Vertical Federated Learning for Multicell Integrated Sensing and Communication Systems

Lai Jiang, Kaitao Meng, and Christos Masouros
Department of Electronic and Electrical Engineering, University College London, London, UK
Emails: {lai.jiang.22, kaitao.meng, c.masouros}@ucl.ac.uk

***Abstract*—Beamforming is a crucial technique to enable dual-functionality for integrated sensing and communication (ISAC) systems by leveraging the multi-antenna arrays. Most existing beamforming strategies focus solely on single-base station (BS) scenarios, neglecting the impact of inter-cell interference (ICI). However, managing ICI often requires additional information exchange, which results in expensive communication overhead and extra latency. To address this problem, we investigate cooperative beamforming in multi-cell ISAC systems and formulate an optimization problem that jointly maximizes the weighted sum of communication rate and radar information rate. Considering the computational efficiency, we propose to apply the vertical federated learning (VFL) technique to solve the optimization problem with limited transmission overhead. In our approach, neural networks are distributed across BSs and jointly trained in an unsupervised manner, then the trained neural networks can perform real-time beamforming with local channel information. Meanwhile, we show that our method can achieve distributed ICI managing and find the globally optimized beamforming solutions for each BS. Through numerical simulations, we demonstrate that our beamforming solution outperforms the benchmarks in terms of both communication rate and radar information rate while significantly reducing computational and communication costs.**

***Index Terms*—Integrated sensing and communication, multi-cell system, federated learning, beamforming.**

I. Introduction

Integrated sensing and communications (ISAC) technology plays a pivotal role in next-generation networks, seamlessly merging radar and communication capabilities [1]. By operating on a shared spectrum and utilizing common waveforms and hardware platforms, ISAC not only enhances spectral efficiency but also reduces system complexity and deployment costs. This integration enables a more compact and energy-efficient infrastructure, which is particularly beneficial for applications requiring real-time environmental awareness and reliable data transmission.

A crucial advantage of ISAC lies in its ability to exploit spatial degrees of freedom (DoFs) through multi-antenna arrays deployed at base stations (BSs). These arrays facilitate advanced transmit beamforming, allowing the system to dynamically adjust resource allocation between communication and sensing tasks [2]–[4]. Unlike traditional systems that treat sensing and communication separately, ISAC leverages their synergy to improve both functionalities, enabling robust target detection while maintaining high communication quality. Optimization-based beamforming techniques remain dominant in both communication-only and ISAC systems [2]–[4]. Despite their effectiveness, these approaches typically rely on iterative algorithms, introducing significant computational overhead and latency. As an alternative, learning-based methods have gained attention for their ability to solve complex optimization problems with lower latency and computational cost [5]–[7]. Unlike traditional mathematical models, deep neural networks (DNNs) trained offline can be deployed in real-time with linear matrix operations. For instance, [7] proposed an unsupervised learning-based beamformer to optimize sum rate under power constraints, achieving performance comparable to the WMMSE method [8]. Similarly, [9] introduced a neural network-based beamforming strategy for ISAC, which enhances target illumination while maintaining the signal-to-interference-plus-noise ratio (SINR) for communication users. However, most existing works fail to effectively mitigate inter-cell interference (ICI). ICI significantly impacts network-level sensing and communication performance in multicell scenarios where most existing beamforming strategies operate on a per-cell basis without considering interference from neighboring cells [2]–[4]. This oversight becomes a critical limitation for sensing as echo signal is especially vulnerable to interference [10]. Consequently, in large-scale multicell networks, ICI emerges as a major bottleneck, degrading overall system performance.

To tackle this challenge, one approach is multi-BS coordinated beamforming, where local information is exchanged among BSs to find globally optimal beamforming matrices. The authors in [11] discussed the necessity of performing joint beamfomring, and proposed an efficient algorithm based on uplink-downlink duality, which can be implemented distributedly in a time-division duplex (TDD) system. Moreover, the authors in [12] proposed a framework where BSs cooperatively serve the users and localize each target for enhancing the ICI management and S&C performance. These motivate us to apply a distribued machine learning framework for coordinated beamfomring design. In this work, we propose an innovative data-driven beamforming framework that leverages vertical federated learning (VFL) to train a deep neural network (DNN) using channel data collected by distributed BSs. Unlike traditional methods that rely on complex iterative computations, our approach enables efficient, real-time beamforming with minimal computational overhead. By collaboratively training the models across BSs, our method learns to generate globally optimized beamforming matrices,

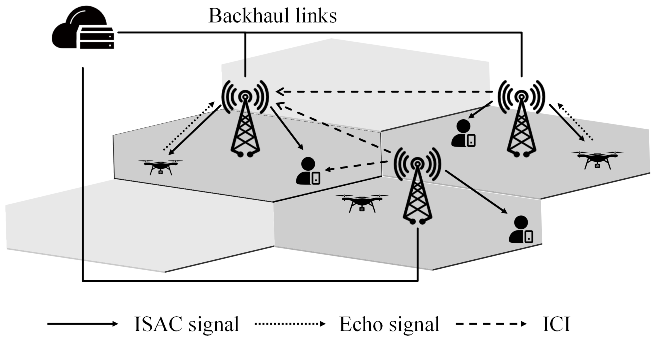

Fig. 1. The considered multi-cell ISAC system.

achieving superior performance while preserving data privacy and reducing communication costs. Through numerical simulations, we demonstrate that the proposed solution achieves performance comparable to the optimization-based algorithms [8] while outperforms the closed-form solutions [13], [14].

II. SYSTEM MODEL

In this paper, we consider a downlink multi-cell ISAC system with M cells. Each BS, $m = 1, 2, ..., M$ is equipped with a uniform linear array (ULA) of N_T transmit antennas and N_R receive antennas spaced at half-wavelength distance, serving K_m single-antenna users and sensing J_m point targets. The index set of the BSs participating in the FL framework is denoted by $\mathcal{M} \triangleq \{1, 2, ..., M\}$. During the training stage, the required information will be exchanged via the backhaul links between the M BSs and a central server.

An illustrative example of the considered system is shown in Fig. 1. In the downlink, the mth BS transmits a unit-power data stream $\mathbf{S}_m \in \mathbb{C}^{K_m \times L}$ with L being the number of timeslots in a frame to the K_m users. The baseband transmitted symbol matrix $\mathbf{X}_m \in \mathbb{C}^{N_T \times L}$ is given by

$$\mathbf{X}_m = \mathbf{W}_m \mathbf{S}_m = \sum_{k=1}^{K_m} \mathbf{w}_{m,k} \mathbf{s}_{m,k}, \tag{1}$$

where $\mathbf{W}_m = [\mathbf{w}_{m,1}, \mathbf{w}_{m,2}, ...\mathbf{w}_{m,K}] \in \mathbb{C}^{N_T \times K_m}$ is the beamforming matrix to be designed and $\mathbf{s}_{m,k}^H \in \mathbb{C}^{L \times 1}$ denotes the data stream intended for the kth user in the mth cell, denoted by U_{mk}. We assume the data streams are orthogonal when L is sufficiently large so that: $(1/L)\mathbf{S}_m\mathbf{S}_m^H = I_{K_m}$. In this case, the received signal at U_{mk} will contain the intended signal and both intra-cell and inter-cell interference, given by

$$\begin{aligned} \mathbf{y}_{m,k}^c = &\mathbf{h}_{m,m,k}^H \mathbf{w}_{m,k} \mathbf{s}_{m,k} + \sum_{l \neq k}^{K_m} \mathbf{h}_{m,m,k}^H \mathbf{w}_{m,l} \mathbf{s}_{m,l} \\ &+ \sum_{n \neq m}^{M} \sum_{i=1}^{K_m} \mathbf{h}_{n,m,k}^H \mathbf{w}_{n,i} \mathbf{s}_{n,i} + \mathbf{z}_{m,k}, \end{aligned} \tag{2}$$

where $\mathbf{z}_{m,k}$ is the additive white Gaussian noise (AWGN) vector, i.e., each entry is independent and identically distributed and follows the Gaussian distribution with zero mean and variance σ_c^2. $\mathbf{h}_{i,j,k} \in \mathbb{C}^{N_T \times 1}$ denotes the block fading channel vector from the ith BS to the kth user in the jth cell, which is modeled as a Rician fading channel given by

$$\mathbf{h}_{i,j,k} = \sqrt{\frac{\upsilon_{i,j,k}}{1 + \upsilon_{i,j,k}}} \mathbf{h}_{i,j,k}^{LOS} + \sqrt{\frac{1}{1 + \upsilon_{i,j,k}}} \mathbf{h}_{i,j,k}^{NLOS}, \tag{3}$$

where $\upsilon_{i,j,k} > 0$ denotes the Rician factor, $\mathbf{h}_{i,j,k}^{LOS}$ and $\mathbf{h}_{i,j,k}^{NLOS}$ are the LOS component and NLOS component respectively.

In this paper, we assume that the J_m sensing targets in each cell can be regarded as point targets relative to the nearest BS, which is commonly adopted in the literature [3], [15]. As the BS works as a monostatic radar, the angle of departure (AoD) and angle of arrival (AoA) are the same. According to the network-level sensing interference model in [10], the ICI channels from neighboring BSs to the serving BS impacts the reception of the target echoes and therefore dominate the network's sensing performance. The received signal at the mth serving BS is the summation of the echo signals reflected by the targets and ICI from the surrounding BSs, given by

$$\begin{aligned} \mathbf{y}_m^s = &\sum_{j}^{J_m} \alpha_{m,j} \mathbf{b}(\theta_{m,j}) \mathbf{a}^H(\theta_{m,j}) \mathbf{W}_m \mathbf{s}_m(t - 2\tau_{m,j}) \\ &+ \sum_{n \neq m}^{M} \mathbf{G}_{n,m} \mathbf{W}_n \mathbf{s}_n(t - \tau_{n,m}) + \mathbf{z}_m, \end{aligned} \tag{4}$$

where $\mathbf{G}_{n,m} \in \mathbb{C}^{N_R \times N_T}$ is the interference channel from the nth BS to the mth BS. In (4), $\alpha_{m,j}$ incorporates the effect of the round-trip pathloss and radar cross section (RCS) of the target, τ denotes the delay experienced by the signal, and $\mathbf{z}_m$ is the AWGN vector with variance σ_s^2. The transmit and receive steering vectors are represented by $\mathbf{a}(\theta_{m,j}) = [1, ..., e^{j\pi(N_T-1)\sin(\theta_{m,j})}]^T \in \mathbb{C}^{N_T \times 1}$ and $\mathbf{b}(\theta_{m,j}) = [1, ..., e^{j\pi(N_R-1)\sin(\theta_{m,j})}]^T \in \mathbb{C}^{N_R \times 1}$ respectively, and θ_{mj} denotes the angle of the target with respect to its nearest BS, which is assumed to be estimated from a previous crude observation. To maximize the received signal-to-noise ratio (SNR) and ensure computational efficiency, the mth BS applies the maximum-ratio combining (MRC) beamformer $\mathbf{v}_{m,j}^H = \mathbf{b}^H(\theta_{m,j}) \in \mathbb{C}^{1 \times N_R}$ to extract the useful information of target j from the received signal, the corresponding signal after processing is

$$\begin{aligned} \mathbf{y}_{m,j}^s &= \mathbf{v}_{m,j}^H \mathbf{y}_m^s \\ &= N_R \alpha_{m,j} \mathbf{a}^H(\theta_{m,j}) \mathbf{W}_m \mathbf{s}_m(t - 2\tau_{m,j}) \\ &+ \sum_{i \neq j}^{J_m} \underbrace{\mathbf{b}^H(\theta_{m,j}) \mathbf{b}(\theta_{m,i}) \mathbf{a}^H(\theta_{m,i})}_{\mathbf{g}_{m,m,i}^H} \mathbf{W}_m \mathbf{s}_m(t - 2\tau_{m,i}) \\ &+ \sum_{n \neq m}^{M} \underbrace{\mathbf{b}^H(\theta_{m,j}) \mathbf{G}_{n,m}}_{\mathbf{g}_{n,m,j}^H} \mathbf{W}_n \mathbf{s}_n(t - \tau_{n,m}) + \tilde{\mathbf{z}}_m. \end{aligned} \tag{5}$$

For convenience, we uniformly represent both intra-cell and inter-cell interference channel as the equivalent interference channel $\mathbf{g}_{n,m,i} \in \mathbb{C}^{1 \times N_T}, \forall n \neq m, i \neq j$. Accordingly, we let $\mathbf{g}_{m,m,j}^H = \alpha_{m,j} \mathbf{a}^H(\theta_{m,j})$ to denote the equivalent sensing channel from the mth BS to the intended target j.

A. Performance Metric

Based on the above model, we can obtain the received SINR of user U_{mk}

$$\gamma^c_{m,k} = \frac{|\mathbf{h}^H_{m,m,k}\mathbf{w}_{m,k}|^2}{\sum_{l\neq k}^{K}|\mathbf{h}^H_{m,m,k}\mathbf{w}_{m,l}|^2 + \sum_{n\neq m}^{M}\sum_{i=1}^{K}|\mathbf{h}^H_{n,m,k}\mathbf{w}_{n,i}|^2 + \sigma_c^2}, \tag{6}$$

the first two terms in the denominator of (6) are the power of intra-cell interference and inter-cell interference, respectively, which are to be minimized to improve the system-level performance. We adopt the achievable communication rate of the network as performance metric, which is written as

$$R_c = \sum_{m=1}^{M}\sum_{k=1}^{K_m}\log_2(1+\gamma^c_{m,k}). \tag{7}$$

Due to the use of receive filters, the optimal beamforming matrix for sensing should match the equivalent channel of the sensing target to maximize the illumination power, while minimizing the interference power from other equivalent interference channels. Therefore, the SINR of sensing target j after applying MRC can be expressed as

$$\gamma^s_{m,j} = \frac{|\mathbf{g}^H_{m,m,j}\mathbf{W}_m|^2}{\sum_{i\neq j}^{J_m}|\mathbf{g}^H_{m,m,i}\mathbf{W}_m|^2 + \sum_{n\neq m}^{M}|\mathbf{g}^H_{n,m,j}\mathbf{W}_n|^2 + \sigma_s^2}. \tag{8}$$

We propose to use the system radar information rate to evaluate the sensing performance in the considered system, as the accuracy of parameter estimation is proportional to the information rate, which is given by

$$R_s = \sum_{m=1}^{M}\sum_{j=1}^{J_m}\log_2(1+\gamma^s_{m,j}). \tag{9}$$

B. Problem Formulation

Without loss of generality, we assume $K_m = K$ and $J_m = J$, $m = 1, 2, ..., M$ for consistency. In this work, we propose to maximize the weighted sum of communication rate and sensing rate, thus the network-level beamforming optimization problem is formulated as

$$\begin{aligned} \max_{\mathbf{W}} \quad & \rho R_c + (1-\rho)R_s \\ s.t. \quad & \text{tr}(\mathbf{W}_m\mathbf{W}_m^H) \leq P_T, \quad \forall m = 1,2,...,M, \end{aligned} \tag{10}$$

where $\mathbf{W} = [\mathbf{W}_1, \mathbf{W}_2, .., \mathbf{W}_M] \in \mathbb{C}^{N_T\times KM}$ is the collection of locally designed beamforming matrices. The performance tradeoff between communication and sensing is achieved by the weighting factor $\rho \in [0, 1]$, and the transmit power at each BS is constrained to P_T. Obviously, solving problem (10) is challenging due to its non-convexity. In the next section, we will apply the deep learning technique to solve (10) efficiently.

III. Vertical FL Framework for Beamforming

Vertical federated learning (VFL) is a collaborative machine learning paradigm designed for scenarios where multiple parties possess different features of the same set of data samples [16]. Unlike horizontal federated learning [17], which deals with data split by users, VFL focuses on data split by features. This makes it especially useful for collaborations between organizations that share users but store different types of information. In this section, we propose to train a multilayer perceptron (MLP) [18] under VFL framework to find the optimal beamformers for each BS which maximizes the weighted sum of communication rate R_c and sensing rate R_s while satisfying the power constraints.

A. DNN architecture

The proposed MLP consists of N_H layers with d_H neurons per hidden layer to approximate the beamforming function.

1) Input layer: During local training at each BS m, $m = 1, ..., M$, the DNN takes communication channels $\mathbf{H}_m = \{(\mathbf{h}_{m,n,k})_{\forall n,k}\} \in \mathcal{C}^{N_T\times MK}$ and sensing channels $\mathbf{G}_m = \{(\mathbf{g}_{m,n})_{\forall n}\} \in \mathcal{C}^{N_T\times M}$ as inputs. Since deep learning frameworks do not natively support complex numbers, these matrices are decomposed into their real and imaginary components before being concatenated into a real-valued input vector.

2) Hidden layers: The depth of a DNN impact its nonlinearity and ability to extract feature information. The proposed network for beamforming contains four fully-connected hidden layers, balancing model complexity and overfitting risk. We apply the LeakyReLU activation function, which allows small negative slopes to handle negative values effectively. Additionally, a dropout layer with a probability of $\zeta = 0.15$ is employed after activation to enhance generalization and prevent overfitting.

3) Output layer: The output layer produces a matrix of size $N_T \times K \times 2$ representing both real and imaginary components of the beamforming solution, which are separated and recombined to reconstruct the complex-valued beamforming matrix $\hat{\mathbf{W}}$ of size $N_T \times K$. Finally, a normalization layer is used to scale the output so that the power constraint is satisfied, which is given by

$$\mathbf{W}^{(o)} = \sqrt{\frac{P_T}{\text{tr}(\hat{\mathbf{W}}\hat{\mathbf{W}}^H)}}\hat{\mathbf{W}}. \tag{11}$$

B. Channel Data Acquisition

In conventional centralized frameworks, local nodes must transmit their collected channel data to a central server, where globally optimized beamforming solutions are computed and then sent back for local utilization. Unlike these approaches, our method only requires channel information uploading during the training process, while the trained models can operate based on the local channel information when deployed online. First, each BS transmits an orthogonal pilot signal $\tilde{\mathbf{X}}$ for channel estimation and initial detection, the covariance matrix of the probing signal is denoted by

$$\mathbf{R}_{\tilde{\mathbf{X}}} = \frac{P_T}{N_T}\mathbf{I}_{N_T}, \tag{12}$$

where the omnidirectional signal is received not only by the intended users and targets but also by interfering devices. BS m can estimates a general direction θ_m of the target of the target within its coverage area based on the echo signal. After receiving the pilots, users and neighboring BSs return the estimated channel data to the transmitting BS. Consequently, BS m constructs its local channel dataset.

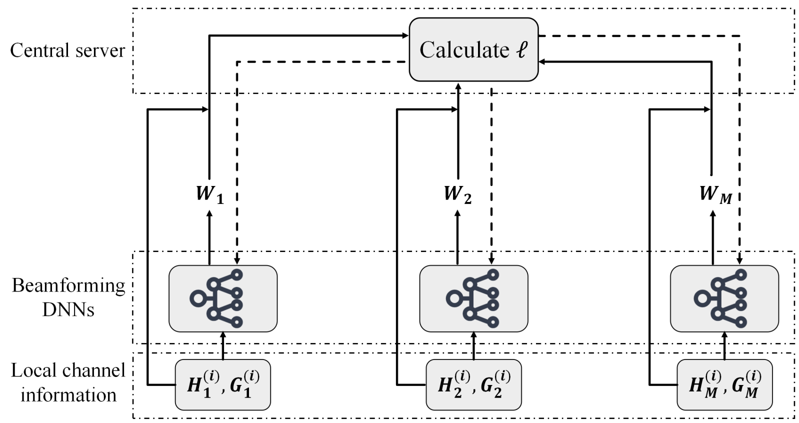

Fig. 2. Proposed training framework.

C. Learning-based Formulation

Fig. 2 shows the proposed VFL-based training framework for cooperative beamforming design. This framework prioritizes minimizing information exchange and reducing the associated overhead. To achieve this, during the training phase, the central server calculates the global communication and sensing loss, which are defined as the negative sum communication rate and radar information rate respectively

$$\mathcal{L}_c(\mathbf{W}) = -\sum_{m=1}^{M}\sum_{k=1}^{K}\log_2(1+\gamma_{m,k}^c), \tag{13}$$

$$\mathcal{L}_s(\mathbf{W}) = -\sum_{m=1}^{M}\sum_{j=1}^{J}\log_2(1+\gamma_{m,j}^s), \tag{14}$$

The global optimization problem becomes minimizing the weighted sum of (13) and (14)

$$\begin{aligned} \min_{[\mathbf{W_1},..,\mathbf{W_M}]} \quad & \rho\mathcal{L}_c(\mathbf{W}) + (1-\rho)\mathcal{L}_s(\mathbf{W}) \\ s.t. \quad & \mathrm{tr}(\mathbf{W}_m\mathbf{W}_m^H) \le P_T, \quad \forall m = 1,2,...,M. \end{aligned} \tag{15}$$

For the CUs, the full feature set of the ith sample sample consists of the global channel information $\mathbf{H}^{(i)} = [\mathbf{H}_1^{(i)}, \mathbf{H}_2^{(i)}, ...\mathbf{H}_M^{(i)}]$, where $\mathbf{H}_m^{(i)}$ represents the partial channel data available at BS m. Likewise, the sensing feature set for the ith sample is given by $\mathbf{G}^{(i)} = [\mathbf{G}_1^{(i)}, \mathbf{G}_2^{(i)}, ...\mathbf{G}_M^{(i)}]$, aggregating the local sensing channel data from all BSs. Within the VFL framework, each BS independently learns to optimize its local beamformer $\mathbf{W}_m^*, \forall m \in \mathcal{M}$ through the following steps.

1) Forward propagation: In the global round T, the local BSs use the ith local channel sample to design the beamforming matrices, the process can be given by

$$\mathbf{W}_m = f([\mathbf{H}_m^{(i)}, \mathbf{G}_m^{(i)}]; \omega_m^{(T)}), \tag{16}$$

where $\omega_m^{(T)}$ is the model parameter set of the mth BS at the Tth global round.

2) Local uploading: Each BS uploads its independently designed beamforming matrices and the corresponding input channel sample $[\mathbf{H}_m^{(i)}, \mathbf{G}_m^{(i)}]$ to the central server.

3) Features aggregation: Upon receiving the designed beamformers and partial channel information from the BSs, the central server can aggregate them into the global channel information $\mathbf{H}^{(i)}$, $\mathbf{G}^{(i)}$ and calculate the global loss

$$\mathcal{L}(\mathbf{W}) = \rho\mathcal{L}_c(\mathbf{W}) + (1-\rho)\mathcal{L}_s(\mathbf{W}), \tag{17}$$

which is then fed back to the local BSs for model updates.

4) Backward propagation: With the global loss information, each BS can perform individual backward propagation using the chain rule, propagating the error from the output layer back to the input layer of its own model. In this work, we adopt the stochastic gradient descent (SGD) algorithm [19] to update the model parameters, which is given by

$$\omega_m^{(T+1)} = w_m^{(T)} - \eta\nabla\ell\big(\mathbf{W}; \omega_m^{(T)}\big), \tag{18}$$

where η represents the learning rate.

D. Complexity Analysis

Deep learning-based beamforming introduces additional offline training costs beyond the computational overhead of online processing. Training complexity mainly rises from forward and backward propagation, involving arithmetic operations in fully connected layers and data processing blocks. The proposed MLP network has N_H hidden layers of dimension d_H, with input and output sizes depending on system configuration. Using Big-O notation, the forward propagation complexity is $O((MK + M + K)N_T d_H + N_H d_H^2)$, while activation layers contribute $O((N_H + 1)d_H)$. The input and output of DNN are processed to generate real-valued vector and complex-valued vector respectively, which results in the computational cost of $O((M + 1)KN_T)$. The computational cost of backward propagation is generally about twice that of forward propagation [20], thus the total training complexity is $O(3((MK + M + K)N_T d_H + N_H d_H^2))$.

Beyond the computational burden mentioned earlier, the joint training process also incurs additional communication overhead. During each training iteration within the VFL framework, participating BSs must transmit their output matrices and corresponding input channel samples to a central server for global loss computation after completing a forward

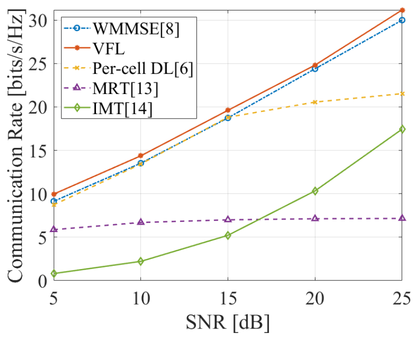

Fig. 3. Communication rate R_c with respect to SNR.

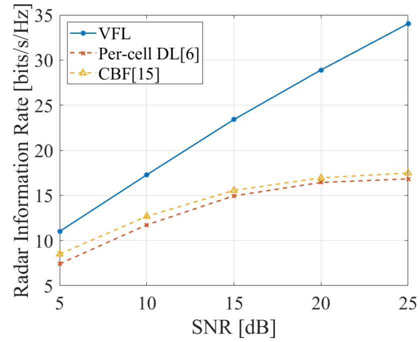

Fig. 4. Radar information rate R_s with respect to SNR.

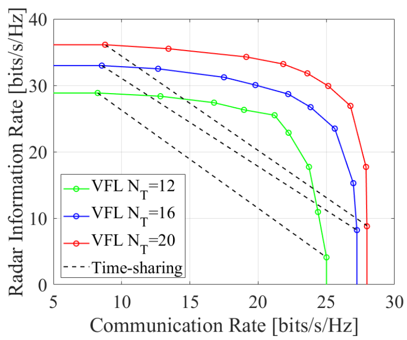

Fig. 5. Tradeoff betweem communication rate R_c and radar information rate R_s when $SNR = 30dB$ and ρ ranges from 0 to 1.

TABLE I
HYPERPARAMETERS

HyperParameter	Value
Cell radius	500(m)
Communication path loss	3.6
Sensing path loss	2
Rician factor υ	3
Target RCS	1.0(m^2)
Num of hidden layers N_H	4
Num of neurons in hidden layers d_H	512
Weight decay factor	10^{-6}
Batch size	128
Learning rate	10^{-4}

pass. This results in a total communication overhead of $(M+1)KN_T$ for each local BS. In the online deployment phase, the primary computational cost arises from forward propagation, which scales proportionally with the number of cells M, communication users K and transmit antennas N_T. In contrast, the WMMSE algorithm has a computational complexity of $O(L_\omega(K^2N_T^2 + KN_T^3))$ according to [21], where L_ω denotes the number of iterations. This highlights a key drawback of traditional optimization approaches—their high complexity and iterative nature, which hinder parallelization and make them less practical for real-time applications.

IV. Performance Evaluation

In this section, we evaluate the proposed beamforming solution for multi-cell ISAC systems. Unless otherwise mentioned, the considered system consists of $M = 3$ cells, where each BS is equipped with $N_T = 12$ transmit antennas and $N_R = 12$ receive antennas, serving $K = 2$ users while sensing $J = 2$ point targets. The positions of users and targets are generated randomly, and both communication and sensing channels are obtained through the channel sounding stage as described in III-B. In our simulations, each BS collects 20,000 channel samples for training and 2,000 channel samples for testing. Model parameters are updated using the Adam optimizer [22], other hyperparameters are summarized in Table. I.

We first evaluate the communication performance of the proposed methods by analyzing the achievable rate across varying signal-to-noise ratios (SNRs). The benchmarks include the optimization-based WMMSE scheme [8], the per-cell deep learning (DL) beamforming method [6], and two closed-form solutions: maximum ratio transmission (MRT) [13] and interference minimizing transmission (IMT) [14]. As shown in Fig. 3, when $\rho = 0.99$, our VFL method consistently outperforms WMMSE, demonstrating the advantage of learning-based approaches, which significantly reduce computational complexity and latency compared to traditional optimization methods. Moreover, leveraging global channel information enables the VFL method to achieve a 45% gain over the per-cell DL method, which suffers from performance saturation at high SNRs. A similar pattern emerges in the MRT-IMT comparison, and this highlights the effectiveness of ICI controlling when interference surpasses noise as the dominant factor. The results underscore the importance of cooperative beamforming in optimizing network-level performance.

For sensing performance, we compare our proposed methods against the conjugate beamforming (CBF) scheme [15], where each BS directs its beamformer towards target angle θ_m to maximize sensing SNR. Additionally, we adapt the per-cell DL approach from [6], modifying its loss function to optimize the radar information rate. Fig. 4 illustrates that when $\rho = 0.01$, the VFL method achieves up to a 94% improvement over CBF as SNR increases. Besides, By comparing the performance differences between our method and two other methods that are performed on a per-cell basis, we can observe that the impact of eliminating ICI is more pronounced in sensing than in communication at low SNRs, as interference is the dominant limiting factor due to round-trip path loss, which is consistent with conclusions drawn in [10].

The tradeoff between communication and sensing performance as a function of the number of transmit antennas is presented in Fig. 5. Each curve represents the achievable communication rate R_c and radar information rate R_s across different values of ρ (0.1–0.9). We introduce a time-sharing scheme, constructed from two extreme operating points, as a reference for performance comparison. The Pareto front, formed by the (R_c, R_s) boundary, expands as the number of antennas increases, demonstrating that the VFL method effectively exploits spatial degrees of freedom (DoFs). Specifically, for a fixed radar information rate, our method achieves up to 58% higher communication rate than the time-sharing scheme with $N_T = 12$, and this advantage grows to 63% when $N_T = 20$. Furthermore, the ability to suppress interference scales with the antenna count, leading to more substantial improvements in sensing performance. These results confirm that interference management benefits grow with available DoFs, significantly enhancing both sensing and communication efficiency at the network level.

V. Conclusion

In this paper, we proposed to use VFL framework for the downlink beamforming design in multi-cell ISAC. We formulated the optimization problem to maximize the weighted sum of the system communication rate and radar information rate, and utilized the neural networks to find the optimal solutions efficiently. The models trained under VFL framework can operate without global channel information exchange between BSs. Through numerical simulations, we demonstrated that our solution can achieve performance comparable to traditional centralized methods while offering significant improvements in computational efficiency and scalability, making it suitable for practical deployments.

References

[1] F. Liu, Y. Cui, C. Masouros, J. Xu, T. X. Han, Y. C. Eldar, and S. Buzzi, "Integrated sensing and communications: Toward dual-functional wireless networks for 6g and beyond," *IEEE journal on selected areas in communications*, vol. 40, no. 6, pp. 1728–1767, 2022.

[2] 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.

[3] F. Liu, Y.-F. Liu, A. Li, C. Masouros, and Y. C. Eldar, "Cramér-rao bound optimization for joint radar-communication beamforming," *IEEE Transactions on Signal Processing*, vol. 70, pp. 240–253, 2021.

[4] J. Li, G. Zhou, T. Gong, and N. Liu, "A framework for mutual information-based MIMO integrated sensing and communication beamforming design," *IEEE Transactions on Vehicular Technology*, 2024.

[5] H. Sun, X. Chen, Q. Shi, M. Hong, X. Fu, and N. D. Sidiropoulos, "Learning to optimize: Training deep neural networks for interference management," *IEEE Transactions on Signal Processing*, vol. 66, no. 20, pp. 5438–5453, 2018.

[6] H. Huang, W. Xia, J. Xiong, J. Yang, G. Zheng, and X. Zhu, "Unsupervised learning-based fast beamforming design for downlink MIMO," *IEEE Access*, vol. 7, pp. 7599–7605, 2018.

[7] W. Xia, G. Zheng, Y. Zhu, J. Zhang, J. Wang, and A. P. Petropulu, "A deep learning framework for optimization of MISO downlink beamforming," *IEEE Transactions on Communications*, vol. 68, no. 3, pp. 1866–1880, 2019.

[8] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted mmse approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4331–4340, 2011.

[9] R. P. Sankar, S. S. Nair, S. Doshi, and S. P. Chepuri, "Learning to precode for integrated sensing and communication systems," in *2023 31st European Signal Processing Conference (EUSIPCO)*. IEEE, 2023, pp. 695–699.

[10] K. Meng, C. Masouros, G. Chen, and F. Liu, "Network-level integrated sensing and communication: Interference management and bs coordination using stochastic geometry," *IEEE Transactions on Wireless Communications*, vol. 23, no. 12, pp. 19 365–19 381, 2024.

[11] H. Dahrouj and W. Yu, "Coordinated beamforming for the multicell multi-antenna wireless system," *IEEE transactions on wireless communications*, vol. 9, no. 5, pp. 1748–1759, 2010.

[12] K. Meng, C. Masouros, A. P. Petropulu, and L. Hanzo, "Cooperative isac networks: Performance analysis, scaling laws, and optimization," *IEEE Transactions on Wireless Communications*, vol. 24, no. 2, pp. 877–892, 2025.

[13] N. Fatema, G. Hua, Y. Xiang, D. Peng, and I. Natgunanathan, "Massive mimo linear precoding: A survey," *IEEE systems journal*, vol. 12, no. 4, pp. 3920–3931, 2017.

[14] R. Zakhour, Z. K. Ho, and D. Gesbert, "Distributed beamforming coordination in multicell MIMO channels," in *VTC Spring 2009-IEEE 69th Vehicular Technology Conference*. IEEE, 2009, pp. 1–5.

[15] U. Demirhan and A. Alkhateeb, "Cell-free ISAC MIMO systems: Joint sensing and communication beamforming," *arXiv preprint arXiv:2301.11328*, 2023.

[16] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: Concept and applications," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 10, no. 2, pp. 1–19, 2019.

[17] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Artificial intelligence and statistics*. PMLR, 2017, pp. 1273–1282.

[18] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural networks*, vol. 2, no. 5, pp. 359–366, 1989.

[19] S.-i. Amari, "Backpropagation and stochastic gradient descent method," *Neurocomputing*, vol. 5, no. 4-5, pp. 185–196, 1993.

[20] S. Li, Y. Zhao, R. Varma, O. Salpekar, P. Noordhuis, T. Li, A. Paszke, J. Smith, B. Vaughan, P. Damania *et al.*, "Pytorch distributed: Experiences on accelerating data parallel training," *arXiv preprint arXiv:2006.15704*, 2020.

[21] Q. Hu, Y. Cai, Q. Shi, K. Xu, G. Yu, and Z. Ding, "Iterative algorithm induced deep-unfolding neural networks: Precoding design for multiuser MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 20, no. 2, pp. 1394–1410, 2020.

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