Distributed Unsupervised Learning for Interference Management in Integrated Sensing and Communication Systems

Haijun Zhang, *Senior Member, IEEE*, Xiangnan Liu, Keping Long, *Senior Member, IEEE*, Arumugam Nallanathan, *Fellow, IEEE*, and Victor C. M. Leung, *Fellow, IEEE*

Abstract

Nowadays, the multi-access interference problem in the ISAC systems can not be ignored. The study on interference management in ISAC has been envisioned as one of key technologies to support ubiquitous sensing functions. Different from the current works, a communications-sensing-intelligence converged network architecture is proposed to coordinate interference in this paper. Each base station equips with the individual deep neural networks to allocate power and beamforming. On this basis, the interference management is transformed into a functional optimization with stochastic constraints. An unsupervised learning algorithm is proposed to allocate power for interference management. Furthermore, a transfer learning method is presented to obtain the interference management in terms of transmit beamforming. Finally, the distributed management is obtained from the local channel state information in the multi-cell scenario. Simulation results verify the effectiveness of the proposed unsupervised learning interference management method in the ISAC systems.

H. Zhang, X. Liu, and K. Long are with Beijing Engineering and Technology Research Center for Convergence Networks and Ubiquitous Services, University of Science and Technology Beijing, Beijing, China, 100083 (email: haijunzhang@ieee.org, xiangnan.liu@xs.ustb.edu.cn, longkeping@ustb.edu.cn).

Arumugam Nallanathan is with the School of Electronic Engineering and Computer Science, Queen Mary, University of London (e-mail: a.nallanathan@qmul.ac.uk)

Victor C. M. Leung is with the Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC V6T 1Z4 Canada (e-mail: vleung@ece.ubc.ca).

Index Terms

Integrated sensing and communication, distributed interference management, unsupervised learning, transfer learning.

I. Introduction

With the development of integrated sensing and communications (ISAC) systems, ubiquitous sensing is required for multi-user communication in future wireless networks [1]. Widespread access and large-scale coordination generate more complicated signal interference problems. These problems include the clutter in harsh environments, the interference from multiple users in multi-cell communications, and self-interference caused by leakage of transmitter. Besides, the waveform design, transceiver design, and signal process algorithm in the ISAC system indicate that the radar subsystem and the communication subsystem are interdependent. On the one hand, the throughput and sum rate are impaired. On the other hand, dynamic radar detection range is decreased. Position precision and coverage capacity are weaken, correspondingly. Therefore, an effective and flexible interference management is worthy of further study in the surging ISAC systems [2].

Nowadays, the radar and communication subsystems are affected by each other due to the same spectrum. The interference in the communication-centric ISAC systems can be derived from two items. One is the multi-user communication signal interference and the other one is generated by the dedicated sensing streams [3]. The multi-user communication signal interfere with sensing and impairs target estimation [4]. Meanwhile, the interference caused by the dedicated sensing leads to a notable degradation in the transmitted symbols. Thus, more attention should be paid to mitigating two kinds of interference in the ISAC systems.

Existing researches on traditional interference management in cellular networks is fruitful, including the weighted minimum-mean-square-error (WMMSE) algorithm [5], semi-definite relaxation [6], and successive-convex-approximation (SCA) algorithm [7]. Meanwhile, there are some existing works on the interference management in the ISAC systems [8]–[11]. By exploiting the constructive multi-user communication signal interference, Liu et. al in [8] minimized the interference from the ISAC base station for a given quality of service and transmit power budget.

Barneto et. al in [9] utilized the combination of antenna separation to obtain the interference cancellation from both analog and digital aspects. With the aid of zero-force beamforming, the radar interference and multi-user communication interference can be eliminated [10]. Wang et. al in [11] proposed utilizing non-orthogonal multiple access technology to cancel interference from these two types. The part of the dedicated sensing signal can be mitigated via successive interference cancellation.

On the other hand, the distributed MIMO radar can offer improved sensing performance due to distributed antennas deployment. To save the unnecessary cost, these participating radars are deployed in the edge base stations. In other words, ISAC network architecture is suitable for the distributed MIMO radar. Ahmed et al. proposed a distributed dual-function radar-communication MIMO system firstly [12]. And the power allocation and target localization could be obtained in this system. Wu et al. designed a novel resource-aware strategy for heterogeneously-distributed joint radar and communications [13]. The network-level ISAC presented in [2] referred to the distributed antenna systems. The distributed ISAC systems can obtain the flexible and diverse performance to support ubiquitous sensing.

The high dimensionality caused by the variable and constraint is intractable. Considering that artificial intelligence (AI) algorithms have learning and memory capabilities, so they have advantages in the identification, avoidance, and utilization of interference signals [14]. AI algorithms obtain the prior information of interference signals to the maximum extent and transform them into sensing areas, so as to improve the anti-interference ability of the system. The unsupervised learning algorithm has gradually obtained recognition and attention from the academics [15].

The interference management in the ISAC system can be transformed into a constrained optimization [8], [10], [11]. The constrained problem could be non-convex, mostly. To solve constrained optimization problems, the primal-dual learning is a promising approach, which is based on Lagrangian relaxation and updates parameters in primal and dual spaces in turn. Besides, the transfer learning is an efficient method, which can transfer the developed models into the related problems [16]. The constrained optimization of interference management is not an exception. The concept of transfer learning is cost-saving in learning wireless resource management [17]. Similar to the wireless resource management, the transfer learning can be

applied to the interference management in the ISAC system.

Without training set and global channel state information (CSI), the distributed unsupervised learning technique is worthy of a try. Al-Abbasi et al. proposed the concept of DeepPool, a distributed model-free learning algorithm [18]. The work in [19] proposed a novel approach for ride sharing. A distributed adaptive unsupervised learning is utilized to minimize the travel distance and the average idle delay. However, the proposed technique can not detect the targets precisely with local CSI. The work in [20] paid attention to the interference channel estimation. The work in [21] provided a robust beamforming solution under the local interference CSI. The authors aimed to maximize the detect probability and to guarantee the constraint of downlink signal-to-interference-plus-noise ratio (SINR).

A large volume of data is generated from multi-cell ISAC systems, which requires the AI method to process properly. The interference management is not an exception [22]. The distributed learning has been envisioned as an efficient solution to distributed learning problems [23]. This shift from the centralized model-free learning technique to the distributed one attracts many researchers' attentions. With the help of the distributed learning concept, Wen et al. investigated multiple ISAC devices in edge AI system [24]. This scheme is superior to the conventional optimization algorithms in terms of communication performance and computing complexity. Utilizing the edge intelligence, a distributed unsupervised interference management can be considered in the ISAC systems.

The ISAC systems generate more information including sensing and communication and more abundant computing resources would be equipped in the future networks. How to utilize AI technology to obtain resource management in ISAC networks is a tendency in 6G networks. In this context, we investigate the distributed interference management for ISAC systems cooperating with edge intelligence. Firstly, we propose an unsupervised algorithm to obtain interference management in ISAC networks. Moreover, a transfer learning is utilized to obtain transmit beamforming, which can reduce the training cost. Furthermore, the more practical situation is considered, in which each base station has the local CSI. Herein, the distributed unsupervised interference management is designed in this paper. In particular, the main contributions of this paper are summarized below.

- *The communications-sensing-intelligence converged network architecture:* We propose a communications-sensing-intelligence converged network architecture where base stations are equipped with the corresponding edge intelligence to save overhead, while each base station also has a radar communication dual function. The multi-user communication interference and the interference imposed by dedicated sensing streams are considered in the proposed scheme.
- *The unsupervised learning method to serve interference management:* Compared with the need for sample training of traditional DNNs, an unsupervised learning method is deployed to obtain the interference management. The cost function iterates via primal-dual learning method. To optimize the spectral efficiency, the proposed unsupervised learning algorithm allocates power to aid the interference management. In this scheme, the developed training set is not required.
- *The transfer learning for solving transmit beamforming problem:* For the existing multi-access interference problem in the ISAC systems, we design a beamforming mechanism that considers the local CSI to reduce interference. We focus on utilizing transfer learning to obtain the interference management in terms of beamforming. Therefore, a transfer learning method is utilized to optimize the transmit beamforming in ISAC systems.
- *The distributed management to be applied in local CSI situation:* To be more realistic, the local CSI is considered. Accordingly, the distributed DNNs that can only needs the local knowledge is designed. The structure consists of a qualifier and an optimizer, which can obtain the distributed interference management in the ISAC systems.

The remainder of this paper is summarized as follows. The system model and the formulation of the interference management are presented in Section II. Section III presents the unsupervised learning method and transfer learning method with global CSI. In Section IV, the case of local CSI is considered and the distributed unsupervised algorithm is presented. The designed algorithms are evaluated by simulation results from a single base station and multiple base stations in Section V. The conclusions are given in Section VI.

II. System Model and Problem Formulation

A. System Model

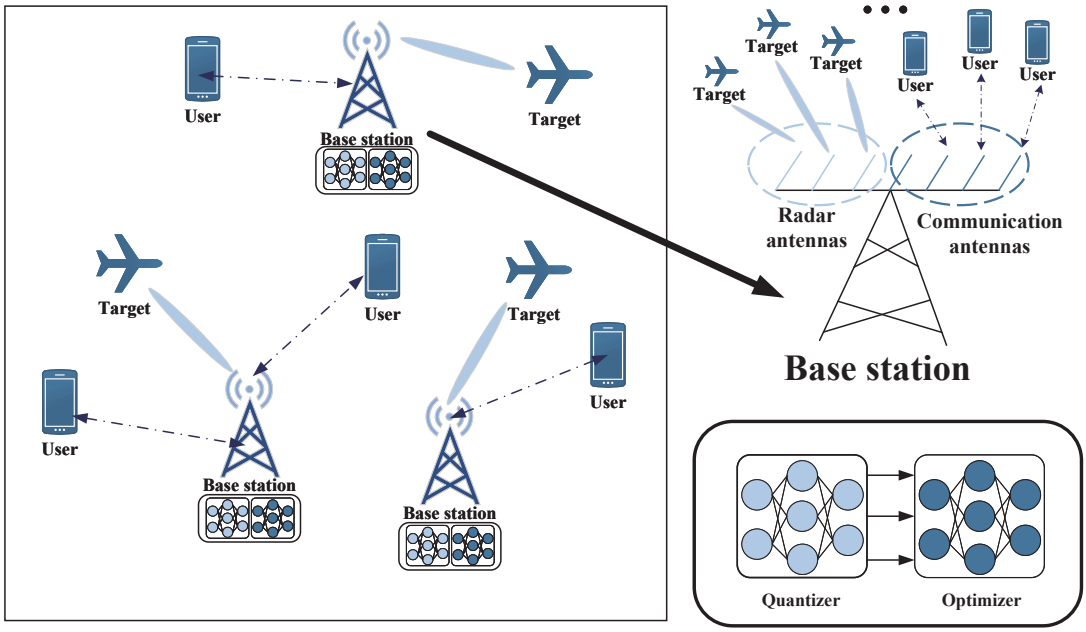

Fig. 1. The multi-user MISO scenario in the ISAC systems.

Fig. 1 presents a multi-user multiple-input single-output (MISO) scenario in the ISAC systems, in which each base station is equipped with edge intelligence to obtain interference management. J base stations with M antennas and K single-antenna users are deployed. The set of base stations, each base station's antennas, and served users are represented by $\mathcal{J} = \{1, 2, ..., J\}$, $\mathcal{M} = \{1, 2, ..., M\}$, and $\mathcal{K} = \{1, 2, ..., K\}$, respectively. Practically, the separated deployment is considered in this paper [8]. M_C and M_R denote the number of communication antennas and radar antennas for the k-th user, respectively. Note that $M_C + M_R = M$. This implies that the transmit antennas of base station j is divided into two dedicated functions. Here, $d_k[t]$ and $n_k[t]$ are the communication symbol and received noise at the time slot t, following $\mathcal{CN}(0, N_0)$. $\mathbf{s}_t$ is the radar signal and its sample covariance matrix $\mathbf{R} \in \mathbb{C}^{M_R \times M_R}$ is

$$\frac{1}{T}\sum_{t=1}^{T} \mathbf{s}_t \mathbf{s}_t^H = \mathbf{R}. \tag{1}$$

T is set as the length of the radar signal. The received signal of the user k after transmit beamforming is expressed as

$$y_k[t] = \mathbf{g}_k^T \sum_{l=1}^{K} \mathbf{F}_l d_l[t] + \mathbf{f}_k^T \mathbf{s}_t + n_k[t], \forall k, \tag{2}$$

where $\mathbf{g}_k \in \mathbb{C}^{M_C \times 1}$ and $\mathbf{f}_k \in \mathbb{C}^{M_R \times 1}$ are the channel vector of communication signal and radar signal, respectively. $\mathbf{F}_k \in \mathbb{C}^{M_C \times 1}$ is the beamforming vector of the user k. The channel vectors are set as $\mathbf{h}_k = [\mathbf{f}_k; \mathbf{g}_k] \in \mathbb{C}^{M \times 1}$. The channel follows flat Rayleigh fading. We assume that the channel $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, ..., \mathbf{h}_K] \in \mathbb{C}^{N \times K}$ is locally estimated via the pilot symbols. However, the global CSI can not be acquired by each base station.

The communication transmit power is constrained within the power budget $P_{\max}$,

$$\sum\nolimits_{k=1}^{K} \tau_k \cdot \|\mathbf{F}_k\|^2 \leq P_{\max}, \tag{3}$$

where $\tau_k \in \{0, 1\}$ is the k-th user's scheduling variable, and $\tau_k = 1$ indicates that the user (target) k is accessed successfully. Otherwise, $\tau_k = 0$. The scheduling variable follows the rules that each user occupies one channel at a time. Therefore, the access scheduling variable can be constrained in the following inequality:

$$\sum\nolimits_{k=1}^{K} \tau_k \leq 1. \tag{4}$$

According to the design concept of the MIMO radar probing signals [25], the design target of the MIMO radar transmit beamforming including two aspects. For one thing, the transmit power of optimized beamforming should be given at directions. It can ensure the transmit power matches the desired beam pattern. For another, the cross correlation pattern among signals should be reduced as soon as possible. These signals at several given directions and the cross correlation pattern of them has a impact on the performance of sensing performance.

The design of beampattern generally is transformed into the design of the covariance matrix $\mathbf{R}$. Herein, a constrained beampattern error minimization is provided,

$$\begin{aligned} \min_{\alpha, \mathbf{R}} \quad & \textstyle\sum_{l=1}^{L} \left| \alpha d(\theta_l) - \mathbf{a}^H(\theta_l) \mathbf{R} \mathbf{a}(\theta_l) \right|^2 \\ s.t. \quad & diag(\mathbf{R}) = \tfrac{P_{\max} \mathbf{1}}{N}, \\ & \mathbf{R} \succcurlyeq 0, \mathbf{R} = \mathbf{R}^H, \\ & \alpha \geqslant 0, \end{aligned} \tag{5}$$

where $\{\theta_l\}_{l=1}^{L}$ is defined as an angular grid that covers the range of $[-\pi/2, \pi/2]$. A uniform linear antenna is assumed to be deployed in the proposed ISAC systems, so the steering vector of the transmit antenna array $\mathbf{a}(\theta_l) \in \mathbb{C}^{M_R \times 1}$ is defined as $\sqrt{\frac{1}{M_R}}\left[1, e^{j2\pi d/\lambda \sin\theta_l}, ..., e^{j2\pi(M_R-1)d/\lambda \sin\theta_l}\right]^T$. The added constraints ensure that the waveform has the same average power. α is a scaling factor and we set α equals to 1. After setting the beam pattern error, another optimization problem for cross correlation has been proposed by Stoica et al. [25], which is given by

$$\begin{aligned} \min_{t,\mathbf{R}} \quad & -t \\ s.t. \quad & \mathbf{a}^H(\theta_0)\mathbf{R}\mathbf{a}(\theta_0) - \mathbf{a}^H(\theta_l)\mathbf{R}\mathbf{a}(\theta_l) \geqslant t, \quad \forall \theta_l \in \Omega, \\ & \mathbf{a}^H(\theta_1)\mathbf{R}\mathbf{a}(\theta_1) = \mathbf{a}^H(\theta_0)\mathbf{R}\mathbf{a}(\theta_0)/2, \\ & \mathbf{a}^H(\theta_2)\mathbf{R}\mathbf{a}(\theta_2) = \mathbf{a}^H(\theta_0)\mathbf{R}\mathbf{a}(\theta_0)/2, \\ & \mathbf{R} \geqslant 0, \mathbf{R} = \mathbf{R}^H, \\ & diag(\mathbf{R}) = \frac{P_0\mathbf{1}}{N}. \end{aligned} \tag{6}$$

For better problem formulation in the proposed ISAC system, the beam pattern error and mean-squared cross correlation pattern are introduced in the problem. The beam pattern error is one of the key sensing performance metrics in the proposed system [10]. It is aimed to optimize beam in provided desired directions. $L_{r,1}(\mathbf{R})$ is defined as the beam pattern error and it can be given by

$$L_{r,1}(\mathbf{R}) = \frac{1}{L}\sum_{l=1}^{L} |d(\theta_l) - P(\theta_l)|^2, \tag{7}$$

where $P(\theta) = \mathbf{a}^H(\theta)\mathbf{R}\mathbf{a}(\theta)$ is transmit power in direction θ and the desired power $d(\theta)$ is calculated by

$$d(\theta) = \begin{cases} 1, & \theta_p - \Delta/2 \leq \theta \leq \theta_p + \Delta/2, p = 1, 2, 3, \\ 0, & otherwise, \end{cases} \tag{8}$$

where $\{\theta_l\}_{l=1}^{L}$ are sampled from -90° to 90° with the resolution of 1°. The subscript of p in (8) symbolizes the index of the ideal beam patterns. Problem (5) is specified by $L_{r,1}(\mathbf{R})$ in formula (7).

Meanwhile, the loss function is also derived from the mean-squared cross correlation pattern $L_{r,2}(\mathbf{R})$, which is calculated as

$$L_{r,2}(\mathbf{R}) = \frac{2}{L^2 - L}\sum_{l=1}^{L-1}\sum_{r=l+1}^{L} |P_c(\theta_l, \theta_r; \mathbf{R})|, \tag{9}$$

where $P_c\left(\theta_1,\theta_2;\mathbf{R}\right)=\mathbf{a}^H\left(\theta_1\right)\mathbf{R}\mathbf{a}\left(\theta_2\right)$ is the cross correlation pattern in directions of θ_1 and θ_2. Problem (6) is specified by $L_{r,2}\left(\mathbf{R}\right)$ in formula (9). Different from the sensing-centric problem in the existing literature [10], the communication-centric problem is considered. The beam pattern error and cross correlation pattern are transformed into the constraint rather than the objective. The beam pattern error and cross correlation pattern can be set as a constraint.

$$L_{r,1}\left(\mathbf{R}\right)+L_{r,2}\left(\mathbf{R}\right)\leqslant\varepsilon. \tag{10}$$

Constraint (10) implies that the error can not exceed the threshold ε, which ensures the resolution of radar detection.

Furthermore, the k-th user's SINR can be given by [26]

$$\gamma_k=\frac{\left|\mathbf{g}_k^T\mathbf{F}_k\right|^2}{\sum\limits_{l\neq k}^{K}\left|\mathbf{g}_k^T\mathbf{F}_l\right|^2+\mathbf{f}_k^T\mathbf{R}\mathbf{f}_k^*+N_0}, \tag{11}$$

where $\sum\limits_{l\neq k}^{K}\left|\mathbf{g}_k^T\mathbf{F}_l\right|^2$ in the denominator is the multi-user interference [1] and $\mathbf{f}_k^T\mathbf{R}\mathbf{f}_k^*$ is the interference imposed by dedicated sensing streams.

B. Problem Formulation

To the best of our knowledge, resource management in wireless communication system can be evaluated by the instantaneous performance $f\left(\mathbf{h},\mathbf{p}\left(\mathbf{h}\right)\right)$, such as sum rate, power consumption, energy efficiency. Here, $\mathbf{h}$ represents unstable and changing wireless environment, such as channel realization, user access situation, service requirement, etc. Practically, the wireless communication environment includes uncertain factors, which results in the the unstable and changing wireless environment. $\mathbf{p}\left(\mathbf{h}\right)$ denotes corresponding instantaneous resource allocation. Nevertheless, the instantaneous system performance varies frequently. The long-term mathematical expectations $\mathbb{E}_{\mathbf{h}}\left[f\left(\mathbf{h},\mathbf{p}\left(\mathbf{h}\right)\right)\right]$ is a more valuable metric. It can be given by

$$\begin{aligned}\max_{\mathbf{p}}\quad & \mathbb{E}_{\mathbf{h}}\left[f\left(\mathbf{h},\mathbf{p}\left(\mathbf{h}\right)\right)\right]\\ s.t.\quad & \mathbb{E}_{\mathbf{h}}\left[g_k\left(\mathbf{h},\mathbf{p}\left(\mathbf{h}\right)\right)\right]\leqslant G_k,\\ & k=1,...,K.\end{aligned} \tag{12}$$

[1]The indexes $k\in\mathcal{K}$, $l\in\mathcal{K}$ symbolize the index of user. The index k denotes the communication user, the index l is symbolized as interference one.

As a state-of-the-art wireless network architecture, ISAC systems also can be applied into the optimization problem. Similarly, the interference management can be designed as a long-term instantaneous performance function [27]. In this paper, $\mathbf{h}$ is the channel state and $\mathbf{p}(\mathbf{h})$ is the transmit power matrix $\mathbf{p}$ or transmit beamforming matrix $\mathbf{F}$. $f(\mathbf{h}, \mathbf{p})$ or $f(\mathbf{h}, \mathbf{F})$ can be the data transmission rate $\log(1+\gamma_k)$. The objective function in problem (12) can be given by

$$f(\mathbf{h}, \mathbf{p}(\mathbf{h})) = \sum\nolimits_{k=1}^{K} \log(1+\gamma_k)\tau_k \prod_{k\neq l}^{K} (1-\tau_l). \tag{13}$$

Since users contend for channel access, transmission of user k in the current time is successful if and only if $\tau_k = 1$ and $\tau_l = 1$ for all $k \neq l$. Combined with constraints (3), (4), and (10), the goal of interference management in this paper is to realize the capacity maximization. The optimization problem of multi-user MISO scenario in the ISAC systems can be formulated as

$$\begin{aligned} \max_{\mathbf{p},\boldsymbol{\tau}} \quad & \mathbb{E}_{\mathbf{h}}\left[\sum\nolimits_{k=1}^{K} \log(1+\gamma_k)\tau_k \prod_{k\neq l}^{K} (1-\tau_l)\right] \\ s.t. \quad & \tau_k |p_k|^2 \leq P_{\max}, \\ & L_r(\mathbf{R}) \leq \varepsilon. \end{aligned} \tag{14}$$

The mean square error (MSE) is unified as $L_r(\mathbf{R}) = L_{r,1}(\mathbf{R}) + L_{r,2}(\mathbf{R})$. Problem (12) is a non-convex problem so the traditional convex method is intractable to problem (12). In this context, an optimization method utilizing deep learning is introduced in the next section.

III. Unsupervised learning with the global CSI

In this section, the feasibility of the unsupervised learning interference management is discussed and the approximation of DNNs is confirmed to be applied to problem (14). Subsequently, we introduce the primal-dual learning optimization to obtain the solution of problem (14). And updating the iterative parameters and dual variables updating is also presented in detail. Finally, a transfer learning algorithm is introduced to solving transmit beamforming, which can obtain interference management.

A. Learning to Interference Management

Deep learning techniques have been applied to wireless resource allocation such as cognitive radio networks [28], interference channels [29], and simultaneous wireless information-power

transmission networks [30]. To solve problem (14), the allocation policy $\mathbf{p}(\mathbf{h})$ is parameterized by a DNN $\boldsymbol{\pi}(\mathbf{h};\boldsymbol{\omega})$ as

$$\mathbf{p}(\mathbf{h}) = \boldsymbol{\pi}(\mathbf{h};\boldsymbol{\omega}). \tag{15}$$

The set of parameters $\boldsymbol{\omega}_r$ consists of weights $\mathbf{w}_r$ and bias $\mathbf{b}_r$ of each layer r. The DNN is feed-forward and composed of R fully-connected layers.

Utilizing enough linear and non-linear operations, it can be accurately approximated by $\mathbf{p}(\mathbf{h})$. For any $\varepsilon > 0$, $\boldsymbol{\pi}(\mathbf{h};\boldsymbol{\omega})$ can obtain the approximation with enough R layers and activation functions.

$$\sup_{\mathbf{h}\in\mathcal{H}} \|\mathbf{p}(\mathbf{h}) - \boldsymbol{\pi}(\mathbf{h};\boldsymbol{\omega})\| \leq \varepsilon. \tag{16}$$

The utilization of the universal function approximation property in DNNs can be suitable for problem (14). Constraint (16) utilizes the universal approximation theorem [31], for a given set $\mathcal{H}$. There exists a set parameter $\boldsymbol{\omega}$ such that the DNNs' structures in (16) can approximate any optimized policy $\boldsymbol{\pi}(\mathbf{h};\boldsymbol{\omega})$ with slight positive error ε.

Utilizing enough linear and non-linear operations, it can be accurately approximated by $\mathbf{p}(\mathbf{h})$. For any $\varepsilon > 0$, $\boldsymbol{\pi}(\mathbf{h};\boldsymbol{\omega})$ can obtain the approximation with enough R layers and activation functions. There exist a set parameter such that the associated DNN of the structure in (16) can approximate any continuous function with arbitrary small positive error ε. The proof of (16) refers to [32].

Proof: Pick R such that $1/R < \varepsilon/2$. For $r \in \{1, 2, ..., R-1\}$ set $\beta_r = 1/R$. Pick $M > 0$ such that $\xi(-M) < \varepsilon/2R$ and $\xi(M) > 1 - \varepsilon/2R$. Because ξ is a sqashing function such an M can be found. For $r \in \{1, 2, ..., R-1\}$ set $o_r = \sup\{\boldsymbol{\omega} : \boldsymbol{\pi}(\mathbf{h};\boldsymbol{\omega}) = o/R\}$. Set $r_R = \sup\{\boldsymbol{\omega} : \boldsymbol{\pi}(\mathbf{h};\boldsymbol{\omega}) = 1 - 1/2R\}$. Because $\mathbf{p}(\cdot)$ is a continuous sqashing function such o_r's exist. For any $r < s$ let $A_{r,s} \in A$ be the unique affine function satisfying $A_{r,s} = M$ and $A_{r,s} = -M$. The desired approximation is then $\boldsymbol{\pi}(\mathbf{h};\boldsymbol{\omega}) = \sum_{r=1}^{R-1} \beta_i \xi(A_r(\lambda))$. It is easy to check each of the intervals $(-\infty, r_1], (r_1, r_2], ..., (r_{R-1}, r_R], (r_R, +\infty]$, we have $\sup_{\mathbf{h}\in\mathcal{H}} \|\mathbf{p}(\mathbf{h}) - \boldsymbol{\pi}(\mathbf{h};\boldsymbol{\omega})\| \leq \varepsilon$.

B. Primal-dual Learning Optimization

The Lagrangian function for problem (12) can be expressed by

$$\begin{aligned}\mathcal{L}(\boldsymbol{\omega},\boldsymbol{\lambda}) = \;& \mathbb{E}_{\mathbf{h}}\left[f\left(\mathbf{h},\boldsymbol{\pi}\left(\mathbf{h};\boldsymbol{\omega}\right)\right)\right] \\ & + \textstyle\sum_{k=1}^{K}\lambda_k\left(\mathbb{E}_{\mathbf{h}}\left[g_k\left(\mathbf{h};\boldsymbol{\pi}\left(\mathbf{h};\boldsymbol{\omega}\right)\right)\right]-G_k\right).\end{aligned} \tag{17}$$

The non-negative variables $\boldsymbol{\lambda} \triangleq \{\lambda_k, \forall k\}$ are associated with constraints $\mathbb{E}_{\mathbf{h}}\left[g_k\left(\mathbf{h},\boldsymbol{\pi}\left(\mathbf{h};\boldsymbol{\omega}\right)\right)\right]-G_k$. The dual function $\mathcal{D}_{\pi}\left(\boldsymbol{\lambda}\right)$ of problem (12) is defined by

$$\mathcal{D}_{\pi}\left(\boldsymbol{\lambda}\right) = \min_{\omega}\mathcal{L}\left(\boldsymbol{\omega},\boldsymbol{\lambda}\right), \tag{18}$$

The dual function $\mathcal{D}\left(\boldsymbol{\lambda}\right)$ is written as

$$\begin{aligned}&\max_{\lambda}\mathcal{D}\left(\boldsymbol{\lambda}\right) \\ &s.t.\lambda_k \geqslant 0 \quad k=1,...,K.\end{aligned} \tag{19}$$

The primal variables can be updated by

$$\begin{aligned}\boldsymbol{\omega}^{t} &= \boldsymbol{\omega}^{t-1}+\eta\nabla_{\boldsymbol{\omega}}\mathcal{L}\left(\boldsymbol{\omega}^{t-1},\boldsymbol{\lambda}^{t-1}\right) \\ &= \boldsymbol{\omega}^{t-1}+\eta\left(\mathbb{E}_{\mathbf{h}}\left[\nabla_{\boldsymbol{\omega}}f\left(\mathbf{h},\boldsymbol{\pi}\left(\mathbf{h};\boldsymbol{\omega}^{t-1}\right)\right)\right]\right. \\ &\quad \left.+\textstyle\sum_{k=1}^{K}\lambda_k^{t-1}\left(\mathbb{E}_{\mathbf{h}}\left[\nabla_{\boldsymbol{\omega}}g_k\left(\mathbf{h},\boldsymbol{\pi}\left(\mathbf{h};\boldsymbol{\omega}^{t-1}\right)\right)\right]-G_k\right)\right.\end{aligned} \tag{20}$$

The dual variables λ_k^t is updated by the subtracting the gradients and projecting onto the positive orthant:

$$\begin{aligned}\lambda_k^{t} &= \left(\lambda_k^{t-1}-\eta\nabla_{\lambda_k}\mathcal{D}\left(\boldsymbol{\lambda}^{t-1}\right)\right)_{+} \\ &= \left(\lambda_k^{t-1}-\eta\left(\mathbb{E}_{\mathbf{h}}\left[\nabla_{\boldsymbol{\omega}}g_k\left(\mathbf{h},\boldsymbol{\pi}\left(\mathbf{h};\boldsymbol{\omega}^{t-1}\right)\right)\right]-G_k\right)\right)_{+}.\end{aligned} \tag{21}$$

The gradients $\nabla_{\boldsymbol{\omega}}f\left(\mathbf{h},\boldsymbol{\pi}\left(\mathbf{h};\boldsymbol{\omega}\right)\right)$ and $\nabla_{\boldsymbol{\omega}}g_k\left(\mathbf{h},\boldsymbol{\pi}\left(\mathbf{h};\boldsymbol{\omega}\right)\right)$ are calculated by the chain rule:

$$\begin{aligned}\nabla_{\boldsymbol{\omega}}f\left(\mathbf{h},\boldsymbol{\pi}\left(\mathbf{h};\boldsymbol{\omega}\right)\right) &= \nabla_{\boldsymbol{\pi}}f\left(\mathbf{h},\boldsymbol{\pi}\left(\mathbf{h};\boldsymbol{\omega}\right)\right)\cdot\nabla_{\boldsymbol{\omega}}\boldsymbol{\pi}\left(\mathbf{h};\boldsymbol{\omega}\right), \\ \nabla_{\boldsymbol{\omega}}g_k\left(\mathbf{h},\boldsymbol{\pi}\left(\mathbf{h};\boldsymbol{\omega}\right)\right) &= \nabla_{\boldsymbol{\pi}}g_k\left(\mathbf{h},\boldsymbol{\pi}\left(\mathbf{h};\boldsymbol{\omega}\right)\right)\cdot\nabla_{\boldsymbol{\omega}}\boldsymbol{\pi}\left(\mathbf{h};\boldsymbol{\omega}\right).\end{aligned} \tag{22}$$

However, the updates of primal variables $\boldsymbol{\omega}^t$ and dual variables $\boldsymbol{\lambda}^t$ need the perfect knowledge of observations from other base stations. Practically, The knowledge of the global CSI can not be available.

Algorithm 1 Iterative Unsupervised Learning Optimization Algorithm

1: **Input** Initialize T_{max}, parameters of DNNs $\boldsymbol{\omega}^0 := \{\omega_k^0, \forall k\}$ and dual variables $\boldsymbol{\lambda}^0 := \{\lambda_k^0, \forall k\}$, set $t = 0$.
2: **repeat**
3: a) Sample a mini-batch set $\mathcal{S} \subset \mathcal{H}$;
4: b) The gradients $\nabla_{\boldsymbol{\omega}} f\left(\mathbf{h}, \boldsymbol{\pi}\left(\mathbf{h}; \boldsymbol{\omega}\right)\right)$ and $\nabla_{\boldsymbol{\omega}} g_k\left(\mathbf{h}, \boldsymbol{\pi}\left(\mathbf{h}; \boldsymbol{\omega}\right)\right)$ are calculated by (22);
5: c) Update parameters of DNNs $\boldsymbol{\omega}^t$ through (20);
6: d) Update dual variables $\boldsymbol{\lambda}^t$ through (21);
7: e) The step size updates $t \to t + 1$.
8: **until** Convergence or $t = T_{max}$

Notably, the pseudocode of the iterative unsupervised learning optimization algorithm is summarized in Algorithm 1. In the stage of initialization, the maximum number of iterations T_{max}, parameters of DNNs $\boldsymbol{\omega}^0 := \{\omega_k^0, \forall k\}$ and the corresponding dual variables $\boldsymbol{\lambda}^t$ are initialized. Then the algorithm steps into iterations. The mini-bath gradient method is deployed for more efficient training. Subsequently, the gradients of performance function and constraint function are calculated by formula (22). The parameters of DNNs are updated in formula (20). The corresponding dual variables are also updated by formula (21). The algorithm does not end until the convergence is obtained or $t = T_{max}$. The analytical results would be verified in Section Simulation Results and Discussions.

C. Transfer Learning Optimization for Transmit Beamforming

With the emergence of more and more application scenarios of machine learning wireless communication networks, the data in ISAC systems can be used to guide power control, bandwidth allocation, and beamforming, and if the existing data and model training can be used to solve different resource allocation problems, transfer learning has obtained more and more attention. Transfer learning can be implied in the hardware heterogeneity for different settings. In other words, if an resource allocation algorithm is well designed, the related resource allocations also can be solved efficiently by the transfer learning method.

Transfer learning applies knowledge or patterns learned in a domain or task to different but related domains or problems. Deep learning requires a large amount of high-quality labelled data, the technique of pre-training combined with fine-tuning is now a very popular trick in deep learning, especially in the field of images, many times will choose pre-trained ImageNet to initialize the model. Interference management of ISAC systems can be obtained through power control, but in functions of communication and sensing, it is necessary to extend power control to beamforming in the complex domain. While power control and beamforming are related, the combination of DNN networks and transfer learning ideas under existing interference management can further realize the beamforming strategy under optimal spectral efficiency [35]. In the scheme of transfer learning, there exist a domain $D = \{F, P(X)\}$ and a task $T = \{Y, f(\cdot)\}$. F represents feature and $P(X)$ is distribution. Let Y and $f(\cdot)$ denote the label and predict function, respectively.

Correspondingly, D_p is denoted as real-domain power control parameter and D_w is denoted as complex-domain beamforming parameter. The transfer learning is utilized to solve transmit beamforming as follows:

$$\begin{aligned} &\arg\min l\left(f_w\left(X_w^t, D_p^t, T_p^t\right), Y_w^t\right) \\ &s.t. D_p^t \neq D_w^t, T_p^t \neq T_w^t. \end{aligned} \tag{23}$$

where f_w^t is the predict function of transmit beamforming, and it can be represented by spectral efficiency. And $f_w\left(X_w^t, D_p^t, T_p^t\right)$ is calculated by the existing domain D_p^t and task T_p^t.

$$f_w^t = \sum\nolimits_{k=1}^{K} \log(1+\gamma_k). \tag{24}$$

Tracing back to the initial optimization problem (14), we can change the local output layer through the power-control DNNs to obtain further control of transmit beamforming.

$$\begin{aligned} \max_{\mathbf{F},\boldsymbol{\tau}} \quad & \mathbb{E}_{\mathbf{h}}\left[\sum\nolimits_{k=1}^{K} f_w(t)\,\tau_k \prod_{k\neq l}^{K}(1-\tau_l)\right] \\ s.t. \quad & \tau_k|\mathbf{F}_k|^2 \leq P_{\max}, \\ & L_r(\mathbf{R}) \leq \varepsilon. \end{aligned} \tag{25}$$

For the corresponding algorithm implementation, we still use the original dual learning optimization, and the same allocation strategy will change accordingly on the original basis $\boldsymbol{\pi}(\mathbf{h};\boldsymbol{\omega})$

$$\boldsymbol{\pi}_w(\mathbf{h};\boldsymbol{\omega};\boldsymbol{\phi}) = \psi_R\left(\xi_{R-1}\left(...\xi_1\left(\xi_0\left(\mathbf{h};\boldsymbol{\omega}_0\right);\boldsymbol{\omega}_1\right)...;\boldsymbol{\omega}_{R-1}\right);\boldsymbol{\phi}_R\right). \tag{26}$$

The output of trained DNNs is replaced by $N \times 2$ transmit beamforming policy $\boldsymbol{\pi}_w(\mathbf{h};\boldsymbol{\omega})$, rather than $N \times 1$ power control policy $\boldsymbol{\pi}(\mathbf{h};\boldsymbol{\omega})$.The last output layer is changed with $\boldsymbol{\phi}_R$ and $\psi_R(\cdot)$. The transfer learning can reduce the overhead of training, because the power control and transmit beamforming have relationship. The developed model is suitable for the transmit beamforming, which is oriented to power control.

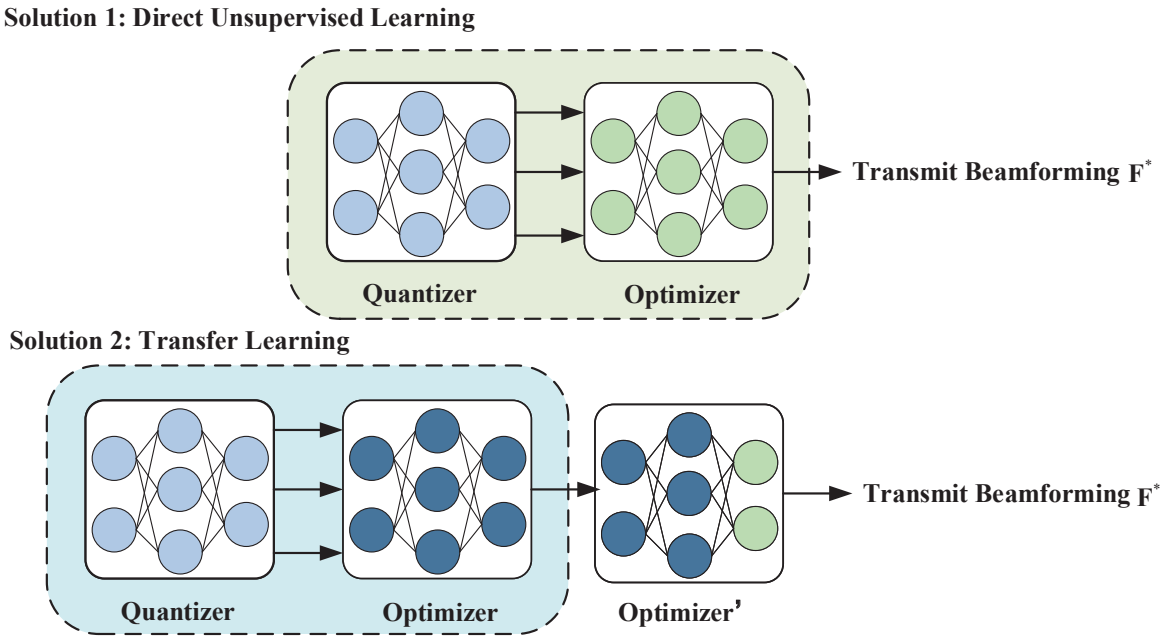

Fig. 2. The principle of the transfer learning for interference management.

As shown in Fig. 2, the interference management for transmit beamforming can be obtained by two solutions. The one is direct unsupervised learning and the other one is transfer learning. There are no obvious differences between power control and transmit beamforming in direct unsupervised learning. A new optimizer is constructed and its output will change from one-dimension power control to two-dimension transmit beamforming. The shortcoming of Solution 1 is that it can not learn more complicated CSI efficiently. To overcome this problem and make full use of the occupied knowledge, the transfer learning is adopted to solve the transmit beamforming. In the simulation results, the last output layer is replaced by a new linear layer, in which the number of elements is set as $2MK$. The rest of settings are in accordance with the interference management in terms of power control. In this context, the beamforming in

ISAC systems is obtained to avoid unnecessary training cost. The analytical result implies that the transfer learning method can obtain the same performance under the less training cost, compared with the direct unsupervised learning method.

IV. Distributed unsupervised learning with the local CSI

The distributed unsupervised learning in the ISAC systems will be introduced in this section. In many cases, a kind of imperfect local CSI should be considered, since the quantitative information of neighbors can save the distributed computing power of individual base stations. To minimize the overhead of the total systems, each base station can only obtain local CSI, and then quantifies its information status and passes it to its neighboring base stations through a link with limited capacity. This quantitative information from other base stations can further save the distributed power consumption of individual base stations. The set of base stations is represented by $\mathcal{J} = \{1, 2, ..., J\}$ and the index of base station is denoted as j. In distributed architecture, each edge base station is equipped with the function of communication and sensing. Meanwhile, due to the distributed computation nature, each base station has the local process capacity.

The proposed distributed framework consist of a stochastic binarization quantizer and a distributed learning optimizer in each base station. As it is shown in Fig. 3, a base station is equipped with two DNNS, including a quantizer and a optimizer. Let $\boldsymbol{\omega}_{Q,j}$ and $\boldsymbol{\omega}_{D,j}$ denote the parameters of quantizer and optimizer, respectively. The channel state $\mathbf{h}_j$ imported to the quantizer and optimizer of the base station. The quantizer produces the quantify information $\mathbf{q}_j$ to concatenate the channel state $\mathbf{h}_j$ to the optimizer. In Section III, the unsupervised learning algorithm and its transfer learning algorithm can be applied into the distributed framework. The difference between these two modes lies on the quantify operation.

A. Stochastic Binarization Quantizer

The quantizer of base station j employs a vector $\mathbf{q}_j(\cdot)$ to quantize the information from neighbouring base stations. $\mathbf{q}_j(\cdot)$ is regarded as a quantization noise and the length is L_j.

$$\mathbf{v}_j = \mathbf{q}_j(\mathbf{h}_j), j = 1, ..., \mathcal{J}. \tag{27}$$

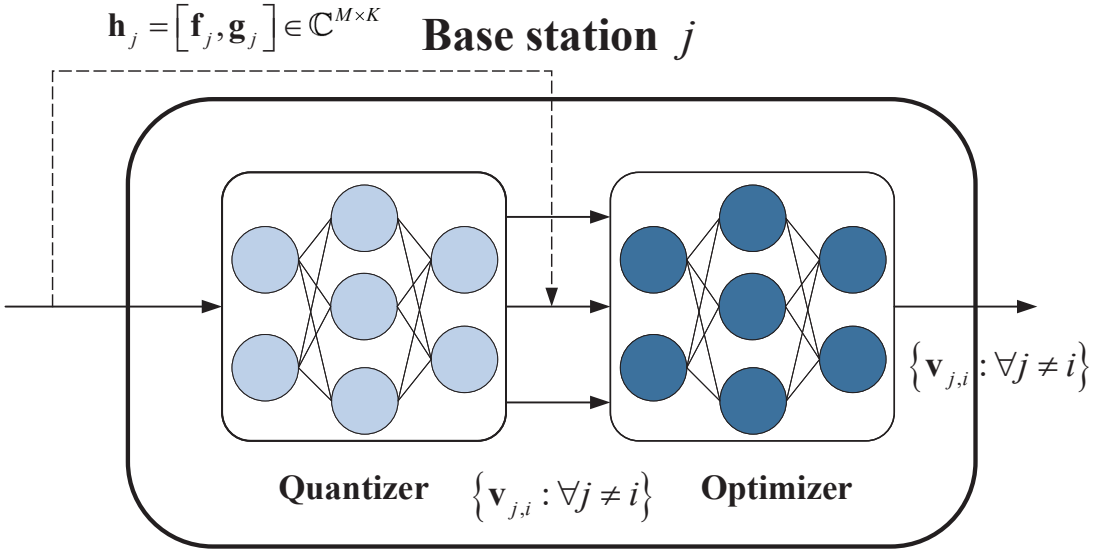

Fig. 3. The DNNs' structure of base station in the ISAC systems.

$\mathbf{v}_j \triangleq [v_{j,1}, ..., v_{j,L}] \in \{-1, +1\}^{L_j}$, which is produced by the hyperbolic tangent function tanh(·). This is because the output of the quantizer DNN is activated by the tanh(·). However, the output $\mathbf{v}_j$ is not discrete. We should transform continuous $\mathbf{v}_j$ into the discrete state. Thus, an affine function $\beta(x) = \frac{1+x}{2}$ is introduced to guarantee that the value of the output is in the two possibilities of -1 and 1. A scalar $q_{j,i}$ is the i-th quantize result in base station j.

$$q_{j,i} = \begin{cases} 1 - v_{j,i}, & (\beta(v_{j,i})). \\ -1 - v_{j,i}, & (1 - \beta(v_{j,i})). \end{cases} \tag{28}$$

Formula (28) is a binarization operation. The relationship between $\mathbf{q}_j$ in (27) and $q_{j,i}$ in (28) is $\mathbf{q}_j \triangleq [q_{j,1}, ..., q_{j,L}]$. Following obtaining the reasonable $q_{j,i}$, a discrete output for the quantizer $\widehat{\mathbf{v}}_j$ can be obtained by

$$\widehat{\mathbf{v}}_j = \mathbf{v}_j + \mathbf{q}_j. \tag{29}$$

Subsequently, according to the status information $\mathbf{h}_j$ and calculation rules $\mathbf{c}_j$ of the transmission, the transmission of the channel state $\mathbf{h}_j$ to the tipping point is obtained. Let $\mathbf{z}_{D,j}(\cdot)$ denote the distributed policy optimization of base station j, which is the output of the optimizer of the base station. The function of $\mathbf{z}_{D,j}$ in (30) is a multi-layer fully-connected DNN.

$$\boldsymbol{\pi}_j = \mathbf{z}_{D,j}(\mathbf{c}_j), j = 1, ..., \mathcal{J}, \tag{30}$$

where $\mathbf{c}_j$ is concatenation of the channel state $\mathbf{h}_j$ and the neighbour association information $\mathbf{v}_{j,i}$. The subscript of $\mathbf{z}_{D,j}(\cdot)$ means distributed optimizer. It can be calculated by

$$\mathbf{c}_j = \mathbf{h}_j \oplus \widehat{\mathbf{v}}_{j,l}, \tag{31}$$

where $\oplus$ is the concatenation operation. $\widehat{\mathbf{v}}_{j,l} \triangleq \widehat{\mathbf{v}}_j, \forall j \neq l$ is discrete the quantize result between base station j and base station l [2]. Since formula (29) requires the local CSI observation and global learning optimization, it can be regarded as a distributed approach. The developed $\mathbf{c}_j$ is imported to the distributed optimizer, similar to $\mathbf{h}$ in the centralized scenario.

B. Distributed Learning Optimizer

By utilizing the quantize operation and concatenation, a distributed mode is designed by the each base station's local CSI. This distributed approach is given by

$$\begin{aligned} \max_{\mathbf{F},\boldsymbol{\tau}} \quad & \mathbb{E}_{\mathbf{h}}\left[f\left(\mathbf{h}_j, \mathbf{z}_{D,j}\left(\mathbf{c}_j\right)\right)\right] \\ \text{s.t.} \quad & \mathbb{E}_{\mathbf{h}}\left[g_k\left(\mathbf{h}_j, \mathbf{z}_{D,j}\left(\mathbf{c}_j\right)\right)\right] \leq G_k, \\ & j = 1, .., J, k = 1, ..., K. \end{aligned} \tag{32}$$

Each base station can be equipped with an optimizer to learn and take charge of updating primal and dual variables. The corresponding updating rules can be given by

$$\begin{aligned} \boldsymbol{\omega}_j^t &= \boldsymbol{\omega}_j^{t-1} - \eta \nabla_{\boldsymbol{\omega}} \mathcal{L}\left(\boldsymbol{\omega}_j^{t-1}, \lambda_j^{t-1}\right) \\ &= \boldsymbol{\omega}_j^{t-1} - \eta \left(\mathbb{E}_{\mathbf{h}}\left[\nabla_{\boldsymbol{\omega}} f\left(\mathbf{h}_j, \boldsymbol{\pi}_j\left(\mathbf{c}_j; \boldsymbol{\omega}_j^{t-1}\right)\right)\right]\right. \\ &\quad \left. + \sum_{k=1}^{K} \lambda_{j,k}^{t-1}\left(\mathbb{E}_{\mathbf{h}}\left[\nabla_{\boldsymbol{\omega}} g_k\left(\mathbf{h}_j, \boldsymbol{\pi}_j\left(\mathbf{c}_j; \boldsymbol{\omega}_j^{t-1}\right)\right)\right] - G_{j,k}\right)\right. . \end{aligned} \tag{33}$$

$$\begin{aligned} \boldsymbol{\lambda}_{j,k}^t &= \left(\boldsymbol{\lambda}_{j,k}^{t-1} + \eta \nabla_{\boldsymbol{\lambda}_{j,k}} \mathcal{D}\left(\boldsymbol{\lambda}_j^{t-1}\right)\right)_+ \\ &= \left(\boldsymbol{\lambda}_{j,k}^{t-1} + \eta\left(\mathbb{E}_{\mathbf{h}}\left[\nabla_{\boldsymbol{\omega}} g_{j,k}\left(\mathbf{h}_j, \boldsymbol{\pi}_j\left(\mathbf{c}_j; \boldsymbol{\omega}_j^{t-1}\right)\right)\right] - G_{j,k}\right)\right)_+ . \end{aligned} \tag{34}$$

If we trace back to the proposed problem, the Lagrangian function of problem (11) can be given by

$$\begin{aligned} \mathcal{L}\left(\boldsymbol{\omega}, \boldsymbol{\lambda}, \boldsymbol{\mu}\right) = \; & \mathbb{E}_{\mathbf{h}}\left[\sum_{k=1}^{K} \log\left(1 + \gamma_k\right)\right] \\ & + \sum_{k=1}^{K} \lambda_{j,k}\left(\|\mathbf{w}_{j,k}\| - P_{\max}\right) + \sum_{k=1}^{K} \mu_{j,k}\left(L_{j,r}\left(\mathbf{R}_j\right) - \varepsilon\right). \end{aligned} \tag{35}$$

[2]The indexes $j \in \mathcal{J}$ and $l \in \mathcal{J}$ denote the base station. The index j denote the current base station, the index l represent the neighbour base stations.

Due to two constraints in the designed problem (11), $\boldsymbol{\mu}_j$ is introduced to act as Lagrangian multiplier, just like $\boldsymbol{\lambda}_j$. Similar to formula (20), the primal variable $\boldsymbol{\omega}_j$ in each base station is updated by

$$\begin{aligned} \boldsymbol{\omega}_j^t &= \boldsymbol{\omega}_j^{t-1} - \eta \left(\mathbb{E}_{\mathbf{h}} \left[\nabla_{\boldsymbol{\omega}} \sum_{k=1}^{K} \log\left(1+\gamma_{j,k}\right) \right] \right. \\ &+ \sum_{k=1}^{K} \lambda_{j,k}^{t-1} \left(\|\mathbf{w}_{j,k}\| - P_{\max} \right) \\ &+ \left. \sum_{k=1}^{K} \mu_{j,k}^{t-1} \left(L_{j,r}\left(\mathbf{R}_j\right) - \varepsilon \right) \right). \end{aligned} \tag{36}$$

The dual variables $\boldsymbol{\lambda}_j$ and $\boldsymbol{\mu}_j$ are given as

$$\lambda_{j,k}^t = \left(\lambda_{j,k}^{t-1} + \eta \left(\|\mathbf{w}_{j,k}\| - P_{\max} \right) \right)_+, \tag{37}$$

$$\mu_{j,k}^t = \left(\mu_{j,k}^{t-1} + \eta \left(L_{j,r}\left(\mathbf{R}_j\right) - \varepsilon \right) \right)_+. \tag{38}$$

C. Algorithm Analysis

In this subsection, the procedures of the algorithm's implementation are provided. Meanwhile, the analysis of the proposed algorithm is elaborated.

Note that in Algorithm 2, the policy is first initialized, such as initial power $\mathbf{p}^0$. Meanwhile the corresponding policy parameters $\boldsymbol{\omega}^0 := \{\mathbf{w}^0; \mathbf{b}^0\}$ are generated. The weight factors $\mathbf{w}^0$ are derived from the truncated Gaussian distribution and the initial bias $\mathbf{b}^0$ is set as a constant. Then the iteration begins and each base station quantizes its local CSI $\mathbf{h}_j$ in formula (29). The obtained local CSI and the neighboring base stations' association information $\mathbf{v}_{j,i}$ are concated. The output $\mathbf{c}_j$ is transformed to each base station's optimizer input to conduct iterative primal-dual learning optimization, which is elaborated in Algorithm 1. After the performance function $f\left(\mathbf{h}_j, \mathbf{z}_{D,j}\left(\mathbf{c}_j\right)\right)$ is obtained, the corresponding base station j's optimizer is saved. The transfer learning for transmit beamforming launches. Firstly, the trained optimizer is loaded and the structure of original DNN is modified in formula (26). Secondly, the last layer's parameters of DNN $\boldsymbol{\phi}_j^t$ are updated. Finally, the step size updates $t \to t+1$. The optimized power control $\mathbf{p}^*$, transmit beamforming $\mathbf{F}^*$ and system performances $f_0^*\left(\mathbf{x}^*\right)$ are obtained.

Finally, two algorithm analyses are discussed, in terms of space and time computational complexity. The calculation for Algorithm 1 entails T_{max} iterations. Besides, the size of mini-batch set is S. In each iteration, the mini-batch is traversed. Therefore, the time complexity analysis is $O(S \times T_{max})$. It can be seen that the proposed Algorithm 1 takes a longer time to

converge. The space complexity depends on the the number of base station's antennas M and the number of users K. The channel state $\mathbf{h}$ is complex number so the input of DNNs is $2MK$. If the interference management in the real-number power control, the output of DNNs is MK. If the interference management in the complex-number transmit beamforming, the output of DNNs is $2MK$. The number of hidden layers is fixed and the number of neurons in each layer is cK. c is a constant. Additionally, the scale of DNN is related to the number of user K [34]. In short, the space complexity of Algorithm 1 is $O((M+c)\times K)$.

Algorithm 2 Distributed Unsupervised Learning Interference Management in ISAC Systems

Input Initial transmit power $\mathbf{p}^0$, policy parameters $\boldsymbol{\omega}^0$, Lagrange multipliers $\boldsymbol{\lambda}^0, \boldsymbol{\mu}^0$.
Iteration begins
For $t = 1, 2, ..., T$ **do**:
  For base stations $j = 1, 2, ..., J$ **do**:
  a) Quantize the base station j's channel state $\mathbf{h}_j$ through (28);
  b) Concate the local CSI $\mathbf{h}_j$ and the neighbor association information $\mathbf{v}_{j,i}$;
    Iterative Unsupervised Learning Optimization
    c) Sample a mini-batch set $\mathcal{S} \subset \mathcal{H}$;
    d) The gradients $\nabla_{\boldsymbol{\omega}} f\left(\mathbf{h}, \boldsymbol{\pi}\left(\mathbf{h}; \boldsymbol{\omega}\right)\right)$ and $\nabla_{\boldsymbol{\omega}} g_k\left(\mathbf{h}, \boldsymbol{\pi}\left(\mathbf{h}; \boldsymbol{\omega}\right)\right)$ are calculated by (22);
    f) Update parameters of DNNs $\boldsymbol{\omega}_j^t$ through (20);
    g) Update dual variables $\boldsymbol{\lambda}^t$ through (21);
    h) Save the trained model for power control;
    Transfer Learning for Transmit Beamforming
    i) Load the trained model and modify the output layer through (26);
    j) Update parameters of DNN's last layer $\boldsymbol{\phi}_j^t$;
    k) The step size updates $t \rightarrow t+1$.
  Execute step a) to k) until convergence or reach the number of iteration T;
  End for
End for
Iteration ends
Output Power control $\mathbf{p}^*$, transmit beamforming $\mathbf{F}^*$, and system performances $f_0^*\left(\mathbf{x}^*\right)$.

Algorithm 2 is the distributed version of the Algorithm 1. The number of the base station is J. Thus, the calculation for Algorithm 2 entails $O(J \times S \times T)$ iterations. To distinguish the Algorithm 1, the number of iterations of Algorithm 2 is set as T. In terms of spatial complexity in Algorithm 2, the main difference lies in the quantizer added in the base station. The scale of optimizer is related on the number of users K and the length of dimensional square box L_j. Therefore, the space complexity of Algorithm 2 is $O((L_j + M + 2c) \times K \times J)$. The algorithm analysis is summarized in Table 1.

TABLE I

SUMMARY OF ALGORITHM ANALYSIS.

Algorithms	**Space Complexity**	**Time Complexity**
Algorithm 1	$O((M + c) \times K)$	$O(S \times T_{max})$
Algorithm 2	$O((L + M + 2c) \times K \times J)$	$O(J \times S \times T)$

V. SIMULATION RESULTS AND DISCUSSIONS

Simulation results are presented to verify the communication and sensing performances, including spectral efficiency, MSE, and beam pattern. We deploy Python v3.9.0 to conduct the simulations and compile them with PyTorch v1.11.0. In the simulation results, the learning rate of each model is set as 1×10^{-4} and the optimizer is Adam. The mini-batch gradient descent method is utilized and the batch size is 32. The channel state in the ISAC systems is generated from Rayleigh distribution, which is generated by the complex Gaussian distribution. Concerning the parameters $\boldsymbol{\omega} := \{\mathbf{w}; \mathbf{b}\}$ of DNNs, the initial weight matrices $\mathbf{w}$ are generated from zero-mean truncated Gaussian distribution, while all DNNs of the initial bias vectors $\mathbf{b}$ are fixed to 0.01. The step size of dual variables η equals the learning rate of DNNs. The structure of DNNs consists of three layers. The first layer is $2K$. The second layer is $3K$. The third layer is $2K$. The number of users is set as 5 and the number of targets is set as 3 in the simulation results. There are 12 radar antennas and 12 communication antennas, respectively.

The unsupervised learning simulation results are presented and discussed. Based on this discussions, the performance of distributed unsupervised learning method are evaluated further.

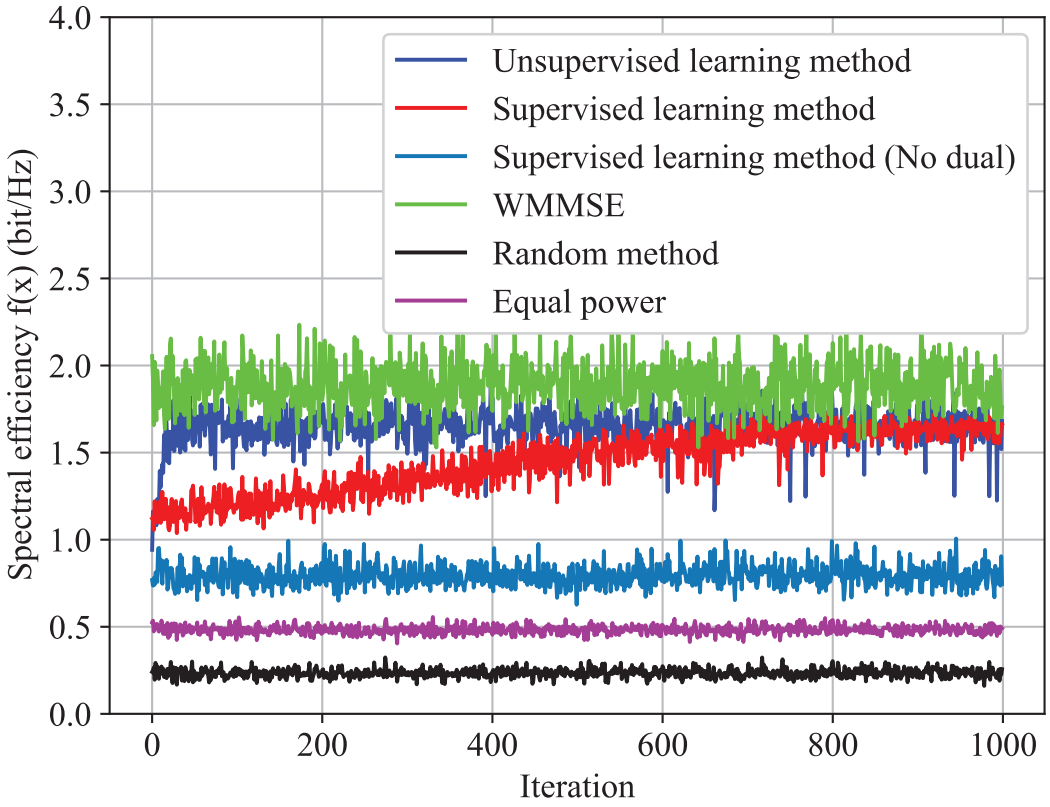

Fig. 4. The spectral efficiency with respect to the power control.

Fig. 4 presents the convergence of different algorithms in terms of system performance function $f(x)$. Besides the proposed unsupervised learning method, the supervised learning method, WMMSE, random method, and equal power are listed. Concerning other five methods, their introductions are listed as follows:

- *Unsupervised learning method (No dual):* The cost function is consistent with the unsupervised learning method. However, the dual variables are not added to optimize the primal problem.
- *WMMSE:* A locally optimal solution is calculated by the conventional iterated optimization [5].
- *Supervised learning method:* The DNN is trained to use supervised learning with the training set, which is provided by the WMMSE algorithm.

- *Equal power:* The transmitters deploy the equal transmit power $p_k = 0.1$.
- *Random method:* The transmit power p_k follows the uniform distribution $[0, 0.1]$.

Obviously, the proposed unsupervised method has excellent performance and convergence. The unsupervised learning method and the supervised learning method obtain the same performance as the WMMSE. It can be seen that the unsupervised learning method is superior to the supervised learning method. These simulation results agree with the analytical results. Without the developed training set, the proposed unsupervised learning method can reach the equivalent performance of the supervised learning method. The conventional methods such as WMMSE, random method, and equal power do not have learning tendency with the iteration increases.

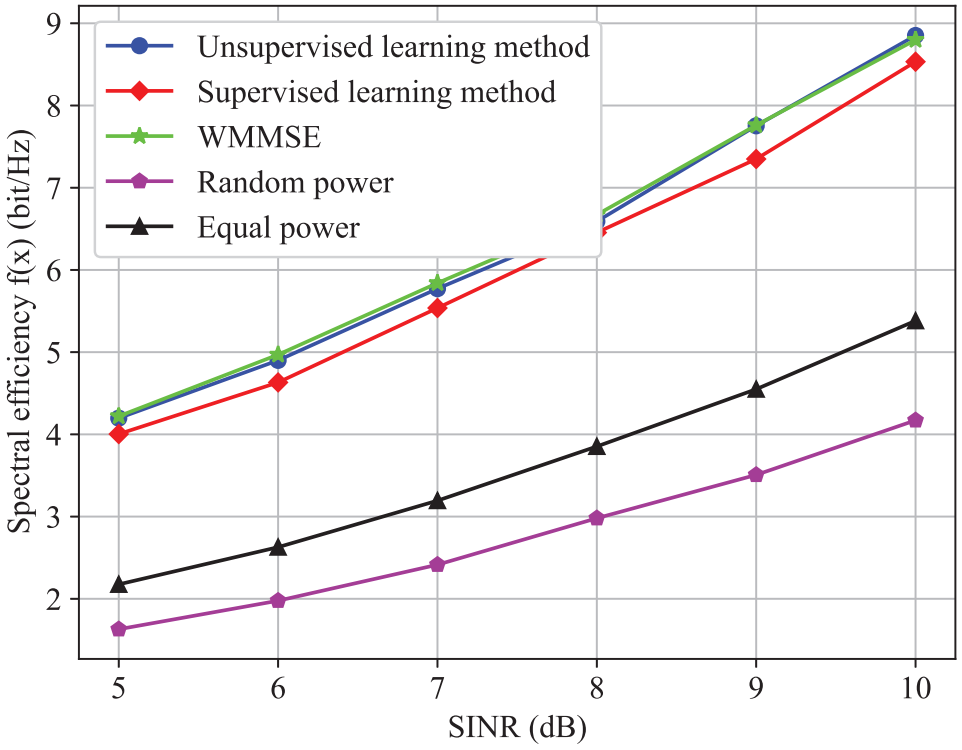

Fig. 5. The spectral efficiency vs different SINRs.

The system performance function $f(x)$ with respect to different SINRs is presented in Fig. 5. The gain of SINR increases from 5 dB to 10 dB. On the basis of Fig. 3, the methods except unsupervised learning method (without dual) are listed to compare the performance. Note that

with the increase of SINR, the spectral efficiencies among five methods increase simultaneously. And the proposed unsupervised learning method has better performance function $f(x)$ than the supervised learning method. Compared with the supervised learning method, there exists 0.5 bit/Hz upgrade in the unsupervised learning method.

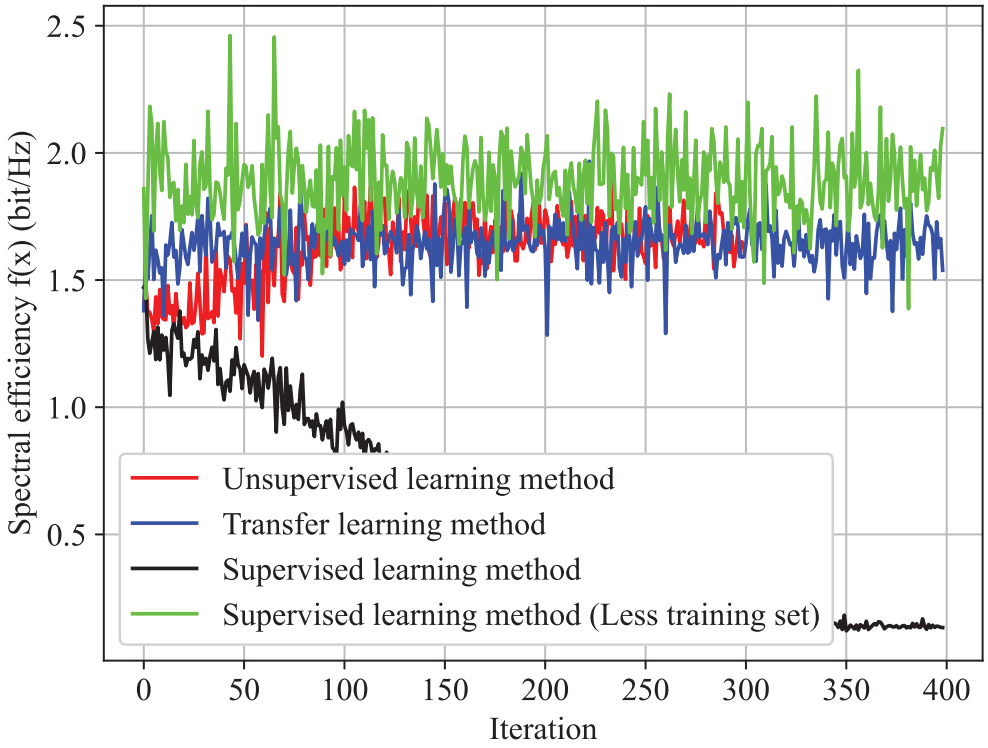

Fig. 6. The spectral efficiency with respect to the beamforming.

To evaluate the performance of interference management discussed in Section III. C. The transfer learning method is compared to the mentioned direct unsupervised learning method and supervised learning method. To utilize the developed model derived from the unsupervised learning method sufficiently, the model is transformed from the original power control into beamforming. The output also extends from the real region into complex region through modifying some layers of former DNNs. Fig. 6 illustrates the advantage of transfer learning method discussed in Section III. C. The policy of the unsupervised learning method and supervised learning method also change into interference management in terms of the beamforming. To

observe directly, the learning rate η changes from 1×10^{-4} to 1×10^{-3}. The unsupervised learning method will converge after 20 iterations and the supervised learning method can not obtain a better learning effect under the same circumstance. Compared with the direct unsupervised learning and supervised learning, the transfer learning does not cost too much train time to manage interference in terms of beamforming. This method can obtain the same convergence performance as the direct unsupervised learning method.

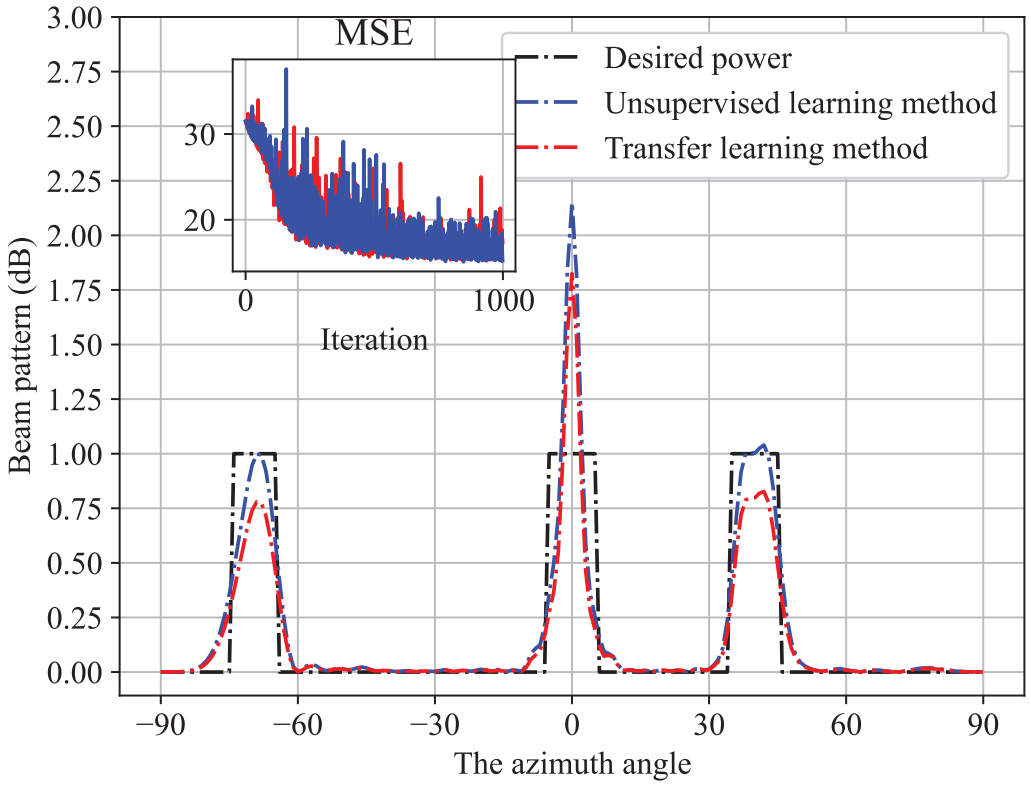

Fig. 7. Transmit beam pattern and its MSE in the ISAC systems.

Fig. 7 depicts the beam patterns in the proposed unsupervised learning method and transfer learning method. The desired detection angle is set as $\{-70°, 0°, 40°\}$. The purpose of this asymmetric settings is to evaluate the training effects of the unsupervised learning method and transfer learning method. Notably, both of beam patterns of these two proposed methods are in accordance with the desired power approximately. The subgraph also is added in Fig. 7, which depicts the MSEs between the desired power and the beam pattern derived from the designed

two methods. It can be concluded that the loss function can converge to the constraints and the good sensing performances can be obtained. The MSE in transfer learning method is more stable than in unsupervised learning method. This is because the developed DNNs are utilized efficiently.

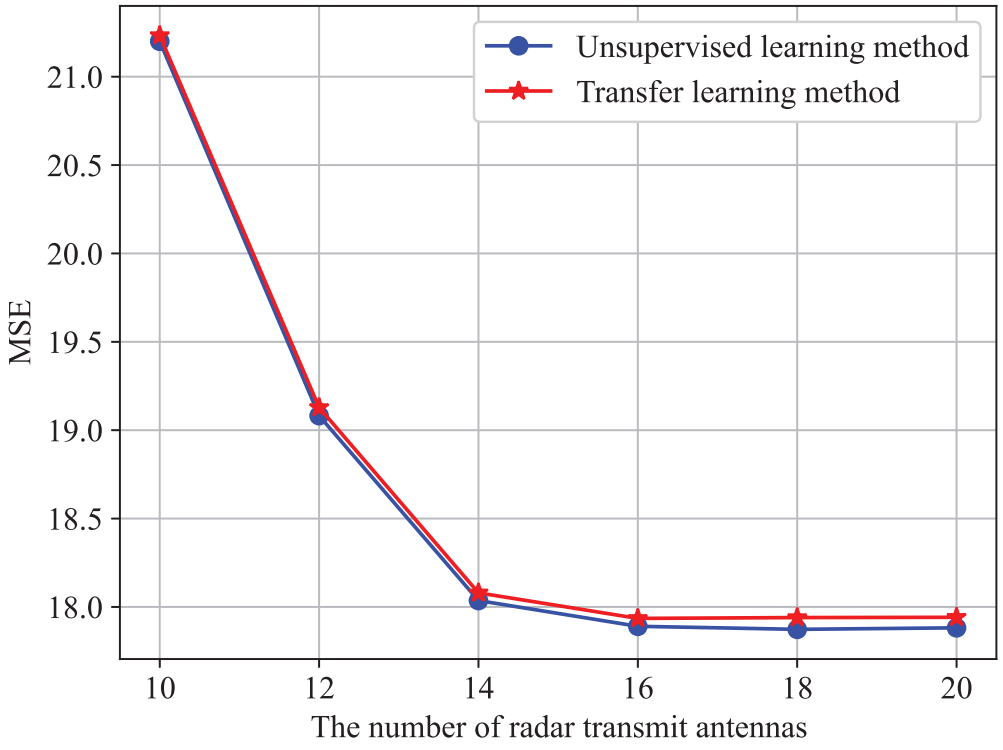

Fig. 8. The comparison between unsupervised learning method and transfer learning method vs different numbers of transmit antennas.

To verify the sensing performances of these two proposed methods, the MSEs are compared with the increasing of the number of the radar transmit antennas M_r. The number of radar antennas ranges from 10 to 20. As shown in Fig. 8, both of MSEs of these two learning methods decrease. And it is obvious that the MSE of transfer learning method can achieve the performance of the unsupervised learning method approximately. It is the consensus of the analytical result discussed before. Although the gap between these two MSEs increases with the large number of transmit radar antennas, the slight gap can be tolerable.

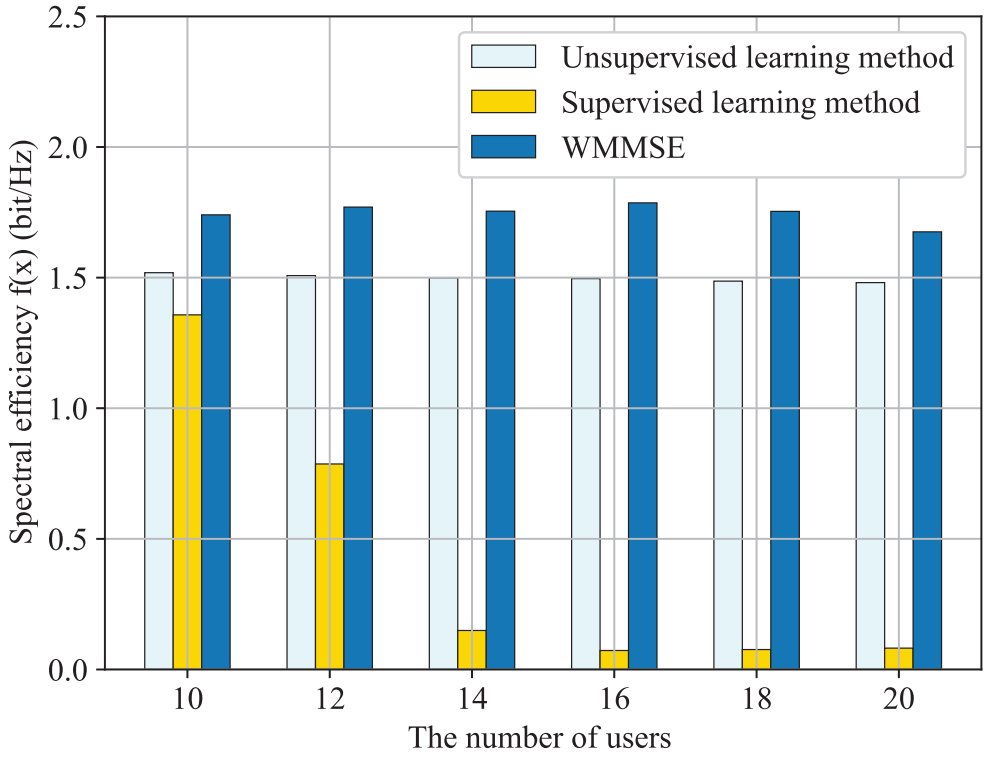

Fig. 9. Spectral efficiency vs different numbers of users.

Fig. 9 illustrates that the proposed unsupervised learning method which is superior to the existing supervised learning method with the increasing number of users. The DNNs' settings are identical to the initial settings. And the supervised DNNs' structures are the same as the unsupervised learning method. It is obvious that with the increasing number of users, the conventional optimized WMMSE method has stable performance. The proposed unsupervised learning method also maintains the well performance. However, the supervised learning method has poor spectral efficiency as the number of users increases. This is because the bigger size of input data results in higher training complexity. The proposed unsupervised learning method only has a slight decrease and it is superior to the training effect of the supervised learning method.

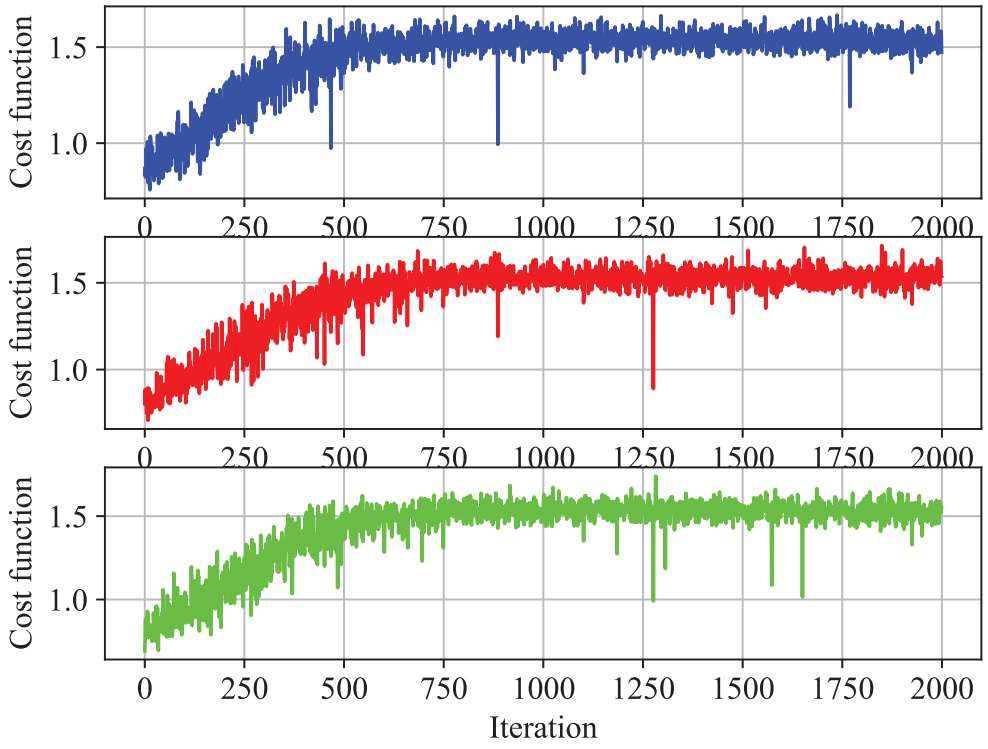

Fig. 10. The different learning performances of different ISAC base stations.

Fig. 10 shows three base stations' cost functions (i.e. spectral efficiency in this paper). The distributed interference management is considered among three base stations with local CSI. The iteration steps are set as 2000 and the learning rate of each base station's DNNs is set as 1×10^{-4}, uniformly. It can be seen that each base station can utilize its local CSI to obtain interference management. With the aid of the combination between optimizer and quantizer, the spectral efficiency can converge after 1000 iterations. Obviously, the proposed distributed interference management can obtain stable performance through enough training iterations.

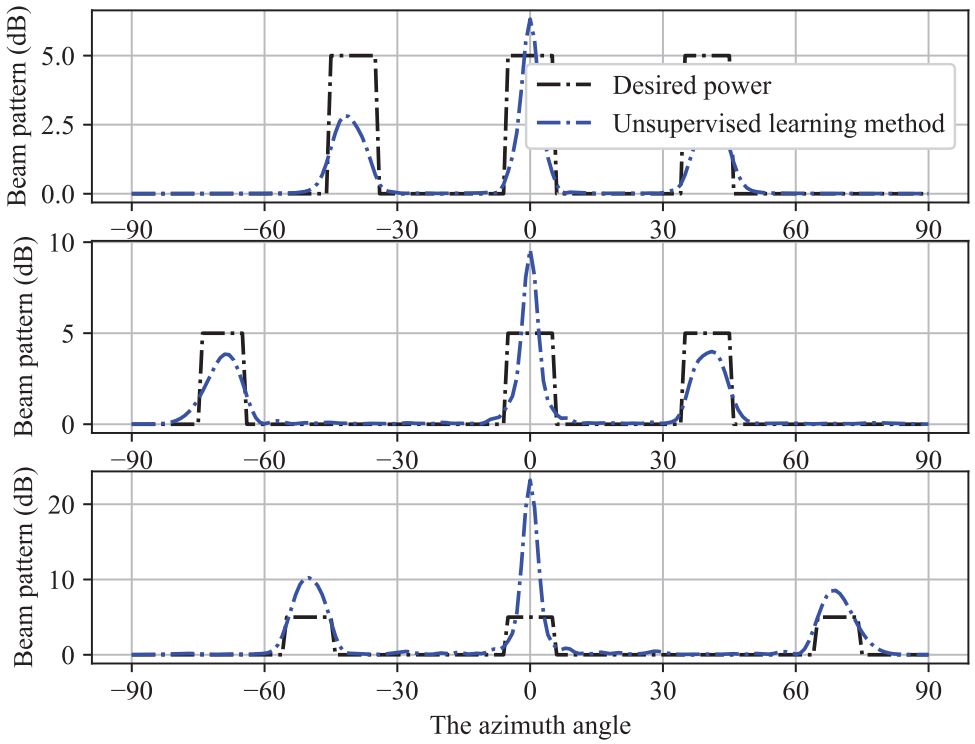

Fig. 11. The different beam patterns of different ISAC base stations.

Fig. 11 depicts the sensing performances of different ISAC base stations. Similar to the settings of Fig. 10, three base stations need to deal with different targets. Base station 1 require to detect the azimuth angle of $\{-40°, 0°, 40°\}$. Base station 2 is required to detect the azimuth angle of $\{-70°, 0°, 40°\}$ and base station 3 is required to detect the azimuth angle of $\{-50°, 0°, 70°\}$. The different azimuth angles settings ensure the flexibility of the proposed method. It can be seen that the unsupervised learning method can obtain stable performance in terms of distributed mode. It is similar to the centralized scenario, which is consistent with the analytical result.

VI. Conclusion

A distributed interference management is provided in the ISAC systems, considering power control and transmit beamforming. An ISAC system cooperating with edge intelligence is proposed to coordinate interference with a unsupervised learning method. The problem is transformed into a functional optimization with stochastic constraints. A unsupervised learning method is proposed to allocate power for interference management. Furthermore, a transfer learning technique is introduced to deal with interference management in terms of transmit beamforming. Then the distributed management is obtained by the local CSI in the multi-user MISO scenario. Each base station is equipped with the individual DNNs. Simulation results demonstrate the effectiveness of the designed methods, in the aspect of distributed interference management in the ISAC systems.

REFERENCES

[1] Y. Cui, F. Liu, X. Jing, and J. Mu, "Integrating sensing and communications for ubiquitous IoT: Applications, trends, and challenges," *IEEE Network*, vol. 35, no. 5, pp. 158–167, Sep. 2021.

[2] A. Zhang, Md. L. Rahman, X. Huang, Y. J. Guo, S. Chen, and R. W. Heath, "Perceptive mobile networks: Cellular networks with radio vision via joint communication and radar sensing," *IEEE Veh. Technol. Mag.*, vol. 16, no. 2, pp. 20–30, Jun. 2021.

[3] F. Liu et al., "Integrated sensing and communications: Towards dual-functional wireless networks for 6G and beyond," arXiv: 2108.07165, Aug. 2021. [Online]. Available: http://arxiv.org/abs/2108.07165.

[4] H. Hua, J. Xu, and T. X. Han, "Optimal transmit beamforming for integrated sensing and communication," arXiv:2104.11871, Dec. 2021. [Online]. Available: http://arxiv.org/abs/2104.11871.

[5] 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 Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, Sep. 2011.

[6] F. Zhu, F. Gao, T. Zhang, K. Sun, and M. Yao, "Physical-layer security for full duplex communications with self-interference mitigation," *IEEE Trans. Wireless Commun.* vol. 15, no. 1, pp. 329–340, 2016.

[7] J. Papandriopoulos and J. S. Evans, "Scale: A low-complexity distributed protocol for spectrum balancing in multiuser dsl networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3711–3724, 2009.

[8] F. Liu, C. Masouros, A. Li, T. Ratnarajah, and J. Zhou, "MIMO Radar and cellular coexistence: A power-efficient approach enabled by interference exploitation," *IEEE Trans. Signal Process.*, vol. 66, no. 14, pp. 3681–3695, Jul. 2018.

[9] C. Baquero Barneto et al., "Full-duplex OFDM radar with LTE and 5G NR waveforms: Challenges, solutions, and measurements," *IEEE Trans. Microwave Theory Tech.*, vol. 67, no. 10, pp. 4042–4054, 2019.

[10] 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 Trans. Signal Process.*, vol. 68, pp. 3929–3944, 2020.

[11] Z. Wang, Y. Liu, X. Mu, and Z. Ding, "NOMA inspired interference cancellation for integrated sensing and communication," arXiv: 2112.09207, Dec. 2021. [Online]. Available: http://arxiv.org/abs/2112.09207.

[12] A. Ahmed, Y. D. Zhang, and B. Himed, "Distributed dual-function radar-communication MIMO system with optimized resource allocation," *2019 IEEE Radar Conference (RadarConf)*, 2019, pp. 1–5.

[13] L. Wu, K. V. Mishra, M. R. B. Shankar, and B. Ottersten, "Resource allocation in heterogeneously-distributed joint radar-communications under asynchronous bayesian tracking framework," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 7, pp. 2026–2042, Jul. 2022.

[14] K. Lee, J.-R. Lee, and H.-H. Choi, "Learning-based joint optimization of transmit power and harvesting time in wireless-powered networks with co-channel interference," *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 3500–3504, Mar. 2020.

[15] D. S. Kalogerias, M. Eisen, G. J. Pappas, and A. Ribeiro, "Model-free learning of optimal ergodic policies in wireless systems," *IEEE Trans. Signal Process.*, vol. 68, pp. 6272–6286, 2020.

[16] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, and Q. He, "A comprehensive survey on transfer learning," *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43–76, 2020.

[17] Q. Peng, A. Gilman, N. Vasconcelos, P. C. Cosman, and L. B. Milstein, "Robust deep sensing through transfer learning in cognitive radio," *IEEE Wireless Commun. Lett.*, vol. 9, no. 1, pp. 38–41, 2019.

[18] A. O. Al-Abbasi, A. Ghosh, and V. Aggarwal, "DeepPool: Distributed model-free algorithm for ride-sharing using deep reinforcement learning," *IEEE Trans. Intell. Transport. Syst.*, vol. 20, no. 12, pp. 4714–4727, Dec. 2019.

[19] M. Haliem, V. Aggarwal, and B. Bhargava, "AdaPool: A diurnal-adaptive fleet management framework using model-free deep reinforcement learning and change point detection," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 3, pp. 2471–2481, Mar. 2022.

[20] B. Li and A. P. Petropulu, "Joint transmit designs for coexistence of MIMO wireless communications and sparse sensing radars in clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 6, pp. 2846–2864, Dec. 2017.

[21] F. Liu, C. Masouros, A. Li, and T. Ratnarajah, "Robust MIMO beamforming for cellular and radar coexistence," *IEEE Wireless Commun. Lett.*, vol. 6, no. 3, pp. 374–377, Jun. 2017.

[22] T. Zhang, S. Wang, G. Li, F. Liu, G. Zhu, and R. Wang, "Accelerating edge intelligence via integrated sensing and communication," *2022 IEEE International Conference on Communications (ICC)*, 2022, pp. 1586–1592.

[23] P. Liu, G. Zhu, W. Jiang, W. Luo, J. Xu, and S. Cui, "Vertical federated edge learning with distributed integrated sensing and communication," *IEEE Commun. Lett.*, vol. 26, no. 9, pp. 2091–2095, Sept. 2022.

[24] D. Wen et al., "Task-Oriented sensing, computation, and communication integration for multi-device edge AI." arXiv: 2207.00969. Accessed: Aug. 2022. [Online]. Available: http://arxiv.org/abs/2207.00969.

[25] P. Stoica, J. Li, and Y. Xie, "On probing signal design for MIMO radar," *IEEE Trans. Signal Process.*, vol. 55, no. 8, pp. 4151–4161, Aug. 2007.

[26] F. Liu and C. Masouros, "A tutorial on joint radar and communication transmission for vehicular networks—part I: Background and fundamentals," *IEEE Commun. Lett.*, vol. 25, no. 2, pp. 322–326, Feb. 2021.

[27] A. Ribeiro, "Ergodic stochastic optimization algorithms for wireless communication and networking," *IEEE Trans. Signal Process.*, vol. 58, no. 12, pp. 6369–6386, Dec. 2010.

[28] M. Liu, T. Song, J. Hu, J. Yang, and G. Gui, "Deep learning-inspired message passing algorithm for efficient resource allocation in cognitive radio networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 641–653, Jan. 2019.

[29] 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 Trans. Signal Process.*, vol. 66, no. 20, pp. 5438–5453, Oct. 2018.

[30] M. Varasteh, J. Hoydis, and B. Clerckx, "Learning to communicate and energize: Modulation, coding, and multiple access designs for wireless information-power transmission," *IEEE Trans. Commun.*, vol. 68, no. 11, pp. 6822–6839, Nov. 2020.

[31] K. G. Kim, "Deep learning," *Healthc Inform Res*, vol. 22, no. 4, p. 351, 2016.

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

[33] H. Lee, S. H. Lee, and T. Q. S. Quek, "Deep learning for distributed optimization: Applications to wireless resource management," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 10, pp. 2251-–2266, Oct. 2019.

[34] M. Eisen, C. Zhang, L. F. O. Chamon, D. D. Lee, and A. Ribeiro, "Learning optimal resource allocations in wireless systems," *IEEE Trans. Signal Process.*, vol. 67, no. 10, pp. 2775–2790, May 2019.

[35] H. Han, H. Liu, C. Yang, and J. Qiao, "Transfer learning algorithm with knowledge division level," *IEEE Trans. Neural Networks Learn. Syst.*, pp. 1–15, 2022.