

Transceiver Design with Vector Perturbation Technique and Iterative Power Loading

Ching-Chih Weng, and P. P. Vaidyanathan

Dept. of Electrical Engineering, MC 136-93

California Institute of Technology, Pasadena, CA 91125, USA

E-mail: cweng@caltech.edu, ppvnath@systems.caltech.edu

Abstract—In this paper we consider the optimization of transceivers which use the nonlinear vector perturbation technique at the transmitter. Since the perturbation vector can be almost totally removed at the receiver, the transmitter can use this extra freedom to reduce the transmitted power while maintaining the performance. The two cases considered in this paper are linear transceivers and transceivers with decision feedback (DFE). For both cases, efficient iterative power loading algorithms are developed to reduce the average bit error rate under the total transmitted power constraint. We present simulation results showing that the proposed technique performs better than the existing state-of-the-art uniform channel decomposition (UCD) system and the vector perturbation (VP) precoder.¹

Index Terms — Decision Feed-Back, BER Optimization, Vector Perturbation, Integer Least Square Problem, MIMO Transceiver.

I. INTRODUCTION

In this paper we revisit the optimization of multiple-input multiple-output (MIMO) communication systems. The focus of this paper will be on combining the vector perturbation technique [5] at the transmitter with the existing optimal linear [10] and DFE transceiver designs [11]. Since the perturbation vector can be almost totally removed at the receiver, the transmitter can use this extra freedom to reduce the transmitted power while the performance is maintained [5]. The MMSE transceiver with vector perturbation is discussed in [4]. The vector perturbation technique requires heavy computational complexity for solving the integer least square problem. The sphere decoding algorithm, which is originally a receiver technique, can be used here [5] to speed up the computation. In [2], Windpassinger et al. adopted the lattice-reduction based approximation for finding the perturbation vector, resulting in a polynomial time algorithm.

However, usually the goal of the communication system is to minimize the average bit error rate (BER). This was not addressed in the literature known to the authors so far, and is the goal of this paper. Therefore, instead of MMSE, we propose the designs which aim at minimizing the average BER under the total transmitted power constraint.

First we modify the linear and DFE transceiver design problems, by incorporating the vector perturbation technique under the total power constraint. Since the optimization problems are intractable, we propose iterative designs, based on existing optimal solutions for linear transceivers and DFE transceivers. We first compute the designs given in [10] and [11]. Based on these existing designs, we compute the perturbation vectors to reduce the transmitted power. Then, we update the power loading matrix in the precoder, which is guaranteed to give lower average bit error rate while satisfying the power constraint. The procedure can be iterated until no significant improvement can be made. The iterative algorithms suggested in this paper will result in better performance and guarantee convergence. The

¹This 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.

resulting system is shown to perform better than the state-of-the-art systems [10] and [7].

This paper is structured as follows. In Section II, we will present the communication models used in the paper, and also give explicit problem formulations. In Section III, we will talk about how to iteratively design the transceivers. Section IV presents the numerical simulation results related to the topics discussed in the paper. The final conclusions of the paper are summarized in section V.

II. SYSTEM MODEL AND PROBLEM FORMULATIONS

We begin by introducing the transceivers for frequency flat MIMO channels. Fig. 1 shows the system with linear receiver and the vector perturbation in front of the linear precoder. Fig. 2 shows the system with DFE, and the vector perturbation in front of the linear precoder. These are the two systems we focus on. Before the precoder matrix \mathbf{F} , the signal \mathbf{s} is modified by the perturbation vector $\tau\mathbf{u}$, where \mathbf{u} is a complex integer vector, and $\tau \geq 0$. The constant τ is chosen to be large enough for the receiver to determine \mathbf{u} without ambiguity [5]. One example of choices is [5]

$$\tau = 2(|c|_{max} + \frac{\delta}{2}), \quad (1)$$

where $|c|_{max}$ is the absolute value of the constellation symbols with largest magnitude, and δ is the spacing between constellation points. The received signal for both cases can be written as

$$\mathbf{y} = \mathbf{H}\mathbf{F}(\mathbf{s} + \tau\mathbf{u}) + \mathbf{n}.$$

For the system in Fig. 1, before the decision device, we use the operation of detecting and removing the perturbation vector $\tau\mathbf{u}$. The detection of \mathbf{u} can be done first by identifying the Voronoi region of the (complex M -D) integer lattice to which $\mathbf{G}\mathbf{y}$ belongs. Since the perturbation vector $\tau\mathbf{u}$ is large compared to the signal vector \mathbf{s} , and the total transfer function $\mathbf{G}\mathbf{H}\mathbf{F}$ is approximately identity in high SNR regime, the detection of $\tau\mathbf{u}$ is almost perfect. After detecting \mathbf{u} , we form $\hat{\mathbf{s}} = \mathbf{G}\mathbf{y} - \tau\mathbf{G}\mathbf{H}\mathbf{F}\mathbf{u}$, which is to be sent to the decision device. The reason for doing this is that, if the transfer function is not unity, τ will not be totally removed and acts like interference to the transmitted symbols. Since $\tau\mathbf{u}$ is large compared to the transmitted symbols, this interference will hurt the system performance. In order to apply our analysis under the assumption that the detection of $\tau\mathbf{u}$ is correct, we adopt this perturbation-vector-remover strategy. The assumption of perfect detection of \mathbf{u} is true when the SNR is high since τ is chosen large enough.

The perturbation vector remover in Fig. 2 operates similar to what is described in the above paragraph, but it operates successively on one component at a time. Recall here that \mathbf{B} is strictly upper triangular [11]. The detected $\hat{\mathbf{u}}$ is fed back

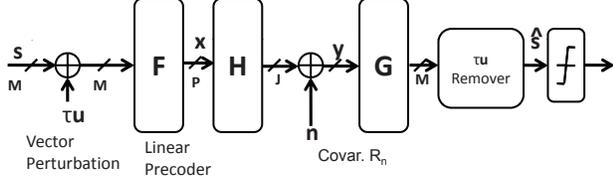


Fig. 1. Linear transceiver with vector perturbation.

with the decision of \tilde{s} to cancel the interference. The analysis in this paper assumes perfect detection of $\tilde{\mathbf{u}}$ and \tilde{s} .

The optimization problem we want to solve, is to minimize the average BER for the systems in Fig. 1 and Fig. 2, under the average transmitted power constraint. The mathematical formulation of the problem is as follows:

$$\begin{aligned} \min_{\mathbf{F}, \mathbf{G}, \mathbf{B}, \mathbf{u}} \quad & \text{average BER} \\ \text{s.t.} \quad & \text{Tr}(\mathbf{F} \cdot E[(\mathbf{s} + \tau\mathbf{u})(\mathbf{s} + \tau\mathbf{u})^\dagger] \cdot \mathbf{F}^\dagger) \leq P_{total}, \end{aligned} \quad (2)$$

where for the system in Fig. 1, the feedback matrix \mathbf{B} is set to be $\mathbf{0}$.

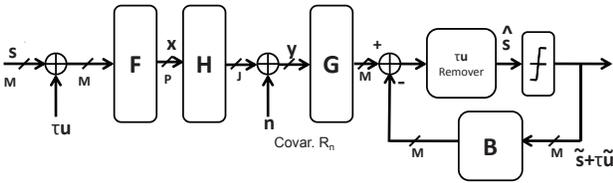


Fig. 2. DFE transceiver with vector perturbation.

It can be seen that the problem (2) is very difficult since the selection of \mathbf{u} is an integer optimization problem. It is also not clear how to optimize $\{\mathbf{F}, \mathbf{G}, \mathbf{B}\}$ and \mathbf{u} jointly. In the later sections, we will develop an iterative algorithm to find the suboptimal solution to this problem. The proposed design technique is guaranteed to converge.

III. ITERATION FOR POWER LOADING MATRICES

In this section we will develop an iterative approach to solve the problem (2). We will start from the optimal solution for linear transceivers [10] and DFE systems [9] with no perturbation vectors \mathbf{u} . Then we will run the integer programming problem for deciding \mathbf{u} , and develop an iterative procedure for updating the power loading matrix of the precoder and the perturbation vector. This iterative procedure will guarantee that the average BER decreases in each iteration. The convergence is also guaranteed.

A. Linear Transceiver

By the assumption that in the receiver $\tau\mathbf{u}$ is removed completely before the decision device, the average BER does not depend on the perturbation vector \mathbf{u} . The vector \mathbf{u} serves only to reduce the transmitted power while maintaining the original symbol detection performance. For the linear transceivers without vector perturbation technique, [10] gives the optimal solution for the design of transmitter and receiver. We will start from the solution given in [10], then by iteratively updating the power loading matrix, reduce the average BER.

The optimal solution in [10] is as follows: the optimal precoder is $\mathbf{F} = \tilde{\mathbf{U}}_h \Phi_f \mathbf{V}_f$, where $\tilde{\mathbf{U}}_h$ is the top-left $P \times$

M matrix of \mathbf{U}_h . Here \mathbf{U}_h is a unitary matrix such that $\mathbf{H}^\dagger \mathbf{R}_n^{-1} \mathbf{H} = \mathbf{U}_h \Lambda_h \mathbf{U}_h^\dagger$, where $\Lambda_h = \text{diag}(\lambda_{h,1}, \dots, \lambda_{h,P})$ is a diagonal matrix, and the entries are the eigenvalues of $\mathbf{H}^\dagger \mathbf{R}_n^{-1} \mathbf{H}$ with non-increasing order. The diagonal power loading matrix is computed as in [10] to minimize the total AM-MSE. Here \mathbf{R}_n represents the noise covariance matrix. The goal of \mathbf{V}_f is to force the diagonal entries of the MSE matrix to be equal. Examples of such \mathbf{V}_f are the $M \times M$ Hadamard matrix or the DFT matrix. The optimal receiver is the corresponding LMMSE solution.

Now suppose we are given the initial value $\mathbf{F} = \tilde{\mathbf{U}}_h \Phi_f \mathbf{V}_f$, we can compute the perturbation vectors $\tau\mathbf{u}$ for each of the transmitted \mathbf{s} , to minimize the transmitted power. This is done by solving the following problem

$$\mathbf{u}_{opt} = \arg \min_{\mathbf{u}} \|\mathbf{F}(\mathbf{s} + \tau\mathbf{u})\|_2^2 \quad (3)$$

This problem is an integer least square problem, which is known to be NP-hard. However, there exists efficient algorithms to compute \mathbf{u} . The interested reader can refer to [3] and [1] for more details about the integer least square problem.

After computing (3), the transmitted covariance matrix will be

$$E[(\mathbf{s} + \tau\mathbf{u})(\mathbf{s} + \tau\mathbf{u})^\dagger] = \mathbf{C}_f. \quad (4)$$

Now we restrict our updated precoder \mathbf{F} to be in the form:

$$\mathbf{F} = \tilde{\mathbf{U}}_h \Phi_{new} \mathbf{V}_f,$$

where Φ_{new} is a $M \times M$ diagonal matrix. This means that we only update the power loading matrix in precoder \mathbf{F} . The goal is to perform this update to reduce the average BER. $\tilde{\mathbf{U}}_h$ serves to diagonalize the channel, and \mathbf{V}_f to distribute the MSE to each substream equally. To minimize average BER, the power loading matrix Φ_{new} should minimize the sum of MSEs. This follows from the logic developed in [10], which uses the fact that the average BER is a Schur-convex function of MSEs. Therefore, the problem of finding the power loading matrix in \mathbf{F} , which minimizes the sum of MSE while satisfying the total power constraint can be formulated as follows:

$$\begin{aligned} \min_{\{\phi_{new,i}\}} \quad & \sum_i^M \frac{1}{1 + \phi_{new,i}^2 \lambda_{h,i}} \\ \text{s.t.} \quad & \sum_i^M c_i \phi_{new,i}^2 \leq P_{total}, \end{aligned} \quad (5)$$

where c_i is the i th diagonal term of $\mathbf{V}_f \mathbf{C}_f \mathbf{V}_f^\dagger$, and $\phi_{new,i}$ is the i th diagonal term of the new power loading matrix Φ_{new} . The inequality constraint comes from $\text{Tr}(\mathbf{F} \mathbf{C}_f \mathbf{F}^\dagger) = \text{Tr}(\tilde{\mathbf{U}}_h \Phi_{new} \mathbf{V}_f \mathbf{C}_f \mathbf{V}_f^\dagger \Phi_{new}^\dagger \tilde{\mathbf{U}}_h^\dagger) = \text{Tr}(\Phi_{new}^\dagger \Phi_{new} \mathbf{V}_f \mathbf{C}_f \mathbf{V}_f^\dagger) = \sum_i c_i \phi_{new,i}^2$.

If we use the substitution $\beta_i = c_i \phi_{new,i}^2$ and $\alpha_i = \lambda_{h,i} / c_i$, then the problem can be cast in exactly the same form as equation (21) in [10]. By the KKT condition [6], the problem can be solved efficiently and has a nice water-filling interpretation:

$$\phi_{new,i}^2 = \left(\sqrt{\frac{D}{c_i \lambda_{h,i}}} - \frac{1}{\lambda_{h,i}} \right)^+ \quad (6)$$

where D is the water-filling level to satisfy the power constraint with equality. Note that every time the updating of the power loading matrix will reduce the sum of MSE, thus reducing the average BER. The next iteration starts from computing the perturbation vectors according to the new precoder \mathbf{F} .

To summarize, the algorithm for updating the linear transceiver with the vector perturbation technique is as follows:

- (1) Compute initial \mathbf{F} according to [10].
- (2) Iteration:
 - Compute \mathbf{u}_{opt} by (3).
 - Replace \mathbf{C}_f by (4).
 - Replace $\phi_{new,i}^2$ by (6).
 - Replace \mathbf{F} by $\tilde{\mathbf{U}}_h \Phi_{new} \mathbf{V}_f$.
 - If MSE is reduced significantly, go to (2).
 - Otherwise, go to (3).
- (3) Compute \mathbf{G} corresponding to the new \mathbf{F} .

B. DFE with Linear Precoder

Based on the assumption of no error propagation, [7] gives the optimal solution for minimizing the average BER in a DFE transceiver under the total power constraint. The solution is

$$\mathbf{F} = \tilde{\mathbf{U}}_h \Phi \mathbf{V},$$

where $\tilde{\mathbf{U}}_h$ is as in Section III-A, Φ is a diagonal matrix as in [11], and \mathbf{V} now is the unitary matrix such that the GTD (generalized triangular decomposition) [8] of $(\mathbf{I} + \Phi^\dagger \tilde{\Lambda}_h \Phi)^{-\frac{1}{2}}$ is

$$(\mathbf{I} + \Phi^\dagger \tilde{\Lambda}_h \Phi)^{-\frac{1}{2}} = \mathbf{Q} \mathbf{R} \mathbf{V}^\dagger,$$

where $\tilde{\Lambda}_h$ is the $M \times M$ leading principal sub-matrix of Λ_h . Here \mathbf{Q} is a unitary matrix, and \mathbf{R} is an upper triangular matrix with equal diagonal terms. The optimal \mathbf{G} and \mathbf{B} are given as in [11]. By this design, the average BER actually depends on the minimization of geometric mean of MSEs only, since the existence of \mathbf{V} will make the MSE in each substream equal to the geometric mean of MSEs. With this given precoder \mathbf{F} , we compute the perturbation vector to reduce the transmitted power. Again, this is done by solving the problem (3), and the result we will get is the covariance matrix of the transmitted signal \mathbf{C}_f as in (4). The transmitted power is reduced to $\text{Tr}(\mathbf{F} \mathbf{C}_f \mathbf{F}^\dagger)$.

Suppose now we restrict our precoder \mathbf{F} to be in the form

$$\mathbf{F} = \tilde{\mathbf{U}}_h \Phi_{new} \mathbf{V},$$

where Φ_{new} is a diagonal matrix, and \mathbf{V} is the same unitary matrix as discussed above. Now let us discuss how to update the power loading matrix in \mathbf{F} to minimize the average BER. Here we introduce an artificial restriction on Φ_{new} , i.e., Φ_{new} is a diagonal matrix such that

$$(\mathbf{I} + \Phi_{new}^\dagger \tilde{\Lambda}_h \Phi_{new})^{-\frac{1}{2}} = \alpha (\mathbf{I} + \Phi^\dagger \tilde{\Lambda}_h \Phi)^{-\frac{1}{2}}. \quad (7)$$

By this set up, the following formula holds:

$$(\mathbf{I} + \Phi_{new}^\dagger \tilde{\Lambda}_h \Phi_{new})^{-\frac{1}{2}} = \mathbf{Q}(\alpha \mathbf{R}) \mathbf{R}^\dagger.$$

Therefore, the resulting MSE matrix will still have equal diagonal terms if we use the corresponding \mathbf{G} and \mathbf{B} as discussed in [11]. By adopting Φ_{new} , the average transmitted power becomes $\text{Tr}(\mathbf{F} \mathbf{C}_f \mathbf{F}^\dagger) = \text{Tr}(\tilde{\mathbf{U}}_h \Phi_{new} \mathbf{V} \mathbf{C}_f \mathbf{V}^\dagger \Phi_{new}^\dagger \tilde{\mathbf{U}}_h^\dagger) = \text{Tr}(\Phi_{new}^\dagger \Phi_{new} \mathbf{V} \mathbf{C}_f \mathbf{V}^\dagger) = \sum_i^M \phi_{new,i}^2 d_i$, where d_i is the i th diagonal terms of $\mathbf{V} \mathbf{C}_f \mathbf{V}^\dagger$. From (7) we have

$$\frac{1}{1 + \phi_{new,i}^2 \lambda_{h,i}} = \frac{\alpha^2}{1 + \phi_i^2 \lambda_{h,i}},$$

i.e.,

$$\phi_{new,i}^2 = \left(\frac{1 + \phi_i^2 \lambda_{h,i}}{\alpha^2} - 1 \right) / \lambda_{h,i}. \quad (8)$$

The problem of deciding α while maintaining the total power constraint can be written as

$$\begin{aligned} \min \quad & \alpha \\ \text{s.t.} \quad & \sum_i^M \left(\frac{1 + \phi_i^2 \lambda_{h,i}}{\alpha^2} - 1 \right) \frac{d_i}{\lambda_{h,i}} \leq P_{total}. \end{aligned} \quad (9)$$

Note that α is also restricted to be less than or equal to 1. The problem of computing α is just a line search to satisfy the equality in (10). The MSE will be α times the original one, and the average BER will decrease correspondingly [11]. This iterative algorithm is guaranteed to converge, and gives lower average BER performance.

To summarize, the algorithm is as follows:

- (1) Compute initial \mathbf{F} according to [11].
- (2) Iteration:
 - Compute \mathbf{u}_{opt} by (3).
 - Replace \mathbf{C}_f by (4).
 - Replace α by solving (9).
 - Replace $\phi_{new,i}^2$ by (8).
 - Replace \mathbf{F} by $\tilde{\mathbf{U}}_h \Phi_{new} \mathbf{V}$.
 - If MSE is reduced significantly, go to (2).
 - Otherwise, go to (3).
- (3) Compute \mathbf{G} and \mathbf{B} corresponding to the new \mathbf{F} .

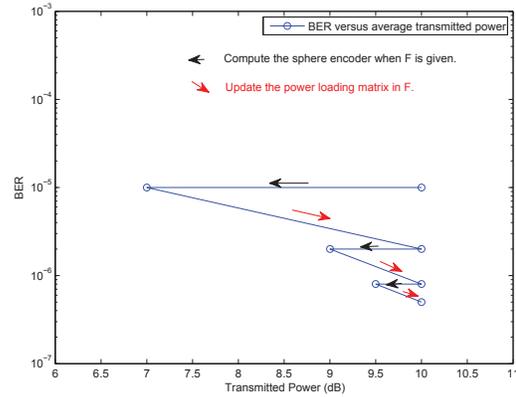


Fig. 3. Illustration of the iteration process.

In a linear transceiver, a fixed Hadamard or DFT matrix \mathbf{V}_f can be used to equalize the MSEs in the substreams. But in the equal-diagonal QR decomposition [7] \mathbf{V} is not fixed and changes with iteration. This is why we introduce the artificial constant α in (7) to facilitate the update procedure.

In both of the cases in this section, the iterating process can be illustrated as in the sketch of Fig. 3. The figure shows the case where total power constraint is 10 dB. The initial value starts at the transmitted power 10 dB (with additive Gaussian noise with power normalized to 0 dB) with average BER performance about 10^{-5} . After computing the perturbation

vector, the transmitted power will decrease while the average BER performance is kept fixed. The updating of the power loading matrix will again increase the transmitted power back to 10 dB, at the same time reducing the average BER. This process will converge eventually. The power loading updating procedure introduced in this paper is not restricted to be used with the vector perturbation scheme, but may be combined with other possible transmitted power reduction schemes as well.

IV. NUMERICAL SIMULATIONS

In this section, we simulate the uncoded BER performance of the proposed MIMO transceiver schemes. Since the overall system is diagonal, which acts like independent scalar Gaussian channels, channel coding may be further added to reduce the BER. We use QPSK gray encoded signaling, and the scalar τ for vector perturbation is calculated by Eq. (1). The channel is chosen randomly. For the error rate performance comparison, we plot the average BER against the average total transmitted power. The noise is assumed to be additive complex white Gaussian noise, with power normalized to 0 dB. To compute Eq. (3), we use Fincke and Pohst approach [3] to optimally solve the integer least square problem.

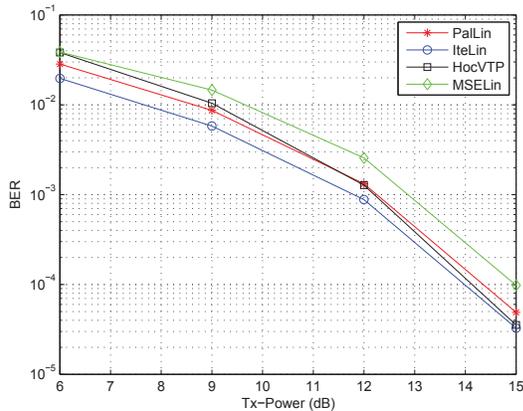


Fig. 4. BER performance with various MIMO transceivers with $J = 4$ and $P = M = 3$. The constellation used is QPSK. The schemes considered are: vector perturbation for inverse channel precoder proposed in [5] (HocVTP); the MMSE linear transceiver without Hadamard or DFT matrix in the precoder [10] (MSELin); the optimal linear transceiver proposed in [10] (PalLin); the proposed iterative scheme in Section III-A (IteLin).

Fig. 4 and Fig. 5 show the Monte Carlo simulation results for different methods relevant to this paper. Fig. 4 is for linear transceivers and Fig. 5 is for DFE transceivers. The schemes considered are: (1) vector perturbation for inverse channel precoder proposed in [5] (HocVTP); here we use $\mathbf{F} = \frac{1}{\delta} \mathbf{H}^\sharp$, where \mathbf{H}^\sharp is the pseudo inverse of \mathbf{H} , and δ is a constant to satisfy the average transmitted power constraint; (2) the optimal linear transceiver proposed in [10] (PalLin); (3) the MMSE linear transceiver without Hadamard or DFT matrix in the precoder [10] (MSELin); (4) the proposed iterative scheme in Section III-A (IteLin); (5) the UCD scheme proposed in [7] or [11] (UCDDFE); and (6) the proposed iterative scheme in Section III-B (IteDFE). For Fig. 4, it can be seen that with the vector perturbation technique and the precoder updating procedures mentioned in Section III, the performance is improved further for linear transceiver schemes. For Fig. 5, about extra 0.6 dB gain can be achieved at the BER of 10^{-5} for the IteDFE scheme compared to the UCD scheme

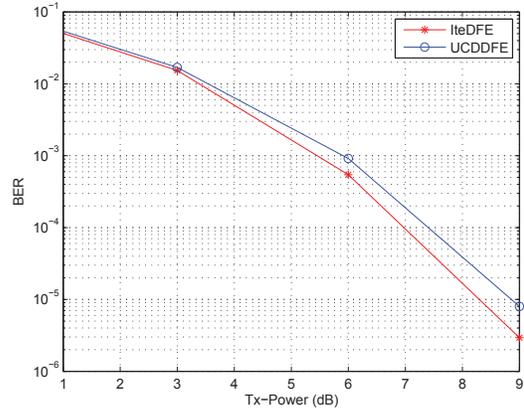


Fig. 5. BER performance with various MIMO transceivers with $J = 5$ and $P = M = 4$. The constellation used is QPSK. The schemes considered are: the UCD scheme proposed in [7] or [11] (UCDDFE); and the proposed iterative scheme in Section III-B (IteDFE).

[7]. Since in [7] the author showed that the UCD scheme is a mutual information maximization scheme, the gain comes from the vector perturbation technique, which acts to modify the constellation to be more suitable for the UCD.

V. CONCLUDING REMARKS

We presented an iterative transceiver design method by incorporating the existing transceiver scheme and the vector perturbation technique. It was shown that the average BER can be further improved under the total power constraint, compared to the state-of-the-art UCD design. Our simulation studies verified our analysis.

REFERENCES

- [1] L. Babai, "On Lovasz' Lattice Reduction and the Nearest Lattice Point Problem" *Combinatorica*, vol. 6, no. 1, pp. 1-13, Mar. 1986.
- [2] C. Windpassinger, R. F. H. Fischer, and J. B. Huber, "Lattice-Reduction-Aided Broadcast Precoding," *IEEE Trans. Commun.* vol. 52, no. 12, pp. 2057-2060, Dec. 2004.
- [3] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, "Closest Point Search in Lattices," *IEEE Trans. Info. Theory*, vol. 48, no. 8, pp. 2201-2214, Aug. 2002.
- [4] D. Schmidt, M. Joham, and W. Utschick, "Minimum Mean Square Error Vector Precoding," *Proc. IEEE PIMRC*, pp. 107 - 111, Sept. 2005.
- [5] B. M. Hochwald, C. B. Peel, and A. Lee Swindlehurst, "A Vector-Perturbation Technique for Near-Capacity Multiantenna Multiuser Communication— Part II: Perturbation," *IEEE Trans. Comm.* pp. 537 - 544, Mar. 2005.
- [6] S. Boyd, and L. Vandenberghe, *Convex optimization*, Cambridge. 2004.
- [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, 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.
- [10] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: a unified framework for convex optimization," *IEEE Trans. Sig. Proc.* pp. 2381 - 2401, Sept. 2003.
- [11] M. B. Shenoouda, T. N. Davidson, "A framework for designing MIMO systems with decision feedback equalization or Tomlinson-Harashima precoding," *Proc. of the ICASSP*, pp. 209-212, April. 2007.
- [12] Y. Jiang, M. K. Varanasi, and D. P. Palomar, "Decision feedback based transceiver optimization for MIMO intersymbol interference channels," *Proc. 41st Asilomar Conference on Signals, Systems and Computers*, Nov. 2007.
- [13] D. P. Palomar, and Y. Jiang, "MIMO transceiver design via majorization theory," *Foundations and Trends in Communications and Information Theory*, Now Publishers, vol. 3, no. 4-5, pp. 331-551, 2007.