Linear Precoder for Codeword Error Minimization in MIMO Systems with Channel Estimation Errors

Boon Sim Thian, Hieu Duy Nguyen, and Sumei Sun

***Abstract*—We propose a precoder design to minimize the codeword error rate of multiple-input multiple-output (MIMO) systems in the presence of channel estimation errors. Our proposed scheme only requires knowledge of the second-order statistics of the channel, noise and the estimation errors; an estimate of the instantaneous channel state information (CSI) is not required. Compared to the instantaneous CSI requirement, our assumption is more practical since obtaining an accurate CSI is challenging if the channel fluctuates rapidly, while the channel statistics are likely to remain unchanged for a much longer period. Furthermore, the feedback overhead from the receiver to the transmitter is greatly reduced. When compared to the case without transmit precoding, our proposed design achieves a significant improvement in error performance: (i) for a real 2×2 MIMO system with 4-PAM modulation and at codeword error rate of 10^{-4}, our proposed precoder design achieves coding gains of up to 6 dB and (ii) for a real 4×4 MIMO system with BPSK modulation and at codeword error rate of 10^{-4}, coding gains of up to 6.5 dB can be achieved.**

***Index Terms*—MIMO communications, imperfect channel state information, robust precoding, optimization**

I. Introduction

The demands for data, multimedia and various other types of services is projected to increase exponentially in the near future. Thus, wireless communication systems will be required to provide substantially higher data rates to a larger number of users. Wireless communication systems with multiple transmit and receive antennas offer significant advantages in terms of increased data rates and reliability than those of single antenna systems [1] [2]. There have been two main approaches that have been proposed to utilize the benefits of multiple-input multiple-output (MIMO) systems. These are the space-time coding scheme which exploits diversity gain to improve communication reliability [3] [4] and the spatial multiplexing method (such as the Bell Laboratories Layered Space-Time (BLAST) scheme [1], [5]) which aims at maximizing the channel throughput. Both approaches require channel state information (CSI) to be available at the receiver only.

If the CSI is also available at the transmitter, then joint transceiver design is considered. Various approaches based on different criteria has been proposed. These include the maximum signal-to-noise ratio (SNR) [6], maximum information rate [7], minimum mean-squared error (MMSE) [8], and bit error rate (BER) [9] transceiver design.

In order to benefit from the advantages of MIMO systems, both the transmitters and receivers must obtain an accurate estimate of the CSI. This is a challenging task in practice, especially for systems with large number of transmit and receiver antennas. Training-symbol based methods [10] [11] and blind channel estimation strategies [12] [13] have been used extensively to estimate the channel. Regardless of the channel estimation techniques used, CSI estimation is subjected to measurement, quantization and other sources of error.

Recent research has moved towards studying transceiver design with imperfect CSI. These include joint channel estimation-signal detection approach [14], transceiver design based on sum-MMSE (SMMSE) [15] [16] and maximin criteria [17] [18]. The aforementioned literature on transceiver designs considered very simple error models, such as only having upper bounds on the magnitude of the errors. However, in practice, channel estimation errors are shown to be correlated due to the channel correlation as well as estimation methods which induce correlation in the errors [11] [21] [22]. Inspired by this observation, a more rigorous correlated multivariate Gaussian channel estimation errors was considered in the problem formulation in [19] and [20].

In this paper, we propose a precoder design for MIMO systems in the presence of channel estimation errors. The objective of our design is to minimize the error probability of the detected signal vector, given that the optimum maximum-likelihood (ML) detector is deployed at the receiver. We demonstrate, via numerical simulations, that significant improvement in error performance can be achieved. Our proposed design *does not* require instantaneous CSI at the transmitter, which leads to less feedback overhead. We assume that the transmitter only has knowledge of the second-order statistics of the channel, noise, and estimation errors. These are the uncertainty attributes of the system and can be obtained from a channel estimator [20] [23]. This assumption on statistical CSI at the transmitter is practical since the actual channel realization may change rapidly, and thus obtaining an accurate CSI estimate is difficult. In contrast, the statistics of the channel are likely to remain unchanged for a much longer period of time. Our design can be viewed as a "semi"-offline approach: The precoder is optimized for each set of uncertainty attributes and is relevant regardless of whether the actual channel realization has changed.

The remainder of this paper is organized as follows. We describe the MIMO system with channel estimation errors in Section II, with a comprehensive treatment of both the channel estimation and data transmission phase. Our precoder design for MIMO systems based on the second-order statistics of the channel, noise, and estimation errors is proposed in Section III. Numerical results and discussions are given in Section IV,

The authors are with the Institute for Infocomm Research (I^{2}R), 1 Fusionopolis Way, #21-01 Connexis (North Tower), Singapore 138632 (email(s): {thianbs, nguyendh, sunsm}@i2r.a-star.edu.sg).

and our conclusions along with future research directions are presented in Section V.

II. System Model

Consider a MIMO system given by

$$\widetilde{\mathbf{y}} = \widetilde{\mathbf{H}}\widetilde{\mathbf{x}} + \widetilde{\mathbf{n}}, \tag{1}$$

where $\widetilde{\mathbf{x}} = [\widetilde{\mathbf{x}}_1, \widetilde{\mathbf{x}}_2, \ldots, \widetilde{\mathbf{x}}_{N'}]^T$ denotes the complex transmitted signal vector of dimension $N' \times 1$; $\widetilde{\mathbf{y}} = [\widetilde{\mathbf{y}}_1, \widetilde{\mathbf{y}}_2, \ldots, \widetilde{\mathbf{y}}_{P'}]^T$ denotes the received signal vector of dimension $P' \times 1$; $\widetilde{\mathbf{H}}$ is the channel matrix of dimension $P' \times N'$; and $\widetilde{\mathbf{n}} = [\widetilde{\mathbf{n}}_1, \widetilde{\mathbf{n}}_2, \ldots, \widetilde{\mathbf{n}}_{P'}]^T$ denotes a vector of circularly symmetric complex Gaussian noise, i.e $\widetilde{\mathbf{n}} \sim \mathcal{CN}(0, 2\sigma_n^2 \mathbf{I}_{P'})$.

The complex system model in (1) can be represented equivalently in the real domain as follows:

$$\begin{bmatrix} \mathrm{Re}\,(\widetilde{\mathbf{y}}) \\ \mathrm{Im}\,(\widetilde{\mathbf{y}}) \end{bmatrix} = \begin{bmatrix} \mathrm{Re}(\widetilde{\mathbf{H}}) & -\mathrm{Im}(\widetilde{\mathbf{H}}) \\ \mathrm{Im}(\widetilde{\mathbf{H}}) & \mathrm{Re}(\widetilde{\mathbf{H}}) \end{bmatrix} \begin{bmatrix} \mathrm{Re}\,(\widetilde{\mathbf{x}}) \\ \mathrm{Im}\,(\widetilde{\mathbf{x}}) \end{bmatrix} + \begin{bmatrix} \mathrm{Re}\,(\widetilde{\mathbf{n}}) \\ \mathrm{Im}\,(\widetilde{\mathbf{n}}) \end{bmatrix} \tag{2}$$

where $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ denote the real and imaginary components, respectively. Letting $\mathbf{y}$, $\mathbf{H}$, $\mathbf{x}$ and $\mathbf{n}$ denote the first, second, third and fourth terms of (2) respectively, we obtain the equivalent real system model,

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}. \tag{3}$$

In this representation, the dimensionality of the system vectors are doubled, i.e $P = 2P'$ and $N = 2N'$. For the rest of the paper, we will work with the real domain representation.

A. Channel Estimation

One of the common methods to estimate the channel in MIMO systems is to transmit training symbols and then estimate the channel based on the received data and training symbols [10] [11]. We consider a block-fading model where the fading coefficients are constant within a coherent interval but may vary from one coherent interval to another.

The received vector during the channel estimation phase is

$$\mathbf{y}_{\mathrm{ce}} = \mathbf{P}\mathbf{h} + \mathbf{n}_{\mathrm{ce}}. \tag{4}$$

where $\mathbf{h} \in \mathbb{R}^{PN\times 1} = \mathrm{vec}\left(\mathbf{H}^T\right)$ denotes the MIMO channel to be estimated, $\mathbf{P} \in \mathbb{R}^{M\times PN}$ denotes a matrix of training symbols (note that $M \geq PN$ so that there will be at least as many equations as unknowns), and $\mathbf{n}_{\mathrm{ce}} \in \mathbb{R}^{M\times 1}$ denotes a vector of Gaussian noise.

We assume that both $\mathbf{h}$ and $\mathbf{n}_{\mathrm{ce}}$ are zero-mean real Gaussian vectors. Their covariance matrices, $\mathbf{\Sigma_h}$ and $\mathbf{\Sigma_n}$, can be estimated via measurements [24] [25]. It is important to note that in this setting, the Bayesian linear-model MMSE estimator which incorporates covariance parameters far outperforms the classical LMMSE channel estimator [26].

Using the Bayesian linear MMSE estimation, the estimate of $\mathbf{h}$ is given by [23]

$$\hat{\mathbf{h}} = \mathbf{K}^{\star}\mathbf{y}_{\mathrm{ce}}, \tag{5}$$

The matrix $\mathbf{K}$ is chosen to minimize the mean-squared error (MSE) between $\mathbf{h}$ and $\hat{\mathbf{h}}$. Therefore, we have

$$\mathbf{K}^{\star} = \arg\min_{\mathbf{K}} \mathbb{E}||\mathbf{K}\left(\mathbf{P}\mathbf{h} + \mathbf{n}_{\mathrm{ce}}\right) - \mathbf{h}||^2. \tag{6}$$

Differentiating the expression in (6) with respect to $\mathbf{K}$ and setting it to zero, it can be shown that [23]

$$\mathbf{K}^{\star} = \mathbf{\Sigma_h}\mathbf{P}^T\left(\mathbf{P}\mathbf{\Sigma_h}\mathbf{P}^T + \mathbf{\Sigma_n}\right)^{-1}. \tag{7}$$

Defining the estimation error as

$$\mathbf{e} = \hat{\mathbf{h}} - \mathbf{h}, \tag{8}$$

it can be shown that $\mathbf{e}$ is a zero-mean Gaussian vector with the following covariance matrix [23]

$$\mathbf{\Sigma_e} \triangleq \mathbb{E}||\hat{\mathbf{h}} - \mathbf{h}||^2 = \left(\mathbf{P}^T\mathbf{\Sigma_n}^{-1}\mathbf{P} + \mathbf{\Sigma_h}^{-1}\right)^{-1}. \tag{9}$$

The covariance matrix of the estimate, $\hat{\mathbf{h}}$ is given by

$$\begin{aligned} \mathbf{\Sigma_{\hat{h}}} &= \mathbb{E}\left[\mathbf{K}^{\star}\mathbf{y}_{\mathrm{ce}}\mathbf{y}_{\mathrm{ce}}^T\mathbf{K}^{\star T}\right] \\ &= \mathbf{\Sigma_h}\mathbf{P}^T\left(\mathbf{P}\mathbf{\Sigma_h}\mathbf{P}^T + \mathbf{\Sigma_n}\right)^{-T}\mathbf{P}\mathbf{\Sigma_h}. \end{aligned} \tag{10}$$

The dimensions of $\mathbf{\Sigma_e}$ and $\mathbf{\Sigma_{\hat{h}}}$ are both $PN \times PN$. Furthermore, by the orthogonality principle, $\mathbf{e}$ is uncorrelated with $\hat{\mathbf{h}}$. The channel estimator described above provides the channel estimate $\hat{\mathbf{h}}$ and the error covariance matrix $\mathbf{\Sigma_e}$.

From (9), it is observed that the error covariance matrix, $\mathbf{\Sigma_e}$, may not necessarily be a scaled identity matrix, which is the assumption of most works in literature. Assuming orthogonal pilot symbols (i.e $\mathbf{P}^T\mathbf{P} = \mathbf{I}$) and white noise (i.e $\mathbf{\Sigma_n} = \sigma_n^2\mathbf{I}$), the error covariance matrix becomes

$$\mathbf{\Sigma_e} = \left(\frac{1}{\sigma_n^2}\mathbf{I} + \mathbf{\Sigma_h}^{-1}\right)^{-1}. \tag{11}$$

Recall that the dimensions of $\mathbf{\Sigma_h}$ is $PN \times PN$, hence there are P^2 blocks of $N \times N$ sub-matrices in $\mathbf{\Sigma_h}$. It is worth noting the structure of $\mathbf{\Sigma_h}$ in the following two cases:

(i) Transmit antennas correlation: Every $N \times N$ sub-matrix of $\mathbf{\Sigma_h}$ is a diagonal matrix.

(ii) Receive antennas correlation: $\mathbf{\Sigma_h}$ is a block diagonal matrix. Except for the main $N \times N$ block diagonals of $\mathbf{\Sigma_h}$, all other elements are zero.

In this paper, we assume only receive antenna correlation; a future work addresses the case with transmit antenna correlation as well as full correlation.

B. System with Channel Estimation Errors

Using (3) and (8), the received signal during the data transmission phase can be expressed as

$$\mathbf{y} = \hat{\mathbf{H}}\mathbf{x} - \mathbf{E}\mathbf{x} + \mathbf{n}, \tag{12}$$

where $\mathbf{E}$ is the error matrix and $\hat{\mathbf{H}}$ is the channel estimate. Denoting $\mathbf{e} = \mathrm{vec}(\mathbf{E}^T)$ and $\hat{\mathbf{h}} = \mathrm{vec}(\hat{\mathbf{H}}^T)$, the second-order statistics of $\mathbf{E}$ and $\hat{\mathbf{H}}$ are given by (9) and (10) respectively. In this paper, we assume that

(i) The receiver has channel estimate $\hat{\mathbf{H}}$ and the error statistics $\mathbf{\Sigma_e}$; and

(ii) The transmitter has knowledge of $\mathbf{\Sigma_{\hat{h}}}$ and $\mathbf{\Sigma_e}$.

For this system model, the signal-to-noise (SNR) ratio of the i^{th} received antenna is

$$\mathrm{SNR[i]} = \frac{\sum_{j=1}^{N}\bar{\mathcal{E}}_{\mathbf{x}_j}\mathbb{E}\left[\mathbf{H}_{i,j}^2\right]}{\sigma_n^2}, \tag{13}$$

where $\bar{\mathcal{E}}_{\mathbf{x}_j} = \mathbb{E}\left[\mathbf{x}_j^2\right]$ is the average energy of the symbols from the j^{th} transmit antenna. The average received SNR of the MIMO system is thus defined as

$$\overline{\mathrm{SNR}} = \frac{1}{P}\sum_{i=1}^{P}\mathrm{SNR}[i]. \tag{14}$$

In a similar fashion, the signal-to-interference-plus-noise ratio (SINR) of the i^{th} received antenna and the average SINR of the MIMO system are respectively given by

$$\mathrm{SINR}\left[i\right] = \frac{\sum_{j=1}^{N}\bar{\mathcal{E}}_{\mathbf{x}_j}\mathbb{E}\left[\mathbf{H}_{i,j}^2\right]}{\sum_{j=1}^{N}\bar{\mathcal{E}}_{\mathbf{x}_j}\mathbb{E}\left[\mathbf{E}_{i,j}^2\right] + \sigma_n^2}, \tag{15}$$

and

$$\overline{\mathrm{SINR}} = \frac{1}{P}\sum_{i=1}^{P}\mathrm{SINR}[i]. \tag{16}$$

C. Optimum ML Detection

The optimum ML detector is given by [20]

$$\begin{aligned}\hat{\mathbf{x}}_{\mathrm{ml}} &= \arg\max_{\mathbf{x}\in\mathcal{X}} \log\Pr\left(\mathbf{y}|\hat{\mathbf{H}},\mathbf{x}\right)\\ &= \arg\min_{\mathbf{x}\in\mathcal{X}}\left(\log\det\boldsymbol{\Sigma}_{\mathbf{x}} + \left(\mathbf{y}-\hat{\mathbf{H}}\mathbf{x}\right)^T\boldsymbol{\Sigma}_{\mathbf{x}}^{-1}\left(\mathbf{y}-\hat{\mathbf{H}}\mathbf{x}\right)\right),\end{aligned} \tag{17}$$

where

$$\boldsymbol{\Sigma}_{\mathbf{x}(i,j)} = \begin{cases}\sigma_n^2 + \mathbf{x}^T\boldsymbol{\Sigma}_{\mathbf{e}(i-1)N+1:iN,(j-1)N+1:jN}\mathbf{x}, & \text{if } i=j;\\ \mathbf{x}^T\boldsymbol{\Sigma}_{\mathbf{e}(i-1)N+1:iN,(j-1)N+1:jN}\mathbf{x}, & \text{otherwise.}\end{cases} \tag{18}$$

is the covariance matrix of $\mathbf{y}$, given that vector $\mathbf{x}$ is transmitted. The optimum detector can be implemented in a reduced complexity fashion (see [20] for the implementation).

III. Precoder Design for MIMO Systems with Channel Estimation Errors

Fig. 1 shows a simplified block diagram of the system model. A message vector $\mathbf{v}$ is generated from the source and encoded through a precoder $\mathbf{T}$ (our proposed precoder design will be described subsequently). It is transmitted via the MIMO channel, $\mathbf{H}$. At the receiver, the optimum ML detector is deployed. Our objective is to design $\mathbf{T}$ to minimize the worst-case pairwise codeword error probability under the constraint of only having statistical CSI at the transmitter. We present a description of the pairwise codeword error probability, our problem formulation and precoder design approach in the subsections below.

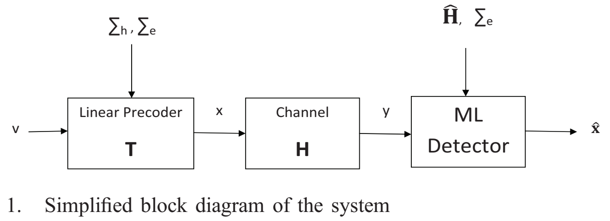

Fig. 1. Simplified block diagram of the system

A. Pairwise Error Probability

Let $\mathbf{v},\mathbf{w}\in\mathcal{V}$ be two distinct uncoded signal vectors, and $\mathbf{x}=\mathbf{T}\mathbf{v}$ and $\mathbf{z}=\mathbf{T}\mathbf{w}$ be the codeword vectors after precoding. Suppose $\mathbf{z}$ is transmitted, an upper bound on the average probability that the optimum detector decides in favor of $\mathbf{x}$ is [20]

$$\begin{aligned}P_{\mathrm{ub}|\mu}\left(\mathbf{z}\to\mathbf{x}\right) =& \frac{1}{2}\left\{\frac{\det\left[\boldsymbol{\Sigma}_{\mathbf{Z}}\right]^{\mu}\det\left[\boldsymbol{\Sigma}_{\mathbf{X}}\right]^{1-\mu}}{\det\left[\mu\boldsymbol{\Sigma}_{\mathbf{Z}} + \left(1-\mu\right)\boldsymbol{\Sigma}_{\mathbf{X}}\right]}\right\}^{1/2}\\ &\cdot\frac{1}{\left\{\det\left[\mathbf{I}_{PN} + 2\mathbf{L}\left(\mu\right)\boldsymbol{\Sigma}_{\hat{\mathbf{h}}}\right]\right\}^{1/2}},\end{aligned} \tag{19}$$

where $\mathbf{X} = \mathbf{I}_P\otimes\mathbf{x}^T$, $\mathbf{Z} = \mathbf{I}_P\otimes\mathbf{z}^T$, $\boldsymbol{\Sigma}_{\mathbf{Z}}$ and $\boldsymbol{\Sigma}_{\mathbf{X}}$ are the covariance matrix of $\mathbf{y}$ given that $\mathbf{z}$ and $\mathbf{x}$ is transmitted, respectively and

$$\mathbf{L}\left(\mu\right) = \frac{\mu\left(1-\mu\right)}{2}\left(\mathbf{Z}-\mathbf{X}\right)^T\left\{\mu\boldsymbol{\Sigma}_{\mathbf{Z}} + \left(1-\mu\right)\boldsymbol{\Sigma}_{\mathbf{X}}\right\}^{-1}\left(\mathbf{Z}-\mathbf{X}\right). \tag{20}$$

The upper bound (19) is valid for $0<\mu<1$. The best upper bound can be found by optimizing over μ. A simpler (but not optimal) bound can be obtained by setting $\mu=\frac{1}{2}$.

As mentioned previously, we focus on the case when there is only receive antenna correlation. In this case, $\boldsymbol{\Sigma}_{\mathbf{h}}$ is a block diagonal matrix and it can easily be shown from (11) that $\boldsymbol{\Sigma}_{\mathbf{e}}$ and $\boldsymbol{\Sigma}_{\hat{\mathbf{h}}}$ are also block diagonal matrices. In addition, $\boldsymbol{\Sigma}_{\mathbf{x}}$ and respectively $\boldsymbol{\Sigma}_{\mathbf{z}}$, are given by

$$\boldsymbol{\Sigma}_{\mathbf{x}(i,j)} = \begin{cases}\sigma_n^2 + \mathbf{x}^T\mathbf{A}^{(i)}\mathbf{x}, & \text{if } i=j;\\ 0, & \text{otherwise,}\end{cases} \tag{21}$$

and

$$\boldsymbol{\Sigma}_{\mathbf{z}(i,j)} = \begin{cases}\sigma_n^2 + \mathbf{z}^T\mathbf{A}^{(i)}\mathbf{z}, & \text{if } i=j;\\ 0, & \text{otherwise.}\end{cases} \tag{22}$$

where we have simplified the notation by defining $\mathbf{A}^{(i)} \triangleq \boldsymbol{\Sigma}_{\mathbf{e}(i-1)N+1:iN,(i-1)N+1:iN}$; $\mathbf{A}^{(i)}$ is the i^{th} main $N\times N$ block diagonal of $\boldsymbol{\Sigma}_{\mathbf{e}}$.

Using (21), (22) and for $\mu=\frac{1}{2}$, we write (19) as

$$\begin{aligned}&P_{\mathrm{ub}|\mu=\frac{1}{2}}\left(\mathbf{z}\to\mathbf{x}\right)\\ &= \frac{\prod_{i=1}^{N}\left(\mathbf{x}^T\mathbf{A}^{(i)}\mathbf{x}+\sigma_n^2\right)^{\frac{1}{4}}\left(\mathbf{z}^T\mathbf{A}^{(i)}\mathbf{z}+\sigma_n^2\right)^{\frac{1}{4}}}{\prod_{i=1}^{N}\left(\mathbf{d}^T\mathbf{B}^{(i)}\mathbf{d} + 2\mathbf{x}^T\mathbf{A}^{(i)}\mathbf{x} + 2\mathbf{z}^T\mathbf{A}^{(i)}\mathbf{z} + 2\sigma_n^2\right)^{\frac{1}{2}}},\end{aligned} \tag{23}$$

where $\mathbf{d}=\mathbf{z}-\mathbf{x}$ and $\mathbf{B}^{(i)} \triangleq \boldsymbol{\Sigma}_{\hat{\mathbf{h}}(i-1)N+1:iN,(i-1)N+1:iN}$ is the i^{th} main $N\times N$ block diagonal of $\boldsymbol{\Sigma}_{\hat{\mathbf{h}}}$.

B. Problem Formulation

The precoder design problem is formulated as

$$\begin{array}{ll}\text{minimize} & \max_{\{\forall\mathbf{v},\mathbf{w}\in\mathcal{V}:\mathbf{w}\neq\mathbf{v}\}} P\left(\mathbf{z}\to\mathbf{x}\right)\\ \text{subject to} & \mathrm{Tr}\left(\mathbf{T}\mathbf{T}^T\right)\le\sum_{j=1}^{N}\bar{\mathcal{E}}_{\mathbf{x}_j}\end{array} \tag{24}$$

with variable $\mathbf{T}\in\mathbb{R}^{P\times P}$. The constraint in (24) restricts the transmit power not to exceed the average transmit power of the original signal vectors. We note that this optimization problem is non-convex in $\mathbf{T}$. In the next subsection, we describe our design approach to overcome the nonconvexity of the problem.

We first present an equivalent formulation of the problem. Taking logarithm of (23), we obtain

$$\begin{aligned}&\log P_{\mathrm{ub}|\mu=\frac{1}{2}}\left(\mathbf{Tw}\rightarrow\mathbf{Tv}\right)\\ =&\frac{1}{4}\sum_{i=1}^{N}\Big[\log\left(\mathbf{v}^T\mathbf{T}^T\mathbf{A}^{(i)}\mathbf{Tv}+\sigma_n^2\right)\\ &\qquad+\log\left(\mathbf{w}^T\mathbf{T}^T\mathbf{A}^{(i)}\mathbf{Tw}+\sigma_n^2\right)\Big]\\ -&\frac{1}{2}\sum_{i=1}^{N}\log\Big(\mathbf{d}^T\mathbf{T}^T\mathbf{B}^{(i)}\mathbf{Td}+2\mathbf{v}^T\mathbf{T}^T\mathbf{A}^{(i)}\mathbf{Tv}\\ &\qquad+2\mathbf{w}^T\mathbf{T}^T\mathbf{A}^{(i)}\mathbf{Tw}+2\sigma_n^2\Big).\end{aligned}\tag{25}$$

Let $\mathbf{t} = \mathrm{vec}\left(\mathbf{T}^T\right)$ and using our earlier definition that $\mathbf{V} = \mathbf{I}_P \otimes \mathbf{v}^T$, we can write $\mathbf{v}^T\mathbf{T}^T\mathbf{A}^{(i)}\mathbf{Tv} = \mathbf{t}^T\mathbf{A}^{(i,\mathbf{v})}\mathbf{t}$, where $\mathbf{A}^{(i,\mathbf{v})} \triangleq \mathbf{V}^T\mathbf{A}^{(i)}\mathbf{V}$. Using similar definitions, $\mathbf{A}^{(i,\mathbf{w})} \triangleq \mathbf{W}^T\mathbf{A}^{(i)}\mathbf{W}$ and $\mathbf{B}^{(i,\mathbf{d})} \triangleq \mathbf{D}^T\mathbf{B}^{(i)}\mathbf{D}$ (where $\mathbf{D} = \mathbf{I}_P \otimes \mathbf{d}^T$), we re-write (25) as

$$\begin{aligned}&f_{\mathbf{Tw},\mathbf{Tv}}(\mathbf{t}) \triangleq \log P_{\mathrm{ub}|\mu=\frac{1}{2}}\left(\mathbf{Tw}\rightarrow\mathbf{Tv}\right)\\ &=\frac{1}{4}\sum_{i=1}^{N}\left[\log\left(\mathbf{t}^T\mathbf{A}^{(i,\mathbf{v})}\mathbf{t}+\sigma_n^2\right)+\log\left(\mathbf{t}^T\mathbf{A}^{(i,\mathbf{w})}\mathbf{t}+\sigma_n^2\right)\right]\\ &-\frac{1}{2}\sum_{i=1}^{N}\log\left(\mathbf{t}^T\mathbf{B}^{(i,\mathbf{d})}\mathbf{t}+2\mathbf{t}^T\mathbf{A}^{(i,\mathbf{v})}\mathbf{t}+2\mathbf{t}^T\mathbf{A}^{(i,\mathbf{w})}\mathbf{t}+2\sigma_n^2\right).\end{aligned}\tag{26}$$

Since logarithm is a monotonic increasing function, (24) can equivalently be formulated as

$$\begin{aligned}\text{minimize}\quad&\tau\\ \text{subject to}\quad&f_{\mathbf{Tw},\mathbf{Tv}}(\mathbf{t})\leq\tau,\quad\forall\,\mathbf{w},\mathbf{v}\in\mathcal{W}:\mathbf{w}\neq\mathbf{v}\\ &\mathbf{t}^T\mathbf{t}\leq\textstyle\sum_{j=1}^{N}\bar{\mathcal{E}}_{\mathbf{x}_j}\end{aligned}\tag{27}$$

with variables $\tau \in \mathbb{R}$ and $\mathbf{t} \in \mathbb{R}^{P^2}$.

C. First-Order Approximation Approach

Unfortunately, (27) is still a non-convex program. We propose using an iterative optimization approach to find a locally-optimal solution of (27) based on the first-order approximation of the function $f_{\mathbf{Tw},\mathbf{Tv}}(\mathbf{t})$. In iteration k, we approximate the function $f_{\mathbf{Tw},\mathbf{Tv}}(\mathbf{t})$ around its previous feasible point $\mathbf{t}^{(k-1)}$. Note that $f_{\mathbf{Tw},\mathbf{Tv}}(\mathbf{t})$ can be linearized at a point $\mathbf{s}$ as

$$f_{\mathbf{Tw},\mathbf{Tv}}(\mathbf{t})\approx f_{\mathbf{Tw},\mathbf{Tv}}(\mathbf{s})+\nabla f_{\mathbf{Tw},\mathbf{Tv}}(\mathbf{s})^T(\mathbf{t}-\mathbf{s}),\tag{28}$$

where, in our case,

$$\begin{aligned}\nabla f_{\mathbf{Tw},\mathbf{Tv}}=&\frac{1}{2}\sum_{i=1}^{N}\left[\frac{\mathbf{A}^{(i,\mathbf{v})}\mathbf{t}}{\mathbf{t}^T\mathbf{A}^{(i,\mathbf{v})}\mathbf{t}+\sigma_n^2}+\frac{\mathbf{A}^{(i,\mathbf{w})}\mathbf{t}}{\mathbf{t}^T\mathbf{A}^{(i,\mathbf{w})}\mathbf{t}+\sigma_n^2}\right]\\ &-\sum_{i=1}^{N}\left[\frac{\mathbf{B}^{(i,\mathbf{d})}\mathbf{t}+2\mathbf{A}^{(i,\mathbf{v})}\mathbf{t}+2\mathbf{A}^{(i,\mathbf{w})}\mathbf{t}}{\mathbf{t}^T\mathbf{B}^{(i,\mathbf{d})}\mathbf{t}+2\mathbf{t}^T\mathbf{A}^{(i,\mathbf{v})}\mathbf{t}+2\mathbf{t}^T\mathbf{A}^{(i,\mathbf{w})}\mathbf{t}+2\sigma_n^2}\right].\end{aligned}\tag{29}$$

Using (28) and (29), the optimization problem in the k^{th} iteration is

$$\begin{aligned}\text{minimize}\quad&\tau\\ \text{subject to}\quad&f_{\mathbf{Tw},\mathbf{Tv}}\left(\mathbf{t}^{(k-1)}\right)+\\ &\nabla f_{\mathbf{Tw},\mathbf{Tv}}\left(\mathbf{t}^{(k-1)}\right)^T\left(\mathbf{t}-\mathbf{t}^{(k-1)}\right)\leq\tau,\forall\,\mathbf{w},\mathbf{v}\in\mathcal{V}:\mathbf{w}\neq\mathbf{v}\\ &\mathbf{t}_j^{(k-1)}-\varepsilon\leq\mathbf{t}_j^{(k)}\leq\mathbf{t}_j^{(k-1)}+\varepsilon,\quad 1\leq j\leq P^2\\ &\mathbf{t}^{(k)T}\mathbf{t}^{(k)}\leq\textstyle\sum_{j=1}^{N}\bar{\mathcal{E}}_{\mathbf{x}_j}\end{aligned}\tag{30}$$

with variables $\tau \in \mathbb{R}$ and $\mathbf{t}^{(k)} \in \mathbb{R}^{P^2}$. In (30), $\mathbf{t}^{(k-1)}$ is the solution obtained in the previous iteration. In addition, the second constraint is the trust region constraint and ε is known as the radius of the trust region. We summarize our proposed approach in Algorithm I below.

Algorithm 1 Proposed precoder design for MIMO systems with channel estimation errors

1: Input: N, P, trust region width ε, initial feasible point $\mathbf{t}^{(0)}$, max. no. of iterations $k_{\max}$, tolerance η.
2: $k \leftarrow 1$
3: **while** $k \leq k_{\max}$ and $||\mathbf{t}^{(k)} - \mathbf{t}^{(k-1)}||^2 > \eta$ **do**
4: Compute $\nabla f_{\mathbf{Tw},\mathbf{Tv}}$ using (29).
5: Use $\nabla f_{\mathbf{Tw},\mathbf{Tv}}$ and previous estimate $\mathbf{t}^{(k-1)}$ to solve optimization problem (30).
6: $k \leftarrow k + 1$.
7: **end while**
8: Output: $\mathbf{t}^{(k)}$.

IV. Numerical Results and Discussion

In this section, we present some numerical simulation results for different system parameters. In all the simulations, we set $\mathbf{\Sigma_n} = \sigma_n^2\mathbf{I}$ and use orthogonal pilot symbols (i.e $\mathbf{P}^T\mathbf{P} = \mathbf{I}$) for channel estimation. For receive antenna correlation, in each of the main sub-blocks of $\mathbf{\Sigma_h}$, the model $\mathbf{\Sigma}_{\mathbf{h}(i,j)} = c\alpha^{|i-j|}$, $0 < \alpha < 1$, is used. This covariance matrix model is frequently used in the literature (see [11], [21], [22], and references therein). Thus, the error covariance $\mathbf{\Sigma_e}$ is given by (11). The definitions of SNR and SINR are given by (14) and (16), respectively.

A. 2×2 MIMO system

In the first set of numerical simulations, we consider a 2×2 real MIMO system with 4-PAM modulation (i.e $\mathcal{V} = \{-3, -1, 1, 3\}$).

1) Convergence of the proposed iterative optimization: First, we study the convergence properties of the proposed iterative optimization method. The top graph in Fig. 2 depicts the evolution of the objective function value with the number of iterations. The bottom graph in Fig. 2 shows the the logarithm of the Euclidean distance of the difference of the estimates versus the number of iterations, i.e $\log||\mathbf{t}^{(k)}-\mathbf{t}^{(k-1)}||^2$. This is a measure of the variation of the estimates across consecutive iterations (smaller value indicates smaller variation and vice versa). Two observations are made from the top graph: (i) The trust region radius needs to be sufficiently small in order for convergence to occur. In this example, the objective function values converges for $\varepsilon = 0.001, 0.05$ and 0.01. This is mathematically intuitive because the linearization model (28) is only "locally" accurate, and (ii) a smaller trust region radius will require more iterations for convergence and vice versa. Furthermore, the amplitude of the fluctuations about the steady state value is also smaller for smaller trust region radius.

It is observed from the bottom graph of Fig. 2 that smaller trust region radius also leads to smaller variation across consecutive estimates; for $\varepsilon = 0.001, 0.005$ and 0.01, $||\mathbf{t}^{(k)} - \mathbf{t}^{(k-1)}||^2 = 10^{-13}, 10^{-10}$ and $10^{-8.5}$ respectively.

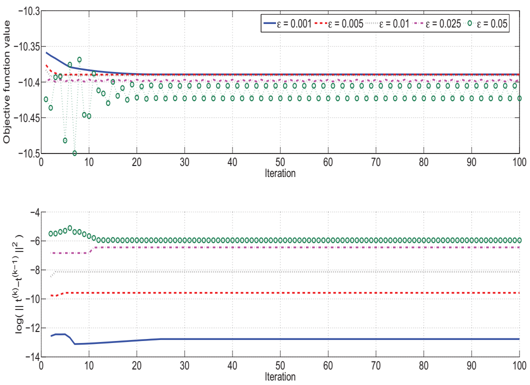

Fig. 2. Convergence of the objective function value and optimization variable

2) Constellation points after optimized precoding:

Next, we present the results for the optimized signal constellation set after precoding. Since 4-PAM modulation is used, there is a total of $4^2 = 16$ signal vectors in the original constellation set. This is illustrated in a 2-dimensional diagram on the left graph of Fig. 3. The horizontal axis shows the first element of the signal vector (corresponding to the first transmit antenna) and the vertical axis shows the second element (corresponding to the second transmit antenna). The right graph in Fig. 3 illustrates the signal constellation set after optimized precoding. It is observed that the constellation set is altered in order to achieve a better error performance.

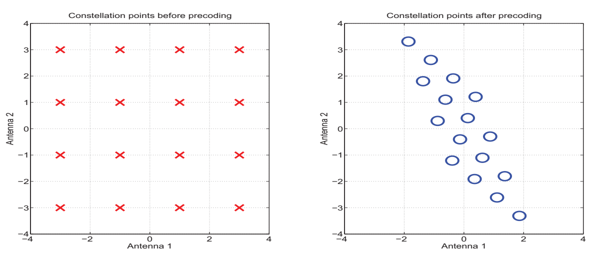

Fig. 3. 2×2 MIMO: Constellation points before and after precoding

3) Error probability performance:

In the third part of this set of simulations, we compare the codeword error probability performance for the same system configuration. For comparison purposes, the following precoder and detector setups are considered: (i) No precoding at the transmitter with Euclidean distance detection $||\mathbf{y} - \hat{\mathbf{H}}\mathbf{x}||^2$ at the receiver, (ii) No precoding at the transmitter with optimum detection (17) at the receiver, (iii) Proposed precoder at the transmitter with Euclidean distance detection $||\mathbf{y} - \hat{\mathbf{H}}\mathbf{x}||^2$ at the receiver, and (iv) Proposed precoder at the transmitter with optimum detection (17) at the receiver.

These different setups are considered in order to illustrate the incremental benefits of each scheme. Fig. 4 illustrates the codeword error rate versus SINR performance. It is observed from Fig. 4 that the lowest error probability is attained using our proposed precoder in conjunction with optimum detection. It is intuitive since the proposed precoder was designed based on the optimum detection at the receiver. Therefore, a mismatched detector (in this case, Euclidean distance based detector) will lead to degradation in error performance. Fig. 4 shows that substantial improvement in error probability can be attained: At a codeword error rate of 10^{-4}, our proposed precoder achieves coding gains of 4.5 to 6 dB over the scenarios without transmit precoding.

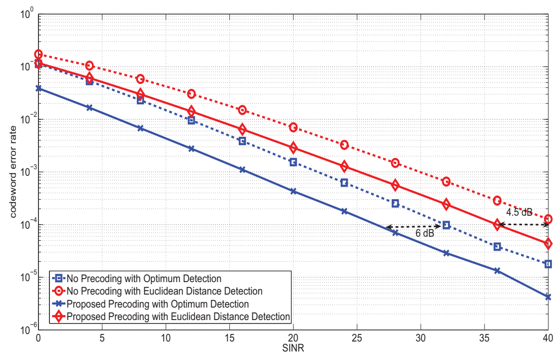

Fig. 4. 2×2 MIMO system with channel estimation errors: Average SER vs SINR for the different setups

B. 4×4 MIMO system

In the second set of numerical simulations, we present the error probability results for a 4×4 MIMO system with BPSK modulation, i.e $\mathcal{V} = \{-1, 1\}$.

Figs. 5 and 6 illustrate the codeword error probability versus SINR and SNR, respectively. In Fig. 6, the power of the channel estimation errors is set to be 5 times of the channel power. Both figures show that the most significant improvement in error probability performance is achieved with our proposed precoder design. It is observed from Fig. 5 that at codeword error rate of 10^{-4}, significant coding gains of 5.5 to 6.5 dB can be achieved by using our proposed precoder.

Figs. 5 and 6 also show that if no precoder is used, then there is noticeable benefit in using optimum detection over Euclidean distance detection. On the other hand, if our proposed precoder is used at the transmitter, then the benefit of using the optimum detector over Euclidean distance decoding is minimal. It is worth noting that error floors are observed in Fig. 6. This is because as SNR increases, the system becomes "interference-limited". The "interference" in this scenario is induced by the presence of channel estimation errors.

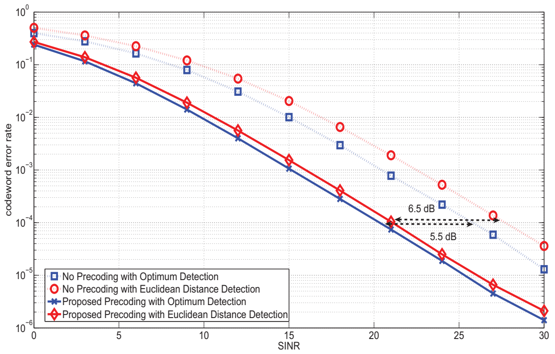

Fig. 5. 4×4 MIMO system with channel estimation errors: Average SER vs SINR for the different setups

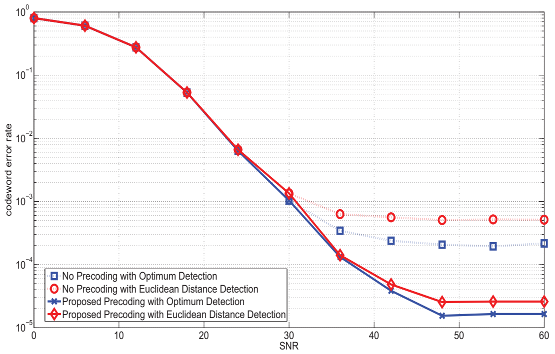

Fig. 6. 4×4 MIMO system with channel estimation errors: Average SER vs SNR for the different setups

V. Conclusion

In practical communication systems, parameter estimation is often erroneous due to imperfect measurement, quantization, and other sources of error. In this paper, we propose a precoder design for error rate minimization in MIMO systems with channel estimation errors. We assume that the receiver knows the instantaneous estimate and the statistics of the channel, while the transmitter only has the knowledge of its statistics. Our design therefore does not require instantaneous CSI feedback, which leads to considerable savings in communication resources. We show, via numerical simulations, that our proposed precoder design leads to significant improvement in error performance. Future studies will consider precoder design with a more general assumption on the detector.

References

[1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antenna," *Wireless Personal Communication*, vol. 6, no. 3, pp. 311-335, Mar. 1998.

[2] R. Health, S. Sandhu, and A. Paulraj, "Antenna selection for spatial multiplexing systems with linear receivers", *IEEE Communication Letters*, vol. 5, no. 4, pp. 142-144, Apr. 2001.

[3] S. M. Alamouti, "A simple transmit diversity techniques for wireless communications, *IEEE J. Sel. Areas Commun.*, vol. 16, no. 10, pp. 14511458, Oct. 1998.

[4] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communications: Performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744765, Mar. 1998.

[5] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas, " *Bell Labs Tech. J.*, vol. 1, pp. 4159, 1996.

[6] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers part I: Unification and optimal designs," *IEEE Trans. Signal Process.*, vol. 47, no. 7, pp. 19882006, Jul. 1999.

[7] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 21982206, Dec. 2001.

[8] J. Yang and S. Roy, "On joint transmitter and receiver optimization for multiple-input-multiple-output (MIMO) transmission systems," *IEEE Trans. Commun.*, vol. 42, no. 12, pp. 32213231, Dec. 1994.

[9] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 10511064, May 2002.

[10] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?, " *IEEE commun. Letters*, vol. 5, no. 4, pp. 142-144, Apr. 2001.

[11] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: A study of estimator tradeoffs and optimal training signals," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 884-893, Mar. 2006

[12] S. Shahbazpanahi, A. B. Gershman and J. H. Manton, "Closed-form blind MIMO channel estimation for orthogonal space-time block codes," *IEEE Trans. Signal Process.*, vol. 53, no. 12, pp. 4506-4517, Dec. 2005.

[13] J. K. Tugnait and B. Huang, "Blind channel estimation and equalization of multiple-input and multiple-output channels," *Proceedings of IEEE Int. Conf. on Personal Wireless Comms.*, pp. 231-235, Feb. 1999

[14] G. Taricco and E. Biglieri, "Space-time decoding with imperfect channel estimation," *IEEE. Trans. Wireless Comms.*, vol. 4, no. 4, pp. 1874-1888, Jul. 2005

[15] H. Li and C. Xu, "Robust optimization of linear precoders/decoders for multiuser MIMO downlink with imperfect CSI at base station," *IEEE WCNC 2007*, pp. 1130-1134, Mar. 2007.

[16] H. Wang, X. Xu, M. Zhao, W. Wu and Y. Yao, "Robust transmission for multiuser MIMO downlink systems with imperfect CSIT, " *IEEE WCNC 2008*, pp. 340-344, Oct. 2008.

[17] A. P-Iserte, D. P. Palomar, A. I. P-Neira and M. A. Lagunas, "A robust maximim appoach for MIMO communications with imperfect channel state information based on convex optimization," *IEEE. Trans. Signal Process.*, vol. 54, no. 1, pp. 346 - 360, Jan. 2006.

[18] A. Farhoodi and M. Biguesh, "Robust ML detection algorithm for MIMO receivers in presence of channel estimation error, " *IEEE PIMRC* 2006.

[19] B. S. Thian and A. Goldsmith, "Decoding for MIMO systems with correlated channel estimation errors," *Allerton 2011*, pp. 524-530, Sep. 2011.

[20] B. S. Thian and A. Goldsmith, "Reduced-Complexity Robust MIMO Decoders," *IEEE Trans. Wireless Comms.*, vol. 12, no. 8, pp. 3783 - 3795, Aug. 2013.

[21] A. B. Gershman, C. F. Mecklenbrauker and J. F. Bohme, "Matrix fitting approach to direction of arrival estimation with imperfect spatial coherence of wavefronts," *IEEE Trans. Signal Process.*, vol. 45, pp. 1894-1899, Jul. 1997

[22] J. Ringelstein, A. B. Gershman and J. F. Bohme, "Direction finding in random inhomogenousmedia in the presence of multiplicative noise fields," *IEEE Sig. Process. Letters*, vol. 7, pp. 269-272, Oct. 2000.

[23] S. Boyd, Lecture notes for EE363: Linear Dynamical Systems, Stanford University, 2008.

[24] A. Dogandzic and J. Jin, "Estimating statistical properties of MIMO fading channels," *IEEE Trans. Signal Process.*, vol. 53, no.8, pp. 3065-3080, Aug. 2005.

[25] S. Wang, A. Abdi, J. Salo, H.M. El-Sallabi, J.W. Wallace, P. Vainikainen and M.A. Jensen, "Time-Varying MIMO Channels: Parametric Statistical Modeling and Experimental Results, " *IEEE Trans. Veh. Tech.*, vol. 55, no. 4, pp. 1949-1963, Jul. 2007.

[26] S. M. Kay, "Fundamentals of Statistical Signal Processing - Estimation Theory.", Englewood Cliffs, NJ: Prentice-Hall, 1993.