Clustered Federated Learning for Massive MIMO Power Allocation

Yushu Yan*, Xiangyu Chang*, Amit K. Roy-Chowdhury*,
Srikanth Krishnamurthy*, Ananthram Swami**, Basak Guler*
*University of California, Riverside, USA
** DEVCOM Army Research Laboratory, USA
Email: *{yyan064@, cxian008@, amitrc@ece., krish@cs., bguler@ece.}ucr.edu, ** ananthram.swami.civ@army.mil

MILCOM 2024 - 2024 IEEE Military Communications Conference (MILCOM) | 979-8-3503-7423-0/24/$31.00 ©2024 IEEE | DOI: 10.1109/MILCOM61039.2024.10773825

***Abstract*—Power allocation in cell-based massive MIMO networks has been studied extensively in recent years. One major bottleneck for traditional optimization techniques for massive MIMO power allocation is the processing time. Recently, federated learning (FL) has been leveraged to address this challenge, by training a machine learning model for optimal power allocation, where the model is trained using the channel dynamics across distributed parties. However, current FL mechanisms train a single model to serve all users, which often fails to perform consistently across all scenarios, particularly when the training dataset exhibits high heterogeneity. To address the demands of real-time systems under heterogeneous network conditions, we propose clustered FL for massive MIMO power allocation, where the training mechanism is aimed at achieving max-min fairness of downlink spectral efficiency across heterogeneous network dynamics. To do so, personalized models are trained for different clusters of user equipments (UEs) with similar channel dynamics. A sample-efficient contextual multi-armed bandit (CMAB) mechanism is then implemented to accelerate the clustered FL training process. Our experiments demonstrate that the proposed distributed learning framework improves the convergence rate significantly compared to heterogeneity-agnostic FL mechanisms, while enhancing robustness to noisy datasets.**

***Index Terms*—Clustered federated learning, contextual multi-armed bandits, power allocation, massive MIMO.**

I. Introduction

Massive multiple-input multiple-output (MIMO) networks represent a specialized form of MIMO technology, characterized by the deployment of a significantly larger number of antennas at the transmitter compared to the number of users they support [1]–[3]. Massive MIMO typically operates in time-division duplexing (TDD) mode, where the uplink and downlink signals are transmitted on the same frequency band but in different time frames, thereby maintaining channel reciprocity. In doing so, the channel can be estimated once for each coherence period. Massive MIMO enhances spectral efficiency through its key characteristic of favorable propagation, where the channels of different users become asymptotically orthogonal [4]. In addition, communication channels are typically influenced by both large-scale and small-scale fading. Large-scale fading is mainly determined by the distance and remains approximately constant over multiple coherence times, whereas small-scale fading changes rapidly with each coherence time. To counteract small-scale fading and improve spectral efficiency, which requires adaptive adjustment of the transmit power, massive MIMO stabilizes the channel by deploying a large number of antennas and implementing receive combining, e.g., maximum ratio (MR) combining. The small-scale fading can be averaged out due to the law of large numbers. This effect, known as channel hardening, is another beneficial property [5].

Even though massive MIMO benefits from channel hardening, resource allocation is still essential to meet user-specific performance constraints. Reference [6] has studied power allocation in cell-free massive MIMO, by maximizing the sensing signal-to-interference-plus-noise ratio (SINR) while ensuring that power consumption does not exceed the maximum power constraints. The problem is then addressed through iterative convex approximations. The key challenge with optimization-based mechanisms is their high latency due to the need for channel state information (CSI), where pilot-based channel estimation is often applied, and that the computational overhead increases notably when optimizing across a large number of variables. Data-driven power allocation mechanisms are proposed in [7] and [8] to address this challenge, where deep neural networks are proposed for power allocation across cellular and cell-free massive MIMO systems, respectively. These mechanisms take large-scale fading information as input and output power allocation decisions, speeding up real-time power allocation significantly. The main challenges in training such deep learning models are two-fold: 1) insufficiency of data at a single data center, as in real world applications, data collection is costly, and a single center may not be able to gather enough data to train an optimal deep neural network model effectively, 2) high costs associated with sharing data across different data centers, primarily due to the significant communication overhead.

To address these challenges, [9] introduces an FL framework, where each edge processor (EP) gathers a local dataset to train a model for power allocation for the access points (APs) within the range of the EP. The local models trained by each EP are then aggregated to develop a more capable global model, by combining the local information gathered

Research was sponsored in part by the OUSD(R&E)/RT&L under Cooperative Agreement Number W911NF-20-2-0267, NSF CAREER Award CCF-2144927, NSF CNS grant 2106982, and the UCR OASIS Fellowship. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the ARL and USD(R&E)/RT&L or the U.S. Government.

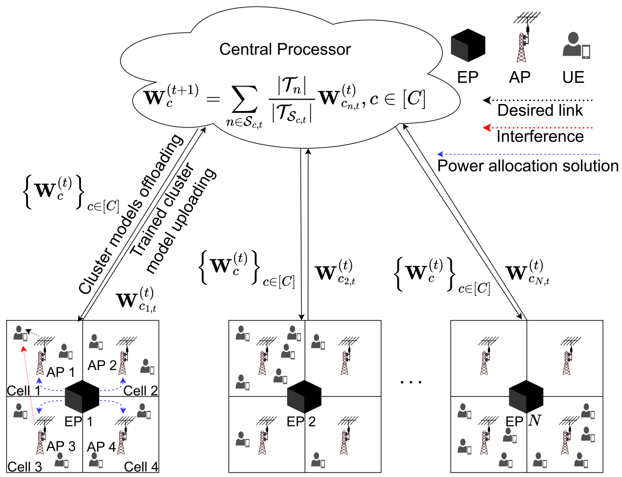

Fig. 1. Clustered FL framework for downlink transmit power allocation

by each local model. The aggregation process is coordinated iteratively by a central processor, where at the start of each round, the central processor distributes the current state of the global model to all EPs. The EPs train the received model using their own locally collected datasets. Selected EPs then send their trained models back to the central processor, which combines the local models to update the global model. In doing so, the FL significantly accelerates model convergence and performance compared to local training, while alleviating the need to collect all data at a central processor. On the other hand, [9] does not address the severe heterogeneity in the environment conditions that may be observed across different EPs in practice. When different sets of users are subject to different network conditions, which may lead to severe heterogeneities across the datasets collected by local EPs, training a single global model to serve all users can significantly deteriorate the network performance [10].

To overcome these challenges, in this work we propose a clustered FL framework for massive MIMO power allocation. Instead of training a single global model to serve all users, we train different global models for groups of users with different data distributions [10]. Due to the lack of prior knowledge about the clusters and the data distributions of the users, clustering is carried out simultaneously during model training. To do so, the central processor sends to each EP a set of C global models. The EP then selects the model that best represents its local dataset, and then trains a local model by using the selected global model. The central processor then updates each global model by aggregating the local models from the corresponding group of users. To accelerate the convergence rate and enhance the robustness of the proposed clustered FL framework, we apply a CMAB approach [11] for EP selection, without requiring prior knowledge of the dataset distribution among EPs.

II. System Model

We consider a multi-cell massive MIMO network illustrated in Fig. 1, with N EPs spread over a broad geographic region. Each EP manages a distinct area divided into L cells. At the center of each cell is an AP with M antennas. With a slight abuse of notation, we use $j \in [L]$ to refer to both the index of each cell and its corresponding AP. AP j serves K_j single-antenna UEs, where $k \in [K_j]$. Each AP j exclusively communicates with the UEs within its own cell, while being subject to interference from neighboring APs. Each EP is linked to its designated APs to manage power allocation strategies. We assume that the EPs are positioned sufficiently far from one another, and the interference across the APs served by different EPs are negligible. Our system model follows the setup considered in [9], but we drop the assumption of network homogeneity and allow the number of EPs with different UE topologies be heterogeneous. The EPs are coordinated by a central processor, which is tasked with coordinating the training of the FL model, which includes EP selection, model offloading, and aggregation in the clustered FL framework as will be described later.

A. Massive MIMO Downlink Communication

We define the channel gain between AP $j \in [L]$ and UE k in cell $l \in [L]$ as $\mathbf{h}_{l,k}^j$, which follows a complex Gaussian distribution $\mathcal{CN}(\mathbf{0}_M, \mathbf{R}_{l,k}^j)$, where $\mathbf{R}_{l,k}^j$ represents the spatial correlation matrix. The channel gains across different users are mutually independent, including users within the same cell. The average channel gain per antenna between AP j and UE k in cell l is given by $\frac{1}{M}\mathrm{Tr}(\mathbf{R}_{l,k}^j) = \beta_{l,k}^j$, which corresponds to the large-scale fading coefficient modeled as in [12],

$$\beta_{l,k}^j = \Upsilon - 10\alpha \log_{10}(d_{l,k}^j) + F_{l,k}^j, \tag{1}$$

where Υ is the median channel gain at a reference distance of 1 km, $d_{l,k}^j$ denotes the distance between AP j and UE k in cell l, α is the path-loss exponent, and $F_{l,k}^j$ denotes the shadow fading which is normally distributed with a variance of σ_{sf}^2. In downlink communication, AP j transmits the downlink signal,

$$\mathbf{x}_j = \sum_{k=1}^{K_j} \boldsymbol{\psi}_{j,k} s_{j,k}, \tag{2}$$

where $\boldsymbol{\psi}_{j,k}$ is the precoding vector, and $s_{j,k} \sim \mathcal{CN}(0, \rho_{j,k})$ is the downlink data signal intended for UE k in cell j, such that $\rho_{j,k}$ denotes the average transmit power. The precoding vector is normalized such that $\mathbb{E}\{\|\boldsymbol{\psi}_{j,k}^{\mathrm{H}}\boldsymbol{\psi}_{j,k}\|^2\} = 1$, to ensure that the expected power of the signal transmitted to UE k is equal to $\rho_{j,k}$, i.e., $\mathbb{E}\{\|\boldsymbol{\psi}_{j,k}^{\mathrm{H}} s_{j,k}\|^2\} = \rho_{j,k}$. We consider imperfect CSI between the AP and UEs, where the channel between AP j and UE k in cell l is modeled as,

$$\widehat{\mathbf{h}}_{l,k}^j = \mathbf{h}_{l,k}^j + \mathbf{n}_{l,k}^j \tag{3}$$

where each element of the noise vector $\mathbf{n}_{l,k}^j$ is assumed to follow a complex Gaussian distribution $\mathcal{CN}(0, \sigma^2)$. We consider maximum ratio combining for the precoder $\boldsymbol{\psi}_{j,k}$,

$$\boldsymbol{\psi}_{j,k} = \frac{\widehat{\mathbf{h}}_{j,k}^j}{\|\widehat{\mathbf{h}}_{j,k}^j\|} \tag{4}$$

The downlink ergodic spectral efficiency of UE k in cell j can then be lower bounded by [12],

$$SE_{j,k} = \frac{\tau_d}{\tau_c}\log_2\left(1 + \frac{\rho_{j,k}a_{j,k}}{\sum_{l=1}^{L}\sum_{i=1}^{K_l}\rho_{l,i}b_{l,i,j,k} + \sigma_{DL}^2}\right) \quad (5)$$

where

$$a_{j,k} = |\mathbb{E}\{\boldsymbol{\psi}_{j,k}^{\mathrm{H}}\mathbf{h}_{j,k}^{j}\}|^2 \quad (6)$$

and

$$b_{l,i,j,k} = \begin{cases} \mathbb{E}\{|\boldsymbol{\psi}_{l,i}^{\mathrm{H}}\mathbf{h}_{j,k}^{l}|^2\}, & (l,i) \neq (j,k) \\ \mathbb{E}\{|\boldsymbol{\psi}_{j,k}^{\mathrm{H}}\mathbf{h}_{j,k}^{j}|^2\} - |\mathbb{E}\{\boldsymbol{\psi}_{j,k}^{\mathrm{H}}\mathbf{h}_{j,k}^{j}\}|^2, & (l,i) = (j,k) \end{cases} \quad (7)$$

where σ_{DL}^2 denotes the downlink noise power, τ_c denotes the total number of samples in one coherence-block, and τ_d represents the number of samples used for downlink communication. Expectations are defined with respect to the channel and noise distributions. Due to the lack of a closed form solution, in our experiments, we consider the empirical average evaluated over multiple channel realizations.

B. Problem Formulation

We consider the downlink transmit power allocation problem aimed at achieving max-min fairness in downlink spectral efficiency among the APs and UEs served by a given EP. The optimization problem can be formulated as follows,

$$\begin{aligned} \max_{\{\rho_{j,k},\forall j,k\}} \quad & \min_{j,k} SE_{j,k} \\ \text{subject to} \quad & \sum_{k=1}^{K_j}\rho_{j,k} \leq \rho_{\max}, \forall j \in [L] \end{aligned} \quad (8)$$

where $\rho_{\max}$ denotes the maximum transmit power of any given AP. We assume that all APs share the same hardware setup.

It has been observed in [9] that solving (8) through conventional methods, such as the bisection approach [12], can result in high latency primarily, failing to meet the real-time requirements of large-scale systems [2], [3]. One could develop a separate DNN to learn the power allocation policy, so as to speed up the allocation problem at operation time. Training across regions (EPs) could help if other regions have similar features. Recently, [9] has proposed FL to speed-up real-time power allocation, but considers a homogeneous user distribution across the geographical regions served by different EPs, which can degrade performance in real-world settings when network topologies are highly heterogeneous, such as rural versus urban environments. In this work, our goal is to address this challenge, by training personalized models to serve users under different network conditions, to support power allocation in highly heterogeneous environments. To do so, we leverage clustered FL, and gradually cluster the groups of EPs with similar characteristics, and adapt the trained models to learn from the data collected from each cluster. In doing so, we enable coordination between EPs with similar network characteristics, while avoiding the potential performance degradation in training performance due to network heterogeneity. As will be detailed later, however, we further consider the impact of imperfect CSI in dataset distribution.

C. Dataset Construction

The large-scale fading coefficient $\boldsymbol{\beta}$ has been proven effective for neural networks to learn optimal power allocations for massive MIMO networks with max-min fairness [2], [3]. The model obviates the need to estimate parameters $a_{j,k}$ and $b_{l,i,j,k}$ from (6) and (7), respectively, allowing it to directly output the transmit power strategies using only the large-scale fading coefficients. Accordingly, for our power allocation task, we consider a training set constructed similar to [2], where the features correspond to the large-scale fading coefficients $\boldsymbol{\beta} = (\beta_{1,1}^1, \ldots, \beta_{L,K_L}^L)$, which is a function of the geographical coordinates, and the labels correspond to the estimated spectral efficiency of the UEs associated with the same EP, namely the $\boldsymbol{\rho}^\star = (\rho_{1,1}^\star, \ldots, \rho_{L,K_L}^\star)$. The labels are then constructed by solving (8) using the CVX toolbox in MATLAB [13]. As will be detailed in our experiments, we study the performance under both noise-free and noisy labels. For the former, it is assumed that APs can obtain perfect CSI, i.e., $\sigma^2 = 0$. For the latter, we consider varying channel estimation qualities at the APs, by varying $\sigma^2 > 0$. The L APs served by a given EP $n \in [N]$ then sends the (imperfect) CSI from (3) to the EP, after which the EP solves (8) using the CVX toolbox to locally generate the labels. We assume that the EP lacks knowledge of the noise power σ^2, and whether the generated data point is noise-free or noisy.

III. Power Allocation with Clustered Federated Learning

In this section, we first introduce our clustered FL framework for power allocation, which supports group-level personalized distributed learning without necessitating dataset transmission. Following this, we present the CMAB mechanism to accelerate the convergence of clustered FL while enhancing robustness against noisy data resulting from imperfect CSI.

A. Clustered FL Framework for Power Allocation

We assume that N EPs support C different UE topologies, such as rural vs urban environments. Our objective is to train C machine learning models, where each model is tailored for power allocation across the cluster of users with a distinct topology. We further assume that the central processor does not have any prior information regarding the specific cluster to which each EP belongs. Unlike traditional FL, which focuses on training a single global model through distributed training and model aggregation, clustered FL addresses the high heterogeneity of local datasets by training different personalized models for users with different data distributions, as suggested by [10]. In doing so, the aim is to train a personalized global model for each cluster, with the assumption that the maximum number of clusters C is known.

The cluster models $\{\mathbf{W}_c^{(0)}\}_{c=1}^{C}$ are initialized randomly at the central processor. At the beginning of each global training round t, the central processor broadcasts the cluster models

$\{\mathbf{W}_c^{(t)}\}_{c=1}^C$ to the EPs. Each EP $n \in [N]$ chooses the model that minimizes the loss with respect to its local training set,

$$c_{n,t} = \arg\min_{c\in[C]} F(\mathcal{T}_n; \mathbf{W}_c^{(t)}) \tag{9}$$

where $c_{n,t}$ denotes the cluster selected by EP n at round t, $\mathcal{T}_n$ is the local training set of EP n, and $F(\mathcal{T}_n; \mathbf{W}_c^{(t)})$ is the local training loss of EP n. In doing so, each EP selects the cluster model that is better suited to its local data distribution. For the power allocation task, we let $F(\cdot)$ be the mean squared error loss. The central processor then selects a subset $\mathcal{S}_t \subseteq [N]$ of EPs for training, which will be detailed in the next section. Each chosen EP $n \in \mathcal{S}_t$ trains their selected cluster model using their local dataset, through E local training rounds using gradient descent,

$$\mathbf{W}_{c_{n,t}}^{(t)}(m+1) \leftarrow \mathbf{W}_{c_{n,t}}^{(t)}(m) - \eta\nabla F(\mathcal{T}_n; \mathbf{W}_{c_{n,t}}^{(t)}(m)) \tag{10}$$

where $0 \leq m \leq E-1$, such that $\mathbf{W}_{c_{n,t}}^{(t)}(0) \triangleq \mathbf{W}_{c_{n,t}}^{(t)}$, and η is the learning rate. After local training, the selected EPs upload their local models to the central processor. Unlike traditional FL, where the central processor aggregates the received local models into a single global model, the central processor in clustered FL aggregates the received models according to their respective clusters. This aggregation strategy ensures that each cluster's model is fine-tuned to the specific characteristics of the data within that cluster. The model aggregation process in clustered FL can be described as follows,

$$\mathbf{W}_c^{(t+1)} = \sum_{n\in\mathcal{S}_{c,t}} \frac{|\mathcal{T}_n|}{|\mathcal{T}_{\mathcal{S}_{c,t}}|}\mathbf{W}_{c_{n,t}}^{(t)}(E), \quad \forall c \in [C] \tag{11}$$

where $\mathcal{S}_{c,t} = \{n \in \mathcal{S}_t : c_{n,t} = c\}$ is the set of participating EPs that select cluster c in round t, $|\mathcal{T}_n|$ is the size of the local training set of EP n, and $|\mathcal{T}_{\mathcal{S}_{c,t}}| \triangleq \sum_{n'\in\mathcal{S}_{c,t}} |\mathcal{T}_{n'}|$ denotes the total size of the local training sets from all participating EPs that select cluster c at round t. Training continues until all cluster models converge. Once training is completed, the central processor distributes the cluster models to the EPs. Each EP then utilizes the cluster model selected using (9) to determine the power allocation policies for the UEs.

B. EP Selection for Training

In real-world resource-limited settings, having all EPs to participate at each round of training is inefficient. To address this, we employ a CMAB mechanism with Thomspon sampling, to sample a subset of $S \leq N$ EPs to participate in each training round [9], [11]. To do so, we seek to maximize the global reward defined as,

$$r^{(t)} = \min_{n\in[N],j\in[L],k\in[K_j]} SE_{j,k}^{(t)}(\mathcal{V}_n, \mathbf{W}_{c_{n,t}}^{(t)}) \tag{12}$$

where $\mathcal{V}_n$ denotes the validation dataset of EP n, where the data samples are distinct from the training set, however, have the same distribution, and $SE_{j,k}^{(t)}(\mathcal{V}_n)$ denotes the minimum spectral efficiency of UE $k \in [K_j]$ served by AP j within the coverage area of EP n at round t, evaluated using the power allocation strategy determined by the cluster model $\mathbf{W}_{c_{n,t}}^{(t)}$ selected by EP n across the validation set $\mathcal{V}_n$. The context vector for EP n at global round t is then defined as,

$$\mathbf{x}_n^{(t)} = [\bar{F}(\mathcal{T}_n; \mathbf{W}_{c_{n,t_n}}^{(t_n)}), \bar{F}(\mathcal{V}_n; \mathbf{W}_{c_{n,t}}^{(t)}), \delta_n^{(t)}]^\top, \tag{13}$$

where,

$$\bar{F}(\mathcal{T}_n; \mathbf{W}_{c_{n,t_n}}^{(t_n)}) = \frac{F(\mathcal{T}_n; \mathbf{W}_{c_{n,t_n}}^{(t_n)})}{F(\mathcal{T}_n; \mathbf{W}_{c_{n,1}}^{(1)})} \tag{14}$$

is the local training loss of EP n normalized by the training loss from the first round, where t_n denotes the most recent round in which EP n is selected by the central processor,

$$\bar{F}(\mathcal{V}_n; \mathbf{W}_{c_{n,t}}^{(t)}) = \frac{F(\mathcal{V}_n; \mathbf{W}_{c_{n,t}}^{(t)})}{F(\mathcal{V}_n; \mathbf{W}_{c_{n,1}}^{(1)})} \tag{15}$$

is the local validation loss of EP n normalized by the validation loss observed in the first round, and $\delta_n^{(t)}$ is the one-hot encoded cluster indicator. For example, if EP n selects the first cluster out of C clusters, this is encoded as $\delta_n^{(t)} = (0, \ldots, 0, 1)$.

Let $\mathbf{X}^{\text{all}}$ and $\mathbf{r}^{\text{all}}$ denote the concatenated history of the context vectors for selected EPs and the corresponding global rewards up to round $t-1$, respectively. At the first global round $t = 1$, the central processor selects all EPs, i.e., $\mathcal{S}_1 = [N]$. Then, at each global round $t \geq 2$, S EPs are selected with the highest scores, where the score of EP n is given by,

$$\text{Score}_n^{(t)} = \boldsymbol{\omega}_{\text{new}}^\top \mathbf{x}_n^{(t)}. \tag{16}$$

where $\boldsymbol{\omega}_{\text{new}}$ is sampled from the Gaussian distribution $\mathcal{N}(\mathbf{V}^{-1}\mathbf{X}^{\text{all}}\mathbf{r}^{\text{all}}, \epsilon^2\mathbf{V}^{-1})$, such that $\mathbf{V} = \mathbf{X}^{\text{all}\top}\mathbf{X}^{\text{all}} + \lambda\mathbf{I}$, and $\mathbf{I}$ denotes the identity matrix, $\lambda > 0$ is a regularization parameter, and $\epsilon > 0$ indicates the exploration parameter. We note that the Sherman-Morrison formula can be used to update the $\mathbf{V}^{-1}$ efficiently, and that the matrix $\mathbf{X}^{\text{all}}$ and $\mathbf{r}^{\text{all}}$ do not need to be explicitly stored.

IV. Numerical Evaluations

We consider a massive MIMO network that consists of 15 EPs, each supporting four adjacent cells. Each cell is structured as a $500m \times 500m$ area with an AP centrally located with fixed position. Each AP is equipped with $M = 100$ antennas. A warp around topology is utilized as in [13]. Communication is done over a 20 MHz bandwidth, with a noise power of -94 dBm at the receiver. We assume 15 EPs are distributed across a vast area, and the distance between each EP is sufficiently large so that interference between different EPs is negligible.

We consider 8 UEs distributed across the 4 cells served by any given EP. We then generate three distinct clusters as illustrated in Fig. 2, with different number of UEs per cell. Specifically, in cluster 1, UEs are uniformly distributed across the four cells. In cluster 2, UEs are concentrated primarily in one cell, with the other cells having fewer and more scattered UEs. In cluster 3, most UEs located in two cells, while the remaining cells are almost empty. Each EP is associated exclusively with one cluster without prior knowledge of the cluster assignments of other EPs.

The distribution of EPs across clusters is unbalanced: out of a total of 15 EPs, two belong to Cluster 1, three to Cluster 2,

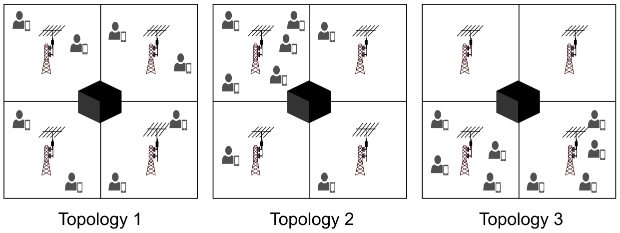

Fig. 2. Heterogeneous UE distribution topologies.

and ten to Cluster 3. As discussed in Section II-C, to simulate the impact of imperfect CSI, we use complex Gaussian noise as shown in (3) before solving (8). The proportion of noisy data points in the local datasets is heterogeneous across the EPs, with the noise power being consistent across all EPs. Specifically, Cluster 1 has a noise-free dataset for the first EP while the second EP has 10% of its data points affected by noise. In Cluster 2, the first two EPs experience the same noise distribution as those in Cluster 1, whereas the third EP has 20% of the data points noisy. In Cluster 3, the noisy data ratio is gradually increased from 0% to 40% across the EPs.

Each EP has a local dataset consisting of 1100 samples, divided into training, validation, and testing sets according to a $9:1:1$ ratio uniformly at random. During training, the learning rate is initially set as $\eta = 10^{-3}$ and is reduced by 0.2% after each round, along with the Adam optimizer [14] to facilitate efficient model training. Each data sample is pre-processed by expressing the features $\boldsymbol{\beta}$ as a 3D matrix of size $K_{\max} \times L \times L$ through zero padding, where $K_{\max}$ denotes the maximum number of UEs that an AP serves across all EPs. We then employ a 3D-ResNet [15] to learn the mapping between $\boldsymbol{\beta}$ and the optimal transmit power $\boldsymbol{\rho}^{\star}$. The architecture of the proposed 3D-ResNet is based on the traditional ResNet18 [16], but modified to include all layers in 3D. The model inputs are the large-scale fading coefficients between all the APs and UEs covered by the same EP. Following the input layer, there is one convolutional layer equipped with 64 filters of size $7 \times 7 \times 7$, stride of 2, and padding of 3, followed by batch normalization and ReLU activation. This is succeeded by a max pooling layer with dimensions $3 \times 3 \times 3$, stride of 2, and padding of 1. The core of the model consists of four consecutive residual blocks; each block contains two convolutional layers, with batch normalization and ReLU activation. The number of filters in each convolutional layer doubles with every block, starting from 64. Each residual block includes a skip-connection, featuring a convolutional layer with batch normalization to align the dimensions of the shortcut output with the main path. The filter size for each convolutional layer is $3 \times 3 \times 3$, except for the first convolutional layer. Finally, the architecture utilizes an average pooling layer and a fully connected layer to map the features to the output layer, which specifies the power allocation for the UEs served by the same EP.

To evaluate the performance of our clustered FL framework with context-based EP selection, we consider the baselines,

- *Clustered FL + CMAB (our work)*: 3 EPs are sampled based on the bandit score computed as described in (16), then each selected EP selects one cluster as described in (9) and performs local training with $E = 1$ local training round. After local training, the central processor aggregates the uploaded models as in (11) to update the global model.
- *Centralized training* [2]: The local datasets from all EPs are gathered at the central processor, who then trains a single model. This baseline represents the conventional centralized training mechanism.
- *Clustered FL+uniform sampling* [10]: In this baseline, 3 EPs are uniformly sampled at each global round (as opposed to CMAB-based EP sampling), and then each selected EP selects one cluster according to (9). After local training, the central processor aggregates the uploaded models according to (11). This baseline studies the impact of uniform sampling for EP selection on the model accuracy.
- *Global aggregation* [17]: This baseline represents the classical FL framework, where a single global model is initialized at the central processor, and all N EPs participate in local training and model updating. This baseline is also the performance upper bound to the FL baseline for massive MIMO power-allocation from [9].
- *Decentralized training* [2]: In this setting, EPs train their own local models using their local datasets only, without any collaboration across different EPs.

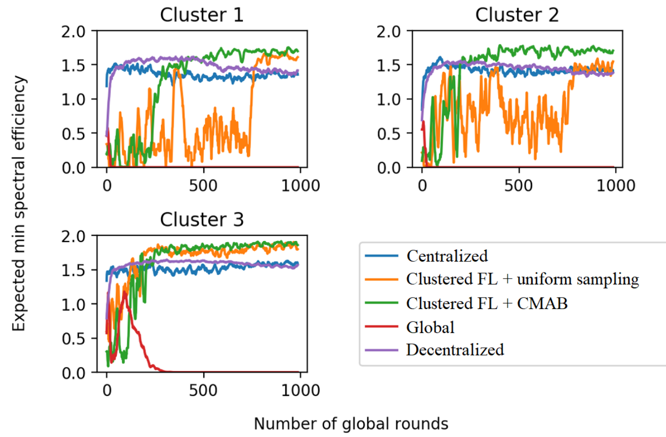

Fig. 3. Expected minimum spectral efficiency for each cluster under perfect CSI.

In Fig. 3, we compare the expected minimum spectral efficiency across each cluster for all baselines, by assuming each EP is equipped with noise-free local datasets with perfect channel estimation. We observe that our framework achieves the best performance across all clusters. Clustered FL with uniform sampling outperforms both decentralized training and global aggregation baselines, yet is outperformed by CMAB-based EP sampling. Clustered FL with CMAB-based EP sampling, on the other hand, where the central processor selects the EPs based on the bandit scores, improves the performance against uniform sampling by 4%, decentralized training by 10%, and centralized training by 13%. As such,

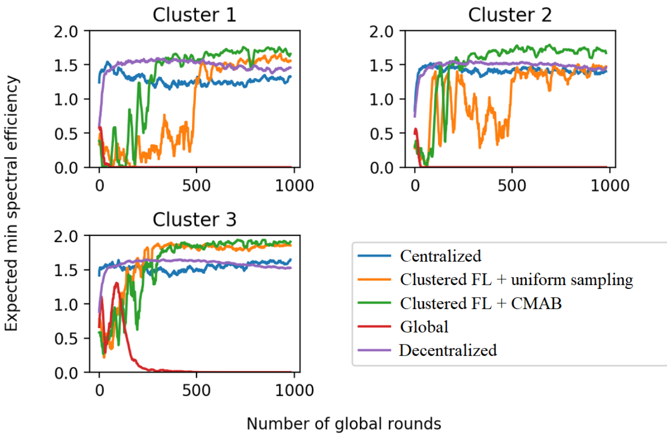

Fig. 4. Expected minimum spectral efficiency for each cluster under mixed dataset with imperfect CSI ($\sigma^2 = 5 \times 10^{-4}$)

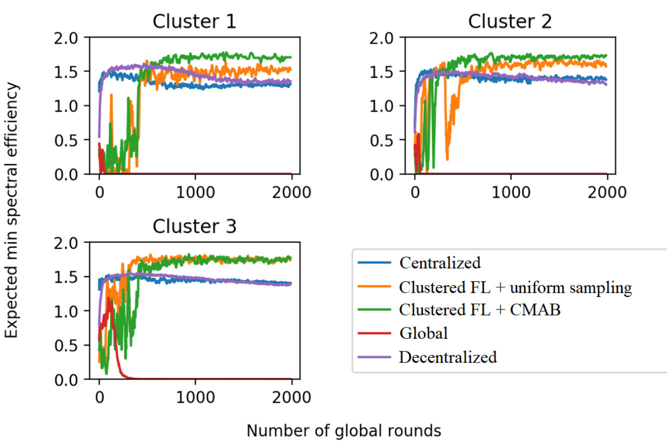

Fig. 5. Expected minimum spectral efficiency for each cluster under mixed dataset with imperfect CSI ($\sigma^2 = 1 \times 10^{-2}$)

personalized models significantly enhance performance compared to conventional baselines.

In Figs. 4 and 5, we consider imperfect CSI using complex Gaussian noise as described in (3), to simulate varying channel estimation qualities, with noise powers $\sigma^2 = 5 \times 10^{-4}$ and $\sigma^2 = 1 \times 10^{-2}$, respectively. In both scenarios, we observe that conventional FL with global aggregation fails to converge, while clustered FL outperforms both centralized and decentralized training baselines. Additionally, we observe that training a single global model fails to achieve consistent performance across all clusters. In decentralized training, the model performance is constrained by the disparate proportions of noisy data across the EPs, which can be attributed to the absence of cooperation enabled by FL. For Cluster 3, the performance of clustered FL with uniform sampling and the proposed clustered FL with CMAB-based sampling are comparable. On the other hand, for Clusters 1 and 2, clustered FL with CMAB-based EP sampling algorithm achieves superior performance. Hence, the models associated with clustered FL with CMAB-based EP selection demonstrate faster convergence and higher expected minimum spectral efficiency across all clusters.

V. Conclusion

In this paper, a clustered FL framework is proposed for massive MIMO power allocation under heterogeneous environments, by utilizing large-scale fading coefficients between APs and UEs. The proposed framework can enhance the training performance in heterogeneous settings, without exchanging local datasets, and significantly speed-up power allocation in real-time delay-sensitive massive MIMO applications. Our experiments show that under highly heterogeneous settings, the proposed clustered FL mechanism achieves highest expected minimum spectral efficiency compared to the conventional baselines. The robustness of the proposed framework has further been verified by considering the impact of noisy channel state estimation during dataset construction and training.

References

[1] E. Castañeda, A. a. Silva, A. Gameiro, and M. Kountouris, "An overview on resource allocation techniques for multi-user MIMO systems," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 1, pp. 239–284, 2017.

[2] S. Chakraborty, E. Björnson, and L. Sanguinetti, "Centralized and distributed power allocation for max-min fairness in cell-free massive MIMO," in *Asilomar Conf. on Signals, Systems, and Computers*, 2019.

[3] M. Zaher, O. Tuğfe Demir, E. Björnson, and M. Petrova, "Distributed DNN power allocation in cell-free massive MIMO," in *Asilomar Conf. on Signals, Systems, and Computers*, 2021, pp. 722–726.

[4] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Aspects of favorable propagation in massive MIMO," in *2014 22nd European Signal Processing Conference (EUSIPCO)*, 2014, pp. 76–80.

[5] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: ten myths and one critical question," *IEEE Communications Magazine*, vol. 54, no. 2, pp. 114–123, 2016.

[6] Z. Behdad, O. T. Demir, K. W. Sung, E. Björnson, and C. Cavdar, "Multi-static target detection and power allocation for integrated sensing and communication in cell-free massive MIMO," *IEEE Transactions on Wireless Communications*, pp. 1–1, 2024.

[7] 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.

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

[9] Y. Yan, X. Chang, B. Guler, A. K. Roy-Chowdhury, S. V. Krishnamurthy, and A. Swami, "Federated learning for massive MIMO power allocation," in *2023 57th Asilomar Conference on Signals, Systems, and Computers*. IEEE, 2023, pp. 651–655.

[10] A. Ghosh, J. Chung, D. Yin, and K. Ramchandran, "An efficient framework for clustered federated learning," *Advances in Neural Information Processing Systems*, vol. 33, pp. 19 586–19 597, 2020.

[11] X. Chang, S. M. Ahmed, S. V. Krishnamurthy, B. Guler, A. Swami, S. Oymak, and A. K. Roy-Chowdhury, "Flash: Federated learning across simultaneous heterogeneities," *arXiv preprint arXiv:2402.08769*, 2024.

[12] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Foundations and Trends® in Signal Processing*, vol. 11, no. 3-4, pp. 154–655, 2017.

[13] L. Sanguinetti, A. Zappone, and M. Debbah, "Deep learning power allocation in massive MIMO," in *52nd Asilomar Conference on Signals, Systems, and Computers*, 2018, pp. 1257–1261.

[14] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[15] H. Kataoka, T. Wakamiya, K. Hara, and Y. Satoh, "Would mega-scale datasets further enhance spatiotemporal 3D CNNs?" *arXiv preprint arXiv:2004.04968*, 2020.

[16] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.

[17] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Int. Conf. on Artificial Int. and Stat. (AISTATS)*, 2017.