatrix}, \; H_2 = \begin{pmatrix} 0 & \alpha_1 \end{pmatrix}, \; H_3 = \begin{pmatrix} \alpha_2 & 0 \\ 0 & \alpha_2 \end{pmatrix},$$

such that their WI capacities are equal.

For this specific case, since the third channel matrix is a scaled identity matrix, 3-JET and 3-GMD coincide. Therefore, the number of channel uses needed to achieve 3-GMD is identical to that of 3-JET.

Table I summarizes achievable fractions of capacity corresponding to different numbers of time extensions. We note that in the table we do not apply power compensation as appears in (41a). Thus, the achievable rates according to (41b) are tabulated. For comparison, with $P \to \infty$, time-sharing between the users achieves 33% of the capacity, whereas both Alamouti modulation and beamforming achieve 50%.[15] We note that Alamouti modulation falls under the framework of space–time triangularization (in this case diagonalization) with two time extensions, see [44, Ch. 1.7.3]. By using more than two time extensions, the proposed scheme achieves a larger fraction of capacity.

# Time extensions	2	3	4	5	6	7	8	10
% Capacity	50	66	75	80	83	85	87	90

TABLE I: Fraction of capacity achievable for different numbers of channel uses processed together, when using 3-GMD and 3-JET (without power compensation) in Example 4.

Remark 17: Note that all the schemes considered here impose a decoding order which is shared among all the users. We will see in Section VIII-B that in this particular example, removing this restriction enables to attain 100% efficiency (with no time extensions!).

Example 5 (A General Three-User 2×2 Case): We assume now three general $n_r^{(k)} \times 2$ channel matrices. The resulting channel canonical matrices (23) are of dimensions 2×2. To be optimal for all three users simultaneously, we need to use 3-JET (which can be done using the same parameters of 2-GMD, as explained in Remark 19). If we further wish to have the same SNR for all the scalar subchannels, then we need to use 3-GMD. Table II summarizes achievable fractions of capacity corresponding to different numbers of time extensions. Again, the achievable rates tabulated are according to (41b). For comparison, with $P \to \infty$, time-sharing between the users achieves 33% of the capacity, whereas both Alamouti modulation and beamforming achieve 50%.[16]

# Time extensions	2	3	4	5	6	10	15	30
GMD % Capacity	–	–	25	40	50	70	80	90
JET % Capacity	50	66	75	80	83	90	93	96

TABLE II: Fraction of capacity achievable for different numbers of channel uses processed together, when using 3-GMD and 3-JET (without power compensation) in Example 5.

A. Preliminaries for the Proof of Theorem 4

We now introduce some definitions and properties that will be used in the proof of Theorem 4 in Appendix I, as well as in its demonstration for the simple 2×2 matrix case in Section VII-B and the demonstrations in Appendices F–H.

Definition 7: Define by $j : m$ the list of consecutive indices between j and m:

$$j : m \triangleq (j, j+1, j+2, \ldots, m) \,.$$

Definition 8: Define the operation of "extraction" of multiple *ordered* indices $n_1\,, n_2\,, \ldots\,, n_k$ from a matrix A by:

$$A \left\lfloor n_1,\, n_2\,, \ldots\,, n_k \right\rceil \triangleq \begin{pmatrix} A_{n_1 n_1} & A_{n_1 n_2} & \cdots & A_{n_1 n_k} \\ A_{n_2 n_1} & A_{n_2 n_2} & \cdots & A_{n_2 n_k} \\ \vdots & \cdots & \ddots & \vdots \\ A_{n_k n_1} & A_{n_k n_2} & \cdots & A_{n_k n_k} \end{pmatrix} .$$

For example, if

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 7 & 8 & 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 & 17 & 18 \\ 19 & 20 & 21 & 22 & 23 & 24 \\ 25 & 26 & 27 & 28 & 29 & 30 \\ 31 & 32 & 33 & 34 & 35 & 36 \end{pmatrix},$$

then,

$$A \left\lfloor 2,\, 5 \right\rceil = \begin{pmatrix} 8 & 11 \\ 26 & 29 \end{pmatrix},$$

$$A \left\lfloor 3 :\, 5 \right\rceil = \begin{pmatrix} 15 & 16 & 17 \\ 21 & 22 & 23 \\ 27 & 28 & 29 \end{pmatrix},$$

$$A \left\lfloor 1,\, 6\,, 2 \right\rceil = \begin{pmatrix} 1 & 6 & 2 \\ 31 & 36 & 32 \\ 7 & 12 & 8 \end{pmatrix}.$$

---

[15] In all the schemes, we assume that the scalar codes used are capacity-achieving.

[16] In all the schemes, we assume that the scalar codes used are capacity-achieving.

Definition 9: Define the "embedding" operation $I_n\left[A\,;\bigcup_j \lfloor m_j,\, n_j\rceil\right]$ as the replacement of the elements in the identity matrix I_n in the index-pairs contained in $\lfloor m_1,\, n_1\rceil \lfloor m_2,\, n_2\rceil \lfloor m_3,\, n_3\rceil \ldots \lfloor m_k,\, n_k\rceil$,[17] with the elements of the 2×2 matrix A.

For example, the embedding $I_4\left[B\,;\lfloor 1,\, 3\rceil \lfloor 2,\, 4\rceil\right]$ of

$$B=\begin{pmatrix} 11 & 2 \\ 3 & 4 \end{pmatrix}$$

into the four-dimensional identity matrix I_4 is

$$\begin{pmatrix} 11 & 0 & 2 & 0 \\ 0 & 11 & 0 & 2 \\ 3 & 0 & 4 & 0 \\ 0 & 3 & 0 & 4 \end{pmatrix} .$$

Definition 10: Define the matrix $\mathcal{J}_n^{\left[\{n_j\}_{j=1}^k\right]}$ as an $n\times k$ matrix, whose columns are the $\{n_j\}_{j=1}^k$ vectors of the standard basis:

$$\mathcal{J}_n^{[\{n_j\}]}=\left(\begin{array}{c|c|c|c} e_n^{n_1} & e_n^{n_2} & \cdots & e_n^{n_k}\end{array}\right) ,$$

where $e_n^{n_j}$ is a column-vector of length n with all entries 0 except for the n_j-th entry which equals 1.

For example,

$$\mathcal{J}_5^{[4,1,5]}=\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

Note that $\left(\mathcal{J}_n^{\left[\{n_j\}_{j=1}^k\right]}\right)^\dagger \mathcal{J}_n^{\left[\{n_j\}_{j=1}^k\right]}=I_k$.

Remark 18: For any matrix A, "extraction" can be materialized via multiplication by a matrix $\mathcal{J}_n^{\{[n_j]\}}$ of Definition 10:

$$A\left\lfloor n_1,\, n_2\,,\ldots,n_k\right\rceil=\left(\mathcal{J}_n^{[\{n_j\}]}\right)^\dagger A\mathcal{J}_n^{[\{n_j\}]} .$$

An important special case is the extraction operation of a submatrix:

$$A\left\lfloor j\,:\,m\right\rceil\triangleq\left(\mathcal{J}_n^{[j:m]}\right)^\dagger A\mathcal{J}_n^{[j:m]} .$$

We now introduce a simple key property that will serve as the main idea in our proofs.

Property 1: Let A be a scaled identity matrix, namely, $A=cI$, for some scalar c. The QR decomposition of the matrix A is invariant to multiplications by unitary matrices on the right. This means that for any unitary matrix V, the resulting triangular matrix after applying the QR decomposition to the matrix AV is the matrix A, and further $Q=V^\dagger$:

$$cI_n=V^\dagger cI_nV \qquad \forall c,n .$$

B. Proof of Theorem 4 for $n=2, K=2$ and General N

We now demonstrate the algorithm for the special case of $n=2$, $K=2$, and general N. The proof is based on $K=2$ steps.

Step 1:

We start by performing 1-GMD on the matrix A_1:

$$\left(U_1^{(1)}\right)^\dagger A_1V^{(1)}=\begin{pmatrix} 1 & x_1 \\ 0 & 1 \end{pmatrix} , \tag{42}$$

where the superscripts denote the step number and the subscripts denote the user index. We now apply the decomposition (42) to each block separately, using:

$$\left(\mathcal{U}_1^{(1)}\right)^\dagger\triangleq I_{2N}\left[\left(U_1^{(1)}\right)^\dagger;\lfloor 1,\, 2\rceil\lfloor 3,\, 4\rceil\cdots\lfloor 2N-1,\, 2N\rceil\right] ,$$

$$\mathcal{V}^{(1)}\triangleq I_{2N}\left[V^{(1)}\,;\lfloor 1,\, 2\rceil\lfloor 3,\, 4\rceil\cdots\lfloor 2N-1,\, 2N\rceil\right] ,$$

which yields the following $2N\times 2N$ extended triangular matrix:

$$\mathcal{T}_1^{(1)}=\left(\mathcal{U}_1^{(1)}\right)^\dagger\mathcal{A}_1\mathcal{V}^{(1)}$$
$$=\begin{pmatrix} 1 & x_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & x_1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & x_1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix} .$$

Note that the same matrix $\mathcal{V}^{(1)}$ has to be applied also to the matrix of the second user (since the encoder is shared by all users). We next decompose the resulting matrix (after multiplying it by $\mathcal{V}^{(1)}$ on the right) according to the QR decomposition, resulting in a unitary matrix $\left(\mathcal{U}_2^{(1)}\right)^\dagger$ such that:

$$\mathcal{T}_2^{(1)}=\left(\mathcal{U}_2^{(1)}\right)^\dagger\mathcal{A}_2\mathcal{V}^{(1)}$$
$$=\begin{pmatrix} r_1 & x_2 & 0 & 0 & \cdots & 0 & 0 \\ 0 & r_2 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & r_1 & x_2 & \cdots & 0 & 0 \\ 0 & 0 & 0 & r_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & r_1 & x_2 \\ 0 & 0 & 0 & 0 & \cdots & 0 & r_2 \end{pmatrix} .$$

Step 2:

Note that the submatrix $\mathcal{T}_1^{(1)}\lfloor 2,\, 3\rceil$ is $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Thus, according to Property 1 we can perform 1-GMD on the corresponding elements of the matrix of user 2, $\mathcal{T}_2^{(1)}\lfloor 2,\, 3\rceil$, without changing $\mathcal{T}_1^{(1)}\lfloor 2,\, 3\rceil$:

$$\left(U_2^{(2)}\right)^\dagger\begin{pmatrix} r_2 & 0 \\ 0 & r_1 \end{pmatrix}V^{(2)}=\begin{pmatrix} 1 & x_2^{(2)} \\ 0 & 1 \end{pmatrix} .$$

---

[17] The notation $\lfloor j,\, m\rceil\lfloor p,\, q\rceil$ stands for $\lfloor j,\, m\rceil\cup\lfloor p,\, q\rceil$.

Hence, by defining

$$\begin{aligned} &\left(\mathcal{U}_2^{(2)}\right)^\dagger \\ &\triangleq I_{2N}\left[\left(U_2^{(2)}\right)^\dagger ; \lfloor 2,\, 3\rceil\, \lfloor 4,\, 5\rceil \cdots \lfloor 2N-2,\, 2N-1\rceil\right], \\ &\mathcal{V}^{(2)} \\ &\triangleq I_{2N}\left[V^{(2)} ; \lfloor 2,\, 3\rceil\, \lfloor 4,\, 5\rceil \cdots \lfloor 2N-2,\, 2N-1\rceil\right], \end{aligned}$$

and applying them to $\mathcal{T}_1^{(1)}$ and $\mathcal{T}_2^{(1)}$, we attain:

$$\begin{aligned} \mathcal{T}_1^{(2)} &= \left(\mathcal{V}^{(2)}\right)^\dagger \mathcal{T}_1^{(1)} \mathcal{V}^{(2)} \\ &= \left(\mathcal{V}^{(2)}\right)^\dagger \left(\mathcal{U}_1^{(1)}\right)^\dagger \mathcal{A}_1 \mathcal{V}^{(1)} \mathcal{V}^{(2)} \\ &= \left(\begin{array}{c|cccccc|c} 1 & \tilde{x}_1 & * & 0 & \cdots & 0 & 0 \\ \cline{2-7} 0 & 1 & 0 & * & \cdots & 0 & 0 \\ 0 & 0 & 1 & \tilde{x}_1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & \tilde{x}_1 \\ \cline{2-7} 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{array}\right), \\ \mathcal{T}_2^{(2)} &= \left(\mathcal{U}_2^{(2)}\right)^\dagger \mathcal{T}_2^{(1)} \mathcal{V}^{(2)} \\ &= \left(\mathcal{U}_2^{(2)}\right)^\dagger \left(\mathcal{U}_2^{(1)}\right)^\dagger \mathcal{A}_2 \mathcal{V}^{(1)} \mathcal{V}^{(2)} \\ &= \left(\begin{array}{c|cccccc|c} r_1 & \tilde{x}_2 & * & 0 & \cdots & 0 & 0 \\ \cline{2-7} 0 & 1 & x_2^{(2)} & * & \cdots & 0 & 0 \\ 0 & 0 & 1 & \tilde{x}_2 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & \tilde{x}_2 \\ \cline{2-7} 0 & 0 & 0 & 0 & \cdots & 0 & r_2 \end{array}\right). \end{aligned}$$

Now, to get the desired decomposition we need to "extract" the middle submatrices (by multiplying on both sides by $\mathcal{J}_{2N}^{[2:2N-1]}$, as explained in Remark 18).

Thus, by defining

$$\begin{aligned} \mathcal{V} &\triangleq \mathcal{V}^{(1)}\mathcal{V}^{(2)}\mathcal{J}_{2N}^{[2:2N-1]} \\ (\mathcal{U}_1)^\dagger &\triangleq \left(\mathcal{J}_{2N}^{[2:2N-1]}\right)^\dagger \left(\mathcal{V}^{(2)}\right)^\dagger \left(\mathcal{U}_1^{(1)}\right)^\dagger \\ (\mathcal{U}_2)^\dagger &\triangleq \left(\mathcal{J}_{2N}^{[2:2N-1]}\right)^\dagger \left(\mathcal{U}_2^{(2)}\right)^\dagger \left(\mathcal{U}_2^{(1)}\right)^\dagger \end{aligned}$$

we arrive at the desired result.

Remark 19: It was shown in Lemma 3 that K-GMD is equivalent to $(K+1)$-JET. Hence, nearly-optimal $(K+1)$-JET can be obtained with the same parameters as in Theorem 4. Alternatively, an explicit algorithm for $(K+1)$-JET can be obtained by performing the K-GMD algorithm as in Appendix I, where in the first step, instead of performing 1-GMD on the matrix A_1, 2-JET on the matrices A_1 and A_2 is performed, and similarly, in step ℓ instead of performing 1-GMD on the matrix $\mathcal{T}_\ell^{(\ell)(1)} \lfloor 1:n\rceil$, 2-JET on the matrices $\mathcal{T}_\ell^{(\ell)(1)} \lfloor 1:n\rceil$ and $\mathcal{T}_{l+1}^{(\ell)(1)} \lfloor 1:n\rceil$ is performed.

VIII. EXTENSIONS

A. Time-Varying Channel

Throughout this paper, we have considered the problem of broadcasting the same information to K different users over *static* Gaussian MIMO channels, described by the matrices H_k. As mentioned in Section III, this problem is equivalent to the problem of transmission over a compound channel [7]–[9], where a transmitter wishes to convey information to a single receiver over a MIMO channel, which can take one out of K realizations, the set of which is known at the transmitter, but the exact realization is known only to the receiver (but not to the transmitter) and remains constant throughout the whole transmission.

For this problem, the schemes of Section V-C and Section VII may be readily used. These schemes may further be extended to the case where the channel varies in time. For $K = 2$, using the JET-based scheme, any arbitrary sequence of channel realizations (within the set $\{H_1, H_2\}$) may be accommodated, provided that this sequence is known to the receiver. The transmitter, in this case, is identical to the one in the "compound scenario", whereas the receiver needs to apply to its received signal, at each time instant, $U_1^\dagger$ or $U_2^\dagger$, depending on the channel realization at this time instant (H_1 or H_2, respectively). The successive decoding process needs to be modified as follows: The last sub-channel is interference-free, as in the "compound scenario", and therefore its interference can be subtracted of the other sub-channels; however, its components in the other sub-channels, differ with the realizations at each time instant ("off-diagonal" coefficients differ with H_k, unlike the diagonal ones which are equal to all channel realizations). The successive decoding process of the other sub-messages needs to be modified in a similar manner.

Note however that for $K > 2$ channel realizations, more channel uses need to be processed together, in general, as explained in Section VII. In the time-varying scenario, this implies that, in order to use the schemes of Section VII, the channel needs to be constant in time for a number of time instants which equals the number of channel uses that are jointly processed together. This requirement is shared by the space–time schemes of [21] and [22].

B. Different Decoding Orders

In the above sections, we discussed the simultaneous decomposition of several matrices into *upper triangular* forms. In terms of the transmission scheme described in Section V-C, all the receivers decode the messages in the same order (starting with the last component; ending with the first one).

This scheme can be generalized, if we allow each receiver to choose *its own* order of decoding. It turns out that this generalized scheme can achieve rates which are strictly higher than the rates achieved using the ordinary scheme (where all the decoders use the same order of decoding).

In the case of two transmit antennas, the channel canonical matrices (23) are 2×2 matrices. Thus, allowing different decoding orders means that some matrices are transformed into *upper triangular* matrices, whereas the others — into *lower triangular* matrices, where all the resulting matrices have equal diagonal values. The following theorem is proved using a similar technique to the one used for the proof of Theorem 2. Again, as explained in Remark 12, we can assume without loss of generality that both matrices have determinants equal to 1.

Theorem 5: Let A_1 and A_2 be *complex-valued* 2×2 matrices with determinants equal to 1. Then, there exist three complex-valued 2×2 unitary matrices U_1, U_2, and V, such that

$$(U_1)^\dagger A_1 V = \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}$$

and

$$(U_2)^\dagger A_2 V = \begin{pmatrix} 1 & 0 \\ * & 1 \end{pmatrix},$$

if and only if the following inequality is satisfied:

$$F_2\left(A_1^\dagger A_1 - I, A_2^\dagger A_2 - I\right) \geq 0,$$

where

$$F_2(S_1, S_2) \triangleq \det\bigl(S_1 S_2 - \operatorname{adj} S_2 \operatorname{adj} S_1\bigr).$$

The proof is given in Appendix J.

This result can be easily generalized, as stated in the following corollary.

Corollary 4: Let A_1 and A_2 be *complex-valued* 2×2 matrices with determinants equal to 1, and let $r > 0$. Then there exist three complex-valued 2×2 unitary matrices U_1, U_2, and V, such that

$$(U_1)^\dagger A_1 V = \begin{pmatrix} r & * \\ 0 & 1/r \end{pmatrix}$$

and

$$(U_2)^\dagger A_2 V = \begin{pmatrix} r & 0 \\ * & 1/r \end{pmatrix}$$

if and only if the following conditions are satisfied:

$$\begin{aligned} \det\left(A_1^\dagger A_1 - r^2 I\right) &\leq 0 \\ \det\left(A_2^\dagger A_2 - 1/r^2 I\right) &\leq 0 \\ F_2\left(A_1^\dagger A_1 - r^2 I, A_2^\dagger A_2 - 1/r^2 I\right) &\geq 0. \end{aligned}$$

The proof of the corollary follows along the same lines as that of Theorem 5 with obvious modifications.

Recall the "degrees-of-freedom mismatch" scenario of Examples 1 and 4. The compound capacity in this case is achieved by a white input covariance matrix. The corresponding channel canonical matrices (23), are

$$\begin{aligned} G_1 &= \begin{pmatrix} 2^{C/2} & 0 \\ 0 & 1 \end{pmatrix}, \\ G_2 &= \begin{pmatrix} 1 & 0 \\ 0 & 2^{C/2} \end{pmatrix}, \\ G_3 &= \begin{pmatrix} 2^{C/4} & 0 \\ 0 & 2^{C/4} \end{pmatrix}. \end{aligned}$$

Since G_3 is a scaled identity matrix, performing 3-GMD on these three matrices is in fact equivalent to 2-GMD of G_1 and G_2, which is not possible according to Theorem 2. However, if we allow generalized triangularization — namely, receiver 1 transforms the channel into upper triangular form, whereas receiver 2 transforms it into lower triangular form — then the decomposition is possible according to Theorem 5, using the following precoding matrix:

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

which gives rise, in turn, to the following triangular matrices:

$$\begin{aligned} T_1 &= \begin{pmatrix} 2^{C/4} & \frac{2^C-1}{2^{C/2}+1} \\ 0 & 2^{C/4} \end{pmatrix}, \\ T_2 &= \begin{pmatrix} 2^{C/4} & 0 \\ -\frac{2^C-1}{2^{C/2}+1} & 2^{C/4} \end{pmatrix}, \\ T_3 &= \begin{pmatrix} 2^{C/4} & 0 \\ 0 & 2^{C/4} \end{pmatrix}. \end{aligned}$$

C. Block GTD

There are certain cases, where triangularity of the resulting matrices is not necessary and block-triangular forms, with blocks satisfying certain relations between their determinants, suffice. In these cases we are interested primarily in deriving information-theoretic bounds, rather than constructing practical communication schemes.

This is the case for the Gaussian MIMO joint source–channel coding (JSCC) problem, where we wish to convey a scalar Gaussian source over Gaussian MIMO links, having different capacities. In this case, pure digital transmission, as in Sections V and VI, is not optimal, as it is restricted to the minimum of the capacities of the different MIMO links. Indeed, better performance may be achieved, using a scheme which better adapts to the different capacities of the different channel links. For more information see [16, Sec. IV].

For this purpose, we first extend the GTD, discussed in Section IV-A, for a block-triangular form, after which we apply this result in the derivation of a block joint triangularization.

Theorem 6 (Block GTD): Let A be an $n \times n$ full-rank matrix. Then, it can be decomposed into a block upper triangular form ($1 \leq M \leq n$):

$$\tilde{R} = \begin{pmatrix} \tilde{R}_{11} & \tilde{R}_{12} & \cdots & \tilde{R}_{1M} \\ 0 & \tilde{R}_{22} & \cdots & \tilde{R}_{2M} \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & \tilde{R}_{MM} \end{pmatrix}, \tag{43}$$

where $\tilde{R}_{j\ell}$ are $n_j \times n_\ell$ blocks, and the matrices $\tilde{R}_{mm}$ have prescribed determinants $\det\left(\tilde{R}_{mm}\right)$, such that $\sum_{m=1}^{M} n_m = n$, if and only if

$$\prod_{m=1}^{q} \left|\det\left(\tilde{R}_{p_m p_m}\right)\right| \leq \prod_{j=1}^{\sum_{m=1}^{q} n_{p_m}} \sigma_j \tag{44}$$

for all $q = 1, 2, ..., M$, and

$$\prod_{m=1}^{M} \left|\det\left(\tilde{R}_{p_m p_m}\right)\right| = \prod_{j=1}^{n} \sigma_j, \tag{45}$$

where σ_j are the singular values of A ordered non-increasingly, $\{p_m\}_{m=1}^{M}$ are the indices satisfying

$$d_{p_1} \geq d_{p_2} \geq \cdots \geq d_{p_M} ,$$

and

$$d_m \triangleq \sqrt[n_m]{\left|\det\left(\tilde{R}_{mm}\right)\right|} , \qquad m = 1, ..., M .$$

Before we prove this theorem, we need the following lemma.

Lemma 4 (GTD with Multiplicities): Let A be an $n \times n$ full-rank matrix with singular values $\{\sigma_j\}$, ordered non-increasingly. Then, it can be decomposed as

$$A = URV^{\dagger} , \tag{46}$$

where R is upper triangular and U, V are unitary, if and only if

$$\prod_{m=1}^{q} r_m^{n_m} \leq \prod_{j=1}^{\sum_{m=1}^{q} n_m} \sigma_j \tag{47}$$

for every q ($q = 1, 2, ..., M$), and

$$\prod_{m=1}^{M} r_m^{n_m} = \prod_{j=1}^{n} \sigma_j , \tag{48}$$

where the absolute values of the diagonal of R take M ($1 \leq M \leq n$) distinct values; these values, ordered non-decreasingly, are denoted by r_m ($m = 1, 2, ..., M$), and the number of occurrences (“multiplicity”) of each value — by n_m.
The proof of this lemma is given in Appendix K.

Note that this lemma suggests that in case of multiplicities of the absolute values of the desired diagonal entries of the triangular matrix, if those entries take only M different values, then it suffices to verify only M conditions (1 condition per distinct value), instead of the n conditions of general GTD.

Proof of Theorem 6: Decompose, according to the GMD, every block matrix $\tilde{R}_{mm}$ in (43) laying on the main diagonal, as

$$\tilde{R}_{mm} = U_{mm} T_{mm} V_{mm}^{\dagger} , \qquad m = 1, 2, ..., K ,$$

where U_{mm} and V_{mm} are unitary and T_{mm} is upper triangular with constant diagonal entries which are equal to

$$[T_{mm}]_j = \sqrt[n_m]{\left|\det\left(\tilde{R}_{mm}\right)\right|} \triangleq d_m , \qquad j = 1, 2, ..., n_m .$$

Hence, applying the unitary matrices $U^{\dagger}$ on the left and V on the right, given by

$$U = \begin{pmatrix} U_{11} & 0 & \cdots & 0 \\ 0 & U_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & U_{MM} \end{pmatrix} ,$$

$$V = \begin{pmatrix} V_{11} & 0 & \cdots & 0 \\ 0 & V_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & V_{MM} \end{pmatrix} ,$$

gives rise to an upper triangular matrix whose diagonal equals to the concatenation of the diagonals of $\{T_{mm}\}$. Therefore, the task of constructing the decomposition (43) is equivalent to decomposing A into triangular form with a diagonal that is equal to the concatenation of the diagonals of $\{T_{mm}\}$. Denote the entries of this diagonal, reordered non-increasingly, by $\boldsymbol{r}(A)$ and the singular values of A by $\boldsymbol{\sigma}(A)$. Then, the aforementioned decomposition is possible if and only if Weyl's condition

$$\boldsymbol{\sigma}(A) \succeq \boldsymbol{r}(A)$$

is satisfied, which in turn is satisfied if and only if (44) and (45) hold, according to Lemma 4. ∎

Corollary 5 (Joint Block Triangularization): Let A_1 and A_2 be two full-rank $n \times n$ complex-valued matrices. Then A_1 and A_2 can be jointly decomposed into block-triangular forms

$$\begin{aligned} A_1 &= U_1 \tilde{R}_1 V^{\dagger} \\ A_2 &= U_2 \tilde{R}_2 V^{\dagger} , \end{aligned}$$

where U_k and V are unitary, and $\tilde{R}_k$ are block-triangular:

$$\tilde{R}_k = \begin{pmatrix} \tilde{R}_{11}^{(k)} & \tilde{R}_{12}^{(k)} & \cdots & \tilde{R}_{1M}^{(k)} \\ 0 & \tilde{R}_{22}^{(k)} & \cdots & \tilde{R}_{2M}^{(k)} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \tilde{R}_{MM}^{(k)} \end{pmatrix} , \qquad k = 1, 2 ,$$

where corresponding blocks $\tilde{R}_{j\ell}^{(1)}$ and $\tilde{R}_{j\ell}^{(2)}$ have the same dimensions $n_j \times n_\ell$, such that $\sum_m n_m = n$, and prescribed determinant ratios of the blocks on the main diagonal, $\det\left(\tilde{R}_{mm}^{(1)}\right) / \det\left(\tilde{R}_{mm}^{(2)}\right)$ if and only if

$$\prod_{m=1}^{q} \left|\det\left(\tilde{R}_{p_m p_m}^{(1)}\right) \Big/ \det\left(\tilde{R}_{p_m p_m}^{(2)}\right)\right| \leq \prod_{j=1}^{\sum_{\ell=1}^{q} n_{k_\ell}} \mu_j$$

for all $q = 1, 2, ..., M$, and

$$\prod_{m=1}^{M} \left|\det\left(\tilde{R}_{p_m p_m}^{(1)}\right) \Big/ \det\left(\tilde{R}_{p_m p_m}^{(2)}\right)\right| = \prod_{j=1}^{n} \mu_j ,$$

where μ_j are the generalized singular values [31], [45] of $(\tilde{R}_1, \tilde{R}_2)$ ordered non-increasingly, $\{p_m\}_{m=1}^{K}$ are the indices satisfying

$$d_{p_1} \geq d_{p_2} \geq \cdots \geq d_{p_M} ,$$

and

$$d_m \triangleq \sqrt[n_m]{\left|\det\left(\tilde{R}_{p_m p_m}^{(1)}\right) \Big/ \det\left(\tilde{R}_{p_m p_m}^{(2)}\right)\right|} , \; m = 1, 2, ..., M .$$

Proof: The proof is similar to the proof of [16, Theorem 1], by replacing the GTD by the block-GTD of Theorem 6 and using the fact that the inverse of a square block-triangular matrix is a matrix of the same block-triangular form with blocks on its main diagonal which are equal to the inverses of the original matrix, and the fact that multiplying two square block-triangular matrices with the same block dimensions results in a matrix of the same block-triangular form with blocks on its main diagonal which are equal to the product of the corresponding blocks of the multiplied matrices. ∎

IX. DISCUSSION AND FURTHER RESEARCH

In this work, we derived new joint triangularizations of several matrices. Specifically, we were interested in designing triangular matrices having equal or constant diagonals, by applying unitary operations, for the construction of a practical scheme for the common-message BC problem, that approaches its capacity. We derived conditions for the existence of such decompositions, for specific cases; conditions for general matrices — remain unknown.

For the general case (even when such exact decompositions are not possible), we introduced a decomposition that nearly achieves this goal for time-extended variants of the channel matrices. However, the number of time extensions required, for this proposed decomposition, grows rapidly with the number of jointly-decomposed matrices. Nonetheless, numerical evidence suggests that this number of required time extensions, can be greatly reduced, and calls for further research.

APPENDIX A
PROOF OF LEMMA 2

Before we turn to the proof of the lemma, we introduce the following lemma, the proof of which is relegated to Appendix B.

Lemma 5: Let S_1 and S_2 be $n \times n$ complex-valued matrices, and let U be an $n \times n$ unitary matrix. Then,

$$F_1\left(U^\dagger S_1 U, U^\dagger S_2 U\right) = F_1(S_1, S_2)\,.$$

Now, let S_1 and S_2 be two complex-valued 2×2 Hermitian matrices. Without loss of generality, we can restrict ourselves to vectors $\boldsymbol{v} \in \mathbb{C}^2$ that have a Euclidean norm of 1. Namely, we are looking for a necessary and sufficient condition for the existence of a solution $v \in \mathbb{C}^2$ to the following three equations:

$$\boldsymbol{v}^\dagger S_1 \boldsymbol{v} = 0 \tag{49a}$$

$$\boldsymbol{v}^\dagger S_2 \boldsymbol{v} = 0 \tag{49b}$$

$$\|\boldsymbol{v}\| = 1\,. \tag{49c}$$

First, note that if $\det(S_1) > 0$ then S_1 is either positive definite or negative definite, and in both cases there is no non-zero solution $\boldsymbol{v}$ to (49a). Similarly, if $\det(S_2) > 0$ there is no non-zero solution to (49b). Therefore, from now on we can assume that $\det(S_1) \leq 0$ and $\det(S_2) \leq 0$.

Note that for any 2×2 unitary matrix U, the decomposition (49) is equivalent to

$$\tilde{\boldsymbol{v}}^\dagger \tilde{S}_1 \tilde{\boldsymbol{v}} = 0 \tag{50a}$$

$$\tilde{\boldsymbol{v}}^\dagger \tilde{S}_2 \tilde{\boldsymbol{v}} = 0 \tag{50b}$$

$$\|\tilde{\boldsymbol{v}}\| = 1\,, \tag{50c}$$

where

$$\begin{aligned} \tilde{\boldsymbol{v}} &\triangleq U^\dagger \boldsymbol{v} \\ \tilde{S}_1 &\triangleq U^\dagger S_1 U \\ \tilde{S}_2 &\triangleq U^\dagger S_2 U\,. \end{aligned}$$

Since S_k are Hermitian, so are $\tilde{S}_k$.

Also, according to Lemma 5, (28) is equivalent to

$$\begin{aligned} \det(\tilde{S}_1) &\leq 0 \\ \det(\tilde{S}_2) &\leq 0 \\ F_1\left(\tilde{S}_1, \tilde{S}_2\right) &\geq 0\,. \end{aligned}$$

Thus, by choosing U that diagonalizes S_1, we can assume without loss of generality that S_1 is real valued and diagonal matrix:

$$\begin{aligned} S_1 &= \begin{pmatrix} a_1 & 0 \\ 0 & c_1 \end{pmatrix} \\ S_2 &= \begin{pmatrix} a_2 & b_2 + i\beta_2 \\ b_2 - i\beta_2 & c_2 \end{pmatrix}, \end{aligned}$$

where $a_1, c_1, a_2, c_2, b_2, \beta_2$ are real-valued. Denoting

$$\boldsymbol{v} = \begin{pmatrix} x_1 + ix_2 \\ y_1 + iy_2 \end{pmatrix},$$

the three equations (50) become:

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ a_1 & 0 & 0 & c_1 \\ a_2 & b_2 & \beta_2 & c_2 \end{pmatrix} \begin{pmatrix} x_1^2 + x_2^2 \\ 2(x_1y_1 + x_2y_2) \\ 2(x_2y_1 - x_1y_2) \\ y_1^2 + y_2^2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \tag{51}$$

We now consider the following cases.

a) Case 1: Assume first that $a_1 \neq c_1$ and $b_2 \neq 0$. Thus, (51) is equivalent to:

$$\begin{aligned} \begin{pmatrix} x_1^2 + x_2^2 \\ 2(x_1y_1 + x_2y_2) \\ 2(x_2y_1 - x_1y_2) \\ y_1^2 + y_2^2 \end{pmatrix} &= \overbrace{\begin{pmatrix} 1 & 0 & 0 & 1 \\ a_1 & 0 & 0 & c_1 \\ a_2 & b_2 & \beta_2 & c_2 \\ 0 & 0 & 1 & 0 \end{pmatrix}^{-1}}^{B^{-1}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ t \end{pmatrix} \\ &\triangleq \frac{1}{\Delta} \begin{pmatrix} f_1(t) \\ f_2(t) \\ f_3(t) \\ f_4(t) \end{pmatrix}, \end{aligned}$$

where t is some real-valued parameter, $f_1(t)$,$f_2(t)$,$f_3(t)$,$f_4(t)$ are four first-degree polynomials in t (with coefficients that depend on the matrices S_1, S_2, and where

$$\Delta \triangleq \det B = b_2(c_1 - a_1) \neq 0\,.$$

Thus, finding a solution $\boldsymbol{v}$ to the original problem is *equivalent* to finding a solution (x_1, x_2, y_1, y_2, t) to the following equations:

$$x_1^2 + x_2^2 = \frac{1}{\Delta} f_1(t) \tag{52a}$$

$$2(x_1y_1 + x_2y_2) = \frac{1}{\Delta} f_2(t) \tag{52b}$$

$$2(x_2y_1 - x_1y_2) = \frac{1}{\Delta} f_3(t) \tag{52c}$$

$$y_1^2 + y_2^2 = \frac{1}{\Delta} f_4(t) \tag{52d}$$

Assertion 1: A solution to (52) exists if and only if the following conditions hold for some $t \in \mathbb{R}$:

$$\frac{1}{\Delta} f_1(t) \geq 0 \tag{53a}$$

$$\frac{1}{\Delta} f_4(t) \geq 0 \tag{53b}$$

$$4 f_1(t) f_4(t) = f_2^2(t) + f_3^2(t)\,. \tag{53c}$$

Proof of Assertion 1: Construct the following three vectors: $\mathbf{p}_1 = (x_2, -x_1)$, $\mathbf{p}_2 = (y_1, y_2)$, $\mathbf{p}_3 = (x_1, x_2)$. Then,

$$||\mathbf{p}_1||^2 = ||\mathbf{p}_3||^2 = x_1^2 + x_2^2 \tag{54a}$$

$$||\mathbf{p}_2||^2 = y_1^2 + y_2^2 \tag{54b}$$

$$2\langle \mathbf{p}_1, \mathbf{p}_2 \rangle = 2(x_2 y_1 - x_1 y_2) \tag{54c}$$

$$2\langle \mathbf{p}_2, \mathbf{p}_3 \rangle = 2(x_1 y_1 + x_2 y_2) \tag{54d}$$

Note that the l.h.s. of (52) and the r.h.s. of (54) coincide. We note that $\mathbf{p}_3$ and $\mathbf{p}_1$ are orthogonal. Hence, the angles between these vectors satisfy

$$\begin{aligned} \cos\theta_1 &= \frac{\langle \mathbf{p}_1, \mathbf{p}_2 \rangle}{||\mathbf{p}_1|| ||\mathbf{p}_2||} \\ \cos\theta_2 &= \frac{\langle \mathbf{p}_3, \mathbf{p}_2 \rangle}{||\mathbf{p}_3|| ||\mathbf{p}_2||} = \frac{\langle \mathbf{p}_3, \mathbf{p}_2 \rangle}{||\mathbf{p}_1|| ||\mathbf{p}_2||} \\ \cos\theta_2 &= \cos\left(\pm\frac{\pi}{2} - \theta_1\right) = \pm \sin\theta_1\,. \end{aligned}$$

Thus, a solution to (52) exists if and only if

$$||\mathbf{p}_1||^2 \geq 0 \tag{55a}$$

$$||\mathbf{p}_2||^2 \geq 0 \tag{55b}$$

$$\cos^2\theta_1 + \sin^2\theta_1 = \frac{\langle \mathbf{p}_1, \mathbf{p}_2 \rangle^2}{||\mathbf{p}_1||^2 ||\mathbf{p}_2||^2} + \frac{\langle \mathbf{p}_3, \mathbf{p}_2 \rangle^2}{||\mathbf{p}_1||^2 ||\mathbf{p}_2||^2} = 1. \tag{55c}$$

where (55c) is equivalent to

$$0 = 4||\mathbf{p}_1||^2 ||\mathbf{p}_2||^2 - (2\langle \mathbf{p}_1, \mathbf{p}_2 \rangle)^2 - (2\langle \mathbf{p}_3, \mathbf{p}_2 \rangle)^2\,,$$

which is equivalent, in turn, to (53). ■

By definition, and using (51), we have $(f_1(t) + f_4(t)) = \Delta$. Therefore, the three conditions of (53) are equivalent to the single condition

$$4 f_1(t) f_4(t) - f_2^2(t) - f_3^2(t) = 0\,.$$

This is a quadratic equation in t:

$$a t^2 + b t + c = 0\,,$$

where the constants a, b, c depend on the matrices S_1, S_2 as follows:

$$a \triangleq -(a_1 - c_1)^2 (b_2^2 + \beta_2^2) \tag{56a}$$

$$b \triangleq 2\beta_2 (a_2 c_1 - a_1 c_2)(a_1 - c_1) \tag{56b}$$

$$c \triangleq -4 a_1 c_1 b_2^2 - (a_2 c_1 - a_1 c_2)^2\,. \tag{56c}$$

Note that since $a_1 \neq c_1$ and $b_2 \neq 0$, the coefficient a is *strictly* negative. Therefore, a necessary and sufficient condition for the existence of a solution is for the discriminant to be non-negative:

$$b^2 - 4ac \geq 0\,.$$

A direct calculation shows that

$$b^2 - 4ac = 4\Delta^2 F_1(S_1, S_2)\,,$$

where

$$F_1(S_1, S_2) \triangleq \det\left(S_1 \operatorname{adj}(S_2) - S_2 \operatorname{adj}(S_1)\right)\,,$$

which completes the proof for this case.

b) Case 2: Assume now that $a_1 = c_1$. Since we assumed $\det(S_1) \leq 0$, this means that $a_1 = c_1 = 0$, namely, $S_1 = 0$. In this case we have

$$F_1(S_1, S_2) = F_1(0, S_2) = 0\,.$$

Thus, condition (26) holds. Since we assumed that $\det(S_2) \leq 0$, S_2 has one non-negative eigenvalue and one non-positive eigenvalue, therefore there necessarily exists $\boldsymbol{v}$ with norm 1 such that $\boldsymbol{v}^\dagger S_2 \boldsymbol{v} = 0$, and therefore there exists a solution to the equations in (49).

c) Case 3: Next, assume that $a_1 \neq c_1$, $b_2 = 0$, and $\beta_2 \neq 0$. Thus, (51) becomes

$$\begin{pmatrix} 1 & 0 & 1 \\ a_1 & 0 & c_1 \\ a_2 & \beta_2 & c_2 \end{pmatrix} \begin{pmatrix} x_1^2 + x_2^2 \\ 2(x_2 y_1 - x_1 y_2) \\ y_1^2 + y_2^2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \tag{57}$$

which reduces to

$$\begin{pmatrix} x_1^2 + x_2^2 \\ 2(x_2 y_1 - x_1 y_2) \\ y_1^2 + y_2^2 \end{pmatrix} = \begin{pmatrix} f_5 \\ f_6 \\ f_7 \end{pmatrix} \triangleq \frac{1}{(a_1 - c_1)\beta_2} \begin{pmatrix} -\beta_2 c_1 \\ a_2 c_1 - a_1 c_2 \\ a_1 \beta_2 \end{pmatrix}. \tag{58}$$

Assertion 2: A solution to (58) exists if and only if the following conditions holds:

$$f_5 \geq 0 \tag{59a}$$

$$f_7 \geq 0 \tag{59b}$$

$$4 f_5 f_7 - f_6^2 \geq 0\,. \tag{59c}$$

Proof of Assertion 2: Construct the following two vectors: $\mathbf{p}_1 = (x_2, -x_1)$, $\mathbf{p}_2 = (y_1, y_2)$. Using the inner product definition, we have

$$||\mathbf{p}_1||^2 = x_1^2 + x_2^2 \tag{60a}$$

$$||\mathbf{p}_2||^2 = y_1^2 + y_2^2 \tag{60b}$$

$$2\langle \mathbf{p}_1, \mathbf{p}_2 \rangle = 2(x_2 y_1 - x_1 y_2)\,, \tag{60c}$$

and the angle between the two vectors satisfies

$$\cos\theta_1 = \frac{\langle \mathbf{p}_1, \mathbf{p}_2 \rangle}{||\mathbf{p}_1|| ||\mathbf{p}_2||}\,.$$

Note that the l.h.s. of (58) coincides with the r.h.s. of (60). Thus, a solution to (58) exists if and only if

$$||\mathbf{p}_1||^2 \geq 0 \tag{61a}$$

$$||\mathbf{p}_2||^2 \geq 0 \tag{61b}$$

$$\frac{\langle \mathbf{p}_1, \mathbf{p}_2 \rangle}{||\mathbf{p}_1|| ||\mathbf{p}_2||} \leq 1\,, \tag{61c}$$

where (61c) is equivalent to

$$4||\mathbf{p}_1||^2||\mathbf{p}_2||^2 - \left(2\left\langle \mathbf{p}_1, \mathbf{p}_2 \right\rangle\right)^2 \geq 0,$$

which is equivalent, in turn, to (59). ■

By definition, and using (57), we have $f_5 + f_7 = 1$. Thus, these three equations are equivalent to the single equation

$$4f_5 f_7 - f_6^2 = \frac{-(a_2c_1 - a_1c_2)^2 - 4a_1c_1\beta_2^2}{\beta_2^2(a_1 - c_1)^2} \geq 0\,.$$

Since the denominator is positive, this is equivalent to

$$-(a_2c_1 - a_1c_2)^2 - 4a_1c_1\beta_2^2 \geq 0\,.$$

On the other hand, we have

$$F_1(S_1, S_2) = -(a_2c_1 - a_1c_2)^2 - 4a_1c_1\beta_2^2\,.$$

Thus, condition (26) holds if and only if there exists a solution to (49).

d) Case 4: We are left with the case where $a_1 \neq c_1$, $b_2 = 0$, and $\beta_2 = 0$. In this case, (51) becomes

$$\begin{pmatrix} 1 & 1 \\ a_1 & c_1 \\ a_2 & c_2 \end{pmatrix} \begin{pmatrix} x_1^2 + x_2^2 \\ y_1^2 + y_2^2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

A necessary condition for the existence of a solution is that the second and the third rows are linearly dependent (or in other words, $a_1c_2 = a_2c_1$), in which case we have

$$\begin{aligned} x_1^2 + x_2^2 &= \frac{c_1}{c_1 - a_1} \\ y_1^2 + y_2^2 &= \frac{-a_1}{c_1 - a_1}\,. \end{aligned}$$

Since we assumed $\det(S_1) \leq 0$, a_1 and c_1 have opposite signs, and therefore $x_1^2 + x_2^2$ and $y_1^2 + y_2^2$ are both non-negative. In conclusion, a necessary and sufficient condition for the existence of a solution to (49) in this case is $a_2c_1 = a_1c_2$. On the other hand, we have

$$F_1(S_1, S_2) = -(a_2c_1 - a_1c_2)^2\,,$$

which is non-negative if and only if $a_2c_1 = a_1c_2$. Thus, (26) is a necessary and sufficient condition for the existence of a solution to (49).

This concludes the proof of the lemma. ■

APPENDIX B
PROOF OF LEMMA 5

Let S_1 and S_2 be $n \times n$ complex-valued matrices, and let U be an $n \times n$ unitary matrix. We have:

$$\begin{aligned} &F_1\left(U^\dagger S_1 U, U^\dagger S_2 U\right) \\ &\quad = \det\left[U^\dagger S_1 U \operatorname{adj}\left(U^\dagger S_2 U\right) - U^\dagger S_2 U \operatorname{adj}\left(U^\dagger S_1 U\right)\right] \\ &\quad = \det\left[U^\dagger S_1 U \operatorname{adj}(U) \operatorname{adj}(S_2) \operatorname{adj}\left(U^\dagger\right)\right. \\ &\qquad\quad \left. - U^\dagger S_2 U \operatorname{adj}(U) \operatorname{adj}(S_1) \operatorname{adj}\left(U^\dagger\right)\right]\,. \end{aligned}$$

Since $U \operatorname{adj}(U) = \det(U) I$, we have

$$\begin{aligned} &F_1\left(U^\dagger S_1 U, U^\dagger S_2 U\right) \\ &\; = [\det(U)]^n \det\left[U^\dagger S_1 \operatorname{adj}(S_2) \operatorname{adj}\left(U^\dagger\right)\right. \\ &\qquad \left. - U^\dagger S_2 \operatorname{adj}(S_1) \operatorname{adj}\left(U^\dagger\right)\right] \\ &\; = [\det(U)]^n \det\left[U^\dagger \left(S_1 \operatorname{adj}(S_2) - S_2 \operatorname{adj}(S_1)\right) \operatorname{adj}\left(U^\dagger\right)\right] \\ &\; = [\det(U)]^n \det\left[U^\dagger \operatorname{adj}\left(U^\dagger\right)\right] \det\left[S_1 \operatorname{adj}(S_2) - S_2 \operatorname{adj}(S_1)\right] \\ &\; = (\det U)^n \left[\det\left(U^\dagger\right)\right]^n \det\left[S_1 \operatorname{adj}(S_2) - S_2 \operatorname{adj}(S_1)\right] \\ &\; = \left[\det\left(UU^\dagger\right)\right]^n \det\left[S_1 \operatorname{adj}(S_2) - S_2 \operatorname{adj}(S_1)\right] \\ &\; = \det(I)^n \det\left[S_1 \operatorname{adj}(S_2) - S_2 \operatorname{adj}(S_1)\right] \\ &\; = \det\left[S_1 \operatorname{adj}(S_2) - S_2 \operatorname{adj}(S_1)\right] \\ &\; = F_1(S_1, S_2)\,. \end{aligned}$$

■

APPENDIX C
REDUCTION FROM 3×3 TO 2×2 IN THE RATELESS PROBLEM

Recall that the original problem was to perform 2-GMD (16) to the following two 3×3 matrices, both having a determinant equal to 1:

$$\begin{aligned} A_1 &= \begin{pmatrix} b^4 & 0 & 0 \\ 0 & b^{-2} & 0 \\ 0 & 0 & b^{-2} \end{pmatrix} \\ A_2 &= \begin{pmatrix} b & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & b^{-2} \end{pmatrix}. \end{aligned}$$

Since these two matrices are diagonal, we can assume, without loss of generality, that the elements in the first column of the matrix V in (16) are positive real-valued (since the phase can be canceled by the matrices U_k). Also, the first columns of $A_1 V$ and of $A_2 V$ must have norms equal to 1, and thus

$$V = \begin{pmatrix} v_{11} & * & * \\ v_{21} & * & * \\ v_{31} & * & * \end{pmatrix},$$

where

$$\begin{aligned} v_{11} &= \frac{1}{\sqrt{b^8 + b^4 + 1}} \\ v_{21} &= \frac{b^3}{\sqrt{b^8 + b^4 + 1}} \\ v_{31} &= \frac{b^2}{\sqrt{b^4 + b^2 + 1}}\,. \end{aligned}$$

The remaining two columns must lay in the orthogonal complement to the subspace spanned by this vector, which is spanned by the two vectors $(v_{12}, v_{22}, v_{32})^T$ and $(v_{13}, v_{23}, v_{33})^T$ where

$$\begin{aligned} v_{12} &= \frac{b^3}{\sqrt{b^6 + 1}} \\ v_{22} &= \frac{-1}{\sqrt{b^6 + 1}} \\ v_{32} &= 0 \\ v_{13} &= \frac{b^2}{\sqrt{(b^2 + 1)(b^8 + b^4 + 1))}} \\ v_{23} &= \frac{b^5}{\sqrt{(b^2 + 1)(b^8 + b^4 + 1)}} \\ v_{33} &= -\frac{\sqrt{1 + b^6}}{\sqrt{b^8 + b^4 + 1}}\,. \end{aligned}$$

In other words, we can represent V as

$$V = V_0 \begin{pmatrix} 1 & 0 & 0 \\ 0 & W_{11} & W_{12} \\ 0 & W_{21} & W_{22} \end{pmatrix} ,$$

where

$$V_0 = \begin{pmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{pmatrix} ,$$

and W is a 2×2 unitary matrix. Thus, the matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & W_{11} & W_{12} \\ 0 & W_{21} & W_{22} \end{pmatrix} ,$$

performs 2-GMD on the two matrices

$$A_1 V_0 = \begin{pmatrix} b^4 v_{11} & b^4 v_{12} & b^4 v_{13} \\ b^{-2} v_{21} & b^{-2} v_{22} & b^{-2} v_{23} \\ b^{-2} v_{31} & b^{-2} v_{32} & b^{-2} v_{33} \end{pmatrix}$$
$$A_2 V_0 = \begin{pmatrix} b v_{11} & b v_{12} & b v_{13} \\ b v_{21} & b v_{22} & b v_{23} \\ b^{-2} v_{31} & b^{-2} v_{32} & b^{-2} v_{33} \end{pmatrix} .$$

or, equivalently, on the same matrices after Gram-Schmidt orthogonalization (i.e., QR decomposition):

$$U_1^\dagger A_1 V_0 = \begin{pmatrix} 1 & * & * \\ 0 & \frac{\sqrt{1-b^2+b^8}}{b^2} & \frac{b^6-1}{b\sqrt{(1-b^2+b^8)(1+b^2+b^4)}} \\ 0 & 0 & \frac{b^2}{\sqrt{1-b^2+b^8}} \end{pmatrix} ,$$
$$U_2^\dagger A_2 V_0 = \begin{pmatrix} 1 & * & * \\ 0 & b & 0 \\ 0 & 0 & b^{-1} \end{pmatrix} .$$

In other words, W performs 2-GMD on the two following matrices:

$$\tilde{A}_1 = \begin{pmatrix} \frac{\sqrt{1-b^2+b^8}}{b^2} & \frac{b^6-1}{b\sqrt{(1-b^2+b^8)(1+b^2+b^4)}} \\ 0 & \frac{b^2}{\sqrt{1-b^2+b^8}} \end{pmatrix}$$
$$\tilde{A}_2 = \begin{pmatrix} b & 0 \\ 0 & b^{-1} \end{pmatrix} ,$$

which is what we wanted to prove. ■

APPENDIX D
PROOF OF THEOREM 3

Let A_1 and A_2 be two complex-valued 2×2 matrices with determinants equal to 1. Define:

$$S_1 \triangleq A_1^\dagger A_1 - I$$
$$S_2 \triangleq A_2^\dagger A_2 - I .$$

Let $N \geq 2$, and define the following extended matrices:

$$\begin{matrix} \mathcal{A}_k \triangleq \lceil A_k \rfloor_{\otimes N} \\ \mathcal{S}_k \triangleq \lceil S_k \rfloor_{\otimes N} \end{matrix} \qquad k = 1, 2 .$$

Now, assume that there exist complex-valued unitary matrices $\mathcal{U}_1, \mathcal{U}_2, \mathcal{V}$ such that

$$\mathcal{U}_k^\dagger \mathcal{A}_k \mathcal{V} = \mathcal{T}_k , \quad k = 1, 2 , \tag{62}$$

where $\mathcal{T}_k$ are upper triangular with all the diagonal values equal 1. In particular, if we denote the first column of $\mathcal{V}$ by ϑ, then necessary (although not sufficient) conditions for the existence of the decomposition (62) are

$$\|\mathcal{A}_1 \vartheta\|^2 = 1$$
$$\|\mathcal{A}_2 \vartheta\|^2 = 1$$
$$\|\vartheta\|^2 = 1 ,$$

or equivalently,

$$\vartheta^\dagger \mathcal{S}_1 \vartheta = 0 \tag{63a}$$
$$\vartheta^\dagger \mathcal{S}_2 \vartheta = 0 \tag{63b}$$
$$\vartheta^\dagger \vartheta = 1 . \tag{63c}$$

As in the proof of Lemma 2, we can assume, without loss of generality, that S_1 is real-valued and diagonal. Denoting

$$\vartheta = \begin{pmatrix} x_1 + i x_2 \\ y_1 + i y_2 \\ \vdots \\ x_{2N-1} + i x_{2N} \\ y_{2N-1} + i y_{2N} \end{pmatrix} ,$$
$$S_1 = \begin{pmatrix} a_1 & 0 \\ 0 & c_1 \end{pmatrix}$$
$$S_2 = \begin{pmatrix} a_2 & b_2 + i\beta_2 \\ b_2 - i\beta_2 & c_2 \end{pmatrix} ,$$

the three equations (63) become

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ a_1 & 0 & 0 & c_1 \\ a_2 & b_2 & \beta_2 & c_2 \end{bmatrix} \begin{bmatrix} X_1 + \cdots + X_N \\ 2(W_1 + \cdots + W_N) \\ 2(Z_1 + \cdots + Z_N) \\ Y_1 + \cdots + Y_N \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \tag{64}$$

where we define

$$X_j \triangleq x_{2j-1}^2 + x_{2j}^2$$
$$W_j \triangleq x_{2j-1} y_{2j-1} + x_{2j} y_{2j}$$
$$Z_j \triangleq x_{2j} y_{2j-1} - x_{2j-1} y_{2j}$$
$$Y_j \triangleq y_{2j-1}^2 + y_{2j}^2 .$$

We now consider the following cases.

a) Case 1: Assume first that $a_1 \neq c_1$ and $b_2 \neq 0$. Thus, (64) is equivalent to:

$$\begin{bmatrix} X_1 + \cdots + X_N \\ 2(W_1 + \cdots + W_N) \\ 2(Z_1 + \cdots + Z_N) \\ Y_1 + \cdots + Y_N \end{bmatrix} = \overbrace{\begin{bmatrix} 1 & 0 & 0 & 1 \\ a_1 & 0 & 0 & c_1 \\ a_2 & b_2 & \beta_2 & c_2 \\ 0 & 0 & 1 & 0 \end{bmatrix}}^{B}{}^{-1} \begin{bmatrix} 1 \\ 0 \\ 0 \\ t \end{bmatrix}$$
$$\triangleq \frac{1}{\Delta} \begin{bmatrix} f_1(t) \\ f_2(t) \\ f_3(t) \\ f_4(t) \end{bmatrix} ,$$

where t is some real-valued parameter, $f_1(t)$,$f_2(t)$,$f_3(t)$,$f_4(t)$ are first-degree polynomials in t (with coefficients that depend on the matrices S_1, S_2), and

$$\Delta \triangleq \det B = b_2(c_1 - a_1) \neq 0 . \tag{65}$$

Thus, finding a solution ϑ to the original problem is *equivalent* to finding a solution $(x_1, \cdots, x_{2N}, y_1, \cdots, y_{2N}, t)$ to the following equations:

$$X_1 + \cdots + X_N = \frac{1}{\Delta} f_1(t) \tag{66a}$$

$$2(W_1 + \cdots + W_N) = \frac{1}{\Delta} f_2(t) \tag{66b}$$

$$2(Z_1 + \cdots + Z_N) = \frac{1}{\Delta} f_3(t) \tag{66c}$$

$$Y_1 + \cdots + Y_N = \frac{1}{\Delta} f_4(t)\,. \tag{66d}$$

Assertion 3: A solution to (66) exists if and only if the following conditions hold for some $t \in \mathbb{R}$:

$$\frac{1}{\Delta} f_1(t) \geq 0 \tag{67a}$$

$$\frac{1}{\Delta} f_4(t) \geq 0 \tag{67b}$$

$$4 f_1(t) f_4(t) \geq f_2^2(t) + f_3^2(t)\,. \tag{67c}$$

Proof of Assertion 3: Construct the following three vectors:

$$\begin{aligned}
\mathbf{p}_1 &= (x_2, -x_1, x_4, -x_3, \cdots, x_{2N}, x_{2N-1}) \\
\mathbf{p}_2 &= (y_1, y_2, y_3, y_4, \cdots, y_{2N-1}, y_{2N}) \\
\mathbf{p}_3 &= (x_1, x_2, x_3, x_4, \cdots, x_{2N-1}, x_{2N})\,.
\end{aligned}$$

Using the inner product definition, we have

$$\begin{aligned}
||\mathbf{p}_1||^2 &= ||\mathbf{p}_3||^2 \\
&= x_1^2 + x_2^2 + x_3^2 + x_4^2 + \cdots + x_{2N-1}^2 + x_{2N}^2 \\
&= X_1 + X_2 \cdots + X_N \\
||\mathbf{p}_2||^2 &= y_1^2 + y_2^2 + y_3^2 + y_4^2 + \cdots + y_{2N-1}^2 + y_{2N}^2 \\
&= Y_1 + Y_2 + \cdots + Y_N \\
2\langle \mathbf{p}_1, \mathbf{p}_2 \rangle &= 2(x_2 y_1 - x_1 y_2 + \cdots + x_{2N} y_{2N-1} - x_{2N-1} y_{2N}) \\
&= 2(Z_1 + Z_2 + \cdots + Z_N) \\
2\langle \mathbf{p}_2, \mathbf{p}_3 \rangle &= 2(x_1 y_1 + x_2 y_2 + \cdots + x_{2N-1} y_{2N-1} + x_{2N} y_{2N}) \\
&= 2(W_1 + W_2 + \cdots + W_N)\,,
\end{aligned} \tag{68}$$

and the angles between these vectors satisfy

$$\begin{aligned}
\cos\theta_1 &= \frac{\langle \mathbf{p}_1, \mathbf{p}_2 \rangle}{||\mathbf{p}_1||||\mathbf{p}_2||} \\
\cos\theta_2 &= \frac{\langle \mathbf{p}_3, \mathbf{p}_2 \rangle}{||\mathbf{p}_3||||\mathbf{p}_2||} = \frac{\langle \mathbf{p}_3, \mathbf{p}_2 \rangle}{||\mathbf{p}_1||||\mathbf{p}_2||}\,.
\end{aligned}$$

Note that the l.h.s. of (66) and the r.h.s. of (68) coincide, and that $\langle \mathbf{p}_3, \mathbf{p}_1 \rangle = 0$. Therefore the angle between them is $\pi/2$. One verifies that the maximum of $\cos^2\theta_1 + \cos^2\theta_2$ is achieved when all three vectors are on the same plane, in which case $\cos\theta_2 = \cos(\pm\pi/2 - \theta_1) = \pm\sin\theta_1$, which implies that $\cos^2\theta_1 + \cos^2\theta_2 = 1$. When the three vectors do not lay on the same plane, $\cos^2\theta_1 + \cos^2\theta_2 < 1$.

Thus, a solution to (66) exists if and only if

$$||\mathbf{p}_1||^2 \geq 0 \tag{69a}$$

$$||\mathbf{p}_2||^2 \geq 0 \tag{69b}$$

$$\cos^2\theta_1 + \cos^2\theta_2 = \frac{\langle \mathbf{p}_1, \mathbf{p}_2 \rangle^2}{||\mathbf{p}_1||^2||\mathbf{p}_2||^2} + \frac{\langle \mathbf{p}_3, \mathbf{p}_2 \rangle^2}{||\mathbf{p}_1||^2||\mathbf{p}_2||^2} \leq 1, \tag{69c}$$

where (69c) is equivalent to

$$4||\mathbf{p}_1||^2||\mathbf{p}_2||^2 - (2\langle \mathbf{p}_1, \mathbf{p}_2 \rangle)^2 - (2\langle \mathbf{p}_3, \mathbf{p}_2 \rangle)^2 \geq 0\,,$$

which is equivalent, in turn, to (67). ■

By definition, and using (65), $(f_1(t) + f_4(t)) = \Delta$. Therefore, these three conditions are equivalent to the following single condition:

$$4 f_1(t) f_4(t) - f_2^2(t) - f_3^2(t) \geq 0\,.$$

This is a quadratic inequality in t,

$$a t^2 + b t + c \geq 0\,, \tag{70}$$

where the constants a, b, c are as in (56). Note that since $a_1 \neq c_1$ and $b_2 \neq 0$, the coefficient a is *strictly* negative. Therefore, a necessary and sufficient condition for the existence of a (real-valued) solution t to the inequality in (70) is for the discriminant to be non-negative:

$$b^2 - 4ac \geq 0\,.$$

A direct calculation shows that

$$b^2 - 4ac = 4\Delta^2 F_1(A_1^\dagger A_1 - I, A_2^\dagger A_2 - I)\,,$$

where F_1 is defined as in (27). This condition is the same as the condition in (26) which completes the proof of Theorem 3 for this case.

b) Case 2: Assume now that $a_1 = c_1$. As in case 2 in the proof of Lemma 2, condition (26) holds, and thus this case is not possible under the assumptions of the theorem.

c) Case 3: Next, assume that $a_1 \neq c_1$, $b_2 = 0$, and $\beta_2 \neq 0$. Thus, (64) becomes

$$\begin{pmatrix} 1 & 0 & 1 \\ a_1 & 0 & c_1 \\ a_2 & \beta_2 & c_2 \end{pmatrix} \begin{pmatrix} X_1 + \cdots + X_N \\ 2(Z_1 + \cdots + Z_N) \\ Y_1 + \cdots + Y_N \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\,,$$

which reduces to

$$\begin{pmatrix} X_1 + \cdots + X_N \\ 2(Z_1 + \cdots + Z_N) \\ Y_1 + \cdots + Y_N \end{pmatrix} \tag{71a}$$

$$= \begin{pmatrix} f_5 \\ f_6 \\ f_7 \end{pmatrix} \triangleq \frac{1}{(a_1 - c_1)\beta_2} \begin{pmatrix} -\beta_2 c_1 \\ a_2 c_1 - a_1 c_2 \\ a_1 \beta_2 \end{pmatrix}\,. \tag{71b}$$

Assertion 4: A solution to (71) exists if and only if the following conditions holds:

$$f_5 \geq 0 \tag{72a}$$

$$f_7 \geq 0 \tag{72b}$$

$$4 f_5 f_7 - f_6^2 \geq 0\,. \tag{72c}$$

Proof of Assertion 4: Construct the following two vectors:

$$\begin{aligned}
\mathbf{p}_1 &= (x_2, -x_1, x_4, -x_3, \cdots, x_{2N}, x_{2N-1}) \\
\mathbf{p}_2 &= (y_1, y_2, y_3, y_4, \cdots, y_{2N-1}, y_{2N})\,.
\end{aligned}$$

Using the inner product definition, we have

$$\begin{aligned}
||\mathbf{p}_1||^2 &= x_1^2 + x_2^2 + x_3^2 + x_4^2 + \cdots + x_{2N-1}^2 + x_{2N}^2 \\
&= X_1 + X_2 \cdots + X_N \\
||\mathbf{p}_2||^2 &= y_1^2 + y_2^2 + y_3^2 + y_4^2 + \cdots + y_{2N-1}^2 + y_{2N}^2 \\
&= Y_1 + Y_2 + \cdots + Y_N \\
2\langle \mathbf{p}_1, \mathbf{p}_2 \rangle &= 2(x_2 y_1 - x_1 y_2 + \cdots + x_{2N} y_{2N-1} - x_{2N-1} y_{2N}) \\
&= 2(Z_1 + Z_2 + \cdots + Z_N) .
\end{aligned} \tag{73}$$

and the angle between the two vectors satisfies

$$\cos\theta_1 = \frac{\langle \mathbf{p}_1, \mathbf{p}_2 \rangle}{||\mathbf{p}_1|| ||\mathbf{p}_2||} .$$

Note that the l.h.s. of (71) and the r.h.s. of (73) coincide. Thus, a solution to (71) exists if and only if

$$||\mathbf{p}_1||^2 \geq 0 \tag{74a}$$

$$||\mathbf{p}_2||^2 \geq 0 \tag{74b}$$

$$\frac{\langle \mathbf{p}_1, \mathbf{p}_2 \rangle}{||\mathbf{p}_1|| ||\mathbf{p}_2||} \leq 1 \tag{74c}$$

where (74c) is equivalent to

$$4||\mathbf{p}_1||^2 ||\mathbf{p}_2||^2 - (2\langle \mathbf{p}_1, \mathbf{p}_2 \rangle)^2 \geq 0 ,$$

which is equivalent, in turn, to (72). ■

From Assertion 4, a necessary condition for the existence of a solution to (71) is

$$4 f_5 f_7 - f_6^2 \geq 0 ,$$

which is equivalent, in turn, to

$$-(a_2 c_1 - a_1 c_2)^2 - 4 a_1 c_1 \beta_2^2 \geq 0 .$$

On the other hand,

$$F_1(S_1, S_2) = -(a_2 c_1 - a_1 c_2)^2 - 4 a_1 c_1 \beta_2^2 .$$

Thus condition (26) must hold true, since otherwise no solution to (63) exists.

d) Case 4: We are left with the case where $a_1 \neq c_1$, $b_2 = 0$, and $\beta_2 = 0$. In this case, (64) reduces to

$$\begin{pmatrix} 1 & 1 \\ a_1 & c_1 \\ a_2 & c_2 \end{pmatrix} \begin{pmatrix} X_1 + \cdots + X_N \\ Y_1 + \cdots + Y_N \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} .$$

A necessary condition for the existence of a solution in this case, is that the second and the third rows are linearly dependent, i.e., $a_1 c_2 = a_2 c_1$. On the other hand,

$$F_1(S_1, S_2) = -(a_2 c_1 - a_1 c_2)^2 .$$

Thus if condition (26) does not hold, no solution to (63) exists.

This concludes the proof of the theorem. ■

APPENDIX E
PROOF OF LEMMA 3

First, assume that statement 2 holds. Namely, There exist $K+2$ matrices with orthonormal columns $U_1, \ldots, U_{K+1}, V$, of dimensions $n \times \tilde{n}$, such that

$$U_k^\dagger A_k V = R_k , \qquad k = 1, \ldots, K+1 ,$$

where $\{R_k\}$ are $\tilde{n} \times \tilde{n}$ upper triangular with equal diagonals. Now, arbitrarily extend V to an $n \times n$ unitary matrix:

$$\tilde{V} = \left(\begin{array}{c|c} V & V^\perp \end{array} \right) .$$

Then, U_k can also be extended to $n \times n$ unitary matrices, by performing Gram-Schmidt process on the columns of $A_k \tilde{V}$:

$$\tilde{U}_k = \left(\begin{array}{c|c} U_k & U_k^\perp \end{array} \right) ,$$

such that

$$\tilde{U}_k^\dagger A_k \tilde{V} = \tilde{R}_k = \left(\begin{array}{c|c} R_k & * \\ \hline 0 & \hat{R}_k \end{array} \right) ,$$

and $\hat{R}_k$ are upper triangular (with diagonal elements that depend on k). Thus, we have:

$$\begin{aligned}
\tilde{U}_k^\dagger B_k \tilde{U}_{K+1} &= \tilde{U}_k^\dagger A_k A_{K+1}^{-1} \tilde{U}_{K+1} \\
&= \tilde{U}_k^\dagger A_k \tilde{V} \tilde{V}^\dagger A_{K+1}^{-1} \tilde{U}_{K+1} \\
&= \tilde{R}_k \tilde{R}_{K+1}^{-1} \\
&= \tilde{T}_k ,
\end{aligned}$$

where $\tilde{T}_k$ is of the form

$$\tilde{T}_k = \left(\begin{array}{c|c} T_k & * \\ \hline 0 & \hat{T}_k \end{array} \right) ,$$

where T_k is upper triangular with all the diagonal elements equal to 1, and $\hat{T}_k$ is upper triangular (with diagonal elements that depend on k). By substitution:

$$\left(\begin{array}{c} U_k^\dagger \\ \hline (U_k^\perp)^\dagger \end{array} \right) B_k \left(\begin{array}{c|c} U_{K+1} & U_{K+1}^\perp \end{array} \right) = \left(\begin{array}{c|c} T_k & * \\ \hline 0 & \hat{T}_k \end{array} \right) .$$

By taking only the first $\tilde{n}$ rows and the first $\tilde{n}$ columns of this equality, we obtain

$$U_k^\dagger B_k U_{K+1} = T_k ,$$

which results in statement 1.

Now, assume that statement 1 holds. Perform the QR decomposition on the matrix $A_{K+1}^{-1} U_{K+1}$:

$$A_{K+1}^{-1} U_{K+1} = V R ,$$

where V is of dimensions $n \times \tilde{n}$ with orthonormal columns, and R is an $\tilde{n} \times \tilde{n}$ upper triangular matrix. Thus, using (39), we obtain the following equalities:

$$\begin{aligned}
U_k^\dagger A_k V R &= U_k^\dagger A_k A_{K+1}^{-1} U_{K+1} \\
&= U_k^\dagger B_k U_{K+1} , \qquad k = 1, \ldots, K ,
\end{aligned}$$

which, according to (40), suggest

$$U_k^\dagger A_k V R = T_k , \qquad k = 1, \ldots, K . \tag{75}$$

On the other hand, we have

$$U_{K+1}^{\dagger} A_{K+1} V R = U_{K+1}^{\dagger} A_{K+1} A_{K+1}^{-1} U_{K+1} \qquad (76a)$$

$$= U_{K+1}^{\dagger} U_{K+1} = I \, . \qquad (76b)$$

Multiplying (75) and (76) by R^{-1} on the right yields:

$$U_k^{\dagger} A_k V = T_k R^{-1} \, , \qquad k = 1, \ldots, K$$

$$U_{K+1}^{\dagger} A_{K+1} V = R^{-1} \, .$$

Since T_k are upper triangular with only 1s on the diagonal, the matrices $R_k \triangleq T_k R^{-1}$ ($k = 1, \ldots, K$) and $R_{K+1} \triangleq R^{-1}$ have equal diagonals, thus statement 2 holds.

This completes the proof. ■

APPENDIX F
PROOF OF THEOREM 4 FOR $n = 2$, $K = 3$, $N = 4$

The proof will be based on $K = 3$ steps.

Denote by $\{\mathcal{A}_k\}$ the extended matrices corresponding to $N = 4$ channel uses.

Step 1:

Start by applying a 1-GMD for each block (corresponding to a single channel use) of the first matrix A_1:

$$\left(U_1^{(1)}\right)^{\dagger} A_1 V^{(1)} = \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} ,$$

which corresponds, in turn, to applying the following extended unitary matrices (recall the definition of the embedding operation)

$$\left(\mathcal{U}_1^{(1)}\right)^{\dagger} \triangleq I_8 \left[\left(U_1^{(1)}\right)^{\dagger} ; \lfloor 1, 2 \rceil \lfloor 3, 4 \rceil \lfloor 5, 6 \rceil \lfloor 7, 8 \rceil \right] ,$$

$$\mathcal{V}^{(1)} \triangleq I_8 \left[V^{(1)} ; \lfloor 1, 2 \rceil \lfloor 3, 4 \rceil \lfloor 5, 6 \rceil \lfloor 7, 8 \rceil \right] ,$$

and results in the following extended triangular matrix

$$\mathcal{T}_1^{(1)} = \left(\mathcal{U}_1^{(1)}\right)^{\dagger} \mathcal{A}_1 \mathcal{V}^{(1)}$$

$$= \begin{pmatrix} 1 & * & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & * & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & * & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} .$$

Note that the same matrix $\mathcal{V}^{(1)}$ has to be applied to all matrices (since the encoder is shared by all users). We decompose the resulting matrices (after multiplying them by $\mathcal{V}^{(1)}$) according to the QR decomposition, resulting in unitary matrices $\left(\mathcal{U}_k^{(1)}\right)^{\dagger}$ such that:

$$\mathcal{T}_k^{(1)} = \left(\mathcal{U}_k^{(1)}\right)^{\dagger} \mathcal{A}_k \mathcal{V}^{(1)}$$

$$= \begin{pmatrix} r_1^k & * & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & r_2^k & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & r_1^k & * & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & r_2^k & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & r_1^k & * & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & r_2^k & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & r_1^k & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & r_2^k \end{pmatrix} ,$$

where $r_1^k \, r_2^k = 1$ and $k = 2, 3$.

Step 2:

In the second step we apply the 1-GMD to the matrices $\mathcal{T}_2^{(1)} \lfloor 2, 5 \rceil$ and $\mathcal{T}_2^{(1)} \lfloor 4, 7 \rceil$. In both cases the two-by-two matrices are of the same form:

$$\left(U_2^{(2)}\right)^{\dagger} \begin{pmatrix} r_2^2 & 0 \\ 0 & r_1^2 \end{pmatrix} V^{(2)} = \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} .$$

Now note that the matrix corresponding to these elements in $\mathcal{T}_1^{(1)}$ have the identity matrix form I_2. Thus, by Property 1, applying $V^{(2)}$ on the right and $\left(V^{(2)}\right)^{\dagger}$ on the left results in the identity matrix, i.e., $\mathcal{T}_1^{(1)} \lfloor 2, 5 \rceil$ and $\mathcal{T}_1^{(1)} \lfloor 4, 7 \rceil$ remain unchanged.

For the third matrix, we apply the QR decomposition with $\left(\mathcal{U}_3^{(2)}\right)^{\dagger}$ (assuming no special structure).

Define

$$\left(\mathcal{U}_2^{(2)}\right)^{\dagger} \triangleq I_8 \left[\left(U_2^{(2)}\right)^{\dagger} ; \lfloor 2, 5 \rceil \lfloor 4, 7 \rceil \right] ,$$

$$\mathcal{V}^{(2)} \triangleq I_8 \left[V^{(2)} ; \lfloor 2, 5 \rceil \lfloor 4, 7 \rceil \right] .$$

Thus, we attain the following matrices after the completion of the second step:

$$\mathcal{T}_2^{(2)} = \left(\mathcal{U}_2^{(2)}\right)^{\dagger} \mathcal{T}_2^{(1)} \mathcal{V}^{(2)}$$

$$= \begin{pmatrix} r_1^2 & * & 0 & 0 & * & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & * & * & 0 & 0 \\ 0 & 0 & r_1^2 & * & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & 1 & * & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & r_2^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & r_2^2 \end{pmatrix} ,$$

$$\mathcal{T}_1^{(2)} = \left(\mathcal{V}^{(2)}\right)^{\dagger} \mathcal{T}_1^{(1)} \mathcal{V}^{(2)}$$

$$= \begin{pmatrix} 1 & * & 0 & 0 & * & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & * & 0 & 0 \\ 0 & 0 & 1 & * & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 1 & * & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} ,$$

$$\mathcal{T}_3^{(2)} = \left(\mathcal{U}_3^{(2)}\right)^{\dagger} \mathcal{T}_3^{(1)} \mathcal{V}^{(2)}$$
$$= \begin{pmatrix} r_1^3 & * & 0 & 0 & * & 0 & 0 & 0 \\ 0 & d_2 & 0 & 0 & * & * & 0 & 0 \\ 0 & 0 & r_1^3 & * & 0 & 0 & * & 0 \\ 0 & 0 & 0 & d_2 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & d_1 & * & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & r_2^3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & d_1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & r_2^3 \end{pmatrix},$$

where $d_1\, d_2 = 1$.

Step 3:
Finally, apply the 1-GMD to $\mathcal{T}_3^{(2)} \lfloor 4,\, 5 \rceil$:

$$\left(U_3^{(3)}\right)^{\dagger} \begin{pmatrix} d_2 & 0 \\ 0 & d_1 \end{pmatrix} V^{(3)} = \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}.$$

Again, note that the corresponding sub-matrices of $\mathcal{T}_1^{(3)}$ and $\mathcal{T}_2^{(3)}$ are equal to I_2. Thus by Property 1, multiplying them by $V^{(3)}$ on the right and $\left(V^{(3)}\right)^{\dagger}$ on the left, gives rise to the identity matrix I_2. By defining

$$\left(\mathcal{U}_3^{(3)}\right)^{\dagger} \triangleq I_8 \left[\left(U_3^{(3)}\right)^{\dagger}; \lfloor 4,\, 5 \rceil\right],$$
$$\mathcal{V}^{(3)} \triangleq I_8 \left[V^{(3)}; \lfloor 4,\, 5 \rceil\right],$$
$$\left(\mathcal{U}_1^{(3)}\right)^{\dagger} = \left(\mathcal{U}_2^{(3)}\right)^{\dagger} \triangleq \left(\mathcal{V}^{(3)}\right)^{\dagger},$$

we arrive to the following three triangular matrices:

$$\mathcal{T}_3^{(3)} = \left(\mathcal{U}_3^{(3)}\right)^{\dagger} \mathcal{T}_3^{(2)} \mathcal{V}^{(3)}$$
$$= \begin{pmatrix} r_1^3 & * & 0 & * & * & 0 & 0 & 0 \\ 0 & d_2 & 0 & * & * & * & 0 & 0 \\ 0 & 0 & r_1^3 & * & * & 0 & * & 0 \\ 0 & 0 & 0 & 1 & * & * & * & * \\ 0 & 0 & 0 & 0 & 1 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & r_2^3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & d_1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & r_2^3 \end{pmatrix},$$

$$\mathcal{T}_2^{(3)} = \left(\mathcal{V}^{(3)}\right)^{\dagger} \mathcal{T}_2^{(2)} \mathcal{V}^{(3)}$$
$$= \begin{pmatrix} r_1^2 & * & 0 & * & * & 0 & 0 & 0 \\ 0 & 1 & 0 & * & * & * & 0 & 0 \\ 0 & 0 & r_1^2 & * & * & 0 & * & 0 \\ 0 & 0 & 0 & 1 & 0 & * & * & * \\ 0 & 0 & 0 & 0 & 1 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & r_2^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & r_2^2 \end{pmatrix},$$

$$\mathcal{T}_1^{(3)} = \left(\mathcal{V}^{(3)}\right)^{\dagger} \mathcal{T}_1^{(2)} \mathcal{V}^{(3)}$$
$$= \begin{pmatrix} 1 & * & 0 & * & * & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & * & 0 & 0 \\ 0 & 0 & 1 & * & * & 0 & * & 0 \\ 0 & 0 & 0 & 1 & 0 & * & 0 & * \\ 0 & 0 & 0 & 0 & 1 & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

By taking the middle rows and columns (rows and columns 4 and 5) we achieve the desired decomposition with diagonal elements equaling to 1 in all three triangular matrices, simultaneously. Formally, we do so by multiplying $\left(\mathcal{I}_8^{[4,5]}\right)^{\dagger}$ on the left and by $\mathcal{I}_8^{[4,5]}$ on the right (see Remark 18) to achieve:

$$\left(\mathcal{I}_8^{[4,5]}\right)^{\dagger} \mathcal{T}_k^{(3)} \mathcal{I}_8^{[4,5]} = \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}.$$

Thus, by defining

$$\mathcal{V} = \mathcal{V}^{(1)} \mathcal{V}^{(2)} \mathcal{V}^{(3)} \mathcal{I}_8^{[4,5]}$$
$$(\mathcal{U}_k)^{\dagger} = \left(\mathcal{I}_8^{[4,5]}\right)^{\dagger} \left(\mathcal{U}_k^{(3)}\right)^{\dagger} \left(\mathcal{U}_k^{(2)}\right)^{\dagger} \left(\mathcal{U}_k^{()}\right)^{\dagger}, \quad k = 1, 2, 3,$$

we arrive at the desired result. ■

APPENDIX G
PROOF OF THEOREM 4 FOR $n = 2$ AND GENERAL K, N

For K users, we use the same idea, i.e., applying two-by-two 1-GMD operations sequentially on the different channel matrices. Thus, stating the indices of the four-tuples for which 1-GMD is applied at each step (for each matrix), suffices to establish the desired construction.
The proof will be based on K steps.

Denote by $\{\mathcal{A}_k\}$ the extended matrices corresponding to N channel uses.

Step 1:
Perform 1-GMD (corresponding to a single channel use) on the matrix A_1: $\left(U_1^{(1)}\right)^{\dagger} A_1 V^{(1)}$.
Then, we apply this decomposition to each block separately, using:

$$\left(\mathcal{U}_1^{(1)}\right)^{\dagger} \triangleq I_{2N} \left[\left(U_1^{(1)}\right)^{\dagger}; \lfloor 1,\, 2 \rceil \lfloor 3,\, 4 \rceil \cdots \lfloor 2N-1,\, 2N \rceil\right],$$
$$\mathcal{V}^{(1)} \triangleq I_{2N} \left[V^{(1)}; \lfloor 1,\, 2 \rceil \lfloor 3,\, 4 \rceil \cdots \lfloor 2N-1,\, 2N \rceil\right].$$

Then, we need to apply the same matrix $\mathcal{V}^{(1)}$ to all matrices (since the encoder is shared by all users). We decompose the resulting matrices (after multiplying them by $\mathcal{V}^{(1)}$) according to the QR decomposition, resulting in unitary matrices $\left(\mathcal{U}_k^{(1)}\right)^{\dagger}$. We denote the resulting extended triangular matrices by $\mathcal{T}_k^{(1)} = \left(\mathcal{U}_k^{(1)}\right)^{\dagger} \mathcal{A}_k \mathcal{V}^{(1)}$.

Step 2:
Perform 1-GMD on the matrix $\mathcal{T}_2^{(1)} \lfloor 2,\, 2^{K-1}+1 \rceil$:

$$\left(U_2^{(2)}\right)^{\dagger} \left(\mathcal{T}_2^{(1)} \lfloor 2,\, 2^{K-1}+1 \rceil\right) V^{(2)}.$$

Then, apply this decomposition to each of the matrices, using:

$$\left(\mathcal{U}_2^{(2)}\right)^\dagger \triangleq I_{2N}\left[\left(U_2^{(2)}\right)^\dagger; \bigcup_q \lfloor 2q,\, 2^{K-1}+2q-1\rceil\right],$$

$$\mathcal{V}^{(2)} \triangleq I_{2N}\left[V^{(2)}; \bigcup_q \lfloor 2q,\, 2^{K-1}+2q-1\rceil\right],$$

for all $q \in \{1\,,2\,,\dots\,,N-2^{K-2}\}$.

Note that the submatrices of $\mathcal{T}_1^{(1)}$ in these indices, $\mathcal{T}_1^{(1)}\lfloor 2,\, 2^{K-1}+1\rceil \cdots \mathcal{T}_1^{(1)}\lfloor 2N-2^{K-1},\, 2N-1\rceil$ are equal to I_2; by Property 1, multiplying them by $V^{(2)}$ on the right and $\left(V^{(2)}\right)^\dagger$ on the left, leaves them unchanged.

Then, we need to apply the same matrix $\mathcal{V}^{(2)}$ to all matrices (since the encoder is shared by all users). We decompose the resulting matrices (after multiplying them by $\mathcal{V}^{(2)}$) according to the QR decomposition, resulting in unitary matrices $\left(\mathcal{U}_k^{(2)}\right)^\dagger$. We denote the resulting extended triangular matrices by $\mathcal{T}_k^{(2)} = \left(\mathcal{U}_k^{(2)}\right)^\dagger \mathcal{T}_k^{(1)} \mathcal{V}^{(2)}$.

Step $3 \le l \le K$:
Perform 1-GMD on the matrix $\mathcal{T}_l^{(l-1)}\lfloor 2^{K-1}-2^{K-(l-1)}+2,\, 2^{K-1}+1\rceil$:

$$\left(U_l^{(l)}\right)^\dagger \left(\mathcal{T}_l^{(l-1)}\left\lfloor 2^{K-1}-2^{K-(l-1)}+2,\, 2^{K-1}+1\right\rceil\right) V^{(l)}.$$

Then, apply this decomposition to each of the extended matrices, using:

$$\left(\mathcal{U}_l^{(l)}\right)^\dagger \triangleq$$
$$I_{2N}\left[\left(U_l^{(l)}\right)^\dagger; \bigcup_q \left\lfloor 2^{K-1}-2^{K-(l-1)}+2q,\, 2^{K-1}+2q-1\right\rceil\right]$$
$$\mathcal{V}^{(l)} \triangleq$$
$$I_{2N}\left[V^{(l)}; \bigcup_q \left\lfloor 2^{K-1}-2^{K-(l-1)}+2q,\, 2^{K-1}+2q-1\right\rceil\right]$$

for all $q \in \{1\,,2\,,\dots\,,N-2^{K-2}\}$.

Note that the submatrices of the matrices $\mathcal{T}_j^{(l-1)}$ ($j = 1\,,\dots\,,l-1$) in the same indices are all equal to I_2; by Property 1, multiplying them by $V^{(l)}$ on the right and $\left(V^{(l)}\right)^\dagger$ on the left, leaves them unchanged.

Then, we need to apply the same matrix $\mathcal{V}^{(l)}$ to all matrices (since the encoder is shared by all users). We decompose the resulting matrices (after multiplying them by $\mathcal{V}^{(l)}$) according to the QR decomposition, resulting in unitary matrices $\left(\mathcal{U}_k^{(l)}\right)^\dagger$. We denote the resulting extended triangular matrices by $\mathcal{T}_k^{(l)} = \left(\mathcal{U}_k^{(l)}\right)^\dagger \mathcal{T}_k^{(l-1)} \mathcal{V}^{(l)}$.

Step K:
After performing the last step (step $l = K$), we are left with K matrices, $\mathcal{T}_k^{(K)}$, the central submatrices of which, $\mathcal{T}_k^{(K)}\lfloor 2^{K-1} : 2N-2^{K-1}+1\rceil$, have diagonals equal to 1. We extract these matrices using the following matrix (see Remark 18):

$$\mathcal{O} \triangleq \mathcal{I}_{2N}^{\left[2^{K-1}:2N-2^{K-1}+1\right]}.$$

Thus, by defining

$$(\mathcal{U}_1)^\dagger \triangleq \mathcal{O}^\dagger \left(\mathcal{V}^{(K)}\right)^\dagger \cdots \left(\mathcal{V}^{(2)}\right)^\dagger \left(\mathcal{U}_1^{(1)}\right)^\dagger$$
$$(\mathcal{U}_k)^\dagger \triangleq \mathcal{O}^\dagger \left(\mathcal{V}^{(K)}\right)^\dagger \cdots \left(\mathcal{V}^{(k+1)}\right)^\dagger \left(\mathcal{U}_k^{(k)}\right)^\dagger \cdots \left(\mathcal{U}_k^{(1)}\right)^\dagger$$
$$(\mathcal{U}_K)^\dagger \triangleq \mathcal{O}^\dagger \left(\mathcal{U}_K^{(K)}\right)^\dagger \cdots \left(\mathcal{U}_K^{(1)}\right)^\dagger$$
$$\mathcal{V} \triangleq \mathcal{V}^{(1)}\mathcal{V}^{(2)}\cdots\mathcal{V}^{(K)}\mathcal{O}\,,$$

we arrive at the desired result. ■

APPENDIX H
PROOF OF THEOREM 4 FOR $K = 2$ AND GENERAL n, N

The proof is composed of $K = 2$ steps, where, in the case of general n, the second step consists of two stages.

Step 1:
We start by performing 1-GMD (corresponding to a single channel use) on the first matrix A_1:

$$\left(U_1^{(1)}\right)^\dagger A_1 V^{(1)} = \begin{pmatrix} 1 & * & \cdots & * & * \\ 0 & 1 & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & * \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

Apply this decomposition to each block separately, on the first extended matrix, $\mathcal{A}_1$, using:

$$\left(\mathcal{U}_1^{(1)}\right)^\dagger \triangleq$$
$$I_{2N}\left[\left(U_1^{(1)}\right)^\dagger; \lfloor 1:\, n\rceil\, \lfloor n+1:\, 2n\rceil \dots \lfloor (N-1)n+1:\, Nn\rceil\right]$$
$$\mathcal{V}^{(1)} \triangleq$$
$$I_{2N}\left[V^{(1)}; \lfloor 1:\, n\rceil\, \lfloor n+1:\, 2n\rceil \dots \lfloor (N-1)n+1:\, Nn\rceil\right].$$

Note that the same matrix $\mathcal{V}^{(1)}$ has to be applied to all matrices (since the encoder is shared by all users). We decompose the resulting matrices (after multiplying them by $\mathcal{V}^{(1)}$) according to the QR decomposition, resulting in unitary matrices $\left(\mathcal{U}_2^{(1)}\right)^\dagger$:

$$\mathcal{T}_k^{(1)} \triangleq \left(\mathcal{U}_k^{(1)}\right)^\dagger \mathcal{A}_k \mathcal{V}^{(1)}$$
$$= \left(\begin{array}{c:c:c:c:c} T_k^{(1)} & 0 & \cdots & 0 & 0 \\ \hdashline 0 & T_k^{(1)} & \cdots & 0 & 0 \\ \hdashline \vdots & \vdots & \vdots & \ddots & \vdots \\ \hdashline 0 & 0 & \cdots & T_k^{(1)} & 0 \\ \hdashline 0 & 0 & \cdots & 0 & T_k^{(1)} \end{array}\right), \qquad k = 1, 2\,,$$

where,

$$T_1^{(1)} \triangleq \begin{pmatrix} 1 & * & \cdots & * & * \\ 0 & 1 & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & * \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix},$$

$$T_2^{(1)} \triangleq \begin{pmatrix} r_1 & * & \cdots & * & * \\ 0 & r_2 & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & r_{n-1} & * \\ 0 & 0 & \cdots & 0 & r_n \end{pmatrix} .$$

<u>Step 2:</u>
This step consists of 2 stages: the first is the reordering stage and the second is application of 1 1-GMD to each block.
<u>Stage 1: Reordering</u>

It is convenient to reorder the columns of $\mathcal{T}_k^{(1)}$ such that the columns

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

are "grouped together" for every k.[18] Formally, we do so by applying the $nN \times n(N-n+1)$ reordering matrix

$$\mathcal{O} = \mathcal{J}_{nN}^{\left[kn, kn+(n-1), kn+2(n-1), \cdots, kn+(n-1)^2\right]} .$$

The reordering stage gives rise to the following matrices of dimensions $n(N-n+1) \times n(N-n+1)$:

$$\begin{aligned} \mathcal{T}_k^{(2)(1)} &\triangleq (\mathcal{O})^\dagger \, \mathcal{T}_k^{(1)} \mathcal{O} \\ &= \left(\begin{array}{c|c|c|c|c} T_k^{(2)(1)} & * & \cdots & * & * \\ \hline 0 & T_k^{(2)(1)} & \cdots & * & * \\ \hline \vdots & \vdots & \vdots & \ddots & \vdots \\ \hline 0 & 0 & \cdots & T_k^{(2)(1)} & * \\ \hline 0 & 0 & \cdots & 0 & T_k^{(2)(1)} \end{array} \right) , \\ & \qquad k = 1\,, 2\,, \end{aligned}$$

where,

$$T_1^{(2)(1)} \triangleq \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix} ,$$

$$T_2^{(2)(1)} \triangleq \begin{pmatrix} r_n & 0 & \cdots & 0 & 0 \\ 0 & r_{n-1} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & r_2 & 0 \\ 0 & 0 & \cdots & 0 & r_1 \end{pmatrix} ,$$

the superscripts denote the step and stage number, and the subscripts denote the user number.

<u>Stage 2: 1-GMD</u>
Perform 1-GMD on the matrix $T_2^{(2)(1)}$:

$$\left(U_2^{(2)}\right)^\dagger T_2^{(2)(1)} V^{(2)} = \begin{pmatrix} 1 & * & \cdots & * & * \\ 0 & 1 & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & * \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix} .$$

Note that the matrix $T_1^{(2)(1)}$ is equal to I_n; by Property 1, multiplying it by $V^{(2)}$ on the right and $\left(V^{(2)}\right)^\dagger$ on the left, leaves it unchanged.

We now apply this decomposition to each block separately, using

$$\left(\mathcal{U}_2^{(2)}\right)^\dagger \triangleq I_{n(N-n+1)} \left[\left(U_2^{(2)}\right)^\dagger ; \bigcup_q \lfloor 1 + n(q-1) : qn \rceil \right] ,$$

$$\mathcal{V}^{(2)} \triangleq I_{n(N-n+1)} \left[V^{(2)} ; \bigcup_{q]} \lfloor 1 + n(q-1) : qn \rceil \right] ,$$

for all $q \in \{1\,, 2\,, \ldots\,, N{-}n{+}1\}$, which results in the extended triangular matrices

$$\begin{aligned} \mathcal{T}_k^{(2)} &\triangleq \left(\mathcal{U}_k^{(2)}\right)^\dagger \mathcal{T}_k^{(2)(1)} \mathcal{V}^{(2)} \\ &= \left(\begin{array}{c|c|c|c|c} T_k^{(2)} & * & \cdots & * & * \\ \hline 0 & T_k^{(2)} & \cdots & * & * \\ \hline \vdots & \vdots & \vdots & \ddots & \vdots \\ \hline 0 & 0 & \cdots & T_k^{(2)} & * \\ \hline 0 & 0 & \cdots & 0 & T_k^{(2)} \end{array} \right) , k = 1\,, 2\,, \end{aligned}$$

where,

$$T_1^{(2)} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix} ,$$

$$T_2^{(2)} = \begin{pmatrix} 1 & * & \cdots & * & * \\ 0 & 1 & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & * \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix} .$$

Thus, by defining

$$\mathcal{V} \triangleq \mathcal{V}^{(1)} \mathcal{O} \mathcal{V}^{(2)}$$

$$(\mathcal{U}_1)^\dagger \triangleq \left(\mathcal{V}^{(2)}\right)^\dagger (\mathcal{O})^\dagger \left(\mathcal{U}_1^{(1)}\right)^\dagger$$

$$(\mathcal{U}_2)^\dagger \triangleq \left(\mathcal{U}_2^{(2)}\right)^\dagger (\mathcal{O})^\dagger \left(\mathcal{U}_2^{(1)}\right)^\dagger$$

we arrive at the desired result. ■

APPENDIX I
PROOF OF THEOREM 4 FOR GENERAL n, N, K

The proof for the case of K users, follows the same principles of the special cases presented in Section VII-B and Appendices F, G, H. The proof is composed of K steps, each of which consists of 2 stages (except for the first step): a reordering stage and a 1-GMD stage.

Denote by $\{\mathcal{A}_k\}$ the extended matrices corresponding to N channel uses.

<u>Step 1:</u>
Perform 1-GMD on the first matrix matrix A_1 (corresponding to to a single channel use): $\left(U_1^{(1)}\right)^\dagger A_1 V^{(1)}$. Apply this

---

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

decomposition to each block separately, on the first extended matrix $\mathcal{A}_1$, using:

$$\left(\mathcal{U}_1^{(1)}\right)^\dagger \triangleq$$
$$I_{nN}\left[\left(U_1^{(1)}\right)^\dagger ; \lfloor 1:\, n\rceil \, \lfloor n+1:\, 2n\rceil \ldots \lfloor (N-1)n+1:\, Nn\rceil\right],$$

$$\mathcal{V}^{(1)} \triangleq$$
$$I_{nN}\left[V^{(1)} ; \lfloor 1:\, n\rceil \, \lfloor n+1:\, 2n\rceil \ldots \lfloor (N-1)n+1:\, Nn\rceil\right].$$

Note that the same matrix $\mathcal{V}^{(1)}$ has to be applied to all matrices (since the encoder is shared by all users). We decompose the resulting matrices (after multiplying them by $\mathcal{V}^{(1)}$) according to the QR decomposition, resulting in unitary matrices $\left(\mathcal{U}_k^{(1)}\right)^\dagger$. The resulting extended triangular matrices are denoted by $\mathcal{T}_k^{(1)} \triangleq \left(\mathcal{U}_k^{(1)}\right)^\dagger \mathcal{A}_k \mathcal{V}^{(1)}$.

Step $2 \leq l \leq K$:

Stage 1: Reordering

We perform the ordering stage using the following ordering matrix, for all $q_1 \in \{1\,, 2\,, \ldots\,, N - n^{K-1} + n^{K-l}\}$ and $q_2 \in \{1\,, 2\,, \ldots\,, n\}$:

$$\mathcal{O}^l \triangleq \mathcal{J}_{nN - n^{(K-l+1)}\left(n^{(l-1)}-1\right)}^{\left[\left\{\{n+(q_1-1)n+(q_2-1)\Delta\}_{q_2}\right\}_{q_1}\right]},$$

where $\Delta = n^{K-l+1} - 1$. Note that the range of q_2 is equal to the dimension n of each block, whereas the range of q_2 is determined by the number of blocks, which depends on l.

Thus, at the end of the first stage, we are left with $\mathcal{T}_k^{(l)(1)} = \left(\mathcal{O}^l\right)^\dagger \mathcal{T}_k^{(l-1)} \mathcal{O}^l$. Note that in each step the size of $\mathcal{T}_k^{(l)(1)}$ is decreasing.

Stage 2: 1-GMD

Perform 1-GMD on the matrix $\mathcal{T}_l^{(l)(1)} \lfloor 1:\, n\rceil$ using:

$$\left(U_l^{(l)}\right)^\dagger \left(\mathcal{T}_l^{(l)(1)} \lfloor 1:\, n\rceil\right) V^{(l)}.$$

Then, apply this decomposition to each of the extended matrices, using:

$$\left(\mathcal{U}_l^{(l)}\right)^\dagger \triangleq I_{nN-n^{(K-l+2)}}\left[\left(U_l^{(l)}\right)^\dagger ; \bigcup_q \lfloor 1+n(q-1),\, nq\rceil\right]$$
$$\mathcal{V}^{(l)} \triangleq I_{nN-n^{(K-l+2)}}\left[V^{(l)} ; \bigcup_q \lfloor 1+n(q-1),\, nq\rceil\right],$$

for all $q \in \{1\,, 2\,, \ldots\,, (N - n^{K-1} + n^{K-l})\}$.

Note that the submatrices of $\mathcal{T}_k^{(l)(1)}$ ($k = 1\,, \ldots\,, l-1$) in the same indices are all equal I_n; by Property 1, multiplying them by $V^{(l)}$ on the right and $\left(V^{(l)}\right)^\dagger$ on the left, leave them unchanged.

The same matrix $\mathcal{V}^{(l)}$ has to be applied to all matrices (since the encoder is shared by all users). We decompose the resulting matrices (after multiplying them by $\mathcal{V}^{(l)}$) according to the QR decomposition, resulting in unitary matrices $\left(\mathcal{U}_k^{(l)}\right)^\dagger$. The resulting extended triangular matrices will be denoted as $\mathcal{T}_k^{(l)} \triangleq \left(\mathcal{U}_k^{(l)}\right)^\dagger \mathcal{T}_k^{(l)(1)} \mathcal{V}^{(l)}$.

Step K:

After performing the last step (step $l = K$) we attain K matrices $\mathcal{T}_k^{(K)}$ which all have 1s on theirs diagonals.

Thus, by defining

$$\left(\mathcal{U}_1\right)^\dagger \triangleq \left(\mathcal{V}^{(K)}\right)^\dagger \left(\mathcal{O}^K\right)^\dagger \cdots \left(\mathcal{V}^{(2)}\right)^\dagger \left(\mathcal{O}^2\right)^\dagger \left(\mathcal{U}_1^{(1)}\right)^\dagger$$
$$\left(\mathcal{U}_k\right)^\dagger \triangleq \left(\mathcal{V}^{(K)}\right)^\dagger \left(\mathcal{O}^K\right)^\dagger \cdots \left(\mathcal{V}^{(k+1)}\right)^\dagger \left(\mathcal{O}^{k+1}\right)^\dagger \cdot$$
$$\cdot \left(\mathcal{U}_k^{(k)}\right)^\dagger \left(\mathcal{O}^k\right)^\dagger \cdots \left(\mathcal{U}_k^{(1)}\right)^\dagger$$
$$\left(\mathcal{U}_K\right)^\dagger \triangleq \left(\mathcal{U}_K^{(K)}\right)^\dagger \left(\mathcal{O}^K\right)^\dagger \cdots \left(\mathcal{U}_K^{(1)}\right)^\dagger$$
$$\mathcal{V} \triangleq \mathcal{V}^{(1)} \mathcal{O}^2 \mathcal{V}^{(2)} \ldots \mathcal{O}^K \mathcal{V}^{(K)},$$

we arrive at the desired result. ■

APPENDIX J
PROOF OF THEOREM 5

We can assume without loss of generality that the matrix V is of the following form:

$$V = \begin{pmatrix} x_1 + ix_2 & y_1 - iy_2 \\ y_1 + iy_2 & -x_1 + ix_2 \end{pmatrix},$$

where x_1, x_2, y_1, y_2 are real numbers satisfying

$$x_1^2 + x_2^2 + y_1^2 + y_2^2 = 1\,.$$

Denote the first column of V by $\boldsymbol{v}_1$ and the second column by $\boldsymbol{v}_2$. Then, there exist unitary matrices U_1, U_2 such that

$$(U_1)^\dagger A_1 V = \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}$$

and

$$(U_2)^\dagger A_2 V = \begin{pmatrix} 1 & 0 \\ * & 1 \end{pmatrix},$$

if and only if the following two vectors have an Euclidean norm of 1:

$$A_1 \boldsymbol{v}_1 = A_1 \begin{pmatrix} x_1 + ix_2 \\ y_1 + iy_2 \end{pmatrix}$$
$$A_2 \boldsymbol{v}_2 = A_2 \begin{pmatrix} y_1 - iy_2 \\ -x_1 + ix_2 \end{pmatrix},$$

or equivalently,

$$\left\|\boldsymbol{v}_1\right\|^2 = 1 \tag{77a}$$
$$\left\|\boldsymbol{v}_2\right\|^2 = 1 \tag{77b}$$
$$\boldsymbol{v}_1^\dagger \left(A_1^\dagger A_1 - I\right) \boldsymbol{v}_1 = 0 \tag{77c}$$
$$\boldsymbol{v}_2^\dagger \left(A_2^\dagger A_2 - I\right) \boldsymbol{v}_2 = 0\,. \tag{77d}$$

By definition, $\left\|\boldsymbol{v}_1\right\|^2 = \left\|\boldsymbol{v}_2\right\|^2$, and also that for any Hermitian 2×2 matrix S:

$$\boldsymbol{v}_2^\dagger S \boldsymbol{v}_2 = \boldsymbol{v}_1^\dagger \operatorname{adj}(S) \boldsymbol{v}_1\,.$$

Thus, (77) is equivalent to

$$\begin{aligned} \|\boldsymbol{v}_1\|^2 &= 1 \\ \boldsymbol{v}_1^\dagger S_1 \boldsymbol{v}_1 &= 0 \\ \boldsymbol{v}_1^\dagger S_2 \boldsymbol{v}_1 &= 0 \,, \end{aligned}$$

where

$$\begin{aligned} S_1 &\triangleq A_1^\dagger A_1 - I \\ S_2 &\triangleq \operatorname{adj}\left(A_2^\dagger A_2 - I\right) \,. \end{aligned}$$

Since $\det(A_1) = \det(A_2) = 1$, we have

$$\begin{aligned} \det(S_1) &\leq 0 \\ \det(S_2) &\leq 0 \,. \end{aligned}$$

Thus, from Lemma 2 it follows that a solution exists if and only if

$$\det\left(S_1 \operatorname{adj}(S_2) - S_2 \operatorname{adj}(S_1)\right) \geq 0 \,. \tag{78}$$

Note that for any 2×2 matrix A,

$$\operatorname{adj}(\operatorname{adj}(A)) = A \,.$$

Hence, the left hand side of condition (78) can be written as

$$\begin{aligned} &\det\left(S_1 \operatorname{adj}(S_2) - S_2 \operatorname{adj}(S_1)\right) \\ &= \det\Big((A_1^\dagger A_1 - I)(A_2^\dagger A_2 - I) \\ &\qquad - \operatorname{adj}(A_2^\dagger A_2 - I)\operatorname{adj}(A_1^\dagger A_1 - I)\Big) \\ &= F_2\left(A_1^\dagger A_1 - I, A_2^\dagger A_2 - I\right) \,, \end{aligned}$$

which completes the proof of the theorem. ■

APPENDIX K
PROOF OF LEMMA 4

Denote the vector consisting of $\{r_m\}$ with their multiplicities, ordered non-increasingly, by $\boldsymbol{r}$ and the vector whose entries are the singular values of A, $\{\sigma_j\}$, ordered non-increasingly, by $\boldsymbol{\sigma}$. According to the GTD [30], the decomposition (46) is possible if and only if Weyl's condition [27], [28],

$$\boldsymbol{\sigma} \succeq \boldsymbol{r} \,, \tag{79}$$

holds true. Namely, n conditions need to be evaluated. We shall show next that when at least some of the absolute values of the desired diagonal R_{jj} are of multiplicity greater than 1, such that there are $M < n$ distinct such (absolute) values, only M of these conditions, (47)-(48), need to be evaluated. The necessity of (47)-(48) is apparent since they constitute the $n_1, n_1 + n_2, ..., n$ conditions in (79).

We shall prove the sufficiency of these conditions by induction.

Basis: We shall show first that the n_1 condition in (79) is sufficient for all the first n_1 conditions in (79) to hold: Assume that

$$r_1^{n_1} \leq \prod_{j=1}^{n_1} \sigma_j \,,$$

holds true. This condition can be rewritten as

$$r_1 \leq \sqrt[n_1]{\prod_{j=1}^{n_1} \sigma_j} \,.$$

Using the fact that the geometric-mean of a set of size n_1 cannot be larger than the geometric-mean of its largest q values ($q = 1, ..., n_1 - 1$), we have

$$r_1 \leq \sqrt[n_1]{\prod_{j=1}^{n_1} \sigma_j} \leq \sqrt[q]{\prod_{j=1}^{q} \sigma_j} \,, \quad q = 1, ..., n_1 - 1 \,,$$

or equivalently,

$$r_1^q \leq \prod_{j=1}^{q} \sigma_j \,, \quad q = 1, ..., n_1 - 1 \,,$$

which are exactly equivalent to the first n_1 conditions of (79).

Inductive step: Assume that the conditions (47)-(48) guarantee that the first $\sum_{m=1}^{k-1} n_m$ conditions in (79) are satisfied. We shall prove that all the first $\sum_{m=1}^{k} n_m$ conditions in (79) hold true. We shall now show that if the $\sum_{m=1}^{k} n_m$ condition in (79) holds true (which is the k-th condition in (47)), then so do the $n_k - 1$ conditions that precede it. Let q be some integer between 1 and M, and assume that

$$\prod_{m=1}^{q} r_m^{n_m} \leq \prod_{j=1}^{\sum_{m=1}^{q} n_m} \sigma_j \,,$$

which can be equivalently written as

$$r_q^{n_q} \leq \gamma \prod_{j=\left(\sum_{m=1}^{q-1} n_m\right)+1}^{\sum_{m=1}^{q} n_m} \sigma_j \,, \tag{80}$$

where γ is defined as

$$\gamma \triangleq \prod_{m=1}^{q-1} r_m^{-n_m} \prod_{j=1}^{\sum_{m=1}^{q-1} n_m} \sigma_j$$

and is equal or larger than 1.

Let l be some integer between 1 and $n_q - 1$, and assume, to contradict, that

$$\left(\prod_{m=1}^{q-1} r_m^{n_m}\right) r_q^l > \prod_{j=1}^{\left(\sum_{m=1}^{q-1} n_m\right)+l} \sigma_j \,,$$

or equivalently,

$$r_q^l > \gamma \prod_{j=\left(\sum_{m=1}^{q-1} n_m\right)+1}^{\left(\sum_{m=1}^{q-1} n_m\right)+l} \sigma_j \,. \tag{81}$$

Dividing (80) by (81) gives rise to

$$r_q^{n_q - l} < \prod_{j=\left(\sum_{m=1}^{q-1} n_m\right)+l+1}^{\sum_{m=1}^{q} n_m} \sigma_j \,,$$

which can be written as

$$r_q < \sqrt[n_q - l]{\prod_{j=\left(\sum_{m=1}^{q-1} n_m\right)+l+1}^{\sum_{m=1}^{q} n_m} \sigma_j} \, .$$

Using the fact that the geometric-mean of the smallest $n_q - l$ values of a set of positive numbers is equal or smaller than the geometric mean of its l largest values, and the fact that $\gamma \geq 1$, we have

$$r_q < \sqrt[n_q - l]{\prod_{j=\left(\sum_{m=1}^{q-1} n_m\right)+l+1}^{\sum_{m=1}^{q} n_m} \sigma_j} \leq \sqrt[l]{\gamma \prod_{j=\left(\sum_{m=1}^{q-1} n_m\right)+1}^{\left(\sum_{m=1}^{q-1} n_m\right)+l} \sigma_j} \, .$$

i.e.,

$$r_q^l < \gamma \prod_{j=\left(\sum_{m=1}^{q-1} n_m\right)+1}^{\left(\sum_{m=1}^{q-1} n_m\right)+l} \sigma_j \, ,$$

in contradiction to (81). ■

REFERENCES

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

[2] J. M. Cioffi and G. D. Forney Jr., "Generalized decision-feedback equalization for packet transmission with ISI and Gaussian noise," in *Comm., Comp., Cont. and Sig. Proc.* US: Springer, 1997, pp. 79–127.

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

[4] J.-K. Zhang, A. Kavčić, and K. M. Wong, "Equal-diagonal QR decomposition and its application to precoder design for successive-cancellation detection," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 154–172, Jan. 2005.

[5] P. Kosowski and A. Smoktunowicz, "On constructing unit triangular matrices with prescribed singular values," *Computing*, vol. 64, no. 3, pp. 279–285, 2000.

[6] S. Lin, W. W. Ho, and Y.-C. Liang, "Block diagonal geometric mean decomposition (BD-GMD) for MIMO broadcast channels," *IEEE Trans. Wireless Comm.*, vol. 7, no. 7, pp. 2778–2789, July 2008.

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

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

[9] J. Wolfowitz, "Simultaneous channels," *Arch. Rational Mech. Anal.*, vol. 4, no. 1, pp. 371–386, Jan. 1959.

[10] 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.

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

[12] A. Hitron, A. Khina, and U. Erez, "Transmission over arbitrarily permuted parallel Gaussian channels," in *Proc. IEEE Int. Symp. on Info. Theory (ISIT)*, Cambridge, MA, USA, July 2012, pp. 2661–2665.

[13] A. Khina, O. Ordentlich, U. Erez, Y. Kochman, and G. W. Wornell, "Decode-and-forward for the Gaussian relay channel via standard AWGN coding and decoding," in *Proc. IEEE Info. Theory Workshop (ITW)*, Lausanne, Switzerland, Sep. 2012, pp. 457–461.

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

[15] ——, "Improved rates and coding for the MIMO two-way relay channel," in *Proc. IEEE Int. Symp. Info. Theory and Its Apps. (ISITA)*, Melbourne, Vic., Australia, Oct. 2014, pp. 658–662.

[16] ——, "Joint unitary triangularization for MIMO networks," *IEEE Trans. Sig. Proc.*, vol. 60, no. 1, pp. 326–336, Jan. 2012.

[17] H. Yao and G. W. Wornell, "Achieving the full MIMO diversity–multiplexing frontier with rotation-based space-time codes," in *Proc. Annual Allerton Conf. on Comm., Control, and Comput.*, Monticello, IL, USA, Oct. 2003, pp. 400–409.

[18] J.-C. Belfiore, G. Rekaya, and E. Viterbo, "The golden code: A 2 x 2 full-rate space-time code with nonvanishing determinants," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1432–1436, Apr. 2005.

[19] F. Oggier, G. Rekaya, J.-C. Belfiore, and E. Viterbo, "Perfect space–time block codes," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3885–3902, Sep 2006.

[20] P. Elia, B. A. Sethuraman, and P. V. Kumar, "Perfect space–time codes for any number of antennas," *IEEE Trans. Inf. Theory*, vol. 53, no. 11, pp. 3853–3868, Nov. 2007.

[21] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Selected Areas in Comm.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[22] V. Tarokh, G. Jafarkhani, and A. R. Calderbank, "Space–time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1456–1467, July 1999.

[23] 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*, Dec. 2003, pp. 2010–2014.

[24] M. J. Lopez, "Multiplexing, scheduling, and multicasting strategies for antenna arrays in wireless networks," Ph.D. dissertation, Massachusetts Institute of Technology, 2002.

[25] S. Tavildar and P. Viswanath, "Approximately universal codes over slow-fading channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3233–3258, July 2006.

[26] D. P. Palomar and Y. Jiang, "MIMO transceiver design via majorization theory," *Found. Trends Comm. Info. Theory*, vol. 3, no. 4, pp. 331–551, Nov. 2006.

[27] H. Weyl, "Inequalities between two kinds of eigenvalues of a linear transformation," in *Proc. Nat. Acad. Sci. USA, 35*, no. 7, May 1949, pp. 408–411.

[28] A. Horn, "On the eigenvalues of a matrix with prescribed singular values," in *Proc. Amer. Math. Soc.*, vol. 5, no. 1, 1954, pp. 4–7.

[29] J.-K. Zhang and K. M. Wong, "Fast QRS decomposition of matrix and its applications to numerical optimization," Dpt. of Elect. and Comp. Engineering, McMaster University, Tech. Rep. [Online]. Available: http://www.ece.mcmaster.ca/~jkzhang/papers/sam_qrs.pdf

[30] Y. Jiang, W. Hager, and J. Li, "The generalized triangular decompostion," *Math. of Comput.*, vol. 77, no. 262, pp. 1037–1056, 2008.

[31] G. H. Golub and C. F. Van Loan, *Matrix Computations, 3rd ed.* Baltimore: Johns Hopkins University Press, 1996.

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

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

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

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

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

[37] A. Hitron, "Linear precoding and coding for Gaussian channels," Master's thesis, Tel-Aviv University, Sep. 2012. [Online]. Available: http://www.eng.tau.ac.il/~uri/theses/hitron_msc.pdf

[38] F. Willems and A. Gorokhov, "Signaling over arbitrarily permuted parallel channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 1374–1382, Mar. 2008.

[39] E. Hof, I. Sason, and S. Shamai, "Polar coding for reliable communications over parallel channels," in *Proc. IEEE Info. Theory Workshop (ITW)*, Dublin, Ireland, Aug/Sep 2010.

[40] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. Inf. Theory*, vol. 48, no. 7, pp. 1804–1824, July 2002.

[41] X.-B. Liang, "Orthogonal designs with maximal rates," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2468–2503, Oct. 2003.

[42] I. Livni, A. Hitron, A. Khina, and U. Erez, “K-matrix geometric mean decomposition (K-GMD): Matlab script,” July 2012. [Online]. Available: http://www.eng.tau.ac.il/~uri/scripts/ggmd.m
[43] ——, “K-matrix geometric mean decomposition (K-GMD): Python script,” July 2012. [Online]. Available: http://www.eng.tau.ac.il/~uri/scripts/ggmd.py
[44] I. Livni, “Perfect and nearly perfect multi-user MIMO transmission via matrix decompositions,” Master’s thesis, Tel-Aviv University, Oct. 2013. [Online]. Available: www.eng.tau.ac.il/~uri/theses/livni_msc.pdf
[45] C. F. Van Loan, “Generalizing the singular value decomposition,” *SIAM J. Numer.*, vol. 13, no. 1, pp. 76–83, 1976.

Anatoly Khina was born in Moscow, USSR, on September 10, 1984. He received the B.Sc. and M.Sc. degrees in electrical engineering (both *summa cum laude*) from Tel Aviv University in 2006 and 2010, respectively, where he is currently working towards completing his Ph.D. degree. His research interests include information theory, signal processing, digital communications and matrix analysis.

In parallel to his studies, Anatoly has been working as an engineer in various algorithms, software and hardware R&D positions. He is a recipient of the Rothschild fellowship, Clore scholarship, Trotsky Award, Weinstein Prize for research in signal processing, and the first prize for outstanding research work of the Advanced Communication Center, Israel.

Idan Livni was born in Tel-Aviv, Israel, on July 30, 1984. He received the B.Sc. and M.Sc. degrees (both *cum laude*) in electrical engineering from Tel Aviv University in 2006 and 2013, respectively. His research interests are in digital communications, signal processing and information theory.

Ayal Hitron received a B.Sc. (*summa cum laude*) in electrical engineering and a B.A. in Physics (*summa cum laude*), both from the Technion — Israel Institute of Technology in 2003, and an M.Sc. in electrical engineering (*summa cum laude*), from Tel Aviv University in 2012.

Ayal is the recipient of a bronze medal in the International Physics Olympiad (IPhO), the Knesset award for outstanding undergraduate student achievements, and the Weinstein Prize for research in signal processing.

Uri Erez (M’09) was born in Tel-Aviv, Israel, on October 27, 1971. He received the B.Sc. degree in mathematics and physics and the M.Sc. and Ph.D. degrees in electrical engineering from Tel-Aviv University in 1996, 1999, and 2003, respectively. During 2003–2004, he was a Postdoctoral Associate at the Signals, Information and Algorithms Laboratory at the Massachusetts Institute of Technology (MIT), Cambridge. Since 2005, he has been with the Department of Electrical Engineering–Systems at Tel-Aviv University. His research interests are in the general areas of information theory and digital communications. He served in the years 2009–2011 as Associate Editor for Coding Techniques for the IEEE TRANSACTIONS ON INFORMATION THEORY.