A Robust Linear Receiver for Multi-Access Space-Time Block-Coded MIMO Systems Based on Probability-Constrained Optimization

Yue Rong Sergiy A. Vorobyov
Dept. of Communication Systems
University of Duisburg-Essen
Bismarckstr. 81, 47057 Duisburg, Germany
E-mail: rong@sent5.uni-duisburg.de

Alex B. Gershman
Dept. of ECE, McMaster University
1280 Main St. West, Hamilton
L8S 4K1 Ontario, Canada
Email: gershman@ieee.org

Abstract— **Traditional multiuser receiver algorithms developed for multiple-input multiple-output (MIMO) wireless communication systems are based on the assumption that the channel state information (CSI) is precisely known at the receiver. However, in practical situations the exact CSI is never available. In this paper, we address the problem of robustness of multi-access space-time block-coded MIMO systems against imperfect CSI. We propose a new linear receiver which guarantees the robustness against CSI errors with a certain selected probability. The proposed receiver has a form of a probability-constrained optimization problem, which can be simplified to a convex nonlinear programming (NLP) problem based on the observation that the CSI mismatch is Gaussian. The later problem can be efficiently solved using standard optimization methods. Numerical simulations demonstrate the robustness of the proposed receiver against CSI errors.**

I. INTRODUCTION

In uplink cellular communications with multiple receiver antennas at the base station (BS), spatial diversity techniques can be employed to enhance the system capacity. Additionally, if mobile users also have multiple antennas, space-time block codes (STBCs) can be used to improve the immunity to fading [1], [2]. Application of space-time block-coded MIMO wireless systems in the multi-access scenario gained recently a significant interest [3], [4].

Both linear and nonlinear receiver algorithms for multi-access MIMO systems have been recently proposed in the literature. Although linear techniques such as zero-forcing (ZF) and minimum mean-square-error (MMSE) receivers are suboptimal, they recently gained much interest due to their low computational complexity. Unfortunately, both the ZF and MMSE receivers require perfect knowledge of the CSI. However, the exact CSI is never available, because of channel estimation errors or outdated training. As a consequence, the performance of the ZF and MMSE receivers may degrade severely. Therefore, linear receivers robust against *imperfect* CSI are of interest. In this paper, we design a linear receiver for multi-access space-time block-coded MIMO systems that is robust against CSI errors.

Recently, several works have addressed the problem of robust linear receiver design. For example, the diagonal loading (DL) minimum variance approach is used in [4] to provide robustness against both the CSI and data covariance matrix mismatches. However, the selection of the DL factor in [4] is *ad hoc*. Moreover, the performance of such receiver still significantly depends on how accurate the CSI is. The latter problem is addressed in [5] by optimizing the worst-case performance of the receiver. However, this strategy may be too pessimistic and, therefore, may lead to unnecessary performance degradation. In this paper, we design a linear receiver which guarantees the robustness against CSI errors with a certain selected probability. The mathematical formulation of the receiver design problem is equivalent to the probability-constrained stochastic optimization problem [6], [7]. The solution to this problem is obtained by converting it into the corresponding deterministic nonlinear programming (NLP) problem [8], which can be efficiently solved using standard optimization software tools.

II. MULTI-ACCESS MIMO LINEAR RECEIVERS

In this section, we review the models for point-to-point and multi-access space-time block-coded MIMO wireless systems. The latter one is used for the problem formulation of multi-access MIMO linear receiver design.

A. Point-to-Point Space-Time Block-Coded MIMO System

The point-to-point MIMO model can be written as [9]

$$\mathbf{Y} = \mathbf{X}\mathbf{H} + \mathbf{N} \tag{1}$$

where $\mathbf{Y}$ is the $T \times M$ complex matrix of the received data, $\mathbf{X}$ is the $T \times N$ complex matrix of the transmitted data, $\mathbf{H}$ is the $N \times M$ complex matrix of quasi-static Rayleigh flat-fading channel whose coherence time is assumed to be longer than T, $\mathbf{N}$ is the $T \times M$ complex additive white Gaussian noise (AWGN) at the receiver, N is the number of transmit antennas, M is the number of receive antennas, and T is the data block length.

0-7803-8255-2/04/$20.00 ©2004 IEEE.

If the user data $s_1, \ldots, s_K$ are encoded by some STBC, then the matrix $\mathbf{X}$ has the following structure [9], [10]

$$\mathbf{X} = \sum_{k=1}^{K}(\mathbf{C}_k \mathrm{Re}\{s_k\} + \mathbf{D}_k \mathrm{Im}\{s_k\}) \quad (2)$$

where

$$\mathbf{C}_k = \mathbf{X}(\mathbf{e}_k), \quad \mathbf{D}_k = \mathbf{X}(j\mathbf{e}_k)$$

$j = \sqrt{-1}$, and $\mathbf{e}_k$ is the $K \times 1$ vector having one in the k-th position and zeros elsewhere. Inserting (2) into (1) yields the following model [4], [9], [10]

$$\underline{\mathbf{Y}} = \mathbf{A}\underline{\mathbf{s}} + \underline{\mathbf{N}} \quad (3)$$

where

$$\mathbf{A} = [\underline{\mathbf{C}_1\mathbf{H}}, \ldots, \underline{\mathbf{C}_K\mathbf{H}}, \underline{\mathbf{D}_1\mathbf{H}}, \ldots, \underline{\mathbf{D}_K\mathbf{H}}]$$

and the "underline" operator is defined as

$$\underline{\mathbf{Y}} = \begin{bmatrix} \mathrm{vec}(\mathrm{Re}\{\mathbf{Y}\}) \\ \mathrm{vec}(\mathrm{Im}\{\mathbf{Y}\}) \end{bmatrix} \quad (4)$$

where $\underline{\mathbf{Y}}$ is the $2MT \times 1$ real vector.

B. Multi-Access Space-Time Block-Coded MIMO System

In the uplink multi-access case, the model (3) can be extended as [4]

$$\underline{\mathbf{Y}} = \sum_{f=1}^{F} \mathbf{A}_f \underline{\mathbf{s}_f} + \underline{\mathbf{N}} \quad (5)$$

where

$$\mathbf{A}_f = [\mathbf{a}_{f,1}, \ldots, \mathbf{a}_{f,2K}], \qquad \mathbf{a}_{f,k} = \underline{\mathbf{F}_k \mathbf{H}_f}$$

$$\mathbf{F}_k = \begin{cases} \mathbf{C}_k & k = 1, \ldots, K \\ \mathbf{D}_{k-K} & k = K+1, \ldots, 2K \end{cases}$$

and $\mathbf{H}_f$ is the channel matrix between f-th user and the BS. Here, F is the number of users.

C. Multi-Access MIMO Linear Receivers

Without loss of generality, we can assume that the first user is the desired one. The data vector $\underline{\hat{\mathbf{s}}_1}$ decoded by a linear receiver has the following form [4]

$$\underline{\hat{\mathbf{s}}_1} = \mathbf{W}^T \underline{\mathbf{Y}} \quad (6)$$

where $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_{2K}]$ is the $2MT \times 2K$ real matrix of weight coefficients. Consequently, the k-th symbol of the desired user is estimated as

$$\begin{aligned} [\hat{\mathbf{s}}_1]_k &= \mathrm{Re}\{[\hat{\mathbf{s}}_1]_k\} + j\mathrm{Im}\{[\hat{\mathbf{s}}_1]_k\} \\ &= [\underline{\hat{\mathbf{s}}_1}]_k + j[\underline{\hat{\mathbf{s}}_1}]_{k+K} \\ &= \mathbf{w}_k^T \underline{\mathbf{Y}} + j\mathbf{w}_{k+K}^T \underline{\mathbf{Y}}. \end{aligned} \quad (7)$$

The problem now is to find such matrix $\mathbf{W}$ which separates the signals from different users. The well-known solution for this problem is given by the ZF receiver [11], provided that exact CSI is available at BS. The receiver coefficient vector can be then written as

$$\mathbf{W}^{\mathrm{ZF}} = \mathbf{M}^{\dagger} \quad (8)$$

where the $2MT \times 2KF$ real matrix $\mathbf{M} = [\mathbf{A}_1, \ldots, \mathbf{A}_F]$ and $(\cdot)^{\dagger}$ denotes pseudo-inverse. Unfortunately, the ZF receiver is quite sensitive to CSI errors.

Another popular linear receiver is the MMSE receiver. Its coefficient vector is given by [11]

$$\mathbf{w}_k^{\mathrm{MMSE}} = \mathbf{R}^{-1}\mathrm{E}\{\underline{\mathbf{Y}} \cdot [\underline{\mathbf{s}_1}]_k\} \quad (9)$$

where $\mathrm{E}\{\cdot\}$ is the statistical expectation, $\mathrm{E}\{\underline{\mathbf{Y}} \cdot [\underline{\mathbf{s}_1}]_k\}$ is the cross-correlation vector between the information-bearing symbol $[\underline{\mathbf{s}_1}]_k$ and the received vector $\underline{\mathbf{Y}}$, and

$$\mathbf{R} = \mathrm{E}\{\underline{\mathbf{Y}}\,\underline{\mathbf{Y}}^T\}$$

is the covariance matrix of received real (vectorized) data.

Although the MMSE receiver (9) does not use any CSI knowledge in an explicit form, it requires the knowledge of second-order statistics (SOSs) of signals, which may not be available at the BS and, therefore, must be estimated using sample data. As a result, the performance of the MMSE receiver highly depends on the accuracy of the estimates of $\mathbf{R}$ and $\mathrm{E}\{\underline{\mathbf{Y}} \cdot [\underline{\mathbf{s}_1}]_k\}$. In order to obtain accurate estimates of the required SOSs, a large number of samples has to be available. Therefore, the performance of the MMSE receiver using the sample data may degrade substantially due to inaccurate estimates of SOSs.

More accurate estimates of the required SOSs can be obtained if the knowledge of approximate CSI, transmit powers of all users, and noise power is available at the BS. Then the estimates are given by

$$\hat{\mathbf{R}} = \sum_{f=1}^{F} \sigma_f^2 \hat{\mathbf{A}}_f \hat{\mathbf{A}}_f^T + \sigma_n^2 \mathbf{I} \quad (10)$$

$$\mathrm{E}\{\underline{\mathbf{Y}} \cdot [\underline{\mathbf{s}_1}]_k\} = \hat{\mathbf{a}}_{1,k} \quad (11)$$

where $\hat{\mathbf{A}}_f$ and $\hat{\mathbf{a}}_{1,k}$ denote the estimated values of $\mathbf{A}_f$ and $\mathbf{a}_{1,k}$, respectively, and σ_f^2 and σ_n^2 are the signal power of the fth user and the noise power, respectively. Note that σ_f^2 and σ_n^2 are assumed to be known. As the ZF receiver, the MMSE receiver is quite sensitive to CSI errors.

III. Design of Robust Linear Receiver via Stochastic Programming

A. Problem Formulation

Let us consider the error matrix $\mathbf{E}_f = \mathbf{H}_f - \hat{\mathbf{H}}_f$ between the true channel matrix $\mathbf{H}_f$ of the fth user and its estimated value $\hat{\mathbf{H}}_f$. Consequently, using the notations of model (5) we can write that

$$\mathbf{e}_{f,q} = \mathbf{a}_{f,q} - \hat{\mathbf{a}}_{f,q} = \underline{\mathbf{F}_q\mathbf{H}_f} - \underline{\mathbf{F}_q\hat{\mathbf{H}}_f} = \underline{\mathbf{F}_q\mathbf{E}_f} \quad (12)$$

where $\mathbf{e}_{f,q}$ is the random mismatch vector between the true $\mathbf{a}_{f,q}$ and its estimate $\hat{\mathbf{a}}_{f,q}$. Note, that the last equality in (12) follows from the linearity of underline operator (4).

The linear receiver design problem is to estimate each entry of $\underline{\mathbf{s}_1}$ by minimizing the noise and *total* interference power while keeping the distortionless response for this entry of $\underline{\mathbf{s}_1}$.

The total interference power consists of the power from *multi-access interference*, which is caused by other users, and *self-interference*, which is caused by other entries of $\underline{s_1}$. In order to provide the robustness of the linear receiver against CSI mismatches, it is necessary to take into account the mismatch vector $\mathbf{e}_{f,q}$ (12), i.e. to consider $\hat{\mathbf{a}}_{f,q} + \mathbf{e}_{f,q}$ instead of $\hat{\mathbf{a}}_{f,q}$. However, the suppression of the total interference power and keeping distortionless response for kth entry of $\underline{s_1}$ will be performed with a certain probability. This is different from the worst-case based robust design [5], where the performance is optimized for the worst-case mismatch and the norm of the mismatch is bounded by some known constant. Our new problem formulation for robust linear receiver design suggests to find the receiver coefficient vector $\mathbf{w}_k$ for the kth entry of $\underline{s_1}$ as the solution to the following optimization problem

$$\min_{\mathbf{w}_k,\mathbf{d}} \quad \|\mathbf{d}\| \tag{13}$$

$$\text{s.t.} \quad \mathrm{P}\{\mathbf{w}_k^T(\hat{\mathbf{a}}_{1,k} + \mathbf{e}_{1,k}) \geq 1\} \geq p \tag{14}$$

$$\mathrm{P}\{|\mathbf{w}_k^T(\hat{\mathbf{a}}_{1,j} + \mathbf{e}_{1,j})| \leq d_{1,j}\} \geq p \tag{15}$$

$$\mathrm{P}\{|\mathbf{w}_k^T(\hat{\mathbf{a}}_{f,q} + \mathbf{e}_{f,q})| \leq d_{f,q}\} \geq p \tag{16}$$

$$j = 1, \cdots, 2K, \quad j \neq k,$$
$$f = 2, \cdots, F, \quad q = 1, \cdots, 2K$$

where $\mathbf{d} = [d_{1,1}, \cdots, d_{F,2K}]^T$ is the $(2FK-1) \times 1$ vector whose values limit the contribution of multi-access interference and self-interference, p is a certain probability value which is selected according to the quality of service (QoS) requirements, $\|\cdot\|$ denotes for Frobenius (Euclidian) norm of a matrix (vector), and $\mathrm{P}\{\cdot\}$ stands for the the probability operator whose form is assumed to be known.

The probability bound p can be selected from the interval $[0, 1]$ and it determines an amount of mismatch that is allowed at the receiver. If $p = 1$ then the receiver does not assume any channel errors (this corresponds to the non-robust design case).

In the formulation (13)-(16), we minimize the total interference power with a certain probability while keeping the probability of the distortionless response to the desired entry of $\underline{s_1}$ larger than p (see the constraint (14)). The constraints (15) and (16) are formulated for self-interference and multi-access interference suppression, correspondingly. In (13)-(16), we do not consider explicitly the noise component due to the following reasons. First, for multi-access scenarios the interference suppression is more important than the noise suppression. Second, the effect of noise is implicitly taken into account by introducing CSI mismatch. We also stress that the problem (13)-(16) belongs to the class of stochastic programming problems [6], [7].

B. Convexity of the Problem (13)-(16)

We assume that the channel mismatch $\mathbf{E}_f$ is Gaussian. The justification of this assumption is that in MIMO wireless communication systems, if nothing is known *a priori* about the channel, the optimal training sequences are orthogonal. It can be proven that for orthogonal training sequences the CSI errors are Gaussian [12].

THEOREM 1: Let vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ have a joint real Gaussian distribution with a covariance matrix $\mathbf{B}$, so that

$$\mathrm{E}\{(\mathbf{v}_i - \mathrm{E}\{\mathbf{v}_i\})(\mathbf{v}_j - \mathrm{E}\{\mathbf{v}_j\})^T\} = r_{ij}\mathbf{B} \quad \forall\, i,j, \quad i,j = 1, \ldots, n \tag{17}$$

where r_{ij} are some constants. Then the set

$$\mathcal{K}(p) = \{\mathbf{x}\,|P\{\mathbf{v}_1^T\mathbf{x} \geq d_1 \ \wedge \ \ldots \ \wedge \ \mathbf{v}_n^T\mathbf{x} \geq d_n\} \geq p\} \tag{18}$$

is convex for $p \geq 0.5$. Here, $\wedge$ denotes the set intersection operation, $0 < p \leq 1$, and d_i are some constants.

PROOF: See [7, p. 312]. □

THEOREM 2: If $[\mathbf{E}_f]_{n,m} \sim \mathcal{CN}(0, \sigma_h^2)$ and $p \in (0.5, 1)$, then the optimization problem (13)-(16) is convex.

PROOF: First of all, the objective function (13) is a simple vector norm that is obviously convex.

Next, we show that the constraint (14) is also convex. Applying the underline operator (4) to (12) and using well-known properties of the Kronecker product (denoted hereafter as $\otimes$) and vec$(\cdot)$ operation [13], we can rewrite the mismatch vector $\mathbf{e}_{f,q}$ as

$$\mathbf{e}_{f,q} = \begin{bmatrix} \mathrm{vec}(\mathrm{Re}\{\mathbf{F}_q\mathbf{E}_f\}) \\ \mathrm{vec}(\mathrm{Im}\{\mathbf{F}_q\mathbf{E}_f\}) \end{bmatrix} \tag{19}$$
$$= \begin{bmatrix} \mathrm{Re}\{\mathbf{I}_M \otimes \mathbf{F}_q\} & -\mathrm{Im}\{\mathbf{I}_M \otimes \mathbf{F}_q\} \\ \mathrm{Im}\{\mathbf{I}_M \otimes \mathbf{F}_q\} & \mathrm{Re}\{\mathbf{I}_M \otimes \mathbf{F}_q\} \end{bmatrix} \begin{bmatrix} \mathrm{vec}(\mathrm{Re}\{\mathbf{E}_f\}) \\ \mathrm{vec}(\mathrm{Im}\{\mathbf{E}_f\}) \end{bmatrix}.$$

It follows from (19) that $\mathbf{e}_{f,q}$ is a linear combination of the real and imaginary parts of the elements of the channel mismatch matrix $\mathbf{E}_f$. Thus, $\mathbf{e}_{f,q}$ has multivariate real Gaussian distribution

$$\mathbf{e}_{f,q} \sim \mathcal{N}(\mathbf{0}_{2MT}, \frac{\sigma_h^2}{2}(\mathbf{I}_{2M} \otimes \mathbf{G}_q\mathbf{G}_q^T)) \tag{20}$$

where

$$\mathbf{G}_q = \begin{cases} \mathbf{C}_q & q = 1, \ldots, K \\ \mathrm{Im}\{\mathbf{D}_{q-K}\} & q = K+1, \ldots, 2K. \end{cases} \tag{21}$$

Since the only random variable in (14) is $\mathbf{e}_{1,k}$, and both $\mathbf{w}_k$ and $\hat{\mathbf{a}}_{1,k}$ are deterministic variables, the random variable $\mathbf{w}_k^T(\hat{\mathbf{a}}_{1,k} + \mathbf{e}_{1,k})$ has also Gaussian distribution

$$\mathbf{w}_k^T(\hat{\mathbf{a}}_{1,k} + \mathbf{e}_{1,k}) \sim \mathcal{N}(\mathbf{w}_k^T\hat{\mathbf{a}}_{1,k}, \frac{\sigma_h^2}{2}\|(\mathbf{I}_{2M} \otimes \mathbf{G}_k^T)\mathbf{w}_k\|^2). \tag{22}$$

Using the standard error function for the Gaussian distribution

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt \tag{23}$$

the left hand side of (14) can be written as

$$\mathrm{P}\{\mathbf{w}_k^T(\hat{\mathbf{a}}_{1,k} + \mathbf{e}_{1,k}) \geq 1\} = \frac{1}{2} - \frac{1}{2}\mathrm{erf}\left(\frac{1 - \mathbf{w}_k^T\hat{\mathbf{a}}_{1,k}}{\sigma_h\|(\mathbf{I}_{2M} \otimes \mathbf{G}_k^T)\mathbf{w}_k\|}\right). \tag{24}$$

Substituting (24) into (14), after some straightforward manipulations, we obtain the following constraint

$$\mathrm{erf}\left(\frac{\mathbf{w}_k^T\hat{\mathbf{a}}_{1,k} - 1}{\sigma_h\|(\mathbf{I}_{2M} \otimes \mathbf{G}_k^T)\mathbf{w}_k\|}\right) \geq 2p - 1 \tag{25}$$

that is convex if $p \in (0.5, 1)$.

The constraints (15) and (16) share the same structure. Thus, it is enough to show that the constraints (15) are convex with respect to the unknown variables $\mathbf{w}_k$ and $d_{1,j}$, $j = 1, \ldots, 2K$, $j \neq k$.

Let us rewrite the constraints (15) in the equivalent form

$$\mathrm{P}\{(\hat{\mathbf{a}}_{1,j}^T + \mathbf{e}_{1,j}^T)\mathbf{w}_k \geq -d_{1,j} \wedge (-\hat{\mathbf{a}}_{1,j}^T - \mathbf{e}_{1,j}^T)\mathbf{w}_k \geq -d_{1,j}\} \geq p \quad j = 1, \cdots, 2K, \quad j \neq k. \tag{26}$$

These constraints are called *joint chance constraints* in stochastic programming literature [6], [7]. Now it is sufficient to prove the convexity of only one constraint from $2K - 1$ such constraints. To be consistent with the notations used in Theorem 1, let us denote

$$\begin{aligned} \mathbf{v} &:= \hat{\mathbf{a}}_{1,j} + \mathbf{e}_{1,j} \\ \mathbf{x} &:= \mathbf{w}_k \\ d &:= -d_{1,j}\,. \end{aligned}$$

Then, the jth constraint can be equivalently written as

$$\mathrm{P}\{\mathbf{v}^T\mathbf{x} \geq d \ \wedge \ -\mathbf{v}^T\mathbf{x} \geq d\} \geq p\,. \tag{27}$$

As the vectors $\mathbf{v}$ and $-\mathbf{v}$ have joint Gaussian distribution with the common covariance matrix

$$\mathbf{B} = \mathrm{E}\{\mathbf{e}_{1,j}\mathbf{e}_{1,j}^T\} = \frac{\sigma_h^2}{2}(\mathbf{I}_{2M} \otimes \mathbf{G}_j\mathbf{G}_j^T) \tag{28}$$

we can see that Theorem 1 can be applied. Thus, the convexity of the constraints (15) and (16) is also proven if $p \in (0.5, 1)$.

Summarizing, the objective function of the problem (13)-(16) is convex, and the constraints are convex if $[\mathbf{E}_f]_{n,m} \sim \mathcal{CN}(0, \sigma_h^2)$ and $p \in (0.5, 1)$. This completes the proof. □

For simplicity, Theorem 2 has been proven for the case of i.i.d. zero-mean channel mismatch. However, it is straightforward to generalize it to the case of correlated non-zero mean channel mismatch because the expression (19) remains valid in such a case and $\mathbf{e}_{f,q}$, $f = 1, \ldots, F$, $q = 1, \ldots, 2K$ still have Gaussian distribution.

C. Nonlinear Programming Approach

Similarly to the constraint (14), the stochastic constraints (15) and (16) can be converted into their deterministic equivalents. Using (22), (23) and (26), the left hand side of (15) can be written as

$$\begin{aligned} &\mathrm{P}\{|\mathbf{w}_k^T(\hat{\mathbf{a}}_{1,j} + \mathbf{e}_{1,j})| \leq d_{1,j}\} \\ &= \mathrm{P}\{\mathbf{w}_k^T(\hat{\mathbf{a}}_{1,j} + \mathbf{e}_{1,j}) \leq d_{1,j}\} \\ &-\mathrm{P}\{\mathbf{w}_k^T(\hat{\mathbf{a}}_{1,j} + \mathbf{e}_{1,j}) \leq -d_{1,j}\} \\ &= \frac{1}{2}\mathrm{erf}\left(\frac{d_{1,j} - \mathbf{w}_k^T\hat{\mathbf{a}}_{1,j}}{\sigma_h\|(\mathbf{I}_{2M} \otimes \mathbf{G}_j^T)\mathbf{w}_k\|}\right) \\ &-\frac{1}{2}\mathrm{erf}\left(\frac{-d_{1,j} - \mathbf{w}_k^T\hat{\mathbf{a}}_{1,j}}{\sigma_h\|(\mathbf{I}_{2M} \otimes \mathbf{G}_j^T)\mathbf{w}_k\|}\right) \\ &j = 1, \cdots, 2K, \ j \neq k\,. \end{aligned} \tag{29}$$

Combining (13), (25) and (29) together, and converting the constraints (16) into their deterministic equivalents, we can rewrite the stochastic programming problem (13)-(16) as the following deterministic NLP problem

$$\min_{\mathbf{w}_k, \mathbf{d}} \quad \|\mathbf{d}\| \tag{30}$$

$$\text{s.t.} \quad \mathrm{erf}\left(\frac{\mathbf{w}_k^T\hat{\mathbf{a}}_{1,k} - 1}{\sigma_h\|(\mathbf{I}_{2M} \otimes \mathbf{G}_k^T)\mathbf{w}_k\|}\right) \geq 2p - 1 \tag{31}$$

$$\mathrm{erf}\left(\frac{d_{1,j} - \mathbf{w}_k^T\hat{\mathbf{a}}_{1,j}}{\sigma_h\|(\mathbf{I}_{2M} \otimes \mathbf{G}_j^T)\mathbf{w}_k\|}\right) - \mathrm{erf}\left(\frac{-d_{1,j} - \mathbf{w}_k^T\hat{\mathbf{a}}_{1,j}}{\sigma_h\|(\mathbf{I}_{2M} \otimes \mathbf{G}_j^T)\mathbf{w}_k\|}\right) \geq 2p \tag{32}$$

$$\mathrm{erf}\left(\frac{d_{f,q} - \mathbf{w}_k^T\hat{\mathbf{a}}_{f,q}}{\sigma_h\|(\mathbf{I}_{2M} \otimes \mathbf{G}_q^T)\mathbf{w}_k\|}\right) - \mathrm{erf}\left(\frac{-d_{f,q} - \mathbf{w}_k^T\hat{\mathbf{a}}_{f,q}}{\sigma_h\|(\mathbf{I}_{2M} \otimes \mathbf{G}_q^T)\mathbf{w}_k\|}\right) \geq 2p \tag{33}$$

$$j = 1, \cdots, 2K, \quad j \neq k$$
$$f = 2, \cdots, F, \quad q = 1, \cdots, 2K\,.$$

Note that although the problems (13)-(16) and (30)-(33) are mathematically equivalent when the channel mismatch is Gaussian, the original stochastic programming problem is computationally intractable, whereas the NLP problem (30)-(33) can be efficiently solved using sequential quadratic programming (SQP) technique. SQP is an iterative technique in which each search direction is the solution of a quadratic programming (QP) subproblem [8]. The computational complexity of solving QP subproblem using, for example, the primal-dual potential reduction method is $\mathcal{O}(M^{4.5}T^{4.5})$ [15]. Overall complexity of the SQP algorithm depends on the number of iterations, which varies depending on problem-specific parameters and the given batch of data. The SQP algorithm has been implemented, for example, in TOMLAB software package [14], which can be applied to solve the problem (30)-(33).

IV. Simulations

In our simulations, we assume an uplink cellular communication scenario with multiple users and $N = 2$ antennas per user. The Alamouti code [1] is employed and the QPSK modulation scheme is used. The MIMO channel between the fth user and BS is assumed to be quasi-static Rayleigh flat fading with $[\mathbf{H}_f]_{n,m} \sim \mathcal{CN}(0, 1)$. The channel mismatch $\mathbf{E}_f$ is assumed to be independent of $\mathbf{H}_f$ with the distribution $[\mathbf{E}_f]_{n,m} \sim \mathcal{CN}(0, \sigma_h^2)$. Then the CSI mismatch level for channel $\mathbf{H}_f$ is characterized by the parameter σ_h. Throughout the simulations, we assume that $\sigma_h = 1/3$. The interference-to-noise-ratio (INR) is set to be equal to 5 dB.

Five methods are compared in terms of the symbol error rate (SER): the proposed method (30)-(33), the ZF receiver (8), the exact MMSE receiver (9)-(11), the sample MMSE receiver, and the diagonally loaded sample MMSE receiver. Note that the exact MMSE receiver has the perfect knowledge of the user and noise powers and is based on (10) and (11), whereas the sample MMSE receiver estimates the matrix $\mathbf{R}$

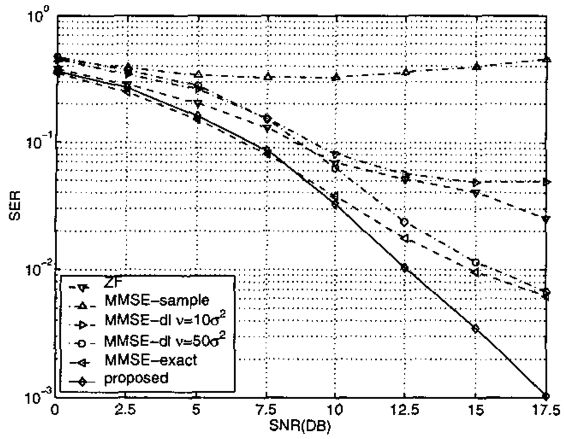

Fig. 1. SER versus SNR, 2 users, $M = 2$.

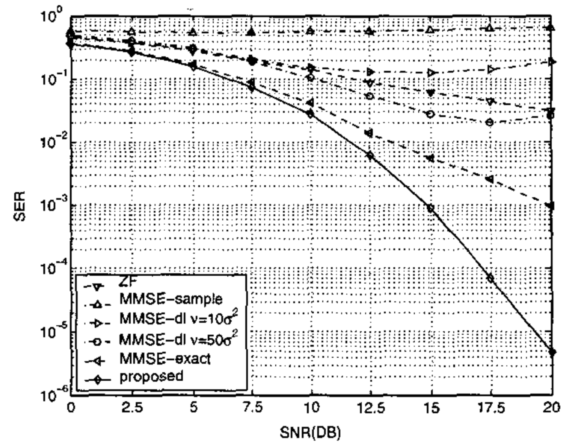

Fig. 2. SER versus SNR, 4 users, $M = 4$.

as

$$\hat{\mathbf{R}} = \frac{1}{L}\sum_{l=1}^{L} \underline{\mathbf{Y}_l}\,\underline{\mathbf{Y}_l}^T$$

where $\underline{\mathbf{Y}_l}$ is the lth received data block and L is the number of available data blocks. The diagonally loaded sample MMSE receiver additionally uses diagonal loading of the sample covariance matrix in the form

$$\hat{\mathbf{R}}_{\rm dl} = \hat{\mathbf{R}} + \nu \mathbf{I}_{2MT}$$

where ν is the DL factor. Two values of $\nu = 10\sigma_n^2$ and $\nu = 50\sigma_n^2$ have been used for the DL sample MMSE receiver. The probability p in the proposed robust methods is set to be equal to 0.95 and 300 Monte Carlo runs are used to obtain each simulated point.

In our first example, we simulate the scenario with $F = 2$ users and $M = 2$ receive antennas at the BS. Figure 1 compares the SERs of the five receivers tested versus the SNR.

In our second example, the scenario with $F = 4$ users and $M = 4$ receive antennas at the BS is considered. Figure 2 displays the SERs of the five receivers tested versus the SNR.

Our simulation figures clearly demonstrate that in both examples, the proposed robust receiver consistently enjoys the best performance among all the methods tested. The performance improvement is especially pronounced at high SNRs.

V. Conclusions

A robust linear receiver for multi-access space-time block-coded MIMO wireless systems has been proposed. Our receiver provides the robustness against CSI errors with a certain selected probability, which can be determined according to a given QoS. The design of the robust receiver boils down to solving the stochastic optimization problem with probability constraints. The implementation of the receiver is based on converting the original stochastic programming problem into the mathematically equivalent deterministic NLP problem if the CSI errors are Gaussian distributed. Numerical simulation results illustrate greatly improved performance of the proposed robust receiver as compared to the traditional ZF and MMSE receivers in the case of imperfectly known (mismatched) CSI.

References

[1] S. M. Alamouti, "A simple transmitter diversity scheme for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1451-1458, Oct. 1998.

[2] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1456-1467, July 1999.

[3] H. Li, X. Lu, and G. B. Giannakis, "Capon multiuser receiver for CSMA systems with space-time coding," *IEEE Trans. Signal Processing*, vol. 50, pp. 1193-1204.

[4] S. Shahbazpanahi, M. Beheshti, A. B. Gershman, M. Gharavi-Alkhansari, and K. M. Wong, "Minimum variance linear receiver for multi-access MIMO wireless systems with space-time block coding," *IEEE Trans. Signal Processing*, accepted (also see *Proc. Asilomar Conference Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2003).

[5] Y. Rong, S. Shahbazpanahi and A. B. Gershman, "Robust linear receivers for space-time block-coded multiple-access MIMO wireless systems," *ICASSP'04*, accepted.

[6] J. R. Birge and F. Louveaux, *Introduction to Stochastic Programming*. New York: Springer-Verlag, 1997.

[7] A. Prékopa, *Stochastic Programming*. Dordrecht, Netherlands: Kluwer Academic Publishers, 1995.

[8] D. P. Bertsekas, *Nonlinear Programming*. Belmont, Athena Scientific, 1995.

[9] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1804-1824, July 2002.

[10] M. Gharavi-Alkhansari and A. B. Gershman, "Contellation space invariance of space-time block codes with application to optimal ML decoding," *Proc. IEEE VTC-Fall*, Orlando, Florida, USA, Oct. 6-9, 2003.

[11] S. Verdú, *Multiuser Detection*. Cambridge: Cambridge university press, 1998.

[12] T. L. Marzetta, "BLAST training: Estimating channel characteristics for high-capacity space-time wireless" in *Proc. 37th Annu. Allerton Conf. Commun., Control, and Comput.*, Sept. 22-24, 1999.

[13] J. W. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. Circuits and Syst.*, vol. 25, pp. 772-781, Sept. 1978.

[14] K. Holmström and A. Göran, "User's guide for Tomlab 4.0.6," *Tomlab Optimization*, Aug. 2003.

[15] Y. Nesterov and A. Nemirovski, *Interior Point Polynomial Algorithms in Convex Programming*, Philadelphia, PA: SIAM, 1994.