Distributed Precoding Using Local CSIT for MU-MIMO Heterogeneous Cellular Networks

Deokhwan Han, *Student Member, IEEE,* and Namyoon Lee, *Senior Member, IEEE,*

Abstract—Cell densification is a key driver to increase area spectral efficiencies in multi-antenna cellular systems. As increasing the densities of base stations (BSs) and users that share the same spectrum, however, both inter-user-interference (IUI) and inter-cell interference (ICI) problems give rise to a significant loss in spectral efficiencies in such systems. To resolve this problem under the constraint of local channel state information per BS, in this paper, we present a novel noncooperative multi-user multiple-input multiple-output (MIMO) precoding technique, called *signal-to-interference-pulse-leakage-pulse-noise-ratio (SILNR) maximization* precoding. The key innovation of our distributed precoding method is to maximize the product of SILNRs of users per cell using local channel state information at the transmitter (CSIT). We show that our precoding technique only using local CSIT can asymptotically achieve the multi-cell cooperative bound attained by cooperative precoding using global CSIT in some cases. We also present a precoding algorithm that is robust to CSIT errors in multi-cell scenarios. By multi-cell system-level simulations, we demonstrate that our distributed precoding technique outperforms all existing noncooperative precoding methods considerably and can also achieve the multi-cell bound very tightly even with not-so-many antennas at BSs.

I. INTRODUCTION

Heterogeneous cellular networks (HetNets) are promising solutions for achieving high data rates ubiquitously [1]–[5]. HetNets are comprised of distinct network tiers, each with different transmission power and number of antennas. By deploying low power base stations (BSs) (e.g., pico and femto BSs) overlaid onto macrocells, HetNets considerably increase area spectral efficiencies by cell-splitting gains. When operating these small BSs using the same frequency/time resources with macro-cells, significant intra-tier and inter-tier interference problems take place. This interference problem makes the area spectral efficiency gains dwindle. As a result, an effective interference mitigation technique is indispensable to obtain the area spectral efficiency gains in HetNets [6]–[8].

Massive multiple-input multiple-output (MIMO) is an effective solution to resolve this interference problem by exploiting a large degree of freedom in the spatial domain [9]–[11]. With time-division-duplexing (TDD) operation, it has shown in [12] that simple precoding techniques with perfect channel state information at the transmitter (CSIT) are sufficient to eliminate both inter-user-interference (IUI) and inter-cell-interference (ICI) when using an infinite number of antennas. This result implies that noncooperative precoding with perfect CSIT of the associated users in a cell can be asymptotically optimal

D. Han and N. Lee are with the Department of Electrical Engineering, POSTECH, Pohang, Gyeongbuk 37673, South Korea (e-mail: {dhhan, nylee}@postech.ac.kr).

when the number of antennas is sufficiently larger than that of active downlink users per cell [13]. In HetNets, however, the number of antennas of BSs in small-cells cannot be many by the cost and hardware limitations [1]–[3]. Besides, the density of active users in the hotspots is relatively high. In this case, the channel hardening effects are not sufficiently pronounced except for the ray-based channel model cases [14]. In general, the simple precoding methods fail to successfully eliminate both ICI and IUI in massive MIMO networks [12], [13].

This paper focuses on a multi-user MIMO (MU-MIMO) HetNet where the macro-cell tier is overlaid with small-cells, each small-cell BS is equipped with a few antennas. The user density of the small-cell areas is higher than that of the other areas. With a limited number of antennas at the BSs, most prior works have focused on cooperative precoding strategies to eliminate both IUI and ICI, which provides considerable gains in the sum-spectral efficiency. In practice, however, these cooperative precoding strategies are undesirable when considering the overheads for the CSI exchange among BSs via backhaul signaling; the gains by the cooperation can disappear [15], [16]. Instead, the precoding strategies using local CSIT have significant merit to simultaneously reduce IUI and interference leakage to other cells' users without causing any signaling overheads.

One of the challenges in designing noncooperative precoding strategies is that the ICI leakage mitigation using local CSIT does not necessarily maximize the sum-spectral efficiency; thereby, the performance gap between the cooperative and noncooperative methods can be consequential. Therefore, finding a noncooperative precoding strategy that can closely attain the performance gain of the cooperative precoding is a significant yet challenging problem. This paper tackles this problem and shows finding such a noncooperative precoding strategy is affirmative in some instances.

A. Prior Works

There are extensive prior studies for multi-cell linear precoding methods using local CSIT. The simplest method is maximum ratio transmission (MRT) [17], which is also known as matched filtering (MF) precoding. To employ MRT, each BS only requires to have the local CSIT of its cell. In a multi-cell massive MIMO setting, where the number of BS antennas is much larger than that of users, it has shown in [12] that this simple precoding can asymptotically eliminate both IUI and ICI under perfect local CSIT assumption. Thanks to its simple precoding structure, the analytical expressions for achievable rates have been derived in closed-forms as

a function of relevant system parameters in massive MIMO settings [18], [19].

Zero-forcing (ZF) [20] is another popular precoding method to eliminate IUI using local CSIT. Unlike MRT precoding, it can entirely remove IUI regardless of the number of BS antennas by selecting the number of users that is not larger than that of BS. Particularly, when the number of users is sufficiently larger than that of BS antennas, ZF precoding with semi-orthogonal user selection [21] has shown to asymptotically achieve the optimal capacity scaling law attained by [22] (DPC). In a massive MIMO setup, in which a BS has an infinite number of antennas, ZF precoding can maximize the sum-spectral efficiency under the perfect CSIT assumption [12]. When the BS has a not-so-large number of antennas compared to the number of users such as HetNets [1], [2], [23], ZF precoding is not effective to mitigate both IUI and ICI simultaneously because of the inefficient utilization of the spatial degrees of freedom.

Signal-to-leakage-plus-noise-ratio (SLNR) precoding [24] is an effective method to suppress both IUI and ICI using local CSIT in multi-cell MIMO networks. It turns out that this precoding is equivalent to the minimum mean square error (MMSE) precoding (or regularized ZF precoding) under uniform power allocation. In a multi-cell massive MIMO setting, the multi-cell MMSE precoding method has been proposed in [25] by taking into account pilot contamination effects. Particularly, in cell-free massive MIMO setting [26], SLNR precoding is an attractive precoding method using local CSIT because of its scalability and ICI mitigation capability. The major limitation is that SLNR maximization precoding does not necessarily maximize sum-spectral efficiency.

Several linear precoding algorithms have been proposed to maximize the sum-spectral efficiency in single-cell and multi-cell multi-user MIMO systems [25], [27]–[29]. Finding the global optimal linear precoder that maximizes the sum-spectral efficiency is NP-hard. The weighted-MMSE (WMMSE) precoding is the most popular sub-optimal precoding technique in the sum-spectral efficiency maximization problem [27]. Thanks to the equivalence between the sum-spectral efficiency maximization problem and the WMMSE minimization, an alternating minimization algorithm has been proposed, which converges to a local-optimal solution. However, this precoding cannot be applicable when the number of antennas is massive at the BS because it requires very high computational complexity. Recently, inspired by principal component analysis (PCA), a novel low-complexity algorithm called generalized power iteration precoding (GPIP) has been presented, which guarantees the first-order optimality for the sum-spectral efficiency maximization problem under perfect and noisy CSIT [28]. This precoding method has also been extended to a multi-cell scenario with pilot contamination effects [29]. This multi-cell precoding method, however, requires to exchange the CSI among BSs, which causes a significant signaling overhead.

B. Contributions

The main contributions of this paper is summarized as follows:

- We introduce a new performance metric to effectively mitigate both IUI and ICI in a distributed manner using local CSIT for MU-MIMO HetNets. The new performance metric is the ratio of signal-to-interference-pulse-leakage-pulse-noise-ratio (SILNR). Intuitively, SILNR measures the ratio between the desired signal power and the sum of IUI, the interference leakage, and the noise power. However, this new metric significantly differs from the conventional SLNR [24], [25] in two aspects. First, while the IUI in the definition of the existing SLNR is treated as the intra-cell leakage interference (i.e., the uplink IUI), the IUI in the SILNR is the exact downlink IUI. Second, the interference leakage term in the SLNR is the sum of interference leakage signals to the individual users in the other cells. In contrast, the interference leakage term in our SILNR is the geometric mean of the interference leakage signals to the other cell users. We show that maximizing the product of SILNRs of users per cell using local CSIT can achieve the identical sum spectral efficiency with the cooperative precoding method using global CSIT in a two-cell MU-MIMO system under certain cases.

- We generalize this precoding method for a multi-cell setting under noisy CSIT assumption. Unfortunately, the product of SILNRs maximization is a non-convex (and even NP-hard) optimization problem similar to the sum-spectral efficiency maximization problem. To design the precoding method, we first derive the first- and second-order necessary conditions for this non-convex optimization problem. Using the derived conditions, we present a low-complexity iterative algorithm that guarantees to converge a locally-optimal solution. The key innovation of the proposed precoding is to identify the joint solutions for the scheduled users per cell, the beamforming vectors of them, and the power allocated to each beam using local and noisy CSIT.

- Using both link-level and system-level simulations, we exhibit that the proposed precoding method significantly outperforms the existing distributed precoding techniques, including MRT, ZF, SLNR, and the sum-spectral efficiency maximization precoding per cell in both perfect and imperfect CSIT scenarios. One remarkable observation is that the proposed method asymptotically achieves the upper bound performance attained by the multi-cell cooperative precoding method when increasing the number of macro BS antennas. This result confirms that our noncooperative precoding technique using local CSIT can be a pragmatic solution to resolving dense cellular networks' interference problems.

C. Notations

Throughout this paper, we use the following notations. The $\mathbb{C}$ denotes a set of a complex number; $\mathbb{R}$ denotes the set of a real number. The $\otimes$ is the Kronecker product operation. We use $\mathbb{E}_{\mathbf{x}}[\mathbf{x}]$ to denote the expectation of a random vector $\mathbf{x}$. The $\mathrm{Re}\{x\}$ means the real part of a complex scalar x. In addition, we use $\lambda_{\mathsf{min}}(\mathbf{A})$ and $\lambda_{\mathsf{max}}(\mathbf{A})$ to denote the minimum and
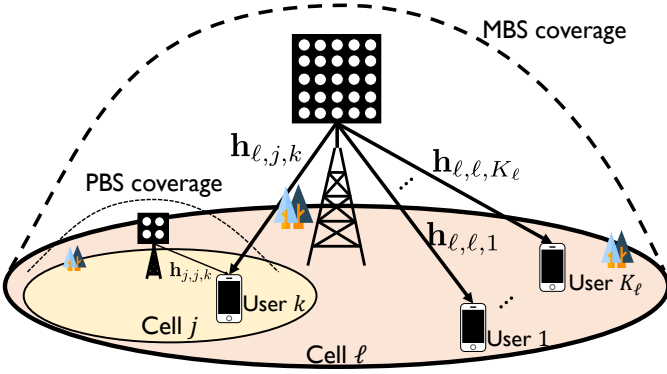
Fig. 1. An illustration of MU-MIMO HetNets.

maximum eigenvalue of matrix $\mathbf{A}$, respectively. A matrix $\mathbf{I}_N$ denote an $N \times N$ identity matrix. Also, $\mathbf{x} \sim \mathcal{CN}(\mathbf{m}, \mathbf{R})$ indicates that the random vector $\mathbf{x}$ is distributed by complex Gaussian distribution with mean vector $\mathbf{m}$ and covariance matrix $\mathbf{R}$.

II. SYSTEM MODEL

We consider a heterogeneous cellular network, which comprises of L BSs each equipped with N_ℓ antennas for $\ell \in \mathcal{L} \triangleq \{1, \ldots, L\}$. The ℓth BS serves K_ℓ users with single antenna.

A. Downlink Channel Model

As illustrated in Fig. 1, we denote the downlink channel vector from the ℓth BS to the kth user in the jth cell by $\mathbf{h}_{\ell,j,k} \in \mathbb{C}^{N_\ell \times 1}$. This downlink channel is modeled as

$$\mathbf{h}_{\ell,j,k} = \beta_{\ell,j,k}^{\frac{1}{2}} \mathbf{g}_{\ell,j,k}, \qquad (1)$$

where $\beta_{\ell,j,k} \in \mathbb{R}$ is a large scale fading coefficient and $\mathbf{g}_{\ell,j,k} \in \mathbb{C}^{N_\ell \times 1}$ is a small scale fading, which is distributed as $\mathbf{h}_{\ell,j,k} \sim \mathcal{CN}(0, \beta_{\ell,j,k} \mathbf{R}_{\ell,j,k})$ with $\mathbf{R}_{\ell,j,k} = \mathbb{E}\left[\mathbf{g}_{\ell,j,k} \mathbf{g}_{\ell,j,k}^{\mathsf{H}}\right]$ for $\forall \ell, j \in \mathcal{L}$ and $\forall k \in \mathcal{K}_j \triangleq \{1, \ldots, K_j\}$. This matrix captures the spatial correlation information on the channel. Under a stationary process assumption, it is typically obtained by using both angle-of-arrival (AoA) vectors of multipaths and the corresponding angular autocorrelation function.

B. Local and Noisy CSIT Acquisition

We present a process of acquiring local CSIT at each BS. For ease of exposition, we focus on the ℓth BS. We define a subset $C_\ell \subset \mathcal{L}$ as the collection of BSs that use mutually orthogonal uplink pilot sequences with those of the ℓth BS for uplink channel training. This BS subset is assumed to be predetermined with a proper cell planning method; thereby, no dynamic BS cooperation is required. Under the premise that ℓth BS has full knowledge of the orthogonal pilots, it can estimate the channel from the users associated with the jth BS where $j \in C_\ell$. For example, let assume that the uplink users served by BS 1, BS 2, BS 3, and BS 4 use mutually orthogonal pilot sequences, then $C_1 = C_2 = C_3 = C_4 = \{1, 2, 3, 4\}$. To guarantee the orthogonality, the uplink pilot length satisfies the condition of $\tau_{\mathsf{u}} \geq \sum_{j=1}^{4} K_j$. Then, BS ℓ can estimate

the channels from the users associated with BS j where $j \in \{1, 2, 3, 4\}$. Thanks to the channel reciprocity in TDD mode, each BS obtains the downlink channel vectors from the uplink channel estimates with a proper RF circuit calibration process.

Applying the MMSE estimator, the estimated downlink channel vector from the ℓth BS to the kth user in the cell is

$$\hat{\mathbf{h}}_{\ell,\ell,k} = \mathbf{h}_{\ell,\ell,k} - \mathbf{e}_{\ell,\ell,k}, \qquad (2)$$

where $\mathbf{e}_{\ell,\ell,k}$ is the estimation error vector and it is distributed as complex Gaussian with zero-mean and covariance matrix $\boldsymbol{\Phi}_{\ell,\ell,k} = \mathbb{E}\left[\mathbf{e}_{\ell,\ell,k} \mathbf{e}_{\ell,\ell,k}^{\mathsf{H}}\right] \in \mathbb{C}^{N_\ell \times N_\ell}$, i.e., $\mathcal{CN}\left(\mathbf{0}, \boldsymbol{\Phi}_{\ell,\ell,k}\right)$. The error covariance matrix is [30], [31]:

$$\boldsymbol{\Phi}_{\ell,\ell,k} = \beta_{\ell,\ell,k} \mathbf{R}_{\ell,\ell,k}$$
$$- \beta_{\ell,\ell,k}^2 \mathbf{R}_{\ell,\ell,k} \left(\sum_{(j,i) \in \bar{C}_\ell} \beta_{\ell,j,i} \mathbf{R}_{\ell,j,i} + \frac{\sigma^2}{\tau_{\mathsf{u}} p_{\mathsf{ul}}} \mathbf{I}_N \right)^{-1} \mathbf{R}_{\ell,\ell,k}, \quad (3)$$

where p_{ul} is the pilot transmission power and $\bar{C}_\ell$ is the collection of users who use the non-orthogonal uplink pilots with that of the kth user in the ℓth cell, i.e., $\bar{C}_\ell = \{(j,i) \mid \forall j \in \mathcal{L} \setminus C_\ell, \forall i \in \mathcal{K}_j\}$.

C. Ergodic Downlink Spectral Efficiency with Noisy CSIT

Suppose τ_{c} be a channel coherence time interval. We also let $\mathbf{x}_\ell[t] \in \mathbb{C}^{N_\ell \times 1}$ be the transmit signal of the ℓth BS using the tth time slot where $t \in [1, \tau_{\mathsf{c}}]$. The ℓth BS transmits K_ℓ independent data symbols $\{s_{\ell,1}[t], \ldots, s_{\ell,K_\ell}[t]\}$ using time slot t along with precoding vectors $\{\mathbf{f}_{\ell,1}, \ldots, \mathbf{f}_{\ell,K_\ell}\}$. The precoding vectors are constructed using noisy and local CSIT at the ℓth BS, i.e., $\{\hat{\mathbf{h}}_{\ell,\ell,1}, \hat{\mathbf{h}}_{\ell,\ell,2}, \ldots, \hat{\mathbf{h}}_{\ell,\ell,K_\ell}\}$ and $\{\hat{\mathbf{h}}_{\ell,j,i}\}$ where $j \in C_\ell$ and $i \in \mathcal{K}_j$. We assume that each data symbol $s_{\ell,k}[t]$ is drawn from a Gaussian codebook with transmit power P_ℓ, i.e., $s_{\ell,k}[t] \sim \mathcal{CN}(0, P_\ell)$. In addition, the linear precoding vectors satisfy the condition of $\sum_{k=1}^{K_\ell} \|\mathbf{f}_{\ell,k}\|_2^2 \leq 1$ to meet the power constraint per BS. Then, the transmit signal of the ℓth BS at the tth time slot is

$$\mathbf{x}_\ell[t] = \sum_{k=1}^{K_\ell} \mathbf{f}_{\ell,k} s_{\ell,k}[t]. \qquad (4)$$

Then, the received signal of the kth user in the ℓth cell is given by

$$y_{\ell,k}[t] = \sum_{j=1}^{L} \mathbf{h}_{j,\ell,k}^{\mathsf{H}} \mathbf{x}_j[t] + n_{\ell,k}[t], \qquad (5)$$

where $n_{\ell,k}[t]$ is the additive complex Gaussian noise with zero-mean and variance σ^2, i.e., $\mathcal{CN}\left(0, \sigma^2\right)$. Then, the signal-to-interference-plus-noise ratio (SINR) of the kth user in the ℓth cell is given by

$$\mathsf{SINR}_{\ell,k} = \frac{\left|\mathbf{h}_{\ell,\ell,k}^{\mathsf{H}} \mathbf{f}_{\ell,k}\right|^2}{\sum_{i \neq k}^{K_\ell} \left|\mathbf{h}_{\ell,\ell,k}^{\mathsf{H}} \mathbf{f}_{\ell,i}\right|^2 + \sum_{j \neq \ell}^{L} \sum_{i=1}^{K_j} \frac{P_j}{P_\ell} \left|\mathbf{h}_{j,\ell,k}^{\mathsf{H}} \mathbf{f}_{j,i}\right|^2 + \frac{\sigma^2}{P_\ell}}. \qquad (6)$$

Let τ_{d} be the downlink channel training length. Using the orthogonal downlink channel training sequence with length $\tau_{\mathrm{d}} \geq \sum_{j \in C_\ell} K_j$, i.e., demodulation reference signals in Long-Term Evolution (LTE) systems, the kth downlink user associated with the ℓth BS can estimate the precoded downlink channel state information at receiver (CSIR), i.e., $\mathbf{h}_{\ell,\ell,k}^{\mathsf{H}} \mathbf{f}_{\ell,k}$. For simplicity, we assume that each downlink user has perfect CSIR for the precoded channel for the ease of exposition. To incorporate the effect of imperfect CSIR, one can use the notion of generalized mutual information introduced in [32], [33] to redefine SINR per user with the channel estimation error variance.

With noisy and local CSIT at the ℓth BS, i.e., $\hat{\mathcal{H}}_\ell = \left\{ \hat{\mathbf{h}}_{\ell,j,i} \mid j \in C_\ell, i \in \mathcal{K}_j \right\}$ corresponding to the $\mathcal{H}_\ell = \left\{ \mathbf{h}_{\ell,j,i} \mid j \in C_\ell, i \in \mathcal{K}_j \right\}$, it can estimate the instantaneous spectral efficiency of the kth downlink user [28], [29]:

$$R_{\ell,k}\left(\hat{\mathcal{H}}_\ell\right) = \mathbb{E}_{\mathcal{H}_\ell \mid \hat{\mathcal{H}}_\ell} \left[\log_2 \left(1 + \mathrm{SINR}_{\ell,k} \right) \mid \hat{\mathcal{H}}_\ell \right], \quad (7)$$

where the expectation is taken over the CSIT error distribution, i.e., $\mathbf{e}_{\ell,\ell,k} \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{\Phi}_{\ell,\ell,k}\right)$ and $\mathbf{e}_{\ell,j,i} \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{\Phi}_{\ell,j,i}\right)$. This quantity measures the average spectral efficiency over the CSIT error distribution for a given estimates of CSIT. Therefore, by taking the expectations over every fading state, the effective ergodic spectral efficiency is given by

$$\begin{aligned} \bar{R}_{\ell,k} &= \left(1 - \frac{\tau_{\mathrm{u}} + \tau_{\mathrm{d}}}{\tau_{\mathrm{c}}} \right) \mathbb{E}_{\hat{\mathcal{H}}_\ell} \left[R_{\ell,k}\left(\hat{\mathcal{H}}_\ell\right) \right] \\ &= \left(1 - \frac{\tau_{\mathrm{u}} + \tau_{\mathrm{d}}}{\tau_{\mathrm{c}}} \right) \mathbb{E} \left[\log_2 \left(1 + \mathrm{SINR}_{\ell,k} \right) \right], \quad (8) \end{aligned}$$

where the pre-log term is a normalization factor by the uplink and downlink channel training overhead. To maximize the ergodic spectral efficiency, we need to optimize the precoding vectors that maximize the instantaneous spectral efficiency using noisy CSIT knowledge in every fading state.

III. Noncooperative Multi-Cell Precoding

In this section, we present a novel noncooperative precoding method using local CSIT. To highlight the idea, we focus on a two-cell MU-MIMO system when each BS has perfect knowledge of local CSIT.

A. From Centralized to Distributed Precoding

We commence by reviewing the multi-cell cooperative precoding method using global CSIT [29]. We then explain when the proposed SILNR maximization precoding using local CSIT can achieve the identical performance to the multi-cell cooperative one.

Cooperative precoding using global CSIT: Let $\mathbf{f}_\ell = \left[\mathbf{f}_{\ell,1}^{\mathsf{H}}, \mathbf{f}_{\ell,2}^{\mathsf{H}}, \ldots, \mathbf{f}_{\ell,K}^{\mathsf{H}} \right]^{\mathsf{H}} \in \mathbb{C}^{N_\ell K_\ell \times 1}$ be the concatenated precoding vector used at the ℓth BS where $\ell \in \{1, 2\}$. We also let $\mathbf{e}_k = [0, \ldots, 1, \ldots, 0]^{\mathsf{T}} \in \mathbb{R}^{K_\ell \times 1}$ be a unit vector with the nonzero value in the kth element. Using this stacked precoding vector, we rewrite the SINR of the kth user in cell $\ell \in \{1, 2\}$ in (6) as

$$\mathrm{SINR}_{\ell,k}(\mathbf{f}_1, \mathbf{f}_2) = \frac{\mathbf{f}_\ell^{\mathsf{H}} \mathbf{S}_{\ell,\ell,k} \mathbf{f}_\ell}{\mathbf{f}_\ell^{\mathsf{H}} \mathbf{U}_{\ell,\ell,k} \mathbf{f}_\ell + \mathbf{f}_{\bar{\ell}}^{\mathsf{H}} \mathbf{C}_{\bar{\ell},\ell,k} \mathbf{f}_{\bar{\ell}}}, \quad (9)$$

where $\bar{\ell} = 3 - \ell \in \{1, 2\}$, $\mathbf{S}_{\ell,\ell,k}$, $\mathbf{U}_{\ell,\ell,k}$, and $\mathbf{C}_{\bar{\ell},\ell,k}$ are defined as

$$\mathbf{S}_{\ell,\ell,k} = \mathbf{e}_k \mathbf{e}_k^{\mathsf{T}} \otimes \mathbf{h}_{\ell,\ell,k} \mathbf{h}_{\ell,\ell,k}^{\mathsf{H}} \in \mathbb{C}^{N_\ell K_\ell \times N_\ell K_\ell}, \quad (10)$$

$$\mathbf{U}_{\ell,\ell,k} = \mathbf{I}_{K_\ell} \otimes \mathbf{h}_{\ell,\ell,k} \mathbf{h}_{\ell,\ell,k}^{\mathsf{H}} - \mathbf{S}_{\ell,\ell,k} + \frac{\sigma^2}{P_\ell} \mathbf{I}_{N_\ell K_\ell} \in \mathbb{C}^{N_\ell K_\ell \times N_\ell K_\ell}, \quad (11)$$

$$\mathbf{C}_{\bar{\ell},\ell,k} = \mathbf{I}_{K_\ell} \otimes \frac{P_\ell}{P_{\bar{\ell}}} \mathbf{h}_{\bar{\ell},\ell,k} \mathbf{h}_{\bar{\ell},\ell,k}^{\mathsf{H}} \in \mathbb{C}^{N_\ell K_\ell \times N_\ell K_\ell}. \quad (12)$$

Then, the sum-spectral efficiency is

$$R^{\mathrm{sum}}(\mathbf{f}_1, \mathbf{f}_2) = R_1^{\mathrm{sum}}(\mathbf{f}_1, \mathbf{f}_2) + R_2^{\mathrm{sum}}(\mathbf{f}_1, \mathbf{f}_2), \quad (13)$$

where

$$R_\ell^{\mathrm{sum}}(\mathbf{f}_1, \mathbf{f}_2) = \sum_{k=1}^{K_\ell} \log_2 \left(\frac{\mathbf{f}_\ell^{\mathsf{H}} \left(\mathbf{S}_{\ell,\ell,k} + \mathbf{U}_{\ell,\ell,k} \right) \mathbf{f}_\ell + \mathbf{f}_{\bar{\ell}}^{\mathsf{H}} \mathbf{C}_{\bar{\ell},\ell,k} \mathbf{f}_{\bar{\ell}}}{\mathbf{f}_\ell^{\mathsf{H}} \mathbf{U}_{\ell,\ell,k} \mathbf{f}_\ell + \mathbf{f}_{\bar{\ell}}^{\mathsf{H}} \mathbf{C}_{\bar{\ell},\ell,k} \mathbf{f}_{\bar{\ell}}} \right) \quad (14)$$

for $\ell \in \{1, 2\}$. Consequently, the sum-spectral efficiency maximization problem is a form:

$$\arg \max_{\mathbf{f}_\ell \in \mathbb{C}^{N_\ell K_\ell}} R^{\mathrm{sum}}(\mathbf{f}_1, \mathbf{f}_2), \quad (15\mathrm{a})$$

$$\text{subject to } \|\mathbf{f}_\ell\|_2^2 = 1, \quad \forall \ell \in \{1, 2\}. \quad (15\mathrm{b})$$

This optimization problem finds a joint solution for a set of scheduled users per cell, the precoding vector, and the power allocation per stream for both BSs to maximize the sum-spectral efficiency. Unfortunately, finding a global optimal solution is infeasible in practice. Besides, global CSIT knowledge is required to obtain a local-optimal solution, as shown in [29]. This global CSIT knowledge requirement underrates the cooperative transmission gains because of the CSIT sharing overheads. This is especially pronounced when the number of cooperative BS increases [15], [16], [34].

Distributed precoding using local CSIT: We propose a novel distributed precoding strategy that harnesses local CSIT only. The central idea is to maximize the sum-spectral efficiency under the local CSIT constraint. To accomplish this, we introduce a new metric named SILNR. The SILNR of the kth user of BS ℓ is defined as

$$\mathrm{SILNR}_{\ell,k}(\mathbf{f}_\ell) = \frac{\mathbf{f}_\ell^{\mathsf{H}} \mathbf{S}_{\ell,\ell,k} \mathbf{f}_\ell}{\mathbf{f}_\ell^{\mathsf{H}} \mathbf{U}_{\ell,\ell,k} \mathbf{f}_\ell + \mathbf{L}_{\ell,\bar{\ell}}(\mathbf{f}_\ell)^{\frac{K_{\bar{\ell}}}{K_\ell}}}, \quad (16)$$

where $\mathbf{L}_{\ell,\bar{\ell}}(\mathbf{f}_\ell)$ is the geometric mean of the interference leakage to users in the other cells by the transmission of BS ℓ, i.e.,

$$\mathbf{L}_{\ell,\bar{\ell}}(\mathbf{f}_\ell) = \left(\prod_{j=1}^{K_{\bar{\ell}}} \mathbf{f}_\ell^{\mathsf{H}} \mathbf{C}_{\ell,\bar{\ell},j} \mathbf{f}_\ell \right)^{\frac{1}{K_{\bar{\ell}}}}. \quad (17)$$

We provide some remarks on this SILNR value.

- $\mathrm{SILNR}_{\ell,k}(\mathbf{f}_\ell)$ is a function of only $\mathbf{f}_\ell$ for $\ell \in \{1, 2\}$; this implies that each BS constructs the aggregated multi-user precoding vector $\mathbf{f}_\ell \in \mathbb{C}^{N_\ell K_\ell}$ without sharing CSIT.
- When $K_{\bar{\ell}} = K_\ell$, the SILNR takes into account the effective interference leakage power as the geometric mean of $\left\{ \mathbf{f}_\ell^{\mathsf{H}} \mathbf{C}_{\ell,\bar{\ell},1} \mathbf{f}_\ell, \ldots, \mathbf{f}_\ell^{\mathsf{H}} \mathbf{C}_{\ell,\bar{\ell},K_{\bar{\ell}}} \mathbf{f}_\ell \right\}$. This geometric mean

structure plays a key role in maximizing the sum-spectral efficiency using local CSIT, which will be explained in the sequel. In addition, when $K_{\bar{\ell}} \neq K_{\ell}$, the exponent of $\mathbf{L}_{\ell,\bar{\ell}}(\mathbf{f}_{\ell})$, $\frac{K_{\bar{\ell}}}{K_{\ell}}$, controls the IUI and the interference leakage toward other cells. For example, in the case when $K_{\ell} \gg K_{\bar{\ell}}$, this exponent $\frac{K_{\bar{\ell}}}{K_{\ell}}$ diminishes the interference leakage power. This implies that the precoder devotes to reduce the IUI power more than the leakage interference power.

- $\mathsf{SILNR}_{\ell,k}(\mathbf{f}_{\ell})$ is always smaller than $\rho_{\ell,k}(\mathbf{f}_{\ell}) = \frac{\mathbf{f}_{\ell}^{\mathsf{H}}\mathbf{S}_{\ell,k}\mathbf{f}_{\ell}}{\mathbf{f}_{\ell}^{\mathsf{H}}\mathbf{U}_{\ell,k}\mathbf{f}_{\ell}}$, because the leakage power term $\mathbf{L}_{\ell,\bar{\ell}}(\mathbf{f}_{\ell})^{\frac{K_{\bar{\ell}}}{K_{\ell}}}$ is positive. Therefore, one can interpret a function $\log_2\left(1 + \mathsf{SILNR}_{\ell,k}\right)$ as a lower bound of the spectral efficiency of $\log_2\left(1 + \rho_{\ell,k}\right)$ by penalizing the interference leakage generated by BS ℓ.

Using this new metric, we define a sum-rate function of BS ℓ as

$$
\begin{aligned}
\hat{R}_{\ell}^{\mathsf{sum}}(\mathbf{f}_{\ell}) &= \sum_{k=1}^{K_{\ell}} \log_2\left(1 + \mathsf{SILNR}_{\ell,k}(\mathbf{f}_{\ell})\right) \\
&= \log_2\left(\prod_{k=1}^{K_{\ell}} \frac{\mathbf{f}_{\ell}^{\mathsf{H}}\left(\mathbf{S}_{\ell,\ell,k} + \mathbf{U}_{\ell,\ell,k}\right)\mathbf{f}_{\ell} + \mathbf{L}_{\ell,\bar{\ell}}(\mathbf{f}_{\ell})^{\frac{K_{\bar{\ell}}}{K_{\ell}}}}{\mathbf{f}_{\ell}^{\mathsf{H}}\mathbf{U}_{\ell,\ell,k}\mathbf{f}_{\ell} + \mathbf{L}_{\ell,\bar{\ell}}(\mathbf{f}_{\ell})^{\frac{K_{\bar{\ell}}}{K_{\ell}}}}\right).
\end{aligned}
\tag{18}
$$

Consequently, each BS independently identifies a joint solution for the user-selection, precoding, and power allocation by exploiting local CSIT. The optimization problem is the following form:

$$
\underset{\mathbf{f}_{\ell}\in\mathbb{C}^{N_{\ell}K_{\ell}\times 1}}{\arg\max} \sum_{k=1}^{K_{\ell}} \log_2\left(1 + \mathsf{SILNR}_{\ell,k}(\mathbf{f}_{\ell})\right),
\tag{19a}
$$

$$
\text{subject to} \quad \|\mathbf{f}_{\ell}\|_2^2 = 1, \quad \forall k \in \mathcal{K}_{\ell}.
\tag{19b}
$$

B. Quasi-Optimal Cases

To shed light on the idea, it is instructive to consider two cases in which the proposed distributed precoding method achieves the cooperative precoding bound very closely.

Case 1 (Zero-IUI condition): Let us consider a multi-cell cooperative precoding strategy that maximizes the tight lower bound under zero-IUI constraint. From (15), the sum-spectral efficiency maximization problem under the zero-IUI constraint becomes

$$
\underset{\mathbf{f}_{\ell}\in\mathbb{C}^{N_{\ell}K_{\ell}}}{\arg\max} \sum_{\ell=1}^{2}\sum_{k=1}^{K_{\ell}} \log_2\left(1 + \mathsf{SINR}_{\ell,k}(\mathbf{f}_1,\mathbf{f}_2)\right),
\tag{20a}
$$

$$
\text{subject to} \quad \mathbf{f}_{\ell}^{\mathsf{H}}\mathbf{U}_{\ell,k}\mathbf{f}_{\ell} = 0, \quad \forall \ell \in \{1,2\}, \forall k \in \mathcal{K}_{\ell},
\tag{20b}
$$

$$
\|\mathbf{f}_{\ell}\|_2^2 = 1, \quad \forall \ell \in \{1,2\}.
\tag{20c}
$$

This optimization finds a joint solution for a set of scheduled users per cell, the precoding vector, and the power allocation per stream for both BSs. Since $\log_2\left(1 + \mathsf{SINR}_{\ell,k}(\mathbf{f}_1,\mathbf{f}_2)\right) \simeq$

$\log_2\left(\mathsf{SINR}_{\ell,k}(\mathbf{f}_1,\mathbf{f}_2)\right)$ for a high SINR regime, when $\mathbf{f}_{\ell}^{\mathsf{H}}\mathbf{U}_{\ell,k}\mathbf{f}_{\ell} = 0$, we approximate the sum-spectral efficiency as

$$
\begin{aligned}
R^{\mathsf{sum}}(\mathbf{f}_1,\mathbf{f}_2) &\simeq \sum_{\ell=1}^{2}\sum_{k=1}^{K_{\ell}} \log_2\left(\mathsf{SINR}_{\ell,k}(\mathbf{f}_1,\mathbf{f}_2)\right) \\
&= \log_2\left(\prod_{k=1}^{K_1} \frac{\mathbf{f}_1^{\mathsf{H}}\mathbf{S}_{1,1,k}\mathbf{f}_1}{\mathbf{f}_2^{\mathsf{H}}\mathbf{C}_{2,1,k}\mathbf{f}_2} \prod_{k=1}^{K_2} \frac{\mathbf{f}_2^{\mathsf{H}}\mathbf{S}_{2,2,k}\mathbf{f}_2}{\mathbf{f}_1^{\mathsf{H}}\mathbf{C}_{1,2,k}\mathbf{f}_1}\right) \\
&= \log_2\left(\frac{\prod_{k=1}^{K_1}\mathbf{f}_1^{\mathsf{H}}\mathbf{S}_{1,1,k}\mathbf{f}_1}{\prod_{k=1}^{K_2}\mathbf{f}_1^{\mathsf{H}}\mathbf{C}_{1,2,k}\mathbf{f}_1}\right) + \log_2\left(\frac{\prod_{k=1}^{K_2}\mathbf{f}_2^{\mathsf{H}}\mathbf{S}_{2,2,k}\mathbf{f}_2}{\prod_{k=1}^{K_1}\mathbf{f}_2^{\mathsf{H}}\mathbf{C}_{2,1,k}\mathbf{f}_2}\right).
\end{aligned}
\tag{21}
$$

Thanks to the zero-IUI condition, the sum-spectral efficiency maximization problem becomes separable with respect to each optimization vector. Therefore, the solution for the joint optimization problem for the aggregated precoding vectors $\mathbf{f}_{\ell}$ is obtained by solving two separate optimization problems independently:

$$
\underset{\mathbf{f}_{\ell}\in\mathbb{C}^{N_{\ell}K_{\ell}\times 1}}{\arg\max} \log_2\left(\frac{\prod_{k=1}^{K_{\ell}}\mathbf{f}_{\ell}^{\mathsf{H}}\mathbf{S}_{\ell,\ell,k}\mathbf{f}_{\ell}}{\prod_{k=1}^{K_{\bar{\ell}}}\mathbf{f}_{\ell}^{\mathsf{H}}\mathbf{C}_{\ell,\bar{\ell},k}\mathbf{f}_{\ell}}\right),
\tag{22a}
$$

$$
\text{subject to} \quad \mathbf{f}_{\ell}^{\mathsf{H}}\mathbf{U}_{\ell,\ell,k}\mathbf{f}_{\ell} = 0, \quad \forall k \in \mathcal{K}_{\ell},
\tag{22b}
$$

$$
\|\mathbf{f}_{\ell}\|_2^2 = 1.
\tag{22c}
$$

Now we turn our attention to the sum-rate function defined in (18). When $\mathbf{f}_{\ell}^{\mathsf{H}}\mathbf{U}_{\ell,\ell,k}\mathbf{f}_{\ell} = 0$, we approximate the sum-rate function as

$$
\begin{aligned}
\hat{R}_{\ell}^{\mathsf{sum}}(\mathbf{f}_{\ell}) &\simeq \sum_{k=1}^{K_{\ell}} \log_2\left(\mathsf{SILNR}_{\ell,k}(\mathbf{f}_{\ell})\right) \\
&= \log_2\left(\frac{\prod_{k=1}^{K_{\ell}}\mathbf{f}_{\ell}^{\mathsf{H}}\mathbf{S}_{\ell,\ell,k}\mathbf{f}_{\ell}}{\prod_{k=1}^{K_{\ell}}\mathbf{L}_{\ell,\bar{\ell}}(\mathbf{f}_{\ell})^{\frac{K_{\bar{\ell}}}{K_{\ell}}}}\right) = \log_2\left(\frac{\prod_{k=1}^{K_{\ell}}\mathbf{f}_{\ell}^{\mathsf{H}}\mathbf{S}_{\ell,\ell,k}\mathbf{f}_{\ell}}{\prod_{j=1}^{K_{\ell}}\mathbf{f}_{\ell}^{\mathsf{H}}\mathbf{C}_{\ell,\bar{\ell},j}\mathbf{f}_{\ell}}\right).
\end{aligned}
\tag{23}
$$

It is remarkable that (22a) and (23) are identical. We conclude that the proposed distributed precoding using local CSIT asymptotically achieves the cooperative bound in the high SINR regime under the zero-IUI condition. This equivalence comes from our SILNR definition using the geometric mean of the interference leakage terms.

Case 2 (Zero-ICI condition): Let us consider the zero-ICI constraint in both sum-spectral efficiency maximization problems in (15) and (18). When $N_{\ell} > K_{\bar{\ell}}$, it is possible to meet the zero-ICI condition by constructing precoding vectors on the nullspace of the column space spanned by the ICI channels. In this manner, we begin by reformulating the problem in (15) under zero-ICI constraint as

$$
\underset{\mathbf{f}_{\ell}\in\mathbb{C}^{N_{\ell}K_{\ell}}}{\arg\max} \sum_{\ell=1}^{2}\sum_{k=1}^{K_{\ell}} \log_2\left(1 + \mathsf{SINR}_{\ell,k}(\mathbf{f}_1,\mathbf{f}_2)\right),
\tag{24a}
$$

$$
\text{subject to} \quad \mathbf{f}_{\ell}^{\mathsf{H}}\mathbf{C}_{\ell,\bar{\ell},k}\mathbf{f}_{\ell} = 0, \quad \forall \ell \in \{1,2\}, \forall k \in \mathcal{K}_{\ell},
\tag{24b}
$$

$$
\|\mathbf{f}_{\ell}\|_2^2 = 1, \quad \forall \ell \in \{1,2\}.
\tag{24c}
$$

Under the zero-ICI constraint, the sum-spectral efficiency

simplifies

$$\sum_{\ell=1}^{2}\sum_{k=1}^{K_\ell}\log_2\left(1+\mathsf{SINR}_{\ell,k}(\mathbf{f}_1,\mathbf{f}_2)\right)$$

$$=\log_2\left(\prod_{k=1}^{K_1}\frac{\mathbf{f}_1^\mathsf{H}\left(\mathbf{S}_{1,1,k}+\mathbf{U}_{1,1,k}\right)\mathbf{f}_1}{\mathbf{f}_1^\mathsf{H}\mathbf{U}_{1,1,k}\mathbf{f}_1}\right)$$

$$+\log_2\left(\prod_{k=1}^{K_2}\frac{\mathbf{f}_2^\mathsf{H}\left(\mathbf{S}_{2,2,k}+\mathbf{U}_{2,2,k}\right)\mathbf{f}_2}{\mathbf{f}_2^\mathsf{H}\mathbf{U}_{2,2,k}\mathbf{f}_2}\right).\qquad(25)$$

This zero-ICI condition also makes the joint precoding design problem separable as

$$\arg\max_{\mathbf{f}_\ell\in\mathbb{C}^{N_\ell K_\ell}}\log_2\left(\prod_{k=1}^{K_\ell}\frac{\mathbf{f}_\ell^\mathsf{H}\left(\mathbf{S}_{\ell,\ell,k}+\mathbf{U}_{\ell,\ell,k}\right)\mathbf{f}_\ell}{\mathbf{f}_\ell^\mathsf{H}\mathbf{U}_{\ell,\ell,k}\mathbf{f}_\ell}\right),\qquad(26a)$$

$$\text{subject to}\quad\mathbf{f}_\ell^\mathsf{H}\mathbf{C}_{\ell,\bar{\ell},k}\mathbf{f}_\ell=0,\quad\forall k\in\mathcal{K}_\ell,\qquad(26b)$$

$$\|\mathbf{f}_\ell\|_2^2=1.\qquad(26c)$$

From (18), when $\mathbf{L}_{\ell,\bar{\ell}}(\mathbf{f}_\ell)=0$, our objective function becomes

$$\hat{R}_\ell^{\mathsf{sum}}(\mathbf{f}_\ell)=\log_2\left(\prod_{k=1}^{K_\ell}\frac{\mathbf{f}_\ell^\mathsf{H}\left(\mathbf{S}_{\ell,\ell,k}+\mathbf{U}_{\ell,\ell,k}\right)\mathbf{f}_\ell}{\mathbf{f}_\ell^\mathsf{H}\mathbf{U}_{\ell,\ell,k}\mathbf{f}_\ell}\right).\qquad(27)$$

This result implies that our precoding strategy exploiting local CSIT is sufficient to achieve the multi-cell cooperation bound very closely under either zero-IUI or zero-ICI constraints. This result, however, holds for the two-cell scenario only. In the sequel, we generalize this idea in a general multi-cell scenario and noisy CSIT assumption, which are more practically relevant.

IV. MULTI-CELL SILNR MAXIMIZATION PRECODING WITH NOISY CSIT

This section generalizes the idea of SILNR maximization precoding introduced in the previous section to a multi-cell noisy CSIT scenario. We start with formulating the SILNR maximization problem under noisy CSIT setting. Then, we establish the first- and second-order optimality condition for the problem. Then, a computationally-efficient algorithm to find a local-optimal solution to the problem is presented.

A. Problem Formulation

Using noisy and local CSIT at the ℓth BS, the received signal of the kth user in the ℓth cell is rewritten as

$$y_{\ell,k}[t]=\hat{\mathbf{h}}_{\ell,\ell,k}^\mathsf{H}\mathbf{f}_{\ell,k}s_{\ell,k}[t]+\sum_{i\neq k}^{K_\ell}\hat{\mathbf{h}}_{\ell,\ell,k}^\mathsf{H}\mathbf{f}_{\ell,i}s_{\ell,i}[t]$$

$$+\sum_{i=1}^{K_\ell}\mathbf{e}_{\ell,\ell,k}^\mathsf{H}\mathbf{f}_{\ell,i}s_{\ell,i}[t]+\tilde{n}_{\ell,k}[t],\qquad(28)$$

where $\tilde{n}_{\ell,k}[t]=\sum_{j\neq\ell}\sum_{k=1}^{K_j}\mathbf{h}_{j,\ell,k}^\mathsf{H}\mathbf{f}_{j,k}s_{j,k}[t]+n_{\ell,k}[t]$ is the effective noise when treating all aggregated ICI as additional noise and distributed by $\tilde{n}_{\ell,k}[t]\sim\mathcal{CN}\left(0,\tilde{\sigma}_{\ell,k}^2\right)$, where $\tilde{\sigma}_{\ell,k}^2=\mathbb{E}[\sum_{j\neq\ell}\sum_{k=1}^{K_j}|\mathbf{h}_{j,\ell,k}^\mathsf{H}\mathbf{f}_{j,k}s_{j,k}[t]|^2]+\sigma^2$. We define the

geometric mean of the interference leakage in (16) for a multi-cell scenario. Let $\mathcal{U}_\ell$ be the collection of the other cell's users who use the orthogonal pilots with the users in the ℓth cell, i.e., $\mathcal{U}_\ell=\left\{(\bar{\ell},j)\mid\forall\bar{\ell}\in\mathcal{C}_\ell\setminus\{\ell\},\forall j\in\mathcal{K}_{\bar{\ell}}\right\}$. Then, the leakage interference $\hat{\mathbf{L}}_\ell(\mathbf{f}_\ell)$ becomes

$$\hat{\mathbf{L}}_\ell(\mathbf{f}_\ell)=\left(\prod_{(\bar{\ell},j)\in\mathcal{U}_\ell}\mathbf{f}_\ell^\mathsf{H}\hat{\mathbf{C}}_{\ell,\bar{\ell},j}\mathbf{f}_\ell\right)^{\frac{1}{|\mathcal{U}_\ell|}},\qquad(29)$$

where $\hat{\mathbf{C}}_{\ell,\bar{\ell},j}=\mathbf{I}_{K_\ell}\otimes\frac{P_{\bar{\ell}}}{P_\ell}\left(\hat{\mathbf{h}}_{\ell,\bar{\ell},j}\hat{\mathbf{h}}_{\ell,\bar{\ell},j}^\mathsf{H}+\mathbf{\Phi}_{\ell,\bar{\ell},j}\right)\in\mathbb{C}^{N_\ell K_\ell\times N_\ell K_\ell}$. Incorporating (28) and (29), we define SILNR of the kth user in the ℓth cell as

$$\mathsf{SILNR}_{\ell,k}(\mathbf{f}_\ell)$$

$$=\frac{\left|\hat{\mathbf{h}}_{\ell,\ell,k}^\mathsf{H}\mathbf{f}_{\ell,k}\right|^2}{\sum_{i\neq k}^{K_\ell}\left|\hat{\mathbf{h}}_{\ell,\ell,k}^\mathsf{H}\mathbf{f}_{\ell,i}\right|^2+\sum_{i=1}^{K_\ell}\mathbf{f}_{\ell,i}^\mathsf{H}\mathbf{\Phi}_{\ell,\ell,k}\mathbf{f}_{\ell,i}+\hat{\mathbf{L}}_\ell(\mathbf{f}_\ell)^{\frac{|\mathcal{U}_\ell|}{K_\ell}}+\frac{\tilde{\sigma}_{\ell,k}^2}{P_\ell}}$$

$$=\frac{\mathbf{f}_\ell^\mathsf{H}\hat{\mathbf{S}}_{\ell,\ell,k}\mathbf{f}_\ell}{\mathbf{f}_\ell^\mathsf{H}\hat{\mathbf{U}}_{\ell,\ell,k}\mathbf{f}_\ell+\mathbf{f}_\ell^\mathsf{H}\mathbf{E}_{\ell,\ell,k}\mathbf{f}_\ell+\hat{\mathbf{L}}_\ell(\mathbf{f}_\ell)^{\frac{|\mathcal{U}_\ell|}{K_\ell}}},\qquad(30)$$

where

$$\hat{\mathbf{S}}_{\ell,\ell,k}=\mathbf{e}_k\mathbf{e}_k^\mathsf{T}\otimes\hat{\mathbf{h}}_{\ell,\ell,k}\hat{\mathbf{h}}_{\ell,\ell,k}^\mathsf{H}\in\mathbb{C}^{N_\ell K_\ell\times N_\ell K_\ell},$$

$$\hat{\mathbf{U}}_{\ell,\ell,k}=\mathbf{I}_{K_\ell}\otimes\hat{\mathbf{h}}_{\ell,\ell,k}\hat{\mathbf{h}}_{\ell,\ell,k}^\mathsf{H}-\hat{\mathbf{S}}_{\ell,\ell,k}+\frac{\tilde{\sigma}_{\ell,k}^2}{P_\ell}\mathbf{I}_{N_\ell K_\ell}\in\mathbb{C}^{N_\ell K_\ell\times N_\ell K_\ell},$$

$$\mathbf{E}_{\ell,\ell,k}=\mathbf{I}_{K_\ell}\otimes\mathbf{\Phi}_{\ell,\ell,k}\in\mathbb{C}^{N_\ell K_\ell\times N_\ell K_\ell}.\qquad(31)$$

Accordingly, our precoding strategy using noisy and local CSIT is to solve the following optimization problem:

$$\arg\max_{\mathbf{f}_\ell\in\mathbb{C}^{N_\ell K_\ell}}\prod_{k=1}^{K_\ell}\frac{\mathbf{f}_\ell^\mathsf{H}\left(\hat{\mathbf{S}}_{\ell,\ell,k}+\hat{\mathbf{U}}_{\ell,\ell,k}+\mathbf{E}_{\ell,\ell,k}+\hat{\mathbf{L}}_\ell(\mathbf{f}_\ell)^{\frac{|\mathcal{U}_\ell|}{K_\ell}}\right)\mathbf{f}_\ell}{\mathbf{f}_\ell^\mathsf{H}\left(\hat{\mathbf{U}}_{\ell,\ell,k}+\mathbf{E}_{\ell,\ell,k}+\hat{\mathbf{L}}_\ell(\mathbf{f}_\ell)^{\frac{|\mathcal{U}_\ell|}{K_\ell}}\right)\mathbf{f}_\ell},\qquad(32a)$$

$$\text{subject to}\quad\|\mathbf{f}_\ell\|_2^2=1,\quad\forall\ell\in\mathcal{L}.\qquad(32b)$$

Since the objective function (32) is highly non-convex, finding even local-optimal solution for $\mathbf{f}_\ell$ is a very challenging task. In the sequel, we derive the first- and second-order optimality conditions for this non-convex optimization problem.

B. Local Optimality Conditions

The following theorems establish the first- and the second-order necessary conditions for the local optimality of the non-convex optimization problem in (32).

Theorem 1. (The first-order necessary condition) *If* $\mathbf{f}_\ell^\star\in\mathbb{C}^{N_\ell K_\ell\times 1}$ *is a stationary point of the non-convex optimization problem* (32)*, it satisfies*

$$\bar{\mathbf{A}}_\ell\left(\mathbf{f}_\ell^\star\right)\mathbf{f}_\ell^\star=\gamma\left(\mathbf{f}_\ell^\star\right)\bar{\mathbf{B}}_\ell\left(\mathbf{f}_\ell^\star,\lambda\right)\mathbf{f}_\ell^\star,\qquad(33)$$

where the functional matrices $\bar{\mathbf{A}}_\ell\left(\mathbf{f}_\ell^\star\right)$ *and* $\bar{\mathbf{B}}_\ell\left(\mathbf{f}_\ell^\star,\lambda\right)$ *are*

$$\tilde{\mathbf{A}}_{\ell,\ell,k}=\hat{\mathbf{S}}_{\ell,\ell,k}+\hat{\mathbf{U}}_{\ell,\ell,k}+\mathbf{E}_{\ell,\ell,k},$$

$$\tilde{\mathbf{B}}_{\ell,\ell,k}=\hat{\mathbf{U}}_{\ell,\ell,k}+\mathbf{E}_{\ell,\ell,k},$$

$$\bar{\mathbf{A}}_\ell\left(\mathbf{f}_\ell\right) = \sum_{i=1}^{K_\ell} \left(\prod_{k \neq i}^{K_\ell} (\mathbf{f}_\ell)^{\mathsf{H}} \left(\tilde{\mathbf{A}}_{\ell,\ell,k} + \hat{\mathbf{L}}_\ell(\mathbf{f}_\ell)^{\frac{|u_\ell|}{K_\ell}} \right) \mathbf{f}_\ell \right)$$
$$\times \left(\tilde{\mathbf{A}}_{\ell,\ell,i} + \tilde{\mathbf{C}}_{\ell,\ell,k} \right) \mathbf{f}_\ell,$$

$$\bar{\mathbf{B}}_\ell\left(\mathbf{f}_\ell, \lambda\right) = \left[\sum_{i=1}^{K_\ell} \left(\prod_{k \neq i}^{K_\ell} (\mathbf{f}_\ell)^{\mathsf{H}} \left(\tilde{\mathbf{B}}_{\ell,\ell,k} + \hat{\mathbf{L}}_\ell(\mathbf{f}_\ell)^{\frac{|u_\ell|}{K_\ell}} \right) \mathbf{f}_\ell \right) \right.$$
$$\left. \times \left(\tilde{\mathbf{B}}_{\ell,\ell,\mathbf{i}} + \tilde{\mathbf{C}}_{\ell,\ell,k} \right) + \frac{\lambda}{\gamma(\mathbf{f}_\ell)} \mathbf{I}_{N_\ell K_\ell} \right] \mathbf{f}_\ell,$$

$$\tilde{\mathbf{C}}_{\ell,\ell,k} = \sum_{(\bar{\ell},j) \in \mathcal{U}_\ell} \frac{\hat{\mathbf{L}}_\ell(\mathbf{f}_\ell)^{\frac{|u_\ell|}{K_\ell}} \hat{\mathbf{C}}_{\ell,\bar{\ell},j}}{K_\ell(\mathbf{f}_\ell)^{\mathsf{H}} \hat{\mathbf{C}}_{\ell,\bar{\ell},j} \mathbf{f}_\ell},$$

$$\gamma(\mathbf{f}_\ell) = \prod_{k=1}^{K_\ell} \frac{\mathbf{f}_\ell^{\mathsf{H}} \left(\hat{\mathbf{S}}_{\ell,\ell,k} + \hat{\mathbf{U}}_{\ell,\ell,k} + \mathbf{E}_{\ell,\ell,k} + \hat{\mathbf{L}}_\ell(\mathbf{f}_\ell)^{\frac{|u_\ell|}{K_\ell}} \right) \mathbf{f}_\ell}{\mathbf{f}_\ell^{\mathsf{H}} \left(\hat{\mathbf{U}}_{\ell,\ell,k} + \mathbf{E}_{\ell,\ell,k} + \hat{\mathbf{L}}_\ell(\mathbf{f}_\ell)^{\frac{|u_\ell|}{K_\ell}} \right) \mathbf{f}_\ell}. \quad (34)$$

In addition, the Lagrange multiplier λ is chosen so that $\mathbf{f}$ satisfies $\|\mathbf{f}_\ell\|_2^2 = 1$.

  Proof: See Appendix A. ∎

Theorem 1 implies that any stationary point of the non-convex problem in (32) is one of the eigenvectors of the functional matrix $[\bar{\mathbf{B}}_\ell\left(\mathbf{f}_\ell, \lambda\right)]^{-1} \bar{\mathbf{A}}_\ell$, i.e.,

$$[\bar{\mathbf{B}}_\ell\left(\mathbf{f}_\ell, \lambda\right)]^{-1} \bar{\mathbf{A}}_\ell\left(\mathbf{f}_\ell\right) \mathbf{f}_\ell = \gamma\left(\mathbf{f}_\ell\right) \mathbf{f}_\ell. \quad (35)$$

As can be seen, the objective function $\gamma\left(\mathbf{f}_\ell\right)$ can be interpreted as the eigenvalue of the functional matrix $[\bar{\mathbf{B}}_\ell\left(\mathbf{f}_\ell, \lambda\right)]^{-1} \bar{\mathbf{A}}_\ell$. Since we are interested in maximizing the objective function $\gamma\left(\mathbf{f}_\ell\right)$, we need to identify the eigenvector corresponding to the maximum eigenvalue, which can be a global optimal solution. Unfortunately, finding such eigenvector is highly non-trivial, because $[\bar{\mathbf{B}}_\ell\left(\mathbf{f}_\ell, \lambda\right)]^{-1} \bar{\mathbf{A}}_\ell$ is a function of $\mathbf{f}_\ell$ and Lagrange multiplier λ. Nevertheless, this principal component analysis helps to understand the global landscape of the non-convex optimization problem; and leads to an algorithm to find a local-optimal solution in a numerically efficient manner.

Although $\mathbf{f}_\ell^\star$ satisfies the first-order necessary condition derived in Theorem 1, we need to check the curvature of the objective function around the stationary point to verify the local optimality. The following theorem gives a testing condition of the negative definiteness of the extended Hessian matrix evaluated at $\mathbf{f}_\ell^\star$ in a closed-form, i.e., $\nabla^2_{\mathbf{f}_\ell^{\mathsf{H}}} \gamma(\mathbf{f}_\ell^\star) < 0$.

Theorem 2. (The second-order necessary condition) *The stationary point $\mathbf{f}_\ell^\star$ is a local-optimal solution, provided that (36), where the functional matrices are defined as $\mathbf{Z}_{\ell,\ell,i} = \sum_{(m,n) \in \mathcal{U}_\ell} \frac{\hat{\mathbf{L}}_\ell(\mathbf{f}_\ell)^{\frac{|u_\ell|}{K_\ell}} \hat{\mathbf{C}}_{\ell,m,n} \mathbf{f}_\ell \mathbf{f}_\ell^{\mathsf{H}} \hat{\mathbf{C}}_{\ell,m,n}}{K_\ell(\mathbf{f}_\ell^{\mathsf{H}} \hat{\mathbf{C}}_{\ell,m,n} \mathbf{f}_\ell)^2}$, $\mathbf{X}^A_{\ell,\ell,i} = \left(\tilde{\mathbf{A}}_{\ell,\ell,i} + \tilde{\mathbf{C}}_{\ell,\ell,i} \right)$, $\mathbf{X}^B_{\ell,\ell,i} = \left(\tilde{\mathbf{B}}_{\ell,\ell,i} + \tilde{\mathbf{C}}_{\ell,\ell,i} \right)$, $\mathbf{Y}^A_{\ell,\ell,i} = \tilde{\mathbf{A}}_{\ell,\ell,i} + \hat{\mathbf{L}}_\ell\left(\mathbf{f}_\ell\right)^{\frac{|u_\ell|}{K_\ell}}$ and $\mathbf{Y}_B = \tilde{\mathbf{B}}_{\ell,\ell,i} + \hat{\mathbf{L}}_\ell\left(\mathbf{f}_\ell\right)^{\frac{|u_\ell|}{K_\ell}}$.*

  Proof: See Appendix B. ∎

Theorem 2 is useful when evaluating the second-order optimality condition, because the direct computation of Hes-

---

Algorithm 1: SILNR Maximization Precoding

Initialization:
  $t = n = 0$, $\mathbf{f}_\ell^{(0)} = \mathsf{ZF}$, $\mathbf{f}_\ell^{(-1)} = \mathbf{0}$, $\lambda^{(0)}$, and ϵ
  while $\|\mathbf{f}_\ell\|_2^2 > 1$ **do**
    $n \leftarrow n + 1$
    $\lambda^{(n)} \leftarrow \lambda^{(n-1)} + \Delta\lambda^{(n)}$
    while $\|\mathbf{f}_\ell^{(t-1)} - \mathbf{f}_\ell^{(t)}\|_2 \geq \epsilon$ **do**
      $t \leftarrow t + 1$
      $\mathbf{f}_\ell^{(t)} \leftarrow \frac{\mathbf{f}_\ell^{(t)}}{\|\mathbf{f}_\ell^{(t)}\|_2}$
      $\mathbf{f}_\ell^{(t)} \leftarrow \left[\bar{\mathbf{B}}_\ell\left(\mathbf{f}_\ell^{(t-1)}, \lambda^{(n)}\right) \right]^{-1} \bar{\mathbf{A}}_\ell\left(\mathbf{f}_\ell^{(t-1)}\right) \mathbf{f}_\ell^{(t-1)}$
    end
  end

---

sian matrix $\nabla^2_{\mathbf{f}_\ell^{\mathsf{H}}} \gamma(\mathbf{f}_\ell^\star)$ is unnecessary. Although Theorem 1 and Theorem 2 provide a guidance for the local optimality conditions for problem (32), finding such a precoding vector is still challenging. To resolve this issue, we propose a computationally-efficient algorithm in the sequel.

C. Algorithm

The proposed algorithm takes two steps: 1) the identification of a stationary point using the generalized power iteration (GPI) technique in [29] for a given Lagrange multiplier $\lambda^{(n)}$ and 2) the Lagrangian multiplier adjustment to be a feasible solution for a given $\mathbf{f}_\ell^{(t)}$. The algorithm starts with an initial precoding solution $\mathbf{f}_\ell^{(0)}$, which is typically chosen as the ZF beamforming method. In the tth iteration, the algorithm evaluates the functional matrices $\bar{\mathbf{A}}_\ell\left(\mathbf{f}_\ell^{(t-1)}\right)$ and $\bar{\mathbf{B}}_\ell\left(\mathbf{f}_\ell^{(t-1)}, \lambda^{(n)}\right)$ defined in (34) using the identified precoding vector in the previous iteration, i.e., $\mathbf{f}_\ell^{(t-1)}$. Then, it is multiplied with $\left[\bar{\mathbf{B}}_\ell\left(\mathbf{f}_\ell^{(t-1)}, \lambda^{(n)}\right) \right]^{-1} \bar{\mathbf{A}}_\ell\left(\mathbf{f}_\ell^{(t-1)}\right)$ to obtain $\mathbf{f}_\ell^{(t)} = \left[\bar{\mathbf{B}}_\ell\left(\mathbf{f}_\ell^{(t-1)}, \lambda^{(n)}\right) \right]^{-1} \bar{\mathbf{A}}_\ell\left(\mathbf{f}_\ell^{(t-1)}\right) \mathbf{f}_\ell^{(t-1)}$. The iteration continues until a stopping condition $\|\mathbf{f}_\ell^{(t-1)} - \mathbf{f}_\ell^{(t)}\|_2 \leq \epsilon$ holds, where ϵ is selected as a small positive number. Once the eigenvector $\mathbf{f}_\ell$ that satisfies Theorem 1 is identified, i.e., the inner loop of Algorithm 1 is converged, Algorithm 1 checks whether the obtained solution is a feasible point or not. If the obtained solution exceeds the power constraint, Algorithm 1 updates the Lagrange multiplier to meet the constraint. The algorithm proceeds until it finds the feasible solution that satisfies the first-order optimality condition.

Remark 1 (The computational complexity of the algorithm): The computational complexity order of the proposed SILNR maximization precoding algorithm is $O\left(J N_\ell^2 K_\ell\right)$, where J is the number of iterations of Algorithm 1. We refer the details of the computational complexity analysis in [29], in which the matrix inverse and the multiplication operations have shown to be performed in a divide and conquer manner by exploiting the block diagonal structure in $\bar{\mathbf{A}}_\ell\left(\mathbf{f}_\ell\right)$ and $\bar{\mathbf{B}}_\ell\left(\mathbf{f}_\ell, \lambda\right)$. For the convergence, the number of iterations for

$$\lambda_{\max}\left(\left[\sum_{i=1}^{K_\ell}\frac{\mathbf{X}_{\ell,\ell,i}^A\mathbf{f}_\ell}{\mathbf{f}_\ell^{\mathsf{H}}\mathbf{Y}_{\ell,\ell,i}^A\mathbf{f}_\ell}-\frac{\mathbf{X}_{\ell,\ell,i}^B\mathbf{f}_\ell}{\mathbf{f}_\ell^{\mathsf{H}}\mathbf{Y}_B\mathbf{f}_\ell}\right]\left[\sum_{i=1}^{K_\ell}\frac{\mathbf{X}_{\ell,\ell,i}^A\mathbf{f}_\ell}{\mathbf{f}_\ell^{\mathsf{H}}\mathbf{Y}_{\ell,\ell,i}^A\mathbf{f}_\ell}-\frac{\mathbf{X}_{\ell,\ell,i}^B\mathbf{f}_\ell}{\mathbf{f}_\ell^{\mathsf{H}}\mathbf{Y}_B\mathbf{f}_\ell}\right]^{\mathsf{H}}+\gamma(\mathbf{f}_\ell)\left[\sum_{i=1}^{K_\ell}\frac{\mathbf{X}_{\ell,\ell,i}^A}{\mathbf{f}_\ell^{\mathsf{H}}\mathbf{Y}_{\ell,\ell,i}^A\mathbf{f}_\ell}+\frac{\mathbf{Z}_{\ell,\ell,i}}{\mathbf{f}_\ell^{\mathsf{H}}\mathbf{Y}_{\ell,\ell,i}^B\mathbf{f}_\ell}+\frac{\mathbf{X}_{\ell,\ell,i}^B\mathbf{f}_\ell\mathbf{f}_\ell^{\mathsf{H}}\mathbf{X}_{\ell,\ell,i}^B}{\left(\mathbf{f}_\ell^{\mathsf{H}}\mathbf{Y}_{\ell,\ell,i}^B\mathbf{f}_\ell\right)^2}\right]\right)$$

$$<\lambda_{\min}\left(\gamma(\mathbf{f}_\ell)\left[\sum_{i=1}^{K_\ell}\frac{\mathbf{X}_{\ell,\ell,i}^B}{\mathbf{f}_\ell^{\mathsf{H}}\mathbf{Y}_{\ell,\ell,i}^B\mathbf{f}_\ell}+\frac{\mathbf{Z}_{\ell,\ell,i}}{\mathbf{f}_\ell^{\mathsf{H}}\mathbf{Y}_{\ell,\ell,i}^A\mathbf{f}_\ell}+\frac{\mathbf{X}_{\ell,\ell,i}^A\mathbf{f}_\ell\mathbf{f}_\ell^{\mathsf{H}}\mathbf{X}_{\ell,\ell,i}^A}{\left(\mathbf{f}_\ell^{\mathsf{H}}\mathbf{Y}_{\ell,\ell,i}^A\mathbf{f}_\ell\right)^2}\right]+\lambda\mathbf{I}\right). \tag{36}$$

the inner loop is five at most in an average sense regarding random channels.

Remark 2 (The second-order optimality condition): From numerical results, we observe that, in every case, the solution obtained from the proposed algorithm satisfies the second-order optimality condition derived in Theorem 2. We conjecture that the proposed algorithm ensures converging to a local-optimal solution.

V. SIMULATION RESULTS

This section provides both the link-level and system-level simulation results to gauge the ergodic sum-spectral efficiency gains of the proposed SILNR maximization precoding compared to the existing precoding schemes.

A. Link-Level Simulations

We consider a two-cell scenario in which each BS equipped with N_ℓ ($=16$) antennas serves single antenna eight users [35]. For link-level simulations, we assume that all channel vectors are drawn from $\mathbf{h}_{j,\ell,k}\sim\mathcal{CN}(0,\mathbf{I}_{N_\ell})$, and each BS has perfect knowledge of local CSIT. In this setting, we compare the ergodic sum-spectral efficiency for different precoding strategies:

- Multi-Cell Coop [28], [29]: This scheme is a cooperative precoding technique that maximizes the sum-spectral efficiency of users in the cooperative area. We use this precoding method as a benchmark for an upper bound of our proposed noncooperative precoding technique.
- Multi-Cell MMSE (or Multi-Cell SLNR) [24]:

$$\arg\max_{\mathbf{f}_{\ell,k}\in\mathbb{C}^{N_\ell}}\frac{\left|\hat{\mathbf{h}}_{\ell,\ell,k}^{\mathsf{H}}\mathbf{f}_{\ell,k}\right|^2}{\mathbf{f}_{\ell,k}^{\mathsf{H}}\left(\sum_{(j,i)\in\mathcal{U}_\ell}\hat{\mathbf{h}}_{\ell,j,i}\hat{\mathbf{h}}_{\ell,j,i}^{\mathsf{H}}+\frac{\bar{\sigma}_{\ell,k}^2}{P_\ell}\mathbf{I}_{N_\ell K_\ell}\right)\mathbf{f}_{\ell,k}}. \tag{37}$$

This scheme uses the downlink precoding solution in (37) by the uplink-downlink duality [25], [36], [37].
- Multi-Cell ZF [20]: This scheme eliminates both IUI and ICI.

Validation of quasi-optimality: As shown in Fig. 2-(b), the proposed algorithm tightly achieves the performance of the multi-cell cooperative precoding scheme. Our simulation result implies that our precoding solution makes the zero-IUI phenomenon occur, even if N_ℓ/K_ℓ is not-so-large.

Ergodic sum-spectral efficiency performance: Fig. 2-(a) shows how the ergodic sum-spectral efficiency changes with
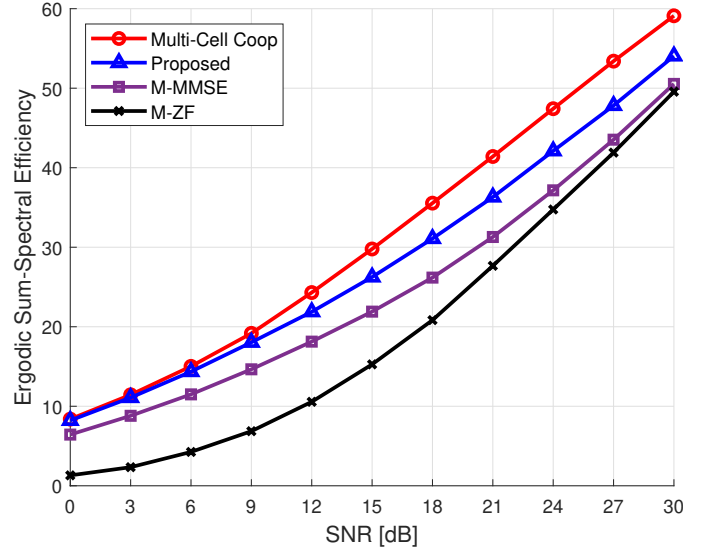
(a)
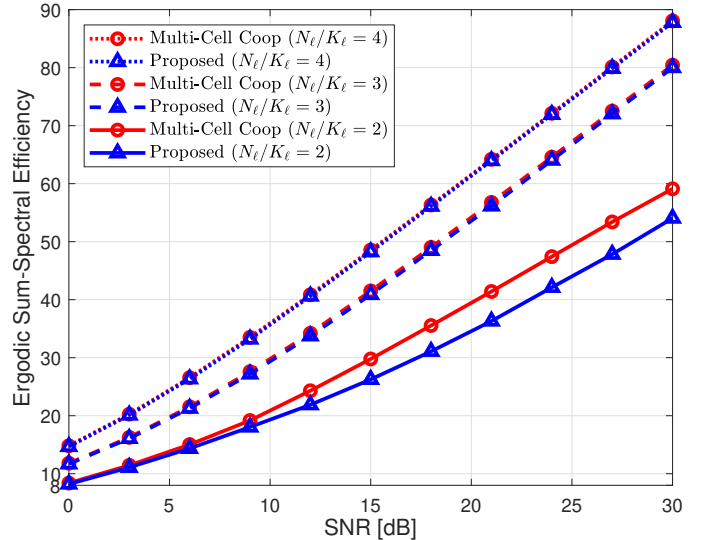
(b)

Fig. 2. (a) The ergodic sum-spectral efficiencies for different precoding strategies and (b) the ergodic sum-spectral efficiencies when increasing the ratio of the number of antennas to the number of users without considering the zero-IUI condition.

increasing SNRs. The proposed precoding yields considerable gains compared to the existing precoding methods in all SNRs. This performance improvement comes from better utilization of spatial degrees of freedom (DoF) to mitigate both IUI and ICI. Specifically, the proposed SILNR jointly finds a set of

TABLE I
COMPLEXITY ORDER ANALYSIS.

List	Computation complexity order
Multi-Cell MMSE	$O\left((\|\mathcal{U}_\ell\| + \|\mathcal{K}_\ell\|)^2 N_\ell \right.$ $\left. + (\|\mathcal{U}_\ell\| + \|\mathcal{K}_\ell\|)^3\right)$
Computation of $\bar{\mathbf{A}}_\ell(\mathbf{f}_\ell^\star)$	$O\left(N_\ell^2 K_\ell\right)$
Computation of $\bar{\mathbf{B}}_\ell(\mathbf{f}_\ell^\star, \lambda)$	$O\left(N_\ell^2 K_\ell\right)$
Computation of $\left[\bar{\mathbf{B}}_\ell(\mathbf{f}_\ell^\star, \lambda)\right]^{-1}$	$O\left(N_\ell^2 K_\ell\right)$
Computation of $\left[\bar{\mathbf{B}}_\ell(\mathbf{f}_\ell^\star, \lambda)\right]^{-1} \bar{\mathbf{A}}_\ell(\mathbf{f}_\ell^\star)\mathbf{f}_\ell^\star$	$O\left(N_\ell^2 K_\ell\right)$
Proposed algorithm	$O\left(J N_\ell^2 K_\ell\right)$

TABLE II
PARAMETERS FOR SYSTEM-LEVEL SIMULATIONS.

Parameters	Value
Topology of BS	28 BSs over 7 hexagonal coverages
# of BSs; (MBS, PBS)	(7,21)
Topology of user	Uniformly distributed per cell
# of UEs per (MBS, PBS)	(16,4)
Bandwidth	20 MHz
Carrier frequency	2 GHz
MBS transmission power	46 dBm
PBS transmission power	23 dBm
Noise power	-113 dB
Spatial channel model	Spatially correlated model
Path-loss model	Okumaura-Hata urban model
BS and user height	32 m/1.5 m
Channel estimation	MMSE estimator in (3)
Channel coherence time slot	$\tau_{\mathrm{c}} = 200$
Uplink training slot	$\tau_{\mathrm{u}} = 28$ and $\tau_{\mathrm{d}} = 0$
Stopping condition	$\epsilon = 0.1$

served users, precoding vectors, and power allocation by optimally balancing IUI and the geometric mean of interference leakage signals. Therefore, it allows each BS to exploit the spatial DoF in a more efficient way to increase the sum-spectral efficiency of the desired cell, while simultaneously reducing leakage interference power towards the other cells.

Complexity order analysis: We provided the computational complexity analysis of the proposed precoding method, which is summarized in Table I. The computational complexity of the proposed precoding method increases in the order of $O\left(N_\ell^2 K_\ell\right)$. To accomplish this analysis, we need to calculate the complexity order for computing functional matrices $\bar{\mathbf{A}}_\ell(\mathbf{f}_\ell^\star) \in \mathbb{C}^{N_\ell K_\ell \times N_\ell K_\ell}$ and $\bar{\mathbf{B}}_\ell(\mathbf{f}_\ell^\star, \lambda) \in \mathbb{C}^{N_\ell K_\ell \times N_\ell K_\ell}$. Exploiting the block diagonal structure of them, one can readily show that the complexity order becomes $O\left(N_\ell^2 K_\ell\right)$. In addition, using the fact that $\bar{\mathbf{B}}_\ell(\mathbf{f}_\ell^\star, \lambda)$ is the sum of rank-one matrices, one can compute the inverse matrix of it successively using Sherman-Morrison Lemma (or also known as Woodbury inverse) [38]. Thanks to the block diagonal structure, computing the inverse of $\bar{\mathbf{B}}_\ell(\mathbf{f}_\ell^\star, \lambda)$ requires the computational complexity order of $O\left(N_\ell^2 K_\ell\right)$. By embracing the number of iterations to converge, J, the total computational complexity of the proposed precoding algorithm is $O\left(J N_\ell^2 K_\ell\right)$.

B. System-Level Simulations

Fig. 3 depicts the network topology for system-level simulations, which consists of seven hexagonal cells, each macro
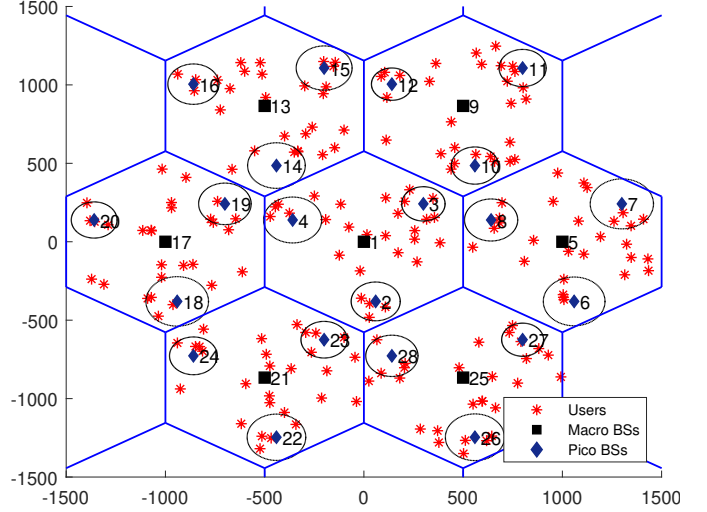

Fig. 3. A snapshots of the network topologies used in system-level simulations.

BS (MBS) coverage area contains three pico BSs (PBSs). See Table II for the details of simulation parameters [39].

For a fair comparison, we also consider the Single-Cell WMMSE [27] precoding method with a leakage level constraint. For this, we consider the following precoding optimization problem:

$$\arg\min_{\{\mathbf{f}_{\ell,k}, w_k, u_k \mid k \in \mathcal{K}_\ell\}} \sum_{k=1}^{K_\ell} (w_k e_k - \log w_k), \tag{38a}$$

$$\text{subject to} \quad \sum_{k=1}^{K_\ell} \|\mathbf{f}_{\ell,k}\|_2^2 \leq 1, \tag{38b}$$

$$\mathbf{f}_{\ell,k}^{\mathsf{H}} \left(\sum_{(m,n) \in \mathcal{U}_\ell} \mathbf{h}_{\ell,m,n} \mathbf{h}_{\ell,m,n}^{\mathsf{H}} \right) \mathbf{f}_{\ell,k} \leq r_k, \tag{38c}$$

where

$$e_k = u_k^{\mathsf{H}} \left(\sum_{i=1}^{K_\ell} \left| \mathbf{h}_{\ell,k}^{\mathsf{H}} \mathbf{f}_{\ell,i} \right|^2 + \sigma^2 \right) u_k - 2\mathsf{Re}\left\{ u_k^{\mathsf{H}} \mathbf{h}_{\ell,k}^{\mathsf{H}} \mathbf{f}_{\ell,k} \right\} + 1,$$

$$w_k = e_k^{-1},$$

$$u_k = \left(\sum_{i=1}^{K_\ell} \left| \mathbf{h}_{\ell,k}^{\mathsf{H}} \mathbf{f}_{\ell,i} \right|^2 + \sigma^2 \right)^{-1} \mathbf{h}_{\ell,k}^{\mathsf{H}} \mathbf{f}_{\ell,k}, \tag{39}$$

denote MSE, MSE weight, and the optimal receiver, respectively. Also, r_k dentoes a target leakage power level. The problem (38) is a combination of the conventional WMMSE optimization (38a), (38b) and leakage power constraints (38c).

Ergodic sum-spectral efficiency performance: We evaluate how the ergodic sum-spectral efficiency behaves when increasing the number of antennas of MBSs. Fig. 4 shows the ergodic sum-spectral efficiency performance under perfect local CSIT. As. can be seen, the proposed precoding method outperforms the existing precoding methods regardless of the number of antennas per MBS. We capitalize that the existing noncooperative precoding can achieve the upper bound when the number of antennas is larger than 50.

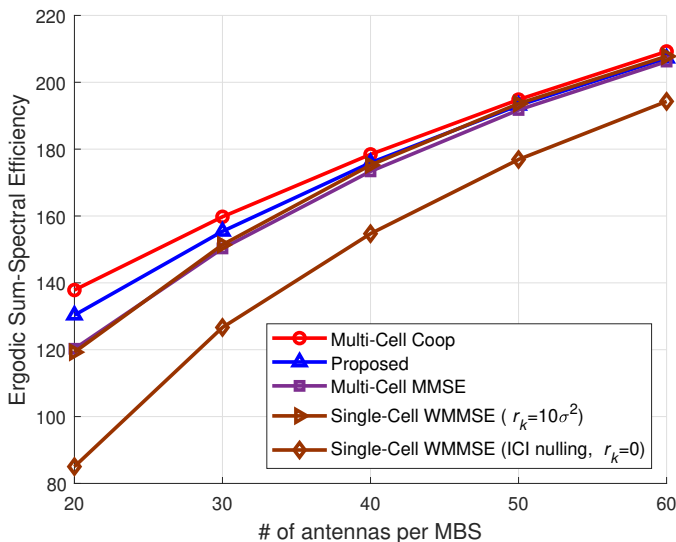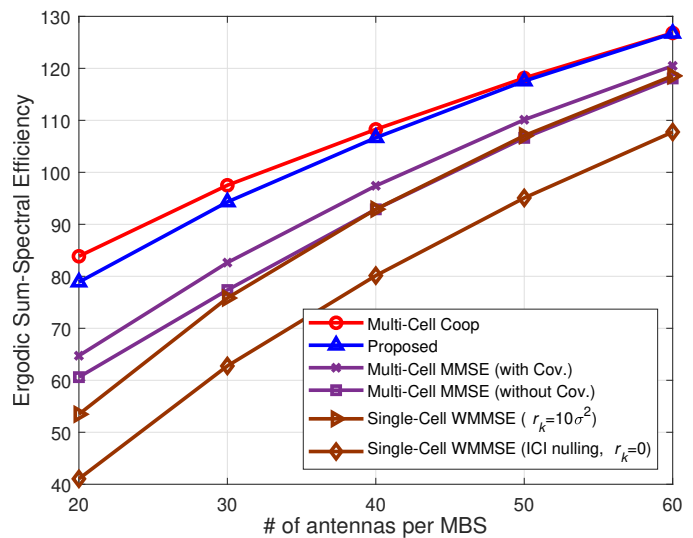Fig. 5 demonstrates the ergodic sum-spectral efficiency per-

Fig. 4. The ergodic sum-spectral efficiencies for different precoding strategies with local and perfect CSIT.

formance under the noisy CSIT setting. For a fair comparison, we compare the performance of the proposed algorithm with multi-cell MMSE exploiting the covariance matrix $\mathbf{\Phi}_{\ell,\ell,k}$ [40] (denoted by Muti-Cell MMSE (with Cov.)). As can be seen, the proposed algorithm obtains a substantial ergodic sum-spectral efficiency gain compared with other precoding methods in all antenna scales. This result implies that our proposed framework is robust to pilot contamination effects, a remarkable aspect of practical cellular systems.

Convergence speed of the proposed SILNR maximization precoding: Fig. 6 illustrates the convergence speed of the proposed algorithm for the SILNR maximization precoding. We consider the cases of $K = \{20, 40, 60\}$ and $N = 64$. We measure the mean square of the difference for the objective functions evaluated at two consecutive precoding solutions during iterations, i.e., $\mathbb{E}\left[\|\mathbf{f}_\ell^{(m)} - \mathbf{f}_\ell^{(m-1)}\|_2^2\right]$ where the average is taken over both the fading channel realizations and the user locations. As depicted in Fig. 6, the number of required iterations to find the solution is at most five in an average sense, when we set the solution accuracy parameter to $\epsilon = 0.1$. As improving the solution accuracy level to $\epsilon = 0.01$, ten iterations are sufficient to end the algorithm for all $K = \{20, 40, 60\}$ and $N = 64$. In addition, when the algorithm starts with the ZF precoding solution as an initial point, we empirically observe that the initially identified solution $\mathbf{f}^\star$ of Algorithm 1 has local optimality in the most of our simulations.

VI. CONCLUSION

In this paper, we presented a novel distributed precoding technique using local CSIT for MU-MIMO HetNets. The central idea of the proposed precoding method was to maximize the downlink sum-spectral efficiency per cell while mitigating the other cell interference leakage using local CSIT. We introduced a new metric called SILNR that measures the ratio between the desired signal power and the superposition of IUI and interference leakage powers towards the other cells.

(a)
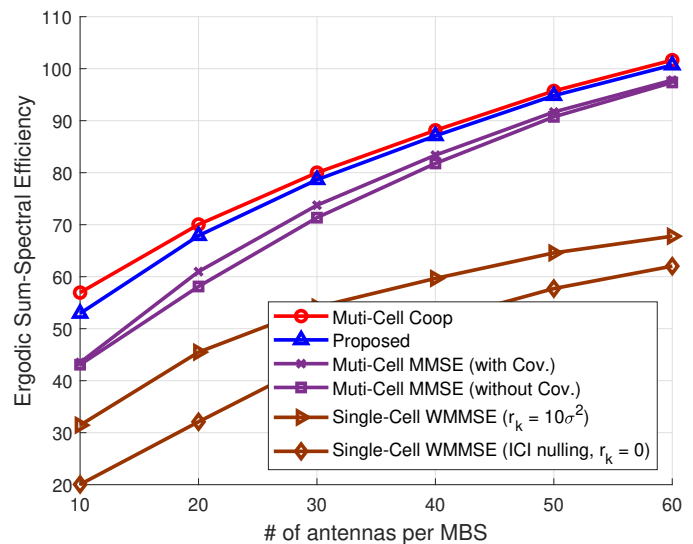
(b)

Fig. 5. The ergodic sum-spectral efficiencies for different precoding strategies with local and noisy CSIT when the number of UEs per (MBS,PBS) = (a) (16,4) and (b) (8,4).

Using this metric, we formulated a maximization problem of the product of SILNRs, which is a non-convex optimization problem. We derived the first- and the second-order necessary conditions for the local-optimal solution of this non-convex optimization problem. Leveraging these conditions, we presented a computationally efficient algorithm that ensures finding the local-optimal solution iteratively. Using both link-level and system-level simulations, we demonstrated our precoding method using local CSIT achieves the same sum-spectral efficiency of the cooperative precoding method that requires global CSIT when increasing the number of antennas BS, confirming that a synergetic gain is possible when massive MIMO meets HetNets.
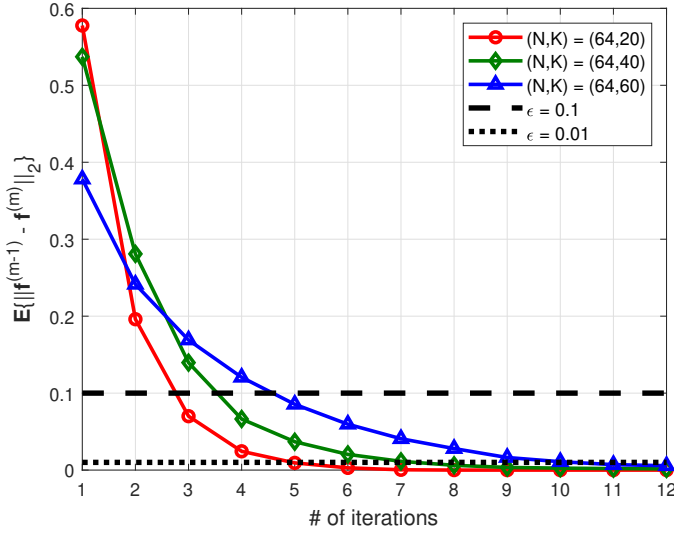
Fig. 6. Convergence speed illustration with $\epsilon = 0.1$ and $\epsilon = 0.01$.

APPENDIX

A. Proof for Theorem 1

We commence by defining the Lagrange function:

$$\mathcal{L}\left(\mathbf{f}_\ell, \lambda\right) = \prod_{k=1}^{K_\ell} \frac{\mathbf{f}_\ell^{\mathsf{H}} \left(\tilde{\mathbf{A}}_{\ell,\ell,k} + \hat{\mathbf{L}}_\ell\left(\mathbf{f}_\ell\right)^{\frac{|u_\ell|}{K_\ell}}\right) \mathbf{f}_\ell}{\mathbf{f}_\ell^{\mathsf{H}} \left(\tilde{\mathbf{B}}_{\ell,\ell,k} + \hat{\mathbf{L}}_\ell\left(\mathbf{f}_\ell\right)^{\frac{|u_\ell|}{K_\ell}}\right) \mathbf{f}_\ell} - \lambda\left(\|\mathbf{f}_\ell\|_2^2 - 1\right). \tag{40}$$

To find a stationary point, we take the partial derivatives of $\mathcal{L}(\mathbf{f}_\ell, \lambda)$ with respective to $\mathbf{f}_\ell^{\mathsf{H}}$ and λ, and set to them zero. When taking the derivative with respective to $\mathbf{f}_\ell^{\mathsf{H}}$, we obtain

$$\nabla_{\mathbf{f}^{\mathsf{H}}}\left\{\mathcal{L}(\mathbf{f}_\ell, \lambda)\right\} = \mathbf{0}$$

$$\Leftrightarrow \sum_{i=1}^{K_\ell} \left(\prod_{k \neq i} \mathbf{f}_\ell^{\mathsf{H}} \left(\tilde{\mathbf{A}}_{\ell,\ell,k} + \hat{\mathbf{L}}_\ell\left(\mathbf{f}_\ell\right)^{\frac{|u_\ell|}{K_\ell}}\right) \mathbf{f}_\ell\right)\left(\tilde{\mathbf{A}}_{\ell,\ell,i} + \nabla_{\mathbf{f}_\ell^{\mathsf{H}}}\hat{\mathbf{L}}_\ell\left(\mathbf{f}_\ell\right)^{\frac{|u_\ell|}{K_\ell}}\right)$$

$$\cdot \prod_{k=1}^{K_\ell} \mathbf{f}_\ell^{\mathsf{H}} \left(\tilde{\mathbf{B}}_{\ell,\ell,k} + \hat{\mathbf{L}}_\ell\left(\mathbf{f}_\ell\right)^{\frac{|u_\ell|}{K_\ell}}\right) \mathbf{f}_\ell \cdot \mathbf{f}_\ell$$

$$= \sum_{i=1}^{K_\ell} \left(\prod_{k \neq i} \mathbf{f}_\ell^{\mathsf{H}} \left(\tilde{\mathbf{B}}_{\ell,\ell,k} + \hat{\mathbf{L}}_\ell\left(\mathbf{f}_\ell\right)^{\frac{|u_\ell|}{K_\ell}}\right) \mathbf{f}_\ell\right)\left(\tilde{\mathbf{B}}_{\ell,\ell,i} + \nabla_{\mathbf{f}_\ell^{\mathsf{H}}}\hat{\mathbf{L}}_\ell\left(\mathbf{f}_\ell\right)^{\frac{|u_\ell|}{K_\ell}}\right)$$

$$\cdot \prod_{k=1}^{K_\ell} \mathbf{f}_\ell^{\mathsf{H}} \left(\tilde{\mathbf{A}}_{\ell,\ell,k} + \hat{\mathbf{L}}_\ell\left(\mathbf{f}_\ell\right)^{\frac{|u_\ell|}{K_\ell}}\right) \mathbf{f}_\ell \cdot \mathbf{f}_\ell - \lambda \mathbf{f}_\ell. \tag{41}$$

By rearranging (41), we obtain:

$$\sum_{i=1}^{K_\ell} \left(\prod_{k \neq i} \mathbf{f}_\ell^{\mathsf{H}} \left(\tilde{\mathbf{A}}_{\ell,\ell,k} + \hat{\mathbf{L}}_\ell\left(\mathbf{f}_\ell\right)^{\frac{|u_\ell|}{K_\ell}}\right) \mathbf{f}_\ell\right)\left(\tilde{\mathbf{A}}_{\ell,\ell,i} + \tilde{\mathbf{C}}_{\ell,\ell,i}\right) \mathbf{f}_\ell$$

$$= \gamma(\mathbf{f}_\ell) \left[\sum_{i=1}^{K_\ell} \left(\prod_{k \neq i} \mathbf{f}_\ell^{\mathsf{H}} \left(\tilde{\mathbf{B}}_{\ell,\ell,k} + \hat{\mathbf{L}}_\ell\left(\mathbf{f}_\ell\right)^{\frac{|u_\ell|}{K_\ell}}\right) \mathbf{f}_\ell\right)\left(\tilde{\mathbf{B}}_{\ell,\ell,i} + \tilde{\mathbf{C}}_{\ell,\ell,i}\right)\right.$$

$$\left. - \frac{\lambda}{\gamma(\mathbf{f}_\ell)}\right] \mathbf{f}_\ell. \tag{42}$$

The condition (42) simplifies to

$$\bar{\mathbf{A}}_\ell\left(\mathbf{f}_\ell\right) \mathbf{f}_\ell = \gamma\left(\mathbf{f}_\ell\right) \bar{\mathbf{B}}_\ell\left(\mathbf{f}_\ell, \lambda\right) \mathbf{f}_\ell. \tag{43}$$

We also take the partial derivatives of $\mathcal{L}(\mathbf{f}, \lambda)$ with respective to Lagrange multiplier λ and set to them zero, which yields the condition:

$$\|\mathbf{f}_\ell\|_2^2 - 1 = 0 \Leftrightarrow \|\mathbf{f}_\ell\|_2^2 = 1. \tag{44}$$

This completes the proof.

B. Proof for Theorem 2

To prove the local optimality claim, it is sufficient to show that the extended Hessian matrix considering constraint sets at a stationary point is negative definite. To accomplish this, we derive the Hessian matrix evaluated at an arbitrary point $\mathbf{f}_\ell \in \mathbb{C}^{N_\ell K_\ell \times 1}$. For ease of exposition, we invoke the objective function $\gamma(\mathbf{f})$ to the gradient of the Lagrange function as

$$\nabla_{\mathbf{f}_\ell^{\mathsf{H}}}\left\{\mathcal{L}\left(\mathbf{f}_\ell, \lambda\right)\right\} = 2\gamma(\mathbf{f})\eta(\mathbf{f}_\ell) - 2\lambda\mathbf{f}_\ell, \tag{46}$$

where

$$\eta(\mathbf{f}_\ell) = \left[\sum_{i=1}^{K_\ell} \frac{\mathbf{X}_{\ell,\ell,i}^A}{\mathbf{f}_\ell^{\mathsf{H}}\mathbf{Y}_{\ell,\ell,i}^A\mathbf{f}_\ell} - \frac{\mathbf{X}_{\ell,\ell,i}^B}{\mathbf{f}_\ell^{\mathsf{H}}\mathbf{Y}_{\ell,\ell,i}^B\mathbf{f}_\ell}\right] \mathbf{f}_\ell. \tag{47}$$

The extended Hessian matrix is obtained by directly calculating the gradient of (46) again, which is given by

$$\nabla_{\mathbf{f}_\ell^{\mathsf{H}}}^2 \mathcal{L}(\mathbf{f}_\ell, \lambda) = \nabla_{\mathbf{f}_\ell^{\mathsf{H}}}\left\{\nabla_{\mathbf{f}_\ell^{\mathsf{H}}}\left\{\mathcal{L}(\mathbf{f}_\ell, \lambda)\right\}\right\}$$

$$= 2\nabla_{\mathbf{f}_\ell^{\mathsf{H}}}\left\{\gamma(\mathbf{f}_\ell)\right\}\eta(\mathbf{f}_\ell)^{\mathsf{H}} + 2\gamma(\mathbf{f}_\ell)\nabla_{\mathbf{f}_\ell^{\mathsf{H}}}\left\{\eta(\mathbf{f}_\ell)\right\} - 2\lambda\mathbf{I}. \tag{48}$$

We compute $\nabla_{\mathbf{f}_\ell^{\mathsf{H}}}^2\left\{\mathcal{L}(\mathbf{f}_\ell, \lambda)\right\}$ and $\nabla_{\mathbf{f}_\ell^{\mathsf{H}}}\left\{\eta(\mathbf{f}_\ell)\right\}$ in terms of the functional matrices defined as the Hermitian and positive definite matrices $\mathbf{X}_{\ell,\ell,i}^A$, $\mathbf{X}_{\ell,\ell,i}^B$, $\mathbf{Y}_{\ell,\ell,i}^A$, $\mathbf{Y}_{\ell,\ell,i}^B$, and $\mathbf{Z}_{\ell,\ell,i}$ in Theorem 2. Since the sum of positive definite (PSD) matrices is also a PSD matrix, the extended Hessian matrix becomes a negative definite matrix, provided that the the minimum eigenvalue of the third term in (45) is greater than the maximum eigenvalue of the sum of the first and second terms in (45). This is a sufficient condition for the local optimality. This ends the proof.

REFERENCES

[1] A. Ghosh, N. Mangalvedhe, R. Ratasuk, B. Mondal, M. Cudak, E. Visotsky, T. A. Thomas, J. G. Andrews, P. Xia, H. S. Jo, H. S. Dhillon, and T. D. Novlan, "Heterogeneous cellular networks: From theory to practice," *IEEE Commun. Mag.*, vol. 50, no. 6, pp. 54–64, 2012.

[2] H. S. Dhillon, M. Kountouris, and J. G. Andrews, "Downlink MIMO HetNets: Modeling, ordering results and performance analysis," *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 5208–5222, 2013.

[3] J. G. Andrews, S. Singh, Q. Ye, X. Lin, and H. S. Dhillon, "An overview of load balancing in HetNets: Old myths and open problems," *IEEE Wireless Commun.*, vol. 21, no. 2, pp. 18–25, 2014.

[4] K. Hosseini, J. Hoydis, S. Ten Brink, and M. Debbah, "Massive MIMO and small cells: How to densify heterogeneous networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, June 2013, pp. 5442–5447.

[5] N. Chen, B. Rong, X. Zhang, and M. Kadoch, "Scalable and flexible massive MIMO precoding for 5G H-CRAN," *IEEE Wireless Commun.*, vol. 24, no. 1, pp. 46–52, 2017.

[6] A. Adhikary, H. S. Dhillon, and G. Caire, "Spatial blanking and inter-tier coordination in massive MIMO heterogeneous cellular networks," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec 2014, pp. 1229–1234.

$$\nabla_{\mathbf{f}_\ell^\mathsf{H}}^2 \mathcal{L}\left(\mathbf{f}_\ell, \lambda\right) = \left[\sum_{i=1}^{K_\ell} \frac{\mathbf{X}_{\ell,\ell,i}^A \mathbf{f}_\ell}{\mathbf{f}_\ell^\mathsf{H} \mathbf{Y}_{\ell,\ell,i}^A \mathbf{f}_\ell} - \frac{\mathbf{X}_{\ell,\ell,i}^B \mathbf{f}_\ell}{\mathbf{f}_\ell^\mathsf{H} \mathbf{Y}_B \mathbf{f}_\ell}\right] \cdot \left[\sum_{i=1}^{K_\ell} \frac{\mathbf{X}_{\ell,\ell,i}^A \mathbf{f}_\ell}{\mathbf{f}_\ell^\mathsf{H} \mathbf{Y}_{\ell,\ell,i}^A \mathbf{f}_\ell} - \frac{\mathbf{X}_{\ell,\ell,i}^B \mathbf{f}_\ell}{\mathbf{f}_\ell^\mathsf{H} \mathbf{Y}_B \mathbf{f}_\ell}\right]^\mathsf{H} + \gamma(\mathbf{f}_\ell)\left[\sum_{i=1}^{K_\ell} \frac{\mathbf{X}_{\ell,\ell,i}^A}{\mathbf{f}_\ell^\mathsf{H} \mathbf{Y}_{\ell,\ell,i}^A \mathbf{f}_\ell} + \frac{\mathbf{Z}_{\ell,\ell,i}}{\mathbf{f}_\ell^\mathsf{H} \mathbf{Y}_{\ell,\ell,i}^B \mathbf{f}_\ell} + \frac{\mathbf{X}_{\ell,\ell,i}^B \mathbf{f}_\ell \mathbf{f}_\ell^\mathsf{H} \mathbf{X}_{\ell,\ell,i}^B}{\left(\mathbf{f}_\ell^\mathsf{H} \mathbf{Y}_{\ell,\ell,i}^B \mathbf{f}_\ell\right)^2}\right]$$
$$- \left[\gamma(\mathbf{f}_\ell)\left\{\sum_{i=1}^{K_\ell} \frac{\mathbf{X}_{\ell,\ell,i}^B}{\mathbf{f}_\ell^\mathsf{H} \mathbf{Y}_{\ell,\ell,i}^B \mathbf{f}_\ell} + \frac{\mathbf{Z}_{\ell,\ell,i}}{\mathbf{f}_\ell^\mathsf{H} \mathbf{Y}_{\ell,\ell,i}^A \mathbf{f}_\ell} + \frac{\mathbf{X}_{\ell,\ell,i}^A \mathbf{f}_\ell \mathbf{f}_\ell^\mathsf{H} \mathbf{X}_{\ell,\ell,i}^A}{\left(\mathbf{f}_\ell^\mathsf{H} \mathbf{Y}_{\ell,\ell,i}^A \mathbf{f}_\ell\right)^2}\right\} + \lambda \mathbf{I}\right]. \tag{45}$$

[7] D. Lopez-Perez, I. Guvenc, G. de la Roche, M. Kountouris, T. Quek, and J. Zhang, "Enhanced intercell interference coordination challenges in heterogeneous networks," *IEEE Wireless Commun.*, vol. 3, no. 18, pp. 22–30, 2011.

[8] A. K. Gupta, H. S. Dhillon, S. Vishwanath, and J. G. Andrews, "Downlink multi-antenna heterogeneous cellular network with load balancing," *IEEE Trans. Commun.*, vol. 62, no. 11, pp. 4052–4067, 2014.

[9] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing—the large-scale array regime," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6441–6463, 2013.

[10] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, 2014.

[11] F. Rusek, D. Persson, B. Lau, E. Larsson, T. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, pp. 40–60, 2013.

[12] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, 2010.

[13] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, 2013.

[14] S. Li, P. J. Smith, P. A. Dmochowski, H. Tataria, M. Matthaiou, and J. Yin, "Massive MIMO for ray-based channels," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2019, pp. 1–7.

[15] N. Lee, D. Morales-Jimenez, A. Lozano, and R. W. Heath Jr., "Spectral efficiency of dynamic coordinated beamforming: A stochastic geometry approach," *IEEE Trans. Wireless Commun.*, vol. 14, no. 1, pp. 230–241, 2014.

[16] J. Park, N. Lee, and R. W. Heath Jr., "Cooperative base station coloring for pair-wise multi-cell coordination," *IEEE Trans. Commun.*, vol. 64, no. 1, pp. 402–415, 2015.

[17] T. K. Lo, "Maximum ratio transmission," in *Proc. IEEE Int. Conf. Commun. (ICC)*, vol. 2, June 1999, pp. 1310–1314.

[18] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Designing multi-user MIMO for energy efficiency: When is massive MIMO the answer?" in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, April 2014, pp. 242–247.

[19] Y.-G. Lim, C.-B. Chae, and G. Caire, "Performance analysis of massive MIMO for cell-boundary users," *IEEE Trans. Wireless Commun.*, vol. 14, no. 12, pp. 6827–6842, 2015.

[20] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, 2004.

[21] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 528–541, 2006.

[22] M. Costa, "Writing on dirty paper (corresp.)," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, 1983.

[23] T. E. Bogale and L. B. Le, "Massive MIMO and mmWave for 5G wireless HetNet: Potential benefits and challenges," *IEEE Veh. Technol. Mag.*, vol. 11, no. 1, pp. 64–75, 2016.

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

[25] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, 2011.

[26] E. Björnson and L. Sanguinetti, "A new look at cell-free massive MIMO: Making it practical with dynamic cooperation," in *Proc. IEEE Annu. Int. Symp. Pers., Indoor and Mobile Radio Commun. (PIMRC)*, Sep. 2019, pp. 1–6.

[27] S. S. Christensen, R. Agarwal, E. De Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792–4799, 2008.

[28] J. Choi, N. Lee, S.-N. Hong, and G. Caire, "Joint user scheduling, power allocation, and precoding design for massive MIMO systems: A principal component analysis approach," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, June 2018, pp. 396–400.

[29] ——, "Joint user selection, power allocation, and precoding design with imperfect CSIT for multi-cell MU-MIMO downlink systems," *IEEE Trans. Wireless Commun. (Early Access)*, 2019.

[30] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, 2013.

[31] J. Hoydis, S. Ten Brink, and M. Debbah, "Massive MIMO: How many antennas do we need?" in *Proc. 49th Annu. Allerton Conf. Commun., Control, and Comput. (Allerton)*, Sep. 2011, pp. 545–550.

[32] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 933–946, 2000.

[33] T. Yoo, E. Yoon, and A. Goldsmith, "MIMO capacity with channel uncertainty: Does feedback help?" in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, vol. 1, 2004, pp. 96–100.

[34] J. Park, N. Lee, J. G. Andrews, and R. W. Heath Jr., "On the optimal feedback rate in interference-limited multi-antenna cellular systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 8, pp. 5748–5762, 2016.

[35] W. Shin, N. Lee, J.-B. Lim, C. Shin, and K. Jang, "On the design of interference alignment scheme for two-cell MIMO interfering broadcast channels," *IEEE Trans. Wireless Commun.*, vol. 10, no. 2, pp. 437–442, 2010.

[36] P. Patcharamaneepakorn, S. Armour, and A. Doufexi, "On the equivalence between SLNR and MMSE precoding schemes with single-antenna receivers," *IEEE Commun. Lett.*, vol. 16, no. 7, pp. 1034–1037, 2012.

[37] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO has unlimited capacity," *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, pp. 574–590, 2017.

[38] J. Sherman, "Adjustment of an inverse matrix corresponding to changes in the elements of a given column or a given row of the original matrix," *Annals of mathematical statistics*, vol. 20, no. 4, p. 621, 1949.

[39] 3GPP TR 36.872, "Small cell enhancements for E-UTRA and E-UTRAN-Physical layer aspect."

[40] X. Li, E. Bjornson, E. G. Larsson, S. Zhou, and J. Wang, "A multi-cell MMSE detector for massive MIMO systems and new large system analysis," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, 2015, pp. 1–6.