Joint Downlink-Uplink Beamforming for Wireless Multi-Antenna Federated Learning

Chong Zhang*, Min Dong†, Ben Liang*, Ali Afana‡, Yahia Ahmed‡

*Dept. of Electrical and Computer Engineering, University of Toronto, Canada, ‡Ericsson Canada, Canada
†Dept. of Electrical, Computer and Software Engineering, Ontario Tech University, Canada

Abstract—We study joint downlink-uplink beamforming design for wireless federated learning (FL) with a multi-antenna base station. Considering analog transmission over noisy channels and uplink over-the-air aggregation, we derive the global model update expression over communication rounds. We then obtain an upper bound on the expected global loss function, capturing the downlink and uplink beamforming and receiver noise effect. We propose a low-complexity joint beamforming algorithm to minimize this upper bound, which employs alternating optimization to breakdown the problem into three subproblems, each solved via closed-form gradient updates. Simulation under practical wireless system setup shows that our proposed joint beamforming design solution substantially outperforms the conventional separate-link design approach and nearly attains the performance of ideal FL with error-free communication links.

I. INTRODUCTION

Federated learning (FL) [1] is a widely recognized machine learning method to process training data locally at multiple worker nodes. In FL, a parameter server organizes the worker nodes to train a machine learning (ML) model collaboratively using their local datasets. In the wireless environment, the parameter server can be taken up by a base station (BS), which exchanges model parameters with participating devices through wireless communication [2]. However, the fluctuation of the wireless link and noisy reception bring distortion, leading to degraded FL performance. Furthermore, practical wireless systems are limited in transmission power and bandwidth. This necessitates efficient communication design to effectively support FL, which requires frequent exchange of a massive number of parameters.

Many existing works have considered improving the communication efficiency of FL over wireless channels [3]–[10]. Various digital transmission-then-aggregation schemes were proposed for uplink acquisition of local parameters from devices to the BS [3]. Such schemes use conventional digital transmission via orthogonal channels and can consume a large bandwidth and incur high latency as the number of devices becomes large. Later, analog transmission-and-aggregation schemes were proposed for the uplink [4]–[8]. These schemes use analog modulation and superposition for over-the-air aggregation of local parameters via the multiple access channel, substantially saving radio resources over the digital schemes. However, these works only focused on the uplink, while assuming an error-free downlink. Subsequently,

noisy downlink transmission for FL was studied in [9] with error-free uplink, where it was shown that since gradient descent in FL is noise resilient, analog transmission can be more efficient than digital transmission, even for the downlink.

In reality, downlink and uplink transmissions are intertwined for parameter exchange in FL. The quality of one link direction affects the other. Furthermore, the noise and distortion in one communication round propagate to all subsequent communication rounds, which brings challenges to tractable design and analysis. The literature on joint downlink-uplink communication design for FL is scarce. The convergence of FL with non-i.i.d. local datasets over noisy downlink and uplink channels was recently studied in [10], where a simple generic signal-in-noise receiver model was used to facilitate analysis without involving actual transmission modeling or design. Analog design was proposed for noisy downlink and uplink in single-cell [11] or multi-cell [12] cases. However, both works considered only single-antenna BSs, and their solutions and convergence analysis are not applicable to the more practical scenario with a multi-antenna BS. For multi-antenna communication, transmit and receive beamforming are key techniques to enhance the communication quality. While beamforming was considered in [9] for FL downlink analog transmission and in [4] for uplink over-the-air aggregation, there is no existing work on joint downlink-uplink beamforming design.

In this paper, we study joint downlink-uplink beamforming design to improve the performance of wireless FL with a multi-antenna BS. We consider noisy analog transmission in both directions and uplink over-the-air aggregation for bandwidth efficiency. We obtain the overall FL global model update over each communication round, capturing the impact of noisy downlink-uplink transmission and local model updates on FL model training. Aiming to maximize the training convergence rate, we then derive an upper bound on the expected global loss function after T rounds, and propose a low-complexity joint downlink-uplink beamforming (JDU-BF) algorithm to minimize the upper bound under transmit power constraints at the BS and devices. JDU-BF employs the alternating optimization (AO) technique to decompose the joint optimization problem into three subproblems and solve each via projected gradient descent (PGD) [13] with fast closed-form updates. Our simulation results under typical wireless network settings show that JDU-BF outperforms the conventional separate-link design and provides learning performance close to ideal FL with error-free communication links.

II. SYSTEM MODEL

A. FL System

We consider FL in a wireless network consisting of a server and K worker devices. Let $\mathcal{K} = \{1, \ldots, K\}$ denote the set of devices. Each device $k \in \mathcal{K}$ holds a local training dataset of size S_k, denoted by $\mathcal{S}_k = \{(\mathbf{s}_{k,i}, v_{k,i}) : 1 \leq i \leq S_k\}$, where $\mathbf{s}_{k,i} \in \mathbb{R}^b$ is the i-th data feature vector and $v_{k,i}$ is the label for this data sample. Using their respective local training datasets, the devices collaboratively train a global model at the server, represented by the parameter vector $\boldsymbol{\theta} \in \mathbb{R}^D$, which predicts the true labels of data feature vectors, while keeping their local datasets private. The local training loss function that represents the training error at device k is defined as

$$F_k(\boldsymbol{\theta}) = \frac{1}{S_k} \sum_{i=1}^{S_k} L(\boldsymbol{\theta}; \mathbf{s}_{k,i}, v_{k,i}) \tag{1}$$

where $L(\cdot)$ is the sample-wise training loss associated with each data sample. The global training loss function is given by the weighted sum of the local loss functions over all K devices:

$$F(\boldsymbol{\theta}) = \sum_{k=1}^{K} \frac{S_k}{S} F_k(\boldsymbol{\theta}) \tag{2}$$

where $S = \sum_{k=1}^{K} S_k$ is the total number of training samples of all devices. The learning objective is to find the optimal global model $\boldsymbol{\theta}^\star$ that minimizes $F(\boldsymbol{\theta})$.

The devices communicate with the server via noisy downlink and uplink wireless channels to exchange the model update information iteratively for model training. The iterative FL model training procedure in each downlink-uplink communication round t is given as follows:

- *Downlink broadcast*: The server broadcasts the current global model parameter vector $\boldsymbol{\theta}_t$ to all K devices via the downlink channels;
- *Local model update*: Each device k performs local training independently using its dataset $\mathcal{S}_k$, based on the received global model $\boldsymbol{\theta}_t$. In particular, the device uses the mini-batch approach to divide $\mathcal{S}_k$ into mini-batches for its local model update, where it performs J iterative local updates and generates the updated local model $\boldsymbol{\theta}_{k,t}^J$;
- *Uplink aggregation*: The devices send their updated local models $\{\boldsymbol{\theta}_{k,t}^J\}_{k \in \mathcal{K}}$ to the server via the uplink channels. The server aggregates $\boldsymbol{\theta}_{k,t}^J$'s to generate an updated global model $\boldsymbol{\theta}_{t+1}$ for the next communication round $t+1$.

B. Wireless Communication Model

We consider a practical wireless communication system where the server is hosted by a BS equipped with N antennas, and each device has a single antenna. The system operates in the time-division duplex (TDD) mode, which is typical for 5G wireless systems. With multiple antennas, the BS uses downlink beamforming to broadcast the global model update $\boldsymbol{\theta}_t$ and applies uplink receiver beamforming to process the received signal from K devices for the global model update.

For the model updating between the BS and devices in the FL system, we consider analog communication for transmitting the updated global/local models. Specifically, the BS and devices send the respective values of $\boldsymbol{\theta}_t$ and $\{\boldsymbol{\theta}_{k,t}^J\}_{k \in \mathcal{K}}$ directly under their transmit power budgets. Furthermore, for the uplink aggregation of the local models, to efficiently use the communication bandwidth, we consider over-the-air computation via analog aggregation over the multiple access channel. Specifically, the devices send their local model $\boldsymbol{\theta}_{k,t}^J$'s to the BS simultaneously over the same frequency resources, and $\boldsymbol{\theta}_{k,t}^J$'s are aggregated over the air and received at the BS. Note that the control and signaling channels of the system are still communicated using digital transmissions and are assumed to be perfect.

Due to the noisy communication channels, the received model updates over downlink and uplink are the distorted noisy versions of $\boldsymbol{\theta}_t$ and $\{\boldsymbol{\theta}_{k,t}^J\}_{k \in \mathcal{K}}$, respectively. The errors in the model updates further propagate over subsequent communication rounds for FL model training, degrading the learning performance. In this paper, we focus on the communication aspect of FL model training. Specifically, our goal is to jointly design downlink and uplink beamforming to maximize the learning performance of FL over wireless transmissions.

III. DOWNLINK-UPLINK TRANSMISSIONS FOR FL

We now formulate the transmission and reception process with downlink and uplink beamforming for the FL model update in one communication round. As mentioned in Section II, each communication round involves three steps. We present each step in detail below.

A. Downlink Broadcast

At the start of round t, the BS has the current global model, denoted by $\boldsymbol{\theta}_t = [\theta_{1,t}, \ldots, \theta_{D,t}]^T$. For efficient transmission, we convert $\boldsymbol{\theta}_t$ into a complex signal vector, whose real and imaginary parts contain half of the elements in $\boldsymbol{\theta}_t$. Specifically, we can re-express $\boldsymbol{\theta}_t = [(\tilde{\boldsymbol{\theta}}_t^{\text{re}})^T, (\tilde{\boldsymbol{\theta}}_t^{\text{im}})^T]^T$, where $\tilde{\boldsymbol{\theta}}_t^{\text{re}} \triangleq [\theta_{1,t}, \ldots, \theta_{\frac{D}{2},t}]^T$, and $\tilde{\boldsymbol{\theta}}_t^{\text{im}} \triangleq [\theta_{\frac{D}{2}+1,t}, \ldots, \theta_{D,t}]^T$. Let $\tilde{\boldsymbol{\theta}}_t$ denote the equivalent complex vector representation of $\boldsymbol{\theta}_t$, which is given by $\tilde{\boldsymbol{\theta}}_t = \tilde{\boldsymbol{\theta}}_t^{\text{re}} + j\tilde{\boldsymbol{\theta}}_t^{\text{im}} \in \mathbb{C}^{\frac{D}{2}}$.

For a TDD system, channel reciprocity holds for downlink and uplink channels. Thus, let $\mathbf{h}_{k,t} \in \mathbb{C}^N$ denote the channel vector between the BS and device $k \in \mathcal{K}$ for both downlink and uplink transmissions in round t. We assume $\{\mathbf{h}_{k,t}\}_{k \in \mathcal{K}}$ remain unchanged during round t and are known perfectly at the BS and the respective devices. The BS sends the complex global model parameter vector $\tilde{\boldsymbol{\theta}}_t$ to the K devices via multicast beamforming. At round t, the received signal vector at device k is given by

$$\mathbf{u}_{k,t} = (\mathbf{w}_t^{\text{dl}})^H \mathbf{h}_{k,t} \tilde{\boldsymbol{\theta}}_t + \mathbf{n}_{k,t}^{\text{dl}}$$

where $\mathbf{w}_t^{\text{dl}} \in \mathbb{C}^N$ is the downlink multicast beamforming vector at round t, $\mathbf{n}_{k,t}^{\text{dl}} \in \mathbb{C}^{\frac{D}{2}}$ is the receiver additive white Gaussian noise (AWGN) vector with i.i.d. elements that are zero mean with variance σ_{d}^2. The beamforming vector is subject to the BS transmit power budget. Let DP^{dl} be the

transmit power budget at the BS for sending the global model $\tilde{\boldsymbol{\theta}}_t$ in D channel uses, where P^{dl} denotes the average transmit power limit per channel use. Then, for transmitting $\tilde{\boldsymbol{\theta}}_t$, $\mathbf{w}_t^{\text{dl}}$ is subject to the transmit power constraint $\|\mathbf{w}_t^{\text{dl}}\|^2 \|\tilde{\boldsymbol{\theta}}_t\|^2 \leq DP^{\text{dl}}$. The BS also sends the scaling factor $\frac{\mathbf{h}_{k,t}^H \mathbf{w}_t^{\text{dl}}}{|\mathbf{h}_{k,t}^H \mathbf{w}_t^{\text{dl}}|^2}$ to device k via the downlink signaling channel to facilitate the receiver processing. Device k post-processes the received signal $\mathbf{u}_{k,t}$ using the received scaling factor and obtains

$$\hat{\tilde{\boldsymbol{\theta}}}_{k,t} = \frac{\mathbf{h}_{k,t}^H \mathbf{w}_t^{\text{dl}}}{|\mathbf{h}_{k,t}^H \mathbf{w}_t^{\text{dl}}|^2} \mathbf{u}_{k,t} = \tilde{\boldsymbol{\theta}}_t + \tilde{\mathbf{n}}_{k,t}^{\text{dl}} \tag{3}$$

where $\tilde{\mathbf{n}}_{k,t}^{\text{dl}} \triangleq \frac{\mathbf{h}_{k,t}^H \mathbf{w}_t^{\text{dl}}}{|\mathbf{h}_{k,t}^H \mathbf{w}_t^{\text{dl}}|^2} \mathbf{n}_{k,t}^{\text{dl}}$ is the post-processed noise vector at device k. By the equivalence of real and complex signal representations between $\boldsymbol{\theta}_t$ and $\tilde{\boldsymbol{\theta}}_t$, device k obtains the estimate of the global model $\boldsymbol{\theta}_t$, denoted by $\hat{\boldsymbol{\theta}}_{k,t}$, given by

$$\hat{\boldsymbol{\theta}}_{k,t} = \left[\mathfrak{Re}\{\hat{\tilde{\boldsymbol{\theta}}}_{k,t}\}^T, \mathfrak{Im}\{\hat{\tilde{\boldsymbol{\theta}}}_{k,t}\}^T \right]^T = \boldsymbol{\theta}_t + \hat{\mathbf{n}}_{k,t}^{\text{dl}} \tag{4}$$

where $\hat{\mathbf{n}}_{k,t}^{\text{dl}} \triangleq [\mathfrak{Re}\{\tilde{\mathbf{n}}_{k,t}^{\text{dl}}\}^T, \mathfrak{Im}\{\tilde{\mathbf{n}}_{k,t}^{\text{dl}}\}^T]^T$.

B. Local Model Update

Based on $\hat{\boldsymbol{\theta}}_{k,t}$ in (4), device k performs local model training. We assume each device adopts the mini-batch stochastic gradient descent (SGD) algorithm to minimize the local training loss function $F_k(\boldsymbol{\theta})$ [14]. The mini-batch SGD is a widely adopted training method for ML tasks. It uses a subset of the training dataset to compute the gradient update at each iteration and achieves a favorable tradeoff between computational efficiency and convergence rate. In particular, assume that each device applies J mini-batch SGD iterations for its local model update in each communication round. Let $\boldsymbol{\theta}_{k,t}^\tau$ be the local model update by device k at iteration $\tau \in \{0, \ldots, J-1\}$, with $\boldsymbol{\theta}_{k,t}^0 = \hat{\boldsymbol{\theta}}_{k,t}$, and let $\mathcal{B}_{k,t}^\tau$ denote the mini-batch, *i.e.,* a subset of $\mathcal{S}_k$, at iteration $\tau + 1$. Then, the local model update is given by

$$\boldsymbol{\theta}_{k,t}^{\tau+1} = \boldsymbol{\theta}_{k,t}^\tau - \eta_t \nabla F_k(\boldsymbol{\theta}_{k,t}^\tau; \mathcal{B}_{k,t}^\tau)$$
$$= \boldsymbol{\theta}_{k,t}^\tau - \frac{\eta_t}{|\mathcal{B}_{k,t}^\tau|} \sum_{(\mathbf{s},v) \in \mathcal{B}_{k,t}^\tau} \nabla L(\boldsymbol{\theta}_{k,t}^\tau; \mathbf{s}, v) \tag{5}$$

where η_t is the learning rate at communication round t, and ∇F_k and ∇L are the gradient functions w.r.t. $\boldsymbol{\theta}_{k,t}^\tau$. After J iterations, device k obtains the updated local model $\boldsymbol{\theta}_{k,t}^J$.

C. Uplink Aggregation

The devices send their updated local models $\{\boldsymbol{\theta}_{k,t}^J\}_{k \in \mathcal{K}}$ to the BS over their uplink channels and perform over-the-air aggregation. For efficient transmission similar to the downlink, we represent $\boldsymbol{\theta}_{k,t}^J$ using a complex vector, with the real and imaginary parts of the vector containing the first and second half of the elements in $\boldsymbol{\theta}_{k,t}^J$, respectively. Specifically, we re-write $\boldsymbol{\theta}_{k,t}^J = [(\tilde{\boldsymbol{\theta}}_{k,t}^{J,\text{re}})^T, (\tilde{\boldsymbol{\theta}}_{k,t}^{J,\text{im}})^T]^T$, where $\tilde{\boldsymbol{\theta}}_{k,t}^{J,\text{re}} \triangleq [\theta_{k1,t}^J, \ldots, \theta_{k\frac{D}{2},t}^J]^T$ and $\tilde{\boldsymbol{\theta}}_{k,t}^{J,\text{im}} \triangleq [\theta_{k(\frac{D}{2}+1),t}^J, \ldots, \theta_{kD,t}^J]^T$. Then, we have the equivalent complex vector representation of $\boldsymbol{\theta}_{k,t}^J$, defined by $\tilde{\boldsymbol{\theta}}_{k,t}^J = \tilde{\boldsymbol{\theta}}_{k,t}^{J,\text{re}} + j\tilde{\boldsymbol{\theta}}_{k,t}^{J,\text{im}} \in \mathbb{C}^{\frac{D}{2}}$.

Transmitting $\tilde{\boldsymbol{\theta}}_{k,t}^J$ from device k to the BS over its uplink channel requires $\frac{D}{2}$ channel uses, one for each element in $\tilde{\boldsymbol{\theta}}_{k,t}^J$. At channel use l, the received signal vector at the BS, denoted by $\mathbf{v}_{l,t}$, is given by

$$\mathbf{v}_{l,t} = \sum_{k=1}^K \mathbf{h}_{k,t} a_{k,t} \tilde{\theta}_{kl,t}^J + \mathbf{u}_{l,t}^{\text{ul}}$$

where $a_{k,t} \in \mathbb{C}$ is the transmit beamforming weight at device k, and $\mathbf{u}_{l,t}^{\text{ul}} \in \mathbb{C}^N$ is the receiver AWGN vector with i.i.d. elements that are zero mean with variance σ_u^2. The BS applies receive beamforming to the received signal $\mathbf{v}_{l,t}$ over N antennas, for $l = 1, \ldots, \frac{D}{2}$, to obtain a weighted sum of $\tilde{\boldsymbol{\theta}}_{k,t}^J$'s from all $k \in \mathcal{K}$. Let $\mathbf{w}_t^{\text{ul}} \in \mathbb{C}^N$ be the unit-norm receive beamforming vector at the BS at round t, with $\|\mathbf{w}_t^{\text{ul}}\|^2 = 1$. The post-processed received signal vector over all $\frac{D}{2}$ channel uses is given by

$$\mathbf{z}_t = \sum_{k=1}^K (\mathbf{w}_t^{\text{ul}})^H \mathbf{h}_{k,t} a_{k,t} \tilde{\boldsymbol{\theta}}_{k,t}^J + \mathbf{n}_t^{\text{ul}} \tag{6}$$

where $\mathbf{n}_t^{\text{ul}} \in \mathbb{C}^{\frac{D}{2}}$ is the post-processed receiver noise with the l-th element being $(\mathbf{w}_t^{\text{ul}})^H \mathbf{u}_{l,t}^{\text{ul}}$, for $l = 1, \ldots, \frac{D}{2}$. Define $\alpha_{k,t}^{\text{ul}} \triangleq (\mathbf{w}_t^{\text{ul}})^H \mathbf{h}_{k,t} a_{k,t}$, which represents the effective channel from device k to the BS after applying transmit and receive beamforming. Following this, we re-write (6) as

$$\mathbf{z}_t = \sum_{k=1}^K \alpha_{k,t}^{\text{ul}} \tilde{\boldsymbol{\theta}}_{k,t}^J + \mathbf{n}_t^{\text{ul}}. \tag{7}$$

We consider uplink joint transmit and receive beamforming, where $\{a_{k,t}\}_{k \in \mathcal{K}}$ and $\mathbf{w}_t^{\text{ul}}$ are designed jointly. For over-the-air aggregation, the local models $\tilde{\boldsymbol{\theta}}_{k,t}^J$'s need to be added up coherently. Thus, the transmit and receive beamforming design should ensure that the resulting effective channels, $\alpha_{k,t}^{\text{ul}}$'s, are phase aligned. For this purpose, the transmit beamforming weight at device k is set to $a_{k,t} = \sqrt{p_{k,t}} \frac{\mathbf{h}_{k,t}^H \mathbf{w}_t^{\text{ul}}}{|\mathbf{h}_{k,t}^H \mathbf{w}_t^{\text{ul}}|}$, where $p_{k,t}$ is the transmit power scaling factor for device k at round t. Following this, the effective channels of all devices are phase aligned to 0 after receive beamforming, *i.e.,* $\alpha_{k,t}^{\text{ul}}$ is real-valued:

$$\alpha_{k,t}^{\text{ul}} = (\mathbf{w}_t^{\text{ul}})^H \mathbf{h}_{k,t} a_{k,t} = \sqrt{p_{k,t}} |\mathbf{h}_{k,t}^H \mathbf{w}_t^{\text{ul}}|, \ k \in \mathcal{K}.$$

Furthermore, each device is subject to transmit power budget. Let DP_k^{ul} be the transmit power budget at device k for sending each local model in D channel uses, where P_k^{ul} denotes the average transmit power budget per channel use. Then, for transmitting $\tilde{\boldsymbol{\theta}}_{k,t}^J$, we have the transmit power constraint $p_{k,t} \|\tilde{\boldsymbol{\theta}}_{k,t}^J\|^2 \leq DP_k^{\text{ul}}$.

At the BS receiver, after receive beamforming, the BS further scales $\mathbf{z}_t$ in (7) to obtain the complex equivalent global model update for the next round $t+1$:

$$\tilde{\boldsymbol{\theta}}_{t+1} = \frac{\mathbf{z}_t}{\sum_{k=1}^K \alpha_{k,t}^{\text{ul}}} = \sum_{k=1}^K \rho_{k,t} \tilde{\boldsymbol{\theta}}_{k,t}^J + \tilde{\mathbf{n}}_t^{\text{ul}} \tag{8}$$

where $\rho_{k,t} \triangleq \frac{\alpha_{k,t}^{\text{ul}}}{\sum_{j=1}^K \alpha_{j,t}^{\text{ul}}}$, $\sum_{k=1}^K \rho_{k,t} = 1$, and $\tilde{\mathbf{n}}_t^{\text{ul}} \triangleq \frac{\mathbf{n}_t^{\text{ul}}}{\sum_{k=1}^K \alpha_{k,t}^{\text{ul}}}$.

From the local model update in Section III-B, let $\Delta \tilde{\boldsymbol{\theta}}_{k,t} = \tilde{\boldsymbol{\theta}}_{k,t}^J - \tilde{\boldsymbol{\theta}}_{k,t}^0$ denote the equivalent complex representation of

the local model change after the local training at device k in round t. Based on this and (3), we can express the global model $\tilde{\boldsymbol{\theta}}_{t+1}$ in (8) in terms of $\tilde{\boldsymbol{\theta}}_t$ in round t as

$$\tilde{\boldsymbol{\theta}}_{t+1} = \tilde{\boldsymbol{\theta}}_t + \sum_{k=1}^{K} \rho_{k,t} \Delta\tilde{\boldsymbol{\theta}}_{k,t} + \sum_{k=1}^{K} \rho_{k,t} \tilde{\mathbf{n}}_{k,t}^{\text{dl}} + \tilde{\mathbf{n}}_t^{\text{ul}}. \quad (9)$$

Finally, the real-valued global model update $\boldsymbol{\theta}_{t+1}$ for round $t+1$ is recovered from $\tilde{\boldsymbol{\theta}}_{t+1}$ as $\boldsymbol{\theta}_{t+1} = [\Re\{\tilde{\boldsymbol{\theta}}_{t+1}\}^T, \Im\{\tilde{\boldsymbol{\theta}}_{t+1}\}^T]^T$.

Remark 1. The global model updating equation in (9) is derived from the entire round-trip FL procedure, including downlink-uplink transmission and the local model update at devices. The second term represents the aggregated update from the local training at K devices obtained via uplink transmission. The third and fourth noise terms reflect how the noisy downlink and uplink transmissions affect the global model update. Overall, the updating equation (9) shows how local model updates contribute to the global model update under the noisy communication channel and transmitter and receiver processing.

IV. JOINT DOWNLINK-UPLINK BEAMFORMING DESIGN

In this paper, we consider the design of the communication aspect of the FL system, aiming to maximize the training convergence rate. In particular, we jointly design the downlink and uplink beamforming to minimize the expected global loss function after T rounds. Let $\mathcal{T} = \{0, \ldots, T-1\}$. The optimization problem is formulated as

$$\mathcal{P}_o : \min_{\{\mathbf{w}_t^{\text{dl}}, \mathbf{w}_t^{\text{ul}}, \mathbf{p}_t\}_{t \in \mathcal{T}}} \mathbb{E}[F(\boldsymbol{\theta}_T)] \quad (10)$$

$$\text{s.t.} \quad \|\mathbf{w}_t^{\text{dl}}\|^2 \|\boldsymbol{\theta}_t\|^2 \leq DP^{\text{dl}}, \quad t \in \mathcal{T}, \quad (11)$$

$$p_{k,t} \|\boldsymbol{\theta}_{k,t}^J\|^2 \leq DP_k^{\text{ul}}, \quad k \in \mathcal{K}, t \in \mathcal{T}, \quad (12)$$

$$\|\mathbf{w}_t^{\text{ul}}\|^2 = 1, \quad t \in \mathcal{T} \quad (13)$$

where $\mathbb{E}[\cdot]$ is the expectation taken w.r.t. receiver noise and mini-batch sampling in local training at each device, and $\mathbf{p}_t \triangleq [p_{1,t}, \ldots, p_{K,t}]^T$ contains the uplink transmit power scaling factors of all K devices at round t. Constraints in (11) and (12) are the transmit power constraints at the BS and each device k, respectively, and constraint in (13) specifies the receive beamforming vector at the BS to be unit-norm.

Problem $\mathcal{P}_o$ is a finite-horizon stochastic optimization problem, which is challenging to solve. To tackle this problem, we develop a more tractable upper bound on $\mathbb{E}[F(\boldsymbol{\theta}_T)]$ by analyzing the convergence rate of the global loss function, and then we develop a joint downlink-uplink beamforming algorithm to minimize this upper bound.

A. Convergence Analysis on Global Training Loss

Let $F^\star$ denote the minimum global loss under the optimal model $\boldsymbol{\theta}^\star$. To examine the expected global loss function $\mathbb{E}[F(\boldsymbol{\theta}_T)]$ in the FL system described in Section II, we can equivalently analyze $\mathbb{E}[F(\boldsymbol{\theta}_T)] - F^\star$, i.e., the expected gap of the global loss function at round T to the minimum global

loss, based on the global model updates $\{\boldsymbol{\theta}_t\}$ obtained in Section III. We first make the following three assumptions on the local loss functions, the SGD, and the difference between the global and weighted average of the local loss functions. These assumptions are commonly adopted for the convergence analysis of the FL model training [9], [11], [12].

Assumption 1. The local loss functions $F_k(\cdot)$'s are differentiable and are L-smooth: $F_k(\mathbf{y}) \leq F_k(\mathbf{x}) + (\mathbf{y}-\mathbf{x})^T \nabla F_k(\mathbf{x}) + \frac{L}{2}\|\mathbf{y}-\mathbf{x}\|^2$, $\forall k \in \mathcal{K}$, $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^D$. Also, $F_k(\cdot)$'s are λ-strongly convex: $F_k(\mathbf{y}) \geq F_k(\mathbf{x}) + (\mathbf{y}-\mathbf{x})^T \nabla F_k(\mathbf{x}) + \frac{\lambda}{2}\|\mathbf{y}-\mathbf{x}\|^2$, $\forall k \in \mathcal{K}$, $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^D$.

Assumption 2. The mini-batch SGD is unbiased: $\mathbb{E}_{\mathcal{B}}[\nabla F_k(\boldsymbol{\theta}_{k,t}^\tau; \mathcal{B}_{k,t}^\tau)] = \nabla F_k(\boldsymbol{\theta}_{k,t}^\tau)$, $\forall k \in \mathcal{K}$, $\forall \tau, t$. The variance of the mini-batch stochastic gradient is bounded by μ: For $\forall k \in \mathcal{K}$, $\forall \tau, t$, $\mathbb{E}[\|\nabla F_k(\boldsymbol{\theta}_{k,t}^\tau; \mathcal{B}_{k,t}^\tau) - \nabla F_k(\boldsymbol{\theta}_{k,t}^\tau)\|^2] \leq \mu$.

Assumption 3. The gradient divergence is bounded by δ: For $\forall k \in \mathcal{K}$, $\forall \tau, t$, $\mathbb{E}[\|\nabla F(\boldsymbol{\theta}_t) - \sum_{k=1}^{K} \phi_k \nabla F_k(\boldsymbol{\theta}_t)\|^2] \leq \delta$, where $\phi_k \in \mathbb{R}$, $\phi_k \geq 0$, and $\sum_{k=1}^{K} \phi_k = 1$.

We now evaluate the expected gap $\mathbb{E}[F(\boldsymbol{\theta}_T)] - F^\star$ through the training loss convergence rate analysis. We point out that although different bounds on this expected gap have been derived in the literature, they are based on either idealized or simplified communication link models without multi-antenna processing effects. Based on the global model update obtained in (9), we first bound the expected change of the global loss function in two consecutive rounds as

$$\mathbb{E}[F(\boldsymbol{\theta}_{t+1}) - F(\boldsymbol{\theta}_t)] = \sum_{k=1}^{K} \frac{S_k}{S} \mathbb{E}[F_k(\boldsymbol{\theta}_{t+1}) - F_k(\boldsymbol{\theta}_t)]$$

$$\overset{(a)}{\leq} \mathbb{E}[(\boldsymbol{\theta}_{t+1} - \boldsymbol{\theta}_t)^T \nabla F(\boldsymbol{\theta}_t)] + \frac{L}{2}\mathbb{E}[\|\boldsymbol{\theta}_{t+1} - \boldsymbol{\theta}_t\|^2] \quad (14)$$

$$\overset{(b)}{=} \Re\{\mathbb{E}[(\tilde{\boldsymbol{\theta}}_{t+1} - \tilde{\boldsymbol{\theta}}_t)^H \nabla \tilde{F}(\boldsymbol{\theta}_t)]\} + \frac{L}{2}\mathbb{E}[\|\tilde{\boldsymbol{\theta}}_{t+1} - \tilde{\boldsymbol{\theta}}_t\|^2]$$

$$\overset{(c)}{=} \Re\left\{\mathbb{E}\left[\underbrace{\left(\sum_{k=1}^{K} \rho_{k,t}\Delta\tilde{\boldsymbol{\theta}}_{k,t} + \sum_{k=1}^{K} \rho_{k,t}\tilde{\mathbf{n}}_{k,t}^{\text{dl}} + \tilde{\mathbf{n}}_t^{\text{ul}}\right)^H \nabla\tilde{F}(\boldsymbol{\theta}_t)}_{\triangleq A_{1,t}}\right]\right\}$$

$$+ \frac{L}{2}\underbrace{\mathbb{E}\left[\left\|\sum_{k=1}^{K} \rho_{k,t}\Delta\tilde{\boldsymbol{\theta}}_{k,t} + \sum_{k=1}^{K} \rho_{k,t}\tilde{\mathbf{n}}_{k,t}^{\text{dl}} + \tilde{\mathbf{n}}_t^{\text{ul}}\right\|^2\right]}_{\triangleq A_{2,t}}. \quad (15)$$

where (a) follows the L-smoothness of $F_k(\cdot)$'s in Assumption 1 and the fact that $\nabla F(\boldsymbol{\theta}) = \sum_{k=1}^{K} \frac{S_k}{S} \nabla F_k(\boldsymbol{\theta})$ following (2), (b) is the equivalent expression of (14) by using the equivalent complex representation $\tilde{\boldsymbol{\theta}}_t$ of $\boldsymbol{\theta}_t$, where $\nabla\tilde{F}(\boldsymbol{\theta}_t)$ denotes the equivalent complex representation of the global loss gradient $\nabla F(\boldsymbol{\theta}_t)$ in round t, and (c) is obtained following the global model update in (9). The upper bound in (15) clearly shows the effects of noisy channels and multi-antenna transmit/receive beamforming processing at both downlink and uplink on the loss function. Note that $A_{1,t}$ and $A_{2,t}$ defined in (15) are functions of the aggregated local model change, the downlink-uplink transmission processing, and the receiver noise at round t. Next, we bound $A_{1,t}$ and $A_{2,t}$ separately.

For $A_{1,t}$, since the receiver noise at the devices and the BS are zero mean and independent of $\nabla \tilde{F}(\boldsymbol{\theta}_t)$, we have

$$A_{1,t} = \mathfrak{Re}\left\{ \mathbb{E}\left[\left(\sum_{k=1}^{K} \rho_{k,t}\Delta\tilde{\boldsymbol{\theta}}_{k,t} \right)^H \nabla\tilde{F}(\boldsymbol{\theta}_t) \right] \right\} \quad (16)$$

$$= \mathbb{E}\left[\left(\sum_{k=1}^{K} \rho_{k,t}\Delta\boldsymbol{\theta}_{k,t} \right)^T \nabla F(\boldsymbol{\theta}_t) \right] \quad (17)$$

where $\Delta\boldsymbol{\theta}_{k,t} \triangleq \boldsymbol{\theta}_{k,t}^J - \boldsymbol{\theta}_{k,t}^0$ is the real-valued local model change after the local training at device k in round t, and (17) is the equivalent expression of (16) by using the real-value parameters. Based on the mini-batch SGD from (5), we have $\Delta\boldsymbol{\theta}_{k,t} = -\eta_t \sum_{\tau=0}^{J-1} \nabla F_k(\boldsymbol{\theta}_{k,t}^\tau; \mathcal{B}_{k,t}^\tau)$.

Based on Assumptions 1–3, we provide an upper bound on $A_{1,t}$, which is stated in the following lemma. Detailed proof is omitted. Part of our proof has used some techniques in [11, Th. 1] (similarly for the proof of Lemma 2).

Lemma 1. Consider the FL system described in Section III and Assumptions 1–3. Let $Q_t \triangleq 1 - 4\eta_t^2 J^2 L^2$ and assume $\eta_t J < \frac{1}{2L}$, $\forall t \in \mathcal{T}$. Then, $A_{1,t}$ is upper bounded as

$$A_{1,t} \le \eta_t J\left(\frac{2}{Q_t} - \frac{5}{2} \right)\mathbb{E}\left[\|\nabla F(\boldsymbol{\theta}_t)\|^2 \right]$$
$$+ \frac{D(1-Q_t)}{4\eta_t J Q_t}\sum_{k=1}^{K} \frac{\rho_{k,t}\sigma_d^2}{|\mathbf{h}_{k,t}^H\mathbf{w}_t^{dl}|^2} + \frac{\eta_t J}{2}\left(\frac{\delta+\mu}{Q_t} + \frac{\delta-\mu}{2} \right). \quad (18)$$

Note that for the bound in (18), η_t and J are parameters set in the SGD for the local model update at each device, and L, μ, δ are parameters specified in Assumptions 1–3.

For $A_{2,t}$, since the receiver noise at the BS is zero mean and independent of $\sum_{k=1}^{K}\rho_{k,t}(\Delta\tilde{\boldsymbol{\theta}}_{k,t} + \tilde{\mathbf{n}}_{k,t}^{dl})$, we have

$$A_{2,t} = \mathbb{E}\left[\left\| \sum_{k=1}^{K}\rho_{k,t}(\Delta\tilde{\boldsymbol{\theta}}_{k,t} + \tilde{\mathbf{n}}_{k,t}^{dl}) \right\|^2 \right] + \mathbb{E}[\|\tilde{\mathbf{n}}_t^{ul}\|^2]$$

$$= \mathbb{E}\left[\left\| \sum_{k=1}^{K}\rho_{k,t}(\Delta\tilde{\boldsymbol{\theta}}_{k,t} + \tilde{\mathbf{n}}_{k,t}^{dl}) \right\|^2 \right] + \frac{D\sigma_u^2}{2(\sum_{k=1}^{K}\alpha_{k,t}^{ul})^2}. \quad (19)$$

Based on Assumptions 1–3, we can upper bound $A_{2,t}$ as shown in the following lemma. Detailed proof is omitted.

Lemma 2. Consider the FL system described in Section III and Assumptions 1–3 and assume $\eta_t J < \frac{1}{2L}$, $\forall t \in \mathcal{T}$. Then, $A_{2,t}$ is upper bounded as

$$A_{2,t} \le \frac{2}{L^2}\left(\frac{1-Q_t}{Q_t} \right)\mathbb{E}[\|\nabla F(\boldsymbol{\theta}_t)\|^2]$$
$$+ D\left(\frac{1-Q_t}{Q_t}\sum_{k=1}^{K}\frac{\rho_{k,t}\sigma_d^2}{|\mathbf{h}_{k,t}^H\mathbf{w}_t^{dl}|^2} + \sum_{k=1}^{K}\frac{\rho_{k,t}^2\sigma_d^2}{|\mathbf{h}_{k,t}^H\mathbf{w}_t^{dl}|^2} \right)$$
$$+ \frac{D\sigma_u^2}{2(\sum_{k=1}^{K}\alpha_{k,t}^{ul})^2} + \frac{1-Q_t}{2L^2Q_t}\left(\left(1 - Q_t + \frac{Q_t}{J}\right)\mu + 4\delta \right). \quad (20)$$

We now analyze the expected gap $\mathbb{E}[F(\boldsymbol{\theta}_T)] - F^\star$ at round T. From (15), the expected gap at round $t+1$ is bounded as

$$\mathbb{E}[F(\boldsymbol{\theta}_{t+1})] - F^\star \le \mathbb{E}[F(\boldsymbol{\theta}_t)] - F^\star + A_{1,t} + \frac{L}{2}A_{2,t}. \quad (21)$$

Using Lemmas 1 and 2, we can further bound the right hand side (RHS) of (21). Summing up both sides over $t \in \mathcal{T}$ and rearranging the terms, we can obtain the upper bound on $\mathbb{E}[F(\boldsymbol{\theta}_T)] - F^\star$, which is stated in Proposition 1 below.

Proposition 1. For the FL system described in Section III, under Assumptions 1–3 and for $\frac{1}{10L} \le \eta_t J < \frac{1}{2L}$, $\forall t \in \mathcal{T}$, the expected gap $\mathbb{E}[F(\boldsymbol{\theta}_T)] - F^\star$ after T communication rounds is upper bounded by

$$\mathbb{E}[F(\boldsymbol{\theta}_T)] - F^\star \le \Gamma\prod_{t=0}^{T-1}G_t + \Lambda + \sum_{t=0}^{T-2}H(\mathbf{w}_t^{dl}, \mathbf{w}_t^{ul}, \mathbf{p}_t)\prod_{s=t+1}^{T-1}G_s$$
$$+ H(\mathbf{w}_{T-1}^{dl}, \mathbf{w}_{T-1}^{ul}, \mathbf{p}_{T-1}) \quad (22)$$

where $\Gamma \triangleq \mathbb{E}[F(\boldsymbol{\theta}_0)] - F^\star$, $\Lambda \triangleq \sum_{t=0}^{T-2}C_t\left(\prod_{s=t+1}^{T-1}G_s\right) + C_{T-1}$ with

$$G_t \triangleq \frac{1-Q_t}{4\eta_t J\lambda Q_t}\left(5(1-Q_t) + 4\sqrt{1-Q_t} - 1\right) + 1,$$
$$C_t \triangleq \frac{\eta_t J}{2}\left(\frac{\delta+\mu}{Q_t} + \frac{\delta-\mu}{2} \right) + \frac{1-Q_t}{2L^2Q_t}\left(\left(1 - Q_t + \frac{Q_t}{J}\right)\mu + 4\delta \right),$$

and $H(\mathbf{w}_t^{dl}, \mathbf{w}_t^{ul}, \mathbf{p}_t)$ is defined in (23).

Proof: See Appendix A. ∎

Note that the upper bound for $\mathbb{E}[F(\boldsymbol{\theta}_T)] - F^\star$ in (22) reflects how the downlink-uplink transmission and the local training affect the convergence of the global model update. In particular, the first term shows the impact of the initial starting point $\boldsymbol{\theta}_0$. The second term Λ is a weighted sum of C_t's, each accounting for the gradient divergence of the local loss function from the global loss function using the mini-batch SGD during the local model updates among K devices in round t.[1] The third term is a weighted sum of $H(\mathbf{w}_t^{dl}, \mathbf{w}_t^{ul}, \mathbf{p}_t)$, where $H(\mathbf{w}_t^{dl}, \mathbf{w}_t^{ul}, \mathbf{p}_t)$ in (23) is in the form of a weighted sum of the inverse of SNRs (*i.e.,* noise-to-signal ratio). Two types of SNRs are shown: the terms with σ_d^2 reflect the post-processing SNR at the BS receiver due to the downlink device receiver noise (after downlink and uplink beamforming and receiver processing), and the term with σ_u^2 shows the post-processing SNR at the BS receiver due to the BS receiver noise in the uplink after receiver beamforming and processing.

The upper bound given in Proposition 1 is in a more tractable form for the expected gap $\mathbb{E}[F(\boldsymbol{\theta}_T)] - F^\star$ that we can use for the optimization design. In the following, we directly minimize this upper bound to obtain a joint downlink and uplink beamforming solution.

B. Joint Downlink-Uplink Beamforming Algorithm

We now replace the objective function in $\mathcal{P}_o$ with the upper bound in (22). Define

$$\Psi(\{\mathbf{w}_t^{dl}, \mathbf{w}_t^{ul}, \mathbf{p}_t\})$$
$$\triangleq \sum_{t=0}^{T-2}H(\mathbf{w}_t^{dl}, \mathbf{w}_t^{ul}, \mathbf{p}_t)\prod_{s=t+1}^{T-1}G_s + H(\mathbf{w}_{T-1}^{dl}, \mathbf{w}_{T-1}^{ul}, \mathbf{p}_{T-1}).$$

[1]Recall that λ in the expression of G_t is specified in Assumption 1.

$$H(\mathbf{w}_t^{\mathrm{dl}}, \mathbf{w}_t^{\mathrm{ul}}, \mathbf{p}_t) \triangleq \frac{LD}{2}\left(\frac{1-Q_t+\sqrt{1-Q_t}}{Q_t}\right)\frac{\sigma_{\mathrm{d}}^2\left(\sum_{k=1}^{K}\frac{\sqrt{p_{k,t}}|\mathbf{h}_{k,t}^H\mathbf{w}_t^{\mathrm{ul}}|}{|\mathbf{h}_{k,t}^H\mathbf{w}_t^{\mathrm{dl}}|^2}\right)}{\sum_{k=1}^{K}\sqrt{p_{k,t}}|\mathbf{h}_{k,t}^H\mathbf{w}_t^{\mathrm{ul}}|} + \frac{LD}{2}\frac{\sigma_{\mathrm{d}}^2\left(\sum_{k=1}^{K}\frac{p_{k,t}|\mathbf{h}_{k,t}^H\mathbf{w}_t^{\mathrm{ul}}|^2}{|\mathbf{h}_{k,t}^H\mathbf{w}_t^{\mathrm{dl}}|^2}\right)+\frac{\sigma_{\mathrm{u}}^2}{2}}{\left(\sum_{k=1}^{K}\sqrt{p_{k,t}}|\mathbf{h}_{k,t}^H\mathbf{w}_t^{\mathrm{ul}}|\right)^2} \tag{23}$$

Omitting the constant terms $\Gamma\prod_{t=0}^{T-1}G_t + \Lambda$ in (22), we can equivalently minimize $\Psi(\{\mathbf{w}_t^{\mathrm{dl}}, \mathbf{w}_t^{\mathrm{ul}}, \mathbf{p}_t\})$. Thus, instead of $\mathcal{P}_o$, we consider the following joint downlink and uplink beamforming optimization problem

$$\mathcal{P}_1 : \min_{\{\mathbf{w}_t^{\mathrm{dl}}, \mathbf{w}_t^{\mathrm{ul}}, \mathbf{p}_t\}_{t=0}^{T-1}} \Psi(\{\mathbf{w}_t^{\mathrm{dl}}, \mathbf{w}_t^{\mathrm{ul}}, \mathbf{p}_t\}) \quad \text{s.t. (11)(12)(13)}.$$

Problem $\mathcal{P}_1$ is a T-horizon joint optimization problem that includes T communication rounds of the model update. Note that in Proposition 1, for $\frac{1}{10L} \leq \eta_t J < \frac{1}{2L}$, we have $G_t > 0$, $\forall\ t \in \mathcal{T}$, and thus, $\prod_{s=t+1}^{T-1} G_s > 0$ in $\Psi(\{\mathbf{w}_t^{\mathrm{dl}}, \mathbf{w}_t^{\mathrm{ul}}, \mathbf{p}_t\})$. Thus, $\mathcal{P}_1$ can be decomposed into T subproblems, one for each round t given by

$$\mathcal{P}_2^t : \min_{\mathbf{w}_t^{\mathrm{dl}}, \mathbf{w}_t^{\mathrm{ul}}, \mathbf{p}_t} H(\mathbf{w}_t^{\mathrm{dl}}, \mathbf{w}_t^{\mathrm{ul}}, \mathbf{p}_t)$$

$$\text{s.t.} \quad \|\mathbf{w}_t^{\mathrm{dl}}\|^2\|\boldsymbol{\theta}_t\|^2 \leq DP^{\mathrm{dl}}, \tag{24}$$

$$p_{k,t}\|\boldsymbol{\theta}_{k,t}^J\|^2 \leq DP_k^{\mathrm{ul}}, \; k \in \mathcal{K}, \tag{25}$$

$$\|\mathbf{w}_t^{\mathrm{ul}}\|^2 = 1. \tag{26}$$

Note that $H(\mathbf{w}_t^{\mathrm{dl}}, \mathbf{w}_t^{\mathrm{ul}}, \mathbf{p}_t)$ in (22) is an involved non-convex function of $(\mathbf{w}_t^{\mathrm{dl}}, \mathbf{w}_t^{\mathrm{ul}}, \mathbf{p}_t)$. It is difficult to find the optimal solution to $\mathcal{P}_2^t$ directly. Instead, we propose to use the AO approach to solve $\mathcal{P}_2^t$ w.r.t. the downlink beamforming $\mathbf{w}_t^{\mathrm{dl}}$, and uplink beamforming $(\mathbf{w}_t^{\mathrm{ul}}, \mathbf{p}_t)$ alternatingly. Furthermore, we propose to solve each AO subproblem via PGD [13].

To facilitate the computation in our algorithm, we express all the complex quantities in $\mathcal{P}_2^t$ using their real and imaginary parts. Define $\mathbf{x}_t^{\mathrm{dl}} \triangleq [\Re\{\mathbf{w}_t^{\mathrm{dl}}\}^T, \Im\{\mathbf{w}_t^{\mathrm{dl}}\}^T]^T$, $\mathbf{x}_t^{\mathrm{ul}} \triangleq [\Re\{\mathbf{w}_t^{\mathrm{ul}}\}^T, \Im\{\mathbf{w}_t^{\mathrm{ul}}\}^T]^T$, and

$$\mathbf{H}_{k,t} \triangleq \begin{bmatrix} \Re\{\mathbf{h}_{k,t}\mathbf{h}_{k,t}^H\} & -\Im\{\mathbf{h}_{k,t}\mathbf{h}_{k,t}^H\} \\ \Im\{\mathbf{h}_{k,t}\mathbf{h}_{k,t}^H\} & \Re\{\mathbf{h}_{k,t}\mathbf{h}_{k,t}^H\} \end{bmatrix}, \; k \in \mathcal{K}, t \in \mathcal{T}.$$

Then, we have $\|\mathbf{w}_t^{\mathrm{dl}}\|^2 = \|\mathbf{x}_t^{\mathrm{dl}}\|^2$, $\|\mathbf{w}_t^{\mathrm{ul}}\|^2 = \|\mathbf{x}_t^{\mathrm{ul}}\|^2$, $|\mathbf{h}_{k,t}^H\mathbf{w}_t^{\mathrm{dl}}|^2 = (\mathbf{x}_t^{\mathrm{dl}})^T\mathbf{H}_{k,t}\mathbf{x}_t^{\mathrm{dl}}$, and $|\mathbf{h}_{k,t}^H\mathbf{w}_t^{\mathrm{ul}}|^2 = (\mathbf{x}_t^{\mathrm{ul}})^T\mathbf{H}_{k,t}\mathbf{x}_t^{\mathrm{ul}}$. Thus, we can express $H(\mathbf{w}_t^{\mathrm{dl}}, \mathbf{w}_t^{\mathrm{ul}}, \mathbf{p}_t)$ in (23) using the corresponding real-valued vectors $(\mathbf{x}_t^{\mathrm{dl}}, \mathbf{x}_t^{\mathrm{ul}}, \mathbf{p}_t)$ by replacing $|\mathbf{h}_{k,t}^H\mathbf{w}_t^{\mathrm{dl}}|$ and $|\mathbf{h}_{k,t}^H\mathbf{w}_t^{\mathrm{ul}}|$ in (23) with $((\mathbf{x}_t^{\mathrm{dl}})^T\mathbf{H}_{k,t}\mathbf{x}_t^{\mathrm{dl}})^{\frac{1}{2}}$ and $((\mathbf{x}_t^{\mathrm{ul}})^T\mathbf{H}_{k,t}\mathbf{x}_t^{\mathrm{ul}})^{\frac{1}{2}}$, respectively. Let $\Phi(\mathbf{x}_t^{\mathrm{dl}}, \mathbf{x}_t^{\mathrm{ul}}, \mathbf{p}_t)$ denote this resulting equivalent converted function from $H(\mathbf{w}_t^{\mathrm{dl}}, \mathbf{w}_t^{\mathrm{ul}}, \mathbf{p}_t)$. Then, $\mathcal{P}_2^t$ can be equivalently transformed into the following problem with all real-valued optimization variables:

$$\mathcal{P}_3^t : \min_{\mathbf{x}_t^{\mathrm{dl}}, \mathbf{x}_t^{\mathrm{ul}}, \mathbf{p}_t} \Phi(\mathbf{x}_t^{\mathrm{dl}}, \mathbf{x}_t^{\mathrm{ul}}, \mathbf{p}_t)$$

$$\text{s.t. } \mathbf{x}_t^{\mathrm{dl}} \in \mathcal{X}_t^{\mathrm{dl}}, \mathbf{x}_t^{\mathrm{ul}} \in \mathcal{X}_t^{\mathrm{ul}}, \mathbf{p}_t \in \mathcal{Y}_t$$

where $\mathcal{X}_t^{\mathrm{dl}} \triangleq \{\mathbf{x}_t^{\mathrm{dl}} : \|\mathbf{x}_t^{\mathrm{dl}}\|^2\|\boldsymbol{\theta}_t\|^2 \leq DP^{\mathrm{dl}}\}$, $\mathcal{X}_t^{\mathrm{ul}} \triangleq \{\mathbf{x}_t^{\mathrm{ul}} : \|\mathbf{x}_t^{\mathrm{ul}}\|^2 = 1\}$, and $\mathcal{Y}_t \triangleq \{\mathbf{p}_t : p_{k,t}\|\boldsymbol{\theta}_{k,t}^J\|^2 \leq DP_k^{\mathrm{ul}}, \; k \in \mathcal{K}\}$.

We use the AO approach to compute a solution to $\mathcal{P}_3^t$ at

Algorithm 1 The JDU-BF Algorithm for $\mathcal{P}_1$

Initialization: Set $t = 0$.
repeat
  Initialization: Set $\mathbf{x}_t^{\mathrm{dl}(0)}, \mathbf{x}_t^{\mathrm{ul}(0)}, \mathbf{p}_t^{(0)}$; Set $i = 0$.
  repeat
    1) Update downlink transmit beamforming vector
$$\mathbf{x}_t^{\mathrm{dl}(i+1)} = \underset{\mathbf{x}_t^{\mathrm{dl}} \in \mathcal{X}_t^{\mathrm{dl}}}{\arg\min} \Phi(\mathbf{x}_t^{\mathrm{dl}}, \mathbf{x}_t^{\mathrm{ul}(i)}, \mathbf{p}_t^{(i)}). \tag{27}$$
    2) Update uplink receive beamforming vector
$$\mathbf{x}_t^{\mathrm{ul}(i+1)} = \underset{\mathbf{x}_t^{\mathrm{ul}} \in \mathcal{X}_t^{\mathrm{ul}}}{\arg\min} \Phi(\mathbf{x}_t^{\mathrm{dl}(i+1)}, \mathbf{x}_t^{\mathrm{ul}}, \mathbf{p}_t^{(i)}). \tag{28}$$
    3) Update uplink device transmit power
$$\mathbf{p}_t^{(i+1)} = \underset{\mathbf{p}_t \in \mathcal{Y}_t}{\arg\min} \Phi(\mathbf{x}_t^{\mathrm{dl}(i+1)}, \mathbf{x}_t^{\mathrm{ul}(i+1)}, \mathbf{p}_t). \tag{29}$$
    4) Set $i \leftarrow i + 1$.
  until convergence // solve $\mathcal{P}_3^t$
  Set $t \leftarrow t + 1$.
until $t = T$

each round $t \in \mathcal{T}$. Our proposed JDU-BF algorithm for $\mathcal{P}_1$ is summarized in Algorithm 1. Note that subproblem (27) is a downlink beamforming problem, subproblem (28) is an uplink receive beamforming problem, and subproblem (29) is an uplink transmit power minimization problem. Thus, our proposed algorithm solves downlink and uplink beamforming problems alternatingly.

For each subproblem in (27)–(29), the objective function is a complicated non-convex function of the optimization variable. Thus, we adopt PGD to solve each subproblem. PGD [13] is an iterative first-order algorithm that uses gradient updates to solve a constrained minimization problem: $\min_{\mathbf{x} \in \mathcal{X}} f(\mathbf{x})$, where $\mathcal{X}$ is the convex feasible set for $\mathbf{x}$. PGD has the following updating procedure: At iteration j,

$$\mathbf{x}_{j+1} = \Pi_{\mathcal{X}}\left(\mathbf{x}_j - \beta\nabla_{\mathbf{x}}f(\mathbf{x}_j)\right) \tag{30}$$

where $\beta > 0$ is the step size and $\Pi_{\mathcal{X}}(\mathbf{x})$ denotes the projection of point $\mathbf{x}$ onto set $\mathcal{X}$. Note that due to the inherent structure of our problem, PGD is particularly suitable for solving subproblems (27)–(29) at each AO iteration. In particular, the projection $\Pi_{\mathcal{X}}(\mathbf{x})$ operation can be expressed in closed-form for each of subproblems (27)–(29):

- For subproblem (27):

$$\Pi_{\mathcal{X}_t^{\mathrm{dl}}}(\mathbf{x}_t^{\mathrm{dl}}) = \begin{cases} \sqrt{\frac{DP^{\mathrm{dl}}}{\|\mathbf{x}_t^{\mathrm{dl}}\|^2\|\boldsymbol{\theta}_t\|^2}}\mathbf{x}_t^{\mathrm{dl}} & \mathbf{x}_t^{\mathrm{dl}} \notin \mathcal{X}_t^{\mathrm{dl}}, \\ \mathbf{x}_t^{\mathrm{dl}} & \mathbf{x}_t^{\mathrm{dl}} \in \mathcal{X}_t^{\mathrm{dl}}. \end{cases}$$

- For subproblem (28): $\Pi_{\mathcal{X}_t^{\mathrm{ul}}}(\mathbf{x}_t^{\mathrm{ul}}) = \frac{\mathbf{x}_t^{\mathrm{ul}}}{\|\mathbf{x}_t^{\mathrm{ul}}\|}$.

- For subproblem (29): $\Pi_{\mathcal{Y}_t}(\mathbf{p}_t)$ is given by

$$p_{k,t} = \begin{cases} \frac{DP_k^{\mathrm{ul}}}{\|\boldsymbol{\theta}_{k,t}^J\|^2} & p_{k,t} > \frac{DP_k^{\mathrm{ul}}}{\|\boldsymbol{\theta}_{k,t}^J\|^2}, \\ p_{k,t} & p_{k,t} \leq \frac{DP_k^{\mathrm{ul}}}{\|\boldsymbol{\theta}_{k,t}^J\|^2}. \end{cases} \quad \forall\, k \in \mathcal{K}$$

Thus, the computation using PGD via (30) has low complexity. Also, PGD is guaranteed to find an approximate stationary point for each subproblem in polynomial time [15]. We summarized our proposed JDU-BF algorithm in Algorithm 1.

C. Separate Downlink and Uplink Beamforming Design

In the above, we have proposed joint downlink-uplink beamforming design for the FL system, which is based on the global model update in (9) derived from each communication round. For comparison purpose, we also consider the conventional approach where downlink and uplink transmission are designed separately for the communication system.

Downlink: We formulate the problem of downlink beamforming and the uplink beamforming separately. At the communication round t, since the BS broadcasts the global model to all devices, the downlink beamforming problem is to maximize the minimum received SNR, which is a single-group multicast beamforming max-min fair problem:

$$\max_{\mathbf{w}_t^{\mathrm{dl}}} \min_{k \in \mathcal{K}} |\mathbf{h}_{k,t}^H \mathbf{w}_t^{\mathrm{dl}}|^2 \quad \text{s.t.} \quad \|\mathbf{w}_t^{\mathrm{dl}}\|^2 \|\boldsymbol{\theta}_t\|^2 \leq DP^{\mathrm{dl}}. \quad (31)$$

The solution to this problem can be efficiently computed using the projected subgradient algorithm proposed in [16] based on the optimal multicast beamforming structure [17].

Uplink: For uplink over-the-air aggregation, the transmit beamforming weight at device k is set to $a_{k,t} = \sqrt{p_{k,t}} \frac{\mathbf{h}_{k,t}^H \mathbf{w}_t^{\mathrm{ul}}}{|\mathbf{h}_{k,t}^H \mathbf{w}_t^{\mathrm{ul}}|}$ to phase-align the transmissions from all devices. Each device uses the maximum transmit power; thus, the power scaling factor is $p_{k,t} = \frac{DP_k^{\mathrm{ul}}}{\|\boldsymbol{\theta}_{k,t}^J\|^2}$, $\forall\, k \in \mathcal{K}$. Then, we design uplink receive beamforming to maximize the received SNR (from the aggregated signal) at the BS:

$$\max_{\mathbf{w}_t^{\mathrm{ul}}: \|\mathbf{w}_t^{\mathrm{ul}}\|^2 = 1} \sum_{k=1}^K p_{k,t} |\mathbf{h}_{k,t}^H \mathbf{w}_t^{\mathrm{ul}}|^2. \quad (32)$$

This problem can be solved using PGD via (30). We name this approach as the separate downlink and uplink beamforming (SDU-BF) algorithm.

V. SIMULATION RESULTS

1) Simulation Setup: We consider the real-world dataset for image classification under an LTE wireless system setting. Following the typical LTE specifications, we set system bandwidth 10 MHz and carrier frequency 2 GHz. The maximum BS transmit power is 47 dBm. The maximum device transmit power is 23 dBm, and we assume the devices use 1 MHz bandwidth for uplink transmission. The path gain between the BS and device k is $G_k[\mathrm{dB}] = -139.2 - 35 \log_{10} d_k - \psi_k$, where $d_k \in (1\,\mathrm{km}, 1.5\,\mathrm{km})$ is the BS-device distance in kilometers, and ψ_k is the shadowing random variable with standard deviation 8 dB. The channel vector is generated as $\mathbf{h}_{k,t} = \sqrt{G_k}\bar{\mathbf{h}}_{k,t}$ with $\bar{\mathbf{h}}_{k,t} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$. Noise power spectral density is $N_0 = -174$ dBm/Hz, and noise figure $N_F = 8$ dB and 2 dB at the device and BS receivers, respectively.

We adopt the MNIST dataset [18] for model training and testing. MNIST consists of 6×10^4 training samples and 1×10^4 test samples from 10 different classes. Each sample is a labeled image of size 28×28 pixels, *i.e.,* $\mathbf{s} \in \mathbb{R}^{784}$ and $v \in \{0, \ldots, 9\}$ indicating the class. We consider training a convolutional neural network with an $8 \times 3 \times 3$ ReLU convolutional layer, a 2×2 max pooling layer, a ReLU fully-connected layer, and a softmax output layer, resulting in $D = 1.361 \times 10^4$ model parameters in total. We use the 1×10^4 test samples to measure the test accuracy of the global model update $\boldsymbol{\theta}_t$ at each round t. The training samples are randomly and evenly distributed over devices, and the local dataset at device k has $S_k = \frac{6 \times 10^4}{K}$ samples. For the local training via the SGD at each device, we set $L = 10$, $J = 30$, mini-batch size $|\mathcal{B}_{k,t}^\tau| = \frac{2 \times 10^3}{K}, \forall k, \tau, t$, and the learning rate $\eta_t = \frac{1}{10JL}, \forall t$.

2) Performance Comparison: For the comparison purpose, we consider the following three schemes: i) **Ideal FL** [1]: Perform FL via the global model update in (9), assuming error-free downlink and uplink and perfect recovery of model parameters at the BS and devices, *i.e.,* receiver noise $\tilde{\mathbf{n}}_{k,t}^{\mathrm{dl}} = \tilde{\mathbf{n}}_t^{\mathrm{ul}} = \mathbf{0}$, receiver post-processing weight $\rho_{k,t} = \frac{1}{K}, \forall k, t$. This benchmark provides the performance upper bound for all schemes. ii) **SDU-BF**: the separate SNR-maximizing design scheme described in Section IV-C. iii) **Random beamforming (RBF)**: Perform FL via (9) with randomly generated downlink and uplink beamforming vectors $\mathbf{w}_t^{\mathrm{dl}}$ and $\mathbf{w}_t^{\mathrm{ul}}$. The devices use the maximum transmit power and do not perform transmit beamforming phase alignment.

Fig. 1 shows the test accuracy performance by the considered methods over communication round T for three system settings for (N, K). All curves are obtained by averaging over 20 channel realizations. The shadowed area over each curve indicates the 90% confidence interval of the curve. Fig. 1-Left shows the test accuracy performance for $(N, K) = (64, 20)$. Our proposed JDU-BF outperforms other alternative schemes: it nearly attains the upper bound under the Ideal FL after 40 communication rounds and achieves an accuracy of $\sim 91\%$ at ~ 100 rounds. SDU-BF has a much slower model training convergence rate. After 100 rounds, it only nearly reaches 80% test accuracy. RBF exhibits the worst performance, where no training convergence is observed, and the accuracy is $\sim 10\%$ for all rounds. This is because that RBF provides no beamforming gain, leading to highly suboptimal communication performance, which affects the learning performance. Fig. 1-Middle shows the test accuracy for $(N, K) = (64, 40)$. We see that as the number of devices K increases from 20 to 40, the learning performance and, thus, the test accuracy of JDU-BF and SDU-BF improves. JDU-BF nearly attains the optimal performance after 30 rounds, while SDU-BF approaches the upper bound slowly and is slightly worse than JDU-BF after 100 rounds. The gain comes from the improved uplink over-the-air aggregation as the result of (distributed) transmit beamforming gain by more devices (*i.e.,* phase alignment via $a_{k,t}$).
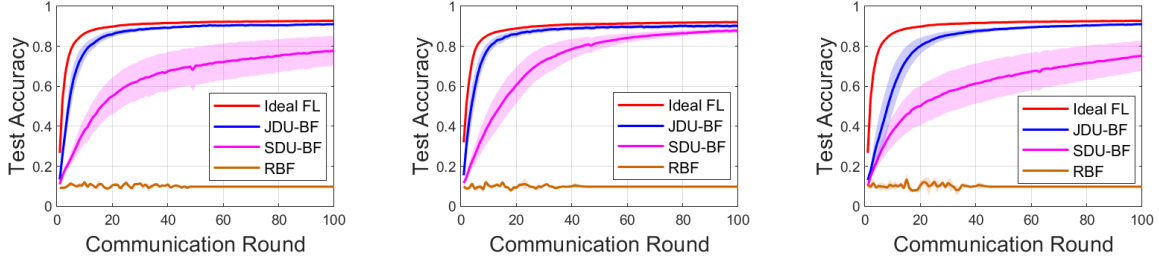
Fig. 1. Test accuracy vs. communication round T. Left: $N = 64, K = 20$. Middle: $N = 64, K = 40$. Right: $N = 16, K = 20$.

In particular, the improvement of SDU-BF over K is more noticeable. RBF is still the worst among all methods, with the test accuracy remaining at 10%, as it does not benefit from more devices since no beamforming gain can be collected.

Fig. 1-Right shows the case for $(N, K) = (16, 20)$, where $N < K$. Compared with Fig. 1-Left, both JDU-BF and SDU-BF perform worse as N reduces. This is expected due to reduced downlink and uplink beamforming gain with fewer antennas, impacting the overall learning performance of FL via wireless communication. Nonetheless, JDU-BF still nearly attains the upper bound after 100 rounds. In summary, our proposed JDU-BF is an effective communication scheme to facilitate FL in a wireless system for achieving fast training convergence and high test accuracy.

VI. Conclusion

In this paper, we have formulated the downlink-uplink transmission process for FL in a wireless system. We obtain the global model update in each round, capturing the impact of transmitter/receiver processing, receiver noise, and local training on the model update. Aiming to optimize downlink-uplink beamforming to maximize the FL training performance, we have derived an upper bound on the expected global loss after T rounds, and proposed an efficient JDU-BF algorithm to minimize this upper bound. JDU-BF is a low-complexity algorithm that uses the AO approach along with PGD to minimize the bound on the loss function per round. Simulation results show that JDU-BF outperforms other alternative schemes and provides a near-optimal learning performance for wireless FL.

Appendix A
Proof of Proposition 1

Proof: Substitute $\rho_{k,t} = \frac{\alpha_{k,t}^{\mathrm{ul}}}{\sum_{j=1}^{K} \alpha_{j,t}^{\mathrm{ul}}}$ with $\alpha_{k,t}^{\mathrm{ul}} = \sqrt{p_{k,t}} |\mathbf{h}_{k,t}^{H} \mathbf{w}_{t}^{\mathrm{ul}}|$ into (18)(20). We apply Lemmas 1 and 2 to (21). Let $M_t = \frac{\eta_t J}{2Q_t} \left(5(1 - Q_t) + 4\sqrt{1 - Q_t} - 1 \right)$ and $R_t = H(\mathbf{w}_t^{\mathrm{dl}}, \mathbf{w}_t^{\mathrm{ul}}, \mathbf{p}_t) + C_t$. For $\frac{1}{10L} \le \eta_t J < \frac{1}{2L}$, we have $Q_t > 0$ and $5(1 - Q_t) + 4\sqrt{1 - Q_t} - 1 > 0$. Thus, $M_t > 0$ and $R_t > 0$. Then, after combining Lemmas 1 and 2 and (21), we have

$$\mathbb{E}[F(\boldsymbol{\theta}_{t+1})] - F^\star \le \mathbb{E}[F(\boldsymbol{\theta}_t)] - F^\star + M_t \mathbb{E}[\|\nabla F(\boldsymbol{\theta}_t)\|^2] + R_t$$

$$\overset{(a)}{\le} \mathbb{E}[F(\boldsymbol{\theta}_t)] - F^\star + L^2 M_t \mathbb{E}[\|\boldsymbol{\theta}_t - \boldsymbol{\theta}^\star\|^2] + R_t$$

$$\overset{(b)}{\le} \mathbb{E}[F(\boldsymbol{\theta}_t)] - F^\star + \frac{2L^2 M_t}{\lambda} (\mathbb{E}[F(\boldsymbol{\theta}_t)] - F^\star) + R_t \quad (33)$$

where (a) and (b) follow from (2) and the L-smoothness and λ-strong-convexity of $F_k(\cdot)$ in Assumption 1, respectively. Summing up both sides of (33) over $t \in \mathcal{T}$ and rearranging the terms, we have (22). ∎

References

[1] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. Y. Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Proc. Int. Conf. Artif. Intell. Statist.*, Apr. 2017, pp. 1273–1282.

[2] G. Zhu, D. Liu, Y. Du, C. You, J. Zhang, and K. Huang, "Toward an intelligent edge: Wireless communication meets machine learning," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 19–25, Jan. 2020.

[3] M. Chen, D. Gündüz, K. Huang, W. Saad, M. Bennis, A. V. Feljan, and H. V. Poor, "Distributed learning in wireless networks: Recent progress and future challenges," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 12, pp. 3579–3605, Dec. 2021.

[4] 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, Mar. 2020.

[5] 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, Jan. 2020.

[6] N. Zhang and M. Tao, "Gradient statistics aware power control for over-the-air federated learning," *IEEE Trans. Wireless Commun.*, vol. 20, no. 8, pp. 5115–5128, Aug. 2021.

[7] 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, Jan. 2022.

[8] X. Fan, Y. Wang, Y. Huo, and Z. Tian, "Joint optimization of communications and federated learning over the air," *IEEE Trans. Wireless Commun.*, vol. 21, no. 6, pp. 4434–4449, Jun. 2022.

[9] M. M. Amiri, D. Gündüz, S. R. Kulkarni, and H. V. Poor, "Convergence of federated learning over a noisy downlink," *IEEE Trans. Wireless Commun.*, vol. 21, no. 3, pp. 1422–1437, Mar. 2022.

[10] X. Wei and C. Shen, "Federated learning over noisy channels: Convergence analysis and design examples," *IEEE Trans. Cogn. Commun.*, vol. 8, no. 2, pp. 1253–1268, Jun. 2022.

[11] W. Guo, R. Li, C. Huang, X. Qin, K. Shen, and W. Zhang, "Joint device selection and power control for wireless federated learning," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 8, pp. 2395–2410, Aug. 2022.

[12] 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, Aug. 2022.

[13] E. S. Levitin and B. T. Polyak, "Constrained minimization methods," *USSR Comput. Math. Math. Phys.*, vol. 6, no. 5, pp. 787–823, 1966.

[14] S. Bubeck, "Convex optimization: Algorithms and complexity," *Found. Trends Mach. Learn.*, vol. 8, no. 3-4, pp. 231–357, 2015.

[15] S. Ghadimi, G. Lan, and H. Zhang, "Mini-batch stochastic approximation methods for nonconvex stochastic composite optimization," *Math. Program.*, vol. 155, no. 1-2, pp. 267–305, 2016.

[16] C. Zhang, M. Dong, and B. Liang, "Fast first-order algorithm for large-scale max-min fair multi-group multicast beamforming," *IEEE Wireless Commun. Lett.*, vol. 11, no. 8, pp. 1560–1564, Aug. 2022.

[17] M. Dong and Q. Wang, "Multi-group multicast beamforming: Optimal structure and efficient algorithms," *IEEE Trans. Signal Process.*, vol. 68, pp. 3738–3753, May 2020.

[18] Y. LeCun, C. Cortes, and C. Burges. The MNIST Database of Handwritten Digits. 1998. [Online]. Available: http://yann.lecun.com/exdb/mnist/.