Coordinated Multi-Point Transmission Strategies for TDD Systems with Non-Ideal Channel Reciprocity

Shengqian Han, Chenyang Yang, Gang Wang, Dalin Zhu, and Ming Lei

***Abstract*—This paper studies transmission strategies for downlink time division duplex coordinated multi-point (CoMP) systems with non-ideal uplink-downlink channel reciprocity, where several multi-antenna base stations (BSs) jointly serve multiple single-antenna users. Due to the inherent antenna calibration errors among different BSs, the uplink-downlink channels are no longer reciprocal such that the channel information at the BSs is with multiplicative noises. To mitigate the performance degradation caused by the imperfect channels, we employ a weighted sum rate estimate as the objective function for robust precoder design. To show the impact of data sharing on the performance of CoMP under the non-ideal channel reciprocity, we provide a unified framework for designing precoder for CoMP systems with different amount of data sharing. The optimal parametric linear precoder structure that maximizes the weighted sum rate estimate under per-BS power constraints is characterized, based on which a closed-form robust signal-to-leakage-plus-noise ratio (SLNR) precoder is proposed to exploit the statistics of the calibration errors. Simulation results show that the proposed precoder in conjunction with user scheduling provides substantial performance gain over Non-CoMP transmission and the CoMP transmission with non-robust precoders.**

***Index Terms*—Coordinated multi-point (CoMP), uplink-downlink channel reciprocity, multi-cell multi-user precoder, robust optimization.**

I. Introduction

INTER-CELL interference (ICI) is a key bottleneck for achieving high spectral efficiency in universal frequency reuse cellular networks, especially when multi-input multi-output (MIMO) techniques are applied. Among various interference mitigation techniques, coordinated multi-point (CoMP) transmission is promising and has attracted much attention [1, 2]. Coherently cooperative transmission, also known as CoMP-JP (joint processing) in the context of 3GPP Long-Term Evolution-Advanced (LTE-A) or network MIMO in the literature [1], can fully exploit the potential of CoMP, where both data and channel state information (CSI) are shared among the base stations (BSs).

The performance gain of CoMP-JP comes at costs of high-capacity and low-latency backhaul links as well as increased signalling overhead. On one hand, in order to reduce the backhaul capacity requirements and signalling overhead, different limited cooperation strategies have been studied in the literature. Coordinated beamforming (CB) is a popular CoMP strategy, which requires no data sharing among the coordinated BSs. By exchanging the channels [3] or control information [4], CoMP-CB can eliminate the ICI and provide a substantial performance gain over Non-CoMP transmission. Partial CoMP transmission together with coordinated BS clustering is another approach for limited cooperation, which can achieve a balance between performance improvement and system overhead. In [5] a partial cooperative precoding scheme was proposed for static BS clustering. In [6] the dynamic clustering strategy with zero-forcing precoding was studied. In [7] dynamic clustering and partial cooperative precoding are jointly optimized. On the other hand, however, even the BSs are connected with high capacity backhaul links that are very expensive, CoMP-JP cannot achieve the expected high throughput when imperfect CSI is directly applied for downlink precoding [8].

With perfect backhaul, CoMP-JP mimics a single-cell multiuser MIMO system in concept, however there are subtle but important differences between the two system settings. Consequently, the well-explored transmission strategies cannot be extended to CoMP in a straightforward manner. First of all, per-BS power constraints (PBPC) should be considered instead of sum power constraints, which yield more complicated optimization [9–11]. Second, CoMP channel is a concatenation of multiple single-cell channels. As a result, the downlink CSI is corrupted by multiplicative noises, which are either caused by non-ideal uplink-downlink channel reciprocity in time division duplex (TDD) systems [12], or by structured codebook in frequency division duplex (FDD) systems [13]. Compared with additive noises such as channel estimation errors, these multiplicative noises are more detrimental because they hinder the co-phasing of coherent CoMP transmission.

The CSI at the BSs is required for all kinds of CoMP transmission to gain their benefits. It is widely recognized that TDD is more desirable for CoMP systems than FDD, because FDD needs large feedback overhead for providing CSI to the BSs [14]. In TDD systems, the downlink channel is usually obtained by the BSs via estimating the uplink channel by exploiting the channel reciprocity. However, this is invalid in practical systems due to the imperfect calibration for analog gains of radio frequency (RF) chains in transmit and receive antennas [15,16], which is especially true for CoMP as will be clear soon. Note that in CoMP systems the interference experienced at the BSs and at the users may also be not reciprocal.

Manuscript received September 5, 2012; revised March 17, 2013. The editor coordinating the review of this paper and approving it for publication was H. Yanikomeroglu.

Part of this work was presented at the IEEE WCNC'12. This work was supported in part by NEC Laboratories, China, by the national key project of next generation wideband wireless communication networks (2011ZX03003-001), and by the Fundamental Research Funds for the Central Universities.

S. Han and C. Yang are with the School of Electronics and Information Engineering, Beihang University, Beijing, China (e-mail: sqhan@ee.buaa.edu.cn; cyyang@buaa.edu.cn).

G. Wang, D. Zhu, and M. Lei are with NEC Laboratories, China (e-mail: {wang_gang, zhu_dalin, lei_ming}@nec.cn).

Digital Object Identifier 10.1109/TCOMM.2013.090313.120667

To solve this problem, the actual signal-to-interference-plus-noise ratio (SINR) to assist downlink modulation and coding selection (MCS) should be estimated at each user and then fed back to the BSs. This, however, is out of the scope of this paper, where we restrict ourselves to the non-ideal reciprocity between uplink and downlink channels.

To ensure the channel reciprocity, antenna calibration is often used to compensate the mismatch of RF chains. Self-calibration is a popular method used in single-cell systems [15], which adjusts all antennas to achieve the same RF analog gain as that of a reference antenna, and hence yields a constant scalar ambiguity between the uplink and downlink channels for all antennas. Such a scalar ambiguity does not affect the performance of single-cell single-user systems [12]. Yet as will be explained in Section III, self-calibration among BSs is hard to implement in CoMP systems. This leads to multiple ambiguity factors between the uplink and downlink channels at different coordinated BSs, unless precise reference antennas are used at different BSs to allow them to achieve identical RF analog gain, which is of high cost in practice. Over-the-air calibration is another method applied in single-cell systems [16], which can be extended to CoMP systems. The calibration accuracy highly depends on the quality of channel estimation. To improve the performance, the measurements of multiple uplink and downlink frames or of multiple users are required.

In this paper, we employ an alternative way to deal with the non-ideal uplink-downlink channel reciprocity in CoMP systems. We resort to robust design for multicell multiuser precoder against the calibration errors. Noting that with the imperfect CSI caused by the multiplicative noises, sharing all data among the coordinated BSs is not always beneficial but will impose heavy burden on existing capacity-limited backhaul links, it is necessary to analyze the impact of data sharing. To this end, we consider a strategy where the data of users are shared only within a subset of coordinated BSs, and provide a unified framework for optimization. In [5–7] cooperative precoding schemes under partial data sharing were studied under the assumption of perfect CSI, which cannot be applied to the scenario with imperfect CSI. In [17] a robust cooperative precoding was optimized against the bounded additive channel errors, while we consider the robust design against the unbounded multiplicative noises, which requires different optimization frameworks. The main contributions are summarized as follows:[1]

- Considering the imperfect downlink channels caused by the inter-BS antenna calibration errors, a weighted sum rate estimate is employed as the objective function for optimizing the robust linear CoMP precoder. A unified framework for designing multicell precoder for CoMP systems with different amount of data sharing is provided, which can bridge the gap between two extreme strategies of CoMP-JP with full data sharing and CoMP-CB with no data sharing. To solve the non-convex optimization problem, the structure of the optimal linear precoder that maximizes the weighted sum rate estimate under PBPC is found, which is governed by finite number of parameters, where spatial scheduling is implicitly included.
- A closed-form suboptimal precoder is proposed by properly selecting the involved parameters given the optimal linear precoder structure, which is decoupled from the scheduling. When the user fairness is not considered and when all the users experience identical interference, the precoder can be regarded as a robust version of the well-known single-to-leakage-plus-noise ratio (SLNR) precoder [19]. Simulation results show that full data sharing may become detrimental in CoMP systems, depending on the employed precoders and the accuracy of antenna calibration. The proposed precoder in conjunction with spatial scheduling provides remarkable performance gain over Non-CoMP transmission and the CoMP transmission with non-robust precoders in realistic system setting.

[1]This work was partly presented in a conference paper [18]. The material in this journal version is substantially extended: the system model and non-ideal channel reciprocity model are more clearly described, partial data sharing is incorporated, the propose precoder is improved, the theoretical results are more rigorous and the proofs are provided, the analysis is more thorough, and all simulation results are updated.

Notation: $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^H$, and $(\cdot)^\dagger$ denote the transpose, the complex conjugate, the conjugate transpose, and the Moore-Penrose inverse, respectively. $\mathrm{tr}(\cdot)$ denotes matrix trace, $\mathrm{diag}\{\cdot\}$ denotes a block diagonal matrix, $\Re\{\cdot\}$ denotes the real part of a complex number, and $\|\cdot\|$ denotes Euclidian norm. $\mathbf{I}_N$, $\mathbf{0}_N$ and $\bar{\mathbf{0}}_N$ denote $N \times N$ identity and zero matrices, and $N \times 1$ zero vector, respectively. For a set $\mathcal{S}$, its elements are $\mathcal{S}(1), \ldots, \mathcal{S}(|\mathcal{S}|)$ where $|\mathcal{S}|$ is the cardinality of $\mathcal{S}$.

II. System Model

Consider a downlink multicell cluster where N_c coordinated BSs each equipped with N_t antennas jointly serve N_u single-antenna users. For notational simplicity, we respectively denote the u-th user and the b-th BS as MS_u and BS_b, $u \in \mathcal{U} = \{1, \ldots, N_u\}$ and $b \in \mathcal{B} = \{1, \ldots, N_c\}$.

In TDD systems, channel reciprocity allows each BS to obtain the downlink channel to each user by estimating the uplink channel from the user. To highlight the multiplicative noises caused by the calibration errors, we assume that uplink channel estimation is perfect in the analysis and assume block flat fading channels, i.e., the channels remain constant at least during one period of uplink and downlink transmission. To emphasize the impact of imperfect channel reciprocity on the benefits of data sharing, we assume that the channels are fully shared among the BSs via noiseless and zero latency backhaul links, and the data intended to the users are partially shared. The sharing of data and channel information among the coordinated BSs needs high capacity backhaul links, which cannot be supported by existing cellular systems. Nonetheless, high capacity backhaul could be deployed in future systems if it is necessary to exploit the potential of CoMP systems, although is very expensive. For example, information sharing between the remote radio heads controlled by a centralized BS via point-to-point fiber is supported in LTE-A systems [20]. Let $\mathcal{D}_u \subseteq \mathcal{B}$ denote a subset of BSs who have the data of MS_u for $u \in \mathcal{U}$. An example of the considered scenario is shown in Fig. 1.

Denote the downlink channel comprising both the propagation channel and analog gains of RF chains from BS_b to

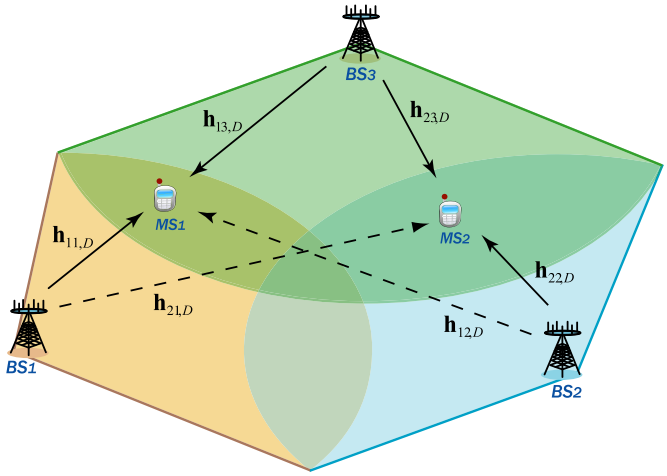


Fig. 1. A downlink multicell cooperative cluster with partial data sharing. In this example, the data of MS_1 are known by BS_1 and BS_3, and the data of MS_2 are known by BS_2 and BS_3. For each user, the solid and dash lines indicate signal and interference links, respectively.

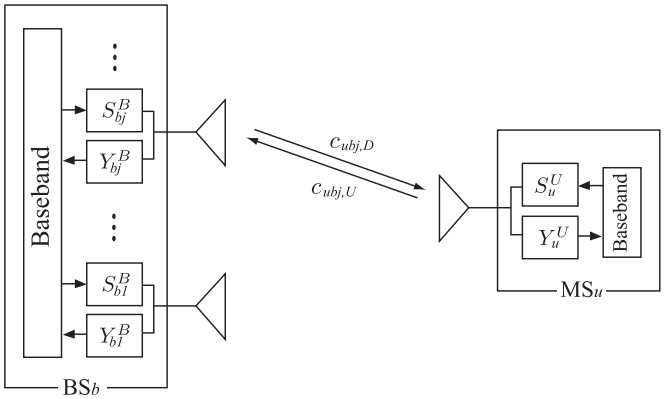


Fig. 2. Illustration of imperfect uplink-downlink channel reciprocity in TDD systems caused by antenna calibration errors. Each antenna is equipped with a high power amplifier (HPA) in the transmitter circuit and a low noise amplifier (LNA) in the receiver circuit. The uplink and downlink channels are $h_{ubj,U} = Y_{bj}^B c_{ubj,U} S_u^U$ and $h_{ubj,D} = Y_u^U c_{ubj,D} S_{bj}^B$, where $c_{ubj,U}$ and $c_{ubj,D}$ are the uplink and downlink propagation channels between the antenna pair, which are reciprocal according to electromagnetic theory, i.e., $c_{ubj,U} = c_{ubj,D}$.

MS_u as $\mathbf{h}_{ub,D} \in \mathbb{C}^{N_t \times 1}$, then the downlink CoMP channel from all BSs to MS_u is a concatenation of multiple single-cell channels, which is $\mathbf{h}_{u,D} = [\mathbf{h}_{u1,D}^T, \ldots, \mathbf{h}_{uN_c,D}^T]^T$. When linear precoding is used, the signal received at MS_u can be expressed as

$$y_u = \mathbf{h}_{u,D}^H \mathbf{w}_u x_u + \mathbf{h}_{u,D}^H \sum_{j \neq u} \mathbf{w}_j x_j + z_u, \tag{1}$$

where x_u is the data symbol for MS_u, the data symbols of all users are assumed as independent and identically distributed (i.i.d.) Gaussian random variables with zero mean and unit variance, z_u is the additive white Gaussian noise (AWGN) with zero mean and variance σ_u^2, $\mathbf{w}_u = [\mathbf{w}_{u1}^T, \ldots, \mathbf{w}_{uN_c}^T]^T \in \mathbb{C}^{N_c N_t \times 1}$ is the precoding vector for MS_u, and $\mathbf{w}_{ub} \in \mathbb{C}^{N_t \times 1}$ represents the precoding vector at BS_b for MS_u.

It is easy to see that the precoder for MS_u, $\mathbf{w}_{ub}$, will be zero if BS_b does not have the data of MS_u, i.e., $b \notin \mathcal{D}_u$. To provide a unified framework for designing multicell precoder for CoMP systems with different amount of data sharing, we express the partial data sharing among the BSs as the following constraint on the precoding vector for MS_u, $\tilde{\mathbf{D}}_u \mathbf{w}_u = \bar{\mathbf{0}}_{N_c N_t}$, where $\tilde{\mathbf{D}}_u \in \mathbb{C}^{N_c N_t \times N_c N_t}$ is block-diagonal with block size N_t, and the b-th diagonal block is $\mathbf{0}_{N_t}$ if $b \in \mathcal{D}_u$ and $\mathbf{I}_{N_t}$ if $b \notin \mathcal{D}_u$, i.e., $\tilde{\mathbf{D}}_u$ sorts out the BSs that do not transmit data x_u to MS_u.

Denote the collection of non-zero precoders of MS_u as $\bar{\mathbf{w}}_u = [\mathbf{w}_{u\mathcal{D}_u(1)}^T, \ldots, \mathbf{w}_{u\mathcal{D}_u(|\mathcal{D}_u|)}^T]^T$. Then by defining $\mathbf{D}_u = \mathbf{I}_{N_c N_t} - \tilde{\mathbf{D}}_u$ and defining $\bar{\mathbf{D}}_u \in \mathbb{C}^{N_c N_t \times |\mathcal{D}_u| N_t}$ as the submatrix of $\mathbf{D}_u$ consisting of all non-zero column vectors of $\mathbf{D}_u$, the relationship between $\bar{\mathbf{w}}_u$ and $\mathbf{w}_u$ is

$$\bar{\mathbf{w}}_u = \bar{\mathbf{D}}_u^H \mathbf{w}_u \quad \text{and} \quad \mathbf{w}_u = \bar{\mathbf{D}}_u \bar{\mathbf{w}}_u. \tag{2}$$

Example: *The notations regarding the precoders with partial data sharing are illustrated as follows. In the example shown in Fig. 1, we have* $\mathcal{D}_1 = \{1, 3\}$, $\mathbf{w}_1 = [\mathbf{w}_{11}^T, \bar{\mathbf{0}}_{N_t}^T, \mathbf{w}_{13}^T]^T$, $\bar{\mathbf{w}}_1 = [\mathbf{w}_{11}^T, \mathbf{w}_{13}^T]^T$, *and*

$$\tilde{\mathbf{D}}_1 = \begin{bmatrix} \mathbf{0}_{N_t} & \mathbf{0}_{N_t} & \mathbf{0}_{N_t} \\ \mathbf{0}_{N_t} & \mathbf{I}_{N_t} & \mathbf{0}_{N_t} \\ \mathbf{0}_{N_t} & \mathbf{0}_{N_t} & \mathbf{0}_{N_t} \end{bmatrix}, \mathbf{D}_1 = \begin{bmatrix} \mathbf{I}_{N_t} & \mathbf{0}_{N_t} & \mathbf{0}_{N_t} \\ \mathbf{0}_{N_t} & \mathbf{0}_{N_t} & \mathbf{0}_{N_t} \\ \mathbf{0}_{N_t} & \mathbf{0}_{N_t} & \mathbf{I}_{N_t} \end{bmatrix},$$

$$\bar{\mathbf{D}}_1 = \begin{bmatrix} \mathbf{I}_{N_t} & \mathbf{0}_{N_t} \\ \mathbf{0}_{N_t} & \mathbf{0}_{N_t} \\ \mathbf{0}_{N_t} & \mathbf{I}_{N_t} \end{bmatrix}.$$

We consider PBPC, since the transmit power cannot be shared among the BSs. Let $\mathbf{B}_b$ be a block-diagonal matrix, where its b-th block is $\mathbf{I}_{N_t}$ and all the other blocks are $\mathbf{0}_{N_t}$. Then the power constraint per BS is expressed as

$$\sum_{u=1}^{N_u} \mathbf{w}_u^H \mathbf{B}_b \mathbf{w}_u \leq P \quad \forall\, b, \tag{3}$$

where P is the maximal transmit power of each BS.

III. Non-perfect Channel Reciprocity

In this section, we model the channel uncertainty caused by imperfect channel reciprocity between uplink and downlink channels in TDD systems.

As shown in Fig. 2, let S_{bj}^B and Y_{bj}^B denote the transmit and receive analog gains of the j-th antenna at BS_b, and S_u^U and Y_u^U denote the transmit and receive analog gains of the antenna at MS_u. Then the uplink channel $h_{ubj,U}$ and downlink channel $h_{ubj,D}$ between the j-th antenna of BS_b and the antenna of MS_u satisfy

$$h_{ubj,D} = \frac{S_{bj}^B Y_u^U}{Y_{bj}^B S_u^U} h_{ubj,U} \triangleq g_{ubj} h_{ubj,U}. \tag{4}$$

We can see that there exists an ambiguity factor g_{ubj} between the uplink and downlink channels for each antenna pair. Since the antennas at both the BS and the user have different RF chains, the ambiguity factors of different pair of uplink and downlink channels differ, which destroys the reciprocity, i.e., $g_{ubi} \neq g_{ubj}$ if $i \neq j$.

Define $\mathbf{h}_{u,U} = [\mathbf{h}_{u1,U}^T, \ldots, \mathbf{h}_{uN_c,U}^T]^T$ as the uplink CoMP channel of MS_u, where $\mathbf{h}_{ub,U} = [h_{ub1,U}, \ldots, h_{ubN_t,U}]^T \in \mathbb{C}^{N_t \times 1}$ is the uplink channel vector from MS_u to BS_b. From (4) the relationship between uplink and downlink CoMP channels is expressed as

$$\mathbf{h}_{u,D} = \mathbf{G}_u \mathbf{h}_{u,U}, \tag{5}$$

where $\mathbf{G}_u = \text{diag}\{\mathbf{g}_u\}$, $\mathbf{g}_u = \left[\mathbf{g}_{u1}^T, \ldots, \mathbf{g}_{uN_c}^T\right]^T$, and $\mathbf{g}_{ub} = \left[g_{ub1}, \ldots, g_{ubN_t}\right]^T$ represents the ambiguity factors between the uplink and downlink channels for BS_b and MS_u.

A. Antenna Self-calibration

To recover the channel reciprocity in TDD systems, antenna calibration is necessary. Self-calibration is a popular antenna calibration method used in single-cell systems [15]. Perfect self-calibration is able to ensure the same ambiguity factor for all antennas at each BS as that of a reference antenna (say the first antenna). Therefore, after perfect self-calibration, the uplink and downlink channels satisfy

$$h_{ubj,D} = g_{ub1} h_{ubj,U} \ \forall \ j. \tag{6}$$

Comparing (6) with (4), it is shown that the self-calibration results in a constant scalar ambiguity factor g_{ub1} for all antennas at each BS, which does not affect the beamforming direction in single-cell single-user systems and hence has no impact on its performance [12].

The self-calibration is easy to implement when the antennas are co-located, since the test signals received by the antennas to be calibrated are with high signal strength. To calibrate the antennas located at different BSs in the same way, additional over-the-air communication protocol between the BSs needs to be designed, and time and frequency domain synchronization among the BSs is required. Even when these are provided, the received calibration signals are weak due to the propagation attenuation, which will induce low-accuracy self-calibration.

When the self-calibration is only performed at each individual BS, different ambiguity factors exist between multiple coordinated BSs. This leads to severe performance loss for CoMP systems, which will be shown by simulations later. Such a problem resembles the phase ambiguity problem in limited feedback CoMP systems with per-cell codebook-based channel quantization [13], where phase ambiguities exist between the independently quantized channel direction information from each coordinated BS to a user. In limited feedback CoMP systems, feeding back the phase ambiguities to the BSs is an immediate way to solve the problem, which however is not applicable to our problem here. This is because neither the BSs nor the users know the ambiguity factors after the self-calibration, which depend on the RF analog gains of both the BSs and the users, as shown in (4).

B. Calibration Error Model

As have been mentioned before, self-calibration within each BS leads to different ambiguity factors at multiple coordinated BSs. We refer to the BS-wise ambiguity as "port error" which is denoted as $g_{ub}^{(1)}$. Theoretically, the ambiguity factors of all antennas at the same BS should be identical after the self-calibration, which however are actually time-varying since the analog gains of RF chains vary with temperature, humidity, etc. We refer to the time-varying ambiguity as "residual error" and denote it as $g_{ubj}^{(2)}$. Then we can express the ambiguity factors for BS_b as

$$\mathbf{g}_{ub} = g_{ub}^{(1)} [g_{ub1}^{(2)}, \ldots, g_{ubN_t}^{(2)}]^T \triangleq g_{ub}^{(1)} \mathbf{g}_{ub}^{(2)}, \tag{7}$$

where $g_{ub}^{(1)}$ and $\mathbf{g}_{ub}^{(2)}$ are independent from each other, and both are usually modeled as random variables with log-uniformly distributed amplitudes and uniformly distributed phases [21].

Note that although we consider self-calibration, the model in (7) is also valid when other antenna calibration methods are applied. For example, with over-the-air calibration the multicell channel estimation errors will also lead to the port errors and the analog gain drift of RF chains will lead to the residual errors.

The resulting multiple ambiguity factors lead to imperfect downlink CoMP channel even if the uplink channel estimation is perfect, as shown in (5). When such a channel with unknown multiplicative factors is used for downlink precoding, the desired signals transmitted from multiple BSs may be added destructively, which turn into the interference.

IV. Optimal Robust Linear Precoder Structure

In order to alleviate the performance degradation caused by the non-ideal channel reciprocity, we resort to the design of robust linear precoder toward maximizing the weighted sum rate estimate of multiple users in the coordinated cells, where the user fairness is reflected in the weights.

The principle of robust design is to exploit *a priori* knowledge of statistics of the uncertainty to make a system less sensitive to the uncertainty. Robust precoders can be designed in various ways against imperfect CSI, depending on how to obtain the objective function that exploits the *a priori* information. The worst-case approach is often used in the scenarios with bounded channel uncertainty, which employs the worst-case estimation of the weighted sum rate as the objective function [17]. Another approach employs statistically modeled channel uncertainty. Bayesian approach is a widely used stochastic method, where the objective function is estimated as the expectation of the weighted sum rate.

In our problem, we assume that *a priori* knowledge of statistics of the calibration errors, $\mathbb{E}_g\{\mathbf{g}_u \mathbf{g}_u^H\}$, is known, where $\mathbb{E}_g\{\cdot\}$ denotes the expectation with respect to the calibration errors. Considering that the closed-form expression of the expectation of the weighted sum rate is very hard to derive, the Bayesian approach is not applicable. For mathematical tractability, an alternative estimate of the weighted sum rate is employed as the design goal. Specifically, we employ the minimum mean-square error (MMSE) estimates of the signal and interference powers to obtain a weighted sum rate estimate as the robust objective function, which gives rise to an explicit expression with respect to the linear precoder. We will compare the estimated data rate with the average data rate via simulations later, which implies that the proposed robust precoder will perform close to the Bayesian robust precoder.

A. Weighted Sum Rate Estimate Maximization Problem

The received power at MS_u of the signals from the BSs to MS_j can be obtained from (1) as

$$p_{uj} = |\mathbf{h}_{u,D}^H \mathbf{w}_j|^2. \tag{8}$$

When $u = j$, p_{uj} is the received power of the desired signal. Otherwise, it is the interference power.

Due to the imperfect channel reciprocity, the BSs do not know the true value of p_{uj}. With the uplink channel, the statistics of the calibration errors, and the model in (5), the BSs can estimate the value of p_{uj}. The MMSE estimate of p_{uj} can be obtained as

$$\hat{p}_{uj}^{\text{mmse}} = \arg\min_{\hat{p}_{uj}} \ \mathbb{E}_g\{|\hat{p}_{uj} - p_{uj}|^2\} = \mathbf{w}_j^H \mathbf{R}_u \mathbf{w}_j, \tag{9}$$

where $\mathbf{R}_u = \mathbb{E}_g\{\mathbf{G}_u \mathbf{h}_{u,U} \mathbf{h}_{u,U}^H \mathbf{G}_u^H\} = \mathbf{H}_{u,U} \mathbb{E}_g\{\mathbf{g}_u \mathbf{g}_u^H\} \mathbf{H}_{u,U}^H$ with $\mathbf{H}_{u,U} = \text{diag}\{\mathbf{h}_{u,U}\}$.

Based on (9), the instantaneous SINR and the data rate of MS_u can be estimated as

$$\widehat{\text{SINR}}_u = \frac{\mathbf{w}_u^H \mathbf{R}_u \mathbf{w}_u}{\sum_{j \neq u} \mathbf{w}_j^H \mathbf{R}_u \mathbf{w}_j + \sigma_u^2}, \ \hat{R}_u = \log(1 + \widehat{\text{SINR}}_u). \tag{10}$$

The robust linear precoder design problem is formulated as follows

$$\max_{\{\mathbf{w}_u\}} \ \sum_{u=1}^{N_u} \alpha_u \hat{R}_u \tag{11a}$$

$$s.t. \ \tilde{\mathbf{D}}_u \mathbf{w}_u = \bar{\mathbf{0}}_{N_c N_t} \ \forall \ u \tag{11b}$$

$$\sum_{u=1}^{N_u} \mathbf{w}_u^H \mathbf{B}_b \mathbf{w}_u \leq P \ \forall \ b, \tag{11c}$$

where the weights α_u, $u = 1, \cdots, N_u$, reflect the priorities of different users, and the constraints (11b) indicate that the precoder for MS_u at BS_b is zero if BS_b does not have the data of MS_u.

Problem (11) maximizes the weighted sum of the data rate estimate subject to PBPC and partial data sharing. The problem is non-convex and NP-hard [22], whose globally optimal solution is very hard to find. In the following subsection, we strive to characterize the structure of the optimal linear precoder, which will be used to develop an efficient suboptimal solution later.

The stochastic approaches including Bayesian approach and the proposed method may lead to an optimistic performance estimation for the users. Therefore, an outage may occur when transmitting over a particular channel realization in deep fading. One usual way to solve this problem is to introduce a margin to the estimated performance. For slow fading channels, which are typical for CoMP, the problem can also be solved with the following *three-phase transmission strategy*.

1) The BSs compute the robust precoder $\{\mathbf{w}_u\}$ based on the data rate estimate, then transmit the precoded pilots with $\{\mathbf{w}_u\}$ in a downlink dedicated training phase, e.g., using the downlink reference signals to assist data demodulation in LTE systems.
2) Based on the received dedicated pilots, each user can estimate the equivalent channels, e.g., MS_u can estimate $\mathbf{h}_u \mathbf{w}_j^H$ for $j = 1, \ldots, N_u$. Thus each user can obtain the accurate SINR, and then report it to the BSs through an uplink feedback channel.
3) Based on the reported SINR, the BSs are able to transmit the data to each user with a proper modulation and coding scheme in the subsequent downlink phase.

Such a three-phase transmission strategy, in fact, is also necessary when considering the uplink-downlink interference non-reciprocity in TDD CoMP systems.

B. Optimal Robust Linear Precoder Structure

In order to find the solution of problem (11), we resort to the relationship between the data rate and mean-square error (MSE). With the assumption of perfect CSI, the equivalence between the weighted sum rate maximization problem and the weighted sum MSE minimization problem was established in [23]. When considering imperfect CSI, the equivalence between the two problems is nontrivial and has not been built so far, where the technique in [23] cannot be extended straightforwardly because we need to get rid of the uncertain calibration errors for both the data rate and MSE. In Section IV-A we have removed the calibration errors in data rate by using the data rate estimate. However, it is still unknown how to remove the calibration errors in MSE such that the resulting weighted sum MSE minimization problem is equivalent to the weighted sum rate estimate maximization problem (11).

Denoting the receive filter at MS_u as v_u, the MSE of its data symbol estimation can be obtained from (1) as

$$\begin{aligned} \epsilon_u \ &\triangleq \ \mathbb{E}_{x,z}\{|v_u^* y_u - x_u|^2\} = 1 - 2\Re\{v_u^* \mathbf{h}_{u,D}^H \mathbf{w}_u\} \\ &\quad + \Big(\sum_{j=1}^{N_u} |\mathbf{h}_{u,D}^H \mathbf{w}_j|^2 + \sigma_u^2\Big)|v_u|^2, \end{aligned} \tag{12}$$

where $\mathbb{E}_{x,z}\{\cdot\}$ denotes the expectation with respect to the data and noise.

Average MSE is a commonly used metric for robust design with imperfect CSI, e.g. in [24]. However, by taking the expectation over the calibration errors involved in $\mathbf{h}_{u,D}^H$, we can see that the average MSE depends on both the first-order and second-order moments of the calibration errors, while the data rate estimate only depends on the second-order moment of the calibration errors as shown in (10). This suggests that problem (11) is not equivalent to a weighed sum average MSE minimization problem. To establish the equivalence, we propose a new metric, named *estimated MSE*, in the following.

With the uplink channel and the statistics of the calibration errors, the MMSE estimate of $|\mathbf{h}_{u,D}^H \mathbf{w}_j|^2$ is $\mathbf{w}_j^H \mathbf{R}_u \mathbf{w}_j$, which is given in (9). Therefore, we can obtain an estimate of $\mathbf{h}_{u,D}^H \mathbf{w}_u$ as $\sqrt{\mathbf{w}_u^H \mathbf{R}_u \mathbf{w}_u} e^{j\theta_u}$ with an arbitrary phase θ_u. Then, the MSE of MS_u's data symbol estimation can be estimated as

$$\begin{aligned} \hat{\epsilon}_u = &1 - 2\Re\{v_u^* \sqrt{\mathbf{w}_u^H \mathbf{R}_u \mathbf{w}_u} e^{j\theta_u}\} + \\ &\Big(\sum_{j=1}^{N_u} \mathbf{w}_j^H \mathbf{R}_u \mathbf{w}_j + \sigma_u^2\Big)|v_u|^2. \end{aligned} \tag{13}$$

Lemma 1. Let $t_u \geq 0$ be a scalar weight for the MSE of MS_u's data symbol. The following weighted estimated sum

MSE minimization problem

$$\begin{aligned} \min_{\mathbf{w},v,t} \quad & \textstyle\sum_{u=1}^{N_u} \alpha_u (t_u \hat{\epsilon}_u - \log t_u) \\ s.t. \quad & \tilde{\mathbf{D}}_u \mathbf{w}_u = \bar{\mathbf{0}}_{N_c N_t} \ \forall\ u \\ & \textstyle\sum_{u=1}^{N_u} \mathbf{w}_u^H \mathbf{B}_b \mathbf{w}_u \le P \ \forall\ b \end{aligned} \tag{14}$$

is equivalent to the weighted sum rate estimate maximization problem (11), in a sense that the two problems have identical globally optimal precoders $\mathbf{w}^{\text{opt}}$.[2]

Proof: See Appendix A. ∎

The estimated MSE in (13), $\hat{\epsilon}_u$, is non-convex with respect to $\mathbf{w}$, while the MSE under perfect CSI in (12), ϵ_u, is convex for $\mathbf{w}$. This indicates that the algorithms proposed in [23] and [7] cannot be used for the considered scenario, which exploit the convexity of ϵ_u with respect to $\mathbf{w}$ for given t and v. To solve the non-convex problem (14), we find the structure of the optimal robust linear precoder from the equivalent relationship shown in Lemma 1. In the following we present and prove the main results of this section.

Theorem 1. *The optimal robust linear precoder that maximizes the weighted sum rate estimate subject to partial data sharing and PBPC has the following structure*

$$\mathbf{w}_u^{\text{opt}} = \sqrt{q_u^{\text{opt}}} \bar{\mathbf{D}}_u \bar{\mathbf{f}}_u^{\text{opt}} \quad \forall u, \tag{15}$$

where $\bar{\mathbf{f}}_u^{\text{opt}}$ *is an eigenvector of the following matrix*

$$\alpha_u \kappa_u \left(\bar{\mathbf{D}}_u^H \left(\sum_{j \ne u} \alpha_j \kappa_j \mathbf{R}_j + \sum_{b=1}^{N_c} \nu_b \mathbf{B}_b \right) \bar{\mathbf{D}}_u \right)^{\dagger} \bar{\mathbf{D}}_u^H \mathbf{R}_u \bar{\mathbf{D}}_u \ \forall u \tag{16}$$

with parameters $\kappa_u, \nu_b \in [0,1]$ *for* $u = 1, \ldots, N_u$ *and* $b = 1, \ldots, N_c$,

$$[q_1^{\text{opt}}, \ldots, q_{N_u}^{\text{opt}}]^T = \mathbf{\Sigma}^{-1} [\sigma_1^2 d_1, \ldots, \sigma_{N_u}^2 d_{N_u}]^T, \tag{17}$$

d_u *denotes the eigenvalue of the matrix (16) corresponding to the eigenvector* $\bar{\mathbf{f}}_u^{opt}$, and $\mathbf{\Sigma}$ is defined as

$$[\mathbf{\Sigma}]_{uj} = \begin{cases} \bar{\mathbf{f}}_u^{\text{opt},H} \bar{\mathbf{D}}_u^H \mathbf{R}_u \bar{\mathbf{D}}_u \bar{\mathbf{f}}_u^{\text{opt}}, & j = u, \\ -d_u \bar{\mathbf{f}}_j^{\text{opt},H} \bar{\mathbf{D}}_j^H \mathbf{R}_u \bar{\mathbf{D}}_j \bar{\mathbf{f}}_j^{\text{opt}}, & j \ne u. \end{cases} \tag{18}$$

Here, $[\mathbf{\Sigma}]_{uj}$ *denotes the element of* $\mathbf{\Sigma}$ *at the* u*-th row and* j*-th column for* $u, j \in \mathcal{U}$.

Proof: See Appendix B. ∎

From the procedure of the proof, we can obtain the following properties with respect to the parameters κ_u, ν_b and d_u, which will be used to develop a closed-form suboptimal precoder in next section.

Property 1. Denote t^{opt}, v^{opt} and λ^{opt} as the optimal solutions and the optimal Lagrange multipliers corresponding to PBPC for the weighted estimated sum MSE minimization problem (14). Then the optimal values of κ_u and ν_b can be expressed as

$$\kappa_u^{\text{opt}} = t_u^{\text{opt}} |v_u^{\text{opt}}|^2 / c \ \text{ and } \ \nu_b^{\text{opt}} = \lambda_b^{\text{opt}} / c, \tag{19}$$

where t^{opt} and v^{opt} are given in (40) in Appendix B, and $c = \max(\max_u t_u^{\text{opt}} |v_u^{\text{opt}}|^2, \max_b \lambda_b^{\text{opt}})$ is only for ensuring the values of $\kappa_u^{\text{opt}}, \nu_b^{\text{opt}} \in [0,1]$, which does not change the matrix in (16).

Property 2. The parameter κ_u^{opt} is an indicator of spatial scheduling for MS_u. If $\kappa_u^{\text{opt}} = 0$, we have $v_u^{\text{opt}} = 0$ from (19) and then $\mathbf{w}_u^{\text{opt}} = \mathbf{0}_{N_c N_t}$ according to (40) in Appendix B, meaning that MS_u is not scheduled. If $\kappa_u^{\text{opt}} > 0$, MS_u is scheduled, and $d_u > 0$ according to (44) in Appendix B, which implies that the optimal beamforming vector of a scheduled user only corresponds to *positive* eigenvalues of the matrix in (16).

Property 3. The optimal Lagrange multipliers λ^{opt} satisfy

$$\sum_{b=1}^{N_c} \lambda_b^{\text{opt}} \le \frac{\sum_{u=1}^{N_u} \alpha_u}{P}, \tag{20}$$

which is proved in Appendix C.

The proposed optimal precoder structure is governed by $N_c + 2N_u$ parameters, i.e., $\{\nu_b\}$, $\{\kappa_u\}$, and $\{d_u\}$. It means that given the optimal $\{\nu_b^{\text{opt}}\}$, $\{\kappa_u^{\text{opt}}\}$ and $\{d_u^{\text{opt}}\}$, we can form the matrix in (16) and find the eigenvector corresponding to the eigenvalue d_u^{opt} as the optimal beamforming vector $\bar{\mathbf{f}}_u^{\text{opt}}$. Then based on $\{\bar{\mathbf{f}}_u^{\text{opt}}\}$ and $\{d_u^{\text{opt}}\}$, the optimal power allocation can be obtained.

In the precoder structure, $\{\nu_b\}$ and $\{\kappa_u\}$ are independent real-valued numbers between 0 and 1, while d_u is a positive eigenvalue of the matrix in (16), having only finite possible choices for any given $\{\nu_b\}$ and $\{\kappa_u\}$. In particular, for the special case of perfect CSI, $\mathbf{R}_u$ is of rank one and hence the matrix in (16) has only one positive eigenvalue, which means that the parameters $\{d_u\}$ are fixed and the precoder structure is governed by only $N_c + N_u$ parameters. This is consistent with the existing results in [8]. Compared to the work in [8] that assumes perfect CSI and exploits a framework of uplink-downlink duality, our work studies a general case with imperfect CSI and exploits the equivalence between the weighted sum rate estimate maximization problem and the weighted estimated sum MSE minimization problem.

C. Impact of Partial Data Sharing

When the non-robust precoder is applied without taking the imperfect CSI into account, full data sharing among the coordinated BSs may lead to severe performance degradation, as shown in [8]. By contrast, when the proposed robust precoder is applied, sharing all the data to support CoMP-JP for all users can always provide the maximal weighted sum rate estimate. This is because with full data sharing we have $\tilde{\mathbf{D}}_u = \mathbf{0}_{N_c N_t}$ according to the definition of $\tilde{\mathbf{D}}_u$ in Section II. Then, the constraints (11b) become inactive, resulting in a relaxed version of problem (11), which yields a better solution than that of the original problem with partial data sharing.

On the other hand, it should be pointed out that full data sharing may not be necessary for maximizing the weighted sum rate estimate. This can be observed from the following example.

Example: *Assume the phase of port error* $g_{ub}^{(1)}$ *in antenna self-calibration as i.i.d. and uniformly distributed between* $[-180°, 180°]$. *In fact, this is the worst case where the antennas among different BS are not calibrated at all.*

[2]In (14), the notation $\mathbf{w}$ is short for $\{\mathbf{w}_u\}$. The notations $v = \{v_u\}$, $t = \{t_u\}$, $\mathbf{w}^{\text{opt}} = \{\mathbf{w}_u^{\text{opt}}\}$, $v^{\text{opt}} = \{v_u^{\text{opt}}\}$, $t^{\text{opt}} = \{t_u^{\text{opt}}\}$, and $\lambda^{\text{opt}} = \{\lambda_b^{\text{opt}}\}$ in the following are defined similarly.

According to the calibration error model in (7), we have

$$\mathbb{E}_g\{\mathbf{g}_{ub}\mathbf{g}_{ub}^H\} = \mathbb{E}_g\{|g_{ub}^{(1)}|^2\}\mathbb{E}_g\{\mathbf{g}_{ub}^{(2)}\mathbf{g}_{ub}^{(2)H}\} \triangleq \mathbf{\Omega}_{ub}\ \forall b, \quad (21)$$

$$\mathbb{E}_g\{\mathbf{g}_{ub}\mathbf{g}_{uj}^H\} = \mathbf{0}_{N_t}\ \forall b \neq j. \quad (22)$$

Considering $\mathbf{g}_u = \left[\mathbf{g}_{u1}^T, \ldots, \mathbf{g}_{uN_c}^T\right]^T$, *we have* $\mathbb{E}_g\{\mathbf{g}_u\mathbf{g}_u^H\} = \mathrm{diag}\{\mathbf{\Omega}_{u1}, \ldots, \mathbf{\Omega}_{uN_c}\}$, *which is a block diagonal matrix. Then from the definitions below (9), we can obtain that*

$$\begin{aligned}\mathbf{R}_u &= \mathrm{diag}\{\mathrm{diag}\{\mathbf{h}_{u1,U}\}\mathbf{\Omega}_{u1}\mathrm{diag}\{\mathbf{h}_{u1,U}^H\}, \ldots, \\ &\qquad \mathrm{diag}\{\mathbf{h}_{uN_c,U}\}\mathbf{\Omega}_{uN_c}\mathrm{diag}\{\mathbf{h}_{uN_c,U}^H\}\} \\ &\triangleq \mathrm{diag}\{\mathbf{\Lambda}_{u1}, \ldots, \mathbf{\Lambda}_{uN_c}\}. \end{aligned} \quad (23)$$

To see the necessary of data sharing, we assume full data sharing in this case, then $\bar{\mathbf{D}}_u = \mathbf{I}_{N_cN_t}$. *Substituting* $\mathbf{R}_u$ *and* $\bar{\mathbf{D}}_u$ *into (16) and considering the definition of* $\mathbf{B}_b$ *in Section II, the matrix in (16) can be expressed as*

$$\begin{aligned}&\alpha_u\kappa_u\mathrm{diag}\Big\{\big(\textstyle\sum_{j\neq u}\alpha_j\kappa_j\mathbf{\Lambda}_{j1} + \nu_1\mathbf{I}_{N_t}\big)^\dagger\mathbf{\Lambda}_{u1}, \ldots, \\ &\qquad \big(\textstyle\sum_{j\neq u}\alpha_j\kappa_j\mathbf{\Lambda}_{jN_c} + \nu_{N_c}\mathbf{I}_{N_t}\big)^\dagger\mathbf{\Lambda}_{uN_c}\Big\}, \end{aligned} \quad (24)$$

which is a block diagonal matrix with N_c *blocks each having the size of* $N_t \times N_t$.

We can obtain the following observations from (24).

1) *According to Theorem 1, the optimal beamforming vector* $\bar{\mathbf{f}}_u^{opt}$ *of MS*$_u$ *is an eigenvector of the matrix in (24). Since the eigen-matrix of a block diagonal matrix is block diagonal,* $\bar{\mathbf{f}}_u^{opt}$ *includes only one nonzero sub-vector, which is an eigenvector of one diagonal block of the matrix in (24). This implies that each user only receives data from one BS, and therefore there is no need to share its data among the BSs. The resulting CoMP transmission is actually CoMP-CB. This is consistent with the intuition: the performance of CoMP-CB will not be destroyed by the phase ambiguity among the BSs, because CoMP-CB can eliminate interference without the need of co-phasing.*
2) *Channel sharing among the BSs is necessary even for this worst case. On one hand, sharing channel will be useful to determine the optimal user access mechanism from a cooperative cluster viewpoint. On the other hand, by jointly designing the parameters* $\{\kappa_u\}$ *in (24) for all users, each BS can judiciously balance the interference suppression to all unserved users and the signal enhancement of its served users.*

In general cases where the dynamic range of phase port errors are less than $[-180°, 180°]$, the required amount of data sharing has an entangled relationship with the performance of the antenna calibration when different precoders are employed. We will evaluate the impact of data sharing on the system performance via simulations in Section VI.

V. Closed-form Suboptimal Multicell Precoder

With the optimal structure we found for the precoder, in this section we will propose a closed-form suboptimal precoder with properly selected parameters according to the properties developed in Section IV-B.

A. Suboptimal Parameter Selection

Based on Property 2, the parameter κ_u reflects the scheduling status of MS$_u$. In reality we can predict that there are always some users who will not be served by using the optimal precoder, i.e., the corresponding κ_u^{opt} should be zeros. This is because in general serving all candidate users in the same time-frequency resources will not achieve the maximal weighted sum rate. Therefore, in the following we decouple the optimal precoder into a spatial scheduling and a suboptimal precoder, and study the selection of the values of κ_u^{opt} and ν_b^{opt} only for the scheduled users.

According to Property 1, the scalar normalization factor c does not affect the optimal precoder. Thus, we can obtain the parameters κ_u^{opt} and ν_b^{opt} by equivalently finding $t_u^{\text{opt}}|v_u^{\text{opt}}|^2$ and λ_b^{opt} as shown in (19).

Based on (40) in Appendix B, $t_u^{\text{opt}}|v_u^{\text{opt}}|^2$ can be expressed as

$$t_u^{\text{opt}}|v_u^{\text{opt}}|^2 = \frac{1}{\underbrace{\sum_{j\neq u}\mathbf{w}_j^{\text{opt},H}\mathbf{R}_u\mathbf{w}_j^{\text{opt}}}_{\text{Interference power}} + \sigma_u^2} \cdot \frac{\mathbf{w}_u^{\text{opt},H}\mathbf{R}_u\mathbf{w}_u^{\text{opt}}}{\underbrace{\mathbf{w}_u^{\text{opt},H}\mathbf{R}_u\mathbf{w}_u^{\text{opt}}}_{\text{Signal power}} + \sum_{j\neq u}\mathbf{w}_j^{\text{opt},H}\mathbf{R}_u\mathbf{w}_j^{\text{opt}} + \sigma_u^2}}. \quad (25)$$

As discussed in Section IV-C, the optimal CoMP transmission strategy under the imperfect CSI depends on the performance of antenna calibration. The optimal precoder can provide high signal power and low interference power, because it can enhance the signal strength by cooperative transmission and at the same time eliminate the inter-cell and inter-user interference. This indicated that we can approximate the first item in the right-hand side of (25) as $\frac{1}{\sigma_u^2}$ and approximate the second item as one. Therefore, it is reasonable to consider the following approximation,

$$t_u^{\text{opt}}|v_u^{\text{opt}}|^2 \approx \frac{1}{\sigma_u^2}\ \forall\ u. \quad (26)$$

We proceed to consider the selection of λ_b^{opt}. Property 3 gives an upper bound of $\sum_{b=1}^{N_c}\lambda_b^{\text{opt}}$ that is $\sum_{j\in\mathcal{S}}\alpha_j/P$ with $\mathcal{S} \subseteq \mathcal{U}$ denoting the set of scheduled users[3], based on which we make a simple suboptimal selection of λ_b^{opt} as

$$\lambda_b^{\text{sub}} = \frac{\sum_{j\in\mathcal{S}}\alpha_j}{N_cP}\ \forall\ b. \quad (27)$$

To evaluate the impact of the approximations for $t_u^{\text{opt}}|v_u^{\text{opt}}|^2$ and λ_b^{opt}, we will compare the performance of the following proposed suboptimal robust precoder with a gradient-based solution to the original problem (11) via simulations later, from which we can see that the approximations lead to minor performance loss.

[3]Herein, since only the users in $\mathcal{S}$ will be served, the term $\sum_{u=1}^{N_u}\alpha_u$ in Property 3 is accordingly replaced by $\sum_{j\in\mathcal{S}}\alpha_j$.

B. Closed-form Suboptimal Linear Precoder

With the selected parameters $t_u^{\text{opt}}|v_u^{\text{opt}}|^2$ and λ_b^{sub} for the scheduled users, the matrix in (16) can be rewritten as

$$\alpha_u\Big(\sum_{j\neq u}\alpha_j\frac{\sigma_u^2}{\sigma_j^2}\bar{\mathbf{D}}_u^H\mathbf{R}_j\bar{\mathbf{D}}_u+\frac{\sigma_u^2\sum_{j\in\mathcal{S}}\alpha_j}{N_cP}\mathbf{I}_{|\mathcal{D}_u|N_t}\Big)^{\dagger}\bar{\mathbf{D}}_u^H\mathbf{R}_u\bar{\mathbf{D}}_u \quad \forall\ u\in\mathcal{S}, \tag{28}$$

where $\bar{\mathbf{D}}_u^H\sum_{b=1}^{N_c}\mathbf{B}_b\bar{\mathbf{D}}_u=\mathbf{I}_{|\mathcal{D}_u|N_t}$ is considered, which is obtained according to the definitions of $\mathbf{B}_b$ and $\bar{\mathbf{D}}_u$.

According to Theorem 1 and Property 2, the beamforming vector of MS_u for $u\in\mathcal{S}$ is an eigenvector corresponding to a positive eigenvalue of the matrix in (28). Let M_u denote the number of positive eigenvalues of the matrix in (28) for $u\in\mathcal{S}$, where $1\leq M_u\leq N_cN_t$. Then there will be $\prod_{u\in\mathcal{S}}M_u$ possible choices for the beamforming vectors of all $|\mathcal{S}|$ scheduled users. The complexity to find the optimal beamforming vectors will be high especially for large N_c, N_t and $|\mathcal{S}|$.

In order to reduce the complexity, we propose to select the beamforming vector for each scheduled user only from the eigenvectors corresponding to the L largest eigenvalues. In this manner, the number of possible choices for the beamforming vectors of all the $|\mathcal{S}|$ scheduled users reduces to $\prod_{u\in\mathcal{S}}\min\{L,M_u\}$.

Let d_{u,l_u} and $\bar{\mathbf{f}}_{u,l_u}$ denote the l_u-th largest eigenvalue and the corresponding eigenvector of the matrix in (28) for MS_u, where $l_u=1,\ldots,\min\{L,M_u\}$ and $u=\mathcal{S}(1),\ldots,\mathcal{S}(|\mathcal{S}|)$, i.e., $u\in\mathcal{S}$. Then we can use $\{l_{\mathcal{S}(1)},\ldots,l_{\mathcal{S}(|\mathcal{S}|)}\}$ to denote one of the $\prod_{u\in\mathcal{S}}\min\{L,M_u\}$ possible choices for all the scheduled users, based on which the proposed closed-form suboptimal robust precoder can be summarized as follows.

1) For every possible choice $\{l_{\mathcal{S}(1)},\ldots,l_{\mathcal{S}(|\mathcal{S}|)}\}$, perform the following procedures:
   - Based on Theorem 1, replace $\bar{\mathbf{f}}_u^{\text{opt}}$ and d_u in $\mathbf{\Sigma}$ defined as (18) with $\bar{\mathbf{f}}_{u,l_u}$ and d_{u,l_u} for all $u\in\mathcal{S}$, and compute the power allocation $\{q_{u,l_u}\}$ according to (17) as
   $$[q_{\mathcal{S}(1),l_{\mathcal{S}(1)}},\ldots,q_{\mathcal{S}(|\mathcal{S}|),l_{\mathcal{S}(|\mathcal{S}|)}}]^T=\mathbf{\Sigma}^{-1}[\sigma_{\mathcal{S}(1)}^2d_{\mathcal{S}(1),l_{\mathcal{S}(1)}},\ldots,\sigma_{\mathcal{S}(|\mathcal{S}|)}^2d_{\mathcal{S}(|\mathcal{S}|),l_{\mathcal{S}(|\mathcal{S}|)}}]^T. \tag{29}$$
   - Construct the initial precoders $\{\mathbf{w}'_{u,l_u}\}$ according to (15) as
   $$\mathbf{w}'_{u,l_u}=\sqrt{q_{u,l_u}}\bar{\mathbf{D}}_u\bar{\mathbf{f}}_{u,l_u}\ \forall\ u\in\mathcal{S}. \tag{30}$$
   - Scale the initial precoders $\{\mathbf{w}'_{u,l_u}\}$ obtained from the suboptimal parameters to ensure PBPC:
   $$\mathbf{w}_{u,l_u}=\sqrt{\rho}\mathbf{w}'_{u,l_u}\ \forall\ u\in\mathcal{S} \tag{31}$$
   with $\rho=P/\max_b\sum_{u\in\mathcal{S}}\mathbf{w}'^H_{u,l_u}\mathbf{B}_b\mathbf{w}'_{u,l_u}$.
   - Compute the weighted sum rate estimate with the precoders $\{\mathbf{w}_{u,l_u}\}$, denoted as $\hat{R}_{sum}(\{l_{\mathcal{S}(1)},\ldots,l_{\mathcal{S}(|\mathcal{S}|)}\})$.
2) Determine the final closed-form robust precoder $\mathbf{w}_u^{\text{RSLNR}}$ as
   $$\mathbf{w}_u^{\text{RSLNR}}=\mathbf{w}_{u,l_u^*}\ \forall\ u\in\mathcal{S}, \tag{32}$$
   where l_u^* is selected as $\{l^*_{\mathcal{S}(1)},\ldots,l^*_{\mathcal{S}(|\mathcal{S}|)}\}=\arg\max\hat{R}_{sum}(\{l_{\mathcal{S}(1)},\ldots,l_{\mathcal{S}(|\mathcal{S}|)}\})$. □

In practice, we can apply any spatial scheduler to select $|\mathcal{S}|$ users from the candidate user set $\mathcal{U}$, and then apply the robust precoder to serve the users in the same time-frequency resource. The proposed robust precoder has a closed-form expression, but requires repeated computations for $\prod_{u\in\mathcal{S}}\min\{L,M_u\}$ possible choices. In Section VI, we will evaluate the impact of L on the performance of the proposed precoder. The results show that when spatial schedulers such as the greedy user scheduler (GUS) [25] are applied together with the proposed precoder, setting $L=1$ can provide close to the maximum performance. This implies that in practice we can simply select the beamforming vector as the eigenvector corresponding to the maximal eigenvalue of the matrix in (28), which leads to very low computational complexity.

C. Closed-form Precoder vs. SLNR Precoder

When all the users have the same weights of priority and noise variances, i.e., $\alpha_{\mathcal{S}(1)}=\cdots=\alpha_{\mathcal{S}(|\mathcal{S}|)}$ and $\sigma^2_{\mathcal{S}(1)}=\cdots=\sigma^2_{\mathcal{S}(|\mathcal{S}|)}$[4], the matrix in (28) can be simplified as

$$\Big(\sum_{j\neq u}\bar{\mathbf{D}}_u^H\mathbf{R}_j\bar{\mathbf{D}}_u+\frac{\sigma_u^2|\mathcal{S}|}{N_cP}\mathbf{I}_{|\mathcal{D}_u|N_t}\Big)^{\dagger}\bar{\mathbf{D}}_u^H\mathbf{R}_u\bar{\mathbf{D}}_u\ \forall\ u\in\mathcal{S}. \tag{33}$$

Furthermore, consider the special case of $L=1$, then to find the beamforming vector $\bar{\mathbf{f}}_u$ of the proposed closed-form precoder (i.e., the eigenvector corresponding to the largest eigenvalue of the matrix in (33)), is equivalent to solve the following Rayleigh quotient maximization problem

$$\max_{\bar{\mathbf{f}}_u}\ \frac{\bar{\mathbf{f}}_u^H\bar{\mathbf{D}}_u^H\mathbf{R}_u\bar{\mathbf{D}}_u\bar{\mathbf{f}}_u}{\bar{\mathbf{f}}_u^H\bar{\mathbf{D}}_u^H\sum_{j\neq u}\mathbf{R}_j\bar{\mathbf{D}}_u\bar{\mathbf{f}}_u+\frac{\sigma_u^2|\mathcal{S}|}{N_cP}}, \tag{34}$$

where the objective function can be regarded as the average signal-to-average leakage-plus-noise ratio of MS_u.

When the channel reciprocity is perfect (i.e., $\mathbf{R}_u=\mathbb{E}_g\{\mathbf{G}_u\mathbf{h}_{u,U}\mathbf{h}_{u,U}^H\mathbf{G}_u^H\}=\mathbf{h}_{u,U}\mathbf{h}_{u,U}^H$ is of rank one), the data are fully shared (i.e., $\bar{\mathbf{D}}_u=\mathbf{I}_{N_cN_t}$), and the weights and the noise power are equal for all users, one can find that the proposed robust precoder will reduce to a conventional SLNR precoder. Therefore, we call the proposed precoder a robust SLNR (RSLNR) precoder.

VI. Simulation Results

In this section, we evaluate the accuracy of the introduced approximations, analyze the impact of calibration errors on the performance of CoMP systems, and evaluate the performance achieved by the proposed RSLNR precoder.

The calibration errors consist of port errors and residual errors. In simulations, we put emphasis on analyzing the impact of phase port errors on the performance of CoMP systems. This is because the residual errors are generally much smaller than the port errors, and the phase port errors will degrade the performance of CoMP systems severely. To this

[4]In practice, there will exist interference from non-cooperative clusters. The inter-cluster interference can be modeled as white noise, which is the worst-case assumption. Therefore, this condition implies that each user experiences the same interference.

end, we introduce a parameter θ and model the phases of the port errors as uniformly distributed random variables within $[-\theta, \theta]$. The phases of the residual errors are modeled as uniformly distributed random variables within $[-10^\circ, 10^\circ]$ [21], and the amplitudes of the port errors and the residual errors are modeled as log-uniformly distributed random variables within [-3, 3] dB and [-1, 1] dB [21], respectively. In addition, it is assumed that the residual errors are independent among all antennas in all coordinated BSs and the port errors are independent among different BSs.

The considered cooperative cluster layout is shown in Fig. 1, where $N_c = 3$ BSs, each equipped with $N_t = 2$ antennas, jointly serve $N_u = 30$ users, i.e., there are 10 users randomly placed in each cell. The cell radius r is set to 250 m, and the interference from non-cooperative clusters is modeled as white noise. Denoting the average receive signal-to-noise ratio (SNR) of users located at the cell edge as SNR_{edge}, then the average receive SNR of a user from a BS with distance d is computed as $\text{SNR}_{\text{edge}} + 37.6\log_{10}(\frac{r}{d}) + \chi$, where χ represents log-normal shadowing with standard deviation of 8 dB, and $d > 50$ m. The i.i.d. Rayleigh flat small-scale fading channels are considered. All the results are averaged over 1000 snapshots, where in each snapshot multiple users are randomly placed in the cluster for a realization of large-scale and small-scale channels.

To capture the user fairness, as in [8], the weights in (11a) are set to $\alpha_u = \eta/\bar{R}_u$ with

$$\bar{R}_u = \mathbb{E}_h\left\{\log\left(1 + \frac{N_c P}{N_u \sigma_u^2}\max_b \|\mathbf{h}_{ub}\|^2\right)\right\},$$

where $\mathbb{E}_h\{\cdot\}$ denotes the expectation with respect to small-scale channels, and η is a scaling factor ensuring $\sum_{u=1}^{N_u}\alpha_u = N_u$. $\bar{R}_u$ can be regarded as the average data rate of MS_u when it is served only by one BS with equal power allocation. Therefore, the weights reflect the proportional fairness among the users.

Unless otherwise specified, the data rates in the following are computed from Shannon capacity formula based on the true value of the SINR of each user when the robust precoder is used, rather than the estimated SINR for deriving the precoder shown in (10). In practice, the true SINR can be obtained via the three-phase transmission strategy suggested in Section IV-A.

A. Performance Evaluation of the Proposed Robust Precoder

First, we evaluate the impact of the parameter L on the performance of the proposed robust precoder. Two associated spatial schedulers are considered. One is the random user scheduler (RUS) which randomly selects one user from each cell, i.e., schedules three users in total in the simulations. The other is the GUS, which achieves near optimal performance in single-cell multiuser MIMO systems [25]. The GUS was originally designed for zero-forcing precoders in [25], which can be applied with the proposed precoder straightforwardly.

In Fig. 3, the weighted sum rates of the RSLNR precoder with RUS and GUS versus the phase port errors θ for $\text{SNR}_{\text{edge}} = 0, 10$ dB are provided. As shown in Fig. 3(a) and Fig. 3(b), when RUS is applied, increasing L can improve the performance for large θ. This can be understood intuitively by considering the extreme case of perfect antenna calibration, where the matrix in (28) will has only one positive eigenvalue so that setting $L = 1$ will be the best choice. The results imply that when the users are randomly scheduled, we should not simply choose the beamforming vector as the eigenvector corresponding to the maximal eigenvalue of the matrix in (28) in the scenarios with large θ. Nevertheless, noting that the performance improvement is negligible when $L > 2$, therefore we can set $L = 2$ to achieve a low-complexity robust precoder when applied together with RUS. By contrast, when GUS is applied, as shown in Fig. 3(c) and Fig. 3(d), setting $L = 1$ is enough to achieve close to the maximum sum rate. Considering that the robust precoder can achieve much better performance when applied with GUS than RUS, in the following simulations in this section, we will employ GUS and choose the beamforming vector as the eigenvector corresponding the maximal eigenvalue of the matrix in (28) by setting $L = 1$.

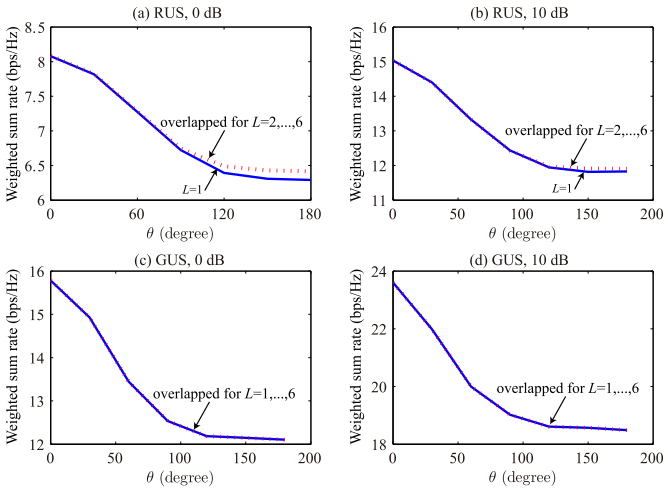


Fig. 3. The impact of L on the performance of the proposed robust precoder together with RUS and GUS. Full data sharing among the BSs is considered, $\theta = 120^\circ$, $\text{SNR}_{\text{edge}} = 0, 10$ dB, $N_c = 3$, and $N_t = 2$. According to the definitions in Section V-B, the matrix in (28) has the rank of $N_c N_t = 6$, $M_u = 6$, and $1 \le L \le 6$.

Then, to understand the impact of using the estimated data rate as the robust objective function, we compare the estimated data rate with the average data rate and the actual data rate in Fig. 4, where $\theta = 120^\circ$ and $\text{SNR}_{\text{edge}} = 10$ dB. We simulate 1000 channel realizations, for each of which we calculate the actual data rate based on the robust precoders and the true downlink channels, calculate the estimated data rate based on (10), and obtain the average data rate by taking the average over 100 calibration errors. We sort the 1000 data rate estimates in an ascending order, then plot it together with the corresponding actual data rate and average data rate in Fig. 4. We can see that the estimated data rate is close to the average data rate, which implies that the proposed robust precoders can perform close to the precoder obtained from the Bayesian robust design. In addition, it can be observed that compared with the actual data rate, both the estimated data rate and the average data rate will lead to the optimistic performance estimate for the users. As mentioned before, this problem can

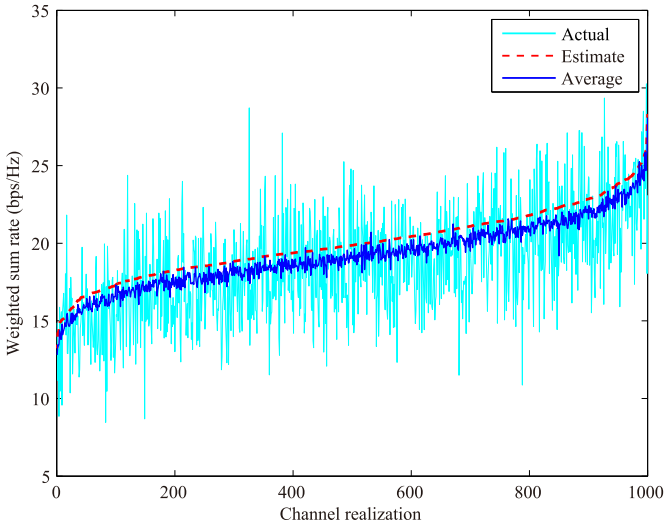


Fig. 4. Comparison of the data rate estimate, the average data rate, and the actual data rate, which are denoted by "Estimate", "Average" and "Actual" in the legend, respectively. Full data sharing among the BSs is considered, $\theta = 120°$, and $\text{SNR}_{\text{edge}} = 10$ dB.

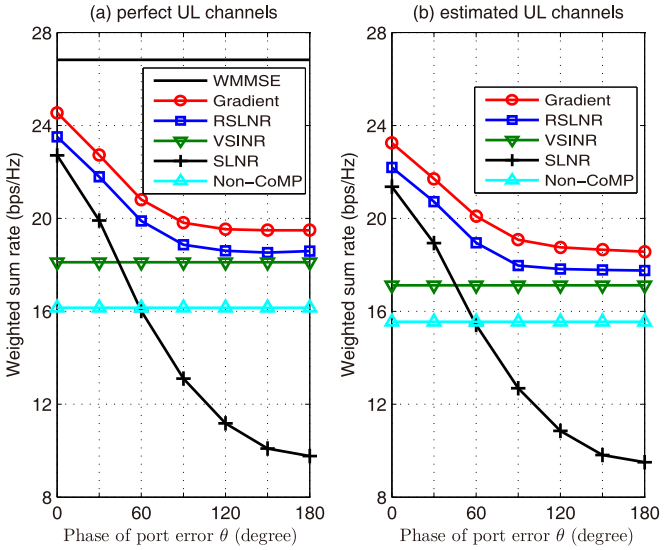


Fig. 5. Weighted sum rates with the six considered strategies versus θ. Full data sharing among the BSs together with both perfect and imperfect uplink channel estimation is considered, and $\text{SNR}_{\text{edge}} = 10$ dB. The WMMSE precoder assumes perfect CSI and hence is not plotted when considering imperfect uplink channel estimation.

be solved with a three-phase transmission strategy.

Next, to show the performance gain from the robust optimization, we compare the proposed RSLNR precoder with the following relevant strategies.

- The non-robust counterpart of the proposed precoder, *the SLNR precoder* [19], who simply assumes perfect channel reciprocity.
- A CoMP-CB scheme, *the virtual SINR (VSINR) precoder* [26], whose performance is not destroyed by the phase ambiguity among the BSs but is limited by no data sharing among the BSs.
- An efficient partial CoMP precoding scheme, *the WMMSE precoder* [7, 23], who is implemented under perfect CSI as a benchmark.
- *A gradient-based solution* (specifically using the function $fmincon$ of the optimization toolbox of MATLAB) to the original problem (11), who exhibits high performance with different initializations but is computationally costly.
- *A Non-CoMP SLNR precoder*, with which each user is served only by its master BS (the BS that provides the maximum receive power) and suffers from the inter-cell interference.

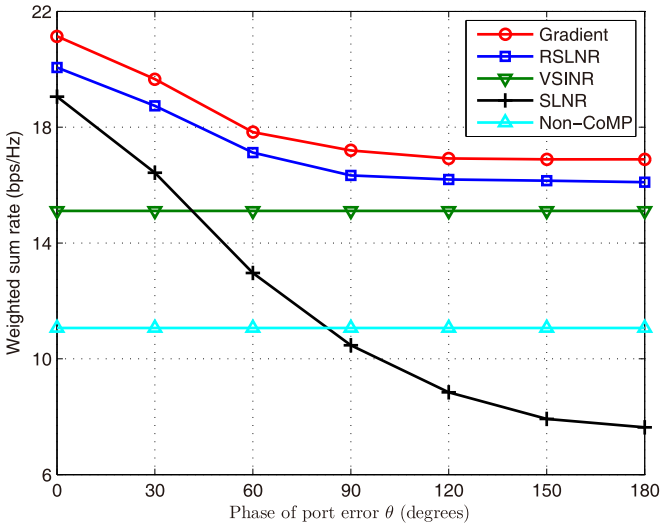
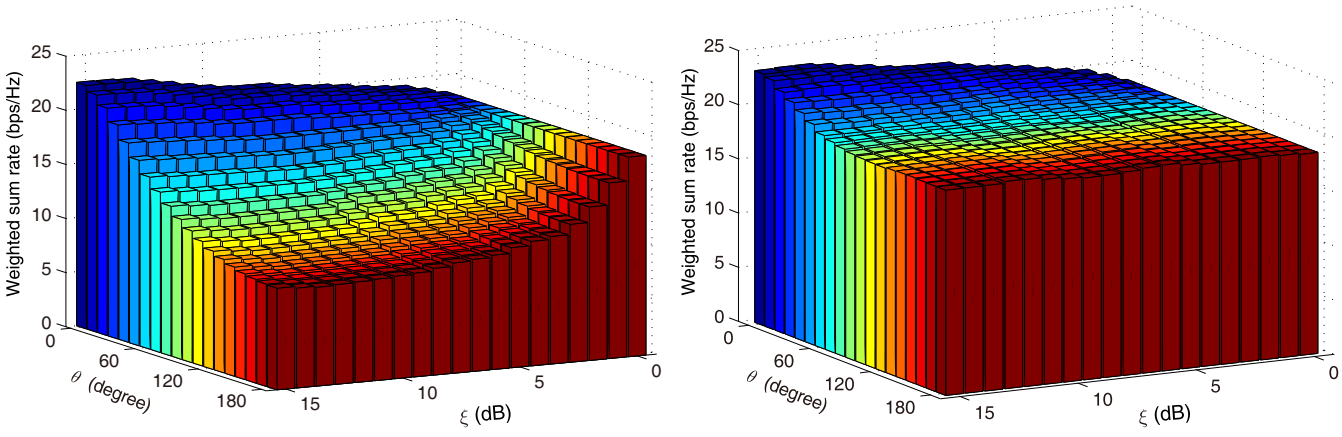


Fig. 6. Weighted sum rates versus θ with feedback latency. Full data sharing and perfect uplink channel estimation are considered, and $\text{SNR}_{\text{edge}} = 10$ dB. The WMMSE precoder assumes perfect CSI and hence is not plotted here.

In Fig. 5, the weighted sum rates with the six considered strategies are shown versus the phase port errors θ, where full data sharing among the BSs is considered. To show that the multiplicative noises are more detrimental than the additive noises in CoMP systems, we consider both perfect and imperfect uplink channel estimations in the simulation. Since the BSs' transmit power is usually much higher than the users', we set the uplink SNR as 5 dB lower than the downlink SNR for each pair of BS and user. The estimation of the uplink channel between MS_u and BS_b is modeled as $\hat{\mathbf{h}}_{ub,U} = \mathbf{h}_{ub,U} + \mathbf{e}_{ub}$, where $\mathbf{e}_{ub}$ consists of i.i.d. zero-mean complex Gaussian variables whose variances are determined by the uplink SNR.

As shown in Fig. 5(a), under the assumption of perfect CSI, the WMMSE precoder with 10 initializations achieves the highest weighted sum rate, which serves as a benchmark for the other strategies considering imperfect CSI. For the gradient-based method, we use 10 initializations to improve its performance. Compared with the gradient-based method, the proposed RSLNR precoder needs much low complexity at the expense of small weighted sum rate loss. Note that when we derive the RSLNR precoder, we have used several approximations in Section V-A. These results imply that the approximations lead to minor performance loss. When considering the non-robust CoMP SLNR precoder with full data sharing, we find that it may be inferior to the VSINR precoder that has no data sharing. This confirms that data sharing is not always beneficial when the channels are imperfect. Moreover, it can be observed that CoMP systems do not always perform better than Non-CoMP systems. With accurate antenna cali-



(a) SLNR precoder (b) RSLNR precoder

Fig. 7. Weighted sum rates achieved by the SLNR precoder and the RSLNR precoder as a function of data sharing threshold ξ and phase port errors θ.

bration (i.e., for small θ), the cooperation of multiple BSs can significantly improve the system performance. Yet with large antenna calibration errors and the SLNR precoder, the CoMP systems are even inferior to the Non-CoMP systems. By using the proposed RSLNR precoder, an evident performance gain can be observed over the SLNR precoder for both CoMP and Non-CoMP systems. By comparing Fig. 5(a) and Fig. 5(b), it can be observed that imperfect uplink channel estimation decreases the performance of all the precoders, but the performance loss is minor for CoMP-JP systems compared with that caused by the multiplicative noises. Note that even when $\theta = 0°$, the RSLNR precoder still outperforms the SLNR precoder and the WMMSE precoder outperforms the gradient-based method although both of them can achieve the local optimum, which is because the amplitude port errors and residual errors still exist in this case.

In addition, it is shown from Fig. 5(a) that to achieve the same weighted sum rate in CoMP systems, say 20 bps/Hz, the SLNR precoder requires θ to be less than $30°$, whereas the proposed RSLNR precoder allows θ to be less than $60°$. This means that a much lower accuracy is required for the reference antenna in self-calibration when using the proposed precoders, which is very attractive since the manufacture of high-accuracy reference antennas is hard and of high cost.

Finally, we show that the employed objective function for the robust optimization, the weighted sum rate estimate, is applicable for practical systems from a more realistic simulation. One issue related to the three-phase transmission strategy mentioned in Section IV-A is the SINR feedback latency in time-varying channels. When the SINR received at the BSs differs from the actual SINR at the users, the outage may occur in the subsequent downlink transmission. In order to ensure an acceptable outage probability, a back off margin was added to the SINR used for MCS selection in [27], which is in fact an often-used strategy in existing systems.

In Fig. 6, the results with perfect uplink channel estimation in Fig. 5(a) are reproduced with 5 ms feedback latency, 3 km/h mobility speed and 2 GHz carrier frequency. To ensure the outage probability to be less than 5%, the back off margin is set to 2 dB and 3 dB for the considered CoMP precoders and the Non-CoMP precoder, respectively.[5] The time-varying channels are based on Jakes' model with the temporal correlation function $J_0(2\pi f_d \tau)$, where $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind, f_d is the Doppler frequency and τ is the feedback latency. Similar to [27], the results do not consider a package sent in outage as a lost, since the received package with error can still be exploited using, for example, a type-II hybrid automatic-repeat request (HARQ). Compared with the ideal scenario where the users perfectly predict future channels and feed back the perfect SINR as shown in Fig. 5(a), the employed back off margin scheme leads to performance degradation for all the precoders in order to keep an acceptable outage probability. However, the relationship between the precoders remains the same.

B. Impact of the Partial Data Sharing

Partial data sharing motivated by capacity-constrained backhaul have been studied in the literature, e.g., [5–7]. Herein, the motivation of analyzing partial data sharing is to show that when considering non-ideal channel reciprocity, even if the backhaul capacity is unlimited, sharing more data among the BSs may be unnecessary or even detrimental for CoMP systems. The results are shown in Fig. 7, where perfect uplink channel estimation is assumed. In the simulations, we consider a simple data sharing strategy, with which the data of a user are shared among the BSs who have large average channel gains to the user. Specifically, for MS_u we denote the average channel gain from its master BS as β_{u0} in dB; then BS_b will be shared with the data of MS_u if the average channel gain β_{ub} from BS_b satisfies $\beta_{u0} - \beta_{ub} \leq \xi$, where ξ is a pre-determined threshold. It is easy to see that $\xi = 0$ means no data sharing (i.e., MS_u is served by all the BSs with CoMP-CB), while $\xi = +\infty$ means full data sharing (i.e., MS_u is served by all the BSs with CoMP-JP).

Figure 7(a) plots the weighted sum rate achieved by the SLNR precoder as a function of data sharing threshold ξ and phase port errors θ, where $\text{SNR}_{\text{edge}} = 10$ dB. It is shown that the performance does not necessarily improve when more

[5]The back off margin is chosen with the method presented in [28]. Specifically, we measure the performance across a range of possible margins through simulations and choose a desired value leading to the outage probability of 5%.

data are shared (i.e., when ξ grows from 0 dB to 15 dB). For large θ, the performance degrades with the growth of ξ. This is because the imperfect CSI turns the desired signals from the coordinated BSs to each user into the inter-cell inference. When the RSLNR precoder is applied, as shown in Fig. 7(b), more data sharing is always beneficial but the performance gain decreases when the calibration errors increase. This agrees with the previous analysis that full data sharing can provide the best performance but may not be necessary for the performance maximization.

C. Robustness to the Statistical Knowledge

In Fig. 8, we investigate the robustness of the proposed RSLNR precoder to the accuracy of the assumed statistics of the antenna calibration errors. We use relative errors to reflect the quality of the statistics, including both residual errors and port errors. Taking θ as an example, we assume $\hat{\theta} = \theta + (\delta\theta)\phi$, where $\delta \in [0, 100\%]$ is the relative error and ϕ is a uniformly distributed variable within $[-1, 1]$ that models the uncertainty. Perfect uplink channel estimation is assumed.

Figure 8 shows the weighted sum rate achieved by the RSLNR precoder versus the relative error of the statistics, where both full data sharing and no data sharing are considered. It is shown that the impact of the mismatch of the statistics is smaller for the case without data sharing (i.e. with smaller slope). This is because in this case the impact of phase port error is minor since the users receive data from only one BS where the amplitude error and the small residual phase error only lead to slight performance loss. For full data sharing, the mismatch of the statistics has a larger impact as expected, yet the performance loss is less than 10% when the relative errors are within 50% for $\text{SNR}_{\text{edge}} = 10$ dB. When $\text{SNR}_{\text{edge}} = 0$ dB, the mismatch of the statistics has a smaller impact because the system becomes noise-limited so that BS cooperation has less performance gain even with perfect statistical knowledge. In addition, it can be observed that with imperfect statistical knowledge, sharing full data among BSs might no longer achieve the maximal weighted sum rate for higher cell-edge SNR. Sharing full data is beneficial when the cell-edge SNR is low, because in this case the "array gain" provided from a "super BS with more antennas" can improve the receive SNR.

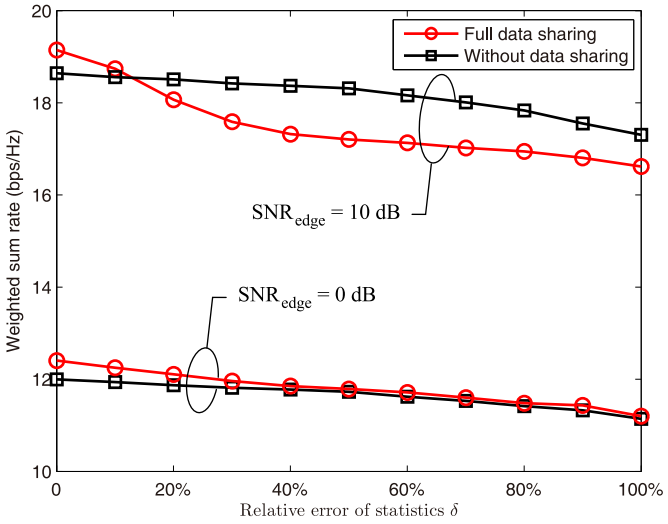


Fig. 8. Weighted sum rate achieved by the RSLNR precoder versus relative errors of the statistics of calibration errors. The true value of θ is 90°, and $\text{SNR}_{\text{edge}} = 0, 10$ dB.

VII. Conclusions

In this paper, we developed robust linear precoders to alleviate the performance degradation caused by the non-ideal uplink-downlink channel reciprocity in TDD CoMP systems. With the knowledge of the statistics of antenna calibration errors, we derived the optimal robust linear precoder structure that maximizes a weighted sum rate estimate. We showed that the optimal robust precoder is determined by finite real-valued parameters. By analyzing the optimal structure and by properly selecting these parameters, we proposed a closed-form suboptimal robust linear precoder. Simulation results showed that the proposed precoder exhibits evident performance gain over the non-robust precoders. Under the inherent non-ideal channel reciprocity, CoMP does not always outperform Non-CoMP and full data sharing among the coordinated BSs is not always beneficial when all the channels are shared among the BSs. Nonetheless, when the proposed robust transmission strategy is applied, CoMP will always outperform Non-CoMP and sharing more data will provide higher data rate.

Appendix A
Proof of Lemma 1

Proof: The proof builds upon the line of the work in [23].

With the expression of $\hat{\epsilon}_u$ in (13), we can rewrite problem (14) as

$$\min_{\mathbf{w}} \min_{t} \min_{v} \sum_{u=1}^{N_u} \alpha_u \left(t_u \left(1 - 2\Re\{v_u^* \sqrt{\mathbf{w}_u^H \mathbf{R}_u \mathbf{w}_u} e^{j\theta_u}\} + \left(\sum_{j=1}^{N_u} \mathbf{w}_j^H \mathbf{R}_u \mathbf{w}_j + \sigma_u^2 \right) |v_u|^2 \right) - \log t_u \right) \tag{35}$$

$$s.t. \quad \tilde{\mathbf{D}}_u \mathbf{w}_u = \bar{\mathbf{0}}_{N_c N_t} \ \forall \ u$$
$$\sum_{u=1}^{N_u} \mathbf{w}_u^H \mathbf{B}_b \mathbf{w}_u \le P \ \forall \ b.$$

The innermost minimization problem with respect to v is an unconstrained convex problem. By the stationary principle of the Lagrangian function, i.e., setting the first order derivation of the Lagrangian function to zero, we can obtain the optimal v_u for given $\mathbf{w}$ and t as

$$v_u^{\text{opt}}(\mathbf{w}) = \frac{\sqrt{\mathbf{w}_u^H \mathbf{R}_u \mathbf{w}_u} e^{j\theta_u}}{\sum_{j=1}^{N_u} \mathbf{w}_j^H \mathbf{R}_u \mathbf{w}_j + \sigma_u^2} \ \forall \ u. \tag{36}$$

By replacing v_u with v_u^{opt} in the objective function of problem (35), we can find that the minimization problem with respect to t is also an unconstrained convex problem. Based on the stationary principle of the Lagrangian function, we can obtained the optimal t_u for given $\mathbf{w}$ as

$$t_u^{\text{opt}}(\mathbf{w}) = 1 + \frac{\mathbf{w}_u^H \mathbf{R}_u \mathbf{w}_u}{\sum_{j \ne u} \mathbf{w}_j^H \mathbf{R}_u \mathbf{w}_j + \sigma_u^2} = 1 + \widehat{\text{SINR}}_u. \tag{37}$$

By replacing v_u and t_u with v_u^{opt} and t_u^{opt} respectively in the objective function of problem (35), we obtain the following

equivalent optimization problem with respect to $\mathbf{w}$

$$\begin{aligned} \min_{\mathbf{w}} \quad & \sum_{u=1}^{N_u} \alpha_u \left(1 - \log(1+\widehat{\mathrm{SINR}}_u)\right) \\ s.t. \quad & \tilde{\mathbf{D}}_u \mathbf{w}_u = \bar{\mathbf{0}}_{N_c N_t} \ \forall\ u \\ & \textstyle\sum_{u=1}^{N_u} \mathbf{w}_u^H \mathbf{B}_b \mathbf{w}_u \le P \ \forall\ b. \end{aligned} \tag{38}$$

Since α_u is a constant priority weight for MS_u, problem (38) is equivalent to the weighted sum rate estimate maximization problem (11). Therefore, problem (11) and problem (14) are equivalent in a sense that they yield the same globally optimal solution. ∎

Appendix B
Proof of Theorem 1

Proof: According to the definition of $\tilde{\mathbf{D}}_u$ and $\mathbf{B}_b$, it is not hard to verify that the Mangasarian-Fromovitz constraint qualification (MFCQ) [29] holds for all feasible vectors satisfying the constraints of problem (14). Therefore, the qualification applies at the global optimum. This implies that the Karush-Kuhn-Tucker (KKT) conditions are necessary conditions for the global optimum, hence the global optimum can be obtained by finding the solutions of the KKT conditions.

The global optimal precoders $\mathbf{w}^{\mathrm{opt}}$ satisfy the following KKT conditions,

$$\begin{aligned} & -\frac{\alpha_u t_u^{\mathrm{opt}} |v_u^{\mathrm{opt}}| \mathbf{R}_u \mathbf{w}_u^{\mathrm{opt}}}{\sqrt{\mathbf{w}_u^{\mathrm{opt},H} \mathbf{R}_u \mathbf{w}_u^{\mathrm{opt}}}} + \left(\sum_{j=1}^{N_u} \alpha_j t_j^{\mathrm{opt}} |v_j^{\mathrm{opt}}|^2 \mathbf{R}_j\right) \mathbf{w}_u^{\mathrm{opt}} \\ & + \tilde{\mathbf{D}}_u \boldsymbol{\mu}_u^{\mathrm{opt}} + \left(\sum_{b=1}^{N_c} \lambda_b^{\mathrm{opt}} \mathbf{B}_b\right) \mathbf{w}_u^{\mathrm{opt}} = \bar{\mathbf{0}}_{N_c N_t} \ \forall\ u, \end{aligned} \tag{39}$$

where $\boldsymbol{\mu}^{\mathrm{opt}}$ and λ^{opt} are the optimal Lagrange multipliers, and v_u^{opt} and t_u^{opt} can be obtained from (36) and (37) as

$$\begin{aligned} v_u^{\mathrm{opt}} &= \frac{\sqrt{\mathbf{w}_u^{\mathrm{opt},H} \mathbf{R}_u \mathbf{w}_u^{\mathrm{opt}}} e^{j\theta_u}}{\sum_{j=1}^{N_u} \mathbf{w}_j^{\mathrm{opt},H} \mathbf{R}_u \mathbf{w}_j^{\mathrm{opt}} + \sigma_u^2} \quad \text{and} \\ t_u^{\mathrm{opt}} &= \frac{\sum_{j=1}^{N_u} \mathbf{w}_j^{\mathrm{opt},H} \mathbf{R}_u \mathbf{w}_j^{\mathrm{opt}} + \sigma_u^2}{\sum_{j\neq u} \mathbf{w}_j^{\mathrm{opt},H} \mathbf{R}_u \mathbf{w}_j^{\mathrm{opt}} + \sigma_u^2} \ \forall\ u. \end{aligned} \tag{40}$$

Multiplying both sides of (39) by $\bar{\mathbf{D}}_u^H$ to eliminate $\boldsymbol{\mu}^{\mathrm{opt}}$, we have

$$\begin{aligned} & \frac{-\alpha_u t_u^{\mathrm{opt}} |v_u^{\mathrm{opt}}| \bar{\mathbf{D}}_u^H \mathbf{R}_u \mathbf{w}_u^{\mathrm{opt}}}{\sqrt{\mathbf{w}_u^{\mathrm{opt},H} \mathbf{R}_u \mathbf{w}_u^{\mathrm{opt}}}} + \left(\sum_{j=1}^{N_u} \alpha_j t_j^{\mathrm{opt}} |v_j^{\mathrm{opt}}|^2 \bar{\mathbf{D}}_u^H \mathbf{R}_j\right) \mathbf{w}_u^{\mathrm{opt}} \\ & + \left(\sum_{b=1}^{N_c} \lambda_b^{\mathrm{opt}} \bar{\mathbf{D}}_u^H \mathbf{B}_b\right) \mathbf{w}_u^{\mathrm{opt}} = \bar{\mathbf{0}}_{|\mathcal{D}_u| N_t} \ \forall\ u. \end{aligned} \tag{41}$$

By substituting (40) into (41) and after some manipulations, (41) can be rewritten as

$$\begin{aligned} \alpha_u \bar{\mathbf{D}}_u^H \mathbf{R}_u \mathbf{w}_u^{\mathrm{opt}} = & \underbrace{\frac{\mathbf{w}_u^{\mathrm{opt},H} \mathbf{R}_u \mathbf{w}_u^{\mathrm{opt}}}{\sum_{j\neq u} \mathbf{w}_j^{\mathrm{opt},H} \mathbf{R}_u \mathbf{w}_j^{\mathrm{opt}} + \sigma_u^2}}_{\widehat{\mathrm{SINR}}_u} \bar{\mathbf{D}}^H \cdot \\ & \left(\sum_{j\neq u} \alpha_j \frac{t_j^{\mathrm{opt}} |v_j^{\mathrm{opt}}|^2}{t_u^{\mathrm{opt}} |v_u^{\mathrm{opt}}|^2} \mathbf{R}_j + \sum_{b=1}^{N_c} \frac{\lambda_b^{\mathrm{opt}}}{t_u^{\mathrm{opt}} |v_u^{\mathrm{opt}}|^2} \mathbf{B}_b\right) \mathbf{w}_u^{\mathrm{opt}} \ \forall\ u. \end{aligned} \tag{42}$$

Let $\bar{\mathbf{w}}_u^{\mathrm{opt}}$ denote the non-zero optimal precoders of MS_u. Then we know from (2) that $\mathbf{w}_u^{\mathrm{opt}} = \bar{\mathbf{D}}_u \bar{\mathbf{w}}_u^{\mathrm{opt}}$. Express $\bar{\mathbf{w}}_u^{\mathrm{opt}}$ as $\bar{\mathbf{w}}_u^{\mathrm{opt}} = \sqrt{q_u^{\mathrm{opt}}} \bar{\mathbf{f}}_u^{\mathrm{opt}}$, where q_u^{opt} and $\bar{\mathbf{f}}_u^{\mathrm{opt}}$ denote the optimal power allocation and non-zero beamforming vector. Then by substituting $\mathbf{w}_u^{\mathrm{opt}} = \sqrt{q_u^{\mathrm{opt}}} \bar{\mathbf{D}}_u \bar{\mathbf{f}}_u^{\mathrm{opt}}$ into (42), we have that the optimal beamforming vector $\bar{\mathbf{f}}_u^{\mathrm{opt}}$ will be an eigenvector of the matrix

$$\begin{aligned} & \alpha_u t_u^{\mathrm{opt}} |v_u^{\mathrm{opt}}|^2 \left(\bar{\mathbf{D}}_u^H \left(\sum_{j\neq u} \alpha_j t_j^{\mathrm{opt}} |v_j^{\mathrm{opt}}|^2 \mathbf{R}_j + \sum_{b=1}^{N_c} \lambda_b^{\mathrm{opt}} \mathbf{B}_b\right) \bar{\mathbf{D}}_u\right)^{\dagger} \cdot \\ & \qquad \bar{\mathbf{D}}_u^H \mathbf{R}_u \bar{\mathbf{D}}_u. \end{aligned} \tag{43}$$

Moreover, the estimated SINR (i.e., $\widehat{\mathrm{SINR}}_u$ in (42)) satisfies $\widehat{\mathrm{SINR}}_u = d_u$, where d_u is the eigenvalue of the matrix in (43) corresponding to the eigenvector $\bar{\mathbf{f}}_u^{\mathrm{opt}}$. Thus, we have N_u linear equations for the optimal power allocation as

$$\frac{q_u^{\mathrm{opt}} \bar{\mathbf{f}}_u^{\mathrm{opt},H} \bar{\mathbf{D}}_u^H \mathbf{R}_u \bar{\mathbf{D}}_u \bar{\mathbf{f}}_u^{\mathrm{opt}}}{\sum_{j\neq u} q_j^{\mathrm{opt}} \bar{\mathbf{f}}_j^{\mathrm{opt},H} \bar{\mathbf{D}}_j^H \mathbf{R}_u \bar{\mathbf{D}}_j \bar{\mathbf{f}}_j^{\mathrm{opt}} + \sigma_u^2} = d_u \ \forall\ u. \tag{44}$$

From (44) the optimal power allocation can be solved as (17).

The optimally allocated power q_u^{opt} and optimal beamforming vector $\bar{\mathbf{f}}_u^{\mathrm{opt}}$ are determined by the matrix shown in (43), which is governed by $N_c + 2N_u$ real-valued parameters, i.e., $\{t_u^{\mathrm{opt}} |v_u^{\mathrm{opt}}|^2\}$, $\{\lambda_b^{\mathrm{opt}}\}$, and $\{d_u\}$. Moreover, the common scaling factor c in $\{t_u^{\mathrm{opt}} |v_u^{\mathrm{opt}}|^2\}$ and $\{\lambda_b^{\mathrm{opt}}\}$ does not affect the matrix in (43), and hence has no impact on q_u^{opt} and $\bar{\mathbf{f}}_u^{\mathrm{opt}}$. Thus, we can define a set of parameters between zero and one as (19). Substituting (19) into (43), we obtain the optimal precoder structure for the weighted estimated sum MSE minimization problem as (15). According to Lemma 1, the structure is also optimal for the weighted sum rate estimate maximization problem (11). ∎

Appendix C
Proof of Property 3

Proof: The optimal Lagrange multipliers λ^{opt} satisfies the KKT conditions in (39). By substituting (40) into (39) and then multiplying both sides by $\mathbf{w}^{\mathrm{opt},H}$, we can rewrite (39) as

$$\begin{aligned} \sum_{b=1}^{N_c} \lambda_b^{\mathrm{opt}} \mathbf{w}_u^{\mathrm{opt},H} \mathbf{B}_b \mathbf{w}_u^{\mathrm{opt}} = & \underbrace{\frac{\alpha_u \mathbf{w}_u^{\mathrm{opt},H} \mathbf{R}_u \mathbf{w}_u^{\mathrm{opt}}}{\sum_{j=1}^{N_u} \mathbf{w}_j^{\mathrm{opt},H} \mathbf{R}_u \mathbf{w}_j^{\mathrm{opt}} + \sigma_u^2}}_{\le \alpha_u} - \\ & \underbrace{\mathbf{w}_u^{\mathrm{opt},H} \left(\sum_{j\neq u} \alpha_j t_j^{\mathrm{opt}} |v_j^{\mathrm{opt}}|^2 \mathbf{R}_j\right) \mathbf{w}_u^{\mathrm{opt}}}_{\ge 0} \ \forall\ u. \end{aligned} \tag{45}$$

From (45), we can obtain $\sum_{b=1}^{N_c} \lambda_b^{\mathrm{opt}} \mathbf{w}_u^{\mathrm{opt},H} \mathbf{B}_b \mathbf{w}_u^{\mathrm{opt}} \le \alpha_u \ \forall\ u$, and hence

$$\sum_{b=1}^{N_c} \lambda_b^{\mathrm{opt}} \sum_{u=1}^{N_u} \mathbf{w}_u^{\mathrm{opt},H} \mathbf{B}_b \mathbf{w}_u^{\mathrm{opt}} \le \sum_{u=1}^{N_u} \alpha_u. \tag{46}$$

Since λ^{opt} needs to satisfy the KKT complementary conditions $\lambda_b^{\mathrm{opt}} \left(\sum_{u=1}^{N_u} \mathbf{w}_u^{\mathrm{opt},H} \mathbf{B}_b \mathbf{w}_u^{\mathrm{opt}} - P\right) = 0$, i.e., $\lambda_b^{\mathrm{opt}} \sum_{u=1}^{N_u} \mathbf{w}_u^{\mathrm{opt},H} \mathbf{B}_b \mathbf{w}_u^{\mathrm{opt}} = \lambda_b^{\mathrm{opt}} P$, we can obtain the property of (20) from (46). ∎

ACKNOWLEDGEMENT

The authors would like to thank Prof. Zhiquan Luo and Dr. Yafeng Liu for helpful discussions.

REFERENCES

[1] M. K. Karakayali, G. J. Foschini, and R. A. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems," *IEEE Wireless Commun. Mag.*, vol. 13, no. 4, pp. 56–61, Aug. 2006.

[2] C. Yang, S. Han, X. Hou, and A. F. Molisch, "How do we design CoMP to achieve its promised potential?" *IEEE Wireless Commun. Mag.*, vol. 20, no. 1, pp. 67–74, Feb. 2013.

[3] D. A. Schmidt, C. Shi, R. A. Berry, M. L. Honig, and W. Utschick, "Comparison of distributed beamforming algorithms for MIMO interference networks," *IEEE Trans. Signal Process.*, vol. 61, no. 13, pp. 3476–3489, July 2013.

[4] Y. Huang, G. Zheng, M. Bengtsson, K.-K. Wong, L. Yang, and B. Ottersten, "Distributed multicell beamforming design approaching pareto boundary with max-min fairness," *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 2921–2933, Aug. 2012.

[5] C. T. K. Ng and H. Huang, "Linear precoding in cooperative MIMO cellular networks with limited coordination clusters," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1446–1454, Sep. 2010.

[6] A. Papadogiannis, D. Gesbert, and E. Hardouin, "A dynamic clustering approach in wireless networks with multi-cell cooperative processing," in *Proc. 2008 IEEE International Commun. Conf.*

[7] M. Hong, R. Sun, H. Baligh, and Z.-Q. Luo, "Joint base station clustering and beamformer design for partial coordinated transmission in heterogeneous networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 226–240, Feb. 2013.

[8] E. Björnson, N. Jalden, M. Bengtsson, and B. Ottersten, "Optimality properties, distributed strategies, and measurement-based evaluation of coordinated multicell OFDMA transmission," *IEEE Trans. Signal Process.*, vol. 59, no. 99, pp. 6086–6101, Dec. 2011.

[9] T. E. Bogale and L. Vandendorpe, "Robust sum MSE optimization for downlink multiuser MIMO systems with arbitrary power constraint: generalized duality approach," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 1862–1875, Apr. 2012.

[10] T. M. Kim, F. Sun, and A. J. Paulraj, "Low-complexity MMSE precoding for coordinated multipoint with per-antenna power constraint," *IEEE Signal Process. Lett.*, vol. 20, no. 4, pp. 395–398, Apr. 2013.

[11] F. Sun and E. de Carvalho, "A leakage-based MMSE beamforming design for a MIMO interference channel," *IEEE Signal Process. Lett.*, vol. 19, no. 6, pp. 368–371, June 2012.

[12] 3GPP R1- 091752, "Performance study on Tx/Rx mismatch in LTE TDD dual-layer beamforming," Nokia, Nokia Siemens Networks, CATT, ZTE, 2009.

[13] F. Yuan and C. Yang, "Bit allocation between per-cell codebook and phase ambiguity quantization for limited feedback coordinated multipoint transmission systems," *IEEE Trans. Commun.*, vol. 60, no. 9, pp. 2546–2559, Sep. 2012.

[14] X. Hou and C. Yang, "Feedback overhead analysis for base station cooperative transmission," *IEEE Trans. Wireless Commun.*, 2013, early access.

[15] S. Hyeon, C. Lee, C. Shin, and S. Choi, "Implementation of a smart antenna base station for mobile WiMAX based on OFDMA," *EURASIP J. Wireless Commun. and Networking*, 2009.

[16] J. Shi, Q. Luo, and M. You, "An efficient method for enhancing TDD over the air reciprocity calibration," in *Proc. 2011 IEEE Wireless Commun. and Networking Conf.*

[17] E. Björnson, G. Zheng, M. Bengtsson, and B. Ottersten, "Robust monotonic optimization framework for multicell MISO systems," *IEEE Trans. Signal Process.*, vol. 60, no. 5, pp. 2508–2523, May 2012.

[18] S. Han, C. Yang, G. Wang, D. Zhu, and M. Lei, "Coordinated multipoint transmission with non-ideal channel reciprocity," in *Proc. 2012 IEEE Wireless Commun. and Networking Conf.*

[19] M. Sadek, A. Tarighat, and A. Sayed, "A leakage-based precoding scheme for downlink multi-user MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, pp. 1711–1721, May 2007.

[20] D. Lee, H. Seo, B. Clerckx, E. Hardouin, D. Mazzarese, S. Nagata, and K. Sayana, "Coordinated multipoint transmission and reception in LTE-advanced: deployment scenarios and operational challenges," *IEEE Commun. Mag.*, vol. 50, no. 2, pp. 148–155, Feb. 2012.

[21] 3GPP R1-091794, "Hardware calibration requirement for dual layer beamforming," Huawei, 2009.

[22] Y.-F. Liu, Y.-H. Dai, and Z.-Q. Luo, "Coordinated beamforming for MISO interference channel: complexity analysis and efficient algorithms," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1142–1157, Mar. 2011.

[23] Q. Shi, M. Razaviyayn, Z. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, Sept. 2011.

[24] T. E. Bogale, B. K. Chalise, and L. Vandendorpe, "Robust transceiver optimization for downlink multiuser MIMO systems," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 446–453, Jan. 2011.

[25] G. Dimic and N. D. Sidiropoulos, "On downlink beamforming with greedy user selection: performance analysis and a simple new algorithm," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3857–3868, Oct. 2005.

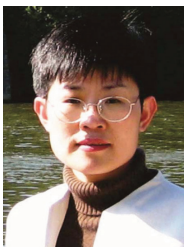
[26] R. Zakhour and D. Gesbert, "Distributed multicell-MISO precoding using the layered virtual SINR framework," *IEEE Trans. Wireless Commun.*, vol. 9, no. 8, pp. 2444–2448, Aug. 2010.

[27] D. Hammarwall, M. Bengtsson, and B. Ottersten, "Utilizing the spatial information provided by channel norm feedback in SDMA systems," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3278–3293, July 2008.

[28] S. V. Hanly, L. L. H. Andrew, and T. Thanabalasingham, "Dynamic allocation of subcarriers and transmit powers in an OFDMA cellular network," *IEEE Trans. Inf. Theory*, vol. 55, no. 12, pp. 5445–5462, Dec. 2009.

[29] O. L. Mangasarian and S. Fromovitz, "The Fritz John necessary optimality conditions in the presence of equality and inequality constraints," *J. Math. Analysis and Applications*, vol. 17, pp. 37–47, 1967.

Shengqian Han (S'05–M'12) received his B.S. degree in communication engineering and Ph.D. degree in signal and information processing from Beihang University (formerly Beijing University of Aeronautics and Astronautics), Beijing, China, in 2004 and 2010 respectively. From 2010 to 2012, he held a postdoctoral research position with the School of Electronics and Information Engineering, Beihang University. He is currently a Lecturer of the same school. His research interests are multiple antenna techniques, cooperative communication and energy efficient transmission in the areas of wireless communications and signal processing.



Chenyang Yang (SM'08) received the M.S.E and Ph.D. degrees in electrical engineering from Beihang University (formerly Beijing University of Aeronautics and Astronautics), Beijing, China, in 1989 and 1997, respectively.

She is currently a Full Professor with the School of Electronics and Information Engineering, Beihang University. She has published various papers and filed many patents in the fields of signal processing and wireless communications. Her recent research interests include signal processing in network MIMO, cooperative communication, energy efficient transmission and interference management.

Prof. Yang was the Chair of the IEEE Communications Society Beijing chapter from 2008 to 2012. She has served as Technical Program Committee Member for many IEEE conferences, such as the IEEE International Conference on Communications and the IEEE Global Telecommunications Conference. She currently serves as an Associate editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, an Associate Editor-in-Chief of the *Chinese Journal of Communications*, and an Associate Editor-in-Chief of the *Chinese Journal of Signal Processing*. She was nominated as an Outstanding Young Professor of Beijing in 1995 and was supported by the First Teaching and Research Award Program for Outstanding Young Teachers of Higher Education Institutions by Ministry of Education (P.R.C. TRAPOYT) during 1999-2004.

Gang Wang received the Ph.D. degree from Tsinghua University in 2006. He joined NEC Laboratories China in 2006 and worked on wireless networking. From 2009, he started research on 3GPP LTE standardization. He visited UC Davis in 2011 and worked with Prof. Zhi Ding on CoMP. He is now a senior researcher at NEC Laboratories China. His research interest includes MIMO, CoMP, dynamic TDD system and coverage enhancement.

Dalin Zhu received his B.E. and M.S. degrees in Information Engineering and Electrical Engineering from Beijing University of Posts and Telecommunications (BUPT) and Kansas State University in 2007 and 2009, respectively. He was a visiting scientist at Stanford University from 2012-2013. Currently, he is a senior research staff member of wireless communications department at NEC Laboratories China (NLC) since 2009. His research interests include 3GPP standardization, LTE/LTE-A, traffic adaptation and interference management, CoMP and green radios. He has published various papers, filed many patents in the fields of signal processing and wireless communications. He also contributed to 3GPP standardization body on CoMP and eIMTA. He has served as Technical Program Committee Member for many IEEE conferences, such as ICC, Globecom and VTC. He received excellent researcher award from NLC in 2011.

Ming Lei received the B.Eng. degree from the Southeast University in 1998 and the Ph.D. degree from BUPT (Beijing University of Posts & Telecommunications) in 2003, all in Electrical Engineering. From April 2003 to February 2008, he was a research scientist with the National Institute of Information and Communications Technology (NICT), Japan, where he contributed to Japan's national projects on 4G mobile communications (MIRAI projects) and IEEE standardization of 60-GHz multi-gigabit WPAN (IEEE 802.15.3c). From March 2008 to May 2009, he was a project lead of Intel Corporation, where he contributed to the standardization of WiGig (60-GHz WPAN), IEEE 802.11ad (60-GHz WLAN) and IEEE 802.16m (mobile WiMAX). Since May 2009, he has been with NEC Laboratories China, NEC Corporation, as the department head managing the wireless research and standardization projects on 4G cellular mobile communications (LTE and LTE-Advanced), 60-GHz, mobile backhaul, etc. Dr. Ming Lei was elected to IEEE Senior Member in 2009.