

Multi-Cell Interference Exploitation: Enhancing the Power Efficiency in Cell Coordination

Zhongxiang Wei, *Member, IEEE*, Christos Masouros, *Senior Member, IEEE*, Kai-Kit Wong, *Fellow, IEEE*, Xin Kang, *Member, IEEE*

Abstract—In this paper, we propose a series of novel coordination schemes for multi-cell downlink communication. Starting from full base station (BS) coordination, we first propose a fully-coordinated scheme to exploit beneficial effects of both inter-cell and intra-cell interference, based on sharing both channel state information (CSI) and data among the BSs. To reduce the coordination overhead, we then propose a partially-coordinated scheme where only intra-cell interference is designed to be constructive while inter-cell is jointly suppressed by the coordinated BSs. Accordingly, the coordination only involves CSI exchange and the need for sharing data is eliminated. To further reduce the coordination overhead, a third scheme is proposed, which only requires the knowledge of statistical inter-cell channels, at the cost of a slight increase on the transmission power. For all the proposed schemes, imperfect CSI is considered. We minimize the total transmission power in terms of probabilistic and deterministic optimizations. Explicitly, the former statistically satisfies the users' signal-to-interference-plus-noise ratio (SINR) while the latter guarantees the SINR requirements in the worst case CSI uncertainties. Simulation verifies that our schemes consume much lower power compared to the existing benchmarks, i.e., coordinated multi-point (CoMP) and coordinated-beamforming (CBF) systems, opening a new dimension on multi-cell coordination.

Index Terms—Multi-Cell Coordination, Constructive Interference, Robust Optimization, Power Efficiency

I. INTRODUCTION

With the exponential growth of the throughput demand in cellular networks, aggressive radio spectrum reuse has emerged as an attractive strategy to enhance the utilization of spectrum resource [1]. However, frequency reuse, if treated inappropriately, may result in severe interference, especially for cell-edge users. By the conventional single-cell based signal processing, the critically increased level of inter-cell interference leads to significant losses in throughput and fairness degradation among the users. To this end, multi-cell cooperation is regarded as a key technique for inter-cell interference management in cellular systems, where multiple base stations (BSs) collaborate with others for joint signal processing, based on the shared channel state information (CSI) and/or intended transmitted data [2] [3]. Depending on the coordination level

among the BSs, multi-cell coordination system can be classified into two categories: coordinated-beamforming (CBF) and coordinated multi-point system (CoMP), and each of them is suited for different scenarios and system configurations.

In the CBF scheme, a user is supported by a single BS and each BS only needs to encode or decode the signal to/from the users within its cell. Hence, the impairment caused by the inter-cell interference is suppressed by cooperative beamforming, and the beamforming at each BS is computed as a function of globally shared CSI, obtained from the users via feedback channels [4]. Since the BSs jointly optimize their beams to suppress the inter-cell gains, the CBF scheme is also named as interference coordination. In this spirit, several designs have been conducted to optimize system performance, e.g., throughput maximization [4]-[10] and power minimization [11] [12]. Downlink throughput [4] [5] [7] [8] [10] or uplink throughput maximization [6] [9] was investigated, where BSs only transmit or decode the signal to/from the users within its cell, and inter-cell interference was mitigated as noise. Most relevant to the approaches considered in this paper, power minimization problem was investigated in [11] and [12]. The authors in [11] minimized power beamformer for a CBF system, where the minimum beamformers can be obtained locally at each BS relying on limited backhaul information exchange among the BSs. The authors in [12] proposed a power efficient transceiver design for multi-cell coordination in cognitive radio networks, where secondary BSs are coordinated to serve secondary users while imposing limited interference for the primary users. In summary, the beamforming design by the CBF scheme typically strikes the trade-off between eliminating the inter-cell interference and maximizing the signal-to-interference-plus-noise ratio (SINR) to/from the user within the cell of interest, which requires a modest amount of coordination overhead.

With a higher level of coordination, the performance of multi-cell systems can be improved by CoMP scheme if BSs are linked by high-capacity delay-free links. Explicitly, the coordinated BSs can share not only CSI, but also the data to be transmitted. This is typically facilitated through a high bandwidth backhaul network implemented by optical fiber or millimeter wave links [2]. Since the intended transmitted data can be available at all the coordinated BSs, CoMP system transforms the multi-cell system into a virtual multi-user multiple-input and multiple-output (MIMO) for which all links (including interfering ones) are exploited to carry useful data [13]. Hence, CoMP is also referred to network-level virtual MIMO, which coordinates the simultaneous in-

Zhongxiang Wei, Christos Masouros, and Kai-Kit Wong are with the Department of Electronic and Electrical Engineering at the University College London, London, UK. Email: {zhongxiang.wei, c.masouros, kai-kit.wong}@ucl.ac.uk

Xin Kang is with Shield Lab., Huawei Singapore Research Center, Singapore. Email: kang.xin@huawei.com

This work was supported by the Engineering and Physical Sciences Research Council, UK, under project EP/R007934/1.

formation transmissions from multiple BSs to the users, at the expense of increased coordination overhead for sharing the intended transmitted data. Based on the principle, the authors in [14] considered the joint optimization of user association, subchannel allocation, and power allocation in a heterogeneous CoMP system, where each user is associated with one macro BS in the center and multiple separated micro BSs. The authors in [15] applied the CoMP into non-orthogonal multiple access systems, where CoMP transmission is to serve users experiencing severe multi-user interference. The authors in [16] investigated beamforming design for CoMP systems with imperfect CSI, where the fully-coordinated BSs jointly serve users through a virtual multi-user multiple-input single-output (MISO) channel. From an information theoretic point of view, CBF scheme can be interpreted as a interference channel while CoMP scheme can be interpreted as a broadcast channel, and the latter generally achieves better performance over the former scheme by exploiting inter-cell channel, at the cost of high overhead.

It is worth noting that, regardless of CBF or CoMP schemes, intra-cell interference (multi-user interference) still needs to be strictly mitigated. Traditional methods cancel multi-user interference by providing users with exclusive access to a fraction of wireless resources, such as time division multiple access (TDMA), frequency division multiple access (FDMA) or frequency reuse. Since each user can only utilize a part of wireless resources, the throughput performance of the orthogonality based approaches is significantly limited. To this end, interference alignment (IA), a degree of freedom optimal approach, was proposed to further improve the system throughput [17] [18]. A well-known example of IA is the MIMO technique, which is also referred to as spatial diversity based multiple access. Nevertheless, there is scope to exploit interference as a source of useful signal based on the concept of constructive interference (CI), which can be utilized as a green source and effectively improves the received SINR [1]. CI was firstly introduced by [19] in code division multiple access systems. Then [20] further proposed that all the interference can be constructive by designing precoding in symbol level [21] [22]. Recently, the concept of CI was applied into beamforming optimization [23] [24], cognitive radio [25], large-scale MIMO [26], multi-user MISO [27], wireless power transfer systems [28], full-duplex [29] and physical layer security-aware systems [30]. However, it should be pointed that all the aforementioned research only focused on single-cell systems, which cannot be directly applied to multi-cell coordination systems. The main challenges come from: (a) In multi-cell systems, both the multi-user and inter-cell interference need to be managed, and the design principle is directly determined by the level of multi-cell coordination. For example, the single cell methods do not apply to the case where only CSI is shared among the cooperative BSs. Indeed, how to manage the rich interference with different levels of coordination, and strike a good trade-off between utilizing CI and overhead is challenging. Importantly, no such overhead issues need to be considered in classical single-cell CI approaches, while the relation between the interference management strategy and the incurred overhead needs to be

addressed in our work. (b) By a low-level coordination, only CSI or statistical inter-cell CSI is shared among the coordinated BSs while the transmitted data is only known by the local BSs. Hence, the incompleteness of the transmitted data or CSI information makes exploiting inter-cell interference as a constructive element impossible. Unfortunately, how to exploit CI for handling multi-user interference while suppressing inter-cell interference, especially in a stochastic manner, is unknown. (c) Considering imperfect CSI in practice, how to guarantee multiple users' SINR requirements under infinite possibilities of CSI error imposes new challenges for CI design in multi-cell systems.

Motivated by the aforementioned open challenges, in this paper, we investigate different multi-cell coordination schemes to fully/partially utilize inter-cell and multi-user interference, in the presence of imperfect or statistical CSI. Our contributions are summarized as follows:

- 1) We propose a series of schemes to fully/partially utilize inter-cell and multi-user interference for multi-cell coordination systems, which require different levels of overhead for coordination. Firstly, a fully-coordinated CI (Full-CI) scheme is proposed to utilize both inter-cell and multi-user interference as beneficial elements, on the basis of CSI and data being shared among the coordinated BSs.
- 2) Then, a partially-coordinated CI (Partial-CI) scheme is proposed to utilize multi-user interference while suppressing inter-cell interference by joint precoding design. Since only CSI needs to be shared among BSs, the coordination overhead is significantly reduced compared to the Full-CI scheme.
- 3) At last, a statistical inter-channel based CI (Stat-CI) scheme is proposed to further reduce the coordination overhead, where BSs only need to know statistical CSI from the adjacent BSs for inter-cell interference suppressing while keeping multi-user interference constructive.
- 4) In all the above designs, we exploit robust precoding design to accommodate scenarios with imperfect CSI. In the presence of CSI errors, we investigate the total power minimization problems in probabilistic and deterministic manners, respectively. Explicitly, multiple users' SINR requirements are issued by outage-probability constrained formulations from the perspective of probabilistic robust optimization, while the users' SINR requirements are guaranteed with all the CSI uncertainties from the perspective of deterministic robust optimization.
- 5) To strike a trade-off between system performance and coordination overhead, five corresponding low-complexity algorithms are proposed to minimize transmit power consumption for the three schemes in terms of probabilistic and deterministic manners. The Full-CI scheme based algorithms consume the least power consumption by utilizing both inter-cell and multi-user interference; the Stat-CI based algorithm requires the smallest coordination overhead; and the Partial-CI scheme

based algorithms make a proper trade-off between the power consumption and coordination overhead. The performance of the proposed algorithms are benchmarked by the existing CBF and CoMP. The complexities and coordination overhead of the algorithms are analytically demonstrated.

Notations: Matrices and vectors are represented by boldface capital and lower case letters, respectively. $| \cdot |$ denotes the absolute value of a complex scalar. $\| \cdot \|$ denotes the Euclidean vector norm. \mathbf{A}^H , \mathbf{A}^T and $\text{Tr}(\mathbf{A})$ denote the Hermitian transpose, transpose and trace of matrix \mathbf{A} . $\text{Rank}(\mathbf{A})$ denotes the rank of matrix \mathbf{A} . $\text{diag}(\mathbf{A})$ returns a diagonal matrix with diagonal elements from matrix \mathbf{A} and $\text{diag}(\mathbf{a})$ stacks the elements of vector \mathbf{a} into a diagonal matrix. $\mathbf{A} \succeq 0$ means \mathbf{A} is a positive semi-definite matrix. Superscript \Re and \Im denote the real and imaginary parts, respectively. $\|\cdot\|_p$ means the p -norm of a vector or a matrix. \mathbf{I}_n means an n -by- n identity matrix. Operator $\text{vec}(\mathbf{A})$ stacks the element of matrix \mathbf{A} into a vector. $\mathbb{C}^{N \times M}$ and $\mathbb{H}^{N \times M}$ denote sets of all $N \times M$ matrices and Hermitian matrices with complex entries.

II. SYSTEM MODEL AND CONSTRUCTIVE INTERFERENCE

In this section, system model is first introduced in II-A and then the concept of CI is briefly discussed in II-B.

A. System Model

We consider a multi-cell system at downlink transmission, where the coordinated BSs exchange CSI and/or data based on different schemes. Assume that each cell has one BS located in the center, and each BS is equipped with M antennas for transmission. Without loss of generality, we assume there are N_{BS} coordinated BSs and K users in each cell. Each user is equipped with one antenna for simplicity. CSI is obtained by channel estimation in the training phase, based on channel reciprocity as in [4] [31] [35]. By classic channel estimation methods, such as minimum mean square error (MMSE), the channel estimation error can be modeled by the uncertainty-unbounded [32]-[35] or uncertainty-bounded models [36]-[39], and the two models correspond to probabilistic and deterministic manners robust optimization, respectively. Hence, to fully investigate the impact of channel estimation error on the robust optimization design, both uncertainty-unbounded and uncertainty-bounded models are exploited in the paper, and the inherited relation of the two models are also been introduced by Lemmas 2 and 3.

Define U_{ik} as the k -th user located in the i -th cell. Then its received signal is expressed as

$$y_{ik} = \sum_{n=1}^N \mathbf{h}_{in}^T \mathbf{w}_{in} s_{in} + \sum_{j \neq i} \sum_{m=1}^{N_{BS}} \mathbf{h}_{jik}^T \mathbf{w}_{jm} s_{jm} + n_{ik}; \quad (1)$$

where $\mathbf{h}_{jik} \in \mathbb{C}^{M \times 1}$ represents the channel from the j -th BS to the user U_{ik} . $\mathbf{w}_{in} \in \mathbb{C}^{M \times 1}$ and s_{in} denote the precoding and transmitted data at the i -th BS to the user U_{in} . $n_{ik} \in \mathbb{C}$ denotes the additive white Gaussian noise (AWGN) at the user U_{ik} , following $n_{ik} \sim \mathcal{CN}(0; \frac{2}{n})$; $\delta_i \in \mathbb{R}$ and $\delta_k \in \mathbb{R}$. Define \mathcal{I} and \mathcal{J} as the sets of users belonging to the i -th and

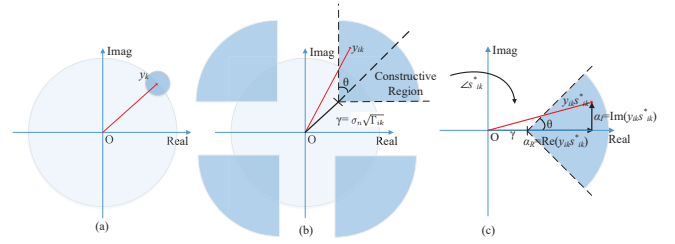


Fig. 1. Optimization region for constructive interference exploitation, QPSK for example. a) Conventional precoding, where classical SINR constraints result in a proximity region around the constellation point. b) By CI design, the received signal y_{ik} falls into a constructive region (dark blue area). The design pushes the resultant symbol away from the original decision threshold of the constellation, where $\theta = \arctan(\frac{\Im\{y_{ik}\}}{\Re\{y_{ik}\}})$. c) After being rotated by $\angle S_{ik}$, the rotated signal (noise excluded) is projected on real axis $R = \Re\{y_{ik} S_{ik}\}$ and imaginary axis $I = \Im\{y_{ik} S_{ik}\}$. According to the geometric interpretation, the received signal falls into constructive region (in Fig. 1 (b)) as long as the inequality $|I| \leq (R - \tau) \tan \theta$ holds (in Fig. 1 (c)), where $\tau = \frac{1}{M}$ and M denotes constellation size.

j -th BSs, respectively. Conventionally by CBF, the SINR of the user U_{ik} is calculated as

$$\text{SINR}_{ik}^{\text{CBF}} = \frac{\mathbb{P}_{j=1}^{N_{BS}} \mathbb{P}_{k \in \mathcal{K}; k \in \mathcal{K} \setminus \mathcal{I}} \mathbf{h}_{jik}^T \mathbf{w}_{jk} \mathbf{w}_{jk}^H}{\mathbb{P}_{j \neq i}^{N_{BS}} \mathbb{P}_{m=1}^K \mathbf{h}_{jik}^T \mathbf{w}_{jm} \mathbf{w}_{jm}^H + \frac{2}{n}}; \quad (2)$$

where the first and second terms in the denominator represent the intra-cell (multi-user) and inter-cell interference, respectively. For the case of CoMP, the system essentially reduces to a network-level coordinated multi-user MISO downlink, in which case the SINR is calculated as

$$\text{SINR}_{ik}^{\text{CoMP}} = \frac{\mathbb{P}_{j=1}^{N_{BS}} \mathbf{h}_{jik}^T \mathbf{w}_{jk} \mathbf{w}_{jk}^H}{\mathbb{P}_{k \in \mathcal{K}; k \in \mathcal{K} \setminus \mathcal{I}} \mathbf{h}_{jik}^T \mathbf{w}_{jk} \mathbf{w}_{jk}^H + \frac{2}{n}}; \quad (3)$$

where the first term in the denominator represents the virtual intra-cell (multi-user) interference and needs to be mitigated. Evidently, CoMP serves multiple users through a broadcast channel.

B. Constructive Interference

In CI, the interference is exploited to push the received signals away from the detection thresholds of the signal constellation. The increased distance to the detection threshold can effectively improve the receiving performance. The CI concept has been thoroughly discussed in the recent literature, and to avoid extensive repetition we refer the readers to [20]. Since the transmitted symbol can be written as $S_{ik} = d_{ik} e^{j\angle(S_{ik})}$, for the purpose of the CI analysis it is convenient to express the symbol by another symbol such that $S_{in} = S_{ik} e^{j(\angle(S_{in}) - \angle(S_{ik}))}$. Hence, the received signal in (1) can be re-written as

$$y_{ik} = \sum_{n=1}^N \mathbf{h}_{in}^T \mathbf{w}_{in} e^{j(\angle(S_{in}) - \angle(S_{ik}))} S_{ik} + \sum_{j \neq i} \sum_{m=1}^{N_{BS}} \mathbf{h}_{jik}^T \mathbf{w}_{jm} e^{j(\angle(S_{jm}) - \angle(S_{ik}))} S_{ik} + n_{ik}; \quad (4)$$

Define $\mathbf{h}_{jik} = \mathbf{h}_{jik} e^{j(\phi_{11} - \phi_{ik})}$, and $\mathbf{w}_j = \sum_{m=1}^K \mathbf{w}_{jm} e^{j(\phi_{jm} - \phi_{11})}$, $\forall j \in \mathcal{N}_{BS}$. Note that the phase ϕ_{11} of the reference symbol s_{11} can be arbitrary. Now, the received signal at the user U_{ik} can be equivalently expressed as

$$y_{ik} = \sum_{j=1}^{N_{BS}} \mathbf{h}_{jik}^T \mathbf{w}_j s_{11} + n_{ik}. \quad (5)$$

The reformulation in (5) indicates that the original broadcast channel reduces to a virtual multicast channel with common messages to all users [28]. Hence, by using the geometrical interpretation in Fig. 1, the requirement of generating CI can be given as

$$\sum_{j=1}^{N_{BS}} \mathbf{h}_{jik}^T \mathbf{w}_j > \gamma_{ik} \sum_{j=1}^{N_{BS}} \mathbf{h}_{jik}^T \mathbf{w}_j + n_{ik} \tan \gamma_{ik}, \quad \forall U_{ik}; \quad (6)$$

where the user's SINR requirement is also embedded. It indicates that generating CI and satisfying the user U_{ik} 's SINR requirement are both guaranteed.

III. FULLY-COORDINATED CI PRECODING WITH IMPERFECT CSI

In Section III, we investigate the power efficient multi-cell CI design for the fully-coordinated systems, where the channel from the j -th BS to the user U_{ik} is given as $\mathbf{h}_{jik} = \hat{\mathbf{h}}_{jik} + \mathbf{e}_{ik}$, $\forall k \in \mathcal{K}; j \in \mathcal{N}_{BS}$. $\hat{\mathbf{h}}_{jik} \in \mathbb{C}^{M \times 1}$ denotes the estimated channel with estimation error $\mathbf{e}_{ik} \in \mathbb{C}^{M \times 1}$. In subsection III-A, we first adopt the uncertainty-unbounded model for the channel estimation error, where the element of channel estimation error vector follows Gaussian distribution such that $[e_{ik}]_m \sim \mathcal{CN}(0, \frac{1}{2} \sigma_{ik}^2)$, $\forall m \in \mathcal{M}$ [32]-[35]. While in subsection III-B, the uncertainty-bounded model is adopted for the channel estimation error, where the error is assumed to be bounded by an ellipsoidal uncertainty region.

A. Probabilistic Robust Optimization

The optimization problem is first formulated in a probabilistic manner by P1. After a series of mathematical manipulations for handling the probabilistic SINR constraints, the problem is finally given by P2.

1) *Problem Formulation*: Define precoding vector $\mathbf{w}_j \in \mathbb{C}^{M \times 1}$, $\forall j \in \mathcal{N}_{BS}$, as the precoding vector at the j -th BS and \mathcal{N}_{BS} as the set of the coordinated BSs. The optimization problem can be given as

$$\begin{aligned} P1 \text{ (Full CI Prob)}: & \arg \min_{\{\mathbf{w}_j\}_{j \in \mathcal{N}_{BS}}} \sum_{j=1}^{N_{BS}} k \mathbf{w}_j^T \mathbf{w}_j; \\ \text{s.t (C1):} & k \mathbf{w}_j^T \mathbf{w}_j \leq p_{max}, \forall j \in \mathcal{N}_{BS}; \\ \text{(C2):} & \Pr \{ \gamma_{ik} \leq \frac{\sum_{j=1}^{N_{BS}} \mathbf{h}_{jik}^T \mathbf{w}_j}{\sum_{j=1}^{N_{BS}} \mathbf{h}_{jik}^T \mathbf{w}_j + n_{ik} \tan \gamma_{ik}} \} \leq \delta, \forall U_{ik}; \end{aligned} \quad (7)$$

where γ_{ik} denotes the SINR requirement for the user U_{ik} . δ denotes the probabilistic threshold for the user U_{ik} . Evidently, (C1) imposes individual transmission power constraint at each BS. (C2) guarantees that the SINR constraints at the user with probabilities δ greater than the SINR requirement γ_{ik} .

2) *Solution to the Problem*: To solve the problem, at first we need to handle the probabilistic constraint (C2). Under the provision of CI, the constraint (C2) equals to

$$\begin{aligned} (C2): & \Pr \{ \gamma_{ik} \leq \frac{\sum_{j=1}^{N_{BS}} \mathbf{h}_{jik}^T \mathbf{w}_j}{\sum_{j=1}^{N_{BS}} \mathbf{h}_{jik}^T \mathbf{w}_j + n_{ik} \tan \gamma_{ik}} \} \\ & \Pr \{ \sum_{j=1}^{N_{BS}} \mathbf{h}_{jik}^T \mathbf{w}_j \geq (\sum_{j=1}^{N_{BS}} \mathbf{h}_{jik}^T \mathbf{w}_j + n_{ik} \tan \gamma_{ik}) \} \\ & \Pr \{ \sum_{j=1}^{N_{BS}} \mathbf{h}_{jik}^T \mathbf{w}_j \geq n_{ik} \tan \gamma_{ik} \} \end{aligned} \quad (8)$$

Removing the absolute operator in (8), (C2) is equivalent to the two following inequalities in (9). We now focus our attention on the first inequality, which can be expanded into

$$\Pr \{ \sum_{j=1}^{N_{BS}} [\hat{\mathbf{h}}_{jik}^T + \mathbf{e}_{ik}^T] \mathbf{w}_j \geq n_{ik} \tan \gamma_{ik} \} \quad (10)$$

For simplicity, we rewrite (10) as $\Pr \{ \sum_{j=1}^{N_{BS}} \mathbf{a}_{jik}^T [\mathbf{w}_j^< ; \mathbf{w}_j^>] \geq n_{ik} \tan \gamma_{ik} \}$, where $\mathbf{a}_{jik} = [\hat{\mathbf{h}}_{jik}^T + \mathbf{e}_{ik}^T; \hat{\mathbf{h}}_{jik}^T \tan \gamma_{ik} + \mathbf{e}_{ik}^T \tan \gamma_{ik}; \hat{\mathbf{h}}_{jik}^T + \mathbf{e}_{ik}^T + \hat{\mathbf{h}}_{jik}^T \tan \gamma_{ik} + \mathbf{e}_{ik}^T \tan \gamma_{ik}]$. It is easy to obtain that the $2M$ -dimensional Gaussian distributed vector \mathbf{a}_{jik} 's expectation is $\bar{\mathbf{a}}_{jik} = [\hat{\mathbf{h}}_{jik}^T; \hat{\mathbf{h}}_{jik}^T \tan \gamma_{ik}; \hat{\mathbf{h}}_{jik}^T + \hat{\mathbf{h}}_{jik}^T \tan \gamma_{ik}]$ with covariance matrix $\Sigma_{jik} = \text{diag}(\underbrace{(1 + \tan^2 \gamma_{ik})^2}_{2M}, \dots, \underbrace{(1 + \tan^2 \gamma_{ik})^2}_{2M})$.

We now stack the vectors \mathbf{w}_j and \mathbf{a}_{jik} , $\forall j \in \mathcal{N}_{BS}$, into two long vectors \mathbf{w} and \mathbf{a}_{ik} , such that $\mathbf{w} = \text{vec}(\mathbf{w}_1^< ; \mathbf{w}_1^> ; \mathbf{w}_2^< ; \mathbf{w}_2^> ; \dots ; \mathbf{w}_{N_{BS}}^< ; \mathbf{w}_{N_{BS}}^>)$ and $\mathbf{a}_{ik} = \text{vec}(\mathbf{a}_{1ik}^< ; \mathbf{a}_{1ik}^> ; \dots ; \mathbf{a}_{N_{BS}ik}^< ; \mathbf{a}_{N_{BS}ik}^>)$. It is observed that the long vector \mathbf{a}_{ik} 's expectation is calculated as $\bar{\mathbf{a}}_{ik} = \text{vec}(\bar{\mathbf{a}}_{1ik}^< ; \bar{\mathbf{a}}_{1ik}^> ; \dots ; \bar{\mathbf{a}}_{N_{BS}ik}^< ; \bar{\mathbf{a}}_{N_{BS}ik}^>)$ with covariance matrix $\Sigma_{ik} = \text{diag}(\underbrace{(1 + \tan^2 \gamma_{ik})^2}_{2M}, \dots, \underbrace{(1 + \tan^2 \gamma_{ik})^2}_{2M})$. Now (10) can be derived as

$$\begin{aligned} & \Pr \{ \mathbf{a}_{ik}^T [\mathbf{w}_1^< ; \mathbf{w}_1^> ; \dots ; \mathbf{w}_{N_{BS}}^< ; \mathbf{w}_{N_{BS}}^>] \geq n_{ik} \tan \gamma_{ik} \} \\ & \Pr \{ \frac{\mathbf{a}_{ik}^T [\mathbf{w}_1^< ; \mathbf{w}_1^> ; \dots ; \mathbf{w}_{N_{BS}}^< ; \mathbf{w}_{N_{BS}}^>]}{\sqrt{\sum_{j=1}^{N_{BS}} \frac{1}{2} [\mathbf{w}_j^< ; \mathbf{w}_j^>]^T \Sigma_{jik} [\mathbf{w}_j^< ; \mathbf{w}_j^>]}} \geq \frac{n_{ik} \tan \gamma_{ik}}{\sqrt{\sum_{j=1}^{N_{BS}} \frac{1}{2} [\mathbf{w}_j^< ; \mathbf{w}_j^>]^T \Sigma_{jik} [\mathbf{w}_j^< ; \mathbf{w}_j^>]}} \} \\ & \Pr \{ \frac{\mathbf{a}_{ik}^T [\mathbf{w}_1^< ; \mathbf{w}_1^> ; \dots ; \mathbf{w}_{N_{BS}}^< ; \mathbf{w}_{N_{BS}}^>]}{\sqrt{\sum_{j=1}^{N_{BS}} \frac{1}{2} [\mathbf{w}_j^< ; \mathbf{w}_j^>]^T \Sigma_{jik} [\mathbf{w}_j^< ; \mathbf{w}_j^>]}} \geq \frac{n_{ik} \tan \gamma_{ik}}{\sqrt{\sum_{j=1}^{N_{BS}} \frac{1}{2} [\mathbf{w}_j^< ; \mathbf{w}_j^>]^T \Sigma_{jik} [\mathbf{w}_j^< ; \mathbf{w}_j^>]}} \} \end{aligned} \quad (11)$$

where $(x) = \frac{1}{\sqrt{2\pi}} \int_1^x e^{-\frac{t^2}{2}} dt$ denotes the cumulative probability function (cdf) of a standard Gaussian distributed variable. Defining $\Phi^{-1}(\cdot)$ as the inverse function of $\Phi(\cdot)$, (11) can be finally derived into a quadratic constraint such that

$$\mathbf{a}_{ik}^T [\mathbf{w}_1^< ; \mathbf{w}_1^> ; \dots ; \mathbf{w}_{N_{BS}}^< ; \mathbf{w}_{N_{BS}}^>] + \Phi^{-1}(\delta) \sqrt{\sum_{j=1}^{N_{BS}} \frac{1}{2} [\mathbf{w}_j^< ; \mathbf{w}_j^>]^T \Sigma_{jik} [\mathbf{w}_j^< ; \mathbf{w}_j^>]} \geq n_{ik} \tan \gamma_{ik} \quad (12)$$

$$\left(\begin{array}{l} \Pr f_1^P \sum_{j=1}^{N_{BS}} (\mathbf{h}_{jik}^<)^T \mathbf{w}_j^< + (\mathbf{h}_{jik}^>)^T \mathbf{w}_j^> \\ \Pr f_2^P \sum_{j=1}^{N_{BS}} (\mathbf{h}_{jik}^<)^T \mathbf{w}_j^< + (\mathbf{h}_{jik}^>)^T \mathbf{w}_j^> \end{array} \right) \begin{array}{l} \sum_{j=1}^{N_{BS}} (\mathbf{h}_{jik}^<)^T \mathbf{w}_j^< \quad (\mathbf{h}_{jik}^>)^T \mathbf{w}_j^> \\ \sum_{j=1}^{N_{BS}} (\mathbf{h}_{jik}^<)^T \mathbf{w}_j^< \quad (\mathbf{h}_{jik}^>)^T \mathbf{w}_j^> \end{array} \begin{array}{l} \rho_{ik} \\ \rho_{ik} \end{array} \tan j e_{ik} g_{ik}; \quad (9)$$

$$\begin{array}{l} \begin{array}{l} \text{8} \\ \text{4} \end{array} \frac{\bar{\mathbf{a}}_{ik}^T [\mathbf{w}_1^<; \mathbf{w}_1^>; \dots; \mathbf{w}_{N_{BS}}^<; \mathbf{w}_{N_{BS}}^>] n^{\rho_{ik} \tan j}}{(\Theta_{ik}^{\frac{1}{2}} [\mathbf{w}_1^<; \mathbf{w}_1^>; \dots; \mathbf{w}_{N_{BS}}^<; \mathbf{w}_{N_{BS}}^>])^T} \begin{array}{l} \Theta_{ik}^{\frac{1}{2}} [\mathbf{w}_1^<; \mathbf{w}_1^>; \dots; \mathbf{w}_{N_{BS}}^<; \mathbf{w}_{N_{BS}}^>] \\ \bar{\mathbf{a}}_{ik}^T [\mathbf{w}_1^<; \mathbf{w}_1^>; \dots; \mathbf{w}_{N_{BS}}^<; \mathbf{w}_{N_{BS}}^>] n^{\rho_{ik} \tan j} \end{array} \begin{array}{l} 3 \\ 5 \\ 0; \end{array} \\ \begin{array}{l} \text{2} \\ \text{4} \end{array} \frac{\bar{\mathbf{b}}_{ik}^T [\mathbf{w}_1^<; \mathbf{w}_1^>; \dots; \mathbf{w}_{N_{BS}}^<; \mathbf{w}_{N_{BS}}^>] n^{\rho_{ik} \tan j}}{(\Theta_{ik}^{\frac{1}{2}} [\mathbf{w}_1^<; \mathbf{w}_1^>; \dots; \mathbf{w}_{N_{BS}}^<; \mathbf{w}_{N_{BS}}^>])^T} \begin{array}{l} \Theta_{ik}^{\frac{1}{2}} [\mathbf{w}_1^<; \mathbf{w}_1^>; \dots; \mathbf{w}_{N_{BS}}^<; \mathbf{w}_{N_{BS}}^>] \\ \bar{\mathbf{b}}_{ik}^T [\mathbf{w}_1^<; \mathbf{w}_1^>; \dots; \mathbf{w}_{N_{BS}}^<; \mathbf{w}_{N_{BS}}^>] n^{\rho_{ik} \tan j} \end{array} \begin{array}{l} 3 \\ 5 \\ 0 \end{array} \end{array} \quad (14)$$

Similarly, the second inequality in (9) can be given as

$$\bar{\mathbf{b}}_{ik}^T [\mathbf{w}_1^<; \mathbf{w}_1^>; \dots; \mathbf{w}_{N_{BS}}^<; \mathbf{w}_{N_{BS}}^>] + \frac{1}{(1 + \tan j)^2} \frac{1}{\Theta_{ik}^{\frac{1}{2}} [\mathbf{w}_1^<; \mathbf{w}_1^>; \dots; \mathbf{w}_{N_{BS}}^<; \mathbf{w}_{N_{BS}}^>]} \frac{1}{n^{\rho_{ik} \tan j}} \quad (13)$$

where $\bar{\mathbf{b}}_{ik} = \text{vec}(\bar{\mathbf{b}}_{1ik}; \bar{\mathbf{b}}_{2ik}; \dots; \bar{\mathbf{b}}_{N_{BS}ik})$ and $\bar{\mathbf{b}}_{jik} = [\hat{\mathbf{h}}_{jik}^> \hat{\mathbf{h}}_{jik}^< \tan; \hat{\mathbf{h}}_{jik}^< + \hat{\mathbf{h}}_{jik}^> \tan]$ with covariance matrix calculated as $\Theta_{ik} = \text{diag}(\underbrace{(1 + \tan j)^2}_{2M} \underbrace{\sigma_{1ik}^2, \dots, \sigma_{N_{BS}ik}^2}_{2N_{BS}})$.

Now, constraint (C2) has been transformed into the two inequalities in (12) and (13). According to Schur Complements that $\frac{1}{\mathbf{e}^T \mathbf{x} + d} \mathbf{A} \mathbf{x} + \mathbf{b} \geq \mathbf{0}$ [40], the above two inequalities in (12) and (13) can be further transformed into two linear matrix inequalities (LMI) as shown in (14). Now we define $\mathbf{W}_j = \mathbf{w}_j \mathbf{w}_j^H; \delta_j \geq 2N_{BS}$. P1 can be finally written as

$$\begin{array}{l} \text{P2 (Full CSI Prob)} : \underset{\mathbf{W}_j; j \geq 2N_{BS}}{\text{argmin}} \text{Tr}(\mathbf{W}_j); \\ \text{s.t (C1)} : \text{Tr}(\mathbf{W}_j) \leq \rho_{max}; \delta_j \geq 2N_{BS}; \\ \text{(C2)} : (14); \delta U_{ik}; \text{(C3)} : \begin{array}{l} \mathbf{W}_j \quad \mathbf{W}_j \\ \mathbf{W}_j^T \quad 1 \end{array} \leq 0; \delta_j \geq 2N_{BS}; \end{array} \quad (15)$$

which is a convex semi-definite programming (SDP) and can be readily solved by CVX solver.

B. Deterministic Robust Optimization

In the previous subsection, we have solved the problem in a probabilistic manner, where the users' QoS requirements are issued by the chance constrained formulations. In this subsection, we handle the CSI uncertainties in a deterministic manner, where the users' QoS requirements are satisfied all the time with the infinite CSI uncertainties. The optimization problem is first formulated in a deterministic manner by P3. Then aided by Lemmas 1 and 2, the problem is finally given by P4.

Define \mathcal{C}_{ik} as the CSI uncertainties set, which contains all the possible CSI uncertainties and specifies an ellipsoidal uncertainty region for the estimated CSI [36]-[39].

1) *Problem Formulation:* To process the power minimization problem in terms of deterministic robust optimization, the formulation is given as

$$\begin{array}{l} \text{P3 (Full CSI Det)} : \underset{\mathbf{w}_j; j \geq 2N_{BS}}{\text{argmin}} \sum_{j=1}^{N_{BS}} k \mathbf{w}_j k^2; \\ \text{s.t (C4)} : k \mathbf{w}_j k^2 \leq \rho_{max}; \delta_j \geq 2N_{BS}; \\ \text{(C5)} : \min_{\mathbf{e}_{ik} \in \mathcal{C}_{ik}} \frac{1}{\rho_{ik}} \leq \delta U_{ik}; \end{array} \quad (16)$$

where (C5) indicates the deterministic SINR requirements for the users, such that the users' worst-case SINR as per the CSI error distribution obey the respective thresholds ρ_{ik} .

2) *Optimization Solution:* In line with the analysis in previous section, constraint (C5) is equivalent to the following two inequalities in (17). We focus our attention on the first inequality of (17), which can be further written as

$$\min_{\mathbf{e}_{ik} \in \mathcal{C}_{ik}} \sum_{j=1}^{N_{BS}} [\mathbf{e}_{ik}^> \mathbf{e}_{ik}^< \tan; \mathbf{e}_{ik}^< + \mathbf{e}_{ik}^> \tan]^T [\mathbf{w}_j^<; \mathbf{w}_j^>] + \rho_{ik} \leq 0; \quad (18)$$

where $\rho_{ik} = \sum_{j=1}^{N_{BS}} (\hat{\mathbf{h}}_{jik}^>)^T \mathbf{w}_j^< \rho_{ik} + (\hat{\mathbf{h}}_{jik}^<)^T \mathbf{w}_j^> \tan + (\hat{\mathbf{h}}_{jik}^<)^T \mathbf{w}_j^> + (\hat{\mathbf{h}}_{jik}^>)^T \mathbf{w}_j^< \tan + n \rho_{ik} \tan$. For simplicity, define vector $\mathbf{c}_{ik} = \text{vec}(\mathbf{c}_{1ik}; \mathbf{c}_{2ik}; \dots; \mathbf{c}_{N_{BS}ik})$, where $\mathbf{c}_{jik} = [\mathbf{e}_{ik}^> \mathbf{e}_{ik}^< \tan; \mathbf{e}_{ik}^< + \mathbf{e}_{ik}^> \tan]$. Then Eq. (18) becomes

$$\min_{\mathbf{e}_{ik} \in \mathcal{C}_{ik}} \mathbf{c}_{ik}^T [\mathbf{w}_1^<; \mathbf{w}_1^>; \dots; \mathbf{w}_{N_{BS}}^<; \mathbf{w}_{N_{BS}}^>] + \rho_{ik} \leq 0; \quad (19)$$

To handle the infinite CSI uncertainties in (19), we transform it into a LMI using the following Lemma 1 (S-procedure):

Lemma 1 (S-Procedure [40]): Let a function $f_m(\mathbf{x})$, $m \geq 1, 2g$, be defined as

$$f_m(\mathbf{x}) = \mathbf{x}^H \mathbf{A}_m \mathbf{x} + 2 \mathbf{b}_m^H \mathbf{x} g + c_m \quad (20)$$

where $\mathbf{A}_m \geq \mathbf{H}^N$, $\mathbf{b}_m \geq \mathbf{C}^N$ and $c_m \geq \mathbf{R}$. The implication $f_1(\mathbf{x}) \leq 0 \wedge f_2(\mathbf{x}) \leq 0$ holds if and only if there exists a $\lambda \geq 0$ such that

$$\begin{array}{l} \mathbf{A}_1; \mathbf{b}_1 \quad \mathbf{A}_2; \mathbf{b}_2 \\ \mathbf{b}_1^H; c_1 \quad \mathbf{b}_2^H; c_2 \end{array} \leq 0; \quad (21)$$

$$\begin{aligned}
& < \min_{\mathbf{e}_{ik}^2} \quad \mathbb{P} \sum_{j=1}^{N_{BS}} (\mathbf{h}_{jik}^<)^T \mathbf{w}_j^< + (\mathbf{h}_{jik}^>)^T \mathbf{w}_j^< & \mathbb{P} \sum_{j=1}^{N_{BS}} (\mathbf{h}_{jik}^<)^T \mathbf{w}_j^< & (\mathbf{h}_{jik}^>)^T \mathbf{w}_j^< & \rho_{ik}^> \tan \delta; \\
& : \min_{\mathbf{e}_{ik}^2} \quad \mathbb{P} \sum_{j=1}^{N_{BS}} (\mathbf{h}_{jik}^<)^T \mathbf{w}_j^< + (\mathbf{h}_{jik}^>)^T \mathbf{w}_j^< & \mathbb{P} \sum_{j=1}^{N_{BS}} (\mathbf{h}_{jik}^<)^T \mathbf{w}_j^< & (\mathbf{h}_{jik}^>)^T \mathbf{w}_j^< & \rho_{ik}^> \tan \delta; \tag{17}
\end{aligned}$$

To utilize S-procedure, we need to construct a premise that guarantees (19) hold. By examining the structure of (19), the premise can be evidently constructed as $(\rho_{ik}^> \mathbf{c}_{ik}^T)^T \mathbf{I}_{2M} \mathbf{I}_{N_{BS}} \rho_{ik}^> \mathbf{c}_{ik}^2$, and the value of $\rho_{ik}^>$ will be given by Lemma 2. By S-procedure, to guarantee the implication $(\rho_{ik}^> \mathbf{c}_{ik}^T)^T \mathbf{I}_{2M} \mathbf{I}_{N_{BS}} \rho_{ik}^> \mathbf{c}_{ik}^2 \leq 0$ (19) holds, the following LMI constraint in (22) should hold with $\rho_{ik}^> = 0$

$$\begin{aligned}
& \mathbf{I}_{2M} \mathbf{I}_{N_{BS}}; \quad 0 \\
& \rho_{ik}^> \quad 0; \quad \rho_{ik}^> \\
& \text{diag}([\mathbf{w}_1^<; \mathbf{w}_1^>; \dots; \mathbf{w}_{N_{BS}}^<; \mathbf{w}_{N_{BS}}^>]); \quad 0 \quad \# \\
& \quad \quad \quad 0; \quad \rho_{ik}^> \quad 0; \tag{22}
\end{aligned}$$

by which the first inequality of (17) containing infinite possibilities is transformed into a deterministic LMI. However, variable $\rho_{ik}^>$ is introduced to bound the term $(\rho_{ik}^> \mathbf{c}_{ik}^T)^T \mathbf{I}_{2M} \mathbf{I}_{N_{BS}} \rho_{ik}^> \mathbf{c}_{ik}^2$, which needs to be connected with the known channel estimation error variance σ_{ik}^2 . Hence, we now introduce Lemma 2 as follows.

Lemma 2 (links $\rho_{ik}^>$ to channel estimation error variance σ_{ik}^2): Provided that the element of CSI error follows Gaussian distribution such that $[e_{ik}]_m \sim \mathcal{CN}(0; \frac{\sigma_{ik}^2}{2})$, $\rho_{ik}^> = \frac{1}{1(\cdot)} \frac{MN_{BS}(1 + \tan^2 \delta)}{\sigma_{ik}^2}$ is obtained. $\frac{1}{1(\cdot)}$ physically represents the outage probability of $(\rho_{ik}^> \mathbf{c}_{ik}^T)^T \mathbf{I}_{2M} \mathbf{I}_{N_{BS}} \rho_{ik}^> \mathbf{c}_{ik}^2$ violating its upper bound σ_{ik}^2 , which can be set close to 1, i.e., $\frac{1}{1(\cdot)} = 0.99$ [35]. $\frac{1}{1(\cdot)}$ denotes the inverse function of the cdf of a standard Gaussian distributed variable.

Proof: Please see Appendix A.

Now with the known value of $\rho_{ik}^>$, the deterministic LMI constraint in (22) is solvable. Similarly, the second inequality in (17) can be transformed as

$$\begin{aligned}
& \mathbf{I}_{2N} \mathbf{I}_{N_{BS}}; \quad 0 \\
& \rho_{ik}^> \quad 0; \quad \rho_{ik}^> \\
& \text{diag}([\mathbf{w}_1^<; \mathbf{w}_1^>; \dots; \mathbf{w}_{N_{BS}}^<; \mathbf{w}_{N_{BS}}^>]); \quad 0 \quad \# \\
& \quad \quad \quad 0; \quad g_{ik} \quad 0; \delta j \geq 2 N_{BS}; \tag{23}
\end{aligned}$$

where $g_{ik} = \mathbb{P} \sum_{j=1}^{N_{BS}} (\hat{\mathbf{h}}_{jik}^>)^T \mathbf{w}_j^< \rho_{ik}^> (\hat{\mathbf{h}}_{jik}^<)^T \mathbf{w}_j^< \tan \delta + (\hat{\mathbf{h}}_{jik}^<)^T \mathbf{w}_j^> + (\hat{\mathbf{h}}_{jik}^>)^T \mathbf{w}_j^> \tan \delta + \rho_{ik}^> \tan \delta$. Now the constraint (C5) containing infinite possibilities is transformed into two deterministic LMI inequalities in (22) and (23), respectively. Defining $\mathbf{W}_j = \mathbf{w}_j \mathbf{w}_j^H$, P3 is finally written as

$$\begin{aligned}
& \text{P4 (Full CI Det)}: \underset{\mathbf{W}_j; \delta j \geq 2 N_{BS}}{\text{argmin}} \quad \text{Tr}(\mathbf{W}_j); \\
& \text{s.t (C7)}: \text{Tr}(\mathbf{W}_j) \leq \rho_{max}; \delta j \geq 2 N_{BS}; \\
& \text{(C8)}: (22) \text{ and } (23); \text{(C9)}: \rho_{ik}^> \leq 0; \\
& \text{(C10)}: \rho_{ik}^> \leq 0; \delta U_{ik}; \\
& \text{(C11)}: \mathbf{W}_j \succeq \mathbf{0}; \delta j \geq 2 N_{BS}; \tag{24}
\end{aligned}$$

which is ready to solve as a standard SDP problem.

IV. PARTIALLY-COORDINATED CI PRECODING WITH IMPERFECT CSI

In the previous section, both the intended data and CSI need to be shared among the coordinated BSs to exploit the constructive inter-cell and multi-user interference, which generally requires that the coordinated BSs are connected with high-capacity and delay-free links. In this section, to reduce the coordination overhead, we propose a scheme that the coordinated BSs only need to share CSI with others yet utilizing multi-user interference as CI, which is referred as Partial-CI scheme. Since only multi-user interference is utilized while inter-cell interference is considered as a harmful element, the received signal at user U_{ik} is written as

$$y_{ik} = \mathbf{h}_{iik}^T \sum_{n=1}^{N_{BS}} \mathbf{w}_{in} s_{in} + \sum_{j \neq i}^{N_{BS}} \mathbf{h}_{jik}^T \sum_{m=1}^{N_{BS}} \mathbf{w}_{jm} s_{jm} + n_{ik}; \tag{25}$$

Since we have $s_{in} = e^{j(\theta_{in} - \theta_{ik})} s_{ik}$ and $s_{jm} = e^{j(\theta_{jm} - \theta_{jk})} s_{jk}$, (25) can be expressed as

$$\begin{aligned}
& y_{ik} = \mathbf{h}_{iik} e^{j(\theta_{i1} - \theta_{ik})} \sum_{n=1}^{N_{BS}} \mathbf{w}_{in} e^{j(\theta_{in} - \theta_{i1})} s_{ik} + \\
& \sum_{j \neq i}^{N_{BS}} \mathbf{h}_{jik} e^{j(\theta_{j1} - \theta_{jk})} \sum_{m=1}^{N_{BS}} \mathbf{w}_{jm} e^{j(\theta_{jm} - \theta_{j1})} s_{jk} + n_{ik}; \tag{26}
\end{aligned}$$

Define $\mathbf{h}_{iik} = \mathbf{h}_{iik} e^{j(\theta_{i1} - \theta_{ik})}$, $\mathbf{h}_{jik} = \mathbf{h}_{jik} e^{j(\theta_{j1} - \theta_{jk})}$, $\mathbf{w}_i = \sum_{n=1}^{N_{BS}} \mathbf{w}_{in} e^{j(\theta_{in} - \theta_{i1})}$, and $\mathbf{w}_j = \sum_{m=1}^{N_{BS}} \mathbf{w}_{jm} e^{j(\theta_{jm} - \theta_{j1})}$. Eq. (26) can be written as

$$y_{ik} = \mathbf{h}_{iik}^T \mathbf{w}_i s_{ik} + \sum_{j \neq i}^{N_{BS}} \mathbf{h}_{jik}^T \mathbf{w}_j s_{jk} + n_{ik}; \tag{27}$$

Hence, the SINR of the Partial-CI scheme can be written as

$$\text{Partial CI}_{ik} = \frac{\mathbf{h}_{iik}^T \mathbf{w}_i^2}{\mathbb{P} \sum_{j \neq i}^{N_{BS}} \mathbf{h}_{jik}^T \mathbf{w}_j^2 + \frac{\sigma_{ik}^2}{\rho_{ik}^>}}; \tag{28}$$

To minimize power consumption in the scenario, we formulate and solve the problems in terms of probabilistic and deterministic manners, respectively.

A. Probabilistic Robust Optimization

In this subsection, we investigate the probabilistic manner optimization for the partially-coordinated CI precoding scheme. The problem formulation is first given by P5. Aided by Lemma 3, the problem is finally given by P6.

$$\begin{aligned}
& \Pr f(\mathbf{h}_{iik}^{\leq})^T \mathbf{w}_i^{\leq} + (\mathbf{h}_{iik}^{\leq})^T \mathbf{w}_i^{\leq} \quad (\mathbf{h}_{iik}^{\leq})^T \mathbf{w}_i^{\leq} \quad (\mathbf{h}_{iik}^{\leq})^T \mathbf{w}_i^{\leq} \\
& : \Pr f(\mathbf{h}_{iik}^{\leq})^T \mathbf{w}_i^{\leq} \quad (\mathbf{h}_{iik}^{\leq})^T \mathbf{w}_i^{\leq} \quad (\mathbf{h}_{iik}^{\leq})^T \mathbf{w}_i^{\leq} \quad (\mathbf{h}_{iik}^{\leq})^T \mathbf{w}_i^{\leq}
\end{aligned} \quad (30)$$

$$\begin{aligned}
& (C12a) : \Pr f(\hat{\mathbf{h}}_{iik}^{\leq} + \mathbf{e}_{ik}^{\leq} \hat{\mathbf{h}}_{iik}^{\leq} \tan \quad \mathbf{e}_{ik}^{\leq} \tan ; \hat{\mathbf{h}}_{iik}^{\leq} + \mathbf{e}_{ik}^{\leq} \hat{\mathbf{h}}_{iik}^{\leq} \tan + \mathbf{e}_{ik}^{\leq} \tan]^T [\mathbf{w}_i^{\leq}; \mathbf{w}_i^{\leq}] \\
& (C12b) : \Pr f(\hat{\mathbf{h}}_{iik}^{\leq} \quad \mathbf{e}_{ik}^{\leq} \hat{\mathbf{h}}_{iik}^{\leq} \tan \quad \mathbf{e}_{ik}^{\leq} \tan ; \hat{\mathbf{h}}_{iik}^{\leq} \quad \mathbf{e}_{ik}^{\leq} \hat{\mathbf{h}}_{iik}^{\leq} \tan + \mathbf{e}_{ik}^{\leq} \tan]^T [\mathbf{w}_i^{\leq}; \mathbf{w}_i^{\leq}] \\
& (C12c) : u_{ik}; \quad (C12d) : u_{ik}
\end{aligned} \quad (31)$$

$$(C12a) : \frac{2}{4} \frac{\bar{\mathbf{F}}_{ik}^T [\mathbf{w}_i^{\leq}; \mathbf{w}_i^{\leq}] \quad (n + u_{ik}) \rho_{ik}^{\leq} \tan \quad I}{\left(\frac{1}{2} [\mathbf{w}_i^{\leq}; \mathbf{w}_i^{\leq}] \right)^T} \quad \frac{3}{5} \quad 0; \quad (33)$$

$$(C12b) : \frac{2}{4} \frac{\bar{\mathbf{d}}_{ik}^T [\mathbf{w}_i^{\leq}; \mathbf{w}_i^{\leq}] \quad (n + u_{ik}) \rho_{ik}^{\leq} \tan \quad I}{\left(\frac{1}{2} [\mathbf{w}_i^{\leq}; \mathbf{w}_i^{\leq}] \right)^T} \quad \frac{3}{5} \quad 0; \quad (34)$$

1) *Problem Formulation:* The problem is formulated as

$$\begin{aligned}
& P5 \text{ (Partial CI Prob)} : \operatorname{argmin}_{\mathbf{w}_j; \delta_j \geq 2N_{BS}} \sum_{j=1}^{N_{BS}} \|\mathbf{w}_j\|^2; \\
& \text{s.t (C11)} : \|\mathbf{w}_j\|^2 \leq P_{max}; \delta_j \geq 2N_{BS}; \\
& \text{(C12)} : \Pr f_{ik} \quad \mathbf{e}_{ik} \mathbf{g}_{ik} \quad u_{ik}; \delta u_{ik};
\end{aligned} \quad (29)$$

2) *Optimization Solution:* Since inter-cell interference is treated as undesired element, the probabilistic constraint (C12) equals to the following two inequalities in (30). To decouple the inter-cell interference at each BS, we introduce two auxiliary variables u_{ik} and U_{ik} . Then, (30) is transformed into (31). Now we handle the constraints in (31) one by one. For constraint (C12a), it can be equivalently written in the form as

$$(C12a) : \bar{\mathbf{F}}_{ik}^T [\mathbf{w}_i^{\leq}; \mathbf{w}_i^{\leq}] + \frac{1}{2} (ik) \sum_{j=1}^{2M} [\mathbf{w}_i^{\leq}; \mathbf{w}_i^{\leq}]^T \quad (32)$$

where $\bar{\mathbf{F}}_{ik} = [\hat{\mathbf{h}}_{iik}^{\leq} \quad \hat{\mathbf{h}}_{iik}^{\leq} \tan ; \hat{\mathbf{h}}_{iik}^{\leq} + \hat{\mathbf{h}}_{iik}^{\leq} \tan]$ and $\rho_{ik}^{\leq} = \operatorname{diag}((1 + \tan^2 \frac{2}{ik}), \dots, (1 + \tan^2 \frac{2}{ik}))$. According to Schur

complement, constraint (C12a) can be given in (33). Similarly, constraint (C12b) can be directly given in (34), where $\bar{\mathbf{d}}_{ik} = [\hat{\mathbf{h}}_{iik}^{\leq} \quad \hat{\mathbf{h}}_{iik}^{\leq} \tan ; \hat{\mathbf{h}}_{iik}^{\leq} + \hat{\mathbf{h}}_{iik}^{\leq} \tan]$. Now we turn to handle the constraint (C12c). By stacking the variables u_{ik} and U_{ik} into a long vector that $\mathbf{t}_{ik} = [u_{ik}; U_{ik}]^T$ and defining $\mathbf{S}_{ik} = \mathbf{t}_{ik} \mathbf{t}_{ik}^H$, (C12c) can be relaxed in the form as

$$(C12c) : \operatorname{Tr}(\mathbf{S}_{ik}) + \mathbf{I} \mathbf{t}_{ik} \leq 0; \text{ and } \begin{matrix} \mathbf{S}_{ik}; & \mathbf{t}_{ik} \\ (\mathbf{t}_{ik})^T; & 1 \end{matrix} \leq 0; \quad (35)$$

where $\mathbf{I} = \begin{bmatrix} 1; & 0 \\ 0; & 0 \end{bmatrix}$ and vector $\mathbf{I} = [0; 1]$. Next we can handle constraint (C12d), which can be further written as

$$\begin{aligned}
& (C12d) : \sum_{j \in i}^{N_{BS}} \mathbf{e}_{ik}^T \mathbf{W}_j \mathbf{e}_{ik} + \sum_{j \in i}^{N_{BS}} \mathbf{e}_{ik}^T \mathbf{W}_j \hat{\mathbf{h}}_{j ik} + \\
& \sum_{j \in i}^{N_{BS}} \hat{\mathbf{h}}_{j ik}^T \mathbf{W}_j \mathbf{e}_{ik} + \sum_{j \in i}^{N_{BS}} \hat{\mathbf{h}}_{j ik}^T \mathbf{W}_j \hat{\mathbf{h}}_{j ik} \leq u_{ik} \quad 0;
\end{aligned} \quad (36)$$

To handle the inequality with S-procedure, we need to construct another premise $\mathbf{e}_{ik}^T \mathbf{I}_M \mathbf{e}_{ik} \leq 2$ to guarantee (36) always hold. We now calculate the value of 2 according to Lemma 3.

Lemma 3 (links 2 to CSI error variance $\frac{2}{ik}$): Provided the element of CSI error follows Gaussian distribution such that $[\mathbf{e}_{ik}]_m \sim \mathcal{CN}(0; \frac{2}{ik})$, the term $\mathbf{e}_{ik}^T \mathbf{I}_M \mathbf{e}_{ik}$ is bounded by $2 = \frac{1}{\Gamma(\cdot)} \frac{2}{ik}$. Explicitly, physically represents the outage probability of $\mathbf{e}_{ik}^T \mathbf{I}_M \mathbf{e}_{ik}$ violating its upper bound, and can be set to approach 1, i.e. $\Gamma(\cdot) = 0.99$ [35]. $\Gamma(\cdot)$ is the inverse function of the cdf of a chi-square distributed variable with M degree of freedom.

Proof: Please see Appendix B.

Now with the known value of 2 , the implication $\mathbf{e}_{ik}^T \mathbf{I}_M \mathbf{e}_{ik} \leq 2 \quad 0$ (36) holds if and only if the LMI in (37) holds. Now, the original constraint (C12) is transformed into equivalent constraints (C12a)-(C12d), as shown by Eqs. (33), (34), (35) and (37). The optimization problem becomes

$$\begin{aligned}
& P6 \text{ (Partial CI Prob)} : \operatorname{argmin}_{\mathbf{w}_j; \delta_j \geq 2N_{BS}} \operatorname{Tr}(\mathbf{W}_j); \\
& \text{s.t (C11)} : \operatorname{Tr}(\mathbf{W}_j) \leq P_{max}; \delta_j \geq 2N_{BS}; \text{ (C12a)} : (33); \\
& \text{(C12b)} : (34); \text{ (C12c)} : (35); \text{ (C12d)} : (37); \delta u_{ik}; \\
& \text{(C13)} : \begin{matrix} \mathbf{W}_j & \mathbf{w}_j \\ \mathbf{w}_j^T & 1 \end{matrix} \leq 0; \delta_j \geq 2N_{BS};
\end{aligned} \quad (38)$$

which is readily solved by CVX.

$$(C12d) : \begin{matrix} \text{"} \\ \text{P} \\ \text{P} \end{matrix} \begin{matrix} I_M \\ \text{P} \\ \text{P} \end{matrix} \begin{matrix} N_{BS} \\ N_{BS} \\ N_{BS} \end{matrix} \begin{matrix} \mathbf{W}_j \\ \mathbf{h}_{jik}^T \mathbf{W}_j \\ \mathbf{h}_{jik}^T \mathbf{W}_j \mathbf{h}_{jik} \end{matrix} \begin{matrix} \# \\ \\ \end{matrix} \quad \text{0; and } i_k \text{ 0; } \quad (37)$$

B. Deterministic Optimization

In this subsection, we now investigate the deterministic manner optimization for the partially-coordinated CI precoding scheme. The problem formulation is first given by P7. After a series of transformations, the optimization problem is finally given by P8. Accordingly, the problem formulation is given as

$$\begin{aligned} P7 \text{ (Partial CI Det)} : & \underset{\mathbf{w}_j; \delta_j \geq 2N_{BS}}{\operatorname{argmin}} \sum_{j=1}^{N_{BS}} k w_j k^2; \\ \text{s.t (C14)} : & k w_j k^2 \leq \rho_{max}; \delta_j \geq 2N_{BS}; \\ \text{(C15)} : & \min_{e_{ik}^2} \sum_{ik} \frac{e_{ik}^2}{e_{ik}^2} \delta U_{ik}; \end{aligned} \quad (39)$$

To make the optimization problem solvable, we need to process constraint (C15) that contains infinite possibilities. Similarly, we introduce two U_{ik} and e_{ik} to decouple the inter-cell interference. Hence, (C15) equals to

$$\begin{aligned} \text{(C15a)} : & \min_{e_{ik}^2} [\hat{\mathbf{h}}_{jik}^T + \mathbf{e}_{ik}^T \hat{\mathbf{h}}_{jik}^T \tan^{-1} \frac{e_{ik}^2}{e_{ik}^2}; \hat{\mathbf{h}}_{jik}^T + \mathbf{e}_{ik}^T \\ & + \hat{\mathbf{h}}_{jik}^T \tan^{-1} \frac{e_{ik}^2}{e_{ik}^2} + \mathbf{e}_{ik}^T \tan^{-1} \frac{e_{ik}^2}{e_{ik}^2}]^T [\mathbf{w}_j^T; \mathbf{w}_j^T] \leq \rho_{max} (n + e_{ik}^2) \tan^{-1} \frac{e_{ik}^2}{e_{ik}^2}; \\ \text{(C15b)} : & \min_{e_{ik}^2} [\hat{\mathbf{h}}_{jik}^T + \mathbf{e}_{ik}^T \hat{\mathbf{h}}_{jik}^T \tan^{-1} \frac{e_{ik}^2}{e_{ik}^2}; \hat{\mathbf{h}}_{jik}^T + \mathbf{e}_{ik}^T \\ & + \hat{\mathbf{h}}_{jik}^T \tan^{-1} \frac{e_{ik}^2}{e_{ik}^2} + \mathbf{e}_{ik}^T \tan^{-1} \frac{e_{ik}^2}{e_{ik}^2}]^T [\mathbf{w}_j^T; \mathbf{w}_j^T] \leq \rho_{max} (n + e_{ik}^2) \tan^{-1} \frac{e_{ik}^2}{e_{ik}^2}; \\ \text{(C15c)} : & e_{ik}^2 \leq U_{ik}; \text{(C15d)} : U_{ik} \leq \sum_{j=1}^{N_{BS}} \mathbf{h}_{jik}^T \mathbf{W}_j \mathbf{W}_j^T \mathbf{h}_{jik}; \end{aligned} \quad (40)$$

Firstly, constraint (C15a) can be written as

$$(C15a) : \min_{e_{ik}^2} [\mathbf{e}_{ik}^T \quad \mathbf{e}_{ik}^T \tan^{-1} \frac{e_{ik}^2}{e_{ik}^2}; \mathbf{e}_{ik}^T + \mathbf{e}_{ik}^T \tan^{-1} \frac{e_{ik}^2}{e_{ik}^2}]^T [\mathbf{w}_j^T; \mathbf{w}_j^T] \leq \rho_{max} (n + e_{ik}^2) \tan^{-1} \frac{e_{ik}^2}{e_{ik}^2}; \quad (41)$$

where $\rho_{max} = (\hat{\mathbf{h}}_{jik}^T)^T \mathbf{w}_j^T \tan^{-1} \frac{e_{ik}^2}{e_{ik}^2} + (\hat{\mathbf{h}}_{jik}^T)^T \mathbf{w}_j^T + (\hat{\mathbf{h}}_{jik}^T)^T \mathbf{w}_j^T \tan^{-1} \frac{e_{ik}^2}{e_{ik}^2} + (n + e_{ik}^2) \tan^{-1} \frac{e_{ik}^2}{e_{ik}^2}$. By applying S-procedure, (41) is equivalent to the following LMI

$$(C15a) : \begin{matrix} \text{"} \\ \text{P} \\ \text{P} \end{matrix} \begin{matrix} I_{2N} \\ \text{P} \\ \text{P} \end{matrix} \begin{matrix} \text{diag}(\mathbf{w}_j^T; \mathbf{w}_j^T); \\ \mathbf{0}; \\ \delta_{ik}^2 \quad \rho_{max} \end{matrix} \begin{matrix} \mathbf{0} \\ \mathbf{0}; \\ \rho_{max} \end{matrix} \begin{matrix} \# \\ \\ \end{matrix} \quad (42)$$

and $\delta_{ik} \geq 0$;

where $\rho_{max} = \frac{1}{M} (1 + \tan^2 \frac{e_{ik}^2}{e_{ik}^2})$ similarly calculated by Lemma 2. Also, constraint (C15b) can be written as

$$(C15b) : \begin{matrix} \text{"} \\ \text{P} \\ \text{P} \end{matrix} \begin{matrix} I_{2N} \\ \text{P} \\ \text{P} \end{matrix} \begin{matrix} \text{diag}(\mathbf{w}_j^T; \mathbf{w}_j^T); \\ \mathbf{0}; \\ \rho_{max} \end{matrix} \begin{matrix} \mathbf{0} \\ \mathbf{0}; \\ \rho_{max} \end{matrix} \begin{matrix} \# \\ \\ \end{matrix} \quad (43)$$

and $\rho_{max} \geq 0$;

where $\rho_{max} = [\hat{\mathbf{h}}_{jik}^T \mathbf{w}_j^T \tan^{-1} \frac{e_{ik}^2}{e_{ik}^2} + \hat{\mathbf{h}}_{jik}^T \mathbf{w}_j^T + \hat{\mathbf{h}}_{jik}^T \mathbf{w}_j^T \tan^{-1} \frac{e_{ik}^2}{e_{ik}^2} + (n + e_{ik}^2) \tan^{-1} \frac{e_{ik}^2}{e_{ik}^2}]$. On the other hand, since (C15c) and (C15d) have similar structures with (C12c) and (C12d), (C15c) can be written in the forms as

$$(C15c) : \operatorname{Tr}(\mathbf{S}_{ik}) + \mathbf{t}_{ik}^T \mathbf{0}; \text{ and } \begin{matrix} \mathbf{S}_{ik}; \\ (\mathbf{t}_{ik})^T; \end{matrix} \begin{matrix} \mathbf{t}_{ik} \\ 1 \end{matrix} \mathbf{0}; \quad (44)$$

while (C15d) can be given in (45). Now, the problem is transformed into

$$\begin{aligned} P8 \text{ (Partial CI Det)} : & \underset{\mathbf{w}_j; \delta_j \geq 2N_{BS}}{\operatorname{argmin}} \sum_{j=1}^{N_{BS}} k w_j k^2; \\ \text{s.t (C14)} : & \operatorname{Tr}(\mathbf{W}_j) \leq \rho_{max}; \delta_j \geq 2N_{BS}; \text{(C15a)} : (42); \\ \text{(C15b)} : & (43); \text{(C15c)} : (44); \text{(C15d)} : (45); \delta U_{ik}; \\ \text{(C16)} : & \begin{matrix} \mathbf{W}_j & \mathbf{w}_j \\ \mathbf{w}_j^T & 1 \end{matrix} \leq \mathbf{0}; \delta_j \geq 2N_{BS}; \end{aligned} \quad (46)$$

which is readily solved by CVX.

V. STATISTICALLY-COORDINATED CI PRECODING

In section IV, instantaneous CSI needs to be shared among the BSs to facilitate the optimization. In this section, to further reduce the coordination overhead, we propose a CI precoding scheme that BSs need to know its local estimated CSI while only sharing statistical inter-cell channel with others. Hence, the coordination overhead can be significantly reduced compare to the above instantaneous CSI based schemes. The problem formulation is first given by P9, and the transformed problem is finally presented by P10.

Let us define \mathbf{R}_{jik} as the correlation matrix of the j -th BS to the user U_{ik} . Assuming white channel covariance [41], the true channel correlation matrix can be written as $\mathbf{R}_{jik} = \mathbf{u}_{jik}^H \mathbf{u}_{jik} + \frac{1}{2} \mathbf{I}_M$; $\delta_j \notin i; j \geq 2N_{BS}$, where \mathbf{u}_{jik} denotes the mean of channel \mathbf{h}_{jik} . Since only multi-user interference can be exploited as constructive elements while inter-cell interference is suppressed as noise, the SINR expression is same to that of the partially-coordinated precoding scheme, as presented by Eq. (28). Accordingly, the problem is formulated as

$$\begin{aligned} P9 \text{ (Stat CI)} : & \underset{\mathbf{w}_j; \delta_j \geq 2N_{BS}}{\operatorname{argmin}} \sum_{j=1}^{N_{BS}} k w_j k^2; \\ \text{s.t (C20)} : & k w_j k^2 \leq \rho_{max}; \delta_j \geq 2N_{BS}; \\ \text{(C21)} : & \min_{e_{ik}^2} \sum_{ik} \frac{e_{ik}^2}{e_{ik}^2} \delta U_{ik}; \end{aligned} \quad (47)$$

We now need to handle constraint (C21), which can be equivalently written as

$$\begin{aligned} \text{(C21c)} : & \min_{e_{ik}^2} [\hat{\mathbf{h}}_{jik}^T + \mathbf{e}_{ik}^T \hat{\mathbf{h}}_{jik}^T \tan^{-1} \frac{e_{ik}^2}{e_{ik}^2}; \hat{\mathbf{h}}_{jik}^T + \mathbf{e}_{ik}^T \\ & + \hat{\mathbf{h}}_{jik}^T \tan^{-1} \frac{e_{ik}^2}{e_{ik}^2} + \mathbf{e}_{ik}^T \tan^{-1} \frac{e_{ik}^2}{e_{ik}^2}]^T [\mathbf{w}_j^T; \mathbf{w}_j^T] \leq \rho_{max} (n + e_{ik}^2) \tan^{-1} \frac{e_{ik}^2}{e_{ik}^2}; \\ \text{(C21d)} : & \min_{e_{ik}^2} [\hat{\mathbf{h}}_{jik}^T + \mathbf{e}_{ik}^T \hat{\mathbf{h}}_{jik}^T \tan^{-1} \frac{e_{ik}^2}{e_{ik}^2}; \hat{\mathbf{h}}_{jik}^T + \mathbf{e}_{ik}^T \\ & + \hat{\mathbf{h}}_{jik}^T \tan^{-1} \frac{e_{ik}^2}{e_{ik}^2} + \mathbf{e}_{ik}^T \tan^{-1} \frac{e_{ik}^2}{e_{ik}^2}]^T [\mathbf{w}_j^T; \mathbf{w}_j^T] \leq \rho_{max} (n + e_{ik}^2) \tan^{-1} \frac{e_{ik}^2}{e_{ik}^2}; \\ \text{(C21e)} : & e_{ik}^2 \leq U_{ik}; \text{(C21d)} : \sum_{j=1}^{N_{BS}} \operatorname{Tr}(\mathbf{w}_j \mathbf{W}_j^H \mathbf{R}_{jik}) \leq U_{ik}; \end{aligned} \quad (48)$$

As can be seen, constraint (C21d) itself is a LMI while (C21a), (C21b) and (C21c) can similarly be solved as presented in Section IV. Hence, the problem can be re-formulated as

$$(C15d): \quad \begin{aligned} & \text{Tr}(\mathbf{W}_j); \\ & \text{Tr}(\mathbf{W}_j \mathbf{h}_{j|ik} \mathbf{W}_j \mathbf{h}_{j|ik}^H); \end{aligned} \quad \text{for } j \in \mathcal{I}, i \in \mathcal{K}; \quad \text{and } u_{ik} \geq 0; \quad \text{and } u_{ik} \geq 0; \quad (45)$$

$$\begin{aligned} P10 \text{ (Stat CI)}: & \quad \underset{\mathbf{W}_j: \delta_j \geq 2N_{BS}}{\text{argmin}} \quad \text{Tr}(\mathbf{W}_j); \\ \text{s.t (C20)}: & \quad \text{Tr}(\mathbf{W}_j) \leq P_{max}; \delta_j \geq 2N_{BS}; \\ (C21a): & \quad (42); (C22b): (43); (C22c): (44); \\ (C21d): & \quad \text{Tr}(\mathbf{w}_j \mathbf{w}_j^H \mathbf{R}_{j|ik}) \leq u_{ik}; \delta u_{ik}; \\ (C22): & \quad \begin{aligned} & \mathbf{W}_j \mathbf{w}_j = \mathbf{0}; \delta_j \geq 2N_{BS}; \\ & \mathbf{w}_j^T \mathbf{1} = 1 \end{aligned} \end{aligned} \quad (49)$$

which can be readily solved by CVX.

In the following Remarks 1 and 2, some insights are revealed for the system design.

Remark 1: As aforementioned, the fully-coordinated CI precoding scheme requires the highest level of overhead, where the transmitted data and explicit CSI need to be shared among the coordinated BSs. Benefiting from this, both the inter-cell interference and multi-user interference can be properly designed as constructive elements. For the partially-coordinated CI precoding scheme, only CSI is shared among the BSs while the transmitted data is only known by the local BSs. Hence, only multi-user interference within each local BS can be utilized while inter-cell interference has to be jointly mitigated. As a result, to achieve a same SINR requirement, the incurred power consumption is higher than the fully-coordinated scheme, as will be demonstrated by Fig. 3. At last, the statistically-coordinated CI precoding scheme only shares the statistical inter-cell CSI among the BSs for jointly suppressing inter-cell interference. Hence, the overhead is the smallest at the cost of high transmission power.

Remark 2: In this remark, we briefly discuss the system design when the users are equipped with multiple antennas. There will be different schemes based on the various possible receiver configurations. (a) If the users are equipped with multiple antennas while no receive combining is applied, this results in a transmit only precoding, and our proposed approach optimizing the transmit precoding applies unchanged. (b) If the receive combining applied at the users is known and fixed, let's denote this as $\mathbf{V} \in \mathbb{C}^{M_r \times 1}$, where M_r denotes the number of receiver applied at the user. We can simply absorb the effect of combining vector into MIMO channel $\mathbf{H} \in \mathbb{C}^{M \times M_r}$ and obtain an equivalent channel from the BSs to the user $\mathbf{H}_e = \mathbf{H}\mathbf{V}$. Then, the proposed optimizations can be applied unchanged by employing the equivalent channel model \mathbf{H}_e . (c) To further improve the system performance, joint precoding and combining design may be conducted. Inspired by the joint precoding/combining design in hybrid beamforming systems, one possible solution is to employ our proposed optimizations to jointly optimize transmit and receive beamforming, by iteratively optimizing the precoding or combining assuming the other being fixed, until convergence is

reached. To be specific, one can first optimize precoding with a fixed receive combining scheme, and then turn to optimize receive beamforming with the obtained precoding scheme. Such an iterative optimization is performed until an optimal point is achieved.

VI. COMPLEXITY AND OVERHEAD ANALYSIS

We present the coordination overhead and complexities for the different algorithms in subsections VI-A and VI-B, respectively. The associated analytical expressions are summarized in Table I.

A. Coordination Overhead

By the Full-CI-Prob/Det algorithms, each BS shares CSI and the transmitted symbols to other coordinated BSs. With K users in each BS and N_{BS} coordinated BSs, the total overhead of sharing CSI and data is given $O(N_{BS}(N_{BS} - 1)K(N_{BS} - c + s))$, where c denotes the required bits for describing one user's CSI while s denotes the required bits for exchanging symbols. Evidently, the Full-CI-Prob/Det algorithms show the same overhead with conventional CoMP systems, which are suited for the scenario where the coordinated BSs are connected with high-capacity and delay-free links. By the Partial-CI-Prob/Det algorithms, only CSI needs to be shared among the BSs. Hence, the overhead is significantly reduced to $O(N_{BS}^2(N_{BS} - 1)K - c)$. By the Stat-CI algorithm, BSs only share statistical CSI of inter-cell channel with others, and the required bits can be ignored compared to other instantaneous CSI acquisition-based schemes.

B. Computational Complexity

Now we analyze the complexities of each algorithm, and the details of analyzing complexity by an interior-point based solver are presented in the footnote ¹.

1) : For the full-CI precoding scheme, P2 handles the optimization in a probabilistic manner. It involves N_{BS} LMI (trace) in (C1) of size 1, $2N_{BS}K$ LMI inequalities in (C2) of size $2N_{BS} + 1$, and $N_{BS}K$ LMI inequalities in (C3) of size $M + 1$. Hence, barrier parameter is given as $\gamma = \frac{1}{N_{BS}(MK + 3K + 1 + 4KN_{BS})}$. On the other hand, P4 handles the problem in a deterministic manner. It involves N_{BS} LMI (trace) in (C7) of size 1, $2N_{BS}K$ LMI inequalities

¹For the interior-point methods based solver, the overall complexity can be given as $\ln(\frac{1}{\epsilon})\sqrt{c_b}(c_f + c_g)$ [42]. Specifically, $\ln(\frac{1}{\epsilon})$ relates to the accuracy setup. $\sqrt{c_b}$ represents the barrier parameter measuring the geometric complexity of the conic constraints. c_f and c_g represent the complexities cost on forming and factorization of $n \times n$ matrix of the optimization problem. A search direction is found by solving a system of linear equations in n unknowns. c_f is calculated as $c_f = n \sum_{j=1}^P k_j^3 + n^2 \sum_{j=1}^P k_j^2 + n \sum_{j=P+1}^m k_j^2$, where k_j presents the size of the j -th constraint. Specifically, the terms $n \sum_{j=1}^P k_j^3 + n^2 \sum_{j=1}^P k_j^2$ come from P LMI constraints while the term $n \sum_{j=P+1}^m k_j^2$ comes from $m - P$ second order cone constraints in the problem formulation. c_g is calculated as $c_g = n^3$ (eq. (18) [42]).

TABLE I. Overhead and complexity analysis, with accuracy factor

Algorithms	Overhead	Complexity
Full-CI-Prob	$\mathcal{O}(N_{BS}^2(N-1)K_c + N_{BS}(N_{BS}-1)K_s)$	$\ln^{(1)}_1 n(N_{BS} + 2N_{BS}K(2N_{BS} + 1)^3 + N_{BS}K(M+1)^3) + n^2(N_{BS} + 2N_{BS}K(2N_{BS} + 1)^2 + N_{BS}K(M+1)^2) + n^3$
Full-CI-Det	$\mathcal{O}(N_{BS}^2(N_{BS}-1)K_c + N_{BS}(N_{BS}-1)K_s)$	$\ln^{(1)}_2 n(N_{BS} + 2N_{BS}K(2N_{BS} + 1)^3 + 2N_{BS}K + N_{BS}K(M+1)^3) + n^2(N_{BS} + 2N_{BS}K(2N_{BS} + 1)^2 + 2N_{BS}K + N_{BS}K(M+1)^2) + n^3$
Partial-CI-Prob	$\mathcal{O}(N_{BS}^2(N_{BS}-1)K_c)$	$\ln^{(1)}_3 n(N_{BS} + 10N_{BS}K + N_{BS}K((N_{BS}-1)M^2 + 1)^3 + 2N_{BS}K(2M+1)^3 + N_{BS}K(M+1)^3) + n^2(N_{BS} + 10N_{BS}K + N_{BS}K((N_{BS}-1)M^2 + 1)^2 + 2N_{BS}K(2M+1)^2 + N_{BS}K(M+1)^2) + n^3$
Partial-CI-Det	$\mathcal{O}(N_{BS}^2(N_{BS}-1)K_c)$	$\ln^{(1)}_4 n(N_{BS} + 31N_{BS}K + 2N_{BS}K(M+1)^3 + 2N_{BS}K(2M+1)^3) + n^2(N_{BS} + 13N_{BS}K + 2N_{BS}K(M+1)^2 + 2N_{BS}K(2M+1)^2) + n^3$
Stat-CI	Negligible	$\ln^{(1)}_5 n(N_{BS} + 31N_{BS}K + 2N_{BS}K(2M+1)^3 + N_{BS}K(M+1)^3) + N_{BS}^2(N_{BS} + 13N_{BS}K + 2N_{BS}K(2M+1)^2 + N_{BS}K(M+1)^2) + n^3$

in (C8) of size $2N_{BS} + 1$, $N_{BS}K$ linear inequality in (C9), $N_{BS}K$ linear inequality in (C10), and $N_{BS}K$ LMI inequalities in (C10) of size $M + 1$. Hence, its barrier parameter is given as $\beta_2 = \frac{1}{N_{BS}(5 + 4N_{BS} + KM + K)}$.

2) : In the partial-CI precoding scheme, P6 first handles the problem in a probabilistic manner. It involves N_{BS} LMI (trace) in (C11) of size 1, $N_{BS}K$ LMI inequalities of size $2M + 1$ in (C12a), $N_{BS}K$ LMI inequalities of size $2M + 1$ in (C12b), $N_{BS}K$ LMI (trace) of size 1 and $N_{BS}K$ LMI inequalities of size 3 in (C12c), $N_{BS}K$ LMI inequalities of size $(N_{BS} - 1)M^2 + 1$ and $N_{BS}K$ linear constraints in (C12d), and $N_{BS}K$ LMI inequalities of size $M + 1$ in (C13). Hence, its barrier parameter is given as $\beta_3 = \frac{1}{N_{BS}(1 + 5K + 2K(3M + 2) + K((N_{BS} - 1)M^2 + 1))}$.

On the other hand, P8 handles the problem in a deterministic manner. It involves N_{BS} LMI (trace) in (C14) of size 1, $N_{BS}K$ LMI inequalities of size $2M + 1$ and $N_{BS}K$ linear constraints in (C15a), $N_{BS}K$ LMI inequalities of size $2M + 1$ and $N_{BS}K$ linear constraints in (C15b), $N_{BS}K$ LMI (trace) in (C15c) of size 1 and $N_{BS}K$ LMI inequalities in (C15c) of size 3, $N_{BS}K$ LMI inequalities of size $M + 1$ and $N_{BS}K$ linear constraints in (C15d), and $N_{BS}K$ LMI inequalities in (C16) of size $M + 1$. Hence, its barrier parameter is given as $\beta_4 = \frac{1}{N_{BS}(2 + M + K(8 + 5M + 2))}$.

3) : Now we check the complexity of the Stat-CI algorithm. (P10) involves N_{BS} LMI (trace) of size 1 in (C20), $N_{BS}K$ inequalities of size $2M + 1$ and $N_{BS}K$ linear constraints in (C21a), $N_{BS}K$ inequalities of size $2M + 1$ and $N_{BS}K$ linear constraints in (C21b), $N_{BS}K$ LMI (trace) of size 1 and $N_{BS}K$ LMI inequalities of size 3 in (C21c), $N_{BS}K$ LMI (trace) of size 1 in (C21d), and N_{BS} LMI inequalities of size $M + 1$ in (C22). Hence, its barrier parameter is given as $\beta_5 = \frac{1}{N_{BS}(2 + 9K + 4KM + M)}$.

VII. SIMULATION RESULTS

Our results employ Monte Carlo simulations of the considered scenarios. The central frequency is set to 2 GHz with 1 MHz bandwidth. The AWGN power spectral density is -174 dBm/Hz. In line with the closely relevant works [35] [36], a 3-cell coordination scenario is considered, and in each cell a BS is located in the center to serve multiple users, as shown in Fig. 2. The number of antennas at each BS is set to $M = 4$. It is assumed there are $K = 3$ users in each cell and are distributed across the map with the exception of the simulation results in Fig. 7, where all the users are placed in the edge

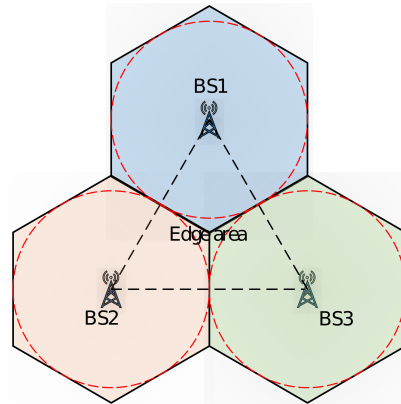


Fig. 2. Illustration of system model, where 3 BSs exchange data or/and CSI for different level of multi-cell coordination.

area to highlight the working modes of the coordinated BSs by the different schemes. QPSK is adopted as constellation scheme and the transmitted data for the multiple users are randomly generated. The maximum transmission power is set to $p_{max} = 40$ dBm. Without loss of generality, the outage probability factor (for the probabilistic manner optimization) and SINR requirement are set to $\bar{\gamma}_{ik} = 80\%$ and $\bar{\gamma}_{ik} = 20$ dB, respectively. The standard deviation of the CSI uncertainty is set to $\sigma_{ik} = 10^{-2}$ for all the users, and hence the channel estimation uncertainty bound for the deterministic manner optimization can be readily calculated by Lemmas 2 and 3. PL model in [35] and Rayleigh fading [39] are adopted for modelling channels. The following schemes are selected as benchmarks. a) CoMP with perfect CSI [2], which serves as a transmit power lower bound of the existing schemes. b) CBF with probabilistic optimization (CBF-Prob) [35], where probabilistic optimization is applied. c) CBF with deterministic optimization (CBF-Det) [36], where deterministic optimization is applied.

Fig. 3 shows the impact of users' SINR requirements $\bar{\gamma}_{ik}$ on the total power consumption. It can be seen that the Full-CI-Prob and Full-CI-Det algorithms consume the lowest power. It is because the Full-CI-Prob/Det utilize both inter-cell and multi-user interference as constructive elements, which help reduce the transmission power to achieve a target SINR. As comparison, although CoMP-Perfect-CSI scheme works as a network-level MISO to obtain a broadcast channel, the multi-user interference still needs to be canceled. As a result,

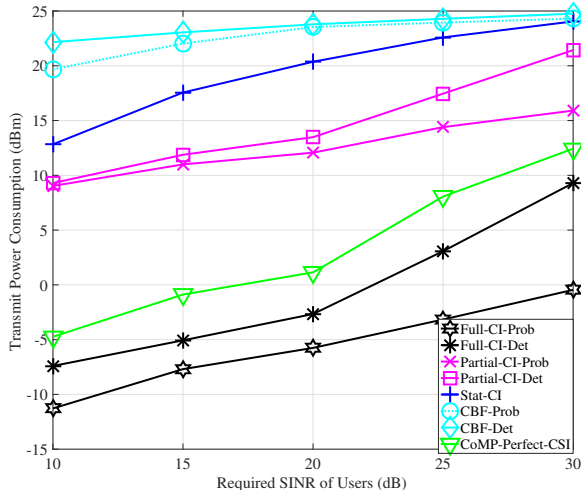


Fig. 3. Impact of the users' SINR requirements $\bar{\gamma}_{ik}$ on the total power consumption.

even with perfect CSI, CoMP-Perfect-CSI is inferior to the proposed Full-CI-Prob/Det algorithms. For the Partial-CI-Prob and Partial-CI-Det algorithms, since the intended transmitted data is not shared among the BSs to reduce the coordination overhead, only multiuser interference is utilized while inter-cell interference is carefully suppressed by joint precoding design. As a result, the Partial-CI-Prob/Det algorithms consume higher power than the Full-CI-Prob/Det algorithms. However, compared to the most relevant CBF benchmarks in [35] and [36], the Partial-CI-Prob/Det algorithms consume much lower power, benefiting from utilizing multi-user interference as beneficial element. At last, the Stat-CI algorithm only needs to share statistical inter-cell channel CSI with its adjacent cells. As a result, the inter-cell interference suppressing is not accurate as that in the instantaneous CSI based Partial-CI-Prob/Det algorithms. Hence, with the lightest complexity, the Stat-CI algorithm demonstrates a slight power consumption increment over the Partial-CI-Prob/Det algorithms. However, it is observed that the Stat-CI algorithm consumes significantly less power compared to the CBF benchmarks in [35] and [36], benefiting from utilizing multi-user interference. Secondly, it is observed that the deterministic robust optimization generally requires more transmission power than the corresponding probabilistic robust optimization. It is because the deterministic robust optimization needs to satisfy the SINR requirements all the time. Differently, the probabilistic robust optimization satisfies the SINR requirement in a statistical way, where SINR outage is allowed to occur in a proper way. Thirdly, the power consumption of all the algorithms increases with a higher value of the users' SINR requirements.

Fig. 4 shows how the channel estimation error affects the transmission power. For the probabilistic manner optimization, a tough channel estimation error increases the norm of covariance matrices, i.e., $\bar{\gamma}_{ik}$ and γ_{ik} , of the estimated channels. According to Eqs. (12) and (13), the amplitude of the precoder (also the transmission power) has to be improved

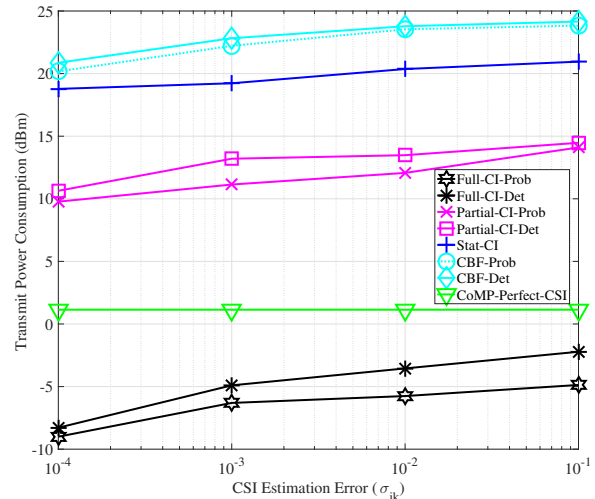


Fig. 4. Impact of channel estimation error on the total power consumption.

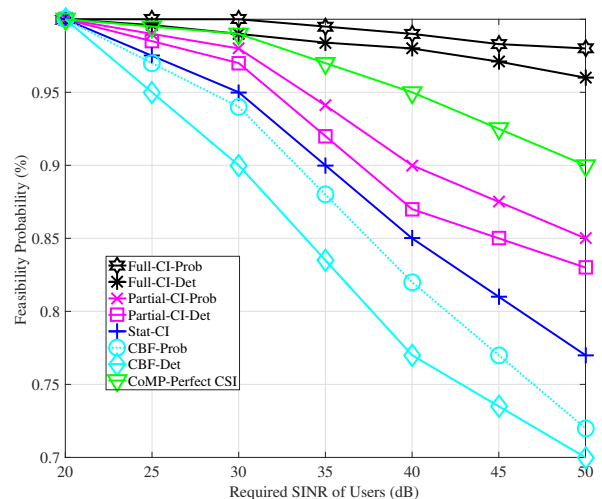


Fig. 5. Feasibility probability (%) versus different SINR requirements $\bar{\gamma}_{ik}$, where the maximum transmission power at each BS $\rho_{max} = 40$ dBm is imposed for all the algorithms.

to make the optimization feasible, leading to a increased transmission power. The same trend is also applicable for the Partial-CI-Prob algorithm. On the other hand, the deterministic manner robust optimization needs to keep the positive semi-definite characteristic for the matrices in Eqs. (22) and (23), which mathematically requires that all the leading principal minors in the matrices to be nonnegative. As a result, the transmission power of all the deterministic robust optimization increases with a tough CSI estimation. At last, since the CoMP-imperfect scheme is designed based on the ideal CSI acquisition, its power consumption is independent with the CSI error.

Fig. 5 shows the feasibility probability with different SINR requirements, where infeasibility occurs when the consumed power violates the maximum power ρ_{max} . As can be seen, the Full-CI-prob and Full-CI-det achieve the highest feasibility

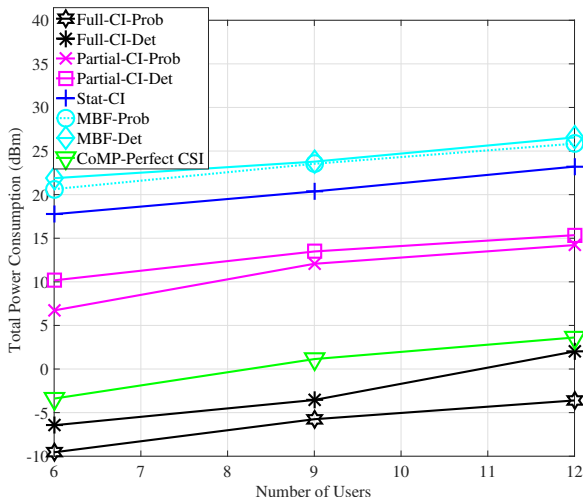


Fig. 6. Impact of the total number of users on the total power consumption.

probabilities, up to 95% even with SINR requirement $\bar{\gamma}_{ik} = 50$ dB. As a comparison, the feasibility of the CoMP-perfect CSI degrades to 90% with $\bar{\gamma}_{ik} = 50$ dB SINR requirement. Besides, the Partial-CI-Prob/Det algorithms achieve lower feasibility probabilities compared to the CoMP-like schemes, as inter-channel gain is considered as harmful element and is suppressed. As a result, they have higher probabilities of violating maximum power budget to achieve a target SINR, e.g., around 85% feasibility probability with $\bar{\gamma}_{ik} = 50$ dB requirement. At last, neither inter-cell or multi-user interference is exploited by the conventional CBF-Prob and CBF-Det benchmarks, and hence their feasibility probabilities sharply degrade to 70%, lower than the 77% achieved by the Stat-CI algorithm.

Fig. 6 demonstrates the power consumption versus different number of users. As can be seen, the proposed algorithms consume less power compared to their corresponding benchmarks with different number of multiple users. Besides, higher power consumption is required with more users. It is because according to the problem formulations, all the users' SINR requirements need to be probabilistically or deterministically satisfied by the constraints. As a result, higher transmission power is led due to the increased number of constraints.

Fig. 7 shows the transmission power for the different levels of coordination algorithms. For illustration, we place 6 users in the BS1 and BS2's edge area while no user in the BS3's coverage. By the Full-CI-Prob/Det, although there is no user in the coverage area of the BS3, the BS3 still coordinates with the BS1 and BS2 as a network-level MISO to serve the edge users. Hence, as presented by Figs. 7 (a) and (b), the BS3 contributes almost identical power compared to the BS1 and BS2. In contrast, by the Partial-CI-Prob/Det algorithms, the BS3 keeps silent when there is no user within its cell, as shown by Figs. 7 (c) and (d). It is because by the Partial-CI-Prob/Det algorithms, BSs only transmits data to the users within its cell, while the inter-cell gain is strictly suppressed by joint beamforming. The same trend is observed by the Stat-

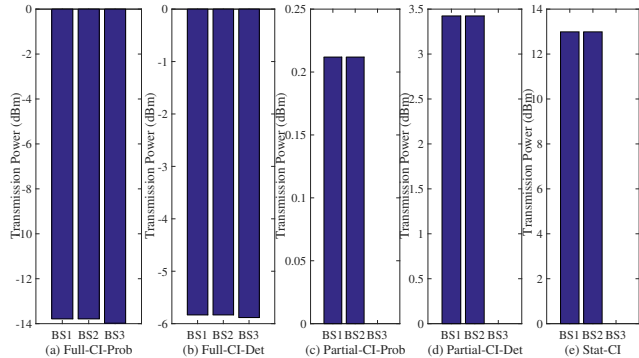


Fig. 7. Transmission power of the BSs under different algorithms.

CI scheme, which suppresses inter-cell interference with the knowledge of statistical inter-cell channels.

Fig. 8 shows the distribution of SINRs obtained by the different algorithms. Firstly, it can be seen from Figs. 8 (b) and (d) that, all the worst-case based optimizations, i.e., Full-CI-Det and Partial-CI-Det, can always guarantee that the achieved SINR higher than the preset SINR requirements. In contrast, the chance-constrained based optimizations, i.e., Full-CI-Prob and Partial-CI-Prob in Figs. 8 (a) and (c), allow proper outage occurs obeying the preset threshold $\bar{\gamma}_{ik}$. Secondly, by comparing Full-CI and Partial-CI schemes, it is observed that the achieved SINR by the Partial-CI scheme is long-tailed, e.g., from 25 dB to 30 dB in Figs. 8 (c) and (d). It is because a BS may radiate high transmit power to compensate the edge users' SINR by the Partial-CI scheme, and hence the users close to the BS may occasionally obtain extreme high SINR, leading to a high pdf value at high SINR regime. Different, the Full-CI scheme efficiently corporates adjacent BSs to serve edge users without radiating high transmission power, and hence all the users' SINR is less expanded and centers on the SINR target. Thirdly, it is observed that the SINR performance of the Stat-CI is impaired due to only knowing the statistical inter-cell channel. However, the Stat-CI still outperforms the conventional CoMP with imperfect CSI.

Fig. 9 shows the average running time of different algorithms versus the number of users. It can be seen proposed algorithms require a lower running time. Among them, the Partial-CI-Prob/Det algorithms require the highest time for obtaining the optimal results. It corresponds to our analysis that the Partial-CI-Prob/Det algorithms are confined by more constraints, and more slack variables are introduced to make the problem solvable. Hence, the Partial-CI-Prob/Det algorithms require longer time to get the optimal results. Besides, all the three benchmarks, CoMP-Perfect-CSI and CBF-Prob/Det, need longer time to get convergence, although they subject to fewer constraints. It is because by the conventional methods, the number of variables (precoders) exactly equals to the number of users. In contrast, the proposed schemes transform the transmission channel into a multi-cast channel by utilizing CI, and hence the number of variables only equals to the number of the coordinated BSs. As a result, even though the proposed algorithms subject to more constraints, a shorter

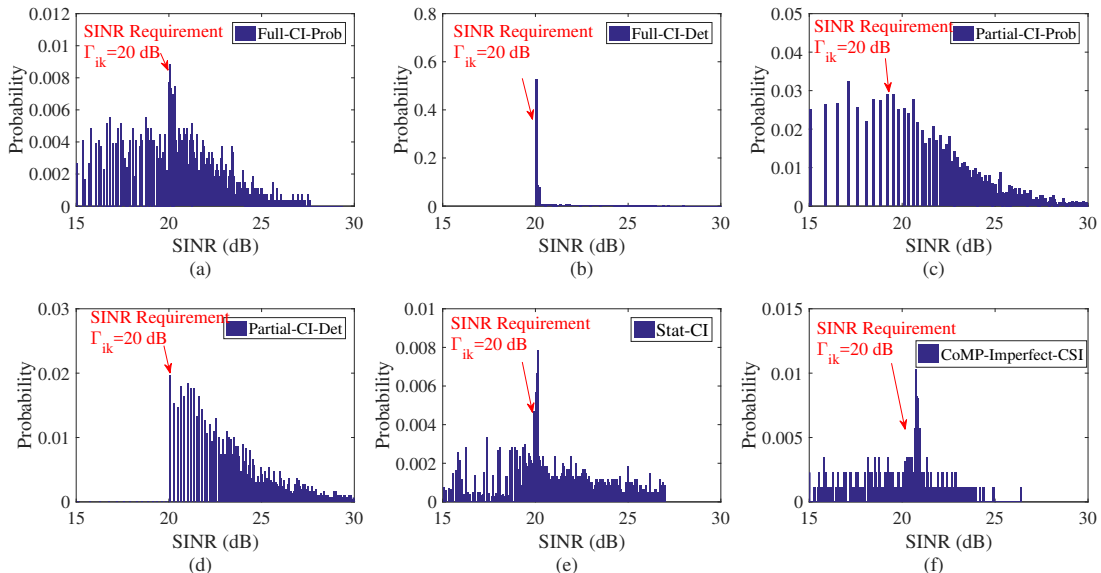


Fig. 8. Probability distribution function (pdf) of the SINRs by the different algorithms a) Full-CI-Prob, b) Full-CI-Det, c) Partial-CI-Prob, d) Partial-CI-Det, e) Stat-CI, f) Conventional CoMP (Imperfect CSI case).

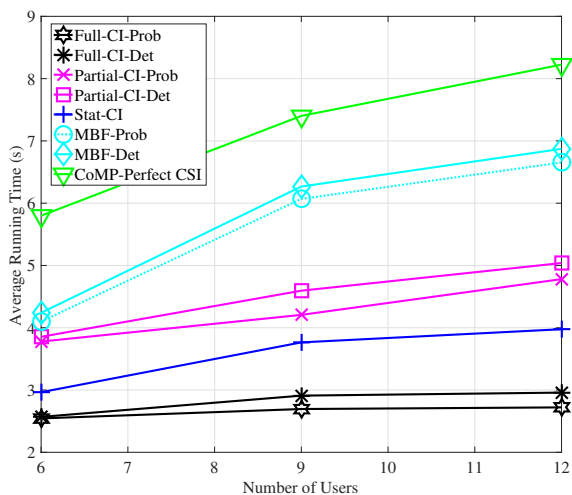


Fig. 9. Impact of the total number of users on the execution time.

running time is provisioned over the conventional schemes. With in-total 12 users served by $N_{BS} = 3$ BSs, there are 12 variables in the conventional algorithms while the number of variables is reduced to 3 by the proposed CI-based algorithms. As a result, the algorithms require 2-5 seconds less than the conventional algorithms.

Fig. 10 demonstrates the total coordination overhead required by the different schemes. As analyzed in Section VI, the total coordination overhead consists of sharing CSI and symbols. For CSI exchange, codebook-based CSI overhead is accounted since it requires less bits to describing CSI over the analog-feedback method. Explicitly, by the codebook-based CSI acquisition, BSs only need to share the indexes of the

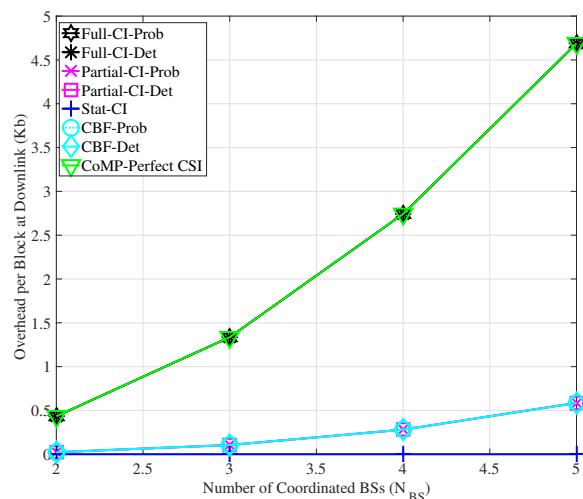


Fig. 10. The total coordination overhead in each sub-frame versus the number of the coordinated BSs.

selected codeword with other BSs, and then each BS locally interprets the codewords which are closest to the projection of the CSI. Generally, around 2-10 bits are needed to describe a MISO channel for each user [43]. For the overhead on sharing symbols, LTE Type 2 downlink frame structure is considered, where 5 out of 10 sub-frames are designed for downlink transmission and each sub-frame contains 14 symbols. It can be seen, all the CoMP-like schemes, i.e., Full-CI-Prob/Det and CoMP-Perfect-CSI, need high overhead, since both the intended transmitted symbols and CSI have to be shared among BSs, and especially sharing the data introduces high coordination overhead. In the lower-level of coordination, Partial-CI-Prob/Det and CBF-Prob/Det require relatively moderate

overhead, where only CSI needs to be shared among BSs. As a result, coordination overhead is reduced to 0.5 Kb when 5 BSs are coordinated. At last, it should be noted that the Stat-CI requires no backhaul overhead for coordination, since only statistical inter-cell CSI needs to be exchanged among the coordinated BSs, which can be ignored compared to other instantaneous CSI acquisition based algorithms.

VIII. CONCLUSIONS

In this paper, we have investigated robust power-efficient precoding design for multi-cell coordination systems, based on exploiting inter-cell and multi-user interference. Three schemes have been proposed to fully/partially utilize inter-cell and multi-user interference, with different levels of coordination overhead. At the presence of CSI estimation error, we have processed the power minimization problems in terms of probabilistic and deterministic robust optimizations. Explicitly, probabilistic optimization satisfies the SINR requirements in a statistical manner while deterministic optimization satisfies the SINR requirement all the time. The simulation results verify that, by sharing the intended symbols and CSI, the Full-CI-Prob/Det algorithms consume the lowest power, and even outperform the conventional CoMP with perfect CSI acquisition. Differently, by only sharing CSI but not the intended symbols, the Partial-CI-Prob/Det algorithms require moderate coordination overhead whereas outperforming the conventional CBF-based schemes in terms of power efficient transmission. To further reduce coordination overhead, the finally proposed Stat-CI algorithm obtains reasonable transmission power consumption with the lightest overhead. Last but not least, the proposed algorithms require low running time, benefiting from the multi-cast transmission characteristics of CI precoding.

APPENDIX A

PROOF OF LEMMA 2

Since the element of CSI error vector e_{ik} follows Gaussian distribution such that $CN(0; \frac{2}{ik})$, the real part $e_{ik}^<$ and imaginary part $e_{ik}^>$ follow $CN(0; \frac{2}{2k})$. Hence, $(\frac{e_{ik}^>}{\sqrt{2}})^T I_{2M} \rho_{\frac{e_{ik}^>}{\sqrt{2}}} \frac{e_{ik}^>}{\sqrt{2}} \leq 0$ can be equivalently written as $\prod_{j=1}^{N_{BS}} (1 + \tan^2) e_{ik}^> + (1 - \tan^2) e_{ik}^<$, which can be approximately seen as the probability distribution function (pdf) of a Gaussian distributed variable such that $\Pr \{ \prod_{j=1}^{N_{BS}} (1 + \tan^2) e_{ik}^> + (1 - \tan^2) e_{ik}^< \leq 0 \} = g$, where g physically represents the probability of the inequality being satisfied. The value of g can be set to close to 1, i.e. $g = 0.99$, meaning the inequality is satisfied with a high probability. By normalizing it into a standard Gaussian distributed variable, we get a cumulative distribution function (cdf) as $\Pr \{ \frac{\prod_{j=1}^{N_{BS}} (1 + \tan^2) e_{ik}^> + (1 - \tan^2) e_{ik}^<}{\sqrt{MN_{BS}(1 + \tan^2)}} \leq 0 \} = g$, where $\Phi(\cdot)$ is the cdf of a standard Gaussian distributed variable. Defining $\Phi^{-1}(g)$ as the inverse function of $\Phi(\cdot)$, finally we get $\frac{\prod_{j=1}^{N_{BS}} (1 + \tan^2) e_{ik}^> + (1 - \tan^2) e_{ik}^<}{\sqrt{MN_{BS}(1 + \tan^2)}} = \Phi^{-1}(g)$.

APPENDIX B

PROOF OF LEMMA 3

The term $e_{ik}^H I e_{ik}$ can be equivalently written as $(\frac{e_{ik}^>}{\sqrt{2}})^2 + \dots + (\frac{e_{ik}^<}{\sqrt{2}})^2$. The left hand follows

chi-square distribution with M degrees of freedom. Hence, $e_{ik}^H I e_{ik}$ can be approximately interpreted as the pdf of a chi-square variable with M degrees of freedom such that $\Pr \{ \frac{e_{ik}^>}{\sqrt{2}} \leq \frac{g}{\sqrt{2}} \} = g$, where g physically represents the probability of the inequality being satisfied. The value of g can be set to close to 1, i.e. $g = 0.99$. Evidently, it can be written as a cdf as $\Pr \{ \frac{e_{ik}^>}{\sqrt{2}} \leq \frac{g}{\sqrt{2}} \} = g$, where $\Phi(\cdot)$ is the cdf of a chi-square variable. Defining $\Phi^{-1}(g)$ as the inverse function of $\Phi(\cdot)$, finally we get $\frac{e_{ik}^>}{\sqrt{2}} = \Phi^{-1}(g)$.

REFERENCES

- [1] C. Masouros, T. Rantajarvi, and A. K. S. Qinetiq, "Known interference in the cellular downlink: a performance limiting factor or a source of green signal power?" *IEEE Commun. Mag.*, vol. 51, no. 10, pp. 162-171, Oct. 2013.
- [2] D. Lee, H. Seo, B. Clerckx, E. Hardouin, 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.
- [3] H. Zeng, X. Zhu, Y. Jiang, and Z. Wei, "A green coordinated multi-cell NOMA system with fuzzy logic multi-criterion user mode selection and resource allocation," *IEEE Sel. Topics Sig. Process.*, vol. 13, no. 3, pp. 480-495, Jun. 2019.
- [4] P. Komulainen, A. Tolli, and M. Juntti, "Effective CSI signaling and decentralized beam coordination in TDD multi-cell MIMO systems," *IEEE Trans. Signal Process.*, vol. 61, no. 9, pp. 2204-2218, May 2013.
- [5] L. Venturino, N. Prasa, and X. Wang, "Coordinated linear beamforming in downlink multi-cell wireless networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 4, pp. 1451-1461, Apr. 2010.
- [6] D. H. N. Nguyen and T. L. Ngoc, "Sum-rate maximization in the multi-cell MIMO multiple-access channel with interference coordination," *IEEE Trans. Wireless Commun.*, vol. 13, no. 1, pp. 36-48, Jan. 2014.
- [7] H. Huh, S. Moon, Y. Kim, I. Lee, and G. Caire, "Energy multi-cell MIMO downlink with cell cooperation and fair scheduling: a large-system limit analysis," *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 7771-7786, Dec. 2011.
- [8] M. Simsek, M. Bennis, and I. Gven, "Learning based frequency- and time-domain inter-cell interference coordination in HetNets," *IEEE Trans. Veh. Technol.*, vol. 64, no. 10, pp. 4589-4602, Oct. 2015.
- [9] H. Wu, X. Gao, X. Wang, and X. You, "Sum-rate-optimal precoding for multi-cell large-scale MIMO uplink based on statistical CSI," *IEEE Trans. Commun.*, vol. 63, no. 8, pp. 2924-2935, Aug. 2015.
- [10] S.-H. Park, H. Park, H. Kong, and I. Lee, "New beamforming techniques based on virtual SINR maximization for coordinated multi-cell transmission," *IEEE Trans. Wireless Commun.*, vol. 11, no. 3, pp. 1034-1044, Mar. 2012.
- [11] A. Tolli, H. Pannanen, and P. Komulainen, "Decentralized minimum power multi-cell beamforming with limited backhaul signaling," *IEEE Trans. Commun.*, vol. 10, no. 2, pp. 570-580, Feb. 2011.
- [12] Y. Kim, H. Lim, M. Song, and G. Im, "Power efficient transceiver designs for multi-cell coordination in MIMO cognitive radio networks," *IEEE Trans. Commun.*, vol. 61, no. 10, pp. 4127-4138, Oct. 2013.
- [13] D. Gesbert, S. Hanly, H. Huang, S. S. Shitz, O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: a new look at interference," *IEEE Journal on Sel. Area in Commun.*, vol. 28, no. 9, pp. 1380-1408, Dec. 2010.
- [14] F. Wang and Q. Wu, "Joint optimization of user association, subchannel allocation, and power allocation in multi-cell multi-association OFDMA heterogeneous networks," *IEEE Trans. Commun.*, vol. 65, no. 6, pp. 2672-2684, Jun. 2017.
- [15] M. Ali, E. Hossain, A. Al-Dweik, and D. I. Kim, "Downlink power allocation for CoMP-NOMA in multi-cell networks," *IEEE Trans. Commun.*, vol. 66, no. 9, pp. 3982-3998, Sep. 2018.
- [16] S. Ararat, and M. R. Nakhai, "Robust distributed beamforming with interference coordination in downlink cellular networks," *IEEE Trans. Commun.*, vol. 62 no. 7, pp. 2411-2421, Jul. 2014.
- [17] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425-3441, Aug. 2008.
- [18] V. R. Cadambe and S. A. Jafar, "Interference Alignment and the degrees of freedom of wireless X networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 3893-3908, Sept. 2009.

- [19] C. Masouros and E. Alsusa, "A novel transmitter-based selective precoding technique for DS/CDMA systems," *IEEE Sig. Process. Lett.*, vol. 14, no. 9, pp. 637-640, Sep. 2007.
- [20] C. Masouros, "Correlation rotation linear precoding for MIMO broadcast communications," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 252-262, Jan. 2011.
- [21] M. Alodeh, S. Chatzinotas, and B. Ottersten, "Energy-efficient symbol-level precoding in multiuser MISO based on relaxed detection region," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3755-3767, May 2016.
- [22] D. Spano, M. Alodeh, S. Chatzinotas, and B. Ottersten, "Symbol level precoding for the non-linear multiuser MISO downlink channel," *IEEE Trans. Signal Process.*, vol. 66, no. 5, pp. 1331-1345, Mar. 2018.
- [23] C. Masouros, M. Sellathurai, and T. Ratnarajah, "Vector perturbation based on symbol scaling for limited feedback MISO downlinks," *IEEE Trans. Sig. Proc.*, vol. 62, no. 3, pp. 562-571, Feb. 2014.
- [24] C. Masouros and G. Zheng, "Exploiting known interference as green signal power for downlink beamforming optimization," *IEEE Trans. Sig. Proc.*, vol. 63, no. 14, pp. 3668-3680, Jul. 2015.
- [25] K. L. Law, C. Masouros, and M. Pesavento, "Transmit precoding for interference exploitation in the underlay cognitive radio Z-channel," *IEEE Trans. Signal Process.*, vol. 65, no. 14, pp. 3617-3630, Jul. 2017.
- [26] P. V. Amadori and C. Masouros, "Large scale antenna selection and precoding for interference exploitation," *IEEE Trans. Commun.*, vol. 65, no. 10, pp. 4529-4542, Oct. 2017.
- [27] A. Li and C. Masouros, "Interference exploitation precoding made practical: optimal closed-form solution for PSK modulations," Dec. 2017. [Online]. Available: <https://arxiv.org/pdf/1712.07846.pdf>
- [28] S. Timotheou, G. Zheng, and C. Masouros, "Exploiting constructive interference for simultaneous wireless information and power transfer in multiuser downlink systems," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 5, pp. 1772-1784, May 2015.
- [29] M. T. Kabir and C. Masouros, "A scalable energy vs latency trade-off in full-duplex mobile edge computing systems," *IEEE Trans. Commun.*, DOI 10.1109/TCOMM.2019.2915833.
- [30] M. R. A. Khandaker, K. Masouros, and K. K. Wong, "Constructive interference based secure precoding: a new dimension in physical layer security," *IEEE Trans. Inf. Foren. Sec.*, vol. 13, no. 9, pp. 2256-2268, Sep. 2018.
- [31] Z. Wei, S. Sun, X. Zhu, Y. Huang, and J. Wang, "Energy-efficient hybrid duplexing strategy for bidirectional distributed antenna systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 6, pp. 5096-5110, Apr. 2018.
- [32] M. Mallick, R. Devarajan, R. A. Loodaricheh, and V. K. Bhargava, "Robust resource optimization for cooperative cognitive radio networks with imperfect CSI," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 907920, Feb. 2015.
- [33] IEEE 802.16 Broadband Wireless Access Working Group. [2007, Nov] Channel estimation modeling for system simulations, [Online] Available: http://ieee802.org/16/tgm/contrib/C80216m-07_208r2.pdf
- [34] S. Boyd, Stanford University [2008, Winter Quarter] Linear dynamical systems, [Online] Available: <https://stanford.edu/class/ee363/lectures/estim.pdf>
- [35] S. Nasserli, M. R. Nakhai, and T. A. Le, "Chance constrained robust downlink beamforming in multicell networks," *IEEE Trans. Mobile Comput.*, vol. 15, no. 11, pp. 2682-2691, Nov. 2016.
- [36] C. Shen, T. Chang, K. Wang, Z. Qiu, and C. Chi, "Distributed robust multi-cell coordinated beamforming with imperfect CSI: an ADMM approach," *IEEE Trans. Sig. Proc.*, vol. 60, no. 6, pp. 2988-3003, Jun. 2012.
- [37] X. Jia, C. Zhang, and I. Kim, "Worst-case robust beamforming design for wireless-powered multi-relay multiuser network with a nonlinear EH model," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 3038-3042, Mar. 2019.
- [38] S. Salari, M. Z. Amirani, I. M. Kim, D. I. Kim, and J. Yang, "Distributed beamforming in two-way relay networks with interference and imperfect CSI," *IEEE Trans. Wireless Commun.*, vol. 15, no. 6, pp. 4455-4469, Jun. 2016.
- [39] E. Boshkovska, D. W. K. Ng, and R. Schober, "Robust resource allocation for MIMO wireless powered communication networks based on a non-linear EH model," *IEEE Trans. Commun.*, vol. 65, no. 5, pp. 1984-1999, May 2017.
- [40] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [41] W. Liao, T. Chang, W. Ma, and C. Chi, "OoS based transmit beamforming in the presence of eavesdroppers: an optimized artificial noise-aided approach," *IEEE Trans. Sig. Proc.*, vol. 59, no. 3, pp. 1202-1015, Mar. 2011.
- [42] K. Wang, W. Ma, and C. Chi, "Outage constrained robust transmit optimization for multiuser MISO downlinks: tractable approximations by conic optimization," *IEEE Trans. Antenna Propagat.*, vol. 62, no. 21, pp. 5690-5715, Nov. 2014.
- [43] P. Huang and Y. Pi, "A novel MIMO channel state feedback scheme and overhead calculation," *IEEE Trans. Commun.*, vol. 16, no. 10, pp. 4550-4562, Oct. 2018.



Zhongxiang Wei (S'15M'17) is research associate of Electrical and Electronics Engineering at University College London (UCL), U.K. He received the Ph.D. degree in Electrical and Electronics Engineering from the University of Liverpool (UOL), Liverpool, U.K., in 2017. From March 2016 to March 2017, he was with the Institution for Info-comm Research, Agency for Science, Technology, and Research (A*STAR), Singapore, as a Research Assistant. From March 2017 to October 2017, he was a visiting student with the Wireless Networks and Communications Group, Harbin Institute of Technology (HIT), Shenzhen, China.

His research interests include constructive interference design, green communications, full-duplex, millimeter-wave communications, and algorithm design. He was the recipient of the Graduate China National Scholarship Award in 2012, and the A*STAR Research Attachment Programme (ARAP) Studentship in 2016. He has acted as a TPC member or session chair for various international conferences, such as ICC 2019 and 2020. He was the recipient of exemplary reviewer of IEEE Transaction on Wireless Communications in 2016.



Christos Masouros (SMIEEE, MIET) received the Diploma degree in Electrical and Computer Engineering from the University of Patras, Greece, in 2004, and MSc by research and PhD in Electrical and Electronic Engineering from the University of Manchester, UK in 2006 and 2009 respectively. In 2008 he was a research intern at Philips Research Labs, UK. Between 2009-2010 he was a Research Associate in the University of Manchester and between 2010-2012 a Research Fellow in Queen's University Belfast. In 2012 he joined University

College London as a Lecturer. He has held a Royal Academy of Engineering Research Fellowship between 2011-2016.

He is currently a Full Professor in the Information and Communications Engineering research group, Dept. Electrical and Electronic Engineering, University College London. His research interests lie in the field of wireless communications and signal processing with particular focus on Green Communications, Large Scale Antenna Systems, Cognitive Radio, interference mitigation techniques for MIMO and multicarrier communications. He was the recipient of the Best Paper Awards in the IEEE GlobeCom 2015 and IEEE WCNC 2019 conferences, and has been recognised as an Exemplary Editor for the IEEE Communications Letters, and as an Exemplary Reviewer for the IEEE Transactions on Communications. He is an Editor for IEEE Transactions on Communications, and IEEE Transactions on Wireless Communications. He has been an Associate Editor for IEEE Communications Letters, and a Guest Editor for IEEE Journal on Selected Topics in Signal Processing issues Exploiting Interference towards Energy Efficient and Secure Wireless Communications and Hybrid Analog / Digital Signal Processing for Hardware-Efficient Large Scale Antenna Arrays. He is currently an elected member of the EURASIP SAT Committee on Signal Processing for Communications and Networking.

Kai-Kit Wong (M'01-SM'08-F'16) received the BEng, the MPhil, and the PhD degrees, all in Electrical and Electronic Engineering, from the Hong Kong University of Science and Technology, Hong Kong, in 1996, 1998, and 2001, respectively. After graduation, he took up academic and research positions at the University of Hong Kong, Lucent Technologies, Bell-Labs, Holmdel, the Smart Antennas Research Group of Stanford University, and the University of Hull, UK. He is Chair in Wireless Communications at the Department of Electronic and Electrical En-

gineering, University College London, UK.

His current research centers around 5G and beyond mobile communications, including topics such as massive MIMO, full-duplex communications, millimetre-wave communications, edge caching and fog networking, physical layer security, wireless power transfer and mobile computing, V2X communications, and of course cognitive radios. There are also a few other unconventional research topics that he has set his heart on, including for example, fluid antenna communications systems, remote ECG detection and etc. He is a co-recipient of the 2013 IEEE Signal Processing Letters Best Paper Award and the 2000 IEEE VTS Japan Chapter Award at the IEEE Vehicular Technology Conference in Japan in 2000, and a few other international best paper awards.

He is Fellow of IEEE and IET and is also on the editorial board of several international journals. He has served as Senior Editor for IEEE Communications Letters since 2012 and also for IEEE Wireless Communications Letters since 2016. He had also previously served as Associate Editor for IEEE Signal Processing Letters from 2009 to 2012 and Editor for IEEE Transactions on Wireless Communications from 2005 to 2011. He was also Guest Editor for IEEE JSAC SI on virtual MIMO in 2013 and currently Guest Editor for IEEE JSAC SI on physical layer security for 5G.

Xin Kang (M'11) received his B.Eng degree in Electrical Engineering from Xian Jiaotong University, China, in 2005. He received his Ph.D. degree in the Electrical and Computer Engineering from National University of Singapore, Singapore, in 2011. He was a research scientist at Institute for Infocomm Research (I2R), A*STAR, Singapore, from 2011 to 2014. After that, he joined Shield Lab, Huawei Singapore as a senior researcher. He has filed more than 30 patents on security protocols, and contributed more than 20 technical proposals to

3GPP SA3. From July 2016, he has joined University of Electronic Science and Technology of China as a full professor.

His research interests include optimization, game theory, energy harvesting, cognitive radio and AI, data offloading, and security protocol designs. He has received Best Paper Award from IEEE ICC 2017, and Best 50 Papers of IEEE GlobeCom 2014.