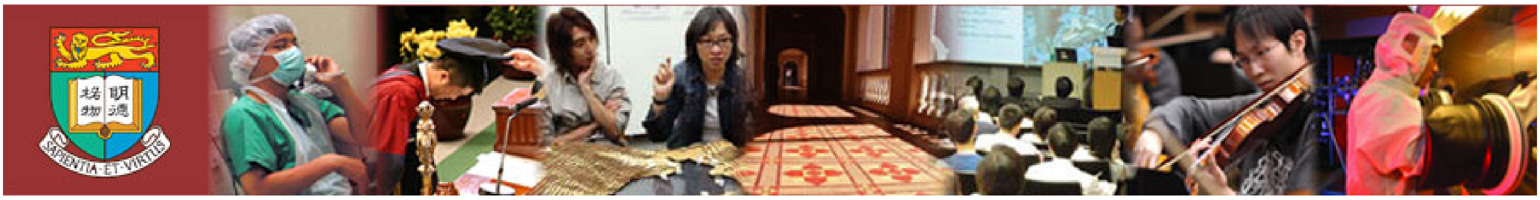

Title	QoS constrained robust MIMO transceiver design under unknown interference
Author(s)	He, X; Chen, J; Wu, YC
Citation	The 7th International ICST Conference on Communications and Networking in China (CHINACOM 2012), Kunming, China, 7-10 August 2012. In Conference Proceedings, 2012, p. 309-314
Issued Date	2012
URL	http://hdl.handle.net/10722/189821
Rights	

QoS Constrained Robust MIMO Transceiver Design Under Unknown Interference

Xin He, Jianwu Chen, and Yik-Chung Wu
Department of Electrical and Electronic Engineering
The University of Hong Kong, Hong Kong
Email: {hexin, jwchen, ycwu}@eee.hku.hk

***Abstract*—We study the robust transceiver optimization in multiple-input multiple-output (MIMO) systems aiming at minimizing transmit power under probabilistic quality-of-service (QoS) requirements. Owing to the unknown distributed interference, the channel estimation error can be arbitrary distributed. Under this situation, the QoS requirements should account for the worst-case channel estimation error distribution. While directly finding the worst-case distribution is challenging, two methods are proposed to solve the robust transceiver design problem. One is based on the Chebyshev inequality, the other is based on a novel duality method. Simulation results show that the QoS requirement is satisfied by both proposed algorithms. Furthermore, among the two proposed methods, the duality method shows a superior performance in transmit power, while the Chebyshev method demonstrates a lower computational complexity.**

***Index Terms*—QoS, Robust MIMO transceiver design, Arbitrary distributed uncertainty, Semidefinite relaxation (SDR).**

I. INTRODUCTION

Traditionally, one popular criterion for MIMO transceiver design is to minimize the mean squre error (MSE) of data subject to power constraint [1], [2]. However, in modern wireless communications, different applications might have different quality of service (QoS) requirements, making purely minimizing MSE not an ideal criterion. Therefore, QoS based transceiver designs received a lot attention recently [3], [4].

Early studies on QoS based transceiver designs [3], [4] assume perfect channel state information (CSI) is known. However, in practice, CSI has to to be estimated and estimation error is unavoidable. Therefore, robust transceiver design, which takes the channel estimation uncertainty into consideration, is important. In [5] and [6], bounded channel estimation error is introduced in the transceiver design problem. Nevertheless, in general, the distribution of estimation error of a random variable is unbounded [7]. To tackle this problem, probabilistic QoS constrained robust transceiver design with Gaussian distributed channel uncertainty is proposed in [8]. However, in the nowadays crowed wireless environment, owing to the interference from unintended users (e.g., co-channel interference in a cellular system or interference from secondary user in cognitive radio systems), the distribution of the interference plus noise might be unknown. Then, the distribution of the channel estimation error under interference cannot be modeled in prior. Therefore, in this paper, we investigate the robust transceiver design problem in MIMO system under probabilistic QoS constraints with arbitrary distributed channel estimation error. Since the distribution of the channel estimation error is unknown, it is difficult to get an appropriate error bound [9]. Furthermore, previous methods assuming a specific distributed channel estimation error are not suitable for solving this problem.

In this paper, we formulate the probabilistic QoS constraints under arbitrary distributed channel estimation error into worst-case probabilistic constraints, and two methods are proposed to tackle this problem. One is based on the Chebyshev's inequality, which provides a upper bound for the worst-case probability. Then an iteration between two convex subproblems is employed to solve the problem. The other is based on a novel duality method, in which the worst-case probability problem is transformed into a deterministic finite constrained problem by using the Lagrange duality and S-Lemma, with strong duality guaranteed. The resulting problem is solved by a convergence-guaranteed iterative algorithm between two subproblems. Simulation results show that the duality method has an excellent performance on the transmit power with QoS guaranteed, while the Chebyshev method exhibits low computational complexity.

The rest of this paper is organized as follows. In Section II, the system model is introduced. In Section III, the robust transceiver design problem is formulated and two different methods are proposed to solve it. Simulation results are presented in Section IV, and conclusions are drawn in Section V.

Notation: In this paper, $\mathbb{E}(\cdot)$, $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^H$ denote statistical expectation, conjugation, transposition and Hermitian, respectively, while $\|\cdot\|_2$ denotes the norm of a vector. In addition, $\mathrm{Tr}(\cdot)$ and $\|\cdot\|_F$ refer to the trace and Frobenius norm of a matrix, respectively. The notations $\mathrm{vec}(\cdot)$ and $\otimes$ stands for the vectorization and Kronecker product, respectively. Symbol $\mathrm{diag}(\mathbf{x})$ denotes a diagonal matrix with vector $\mathbf{x}$ on its diagonal. Finally, $\mathbf{I}_K$ is a $K \times K$ identity matrix.

II. SYSTEM MODEL

The MIMO system under consideration consists of one transmitter equipped with N antennas, and one receiver equipped with M antennas. It is assumed that L independent data streams are transmitted to the receiver. In order to guarantee data recovery at the receiver, it is required that $L \leq \min\{M, N\}$. Let $\mathbf{s}$ be the $L \times 1$ data vector transmitted to the receiver, we have $\mathbb{E}(\mathbf{s}\mathbf{s}^H) = \mathbf{I}_L$. Let $\mathbf{G}$ be the $N \times L$

precoding matrix at the transmitter, then the received $M \times 1$ signal at the receiver is

$$\mathbf{y} = \mathbf{HGs} + \mathbf{n}, \tag{1}$$

where $\mathbf{H}$ and $\mathbf{n}$ are the $M \times N$ channel matrix and the received $M \times 1$ interference plus noise vector, respectively. It is assumed that the interference plus noise can be arbitrary distributed with its first two moments being known. Without loss of generality, we assume $\mathbb{E}(\mathbf{n}) = \mathbf{0}$ and $\mathbb{E}(\mathbf{nn}^H) = \mathbf{R}$, and we write $\mathbf{n} \sim \mathcal{A}(\mathbf{0}, \mathbf{R})$, where $\mathcal{A}$ denotes an arbitrary distribution.

At the receiver, an $L \times M$ equalizer $\mathbf{F}$ is used to process the received signal. The recovered $L \times 1$ data vector is

$$\hat{\mathbf{s}} = \mathbf{FHGs} + \mathbf{Fn}. \tag{2}$$

Since the transmitted data are independent with the interference and noise, the total data MSE can be calculated as

$$\begin{aligned} \mathrm{MSE} &= \mathbb{E}_{\mathbf{s},\mathbf{n}}\left[\mathrm{Tr}\{(\mathbf{s}-\hat{\mathbf{s}})(\mathbf{s}-\hat{\mathbf{s}})^H\}\right] & (3)\\ &= \|\mathbf{FHG}-\mathbf{I}\|_F^2 + \mathrm{Tr}(\mathbf{FRF}^H), & (4) \end{aligned}$$

It is obvious that, in addition to the precoder $\mathbf{G}$ and equalizer $\mathbf{F}$, the MSE also depends on the channel realization $\mathbf{H}$, which has to be estimated in practice.

In general, the channel estimation error can be modeled as

$$\mathbf{H} = \hat{\mathbf{H}} + \mathbf{\Delta}, \tag{5}$$

where $\hat{\mathbf{H}}$ and $\mathbf{\Delta}$ are the estimated channel and channel estimation error, respectively. Since the distribution of $\mathbf{\Delta}$ depends on the unknown the interference plus noise, we can only model the channel estimation uncertainty $\mathbf{\Delta}$ as arbitrary distributed with its first two moments known, i.e., $\mathrm{vec}(\mathbf{\Delta}) \sim \mathcal{A}(\mathbf{0}, \mathbf{\Sigma})$ [10].

After substituting (5) into (4), the total MSE is

$$\mathrm{MSE}(\mathbf{\Delta}) = \|\mathbf{F}(\hat{\mathbf{H}}+\mathbf{\Delta})\mathbf{G}-\mathbf{I}\|_F^2 + \mathrm{Tr}(\mathbf{FRF}^H). \tag{6}$$

Note that the distribution of the MSE depends on that of the $\mathbf{\Delta}$. Furthermore, $\mathbf{G}$ and $\mathbf{F}$ are unknown, and in general depends on the statistical information of $\mathbf{\Delta}$. Therefore, the distribution of the MSE cannot be obtained or approximated in advance.

III. Robust MIMO Transceiver Design

In this paper, we impose a probabilistic QoS constraint in the data MSE in the form of $\mathrm{Prob}\{\mathrm{MSE}(\mathbf{\Delta}) \geq \varepsilon\} \leq p$. After the QoS requirement is satisfied, it is crucial to save transmission power at the transmitter. Therefore, the probabilistic QoS constrained transceiver design problem can be formulated as

$$\begin{aligned} \min_{\mathbf{G},\mathbf{F}} \quad & \mathrm{Tr}(\mathbf{GG}^H) \\ \text{s.t.} \quad & \sup_{\mathrm{vec}(\mathbf{\Delta})\sim\mathcal{A}(\mathbf{0},\mathbf{\Sigma})} \mathrm{Prob}\{\mathrm{MSE}(\mathbf{\Delta}) \geq \varepsilon\} \leq p. \end{aligned} \tag{$\mathbb{P}0$}$$

In the QoS constraint, due to the arbitrary distributed channel estimation errors, the supremum of the outage probability is used to guarantee the QoS performance is satisfied even at the worst case.

Note that the problem ($\mathbb{P}0$) is a bilevel optimization problem, and the lower level problem involves finding the supremum of the outage probability. Owing to the unknown distribution of $\mathbf{\Delta}$, it is difficult to get an analytical solution for the lower level problem. Below, we consider two methods to solve this problem. The first one is to use the Chebyshev upper bound to eliminate the supremum. Another method is the proposed duality method, in which the lower level problem is transformed from a stochastic problem into a solvable deterministic problem.

A. The Chebyshev Inequality Based Method

Since the distribution of the MSE is unknown, the one-side Chebyshev inequality [11] can be used to get an upper bound of the outage probability

$$\mathrm{Prob}\{\mathrm{MSE}(\mathbf{\Delta}) \geq \varepsilon\} \leq \frac{\mathbb{E}_{\mathbf{\Delta}}\{\mathrm{MSE}(\mathbf{\Delta})\}}{\varepsilon}, \tag{7}$$

where $\mathbb{E}_{\mathbf{\Delta}}\{\mathrm{MSE}(\mathbf{\Delta})\}$ can be derived as

$$\mathbb{E}_{\mathbf{\Delta}}\{\mathrm{MSE}\} = \mathbb{E}_{\mathbf{\Delta}}\{\|\mathbf{F\Delta G}\|_F^2\} + \|\mathbf{F}\hat{\mathbf{H}}\mathbf{G}-\mathbf{I}\|_F^2 + \|\mathbf{R}^{1/2}\mathbf{F}^H\|_F^2 \tag{8}$$

$$= \|[\mathrm{vec}\left(\mathbf{\Sigma}^{\frac{1}{2}}(\mathbf{G}^T\otimes\mathbf{F})\right)^T \mathrm{vec}(\mathbf{F}\hat{\mathbf{H}}\mathbf{G}-\mathbf{I})^T \mathrm{vec}(\mathbf{R}^{1/2}\mathbf{F}^H)^T]\|_2^2. \tag{9}$$

Since the right hand side of (7) is independent of the exact distribution, the problem ($\mathbb{P}0$) can be approximated as

$$\begin{aligned} \min_{\mathbf{G},\mathbf{F}} \quad & \mathrm{Tr}(\mathbf{GG}^H) \\ \text{s.t.} \quad & \|[\mathrm{vec}\left(\mathbf{\Sigma}^{\frac{1}{2}}(\mathbf{G}^T\otimes\mathbf{F})\right)^T \mathrm{vec}(\mathbf{F}\hat{\mathbf{H}}\mathbf{G}-\mathbf{I})^T \\ & \quad \mathrm{vec}(\mathbf{R}^{1/2}\mathbf{F}^H)^T]\|_2^2 \leq p\varepsilon. \end{aligned} \tag{10}$$

Note that this problem becomes a convex problem of $\mathbf{G}$ when the equalizer $\mathbf{F}$ is fixed. Furthermore, when the precoder $\mathbf{G}$ is fixed, the objective function is not related to the equalizer $\mathbf{F}$. In order to provide a larger feasible space for the next round precoder design, the equalizer $\mathbf{F}$ can be used to minimize the left side of the QoS constraints in (10). Therefore, the problem (10) can be solved by iterations between two convex subproblems as follows.

More specifically, in the first subproblem, with the equalizer $\mathbf{F}$ fixed, the optimum precoder $\mathbf{G}$ can be obtained from the second-order cone programming (SOCP) problem

$$\begin{aligned} \min_{\mathbf{G},P} \quad & P \\ \text{s.t.} \quad & \|\mathrm{vec}(\mathbf{G})\|_2 \leq P \\ & \|[\mathrm{vec}\left(\mathbf{\Sigma}^{\frac{1}{2}}(\mathbf{G}^T\otimes\mathbf{F})\right)^T \mathrm{vec}(\mathbf{F}\hat{\mathbf{H}}\mathbf{G}-\mathbf{I})^T \\ & \quad \mathrm{vec}(\mathbf{R}^{1/2}\mathbf{F}^H)^T]\|_2 \leq \sqrt{p\varepsilon}, \end{aligned} \tag{11}$$

where P is a slack variable.

In the second subproblem, with the precoder $\mathbf{G}$ fixed, the equalizer $\mathbf{F}$ can be used to minimize the left side of the QoS constraint. Expressing the left side of the QoS constraint in (10) with the Frobenius norm, the equalizer $\mathbf{F}$ can be updated from the following problem

$$\min_{\mathbf{F}} \quad \|\mathbf{\Sigma}^{\frac{1}{2}}(\mathbf{G}^T\otimes\mathbf{F})\|_F^2 + \|\mathbf{F}\hat{\mathbf{H}}\mathbf{G}-\mathbf{I}\|_F^2 + \|\mathbf{R}^{1/2}\mathbf{F}^H\|_F^2. \tag{12}$$

Furthermore, writing the first term of the cost function (12) as

$$\begin{aligned} \|\mathbf{\Sigma}^{\frac{1}{2}}(\mathbf{G}^T\otimes\mathbf{F})\|_F^2 &= \mathrm{Tr}\left(\mathbf{\Sigma}\left((\mathbf{G}^*\mathbf{G}^T)\otimes(\mathbf{F}^H\mathbf{F})\right)\right) & (13)\\ &= \mathrm{Tr}(\sum_{j=1}^N\sum_{i=1}^N g_{ij}\mathbf{\Sigma}^{ji}\mathbf{F}^H\mathbf{F}), & (14) \end{aligned}$$

where g_{ij} is the $(i,j)^{\text{th}}$ element of the matrix $\mathbf{G}^*\mathbf{G}^T$, $\mathbf{\Sigma}^{ji}$ is the $(i,j)^{\text{th}}$ $M \times M$ subblock of the matrix $\mathbf{\Sigma}$. Putting (14) into (12), and setting the derivative of the cost function (12) with respect to $\mathbf{F}^*$ to zero, the optimum equalizer $\mathbf{F}$ is

$$\mathbf{F} = (\hat{\mathbf{H}}\mathbf{G})^H(\hat{\mathbf{H}}\mathbf{G}\mathbf{G}^H\hat{\mathbf{H}}^H + \mathbf{R} + \sum_{j=1}^{N}\sum_{i=1}^{N} g_{ij}\mathbf{\Sigma}^{ji})^{-1}. \quad (15)$$

It is observed that the obtained equalizer is a conventional MMSE equalizer with additional regularization from the weighted uncertainty $\sum_{j=1}^{N}\sum_{i=1}^{N} g_{ij}\mathbf{\Sigma}^{ji}$.

The iterative algorithm between the two subproblems is summarized at Table I. Note that with a feasible initialization, the transmit power in each iteration obtained by the iterative algorithm decreases monotonically and therefore the proposed algorithm converges.

With regard to the initialization, it is a common problem for the QoS based MIMO transceiver design since a feasible initial transceiver pair is required [4] [6]. Conventionally, the receiver $\mathbf{F}$ is initialized with an identity or a randomly generated matrix. However, these initializations are not guaranteed to satisfy the QoS constraints. It is observed in (11) that if the elements of $\mathbf{F}$ are small, we have a better chance of satisfying the QoS constraints. Based on this observation, it is suggested in [6] that scaling factors $1/\gamma$ are introduced into the initial chosen equalizer in (11), and the scaling factors can be obtained from the following SOCP problem

$$\begin{aligned} \min_{\mathbf{G},P,\gamma} \quad & P \\ \text{s.t.} \quad & \|\text{vec}(\mathbf{G})\|_2 \le P \\ & \|[\text{vec}\Big(\mathbf{\Sigma}^{\frac{1}{2}}(\mathbf{G}^T \otimes \mathbf{F}_o)\Big)^T \text{vec}(\mathbf{F}_o\hat{\mathbf{H}}\mathbf{G} - \gamma\mathbf{I})^T \\ & \qquad \text{vec}(\mathbf{R}^{1/2}\mathbf{F}_o^H)^T]\|_2 \le \gamma\sqrt{p\varepsilon}, \end{aligned} \quad (16)$$

where $\mathbf{F}_o$ is the initial chosen equalizer. If this problem is feasible, the result $(\mathbf{G}, \frac{1}{\gamma}\mathbf{F}_o)$ is used as the initial starting point for the iterative algorithm. Otherwise, another $\mathbf{F}_o$ with a different beamforming direction may be chosen and (16) is solved again.

B. The Duality Based Method

In the MSE expression (6), the random variable $\mathbf{\Delta}$ is weighted by unknown $\mathbf{F}$ and $\mathbf{G}$, whose values in general depend on $\mathbf{\Delta}$. Therefore, the MSE is a sum of correlated elements. According to the generalized weak-convergence theorem [12], a sum of many random variables with dependence will tend to be distributed according to one of a small set of stable distributions. This means that although the channel estimation uncertainty $\mathbf{\Delta}$ is arbitrary distributed, the MSE in (6) is in fact not arbitrary distributed. Therefore, the Chebyshev bound in (7) is quite loose [11], and the QoS and power saving performance of the Chebyshev method is expected to be conservative. In this subsection, the exact solution of the lower level problem is derived by the proposed duality method.

Since the MSE is a function of the channel estimation uncertainty, let $\psi(\mathbf{x}) = \text{MSE}(\mathbf{\Delta})$, where $\mathbf{x} \triangleq \text{vec}(\mathbf{\Delta})$. The lower level problem $\sup_{\text{vec}(\mathbf{\Delta})\sim\mathcal{A}(\mathbf{0},\mathbf{\Sigma})} \text{Prob}\{\text{MSE}(\mathbf{\Delta}) \ge \varepsilon\}$ can be reformulated as

$$\begin{aligned} \sup_{f(\mathbf{x})} \quad & \text{Prob}\{\psi(\mathbf{x}) \ge \varepsilon\} \\ \text{s.t.} \quad & \int_{\mathbf{x}\in\mathbb{C}^{NM}} f(\mathbf{x})d\mathbf{x} = 1, \quad \mathbb{E}(\mathbf{x}) = \mathbf{0} \\ & \mathbb{E}(\mathbf{x}\mathbf{x}^H) = \mathbf{\Sigma}, \end{aligned} \quad (17)$$

where $f(\mathbf{x})$ is the probability density function (PDF) of the vectorized uncertainty $\mathbf{x}$.

In order to solve the problem (17), the lagrangian of this problem is presented as

$$\begin{aligned} & \mathcal{L}\left(f(\mathbf{x}), \alpha, \boldsymbol{\eta}, \mathbf{\Xi}\right) \\ & = \alpha + \text{Tr}(\mathbf{\Xi}^H\mathbf{\Sigma}) \\ & \quad + \int_{\psi(\mathbf{x})\ge\varepsilon} \Big(1 - \alpha - \boldsymbol{\eta}^H\mathbf{x} - \text{Tr}(\mathbf{\Xi}^H\mathbf{x}\mathbf{x}^H)\Big) \cdot f(\mathbf{x})d\mathbf{x} \\ & \quad + \int_{\psi(\mathbf{x})<\varepsilon} \Big(0 - \alpha - \boldsymbol{\eta}^H\mathbf{x} - \text{Tr}(\mathbf{\Xi}^H\mathbf{x}\mathbf{x}^H)\Big) \cdot f(\mathbf{x})d\mathbf{x}. \end{aligned} \quad (18)$$

where $\alpha, \boldsymbol{\eta}, \mathbf{\Xi}$ are the lagrangian multipliers, and $\mathbf{\Xi} = \mathbf{\Xi}^H$. With the implicit PDF constraint $f(\mathbf{x}) \ge 0$, the lagrange dual function of the problem (17) is

$$g\left(\alpha, \boldsymbol{\eta}, \mathbf{\Xi}\right) = \sup_{f(\mathbf{x})\ge 0} \mathcal{L}\left(f(\mathbf{x}), \alpha, \boldsymbol{\eta}, \mathbf{\Xi}\right). \quad (19)$$

Note that the supremum of the first integral in (18) with the nonnegative PDF constraint is zero if $\alpha + \boldsymbol{\eta}^H\mathbf{x} + \text{Tr}(\mathbf{\Xi}^H\mathbf{x}\mathbf{x}^H) > 1$, otherwise the supremum is infinity. Similarly, the supremum of the second integral with the nonnegative PDF constraint is also zero if $\alpha + \boldsymbol{\eta}^H\mathbf{x} + \text{Tr}(\mathbf{\Xi}^H\mathbf{x}\mathbf{x}^H) \ge 0$, otherwise the supremum is infinity. Therefore, the lagrange dual function $g\left(\alpha, \boldsymbol{\eta}, \mathbf{\Xi}\right)$ is

$$g\left(\alpha, \boldsymbol{\eta}, \mathbf{\Xi}\right) = \begin{cases} \alpha + \text{Tr}(\mathbf{\Xi}^H\mathbf{\Sigma}) & \text{if} \quad \begin{array}{c} \alpha + \boldsymbol{\eta}^H\mathbf{x} + \text{Tr}(\mathbf{\Xi}^H\mathbf{x}\mathbf{x}^H) \ge 0, \\ \forall\mathbf{x}: \psi(\mathbf{x}) < \varepsilon; \quad \text{and} \\ \alpha + \boldsymbol{\eta}^H\mathbf{x} + \text{Tr}(\mathbf{\Xi}^H\mathbf{x}\mathbf{x}^H) > 1, \\ \forall\mathbf{x}: \psi(\mathbf{x}) \ge \varepsilon \end{array} \\ +\infty & \text{otherwise.} \end{cases} \quad (20)$$

Combining the two conditions that make $g\left(\alpha, \boldsymbol{\eta}, \mathbf{\Xi}\right) = \alpha + \text{Tr}(\mathbf{\Xi}^H\mathbf{\Sigma})$, the first condition can be replaced by $\alpha + \boldsymbol{\eta}^H\mathbf{x} + \text{Tr}(\mathbf{\Xi}^H\mathbf{x}\mathbf{x}^H) \ge 0$ for all $\mathbf{x} \in \mathbb{C}^{NM}$. Therefore, the dual of the problem (17) can be formulated as

$$\begin{aligned} \min_{\alpha,\boldsymbol{\eta},\mathbf{\Xi}} \quad & \alpha + \text{Tr}(\mathbf{\Xi}^H\mathbf{\Sigma}) \\ \text{s.t.} \quad & \alpha + \boldsymbol{\eta}^H\mathbf{x} + \text{Tr}(\mathbf{\Xi}^H\mathbf{x}\mathbf{x}^H) \ge 0, \quad \forall\mathbf{x}: \mathbf{x} \in \mathbb{C}^{NM} \\ & \alpha + \boldsymbol{\eta}^H\mathbf{x} + \text{Tr}(\mathbf{\Xi}^H\mathbf{x}\mathbf{x}^H) > 1, \quad \forall\mathbf{x}: \psi(\mathbf{x}) \ge \varepsilon \\ & \mathbf{\Xi} = \mathbf{\Xi}^H. \end{aligned} \quad (21)$$

TABLE I
CHEBYSHEV INEQUALITY BASED ROBUST TRANSCEIVER DESIGN

1.	Set iteration number $j = 1$, initialize with a feasible transceiver set $(\mathbf{G}[0], \mathbf{F}[0])$, define $P[0] = \text{Tr}(\mathbf{G}[0]^H\mathbf{G}[0])$
2.	Update $\mathbf{G}[j]$ using the solution of (11), calculate $P[j] = \text{Tr}(\mathbf{G}[j]^H\mathbf{G}[j])$
3.	Update $\mathbf{F}[j]$ using (15)
4.	If $P[j-1] - P[j] \le \epsilon$ (ϵ is a pre-defined threshold) then stop, otherwise increment j and go to step 2

Proposition 1: The strong duality holds between the primal problem (17) and the dual problem (21) under the condition $\mathbf{\Sigma} \succ 0$.

Proof: First, it is recognized that the problem (17) is known as the moment problem [13]. Since only the first two moments of the random vector $\mathbf{x}$ are used in (17), the feasible moment vector set of $\mathbf{x}$ is $\mathcal{M} = \{(\mathbf{0}, \mathbf{\Sigma})|\mathbf{\Sigma} \succeq 0\}$. According to the general theory on the moment problem [13, p.812], strong duality holds between the primal moment problem (17) and its dual problem (21) when the moment vector of $\mathbf{x}_k$ is an interior point of $\mathcal{M}$, i.e., $\mathbf{\Sigma} \succ 0$. ■

In order to get a compact form of (21), we define $\mathbf{Z} \triangleq \begin{bmatrix} \mathbf{\Xi}^H & \frac{1}{2}\boldsymbol{\eta} \\ \frac{1}{2}\boldsymbol{\eta}^H & \alpha \end{bmatrix}$, $\tilde{\mathbf{\Sigma}} \triangleq \begin{bmatrix} \mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix}$ and $\mathbf{u} \triangleq \begin{bmatrix} \mathbf{x}^T & 1 \end{bmatrix}^T$, then we get

$$\begin{aligned} \min_{\mathbf{Z}} \quad & \mathrm{Tr}(\mathbf{Z}\tilde{\mathbf{\Sigma}}) \\ \text{s.t.} \quad & \mathbf{u}^H\mathbf{Z}\mathbf{u} \geq 0, \qquad \forall \mathbf{x} : \mathbf{x} \in \mathbb{C}^{NM} \\ & \mathbf{u}^H\mathbf{Z}\mathbf{u} - 1 > 0, \quad \forall \mathbf{x} : \psi(\mathbf{x}) \geq \varepsilon \\ & \mathbf{Z} = \mathbf{Z}^H. \end{aligned} \tag{22}$$

Note that the first and the third constraints of (22) can be combined into $\mathbf{Z} \succeq 0$. Furthermore, after replacing the MSE term $\psi(\mathbf{x})$ with $\mathrm{MSE}(\mathbf{\Delta})$ in (6), the problem (22) becomes

$$\begin{aligned} \min_{\mathbf{Z}} \quad & \mathrm{Tr}(\mathbf{Z}\tilde{\mathbf{\Sigma}}) \\ \text{s.t.} \quad & \mathbf{Z} \succeq 0 \\ & \mathbf{u}^H\mathbf{Z}\mathbf{u} - 1 > 0, \forall \mathbf{\Delta} : \|\mathbf{F}(\hat{\mathbf{H}} + \mathbf{\Delta})\mathbf{G} - \mathbf{I}\|_F^2 \\ & +\mathrm{Tr}(\mathbf{F}\mathbf{R}\mathbf{F}^H) \geq \varepsilon. \end{aligned} \tag{23}$$

It is observed that the problem (23) is a deterministic problem, but it is still an infinite constrained problem.

In order to transform the infinite constrained problem (23) into a finite constrained one, we first reformulate the Frobenius norm as a spectral norm as follows

$$\begin{aligned} & \|\mathbf{F}(\hat{\mathbf{H}} + \mathbf{\Delta})\mathbf{G} - \mathbf{I}\|_F^2 \\ & = \|\mathrm{vec}(\mathbf{F}\hat{\mathbf{H}}\mathbf{G} - \mathbf{I}) + \mathrm{vec}(\mathbf{F}\mathbf{\Delta}\mathbf{G})\|_2^2 && (24) \\ & = \left\| [\mathbf{G}^T \otimes \mathbf{F} \;\; \mathrm{vec}(\mathbf{F}\hat{\mathbf{H}}\mathbf{G} - \mathbf{I})] \begin{bmatrix} \mathrm{vec}(\mathbf{\Delta}) \\ 1 \end{bmatrix} \right\|_2^2 && (25) \\ & = \mathbf{u}^H\mathbf{Q}^H\mathbf{Q}\mathbf{u}, && (26) \end{aligned}$$

where $\mathbf{Q} \triangleq [\mathbf{G}^T \otimes \mathbf{F} \;\; \mathrm{vec}(\mathbf{F}\hat{\mathbf{H}}\mathbf{G} - \mathbf{I})]$. Putting (26) into the condition of the second constraint of (23), we get the quadratic form

$$\begin{aligned} & \|\mathbf{F}(\hat{\mathbf{H}} + \mathbf{\Delta})\mathbf{G} - \mathbf{I}\|_F^2 + \mathrm{Tr}(\mathbf{F}\mathbf{R}\mathbf{F}^H) - \varepsilon \\ & = \mathbf{u}^H\left(\mathbf{Q}^H\mathbf{Q} - \begin{bmatrix} \mathbf{0}_{NM} & \mathbf{0}_{NM\times 1} \\ \mathbf{0}_{1\times NM} & \varepsilon - \mathrm{Tr}(\mathbf{F}\mathbf{R}\mathbf{F}^H) \end{bmatrix}\right)\mathbf{u} && (27) \\ & = \mathbf{u}^H(\mathbf{Q}^H\mathbf{Q} - \mathrm{diag}([\mathbf{0}, \varepsilon - \mathrm{Tr}(\mathbf{F}\mathbf{R}\mathbf{F}^H)]))\mathbf{u}. && (28) \end{aligned}$$

Therefore, the second constraint of the problem (23) can be transformed as

$$\begin{aligned} \mathbf{u}^H(\mathbf{Z} - \mathrm{diag}([\mathbf{0}, 1]))\mathbf{u} > 0, \quad \forall \mathbf{x} : \mathbf{u}^H\left(\mathbf{Q}^H\mathbf{Q}\right. \\ \left. - \mathrm{diag}([\mathbf{0}, \varepsilon - \mathrm{Tr}(\mathbf{F}\mathbf{R}\mathbf{F}^H)])\right)\mathbf{u} \geq 0. \end{aligned} \tag{29}$$

By using the S-Lemma in control theory [14], (29) can be equivalently formulated as

$$\begin{aligned} \exists \lambda \geq 0 : \; \mathbf{Z} - \mathrm{diag}([\mathbf{0}, 1]) - \lambda\left(\mathbf{Q}^H\mathbf{Q}\right. \\ \left. - \mathrm{diag}([\mathbf{0}, \varepsilon - \mathrm{Tr}(\mathbf{F}\mathbf{R}\mathbf{F}^H)])\right) \succeq 0. \end{aligned} \tag{30}$$

Note that we can omit $\lambda = 0$ in (30) to arrive at a stronger condition, that is still equivalent to (29). This is because the eigenvalue of $\mathbf{Q}^H\mathbf{Q} - \mathrm{diag}([\mathbf{0}, \varepsilon - \mathrm{Tr}(\mathbf{F}\mathbf{R}\mathbf{F}^H)])$ is finite due to the finite transmission power and finite value of the equalizers. Therefore, there exists a small enough $\lambda \neq 0$ such that the right side of (30) is satisfied. Based on this observation, letting $\beta \triangleq 1/\lambda > 0$, the constraint (30) can be reformulated as

$$\exists \beta > 0 : \; \beta\mathbf{Z} + \mathrm{diag}([\mathbf{0}, \varepsilon - \mathrm{Tr}(\mathbf{F}\mathbf{R}\mathbf{F}^H) - \beta]) - \mathbf{Q}^H\mathbf{Q} \succeq 0. \tag{31}$$

According to the Schur's Complement [15], (31) can be transformed into an linear matrix inequality (LMI) form of β as

$$\exists \beta > 0 : \; \begin{bmatrix} \beta\mathbf{Z} + \mathrm{diag}([\mathbf{0}, \varepsilon - \mathrm{Tr}(\mathbf{F}\mathbf{R}\mathbf{F}^H) - \beta]) & \mathbf{Q}^H \\ \mathbf{Q} & \mathbf{I}_L \end{bmatrix} \succeq 0. \tag{32}$$

After replacing the second constraint of the problem (23) with its equivalent form (32), the problem (23) becomes

$$\begin{aligned} \min_{\beta, \mathbf{Z}} \quad & \mathrm{Tr}(\mathbf{Z}\tilde{\mathbf{\Sigma}}) \\ \text{s.t.} \quad & \mathbf{Z} \succeq 0, \quad \beta > 0 \\ & \begin{bmatrix} \beta\mathbf{Z} + \mathrm{diag}([\mathbf{0}, \varepsilon - \mathrm{Tr}(\mathbf{F}\mathbf{R}\mathbf{F}^H) - \beta]) & \mathbf{Q}^H \\ \mathbf{Q} & \mathbf{I}_L \end{bmatrix} \succeq 0. \end{aligned} \tag{33}$$

Defining a new variable $\tilde{\mathbf{Z}} \triangleq \beta\mathbf{Z}$, then the problem (33) further becomes an LMI problem

$$\begin{aligned} \min_{\beta, \tilde{\mathbf{Z}}} \quad & \mathrm{Tr}(\tilde{\mathbf{Z}}\tilde{\mathbf{\Sigma}})/\beta \\ \text{s.t.} \quad & \tilde{\mathbf{Z}} \succeq 0, \quad \beta > 0 \\ & \begin{bmatrix} \tilde{\mathbf{Z}} + \mathrm{diag}([\mathbf{0}, \varepsilon - \mathrm{Tr}(\mathbf{F}\mathbf{R}\mathbf{F}^H) - \beta]) & \mathbf{Q}^H \\ \mathbf{Q} & \mathbf{I}_L \end{bmatrix} \succeq 0. \end{aligned} \tag{34}$$

Therefore, the stochastic problem (17) is transformed into the deterministic finite constrained problem (34). Since the strong duality holds between the primal and dual problem, after replacing the lower level problem of ($\mathbb{P}0$) with (34), we get the following results.

Proposition 2: The original bilevel optimization problem ($\mathbb{P}0$) is equivalent to the following single-level problem

$$\begin{aligned} \min_{\mathbf{G}, \mathbf{F}, \beta, \tilde{\mathbf{Z}}} \quad & \mathrm{Tr}(\mathbf{G}\mathbf{G}^H) \\ \text{s.t.} \quad & \mathrm{Tr}(\tilde{\mathbf{Z}}\tilde{\mathbf{\Sigma}})/\beta \leq p, \quad \beta > 0, \quad \tilde{\mathbf{Z}} \succeq 0, \\ & \begin{bmatrix} \tilde{\mathbf{Z}} + \mathrm{diag}([\mathbf{0}, \varepsilon - \mathrm{Tr}(\mathbf{F}\mathbf{R}\mathbf{F}^H) - \beta]) & \mathbf{Q}^H \\ \mathbf{Q} & \mathbf{I}_L \end{bmatrix} \succeq 0. \end{aligned} \tag{$\mathbb{P}1$}$$

Similar to the proposed Chebyshev method, the problem ($\mathbb{P}1$) can be solved by an iteration between two subproblems as follows.

In the first subproblem, when the equalizer $\mathbf{F}$ is fixed, the subproblem becomes a convex problem of $\mathbf{G}$, and the precoder design problem is

$$\begin{aligned} \min_{\mathbf{G}, \beta, \tilde{\mathbf{Z}}} \quad & \mathrm{Tr}(\mathbf{G}\mathbf{G}^H) \\ \text{s.t.} \quad & \mathrm{Tr}(\tilde{\mathbf{Z}}\tilde{\mathbf{\Sigma}}) \leq p \cdot \beta, \quad \beta > 0, \quad \tilde{\mathbf{Z}} \succeq 0, \\ & \begin{bmatrix} \tilde{\mathbf{Z}} + \mathrm{diag}([\mathbf{0}, \varepsilon - \mathrm{Tr}(\mathbf{F}\mathbf{R}\mathbf{F}^H) - \beta]) & \mathbf{Q}^H \\ \mathbf{Q} & \mathbf{I}_L \end{bmatrix} \succeq 0. \end{aligned} \tag{35}$$

This convex problem can be efficiently solved by the interior point method in [16].

For the second subproblem, the equalizer $\mathbf{F}$ is chosen to minimize the guaranteed data MSE $\tilde{\varepsilon}$ to create a larger feasible space for the next round precoder design. Therefore, $\mathbf{F}$ is updated using the following problem

$$\begin{aligned} \min_{\mathbf{F},\beta,\tilde{\mathbf{Z}},\tilde{\varepsilon}} \quad & \tilde{\varepsilon} \\ \text{s.t.} \quad & \mathrm{Tr}(\tilde{\mathbf{Z}}\tilde{\mathbf{\Sigma}}) \leq p \cdot \beta, \quad \beta > 0, \quad \tilde{\mathbf{Z}} \succeq 0 \\ & \begin{bmatrix} \tilde{\mathbf{Z}} + \mathrm{diag}([\mathbf{0}, \tilde{\varepsilon} - \mathrm{Tr}(\mathbf{R}\mathbf{F}^H\mathbf{F}) - \beta]) & \mathbf{Q}^H \\ \mathbf{Q} & \mathbf{I}_L \end{bmatrix} \succeq 0. \end{aligned} \tag{36}$$

In contrast to the problem (12), we cannot get a closed-form solution from (36). However, by using semidefinite relaxation (SDR) [17], this problem can be relaxed to a convex one. In particular, let $\mathbf{C} = \mathbf{F}^H\mathbf{F}$, then the equalizer $\mathbf{F}$ can be efficiently obtained from the semidefinite programming (SDP) problem

$$\begin{aligned} \min_{\mathbf{F},\mathbf{C},\beta,\tilde{\mathbf{Z}},\tilde{\varepsilon}} \quad & \tilde{\varepsilon} \\ \text{s.t.} \quad & \mathrm{Tr}(\tilde{\mathbf{Z}}\tilde{\mathbf{\Sigma}}) \leq p \cdot \beta, \quad \beta > 0, \quad \tilde{\mathbf{Z}} \succeq 0 \\ & \begin{bmatrix} \tilde{\mathbf{Z}} + \mathrm{diag}([\mathbf{0}, \tilde{\varepsilon} - \mathrm{Tr}(\mathbf{R}\mathbf{C}) - \beta]) & \mathbf{Q}^H \\ \mathbf{Q} & \mathbf{I}_L \end{bmatrix} \succeq 0 \\ & \begin{bmatrix} \mathbf{C} & \mathbf{F}^H \\ \mathbf{F} & \mathbf{I}_L \end{bmatrix} \succeq 0, \end{aligned} \tag{37}$$

where the last constraint is a convex relaxation from $\mathbf{C} = \mathbf{F}^H\mathbf{F}$ to $\mathbf{C} \succeq \mathbf{F}^H\mathbf{F}$. Note that the SDR is tight [18].

The proposed duality based method starts from a feasible initial point and iterates between (35) and (37) until convergence. For initialization, the conventional identity or randomly generated receiver may not be feasible. Although it is appealing to have a feasibility enhancement method tailored for the QoS constraints in (35), but the resulting feasibility enhancement problem is nonconvex, and is difficult to solve. However, the initialization results of the Chebyshev method in (16) can also be taken as an initialization for the duality method, because the approximated QoS constraints in the Chebyshev method are more stringent than the original requirement. Therefore, the iterative algorithm for the duality method can be represented as in Table I, except that the two subproblems (11) and (15) are replaced by (35) and (37), respectively.

IV. Simulation Results and Discussions

In this section, the performance of the proposed two algorithms is illustrated by simulations. In the simulations, a 2×2 MIMO system with spatial multiplexing is implemented, i.e., $N = 2, M = 2, L = 2$. The estimated channel and the estimation error have complex entries with zero mean and covariance matrix $(1 - \sigma^2)\mathbf{I}_4$ and $\mathbf{\Sigma} = \sigma^2\mathbf{I}_4$, respectively. In this paper, σ^2 is fixed at 0.005. The elements of the estimated channel $\hat{\mathbf{H}}$ are generated as Gaussian distributed random variables, while that of the uncertainty $\mathbf{\Delta}$ are assumed to be Laplace distributed. For simplicity, the interference plus noise covariance is fixed as $\mathbf{R} = 0.01 \times \mathbf{I}_2$. The required outage probability is $p = 0.15$, the QoS target is varied between 0.1 and 0.5, and the threshold to terminate the iteration is $\epsilon = 5 \times 10^{-4}$. All simulation results are averaged over 10^3 randomly chosen channel realizations. For fair comparison, the randomly chosen channel realizations are feasible for all algorithms under consideration in Figs. 1 and 2.

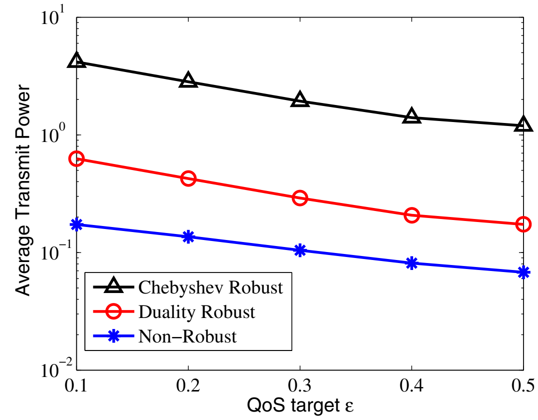

Fig. 1. The transmission power with different QoS requirement.

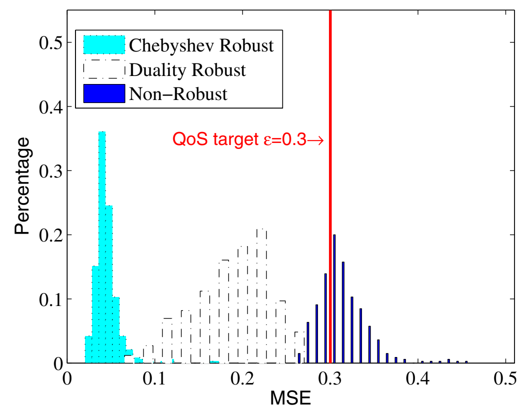

Fig. 2. The residual MSE with QoS requirement $p = 0.15, \varepsilon = 0.3$.

In Fig. 1, the relative transmit power of the two proposed algorithms is illustrated. It is observed that over a wide range of QoS target, the duality method consistently requires a significantly lower transmission power than that of the Chebyshev method, saving more than 10 dB in transmission power. This is owing to the loose Chebyshev inequality which makes the approximated QoS requirements in (10) over conservative. The transmission power of the non-robust method in [4] is also plotted and is founded to be the lowest. However, the QoS performance of the non-robust method cannot be guaranteed, and is shown in Fig. 2.

In Fig. 2, the histograms of the recovered data MSEs of the proposed two methods and the non-robust method are compared. It is observed that the residual MSEs of the two proposed robust methods are smaller than the prefixed QoS target, while that of the non-robust method exceeds the QoS target for about 70% of the channel realizations. Furthermore, it is also noted that the residual MSEs of the duality method are very close to the QoS target, while that of the Chebyshev method are much smaller than the target. This shows the conservative nature of the Chebyshev inequality based design.

From Figs. 1 and 2, it can be concluded that the proposed duality method guarantees the QoS requirement, while maintaining a low transmission power. The additional required transmit power of the duality method compared to the non-robust method is about 3 dB, which is the cost of the guaranteed QoS performance against the channel uncertainty. While the Chebyshev method also guarantee the QoS constraints, the duality method achieves a good balance between the QoS requirement and the power saving requirement. Similar conclusions can be drawn in other simulation settings with different estimation error distributions, e.g., Gaussian and uniform distribution. The corresponding figures are not repeated in this paper.

V. Conclusions

In this paper, probabilistic QoS constrained robust MIMO transceiver design was investigated. The objective of the proposed design is to minimize the transmit power, while still guarantees a probabilistic QoS requirement under arbitrary distributed channel estimation error. Chebyshev inequality based method and a novel duality method were proposed to solve the problem. Simulation results showed that the QoS requirement is guaranteed for both proposed methods. For the minimized transmit power, the duality method showed superior performance than the Chebyshev method, due to its tight reformulation. But the Chebyshev method exhibited a much lower computational complexity.

References

[1] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. Commu.*, vol. 49, no. 12, pp. 2198-2206, 2001.

[2] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1051-1064, 2002.

[3] D. P. Palomar, M. A. Lagunas, and J. M. Cioffi, "Optimum linear joint transmit-receive processing for MIMO channels with QoS constraints," *IEEE Trans. Signal Process.*, vol. 52, no. 5, pp. 1179-1197, 2004.

[4] S. Shi, M. Schubert, and H. Boche, "Downlink MMSE transceiver optimization for multiuser MIMO systems: MMSE balancing," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3702-3712, 2008.

[5] Y. C. Eldar, A. Nehorai, and P. S. La Rosa, "A Competitive Mean-Squared Error Approach to Beamforming," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5143-5154, 2007.

[6] N. Vucic, H. Boche, and S. Shi, "Robust transceiver optimization in downlink multiuser MIMO systems," *IEEE Trans. Signal Process.*, vol. 57, no. 9, pp. 3576-3587, 2009.

[7] S. M. Kay, "*Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*," Prentice Hall, 1993.

[8] G. Zheng, K. K. Wong, and T. S. Ng, "Energy-efficient multiuser SIMO: achieving probabilistic robustness with gaussian channel uncertainty," *IEEE Trans. Commu.*, vol. 57, no. 6, pp. 1866-1878, 2009.

[9] M. Payaro, A. Pascual-Iserte, and M. A. Lagunas, "Robust power allocation designs for multiuser and multiantenna downlink communication systems through convex optimization," *IEEE J. Select. Areas Commun.*, vol. 25 no. 7, pp. 1390-1401, 2007.

[10] Y. Liu, T. F. Wong, and W. W. Hager, "Training signal design for estimation of correlated MIMO channels with colored interference," *IEEE Trans. Signal Process.*, vol. 55, no. 4, pp. 1486-1497, 2007.

[11] H. Stark and J.W. Woods, "*Probability and Random Processes with Applications to Signal Processing*," Prentice Hall, 3rd Edition, 2001

[12] S. Umarov, C. Tsallis, and S. Steinberg, "On a q-central limit theorem consistent with nonextensive statistical mechanics," *Milan J. math.*, vol. 76, pp. 307-328, 2008.

[13] J. Smith, "Generalized chebychev inequalities: theory and applications in decision analysis," *Oper. Res.*, vol. 43, pp. 807-825, 1995.

[14] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, "*Linear Matrix Inequalities in System and Control Theory*," SIAM Studies in Applied Mathematics, SIAM, Philadelphia, 1994.

[15] R. A. Horn and C. R. Johnson, "*Matrix Analysis*," Cambridge University Press, 1990.

[16] Y. Nesterov and A. Nemirovsky, "*Interior Point Polynomial Methods in Convex Programming*," SIAM, 1994.

[17] A. d'Aspremont and S. Boyd, "Relaxations and randomized methods for nonconvex QCQPs," *lecture notes at Stanford University*, Autumn, 2003.

[18] N. D. Sidiropoulos, T. N. Davidson, and Z.-Q. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2239-2251, 2006.