Equal-Diagonal QR Decomposition and its Application to Precoder Design for Successive-Cancellation Detection

Jian-Kang Zhang, Aleksandar Kavčić and Kon Max Wong

***Abstract*—In multiple-input and multiple-output (MIMO) multiuser detection theory, the QR decomposition of the channel matrix H can be used to form the back-cancellation detector. In this paper we propose an optimal QR decomposition, which we call the equal-diagonal QR decomposition, or shortly the QRS decomposition. We apply the decomposition to precoded successive-cancellation detection, where we assume that both the transmitter and the receiver have perfect channel knowledge. We show that, for any channel matrix H, there exists a unitary precoder matrix S, such that $\mathrm{HS} = \mathrm{QR}$, where the non-zero diagonal entries of the upper triangular matrix R in the QR decomposition of HS are all equal to each other. The precoder and the resulting successive-cancellation detector have the following properties. a) The minimum Euclidean distance between two signal points at the channel output is equal to the minimum Euclidean distance between two constellation points at the precoder input up to a multiplicative factor that equals the diagonal entry in the R-factor. b) The superchannel HS naturally exhibits an optimally ordered column permutation, i.e., the optimal detection order for the V-BLAST detector is the natural order. c) The precoder S minimizes the block error probability of the QR successive cancellation detector. d) A lower and an upper bound for the free distance at the channel output is expressible in terms of the diagonal entries of the R-factor in the QR decomposition of a channel matrix. e) The precoder S maximizes the lower bound of the channel's free distance subject to a power constraint. f) For the optimal precoder S, the performance of the QR detector is asymptotically (at large signal-to-noise ratios) equivalent to that of the maximum likelihood detector (MLD) that uses the same precoder. Further, in this paper we consider two multiplexing schemes: time division multiple access (TDMA) and orthogonal frequency division multiplexing (OFDM). We design the optimal precoder for binary phase shift keying (BPSK) with these multiplexing schemes, but outline the procedure to extend the method to non-binary schemes such as PAM, PSK and QAM. Finally, examples are given that illustrate the performance of the precoder and the corresponding successive cancellation detector.**

***Index Terms*— maximum likelihood detection, minimum distance, multiple-input multiple-output (MIMO) systems, orthogonal frequency division multiplexing (OFDM), precoders, QR-decomposition, successive cancellation detection, time-division multiple access (TDMA)**

I. Introduction

The QR decomposition [1], [2] is a commonly used tool in various signal processing applications [3], [4] [5]. The QR decomposition of a matrix $\mathbf{H}$ is a factorization $\mathbf{H} = \mathbf{QR}$, where $\mathbf{Q}$ is a unitary matrix and $\mathbf{R}$ is an upper triangular matrix. In this paper we study the application of the QR decomposition in precoded successive-cancellation detection and we propose an optimal QR decomposition. In multiple-input and multiple-output multiuser detection theory [3], [5], [6], [7], [8], [9] [10], [11], [12], [13], [14], [15] [16], [17], the QR decomposition can be used to form the back-cancellation detector. Though it is easily shown that such a detector is not optimal, the QR decomposition is still very appealing because of its implementation simplicity and numerical stability [2].

In [18], Golden et al. introduced an optimally ordered successive cancellation detector. We show that this detector may be equivalently represented by a permutation matrix $\mathbf{P}$ followed by a QR-decomposition-based detector. Thereby, the algorithm presented by Golden et al. [18] has an equivalent interpretation: it is an efficient algorithm to determine the permutation matrix $\mathbf{P}$, such that the QR decomposition of $\mathbf{HP}$ (that is, $\mathbf{HP} = \mathbf{QR}$) gives rise to an optimal back-cancellation detector[1]. In this paper we extend the work of Golden et al. [18] to a general QR decomposition $\mathbf{HS} = \mathbf{QR}$, where $\mathbf{S}$ is a unitary matrix. We shall show that for a wide class of detection problems, the optimal matrix $\mathbf{S}$ is one that delivers an upper triangular matrix $\mathbf{R}$ whose diagonal entries are all *equal* to each other. We call this special matrix $\mathbf{R}$ the *equal-diagonal R-factor*, and the resulting decomposition $\mathbf{HS} = \mathbf{QR}$, the *equal-diagonal* QR decomposition, or simply the *QRS-decomposition*.

The paper proceeds to show that the QRS-decomposition has an important role to play in precoded block transmission over dispersive channels under the assumption that both the transmitter and the receiver have perfect channel knowledge. The filterbank precoding framework by Xia [20], and Scaglione, Giannakis and Barbarossa [21] unifies several existing modulation schemes, including orthogonal frequency division multiplexing (OFDM) [22], [23], [24], [25], [26], [27], discrete multitone

Jian-Kang Zhang was with the Division of Engineering and Applied Sciences, Harvard University, Cambridge, MA when this work was performed. He is presently with the Department of Electrical and Computer Engineering, McMaster University, Ontario, Canada. Aleksandar Kavčić is with the Division of Engineering and Applied Sciences, Harvard University, Cambridge, MA. Kon Max Wong is with the Department of Electrical and Computer Engineering, McMaster University, Ontario, Canada.

Partial results of this paper were published in ICASSP 2002. This work was supported in part by the National Science Foundation under Grant # CCR–0118701.

[1] The reader should also see [19] where the link to a fast algorithm involving QR decompositions is explicitly made.

(DMT) [28], [29], [30], [31], [32], time division multiplex access (TDMA) [33] and code division multiplex access (CDMA) [3], [33], [34], both for single user and multiuser communications. Scaglione et al. [35] proposed precoders that maximize the output signal to noise ratio (SNR) or minimize the mean-square error under zero-forcing or fixed transmitted power constraints. The block transmitter that maximizes the information rate was derived in [21] by the same authors. Milanovic et al. [36] studied the design of robust redundant precoding filter banks with zero-forcing equalization for unknown frequency-selective channels. Recently, Ding et al. [37] designed a precoder for TDMA systems that minimizes the average bit error probability when using zero-forcing equalizers.

A universally optimal precoder for detection applications[2] is one that minimizes the detection error probability of the maximum likelihood detector [3], [38], [39], [40]. It is known that at high signal-to-noise ratios (SNRs), the average probability of error over all blocks is dominated by the free distance term [39], [40]. This suggests that maximizing the free distance may be a good precoder design strategy. However, directly maximizing the free distance results in a precoder design and a detection strategy that are too complex to be affordable. In this paper, under some fairly loose assumptions, we derive the upper and lower bounds on the channel's free distance in terms of the diagonal entries of the R-factor in the QR decomposition [41]. Subject to a power constraint, we design the optimal precoder that maximizes the lower-bound of the channel's free distance. The resulting precoder is exactly the $\mathbf{S}$ matrix in the QRS decomposition, $\mathbf{HS} = \mathbf{QR}$.

Curiously, the same optimal precoder matrix can be derived under a different optimization criterion. Namely, if we design the precoder to minimize the block error probability of a successive cancellation detector that employs QR decomposition, the optimal precoder is the S-factor of the QRS-decomposition, $\mathbf{HS} = \mathbf{QR}$. For this reason we call the S-factor the *QR-optimal* precoder.

In this paper, we apply the QRS decomposition to derive the optimal QR-precoder for TDMA and OFDM systems used over channels with intersymbol interference. Examples are given to illustrate the performance of the precoder and the corresponding successive cancellation detector.

The paper is structured as follows. In Section II, we review the QR-decomposition-based successive cancellation detector and the optimally ordered cancellation algorithm of Golden et al. [18]. In Section III, we develop the underlying QR decomposition theory by introducing the QRS decomposition and proving its properties. In Section IV, we introduce the channel models. Section V is devoted to deriving the QR-optimal precoder for the successive-cancellation detector. In Section VI, we show that the optimal precoder can also be derived using a free-distance criterion. Extensive simulation results, comparing the precoder to other precoders available in the literature, are presented in Section VII. Section VIII concludes the paper.

Notation: Matrices are denoted by uppercase boldface characters (e.g., $\mathbf{A}$), while column vectors are denoted by lowercase boldface characters (e.g., $\mathbf{b}$). The (i,j)-th entry of $\mathbf{A}$ is denoted by $A_{i,j}$. The i-th entry of $\mathbf{b}$ is denoted by b_i. The columns of an $M \times N$ matrix $\mathbf{A}$ are denoted by $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_N$. Notation $\mathbf{A}_k$ denotes a matrix consisting of the first k columns of $\mathbf{A}$, i.e., $\mathbf{A}_k = [\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_k]$. By convention, $\mathbf{A}_0 = 1$. The remaining matrix after deleting columns $\mathbf{a}_{k_1}, \mathbf{a}_{k_2}, \cdots, \mathbf{a}_{k_i}$ from $\mathbf{A}$ is denoted by $\overline{\mathbf{A}}_{k_1,k_2\cdots,k_i}$. The j-th diagonal entry of a matrix $\mathbf{A}$ is denoted by $[\mathbf{A}]_j = A_{j,j}$. Notation $\mathbf{A}^\perp$ denotes the orthonormal complement of a matrix $\mathbf{A}$ in $\mathbb{C}^N$. The transpose of $\mathbf{A}$ is denoted by $\mathbf{A}^T$. The Hermitian transpose of $\mathbf{A}$ (i.e., the conjugate and transpose of $\mathbf{A}$) is denoted by $\mathbf{A}^H$. The matrix $\mathbf{A}^+$ stands for the pseudo-inverse of $\mathbf{A}$; i.e., $\mathbf{A}^+ = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$. The notation $\mathcal{P}_{\mathbf{A}} = \mathbf{I} - \mathbf{A}\mathbf{A}^+$ denotes the projection matrix (that projects an arbitrary vector to the null space of $\mathbf{A}^H$).

II. Review of the QR decomposition for successive cancellation detection

We first briefly review the successive cancellation detection algorithm that uses the QR decomposition, then we review the optimally ordered detector developed by Golden et al. [18], and finally we show how to equivalently represent this detector as a precoded QR-decomposition cancellation detector.

Let $\mathbf{x} = [x_1, \cdots, x_N]^T$ be an $N \times 1$ vector of symbols to be transmitted over a noisy channel. Each symbol x_i is chosen from a finite-size alphabet $\mathcal{X}$. Consider a general multiple-input and multiple-output (MIMO) channel model

$$\mathbf{r} = \mathbf{Hx} + \boldsymbol{\xi}, \tag{1}$$

where $\mathbf{H}$ is an $M \times N$ full column rank channel matrix (known to the receiver) with $M \geq N$, $\boldsymbol{\xi} = [\xi_1, \cdots, \xi_M]^T$ is a white Gaussian noise vector where $E(\boldsymbol{\xi}\boldsymbol{\xi}^H) = \sigma^2\mathbf{I}$, and $\mathbf{r} = [r_1, \cdots, r_M]^T$ is the observed received vector. Our task is to detect (estimate) the vector $\mathbf{x} \in \mathcal{X}^N$ given the noisy observation $\mathbf{r}$. We denote the estimate of $\mathbf{x}$ by $\hat{\mathbf{x}} = [\hat{x}_1, \cdots, \hat{x}_N]^T$.

A. Successive cancellation detection using QR decomposition

The QR-decomposition-based successive cancellation detector is captured by the following three steps:

Algorithm 1 (QR-decomposition-based successive cancellation):

1) *QR-decomposition*. Perform the QR-decomposition, $\mathbf{H} = \mathbf{QR}$, where $\mathbf{Q}$ is a tall[3] $M \times N$ column-wise orthonormal matrix and $\mathbf{R}$ is an upper triangular square matrix,

$$\mathbf{R} = \begin{pmatrix} R_{1,1} & R_{1,2} & \dots & R_{1,N} \\ 0 & R_{2,2} & \dots & R_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & R_{N,N} \end{pmatrix},$$

[2]When an error correction code is used, it does not make sense to talk about the optimality of a precoder, because, for example, any invertible precoder would allow the capacity to be achieved, provided that an optimal code for that precoder is used. Therefore, to talk about the optimality of a precoder, one must use constraints other than just capacity achievability.

[3]A tall matrix is a matrix whose number of columns does not exceed the number of rows.

where $R_{k,k} > 0$ for $k = 1, 2, \cdots, N$. Left-multiplying (1) by $\mathbf{Q}^H$, we get

$$\begin{pmatrix} \tilde{r}_1 \\ \tilde{r}_2 \\ \vdots \\ \tilde{r}_N \end{pmatrix} = \begin{pmatrix} R_{1,1} & R_{1,2} & \dots & R_{1,N} \\ 0 & R_{2,2} & \dots & R_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & R_{N,N} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} + \begin{pmatrix} \tilde{\xi}_1 \\ \tilde{\xi}_2 \\ \vdots \\ \tilde{\xi}_N \end{pmatrix}, \tag{2}$$

where $\tilde{\mathbf{r}} = [\tilde{r}_1, \cdots, \tilde{r}_N]^T = \mathbf{Q}^H \mathbf{r}$ and $\tilde{\boldsymbol{\xi}} = [\tilde{\xi}_1, \cdots, \tilde{\xi}_N]^T = \mathbf{Q}^H \boldsymbol{\xi}$. Equation (2) is equivalently written as

$$\tilde{r}_k = [\mathbf{R}]_k x_k + \sum_{m=k+1}^{N} R_{k,m} x_m + \tilde{\xi}_k,$$

where $[\mathbf{R}]_k$ denotes the k-th diagonal entry of $\mathbf{R}$.

2) *Hard decision.* From the last row in (2) we first estimate the symbol x_N by making the hard decision $\hat{x}_N = \mathrm{Quant}\,[\tilde{r}_N/[\mathbf{R}]_N]$. The function $q = \mathrm{Quant}(t)$ sets q to the element of $\mathcal{X}$ that is closest (in terms of Euclidean distance) to t.

3) *Cancellation.* Substitute the estimated symbol $\hat{x}_N$ back into the $(N-1)$-th row in (2) so as to remove the interference term in $\tilde{r}_{N-1}$ and then estimate x_{N-1}. Continue this procedure until we obtain the estimate of the first symbol x_1. The above procedure is described by the following recursive algorithm,

$$\begin{aligned} \hat{x}_N &= \mathrm{Quant}\left[\frac{\tilde{r}_N}{[\mathbf{R}]_N}\right], \\ \hat{x}_k &= \mathrm{Quant}\left[\frac{\tilde{r}_k - \sum_{m=k+1}^{N} R_{k,m}\hat{x}_m}{[\mathbf{R}]_k}\right], \end{aligned}$$

for $k = N-1, N-2, \cdots, 1$.

■

B. Optimally ordered detection

Golden et al. [18] proposed a vertical Bell Laboratories layered space-time (V-BLAST) system with an optimal ordered detection algorithm that maximizes the SNR.

Algorithm 2 (see Golden et al. [18]):

1) *Initial nulling.* Find an initial nulling vector with the smallest norm using zero-forcing. That is, find the index k_N, as the position of the smallest diagonal entry of $\left(\mathbf{H}^H\mathbf{H}\right)^{-1}$,

$$k_N = \arg\min_{1\le j\le N} \left[\left(\mathbf{H}^H\mathbf{H}\right)^{-1}\right]_j. \tag{3}$$

Then, project the received signal $\mathbf{r}$ onto the nulling direction and perform the hard decision to detect the symbol x_{k_N}. That is, set $\hat{x}_{k_N} = \mathrm{Quant}\left[\left(\mathbf{e}_{k_N}^{(N)}\right)^H \mathbf{r}\right]$, where $\mathbf{e}_{k_N}^{(N)}$ is the k_N-th column of $\mathbf{E}^{(N)} = \left(\mathbf{H}^+\right)^H$.

2) *Cancellation.* Subtract the detected signal from the received signal to get

$$\mathbf{r}^{(N-1)} = \mathbf{r} - \mathbf{h}_{k_N}\hat{x}_{k_N},$$

where $\mathbf{h}_{k_N}$ denotes the k_N-th column of $\mathbf{H}$.

3) *Recursion.* Repeat the above two steps until all symbols are detected,

$$\begin{aligned} k_i &= \arg\min_{1\le j\le i} \left[\left(\overline{\mathbf{H}}_{k_{i+1},k_{i+2},\cdots,k_N}^H \overline{\mathbf{H}}_{k_{i+1},k_{i+2},\cdots,k_i}\right)^{-1}\right]_j, \\ \hat{x}_{k_i} &= \mathrm{Quant}\left[\left(\mathbf{e}_{k_i}^{(i)}\right)^H \mathbf{r}^{(i)}\right], \\ \mathbf{r}^{(i-1)} &= \mathbf{r}^{(i)} - \mathbf{h}_{k_i}\hat{x}_{k_i} \qquad \text{for } i = N-1, N-2, \cdots, 1, \end{aligned}$$

where $\mathbf{e}_{k_i}^{(i)}$ denotes the k_i-th column of $\mathbf{E}^{(i)} = \left(\overline{\mathbf{H}}_{k_{i+1},k_{i+2},\cdots,k_N}^+\right)^H$ and $\mathbf{h}_{k_i}$ denotes the k_i-th column of $\overline{\mathbf{H}}_{k_{i+1},k_{i+2},\cdots,k_N}$. ■

C. QR interpretation of Algorithm 2

We use the QR decomposition to interpret the algorithm of Golden et al. [18] given in Section II-B. The first step (3) is equivalent to finding a subchannel whose SNR is the highest among all N possible subchannels. If we look at this problem from the viewpoint of the signal space that is spanned by the column vectors of $\mathbf{H}$, then the first step (3) is to maximize the difference between the column being projected and its projection. Repeat the above procedure for the remaining columns. Finally, Algorithm 2 actually finds the optimal order. That is, it finds a permutation matrix $\mathbf{P} = [\mathbf{p}_{k_1}, \cdots, \mathbf{p}_{k_N}]$, where $\mathbf{p}_i$ denotes an $N \times 1$ vector whose i-th element is one, but others are zeros, such that the QR decomposition of $\mathbf{HP}$ gives rise to the optimally ordered successive cancellation detector.

Algorithm 3 (QR interpretation of Algorithm 2):

1) *Initialization.* Find the column vector of $\mathbf{H}$ that has the longest distance between itself and its projection on the space spanned by all other columns of $\mathbf{H}$. The index of this column is

$$k_N = \arg\max_{1\le k\le N} \left\| \left(\mathbf{I} - \overline{\mathbf{H}}_k \overline{\mathbf{H}}_k^+\right) \mathbf{h}_k \right\|^2,$$

and the column is $\mathbf{h}_{k_N}$. Let $\boldsymbol{\alpha}_N = \left(\mathbf{I} - \overline{\mathbf{H}}_{k_N} \overline{\mathbf{H}}_{k_N}^+\right) \mathbf{h}_{k_N}$ and $\mathbf{q}_N = \boldsymbol{\alpha}_N / \parallel \boldsymbol{\alpha}_N \parallel$.

2) *Recursion.* Repeat the first step by stripping off the column vectors one by one

$$k_i = \arg\max_{\substack{1\le k\le N \\ k\ne k_{i+1},\cdots,k_N}} \left\| (\mathbf{I} - \overline{\mathbf{H}}_{k,k_{i+1},\cdots,k_N} \overline{\mathbf{H}}_{k,k_{i+1},\cdots,k_N}^+)\mathbf{h}_k \right\|^2$$

for $i = N-1, N-2, \cdots, 1$. Define $\boldsymbol{\alpha}_i$ as $\boldsymbol{\alpha}_i = \left(\mathbf{I} - \overline{\mathbf{H}}_{k_i,k_{i+1},\cdots,k_N} \overline{\mathbf{H}}_{k_i,k_{i+1},\cdots,k_N}^+\right) \mathbf{h}_{k_i}$ and let $\mathbf{q}_i = \boldsymbol{\alpha}_i / \parallel \boldsymbol{\alpha}_i \parallel$.

3) *Permutation matrix formulation.* Finally, we obtain the optimal matrix $\mathbf{P} = [\mathbf{p}_{k_1}, \cdots, \mathbf{p}_{k_N}]$ and $\mathbf{Q} = [\mathbf{q}_1, \cdots, \mathbf{q}_N]$ such that $\mathbf{HP} = \mathbf{QR}$. ■

It is easily verified that Algorithm 2 is equivalent to applying Algorithm 1 to $\mathbf{r} = \mathbf{H}'\mathbf{x}' + \boldsymbol{\xi}$, where $\mathbf{H}' = \mathbf{HP}$ and $\mathbf{x} = \mathbf{P}\mathbf{x}'$. Therefore, if we precode a vector $\mathbf{x}'$ with the permutation matrix $\mathbf{P}$, and apply Algorithm 1 to detect $\mathbf{x}'$, we get the optimally ordered successive-cancellation detector of Golden et al. [18]. In the remainder of the paper, we do *not* confine the precoder matrix to be a permutation matrix, and we derive the optimal QR decomposition and the corresponding successive-cancellation detector.

III. EQUAL-DIAGONAL QRS DECOMPOSITION

In this section we develop the equal-diagonal QR decomposition, which we call simply the QRS decomposition. This decomposition will hold the key to the design of the optimal (not necessarily permutation) precoder matrix.

A. QRS decomposition

Lemma 1: Let $\mathbf{H}$ be an $M \times N$ full column rank matrix ($M \geq N$). Then, the diagonal entries in the R-factor $\mathbf{R}$ of the QR decomposition of $\mathbf{H}$ are

$$[\mathbf{R}]_k = \sqrt{\frac{det\left(\mathbf{H}_k^H \mathbf{H}_k\right)}{det\left(\mathbf{H}_{k-1}^H \mathbf{H}_{k-1}\right)}} \qquad \text{for } k = 1, 2, \cdots, N. \quad (4)$$

■

*P*roof: The proof follows directly once the following equation is verified

$$det(\mathbf{H}_k^H \mathbf{H}_k) = det(\mathbf{R}_k^H \mathbf{R}_k) = \prod_{i=1}^{k} [\mathbf{R}]_i^2.$$

□

Definition 1: If a full column rank matrix $\mathbf{H}$ has a QR decomposition $\mathbf{H} = \mathbf{QR}$, where the matrix $\mathbf{R}$ has equal elements along the diagonal, then the matrix $\mathbf{R}$ is called an *equal-diagonal R-factor*. ■

Theorem 1: Let $\mathbf{H}$ be an $M \times N$ full column rank matrix ($M \geq N$). Then, $\mathbf{H}$ has an equal-diagonal R-factor if and only if the submatrix $\mathbf{H}_k$ of $\mathbf{H}$ satisfies

$$det\left(\mathbf{H}_k^H \mathbf{H}_k\right) = det\left(\mathbf{H}^H \mathbf{H}\right)^{\frac{k}{N}} \qquad \text{for } k = 1, 2, \cdots, N. \quad (5)$$

■

Proof. By Lemma 1, the series of equalities $[\mathbf{R}]_1 = [\mathbf{R}]_2 = \cdots = [\mathbf{R}]_N$ is equivalent to

$$\frac{det\left(\mathbf{H}_1^H \mathbf{H}_1\right)}{det\left(\mathbf{H}_0^H \mathbf{H}_0\right)} = \frac{det\left(\mathbf{H}_2^H \mathbf{H}_2\right)}{det\left(\mathbf{H}_1^H \mathbf{H}_1\right)} = \cdots = \frac{det\left(\mathbf{H}_N^H \mathbf{H}_N\right)}{det\left(\mathbf{H}_{N-1}^H \mathbf{H}_{N-1}\right)}.$$

This shows that $det\left(\mathbf{H}_k^H \mathbf{H}_k\right)$ is a geometric sequence in k. Since $det\left(\mathbf{H}_0^H \mathbf{H}_0\right) = 1$ and $det\left(\mathbf{H}_N^H \mathbf{H}_N\right) = det\left(\mathbf{H}^H \mathbf{H}\right)$, we have $det\left(\mathbf{H}_k^H \mathbf{H}_k\right) = det\left(\mathbf{H}^H \mathbf{H}\right)^{k/N}$. □

Theorem 1 gives us a necessary and sufficient condition to check if a matrix $\mathbf{H}$ has an equal-diagonal R-factor. But it does not tell us how to transform an arbitrary matrix into a matrix with an equal-diagonal R-factor. We next develop the apparatus that will enable this transformation.

Lemma 2: Let $\mathbf{A}$ be a positive definite matrix. Then, the equations

$$\boldsymbol{\gamma}^H \mathbf{A} \boldsymbol{\gamma} = 1 \quad (6)$$
$$\boldsymbol{\gamma}^H \boldsymbol{\gamma} = 1 \quad (7)$$

have a solution vector $\boldsymbol{\gamma}$ if and only if $\lambda_{\min}(\mathbf{A}) \leq 1$ and $\lambda_{\max}(\mathbf{A}) \geq 1$, where $\lambda_{\max}(\mathbf{A})$ is the maximal eigenvalue of $\mathbf{A}$, and $\lambda_{\min}(\mathbf{A})$ is the minimal eigenvalue of $\mathbf{A}$. ■

Proof: For any unit norm vector $\boldsymbol{\gamma}$, we have $\lambda_{\min}(\mathbf{A}) \leq \boldsymbol{\gamma}^H \mathbf{A} \boldsymbol{\gamma} \leq \lambda_{\max}(\mathbf{A})$. Hence, equations (6) and (7) have a solution if and only if $\lambda_{\min}(\mathbf{A}) \leq 1$ and $\lambda_{\max}(\mathbf{A}) \geq 1$. □

Employing Lemma 2 we further get the following result.

Corollary 1.1: Let $\mathbf{A}$ be a $K \times K$ positive definite matrix, and $\mathbf{B}$ be an $N \times N$ positive definite matrix, where $K \leq N$. If $det(\mathbf{A})^{1/K} = det(\mathbf{B})^{1/N}$, then, the following equations

$$\boldsymbol{\gamma}^H \mathbf{A} \boldsymbol{\gamma} = det(\mathbf{B})^{\frac{1}{N}}$$
$$\boldsymbol{\gamma}^H \boldsymbol{\gamma} = 1$$

always have a solution vector $\boldsymbol{\gamma}$. ■

Proof. Let $\widetilde{\mathbf{A}} = \mathbf{A}/det(\mathbf{B})^{1/N}$. Then, $\lambda_{\min}(\widetilde{\mathbf{A}}) \leq 1$ and $\lambda_{\max}(\widetilde{\mathbf{A}}) \geq 1$. Applying Lemma 2 to $\widetilde{\mathbf{A}}$ yields the desired result. □

Canonical equations: The following system of equations will prove to be crucial for the remainder of the paper. Let $\mathbf{H}$ be a full column rank tall $M \times N$ matrix and $\sqrt{\lambda_k}$ ($k = 1, 2, \cdots, N$) be its singular values that we arrange into a matrix $\boldsymbol{\Lambda} = [diag(\sqrt{\lambda_1}, \cdots, \sqrt{\lambda_N}), \mathbf{0}_{N\times(M-N)}]^T$ (here $\lambda_1, \ldots, \lambda_N$ are the eigenvalues of $\mathbf{H}^H\mathbf{H}$). We seek to find a matrix $\mathbf{V}$ of orthonormal vectors $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_N]$, such that the vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_N$ satisfy the following constraints:

1) The vector $\mathbf{v}_1$ satisfies

$$\mathbf{v}_1^H \boldsymbol{\Lambda}^H \boldsymbol{\Lambda} \mathbf{v}_1 = det\left(\mathbf{H}^H \mathbf{H}\right)^{1/N} \quad (8)$$
$$\mathbf{v}_1^H \mathbf{v}_1 = 1. \quad (9)$$

2) For $k \geq 1$, the vector $\mathbf{v}_{k+1}$ may be expressed as $\mathbf{v}_{k+1} = \mathbf{V}_k^{\perp} \mathbf{z}_{k+1}$, where $\mathbf{z}_{k+1}$ satisfies

$$\mathbf{z}_{k+1}^H \mathbf{A}^{(k)} \mathbf{z}_{k+1} = det\left(\mathbf{H}^H \mathbf{H}\right)^{1/N} \quad (10)$$
$$\mathbf{z}_{k+1}^H \mathbf{z}_{k+1} = 1, \quad (11)$$

with

$$\mathbf{A}^{(k)} = (\boldsymbol{\Lambda} \mathbf{V}_k^{\perp})^H \mathcal{P}_{\boldsymbol{\Lambda}\mathbf{V}_k} \boldsymbol{\Lambda} \mathbf{V}_k^{\perp}. \quad (12)$$

Theorem 2: Let $\mathbf{H}$ be a full column rank tall $M \times N$ matrix and $\sqrt{\lambda_k}$ ($k = 1, 2, \cdots, N$) be its singular values arranged into a matrix $\boldsymbol{\Lambda} = [diag(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \cdots, \sqrt{\lambda_N}), \mathbf{0}_{N\times(M-N)}]^T$. There always exists a unitary matrix $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_N]$ whose columns $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_N$ satisfy the canonical equations (8)–(12). ■

The proof of Theorem 2 is given in the Appendix.

We next show that the matrix $\mathbf{V}$ in Theorem 2 is comprised of the eigenvectors of $\mathbf{H}^H\mathbf{H}$ if and only if $\mathbf{H}$ has an equal-diagonal R-factor.

Theorem 3: Let $\mathbf{H}$ be an $M \times N$ full column rank matrix ($M \geq N$) and let $\mathbf{H} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{V}$ be the singular value decomposition (SVD) of $\mathbf{H}$ with $\boldsymbol{\Lambda} = [diag(\sqrt{\lambda_1}, \cdots, \sqrt{\lambda_N}), \mathbf{0}_{N\times(M-N)}]^T$, where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N > 0$. Then, $\mathbf{H}$ has an equal-diagonal R-factor if and only if the vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_N$ of the unitary matrix $\mathbf{V}$ and the singular value matrix $\boldsymbol{\Lambda}$ satisfy the canonical equations (8)-(12). ■

The proof of Theorem 3 is given in the Appendix.

Theorem 3 not only provides a criterion to judge if the diagonal elements of the R-factor are equal, but also implicitly provides a recursive algorithm to construct a matrix with such R-factor. This recursive algorithm is obtained by successively constructing the vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_N$ to satisfy equations (8)-(11). We provide the algorithm in Sections III–B and III–C.

Theorem 4: (Equal-diagonal QRS decomposition) For an arbitrary $M \times N$ matrix $\mathbf{H}$ with rank r, there exists a unitary matrix $\mathbf{S}$ such that $\mathbf{HS}$ has an equal-diagonal R-factor, i.e.

$$\mathbf{HS} = \mathbf{QR}, \tag{13}$$

where $\mathbf{Q}$ is an $M \times r$ column-wise orthonormal matrix and $\mathbf{R} = \left[\begin{array}{cc} \mathbf{R}_{r\times r} & \mathbf{0}_{r\times(N-r)} \end{array}\right]$ with $\mathbf{R}_{r\times r}$ being the equal-diagonal R-factor. ■

Proof: Let the singular-value decomposition (SVD) of the matrix $\mathbf{H}$ be $\mathbf{H} = \mathbf{U\Lambda V}$, where $\mathbf{U}$ is an $M \times M$ unitary matrix, $\mathbf{V}$ is an $N \times N$ unitary matrix and

$$\mathbf{\Lambda} = \left(\begin{array}{cc} \mathbf{\Lambda}_{r\times r} & \mathbf{0}_{r\times(N-r)} \\ \mathbf{0}_{(M-r)\times r} & \mathbf{0}_{(M-r)\times(N-r)} \end{array} \right), \tag{14}$$

with $\mathbf{\Lambda}_{r\times r} = diag\left(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \cdots, \sqrt{\lambda_r}\right), \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$. Let $\widetilde{\mathbf{H}} = \mathbf{H} \cdot \left(\mathbf{V}^H\right)_r$, i.e.,

$$\widetilde{\mathbf{H}} = \mathbf{U} \left(\begin{array}{c} \mathbf{\Lambda}_{r\times r} \\ \mathbf{0}_{(M-r)\times r} \end{array} \right). \tag{15}$$

Now the key to the proof is to show that there exists a unitary matrix $\widetilde{\mathbf{S}}$ such that the diagonal entries in the R-factor of the QR decomposition of $\widetilde{\mathbf{H}}\widetilde{\mathbf{S}}$ are equal. If $\widetilde{\mathbf{H}}\widetilde{\mathbf{S}}$ is to have an equal-diagonal R-factor, then by Theorem 3, the matrix of its right singular vectors must satisfy the canonical equations (8)–(12). (Notice that by multiplying $\widetilde{\mathbf{H}}$ by a unitary matrix $\widetilde{\mathbf{S}}$, we do not change the singular values, so the equations (8)–(12) keep their form.) Since (by Theorem 2) the canonical equations (8)–(12) always have a solution, we conclude that the desired matrix of right singular vectors exists, and hence the matrix $\widetilde{\mathbf{S}}$ with the designed property must exist too. Therefore, we can write $\widetilde{\mathbf{H}}\widetilde{\mathbf{S}} = \mathbf{QR}_{r\times r}$ with $[\mathbf{R}_{r\times r}]_1 = [\mathbf{R}_{r\times r}]_2 = \cdots = [\mathbf{R}_{r\times r}]_r = \sqrt{\lambda_1\lambda_2\cdots\lambda_r}^{1/r}$. Let $\mathbf{S} = [(\mathbf{V}^H)_r\widetilde{\mathbf{S}}, (\overline{\mathbf{V}^H})_{1,2,\cdots,r}]$. Then, $\mathbf{S}$ is an $N \times N$ unitary matrix satisfying $\mathbf{HS} = [\widetilde{\mathbf{H}}\widetilde{\mathbf{S}}, \mathbf{H}(\overline{\mathbf{V}^H})_{1,2,\cdots,r}] = \mathbf{Q}[\mathbf{R}_{r\times r}, \mathbf{0}_{r\times(N-r)}] = \mathbf{QR}$. □

B. Construction of the S-factor

Implicitly stated in the proof of Theorem 4 is the following recursive algorithm to find the S-factor of the QRS decomposition $\mathbf{HS} = \mathbf{QR}$.

Algorithm 4 (Construction of the S-factor):

1) *SVD*. Perform the SVD of $\mathbf{H} = \mathbf{U\Lambda V}$ and form $\widetilde{\mathbf{H}}$ according to (15).
2) *Initialization*. Determine the first column of $\widetilde{\mathbf{S}}$, i.e., $\widetilde{\mathbf{s}}_1 = (\widetilde{S}_{1,1}, \cdots, \widetilde{S}_{r,1})^T$, such that constraints

$$\widetilde{\mathbf{s}}_1^H\widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}}\widetilde{\mathbf{s}}_1 = det\left(\widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}}\right)^{1/r} \tag{16}$$

$$\widetilde{\mathbf{s}}_1^H\widetilde{\mathbf{s}}_1 = 1 \tag{17}$$

are satisfied.
3) *Recursion* (reduce the dimension and decouple constraints). Set $\widetilde{\mathbf{s}}_{k+1} = \widetilde{\mathbf{S}}_k^\perp\mathbf{z}_{k+1}$, where $\mathbf{z}_{k+1}$ is any vector that satisfies

$$\mathbf{z}_{k+1}^H\mathbf{C}^{(k)}\mathbf{z}_{k+1} = det\left(\widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}}\right)^{1/r} \tag{18}$$

$$\mathbf{z}_{k+1}^H\mathbf{z}_{k+1} = 1, \tag{19}$$

with $\mathbf{C}^{(k)} = (\widetilde{\mathbf{H}}\widetilde{\mathbf{S}}_k^\perp)^H\mathcal{P}_{\widetilde{\mathbf{H}}\widetilde{\mathbf{S}}_k}\widetilde{\mathbf{H}}\widetilde{\mathbf{S}}_k^\perp$.
4) *Complete the S-factor*. $\mathbf{S} = [\mathbf{V}_r\widetilde{\mathbf{S}}, (\overline{\mathbf{V}^H})_{1,\cdots,r}]$. ■

Remark: We would like to make some comments on the recursive algorithm.

- The first column of $\widetilde{\mathbf{S}}$ must lie both on the hyper-ellipse and the hyper-sphere, which are determined by the quadratic equations (16) and (17), respectively. These equations do have a solution (see Corollary 1.1). In fact, they have an infinite number of solutions when $r > 2$. Additional constraints may be imposed to narrow the solution space, but we do not pursue this option here. For us, all possible solutions of (16)-(19) are equally good.
- From Theorem 1 we know that $det\left(\mathbf{C}^{(k)}\right) = det\left(\widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}}\right)^{(r-k)/r}$. This, in fact, is equivalent to $det(\mathbf{C}^{(k)})^{1/(r-k)} = det(\widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}})^{1/r}$. Therefore, in each recursion, equations (18) and (19) keep the same structure. This structure ensures (see Corollary 1.1) that the equations (18) and (19) have a solution for every $k \geq 1$.
- Once we have determined the first column $\widetilde{\mathbf{s}}_1$ of $\widetilde{\mathbf{S}}$, we look for the second column $\widetilde{\mathbf{s}}_2$ of $\widetilde{\mathbf{S}}$ in the space orthogonal to $\widetilde{\mathbf{s}}_1$ via the transform $\widetilde{\mathbf{s}}_2 = \widetilde{\mathbf{S}}_1^\perp\mathbf{z}_2$. Thereby, the orthogonality $\widetilde{\mathbf{s}}_1^H\widetilde{\mathbf{s}}_2 = 0$ is automatically satisfied; further, the constraints on $\widetilde{\mathbf{s}}_2$ are transformed into constraints (18)-(19) which involve only $\mathbf{z}_2$. More importantly, we transform joint constraints on $\widetilde{\mathbf{s}}_1$ and $\widetilde{\mathbf{s}}_2$, into disjoint constraints: one for $\widetilde{\mathbf{s}}_1$, and the other for $\mathbf{z}_2$ alone. Thereby, we also preserve the structure of the recursive algorithm, i.e., the constraints (16) and (17) associated with $\widetilde{\mathbf{s}}_1$ have the same structure as the constraints (18)-(19) associated with $\mathbf{z}_2$, which is the key to the successful design of the unitary matrix $\widetilde{\mathbf{S}}$ in question. We continue the procedure until we determine the last column $\widetilde{\mathbf{s}}_r$.

C. An explicit S-factor

In Section III-B we established that the solution for the S-factor is not unique. In the following we show how to find an explicit special solution of equations (16)-(19). Since the recursive step in Algorithm 4 involves two quadratic equations with the same structure, the key to solving (16)-(19) is to solve for $\mathbf{z}$ the following equations.

$$\mathbf{z}^H\mathbf{C}\mathbf{z} = det(\widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}})^{1/r} \tag{20}$$

$$\mathbf{z}^H\mathbf{z} = 1, \tag{21}$$

where $\mathbf{C}$ is a positive definite matrix. Let the eigenvalue decomposition of $\mathbf{C}$ be $\mathbf{C} = \mathbf{V}^{(\mathbf{C})}\mathbf{\Lambda}^{(\mathbf{C})}(\mathbf{V}^{(\mathbf{C})})^H$, where $\mathbf{\Lambda}^{(\mathbf{C})} = diag\left(\lambda_1^{(\mathbf{C})}, \cdots, \lambda_\rho^{(\mathbf{C})}\right)$ with $\lambda_1^{(\mathbf{C})} \geq \lambda_2^{(\mathbf{C})} \geq \cdots \geq \lambda_\rho^{(\mathbf{C})} > 0$.

Let $\mathbf{z} = \mathbf{V}^{(\mathbf{C})}\mathbf{y}$, where $\mathbf{y} = [y_1, \cdots, y_\rho]^T$. Substituting this into (20) and (21), respectively, yields

$$\begin{aligned} \mathbf{y}^H \mathbf{\Lambda}^{(\mathbf{C})} \mathbf{y} &= det(\widetilde{\mathbf{H}}^H \widetilde{\mathbf{H}})^{1/r} &(22)\\ \mathbf{y}^H \mathbf{y} &= 1. &(23)\end{aligned}$$

Now we consider two cases:

Case 1 $\rho = 2$. Equations (22) and (23) have four real solutions, one of which is

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \sqrt{\frac{det(\widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}})^{1/r} - \lambda_2^{(\mathbf{C})}}{\lambda_1^{(\mathbf{C})} - \lambda_2^{(\mathbf{C})}}} \\ \sqrt{\frac{\lambda_1^{(\mathbf{C})} - det(\widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}})^{1/r}}{\lambda_1^{(\mathbf{C})} - \lambda_2^{(\mathbf{C})}}} \end{pmatrix}. \qquad (24)$$

Case 2 $\rho > 2$. In this case, there is an infinite number of solutions to equations (22) and (23). We can obtain a special solution as follows,

$$\begin{aligned} y_1 &= \sqrt{\frac{det(\widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}})^{1/r} - \lambda_\rho^{(\mathbf{C})}}{\lambda_1^{(\mathbf{C})} - \lambda_\rho^{(\mathbf{C})}}} \\ y_\ell &= 0 \qquad \textit{for } \ell = 2, \cdots, \rho - 1 &(25)\\ y_\rho &= \sqrt{\frac{\lambda_1^{(\mathbf{C})} - det(\widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}})^{1/r}}{\lambda_1^{(\mathbf{C})} - \lambda_\rho^{(\mathbf{C})}}}. \end{aligned}$$

Using the above specific solutions (24) and (25) in each recursion of Algorithm 4, we get the following algorithm to compute an explicit S-factor.

Algorithm 5 (Construction of an explicit S-factor):

1) *SVD*. Perform the SVD of $\mathbf{H} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}$ and form $\widetilde{\mathbf{H}}$ according to (15).
2) *Initialization*. An explicit solution for the first column of $\widetilde{\mathbf{S}}$, i.e., $\widetilde{\mathbf{s}}_1 = (\widetilde{S}_{1,1}, \cdots, \widetilde{S}_{r,1})^T$, is

$$\begin{aligned} \widetilde{S}_{1,1} &= \sqrt{\frac{det(\widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}})^{1/r} - \lambda_r}{\lambda_1 - \lambda_r}} \\ \widetilde{S}_{k,1} &= 0 \qquad \text{for } k = 2, \cdots, r-1 \\ \widetilde{S}_{r,1} &= \sqrt{\frac{\lambda_1 - det(\widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}})^{1/r}}{\lambda_1 - \lambda_r}}. \end{aligned}$$

In this case, it is not hard to verify that

$$\widetilde{\mathbf{S}}_1^\perp = \begin{pmatrix} -\widetilde{S}_{r,1} & \mathbf{0}_{1\times(r-2)} \\ \mathbf{0}_{(r-2)\times 1} & \mathbf{I}_{(r-2)\times(r-2)} \\ \widetilde{S}_{1,1} & \mathbf{0}_{1\times(r-2)} \end{pmatrix}.$$

3) *Form the positive definite matrix for recursion*. Set (starting initially with $k = 1$)

$$\mathbf{C}^{(k)} = (\widetilde{\mathbf{H}}\widetilde{\mathbf{S}}_k^\perp)^H \mathcal{P}_{\widetilde{\mathbf{H}}\widetilde{\mathbf{S}}_k} \widetilde{\mathbf{H}}\widetilde{\mathbf{S}}_k^\perp.$$

4) *Eigenvalue decomposition*. Perform the eigenvalue decomposition of $\mathbf{C}^{(k)}$,

$$\mathbf{C}^{(k)} = \mathbf{V}^{(k)}\mathbf{\Lambda}^{(k)}\left(\mathbf{V}^{(k)}\right)^H,$$

where $\mathbf{V}^{(k)}$ is a $(r-k) \times (r-k)$ unitary matrix and $\mathbf{\Lambda}^{(k)} = diag(\lambda_1^{(k)}, \cdots, \lambda_{r-k}^{(k)})$ with $\lambda_1^{(k)} \geq \lambda_2^{(k)} \geq \cdots \geq \lambda_{r-k}^{(k)} > 0$.
5) *Recursion*. Set $\widetilde{\mathbf{s}}_{k+1} = \widetilde{\mathbf{S}}_k^\perp \mathbf{V}^{(k)}\mathbf{y}^{(k)}$ for $k = 1, \cdots, r-2$, where $\mathbf{y}^{(k)} = \left[y_1^{(k)}, y_2^{(k)}, \cdots, y_{r-k}^{(k)}\right]^T$ is determined by

$$\begin{aligned} y_1^{(k)} &= \sqrt{\frac{det(\widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}})^{1/r} - \lambda_{r-k}^{(k)}}{\lambda_1^{(k)} - \lambda_{r-k}^{(k)}}} \\ y_\ell^{(k)} &= 0 \qquad \text{for } \ell = 2, \cdots, r-k-1 \\ y_{r-k}^{(k)} &= \sqrt{\frac{\lambda_1^{(k)} - det(\widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}})^{1/r}}{\lambda_1^{(k)} - \lambda_{r-k}^{(k)}}}. \end{aligned}$$

When $k = r-1$, set $\widetilde{\mathbf{s}}_r = \widetilde{\mathbf{S}}_{r-2}^\perp \mathbf{V}^{(r-2)}\mathbf{y}^{(r-1)}$ where

$$\mathbf{y}^{(r-1)} = \begin{pmatrix} -\sqrt{\frac{det(\widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}})^{1/r} - \lambda_2^{(r-2)}}{\lambda_1^{(r-2)} - \lambda_2^{(r-2)}}} \\ \sqrt{\frac{\lambda_1^{(r-2)} - det(\widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}})^{1/r}}{\lambda_1^{(r-2)} - \lambda_2^{(r-2)}}} \end{pmatrix}.$$

6) *Complete the explicit S-factor*.
$\mathbf{S} = [(\mathbf{V}^H)_r \widetilde{\mathbf{S}}, (\mathbf{V}^H)_{1,\cdots,r}]$. ■

For example, when $\mathbf{H} = diag(\sqrt{\lambda_1}, \sqrt{\lambda_2})$, the S-factor is

$$\mathbf{S} = \begin{pmatrix} \sqrt{\frac{\sqrt{\lambda_2}}{\sqrt{\lambda_1}+\sqrt{\lambda_2}}} & -\sqrt{\frac{\sqrt{\lambda_1}}{\sqrt{\lambda_1}+\sqrt{\lambda_2}}} \\ \sqrt{\frac{\sqrt{\lambda_1}}{\sqrt{\lambda_1}+\sqrt{\lambda_2}}} & \sqrt{\frac{\sqrt{\lambda_2}}{\sqrt{\lambda_1}+\sqrt{\lambda_2}}} \end{pmatrix}.$$

D. Properties of equal-diagonal R-factors

Definition 2: Define the minimum distance of the constellation $\mathcal{X}$ as

$$d_{\min}(\mathcal{X}) = \min_{x,x'\in\mathcal{X}, x\neq x'} |x - x'| = \sqrt{\min_{\mathbf{x},\mathbf{x}'\in\mathcal{X}^N, \mathbf{x}\neq\mathbf{x}'} ||\mathbf{x} - \mathbf{x}'||^2}. \qquad (26)$$

Definition 3: Define the free distance of an $M \times N$ channel matrix $\mathbf{H}$ as

$$d_{\text{free}}(\mathbf{H}) = \sqrt{\min_{\mathbf{x},\mathbf{x}'\in\mathcal{X}^N, \mathbf{x}\neq\mathbf{x}'} (\mathbf{x} - \mathbf{x}')^H \mathbf{H}^H \mathbf{H} (\mathbf{x} - \mathbf{x}')}. \quad (27)$$

The following theorem shows that the free distance can be bounded in terms of the diagonal entries of the R-factor in the QR decomposition of a channel matrix.

Theorem 5: Let $\mathbf{H}$ be an $M \times N$ full column rank tall matrix and $\mathbf{R}$ be the R-factor in the QR decomposition of $\mathbf{H} = \mathbf{QR}$. Denoting the diagonal elements of $\mathbf{R}$ by $[\mathbf{R}]_k$, we have,

$$\left[\min_{1\leq k\leq N} [\mathbf{R}]_k\right] \cdot d_{\min}(\mathcal{X}) \leq d_{\text{free}}(\mathbf{H}) \leq [\mathbf{R}]_1 \cdot d_{\min}(\mathcal{X}). \quad (28)$$

■

Proof: Consider two different signal vectors: $\mathbf{x} = [x_1, x_2, \cdots, x_N]^T$ and $\mathbf{x}' = [x_1', x_2', \cdots, x_N']^T$. If $x_k = x_k'$ for $k = 2, \cdots, N$, but $x_1 \neq x_1'$, then

$$(\mathbf{x} - \mathbf{x}')^H \mathbf{H}^H \mathbf{H} (\mathbf{x} - \mathbf{x}') = [\mathbf{R}]_1^2 |x_1 - x_1'|^2. \qquad (29)$$

Hence, by taking the minima of both sides of (29), we get

$$d^2_{\text{free}}(\mathbf{H}) \le \min_{x_1,x'_1\in\mathcal{X},x_1\neq x'_1} [\mathbf{R}]_1^2 \cdot |x_1 - x'_1|^2 = [\mathbf{R}]_1^2 \cdot d^2_{\min}(\mathcal{X}),$$

which leads to $d_{\text{free}}(\mathbf{H}) \le [\mathbf{R}]_1 \cdot d_{\min}(\mathcal{X})$. This completes the proof of the right-hand side of inequality (28). To prove the left hand side of (28), note that

$$(\mathbf{x}-\mathbf{x}')^H\mathbf{H}^H\mathbf{H}(\mathbf{x}-\mathbf{x}') = \sum_{i=1}^{N}\left|\sum_{j=i}^{N} R_{i,j}\cdot(x_j - x'_j)\right|^2. \quad (30)$$

Assume $\mathbf{x} \neq \mathbf{x}'$. Let k be an integer such that $x_i = x'_i$, for $i > k$, but $x_k \neq x'_k$. Then, from (30), using the upper triangularity of $\mathbf{R}$, we have

$$\begin{aligned}(\mathbf{x}-\mathbf{x}')^H\mathbf{H}^H\mathbf{H}(\mathbf{x}-\mathbf{x}') &= \sum_{i=1}^{k}\left|\sum_{j=i}^{k} R_{i,j}(x_j - x'_j)\right|^2 \\ &\ge [\mathbf{R}]_k^2|x_k - x'_k|^2 \\ &\ge [\mathbf{R}]_k^2 \cdot d^2_{\min}(\mathcal{X}). \quad (31)\end{aligned}$$

Taking the minima of both sides of (31) yields $d_{\text{free}}(\mathbf{H}) \ge \left[\min_{1\le k\le N}[\mathbf{R}]_k\right]\cdot d_{\min}(\mathcal{X})$, which completes the proof. □

From Theorem 5 we immediately obtain:

Property 1: Let $\mathbf{H}$ be an $M \times N$ full column rank tall matrix that has an equal-diagonal R-factor. Then,

$$d_{\text{free}}(\mathbf{H}) = \left(det\left(\mathbf{H}^H\mathbf{H}\right)\right)^{1/2N} d_{\min}(\mathcal{X}) = [\mathbf{R}]_k \cdot d_{\min}(\mathcal{X}). \quad (32)$$

■

Property 1 shows that for a channel matrix $\mathbf{H}$ with an equal-diagonal R-factor, the minimum Euclidean distance of the signal lattice before the channel ($d_{\min}(\mathcal{X})$) and the minimum Euclidean distance of the signal lattice after the channel ($d_{\text{free}}(\mathbf{H})$), are equivalent up to a multiplicative factor $[\mathbf{R}]_k$. We know that the free distance determines the detection performance of the maximum likelihood detector when the signal to noise ratio is high [39]. From Property 1 we conclude that if the channel $\mathbf{H}$ has an equal-diagonal R-factor, then the free distance is computed by the QR decomposition. Therefore, this suggests that if the channel matrix has an equal-diagonal R-factor, the detection performance of the QR successive cancellation detector is asymptotically equivalent to that of the maximum likelihood detector as SNR $\to \infty$ (which we will demonstrate in Section V and Section VI).

Suppose we wish to use the VBLAST detector [18] on a channel that has the equal-diagonal R-factor. A natural question is: What is the optimal detection order? The following property gives the answer.

Property 2: If a channel matrix has an equal-diagonal R-factor, the optimal detection order (that ensures that the high SNR components are detected first) is the natural order, i.e., $x_N \to x_{N-1} \to \cdots \to x_1$, in other words, the i-th symbol to be detected is the symbol x_{N+1-i}. ■

Proof of Property 2: The Gram-Schmidt orthogonalization procedure of $\mathbf{H}$ is described as follows:

$$\boldsymbol{\beta}_1 = \mathbf{h}_1 \quad (33)$$

$$\boldsymbol{\beta}_j = \mathbf{h}_j - \sum_{k=1}^{j-1} c_{j,k}\boldsymbol{\beta}_k, \qquad \text{for } j = 2,\cdots,N, \quad (34)$$

$$c_{j,k} = \frac{\mathbf{h}_j^H\boldsymbol{\beta}_k}{\boldsymbol{\beta}_k^H\boldsymbol{\beta}_k}. \quad (35)$$

First, we consider the N-th column $\mathbf{h}_N$. Suppose that we exchange an arbitrary K-th column $\mathbf{h}_K$ of $\mathbf{H}$ with the N-th column $\mathbf{h}_N, K \neq N$. Let $\acute{\mathbf{H}} = [\mathbf{H}_{K-1}, \mathbf{h}_N, \mathbf{h}_{K+1}, \mathbf{h}_{K+2}, \cdots, \mathbf{h}_{N-1}, \mathbf{h}_K]$. Then, the Gram-Schmidt orthogonalization procedure of the first $K-1$ columns of $\acute{\mathbf{H}}$ is the same as that of the first $K-1$ columns of $\mathbf{H}$, but beyond the K-th column, they are different, i.e.,

$$\acute{\boldsymbol{\beta}}_j = \boldsymbol{\beta}_j \qquad \text{for } j = 1,2,\cdots,K-1 \quad (36)$$

$$\acute{\boldsymbol{\beta}}_j = \acute{\mathbf{h}}_j - \sum_{k=1}^{j-1}\acute{c}_{j,k}\acute{\boldsymbol{\beta}}_k \qquad \text{for } j = K, K+1,\cdots,N \quad (37)$$

$$\acute{c}_{j,k} = \frac{\acute{\mathbf{h}}_j^H\acute{\boldsymbol{\beta}}_j}{\acute{\boldsymbol{\beta}}_j^H\acute{\boldsymbol{\beta}}_j}. \quad (38)$$

Using (37) for $j \ge K$, we obtain

$$\acute{\boldsymbol{\beta}}_j^H\acute{\boldsymbol{\beta}}_j = \acute{\mathbf{h}}_j^H\acute{\mathbf{h}}_j - \sum_{k=1}^{j-1}|\acute{c}_{j,k}|^2\acute{\boldsymbol{\beta}}_k^H\acute{\boldsymbol{\beta}}_k. \quad (39)$$

Utilizing (39), once for $j = N$ and again for $j = K$, yields

$$\begin{aligned}[\acute{\mathbf{R}}]_N^2 &= \acute{\boldsymbol{\beta}}_N^H\acute{\boldsymbol{\beta}}_N \\ &= \mathbf{h}_K^H\mathbf{h}_K - \sum_{k=1}^{K-1}|c_{N,k}|^2\boldsymbol{\beta}_k^H\boldsymbol{\beta}_k - \sum_{k=K}^{N-1}|\acute{c}_{N,k}|^2\acute{\boldsymbol{\beta}}_k^H\acute{\boldsymbol{\beta}}_k \\ &= \boldsymbol{\beta}_K^H\boldsymbol{\beta}_K - \sum_{k=K}^{N-1}|\acute{c}_{N,k}|^2\acute{\boldsymbol{\beta}}_k^H\acute{\boldsymbol{\beta}}_k \\ &= [\mathbf{R}]_K^2 - \sum_{k=K}^{N-1}|\acute{c}_{N,k}|^2[\acute{\mathbf{R}}]_k^2, \quad (40)\end{aligned}$$

where we have used $\mathbf{h}_j = \acute{\mathbf{h}}_j$ for $1 \le j \le K-1$, $\acute{\mathbf{h}}_N = \mathbf{h}_K$, $[\mathbf{R}]_k = \sqrt{\boldsymbol{\beta}_k^H\boldsymbol{\beta}_k}$ and $[\acute{\mathbf{R}}]_k = \sqrt{\acute{\boldsymbol{\beta}}_k^H\acute{\boldsymbol{\beta}}_k}$. Equation (40) implies that $[\acute{\mathbf{R}}]_N^2 \le [\mathbf{R}]_K^2$, i.e., $[\acute{\mathbf{R}}]_N \le [\mathbf{R}]_K = [\mathbf{R}]_N$, since our assumption is that $\mathbf{H}$ possesses the equal-diagonal R-factor. We have thus proved that if the K-th column and the N-th column of the matrix $\mathbf{H}$ are exchanged, then, the N-th diagonal entry in the R-factor is not increased. Therefore, the N-th symbol x_N should be detected first. By induction, we can complete the proof of Property 2. □

Property 2 essentially characterizes a geometric property of a channel with the equal-diagonal R-factor. Namely, among all column vectors of $\mathbf{H}$, the last column vector $\mathbf{h}_N$ has the maximal distance from the space spanned by all the remaining column vectors. After we have eliminated $\mathbf{h}_N$, among all remaining column vectors of $\mathbf{H}$, the second to last column vector

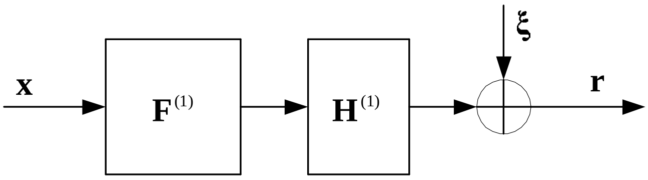

Fig. 1. Equivalent matrix receiver model for a linearly precoded TDMA system.

$\mathbf{h}_{N-1}$ has the maximum distance from the space spanned by all the remaining column vectors (except $\mathbf{h}_N$). We continue this procedure until we reach the first column vector $\mathbf{h}_1$.

Recently, Hassibi [19] derived a fast square root algorithm, which, just like the algorithm in this paper, is based on a certain QR decomposition and the Schur decomposition. It is therefore natural to compare Hassibi's algorithm [19] to our method. Basically, the method in [19], just like ours, efficiently implements the optimal permutation order at the detector. The main difference between our algorithm and Hassibi's algorithm is that our decomposition not only directly gives the optimal ordering, but also guarantees that the minimum diagonal entry in the R-factor achieves the upper bound, which in turn reduces the probability of error (we prove this in Section V). Hassibi's decomposition [19] does not have this property. On the other hand, to implement our decomposition, both the transmitter and the receiver need to know the channel matrix, whereas Hassibi's decomposition requires that the channel matrix be known only to the receiver. So, there is a tradeoff between the two methods.

IV. PRECODED TRANSMISSION MODEL

In this section, we consider two kinds of transmission techniques with linear precoders: the TDMA system and the OFDM system.

A. Linearly precoded TDMA

Figure 1 shows the discrete-time equivalent model of the baseband communication system using filter-bank precoders proposed by Scaglione, et al. [21], [35]. The signal vector $\mathbf{x} = [x_1, x_2, \cdots, x_N]^T$ is precoded by an $N \times N$ matrix $\mathbf{F}^{(1)}$ and then transmitted through the channel, where the $M \times N$ channel matrix is

$$\mathbf{H}^{(1)} = \begin{pmatrix} h(0) & 0 & \dots & 0 \\ h(1) & h(0) & \dots & 0 \\ \vdots & h(1) & \ddots & \vdots \\ h(L) & \ddots & \ddots & h(0) \\ 0 & \ddots & \ddots & h(1) \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \ddots & 0 & h(L) \end{pmatrix}.$$

Here, $M = N + L$. In this case, the received signal $\mathbf{r}$ can be expressed as

$$\mathbf{r} = \mathbf{H}^{(1)}\mathbf{F}^{(1)}\mathbf{x} + \boldsymbol{\xi}. \tag{41}$$

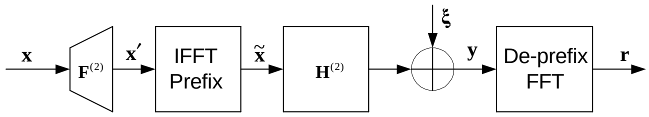

Fig. 2. Linearly precoded OFDM system

In this model, we assume that for $k < 0$ or $k > L$, the channel impulse response coefficients are $h(k) = 0$ and $h(0) \neq 0$. Also, we assume that the $N \times N$ matrix $\mathbf{F}^{(1)}$ is invertible. It is not hard to verify that the matrix $\left(\mathbf{H}^{(1)}\right)^H \mathbf{H}^{(1)}$ is an $N \times N$ Toeplitz matrix, whose entries consist of samples of the autocorrelation function of the channel impulse response.

B. Precoded OFDM system

Figure 2 depicts an OFDM system precoded by a tall $M \times N$ matrix $\mathbf{F}^{(2)}$, see [42], [20]. The serially transmitted symbols are first converted into a signal vector $\mathbf{x}$ of size N. This signal vector is multiplied by the precoder matrix $\mathbf{F}^{(2)}$ to produce another signal vector $\mathbf{x}'$ of size $M = N + K$. The inverse fast Fourier transform (IFFT) is performed on $\mathbf{x}'$. Finally, L prefix symbols are padded to the output of the IFFT to form a new signal vector $\tilde{\mathbf{x}}$ of size $M + L - 1$ for parallel transmission. The purpose of introducing redundancy to $\mathbf{x}$ is to protect the signal from channel nulls. Therefore, a requirement is that K be greater than or equal to the number of nulls on the interval $[0\ 2\pi)$ of the frequency response $H(\omega) = \sum_{k=0}^{L-1} h(k)e^{-jk\omega}$, where $h(k)$ is the channel's k-th impulse response coefficient. The purpose of padding the prefix to $\tilde{\mathbf{x}}$ is to transform the original Toeplitz intersymbol interference (ISI) channel matrix into a cyclic matrix, which is diagonalized by the fast Fourier transform (FFT). Hence, the prefix is discarded at the receiver before the FFT is implemented. Consequently, the ISI channel with additive white Gaussian noise is transformed into parallel ISI-free sub-channels, each with gain equal to the channel's frequency response value at the corresponding FFT bin. The result of these operations is a discrete-time channel model (for details see [42], [20])

$$\mathbf{r} = \mathbf{H}^{(2)}\mathbf{F}^{(2)}\mathbf{x} + \boldsymbol{\xi}. \tag{42}$$

Here, $\boldsymbol{\xi}$ is a white Gaussian noise vector and $\mathbf{H}^{(2)} = diag\left(H_0, H_1, \cdots, H_{M-1}\right)$ is the channel matrix with $H_k = H(\omega)|_{\omega = \frac{2k\pi}{M}}, k = 0, 1, \cdots, M-1$.

C. Unified model

Models (41) and (42) are both precoded parallel transmission models. They are unified into a single model as

$$\mathbf{r} = \mathbf{H}\mathbf{F}\mathbf{x} + \boldsymbol{\xi}, \tag{43}$$

where $\mathbf{H}$ is the channel matrix, $\mathbf{F}$ is the precoder, $\mathbf{x}$ is the transmitted vector and $\boldsymbol{\xi}$ is a white Gaussian noise vector.

V. PRECODER DESIGN FOR MINIMIZING THE PROBABILITY OF BLOCK ERROR

The purpose of this section is to design the precoder matrix such that the block error probability of the QR successive cancellation detector is minimized. We assume that the product $\mathbf{H}^H\mathbf{H}$ is known to the transmitter and that the channel $\mathbf{H}$ is known to the receiver. Also, we assume that the elements of $\mathbf{x}$ are independent identically distributed (i.i.d.) equally likely binary phase-shift keying (BPSK) signal points (the results are straightforwardly extended to quaternary phase-shift keying (QPSK) signal constellations). Extensions to other constellations are briefly discussed at the end of the section.

A. Problem formulation

First, we derive an expression for $P_c(\mathbf{F})$, the probability of correctly detecting a block using the QR successive cancellation detector. By the chain rule, we have

$$\begin{aligned} P_c(\mathbf{F}) &= Pr(x_1^c, x_2^c, \cdots, x_N^c) \\ &= \prod_{k=1}^{N} Pr(x_k^c | x_{k+1}^c, \cdots, x_N^c), \end{aligned} \quad (44)$$

where x_k^c denotes the *event* that the k-th detected symbol is correct, i.e., $x_k = \hat{x}_k$. If x_k^e denotes the event that the k-th symbol is not detected correctly, then (44) can be rewritten as

$$P_c(\mathbf{F}) = \prod_{k=1}^{N} \left\{1 - Pr(x_k^e | x_{k+1}^c, \cdots, x_N^c)\right\}. \quad (45)$$

For the model in (43), we deduce (see, e.g., [38]) that

$$Pr(x_k^e | x_{k+1}^c, \cdots, x_N^c) = Q\left(\frac{[\mathbf{R}]_k}{\sigma}\right), \quad (46)$$

where $Q(t) = 1/\sqrt{2\pi}\int_t^{+\infty} exp(-z^2/2)dz$, and (from Lemma 1)

$$[\mathbf{R}]_k = \sqrt{\frac{det\left(\mathbf{F}_k^H\mathbf{H}^H\mathbf{H}\mathbf{F}_k\right)}{det\left(\mathbf{F}_{k-1}^H\mathbf{H}^H\mathbf{H}\mathbf{F}_{k-1}\right)}} \quad (47)$$

is the k-th diagonal entry in the R-factor of the QR decomposition of $\mathbf{HF}$. Substituting (46) into (45) yields

$$P_c(\mathbf{F}) = \prod_{k=1}^{N}\left(1 - Q\left(\frac{[\mathbf{R}]_k}{\sigma}\right)\right). \quad (48)$$

Our problem is now stated as:

Problem 1: Let $P > 0$ be the power budget. Find the precoder matrix $\mathbf{F}$ that minimizes the block error probability of the QR successive cancellation detector, subject to the power constraint, $\mathrm{tr}\left(\mathbf{F}^H\mathbf{F}\right) \le P$. More precisely, it is formulated as the following equivalent optimization problems,

$$\mathbf{F}^\star = \arg \min_{\mathrm{tr}(\mathbf{F}^H\mathbf{F}) \le P} P_e(\mathbf{F}),$$

$$\mathbf{F}^\star = \arg \max_{\mathrm{tr}(\mathbf{F}^H\mathbf{F}) \le P} P_c(\mathbf{F}),$$

where $P_e(\mathbf{F}) = 1 - P_c(\mathbf{F})$. ■

In the following we show that under a mild SNR constraint, Problem 1 is equivalent to finding a precoder whose superchannel matrix $\mathbf{HF}$ exhibits equal diagonal entries in the R-factor. In other words, we will show that the optimal precoder $\mathbf{F}^\star$ is the S-factor in the QRS decomposition of $\mathbf{H}$ (up to a scaling constant).

Let $f(t) = Q\left(\frac{1}{\sigma\sqrt{t}}\right)$ for $t > 0$. The second derivative of $f(t)$ is

$$\frac{d^2 f(t)}{d^2 t} = \frac{1}{\sqrt{2\pi\sigma^2 t^5}}\left(\frac{1}{2\sigma^2 t} - \frac{3}{2}\right)\exp\left(-\frac{1}{2\sigma^2 t}\right).$$

If $t < 1/(3\sigma^2)$, then $d^2 f(t)/d^2 t > 0$, and thus, in this case, $f(t)$ is a convex function. In addition, we define $\mathcal{F}(t) = \ln\left[1 - f(t)\right]$. The second derivative of $\mathcal{F}(t)$ is

$$\frac{d^2\mathcal{F}(t)}{d^2 t} = -\frac{\frac{d^2 f(t)}{d^2 t}\left(1 - f(t)\right) + \left(\frac{df(t)}{dt}\right)^2}{\left(1 - f(t)\right)^2}.$$

If $t < 1/(3\sigma^2)$, then $d^2 f(t)/d^2 t > 0$, and as a result, $d^2\mathcal{F}(t)/d^2 t < 0$, since $f(t) < 1$. Hence, for small values of t, the function $\mathcal{F}(t)$ is concave. Therefore, if

$$\frac{[\mathbf{R}]_k^2}{\sigma^2} \ge 3 \qquad \text{for } k = 1, 2, \cdots, N, \quad (49)$$

then, by combining the concavity of $\mathcal{F}(t)$ and (48), we get (Jenssen's inequality [43])

$$\ln P_c(\mathbf{F}) = \sum_{k=1}^{N}\mathcal{F}([\mathbf{R}]_k^{-2}) \le N \cdot \mathcal{F}\left(\frac{\sum_{k=1}^{N}[\mathbf{R}]_k^{-2}}{N}\right). \quad (50)$$

Due to the concavity of $\mathcal{F}(t)$, the equality in (50) holds if and only if

$$[\mathbf{R}]_1 = [\mathbf{R}]_2 = \cdots = [\mathbf{R}]_N. \quad (51)$$

On the other hand, from the general relationship between the arithmetic and the geometric mean, we get

$$\frac{\sum_{k=1}^{N}[\mathbf{R}]_k^{-2}}{N} \ge \left(\prod_{k=1}^{N}[\mathbf{R}]_k^{-2}\right)^{1/N} = \frac{1}{\left(det\left(\mathbf{F}^H\mathbf{H}^H\mathbf{H}\mathbf{F}\right)\right)^{1/N}}, \quad (52)$$

where we have utilized Lemma 1 to obtain the last equality in (52). Fortunately, the equality in (52) holds if and only if (51) also holds. So, the equalities in (50) and (52) both hold under the same condition (51). Combining (50) with (52), and using the fact that $\mathcal{F}(t)$ is monotonically decreasing, yields

$$\ln P_c(\mathbf{F}) \le N \ln\left(1 - Q\left(\frac{\left(det(\mathbf{F}^H\mathbf{H}^H\mathbf{H}\mathbf{F})\right)^{1/2N}}{\sigma}\right)\right). \quad (53)$$

The equality in (53) holds if and only if (51) holds. Therefore, we consider the following optimization problem that maximizes the upper bound in (53).

Formulation 1: Find $\breve{\mathbf{F}}$ such that

$$\breve{\mathbf{F}} = \arg \max_{\mathbf{F}} det\left(\mathbf{F}^H\mathbf{H}^H\mathbf{H}\mathbf{F}\right), \quad (54)$$

where the maximum in (54) is taken subject to the power constraint,

$$\mathrm{tr}\left(\mathbf{F}^H\mathbf{F}\right) \le P. \quad (55)$$

B. Problem solution

We solve the problem in Formulation 1 for two cases.

Case 1 $\mathbf{H} = \mathbf{H}^{(1)}$, see Section IV-A. In this case, $\mathbf{F} = \mathbf{F}^{(1)}$ is a square matrix and $det\left(\mathbf{F}^H\mathbf{H}^H\mathbf{H}\mathbf{F}\right) = det\left(\mathbf{H}^H\mathbf{H}\right) det\left(\mathbf{F}^H\mathbf{F}\right)$. Thus, maximizing $det\left(\mathbf{F}^H\mathbf{H}^H\mathbf{H}\mathbf{F}\right)$ over all choices of the matrix $\mathbf{F}$ is equivalent to maximizing $det(\mathbf{F}^H\mathbf{F})$ over the matrix $\mathbf{F}$. Applying Hadamard's inequality [1] leads to

$$det\left(\mathbf{F}^H\mathbf{F}\right) \leq \prod_{k=1}^{N} \mathbf{f}_k^H\mathbf{f}_k, \tag{56}$$

where $\mathbf{f}_k$ is the k-th column of $\mathbf{F}$. The equality in (56) holds if and only if $\mathbf{F}^H\mathbf{F}$ is a diagonal matrix. Furthermore, under the power constraint $\mathrm{tr}(\mathbf{F}^H\mathbf{F}) \leq P$, we have

$$\prod_{k=1}^{N} \mathbf{f}_k^H\mathbf{f}_k \leq \left(\frac{\sum_{k=1}^{N}\mathbf{f}_k^H\mathbf{f}_k}{N}\right)^N \leq \left(\frac{P}{N}\right)^N. \tag{57}$$

Both equalities in (56) and (57) hold if and only if the diagonal entries of $\mathbf{F}^H\mathbf{F}$ are equal, i.e., if and only if

$$\mathbf{f}_1^H\mathbf{f}_1 = \mathbf{f}_2^H\mathbf{f}_2 = \cdots = \mathbf{f}_N^H\mathbf{f}_N = \frac{P}{N}. \tag{58}$$

Hence, in this case, we easily obtain the solution of Formulation 1,

$$\breve{\mathbf{F}}^{(1)} = \sqrt{\frac{P}{N}}\breve{\mathbf{W}}^{(1)}, \tag{59}$$

where $\breve{\mathbf{W}}^{(1)}$ is an arbitrary $N \times N$ unitary matrix. ■

Case 2 $\mathbf{H} = \mathbf{H}^{(2)}$, see Section IV-B. In this case, the Lagrangian formulation [44] of this problem admits a closed-form analytic solution $\mathbf{F} = \breve{\mathbf{F}}^{(2)}$. The Lagrangian here is

$$\mathcal{L}(\mathbf{F}) = \ln\left(det\left(\mathbf{F}^H\mathbf{H}^H\mathbf{H}\mathbf{F}\right)\right) - \mu\left(tr\left(\mathbf{F}^H\mathbf{F}\right) - P\right),$$

where μ is the Lagrange multiplier. The necessary condition for achieving the maximum is that the gradient of $\mathcal{L}(\mathbf{F})$ with respect to $\mathbf{F}$ equals to zero. After appropriate matrix manipulations [44], this condition yields the following equivalent condition:

$$\mathbf{H}^H\mathbf{H}\mathbf{F}\left(\mathbf{F}^H\mathbf{H}^H\mathbf{H}\mathbf{F}\right)^{-1} - \mu\mathbf{F} = 0. \tag{60}$$

Left multiplying the two sides of equation (60) by $\mathbf{F}^H$ yields $\mathbf{F}^H\mathbf{F} = \mu^{-1}\mathbf{I}$, which shows that the optimal solution $\mathbf{F}$ must have orthogonal columns. Considering the power constraint (55), we have that $\mu \geq \frac{N}{P}$. Therefore, in this case, the solution of Formulation 1 is

$$\breve{\mathbf{F}}^{(2)} = \sqrt{\frac{P}{N}}\breve{\mathbf{P}}^{(2)}\breve{\mathbf{W}}^{(2)}. \tag{61}$$

Here, $\breve{\mathbf{W}}^{(2)}$ is an arbitrary $N \times N$ unitary matrix, and $\breve{\mathbf{P}}^{(2)}$ is an $M \times N$ *selection matrix* such that $\mathbf{H}^{(2)}\breve{\mathbf{P}}^{(2)}$ consists of those N columns of $\mathbf{H}^{(2)}$ that have the largest N magnitudes (note $\mathbf{H}^{(2)}$ is a diagonal matrix). ■

Regardless which case we consider (Case 1 or Case 2), the optimal solution of Formulation 1 can always be expressed as

$$\breve{\mathbf{F}} = \sqrt{\frac{P}{N}}\breve{\mathbf{P}}\breve{\mathbf{W}}, \tag{62}$$

where $\breve{\mathbf{P}}$ is an $N \times N$ identity matrix in Case 1 and $\breve{\mathbf{P}} = \breve{\mathbf{P}}^{(2)}$ in Case 2. In the remainder of this section we show that the optimal choice for $\breve{\mathbf{W}}$ is

$$\breve{\mathbf{W}} = \mathbf{S}, \tag{63}$$

that is, the unitary factor $\breve{\mathbf{W}}$ of the optimal precoder is exactly the S-factor in the QRS decomposition of the superchannel matrix $\mathbf{H}\breve{\mathbf{P}}$.

At this point, a natural question is when is the upper bound in (53) achieved? If we desire that the maximum probability of correct decisions

$$P_c(\mathbf{F}) = \left(1 - Q\left(\frac{\left(det(\breve{\mathbf{F}}^H\mathbf{H}^H\mathbf{H}\breve{\mathbf{F}})\right)^{1/2N}}{\sigma}\right)\right)^N \tag{64}$$

be attainable, equations (51) must hold. In other words, the matrix $\mathbf{H}\breve{\mathbf{F}}$ must have an equal-diagonal R-factor. Luckily, in the solution $\breve{\mathbf{F}}$ in (62), we have the freedom of choosing $\breve{\mathbf{W}}$ to meet constraint (51). An obvious choice is therefore to set

$$\breve{\mathbf{W}} = \mathbf{S}, \tag{65}$$

where $\mathbf{S}$ is the S-factor in the QRS decomposition of $\sqrt{P/N}\mathbf{H}\breve{\mathbf{P}}$, i.e., $\left(\sqrt{P/N}\mathbf{H}\breve{\mathbf{P}}\right)\mathbf{S} = \mathbf{Q}\mathbf{R}$. The factor $\mathbf{S}$ is computable using Algorithm 5. The precoder

$$\mathbf{F}^{\star} = \sqrt{\frac{P}{N}}\breve{\mathbf{P}}\mathbf{S} \tag{66}$$

is also the solution of Problem 1 if the constraint (49) holds as well. For the precoder matrix given in (66), the diagonal entries $[\mathbf{R}]_k$ are all equal, and an equivalent formulation of constraint (49) is

$$\begin{aligned}
\left(\frac{P}{N}\right)^N det\left(\breve{\mathbf{P}}\mathbf{H}^H\mathbf{H}\breve{\mathbf{P}}\right) &= det\left(\left(\mathbf{F}^{\star}\right)^H\mathbf{H}^H\mathbf{H}\mathbf{F}^{\star}\right) \\
&= \prod_{k=1}^{N}[\mathbf{R}]_k^2 = \left([\mathbf{R}]_1^2\right)^N \\
&\geq \left(3\sigma^2\right)^N,
\end{aligned}$$

which leads to

$$\frac{P}{N\sigma^2} \geq \frac{3}{det\left(\breve{\mathbf{P}}^H\mathbf{H}^H\mathbf{H}\breve{\mathbf{P}}\right)^{1/N}}. \tag{67}$$

It is now clear that constraint (49), or equivalently (67), is a constraint on SNR since P is the power budget, σ^2 is the noise variance and $P/\left(N\sigma^2\right)$ is the SNR. The next theorem (Theorem 6) summarizes these results

Theorem 6: Let the SNR constraint (67) hold. Let $\breve{\mathbf{P}} = \mathbf{I}$ if our model is (41); otherwise if our model is (42), let $\breve{\mathbf{P}} = \breve{\mathbf{P}}^{(2)}$ be an $M \times N$ selection matrix such that $\mathbf{H}^{(2)}\breve{\mathbf{P}}^{(2)}$ consists of those N columns of $\mathbf{H}^{(2)}$ with the largest N magnitudes. Denote by $\mathbf{F}^{\star}$ the solution to Problem 1. Let the QRS decomposition of $\sqrt{P/N}\mathbf{H}\breve{\mathbf{P}}$ be $\left(\sqrt{P/N}\mathbf{H}\breve{\mathbf{P}}\right)\mathbf{S} = \mathbf{Q}\mathbf{R}$, where $\mathbf{S}$ can be computed by Algorithm 5. Then, we have:

1) The optimal solution of Problem 1, $\mathbf{F}^\star$, has the following structure
$$\mathbf{F}^\star = \sqrt{\frac{P}{N}}\breve{\mathbf{P}}\mathbf{S}. \quad (68)$$
2) The QR decomposition of $\mathbf{HF}^\star$ exhibits an equal-diagonal R-factor.
3) For the optimal precoder $\mathbf{F}^\star$, the block error probability of the QR successive-cancellation detector is minimized and equals
$$P_e^{(\min)} = 1 - \left(1 - Q\left(\sqrt{snr}\cdot det\left(\breve{\mathbf{P}}^H\mathbf{H}^H\mathbf{H}\breve{\mathbf{P}}\right)^{1/2N}\right)\right)^N, \quad (69)$$
where $snr = \frac{P}{N\sigma^2}$.
4) The free distance of the precoded channel matrix $\mathbf{HF}^\star$ is
$$d_{\text{free}}(\mathbf{HF}^\star) = \sqrt{\frac{P}{N}}\left(det\left(\breve{\mathbf{P}}^H\mathbf{H}^H\mathbf{H}\breve{\mathbf{P}}\right)\right)^{1/2N} d_{\min}(\mathcal{X}). \quad (70)$$
■

Proof: Most of the proof has already been completed through the derivation of equations (56)–(67). Statement 1 follows from the derivation of equations (62)–(66). Statement 2 is triviality satisfied by the construction of $\mathbf{F}^\star$ in Statement 1. To prove statement 3, we note that by Lemma 1 and Theorem 1 an equal-diagonal R-factor implies $[\mathbf{R}]_k = \sqrt{P/N}det(\breve{\mathbf{P}}^H\mathbf{H}^H\mathbf{H}\breve{\mathbf{P}})^{1/2N}$ for $k = 1, \cdots, N$, and the result follows directly from (48). Statement 4 is proved using Property 1 in Section III-D. □

Remarks: To understand the physical meaning of the solution $\mathbf{F}^\star$, we make the following comments.

- Condition (67) is an SNR condition. When it is satisfied, i.e., when the SNR is relatively high, we can obtain our optimal solution. Otherwise, when the SNR is low, we do not know how to solve the problem. In fact, condition (67) is equivalent to
$$snr(dB) \geq 10log3 - \frac{10}{N}\sum_{k=1}^{N} log\lambda_k, \quad (71)$$
where λ_k are the eigenvalues of the matrix $\breve{\mathbf{P}}^H\mathbf{H}^H\mathbf{H}\breve{\mathbf{P}}$. For N large, $\frac{1}{N}\sum\limits_{k=1}^{N} log\lambda_k$ tends to $\frac{1}{2\pi}\int\limits_{-\pi}^{\pi} log|H(\omega)|^2 d\omega$. Hence, using (71), condition (67) becomes
$$snr(dB) \geq 6\left(h[3] - h\left[|H(\omega)|^2\right]\right),$$
where using Kolmogorov's result ([43], p. 274), we denote the entropy rate of a Gaussian process with power spectral density $S(\omega)$ by $h[S(\omega)] = \frac{1}{4\pi}\int_{-\pi}^{\pi} log_2\left(2\pi e S(\omega)\right) d\omega$.
- From the structure of the optimal precoder $\mathbf{F}^\star$ we see that $\breve{\mathbf{P}}$ selects N good sub-channels from the existing M sub-channels. The role of $\mathbf{S}$ is to shape the precoded channel matrix so as to equalize all the diagonal entries of the upper triangular matrix $\mathbf{R}$ in the QR decomposition. As a result, the conditional error probability of each symbol is equal, which minimizes the block error probability of the QR successive cancellation detector.
- At high SNRs, the error probability of the maximum likelihood detector is dominated by the free distance. A comparison of (69) to (70) shows that the detection performance of the QR successive cancellation detector with the optimal precoder $\mathbf{F}^\star$ is asymptotically equivalent to that of the MLD when the SNRs are large.

C. Sketch of extensions to non-binary constellations

In this subsection, we briefly outline the procedure of how to extend the results in this paper to some frequently-used non-binary constellations. The key to generalizing this work to PAM, PSK and QAM constellations is to evaluate the symbol error probability, see Simon and Alouini [45]. In the following $[\mathbf{R}]_k$ is the k-th diagonal entry in the R-factor of the QR-decomposition.

PAM signals: The symbol error probability for q-ary PAM symbol x_k [45] is
$$\text{P}_{\text{PAM}}(x_k^e | x_{k+1}^c, \cdots, x_N^c) = \frac{2(q-1)}{q} Q\left(\frac{[\mathbf{R}]_k\sqrt{3}}{\sigma\sqrt{q^2-1}}\right).$$

PSK signals: The symbol error probability for q-ary PSK signal x_k [45] is
$$\text{P}_{\text{PSK}}(x_k^e | x_{k+1}^c, \cdots, x_N^c) = \frac{1}{\pi}\int_0^{(q-1)\pi/q} \exp\left(-\frac{[\mathbf{R}]_k^2 \sin^2(\pi/q)}{2\sigma^2\sin^2\theta}\right) d\theta.$$

QAM signals: The symbol error probability for q-ary QAM signal x_k [45] with $q = 2^\ell \times 2^\ell$ is
$$\text{P}_{\text{QAM}}(x_k^e | x_{k+1}^c, \cdots, x_N^c) = 4\left(1 - \frac{1}{\sqrt{q}}\right) Q\left(\frac{[\mathbf{R}]_k\sqrt{3}}{\sigma\sqrt{(q-1)}}\right) - 4\left(1 - \frac{1}{\sqrt{q}}\right)^2 Q^2\left(\frac{[\mathbf{R}]_k\sqrt{3}}{\sigma\sqrt{(q-1)}}\right). \quad (72)$$

For our purpose, let
$$\begin{aligned} f(t) &= Q\left(\sqrt{\frac{a}{t}}\right) \\ g(t) &= \int_0^{(q-1)\pi/q} \exp\left(-\frac{b}{t\sin^2\theta}\right) d\theta \\ h(t) &= Q\left(\sqrt{\frac{c}{t}}\right) - \left(1 - \frac{1}{\sqrt{q}}\right) Q^2\left(\sqrt{\frac{c}{t}}\right), \end{aligned}$$
where $a = \frac{3}{\sigma^2(q^2-1)}, b = \frac{\sin^2(\pi/q)}{2\sigma^2}$ and $c = \frac{3}{\sigma^2(q-1)}$. Let $x = \sqrt{T}$ be the root of following equation,
$$Q(x) + \frac{x}{\sqrt{2\pi}(x^2-3)}\exp\left(-\frac{x^2}{2}\right) - \frac{\sqrt{q}}{2(\sqrt{q}-1)} = 0.$$
Then, $h(t)$ is convex in t for $0 < t \leq \frac{c}{T}$, while it is concave in t for $t > \frac{c}{T}$. The function $f(t)$ is convex in the region $0 < t \leq \frac{a}{3}$, while it is concave in the region $t > \frac{a}{3}$. Similarly, we can get the convex/concave regions for the function $g(t)$. Identifying the concave/convex region is the key to finding formulations analogous to Formulation 1, but for non-binary constellations (see Section IV-A). We do not pursue the details further.

VI. Precoder design for optimizing a free distance criterion

In order to better appreciate the physical meaning of our optimal precoder designed in Section V, in this section we take another point of view. Ideally, an optimal precoder should minimize the detection error probability when MLD is used. But we know that for the MLD, at high SNRs the average probability of error over all blocks is dominated by the free distance term [39]. However, directly maximizing the free distance is too expensive for the complexities of both its design and detection to be affordable. Here, we are interested in designing the precoder that maximizes the lower-bound on the free distance. Utilizing Theorem 5, the lower bound on the free distance is

$$\left[\min_{1\le k\le N}[\mathbf{R}]_k\right]\cdot d_{\min}(\mathcal{X})\le d_{\text{free}}(\mathbf{H}) \tag{73}$$

and our problem may now be stated as:

Problem 2: Find the precoder matrix $\mathbf{F}$ that maximizes the lower bound on the free distance (73), that is, maximize $\min[\mathbf{R}]_k$, subject to the power constraint, $\operatorname{tr}\left(\mathbf{F}^H\mathbf{F}\right)\le P$. More precisely, it is formulated as the following optimization problem,

$$\mathbf{F}^\star=\arg\max_{\operatorname{tr}(\mathbf{F}^H\mathbf{F})\le P}\left[\min_{1\le k\le N}[\mathbf{R}]_k\right],$$

where $[\mathbf{R}]_k$ is determined by equation (47). ■

First we note that

$$\min_{1\le k\le N}[\mathbf{R}]_k\le\left(\prod_{k=1}^{N}[\mathbf{R}]_k^2\right)^{1/2N}=det\left(\mathbf{F}^H\mathbf{H}^H\mathbf{H}\mathbf{F}\right). \tag{74}$$

The equality here holds if and only if

$$[\mathbf{R}]_1=[\mathbf{R}]_2=\cdots=[\mathbf{R}]_N. \tag{75}$$

Therefore, Problem 2 is reduced to first solving the optimization problem in Formulation 1 and then finding a unitary matrix $\breve{\mathbf{W}}$ that enforces condition (75).

Theorem 7: Let $\breve{\mathbf{P}}=\mathbf{I}$ if our model is (41); otherwise if our model is (42), let $\breve{\mathbf{P}}=\breve{\mathbf{P}}^{(2)}$ be an $M\times N$ selection matrix such that $\mathbf{H}^{(2)}\breve{\mathbf{P}}^{(2)}$ consists of those N columns of $\mathbf{H}^{(2)}$ with the largest N magnitudes. Denote the QRS decomposition of $\sqrt{P/N}\mathbf{H}\breve{\mathbf{P}}$ by $\left(\sqrt{P/N}\mathbf{H}\breve{\mathbf{P}}\right)\mathbf{S}=\mathbf{QR}$, where the S-factor $\mathbf{S}$ is computed by Algorithm 5. The solution $\mathbf{F}^\star$ to Problem 2 is

$$\mathbf{F}^\star=\sqrt{\frac{P}{N}}\breve{\mathbf{P}}\mathbf{S}. \tag{76}$$

Moreover, for this optimal precoder, we have

$$d_{\text{free}}(\mathbf{H}\mathbf{F}^\star)=\sqrt{\frac{P}{N}}\left(det\left(\breve{\mathbf{P}}^H\mathbf{H}^H\mathbf{H}\breve{\mathbf{P}}\right)\right)^{1/2N}d_{\min}(\mathcal{X}). \tag{77}$$

■

Proof: Combining (74) and (75) with the proof of Statement 1 in Theorem 6, we can complete the proof of equation (76). Statement (77) is proved using Property 1 in Section III-D. □

The only difference between Theorem 6 and Theorem 7 is that in Theorem 7 the SNR condition (67) need not hold. Essentially, the two theorems are equivalent when the SNR is high. This equivalence is justified as follows. We derive from (48) that

$$\left(1-Q\left(\frac{\min[\mathbf{R}]_k}{\sigma}\right)\right)^N\le P_c(\mathbf{F})\le 1-Q\left(\frac{\min[\mathbf{R}]_k}{\sigma}\right), \tag{78}$$

which, utilizing $P_c(\mathbf{F})=1-P_e(\mathbf{F})$, is equivalent to

$$\begin{aligned}Q\left(\frac{\min[\mathbf{R}]_k}{\sigma}\right)\le P_e(\mathbf{F}) &\le 1-\left(1-Q\left(\frac{\min[\mathbf{R}]_k}{\sigma}\right)\right)^N\\ &\le N\cdot Q\left(\frac{\min[\mathbf{R}]_k}{\sigma}\right).\end{aligned} \tag{79}$$

This shows that when the SNR is high, the minimum diagonal entry in the R-factor of the QR decomposition of the channel matrix determines (up to a constant of proportionality) the block error probability of the QR successive cancellation detector. In other words, the minimal diagonal entry accounts for the dominant term of the QR detector error probability just as the free distance is the dominant term of the MLD error probability. (See Varanasi [6] for an analysis of the performance of the cancellation detector.) Hence, increasing the minimum diagonal entry can decrease the block error probability of the QR successive cancellation detector. Consequently, when the minimum diagonal entry becomes equal to the maximum diagonal entry, i.e., when the precoded channel matrix exhibits an equal-diagonal R-factor, the block error probability reaches its minimum. Using Property 1 in Section III-D, we conclude that for BPSK (and QPSK) signals the detection performance of the QR successive cancellation detector is asymptotically equivalent to that of the maximum likelihood detector when the signal to noise ratio (SNR) is large.

VII. Simulations

A. Comparison of the precoders for the TDMA system

In this subsection we adopt the model in equation (41) and compare our QR-optimal precoder (**QR-OP**) followed by a successive-cancellation detector to other precoders and detectors available in the literature. We compare the following precoders/detectors:

1) The zero-forcing minimum mean square error (**ZF-MMSE**) precoder, followed by a threshold detector. The precoder in this scenario was determined in [21] to be

$$\mathbf{F}=\sqrt{\frac{P}{\operatorname{tr}\left(\mathbf{D}^{1/2}\right)}}\mathbf{W}\mathbf{D}^{1/4}\mathbf{U}, \tag{80}$$

where $\left((\mathbf{H}^{(1)})^H\mathbf{H}^{(1)}\right)^{-1}=\mathbf{W}\mathbf{D}\mathbf{W}^H$ is the eigen decomposition of $\left((\mathbf{H}^{(1)})^H\mathbf{H}^{(1)}\right)^{-1}$ and $\mathbf{U}$ is an arbitrary unitary matrix. In our simulation, $\mathbf{U}$ is taken as the identity matrix $\mathbf{I}$.

2) The minimum bit error rate zero-forcing precoder (**MBER-ZF**), followed by a threshold detector. In [37], it was shown that this precoder has the same structure as

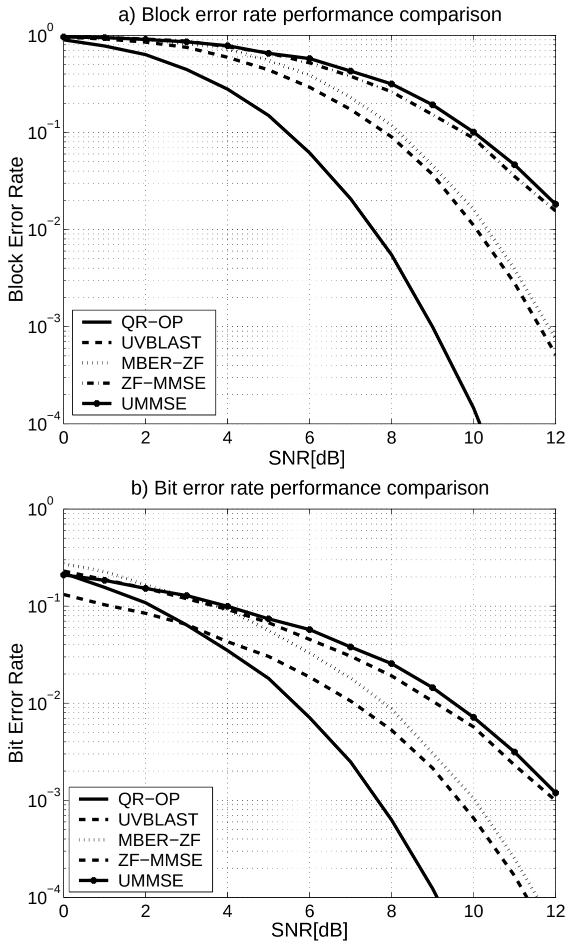

Fig. 3. Performance comparison of the optimal QR precoder with other precoders for Example 1. (a) Block error rate performance comparison. (b) Bit error rate performance comparison.

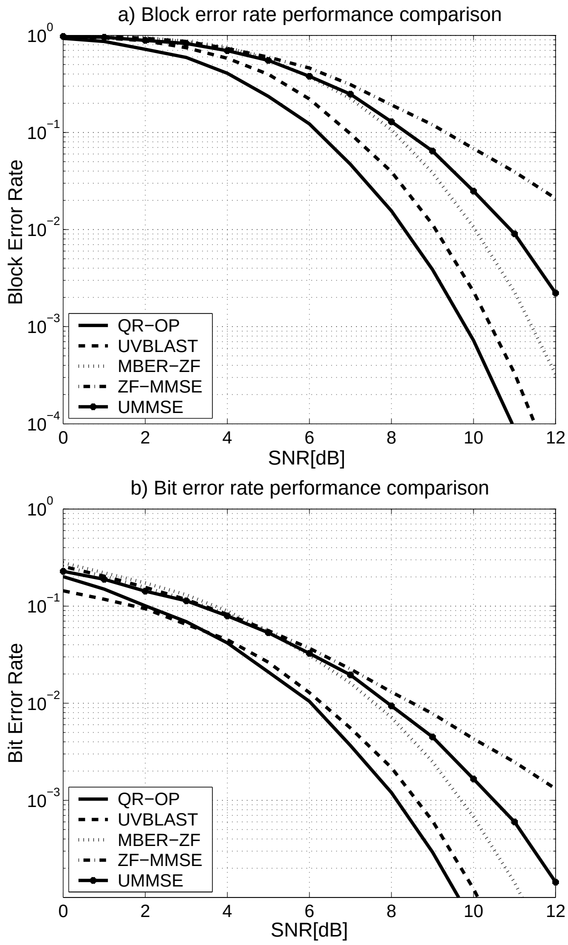

Fig. 4. Performance comparison of the optimal QR precoder with other precoders for Example 2. (a) Block error rate performance comparison. (b) Bit error rate performance comparison.

in (80), where the unitary matrix $\mathbf{U} = \mathbf{U}_{opt}$ is the normalized $N \times N$ the discrete Fourier transform matrix. If N is a power of 2, $\mathbf{U}_{opt}$ can be the normalized Hadamard matrix.

3) The unprecoded mean square error detector [46] (**UMMSE**).
4) The unprecoded V-BLAST detector [18] (**UVBLAST**).

We compare the performance of these precoders to the block error rate for the QR-optimal precoder given in equation (68). Note that some of these comparisons are not really fair. For example, it is not fair to compare precoded to unprecoded detectors. Similarly, it is not fair to compare the V-BLAST detector (which operates under the assumption that the transmitter does not know the channel coefficients) to other precoders that utilize the knowledge of the channel coefficients at the transmitter. Nonetheless, we still conduct the comparisons because they give insight into the performances of methods available in the literature.

Example 1:. We consider a finite impulse response (FIR) channel of order $L = 3$, whose impulse response coefficients are $h(0) = 0.0838 - 0.0633j, h(1) = 0.6028j, h(2) = -0.8709 - 0.7320j, h(3) = 1$. The channel has $L = 3$ zeros at $0.7, 0.5\exp(\mathrm{j}2\pi 0.256)$ and $0.3\exp(\mathrm{j}2\pi 0.141)$. The block length of the precoder was chosen to be $N = 32$, resulting in $M = N + L = 35$. Simulated *block* error rates and *bit* error rates are shown in Figures 3 (a) and (b), respectively. We observe that our QR-optimal precoder (**QR-OP**) performs about 3dB better than best precoder/detector known.

Example 2:. The scenario of this example is similar to that of Example 1. An FIR channel of order $L = 4$ is considered, where channel response coefficients are $h(0) = -0.7920, h(1) = 0.0523 - 0.1580j, h(2) = 0.2526 - 0.0395j, h(3) = -0.5129 + 0.1975j, h(4) = 1$. The zeros of this channel response are $z_1 = 1, z_2 = 0.9\exp(\mathrm{j}9\pi/20), z_3 = 1.1\exp(-\mathrm{j}9\pi/20), z_4 = -0.8$. Notice that this channel has a spectral null at zero-frequency (since $z_1 = 1$), while the other roots are also very close to the unit circle. For this channel, the block length was chosen to be $N = 32$ and $M = N + L = 35$. The block and bit error rates are plotted in Figures 4 (a) and (b), respectively, showing that the QR-optimal precoder (**QR-OP**) outperforms the minimum bit error zero-forcing precoder (**MBER-ZF**) by 2dB and outperforms the unprecoded V-BLAST detector (**UVBLAST**) by 0.5 dB.

B. Comparison of the precoders for the OFDM system

In this subsection, we compare the QR-optimal precoder (**QR-OP**) to other known precoder/detectors for the OFDM system. Thereby, we adopt the transmission model

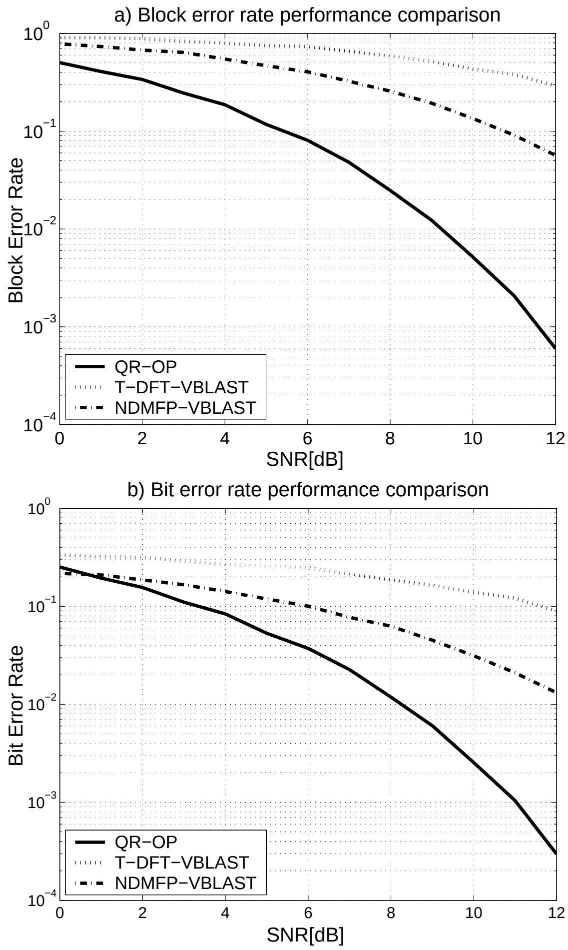

Fig. 5. Performance comparison of the optimal QR precoder with other precoders for Example 3. (a) Block error rate performance comparison. (b) Bit error rate performance comparison.

given in equation (42). We compare the following precoders/detectors

1) The truncated DFT precoder [42], followed by the V-BLAST detector [18], which we abbreviate by (**T-DFT-VBLAST**).
2) The non-maximally decimated multirate filterbank precoder (**NDMFP**) [20], followed by the V-BLAST detector [18],which we abbreviate by (**NDMFP-VBLAST**).

We also show comparisons to the theoretical performance of the QR-optimal precoder followed by the successive cancellation detector, given in equation (69).

Example 3:. An FIR channel of order $L = 4$ is assumed. The channel coefficients are $h(0) = 0.1000 + 0.1732j, h(1) = -0.5550 - 0.5550j, h(2) = 1.0131 + 0.5849j, h(3) = -0.7581 - 0.2031j, h(4) = 0.2000$ with zeros $z_k = \exp(j2\pi k/M), k = 0, 1, 2, 3$. For this scenario, we chose $N = 8, K = 4$ and $M = N + K = 12$. The block and bit error rate simulations are given in Fig. 5(a) and Fig. 5(b), respectively.

Example 4:. The scenario in this example is a bit more realistic than Example 3 in that we choose a longer block length $N = 32$. We choose an FIR channel of order $L = 4$ with channel coefficients $h(0) = -0.1895 + 0.1432j, h(1) = 0.3535-1.1414j, h(2) = 0.6179+2.2021j, h(3) = -1.7115-1.2342j, h(4) = 0.9428$, and channel zeros at $z_1 = 0.9, z_2 = 0.7, z_3 = \exp(j2\pi 0.256), z_4 = 0.4\exp(j2\pi 0.141)$. We take $K = 1$, to get $M = N + K = 35$. The block and bit error rates are plotted in Figures 6(a) and 6(b), respectively.

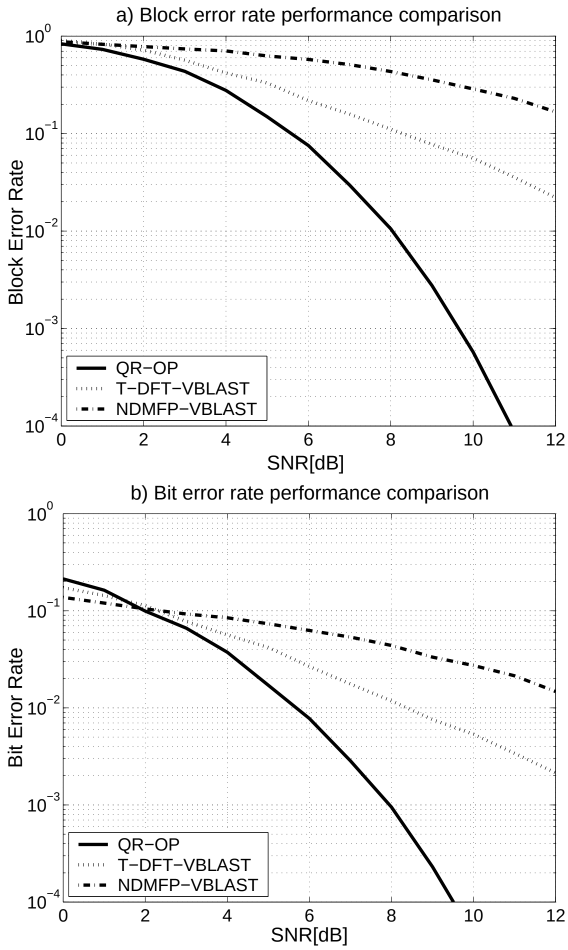

Fig. 6. Performance comparison of the optimal QR precoder with other precoders for Example 4. (a) Block error rate performance comparison. (b) Bit error rate performance comparison.

C. *Comparison of the optimal QR detector with MLD*

In the discussion that followed equation (79), we asserted that the block error rate of the QR-optimal precoder followed by the successive cancellation detector is asymptotically (as SNR$\rightarrow \infty$) equivalent (up to a constant of proportionality) to the block error rate of the maximum likelihood detector when using the same precoder. In this subsection we test this assertion through a simulation as well as by an upper-bound error-rate analysis.

Example 5:. The scenario of this example is similar to Example 2, with the only difference being that we now choose the block length to be smaller ($N = 8$) so that we can actually simulate the maximum likelihood detector. We determine the QR-optimal precoder and simulate the performance of two detectors for this precoder: 1) The successive-cancellation detector, and 2) the maximum likelihood detector, which can be efficiently implemented using the sphere decoder if the radius

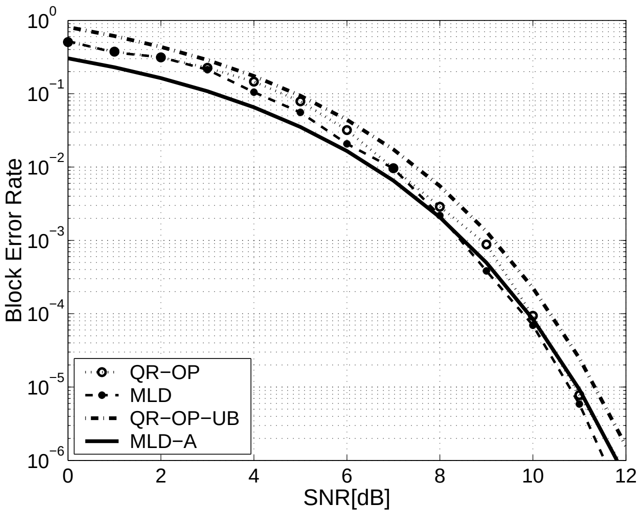

Fig. 7. Performance comparison of the optimal QR detector with MLD for the optimally precoded channel under the scenario in Example 5.

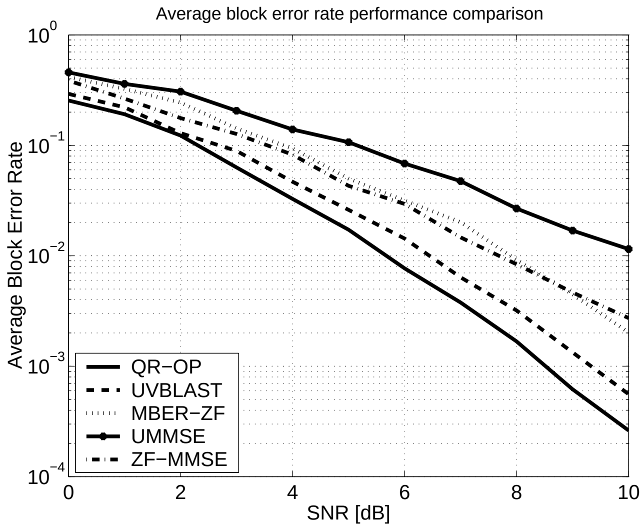

Fig. 8. Comparison of the successive cancellation detector with optimal precoding (QR-OP) to other precoders in a stochastic channel scenario.

is properly set (see [47] and the references therein). The results are shown in Figure 7. Notice that at very high SNRs, the two curves (QR-OP and MLD) are parallel, which confirms the assertion that the two error rates differ only by a constant of proportionality. Also, notice that there is a gap of about 0.3dB between the two curves at error rates of 10^{-6}. We next propose a quick method to analytically estimate this gap.

It is not hard to show that the union bound of the probability of error of the maximum likelihood detector at high SNRs is approximated by (we omit the derivation)

$$\begin{aligned} P_e^{(\text{MLD-A})}(\mathbf{F}^\star) &\approx \kappa \cdot Q\left(\frac{d_{\text{free}}(\mathbf{H}\mathbf{F}^\star)}{2\sigma}\right) \\ &= \kappa \cdot Q\left(\sqrt{snr} \cdot det\left(\mathbf{H}^H\mathbf{H}\right)^{1/2N}\right), \end{aligned} \tag{81}$$

where κ is the number of weakest subchannels in the optimally precoded channel $\mathbf{H}\mathbf{F}^\star$. On the other hand, we have from (79) that the probability of error of the successive cancellation detector (SCD) for the QR-optimal precoder is

$$\begin{aligned} P_e^{(\text{QR-OP})}(\mathbf{F}^\star) &\leq N \cdot Q\left(\sqrt{snr} \cdot det\left(\mathbf{H}^H\mathbf{H}\right)^{1/2N}\right) \\ &= P_e^{(\text{QR-OP-UB})}. \end{aligned} \tag{82}$$

Assuming that the equalities in (81) and (82) hold, and further requiring that $P_e^{(\text{QR-OP})} = P_e^{(\text{MLD})}$, we can compute that the SNR difference between the two detectors is approximated by

$$\begin{aligned} &\text{SNR}_{(\text{QR-OP})}[\text{dB}] - \text{SNR}_{(\text{MLD})}[\text{dB}] \\ &\approx \text{SNR}_{(\text{QR-OP-UB})}[\text{dB}] - \text{SNR}_{(\text{MLD-A})}[\text{dB}] \\ &= 20\log_{10}\frac{Q^{-1}\left(\frac{P_e}{\kappa}\right)}{Q^{-1}\left(\frac{P_e}{N}\right)}, \end{aligned} \tag{83}$$

where $Q^{-1}(\cdot)$ is the inverse of the $Q(\cdot)$ function, i.e., $Q^{-1}(Q(x)) = x$. Now, for our Example 5, we have $N = 8$ and $\kappa = 3$. If we now substitute $P_e = 10^{-6}$ into (83), we get the SNR gap to be $\text{SNR}_{(\text{QR-OP})}[\text{dB}] - \text{SNR}_{(\text{MLD})}[\text{dB}] \approx \text{SNR}_{(\text{QR-OP-UB})}[\text{dB}] - \text{SNR}_{(\text{MLD-A})}[\text{dB}] = 0.3205\,\text{dB}$, which agrees with the simulation results shown in Figure 7. Note that in Figure 7, the notation QR-OP-UB denotes the upper bound of the optimal QR detector on the right hand side of (82) and MLD-A denotes the approximation of the MLD in (81).

D. Comparison in stochastic channel scenarios

The method derived in this paper applies only to channels that are simultaneously known to the transmitter and the receiver. Often the channel is stochastic, i.e., the channel is a realization of a random process (e.g., fading channels). We compare the performance of our precoded successive-cancellation detector to other detectors in stochastic channel scenarios. Thereby, we have to make the assumption that both the transmitter and the receiver instantaneously receive the knowledge of the channel realization immediately after the realization occurs.

Example 6: The channel model was chosen as a TDMA channel model, which corresponds to the scenario in Example 1, but the channel coefficients are complex random i.i.d. white Gaussian variables with unit variance. The precoder and detector performance comparison is shown in Figure 8. The results show that the average performance of the precoder proposed here outperforms the best precoder-based cancellation detector known in the literature by about 1dB.

Example 7: The channel model is a MIMO flat fading channel with four transmitter and six receiver antennas. The channel coefficients are complex random i.i.d. white Gaussian variables with unit variance. In Figure 9, the performance of our precoder-based successive cancellation detector (QR-OP) is compared to the performances of the maximum likelihood detector without precoding (MLD-UP), and the maximum likelihood detector after our optimal precoder is applied at transmitter (MLD-OP). Interestingly, the QR-OP detector outperforms

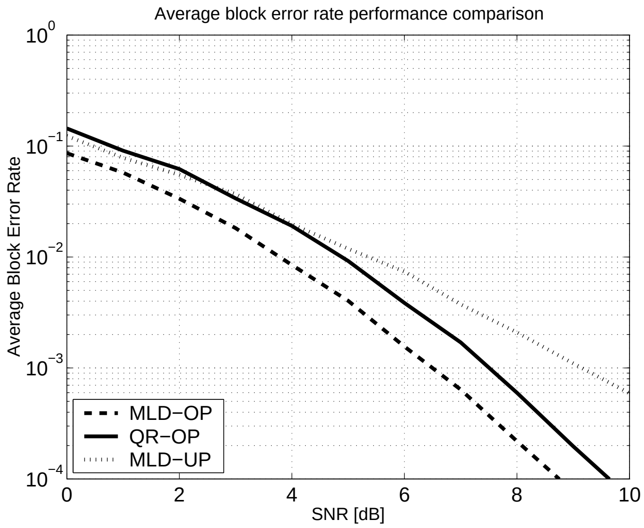

Fig. 9. Comparison of the successive cancellation detector with optimal precoding (QR-OP) to the maximum likelihood detector that uses the same precoder (MLD-OP) and the maximum likelihood detector that uses no precoder (MLD-UP) in a stochastic MIMO channel scenario.

the MLD-UP detector, which shows that there are advantages to applying the precoder. Naturally, the MLD-OP detector outperforms the QR-OP detector, but asymptotically they are equivalent as seen by their equal slopes in Figure 9, which was of course predicted by our (crude) analysis in Sections VI and VII-C. This suggests that the derived optimal precoder for cancellation detection, may be optimal even if used in conjunction with the MLD.

VIII. Conclusion and Discussion

In this paper, we proposed the QRS decomposition of a matrix $\mathbf{H}$. The decomposition has the following form $\mathbf{HS} = \mathbf{QR}$, where $\mathbf{S}$ is a unitary matrix, and $\mathbf{Q}$ and $\mathbf{R}$ are the factors of the QR-decomposition [1], [2] of the matrix $\mathbf{HS}$. The special property of the QRS decomposition is that it exhibits an *equal-diagonal* R-factor, that is, the diagonal entries of the upper triangular factor $\mathbf{R}$ are all equal to each other. We proved that for any matrix $\mathbf{H}$, there exists a QRS decomposition. However, the S-factor $\mathbf{S}$ is not unique. We also provided a recursive algorithm to compute an explicit solution for the factor $\mathbf{S}$.

We further revealed that the QRS decomposition has an important role to play in precoded MIMO detection theory. Namely, the S-factor $\mathbf{S}$ is the optimal precoder for the successive cancellation detector that uses the QR decomposition. We showed the following properties of the precoder and the resulting QR-detector

1) The precoder $\mathbf{S}$ minimizes the block error probability of the QR successive cancellation detector.
2) The superchannel $\mathbf{HS}$ naturally exhibits an optimally ordered column permutation, i.e., the optimal detection order for the V-BLAST detector [18] is the natural order.
3) The minimum Euclidean distance between two signal points at the channel output is equal to the minimum Euclidean distance between two constellation points at the precoder input up to a multiplicative factor that equals the diagonal entry in the $\mathbf{R}$-factor.
4) A lower and an upper bound for the free distance at the channel output is expressible in terms of the diagonal entries of the R-factor in the QR decomposition of a channel matrix.
5) The precoder $\mathbf{S}$ maximizes the lower bound of the channel's free distance subject to a power constraint.
6) The detection performance of the QR detector for a channel with the equal-diagonal R-factor $\mathbf{R}$ is asymptotically equivalent to that of the maximum likelihood detector used over the same channel when the signal to noise ratio (SNR) is large.

We applied the QRS decomposition to design the optimal QR-precoder for BPSK signals transmitted using two multiplexing schemes, TDMA and OFDM, and we presented simulation results to demonstrate the superior performance of the optimal precoder.

The presented technique for solving the optimization problem in this paper can be universally applied to zero-forcing decision-feedback equalization (ZF-DFE). Consider the following class of optimization problems involving ZF-DFE. Given are a matrix $\mathbf{B}$, a constraint P and a positive constant p. We need to find a matrix $\mathbf{F}$ that achieves the maximum $\max_{\mathrm{tr}(\mathbf{F}^H\mathbf{F})\le P} \sum_{k=1}^{N} G([\mathbf{R}]_k^{-p})$ with $[\mathbf{R}]_k$ being the k-th diagonal entry of the R-factor of QR decomposition of the matrix $\mathbf{BF}$, where the function $G(t)$ has two key properties: (a) $G(t)$ is monotonically decreasing with t; (b) $G(t)$ is concave with respect to the variable t. This class of optimization problems has a closed form solution, which is exactly the S-factor of the QRS decomposition of $\mathbf{B}$. Thus, the solution technique depends only on the features of ZF-DFE itself, but does not depend on the specifics of the function $G(\cdot)$.

In this paper, we assumed that the autocorrelation sequence of the channel coefficients (i.e., the matrix $\mathbf{H}^H\mathbf{H}$) is known to the transmitter. However, this assumption is not realistic in many wireless communication systems. Extending the design method that was presented here for a deterministic channel to the design of a precoder for a statistical (i.e., fading) channel model (for which only the second order statistics of the channel coefficients are available at the transmitter site) is still an open problem. Recently, however, progress has been made by Wang, Ma and Giannakis [48], who have shown that the truncated DFT precoders minimize the average symbol error rate at high SNR for i.i.d. Rayleigh distributed multipath channels, among a certain class of precoded OFDM transmissions.

Acknowledgment

We would like to thank Dr. Xiao Ma, Dr. Z.-Q. Luo, Dr. T. N. Davidson and Dr. A. Gershman for helpful discussions.

IX. Appendix

A. Proof of Theorem 2

Proof: The proof consists of the following three steps.

- First we show that $\mathbf{v}_1$ exists. Since $\sqrt{\lambda_k}$ are the singular values of $\mathbf{H}$, we have $det\left(\mathbf{\Lambda}^H\mathbf{\Lambda}\right)^{1/N} = det\left(\mathbf{H}^H\mathbf{H}\right)^{1/N}$. Substituting $\mathbf{A} = \mathbf{\Lambda}^H\mathbf{\Lambda}, \mathbf{B} = \mathbf{H}^H\mathbf{H}$ and $K = N$ into Corollary 1.1, we obtain that there exists a vector $\mathbf{v}_1$ that satisfies (8) and (9).
- Next we show that $\mathbf{v}_2$ exists. The key here is to show that the $(N-1)\times(N-1)$ matrix $\mathbf{A}^{(1)}$ satisfies $det\left(\mathbf{A}^{(1)}\right)^{1/(N-1)} = det\left(\mathbf{H}^H\mathbf{H}\right)^{1/N}$. Then, by substituting $\mathbf{A} = \mathbf{A}^{(1)}, \mathbf{B} = \mathbf{H}^H\mathbf{H}, K = N-1$ and $\boldsymbol{\gamma} = \mathbf{z}_2$ into Corollary 1.1, we conclude that (10)–(12) have a solution for $\mathbf{z}_2$, and therefore $\mathbf{v}_2$ exists. So all we need to prove is

$$det\left(\mathbf{A}^{(1)}\right)^{1/(N-1)} = det\left(\mathbf{H}^H\mathbf{H}\right)^{1/N}. \qquad (84)$$

To prove (84), we note that

$$\begin{aligned}&\left[\mathbf{v}_1, \mathbf{V}_1^{\perp}\right]^H \mathbf{\Lambda}^H\mathbf{\Lambda}\left[\mathbf{v}_1, \mathbf{V}_1^{\perp}\right]\\ &= \begin{pmatrix} \mathbf{v}_1^H\mathbf{\Lambda}^H\mathbf{\Lambda}\mathbf{v}_1 & \mathbf{v}_1^H\mathbf{\Lambda}^H\mathbf{\Lambda}\mathbf{V}_1^{\perp} \\ (\mathbf{V}_1^{\perp})^H\mathbf{\Lambda}^H\mathbf{\Lambda}\mathbf{v}_1 & (\mathbf{V}_1^{\perp})^H\mathbf{\Lambda}^H\mathbf{\Lambda}\mathbf{V}_1^{\perp}\end{pmatrix}.\end{aligned}$$

Hence, by Schur decomposition ([1], pp. 21-22) we get

$$\begin{aligned} det\big(\mathbf{H}^H\mathbf{H}\big) &= det(\mathbf{\Lambda}^H\mathbf{\Lambda})\\ &= det\Big(\left[\mathbf{v}_1, \mathbf{V}_1^{\perp}\right]^H \mathbf{\Lambda}^H\mathbf{\Lambda}\left[\mathbf{v}_1, \mathbf{V}_1^{\perp}\right]\Big)\\ &= (\mathbf{v}_1^H\mathbf{\Lambda}^H\mathbf{\Lambda}\mathbf{v}_1) det\Big((\mathbf{V}_1^{\perp})^H\mathbf{\Lambda}^H\mathcal{P}_{\mathbf{\Lambda}\mathbf{v}_1}\mathbf{\Lambda}\mathbf{V}_1^{\perp}\Big)\\ &= (\mathbf{v}_1^H\mathbf{\Lambda}^H\mathbf{\Lambda}\mathbf{v}_1) det(\mathbf{A}^{(1)}). \qquad (85)\end{aligned}$$

Now, since we proved that $\mathbf{v}_1$ exists such that (8) and (9) are satisfied, we get from (8) that $\mathbf{v}_1\mathbf{\Lambda}^H\mathbf{\Lambda}\mathbf{v}_1 = det\left(\mathbf{H}^H\mathbf{H}\right)^{1/N}$. Substituting this into (85) yields (84).
- By induction we easily prove that $\mathbf{v}_k$ exists for $k > 2$. Again, what is needed is to show that

$$det\left(\mathbf{A}^{(k)}\right)^{1/(N-k)} = det\left(\mathbf{H}^H\mathbf{H}\right)^{1/N}, \qquad (86)$$

and then apply Corollary 1.1. The proof is similar to proving (84), so we omit the details. □

B. Proof of Theorem 3

Proof: We first derive an expression that hods for any matrix $\mathbf{H}$ (not just for those matrices that have an equal-diagonal R-factor). Since $\mathbf{H}_k = \mathbf{U}\mathbf{\Lambda}\mathbf{V}_k$, we have $det\left(\mathbf{H}_k^H\mathbf{H}_k\right) = det\left(\mathbf{V}_k^H\mathbf{\Lambda}^H\mathbf{\Lambda}\mathbf{V}_k\right)$. Now applying Schur's decomposition ([1], pp. 21-22) to

$$\begin{aligned}\mathbf{V}_{n+1}^H\mathbf{\Lambda}^H\mathbf{\Lambda}\mathbf{V}_{n+1} &= \left[\mathbf{V}_n, \mathbf{v}_{n+1}\right]^H \mathbf{\Lambda}^H\mathbf{\Lambda}\left[\mathbf{V}_n, \mathbf{v}_{n+1}\right]\\ &= \begin{pmatrix}\mathbf{V}_n^H\mathbf{\Lambda}^H\mathbf{\Lambda}\mathbf{V}_n & \mathbf{V}_n^H\mathbf{\Lambda}^H\mathbf{\Lambda}\mathbf{v}_{n+1}\\ \mathbf{v}_{n+1}^H\mathbf{\Lambda}^H\mathbf{\Lambda}\mathbf{V}_n & \mathbf{v}_{n+1}^H\mathbf{\Lambda}^H\mathbf{\Lambda}\mathbf{v}_{n+1}\end{pmatrix},\end{aligned}$$

we get

$$\begin{aligned} det\left(\mathbf{H}_{n+1}^H\mathbf{H}_{n+1}\right) &= det\left(\mathbf{V}_{n+1}^H\mathbf{\Lambda}^H\mathbf{\Lambda}\mathbf{V}_{n+1}\right)\\ &= det(\mathbf{V}_n^H\mathbf{\Lambda}^H\mathbf{\Lambda}\mathbf{V}_n)\mathbf{v}_{n+1}^H\mathbf{\Lambda}^H\mathcal{P}_{\mathbf{\Lambda}\mathbf{V}_n}\mathbf{\Lambda}\mathbf{v}_{n+1}\\ &= det\left(\mathbf{H}_n^H\mathbf{H}_n\right)\mathbf{v}_{n+1}^H\mathbf{\Lambda}^H\mathcal{P}_{\mathbf{\Lambda}\mathbf{V}_n}\mathbf{\Lambda}\mathbf{v}_{n+1}. \end{aligned} \qquad (87)$$

Denote by $\mathbf{V}_n^{\perp}$ the orthonormal complement of $\mathbf{V}_n = [\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_n]$. Note that the vectors that comprise the matrix $\mathbf{V}_n^{\perp}$ are not necessarily the eigenvectors $\mathbf{v}_{n+1}$ through $\mathbf{v}_N$, but they do span the same subspace, and are orthonormal. Since the columns of $\mathbf{V}_n^{\perp}$ span the same subspace as $[\mathbf{v}_{n+1}, \mathbf{v}_{n+2}, \ldots \mathbf{v}_N]$, there must exist a unit-norm vector, call it $\mathbf{z}_{n+1}$, such that $\mathbf{v}_{n+1} = \mathbf{V}_n^{\perp}\mathbf{z}_{n+1}$. Substituting this into equation (87), and utilizing (12) yields

$$\mathbf{z}_{n+1}^H\mathbf{A}^{(n)}\mathbf{z}_{n+1} = \frac{det\left(\mathbf{H}_{n+1}^H\mathbf{H}_{n+1}\right)}{det\left(\mathbf{H}_n^H\mathbf{H}_n\right)}. \qquad (88)$$

Note that so far we made no assumption on the equal-diagonal properties of the R-factor of $\mathbf{H}$, i.e., there exist unit-norm vectors $\mathbf{z}_{n+1}$ such that (88) holds for *any* matrix $\mathbf{H}$. We next concentrate on matrices with equal-diagonal R-factors.

By Theorem 1, the right-hand side of (88) equals $det(\mathbf{H}^H\mathbf{H})^{1/N}$ if an only if the matrix $\mathbf{H}$ has an equal-diagonal R-factor, which completes the proof. □

REFERENCES

[1] R. A. Horn and C. Johnson, *Matrix Analysis*. Cambridge, MA: Cambridge University Press, 1985.
[2] G. Golub and C. F. van Loan, *Matrix Computations (The third edition)*. 2715 North Charles street, Baltimore Maryland: The John Hopkins University Press, 1996.
[3] S. Verdú, *Multiuser Detection*. Cambridge, MA: Cambridge University Press, 1998.
[4] L. Scharf, *Statistical Signal Processing: Detection, Estimation And Time Series Analysis*. Reading, MA: Addison-Wesley Publishing Company, Inc, 1991.
[5] V. Poor, *An Introduction to Signal Detection and Estimation (2nd ed.)*. New York: Spinger-Verlag, 1989.
[6] M. K. Varanasi, "Decision feedback multiuser detection: a systematic approach," *IEEE Trans. Inform. Theory*, vol. 45, pp. 219–240, Jan. 1999.
[7] M. K. Varanasi, "Group detection for synchronous Gaussain code-division multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 41, pp. 1083–1084, July 1995.
[8] A. Duel-Hallen, J. Holtzman, and Z. Zvonar, "Multiuser detection for CDMA systems," *IEEE Personal Commun.*, pp. 46–58, Apr. 1995.
[9] S. Moshavi, "Multi-user detection for DS-CDMA communications," *IEEE Commun. Mag.*, pp. 124–136, Oct. 1996.
[10] S. Marinkovic, B. Vucetic, and J. Evants, "Improved iterative parallel interference cancellation for coded CDMA," in *Proceedins of IEEE International Symposium on Information Theory*, (Washington DC), p. 34, June 2001.
[11] W. Phoel and M. L. Honig, "Transmitter diversity for DS-CDMA with MMSE receivers decision feedback detection," *Proc. IEEE Globecom Conference, San Francisco, CA*, vol. 35, pp. 14–15, Dec. 2000.
[12] M. L. Honig and J. S. Goldstein, "Adaptive reduced-rank residual correlation algorithms for DS-CDMA interference suppression," in *Proc. 32nd Asilomar Conf. Syst. Comput.*, (Pacific Grove, CA), Nov. 1998.
[13] M. F. Madkour, S. C. Gupta, and Y.-P. E. Wang, "Successive interference cancellation algorithms for downlink W-CDMA communications," *IEEE Trans. Wireless Comm.*, vol. 35, pp. 14–15, Jan. 2002.
[14] N. Al-Dhahir and J. Cioffi, "Efficiently computed reduced-parameter input-aided MMSE equalizer for ML detection: A unified approach," *IEEE Trans. Inform. Theory*, vol. 42, pp. 905–915, May 1996.
[15] N. Al-Dhahir and J. M. Cioffi, "MMSE decision-feedback equalizers: finite-length results," *IEEE Trans. Inform. Theory*, pp. 961–975, Jul. 1995.
[16] N. Al-Dhahir and A. H. Sayed, "The finite-length MIMO MMSE-DFE," *IEEE Trans. Signal Processing*, pp. 2921–2936, Oct. 2000.
[17] M. O. Damen, K. Abed-Meriam, and S. Burykh, "Iterative QR detection for an uncoded space-time communication architecture," in *Proc. Allerton Conf. on Comm. Control, and Comput.*, Oct. 2000.
[18] G. D. Golden, G. J. Foschini, R. A. Valenzuela, and P.-W. Wolniansky, "Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture," *Electronics Letters*, vol. 35, pp. 14–15, Jan. 1999.
[19] B. Hassibi, "An fast square root algorithm for BLAST," in *Signal, Systems and Computers Asilomar Conference*, pp. 1255–1259, Nov. 2000.

[20] X.-G. Xia, "New precoding for intersymbol interference cancellation using nonmaximally decimated multirate filterbanks with ideal FIR equalizers," *IEEE Trans. Signal Processing*, vol. 45, pp. 2413–2440, Apr. 1997.

[21] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Filterbank transceivers optimizing information rate in block transmissions over dispersive channels," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1019–1032, Apr. 1999.

[22] R. W. Chang, "Synthesis of band limited orthogonal signals for multichannel data transmission," *Bell Syst. Tech. J.*, vol. 45, pp. 1775–1796, Dec. 1996.

[23] J. Fliege, "Orthogonal multiplex carrier data transmission," *Euro. Trans. Telecommun.*, vol. 3, pp. 255–264, May 1992.

[24] B. R. Salzberg, "Performance of an efficient parallel data transmission system," *IEEE Trans. Comm.*, vol. 15, pp. 805–813, Dec. 1967.

[25] S. B. Weinstein and P. M. Ebert, "Data transmission by frequency division multiplexing using the discrete fourier transform," *IEEE Trans. Comm.*, vol. 19, pp. 628–634, Oct. 1971.

[26] A. N. Akansu, P. Duhamel, X. Lin, and M. de Courville, "Orthogonal transmultiplexers in communication: A review," *IEEE Trans. Signal Processing*, vol. 46, pp. 979–995, Apr 1998.

[27] R. V. Nee and R. Prasad, *OFDM for Wireless Multimedia Communications*. Boston, London: Artech House, 2000.

[28] J. S. Chow, J. C. Tu, and J. M. Cioffi, "A discrete multione tranceiver system for HDSL applications," *IEEE J. Select. Areas Commun.*, vol. 9, pp. 895–908, August 1991.

[29] J. S. Chow, J. C. Tu, and J. M. Cioffi, "Performance evaluation of a multichannel transceiver system for ADSL and VHDSL receivers," *IEEE J. Select. Areas Commun.*, vol. 9, pp. 909–919, August 1991.

[30] A. Ruiz, J. M. Cioffi, and S. Kasturia, "Discrete multiple tone modulation with coset coding for the spectrally shaped channel," *IEEE Trans. Commun.*, vol. 40, pp. 1012–102, June 1992.

[31] I. Kalet, "The multione channelthe multione channel," *IEEE Trans. Commun.*, vol. 37, pp. 119–124, Feb 1989.

[32] S. D. Sandberg and M. A. Tzanets, "Overlaped discrete multitone modulation for high speed copper wire communications," *IEEE J. Select. Areas Commun.*, vol. 13, pp. 1571–1585, December 1995.

[33] T. S. Rappaport, *Wireless Communications Principles And Practice*. Upper Saddle River, New Jersey: Prentice-Hall, Inc, 1996.

[34] G. Wornell, "Emerging applications of multirate signal processing and wavelets in digital communications," *Proc. IEEE*, vol. 84, pp. 586–603, April 1996.

[35] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers part I: unification and optimal designs," *IEEE Trans. Signal Processing*, vol. 47, pp. 1988–2005, July 1999.

[36] J. Milanovic, T. N. Davidson, A.-Q. Luo, and K. M. Wong, "Design of robust redundant precoding filter banks with zero-forcing equalizers for unknown frequency-selective channels," in *Int. Conf. Acoust., Speech, Signal Process.*, (Istanbul Turkey), pp. 2761–2764, June 2000.

[37] Y.-W. Ding, T. N. Davidson, J.-K. Zhang, A.-Q. Luo, and K. M. Wong, "Minimum ber block precoders for zero-forcing equalization," in *Int. Conf. Acoust., Speech, Signal Process.*, (Orlando USA), pp. 2261–2264, May 2002.

[38] J. G. Proakis, *Digital Comunications*. New York: Prentice-Hall, Inc, 1995.

[39] G. D. Forney Jr. and G. U. Ungerboeck, "Modulation and coding for linear Gaussian channel," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2384–2415, May 1998.

[40] G. D. Forney Jr., "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Transactions on Information Theory*, vol. 18, pp. 363–378, March 1972.

[41] J.-K. Zhang, A. Kavcic, X. Ma, and K. M. Wong, "Unitary precoder design for ISI channels," in *Int. Conf. Acoust., Speech, Signal Process.*, (Orlando USA), pp. 2265–2268, May 2002.

[42] Z. Wang and G. B. Giannakis, "Linearly precoded or coded OFDM against wireless channel fades," in *Third IEEE Signal Processing Workshop on Signal Advances in Wireless Communications*, (Taoyuan, Taiwan), pp. 20–23, March 2001.

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

[44] A. Graham, *Kroneker Products and Matrix Calculus*. New York: John Wiley & Sons, INC, 1981.

[45] M. K. Simon and M.-S. Alouini, "A unified approach to the perfromance analysis of digital communication over generalized fading channels," *Proceedings of the IEEE*, vol. 86, pp. 1860–1877, Sept. 1998.

[46] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. Prentice-Hall, Inc., 1993.

[47] M. O. Damen, A. Chkeif, and J. C. Belfiore, "Lattice codes detector for space-time codes," *IEEE Commun. Letters*, vol. 4, pp. 161–163, May 2000.

[48] Z. Wang, X. Ma, and G. B. Giannakis, "Optimality of single-carrier zero-padded block transmission," in *Proc. of Wireless Commun. Networking Conf.*, pp. 660–664, Mar. 2002.

PLACE PHOTO HERE

Jian-Kang Zhang received the B.S. degree in Mathematics from Shaanxi Normal University, Xi'an, China, M.S. degree in Mathematics from Northwest University, Xi'an, China, and the Ph.D. degree in Electrical Engineering from Xidian University, Xi'an, China.

He is now with the Department of Electrical and Computer Engineering at McMaster University, Hamilton, Ont., Canada. His research interests include multirate filterbanks, wavelet and multiwavelet transforms and their applications, number theory transform and fast algorithm. His current research focuses on multiscale wavelet transform, random matrices, precoding and space-time coding for block transmissions and multicarrier and wideband wireless communication systems.

PLACE PHOTO HERE

Aleksandar Kavčić (S'93–M'98–SM'04) received the Dipl.-Ing. degree in Electrical Engineering from Ruhr-University, Bochum, Germany in 1993, and the Ph.D. degree in Electrical and Computer Engineering from Carnegie Mellon University, Pittsburgh, Pennsylvania in 1998.

Since 1998, he has been with the Division of Engineering and Applied Sciences of Harvard University where he is John L. Loeb Associate Professor of Natural Sciences. From 1998 to 2002 he was an Assistant Professor of Electrical Engineering at Harvard University. He held short-term research positions at Seagate Technology in 1995, Read-Rite Corporation in 1996, and Quantum Corporation from 1997 to 1998, and has served as a technical consultant for Quantum Corporation and Link-A-Media Devices. His research spans topics in Communications, Signal Processing, Information Theory and Magnetic Recording.

Dr. Kavčić received the IBM Partnership Award in 1999 and the NSF CAREER Award in 2000. He served on the Editorial Board of the IEEE TRANSACTIONS ON INFORMATION THEORY as Associate Editor for Detection and Estimation from 2001 to 2004.

PLACE PHOTO HERE

Kon-Max Wong was born in Macau. He received his BSc(Eng), DIC, PhD, and DSc(Eng) degrees, all in electrical engineering, from the University of London, England, in 1969, 1972, 1974 and 1995, respectively. He started working at the Transmission Division of Plessey Telecommunications Research Ltd., England, in 1969. In October 1970 he was on leave from Plessey pursuing postgraduate studies and research at Imperial College of Science and Technology, London. In 1972, he rejoined Plessey as a research engineer and worked on digital signal processing and signal transmission. In 1976, he joined the Department of Electrical Engineering at the Technical University of Nova Scotia, Canada, and in 1981, moved to McMaster University, Hamilton, Canada, where he has been a Professor since 1985 and served as Chairman of the Department of Electrical and Computer Engineering in 1986-87, and 1988-94. Professor Wong was on leave as Visiting Professor at the Department of Electronic Engineering of the Chinese University of Hong Kong from 1997 to1999. At present, he holds the NSERC-Mitel Professorship of Signal Processing and is the Director of the Communication Technology Research Centre at McMaster University. He is also serving another five-year term as Chair of the Department. His research interest is in signal processing and communication theory and has published over 190 papers in the area.

Professor Wong was the recipient of the IEE Overseas Premium for the best paper in 1989, and is a Fellow of IEEE, a Fellow of the Institution of Electrical Engineers, a Fellow of the Royal Statistical Society, and a Fellow of the Institute of Physics. He also served as an Associate Editor of the IEEE Transaction on Signal Processing, 1996-98 and is the chairman of the Sensor Array and Multi-channel Signal Processing Technical Committee of the Signal Processing Society. Professor Wong was the recipient of a medal presented by the International Biographical Centre, Cambridge, England, for his "outstanding contributions to the research and education in signal processing" in May 2000, and was honored with the inclusion of his biography in the two books: Outstanding People of the 20th Century and 2000 Outstanding Intellectuals of the 20th Century published by IBC to celebrate the arrival of the new millennium.