Online Optimization for Over-the-Air Federated Learning with Energy Harvesting

Qiaochu An, Yong Zhou, *Senior Member, IEEE*, Zhibin Wang, Hangguan Shan, *Senior Member, IEEE*, Yuanming Shi, *Senior Member, IEEE*, and Mehdi Bennis, *Fellow*, *IEEE*

***Abstract*—Federated learning (FL) is recognized as a promising privacy-preserving distributed machine learning paradigm, given its potential to enable collaborative model training among distributed devices without sharing their raw data. However, supporting FL over wireless networks confronts the critical challenges of periodically executing power-hungry training tasks on energy-constrained devices and transmitting high-dimensional model updates over spectrum-limited channels. In this paper, we reap the benefits of both energy harvesting (EH) and over-the-air computation (AirComp) to alleviate the battery limitation by harvesting ambient energy to support both the training and transmission of local models, and to achieve low-latency model aggregation by concurrently transmitting local gradients via AirComp. We characterize the convergence of the proposed FL by deriving an upper bound of the expected optimality gap, revealing that the convergence depends on the accumulated errors due to partial device participation and model distortion, both of which further depend on dynamic energy levels. To accelerate the convergence, we formulate a joint AirComp transceiver design and device scheduling problem, which is then tackled by developing an efficient Lyapunov-based online optimization algorithm. Simulations demonstrate that, by appropriately scheduling devices and allocating energy across multiple communication rounds, our proposed algorithm achieves a much better learning performance than benchmarks.**

***Index Terms*—Federated learning, Lyapunov optimization, energy harvesting, over-the-air computation.**

I. Introduction

The commercial success of diverse intelligent applications (e.g., target advertising, autonomous driving, smart manufacturing) boosts the advancement of various machine learning (ML) algorithms in both industry and academia. The conventional centralized training methodology that involves raw data exchange typically incurs huge communication load and severe privacy leakage [1]. To tackle these issues, federated learning (FL) emerges as a novel distributed ML paradigm that enables the collaborative model training among geographically dispersed edge devices without sharing their raw data [2]. As only gradients/models are exchanged between the edge server and devices, FL is capable of maintaining low communication load and mitigating privacy leakage.

Despite the great promise, the performance of FL over wireless networks is severely affected by various communication issues, necessitating the joint communication and learning design to facilitate wireless FL [3]–[7]. Specifically, the authors in [3] investigate how the number of local updates per communication round affects the learning performance under limited resource budgets. In [4], a device scheduling policy that accounts for both channel condition and model importance is proposed. Under the communication and computation latency constraint, the authors in [5] propose an energy-efficient FL scheme by jointly optimizing the allocation of time, bandwidth, computation frequency, and transmit power. The authors in [6] design a joint transmission and scheduling policy to speed up the training convergence. The joint communication-and-learning design is further extended to the multiple-input–multiple-output (MIMO) scenario [7]. However, all aforementioned studies transmit high-dimensional gradient/model with orthogonal multiple access (OMA) schemes, leading to excessive transmission latency when the edge devices are large in quantity.

Given the potential to enable concurrent non-orthogonal transmission, over-the-air computation (AirComp) can be leveraged to achieve low-latency wireless FL. AirComp-assisted FL is initially studied in [8]–[11], where the model/gradient updates are simultaneously uploaded by edge devices and then directly aggregated at the edge server by leveraging the waveform-superposition nature of radio channels. The transceiver beamforming and device scheduling are jointly optimized for AirComp-assisted FL in [11], while reconfigurable intelligent surface is utilized in [12], [13] to assist the gradient transmission and hence improve the aggregation accuracy. In [14], a power control strategy is developed to reduce the model aggregation error due to receiver noise. The authors in [15] optimize the learning performance of over-the-air FL by managing the interference in both uplink and downlink transmission. As synchronization is an important issue of AirComp, the authors in [16] propose a filtering and sampling scheme to combat misaligned model updates at the edge server. In addition, over-the-air FL is also considered in many other scenarios, such as hierarchical networks [17], device-to-device networks [18], and millimeter wave networks

The work of Yong Zhou was supported in part by the National Natural Science Foundation of China under Grants 62001294, U20A20159, and 61971286, and in part by the Natural Science Foundation of Shanghai under Grant 23ZR1442800. The work of Hangguan Shan was supported in part by the National Natural Science Foundation of China under Grants U21B2029 and U21A20456, and the Zhejiang Provincial Natural Science Foundation of China under Grant LR23F010006. The work of Yuanming Shi was supported in part by the Natural Science Foundation of Shanghai under Grant No. 21ZR1442700, the National Nature Science Foundation of China under Grant 62271318, and the Shanghai Rising-Star Program under Grant No. 22QA1406100. (*Corresponding Author: Yong Zhou.*)

Q. An, Y. Zhou, Z. Wang, and Y. Shi are with the School of Information Science and Technology, ShanghaiTech University, Shanghai, 201210, China (e-mail: {anqch, zhouyong, wangzhb, shiym}@shanghaitech.edu.cn).

H. Shan is with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, 310027, China (e-mail: hshan@zju.edu.cn).

M. Bennis is with the Centre for Wireless Communication, University of Oulu, Finland (email: mehdi.bennis@oulu.fi).

[19]. However, all the aforementioned studies overlook the effect of high energy consumption on the FL performance due to periodic local training and model uploading, which is critically important especially in the scenario with battery-limited edge devices. When the battery energy is exhausted, the edge device will not be able to participate in the training process and in turn degrade the learning performance. To resolve this issue, energy harvesting (EH) is a promising technology that can be utilized to prolong the service time of edge devices and in turn enhance the learning performance.

EH is a disruptive technology for enabling devices with EH capabilities to collect ambient energy, such as solar, thermoelectric, and electromagnetic radiation, to perform computation and transmission tasks [20]. With renewable energy, the battery limitation of edge devices can be mitigated, which motivates recent studies on EH-enabled FL [21]–[25]. Specifically, the authors in [21] develop a wireless-powered FL framework, where devices harvest electromagnetic energy from the edge server and dedicated charging stations. In [22] and [23], simultaneous wireless information and power transfer (SWIPT) is leveraged to collect energy from downlink beamforming to support subsequent uplink model aggregation. To fully utilize the harvested energy, various device scheduling strategies are investigated in EH-enabled FL systems [22]–[25]. Specifically, the authors in [22] formulate a stochastic shortest path problem to schedule devices for reducing the long-term energy consumption. The authors in [23] investigate the joint design of device scheduling and resource allocation for SWIPT-enabled unmanned aerial vehicle swarms. Besides, the authors in [24] schedule EH-enabled devices to minimize the training and uploading duration, while the authors in [25] propose a greedy device scheduling policy based on the available energy and channel coefficients.

A. Motivation and Contributions

Most prior works [21]–[25] on EH-enabled FL adopt heuristic performance metrics such as energy consumption, number of participating devices, and training time to guide resource allocation, without explicitly accounting for the impact of EH on the performance of wireless FL, which inevitably leads to performance loss. Different from the existing studies, we consider the joint communication-and-learning design and directly minimize the expected optimality gap of the global loss function, while taking into account the impact of EH. Besides, most existing works on over-the-air FL only focus on designing transmission strategies for a typical communication round, but the ultimate FL performance is affected by the accumulated model aggregation error over all communication rounds. Instead of treating each communication round equally, we optimize the EH-enabled over-the-air FL from a long-term perspective. We note that [14] is an early attempt to consider over-the-air FL from a long-term perspective, which however does not account for energy-constrained devices and assumes that future channel state information (CSI) is known beforehand.

In this paper, we leverage the superiority of AirComp and EH to enable communication- and energy-efficient wireless FL, where ambient energy harvested by distributed devices is first stored and then fully utilized to support both local model training and uplink gradient transmission. To minimize the expected optimality gap of the global loss function over the entire FL training process, we jointly optimize the device scheduling and AirComp transceiver design, while taking into account dynamic energy arrival at each device. To tackle the formulated stochastic optimization problem, we develop an effective online optimization algorithm that does not depend on any prior knowledge on future CSI and energy arrivals. The main contributions are summarized as follows.

- We develop a novel wireless FL framework that integrates AirComp and EH to enable low-latency gradient aggregation and mitigate battery limitation. Under such an FL framework, we derive the convergence rate and optimality gap that explicitly characterize the impact of the device scheduling as well as the accumulated model aggregation error, both of which further depend on the dynamic energy levels.
- Considering the dynamic energy arrival at each device, we formulate a stochastic optimization problem that requires the joint optimization of the transmit power of edge devices, aggregation beamformer of the base station (BS), and device scheduling to enhance the test accuracy. We develop an online optimization algorithm by using tools from Lyapunov optimization. In each communication round, we adopt Gibbs sampling for effective device scheduling, and optimize the transmit power and receive beamforming vector by using the derived closed-form expressions. Through theoretical analysis, we reveal that the gap between the proposed online algorithm and the optimal solution depends on the maximum transmit power and the energy charging rate.
- We conduct extensive simulations to illustrate the superiority of both the proposed over-the-air FL framework and online optimization algorithm. Simulations show that the proposed algorithm is capable of achieving a near-optimal performance and outperforms the benchmarks in terms of test accuracy. Compared with myopic benchmarks that optimize FL systems from the perspective of individual communication rounds, the test accuracy gain verifies that the proposed framework is capable of better utilizing the limited battery energy.

B. Organization and Notation

We describe the learning model, over-the-air gradient aggregation, and energy update model in Section II. We conduct the convergence analysis and formulate a joint communication-and-learning design problem in Section III. We elaborate the proposed online optimization algorithm and conduct the performance analysis in Section IV. Simulation results are illustrated in Section V. Section VI concludes this paper.

Scalar and vector are represented by italic and boldface lowercase symbols, respectively. $\mathbb{C}^{m\times n}$ represents the space of $m \times n$ complex matrices. Superscripts $(\cdot)^*$, $(\cdot)^{\mathsf{T}}$, $(\cdot)^{\mathsf{H}}$, and $(\cdot)^{-1}$ refer to conjugate, transpose, Hermitian transpose, and inverse operations, respectively. Operators $\mathbb{E}[\cdot]$ and ∇ stand for

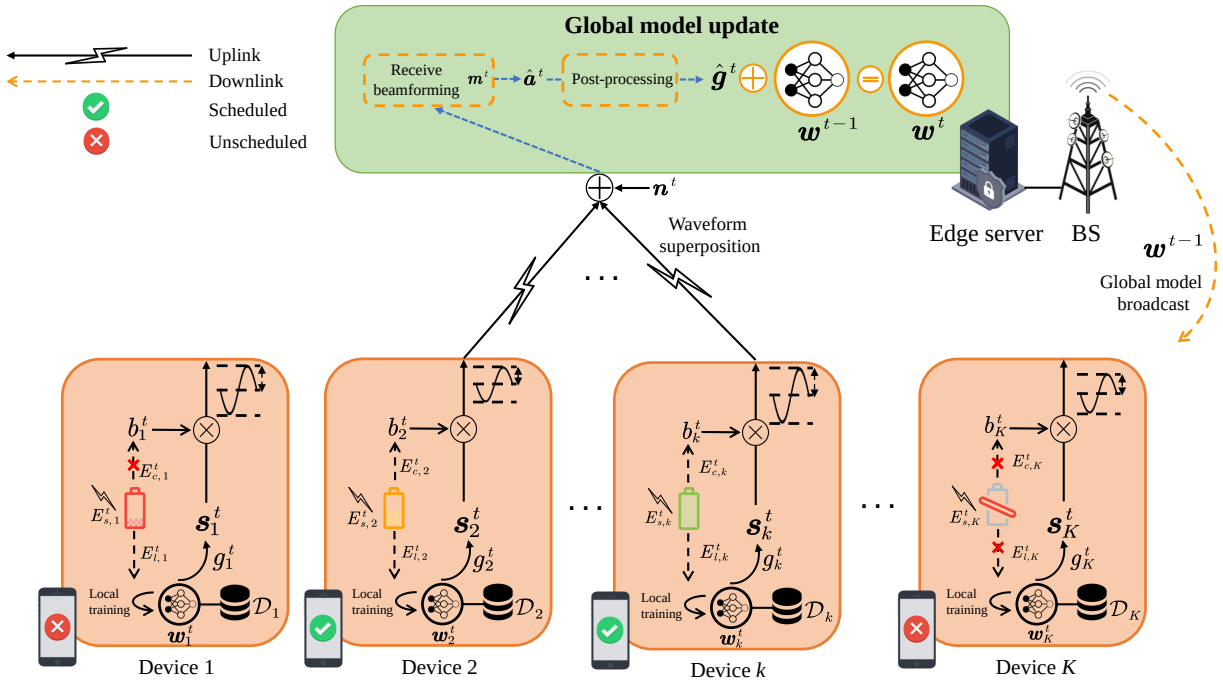


Fig. 1. An illustration of the over-the-air FL system with EH devices.

the expectation and gradient, respectively. Operators $|\cdot|$ and $\|\cdot\|$ denote the absolute value of scalar numbers and Euclidean norm of vectors, respectively.

II. System Model

This section presents the system model of EH-enabled over-the-air FL, where the devices harvest energy to support local model training and uplink gradient transmission.

A. FL Model

Consider FL over a single-cell wireless network, where K single-antenna devices with energy harvesting and storage capabilities cooperatively train a statistical global model parameterized by $\boldsymbol{w} = [w_1, \ldots, w_d]^{\mathsf{T}} \in \mathbb{R}^d$ under the coordination of an edge server co-located with an M-antenna BS, as depicted in Fig. 1. The terms edge server and BS are used interchangeably hereafter. The set of devices is denoted as $\mathcal{K} = \{1, \ldots, K\}$ and each device $k \in \mathcal{K}$ owns a local dataset $\mathcal{D}_k = \{(\boldsymbol{x}_i, y_i)\}_{i=1}^{D_k}$ with $D_k = |\mathcal{D}_k|$ data pairs. Each data pair $(\boldsymbol{x}_i, y_i)$ consists of sample $\boldsymbol{x}_i$ and its corresponding ground-truth label y_i. We assume that the local datasets are non-overlapping and equal-sized, i.e., $\mathcal{D}_k \cap \mathcal{D}_j = \emptyset$ and $D_k = D_j$, $\forall\, k \neq j$. All local datasets constitute a global dataset $\mathcal{D} = \cup_{k\in\mathcal{K}} \mathcal{D}_k$ with $D = |\mathcal{D}|$ data pairs.

For a learning task, the global loss function is defined as

$$F(\boldsymbol{w}) = \frac{1}{D} \sum_{(\boldsymbol{x}_i, y_i)\in\mathcal{D}} f(\boldsymbol{w}; \boldsymbol{x}_i, y_i), \tag{1}$$

where sample-wise loss function $f(\boldsymbol{w}; \boldsymbol{x}_i, y_i) \in \mathbb{R}$ evaluates the prediction error on data pair $(\boldsymbol{x}_i, y_i)$ and is denoted as $f_i(\boldsymbol{w})$ for brevity hereafter. By denoting $\boldsymbol{w}^\star$ as the optimal model parameter that minimizes $F(\boldsymbol{w})$, we have

$$\boldsymbol{w}^\star = \arg\min_{\boldsymbol{w}\in\mathbb{R}^d} F(\boldsymbol{w}). \tag{2}$$

Following the principle of FL, problem (2) can be solved by performing local training on the devices in parallel, each of which applies the first-order optimization algorithm to minimize its local loss function. By denoting the local loss function $F_k(\boldsymbol{w})$ at device k as

$$F_k(\boldsymbol{w}) = \frac{1}{D_k} \sum_{(\boldsymbol{x}_i, y_i)\in\mathcal{D}_k} f_i(\boldsymbol{w}), \forall\ k \in \mathcal{K}, \tag{3}$$

we can rewrite $F(\boldsymbol{w})$ as

$$F(\boldsymbol{w}) = \frac{1}{D} \sum_{k\in\mathcal{K}} \sum_{(\boldsymbol{x}_i, y_i)\in\mathcal{D}_k} f_i(\boldsymbol{w}) = \sum_{k\in\mathcal{K}} \frac{D_k}{D} F_k(\boldsymbol{w}). \tag{4}$$

The overall training process for minimizing the global loss in (4) is assumed to involve T rounds indexed by $\mathcal{T} = \{1, \ldots, T\}$. In the t-th round, the following steps are performed:

- The edge server distributes global model $\boldsymbol{w}^{t-1}$ to the set of scheduled devices, denoted as $\mathcal{S}^t$ (the device scheduling policy shall be elaborated in Section V-B). The downlink transmission of $\boldsymbol{w}^{t-1}$ is assumed to be error-free[1].

[1] As in most of the existing studies [26], the edge server combats the channel fading and receiver noise by utilizing a large transmit power.

- After receiving $\boldsymbol{w}^{t-1}$, each device $k \in \mathcal{S}^t$ initializes its current local model parameter by setting $\boldsymbol{w}_k^{t-1} = \boldsymbol{w}^{t-1}$ and then computes the local gradient as follows

$$\boldsymbol{g}_k^t = \nabla F_k(\boldsymbol{w}^{t-1}) = \frac{1}{D_k} \sum_{(\boldsymbol{x}_i, y_i) \in \mathcal{D}_k} \nabla f_i(\boldsymbol{w}^{t-1}),\ \forall\, k \in \mathcal{S}^t. \tag{5}$$

- Each device $k \in \mathcal{S}^t$ uploads local gradient $\boldsymbol{g}_k^t$ to the BS via multiple-access fading channels. With limited spectrum and power resources, the local gradients uploading suffers from inevitable distortion. Hence, the local gradient of device k received at the edge server is denoted as $\hat{\boldsymbol{g}}_k^t$. After collecting the updated local gradients $\{\hat{\boldsymbol{g}}_k^t\}$, the edge server updates the global model as follows

$$\boldsymbol{w}^t = \boldsymbol{w}^{t-1} - \gamma^t \frac{\sum_{k \in \mathcal{S}^t} D_k \hat{\boldsymbol{g}}_k^t}{\sum_{k \in \mathcal{S}^t} D_k} = \boldsymbol{w}^{t-1} - \gamma^t \hat{\boldsymbol{g}}^t, \tag{6}$$

  where $\hat{\boldsymbol{g}}^t$ is the aggregated global gradient and γ^t is the learning rate in the t-th round.

The above training procedures are executed repeatedly until either a specific learning accuracy is achieved or the maximum number of rounds is reached.

B. Over-the-Air Gradient Aggregation

The communication efficiency of FL is mainly bottlenecked by the gradient uploading process. With OMA, the edge server can only perform gradient aggregation after successfully decoding all local gradients transmitted by the selected devices. This may lead to excessive communication latency especially when K is large. As can be observed from (6), performing the global gradient aggregation only requires the weighted average of local gradients, without the need of decoding each local gradient. Thus, we resort to adopting AirComp, which allows the concurrent transmission of local gradients from all selected devices over the same wireless channel, thereby reducing the communication latency. Specifically, all devices in $\mathcal{S}^t$ concurrently transmit their updated gradients over the same time-frequency block. By superposing the signals over the air, the weighted average of local gradients can be directly obtained at the edge server, thereby avoiding the process of decoding each local gradient. AirComp-assisted uplink gradient aggregation is elaborated in the following.

Note that each device is only access to its own data statistics. Before the uplink transmission, we first normalize the local gradient at each device to facilitate its power control, i.e.,

$$\boldsymbol{s}_k^t = \frac{\boldsymbol{g}_k^t - \bar{g}_k^t \mathbf{1}}{\zeta_k^t}, \forall\, k \in \mathcal{S}^t, \tag{7}$$

where the mean and standard deviation of the local gradient $\boldsymbol{g}_k^t$ are respectively defined as

$$\bar{g}_k^t = \frac{1}{d} \sum_{j=1}^{d} g_{k,j}^t,\ \forall\, k \in \mathcal{S}^t, \tag{8a}$$

$$\zeta_k^t = \sqrt{\frac{1}{d} \sum_{j=1}^{d} \left(g_{k,j}^t - \bar{g}_k^t\right)^2},\ \forall\, k \in \mathcal{S}^t. \tag{8b}$$

Thus, every entry of the normalized local gradient vector $\boldsymbol{s}_k^t$ has zero mean and unit variance, i.e., $\mathbb{E}[s_{k,j}^t] = 0$ and $\mathbb{E}[s_{k,j}^t (s_{k,j}^t)^*] = 1,\ \forall\ j \in \{1, 2, \ldots, d\}$.

The channel coefficient between device k and the BS in the t-th round is denoted as $\boldsymbol{h}_k^t \in \mathbb{C}^{M \times 1}$. By assuming that all devices are well synchronized[2], the signal received at the BS in the t-th round is

$$\boldsymbol{r}^t = \sum_{k \in \mathcal{S}^t} \boldsymbol{h}_k^t b_k^t s_{k,j}^t + \boldsymbol{n}^t, \tag{9}$$

where $b_k^t \in \mathbb{C}$ is the transmit scalar of device k and $\boldsymbol{n}^t \sim \mathcal{CN}(\mathbf{0}, \sigma_n^2 \boldsymbol{I})$ represents the additive white Gaussian noise (AWGN) vector. Each device k has the maximum transmit power $P_k^{\max}$, i.e.,

$$\mathbb{E}\left[|b_k^t s_{k,j}^t|^2\right] = |b_k^t|^2 = P_k^t \le P_k^{\max}, \forall\, k \in \mathcal{S}^t, \tag{10}$$

where P_k^t is the transmit power of device k in the t-th round. The signal after receive beamforming is

$$\hat{a}_j^t = (\boldsymbol{m}^t)^{\mathsf{H}} \boldsymbol{r}^t = (\boldsymbol{m}^t)^{\mathsf{H}} \sum_{k \in \mathcal{S}^t} \boldsymbol{h}_k^t b_k^t s_{k,j}^t + (\boldsymbol{m}^t)^{\mathsf{H}} \boldsymbol{n}^t, \tag{11}$$

where $\boldsymbol{m}^t \in \mathbb{C}^M$ denotes the BS's receive beamforming vector. After receiving all d entries of the gradients, the global gradient can be obtained at the edge server by performing the post-processing as follows

$$\hat{\boldsymbol{g}}^t = \frac{1}{|\mathcal{S}^t|} \Big(\hat{\boldsymbol{a}}^t + \sum_{k \in \mathcal{S}^t} \bar{g}_k \mathbf{1}\Big), \tag{12}$$

where $\hat{\boldsymbol{a}}^t = [\hat{a}_1^t, \ldots, \hat{a}_d^t]$.

C. Energy Update Model

With limited battery capacity, it is challenging to support long-term services without battery recharging at the devices. To tackle this issue, we consider that the devices in FL systems are powered by the energy gathered from ambient energy sources (e.g., thermoelectric, electromagnetic radiation). Specifically, by denoting device k's energy level at the beginning of the t-th communication round as $E_{b,k}^t$, we have

$$E_{b,k}^t \le E_{\max},\ \forall\, k \in \mathcal{K}, \forall\ t \in \mathcal{T}, \tag{13}$$

where $E_{\max}$ denotes the maximum battery capacity. We denote the total amount of energy arrived at device k in the t-th round as $E_{a,k}^t$. Considering inevitable energy loss, the actual energy stored at the end of the t-th round, denoted as $E_{s,k}^t$, should not be greater than the arrived energy (i.e., $E_{a,k}^t$), i.e., $E_{s,k}^t \le E_{a,k}^t$. As in [29], we consider the scenario that all harvested energy is utilized to support the local training and uplink gradient transmission. Thus, the dynamics of the battery level at device k over two adjacent communication rounds is updated as follows

$$E_{b,k}^{t+1} = E_{b,k}^t - E_{c,k}^t - E_{l,k}^t + E_{s,k}^t,\ \forall\, k \in \mathcal{K}, \forall\ t \in \mathcal{T}, \tag{14}$$

[2]The symbol-level synchronization at the BS can be guaranteed by employing the timing advance (TA) technique in 5G new radio [27] which adjusts the transmission time of each device according to the received TA command via referring to the common clock [28].

where $E_{c,k}^t = dP_k^t\tau^t$ represents the energy consumption of device k for the local gradient transmission, τ^t denotes the transmission duration in the t-th round, and $E_{l,k}^t$ is the local training energy consumption, which is mainly determined by the data size, number of CPU cycles required for computing one sample, and CPU frequency [30], [31]. We assume that all EH-enabled devices follow the harvest-store-use rule [32], which means that the energy harvested in each round is stored and can only be used in later communication rounds. Thus, the energy consumption of device k in the t-th round should satisfy

$$E_{c,k}^t + E_{l,k}^t \le E_{b,k}^t,\ \forall\, k \in \mathcal{S}^t, \forall\ t \in \mathcal{T}, \tag{15}$$

and the stored energy $E_{s,k}^t$ in each communication round is thus given by

$$E_{s,k}^t = \min\{\rho_c E_{a,k}^t, E_{\max} - (E_{b,k}^t - E_{c,k}^t - E_{l,k}^t), \quad E_{c,\max}\},\ \forall\, k \in \mathcal{K}, \forall\ t \in \mathcal{T}, \tag{16}$$

where $E_{c,\max}$ is the maximum charging energy in each round and $\rho_c \in (0,1]$ denotes the charging efficiency. It is observed that the stored energy in each round (i.e., $E_{s,k}^t$) depends on the amount of arrived energy (i.e., $E_{a,k}^t$), the remaining battery capacity (i.e., $E_{\max}-(E_{b,k}^t-dP_k^t\tau^t-E_{l,k}^t)$), and the maximum charging energy in each round (i.e., $E_{c,\max}$) [33].

III. Convergence Analysis and Problem Formulation

Analyzing the convergence behavior of the proposed over-the-air FL framework is essential for characterizing the impact of device scheduling and transceiver design on the learning performance. In this section, we first derive an explicit expression of the upper bound of the difference between the expected value of global loss function and its optima after T communication rounds, and then formulate a non-convex mixed-integer resource allocation problem for enhancing the ultimate learning performance.

A. Convergence Analysis

To facilitate the convergence analysis, we make several common assumptions as follows.

Assumption 1 (Smoothness)**.** *The global loss function $F(\cdot)$ is L-smooth, i.e.,*

$$F(\boldsymbol{w}') \le F(\boldsymbol{w}) + \nabla F(\boldsymbol{w})^T(\boldsymbol{w}'-\boldsymbol{w}) + \frac{L}{2}\|\boldsymbol{w}'-\boldsymbol{w}\|^2, \quad \forall\, \boldsymbol{w}, \boldsymbol{w}'. \tag{17}$$

Assumption 2 (Bounded Gradient)**.** *For any model parameter vector $\boldsymbol{w} \in \mathbb{R}^d$ and data pair $(\boldsymbol{x}, y) \in \mathbb{R}^d \times \mathbb{R}$, there is a constant $\kappa \ge 0$ satisfying*

$$\|\nabla f_i(\boldsymbol{w})\|^2 \le \kappa. \tag{18}$$

Assumption 3 (μ-Polyak-Łojasiewicz (PL) Condition [34])**.** *For any differentiable function $F(\cdot) : \mathbb{R}^d \to \mathbb{R}$, there is a constant $\mu > 0$ satisfying*

$$\|\nabla F(\boldsymbol{w})\|^2 \ge 2\mu\big(F(\boldsymbol{w}) - F(\boldsymbol{w}^\star)\big), \forall\, \boldsymbol{w} \in \mathbb{R}^d. \tag{19}$$

It is noteworthy that Assumption 3 is a relaxed version of the strong convexity. Here we make this assumption because the objective functions in many practical applications only satisfy the μ-PL condition, rather than being strongly convex. Typical examples include least-squares and logistic regression. According to the above assumptions, we derive the propositions as follows.

Proposition 1. *With the set of scheduled devices $\mathcal{S}^t$ and learning rate $0 < \gamma^t \equiv \gamma < \frac{1}{L}$, the gap of $F(\cdot)$ across two adjacent communication rounds is bounded as*

$$\mathbb{E}\left[F\left(\boldsymbol{w}^{t+1}\right)\right] - \mathbb{E}\left[F\left(\boldsymbol{w}^{t}\right)\right] \le -\frac{\gamma}{2}\mathbb{E}\left[\left\|\nabla F\left(\boldsymbol{w}^t\right)\right\|^2\right] + 4\gamma^3\kappa\Big(1-\frac{|\mathcal{S}^t|}{K}\Big)^2 + \gamma\sum_{j=1}^{d}\mathsf{MSE}^t,\ \forall\ t \in \mathcal{T}, \tag{20}$$

where $\mathbb{E}[\cdot]$ refers to an expectation taken over the receiver noise and transmit symbols. Besides, MSE^t denotes the mean-squared-error of the estimated gradient $\hat{g}_j^t$ with respect to the ground-truth gradient g_j^t, and is defined as

$$\begin{aligned}\mathsf{MSE}^t &= \mathbb{E}\Big[|\hat{g}_j^t - g_j^t|^2\Big] \\ &= \frac{1}{|\mathcal{S}^t|}\sum_{k\in\mathcal{S}^t}\mathbb{E}\left[\left|\left[(\boldsymbol{m}^t)^{\mathsf{H}}\boldsymbol{h}_k^t b_k^t - \zeta_k^t\right]s_{k,j}^t\right|^2\right] + \frac{\sigma_n^2}{|\mathcal{S}^t|}\|\boldsymbol{m}^t\|^2 \\ &= \frac{1}{|\mathcal{S}^t|}\sum_{k\in\mathcal{S}^t}\left|(\boldsymbol{m}^t)^{\mathsf{H}}\boldsymbol{h}_k^t b_k^t - \zeta_k^t\right|^2 + \frac{\sigma_n^2}{|\mathcal{S}^t|}\|\boldsymbol{m}^t\|^2.\end{aligned} \tag{21}$$

Proof. See Appendix A. □

Remark 1. *The first term in the right-hand-side of (20) denotes the expected norm of the ideal global gradient, while the second and third terms account for the errors due to partial device participation and transmission distortion. In particular, the convergence rate of the AirComp-assisted FL has the potential to be accelerated by scheduling as many devices as possible and reducing the model aggregation error. Achieving such a goal is non-trivial due to the following reasons. First, there generally exists a tradeoff between scheduling more devices and achieving a lower model aggregation error. Second, both device scheduling and aggregation distortion depend on the time-varying battery energy level. Insufficient battery may lead to severe gradient distortion or even less scheduled devices. To accelerate the convergence of FL, the device energy should be carefully managed by jointly designing device scheduling, transmit power, and receive beamforming.*

With Proposition 1, the expected performance gap to the optima in terms of the global loss function $F(\cdot)$ after T communication rounds is presented in the following proposition.

Proposition 2. *With any device scheduling set $\{\mathcal{S}^t, t = 1, \ldots, T\}$, the gap of global loss functions $F(\cdot)$ between the optima and that after T rounds can be bounded as*

$$\mathbb{E}\left[F\left(\boldsymbol{w}^T\right)\right] - \mathbb{E}\left[F\left(\boldsymbol{w}^\star\right)\right] \le \Delta^T\Big(\mathbb{E}\left[F\left(\boldsymbol{w}^0\right)\right] - \mathbb{E}\left[F\left(\boldsymbol{w}^\star\right)\right]\Big) + \sum_{t=0}^{T-1}\Delta^{T-1-t}\Big[4\gamma\kappa\Big(1-\frac{|\mathcal{S}^t|}{K}\Big)^2 + \gamma d\mathsf{MSE}^t\Big], \tag{22}$$

where $\Delta = 1-\gamma\mu < 1$ *and* $\boldsymbol{w}^0$ *is the initial model parameter vector.*

Proof. See Appendix B. □

Remark 2. *According to Proposition 2, the FL performance after* T *rounds is determined by the initial gap to the optima, partial device participation, as well as the aggregation error. The adverse effect of the initial gap can be alleviated as the number of communication rounds increases. According to (22), the weight of the second term (i.e.,* Δ^{T-1-t}*) has a negative exponential relationship with the index of communication rounds (i.e.,* t*), which indicates that the earlier communication rounds are "less important" learning stages than the latter communication rounds. That is because the errors caused by partial device participation and gradient transmission distortion affect the ultimate learning performance sightly when* t *is small, while the performance is more vulnerable to these errors in the latter rounds. Thus, it is necessary to jointly optimize device scheduling and transceiver design across all communication rounds to effectively allocate limited resources and achieve a better FL performance, rather than treating each round equally and focusing on the per-round performance.*

B. Problem Formulation

We aim to minimize the expected global loss after T communication rounds, while taking into account the battery operation constraints and the maximum power budget. Mathematically, the formulated optimization problem is as follows

$$\mathscr{P}_1: \underset{\boldsymbol{m}^t,\{b_k^t\},\mathcal{S}^t}{\text{minimize}} \quad \mathbb{E}\{F(\boldsymbol{w}^T)\} \tag{23a}$$

$$\text{subject to} \quad (10),(14),(15). \tag{23b}$$

Remark 3. *Equation (16) is only a formula to calculate the stored energy* $E_{s,k}^t$ *that is determined by the arrived energy* $E_{a,k}^t$ *and the maximum battery capacity* $E_{\max}$ *, and is incorporated in (14). Thus, (16) is not explicitly incorporated as a constraint of* $\mathscr{P}_1$*. Additionally, different from [35], we consider a multi-antenna scenario, where both the transmit scalar and receive beamforming are optimized to enhance the performance of AirComp-assisted FL.*

Note that solving problem $\mathscr{P}_1$ confronts the following challenges. First, problem $\mathscr{P}_1$ is a non-convex mixed-integer programming problem, involving coupled optimization variables $\boldsymbol{m}^t$ and $\{b_k^t\}$, as well as combinatorial variable $\mathcal{S}^t$. Second, constraint (15) is limited by the battery energy level $E_{b,k}^t$ that obeys the time-related evolution in (14), resulting in the coupling between transmit power and device scheduling over different communication rounds. Moreover, the objective function of problem $\mathscr{P}_1$ is hard to be expressed in a closed form in terms of the optimization variables because of the complicated model update process. To handle these challenges, we first approximate the objective function of problem $\mathscr{P}_1$ in Section IV according to its upper bound derived in Proposition 2, and then propose a joint design of AirComp transceiver and device scheduling.

IV. Proposed Online Optimization Algorithm

We propose an online optimization algorithm that effectively solves problem $\mathscr{P}_1$ in this section. We first approximate the objective function using its upper bound, and then formulate a series of joint optimization subproblems by applying the Lyapunov optimization. Finally, we present the analytical results for the proposed online optimization algorithm.

A. Problem Transformation

By approximating the objective of problem $\mathscr{P}_1$ using the upper bound according to Proposition 2, the reformulated problem is given by

$$\mathscr{P}_2: \underset{\boldsymbol{m}^t,\{b_k^t\},\mathcal{S}^t}{\text{minimize}} \quad \sum_{t=0}^{T-1}\Delta^{T-1-t}\Big[4\gamma\kappa\Big(1-\frac{|\mathcal{S}^t|}{K}\Big)^2+\gamma d\mathsf{MSE}^t\Big]$$
$$\text{subject to} \quad (10),(14),(15). \tag{24}$$

Note that problem $\mathscr{P}_2$ falls into the category of stochastic optimization problems and has two unique challenges. Firstly, this problem involves multiple training rounds and can be tackled under the assumption that the complete knowledge of random processes $\{h_k^t\}$ and $\{E_{a,k}^t\}$ is known beforehand [36], which, however is impractical to be acquired. Moreover, although the optimal power control for AirComp-assisted FL proposed in [37] can be derived in an off-line manner, this method is only suitable for the single-antenna scenario, but cannot be generalized to its multi-antenna counterpart because the variables are highly coupled, making the corresponding transceiver design problem NP-hard [38]. The second unique challenge is induced by time-coupled constraint (14) over different communication rounds, which is difficult to be tackled since the current action will have an impact on the future decisions. To address these challenges, we resort to designing an online optimization algorithm for problem $\mathscr{P}_2$ without any prior information of future CSI and energy arrival [39]. Specifically, we apply Lyapunov optimization to solve this stochastic optimization problem $\mathscr{P}_2$ in an online manner by dividing it into multiple single-round subproblems, each of which enables variables $\{b_k^t\}$, $\boldsymbol{m}^t$, and scheduling subset $\mathcal{S}^t$ to be optimized only based on the current system state (i.e., $\{\boldsymbol{h}_k^t\}$ and $\{E_{a,k}^t\}$).

As constraints (14) and (15) involve optimization variables over two adjacent communication rounds, the Lyapunov optimization method cannot be directly applied. Thus, we first recast problem $\mathscr{P}_2$ into a standard stochastic optimization problem by relaxing constraint (15) and transforming (14) to the long-term time-averaged counterpart. In particular, by recursively summing up (14) across rounds $t \in \{0,1,\ldots,T-1\}$, the energy level in each local device over T rounds is given by

$$\mathbb{E}[E_{b,k}^T - E_{b,k}^0] = \sum_{t=0}^{T-1}\mathbb{E}[E_{s,k}^t - E_{c,k}^t - E_{l,k}^t],\ \forall\, k \in \mathcal{K}. \tag{25}$$

We divide both sides of (25) by T and take the limit $T \to \infty$, yielding the following equation

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}[E_{c,k}^t + E_{l,k}^t] = \lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}[E_{s,k}^t]. \tag{26}$$

Equation (26) reveals that the energy consumed for gradient uploading and local training is supplied by the harvested energy.

By further dividing the objective function by T, we obtain a standard stochastic optimization problem as follows

$$\mathscr{P}_3: \underset{\boldsymbol{m}^t,\{b_k^t\},\mathcal{S}^t}{\text{minimize}} \quad \frac{1}{T}\sum_{t=0}^{T-1}\Delta^{T-1-t}\Big[4\gamma\kappa\Big(1-\frac{|\mathcal{S}^t|}{K}\Big)^2+\gamma d\mathsf{MSE}^t\Big] \quad \text{subject to} \quad (10),(26). \tag{27}$$

This problem is a relaxation of problem $\mathscr{P}_2$ and can be solved by Lyapunov optimization in an online manner. In Section IV-B, we will show that the solution obtained by the proposed algorithm is feasible to problem $\mathscr{P}_2$.

B. Lyapunov Online Optimization

The long-term minimization problem can be transformed into an online optimization problem by applying Lyapunov optimization [40]. Specifically, we introduce a virtual queue Q_k^t to record the state of the battery level at each device k. The queuing dynamics of Q_k^t is updated as

$$Q_k^{t+1}=Q_k^t-E_{c,k}^t-E_{l,k}^t+E_{s,k}^t, \tag{28}$$

which incorporates the dynamics of the energy level at device k. Note that the value of virtual queue Q_k^t can be negative. Q_k^t is evolved based on the energy update of the t-th round in (26). Preserving the stability of Q_k^t is identical to satisfying constraint (26).

To measure the congestion of queue Q_k^t, the Lyapunov function is defined as

$$L(\boldsymbol{Q}^t)=\frac{1}{2}\sum_{k\in\mathcal{K}}(Q_k^t)^2, \tag{29}$$

where $\boldsymbol{Q}^t=(Q_1^t,\ldots,Q_K^t)$ is the vector of virtual queues. The corresponding single-round Lyapunov drift conditioned on $\boldsymbol{Q}^t$ is

$$\Delta(\boldsymbol{Q}^t)=\mathbb{E}[L(\boldsymbol{Q}^{t+1})-L(\boldsymbol{Q}^t)|\boldsymbol{Q}^t], \tag{30}$$

where the expectation is taken over the random system state (i.e., $\{\boldsymbol{h}_k^t\}$ and $\{E_{a,k}^t\}$). According to Lyapunov optimization [40], our objective is to force the Lyapunov function (29) into a low congestion region so as to stabilize queue Q_k^t when solving problem $\mathscr{P}_3$. Therefore, we minimize the following drift-plus-penalty metric

$$\Delta(\boldsymbol{Q}^t)+V\mathbb{E}\Big[\Delta^{T-1-t}\Big[4\gamma\kappa\Big(1-\frac{|\mathcal{S}^t|}{K}\Big)^2+\gamma d\mathsf{MSE}^t\Big]\Big|\boldsymbol{Q}^t\Big], \tag{31}$$

where $V>0$ is an adjustable parameter that represents how much we emphasize on minimizing the objective of problem $\mathscr{P}_3$.

According to (30) and the dynamics of the virtual queue, the difference of the Lyapunov function between two adjacent rounds is

$$\begin{aligned}
&L(\boldsymbol{Q}^{t+1})-L(\boldsymbol{Q}^t)\\
&=\frac{1}{2}\sum_{k\in\mathcal{S}^t}\big[(Q_k^{t+1})^2-(Q_k^t)^2\big]\\
&=\sum_{k\in\mathcal{S}^t}[Q_k^t(E_{s,k}^t-E_{c,k}^t-E_{l,k}^t)\frac{1}{2}(E_{s,k}^t-E_{c,k}^t-E_{l,k}^t)^2]\\
&\le B+\sum_{k\in\mathcal{S}^t}[Q_k^t(E_{s,k}^t-E_{c,k}^t-E_{l,k}^t)],
\end{aligned} \tag{32}$$

where $B=\sum_{k\in\mathcal{S}^t}\Big(\max\{E_{c,\max},E_{l,k}^t+d\tau P_k^{\max}\}\Big)^2/2$, and the inequality holds because the harvested energy $E_{s,k}^t$ in each communication round cannot exceed the maximum charging rate $E_{c,\max}$ and each device is limited by its maximum power $P_k^{\max}$. Hence, we have

$$(E_{s,k}^t-E_{c,k}^t-E_{l,k}^t)^2\le\Big(\max\big\{E_{l,k}^t+d\tau^t P_k^{\max},E_{c,\max}\big\}\Big)^2. \tag{33}$$

Accordingly, for any control policy with $V\ge 0$, the drift-plus-penalty metric (31) is upper-bounded as follows

$$\begin{aligned}
&\Delta(\boldsymbol{Q}^t)+V\mathbb{E}\Big[\Delta^{T-1-t}\Big[4\gamma\kappa\Big(1-\frac{|\mathcal{S}^t|}{K}\Big)^2+\gamma d\mathsf{MSE}^t\Big]\Big|\boldsymbol{Q}^t\Big]\\
&\le B+\sum_{k\in\mathcal{S}^t}Q_k^t\mathbb{E}\Big[E_{s,k}^t-E_{c,k}^t-E_{l,k}^t|\boldsymbol{Q}^t\Big]+V\mathbb{E}\Big[\Delta^{T-1-t}\\
&\Big[4\gamma\kappa\Big(1-\frac{|\mathcal{S}^t|}{K}\Big)^2+\gamma d\mathsf{MSE}^t\Big]\Big|\boldsymbol{Q}^t\Big].
\end{aligned} \tag{34}$$

Following the drift-plus-penalty minimization principle, the system can be stabilized by minimizing the bound on $\Delta(\boldsymbol{Q}^t)$ in every round. Thus, we attempt to optimize the right hand side of (34) in each communication round. By omitting the constant terms (i.e.,$B+\sum_{k\in\mathcal{S}^t}Q_k^t(E_{s,k}^t-E_{l,k}^t)$), the resulting online sub-problem is given by

$$\begin{aligned}
\underset{\mathcal{S}^t,\{b_k^t\},\boldsymbol{m}^t}{\text{minimize}} \quad & V\Delta^{T-1-t}\Big[4\gamma\kappa\Big(1-\frac{|\mathcal{S}^t|}{K}\Big)^2+\gamma d\mathsf{MSE}^t\Big]\\
&-\sum_{k\in\mathcal{S}^t}Q_k^t\tau^t d|b_k^t|^2\\
\text{subject to} \quad & (10).
\end{aligned} \tag{35}$$

However, problem (35) remains challenging to be solved since it involves integer variable (i.e., $\mathcal{S}^t$) and is a mixed integer non-linear programming problem.

C. Theoretical Performance Analysis

We now analyze the performance of the proposed Lyapunov online optimization algorithm. The objective value of problem $\mathscr{P}_2$ in the t-th round is denoted as C_t and there exists a minimum value $C_t^{\min}$ of C_t without considering any constraint. Let $\sum_{t=0}^{T-1}C_t^{\dagger}$ be the objective value of $\mathscr{P}_2$ based on the optimal solutions of online problems (35) and $\sum_{t=0}^{T-1}C_t^{\text{opt}}$ stands for the objective value of $\mathscr{P}_2$ based on the optimal offline solutions. The performance analysis is given as follows.

Proposition 3. *By denoting $C^{\min}$ as the minimum among $\{C_t^{\min}\}_{t=1}^T$ and $C^{\max}$ as the maximum among $\{C_t^{\text{opt}}\}_{t=1}^T$, we have the following statements:*

a) *Queues $\{Q_k^T\}$ are mean rate stable, i.e.,*

$$\lim_{T\to\infty} \frac{\mathbb{E}[|Q_k^T|]}{T} = 0, \forall\, k \in \mathcal{K}. \tag{36}$$

b) *The proposed online optimization algorithm can achieve a performance with gap in the order of $\mathcal{O}(1/V)$ to the optimal solution, i.e.,*

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} C_t^{\dagger} \le \frac{B}{V} + \lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} C_t^{\text{opt}}, \tag{37}$$

where B is defined in (32).

c) *The average queue length is bounded by*

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1}\sum_{k=1}^{K} \mathbb{E}\left[|Q_k^t|\right] \le \frac{B + V(C^{\max} - C^{\min})}{\epsilon}. \tag{38}$$

Proof. See Appendix C. □

Proposition 3(a) demonstrates that the virtual queues introduced by Lyapunov optimization can be stable. Proposition 3(b) reveals that the learning performance gap monotonously decreases with parameter V, and thus the proposed online optimization approach can achieve an asymptotically optimal solution of problem $\mathscr{P}_2$. There is a linear relationship between the time average queue backlog and V shown in Proposition 3(c), which presents a $[\mathcal{O}(1/V), \mathcal{O}(V)]$ performance-backlog tradeoff. Thus, we should appropriately set parameter V to balance the tradeoff between the queue backlog and learning performance.

In the following section, we decouple problem (35) into two subproblems with respect to the transceiver design and device scheduling.

V. Proposed Alternating Algorithm for Online Design

Although we have designed a lyapunov-based online algorithm for solving $\mathscr{P}_2$, variables are highly coupled in problem (35). This motivates us to develop an efficient alternating optimization approach to optimize variables by decoupling problem (35) into two subproblems.

A. Optimizing Transmit Scalars and Receive Beamforming Vector with Fixed $\mathcal{S}^t$

We reformulate problem (35) as

$$\begin{aligned}\underset{\{b_k^t\},\boldsymbol{m}^t}{\text{minimize}} \quad & U^t\Big[\sum_{k\in\mathcal{S}^t}\Big|(\boldsymbol{m}^t)^{\mathsf{H}}\boldsymbol{h}_k^t b_k^t - \zeta_k^t\Big|^2 + \sigma_n^2\|\boldsymbol{m}^t\|^2\Big] \\ & - \sum_{k\in\mathcal{S}^t} Q_k^t\tau^t d|b_k^t|^2 \\ \text{subject to} \quad & (10),\end{aligned} \tag{39}$$

where $U^t = V\Delta^{T-1-t}\gamma d/|\mathcal{S}^t|^2$. However, this problem remains non-convex because of the coupled $\{b_k^t\}$ and $\boldsymbol{m}^t$. We develop an alternating optimization algorithm to tackle this issue.

1) Optimizing $\{b_k^t\}$ with fixed $\boldsymbol{m}^t$: With fixed receive beamforming vector $\boldsymbol{m}^t$, we set the transmit scalar of each device based on maximum-ratio transmission [41]. Specifically, we set $b_k^t = \sqrt{P_k^t}\frac{(\boldsymbol{m}^{\mathsf{H}}\boldsymbol{h}_k)^{\mathsf{H}}}{|\boldsymbol{m}^{\mathsf{H}}\boldsymbol{h}_k|}$. Thus, the estimated gradient at the BS is given by

$$\hat{g}_j^t = \sum_{k\in\mathcal{S}^t} |(\boldsymbol{m}^t)^{\mathsf{H}}\boldsymbol{h}_k^t|\sqrt{P_k^t}s_{k,j}^t + (\boldsymbol{m}^t)^{\mathsf{H}}\boldsymbol{n}^t, \tag{40}$$

and problem (39) can be reformulated as

$$\begin{aligned}\underset{\{P_k^t\}}{\text{minimize}} \quad & U^t\Big[\sum_{k\in\mathcal{S}^t} P_k^t(\boldsymbol{m}^t)^{\mathsf{H}}\boldsymbol{H}_k^t\boldsymbol{m}^t - 2\sum_{k\in\mathcal{S}^t}|(\boldsymbol{m}^t)^{\mathsf{H}}\boldsymbol{h}_k^t|\zeta_k^t \\ & \sqrt{P_k^t} + \sigma_n^2\|\boldsymbol{m}^t\|^2 + |\mathcal{S}^t|(\zeta_k^t)^2\Big] - \sum_{k\in\mathcal{S}^t} Q_k^t\tau^t d P_k^t \\ \text{subject to} \quad & (10),\end{aligned} \tag{41}$$

where $\boldsymbol{H}_k^t = \boldsymbol{h}_k^t(\boldsymbol{h}_k^t)^{\mathsf{H}}$.

Problem (41) can be divided into $|\mathcal{S}^t|$ subproblems, each of which corresponds to the transmit power optimization of a single device. Since the convexity of each subproblem is determined by the value of virtual queue Q_k^t and constraint (10) is linear in P_k^t, we derive the optimal transmit power $(P_k^t)^\star$ of each device k as a function of Q_k^t as follows.

Proposition 4. *The optimal transmit power $(P_k^t)^\star$ of each device k in the t-th round for problem (41) is given in (42), shown on the top of this page.*

Proof. See Appendix D. □

In each round t, the transmit power P_k^t only relies on the values of $\boldsymbol{h}_k^t$ and $E_{a,k}^t$. After obtaining P_k^t based on (42), Q_k^t can be updated based on (28).

Note that problem $\mathscr{P}_3$ is relaxed from problem $\mathscr{P}_2$ and we should ensure that the transmit power derived from (42) is also feasible to problem $\mathscr{P}_2$. Hence, after obtaining the solution based on (42), the solution that violates (15) is updated as follows

$$P_k^t = \left[\frac{E_{b,k}^t - \frac{1}{t}\sum_{j=1}^t (E_{b,k}^j - E_{c,k}^j - E_{l,k}^j) - E_{l,k}^t}{\tau^t d}\right]^+, \tag{43}$$

where $[x]^+ \triangleq \max\{x, 0\}$. As (43) satisfies (15) and (26), the solution in (43) is feasible to both problems $\mathscr{P}_2$ and $\mathscr{P}_3$. In addition, (43) also reveals that the transmit power P_k^t is determined by the historical time-averaged battery energy level.

2) Optimizing $\boldsymbol{m}^t$ with fixed $\{b_k^t\}$: With fixed transmit scalar $\{b_k^t\}$, problem (39) is reformulated as an unconstrained convex optimization problem with respect to $\boldsymbol{m}^t$. In each round t, the optimal receive beamforming vector, denoted as $(\boldsymbol{m}^t)^\star$, is

$$(\boldsymbol{m}^t)^\star = \Big(\sum_{k\in\mathcal{S}^t}|b_k^t|^2\boldsymbol{H}_k^t + \sigma_n^2\boldsymbol{I}_M\Big)^{-1}\sum_{k\in\mathcal{S}^t}\boldsymbol{h}_k^t b_k^t\zeta_k^t, \tag{44}$$

which is derived by calculating its first-order optimal condition with respect to $\boldsymbol{m}^t$.

$$(P_k^t)^\star = \begin{cases} \left(\frac{U^t |\boldsymbol{h}_k^{\mathsf{H}} \boldsymbol{m}| \zeta_k^t}{U^t \boldsymbol{m}^{\mathsf{H}} \boldsymbol{H}_k \boldsymbol{m} - Q_k^t \tau^t d} \right)^2, & \text{if } Q_k^t < \frac{\sqrt{P_k^{\max}} U^t \boldsymbol{m}^{\mathsf{H}} \boldsymbol{H}_k \boldsymbol{m} - U^t |\boldsymbol{h}_k^{\mathsf{H}} \boldsymbol{m}| \zeta_k^t}{\sqrt{P_k^{\max}} \tau^t d}, \\ P_k^{\max}, & \text{if } Q_k^t \geq \frac{\sqrt{P_k^{\max}} U^t \boldsymbol{m}^{\mathsf{H}} \boldsymbol{H}_k \boldsymbol{m} - U^t |\boldsymbol{h}_k^{\mathsf{H}} \boldsymbol{m}| \zeta_k^t}{\sqrt{P_k^{\max}} \tau^t d}. \end{cases} \tag{42}$$

Algorithm 1: Online optimization algorithm for beamforming variables and device scheduling.

Input: Initial arbitrary $\mathcal{S}^0$, $\boldsymbol{m}^0$, $\{b_k, \forall k\}^0$, T, $J_{\max}$ and $\boldsymbol{s}^0 = [1, \ldots, 1]$
for $t = 1, 2, \ldots, T$ **do**
  $\boldsymbol{s}^{t,0} \leftarrow \boldsymbol{s}^{t-1}$
  for $j = 1, 2, \ldots, J_{\max}$ **do**
    Generate sampling set $\mathcal{S}^{t,j}$
    foreach $\boldsymbol{s}^{t,j} \in \mathcal{S}^{t,j}$ **do**
      Compute $J(\boldsymbol{s}^{t,j})$
    end
    Generate sample $\boldsymbol{s}^{t,j}$ based on distribution (45)
  end
  $\boldsymbol{s}^t \leftarrow \boldsymbol{s}^{t,R_s}$, $\mathcal{S}^t \leftarrow \boldsymbol{s}^t$
  Compute $\{P_k, \forall k\}$ according to (42) with given $\mathcal{S}^t$
  if P_k^t *violates* (15) **then**
    Compute P_k according to (43)
  end
  Update $\{b_k, \forall k\}$ according to $b_k^t = \sqrt{P_k^t}(\boldsymbol{m}^{\mathsf{H}} \boldsymbol{h}_k)^H / |\boldsymbol{m}^{\mathsf{H}} \boldsymbol{h}_k|$
  Update $\boldsymbol{m}^t$ according to (44) with given $\mathcal{S}^t$
end
Output: $\mathcal{S}^t$, $\{b_k^t\}$, $\boldsymbol{m}^t$

B. Device Scheduling via Gibbs Sampling

In each communication round, we adopt the Gibbs sampling to approximate the distribution of scheduling variables. Specifically, we use a binary indicator vector $\boldsymbol{s}^t$ to represent the device selection set $\mathcal{S}^t$, i.e., $\boldsymbol{s}^t = [s_1^t, s_2^t, \ldots, s_K^t]^{\mathsf{T}}$ with the k-th entry $s_k^t = 1$ meaning that the k-th device is selected and $s_k^t = 0$ otherwise. We obtain $\boldsymbol{s}^t$ from a series of samples $\{\boldsymbol{s}^{t,j}\}$, where $\boldsymbol{s}^{t,j}, \forall\, j \in \{1, \ldots, J_{\max}\}$, denotes the j-th sample and $J_{\max}$ is the maximum number of samples. In particular, $\boldsymbol{s}^{t,j}$ is sampled from set $\mathcal{S}^{t,j} = \{\boldsymbol{s}^{t,j-1}\} \cup \{\boldsymbol{s}_{(i)}^{t,j-1}\}$, based on the following distribution

$$\pi(\boldsymbol{s}^{t,j}) = \frac{\exp(-J(\boldsymbol{s}^{t,j})/\beta)}{\sum_{i=1}^K \exp(-J(\boldsymbol{s}_{(i)}^{t,j-1})/\beta) + \exp(-J(\boldsymbol{s}^{t,j-1})/\beta)}, \tag{45}$$

where $\boldsymbol{s}_{(i)}^{t,j-1}$ denotes the indicator vector that differs from $\boldsymbol{s}^{t,j-1}$ only at the i-th element, $J(\boldsymbol{s}^{t,j})$ is the objective value of problem (35) with respect to scheduling vector $\boldsymbol{s}^{t,j}$, transceiver variables $\boldsymbol{m}^t$ and $\{b_k^t\}$, and $\beta > 0$ is a "temperature parameter" to accelerate the convergence. We iteratively sample $\{\boldsymbol{s}^{t,j}\}$ until either the results stabilize or the number of iterations reaches $J_{\max}$ in each communication round. Then, we update the scheduling vector as $\boldsymbol{s}^t = \boldsymbol{s}^{t,j^\star}$, where $j^\star$ is the last sample in $\{\boldsymbol{s}^{t,j}\}$. The proposed online optimization algorithm is presented in Algorithm 1.

C. Computation Complexity

In each communication round, the worst-case complexity for computing $\{b_k^t\}$ and $\boldsymbol{m}^t$ with given scheduled device set $\mathcal{S}^t$ in Algorithm 1 is $\mathcal{O}(M^3)$. Since the Gibbs sampling method involves the computation for the objective value of (35) $J_{\max} K$ times, the computation complexity of the proposed Lyapunov online optimization method in each communication round is $\mathcal{O}(J_{\max} K M^3)$.

For the baseline, we utilize a semidefinite relaxation [42] to solve problem (48). The worst-case computation complexity for solving each semidefinite programming problem is $\mathcal{O}(\max\{K, M\}^4\ M^{1/2} \log(1/\epsilon))$, where $\epsilon > 0$ is the given solution precision. Thus, the overall computation complexity is $\mathcal{O}(\max\{K, M\}^4 J_{\max} M^{3/2} \log(1/\epsilon))$, which is much greater than the proposed algorithm.

VI. Simulation Results

In this section, extensive simulations are conducted to illustrate the superiority of the proposed over-the-air FL framework and the online optimization algorithm.

A. Simulation Setup

1) Learning model setting: The proposed online optimization algorithm is evaluated by implementing the image classification tasks on the MNIST and CIFAR-10 datasets under independent and identically distributed configurations. The multinomial logistic regression is leveraged in this paper to train the models with sample-wise loss functions defined as follows

$$f_i(\boldsymbol{w}) = -\sum_{c=1}^{C} \mathbb{I}\{y_i = c\} \log \left(\frac{\exp\left(\boldsymbol{w}_c^\top \boldsymbol{x}_i\right)}{\sum_{j=1}^C \exp\left(\boldsymbol{w}_j^\top \boldsymbol{x}_i\right)} \right), \tag{46}$$

where C represents the total number of different label categories in datasets, and $\boldsymbol{w}_c$ denotes the model parameter vector associated with label category $c \in \{1, \ldots, C\}$. The entire model parameter vector is defined as $\boldsymbol{w} = [\boldsymbol{w}_1^{\mathsf{T}}, \boldsymbol{w}_2^{\mathsf{T}}, \ldots, \boldsymbol{w}_C^{\mathsf{T}}]^{\mathsf{T}}$. Besides, for each $\boldsymbol{w}_c$, the partial gradient is computed as

$$\nabla f_i(\boldsymbol{w}_c) = -\left(\mathbb{I}_{\{y_i = c\}} - \frac{\exp\left(\boldsymbol{w}_c^\top \boldsymbol{x}_i\right)}{\sum_{j=1}^C \exp\left(\boldsymbol{w}_j^\top \boldsymbol{x}_i\right)} \right) \boldsymbol{x}_i, \tag{47}$$

and the entire gradient is computed as $\nabla f_i(\boldsymbol{w}) = [\nabla f_i(\boldsymbol{w}_1)^{\mathsf{T}}, \ldots, \nabla f_i(\boldsymbol{w}_C)^{\mathsf{T}}]^{\mathsf{T}}$. The training dataset of MNIST (or CIFAR-10) contains 60000 samples with 784 features and 10 labels (or 50000 samples with 3072 features and 10 labels), and is randomly and uniformly partitioned into K disjoint subsets. We set the learning rate γ as 0.05 for the MNIST dataset and 0.15 for the CIFAR-10 dataset.

2) Communication model setting: All devices are randomly placed in a circular area with radius $R = 25$ meters, which is centered around the BS. Each link from a device to the BS suffers from quasi-static Rician fading. We denote the path loss as $L(d) = L_0(d/d_0)^{-\alpha}$, where L_0 is the path loss at distance $d_0 = 1$ meter, d denotes the link distance, and α refers to the path loss exponent.

The energy blocks arrive at each device following the Poisson process with rate $\lambda = 2$ and the energy of each block obeys the uniform distribution within $[0, 1]$ J. Besides, the maximum charging energy in each round is set to $E_{c,\max} = 2$ J. The default initial battery energy level at each device is 2 J. The transmission duration τ^t is set to 0.01 second. The energy consumption for local training $E^t_{l,k}$ is set according to [43]. Unless specified otherwise, we set other parameters as $E_{\max} = 10$ J, $P_k^{\max} = 200$ mW, $L_0 = -30$ dB, $\sigma^2 = -90$ dBm, $\rho_c = 0.9$, and $\alpha = 2.2$.

B. Performance Comparison

We consider four benchmark schemes to validate the superiority of the proposed online algorithm. The benchmarks are described below:

- **Optimal aggregation**: All devices have sufficient battery energy to successfully perform the local training and transmit local gradients to the edge server through error-free channels, while the edge server aggregates these error-free gradients as in (6) to update the global model.
- **Greedy MSE minimization**: In each communication round, edge devices are scheduled according to the proposed Gibbs sampling method, i.e., devices with better channel conditions and more energy budget are more likely to be scheduled. Each selected device utilizes all of its currently available energy for local model training and transmission to minimize the MSE by solving the following optimization problem

$$\begin{aligned} \underset{\boldsymbol{m}^t, \{b_k^t\}}{\text{minimize}} \quad & \mathsf{MSE}^t \\ \text{subject to} \quad & P_k^t \le \min\left\{E_{b,k}^t/(d\tau^t), P_k^{\max}\right\}, \forall\, k \in \mathcal{S}^t. \end{aligned} \tag{48}$$

- **Myopic power allocation**: In each communication round, edge devices are scheduled according to the proposed Gibbs sampling method and the maximum energy budget of each selected devices is the currently available energy averaged by the number of remaining rounds, i.e., $\frac{E_{b,k}^t}{T-t+1}$.
- **Greedy scheduling**: With fixed transmit power and receive beamforming vector, problem (35) is reformulated as

$$\begin{aligned} \underset{\mathcal{S}^t}{\text{minimize}}\ & V\Delta^{T-1-t}\Big[4\gamma\kappa\Big(1-\frac{|\mathcal{S}^t|}{K}\Big)^2 + \frac{\gamma d}{|\mathcal{S}^t|^2}\sum_{k\in\mathcal{S}^t} \\ & \Big|(\boldsymbol{m}^t)^{\mathsf{H}}\boldsymbol{h}_k^t b_k^t - 1\Big|^2\Big] - \sum_{k\in\mathcal{S}^t} Q_k^t \tau^t d P_k. \end{aligned} \tag{49}$$

By sorting $\mathcal{G}^t = \{U^t|(\boldsymbol{m}^t)^{\mathsf{H}}\boldsymbol{h}_k^t b_k^t - 1|^2 - Q_k^t\tau^t d P_k, \forall\, k\}$ in an ascending order, we represent the $k-$th smallest value of $\mathcal{G}^t$ as G_k^t. Then the objective function in (49) can take at most K different values as

$$\mathcal{R}^t \triangleq \Big\{4V\Delta^{T-1-t}\gamma\kappa\Big(1-\frac{k}{K}\Big)^2 + \sum_{n=1}^{k} G_k^t, \forall\, k \in \mathcal{K}\Big\}. \tag{50}$$

Thus, the scheduling is performed on the k^* devices with the smallest $U^t|(\boldsymbol{m}^t)^{\mathsf{H}}\boldsymbol{h}_k^t b_k^t - 1|^2 - Q_k^t\tau^t d P_k$, where the number of scheduling devices in the greedy policy (i.e., k^*) is the value of k corresponding to the minimum element in $\mathcal{R}^t$. Besides, we still utilize the same transceiver design as in our proposed method.
- **Random scheduling**: In each communication round, edge devices participating in the gradients uploading are randomly scheduled (i.e., both the number and the indices of devices are randomly determined).

In Fig. 2, we implement an image classification task to compare the test accuracy of the proposed algorithm with the benchmarks when $V = 1$, $K = 5$, and $N = 4$ on MNIST dataset. The proposed algorithm attains a learning performance close to the optimal aggregation and outperforms other benchmarks. This is because the harvested energy is fully utilized through appropriate transceiver design of the transmit power and device scheduling that accounts for the long-term energy constraint. In the greedy MSE minimization scheme, all remaining energy is utilized in each communication round, which may lead to temporal pause of training because of insufficient energy left for the training and transmission in the following rounds. This coincides with its stepwise growth of the test accuracy in Fig. 2. Since the energy may also be wasted in the rounds with deep fading channels, its performance is much worse than that of the proposed algorithm. Besides, although the myopic power allocation scheme that averages the remaining energy for the following round can ensure the continuity of the whole training process, such a short-sighted power allocation scheme makes the uploaded local models more vulnerable to the channel fading and receiver noise than other schemes, thus degrading the overall training performance. With the same setting of simulation parameters, we also compare the test accuracy of the proposed algorithm with that of the benckmarks on the CIFAR-10 dataset. As shown in Fig. 3, similar performance trends can be observed and the proposed algorithm still outperforms other benchmarks except optimal aggregation.

Fig. 4 illustrates the performance of the proposed algorithm and other device scheduling schemes when $V = 1$, $K = 5$, and $N = 4$. Note that the proposed device scheduling approach based on Gibbs sampling achieves a greater test accuracy than the greedy scheduling counterpart throughout the entire training process. This is because the second term of (49) is related to the MSE of gradient and it may vanish with the proper setting of optimization variables $\boldsymbol{m}$ and $\{P_k^t\}$. By increasing the number of participating devices, the value of the remaining terms in (49) monotonically decreases, which results in a situation that all devices are scheduled in most cases and causes energy to be wasted in deep channel fading. Therefore,

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2023.3339298

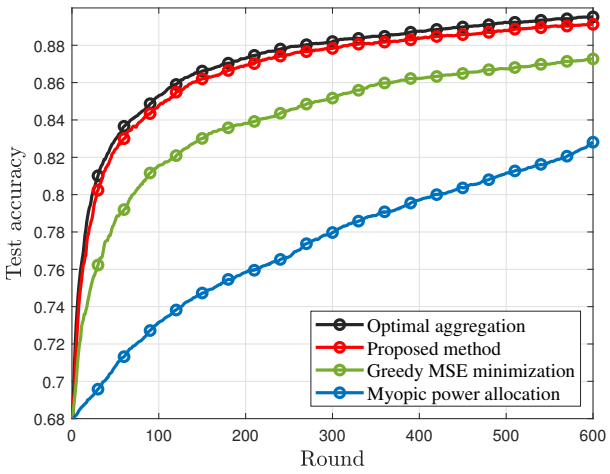


Fig. 2. Test accuracy versus the number of training rounds with different transmission strategies on MNIST dataset.

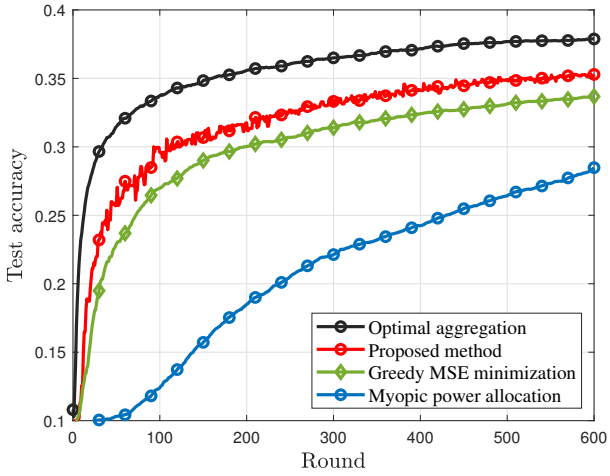


Fig. 3. Test accuracy versus the number of training rounds with different transmission strategies on CIFAR-10 dataset.

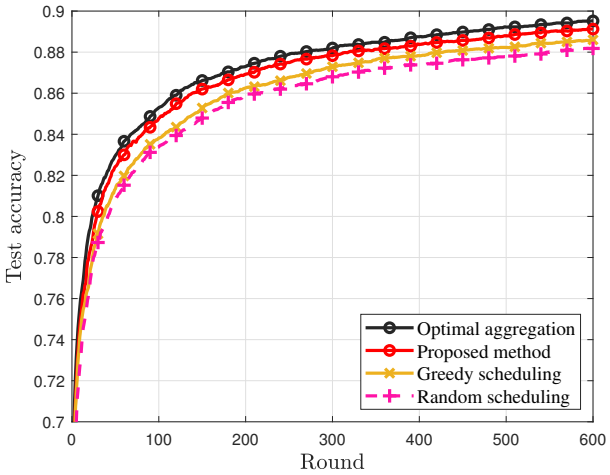


Fig. 4. Test accuracy versus the number of training rounds with different scheduling strategies on MNIST dataset.

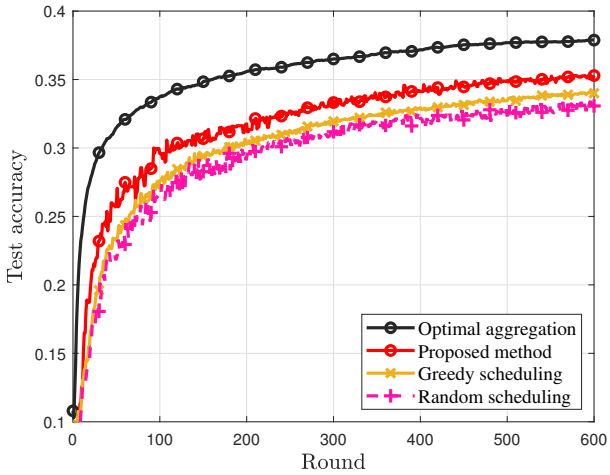


Fig. 5. Test accuracy versus the number of training rounds with different scheduling strategies on CIFAR-10 dataset.

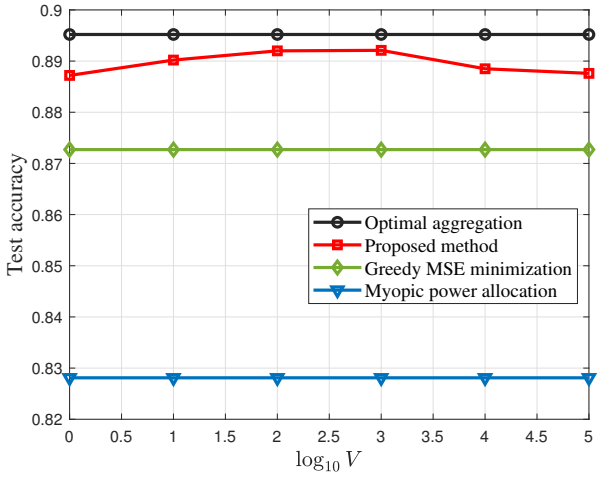


Fig. 6. Impact of weight parameter V on test accuracy.

the proposed method that can reserve energy for later favorable propagation conditions achieves a better performance than the greedy scheduling scheme and maintains a near-optimal performance. Fig. 5 illustrates similar performance trends on the CIFAR-10 dataset.

To evaluate the performance of the proposed algorithm more comprehensively, we have compared the proposed method with other benchmarks with different number of devices on the MNIST dataset, as shown in Table I. All experiments are conducted on the same global dataset with a fixed size, and the more devices there are, the less data each device holds. It can be observed that the proposed method outperforms all other benchmarks under different number of devices. Besides, the test accuracy decreases monotonically with the increase in the number of edge devices. With a fixed global dataset, the growth in the number of participating edge devices leads to severer signal distortion, thus leading to larger gradients aggregation errors and performance degradation.

Fig. 6 shows the impact of parameter V in the drift-plus-

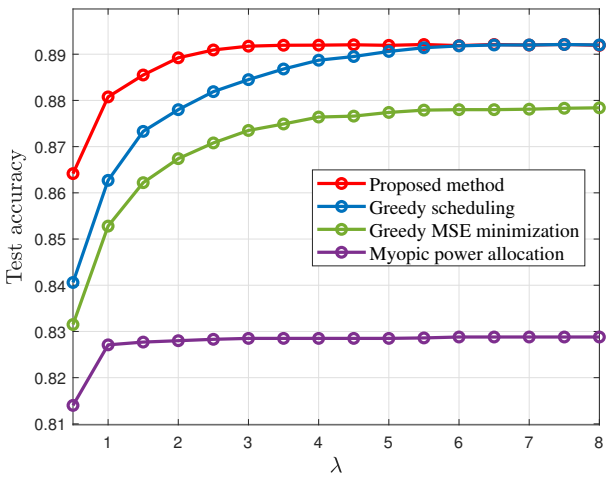


Fig. 7. Test accuracy versus the expected energy arrival rate under different device scheduling schemes on MNIST dataset.

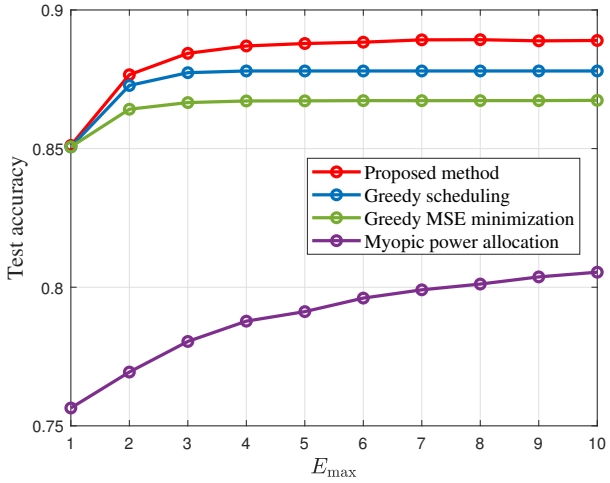


Fig. 8. Test accuracy versus the maximum battery capacity under different device scheduling schemes on MNIST dataset.

TABLE I
TEST ACCURACY VERSUS NUMBER OF EDGE DEVICES ON MNIST DATASET

Benchmark	$K = 5$	$K = 10$	$K = 50$	$K = 50$
Proposed method	86.97%	86.83%	86.33%	84.34%
Greedy scheduling	85.23%	85.13%	84.44%	82.44%
Greedy MSE minimization	83.82%	83.67%	83.60%	83.59%
Myopic power allocation	80.62%	80.60%	80.59%	80.55%

penalty minimization problem on the training performance when $K = 5$, $N = 4$, and $T = 600$. Note that the test accuracy refers to the classification accuracy achieved by the global model on the test dataset after T training rounds. As can be observed, there is a positive correlation between the test accuracy and the value of parameter V when $V \leq 10^3$. A larger value of V implies the algorithm puts more emphasis on the FL training task, leading to a better learning performance, and the achieved accuracy after T training rounds becomes stable after an obvious increment with V. The larger the value of V, the smaller performance gap from the optimal solution in (37). Furthermore, other benchmarks are not related to V, and thus their test performance remains the same. In addition, we find that there is a performance degradation when $V \geq 10^4$. This is because the average queue length increases with V monotonically according to Proposition 3(c), which makes these queues difficult to be stable within limited training rounds, thereby violating the original constraints and degrading the learning performance. Thus, we should set an appropriate value of V to achieve a near-optimal performance with an acceptable buffer size.

In Fig. 7, we show the test accuracy versus the expected energy arrival rate on the MNIST dataset when $V = 1$, $K = 5$, $M = 4$, $E_{\max} = 20$ J, and $T = 600$. The achieved test accuracy after T training rounds monotonically increases with energy arrival rate λ in all schemes. This is because, with more harvested energy, more devices can be scheduled to participate in gradient aggregation and higher transmit powers can be used to mitigate the transmission distortion. Compared with the greedy MSE minimization scheme that directly minimizes the MSE in each round, our proposed Lyapunov-based optimization algorithm shows a considerable performance gain, which verifies that the derived theoretical performance gap in (22) is a better performance metric and the long-term energy allocation is effective. Furthermore, the proposed algorithm outperforms the greedy scheduling counterpart when the expected energy arrival rate λ is less than 6, which demonstrates that our device scheduling scheme can make full use of the limited energy. If some devices have unfavorable channel conditions, they are not allowed to transmit and the energy is saved for later gradients uploading. When λ is larger than 6, the battery level gradually becomes saturated and each device has sufficient budget to overcome the signal distortion. Therefore, with the increase of the expected energy arrival rate λ, almost all devices are scheduled, and the performance gain brought by device scheduling will gradually diminish.

Fig. 8 shows the achieved test accuracy after $T = 600$ training rounds versus the maximum battery capacity on the MNIST dataset when $V = 1$, $K = 5$, and $M = 4$. The test accuracy increases monotonically with the increase in battery capacity for all schemes. However, when $E_{\max}$ is grater than 5 J, the performance of other schemes remains stable, except for the myopic power allocation scheme. This is because the battery is easy to be saturated and the learning performance is mainly constrained by the maximum battery capacity when $E_{\max}$ is less than 5 J. A larger battery capacity means more energy can be used for model training and transmission. However, when $E_{\max}$ is greater than 5 J, the battery may not be fully charged with high probabilities and the learning performance mainly depends on the stored energy $E_{s,k}^t$ and will not be affected by the maximum battery capacity. For the myopic power allocation scheme, the energy budget in each round is evenly distributed, and the majority of energy is saved for subsequent rounds. Thus, its training performance shows

a positive correlation with the battery capacity.

VII. Conclusions

In this paper, we propose an AirComp and EH empowered communication- and energy-efficient FL framework, where AirComp and EH facilitate low-latency gradient aggregation and alleviate battery limitation, respectively. The theoretical analysis of the proposed FL framework reveals that the convergence performance can be enhanced by mitigating the aggregation error due to partial device participation and model distortion. We also observe that the weights of aggregation errors differ in different rounds, which follows the "latter-is-better" principle. Based on the convergence analysis, we formulate a long-term stochastic optimization problem and propose an effective Lyapunov-based online optimization algorithm that only relies on the current energy arrival and CSI. Simulation results demonstrate that our proposed algorithm outperforms the benchmarks by appropriately scheduling devices and allocating energy across different communication rounds.

Appendix

A. Proof of Proposition 1

With local gradient in (5), the global model aggregation via AirComp is

$$\begin{aligned}\boldsymbol{w}^{t+1} &= \boldsymbol{w}^t - \gamma^t \hat{\boldsymbol{g}}^t \\ &= \boldsymbol{w}^t - \gamma^t \Bigg[\frac{\sum_{k\in\mathcal{K}} D_k \boldsymbol{g}_k^t}{\sum_{k\in\mathcal{K}} D_k} + \\ &\quad \underbrace{\left(\hat{\boldsymbol{g}}^t - \boldsymbol{g}^t\right)}_{\boldsymbol{e}_{\text{cmm}}^t} + \underbrace{\left(\boldsymbol{g}^t - \frac{\sum_{k\in\mathcal{K}} D_k \boldsymbol{g}_k^t}{\sum_{k\in\mathcal{K}} D_k}\right)}_{\boldsymbol{e}_{\text{sel}}^t} \Bigg] \\ &= \boldsymbol{w}^t - \gamma^t \left(\frac{\sum_{k\in\mathcal{K}} D_k \boldsymbol{g}_k^t}{\sum_{k\in\mathcal{K}} D_k} + \boldsymbol{e}_{\text{cmm}}^t + \boldsymbol{e}_{\text{sel}}^t \right) \\ &= \boldsymbol{w}^t - \gamma^t \Big(\nabla F(\boldsymbol{w}^t) + \boldsymbol{e}^t \Big),\end{aligned} \tag{51}$$

where $\boldsymbol{e}_{\text{sel}}^t$ and $\boldsymbol{e}_{\text{cmm}}^t$ refer to the model aggregation errors induced by device scheduling and communication distortion, respectively, and the total error is $\boldsymbol{e}^t = \boldsymbol{e}_{\text{cmm}}^t + \boldsymbol{e}_{\text{sel}}^t$. Specifically, $\boldsymbol{e}_{\text{sel}}^t$ can be computed as

$$\boldsymbol{e}_{\text{sel}}^t = \boldsymbol{g}^t - \nabla F(\boldsymbol{w}^t) = \frac{1}{|\mathcal{S}^t|} \sum_{k\in\mathcal{S}^t} \boldsymbol{g}_k^t - \frac{1}{K} \sum_{k\in\mathcal{K}} \boldsymbol{g}_k^t, \tag{52}$$

where $\boldsymbol{g}^t = \frac{1}{|\mathcal{S}^t|} \sum_{k\in\mathcal{S}^t} \boldsymbol{g}_k^t$ denotes the ideal aggregated gradients from scheduled devices. In addition, $\boldsymbol{e}_{\text{cmm}}^t$ is given by

$$\begin{aligned}\boldsymbol{e}_{\text{cmm}}^t &= \hat{\boldsymbol{g}}^t - \boldsymbol{g}^t \\ &= \frac{1}{|\mathcal{S}^t|} \left[\sum_{k\in\mathcal{S}^t} (\boldsymbol{m}^t)^{\mathsf{H}} \boldsymbol{h}_k^t b_k^t \boldsymbol{s}_k^t + (\boldsymbol{m}^t)^{\mathsf{H}} \boldsymbol{n}^t \right] \\ &\quad - \frac{1}{|\mathcal{S}^t|} \sum_{k\in\mathcal{S}^t} \boldsymbol{g}_k^t + \frac{1}{|\mathcal{S}^t|} \sum_{k\in\mathcal{S}^t} \bar{g}_k \mathbf{1} \\ &= \frac{1}{|\mathcal{S}^t|} \sum_{k\in\mathcal{S}^t} \Big[(\boldsymbol{m}^t)^{\mathsf{H}} \boldsymbol{h}_k^t b_k^t - \zeta_k \Big] \boldsymbol{s}_k^t + \frac{1}{|\mathcal{S}^t|} (\boldsymbol{m}^t)^{\mathsf{H}} \boldsymbol{n}^t.\end{aligned} \tag{53}$$

According to (51) and Assumption 1, we have

$$\begin{aligned}&F\left(\boldsymbol{w}^{t+1}\right) - F\left(\boldsymbol{w}^t\right) \\ &\le \left(\frac{\gamma^2 L}{2} - \gamma\right) \left\|\nabla F\left(\boldsymbol{w}^t\right)\right\|^2 + \left(\gamma^2 L - \gamma\right) \\ &\quad \left\langle \nabla F\left(\boldsymbol{w}^t\right), \boldsymbol{e}^t \right\rangle + \frac{\gamma^2 L}{2} \left\|\boldsymbol{e}^t\right\|^2 \\ &\overset{\text{(I)}}{\le} -\frac{\gamma}{2} \left\|\nabla F\left(\boldsymbol{w}^t\right)\right\|^2 + \frac{\gamma}{2} \left\|\boldsymbol{e}^t\right\|^2 \\ &\overset{\text{(II)}}{\le} -\frac{\gamma}{2} \left\|\nabla F\left(\boldsymbol{w}^t\right)\right\|^2 + \gamma \left(\left\|\boldsymbol{e}_{\text{cmm}}^t\right\|^2 + \left\|\boldsymbol{e}_{\text{sel}}^t\right\|^2 \right),\end{aligned} \tag{54}$$

where inequalities (I) and (II) are based on $-\left\langle \nabla F\left(\boldsymbol{w}^t\right), \boldsymbol{e}^t \right\rangle \le \left(\left\|\nabla F\left(\boldsymbol{w}^t\right)\right\|^2 + \left\|\boldsymbol{e}^t\right\|^2\right)/2$ and $\left\|\boldsymbol{e}^t\right\|^2 \le 2\left(\left\|\boldsymbol{e}_{\text{cmm}}^t\right\|^2 + \left\|\boldsymbol{e}_{\text{sel}}^t\right\|^2 \right)$, respectively.

According to Assumption 2, the norm of the device selection error is upper bounded by $\left\|\boldsymbol{e}_{\text{sel}}^{[t]}\right\|^2 \le 4\gamma^2\kappa \left(1 - |\mathcal{S}^t|/K\right)^2$. Thus, we have

$$\begin{aligned}&\mathbb{E}\left[F\left(\boldsymbol{w}^{t+1}\right)\right] - \mathbb{E}\left[F\left(\boldsymbol{w}^t\right)\right] \\ &\le -\frac{\gamma}{2} \mathbb{E}\left[\left\|\nabla F\left(\boldsymbol{w}^t\right)\right\|^2\right] + 4\gamma^3\kappa \left(1 - \frac{|\mathcal{S}^t|}{K}\right)^2 \\ &\quad + \gamma \mathbb{E}\left[\left\|\hat{\boldsymbol{g}}^t - \boldsymbol{g}^t\right\|^2\right] \\ &= -\frac{\gamma}{2} \mathbb{E}\left[\left\|\nabla F\left(\boldsymbol{w}^t\right)\right\|^2\right] + 4\gamma^3\kappa \left(1 - \frac{|\mathcal{S}^t|}{K}\right)^2 \\ &\quad + \gamma \sum_{j=1}^{d} \mathsf{MSE}^t.\end{aligned} \tag{55}$$

B. Proof of Proposition 2

Based on Assumption 3, inequality (55) can be further derived as

$$\begin{aligned}\mathbb{E}\left[F\left(\boldsymbol{w}^{t+1}\right)\right] - \mathbb{E}\left[F\left(\boldsymbol{w}^t\right)\right] &\le -\gamma\mu \Big(\mathbb{E}\left[F\left(\boldsymbol{w}^t\right)\right] - \\ \mathbb{E}\left[F\left(\boldsymbol{w}^\star\right)\right] \Big) &+ 4\gamma^3\kappa \left(1 - \frac{|\mathcal{S}^t|}{K}\right)^2 + \gamma d \mathsf{MSE}^t.\end{aligned} \tag{56}$$

By subtracting $\mathbb{E}\left[F\left(\boldsymbol{w}^\star\right)\right]$ on both sides of (56), we have

$$\begin{aligned}\mathbb{E}\left[F\left(\boldsymbol{w}^{t+1}\right)\right] - \mathbb{E}\left[F\left(\boldsymbol{w}^\star\right)\right] &\le (1 - \gamma\mu) \Big(\mathbb{E}\left[F\left(\boldsymbol{w}^t\right)\right] \\ -\mathbb{E}\left[F\left(\boldsymbol{w}^\star\right)\right] \Big) &+ 4\gamma^3\kappa \left(1 - \frac{|\mathcal{S}^t|}{K}\right)^2 + \gamma d \mathsf{MSE}^t.\end{aligned} \tag{57}$$

Recursively applying (57) for $t \in \{0, 1, \ldots, T-1\}$ and letting $\Delta = (1 - \varsigma\mu)$, we have

$$\begin{aligned}\mathbb{E}\left[F(\boldsymbol{w}^T)\right] - \mathbb{E}\left[F(\boldsymbol{w}^\star)\right] &\le \Delta^T \Big(\mathbb{E}\left[F(\boldsymbol{w}^0)\right] - \mathbb{E}\left[F(\boldsymbol{w}^\star)\right] \Big) \\ + \sum_{t=0}^{T-1} \Delta^{T-1-t} \Big[4\gamma^3\kappa \Big(1 - \frac{|\mathcal{S}^t|}{K}\Big)^2 + \gamma d \mathsf{MSE}^t \Big].\end{aligned} \tag{58}$$

C. Proof of Proposition 3

Based on the upper bound of drift-plus-penalty metric (34), we have

$$\begin{aligned}\Delta(\boldsymbol{Q}^t) + V C_t^\dagger &\le B + \sum_{k\in\mathcal{S}^t} Q_k^t \mathbb{E}\Big[E_{s,k}^t - d P_k^t \tau^t - E_{l,k}^t | \boldsymbol{Q}^t \Big] \\ + V C_t^\dagger &\overset{\text{(I)}}{=} B + V C_t^\dagger \overset{\text{(II)}}{\le} B + V C_t^{opt},\end{aligned} \tag{59}$$

where (I) follows according to Lemma 5 in [44] and (II) follows because problem $\mathscr{P}_3$ is relaxed from problem $\mathscr{P}_2$.

Based on (59) and $C_t^{\min} \leq C_t^{\dagger}$, we have

$$\Delta(\boldsymbol{Q}^t) \leq B + V(C_t^{opt} - C_t^{\min}), \tag{60}$$

and by summing up the both sides of above inequality over $\mathcal{T}$, we obtain

$$\mathbb{E}\left[L(\boldsymbol{Q}^T)\right] \leq TB + \mathbb{E}\left[L(\boldsymbol{Q}^0)\right] + V\sum_{t=0}^{T-1}(C_t^{opt} - C_t^{\min}), \tag{61}$$

when we have

$$\mathbb{E}\left[(Q_k^T)^2\right] \leq 2TB + 2\mathbb{E}\left[L(\boldsymbol{Q}^0)\right] + 2V\sum_{t=0}^{T-1}(C_t^{opt} - C_t^{\min}). \tag{62}$$

Since $\mathbb{E}\left[|Q_k^T|\right]^2 \leq \mathbb{E}\left[(Q_k^T)^2\right]$, we can further have

$$\mathbb{E}\left[|Q_k^T|\right] \leq \sqrt{2TB + 2\mathbb{E}\left[L(\boldsymbol{Q}^0)\right] + 2V\sum_{t=0}^{T-1}(C_t^{opt} - C_t^{\min})}. \tag{63}$$

Dividing both sides of the above inequality by T and taking a limits as $T \to \infty$, we prove that

$$\begin{aligned}&\lim_{T\to\infty}\frac{1}{T}\mathbb{E}\left[|Q_k^T|\right]\\ &\leq \lim_{T\to\infty}\sqrt{\frac{2B}{T} + \frac{2\mathbb{E}\left[L(\boldsymbol{Q}^0)\right]}{T^2} + \frac{2V}{T^2}\sum_{t=0}^{T-1}(C_t^{opt} - C_t^{\min})}\\ &= 0.\end{aligned} \tag{64}$$

Thus, the virtual queues $\{Q_k^T\}$ are mean rate stable and Proposition 3(a) is proved.

Summing up both sides of inequality (59) over $\mathcal{T}$, we have

$$\sum_{t=0}^{T-1}\Delta(\boldsymbol{Q}^t) + V\sum_{t=0}^{T-1}C_t^{\dagger} \leq TB + V\sum_{t=0}^{T-1}C_t^{opt}, \tag{65}$$

and

$$\begin{aligned}V\sum_{t=0}^{T-1}C_t^{\dagger} &\leq TB + V\sum_{t=0}^{T-1}C_t^{opt} + \mathbb{E}\left[L(\boldsymbol{Q}^0)\right] - \mathbb{E}\left[L(\boldsymbol{Q}^T)\right]\\ &\leq TB + V\sum_{t=0}^{T-1}C_t^{opt} + \mathbb{E}\left[L(\boldsymbol{Q}^0)\right].\end{aligned} \tag{66}$$

Dividing the both sides of the above inequality by VT and taking the limit as T goes to infinity, we have

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}C_t^{\dagger} \leq \frac{B}{V} + \lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}C_t^{opt}, \tag{67}$$

where the inequality holds under the assumption that $\mathbb{E}\left[L(\boldsymbol{Q}^0)\right] < \infty$. The proof of Proposition 3(b) is completed.

According to the Theorem 4.2 in [40], we have

$$\Delta(\boldsymbol{Q}^t) + VC_t^{\min} \leq B + VC_t^{\text{opt}} - \epsilon\sum_{k=1}^{K}\mathbb{E}\left[|Q_k^t|\right]. \tag{68}$$

Since $\{\mathsf{MSE}^t\}_{t=1}^T$ are upper-bounded, $\{C_t^{\text{opt}}\}_{t=1}^T$ are also upper-bounded. Thus, there always exists a constant $C^{\max}$ such that $C^{\max} = \max\{C_t^{\text{opt}}\}$. By applying the law of telescoping sums over $\mathcal{T}$, we have

$$\begin{aligned}\epsilon\sum_{t=0}^{T-1}\sum_{k=1}^{K}\mathbb{E}\left[|Q_k^t|\right] &\leq T[B + V(\max\{C_t^{\text{opt}}\}\\ &\quad - \min\{C_t^{\min}\})] + \mathbb{E}\left[L(\boldsymbol{Q}^0)\right] - \mathbb{E}\left[L(\boldsymbol{Q}^T)\right].\end{aligned} \tag{69}$$

Dividing (69) by ϵT and taking a limit as $T \to \infty$, we finally obtain

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\sum_{k=1}^{K}\mathbb{E}\left[|Q_k^t|\right] \leq \frac{B + V(C^{\max} - C^{\min})}{\epsilon}. \tag{70}$$

The proof of Proposition 3(c) is completed.

D. Proof of Proposition 4

When $U^t\boldsymbol{m}^{\mathsf{H}}\boldsymbol{H}_k\boldsymbol{m} + Q_k^t\tau^t d \geq 0$, problem (41) is convex with respect to $\sqrt{P_k}$ of device $k \in \mathcal{S}^t$. We consider the following two cases:

(i) If $\frac{U^t|\boldsymbol{m}^{\mathsf{H}}\boldsymbol{h}_k|\zeta_k^t}{U^t\boldsymbol{m}^{\mathsf{H}}\boldsymbol{H}_k\boldsymbol{m} + Q_k^t\tau^t d} \geq \sqrt{P_k^{\max}}$, then the objective function of problem (41) is monotonic within constraint (10). Thus, the optimal solution is $\sqrt{P_k^{\max}}$, i.e., $(P_k^t)^\star = P_k^{\max}$.

(ii) If $\frac{U^t|\boldsymbol{m}^{\mathsf{H}}\boldsymbol{h}_k|\zeta_k^t}{U^t\boldsymbol{m}^{\mathsf{H}}\boldsymbol{H}_k\boldsymbol{m} + Q_k^t\tau^t d} \leq \sqrt{P_k^{\max}}$, then the minimum value of problem (41) is given by $(P_k^t)^\star = \left(\frac{U^t|\boldsymbol{h}_k^{\mathsf{H}}\boldsymbol{m}|\zeta_k^t}{U^t\boldsymbol{m}^{\mathsf{H}}\boldsymbol{H}_k\boldsymbol{m} + Q_k^t\tau^t d}\right)^2$.

When $U^t\boldsymbol{m}^{\mathsf{H}}\boldsymbol{H}_k\boldsymbol{m} + Q_k^t\tau^t d \leq 0$, problem (41) is concave with respect to $\sqrt{P_k}$ and the objective function is also monotonically decreasing within constraint (10). Hence, the optimal solution is $\sqrt{P_k^{\max}}$, i.e., $(P_k^t)^\star = P_k^{\max}$.

In summary, the transmit power of each device k is shown in (42).

REFERENCES

[1] K. B. Letaief, Y. Shi, J. Lu, and J. Lu, "Edge artificial intelligence for 6G: Vision, enabling technologies, and applications," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 1, pp. 5–36, 2022.

[2] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: Concept and applications," vol. 10, no. 2, pp. 1–19, 2019.

[3] S. Wang, T. Tuor, T. Salonidis, K. K. Leung, C. Makaya, T. He, and K. Chan, "Adaptive federated learning in resource constrained edge computing systems," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 6, pp. 1205–1221, 2019.

[4] M. M. Amiri, D. Gündüz, S. R. Kulkarni, and H. V. Poor, "Convergence of update aware device scheduling for federated learning at the wireless edge," *IEEE Trans. Wireless Commun.*, vol. 20, no. 6, pp. 3643–3658, 2021.

[5] Z. Yang, M. Chen, W. Saad, C. S. Hong, and M. Shikh-Bahaei, "Energy efficient federated learning over wireless communication networks," *IEEE Trans. Wireless Commun.*, vol. 20, no. 3, pp. 1935–1949, 2020.

[6] M. Chen, Z. Yang, W. Saad, C. Yin, H. V. Poor, and S. Cui, "A joint learning and communications framework for federated learning over wireless networks," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 269–283, 2021.

[7] T. T. Vu, D. T. Ngo, N. H. Tran, H. Q. Ngo, M. N. Dao, and R. H. Middleton, "Cell-free massive mimo for wireless federated learning," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6377–6392, 2020.

[8] Y. Zou, Z. Wang, X. Chen, H. Zhou, and Y. Zhou, "Knowledge-guided learning for transceiver design in over-the-air federated learning," *IEEE Transactions on Wireless Communications*, vol. 22, no. 1, pp. 270–285, Jan. 2023.

[9] G. Zhu, Y. Wang, and K. Huang, "Broadband analog aggregation for low-latency federated edge learning," *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 491–506, 2020.

[10] M. M. Amiri and D. Gündüz, "Federated learning over wireless fading channels," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 3546–3557, 2020.

[11] K. Yang, T. Jiang, Y. Shi, and Z. Ding, "Federated learning via over-the-air computation," *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 2022–2035, 2020.

[12] Z. Wang, J. Qiu, Y. Zhou, Y. Shi, L. Fu, W. Chen, and K. B. Letaief, "Federated learning via intelligent reflecting surface," *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 808–822, 2021.

[13] Z. Wang, Y. Zhou, Y. Zou, Q. An, Y. Shi, and M. Bennis, "A graph neural network learning approach to optimize RIS-assisted federated learning," *IEEE Trans. Wireless Commun.*, to appear.

[14] X. Cao, G. Zhu, J. Xu, Z. Wang, and S. Cui, "Optimized power control design for over-the-air federated edge learning," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 1, pp. 342–358, 2022.

[15] Z. Wang, Y. Zhou, Y. Shi, and W. Zhuang, "Interference management for over-the-air federated learning in multi-cell wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 8, pp. 2361–2377, 2022.

[16] Y. Shao, D. Gündüz, and S. C. Liew, "Federated edge learning with misaligned over-the-air computation," *IEEE Trans. Wireless Commun.*, vol. 21, no. 6, pp. 3951–3964, 2022.

[17] L. Su and V. K. N. Lau, "Hierarchical federated learning for hybrid data partitioning across multitype sensors," *IEEE Internet Things J.*, vol. 8, no. 13, pp. 10 922–10 939, 2021.

[18] H. Xing, O. Simeone, and S. Bi, "Federated learning over wireless device-to-device networks: Algorithms and convergence analysis," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 12, pp. 3723–3741, 2021.

[19] L. Hu, Z. Wang, H. Zhu, and Y. Zhou, "RIS-assisted over-the-air federated learning in millimeter wave mimo networks," *J. Commun. Netw.*, vol. 7, no. 2, pp. 145–156, 2022.

[20] M.-L. Ku, W. Li, Y. Chen, and K. R. Liu, "Advances in energy harvesting communications: Past, present, and future challenges," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1384–1412, 2015.

[21] Q. Zeng, Y. Du, and K. Huang, "Wirelessly powered federated edge learning: Optimal tradeoffs between convergence and power transfer," *IEEE Trans. Wireless Commun.*, vol. 21, no. 1, pp. 680–695, 2022.

[22] Y. Li, Y. Wu, Y. Song, L. Qian, and W. Jia, "Dynamic user-scheduling and power allocation for swipt aided federated learning: A deep learning approach," *IEEE Trans. Mob. Comput.*, 2022.

[23] W. Wen, Y. Jia, and W. Xia, "Joint scheduling and resource allocation for federated learning in swipt-enabled micro uav swarm networks," *China Commun.*, vol. 19, no. 1, pp. 119–135, 2022.

[24] Q. V. Do, Q.-V. Pham, and W.-J. Hwang, "Deep reinforcement learning for energy-efficient federated learning in uav-enabled wireless powered networks," *IEEE Commun. Lett.*, vol. 26, no. 1, pp. 99–103, 2021.

[25] R. Hamdi, M. Chen, A. B. Said, M. Qaraqe, and H. V. Poor, "Federated learning over energy harvesting wireless networks," *IEEE Internet Things J.*, vol. 9, no. 1, pp. 92–103, 2021.

[26] D. Liu and O. Simeone, "Privacy for free: Wireless federated learning via uncoded transmission with adaptive power control," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 1, pp. 170–185, 2021.

[27] A. Mahmood, M. I. Ashraf, M. Gidlund, J. Torsner, and J. Sachs, "Time synchronization in 5G wireless edge: Requirements and solutions for critical-mtc," *IEEE Commun. Mag.*, vol. 57, no. 12, pp. 45–51, 2019.

[28] O. Abari, H. Rahul, D. Katabi, and M. Pant, "Airshare: Distributed coherent transmission made seamless," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, 2015, pp. 1742–1750.

[29] C. Qiu, Y. Hu, Y. Chen, and B. Zeng, "Lyapunov optimization for energy harvesting wireless sensor communications," *IEEE Internet Things J.*, vol. 5, no. 3, pp. 1947–1956, 2018.

[30] W. Zhang, Y. Wen, K. Guan, D. Kilper, H. Luo, and D. O. Wu, "Energy-optimal mobile cloud computing under stochastic wireless channel," *IEEE Trans. Wireless Commun.*, vol. 12, no. 9, pp. 4569–4581, 2013.

[31] Y. Wang, M. Sheng, X. Wang, L. Wang, and J. Li, "Mobile-edge computing: Partial computation offloading using dynamic voltage scaling," *IEEE Trans. Commun.*, vol. 64, no. 10, pp. 4268–4282, 2016.

[32] S. Sudevalayam and P. Kulkarni, "Energy harvesting sensor nodes: Survey and implications," *IEEE Commun. Surveys Tuts.*, vol. 13, no. 3, pp. 443–461, 2011.

[33] M. Ashraf, J. Jung, H. M. Shin, and I. Lee, "Energy efficient online power allocation for two users with energy harvesting," *IEEE Signal Process. Lett.*, vol. 26, no. 1, pp. 24–28, 2019.

[34] H. Karimi, J. Nutini, and M. Schmidt, "Linear convergence of gradient and proximal-gradient methods under the polyak-łojasiewicz condition," in *Joint European conference on machine learning and knowledge discovery in databases*. Springer, 2016, pp. 795–811.

[35] Y. Sun, S. Zhou, Z. Niu, and D. Gündüz, "Dynamic scheduling for over-the-air federated edge learning with energy constraints," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 1, pp. 227–242, 2022.

[36] D. Bertsekas, *Dynamic programming and optimal control: Volume I*. Athena Scientific, 2012, vol. 1.

[37] X. Cao, G. Zhu, J. Xu, Z. Wang, and S. Cui, "Optimized power control design for over-the-air federated edge learning," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 1, Jan. 2022.

[38] S. Jing and C. Xiao, "Transceiver beamforming for over-the-air computation in massive mimo systems," *IEEE Trans. Wireless Commun.*, vol. 22, no. 10, pp. 6978–6992, Oct. 2023.

[39] Z. Zhou, Q. Wu, and X. Chen, "Online orchestration of cross-edge service function chaining for cost-efficient edge computing," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 8, pp. 1866–1880, 2019.

[40] M. J. Neely, "Stochastic network optimization with application to communication and queueing systems," *Synth. Lect. on Commun. Netw.*, vol. 3, no. 1, pp. 1–211, 2010.

[41] D. Tse and P. Viswanath, *Fundamentals of Wireless Wommunication*. Cambridge university press, 2005.

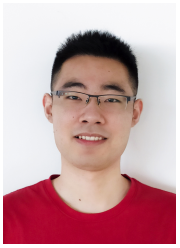
[42] Z. Luo, W. Ma, A. M. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, 2010.

[43] S. Luo, X. Chen, Q. Wu, Z. Zhou, and S. Yu, "Hfel: Joint edge association and resource allocation for cost-efficient hierarchical federated edge learning," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6535–6548, 2020.

[44] F. Amirnavaei and M. Dong, "Online power control optimization for wireless transmission with energy harvesting and storage," *IEEE Trans. Wireless Commun.*, vol. 15, no. 7, pp. 4888–4901, 2016.

Qiaochu An (Graduate Student Member, IEEE) received the B.S. degree from the School of communications engineering, Zhengzhou University, Zhengzhou, China, in 2019. He is currently pursuing the Ph.D. degree with the School of Information Science and Technology, ShanghaiTech University, Shanghai, China. His research interests include Internet of Things, intelligent reflecting surface, and federated learning.

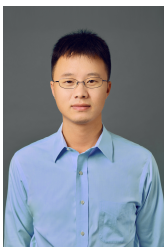
Yong Zhou (Senior Member, IEEE) received the B.Sc. and M.Eng. degrees from Shandong University, Jinan, China, in 2008 and 2011, respectively, and the Ph.D. degree from the University of Waterloo, Waterloo, ON, Canada, in 2015. From Nov. 2015 to Jan. 2018, he worked as a postdoctoral research fellow in the Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, Canada. He is currently an Assistant Professor in the School of Information Science and Technology, ShanghaiTech University, Shanghai, China. He serves as an Associate Editor of IEEE Transactions on Wireless Communications and IEEE Open Journal of the Communications Society. He was the track co-chair of IEEE VTC 2020 Fall and IEEE VTC 2023 Spring, and the general co-chair of IEEE ICC 2022 workshop on edge artificial intelligence for 6G. His research interests include 6G communications, edge intelligence, and Internet of Things.



Zhibin Wang (Graduate Student Member, IEEE) received the B.S. degree in telecommunications engineering from Xidian University, Xidian University, Xi'an, China, in 2019. He is currently pursuing the Ph.D. degree with the School of Information Science and Technology, ShanghaiTech University, Shanghai, China. His research interests include Internet of Things, intelligent reflecting surface, and federated learning.

Mehdi Bennis (IEEE Fellow) is a full (tenured) Professor at the Centre for Wireless Communications, University of Oulu, Finland, IEEE Fellow and head of the intelligent connectivity and networks/systems group (ICON). His main research interests are in radio resource management, game theory and distributed AI in 5G/6G networks. He has published more than 300 research papers in international conferences, journals and book chapters. He has been the recipient of several prestigious awards including the 2015 Fred W. Ellersick Prize from the IEEE Communications Society, the 2016 Best Tutorial Prize from the IEEE Communications Society, the 2017 EURASIP Best paper Award for the Journal of Wireless Communications and Networks, the all-University of Oulu award for research, the 2019 IEEE ComSoc Radio Communications Committee Early Achievement Award and the 2020-2023 Clarviate Highly Cited Researcher by the Web of Science.

Hangguan Shan (Senior Member, IEEE) received the B.Sc. degree in electrical engineering from Zhejiang University, Hangzhou, China, in 2004, and the Ph.D. degree in electrical engineering from Fudan University, Shanghai, China, in 2009. From 2009 to 2010, he was a Postdoctoral Research Fellow with the University of Waterloo, Waterloo, ON, Canada. Since 2011, he has been with the College of Information Science and Electronic Engineering, Zhejiang University, where he is currently an Associate Professor. He is also with the Zhejiang Provincial Key Laboratory of Information Processing and Communication Networks, Zhejiang University. His current research interests include machine learning-enabled resource allocation and quality-of-service provisioning in wireless networks. Dr. Shan has co-received the Best Industry Paper Award from the IEEE WCNC'11 and the Best Paper Award from the IEEE WCSP'23. He has served as a Technical Program Committee Member of various international conferences. He was an Editor of the IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING.



Yuanming Shi (Senior Member, IEEE) received the B.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 2011. He received the Ph.D. degree in electronic and computer engineering from The Hong Kong University of Science and Technology (HKUST), in 2015. Since September 2015, he has been with the School of Information Science and Technology in ShanghaiTech University, where he is currently a tenured Associate Professor. He visited University of California, Berkeley, CA, USA, from October 2016 to February 2017. His research areas include edge AI, federated edge learning, task-oriented communications, and satellite networks. He was a recipient of the IEEE Marconi Prize Paper Award in Wireless Communications in 2016, the Young Author Best Paper Award by the IEEE Signal Processing Society in 2016, the IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award in 2021, and the Chinese Institute of Electronics First Prize in Natural Science in 2022. He is also an editor of IEEE Transactions on Wireless Communications, IEEE Journal on Selected Areas in Communications, and Journal of Communications and Information Networks. He is an IET Fellow.