

Unsupervised Learning-Based Joint Power Control and Fronthaul Capacity Allocation in Cell-Free Massive MIMO With Hardware Impairments

Nuwanthika Rajapaksha¹, K. B. Shashika Manosha¹, Nandana Rajatheva¹, *Senior Member, IEEE*,
and Matti Latva-aho², *Senior Member, IEEE*

Abstract—A deep learning-based resource allocation algorithm that maximizes the sum rate of a limited fronthaul cell-free massive MIMO network with transceiver hardware impairments is proposed in this letter. The sum rate maximization problem with user power constraints and total fronthaul capacity constraints for channel state information (CSI) and data transmission is considered. The deep neural network (DNN) *PowerNet* is proposed to learn solutions to the joint power control and capacity allocation problem in a low-complex, flexible, and scalable way. An unsupervised learning approach is used which eliminates the need of knowing the optimal resource allocation vectors during model training, hence having a simpler and more flexible model training stage. Numerical simulations show that *PowerNet* achieves close sum rate performance compared to the existing optimization-based approach, with a significantly lower time complexity which does not exponentially scale with the number of users and access points (APs) in the network. Furthermore, the addition of the online learning stage resulted in a better sum rate than the optimization-based method.

Index Terms—Cell-free massive MIMO, sum rate maximization, power control, hardware impairments, limited fronthaul capacity, deep learning, unsupervised learning.

I. INTRODUCTION

CELL-FREE massive MIMO utilizes a large number of distributed APs which are connected to a central processing unit (CPU) via fronthaul links to serve a much smaller number of users distributed over a wide area using the same time-frequency resources, without cells or cell-boundaries [1], [2], [3]. In cell-free massive MIMO, the efficient utilization of the limited fronthaul links is essential in achieving improved performance [4], [5], [6]. Also, when looking from a practical implementation perspective with lower deployment costs, it is more realistic to consider cell-free massive MIMO systems with non-ideal transceiver components which may introduce hardware impairments to the network [6], [7], [8]. Furthermore, proper radio resource management (RRM) improves the performance of cell-free massive MIMO networks. Conventionally, RRM problems are solved

using optimization methods, which are often infeasible in practice due to the processing complexities, timing overhead, parameter sensitivity, and lack of flexibility [9]. On the other hand, the universal function approximation property of artificial neural networks has enabled deep learning (DL)-based techniques to be used for RRM with lower complexity than traditional optimization-based approaches [10], [11], [12], [13]. In this letter, we propose a DL-based novel low-complexity RRM method in order to perform uplink power control and fronthaul capacity allocation for sum rate maximization in a limited fronthaul cell-free massive MIMO network with hardware impairments. We extend the DNN-based unsupervised learning approach in our previous work [13] and show its potential to solve joint optimization tasks, which is not investigated in prior work according to our best knowledge. Unsupervised learning allows the DNN to directly optimize the max-sum rate objective as a custom loss function during training to produce power control and fronthaul capacity allocation outputs. We show that the DNN converges into acceptable solutions even with a more complex loss function consisting of the joint optimization variables, distortions due to hardware impairments in the transceivers, and quantization errors due to the limited fronthaul. The performance, time-complexity, and flexibility of the proposed DL-based solution are compared with those of the optimization-based solution proposed in [6].

II. SYSTEM MODEL

We consider the uplink data transmission of a cell-free massive MIMO system where M single-antenna APs are distributed in a given area and simultaneously serve randomly distributed K single-antenna users. The m th AP, $m \in \{1, 2, \dots, M\}$ is connected to a CPU via a fronthaul link with limited capacity C_m [bits/s/Hz]. We assume that wireless channels between APs and users are modeled as block-fading ones with coherence interval T [samples]. Thus, the channel coefficient between k th user and m th AP over a coherence interval is modeled as $g_{mk} = \sqrt{\beta_{mk}} h_{mk}$, for $m = 1, 2, \dots, M$ and $k = 1, 2, \dots, K$. 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.

We assume the network to have non-ideal transceivers at the users and the APs. In order to analyze the impact of (residual) hardware impairments on communication, we use the hardware impairment model proposed in [6] and [14], where the desired signal is scaled by a deterministic factor and an uncorrelated memoryless distortion term is added, which is Gaussian

Manuscript received 5 December 2022; revised 15 February 2023; accepted 31 March 2023. Date of publication 6 April 2023; date of current version 11 July 2023. This work was supported in part by the Academy of Finland, 6G Flagship Program under Grant 346208, and in part by the Hexa-X Project funded by the European Union's Horizon 2020 Research and Innovation Programme under Grant 101015956. The associate editor coordinating the review of this article and approving it for publication was H. Q. Ngo. (Corresponding author: Nuwanthika Rajapaksha.)

The authors are with the Centre for Wireless Communications, University of Oulu, 90014 Oulu, Finland (e-mail: nuwanthika.rajapaksha@oulu.fi; manoshad@gmail.com; nandana.rajatheva@oulu.fi; matti.latva-aho@oulu.fi). Digital Object Identifier 10.1109/LWC.2023.3265348

distributed in the worst case. We assume all the users have the same transmitter hardware quality factor ξ_t and all the APs have the same receiver hardware quality factor ξ_r [6].

The users first undergo a pilot transmission phase in order to estimate the uplink channel coefficients. We assume that τ -length orthogonal pilots are assigned to the users, where $\tau = K \leq T$. Then, the uplink data transmission happens where all the users simultaneously send their signals to the APs. We assume the compress-forward-estimate (CFE) strategy for transmission from the APs to the CPU [6]. Specifically, upon receiving the pilot/data signals, each AP compresses the received vector element-wise and transmits the quantized version to the CPU. The CPU then performs maximum ratio combining using the received quantized data signals from all the APs and the estimated channel coefficients. The use-and-then-forget approach considered in [6] is used to obtain the effective received signal for the k th user. Detailed equations in each of the above steps are omitted due to the space constraints and could be referred to from [6]. Following the derivations and notations in [6], for the CFE strategy, the achievable uplink user rate can be obtained as $R_k^{CFE} = \log_2(1 + \text{SINR}_k^{CFE})$ where SINR_k^{CFE} is given by (1), shown at the bottom of the page.

III. DEEP LEARNING-BASED JOINT POWER CONTROL AND FRONTHAUL CAPACITY ALLOCATION

As shown in [6], proper power control and fronthaul capacity allocation for pilots/CSI and data provide spectral efficiency gains in the uplink transmission. The proposed optimization-based solution, however, is sub-optimal and has a higher computational complexity to be feasible for practical implementation. In this letter, we propose a DL-based solution for the following sum rate optimization problem and compare its performance against the optimization-based solution proposed in [6].

$$P1 : \max_{\eta_k} \sum_{k=1}^K \log_2 \left(1 + \text{SINR}_k^{CFE} \right),$$

$$\text{s. t. } 0 \leq \eta_k \leq 1, \quad k = 1, 2, \dots, K,$$

$$C_{p,m} + C_{d,m} = C_m, \quad m = 1, 2, \dots, M. \quad (2)$$

In (2), the first set of constraints $\eta_k \in [0, 1], \forall k$ is for the power control coefficient of k th user. $C_{p,m}$ and $C_{d,m}$ are the fronthaul rates assigned for forwarding quantized pilot and data signals respectively which adds up to total fronthaul capacity C_m . The pilot and data signal quantization noises $Q_{p,m}$ and $Q_{d,m}$ which are needed for channel estimation and data detection are linked to $C_{p,m}$ and $C_{d,m}$ respectively, by [6, (13) and (14)].

We design and implement *PowerNet*, a DNN to solve the non-convex and NP-hard optimization problem P1 in (2), where optimal user power allocations and fronthaul capacity allocations are learned in a data-driven manner via unsupervised learning. The network consists of $L + 1$ sequentially connected layers to produce the mapping $f(\mathbf{x}_0; \boldsymbol{\theta}) : \mathbb{R}^{N_0} \mapsto \mathbb{R}^{N_L}$ of an input vector $\mathbf{x}_0 \in \mathbb{R}^{N_0}$ to an output vector $\mathbf{x}_L \in \mathbb{R}^{N_L}$. Large-scale fading coefficients vector $\mathbf{x}_0 = \boldsymbol{\beta} \in \mathbb{R}^{MK}$ is used as the input.¹ Since we need to learn optimal values for both variables: power control coefficients $\eta_k, \forall k$ and the capacity allocations $C_{p,m}, \forall m$, the hidden layers are constructed to enable learning their corresponding $\boldsymbol{\theta}$ values independently. Thus, each hidden layer and the output layer consist of two components corresponding to power coefficients and capacity allocations as $\mathbf{x}_l^{(\eta)} = f_l^{(\eta)}(\mathbf{x}_{l-1}^{(\eta)}; \boldsymbol{\theta}_l^{(\eta)})$ and $\mathbf{x}_l^{(c)} = f_l^{(c)}(\mathbf{x}_{l-1}^{(c)}; \boldsymbol{\theta}_l^{(c)}), l = 1, \dots, L$. Here $f_l^{(\eta)}(\mathbf{x}_{l-1}^{(\eta)}; \boldsymbol{\theta}_l^{(\eta)}) : \mathbb{R}^{N_{l-1}^{(\eta)}} \mapsto \mathbb{R}^{N_l^{(\eta)}}$ is the mapping performed by the l th layer corresponding to the power coefficient vector $\boldsymbol{\eta}$. Similarly, $f_l^{(c)}(\mathbf{x}_{l-1}^{(c)}; \boldsymbol{\theta}_l^{(c)}) : \mathbb{R}^{N_{l-1}^{(c)}} \mapsto \mathbb{R}^{N_l^{(c)}}$ is the mapping corresponding to $C_{p,m}, \forall m$, the fronthaul capacity allocation for the CSI. The output vector at the l th layer is constructed by concatenating those two components as $\mathbf{x}_l = [\mathbf{x}_l^{(\eta)}, \mathbf{x}_l^{(c)}]$. Furthermore, the mapping function $f_l^{(\eta)}(\mathbf{x}_{l-1}^{(\eta)}; \boldsymbol{\theta}_l^{(\eta)})$ has the form $f_l^{(\eta)}(\mathbf{x}_{l-1}^{(\eta)}; \boldsymbol{\theta}_l^{(\eta)}) = \sigma(\mathbf{W}_l^{(\eta)} \mathbf{x}_{l-1}^{(\eta)} + \mathbf{b}_l^{(\eta)})$ where $\mathbf{W}_l^{(\eta)} \in \mathbb{R}^{N_l^{(\eta)} \times N_{l-1}^{(\eta)}}$ is the weight matrix and $\mathbf{b}_l^{(\eta)} \in \mathbb{R}^{N_l^{(\eta)}}$ is the bias vector in the l th layer corresponding to power coefficients vector $\boldsymbol{\eta} = [\eta_1, \eta_2, \dots, \eta_K]^T$. The set of learnable parameters is thus $\boldsymbol{\theta}_l^{(\eta)} = \{\mathbf{W}_l^{(\eta)}, \mathbf{b}_l^{(\eta)}\}$. In the hidden layers, we have used *eLU* (exponential linear unit) for the *activation function* $\sigma(\cdot)$.² In the output layer, *Sigmoid* activation was used to guarantee that the power coefficients are in the range $[0, 1]$. Similarly, the capacity allocations $C_{p,m}, \forall m$ have the mapping $f_l^{(c)}(\mathbf{x}_{l-1}^{(c)}; \boldsymbol{\theta}_l^{(c)}) = \sigma(\mathbf{W}_l^{(c)} \mathbf{x}_{l-1}^{(c)} + \mathbf{b}_l^{(c)})$ where $\mathbf{W}_l^{(c)} \in \mathbb{R}^{N_l^{(c)} \times N_{l-1}^{(c)}}$ is the weight matrix and $\mathbf{b}_l^{(c)} \in \mathbb{R}^{N_l^{(c)}}$ is the bias vector corresponding to fronthaul capacity allocations

¹Since the objective of the DNN is sum rate optimization, giving an input to the DNN which has a direct impact on the user rates is important. We have selected large-scale fading coefficients (which consist of both the pathloss and shadow fading) as the DNN input since the user rates are mainly dependent on the large-scale fading channel values.

²ReLU activation can also be used for the hidden layer activations, but it is preferable to eLU activation since it avoids the dying ReLU problem associated with ReLU. Also, both positive and negative outputs produced by eLU help the network direct weights and biases in the right direction. From the simulations also it was observed that using eLU activations gave better results compared to ReLU.

$$\text{SINR}_k^{CFE} = \frac{\rho_u \eta_k \xi_t \xi_r \Gamma_{kk}}{\sum_{k'=1}^K \rho_u \eta_{k'} \left[\Omega_{kk'} + \xi_r (1 - \xi_t) \left(\boldsymbol{\varphi}_k^H \boldsymbol{\varphi}_{k'} + \frac{1}{\tau \xi_t} \right) \Gamma_{kk'} + \rho_p (1 - \xi_r) \left(\tau \xi_t \xi_r \boldsymbol{\varphi}_k^H \boldsymbol{\varphi}_{k'} + (1 + \xi_r - \xi_t \xi_r) \right) \Lambda_{kk'} \right] + \mathcal{E}_k}$$

$$\Gamma_{kk'} = \left(\sum_{m=1}^M \gamma_{mk} \frac{\beta_{mk'}}{\beta_{mk}} \right)^2, \quad \Omega_{kk'} = \sum_{m=1}^M \gamma_{mk} \beta_{mk'}, \quad \Lambda_{kk'} = \sum_{m=1}^M \lambda_{mk}^2 \beta_{mk'}^2, \quad \mathcal{E}_k = \sum_{m=1}^M (N + Q_{d,m}) \gamma_{mk} \quad (1)$$

$C_{p,m}$ in the l th layer. The set of learnable parameters is thus $\theta_l^{(c)} = \{\mathbf{W}_l^{(c)}, \mathbf{b}_l^{(c)}\}$. Here also, *eLU* and *Sigmoid* activations were used for the hidden layers and output layer respectively. The model output $\mathbf{x}_L \in \mathbb{R}^{(M+K)}$ is constructed combining the two components as $\mathbf{x}_L = [\mathbf{x}_L^{(\eta)}, \mathbf{x}_L^{(c)}]$, where $\mathbf{x}_L^{(\eta)} = \boldsymbol{\eta} = [\eta_1, \eta_2, \dots, \eta_K]^T$ with $\eta_k \in [0, 1]$ is the power allocation vector and $\mathbf{x}_L^{(c)} = \mathbf{c} = [c_1, c_2, \dots, c_M]^T$ with the coefficient $c_m = C_{p,m}/C_m \in [0, 1]$. Thus, the fronthaul capacity allocation at each AP for transmitting CSI, $C_{p,m}, \forall m$ can be obtained by multiplying the output from *PowerNet* by the total fronthaul capacity C_m . We implemented *PowerNet* consisting of 5 layers ($L = 4$) and $\{MK, M, M, M, K\}$ number of neurons in each layer to learn the power control coefficient vector and $\{MK, M, M, M, M\}$ number of neurons corresponding to the capacity allocation vector. Thus, the overall structure of the *PowerNet* becomes $\{MK, 2M, 2M, 2M, (M + K)\}$. Since we use an unsupervised learning approach, we use the custom loss function given by $loss = -\mathbb{E}_{\boldsymbol{\beta}}[R(\boldsymbol{\beta}, \boldsymbol{\theta})_{sum}]$ for model training, where $R(\boldsymbol{\beta}, \boldsymbol{\theta})_{sum} = \sum_{k=1}^K R(\boldsymbol{\beta}, \boldsymbol{\theta})_k^{CFE}$ is the sum rate of all the K users for a given channel realization with large-scale fading of $\boldsymbol{\beta}$ and given set of trainable parameters $\boldsymbol{\theta}$. This loss function is differentiable with $\boldsymbol{\theta}$ 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

$$loss \approx -\frac{1}{|\mathcal{B}|} \sum_{\boldsymbol{\beta} \in \mathcal{B}} R(\boldsymbol{\beta}, \boldsymbol{\theta})_{sum}, \quad (3)$$

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 $\boldsymbol{\theta}$ to minimize the loss given in (3) which maximizes the sum rate as expected.

IV. SIMULATIONS AND RESULTS

A. Simulation Setup and DNN Model Training

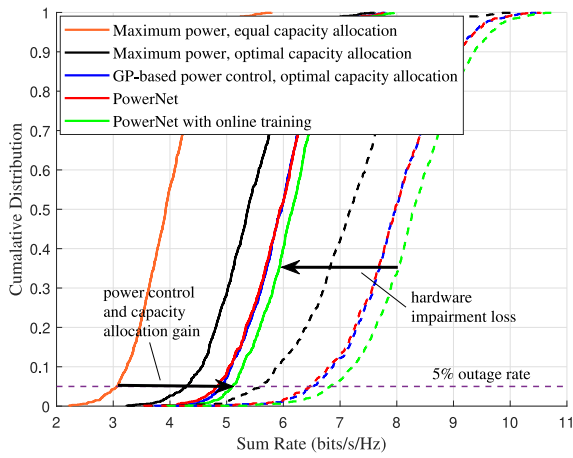
We evaluate the performance of a cell-free MIMO system in a simulation area of 1×1 km² with different numbers of AP and user configurations. All the simulation parameters and dataset preparation for the DNN model training are similar to [13]. The three-slope channel model in [2] is used for pathloss. Full details of the channel model can be found in [6]. The one-dimensional line search algorithm along with the geometric programming (GP) solution proposed in [6] for the high-SINR approximation of problem P1 is used as the optimization-based baseline method to find fronthaul capacity and user power allocations respectively. It is implemented in MATLAB using CVX convex optimization software package [15]. The DNN model implementation, training, and testing are done in TensorFlow. Both implementations are done on the same platform with a 4-core Intel(R) Core(TM) i5-8250U CPU with a 1.6 GHz frequency. We used three different datasets for training, validation, and testing the DNN model, 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 20000 iterations, using mini-batch gradient descent along with the ADAM optimizer with a learning rate of 0.005. In each iteration, a random mini-batch of size 100 is selected from the training dataset. In every 50 iterations, the 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 and capacity allocations for the test dataset and to calculate per-user rates. Furthermore, we exploit the unsupervised learning capability of the DNN and perform online learning to further improve the performance. Specifically, during the inference stage, in each channel realization, the trained model parameters are loaded to *PowerNet* and the model is retrained for 100 iterations with a learning rate of 0.005 using the custom loss function in (3) with mini-batch size $|\mathcal{B}| = 1$ (single sample) to maximize the sum rate for the given channel realization to produce power control and capacity allocation outputs. Thus, performing online learning allows further customization and fine-tuning of model parameters based on large-scale channel inputs in each channel realization, to further optimize the sum rate performance.

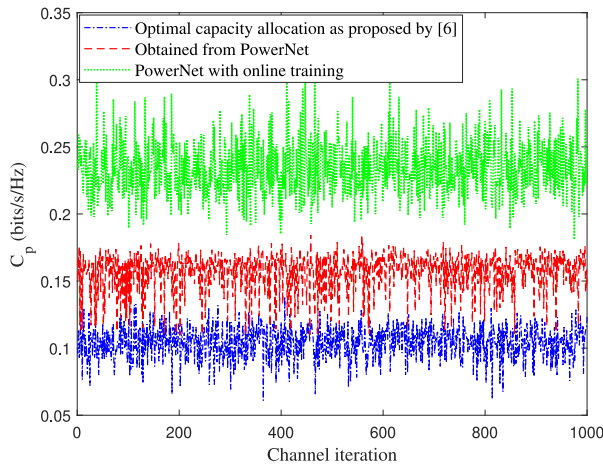
B. Results: Performance, Flexibility, and Scalability

Fig. 1(a) compares the sum spectral efficiency (SE) performance of the baseline and *PowerNet*, with different hardware quality factors. The results show that *PowerNet* achieves baseline sum-SE in both perfect and imperfect hardware scenarios. Furthermore, online learning enables outperforming the baseline approach which used GP for power control and the capacity allocation algorithm proposed in [6]. Better sum rate performance resulting from *PowerNet* with online learning doesn't necessarily mean it has achieved the global optimal solution for problem P1, but it means that it has outperformed the optimization-based baseline algorithm by converging to a better sub-optimal solution than the baseline solution. It can be explained as a result of the instantaneous model parameter optimization via online learning to achieve the sum rate maximization objective in contrast to the sub-optimal baseline solution obtained by splitting the optimization task into two sub-problems and solving them sub-optimally. Fig. 1(b) depicts the obtained optimal CSI fronthaul capacity allocations $C_{p,m}, \forall m$ from the three different methods. Note that according to the one-dimensional search algorithm proposed by [6], all the APs are allocated the same $C_{p,m} = C_p, \forall m$ which is plotted in the figure. However, since *PowerNet* outputs different $C_{p,m}$ values for each AP, the plotted C_p value for each iteration is the average CSI fronthaul capacity allocation per AP. Thus, *PowerNet* allows determining $C_{p,m}$ for each AP depending on its channel conditions which helps improve the sum rate performance as seen from Fig. 1(a).

Fig. 2 and Fig. 3 show the flexibility of *PowerNet* to be used in either ideal or imperfect hardware conditions and in



(a) Sum-SE performance comparison between *PowerNet* and baseline. Solid lines: $\xi_t = \xi_r = 0.9$ and dashed lines: $\xi_t = \xi_r = 1$ (perfect hardware).



(b) Variation of optimal CSI fronthaul capacity (C_p) allocation over channel realizations, obtained from *PowerNet*, and as proposed by [6].

Fig. 1. Joint power control and fronthaul capacity allocation for $M = 50$, $K = 10$ with CFE strategy. Total capacity $C_m = 1$ bits/s/Hz.

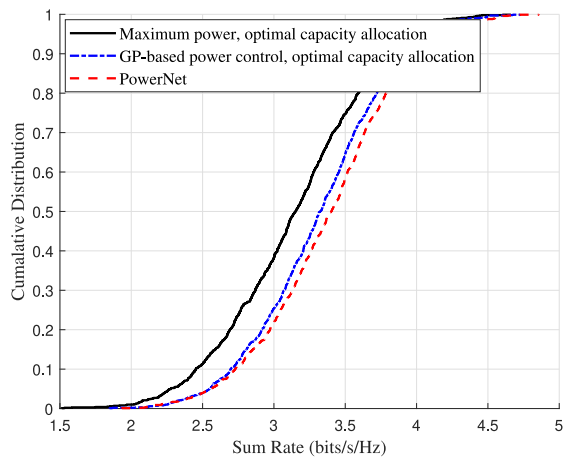


Fig. 2. Cumulative distribution of the sum rate for $M = 50$, $K = 5$, obtained using the model trained with $M = 50$, $K = 10$. Total capacity $C_m = 1$ bits/s/Hz and $\xi_t = \xi_r = 0.9$.

varying system parameters such as the number of users and total fronthaul capacity. Fig. 2 shows the cumulative distribution of the sum rate for $M = 50$, $K = 5$ where the pre-trained

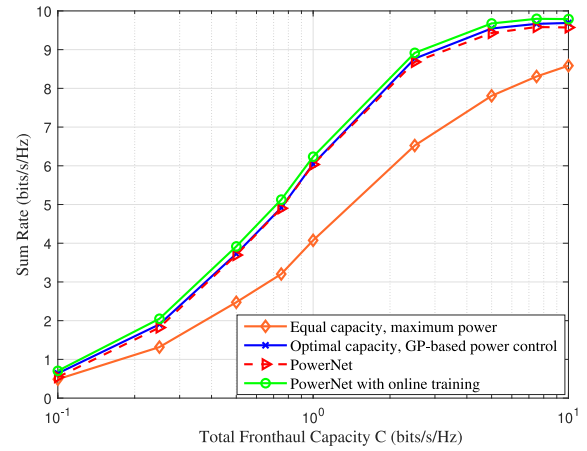


Fig. 3. Average sum rate performance with different power control and capacity allocation algorithms for $M = 50$, $K = 10$ with $\xi_t = \xi_r = 0.9$. The DNN results are obtained using the model trained with total capacity $C = 1$ bits/s/Hz.

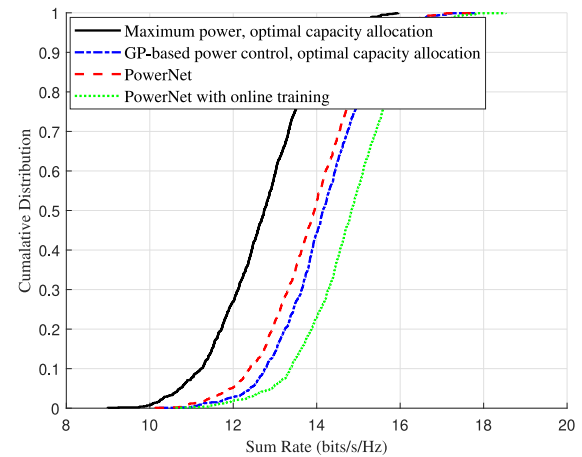


Fig. 4. Cumulative distribution of the sum rate for $M = 120$, $K = 20$ for CFE strategy. Total capacity $C_m = 1$ bits/s/Hz and $\xi_t = \xi_r = 0.9$.

model with $M = 50$, $K = 10$ is used to obtain the user power allocations and fronthaul capacity allocations. Furthermore, Fig. 3 shows the average sum rate performance for different total fronthaul capacity values $C_m = C, \forall m$. There, we have used the model trained for $C = 1$ bits/s/Hz scenario to generate the results of other instances. From the plots, we can see that *PowerNet* produces similar results to the baseline solution and online learning slightly improves the average sum rate across the whole fronthaul capacity range considered. This flexibility of the DL-based approach overcomes the challenge associated with conventional optimization techniques for resource allocation where it is needed to reformulate and solve the problem when the network parameters or system parameters change. Fig. 4 shows the sum rate performance comparison between *PowerNet* and baseline for $M = 120$, $K = 20$, showing that the proposed simple DNN structure can achieve expected results when the optimization problem is getting complex due to the increased number of users and APs. While the offline trained model performs slightly worse than the baseline, online learning outperformed the baseline, showing the scalability and learning capacity of *PowerNet* when problem complexity is increased.

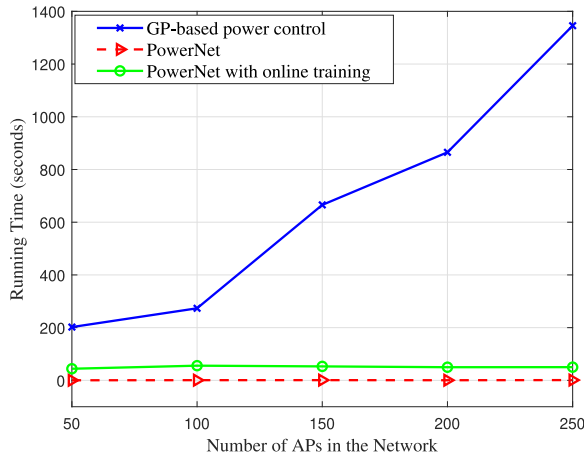


Fig. 5. Recorded CPU timing for the CVX solver and the *PowerNet* to produce outputs for 100 channel realizations for different network configurations with $M/K = 5$.

C. Computational Complexity Comparison

The GP algorithm for uplink power control has a $\mathcal{O}(K^{7/2})$ algorithmic complexity which scales with the number of users [16]. The *PowerNet* only does a one-shot calculation performing a series of matrix multiplications and additions and function mappings in each layer to produce the outputs and hence has a fixed algorithmic complexity. The total floating point operations (FLOP) count for the proposed DNN model is $4M^2K + 8M^2 + 2M(M + K)$. Fig. 5 compares the time complexity of the GP-based power control algorithm implemented using the CVX solver and the proposed *PowerNet* for solving the sum rate maximization problem. A set of 100 channel realizations are used and different network configurations are used with a different number of users and APs, keeping $M/K = 5$. The processing time for the baseline method exponentially increases with increasing M and K , whereas the processing time in *PowerNet* is only slightly increased with M and K . This is advantageous when considering practical implementations with a large number of users and APs. Furthermore, it should be noted that *PowerNet* takes less than 0.01s per sample which is very low compared to the processing time of the algorithm in [6].

V. CONCLUSION

In this letter, we proposed *PowerNet*, a DNN trained in an unsupervised manner to perform joint power control and fronthaul capacity allocation to achieve sum rate maximization in the uplink of a limited fronthaul cell-free massive MIMO system. Simulations showed that the outputs obtained from *PowerNet* along with online learning result in better sum rate performance than the optimization-based solution in [6]. *PowerNet* has a significantly lower computational complexity than the baseline and it does not exponentially scale with the number of users and APs as the existing algorithm, which is an advantage when considering practical implementations with a large number of

users and APs. The performance, low complexity, flexibility, and scalability of the proposed DL-based approach make it a potential candidate for resource allocation in a cell-free massive MIMO network considering practical scenarios.

REFERENCES

- [1] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO: Uniformly great service for everyone," in *Proc. IEEE 16th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, 2015, pp. 201–205.
- [2] 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.
- [3] E. Nayebi, A. Ashikhmin, T. L. Marzetta, H. Yang, and B. D. Rao, "Precoding and power optimization in cell-free massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4445–4459, Jul. 2017.
- [4] M. Bashar, K. Cumanan, A. G. Burr, H. Q. Ngo, E. G. Larsson, and P. Xiao, "On the energy efficiency of limited-backhaul cell-free massive MIMO," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2019, pp. 1–7.
- [5] M. Bashar, K. Cumanan, A. G. Burr, H. Q. Ngo, M. Debbah, and P. Xiao, "Max-min rate of cell-free massive MIMO uplink with optimal uniform quantization," *IEEE Trans. Commun.*, vol. 67, no. 10, pp. 6796–6815, Oct. 2019.
- [6] H. Masoumi and M. J. Emadi, "Performance analysis of cell-free massive MIMO system with limited fronthaul capacity and hardware impairments," *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 1038–1053, Feb. 2020.
- [7] J. Zhang, Y. Wei, E. Björnson, Y. Han, and X. Li, "Spectral and energy efficiency of cell-free massive MIMO systems with hardware impairments," in *Proc. 9th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, 2017, pp. 1–6.
- [8] J. Zhang, Y. Wei, E. Björnson, Y. Han, and S. Jin, "Performance analysis and power control of cell-free massive MIMO systems with hardware impairments," *IEEE Access*, vol. 6, pp. 55302–55314, 2018.
- [9] F. Hussain, S. A. Hassan, R. Hussain, and E. Hossain, "Machine learning for resource management in cellular and IoT networks: Potentials, current solutions, and open challenges," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 2, pp. 1251–1275, 2nd Quart., 2020.
- [10] 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. 87185–87200, 2020.
- [11] C. D'Andrea, A. Zappone, S. Buzzi, and M. Debbah, "Uplink power control in cell-free massive MIMO via deep learning," in *Proc. IEEE 8th Int. Workshop Comput. Adv. Multi-Sens. Adapt. Process. (CAMSAP)*, 2019, pp. 554–558.
- [12] T. Van Chien, T. N. Canh, E. Björnson, and E. G. Larsson, "Power control in cellular massive MIMO with varying user activity: A deep learning solution," *IEEE Trans. Wireless Commun.*, vol. 19, no. 9, pp. 5732–5748, Sep. 2020.
- [13] N. Rajapaksha, K. B. S. Manosha, N. Rajatheva, and M. Latva-Aho, "Deep learning-based power control for cell-free massive MIMO networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2021, pp. 1–7.
- [14] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Found. Trends Signal Process.*, vol. 11, nos. 3–4, pp. 154–655, 2017. [Online]. Available: <http://dx.doi.org/10.1561/20000000093>
- [15] M. Grant and S. Boyd. "CVX: MATLAB software for disciplined convex programming, version 2.1." Mar. 2014. [Online]. Available: <http://cvxr.com/cvx>
- [16] 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 Trans. Wireless Commun.*, vol. 18, no. 4, pp. 2021–2036, Apr. 2019.