Robust Joint Transceiver Design for Multiuser MIMO Systems with Calibration Errors

Mohammad Kazemi
EEE Dept., Bilkent University
Ankara, Turkey
kazemi@ee.bilkent.edu.tr

Çağrı Göken
ASELSAN A.Ş
Ankara, Turkey
cgoken@aselsan.com.tr

Tolga M. Duman
EEE Dept., Bilkent University
Ankara, Turkey
duman@ee.bilkent.edu.tr

ICC 2022 - IEEE International Conference on Communications | 978-1-5386-8347-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICC45855.2022.9838519

***Abstract*—We consider the downlink of a multiuser multiple-input multiple-output (MIMO) system operating in the time-division duplexing (TDD) mode. In this mode, assuming reciprocity, the channel coefficients estimated during the uplink channel training are utilized by the base station (BS) in the downlink data transmission. However, due to hardware mismatches, the uplink and downlink channels are not exactly the same, and therefore, there are calibration errors, which degrade the system performance. In this paper, our goal is to provide a transceiver design which has a robust performance under imperfect channel reciprocity. To this end, we first formulate a robust joint precoder and combiner design as a stochastic minimum mean square error (MMSE) optimization problem. Then, employing an alternating optimization approach, we propose an algorithm to obtain the precoding and combining matrices assuming imperfect CSI and calibration errors at both the BS and user sides. Extensive simulation results show that the proposed robust joint precoder/combiner outperforms the existing solutions in the literature.**

***Index Terms*—Multiuser MIMO, precoding, combining, calibration errors, robust design, imperfect CSI.**

I. Introduction

In time-division duplexing (TDD) transmission mode, assuming channel reciprocity, estimated channel coefficients in the uplink channel training are utilized in the downlink precoding by the base station (BS). However, in practice, due to hardware mismatches between the transmitter and receiver sides of the BS and the users, even after calibrations, the uplink channel coefficients are not exactly the same as the downlink ones. This necessitates addressing of the calibration errors in data transmission as these errors may result in significant performance loss in practical systems.

There are some existing results on the effects of calibration errors on multiuser systems in the form of either system performance analysis [1]–[4] or development of robust transmission schemes [5]–[8], all of them assuming single-antenna users. However, utilizing multiple antennas at the user side is vital in many emerging 5G and beyond systems, especially those employing higher frequency bands such as millimeter wave (mm-wave) massive multiple-input multiple-output (MIMO) systems. With multiple antennas, the users need to employ combining techniques at the receiver (also known as receive beamforming) to recover the transmitted data. Both problems of precoder design with fixed combiners and joint design of BS precoder and users' combiners have been considered in the literature, but without taking into account the effects of calibration errors. For the case of fixed combiners, for instance, two precoding schemes are developed based on minimum mean-square error (MMSE) and minimum sum SINR criteria in [9] assuming maximum ratio combining (MRC). Also, a block diagonalization (BD) with a vector perturbation (BD-VP) scheme is combined with the MMSE precoding in [10] in order to have a balance between the interference suppression and the noise enhancement.

For the joint precoder/combiner design, for instance, the authors in [11] employ an iterative (alternating) optimization approach based on a total mean square error (T-MMSE) criterion. In [12], the authors employ geometric mean decomposition (GMD) to improve the performance of zero-forcing vertical Bell-labs layered space-time (VBLAST) detection and zero-forcing dirty paper (ZFDP) transceiver schemes. Antenna selection is considered alongside the joint precoder/combiner design in [13] to lower the BS energy consumption. A mm-wave MIMO system for simultaneous wireless information and power multicast (SWIPM) is considered in [14], which employs a max-min criterion for the joint design while satisfying the energy requirements of the energy harvesting users.

In multiuser MIMO systems with calibration errors, while there are works on the precoder design for single-antenna users, to the best of our knowledge, the transceiver design for multiple-antenna users has not been considered in the literature. In this paper, considering multiple antennas and calibration errors at both the BS and user sides as well as imperfect CSI, we propose a robust joint precoding/combining algorithm, and demonstrate that it significantly improves the system performance compared to the non-robust solutions, while having the same order of complexity.

The rest of the paper is organized as follows. The system model is introduced in Section II. In Section III, we present the robust precoding/combining algorithm. Numerical examples are provided in Section IV, and finally, the paper is concluded in Section V.

II. System Model

We consider downlink of a multiuser MIMO system with a BS equipped with N_t antennas serving K users, with the k-th user having $N_{r,k}$ antennas and r_k data streams. The system

This research was supported by ASELSAN Inc. under the "5G Platform" project which is partly funded by the Scientific and Technological Research Council of Turkey (TUBITAK) under the project number 1160206.

works in the TDD mode, that is, the BS estimates the users' channel coefficients via uplink pilot training, which are then used in the downlink precoding process assuming channel reciprocity.

The received signal at the k-th user can be written as

$$\mathbf{y}_k = \beta \mathbf{H}_{DL,k} \sum_{u=1}^{K} \mathbf{A}_u \mathbf{s}_u + \mathbf{n}_k, \tag{1}$$

where $\mathbf{H}_{DL,k}$ is the $N_{r,k} \times N_t$ matrix of downlink channel coefficients of the k-th user, $\mathbf{A}_u$ is the $N_t \times r_u$ precoding matrix of the u-th user, $\mathbf{s}_u$ is the $r_u \times 1$ vector of transmitted data symbols of the u-th user, and $\mathbf{n}_k$ is the $N_{r,k} \times 1$ complex circularly symmetric AWGN noise vector at the k-th user with zero-mean and covariance matrix $\mathbf{R}_{n,k}$; i.e., $\mathrm{E}[\mathbf{n}_k] = \mathbf{0}_{N_{r,k}\times 1}$ and $\mathrm{E}\left[\mathbf{n}_k \mathbf{n}_k^H\right] = \mathbf{R}_{n,k}, \forall k \in \mathscr{K}$, where $\mathbf{0}_{i\times j}$ denotes an $i \times j$ all-zero matrix and $\mathscr{K}$ is the set of all users. Also, β is a common scaling coefficient known at the users, employed to limit the transmit power to a certain level. We assume that the users' data vectors are independent of each other with zero mean and covariance matrix $\mathbf{R}_{s,k}$; i.e., $\mathrm{E}[\mathbf{s}_k] = \mathbf{0}_{r_k\times 1}$ and $\mathrm{E}\left[\mathbf{s}_k \mathbf{s}_k^H\right] = \mathbf{R}_{s,k}, \forall k \in \mathscr{K}$, and $\mathrm{E}\left[\mathbf{s}_k \mathbf{s}_{k'}^H\right] = \mathbf{0}_{r_k \times r_{k'}}, \forall k \neq k' \in \mathscr{K}$. After rescaling and combining, the received signal at the k-th user becomes

$$\mathbf{y}'_k = \mathbf{W}_k \mathbf{H}_{DL,k} \sum_{u=1}^{K} \mathbf{A}_u \mathbf{s}_u + \beta^{-1} \mathbf{W}_k \mathbf{n}_k, \tag{2}$$

where $\mathbf{W}_k$ is the $r_k \times N_{r,k}$ combining matrix of the k-th user.

Note that due to hardware mismatches, the uplink and downlink channel coefficients are not exactly the same even after calibration, which calls for calibration error aware transmission schemes. After channel calibration, the uplink and downlink channel matrices are modelled as [1]

$$\mathbf{H}_{UL,k} = \mathrm{diag}\left\{\mathbf{c}_{rb}\right\} \mathbf{H}_k^T \mathrm{diag}\left\{\mathbf{c}_{tu,k}\right\}, \tag{3}$$

$$\mathbf{H}_{DL,k} = \mathrm{diag}\left\{\mathbf{c}_{ru,k}\right\} \mathbf{H}_k \,\mathrm{diag}\left\{\mathbf{c}_{tb}\right\}, \tag{4}$$

where $\mathbf{H}_k$ is the $N_{r,k} \times N_t$ matrix of physical channel coefficients (air medium) between the BS and the k-th user, $\mathbf{c}_{tb}$ and $\mathbf{c}_{rb}$ are the $N_t \times 1$ vectors of calibration errors of the transmitter and receiver parts of the BS antennas, respectively, and $\mathbf{c}_{tu,k}$ and $\mathbf{c}_{ru,k}$ are the $N_{r,k} \times 1$ vectors of calibration errors of the transmitter and receiver parts of the k-th user's antennas, respectively. We assume imperfect CSI at the BS; i.e., the BS only knows estimates of the channel coefficients as

$$\hat{\mathbf{H}}_{UL,k} = \mathbf{H}_{UL,k} + \tilde{\mathbf{H}}_{UL,k}, \tag{5}$$

where $\tilde{\mathbf{H}}_{UL,k}$ is the matrix of channel estimation errors of the k-user, with independent and identically distributed (i.i.d.) entries with zero mean and variance σ_h^2. Combining (3), (4) and (5), we have

$$\mathbf{H}_{DL,k} = \mathbf{C}_{u,k} \left(\hat{\mathbf{H}}_{UL,k}^T - \tilde{\mathbf{H}}_{UL,k}^T\right) \mathbf{C}_b, \tag{6}$$

where $\mathbf{C}_b \coloneqq \mathrm{diag}\left\{\mathbf{c}_b\right\}$ and $\mathbf{C}_{u,k} \coloneqq \mathrm{diag}\left\{\mathbf{c}_{u,k}\right\}$, with $\mathbf{c}_b \coloneqq \mathbf{c}_{tb} \oslash \mathbf{c}_{rb}$ and $\mathbf{c}_{u,k} \coloneqq \mathbf{c}_{ru,k} \oslash \mathbf{c}_{tu,k}$ representing the ratio of the calibration errors at Rx and Tx of the BS and the k-th user, respectively, with means of $\mathbf{m}_b$ and $\mathbf{m}_{u,k}$ (i.e., $\mathrm{E}[\mathbf{c}_b] \coloneqq \mathbf{m}_b$ and $\mathrm{E}\left[\mathbf{c}_{u,k}\right] \coloneqq \mathbf{m}_{u,k}$) and correlation matrices of $\mathbf{R}_b$ and $\mathbf{R}_{u,k}$ (i.e., $\mathrm{E}\left[\mathbf{c}_b \mathbf{c}_b^H\right] \coloneqq \mathbf{R}_b$ and $\mathrm{E}\left[\mathbf{c}_{u,k} \mathbf{c}_{u,k}^H\right] \coloneqq \mathbf{R}_{u,k}$), respectively, where $\oslash$ denotes the Hadamard (point-wise) division.

We assume that the BS (or any node that performs the joint optimization) knows the estimated uplink channel matrix ($\hat{\mathbf{H}}_{UL,k}$), the mean vectors and covariance matrices of the calibration errors at the BS ($\mathbf{m}_b$ and $\mathbf{R}_b$) and the users ($\mathbf{m}_{u,k}$ and $\mathbf{R}_{u,k}$), in addition to the covariance matrices of noise ($\mathbf{R}_{n,k}$) and user data ($\mathbf{R}_{s,k}$). This is a practical assumption in the sense that all these vectors and matrices vary much more slowly than the channel coefficients in practice.

III. Robust Precoder/Combiner Design

In this section, we design a robust joint precoder/combiner against the calibration errors and imperfect CSI. To this end, we formulate the following stochastic optimization problem with the sum of MSEs at the users as the cost function:

$$\begin{aligned} &(\mathbf{A}_j, \mathbf{W}_j, \beta) = \underset{(\mathbf{A}_j, \mathbf{W}_j, \beta)}{\arg\min} \sum_{k=1}^{K} \varepsilon_k, \quad \forall j \in \mathscr{K} \\ &s.t. \qquad \beta^2 \sum_{u=1}^{K} \mathrm{tr}\left\{\mathbf{A}_u \mathbf{R}_{s,u} \mathbf{A}_u^H\right\} = P_T, \end{aligned} \tag{7}$$

where P_T is the total transmit power at the BS, and $\varepsilon_k \coloneqq \mathrm{E}_{n_k, s_k, \tilde{H}_{UL,k}, C_b, C_{u,k}}\left[\left\|\mathbf{y}'_k - \mathbf{s}_k\right\|^2\right]$ is the MSE at the k-th user. In other words, in order to have the best estimates of the transmitted data, the sum of the MSEs at the users is jointly minimized over the precoding and combining matrices, and the scaling coefficient, with the only constraint of the transmit power being fixed to a certain value.

To solve the optimization problem in (7), we first obtain a closed-form expression for the MSE at the k-th user, which is given in the following proposition.

Proposition 1. *The closed-form expression for the MSE at the k-th user in the presence of calibration errors and imperfect CSI can be obtained as*

$$\begin{aligned} \varepsilon_k =& \mathrm{tr}\left\{\mathbf{R}_{s,k}\right\} - \mathrm{tr}\left\{\mathbf{R}_{s,k} \mathbf{A}_k^H \mathbf{M}_b^H \hat{\mathbf{H}}_{UL,k}^* \mathbf{M}_{u,k}^H \mathbf{W}_k^H\right\} \\ &- \mathrm{tr}\left\{\mathbf{W}_k \mathbf{M}_{u,k} \hat{\mathbf{H}}_{UL,k}^T \mathbf{M}_b \mathbf{A}_k \mathbf{R}_{s,k}\right\} \\ &+ \mathrm{tr}\left\{\mathbf{W}_k \mathbf{D}_{c,k} \mathbf{W}_k^H\right\} + \beta^{-2} \mathrm{tr}\left\{\mathbf{W}_k \mathbf{R}_{n,k} \mathbf{W}_k^H\right\}, \end{aligned} \tag{8}$$

where $\mathbf{M}_b \coloneqq \mathrm{diag}\left\{\mathbf{m}_b\right\}$, $\mathbf{M}_{u,k} \coloneqq \mathrm{diag}\left\{\mathbf{m}_{u,k}\right\}$ *and*

$$\begin{aligned} \mathbf{D}_{c,k} \coloneqq& \mathbf{R}_{u,k} \odot \left(\hat{\mathbf{H}}_{UL,k}^T \left(\mathbf{R}_b \odot \left(\sum_{u=1}^{K} \mathbf{A}_u \mathbf{R}_{s,u} \mathbf{A}_u^H\right)\right) \hat{\mathbf{H}}_{UL,k}^*\right) \\ &+ \sigma_h^2 \mathrm{tr}\left\{\mathbf{R}_b \odot \left(\sum_{u=1}^{K} \mathbf{A}_u \mathbf{R}_{s,u} \mathbf{A}_u^H\right)\right\} \left(\mathbf{R}_{u,k} \odot \mathbf{I}_{N_{r,k}}\right), \end{aligned} \tag{9}$$

where $\odot$ denotes the Hadamard (point-wise) product.

Also, the MSE at the k-th user can be equivalently given as

$$\begin{aligned}\varepsilon_k =&\mathrm{tr}\left\{\mathbf{R}_{s,k}\right\} - \mathrm{tr}\left\{\mathbf{M}_b^H\hat{\mathbf{H}}_{UL,k}^*\mathbf{M}_{u,k}^H\mathbf{W}_k^H\mathbf{R}_{s,k}\mathbf{A}_k^H\right\} \\ &- \mathrm{tr}\left\{\mathbf{A}_k\mathbf{R}_{s,k}\mathbf{W}_k\mathbf{M}_{u,k}\hat{\mathbf{H}}_{UL,k}^T\mathbf{M}_b\right\} \\ &+ \mathrm{tr}\left\{\mathbf{D}_{p,k}\sum_{u=1}^{K}\mathbf{A}_u\mathbf{R}_{s,u}\mathbf{A}_u^H\right\} + \beta^{-2}\mathrm{tr}\left\{\mathbf{W}_k\mathbf{R}_{n,k}\mathbf{W}_k^H\right\},\end{aligned} \tag{10}$$

where

$$\begin{aligned}\mathbf{D}_{p,k} :=&\mathbf{R}_b^T\odot\left(\hat{\mathbf{H}}_{UL,k}^*\left(\mathbf{R}_{u,k}^T\odot\left(\mathbf{W}_k^H\mathbf{W}_k\right)\right)\hat{\mathbf{H}}_{UL,k}^T\right) \\ &+\sigma_h^2\mathrm{tr}\left\{\mathbf{W}_k\left(\mathbf{R}_{u,k}\odot\mathbf{I}_{N_{r,k}}\right)\mathbf{W}_k^H\right\}\left(\mathbf{R}_b\odot\mathbf{I}_{N_t}\right).\end{aligned} \tag{11}$$

Proof: The proof is given in Appendix A. ■

With the closed-form MSE expressions in Proposition 1, the main optimization in (7) turns into a non-stochastic problem, which can be solved by employing standard optimization techniques. Next, to make the optimization problem more tractable, we break it into two sub-problems, and utilize an alternating optimization approach, which separates the precoder and combiner design problems, and operates in an iterative manner.

A. Precoder Design

Employing the alternating optimization approach, we first fix the combining matrices and solve for the precoding matrices and the scaling coefficient. That is, for given combining matrices, we formulate the joint precoding and scaling coefficient optimization problem as:

$$\begin{aligned}(\mathbf{A}_j,\beta) &= \underset{(\mathbf{A}_j,\beta)}{\arg\min}\sum_{k=1}^{K}\varepsilon_k \quad \forall j\in\mathscr{K} \\ s.t. \quad &\beta^2\sum_{u=1}^{K}\mathrm{tr}\left\{\mathbf{A}_u\mathbf{R}_{s,u}\mathbf{A}_u^H\right\} = P_T,\end{aligned} \tag{12}$$

with a closed-form solution given in the following proposition.

Proposition 2. *As a solution to the robust optimization problem in* (12)*, the precoding matrix of the j-th user is obtained as*

$$\mathbf{A}_j=\left(\mathbf{D}_p+\frac{1}{P_T}\sum_{k=1}^{K}\mathrm{tr}\{\mathbf{W}_k^H\mathbf{W}_k\mathbf{R}_{n,k}\}\mathbf{I}_{N_t}\right)^{-1}\mathbf{M}_b^*\hat{\mathbf{H}}_{UL,j}^*\mathbf{M}_{u,j}^*\mathbf{W}_j^H, \tag{13}$$

where $\mathbf{D}_p := \sum_{k=1}^{K}\mathbf{D}_{p,k}$. Also, for β, we have

$$\beta=\sqrt{P_T\left(\mathbf{A}_j\mathbf{R}_{s,j}\mathbf{A}_j^H\right)^{-1}}. \tag{14}$$

Proof: The proof is given in Appendix B. ■

Based on Proposition 2, the proposed precoder is an extended version of the MMSE precoder, which takes the effects of calibration errors and imperfect CSI into account in addition to noise. Also, note that in the precoding expression (13), the matrix to be inverted is independent of the user index, which reduces the computational complexity.

B. Combiner Design

We next fix the precoding matrices $\mathbf{A}_j$ and β, and solve for the combining matrices $\mathbf{W}_j$, which turns the optimization problem in (7) into the following unconstrained one

$$\mathbf{W}_j = \underset{\mathbf{W}_j}{\arg\min}\sum_{k=1}^{K}\varepsilon_k, \quad \forall j\in\mathscr{K}, \tag{15}$$

with a closed-form solution given in the following proposition.

Proposition 3. *Letting β and the precoding matrices $\mathbf{A}_j$ to be fixed, the optimal combining matrix (the solution to the optimization problem in* (15)*) is obtained as*

$$\mathbf{W}_j=\mathbf{R}_{s,j}\mathbf{A}_j^H\mathbf{M}_b^*\hat{\mathbf{H}}_{UL,j}^*\mathbf{M}_{u,j}^*\left(\mathbf{D}_{c,j}+\beta^{-2}\mathbf{R}_{n,j}\right)^{-1}, \tag{16}$$

where $\mathbf{D}_{c,j}$ and β are given in (9) *and* (14)*, respectively.*

Proof: Using (8), the derivative of the cost function of (15) with respect to the combining matrix $\mathbf{W}_j$ becomes

$$\frac{\partial}{\partial\mathbf{W}_j}\sum_{k=1}^{K}\varepsilon_k=-\mathbf{R}_{s,j}^T\mathbf{A}_j^T\mathbf{M}_b\hat{\mathbf{H}}_{UL,j}\mathbf{M}_{u,j}+\mathbf{W}_j^*\left(\mathbf{D}_{c,j}+\beta^{-2}\mathbf{R}_{n,j}\right)^T, \tag{17}$$

where we have used the fact that $\frac{\partial}{\partial\mathbf{A}}\mathrm{tr}\{\mathbf{XA}\}=\mathbf{X}^T$, $\frac{\partial}{\partial\mathbf{A}}\mathrm{tr}\{\mathbf{XA}^H\}=\mathbf{0}$, and $\frac{\partial}{\partial\mathbf{A}}\mathrm{tr}\{\mathbf{X}_1\mathbf{A}\mathbf{X}_2\mathbf{A}^H\}=\mathbf{X}_1^T\mathbf{A}^*\mathbf{X}_2^T$ [15].

Finally, setting the derivative to zero gives us the combining matrix in Proposition (3). ■

Similar to what we observed in Proposition 2, the proposed combiner in Proposition 3 is an extended version of the MMSE combiner, which takes into account the effects of calibration errors and the imperfect CSI as well as noise. However, contrary to the precoding expression, the matrix to be inverted in the combining expression (16) is user dependent, i.e., the matrix inversion needs to be performed for each user separately. However, since the matrices to be inverted are $N_{r,k}\times N_{r,k}$, the complexity of these operations is negligible in comparison with the single $N_t\times N_t$ matrix inversion in the precoder design. Note that the non-robust precoder and combiner can be readily obtained by setting $\sigma_h^2=0$, $\mathbf{M}_b=\mathbf{I}_{N_t}$, $\mathbf{M}_{u,k}=\mathbf{I}_{N_{r,k}}$, $\mathbf{R}_b=\mathbf{1}_{N_t\times N_t}$ and $\mathbf{R}_{u,k}=\mathbf{1}_{N_{r,k}\times N_{r,k}}$ in (13) and (16), respectively, where $\mathbf{1}_{i\times j}$ is the $i\times j$ all-one matrix.

C. Proposed Joint Precoding and Combining Algorithm

Combining the results of the previous two subsections, we propose an algorithm to iteratively solve for the optimal precoding and combining matrices as follows. Knowing the initial combining matrices (or those from the previous iteration), the precoding matrices are obtained as in Proposition 2. Next, utilizing the obtained precoding matrices, the combining matrices are obtained as in Proposition 3. This process is repeated until a maximum number of iterations or a performance threshold is reached. The proposed algorithm is summarized in Algorithm 1, where $iter_{\max}$ is the maximum number of iterations, $\mathbf{W}_{init,j}$ is the initial value for the combining matrix, and $\delta_{W,\min}$ is a selected threshold for the difference of the combining matrices in two subsequent iterations. Also, $\mathbf{A}_j^i$, $\mathbf{W}_j^i$ and β^i are the precoding matrix, the combining matrix and the scaling coefficient of the j-th user in the i-th iteration, respectively.

Algorithm 1: Steps of the proposed robust joint transceiver design.

Initialization: Set $i = 1$ and $\mathbf{W}_j^0 = \mathbf{W}_{init,j}, \quad \forall j \in \mathcal{K}$;
while $i \leq iter_{\max}$ & $\delta_W \geq \delta_{W,\min}$ **do**
  1. Knowing $\mathbf{W}_j^{i-1}$, calculate the precoding matrix $\mathbf{A}_j^i$ using (13).
  2. Knowing $\mathbf{A}_j^i$, calculate the combining matrix $\mathbf{W}_j^i$ using (14) and (16)
  3. Stopping criteria: $\delta_W = \max\limits_{j \in \mathcal{K}} \left\| \mathbf{W}_j^i - \mathbf{W}_j^{i-1} \right\|_F$;
  $i = i + 1$;
Calculate the scaling coefficient β^i using (14).

IV. Simulation Results

We now investigate the effects of calibration and channel estimation errors on the system performance via simulations. Note that since there is no robust precoding/combining solutions against the calibration errors in the literature, we only compare our results with those of the non-robust solutions. We have set the default values of the numbers of BS antennas, users, user antennas, and streams as $N_t = 15$, $K = 5$, $N_{r,k} = 3$ and $r_k = 3, \forall k \in \mathcal{K}$, respectively. We have also set $\sigma_h^2 = 0$, $iter_{\max} = 5$ and $\delta_{W,\min} = 10^{-4}$. Moreover, the users' data are i.i.d. and from a binary phase-shift keying (BPSK) constellation, and $\mathbf{R}_{n,k} = \mathbf{I}_{N_{r,k}}, \forall k \in \mathcal{K}$. Also, the uplink channel coefficients are i.i.d. complex Gaussian distributed with zero mean and unit variance.

For the simulations, we assume that the amplitude and phase of the calibration errors at the BS are i.i.d. random variables with log-normal and uniform distributions, respectively; i.e., $[\mathbf{C}_b]_{i,i} = \rho_{b,i} e^{j\phi_{b,i}}$, where $\rho_{b,i} \sim \text{Lognormal}(\mu_{\rho_b}, \sigma_{\rho_b}^2)$ and $\phi_{b,i} \sim \mathcal{U}(-\phi_{\max,b}, \phi_{\max,b})$ [16]. We have similar assumptions for the calibration error at the users; i.e., $\left[\mathbf{C}_{u,k}\right]_{i,i} = \rho_{u,k,i} e^{j\phi_{u,k,i}}$, where $\rho_{u,k,i} \sim \text{Lognormal}(\mu_{\rho_u}, \sigma_{\rho_u}^2)$ and $\phi_{u,k,i} \sim \mathcal{U}(-\phi_{\max,u}, \phi_{\max,u})$ [16]. The default values for the simulations are $\mu_{\rho_b} = \frac{-1}{2}\sigma_{\rho_b}^2$ (to ensure that $\mathrm{E}\left[\rho_{b,i}\right] = 1$), $\mu_{\rho_{u,k}} = \frac{-1}{2}\sigma_{\rho_{u,k}}^2$ (similarly, to ensure that $\mathrm{E}\left[\rho_{u,k,i}\right] = 1$), $\sigma_{\rho_b}^2 = \sigma_{\rho_u}^2 = 0.1$, and $\phi_{\max,b} = \phi_{\max,u} = \frac{\pi}{9}$.

We have compared the performance of the joint precoder/combiner with that of precoding plus an identity combining at the users (i.e., $\mathbf{W}_k$ is an $a \times b$ matrix with ones on its main diagonal and zeros elsewhere) and the ideal case (joint design with no calibration errors). The non-robust algorithm refers to employing Algorithm 1 while ignoring the calibration errors (i.e., setting $\sigma_h^2 = 0$, $\mathbf{M}_b = \mathbf{I}_{N_t}$, $\mathbf{M}_{u,k} = \mathbf{I}_{N_{r,k}}$, $\mathbf{R}_b = \mathbf{1}_{N_t \times N_t}$ and $\mathbf{R}_{u,k} = \mathbf{1}_{N_{r,k} \times N_{r,k}}$). The BER performance of the proposed robust algorithm versus signal-to-noise ratio (SNR) ($SNR := \frac{P_T}{\sigma_n^2} = P_T$) for the single-stream case (i.e., $r_k = 1$) is presented in Fig. 1. As can be seen, the proposed robust solution outperforms the non-robust one. For instance, there is a performance gap of about 3 dB at a BER of 10^{-5}. The BER performance for the multiple-stream case of $r_K = 2$ is presented in Fig. 2. Similar to the single-stream case, the proposed robust solution outperforms the non-robust one, however, this time with a performance gap of about 2 dB at the same BER.

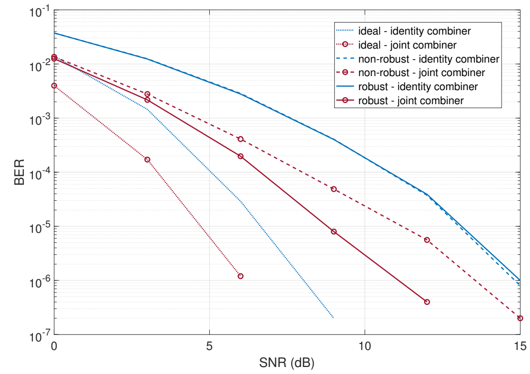

Fig. 1: BER versus SNR for the single-stream case ($r_k = 1$).

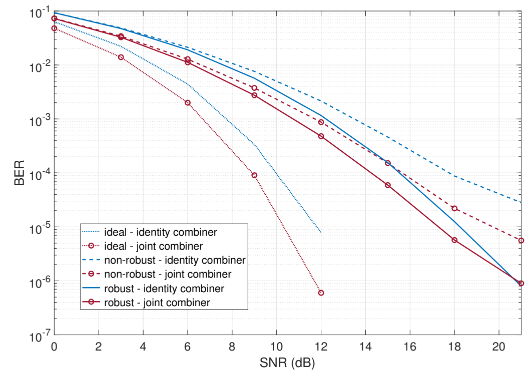

Fig. 2: BER versus SNR for the case of $r_k = 2$.

The results for the full-stream case (i.e., $r_k = 3$) are provided in Fig. 3. It can be seen that the performance of the non-robust solution is almost the same as that of T-MMSE (only supports the full-stream case) in [11]; both saturate at a BER of about 10^{-2} at an SNR of around 15 dB, while the performance of the proposed robust solution decreases with SNR. Also, note that the joint precoder/combiner solution has almost the same performance as the precoder plus the identity combiner solution, which eliminates the need for combiner feedback or extra computational complexity due to the combiner design.

A user can know its combining matrix in two ways: by calculating its own combining matrix via solving for the precoding and combining matrices (through the same procedure as the BS) using Algorithm 1, which increases the computational complexity at the user side; or through the combining matrices at the BS being fed back to the user, which leads to a communication overhead. Note that for the latter approach, a quantized version or its codebook index can be used. In Fig. 4, employing the second approach, the performance of the quantized combiner is analyzed, which shows that even with a simple QPSK quantizer, we can obtain an acceptable performance.

A. Computational Complexity Analysis

Based on (11) and (13), the computational complexity of calculating the precoding matrix is dominated by its corresponding matrix inversion operation; hence it is $O(N_t^3)$,

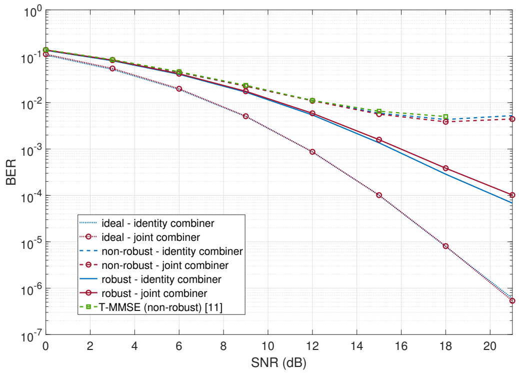

Fig. 3: BER versus SNR for the full-stream case ($r_k = 3$).

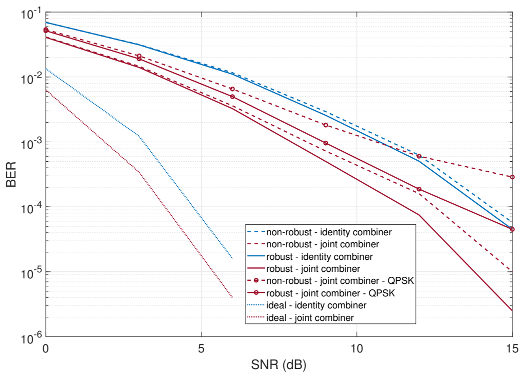

Fig. 4: BER for the single-stream with quantized combiner.

where $O(\cdot)$ is the standard big-O notation. The complexity of calculating the combining matrices, based on (9) and (16), is, however, dominated by the calculation of $\mathbf{D}_{c,k}$, which has an order of complexity of $O(KN_t^2N_r)$. Calculation of the scaling coefficient also involves similar matrix multiplications and has a complexity of $O(KN_t^2 r)$. Therefore, the overall complexity of the proposed robust algorithm becomes $O(KN_t^3)$ (also holds for its non-robust version), which is at the same order as the commonly used non-robust joint transceiver design algorithms such as T-MMSE [11] as well as BD and its regularized version (RBD) [17].

V. Conclusions

We have considered the downlink of a multiuser MIMO system with multiple-antenna users operating in the TDD mode. Assuming calibration errors and imperfect CSI, we proposed a robust algorithm to iteratively solve for the optimal precoding and combining matrices. Basically, the proposed precoder and combiners are extended versions of the conventional MMSE precoder and combiner, respectively, which take into account the effects of calibration errors and the imperfectness of CSI in addition to noise. Simulation results demonstrate that the proposed robust solution has a superior performance compared to the existing algorithms, especially in the full-stream case, while having the same order of computational complexity.

Appendix A: Proof of Proposition 1

By expanding the objective function of (7) using (2) and taking the expectation over the noise and the users' data, we have

$$\begin{aligned}&\mathrm{E}_{n_k,s_k}\left[\left\|\mathbf{y}'_k-\mathbf{s}_k\right\|^2\right]=\beta^{-2}\mathrm{tr}\left\{\mathbf{W}_k^H\mathbf{W}_k\mathbf{R}_{n,k}\right\}+\mathrm{tr}\left\{\mathbf{R}_{s,k}\right\}\\&-\mathrm{tr}\left\{\mathbf{R}_{s,k}\mathbf{A}_k^H\mathbf{C}_b^H\left(\hat{\mathbf{H}}^*_{UL,k}-\tilde{\mathbf{H}}^*_{UL,k}\right)\mathbf{C}_{u,k}^H\mathbf{W}_k^H\right\}\end{aligned}\tag{18}$$

$$\begin{aligned}&-\mathrm{tr}\left\{\mathbf{W}_k\mathbf{C}_{u,k}\left(\hat{\mathbf{H}}^T_{UL,k}-\tilde{\mathbf{H}}^T_{UL,k}\right)\mathbf{C}_b\mathbf{A}_k\mathbf{R}_{s,k}\right\}\\&+\mathrm{tr}\Bigg\{\mathbf{W}_k\mathbf{C}_{u,k}\left(\hat{\mathbf{H}}^T_{UL,k}-\tilde{\mathbf{H}}^T_{UL,k}\right)\mathbf{C}_b\left(\sum_{u=1}^{K}\mathbf{A}_u\mathbf{R}_{s,u}\mathbf{A}_u^H\right)\\&\qquad\mathbf{C}_b^H\left(\hat{\mathbf{H}}^*_{UL,k}-\tilde{\mathbf{H}}^*_{UL,k}\right)\mathbf{C}_{u,k}^H\mathbf{W}_k^H\Bigg\},\end{aligned}\tag{19}$$

knowing that $\mathrm{E}[\mathbf{n}_k]=\mathbf{0}_{N_{r,k}\times 1}$, $\mathrm{E}\left[\mathbf{n}_k\mathbf{n}_k^H\right]=\mathbf{R}_{n,k}$, $\mathrm{E}[\mathbf{s}_k]=\mathbf{0}_{r_k\times 1}$ and $\mathrm{E}\left[\mathbf{s}_k\mathbf{s}_k^H\right]=\mathbf{R}_{s,k}, \forall k\in\mathscr{K}$, and that $\mathrm{E}\left[\mathbf{s}_k\mathbf{s}_{k'}^H\right]=\mathbf{0}_{r_k\times r_{k'}}, \forall k\neq k'\in\mathscr{K}$.

Lemma 1. *Let* $\mathbf{Q}$ *and* $\mathbf{X}$ *be arbitrary* $n\times n$ *and random* $m\times n$ *matrices, respectively, with i.i.d. entries of zero mean and variance* σ^2. *Then, we have* $\mathrm{E}\left[\mathbf{X}\mathbf{Q}\mathbf{X}^H\right]=\sigma^2\mathrm{tr}\left\{\mathbf{Q}\right\}\mathbf{I}_m$.

Proof: Let $\mathbf{Z}:=\mathrm{E}\left[\mathbf{X}\mathbf{Q}\mathbf{X}^H\right]$, then, its (i,j)-th entry can be written as $z_{i,j}=\mathrm{E}\left[\mathbf{X}_i\mathbf{Q}\mathbf{X}_j^H\right]$, where $\mathbf{X}_i$ is the i-th row of $\mathbf{X}$. Since $\mathbf{X}$ has i.i.d. zero mean entries, for the off-diagonal entries ($i\neq j$) of $\mathbf{Z}$, we have $z_{i,j}=\mathrm{E}\left[\mathbf{X}_i\right]\mathbf{Q}\mathrm{E}\left[\mathbf{X}_j^H\right]=0$. For the diagonal entries, we have

$$\begin{aligned}z_{i,i}=&\mathrm{E}\left[\mathbf{X}_i\mathbf{Q}\mathbf{X}_i^H\right]\\=&\sum_{r=1}^{n}\sum_{\substack{s=1\\s\neq r}}^{n}\mathrm{E}\left[x_{i,r}\right]q_{r,s}\mathrm{E}\left[x_{i,s}^*\right]+\sum_{r=1}^{n}\mathrm{E}\left[\left|x_{i,r}\right|^2\right]q_{r,r}\\=&\sigma^2\mathrm{tr}\left\{\mathbf{Q}\right\},\end{aligned}\tag{20}$$

where $x_{i,j}$ and $q_{i,j}$ are the (i,j)-th entries of $\mathbf{X}$ and $\mathbf{Q}$, respectively. This concludes the proof. ■

Since $\tilde{\mathbf{H}}_{UL,k}$ has i.i.d. entries of zero mean and variance σ_h^2, by taking the expectation of (19) over the channel estimation error and using Lemma 1, we obtain

$$\begin{aligned}&\mathrm{E}_{n_k,s_k,\tilde{H}_{UL,k}}\left[\left\|\mathbf{y}'_k-\mathbf{s}_k\right\|^2\right]\\&=\mathrm{tr}\left\{\mathbf{R}_{s,k}\right\}-\mathrm{tr}\left\{\mathbf{R}_{s,k}\mathbf{A}_k^H\mathbf{C}_b^H\hat{\mathbf{H}}^*_{UL,k}\mathbf{C}_{u,k}^H\mathbf{W}_k^H\right\}\\&-\mathrm{tr}\left\{\mathbf{W}_k\mathbf{C}_{u,k}\hat{\mathbf{H}}^T_{UL,k}\mathbf{C}_b\mathbf{A}_k\mathbf{R}_{s,k}\right\}+\beta^{-2}\mathrm{tr}\left\{\mathbf{W}_k\mathbf{R}_{n,k}\mathbf{W}_k^H\right\}\\&+\mathrm{tr}\left\{\mathbf{W}_k\mathbf{C}_{u,k}\hat{\mathbf{H}}^T_{UL,k}\mathbf{C}_b\left(\sum_{u=1}^{K}\mathbf{A}_u\mathbf{R}_{s,u}\mathbf{A}_u^H\right)\mathbf{C}_b^H\hat{\mathbf{H}}^*_{UL,k}\mathbf{C}_{u,k}^H\mathbf{W}_k^H\right\}\\&+\sigma_h^2\mathrm{tr}\left\{\mathbf{C}_b\left(\sum_{u=1}^{K}\mathbf{A}_u\mathbf{R}_{s,u}\mathbf{A}_u^H\right)\mathbf{C}_b^H\right\}\mathrm{tr}\left\{\mathbf{W}_k\mathbf{C}_{u,k}\mathbf{C}_{u,k}^H\mathbf{W}_k^H\right\}.\end{aligned}\tag{21}$$

Lemma 2. *Let* $\mathbf{Q}$ *and* $\mathbf{X}$ *be arbitrary and random diagonal (with diagonal elements having covariance matrix* $\mathbf{R}_X$*) matrices, respectively. Then, we have* $\mathrm{E}\left[\mathbf{X}\mathbf{Q}\mathbf{X}^H\right]=\mathbf{R}_X\odot\mathbf{Q}$, *and* $\mathrm{E}\left[\mathbf{X}^H\mathbf{Q}\mathbf{X}\right]=\mathbf{R}_X^T\odot\mathbf{Q}$.

Proof: Let $\mathbf{Z}:=\mathrm{E}\left[\mathbf{X}\mathbf{Q}\mathbf{X}^H\right]$, then, since $\mathbf{X}$ is diagonal, the (i,j)-th entry of $\mathbf{Z}$, $z_{i,j}$, can be written as $z_{i,j}=q_{i,j}\mathrm{E}\left[x_{i,i}x_{j,j}^*\right]=q_{i,j}(\mathbf{R}_X)_{i,j}$, where $(\mathbf{R}_X)_{i,j}$ is the (i,j)-th entry of $\mathbf{R}_X$, which proves the first equation in Lemma 2. Next, let $\mathbf{Y} := \mathrm{E}\left[\mathbf{X}^H\mathbf{Q}\mathbf{X}\right]$. Since covariance matrix is Hermitian, for the (i,j)-th entry of $\mathbf{Y}$, $y_{i,j}$, we have $y_{i,j} = q_{i,j}\mathrm{E}\left[x_{i,i}^* x_{j,j}\right] = q_{i,j}(\mathbf{R}_X)_{j,i}$, which concludes the proof. ■

We know that $\mathbf{C}_b$ and $\mathbf{C}_{u,k}$ are diagonal matrices with diagonal entries with covariance matrices $\mathbf{R}_b$ and $\mathbf{R}_{u,k}$, respectively. Hence, by taking the expectation of (21) over the calibration errors and using Lemma 2, the closed-form expression for the MSE at the k-th user can be obtained as in (8).

Using the fact that trace is invariant under cyclic permutations, i.e., $\mathrm{tr}\{\mathbf{ABCD}\}=\mathrm{tr}\{\mathbf{DABC}\}$, (21) can be rewritten as

$$\begin{aligned}&\mathrm{E}_{n_k,s_k,\tilde{H}_{UL,k}}\left[\left\|\mathbf{y}_k' - \mathbf{s}_k\right\|^2\right] \quad (22)\\ &= \mathrm{tr}\left\{\mathbf{R}_{s,k}\right\} - \mathrm{tr}\left\{\mathbf{C}_b^H\hat{\mathbf{H}}_{UL,k}^*\mathbf{C}_{u,k}^H\mathbf{W}_k^H\mathbf{R}_{s,k}\mathbf{A}_k^H\right\}\\ &- \mathrm{tr}\left\{\mathbf{R}_{s,k}\mathbf{W}_k\mathbf{C}_{u,k}\hat{\mathbf{H}}_{UL,k}^T\mathbf{C}_b\mathbf{A}_k\right\} + \beta^{-2}\mathrm{tr}\left\{\mathbf{W}_k\mathbf{R}_{n,k}\mathbf{W}_k^H\right\}\\ &+\mathrm{tr}\left\{\mathbf{C}_b^H\hat{\mathbf{H}}_{UL,k}^*\mathbf{C}_{u,k}^H\mathbf{W}_k^H\mathbf{W}_k\mathbf{C}_{u,k}\hat{\mathbf{H}}_{UL,k}^T\mathbf{C}_b\left(\sum_{u=1}^K\mathbf{A}_u\mathbf{R}_{s,u}\mathbf{A}_u^H\right)\right\}\\ &+ \sigma_h^2\mathrm{tr}\left\{\mathbf{C}_b^H\mathbf{C}_b\left(\sum_{u=1}^K\mathbf{A}_u\mathbf{R}_{s,u}\mathbf{A}_u^H\right)\right\}\mathrm{tr}\left\{\mathbf{W}_k\mathbf{C}_{u,k}\mathbf{C}_{u,k}^H\mathbf{W}_k^H\right\}.\end{aligned}$$

Now, taking the expectation of (22) over the calibration errors and using Lemma 2 results in the equivalent expression (10).

Appendix B: Proof of Proposition 2

The Lagrangian of the optimization problem in (12) is

$$\mathrm{L}(\mathbf{A}_k,\beta) = \sum_{k=1}^K\varepsilon_k - \lambda\left(\beta^2\sum_{u=1}^K\mathrm{tr}\left\{\mathbf{A}_u\mathbf{R}_{s,u}\mathbf{A}_u^H\right\} - P_T\right). \quad (23)$$

Using (10), and taking the derivative of the Lagrangian with respect to the precoding matrices, we have

$$\frac{\partial}{\partial\mathbf{A}_j}\mathrm{L}(\mathbf{A}_k,\beta) = -\mathbf{M}_b\hat{\mathbf{H}}_{UL,j}\mathbf{M}_{u,j}\mathbf{W}_j^T\mathbf{R}_{s,j}^T + \left(\mathbf{D}_p^T - \lambda\beta^2\mathbf{I}_{N_t}\right)\mathbf{A}_j^*\mathbf{R}_{s,j}^T, \quad (24)$$

where $\mathbf{D}_p := \sum_{k=1}^K\mathbf{D}_{p,k}$ and we have used the fact that $\frac{\partial}{\partial\mathbf{A}}\mathrm{tr}\{\mathbf{XA}\} = \mathbf{X}^T$, $\frac{\partial}{\partial\mathbf{A}}\mathrm{tr}\{\mathbf{XA}^H\} = \mathbf{0}$ and $\frac{\partial}{\partial\mathbf{A}}\mathrm{tr}\left\{\mathbf{X}_1\mathbf{A}\mathbf{X}_2\mathbf{A}^H\right\} = \mathbf{X}_1^T\mathbf{A}^*\mathbf{X}_2^T$ [15].

Setting the derivative to zero; after some calculations, we obtain

$$\mathbf{A}_j = \left(\mathbf{D}_p - \lambda\beta^2\mathbf{I}_{N_t}\right)^{-1}\mathbf{M}_b^*\hat{\mathbf{H}}_{UL,j}^*\mathbf{M}_{u,j}^*\mathbf{W}_j^H, \quad (25)$$

using the fact that $\mathbf{D}_p$ is a Hermitian matrix. Also, Using the power constraint of (12), β is readily obtained as (14). Substituting (25) into the optimization problem in (12), we reach the following unconstrained problem.

$$\xi = \arg\min_{\xi}\sum_{k=1}^K\varepsilon_k, \quad (26)$$

where $\xi := -\lambda\beta^2$. Using (10) and (25), the derivative of the objective function in (26), after some calculations, becomes

$$\begin{aligned}&\frac{\partial}{\partial\xi}\sum_{k=1}^K\varepsilon_k = 2\left(\xi - \frac{1}{P_T}\sum_{k=1}^K\mathrm{tr}\left\{\mathbf{W}_k\mathbf{R}_{n,k}\mathbf{W}_k^H\right\}\right)\times \quad (27)\\ &\mathrm{tr}\left\{\mathbf{M}_b\left(\mathbf{D}_p + \xi\mathbf{I}_{N_t}\right)^{-3}\mathbf{M}_b^*\sum_{j=1}^K\hat{\mathbf{H}}_{UL,j}^*\mathbf{M}_{u,j}^*\mathbf{W}_j^H\mathbf{R}_{s,j}\mathbf{W}_j\mathbf{M}_{u,j}\hat{\mathbf{H}}_{UL,j}^T\right\}.\end{aligned}$$

Setting this expression to zero, we obtain $\xi_{opt} = \frac{1}{P_T}\sum_{k=1}^K\mathrm{tr}\left\{\mathbf{W}_k^H\mathbf{W}_k\mathbf{R}_{n,k}\right\}$, which concludes the proof.

References

[1] H. Wei, D. Wang, H. Zhu, J. Wang, S. Sun and X. You, "Mutual coupling calibration for multiuser massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 606-619, Jan. 2016.

[2] W. Ma, D. Liu, Y. Liu, W. Pan and H. Zhao, "On the capacity of ZF beamforming in massive MIMO systems with imperfect reciprocity calibration," in *P*roc. IEEE Global Commun. Conf. (GLOBECOM), Singapore, 2017, pp. 1-5.

[3] D. Mi, M. Dianati, L. Zhang, S. Muhaidat and R. Tafazolli, "Massive MIMO performance with imperfect channel reciprocity and channel estimation error," *IEEE Trans. Commun.*, vol. 65, no. 9, pp. 3734-3749, Sept. 2017.

[4] R. Chopra, C. R. Murthy, H. A. Suraweera and E. G. Larsson, "Blind channel estimation for downlink massive MIMO systems with imperfect channel reciprocity," *IEEE Trans. Signal Process.*, vol. 68, pp. 3132-3145, Apr. 2020.

[5] R. Habendorf and G. Fettweis, "Pre-equalization for TDD systems with imperfect transceiver calibration," in *Proc. IEEE Veh. Technol. Conf. VTC Spring*, Singapore, 2008, pp. 1369—1373.

[6] S. Bazzi and W. Xu, "Robust Bayesian precoding for mitigation of TDD hardware calibration errors," *IEEE Signal Process. Lett.*, vol. 23, no. 7, pp. 929-933, July 2016.

[7] A. Minasian, S. Shahbazpanahi and R. S. Adve, "Distributed massive MIMO systems with non-reciprocal channels: Impacts and robust beamforming," *IEEE Trans. Commun.*, vol. 66, no. 11, pp. 5261-5277, Nov. 2018.

[8] S. Han, C. Yang, G. Wang, D. Zhu and M. Lei, "Coordinated multipoint transmission strategies for TDD systems with non-ideal channel reciprocity," *IEEE Trans. Commun.*, vol. 61, no. 10, pp. 4256-4270, Oct. 2013.

[9] A. Coskun and C. Candan, "Transmit precoding for flat-fading MIMO multiuser systems with maximum ratio combining receivers," *IEEE Trans. Veh. Technol.*, vol. 60, no. 2, pp. 710-716, Feb. 2011.

[10] J. Park, B. Lee and B. Shim, "A MMSE vector precoding with block diagonalization for multiuser MIMO downlink," *IEEE Trans. Commun.*, vol. 60, no. 2, pp. 569-577, Feb. 2012.

[11] J. Zhang, Y. Wu, Sh. Zhou and J. Wang, "Joint linear transmitter and receiver design for the downlink of multiuser MIMO systems," *IEEE Commun. Lett.* , vol. 9, no. 11, pp. 991-993, Nov. 2005.

[12] Y. Jiang, J. Li and W. W. Hager, "Joint transceiver design for MIMO communications using geometric mean decomposition," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3791-3803, Oct. 2005.

[13] X. Zhai, Y. Cai, Q. Shi, M. Zhao, G. Y. Li and B. Champagne, "Joint transceiver design with antenna selection for large-scale MU-MIMO mmWave systems," *IEEE J. Sel. Areas Commun.* , vol. 35, no. 9, pp. 2085-2096, Sept. 2017.

[14] Q. Yue, J. Hu, K. Yang and C. Huang, "Transceiver design for simultaneous wireless information and power multicast in multi-User mmWave MIMO system," *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 11394-11407, Oct. 2020.

[15] A. Hjorungnes and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2740-2746, June 2007.

[16] InterDigital Communications, "R1-1609904: Impact of UE calibration in a reciprocity-based UL MIMO transmission," 3GPP TSG RAN WG1 Meeting 86, Oct. 2016.

[17] K. Zu, R. C. de Lamare and M. Haardt, "Generalized design of low-complexity block diagonalization type precoding algorithms for multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 10, pp. 4232-4242, Oct. 2013.