

# Optimal Massive-MIMO-Aided Clustered Base-station Coordination

Xueru Li, *Member, IEEE*, Xu Zhang, *Member, IEEE*, Yongxing Zhou, *Senior Member, IEEE* and Lajos Hanzo, *Fellow, IEEE*

**Abstract**—A large-scale clustered massive MIMO network is proposed for improving the spectral efficiency of the next-generation wireless infrastructure by maximizing its sum-rate. Our solution combines the advantages of the centralized processing architecture and massive MIMO. Explicitly, the network is divided into multiple clusters; each cluster is handled by a centralized processing unit, which connects to a certain number of massive MIMO-aided BSs, where only limited information is exchanged among the clusters; each user of a cluster can be served by several nearby BSs in a user-centric way.

We analyze the maximum sum-rate of the network with multiple antennas at BSs and UEs, relying on the optimal transmit precoder matrix of each BS configured for each user, and on the optimal frequency-domain power sharing scheme of each cluster. The optimizations are conceived for multiple coordination schemes that were widely studied in literature, namely the coherent-joint-transmission (CJT) scheme, the non-coherent-joint-transmission (NCJT) scheme and the coordinated-beamforming/scheduling (CBF/CS) scheme. Our simulations show that the optimal CJT achieves 2.2 – 4.5 times higher average sum-rate than its non-cooperative massive MIMO network counterpart, while the optimal NCJT and the optimal CBF/CS achieve at most a factor 1.3 average sum-rate gain. The popular signal-to-leakage-and-noise-ratio (SLNR) scheme is also extended to the multi-antenna UE scenario and achieves a factor 1.1 – 1.2 gain.

**Index Terms**—Large-scale clustered MIMO network, cell free MIMO, weighted sum rate maximization, optimal precoding matrix, optimal power sharing.

## I. INTRODUCTION

In the next-generation era, operators aim for a substantially improved user experience with an extremely high data rate and low latency, which are far better than those provided by previous generations of mobile broadband systems. In this context, the first launched use-case, namely the Enhanced Mobile Broadband scenario, is expected to provide flawless access to multi-media content and to emerging applications such as high quality video monitoring, augmented and virtual reality, just to name a few. It is expected that compelling new applications will attract new subscribers faster than in LTE.

Coordinated multi-point transmission (CoMP) [1] is known as a key technique of supporting high spectral efficiency in scenarios generating heavy tele-traffic. In LTE, CoMP has

been proposed and discussed in the 3rd Generation Partnership Project (3GPP) since Release-11 [2]. Furthermore, CoMP was also widely discussed in [1], [3], [4], which coordinates the transmission and/or reception of multiple BSs to form a large virtual antenna array for achieving macro-diversity gain, beamforming gain and for mitigating the intercell interference. Sophisticated coordinated scheduling/beamforming (CS/CBF), dynamic point selection (DPS), non-coherent joint transmission (NCJT) and coherent joint transmission (CJT) were also proposed in 3GPP LTE [5]. However, due to the associated implementation issues, such as excessive complexity and information exchange delay, as well as the limited exchange capacity of the different BSs, these sophisticated technologies often fail to achieve much improved spectral efficiency owing to these practical limitations [6].

Bearing in mind the practical limitations mentioned above, there are two directions for researchers to further improve the spectral efficiency. One of these is the further enhancement of the massive MIMO solution by primarily increasing the number of co-located antennas at each BS for improving the spatial multiplexing gain and the beamforming gain [7], [8]. This solution is capable of avoiding a high-capacity backhaul among the BSs. However, there is a practical limit on the number of antennas at each BS, owing to the increased weight and size of the antenna panel, which further increases the deployment cost.

The other direction relies on the cell-free MIMO concept [9], [10], [11], [12], [13], [14], which is similar to the early-proposed concept of distributed-antenna-system (DAS) [15], [16], [17]. The cell-free MIMO system represents a cooperative network relying on a large number of distributed access points (APs), where each UE is connected to and co-served by the several nearby APs. Each AP only has a small number of antennas (e.g., 2), and the number of APs can be much higher than the number of UEs in this area. Since a UE has a short distance to at least one of its serving APs, the quality of UE experience in this area is more uniform than that in macro cellular networks. To avoid the centralized calculation of the transmit precoder (TPC) weights, Björnson and Sanguinetti [18] proposed a scalable framework relying on the distributed SLNR precoding scheme and a UE-centric clustering method. This method mitigates the backhaul rate and reduces the computational burden imposed on the CPU.

Although the cell-free MIMO system has an appealingly light-weight architecture for the APs, when relying on distributed TPC matrix calculation and a low number of antennas per AP, having a CPU becomes necessary for carrying out

Xueru Li, Xu Zhang and Yongxing Zhou are with Huawei Technologies Co., Ltd., Shenzhen, 310051, China, (e-mail: lixueru2@huawei.com, criss.zhang@huawei.com, yongxing.zhou@huawei.com). Lajos Hanzo is with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK (e-mail: lh@ecs.soton.ac.uk).

L. Hanzo would like to acknowledge the financial support of the Engineering and Physical Sciences Research Council projects EP/N004558/1, EP/P034284/1, EP/P034284/1, EP/P003990/1 (COALESCE) as well as of the European Research Council's Advanced Fellow Grant QuantCom.

the associated scheduling strategy and resource allocation (RA) decisions, so that the SLNR precoder weight can be determined. This implies that the cell-free MIMO system is also a centralized system, and the computational complexity incurred by the scheduling and RA decisions can still be heavy when the number of UEs is high. This also means that the cell-free MIMO is not readily a system without cell boundaries, since having instantaneous information exchange among CPUs is not practical, which hence effectively results in cell boundaries defined by the APs connected to the same CPUs. Finally, constructing more APs may become costly in metropolitan areas.

But thanks to the advanced cloud/centralized-RAN (C-RAN) structure [19], constructing a centralized cooperation-aided network is feasible in practical scenarios [20], which facilitates the cooperation among a large number of BSs. Given a C-RAN structure [19], a CPU responsible for baseband signal processing is used for facilitating connection to a large number of distributed BSs. Thus, the baseband signal gleaned from all connected BSs can be processed in the CPU and excessive information exchanges among the BSs through the backhubs can be avoided. This type of network architecture has drawn attention of industry and is believed to be a promising direction of wireless network evolution.

Thus, we set out to reconsider cooperative transmissions for increasing the spectral efficiency of next-generation systems, under the concept of a clustering and co-processing aided C-RAN structure. In this paper, we target such a system, and refer to it as a large-scale clustered massive-MIMO(m-MIMO)-aided network. Specifically, the network is partitioned into multiple clusters, and each cluster is managed by a centralized processing unit, which is connected to a large number of massive MIMO-aided BSs. The connection between the CPU and the BSs is constrained by practical aspects such as a limited processing capability, the specific CPU and BS locations, and how long the associate fiber connection is. Thus, this cluster has to be formed in a CPU-centric way. This concept was inspired by [21], [28]. The association between a UE and a nearby BS is based on their channel condition and thus can be UE-centric. The associated BSs of a UE can co-process the UE's data so that a significantly increased signal

power is received by a large virtual antenna array, and the cooperation across the entire CPU-controlled area substantially reduces the interference with the aid of the TPC design, the power sharing and UE scheduling. Since a CPU can only coordinate a certain number of BSs, encountering some inter-cluster interference remains inevitable. Nonetheless, given a high spatial multiplexing gain/beamforming gain per massive MIMO-aided BS and a well-designed cooperation among BSs, the clustered massive MIMO network has a great potential.

To achieve the potential of the clustered massive MIMO network, several aspects need to be considered well. Existing literatures have shed light on the design for uplink (UL) and downlink (DL) cooperative networks. The authors of [21], [28], [22] proposed a Fog-RAN network structure and discussed the optimal UL RA in the frequency domain, relying on both an independent and a joint decoding scheme. These impressive contributions have also quantified the impact of cluster-size on the system performance: full scale cooperation may attain a factor of three UL throughput enhancement. The cell-free architecture relying on the above-mentioned SLNR precoder was characterized in terms of its achievable DL rate in [18]. The zero-forcing (ZF) precoder's DL energy efficiency was studied in [23], [24], while the minimum mean squared error (MMSE) receiver and the large scale fading detection (LSFD) receiver was investigated in [25] in terms of its UL rate. However, the optimal scheme that maximizes the system's sum-rate is unknown at the time of writing. The precoder design of [26], [27] conceived for the C-RAN architecture succeeded in maximizing the sum-rate with the aid of search algorithms, but without providing any closed-form expression [26]. By contrast, the authors of [27] relied on experience-based parameter selection for finding a stationary point of the sum-rate optimization problem. Their precoder designs are also based on a sparse prerequisite that each UE is served by only a small proportion of BSs in a cluster, which cannot exploit the full potential of cooperative networks. They also assumed having a single data stream per UE, while

<sup>1</sup>"Max" and "THP" represent "Maximizing" and "throughput", respectively.

<sup>2</sup>"EE" represents "Energy efficiency".

<sup>3</sup>"CF exp" represent "Closed-form expression".

<sup>4</sup>"C" and "D" represent "Centralized" and "Distributed", respectively.

TABLE I: Summary of literature survey on the design of cooperative network.

|   | [21]                    | [22]       | [18]               | [23], [24]             | [25]              | [26], [27] | Proposed   |
|---|-------------------------|------------|--------------------|------------------------|-------------------|------------|------------|
| UL/DL   | UL                      | UL         | DL                 | DL                     | UL                | DL         |            |
| Target  | Max THP <sup>1</sup> UL | Max THP UL | Scalable framework | Max EE <sup>2</sup> DL | Low complexity Rx | Max THP DL | Max THP DL |
| Multi-antennas at BS                          | ✓                       | ×          | ✓                  | ✓                      | ×                 | ✓          | ✓          |
| Multi-streams for a UE                        | ×                       | ×          | ×                  | ×                      | ×                 | ×          | ✓          |
| CF exp <sup>3</sup> for optimal DL TPC design | ×                       | ×          | ×                  | ×                      | ×                 | ×          | ✓          |
| C or D <sup>4</sup> DL TPC calculation        | C                       | D          | D                  | D, C                   | C                 | C          | C          |
| Joint TPC and RA optimization                 | ×                       | ×          | ×                  | ×                      | ×                 | ×          | ✓          |

next-generation solutions aim for the more ambitious goal of supporting multiple streams.

Against this background, the main contributions of this paper are as follows:

- We solve the weighted sum-rate (WSR) optimization problem of a large-scale clustered m-MIMO aided network, considering two dimensions: the spatial dimension (SD) in each frequency resource block (RB) and the power sharing in the frequency domain (FD). For the SD, the optimal CJT (O-CJT) TPC matrices are derived, assuming the CJT coordination scheme, where all BSs are connected to the same CPU for coherently precoding their data streams for serving a UE [26], [27]. As for the FD, the judicious sharing of the total transmit power of each BS among the FD RBs within the communication bandwidth is considered. To approach the performance of the O-CJT scheme, the optimization is carried out by iterating between the two dimensions.
- The TPC matrix optimization problem is converted into a convex problem, and the optimal TPC matrix expression is obtained, explicitly showing the inter-dependence of the TPCs of different BSs. Furthermore, an algorithm is proposed for calculating the numerical results. The optimal TPC scheme is also responsible for the power-sharing among the multiple data streams and UEs, as well as for determining the beamforming direction. Thus, the optimal TPC matrix represents the joint optimization of the rank decision per UE, the UE scheduling and the TPC design.
- We also investigate the performance of both the optimal CBF/CS (O-CBF/CS) and of the optimal NCJT (O-NCJT) schemes by using similar methods to those above. The CBF/CS and the NCJT were widely discussed in [5] and were also implemented in some operational networks. Therefore, it is necessary to critically appraise all the cooperative techniques to show, which should be the main technical direction for system capacity boosting in next-generation systems.
- We will demonstrate that the O-CJT scheme attains a factor 3 – 4.5 average sum-rate enhancement over a non-cooperative massive MIMO-aided network, while the other schemes achieve at most a factor 1.3 enhancement. We also simulate the performance of the SLNR TPC of [18] by further developing it to multi-antenna UEs.

The comparison between the existing studies and our proposed method are summarized in Table I, where the explanations of the acronyms are added as footnotes on the bottom of the last page.

The structure of this paper is as follows. Section II describes the system and our signal model. Section III solves the WSR maximization problem by deriving the O-CJT TPC matrix and power sharing. Section IV optimizes the weighted sum-rate both of the O-CBF/CS scheme. Finally, Section V provides our simulation results and Section VI concludes the paper.

## II. SYSTEM AND SIGNAL MODEL

### A. System model

Consider the downlink (DL) of a time-division duplex large-scale clustered MIMO network, which consists of one or more large-scale clusters, and each cluster consists of  $B$  BSs. The  $B$  BSs in the same cluster are connected to a central CPU via a flawless infinite-bandwidth and low-latency backhaul, and the CPUs of different clusters are assumed to have limited information exchange. Therefore, the TPC matrices and the power sharing of BSs in the same cluster can be calculated in a centralized way, in order to glean a beneficial coordination gain. The specific cluster a UE will connect to is selected based on channel condition. Without loss of generality, we assume that  $U$  UEs connect to each cluster. For each cluster,  $\{1, 2, \dots, B\}$  and  $\{1, 2, \dots, U\}$  represent the BS IDs and UE IDs, respectively. Each base station has  $T$  TAs and each UE has  $R$  RAs.

The DL communication bandwidth hosts a large number of sub-carriers and can be partitioned into  $K$  RBs (RB). The channel's coherence bandwidth is assumed to be as high as the bandwidth covered by a RB. Therefore, the DL channel state information (CSI) within a RB can be regarded to be similar and can be obtained at the BS by exploiting the UL-DL channel reciprocity. In OFDM systems, the sub-carriers can be treated as parallel channels, which theoretically enables independent precoding and power sharing across the different sub-carriers. However, since the FD channel is near-constant across a RB, it is reasonable to use a RB or even a few RBs for which a constant TPC matrix and power sharing design is employed. This strikes a compelling performance vs. complexity trade-off. Without loss of generality, we use one RB as the basis of our analysis. This system model applies to all the schemes that will be discussed in this paper, and a illustrative figure is added in Figure 1 and 2 to show the network architecture and the time/frequency resource definition in our paper.

### B. Signal model for the O-CJT scheme

We first describe the signal model of the O-CJT scheme, which will be shown to provide the most promising performance for the network considered. The signal model is the basis of our optimal precoding/power sharing scheme discussion in the next section. The signal model of the O-CBF/CS scheme will be given in Section IV together with the discussion of the corresponding optimal scheme.

For each UE supported by the O-CJT scheme, some or all of the  $B$  BSs in its connected cluster act as its CJT set. Let set  $\mathcal{B}_u \subseteq \{1, 2, \dots, B\}$  denote the CJT set of UE  $u$ . Then each data stream of UE  $u$  is transmitted from the BSs belonging to  $\mathcal{B}_u$ , so that the signal from these coordinated BSs arrive at the UE constructively. The BSs that do not belong to  $\mathcal{B}_u$  do not contribute to this transmission. Depending on the UE locations, the CJT sets of different UEs can be different, i.e. the determination of  $\mathcal{B}_u$  is UE-centric. In the following analysis, we assume that all BSs in a cluster act as the CJT cluster for all UEs connected to this cluster, since this allows the maximum degrees of design freedom, and we will quantify the influence

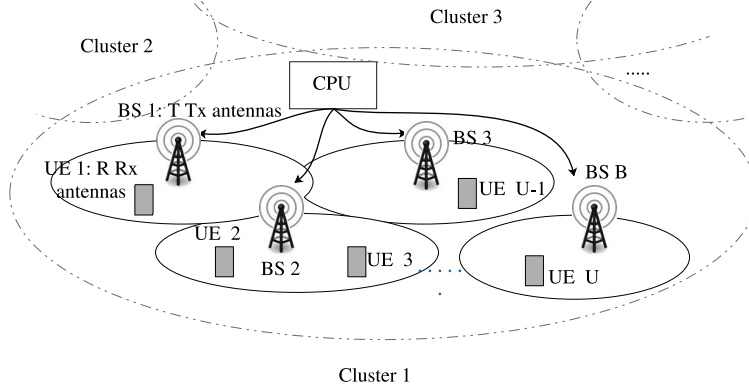


Fig. 1: The large-scale clustered MIMO-aided network.

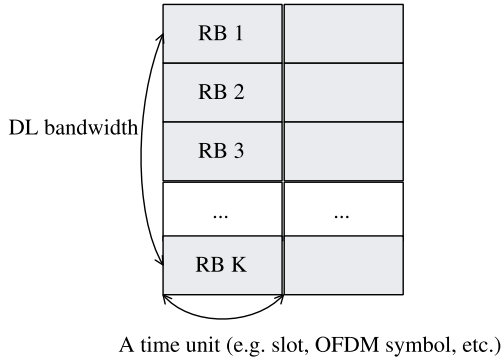


Fig. 2: An illustrative figure of the time/frequency resource definition.

of  $|\mathcal{B}_u|$  with the aid of simulations, where  $|\mathcal{X}|$  represents the number of elements in set  $\mathcal{X}$ .

Based on the above system model, the signal received at UE  $u$  on RB  $k$  is formulated as:

$$\begin{aligned}
 \mathbf{y}_{k,u} = & \underbrace{\sum_{b=1}^B \sqrt{P_{k,b}} \mathbf{H}_{k,b,u} \mathbf{W}_{k,b,u} \mathbf{s}_{k,u}}_{\text{signal part}} \\
 & + \underbrace{\sum_{b=1}^B \sum_{u' \neq u} \sqrt{P_{k,b}} \mathbf{H}_{k,b,u} \mathbf{W}_{k,b,u'} \mathbf{s}_{k,u'}}_{\text{MU interference within the cluster}} \\
 & + \underbrace{\mathbf{I}_{k,u}^{\text{inf}} + \mathbf{n}_{k,u}}_{\text{Inter-cluster interference and noise}} \in \mathbb{C}^{R \times 1}
 \end{aligned} \quad (1)$$

where

- $P_{k,b} \geq 0$  is the transmit power of BS  $b$  on RB  $k$ , and satisfies  $\sum_k P_{k,b} \leq P_{b,\max}$ , with  $P_{b,\max}$  being the maximum total transmit power of BS  $b$ .
- Matrix  $\mathbf{H}_{k,b,u} \in \mathbb{C}^{R \times T}$  represents the DL channel matrix of the link spanning from BS  $b$  to UE  $u$  on RB  $k$ . We have

$\mathbf{H}_{k,b,u} = \sqrt{\eta_{b,u}} \mathbf{H}_{k,b,u}^{\text{sf}}$  where  $\eta_{b,u}$  denotes large scale fading and  $\mathbf{H}_{k,b,u}^{\text{sf}}$  is small scale fading.

- Matrix  $\mathbf{W}_{k,b,u} \in \mathbb{C}^{T \times L_u}$  is the TPC matrix used by BS  $b$  for UE  $u$  on RB  $k$ , where  $L_u \leq R$  is the maximum number of data layers that can be scheduled for UE  $u$ .
- Vector  $\mathbf{s}_{k,u} \sim \mathcal{CN}(0, \mathbf{I}_R)$  consists of the symbols from  $L_u$  layers destined for UE  $u$  on RB  $k$ , where  $\mathbf{I}_R$  is an identity matrix of dimension  $R$ .
- $\mathbf{I}_{k,u}^{\text{inf}} \in \mathbb{C}^{R \times 1}$  and  $\mathbf{n}_{k,u} \in \mathbb{C}^{R \times 1}$  denotes the inter-cluster interference and the thermal noise at UE  $u$  on RB  $k$ , respectively.  $\mathbf{n}_{k,u} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_R)$  is assumed.

Based on (1), the DL achievable rate of UE  $u$  on RB  $k$  in its connected cluster can be expressed as:

$$R_{k,u}^{\text{cjt}} = \log_2 \det \left( \mathbf{I}_R + \bar{\mathbf{H}}_{k,u} \bar{\mathbf{W}}_{k,u} \bar{\mathbf{W}}_{k,u}^H \bar{\mathbf{H}}_{k,u}^H \Sigma_{k,u}^{-1} \right), \quad (2)$$

where  $\Sigma_{k,u}$  is the sum covariance matrix of the MU interference, inter-cluster interference and noise, formulated as:

$$\Sigma_{k,u} = \sum_{u' \neq u} \bar{\mathbf{H}}_{k,u} \bar{\mathbf{W}}_{k,u'} \bar{\mathbf{W}}_{k,u'}^H \bar{\mathbf{H}}_{k,u}^H + \mathbf{R}_{k,u}^{\text{inf}} + \sigma^2 \mathbf{I}_R, \quad (3)$$

where  $\mathbf{R}_{k,u}^{\text{inf}} = \mathbb{E}\{\mathbf{I}_{k,u}^{\text{inf}} (\mathbf{I}_{k,u}^{\text{inf}})^H\}$  is the inter-cluster interference covariance matrix. In (2) and (3), matrices  $\bar{\mathbf{H}}_{k,u}$  and  $\bar{\mathbf{W}}_{k,u}$  are defined as:

$$\bar{\mathbf{H}}_{k,u} = [\mathbf{H}_{k,1,u}, \mathbf{H}_{k,2,u}, \dots, \mathbf{H}_{k,B,u}] \in \mathbb{C}^{R \times BT}, \quad (4)$$

$$\bar{\mathbf{W}}_{k,u} = \left[ \sqrt{P_{k,1}} \mathbf{W}_{k,1,u}, \dots, \sqrt{P_{k,B}} \mathbf{W}_{k,B,u} \right]^H \in \mathbb{C}^{BT \times L_u}. \quad (5)$$

Based on (2), it can be seen that to maximize the achievable rate of UE  $u$ , namely  $\sum_k R_{k,u}^{\text{cjt}}$ , the O-CJT TPC matrices  $\mathbf{W}_{k,b,u}$  and the power sharing scheme  $P_{k,b}$  of each BS has to be well designed, so as to maximize the signal power contributed by all the coordinated BSs, and to minimize the MU interference in the cluster. The inter-cluster interference  $\mathbf{I}_{k,u}^{\text{inf}}$  will change when carrying on the optimization independently for each cluster. However, in this paper we assume that the inter-cluster interference remains constant

during the optimization, to avoid information exchange among the clusters. The initial information concerning the level of the inter-cluster interference can be obtained from the UE reports, based on measuring the interference level during unallocated time-slots.

### III. JOINT OPTIMIZATION PROBLEM FOR O-CJT BASED LARGE-SCALED CLUSTERED MASSIVE MIMO SYSTEM

The optimization problem we target in this section is to design a joint optimization scheme for the O-CJT TPC matrix and power sharing for each BS in each cluster, so that the achievable weighted sum rate of all UEs in this cluster can be maximized. Explicitly, the weighted sum rate optimization problem can be modeled as:

$$(\mathcal{P}_1) : \begin{aligned} & \max_{\{\mathbf{W}_{k,b,u}, P_{k,b}\}} \sum_{u=1}^U \sum_k w_u R_{k,u}^{\text{cjt}} \\ & s.t. \quad \sum_{k=1}^K P_{k,b} \leq P_{b,\max}, \quad P_{k,b} \geq 0, \quad \forall b, k \\ & \quad \sum_{u=1}^U \|\mathbf{W}_{k,b,u}\|^2 \leq P_{k,b}, \quad \forall b, k \end{aligned} \quad (6)$$

where  $R_{k,u}^{\text{cjt}}$  is as in (2), and  $w_u \geq 0$  is the weight of UE  $u$ .

It can be seen that in  $(\mathcal{P}_1)$ ,  $P_{k,b}$  plays the role of judiciously sharing the total power  $P_{b,\max}$  in the FD, while  $\mathbf{W}_{k,b,u}$  plays the role of both beamforming-aided transmission to the target UE, and sharing  $P_{k,b}$  among the serving UEs and the layers of each serving UE on RB  $k$ . If BS  $b$  decides not to serve UE  $u$ , then we have  $\|\mathbf{W}_{k,b,u}\|^2 = 0$ . Similarly, if BS  $b$  decides to schedule  $L < R$  layers for UE  $u$ , then the  $l$ -th column of  $\mathbf{W}_{k,b,u}$  satisfies  $\|\mathbf{w}_{k,b,u,l}\|^2 = 0$ , ( $l = L_u + 1, \dots, R$ ), where  $\mathbf{w}_{k,b,u,l}$  is the  $l$ -th column of matrix  $\mathbf{W}_{k,b,u}$ .

To optimize the sum-rate, the power sharing and the TPC matrix have to be optimized jointly. However, their direct joint optimization is a non-convex and non-trivial problem. Thus, iterative optimization is carried out by iterating between power sharing and TPC matrix. In Sub-section III-A, the O-CJT TPC matrix is derived for a fixed power sharing scheme  $P_{k,b}$ . Then in Sub-section III-B, the optimal power sharing scheme is derived for a fixed O-CJT TPC matrix.

#### A. Optimization of the O-CJT TPC Matrices

For a fixed power sharing scheme  $P_{k,b}$ , the O-CJT TPC matrix optimization under per-BS power constraints becomes

$$(\mathcal{P}_2) : \begin{aligned} & \max_{\{\mathbf{W}_{k,b,u}\}} \sum_u \sum_k w_u R_{k,u}^{\text{cjt}} \\ & s.t. \quad \sum_{u=1}^U \|\mathbf{W}_{k,b,u}\|^2 \leq P_{k,b}, \quad \forall b, k. \end{aligned} \quad (7)$$

Note that the optimization is not convex w.r.t  $\mathbf{W}_{k,b,u}$ , thus we have to transform it into a convex problem.

1) *Converting  $(\mathcal{P}_2)$  into a convex problem:* The idea of converting  $(\mathcal{P}_2)$  into a convex problem was inspired by the method proposed in [29]. The detailed procedure is illustrated as follows.

According to Lemma 13.8 of [30], the achievable rate of UE  $u$  on RB  $k$  can be expressed as

$$R_{k,u}^{\text{cjt}} = I(\mathbf{y}_{k,u}; \mathbf{s}_{k,u}) = \max_{q(\mathbf{s}_{k,u}|\mathbf{y}_{k,u})} \mathbb{E} \left[ \log_2 \frac{q(\mathbf{s}_{k,u}|\mathbf{y}_{k,u})}{p(\mathbf{s}_{k,u})} \right], \quad (8)$$

where  $I(\mathbf{y}_{k,u}; \mathbf{s}_{k,u})$  is the mutual information between  $\mathbf{y}_{k,u}$  and  $\mathbf{s}_{k,u}$ , and  $q(\mathbf{s}_{k,u}|\mathbf{y}_{k,u})$  is the test *a posteriori* probability. For a fixed O-CJT TPC  $\mathbf{W}_{k,b,u}$ , the optimal  $q(\mathbf{s}_{k,u}|\mathbf{y}_{k,u})$  is:

$$q^{\text{opt}}(\mathbf{s}_{k,u}|\mathbf{y}_{k,u}) = p(\mathbf{s}_{k,u}|\mathbf{y}_{k,u}) = \mathcal{CN}(\hat{\mathbf{s}}_{k,u}, \mathbf{\Xi}_{k,u}), \quad (9)$$

where  $\hat{\mathbf{s}}_{k,u} = \mathbf{G}_{k,u}^H \mathbf{y}_{k,u} \in \mathbb{C}^{L_u \times 1}$  is the symbol vector post-processed by using the receiver matrix  $\mathbf{G}_{k,u}^H \in \mathbb{C}^{L_u \times R}$ , and  $\mathbf{\Xi}_{k,u}$  is the covariance matrix of  $\hat{\mathbf{s}}_{k,u}$ . For a fixed TPC matrix, the optimal linear receiver matrices is MMSE receiver matrix that can be written as

$$\mathbf{G}_{k,u}^H = \bar{\mathbf{W}}_{k,u}^H \bar{\mathbf{H}}_{k,u}^H \left( \sum_{u'=1}^U \bar{\mathbf{H}}_{k,u'} \bar{\mathbf{W}}_{k,u'} \bar{\mathbf{W}}_{k,u'}^H \bar{\mathbf{H}}_{k,u'}^H + \sigma^2 \mathbf{I}_R \right)^{-1}. \quad (10)$$

Given the above MMSE receiver matrix, the covariance matrix  $\mathbf{\Xi}_{s,u}$  is formulated as

$$\begin{aligned} \mathbf{\Xi}_{s,u} &= \mathbb{E} \left[ (\hat{\mathbf{s}}_{k,u} - \mathbf{s}_{k,u}) (\hat{\mathbf{s}}_{k,u} - \mathbf{s}_{k,u})^H \right] \\ &= \mathbf{I}_{L_u} - \bar{\mathbf{W}}_{k,u}^H \bar{\mathbf{H}}_{k,u}^H \mathbf{G}_{k,u}. \end{aligned} \quad (11)$$

Then based on [29], the achievable rate of UE  $u$  on RB  $k$  can be written as

$$\begin{aligned} R_{k,u}^{\text{cjt}} &= \mathbb{E} \left[ \log_2 \frac{q^{\text{opt}}(\mathbf{s}_{k,u}|\mathbf{y}_{k,u})}{p(\mathbf{s}_{k,u})} \right] \\ &= -\text{tr} \left( \mathbf{G}_{k,u} \mathbf{\Xi}_{k,u}^{-1} \mathbf{G}_{k,u}^H \mathbf{J}_{k,u} \right) - \text{tr} \left( \mathbf{\Xi}_{k,u}^{-1} \right) - \log_2 \det(\mathbf{\Xi}_{k,u}) \\ &\quad + L_u + 2\text{Re} \left\{ \text{tr} \left( \mathbf{\Xi}_{k,u}^{-1} \mathbf{G}_{k,u}^H \bar{\mathbf{H}}_{k,u} \bar{\mathbf{W}}_{k,u} \right) \right\} \triangleq \bar{R}_{k,u}^{\text{cjt}} \end{aligned} \quad (12)$$

based on the optimal test *a posteriori* probability, given a fixed TPC matrix and the assumption of  $\mathbf{s}_{k,u} \sim \mathcal{CN}(0, \mathbf{I}_{L_u})$ , where  $\mathbf{J}_{k,u} = \sum_{u'=1}^U \bar{\mathbf{H}}_{k,u'} \bar{\mathbf{W}}_{k,u'} \bar{\mathbf{W}}_{k,u'}^H \bar{\mathbf{H}}_{k,u'}^H + \sigma^2 \mathbf{I}_R$ .

Equation (12) shows that the maximal achievable rate depends on TPC matrix, receiver matrix and  $\mathbf{\Xi}_{k,u}$ . However, this joint optimization problem is non-convex. To handle this kind of problem, an iterative optimization procedure alternating between the TPC matrix and the receiver matrix was widely used in the literature [29], [31]. We apply this alternating optimization method to solve  $(\mathcal{P}_2)$ . Given a fixed receiver matrix  $\mathbf{G}_{k,u}^H$  and the covariance matrix  $\mathbf{\Xi}_{k,u}$ ,  $(\mathcal{P}_2)$  can formally be proved to be a convex optimization problem w.r.t  $\mathbf{W}_{k,b,u}$ . Thus, the classic KarushCKuhnCTucker (KKT) conditions can be used for obtaining the globally O-CJT TPC matrix [32].

2) *KKT conditions and the algorithm proposed for solving problem  $(\mathcal{P}_2)$ :* To arrive at the KKT conditions, the Lagrangian function of  $(\mathcal{P}_2)$  is obtained first as (13) on top of the next page, where  $\lambda_{k,b} \geq 0$  is the Lagrangian multiplier

$$F_1 = \sum_k \sum_u w_u \left[ \text{tr} \left( \mathbf{G}_{k,u} \mathbf{\Xi}_{k,u}^{-1} \mathbf{G}_{k,u}^H \mathbf{J}_{k,u} \right) \text{tr} \left( \mathbf{\Xi}_{k,u}^{-1} \right) - 2\text{Re} \left\{ \text{tr} \left( \mathbf{\Xi}_{k,u}^{-1} \mathbf{G}_{k,u}^H \bar{\mathbf{H}}_{k,u} \bar{\mathbf{W}}_{k,u} \right) \right\} + \log_2 \det \left( \mathbf{\Xi}_{k,u} \right) - L_u \right] \\ + \sum_{b,k} \lambda_{k,b} \left( \sum_u \text{tr} \left( \mathbf{W}_{k,b,u} \mathbf{W}_{k,b,u}^H \right) - 1 \right). \quad (13)$$

related to the sum power constraint of TPC matrix on each RB for each BS:

Then the partial derivative of  $F_1$  w.r.t  $\mathbf{W}_{k,b,u_0}$  can be obtained as:

$$\frac{\partial F_1}{\partial \mathbf{W}_{k,b,u_0}} = \left( \sum_u w_u P_{k,b} \mathbf{X}_{k,b,b',u} + \lambda_{k,b} \mathbf{I}_T \right) \mathbf{W}_{k,b,u_0} \\ + \sum_u \sum_{b' \neq b} w_u \sqrt{P_{k,b} P_{k,b'}} \mathbf{X}_{k,b,b',u} \mathbf{W}_{k,b',u_0} - w_{u_0} \sqrt{P_{k,b}} \mathbf{Z}_{k,b,u_0}, \quad (15)$$

where

$$\begin{cases} \mathbf{X}_{k,b,b',u} = \mathbf{H}_{k,b,u}^H \mathbf{G}_{k,u} \mathbf{\Xi}_{k,u}^{-1} \mathbf{G}_{k,u}^H \mathbf{H}_{k,b',u} & \in \mathbb{C}^{T \times T} \\ \mathbf{Z}_{k,b,u} = \mathbf{H}_{k,b,u}^H \mathbf{G}_{k,u} \mathbf{\Xi}_{k,u}^{-1} & \in \mathbb{C}^{T \times L}. \end{cases} \quad (16)$$

Setting  $\frac{\partial F_1}{\partial \mathbf{W}_{k,b,u_0}} = \mathbf{0}$ , we obtain the condition that the O-CJT TPC matrix should satisfy:

$$\mathbf{W}_{k,b,u_0}^* = (P_{k,b} \mathbf{Y}_{k,b,b'} + \lambda_{k,b} \mathbf{I}_T)^{-1} \times \\ \left( w_{u_0} \sqrt{P_{k,b}} \mathbf{Z}_{k,b,u_0} - \sum_{b' \neq b} \sqrt{P_{k,b} P_{k,b'}} \mathbf{Y}_{k,b,b'} \mathbf{W}_{k,b',u_0}^* \right), \quad (17)$$

where  $\mathbf{Y}_{k,b,b'} = \sum_u w_u \mathbf{X}_{k,b,b',u}$ . According to the KKT condition,  $\lambda_{k,b}$  and  $\mathbf{W}_{k,b,u}$  should also satisfy:

$$\begin{cases} \sum_{u=1}^U \|\mathbf{W}_{k,b,u}\|^2 \leq P_{k,b}, \\ \left( \sum_{u=1}^U \|\mathbf{W}_{k,b,u}\|^2 - P_{k,b} \right) \lambda_{k,b} = 0, \\ \lambda_{k,b} \geq 0. \end{cases} \quad (18)$$

Equation (17) shows that for UE  $u_0$ , the O-CJT TPC matrix of BS  $b$  is determined by the channel matrix  $\mathbf{H}_{k,b,u}$ , receiver matrix  $\mathbf{G}_{k,u}^H$ , covariance matrix  $\mathbf{\Xi}_{k,u}$ , and most importantly, the TPC matrix of the other BSs  $b'$ . Hence, for a specific UE, the O-CJT TPC matrices of the BSs belonging to the same CJT cluster are mutually coupled. That is not unexpected since the BSs of the same CJT cluster coherently transmit a data layer to a specific UE, and a change in the TPC matrix of a BS will definitely have an impact on the TPC matrix of another BS. Therefore, alternatively calculating the O-CJT TPC matrix of each BS is proposed to get numerical results. At the  $t$ -th iteration, the TPC matrix of each BS  $\mathbf{W}_{k,b,u}^{(t)}$  is updated to  $\mathbf{W}_{k,b,u}^{(t+1)}$  according to the following equation:

$$\mathbf{W}_{k,b,u_0}^{(t+1)} = \left( P_{k,b} \mathbf{Y}_{k,b,b'} + \lambda_{k,b}^{(t+1)} \mathbf{I}_T \right)^{-1} \left( w_{u_0} \sqrt{P_{k,b}} \mathbf{Z}_{k,b,u_0} \right. \\ \left. - \sum_{b' \neq b} \sqrt{P_{k,b} P_{k,b'}} \mathbf{Y}_{k,b,b'} \mathbf{W}_{k,b',u_0}^{(t)} \right), \quad (19)$$

where  $\lambda_{k,b}^{(t+1)}$  should be selected for ensuring that the constraints in (18) are satisfied. For example, if  $\sum_u \|\mathbf{W}_{k,b,u}^{(t+1)}\|^2 \leq P_{k,b}$  is satisfied with  $\lambda_{k,b}^{(t+1)} = 0$ , then  $\lambda_{k,b}^{(t+1)} = 0$ . Otherwise, a bisection algorithm is used for finding the numerical value

**Algorithm 1** Iterative algorithm for finding the O-CJT TPC matrix and receiver matrix

**Input:**

- 1: Channel matrix  $\mathbf{H}_{k,b,u}$  for all  $(k, b, u)$ ; UE weights  $w_u$  for all  $u$ ; power sharing scheme  $P_{k,b}$  for all  $(k, b)$ .
- 2: Initial values of  $\mathbf{W}_{k,b,u}^{(n=0)}$  for all  $(k, b, u)$ ; initial values  $\delta_1 = 10$  and  $\delta_2 = 10$ ; thresholds  $\Delta_1$  and  $\Delta_2$ .

**Output:** Optimal TPC matrix  $\mathbf{W}_{k,b,u}$  and optimal receiver matrix  $\mathbf{G}_{k,u}^H$  for all  $(k, b, u)$ .

- 3: Calculate the sum rate  $R^{(n)} = \sum_u R_{k,u}^{(\text{cjt})^{(n)}}$  based on (2) with  $\mathbf{W}_{k,b,u}^{(n)}$ .
- 4: **while**  $\delta_1 > \Delta_1$  **do**
- 5:   Update receiver matrix  $\mathbf{G}_{k,u}^{H,(n)}$  and covariance matrix  $\mathbf{\Xi}_{k,u}^{(n)}$  with  $\mathbf{W}_{k,b,u}^{(n)}$ , according to (10) and (11).
- 6:   Update  $\mathbf{Z}_{k,b,u}^{(n)}$  and  $\mathbf{Y}_{k,b,b'}^{(n)}$  based on (16), with  $\mathbf{G}_{k,u}^{H,(n)}$  and  $\mathbf{\Xi}_{k,u}^{(n)}$ .
- 7:   Initialize  $t = 0$ , and  $\mathbf{W}_{k,b,u}^{(n,t)} = \mathbf{W}_{k,b,u}^{(n)}$ .
- 8:   **while**  $\delta_2 > \Delta_2$  **do**
- 9:     **for**  $b = 1 \rightarrow B$  **do**
- 10:       Update  $\mathbf{W}_{k,b,u}^{(n,t+1)}$  based on (19), with  $\mathbf{Z}_{k,b,u}^{(n)}$  and  $\mathbf{Y}_{k,b,b'}^{(n)}$ .  $\lambda_{k,b}^{(t+1)}$  is obtained from the above bisection method.
- 11:     **end for**
- 12:     Update  $\delta_2 = \max_b \sum_u \left\| \mathbf{W}_{k,b,u}^{(n,t+1)} - \mathbf{W}_{k,b,u}^{(n,t)} \right\|_F^2$ , and update  $t = t + 1$ .
- 13:   **end while**
- 14:   Update  $\mathbf{W}_{k,b,u}^{(n+1)} = \mathbf{W}_{k,b,u}^{(n,t \rightarrow \infty)}$ , where  $\mathbf{W}_{k,b,u}^{(n,t \rightarrow \infty)}$  is the converged  $\mathbf{W}_{k,b,u}^{(n,t)}$  from the above while loop.
- 15:   Update sum rate  $R^{(n+1)} = \sum_u R_{k,u}^{(\text{cjt})^{(n+1)}}$  based on (2) with  $\mathbf{W}_{k,b,u}^{(n+1)}$ .
- 16:   Update  $\delta_1 = R^{(n+1)} - R^{(n)}$ , and update  $n = n + 1$ .
- 17: **end while**

of  $\lambda_{k,b}^{(t+1)}$  until (18) is satisfied. The iterations in (19) continue until convergence is achieved.

Having acquired a converged  $\mathbf{W}_{k,b,u}^{(t+1)}$ , receiver matrix  $\mathbf{G}_{k,u}^H$  and covariance matrix  $\mathbf{\Xi}_{k,u}$  are updated according to (10) and (11). The alternating update between the TPC matrix and the receiver matrix continues until convergence is achieved. Our procedure conceived for obtaining the O-CJT TPC matrix and receiver matrix is summarized in Algorithm 1.

The convergence of Algorithm 1 is not formally proved here, but it is demonstrated by our simulations. In Section V, we show that the O-CJT TPC scheme attains a significant system throughput gain over a non-cooperative system. Ad-

ditionally, it is found that most of the maximum attainable performance gain can be achieved by as few as 2–4 iterations, which limits the computational complexity of this algorithm.

### B. Optimal Power Allocation Scheme

Having obtained the O-CJT TPC matrix and receiver matrix in Sub-section III-A, the optimal power sharing in the frequency domain can further be designed to release more potentials. The power sharing scheme deals with the problem of distributing the total transmit power of a BS among the  $K$  RBs, i.e. determines  $P_{k,b}$  under the constraint of  $\sum_k P_{k,b} \leq P_{b,\max}$ .

This optimization problem is transformed into a convex optimization problem, and an algorithm is proposed according to the KKT conditions. To make this paper more concise, we put the derivation and algorithm in our arxiv version [33]. We summarize the main conclusions here.

- When a naive TPC is used, the proposed power sharing can bring a large performance gain. For example, if a local ZF TPC is used, 7% – 30% gain over equal power sharing can be achieved. The reason is that the local ZF TPC can neither increase the coherent signal power nor reduce the intercell interference either, a well designed power sharing scheme is very helpful for enhancing the SINR.
- However, if the proposed O-CJT TPC is used, the power sharing can only provide marginal rate-gain. According to our simulations, the proposed scheme provides only 1.4% gain over equal-power sharing. This is not surprising since the O-CJT TPC already significantly increases the SINR.

Therefore, it can be concluded that when the O-CJT TPC is applied, further optimizing the power sharing in the frequency domain is not needed. This is good news since the computational complexity of the power sharing can be saved.

### C. Iterative optimization of the O-CJT TPC matrix and of the power sharing scheme

Based on the O-CJT precoding and power sharing algorithm in Subsection III-A and III-B, respectively, alternating optimization can further be proposed for the joint optimization of the O-CJT TPC and power sharing. Explicitly, the O-CJT TPC matrix is obtained based on a fixed power allocation strategy. Then, the power allocation strategy is updated based on the obtained O-CJT. The iteration goes until convergence is achieved or the maximal iteration number is reached. The fixed power allocation strategy can be initialized as equal power allocation strategy. The detailed algorithm is omitted here due to the lack of space, and can be found in our arxiv version [33].

## IV. OPTIMIZATION PROBLEM FOR O-CBF/CS BASED LARGE-SCALED CLUSTERED MASSIVE MIMO SYSTEM

In order to characterize the performance of different coordination schemes in our large-scale clustered massive MIMO system, a pair of popular schemes [34] will be analyzed in this section. They are coordinated beamforming (CBF)

and coordinated scheduling (CS). These two schemes were extensively studied in LTE [35], [36]. Briefly, in CBF a BS refrains from transmitting a certain beam relying on a TPC matrix, if this beam will inflict serious interference on the UEs in neighboring cells. By contrast, in CS a BS will assign zero power to a UE on a RB, if it will inflict serious interference on the UEs in neighboring cells. Thus, CBF and CS aim for interference coordination.

Usually, these schemes assume some level of information exchange among the BSs, in order to achieve interference coordination. However, since the backhaul typically has limited bandwidth and imposes latency, only some simple statistics based information-exchanging schemes have been discussed in the 3GPP during LTE [5], which fail to approach the optimal achievable weighted sum-rate. In this section, these schemes are revisited to obtain the O-CBF and O-CS schemes, with the motivation approaching the optimal achievable weighted sum-rate, assuming perfect centralized processing and having perfect global DL CSI just like the O-CJT scheme. Furthermore, in order to approach the system's full potential, the O-CBF and O-CS schemes will be intrinsically amalgamated by relying on the jointly optimal TPC matrices and optimal power sharing of each BS. More explicitly, the former is closely related to CBF, while the latter to CS.

The main difference between the amalgamated O-CBF/CS scheme and the O-CJT scheme is that in the former scheme a UE is only scheduled by a single BS, while in the latter scheme more than one BSs participate in the scheduling. In the operational commercial networks, a UE selects its serving BS according to the received signal power (RSP). In this paper, we follow this method and define RSP as  $\eta_{b,u} \sum_k \sqrt{P_{k,b}}/K$ .

Therefore, upon denoting the serving BS for UE  $u$  by  $b_u$ , the signal received at UE  $u$  on RB  $k$  is:

$$\begin{aligned} \mathbf{y}_{k,u} = & \sqrt{P_{k,b_u}} \mathbf{H}_{k,b_u,u} \tilde{\mathbf{W}}_{k,b_u,u} \mathbf{s}_{k,u} \\ & + \sum_{(b',u') \neq (b_u,u)} \sqrt{P_{k,b'}} \mathbf{H}_{k,b',u} \tilde{\mathbf{W}}_{k,b',u'} \mathbf{s}_{k,u'} \\ & + \tilde{\mathbf{I}}_{k,u}^{\text{inf}} + \mathbf{n}_{k,u}, \end{aligned} \quad (20)$$

where  $\tilde{\mathbf{W}}_{k,b,u} \in \mathbb{C}^{T \times L_u}$  is the O-CBF/CS TPC matrix of BS  $b$  for UE  $u$  on RB  $k$ , and  $\|\tilde{\mathbf{W}}_{k,b,u}\|^2 = 0$  if  $b \neq b'$ , while  $\tilde{\mathbf{I}}_{k,u}^{\text{inf}}$  is the inter-cluster interference under the CBF/CS scheme. The other parameters have the same meaning as in (1).

Then the achievable rate of UE  $u$  on RB  $k$  using the amalgamated O-CBF/CS scheme is:

$$\begin{aligned} R_{k,u}^{\text{cbf}} = & \log_2 \det \left( \mathbf{I}_R + P_{k,b_u} \mathbf{H}_{k,b_u,u} \tilde{\mathbf{W}}_{k,b_u,u} \tilde{\mathbf{W}}_{k,b_u,u}^H \mathbf{H}_{k,b_u,u}^H \tilde{\Sigma}_{k,u}^{-1} \right), \end{aligned} \quad (21)$$

where

$$\begin{aligned} \tilde{\Sigma}_{k,b,u}^{-1} = & \sum_{(b',u') \neq (b,u)} P_{k,b'} \mathbf{H}_{k,b',u} \tilde{\mathbf{W}}_{k,b',u'} \tilde{\mathbf{W}}_{k,b',u'}^H \mathbf{H}_{k,b',u}^H + \tilde{\mathbf{R}}_{k,u}^{\text{inf}} + \sigma^2 \mathbf{I}_R, \end{aligned} \quad (22)$$

and  $\tilde{\mathbf{R}}_{k,u}^{\text{inf}} = \mathbb{E}\{\tilde{\mathbf{I}}_{k,u}^{\text{inf}} (\tilde{\mathbf{I}}_{k,u}^{\text{inf}})^H\}$ .

Hence, the weighted sum-rate maximization problem of the O-CBF/CS scheme can be modelled as:

$$\begin{aligned}
 (\mathcal{P}_5) : \quad & \max_{\{\mathbf{w}_{k,b,u}, P_{k,b}\}} \sum_{u=1}^U \sum_{k=1}^K w_u R_{k,u}^{\text{cbf}} \\
 \text{s.t.} \quad & \sum_{k=1}^K P_{k,b} \leq P_{b,\max}, \quad P_{k,b} \geq 0, \forall b, \\
 & \sum_{u=1}^U \|\tilde{\mathbf{w}}_{k,b,u}\|^2 \leq P_{k,b}, \forall b, k
 \end{aligned} \quad (23)$$

Although the expression of  $R_{k,u}^{\text{cbf}}$  is different from that of  $R_{k,u}^{\text{cjt}}$ , the derivation of the optimal TPC matrix and receiver matrix can also be based on a procedure similar to that proposed in Section III-A. Specifically, the optimal receiver matrix can be obtained based on similar analysis to (9) and (10), given a TPC matrix, and it is written as:

$$\begin{aligned}
 \tilde{\mathbf{G}}_{k,b,u}^H &= \tilde{\mathbf{W}}_{k,b,u}^H \mathbf{H}_{k,b,u}^H \times \\
 & \left( \sum_{(b',u')} P_{k,b'} \mathbf{H}_{k,b',u'} \tilde{\mathbf{W}}_{k,b',u'} \tilde{\mathbf{W}}_{k,b',u'}^H \mathbf{H}_{k,b',u'}^H + \sigma^2 \mathbf{I}_R \right)^{-1}.
 \end{aligned} \quad (24)$$

Given the above receiver matrix, the O-CBF/CS TPC matrix can be derived based on a procedure similar to (12) - (17), and can be written as:

$$\begin{aligned}
 \tilde{\mathbf{W}}_{k,b,u} &= w_u \left( P_{k,b} \sum_{b',u'} w_{u'} \mathbf{H}_{k,b,u'}^H \tilde{\mathbf{G}}_{k,b',u'} \tilde{\mathbf{\Xi}}_{k,b',u'}^{-1} \tilde{\mathbf{G}}_{k,b',u'}^H \mathbf{H}_{k,b,u'} \right. \\
 & \quad \left. + \lambda_{k,b} \mathbf{I}_{Tx} \right)^{-1} \mathbf{H}_{k,b,u}^H \tilde{\mathbf{G}}_{k,b,u} \tilde{\mathbf{\Xi}}_{k,b,u}^{-1},
 \end{aligned} \quad (25)$$

where  $\tilde{\mathbf{\Xi}}_{k,b,u} = \mathbf{I}_{L_u} - \tilde{\mathbf{W}}_{k,b,u}^H \mathbf{H}_{k,b,u}^H \tilde{\mathbf{G}}_{k,b,u}$ .

In contrast to the O-CJT TPC matrix condition of (17), the O-CBF/CS TPC matrix of a BS does not explicitly rely on the TPC matrices of other BSs. But this mutual impact still exists, owing to the receiver matrix: when the TPC matrix of a BS changes, the receiver matrix of its serving UE will change accordingly. Then the precoding matrices of other BSs have to be adjusted in order to avoid strong interference.

According to (24) and (25), an iterative algorithm may be conceived for calculating the optimal  $\tilde{\mathbf{W}}_{k,b,u}^*$  and the optimal  $\tilde{\mathbf{G}}_{k,b,u}^{H*}$  can be obtained, which is similar to Algorithm 1. Then the optimal power sharing scheme can be obtained based on a similar procedure to that of Subsection III-B. Then the correspondingly joint optimization can be readily conceived, but it has to be omitted owing to the lack of space.

## V. SIMULATION RESULTS

In this section, we will characterize the performance of our large-scale clustered MIMO network by evaluating the proposed O-CJT, O-CBF/CS and O-NCJT optimization schemes. The SLNR scheme of [18] was also extended, because in [18] it was proposed for a single-RA receiver, while we extend it to a multiple-RA antenna UE scenario. The O-NCJT optimization is similar to the procedure of the O-CBF/CS scheme, and is omitted in the paper due to the lack of space. The detailed

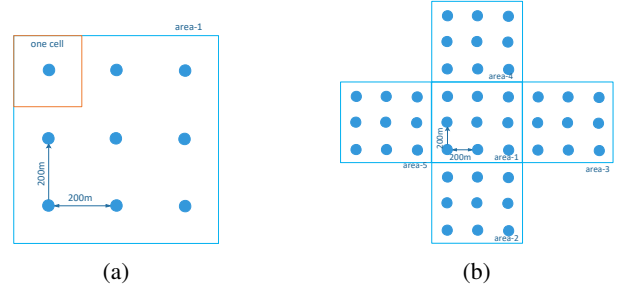


Fig. 3: An example of BS positions for (a) single-cluster scenario (b) multiple-cluster scenario, with  $B_1 = B_2 = 3$ .

optimization procedure can be found in our full version in [33]. The non-cooperative multi-cell MIMO system is evaluated as our baseline. Simulation details are as follows:

- The distance between the adjacent BSs is set to 200 meters and the minimum distance between a UE and a BS is 15 meters. Each cell is a rectangular one centered on its BS;
- Each cluster covers a rectangular area, which contains  $B$  BSs, organized by  $B_1$  columns and  $B_2$  rows;
- The scenarios of  $T = 8$  or  $T = 32$  TAs,  $R = 1$ ,  $R = 2$  or  $R = 4$  RAs are evaluated;  $L_u = R$  is assumed for all UEs.
- Two scenarios are considered: a single-cluster and a multiple-cluster scenario. The single-cluster scenario is a case where all BSs can be connected to a CPU and all calculations can be carried out in a centralized way. Naturally, this would have the best performance, since all the intercell interference can be substantially mitigated and the coherent TPC gain can be maximized. The multi-cluster scenario represents a more realistic case where only some BSs can benefit from a centralized processing. In this case, only the interference within a single cluster can be mitigated, while the inter-cluster interference is hard to coordinate. Consider the above  $B = 9$  BSs per cluster as an example, where the single-cluster scenario and multi-cluster scenario considered are shown in Figures 3a and 3b;
- The large scale fading  $\eta_{b,u}$  includes both path-loss and shadow-fading. The path loss model of the 3GPP TS 36.873 UMa is used at a carrier frequency of 3.5 GHz. The shadow fading  $\text{SF}_{b,u}$  follows a distribution of  $\log_{10} \text{SF}_{b,u} \sim \mathcal{CN}(0, \sigma_{\text{SF}}^2)$ , where  $\sigma_{\text{SF}}^2 = 6$ . Each element of  $\mathbf{H}_{k,b,u}^{\text{sf}}$  follows identical and independent Gaussian distribution with zero mean and unit variance;
- For the DL bandwidth,  $K = 5$  RBs are assumed, and the maximum total DL transmission power of a BS is  $\log_{10}(P_{b,\max}) = 25.6$  dBm. The value of 25.6 dBm is based on a commercial network configuration: 100 MHz communication bandwidth, 30 KHz subcarrier spacing (SCS), and thus 277 RB (each of RB has 12 subcarriers); the maximum DL transmission power of a BS for 100 MHz transmission is usually 53 dBm; thus we have  $53 - 10 \log_{10}(277/5) = 25.6$  (dBm). The reason to use  $K = 5$  here is to speed up the simulation. To achieve

a nice trade off between complexity and performance in a wideband system, a concept of subband is used, i.e. several (e.g. 2, 4 or 8) consecutive RBs are regarded as one unit to share a common TPC matrix.

- For all evaluated schemes, the initial power sharing is the average power, i.e.,  $P_{k,b} = \frac{P_{b,\max}}{K}$ . The noise power at UE is  $\sigma^2 = 1.1384 \times 10^{-11}$ , which is calculated based on  $\sigma^2 = 10^{\frac{-165}{10}} \times \text{SCS} \times 12$ , where  $-165$  dBm/Hz is the noise power density at the UE after taking into account the noise figure of the UE receiver (usually 9dB is used).
- The O-CJT TPC matrix  $\mathbf{W}_{k,b,u}$  is initialized as follows: let  $\tilde{\mathbf{V}}_{k,u} \in \mathbb{C}^{BT \times L_u}$  be the dominant right singular matrix of the concatenated channel matrix  $\tilde{\mathbf{H}}_{k,u}$  defined in (4), then  $\mathbf{W}_{k,b,u}$  is obtained by first getting the corresponding rows of the right singular vectors, and then using per-column normalization. Finally,  $\mathbf{W}_{k,b,u}$  will be normalized as  $\sqrt{P_{k,b}} \mathbf{W}_{k,b,u} / \sqrt{\|\mathbf{W}_{k,b,u}\|^2}$ , where  $\mathbf{W}_{k,b} = [\mathbf{W}_{k,b,1}, \dots, \mathbf{W}_{k,b,U}]$ , so as to satisfy the sum power constraint on this RB.
- The O-CBF/CS TPC matrix  $\tilde{\mathbf{W}}_{k,b,u}$  is initialized as:  $\tilde{\mathbf{W}}_{k,b,u} = \frac{1}{L_u} \mathbf{V}_{k,b,u}$ , where  $\mathbf{V}_{k,b,u}$  contains  $L_u$  right eigenvectors of the DL channel between a UE  $u$  and its serving BS  $b_u$ .
- For the extended SLNR (E-SLNR) scheme, we assume a receiver matrix of  $\mathbf{G}_{k,u}^{\text{slnr}} \in \mathbb{C}^{L_u \times R}$  for each multi-antenna UE. Explicitly,  $\mathbf{G}_{k,u}^{\text{slnr}}$  is assumed to contain  $L_u$  dominant left eigenvectors of the concatenated channel matrix  $\mathbf{H}_{k,u}^{\text{slnr}} = [\mathbf{H}_{k,b_1,u}, \dots, \mathbf{H}_{k,b_n,u}]$ , where  $\mathcal{B}_u^{\text{slnr}} = \{b_1, \dots, b_n\}$  represents the serving BSs of UE  $u$ . Then each row of the matrix  $\mathbf{G}_{k,u}^{\text{slnr}} \mathbf{H}_{k,b,u}$  can be viewed as the channel vector of a single-RA UE, and thus the multi-antenna UE based problem is turned to a single-antenna problem. Equal power sharing is assumed for all scheduled layers of the E-SLNR scheme on each RB.
- To determine  $\mathcal{B}_u$  for the O-CJT and  $\mathcal{B}_u^{\text{slnr}}$  for the E-SLNR, we assume similar criteria: BSs that satisfy  $\frac{P_{b,\max}}{K} \eta_{b,u} \geq X_{\text{thr}}$  belong to the set  $\mathcal{B}_u$  and  $\mathcal{B}_u^{\text{slnr}}$ . In our simulations, three values of  $X_{\text{thr}}$  are evaluated, so that the average  $|\mathcal{B}_u|$  is 9 BSs, 6 BSs and 3 BSs, respectively.
- For the non-cooperative multi-cell MIMO benchmark system, each BS locally determines its ZF TPC on each RB for its serving UEs. For example, let  $\mathcal{U}_b = \{u_1, u_2, \dots, u_{|\mathcal{U}_b|}\}$  be the serving UE set of BS  $b$ . Then the local ZF TPC on RB  $k$  is

$$\mathbf{W}_{k,b}^{\text{local}} = \frac{1}{\sqrt{\sum_{u \in \mathcal{U}_b} L_u}} \mathbf{F}_{k,b} (\mathbf{F}_{k,b}^H \mathbf{F}_{k,b})^{-1} \mathbf{\Lambda}_{k,b}, \quad (26)$$

where  $\mathbf{F}_{k,b} = [\mathbf{F}_{k,b,u_1}, \mathbf{F}_{k,b,u_2}, \dots, \mathbf{F}_{k,b,u_{|\mathcal{U}_b|}}]$ , and  $\mathbf{F}_{k,b,u} \in \mathbb{C}^{T \times L_u}$  contains  $L_{u_i}$  dominant right singular vectors of  $\mathbf{H}_{k,b,u}$ , with  $\mathbf{\Lambda}_{k,b}$  being a diagonal matrix used for normalizing each column of  $\mathbf{F}_{k,b} (\mathbf{F}_{k,b}^H \mathbf{F}_{k,b})^{-1}$ .

- The determination of  $\mathcal{U}_b$  for the baseline scheme assumes that UE  $u$  is attached to a BS which has the largest  $\frac{P_{b,\max}}{K} \eta_{b,u}$ . If the selected BS cannot serve this UE – for example because number of serving layers will be higher than the number of TAs – then UE  $u$  is attached to the next BS which has the second largest  $\frac{P_{b,\max}}{K} \eta_{b,u}$  and the

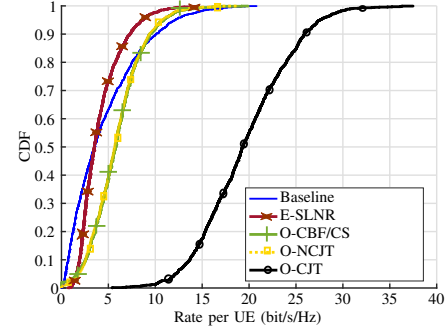


Fig. 4: CDF curve of achievable rate per UE for single cluster with  $B_1 = B_2 = 3$ ,  $T = 8$ ,  $R = 2$  and average  $|\mathcal{B}_u| = 9$ .

ability to serve this UE. Otherwise, this UE is not served.

- All schemes are simulated using 100 channel realizations. In each realization,  $U_i$  UEs are randomly dropped in cell  $i$  and the large-scale fading/small-scale fading are refreshed, where  $U_i$  is chosen randomly from  $\{1, 2, \dots, \frac{T}{R}\}$ , unless specifically mentioned. In all the simulations, perfect DL CSI knowledge is assumed.

#### A. The performance of the TPC matrix

In this subsection, we characterize the performance of the proposed optimal TPC matrix optimization for all the schemes discussed, using the average power sharing in the FD.

1) *Single-cluster scenario:* The cumulative distribution function (CDF) curves of the achievable rate per UE are shown in Figure 4 for a single-cluster scenario, where  $|\mathcal{B}_u| = 9 (\forall u)$ .  $T = 8$  and  $R = 2$  are assumed. Given perfect CSI knowledge, the performance of the above mentioned schemes are portrayed. We will analyze them one by one.

**The O-CJT TPC scheme:** the O-CJT coordination system provides a very promising gain over the baseline: an average sum-rate gain of 4.3, a factor 28.3 gain at the 5 percentile, and a factor 2.3 gain at the 95 percentile. The significant gain accrued from the following aspects:

- Significantly increased beamforming gain, since all BSs coherently transmit each data layer.
- Substantially reduced intercell interference. Based on the proposed O-CJT scheme, the original intercell interference is transformed to useful coherent signal power, which is extremely helpful for achieving a low interference. For the baseline scheme, the intercell interference is the performance bottleneck, which is the pain in commercial networks.
- A compelling balance between the signal power improvement and MU interference reduction, bearing in mind that totally nulling out inter-user interference typically leads to signal power reduction when the interfering UE and target UE are in similar directions. According to (17), the CJT TPC matrix does not target nulling out the inter-user interference, but aims for increasing the rate and thus it strikes a compromise between maximal signal power and minimal interference.

**The E-SLNR scheme:** Figure 4 also portrays the performance of the E-SLNR scheme. Observe that compared to the baseline, the E-SLNR scheme beneficially increases the UE rate at cell-edge, but at the cost of eroding the performance of the cell-center UEs. Specifically, it achieves a factor 3.7 sum-rate gain at the 5 percentile but suffers from 30% performance loss at the 95 percentile, which leads to 10% average performance loss. One of the reasons is that emphasizing the connection between the UE and BS leads to over-suppression of the interference leakage, which ultimately leads to signal power loss. Therefore, supporting less UE connections might be helpful for improving the performance of the E-SLNR, as it will be shown by our simulations. Note that upon extending the SLNR scheme, a potentially suboptimal receiver matrix is used for each UE. Therefore, using an iterative calculation between the receiver matrix and the SLNR TPC matrices may be beneficial. However, our simulations show that this iteration has marginal benefits, which are omitted due to lack of space.

We further compare the performance of the E-SLNR precoder and that of the O-CJT precoder. In the single-cluster scenario, the O-CJT achieves much higher performance for both cell edge and cell center: a factor 7.6 cell edge gain and a factor of 3.3 cell center gain, respectively, which contribute to a factor 4.8 average sum-rate gain. Since the E-SLNR is the state of the art precoder proposed for cell-free MIMO system, and enjoy the advantages of low complexity, it is worth spending more time to compare the pros and cons of the proposed O-CJT precoder and those of E-SLNR precoder. The detailed analysis can be found in the next subsection “*Multiple-cluster scenario*” to make a comprehensive comparison.

**The NCJT and the CBF/CS scheme:** For the NCJT and CBF/CS scheme, it can be seen in Figure 4 that they achieve similar performance: about a factor 1.3 average sum-rate gain. This gain accrues from the optimized TPC which targets reducing the intercell (for CBF/CS) and inter-layer (for NCJT) interference. However, since the beamforming gain of each layer is only contributed from a single BS, non-negligible intercell and inter-layer interference still exist, and hence the sum-rate gain remains quite limited, as in Table II.

TABLE II: Statistical performance comparison for evaluations in Figure 4.

|              | Baseline | E-SLNR | O-CBF/CS | O-NCJT | O-CJT |
|--------------|----------|--------|----------|--------|-------|
| Average gain | 1        | 0.9    | 1.3      | 1.3    | 4.3   |
| 5% gain      | 1        | 3.7    | 3.7      | 3.7    | 28.3  |
| 95% gain     | 1        | 0.7    | 0.9      | 0.9    | 2.3   |

The impact of average cluster size  $|\mathcal{B}_u|$  on the sum-rate gain is shown in Figure 5. For Figure 5a and 5b,  $X_{\text{thr}}$  is selected for ensuring that the average  $|\mathcal{B}_u|$  is 6 and 3, respectively. The same  $X_{\text{thr}}$  values are applied to the E-SLNR scheme. It can be seen that the sum-rate of the O-CJT scheme having smaller cluster sizes degrades compared to that of Figure 4. This is because a wider cooperation cluster attains an improved interference suppression and a higher beamforming gain at a given signal power. As opposed to the O-CJT scheme, the E-SLNR scheme has a better performance than in Figure 4.

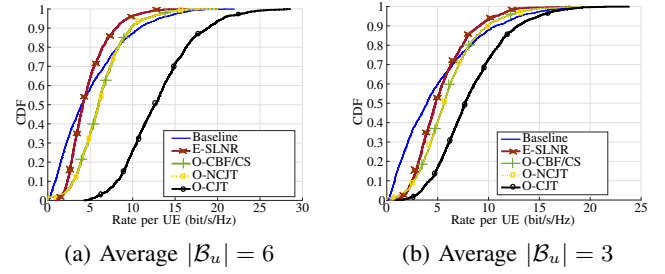


Fig. 5: CDF curve of achievable rate per UE for single cluster, with  $B_1 = B_2 = 3$ ,  $T = 8$ ,  $R = 2$ .

Since the E-SLNR scheme has to strike a beneficial balance between interference-leakage suppression and signal-power maximization, having a moderate cooperation cluster size is important. The detailed performance gains are shown in Table III. The slightly different gains of O-CBF/CS and O-NCJT between Table II and Table III may be due to having in-sufficient channel realizations.

TABLE III: Statistical per-UE rate gain comparison of Figure 5a/5b.

|              | Baseline | E-SLNR      | O-CBF/CS  | O-NCJT    | O-CJT      |
|--------------|----------|-------------|-----------|-----------|------------|
| Average gain | 1        | 0.94 / 1.08 | 1.2 / 1.2 | 1.2 / 1.2 | 2.6 / 1.66 |
| 5% gain      | 1        | 3.9 / 4.3   | 4.0 / 4.0 | 4.1 / 4.1 | 13.9 / 7.1 |
| 95% gain     | 1        | 0.72 / 0.87 | 0.9 / 0.9 | 0.9 / 0.9 | 1.6 / 1.2  |

The per-UE rate gain of the massive MIMO case is further characterized in Figure 6. In the simulations, we consider different numbers of UEs per cell, to show how well the large-scale clustered massive MIMO can handle heavy tele-traffic case.  $T = 32$  and  $R = 4$  are assumed. This antenna configuration is common for the commercial LTE market in East Asia, and a higher  $T$  will be used in the emerging 5G NR network, e.g.,  $T = 64$ . The average  $|\mathcal{B}_u| = 9$  is used. The other parameters are the same as in Figure 4.

Observe from Figure 6 that, for massive MIMO-aided BSs, the O-CJT scheme achieves a more significant per-UE rate gain for the single-cluster scenario: a factor 3 – 4.5 average sum-rate gain and 6.3 – 65.6 gain at the 5 percentile. Figure 6 also shows that the gain grows as the traffic-load per cell increases, which means that the O-CJT TPC based on our clustered MIMO network supports a higher traffic load than its non-cooperative counterpart. Quantitatively, a factor 4.5 average sum-rate gain is attained by 8 UE/cell, where we have  $L_u = R = 4$ . But the full spatial multiplexing capability has to be exploited for both the baseline and the CJT schemes. Therefore, no additional beamforming gain is attained by the O-CJT scheme. However, the O-CJT scheme substantially reduces the intercell interference, and thus the performance gain is notable. This observation again shows that the bottleneck of a non-cooperative MIMO network is the intercell interference. The sum rate gain of the O-NCJT scheme remains limited: to about 20% – 30%. The detailed gains are shown in Table IV.

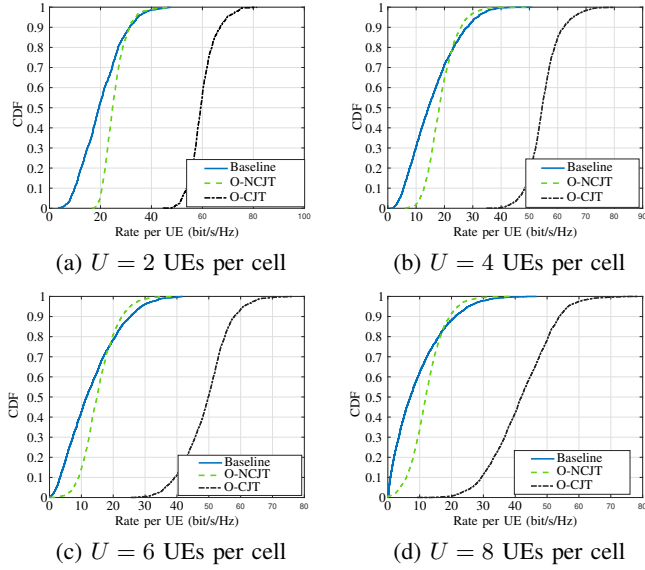


Fig. 6: CDF curve of achievable rate per UE for single-cluster, for  $B_1 = B_2 = 3$ ,  $T = 32$ ,  $R = 4$ , average  $|\mathcal{B}_u| = 9$ .

TABLE IV: Statistical performance comparison for evaluations in Figure 6.

|              | Baseline | O-NCJT w/ $U = 2/4/6/8$    | O-CJT w/ $U = 2/4/6/8$  |
|--------------|----------|----------------------------|-------------------------|
| Average gain | 1        | 1.26 / 1.2 / 1.17 / 1.33 / | 2.97 / 3.5 / 3.75 / 4.5 |
| 5% gain      | 1        | 2.4 / 2.7 / 3.5 / 9.6      | 6.3 / 11 / 16.7 / 65.6  |
| 95% gain     | 1        | 0.97 / 0.9 / 0.88 / 0.87   | 2 / 2.1 / 2.1 / 2.24    |

For the mMTC networks, chip single-antenna UE are commonly used. For this case, the simulation results are also provided to compare the performance of the E-SLNR precoder and the O-CJT precoder. In this case, the E-SLNR precoder rollbacks to the existing SLNR precoder. The CDF curves of per-UE rate are shown in Figure 7.

Based on Figure 7, we can see that the O-CJT scheme achieves a substantial performance gain: a factor 1.38 average sum-rate gain and 1.7 gain at the 5 percentile. At the 95 percentile, the O-CJT scheme has almost the same rate as the baseline. The main reason is that with only 4 UEs and

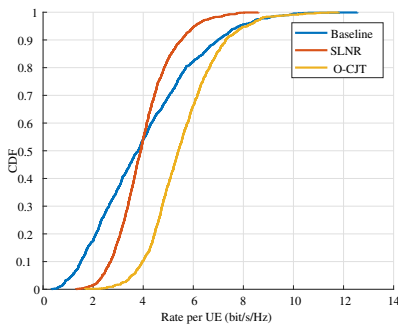


Fig. 7: CDF curve of achievable rate per UE for single-cluster, for  $B_1 = B_2 = 3$ ,  $T = 8$ ,  $R = 1$ ,  $U = 4$ , average  $|\mathcal{B}_u| = 9$ .

a single stream per UE in each cell, the interference level is not dominant. The SLNR precoder achieves the same average sum-rate as the baseline, a factor 1.36 rate gain at the 5 percentile, but it has a rate loss at the 95 percentile. Therefore, the O-CJT precoder can help to increase the per-UE rate both in mMTC as well as MBB networks.

Although our simulations above consider perfect DL CSI, in real networks, the thermal noise and interference imposed on the SRS introduces CSI imperfections and thus will inflict a performance loss on all the above-mentioned schemes. How to guarantee a sufficiently accurate CSI is beyond the scope of this paper, but we would like to provide a few options. For example, a narrow SRS hopping bandwidth can be used for increasing the SRS power density so as to guarantee the target SNR; SRS repetition can be applied in the time domain so that the noise/interference can be averaged out. More advanced methods relying on Wiener filtering exploiting the channel statistics in the time/frequency-domain may also be used for further reducing the noise. For the proposed O-CJT precoder, reducing the cluster-size for each serving UE is also an option, since near-by BSs may acquire an accurate CSI more easily.

2) *Multiple-cluster scenario*: Observe from the single-cluster evaluations that the O-CJT scheme achieves a very promising performance gain over the non-cooperative multicell MIMO system, by turning the intercell interference into useful signal power. However, in realistic networks, it is almost impossible to connect all BSs in a city to a single CPU for a fully centralized optimization, owing to excessive fronthaul rate. The computational complexity of a centralized optimization is another restriction. Therefore, it is necessary to evaluate the performance gain of the O-CJT scheme for realistic networks relying on BS clusters, each having a CPU and carrying out optimization independently.

In the following simulations, the UE-BS association is determined as follows. For the O-CJT, O-CBF/CS and O-NCJT schemes, each UE can only be attached to a single cluster based on the highest  $\frac{P_{b,\max}}{K} \eta_{b,u}$  among all clusters. According to the  $X_{\text{thr}}$  value, serving BSs can be selected from the serving cluster. For our E-SLNR scheme, the serving BSs are also allowed to belong to multiple clusters, have it may also be reviewed as a distributed scheme w.r.t. TPC matrix calculation. However, even though the TPC is calculated independently, the scheduling actually requires centralized decisions for gleaning beamforming gains from multiple BSs. Therefore, the SLNR scheme evaluated here is essentially a centralized method across multiple clusters.

For the multi-cluster scenario, the BS locations and cluster partitions are shown in Figure 3b having five separate clusters, and the achievable rate of the UEs attached to cluster-1 is calculated and plotted. The same three values of  $X_{\text{thr}}$  as in the single-cluster scenario of Figures 4 and 5 are used, to guarantee the average  $|\mathcal{B}_u| = 9, 6, 3$ , respectively. The CDF curve for  $|\mathcal{B}_u| = 9$ ,  $T = 8$  and  $R = 2$  is shown in Figure 8. Since similar trend can be found for  $|\mathcal{B}_u| = 6, 3$ , their CDF curves are omitted, and the main results are summarized in Table V.

Compared to the performance of the single-cluster scenario of Figure 4 and 5, the per-UE rate gain of the CJT over the

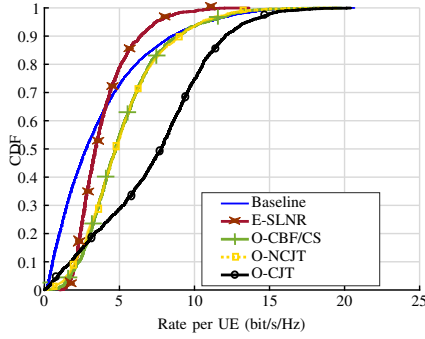


Fig. 8: CDF curve of achievable rate per UE for multi-cluster with  $B_1 = B_2 = 3$ ,  $T = 8$ ,  $R = 2$  and average  $|\mathcal{B}_u| = 9$ .

TABLE V: Statistical performance comparison for evaluations with  $B_1 = B_2 = 3$ ,  $T = 8$ ,  $R = 2$ ,  $|\mathcal{B}_u| = 9/6/3$ .

|              | Baseline | E-SLNR           | O-CBF/CS         | O-NCJT           | O-CJT           |
|--------------|----------|------------------|------------------|------------------|-----------------|
| Average gain | 1        | 1.1 / 0.94 / 1.2 | 1.2 / 1.3 / 1.3  | 1.2 / 1.3 / 1.3  | 2.1 / 2 / 1.7   |
| 5% gain      | 1        | 2.3 / 5.8 / 7.8  | 5/ 3.8 / 3.8     | 5/ 3.8 / 3.8     | 2.6 / 2 / 2     |
| 95% gain     | 1        | 0.7 / 0.7 / 0.8  | 0.92 / 0.9 / 0.9 | 0.92 / 0.9 / 0.9 | 1.3 / 1.2 / 1.1 |

baseline scheme drops dramatically: the average gain drops from 2.6 to 2.1 and the 5 percentile gain drops from 13.9 to 2.6. The main reason is that the inter-cluster interference cannot be mitigated during the optimization, which will impose a grave impact on cell-edge UE. This significant impact of the interference shows again that reducing the interference is the most crucial task in our network.

The O-CJT scheme still achieves the highest average rate-gain of 2.1, while the O-CBF/CS and O-NCJT have a factor of 1.2–1.3 gain. For cell-edge UEs, however, the O-CBF/CS and the O-NCJT schemes attain a better performance, since both schemes reduce the interference as much as possible for increasing the sum rate. As opposed to these two schemes, the O-CJT may reduce the power sharing toward the cluster-edge UEs since they suffer from serious inter-cluster interference, and allocate more power to the cluster-center UEs, to achieve a sum-rate maximization. Hence, the cluster-edge UEs in the O-CJT scheme have lower rates than those in the O-CBF/CS and O-NCJT schemes.

If we take the E-SLNR as baseline, the performance gain of the O-CJT scheme over the E-SLNR can be calculated as shown in Table VI.

TABLE VI: Statistical performance comparison between the E-SLNR and the O-CJT schemes.

|              | E-SLNR | O-CJT              |
|--------------|--------|--------------------|
| Average gain | 1      | 1.9 / 2.1 / 1.4    |
| 5% gain      | 1      | 1.13 / 0.34 / 0.26 |
| 95% gain     | 1      | 1.86 / 1.7 / 1.4   |

Table VI shows that in multi-cluster scenarios, the O-CJT scheme still achieves a factor of 1.4–1.86 gain for cell-center UEs since more BSs contribute to the signal power. However, since the inter-cluster interference cannot be tightly controlled,

the O-CJT has a performance loss compared to the E-SLNR scheme for cell-edge UEs. This is not unexpected since the E-SLNR is explicitly designed for suppressing all kinds of interference, which becomes an advantage for interference-infested UEs. However, we have to bear in mind that the E-SLNR scheme evaluated here assumes a global knowledge of the scheduling decisions across all clusters, which is an important prerequisite for improving the performance gain. Without this global knowledge, the E-SLNR scheme cannot guarantee a good performance for cell-edge UEs either.

TABLE VII: A full comparison between the E-SLNR scheme and the O-CJT scheme.

|                          | E-SLNR   | O-CJT  |
|--------------------------|--|--|
| TPC matrix calculation   | distributed and no iteration involved.   | centralized and involves iteration, but iteration can be reduced to 3.   |
| computational complexity | moderate   | high   |
| scheduling decision      | centralized  | centralized  |
| performance              | small average gain over non-cooperative m-MIMO network, but can provide uniform user experiences.                        | large average gain over non-cooperative m-MIMO network, suitable for throughput-demanding network.   |
| closed form expression   | available, but it is hard to find a good balance between maximization of signal and suppression of interference leakage. | available, and the optimal balance between maximization of signal and suppression of interference leakage can be achieved during iterations. |

Collecting the conclusions from the single-cluster and multi-cluster scenarios, we list the pros and cons of the E-SLNR scheme and the O-CJT scheme in Table VII, where “SC” represents “single-cluster”, and “MC” represents “multi-cluster”. The per-UE rate gain of the massive MIMO scenario is also shown in Figure 9, corresponding to different numbers of UEs per cell. The numbers of TAs and RAs are  $T = 32$  and  $R = 4$ , respectively. The average  $|\mathcal{B}_u| = 9$ . The other parameters are the same as in Figure 8. It is shown that in the multi-cluster scenario, having uncontrolled inter-cluster interference, the O-CJT TPC based on the clustered network can still support a high traffic load at a gain of 2.2 on average and a factor 5.1 at the 5 percentile, with 8 UE/cell. Even when subjected to uncontrolled inter-cluster interference, the per-cluster O-CJT scheme substantially reduces the intercell interference within each cluster, and thus the performance gain is quite notable. The gain of the O-NCJT scheme is still limited to about 10%. The detailed rate-gain values are in Table VIII.

TABLE VIII: Statistical performance comparison for evaluations in Figure 9.

|              | Baseline | O-NCJT w/ $U = 2/4/6/8$ | O-CJT w/ $U = 2/4/6/8$   |
|--------------|----------|-------------------------|--------------------------|
| Average gain | 1        | 1.14 / 1.1 / 1.15 / –   | 1.4 / 1.6 / 1.84 / 2.2   |
| 5% gain      | 1        | 1.5 / 1.8 / 3.2 / –     | 1.7 / 2.1 / 2.45 / 5.1   |
| 95% gain     | 1        | 0.96 / 0.94 / 0.85 / –  | 1.04 / 1.2 / 1.23 / 1.35 |

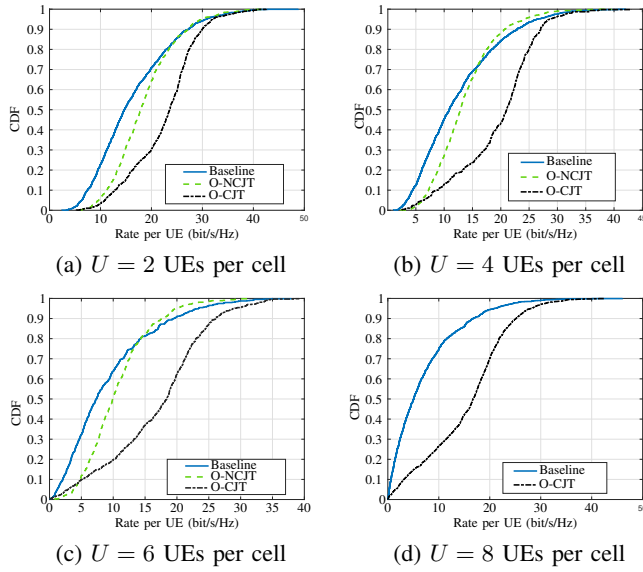


Fig. 9: CDF curve of achievable rate per UE for multi-cluster, for  $B_1 = B_2 = 3$ ,  $T = 32$ ,  $R = 4$ , average  $|B_u| = 9$ .

## VI. CONCLUSIONS

In this paper, we proposed a new network structure based on C-RAN relying on a large-scale clustered massive MIMO network and solved the optimization problem of the weighted achievable sum-rate maximization. Explicitly, we designed the O-CJT TPC matrix on each RB and the optimal power sharing among all RBs, under the per-BS total power constraint in the face of inter-cluster interference. Our simulations show that for a single-cluster scenario, a factor 4.5 average sum-rate gain is attained over the non-cooperative MIMO system, using  $T = 32$ ,  $R = 4$  and 9 cooperative BSs. For a multi-cluster scenario, a factor 2.2 average sum-rate gain is achieved. The 5 percentile gain is more significant. The power sharing scheme further increased the sum-rate, but for the O-CJT TPC, the additional gain remains limited. This shows that for the O-CJT TPC, equal power sharing in the frequency domain is sufficient, no need for joint optimization.

To compare the potential of different coordination schemes, both the NCJT and the CBF/CS schemes were further optimized. Since no coherent beamforming gain can be harvested from different BSs, the sum-rate gains remain limited: to 1.3 and 1.15 for the single-cluster and multi-cluster scenarios, respectively, at  $T = 32$ ,  $R = 4$ . The popular single-antenna SLNR scheme is also extended to multi-antenna UEs, and our simulations show a factor 1.1–1.2 gain.

Therefore, to satisfy the ever-growing demand for high capacity, we believe that O-CJT based large-scale clustered massive MIMO networks constitute promising next-generation techniques.

## REFERENCES

- [1] M. Boldi, A. Tölli, M. Olsson, E. Hardouin, T. Svensson, F. Boccardi, L. Thiele, and V. Jungnickel. *Mobile and Wireless Communications for IMT-Advanced and Beyond*. WileyChichester, 2011, 2011.
- [2] 3GPP. TR 36.814: Further Advancements for E-UTRA; Physical Layer Aspects.

- [3] P. Marsch, S. Brück, A. Garavaglia, M. Schulist, R. Weber, and A. Dekorsy. *Coordinated multi-point in mobile communications: From theory to practice*. P. Marsch, G. Fettweis. Clustering (Cambridge University PressNew York, 2011), 2012.
- [4] R. Irmer, H. Droste, P. Marsch, M. Grieger, G. Fettweis, S. Brueck, H. P. Mayer, L. Thiele, and V. Jungnickel. Coordinated multipoint: Concepts, performance, and field trial results. *IEEE Commun. Mag.*, 49(2):102–111, 2011.
- [5] J. Lee et al. Coordinated multipoint transmission and reception in LTE-advanced systems. *IEEE Commun. Mag.*, 50(11):44–50, 2012.
- [6] A. Lozano, R.W. Heath, and J. Andrews. Fundamental limits of cooperation. *IEEE transactions on information theory*, 59(9):5213–5226, Sept. 2013.
- [7] <https://www.ericsson.com/en/networks/trending/hot-topics/5g-radio-access/massive-mimo>.
- [8] <https://www.huawei.com/en/press-events/news/2018/2/Huawei-Scenario-based-Massive-MIMO-AAUs>.
- [9] 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.*, 16(3):1834–1850, 2017.
- [10] E. Nayeibi, A. Ashikhmin, T. L. Marzetta, and H. Yang. Cell-free massive MIMO systems. In *2015 49th Asilomar Conference on Signals, Systems and Computers*, pages 695–699, Nov. 2015.
- [11] E. Björnson and L. Sanguinetti. Making cell-free massive MIMO competitive with MMSE processing and centralized implementation. *IEEE Transactions on Wireless Communications*, 19(1):77–90, Jan. 2020.
- [12] A. Burr, M. Bashar, and D. Maryopi. Cooperative access networks: Optimum fronthaul quantization in distributed massive MIMO and cloud RAN - invited paper. In *2018 IEEE 87th Vehicular Technology Conference*, pages 1–5, June 2018.
- [13] G. Interdonato, E. Björnson, and H. Q. Ngo et al. Ubiquitous cell-free massive MIMO communications. *EURASIP Journal on Wireless Communications and Networking*, 197, 2019.
- [14] T. C. Mai, H. Q. Ngo, M. Egan, and T. Q. Duong. Pilot power control for cell-free massive MIMO. *IEEE Transactions on Vehicular Technology*, 67(11):11264–11268, Nov. 2018.
- [15] S. Yang, X. Xu, D. Alanis, S. Xin Ng, and L. Hanzo. Is the low-complexity mobile-relay-aided FFR-DAS capable of outperforming the high-complexity CoMP? *IEEE Transactions on Vehicular Technology*, 65(4):2154–2169, April 2016.
- [16] X. Xu, R. Zhang, and L. Hanzo. Digital RoF aided cooperative distributed antennas with FFR in multicell multiuser networks. In *2011 IEEE Vehicular Technology Conference*, pages 1–5, Sep. 2011.
- [17] J. Joung, H. D. Nguyen, and S. Sun. Distributed precoder design with direct-link channel information for distributed antenna systems. *IEEE Transactions on Vehicular Technology*, 67(4):3679–3684, April 2018.
- [18] E. Björnson and L. Sanguinetti. Scalable cell-free massive MIMO systems. <https://arxiv.org/abs/1908.03119>, 2019.
- [19] A. Checko, H. L. Christiansen, Y. Yan, L. Scolari, G. Kardaras, M. S. Berger, and L. Dittmann. Cloud RAN for mobile networks: A technology overview. *IEEE Communications Surveys Tutorials*, 17(1):405–426, 2015.
- [20] <https://www.ericsson.com/en/press-releases/2016/2/ericsson-and-kpn-launch-c-ran-small-cells-in-bus-stop>.
- [21] M. Yemini and A. J. Goldsmith. Optimal resource allocation for cellular networks with virtual cell joint decoding. In *2019 IEEE International Symposium on Information Theory (ISIT)*, pages 2519–2523, July 2019.
- [22] A. Burr, M. Bashar, and D. Maryopi. Ultra-dense radio access networks for smart cities: Cloud-RAN, Fog-RAN and “cell-free” massive MIMO. <https://arxiv.org/abs/1811.11077>, 2018.
- [23] H. Q. Ngo, L. Tran, T. Q. Duong, M. Matthaiou, and E. G. Larsson. On the total energy efficiency of cell-free massive MIMO. *IEEE Transactions on Green Communications and Networking*, 2(1):25–39, Mar. 2018.
- [24] L. D. Nguyen, T. Q. Duong, H. Q. Ngo, and K. Tourki. Energy efficiency in cell-free massive MIMO with zero-forcing precoding design. *IEEE Communications Letters*, 21(8):1871–1874, Aug. 2017.
- [25] E. Nayeibi, A. Ashikhmin, T. L. Marzetta, and B. D. Rao. Performance of cell-free massive MIMO systems with MMSE and LSFD receivers. In *2016 50th Asilomar Conference on Signals, Systems and Computers*, pages 203–207, Nov. 2016.
- [26] B. Dai and W. Yu. Sparse beamforming and user-centric clustering for downlink cloud radio access network. *IEEE Access*, 2:1326–1339, 2014.
- [27] M. Hong, R. Sun, H. Baligh, and Z. Luo. Joint base station clustering and beamformer design for partial coordinated transmission in heterogeneous networks. *IEEE Journal on Selected Areas in Communications*, 31(2):226–240, Feb. 2013.
- [28] M. Yemini and A. J. Goldsmith. Virtual cell clustering with optimal resource allocation to maximize cellular system capacity. In *2019 IEEE*

*Global Communications Conference (GLOBECOM)*, pages 1–7, Dec. 2019.

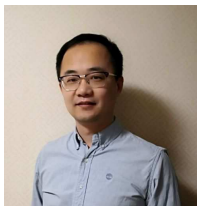
- [29] S. S. Christensen, R. Agarwal, E. De Carvalho, and J. M. Cioffi. Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design. *IEEE Transactions on Wireless Communications*, 7(12):4792–4799, Dec. 2008.
- [30] T. M. Cover and J. A. Thomas. *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, New York, NY, USA, 2006.
- [31] CN. Na, XL. Hou, and A. Harada. Enhanced iterative max-sum-rate algorithm for linear MU-MIMO precoding. In *2011 IEEE 22nd International Symposium on Personal, Indoor and Mobile Radio Communications*, pages 1495–1499, Sep. 2011.
- [32] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, New York, NY, USA, 2004.
- [33] X. Li, X. Zhang, Y. Zhou, and L. Hanzo. Optimal massive-MIMO-aided clustered base-station coordination. to-be-uploaded-to-arxiv, 2020.
- [34] T. Q. S. Quek, M. Peng, O. Simeone, and W. Yu. *Cloud Radio Access Networks: Principles, Technologies, and Applications*. Cambridge University Press, 2017.
- [35] S. Brueck, L. Zhao, J. Giese, and M. A. Amin. Centralized scheduling for joint transmission coordinated multi-point in LTE-Advanced. In *2010 International ITG Workshop on Smart Antennas*, pages 177–184, Feb. 2010.
- [36] H. Dahrouj and W. Yu. Coordinated beamforming for the multicell multi-antenna wireless system. *IEEE Transactions on Wireless Communications*, 9(5):1748–1759, May 2010.



**Lajos Hanzo** FEng, FIEEE, FIET, Fellow of EURASIP, DSc holds an honorary doctorate by the Technical University of Budapest (2009) and by the University of Edinburgh (2015). He is a Foreign Member of the Hungarian Academy of Sciences and a former Editor-in-Chief of the IEEE Press. He has served as Governor of both IEEE ComSoc and of VTS. He has published 1900+ contributions at IEEE Xplore, 19 Wiley-IEEE Press books and has helped the fast-track career of 119 PhD students. (<http://www-mobile.ecs.soton.ac.uk>, [https://en.wikipedia.org/wiki/Lajos\\_Hanzo](https://en.wikipedia.org/wiki/Lajos_Hanzo)).



**Xueru Li** received the B.S. degree in electronic information engineering from Beijing University of Posts and Telecommunications, Beijing, China, in 2011, and the Ph.D. degree in information and communication engineering from Tsinghua University, Beijing, China, in 2016. She is currently a senior engineer of Huawei Wireless Radio Access Technology Department. Her research interests includes MIMO, CoMP, 5G/B5G mobile communication systems.



**Xu Zhang** received the M.S. degree in information and communication engineering from Beihang University, Beijing, China, in 2010, and the Ph.D. degree in information and communication engineering from Tsinghua University, Beijing, China, in 2015. He is currently a senior engineer of Huawei Wireless Radio Access Technology Department. His research interests include CoMP, 5G/B5G mobile communication systems.



**Yongxing Zhou** Yongxing Zhou is now Head of Huawei Device Communication Standard & Research. He is currently working on 5G and beyond technologies. Dr. Yongxing Zhou has more than 100 issued patents.

Before joining Huawei, he was with Samsung from 2002 to 2009 working on IEEE 802.22, IEEE 802.11n standard and implementations as well as TDD related research. He received his Ph.D degree from Tsinghua University, China.