

Deep Learning-based Power Control for Cell-Free Massive MIMO Networks

Nuwanthika Rajapaksha, K. B. Shashika Manosha, Nandana Rajatheva, and Matti Latva-aho
Centre for Wireless Communications, University of Oulu, Finland

E-mail: {nuwanthika.rajapaksha, nandana.rajatheva, matti.latva-aho}@oulu.fi, manoshadt@gmail.com

Abstract—A deep learning (DL)-based power control algorithm that solves the max-min user fairness problem in a cell-free massive multiple-input multiple-output (MIMO) system is proposed. Max-min rate optimization problem in a cell-free massive MIMO uplink setup is formulated, where user power allocations are optimized in order to maximize the minimum user rate. Instead of modeling the problem using mathematical optimization theory, and solving it with iterative algorithms, our proposed solution approach is using DL. Specifically, we model a deep neural network (DNN) and train it in an unsupervised manner to learn the optimum user power allocations which maximize the minimum user rate. This novel unsupervised learning-based approach does not require optimal power allocations to be known during model training as in previously used supervised learning techniques, hence it has a simpler and flexible model training stage. Numerical results show that the proposed DNN achieves a performance-complexity trade-off with around 400 times faster implementation and comparable performance to the optimization-based algorithm. An online learning stage is also introduced, which results in near-optimal performance with 4-6 times faster processing.

Index Terms—cell-free massive MIMO, max-min user fairness, power control, deep learning, unsupervised learning

I. INTRODUCTION

Massive MIMO, where a base station with a large number of antennas simultaneously serves many users, has become a key technology in fifth generation (5G) networks due to their high throughput and reliability [1]–[3]. Cell-free massive MIMO combines both massive MIMO and distributed MIMO and has the potential of providing uniformly good throughput for all users [4]–[6]. In cell-free massive MIMO, a large number of distributed access points (APs) serve a much smaller number of users distributed over a wide area, and there are no cells or cell-boundaries [4]. All the APs are connected to a central processing unit (CPU) via the backhaul network, and they coherently cooperate to serve all users using same time-frequency resources via time-division duplex (TDD) [5]. Different scalable cell-free massive MIMO architectures and receiver/combiner schemes for uplink/downlink signal processing are studied in [5]–[8].

Proper power allocation helps controlling the inter-user interference and optimizing the network performance, which often involves advanced optimization techniques [9]. In [5], authors show that max-min power control enables cell-free massive MIMO to provide uniformly good service to all users,

regardless of their locations. The channel hardening property of cell-free MIMO allows neglecting small-scale fading, causing long-term fading to determine the power controlling time [9]. Higher time-complexity of optimization-based power control becomes a challenge to meet these time constraints and limits their practical implementation.

Owing to the universal function approximation property of artificial neural networks (ANNs) [10], DL-based techniques have enabled radio resource allocation with a lower complexity than traditional optimization-based approaches. Several studies have proposed DL-based power control for cellular and cell-free massive MIMO systems [9], [11]–[14]. Most of the existing studies focus on supervised learning approach where a DNN is trained to learn the mapping between the inputs (user locations or channel statistics) and the optimal power allocations obtained by an optimization algorithm. The unsupervised learning algorithm proposed in [13] for K -user interference channel power control problem eliminates the need of knowing the optimal power allocations during model training. In this study, we are interested in such an unsupervised learning algorithm for cell-free massive MIMO power control which will simplify the data preparation and model training stages. The contributions of the paper are as follows:

- In this paper, we consider the max-min rate problem in a cell-free massive MIMO system. We propose, design and implement a DNN to learn user power allocations using channel statistics to achieve max-min user fairness in an unsupervised manner. The method consists of an offline model training stage and an online prediction stage.
- In contrast to previous work on supervised learning-based power control, we introduce unsupervised learning for cell-free massive MIMO power control. It does not require the optimal power allocations to be known during model training as in supervised learning, which makes the data preparation and model training simpler, more practical and flexible, since the DNN can be easily retrained in a changing environment over the time.
- Furthermore, we introduce an online training stage to improve the performance. The model is customized and fine-tuned in each channel realization during the online implementation in order to improve minimum user rate.
- Simulation results show that the proposed DNN achieves close performance to the conventional optimization-based max-min power control, with a significantly lower time-

complexity. The performance-complexity trade-off of the proposed DL-based approach makes it a potential candidate for practical implementation.

II. SYSTEM MODEL

We consider a cell-free massive MIMO system with M single-antenna APs and K single-antenna users randomly distributed in a $D \times D$ geographic area. The APs are connected to a CPU via backhaul connections. The channel coefficient between k th user and m th AP is modeled as $g_{mk} = \sqrt{\beta_{mk}}h_{mk}$ [5]. Here, β_{mk} is the large-scale fading consisting of pathloss and shadowing, and $h_{mk} \sim \mathcal{CN}(0, 1)$ represents the small-scale fading between k th user and m th AP. The uplink of the network is considered which consists of pilot transmission, channel estimation, and uplink data transmission phases.

A. Pilot Transmission and Channel Estimation

Initially, all the users undergo a pilot transmission phase in order to estimate the uplink channel coefficients. During this stage, all K users simultaneously transmit their pilot sequences of length τ symbols to the APs. Let $\sqrt{\tau}\phi_k \in \mathbb{C}^{\tau \times 1}$ be the pilot sequence assigned to k th user with $\|\phi_k\|^2 = 1$. The received signal at m th AP is then given by

$$\mathbf{y}_{p,m} = \sqrt{\tau\rho_p} \sum_{k=1}^K g_{mk}\phi_k + \mathbf{w}_{p,m}, \quad (1)$$

where $\mathbf{w}_{p,m} \in \mathbb{C}^{\tau \times 1}$ is the additive noise at m th AP with i.i.d $\mathcal{CN}(0, 1)$ elements. Then, m th AP estimates the channel $g_{mk}, \forall k$ by projecting the received signal $\mathbf{y}_{p,m}$ onto pilot sequence ϕ_k^H as $\tilde{y}_{p,mk} = \phi_k^H \mathbf{y}_{p,m}$ [5]. Thus,

$$\tilde{y}_{p,mk} = \sqrt{\tau\rho_p} \left(g_{mk} + \sum_{k' \neq k} g_{mk'} \phi_k^H \phi_{k'} \right) + \phi_k^H \mathbf{w}_{p,m}, \quad (2)$$

where the linear minimum mean-squared error (MMSE) estimate of g_{mk} given $\tilde{y}_{p,mk}$ is

$$\hat{g}_{mk} = \frac{\mathbb{E}\{\tilde{y}_{p,mk}^* g_{mk}\}}{\mathbb{E}\{|\tilde{y}_{p,mk}|^2\}} \tilde{y}_{p,mk} = c_{mk} \tilde{y}_{p,mk}, \quad (3)$$

where c_{mk} is obtained as [5]

$$c_{mk} = \frac{\sqrt{\tau\rho_p}\beta_{mk}}{\tau\rho_p \sum_{k'=1}^K \beta_{mk'} + |\phi_k^H \phi_{k'}|^2 + 1}. \quad (4)$$

B. Uplink Data Transmission

After the training phase, the actual uplink data transmission begins where all the users simultaneously send their signals to the APs. Let $x_k = \sqrt{\rho} q_k s_k$ be the transmit signal from k th user, where s_k is the transmit symbol with $\mathbb{E}\{|s_k|^2\} = 1$. Normalized uplink SNR is denoted by ρ and q_k is the power control coefficient of k th user, where $0 \leq q_k \leq 1$. The received signal at m th AP from all the users is given by

$$y_m = \sum_{k=1}^K g_{mk} x_k + w_m = \sqrt{\rho} \sum_{k=1}^K g_{mk} \sqrt{q_k} s_k + w_m, \quad (5)$$

where $w_m \sim \mathcal{CN}(0, 1)$ is the additive noise at m th AP. Then match filtering is done at each AP using the locally obtained

channel estimate \hat{g}_{mk} , and the scaled received signals are sent to the CPU for joint detection. The aggregated received signal r_k (6) at the CPU is used to detect s_k . We assume the large scale fading β_{mk} to be known [5].

$$r_k = \sqrt{\rho} \sum_{k'=1}^K \sum_{m=1}^M \hat{g}_{mk'}^* g_{mk'} \sqrt{q_{k'}} s_{k'} + \sum_{m=1}^M \hat{g}_{mk}^* w_m. \quad (6)$$

C. Max-Min User Rate Scheme

In this work, we consider a max-min user fairness scheme where the objective is to maximize the minimum user rate by optimizing the user power allocations. We assume that only the knowledge of channel statistics is used at the CPU when deriving the achievable rate of each user. The uplink rate for the k th user can be derived as (7) [5]. There, $\gamma_{mk} = \mathbb{E}\{|\hat{g}_{mk}|^2\} = \sqrt{\tau\rho_p}\beta_{mk}c_{mk}$. Therefore, we can see that the achievable rate in (7) is a function of only the large-scale fading β_{mk} and user transmit power coefficients q_k , and does not involve instantaneous channel values. Then, the max-min rate problem can be formulated as

$$\begin{aligned} P1 : \max_{q_k} \quad & \min_{k=1,2,\dots,K} R_k^{UP}, \\ \text{s.t.} \quad & 0 \leq q_k \leq 1, \quad k = 1, 2, \dots, K. \end{aligned} \quad (8)$$

In [5], an algorithm using bisection and solving a sequence of linear feasibility problems is proposed to solve problem P1. In [15], a less complex algorithm is proposed by reformulating the original problem into a geometric programming (GP) problem and solving it using a convex optimization software to obtain optimum power allocations. Instead of using such an analytical method, we propose a data-driven approach to learn the optimum solutions of the max-min problem.

III. DEEP LEARNING-BASED POWER CONTROL

As mentioned earlier, we are interested in an unsupervised learning-based approach which does not require optimal ground truth outputs for model training. For the considered power control problem, a supervised learning approach complicates data preparation and model training due to the complexity of generating ground truth power allocations using an optimization algorithm, especially when M and K are large. In contrast, an unsupervised DNN can be directly fed with inputs in order to learn the optimum solutions minimizing a given loss function during the training process. Such approach is more flexible and adaptable to be practically implemented in a changing wireless communications environment.

We propose a feedforward DNN as illustrated in Fig. 1 in order to address the power control problem P1 (8). The network consists of $L+1$ layers that are sequentially connected to produce the mapping $f(\mathbf{x}_0; \theta) : \mathbb{R}^{N_0 \times 1} \mapsto \mathbb{R}^{N_L \times 1}$ of an input vector $\mathbf{x}_0 \in \mathbb{R}^{N_0 \times 1}$ to an output vector $\mathbf{x}_L \in \mathbb{R}^{N_L \times 1}$ through L iterative processing steps

$$\mathbf{x}_l = f_l(\mathbf{x}_{l-1}; \theta_l), \quad l = 1, 2, \dots, L. \quad (9)$$

Here $f_l(\mathbf{x}_{l-1}; \theta_l) : \mathbb{R}^{N_{l-1} \times 1} \mapsto \mathbb{R}^{N_l \times 1}$ is the mapping performed by the l th layer which depends on the output

$$R_k^{UP} = \log_2 \left(1 + \frac{q_k \left(\sum_{m=1}^M \gamma_{mk} \right)^2}{\sum_{k' \neq k}^K q_{k'} \left(\sum_{m=1}^M \gamma_{mk} \frac{\beta_{mk'}}{\beta_{mk}} \right)^2 + |\phi_k^H \phi_{k'}|^2 + \sum_{k'=1}^K q_{k'} \sum_{m=1}^M \gamma_{mk} \beta_{mk'} + \frac{1}{\rho} \sum_{m=1}^M \gamma_{mk}} \right). \quad (7)$$

vector \mathbf{x}_{l-1} from the previous layer and the set of learnable parameters θ_l in the l th layer. The set $\theta = \{\theta_1, \theta_2, \dots, \theta_L\}$ denotes the set of all the parameters of the network which are learnt through model training.

Input to the network is the large-scale channel coefficients β_{mk} of all the APs and users, aligned as a column vector denoted by β with dimension $N_0 = MK$. Thus, the DNN input is $\mathbf{x}_0 = \beta \in \mathbb{R}^{MK \times 1}$. The model outputs the estimated power allocation vector $\mathbf{q} = [q_1, q_2, \dots, q_K]^T \in \mathbb{A}^{K \times 1}$, where $\mathbb{A} = \{a \in \mathbb{R} : 0 \leq a \leq 1\}$. We have implemented a fully connected neural network with *Dense* layers where $f_l(\mathbf{x}_{l-1}; \theta_l)$ has the form

$$f_l(\mathbf{x}_{l-1}; \theta_l) = \sigma(\mathbf{W}_l \mathbf{x}_{l-1} + \mathbf{b}_l), \quad l = 1, 2, \dots, L, \quad (10)$$

where $\mathbf{W}_l \in \mathbb{R}^{N_l \times N_{l-1}}$ is the weight matrix, $\mathbf{b}_l \in \mathbb{R}^{N_l \times 1}$ is the bias vector. Then, the set of learnable parameters is $\theta_l = \{\mathbf{W}_l, \mathbf{b}_l\}$. In (10), $\sigma(\cdot)$ is called an *activation function* such as *ReLU*, *eLU*, *Sigmoid* etc. which introduces non-linearity to the network. For the L hidden layers in the model, we have used the *eLU* (exponential linear unit) activation function. For the output layer, *Sigmoid* activation function is used to guarantee that the outputs are in the range $[0, 1]$ adhering to the transmit power constraints $0 \leq q_k \leq 1, \forall k$ users.

We implemented a DNN with 4 layers ($L = 3$) which has $\{MK, K, M, K\}$ number of neurons in each layer and $\{eLU, eLU, eLU, Sigmoid\}$ activations respectively. This DNN has a considerably simpler structure than the DNN proposed in [12] in terms of network dimensions. Therefore, it has a lower training complexity and can produce outputs with a lower online complexity.

Given that the goal of problem P1 in (8) is to maximize the minimum user rate, we apply following loss function for model training as

$$loss = -\mathbb{E}_{\beta} [R(\beta, \theta)_{min}], \quad (11)$$

where θ denotes the set of trainable parameters in the model. There, $R(\beta, \theta)_{min} = \min_{k=1,2,\dots,K} R(\beta, \theta)_k^{UP}$ is the minimum user rate among all the K users for a given channel realization with large-scale fading of β and given θ . For each user k , $R(\beta, \theta)_k^{UP}$ is calculated from (7) using β and $\mathbf{q}(\theta)$ where $\mathbf{q}(\theta)$ is the output from DNN for given θ .

This loss function is differentiable with θ which allows training the network via stochastic gradient descent (SGD). We adopt mini-batch gradient descent approach to reduce the complexity of the SGD. In each iteration of the training, a set of channel realizations are generated from its distribution. Thus, the training loss is approximated as

$$loss \approx -\frac{1}{|\mathcal{B}|} \sum_{\beta \in \mathcal{B}} [R(\beta, \theta)_{min}], \quad (12)$$

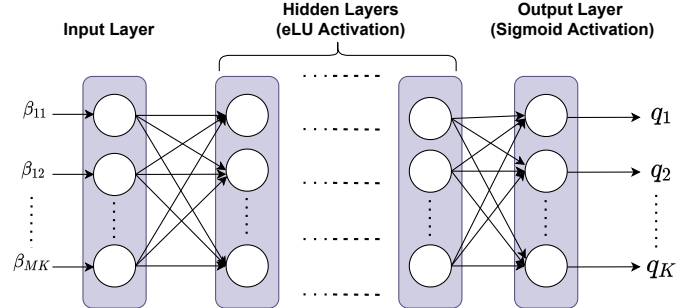


Fig. 1. Model layout of power control DNN.

where \mathcal{B} denotes the set of channel realizations in each iteration and $|\mathcal{B}|$ is the mini-batch size. Thus, during training, the model learns parameters θ to minimize the loss given in (12) which maximizes the minimum user rate as expected.

IV. SIMULATIONS AND RESULTS

In this section we present the numerical simulations to evaluate the performance of the proposed DL-based max-min fairness scheme in comparison with existing optimization-based techniques.

A. Simulation Parameters

We consider a cell-free MIMO system in a simulation area of 1×1 km^2 with different number of AP and user configurations. This square area is wrapped around at the edges to avoid boundary effects and to emulate a cell-free network with an infinite area [5]. We refer to [5] and Table I for more details about simulation parameters.

The large-scale fading coefficient β_{mk} from k th user to m th AP is given by [5]

$$\beta_{mk} = PL_{mk} 10^{\frac{\sigma_{sh} z_{mk}}{10}}, \quad (13)$$

where PL_{mk} is the pathloss from k th user to m th AP, calculated using the three-slope model used in [5] with parameters as in Table I. The shadow fading with standard deviation σ_{sh} , and $z_{mk} \sim \mathcal{N}(0, 1)$ is denoted by $10^{\frac{\sigma_{sh} z_{mk}}{10}}$.

When evaluating the spectral efficiencies, per-user net throughputs are considered as follows accounting for the channel estimation overhead as well.

$$R_k^{net} = B \frac{(1 - \tau/\tau_c)}{2} R_k^{UP}, \quad (14)$$

where R_k^{UP} is the per-user rate in (7), B is the spectral bandwidth, and τ_c is the coherence interval in samples. We have used $\tau_c = 200$ corresponding to a coherence bandwidth

TABLE I
SIMULATION PARAMETERS

Parameter	Value
Carrier frequency (f)	1.9 GHz
Simulation area length (D)	1 km
AP antenna height (h_{AP})	15 m
User antenna height (h_u)	1.65 m
d_0, d_1	10, 50 m
Bandwidth (B)	20 MHz
Noise figure	9 dB
σ_{sh}	8 dB
Pilot and data transmit powers (\bar{p}_p, \bar{p})	100, 100 mW

of 200 kHz and a coherence time of 1 ms. Orthogonal pilot assignment is considered in all simulations where $\tau = K$ so that each user is assigned with a unique orthogonal pilot sequence.

B. Simulation Setup

DNN model implementation, training and testing is done in TensorFlow [16]. Optimization-based baseline implementation is done in Matlab using CVX convex optimization software package [17], [18]. Both implementations are done on the same platform with a 4-core Intel(R) Core(TM) i5-8250U CPU with 1.6 GHz frequency.

We used three different datasets for DNN model training, validation and testing, consisting of 10^5 , 1000 and 1000 different samples respectively. For each sample, different AP and user distributions were considered with randomly generated large-scale fading channel coefficients. Input to the DNN is normalized using the training dataset mean and variance. The network is trained for 10000 iterations, using mini-batch gradient descent along with the ADAM optimizer with learning rate 0.01. In each iteration, a random mini-batch of size 100 is selected from the training dataset. In every 50 iterations, validation dataset is used to evaluate the model, where the model parameters corresponding to the minimum validation loss are preserved along the training. After training, performance is evaluated for the test dataset where the trained model is used to produce the power allocations for the test dataset and to calculate per-user rates using (7).

The GP optimization algorithm proposed in [15] is used as the baseline for performance comparison and solved using CVX. Maximum power transmission where all users transmit with full power (i.e. $q_k = 1, \forall k$) is also considered. CVX-based max-min power control results and maximum power transmission results for the test dataset are denoted as “baseline” and “maximum-power” respectively, in the results section.

C. Results and Discussion

For performance comparison, we consider two scenarios with $M = 30, K = 5$ and $M = 50, K = 10$ randomly distributed over the 1×1 km² simulation area. Obtained cumulative distribution curves are presented in Fig. 2. The DNN has close performance to the baseline in the lower net throughput range. Both the baseline and DNN have almost

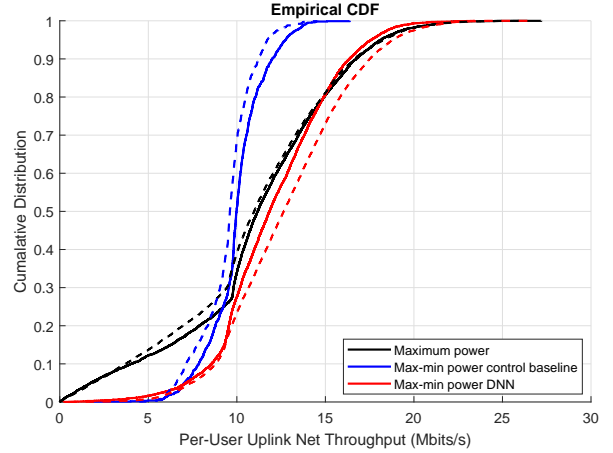


Fig. 2. Cumulative distribution of the per-user net throughput for $M = 30, K = 5$ (solid lines) and $M = 50, K = 10$ (dashed lines).

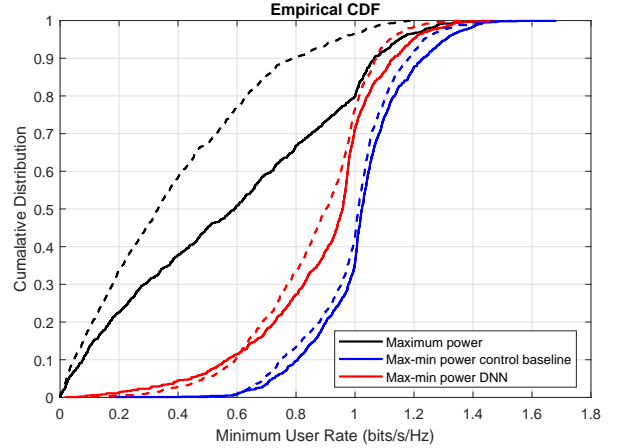


Fig. 3. Cumulative distribution of the minimum user rate for $M = 30, K = 5$ (solid lines) and $M = 50, K = 10$ (dashed lines).

same 95% likely per-user net throughput in both network configurations. However, the difference in the baseline and DNN performance suggests that the DNN may have learnt a sub-optimal power allocation scheme. In order to assess how well the DNN has achieved the desired objective of max-min user fairness, we have also evaluated the minimum user rate performance, which is shown in Fig. 3. Even though the DNN has a lower minimum user rate performance than the optimal baseline solution, it has a significant improvement over the maximum power scenario.

Furthermore, exploiting the unsupervised learning capability of the model, we have introduced online training to improve the performance of the DNN. There, the originally trained model (with the training set) is retrained for 100 iterations with learning rate 0.01 for each input sample in the test set. Performing online training allows further customization and fine-tuning of model parameters based on large-scale channel inputs in each channel realization, further improving the minimum user rate. Fig. 4 and Fig. 5 show the improved

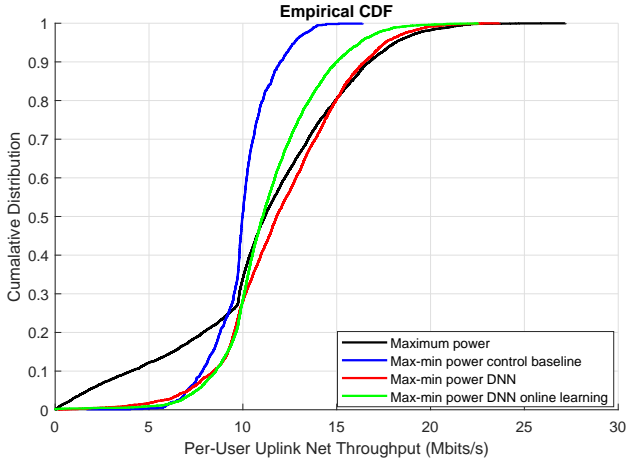


Fig. 4. Cumulative distribution of the per-user net throughput for $M = 30, K = 5$, with online training.

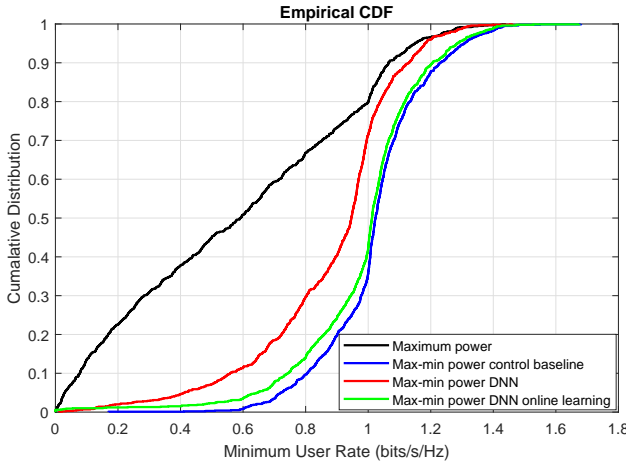


Fig. 5. Cumulative distribution of the minimum user rate for $M = 30, K = 5$ with online training.

results for the per-user rate and minimum user rate cumulative distributions with online training. It can be seen that the minimum user rates are significantly improved with online training, resulting in worst case user performance much closer to the optimal. Obtained average minimum user rates over the training and test datasets for baseline and DNN implementations in all simulation scenarios are summarized in Table II.

We also evaluated the performance of the proposed DNN model with a fixed AP setup and moving user scenario with $M = 30, K = 5$. The same DNN architecture and training process is used, but with a different data set. Here, we consider that the APs are located in a regular grid. Each user starts with a random initial position which is uniformly distributed in the coverage area, and moves in a random direction (left, right, up and down) in a random speed uniformly distributed between 0 and 20 m/s. The moving speed and direction are changed in every 5s. If a user reaches the boundary of the coverage area, then the direction is reversed so that it remains

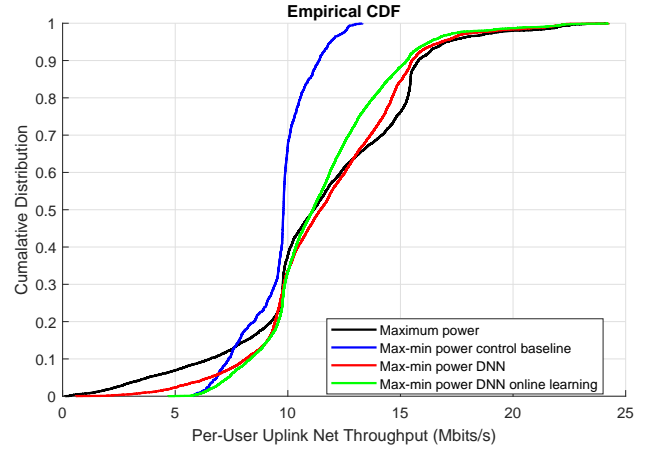


Fig. 6. Cumulative distribution of the per-user net throughput for $M = 30, K = 5$, with fixed APs and moving users.

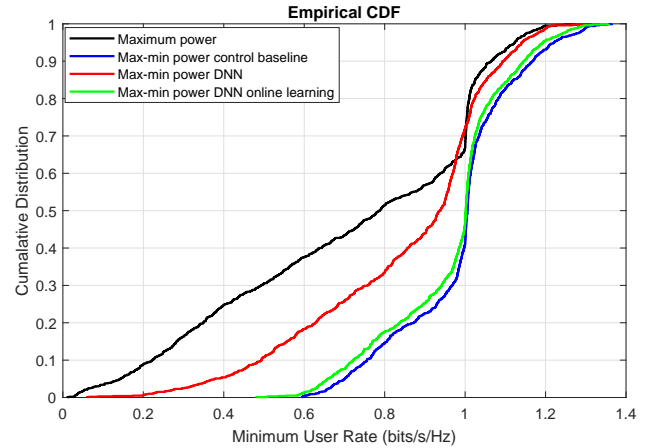


Fig. 7. Cumulative distribution of the minimum user rate for $M = 30, K = 5$, with fixed APs and moving users.

inside the area. A dataset of 12000 samples corresponding to 12000s is generated in this manner, and divided into 10000, 1000 and 1000 samples to be used as training, validation and test sets respectively. Fig. 6 and Fig. 7 show the per-user rate and minimum user rate cumulative distributions obtained for this setup. DNN implementations show better results than in the earlier random AP/user scenario, achieving near-optimal minimum user rate performance with online training. Note that here the reported performance was achieved only using a training set of 10000 samples and spending less than 10 minutes for offline model training. This shows the potential of the proposed unsupervised DNN for quick and easy deployment in a practical setup due to its low complexity offline training and online processing, and the comparable performance.

V. COMPLEXITY ANALYSIS

Here we compare the computational complexity of the baseline methods and the proposed DNN implementation for solving problem P1. Bisection-based algorithm proposed in

TABLE II
AVERAGE MINIMUM USER RATE COMPARISON FOR DNN AND BASELINE.

System Setup	Baseline	DNN	DNN with Online Training
$M = 30$ $K = 5$	Train: 1.0218 Test: 1.0221	Train: 0.8747 Test: 0.8653	Train: 0.8747 Test: 0.9817
$M = 50$ $K = 10$	Train: 0.9897 Test: 0.9940	Train: 0.7966 Test: 0.7816	Train: 0.7966 Test: 0.9199
$M = 30$ $K = 5$ Moving users	Train: 0.9889 Test: 0.9854	Train: 0.8842 Test: 0.8473	Train: 0.8842 Test: 0.9664

[5] has a complexity of $\log_2\left(\frac{t_{max}-t_{min}}{\epsilon}\right)\mathcal{O}(K^4)$ [15]. The GP-based low complexity approach proposed in [15] has $\mathcal{O}(K^{7/2})$ complexity. The newly proposed DNN for approximating solutions for the problem P1 has a complexity of $\mathcal{O}(K^2M)$ considering the dimensions of the proposed DNN model.

The recorded CPU timing for CVX solver and the DNN with and without online training, to produce outputs for 100 channel realizations for $M = 30, K = 5$ are 46.47s, 10.61s and 0.12s respectively. For $M = 50, K = 10$, respective computational times are 78.95s, 12.87s and 0.17s. Thus, the DNN is around 400 times faster than the baseline. Furthermore, DNN with online training which has near-optimal rate performance is also around 4–6 times faster than the baseline. This computational complexity of DNNs can be significantly reduced by GPU aided parallel processing implementations which are often used in DL implementations. Furthermore, it is apparent that the fast processing of DNN implementations become more significant with increasing network dimensions M and K .

VI. CONCLUSION

In this study, we have proposed an unsupervised DL-based algorithm for max-min power control for the uplink of a cell-free massive MIMO system. The proposed DNN produces sub-optimal power allocations resulting in close per-user net throughput performance compared to the GP-based optimal solution. Performing online training to customize the learnt model parameters in each channel realization to further improve the max-min performance results in near-optimal per-user and min-user rate performance at the expense of processing complexity. Nevertheless, the proposed DNN implementations are much less computationally complex than the GP-based optimization algorithm, specially for larger AP and user configurations. Furthermore, the proposed unsupervised learning approach has a lower training complexity than the supervised learning implementation in [12], and also has a much lower online complexity due to its simpler network structure compared to [12].

While this is the first time unsupervised learning-based power control is implemented for cell-free massive MIMO, it should be noted that we have analyzed the performance for a fairly a simpler network configuration with less number of APs and users than in a practical network setup. Therefore,

further investigations need to be done to understand the best DNN architectures and hyper-parameters for the proposed unsupervised learning approach to get better results in such complex setups. Nevertheless, DL-based power control in cell-free massive MIMO has research potential, especially when considering complex scenarios such as joint AP selection/user assignment and power control where conventional approaches might be sub-optimal.

REFERENCES

- [1] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 74–80, 2014.
- [2] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, 2010.
- [3] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, 2013.
- [4] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO: Uniformly great service for everyone," in *2015 IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2015, pp. 201–205.
- [5] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1834–1850, 2017.
- [6] E. Nayebi, A. Ashikhmin, T. L. Marzetta, H. Yang, and B. D. Rao, "Precoding and power optimization in cell-free massive MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 16, no. 7, pp. 4445–4459, 2017.
- [7] E. Björnson and L. Sanguinetti, "Making cell-free massive MIMO competitive with MMSE processing and centralized implementation," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 77–90, 2020.
- [8] E. Björnson and L. Sanguinetti, "Scalable cell-free massive MIMO systems," *IEEE Transactions on Communications*, vol. 68, no. 7, pp. 4247–4261, 2020.
- [9] Y. Zhao, I. G. Niemegeers, and S. H. De Groot, "Power allocation in cell-free massive MIMO: A deep learning method," *IEEE Access*, vol. 8, pp. 87 185–87 200, 2020.
- [10] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks*, vol. 2, no. 5, pp. 359 – 366, 1989.
- [11] L. Sanguinetti, A. Zappone, and M. Debbah, "Deep learning power allocation in massive MIMO," in *2018 52nd Asilomar Conference on Signals, Systems, and Computers*, 2018, pp. 1257–1261.
- [12] C. D'Andrea, A. Zappone, S. Buzzi, and M. Debbah, "Uplink power control in cell-free massive MIMO via deep learning," in *2019 IEEE 8th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, 2019, pp. 554–558.
- [13] F. Liang, C. Shen, W. Yu, and F. Wu, "Towards optimal power control via ensembling deep neural networks," *IEEE Transactions on Communications*, vol. 68, no. 3, pp. 1760–1776, 2020.
- [14] T. Van Chien, T. Nguyen Canh, E. Björnson, and E. G. Larsson, "Power control in cellular massive MIMO with varying user activity: A deep learning solution," *IEEE Transactions on Wireless Communications*, vol. 19, no. 9, pp. 5732–5748, 2020.
- [15] M. Bashar, K. Cumanan, A. G. Burr, M. Debbah, and H. Q. Ngo, "On the uplink max–min SINR of cell-free massive MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 18, no. 4, pp. 2021–2036, 2019.
- [16] M. Abadi *et al.*, "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015, software available from tensorflow.org. [Online]. Available: <https://www.tensorflow.org/>
- [17] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," <http://cvxr.com/cvx>, Mar. 2014.
- [18] M. Grant and S. Boyd, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control*, ser. Lecture Notes in Control and Information Sciences, V. Blondel, S. Boyd, and H. Kimura, Eds. Springer-Verlag Limited, 2008, pp. 95–110.