

Research Article

Robust MMSE Transceiver Designs for Downlink MIMO Systems with Multicell Cooperation

Jialing Li, I-Tai Lu, and Enoch Lu

Department of ECE, Polytechnic Institute of NYU, 6 Metrotech Center, Brooklyn, NY 11201, USA

Correspondence should be addressed to Jialing Li, jialing.li.phd2@gmail.com

Received 1 July 2009; Accepted 16 December 2009

Academic Editor: Hongxiang Li

Copyright © 2010 Jialing Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The *robust-generalized iterative approach (Robust-GIA)*, *robust-fast iterative approach (Robust-FIA)*, and *robust-decoder covariance optimization approach (Robust-DCOA)* are proposed for designing MMSE transceivers of downlink multicell multiuser MIMO systems with per-cell and per-antenna power constraints and possibly imperfect channel state information. The *Robust-DCOA* is the most restrictive but is always optimum, the *Robust-GIA* is the most general, and the *Robust-FIA* is the most efficient. When the *Robust-DCOA* is applicable and the decoder covariance matrices are full rank, the three proposed approaches are equivalent and all provide the optimum solution. Numerical results show that the proposed robust approaches outperform their non-robust counterparts in various single-cell and multicell examples with different system configurations, channel correlations, power constraints, and cooperation scenarios. Moreover, performances of the robust approaches are insensitive to estimation errors of channel statistics (correlations and path loss). With cell-cooperation, cell edge interference problems can be remedied without reducing the number of data streams by using the proposed robust approaches.

1. Introduction

Joint transceiver designs with criteria such as minimum mean square error (MMSE), maximum sum capacity, and minimum bit error rate (BER), and so forth, for multiple-input-multiple-output (MIMO) systems, with both uplink and downlink configurations, have been studied intensively in recent literature (e.g., see [1, 2]). Discussed in this paper is the robust MMSE transceiver design with respect to channel estimation errors for downlink multicell multiuser MIMO systems.

Assuming perfect channel state information (CSI), joint MMSE transceiver design has been studied by many researchers. A closed form design subject to the total power constraint for a single-user MIMO system is derived in [3]. Unfortunately though, this closed form design cannot be extended either to the multiuser case or to the per-antenna power constraint. For multiuser uplink MIMO problems subject to per-user power constraint, numerical solutions are provided mainly by the *transmit covariance optimization approach (TCOA)* [4, 5] and iterative approaches such as in [4]. We have developed a *generalized iterative approach (GIA)* for the uplink to deal with arbitrary linear power constraints

(including the more practical per-antenna power constraint) [6]. Recently, we have also extended the *TCOA* to deal with arbitrary linear power constraints and have shown that the *GIA* and the *TCOA* are equivalent and optimum when the source covariance matrices are all projection matrices multiplied by the same constant and the transmit covariance matrices are all full rank [7].

For the downlink configuration, iterative approaches such as in [8] and a dual uplink approach in [9, 10] are employed to provide numerical MMSE solutions for multiuser MIMO systems subject to the total power constraint. The extension to deal with the per-antenna and per-cell power constraints for the downlink scenario is achieved by an iterative approach using a second-order cone programming (SOCP) [11] and by our *GIA* for the downlink [12]. Recently, we have also developed the *decoder covariance optimization approach (DCOA)* [13] to deal with arbitrary linear power constraints (including the per-antenna and per-cell power constraints). Furthermore, we have shown that the *GIA* and the *DCOA* are equivalent and optimum when the source covariance matrices are all identity matrices multiplied by the same constant and the decoder covariance matrices are all full rank [13].

All of the above mentioned MMSE transceiver designs are based on perfect CSI. However, the CSI is usually estimated in practice and is therefore subject to CSI estimation errors and possibly quantized CSI feedback errors. Hence, in practice, joint transceiver design has to be based on imperfect CSI. One option is to ignore that the CSI is imperfect. This type of approach is herein called non-robust. Unfortunately, the system performances derived from the non-robust approaches depend strongly on the quality of the available CSI (performances get worse quickly if the CSI quality deteriorates). Moreover, an optimum design based on poor CSI could be worse than suboptimum designs using the same CSI. Therefore, a more appealing option is to model the CSI error and to incorporate the error model into the transceiver design. This type of approach is herein called robust. The robust approaches can better mitigate the degradation of system performances due to imperfect CSI than the non-robust approaches if the CSI error is modeled correctly. Two classes of imperfect CSI models are usually employed: the stochastic model for the CSI estimation errors and the deterministic model for the CSI feedback errors. If a stochastic model is used, a statistically robust design is usually performed to optimize some system performance functions. If a deterministic model is used, a minimax or maximin design aiming at optimizing the worst-case system performance is usually carried out.

To cope with CSI estimation errors, closed form solutions for the robust joint MMSE transceiver design subject to the total power constraint are developed for single-user MIMO systems in [14–16]. But, similar to the perfect CSI case, no closed form solution is found when the problem is extended to deal with either multiuser applications or the per-antenna power constraint. For multiuser uplink MIMO problems subject to the inequality per-user and sum power constraints, the *robust transmit covariance optimization approach (Robust-TCOA)* is developed for independent identically distributed (i.i.d.) MIMO channels with CSI estimation errors in [17]. For multiuser downlink MIMO problems, when the CSI errors are bounded, the worst-case design under arbitrary power constraints is made based on SOCP in [18]; when the CSI errors are statistical errors, the robust design under the total power constraint is solved numerically by a dual uplink approach in [18].

So far, no statistically robust approach has been shown optimum in the MMSE sense for the downlink MIMO systems (either single-cell or multicell) under the per-antenna power constraint. Proposed in this paper is the robust MMSE transceiver design with respect to CSI estimation errors for downlink multicell MIMO systems subject to arbitrary linear power constraints. Specifically, the per-antenna and per-cell power constraints are considered. The work is relevant to frequency division duplex (FDD) systems where channel estimation is done at each user equipment (UE) and then fed back to the base station(s) (denoted as evolved Node B or eNB) via a zero-delay and error free communication link. Note that CSI feedback errors are not considered in this paper. The work may possibly also be extended to time division duplex (TDD) systems where channel estimation is done at the eNBs.

We first extend the statistical model of imperfect CSI in [16] to take into account the path loss effects. This extension is very crucial for practical multicell systems because the variances of CSI estimation errors depend on the distances between the UEs and the eNBs. The CSI estimation error of a UE near the eNB is much smaller than the CSI estimation error of a UE that is far away from the eNB. With the extended imperfect CSI model in hand, we herein propose three robust approaches to deal with arbitrary linear equality power constraints. The first, the *robust-generalized iterative approach (Robust-GIA)*, is an extension of the GIA [12, 13] to the imperfect CSI case. The second, the *robust-decoder covariance optimization approach (Robust-DCOA)*, is an extension of the DCOA [13]. The third, the *robust-fast iterative approach (Robust-FIA)* is completely new. Though the first two are both extensions, their complexities are still similar to those of their predecessors!

The DCOA requires that the numbers of data streams are not prespecified and that all the source covariance matrices are identity matrices multiplied by the same constant. The *Robust-DCOA* is even more restricted since not only it requires all of the conditions of the DCOA but also that the transmit correlation matrix for each user is an identity matrix. The statistics of the CSI estimation error also need to be the same for all users if the power constraints of the users are interdependent. The GIA and the *Robust-GIA*, on the other hand, do not require any of the above mentioned conditions. The *Robust-FIA* has the same requirements as the DCOA, but not the additional restrictions of the *Robust-DCOA*.

The relationships between the *Robust-GIA*, the *Robust-FIA*, and the *Robust-DCOA* are very interesting. The *Robust-GIA* is the most general and can provide tradeoff between diversity and multiplexing gains. The *Robust-FIA* is the most efficient. Even though the *Robust-DCOA* is the most restricted, it always gives the optimum solution when it is applicable. But whenever the *Robust-DCOA* is applicable and all the decoder covariance matrices are full rank, the solutions obtained by the three robust approaches are actually equivalent (i.e., the *Robust-FIA* and the *Robust-GIA* are also optimum)! Interestingly, the *Robust-GIA* and the *Robust-DCOA* actually become the GIA and the DCOA, respectively, when the CSI is perfectly known, thereby providing a unified framework to take care of both perfect and imperfect CSI cases! We also denote the special case of the *Robust-FIA* when the CSI is perfectly known as the *fast iterative approach (FIA)* for convenience.

MMSE transceiver designs using the proposed robust approaches are performed for various single-cell and multicell examples with different system configurations, power constraints, channel correlations, and cooperation scenarios. System performances in terms of MSE and BER of various numerical examples are compared. Computational efficiency for various approaches is studied. Sensitivity studies with respect to channel statistics (channel correlations and path loss, estimated independently from channel estimation) are also investigated. The numerical results show that the proposed robust approaches are indeed superior to the non-robust approaches. Moreover, accurate channel correlations

and path loss are not required in the robust approaches. With cell cooperation, the cell edge UEs perform as well as those UEs without inter-cell interferences.

Notations are as follows. All boldface letters indicate vectors (lower case) or matrices (upper case). \mathbf{A}' , \mathbf{A}^* , \mathbf{A}^{-1} , $\text{tr}(\mathbf{A})$, $\langle \mathbf{A} \rangle$, and $\|\mathbf{A}\|$ stand for the transpose, Hermitian, inverse, trace, expectation, and Frobenius norm of \mathbf{A} , respectively. Matrix \mathbf{I}_a signifies an identity matrix with rank a . $\text{diag}[\dots]$ denotes the diagonal matrix with elements $[\dots]$ on the main diagonal. $\mathbf{A} \geq \mathbf{B}$ means that $\mathbf{A} - \mathbf{B}$ is positive semidefinite. $\mathbf{A} * \mathbf{B}$ denotes the Schur product of \mathbf{A} and \mathbf{B} (elementwise product of \mathbf{A} and \mathbf{B}). $\text{CN}(\mu, \sigma^2)$ denotes a complex normal random variable with mean μ and variance σ^2 .

2. Formulation

2.1. Downlink Multicell Multiuser MIMO Systems. Consider the downlink of a multicell multiuser MIMO system with C eNBs and K UEs. Denote the number of transmit antennas at the c th eNB by t_c and the total number of transmit antennas by t , that is, $t = \sum_{c=1}^C t_c$. Also denote the number of receive antennas at the i th UE by r_i and the number of data streams intended for the i th UE by m_i .

In this system, there may be multiple groups where each group jointly designs its precoders and decoders but does so independently of the other groups. In the with-cooperation scenario (there is full cooperation among all eNBs), system-wide design is performed and there is only one group. In the without-cooperation scenario (there is no cooperation among eNBs), the eNB and UEs in a cell are one group. Let D_a (D_b) and S_a (S_b) define one such group in the with-cooperation (without-cooperation) scenario, D_a (D_b) being the set of indices of all eNBs in the group and S_a (S_b) being the set of indices of all UEs in the group.

At the i th UE, let \mathbf{y}_i and \mathbf{a}_i denote the received signal and noise, respectively. At eNB c , let \mathbf{s}_i and \mathbf{F}_{ic} denote the data and the precoder for the i th UE, respectively. Also let \mathbf{H}_{ic} denote the channel matrix from eNB c to the i th UE. In the with-cooperation scenario, the data \mathbf{s}_i for the i th UE are jointly transmitted by *all* eNBs. Thus, the received signal at the i th UE is:

$$\mathbf{y}_i = \sum_{c \in D_a} \mathbf{H}_{ic} \sum_{j \in S_a} \mathbf{F}_{jc} \mathbf{s}_j + \mathbf{a}_i, \quad (1a)$$

where $D_a = \{1, \dots, C\}$ and $S_a = \{1, \dots, K\}$.

In the without-cooperation scenario, let the eNB serving the i th UE be denoted by the index c_i where $c_i \in \{1, 2, \dots, C\}$. Thus, the data \mathbf{s}_i for the i th UE are only transmitted by eNB c_i . Therefore, a system with C eNB's can be decoupled into C single-cell downlink groups. In the c_i th group, the received signal vector \mathbf{y}_i at the i th UE becomes

$$\mathbf{y}_i = \sum_{c \in D_b} \mathbf{H}_{ic} \sum_{j \in S_b} \mathbf{F}_{jc} \mathbf{s}_j + \mathbf{b}_i + \mathbf{a}_i, \quad \mathbf{b}_i = \sum_{c_j \notin D_b} \sum_{j \notin S_b} \mathbf{H}_{ic_j} \mathbf{F}_{j c_j} \mathbf{s}_{j}, \quad (1b)$$

where $D_b = \{c_i\}$ consists of the index of eNB c_i and $S_b = \{j \mid c_j = c_i\}$ consists of the indexes of all UEs served by eNB c_i .

Since there are multiuser precodings at eNB c_i , the data transmitted by eNB c_i to the UEs other than the i th UE are not considered as interference to the i th UE. But, the data transmitted from the eNBs other than eNB c_i result in the inter-cell interference (denoted as \mathbf{b}_i) to the i th UE.

In order to unify (1a) and (1b), let D and S define a group in the system, D being the set of indices of all eNBs in the group and S being the set of indices of all UEs in the group. For the i th UE, $i \in S$, the received signal vector, \mathbf{y}_i , can thus be expressed as

$$\mathbf{y}_i = \sum_{c \in D} \mathbf{H}_{ic} \sum_{j \in S} \mathbf{F}_{jc} \mathbf{s}_j + \mathbf{n}_i = \mathbf{H}_i \sum_{j \in S} \mathbf{F}_j \mathbf{s}_j + \mathbf{n}_i. \quad (2)$$

When there is full eNB cooperation,

$$\mathbf{H}_i = [\mathbf{H}_{i1} \cdots \mathbf{H}_{iC}], \quad \mathbf{F}_j = [\mathbf{F}'_{j1} \cdots \mathbf{F}'_{jC}]', \quad (2a)$$

$$\mathbf{n}_i = \mathbf{a}_i, \quad D = D_a, \quad S = S_a.$$

When there is no eNB cooperation,

$$\mathbf{H}_i = \mathbf{H}_{ic_i}, \quad \mathbf{F}_i = \mathbf{F}_{ic_i}, \quad (2b)$$

$$\mathbf{n}_i = \mathbf{b}_i + \mathbf{a}_i, \quad D = D_b, \quad S = S_b.$$

In (2b), \mathbf{H}_i denotes the channel matrix from the eNB c_i to the i th UE, \mathbf{F}_i denotes the precoder for the i th UE at eNB c_i , and \mathbf{n}_i is the interference plus noise vector at the i th UE. But, in (2a), \mathbf{H}_i denotes the composite channel matrix from all eNBs to the i th UE, \mathbf{F}_i denotes the composite precoder for the i th user at all eNBs, and \mathbf{n}_i is the noise vector. Equation (2) is essentially the same as the formulation in [11].

2.2. Extended Imperfect CSI Model [16]. In order to account for path loss and spatial correlation, the channel \mathbf{H}_{ic} from eNB c to the i th UE is modeled as

$$\mathbf{H}_{ic} = \alpha d_{ic}^{-\beta} \mathbf{R}_{R,i}^{1/2} \mathbf{H}_{W,ic} \mathbf{R}_{T,c}^{1/2}. \quad (3)$$

Here, α is a constant, d_{ic} denotes the distance between the i th UE and eNB c , and 2β is the path loss exponent. In (3), $\mathbf{R}_{R,i}$ and $\mathbf{R}_{T,c}$ are known, normalized (unit diagonal entries), and full rank. They represent receive and transmit correlation matrices, respectively. The entries of $\mathbf{H}_{W,ic}$ are i.i.d. $\text{CN}(0, 1)$. Here, the subscript W represents spatially white.

In practice, the CSI is estimated, resulting in estimation error. Thus,

$$\mathbf{H}_{ic} = \hat{\mathbf{H}}_{ic} + \mathbf{E}_{ic}, \quad (4)$$

where $\hat{\mathbf{H}}_{ic}$ is the channel estimate and \mathbf{E}_{ic} is the CSI estimation error. By using an orthogonal training method and MMSE channel estimation, $\hat{\mathbf{H}}_{ic}$ and \mathbf{E}_{ic} have been shown in [16] to be independent and

$$\mathbf{E}_{ic} = \alpha d_{ic}^{-\beta} \mathbf{R}_{E,ic}^{1/2} \mathbf{E}_{W,ic} \mathbf{R}_{T,c}^{1/2}, \quad \mathbf{R}_{E,ic} = [\mathbf{I}_{r_i} + \sigma_{E,ic}^2 \cdot \mathbf{R}_{R,i}^{-1}]^{-1}, \quad (4a)$$

$$\sigma_{E,ic}^2 = \alpha^{-2} d_{ic}^{2\beta} \text{tr}(\mathbf{R}_{T,c}^{-1}) \cdot \frac{\sigma_a^2}{T_c}, \quad (4b)$$

where the entries of $\mathbf{E}_{W,ic}$ are i.i.d. $\text{CN}(0, \sigma_{E,ic}^2)$. Here, σ_a^2 is the noise variance at each of the receive antennas and T_c is the total training power transmitted from eNB c . Note that there is no inter-cell interference when the channel is estimated. Also note that the estimated channels $\{\hat{\mathbf{H}}_{ic}\}_{i \in S}$ are fed back by the UEs to a central processing unit in the with-cooperation scenario and to eNB c in the without-cooperation scenario via a zero-delay and error free communication link. As to be expected, $\hat{\mathbf{H}}_i$, $i \in S$, is equal to $[\hat{\mathbf{H}}_{i1} \cdots \hat{\mathbf{H}}_{iC}]$ in the with-cooperation scenario and to $\hat{\mathbf{H}}_{ic}$, $c \in D$, in the without-cooperation scenario. Also, when perfect CSI is available, $\hat{\mathbf{H}}_{ic} = \mathbf{H}_{ic}$ and $\sigma_{E,ic}^2 = 0$.

2.3. Joint MMSE Precoder and Decoder Design Formulation. For a given group and thus a given D and S , the following is the problem formulation. Define the mean square error (MSE) of the data streams intended for the i th UE, $i \in S$, as

$$\eta_i(\mathbf{F}_i, \mathbf{G}_i) = \text{tr} \left\langle (\mathbf{G}_i \mathbf{y}_i - \mathbf{s}_i) (\mathbf{G}_i \mathbf{y}_i - \mathbf{s}_i)^* \right\rangle, \quad (5a)$$

where \mathbf{G}_i is the decoder at the i th UE, \mathbf{F}_i is the precoder in (2a) or (2b), and \mathbf{y}_i and \mathbf{s}_i are given in (2). Equation (5a), using the actual channel \mathbf{H}_i and actual noise vector \mathbf{n}_i , is the metric for MSE evaluation for the perfect-CSI, non-robust, and robust approaches. However, in the robust MMSE transceiver design, the following conditional MSE

$$\begin{aligned} \eta_i(\mathbf{F}_i, \mathbf{G}_i | \hat{\mathbf{H}}_i) &= \text{tr} \left\langle (\mathbf{G}_i \mathbf{y}_i - \mathbf{s}_i) (\mathbf{G}_i \mathbf{y}_i - \mathbf{s}_i)^* \right\rangle, \\ \mathbf{y}_i &= \hat{\mathbf{H}}_i \sum_{j \in S} \mathbf{F}_j \mathbf{s}_j + \tilde{\mathbf{n}}_i, \quad \tilde{\mathbf{n}}_i = \mathbf{n}_i + \mathbf{E}_{ic} \sum_{j \in S} \mathbf{F}_j \mathbf{s}_j \end{aligned} \quad (5b)$$

is used. Here, when perfect CSI is available, $\hat{\mathbf{H}}_i$ and $\tilde{\mathbf{n}}_i$ represent the actual channel and actual noise vector, that is, $\hat{\mathbf{H}}_i = \mathbf{H}_i$ and $\tilde{\mathbf{n}}_i = \mathbf{n}_i$, respectively; otherwise, $\hat{\mathbf{H}}_i$ represents the channel estimate, that is, $\hat{\mathbf{H}}_i \neq \mathbf{H}_i$, and $\tilde{\mathbf{n}}_i$ represents the equivalent interference plus noise vector. We will jointly choose the decoders and precoders corresponding to all UEs in S and all eNBs in D to minimize the sum MSE η :

$$\{\mathbf{F}_i, \mathbf{G}_i\}_{\text{MMSE}} = \underset{\{\mathbf{F}_i, \mathbf{G}_i | i \in S\}}{\text{argmin}} \{\eta\}, \quad \eta = \sum_{j \in S} \eta_j(\mathbf{F}_j, \mathbf{G}_j | \hat{\mathbf{H}}_j). \quad (6)$$

Define the positive definite source covariance matrix as $\Phi_{si} = \langle \mathbf{s}_i \mathbf{s}_i^* \rangle$ for the i th UE. The eNB(s) are subject to either the per-antenna or per-cell power constraints. For the per-antenna power constraint, the d th antenna of the c th eNB, $c \in D$, has power

$$P_{ad} = \mathbf{e}'_d \left(\sum_{j \in S} \mathbf{F}_j \Phi_{sj} \mathbf{F}_j^* \right) \mathbf{e}_d. \quad (7a)$$

For the per-cell power constraint, the c th eNB, $c \in D$, has power

$$P_{bc} = \text{tr} \left(\mathbf{Q}'_c \left(\sum_{j \in S} \mathbf{F}_{jc} \Phi_{sj} \mathbf{F}_{jc}^* \right) \mathbf{Q}_c \right). \quad (7b)$$

Here, \mathbf{e}_d in (7a) are an $l \times 1$ unit vector with the d th entry equal to one and all other entries equal to zero, \mathbf{Q}_c in (7b) is an $l \times l$ matrix whose entries are all equal to zero except for the diagonal elements corresponding to the antennas of the c th eNB which are equal to one, and $l = \sum_{c \in D} t_c$.

2.4. Augmented Cost Function. To solve (6) subject to (7a) or (7b), one can use the *method of Lagrange multipliers* to set up the augmented cost function:

$$\xi = \eta + \text{tr} \left(\Lambda \left(\sum_{j \in S} \mathbf{F}_j \Phi_{sj} \mathbf{F}_j^* - \mathbf{P} \right) \right), \quad (8)$$

where Λ is an unknown diagonal matrix, representing the Lagrange multipliers. For the per-antenna power constraint in (7a),

$$\Lambda = \text{diag}[\lambda_1, \dots, \lambda_l], \quad \mathbf{P} = \text{diag}[P_{a1}, \dots, P_{al}]. \quad (9a)$$

For the per-cell power constraint in (7b), define $\Lambda_k = \mathbf{I}_{t_k} \lambda_k$ and $\mathbf{P}_k = \mathbf{I}_{t_k} P_{bk}/t_k$, $k=1, \dots, C$. Thus,

$$\Lambda = \Lambda_c, \quad \mathbf{P} = \mathbf{P}_c, \quad c \in D, \quad \text{without cooperation};$$

$$\Lambda = \text{diag}[\Lambda_1 \dots \Lambda_C], \quad \mathbf{P} = \text{diag}[\mathbf{P}_1 \dots \mathbf{P}_C], \quad \text{with cooperation.} \quad (9b)$$

2.5. Robust Design. The central processing unit is assumed to have knowledge about the channel estimate(s), $\hat{\mathbf{H}}_j$, $j \in S$, and the channel statistics. Define the interference-plus-noise covariance matrix at the i th UE as $\Phi_{ni} = \langle \mathbf{n}_i \mathbf{n}_i^* \rangle$. Thus,

$$\Phi_{ni} = \begin{cases} \Phi_{ai} = \langle \mathbf{a}_i \mathbf{a}_i^* \rangle & \text{with cooperation;} \\ \Phi_{bi} + \Phi_{ai} = \langle \mathbf{b}_i \mathbf{b}_i^* \rangle + \langle \mathbf{a}_i \mathbf{a}_i^* \rangle & \text{without cooperation.} \end{cases} \quad (10)$$

Also define the equivalent interference-plus-noise covariance matrix at the i th UE as $\tilde{\Phi}_{ni} = \langle \tilde{\mathbf{n}}_i \tilde{\mathbf{n}}_i^* \rangle$. After some math manipulations, (5b) for the i th UE, $i \in S$, becomes

$$\begin{aligned} \eta_i(\mathbf{F}_i, \mathbf{G}_i | \hat{\mathbf{H}}_i) &= -\mathbf{G}_i \hat{\mathbf{H}}_i \mathbf{F}_i \Phi_{si} - \Phi_{si} \mathbf{F}_i^* \hat{\mathbf{H}}_i^* \mathbf{G}_i^* \\ &\quad + \Phi_{si} + \mathbf{G}_i \left[\hat{\mathbf{H}}_i \left(\sum_{j \in S} \mathbf{F}_j \Phi_{sj} \mathbf{F}_j^* \right) \hat{\mathbf{H}}_i^* + \tilde{\Phi}_{ni} \right] \mathbf{G}_i^*, \end{aligned} \quad (11)$$

where $\tilde{\Phi}_{ni} = \Xi_i + \Phi_{ni}$ and

$$\Xi_i = \begin{cases} \sum_{c \in D} \sigma_c^2 P_{bc} \cdot \mathbf{R}_{E,ic}, & \text{if } \mathbf{R}_{T,c} = \mathbf{I}_{t_c}, \quad c \in D, \\ \sum_{c \in D} \sigma_c^2 \sum_{j \in S} \text{tr}(\mathbf{F}_{jc} \Phi_{sj} \mathbf{F}_{jc}^* \mathbf{R}_{T,c}) \mathbf{R}_{E,ic}, & \text{if } \mathbf{R}_{T,c} \neq \mathbf{I}_{t_c}, \quad c \in D, \end{cases} \quad (12)$$

$$\sigma_c^2 = \text{tr}(\mathbf{R}_{T,c}^{-1}) \cdot \frac{\sigma_a^2}{T_c}. \quad (13)$$

For a given set of precoders $\{\mathbf{F}_j\}_{j \in S}$, setting the gradient of ξ in (8) with respect to \mathbf{G}_i equal to zero, we yield the MMSE decoder for the i th UE, $i \in S$:

$$\mathbf{G}_i = \Phi_{s_i} \mathbf{F}_i^* \hat{\mathbf{H}}_i^* \mathbf{M}_i, \quad \mathbf{M}_i = \left(\hat{\mathbf{H}}_i \left(\sum_{j \in S} \mathbf{F}_j \Phi_{s_j} \mathbf{F}_j^* \right) \hat{\mathbf{H}}_i^* + \tilde{\Phi}_{n_i} \right)^{-1}. \quad (14)$$

Substituting (11) and (14) into (8), the augmented cost function ξ in (8) is reduced to

$$\begin{aligned} \xi_1 = & \sum_{j \in S} \text{tr} \left(-\Phi_{s_j} \mathbf{F}_j^* \hat{\mathbf{H}}_j^* \mathbf{M}_j \hat{\mathbf{H}}_j \mathbf{F}_j \Phi_{s_j} + \Phi_{s_j} \right) \\ & + \text{tr} \left(\Lambda \left(\sum_{j \in S} \mathbf{F}_j \Phi_{s_j} \mathbf{F}_j^* - \mathbf{P} \right) \right). \end{aligned} \quad (15)$$

Note that the ξ_1 in (15) no longer depends on $\{\mathbf{G}_j\}_{j \in S}$ explicitly.

On the other hand, for a given set of decoders $\{\mathbf{G}_j\}_{j \in S}$ and Lagrange multipliers Λ , setting the gradient of ξ in (8) with respect to \mathbf{F}_i equal to zero, we have the MMSE precoder for the i th UE, $i \in S$:

$$\mathbf{F}_i = \mathbf{N} \hat{\mathbf{H}}_i^* \mathbf{G}_i^*, \quad \mathbf{N} = \left(\sum_{j \in S} \hat{\mathbf{H}}_j^* \mathbf{G}_j^* \mathbf{G}_j \hat{\mathbf{H}}_j + \Lambda + \Theta \right)^{-1}, \quad (16)$$

where

$$\begin{aligned} \Theta &= \begin{cases} 0, & \text{if } \mathbf{R}_{T,c} = \mathbf{I}_{t_c}, c \in D, \\ \Upsilon \mathbf{R}_T, & \text{if } \mathbf{R}_{T,c} \neq \mathbf{I}_{t_c}, c \in D, \end{cases} \\ \Upsilon &= \begin{cases} \mathbf{Y}_c, & c \in D \text{ without cooperation,} \\ \text{diag}[\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_C] & \text{with cooperation,} \end{cases} \\ \mathbf{R}_T &= \begin{cases} \mathbf{R}_{T,c}, & c \in D \text{ without cooperation,} \\ \text{diag}[\mathbf{R}_{T,1}, \mathbf{R}_{T,2}, \dots, \mathbf{R}_{T,C}] & \text{with cooperation.} \end{cases} \end{aligned} \quad (17)$$

Here, $\mathbf{Y}_c = \sigma_c^2 \sum_{i \in S} \text{tr}(\mathbf{G}_i^* \mathbf{G}_i \mathbf{R}_{E,ic}) \mathbf{I}_{t_c}$ and σ_c^2 is given in (13). Substituting (11) and (16) into (8), the augmented cost function ξ in (8) is reduced to

$$\begin{aligned} \xi_2 = & \text{tr} \left(- \left(\sum_{j \in S} \hat{\mathbf{H}}_j^* \mathbf{G}_j^* \Phi_{s_j} \mathbf{G}_j \hat{\mathbf{H}}_j \right) \mathbf{N} - \Lambda \mathbf{P} \right) \\ & + \sum_{j \in S} \text{tr} \left(\Phi_{s_j} + \mathbf{G}_j \bar{\Phi}_{n_j} \mathbf{G}_j^* \right), \end{aligned} \quad (18)$$

where

$$\bar{\Phi}_{n_j} = \begin{cases} \tilde{\Phi}_{n_j}, & \text{if } \mathbf{R}_{T,c} = \mathbf{I}_{t_c}, c \in D, \\ \Phi_{n_j}, & \text{if } \mathbf{R}_{T,c} \neq \mathbf{I}_{t_c}, c \in D. \end{cases} \quad (19)$$

The ξ_2 in (18) no longer depends on $\{\mathbf{F}_j\}_{j \in S}$ explicitly.

3. Robust Iterative Approaches

3.1. Robust-Generalized Iterative Approach (Robust-GIA). By setting the gradients of (15) with respect to \mathbf{F}_j equal to zero, left multiplying the resulting equation with \mathbf{F}_j , and summing up the resulting equation over j , we obtain

$$\sum_{j \in S} \left(\mathbf{F}_j \Phi_{s_j} \mathbf{F}_j^* \right) \Lambda = \mathbf{B}, \quad (20)$$

where

$$\begin{aligned} \mathbf{B} = & \sum_{j \in S} \mathbf{F}_j \Phi_{s_j}^2 \mathbf{F}_j^* \hat{\mathbf{H}}_j^* \mathbf{M}_j \hat{\mathbf{H}}_j - \Omega \\ & - \sum_{j \in S} \left(\mathbf{F}_j \Phi_{s_j} \mathbf{F}_j^* \right) \sum_{k \in S} \left(\hat{\mathbf{H}}_k^* \mathbf{M}_k \hat{\mathbf{H}}_k \mathbf{F}_k \Phi_{s_k}^2 \mathbf{F}_k^* \hat{\mathbf{H}}_k^* \mathbf{M}_k \hat{\mathbf{H}}_k \right), \end{aligned} \quad (21)$$

$$\Omega = \begin{cases} 0, & \text{if } \mathbf{R}_{T,c} = \mathbf{I}_{t_c}, c \in D, \\ \left(\sum_{j \in S} \mathbf{F}_j \Phi_{s_j} \mathbf{F}_j^* \right) \mathbf{R}_T \sum_{j \in S} \Sigma_j, & \text{if } \mathbf{R}_{T,c} \neq \mathbf{I}_{t_c}, c \in D, \end{cases} \quad (22)$$

$$\Sigma_j = \begin{cases} \Sigma_{j,c}, & c \in D \text{ without cooperation,} \\ \text{diag}[\Sigma_{j,1}, \Sigma_{j,2}, \dots, \Sigma_{j,C}] & \text{with cooperation} \end{cases} \quad (23)$$

with $\Sigma_{j,c} = \sigma_c^2 \text{tr}(\mathbf{R}_{E,jc} \mathbf{M}_j \hat{\mathbf{H}}_j \mathbf{F}_j \Phi_{s_j}^2 \mathbf{F}_j^* \hat{\mathbf{H}}_j^* \mathbf{M}_j) \mathbf{I}_{t_c}$. Utilizing (9a) and (9b), we can obtain explicit expressions for the Lagrange multipliers Λ as follows. For the per-antenna power constraint in (7a), we have

$$\Lambda = \mathbf{P}^{-1} (\mathbf{I}_l * \mathbf{B}), \quad (24a)$$

and for the per-cell power constraint in (7b), we have

$$\lambda_c = P_{bc}^{-1} \text{tr}[\mathbf{Q}_c' (\mathbf{I}_l * \mathbf{B}) \mathbf{Q}_c]. \quad (24b)$$

With the explicit expression for the Lagrange multipliers Λ in (24a) or (24b) in hand, a *Robust-GIA* can be developed using the MMSE decoder in (14) and MMSE precoder in (16). There are three steps in each iteration of the *Robust-GIA*.

Step 1. Given $\{\mathbf{F}_i\}_{i \in S}$, obtain $\{\mathbf{G}_i\}_{i \in S}$ by (14).

Step 2. Given $\{\mathbf{F}_i\}_{i \in S}$, obtain Λ using (24a) or (24b).

Step 3. Given $\{\mathbf{G}_i\}_{i \in S}$ and Λ , obtain $\{\mathbf{F}_i\}_{i \in S}$ by (16).

Note that the *Robust-GIA* can allow tradeoff between diversity and multiplexing gains because it can deal with various sets of prespecified numbers of data streams intended for the UEs.

3.2. *Robust-Fast Iterative Approach (Robust-FIA) When $\Phi_{si} = \sigma^2 \mathbf{I}_{m_i}$.* The *Robust-FIA* can be developed based on the *Robust-GIA* when the source covariance matrices are all identity matrices multiplied by the same constant, that is, $\Phi_{si} = \sigma^2 \mathbf{I}_{m_i}$, $i \in S$. For convenience and without loss of generality, we assume $\sigma^2 = 1$. Define the transmit covariance matrices as

$$\mathbf{U}_i = \mathbf{F}_i \mathbf{F}_i^*, \quad (25)$$

and the decoder covariance matrices as

$$\mathbf{V}_i = \mathbf{G}_i^* \mathbf{G}_i. \quad (26)$$

Substituting (14) into (26) and using (25), we obtain

$$\mathbf{V}_i = \mathbf{M}_i \hat{\mathbf{H}}_i \mathbf{U}_i \hat{\mathbf{H}}_i^* \mathbf{M}_i, \quad \mathbf{M}_i = \left(\hat{\mathbf{H}}_i \left(\sum_{j \in S} \mathbf{U}_j \right) \hat{\mathbf{H}}_i^* + \tilde{\Phi}_{ni} \right)^{-1}. \quad (27)$$

Similarly, substituting (16) into (25) and using (26), we obtain

$$\mathbf{U}_i = \mathbf{N} \hat{\mathbf{H}}_i^* \mathbf{V}_i \hat{\mathbf{H}}_i \mathbf{N}, \quad \mathbf{N} = \left(\sum_{j \in S} \hat{\mathbf{H}}_j^* \mathbf{V}_j \hat{\mathbf{H}}_j + \Lambda + \Theta \right)^{-1}, \quad (28)$$

where Θ is given in (17). Substituting (26) into Υ_c , we can express Θ in (17) in terms of $\{\mathbf{V}_i\}_{i \in S}$ in (26). To remove the dependence of $\{\mathbf{V}_i\}_{i \in S}$ on $\{\mathbf{U}_i\}_{i \in S}$, substitute (28) into (27) to yield

$$\mathbf{V}_i = \mathbf{M}_i \hat{\mathbf{H}}_i \mathbf{N} \hat{\mathbf{H}}_i^* \mathbf{V}_i \hat{\mathbf{H}}_i \mathbf{N} \hat{\mathbf{H}}_i^* \mathbf{M}_i, \quad (29)$$

$$\mathbf{M}_i = \left(\hat{\mathbf{H}}_i \mathbf{N} \left(\sum_{j \in S} \hat{\mathbf{H}}_j^* \mathbf{V}_j \hat{\mathbf{H}}_j \right) \mathbf{N} \hat{\mathbf{H}}_i^* + \tilde{\Phi}_{ni} \right)^{-1},$$

where \mathbf{N} is given in (28). Similarly, using the fact that $\Phi_{si} = \sigma^2 \mathbf{I}_{m_i}$ and substituting (25) and (28) into (21)–(23), we can express the Lagrange multipliers Λ in (24a) or (24b) in terms of $\{\mathbf{V}_i\}_{i \in S}$.

With (24a), (24b), and (29) being available, the *Robust-FIA* can be readily developed. There are two steps in each iteration of the *Robust-FIA*.

Step 1. Given $\{\mathbf{V}_i\}_{i \in S}$, obtain Λ using (24a) or (24b).

Step 2. Given Λ , use (29) to obtain $\{\mathbf{V}_i\}_{i \in S}$ for the next iteration.

Note that the number of data streams intended for the UE's $\{m_i\}_{i \in S}$, has to be equal to the ranks of $\{\mathbf{V}_i\}_{i \in S}$ that the *Robust-FIA* returns and thus cannot be prespecified when using the *Robust-FIA*. When the *Robust-FIA* converges, the decoders $\{\mathbf{G}_i\}_{i \in S}$ can be obtained by the decomposition defined in (26) and the precoders $\{\mathbf{F}_i\}_{i \in S}$ can then be obtained from (16). Note that the decomposition in (26) is not unique: $\mathbf{V}_i = \mathbf{G}_i^* \mathbf{G}_i = \mathbf{G}_i^* \mathbf{A}_i \mathbf{A}_i^* \mathbf{G}_i$ where \mathbf{A}_i is an arbitrary unitary matrix. One can easily show that when $\Phi_{si} = \sigma^2 \mathbf{I}_{m_i}$, $i \in S$, if $(\mathbf{F}_i, \mathbf{G}_i)$ is a pair of joint MMSE precoder and

decoder, so is $(\mathbf{F}_i \mathbf{A}_i, \mathbf{A}_i^* \mathbf{G}_i)$. Both $(\mathbf{F}_i, \mathbf{G}_i)$ and $(\mathbf{F}_i \mathbf{A}_i, \mathbf{A}_i^* \mathbf{G}_i)$ give the same MMSE η . However, different choices of $\{\mathbf{A}_i\}$ matrices may lead to different BER results.

Note that when the CSI is perfectly known, the *Robust-FIA* is reduced to the *FIA* by replacing $\{\hat{\mathbf{H}}_i\}_{i \in S}$ by $\{\mathbf{H}_i\}_{i \in S}$ in (24a), (24b), and (29), setting $\tilde{\Phi}_{ni} = 0$ in \mathbf{M}_i in (29) and setting $\Theta = 0$ in \mathbf{N} in (28).

4. Robust-Decoder Covariance Optimization Approach (*Robust-DCOA*)

When the source covariance matrices are all identity matrices multiplied by the same constant, that is, $\Phi_{si} = \sigma^2 \mathbf{I}_{m_i}$, $i \in S$, and when the transmit correlation matrices are all identity matrices, that is, $\mathbf{R}_{T,c} = \mathbf{I}_c$, $c \in D$, a *robust-decoder covariance optimization approach (Robust-DCOA)* can be used for jointly designing the MMSE transceivers. For convenience, we assume $\sigma^2 = 1$. The augmented cost function in (18) becomes

$$\xi_2 = \text{tr} \left[\left(\sum_{i \in S} \hat{\mathbf{H}}_i^* \mathbf{V}_i \hat{\mathbf{H}}_i + \Lambda \right)^{-1} \Lambda \right] + \sum_{i \in S} \text{tr}(\mathbf{V}_i \tilde{\Phi}_{ni}) - \text{tr}[\Lambda \mathbf{P}] + \sum_{i \in S} m_i - l. \quad (30)$$

The robust MMSE transceiver design problem becomes

$$\begin{aligned} & \min_{\{\mathbf{V}_i\}_{i \in S}} \max_{\Lambda} \xi_2, \\ & \text{subject to } \mathbf{V}_i \geq 0, \text{ rank}(\mathbf{V}_i) = m_i, i \in S, \\ & \Lambda \geq 0. \end{aligned} \quad (31a)$$

The problem in (31a) is not convex because of the implicit rank constraints dealing with the numbers of data streams, that is, $\text{rank}(\mathbf{V}_i) = m_i$. Allowing $\{m_i\}_{i \in S}$ to be unspecified and noting that l is a known constant, we obtain the rank-relaxed decoder covariance optimization problem:

$$\begin{aligned} & \min_{\mathbf{V}_i \geq 0, i \in S} \max_{\Lambda \geq 0} \xi_{2,\text{rel}}, \\ & \xi_{2,\text{rel}} = \text{tr} \left[\left(\sum_{i \in S} \hat{\mathbf{H}}_i^* \mathbf{V}_i \hat{\mathbf{H}}_i + \Lambda \right)^{-1} \Lambda - \Lambda \mathbf{P} \right] + \sum_{i \in S} \text{tr}(\mathbf{V}_i \tilde{\Phi}_{ni}). \end{aligned} \quad (31b)$$

The cost function $\xi_{2,\text{rel}}$ in (31b) is convex with respect to $\{\mathbf{V}_i\}_{i \in S}$ and concave with respect to Λ . Define $\min_{\{\mathbf{V}_i \geq 0, i \in S\}} \max_{\{\Lambda \geq 0\}} \xi_{2,\text{rel}}$ as the primal problem and $\max_{\{\Lambda \geq 0\}} \min_{\{\mathbf{V}_i \geq 0, i \in S\}} \xi_{2,\text{rel}}$ as the dual problem. Since both the primal problem and the dual problem are convex and strictly feasible, strong duality holds; that is, the optimum values of $\{\mathbf{V}_i\}_{i \in S}$, Λ , and $\xi_{2,\text{rel}}$ obtained from the primal problem are the same as those obtained from the dual problem.

4.1. Primal-Dual Algorithm. We propose a novel *primal-dual algorithm* to solve the rank-relaxed decoder covariance optimization problem in (31b). Denote the feasible set of values for $\{\mathbf{V}_i\}_{i \in S}$ as the primal domain and the feasible set of values for Λ as the dual domain. In short, the approach consists of iterating between a primal domain step (looking in the primal domain for the best $\{\mathbf{V}_i\}_{i \in S}$ for a given Λ) and a dual domain step (looking in the dual domain for the best Λ for a given $\{\mathbf{V}_i\}_{i \in S}$). The iterative procedure stops when the $\xi_{2,\text{rel}}$'s corresponding to the primal domain step and the dual domain step converge to the same value and when $\{\mathbf{V}_i\}_{i \in S}$ converge and Λ converge. The two steps of the $(j+1)$ th iteration are as follows.

Step 1 (Primal domain step). Given $\Lambda = \Lambda^{(j)}$, find the $\{\mathbf{V}_i\}_{i \in S}$ which solves (32). Denote them as $\{\mathbf{V}_i^{(j+1)}\}_{i \in S}$:

$$\min_{\{\mathbf{V}_i\}_{i \in S}} \text{tr} \left[\left(\sum_{i \in S} \hat{\mathbf{H}}_i^* \mathbf{V}_i \hat{\mathbf{H}}_i + \Lambda \right)^{-1} \Lambda \right] + \sum_{i \in S} \text{tr}(\mathbf{V}_i \tilde{\Phi}_{ni}), \quad (32)$$

subject to $\mathbf{V}_i \geq 0$, $i \in S$.

Step 2 (Dual domain step). Given $\{\mathbf{V}_i\}_{i \in S} = \{\mathbf{V}_i^{(j+1)}\}_{i \in S}$, find the Λ which solves (33). Denote it as $\Lambda^{(j+1)}$:

$$\max_{\Lambda} \text{tr} \left[\left(\sum_{i \in S} \hat{\mathbf{H}}_i^* \mathbf{V}_i \hat{\mathbf{H}}_i + \Lambda \right)^{-1} \Lambda - \Lambda \mathbf{P} \right], \quad (33)$$

subject to $\Lambda \geq 0$.

Both subproblems, defined in (32) and (33), are convex because their cost functions are convex and concave, respectively, and their constraints are all linear matrix inequalities. Therefore the resulting solution of each subproblem is optimum. Furthermore, the convexity of the rank-relaxed decoder covariance optimization problem guarantees that the resulting solution provided by the *primal-dual algorithm* is global optimum. Once the optimal Λ and $\{\mathbf{V}_i\}_{i \in S}$ are obtained, the optimum numbers of data streams $\{m_i\}_{i \in S}$, the optimum decoders $\{\mathbf{G}_i\}_{i \in S}$, and the corresponding optimum precoders $\{\mathbf{F}_i\}_{i \in S}$ are obtained in the same way as in the *Robust-FIA*. And, in all this, the power constraints have been accounted for by the Lagrange multipliers.

In practice, the *Robust-DCOA* given by solving (32) and (33) is considered to have converged at the $(j+1)$ th iteration when $\{\|\mathbf{V}_i^{(j+1)} - \mathbf{V}_i^{(j)}\|\}_{i \in S}$, $\|\Lambda^{(j+1)} - \Lambda^{(j)}\|$, and the duality gap of the values of $\xi_{2,\text{rel}}$ derived from the two steps

$$\text{gap}^{(j+1)} = \xi_{2,\text{rel}} \left(\{\mathbf{V}_i^{(j+1)}\}_{i \in S}, \Lambda^{(j+1)} \right) - \xi_{2,\text{rel}} \left(\{\mathbf{V}_i^{(j+1)}\}_{i \in S}, \Lambda^{(j)} \right), \quad (34)$$

is less than some prespecified thresholds.

4.2. Semidefinite Programming (SDP) Procedure. Similar to the TCOA [5, 7] in uplink MIMO systems, (32) and (33) can be reformulated as SDP formulations, which can be solved numerically using existing codes such as SeDuMi [19]

and Yalmip [20]. Equation (32) can be reformulated as SDP formulation:

$$\begin{aligned} & \min_{\mathbf{W}_p, \{\mathbf{V}_i\}_{i \in S}} \text{tr}[\mathbf{W}_p \Lambda] + \sum_{i \in S} \text{tr}(\mathbf{V}_i \tilde{\Phi}_{ni}), \\ & \text{subject to } \mathbf{V}_i \geq 0, \quad i \in S, \\ & \begin{bmatrix} \mathbf{W}_p & \mathbf{I} \\ \mathbf{I} & \sum_{i \in S} \hat{\mathbf{H}}_i^* \mathbf{V}_i \hat{\mathbf{H}}_i + \Lambda \end{bmatrix} \geq 0. \end{aligned} \quad (35)$$

Since

$$\begin{aligned} \left(\sum_{i \in S} \hat{\mathbf{H}}_i^* \mathbf{V}_i \hat{\mathbf{H}}_i + \Lambda \right)^{-1} \Lambda &= \mathbf{I} - \left(\sum_{i \in S} \hat{\mathbf{H}}_i^* \mathbf{V}_i \hat{\mathbf{H}}_i + \Lambda \right)^{-1} \\ & \quad \left(\sum_{i \in S} \hat{\mathbf{H}}_i^* \mathbf{V}_i \hat{\mathbf{H}}_i \right), \end{aligned} \quad (36)$$

(33) can also be reformulated as SDP formulation:

$$\begin{aligned} & \min_{\mathbf{W}_d, \Lambda} \text{tr} \left[\mathbf{W}_d \left(\sum_{i \in S} \hat{\mathbf{H}}_i^* \mathbf{V}_i \hat{\mathbf{H}}_i \right) \right] + \text{tr}(\Lambda \mathbf{P}), \\ & \text{subject to } \Lambda \geq 0, \end{aligned} \quad (37)$$

$$\begin{bmatrix} \mathbf{W}_d & \mathbf{I} \\ \mathbf{I} & \sum_{i \in S} \hat{\mathbf{H}}_i^* \mathbf{V}_i \hat{\mathbf{H}}_i + \Lambda \end{bmatrix} \geq 0.$$

4.3. Numerically Efficient Procedure. We observe poor convergence behavior of the SDP procedure for the *Robust-DCOA* at high SNRs due to numerical errors introduced by SDP solvers. We therefore use the explicit closed form expression of Λ given in (24a) and (24b) in the *Robust-FIA* for the dual domain step in (33). The SDP procedure in (35) is still employed for the primal domain step in (32). This improves the convergence of the *Robust-DCOA* greatly.

4.4. Equivalence of the Robust-DCOA, the Robust-FIA, and the Robust-GIA. When the *Robust-DCOA* is applicable and the decoder covariance matrices $\{\mathbf{V}_i\}_{i \in S}$ are full rank, we claim that the *Robust-DCOA*, the *Robust-FIA*, and the *Robust-GIA* are equivalent. Thus, the solution of the *Robust-FIA* or the *Robust-GIA* is optimum under the above mentioned conditions because the solution given by the *Robust-DCOA* is always optimal (due to convexity).

Note that the *Robust-FIA* is equivalent to the *Robust-GIA* because the *Robust-FIA* is a special case of the *Robust-GIA* when $\Phi_{si} = \sigma^2 \mathbf{I}_{m_i}$, $i \in S$. To prove the equivalence between the *Robust-DCOA* and the *Robust-FIA*, it suffices to show that the KKT conditions of the two approaches are equivalent. The KKT conditions common to both approaches are (16) and the power constraint, (7a) or (7b).

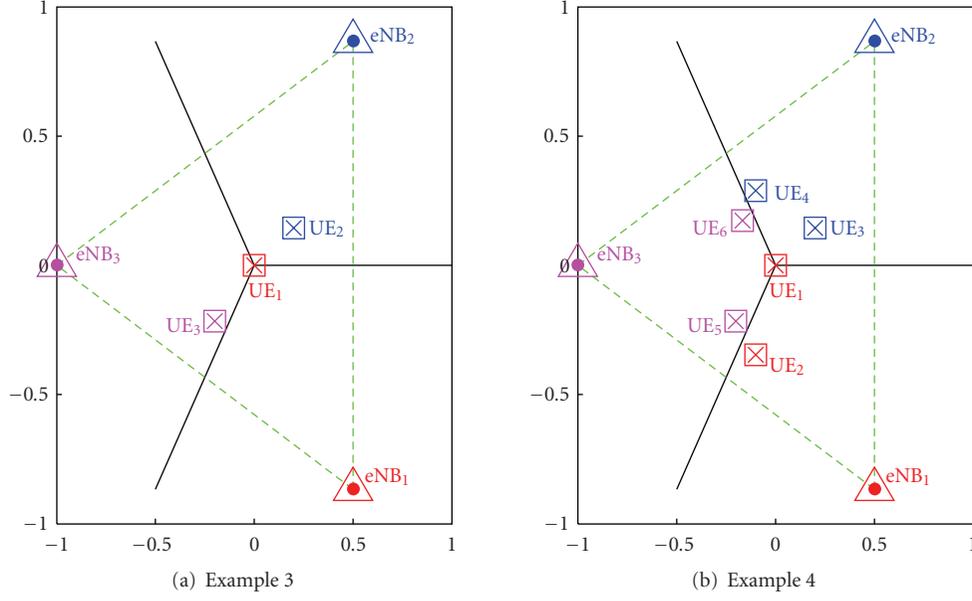


FIGURE 1: System configurations of two multicell examples: (a) 3 eNBs and 3 UEs. (b) 3 eNBs and 6 UEs. The coordinates of relevant eNBs and UEs shown here are employed for simulations. Note that UE₁ is right on the edge of three cells in both systems. In example 3, UE₃ is near the cell edge between cells 1 and 3. In example 4, UE₂ and UE₅ are near the cell edge between cells 1 and 3, and UE₄ and UE₆ are near the cell edge between cells 2 and 3. (Legends: Each UE is associated with the eNB in the same color.)

For the *Robust-DCOA*, we set up the following augmented cost function from (31b) to include the nonnegative constraint on $\{\mathbf{V}_i\}_{i \in S}$:

$$\zeta = \text{tr} \left[\left(\sum_{i \in S} \hat{\mathbf{H}}_i^* \mathbf{V}_i \hat{\mathbf{H}}_i + \Lambda \right)^{-1} \Lambda - \Lambda \mathbf{P} \right] + \sum_{i \in S} \text{tr} (\mathbf{V}_i \tilde{\Phi}_{ni} - \Psi_i \mathbf{V}_i), \quad (38)$$

where $\{\Psi_i\}_{i \in S}$ are the Lagrange Multipliers satisfying

$$\text{tr}(\Psi_i \mathbf{V}_i) = 0, \quad \Psi_i \geq 0, \quad i \in S. \quad (39)$$

Setting the gradient of (38) with respect to \mathbf{V}_i equal to zero, we have

$$\hat{\mathbf{H}}_i \mathbf{N} \mathbf{N} \hat{\mathbf{H}}_i^* = \tilde{\Phi}_{ni} - \Psi_i, \quad i \in S. \quad (40a)$$

When $\{\mathbf{V}_i\}_{i \in S}$ are full rank, the Lagrangian variable $\{\Psi_i\}_{i \in S}$ is zero matrices and (40a) becomes

$$\hat{\mathbf{H}}_i \mathbf{N} \mathbf{N} \hat{\mathbf{H}}_i^* = \tilde{\Phi}_{ni}, \quad i \in S. \quad (40b)$$

The task of showing the equivalence of the KKT conditions of the two approaches which boils down to showing the above KKT condition of the *Robust-DCOA* (40b), can be derived from (and can be used to derived to) the KKT condition unique to the *Robust-FIA*, which are (14), (24a), and (24b). Substitute (16) and (28) into (14) to obtain

$$\mathbf{G}_i \hat{\mathbf{H}}_i \mathbf{N} \hat{\mathbf{H}}_i^* \left(\hat{\mathbf{H}}_i \mathbf{N} \left(\sum_{j \in S} \hat{\mathbf{H}}_j^* \mathbf{V}_j \hat{\mathbf{H}}_j \right) \mathbf{N} \hat{\mathbf{H}}_i^* + \tilde{\Phi}_{ni} \right)^{-1} = \mathbf{G}_i. \quad (41a)$$

Then left multiply (41a) by $\mathbf{V}_i^{-1} \mathbf{G}_i^*$ to get

$$\hat{\mathbf{H}}_i \mathbf{N} \hat{\mathbf{H}}_i^* \left(\hat{\mathbf{H}}_i \mathbf{N} \left(\sum_{j \in S} \hat{\mathbf{H}}_j^* \mathbf{V}_j \hat{\mathbf{H}}_j \right) \mathbf{N} \hat{\mathbf{H}}_i^* + \tilde{\Phi}_{ni} \right)^{-1} = \mathbf{I}_r. \quad (41b)$$

Summing up (41b) over $i \in S$ and using some matrix manipulations, we can show that the resulting equation and (40b) are equivalent. To get (24a) and (24b) from (40b), note that (24a) and (24b) can be obtained by using (20) and the power constraints. In turn, (20) can be obtained by substituting (14) and (26) into (16). Since (14) and (40b) can be derived from each other, this proof is complete. The above proof is done assuming $\Phi_{si} = \mathbf{I}_{m_i}$, $i \in S$. It is also applicable when $\Phi_{si} = \sigma^2 \mathbf{I}_{m_i}$, $i \in S$, with $\sigma^2 \neq 1$.

5. Numerical Results

Without loss of generality, let $\alpha = 1$ and $\beta = 3.5$ (i.e., $2\beta = 7$) in the path loss model. Each cell is a hexagon with sides normalized to be 1 in length. The noise and source covariance matrices Φ_{ai} and Φ_{si} are all identity matrices of dimension r_i and m_i , respectively. Let the per-antenna power constraint P_{ad} for antenna d in cell c ($d = 1, 2, \dots, t_c$) be equal to P (see (7a)) and let the per-cell power constraint P_{bc} for cell c , $c = 1, 2, \dots, C$, be equal to $t_c P$ (see (7b)). Therefore, the total transmission power from eNB c is $P_{bc} = t_c P$ under both per-antenna and per-cell power constraints.

Four examples (two single-cell and two 3-cell examples) will be considered. Their system parameters are shown in Table 1. The configurations of the two single-cell examples will be detailed later on while the configurations of the two 3-cell examples are shown here in Figure 1. In example 3

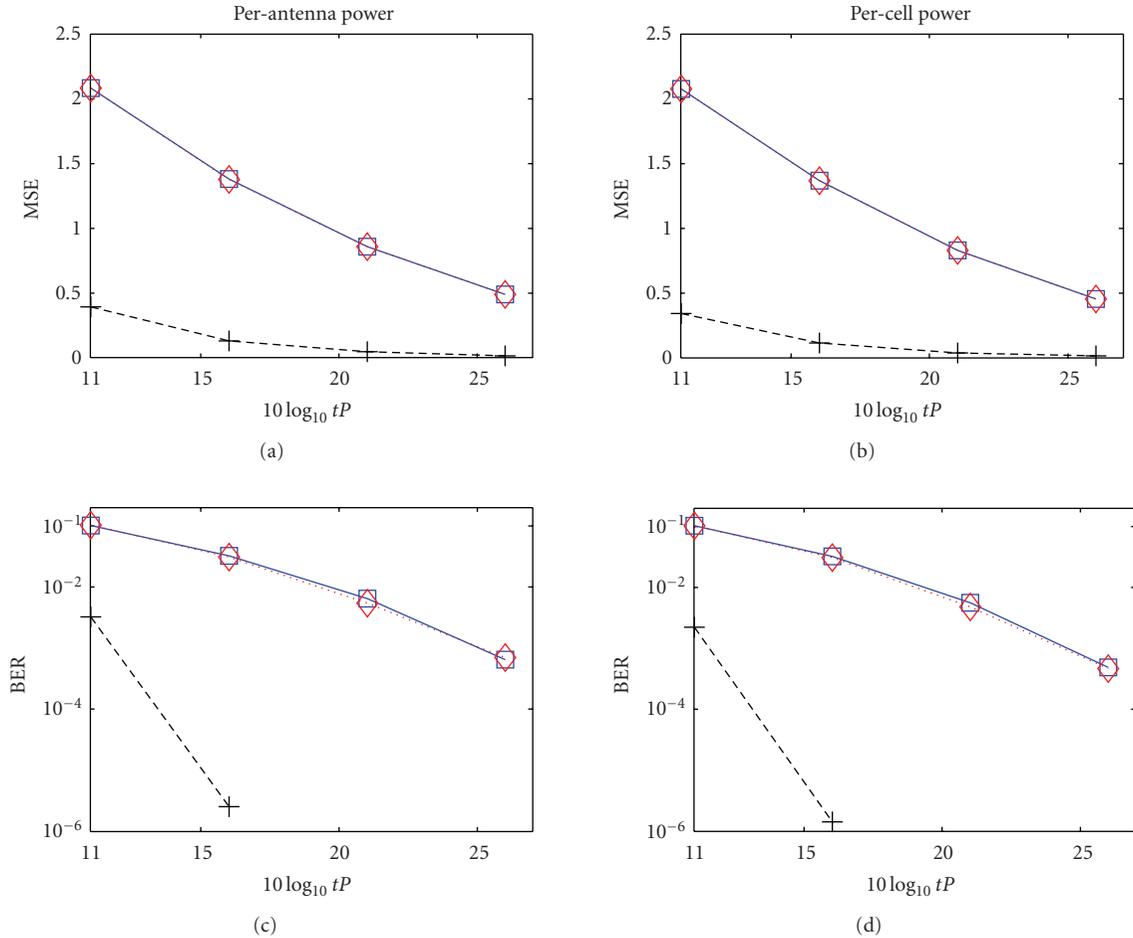


FIGURE 2: MSE and BER as functions of $10\log_{10}tP$ under per-antenna and per-cell power constraints. System parameters are shown in examples 1 and 2 in Table 1. (Legends: Solid square lines: the DCOA with $(m_1, m_2) = (2, 2)$; Dotted diamond lines: the GIA (FIA) with $(m_1, m_2) = (2, 2)$; Dash “+” lines: the GIA with $(m_1, m_2) = (1, 1)$.)

TABLE 1: System parameters.

Example	1	2	3	4
No. of cells (eNBs): C	1	1	3	3
No. of Tx Antennas per eNB: $t_c, c = 1, \dots, C$	4	4	2	4
Total No. of Tx Antennas of the System: t	4	4	6	12
Number of UEs associated per eNB	2	2	1	2
Total No. of UEs of the System: K	2	2	3	6
No. of Rx Antennas per UE: $r_j (j = 1, \dots, K)$	2	2	2	2
No. of data streams per UE: $m_j (j = 1, \dots, K)$	2	1	2	2

(shown in Figure 1(a)), only one UE is associated with each eNB, and therefore, there are three UEs in total ($K = 3$). In example 4 (shown in Figure 1(b)), two UEs are associated with each eNB and there are 6 UEs in total ($K = 6$).

In the simulation, no CSI feedback error is assumed. The only CSI error is the CSI estimation error. $\sigma_a^2 = 1$ and T_c , the total transmission power of the c th eNBs training signal used for channel estimation, is the same as the total transmission power of the data signal, $t_c P$. Three types of designs (perfect-CSI, robust, and non-robust) will be performed. Take the

family of *generalized iterative approaches* as an example. For the perfect-CSI design (denoted as the GIA), there is no CSI estimation error and the perfect CSI is employed for the joint MMSE design of precoders and decoders. On the other hand, there are CSI estimation errors for the non-robust design, *Non-robust-GIA*, and the robust design, *Robust-GIA*; only an estimated CSI is available to them. The difference between the non-robust and robust designs is simple; the non-robust design is unaware that the CSI it has is estimated and thus treats it as if it were perfect while the robust design is aware and thus incorporates the statistics of the CSI estimation error and the CSI into its design.

5.1. Equivalence of the Various Proposed Approaches. Without loss of generality, we will numerically show the equivalence of the *Robust-GIA*, *Robust-FIA*, and *Robust-DCOA* when the CSI is perfect (recall that the *Robust-GIA*, *Robust-FIA*, and *Robust-DCOA* are actually the *GIA*, *FIA*, and *DCOA*, resp., when the CSI is perfect!). To this end, consider two single-cell examples: examples 1 and 2 of Table 1. Also, for convenience, consider $d_{ic} = 1$, $\mathbf{R}_{T,c} = \mathbf{I}_{t_c}$, and $\mathbf{R}_{R,i} = \mathbf{I}_{r_i}$ for $i = 1, 2$ and $c = 1$.

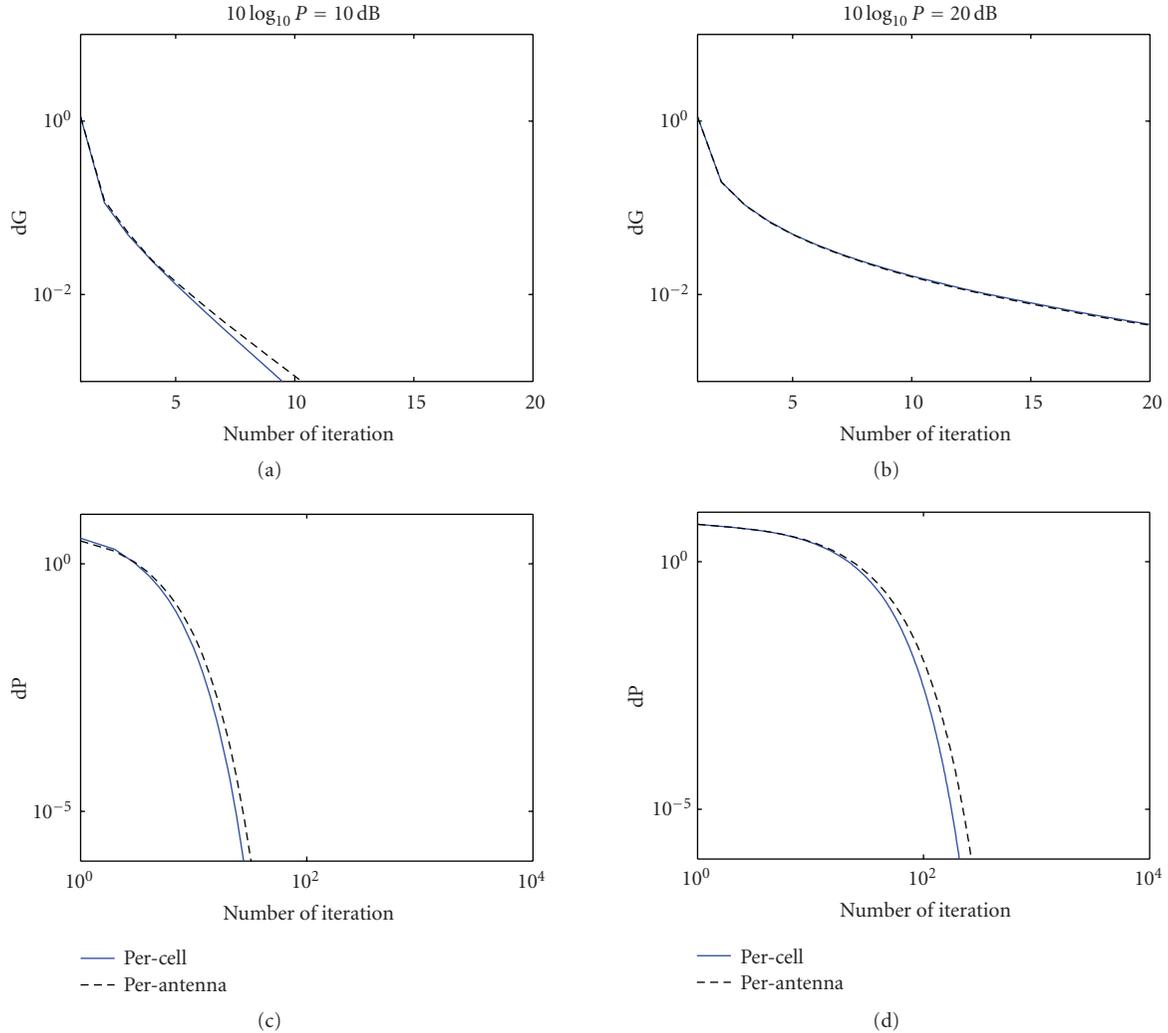


FIGURE 3: Convergence of the *GIA* (*FIA*): the differences dG and dP , defined in (42), as functions of the number of iterations under the per-antenna and per-cell power constraints. System parameters are given in example 1 of Table 1.

Figure 2 shows the MSE's and BER's as functions of the sum power of the system, that is, $10 \log_{10} tP$ where $t = 4$. The results are obtained by averaging over 20 channel realizations. When two data streams are transmitted for both users (i.e., example 1 in Table 1 where $(m_1, m_2) = (2, 2)$ and $\{\mathbf{V}_i\}$ are full rank), both the *GIA* (or the *FIA*) and the *DCOA* can be employed to find the globally optimum precoders and decoders. Comparing the two 2-data-stream curves in each subplot of Figure 2, we observe that the *GIA* (or the *FIA*) indeed has the same globally optimum performance as the *DCOA*. It is remarkable that the performances for the per-cell and per-antenna power constraints are similar to each other.

For the 1-data-stream scenario (i.e., example 2 in Table 1 where $(m_1, m_2) = (1, 1)$ and $\{\mathbf{V}_i\}$ are not full rank), only the *GIA* can be employed because both the *DCOA* and the *FIA* result in $(m_1, m_2) = (2, 2)$ and thus are not applicable here. Comparing the 1-data-stream curve against the 2-data-stream curves in each subplot of Figure 2, the MSE and BER performances, as predicted, are improved by transmitting

fewer data streams than transmit antennas. But the increase in diversity gain is accompanied by a reduction in the multiplexing gain. For the 1-data-stream scenario, only the local optimality of the *GIA* can be guaranteed.

5.2. Computational Efficiency: The *GIA* (*FIA*) versus the *DCOA*. Without loss of generality, we will compare the computational efficiency of the various proposed approaches with perfect CSI. Consider example 1 in Table 1. The number of data streams is two for each of the two UEs so that the *DCOA* can be applicable. For convenience, we also choose $d_{ic} = 1$, $\mathbf{R}_{T,c} = \mathbf{I}_t$, $\mathbf{R}_{R,i} = \mathbf{I}_r$ for $i = 1, 2$ and $c = 1$.

Note that the *GIA* and the *FIA* have the same convergence property because the *FIA* is a special case of the *GIA* when the source covariance matrices are all identity matrices multiplied by the same constant. The *FIA* is slightly more efficient than the *GIA* because it combines, into one step, two of the three steps in each iteration of the *GIA*. The convergence property (expressed as dG and dP) of the *GIA* (or the *FIA*) for both per-antenna and per-cell power

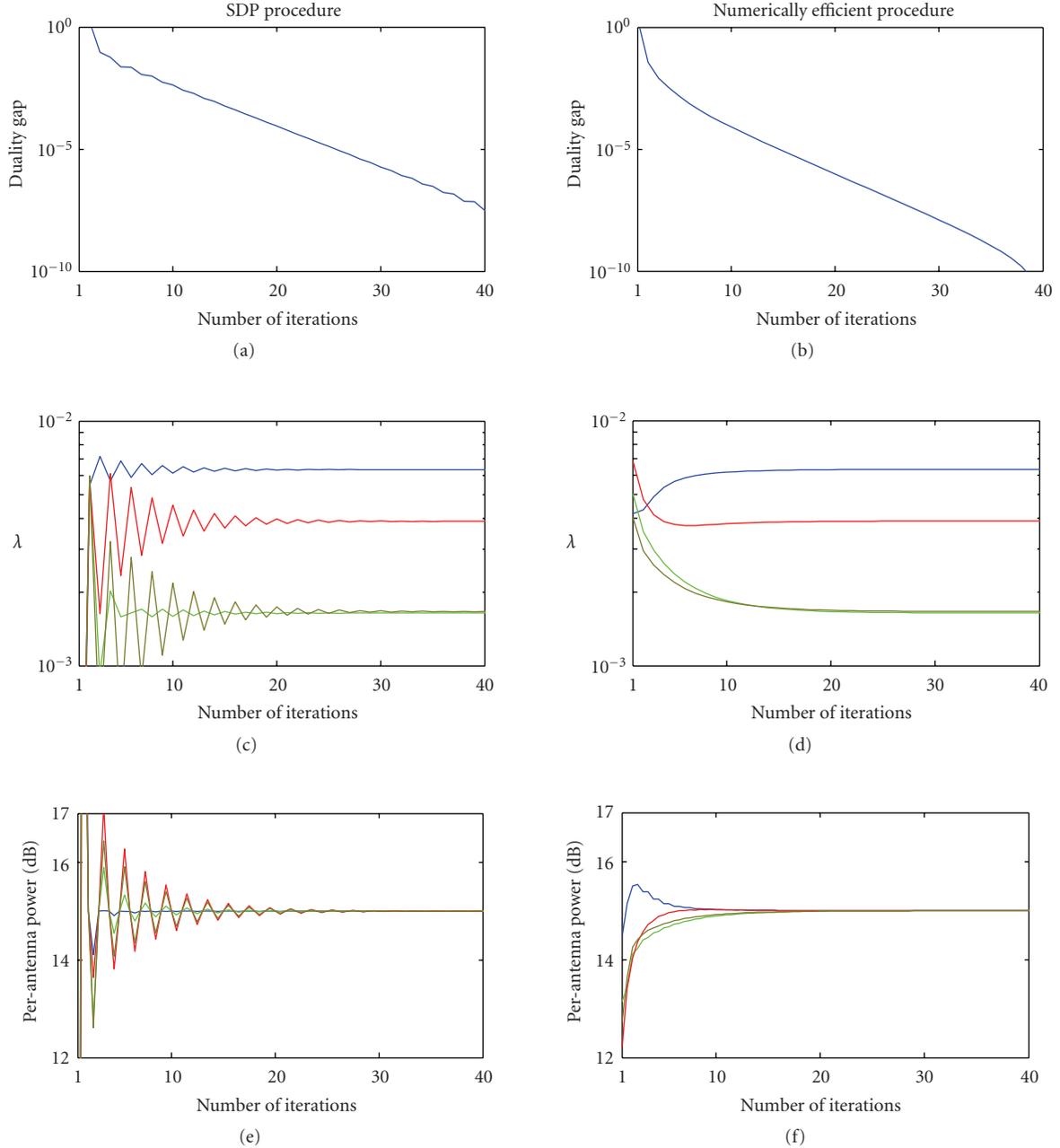


FIGURE 4: Convergence of the DCOA: duality gap, four Lagrange multipliers, and four per-antenna transmission power, as functions of the number of iterations under the per-antenna power constraint with $10 \log_{10} P = 15$ dB. (Legend: the 4 different colors in the two middle subplots and two lower subplots correspond to the 4 transmit antennas.)

constraints is shown in Figure 3. The difference in the decoders between the j th iteration and the $(j+1)$ th iteration and the “distance” from the power constraints at the j th iteration are defined as, respectively,

$$\begin{aligned}
 dG^{(j)} &= \sum_{i=1}^2 \left\| \mathbf{G}_i^{(j+1)} - \mathbf{G}_i^{(j)} \right\|, \\
 dP^{(j)} &= \frac{1}{tP} \operatorname{tr} \left(\left| \sum_{i=1}^2 \mathbf{F}_i^{(j)} (\mathbf{F}_i^{(j)})^* - \mathbf{P} \right| \right).
 \end{aligned} \tag{42}$$

In Figure 3, the convergence rates for both power constraints are similar. It is remarkable that the GIA (or the FIA) converges much slower in higher power. This is due to the fact that, when P increases, the Lagrange multipliers λ_1 decrease quickly. For large P 's, the Lagrange multipliers are very small. For example, when $10 \log_{10} P = 30$ dB, the Lagrange multipliers can be as small as 10^{-10} . Under such a situation, the equality power constraints in (7a) and (7b) are difficult to be met because the usage of (24a) or (24b) merely enforces the corresponding complementary slackness

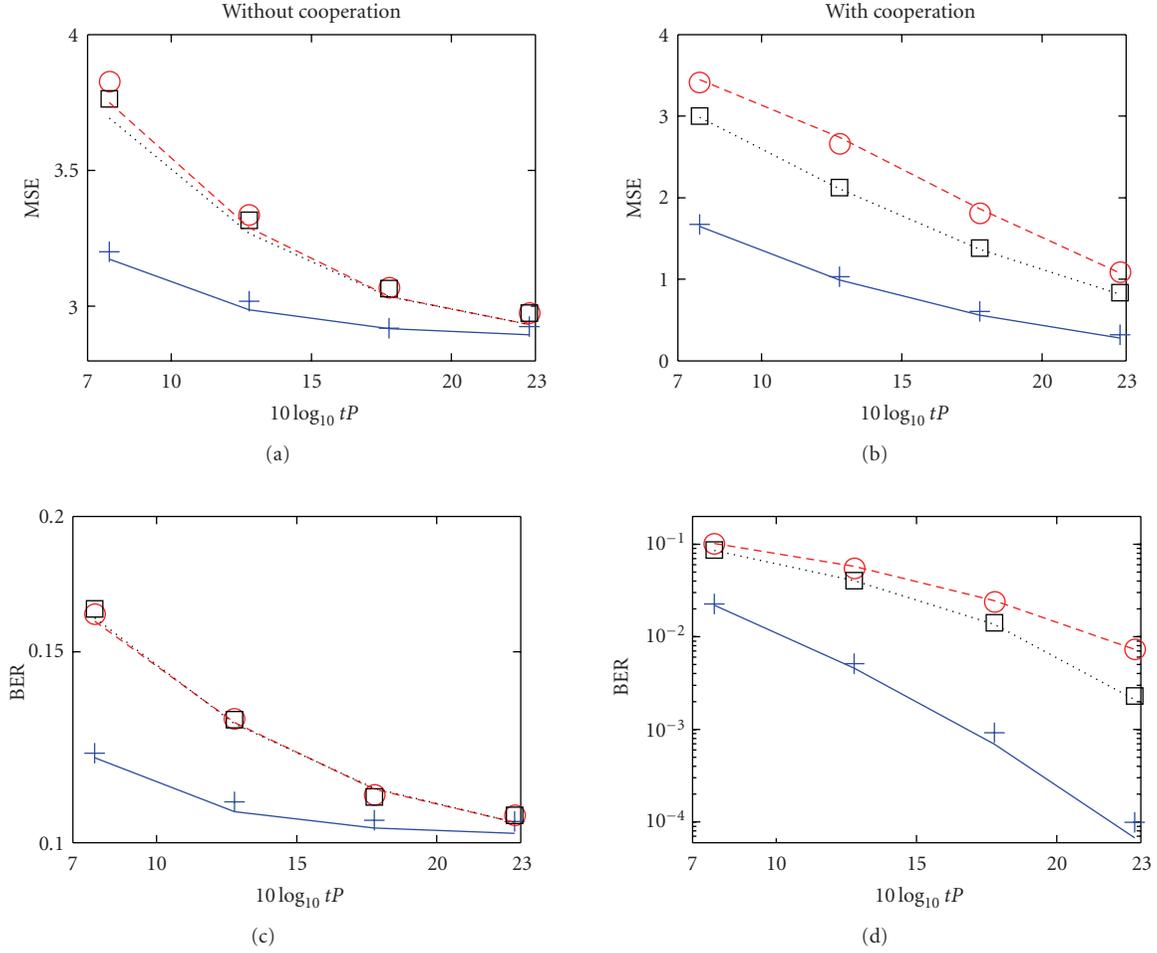


FIGURE 5: MSE and BER as functions of $10 \log_{10} tP$ under per-antenna and per-cell power constraints. System parameters and configurations are shown in example 3 of Table 1 and Figure 1(a). (Legends: the blue solid lines, red dashed lines, and black dotted lines represent, resp., the GIA, the Non-robust-GIA, and the Robust-GIA results under the per-cell power constraint. And the blue plus markers, red circle markers, and black square markers represent, resp., the GIA, the Non-robust-GIA, and the Robust-GIA results under the per-antenna power constraint.)

conditions:

$$\Lambda \left[\mathbf{I}_t * \left(\sum_{j=1}^2 \mathbf{F}_j \Phi_{s_j} \mathbf{F}_j^* \right) - \mathbf{P} \right] = 0, \quad (43a)$$

$$\lambda_1 \left[\text{tr} \left(\sum_{j=1}^2 \mathbf{F}_j \Phi_{s_j} \mathbf{F}_j^* \right) - P_{b1} \right] = 0. \quad (43b)$$

Thus, the number of iteration increases drastically as P increases if equality in the power constraints in (7a) or (7b) is insisted. However, if the equality constraints are relaxed and only inequality constraints (the per-antenna or per-cell transmission powers are allowed to be less than the corresponding power constraints) are required, the convergence rate at high power will be improved greatly.

Using the same single-cell example, the convergent properties of the SDP Procedure and the Numerically Efficient Procedure of the DCOA are shown in Figure 4. Here, $10 \log_{10} P = 15$ dB and, for convenience, only the per-antenna power constraint is considered. Observing the

convergence rates of the duality gap in (34), the Lagrange multipliers in (9a), and the per-antenna transmission power from Figure 4, we conclude that the Numerical Efficient Procedure converges faster than the SDP procedure.

Comparing the DCOA with the GIA (or the FIA), the GIA (or the FIA) is numerically more efficient than the DCOA. This is because, for the GIA (or the FIA), closed form expressions are available for the precoders, decoders and Lagrange multipliers; but for the DCOA, a numerical optimization procedure has to be carried out to find the decoder covariance matrices in the primal step. Note that, just like the GIA (or the FIA), the number of iterations between the primal and dual steps of the DCOA increases drastically as P increases. This is because the convergence problem due to very small Lagrange multipliers at high power exists for both the DCOA and the GIA (or the FIA). In fact, the DCOA does not even converge at times due to the lack of numerical precision of the numerical solvers used. Thus, both the DCOA and the GIA (or the FIA) have difficulty in convergence at high power. Fortunately, within the practical power range, both the DCOA and the GIA

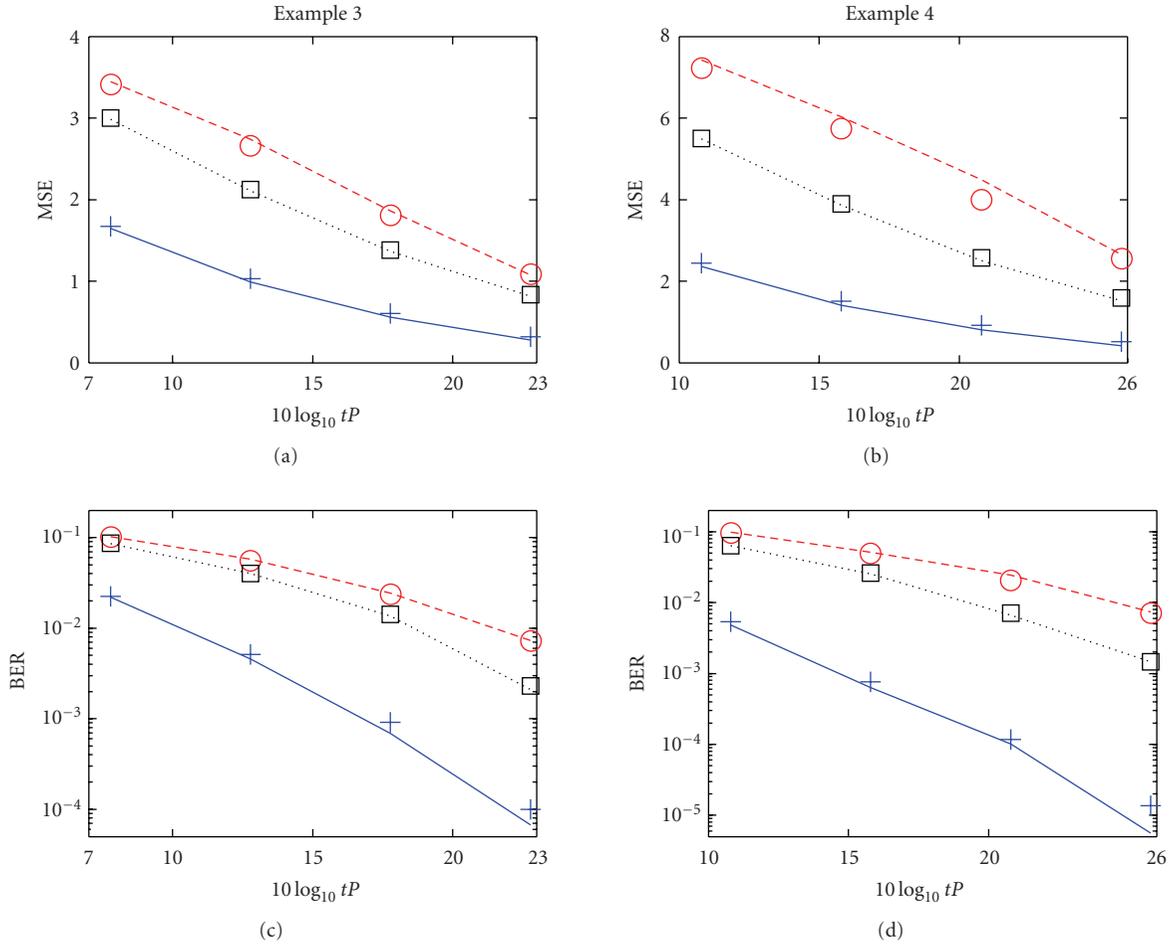


FIGURE 6: MSE and BER as functions of $10 \log_{10} tP$ under per-antenna and per-cell power constraints for examples 3 and 4 in Table 1 and Figure 1. (Legends are the same as in Figure 5.)

(or the *FIA*) worked fine as long as some attention was paid to the selection of the initial values of the iteration process at high power. Note that the robust approaches have only a small increase in complexity compared to their perfect-CSI counterparts and the conclusion made here for complexity analysis is also applicable to the robust approaches.

In the following sections, we will consider the situation where $\{\Phi_{si}\}$ are all identity matrices and $\{\mathbf{V}_i\}$ are full rank. Under such a situation, the *DCOA*, the *FIA*, and the *GIA* are equivalent. Moreover, the *Non-robust-DCOA*, the *Non-robust-FIA*, and the *Non-robust-GIA* are equivalent, and the *Robust-FIA*, and the *Robust-GIA* are also equivalent. If the *Robust-DCOA* is applicable, the *Robust-DCOA*, the *Robust-FIA*, and the *Robust-GIA* are equivalent. Thus, only the *GIA*, the *Robust-GIA*, and the *Non-robust-GIA* results are presented for convenience.

5.3. Multicell: With Cooperation versus without Cooperation. Using the 3-cell configuration in Figure 1(a) and the system parameters of example 3 in Table 1, two different cooperation scenarios will be simulated. In the first scenario, there is no cooperation among the eNBs. In the second scenario,

there is full cooperation among the three eNBs. Note that the interference-plus-noise covariance matrix in (10) needs to be estimated in the without-cooperation scenario. With some derivations, we can show that (10) can be approximated as

$$\Phi_{ni} \approx \sum_{c \notin D_b} \alpha^2 d_{ic}^{-2\beta} P_{bc} \mathbf{R}_{R,i} + \Phi_{ai} \quad (44)$$

for the without-cooperation scenario. For convenience, we choose $\mathbf{R}_{T,c} = \mathbf{I}_{t_c}$ and $\mathbf{R}_{R,i} = \mathbf{I}_{r_i}$ for $i = 1, 2, 3$ and $c = 1, 2, 3$. Channel matrices are estimated and $\hat{\mathbf{H}}_{ic} \neq \mathbf{H}_{ic}$ and $\sigma_{E,ic}^2 \neq 0$. Figure 5 shows the MSE and BER results derived with and without eNB cooperation. All the MSE and BER results are obtained by averaging over 30 channel realizations.

It is not surprising to see that the BER and the MSE of the without-cooperation scenario are much larger (worse) than the BER and the MSE of the with-cooperation scenario, respectively. Even with perfect CSI, the without-cooperation BER is larger than 10% even at high power. It is obvious that some kinds of time/frequency scheduling or code spreading are needed in order to reduce the cell edge interferences if no cooperation among eNBs is available. On the other hand, in the with-cooperation scenario, the BER of the *GIA* is below

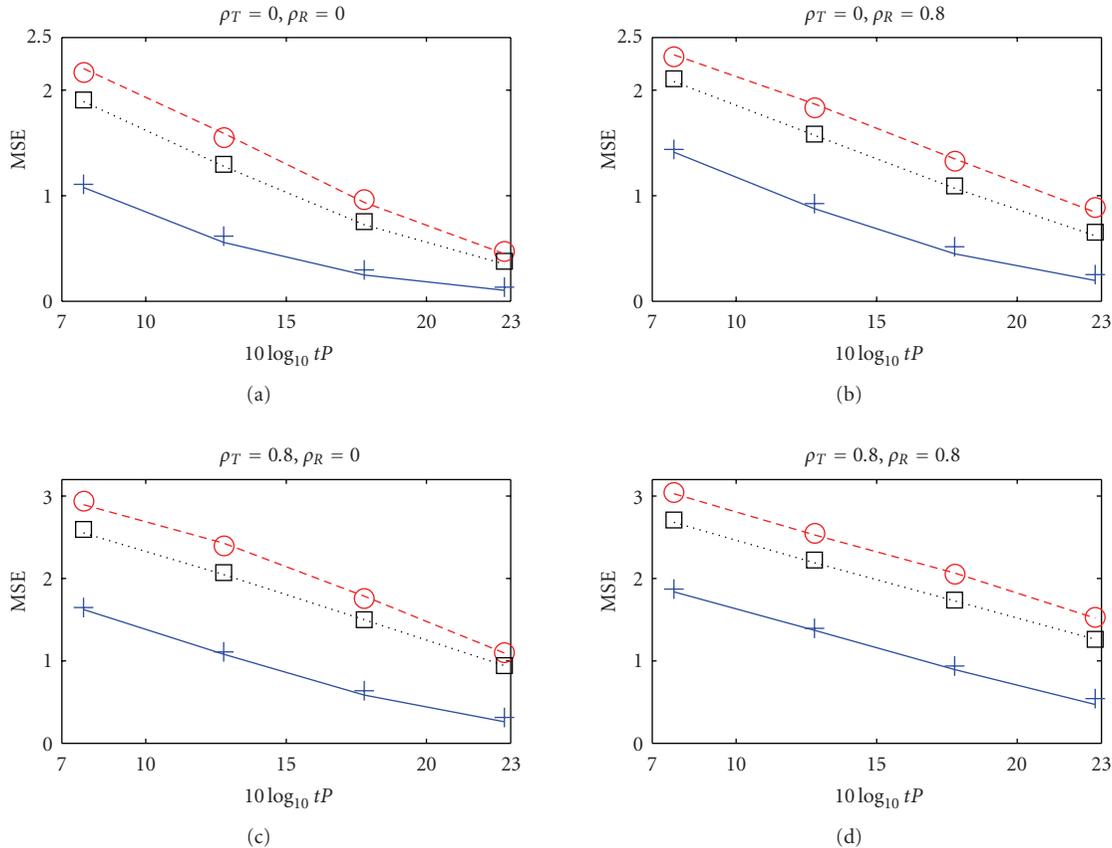


FIGURE 7: MSE's as functions of $10 \log_{10} tP$ under per-antenna and per-cell power constraints for different values of ρ_T and ρ_R . System parameters are shown in example 1 of Table 1. (Legends are the same as in Figure 5.)

1% at the low transmission power ($10 \log_{10} tP = 10 \log_{10} 6 + 5$ dB) when the perfect CSI is available. When the perfect CSI is not available, the *Robust-GIA* result is decent. Even the *Non-robust-GIA* result in the with-cooperation scenario is better than the *GIA* result in the without-cooperation scenario. The *Robust-GIA* result loses around 8 dB in SNR with respect to the *GIA* result and has around a 3 dB gain in SNR with respect to the *Non-robust-GIA*. Note that results obtained from both per-antenna and per-cell power constraints are similar.

5.4. Multicell: Example 3 versus Example 4. We now compare the results of example 3 with the results of example 4 in Table 1. The system configurations of examples 3 and 4 are shown in Figures 1(a) and 1(b), respectively. Note that there is one UE per cell in example 3 but there are two UEs per cell in example 4. For convenience, we choose $\mathbf{R}_{T,c} = \mathbf{I}_{t_c}$ and $\mathbf{R}_{R,i} = \mathbf{I}_{r_i}$ for $i = 1, 2, \dots, K$ and $c = 1, 2, 3$. Channel matrices are estimated and $\hat{\mathbf{H}}_{ic} \neq \mathbf{H}_{ic}$ and $\sigma_{E,ic}^2 \neq 0$. Note that the coordinates of the eNBs and UEs are shown in Figures 1(a) and 1(b); UE₁ is right on the 3-cell edge and each of the other UEs is near at least one of the 2-cell edges.

Figure 6 shows the MSE and BER results of examples 3 and 4 with full cooperation among 3 eNBs. All the MSE and BER results are obtained by averaging over 25 channel

realizations. Note that the average per-antenna power P in examples 3 and 4 is the same. But the total power of example 4 is twice of (3 dB larger than) the total power of example 3 since the t_c in example 4 is twice of the t_c in example 3. Therefore, there is a 3 dB difference in the scales of the x -axes of examples 3 and 4 in Figure 6.

We make four main observations. First, the results for the per-cell and per-antenna power constraints are more or less the same for all of the approaches (the *GIA*, the *Robust-GIA*, and the *Non-robust-GIA*) in both examples 3 and 4. This is remarkable because the per-antenna power constraint, though more practical, is much stricter than the per-cell power constraint.

Secondly, as expected, the *Robust-GIA* yields better MSE and BER performances than the *Non-robust-GIA*. In the power ranges shown in Figure 6, the performance gain of the *Robust-GIA* over the *Non-robust-GIA* for the MSE results is around 5 dB for example 4 and 3 dB for example 3. The performance gain of the *Robust-GIA* over the *Non-robust-GIA* for the BER results is around 2–5 dB for example 4 and 0–3 dB for example 3. The performance gain for the MSE results decreases as power P increases. This is due to the fact that CSI estimation errors decrease as P increases ($T_c = t_c P$). However, the performance gain for the BER results increases as P increases. This is because the BER is expressed in log scale. We conclude that the robust approach is more crucial

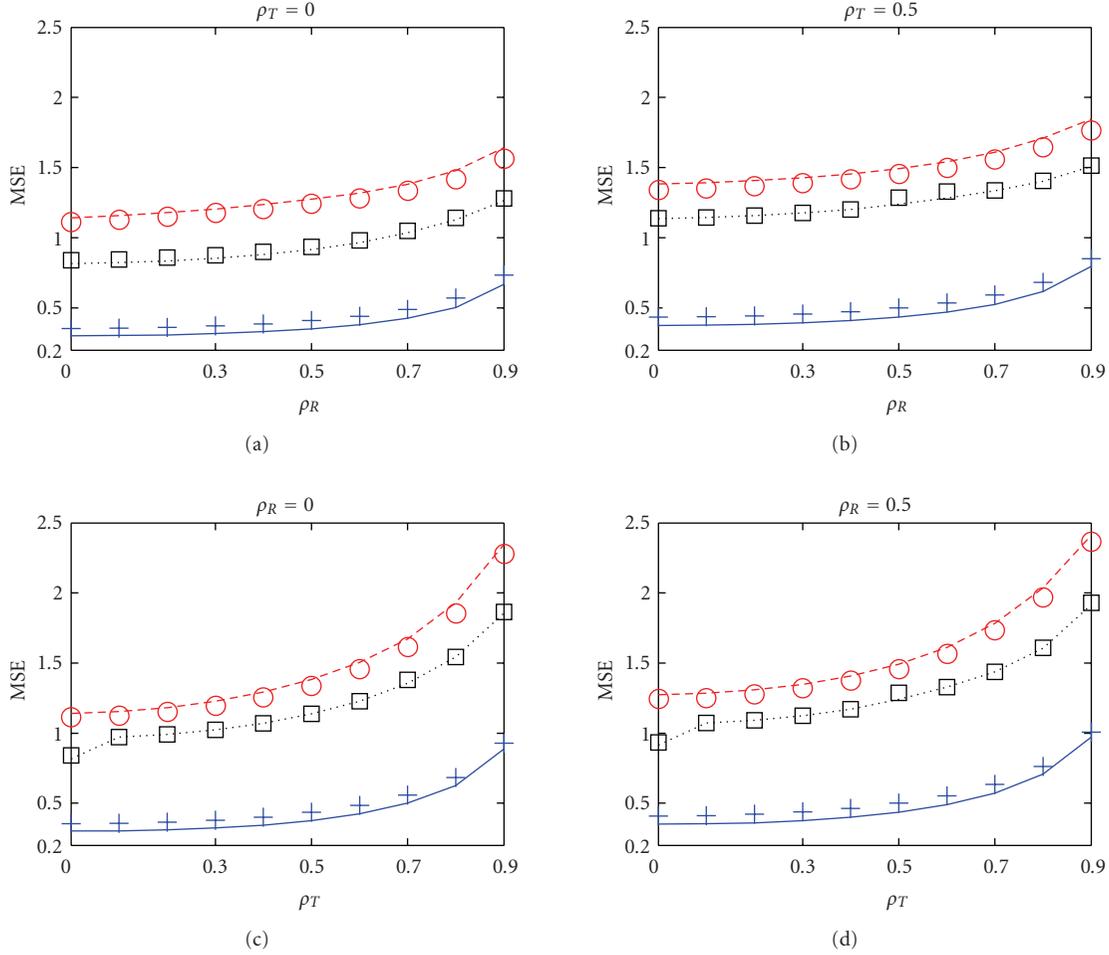


FIGURE 8: MSE as functions of ρ_T (with fixed ρ_R) and functions of ρ_R (with fixed ρ_T) under per-antenna and per-cell power constraints. The total transmission power is 16 dB. System parameters are shown in example 1 of Table 1. (Legends are the same as in Figure 5.)

to larger MIMO systems such as example 4 than smaller MIMO systems such as example 3.

Thirdly, also as expected, the *Robust-GIA* yields larger (worse) MSE and BER than the *GIA*. In the power ranges shown in Figure 6, the performance degradation of the *Robust-GIA* with respect to the *GIA* for the MSE results is around 9–10 dB for example 4 and around 7–8 dB for example 3. The performance degradation of the *Robust-GIA* with respect to the *GIA* for the BER results is around 9–12 dB for example 4 and 9 dB for example 3. For the same reasons stated before, the performance degradation of MSE results decreases as power P increases, but the performance degradation of BER results increases as power P increases. We conclude that larger MIMO systems such as example 4 are more sensitive to the CSI estimation errors than smaller MIMO systems such as example 3.

Lastly, compared to the results in example 3, the MSE results for all the approaches are noticeably higher in example 4, but the degradation of BER results in example 4 compared to example 3 is not significant if the per-antenna power P is the same in both examples. We conclude that cooperation among the eNBs is very effective in mitigating inter-cell interferences at cell edges. And, increasing the

antenna numbers is an effective way to increase the system capacity even at cell edges as long as full eNB cooperation is allowed for the joint design of robust precoders and decoders.

5.5. Spatial Channel Correlations. Using the example 1 in Table 1, system performances of various approaches under different antenna correlation conditions are studied. The channel correlation matrices are defined as

$$\mathbf{R}_{T,c} = \begin{bmatrix} 1 & \rho_T \\ \rho_T & 1 \end{bmatrix}, \quad \mathbf{R}_{R,i} = \begin{bmatrix} 1 & \rho_R \\ \rho_R & 1 \end{bmatrix}, \quad i = 1, 2, \quad c = 1. \quad (45)$$

We choose $d_{1c} = 1$ and $d_{2c} = 0.78$. Channel matrices are estimated and $\hat{\mathbf{H}}_{ic} \neq \mathbf{H}_{ic}$ and $\sigma_{E,ic}^2 \neq 0$.

Figures 7 and 8 show the MSE results for various values of ρ_T and ρ_R . In Figure 7, the MSE is plotted against the sum power; in Figure 8, the MSE is plotted against either ρ_T or ρ_R . The MSE results are obtained by averaging over 20 channel realizations. Again, we observe that the results for the per-cell and per-antenna power constraints are more or less the same for all of the approaches (the *GIA*, the *Robust-GIA*,

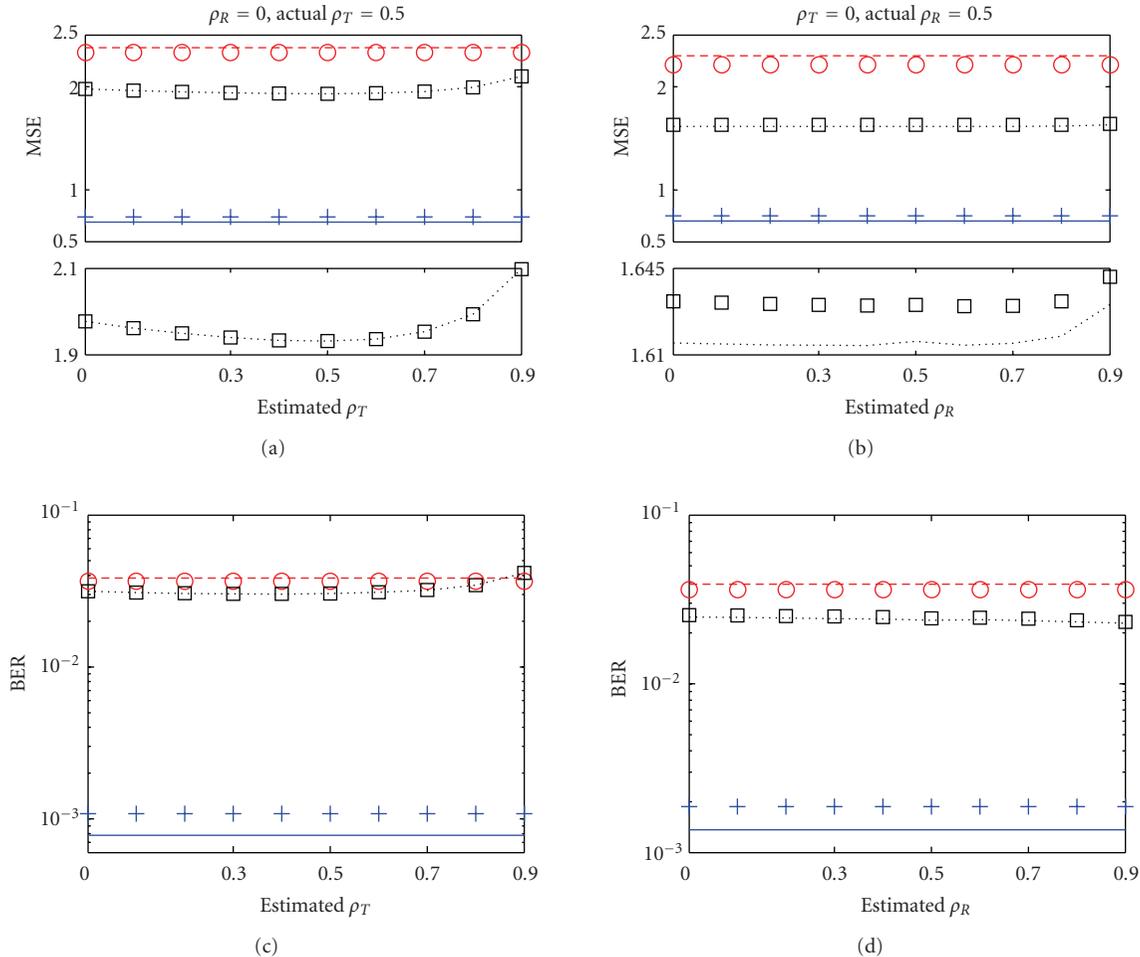


FIGURE 9: MSE and BER as functions of estimated ρ_T or ρ_R under per-antenna and per-cell power constraints ($10 \log_{10} tP = 10 \log_{10} 3 + 10$ dB, averaging over 20 channel realizations). (Legends are the same as in Figure 5.)

and the *Non-robust-GIA*). As ρ_T and/or ρ_R increase, the MSE increases, the performance gain of the *Robust-GIA* over the *Non-robust-GIA* decreases, and the performance degradation of the *Robust-GIA* with respect to the *GIA* increases. The effect due to increasing ρ_T is more profound than the effect due to increasing ρ_R . We conclude that the robust approaches work satisfactorily in wireless channels with high channel correlations.

5.6. Sensitivity with Respect to Estimation Errors of ρ_T or ρ_R . Using the 3-cell configuration in Figure 1(a) and the system parameters of example 3 in Table 1, the sensitivity of MSE and BER performances with respect to estimation errors of ρ_T or ρ_R is studied. Channel matrices are estimated and $\hat{\mathbf{H}}_{ic} \neq \mathbf{H}_{ic}$ and $\sigma_{E,ic}^2 \neq 0$. Full cell cooperation is assumed. In Figure 9, MSE and BER are plotted against the estimated ρ_T for a fixed ρ_R or against the estimated ρ_R for a fixed ρ_T . The enlarged MSE results of the *Robust-GIA* are shown in the middle two subplots. First of all, the *GIA* results are independent of estimation errors of ρ_T and ρ_R because the perfect CSI is employed in the design. Similarly, the *Non-robust-GIA* results are also independent of estimation errors

of ρ_T and ρ_R because channel correlation statistics are not needed in estimating the instantaneous channel matrices in practice. Secondly, the *Robust-GIA* outperforms the *Non-robust-GIA* in terms of both MSE and BER regardless of the estimation error in ρ_T or ρ_R . Thirdly, the performance degradation due to the estimation error in ρ_R (for a fixed ρ_T) is less profound than that due to the estimation error in ρ_T (for a fixed ρ_R). This is because the variance $\sigma_{E,ic}^2$ (see (4b)) of $\mathbf{E}_{W,ic}$ depends only on ρ_T and the accuracy of $\mathbf{R}_{E,ic}$ in (4a) is not significantly affected by the estimation error in ρ_R if SNR is sufficiently large. Lastly, the *Robust-GIA* is less sensitive to underestimates of ρ_T or ρ_R than overestimates of ρ_T or ρ_R . The same observations as above are made from sensitivity studies for various nonzero values of actual ρ_T or ρ_R . We conclude that effects of estimation errors (especially underestimates) of channel correlations ρ_T or ρ_R on the system performances of the *Robust-GIA* are very small.

5.7. Sensitivity with Respect to Estimation Errors of Path Loss. Using the 3-cell configuration in Figure 1(a) and the system parameters of example 3 in Table 1, the sensitivity of MSE and BER performances with respect to estimation

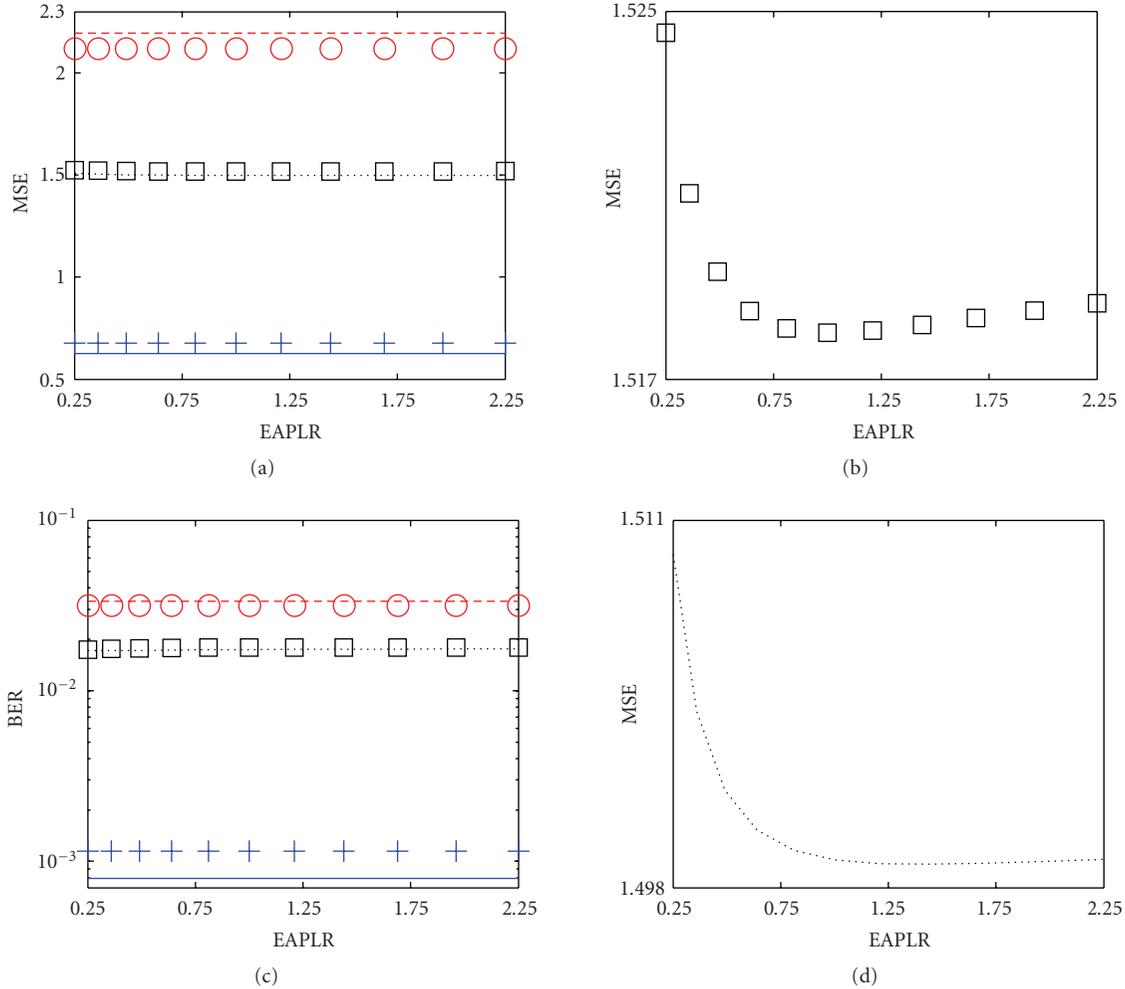


FIGURE 10: MSE and BER as functions of the estimated-to-actual-path loss ratio (EAPLR) under the per-antenna and per-cell power constraints ($10 \log_{10} tP = 10 \log_{10} 3 + 10$ dB, averaging over 20 channel realizations). (Legends are the same as in Figure 5.)

errors of path loss ($PL = \alpha d_{ic}^{-2\beta}$ in (4a)) is studied. Full cell cooperation is assumed. For convenience, we choose $\mathbf{R}_{T,c} = \mathbf{I}_c$ and $\mathbf{R}_{R,i} = \mathbf{I}_i$ for $i = 1, 2, 3$ and $c = 1, 2, 3$. Channel matrices are estimated and $\hat{\mathbf{H}}_{ic} \neq \mathbf{H}_{ic}$ and $\sigma_{E,ic}^2 \neq 0$. Define the estimated-to-actual-path-loss ratio (EAPLR) as $PL_{\text{estimated}}/PL_{\text{actual}}$. In Figure 10, MSE and BER are plotted against the EAPLR ranging from 0.25 to 2.25. The enlarged MSE results of the *Robust-GIA* under the per-antenna and per-cell power constraints are shown in the two right subplots, respectively. First of all, the *GIA* results are independent of estimation errors of path loss because the perfect CSI is employed in the design. Similarly, the *Non-robust-GIA* results are also independent of estimation errors of path loss because channel statistics are not used in estimating the instantaneous channel matrices. Secondly, the *Robust-GIA* outperforms the *Non-robust-GIA* in terms of both MSE and BER regardless of estimation errors of path loss. Thirdly, the degradation of MSE due to estimation errors of path loss is negligible. This is due to the fact that σ_c^2 in (13) is independent of path loss, and the accuracy of $\mathbf{R}_{E,ic}$ in (4a) is not significantly affected by path loss errors if SNR is sufficiently large. We conclude that the effects of estimation

errors of path loss on the system performances of the *Robust-GIA* are negligible.

6. Conclusion

Three robust approaches, the *Robust-GIA*, the *Robust-FIA*, and the *Robust-DCOA*, are proposed for designing MMSE transceivers in the downlink of multicell multiuser MIMO Systems under general linear equality power constraints and with CSI estimation errors. The *GIA*, the *FIA*, and the *DCOA* (the perfect CSI approaches) are special cases of the *Robust-GIA*, the *Robust-FIA*, and the *Robust-DCOA* thus giving a general framework to deal with both perfect and imperfect CSI! Note that the robust approaches have only a small increase in complexity compared to their perfect-CSI counterparts.

The *Robust-DCOA* always gives optimum solutions but is only applicable when the rank constraints on the precoders are relaxed, the transmit correlation matrix of each user is an identity matrix, and the source covariance matrices are all identity matrices multiplied by the same constant. The statistics of the CSI estimation error also need to be the

same for all users if the power constraints of the users are interdependent. The *Robust-GIA*, on the other hand, has no such restrictions and is the most general among the three proposed robust approaches. It allows tradeoff between diversity and multiplexing gains, which is not possible in the *Robust-DCOA* or the *Robust-FIA*. The multiplexing gains of the *Robust-DCOA* or the *Robust-FIA* are determined by the ranks of the decoder covariance matrices. The *Robust-FIA* is a special case of the *Robust-GIA*. It, requiring that the source covariance matrices are identity matrices multiplied by the same constant, is a bit less flexible than the *Robust-GIA*. But, it is much more flexible than the *Robust-DCOA* since it does not require all of the transmit correlation matrices to be identity matrices. Both the *Robust-GIA* and the *Robust-FIA* are numerically more efficient than the *Robust-DCOA*. The *Robust-FIA* is slightly more efficient than the *Robust-GIA* because it combines two of the three steps in each iteration of the *Robust-GIA* into one step. All approaches show difficulties in convergence when the transmit power is very high. Relaxing the equality power constraints will improve the numerical efficiency greatly. Both the *Robust-GIA* and the *Robust-FIA* can only guarantee local optimality. But, whenever the *Robust-DCOA* is applicable and all the decoder covariance matrices are full rank, the three robust approaches are actually equivalent (i.e., the *Robust-GIA* and the *Robust-FIA* are also optimum).

MMSE transceiver designs using the three proposed approaches are performed for various single-cell and multicell examples with different system configurations, power constraints, channel spatial correlations, and cooperation scenarios. System performances in terms of MSE and BER are investigated. Important concluding remarks made from these numerical examples are list below. First of all, the robust approaches outperform their non-robust counterparts in most of the numerical simulations (even when the channel is highly correlated, when the CSI estimation errors are large, and when there exist estimation errors in statistics of channel parameters). Secondly, the performance of the with-cooperation scenario is much better than that of the without-cooperation scenario. With cell cooperation, the cell edge UEs perform as well as those UEs without inter-cell interferences and therefore the cell edge difficulties can be remedied. Thus, with full cell cooperation, the system throughput can increase linearly with the numbers of antennas for both transmission and reception. Thirdly, the robust approaches are insensitive to the estimation errors of the channel statistics (e.g., to channel correlations and path loss). This important feature makes robust approaches practical. Fourthly, the system performances derived under the more practical per-antenna power constraint are very similar to those with the per-cell power constraint. Thus, the practical per-antenna power constraint inflicts little performance losses compared to the optimum per-cell power constraint. Fifthly, the performance gain of the robust approaches over the non-robust approaches is more profound in larger MIMO systems. Sixthly, the performance gain of the robust approaches over the non-robust approaches is reduced if the channel correlations increase.

In short, we have herein proposed, for joint MMSE transceiver designs, three novel robust approaches: the *Robust-GIA* (the most general), *Robust-FIA* (the most efficient), and the *Robust-DCOA* (which guarantees the global optimality). The proposed approaches are indeed robust with respect to different system configurations, CSI estimation errors, channel correlations, and channel modeling errors. When cell cooperation is available, the robust approaches provide a remedy for solving the cell edge problem without reducing the number of data streams.

Acknowledgment

The authors would like to thank InterDigital Communications Corporation for its financial support.

References

- [1] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: a unified framework for convex optimization," *IEEE Transactions on Signal Processing*, vol. 51, no. 9, pp. 2381–2401, 2003.
- [2] M. Codreanu, A. Tölli, M. Juntti, and M. Latva-aho, "Joint design of Tx-Rx beamformers in MIMO downlink channel," *IEEE Transactions on Signal Processing*, vol. 55, no. 9, pp. 4639–4655, 2007.
- [3] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Transactions on Signal Processing*, vol. 50, no. 5, pp. 1051–1064, 2002.
- [4] S. Serbetli and A. Yener, "Transceiver optimization for multiuser MIMO systems," *IEEE Transactions on Signal Processing*, vol. 52, no. 1, pp. 214–226, 2004.
- [5] Z.-Q. Luo, T. N. Davidson, G. B. Giannakis, and K. M. Wong, "Transceiver optimization for block-based multiple access through ISI channels," *IEEE Transactions on Signal Processing*, vol. 52, no. 4, pp. 1037–1052, 2004.
- [6] I.-T. Lu, "Joint MMSE precoder and decoder subject to arbitrary power constraints for uplink multiuser MIMO systems," in *Proceedings of the 70th IEEE Vehicular Technology Conference (VTC '09)*, Anchorage, Alaska, USA, September 2009.
- [7] I.-T. Lu, J. Li, and E. Lu, "Novel MMSE precoder and decoder designs subject to per-antenna power constraint for uplink multiuser MIMO systems," in *Proceedings of the 3rd International Conference on Signal Processing and Communication Systems (ICSPCS '09)*, Omaha, Neb, USA, 2009.
- [8] J. Zhang, Y. Wu, S. Zhou, and J. Wang, "Joint linear transmitter and receiver design for the downlink of multiuser MIMO systems," *IEEE Communications Letters*, vol. 9, no. 11, pp. 991–993, 2005.
- [9] M. Schubert, S. Shi, E. A. Jorswieck, and H. Boche, "Downlink sum-MSB transceiver optimization for linear multi-user MIMO systems," in *Proceedings of Annual Asilomar Conference on Signals, Systems and Computers*, pp. 1424–1428, Pacific Grove, Calif, USA, October 2005.
- [10] G. Zheng, T.-S. Ng, and K.-K. Wong, "Optimal beamforming for sum-MSE minimization in MIMO downlink channels," in *Proceedings of the 63rd IEEE Vehicular Technology Conference (VTC '06)*, vol. 4, pp. 1830–1834, Melbourne, Canada, May 2006.
- [11] S. Shi, M. Schubert, N. Vucic, and H. Boche, "MMSE optimization with per-base-station power constraints for

- network MIMO systems,” in *Proceedings of IEEE International Conference on Communications (ICC '08)*, pp. 4106–4110, Beijing, China, May 2008.
- [12] I.-T. Lu, “Joint MMSE precoder and decoder design for downlink multiuser MIMO systems with arbitrary transmit power constraints,” in *Proceedings of IEEE Sarnoff Symposium (SARNOFF '09)*, pp. 1–5, Princeton, NJ, USA, March 2009.
- [13] J. Li, I.-T. Lu, and E. Lu, “Optimum MMSE transceiver designs for the downlink of multicell MIMO systems,” in *Proceedings of Military Communications Conference (Milcom '09)*, Boston, Mass, USA, October 2009.
- [14] N. Khaled, G. Leus, C. Desset, and H. De Man, “A robust joint linear precoder and decoder MMSE design for slowly time-varying MIMO channels,” in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '04)*, vol. 4, pp. 485–488, Montreal, Canada, May 2004.
- [15] S. Serbetli and A. Yener, “MMSE transmitter design for correlated MIMO systems with imperfect channel estimates: power allocation trade-offs,” *IEEE Transactions on Wireless Communications*, vol. 5, no. 8, pp. 2295–2304, 2006.
- [16] M. Ding and S. D. Blostein, “MIMO minimum total MSE transceiver design with imperfect CSI at both ends,” *IEEE Transactions on Signal Processing*, vol. 57, no. 3, pp. 1141–1150, 2009.
- [17] X. Zhang, D. P. Palomar, and B. Ottersten, “Robust MAC MIMO transceiver design with partial CSIT and CSIR,” in *Proceedings of the 41st Asilomar Conference on Signals, Systems and Computers*, pp. 324–328, Pacific Grove, Calif, USA, November 2007.
- [18] M. B. Shenoouda and T. N. Davidson, “On the design of linear transceivers for multiuser systems with channel uncertainty,” *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 6, pp. 1015–1024, 2008.
- [19] J. F. Sturm, “Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones,” *Optimization Methods and Software*, vol. 11, no. 1–4, pp. 625–653, 1999.
- [20] J. Löfberg, “YALMIP: a toolbox for modeling and optimization in MATLAB,” in *Proceedings of the IEEE International Symposium on Computer-Aided Control System Design (CACSD '04)*, pp. 284–289, Taipei, Taiwan, September 2004.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

