

Transmission over Arbitrarily Permuted Parallel Gaussian Channels

Ayal Hitron, Anatoly Khina, Uri Erez*
 {ayal,anatolyk,uri}@eng.tau.ac.il

EE-Systems Dept., Tel Aviv University, Tel Aviv, Israel

Abstract—We address the problem of communication over arbitrarily permuted parallel Gaussian channels, where the permutation is known only to the receiver. We present a practical transmission scheme, that allows to transmit over this channel using off-the-shelf codes, in conjunction with linear processing and successive interference cancellation. The scheme is based on the approach of joint matrix triangularization. Explicit precoding matrices are derived for up to six parallel channels.

I. INTRODUCTION

The problem of transmitting information over arbitrarily permuted parallel channels was studied by Willems and Gorokhov [1] and by Hof et al. [2]. In this point-to-point scenario, the transmitter is connected to the receiver via K parallel memoryless channels (see Figure 1 for the case of $K = 3$), sharing the same input alphabet, the transition matrices of which are known but not their order. Namely, at each time instance, the transmitter generates K input symbols to be sent over the K parallel channels, and these symbols are then *permuted* by a one-to-one-mapping (permutation) π from $\{1, \dots, K\}$ onto itself.

The permutation π is arbitrary, yet constant,¹ 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; thus, the histogram (when the number of channels is large) is known, but the permutation is not. For details see [1, Section VII].

When the permutation is constant during the whole transmission period (as considered in [1], [2]), this setting falls under the framework of compound channels,² the capacity of which is well known (see, e.g., [3]).

Willems and Gorokhov [1] constructed an MDS-like scheme for treating this case, where the receiver uses joint typicality decoding. In [2] a scheme based on polar codes along with MDS codes was proposed for the case where the channels are binary-input and output-symmetric (BIOS). Another approach to tackle this problem in the Gaussian case is using approximately universal codes [4], which assumes high signal-to-noise ratio (SNR) regime.

* This work was supported in part by the Israel Science Foundation under Grant No. 1557/12.

¹The case of time-varying permutation will be addressed in Section VI.

²In the classical compound channel setting the receiver is not aware of the compound parameter as well. However, the capacity in both cases is the same; see [3, Sec. 8].

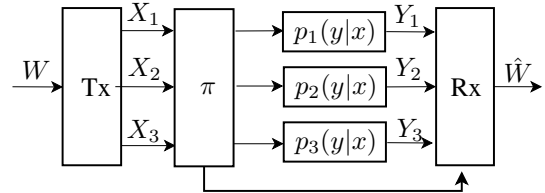


Fig. 1. Permuted Parallel Channels

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

$$y_i = \alpha_i x_i + z_i, \quad i = 1, 2, \dots, K, \quad (1)$$

where x_i is the input to the i -th channel which is subject to a power constraint³

$$\mathbb{E}(|x_i|^2) \leq 1,$$

y_i is the output of the i -th channel, and $\{z_i\}$ are i.i.d. circularly-symmetric Gaussian variables with unit variance, independent of $\{x_i\}$. The gains $\{\alpha_i\}$ are known to the receiver, whereas the transmitter knows the gains up to an unknown permutation. Namely, the transmitter knows the gains but does not know their *order*. The scheme is based on codes designed for a point-to-point additive white Gaussian noise (AWGN) channel with known (fixed) SNR, in conjunction with successive interference cancellation.

Since the capacity-achieving input distribution for all the channels is identical (circularly-symmetric complex normal with unit variance), the capacity of the compound channel is:

$$C = \sum_{i=1}^K \log(1 + |\alpha_i|^2). \quad (2)$$

An interesting special case of the permuted parallel channels problem is that of parallel Gaussian erasure channels, where a constant number of channels is “erased” at each time instance. In the notation of the Gaussian permutation channel, this corresponds to the case of the coefficients α_i (1) equaling either the same constant channel gain α or 0 (in case of an erasure). For this special case, capacity in the compound setting can be efficiently achieved using MDS codes concatenated with codes which are good for scalar AWGN channels. However, except

³Alternatively, the individual power constraints can be replaced by a common power constraint. However, in our case of interest, both cases reduce to the same result.

for this extreme case, practical capacity-achieving schemes are not known.

In this paper, we develop a space–time modulation technique that, in conjunction with successive interference cancellation (SIC), gives rise to effective (scalar) parallel AWGN channels with the same gains (*and order*) for all possible permutations. Thus, by using off-the-shelf (fixed-rate) AWGN codes over the effective channels, capacity is achieved. We present exact solutions for $K \leq 6$ parallel channels, and discuss the generalization for larger K .

II. BACKGROUND: TWO-USER MULTICAST VIA JOINT MATRIX TRIANGULARIZATION

In this section we review the scheme proposed in [5] for multicasting a common message over MIMO Gaussian channels. This scheme uses joint matrix triangularization, along with SIC, to transform the channel into parallel single-input single-output (SISO) AWGN channels, with rates that are *known at the transmitter*.

The channel model in [5] is the two-user common-message Gaussian broadcast channel:

$$\mathbf{y}_k = H_k \mathbf{x} + \mathbf{z}_k, \quad k = 1, 2,$$

where \mathbf{x} is the channel input of dimensions $n_t \times 1$ subject to an average power constraint P , \mathbf{y}_k is the channel output vector of receiver k ($k = 1, 2$) of dimensions $n_r^{(k)} \times 1$; H_k is the channel matrix to receiver k of dimensions $n_r^{(k)} \times n_t$ and \mathbf{z}_k is an additive circularly-symmetric Gaussian noise vector of dimensions $n_r^{(k)} \times 1$.

The aim of the transmitter is to multicast the same (common) message to all the receivers. The capacity of this scenario equals the (worst-case) capacity of the compound channel (see, e.g., [3]), with the compound parameter being the channel matrix index:

$$C = \max_{C_{\mathbf{x}}} \min_{k=1,2} \log \det \left(I + H_k C_{\mathbf{x}} H_k^\dagger \right), \quad (3)$$

where maximization is carried over all admissible channel input covariance matrices $C_{\mathbf{x}} \succeq 0$, subject to the power constraint.

The transmission scheme of [5] is based on applying a unitary triangularization to two *augmented matrices*:

$$G_k \triangleq \begin{pmatrix} H_k \sqrt{C_{\mathbf{x}}} \\ I \end{pmatrix}, \quad k = 1, 2 \quad (4)$$

$$T_k = U_k^\dagger G_k V,$$

where V is an $n_t \times n_t$ precoding unitary matrix applied at the transmitter (and thus, cannot depend on k), U_k^\dagger is an $n_t \times (n_r^{(k)} + n_t)$ matrix with orthonormal rows which is known at the receiver (and may differ between the receivers), and T_k is an $n_t \times n_t$ upper-triangular matrix.

The transmitter sends the signal:

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

where the vector $\tilde{\mathbf{x}}$ is formed by taking one symbol from each of n_t parallel codebooks, of equal powers $1/n_t$.

Receiver k applies the matrix \tilde{U}_k^\dagger to the channel output, where \tilde{U}_k^\dagger is an $n_t \times n_r^{(k)}$ matrix consisting of the left $n_r^{(k)}$ columns of U_k^\dagger ,

$$\tilde{\mathbf{y}}_k = \tilde{U}_k^\dagger \mathbf{y}_k.$$

This results in the equivalent channel:

$$\begin{aligned} \tilde{\mathbf{y}}_k &= \tilde{U}_k^\dagger (H_k \sqrt{C_{\mathbf{x}}} V \tilde{\mathbf{x}} + \mathbf{z}) \\ &= \tilde{U}_k^\dagger H_k \sqrt{C_{\mathbf{x}}} V \tilde{\mathbf{x}} + \tilde{U}_k^\dagger \mathbf{z} \\ &\triangleq \tilde{T}_k \tilde{\mathbf{x}} + \tilde{\mathbf{z}}. \end{aligned}$$

Finally, SIC is performed, i.e., the codebooks are decoded from last ($j = n_t$) to first ($j = 1$), where each codebook is recovered from $y'_j \triangleq \tilde{y}_j - \sum_{l=j+1}^{n_t} \tilde{T}_{j,l} \hat{x}_l$, where \hat{x}_l is the decoded symbol from the l -th codebook.

As is shown in [5], this scheme achieves the rate:

$$R = 2 \sum_{i=1}^{n_t} \min_{k=1,2} \{ \log |[T_k]_{ii}| \},$$

where $[A]_{ij}$ denotes the (i, j) -th element of the matrix A . If all the matrices T_k have *the same* diagonal values, then:

$$R = 2 \sum_{i=1}^{n_t} \log |[T_k]_{ii}| = \log \det (I + H_k^\dagger C_{\mathbf{x}} H_k),$$

namely, the scheme achieves the capacity (3).

Finally, it is shown in [5] that there *always* exists such a decomposition, which is coined *joint equi-diagonal triangularization* (JET). Moreover, an explicit construction of the matrices V, U_1, U_2 is given. As a result, this scheme can always be used in order to achieve the capacity of the two-user common message Gaussian MIMO broadcast channel (3). However, an exact JET is known to exist only for $K = 2$ for general matrices, and it is not known how to generalize it to larger values of K .

III. PROBLEM REFORMULATION

In this section we reformulate the problem of arbitrarily permuted parallel Gaussian channels in terms of a common-message MIMO broadcast problem for $K!$ users, where K is the number of parallel channels.

The K parallel channels (1) can be regarded as a single MIMO channel:

$$\mathbf{y} = H \mathbf{x} + \mathbf{z}$$

where \mathbf{x} is the channel input vector of length K , and \mathbf{z} is a circularly-symmetric white Gaussian random vector of length K and identity covariance matrix.

The channel matrix H is a $K \times K$ diagonal matrix, which is known at the receiver:

$$H = \begin{pmatrix} \alpha_1 & 0 & \cdots & 0 \\ 0 & \alpha_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_K \end{pmatrix}. \quad (5)$$

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

The latter is, in turn, equivalent to multicasting the same (common) message to $K!$ receivers simultaneously, where the channel matrices of the various users are all of the form:

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

and $\pi_k \in S_K$ is a permutation which is different for each user.

The transmission scheme is obtained, similarly to the two-user scheme (4), by applying a joint unitary triangularization to the augmented matrices, with $C_{\mathbf{x}} = I$:

$$G_k \triangleq \begin{pmatrix} H_k \\ I \end{pmatrix}, \quad T_k = U_k^\dagger G_k V, \quad (6)$$

where V is a precoding unitary matrix which is known at the transmitter (and thus, cannot depend on k), U_k is a matrix with orthogonal columns which is applied by receiver k , and T_k is an upper-triangular matrix.

As discussed in Section II, this scheme achieves the rate:

$$R = 2 \sum_{i=1}^K \min_i \{ \log [T_k]_{ii} \}.$$

In the case where all the matrices $\{T_k\}$ have the same diagonal values, the following rate is achieved:

$$R = 2 \sum_{i=1}^K \log [T_k]_{ii} = \log \det(I + H_k^\dagger H_k) = \sum_{i=1}^K \log(1 + |\alpha_i|^2),$$

which, according to (2), is equal to the capacity of the channel.

Note that U_k and T_k in (6) constitute the QR decomposition of the matrix $G_k V$. Consequently, the matrix T_k is the Cholesky factor of the matrix $V^\dagger(I + H_k^\dagger H_k)V$:

$$V^\dagger(I + H_k^\dagger H_k)V = T_k^\dagger T_k, \quad k = 1, 2, \dots, K!. \quad (7)$$

To summarize, the proposed scheme allows to achieve the capacity (2) using codes designed for scalar AWGN channels with known SNRs, provided that there exists a unitary matrix V such that the Cholesky decomposition (7) holds for every k , $\{T_k\}$ being upper-triangular with the same diagonal values.

For $K = 2$ this is achievable using the JET, as discussed in Section II. However, as mentioned above, JET is known to exist for only $K = 2$ (for general matrices), and it is not known how to generalize it to larger values of K . In the following section, we show that in the special case of arbitrarily permuted channels (5), the decomposition (7) exists for $K = 2$ and $K = 3$. Later, we review the concept of *space-time triangularization*, introduced in [6], and show that it can be used to obtain a practical scheme which achieves the capacity for $4 \leq K \leq 6$.

IV. SPACE-ONLY TRIANGULARIZATION

A. Two Parallel Channels ($K = 2$)

We now consider the special case of two parallel channels ($K = 2$). The channel can be in one of two “states”:

$$I + H_1^\dagger H_1 = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix},$$

$$I + H_2^\dagger H_2 = \begin{pmatrix} b & 0 \\ 0 & a \end{pmatrix},$$

where $a, b \geq 1$ are known.

The special case $a = 1$ (or $b = 1$) corresponds to a Gaussian erasure, where there occurs exactly one erasure in every two symbols, but the location of the erasure is unknown at the transmitter.

Since there are only two options for the channel matrix H , the capacity in this case can be achieved by using the JET described in Section II.

Specifically, we show next that the decomposition (7) is obtained by choosing the precoding matrix to be the (scaled) Hadamard matrix (which coincides with the 2×2 DFT matrix):

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

In our case, there are two possible channel matrices H :

$$V^\dagger(I + H_1^\dagger H_1)V = \frac{1}{2} \begin{pmatrix} a+b & a-b \\ a-b & a+b \end{pmatrix} \quad (8)$$

$$V^\dagger(I + H_2^\dagger H_2)V = \frac{1}{2} \begin{pmatrix} a+b & b-a \\ b-a & a+b \end{pmatrix}. \quad (9)$$

Cholesky decompositions (7) of the above two matrices yield

$$T_1 = \begin{pmatrix} \sqrt{\frac{a+b}{2}} & \frac{a-b}{\sqrt{2(a+b)}} \\ 0 & \sqrt{\frac{2ab}{a+b}} \end{pmatrix} \quad (10)$$

$$T_2 = \begin{pmatrix} \sqrt{\frac{a+b}{2}} & \frac{b-a}{\sqrt{2(a+b)}} \\ 0 & \sqrt{\frac{2ab}{a+b}} \end{pmatrix}. \quad (11)$$

Alternatively, we can use the fact that the diagonal values of the Cholesky factor of a matrix can be expressed as the ratios between the determinants of its principal minors (see, e.g., [7]). In our case, the determinants of the principal minors of both (8) and (9) are $M_1 = \frac{1}{2}(a+b)$ and $M_2 = ab$. Therefore, the diagonal values of the Cholesky factors are

$$[T_k]_{11} = \sqrt{M_1}, \quad [T_k]_{22} = \sqrt{\frac{M_2}{M_1}},$$

which coincide with (10) and (11), respectively.

Since T_1 and T_2 have the same diagonal values, the same SISO codewords can be used to simultaneously achieve the capacities of both channels H_1 and H_2 .

Note that the precoding matrix V used by the transmitter does not depend on a or b . Nonetheless, the rates of the SISO codebooks need to be known at both transmission ends.

B. Three Parallel Channels ($K = 3$)

We now treat the case of three parallel channels ($K = 3$). In this case, we have:

$$I + H^\dagger H = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix},$$

where $a, b, c \geq 1$ are known, up to an unknown permutation.

In this case, we propose 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 $e \triangleq e^{2\pi i/3}$.

Using the matrix V in (7) yields

$$V^\dagger(I + H^\dagger H)V = \frac{1}{3} \begin{pmatrix} a+b+c & a+be+ce^{-1} & a+be^{-1}+ce \\ a+be^{-1}+ce & a+b+c & a+be+ce^{-1} \\ a+be+ce^{-1} & a+be^{-1}+ce & a+b+c \end{pmatrix}. \quad (12)$$

The diagonal values of the Cholesky factors of (12) are

$$[T_k]_{11} = \sqrt{M_1}, \quad [T_k]_{22} = \sqrt{\frac{M_2}{M_1}}, \quad [T_k]_{33} = \sqrt{\frac{M_3}{M_2}},$$

where M_1, M_2, M_3 are the determinants of the principal minors of (12):

$$\begin{aligned} M_1 &= \frac{1}{3}(a+b+c) \\ M_2 &= \frac{1}{3}(ab+ac+bc) \\ M_3 &= abc. \end{aligned}$$

Since M_1, M_2, M_3 are invariant to the order of (a, b, c) , so are the diagonal values d_1, d_2, d_3 . Thus, the proposed scheme, using the precoding matrix V , achieves the capacity (2) simultaneously for all the possible permutations π_k .

Note that a 3×3 linear transformation over the complex field can be regarded as a 6×6 transformation over the reals. Thus, in the case that the parallel channels (1) are real-valued (rather than complex valued), we can, nonetheless, use the proposed scheme, by treating two consecutive real-valued input symbols as a single complex-valued symbol. In fact, as explained in the next section, this is a special case of the *time-space triangularization* approach.

Unfortunately, the above two special cases do not carry over to the case of $K = 4$: using the 4×4 complex-valued DFT matrix does *not* yield equal diagonal values after Cholesky decomposition of $V^\dagger(I + H_k^\dagger H_k)V$. For this reason, we need to extend the space-time approach beyond the complex field, using *quaternions*, as detailed in the next section.

V. SPACE-TIME TRIANGULARIZATION

In order to obtain a transmission scheme for more than three parallel channels, we utilize *space-time triangularization*, which was proposed in [6] for the general case.

For illustration, assume we have three parallel channels with gains $\alpha_1, \alpha_2, \alpha_3$. The parallel channels can be represented by the 3×3 channel matrix

$$H = \text{diag}(\alpha_1, \alpha_2, \alpha_3) \triangleq \begin{pmatrix} \alpha_1 & 0 & 0 \\ 0 & \alpha_2 & 0 \\ 0 & 0 & \alpha_3 \end{pmatrix}.$$

However, we can also regard it is a 6×6 MIMO channel, where in each channel-use six symbols are being transmitted, two symbols on each one of the parallel channels. Thus, the channel matrix of this 6×6 channel is:

$$\mathcal{H} = \text{diag}(\alpha_1, \alpha_1, \alpha_2, \alpha_2, \alpha_3, \alpha_3).$$

This will be referred to as an *extended channel matrix with two duplications*. Extended matrices with more than two duplications are defined in a similar manner.

Using extended channel matrices, the scheme described in Section II can be employed, where the channel matrices H_k are replaced by their extended versions \mathcal{H}_k . As we shall see in the sequel, this allows to achieve the channel capacity of K parallel channels beyond the case of $K = 3$ channels.

In order to achieve the capacity for $4 \leq K \leq 6$, we will use extended matrices with two duplications, where each pair of two consecutive complex-valued symbols will be regarded as a single *quaternion-valued* symbol.

A *quaternion* $w \in \mathbb{H}$ can be regarded as a collection of four real-valued symbols [8]

$$w = a + bi + cj + dk.$$

The quaternions form a vector space over the real numbers. There also exists an associative product of two quaternions, which is defined by the products of the basis elements $1, i, j, k$:

$$\begin{aligned} 1^2 &= 1, \quad i^2 = j^2 = k^2 = -1, \\ 1i &= i1 = i, \quad 1j = j1 = j, \quad 1k = k1 = k, \\ ij &= -ji = k, \quad jk = -kj = i, \quad ki = -ik = j. \end{aligned}$$

The transformation of multiplying a quaternion on the left by the quaternion w , $W(x) \triangleq wx$, is a linear transformation from \mathbb{H} to \mathbb{H} over the field \mathbb{R} . Using the basis $\{1, i, j, k\}$, this transformation is given by the following matrix:

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

Alternatively, the same quaternion can be regarded as a collection of two complex-valued symbols:

$$w = (a + bi) + j(c - di).$$

Using the basis $\{1, j\}$, the transformation $W(x) = wx$ over the field \mathbb{C} is also linear, and is represented by:

$$W = \begin{pmatrix} a + bi & -c - di \\ c - di & a - bi \end{pmatrix}.$$

A. Four Parallel Channels ($K = 4$)

We now show how quaternion-valued matrices allow the design of a precoding matrix for four permuted parallel channels ($K = 4$). Using extended matrices with two duplications, the channel matrix (over the complex field) is:

$$I + H^\dagger H = \text{diag}(a, a, b, b, c, c, d, d),$$

which is an 8×8 matrix. By combining every two complex-valued symbols into a single quaternion-valued symbol, we arrive at the following 4×4 matrix over the quaternion division ring:

$$I + H^\dagger H = \begin{pmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & d \end{pmatrix}.$$

Consider the following quaternion-valued precoding matrix V :

$$V = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & x & i & iy \\ 1 & z & -1 & -z \\ 1 & y & -i & -ix \end{pmatrix},$$

where

$$\begin{aligned} x &= \frac{1}{3}(-1 - 2i - \sqrt{2}j + \sqrt{2}k) \\ y &= \frac{1}{3}(-1 + 2i - \sqrt{2}j - \sqrt{2}k) \\ z &= \frac{1}{3}(-1 + 2\sqrt{2}j). \end{aligned}$$

Applying the Cholesky decomposition (7) with this V yields

$$T = \begin{pmatrix} d_1 & * & * & * \\ 0 & d_2 & * & * \\ 0 & 0 & d_3 & * \\ 0 & 0 & 0 & d_4 \end{pmatrix},$$

where

$$\begin{aligned} d_1^2 &= \frac{1}{4}(a + b + c + d) \\ d_1^2 d_2^2 &= \frac{1}{6}(ab + ac + ad + bc + bd + cd) \\ d_1^2 d_2^2 d_3^2 &= \frac{1}{4}(abc + abd + acd + bcd) \\ d_1^2 d_2^2 d_3^2 d_4^2 &= abcd \end{aligned}$$

and $*$ represents some value (which may differ from entry to entry). Again, the diagonal values are invariant to reordering of the values (a, b, c, d) , hence this scheme achieves capacity.

B. Five Parallel Channels ($K = 5$)

For five parallel channels, there also exists a capacity-achieving precoding matrix over the quaternions. The capacity (2) is achieved using the following quaternion precoding matrix:

$$V = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & w & q & q^2 & T(w) \\ 1 & x & q^2 & q^4 & T(x) \\ 1 & y & q^3 & q^6 & T(y) \\ 1 & z & q^4 & q^8 & T(z) \end{pmatrix},$$

where $T(w + xi + yj + zk) \triangleq w + xi - yj - zk$ and $q \triangleq e^{\frac{2\pi i}{5}}$. The values w, x, y, z were obtained numerically:

$$\begin{aligned} w &\approx -0.25 - 0.7694i - 0.1904j - 0.5561k \\ x &\approx -0.25 + 0.1816i - 0.7606j + 0.5710k \\ y &\approx -0.25 - 0.1816i + 0.7781j + 0.5469k \\ z &\approx -0.25 + 0.7694i + 0.1729j - 0.5618k. \end{aligned}$$

The diagonal values after Cholesky factorization are

$$\begin{aligned} d_1^2 &= \frac{1}{5}(a + b + c + d + e) \\ d_1^2 d_2^2 &= \frac{1}{10}(ab + ac + ad + \dots) \\ d_1^2 d_2^2 d_3^2 &= \frac{1}{10}(abc + abd + abe + \dots) \\ d_1^2 d_2^2 d_3^2 d_4^2 &= \frac{1}{5}(abcd + abce + \dots) \\ d_1^2 d_2^2 d_3^2 d_4^2 d_5^2 &= abcde, \end{aligned} \quad (13)$$

which are again invariant to permutations of a, b, c, d, e .

C. Six Parallel Channels ($K = 6$)

For $K = 6$, a 6×6 quaternion precoding matrix was found numerically, but we do not have explicit expressions for the entries of this matrix. The matrix, which is given in [9], yields diagonal values of a form similar to (13).

VI. EXTENSIONS

For a larger number of parallel channels, algebras of higher dimensions need to be considered, to achieve a Cholesky decomposition with a form similar to (13). Another approach could be to use nearly-optimal precoding matrices, as described in [10] for general channel matrices. However, this is not practical, even for relatively small number of channels. For example, for $K = 7$ there can be $7! = 5040$ possible matrices $\{H_k\}$ (in case that all the channel gains α_i are different), so that using this scheme for joint triangularization would suggest an enormous amount of duplications, which becomes impractical. Nearly optimal schemes for permuted parallel Gaussian channels (i.e., this special class of matrices) is an interesting avenue for future research.

Finally, we note that the proposed scheme can also be applied to the case of arbitrarily varying permuted Gaussian channels, where the permutation π may vary between channel uses. Suppose that we obtained a joint triangularization (6), where T_k is upper-triangular with diagonal elements not depending on k . Since the receiver multiplies the received input signal \mathbf{y} by U_π^\dagger and performs SIC of the various codewords, the scheme allows the channel H_k (and therefore the unitary matrix U_k) to vary from symbol to symbol. Note, however, that this applies only for $K = 2, 3$, where one sample is processed at a time. In cases where several channel outputs ought to be grouped and processed together, the permutation π needs to be constant during several consecutive channel uses. This requirement is reminiscent of the requirement in Alamouti space-time coding [11]. If, for example, quaternion precoding is being used, then π need to be constant during every two consecutive channel uses.

REFERENCES

- [1] F. Willems and A. Gorokhov, "Signaling over arbitrarily permuted parallel channels," *IEEE Trans. Info. Theory*, vol. 54, pp. 1374–1382, March 2008.
- [2] E. Hof, I. Sason, and S. S. (Shitz), "Polar coding for reliable communications over parallel channels," in *2010 IEEE Information Theory Workshop (ITW'2010)*, Cambridge, MA, USA, Aug. 30–Sep. 3, 2012.
- [3] J. Wolfowitz, "Simultaneous channels," *Arch. Rational Mech. Anal.*, vol. 4, pp. 371–386, 1960.
- [4] S. Tavildar and P. Viswanath, "Approximately universal codes over slow-fading channels," *IEEE Trans. Info. Theory*, vol. 52, pp. 3233–3258, July 2006.
- [5] A. Khina, Y. Kochman, and U. Erez, "Joint unitary triangularization for MIMO networks," *IEEE Trans. Signal Proc.*, vol. 60, pp. 326–336, Jan. 2012.
- [6] A. Khina, A. Hitron, and U. Erez, "Modulation for MIMO networks with several users," in *2011 IEEE International Symposium on Information Theory (ISIT'2011)*, St. Petersburg, Russia, July 31–Aug. 5, 2011, pp. 2452–2456.
- [7] A. S. Householder, *The Theory of Matrices in Numerical Analysis*, 1964, reprinted by Dover, New York, 1975.
- [8] W. R. Hamilton, *On Quaternions, or On a New System of Imaginaries in Algebra*. Philosophical Magazine, 1844, vol. 25, no. 3.
- [9] [Online]. Available: <http://www.eng.tau.ac.il/~uri/V6.txt>
- [10] I. Livni, A. Khina, A. Hitron, and U. Erez, "Space-time MIMO multicasting," in *2012 IEEE International Symposium on Information Theory (ISIT'2012)*, Cambridge, MA, USA, July 2012.
- [11] S. M. Alamouti, "Multidimensional constellations - Part II: Voronoi Constellations," *A simple transmit diversity technique for wireless communications*, vol. 16, pp. 1451–1458, Oct. 1998.