DESIGN OF OPTIMAL ORTHOGONAL LINEAR CODES IN MIMO SYSTEMS FOR MMSE RECEIVER

Jing Liu, Jian-Kang Zhang and K. Max Wong

Department of Electrical and Computer Engineering
McMaster University , Hamilton, ON, Canada

ABSTRACT

In this paper, we present a design of linear codes which improves the Bit Error Rate (BER) performance for multiple-input multiple-output (MIMO) communication system under a flat-fading environment in which the channel information is unknown for the transmitter but known at the receiver. Recently, space-time codes have been designed to achieve capacity and/or minimize the overall mean square error. Our codes here simultaneously achieve the lower bound of BER, the upper bound of the overall signal to interference plus noise ratio (SINR) and the lower bound of *substream* minimum mean square error. Our code structure is orthogonal in both overall and individual senses.

1. INTRODUCTION

MIMO wireless systems are important due to their potential for high data rate and/or diversity. Many existing space time codes using large number of antennas suffer from being impractical either because of their complexity or because of their inferior performance at high data rates. One approach which attempts to achieve high data rate while having reasonable complexity is VBLAST [1]. However, this is done at the expense of diversity gain. The design of linear space time block codes (STBC) by maximizing the mutual information based on Maximum Likelihood (ML) detection has been studied numerically [2] and theoretically [3], however, there was no guarantee for full diversity. Another important scheme is the Orthogonal STBC which ensures full diversity but suffers from low transmission data rate[4]. A full diversity full rate coding structure is presented in [5][6][7].

This paper considers a MIMO system which uses a linear detector for its simplicity. The goal here is to design codes that satisfy certain optimum criteria while maintaining the same transmission data rate as VBLAST. To ensure that the transmission data rate is maintained, we constrain our code to be a linear combination of the data. We then obtain the optimal design of the codes based on our criteria which are developed in two stages. First, we maximize a special SINR resulting in codes having an overall orthogonal structure that ensures high diversity gain. Second, we minimize the mean-square error (MSE) of each *substream* resulting in the coding matrix having an individual orthogonal structure. By analyzing its performance, we show that the codes so designed are optimum in that they achieve the lower bound of the average BER. Simulation results show that our design outperforms VBLAST and OSTBC in BER while maintaining the same transmission data rate.

2. MIMO FLAT-FADING SYSTEMS WITH LINEAR CODING

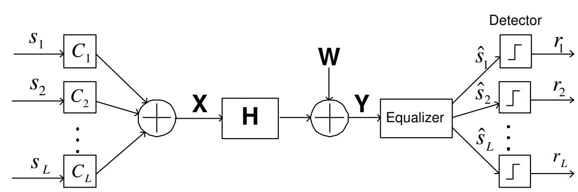

Fig.1 Original System Model

Consider the MIMO communication system shown in Fig.1 in which there are M transmitter antennas and N ($N \geq M$) receiver antennas. The channel is assumed to be flat-fading having zero-mean normalized iid random Gaussian coefficients h_{ij}, $i = 1, \cdots, M; j = 1, \cdots, N$ which remain constant for T ($M = T$) time slots. There are, in total, L ($L = MT$) data symbols s_i, $i = 1, \cdots, L$, randomly selected from a given constellation with zero mean and unity covariance (e.g. M-ary PSK). Each symbol is processed by an $M \times T$ coding matrix $\mathbf{C}_i$, $i = 1, \cdots, L$, and is then transmitted during the T time slots in the form: $\mathbf{X} = \sum_{i=1}^{L} s_i \mathbf{C}_i$, where the $M \times T$ matrix $\mathbf{X}$ is the combined coded signal. The total power assigned to all the coding matrices is a constant c, such that $\sum_{i=1}^{L} \text{tr}\, \mathbf{C}_i^H \mathbf{C}_i = c$, where "tr" denotes trace.

Let the $N \times T$ matrix $\mathbf{W}$ be the white Gaussian noise of distribution $\mathcal{CN}(0, 1)$ at the receiver antennas. Then the received signal $\mathbf{Y}$ with dimension $N \times T$ can be expressed in the following matrix form,

$$\mathbf{Y} = \sqrt{\frac{\rho}{M}} \mathbf{H}\mathbf{X} + \mathbf{W} \tag{1}$$

where $\mathbf{H}$ is the $N \times M$ channel matrix with elements h_{ij}, and ρ is the signal to noise ratio per receiver antenna.

By vectorizing both sides of (1) and defining two new matrices: $\mathbf{F} = [\text{vec}(\mathbf{C}_1), \cdots, \text{vec}(\mathbf{C}_L)]$, $\mathcal{H} = (\mathbf{I}_T \otimes \mathbf{H})\mathbf{F}$, where $\mathbf{I}$ is the identity matrix and T stands for the dimension, (1) can be written in the following equivalent form,

$$\mathbf{y} = \sqrt{\frac{\rho}{M}}\mathcal{H}\mathbf{s} + \mathbf{w} \tag{2}$$

where $\mathbf{y} = \text{vec}(\mathbf{Y})$, $\mathbf{s} = [s_1, s_2, \cdots, s_L]$, and $\mathbf{w} = \text{vec}(\mathbf{W})$.

3. MMSE EQUALIZER: PERFORMANCE BOUNDS

For the equivalent system model (2), we employ a Minimum MSE equalizer $\mathbf{G} = \sqrt{\frac{M}{\rho}}(\frac{M}{\rho}\mathbf{I} + \mathcal{H}^H\mathcal{H})^{-1}\mathcal{H}^H$ to facilitate threshold detection.

3.1. Coding Structure by Maximizing a Special SINR

Consider the SINR at the receiver after the equalizer of the MIMO system in (2). The received signal is given by $\hat{\mathbf{s}} = \sqrt{\frac{\rho}{M}}[\text{diag}(\mathbf{G}\mathcal{H})\mathbf{s}]$ whereas the interference and noise are given by $\mathbf{u} = \sqrt{\frac{\rho}{M}}[\mathbf{G}\mathcal{H} - \text{diag}(\mathbf{G}\mathcal{H})\mathbf{s}]$ and $\nu = \mathbf{G}\mathbf{w}$ respectively, where diag$(\cdot)$ denotes a diagonal matrix formed with the diagonal elements of a square matrix. Therefore, the SINR is given by

$$\text{SINR}_\text{T} = \frac{\text{tr } \mathbf{R}_{\hat{\mathbf{s}}}}{\text{tr } (\mathbf{R}_\mathbf{u} + \mathbf{R}_\nu)} \tag{3}$$

where $\mathbf{R}_{\hat{\mathbf{s}}}$, $\mathbf{R}_\mathbf{u}$ and $\mathbf{R}_\nu$ are the covariance matrices of the received signal $\hat{\mathbf{s}}$, the received interference, and the received noise respectively. However, to arrive at an optimum code design which is tractable, we rewrite the equalized signal as $\mathbf{G}\mathbf{y} = \mathbf{s} + \sqrt{\frac{\rho}{M}}(\mathbf{G}\mathcal{H} - \mathbf{I})\mathbf{s} + \mathbf{G}\mathbf{w}$ so that only the perfectly reconstructed symbols $\mathbf{s}$ is considered as the received signal and the rest of the distortion on the symbols is considered as interference. Thus, we arrive at an alternative definition of the SINR:

$$\text{SINR}_\text{A} = \frac{\text{tr } \mathbf{R}_\mathbf{s}}{\text{tr } \mathbf{R}_\text{int+n}} \tag{4}$$

where $\mathbf{R}_\mathbf{s}$ is the covariance matrix of $\mathbf{s}$, being equal to $\mathbf{I}_L$ in this case, and the covariance matrix for interference plus noise $\mathbf{R}_\text{int+n}$ is given by

$$\mathbf{R}_\text{int+n} = \text{E}\left\{(\mathbf{I} + \frac{\rho}{M}\mathcal{H}^H\mathcal{H})^{-1}\right\} \tag{5}$$

To arrive at an optimum code, we seek to maximize the SINR_A in (4). Since tr $\mathbf{R}_\mathbf{s}$ is a constant, we can equivalently minimize tr $\mathbf{R}_\text{int+n}$. Defining $\mathbf{A} \triangleq \mathbf{F}\mathbf{F}^H$, the power of interference plus noise in (5) can be written as

$$\text{tr } \mathbf{R}_\text{int+n} = \text{E}\left\{\text{tr } [\mathbf{I} + \frac{\rho}{M}\mathbf{A}^{1/2}(\mathbf{I} \otimes \mathbf{H}^H\mathbf{H})\mathbf{A}^{1/2}]^{-1}\right\} \tag{6}$$

Lemma 1 *Given any nonsingular Hermitian symmetric Positive Semi-Definite (PSD) matrix $\mathbf{Z}$ in block form: $\mathbf{Z} = \begin{bmatrix} \mathbf{D} & \mathbf{B} \\ \mathbf{B}^H & \mathbf{C} \end{bmatrix}$ with $\mathbf{D}$ and $\mathbf{C}$ being nonsingular Hermitian symmetric PSD, we have*

$$tr\, \mathbf{Z}^{-1} \geq tr\, \mathbf{D}^{-1} + tr\, \mathbf{C}^{-1}$$

where equality holds iff $\mathbf{B} = \mathbf{0}$, i.e., $\mathbf{Z}$ is block diagonal.

Let $M \times M$ matrix $\mathbf{A}_i$, $i = 1, \cdots, T$ denote the submatrix on the diagonal of $\mathbf{A}$, then according to **Lemma 1**, (6) is lower-bounded by

$$\text{tr } \mathbf{R}_\text{int+n} \geq \sum_{i=1}^{T} \text{E}\left\{\text{tr } [\mathbf{I} + \frac{\rho}{M}\mathbf{A}_i^{1/2}\mathbf{H}^H\mathbf{H}\mathbf{A}_i^{1/2}]^{-1}\right\} \tag{7}$$

Applying eigen-decomposition (ED) to each submatrix such that $\mathbf{A}_i = \mathbf{U}_i\mathbf{V}_i\mathbf{U}_i^H$ with $\mathbf{U}_i$ unitary and $\mathbf{V}_i$ nonnegative diagonal, (7) becomes,

$$\text{tr } \mathbf{R}_\text{int+n} \geq \sum_{i=1}^{T} \text{E}\left\{\text{tr } [\mathbf{I} + \frac{\rho}{M}\mathbf{H}^H\mathbf{H}\mathbf{V}_i]^{-1}\right\} \tag{8}$$

where we have used the invariance of the stochastic property of $\mathbf{H}$ when right-multiplied by a unitary matrix [8].

Now let's consider the following optimization problem,

$$\begin{aligned} \min_{\mathbf{V}_i} \quad &: \quad \sum_{i=1}^{T} \text{E}\left\{\text{tr } [\mathbf{I} + \frac{\rho}{M}\mathbf{H}^H\mathbf{H}\mathbf{V}_i]^{-1}\right\} \\ \text{s.t.} \quad &: \quad \sum_i \text{tr } \mathbf{V}_i = c. \end{aligned} \tag{9}$$

where the power constraint is equivalent to the original one since $\sum$ tr $\mathbf{V}_i$ = tr $\mathbf{A}$ = $\sum$ tr $\mathbf{C}_i^H\mathbf{C}_i = c$. We introduce the permutation matrices $\mathbf{P}_p$, $p = 1, \cdots, M!$, and modify the second term in the bracket of the cost function to $\frac{\rho}{M}\mathbf{P}_p^H\mathbf{H}^H\mathbf{H}\mathbf{P}_p\mathbf{V}_i$ without changing the value of cost function. Now, we perform Cholesky decomposition on $\mathbf{H}^H\mathbf{H} = \mathbf{\Phi}^H\mathbf{\Phi}$ where $\mathbf{\Phi}$ is $M \times M$ upper triangular matrix. Thus, the cost function in (9) becomes

$$\sum_{i=1}^{T} \text{E}\left\{\text{tr } [\mathbf{I} + \frac{\rho}{M}\mathbf{\Phi}\mathbf{P}_p\mathbf{V}_i\mathbf{P}_p^H\mathbf{\Phi}^H]^{-1}\right\} \tag{10}$$

The rearrangement of the diagonal elements of $\mathbf{V}_i$ by $\mathbf{P}_p$ does not change the value of cost function. If we find the average of all the expected values in (9) over all the possible ways of permuting (i.e. $M!$), the value will remain unchanged. However, since (10) is a convex function over the matrix inside the inverse computation, by Jensen's inequality, the quantity inside the first sum in (10) will be lower-bounded by the value when $\mathbf{V}_i = \frac{1}{M!}\sum_{p=1}^{M!}\mathbf{P}_p\mathbf{V}_i\mathbf{P}_p^H = \frac{\text{tr } \mathbf{V}_i}{M}\mathbf{I}$, i.e., the lower bound for (10) is

$$\sum_{i=1}^{T} \text{E}\left\{\text{tr } [\mathbf{I} + \frac{\rho}{M^2}(\text{tr } \mathbf{V}_i)\mathbf{H}^H\mathbf{H}]^{-1}\right\} \tag{11}$$

Since (11) is still a convex function, we can use Jensen's inequality again to obtain the lower bound for (11) which can be achieved when tr $\mathbf{V}_i = \frac{c}{T}$, i.e.,

$$T\mathrm{E}\left\{\text{tr}\,[\mathbf{I} + \frac{\rho\, c}{M^2 T}\mathbf{H}^H\mathbf{H}]^{-1}\right\} \tag{12}$$

The type of problem as shown in (12) has been solved [9] and the quantity in (12) is given by

$$\frac{T}{M}\int_0^\infty \frac{1}{1+\frac{c\rho}{M^2 T}\lambda}\sum_{i=1}^{M}\varphi_i(\lambda)^2\lambda^{N-M}e^{-\lambda}d\lambda \tag{13}$$

where, $\varphi(\lambda) = \left[\frac{k!}{(k+N-M)!}\right]^{1/2}\mathcal{L}_k^{N-M}(\lambda),\ k = 0, 1, \cdots,$ $M-1$, and $\mathcal{L}_k^{N\,-\,M}(x) = \frac{1}{k!}e^x x^{M-N}\frac{d^k}{dx^k}(e^{-x}x^{N-M+k})$ is the associated Laguerre polynomial of order k.

In summary, the optimal solution to (9) is $\mathbf{V}_{i\text{opt}} = \frac{c}{MT}\mathbf{I}_M$, $i = 1, \cdots, L$. (13) yields the optimal value (i.e., a reachable constant lower bound) of cost function in (9), which is a further lower bound of tr $\mathbf{R}_{\text{int+n}}$ in (8). Thus, when equality in (8) holds, i.e., $\mathbf{A}$ is block diagonal, and $\mathbf{V}_i = \frac{c}{MT}\mathbf{I}_M$, (13) becomes the minimum value of tr $\mathbf{R}_{\text{int+n}}$. Recall that $\mathbf{V}_i$ is the eigen-value matrix of $\mathbf{A}_i$ which is the ith block sub-matrix forming the diagonal of $\mathbf{A} = \mathbf{F}\mathbf{F}^H$. Therefore, $\mathbf{A}_{\text{opt}} = \frac{c}{MT}\mathbf{I}_L$, and $\mathbf{F}$ is a unitary matrix multiplied by a constant governed by the power constraint. By definition, $\mathbf{F} = [\text{vec}(\mathbf{C}_1), \cdots, \text{vec}(\mathbf{C}_L)]$. Hence, the overall structure for the optimal codes is

$$\text{tr}\,\mathbf{C}_i^H\mathbf{C}_j = \frac{c}{L}\,\delta(i-j), \qquad i, j = 1, \cdots, L \tag{14}$$

where $\delta(x) = \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{else} \end{cases}$.

3.2. Optimal Coding Matrix Structure by Substream Performance

Eq.(14) tells us that to maximize the SINR_A, each $\mathbf{C}_i$ should be allocated equal power. However, the inside structure of each $\mathbf{C}_i$ is still undetermined. We will now explore this by considering the MSE performance of the substreams.

The MSE matrix of the system (2) employing the coding structure in (14) is [10]

$$\mathbf{E}_{\text{rr}} = \frac{L}{c}\mathrm{E}\{\mathbf{F}^H[\mathbf{I} + \frac{c\rho}{ML}\mathbf{I}\otimes\mathbf{H}^H\mathbf{H}]^{-1}\mathbf{F}\} \tag{15}$$

The MSE of ith substream e_i is the iith element of $\mathbf{E}_{\text{rr}}$,

$$e_i = \frac{L}{c}\mathrm{E}\{\text{tr}\,[\mathbf{C}_i^H(\mathbf{I} + \frac{c\rho}{ML}\mathbf{H}^H\mathbf{H})^{-1}\mathbf{C}_i]\} \tag{16}$$

Here, we want to minimize the largest MSE among all the substreams. We note that,

$$\max\{e_i\} \geq \frac{\sum_{i=1}^{L} e_i}{L} = \frac{\text{tr}\,\mathbf{E}_{\text{rr}}}{L} \tag{17}$$

By substituting (15) into (17), the right hand side of (17) can be written as,

$$\frac{T\mathrm{E}\{\sum_{i=1}^{M}\frac{1}{1+\frac{c\rho}{ML}\lambda_i}\}}{L} \tag{18}$$

where λ_i is the eigen-value of $\mathbf{H}^H\mathbf{H}$. Through similar steps as those for (13), we can obtain a close form expression for (18) which is a constant. Thus, the optimal value is achieved when equality in (17) holds, i.e. $e_i = e_j$ for all i and j. Now, from the expression of e_i in (16) together with the consideration of the code structure in (14), to ensure that all e_i are equal for all i, we see that one optimal solution is

$$\mathbf{C}_i^H\mathbf{C}_i = \frac{c}{ML}\,\mathbf{I}_M, \qquad i = 1, \cdots, L \tag{19}$$

3.3. BER Performance of the Optimal Codes

We now examine the BER performance of our codes when $\mathbf{s}$ is BPSK. For the ith symbol of $\mathbf{s}$ in the equivalent model of (2), if there is a large number of interference sources, the interference plus noise for ith symbol can be treated as Gaussian (*Central Limit Theorem*). Then, the error probability for ith symbol is

$$\mathcal{P}_{ei} = Q\left(\sqrt{2\text{SINR}_{\text{T}i}}\right)$$

where $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-z^2}dz$ and SINR_i is the SINR_T associated with the ith received symbol of $\mathbf{s}$ that is given by

$$\text{SINR}_i = \frac{P_{\text{sig}}}{P_{\text{all}} - P_{\text{sig}}} = \frac{1}{\frac{P_{\text{all}}}{P_{\text{sig}}} - 1}$$

where P_{sig} and P_{all} are the power of the ith received signal $\hat{s}_i$ and the total power of signal, interference and noise at ith receiver. From [10],

$$P_{\text{sig}} = \{[\sqrt{\frac{\rho}{M}}\mathbf{G}\mathcal{H}]_{ii}\}^2$$

Let $a \triangleq [\sqrt{\frac{\rho}{M}}\mathbf{G}\mathcal{H}]_{ii}$, then

$$a = \frac{L}{c}\text{tr}\,\{\mathbf{C}_i^H\boldsymbol{\Upsilon}\mathbf{C}_i\} \tag{20}$$

where $\boldsymbol{\Upsilon} = (\mathbf{I} + \frac{c\rho}{ML}\mathbf{H}^H\mathbf{H})^{-1}(\frac{c\rho}{ML}\mathbf{H}^H\mathbf{H})$ is a symmetric PSD matrix. At the same time, in system model (2),

$$P_{\text{all}} = [\frac{\rho}{M}\mathbf{G}\mathcal{H}\mathcal{H}^H\mathbf{G}^H + \mathbf{G}\mathbf{G}^H]_{ii}$$

After some simplifications, we have

$$\text{SINR}_i = \frac{1}{\frac{1}{a} - 1}$$

The average BER of the detected signal is thus,

$$\mathcal{P}_e = \frac{1}{L}\sum_{i=1}^{L} Q\left(\sqrt{\frac{2}{\frac{1}{\frac{L}{c}\mathrm{tr}\,\{\mathbf{C}_i^H \boldsymbol{\Upsilon} \mathbf{C}_i\}} - 1}}\right) \qquad (21)$$

Now, $Q(\sqrt{\frac{2}{x}})$ is a convex function of x when $x < \frac{4}{3}$, i.e., when $a > \frac{3}{7}$. If this condition is satisfied, we can find the lower bound for $\mathcal{P}_e$ by Jensen's inequality such that

$$\mathcal{P}_e \geq Q\left(\sqrt{\frac{2}{\frac{1}{L}\sum_{i=1}^{L}\frac{1}{\frac{L}{c}\mathrm{tr}\,\{\mathbf{C}_i^H \boldsymbol{\Upsilon} \mathbf{C}_i\}} - 1\}}}\right) \qquad (22)$$

with the equality reached when all the terms under summation are equal for all i. This can be achieved if $\mathbf{C}_i$ has the unitary structure as in (19). To obtain the condition for which convexity of (21) is satisfied, we substitute the values of a from (20). Using the usual manipulations of ED and simplifying, we obtain

$$\frac{c\rho}{ML}\lambda > \frac{3}{4} \qquad (23)$$

(23) is a sufficient condition for (21)to be convex. It implies that when the total coding power c or SNR per receiver antenna ρ is high or the channel attenuation is low, such that the inequality in (23) is satisfied, our codes achieve the lower bound of BER.

4. SIMULATIONS

In this section, we provide simulation results comparing the BER performance between VBLAST [2], Orthogonal design [4] and our optimal codes under the same transmission data rate $R = 6$ bits per channel use where $M = 3$ transmitter antennas and $N = 4$ receiver antennas are used. Our flat-fading channel is generated by selecting from normalized Gaussian random numbers, and the transmitted signal is of Q-PSK. Fig.2 shows the BER performance under different SNR. It can be observed that the optimal code designed from our algorithm is superior in performance to both V-BLAST and the Orthogonal code in [4].

5. CONCLUSION

In this paper, we present a new approach in designing orthogonal linear codes for a flat-fading MIMO system in which an MMSE detector is employed. The overall coding structure is obtained by maximizing the alternative SINR and the specific structure for the individual coding matrices is obtained by considering substream performance. It has also been shown that this code structure achieves the lower bound of the BER at high SNR.

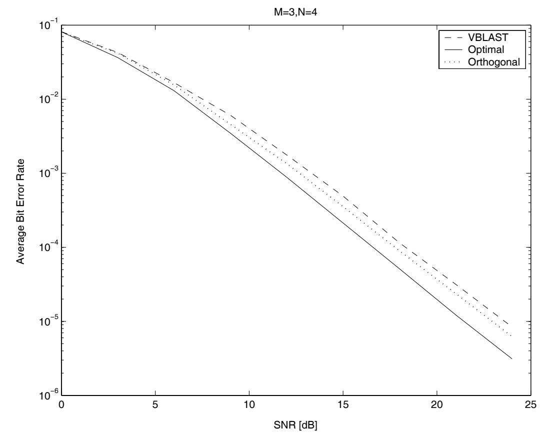

Fig.2 BER performance comparison.

6. REFERENCES

[1] G.D. Golden, G.J. Foschini, R.A. Valenzuala, and P.W. Wolnaiansky, "Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture," *Electron. Lett.*, vol. 17, Nov. 1998.

[2] B. Hassibi and B.M.Hochwald, "High-Rate codes that are linear in space and time," *IEEE Trans. Inf. Th.*, vol. 48, pp. 1804–1824, July 2002.

[3] A.R. Ghaderipoor, L. Beygi, and S.H. Jamali, "Analytical survey on the design of linear dispersion space-time codes," *Acous., Speech, and Sig. Proc.*, vol. 3, pp. 2433–2436, May 2002.

[4] E. G. Larsson and P. Stoica, *Space-time Block Coding for wireless communications*, Cambridge University Press, 2003.

[5] Xiaoli Ma and G. Giannakis, "Complex field coded MIMO systems: performance, rate, and trade-offs," *Wireless communications and mobile computing*, vol. 2, pp. 693–717, 2002.

[6] Xiaoli Ma and G. B. Giannakis, "Full-diversity full rate complex-field space-time coding," *IEEE Trans. Signal Processing*, vol. 51, pp. 2917–2930, Nov. 2003.

[7] H. El. Gamal and M. O. Damen, "Universal space-time coding," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1097–1119, May 2003.

[8] R. J. Muirhead, *Aspects of multivariate statistical theory*, John Wiley and Sons, Inc., 1982.

[9] İ. E. Telatar, "Capacity of Multi-antanna Gaussian Channels," *Euro. Trans. on Telecomm.*, vol. 10, No. 6, pp. 585–59, Nov/Dec 1999.

[10] D. P.érez 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.*, vol. 51, No.9, pp. 2381–2401, Sept. 2003.