Channel-aware Joint AoI and Diversity Optimization for Client Scheduling in Federated Learning with Non-IID Datasets

Manyou Ma, *Student Member, IEEE*, Vincent W.S. Wong, *Fellow, IEEE*, and Robert Schober, *Fellow, IEEE*

***Abstract*—Federated learning (FL) is a distributed learning framework where clients jointly train a global model without sharing their local datasets. In each communication round of FL, a subset of clients are scheduled to participate in training. Recent research has shown that diversity-based FL can improve the convergence performance of FL, especially when the client datasets are not independent and identically distributed (non-IID). In this paper, we show that by considering the channel state information and age of information (AoI) of each client, the convergence of FL can further be improved. We formulate a channel-aware joint AoI and diversity-based client scheduling problem as a constrained Markov decision process (CMDP). By using Lagrangian index and one-step lookahead approaches, we develop a two-stage online algorithm which is scalable and has a low computational complexity. For FL tasks with non-IID client datasets, our results show that the proposed algorithm can speed up the convergence of FL by up to 71%, through reducing the duration of uplink transmission, when compared with three state-of-the-art FL algorithms.**

***Index Terms*—Age of information (AoI), constrained Markov decision process (CMDP), diversity, federated learning, index policy, Lagrangian index.**

I. Introduction

Federated learning (FL) [2] is a distributed learning framework, where multiple mobile clients are orchestrated by a parameter server (PS) to train a deep learning model. For the federated averaging (FedAvg) algorithm proposed in [2], multiple clients are connected to the PS through wireless links. The training phase of FL involves multiple communication rounds. At the beginning of each communication round, the PS broadcasts the updated model and schedules a subset of clients to participate in training (step ① in Fig. 1). The scheduled clients perform gradient-based learning using their local datasets (step ② in Fig. 1) and transmit the updated model back to the PS (step ③ in Fig. 1). Finally, the PS aggregates the received model updates by averaging them (step ④ in Fig. 1). In FL, client privacy is preserved since local datasets are not revealed to the PS. FL is also communication-efficient since only a subset of clients are scheduled to transmit over the wireless links in each communication round.

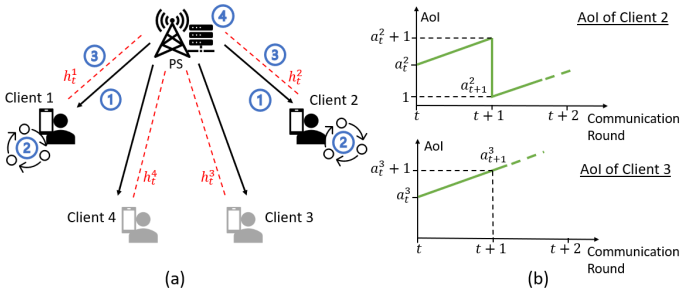


Fig. 1: (a) Illustration of a federated learning (FL) system with one parameter server (PS) and four clients. In communication round t, two clients are scheduled to participate in FL training. The solid arrows represent the downlink broadcasting of the model parameters. The red dashed lines represent the wireless channels of the clients. (b) Plots of the evolution of two clients' age of information (AoI).

Client scheduling is crucial for the convergence of FL. Some of the previous works on FL use probability-based client scheduling. In the FedAvg algorithm [2], a subset of clients are selected uniformly at random from the set of clients in each communication round. In the multinomial distribution (MD) sampling scheme [3], the probability of a client being selected is proportional to the size of its local dataset. However, probability-based client scheduling has two potential limitations: (a) the channel conditions of the clients are not considered, which may lead to a long uplink communication duration for clients with poor channel conditions; (b) the speed of convergence of the training may be slow, when the local datasets are not independent and identically distributed (non-IID).

To address the first limitation, recent works have shown that channel-aware scheduling can reduce the duration of uplink transmission in each communication round, which in turn improves convergence. A Lyapunov optimization-based algorithm using the instantaneous channel state information (CSI) of the clients has been proposed in [4]. For the case when the data distributions of the local client datasets are

Manuscript received on Nov. 1, 2022; revised on Apr. 23, 2023 and Jul. 31, 2023; accepted on Nov. 3, 2023. The work of Manyou Ma and Vincent W.S. Wong was supported in part by the Natural Sciences and Engineering Research Council of Canada and the Digital Research Alliance of Canada (alliancecan.ca). Robert Schober's work was (partly) funded by the German Ministry for Education and Research (BMBF) under the program of "Souverän. Digital. Vernetzt." joint project 6G-RIC (Project-ID 16KISK023).

This paper has been published in part in the *Proceedings of the IEEE International Conference on Communications (ICC)*, May 2023 [1]. The editor coordinating the review of this paper and approving it for publication was Jing Yang. *(Corresponding author: Vincent W.S. Wong)*

M. Ma and V. W.S. Wong are with the Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC, V6T 1Z4, Canada (e-mail: {manyoum, vincentw}@ece.ubc.ca). R. Schober is with the Institute for Digital Communications, Friedrich-Alexander University of Erlangen-Nuremberg, Erlangen 91058, Germany (e-mail: robert.schober@fau.de).

Color versions of one or more of the figures in this paper are available online at https://ieeexplore.ieee.org.

known, a client scheduling algorithm exploiting both the CSI and the total data distribution distance is proposed in [5]. In [6], a client scheduling algorithm based on each client's local gradient and instantaneous CSI is proposed. However, due to the temporal and spatial correlation of wireless channels, scheduling algorithms based on CSI and the distribution of the local datasets may lead to the same subset of clients being chosen multiple times in a short time interval, which can degrade the convergence performance of FL.

To address the second limitation, recent research has considered client scheduling based on the model updates received from all clients [7]–[9]. The premise is that clients with similar datasets will render similar model updates in each communication round. Although the data distribution for each client dataset is unknown due to privacy concerns, it can be inferred from the model update received from each client. In this way, sampling a subset of clients that possess similar datasets in each communication round can be avoided and the convergence of the learning toward the optimal global model can be ensured. In the clustered sampling approach [8], different clusters of clients are formed based on their recent model updates, and clients from different clusters are scheduled in each communication round. In [7], an algorithm called federated averaging with diverse client selection (DivFL) is proposed to schedule a diverse subset of clients that can approximate the model update from the full client participation scenario. Deep reinforcement learning [9] has also been employed to perform gradient information-based client scheduling. Since the PS does not have information on clients that did not participate in FL in the previous communication round, recent works in [7] and [8] have proposed to store the most recent model update from each client, which is called the *representative gradient*, and used this information for client scheduling. Since the aforementioned works aim to schedule a diverse subset of clients in the representative gradient space in each communication round, we will refer to them as *diversity-based* FL. In this work, we theoretically prove that stale versions of the representative gradients of the clients affect the convergence rate of FL, and hence the PS needs to ensure the freshness of the representative gradients. To the best of our knowledge, none of the existing works on client scheduling algorithm design has considered this aspect.

In this paper, we show that by considering the age of information (AoI) of each client for client scheduling, the convergence speed of diversity-based and channel-aware FL can further be improved. In the context of client scheduling in FL, the AoI of a client represents the time that has elapsed since the last time the client has been chosen to participate in FL. In [10], the average AoI of all clients is proposed as a regularization term to avoid scheduling a subset of clients exclusively in channel-aware FL. In the literature (*e.g.*, [11]–[14]), AoI has been used as an objective function or performance metric for scheduling algorithm design for different use cases and applications. These works consider multiuser scheduling problems with AoI as the optimization objective under resource constraints. In [15], the age of update is proposed as a metric to accelerate FL training. However, the diversity information of the clients' local datasets is not considered in [15].

In this work, we consider the case where the client data distribution is not known *a priori*, and the diversity information is inferred from the representative gradients of all clients. We formulate the client scheduling problem as a finite-horizon constrained Markov decision process (CMDP), which jointly optimizes the AoI, diversity, and uplink transmission duration of the scheduled clients. This is a hard problem to solve, for two main reasons:

- First, due to the temporal dependencies of the decisions made in different time instants, optimization problems with AoI as part of the objective function or constraints are sequential decision problems, which are often formulated as CMDPs. When the number of clients is large, the computational complexity for determining the optimal CMDP policy becomes high. Some recent works have proposed scalable suboptimal scheduling algorithms (*e.g.*, [16], [17]) based on the Whittle index approach [18], which has been proven to be asymptotically optimal when the number of clients approaches infinity and the percentage of scheduled clients remains constant. However, only a special class of problems (known as Whittle indexable) can be solved by the Whittle index approach. Even for a problem that can be proven to be Whittle indexable, deriving the Whittle index is still hard, and is often impossible for practical problems with large state space [19].
- Second, the inclusion of the diversity term in the optimization objective, *i.e.,* as part of the reward/cost function of the CMDP, introduces two challenges. The part of the reward/cost related to diversity can only be computed for a group of clients, and the cost/reward function cannot be decomposed into individual cost/reward functions of the clients[1]. Therefore, it is not straightforward to decompose the formulated CMDP into subproblems, where each individual cost/reward is a function of the local state and action of a client, which is a necessary condition to obtain a solution based on Whittle index. A new scalable solution approach needs to be developed to handle this cost/reward structure.

To address the first issue, in this paper, we propose a Lagrangian index-based solution [21], [22] to solve a diversity-agnostic version of the formulated CMDP, which jointly minimizes the weighted AoI and uplink transmission duration of all clients. The proposed Lagrangian index-based approach is proven to be asymptotically optimal [21], enjoys similar scalability as the Whittle index-based approach, and does not require the underlying CMDP problem to satisfy any special property. To address the second issue, we propose a two-stage online algorithm to tackle the cost function that includes diversity. In the first stage, the Lagrangian indices of all clients are determined for a diversity-agnostic variant of the formulated CMDP problem. In the second stage, we select a subset of clients to participate in FL in each time instant by jointly considering the diversity information and

[1]This problem is also referred to as the credit assignment problem [20] in the reinforcement learning literature.

the Lagrangian indices of all clients in an online manner. The contributions of this paper are as follows:

- We study the impact of the AoI of the representative gradients of the clients on the convergence performance of diversity-based FL. We show that a small AoI can improve the convergence speed of diversity-based FL algorithms.
- We formulate a channel-aware joint AoI and diversity-based client scheduling problem in FL as a CMDP, and refer to it as the *diversity-based client scheduling problem*.
- We propose a Lagrangian index-based approach to solve the diversity-agnostic variant of the client scheduling problem. The proposed Lagrangian index-based approach achieves asymptotically optimal performance.
- We also propose a two-stage online algorithm to solve the CMDP problem. In the first stage, the Lagrangian index policy for the diversity-agnostic client scheduling problem is used as the base policy. In the second stage, a one-step lookahead policy improvement [23] enables us to obtain a suboptimal solution to the diversity-based client scheduling problem.
- We evaluate the performance of the proposed two-stage online algorithm in large-scale FL experiments with 100 clients using MNIST and CIFAR-10 datasets. Simulation results show that the proposed two-stage algorithm outperforms MD sampling [3], clustered sampling [8], and DivFL [7] algorithms by up to 71% in terms of the average uplink transmission duration, achieving a better convergence performance.

The rest of this paper is organized as follows. In Section II, we introduce the FL system model and formulate the channel-aware joint AoI and diversity-based client scheduling problem as a CMDP. In Section III, we propose a Lagrangian index-based algorithm for the diversity-agnostic variant of the problem. In Section IV, we propose a two-stage online algorithm for solving the formulated CMDP. In Section V, we present numerical results and performance comparisons. Section VI concludes the paper. The analytical results related to the AoI and diversity on the convergence performance of FL are presented in the Appendix.

Notations: We use $\mathbb{C}$, $\mathbb{R}$, and $\mathbb{N}_+$ to denote the set of complex numbers, real numbers, and positive integers, respectively. We use $(\cdot)^{\mathrm{T}}$ to denote the transpose of a vector or matrix, $\mathbb{E}[\cdot]$ to denote the expectation of a random variable, and $\mathbb{1}(\cdot)$ to denote the indicator function. $\|\cdot\|$ denotes the norm of a vector. $|\cdot|$ denotes the cardinality of a set. $\mathcal{O}(\cdot)$ denotes the big-O notation for algorithmic complexity analysis.

II. System Model and Problem Formulation

A. System Model

We consider a system with a PS and N clients, *cf.* Fig. 1. The set of clients is denoted by $\mathcal{N} = \{1, 2, \ldots, N\}$. Each client has its own local dataset. We consider a time-slotted system with a finite time horizon, where the set of communication rounds is denoted by $\mathcal{T} = \{1, 2, \ldots, T\}$.

1) Scheduling Decision Vector: In communication round $t \in \mathcal{T}$, let $\mathbf{u}_t = (u_t^1, \ldots, u_t^N) \in \{0,1\}^N$ denote the scheduling decision vector determined by a decision-making module located in the PS, where

$$\text{C1: } u_t^n \in \mathcal{U}_n \triangleq \{0,1\}, \quad n \in \mathcal{N}, t \in \mathcal{T}. \tag{1}$$

Client n participates in training when u_t^n is equal to one and does not participate in training when u_t^n is equal to zero. The PS has limited capacity and can aggregate parameters of the neural network from at most M clients in each communication round. Thus, the scheduling decision vector has to be chosen from the following feasible set:

$$\text{C2: } \mathbf{u}_t \in \mathcal{U} \triangleq \big\{\mathbf{u} = (u^1, \ldots, u^N) \in \{0,1\}^N \mid \sum_{n=1}^{N} u^n \leq M\big\}, \quad t \in \mathcal{T}. \tag{2}$$

To simplify the notation, we define $\mathcal{M}_t \triangleq \{n \in \mathcal{N} \mid u_t^n = 1\}$ as the subset of clients that are scheduled to participate in FL in communication round $t \in \mathcal{T}$. We will use $\mathcal{M}_t$ and $\mathbf{u}_t$ interchangeably in the rest of the paper.

2) Federated Learning Model: We follow the notations adopted in [3]. Let $\mathcal{X}_n$ denote the set of training data samples in the local dataset of client $n \in \mathcal{N}$. Let θ denote the model being trained by the FL system and $|\theta|$ denote the number of parameters. We use vector $\mathbf{w} \in \mathcal{W} \triangleq \mathbb{R}^{|\theta|}$ to denote the weights or parameters in model θ. The goal of FL is to find the optimal weights to the optimization problem

$$\underset{\mathbf{w}}{\text{minimize}} \quad F(\mathbf{w}) \triangleq \sum_{n=1}^{N} p_n F_n(\mathbf{w}), \tag{3}$$

where $p_n \triangleq \frac{|\mathcal{X}_n|}{\sum_{i \in \mathcal{N}} |\mathcal{X}_i|}$, $n \in \mathcal{N}$. The function $F_n(\cdot)$ denotes the local objective of client n, which is defined as

$$F_n(\mathbf{w}) \triangleq \frac{1}{|\mathcal{X}_n|} \sum_{x_n \in \mathcal{X}_n} \mathcal{L}(\mathbf{w} \mid x_n), \quad n \in \mathcal{N}, \tag{4}$$

where $\mathcal{L}(\cdot)$ denotes the loss function of a client.

Before the first communication round, the weight vector is initialized to $\mathbf{w}_0$. Then, at the beginning of communication round $t \in \mathcal{T}$, the PS schedules a subset of clients $\mathcal{M}_t \subseteq \mathcal{N}$ to participate in FL and broadcasts the current global model to these clients. After obtaining the current global model, the participating clients perform $E \in \mathbb{N}_+$ local stochastic gradient descent (SGD) steps, and then forward the updated model to the PS. That is, in communication round $t \in \mathcal{T}$, the PS first broadcasts the global model $\mathbf{w}_{(t-1)E}$. When client $n \in \mathcal{M}_t$ has received the global model, it sets its local model to $\mathbf{w}_{(t-1)E}^n = \mathbf{w}_{(t-1)E}$, and performs SGD as follows

$$\mathbf{w}_{(t-1)E+i+1}^n \leftarrow \mathbf{w}_{(t-1)E+i}^n - \eta_{(t-1)E+i} \times \nabla F_n(\mathbf{w}_{(t-1)E+i}^n \mid \Xi_{(t-1)E+i}^n), \quad i \in 0, \ldots, E-1, \tag{5}$$

where η_k denotes the learning rate for the k-th SGD step, $k \in \mathcal{K} \triangleq \{0, 1, 2, \ldots, TE\}$ and Ξ_k^n is a data sample uniformly sampled from local dataset $\mathcal{X}_n$. After training, each client $n \in$

$\mathcal{M}_t$ sends its model update

$$\mathbf{q}_t^n \triangleq -(\mathbf{w}_{tE}^n - \mathbf{w}_{(t-1)E}^n), \quad t \in \mathcal{T}, \tag{6}$$

to the PS over a wireless link. When the PS has received the model updates from all scheduled clients, it updates the global model as follows

$$\mathbf{w}_{tE} \leftarrow \mathbf{w}_{(t-1)E} - \frac{N}{|\mathcal{M}_t|} \sum_{n \in \mathcal{M}_t} p_n \mathbf{q}_t^n, \quad t \in \mathcal{T}. \tag{7}$$

3) Update of Clients' AoI: In communication round $t \in \mathcal{T}$, let a_t^n denote the AoI of client $n \in \mathcal{N}$. It represents the number of communication rounds that have elapsed since client n sent its updated model to the PS. That is,

$$a_t^n = \min\{i \mid n \in \mathcal{M}_{t-i}, i \in \{1, \ldots, t\}\}, \tag{8}$$

for $n \in \mathcal{N}$, $t \in \mathcal{T}$, where we additionally define $\mathcal{M}_0 \triangleq \mathcal{N}$.

4) Client Sampling based on Diversity: Similar to [7], [8], [24], we perform client selection based on the recent model updates received from the clients. Following [7], we define the *diversity* of a subset of clients as how well can the average of their model updates approximate the average model update from the full participation scenario. To explain this term properly, for all $k \in \mathcal{K}$, $t \in \mathcal{T}$, we define the gradient approximation error $\epsilon_k(\mathcal{M}_t)$ as

$$\epsilon_k(\mathcal{M}_t) \triangleq \left\| \sum_{n \in \mathcal{N}} p_n \nabla F_n(\mathbf{w}_k^n) - \frac{1}{|\mathcal{M}_t|} \sum_{n \in \mathcal{M}_t} \nabla F_n(\mathbf{w}_k^n) \right\|. \tag{9}$$

The gradient approximation error is small when the mean $\frac{1}{|\mathcal{M}_t|} \sum_{n \in \mathcal{M}_t} \nabla F_n(\mathbf{w}_k^n)$ of the sampled clients is close to the weighted mean $\sum_{n \in \mathcal{N}} p_n \nabla F_n(\mathbf{w}_k^n)$ of all clients. Note that $\epsilon_k(\mathcal{M}_t)$ is generally unknown in practical FL systems, since $\nabla F_n(\mathbf{w}_k^n)$ represents the true gradient of $F_n(\cdot)$ and is not computationally feasible to be obtained for neural networks we consider.

Suppose we consider the ideal case where $\nabla F_n(\mathbf{w}_{tE}^n)$, $n \in \mathcal{N}$, is known at the beginning of each communication round $t \in \mathcal{T}$. Then, a subset of scheduled clients $\mathcal{M}_t$ is considered to be more diverse, when its corresponding *gradient approximation error* $\epsilon_{tE}(\mathcal{M}_t)$ is smaller.

When diversity-based FL is adopted, in each communication round t, given the scheduling decision $\mathcal{M}_t$, the PS updates the global model according to

$$\mathbf{w}_{tE} \leftarrow \mathbf{w}_{(t-1)E} - \frac{1}{|\mathcal{M}_t|} \sum_{n \in \mathcal{M}_t} \mathbf{q}_t^n, \quad t \in \mathcal{T}. \tag{10}$$

Suppose $\nabla F_n(\mathbf{w}_{tE}^n)$ is known, based on Lemma 1 in the Appendix, it can be used to estimate $\mathbf{q}_t^n$, $n \in \mathcal{N}$, $t \in \mathcal{T}$. We can further show that the difference between the update equation in (10) and the full client participation scenario (*i.e.*, by setting $\mathcal{M}_t = \mathcal{N}$) in (7) is bounded. The bound on this difference is smaller when the gradient approximation error $\epsilon_{tE}(\mathcal{M}_t)$ is smaller.

However, since $\nabla F_n(\mathbf{w}_{tE}^n)$ is not known at the beginning of communication round $t \in \mathcal{T}$ in general, scheduling decisions based on $\epsilon_{tE}(\mathcal{M}_t)$ cannot be implemented in practical systems. In [7], the authors proposed to use the most recent model update received from each client n, which is referred to as the *representative gradient*, to estimate $\nabla F_n(\mathbf{w}_{tE}^n)$, $n \in \mathcal{N}$, $t \in \mathcal{T}$. In practice, the PS can store the most recent model update received from each client in a lookup table. Let $\hat{\mathbf{q}}_t^n \in \mathbb{R}^{|\theta|}$ denote the representative gradient from client n in communication round t. Before the first communication round, we initialize $\hat{\mathbf{q}}_0^n = \mathbf{0}$, for all $n \in \mathcal{N}$. In communication round $t+1$, we have

$$\hat{\mathbf{q}}_{t+1}^n = \mathbb{1}(u_t^n = 1)\mathbf{q}_t^n + \mathbb{1}(u_t^n = 0)\hat{\mathbf{q}}_t^n, \quad n \in \mathcal{N}. \tag{11}$$

As a result, we have $\hat{\mathbf{q}}_t^n = \mathbf{q}_{t-a_t^n}^n$, $n \in \mathcal{N}$, $t \in \mathcal{T}$. Based on $\hat{\mathbf{q}}_t^n$, we define an estimated version of $\epsilon_{tE}(\mathcal{M}_t)$ as follows

$$\hat{\epsilon}_t(\mathcal{M}_t) \triangleq \left\| \sum_{n \in \mathcal{N}} p_n \hat{\mathbf{q}}_t^n - \frac{1}{|\mathcal{M}_t|} \sum_{n \in \mathcal{M}_t} \hat{\mathbf{q}}_t^n \right\|, \quad t \in \mathcal{T}. \tag{12}$$

In Lemma 3 in the Appendix, we show that an upper bound on the gap between the expected norm of a scaled version of $\hat{\epsilon}_t(\mathcal{M}_t)$ and $\epsilon_{tE}(\mathcal{M}_t)$ depends on the largest AoI of the clients' representative gradients stored at the PS. In Theorem 3 in the Appendix, we further show that the largest AoI of the clients' representative gradients has an impact on the upper bound on the convergence rate of FL.

5) Diversity Cost: To simplify the notation, we define the representative gradient matrix in communication round $t \in \mathcal{T}$ as

$$\mathbf{G}_t = \left[\hat{\mathbf{q}}_t^1 \cdots \hat{\mathbf{q}}_t^N\right] \in \mathbb{R}^{|\theta| \times N}. \tag{13}$$

Let us define vector $\mathbf{p} = (p_1, \ldots, p_N) \in \mathbb{R}^N$. Let vector $\mathbf{g}_t \in \mathcal{G} \triangleq \mathbb{R}^{|\theta|N}$ denote the concatenation of the columns of matrix $\mathbf{G}_t$. In communication round t, given $\mathbf{g}_t$ and $\mathbf{u}_t$, we define the diversity cost as the gradient approximation error

$$\epsilon_t^c(\mathbf{g}_t, \mathbf{u}_t) = \hat{\epsilon}_t(\mathcal{M}_t) = \left\| \mathbf{G}_t \left(\mathbf{p} - \frac{\mathbf{u}_t}{\sum_{n \in \mathcal{N}} u_t^n} \right) \right\|, \quad t \in \mathcal{T}. \tag{14}$$

6) Duration of Uplink Transmission in Each Communication Round: In FL, after local training in communication round $t \in \mathcal{T}$, each scheduled client $n \in \mathcal{M}_t$ needs to send its updated model parameters to the PS. Similar to [4], in this paper, we consider time-division multiple access (TDMA), where the scheduled clients perform uplink transmission sequentially using a fixed transmit power. The time it takes for client n to send its updated parameters successfully to the PS depends on the CSI between client n and the PS. In communication round t, let $h_t^n \in \mathbb{C}$ denote the instantaneous CSI between client n and the PS. Given the system bandwidth W and transmit power P_n of client n, its instantaneous transmission rate in communication round t can be expressed as

$$r(h_t^n) = W \log_2 \left(1 + \frac{|h_t^n|^2 P_n}{\sigma_{\text{noise},n}^2} \right), \quad n \in \mathcal{N}, t \in \mathcal{T}, \tag{15}$$

where $\sigma_{\text{noise},n}^2$ denotes the noise variance of client n. We discretize the possible values of h_t^n into a finite set $\mathcal{H}_n = \{h^{n,1}, \ldots, h^{n,\max}\}$, $n \in \mathcal{N}$, and define the CSI vector $\mathbf{h}_t = (h_t^1, \ldots, h_t^N)$, $t \in \mathcal{T}$, where $h_t^n \in \mathcal{H}_n$. Let ζ denote the packet size (in bits) required to transmit the parameters.

In an FL system, the updates from all clients have the same size ζ. The time it takes for scheduled client n to send its model update to the PS is given by

$$y(h_t^n) = \frac{\zeta}{r(h_t^n)}, \quad n \in \mathcal{N}, t \in \mathcal{T}. \tag{16}$$

B. Problem Formulation

In this subsection, we formulate the channel-aware joint AoI and diversity-based client scheduling problem as a finite-horizon CMDP. We first introduce the set of decision epochs, actions, states, state transition probability, cost function, objective function, and constraints of the CMDP problem.

1) Decision Epochs and Actions: We consider a finite-horizon CMDP, where each decision epoch corresponds to a communication round. We use the set of communication rounds $\mathcal{T} = \{1, \ldots, T\}$ as the set of decision epochs of the CMDP. In decision epoch $t \in \mathcal{T}$, the action vector corresponds to a feasible scheduling decision for all N clients. That is, the action vector is $\mathbf{u}_t \in \mathcal{U}$, where the feasible action set $\mathcal{U}$ is defined in constraint C2.

2) States: To obtain a finite state space, we set an upper limit $A_{\max}$ for the AoI. That is, we set $a_t^n = A_{\max}$ for any client n that has not updated its model for more than $A_{\max}$ communication rounds. We have $a_t^n \in \mathcal{A} = \{1, 2, \ldots, A_{\max}\}$. Let $\mathbf{a}_t = (a_t^1, \ldots, a_t^N)$ denote the AoI vector in decision epoch t. The state vector in decision epoch t can be represented as

$$\mathbf{s}_t = (\mathbf{a}_t, \mathbf{h}_t, \mathbf{g}_t) \in \mathcal{S} \triangleq \left[\prod_{n \in \mathcal{N}} (\mathcal{A} \times \mathcal{H}_n)\right] \times \mathcal{G}, \quad t \in \mathcal{T}.$$

3) State Transition Probability: Given a_t^n and u_t^n, the AoI of client n in the next decision epoch, a_{t+1}^n, is a deterministic value, which is given by

$$\mathbb{P}(a_{t+1}^n \mid a_t^n, u_t^n) = \begin{cases} 1, & \text{if } a_{t+1}^n = 1 \text{ and } u_t^n = 1, \\ 1, & \text{if } a_{t+1}^n = \min(A_{\max}, a_t^n + 1) \\ & \quad \text{and } u_t^n = 0, \\ 0, & \text{otherwise,} \end{cases}$$

where $n \in \mathcal{N}$, $t \in \mathcal{T}$. The first line corresponds to the case when client n is selected to participate in training in decision epoch t. The second line accounts for the alternative case. We consider a wireless channel that evolves according to a stochastic random process. At the beginning of each decision epoch t, the CSI of client n, h_t^n, is revealed to the decision-making module. We consider the case where $h_t^n \in \mathcal{H}_n$ is distributed according to probability $\mathbb{P}(h_t^n)$, $n \in \mathcal{N}$, $t \in \mathcal{T}$. For the representative gradient matrix, we make a simplifying assumption that it follows an unknown probability $\mathbb{P}(\mathbf{g}_{t+1})$, which depends on the compositions in the local datasets. The state transition probability function can be expressed as

$$\mathbb{P}(\mathbf{s}_{t+1} \mid \mathbf{s}_t, \mathbf{u}_t) = \prod_{n=1}^{N} \left[\mathbb{P}(a_{t+1}^n \mid a_t^n, u_t^n)\mathbb{P}(h_{t+1}^n)\right] \times \mathbb{P}(\mathbf{g}_{t+1}), \quad t \in \{0, 1, \ldots, T-1\}.$$

4) Client Scheduling Policy: A client scheduling policy π is defined as a mapping from state space $\mathcal{S}$ and the set of decision epochs $\mathcal{T}$ to action space $\mathcal{U}$. Let $(\mathbf{s}_1^\pi, \ldots, \mathbf{s}_T^\pi)$ denote the state evolution under policy π, where $\mathbf{s}_t^\pi = (\mathbf{a}_t^\pi, \mathbf{h}_t, \mathbf{g}_t^\pi)$, $t \in \mathcal{T}$. Let $(\mathbf{u}_1^\pi, \ldots, \mathbf{u}_T^\pi)$ denote the action taken under policy π, where $\mathbf{u}_t^\pi = (u_t^{1,\pi}, \ldots, u_t^{N,\pi})$, $t \in \mathcal{T}$. In decision epoch $t \in \mathcal{T}$, given state vector $\mathbf{s}_t \in \mathcal{S}$, the decision-making module chooses an action $\pi_t(\mathbf{s}_t) = \mathbf{u}_t^\pi$. Let Π denote the set of all deterministic policies that satisfy constraint C2. Thus, $\pi \in \Pi$ if and only if $\mathbf{u}_t^\pi \in \mathcal{U}$, for all $t \in \mathcal{T}$.

5) Cost Function: In decision epoch $t \in \mathcal{T}$, we use the following cost function that jointly accounts for the weighted aggregate AoI, the uplink transmission time, and the diversity:

$$c(\mathbf{s}_t, \mathbf{u}_t) = \sum_{n=1}^{N} \left[N p_n a_t^n + \xi y(h_t^n) u_t^n\right] + \rho \epsilon_t^c(\mathbf{g}_t, \mathbf{u}_t), \tag{17}$$

where ξ and ρ are non-negative weight coefficients[2]. In (17), the weight coefficient Np_n places a higher weight on clients with more data samples, which encourages these clients to be scheduled more frequently, in order to reduce their AoI. Note that in the case when all the clients have the same amount of data, *i.e.*, $p_n = \frac{1}{N}$, $n \in \mathcal{N}$, we have Np_n equals to one for all clients.

6) CMDP Problem Formulation: The optimal policy π^* is defined as the policy that minimizes the expected total cost. The CMDP can be formulated as follows

$$\begin{aligned} \underset{\pi}{\text{minimize}} \quad & \sum_{t=1}^{T} \mathbb{E}\left[c(\mathbf{s}_t^\pi, \mathbf{u}_t^\pi)\right] \\ \text{subject to} \quad & \text{C2a: } \mathbf{u}_t^\pi \in \mathcal{U}, \quad t \in \mathcal{T}. \end{aligned} \tag{18}$$

The objective function in (18) favors client scheduling policies that simultaneously lead to lower weighted aggregate AoI of all clients, shorter uplink transmission duration, and smaller gradient estimation error. In this way, our channel-aware joint AoI and diversity-based client scheduling problem formulation generalizes some of the previous works and combines the advantages of several types of FL client scheduling approaches in the literature, including those which are based on CSI [4]–[6], AoI [10], [15], and diversity [7], [8]. Problem (18) is a finite-horizon CMDP. Let $V_t^*(\mathbf{s}_t)$ denote the cost-to-go function of the optimal policy, which corresponds to the minimum expected total cost obtained between decision epochs t and T, given that the current system state is $\mathbf{s}_t \in \mathcal{S}$, $t \in \mathcal{T}$. By assuming that $\mathbb{P}(\mathbf{g}_t)$ follows a discrete probability distribution that satisfies the Markov property, $t \in \mathcal{T}$, $V_t^*(\mathbf{s}_t)$ and the optimal solution (for the assumed distribution of $\mathbf{g}_t$) can be found by solving the Bellman equation iteratively using value iteration [23]. Note that this assumption is not required in Sections III and V. However, the value iteration approach cannot be applied to problems with a large number of clients, due to its high computational complexity $\mathcal{O}(T|\mathcal{S}||\mathcal{U}|^2) = \mathcal{O}(T(H_{\max}|\mathcal{A}|)^N)$ [25], where $H_{\max} = \max_{n \in \mathcal{N}} |\mathcal{H}_n|$.

[2]Here, we use a linear combination of AoI, CSI and diversity terms due to its simplicity. In practice, nonlinear cost functions can also be considered based on the system requirement.

III. Diversity-agnostic Problem and the Lagrangian Index Algorithm

As discussed in Section I, directly solving the CMDP (18) with the diversity term is a hard problem. This is partially due to the fact that the diversity part of the cost function cannot be decomposed into individual cost functions for each client. To tackle this issue, in this paper, we propose a two-stage low-complexity algorithm to solve problem (18).

In this section, we focus on the first stage of the solution approach, and consider the special case when parameter ρ in (17) is equal to zero. In this case, the cost function in (17) only depends on the AoI, the uplink transmission duration of the scheduled clients, and the scheduling decision vector. We refer to this problem as the *diversity-agnostic* client scheduling problem, and propose a *Lagrangian index*-based algorithm to solve this problem. In Section IV, we will use the Lagrangian index-based policy developed in this section as a base policy to solve the diversity-based client scheduling problem through one-step lookahead policy improvement [23].

A. Diversity-agnostic Client Scheduling Problem Formulation

By only considering the AoI and uplink transmission duration, the cost function of the diversity-agnostic client scheduling problem becomes

$$c^{\mathrm{DA}}(\mathbf{s}_t, \mathbf{u}_t) = \sum_{n=1}^{N} \left[N p_n a_t^n + \xi y(h_t^n) u_t^n \right], \quad t \in \mathcal{T}. \tag{19}$$

The diversity-agnostic client scheduling problem can be formulated as the following CMDP

$$\begin{aligned} \underset{\pi}{\text{minimize}} \quad & \sum_{t=1}^{T} \mathbb{E}\left[c^{\mathrm{DA}}(\mathbf{s}_t^{\pi}, \pi(\mathbf{s}_t^{\pi})) \right] \\ \text{subject to} \quad & \text{constraint C2a.} \end{aligned} \tag{20}$$

Since diversity is not considered in this section, we define the local state of client $n \in \mathcal{N}$ in decision epoch $t \in \mathcal{T}$ as

$$\mathbf{s}_t^n = (a_t^n, h_t^n) \in \mathcal{S}_n \triangleq \mathcal{A} \times \mathcal{H}_n. \tag{21}$$

Given a policy π, let $\mathbf{s}_t^{n,\pi}$ denote the state evolution of client $n \in \mathcal{N}$ under policy π in decision epoch $t \in \mathcal{T}$. Now, the objective function of problem (20) is separable, and the cost function related to client n in decision epoch t becomes

$$c^{n,\mathrm{DA}}(\mathbf{s}_t^n, u_t^n) = N p_n a_t^n + \xi y(h_t^n) u_t^n, \quad n \in \mathcal{N}, t \in \mathcal{T}. \tag{22}$$

B. Lagrangian Index-based Solution

It has been shown in [21] that by relaxing constraint C2a so that it holds in expectation, problem (20) can be decomposed into N client-specific CMDP problems. In this subsection, we will adopt this approach and obtain a lower bound for the solution of problem (20).

1) Relaxation of Constraint C2a: Let Φ denote the class of (possibly randomized) policies that satisfy constraint C2a in expectation. Note that Φ is different from the class of policies $\Pi \subset \Phi$ that satisfy constraint C2a in each decision epoch. Given a policy ϕ, we have $\phi \in \Phi$ if and only if

$$\text{C1a:} \ \ \mathbf{u}_t^{\phi} \in \{0,1\}^N, \quad t \in \mathcal{T}$$

$$\text{C2b:} \ \ \mathbb{E}\left[\sum_{n=1}^{N} u_t^{n,\phi} \right] \leq M, \quad t \in \mathcal{T}.$$

The expectation in constraint C2b is taken with respect to the stochasticity of state $\mathbf{s}_t^{\phi}$ under policy ϕ. After relaxing constraint C2a so that it holds in expectation, problem (20) becomes

$$\underset{\phi \in \Phi}{\text{minimize}} \quad \sum_{t=1}^{T} \mathbb{E}\left[c^{\mathrm{DA}}(\mathbf{s}_t^{\phi}, \mathbf{u}_t^{\phi}) \right]. \tag{23}$$

Since any policy that satisfies constraint C2a also satisfies constraints C1a and C2b, the optimal value of problem (23) is a lower bound of the optimal value of problem (20).

2) Optimal Solution to Problem (23) via Linear Programming: Let Φ_n denote the class of policies with scheduling decision $u_t^n = \phi_t^n(\mathbf{s}_t^n)$ for client n in state $\mathbf{s}_t^n \in \mathcal{S}_n$, and satisfy $u_t^n \in \{0,1\}$, $t \in \mathcal{T}$, $n \in \mathcal{N}$. In this way, a policy $\phi = \prod_{n \in \mathcal{N}} \phi^n \in \prod_{n \in \mathcal{N}} \Phi_n \subset \Phi$ can be constructed by combining the policies of all N clients, where $\mathbf{u}_t = \phi_t(\mathbf{s}_t) = (\phi_t^1(\mathbf{s}_t^1), \cdots, \phi_t^N(\mathbf{s}_t^N))$, $t \in \mathcal{T}$.

In the following, we adopt the linear programming approach for obtaining the optimal solution to problem (23), by first finding the optimal policies for N client-specific CMDPs.

For each client-specific CMDP $n \in \mathcal{N}$, we use the expected fraction of time that client n sojourns in state $\mathbf{s}_t^n$ and selects action u_t^n, for all $\mathbf{s}_t^n \in \mathcal{S}_n$, $u_t^n \in \mathcal{U}_n$, $n \in \mathcal{N}$, $t \in \mathcal{T}$, as the optimization variables. Given policies $\phi^n \in \Phi_n$, $n \in \mathcal{N}$, and $\phi = \prod_{n \in \mathcal{N}} \phi^n$, we define the expected sojourn time $\nu_t^{n,\phi}(\mathbf{s}_t^n, u_t^n)$ as the probability that client $n \in \mathcal{N}$ is in state $\mathbf{s}_t^n$ and selects action u_t^n in decision epoch $t \in \mathcal{T}$. We have

$$\nu_t^{n,\phi}(\mathbf{s}_t^n, u_t^n) = \mathbb{E}\left[\mathbb{1}(\mathbf{s}_t^{n,\phi} = \mathbf{s}_t^n, u_t^{n,\phi} = u_t^n) \right], \tag{24}$$

where $\mathbf{s}_t^{n,\phi}$ and $u_t^{n,\phi}$ denote, respectively, client n's state evolution and action taken in decision epoch t under policy ϕ. In this way, problem (23) is equivalent to the following linear program [21]

$$\underset{\substack{\nu_t^{n,\phi}(\mathbf{s}_t^n, u_t^n), \\ \mathbf{s}_t^n \in \mathcal{S}_n, \\ u_t^n \in \mathcal{U}_n, \\ n \in \mathcal{N}, t \in \mathcal{T}}}{\text{minimize}} \sum_{t \in \mathcal{T}} \sum_{n \in \mathcal{N}} \sum_{\mathbf{s}_t^n \in \mathcal{S}_n} \sum_{u_t^n \in \mathcal{U}_n} c^{n,\mathrm{DA}}(\mathbf{s}_t^n, u_t^n) \nu_t^{n,\phi}(\mathbf{s}_t^n, u_t^n) \tag{25}$$

$$\begin{aligned} \text{subject to} \quad & \text{C2c:} \ \sum_{n \in \mathcal{N}} \sum_{\mathbf{s}_t^n \in \mathcal{S}_n} \nu_t^{n,\phi}(\mathbf{s}_t^n, 1) \leq M, \quad t \in \mathcal{T}, \\ & \text{C3:} \ \sum_{u_t^n \in \mathcal{U}_n} \nu_t^{n,\phi}(\mathbf{s}_t^n, u_t^n) = \\ & \sum_{\mathbf{s}_{t-1}^n \in \mathcal{S}_n} \sum_{u_{t-1}^n \in \mathcal{U}_n} \mathbb{P}(\mathbf{s}_t^n \mid \mathbf{s}_{t-1}^n, u_{t-1}^n) \nu_{t-1}^{n,\phi}(\mathbf{s}_{t-1}^n, u_{t-1}^n), \\ & \qquad \mathbf{s}_t^n \in \mathcal{S}_n, n \in \mathcal{N}, t \in \mathcal{T} \setminus \{1\}, \\ & \text{C4:} \ \sum_{u_1^n \in \mathcal{U}_n} \nu_1^{n,\phi}(\mathbf{s}_1^n, u_1^n) = 1, \quad \mathbf{s}_1^n \in \mathcal{S}_n, n \in \mathcal{N}, \\ & \text{C5:} \ \nu_t^{n,\phi}(\mathbf{s}_t^n, u_t^n) \geq 0, \mathbf{s}_t^n \in \mathcal{S}_n, u_t^n \in \mathcal{U}_n, n \in \mathcal{N}, t \in \mathcal{T}. \end{aligned}$$

The objective function of problem (25) corresponds to the expected total cost. The left-hand side of constraint C2c corresponds to the expected number of clients that are sched-

uled to participate in FL in decision epoch t. Thus, constraints C2b and C2c are equivalent. Constraints C3 and C4 are the flow conservation conditions that ensure the solution satisfies the state transition probability. Constraint C5 ensures that the optimization variables, which are probabilities of events, are non-negative. Problem (25) involves $2TN|\mathcal{S}_n|$ variables and has a computational complexity of $\mathcal{O}((TNH_{\max}|\mathcal{A}|)^{2.5}\log(TNH_{\max}|\mathcal{A}|/\delta))$, where δ denotes the relative accuracy of the employed solver [26].

Given the optimal solution of problem (25) $\nu_t^{n,\phi^*}(\mathbf{s}_t^n, u_t^n)$, we can design a randomized policy $\phi^* = \prod_{n\in\mathcal{N}} \phi^{n,*}$, where $\phi_t^{n,*}(\mathbf{s}_t^n)$ is a random variable with probability distribution

$$\mathbb{P}(\phi_t^{n,*}(\mathbf{s}_t^n) = u_t^n) = \begin{cases} \frac{\nu_t^{n,\phi^*}(\mathbf{s}_t^n,u_t^n)}{\nu_t^{n,\phi^*}(\mathbf{s}_t^n,1)\ +\ \nu_t^{n,\phi^*}(\mathbf{s}_t^n,0)}, \text{if } u_t^n \in \mathcal{U}_n \\ \qquad\qquad \text{and } \mathbf{s}_t^n \in \mathcal{S}_{t,\text{visited}}^n, \\ \frac{1}{2}, \text{if } u_t^n \in \mathcal{U}_n \text{ and } \mathbf{s}_t^n \in \mathcal{S}_n \setminus \mathcal{S}_{t,\text{visited}}^n, \\ 0, \quad \text{otherwise.} \end{cases}$$

The set $\mathcal{S}_{t,\text{visited}}^n = \left\{\mathbf{s}_t^n \in \mathcal{S}_n \mid \sum_{u_t^n\in\mathcal{U}_n} \nu_t^{n,\phi^*}(\mathbf{s}_t^n, u_t^n) > 0\right\}$ represents the states that are likely being visited under the randomized policy ϕ^*.

Theorem 1. The client scheduling policy ϕ^* is the optimal solution to problem (23).

Proof: Problem (23) falls into the class of problems considered in [21]. The proof of Theorem 1 can be found in [21, Sections EC 1.1−1.3]. ∎

3) Lagrangian Index-based Feasible Solution: Although policy ϕ^* achieves the optimal solution to problem (23), it may not be a feasible solution to problem (20) since constraint C2a may not always be satisfied. In [21], the authors proposed a low-complexity asymptotically-optimal feasible heuristic solution to problem (20) called the *Lagrangian index* scheduling policy. In this subsection, we introduce the steps for finding the Lagrangian index by first solving the dual problem of problem (23).

Let $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_T) \succeq \mathbf{0}$ denote the vector of Lagrange multipliers corresponding to constraint C2b. Then, the dual problem of problem (23) can be expressed as follows

$$\underset{\boldsymbol{\lambda}\succeq\mathbf{0}}{\text{maximize}} \ \min_{\psi\in\Phi'} \sum_{t=1}^{T} \Bigg\{ \mathbb{E}\left[c^{\text{DA}}(\mathbf{s}_t^{\psi}, \mathbf{u}_t^{\psi})\right] + \lambda_t \left(\sum_{n=1}^{N} \mathbb{E}[u_t^{n,\psi}] - M\right) \Bigg\}, \tag{26}$$

where Φ' denotes the class of policies that satisfy constraint C1a. Given the optimal Lagrange multiplier vector $\boldsymbol{\lambda}^* = (\lambda_1^*, \ldots, \lambda_T^*)$, the optimal value of problem (26) can be found recursively by backward induction [21][3]. First, we initialize $L_{T+1}^{\text{DA},\boldsymbol{\lambda}^*}(\mathbf{s}_{T+1}) = 0$, for $\mathbf{s}_{T+1} \in \mathcal{S}$, and then recursively compute

$$L_t^{\text{DA},\boldsymbol{\lambda}^*}(\mathbf{s}_t) = \min_{\mathbf{u}_t\in\mathcal{U}} \Bigg\{ c^{\text{DA}}(\mathbf{s}_t, \mathbf{u}_t) + \mathbb{E}\left[L_{t+1}^{\text{DA},\boldsymbol{\lambda}^*}(\mathbf{s}_{t+1}) \mid \mathbf{s}_t, \mathbf{u}_t\right] + \lambda_t^* \left(\sum_{n=1}^{N} u_t^n - M\right) \Bigg\},$$

for all $t \in \mathcal{T}$, $\mathbf{s}_t \in \mathcal{S}$, where we use the following short-hand notation

$$\begin{aligned} &\mathbb{E}\left[f(\mathbf{s}_{t+1}) \mid \mathbf{s}, \mathbf{u}\right] \\ &\triangleq \sum_{\mathbf{s}_+\in\mathcal{S}} f(\mathbf{s}_+)\mathbb{P}(\mathbf{s}_{t+1} = \mathbf{s}_+ \mid \mathbf{s}_t = \mathbf{s}, \mathbf{u}_t = \mathbf{u}), \end{aligned} \tag{27}$$

and $f(\cdot)$ is any function of $\mathbf{s}_{t+1}$. In this way, given the optimal Lagrange multiplier vector $\boldsymbol{\lambda}^*$ and the initial state $\mathbf{s}_1$, $L_1^{\text{DA},\boldsymbol{\lambda}^*}(\mathbf{s}_1)$ is the optimal value of problem (26). Furthermore, given $\boldsymbol{\lambda}^*$, $L_t^{\text{DA},\boldsymbol{\lambda}^*}(\mathbf{s})$ can be expressed as follows

$$L_t^{\text{DA},\boldsymbol{\lambda}^*}(\mathbf{s}_t) = -\sum_{i=t}^{T} \lambda_i^* M + \sum_{n=1}^{N} V_t^{n,\text{DA},\boldsymbol{\lambda}^*}(\mathbf{s}_t^n), \tag{28}$$

where $\mathbf{s}_t \in \mathcal{S}$, $t \in \mathcal{T}$, and $V_t^{n,\text{DA},\boldsymbol{\lambda}^*}(\mathbf{s}_t^n)$ denotes the value function of the client-specific MDP. It can similarly be derived using backward induction, by first initializing $V_{T+1}^{n,\text{DA},\boldsymbol{\lambda}^*}(\mathbf{s}_{T+1}^n) = 0$, for $\mathbf{s}_{T+1}^n \in \mathcal{S}_n$, $n \in \mathcal{N}$. Then, $V_t^{n,\text{DA},\boldsymbol{\lambda}^*}(\mathbf{s}_t^n)$, $\mathbf{s}_t^n \in \mathcal{S}_n$, are found iteratively for decision epoch $t = T, T-1, \ldots, 1$, by calculating

$$V_t^{n,\text{DA},\boldsymbol{\lambda}^*}(\mathbf{s}_t^n) = \min_{u_t^n\in\mathcal{U}_n} \Bigg\{ c^{n,\text{DA}}(\mathbf{s}_t^n, u_t^n) + \mathbb{E}\left[V_{t+1}^{n,\text{DA},\boldsymbol{\lambda}^*}(\mathbf{s}_{t+1}^n) \mid \mathbf{s}_t^n, u_t^n\right] + \lambda_t^* u_t^n \Bigg\}. \tag{29}$$

In decision epoch $t \in \mathcal{T}$, given client $n \in \mathcal{N}$ is in state $\mathbf{s}_t^n \in \mathcal{S}_n$, the optimal scheduling policy, $\psi^{n,*} \in \Phi_n$, can be expressed as follows

$$\psi_t^{n,*}(\mathbf{s}_t^n) = \arg\min_{u_t^n\in\mathcal{U}_n} \Bigg\{ c^{n,\text{DA}}(\mathbf{s}_t^n, u_t^n) + \mathbb{E}\left[V_{t+1}^{n,\text{DA},\boldsymbol{\lambda}^*}(\mathbf{s}_{t+1}^n) \mid \mathbf{s}_t^n, u_t^n\right] + \lambda_t^* u_t^n \Bigg\}, \quad \mathbf{s}_t^n \in \mathcal{S}_n, t \in \mathcal{T}.$$

To handle ties, we assume that a deterministic rule is designed, *e.g.,* assign $u_t^n = 0$ whenever a tie happens. In this way, the policy $\psi^* = \prod_{n\in\mathcal{N}} \psi^{n,*}$ that combines the deterministic policies of all N clients can be used to determine the optimal value of the dual problem (26). The value functions of the client-specific MDPs can be utilized to derive the following Lagrangian index for each client.

Definition 1 (Lagrangian Index)**.** In decision epoch $t \in \mathcal{T}$, the Lagrangian index for client $n \in \mathcal{N}$ in state $\mathbf{s}_t^n \in \mathcal{S}_n$ is defined as

$$i_t^{n,\text{DA}}(\mathbf{s}_t^n) \triangleq \left(c^{n,\text{DA}}(\mathbf{s}_t^n, 1) + \mathbb{E}\left[V_{t+1}^{n,\text{DA},\boldsymbol{\lambda}^*}(\mathbf{s}_{t+1}^n) \mid \mathbf{s}_t^n, 1\right]\right) - \left(c^{n,\text{DA}}(\mathbf{s}_t^n, 0) + \mathbb{E}\left[V_{t+1}^{n,\text{DA},\boldsymbol{\lambda}^*}(\mathbf{s}_{t+1}^n) \mid \mathbf{s}_t^n, 0\right]\right).$$

[3]Since constraints C2b and C2c are equivalent, $\boldsymbol{\lambda}^*$ can be found by solving the dual problem of problem (25), for the dual variables associated with constraint C2c.

In decision epoch t, we define $\mathbf{i}_t^{\mathrm{DA}}(\mathbf{s}_t) \triangleq (i_t^{1,\mathrm{DA}}(\mathbf{s}_t^1), \ldots, i_t^{N,\mathrm{DA}}(\mathbf{s}_t^N))$ as the vector of Lagrangian indices of all clients. We now introduce the Lagrangian index scheduling policy.

Definition 2 (Lagrangian Index Scheduling)**.** Under the Lagrangian index scheduling policy, in decision epoch $t \in \mathcal{T}$, those M clients with the smallest non-positive Lagrangian indices are scheduled to participate in FL. If multiple clients have the same Lagrangian indices, then policy ϕ^* is utilized to break the tie. Let $\tilde{\pi} \in \Pi$ denote the Lagrangian index policy. In each decision epoch, we have

$$\tilde{\pi}_t(\mathbf{s}_t) = \arg\min_{\mathbf{u}_t \in \mathcal{U}} \mathbf{u}_t^{\mathsf{T}} \mathbf{i}_t^{\mathrm{DA}}(\mathbf{s}_t). \tag{30}$$

Since constraint C2a is satisfied, the optimal solution of problem (30) is a feasible solution to CMDP problem (18).

We use $\tilde{V}_t^{\mathrm{DA}}(\mathbf{s}_t)$ to denote the cost-to-go function of state $\mathbf{s}_t$ under policy $\tilde{\pi}$. By definition, $\tilde{V}_t^{\mathrm{DA}}(\mathbf{s}_t)$ can be found using policy iteration by first initializing $\tilde{V}_{T+1}^{\mathrm{DA}}(\mathbf{s}_{T+1}) = 0$, $\mathbf{s}_{T+1} \in \mathcal{S}$, and then iteratively calculating for $t = T, T-1, \ldots, 1$, for all $\mathbf{s}_t \in \mathcal{S}$ using the following equation

$$\tilde{V}_t^{\mathrm{DA}}(\mathbf{s}_t) = c^{\mathrm{DA}}(\mathbf{s}_t, \tilde{\pi}_t(\mathbf{s}_t)) + \mathbb{E}\left[\tilde{V}_{t+1}^{\mathrm{DA}}(\mathbf{s}_{t+1}) \mid \mathbf{s}_t, \tilde{\pi}_t(\mathbf{s}_t)\right].$$

However, the computational complexity of this approach becomes high as N increases. In Section IV, we will introduce a method to estimate $\tilde{V}_t^{\mathrm{DA}}(\mathbf{s}_t)$. Let $V_t^{\mathrm{DA},*}(\mathbf{s}_t)$ denote the optimal cost-to-go function of state $\mathbf{s}_t$ of the diversity-agnostic client scheduling problem obtained by the value iteration algorithm [23]. Then, it can be proven that by increasing the number of clients in the FL system, while a fixed percentage of clients are scheduled in each decision epoch, the Lagrangian index-based scheduling policy achieves asymptotically optimal performance. We state this property in the following theorem.

Theorem 2. Consider a sequence of the diversity-agnostic client scheduling problems with N clients (indexed by N), and a fixed portion of zN clients ($0 < z < 1$) are scheduled for participating in FL. Let $V_1^{\mathrm{DA},*}(\mathbf{s}_1; N)$, $L_1^{\mathrm{DA},\boldsymbol{\lambda}^*}(\mathbf{s}_1; N)$, and $\tilde{V}_1^{DA}(\mathbf{s}_1; N)$ denote the optimal value function, the optimal Lagrangian, and the value function using the Lagrangian index-based policy $\tilde{\pi}$, respectively. We have

$$\lim_{N\to\infty} \frac{\tilde{V}_1^{\mathrm{DA}}(\mathbf{s}_1; N)}{V_1^{\mathrm{DA},*}(\mathbf{s}_1; N)} = 1 \quad \text{and} \quad \lim_{N\to\infty} \frac{L_1^{\mathrm{DA},\boldsymbol{\lambda}^*}(\mathbf{s}_1; N)}{V_1^{\mathrm{DA},*}(\mathbf{s}_1; N)} = 1.$$

Proof: Problem (20) falls into the class of problems considered in [21]. The detailed proof of Theorem 2 follows similar steps as the proof provided in [21, Section EC 3.1]. ■

C. Infinite-horizon Approximation of the Lagrangian Index

Some of the FL tasks may have a long time-horizon T, which will increase the dimensionality of problem (25), making it more computationally complex to solve. In this subsection, we propose a *stationary Lagrangian index scheduling* method, which has a lower complexity for problems with long time-horizon. It can be interpreted as an infinite-horizon approximation of the original Lagrangian index scheduling approach. Given a stationary policy $\bar{\phi}$, let us define the probability that client $n \in \mathcal{N}$ is in state $\mathbf{s}^n$ and takes action u^n in any decision epoch $t \in \mathcal{T}$ as

$$\bar{\nu}^{n,\bar{\phi}}(\mathbf{s}^n, u^n) = \mathbb{E}\left[\mathbb{1}(\mathbf{s}_t^{n,\bar{\phi}} = \mathbf{s}^n, u_t^{n,\bar{\phi}} = u^n)\right]. \tag{31}$$

In this way, the equivalent linear program for problem (25) can be formulated as follows

$$\begin{aligned}
&\underset{\substack{\bar{\nu}^{n,\bar{\phi}}(\mathbf{s}^n,u^n),\\ \mathbf{s}^n\in\mathcal{S}_n, u^n\in\mathcal{U}_n,\\ n\in\mathcal{N}}}{\text{minimize}} \sum_{n\in\mathcal{N}} \sum_{\mathbf{s}^n\in\mathcal{S}_n} \sum_{u^n\in\mathcal{U}_n} c^{n,\mathrm{DA}}(\mathbf{s}^n, u^n)\bar{\nu}^{n,\bar{\phi}}(\mathbf{s}^n, u^n) \\
&\text{subject to} \quad \text{C2d:} \sum_{n\in\mathcal{N}} \sum_{\mathbf{s}^n\in\mathcal{S}_n} \bar{\nu}^{n,\bar{\phi}}(\mathbf{s}^n, 1) \le M, \\
&\quad \text{C3a:} \sum_{u^n\in\mathcal{U}_n} \bar{\nu}^{n,\bar{\phi}}(\mathbf{s}^n, u^n) = \\
&\quad \sum_{\mathbf{s}_{t-1}^n\in\mathcal{S}_n} \sum_{u_{t-1}^n\in\mathcal{U}_n} \mathbb{P}(\mathbf{s}^n \mid \mathbf{s}_{t-1}^n, u_{t-1}^n)\bar{\nu}^{n,\bar{\phi}}(\mathbf{s}_{t-1}^n, u_{t-1}^n), \\
&\qquad\qquad \mathbf{s}^n \in \mathcal{S}_n, n \in \mathcal{N}, \\
&\quad \text{C4a:} \sum_{\mathbf{s}^n\in\mathcal{S}_n} \sum_{u^n\in\mathcal{U}_n} \bar{\nu}^{n,\bar{\phi}}(\mathbf{s}^n, u^n) = 1, \quad n \in \mathcal{N}, \\
&\quad \text{C5a:} \ \bar{\nu}^{n,\bar{\phi}}(\mathbf{s}^n, u^n) \ge 0, \quad \mathbf{s}^n \in \mathcal{S}_n, u^n \in \mathcal{U}_n, n \in \mathcal{N}.
\end{aligned} \tag{32}$$

Let λ_{inf}^* denote the optimal Lagrange multiplier related to constraint C2d, which can be found by solving the dual problem of problem (32). Problem (32) has $2N|\mathcal{S}_n|$ variables and has a computational complexity of $\mathcal{O}((NH_{\max}|\mathcal{A}|)^{2.5}\log(NH_{\max}|\mathcal{A}|/\delta))$ [26].

Let $V^{n,\mathrm{DA},\lambda_{\mathrm{inf}}^*}(\mathbf{s}^n)$ denote the infinite-horizon average cost value function of state $\mathbf{s}^n$ of the client-specific MDP, given λ_{inf}^*. $V^{n,\mathrm{DA},\lambda_{\mathrm{inf}}^*}(\mathbf{s}^n)$ can be obtained using the relative value iteration algorithm (RVIA) [27, Proposition 5.3.2], by first initializing $V_0^{n,\mathrm{DA},\lambda_{\mathrm{inf}}^*}(\mathbf{s}^n) = 0$, $\mathbf{s}^n \in \mathcal{S}_n, n \in \mathcal{N}$, and then iteratively calculating, for all $j = 1, 2, \ldots$

$$\begin{aligned}
V_j^{n,\mathrm{DA},\lambda_{\mathrm{inf}}^*}(\mathbf{s}^n) = \min_{u^n\in\mathcal{U}_n} \Big\{ & c^{n,\mathrm{DA}}(\mathbf{s}^n, u^n) \\
& + \mathbb{E}\left[V_{j-1}^{n,\mathrm{DA},\lambda_{\mathrm{inf}}^*}(\mathbf{s}_{\mathrm{next}}^n) \mid \mathbf{s}^n, u^n\right] + \lambda_{\mathrm{inf}}^* u^n - V_j^{n,\mathrm{DA},\lambda_{\mathrm{inf}}^*}(\mathbf{s}_{\mathrm{ref}}^n)\Big\},
\end{aligned} \tag{33}$$

for all $\mathbf{s}^n \in \mathcal{S}_n$, $n \in \mathcal{N}$, where $\mathbf{s}_{\mathrm{ref}}^n$ is a fixed reference state. Any state $\mathbf{s}_{\mathrm{ref}}^n \in \mathcal{S}_n$ can be chosen as the reference state, but it needs to be fixed across the RVIA iterations. At the beginning of each iteration j, $V_j^{n,\mathrm{DA},\lambda_{\mathrm{inf}}^*}(\mathbf{s}_{\mathrm{ref}}^n)$ is first calculated by omitting the last term on the right-hand side of (33). Since RVIA is guaranteed to converge [27], we let $V^{n,\mathrm{DA},\lambda_{\mathrm{inf}}^*}(\mathbf{s}^n) = \lim_{j\to\infty} V_j^{n,\mathrm{DA},\lambda_{\mathrm{inf}}^*}(\mathbf{s}^n)$. We subsequently define the stationary Lagrangian index in Definition 3 and its corresponding scheduling policy in Definition 4.

Definition 3 (Stationary Lagrangian Index)**.** The stationary Lagrangian index for client $n \in \mathcal{N}$ in state $\mathbf{s}^n \in \mathcal{S}_n$ is defined as

$$\begin{aligned}
i_{\mathrm{inf}}^{n,\mathrm{DA}}(\mathbf{s}^n) \triangleq & \left(c^{n,\mathrm{DA}}(\mathbf{s}^n, 1) + \mathbb{E}\left[V^{n,\mathrm{DA},\lambda_{\mathrm{inf}}^*}(\mathbf{s}_{\mathrm{next}}^n) \mid \mathbf{s}^n, 1\right]\right) \\
& - \left(c^{n,\mathrm{DA}}(\mathbf{s}^n, 0) + \mathbb{E}\left[V^{n,\mathrm{DA},\lambda_{\mathrm{inf}}^*}(\mathbf{s}_{\mathrm{next}}^n) \mid \mathbf{s}^n, 0\right]\right).
\end{aligned}$$

Definition 4 (Stationary Lagrangian Index Scheduling)**.** Under the stationary Lagrangian index scheduling policy, in each decision epoch, given the vector of stationary Lagrangian indices $\mathbf{i}_{\text{inf}}^{\text{DA}}(\mathbf{s}) \triangleq (i_{\text{inf}}^{1,\text{DA}}(\mathbf{s}^1), \ldots, i_{\text{inf}}^{N,\text{DA}}(\mathbf{s}^N))$, the M clients with the smallest non-positive stationary Lagrangian indices are scheduled to participate in FL. Let $\tilde{\pi}^{\text{inf}} \in \Pi$ denote the stationary Lagrangian index scheduling policy. Then, given current state $\mathbf{s} \in \mathcal{S}$, in each decision epoch,

$$\tilde{\pi}^{\text{inf}}(\mathbf{s}) = \underset{\mathbf{u}\in\mathcal{U}}{\arg\min}\ \mathbf{u}^{\text{T}}\mathbf{i}_{\text{inf}}^{\text{DA}}(\mathbf{s}). \tag{34}$$

Compared with problem (30), we can obtain a stationary policy for each state $\mathbf{s} \in \mathcal{S}$ from (34). This allows us to solve problems with long time-horizon since the dimension of problem (32) does not increase with T. In Section V, we will show that the infinite-horizon approximation renders similar performance as the original Lagrangian index-based approach.

IV. Diversity-based Lagrangian Index Solution

In this section, we propose a low-complexity suboptimal solution to the original diversity-based CMDP (18). We consider the case when the representative gradient matrix $\mathbf{G}_t$ is revealed at the beginning of communication round $t \in \mathcal{T}$, without making any assumption on $\mathbb{P}(\mathbf{g}_t)$, $t \in \mathcal{T}$. We use the Lagrangian index-based scheduling policy for the diversity-agnostic client scheduling problem as a base policy, and derive the diversity-based client scheduling policy based on *one-step lookahead policy improvement*.

A. One-step Lookahead Policy Improvement

In large-scale MDP problems, it is often hard to obtain the optimal policy using the value iteration algorithm, due to the large state and action spaces of the problem. One-step and multi-step lookahead approaches have been proposed to obtain low-complexity heuristic policies. In this section, we adopt the one-step lookahead policy proposed in [23, Ch. 6] to find a heuristic policy for the diversity-based client scheduling problem.

Definition 5 (One-step Lookahead Policy)**.** In the one-step lookahead scheme, starting from a base policy π^{base} with value function $V^{\text{base}}(\cdot) : \mathcal{S}\times\mathcal{T} \mapsto \mathbb{R}$, given state $\mathbf{s}_t$ in decision epoch $t \in \mathcal{T}$, the action vector $\mathbf{u}_t$ is selected based on the following optimization objective

$$\min_{\mathbf{u}_t\in\mathcal{U}} \Big\{ c(\mathbf{s}_t, \mathbf{u}_t) + \mathbb{E}\left[V_{t+1}^{\text{base}}(\mathbf{s}_{t+1}) \mid \mathbf{s}_t, \mathbf{u}_t\right] \Big\}.$$

The one-step lookahead policy $\hat{\pi} \in \Pi$ is defined as

$$\hat{\pi}_t(\mathbf{s}_t) = \underset{\mathbf{u}_t\in\mathcal{U}}{\arg\min} \Big\{ c(\mathbf{s}_t, \mathbf{u}_t) + \mathbb{E}\left[V_{t+1}^{\text{base}}(\mathbf{s}_{t+1}) \mid \mathbf{s}_t, \mathbf{u}_t\right] \Big\}.$$

In [23, Section 6.4], it was proved that the policy achieved by the one-step lookahead scheme achieves better performance compared to the base policy. Recent work has further shown that the one-step lookahead scheme typically improves the performance considerably [23]. In this paper, we use the Lagrangian index-based policy $\tilde{\pi}$ (obtained from the diversity-agnostic case) as the base policy for the diversity-based client scheduling problem and obtain an improved policy $\overline{\pi} \in \Pi$ through the one-step lookahead scheme. From the results in Theorem 2 and considering the client scheduling problem with a truncated time-horizon $\mathcal{T}' = \{t, \ldots, T\}$, we can estimate the cost-to-go functions under the Lagrangian index-based policy $\tilde{V}_1^{\text{DA}}(\mathbf{s}_t)$, for all $\mathbf{s}_t \in \mathcal{S}$, $t \in \mathcal{T}$ as follows

$$\tilde{V}_t^{\text{DA}}(\mathbf{s}_t) \approx L_t^{\text{DA},\boldsymbol{\lambda}^*}(\mathbf{s}_t) = -\sum_{i=t}^{T}\lambda_i^* M + \sum_{n=1}^{N} V_t^{n,\text{DA},\boldsymbol{\lambda}^*}(\mathbf{s}_t^n). \tag{35}$$

Under the conditions given in Theorem 2, approximation (35) is asymptotically accurate. Since the future diversity information is unknown, we do not consider the portion of the expected future cost corresponding to diversity for the cost-to-go function estimation[4]. In this way, for all $\mathbf{s}_t \in \mathcal{S}$, $t \in \mathcal{T}$, the cost-to-go function of the Lagrangian index-based policy $\tilde{\pi}$ for the diversity-based client scheduling problem can be approximated as follows

$$\tilde{V}_t(\mathbf{s}_t) \approx -\sum_{i=t}^{T}\lambda_i^* M + \sum_{n=1}^{N} V_t^{n,\text{DA},\boldsymbol{\lambda}^*}(\mathbf{s}_t^n). \tag{36}$$

B. Per-time Slot Decision Problem

Given the cost-to-go function under the Lagrangian index-based policy is approximated by (36), the one-step lookahead policy for the diversity-based client scheduling problem becomes

$$\begin{aligned}
\overline{\pi}_t(\mathbf{s}_t) = &\ \underset{\mathbf{u}_t\in\mathcal{U}}{\arg\min} \Bigg\{ c(\mathbf{s}_t, \mathbf{u}_t) - \sum_{i=t}^{T}\lambda_i^* M \\
&+ \sum_{n=1}^{N} \mathbb{E}\left[V_{t+1}^{n,\text{DA},\boldsymbol{\lambda}^*}(\mathbf{s}_{t+1}^n) \mid \mathbf{s}_t^n, u_t^n\right] \Bigg\} \\
= &\ \underset{\mathbf{u}_t\in\mathcal{U}}{\arg\min} \left\{ c^{\text{DA}}(\mathbf{s}_t, \mathbf{u}_t) + \rho \left\| \mathbf{G}_t \left(\mathbf{p} - \frac{\mathbf{u}_t}{\sum_{n\in\mathcal{N}} u_t^n} \right) \right\| \right\} \\
&+ \sum_{n=1}^{N} \mathbb{E}\left[V_{t+1}^{n,\text{DA},\boldsymbol{\lambda}^*}(\mathbf{s}_{t+1}^n) \mid \mathbf{s}_t^n, u_t^n\right] \\
\overset{\text{(a)}}{=} &\ \underset{\mathbf{u}_t\in\mathcal{U}}{\arg\min} \left[\mathbf{u}_t^{\text{T}}\mathbf{i}_t^{\text{DA}}(\mathbf{s}_t) + \rho \left\| \mathbf{G}_t \left(\mathbf{p} - \frac{\mathbf{u}_t}{\sum_{n\in\mathcal{N}} u_t^n} \right) \right\| \right].
\end{aligned} \tag{37}$$

We can obtain equality (a) in (37) as follows. Let us define $\Theta_1(\mathbf{s}_t, \mathbf{u}_t) \triangleq \mathbf{u}_t^{\text{T}}\mathbf{i}_t^{\text{DA}}(\mathbf{s}_t)$, $\Omega(\mathbf{s}_t) \triangleq \sum_{n=1}^{N}\left(c^{n,\text{DA}}(\mathbf{s}_t^n, 0) + \mathbb{E}\big[V_{t+1}^{n,\text{DA},\boldsymbol{\lambda}^*}(\mathbf{s}_{t+1}^n) \mid \mathbf{s}_t^n, 0\big]\right)$, and $\Theta_2(\mathbf{s}_t, \mathbf{u}_t) \triangleq \Omega(\mathbf{s}_t) + \Theta_1(\mathbf{s}_t, \mathbf{u}_t)$.

Based on Definition 1, we have

$$\begin{aligned}
\Theta_1(\mathbf{s}_t, \mathbf{u}_t) = &\ \sum_{n=1}^{N} u_t^n \left(c^{n,\text{DA}}(\mathbf{s}_t^n, 1) - c^{n,\text{DA}}(\mathbf{s}_t^n, 0) \right) \\
&+ \sum_{n=1}^{N} u_t^n \Bigg(\mathbb{E}\left[V_{t+1}^{n,\text{DA},\boldsymbol{\lambda}^*}(\mathbf{s}_{t+1}^n) \mid \mathbf{s}_t^n, 1\right]
\end{aligned}$$

[4]Although diversity is not considered in the cost-to-go function estimation, it is still accounted for in the one-step lookahead scheme, through cost function $c(\mathbf{s}_t, \mathbf{u}_t)$.

$$- \mathbb{E}\left[V_{t+1}^{n,\mathrm{DA},\boldsymbol{\lambda}^*}(\mathbf{s}_{t+1}^n) \mid \mathbf{s}_t^n, 0\right]\Big),$$

$$\begin{aligned}\Theta_2(\mathbf{s}_t,\mathbf{u}_t) = \;& \Theta_1(\mathbf{s}_t,\mathbf{u}_t) + \Omega(\mathbf{s}_t)\\ =\;& \sum_{n=1}^{N}\left(\mathbb{1}(u_t^n=1)c^{n,\mathrm{DA}}(\mathbf{s}_t^n,1) + \mathbb{1}(u_t^n=0)c^{n,\mathrm{DA}}(\mathbf{s}_t^n,0)\right)\\ &+\sum_{n=1}^{N}\Bigg(\mathbb{1}(u_t^n=1)\mathbb{E}\left[V_{t+1}^{n,\mathrm{DA},\boldsymbol{\lambda}^*}(\mathbf{s}_{t+1}^n) \mid \mathbf{s}_t^n, 1\right]\\ &\qquad + \mathbb{1}(u_t^n=0)\mathbb{E}\left[V_{t+1}^{n,\mathrm{DA},\boldsymbol{\lambda}^*}(\mathbf{s}_{t+1}^n) \mid \mathbf{s}_t^n, 0\right]\Bigg)\\ =\;& c^{\mathrm{DA}}(\mathbf{s}_t,\mathbf{u}_t) + \sum_{n=1}^{N}\mathbb{E}\left[V_{t+1}^{n,\mathrm{DA},\boldsymbol{\lambda}^*}(\mathbf{s}_{t+1}^n) \mid \mathbf{s}_t^n, u_t^n\right].\end{aligned}$$

Since $\Omega(\mathbf{s}_t)$ does not depend on $\mathbf{u}_t$, we have $\arg\min_{\mathbf{u}_t\in\mathcal{U}} \Theta_1(\mathbf{s}_t,\mathbf{u}_t) = \arg\min_{\mathbf{u}_t\in\mathcal{U}} \Theta_2(\mathbf{s}_t,\mathbf{u}_t)$, which is equivalent to equality (a) in (37).

When the time horizon T is large, the infinite-horizon approximation obtained in Section III-C can be adopted, and the stationary Lagrangian index vector $\mathbf{i}_{\mathrm{inf}}^{\mathrm{DA}}(\mathbf{s}_t)$ is used in place of $\mathbf{i}_t^{\mathrm{DA}}(\mathbf{s}_t)$. The optimal solution to problem (37) can be found by first introducing an equality constraint $\sum_{n=1}^{N} u_t^n = \tilde{M}_t$, and then solving the optimization problem M times with $\tilde{M}_t = 1, 2, \ldots, M$. In this way, the optimization problem becomes

$$\begin{aligned}\underset{\mathbf{u}_t\in\{0,1\}^N}{\text{minimize}} \quad & -\mathbf{u}_t^{\mathrm{T}}\mathbf{i}_t^{\mathrm{DA}}(\mathbf{s}_t) + \rho\left\|\mathbf{G}_t\left(\mathbf{p}-\mathbf{u}_t/\tilde{M}_t\right)\right\| \qquad (38)\\ \text{subject to} \quad & \sum_{n=1}^{N} u_t^n = \tilde{M}_t.\end{aligned}$$

Given a fixed $\tilde{M}_t$, problem (38) is a binary optimization problem. In practice, to obtain a low-complexity solution that can be implemented in real time, we can set $\tilde{M}_t = M$, $t \in \mathcal{T}$, relax the binary variables to continuous ones, and use successive convex approximation (SCA) [28] to obtain a suboptimal solution to problem (38). We initialize $\mathbf{u}_t^{(0)}$ to a random binary vector and employ SCA for I^{SCA} iterations. In the i-th iteration, we obtain $\mathbf{u}_t^{(i)}$ by solving the following optimization problem

$$\begin{aligned}\underset{\mathbf{u}_t^{(i)}}{\text{minimize}} \quad & -\left(\mathbf{u}_t^{(i)}\right)^{\mathrm{T}}\mathbf{i}_t^{\mathrm{DA}}(\mathbf{s}_t) + \rho\left\|\mathbf{G}_t\left(\mathbf{p}-\mathbf{u}_t^{(i)}/M\right)\right\|\\ & + 2^{i+1}\rho'\left(\mathbf{u}_t^{(i)}-\mathbf{u}_t^{(i-1)}\right)^{\mathrm{T}}\mathbf{u}_t^{(i-1)}\\ \text{subject to} \quad & \sum_{n=1}^{N} u_t^{n,(i)} = M, \qquad (39)\end{aligned}$$

where ρ' denotes a penalty factor. The values for ρ and ρ' are chosen manually before FL training. In each iteration of SCA, the optimization problem involves N continuous variables and one constraint, and has a computational complexity of $\mathcal{O}(N^3)$ [29]. The problem can be solved by using optimization solvers such as CVX [30].

To solve problem (39), the Lagrangian indices of each client in different states are first obtained during the planning stage. During the deployment stage, problem (39) is solved in an online manner, based on the representative gradient matrix revealed at the beginning of communication round t. Henceforth, we will refer to this proposed approach for solving problem (18) as the two-stage online algorithm. The key steps in the planning and deployment stages are presented in Algorithm 1.

Algorithm 1 Two-stage Online Client Scheduling Algorithm

Planning Stage:
Solve the dual problem of problem (32) to obtain λ_{inf}^*, $\forall t \in \mathcal{T}$.
Initialization: T_{iter}, $V_0^{n,\mathrm{DA},\lambda_{\mathrm{inf}}^*}(\mathbf{s}^n) := 0$, $\forall \mathbf{s}^n \in \mathcal{S}_n, n \in \mathcal{N}$.
for $n = 1$ to N **do**
  Set $j := 0$.
  while $j \le T_{\mathrm{iter}}$ **do**
    Calculate $V_j^{n,\mathrm{DA},\lambda_{\mathrm{inf}}^*}(\mathbf{s}^n)$, $\forall\, \mathbf{s}^n \in \mathcal{S}_n$ from (33).
    Set $j := j + 1$.
  end while
  Compute and store the stationary Lagrangian index $i_{\mathrm{inf}}^{n,\mathrm{DA}}(\mathbf{s}^n)$ from Definition 3, $\forall \mathbf{s}^n \in \mathcal{S}_n$.
end for
Deployment Stage:
Set $t := 1$
while $t \le T$ **do**
  Observe $\mathbf{G}_t$.
  for $n = 1$ to N **do**
    Observe $\mathbf{s}_t^n := (a_t^n, h_t^n)$ and retrieve the stored stationary Lagrangian index $i_{\mathrm{inf}}^{n,\mathrm{DA}}(\mathbf{s}_t^n)$.
  end for
  Solve problem (39) for I^{SCA} iterations, and set $\mathbf{u}_t := \mathbf{u}_t^{(I^{\mathrm{SCA}})}$
  Set $t := t + 1$.
end while

V. Performance Evaluation and Comparison

In this section, we conduct simulation experiments to evaluate the performance of the proposed algorithm. We first introduce the simulation setup and then present the results.

1) Simulation Parameters: We consider a communication scenario where the set of N users are uniformly distributed within a ring with inner radius $L_i = 10$ m and outer radius $L_o = 1.5$ km. The PS is located in the centre of the ring. The system bandwidth W is equal to 50 MHz. The transmit power of each client P_n is set to 28 dBm, and the noise variance $\sigma_{\mathrm{noise},n}^2$ is set to -97 dBm. We consider a channel model where the pathloss for client $n \in \mathcal{N}$ can be expressed as $128.1 + 37.6\log_{10}(l^n)$, and l^n denotes the distance between client n and the PS in kilometers [5]. We assume Rayleigh fading when computing the instantaneous channel gain of each client. We discretize the CSI into $|\mathcal{H}_n| = \bar{H}$ levels, for all $n \in \mathcal{N}$, based on its empirical cumulative distribution function. For stationary Lagrangian indices, we set $T_{\mathrm{iter}} = 1000$.

We consider two popular machine learning datasets. The MNIST dataset consists of pictures of hand-written digits $\{0, \ldots, 9\}$ and the CIFAR-10 dataset consists of photos of 10 classes of objects. Each client performs 50 SGD steps before sending the updated model back to the PS. We consider a convolutional neural network (CNN) with three convolutional layers and two fully-connected layers for the MNIST dataset. The size of the neural network is $\zeta = 159.8$ kB. For the CIFAR-10 dataset, we further include a dropout layer in the model. The number of samples owned by the set of clients $\mathcal{N}$ follows a Zipf distribution with parameter κ. That is, the number of samples owned by client n satisfies

$|\mathcal{X}_n| = \left\lceil \frac{n^{-\kappa}\sum_{j\in\mathcal{N}}|\mathcal{X}_j|}{\sum_{i\in\mathcal{N}} i^{-\kappa}} \right\rceil$, $n \in \mathcal{N}$, where κ represents the degree of difference between the amount of data owned by different clients. When κ is equal to zero, it corresponds to the case where all clients have the same amount of data. Among the $|\mathcal{X}_n|$ samples owned by client n, the composition of classes follows a Dirichlet distribution with parameter α. We have $\mathbb{P}(\mathbf{x}^n) \propto \prod_{i=1}^{10}(x_i^n)^{\alpha-1}$, and $\sum_{i=1}^{10} x_i^n = |\mathcal{X}_n|$, $n \in \mathcal{N}$. A smaller value of α corresponds to more severe disparity between data shared among all clients[5]. The FL experiments were constructed using the PyTorch code package, based on the source code of an existing FL project [8]. Simulations were conducted on an NVIDIA RTX2060 GPU.

2) Performance Comparison: In Fig. 2, we compare the performance of the finite-horizon Lagrangian index-based algorithm and the proposed infinite-horizon approximation stationary Lagrangian index-based algorithm. We use the performance given by the optimal solution to problem (23), *i.e.*, policy ϕ^*, as a lower bound for the optimal value of problem (20). In Fig. 2(a), we compare the Lagrange multiplier vector $\boldsymbol{\lambda}^* = (\lambda_1^*, \ldots, \lambda_T^*)$ obtained from problem (25) and the Lagrange multiplier λ_{inf}^* obtained from the infinite-horizon approximation from problem (32), when $N = 25$ and $M = 5$. The results show that the values of λ_t^* differ from λ_{inf}^* significantly only at the beginning and toward the end of the entire time-horizon. This justifies the choice of using the infinite-horizon approximation. In Fig. 2(b), we vary the number of clients N. One fifth of the clients are scheduled to participate in FL in each communication round. Each experiment was repeated 1,000 times with different random seeds. The expected total cost obtained using the finite-horizon and infinite-horizon algorithms are compared. They show similar performance. These results show that the proposed infinite-horizon approximation achieves similar performance as the finite-horizon approach.

Next, we deploy the proposed algorithm in a large-scale FL system. We consider an FL system with $N = 100$ potential clients and $M = 10$ users are scheduled to send their updated model during each communication round. We set $\bar{H} = 20$, $A_{\max} = 50$,[6] $\rho = 5$, $I^{\text{SCA}} = 10$, $\xi = 10000$, and $\rho' = 1$. The infinite-horizon approximation of the Lagrangian index was adopted. In Figs. 3 and 4, we show the performance of the proposed algorithm in terms of the following three metrics: The *global loss* captures the cross-entropy loss of the model during training. The *average duration of one communication round* corresponds to $\frac{1}{t}\sum_{\tau=1}^{t}\sum_{n=1}^{N} y(h_\tau^n, u_\tau^n)$.[7] The *testing accuracy* corresponds to the mean accuracy of all clients on the testing dataset. For the MNIST dataset, we split the dataset evenly across all clients (*i.e.*, $\kappa = 0$). For the CIFAR-10 dataset, we set $\kappa = 0.2$. We observe that the classification

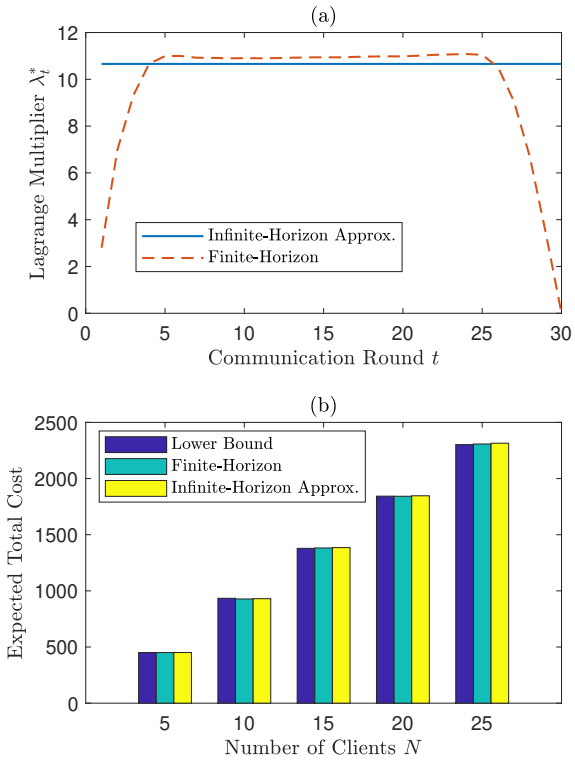


Fig. 2: Performance comparison between the proposed Lagrangian index-based algorithm (finite-horizon) and the proposed stationary Lagrangian index-based algorithm (infinite-horizon approximation) in an FL system with $\bar{H} = 5$, $T = 30$, $\kappa = 0$, $\xi = 200$, and $\rho = 0$. One fifth of the clients are scheduled in each decision epoch: (a) The Lagrangian index vector $\boldsymbol{\lambda}^*$ obtained from the dual problem of problem (25), and λ_{inf}^* obtained from the dual problem of problem (32), where $N = 25$ and $M = 5$. (b) The expected total cost obtained by the two algorithms and the lower bound.

task using the CIFAR-10 dataset takes longer to converge. We set $T = 200$ for the MNIST dataset and $T = 3000$ for the CIFAR-10 dataset. We compare the results obtained with the proposed algorithm with three baseline algorithms: The **MD sampling** algorithm [3][8] samples clients according to their local dataset size in each communication round. It has a computational complexity of $\mathcal{O}(M)$. In the **clustered sampling** algorithm [8], clustering and random sampling are performed in each communication round. It has a computational complexity of $\mathcal{O}(N^2\log(N))$. The **DivFL** algorithm [7] samples clients only based on the diversity cost, and adopts a submodular heuristic with a computational complexity of $\mathcal{O}(N^2)$. These three algorithms are chosen since they are state-of-the-art client scheduling algorithms in the literature. The global loss with respect to the number of communication rounds, achieved by the four algorithms, are plotted in the first column of Figs. 3 and 4. We observe that the proposed two-stage online algorithm achieves slightly better performance compared to the clustered sampling algorithm, which has the best performance among the three baselines. This is due to the inclusion of the diversity term in the objective function. The

[5]In real-life FL scenarios, the datasets owned by different clients are often non-IID. Therefore, in this paper, we focus on non-IID local datasets of clients. We note that when the client local datasets are IID, having the diversity term in the cost function may not improve the convergence performance of FL.

[6]This value is chosen such that a_t^n never exceeds $A_{\max}$, $t \in \mathcal{T}$, $n \in \mathcal{N}$, when using the proposed algorithm.

[7]Similar to [4], we assume that the duration of a communication round can be approximated by the total uplink transmission time of all clients that are scheduled to participate in FL.

[8]Since the initial gradient estimate for the proposed two-stage online algorithm can be inaccurate, a round robin sampling scheme was adopted in the first $\frac{N}{M}$ communication rounds to ensure that each client is sampled at least once in the beginning.

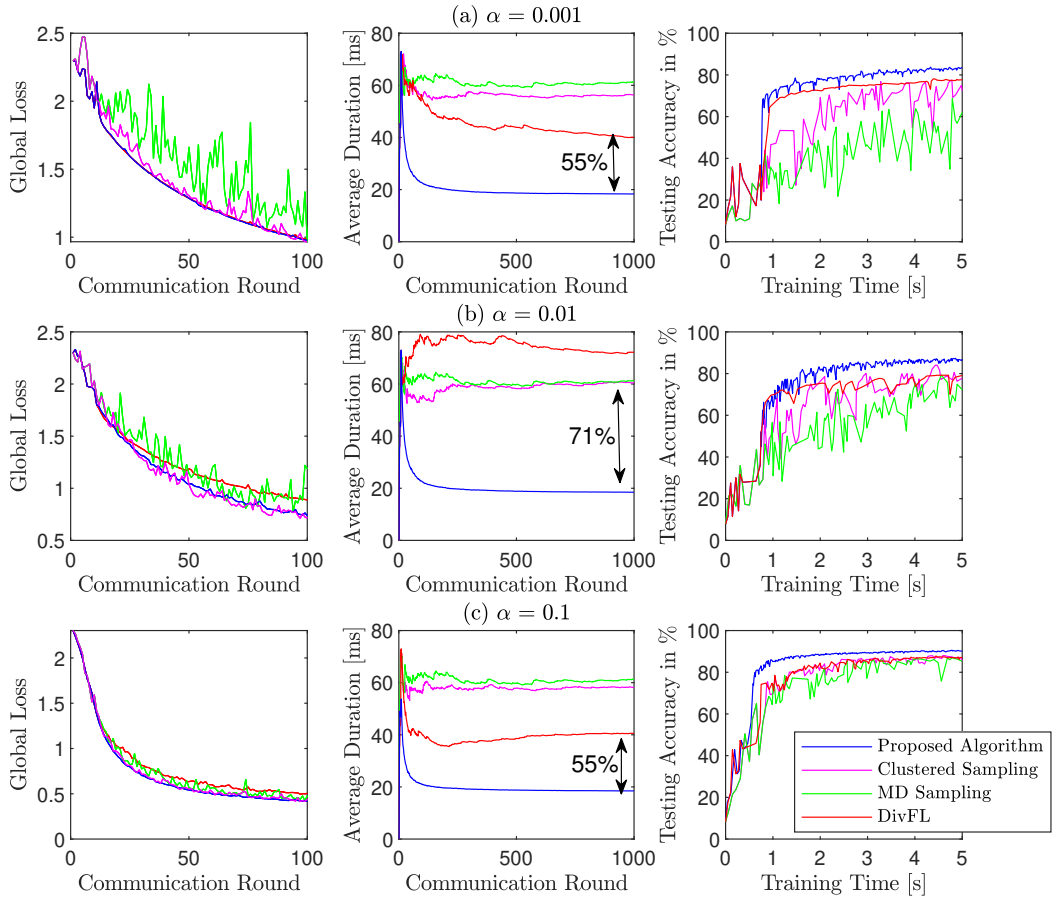


Fig. 3: Comparison of the global loss, average duration of each communication round, and the testing accuracy on the MNIST dataset between the proposed two-stage online algorithm, clustered sampling [8], MD sampling [3], and DivFL [7] algorithms when (a) $\alpha = 0.001$, (b) $\alpha = 0.01$, and (c) $\alpha = 0.1$.

average duration of each communication round is shown in the second column of Figs. 3 and 4. The proposed algorithm can reduce the average duration of each communication round by up to 71%, compared with the other three baseline algorithms. This is due to the inclusion of the duration of uplink transmission time in the objective function. The testing accuracy with respect to the elapsed time since training began is shown in the third column of Figs. 3 and 4. We observe that the testing accuracy of the proposed algorithm improves faster compared to the baseline algorithms, due to the shorter duration of the communication rounds. The performance of the proposed algorithm consistently outperforms the three baselines on local datasets with different values of α's.

To investigate the necessity of incorporating the AoI, uplink transmission time, and diversity terms in the cost function (17), we perform ablation studies and consider three cases where only two out of three terms are present in the cost function. In Fig. 5, we show the performance comparison between these three cases and the original proposed two-stage online algorithm, using the same simulation settings as in Fig. 4(a). From the global loss curves, we notice that the training converges slower without the AoI term in the cost function. This happens because a small subset of clients with good CSI and low diversity cost are repeatedly sampled to participate in FL training and the stored representative gradients at the PS are not up-to-date. Without the diversity term in the objective function, the training converges more slowly, with a larger variance. Without the CSI term, the training is also slower, due to the longer duration of the communication rounds. Thus, all three terms in (17) contributed to the overall improvement of the convergence of FL.

VI. Conclusion

In this paper, we designed a channel-aware joint AoI and diversity optimization framework to address the client scheduling problem in FL with non-IID client datasets. First, we formulated the client scheduling problem as a CMDP and derived the optimal solution using the value iteration algorithm. Then, we considered a diversity-agnostic variant of the problem and proposed a low-complexity Lagrangian index solution. The proposed Lagrangian index-based approach has the potential of being applied to other optimization problems with AoI as part of the objective function or constraints. Based on the Lagrangian index solution, we designed a two-stage online algorithm for solving the formulated CMDP. We applied the proposed algorithm to FL training for the MNIST and CIFAR-10 datasets. The results showed that the proposed algorithm can improve the training performance by up to

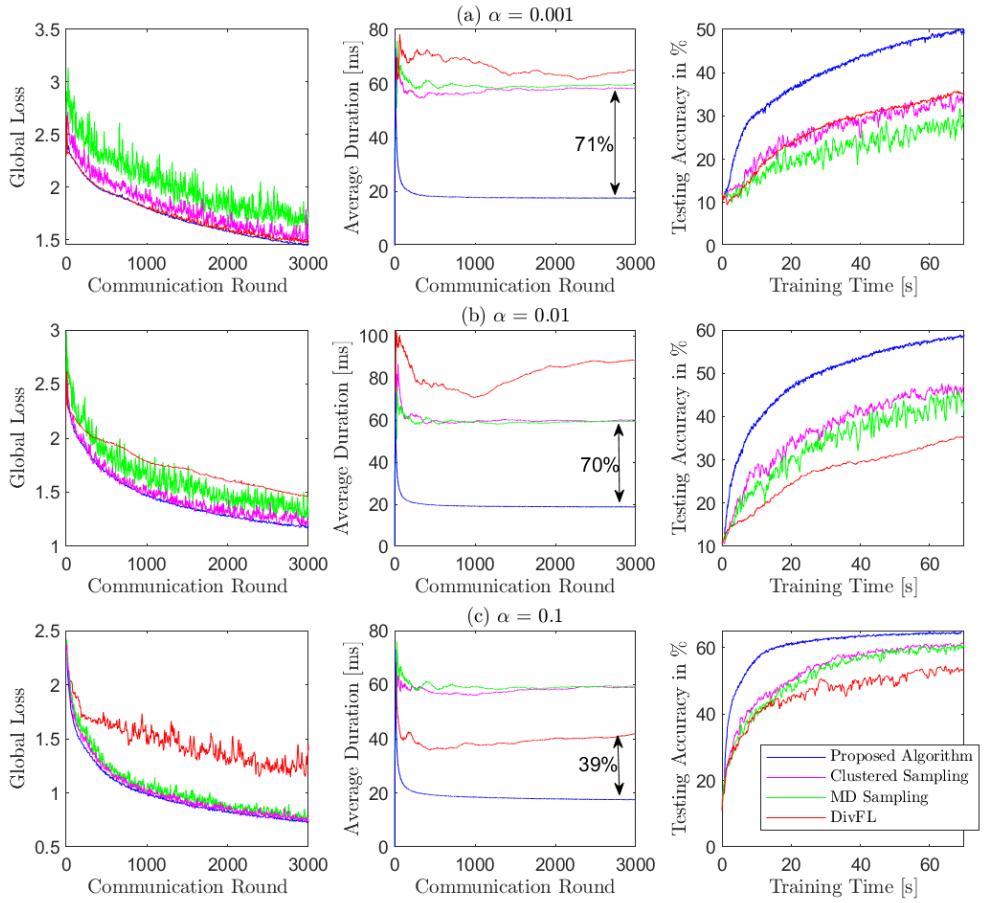


Fig. 4: Comparison of the global loss, average duration of each communication round, and the testing accuracy on the CIFAR-10 dataset between the proposed two-stage online algorithm, clustered sampling [8], MD sampling [3], and DivFL [7] algorithms when (a) $\alpha = 0.001$, (b) $\alpha = 0.01$, and (c) $\alpha = 0.1$.

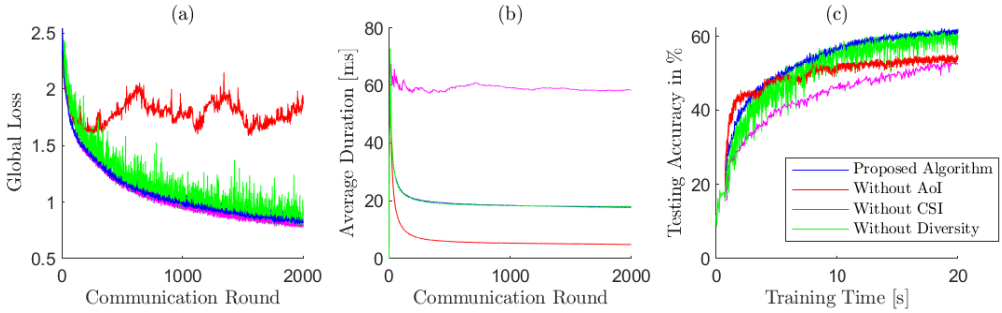


Fig. 5: Performance comparison between the proposed two-stage online algorithm with three cases, where only two out of the three terms in equation (17) are considered. The simulation settings are the same as in Fig. 4(c).

71%, in terms of uplink transmission time. Future work will study client scheduling in FL systems by considering adaptive transmit power allocation and beamforming design.

APPENDIX

In this Appendix, we show the effect of AoI on the convergence rate of diversity-based FL. Similar to [3], [7], we define $\Gamma \triangleq F^* - \sum_{n=1}^{N} p_n F_n^*$ to model the degree of heterogeneity of the client local datasets, where F^* and F_n^* denote the optimal values of F and F_n, respectively. We require function $F_n(\cdot)$, $n \in \mathcal{N}$, to satisfy the following standard assumptions:

- **Assumption 1.** The function $F_n(\cdot)$, $n \in \mathcal{N}$, is L-smooth.
- **Assumption 2.** The function $F_n(\cdot)$, $n \in \mathcal{N}$, is μ-strongly convex.
- **Assumption 3.** The variance of the stochastic gradient is bounded, such that

$$\mathbb{E}\left[||\nabla F_n(\mathbf{w}_k^n \mid \Xi_k^n) - \nabla F_n(\mathbf{w}_k^n)||^2\right] \leq \sigma_n^2,$$

  for all $n \in \mathcal{N}$, $k \in \mathcal{K}$. To simplify the notation, we also define $\sigma \triangleq \max_{n \in \mathcal{N}} \sigma_n$.
- **Assumption 4.** The expected squared norm of the

stochastic gradient is uniformly bounded, such that

$$\mathbb{E}\big[||\nabla F_n(\mathbf{w}_k^n \mid \Xi_k^n)||^2\big] \leq G^2, \quad \forall n \in \mathcal{N}, k \in \mathcal{K}.$$

Based on the above assumptions, we will show that the AoI has an impact on the convergence performance of FL.

Theorem 3. Let Assumptions 1−4 hold and L, Γ, μ, σ_n, σ, G be defined therein. Choose $\gamma = \max(8L/\mu, E) - 1$ and learning rate $\eta_k = \frac{2}{\mu(\gamma+k)}$, $k \in \mathcal{K}$[9]. Let $C_t^\eta \triangleq \sum_{i=0}^{E-1} \eta_{(t-1)E+i}$, $t \in \mathcal{T}$. Given the AoI of each client $a_t^n \leq \beta$, $n \in \mathcal{N}$ and the diversity cost in (14), $\epsilon_t^c(\mathbf{g}_t, \mathbf{u}_t)/C_{(t-\beta)}^\eta \leq \epsilon$, $t \in \mathcal{T}$, we have

$$\mathbb{E}[F(\mathbf{w}_{tE})] - F^* < \frac{L}{(\gamma + tE)}\left[\frac{2(B_1+B_2)}{\mu^2} + \frac{\gamma}{2}\mathbb{E}\left\|\mathbf{w}_0 - \mathbf{w}^*\right\|^2\right] + \frac{B_3 L}{\mu}, \quad t \in \mathcal{T}, \tag{40}$$

where B_1, B_2, and B_3 are defined as

$$B_1 \triangleq \sum_{n=1}^{N} p_n^2 \sigma_n^2 + 6L\Gamma + 8(E-1)^2 G^2,$$
$$B_2 \triangleq 4(\epsilon + 4\sigma)^2 E^2 + 2^{\beta+3}(\beta+2)LGE^2\big(2^{\beta+3} \times (\beta+2)LGE^2 + 4(\epsilon+4\sigma)E + 2(G/\mu + J)\big),$$
$$B_3 \triangleq 4(G/\mu + J)(\epsilon + 4\sigma)E,$$

and J is a positive constant.

The term inside the square bracket in (40) converges to zero when $t \to \infty$ and affects the speed of convergence. The term outside of the square bracket persists during training and is referred to as a *bias term*. Similar to [7], we observe that ϵ appeared in B_2 and B_3, and it affects both the convergence speed and the bias term. Moreover, we notice that the AoI term β only appeared in B_2. Hence, a larger AoI will lead to slower convergence of FL. To facilitate the proof of Theorem 3, we begin by introducing the following three lemmas:[10]

Lemma 1. Under Assumptions 1, 3, and 4, consider a learning rate η_k that is non-increasing and satisfies $\eta_k \leq 2\eta_{k+E}$, $k \in \mathcal{K}$. Then, for all $n \in \mathcal{N}$, $t \in \mathcal{T}$, we have

$$\mathbb{E}\big[\,\|\mathbf{q}_t^n/C_t^\eta - \nabla F_n(\mathbf{w}_{tE}^n)\|\,\big] \leq 2\eta_{(t-1)E} ELG + 2\sigma_n. \tag{41}$$

Proof: We start by defining

$$\tilde{\mathbf{q}}_t^n \triangleq C_t^\eta \nabla F_n(\mathbf{w}_{tE}^n) = \sum_{i=0}^{E-1} \eta_{(t-1)E+i} \nabla F_n(\mathbf{w}_{tE}^n), \tag{42}$$

and recall the definition of $\mathbf{q}_t^n$ from (5) and (6)

$$\begin{aligned}\mathbf{q}_t^n &= -(\mathbf{w}_{tE}^n - \mathbf{w}_{(t-1)E}^n) \\ &= \sum_{i=0}^{E-1} \eta_{(t-1)E+i} \nabla F_n(\mathbf{w}_{(t-1)E+i}^n \mid \Xi_{(t-1)E+i}^n).\end{aligned}$$

Then, we can bound the expected norm of the difference between $\mathbf{q}_t^n$ and $\tilde{\mathbf{q}}_t^n$ by

$$\begin{aligned}\mathbb{E}\left[\|\mathbf{q}_t^n - \tilde{\mathbf{q}}_t^n\|\right] \leq & \sum_{i=0}^{E-1} \eta_{(t-1)E+i} \\ & \times \mathbb{E}\left[||\nabla F_n(\mathbf{w}_{tE}^n) - \nabla F_n(\mathbf{w}_{(t-1)E+i}^n \mid \Xi_{(t-1)E+i}^n)||\right] \\ = & \sum_{i=0}^{E-1} \eta_{(t-1)E+i} \mathbb{E}\Big[||\nabla F_n(\mathbf{w}_{tE}^n) - \nabla F_n(\mathbf{w}_{(t-1)E+i}^n) \\ & + \nabla F_n(\mathbf{w}_{(t-1)E+i}^n) - \nabla F_n(\mathbf{w}_{(t-1)E+i}^n \mid \Xi_{(t-1)E+i}^n)||\Big] \\ \overset{\text{(T)}}{\leq} & \sum_{i=0}^{E-1} \eta_{(t-1)E+i} \mathbb{E}\Big[||\nabla F_n(\mathbf{w}_{tE}^n) - \nabla F_n(\mathbf{w}_{(t-1)E+i}^n)|| \\ & + ||\nabla F_n(\mathbf{w}_{(t-1)E+i}^n) - \nabla F_n(\mathbf{w}_{(t-1)E+i}^n \mid \Xi_{(t-1)E+i}^n)||\Big] \\ \overset{\text{(a)}}{\leq} & \sum_{i=0}^{E-1} \eta_{(t-1)E}(\mathbb{E}||\nabla F_n(\mathbf{w}_{(t-1)E+i}^n) - \nabla F_n(\mathbf{w}_{tE}^n)|| + \sigma_n) \\ \overset{\text{(T)}}{\leq} & \Bigg[\sum_{i=0}^{E-1} \eta_{(t-1)E} \sum_{j=i}^{E-1} \mathbb{E}\Big|\Big|\nabla F_n(\mathbf{w}_{(t-1)E+j}^n) \\ & - \nabla F_n(\mathbf{w}_{(t-1)E+j+1}^n)\Big|\Big|\Bigg] + E\sigma_n \eta_{(t-1)E} \\ \overset{\text{(A4)}}{\leq} & \eta_{(t-1)E} \sum_{i=0}^{E-1} \sum_{j=i}^{E-1} \eta_{(t-1)E+j} LG + E\sigma_n \eta_{(t-1)E} \\ \leq & \eta_{(t-1)E}^2 E^2 LG + E\sigma_n \eta_{(t-1)E},\end{aligned}$$

where inequality (a) is due to Assumption 3 and Jensen's inequality. In this way, the expected norm of the difference between $\mathbf{q}_t^n/C_t^\eta$ and $\nabla F_n(\mathbf{w}_{tE}^n)$ can be bounded by

$$\begin{aligned}\mathbb{E}\left\|\mathbf{q}_t^n/C_t^\eta - \nabla F_n(\mathbf{w}_{tE}^n)\right\| &\leq \frac{\eta_{(t-1)E}^2 E^2 LG + E\sigma_n \eta_{(t-1)E}}{\sum_{i=0}^{E-1} \eta_{(t-1)E+i}} \\ &\overset{\text{(L)}}{\leq} \frac{\eta_{(t-1)E}^2 E^2 LG + E\sigma_n \eta_{(t-1)E}}{E\eta_{tE}} \overset{\text{(b)}}{\leq} 2\eta_{(t-1)E} ELG + 2\sigma_n.\end{aligned}$$

Inequality (b) is due to the inequality $\eta_{(t-1)E}/\eta_{tE} \leq 2$, $t \in \{2, \ldots, T\}$. ∎

Lemma 2 (Impact of Using Stale Model Updates)**.** Under Assumptions 1 and 4, consider a learning rate η_k, $k \in \mathcal{K}$, which is non-increasing. In communication round $t \in \mathcal{T}$, given the AoI of client $n \in \mathcal{N}$ is a_t^n and $k \in \{i \mid i \geq (t - a_t^n)E, i \in \mathcal{K}\}$, we have

$$\mathbb{E}\left[\left\|\hat{\mathbf{q}}_t^n/C_{(t-a_t^n)}^\eta - \nabla F_n(\mathbf{w}_k^n)\right\|\right] \leq \eta_{(t-a_t^n-1)E} LG[k - (t-a_t^n)E + 2E] + 2\sigma_n. \tag{43}$$

Proof: Based on the L-smooth property of $F_n(\cdot)$ (*i.e.,* Assumption 1), we can show that

$$\mathbb{E}\left[\left\|\nabla F_n(\mathbf{w}_k^n) - \nabla F_n(\mathbf{w}_{(t-a_t^n)E}^n)\right\|\right] \overset{\text{(A1)}}{\leq} L\mathbb{E}\left[\left\|\mathbf{w}_k^n - \mathbf{w}_{(t-a_t^n)E}^n\right\|\right]$$

[9]To simplify the notation, we also define $\eta_{-k} \triangleq \frac{2}{\mu\gamma}$, $k \in \mathcal{K}$.

[10]To improve the readability of the proofs, we placed (T), (L), and (A1)−(A4) above an inequality sign when the inequality is due to the triangle inequality, the learning rate η_k, and Assumptions 1−4, respectively.

$$\overset{\text{(T)}}{\leq} L\mathbb{E}\left[\sum_{i=0}^{k-1-(t-a_t^n)E}\left\|\mathbf{w}_{k-i}^n-\mathbf{w}_{k-i-1}^n\right\|\right]$$
$$= L\sum_{i=0}^{k-1-(t-a_t^n)E}\mathbb{E}\left[\left\|\mathbf{w}_{k-i}^n-\mathbf{w}_{k-i-1}^n\right\|\right]$$
$$\overset{\text{(a)}}{\leq} L\sum_{i=0}^{k-1-(t-a_t^n)E}\eta_{k-i-1}\mathbb{E}\left[\left\|\nabla F_n(\mathbf{w}_{k-i-1}^n\mid\Xi_{k-i-1}^n)\right\|\right]$$
$$\overset{\text{(A4), (L)}}{\leq} LG\sum_{i=0}^{k-1-(t-a_t^n)E}\eta_{(t-a_t^n)E}$$
$$= \eta_{(t-a_t^n)E}LG(k-(t-a_t^n)E), \tag{44}$$

where inequality (a) comes from (5). Then, we can show

$$\mathbb{E}\left[\left\|\hat{\mathbf{q}}_t^n/C_{(t-a_t^n)}^{\eta}-\nabla F_n(\mathbf{w}_k^n)\right\|\right]$$
$$= \mathbb{E}\Bigg[\Bigg\|\hat{\mathbf{q}}_t^n/C_{(t-a_t^n)}^{\eta}-\nabla F_n(\mathbf{w}_{(t-a_t^n)E}^n)$$
$$+\nabla F_n(\mathbf{w}_{(t-a_t^n)E}^n)-\nabla F_n(\mathbf{w}_k^n)\Bigg\|\Bigg]$$
$$\overset{\text{(b)}}{\leq} \mathbb{E}\left[\left\|\mathbf{q}_{(t-a_t^n)}^n/C_{(t-a_t^n)}^{\eta}-\nabla F_n(\mathbf{w}_{(t-a_t^n)E}^n)\right\|\right]$$
$$+\mathbb{E}\left[\left\|\nabla F_n(\mathbf{w}_{(t-a_t^n)E}^n)-\nabla F_n(\mathbf{w}_k^n)\right\|\right]$$
$$\overset{\text{(c)}}{\leq} 2\eta_{(t-a_t^n-1)E}ELG+2\sigma_n+\eta_{(t-a_t^n)E}LG(k-(t-a_t^n)E)$$
$$\overset{\text{(L)}}{\leq} \eta_{(t-a_t^n-1)E}LG[k-(t-a_t^n)E+2E]+2\sigma_n,$$

where inequality (b) is due to triangle inequality and the definition of $\hat{\mathbf{q}}_t^n$ from (11). Inequality (c) is due to Lemma 1 and inequality (44). ■

Lemma 3. Under Assumptions 1, 3, and 4, consider a subset of clients $\mathcal{M}_t$ with AoI a_t^n, $n\in\mathcal{M}_t$, where $a_t^n\leq\beta$. Given a sequence of non-increasing learning rates η_k such that $\eta_k\leq 2\eta_{k+E}$, for $k\in\{i\mid i>(t-\beta)E, i\in\mathcal{K}\}$, we have

$$\mathbb{E}\left[\epsilon_k(\mathcal{M}_t)\right]\leq 1/C_{(t-\beta)}^{\eta}\mathbb{E}\left[\hat{\epsilon}_t(\mathcal{M}_t)\right]$$
$$+2\eta_{(t-\beta-1)E}LG[k-(t-\beta)E+2E]+4\sigma. \tag{45}$$

Proof: Based on triangle inequality, we have

$$\underbrace{\mathbb{E}\left[\epsilon_k(\mathcal{M}_t)\right]}_{\|T_1-T_4\|}$$
$$\overset{\text{(T)}}{\leq} \underbrace{\mathbb{E}\left[\left\|\sum_{n\in\mathcal{N}}p_n\nabla F_n(\mathbf{w}_k^n)-\sum_{n\in\mathcal{N}}\frac{1}{C_{(t-\beta)}^{\eta}}p_n\hat{\mathbf{q}}_t^n\right\|\right]}_{\|T_1-T_2\|}$$
$$+\underbrace{\mathbb{E}\left[\left\|\sum_{n\in\mathcal{N}}\frac{1}{C_{(t-\beta)}^{\eta}}p_n\hat{\mathbf{q}}_t^n-\frac{1}{|\mathcal{M}_t|}\sum_{n\in\mathcal{M}_t}\frac{1}{C_{(t-\beta)}^{\eta}}\hat{\mathbf{q}}_t^n\right\|\right]}_{\|T_2-T_3\|}$$
$$+\underbrace{\mathbb{E}\left[\frac{1}{|\mathcal{M}_t|}\left\|\sum_{n\in\mathcal{M}_t}\frac{1}{C_{(t-\beta)}^{\eta}}\hat{\mathbf{q}}_t^n-\sum_{n\in\mathcal{M}_t}\nabla F_n(\mathbf{w}_k^n)\right\|\right]}_{\|T_3-T_4\|}. \tag{46}$$

The first term on the right-hand side of (46) can be bounded by using inequality (43), such that

$$\mathbb{E}\left[\left\|\sum_{n\in\mathcal{N}}p_n\nabla F_n(\mathbf{w}_k^n)-\sum_{n\in\mathcal{N}}\frac{1}{C_{(t-\beta)}^{\eta}}p_n\hat{\mathbf{q}}_t^n\right\|\right]$$
$$= \mathbb{E}\left[\left\|\sum_{n\in\mathcal{N}}p_n\left(\nabla F_n(\mathbf{w}_k^n)-\frac{1}{C_{(t-\beta)}^{\eta}}\hat{\mathbf{q}}_t^n\right)\right\|\right]$$
$$\overset{\text{(T)}}{\leq} \sum_{n\in\mathcal{N}}p_n\mathbb{E}\left[\left\|\nabla F_n(\mathbf{w}_k^n)-\frac{1}{C_{(t-\beta)}^{\eta}}\hat{\mathbf{q}}_t^n\right\|\right]$$
$$\leq \sum_{n\in\mathcal{N}}p_n(\eta_{(t-\beta-1)E}LG[k-(t-\beta)E+2E]+2\sigma)$$
$$= \eta_{(t-\beta-1)E}LG[k-(t-\beta)E+2E]+2\sigma.$$

The second term on the right-hand side of (46) can be simplified as $1/C_{(t-\beta)}^{\eta}\mathbb{E}\left[\hat{\epsilon}_t(\mathcal{M}_t)\right]$. The third term on the right-hand side of (46) can be bounded in a similar way as the first term. ■

Using the results from these lemmas, we can now prove Theorem 3.

Proof: The proof is similar to the proof of [7, Theorem 1]. The key difference between these two theorems is the inclusion of the AoI term in our work. Similar to [3], [7], we define a sequence of $\mathbf{v}_k^n$, where $\mathbf{v}_{k+1}^n=\mathbf{w}_k^n-\eta_k\nabla F_n(\mathbf{w}_k^n\mid\Xi_k^n)$, $n\in\mathcal{N}$, $k\in\mathcal{K}$. We also define $\bar{\mathbf{w}}_k\triangleq\sum_{n\in\mathcal{N}}p_n\mathbf{w}_k^n$ and $\bar{\mathbf{v}}_k\triangleq\sum_{n\in\mathcal{N}}p_n\mathbf{v}_k^n$, $k\in\mathcal{K}$. $\bar{\mathbf{w}}_k$ and $\bar{\mathbf{v}}_k$ only differ when $k=tE$, $t\in\mathcal{T}$. Since $\bar{\mathbf{w}}_{(t-1)E+1}=\bar{\mathbf{v}}_{(t-1)E+1}$, we can bound $\mathbb{E}\left\|\bar{\mathbf{w}}_{tE}-\bar{\mathbf{v}}_{tE}\right\|$ by [7, eqns. (20)−(29)]:

$$\mathbb{E}\left[\|\bar{\mathbf{w}}_{tE}-\bar{\mathbf{v}}_{tE}\|\right]\leq\sum_{i=0}^{E-1}\eta_{(t-1)E+i}\mathbb{E}\left[\epsilon_{(t-1)E+i}(\mathcal{M}_t)\right]$$
$$\overset{\text{(L)}}{\leq} \eta_{(t-1)E}\sum_{i=0}^{E-1}\mathbb{E}\left[\epsilon_{(t-1)E+i}(\mathcal{M}_t)\right]$$
$$\leq \eta_{(t-1)E}\sum_{i=0}^{E-1}\left(\epsilon+4\sigma+2\eta_{(t-\beta-1)E}LG[(\beta+1)E+i]\right)$$
$$\overset{\text{(L)}}{\leq} A_1(tE)\triangleq 2(\epsilon+4\sigma)E\eta_{tE}+2^{\beta+3}(\beta+2)LGE^2\eta_{tE}^2. \tag{47}$$

Since we did not make any assumption on the value of $\bar{\mathbf{v}}_{tE}$, we also have

$$\mathbb{E}\left[\|\bar{\mathbf{w}}_{tE}-\bar{\mathbf{v}}_{tE}\|\mid\bar{\mathbf{v}}_{tE}=\bar{\mathbf{v}}'\right]\leq A_1(tE). \tag{48}$$

Similar to [7, eqn. (33)], we can derive a recursion for $\left\|\bar{\mathbf{w}}_{k+1}-\mathbf{w}^*\right\|^2$, where

$$\mathbb{E}\left[\left\|\bar{\mathbf{w}}_{k+1}-\mathbf{w}^*\right\|^2\right]\leq 2\mathbb{E}\left[\left\|\bar{\mathbf{w}}_{k+1}-\bar{\mathbf{v}}_{k+1}\right\|\left\|\bar{\mathbf{v}}_{k+1}-\mathbf{w}^*\right\|\right]$$
$$+\mathbb{E}\left[\left\|\bar{\mathbf{w}}_{k+1}-\bar{\mathbf{v}}_{k+1}\right\|^2\right]+\mathbb{E}\left[\left\|\bar{\mathbf{v}}_{k+1}-\mathbf{w}^*\right\|^2\right]$$
$$\overset{\text{(a)}}{\leq} 2A_1(k+1)\underbrace{\mathbb{E}\left[\left\|\bar{\mathbf{v}}_{k+1}-\mathbf{w}^*\right\|\right]}_{A_2}+\underbrace{\mathbb{E}\left[\left\|\bar{\mathbf{w}}_{k+1}-\bar{\mathbf{v}}_{k+1}\right\|^2\right]}_{A_3}$$
$$+\underbrace{\mathbb{E}\left[\left\|\bar{\mathbf{v}}_{k+1}-\mathbf{w}^*\right\|^2\right]}_{A_4}. \tag{49}$$

Here, inequality (a) is due to the fact that

$$\begin{aligned}
&\mathbb{E}\left[\|\bar{\mathbf{w}}_{k+1}-\bar{\mathbf{v}}_{k+1}\|\,\|\bar{\mathbf{v}}_{k+1}-\mathbf{w}^*\|\right]\\
&\overset{\text{(b)}}{=}\mathbb{E}_{\bar{\mathbf{v}}'}\left[\mathbb{E}\left[\|\bar{\mathbf{w}}_{k+1}-\bar{\mathbf{v}}_{k+1}\|\,\|\bar{\mathbf{v}}_{k+1}-\mathbf{w}^*\| \;\middle|\; \bar{\mathbf{v}}_{k+1}=\bar{\mathbf{v}}'\right]\right]\\
&=\mathbb{E}_{\bar{\mathbf{v}}'}\left[\,\|\bar{\mathbf{v}}'-\mathbf{w}^*\|\,\mathbb{E}\left[\|\bar{\mathbf{w}}_{k+1}-\bar{\mathbf{v}}_{k+1}\| \;\middle|\; \bar{\mathbf{v}}_{k+1}=\bar{\mathbf{v}}'\right]\right]\\
&\le A_1(k+1)\mathbb{E}_{\bar{\mathbf{v}}'}[\|\bar{\mathbf{v}}'-\mathbf{w}^*\|]=A_1(k+1)\mathbb{E}[\|\bar{\mathbf{v}}_{k+1}-\mathbf{w}^*\|].
\end{aligned}$$

Equality (b) is often referred to as the law of iterated expectations. The term A_3 can be bounded by (47). From [3, Theorem 1], the term A_4 can be bounded by

$$\mathbb{E}\left[\|\bar{\mathbf{v}}_{k+1}-\mathbf{w}^*\|^2\right]\le(1-\eta_k\mu)\mathbb{E}\left[\|\bar{\mathbf{w}}_k-\mathbf{w}^*\|^2\right]+\eta_k^2B_1.$$

From [7, eqn. (38)], the term A_2 is bounded by

$$\mathbb{E}[\|\bar{\mathbf{v}}_{k+1}-\mathbf{w}^*\|]\le G/\mu+J.$$

By defining $\Delta_k\triangleq\mathbb{E}[\|\bar{\mathbf{w}}_k-\mathbf{w}^*\|^2]$ and expanding the terms, we have

$$\Delta_{k+1}\le(1-\mu\eta_k)\Delta_k+(B_1+B_2)\eta_k^2+B_3\eta_k. \tag{50}$$

Then, let us define

$$v_k\triangleq\max\left\{\frac{4(B_1+B_2)}{\mu^2}+\frac{2B_3}{\mu}(\gamma+k),\gamma\Delta_0\right\}, \tag{51}$$

and we have $\Delta_0\le\frac{v_0}{\gamma}$. Then, given $\Delta_k\le\frac{v_k}{k+\gamma}$, we can prove that $\Delta_{k+1}\le\frac{v_{k+1}}{k+1+\gamma}$. Starting from (50), we have

$$\begin{aligned}
&\Delta_{k+1}\\
&\le\left(1-\frac{2\mu}{\mu(k+\gamma)}\right)\frac{v_k}{k+\gamma}+\frac{4(B_1+B_2)}{\mu^2(k+\gamma)^2}+\frac{2B_3}{\mu(k+\gamma)}\\
&=\frac{k+\gamma-1}{(k+\gamma)^2}v_k+\left[\frac{4(B_1+B_2)}{\mu^2(k+\gamma)^2}+\frac{2B_3}{\mu(k+\gamma)}-\frac{v_k}{(k+\gamma)^2}\right]\\
&\overset{\text{(c)}}{\le}\frac{k+\gamma-1}{(k+\gamma)^2}v_k\le\frac{v_{k+1}}{k+\gamma+1},
\end{aligned}$$

where inequality (c) is due the definition (51), hence the term inside the square bracket is non-positive. From the L-smooth property of $F(\cdot)$ (Assumption 1), we further have

$$\begin{aligned}
&\mathbb{E}[F(\bar{\mathbf{w}}_k)]-F^*\le\frac{L}{2}\Delta_k\le\frac{Lv_k}{2(k+\gamma)}\\
&<\frac{L}{2(k+\gamma)}\left[\frac{4(B_1+B_2)}{\mu^2}+\frac{2B_3}{\mu}(\gamma+k)+\gamma\Delta_0\right]\\
&\le\frac{L}{(k+\gamma)}\left[\frac{2(B_1+B_2)}{\mu^2}+\frac{\gamma}{2}\mathbb{E}\left\|\bar{\mathbf{w}}_0-\mathbf{w}^*\right\|^2\right]+\frac{B_3L}{\mu}.
\end{aligned}$$

Then, inequality (40) follows from the fact that $\bar{\mathbf{w}}_{tE}=\mathbf{w}_{tE}$, for $t\in\mathcal{T}$. ∎

REFERENCES

[1] M. Ma, V. W. S. Wong, and R. Schober, "AoI-driven client scheduling for federated learning: A Lagrangian index approach," in *Proc. of IEEE Int'l Conf. on Commun. (ICC)*, Rome, Italy, May 2023.

[2] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Proc. Int'l Conf. Artificial Intelligence and Statistics (AISTATS)*, Ft. Lauderdale, FL, Apr. 2017.

[3] X. Li, K. Huang, W. Yang, S. Wang, and Z. Zhang, "On the convergence of FedAvg on non-IID data," in *Proc. Int'l Conf. Learning Representations (ICLR)*, Apr. 2020.

[4] J. Perazzone, S. Wang, M. Ji, and K. Chan, "Communication-efficient device scheduling for federated learning using stochastic optimization," in *Proc. IEEE Int'l Conf. on Comput. Commun. (INFOCOM)*, Jul. 2022.

[5] S. Liu, G. Yu, X. Chen, and M. Bennis, "Joint user association and resource allocation for wireless hierarchical federated learning with IID and non-IID data," *IEEE Trans. Wireless Commun.*, vol. 21, no. 10, pp. 7852–7866, Oct. 2022.

[6] J. Ren, Y. He, D. Wen, G. Yu, K. Huang, and D. Guo, "Scheduling for cellular federated edge learning with importance and channel awareness," *IEEE Trans. Wireless Commun.*, vol. 19, no. 11, pp. 7690–7703, Nov. 2020.

[7] R. Balakrishnan, T. Li, T. Zhou, N. Himayat, V. Smith, and J. Bilmes, "Diverse client selection for federated learning via submodular maximization," in *Proc. Int'l Conf. Learning Representations (ICLR)*, Apr. 2022.

[8] Y. Fraboni, R. Vidal, L. Kameni, and M. Lorenzi, "Clustered sampling: Low-variance and improved representativity for clients selection in federated learning," in *Proc. Int'l Conf. Machine Learning (ICML)*, Jul. 2021.

[9] H. Wang, Z. Kaplan, D. Niu, and B. Li, "Optimizing federated learning on non-IID data with reinforcement learning," in *Proc. IEEE Int'l Conf. on Comput. Commun. (INFOCOM)*, Jul. 2020.

[10] E. Ozfatura, D. Gündüz, and H. V. Poor, "Collaborative learning over wireless networks: An introductory overview," in *Machine Learning and Wireless Communications*. Cambridge University Press, 2022.

[11] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *Proc. IEEE Int'l Conf. on Comput. Commun. (INFOCOM) Mini-Conf.*, Orlando, FL, Mar. 2012.

[12] R. D. Yates, P. Ciblat, A. Yener, and M. Wigger, "Age-optimal constrained cache updating," in *Proc. IEEE Int'l Symp. on Inf. Theory (ISIT)*, Aachen, Germany, Jun. 2017.

[13] M. Ma and V. W. S. Wong, "Age of information driven cache content update scheduling for dynamic contents in heterogeneous networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 8427–8441, Dec. 2020.

[14] M. Bastopcu and S. Ulukus, "Age of information for updates with distortion: Constant and age-dependent distortion constraints," *IEEE/ACM Trans. Netw.*, vol. 29, no. 6, pp. 2425–2438, Dec. 2021.

[15] H. H. Yang, A. Arafa, T. Q. Quek, and H. V. Poor, "Age-based scheduling policy for federated learning in mobile edge networks," in *Proc. Int'l Conf. Acoustics, Speech and Sig. Processing (ICASSP)*, May 2020.

[16] Y.-P. Hsu, E. Modiano, and L. Duan, "Scheduling algorithms for minimizing age of information in wireless broadcast networks with random arrivals," *IEEE Trans. Mobile Comput.*, vol. 19, no. 12, pp. 2903–2915, Dec. 2019.

[17] V. Tripathi and E. Modiano, "A Whittle index approach to minimizing functions of age of information," in *Proc. IEEE Allerton Conf. on Commun., Contr., and Comp. (Allerton)*, Urbana, IL, Sep. 2019.

[18] P. Whittle, "Restless bandits: Activity allocation in a changing world," *Journal of Applied Probability*, vol. 25, pp. 287–298, 1988.

[19] K. Nakhleh, S. Ganji, P.-C. Hsieh, I.-H. Hou, and S. Shakkottai, "NeurWIN: Neural Whittle index network for restless bandits via deep RL," in *Proc. Conf. Neural Inf. Process. Syst. (NeurIPS)*, Dec. 2021.

[20] D. T. Nguyen, A. Kumar, and H. C. Lau, "Credit assignment for collective multiagent RL with global rewards," in *Proc. Conf. Neural Inf. Process. Syst. (NeurIPS)*, Montreal, Canada, Dec. 2018.

[21] D. B. Brown and J. E. Smith, "Index policies and performance bounds for dynamic selection problems," *Management Science*, vol. 66, no. 7, pp. 3029–3050, Jul. 2020.

[22] L. Hao, Y. Xu, and L. Tong, "Asymptotically optimal Lagrangian priority policy for deadline scheduling with processing rate limits," *IEEE Trans. Autom. Control*, vol. 67, no. 1, pp. 236–250, Jan. 2022.

[23] D. P. Bertsekas, *Dynamic Programming and Optimal Control, vol. 1, 4th ed.* Belmont, MA: Athena Scientific, 2017.

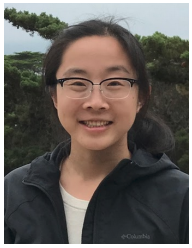
[24] B. Mirzasoleiman, J. Bilmes, and J. Leskovec, "Coresets for data-efficient training of machine learning models," in *Proc. Int'l Conf. on Machine Learning (ICML)*, Jul. 2020.

[25] M. L. Littman, T. L. Dean, and L. P. Kaelbling, "On the complexity of solving Markov decision problems," in *Proc. of Conf. on Uncertainty in Artificial Intelligence*, San Francisco, CA, Aug. 1995.

[26] J. van den Brand, "A deterministic linear program solver in current matrix multiplication time," in *Proc. Annual ACM-SIAM Symp. on Discrete Algorithms (SODA)*, Salt Lake City, UT, Jan. 2020.

[27] D. P. Bertsekas, *Dynamic Programming and Optimal Control, vol. 2, 4th ed.* Belmont, MA: Athena Scientific, 2012.

[28] Y. Sun, D. W. K. Ng, Z. Ding, and R. Schober, "Optimal joint power and subcarrier allocation for full-duplex multicarrier non-orthogonal multiple access systems," *IEEE Trans. Commun.*, vol. 65, no. 3, pp. 1077–1091, Mar. 2017.
[29] E. Che, H. D. Tuan, and H. H. Nguyen, "Joint optimization of cooperative beamforming and relay assignment in multi-user wireless relay networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 10, pp. 5481–5495, Oct. 2014.
[30] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.2," http://cvxr.com/cvx, Jan. 2020.



Manyou Ma (S'19) received the B.Eng degree from McGill University, Montreal, Canada, in 2014 and the M.A.Sc. degree from the University of British Columbia (UBC), Vancouver, Canada, in 2016. She is currently pursuing the Ph.D. degree in electrical and computer engineering at UBC. From February 2020 to July 2021, she worked as a Mitacs Accelerate intern with Rogers Communications Canada Inc., Vancouver, Canada. From July 2021 to January 2022, she worked as a research intern with Samsung Research America, Montreal, Canada. Her research interests include deep learning-based wireless scheduling algorithms design and age of information.

Vincent W.S. Wong (S'94, M'00, SM'07, F'16) received the B.Sc. degree from the University of Manitoba, Canada, in 1994, the M.A.Sc. degree from the University of Waterloo, Canada, in 1996, and the Ph.D. degree from the University of British Columbia (UBC), Vancouver, Canada, in 2000. From 2000 to 2001, he worked as a systems engineer at PMC-Sierra Inc. (now Microchip Technology Inc.). He joined the Department of Electrical and Computer Engineering at UBC in 2002 and is currently a Professor. His research areas include protocol design, optimization, and resource management of communication networks, with applications to 5G/6G wireless networks, Internet of things, mobile edge computing, smart grid, and energy systems. Dr. Wong is the Editor-in-Chief of *IEEE Transactions on Wireless Communications*. He has served as an Area Editor of *IEEE Transactions on Communications* and *IEEE Open Journal of the Communications Society*, an Associate Editor of *IEEE Transactions on Mobile Computing* and *IEEE Transactions on Vehicular Technology*, and a Guest Editor of *IEEE Journal on Selected Areas in Communications*, *IEEE Internet of Things Journal*, and *IEEE Wireless Communications*. Dr. Wong is the General Co-chair of *IEEE INFOCOM* 2024. He was a Tutorial Co-Chair of *IEEE GLOBECOM*'18, a Technical Program Co-chair of *IEEE VTC*2020-*Fall* and *IEEE SmartGridComm*'14, and a Symposium Co-chair of *IEEE ICC*'18, *IEEE SmartGridComm* ('13, '17) and *IEEE GLOBECOM*'13. He received the Best Paper Award at the *IEEE ICC* 2022 and *IEEE GLOBECOM* 2020. He is the Chair of the IEEE Vancouver Joint Communications Chapter and has served as the Chair of the IEEE Communications Society Emerging Technical Sub-Committee on Smart Grid Communications. Dr. Wong is an IEEE Vehicular Technology Society Distinguished Lecturer (2023−2025) and was an IEEE Communications Society Distinguished Lecturer (2019−2020).

Robert Schober (S'98, M'01, SM'08, F'10) received the Diplom (Univ.) and the Ph.D. degrees in electrical engineering from Friedrich-Alexander University of Erlangen-Nuremberg (FAU), Germany, in 1997 and 2000, respectively. From 2002 to 2011, he was a Professor and Canada Research Chair at the University of British Columbia (UBC), Vancouver, Canada. Since January 2012 he is an Alexander von Humboldt Professor and the Chair for Digital Communication at FAU. His research interests fall into the broad areas of Communication Theory, Wireless and Molecular Communications, and Statistical Signal Processing.

Robert received several awards for his work including the 2002 Heinz Maier-Leibnitz Award of the German Science Foundation (DFG), the 2004 Innovations Award of the Vodafone Foundation for Research in Mobile Communications, a 2006 UBC Killam Research Prize, a 2007 Wilhelm Friedrich Bessel Research Award of the Alexander von Humboldt Foundation, the 2008 Charles McDowell Award for Excellence in Research from UBC, a 2011 Alexander von Humboldt Professorship, a 2012 NSERC E.W.R. Stacie Fellowship, a 2017 Wireless Communications Recognition Award by the IEEE Wireless Communications Technical Committee, and the 2022 IEEE Vehicular Technology Society Stuart F. Meyer Memorial Award. Furthermore, he received numerous Best Paper Awards for his work including the 2022 ComSoc Stephen O. Rice Prize and the 2023 ComSoc Leonard G. Abraham Prize. Since 2017, he has been listed as a Highly Cited Researcher by the Web of Science. Robert is a Fellow of the Canadian Academy of Engineering, a Fellow of the Engineering Institute of Canada, and a Member of the German National Academy of Science and Engineering.

He served as Editor-in-Chief of the *IEEE Transactions on Communications*, VP Publications of the IEEE Communication Society (ComSoc), ComSoc Member at Large, and ComSoc Treasurer. Currently, he serves as Senior Editor of the *Proceedings of the IEEE* and as ComSoc President-Elect.