Joint Transceiver Design for MIMO Communications Using Geometric Mean Decomposition

Yi Jiang, *Student Member, IEEE*, Jian Li, *Fellow, IEEE*, and William W. Hager

***Abstract*—In recent years, considerable attention has been paid to the joint optimal transceiver design for multi-input multi-output (MIMO) communication systems. In this paper, we propose a joint transceiver design that combines the geometric mean decomposition (GMD) with either the conventional zero-forcing VBLAST decoder or the more recent zero-forcing dirty paper precoder (ZFDP). Our scheme decomposes a MIMO channel into multiple *identical* parallel subchannels, which can make it rather convenient to design modulation/demodulation and coding/decoding schemes. Moreover, we prove that our scheme is asymptotically optimal for (moderately) high SNR in terms of both channel throughput and bit error rate (BER) performance. This desirable property is not shared by any other conventional schemes. We also consider the subchannel selection issues when some of the subchannels are too poor to be useful. Our scheme can also be combined with orthogonal frequency division multiplexing (OFDM) for intersymbol interference (ISI) suppression. The effectiveness of our approaches has been validated by both theoretical analyses and numerical simulations.**

***Index Terms*—Channel capacity, dirty paper precoding, intersymbol interference suppression, joint transceiver design, matrix decomposition, MIMO, VBLAST, water filling.**

I. Introduction

COMMUNICATIONS over multiple-input multiple-output (MIMO) channels have been the subject of intense research over the past several years because MIMO channels can support much greater data rate and higher reliability over their single-input single-output (SISO) counterpart [1]–[4]. Two main approaches have been proposed to exploit the many advantages of MIMO channels. One is the space-time coding method that aims at improving communication reliabilities by exploiting the diversity gain (e.g., [5]–[7]). The other is the spatial multiplexing method, e.g., the Bell Labs Space-Time Architecture (BLAST) [2], [8], [9], which focuses on maximizing the channel throughput.

Both methods assume that the channel state information (CSI) is available at the receiver (CSIR) only. However, if the communication environment is slowly time varying, such as communications via indoor wireless local area networks (WLANs) or the bonded digital subscriber lines (DSLs), the availability of CSI at the transmitter (CSIT) is also possible via feedback or the reciprocal principle when time division duplex (TDD) is used. In fact, in the third-generation WCDMA standard [10], both CSIR and CSIT are assumed available and are referred to as the closed-loop transmit diversity or transmit adaptive array (TxAA) technique. Based on this assumption, the joint optimal transceiver design (which is also referred to as precoding at the transmitter and equalization at the receiver) has recently attracted considerable attention [11]–[20].

Several designs have been proposed based on a variety of criteria, including minimum mean-squared-error (MMSE) [11], [14], [15], maximum signal-to-noise ratio (SNR) [14], maximum information rate [12], [13], [15], and bit error rate (BER) based criteria [16]–[18]. More recently, a unified framework has been presented to accommodate all these criteria, under which the design problems can be solved via convex optimization methods [19].

The aforementioned literature on joint transceiver design considered linear transformations only. It is widely understood that the singular value decomposition (SVD), which decomposes a MIMO channel into multiple parallel subchannels, and water filling can be used to achieve the channel capacity [3]. However, due to the very different SNRs of the subchannels, this apparently simple scheme requires careful bit allocation (see, e.g., [12], [13], and [16]) to match the subchannel capacity and achieve a prescribed BER. Bit allocation not only increases the coding/decoding complexity but is also inherently capacity lossy because of the finite constellation granularity. An alternative is to use the same constellation in all the subchannels (or subcarriers), like the schemes adopted by the European standard HIPERLAN/2 and the IEEE 802.11 standards for WLANs. However, for this alternative, the BER is dominated by the subchannels with the lowest SNRs. To optimize the BER performance, more signal power could be allocated to the poorer subchannels. Yet this approach causes significant capacity loss due to "inverse water filling" like power allocation. There is apparently a fundamental tradeoff between the capacity and the BER performance.

In this paper, we propose a novel transceiver design based on the geometric mean decomposition (GMD) [21]. By combining GMD with either the conventional VBLAST decoder [8], which is in fact a generalized decision feedback equalizer (GDFE) [22], or the more recent zero-forcing dirty paper precoder (ZFDP),[1] our scheme decomposes a MIMO channel into multiple *identical* parallel subchannels. This desirable property can

Manuscript received March 5, 2004; revised October 15, 2004. This work was supported in part by the National Science Foundation Grant CCR-0104887. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Ananthram Swami.

Y. Jiang was with Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611-6130 USA. He is now with the University of Colorado, Boulder, CO 80309 USA.

J. Li is with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611-6130 USA (e-mail: li@dsp.ufl.edu).

W. W. Hager is with the Department of Mathematics, University of Florida, Gainesville, FL 32611-8105 (e-mail: hager@math.ufl.edu).

Digital Object Identifier 10.1109/TSP.2005.855398

[1]In the sequel, we refer to the two versions of this method as GMD-VBLAST and GMD-ZFDP, respectively, and refer to either of them as the GMD scheme.

bring about much convenience in coding/decoding and modulation/demodulation processes. Moreover, we prove that our scheme is asymptotically optimal for (moderately) high SNR in terms of both channel throughput and BER performance. Hence, the GMD scheme does not make tradeoffs between the throughput and the BER performance. Instead, it attempts to get the best of the both worlds simultaneously.

The remainder of the paper is organized as follows. Section II introduces the MIMO flat-fading channel model and some relevant results on channel capacity. The capacity loss due to the linearity constraints of transceiver designs is analyzed in Section III. After reviewing the VBLAST and ZFDP briefly, we propose the GMD scheme in Section IV. Section V addresses the performance analyses and implementation issues of the GMD scheme. The relationship between GMD and the paper [23] is also discussed. Section VI presents several numerical examples showing the superior performance of the GMD scheme from both the information theoretic and the BER aspects. The advantages of the GMD scheme over its linear counterparts and the open-loop counterpart, i.e., the conventional VBLAST, are clearly demonstrated. Section VII gives the conclusions of this paper.

II. Channel Model and Channel Capacity

In this section, we introduce the flat fading channel model. We also discuss several results of channel capacity needed by the GMD scheme.

A. Channel Model

We first consider a communication system with M_t transmitting and M_r receiving antennas in a frequency flat-fading channel. The sampled baseband signal is given by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z} \tag{1}$$

where $\mathbf{x} \in \mathbb{C}^{M_t \times 1}$ is the transmitted signal, $\mathbf{y} \in \mathbb{C}^{M_r \times 1}$ is the received signal, and $\mathbf{H} \in \mathbb{C}^{M_r \times M_t}$ is the channel matrix with the (i,j)th element denoting the fading coefficient between the jth transmitting and ith receiving antennas. Throughout this paper, we let K denote the rank of $\mathbf{H}$. We assume that $\mathbf{z} \sim N\left(0, \sigma_z^2 \mathbf{I}_{M_r}\right)$ is zero-mean circularly symmetric complex Gaussian noise, where $\mathbf{I}_{M_r}$ denotes the identity matrix with dimension M_r. We define the SNR as

$$\rho = \frac{E[\mathbf{x}^*\mathbf{x}]}{\sigma_z^2} \tag{2}$$

where $(\cdot)^*$ denotes the conjugate transpose, and $E[\cdot]$ stands for the expected value. Throughout this paper, we assume perfect CSIR and CSIT, i.e., $\mathbf{H}$ is known exactly at both the transmitter and receiver. Yet a combination of the GMD scheme with blind channel subspace tracking technique is also possible, as we will discuss in Section V-B.

B. Channel Capacity

The capacity of the channel of (1) with CSIT is

$$C_{IT} = \max_{\mathbf{R_x}} \log_2 \frac{\left|\sigma_z^2 \mathbf{I} + \mathbf{H}\mathbf{R_x}\mathbf{H}^*\right|}{\left|\sigma_z^2 \mathbf{I}\right|} \tag{3}$$

where $\mathbf{R_x} = E[\mathbf{x}\mathbf{x}^*]$, and $|\cdot|$ denotes the determinant of a matrix. Here, the subscript of C_{IT} stands for "informed transmitter." We assume that the transmitted signal is power constrained, say $\mathrm{Tr}\{\mathbf{R_x}\} \le P$, where $\mathrm{Tr}\{\mathbf{R}\}$ denotes the trace of $\mathbf{R}$. The SNR is $\rho = \left(P/\sigma_z^2\right)$. Calculating the channel capacity is equivalent to solving the following convex optimization problem:

$$C_{IT} = \max_{\mathrm{Tr}\{\mathbf{S}\} \le \rho} \log_2 |\mathbf{I} + \mathbf{H}\mathbf{S}\mathbf{H}^*|. \tag{4}$$

The solution to (4) is (cf. [3])

$$C_{IT} = \sum_{n=1}^{K} \log_2 \left(1 + \lambda_{S,n}\lambda_{H,n}^2\right) \text{ bit/s/Hz} \tag{5}$$

where K is the rank of $\mathbf{H}$, $\{\lambda_{H,n}\}_{n=1}^K$ are the K nonzero singular values of $\mathbf{H}$, and $\lambda_{S,n}$ is found via "water filling" to be

$$\lambda_{S,n}(\mu) = \left(\mu - \frac{1}{\lambda_{H,n}^2}\right)^+ \tag{6}$$

where μ is chosen such that $\sum_{n=1}^K \lambda_{S,n}(\mu) = \rho$ and $(a)^+ = \max\{0, a\}$.

If the CSIT is not available, the optimal transmission strategy is to evenly allocate power to each antenna [3]. For this case, $\mathbf{S} = (\rho/M_t)\mathbf{I}_{M_t}$, and the channel capacity with a uninformed transmitter (UT) is

$$C_{\mathrm{UT}} = \sum_{n=1}^{M_t} \log_2 \left(1 + \frac{\rho\lambda_{H,n}^2}{M_t}\right) \text{ bit/s/Hz}. \tag{7}$$

It is proven in [24] that if $K = M_t$

$$\frac{C_{IT}}{C_{\mathrm{UT}}} \to 1 \text{ as } \rho \to \infty. \tag{8}$$

We claim a stronger relationship as follows.[2]

Lemma II.1: For the data model in (1), if the channel matrix $\mathbf{H}$ is of full column rank, i.e., $K = M_t$, then

$$C_{IT} - C_{\mathrm{UT}} \to 0 \text{ as } \rho \to \infty. \tag{9}$$

Proof: Inserting (6) into (5) yields

$$C_{IT} = \sum_{n=1}^{K} \log_2 \left(\mu\lambda_{H,n}^2\right)^+. \tag{10}$$

Observe that $\lambda_{S,n}(\mu)$ is an increasing function of μ. Assuming that ρ is large enough that $\lambda_{S,n}(\mu)$ is strictly positive for each n when μ is chosen such that

$$K\mu - \sum_{n=1}^{K} \frac{1}{\lambda_{H,n}^2} = \rho \tag{11}$$

or

$$\mu = \frac{\rho}{K} + \frac{1}{K}\sum_{n=1}^{K} \frac{1}{\lambda_{H,n}^2} \tag{12}$$

from (7), (10), and (12) and using the fact that $K = M_t$, we have

$$C_{IT} - C_{\mathrm{UT}} = \sum_{n=1}^{K} \log_2 \left(\frac{\rho + \sum_{n=1}^{K} \frac{1}{\lambda_{H,n}^2}}{\rho + \frac{K}{\lambda_{H,n}^2}}\right). \tag{13}$$

[2]A similar, but somewhat vague, statement is found in [9].

Note that

$$\lim_{\rho\to\infty}\frac{\rho+\sum_{n=1}^{K}\frac{1}{\lambda_{H,n}^2}}{\rho+\frac{K}{\lambda_{H,n}^2}}=1,\ \text{for } 1\le n\le K \tag{14}$$

and that $f(x)=\log x$ is a continuous function of x if $x>0$. The lemma follows immediately from (13). ∎

We note that this lemma is more informative than (8) since C_{IT} tends to infinity as ρ increases. Lemma II.1 shows that the CSIT becomes useless for high SNR from the capacity perspective.

On the other hand, we note that CSIT can be very helpful in the following cases:

A) The SNR is low or moderate.
B) $\mathbf{H}$ is rank deficient or ill-conditioned.
C) There are more transmitting antennas than receiving ones, i.e., $M_t > M_r$.

Moreover, the availability of CSIT provides more freedom, which makes it easier to design a joint transceiver to achieve the underlying channel capacity.

III. Rate Performance of Linear Transceivers

To gain insights into the limitations of the linear transceiver designs, we analyze the asymptotic rate performances of two typical linear transceiver designs for high SNR. We will show that the linear transceivers may suffer from considerable capacity loss, and there is apparently a fundamental tradeoff between the throughput and the BER performance.

For all the linear transceiver designs (cf. [19]), the information symbol $\mathbf{s}$ is precoded to be $\mathbf{x}=\mathbf{Fs}$, and without loss of generality (w.l.o.g.), we assume $E[\mathbf{ss}^*]=\mathbf{I}$. According to the channel model of (1), the received data vector is

$$\mathbf{y}=\mathbf{HFs}+\mathbf{z}. \tag{15}$$

The optimal linear receiver is always the Wiener filter (see, e.g., [16])

$$\mathbf{G}_{opt}=\mathbf{F}^*\mathbf{H}^*\left(\mathbf{HFF}^*\mathbf{H}^*+\sigma_z^2\mathbf{I}\right)^{-1} \tag{16}$$

which yields the optimal estimate of the information symbol $\hat{\mathbf{s}}=\mathbf{G}_{opt}\mathbf{y}$. The mean-squared-error (MSE) matrix of $\hat{\mathbf{s}}$ is

$$\mathbf{E}=\left(\mathbf{I}+\mathbf{F}^*\mathbf{H}^*\mathbf{HF}\frac{1}{\sigma_z^2}\right)^{-1}. \tag{17}$$

Note that $\mathbf{E}$ is a function of the linear precoder $\mathbf{F}$. In the following, we analyze two linear precoder designs based on the minimization of the trace of the MSE matrix (MTM) and the minimization of the maximum diagonal elements of MSE matrix (MMD) criteria, which are referred to as ARITH-MSE and MAX-MSE in [19], respectively. We choose these two schemes because they appear to be the most typical ones, and the MMD scheme yields the optimal (or very close to the optimal) performance among all the linear transceivers. Indeed, the MMD is equivalent to the linear MIN-BER scheme in the flat-fading channel case (see [19]). We do not consider the SVD plus water filling scheme herein since it requires complicated bit loading.

The MTM scheme, or ARITH-MSE, which has appeared in several linear transceiver design papers (see, e.g., [15], [16], and [19]), attempts to minimize $\mathrm{Tr}(\mathbf{E})$ with respect to $\mathbf{F}$. The MTM precoder turns out to be

$$\mathbf{F}_{\mathrm{MTM}}=\mathbf{V}\mathbf{\Phi}^{1/2} \tag{18}$$

where $\mathbf{V}$ is as defined in the SVD $\mathbf{H}=\mathbf{U}\mathbf{\Lambda}\mathbf{V}^*$, and $\mathbf{\Phi}$ is a diagonal matrix whose ith diagonal element ϕ_i denotes the signal power loaded to the ith subchannel. According to the literature (see, e.g., [16, Sec.III-A])

$$\phi_i=\left(\frac{\sigma_z}{\mu^{1/2}\lambda_{H,i}}-\frac{\sigma_z^2}{\lambda_{H,i}^2}\right)^+ \tag{19}$$

where μ is the Lagrange multiplier that controls the loaded power such that $\sum_{i=1}^{K}\phi_i=\rho\sigma_z^2$. Suppose ρ is sufficiently large. Then, all the K subchannels are used, and

$$\sum_{i=1}^{K}\left(\frac{\sigma_z}{\mu^{1/2}\lambda_{H,i}}-\frac{\sigma_z^2}{\lambda_{H,i}^2}\right)=\rho\sigma_z^2 \tag{20}$$

or

$$\mu^{-1/2}=\frac{\rho+\sum_{i=1}^{K}\lambda_{H,i}^{-2}}{\sum_{i=1}^{K}\lambda_{H,i}^{-1}\sigma_z^{-1}}. \tag{21}$$

Substituting (18), (19), and (21) into (17), we see that $\mathbf{E}$ is diagonal with the ith diagonal element

$$E_i=\frac{\sum_{k=1}^{K}\lambda_{H,k}^{-1}}{\left(\rho+\sum_{k=1}^{K}\lambda_{H,k}^{-2}\right)\lambda_{H,i}} \tag{22}$$

Then [cf. (28) of [19]]

$$C_i=-\log_2 E_i \tag{23}$$

$$=\log_2\left(\frac{\rho+\sum_{k=1}^{K}\lambda_{H,k}^{-2}}{\sum_{i=k}^{K}\lambda_{H,k}^{-1}}\right)+\log_2\lambda_{H,i}. \tag{24}$$

Hence, the sum rate of the channel using the MTM scheme is

$$C_{\mathrm{MTM}}=\sum_{i=1}^{K}C_i$$

$$=K\log_2\left(\frac{\rho+\sum_{k=1}^{K}\lambda_{H,k}^{-2}}{\sum_{k=1}^{K}\lambda_{H,k}^{-1}}\right)+\sum_{i=1}^{K}\log_2\lambda_{H,i}. \tag{25}$$

The channel capacity with uniform power loading in the K subchannels is

$$C_{\mathrm{UPL}}=\sum_{i=1}^{K}\log_2\left(1+\frac{\rho}{K}\lambda_{H,i}^2\right). \tag{26}$$

Here, C_{UPL} is different from C_{UT} defined in (7) in that C_{UPL} corresponds to the channel with the transmitter knowing the range space of $\mathbf{H}$.

It follows from (25) and (26) that

$$C_{\mathrm{UPL}} - C_{\mathrm{MTM}} = \sum_{i=1}^{K} \log_2 \left(1 + \frac{\rho}{K}\lambda_{H,i}^2\right) - K \log_2 \left(\frac{\rho + \sum_{i=1}^{K} \lambda_{H,i}^{-2}}{\sum_{i=1}^{K} \lambda_{H,i}^{-1}} \right) - \sum_{i=1}^{K} \log_2 \lambda_{H,i}. \quad (27)$$

After some straightforward calculations, we have

$$\lim_{\rho\to\infty} C_{\mathrm{UPL}} - C_{\mathrm{MTM}} = K \log_2 \frac{\frac{1}{K}\sum_{i=1}^{K} \lambda_{H,i}^{-1}}{\left(\prod_{i=1}^{K} \lambda_{H,i}^{-1}\right)^{1/K}} \frac{\frac{\mathrm{bit}}{\mathrm{s}}}{\mathrm{Hz}}. \quad (28)$$

Note that for any real-valued sequence $\{\lambda_{H,i}\}_{i=1}^{K} > 0$, the arithmetic mean is greater than or equal to the geometric mean, or $(1/K)\sum_{i=1}^{K} \lambda_{H,i}^{-1} \geq \left(\prod_{i=1}^{K} \lambda_{H,i}^{-1}\right)^{1/K}$. Hence, we conclude that $\lim_{\rho\to\infty} C_{\mathrm{UPL}} - C_{\mathrm{MTM}} \geq 0$, and the equality holds if and only if $\{\lambda_{H,i}\}_{i=1}^{K}$ are all the same. We infer from (28) that the capacity loss of the MTM transceiver can be quite large if the channel matrix $\mathbf{H}$ has a large condition number, which is verified in Section VI.

If the same constellation is used for each subchannel, then the substream corresponding to the largest E_i dominates the overall BER performance. Recall that $E_i = \left(\sum_{k=1}^{K} \lambda_{H,k}^{-1} / \left(\rho + \sum_{k=1}^{K} \lambda_{H,k}^{-2}\right) \lambda_{H,i}\right)$, which is proportional to the inverse of $\lambda_{H,i}$. Hence, the subchannels may have very different SNRs, especially when $\mathbf{H}$ has a large condition number. To mitigate this undesirable effect, one can use the MMD transceiver, or MAX-MSE (cf. [19, Sec. V-A5], with

$$\mathbf{F}_{\mathrm{MMD}} = \mathbf{F}_{\mathrm{MTM}} \boldsymbol{\Theta} \quad (29)$$

where $\boldsymbol{\Theta}$ is a unitary matrix that makes all the diagonal elements of $\mathbf{E}$ in (17) the same, that is

$$\bar{E} = \frac{1}{K} \sum_{i=1}^{K} E_i. \quad (30)$$

According to (23), the capacity of the channel using the MMD linear transceiver is

$$C_{\mathrm{MMD}} = -K \log_2 \bar{E} = -K \log_2 \frac{1}{K} \sum_{i=1}^{K} E_i. \quad (31)$$

Thus

$$C_{\mathrm{MTM}} - C_{\mathrm{MMD}} = K \log_2 \frac{1}{K} \sum_{i=1}^{K} E_i - \sum_{i=1}^{K} \log_2 E_i \quad (32)$$

$$= K \log_2 \frac{\frac{1}{K}\sum_{i=1}^{K} E_i}{\left(\prod_{i=1}^{K} E_i\right)^{1/K}} \quad (33)$$

$$= K \log_2 \frac{\frac{1}{K}\sum_{i=1}^{K} \lambda_{H,i}^{-1}}{\left(\prod_{i=1}^{K} \lambda_{H,i}^{-1}\right)^{1/K}} \quad (34)$$

where, to get (34) from (33), we have used (22). Note that the relative capacity loss of MMD compared with MTM is independent of SNR, given that all the subchannels are used. Interestingly, we can see from (28) and (34) that $C_{\mathrm{MTM}} - C_{\mathrm{MMD}} = \lim_{\rho\to\infty} C_{\mathrm{UPL}} - C_{\mathrm{MTM}}$. We conclude that asymptotically, for high SNR, the MMD transceiver has twice the capacity loss of MTM, i.e.,

$$\lim_{\rho\to\infty} C_{\mathrm{UPL}} - C_{\mathrm{MMD}} = 2K \log_2 \frac{\frac{1}{K}\sum_{i=1}^{K} \lambda_{H,i}^{-1}}{\left(\prod_{i=1}^{K} \lambda_{H,i}^{-1}\right)^{1/K}} \ \mathrm{bit/s/Hz} \quad (35)$$

although it may yield better BER performance. An intuitive explanation of the capacity loss of the MMD transceiver is as follows. Note that the only difference between MTM and MMD is the prerotation matrix $\boldsymbol{\Theta}$, which is an invariant operator in terms of information capacity. However, $\boldsymbol{\Theta}$ makes the MSE matrix $\mathbf{E}$ nondiagonal, which means that the elements of $\hat{\mathbf{s}} = \mathbf{G}_{opt}\mathbf{y}$ are correlated. Clearly, the correlation contains useful information for symbol detection and decoding. However, the linear equalizer ignores the correlation, which results in the additional capacity loss quantified in (34).

In summary, the MTM transceiver suffers from capacity loss of (28) due to the information theoretically nonoptimal power loading defined in (19). The MMD transceiver suffers from additional capacity loss because it makes the MSE matrix nondiagonal. Hence, there is an apparently inevitable tradeoff between the information rate and BER performance if the same symbol constellation is used in the different subchannels. The main contribution of this paper is to introduce the GMD scheme and clarify that there is not necessarily a tradeoff between BER performance and channel capacity. Indeed, the GMD scheme attempts to achieve the best of both worlds simultaneously.

IV. Geometric Mean Decomposition for Transceiver Design

In this section, we first give a brief introduction to the VBLAST architecture [8], which is equivalent to the GDFE [22]. We also introduce the more recent ZFDP applied to the MIMO broadcast channels [25], [26]. Then, we introduce the novel GMD algorithm, which can be combined with either VBLAST or ZFDP.

A. VBLAST and ZFDP

VBLAST is a simple suboptimal receiver interface that is used in the MIMO system, assuming that only CSIR is available. For a MIMO system (1) with $M_t \leq M_r$ and rank $K = M_t$, the transmitter allocates independent bit streams across the M_t transmitting antennas with no precoding. To decode the transmitted information symbol, VBLAST first estimates the signal with the spatial structure $\mathbf{h}_{M_t}$, where $\mathbf{h}_i$ denotes the ith column of $\mathbf{H}$ and then cancels it out from the received signal vector. Next, it estimates the signal with spatial structure $\mathbf{h}_{M_t-1}$, and so on. The signal estimator can be either the ZF or MMSE estimator. Some proper reordering of the columns of $\mathbf{H}$ is helpful

to improve the BER performance [8]. This decoding scheme involves *sequential nulling and cancellation*, which is proven to be equivalent to the GDFE [22].

The ZF nulling step in the VBLAST scheme can be represented by the QR decomposition $\mathbf{H} = \mathbf{QR}$, where $\mathbf{Q}$ is an $M_r \times K$ matrix with orthonormal columns, and $\mathbf{R}$ is a $K \times K$ upper triangular matrix. Let us rewrite (1) as

$$\mathbf{y} = \mathbf{QRx} + \mathbf{z}. \tag{36}$$

Multiplying $\mathbf{Q}^*$ with both sides of (36) yields

$$\tilde{\mathbf{y}} = \mathbf{Rx} + \tilde{\mathbf{z}} \tag{37}$$

or

$$\begin{bmatrix} \tilde{y}_1 \\ \tilde{y}_2 \\ \vdots \\ \tilde{y}_K \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1K} \\ 0 & r_{22} & \dots & r_{2K} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & r_{KK} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_K \end{bmatrix} + \begin{bmatrix} \tilde{z}_1 \\ \tilde{z}_2 \\ \vdots \\ \tilde{z}_K \end{bmatrix}. \tag{38}$$

The sequential signal detection is as follows:

```
for i = K : −1 : 1
  x̂_i = C[(ỹ_i − Σ_{j=i+1}^{K} r_ij x̂_j)/r_ii]
end
```

Here, $\mathcal{C}$ stands for mapping to the nearest symbol in the symbol constellation. Ignoring the error-propagation effect, we see that the MIMO channel is decomposed into K parallel scalar subchannels

$$\tilde{y}_i = r_{ii}x_i + \tilde{z}_i, \quad i = 1, 2, \dots, K. \tag{39}$$

Next, we consider a broadcast MIMO channel with M_t transmitting antennas and M_r receiving antennas $(M_t \geq M_r)$. The channel model is exactly the same as (1), and the CSIT is available. However, the receiving antennas cannot cooperate with each other. A vector transmission scheme was proposed in [27], which combines the QR decomposition and "dirty paper" precoding. We refer to this approach as the ZFDP. (The use of the "dirty paper" phrase is due to Costa [28].)

The ZFDP scheme resembles the zero-forcing VBLAST method. It also goes through the *sequential nulling and cancellation* procedure. The only difference is that all these operations are done by the transmitter.

By assuming $\mathbf{H}$ to be of full row rank, i.e., $K = M_r$, ZFDP also begins with the QR decomposition $\mathbf{H}^* = \breve{\mathbf{Q}}\breve{\mathbf{R}}$. Let us rewrite (1) as

$$\mathbf{y} = \breve{\mathbf{R}}^*\breve{\mathbf{Q}}^*\mathbf{x} + \mathbf{z}. \tag{40}$$

Denoting $\mathbf{x} = \breve{\mathbf{Q}}\breve{\mathbf{x}}$ yields

$$\mathbf{y} = \breve{\mathbf{R}}^*\breve{\mathbf{x}} + \mathbf{z} \tag{41}$$

or

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_K \end{bmatrix} = \begin{bmatrix} \breve{r}_{11} & 0 & \dots & 0 \\ \breve{r}_{21} & \breve{r}_{22} & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \breve{r}_{K1} & \dots & \dots & \breve{r}_{KK} \end{bmatrix} \begin{bmatrix} \breve{x}_1 \\ \breve{x}_2 \\ \vdots \\ \breve{x}_K \end{bmatrix} + \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_K \end{bmatrix}. \tag{42}$$

Denote $\mathbf{s} \in \mathbb{C}^{K\times 1}$ as the symbol vector destined for the K receivers. We wish to have $\breve{\mathbf{x}}$, satisfying

$$\begin{bmatrix} \breve{r}_{11}s_1 \\ \breve{r}_{22}s_2 \\ \vdots \\ \breve{r}_{KK}s_r \end{bmatrix} = \begin{bmatrix} \breve{r}_{11} & 0 & \dots & 0 \\ \breve{r}_{21} & \breve{r}_{22} & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \breve{r}_{K1} & \dots & \dots & \breve{r}_{KK} \end{bmatrix} \begin{bmatrix} \breve{x}_1 \\ \breve{x}_2 \\ \vdots \\ \breve{x}_K \end{bmatrix}. \tag{43}$$

The solution to (43) is

$$\breve{\mathbf{x}} = \breve{\mathbf{R}}^{-*}\,\mathrm{diag}\{\breve{\mathbf{R}}\}\mathbf{s}. \tag{44}$$

However, the matrix inversion can amplify the norm of $\breve{\mathbf{x}}$ significantly, which can lead to additional power consumption at the transmitter. By exploiting the finite alphabet property of the communication signals, the modulo arithmetic precoder (more recently known as the Tomlinson–Harashima Precoder [29], [30]) can be applied to bound the value of the transmitted signal. Moreover, the trellis precoding can be used to eliminate the 1.53-dB shape-loss of Tomlinson–Harashima precoding [31]. The ZFDP transmission scheme decomposes the MIMO channel into K parallel scalar channels (see [27] for more details)

$$y_i = \breve{r}_{ii}\breve{x}_i + z_i \quad i = 1, 2, \dots, K. \tag{45}$$

Several remarks are now in order. a) VBLAST is shown to be able to achieve only about 72% of the capacity [8]. That is because imposing the same rate of transmission on all the transmitters makes the channel capacity limited by the worst of the K scalar subchannels. b) ZFDP can achieve the broadcast channel capacity for high SNR [26], but the subchannels have different fading levels. Hence, the transmitter, just like the aforementioned linear transceivers, has to consider the tradeoffs between the BER performance and the channel throughput. c) The ZFDP scheme causes no error propagation, and thus, (45) is precise. d) Both VBLAST and ZFDP involve nonlinear operations.

B. Geometric Mean Decomposition

Note that VBLAST assumes no cooperation among transmitting antennas, and ZFDP assumes no cooperation at the receivers. Then, a natural question arises: Can we exploit both the CSIR and CSIT to make things better if both CSIR and CSIT are available? We attempt to address this question next.

In the sequel, we assume that the same signal constellation is used in all the independent symbol streams to reduce the system complexity. This is consistent with the HIPERLAN/2 and IEEE 802.11 standards. Then the overall BER performance of the system will be limited by the subchannel with the lowest SNR. To mitigate this problem, based on (39) and (45), we consider the following optimization problem

$$\begin{aligned} \max_{\mathbf{Q},\mathbf{P}} \min \quad & \{r_{ii} : 1 \leq i \leq K\} \\ \text{subject to} \quad & \mathbf{R} = \mathbf{Q}^*\mathbf{HP} \\ & \mathbf{R} \in \mathbb{R}^{K\times K}, \quad r_{ij} = 0, \text{ for } i > j \\ & r_{ii} > 0, \text{ for } 1 \leq i \leq K \\ & \mathbf{Q}^*\mathbf{Q} = \mathbf{P}^*\mathbf{P} = \mathbf{I}_K \end{aligned} \tag{46}$$

where the semi-unitary matrices $\mathbf{Q}$ and $\mathbf{P}$ denote the linear operations at the receiver and transmitter, respectively.

Since both $\mathbf{Q}$ and $\mathbf{P}$ are semi-unitary matrices, we have $\prod_{n=1}^{K} r_{nn} = \prod_{n=1}^{K} \lambda_{H,n}$, where $\{\lambda_{H,n}\}_{n=1}^{K}$ are the K nonzero singular values of $\mathbf{H}$. In [21], we show that if there exist semi-unitary matrices $\mathbf{P}$ and $\mathbf{Q}$ satisfying

$$\mathbf{H} = \mathbf{Q}\mathbf{R}\mathbf{P}^*, \text{ or equivalently, } \mathbf{R} = \mathbf{Q}^*\mathbf{H}\mathbf{P} \tag{47a}$$

where the diagonal elements of $\mathbf{R}$ are given by

$$r_{ii} = \bar{\lambda}_H \triangleq \left(\prod_{n=1}^{K} \lambda_{H,n}\right)^{1/K}, \quad 1 \le i \le K \tag{47b}$$

then the $\mathbf{R}$ in (47) is the solution to (46). The existence of the decomposition (47) is a corollary to the following theorem due to Horn [32].

Theorem IV.1: If $\mathbf{r} \in \mathbb{R}^n$ and $\boldsymbol{\lambda} \in \mathbb{R}^n$ satisfy

$$\prod_{i=1}^{k} |r_i| \le \prod_{i=1}^{k} \lambda_i, \quad 1 \le k < K \tag{48a}$$

$$\prod_{i=1}^{K} |r_i| = \prod_{i=1}^{K} \lambda_i \tag{48b}$$

where r_i is the ith largest (in magnitude) element of $\mathbf{r}$, and $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_K > 0$, then there exists an upper triangular matrix $\mathbf{R} \in \mathbb{R}^{K\times K}$ with singular values λ_i, $1 \le i \le n$ and with $\mathbf{r}$ on the diagonal of $\mathbf{R}$.

From Theorem IV.1, we can prove the following lemma.

Lemma 1: For any rank K matrix $\mathbf{H} \in \mathbb{C}^{M_r\times M_t}$ with singular values $\lambda_{H,1} \ge \lambda_{H,2} \ge \cdots \ge \lambda_{H,K} > 0$, there exists an upper triangular matrix $\mathbf{R} \in \mathbb{R}^{K\times K}$ with identical diagonal elements

$$r_{ii} = \bar{\lambda}_H \triangleq \left(\prod_{n=1}^{K} \lambda_{H,n}\right)^{1/K}, \quad 1 \le i \le K \tag{49}$$

and semi-unitary matrices $\mathbf{Q}$ and $\mathbf{P}$ such that $\mathbf{H} = \mathbf{Q}\mathbf{R}\mathbf{P}^*$.

Proof: It can be readily verified that $\{r_{ii} = \bar{\lambda}_H\}_{i=1}^{K}$ and $\{\lambda_{H,i}\}_{i=1}^{K}$ satisfy the conditions given in (48). According to Theorem IV.1, there exists an upper triangular matrix $\mathbf{R} \in \mathbb{R}^{K\times K}$ with SVD

$$\mathbf{R} = \mathbf{U}_R \boldsymbol{\Lambda} \mathbf{V}_R^* \tag{50}$$

with the ith diagonal element $r_{ii} = \bar{\lambda}$. Here, $\boldsymbol{\Lambda}$ is a diagonal matrix whose diagonal elements are equal to $\{\lambda_{H,i}\}_{i=1}^{K}$. On the other hand

$$\mathbf{H} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^*. \tag{51}$$

Combining (50) and (51) yields $\mathbf{H} = \mathbf{U}\mathbf{U}_R^*\mathbf{R}\mathbf{V}_R\mathbf{V}^* \triangleq \mathbf{Q}\mathbf{R}\mathbf{P}^*$. ■

Hence, the matrix decomposition of (47) exists, which we refer to as the GMD since the diagonal elements of $\mathbf{R}$ are the geometric mean of $\{\lambda_{H,n}\}_{n=1}^{K}$. A computationally efficient and numerically stable algorithm is proposed in [21] to calculate the decomposition. To make this paper self-contained, we include the algorithm in the Appendix.

It seems reasonable to constrain the linear equalizer $\mathbf{Q}$ to be semi-unitary since it will keep the background noise white. Yet it seems unnecessary to constrain $\mathbf{P}$ to be semi-unitary as well. Indeed, the constraint that $\mathbf{P}$ and $\mathbf{Q}$ should be semi-unitary is in fact *inactive*, as shown in the following lemma established in [21].

Lemma 2: The GMD of (47) is also the solution to the following optimization problem with relaxed constraints:

$$\begin{aligned} \max_{\mathbf{P},\mathbf{Q}} \min \quad & \{r_{ii} : 1 \le i \le K\} \\ \text{subject to} \quad & \mathbf{R} = \mathbf{Q}^*\mathbf{H}\mathbf{P}, \quad r_{ij} = 0, \text{ for } i > j \\ & \mathbf{R} \in \mathbb{R}^{K\times K} \\ & r_{ii} > 0, \quad 1 \le i \le K \\ & \text{tr}\,(\mathbf{Q}^*\mathbf{Q}) \le K, \quad \text{tr}\,(\mathbf{P}^*\mathbf{P}) \le K. \end{aligned} \tag{52}$$

Proof: The proof is omitted. See [21] for details. ■

The GMD, which can be viewed as an extended QR decomposition, can be readily combined with the aforementioned VBLAST (GDFE) or ZFDP. GMD-VBLAST is implemented as follows: We first calculate the GMD $\mathbf{H} = \mathbf{Q}\mathbf{R}\mathbf{P}^*$. Next, we encode the information symbol $\mathbf{s}$ via the linear precoder $\mathbf{P}$ to be $\mathbf{x} = \mathbf{P}\mathbf{s}$. Then, the equivalent data model is

$$\mathbf{y} = \mathbf{Q}\mathbf{R}\mathbf{s} + \mathbf{z}. \tag{53}$$

The next step is nothing but the VBLAST decoder.

Ignoring the error propagation effect, we can regard the resulting subchannels as K independent and *identical* subchannels

$$y_i = \bar{\lambda}_H x_i + z_i, \text{ for } i = 1, \ldots, K. \tag{54}$$

The GMD-ZFDP scheme is similar to GMD-VBLAST because of the duality between VBLAST and ZFDP.

V. Performance Analyses and Implementation Issues

In this section, we first present the performance analyses of the GMD scheme from a capacity perspective, from which we demonstrate the advantages of our GMD scheme over the linear transceivers. Next, we consider combining the GMD scheme with the blind two-way channel subspace tracking in the TDD scenario. To achieve close to optimal performance at low SNR, we propose to combine GMD with subchannel selection. Finally, we discuss the relationship between our GMD scheme and [23].

A. Performance Analyses

As we have mentioned earlier, the overall BER performance of a MIMO communication system is dominated by the worst subchannels asymptotically for high SNR. Hence, the scheme optimizing the worst subchannel can enjoy the optimal BER performance for high SNR. This observation is also the motivation of the aforementioned MMD scheme. As a major advantage over the linear transceiver schemes, the GMD scheme is also asymptotically optimal in terms of the channel capacity for high SNR, as we will show below.

If the signal power is allocated evenly to the K subchannels, then based on (54), we get

$$C_{\mathrm{GMD}} = K \log_2 \left(1 + \frac{\rho}{K} \bar{\lambda}_H^2\right) \tag{55}$$

where ρ is defined in (2). The channel capacity with uniform power loading on the K subchannels is [see (26)]

$$C_{\mathrm{UPL}} = \sum_{n=1}^{K} \log_2 \left(1 + \frac{\rho}{K} \lambda_{H,n}^2\right). \tag{56}$$

It follows from (55) and (56) that

$$C_{\mathrm{UPL}} - C_{\mathrm{GMD}} = \log_2 \frac{\prod_{n=1}^{K} \left(1 + \rho \lambda_{H,n}^2\right)}{\left(1 + \rho \bar{\lambda}_H^2\right)^K}. \tag{57}$$

From (47b) and (57), we have

$$\lim_{\rho \to \infty} C_{\mathrm{UPL}} - C_{\mathrm{GMD}} = 0. \tag{58}$$

Based on Lemma II.1

$$\lim_{\rho \to \infty} C_{IT} - C_{\mathrm{UPL}} = 0. \tag{59}$$

Hence, it follows from (58) and (59) that

$$\lim_{\rho \to \infty} C_{IT} - C_{\mathrm{GMD}} = 0 \tag{60}$$

i.e., for high SNR the GMD scheme is asymptotically optimal.

Hence, the GMD scheme does not need to make the tradeoff between the information rate and BER performance as the conventional linear transceivers. Instead, our GMD scheme can achieve the optimum on both aspects simultaneously for high SNR.

As we have mentioned before, VBLAST may suffer from error propagation. Hence, the BER performance of GMD-VBLAST will be inferior to the scalar equivalence in (54). We calculate the upper bound of the GMD-VBLAST BER as follows. For a fixed SNR ρ, we assume that the system of (54) has symbol error rate (SER) P_e, i.e., each subchannel has SER P_e/K. We consider the worst case that decoding errors in some subchannels will cause the failure of the decoding in all the subsequent subchannels. The SER upper bound is readily calculated as

$$\begin{aligned} P_{e,\text{GMD-VBLAST}} &= \frac{1}{K} \sum_{n=0}^{K-1} (1 - P_e)^n (K - n) P_e \\ &< \frac{1}{K} \sum_{n=0}^{K-1} (K - n) P_e \\ &= \frac{K+1}{2} P_e. \end{aligned} \tag{61}$$

For a moderate K, say $K \leq 10$, the performance loss caused by the error propagation is rather small. For a system with high dimensionality, GMD-ZFDP is a better choice since it causes no error propagation. On the other hand, the Tomlinson–Harashima precoder leads to an input power increase of $M/(M-1)$ for M-quadrature amplitude modulation (QAM).

B. Combination of GMD With Two-Way Channel Subspace Tracking

In TDD systems, the GMD scheme may be combined with two-way channel subspace tracking techniques. Our algorithm for computing the GMD of $\mathbf{H}$, given in the Appendix, starts with the SVD. To calculate the matrix $\mathbf{P}$ [cf. (46)], we only need to know the singular values $\mathbf{\Lambda}$ and the right singular vectors $\mathbf{V}$ (cf. the Appendix). Similarly, only $\mathbf{\Lambda}$ and $\mathbf{U}$ are used to calculate $\mathbf{Q}$. Rewriting (1) with the transmitted signal $\mathbf{x}$ precoded as $\mathbf{x} = \mathbf{P}\mathbf{s}$ yields

$$\mathbf{y} = \mathbf{H}\mathbf{P}\mathbf{s} + \mathbf{z}. \tag{62}$$

Since the GMD scheme uses the same signal constellation and uniform power allocation, the covariance matrix of $\mathbf{s}$ is a scaled identity matrix, i.e., $E[\mathbf{s}\mathbf{s}^*] = \sigma_s^2 \mathbf{I}$. Hence

$$\mathbf{R}_{\mathbf{y}} = E[\mathbf{y}\mathbf{y}^*] = \mathbf{H}\mathbf{H}^* \sigma_s^2 + \sigma_z^2 \mathbf{I}. \tag{63}$$

If the signal power σ_s^2 and the noise power σ_z^2 are known *a priori*, we have $\mathbf{H}\mathbf{H}^* = \left(\mathbf{R}_{\mathbf{y}} - \sigma_z^2 \mathbf{I}\right)/\sigma_s^2$. Applying SVD to $\mathbf{H}\mathbf{H}^*$, we get

$$\mathbf{H}\mathbf{H}^* = \mathbf{U}\mathbf{\Lambda}^2\mathbf{U}^* \tag{64}$$

where $\mathbf{U}$ and $\mathbf{\Lambda}$ are the same as those defined in (51). The GMD algorithm can be applied based on $\mathbf{U}$ and $\mathbf{\Lambda}$ to get the matrices $\mathbf{Q}$ and $\mathbf{R}$, which are sufficient for decoding. If a TDD system is used, the reverse channel, where the roles of previous transmitter and receiver are exchanged, can be modeled as

$$\mathbf{y}_{\mathrm{rev}} = \mathbf{H}^T \mathbf{Q}^* \mathbf{s}_{\mathrm{rev}} + \mathbf{z}_{\mathrm{rev}} \tag{65}$$

where $(\cdot)^T$ denotes the transpose, and the subscript "rev" means "reverse channel." Define

$$\mathbf{R}_{\mathbf{y}_{\mathrm{rev}}} = E\left[\bar{\mathbf{y}}_{\mathrm{rev}} \mathbf{y}_{\mathrm{rev}}^T\right] \tag{66}$$

where $\bar{\mathbf{y}}$ denotes the complex conjugate of $\mathbf{y}$. Using the similar argument, we have

$$\mathbf{H}^*\mathbf{H} = \mathbf{V}\mathbf{\Lambda}^2\mathbf{V}^*. \tag{67}$$

Then, the *reverse* receiver, i.e., the previous transmitter, can calculate $\mathbf{R}$ and $\mathbf{P}$ from $\mathbf{V}$ and $\mathbf{\Lambda}$. Channel subspace tracking techniques (see, e.g., [33] and [34]) can be used to estimate $\mathbf{U}$, $\mathbf{V}$, and $\mathbf{\Lambda}$ efficiently. Hence, our GMD scheme can be applied without the need to use training symbols for channel estimation. We note that this merit of GMD is not shared by the conventional transceiver schemes introduced in Section I since all those methods allocate different powers to different subchannels, which makes it difficult, if not impossible, to estimate the singular values in $\mathbf{\Lambda}$. Of course, if the same power is allocated to each eigen-subchannel, this blind two-way channel subspace tracking idea can also be combined with the SVD-based schemes at the cost of significant capacity loss.

The GMD scheme can be made backward compatible with the TDD systems using VBLAST decoders. By using CSIT or blind subspace tracking techniques, the transmitter can calculate the linear precoder $\mathbf{P}$. Hence, it can always precode the transmitted data $\mathbf{s}$ to be $\mathbf{x} = \mathbf{P}\mathbf{s}$, even when sending the training data. Thus, the receiver is "fooled" to believe that the channel is the virtual one $\mathbf{H}_{vt} = \mathbf{H}\mathbf{P} = \mathbf{Q}\mathbf{R}$. Although the linear precoder $\mathbf{P}$ is made transparent to the VBLAST decoder, the decoder still enjoys the multiple *identical* subchannels due to the linear precoder $\mathbf{P}$.

C. Subchannel Selection

The previous discussion is based on the assumption that all the subchannels corresponding to positive singular values are used for signal transmission. However, in practical scenarios, some of the positive singular values of the channel matrix $\mathbf{H}$ can be very small. This situation occurs for spatially correlated flat-fading channels, or even i.i.d. Rayleigh flat-fading channels with $M_r \approx M_t \gg 1$. From (47b), we see that it will influence the overall channel quality, and hence, subchannel selection is helpful. The other situation where subchannel selection is needed is the case when the input power is low or moderate. In this section, we propose a simple algorithm to select the subchannels, which is numerically verified to be able to achieve near optimal capacity, even at low SNR.

Let us sort the singular values of $\mathbf{H}$ as $\lambda_{H,1} \geq \lambda_{H,2} \cdots \geq \lambda_{H,K} > 0$. If GMD is constrained to the first $n \leq K$ eigen subchannels, we obtain n identical subchannels

$$y_i = \bar{\lambda}_n x_i + z_i, \text{ for } i = 1, \ldots, n \tag{68}$$

where

$$\bar{\lambda}_n = \sqrt[n]{\prod_{i=1}^{n} \lambda_{H,i}}. \tag{69}$$

To maximize the channel throughput with our GMD scheme, we need to solve the following problem:

$$\max_{1 \leq n \leq K} n \log \left(1 + \sqrt[n]{\prod_{i=1}^{n} \lambda_{H,i}^2} \frac{\rho}{n} \right) \tag{70}$$

or

$$\max_{1 \leq n \leq K} \left(1 + \sqrt[n]{\prod_{i=1}^{n} \lambda_{H,i}^2} \frac{\rho}{n} \right)^n. \tag{71}$$

The solution to this problem is straightforward. We can use either linear search or bisection method to find the optimal n.

Several remarks are in order. i) It is straightforward to incorporate the channel selection into the GMD algorithm. In the Appendix, we show that GMD starts from SVD $\mathbf{H} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^*$ and then applies a series of Givens transformation to $\mathbf{\Lambda}$ to make it upper triangular (see Appendix). The Givens transformation can be constrained to the first $n \leq K$ diagonal elements of $\mathbf{\Lambda}$. ii) The blind channel subspace tracking can be combined with the subchannel selection strategy seamlessly. If only the subchannels corresponding to the largest $n < K$ singular values are selected, the blind channel tracking technique will track the n dimensional subspace automatically. iii) The performance loss of the GMD scheme at a low SNR region is due to the well-known fact that the zero-forcing equalizer is inherently suboptimal. In a subsequent paper, we propose a so-called uniform channel decomposition (UCD) scheme, which can decompose a MIMO channel into multiple identical subchannels in a *strictly* capacity lossless manner [35].

D. Further Remarks

When this paper was under review, we noticed [23], in which an idea similar to GMD was proposed to approach the performance of the ML detector in the ISI suppression scenario. For a SISO ISI channel, if symbols are precoded and transmitted in a block manner, then the data model (15) can be used to represent the received block data. Note that for this case, $\mathbf{H}$ is a Toeplitz matrix due to the time-invariant property of the ISI channel. A linear precoder design $\mathbf{F}$ was proposed in [23] such that the virtual channel $\mathbf{H}_{vt} = \mathbf{HF}$ can be decomposed, via QR decomposition, to be $\mathbf{H}_{vt} = \mathbf{QR}$, where $\mathbf{R}$ has equal diagonal elements. We see that this equal diagonal idea is equivalent to GMD. However, our GMD scheme, which is independently motivated by the MIMO transceiver design problem, has several major advantages over the algorithm in [23].

1) Our GMD scheme represents a paradigm shift from the conventional linear transceiver designs to nonlinear de-

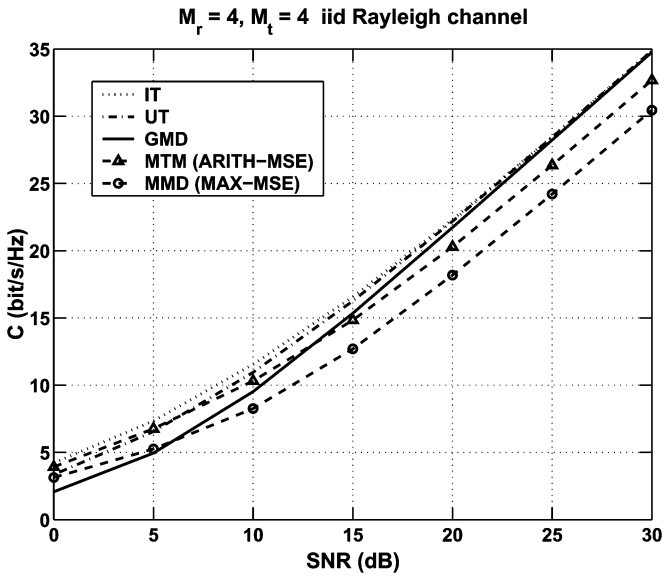


Fig. 1. Average capacity over 1000 Monte Carlo trials versus SNR with $M_t = 4$ and $M_r = 4$ for i.i.d. Rayleigh flat-fading channels.

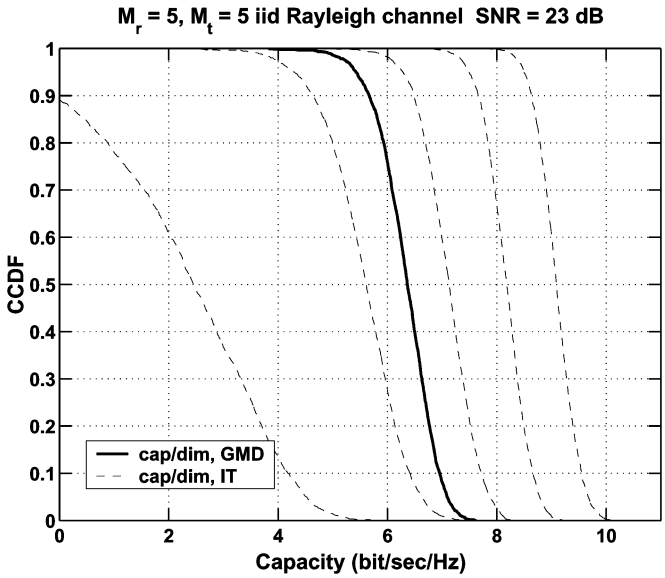


Fig. 2. Complementary cumulative distribution functions of the capacities of five subchannels of the i.i.d. Rayleigh flat-fading channel with $M_t = 5$ and $M_r = 5$. Results based on 2000 Monte Carlo trials.

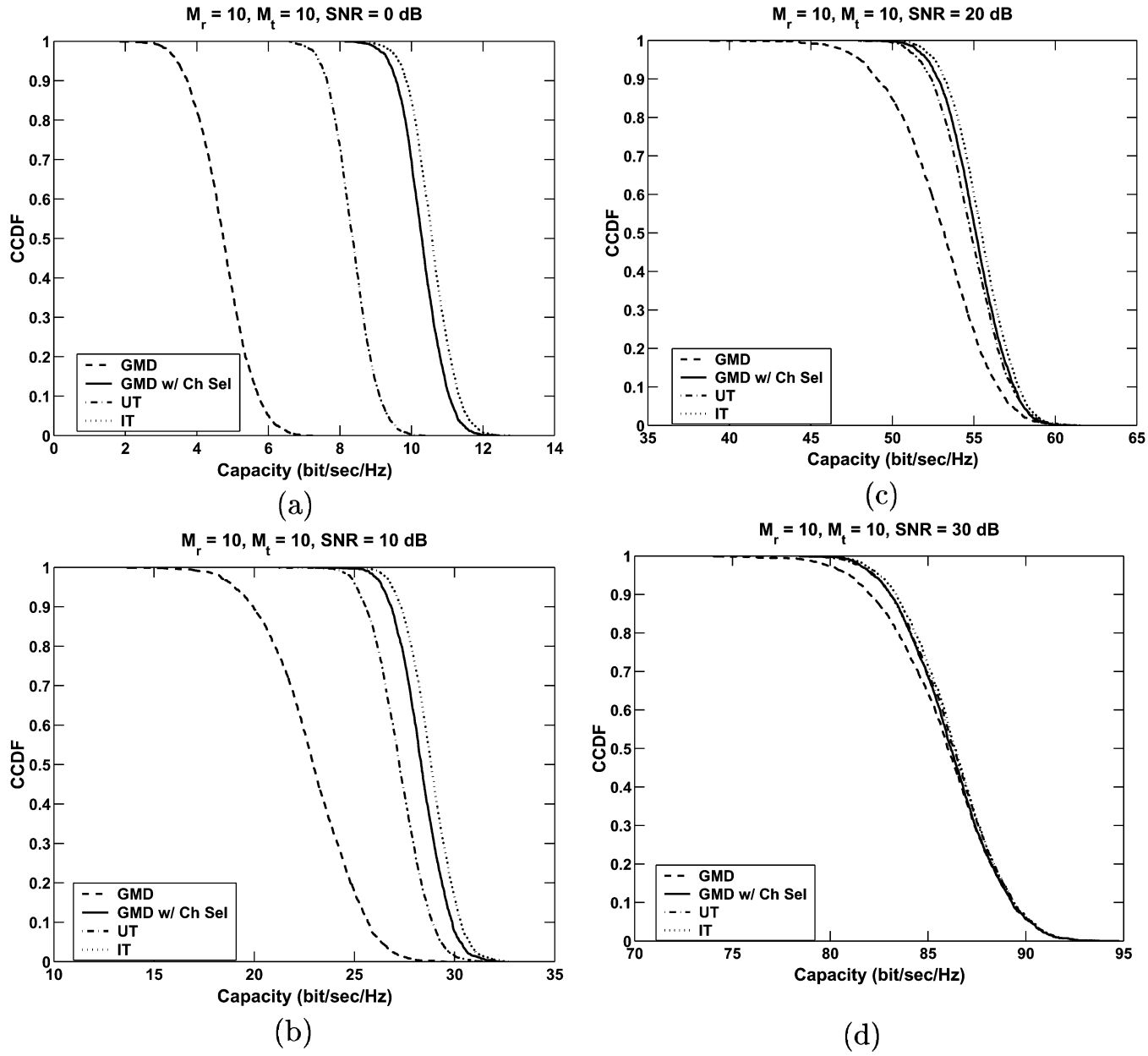


Fig. 3. Complementary cumulative distribution function of the capacity of an i.i.d. Rayleigh flat fading channel with $M_t = 10$ and $M_r = 10$. Results based on 1000 Monte Carlo trials. SNR = (a) 0 dB, (b) 10 dB, (c) 20 dB, and (d) 30 dB.

signs and can be proven, both numerically and theoretically, to have superior performance from both BER and information theoretic aspects.

2) Our GMD algorithm is computationally much more efficient than that of [23]. Both algorithms start from the SVD of $\mathbf{H}$, which is followed by $K-1$ iterations. The GMD involves $2K-2$ fast Givens rotations. For a channel $\mathbf{H}$ with $M_t = M_r = K$, the SVD requires $O(K^3)$ flops, whereas the GMD requires additional $O(K^2)$ flops. Thus, the computational complexity of the GMD scheme is comparable to the conventional linear transceiver schemes. However, the algorithm in [23] involves multiplications and inversions of matrices in each iteration, and the overall computational burden turns out to be an additional $O(K^4)$ flops.

3) For the GMD algorithm, only the information for $\mathbf{HH}^*$, and hence $\mathbf{\Lambda}$ and $\mathbf{U}$, are needed to calculate $\mathbf{Q}$. However, for the algorithm in [23], the equalizer needs to know both the precoder $\mathbf{F}$ and $\mathbf{H}$ and, hence, $\mathbf{H}_{vt} = \mathbf{HF}$, in order to apply the traditional QR to $\mathbf{H}_{vt}$. Hence, it cannot be combined with the aforementioned blind two-way channel subspace tracking algorithm introduced in Section V-B.

4) The techniques proposed in the GMD algorithm, i.e., permutations and Givens rotations, can be used to achieve any possible upper triangular decomposition predicted by Theorem IV.1, which we refer to as the generalized triangular decomposition (GTD) [36]. The GTD can be used in MIMO transceiver designs with QoS constraints [37].

Like the algorithm in [23], the GMD scheme can also be combined with orthogonal frequency division multiplexing (OFDM) for ISI suppression. For a SISO ISI channel with memory L

$$y(n) = \sum_{l=0}^{L-1} h_l x(n-l) + z(n) \tag{72}$$

after applying OFDM with block length N, we get a MIMO channel

$$\mathbf{y} = \mathbf{D}\mathbf{x} + \mathbf{z} \tag{73}$$

where $\mathbf{D}$ is a diagonal matrix with the diagonal elements equal to the N-point fast Fourier transform (FFT) of $\mathbf{h} = [h_0, h_1, \ldots, h_{L-1}]^T$. Hence, the GMD scheme can be applied directly. We expect that GMD-ZFDP may have better BER performance than GMD-VBLAST if $N \gg 1$, in which case, the GMD-VBLAST may suffer from considerable performance degradation due to error propagations.

VI. Performance Examples

We present next several numerical examples to demonstrate the effectiveness of the GMD scheme. In all the examples, we assume Rayleigh independent flat-fading channels.

In the first example, we consider a Rayleigh flat-fading channel with $M_t = 4$ and $M_r = 4$. We compute the Shannon capacities of the channel with both CSIR and CSIT [C_{IT}, (5)], the channel with uninformed transmitter [C_{UT}, (7)], the channel using the GMD scheme [C_{GMD}, (55)[, the channel using the MTM scheme [C_{MTM}, (25)], and the channel using the MMD scheme [C_{MMD}, (31)]. We average the capacities of 1000 Monte Carlo-generated $\mathbf{H}$ realizations. The result is presented in Fig. 1. We note that the capacity loss of the MMD scheme is about twice that of the MTM scheme at high SNR, as predicted in Section III. The relative capacity loss of the MMD scheme compared with MTM is smaller at low SNR because some subchannels are not used at low SNR. The GMD scheme outperforms the linear transceiver designs when the SNR is moderate or high and is asymptotically capacity lossless at high SNR.

Fig. 2 shows the complementary cumulative distribution functions (CCDFs) of the channel capacities of a 5×5 independent Rayleigh flat-fading channel with SNR equal to 23 dB. The five thin dashed curves denote the channel capacities of the five subchannels obtained via SVD plus water filling. Note that the left-most thin curve crosses the vertical axis at a value less than one, which means that the worst subchannel (corresponding to the smallest singular value of the channel matrix) is sometimes discarded by water filling. The thick line is the CCDF of each subchannel capacity obtained via GMD. Fig. 2 further illustrates the disadvantages of the conventional "SVD plus bit allocation" scheme (see, e.g., [12], [13], and [16]). The channel capacities of the five subchannels obtained via SVD plus water filling range from 0 to about 10 bit/s/Hz, which suggests that the binary phase shift keying (BPSK) or quaternary phase shift keying (QPSK) modulation should be used to match the capacity of the worst subchannel and something like 512 or 1024 QAM to the best subchannel. This bit allocation significantly increases the modulation/demodulation complexity.

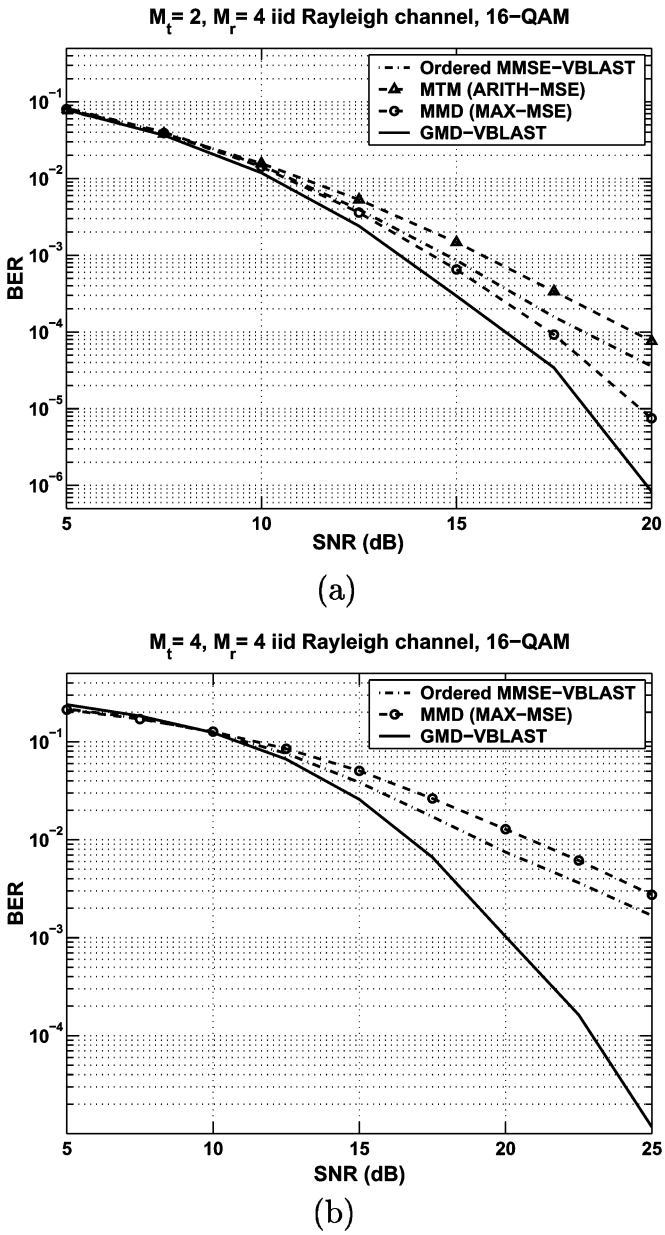


Fig. 4. BER performance averaged over 1000 Monte Carlo trials of i.i.d. Rayleigh flat-fading channel versus SNR with (a) $M_t = 2$ and $M_r = 4$ and (b) $M_t = 4$ and $M_r = 4$.

Moreover, using a constellation with size greater than 256 is impractical for the current RF circuit design technology. For the GMD scheme, on the other hand, the same constellation with a moderate size, say 64-QAM, can be applied to reap most of the channel capacity.

To demonstrate the effectiveness of the subchannel selection approach, we consider a 10×10 independent Rayleigh flat-fading channel. The channel is usually ill-conditioned since some singular values of $\mathbf{H}$ are very close to zero. Without the subchannel selection strategy, GMD suffers from performance degradation, especially at low SNR, as seen in Fig. 3. On the other hand, with the subchannel selection scheme, there is only about 0.2 bit/sec/Hz rate loss compared with the C_{IT}, even at very low SNR.

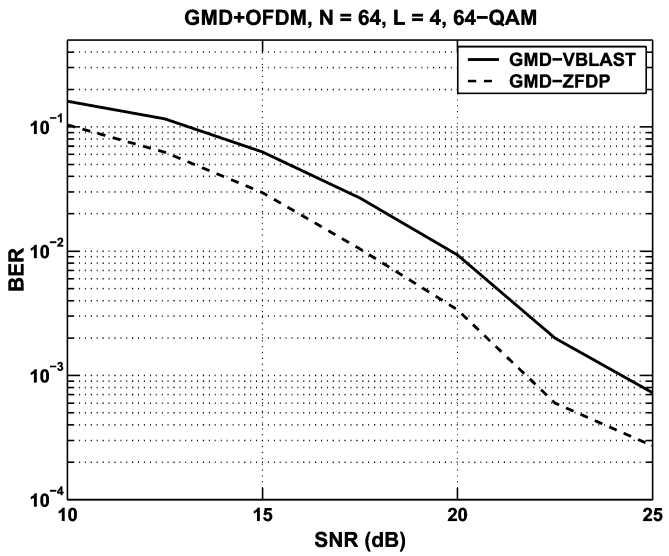


Fig. 5. BER performances of GMD-VBLAST and GMD-ZFDP. Both are combined with OFDM for ISI suppression.

We compare the BER performance of the GMD-VBLAST scheme with the unprecoded MMSE-VBLAST scheme with the optimal detection ordering, the MTM scheme, and the MMD scheme. No error correcting code is used in the simulations. In Fig. 4(a), $\mathbf{H} \in \mathbb{C}^{4\times 2}$ has identically independent Rayleigh fading elements. Hence, the channel matrix is usually well-conditioned. Two independent symbol streams modulated as 16-QAM are transmitted. The figure is obtained by averaging 1000 Monte Carlo trials of $\mathbf{H}$. We see that the GMD scheme has more than a 1-dB improvement over the MMD scheme at moderate to high SNR. In Fig. 4(b), $\mathbf{H} \in \mathbb{C}^{4\times 4}$ usually has a large condition number, in which case, the MMD scheme is subject to more capacity loss, as analyzed in Section III. Four independent symbol streams are transmitted. The BER performance of the GMD scheme is much better than the others. We did not include MTM because it discards some bad subchannels and, hence, cannot be used to transmit four independent data streams.

In the final example, we combine the GMD scheme with 64-point FFT-based OFDM for ISI suppression in a SISO channel (see Fig. 5). We assume that the channel responses $h_l, l = 0, 1, \ldots, L-1$ are independent zero-mean circularly symmetric Gaussian random variables with unit variance. The channel length is $L = 4$. The GMD-ZFDP is about 2 dB better than GMD-VBLAST. This is because GMD-VBLAST suffers from considerable error propagation effect. This result suggests that GMD-ZFDP may be preferred over GMD-VBLAST if the channel has a large dimensionality.

VII. Conclusions

We have introduced a novel joint transceiver design scheme when the CSI is available at both the transmitter and receiver of a MIMO communication system. We show that the geometric mean decomposition (GMD), combined with the VBLAST decoder or dirty paper precoder, can decompose a MIMO channel into multiple *identical* scalar subchannels. This desirable property can bring about much convenience to the practical system design, particularly the symbol constellation selection. Moreover, we have shown that the GMD scheme is optimal asymptotically for high SNR in terms of both information rate and BER performance, whereas the computational complexity of our scheme is comparable with the conventional linear transceiver scheme. Furthermore, we have shown that the GMD scheme can be applied without the need to use training symbols for channel estimation if combined with subspace tracking techniques. We have also considered the issue of subchannel selection when some of the subchannels are too poor to be useful. The GMD scheme can also be combined with OFDM for ISI suppression. Both the theoretical analyses and empirical simulations have been provided to validate the effectiveness of our approaches.

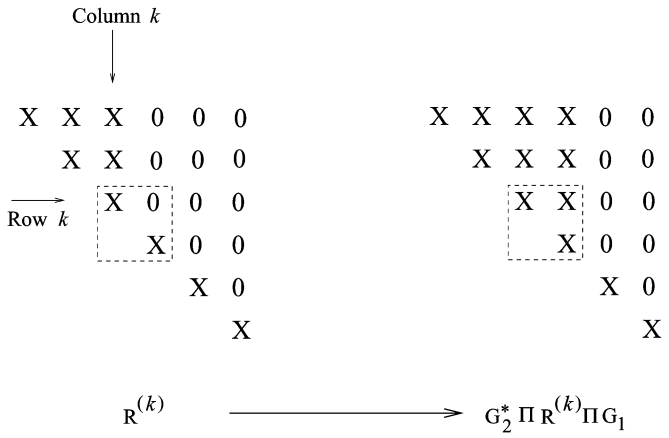


Fig. 6. Operation displayed in (74).

Appendix
Geometric Mean Decomposition

We now give an algorithm that evaluates the GMD that starts with the SVD $\mathbf{H} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^*$. The algorithm generates a sequence of upper triangular matrices $\mathbf{R}^{(\tilde{K})}$, $1 \le \tilde{K} < K$, with $\mathbf{R}^{(1)} = \mathbf{\Lambda}$. Each matrix $\mathbf{R}^{(\tilde{K})}$ has the following properties.

a) $r_{ij}^{(\tilde{K})} = 0$ when $i > j$ or $j > \max\{\tilde{K}, i\}$.
b) $r_{ii}^{(\tilde{K})} = \bar{\lambda}_H$ for all $i < \tilde{K}$, and the geometric mean of $r_{ii}^{(\tilde{K})}$, $\tilde{K} \le i \le K$ is $\bar{\lambda}_H$.

We express $\mathbf{R}^{(k+1)} = \mathbf{Q}_k^T \mathbf{R}^{(k)} \mathbf{P}_k$, where $\mathbf{Q}_k$ and $\mathbf{P}_k$ are orthogonal for each k.

These orthogonal matrices are constructed using a symmetric permutation and a pair of Givens rotations. Suppose that $\mathbf{R}^{(k)}$ satisfies a) and b). If $r_{kk}^{(k)} \ge \bar{\lambda}_H$, then let $\mathbf{\Pi}$ be a permutation matrix with the property that $\mathbf{\Pi}\mathbf{R}^{(k)}\mathbf{\Pi}$ exchanges the $(k+1)$st diagonal element of $\mathbf{R}^{(k)}$ with any element r_{pp}, $p > k$, for which $r_{pp} \le \bar{\lambda}_H$. If $r_{kk}^{(k)} < \bar{\lambda}_H$, then let $\mathbf{\Pi}$ be chosen to exchange the $(k+1)$st diagonal element with any element r_{pp}, $p > k$ for which $r_{pp} \ge \bar{\lambda}_H$. Let $\delta_1 = r_{kk}^{(k)}$ and $\delta_2 = r_{pp}^{(k)}$ denote the new diagonal elements at locations k and $k+1$ associated with the permuted matrix $\mathbf{\Pi}\mathbf{R}^{(k)}\mathbf{\Pi}$.

Next, we construct orthogonal matrices $\mathbf{G}_1$ and $\mathbf{G}_2$ by modifying the elements in the identity matrix that lie at the intersection of rows k and $k+1$ and columns k and $k+1$. We multiply the permuted matrix $\mathbf{\Pi}\mathbf{R}^{(k)}\mathbf{\Pi}$ on the left by $\mathbf{G}_2^T$ and on the right

$$\underset{(\mathbf{G}_2^T)}{\bar{\lambda}_H^{-1}\begin{bmatrix} c\delta_1 & s\delta_2 \\ -s\delta_2 & c\delta_1 \end{bmatrix}} \quad \underset{(\mathbf{\Pi}\mathbf{R}^{(k)}\mathbf{\Pi})}{\begin{bmatrix} \delta_1 & 0 \\ 0 & \delta_2 \end{bmatrix}} \quad \underset{(\mathbf{G}_1)}{\begin{bmatrix} c & -s \\ s & c \end{bmatrix}} = \underset{(\mathbf{R}^{(k+1)})}{\begin{bmatrix} \bar{\lambda}_H & x \\ 0 & y \end{bmatrix}}. \tag{74}$$

by $\mathbf{G}_1$. These multiplications will change the elements in the 2 by 2 submatrix at the intersection of rows k and $k+1$ with columns k and $k+1$. Our choice for the elements of $\mathbf{G}_1$ and $\mathbf{G}_2$ is shown below, where we focus on the relevant 2 by 2 submatrices of $\mathbf{G}_2^T$, $\mathbf{\Pi}\mathbf{R}^{(k)}\mathbf{\Pi}$, and $\mathbf{G}_1$, as in (74), shown at the top of the page. If $\delta_1 = \delta_2 = \bar{\lambda}_H$, we take $c = 1$ and $s = 0$; if $\delta_1 \neq \delta_2$, we take

$$c = \sqrt{\frac{\bar{\lambda}_H^2 - \delta_2^2}{\delta_1^2 - \delta_2^2}} \text{ and } s = \sqrt{1 - c^2}. \tag{75}$$

Since $\bar{\lambda}_H$ lies between δ_1 and δ_2, s and c are non-negative real-valued scalars.

Fig. 6 depicts the transformation from $\mathbf{R}^{(k)}$ to $\mathbf{G}_2^T\mathbf{\Pi}\mathbf{R}^{(k)}\mathbf{\Pi}\mathbf{G}_1$. The dashed box is the 2 by 2 submatrix displayed in (74). Note that c and s, which are defined in (75), are real-valued scalars that are chosen so that

$$c^2 + s^2 = 1 \text{ and } (c\delta_1)^2 + (s\delta_2)^2 = \bar{\lambda}_H^2.$$

With these identities, the validity of (74) follows by direct computation. Defining $\mathbf{Q}_k = \bar{\mathbf{\Pi}}\mathbf{G}_2$ and $\mathbf{P}_k = \bar{\mathbf{\Pi}}\mathbf{G}_1$, we set

$$\mathbf{R}^{(k+1)} = \mathbf{Q}_k^T\mathbf{R}^{(k)}\mathbf{P}_k. \tag{76}$$

It follows from Fig. 6, (74), and the identity $|\mathbf{R}^{(k+1)}| = |\mathbf{R}^{(k)}|$ that a) and b) hold for $\tilde{K} = k+1$. Thus, there exists a real-valued upper triangular matrix $\mathbf{R}^{(K)}$, with $\bar{\lambda}_H$ on the diagonal, and unitary matrices $\mathbf{Q}_i$ and $\mathbf{P}_i$, $i = 1, 2, \ldots, K-1$ such that

$$\mathbf{R}^{(K)} = (\mathbf{Q}_{K-1}^T \ldots \mathbf{Q}_2^T\mathbf{Q}_1^T)\mathbf{\Lambda}(\mathbf{P}_1\mathbf{P}_2 \ldots \mathbf{P}_{K-1}).$$

Combining this identity with the SVD, we obtain $\mathbf{H} = \mathbf{Q}\mathbf{R}\mathbf{P}^*$, where

$$\mathbf{Q} = \mathbf{U}\left(\prod_{i=1}^{K-1}\mathbf{Q}_i\right), \quad \mathbf{R} = \mathbf{R}^{(K)}, \text{ and } \mathbf{P} = \mathbf{V}\left(\prod_{i=1}^{K-1}\mathbf{P}_i\right).$$

A Matlab implementation of this algorithm for the GMD is posted at the following website: http://www.sal.ufl.edu/yjiang/papers/gmd.m.

Given the SVD, this algorithm for the GMD requires $O((M_r + M_t)K)$ flops. For comparison, a reduction of $\mathbf{H}$ to bidiagonal form by the Golub–Kahan bidiagonalization scheme [38], which is often the first step in the computation of the SVD, requires $O(M_rM_tK)$ flops.

REFERENCES

[1] I. E. Telatar, "Capacity of multiple antenna Gaussian channels," *AT&T Tech. Memo.*, Jun. 1995.

[2] G. J. Foschini Jr, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, vol. 1, pp. 41–59, 1996.

[3] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Tran. Telecommun.*, vol. 10, no. 6, pp. 585–595, 1999.

[4] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, no. 3, pp. 311–335, 1999.

[5] S. M. Alamouti, "A simple transmit diversity techniques for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 10, pp. 1451–1458, Oct. 1998.

[6] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communications: Performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

[7] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, no. 4, pp. 1456–1467, Jul. 1999.

[8] G. J. Foschini, G. D. Golden, R. A. Valenzuela, and P. W. Wolniansky, "Simplified processing for high spectral efficiency wireless communication employing multiple-element arrays," *Wireless Pers. Commun.*, vol. 6, pp. 311–335, Mar. 1999.

[9] G. J. Foschini, D. Chizhik, M. J. Gans, C. Papadias, and R. A. Valenzuela, "Analysis and performance of some basic space time architectures," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 4, pp. 303–320, Apr. 2003.

[10] . [Online]. Available: http://www.3gpp.org

[11] J. Yang and S. Roy, "On joint transmitter and receiver optimization for multiple-input-multiple-output (MIMO) transmission systems," *IEEE Trans. Commun.*, vol. 42, no. 12, pp. 3221–3231, Dec. 1994.

[12] G. G. Raleigh and J. M. Cioffi, "Spatial-temporal coding for wireless communication," *IEEE Trans. Commun.*, vol. 46, no. 3, pp. 357–366, Mar. 1998.

[13] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Filterbank transceiver optimizating information rate in block transmissions over dispersive channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 1019–1032, Apr. 1999.

[14] ——, "Redundant filterbank precoders and equalizers part I: Unification and optimal designs," *IEEE Trans. Signal Process.*, vol. 47, no. 7, pp. 1988–2006, Jul. 1999.

[15] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2198–2206, Dec. 2001.

[16] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1051–1064, May 2002.

[17] E. Onggosanusi, A. Sayeed, and B. V. Veen, "Optimal antenna diversity signaling for wideband space-time wireless channels utilizing channel state information," *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 341–353, Feb. 2002.

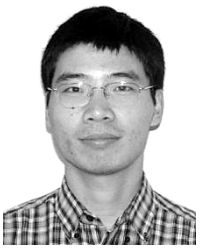
[18] ——, "Efficient signaling schemes for wideband space-time wireless channels using channel state information," *IEEE Trans. Veh. Technol.*, vol. 52, no. 1, pp. 1–13, Jan. 2003.

[19] D. Palomar, J. Cioffi, and M. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2381–2401, Sep. 2003.

[20] D. Palomar and M. Lagunas, "Joint transmit-receive space-time equalization in spatially correlated MIMO channels: A beamforming approach," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 6, pp. 730–743, Jun. 2003.

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

[22] G. Ginis and J. M. Cioffi, "On the relationship between V-BLAST and the GDFE," *IEEE Commun. Lett.*, vol. 5, no. 9, pp. 364–366, Sep. 2001.
[23] J.-K. Zhang, A. Kavčić, X. Ma, and K. M. Wong, "Design of unitary precoders for ISI channels," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. III, Orlando, FL, 2002, pp. 2265–2268.
[24] D. W. Bliss, K. W. Forsythe, A. O. Hero, and A. F. Yegulalp, "Environmental issues for MIMO capacity," *IEEE Trans. Signal Process.*, vol. 50, no. 9, pp. 2128–2142, Sep. 2002.
[25] G. Ginis and J. M. Cioffi, "Vectored transmission for digital subscriber line systems," *IEEE J. Sel. Areas Commun.*, vol. 20, pp. 1085–1104, Jun. 2002.
[26] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1691–1706, Jul. 2003.
[27] G. Ginis and J. M. Cioffi, "A multi-user precoding scheme achieving crosstalk cancellation with application to DSL systems," in *Proc. 34th Asilomar Conf. Signals, Syst., Comput.*, vol. 2, Asilomar, CA, Oct.–Nov. 29–1, 2000, pp. 1627–1631.
[28] M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. IT–29, pp. 439–441, May 1983.
[29] M. Tomlinson, "New automatic equaliser employing modulo arithmetic," *Electron. Lett.*, vol. 7, pp. 138–139, Mar. 1971.
[30] H. Harashima and H. Miyakawa, "Matched-transmission technique for channels with intersymbol interference," *IEEE Trans. Commun.*, vol. COM-20, pp. 774–780, Aug. 1972.
[31] W. Yu and J. M. Cioffi, "Trellis precoding for the broadcast channel," in *Proc. Global Telecommun. Conf.*, vol. 2, Nov. 2001, pp. 1344–1348.
[32] A. Horn, "On the eigenvalues of a matrix with prescribed singular values," *Proc. Amer. Math. Soc.*, vol. 5, pp. 4–7, 1954.
[33] B. C. Banister and J. Zeidler, "Feedback assisted transmission subspace tracking for MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 4, pp. 452–463, Apr. 2003.
[34] A. Poon, D. Tse, and R. W. Brodersen, "An adaptive multiantenna transceiver for slowly flat fading channels," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1820–1827, Nov. 2003.
[35] Y. Jiang and J. Li, "Uniform channel decomposition for MIMO communications," in *Proc. 38th Asilomar Conf. Signals, Syst., Comput.*, Asilomar, CA, Nov. 7–10, 2004.
[36] Y. Jiang, W. Hager, and J. Li, "The generalized triangular decomposition," SIAM J. Matrix Anal. Applicat., submitted for publication.
[37] Y. Jiang, J. Li, and W. Hager, "Transceiver design using generalized triangular decomposition for MIMO communications with QoS constraints," in *Proc. 38th Asilomar Conf. Signals, Syst., Comput.*, Asilomar, CA, Nov. 7–10, 2004.
[38] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD: Johns Hopkins Univ. Press, 1983.



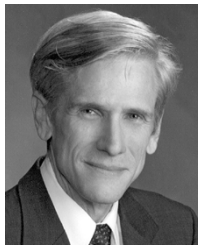
Yi Jiang (S'02) received the B.S. degree in electrical engineering and information science from the University of Science and Technology of China (USTC), Hefei, in 2001 and the M.S. and Ph.D. degrees from the University of Florida, Gainesville, in 2003 and 2005, respectively, both in electrical engineering.

In the summer of 2005, he was a research consultant with Information System Technology, Inc. (ISTI), Fort Collins, CO. He is now a Postdoctoral researcher with the University of Colorado, Boulder. His research interests are in the areas of signal processing, wireless communications, and information theory.

Jian Li (S'87–M'91–SM'97–F'05) received the M.Sc. and Ph.D. degrees in electrical engineering from The Ohio State University (OSU), Columbus, in 1987 and 1991, respectively.

From April 1991 to June 1991, she was an Adjunct Assistant Professor with the Department of Electrical Engineering, OSU. From July 1991 to June 1993, she was an Assistant Professor with the Department of Electrical Engineering, University of Kentucky, Lexington. Since August 1993, she has been with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, where she is currently a Professor. Her current research interests include spectral estimation, array signal processing, sensor networks, machine learning, and their applications.

Dr. Li is a Fellow of the IEE. She is a member of Sigma Xi and Phi Kappa Phi. She received the 1994 National Science Foundation Young Investigator Award and the 1996 Office of Naval Research Young Investigator Award. She was an Executive Committee Member of the 2002 International Conference on Acoustics, Speech, and Signal Processing, Orlando, FL, May 2002. She was an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 1999 to 2005 and has been an Associate Editor of the IEEE SIGNAL PROCESSING MAGAZINE since 2003. She is presently a member of two of the IEEE Signal Processing Society technical committees: the Signal Processing Theory and Methods (SPTM) Technical Committee and the Sensor Array and Multichannel (SAM) Technical Committee.



William W. Hager received the B.S. degree in mathematics from Harvey Mudd College, Claremont, CA, in 1970. He was awarded a National Science Foundation Fellowship for graduate studies in mathematics at the Massachusetts Institute of Technology, Cambridge, where he received the Masters and Ph.D. degrees in 1971 and 1974, respectively.

He was an Assistant Professor at the University of South Florida, Tampa, from 1974 to 1976, an Assistant Professor at Carnegie Mellon University, Pittsburgh, PA, from 1976 to 1980, an Associate Professor from 1980 to 1986 and a Professor from 1986 to 1988 at Pennsylvania State University, University Park, and a Professor at the the University of Florida, Gainesville, from 1988 to the present. He has been editor-in-chief of *Computational Optimization and Applications* since 1992 and co-Director of the Center for Applied Optimization at the University of Florida since 1992. He was Program Director of the SIAM Activity Group on Control and System Theory from 1998 to 2001.

Dr. Hager was co-chair of the 2001 SIAM Conference on Control and Its Applications.