2025 IEEE 26th International Workshop on Signal Processing and Artificial Intelligence for Wireless Communications (SPAWC) | 978-1-6654-7776-5/25/$31.00 ©2025 IEEE | DOI: 10.1109/SPAWC66079.2025.11143524

DNN-Based Distributed Downlink Power Control in User-Centric Cell-Free Massive MIMO Systems

Xuan Liao*, Yue Zhang*, Pei Liu†, Junyuan Wang‡, Wen Zhan§, Giovanni Interdonato¶, Stefano Buzzi¶
*Department of Electronic Engineering, Shantou University, Shantou 515063, China
†School of Information Engineering, Wuhan University of Technology, Wuhan 430070, China & Wuhan University of Technology Advanced Engineering Technology Research Institute of Zhongshan City, Zhongshan 528437, China
‡Department of Electronic and Information Engineering, Tongji University, Shanghai 201804, China
§Department of Electronics and Communication Engineering, Shenzhen Campus of Sun Yat-sen University, Shenzhen 518107, China
¶Department of Electrical and Information Engineering, University of Cassino and Southern Lazio, 03043 Cassino, Italy
(e-mail: {23xliao, yuezhang}@stu.edu.cn, pei.liu@ieee.org, junyuanwang@tongji.edu.cn, zhanw6@mail.sysu.edu.cn, {giovanni.interdonato, s.buzzi}@unicas.it)

(*Invited Paper*)

***Abstract*—This paper studies the distributed downlink power control problem in user-centric cell-free systems, where each user equipment (UE) is served by a subset of access points (APs) to reduce fronthaul requirements and computational complexity at the AP side. We propose a distributed downlink power control method based on a deep neural network (DNN) to enhance the long-term downlink data rate across the entire network. This method relies solely on locally collected large-scale fading information as the DNN input. Also, the method can adapt to dynamic scenarios with varying numbers of associated UEs while meeting real-time power control requirements. Simulation results demonstrate that the proposed scheme outperforms benchmark approaches in terms of network performance.**

***Index Terms*—Cell-free massive MIMO, distributed wireless system, deep learning, power control, user-centric network.**

I. Introduction

Cell-free massive MIMO, a key sixth generation (6G) technology, employs distributed access points (APs) to provide uniform quality of service (QoS) through cooperative transmission [1], [2]. While canonical implementations rely on centralized central processing unit (CPU) coordination [1], [3], practical limitations emerge as fronthaul/computational costs scale linearly with user equipment (UE) numbers [4]. As path loss increases exponentially with the signal transmission distance, full AP coordination could be unnecessary [5], motivating user-centric architectures where UEs are served by localized AP clusters [6]. This approach reduces complexity while maintaining performance via partial AP cooperation [7].

In cell-free massive MIMO systems, the dense distribution of APs and UEs leads to a highly interference-prone environment, where effective power allocation remains critical [8]. Traditional iterative optimization methods [9], [10] suffer from polynomial complexity growth, while supervised deep neural network (DNN) approaches [11], [12] depend on computationally expensive label generation. Although unsupervised learning eliminates labeling requirements [8], [13], its centralized implementations lack scalability. Recent distributed DNN designs [14], [15] address this issue through localized processing. However, their network design is based on single-variable optimization, which may limit model performance and can not achieve global optimal solutions.

In this paper, we propose a distributed DNN architecture employing unsupervised learning for joint multi-variable power optimization. Our method operates without labeled training data, using only locally available large-scale fading information to simultaneously improve the average sum rate and reduce transmit power.

The remainder of this paper is organized as follows: Section II introduces the system model and formulates the downlink power control problem. Section III discusses the design of the distributed DNN. Section IV presents the numerical evaluations. Finally, Section V concludes this article.

Notation: Throughout this paper, $x \sim \mathcal{CN}\left(u, \sigma^2\right)$ represents a complex Gaussian random variable with mean u and variance σ^2. $|\mathcal{X}|$ denotes the cardinality of set $\mathcal{X}$. $\mathbf{I}_N$ denotes the $N \times N$ identity matrix. The expectation is denoted by $\mathbb{E}\left[\cdot\right]$. The superscripts $\left(\cdot\right)^*$ and $\left(\cdot\right)^H$ indicate the conjugate-transpose and Hermitian transpose, respectively. L2 vector norm is denoted by $\|\cdot\|$. The operator $\mathrm{diag}\left(\cdot\right)$ represents either the creation of a diagonal matrix from a given vector or the

The work of Y. Zhang and X. Liao was supported in part by the National Natural Science Foundation of China (No. 62201336), in part by the Guangdong Basic and Applied Basic Research Foundation (No. 2024A1515011541). The work of P. Liu was supported in part by the National Natural Science Foundation of China (No. 62471346), in part by the Guangdong Basic and Applied Basic Research Foundation (No. 2024A1515010247), and in part by the China Scholarship Council (CSC) (No. 202306950052). The work of J. Wang was supported by the National Natural Science Foundation of China (No. 62371344). The work of W. Zhan was supported in part by Guangdong Basic and Applied Basic Research Foundation (No. 2025A1515010235), in part by the Fundamental Research Funds for the Central Universities, Sun Yat-sen University (No. 24qnpy204). The work of S. Buzzi and G. Interdonato was supported by the European Union under the Italian National Recovery and Resilience Plan (NRRP) of NextGenerationEU, partnership on "Telecommunications of the Future" (PE00000001 - program "RESTART", Structural Projects 6GWINET and SRE, cascade call SPARKS). S. Buzzi was also supported by the European Commission through the JU SNS project CENTRIC (Grant No. 101096379). *(Corresponding authors: Yue Zhang, Giovanni Interdonato.)*

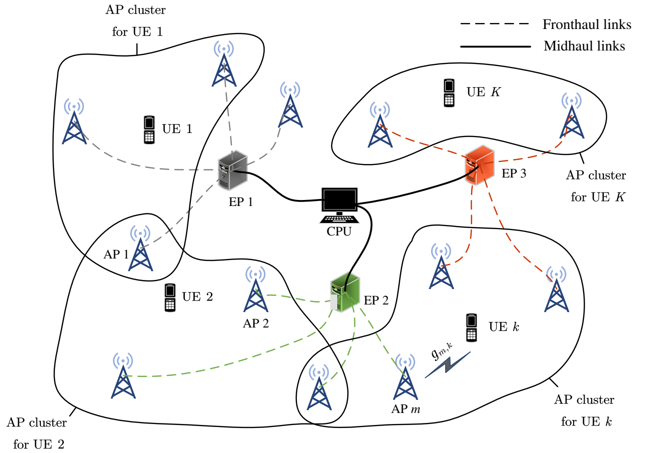


Fig. 1. A cell-free massive MIMO network architecture with user-centric clustering, where each UE is served by several surrounding APs.

extraction of the diagonal elements of a square matrix as a vector, depending on its input.

II. System Model and Problem Formulation

We consider a distributed cell-free architecture with K single-antenna UEs indexed by the set $\mathcal{K}$ and M N-antenna APs indexed by the set $\mathcal{M}$, uniformly deployed as shown in Fig. 1 [16]. Each AP connects via fronthaul to an edge processor (EP), with EPs linked to the CPU through midhaul connection. This design decentralizes power control to EPs, but introduces coordination challenges when a UE is served by APs connected to multiple EPs. For example, in Fig. 1, AP 1 and AP 2 are connected to EP 1 and EP 2, respectively, and hence their transmit power can not be fully coordinated, although they all serve UE 2.

The channel gain between AP m and UE k is modeled as [1]

$$\mathbf{g}_{m,k} = \beta_{m,k}^{1/2}\mathbf{h}_{m,k}, \tag{1}$$

where $\beta_{m,k}$ and $\mathbf{h}_{m,k} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$ represent large-scale and small-scale fadings, respectively. Operating in time division duplex (TDD) mode with block fading, each coherence block contains two phases: uplink channel estimation and downlink data transmission, which are detailed in subsequent sections.

A. Uplink Channel Estimation

In the uplink channel estimation phase, each UE will first send a pilot signal to APs. Let τ_{p} represent the length of each pilot sequence and the number of pilots in the system. The pilot sequence sent by UE k, denoted by $\boldsymbol{\varphi}_k \in \mathbb{C}^{\tau_{\mathrm{p}}\times 1}$, satisfies

$$\boldsymbol{\varphi}_k^H \boldsymbol{\varphi}_j = \begin{cases} 1, & \text{if } \boldsymbol{\varphi}_k = \boldsymbol{\varphi}_j, \\ 0, & \text{otherwise}, \end{cases} \tag{2}$$

$\forall k, j \in \mathcal{K}$. By employing the minimum mean square error (MMSE) estimator at AP m, the estimate $\hat{\mathbf{g}}_{m,k}$ of $\mathbf{g}_{m,k}$ and the estimation error $\mathbf{e}_{m,k} = \mathbf{g}_{m,k} - \hat{\mathbf{g}}_{m,k}$ follow the distributions

$$\hat{\mathbf{g}}_{m,k} \sim \mathcal{CN}\left(\mathbf{0}, \phi_{m,k}\beta_{m,k}\mathbf{I}_N\right) \tag{3}$$

and

$$\mathbf{e}_{m,k} \sim \mathcal{CN}\left(\mathbf{0}, \left(1-\phi_{m,k}\right)\beta_{m,k}\mathbf{I}_N\right), \tag{4}$$

respectively, where

$$\phi_{m,k} = \frac{\tau_{\mathrm{p}}\eta_{\mathrm{u}}\beta_{m,k}}{\tau_{\mathrm{p}}\eta_{\mathrm{u}}\sum_{j=1}^{K}\beta_{m,j}|\boldsymbol{\varphi}_k^H\boldsymbol{\varphi}_j|^2 + \sigma^2} \tag{5}$$

with η_{u} and σ^2 denoting the uplink transmit power of each UE and the noise power, respectively [17].

B. Downlink Data Transmission

We let $\mathcal{K}_m \subset \mathcal{K}$ denote the set of UEs served by AP m. The maximum ratio (MR) precoding scheme is employed, i.e., the normalized precoding vector between UE k and AP m is

$$\mathbf{w}_{m,k} = \frac{\hat{\mathbf{g}}_{m,k}}{\|\hat{\mathbf{g}}_{m,k}\|}. \tag{6}$$

AP m's transmitted signal can then be expressed as

$$\mathbf{x}_m = \sum\nolimits_{k\in\mathcal{K}_m} \eta_{m,k}^{1/2}\mathbf{w}_{m,k}q_k, \tag{7}$$

where q_k, $k \in \mathcal{K}$ denotes the symbol intended for UE k, which satisfies $\mathbb{E}[|q_k|^2] = 1$ and $\mathbb{E}\left[q_k q_j^*\right] = 0, \forall k \neq j$. Additionally, $\eta_{m,k}$ denotes the transmit power assigned to UE k by AP m. Let $\mathcal{M}_k \subset \mathcal{M}$ represent the set of APs serving UE k. Then UE k's received signal is given by

$$\begin{aligned} y_k &= \sum\nolimits_{m=1}^{M} \mathbf{g}_{m,k}^H \mathbf{x}_m + n_k \\ &= \underbrace{\sum\nolimits_{m\in\mathcal{M}_k} \mathbf{g}_{m,k}^H \mathbf{w}_{m,k}\eta_{m,k}^{1/2} q_k}_{\text{desired signal}} \\ &\quad + \underbrace{\sum\nolimits_{m=1}^{M} \mathbf{g}_{m,k}^H \sum\nolimits_{k'\neq k, k'\in\mathcal{K}_m} \eta_{m,k'}^{1/2}\mathbf{w}_{m,k'}q_{k'}}_{\text{multiuser interference}} + n_k, \end{aligned} \tag{8}$$

where $n_k \sim \mathcal{CN}\left(0, \sigma^2\right)$ is the additive white Gaussian noise (AWGN). By assuming that each UE has perfect channel state information (CSI), the signal-to-interference-plus-noise ratio (SINR) of UE k can be obtained as

$$\mathrm{SINR}_k = \frac{\left|\sum_{m\in\mathcal{M}_k} \mathbf{g}_{m,k}^H \mathbf{w}_{m,k}\eta_{m,k}^{1/2}\right|^2}{\sum_{k'\neq k, k'\in\mathcal{K}} \left|\sum_{m\in\mathcal{M}_{k'}} \mathbf{g}_{m,k}^H \mathbf{w}_{m,k'}\eta_{m,k'}^{1/2}\right|^2 + \sigma^2}. \tag{9}$$

And the instantaneous downlink rate $\tilde{R}_k$ of UE k is given by[1]

$$\tilde{R}_k = B\log_2\left(1+\mathrm{SINR}_k\right), \text{ [bit/s]} \tag{10}$$

where B is the communication bandwidth. The achievable average ergodic downlink rate R_k of UE k is written as

$$R_k = \mathbb{E}[\tilde{R}_k], \text{ [bit/s]}, \tag{11}$$

where the expectation is taken over the channel realizations. The average ergodic sum rate is then defined as

$$R = \sum\nolimits_{k=1}^{K} R_k = B\sum\nolimits_{k=1}^{K} \mathbb{E}[\log_2\left(1+\mathrm{SINR}_k\right)], \text{ [bit/s]}. \tag{12}$$

C. Problem Formulation

In this paper, our objective is to find the optimal downlink transmit power that maximizes the sum rate under the power

[1] With channel estimation, the data rate could be corrupted by a constant, representing the time proportion occupied by uplink pilot transmission. Without causing unfair comparison, we omit this overhead to simplify notation.

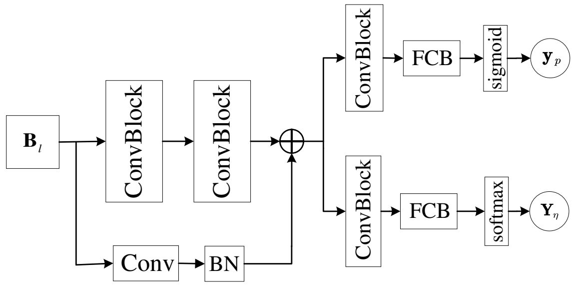


Fig. 2. The designed DNN architecture for joint optimization of AP total transmit power and downlink power allocation.

constraints of each AP based on the channel gains. The sum rate maximization problem can be expressed as

$$\begin{aligned} \text{P1:} \quad & \max_{\{\eta_{m,k}: k\in\mathcal{K}_m, m\in\mathcal{M}\}} \quad R \\ \text{s.t.} \quad & \textstyle\sum_{k\in\mathcal{K}_m} \eta_{m,k} \le P_{\max}, \ \ m\in\mathcal{M}, \\ & \eta_{m,k} \ge 0, \quad k\in\mathcal{K}_m, \, m\in\mathcal{M}, \end{aligned} \tag{13}$$

where $P_{\max}$ is the maximum transmit power for each AP. It is worth noting that UEs that are not served by any APs will be allocated zero power, and hence we only consider the power allocation of the UEs in set $\bigcup_{m\in\mathcal{M}} \mathcal{K}_m$. From (9), we can see that Problem P1 is non-convex. In any non-trivial setup, the search for the optimal solution is very complex. Moreover, since deriving R_k in a closed form is analytically intractable unless bounds and/or approximations are used [1]. However, approximations and bounds may fail to precisely capture the real objective function, resulting in certain performance losses. This issue can be addressed through the data-driven optimization capabilities of deep learning techniques. According to the universal approximation theorem [18], deep learning can approximate functions lacking closed-form expressions, making it a useful tool for solving complex optimization problems.

III. Distributed Deep Learning-Based Power Control Method

In this section, we propose a distributed DNN framework to solve Problem P1, mitigating fronthaul overhead from centralized DNNs and conventional optimization methods. Each EP employs a dedicated DNN that locally manages power allocation across its connected APs using local CSI, with network training conducted via unsupervised learning.

A. Network Design

In the proposed scheme, each EP uses locally collected large-scale fading information as input to the DNN, capturing key channel propagation and interference characteristics, which can be practically measured through received signal strength. No channel gain exchange occurs between EPs, ensuring scalability. To improve downlink power control, we design a DNN that jointly optimizes total transmit power and downlink power allocation, as shown in Fig. 2. This design can also be extended to deal with other multi-objective optimization problems by adding output blocks.

The set of local large-scale fading coefficients collected by EP l is denoted as $\mathcal{B}_l$, and the number of APs connected to EP l is defined as M_l. To solve Problem P1 using the DNN approach, we need to find an unknown mapping from $\{\mathcal{B}_1, \mathcal{B}_2, ..., \mathcal{B}_L\}$ to the AP total transmit power and downlink power allocation.

The network design processes across EPs are generally consistent. In the rest of this subsection, we will take EP l as an example. For each AP connected to EP l, the collected large-scale fading coefficients $\beta_{m,k} \in \mathcal{B}_l$ need to be pre-processed before input to the model, as follows:

$$\tilde{\beta}_{m,k} = \frac{\beta_{m,k}}{\sum_{i\in\mathcal{K}_m} \beta_{m,i}}. \tag{14}$$

This step accelerates the training process and contributes to the improved performance of the distributed DNN [11]. The pre-processed large-scale fading coefficients collected in tensor $\mathbf{B}_l \in \mathbb{R}^{M_l\times K\times T}$ are used as inputs to the model, where T denotes the number of samples in a batch. It is worth noting that the dynamic changes in input and output data sizes, caused by varying AP-UE associations, are addressed by designing the DNN to accommodate the maximum number of UEs K supported by the system. In this design, $\tilde{\beta}_{m,k} \in \mathbf{B}_l^{(t)}$, $1 \le t \le T$ are set to zero whenever $k \notin \mathcal{K}_m$. Consequently, the input and output sizes of the DNN are fixed, making them independent of different AP-UE associations. This allows a single DNN at each EP to handle scenarios with dynamic AP-UE service relationships effectively.

The proposed model consists of a feature extraction network and two parallel output networks. The feature extraction network is composed of multiple convolutional layers to extract features, where the convolution operation can be expressed as:

$$\mathbf{C}^{(e)} = \mathrm{Conv}(\mathbf{B}_l^{(t)}; \{\mathbf{W}^{(e-1)}, \mathbf{b}^{(e-1)}\}), \tag{15}$$

in which $\mathrm{Conv}(\cdot)$ represents the convolution operation with a convolution kernel $\mathbf{W}^{(e-1)}$ using a sliding step of 1 and zero padding of 1, along with bias $\mathbf{b}^{(e-1)}$. The variable e represents the epoch index. The convolution block, referred to as $\mathrm{ConvBlock}$, processes the input as follows:

$$\mathrm{ConvBlock}(\mathbf{B}_l^{(t)}) = \mathrm{BN}\big(\delta(\mathbf{C}^{(e)})\big), \tag{16}$$

where $\delta(\cdot)$ denotes the LeakyReLU (LReLU) activating function, and $\mathrm{BN}(\cdot)$ is the batch normalization operation.

The operations of the feature extraction network can be expressed as:

$$\begin{aligned} \mathbf{X}^{(e)} = & \,\mathrm{ConvBlock}\big(\mathrm{ConvBlock}(\mathbf{B}_l^{(t)})\big) \\ & + \mathrm{BN}\big(\mathrm{conv}\big(\mathbf{B}_l^{(e)}; \{\mathbf{W}^{(e-1)}, \mathbf{b}^{(e-1)}\}\big)\big). \end{aligned} \tag{17}$$

We design two parallel output networks to learn two decision variables, respectively used to decide the total transmit power of APs and the downlink power allocation. These networks consist of convolution blocks and fully connected blocks (FCBs), the output of the FCBs can be expressed as:

$$\mathbf{C}_p^{(e)} = \delta\big(\mathrm{BN}\big(\mathrm{ConvBlock}(\mathbf{X}^{(e)}); \theta_p^{(e-1)}\big)\big), \tag{18}$$

$$\mathbf{C}_\eta^{(e)} = \delta\big(\mathrm{BN}\big(\mathrm{ConvBlock}(\mathbf{X}^{(e)}); \theta_\eta^{(e-1)}\big)\big), \tag{19}$$

where $\theta_p^{(e-1)}$ and $\theta_\eta^{(e-1)}$ are the learned parameters of the fully connected layers. The network outputs are converted to power control coefficients through the sigmoid and softmax functions in the output layer, respectively. Thus, the outputs are $\mathbf{y}_p^{(e)} = \text{sigmoid}(\mathbf{C}_p^{(e)})$ and $\mathbf{Y}_\eta^{(e)} = \text{softmax}(\mathbf{C}_\eta^{(e)})$, where $\mathbf{y}_p^{(e)} \in \mathbb{R}^{M_l}$ denotes the power control coefficient used for predicting AP total transmit power, $\mathbf{Y}_\eta^{(e)} \in \mathbb{R}^{M_l \times K}$ denotes the power control coefficient used for predicting downlink power allocation. The choice of the sigmoid and softmax functions ensures that the outputs comply with the constraints in Problem P1. Specifically, the sigmoid function limits $\mathbf{y}_p^{(e)}$ to the range $(0, 1]$, thus ensuring that the predicted AP total transmit power does not exceed the maximum power constraint. Meanwhile, the softmax function enables each AP to proportionally allocate its total transmit power among the UEs it serves. Finally, the downlink transmit power $\mathbf{P}_l$ at EP l of epoch e can be obtained as

$$\mathbf{P}_l^{(e)} = \text{diag}(\mathbf{y}_p^{(e)} P_{\max}) \cdot \mathbf{Y}_\eta^{(e)}. \tag{20}$$

B. Network Training

To maximize the network-wide long-term average ergodic downlink rate, the DNN at each EP collaboratively minimizes a common loss function defined as:

$$\text{Loss} = -\mathbb{E}_{\mathbf{g}(\{\beta_{m,k}\})}\left[\sum_{k=1}^{K} \tilde{R}_k\big(\mathbf{g}(\{\beta_{m,k}\}), \boldsymbol{\Theta}\big)\right], \tag{21}$$

where $\boldsymbol{\Theta} = \{\Theta_1, \Theta_2, \ldots, \Theta_L\}$ denotes the set of training parameters for DNNs, and $\mathbf{g}(\{\beta_{m,k}\})$ represents the channel gains with given large-scale fading coefficients. The expectation in (21) is averaged over the small-scale fading coefficients. However, it is challenging to derive its closed-form expression. To address this issue, we utilize the commonly employed mini-batch gradient descent method for training the network [14], [19]. During each training iteration, large-scale fading information $\{\beta_{m,k}\}$ is collected based on the known AP-UE deployment scenario. Subsequently, T synthetic samples of channel estimate and channel estimation errors, denoted as $\{\check{\mathbf{g}}^{(t)}\}_{t=1,\ldots,T}$ and $\{\check{\mathbf{e}}^{(t)}\}_{t=1,\ldots,T}$, respectively, are generated according to the known distributions in (3) and (4) using Monte Carlo sampling with given $\{\beta_{m,k}\}$ and pilot SINR $\{\phi_{m,k}\}$. Using these T samples, the loss function (21) can be approximated as:

$$\begin{aligned}\text{Loss} \approx -\tfrac{1}{T}\textstyle\sum_{t=1}^{T}\sum_{k=1}^{K} \tilde{R}_k\big(&\check{\mathbf{g}}^{(t)}(\{\beta_{m,k}\}, \{\phi_{m,k}\}),\\ &\check{\mathbf{e}}^{(t)}(\{\beta_{m,k}\}, \{\phi_{m,k}\}), \boldsymbol{\Theta}\big).\end{aligned} \tag{22}$$

The approximate expression in (22) is employed during each iteration to compute the gradient and update the model parameters $\boldsymbol{\Theta}$. Via the above approach, the training process efficiently handles the stochastic nature of small-scale fading and facilitates the network's long-term performance optimization.

IV. Simulation and Numerical Results

To verify the performance of the proposed scheme, we conduct simulations in this section.

A. Simulation Setup

We focus on a scenario where APs and UEs are randomly distributed within a 1000 m × 1000 m square area following a uniform distribution. We use 4 EPs, each connected to 10 randomly selected APs, with each AP assigned to a single EP. The number of antennas for each AP is $N = 2$. The maximum transmit power per AP $P_{\max} = 1$ W, and the uplink transmit power of pilot signals $\eta_{\text{u}} = 100$ mW. The received noise power σ^2 is set to -94 dBm. We normalize the bandwidth B to unity. The large-scale fading between AP m and UE k is [20]

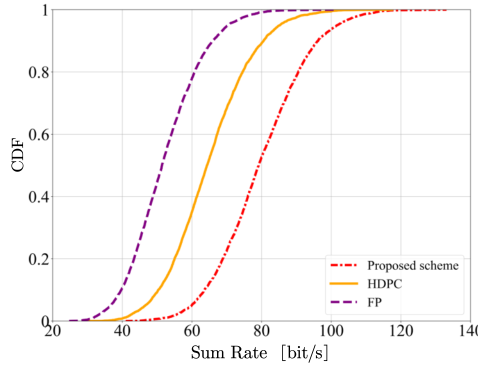
$$\beta_{m,k} = -30.5 - 36.7 \log_{10}(d_{m,k}/1\,\text{m}) + F_{m,k}\ \text{dB}, \tag{23}$$

where $d_{m,k}$ is the distance from UE k to AP m. $F_{m,k} \sim \mathcal{N}\left(0, 4^2\right)$ is the shadow fading. The proposed DNN scheme is trained with $M = 40$ APs and for $K \in \{10, 20\}$ UEs. $\tau_{\text{p}} = K/2$ orthogonal pilots are randomly allocated to each UE. The method in [7] is introduced for UE association. We trained the model using the Adam optimizer with a learning rate of 0.001 on 200,000 samples, split into 90% for training and 10% for testing, with a batch size of 100. Early stopping based on test set performance was used to prevent overfitting.

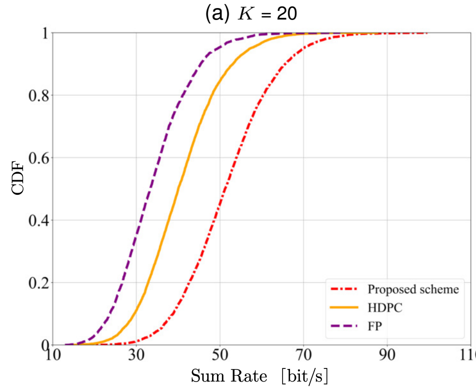
The following approaches are compared:

- *Heuristic downlink power control (HDPC) method*: This method, proposed in [11], employs a modified version of the fully connected network to optimize the power control.
- *Fractional power (FP) allocation*: The system employs a scalable power control method proposed in [16], where the total transmit power of each AP m is fixed at $P_{\max}$, and the downlink power allocation is given by

$$\eta_{m,k} = P_{\max} \cdot \frac{\beta_{m,k}}{\sum_{i \in \mathcal{K}_m} \beta_{m,i}}, k \in \mathcal{K}_m. \tag{24}$$



(a) K = 20



(b) K = 10

Fig. 3. CDF of the sum rate. $M = 40$, $K \in \{10, 20\}$.

Fig. 3 shows the cumulative distribution function (CDF) of the sum rate $\sum_{k=1}^{K} \tilde{R}_k$. It can be observed that for both $K =$

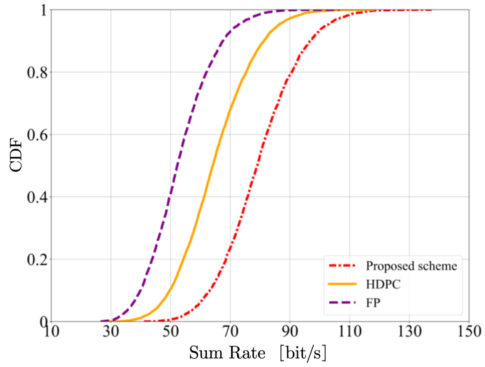


Fig. 4. CDF of the sum rate with different UE association algorithms for generating training and test data. $M = 40$, $K = 20$.

TABLE I
AVERAGE TRANSMIT POWER PER AP ($M = 40$, $K = 20$)

Method	Average transmit power (normalized by $P_{\max}$)
Proposed Scheme	31.4%
HDPC	48.7%
FP	100%

10 and $K = 20$, the proposed scheme consistently outperforms other benchmark methods, demonstrating its adaptability to systems with varying numbers of UEs. Furthermore, as the number of UEs decreases, the per-UE rates achieved by all methods improve. This phenomenon is attributed to the reduction in interference as the number of UEs decreases, resulting in higher per-UE data rates.

In Fig. 4, we generate test data using the association algorithm from [21] to evaluate the impact of different UE association algorithms on the proposed scheme. While during the training phase, the UE association method proposed in [7] is employed. The results presented in Fig. 4 demonstrate that the proposed scheme outperforms the two benchmarks. Despite the different association algorithms used during the training and testing phases, the proposed scheme effectively adapts to the varying association conditions encountered during testing, confirming its robustness.

Table I illustrates the average transmit power per AP $\frac{1}{M}\sum_{m=1}^{M}\sum_{k\in\mathcal{K}_m}\eta_{m,k}$ for different methods under $M = 40$, $K = 20$. The proposed scheme can, on average, reduce the transmit power by 17.9% and 68.6% compared to the HDPC method and the FP schemes, respectively. The result shows that the proposed method can not only achieve higher data rates but also incur lower power overhead, demonstrating its superiority in improving the overall network performance.

V. CONCLUSION

In this paper, we propose a distributed downlink power control method based on DNN to improve the network-wide performance of a cell-free massive MIMO system. It jointly optimizes each AP's total transmit power and their downlink power allocation to their respective serving UEs, using only local large-scale fading information as input for the DNN. This eliminates the need for frontend information exchange, allowing power control tasks to be independently completed on each EP. Simulation results validate the effectiveness of our proposed schemes. Furthermore, our power control scheme effectively adapts to the dynamic service relationships between APs and UEs induced by different UE association algorithms, demonstrating its generalization and robustness.

REFERENCES

[1] 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.

[2] H. Q. Ngo, G. Interdonato, E. G. Larsson, G. Caire, and J. G. Andrews, "Ultradense cell-free massive MIMO for 6G: Technical overview and open questions," *Proc. IEEE*, vol. 112, no. 7, pp. 805–831, Jul. 2024.

[3] S. Buzzi and A. Zappone, "Downlink power control in user-centric and cell-free massive MIMO wireless networks," in *Proc. IEEE PIMRC*, Oct. 2017.

[4] C. Hao, T. T. Vu, H. Q. Ngo, M. N. Dao, X. Dang, C. Wang, and M. Matthaiou, "Joint user association and power control for cell-free massive MIMO," *IEEE Internet Things J.*, vol. 11, no. 9, pp. 15823–15841, May 2024.

[5] J. Wang, L. Dai, L. Yang, and B. Bai, "Clustered cell-free networking: A graph partitioning approach," *IEEE Trans. Wireless Commun.*, vol. 22, no. 8, pp. 5349–5364, Aug. 2023.

[6] H. A. Ammar, R. Adve, S. Shahbazpanahi, G. Boudreau, and K. V. Srinivas, "User-centric cell-free massive MIMO networks: A survey of opportunities, challenges and solutions," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 1, pp. 611–652, 1st Quart., 2022.

[7] S. Buzzi and C. D'Andrea, "Cell-free massive MIMO: User-centric approach," *IEEE Wireless Commun. Lett.*, vol. 6, no. 6, pp. 706–709, Dec. 2017.

[8] F. Liang, C. Shen, W. Yu, and F. Wu, "Towards optimal power control via ensembling deep neural networks," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1760–1776, Mar. 2020.

[9] S. Chakraborty, Ö. T. Demir, E. Björnson, and P. Giselsson, "Efficient downlink power allocation algorithms for cell-free massive MIMO systems," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 168–186, 2021.

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

[11] M. Zaher, Ö. T. Demir, E. Björnson, and M. Petrova, "Learning-based downlink power allocation in cell-free massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 22, no. 1, pp. 174–188, Jan. 2023.

[12] Y. Zhao, I. G. Niemegeers, and S. H. D. Groot, "Power allocation in cell-free massive MIMO: A deep learning method," *IEEE Access*, vol. 8, pp. 87185–87200, 2020.

[13] N. Rajapaksha, K. B. S. Manosha, N. Rajatheva, and M. Latva-aho, "Unsupervised learning-based joint power control and fronthaul capacity allocation in cell-free massive MIMO with hardware impairments," *IEEE Wireless Commun. Lett.*, vol. 12, no. 7, pp. 1159–1163, Jul. 2023.

[14] D. Yu, H. Lee, S. -E. Hong, and S. -H. Park, "Learning decentralized power control in cell-free massive MIMO networks," *IEEE Trans. Veh. Technol.*, vol. 72, no. 7, pp. 9653–9658, Jul. 2023.

[15] Y. Zhang, J. Zhang, S. Buzzi, H. Xiao, and B. Ai, "Unsupervised deep learning for power control of cell-free massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 72, no. 7, pp. 9585–9590, Jul. 2023.

[16] G. Interdonato, P. Frenger, and E. G. Larsson, "Scalability aspects of cell-free massive MIMO," in *Proc. IEEE ICC*, 2019, pp. 1–6.

[17] G. Interdonato, H. Q. Ngo, and E. G. Larsson, "Enhanced normalized conjugate beamforming for cell-free massive MIMO," *IEEE Trans. Commun.*, vol. 69, no. 5, pp. 2863–2877, May 2021.

[18] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Netw.*, vol. 2, no. 5, pp. 359–366, Jan. 1989.

[19] Q. Qian, R. Jin, J. Yi, L. Zhang, and S. Zhu, "Efficient distance metric learning by adaptive sampling and mini-batch stochastic gradient descent (SGD)," *Mach. Learn.*, vol. 99, no. 3, pp. 353–372, Jun. 2015.

[20] Ö. T. Demir, E. Björnson, and L. Sanguinetti, "Foundations of user-centric cell-free massive MIMO," *Found. Trends Signal Process.*, vol. 14, nos. 3–4, pp. 162–472, 2021.

[21] E. Björnson and L. Sanguinetti, "Scalable cell-free massive MIMO systems," *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 4247–4261, Jul. 2020.