Statistical Precoding for MIMO Systems With Channel Estimation Errors

Boon Sim Thian, Hieu Duy Nguyen, and Sumei Sun

***Abstract*—Obtaining accurate instantaneous channel state information (CSI) is challenging for multiple-input–multiple-output (MIMO) systems, particularly if the channel fluctuates rapidly. A more practical assumption is statistical CSI as the channel statistics are likely to remain unchanged for a longer period. In this letter, we propose a precoder design, using only statistical CSI, to minimize error probability for practical MIMO systems with channel estimation errors. Our proposed precoder design is shown to achieve a significant improvement in error performance when compared with other precoding schemes in literature. For example, for a real 4 × 4 MIMO system with BPSK modulation and at a codeword error rate of 10^{-3}, coding gains of up to 12 dB can be achieved.**

***Index Terms*—Multiple-input–multiple-output (MIMO), error probability, imperfect channel state information (CSI), robust precoding.**

I. Introduction

With significant improvement in data rate and reliability over their single antenna counterparts, multi-antenna systems have emerged as candidates for next generation communication networks. The two main methods to harness the benefits of multiple-input multiple-output (MIMO) systems are spatial multiplexing [1] and space-time coding [2].

Initial studies assumed perfect channel state information (CSI) at both the transmitter and receiver. While the assumption of perfect CSI provides a benchmark on the achievable system performance, it is extremely difficult to achieve in practice. The system design under perfect CSI assumption is not robust to CSI imperfection, and thus suffers from significant performance degradation even with modest CSI mismatch at the transceiver [3].

Transceiver designs with imperfect CSI have therefore attracted much attention in literature, including schemes based on sum-MMSE (SMMSE) [4] and maximin criteria [5]. These studies considered simple channel estimation error models, such as independent and identically distributed (i.i.d.) errors or magnitude-bounded errors. However, in practice, channel estimation errors are proven to be correlated; the correlation is induced by the channel correlation and channel estimation methods [6]. A more rigorous correlated multivariate Gaussian channel estimation error model was considered in [7].

Furthermore, if the channel is rapidly varying, the transmitter cannot obtain a reliable instantaneous channel estimate but only its statistical properties. Prior research have primarily focused on transceiver design by indirectly optimizing proxy metrics, namely mean-squared error (MSE) [8], [9] or signal-to-noise ratio (SNR). In this letter, we propose a precoder design, based on statistical CSI, for MIMO systems with imperfect CSI. Our objective is to *directly* minimize the error probability of the detected signal vector when the optimum maximum-likelihood (ML) detector is deployed at the receiver.

The remainder of this letter is organized as follows. We describe the MIMO system with channel estimation errors in Section II and propose the precoder design in Section III. Numerical results and discussions are given in Section IV. Finally, conclusions 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 $\widetilde{\mathbf{x}} = [\widetilde{x}_1, \ldots, \widetilde{x}_{N_t'}]^T$ denotes the transmitted signal vector; $\widetilde{y} = [\widetilde{y}_1, \ldots, \widetilde{y}_{N_r'}]^T$ denotes the received signal vector; $\widetilde{\mathbf{H}}$ is the Rayleigh fading channel matrix of dimension $N_r' \times N_t'$; and $\widetilde{\mathbf{n}} = [\widetilde{n}_1, \ldots, \widetilde{n}_{N_r'}]^T$ denotes a vector of Gaussian noise, i.e. $\widetilde{\mathbf{n}} \sim \mathcal{CN}(0, 2\sigma_n^2 \mathbf{I}_{N_r'})$. The complex system model (1) can be represented equivalently in the real domain as follows:

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

Let $\mathbf{y}$, $\mathbf{H}$, $\mathbf{x}$ and $\mathbf{n}$ denote the first, second, third and fourth terms of (2), respectively, we obtain the equivalent real system model, $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$. In this form, the dimensions of the system are doubled, i.e. $N_r = 2N_r'$ and $N_t = 2N_t'$. In this letter, we will work with the real domain representation.

Let $\mathbf{h} = \mathrm{vec}(\mathbf{H}^T)$, then $\mathbf{h}$ is a real Gaussian random vector with mean $\mu_{\mathbf{h}} = \mathbb{E}[\mathbf{h}] = \mathbf{0}$ and covariance matrix $\Sigma_{\mathbf{h}} = \mathbb{E}[\mathbf{h}\mathbf{h}^T]$, respectively. The MIMO system with channel estimation errors can be expressed as follows:

$$\mathbf{y} = \hat{\mathbf{H}}\mathbf{x} - \mathbf{E}\mathbf{x} + \mathbf{n}, \tag{3}$$

where $\mathbf{E}$ is the error matrix and $\hat{\mathbf{H}}$ is the channel estimate. In this letter, we assume that (i) both the transmitter and receiver have knowledge of $\Sigma_{\hat{\mathbf{h}}} \overset{\Delta}{=} \mathbb{E}[\mathrm{vec}(\hat{\mathbf{H}}^T)\mathrm{vec}(\hat{\mathbf{H}}^T)^T]$ and $\Sigma_{\mathbf{e}} \overset{\Delta}{=} \mathbb{E}[\mathrm{vec}(\mathbf{E}^T)\mathrm{vec}(\mathbf{E}^T)^T]$ [1]; and (ii) the receiver has an instantaneous channel estimate, $\hat{\mathbf{H}}$ and uses the optimum maximum-likelihood (ML) detector, given by [7]

$$\hat{\mathbf{x}}_{\mathrm{ml}} = \arg\min_{\mathbf{x} \in \mathcal{X}} \left(\log\det \Sigma_{\mathbf{x}} + (\mathbf{y} - \hat{\mathbf{H}}\mathbf{x})^T \Sigma_{\mathbf{x}}^{-1} (\mathbf{y} - \hat{\mathbf{H}}\mathbf{x}) \right), \tag{4}$$

where $\Sigma_{\mathbf{x}(i,j)} = \mathbf{x}^T \Sigma_{\mathbf{e}(i-1)N_t+1:iN_t,(j-1)N_t+1:jN_t} \mathbf{x} + \delta(i-j)\sigma_n^2$ (and δ is the delta-Dirac function) is the covariance matrix of $\mathbf{y}$, given that $\mathbf{x}$ is transmitted. In (4), $\mathcal{X}$ is the constellation set from which the transmitted signal vector $\mathbf{x}$ is drawn from.

Manuscript received July 23, 2014; revised November 4, 2014; accepted November 4, 2014. Date of publication November 7, 2014; date of current version February 18, 2015. The associate editor coordinating the review of this paper and approving it for publication was K. K. Wong.

The authors are with the Institute for Infocomm Research (I^{2}R), Singapore 138632 (e-mail: thianbs@i2r.a-star.edu.sg; nguyendh@i2r.a-star.edu.sg; sunsm@i2r.a-star.edu.sg).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LWC.2014.2368570

[1]For example, if Bayesian linear minimum mean-squared (LMMSE) estimators are used, the covariance matrices can be computed as given by [7].

III. Precoder Design With Statistical CSI

In this section, we propose a precoder design to minimize the worst-case codeword error probability when the optimum ML detector is used. The codeword error probability is defined as the probability that the detected signal vector is different from the transmitted signal vector.

A. Precoder as a Set of Functions

We propose a general design where different mappings are employed for different source messages. Let $\mathcal{V}$ be the source constellation set consisting of messages $\mathbf{v}^{(m)}, m = 1, \ldots, |\mathcal{V}|$. Consider $\overline{\mathcal{S}} = \{\mathcal{S}_{I(m)} : 1 \le I(m) \le N_{\mathrm{pf}}\}$ to be a partition of $\mathcal{V}$, where $\mathcal{S}_{I(m)}$ is associated with a precoding function $\mathbf{g}_{I(m)}$; and $I(m)$ denotes the index of the subset to which $\mathbf{v}^{(m)}$ belongs. The coded vector $\mathbf{x}^{(m)}$ is given by

$$\mathbf{x}^{(m)} = \mathbf{g}_{I(m)}(\mathbf{v}^{(m)}) = \mathbf{T}^{(I(m))}\mathbf{v}^{(m)}, \tag{5}$$

where we have adopted a linear functional form for $\mathbf{g}_{I(m)}$ and $\mathbf{T}^{(I(m))}$ is a $N_t \times N_t$ real matrix. Hence, the precoder design problem can be equivalently described as the design of the set of matrices $\{\mathbf{T}^{(I(m))} : 1 \le I(m) \le N_{\mathrm{pf}}\}$. Note that $N_{\mathrm{pf}} = |\mathcal{V}|$ corresponds to the case when every source message vector is assigned to a different precoding function and $N_{\mathrm{pf}} = 1$ corresponds to the case when all the source message vectors are assigned to one precoding function.

B. Upper Bound on Pairwise Error Probability

Suppose that $\mathbf{x}^{(m)}$ is transmitted. An upper bound on the average probability that the optimum detector (4) decides in favor of $\mathbf{x}^{(n)}$ is given by [7]

$$\mathrm{P}\left(\hat{\mathbf{x}}=\mathbf{x}^{(n)}|\mathbf{x}^{(m)}\right)=\left\{\frac{\det\left[\Sigma_{\mathbf{x}^{(m)}}\right]^{\mu}\det\left[\Sigma_{\mathbf{x}^{(n)}}\right]^{1-\mu}}{\det\left[\mu\Sigma_{\mathbf{x}^{(m)}}+(1-\mu)\Sigma_{\mathbf{x}^{(n)}}\right]}\right\}^{1/2}\cdot \frac{1}{2\left\{\det\left[\mathbf{I}_{N_tN_r}+2\mathbf{L}(\mu)\Sigma_{\hat{\mathbf{h}}}\right]\right\}^{1/2}}, \tag{6}$$

$$\mathbf{L}(\mu) = \frac{\mu(1-\mu)}{2}\left(\mathbf{X}^{(m)}-\mathbf{X}^{(n)}\right)^T \left\{\mu\Sigma_{\mathbf{x}^{(m)}}+(1-\mu)\Sigma_{\mathbf{x}^{(n)}}\right\}^{-1}\left(\mathbf{X}^{(m)}-\mathbf{X}^{(n)}\right), \tag{7}$$

where $\mathbf{X}^{(m)} = \mathbf{I}_{N_r} \otimes \mathbf{x}^{(m)T}$; and $\otimes$ denotes the Kronecker product. While (6) is valid for $0 < \mu < 1$, we obtain a simple bound by setting $\mu = \frac{1}{2}$. To simplify (6), we approximate $\Sigma_{\hat{\mathbf{h}}}$ and $\Sigma_{\mathbf{e}}$ by considering their main $N_t \times N_t$ block diagonals. With this approximation, $\Sigma_{\mathbf{x}^{(m)}}$ is also a diagonal matrix, with $\Sigma_{\mathbf{x}^{(m)}}(i,i) = \mathbf{x}^{(m)T}\mathbf{A}^{(i)}\mathbf{x}^{(m)} + \sigma_n^2$, where $\mathbf{A}^{(i)} \triangleq \Sigma_{\mathbf{e}(i-1)N_t+1:iN_t,(i-1)N_t+1:iN_t}$ is the i^{th} main $N_t \times N_t$ block diagonal of $\Sigma_{\mathbf{e}}$. Substituting the approximated $\Sigma_{\mathbf{x}^{(m)}}$ into (6) and simplifying it, we have

$$\mathrm{P}\left(\hat{\mathbf{x}}=\mathbf{x}^{(n)}|\mathbf{x}^{(m)}\right) = \frac{\prod_{i=1}^{N_t}\prod_{j=m,n}\rho_{i,j}\left(\mathbf{A}^{(i)},\mathbf{x}^{(j)}\right)}{\prod_{i=1}^{N_t}\varphi_i\left(\mathbf{A}^{(i)},\mathbf{B}^{(i)},\mathbf{x}^{(m)},\mathbf{x}^{(n)}\right)}, \tag{8a}$$

$$\rho_{i,j}\left(\mathbf{A}^{(i)},\mathbf{x}^{(j)}\right) = \left(\mathbf{x}^{(j)T}\mathbf{A}^{(i)}\mathbf{x}^{(j)}+\sigma_n^2\right)^{\frac{1}{4}}, \tag{8b}$$

$$\varphi_i\left(\mathbf{A}^{(i)},\mathbf{B}^{(i)},\mathbf{x}^{(m)},\mathbf{x}^{(n)}\right) = \Big((\mathbf{x}^{(m)}-\mathbf{x}^{(n)})^T \mathbf{B}^{(i)}(\mathbf{x}^{(m)}-\mathbf{x}^{(n)}) + 2\mathbf{x}^{(m)T}\mathbf{A}^{(i)}\mathbf{x}^{(m)} + 2\mathbf{x}^{(n)T}\mathbf{A}^{(i)}\mathbf{x}^{(n)} + 2\sigma_n^2\Big)^{\frac{1}{2}}, \tag{8c}$$

where $\mathbf{B}^{(i)} \triangleq \Sigma_{\hat{\mathbf{h}}(i-1)N_t+1:iN_t,(i-1)N_t+1:iN_t}$ is the i^{th} main $N_t \times N_t$ block diagonal of $\Sigma_{\hat{\mathbf{h}}}$.

C. Problem Formulation

We propose to minimize the codeword error probability via minimizing the worst-case pairwise error probability. We make substitutions in (8a): Let $\mathbf{x}^{(m)} = \mathbf{T}^{(I(m))}\mathbf{v}^{(m)}$, $\mathbf{V}^{(m)} = \mathbf{I}_{N_r} \otimes \mathbf{v}^{(m)T}$ and $\mathbf{t}^{(I(m))} = \mathrm{vec}(\mathbf{T}^{(I(m))T})$, we have $\mathbf{T}^{(I(m))}\mathbf{v}^{(m)} = \mathbf{V}^{(m)}\mathbf{t}^{(I(m))}$ and $\mathbf{v}^{(m)T}\mathbf{T}^{(I(m))T}\mathbf{A}^{(i)}\mathbf{T}^{(I(m))}\mathbf{v}^{(m)} = \mathbf{t}^{(I(m))T}\mathbf{A}^{(i,\mathbf{V}^{(m)})}\mathbf{t}^{(I(m))}$, where $\mathbf{A}^{(i,\mathbf{V}^{(m)})} \triangleq \mathbf{V}^{(m)T}\mathbf{A}^{(i)}\mathbf{V}^{(m)}$. Similarly, we define $\mathbf{B}^{(i,\mathbf{V}^{(m)},\mathbf{V}^{(n)})} \triangleq \mathbf{V}^{(m)T}\mathbf{B}^{(i)}\mathbf{V}^{(n)}$. Taking logarithm of (8a),

$$\begin{aligned}
&\Lambda_{\left(\mathbf{v}^{(m)},\mathbf{v}^{(n)}\right)}\left(\mathbf{t}^{(I(m))},\mathbf{t}^{(I(n))}\right) \triangleq \log \mathrm{P}\left(\hat{\mathbf{x}}=\mathbf{x}^{(n)}|\mathbf{x}^{(m)}\right)\\
&= \frac{1}{4}\sum_{i=1}^{N_t}\sum_{j=m,n}\left[\log\left(\mathbf{t}^{(I(j))T}\mathbf{A}^{\left(i,\mathbf{V}^{(j)}\right)}\mathbf{t}^{(I(j))}+\sigma_n^2\right)\right.\\
&\quad - \frac{1}{2}\sum_{i=1}^{N_t}\log\Big(\mathbf{t}^{(I(m))T}\mathbf{B}^{\left(i,\mathbf{V}^{(m)},\mathbf{V}^{(m)}\right)}\mathbf{t}^{(I(m))}\\
&\quad - \mathbf{t}^{(I(m))T}\mathbf{B}^{\left(i,\mathbf{V}^{(m)},\mathbf{V}^{(n)}\right)}\mathbf{t}^{(I(n))}\\
&\quad - \mathbf{t}^{(I(n))T}\mathbf{B}^{\left(i,\mathbf{V}^{(n)},\mathbf{V}^{(m)}\right)}\mathbf{t}^{(I(m))}\\
&\quad + \mathbf{t}^{(I(n))T}\mathbf{B}^{\left(i,\mathbf{V}^{(n)},\mathbf{V}^{(n)}\right)}\mathbf{t}^{(I(n))} + 2\mathbf{t}^{(I(m))T}\mathbf{A}^{\left(i,\mathbf{V}^{(m)}\right)}\mathbf{t}^{(I(m))}\\
&\quad + 2\mathbf{t}^{(I(n))T}\mathbf{A}^{\left(i,\mathbf{V}^{(n)}\right)}\mathbf{t}^{(I(n))} + 2\sigma_n^2\Big).
\end{aligned} \tag{9}$$

Formulating our minimax problem in its epigraph form and using the fact that logarithm is a monotonic increasing function, we can write our problem as

$$\begin{aligned}
&\text{minimize} \quad \alpha\\
&\text{subject to} \quad \Lambda_{\left(\mathbf{v}^{(m)},\mathbf{v}^{(n)}\right)}\left(\mathbf{t}^{(I(m))},\mathbf{t}^{(I(n))}\right) \le \alpha,\\
&\qquad \forall\, \mathbf{v}^{(m)},\mathbf{v}^{(n)} \in \mathcal{V} : \mathbf{v}^{(m)} \neq \mathbf{v}^{(n)}\\
&\qquad \frac{1}{|\mathcal{V}|}\sum_{i=1}^{|\mathcal{V}|}\mathbf{t}^{(i)T}\mathbf{V}^{(i)T}\mathbf{V}^{(i)}\mathbf{t}^{(i)} \le \mathcal{E}_{\mathrm{avg}}
\end{aligned} \tag{10}$$

with variables $\alpha \in \mathbb{R}$ and $\{\mathbf{t}^{(i)} \in \mathbb{R}^{N_t^2} : 1 \le i \le N_{\mathrm{pf}}\}$. The second constraint is a mean power constraint and $\mathcal{E}_{\mathrm{avg}} = \frac{1}{|\mathcal{V}|}\sum_{i=1}^{|\mathcal{V}|}\mathbf{v}^{(i)T}\mathbf{v}^{(i)}$ is the mean power of the source vectors.

D. First-Order Approximation Approach

As (10) is non-convex, we propose an iterative approach, based on first-order approximation of $\Lambda_{\left(\mathbf{v}^{(m)},\mathbf{v}^{(n)}\right)}$, to find a good locally optimal solution of (10). In iteration l, we approximate the function $\Lambda_{\left(\mathbf{v}^{(m)},\mathbf{v}^{(n)}\right)}$ around its previous feasible point. Note that $\Lambda_{\left(\mathbf{v}^{(m)},\mathbf{v}^{(n)}\right)}$ is a function of two variables if $I(m) \neq I(n)$ and a function of one variable if $I(m) = I(n)$. Let $\mathbf{u} \triangleq \mathbf{v}^{(m)}$, $\mathbf{w} \triangleq \mathbf{v}^{(n)}$, $\theta \triangleq \mathbf{t}^{(I(m))}$, $\phi \triangleq \mathbf{t}^{(I(n))}$. $\Lambda_{(\mathbf{u},\mathbf{w})}(\theta,\phi)$ can be linearized at a point $(\tilde{\theta},\tilde{\phi})$ as

$$\Lambda_{(\mathbf{u},\mathbf{w})}(\theta,\phi) \approx \Lambda_{(\mathbf{u},\mathbf{w})}(\tilde{\theta},\tilde{\phi}) + \nabla_{(\theta,\phi)}\Lambda_{(\mathbf{u},\mathbf{w})}(\theta,\phi)^T\left((\theta,\phi)-(\tilde{\theta},\tilde{\phi})\right), \tag{11}$$

where $\nabla_{(\theta,\phi)}\Lambda_{(\mathbf{u},\mathbf{w})}(\theta,\phi)$ is the gradient function of $\Lambda_{(\mathbf{u},\mathbf{w})}(\theta,\phi)$. For the case when $I(m)\neq I(n)$ (or equivalently $\theta\neq\phi$), the gradient function is given by

$$\nabla_{(\theta,\phi)}\Lambda_{(\mathbf{u},\mathbf{w})}(\theta,\phi)=\left[\nabla_{\theta}\Lambda_{(\mathbf{u},\mathbf{w})}(\theta,\phi),\nabla_{\phi}\Lambda_{(\mathbf{u},\mathbf{w})}(\theta,\phi)\right]^T, \tag{12a}$$

$$\nabla_{\theta}\Lambda_{(\mathbf{u},\mathbf{w})}(\theta,\phi)=\sum_{i=1}^{N_t}\left[\frac{1}{4}\frac{\frac{d\zeta_i^{(1)}}{d\theta}}{\zeta_i^{(1)}(\theta)}-\frac{1}{2}\frac{\frac{d\kappa_i}{d\theta}}{\kappa_i(\theta,\phi)}\right], \tag{12b}$$

$$\nabla_{\phi}\Lambda_{(\mathbf{u},\mathbf{w})}(\theta,\phi)=\sum_{i=1}^{N_t}\left[\frac{1}{4}\frac{\frac{d\zeta_i^{(2)}}{d\phi}}{\zeta_i^{(2)}(\phi)}-\frac{1}{2}\frac{\frac{d\kappa_i}{d\phi}}{\kappa_i(\theta,\phi)}\right], \tag{12c}$$

$$\zeta_i^{(1)}(\theta)=\theta^T\mathbf{A}^{(i,\mathbf{U})}\theta+\sigma_n^2, \tag{12d}$$

$$\zeta_i^{(2)}(\phi)=\phi^T\mathbf{A}^{(i,\mathbf{W})}\phi+\sigma_n^2, \tag{12e}$$

$$\begin{aligned}\kappa_i(\theta,\phi)=&\,\theta^T\mathbf{B}^{(i,\mathbf{U},\mathbf{U})}\theta-\theta^T\mathbf{B}^{(i,\mathbf{U},\mathbf{W})}\phi\\&-\phi^T\mathbf{B}^{(i,\mathbf{W},\mathbf{U})}\theta\\&+\phi^T\mathbf{B}^{(i,\mathbf{W},\mathbf{W})}\phi+2\theta^T\mathbf{A}^{(i,\mathbf{U})}\theta\\&+2\phi^T\mathbf{A}^{(i,\mathbf{W})}\phi+2\sigma_n^2,\end{aligned} \tag{12f}$$

where $\mathbf{A}^{(i,\mathbf{U})}\triangleq\mathbf{U}^T\mathbf{A}^{(i)}\mathbf{U}$ (where $\mathbf{U}\triangleq\mathbf{I}_{N_r}\otimes\mathbf{u}^T$) and $\mathbf{B}^{(i,\mathbf{W},\mathbf{U})}\triangleq\mathbf{W}^T\mathbf{B}^{(i)}\mathbf{U}$. For the case when $I(m)=I(n)$ (or equivalently $\theta=\phi$), the gradient function is given by

$$\nabla_{\theta}\Lambda_{(\mathbf{u},\mathbf{w})}(\theta)=\sum_{i=1}^{N_t}\left[\frac{1}{4}\frac{\frac{d\zeta_i^{(1)}}{d\theta}}{\zeta_i^{(1)}(\theta)}+\frac{1}{4}\frac{\frac{d\zeta_i^{(2)}}{d\phi}}{\zeta_i^{(2)}(\phi)}-\frac{1}{2}\frac{\frac{d\kappa_i}{d\theta}}{\kappa_i(\theta,\phi)}\right]. \tag{13}$$

Using (11)–(13), the problem in the l^{th} iteration is

$$\begin{aligned}\text{minimize}\quad&\alpha\\\text{subject to}\quad&\Lambda_{(\mathbf{v}^{(m)},\mathbf{v}^{(n)})}\left(\mathbf{t}_{[l-1]}^{(I(m))},\mathbf{t}_{[l-1]}^{(I(n))}\right)\\&+\left[\nabla_{\left(\mathbf{t}_{[l]}^{(I(m))},\mathbf{t}_{[l]}^{(I(n))}\right)}\Lambda_{(\mathbf{v}^{(m)},\mathbf{v}^{(n)})}\right]^T\\&\left(\left(\mathbf{t}_{[l]}^{(I(m))},\mathbf{t}_{[l]}^{(I(n))}\right)-\left(\mathbf{t}_{[l-1]}^{(I(m))},\mathbf{t}_{[l-1]}^{(I(n))}\right)\right)\le\alpha,\\&\forall\,\mathbf{v}^{(m)},\mathbf{v}^{(n)}\in\mathcal{V}:\mathbf{v}^{(m)}\neq\mathbf{v}^{(n)},\,I(m)\neq I(n)\\&\Lambda_{(\mathbf{v}^{(m)},\mathbf{v}^{(n)})}\left(\mathbf{t}_{[l-1]}^{(I(m))}\right)\\&+\left[\nabla_{\left(\mathbf{t}_{[l]}^{(I(m))}\right)}\Lambda_{(\mathbf{v}^{(m)},\mathbf{v}^{(n)})}\right]^T\\&\left(\mathbf{t}_{[l]}^{(I(m))}-\mathbf{t}_{[l-1]}^{(I(m))}\right)\le\alpha,\\&\forall\mathbf{v}^{(m)},\mathbf{v}^{(n)}\in\mathcal{V}:\mathbf{v}^{(m)}\neq\mathbf{v}^{(n)},\,I(m)=I(n)\\&\mathbf{t}_{[l-1]}^{(i)}-\varepsilon\le\mathbf{t}_{[l]}^{(i)}\le\mathbf{t}_{[l-1]}^{(i)}+\varepsilon,\ 1\le i\le N_{\text{pf}}\\&\frac{1}{|\mathcal{V}|}\sum_{i=1}^{|\mathcal{V}|}\mathbf{t}_{[l]}^{(I(i))T}\mathbf{V}^{(i)T}\mathbf{V}^{(i)}\mathbf{t}_{[l]}^{(I(i))}\le\mathcal{E}_{\text{avg}}\end{aligned} \tag{14}$$

with variables $\alpha\in\mathbb{R}$ and $\{\mathbf{t}_{[l]}^{(i)}\in\mathbb{R}^{N_t^2}:1\le i\le N_{\text{pf}}\}$. In (14), $\{\mathbf{t}_{[l-1]}^{(i)}\in\mathbb{R}^{N_t^2}:1\le i\le N_{\text{pf}}\}$ is the solution obtained in the $(l-1)^{\text{th}}$ iteration. The third constraint in (14) is the trust region constraint with radius ε, indicating the region in which the linearization (11) is accurate. We summarize our proposed approach in Algorithm I below.

Algorithmic 1 Proposed Precoder Design for MIMO systems with Channel Estimation Errors

---

1: Input: $N_r,N_t,N_{\text{pf}},\varepsilon$, $\{\mathbf{t}_{[0]}^{(i)}\in\mathbb{R}^{N_t^2}:1\le i\le N_{\text{pf}}\}$, maximum number of iterations $l_{\max}$, tolerance η.
2: Assign the source message vectors into N_{pf} subsets (using a pre-specified criteria).
3: **for** $l=1$ to $l_{\max}$ **do**
4: **for** $i=1$ to $|\mathcal{V}|$ and $j=i+1$ to $|\mathcal{V}|$ **do**
5: Compute $\nabla_{(\mathbf{t}^{(I(i))},\mathbf{t}^{(I(j))})}\Lambda_{(\mathbf{v}^{(i)},\mathbf{v}^{(j)})}\left(\mathbf{t}_{[l]}^{(I(i))},\mathbf{t}_{[l]}^{(I(j))}\right)$ using (12a)–(12f), if $I(i)\neq I(j)$.
6: Compute $\nabla_{(\mathbf{t}^{(I(i))})}\Lambda_{(\mathbf{v}^{(i)},\mathbf{v}^{(j)})}\left(\mathbf{t}_{[l]}^{(I(j))}\right)$ using (13), if $I(i)=I(j)$.
7: **end for**
8: Use gradient functions (in lines 5–6), and previous estimate $\left\{\mathbf{t}_{[l-1]}^{(i)}\in\mathbb{R}^{N_t^2}:1\le i\le N_{\text{pf}}\right\}$, to solve (14).
9: Exit the loop if $\sum_{i=1}^{N_{\text{pf}}}\|\mathbf{t}_{[l]}^{(i)}-\mathbf{t}_{[l-1]}^{(i)}\|^2\le\eta$.
10: **end for**
11: Output: $\left\{\mathbf{t}_{[l]}^{(i)}\in\mathbb{R}^{N_t^2}:1\le i\le|N_{\text{pf}}|\right\}$.

E. Computational Complexity

Problem (14) is a linear program with well known methods of solving it. The Karmarkar's algorithm [10] has been proven to be more efficient than other methods, with its complexity given by [10] $O(N_{\text{pf}}^{3.5}N_t^7L^2\cdot\log L\cdot\log\log L)$, where L is the number of input bits.

IV. Numerical Results and Discussion

In this section, we evaluate the performance of our proposed precoder design for different system parameters. We use two models as the channel correlation matrix: (i) For correlated channels, we set $\Sigma_{\mathbf{h}(i,j)}=c\beta^{|i-j|}$; (ii) for uncorrelated channels, we set $\Sigma_{\mathbf{h}(i,i)}=ci$. These correlation matrix models are frequently used in the literature (see [6] and references therein). In the simulations, the definition of average signal-to-interference-plus-noise ratio is $\overline{\text{SINR}}=\frac{1}{N_r}\sum_{i=1}^{N_r}\text{SINR}[i]$, where $\text{SINR}[i]=\frac{\sum_{j=1}^{N_t}\frac{\mathcal{E}_{\text{avg}}}{N_t}\mathbb{E}[\mathbf{H}_{i,j}^2]}{\sum_{j=1}^{N_t}\frac{\mathcal{E}_{\text{avg}}}{N_t}\mathbb{E}[\mathbf{E}_{i,j}^2]+\sigma_n^2}$ [7]. We compare our algorithm with those proposed in [8] and [9].

It is worth noting that the schemes given in [8] and [9] require instantaneous channel estimate. For frequency division duplex (FDD) systems where the uplink and downlink channels are not reciprocal, the channel estimates need to be transmitted back from the receiver to the transmitter. It is therefore a much more complex system than ours, which utilizes only the channel statistical information and does not need channel feedback. Both [8] and [9] use iterative algorithms to design their precoders. Though the computational complexity per iteration of their algorithms, given by $O(N_r^3L\cdot\log L\cdot\log\log L)$, is lower

T

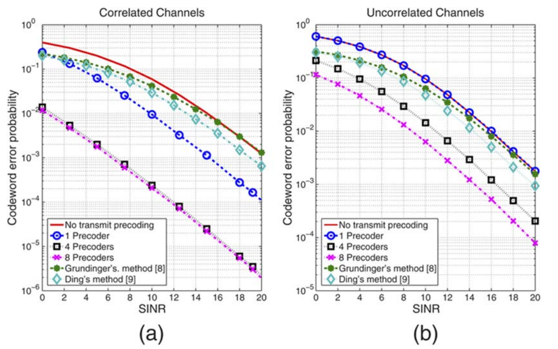

Fig. 1. Comparisons of the codeword error rate of 4 × 4 real MIMO system with different number of precoder functions.

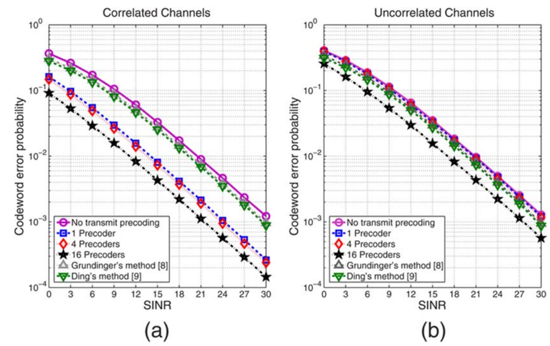

Fig. 2. Comparisons of the codeword error rate of 2 × 2 real MIMO system with different number of precoder functions.

than ours, but our proposed scheme is much more advantageous where system-level complexity is concerned.

A. Error Probability Performance of 4 × 4 MIMO System

In the setup, we consider a 4 × 4 real MIMO system with BPSK-modulation per transmit antenna, and ML detection at the receiver. We illustrate the codeword error probability versus SINR performances for our proposed precoder (for $N_{pf} = 1$, 4, 8), Grundinger's [8] and Ding's [9] methods in Fig. 1. As observed, the systems without transmit precoding have the worst error probability performance. Interestingly, it is noted that correlated channels benefit more from our proposed precoder design than uncorrelated channels. It is also observed from Fig. 1(a) that at a codeword error rate of 10^{-3}, our proposed design with just 1 common precoder achieves a coding gain of 5 dB over the system without transmit precoding. The codeword error probability performance can be further improved via the use of more precoders. At a codeword error rate of 10^{-3}, our proposed design using 8 precoders achieves a coding gain of more than 12 dB gain over the system without transmit precoding.

It is also noted that the schemes in [8] and [9] have only a slight improved error performance over the systems without transmit precoding; the improvement is small as SINR increases. The error performance of their methods is still outperformed by our proposed precoder. This is because the schemes in [8] and [9] did not utilize the full information from the channel and error correlation matrix.

B. Error Probability Performance of 2 × 2 MIMO System

In the second setup, we consider a 2 × 2 real MIMO system with 8-PAM modulation per transmit antenna, and ML detection at the receiver. We illustrate the codeword error probability versus SINR performances for our proposed precoder (for $N_{pf} =$ 1, 4, 16), Grundinger's [8] and Ding's [9] methods in Fig. 2. The same relative performance is observed; correlated channels achieved more significant gains with our proposed design than uncorrelated channels. However, different from Fig. 1, the gain is realized only with large number of precoder functions, i.e., $N_{pf} = 16$ for uncorrelated channels. It is observed from Fig. 2(a) that our proposed design with 1 precoder achieves significant improvement in performance. At a codeword error rate of 10^{-3}, the proposed design using 1 precoder attains a coding gain of 5.5 dB over the system without transmit precoding. A significant improvement can be gained by using 16 precoders. At a codeword error rate of 10^{-3}, our proposed design using 16 precoders achieves a coding gain of more than 8 dB gain over the system without the transmit precoding.

Finally, we note that the schemes in [8] and [9] have negligible improvement in error performance over the systems without transmit precoding.

V. Conclusion

In this letter, we have proposed a precoder design to minimize codeword error rate in practical MIMO systems with imperfect CSI. We assume that the receiver knows the instantaneous estimate and the statistics of the channel, while the transmitter has the knowledge of its statistics only. Our design thus does not require instantaneous channel feedback, which leads to considerable savings in resources. We show, via numerical simulations, that our proposed design achieves a significant improvement in error rate performance.

References

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

[2] 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. 744–765, Mar. 1998.

[3] A. J. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 684–702, Jun. 2003.

[4] M. Ding and S. D. Blostein, "Maximum mutual information design for MIMO systems with imperfect channel knowledge," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 4793–4801, Oct. 2010.

[5] A. P-Iserte, D. P. Palomar, A. I. P-Neira, and M. A. Lagunas, "A robust maximum approach 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.

[6] J. Ringelstein, A. B. Gershman, and J. F. Bohme, "Direction finding in random inhomogeneous media in the presence of multiplicative noise fields," *IEEE Signal Process. Lett.*, vol. 7, no. 10, pp. 269–272, Oct. 2000.

[7] B. S. Thian and A. J. Goldsmith, "Reduced-complexity robust MIMO decoders," *IEEE Trans. Wireless Commun.*, vol. 12, no. 8, pp. 3783–3795, Aug. 2013.

[8] A. Grundinger, M. Joham, and W. Utstick, "Stochastic transceiver design in point-to-point MIMO channels with imperfect CSI," in *Proc. Int. ITG workshop Smart Antennas*, Feb. 2011, pp. 1–8.

[9] M. Ding and S. D. Blostein, "MIMO minimum total MSE transceiver design with imperfect CSI at both ends," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 1141–1150, Mar. 2009.

[10] N. Karmarkar, "A new polynomial time algorithm for linear programming," *Combinatorica*, vol. 4, no. 4, pp. 373–395, Apr. 1984.