GTD-Based Transceivers for Decision Feedback and Bit Loading

Ching-Chih Weng, Chun-Yang Chen and P. P. Vaidyanathan
Dept. of Electrical Engineering, MC 136-93
California Institute of Technology, Pasadena, CA 91125, USA
E-mail: cweng@caltech.edu, cyc@caltech.edu, ppvnath@systems.caltech.edu

***Abstract*— We consider new optimization problems for transceivers with DFE receivers and linear precoders, which also use bit loading at the transmitter. First, we consider the MIMO QoS (quality of service) problem, which is to minimize the total transmitted power when the bit rate and probability of error of each data stream are specified. The developments of this paper are based on the generalized triangular decomposition (GTD) recently introduced by Jiang, Li, and Hager. It is shown that under some multiplicative majorization conditions there exists a custom GTD-based transceiver which achieves the minimal power. The problem of maximizing the bit rate subject to the total power constraint and given error probability is also considered in this paper. It is shown that the GTD-based systems also give the optimal solutions to the bit rate maximization problem.**[1]

***Index Terms* — Decision Feed-Back, BER Optimization, Generalized Triangular Decomposition, Bit Allocation, MIMO Transceiver.**

I. INTRODUCTION

In this paper we consider the optimization of multiple-input multiple-output (MIMO) communication systems with perfect channel state information (CSI) at both sides of the link. The focus of this paper will be on the system with a decision feedback equalizer at the receiver, and a linear precoder at the transmitter. The design methods for such a system have been considered by many authors when the bit constellations are fixed and identical for each sub-stream [13], [7], [8], [10], [14], [12], [9]. Similarly, when the channel and DFE are given, the bit loading scheme is a well treated problem [5]. Another subclass of optimization problems for such transceivers was considered in [2].

We consider two optimization problems for MIMO communication, both based on GTD (generalized triangular decomposition) reported in [8]. The first problem is to minimize the total transmitted power when the error probabilities and the bit rates of the substreams are fixed. The similar problem setting was discussed in [15] when each user is assigned the same number of sub-channels in the DMT system. The problem we are considering can be seen as the case in [15] where each user is assigned one sub-channel. We show that the optimal system can be designed by representing the channel in terms of some custom GTD, and choosing the transceiver matrices appropriately in terms of the GTD. Also, we consider the bit rate maximization problem while the transmitted power and the bit error rate are kept fixed. We also show the GTD system is the optimal solution for this problem.

This paper is structured as follows. In Section II, we will introduce the communication models and give explicit problem formulations. Section III gives the transceiver structure based on the generalized triangular decomposition of the channel matrix. Section IV proves the optimality of the GTD-based system for the two problems considered in the paper. Section V presents the numerical simulation results related to the topics discussed in the paper. The final conclusions of the paper are summarized in section VI.

II. SYSTEM MODEL AND PROBLEM FORMULATIONS

The transceiver considered in this paper is shown in Fig. 1, with the sizes of matrices indicated (e.g., $\mathbf{F}$ is $P \times M$, etc.). The additive channel noise is assumed to have covariance $\sigma_n^2\mathbf{I}$. Here $\mathbf{F}$ is the linear precoder, $\mathbf{H}$ is the channel, $\mathbf{G}$ is the feedforward part of the equalizer, and $\mathbf{B}$ is the feedback part. The decision device processes the vector $\widehat{\mathbf{s}}$ bottom-up sequentially, and the past decisions within a block are fedback via $\mathbf{B}$ to correct future decisions in the block. This causality of decision feedback is ensured by restricting $\mathbf{B}$ to be strictly upper triangular.

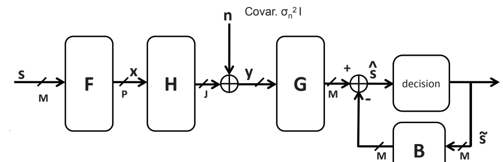

Fig. 1. The MIMO transceiver with linear precoder and DFE.

To understand how the problems of bit allocation and power minimization arise, we first examine the relationships between the error probabilities, bit rates and data stream powers. Assume the input signals are zero-mean uncorrelated processes representing independent data streams with powers P_k so that the input covariance is

$$\mathbf{\Lambda}_s = \text{diag}(P_1, P_2, \cdots, P_M). \tag{1}$$

Consider the situation where each data stream is represented with a different constellation size. Let us say the kth data stream uses b_k-bit QAM symbols with average power P_k. If the error at the kth sub-stream has variance $\sigma_{e_k}^2$, based on the low error and high bit rate assumption, it can be shown [11] that

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

where $P_e(k)$ are the symbol error probabilities. This equation expresses the average power to noise ratio required for the kth data stream to operate at the probability of error $P_e(k)$ with b_k-bit QAM constellation. Note that this formulation can be easily specialized to the single user DMT system when we set all $P_e(k)$ to be equal [11]. The total power transmitted can be written as $P_{trans} = \text{Tr}(\mathbf{F}\mathbf{\Lambda}_s\mathbf{F}^\dagger) = \text{Tr}(\mathbf{F}^\dagger\mathbf{F}\mathbf{\Lambda}_s) = \sum_{k=1}^{M} P_k[\mathbf{F}^\dagger\mathbf{F}]_{kk}$. Substituting from (2) we can rewrite this as

$$P_{trans} = \sum_{k=1}^{M} d_k 2^{b_k}\sigma_{e_k}^2[\mathbf{F}^\dagger\mathbf{F}]_{kk}, \tag{3}$$

where $d_k = \frac{1}{3}(Q^{-1}(\frac{P_e(k)}{4}))^2$, which is determined by the specified probability of error. It is usually assumed that the previous detected symbols $\widetilde{\mathbf{s}}$ in Fig. 1 are always correct. When we assume there is no error propagation, the zero forcing constraint can be written as

$$\mathbf{GHF} - \mathbf{B} = \mathbf{I} \tag{4}$$

[1]This work is supported in parts by the ONR grant N00014-08-1-0709, California Institute of Technology, and the TMS scholarship 94-2-A-018 of the National Science Council of Republic of China, Taiwan.

This means that the interference from other transmitted symbols is canceled out completely. Under the zero-forcing constraint, the error before the decision device for each sub-stream entirely comes from the channel noise. Since the channel noise has covariance $\sigma_n^2 \mathbf{I}$, the error variance before the kth input of the decision device is given by

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

From (3) the transmitted power can then be written as

$$P_{trans} = \sum_{k=1}^{M} c_k 2^{b_k} [\mathbf{F}^\dagger \mathbf{F}]_{kk} [\mathbf{G}\mathbf{G}^\dagger]_{kk}, \quad (6)$$

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

In this paper we consider two MIMO transceiver problems. The first one is the quality of service problem: to minimize the transmitted power subject to the given bit rate and probability of error constraints. The mathematical formulation of the problem is as follows:

$$\begin{aligned} \min_{\mathbf{F},\mathbf{G},\mathbf{B}} \quad & P_{trans} & (7)\\ \text{s.t.} \quad (a) \quad & \mathbf{GHF} = \mathbf{I} + \mathbf{B} \\ (b) \quad & \{c_k, b_k\} \text{ QoS for data stream } k. \end{aligned}$$

The second problem we want to address is bit rate maximization for fixed bit error rate and fixed transmitted power. Consider again the system shown in Fig. 1, with the zero-forcing constraint. For QAM modulation and under the high bit rate assumption [1], the bit loading formula can be approximated as

$$b_k \approx \log_2(P_k/(\sigma_{ek}^2 d_k)). \quad (8)$$

The problem of maximizing the average bit rate for fixed bit error rate and fixed transmitted power can be formulated as

$$\begin{aligned} \max_{\mathbf{F},\mathbf{G},\mathbf{B},\{P_k\}} \quad & b = \frac{1}{M}\sum_{k=1}^{M} \log_2(\frac{P_k}{\sigma_{ek}^2 d_k}) & (9)\\ \text{s.t.} \quad (a) \quad & \text{Tr}(\mathbf{F}\mathbf{\Lambda}_s \mathbf{F}^\dagger) \le P_{total} \\ (b) \quad & \mathbf{GHF} = \mathbf{I} + \mathbf{B}. \end{aligned}$$

We will see that the theorem of generalized triangular decomposition (GTD) helps to solve the above two problems.

III. GTD Systems

First, let us review the GTD theorem, which was recently introduced by Jiang, Li, and Hager [8].

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

$$\mathbf{a} \prec_\times \mathbf{h}, \quad (10)$$

where $\mathbf{a} = [|r_1|, |r_2|, \cdots, |r_K|]$, $\mathbf{h} = [\sigma_{h,1}, \sigma_{h,2}, \cdots, \sigma_{h,K}]$, and " $\prec_\times''$ stands for multiplicative majorization [8]. Then there exist matrices $\mathbf{R}$, $\mathbf{Q}$, and $\mathbf{P}$ such that

$$\mathbf{H} = \mathbf{Q}\mathbf{R}\mathbf{P}^\dagger, \quad (11)$$

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.

Proof: See [8]. ■

Let us first discuss the general GTD-based system. We will focus on the case with orthonormal precoders. It is shown in [3] there is no loss of optimality in designing the precoder to have orthonormal columns.

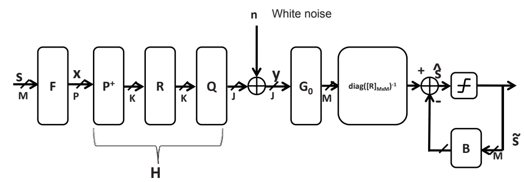

Fig. 2. The system with orthonormal linear precoding and DFE.

With the channel decomposed using the GTD as in (11) we now describe a method to construct the matrices $\{\mathbf{F}, \mathbf{G}, \mathbf{B}\}$. This design, with appropriate bit allocation, will be optimal in the sense described in Theorem 2 below. In (11) the matrix $\mathbf{R}$ is a $K \times K$ upper triangular matrix with the vector $\{[\mathbf{R}]_{M+1,M+1}, \cdots, [\mathbf{R}]_{K,K}\}$ equal to some permutation of the vector $\{\sigma_{h,M+1}, \cdots, \sigma_{h,K}\}$, which contains the smallest $K - M$ singular values of $\mathbf{H}$. The first M diagonal elements of $\mathbf{R}$, $\mathbf{r} = \{[\mathbf{R}]_{1,1}, \cdots, [\mathbf{R}]_{M,M}\} \in \mathcal{R}^+$, is multiplicatively majorized by the vector $\sigma = \{\sigma_{h,1}, \cdots, \sigma_{h,M}\}$, which contains the first M dominant singular values of $\mathbf{H}$. Here we assume the rank of the channel matrix $\mathbf{H}$ is K, and $K \ge M$. Note that this decomposition is possible because of the GTD theory [8]. Also we want to point out that with this decomposition we have

$$\prod_{k=1}^{M} [\mathbf{R}]_{kk}^2 = \prod_{k=1}^{M} \sigma_{h,k}^2, \quad (12)$$

which is a direct consequence of the multiplicative majorization relationship. This fact will be useful in later discussions.

Now consider Fig. 2. Suppose we choose the precoder $\mathbf{F}$ to be such that

$$\mathbf{P}^\dagger \mathbf{F} = \begin{pmatrix} \mathbf{I_M} \\ \mathbf{0} \end{pmatrix}, \text{ i.e., } \mathbf{F} = [\mathbf{P}]_{P \times M}. \quad (13)$$

Since $\mathbf{P}$ has orthonormal columns, $\mathbf{F}$ has orthonormal columns as well. The transmitted power will be

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

The matrix $\mathbf{G}_0$ will be chosen so that

$$\mathbf{G}_0 \mathbf{Q} = \begin{pmatrix} \mathbf{I_M} & \mathbf{0} \end{pmatrix}, \text{ i.e., } \mathbf{G}_0 = [\mathbf{Q}^\dagger]_{M \times J}. \quad (14)$$

Since $\mathbf{Q}$ has orthonormal columns, $\mathbf{G}_0$ has orthonormal rows, therefore the noise covariance after the filter $\mathbf{G}_0$ will be

$$E[\mathbf{G}_0 \mathbf{n}\mathbf{n}^\dagger \mathbf{G}_0^\dagger] = \mathbf{G}_0 E[\mathbf{n}\mathbf{n}^\dagger]\mathbf{G}_0^\dagger = \sigma_n^2 \mathbf{I}.$$

Thus the noise remains white after passing through the filter $\mathbf{G}_0$. The signal sub-streams then will pass through some multipliers $\{[\mathbf{R}]_{ij}^{-1}\}$ before the decision devices. Those multipliers can be equivalently viewed as a diagonal matrix multiplied with the signal vector. Thus the feedforward filter can be written as

$$\mathbf{G} = (\text{diag}([\mathbf{R}]_{M \times M}))^{-1} \mathbf{G}_0. \quad (15)$$

Therefore the signal transfer function from $\mathbf{s}$ to $\hat{\mathbf{s}}$ without the decision feedback will be

$$\mathbf{GHF} = (\text{diag}([\mathbf{R}]_{M \times M}))^{-1} [\mathbf{R}]_{M \times M}.$$

The feedback filter $\mathbf{B}$ is the one that makes the zero-forcing constraint satisfied, i.e.,

$$\mathbf{B} = \mathbf{GHF} - \mathbf{I} = (\text{diag}([\mathbf{R}]_{M \times M}))^{-1} [\mathbf{R}]_{M \times M} - \mathbf{I}. \quad (16)$$

Since $\mathbf{R}$ is an upper triangular matrix, it can be seen that $\mathbf{B}$ in (16) will be strictly upper triangular. In this scenario, the noise variance in the k-th substream will be

$$\sigma_{e_k}^2 = \sigma_n^2/[\mathbf{R}]_{kk}^2. \quad (17)$$

Substituting this into equation (3), the transmitted power needed to satisfy the specified QoS and bit rate constraints can be expressed as

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

IV. Optimality of the GTD-Based System

In this section, we will show that the GTD-based system is actually optimal for both of the problems discussed in Section II under some mild conditions.

A. Quality of Service (QoS) Problem

Now it is time to solve the QoS problem. The solution is given in the following theorem:

Theorem 2: Consider problem (7), then

(a) the minimum required power to achieve the specification will be greater or equal to

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

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

(b) This P_{min} is achievable if

$$\frac{\{c_1 2^{b_1}, \cdots, c_M 2^{b_M}\}}{c2^b/M} \prec_x \frac{\{\sigma_{h,1}^2, \cdots, \sigma_{h,M}^2\}}{(\prod_{k=1}^M \sigma_{h,k}^2)^{\frac{1}{M}}}. \quad (18)$$

Proof: Part (a) is clearly true since the problem (7) discussed in [2] is a relaxed version of the current problem (7).

Now let us prove part (b). Assume the rank of channel matrix is K. Suppose the given $\{c_k, b_k\}$ satisfies (18), then there exists a $K \times K$ upper triangular matrix $\mathbf{R}$, such that the decomposition

$$\mathbf{H} = \mathbf{Q}\mathbf{R}\mathbf{P}^\dagger$$

is true, where $\mathbf{Q}$ and $\mathbf{P}$ have orthonormal columns and the diagonal terms of $\mathbf{R}$ satisfy

$$[\mathbf{R}]_{kk} = \begin{cases} \frac{Mc_k 2^{b_k}(\prod_{k=1}^M \sigma_{h,k}^2)^{\frac{1}{M}}}{c2^b}, & \text{for } k = 1, 2, \cdots, M. \\ \sigma_{h,k}^2 & \text{, otherwise.} \end{cases} \quad (19)$$

Note that this factorization is possible because from (18) we have

$$\{[\mathbf{R}]_{kk}^2\}_{k=1}^M \prec_x \{\sigma_{h,k}^2\}_{k=1}^M$$

and by GTD theorem (theorem 1), such $\mathbf{R}$ exists. By using the precoder and equalizer in (13) and (15) as discussed in Sec. III, we are able to achieve P_{min} with equality. ∎

The intuition behind (18) is that, if the QoS constraint is less spread out than the channel singular values, it is possible to achieve the minimal P_{min} with equality. This system, which achieves the minimal P_{min} is called *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 $\{c_k, b_k\}$. Now it is clear that the GTD-based system has much more flexibility than the linear transceiver system. The custom GTD-based transceiver is computed from the channel $\mathbf{H}$ and the given QoS constraints $\{c_k, b_k\}$.

B. Max-Bit-Rate with Fixed-Power Problem

First, we observe that the power constraint can be rewritten as

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

Also, since the zero-forcing constraint is imposed, the noise comes entirely from the channel Gaussian noise. The noise variance can be written as $\sigma_{e_k}^2 = \sigma_n^2[\mathbf{G}\mathbf{G}^\dagger]_{kk}$. Now we will first find the optimal power P_k for given $\mathbf{F}$ and $\mathbf{G}$, under the total transmit power constraint. Based on the optimal power allocation P_k, we will then derive the optimal transceiver for maximizing the bit rate.

First of all, we observe that if $\{\mathbf{F}, \mathbf{G}, \mathbf{B}\}$ are given, the problem (9) is a convex problem in P_k for all k. Since the problem is convex and Slater's condition [4] is easily checked to be true, the duality gap is zero. Thus, we can first solve the optimal P_k, which gives the global optimum, and then substitute the formula of P_k and further solve for optimal $\{\mathbf{F}, \mathbf{G}, \mathbf{B}\}$. Similar to [1], to solve for the optimal P_k we first check the Karush-Kuhn-Tucker (KKT) condition [4]. Suppose P_k^* is the optimal value for problem (9), then the KKT condition states that there exists a constant α such that

$$\alpha \le 0, \quad (20)$$

$$\frac{\partial}{\partial P_k}\ \{\tfrac{1}{M}\textstyle\sum_{k=1}^M \log_2(\frac{P_k}{\sigma_{e_k}^2 d_k}) + \alpha(\sum_{k=1}^M P_k[\mathbf{F}^\dagger\mathbf{F}]_{kk} - P_{total})\}\ |_{P_k=P_k^*} = 0, \quad (21)$$

$$\alpha(\sum_{k=1}^M P_k[\mathbf{F}^\dagger\mathbf{F}]_{kk} - P_{total})|_{P_k=P_k^*} = 0. \quad (22)$$

By solving these equations, we can get the optimal power allocation

$$P_k^* = P_{total}/(M[\mathbf{F}^\dagger\mathbf{F}]_{kk}). \quad (23)$$

By substituting this into the bit rate, we have

$$\begin{aligned} b &= \sum_{k=1}^M \log_2\left(\frac{P_{total}}{Md_k[\mathbf{F}^\dagger\mathbf{F}]_{kk}\sigma_{ek}^2}\right)^{\frac{1}{M}} \\ &= \log_2\left(\prod_{k=1}^M \frac{P_{total}}{Mc_k[\mathbf{F}^\dagger\mathbf{F}]_{kk}[\mathbf{G}\mathbf{G}^\dagger]_{kk}}\right)^{\frac{1}{M}}. \end{aligned} \quad (24)$$

Therefore, the problem of maximizing bit rate now is reduced to maximizing (24) with the zero forcing constraint. Note that to maximize (24) is the same as to minimize Eq. (10) in [2]. This problem is already treated in detail in [2]. The maximized bit rate can be calculated from (24):

$$b_{max} = \log_2\left(\frac{P_{total}}{c}\left(\prod_{k=1}^M \sigma_{h,k}^2\right)^{\frac{1}{M}}\right). \quad (25)$$

As before, with the GTD-based system the maximum bit rate can be achieved. We explain this briefly in the following. Let us look at the system in Fig. 2. Choose $\mathbf{F} = [\mathbf{P}]_{P\times M}$, $\mathbf{G}_0 = [\mathbf{Q}^\dagger]_{M\times J}$, $\mathbf{G}$ as in (15), and $\mathbf{B}$ as in (16). Then the power in kth sub-stream will be

$$P_k = P_{total}/(M[\mathbf{F}^\dagger\mathbf{F}]_{kk}) = P_{total}/M,$$

and the noise variance will be $\sigma_{e_k}^2 = \sigma_n^2/[\mathbf{R}]_{kk}^2$. Substituting those back into (24), the average bit rate of this system will be

$$b = \log_2\left(\frac{P_{total}}{c}(\prod_{k=1}^{M}\sigma_{h,k}^2)^{\frac{1}{M}}\right) = b_{max}. \quad (26)$$

Therefore, the GTD-based system achieves the maximum bit rate. Thus, all the special cases of the systems discussed [2] are optimal in this sense as well.

V. Numerical Simulations

In this section we consider wireless communication systems with multiple antennas at both sides of the link with perfect channel state information. We use 100 randomly generated MIMO channels for the simulation. The matrix channel is of size 5×4, and normalized so that $E[|[\mathbf{H}]_{i,j}|^2] = 1$.

We implement five methods in the numerical simulations. "SVD", "GMD", "QR", "BID" and stand for the special cases of the GTD-based transceiver structures discussed in Section. IV of [2]; "GB" is the custom GTD-based system derived from the section III, where $[\mathbf{R}]_{kk}$ is obtained from equality (19) with given equal c_k. The additive noise is complex circulant Gaussian with average power normalized to 0 dB. The results are given in terms of the uncoded bit error rate versus transmitted power. Since the resulting optimal system acts like parallel independent Gaussian channels, scalar channel coding can be further added in each sub-stream to reduce the probability of error.

Fixed, identical constellation: In Fig. 3 we consider the system with a given fixed and identical constellation in each sub-stream. In this simulation, each channel is given 6 bits, therefore, a 64-QAM constellation is imposed. It can be observed that the "GB" method performs similar to "GMD", and those two outperform all other methods. This is because with equal constellation, our custom "GB" actually reduces to the "GMD" system, which is optimal in terms of BER performance [6]. For all other schemes, since the resulting sub-channel gains are quite different, it is not surprising that the equal bit allocation scheme would perform badly.

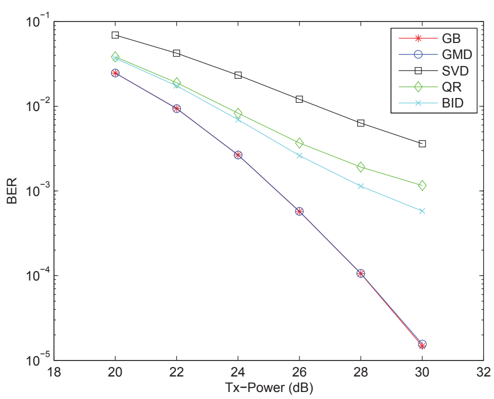

Fig. 3. BER versus Tx-Power when all constellations are fixed as 64-QAM.

Fixed, nonidentical constellations: In Fig. 4 we consider the system with fixed but different constellations in different sub-streams. In this simulation, the bits allocated to the sub-streams are forced to be $[8,\ 8,\ 6,\ 6]$, which is 256-QAM, 256-QAM, 64-QAM, and 64-QAM, respectively. It can be observed that "GB" outperforms all other methods significantly. However, among the other four methods, there is no theory about which one performs better than which.

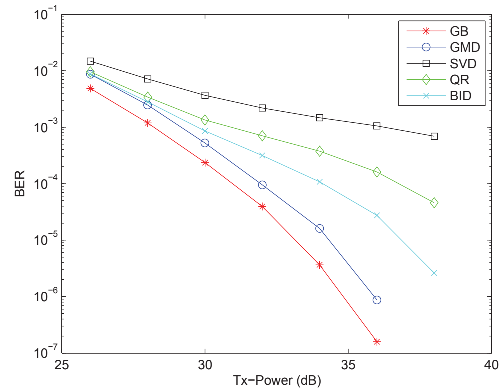

Fig. 4. BER versus Tx-Power when constellations are fixed as 256-QAM, 256-QAM, 64-QAM, and 64-QAM, respectively.

VI. 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 is 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 transmitted power is identical for a given bit rate and error probability. We also proved that any GTD-based system achieves the optimal performance. We also considered the quality of services problem and showed that there is a custom GTD-based system which gives the minimal power. Both the theoretical analysis and numerical simulations have been provided to validate the effectiveness of our results.

References

[1] C. C. Li, S. H. Tsai, Y. P. Lin, and P. P. Vaidyanathan, "Optimal Zero-Forcing Transceiver Design for Maximizing Bit Rate Subject to a Total Transmit Power Constraint" *Proc. 16th European Signal Processing Conference.* Lausanne, 2008.

[2] C. C. Weng, C. Y. Chen, and P. P. Vaidyanathan, "Joint Optimization of Transceivers with Decision Feedback and Bit Loading," *IEEE Asilomar Conference.* Monterey Bay, USA, Nov. 2008. http://www.systems.caltech.edu/cyc/dfe-bitloading.pdf

[3] C. C. Weng, C. Y. Chen, and P. P. Vaidyanathan, "MIMO Transceivers with Decision Feedback and Bit Loading: Theory and Optimization," *submitted to IEEE Trans. Sig. Proc.*

[4] S. Boyd, and L. Vandenberghe, *Convex optimization,* Cambridge. 2004.

[5] J. Campello, "Optimal Discrete Bit Loading for Multicarrier Modulation Sysrtems" *IEEE ISIT.* pp. 193, Aug. 1998.

[6] Y. Jiang, J. Li, and W. W. Hager "Joint transceiver design for MIMO communications using geometric mean decomposition," *IEEE Trans. Sig. Proc.,* pp.3791-3803, Oct. 2005.

[7] Y. Jiang, J. Li, and W. W. Hager "Uniform channel decomposition for MIMO communications," *IEEE Trans. Sig. Proc.,* pp.4283-4294, Nov. 2005.

[8] Y. Jiang, W. W. Hager, and J. Li "Generalized triangular decomposition," *Mathematics of computation.,* Oct. 2007.

[9] Y. Jiang, W. W. Hager, and J. Li "Tunable channel decomposition for MIMO communications using channel state information," *IEEE Trans. Sig. Proc.,* pp. 4405 - 4418, Nov. 2006.

[10] Y. Jiang, D. P. Palomar, and M. K. Varanasi, "Precoder Optimization for Nonlinear MIMO Transceiver Based on Arbitrary Cost Function," *Proc. Conference on Information Sciences and Systems.* March, 2007.

[11] Y. P. Lin, and S. M. Phoong, "Optimal ISI-Free DMT transceivers for distorted channels with colored noise," *IEEE Trans. Sig. Proc.* pp. 2702 - 2712, Nov. 2001.

[12] M. B. Shenouda, T. N. Davidson, "A framework for designing mimo systems with decision feedback equalization or tomlinson-harashima precoding," *IEEE Journal on Selected Areas in Communications,* pp. 401 - 411, Feb. 2008.

[13] F. Xu, T. N. Davidson, J. K. Zhang, S. S. Chan, and K. M. Wong, "Design of block transceivers with decision feedback detection," *IEEE Trans. Sig. Proc.,* pp. 965 - 978, March. 2006.

[14] J. Zhang, A. Kavcic, and K. M. Wong, "Equal-diagonal QR decomposition and its application to precoder design for successive-cancellation detection," *IEEE Trans. Info. Theory,* pp. 154 - 172, Jan. 2005.

[15] A. Pandharipande, and S. Dasgupta, "Optimum DMT-based transceivers for multiuser communications," *IEEE Trans. Communications,* pp. 2038 - 2046, Dec. 2003.