IEEE INFOCOM 2024 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS) | 979-8-3503-8447-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/INFOCOMWKSHPS61880.2024.10620738

Federated Learning Based Integrated Sensing, Communications, and Powering Over 6G Massive-MIMO Mobile Networks

Xi Zhang and Qixuan Zhu

Networking and Information Systems Laboratory
Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843, USA
E-mail: {*xizhang@ece.tamu.edu*, *qixuan@tamu.edu*}

***Abstract*—To sense mobile users' positions and channel states while simultaneously conveying information for supporting the 6G traffic, the *integrated sensing and communication (ISAC)* has been proposed as a key technique for sensing and communicating using the same radio frequency and hardware. Moreover, to support energy-constrained and battery-limited wireless networks, *simultaneous wireless information and power transfer* (SWIPT) has also emerged to simultaneously deliver information and energy to a receiver. To integrate these two techniques, *integrated sensing, communications, and powering* (ISACP) has been developed to simultaneously sense the targeted mobile users, transmit the information, and deliver the power. However, how to efficiently optimize the power allocation in ISACP schemes to support the 6G traffic has imposed many new challenges not encountered before. To overcome these challenges, in this paper we propose a federated-learning (FL) enabled ISACP scheme to maximize the energy-efficiency over 6G massive MIMO wireless networks. First, we establish the system models for our proposed ISACP and FL schemes. Second, we design the training data set collection in the FL algorithm by sending the pilot signal from mobile users to the massive MIMO base station, and formulate an energy-efficiency maximization problem for the developed ISACP scheme. Third, To solve the formulated energy-efficiency maximization problem, we propose an actor-critic enabled multi-agent based FL mechanism. Finally, we use the numerical analyses to validate and evaluate our developed schemes.**

***Index Terms*—6G wireless networks, integrated sensing, communications, and powering (ISACP), federated learning, massive MIMO, channel state estimation, energy-efficiency.**

I. Introduction

The massive ultra-reliable and low-latency communications (mURLLC) services are emerging as a new but dominant traffic type of the 6G mobile wireless networks that support a massive number of mobile users (MUs) demanding the stringent quality of service (QoS) such as high data rate and low error probability. The integrated sensing and communications (ISAC) has been proposed to fulfill both channel/position-sensing and data-communication functionalities by using the same spectrum and hardware for improving various QoS performances of 6G traffics. To address the issue of limited battery capacity for MUs, simultaneous wireless information and power transfer (SWIPT) has been proposed as a promising solution to improve the energy efficiency by MUs applying the energy harvesting (EH) techniques for receiving the energy and information from radio-frequency (RF) signals. Due to their advancements, researchers envision that 6G wireless networks will integrate ISAC with SWIPT to provide new multi-functional wireless systems with integrated sensing, communications, and powering (ISACP) to simultaneously sense the wireless channel, and transfer the information and power.

Recent works have studied various ISAC, SWIPT, and ISACP schemes. The work of [1] designed an accurate ISAC-enabled beam tracking scheme in high-mobility vehicular networks to establish the reliable communication link. The authors of [2] investigated the application of ISAC in space-air-ground-sea integrated network using massive multiple-input multiple-output (MIMO) communication, through proposing a beam squint-aware ISAC technique for hybrid analog/digital massive MIMO enabled low earth orbit (LEO) satellite systems. The work of [3] maximized the sum energy efficiency over all device-to-device (D2D) links in a D2D underlaid cellular network by optimizing the resource and power allocation based on a nonlinear EH model. The authors of [4] addressed the secure communications over EH cooperative cognitive radio networks by proposing the polarization-enabled two-phase mechanism. The ISACP scheme has been proposed in [5] by allowing a multi-antenna hybrid access point to transmit wireless signals to communicate with a multi-antenna information decoding receiver, wirelessly charging a multi-antenna EH receiver and performing radar target sensing based on the echo signal concurrently.

However, due to the time-varying of the wireless channel and dynamic states for MUs, how to efficiently optimize the ISACP scheme, while satisfying the stringent QoS requirements for 6G traffics has not been thoroughly studied. To overcome this challenge, in this paper we propose a federated learning (FL)-base ISACP mechanism to support the QoS provisioning by using massive MIMO communications. First, we establish the system models for our proposed ISACP and FL schemes to estimate channel states through the radar echo. Considering the computational complexity of channel estimations, we then design the training data set collection for

This work of Xi Zhang and Qixuan Zhu was supported in part by the U.S. National Science Foundation under Grants CCF-2142890, CCF-2008975, ECCS-1408601, and CNS-1205726, and the U.S. Air Force under Grant FA9453-15-C-0423.

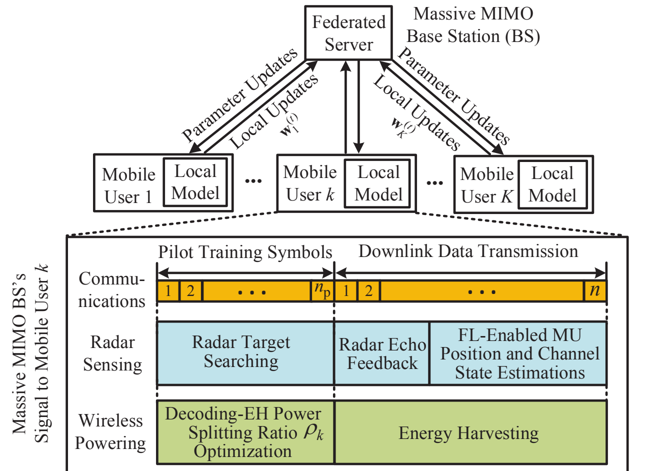

Fig. 1. The system architecture models of our proposed federated learning based channel estimation using massive MIMO base station.

the FL algorithm by sending the pilot signal from MUs to the massive MIMO base station (BS) and formulate an energy-efficiency maximization problem for the developed ISACP scheme. Third, to solve the formulated energy-efficiency maximization problem, we also propose an actor-critic enabled multi-agent based FL mechanism.

The rest of this paper is organized as follows. Section II establishes the system models for our proposed ISACP architectures. Section III develops the FL algorithm for estimating the channel state information at the receiver (CSIR). Section IV proposes a multi-agent based FL algorithm to maximize the energy-efficiency of our proposed ISACP scheme through estimating the channel state information at the transmitter (CSIT). Section V validates and evaluates our developed ISACP schemes. This paper concludes with Section VI.

II. System Models for Our Proposed ISACP-Based Architectures Using Massive MIMO

A. Our Proposed ISACP-Based Architectures Using Massive MIMO For Estimating CSIT

As shown in Fig. 1, we consider a cellular network consisting of a massive MIMO BS and totally K moving targeted MUs, where the massive MIMO BS sends the radar sensing and the downlink communication signals simultaneously to these K targeted MUs. Assume that there are M_T antennas on the BS and there are M_R antennas on each targeted MU, where $M_T \gg M_R$. The kth MU, $\forall k$, splits the received RF signals into two parts with decoding-EH power splitting ratios ρ_k and $(1-\rho_k)$ for decoding the information and energy harvesting, respectively, where $0<\rho_k<1$.

Let $B_k, \forall k$, be the total transmit power allocation of the sensing and communication signal (i.e., ISAC signal) sending from the massive MIMO BS to the kth MU. The massive MIMO BS transmits the ISAC signals to all K MUs, which are assumed to be EH devices. Each MU k, $\forall k$, receives the ISAC signal, decodes the received information with the energy $\rho_k B_k$, and harvests the rest $(1-\rho_k)B_k$ energy. The echo of the transmitted ISAC signal is then reflected back to the massive MIMO BS. By estimating the kth MU's angle-of-arrival (AoA), denoted by $\phi_k, \forall k$, using the radar echo, the massive MIMO BS is able to obtain the kth MU's position. Assume that the wireless fading channels between the massive MIMO BS and MUs only depends on the multipaths which are determined by MUs' positions. Thus, by estimating the MUs' positions (i.e., AoAs), the massive MIMO BS is able to obtain the *channel state information at the transmitter (CSIT)* to satisfy MUs' stringent QoS requirements.

B. Our Proposed Federated Learning Based Algorithm For Estimating CSIR

In addition to the CSIT estimated by the massive MIMO BS, each MU also needs to estimate its *channel state information at the receiver (CSIR)*. Due to the constrained MU's local computation resources and the computational complexity for estimating MUs' positions and channel states, we propose to apply the FL based algorithm for MUs' position/channel-state estimation. As a decentralized machine learning approach, a local model of the FL can be trained in each of these K MUs using local data samples without exchanging them. We propose that MUs train their model using pilot signals, which are known by both BS and MUs, as local data samples.

III. Federated Learning Based CSIR Estimation

In our proposed FL based algorithm, the local model at each MU agent is being initialized by collecting the data set from its own MU and training its own local model to obtain the learnable parameter by using its own data set. We define the *learnable parameter* as the parameter to solve the estimated channel state, and the local model aims at minimizing the loss function of each MU's channel state estimation under the learnable parameter. Each MU agent computes the gradient of the local loss function, updates its local model's parameters, and sends its updated model gradient to the central server located at the massive MIMO BS. The central server aggregates the received model parameters using a predefined federated averaging (FedAvg) approach [6] and updates the global model, whose parameter is then sent back to all MU agents. Each MU agent receives the updated parameter of the global model and incorporates it into its local model to update their model parameters. The FL based algorithm trains the global model without explicitly exchanging the data, which is advantageous in terms of privacy considerations and communication resources. We treat each MU as a learning agent and deploy the central server at the massive MIMO BS, and thus, our proposed FL consists of K MU agents and one central server.

A. Training Data Set Collection Using Pilot Signals For Estimating CSIR

Let $\mathbf{s}_k^{\text{com}} \in \mathbb{C}^{M_R \times 1}$ be the actual channel state for the channel between the massive MIMO BS and the kth MU and let $\widehat{\mathbf{s}}_k^{\text{com}}$ be CSIR for $\mathbf{s}_k^{\text{com}}$ estimated the kth MU. Denote the learnable parameter by $\mathbf{w}_k^{(t)}$ at the tth iteration of the FL based model for the kth MU agent, $\forall k \in \{1, 2, \cdots, K\}$. Denote the local data set of the kth MU agent by $\mathcal{D}_k$ and assume that $\mathcal{D}_k \cap \mathcal{D}_{\widetilde{k}} = \emptyset$ if $k \neq \widetilde{k}$. The local data set $\mathcal{D}_k$ contains the input-output pairs $\mathcal{D}_k^{(\nu)} = \left(\mathbf{p}_k^{(\nu)}, b_k^{(\nu)}\right), \forall k$, and $\forall \nu \in \{1, 2, \cdots, |\mathcal{D}_k|\}$, where $\mathbf{p}_k^{(\nu)} \triangleq \left[p_k^{(1,\nu)}, \cdots, p_k^{(n_\mathrm{p},\nu)}\right]^\intercal \in \mathbb{C}^{n_\mathrm{p} \times 1}, \forall \nu$, is the input data representing the pilot signals sent from the massive MIMO BS to the kth MU agent and is known by both the MU agent and the central server located at the massive MIMO BS, and $b_k^{(\nu)}, \forall \nu$, is the output/label data, which is the estimated channel state $\widehat{\mathbf{s}}_k^{\mathrm{com}}$ by the kth MU. The objective of the FL based model is to obtain an optimal learnable parameter, denoted by $\mathbf{w}^*$, such that $\Lambda\big(\mathbf{p}_k^{(\nu)}\big|\mathbf{w}^*\big) = b_k^{(\nu)}$.

We assign the massive MIMO BS with an input data set (i.e., pilot signals) $\mathbf{p}_k$, where we select the input data set such that $\big(\mathbf{p}_k^{(\nu)}\big)^H \mathbf{p}_k^{(\nu)} = 1$. The massive MIMO BS sends the signal, denoted by $\mathbf{x}_k^{(\mathrm{p},\nu)} \in \mathbb{C}^{1 \times n_\mathrm{p}}$ to the kth MU agent, which is given as follows:

$$\mathbf{x}_k^{(\mathrm{p},\nu)} = \sqrt{n_\mathrm{p} B^{\mathrm{dl}}} \big(\mathbf{p}_k^{(\nu)}\big)^H, \tag{1}$$

where B^{dl} is the transmit power for the input data. Then the kth MU's received signal, denoted by $\mathbf{Y}_k^{(\mathrm{p},\nu)} \in \mathbb{C}^{M_R \times n_\mathrm{p}}$, from the kth MU agent is given as follows:

$$\mathbf{Y}_k^{(\mathrm{p},\nu)} = \mathbf{s}_k^{\mathrm{com}} \mathbf{x}_k^{(\mathrm{p},\nu)} + \mathbf{N}_k^{\mathrm{com}} \tag{2}$$

where $\mathbf{N}_k^{\mathrm{com}} \in \mathbb{C}^{M_R \times n_\mathrm{p}}$ is the additive white Gaussian noise (AWGN) for the kth MU agent. The kth MU agent performs a de-spreading operation by correlating the received signals with the kth input data set, which yields the kth MU agent's received signal after de-spreading, denoted by $\widetilde{\mathbf{Y}}_k^{(\mathrm{p},\nu)} \in \mathbb{C}^{M_R \times 1}$, as follows:

$$\begin{aligned}\widetilde{\mathbf{Y}}_k^{(\mathrm{p},\nu)} &= \mathbf{Y}_k^{(\mathrm{p},\nu)} \mathbf{p}_k^{(\nu)} = \sqrt{n_\mathrm{p} B^{\mathrm{dl}}} \mathbf{s}_k^{\mathrm{com}} \big(\mathbf{p}_k^{(\nu)}\big)^H \mathbf{p}_k^{(\nu)} + \mathbf{N}_k^{\mathrm{com}} \mathbf{p}_k^{(\nu)} \\ &= \sqrt{n_\mathrm{p} B^{\mathrm{dl}}} \mathbf{s}_k^{\mathrm{com}} + \mathbf{N}_k^{\mathrm{com}} \mathbf{p}_k^{(\nu)}. \end{aligned} \tag{3}$$

Since the channel estimation $\widehat{s}_{k,l}^{\mathrm{com}}$, which is the lth element of $\widehat{\mathbf{s}}_k^{\mathrm{com}}$, under the minimum mean square error (MMSE) estimation is given by

$$\widehat{s}_{k,l}^{\mathrm{com}} = \mathbb{E}\left[s_{k,l}^{\mathrm{com}} \middle| \widetilde{\mathbf{Y}}_k^{(\mathrm{p},\nu)}, \mathbf{w}_k^{(t)} \right] \triangleq \Lambda\left(\mathbf{p}_k^{(\nu)} \middle| \mathbf{w}_k^{(t)} \right) \tag{4}$$

where the expectation operation is taken under the multipath fading, the kth MU agent trains $\mathbf{w}_k^{(t)}$ as the parameter to derive the multipath fading factor which affects the expectation operation in Eq. (4). Thus, $\mathbf{w}_k^{(t)}$ is the parameter such that

$$\mathbf{w}_k^{(t)} = \arg\min \left\{ \Lambda\left(\mathbf{p}_k^{(\nu)} \middle| \mathbf{w}_k^{(t)} \right) - b_k^{(\nu)} \right\} \tag{5}$$

where $b_k^{(\nu)}$ is given by the local data set $\mathcal{D}_k^{(\nu)} = \left(\mathbf{p}_k^{(\nu)}, b_k^{(\nu)}\right)$.

B. Federated Learning Parameter Updating

We specify that each communication round between the central server at the massive MIMO BS and MU agents in the tth iteration of the FL consists of four major steps as shown in the followings.

Step 1: *Local Training*. Each MU agent k trains a local model to minimize a local loss function $L_k\left(\mathbf{w}_k^{(t)}, \mathcal{D}_k\right)$ by using its own local data set $\mathcal{D}_k$. During this local training, each agent k obtains the weight of its local model $\alpha_k^{(t)}$ and the gradient of the local loss function $\nabla L_k(\mathbf{w}_k^{(t)}, \mathcal{D}_k)$.

Define a loss function $L_k\left(\mathbf{w}_k^{(t)}, \mathcal{D}_k\right), \forall k$, as follows:

$$L_k\left(\mathbf{w}_k^{(t)}, \mathcal{D}_k\right) \triangleq \frac{1}{|\mathcal{D}_k|} \sum_{\nu=1}^{|\mathcal{D}_k|} \left| \Lambda\left(\mathbf{p}_k^{(\nu)} \middle| \mathbf{w}_k^{(t)} \right) - b_k^{(\nu)} \right|^2. \tag{6}$$

The FL-based model training is performed at all MU agents as follows:

$$\min_{\mathbf{w}_k^{(t)}} \left\{ \overline{L}\left(\mathbf{w}_k^{(t)}, \mathcal{D}\right) \right\} \triangleq \min_{\mathbf{w}_k^{(t)}} \left\{ \sum_{k=1}^{K} \alpha_k^{(t)} L_k\left(\mathbf{w}_k^{(t)}, \mathcal{D}_k\right) \right\} \tag{7}$$

where $\mathcal{D} = \{\mathcal{D}_1, \cdots, \mathcal{D}_K\}$, $\alpha_k^{(t)}$ is the weight of the kth MU agent at the tth iteration, and $\sum_{k=1}^{K} \alpha_k^{(t)} = 1$. Each MU agent k computes the gradient $\nabla L_k\big(\mathbf{w}_k^{(t)}, \mathcal{D}_k\big)$ to solve Eq. (7), where the gradient operation is taken with respect to $\mathbf{w}_k^{(t)}$.

Step 2: *Global Aggregation*. Each MU agent k uploads its obtained weight $\alpha_k^{(t)}$ and gradient $\nabla L_k\big(\mathbf{w}_k^{(t)}, \mathcal{D}_k\big)$ to the central server. The central server aggregates weights and gradients received from all MU agents, updates the weight for each MU agent $\alpha_k^{(t+1)}$ for the next iteration, and computes the average gradient $\nabla\overline{L}$ as follows:

$$\nabla\overline{L} = \sum_{k=1}^{K} \alpha_k^{(t+1)} \nabla L_k\left(\mathbf{w}_k^{(t)}, \mathcal{D}_k\right). \tag{8}$$

Step 3: *Model Updating*. The central server uses the updated weight $\alpha_k^{(t+1)}$ and average gradient $\nabla\overline{L}$ to update the parameters for the global model to improve the accuracy of the global model. The updated global model with these updated parameters are then sent back to all MU agents for updating their local models. Then MU agents update their learnable parameter $\mathbf{w}_k^{(t+1)}$ for the next iteration as follows:

$$\mathbf{w}_k^{(t+1)} = \mathbf{w}_k^{(t)} - \eta \sum_{k=1}^{K} \alpha_k^{(t+1)} \nabla\overline{L}, \quad \forall k, \tag{9}$$

where η is the step size.

Step 4: *Control and Feedback*. Each MU agent periodically communicates with the central server to upload and update the parameters of its local model. The central server sends the control and feedback information, e.g., new learning rate schedules and regularization parameters of the global model, to all MU agents, so that each agent is able to improve its local model.

IV. Energy-Efficiency Maximization for ISACP Through Estimating CSIT Using Multi-Agent Based Algorithm

A. The ISAC Signal Sent by the Massive MIMO BS

For the massive MIMO downlink integrated sensing and communications, define $\mathbf{B} \triangleq [B_1, B_2, \cdots, B_K]^\intercal$ as the power allocation vector for all K MUs, where B_k is the total power allocation for the kth MU's sensing and communication signal. Let $\mathbf{x}_i(\mathbf{B}) \in \mathbb{C}^{M_T \times 1}$ be the transmitted signal under the total power allocation vector $\mathbf{B}$ for the ith symbol, $\forall i \in \{1, \cdots, n\}$, from the massive MIMO BS to all K MUs, which is given by [7, Eq. (1)] [8, Eq. (1)]

$$\mathbf{x}_i(\mathbf{B}) = \mathbf{U}^{\text{com}}(\mathbf{B})\mathbf{q}_i^{\text{com}} + \mathbf{U}^{\text{sen}}(\mathbf{B})\mathbf{q}_i^{\text{sen}} \tag{10}$$

where

$$\begin{cases} \mathbf{U}^{\text{com}}(\mathbf{B}) = \left[\sqrt{\dfrac{\rho_1 B_1}{2M_T}}\mathbf{U}_1^{\text{com}}, \cdots, \sqrt{\dfrac{\rho_K B_K}{2M_T}}\mathbf{U}_K^{\text{com}}\right] \\ \mathbf{U}^{\text{sen}}(\mathbf{B}) = \left[\sqrt{\dfrac{\rho_1 B_1}{2M_T}}\mathbf{U}_1^{\text{sen}}, \cdots, \sqrt{\dfrac{\rho_K B_K}{2M_T}}\mathbf{U}_K^{\text{sen}}\right] \end{cases} \tag{11}$$

are the communication and sensing signals beamforming matrixes (i.e., precoders) for all K targeted MUs, respectively, under the total power allocation vector $\mathbf{B}$; $\mathbf{U}_k^{\text{com}} \in \mathbb{C}^{M_T \times M_R}$ and $\mathbf{U}_k^{\text{sen}} \in \mathbb{C}^{M_T \times M_R}$ are the communication and sensing precoders for the kth targeted MU, respectively; and $\mathbf{q}_i^{\text{com}} \in \mathbb{C}^{KM_R \times 1}$ and $\mathbf{q}_i^{\text{sen}} \in \mathbb{C}^{KM_R \times 1}$ are the ith communication symbol and the ith probing sensing symbol, respectively. Since all MUs are EH devices, the kth MU, $\forall k$, uses ρ_k portion of the energy to receive information and harvests the rest $(1-\rho_k)$ portion of the energy. Denote by $\widetilde{\mathbf{B}} \triangleq [\rho_1 B_1, \rho_2 B_2, \cdots, \rho_K B_K]^\intercal$ the vector for the signal power after MUs' energy harvesting. After MUs' energy harvesting, the reflected radar echo signal for the ith symbol under the signal power vector $\widetilde{\mathbf{B}}$, denoted by $\mathbf{y}_i^{\text{sen}}(\widetilde{\mathbf{B}}) \in \mathbb{C}^{M_T \times 1}$, received by the massive MIMO BS is given by

$$\mathbf{y}_i^{\text{sen}}(\widetilde{\mathbf{B}}) = \mathbf{S}^{\text{sen}}\mathbf{x}_i(\widetilde{\mathbf{B}}) + \mathbf{v}^{\text{sen}}, \ \forall i \in \{1, 2, \cdots n\} \tag{12}$$

where $\mathbf{S}^{\text{sen}} \in \mathbb{C}^{M_T \times M_T}$ is the sensing signal channel, $\mathbf{x}_i(\widetilde{\mathbf{B}}) \in \mathbb{C}^{M_T \times 1}$ can be obtained by replacing $\mathbf{B}$ in $\mathbf{x}_i(\mathbf{B})$ of Eq. (10) by $\widetilde{\mathbf{B}}$, $\mathbf{v}^{\text{sen}} \in \mathbb{C}^{M_T \times 1}$ is AWGN following a complex Gaussian distribution $\mathcal{CN}(\mathbf{0}, \sigma_w^2 \mathbf{I}_{M_T})$ at all M_T antennas deployed in the massive MIMO BS, where σ_w^2 is the variance of any element in $\mathbf{v}^{\text{sen}}$ and $\mathbf{I}_{M_T}$ is the identity matrix of size M_T, and n is the length of a data block.

Using the obtained $\mathbf{y}_i^{\text{sen}}(\widetilde{\mathbf{B}})$, the massive MIMO BS is able to update all K MUs' positions, represented by their estimated AoA vectors, denoted by $\widehat{\boldsymbol{\phi}}(\widetilde{\mathbf{B}}) \triangleq \left[\widehat{\phi}_1(\widetilde{\mathbf{B}}), \cdots, \widehat{\phi}_K(\widetilde{\mathbf{B}})\right]$, where $\widehat{\phi}_k(\widetilde{\mathbf{B}}), \forall k \in \{1, \cdots, K\}$, is the estimated AoA for the kth MU as a function of the power after energy harvesting $\widetilde{\mathbf{B}}$. Applying $\widehat{\boldsymbol{\phi}}(\widetilde{\mathbf{B}})$, the massive MIMO BS obtains the estimated CSIT, denoted by $\widehat{\mathbf{S}}^{\text{com}}\left(\widehat{\boldsymbol{\phi}}(\widetilde{\mathbf{B}})\right) \in \mathbb{C}^{KM_R \times M_T}$, of all K MUs.

Under the ISACP scheme with massive MIMO communication, the received ISAC signal for the ith symbol ($\forall i \in \{1, 2, \cdots, n\}$) through communication channels at all antennas on K MUs, denoted by $\mathbf{y}_i^{\text{com}}\left(\widehat{\boldsymbol{\phi}}(\widetilde{\mathbf{B}})\right) \in \mathbb{C}^{KM_R \times 1}$, is given by

$$\mathbf{y}_i^{\text{com}}\left(\widehat{\boldsymbol{\phi}}(\widetilde{\mathbf{B}})\right) = \mathbf{S}^{\text{com}}\mathbf{x}_i(\widetilde{\mathbf{B}}) + \mathbf{v}^{\text{com}} \tag{13}$$

where $\mathbf{S}^{\text{com}} \in \mathbb{C}^{KM_R \times M_T}$ is the communication channel state matrix, representing the channel state between all antennas of the kth targeted MU and all antennas on the massive MIMO BS, and $\mathbf{v}^{\text{com}} \in \mathbb{C}^{KM_R \times 1}$ is the AWGN at all antennas of all MUs. The mth element of $\mathbf{y}_i^{\text{com}}(\widehat{\boldsymbol{\phi}}(\widetilde{\mathbf{B}}))$, denoted by $y_{m,i}^{\text{com}}\left(\widehat{\boldsymbol{\phi}}(\widetilde{\mathbf{B}})\right)$, representing the received signal on the mth antenna of all K MUs, is given by

$$y_{m,i}^{\text{com}}\left(\widehat{\boldsymbol{\phi}}(\widetilde{\mathbf{B}})\right) = \sqrt{\frac{\rho_k B_k}{2M_T}}\widehat{\mathbf{s}}_m^{\text{com}}\left(\widehat{\boldsymbol{\phi}}(\widetilde{\mathbf{B}})\right)\mathbf{u}_m^{\text{com}} q_{m,i}^{\text{com}} + \Omega_{m,i}\left(\widehat{\boldsymbol{\phi}}(\widetilde{\mathbf{B}})\right) \tag{14}$$

where $\forall m \in \{1, \ldots, KM_R\}$ with assuming that m is the index for the antenna belonging to the kth targeted mobile user, $\widehat{\mathbf{s}}_m^{\text{com}}\left(\widehat{\boldsymbol{\phi}}(\widetilde{\mathbf{B}})\right) \in \mathbb{C}^{1 \times M_T}$ is the mth row of the matrix $\widehat{\mathbf{S}}^{\text{com}}\left(\widehat{\boldsymbol{\phi}}(\widetilde{\mathbf{B}})\right)$, $\mathbf{u}_m^{\text{com}} \in \mathbb{C}^{M_T \times 1}$ is the mth column of the matrix $\mathbf{U}^{\text{com}}(\widetilde{\mathbf{B}}^{\text{com}})$, $q_{m,i}^{\text{com}}$ and v_m^{com} are the mth element of $\mathbf{q}_i^{\text{com}}$ and $\mathbf{v}^{\text{com}}$, respectively, and $\Omega_{m,i}\left(\widehat{\boldsymbol{\phi}}(\widetilde{\mathbf{B}})\right)$ is the *effective additive noise* on the mth antenna of all MUs, including the AWGN noise and negligible inter-user interference.

B. Energy-Efficiency Maximization for Our Proposed ISACP Schemes

We propose to maximize the energy-efficiency for our proposed ISACP schemes. Define the power splitting ratio vector, denoted by $\boldsymbol{\rho}$, for all K MUs as follows:

$$\boldsymbol{\rho} \triangleq [\rho_1, \rho_2, \cdots, \rho_K]^\intercal. \tag{15}$$

Denote by $B(\boldsymbol{\rho})$ the total power consumption (Joule/sec) for the ISACP-based massive MIMO communications system under the power splitting ratio vector $\boldsymbol{\rho}$, which is given by:

$$\begin{aligned} B(\boldsymbol{\rho}) = & B_{\text{T}} + KB_{\text{R}} + \sum_{k=1}^{K}\rho_k B_k - \sum_{k=1}^{K}(1-\rho_k)B_k \\ & - \sum_{k=1}^{K}\Gamma_k(\rho_k) - \sum_{k=1}^{K}\sum_{\substack{m= \\ (k-1)M_R+1}}^{kM_R}\Omega_{m,i}\left(\widehat{\boldsymbol{\phi}}(\widetilde{\mathbf{B}})\right) \end{aligned} \tag{16}$$

where B_{T} is a constant signal processing circuit power consumption in the massive MIMO BS, B_{R} is a constant circuit power consumption of each MU, $\Gamma_k(\rho_k)$ is the energy harvested from the information signal at the kth MU, derived from Eq. (14), given by:

$$\Gamma_k(\rho_k) = \sum_{m=(k-1)M_R+1}^{kM_R}\frac{\rho_k B_k}{2M_T}\text{Var}\left[\widehat{\mathbf{s}}_m^{\text{com}}\left(\widehat{\boldsymbol{\phi}}(\widetilde{\mathbf{B}})\right)\mathbf{u}_m^{\text{com}} q_{m,i}^{\text{com}}\right], \tag{17}$$

and $\Omega_{m,i}\left(\widehat{\boldsymbol{\phi}}(\widetilde{\mathbf{B}})\right)$ in Eq. (16) is defined following Eq. (14), representing the energy harvested from the effective additive noise by the mth antenna of all MUs.

We define the overall *energy-efficiency* (bits/Joule/Hz), denoted by $E(\boldsymbol{\rho})$, for the ISACP-based massive MIMO communications system by a function of power splitting ratio vector $\boldsymbol{\rho}$ defined in Eq. (15) as the total number of bits successfully conveyed to all K MUs per Joule consumed energy under the QoS requirements for 6G traffics, which is given by

$$E(\boldsymbol{\rho}) \triangleq \frac{\sum_{k=1}^{K} C_k(\rho_k B_k)}{B(\boldsymbol{\rho})} \tag{18}$$

where $C_k(\rho_k B_k)$ is the kth MU's channel capacity given by:

$$C_k(\rho_k B_k) = \sum_{\substack{m= \\ (k-1)M_R+1}}^{kM_R} \log_2\left(1+\frac{\rho_k \beta_k B_k \mathrm{Var}\left[\widehat{\mathbf{s}}_m^{\mathrm{com}}\left(\widehat{\boldsymbol{\phi}}(\widetilde{\mathbf{B}})\right)\mathbf{u}_m^{\mathrm{com}} q_{m,i}^{\mathrm{com}}\right]}{2M_T \mathrm{Var}\left[\Omega_{m,i}\left(\widehat{\boldsymbol{\phi}}(\widetilde{\mathbf{B}})\right)\right]}\right) \tag{19}$$

and $B(\boldsymbol{\rho})$ is the total power consumption given by Eq. (16). Then, we maximize the energy-efficiency $E(\boldsymbol{\rho})$ as follows:

$$\max_{\boldsymbol{\rho}}\{E(\boldsymbol{\rho})\} \tag{20}$$

$$\text{s.t.: } \mathrm{C1}: \mathbb{E}\left[\left(\widehat{\phi}_k(\widetilde{\mathbf{B}})-\phi_k\right)^2\right] \leq \epsilon_{\max}, \forall k$$

$$\mathrm{C2}: C_k(\rho_k B_k) \geq C_{\min}, \forall k$$

where $\epsilon_{\max}$ in C1 is the maximum tolerable AoA estimation error, measuring the sensing performance and $C_k(\rho_k B_k)$ in C2 is the channel capacity for the kth MU, measuring the communication performance. Solving Eq. (20), we can obtain the optimal power splitting ratio vector, denoted by $\boldsymbol{\rho}^* \triangleq [\rho_1^*, \cdots, \rho_K^*]$, for our proposed ISACP scheme.

Observing the objective function specified by Eq. (20) and Eqs. (18)-(19), we obtain that the energy-efficiency is a function of $\widehat{\mathbf{s}}_m^{\mathrm{com}}\left(\widehat{\boldsymbol{\phi}}(\widetilde{\mathbf{B}})\right)$, which makes it difficult to derive the closed-form solution for Eq. (20), we propose to use the machine learning technique to obtain $\widehat{\mathbf{s}}_m^{\mathrm{com}}\left(\widehat{\boldsymbol{\phi}}(\widetilde{\mathbf{B}})\right)$ and solve Eq. (20). Observing the large number of MUs in the ISACP network, we can obtain $\widehat{\mathbf{s}}_m^{\mathrm{com}}\left(\widehat{\boldsymbol{\phi}}(\widetilde{\mathbf{B}})\right)$ by applying the *multi-agent based federated learning* (MAFL) based algorithm, which enables to train a learning model without extracting the data directly from the MUs. Instead of imposing the huge communication overhead in the conventional centralized learning approach, each agent using the MAFL approach sends only gradients of the model parameters to the central server, generating a more rapid response for optimal solution.

C. The Actor-Critic Algorithm Based Multi-Agent Federated Learning Mechanism

We then describe the maximization of ISACP-based energy-efficiency given by Eq. (20) through *actor-critic algorithm* [9], which integrates the actor-only method with critic-only method of machine learning. In actor-critic algorithm, the actor decides which action optimizes the objective function given by Eq. (20) according to the current state. The critic evaluates this action through a value function, and then, informs the actor how good the action is and how to improve the action.

Each MU agent k uses the *Markov decision process (MDP)* to solve Eq. (20). We define the finite state set, denoted by $\mathcal{Z}$, to characterize the MDP's *states* describing the channel feedback obtained by the radar echo. The channel feedback $\mathbf{Z}(t)$ is defined as the received radar echo $\mathbf{Z}(t) \triangleq \mathbf{y}_i^{\mathrm{sen}}(\widetilde{\mathbf{B}})$. Define $\mathbf{A}(t) \triangleq [A_1(t), A_2(t), \ldots, A_{KM_R}(t)] \in \widehat{\mathcal{S}}^{\mathrm{com}}$, where $\widehat{\mathcal{S}}^{\mathrm{com}}$ is the *action* set of this MDP for all MUs, and $A_m(t) \triangleq \widehat{\mathbf{s}}_m^{\mathrm{com}}\left(\widehat{\boldsymbol{\phi}}(\widetilde{\mathbf{B}})\right)$ with $m \in \{1, 2, \ldots, KM_R\}$ denotes the channel estimation at time t for the mth antenna of all MUs. We also denote the actor parameter by τ over the parameter set $\mathcal{T}$, and denote the critic parameter by q. We then define the *policy* π_k of the MU agent k as a mapping function $\pi_k : \mathcal{Z} \times \mathcal{T} \to \widehat{\mathcal{S}}^{\mathrm{com}}$ that assigns each state-parameter pair $(\mathbf{Z}(t), \tau) \in \mathcal{Z} \times \mathcal{T}$ to the channel estimation schemes $\mathbf{A}(t) \in \widehat{\mathcal{S}}^{\mathrm{com}}$. Define the *reward* r_t of the actor-critic algorithm at time t as the gain of the *ISACP-based energy-efficiency* $E(\boldsymbol{\rho})$, calculated by solving Eq. (20). We also define the *value* of the current policy under the critic parameter q as $Q_q(\mathbf{Z}(t), \mathbf{A}(t))$. This actor-critic algorithm is summarized in **Algorithm 1**, where $\widetilde{\eta}$ represents the importance of future rewards, and ϕ_τ and ϕ_q measure how quickly the local model learns for the actor and the critic, respectively.

Algorithm 1 Actor-Critic Based Algorithm for Maximizing Aggregate ISACP-Based Energy-Efficiency

1: **Initialize:** Actor parameter τ, critic parameter q, $\mathbf{Z}(0)$, learning rates ϕ_τ and ϕ_q.
2: **for** time slot t in 1, 2, ... **do**
3: Each MU agent k selects an action $\mathbf{A}(t)$ from action set $\widehat{\mathcal{S}}^{\mathrm{com}}$ according to policy $\pi_k(\mathbf{Z}(t), \tau)$.
4: Use the current state to obtain a reward $r_t = E(\boldsymbol{\rho})$ by solving Eq. (20).
5: Obtain the next state $\mathbf{Z}(t+1)$ and the next action $\mathbf{A}(t+1)$.
6: Update the actor parameter $\tau \leftarrow \tau + \phi_\tau Q_q(\mathbf{Z}(t), \mathbf{A}(t)) \nabla_\tau \log \pi_k(\mathbf{Z}(t), \tau)$.
7: Compute the correction for action values $\delta \triangleq r_t + \widetilde{\eta} Q_q(\mathbf{Z}(t+1), \mathbf{A}(t+1)) - Q_q(\mathbf{Z}(t), \mathbf{A}(t))$.
8: Use δ to update the critic parameter of value function $q \leftarrow q + \phi_q \delta \nabla_q Q_q(\mathbf{Z}(t), \mathbf{A}(t))$.
9: **end for**
10: **Output:** optimal aggregate ISACP-based energy-efficiency $E(\boldsymbol{\rho})$.

V. Performance Evaluations

In Fig. 2, we compare the error of channel state estimation $\left|s_{k,l}^{\mathrm{com}} - \widehat{s}_{k,l}^{\mathrm{com}}\right|$ under different values of step size η in the federated learning based algorithm in Eq. (9). We assume that the number of MUs is $K = 250$. We set the step size η as 0.2β, 0.4β, and 0.45β, respectively, where β is the β-strong convexity of the learning approach. We can observe from Fig. 2 that a larger step size yields the faster convergency for the communication rounds between the central server at the massive MIMO BS and MU agents. We can also observe from Fig. 2 that there always exists the error between the actual channel state and the estimated channel state, due to the non-vanishing wireless fading.

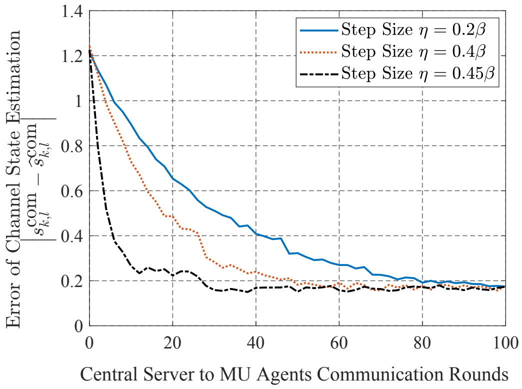

Fig. 2. The error of channel state estimation under different values of step size η in the federated learning based algorithm.

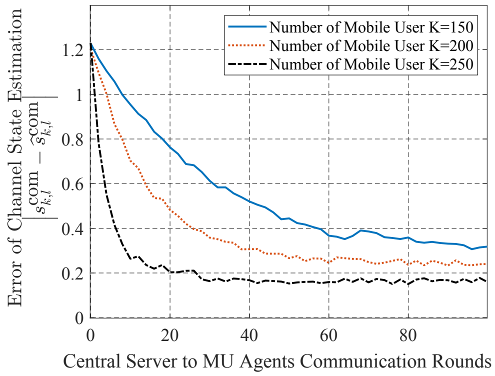

Fig. 3. The error of channel state estimation under different numbers of MUs using the federated learning based algorithm.

In Fig. 3, we compare the error of channel state estimation $\left|s_{k,l}^{\text{com}} - \widehat{s}_{k,l}^{\text{com}}\right|$ under different numbers of MUs using the federated learning based algorithm. We set that the step size $\eta = 0.45\beta$. We set the numbers of MUs as 150, 200, and 250, respectively. Figure 3 reveals that the converged error of channel state estimation monotonically decreases as the numbers of MUs increases. This is because a larger number of MUs implies a more accurate global model, and thus, results in a smaller channel state estimation error. Figure 3 also shows that a larger number of MUs yields a faster convergency since a larger number of MUs can improve the convergency speed of the average gradient $\nabla\overline{L}$ at the central server.

Figure 4 shows the ISACP-based energy-efficiency, denoted by $E(\rho_k)$, of the kth MU under different power splitting ratio ρ_k and different numbers of antennas on the massive MIMO BS. We set the total power allocation for the kth MU $B_k = 3$W and the actual AoA for kth MU $\phi_k = 30°$. Figure 4 reveals that there always exists an optimal power splitting ratio ρ_k for the kth MU to maximize its energy-efficiency, and thus, there exists an optimal power splitting ratio vector $\boldsymbol{\rho}$ for all MUs, to maximize the overall energy-efficiency $E(\boldsymbol{\rho})$ given by Eq. (20). We also observe from Fig. 4 that the maximum ISACP-based energy-efficiency monotonically increases as the number of antennas on the massive MIMO BS increases, showing that a large number of transmit antennas can improve the performance of our proposed ISACP scheme.

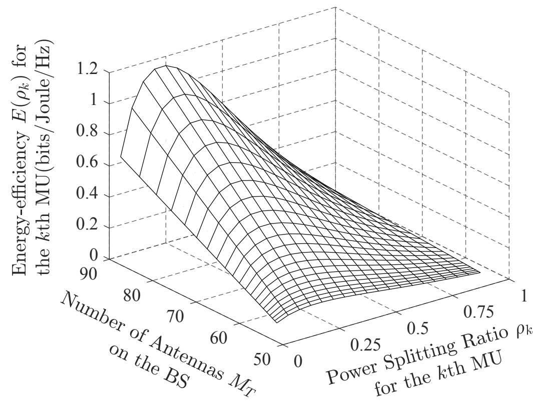

Fig. 4. The function of ISACP-based energy-efficiency, denoted by $E(\rho_k)$, of the kth MU under different power splitting ratio ρ_k and different numbers of antennas on the massive MIMO BS.

VI. Conclusions

We have applied the ISACP to integrate the simultaneous sensing, communication, and powering by using the massive MIMO downlink signal. To estimate the CSIR, we have applied the federated learning based algorithm to train the learning model through a local data set, where the input data is the pilot signals and the output data is MUs' estimated CSIR. Having formulated an energy-efficiency maximization problem, we have then applied the actor-critic enabled multi-agent algorithm to derive the optimal power splitting ratio by using the CSIT estimated by the massive MIMO BS.

References

[1] J. Mu, Y. Gong, F. Zhang, Y. Cui, F. Zheng, and X. Jing, "Integrated sensing and communication-enabled predictive beamforming with deep learning in vehicular networks," *IEEE Communications Letters*, vol. 25, no. 10, pp. 3301–3304, 2021.

[2] L. You, X. Qiang, C. G. Tsinos, F. Liu, W. Wang, X. Gao, and B. Ottersten, "Beam squint-aware integrated sensing and communications for hybrid massive MIMO LEO satellite systems," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 10, pp. 2994–3009, 2022.

[3] H. Yang, Y. Ye, X. Chu, and M. Dong, "Resource and power allocation in SWIPT-enabled device-to-device communications based on a nonlinear energy harvesting model," *IEEE Internet of Things Journal*, vol. 7, no. 11, pp. 10 813–10 825, 2020.

[4] F. Wang and X. Zhang, "Secure resource allocations for polarization-enabled multiple-access cooperative cognitive radio networks with energy harvesting capability," *IEEE Transactions on Wireless Communications*, vol. 22, no. 8, pp. 4990–5004, 2023.

[5] Y. Chen, H. Hua, J. Xu, and D. W. K. Ng, "ISAC meets SWIPT: Multi-functional wireless systems integrating sensing, communication, and powering," *IEEE Transactions on Wireless Communications*, Early Access, 2024.

[6] H. Guo, A. Liu, and V. K. N. Lau, "Analog gradient aggregation for federated learning over wireless networks: Customized design and convergence analysis," *IEEE Internet of Things Journal*, vol. 8, no. 1, pp. 197–210, 2021.

[7] X. Liu, T. Huang, N. Shlezinger, Y. Liu, J. Zhou, and Y. C. Eldar, "Joint transmit beamforming for multiuser MIMO communications and MIMO radar," *IEEE Transactions on Signal Processing*, vol. 68, pp. 3929–3944, 2020.

[8] M. Zhu, L. Li, S. Xia, and T.-H. Chang, "Information and sensing beamforming optimization for multi-user multi-target MIMO ISAC systems," *EURASIP Journal on Advances in Signal Processing*, vol. 15, pp. 1–29, 2023.

[9] V. Konda and J. Tsitsiklis, "Actor-critic algorithms," *Advances in neural information processing systems*, vol. 12, 1999.