Distributed Beamforming for Cooperative Multi-cell ISAC: A Federated Learning Approach

Lai Jiang, Kaitao Meng, Murat Temiz, Jiaming Hu, and Christos Masouros
Department of Electronic and Electrical Engineering, University College London, London, UK
Emails: {lai.jiang.22, kaitao.meng, m.temiz, jiaming.hu.19, c.masouros}@ucl.ac.uk

***Abstract*—In this work, we propose a distributed framework based on the federated learning (FL) for beamforming design in multicell integrated sensing and communications (ISAC) systems. Our aim is to address the following dilemma: 1) Beamforming strategies based on solely local information may cause severe inter-cell interference (ICI) affecting both communication users and sensing receivers in adjacent cells, leading to degraded performance in communication and sensing, 2) Centralized beamforming requires global knowledge of global communication and sensing channel information, which incurs additional transmission overhead and latency. In the proposed framework, multiple base stations (BSs) jointly train a deep neural network (DNN) to cooperatively design the optimal beamforming matrices, aiming at maximizing the weighted sum of communication rate and radar information rate. To implement a fully decentralized design without channel information exchange among BSs, we develop a novel loss function to manage the interference leakage, which can be computed by only using local channel information. Numerical results demonstrate that the proposed method achieves performance comparable to optimization-based algorithms and surpasses closed-form solutions in terms of both communication rate and radar information rate.**

***Index Terms*—Integrated sensing and communication, multicell system, federated learning, beamforming.**

I. Introduction

Integrated sensing and communications (ISAC) has been recognized as a key enabler for the next-generation networks [1]. By employing unified radar and communication spectrum/waveform/platforms, ISAC systems greatly improve the utilization efficiency of energy and hardware resources. With the deployment of multi-antenna arrays on base stations (BSs), transmit beamforming can enable tradeoffs and mutual benefits between the communication and sensing functionalities by leveraging the available spatial degrees of freedom (DoFs). Most existing beamforming strategies [2]–[4] are performed on a per-cell basis, while the effect of inter-cell interference (ICI) is ignored. However, ICI has been proven detrimental to the network-level S&C performance in [5], with sensing being particularly sensitive due to the round-trip path loss of echo signals. One potential solution to this problem involves the exchange of local information for multi-BS coordinated beamforming, or alternatively, the collection of global information to design the optimal beamformer in a centralized manner [6]. Specifically, the authors in [7] proposed a framework where BSs cooperatively serve the users and localize each target for enhancing the ICI management and S&C performance. Nevertheless, centralized strategies suffer from expensive transmission overhead and additional latency, especially for large-scale networks. Thus, it is critical to design a distributed beamforming method to manage the ICI without requiring information exchange.

In the realm of beamforming/precoding design, whether for communication-only systems or ISAC systems, the majority of existing works [2]–[4], [8] are based on optimization techniques and rely heavily on iterative algorithms with high computational complexity and latency, making it challenging to implement them in practical systems. As low latency and low cost are generally demanded in real-time applications, learning-based methods provide a novel approach to solve complex optimization problems [9]–[11]. Unlike the methods based on rigorous mathematical models, the offline-trained deep neural network (DNN) can be deployed online, operating with limited matrix multiplications and additions. Specifically, the authors in [11] designed a communication-only beamformer by training a DNN in an unsupervised manner to maximize the sum rate under the transmit power constraints, which achieves the performance close to WMMSE method [8]. Moreover, in [12], the authors proposed a beamforming neural network for ISAC systems to maximize the target illumination power while ensuring the signal-to-interference-plus-noise ratio (SINR) for communication users. Despite achieving low computational complexity, the aforementioned references often fall short in effectively managing the ICI.

The above discussion motivates us to adopt the federated learning (FL) technique for beamforming design in multicell ISAC systems. FL [13] is a paradigm of distributed learning framework that leverages the datasets spread across numerous local nodes without explicitly exchanging them. Recent research efforts are applying FL for low-overhead precoding design in massive MIMO systems [14] [15]. However, these methods can not be directly extended to ISAC systems where sensing functionality should also be considered.

In this work, we propose a distributed beamforming framework for multicell ISAC systems based on FL. During offline training, multiple dual-functional BSs use their local datasets to update the global DNN model under the coordination of a central server. To support a fully decentralized implementation, we design a novel loss function which maximizes the communication rate and radar information rate while mitigating interference leakage to undesired receivers. The loss function can be independently optimized at each BS, thereby eliminating the need for exchanging global channel

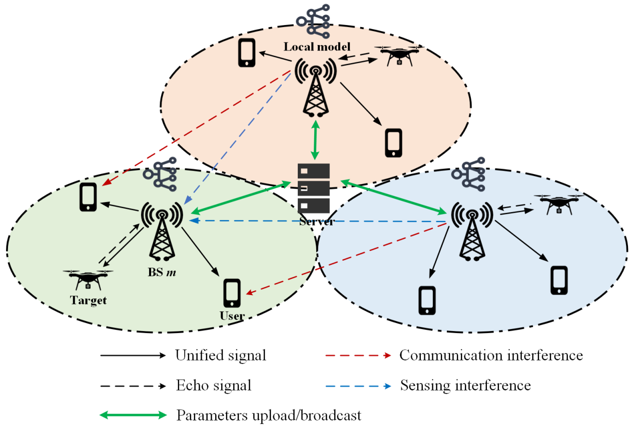

Fig. 1. The considered multicell ISAC system.

information during both training and inference stages. Numerical simulations demonstrate that the proposed solution achieves performance comparable to centralized methods, offering substantial improvements in efficiency and feasibility for practical deployments. To the best of our knowledge, this study represents the first application of FL for beamforming design in multicell ISAC systems.

II. SYSTEM MODEL

In the considered multicell ISAC system, each dual-functional BS is equipped with a half-wavelength spaced uniform linear array (ULA) of N_T transmit antennas and N_R receive antennas, serving K downlink single-antenna users while sensing a target simultaneously. The FL framework is carried out by M BSs and a central server. During the training stage, the participating BSs use local communication and sensing channel data to update their model and upload the parameters to the server via backhaul links for model aggregation. Subsequently, the updated model is fed back to the BSs for the next iteration until convergence.

A. Communication model

In the downlink, the mth BS transmits a unit-power data stream $\mathbf{S}_m \in \mathbb{C}^{K\times L}$ with length L to the K communication users (CUs) which are uniformly distributed within the coverage area of each cell. The baseband transmitted symbol matrix $\mathbf{X}_m \in \mathbb{C}^{N_T\times L}$ at the mth BS is denoted by

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

where $\mathbf{W}_m = [\mathbf{w}_{m,1}, \mathbf{w}_{m,2}, ...\mathbf{w}_{m,K}] \in \mathbb{C}^{N_T\times K}$ 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. We assume the data streams are orthogonal when L is sufficiently large so that: $(1/L)\mathbf{S}_m\mathbf{S}_m^H = I_K$. In this case, the received signal at the kth user in the mth cell is the summation of the intended signal and both intra-cell and inter-cell interference, given by

$$\begin{aligned}\mathbf{y}_{m,k}^c = &\underbrace{\mathbf{h}_{m,m,k}^H\mathbf{w}_{m,k}\mathbf{s}_{m,k}}_{\text{intended signal}} + \underbrace{\sum_{l\neq k}^{K}\mathbf{h}_{m,m,k}^H\mathbf{w}_{m,l}\mathbf{s}_{m,l}}_{\text{intracell interference}} \\ &+ \underbrace{\sum_{n\neq m}^{M}\sum_{i=1}^{K}\mathbf{h}_{n,m,k}^H\mathbf{w}_{n,i}\mathbf{s}_{n,i}}_{\text{intercell interference}} + \mathbf{z}_{m,k},\end{aligned} \tag{2}$$

where $\mathbf{h}_{i,j,k} \in \mathbb{C}^{N_T\times 1}$ denotes the block fading channel vector from the ith BS to the kth user in the jth cell, and $\mathbf{z}_{m,k}$ 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. The received SINR of user k is

$$\gamma_{m,k}^c = \frac{|\mathbf{h}_{m,m,k}^H\mathbf{w}_{m,k}|^2}{\sum_{l\neq k}^{K}|\mathbf{h}_{m,m,k}^H\mathbf{w}_{m,l}|^2 + \sum_{n\neq m}^{M}\sum_{i=1}^{K}|\mathbf{h}_{n,m,k}^H\mathbf{w}_{n,i}|^2 + \sigma_c^2}. \tag{3}$$

The first two terms in the denominator of (3) are the power of intra-cell interference and inter-cell interference, respectively, which are to be minimized to improve the system-level performance. As the communication performance metric adopted in this work, the achievable global rate of the ISAC network is written as

$$R_c = \sum_{m=1}^{M}\sum_{k=1}^{K}\log_2(1+\gamma_{m,k}^c). \tag{4}$$

B. Sensing model

We consider a point target located in the far field of each cell and sensed by its nearest BS, which is commonly adopted in the literature [3], [16]. As the BS works as a monostatic radar, the angle of departure (AoD) and angle of arrival (AoA) are the same. Following the network-level sensing interference model in [5], it is the ICI channels from neighboring BSs to the serving BS that impact 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 signal reflected by the target and ICI from the surrounding BSs, given by

$$\begin{aligned}\mathbf{y}_m^s = &\underbrace{\alpha_m\mathbf{b}(\theta_m)\mathbf{a}^H(\theta_m)\mathbf{W}_m\mathbf{s}_m(t-2\tau_m)}_{\text{target echo signal}} \\ &+ \underbrace{\sum_{n\neq m}^{M}\mathbf{G}_{n,m}\mathbf{W}_n\mathbf{s}_n(t-\tau_{n,m})}_{\text{intercell inteference}} + \mathbf{z}_m,\end{aligned} \tag{5}$$

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 (5), α_m incorporates the effect of the round-trip pathloss and radar cross section (RCS) of the target, and $\mathbf{z}_m$ is the AWGN vector with variance σ_s^2. The transmit and receive steering vectors are represented by $\mathbf{a}(\theta_m) = [1, ..., e^{j\pi(N_T-1)\sin(\theta_m)}]^T \in \mathbb{C}^{N_T\times 1}$ and $\mathbf{b}(\theta_m) = [1, ..., e^{j\pi(N_R-1)\sin(\theta_m)}]^T \in \mathbb{C}^{N_R\times 1}$ respectively, and θ_m 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 $\mathbf{v}_m^H = \mathbf{b}^H(\theta_m) \in \mathbb{C}^{1\times N_R}$ to process the received signal, the echo after processing is

$$\begin{aligned}\tilde{\mathbf{y}}_m^s &= \mathbf{v}_m^H \mathbf{y}_m^s \\ &= N_R \alpha_m \mathbf{a}^H(\theta_m)\mathbf{W}_m \mathbf{s}_m(t-2\tau_m) \\ &+ \sum_{n\neq m}^{M} \beta_{n,m} \mathbf{g}_{n,m}^H \mathbf{W}_n \mathbf{s}_n(t-\tau_{n,m}) + \tilde{\mathbf{z}}_m,\end{aligned} \tag{6}$$

where $\mathbf{g}_{n,m}^H = \mathbf{v}_m^H \mathbf{G}_{n,m} \in \mathbb{C}^{1\times N_T}$ is the equivalent interference channel. Let $\mathbf{g}_{m,m}^H = \alpha_m \mathbf{a}^H(\theta_m)$ denote the equivalent sensing channel from the mth BS to the intended target, the SINR of received signal at the mth BS can be denoted by

$$\gamma_m^s = \frac{\sum_{k=1}^K |\mathbf{g}_{m,m}^H \mathbf{w}_{m,k}|^2}{\sum_{n\neq m}^M \sum_{l=1}^K |\mathbf{g}_{n,m}^H \mathbf{w}_{n,l}|^2 + \sigma_s^2}. \tag{7}$$

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

$$R_s = \sum_{m=1}^M \log_2(1+\gamma_m^s). \tag{8}$$

In (7), the numerator represents the illumination power for the intended target, which is expected to be improved for achieving better sensing performance [17]. Apparently, to maximize the performance from the network perspective, we also need to eliminate the sensing ICI received at each BS.

C. Problem Description

To achieve good performance tradeoff between communications and sensing, we aim to solve the following global optimization problem

$$\begin{aligned}\max_{\mathbf{W}} \quad & \rho \mathbf{R}_c + (1-\rho)\mathbf{R}_s \\ s.t. \quad & \text{tr}(\mathbf{W}_m \mathbf{W}_m^H) \leq P_T, \quad \forall m,\end{aligned} \tag{9}$$

where $\rho \in [0,1]$ is the weighting factor to select between the communication metric and sensing metric, and P_T denotes the transmit power constraint at each BS. However, finding the optimal solution of problem (9) is challenging not only due to its non-convexity but also because it requires the knowledge of global channel information, such as ICI channels from other BSs to its intended users and beamforming vectors designed by other BSs. To avoid the need for channel information exchange, we design a distributed beamforming method to solve the problem (9) based on the FL framework, which is detailed in the next section.

III. FL framework for Beamforming

In this section, we propose to train a DNN to fit the mapping from channel information to the optimal beamformer $\mathbf{W}^*$ which maximizes the weighted sum of communication rate R_c and sensing rate R_s.

A. The proposed loss function

As mentioned in the previous discussion, it is essential to understand the inherent complexity for calculating ICI (3) and (7) at each local BS due to the lack of global information, i.e. the local information of transmitter m is limited to the channel between itself and all downlink receivers $\mathbf{g}_{m,m}$, $\mathbf{h}_{m,n,k}, \forall n, k$ and $\mathbf{W}_m, m = 1, ..., M$. This makes it infeasible to directly define the performance metric as the object function of learning as existing methods [10], [12]. Given the constraint of local information, we propose a loss function to control the ICI that does not rely on direct access to global channel information from other BSs. Following the workaround mentioned in [18], we rewrite the ICI as the communication interference leakage (CIL) to others, which can be controlled by minimizing a penalty term

$$\mathbf{\Phi}_m^c(\mathbf{W}_m) = \sum_{n\neq m}^M \sum_{i=1}^K |\mathbf{h}_{m,n,i}^H \mathbf{W}_m|^2. \tag{10}$$

Specifically, the aim is to force the interference caused to undesired receivers to zero via generating the beamformer from the null space spanned by the interfered channels. By explicitly minimizing the CIL term above, we can approximate the sum rate in the cell m as

$$\tilde{R}_{c,m} = \sum_{k=1}^K \log_2\Big(1 + \frac{|\mathbf{h}_{m,m,k}^H \mathbf{w}_{m,k}|^2}{\sum_{l\neq k}^K |\mathbf{h}_{m,m,k}^H \mathbf{w}_{m,l}|^2 + \sigma_c^2}\Big). \tag{11}$$

It can be revealed that (11) can achieve a good approximation of the original objective function in (4) when the interference caused to the mth cell is eliminated by other BSs. Based on the above approximation, the communication loss function $\mathcal{L}_c(\mathbf{W}_m)$ is formulated by combining the approximated sum rate and inter-cell interference leakage, which is given in (12). The weighting factor α decides the importance of eliminating the interference leakage to other BSs. Maximizing this approximated sum rate (11) is readily verified to achieve a good performance through simulations as shown in Section IV-B

Similarly, to improve the system sensing performance, each BS should also avoid the interference caused to other BSs when sensing the target, which can be achieved by minimizing the sensing interference leakage (SIL)

$$\mathbf{\Phi}_m^s(\mathbf{W}_m) = \sum_{n\neq m}^M |\mathbf{g}_{m,n}^H \mathbf{W}_m|^2. \tag{13}$$

Meanwhile, we maximize the illumination for intended target to ensure the sensing performance, so the sensing loss function $\mathcal{L}_s(\mathbf{W}_m)$ is formulated as the summation of negative target illumination power and sensing interference leakage power weighted by a factor β, which is given in (14). Both (12) and (14) can be calculated using purely local information, thus a fully decentralized training framework can be implemented.

B. DNN architecture

Multilayer perceptrons (MLP) are recognized as universal function approximators [19] and widely used in approaching complex nonlinear functions. In this work, we propose to use a 6-layers MLP with 512 neurons in each hidden layer, the details are as follows.

$$\mathcal{L}_c(\mathbf{W}_m) = -\sum_{k=1}^{K}\left(\log_2(1+\frac{|\mathbf{h}_{m,m,k}^H\mathbf{w}_{m,k}|^2}{\sum_{l\neq k}^{K}|\mathbf{h}_{m,m,k}^H\mathbf{w}_{m,l}|^2+\sigma_c^2}) - \alpha\sum_{n\neq m}^{M}\sum_{i=1}^{K}|\mathbf{h}_{m,n,i}^H\mathbf{w}_{m,k}|^2\right). \quad (12)$$

$$\mathcal{L}_s(\mathbf{W}_m) = -\sum_{k=1}^{K}\left(|\mathbf{g}_{m,m}^H\mathbf{w}_{m,k}|^2 - \beta\sum_{n\neq m}^{M}|\mathbf{g}_{m,n}^H\mathbf{w}_{m,k}|^2\right). \quad (14)$$

1) Input layer: During local training at each BS m, $m = 1, ..., M$, the DNN takes the locally collected communication channels $\mathbf{H}_m = \{(\mathbf{h}_{m,n,k})_{\forall n,k}\} \in \mathcal{C}^{MK\times N_T}$ and sensing channels realizations $\mathbf{G}_m = \{(\mathbf{g}_{m,n})_{\forall n}\} \in \mathcal{C}^{M\times N_T}$ as the input. As complex number operations are not supported by the current DNN software, we transform the channel vector to real-valued coefficients before feeding it into the input layer. Specifically, the communication channel matrix $\mathbf{H}_m$ is split into the real part $\Re(\mathbf{H}_m)$ and imaginary part $\Im(\mathbf{H}_m)$, which are then stacked to form a real-valued input vector $\mathbf{H}_m^{(i)} = [\Re(\mathbf{H}_m), \Im(\mathbf{H}_m)]$. Similarly, we process the sensing channel matrix $\mathbf{G}_m$ using the same pipeline and get $\mathbf{G}_m^{(i)} = [\Re(\mathbf{G}_m), \Im(\mathbf{G}_m)]$. The concatenated matrix $\mathbf{X}^{(i)} = [\mathbf{H}_m^{(i)}, \mathbf{G}_m^{(i)}]$ is then flattened and fed into the first layer of the neural network. The learning model can be represented by $f(\mathbf{X}^{(i)}; \omega)$, where ω is the model parameter set.

2) Hidden layers: The proposed MLP contains 4 fully-connected hidden layers. Each hidden layer is followed by an activation layer and a dropout layer. As some elements of the beamforming matrix can be negative, we adopts the LeakyReLu function as the activation function which provides a non-zero slope to negative values. The dropout layer with probability factor $\zeta = 0.15$ can help to alleviate overfitting by randomly setting input units to zero.

3) Output layer: The output layer is of size $N_T \times K \times 2$ and followed by a normalization layer to scale the output so that the power constraint is satisfied, which can be denoted by

$$\mathbf{W}^{(o)} = \sqrt{\frac{P_T}{\text{tr}(\mathbf{W}\mathbf{W}^H)}}\hat{\mathbf{W}}. \quad (15)$$

Finally, we split the output $\mathbf{W}^{(o)} \in \mathbb{C}^{N_T\times K\times 2}$ in the last dimension to obtain the real and imaginary parts and recover the designed beamformer $\mathbf{W}^*$ of size $N_T \times K$ by the $\mathbb{C}2\mathbb{R}$ block, given by

$$\mathbf{W}^* = \mathbf{W}^{(o)}[:, :, 1] + j\mathbf{W}^{(o)}[:, :, 2]. \quad (16)$$

C. Data Acquisition and Distributed training

To ensure the channel state information (CSI) can be estimated via pilot symbols, we assume the time-division duplex (TDD) scheme is applied in the considered system. In traditional centralized approaches, the local nodes have to upload collected channel data to the central server, where global optimal beamformers are designed and then feed back for local use. In contrast to the centralized methods, the data acquisition process is introduced before the learning process for the BS to collect local training data. 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

$$\mathbf{R}_{\tilde{\mathbf{X}}} = \frac{P_T}{N_T}\mathbf{I}_{N_T}, \quad (17)$$

this omnidirectional signal are received by both intended users and targets, as well as the interference devices. The BS m can estimate a general direction θ_m of the target in its cell based on the echo signal. Upon receiving the pilots, the users and neighboring BSs feed back the estimated channel data to the transmitting BS. Finally, BS m obtains a collection of channel realizations $\mathbf{H}_m = \left\{\left(\mathbf{h}_{m,n,k}\right)_{\forall n,k}\right\}$ and $\mathbf{G}_m = \left\{\left(\mathbf{g}_{m,n}\right)_{\forall n}\right\}$, and the local dataset can be represented by $\mathcal{D}^{(m)} = \left\{(\mathbf{H}_m), (\mathbf{G}_m)\right\}$. Specifically, it is reasonable to assume that BSs can obtain accurate channel estimation because DNNs have strong robustness towards input data.

Based on the previous discussion, we aim to solve the following global optimization problem corresponding to fit the whole dataset $\mathcal{D} = \{(\mathcal{D}^{(m)})_{\forall m}\}$ across the BSs

$$\begin{aligned} \min_{\mathbf{W}} \quad & \sum_{m=1}^{M}\left(\rho\mathcal{L}_c(\mathbf{W_m}) + (1-\rho)\mathcal{L}_s(\mathbf{W_m})\right) \\ s.t. \quad & \mathbf{tr}(\mathbf{W}_m\mathbf{W}_m^H) \leq P_T, \quad \forall m = 1, 2, ...M. \end{aligned} \quad (18)$$

With the FL framework and proposed loss function, the above problem can be solved in a decentralized manner with local dataset $\mathcal{D}^{(m)}$, by training with the following local loss function at BS m, $\forall m \subset M$

$$\mathcal{L}(\mathbf{W}_m) = \rho\mathcal{L}_c(\mathbf{W}_m) + (1-\rho)\mathcal{L}_s(\mathbf{W}_m), \quad \forall m = 1...M. \quad (19)$$

IV. PERFORMANCE EVALUATION

In this section, we present the simulation results to evaluate the performance of the proposed FL-based beamforming algorithm. The proposed framework is implemented in Python 3.11.5 and Pytorch 2.1.0 on a PC with one NVIDIA RTX 3060 GPU and 8 Intel i7-11800 CPU cores.

A. System Setup and Sample Generation

Unless otherwise mentioned, we consider the number of cells $M = 3$, while each BS has $N_T = 12$ transmit antennas and $N_R = 12$ receive antennas. We produce channel samples by randomly generating the positions of users and targets. The distance between each BS is 500m. Specifically, the targets are assumed to be located within the range of $[-\pi/2, \pi/2]$ in the angular domain with respect to each BS, while the communication users are randomly distributed within the cell. Moreover, both communication channels and sensing inter-cell interference channels are assumed to be none-line-of-sight (NLOS) Rayleigh fading channels with large-scale pathloss factor $\alpha_{pl} = 3.6$, while sensing channels are line-of-sight (LOS) channels with pathloss factor $\beta_{pl} = 2$. The AWGN

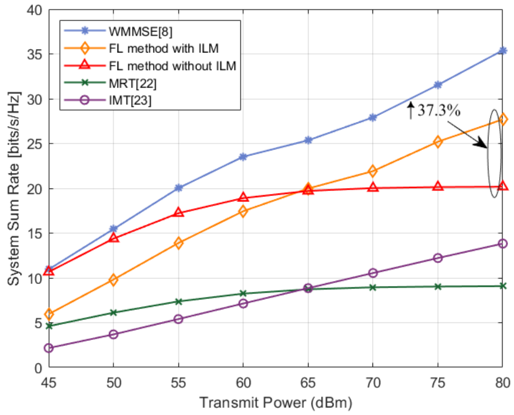

Fig. 2. System sum rate R_c with respect to transmit power P_T.

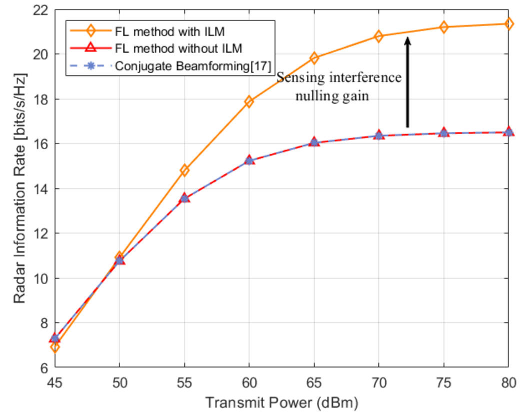

Fig. 3. Radar information rate R_s with respect to transmit power P_T.

powers are set to be $\sigma_c^2 = \sigma_s^2 = 30$ dBm. Each local training dataset and test dataset contains 30,000 and 3000 channel realizations respectively. During training, we use the Adam optimizer to train the model and set weight decay factor to be 10^{-6} for further alleviating the risk of overfitting. The mini-batch size is set to be 128 and the learning rate is 0.0001.

B. Numerical results

Fig. 2 and Fig. 3 show the achievable system sum rate and information rate when the number of cells $M = 3$ and each BS has $N_T = N_R = 12$ antennas, serving $K = 1$ communication user [1] and one target. For evaluating communication performance, we adopt the optimization-based scheme WMMSE [8] as well as the closed-form solutions maximum ratio transmission (MRT) [20] and the interference minimizing transmission (IMT) [21] as benchmarks. To demonstrate the effectiveness of eliminating the ICI, we show the performance of the FL-based method without interference leakage minimization by setting the factor α in (10) to be zero. It can be observed that the FL method without ILM and the MRT scheme both exhibit saturation in communication rate as transmit power increases. In contrast, the FL method with ILM demonstrates superior performance and improves the achievable rate up to 37% at high SINRs, which is referred to as the interference nulling gain. Throughout the increasing SINRs, WMMSE achieves the highest performance. However, WMMSE involves intensive computations and information exchange to manage the ICI, making it less efficient in practical scenarios. On the other hand, the distributed schemes FL method, MRT and IMT operate with lower computational cost and latency, while the proposed FL method outperforms the latter two schemes as it directly optimizes the communication rate.

Fig. 3 illustrates the sensing performance of the proposed method. We compare the sensing performance of our method with the closed-form scheme conjugate beamforming [16], where the transmit beamformer of the mth BS is towards the direction θ_m of the target to maximize the illumination power.

[1]Here, we study the case when ICI dominates the received SINR of CUs, the result can also be extended to the multi-users scenario.

It can be observed that the proposed FL method with ILM outperforms the closed-form solution by up to 29% in terms of the information rate as transmit power increases, and the benefits brought by ICI elimination manifest earlier compared to that in communications. The result indicates that sensing ICI becomes a more dominant factor than noises and useful signal which suffers from round-trip pathloss affecting the sensing performance, which aligns with the conclusion drawn in [5]. It is also worth noting that the curve of FL-based scheme without ILM overlaps with the curve of conjugate beamforming strategy. That is because when β in (14) is set to be zero, minimizing the sensing loss becomes equivalent to maximizing the target illumination power, which provides slightly better performance at low SINRs. Overall, these findings highlight the importance of incorporating ILM in FL-based beamforming methods for optimizing both communication and sensing performance in multicell ISAC systems. The beampatterns obtained by FL method with varying ICI factor β are shown in Figure. 4. We assume a point target located at $\theta = 15°$, while two interfered BSs are located at $-30°$ and $30°$ respectively. The mainlobes of the transmit waveforms obtained using the FL method align with the direction of the intended target, thereby maximizing the target illumination power to ensure the performance of detection and estimation. When β is set to be one, the beampattern of FL method closely resembles that obtained by conjugate beamforming method, which disregards the effect of sensing ICI towards the interfered directions. In contrast, the sidelobes of the beampattern with decreased β exhibit a gradual attenuation at both $-30°$ and $30°$, which results in less power leakage to the directions of interfered BSs so that the sensing interference nulling gain is achieved. The tradeoff profile of the communication performance and sensing performance under different transmit powers is shown in Fig. 5. Each point represents the achievable communication rate R_c and radar information rate R_s as ρ in (9) ranges from 0 to 1. The time-sharing scheme is introduced here to evaluate the performance of our method. The boundary of the (R_c, R_s) region is viewed as the Pareto front constituted by the achieved

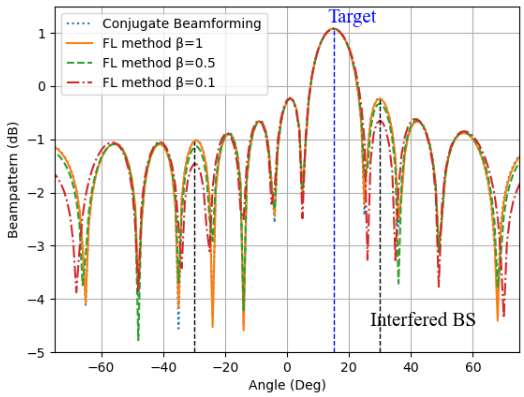

Fig. 4. Beampatterns for the scenario of a point target located at 15° when $K = 2$ and $\rho = 0.1$.

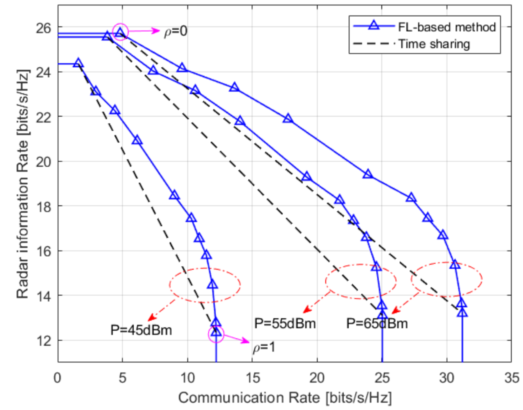

Fig. 5. Tradeoff betweem sum communication rate R_c and radar information rate R_s when $K = 2$ and ρ ranges from 0 to 1.

performance tradeoff. It can be observed that the boundaries exhibit a gradual expansion as transmit power increases, thus the (R_c, R_s) region of proposed method increases noticeably compared to the time-sharing scheme at high SINRs, which implies the increased performance gain with more allocated power.

V. Conclusion

In this paper, we proposed a FL-based beamforming solution for multicell ISAC systems to reduce both the computational complexity and transmission overhead. The DNN is jointly trained by multiple BSs with low communication cost. When deployed online, the trained model designs the beamformer with local channel data and operates with high efficiency. Through numerical simulations, we demonstrated that our FL-based beamforming solution achieves performance comparable to traditional centralized methods while offering significant improvements in computational efficiency and scalability, making it suitable for practical deployments in multicell ISAC systems.

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] K. Meng, C. Masouros, G. Chen, and F. Liu, "Network-level integrated sensing and communication: Interference management and bs coordination using stochastic geometry," *arXiv preprint arXiv:2311.09052*, 2023.

[6] K. Meng, C. Masouros, A. P. Petropulu, and L. Hanzo, "Cooperative ISAC networks: Opportunities and challenges," *arXiv preprint arXiv:2405.06305*, 2024.

[7] ——, "Cooperative isac networks: Performance analysis, scaling laws and optimization," *arXiv preprint arXiv:2404.14514*, 2024.

[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] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] A. M. Elbir and S. Coleri, "Federated learning for hybrid beamforming in mm-wave massive MIMO," *IEEE Communications Letters*, vol. 24, no. 12, pp. 2795–2799, 2020.

[15] D. Wang, M. Tao, X. Zeng, and J. Liang, "Federated learning for precoding design in cell-free massive MIMO systems," *IEEE Open Journal of the Communications Society*, 2023.

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

[17] P. Stoica, J. Li, and Y. Xie, "On probing signal design for MIMO radar," *IEEE Transactions on Signal Processing*, vol. 55, no. 8, pp. 4151–4161, 2007.

[18] N. Babu and C. Masouros, "Multi-cell coordinated joint sensing and communications," in *2023 57th Asilomar Conference on Signals, Systems, and Computers*. IEEE, 2023, pp. 763–767.

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

[20] 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.

[21] 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.