

Distributed Precoding Design for Cell-Free Massive MIMO Systems

Bikshapathi Gouda, Italo Atzeni, and Antti Tölli

Centre for Wireless Communications, University of Oulu, Finland

Emails: {bikshapathi.gouda, italo.atzeni, antti.tolli}@oulu.fi

Abstract—We propose a distributed precoding design for downlink cell-free massive multiple-input multiple-output (MIMO) systems where the channel state information (CSI) exchange via backhaul signaling is entirely replaced by a novel over-the-air (OTA) signaling mechanism. The proposed method enjoys desirable flexibility and scalability properties, as the amount of OTA signaling does not scale with the number of base stations or user equipments. Numerical results show remarkable gains over uncoordinated precoding design, even after very few iterations. The proposed method may also outperform the centralized precoding design under realistic CSI acquisition.

I. INTRODUCTION

Cell-free massive multiple-input multiple-output (MIMO) is a recent concept that conveniently combines elements from small cells, massive MIMO, and user-centric joint transmission coordinated multi-point (JT-CoMP) [1]. In cell-free massive MIMO, a large number of antenna elements are distributed across the network to provide overall reduced pathloss and enhanced coverage. Furthermore, a user-centric coherent transmission extended to the whole network, where each user equipment (UE) is served jointly by all the base stations (BSs), allows to entirely eliminate the inter-cell interference [2]. To this end, all the BSs are connected to a central processing unit (CPU) that provides the UE-specific data via backhaul links.

Cell-free massive MIMO systems have been shown to outperform traditional small-cell and cellular massive MIMO networks in several practical scenarios [3], [4], while the sum-rate performance depends on the level of coordination among the BSs [5]. Simple uncoordinated precoding strategies, such as matched filtering and local minimum mean squared error (MMSE), are typically assumed at the BSs as they can be optimized based on locally acquired channel state information (CSI) without any CSI exchange via backhaul signaling. However, the fact that the channel hardening effect is less pronounced in cell-free massive MIMO than in cellular massive MIMO [1] suggests that cooperative precoding design can bring considerable gains over its non-cooperative counterpart.

In the centralized precoding design, the BSs convey their locally acquired CSI to the CPU, which then feeds back the optimized precoding vectors to the BSs via backhaul signaling. As the number of BSs and UEs involved in the joint processing increases, both the computational complexity at the CPU and the amount of CSI exchange via backhaul signaling become impractical. In [6], a distributed iterative scheme for JT-CoMP was proposed to optimize the precoding strategies locally at

This work was supported by the Academy of Finland under grant no. 318927 (6Genesis Flagship).

each BS using bi-directional training between the BSs and the UEs [7] along with periodic exchange of cross-term information among the BSs via backhaul signaling. Still, the amount of such CSI exchange makes the implementation of [6] challenging, especially in cell-free massive MIMO contexts.

This paper proposes a distributed precoding design for cell-free massive MIMO (and, more generally, for JT-CoMP) systems that entirely eliminates the need for backhaul signaling for CSI exchange. To this end, a novel over-the-air (OTA) signaling mechanism allows each BS to acquire the same cross-term information that was exchanged among the BSs via backhaul signaling in [6]. Remarkably, the amount of OTA signaling does not scale with the number of BSs or UEs, and depends only on the pilot length and on the number of bi-directional training iterations. The proposed distributed precoding design via OTA signaling brings significant gains over local MMSE precoding, even after very few iterations. Furthermore, it outperforms the centralized precoding design under realistic CSI acquisition thanks to its robustness against imperfect channel estimation.

II. SYSTEM MODEL

Consider a downlink cell-free massive MIMO network where a set of BSs $\mathcal{B} \triangleq \{1, \dots, B\}$, each equipped with M antennas, serve a set of UEs $\mathcal{K} \triangleq \{1, \dots, K\}$, each equipped with N antennas. Assuming a time division duplex (TDD) setting with channel reciprocity and a single data stream per UE, let $\mathbf{H}_{b,k} \in \mathbb{C}^{M \times N}$ be the uplink channel matrix between UE $k \in \mathcal{K}$ and BS $b \in \mathcal{B}$, with $\mathbf{H}_k \triangleq [\mathbf{H}_{1,k}^T, \dots, \mathbf{H}_{B,k}^T]^T \in \mathbb{C}^{BM \times N}$ denoting the global uplink channel matrix seen by UE k . Likewise, let $\mathbf{w}_{b,k} \in \mathbb{C}^{M \times 1}$ be the BS-specific precoding vector used by BS b for UE k , with $\mathbf{w}_k \triangleq [\mathbf{w}_{1,k}^T, \dots, \mathbf{w}_{B,k}^T]^T \in \mathbb{C}^{BM \times 1}$ denoting the global precoding vector used for UE k . Hence, the receive signal at UE k is given by

$$\mathbf{y}_k \triangleq \sum_{b \in \mathcal{B}} \sum_{\bar{k} \in \mathcal{K}} \mathbf{H}_{b,k}^H \mathbf{w}_{b,\bar{k}} d_{\bar{k}} + \mathbf{z}_k \in \mathbb{C}^{N \times 1} \quad (1)$$

where $d_k \sim \mathcal{CN}(0, 1)$ is the transmit data symbol for UE k , and $\mathbf{z}_k \sim \mathcal{CN}(\mathbf{0}, \sigma_k^2 \mathbf{I}_N)$ is the average white Gaussian noise (AWGN) at UE k . Upon receiving \mathbf{y}_k , UE k uses the combining vector $\mathbf{v}_k \in \mathbb{C}^{N \times 1}$ and the resulting signal-to-interference-plus-noise ratio (SINR) is

$$\text{SINR}_k \triangleq \frac{|\sum_{b \in \mathcal{B}} \mathbf{v}_k^H \mathbf{H}_{b,k} \mathbf{w}_{b,k}|^2}{\sum_{\bar{k} \in \mathcal{K} \setminus \{k\}} |\sum_{b \in \mathcal{B}} \mathbf{v}_k^H \mathbf{H}_{b,k} \mathbf{w}_{b,\bar{k}}|^2 + \|\mathbf{v}_k\|^2 \sigma_k^2}. \quad (2)$$

Finally, the sum rate (measured in bps/Hz) is given by $R \triangleq \sum_{k \in \mathcal{K}} \log_2(1 + \text{SINR}_k)$. In the following, we describe realistic pilot-aided CSI acquisition at both the BSs and the UEs.

A. Uplink Pilot-Aided Channel Estimation

Let $\mathbf{h}_{b,k} \triangleq \mathbf{H}_{b,k} \mathbf{v}_k \in \mathbb{C}^{M \times 1}$ be the effective uplink channel vector between UE k and BS b , and let $\mathbf{p}_k \in \mathbb{C}^{\tau \times 1}$ be the pilot assigned to UE k , with $\|\mathbf{p}_k\|^2 = \tau$. During the uplink pilot-aided channel estimation, each UE k simultaneously¹ transmits its pilot precoded with its combining vector, i.e.,

$$\mathbf{X}_k^{\text{UL-1}} \triangleq \mathbf{v}_k \mathbf{p}_k^H \in \mathbb{C}^{N \times \tau}. \quad (3)$$

Note that, when used as precoder, \mathbf{v}_k must comply with the transmit power constraint of UE k (as detailed in Section V). Then, each BS b receives

$$\mathbf{Y}_b^{\text{UL-1}} \triangleq \sum_{k \in \mathcal{K}} \mathbf{h}_{b,k} \mathbf{p}_k^H + \mathbf{Z}_b^{\text{UL-1}} \in \mathbb{C}^{M \times \tau} \quad (4)$$

where $\mathbf{Z}_b^{\text{UL-1}} \in \mathbb{C}^{M \times \tau}$ is the AWGN at BS b with elements distributed as $\mathcal{CN}(0, \sigma_b^2)$, and the least-squares (LS) estimate of $\mathbf{h}_{b,k}$ is obtained as

$$\hat{\mathbf{h}}_{b,k} \triangleq \frac{1}{\tau} \mathbf{Y}_b^{\text{UL-1}} \mathbf{p}_k \quad (5)$$

$$= \mathbf{h}_{b,k} + \frac{1}{\tau} \sum_{\bar{k} \in \mathcal{K} \setminus \{k\}} \mathbf{h}_{b,\bar{k}} \mathbf{p}_{\bar{k}}^H \mathbf{p}_k + \frac{1}{\tau} \mathbf{Z}_b^{\text{UL-1}} \mathbf{p}_k. \quad (6)$$

For the centralized precoding design, the estimation of the channel matrix $\mathbf{H}_{b,k}$ requires N antenna-specific pilots for UE k . In this context, let $\mathbf{P}_k \in \mathbb{C}^{\tau \times N}$ be the pilot matrix assigned to UE k , with $\|\mathbf{P}_k\|_F^2 = \tau N$. Similarly to (3), each UE k simultaneously transmits its pilot matrix and BS b obtains the LS estimate of $\mathbf{H}_{b,k}$ as (cf. (6))

$$\hat{\mathbf{H}}_{b,k} \triangleq \frac{1}{\tau} \sum_{\bar{k} \in \mathcal{K}} \mathbf{H}_{b,\bar{k}} \mathbf{p}_{\bar{k}}^H \mathbf{P}_k + \frac{1}{\tau} \mathbf{Z}_b^{\text{UL-1}} \mathbf{P}_k. \quad (7)$$

B. Downlink Pilot-Aided Channel Estimation

Let $\mathbf{g}_k \triangleq \sum_{b \in \mathcal{B}} \mathbf{H}_{b,k}^H \mathbf{w}_{b,k} \in \mathbb{C}^{N \times 1}$ be the effective downlink channel vector between all the BSs and UE k . During the downlink pilot-aided channel estimation, each BS b simultaneously transmits a superposition of pilots after precoding them with the corresponding precoding vectors, i.e.,

$$\mathbf{X}_b^{\text{DL}} \triangleq \sum_{k \in \mathcal{K}} \mathbf{w}_{b,k} \mathbf{p}_k^H \in \mathbb{C}^{M \times \tau}. \quad (8)$$

Then, each UE k receives

$$\mathbf{Y}_k^{\text{DL}} \triangleq \sum_{b \in \mathcal{B}} \sum_{\bar{k} \in \mathcal{K}} \mathbf{H}_{b,k}^H \mathbf{w}_{b,\bar{k}} \mathbf{p}_{\bar{k}}^H + \mathbf{Z}_k^{\text{DL}} \in \mathbb{C}^{N \times \tau} \quad (9)$$

where $\mathbf{Z}_k^{\text{DL}} \in \mathbb{C}^{N \times \tau}$ is the AWGN at UE k with elements distributed as $\mathcal{CN}(0, \sigma_k^2)$, and the LS estimate of \mathbf{g}_k is obtained as

$$\hat{\mathbf{g}}_k \triangleq \frac{1}{\tau} \mathbf{Y}_k^{\text{DL}} \mathbf{p}_k \quad (10)$$

$$= \mathbf{g}_k + \frac{1}{\tau} \sum_{b \in \mathcal{B}} \sum_{\bar{k} \in \mathcal{K} \setminus \{k\}} \mathbf{H}_{b,k}^H \mathbf{w}_{b,\bar{k}} \mathbf{p}_{\bar{k}}^H \mathbf{p}_k + \frac{1}{\tau} \mathbf{Z}_k^{\text{DL}} \mathbf{p}_k. \quad (11)$$

III. PROBLEM FORMULATION WITH PERFECT CSI

In this paper, we target the weighted sum mean squared error (MSE) minimization problem to optimize the precoding vectors $\{\mathbf{w}_{b,k}\}_{b \in \mathcal{B}, k \in \mathcal{K}}$ and the combining vectors $\{\mathbf{v}_k\}_{k \in \mathcal{K}}$ (see [8] for more details). For the sake of clarity, we begin by illustrating

¹Although perfect synchronization is infeasible, quasi-synchronous transmissions among UEs and BSs can be achieved in practice as described in [1], [5].

the case of perfect channel estimation; then, we focus on the case of realistic pilot-aided CSI acquisition in Section IV.

Building on (1), let us introduce the MSE at UE k as

$$\text{MSE}_k \triangleq \mathbb{E}[|\mathbf{v}_k^H \mathbf{y}_k - d_k|^2] \quad (12)$$

$$= \sum_{\bar{k} \in \mathcal{K}} \left| \sum_{b \in \mathcal{B}} \mathbf{v}_k^H \mathbf{H}_{b,k}^H \mathbf{w}_{b,\bar{k}} \right|^2 - 2 \text{Re} \left[\sum_{b \in \mathcal{B}} \mathbf{v}_k^H \mathbf{H}_{b,k}^H \mathbf{w}_{b,k} \right] + \sigma_k^2 \|\mathbf{v}_k\|^2 + 1 \quad (13)$$

which is convex with respect to either the transmit or the receive strategies, but not jointly convex with respect to both. Hence, we use *alternate optimization*, whereby the precoding vectors are optimized for fixed combining vectors and vice versa in an iterative best-response fashion [8].

- **Computation of the combining vectors.** The combining vectors $\{\mathbf{v}_k\}_{k \in \mathcal{K}}$ are computed locally and independently by the UEs such that each UE k minimizes MSE_k in (13). From the point of view of UE k , we can rewrite the MSE as

$$\text{MSE}_k = \mathbf{v}_k^H (\boldsymbol{\Psi}_k + \sigma_k^2 \mathbf{I}_N) \mathbf{v}_k - 2 \text{Re}[\mathbf{v}_k^H \mathbf{g}_k] + 1 \quad (14)$$

where we have defined

$$\boldsymbol{\Psi}_k \triangleq \sum_{\bar{k} \in \mathcal{K}} \left(\sum_{b \in \mathcal{B}} \mathbf{H}_{b,k}^H \mathbf{w}_{b,\bar{k}} \right) \left(\sum_{b \in \mathcal{B}} \mathbf{w}_{b,\bar{k}}^H \mathbf{H}_{b,k} \right). \quad (15)$$

The combining vector \mathbf{v}_k that minimizes (14) is the well-known MMSE receiver, which may be written as

$$\mathbf{v}_k = (\boldsymbol{\Psi}_k + \sigma_k^2 \mathbf{I}_N)^{-1} \mathbf{g}_k. \quad (16)$$

Observe that $\boldsymbol{\Psi}_k$ in (15) and the effective downlink channel \mathbf{g}_k must be known by UE k .

- **Computation of the precoding vectors.** The precoding vectors $\{\mathbf{w}_{b,k}\}_{b \in \mathcal{B}, k \in \mathcal{K}}$ are computed as the solutions of the weighted sum MSE minimization problem with per-BS power constraints $\{\sum_{k \in \mathcal{K}} \|\mathbf{w}_{b,k}\|^2 \leq \rho_{\text{BS}}\}_{b \in \mathcal{B}}$, where ω_k is the weight of UE k . In this respect, we introduce some preliminary definitions: $\mathbf{h}_k \triangleq [\mathbf{h}_{1,k}^T, \dots, \mathbf{h}_{B,k}^T]^T \in \mathbb{C}^{BM \times 1}$, $\mathbf{H} \triangleq [\mathbf{h}_1, \dots, \mathbf{h}_K] \in \mathbb{C}^{BM \times K}$, $\mathbf{W} \triangleq [\mathbf{w}_1, \dots, \mathbf{w}_K] \in \mathbb{C}^{BM \times K}$, $\boldsymbol{\Omega} \triangleq \text{diag}([\omega_1, \dots, \omega_K]) \in \mathbb{R}^{K \times K}$, and $\boldsymbol{\Phi} \triangleq \sum_{k \in \mathcal{K}} \omega_k \mathbf{h}_k \mathbf{h}_k^H \in \mathbb{C}^{BM \times BM}$, where the latter may be rewritten as

$$\boldsymbol{\Phi} \triangleq \begin{bmatrix} \boldsymbol{\Phi}_{11} & \dots & \boldsymbol{\Phi}_{1B} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Phi}_{1B}^H & \dots & \boldsymbol{\Phi}_{BB} \end{bmatrix} \quad (17)$$

with $\boldsymbol{\Phi}_{b\bar{b}} \triangleq \sum_{k \in \mathcal{K}} \omega_k \mathbf{h}_{b,k} \mathbf{h}_{\bar{b},k}^H \in \mathbb{C}^{M \times M}$. Finally, the weighted sum MSE can be expressed as

$$\sum_{k \in \mathcal{K}} \omega_k \text{MSE}_k = \text{tr}(\mathbf{W}^H \boldsymbol{\Phi} \mathbf{W}) - 2 \text{Re}[\text{tr}(\boldsymbol{\Omega} \mathbf{H}^H \mathbf{W})] + \sum_{k \in \mathcal{K}} \omega_k (\sigma_k^2 \|\mathbf{v}_k\|^2 + 1). \quad (18)$$

Next, we describe the centralized and the distributed precoding designs in Sections III-A and III-B, respectively.

A. Centralized Precoding Design

In the centralized precoding design, the alternate optimization of the precoding and the combining vectors takes place transparently at the CPU. For fixed combining vectors, for each UE k , the centralized precoding solution is given by

$$\mathbf{w}_k = \omega_k (\Phi + \Lambda)^{-1} \mathbf{h}_k \quad (19)$$

where $\Lambda \triangleq \text{blkdiag}(\lambda_1 \mathbf{I}_M, \dots, \lambda_K \mathbf{I}_M) \in \mathbb{R}^{BM \times BM}$ contains the dual variables $\{\lambda_b\}_{b \in \mathcal{B}}$ related with the per-BS power constraints (which can be computed using the ellipsoid method).

The centralized precoding design is carried out as follows: 1) each BS b acquires the channel matrices $\{\mathbf{H}_{b,k}\}_{k \in \mathcal{K}}$ and transmits them to the CPU via backhaul signaling; 2) the CPU computes the global precoding vectors $\{\mathbf{w}_k\}_{k \in \mathcal{K}}$ as in (19) and the combining vectors $\{\mathbf{v}_k\}_{k \in \mathcal{K}}$ as in (16) via alternate optimization; 3) the CPU feeds back the BS-specific precoding vectors $\{\mathbf{w}_{b,k}\}_{k \in \mathcal{K}}$ to each BS b via backhaul signaling; 4) each UE k acquires Ψ_k in (15) and the effective downlink channel \mathbf{g}_k , and computes its combining vector \mathbf{v}_k as in (16).

B. Distributed Precoding Design

In the distributed precoding design, the alternate optimization of the precoding and the combining vectors takes place between the BSs and the UEs by means of bi-directional training [7]. For fixed combining vectors, for each BS b and for each UE k , the distributed precoding solution is given by

$$\mathbf{w}_{b,k} = (\Phi_{bb} + \lambda_b \mathbf{I}_M)^{-1} (\omega_k \mathbf{h}_{b,k} - \boldsymbol{\eta}_{b,k}) \quad (20)$$

where we have defined

$$\boldsymbol{\eta}_{b,k} \triangleq \sum_{\bar{b} \in \mathcal{B} \setminus \{b\}} \Phi_{b\bar{b}} \mathbf{w}_{\bar{b},k} \quad (21)$$

and where λ_b can be computed via bisection. At each iteration i , each BS b can compute $\mathbf{w}_{b,k}$ locally as in (20) for fixed $\{\mathbf{w}_{\bar{b},k}\}_{\bar{b} \in \mathcal{B} \setminus \{b\}}$ in parallel with the other BSs and update it as

$$\mathbf{w}_{b,k}^{(i)} = (1 - \alpha) \mathbf{w}_{b,k}^{(i-1)} + \alpha \mathbf{w}_{b,k} \quad (22)$$

with $\alpha \in (0, 1]$ (see [6] for more details).

Remark 1. The vector $\boldsymbol{\eta}_{b,k}$ in (21) contains implicit information about the channel correlation between BS b and the other BSs and about the precoding vectors adopted by the latter for UE k . Such cross-term information allows to iteratively adjust the distributed precoding solution so that it converges to the centralized precoding solution in (19). Note that omitting $\boldsymbol{\eta}_{b,k}$ from (20) yields the highly suboptimal *local MMSE precoding*. In this regard, while the effective uplink channels $\{\mathbf{h}_{b,k}\}_{k \in \mathcal{K}}$ can be acquired locally by each BS b , extensive CSI exchange among the BSs via backhaul signaling is still required to obtain $\{\boldsymbol{\eta}_{b,k}\}_{k \in \mathcal{K}}$ [6]. In Section V, we propose a practical method to acquire an estimate of such terms solely via OTA signaling.

The distributed precoding design iterates the following process until a predefined termination criterion is met: 1) for fixed combining vectors $\{\mathbf{v}_k\}_{k \in \mathcal{K}}$, each BS b acquires the effective uplink channels $\{\mathbf{h}_{b,k}\}_{k \in \mathcal{K}}$ and the vectors $\{\boldsymbol{\eta}_{b,k}\}_{k \in \mathcal{K}}$; 2) each BS b computes its BS-specific precoding vectors $\{\mathbf{w}_{b,k}\}_{k \in \mathcal{K}}$ locally as in (20) and updates them as in (22); 3) each UE k acquires Ψ_k in (15) and the effective downlink channel \mathbf{g}_k , and computes its combining vector \mathbf{v}_k as in (16).

IV. PROBLEM FORMULATION WITH IMPERFECT CSI

We now consider the centralized and the distributed precoding designs described in Sections III-A and III-B, respectively under realistic pilot-aided CSI acquisition at both the BSs and

the UEs (see Sections II-A and II-B). For notational simplicity, and without loss of generality, we assume $\{\omega_k = 1\}_{k \in \mathcal{K}}$.

A. Centralized Precoding Design

In the centralized precoding design, the CPU computes the combining and the precoding vectors for each UE k as

$$\mathbf{v}_k = \left(\sum_{\bar{k} \in \mathcal{K}} \left(\sum_{b \in \mathcal{B}} \hat{\mathbf{H}}_{b,k}^H \mathbf{w}_{b,\bar{k}} \right) \left(\sum_{b \in \mathcal{B}} \mathbf{w}_{b,\bar{k}}^H \hat{\mathbf{H}}_{b,k} \right) + \sigma_k^2 \mathbf{I}_N \right)^{-1} \times \sum_{b \in \mathcal{B}} \hat{\mathbf{H}}_{b,k}^H \mathbf{w}_{b,k}, \quad (23)$$

$$\mathbf{w}_k = \left(\sum_{\bar{k} \in \mathcal{K}} \hat{\mathbf{H}}_{\bar{k}} \mathbf{v}_{\bar{k}} \mathbf{v}_{\bar{k}}^H \hat{\mathbf{H}}_{\bar{k}}^H + \Lambda \right)^{-1} \hat{\mathbf{H}}_k \mathbf{v}_k \quad (24)$$

respectively. More specifically, (23) and (24) are obtained by replacing the channels $\{\mathbf{H}_{b,k}\}_{b \in \mathcal{B}}$ with the estimated channels $\{\hat{\mathbf{H}}_{b,k}\}_{b \in \mathcal{B}}$ (obtained as in (7)) in (16) and (19), respectively.

B. Distributed Precoding Design

In the distributed precoding design, after the downlink pilot-aided channel estimation phase, each UE k obtains

$$\frac{1}{\tau} \mathbf{Y}_k^{\text{DL}} (\mathbf{Y}_k^{\text{DL}})^H = \Psi_k + \frac{1}{\tau} \sum_{\substack{\bar{k}, j \in \mathcal{K} \\ \bar{k} \neq j}} \left(\sum_{b \in \mathcal{B}} \mathbf{H}_{b,k}^H \mathbf{w}_{b,\bar{k}} \right) \times \left(\sum_{b \in \mathcal{B}} \mathbf{w}_{b,j}^H \mathbf{H}_{b,k} \right) \mathbf{p}_{\bar{k}}^H \mathbf{p}_j + \mathbf{N}_k^{\text{DL}} \quad (25)$$

with \mathbf{Y}_k^{DL} and Ψ_k defined in (9) and (15), respectively, and

$$\mathbf{N}_k^{\text{DL}} \triangleq \frac{1}{\tau} \left(\sum_{b \in \mathcal{B}} \sum_{\bar{k} \in \mathcal{K}} \left(\mathbf{H}_{b,k}^H \mathbf{w}_{b,\bar{k}} \mathbf{p}_{\bar{k}}^H (\mathbf{Z}_k^{\text{DL}})^H + \mathbf{Z}_k^{\text{DL}} \mathbf{p}_{\bar{k}} \mathbf{w}_{b,\bar{k}}^H \mathbf{H}_{b,k} \right) + \mathbf{Z}_k^{\text{DL}} (\mathbf{Z}_k^{\text{DL}})^H \right). \quad (26)$$

Here, perfect channel estimation would imply that: *i*) the pilot contamination in the second term of (25) is eliminated; *ii*) as $\tau \rightarrow \infty$, $\mathbf{N}_k^{\text{DL}} \rightarrow \sigma_k^2 \mathbf{I}_N$. Hence, (25) can be used as an estimate of $\Psi_k + \sigma_k^2 \mathbf{I}_N$ and, consequently, we can obtain an estimate of MSE_k in (14) as

$$\text{MSE}_k \simeq \frac{1}{\tau} \mathbf{v}_k^H \mathbf{Y}_k^{\text{DL}} (\mathbf{Y}_k^{\text{DL}})^H \mathbf{v}_k - \frac{2}{\tau} \text{Re}[\mathbf{v}_k^H \mathbf{Y}_k^{\text{DL}} \mathbf{p}_k] + 1. \quad (27)$$

Finally, the combining vector \mathbf{v}_k can be computed locally by UE k as

$$\mathbf{v}_k = (\mathbf{Y}_k^{\text{DL}} (\mathbf{Y}_k^{\text{DL}})^H)^{-1} \mathbf{Y}_k^{\text{DL}} \mathbf{p}_k \quad (28)$$

which is equal to (16) for perfect channel estimation.

On the other hand, after the uplink pilot-aided channel estimation phase, for each BS pair b and \bar{b} , we have

$$\frac{1}{\tau} \mathbf{Y}_b^{\text{UL-1}} (\mathbf{Y}_b^{\text{UL-1}})^H = \Phi_{b\bar{b}} + \frac{1}{\tau} \sum_{\substack{k, \bar{k} \in \mathcal{K} \\ k \neq \bar{k}}} \mathbf{h}_{b,k} \mathbf{h}_{\bar{b},\bar{k}}^H \mathbf{p}_k^H \mathbf{p}_{\bar{k}} + \mathbf{N}_{b\bar{b}}^{\text{UL-1}} \quad (29)$$

with $\mathbf{Y}_b^{\text{UL-1}}$ defined in (4) and

$$\mathbf{N}_{b\bar{b}}^{\text{UL-1}} \triangleq \frac{1}{\tau} \left(\sum_{k \in \mathcal{K}} \left(\mathbf{h}_{b,k} \mathbf{p}_k^H (\mathbf{Z}_b^{\text{UL-1}})^H + \mathbf{Z}_b^{\text{UL-1}} \mathbf{p}_k \mathbf{h}_{\bar{b},k}^H \right) + \mathbf{Z}_b^{\text{UL-1}} (\mathbf{Z}_b^{\text{UL-1}})^H \right). \quad (30)$$

Here, perfect channel estimation would imply that: *i*) the pilot

$$\begin{aligned} & \frac{1}{\tau} (\mathbf{Y}_b^{\text{UL-2}} \mathbf{p}_k - (\mathbf{Y}_b^{\text{UL-1}} (\mathbf{Y}_b^{\text{UL-1}})^{\text{H}} - \tau \sigma_b^2 \mathbf{I}_M) \mathbf{w}_{b,k}) \\ &= \boldsymbol{\eta}_{b,k} + \frac{1}{\tau} \left(\sum_{\bar{k} \neq k} \sum_{\bar{b} \in \mathcal{B}} \Phi_{\bar{b}\bar{k}} \mathbf{w}_{\bar{b},\bar{k}} \mathbf{p}_k^{\text{H}} \mathbf{p}_k - \sum_{\substack{\bar{k},j \in \mathcal{K} \\ \bar{k} \neq j}} \mathbf{h}_{b,\bar{k}} \mathbf{h}_{b,j}^{\text{H}} \mathbf{w}_{b,k} \mathbf{p}_k^{\text{H}} \mathbf{p}_j \right) + \mathbf{n}_{b,k}^{\text{UL-2}} + (\sigma_b^2 \mathbf{I}_M - \mathbf{N}_{bb}^{\text{UL-1}}) \mathbf{w}_{b,k} \quad (37) \end{aligned}$$

contamination in the second term of (29) is eliminated; *ii*) as $\tau \rightarrow \infty$, $\mathbf{N}_{bb}^{\text{UL-1}} \rightarrow \mathbf{0}$ if $\bar{b} \neq b$ and $\mathbf{N}_{bb}^{\text{UL-1}} \rightarrow \sigma_b^2 \mathbf{I}_M$. Hence, (29) can be used as an estimate of $\Phi_{\bar{b}\bar{k}}$ if $\bar{b} \neq b$ or of $\Phi_{bb} + \sigma_b^2 \mathbf{I}_M$ if $\bar{b} = b$. This can be exploited to obtain an estimate of the sum MSE in (18) as

$$\begin{aligned} \sum_{k \in \mathcal{K}} \text{MSE}_k &\simeq \frac{1}{\tau} \text{tr}(\mathbf{W}^{\text{H}} (\mathbf{Y}^{\text{UL-1}} (\mathbf{Y}^{\text{UL-1}})^{\text{H}} - \tau \boldsymbol{\Sigma}) \mathbf{W}) \\ &- \frac{2}{\tau} \text{Re}[\text{tr}(\mathbf{P}^{\text{H}} (\mathbf{Y}^{\text{UL-1}})^{\text{H}} \mathbf{W})] + \sum_{k \in \mathcal{K}} \sigma_k^2 \|\mathbf{v}_k\|^2 + K \quad (31) \end{aligned}$$

where we have defined $\mathbf{Y}^{\text{UL-1}} \triangleq [(\mathbf{Y}_1^{\text{UL-1}})^{\text{T}}, \dots, (\mathbf{Y}_K^{\text{UL-1}})^{\text{T}}]^{\text{T}} \in \mathbb{C}^{BM \times \tau}$, $\boldsymbol{\Sigma} \triangleq \text{blkdiag}(\sigma_1^2 \mathbf{I}_M, \dots, \sigma_B^2 \mathbf{I}_M) \in \mathbb{R}^{BM \times BM}$, and $\mathbf{P} \triangleq [\mathbf{p}_1, \dots, \mathbf{p}_K] \in \mathbb{C}^{\tau \times K}$. Finally, for each BS b and for each UE k , the distributed precoding solution is given by

$$\begin{aligned} \mathbf{w}_{b,k} &= (\mathbf{Y}_b^{\text{UL-1}} (\mathbf{Y}_b^{\text{UL-1}})^{\text{H}} + \tau (\lambda_b - \sigma_b^2) \mathbf{I}_M)^{-1} \\ &\times \mathbf{Y}_b^{\text{UL-1}} \left(\mathbf{p}_k - \sum_{\bar{b} \in \mathcal{B} \setminus \{b\}} (\mathbf{Y}_{\bar{b}}^{\text{UL-1}})^{\text{H}} \mathbf{w}_{\bar{b},k} \right). \quad (32) \end{aligned}$$

which is equal to (20) for perfect channel estimation. To compute $\mathbf{w}_{b,k}$ as in (32), BS b needs to acquire the vectors $\{(\mathbf{Y}_{\bar{b}}^{\text{UL-1}})^{\text{H}} \mathbf{w}_{\bar{b},k}\}_{\bar{b} \in \mathcal{B} \setminus \{b\}}$ via backhaul signaling [6]. This hinders the implementation of the distributed precoding design, as the amount of such signaling scales not only with the pilot length τ but also with the number of BSs B and UEs K .

V. DISTRIBUTED PRECODING DESIGN VIA OTA SIGNALING

To efficiently implement the distributed precoding design without any CSI exchange among the BSs via backhaul signaling, we introduce a *new uplink signaling resource* together with a *new CSI combining mechanism* that complement the existing uplink and downlink signaling described in Sections II-A and II-B, respectively.

To enable each BS b to acquire an estimate of $\boldsymbol{\eta}_{b,k}$ in (21), each UE k simultaneously transmits \mathbf{Y}_k^{DL} in (9) precoded with the rank-1 matrix $\mathbf{v}_k \mathbf{v}_k^{\text{H}}$, i.e.,

$$\mathbf{X}_k^{\text{UL-2}} \triangleq \mathbf{v}_k \mathbf{v}_k^{\text{H}} \mathbf{Y}_k^{\text{DL}} \in \mathbb{C}^{N \times \tau}. \quad (33)$$

Then, each BS b receives

$$\mathbf{Y}_b^{\text{UL-2}} \triangleq \sum_{k \in \mathcal{K}} \mathbf{h}_{b,k} \mathbf{v}_k^{\text{H}} \left(\sum_{\bar{b} \in \mathcal{B}} \sum_{\bar{k} \in \mathcal{K}} \mathbf{H}_{\bar{b},\bar{k}}^{\text{H}} \mathbf{w}_{\bar{b},\bar{k}} \mathbf{p}_k^{\text{H}} + \mathbf{Z}_k^{\text{DL}} \right) + \mathbf{Z}_b^{\text{UL-2}} \quad (34)$$

where $\mathbf{Z}_b^{\text{UL-2}} \in \mathbb{C}^{M \times \tau}$ is the AWGN at BS b with elements distributed as $\mathcal{CN}(0, \sigma_b^2)$. Hence, each BS b obtains

$$\frac{1}{\tau} \mathbf{Y}_b^{\text{UL-2}} \mathbf{p}_k = \sum_{\bar{b} \in \mathcal{B}} \Phi_{\bar{b}\bar{k}} \mathbf{w}_{\bar{b},k} + \frac{1}{\tau} \sum_{\bar{k} \neq k} \sum_{\bar{b} \in \mathcal{B}} \Phi_{\bar{b}\bar{k}} \mathbf{w}_{\bar{b},\bar{k}} \mathbf{p}_k^{\text{H}} \mathbf{p}_k + \mathbf{n}_{b,k}^{\text{UL-2}} \quad (35)$$

where we have defined

$$\mathbf{n}_{b,k}^{\text{UL-2}} \triangleq \frac{1}{\tau} \left(\sum_{\bar{k} \in \mathcal{K}} \mathbf{H}_{b,\bar{k}} \mathbf{v}_{\bar{k}} \mathbf{v}_{\bar{k}}^{\text{H}} \mathbf{Z}_{\bar{k}}^{\text{DL}} + \mathbf{Z}_b^{\text{UL-2}} \right) \mathbf{p}_k. \quad (36)$$

Algorithm 1 (Distributed-OTA)

Initialization: Each BS b initializes $\{\mathbf{w}_{b,k}\}_{k \in \mathcal{K}}$; set $i = 0$.

Until a predefined termination criterion is met, **do**:

(S.0) $i \leftarrow i + 1$.

(S.1) **DL:** Each BS b transmits \mathbf{X}_b^{DL} in (8); each UE k receives \mathbf{Y}_k^{DL} in (9).

(S.2) Each UE k computes its combining vector as in (28).

(S.3) **UL-1:** Each UE k transmits $\mathbf{X}_k^{\text{UL-1}}$ in (3); each BS b receives $\mathbf{Y}_b^{\text{UL-1}}$ in (4).

(S.4) **UL-2:** Each UE k transmits $\mathbf{X}_k^{\text{UL-2}}$ in (33); each BS b receives $\mathbf{Y}_b^{\text{UL-2}}$ in (34).

(S.5) For each UE k , each BS b computes its precoding vectors as in (38) and updates them as in (22).

End

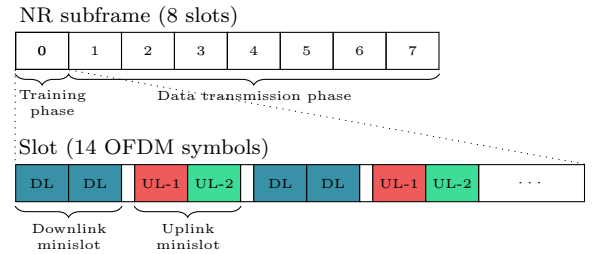


Figure 1: Example of how the proposed OTA signaling can be integrated into the flexible NR frame/slot structure.

Here, perfect channel estimation would imply that: *i*) the pilot contamination in the second term of (35) is eliminated; *ii*) as $\tau \rightarrow \infty$, the noise term $\mathbf{n}_{b,k}^{\text{UL-2}}$ in (36) is eliminated. Therefore, (35) can be used as an estimate of $\sum_{\bar{b} \in \mathcal{B}} \Phi_{\bar{b}\bar{k}} \mathbf{w}_{\bar{b},k}$ and each BS b can acquire an estimate of $\boldsymbol{\eta}_{b,k}$ in (21) as in (37) at the top of the page. Finally, for each BS b and for each UE k , the distributed precoding solution via OTA signaling is obtained as

$$\begin{aligned} \mathbf{w}_{b,k} &= (\mathbf{Y}_b^{\text{UL-1}} (\mathbf{Y}_b^{\text{UL-1}})^{\text{H}} + \tau (\lambda_b - \sigma_b^2) \mathbf{I}_M)^{-1} (\mathbf{Y}_b^{\text{UL-1}} \\ &\times (\mathbf{p}_k + (\mathbf{Y}_b^{\text{UL-1}})^{\text{H}} \mathbf{w}_{b,k}) - \mathbf{Y}_b^{\text{UL-2}} \mathbf{p}_k - \tau \sigma_b^2 \mathbf{w}_{b,k}) \quad (38) \end{aligned}$$

which is equal to (20) for perfect channel estimation. The iterative implementation of the distributed precoding design via OTA signaling is formalized in Algorithm 1. In [9], the proposed OTA signaling mechanism has been adapted for the uplink scenario to enable distributed joint receiver design.

Remark 2. The amount of OTA signaling does not scale with the number of BSs B and UEs K , and depends only on the pilot length τ and on the number of bi-directional training iterations. Note that the signaling overhead becomes negligible for a sufficiently large scheduling block size (see [7] for more details). The optimal tradeoff between OTA signaling and performance has been studied in the longer version of this paper [8].

Remark 3. The overall bi-directional training (comprising DL, UL-1, and UL-2) can be easily integrated into the New Radio (NR) frame/slot structure [7], [8]. Each of the 10 NR

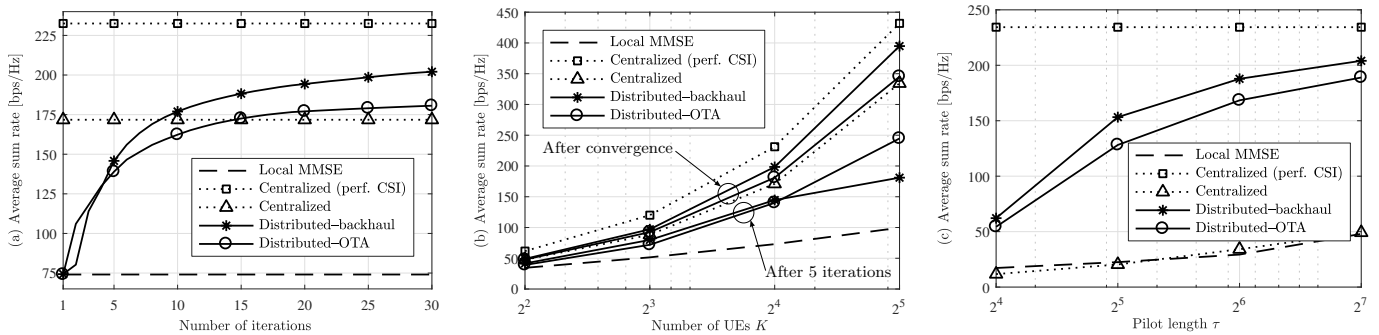


Figure 2: Average sum rate versus: (a) number of iterations with orthogonal pilots and $K = 16$; (b) number of UEs with orthogonal pilots; and (c) pilot length with non-orthogonal random pilots and $K = 16$.

subframes consists of 8 slots of 14 OFDM symbols and can be conveniently divided into training and data transmission phase. Here, minislots of minimum 2 OFDM symbols can be flexibly constructed to accommodate either uplink (i.e., UL-1 and/or UL-2) or downlink (i.e., DL) training, as in the example in Figure 1. The number and placement of training minislots can be adjusted based on rate and latency demands.

Remark 4. To comply with the UE transmit power constraint during the uplink training, $\mathbf{X}_k^{\text{UL-1}}$ in (3) and $\mathbf{X}_k^{\text{UL-2}}$ in (33) can be multiplied by a scaling factor based on, e.g., the average received power at all the UEs, and scaled back at each BS b upon receiving $\mathbf{Y}_b^{\text{UL-1}}$ in (4) and $\mathbf{Y}_b^{\text{UL-2}}$ in (34) (as done in Section VI).

VI. NUMERICAL RESULTS

The simulation setup consists of $B = 25$ BSs with $M = 4$ antennas and transmit power $\rho_{\text{BS}} = 30$ dBm, placed on a square grid with inter-site distance of 100 m, serving $K = 16$ randomly placed UEs with $N = 2$ antennas and transmit power $\rho_{\text{UE}} = 20$ dBm. The channel model includes i.i.d. Rayleigh fading and power-law pathloss with pathloss exponent 3.67. The AWGN powers are fixed to $\sigma_b^2 = -95$ dBm, $\forall b \in \mathcal{B}$, and $\sigma_k^2 = -95$ dBm, $\forall k \in \mathcal{K}$. We evaluate the average sum rate obtained via Monte Carlo simulations with 10^3 independent UE drops and compare the distributed precoding design via OTA signaling (Distributed-OTA) with the local MMSE precoding (Local MMSE), the centralized precoding design (Centralized), and the distributed precoding design via backhaul signaling [6] (Distributed-backhaul); for the latter, we consider a delay of one iteration in the CSI exchange among the BSs.

Let us begin by considering orthogonal pilots. Figure 2(a) shows that the proposed Distributed-OTA achieves a 100% performance increase with respect to the Local MMSE after just 6 iterations and reaches the performance of the Centralized within 15 iterations. Here, the distributed precoding designs can outperform the Centralized under imperfect CSI: this is mainly due to the noise averaging during the bi-directional training, whereas the Centralized relies on a single pilot-aided CSI acquisition. Furthermore, the faster convergence of the Distributed-OTA during the first few iterations with respect to the Distributed-backhaul is due to delayed backhaul update. Figure 2(b) shows that the performance gap between the proposed Distributed-OTA (even after just 5 iterations) and the Local MMSE increases with the number of UEs K . In

particular, in case of very high spatial load, the Distributed-OTA also outperforms the Distributed-backhaul. Considering non-orthogonal random pilots, Figure 2(c) shows that the Centralized (with orthogonal antenna-specific pilots within each UE) has a similar performance to the Local MMSE due to the pilot contamination in (7). Differently, in the distributed precoding designs, the combining and precoding vectors are directly estimated at each bi-directional training iteration as detailed in Section IV. This provides greatly improved resilience against pilot contamination (as also demonstrated in [6]) and the ideal performance can be approached by increasing the pilot length.

VII. CONCLUSIONS

We proposed a distributed precoding design for cell-free massive MIMO that does not require any CSI exchange among the BSs via backhaul signaling. Remarkably, the amount of involved OTA signaling does not scale with the number of BSs or UEs. The proposed method is shown to outperform the local MMSE precoding (providing a 100% gain in the average sum rate after just 6 iterations for the considered scenario) and even the centralized precoding design under realistic CSI acquisition.

REFERENCES

- [1] G. Interdonato, E. Björnson, H. Q. Ngo, P. Frenger, and E. G. Larsson, "Ubiquitous cell-free massive MIMO communications," *EURASIP J. Wireless Commun. and Networking*, vol. 2019, pp. 197–209, Aug. 2019.
- [2] S. Buzzi, C. D'Andrea, A. Zappone, and C. D'Elia, "User-centric 5G cellular networks: Resource allocation with the cell-free massive MIMO approach," *IEEE Trans. Wireless Commun.*, pp. 1–1, 2019.
- [3] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1834–1850, Mar. 2017.
- [4] E. Nayebi, A. Ashikhmin, T. L. Marzetta, H. Yang, and B. D. Rao, "Precoding and power optimization in cell-free massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4445–4459, July 2017.
- [5] E. Björnson and L. Sanguinetti, "Making cell-free massive MIMO competitive with MMSE processing and centralized implementation," *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 77–90, Jan. 2020.
- [6] J. Kaleva, A. Tölli, M. Juntti, R. A. Berry, and M. L. Honig, "Decentralized joint precoding with pilot-aided beamformer estimation," *IEEE Trans. Signal Process.*, vol. 66, no. 9, pp. 2330–2341, May 2018.
- [7] A. Tölli, H. Ghauch, J. Kaleva, P. Komulainen, M. Bengtsson, M. Skoglund, M. Honig, E. Lahetkangas, E. Tirola, and K. Pajukoski, "Distributed coordinated transmission with forward-backward training for 5G radio access," *IEEE Commun. Mag.*, vol. 57, no. 1, pp. 58–64, Jan. 2019.
- [8] I. Atzeni, B. Gouda, and A. Tölli, "Distributed precoding design via over-the-air signaling for cell-free massive MIMO," 2020. [Online]. Available: <https://arxiv.org/pdf/2004.00299.pdf>
- [9] —, "Distributed joint receiver design for uplink cell-free massive MIMO," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Dublin, Ireland, June 2020.