MIMO Transceivers With Decision Feedback and Bit Loading: Theory and Optimization

Ching-Chih Weng, *Student Member, IEEE*, Chun-Yang Chen, *Student Member, IEEE*, and P. P. Vaidyanathan, *Fellow, IEEE*

***Abstract*—This paper considers MIMO transceivers with linear precoders and decision feedback equalizers (DFEs), with bit allocation at the transmitter. Zero-forcing (ZF) is assumed. Considered first is the minimization of transmitted power, for a given total bit rate and a specified set of error probabilities for the symbol streams. The precoder and DFE matrices are optimized jointly with bit allocation. It is shown that the generalized triangular decomposition (GTD) introduced by Jiang, Li, and Hager offers an optimal family of solutions. The optimal linear transceiver (which has a linear equalizer rather than a DFE) with optimal bit allocation is a member of this family. This shows formally that, under optimal bit allocation, linear and DFE transceivers achieve the same minimum power. The DFE transceiver using the geometric mean decomposition (GMD) is another member of this optimal family, and is such that optimal bit allocation yields identical bits for all symbol streams—no bit allocation is necessary—when the specified error probabilities are identical for all streams. The QR-based system used in VBLAST is yet another member of the optimal family and is particularly well-suited when limited feedback is allowed from receiver to transmitter. Two other optimization problems are then considered: a) minimization of power for specified set of bit rates and error probabilities (the QoS problem), and b) maximization of bit rate for fixed set of error probabilities and power. It is shown in both cases that the GTD yields an optimal family of solutions.**

***Index Terms*—BER optimization, bit allocation, decision feed-back equalizers, generalized triangular decomposition, limited feedback, MIMO transceivers.**

I. INTRODUCTION

IN this paper we consider MIMO transceivers with a linear precoder and a decision feedback equalizer (DFE), with bit allocation allowed at the transmitter end. The focus is on the joint optimization of the precoder, the DFE matrices, and bit allocation. Zero-forcing (ZF) and QAM signaling are considered throughout, and the perfect channel information is assumed to be known to the transmitter and the receiver.

While this joint optimization has not been addressed in the past, a variety of related transceiver designs have been studied previously. When the transmitter has perfect channel information, four major optimization problems have been considered. First, for fixed precoder and DFE matrices, the optimal bit loading problem has been studied in [3]. Second, for the case where the bit allocation is fixed to be uniform, joint optimization of the precoder and DFE matrices is a well studied problem [10]–[12], [22], [26], [31], [32]. Third, for the case of linear transceivers, the joint optimization of the precoder, the (linear) equalizer, and bit allocation has been studied in [17] (under ZF constraint), and in [21] (without ZF constraint). Fourth, if the precoding matrix is restricted to be a diagonal matrix where only power loading applies, the optimization problem of rate and power allocation for the systems with DFE receiver has been discussed in [4]. If no perfect channel state information is present (only channel statistics known at the transmitter), the optimization of power and rate allocation for the system with DFE receiver was addressed in [24].

As summarized above, bit loading, precoder, and receiver design optimizations have been studied extensively. However, current literature lacks a discussion that reviews bit loading, linear precoder, and DFE *jointly* when perfect channel information is available at both ends of the transceiver. To begin solving this problem, we start with the minimization of transmitted power for a specified set of error probabilities for the symbol streams. We show that the generalized triangular decomposition (GTD) introduced in [12] offers an optimal solution. The GTD in fact gives rise to a family of solutions, with the bit allocation details changing from solution to solution. We will see in particular that the optimal linear transceiver with optimal bit allocation, which has a *linear* equalizer rather than a DFE, is a member of this family of solutions. This shows formally that, under optimal bit allocation, optimum linear transceivers achieve the same transmitted power as optimum DFE transceivers with bit allocation. These discussions assume that the bit allocation formula is realizable (i.e, the bits are nonnegative integers). The DFE transceiver based on the geometric mean decomposition (GMD) [10] is another member of the above family of optimal solutions, and is such that the optimal bit allocation formula yields *identical* bits for all symbol substreams, when the specified error probabilities are identical for the substreams. DFE with GMD, therefore, achieves minimum power even without the need for bit allocation. In a way this complements one of the results in [14], namely, when all symbol streams are constrained to have identical bits, the average biterror rate (BER) for fixed power is minimized by the GMD. Other special cases arising from the GTD family of optimal DFE systems include the VBLAST system

[1]While this papers was in the galley stage, we became aware of [33], which has some overlap with our development.

Manuscript received December 09, 2008; accepted September 08, 2009. First published September 29, 2009; current version published February 10, 2010. This work was supported in part by the Office of Naval Researc (ONR) under Grant N00014-08-1-0709 and in part by the TMS scholarship 94-2-A-018 of the National Science Council of R. O. C., Taiwan. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Anna Scaglione.

The authors are with the California Institute of Technology (Caltech), Pasadena, CA 91125 USA (e-mail: cweng@caltech.edu; cyc@caltech.edu; ppvnath@systems.caltech.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2009.2033311

[30], and a new solution called the Bi-diagonal (BID) transceiver.

This paper is structured as follows. In Section II, we formulate the power minimization problem. In Section III, we show that under optimal bit allocation, optimum linear transceivers achieve the same minimum value for transmitted power as optimum DFE transceivers with bit allocation. Section IV presents a transceiver structure based on generalized triangular decomposition of the channel matrix, and proves that such a system always achieves the minimum power given in Section III. In Section V we report several special cases of the optimal solutions developed in Section IV. Some of these are known structures (SVD, GMD, and VBLAST or QR-based) and some are new (e.g., the bi-diagonal structure). The special case based on a QR receiver has the property that the transmitter only has bit allocation but no precoding matrix. The usefulness of this in limited feedback systems is revisited in Section V-C. Two other optimizations are considered in Section VI: a) minimization of power for fixed set of bit rates and error probabilities (the QoS problem), and b) maximization of bit rate for fixed power and error probabilities. It is shown in both cases that the GTD yields optimal solutions. Sections Sections VII and VIII present simulation results and Section IX concludes the paper.[2] [3]

II. Formulating the Power Minimization Problem

The transceiver we consider is shown in Fig. 1. Here, $\mathbf{H}$ is a $J \times N$ memoryless channel matrix, and the additive Gaussian noise $\mathbf{n}$ is assumed to have zero-mean and covariance matrix $\mathrm{E}\{\mathbf{n}\mathbf{n}^{\dagger}\} = \sigma_n^2\mathbf{I}$. It is assumed throughout that $\mathbf{H}$ is deterministic and known to the transmitter and receiver except in Section VIII. The linear precoder matrix is denoted as $\mathbf{F}$. The vector $\mathbf{s}(n)$ represents the M transmitted symbol streams $s_k(n)$ (with time argument n deleted in all discussions). The received signal is $\mathbf{y} = \mathbf{Hx} + \mathbf{n}$, where $\mathbf{x} = \mathbf{Fs}$. The DFE equalizer consists of the feedforward part $\mathbf{G}$ and feedback part $\mathbf{B}$. Causality of decision feedback is ensured by restricting $\mathbf{B}$ to be strictly upper triangular. With $\tilde{\mathbf{s}}$ denoting the signal vector after the decision device, the input to the decision device has the form $\hat{\mathbf{s}} = \mathbf{GHFs} - \mathbf{B}\tilde{\mathbf{s}} + \mathbf{Gn}$. Under the assumption of correct past decisions, i.e., $\tilde{\mathbf{s}} = \mathbf{s}$ (a good assumption in the high SNR regime), this yields

$$\hat{\mathbf{s}} = (\mathbf{GHF} - \mathbf{B})\mathbf{s} + \mathbf{Gn}. \quad (1)$$

Equation (1) shows that the system described above has an effective transfer matrix $\mathbf{GHF} - \mathbf{B}$ from $\mathbf{s}$ to $\hat{\mathbf{s}}$, and an additive noise term $\mathbf{Gn}$. It, therefore, has the ZF property if

$$\mathbf{GHF} - \mathbf{B} = \mathbf{I}. \quad (2)$$

ZF will be assumed throughout the paper, so that

$$\hat{\mathbf{s}} = \mathbf{s} + \mathbf{Gn}. \quad (3)$$

[2]The following notations are used in the paper. Boldface upper-case letters denote matrices, boldface lower-case letters denote column vectors, and italics denote scalars. Superscript $(\cdot)^{\dagger}$ and $(\cdot)^T$ denote transpose conjugation and transpose respectively. $[\mathbf{A}]_{ij}$ denotes the (i, j)th element of the matrix $\mathbf{A}$. $[\mathbf{A}]_{i\times j}$ denotes the $i \times j$ matrix containing the first i rows and j columns of the matrix $\mathbf{A}$. For vector $\mathbf{x}$, the notation $\mathrm{diag}(\mathbf{x})$ denotes the diagonal matrix with diagonal terms equal to the elements in $\mathbf{x}$. For matrix $\mathbf{X}$, the notation $\mathrm{diag}(\mathbf{X})$ denotes the column vector whose elements are the diagonal terms of $\mathbf{X}$.

[3]Preliminary version of these results were presented in [34] and [35].

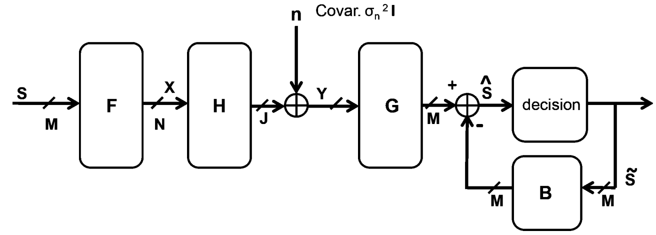


Fig. 1. The MIMO transceiver with linear precoder and DFE.

Without this assumption the problems to be addressed are more difficult, and will be left for the future. Since $\mathbf{GHF} = \mathbf{I} + \mathbf{B}$ is upper triangular with unit diagonal elements, it has rank M. To make the zero forcing assumption possible, $\mathbf{H}$ is assumed to have rank $K \geq M$.

In the following sections, we will first discuss the problem of minimizing the transmitted power subject to a specified total bit rate and a specified error probability in each sub-stream. Assume the components s_k of $\mathbf{s}$ are zero-mean uncorrelated processes representing independent data streams with power P_k so that the input covariance is

$$\mathbf{\Lambda}_s = \mathrm{E}\{\mathbf{s}\mathbf{s}^{\dagger}\} = \mathrm{diag}(P_1, P_2, \ldots, P_M). \quad (4)$$

We assume the $k\mathrm{th}$ data stream is a b_k-bit QAM constellation. From (3), since the error vector at the input of the decision device is $\mathbf{e} = \hat{\mathbf{s}} - \mathbf{s} = \mathbf{Gn}$ where $\mathbf{n}$ is zero-mean Gaussian, the error components e_k are zero-mean Gaussian with variance

$$\sigma_{e_k}^2 = \sigma_n^2[\mathbf{G}\mathbf{G}^{\dagger}]_{kk}. \quad (5)$$

The probability of error for the $k\mathrm{th}$ symbol stream is then [25]

$$P_e(k) \approx 4(1 - 2^{-b_k/2})Q\left(\sqrt{\frac{3P_k}{(2^{b_k} - 1)\sigma_{e_k}^2}}\right) \quad (6)$$

where $Q(\tau) = \int_{\tau}^{\infty} e^{-t^2/2}dt/\sqrt{2\pi}$. Under the high bit rate assumption ($b_k \gg 1$) we have $2^{b_k} - 1 \approx 2^{b_k}$ and $1 - 2^{-b_k/2} \approx 1$. By rearranging (6) we then get

$$\frac{P_k}{\sigma_{e_k}^2} \approx \frac{2^{b_k}}{3}\left(Q^{-1}\left(\frac{P_e(k)}{4}\right)\right)^2 \quad (7)$$

where $Q^{-1}(\cdot)$ denotes the inverse function of $Q(\cdot)$. This is the average power to noise ratio required for the $k\mathrm{th}$ QAM stream to operate at error probability $P_e(k)$ with b_k-bits.

In this section we will regard the error probability $P_e(k)$ as the quality of service (QoS) specification. For the special case of DMT systems one takes all $P_e(k)$ to be equal [17]. The total power transmitted on the channel can be written as

$$P_{\mathrm{trans}} = \mathrm{Tr}(\mathbf{F}\mathbf{\Lambda}_s\mathbf{F}^{\dagger}) = \mathrm{Tr}(\mathbf{F}^{\dagger}\mathbf{\Lambda}_s\mathbf{F}) = \sum_{k=1}^{M} P_k[\mathbf{F}^{\dagger}\mathbf{F}]_{kk}.$$

Substituting from (7) we can rewrite this as

$$P_{\mathrm{trans}} = \sum_{k=1}^{M} d_k 2^{b_k}\sigma_{e_k}^2[\mathbf{F}^{\dagger}\mathbf{F}]_{kk} \quad (8)$$

where

$$d_k = \frac{1}{3}\left(Q^{-1}\left(\frac{P_e(k)}{4}\right)\right)^2 \tag{9}$$

which is determined by the specified probability of error. From (8) and (5) the transmitted power can then be written as

$$P_{\text{trans}} = \sum_{k=1}^{M} c_k 2^{b_k} [\mathbf{F}^\dagger \mathbf{F}]_{kk} [\mathbf{G}\mathbf{G}^\dagger]_{kk} \tag{10}$$

where

$$c_k = \sigma_n^2 d_k = \frac{\sigma_n^2}{3}\left(Q^{-1}\left(\frac{P_e(k)}{4}\right)\right)^2. \tag{11}$$

Therefore, the problem of minimizing the transmitted power subject to the specified BER and total bit rate constraints, and the ZF constraint can be written as follows:

$$\begin{aligned} \min_{\mathbf{F},\mathbf{G},\mathbf{B},\{b_k\}} \quad & P_{\text{trans}} = \sum_{k=1}^{M} c_k 2^{b_k} [\mathbf{F}^\dagger \mathbf{F}]_{kk} [\mathbf{G}\mathbf{G}^\dagger]_{kk} \\ \text{s.t.} \quad & a)\ \frac{1}{M}\sum_{k=1}^{M} b_k = b \\ & b)\ \mathbf{GHF} - \mathbf{B} = \mathbf{I}. \end{aligned} \tag{12}$$

Ideally, we should also impose the constraint that b_k be nonnegative integers. But the problem is not analytically tractable in that case. For the high bit rate case (large b) the optimal bit allocation formula derived next yields positive b_k which can be rounded to integers without severe loss of optimality.

III. Optimal Bit-Loaded DFE Transceivers

Returning to the minimization of (12), we first observe that

$$\begin{aligned} P_{\text{trans}} &= \sum_{k=1}^{M} c_k 2^{b_k} [\mathbf{F}^\dagger \mathbf{F}]_{kk} [\mathbf{G}\mathbf{G}^\dagger]_{kk} \\ &\geq M \prod_{k=1}^{M} \left(c_k 2^{b_k} [\mathbf{F}^\dagger \mathbf{F}]_{kk} [\mathbf{G}\mathbf{G}^\dagger]_{kk}\right)^{1/M} \\ &= c 2^b \left(\prod_{k=1}^{M} [\mathbf{F}^\dagger \mathbf{F}]_{kk}\right)^{1/M} \left(\prod_{k=1}^{M} [\mathbf{G}\mathbf{G}^\dagger]_{kk}\right)^{1/M} \end{aligned}$$

where we have used the AM-GM inequality, and the fact that

$$b = \frac{1}{M}\sum_{k=1}^{M} b_k. \tag{13}$$

Here

$$c = M\left(\prod_{k=1}^{M} c_k\right)^{1/M}. \tag{14}$$

Equality can be achieved in the AM-GM inequality if and only if the terms are identical for all k, that is

$$c_k 2^{b_k} [\mathbf{F}^\dagger \mathbf{F}]_{kk} [\mathbf{G}\mathbf{G}^\dagger]_{kk} = A$$

for some constant A. Taking logarithms on both sides we get

$$b_k = D - \log_2 c_k - \log_2 [\mathbf{F}^\dagger \mathbf{F}]_{kk} - \log_2 [\mathbf{G}\mathbf{G}^\dagger]_{kk} \tag{15}$$

where D is a constant, chosen such that (13) is satisfied. Equation (15) is called the optimum bit loading formula.[4] For any fixed precoder $\mathbf{F}$, receiver $\{\mathbf{G}, \mathbf{B}\}$, and specified probabilities of error $P_e(k)$, the bit allocation that minimizes the transmitted power is given by (15). With this b_k, the quantities P_k are computed from (7) where $\sigma_{e_k}^2$ is as in (5). With P_k so chosen, the specified probabilities of error are met, and the total power P_{trans} is minimized. This minimized power is

$$P_{\text{trans}} = c 2^b \left(\prod_{k=1}^{M} \left[\mathbf{F}^\dagger \mathbf{F}\right]_{kk}\right)^{1/M} \left(\prod_{k=1}^{M} \left[\mathbf{G}\mathbf{G}^\dagger\right]_{kk}\right)^{1/M} \tag{16}$$

and depends only on $\mathbf{F}$ and $\mathbf{G}$, which will be chosen to minimize (16) further. First we derive the optimal $\mathbf{G}$:

Lemma 1: When the precoder $\mathbf{F}$ and the feedback filter $\mathbf{B}$ are given, the optimal feed-forward filter $\mathbf{G}$ for minimizing the transmitted power in (16) subject to the zero forcing condition (2) is

$$\mathbf{G}_{opt} = (\mathbf{I} + \mathbf{B})(\mathbf{HF})^\sharp \tag{17}$$

where $(\mathbf{HF})^\sharp = (\mathbf{F}^\dagger \mathbf{H}^\dagger \mathbf{HF})^{-1}\mathbf{F}^\dagger \mathbf{H}^\dagger$, which is the minimum-norm pseudo inverse of $(\mathbf{HF})$.

Proof: First note that the ZF constraint is satisfied by (17):

$$\mathbf{G}_{opt}\mathbf{HF} - \mathbf{B} = (\mathbf{I} + \mathbf{B})(\mathbf{HF})^\sharp \mathbf{HF} - \mathbf{B} = \mathbf{I}.$$

Suppose there is another $\mathbf{G}'$ satisfying the zero forcing constraint with the given $\mathbf{F}$ and $\mathbf{B}$, i.e., $\mathbf{G}'\mathbf{HF} = \mathbf{I} + \mathbf{B}$. Define $\boldsymbol{\Delta} = \mathbf{G}_{opt} - \mathbf{G}'$. Since both $\mathbf{G}_{opt}$ and $\mathbf{G}'$ satisfy the ZF constraint, it follows that

$$\begin{aligned} \boldsymbol{\Delta}\mathbf{G}_{opt}^\dagger &= \boldsymbol{\Delta}\mathbf{HF}(\mathbf{F}^\dagger \mathbf{H}^\dagger \mathbf{HF})^{-\dagger}(\mathbf{I} + \mathbf{B})^\dagger \\ &= (\mathbf{G}_{opt}\mathbf{HF} - \mathbf{G}'\mathbf{HF})(\mathbf{F}^\dagger \mathbf{H}^\dagger \mathbf{HF})^{-\dagger}(\mathbf{I} + \mathbf{B})^\dagger \\ &= \mathbf{0}. \end{aligned}$$

Therefore

$$\begin{aligned} [\mathbf{G}'\mathbf{G}'^\dagger]_{kk} &= [(\mathbf{G}_{opt} - \boldsymbol{\Delta})(\mathbf{G}_{opt} - \boldsymbol{\Delta})^\dagger]_{kk} \\ &= [(\mathbf{G}_{opt}\mathbf{G}_{opt}^\dagger + \boldsymbol{\Delta}\boldsymbol{\Delta}^\dagger]_{kk} \\ &\geq [\mathbf{G}_{opt}\mathbf{G}_{opt}^\dagger]_{kk} \end{aligned}$$

where we have used $\boldsymbol{\Delta}\mathbf{G}_{opt}^\dagger = \mathbf{0}$ in these inequalities. Therefore we have smaller sub-channel noise variances if we replace $\mathbf{G}'$ with $\mathbf{G}_{opt}$; hence, with given bit rate and probabilities of error, a lower transmitted power can be achieved. ■

The ZF constraint yields the form (17), which can also be found in other references such as [31], where a different problem is solved (mean square error minimized suject to ZF, without bit

[4]In general (15) can yield noninteger or even negative b_k. However in the high-bit-rate case (large b), b_k are large enough to be replaced with integer values without compromising optimality severely. The conclusions derived in the following discussions are valid only under this assumption which has often been made in other papers [16], [17]. Incorporating the positive integer constraint directly into the problem makes it analytically non tractable.

allocation). The main point of the lemma is that the pseudoinverse $(\mathbf{HF})^\sharp$ should be taken to be the *minimum norm* pseudoinverse. This also happens in [31] but the proof techniques for the two problems are different.

Substitute for $\mathbf{G}_{opt}$ from (17) into (16) we get

$$P_{\text{trans}} = c2^b \left(\prod_{k=1}^{M}[\mathbf{F}^\dagger\mathbf{F}]_{kk}\right)^{1/M} \times \left(\prod_{k=1}^{M}[(\mathbf{I}+\mathbf{B})(\mathbf{F}^\dagger\mathbf{H}^\dagger\mathbf{HF})^{-1}(\mathbf{I}+\mathbf{B})^\dagger]_{kk}\right)^{1/M}.$$

Hadamard's inequality for positive definite matrices yields

$$\prod_{k=1}^{M}[\mathbf{F}^\dagger\mathbf{F}]_{kk} \geq \det(\mathbf{F}^\dagger\mathbf{F})$$

and

$$\begin{aligned}\prod_{k=1}^{M}&[(\mathbf{I}+\mathbf{B})(\mathbf{F}^\dagger\mathbf{H}^\dagger\mathbf{HF})^{-1}(\mathbf{I}+\mathbf{B})^\dagger]_{kk}\\ &\geq \det((\mathbf{I}+\mathbf{B})(\mathbf{F}^\dagger\mathbf{H}^\dagger\mathbf{HF})^{-1}(\mathbf{I}+\mathbf{B})^\dagger)\\ &= \det((\mathbf{F}^\dagger\mathbf{H}^\dagger\mathbf{HF})^{-1})\end{aligned}$$

where we use the fact that $\det(\mathbf{I}+\mathbf{B}) = 1$ since $\mathbf{I}+\mathbf{B}$ is upper triangular with diagonal terms all equal to unity. Substituting the above result into the transmitted power, we have

$$P_{\text{trans}} \geq c2^b \left(\frac{\det(\mathbf{F}^\dagger\mathbf{F})}{\det(\mathbf{F}^\dagger\mathbf{H}^\dagger\mathbf{HF})}\right)^{1/M}.$$

In the Appendix we prove that

$$P_{\text{trans}} \geq P_{\min} = c2^b \left(\frac{1}{\prod_{k=1}^{M}\sigma_{h,k}^2}\right)^{1/M} \tag{18}$$

where $\{\sigma_{h,k}\}_{k=1}^{M}$ are the first M dominant channel singular values. Note that $P_{\min}$ in (18) is exactly equal to the form derived for a linear transceiver with optimal bit loading [17]. This means, the extra freedom provided by the decision feedback receiver structure does not reduce the power needed to achieve the specified bit rate and probability of error. So we have proved:

Theorem 1: Linear versus DFE transceiver: Consider the DFE system of Fig. 1 and assume the bit rate b and error probabilities $P_e(k)$ are fixed. Then under optimal bit allocation and ZF, the minimum transmitted power obtained by optimizing $\mathbf{F}, \mathbf{G}$, and $\mathbf{B}$ is given by $P_{\min}$ defined in (18). This same minimum power can also be achieved by a linear transceiver (a transceiver with $\mathbf{B} = \mathbf{0}$) by optimizing $\mathbf{F}, \mathbf{G}$, and the bit allocation under the ZF constraint.

Thus, when bit loading is allowed, *DFE with linear precoding has the same performance as linear transceivers!* However, the DFE system with linear precoding actually provides more choices of possible configurations that achieve the $P_{\min}$ in (18). This interesting observation will be elaborated further in the following sections.

IV. GTD-Based Systems

We now show that the generalized triangular decomposition (GTD), proposed in [12], can be used to construct optimal solutions to the problem (12). First we need a definition:

Definition 1: Multiplicative Majorization [9], [20]: Given two vectors $\mathbf{a} = [a_1, a_2, \ldots, a_n]$ and $\mathbf{b} = [b_1, b_2, \ldots, b_n]$ where a_i and b_i are all positive, we say $\mathbf{a}$ is multiplicatively majorized by $\mathbf{b}$, or $\mathbf{b}$ multiplicatively majorizes $\mathbf{a}$, and we write $\mathbf{a} \prec_\times \mathbf{b}$ if

$$\prod_{i=1}^{k} a_{[i]} \leq \prod_{i=1}^{k} b_{[i]}, \quad \text{for } 1 \leq k \leq n$$

and equality holds when $k = n$. Here, "$[i]$" denotes the component of the vector with i-th largest magnitude.

The following result was proved in [12]:

Lemma 2: The generalized triangular decomposition (GTD): Let $\mathbf{H} \in \mathcal{C}^{m\times n}$ be a rank-K matrix with singular values $\sigma_{h,1}, \sigma_{h,2}, \ldots, \sigma_{h,K}$ in descending order. Let $\mathbf{r} = [r_1, r_2, \ldots, r_K]$ be a given vector which satisfies

$$\mathbf{a} \prec_\times \mathbf{h} \tag{19}$$

where $\mathbf{a} = [|r_1|, |r_2|, \ldots, |r_K|]$ and $\mathbf{h} = [\sigma_{h,1}, \sigma_{h,2}, \ldots, \sigma_{h,K}]$. Then there exist matrices $\mathbf{R}$, $\mathbf{Q}$, and $\mathbf{P}$ such that

$$\mathbf{H} = \mathbf{QRP}^\dagger \tag{20}$$

where $\mathbf{R}$ is a $K \times K$ upper triangular matrix with diagonal terms equal to r_k, and $\mathbf{Q} \in \mathcal{C}^{m\times K}$ and $\mathbf{P} \in \mathcal{C}^{n\times K}$ both have orthonormal columns.

This decomposition is the extended version of the results by Weyl in 1949 [29] and Horn in 1954 [7], which give the complete relationship between the matrix singular values and eigenvalues. Special instances of the GTD include the following.

1) The singular value decomposition (SVD) [5].
2) The Schur decomposition [5].
3) The QR decomposition [5].
4) The complete orthogonal decomposition [5].
5) The geometric mean decomposition (GMD) [10].
6) The bi-diagonal decomposition [5].

In all these cases, the majorization property (19) can be verified to be true. Before diving into any specific realization, we describe in detail the GTD-based method to construct the transceiver matrices $\mathbf{F}, \mathbf{G}, \mathbf{B}$. The following are the steps involved.

1) Given the channel $\mathbf{H}$, we first choose a set of diagonal elements r_k for $\mathbf{R}$ such that (19) holds, and express $\mathbf{H}$ in the GTD form (20), thereby determining a set of matrices $\mathbf{P}, \mathbf{Q}$ and $\mathbf{R}$.
2) We then show how to choose the precoder $\mathbf{F}$, the receiver matrices $\mathbf{G}$ and $\mathbf{B}$, and the bit allocation such that the transmitted power achieves the minimum value $P_{\min}$ given by (18).

The first step offers considerable freedom, since any choice for the diagonal elements $r_k = [\mathbf{R}]_{kk}$ is acceptable as long as (19) holds. We will choose the K elements r_k as follows: a) Choose $r_1, r_2, \ldots r_M$ to be any set of positive numbers multiplicatively majorized by the first M dominant singular values $\sigma_{h,1} \geq \cdots \geq \sigma_{h,M}$ of the channel. b) Choose $r_{M+1}, \ldots r_K$ to

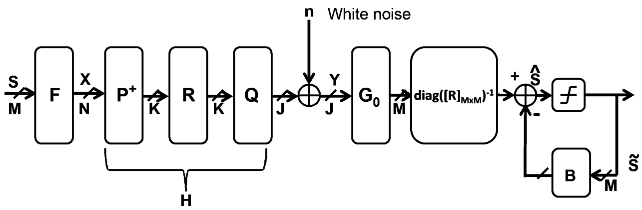


Fig. 2. The proposed form of optimal solution for the DFE transceiver.

be $\sigma_{h,M+1},\ldots,\sigma_{h,K}$ or any permutation thereof. The choice in a) implies in particular that

$$\prod_{k=1}^{M}[\mathbf{R}]_{kk}^2=\prod_{k=1}^{M}\sigma_{h,k}^2. \tag{21}$$

With $[\mathbf{R}]_{kk}$ chosen as above, assume the channel has been expressed as in (20). We are now ready for the second step. We begin by choosing $N\times M$ precoder as

$$\mathbf{F}=[\mathbf{P}]_{N\times M}. \tag{22}$$

Thus the columns of the precoder are the first M columns of $\mathbf{P}$. We then choose the feedforward matrix as

$$\mathbf{G}=\left(\mathrm{diag}([\mathbf{R}]_{M\times M})\right)^{-1}\mathbf{G}_0. \tag{23}$$

where

$$\mathbf{G}_0=[\mathbf{Q}^\dagger]_{M\times J}. \tag{24}$$

Since $\mathbf{P}$ and $\mathbf{Q}$ have orthonormal columns, the columns of $\mathbf{F}$ are orthonormal, and so are the rows of $\mathbf{G}_0$. Finally, the feedback matrix $\mathbf{B}$ is determined by the zero forcing condition $\mathbf{B}=\mathbf{GHF}-\mathbf{I}$. To simplify this, observe first that

$$\begin{aligned}\mathbf{GHF}&=\left(\mathrm{diag}([\mathbf{R}]_{M\times M})\right)^{-1}\mathbf{G}_0\mathbf{QRP}^\dagger\mathbf{F}\\&=\left(\mathrm{diag}([\mathbf{R}]_{M\times M})\right)^{-1}(\mathbf{I_M}\ \mathbf{0})\,\mathbf{R}\begin{pmatrix}\mathbf{I_M}\\ \mathbf{0}\end{pmatrix}\\&=\left(\mathrm{diag}([\mathbf{R}]_{M\times M})\right)^{-1}[\mathbf{R}]_{M\times M}.\end{aligned}$$

Here, we have used the facts that

$$\mathbf{G}_0\mathbf{Q}=(\mathbf{I}_M\ \ \mathbf{0})\quad\text{and}\quad\mathbf{P}^\dagger\mathbf{F}=\begin{pmatrix}\mathbf{I_M}\\ \mathbf{0}\end{pmatrix}$$

which follow from the choices (22) and (24), and the column orthonormality of $\mathbf{P}$ and $\mathbf{Q}$. Thus the expression for the feedback matrix becomes

$$\mathbf{B}=\mathbf{GHF}-\mathbf{I}=\left(\mathrm{diag}([\mathbf{R}]_{M\times M})\right)^{-1}[\mathbf{R}]_{M\times M}-\mathbf{I}. \tag{25}$$

This is strictly upper triangular since $\mathbf{R}$ is upper triangular. Fig. 2 shows the structure of the GTD transceiver just described.

With the above choice of transceiver matrices the error variance (5) in the k-th substream becomes

$$\sigma_{e_k}^2=\frac{\sigma_n^2}{[\mathbf{R}]_{kk}^2}. \tag{26}$$

Substituting into (8) the transmitted power needed to satisfy the specified QoS and bit rate constraints can be expressed as

$$P_{\mathrm{trans}}=\sum_{k=1}^{M}d_k2^{b_k}[\mathbf{F}^\dagger\mathbf{F}]_{kk}\sigma_{e_k}^2=\sum_{k=1}^{M}\frac{d_k2^{b_k}}{[\mathbf{R}]_{kk}^2}\sigma_n^2.$$

Since $\sigma_n^2d_k=c_k$ (from (11)), this simplifies to

$$P_{\mathrm{trans}}=\sum_{k=1}^{M}\frac{c_k2^{b_k}}{[\mathbf{R}]_{kk}^2}. \tag{27}$$

We now show that the system in Fig. 2 with $\mathbf{F}$, $\mathbf{G}$, and $\mathbf{B}$ chosen as described achieves optimality for problem (12), provided the bit allocation is chosen appropriately:

Theorem 2: With the bit allocation chosen as

$$b_k=\log_2\left(\frac{c}{M}2^b\left(\frac{1}{\prod_{k=1}^{M}\sigma_{h,k}^2}\right)^{1/M}\right)-\log_2(c_k)+\log_2([\mathbf{R}]_{kk}^2) \tag{28}$$

for $1\le k\le M$, the system in Fig. 2 with $\mathbf{F}$ as in (22), $\mathbf{G}$ as in (23), and $\mathbf{B}$ as in (25), achieves the minimized power for the specified $\{P_e(k)\}$ and bit rate constraint.

Proof: Observe first that (28) satisfies the total bit constraint because

$$\begin{aligned}\sum_{k=1}^{M}b_k&=\log_2\left(\left(\frac{c}{M}\right)^M\frac{2^{Mb}}{\prod_{k=1}^{M}\sigma_{h,k}^2}\right)\\&\quad-\log_2\prod_{k=1}^{M}c_k+\log_2\left(\prod_{k=1}^{M}[\mathbf{R}]_{kk}^2\right)\\&=Mb-\log_2\left(\frac{1}{\prod_{k=1}^{M}\sigma_{h,k}^2}\right)+\log_2\left(\prod_{k=1}^{M}[\mathbf{R}]_{kk}^2\right)\\&=Mb\end{aligned}$$

using (21) and $c=M(\prod_{k=1}^{M}c_k)^{1/M}$. Next, (28) implies

$$\frac{c_k2^{b_k}}{[\mathbf{R}]_{kk}^2}=\frac{c}{M}2^b\left(\frac{1}{\prod_{k=1}^{M}\sigma_{h,k}^2}\right)^{1/M}. \tag{29}$$

Substituting into (27) we get

$$P_{\rm trans} = \sum_{k=1}^{M} \frac{c_k 2^{b_k}}{[\mathbf{R}]_{kk}^2} = M \times \frac{c}{M} 2^b \left(\frac{1}{\prod_{k=1}^{M} \sigma_{h,k}^2} \right)^{1/M}. \quad (30)$$

Since this is the minimum achievable power $P_{\min}$ (see discussion leading to (18)), the proof is complete. ■

The extra flexibility in designing the transceivers, offered by this GTD-based DFE system, must be carefully understood. Recall that the bit loading formula for the linear transceiver to achieve the minimum transmitted power is [17]

$$b_k = D - \log_2 c_k + \log_2(\sigma_{h,k}^2) \quad (31)$$

where $\sigma_{h,k}$ are fixed numbers given to us by the channel. The values computed from (31) are not guaranteed to be integers, or even nonnegative. For the GTD-based DFE system, the bit loading scheme (28) can be written as

$$b_k = D - \log_2 c_k + \log_2([\mathbf{R}]_{kk}^2). \quad (32)$$

The freedom of the GTD-based system allows us to reshape the value of $[\mathbf{R}]_{kk}$ as long as the multiplicative majorization property (19) is satisfied. This flexibility may be used, for example, to ensure that the bit loading scheme in (32) is realizable. So, even though the linear transceiver with bit allocation (31) can achieve the same minimum power (30) as any optimal DFE-transceiver, the bit allocation formula in the GTD-based DFE opens up more freedom.

We now make an interesting observation about the powers P_k in the optimal system. Substituting (26) into (7) and using the definition of c_k in (11) we find

$$P_k = \frac{2^{b_k} c_k}{[\mathbf{R}]_{kk}^2}. \quad (33)$$

Substituting from (32) it then follows that

$$P_k = 2^D \quad (34)$$

for all k. Thus in the optimal system which has orthonormal columns for the precoder $\mathbf{F}$, the powers P_k are identical for all k. Since $P_{\rm trans} = \sum_k P_k$ from (33) and (27), we have $P_k = P_{\rm trans}/M$ for all k.

V. Examples of GTD-Based Configurations

In Section IV we mentioned many examples of the GTD, such as SVD, Schur decomposition, GMD, and so on. Some of these have already appeared in the literature in different contexts. Each of these serves as a specific realization of the optimal DFE transceiver acheiving minimum transmitted power, provided the bits are allocated as in (28). Each realization has a different choice of $r_k(=[\mathbf{R}]_{kk})$ satisfying the majorization condition (19), and in all cases, we restrict the precoder $\mathbf{F}$ to be the orthonormal choice (22). $\mathbf{G}$ is chosen as in (23), and $\mathbf{B}$ as in (25). We now elaborate on these different realizations arising from different GTD forms $\mathbf{H} = \mathbf{QRP}^\dagger$.

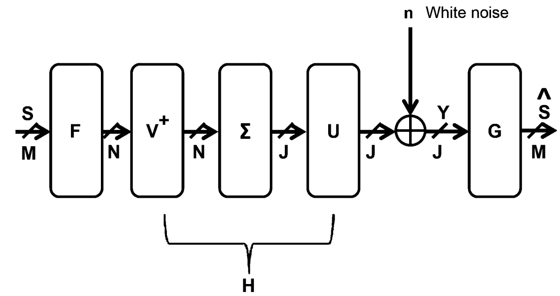


Fig. 3. The linear transceiver with channel represented using SVD.

A. SVD Transceiver—the Linear Transceiver

The singular value decomposition (SVD) of the channel matrix can be written as $\mathbf{H} = \mathbf{U\Sigma V}^\dagger$, where $\mathbf{U}$ and $\mathbf{V}$ are unitary and Σ is a diagonal matrix. Since $\mathbf{R} = \mathbf{\Sigma}$ is diagonal, the feedback matrix $\mathbf{B} = \mathbf{0}$ from (25), and the system reduces to a linear transceiver as in Fig. 3. This optimal solution for linear transceivers was proposed in [17].

B. GMD Transceiver

The geometric mean decomposition (GMD) was introduced in [10]. The GMD of the channel $\mathbf{H}$ has the form $\mathbf{H} = \mathbf{QRP}^\dagger$, where $\mathbf{Q}$ and $\mathbf{P}$ have orthonormal columns, and $\mathbf{R}$ is an upper triangular matrix. Furthermore *the first M diagonal elements of $\mathbf{R}$ are identical*, and equal to the geometric mean of the M dominant channel singular values. For the case where the specified error probabilities $P_e(k)$ (hence, c_k) are identical for all k, it follows from (28) that there is no need for bit allocation, that is, $b_k = b$ for all k. Unlike other special cases of the GTD such as the SVD, the question of b_k becoming unrealizable (i.e., taking noninteger or negative values), therefore, does not arise.

C. QR Transceiver—ZF-VBLAST System

The QR decomposition of the channel matrix can be written as $\mathbf{H} = \mathbf{QR}$, where $\mathbf{Q}$ has orthonormal columns, and $\mathbf{R}$ is upper triangular. This yields a special case of the GTD transceiver, where the precoder is $\mathbf{F} = \begin{pmatrix} \mathbf{I}_M \\ \mathbf{0} \end{pmatrix}$, and can be implemented at no cost. See Fig. 4. This system leads to the ZF-VBLAST system, widely used in MIMO wireless communication [30].

The optimal transceiver design usually assumes that $\mathbf{H}$ is known at the transmitter side. This assumption is not generally true. The more practical scheme would be the so called *limited feedback scheme*, in which the receiver uses a low rate feedback to tell the transmitter to use one of the precoders in a pre-determined codebook of precoders [18].

The QR based transceiver with bit loading is very suitable in limited feedback systems because the precoder matrix is identity, and only the bit loading vector $[b_1 \dots b_M]$ needs to be known.[5] The receiver can compute $\{b_k\}$ from (15), quantize it to the bit loading vector nearest to the vectors in a pre-determined codebook, and feed back the index of that vector to the transmitter. The design of this codebook is an interesting

[5]In the scheme described in [4], the power allocation P_k also should be fedback, but in the GTD based optimal system $P_k = P_{\rm trans}/M$ for all k as shown at the end of Section IV.

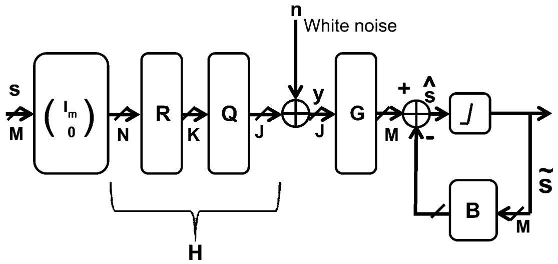


Fig. 4. The QR transceiver, which has the lazy precoder. This is identical to the ZF-VBLAST system.

problem, but is beyond the scope of this paper. Intuitively, this scheme will perform better than limited feedback schemes using Grassmann codebook [18], [27], since the Grassmann codebook aims to cover the Grassmann manifold [28] while the bit loading codebook only tries to cover M-vectors with integer valued entries. This intuition is supported by Monte Carlo simulations in Section VIII. It is assuring to know that since all GTD-based systems are optimal when the bit loading formula is realizable, this QR based special case has no loss of optimality even though it offers a simple precoder and a simple way to perform limited feedback.

D. BI-Diagonal Transceiver

It is well-known [5] that any $J \times P$ matrix $\mathbf{H}$ can be factored as $\mathbf{H} = \mathbf{QRP}^{\dagger}$, where $\mathbf{Q}$ and $\mathbf{P}$ have orthonormal columns, and $\mathbf{R}$ has the bi-diagonal form

$$\mathbf{R} = \begin{pmatrix} d_1 & f_1 & 0 & \cdots & 0 \\ 0 & d_2 & f_2 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & d_{P-1} & f_{P-1} \\ 0 & \cdots & \cdots & 0 & d_P \\ & & \mathbf{0} & & \end{pmatrix}.$$

With the channel represented in this bi-diagonal form, the feedback matrix given in (25) becomes

$$\mathbf{B} = \begin{pmatrix} 0 & f_1 & 0 & \cdots & 0 \\ 0 & 0 & f_2 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & f_{M-1} \\ 0 & \cdots & \cdots & 0 & 0 \end{pmatrix}.$$

Therefore the implementation of the DFE will be very simple since we need only to feedback *one* previous decision for detecting the current symbol. Also, the computation of the bidiagonal decomposition is inexpensive [5]. To the authors' knowledge, this kind of system has not previously been reported in transceiver literature.

Summarizing, any of the above four GTD-based systems achieves optimality. However, each one of them has some special features, which might be useful in different situations. Also, it is possible that other GTD-based systems exist with potential benefits in specific situations.

VI. Other Transceiver Problems Solved by GTD

We now consider two variations of the transceiver optimization problem, both of which have solutions based on the GTD.

A. Quality of Service (QoS) Problem

The quality of service problem in MIMO communication has been considered by a number of authors [6], [13], [23]. In these papers the QoS is defined in the output SINR sense, and furthermore there is no bit allocation. In fact, [6] addresses a special case of the problem discussed in [13], namely the case where the channel $\mathbf{H} = \mathbf{I}$. Here, we consider a different situation where the error probability $P_e(k)$ [equivalently the constants c_k in (11)] and bit rate b_k of each sub-stream are specified to be the QoS parameters. We will show that under some multiplicative majorization condition, we can customize the GTD-based transceiver to obtain an optimal solution which minimizes power subject to the QoS specifications $\{c_k, b_k\}$. More precisely, the problem considered here is

$$\begin{aligned} \min_{\mathbf{F},\mathbf{G},\mathbf{B}} \quad & P_{\text{trans}} \\ \text{s.t.} \quad & \text{(a) } \mathbf{GHF} = \mathbf{I} + \mathbf{B} \\ & \text{(b) } \{\mathrm{c_k}, \mathrm{b_k}\} \text{fixed (QoS for the kth data stream).} \end{aligned} \tag{35}$$

The solution is based on the following result:

Theorem 3: For the QoS problem (35), the following are true.

a) The minimum required power to achieve the specification will be at least as large as

$$P_{\min} = c2^b \left(\frac{1}{\prod_{k=1}^{M} \sigma_{h,k}^2} \right)^{1/M}$$

where $c = M(\prod_{k=1}^{M} c_k)^{1/M}$ and $b = \sum_{k=1}^{M} b_k/M$.

b) This $P_{\min}$ is achievable if

$$\frac{\{c_1 2^{b_1}, \ldots, c_M 2^{b_M}\}}{\frac{c2^b}{M}} \prec_{\times} \frac{\{\sigma_{h,1}^2, \ldots, \sigma_{h,M}^2\}}{\left(\prod_{k=1}^{M} \sigma_{h,k}^2\right)^{1/M}} \tag{36}$$

that is, if the vector on the left which is determined by the QoS constraints, is multiplicatively majorized by the vector on the right which is determined by the channel.

Proof: Part a) is true because the problem (12) discussed in previous sections is a relaxed version of the current problem (35). We prove part b) by constructing a system that achieves $P_{\min}$ when (36) holds. If (36) holds then the majorization condition (19) required in Lemma 2 can be satisfied by choosing $[\mathbf{R}]_{kk}$ to be positive square roots of

$$[\mathbf{R}]_{kk}^2 = \begin{cases} \dfrac{Mc_k 2^{b_k} \left(\prod_{k=1}^{M} \sigma_{h,k}^2 \right)^{1/M}}{c2^b}, & \text{for } k = 1, 2, \ldots, M. \\ \sigma_{h,k}^2, & \text{for } k = M+1, \ldots K \end{cases} \tag{37}$$

where K is the rank of $\mathbf{H}$. Then by Lemma 2 there exists a $K \times K$ upper triangular matrix $\mathbf{R}$, such that the decomposition $\mathbf{H} = \mathbf{QRP}^\dagger$ is true, where $\mathbf{Q}$ and $\mathbf{P}$ have orthonormal columns. Now choose the transceiver matrices $\mathbf{F}, \mathbf{G}$, and $\mathbf{B}$ as in (22), (23), and (25). Then P_{trans} is as in (27). Substituting from (37) we get $P_{\text{trans}} = P_{\min}$ indeed. ∎

This system, which achieves $P_{\text{trans}} = P_{\min}$, will be referred to as the custom GTD-based system, since the value of the precoder and equalizer are not computed solely depending on $\mathbf{H}$, but also depending on the given QoS $\{c_k, b_k\}$. This example shows that the GTD-based system has much more flexibility than the linear transceiver system.

It should be pointed out here that when the QoS specification $\{c_k, b_k\}$ is identical for all k, the custom GTD reduces to the GMD. This is because the multiplicative majorization relation (36) always holds in this case.

B. *Maximizing the Bit Rate for Fixed Power and* $\{P_e(k)\}$

The bit rate maximization problem subject to transmitted power constraint is the dual of the problem described in (12). It will be shown that the GTD transceiver gives the optimal solution. For the special case of linear transceivers this problem was considered in [15]. Consider again the system shown in Fig. 1 with ZF. Under the high bit rate assumption, (7) holds, and can be rearranged as

$$b_k \approx \log_2\left(\frac{P_k}{\sigma_{ek}^2 d_k}\right) \tag{38}$$

where d_k represents the BER via (9). Therefore, the problem of maximizing the average bit rate for fixed set of BERs $\{d_k\}$ and total power $\leq P_{\text{total}}$ can be written as

$$\begin{aligned} \max_{\mathbf{F},\mathbf{G},\mathbf{B},\{P_k\}} \quad & b = \frac{1}{M}\sum_{k=1}^{M}\log_2\left(\frac{P_k}{\sigma_{ek}^2 d_k}\right) \\ \text{s.t.} \quad & a)\ \mathrm{Tr}(\mathbf{F}\boldsymbol{\Lambda}_s\mathbf{F}^\dagger) \leq P_{\text{total}} \\ & b)\ \mathbf{GHF} = \mathbf{I} + \mathbf{B}(\text{zero-forcing}) \end{aligned} \tag{39}$$

where $\boldsymbol{\Lambda}_s = \mathrm{diag}(P_1, P_2, \ldots, P_M)$. The power constraint can be rewritten as

$$\sum_{k=1}^{M} P_k[\mathbf{F}^\dagger\mathbf{F}]_{kk} \leq P_{\text{total}}.$$

We solve the above optimization problem in two stages. First we find the optimal power P_k for given $\mathbf{F}$, $\mathbf{G}$, and $\mathbf{B}$, under the power constraint. We then derive the optimal transceiver matrices. Suppose $\{P_k^*\}$ are optimal for problem (39), then the KKT condition [2] states that there exists α such that (40) and (41), shown at the bottom of the page hold, and

$$\alpha\left(\sum_{k=1}^{M} P_k[\mathbf{F}^\dagger\mathbf{F}]_{kk} - P_{\text{total}}\right)\Big|_{P_k=P_k^*} = 0. \tag{42}$$

By solving these equations, we get

$$\alpha = \frac{-M}{P_{\text{total}}\log_e 2}$$

and the optimal power allocation

$$P_k^* = \frac{P_{\text{total}}}{M[\mathbf{F}^\dagger\mathbf{F}]_{kk}}. \tag{43}$$

Observe that when the triplet $\{\mathbf{F}, \mathbf{G}, \mathbf{B}\}$ is fixed, (39) is concave in the vector $[P_1 \ldots P_M]$. So the preceding solution represents a maximum (rather than minimum) of (39). The derivation of (43) is similar to the one in [15] for linear transceivers. Using (43) in (39) and simplifying, we have

$$b = \log_2\left(\prod_{k=1}^{M}\frac{P_{\text{total}}}{Mc_k[\mathbf{F}^\dagger\mathbf{F}]_{kk}[\mathbf{GG}^\dagger]_{kk}}\right)^{1/M}. \tag{44}$$

Thus, the problem of maximizing the bit rate is reduced to maximizing (44) subject to zero forcing. But maximizing (44) is equivalent to minimizing (16). The latter minimization can be achieved with the GTD and results in $P_{\text{trans}} = P_{\min}$ given by (18). So it follows that the optimal solution is such that

$$\prod_{k=1}^{M}[\mathbf{F}^\dagger\mathbf{F}]_{kk}[\mathbf{GG}^\dagger]_{kk} = \frac{1}{\prod_{k=1}^{M}\sigma_{h,k}^2}. \tag{45}$$

Substituting this into (44) the maximized bit rate becomes:

$$b_{\max} = \log_2\left(\frac{P_{\text{total}}}{c}\left(\prod_{k=1}^{M}\sigma_{h,k}^2\right)^{1/M}\right). \tag{46}$$

This is exactly the maximum bit rate that has been achieved with linear transceivers, as shown in [15]. This is not surprising because the GTD yields an entire family of solutions, of which the linear SVD transceiver is a special case (Section V). Thus, whenever bit allocation is permitted, the DFE transceiver offers no advantage over the linear transceiver, as far as maximizing the bit rate is concerned.

For completeness recall that the GTD based optimal solution has matrices $\mathbf{F} = [\mathbf{P}]_{P\times M}$, $\mathbf{G}_0 = [\mathbf{Q}^\dagger]_{M\times J}$, $\mathbf{G}$ as in (23), and $\mathbf{B}$ as in (25). Since this achieves the maximum bit rate, all the special cases discussed in Section V maximize bit rate. Jiang et.

$$\alpha \leq 0 \tag{40}$$

$$\frac{\partial}{\partial P_k}\left(\frac{1}{M}\sum_{k=1}^{M}\log_2\left(\frac{P_k}{\sigma_{ek}^2 d_k}\right) + \alpha\left(\sum_{k=1}^{M}P_k[\mathbf{F}^\dagger\mathbf{F}]_{kk} - P_{total}\right)\right)\Big|_{P_k=P_k^*} = 0 \tag{41}$$

al. considered a different problem in [22] where they showed that the GMD system achieves maximum channel throughput (defined in terms of mutual information) with uniform bit allocaiton, for the case of big SNR. This result is consistent with our result in this section for the actual bit rate, which holds for any GTD. The mathematics used in this section is similar to that in [22]. It is also shown in [22] that when b_k are constrained to be equal for all k, the GMD system minimizes average BER .

VII. Simulation Results With Perfect CSI

In this section we present simulations for the case where the channel is known to the transmitter and the receiver. We consider a number of methods in the comparison. These include the linear transceiver based on SVD (Section V-A), and DFE-based transceivers based on GMD (Section V-B), QR decomposition (Section V-C), and bidiagonal decomposition BID (Section V-D). For these methods, whenever the bit loading formula (15) is not realizable due to finite constellation granularity, we replace it with the optimal bit loading algorithm in [3], and take the precoder and equalizer matrices to be the optimum ones determined by the theory.

In addition, we introduce a new procedure which allows us to achieve the minimum power $P_{\min}$ of (18) with integer bit allocation for the special case of equal c_k (equal error probabilities for all k). This procedure will be denoted as the GB method (generalized bit allocation method) in all simulations. It exploits the freedom offered by the GTD in choosing the diagonal elements $\mathbf{R}_{kk}$ of the lower triangular matrix $\mathbf{R}$. Since the method is somewhat involved we first describe it briefly before proceeding with the simulation examples.

A. Achieving Integer Bits and Minimum Power Together

Assume c_k is identical for all k. The method which we refer to as generalized bit allocation (GB) proceeds as follows. First we compute b_k using (31), and truncate it to the nearest even integer to get a square QAM constellation (replacing b_k with zero if it turns out to be negative). We then check if the bit rate constraint $\sum_k b_k = Mb$ is satisfied with equality. If this is not the case, then we adjust b_k in one of two possible ways depending on the situation. For convenience assume b_k is renumbered such that $b_k \geq b_{k+1}$.

1) If $\sum_k b_k < Mb$ we replace b_M with $b_M + 2$ until either $\sum_k b_k = Mb$ or $b_{M-1} = b_M$. In the former event we stop. If the latter is true but $\sum_k b_k < Mb$ still prevails, we replace b_{M-1} with $b_{M-1} + 2$, and continue the process. If we reach a point where

$$b_{M-n-1} > b_{M-n} = \cdots = b_{M-1} = b_M$$

for some $n > 1$, with $\sum_k b_k < Mb$ still prevailing, then we replace b_{M-n} with $b_{M-n} + 2$. Repeated application of this procedure leads to a bit allocation that satisfies $\sum_k b_k = Mb$.

2) If $\sum_k b_k > Mb$ we modify the preceding in the obvious way: we replace b_1 with $b_1 - 2$ until either $\sum_k b_k = Mb$ or $b_1 = b_2$. In the former event we stop. If the latter event is true but $\sum_k b_k > Mb$ still prevails, we replace b_2 with $b_2 - 2$, and continue the process. If we reach a point where

$$b_1 = b_2 = \cdots = b_n > b_{n+1}$$

with $\sum_k b_k > Mb$ still prevailing, then we replace b_n with $b_n - 2$. Repeated application of this procedure leads to a bit allocation that satisfies $\sum_k b_k = Mb$.

Let $\{b_1^i, b_2^i \dots b_M^i\}$ denote the final bit allocation resulting from this algorithm (superscript i is for "integer") and let $\{b_1, b_2 \dots b_M\}$ denote the initial allocation from (31). We have $\sum_k b_k = \sum_k b_k^i$ by construction. Furthermore, if $\{\sigma_{h,k}\}$ has a wide distribution, then the final bit allocation satisfies

$$[\, b_1^i \quad b_2^i \quad \dots \quad b_M^i \,] \prec_+ [\, b_1 \quad b_2 \quad \dots \quad b_M \,]. \tag{47}$$

The notation $\prec_+$ means that the vector on the left is additively majorized by that on the right [9], [20]. The next step depends upon whether this happens or not. Suppose (47) indeed holds (which is often the case as seen through simulations). If $c_i = c/M$ for all i then by using (31) we verify that this is equivalent to the multiplicative majorization condition (36). Now, with $[\mathbf{R}]_{kk}$ defined as in (32) or more precisely

$$b_k^i = D - \log_2 c_k - \log_2([\mathbf{R}]_{kk}^2). \tag{48}$$

Equation (36) (hence, (47)) is equivalent to the condition (19) demanded by Theorem 1. This means that there exists a GTD for the channel $\mathbf{H}$ such that both (36) and the integer bit allocation (48) hold simultaneously.

According to Theorem 2, this design, therefore, achieves minimum power while at the same time satisfies the integer bit rate constraint for the case where $c_i = c/M$ for all i. This is precisely the beauty of the GTD. We have successfully exploited the flexibility in bit allocation offered by the freedom to choose the diagonal elements $[\mathbf{R}]_{kk}$ in the GTD.

There remains one more case to be considered, namely the situation where the majorization relation (47) does not hold. In this case we have observed that the SVD transceiver (linear transceiver) with integer bit allocation (48) typically yields a smaller BER than all the other GTD methods. So we simply use the SVD system whenever the second situation prevails.

B. Simulations Examples

Throughout this section we assume $M = N = 4$ and $J = 5$ (in the notation of Fig. 1). So the channel matrix $\mathbf{H}$ is of size 5×4; each of its entries $[\mathbf{H}]_{km}$ is drawn from a iid Gaussian distribution with zero mean and unit variance. For each realization of this random $\mathbf{H}$ we compute the BER, and average it over 1000 such realizations. The additive noise is complex circular Gaussian with average power normalized to 0 dB. Gray encoded bits are adopted. The results are given in terms of BER versus transmitted power. Here, we compare the uncoded BER. Since in all our designs the MSE matrix is diagonal, this makes the overall systems act like a set of parallel AWGN channels. Channel coding may be further added to provide coding gain independent of the transceiver designs discussed in the paper. Decision feedback is operative in all the systems being compared, except in the special case of the SVD system.

1) Example 1: High Bit Rate Case: In this example we consider GTD transceivers with bit allocation approximating (28). We assume $c_k = c/M$ (identical error probabilities $P_e(k)$) for all k. The GTD system minimizes the required power to the value $P_{\min}$ given in (18). Fig. 5 shows the simulated BER plots

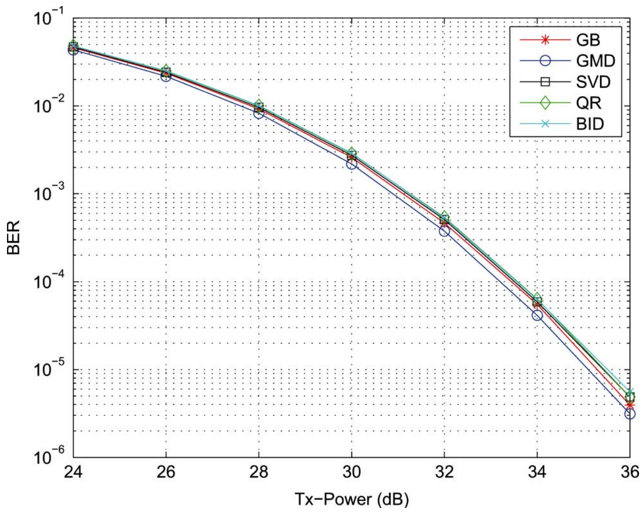


Fig. 5. Example 1: BER versus Tx-Power for $\sum_k b_k = 32$.

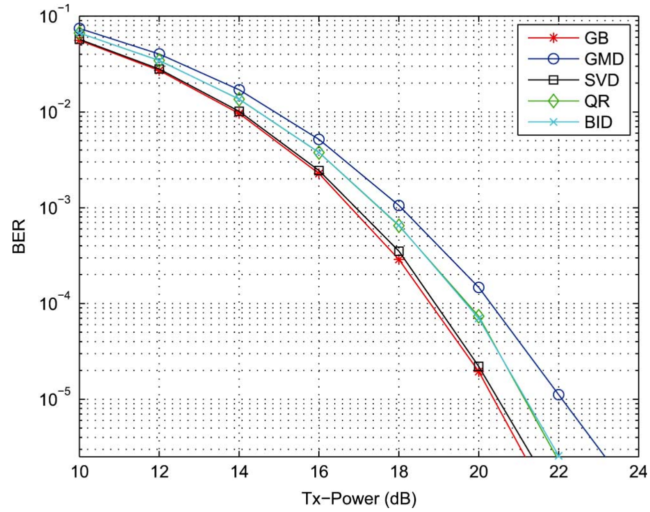


Fig. 6. Example 2. BER versus Tx-Power for $\sum_k b_k = 14$.

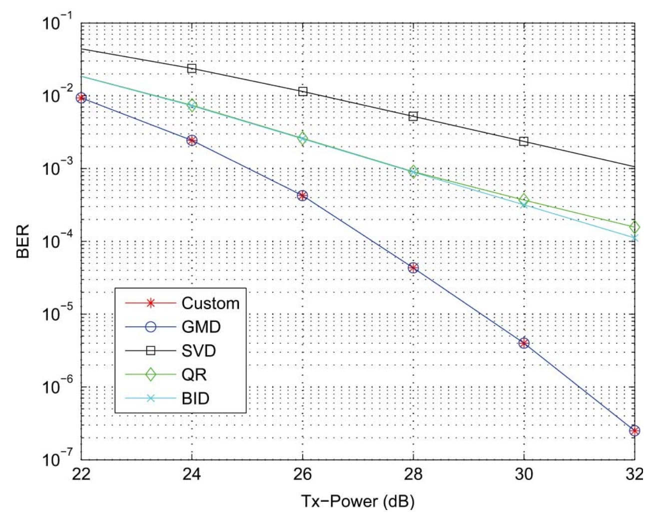


Fig. 7. Example 3. BER versus Tx-Power when $b_k = 6$ for all k.

for the case where $\sum_k b_k = 32$, that is, there are 40 bits to be allocated into the four signal sub-streams. It can be observed that all systems perform about the same. This is consistent with Theorems 1 and 2 under the assumption of high bit rate. Notice in particular that the SVD system without DFE is almost as good as the systems with DFE. For the GB method, integer bit allocation is handled as described in (Section VII-A). For all other methods, whenever the bit loading formula (15) is not realizable due to finite constellation granularity, we replace it with the optimal bit loading algorithm in [3], and take the precoder and equalizer matrices to be the optimum ones determined by the theory. Forcing b_k to be integers usually results in $P_e(k)$ being only approximately equal; the plots are based on BER values averaged over all k. As explained at the end of Section IV, the powers P_k are identical for all k.

2) Example 2: Low Bit Rate Case: For the methods compared above, the theory in this paper predicts identical performance under the "high bit rate" assumption. This was essentially confirmed in the preceding example. In the present example we will see that the performances are quite different from each other in the low bit rate case. This example is similar to Example 1 with the difference that $\sum_k b_k = 14$. Fig. 6 shows the BER plots. In this case, oftentimes the SVD will drop the sub-streams for which the corresponding singular values are too small (by not allocating any bits for them). However, the GMD system will never drop any sub-stream; instead it will force each of the sub-streams to have equal error variance and allocate about the same number of bits. If a sub-stream is very bad (noisy), this strategy will seriously degrade the performance. But the SVD system simply drops the bad sub-channels, therefore, retaining good performance. Note that this behavior of GMD is due to the ZF constraint enforced throughout the paper. For the MMSE receiver without ZF constraint, this effect may disappear. For the "GB" method (Section VII-A) we drop the bad sub-streams as in SVD. This is why both the "GB" and the SVD systems outperform other methods when there are some very bad sub-channels. Note that this effect is not so noticeable in the high bit rate case. Also the "GB" method does not have the non-integer bit allocation problem which all other methods suffer from (unless (47) fails in which case we replace it with SVD as explained at the end of Section VII-A). This is why our GB method performs the best among all the systems.

3) Example 3. Fixed, Identical Constellations: In this example we fix $b_k = 6$ bits for each k (64-QAM streams), and all c_k (i.e., error probabilities $P_e(k)$) are identical. The term "custom" stands for the custom-GTD system with $[\mathbf{R}]_{kk}$ obtained from (32). In this example since $P_e(k)$ and b_k are identical for all k, the custom GTD system reduces to the "GMD" system, which is known to be optimal in terms of BER [10]. Fig. 7 shows the performances of various GTD systems. Clearly GMD and custom GTD outperform other GTDs.

4) Example 4. Fixed, Nonidentical Constellations: This is similar to Ex. 3 with the difference that the fixed constellations have non identical bits: $[8,\ 8,\ 6,\ 6]$ (i.e., 256-QAM, 256-QAM, 64-QAM, and 64-QAM). Fig. 8 shows the BER plots. Again, "custom" denotes the custom-GTD system with $[\mathbf{R}]_{kk}$ obtained from (32), and so it has minimum power for fixed BER. It can be observed from the plots that the custom GTD significantly outperforms all other methods including the GMD. This clearly demonstrates the advantage offered by the flexibility of

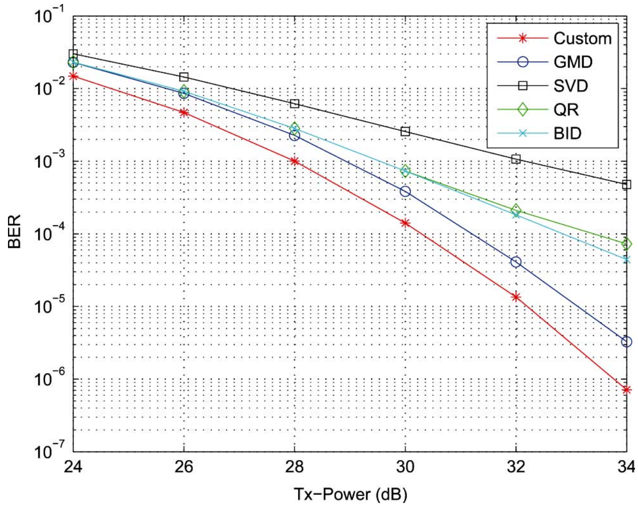


Fig. 8. Example 4. BER versus Tx-Power when bit vector is fixed as $[8,\ 8,\ 6,\ 6]$.

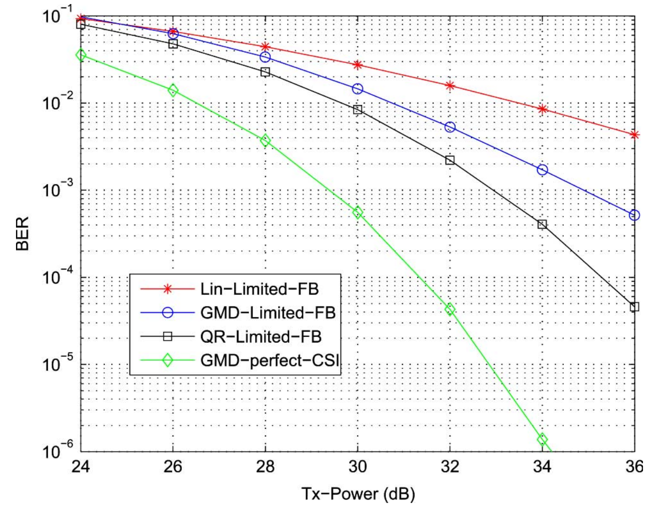


Fig. 9. BER versus Tx-Power with limited feedback (8 feedback bits per block, and 32 bits transmitted per block).

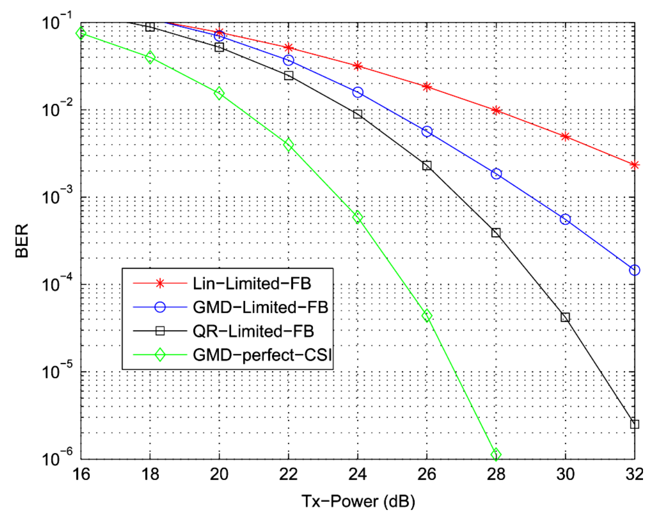


Fig. 10. BER versus Tx-Power with limited feedback (8 feedbck bits per block, and 24 bits transmitted per block).

the GTD. However, among the other four methods, there is no theory as to which one performs better.

VIII. Simulation Results With Limited Feedback

We now consider the limited feedback scheme. As in earlier sections, ZF is assumed, and c_k (equivalently error probabilities $P_e(k)$) are identical for all k. We assume $M = 4$, and $N = J = 5$ so that the 5×4 orthonormal precoders in the Grassman codebook published in [19], [28] can be used. It is assumed that feedback from the receiver to the transmitter is error free. As in the previous section, each of the channel entries $[\mathbf{H}]_{km}$ is drawn from an iid Gaussian distribution with zero mean and unit variance. For each realization of this random $\mathbf{H}$ we compute the BER, and average it over 1000 such realizations. The schemes considered are as follows:

1) The scheme proposed in [18] based on the so-called projection 2-norm criterion. This is a linear transceiver with an orthonormal precoder, with no bit allocation or power allocation. It uses a 8-bit Grassmann codebook [19], [28] to represent a set of 256 precoder matrices. The receiver feeds back the 8 bits to the transmitter to tell which precoder ought to be used. This system is referred to as "Lin-limited-FB".
2) The minimun-BER DFE design proposed in [27] which uses a 8-bit Grassmann codebook in conjunction with GMD. This is referred to as "GMD-limited-FB". It has $b_k = b$ for all k.
3) The QR based DFE design (Section V-C). The precoder is identity, and the receiver feeds back only bit loading information $\{b_k\}$ as described in Section V-C. This will be referred to as "QR-limited-FB".
4) We also show the BER plots for the optimal DFE system based on GMD with perfect CSI at the transmitter, described in Section V-B. This ideal system has the smallest BER, which is shown for reference. This system is referred to as "GMD-perfect-CSI".

Like the first method, the last three methods also have identical powers P_k for all k, but for a different reason as described at the end of Section IV. We present BER plots for two cases: the case where $\sum_k b_k = 32$ (Fig. 9) and where $\sum_k b_k = 24$ (Fig. 10). From the plots we see that the proposed "QR-limited-FB" scheme performs significantly better than the state-of-the-art limited feedback schemes [18], [27], and comes close to the optimal "GMD-perfect-CSI" scheme. Note that the Grassmann codebook aims to cover the Grassmann manifold of orthonormal precoder matrices [1], [28] while the bit loading codebook in the "QR-limited-FB" scheme only has to cover integer valued vectors $[\,b_1 \quad b_2 \quad \ldots \quad b_M\,]$.

We now discuss some details about the "QR-limited-FB" scheme. As described in Section V-C, the codebook here is a set of integer vectors which specifies to the transmitter what b_k are. After the receiver calculates b_k from (15), it quantizes the vector $[\,b_1 \quad b_2 \quad \ldots \quad b_M\,]$ to the nearest vector in the codebook. In the simulation we also restrict the codebook to have vectors with each b_k no less than 4 for Fig. 9 (and 2 for Fig. 10). Also, we use square QAM, so the possible number of bits in each substream will be even. The size of the codebook is, therefore, $(11 \times 10 \times 9)/(3 \times 2) = 165$. This requires less than 8 bits of feedback from receiver to transmitter. Even with such limited feed back, the proposed "QR-limited-FB" scheme performs very well indeed.

IX. Concluding Remarks

We have presented a method for the joint optimization of the matrices $\{\mathbf{F}, \mathbf{G}, \mathbf{B}\}$ and the bits $\{b_k\}$ in a transceiver with DFE. It was formally shown that when the bit allocation, precoder, and equalizer are jointly optimized, linear transceivers and transceivers with DFE have identical performance in the sense that the transmitted power is identical for a given bit rate and error probability. We also proved that any GTD-based system achieves the optimal performance. The GTD family also yields optimum solutions for the QoS problem and the bit rate maximization problem. Many existing systems are identified to be special cases of the GTD-based system, and some new GTD-based transceivers were also indicated. The QR-based GTD has the advantage of offering a simple way to do limited-feedback by sending the bit allocation information from the receiver to transmitter. Further work is necessary to examine how the results can be extended to the case where ZF is not imposed.

Appendix

In this appendix we prove (18). For any $J \times P$ matrix $\mathbf{H}$ with rank $\geq M$ and $P \times M$ matrix $\mathbf{F}$, we show

$$\psi \doteq \frac{\det(\mathbf{F}^\dagger \mathbf{F})}{\det(\mathbf{F}^\dagger \mathbf{H}\mathbf{H}^\dagger \mathbf{F})} \geq \frac{1}{\prod_{k=1}^{M} \sigma_{h,k}^2}$$

where $\sigma_{h,k}$ are the M dominant singular values of $\mathbf{H}$. For this express $\mathbf{F}$ in SVD form:

$$\mathbf{F} = \mathbf{U}_f \begin{pmatrix} \Sigma_f \\ \mathbf{0} \end{pmatrix} \mathbf{V}_f^\dagger.$$

Thus we have

$$\mathbf{F}^\dagger \mathbf{F} = \mathbf{V}_g \begin{pmatrix} \Sigma_f^\dagger & \mathbf{0} \end{pmatrix} \begin{pmatrix} \Sigma_f \\ \mathbf{0} \end{pmatrix} \mathbf{V}_g^\dagger.$$

Therefore $\det(\mathbf{F}^\dagger \mathbf{F}) = \det(\mathbf{\Sigma}_f^\dagger \mathbf{\Sigma}_f)$, and

$$\mathbf{F}^\dagger \mathbf{H}\mathbf{H}^\dagger \mathbf{F} = \mathbf{V}_f (\Sigma_f^\dagger \ \mathbf{0}) \mathbf{U}_f^\dagger \mathbf{H}\mathbf{H}^\dagger \mathbf{U}_f \begin{pmatrix} \Sigma_f \\ \mathbf{0} \end{pmatrix} \mathbf{V}_f^\dagger$$

so that

$$\det(\mathbf{F}^\dagger \mathbf{H}\mathbf{H}^\dagger \mathbf{F}) = \det(\mathbf{\Sigma}_f^\dagger \mathbf{\Sigma}_f) \det([\mathbf{U}_f^\dagger \mathbf{H}\mathbf{H}^\dagger \mathbf{U}_f]_{M \times M}).$$

Thus $\psi = 1/\det([\mathbf{U}_f^\dagger \mathbf{H}\mathbf{H}^\dagger \mathbf{U}_f]_{M \times M})$. Next we know that

$$\det([\mathbf{U}_f^\dagger \mathbf{H}\mathbf{H}^\dagger \mathbf{U}_f]_{M \times M}) = \prod_{k=1}^{M} \mu_k$$

where μ_k are the singular values of $[\mathbf{U}_f^\dagger \mathbf{H}\mathbf{H}^\dagger \mathbf{U}_f]_{M \times M}$ in descending order. The $J \times J$ matrix $\mathbf{H}\mathbf{H}^\dagger$ has eigenvalues $\sigma_{h,1}^2 \geq \sigma_{h,2}^2 \geq \cdots \geq \sigma_{h,J}^2$, so it follows from the interlacing property for Hermitian matrices [8] that

$$\sigma_{h,1}^2 \geq \mu_1,\ \sigma_{h,2}^2 \geq \mu_2,\ \cdots \sigma_{h,M}^2 \geq \mu_M.$$

So $\psi = 1/\prod_{k=1}^{M} \mu_k \geq 1/\prod_{k=1}^{M} \sigma_{h,k}^2$, completing the proof.

References

[1] A. Barg and D. Y. Nogin, "Bounds on packings of spheres in the Grassmann manifold," *IEEE Trans. Inform. Theory*, vol. 48, no. 9, pp. 2450–2454, Sep. 2002.

[2] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge: Cambridge Univ. Press, 2004.

[3] "Optimal discrete bit loading for multicarrier modulation systems," in *Proc. IEEE 1998 IEEE Int. Symp. Information Theory (ISIT'98)*, Aug. 1998, p. 193.

[4] S. T. Chung, A. Lozano, H. C. Huang, A. Sutivong, and J. M. Cioffi, "Approaching the MIMO capacity with a low-rate feedback channel in V-BLAST," *EURASIP J. Applied Signal Process.*, vol. 5, pp. 762–771, 2004.

[5] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD: The Johns Hopkins Univ. Press, 1996.

[6] T. Guess, "Optimal sequences for CDMA with decision-feedback receivers," *IEEE Trans. Inform. Theory*, vol. 49, no. 4, pp. 886–900, Apr. 2003.

[7] R. A. Horn, "On the eigenvalues of a matrix with prescribed singular values," *Proc. Amer. Mathematical Soc.*, vol. 5, no. 1, pp. 4–7, Feb. 1954.

[8] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.

[9] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1991.

[10] Y. Jiang, J. Li, and W. W. Hager, "Joint transceiver design for MIMO communications using geometric mean decomposition," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3791–3803, Oct. 2005.

[11] Y. Jiang, J. Li, and W. W. Hager, "Uniform channel decomposition for MIMO communications," *IEEE Trans. Signal. Process.*, vol. 53, no. 11, pp. 4283–4294, Nov. 2005.

[12] Y. Jiang, W. W. Hager, and J. Li, "Generalized triangular decomposition," *Math. Comput.*, pp. 1037–1056, Oct. 2007.

[13] Y. Jiang, W. W. Hager, and J. Li, "Tunable channel decomposition for MIMO communications using channel state information," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4405–4418, Nov. 2006.

[14] Y. Jiang, W. W. Hager, and J. Li, "The geometric mean decomposition," *Linear Algebra and its Applications*, pp. 373–384, 2005.

[15] C. C. Li, Y. P. Lin, S. H. Tsai, and P. P. Vaidyanathan, "Optimization of transceivers with bit allocation to maximize bit rate for MIMO transmission," *IEEE Trans. Commun.*, vol. 57, no. 11, pp. 3556–3560, Dec. 2009.

[16] Y. P. Lin and S. M. Phoong, "Perfect discrete multitone modulation with optimal transceivers," *IEEE Trans. Signal Process.*, vol. 48, no. 6, pp. 1702–1711, Jun. 2000.

[17] Y. P. Lin and S. M. Phoong, "Optimal ISI-free DMT transceivers for distorted channels with colored noise," *IEEE Trans. Signal Process.*, vol. 49, no. 11, pp. 2702–2712, Nov. 2001.

[18] D. Love and R. W. Heath, Jr, "Limited feedback unitary precoding for spatial multiplexing systems," *IEEE Trans, Inf. Theory*, vol. 51, no. 8, pp. 2967–2976, Aug. 2005.

[19] D. Love, Tables of Complex Grassmannian Packings [Online]. Available: http://www.ece.purdue.edu/~djlove/grass.html

[20] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and its Applications*. New York: Academic Press, 1979.

[21] D. P. Palomar and S. Barbarossa, "Designing MIMO communication systems: Constellation choice and linear transceiver design," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3804–3818, Oct. 2005.

[22] D. P. Palomar and Y. Jiang, "MIMO transceiver design via majorization theory," *Foundations Trends Commun. Inf. Theory*, vol. 3, no. 4–5, pp. 331–551, 2006.

[23] D. P. Palomar, M. A. Lagunas, and J. M. Cioffi, "Optimum linear joint transmit-receive processing for MIMO channels with QoS constraint," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2381–2401, Sep. 2003.

[24] N. Prasad and M. K. Varanasi, "Analysis of decision feedback detection for MIMO rayleigh-fading channels and the optimization of power and rate allocations," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1009–1025, Jun. 2004.

[25] J. G. Proakis, *Digital Communicatins*. New York: McGraw-Hill, 2001.

[26] M. B. Shenouda and T. N. Davidson, "A framework for designing MIMO systems with decision feedback equalization or tomlinson-harashima precoding," *IEEE Select. Areas. Commun.*, vol. 26, no. 2, pp. 401–411, Feb. 2008.

[27] M. B. Shenouda and T. N. Davidson, "Limited feedback design of MIMO systems with zero-forcing DFE using Grassmann codebooks," in *Proc. IEEE Can. Workshop Information Theory*, Edmonton, ON, Canada, Jun. 2007, pp. 118–123.

[28] P. W. Shor and N. J. A. Sloane, "A family of optimal packings in the Grassmannian manifolds," *J. Algebraic Combin.*, vol. 7, pp. 157–163, 1998.

[29] H. Weyl, "Inequalities between the two kinds of eigenvalues of linear transformations," *Proc. Nat. Acad. Sci.*, vol. 35, pp. 408–411, Jul. 1949.

[30] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. IEEE Symp. Signals, Systems, and Electronics*, Pisa, Italy, Oct. 1998, pp. 295–300.

[31] F. Xu, T. N. Davidson, J. K. Zhang, and K. M. Wong, "Design of block transcivers with decision feedback detection," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 965–978, Mar. 2006.

[32] J. Zhang, A. Kavcic, and K. M. Wong, "Equal-diagonal QR decomposition and its application to precoder design for successive-cancellation detection," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 154–172, Jan. 2005.

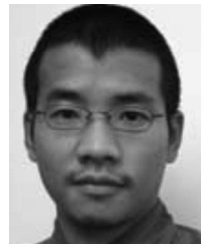
[33] S. Bergman, D. P. Palomar, and B. Ottersten, "Joint bit allocation and precoding for MIMO systems with decision feedback detection," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4509–4521, Nov. 2009.

[34] C. C. Weng, C. Y. Chen, and P. P. Vaidyanathan, "Joint optimization of transceivers with decision feedback and bit loading," in *Proc. 42nd Asilomar Conf. Signals, Systems, and Computers*, Monterey, CA, Oct. 2008, pp. 1310–1314.

[35] C. C. Weng, C. Y. Chen, and P. P. Vaidyanathan, "GTD-based transceivers for decision feedback and Bit Loading," in *Proc. IEEE Int. Conf. Acoustics Speech, and Signal Processing*, Taipai, Taiwan, R.O.C., Apr. 2009, pp. 2661–2664.

Ching-Chih Weng (S'07) was born in Taipei, Taiwan, R.O.C., on October 15, 1981. He received the B.S. degree from National Taiwan University (NTU), Taipei, in 2004 and the M.S. degree from California Institute of Technology (Caltech), Pasadena, CA, in 2007, both in electrical engineering. He is currently pursuing the Ph.D. degree in the field of digital signal processing at Caltech. His research interests include digital communications and radar signal processing.

In summer 2009, he was with Qualcomm Inc., CA, as an intern system engineer.



Chun-Yang Chen (S'06) was born in Taipei, Taiwan, R.O.C., on November 22, 1977. He received the B.S. and M.S. degrees in electrical engineering and communication engineering, both from National Taiwan University (NTU), Taipei, in 2000 and 2002, respectively. He received the Ph.D. degree in electrical engineering in the field of digital signal processing at California Institute of Technology (Caltech), Pasadena.

He is currently working for Facebook.

P. P. Vaidyanathan (S'80–M'83–SM'88–F'91) was born in Calcutta, India, on October 16, 1954. He received the B.Sc. (Hons.) degree in physics and the B.Tech. and M.Tech. degrees in radiophysics and electronics, all from the University of Calcutta, India, in 1974, 1977 and 1979, respectively, and the Ph.D. degree in electrical and computer engineering from the University of California at Santa Barbara in 1982.

He was a Postdoctoral Fellow at the University of California, Santa Barbara from September 1982 to March 1983. In March 1983, he joined the Electrical Engineering Eepartment ofthe California Institute of Technology as an Assistant Professor, and since 1993 has been Professor of Electrical Engineering there. His main research interests are in digital signal processing, multirate systems, wavelet transforms and signal processing for digital communications.

Dr. Vaidyanathan served as Vice-Chairman of the Technical Program committee for the 1983 IEEE International Symposium on Circuits and Systems, and as the Technical Program Chairman for the 1992 IEEE International Symposium on Circuits and Systems. He was an Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS for the period 1985–1987, and is currently an Associate Editor for IEEE SIGNAL PROCESSING LETTERS, and a Consulting Editor for the journal *Applied and Computational Harmonic Analysis*. He has been a Guest Gditor in 1998 for special issues of the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II, on the topics of filter banks, wavelets and subband coders. He has authored more than 360 papers in journals and conferences, and the book *Multirate Systems and Filter Banks* (Englewood Cliffs, NJ: Prentice-Hall, 1993). He has written several chapters for various signal processing handbooks. He was a recepient of the Award for Excellence in Teaching at the California Institute of Technology for the years 1983–1984, 1992–1993 and 1993–1994. He also received the NSF's Presidential Young Investigator award in 1986. In 1989, he received the IEEE ASSP Senior Award for his paper on multirate perfect-reconstruction filter banks. In 1990, he was recipient of the S. K. Mitra Memorial Award from the Institute of Electronics and Telecommuncations Engineers, India, for his joint paper in the *IETE* journal. In 2009, he was chosen to receive the IETE students' journal award for his tutorial paper in the *IETE Journal of Education*. He was also the coauthor of a paper on linear-phase perfect reconstruction filter banks in the IEEE TRANSACTIONS ON SIGNAL PROCESSING, for which the first author (T. Nguyen) received the Young Outstanding Author Award in 1993. He received the 1995 F. E. Terman Award of the American Society for Engineering Education, sponsored by Hewlett Packard Co., for his contributions to engineering education, especially the book *Multirate Systems and Filter Banks*. He has given several plenary talks including at the IEEE ISCAS-04, Sampta-01, Eusipco-98, SPCOM-95, and Asilomar-88 conferences on signal processing. He has been chosen a distinguished lecturer for the IEEE Signal Processing Society for the year 1996–1997. In 1999, he was chosen to receive the IEEE CAS Society's Golden Jubilee Medal. He is a recepient of the IEEE Signal Processing Society's Technical Achievement Award for the year 2002.