ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) | 978-1-7281-6327-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICASSP49357.2023.10097184

JOINT ANTENNA SELECTION AND BEAMFORMING IN INTEGRATED AUTOMOTIVE RADAR SENSING-COMMUNICATIONS WITH QUANTIZED DOUBLE PHASE SHIFTERS

Lifan Xu[†], *Shunqiao Sun*[†], *Yimin D. Zhang*[‡] *and Athina Petropulu*[§]

[†]Department of Electrical and Computer Engineering, The University of Alabama, Tuscaloosa, AL, USA

[‡]Department of Electrical and Computer Engineering, Temple University, Philadelphia, PA, USA

[§]Department of Electrical and Computer Engineering, Rutgers University, Piscataway, NJ, USA

ABSTRACT

We consider an integrated sensing-communication system operating in a dynamic environment, such as an autonomous vehicle scenario. We propose a novel, low-cost, low power consumption and low-computation approach for designing a beam that can simultaneously reach the radar target of interest and the desired communication destination. The transmitter is a uniform linear array, equipped with quantized double phase shifters, which enables a flexible beam design while using analog only processing. Only a small number of antennas are selected to transmit in each channel use, in order to save system power and reduce antenna coupling. We propose a deep reinforcement learning approach to adaptively adjust the double phase shifters and select the active antennas in order to optimize the transmit beamforming, through a transmission and feedback trail. The actor-critic network strategy together with the Wolpertinger policy is adopted to obtain the optimal solutions efficiently and effectively. Numerical results demonstrate the feasibility of the proposed method.

Index Terms— Deep reinforcement learning, integrated sensing and communication (ISAC), automotive radar, sparse array, adaptive beamforming

I. INTRODUCTION

Millimeter wave (mmWave) radar can cope well with various weather and lighting conditions, and achieve reliable target perception at a lower cost than LiDAR. As a result, they are viewed as a key enabling technology to support autonomous driving [1]–[3]. In addition to target sensing, autonomous vehicles are also required to communicate with road infrastructure and other nearby vehicles for operational coordination, especially in vehicle platooning [4]–[10]. Traditionally, automotive radar sensing and inter-vehicle communication functions have been implemented through separate hardware. More recently, integrated sensing and communication (ISAC) systems [11]–[14] combine both functions on a single hardware platform. Dual-function radar-communication (DFRC) systems, a special class of ISAC systems, use the same waveform as well as the same hardware platform for both sensing and communication. DFRC systems are already been used in autonomous vehicles [15], [16].

A variety of different ways are used to design ISAC systems [5], [17]–[21]. When designing the transmit array of the ISAC system, sparse arrays are more attractive than uniform linear arrays, because of reduced hardware complexity and lower power consumption. The structure of the sparse array can greatly affect the transmit beampatterns, including the sidelobe levels. Excessive sidelobe levels can result in reduced radar perception and communication capabilities. In order to avoid interference to other radars, it is also desirable that the transmit array can form a null in the direction of the users of no interest. [22], [23] utilize optimization and greedy search methodologies to design the sparse array, or equivalently, select the active antennas from a uniform linear array. [19] resorts to artificial intelligence technique to solve antenna selection problem while avoiding large latency.

This work has been funded in part by U.S. National Science Foundation (NSF) under Grants CCF-2153386 and ECCS-2033433.

In low-cost transmitters, phase shifters are designed to take discrete values out of a set. In this case, the available transmit signals cannot be arbitrarily designed but are rather limited by the degrees of freedom of the phase shifter. For example, the Texas Instruments (TI) imaging radar cascaded by four AWR2243 radar chipsets supports 6-bit phase shifters [24]. Another approach to synthesize more complex transmit beams is via the use of double-phase shifters [25], [26]. Joint design of the phase shifter phases and selection of the active antennas is a difficult problem, which becomes more challenging as the number of the bits of phase shifters and the number of antenna elements increase. For those problems, one could use value-based reinforcement learning, such as deep Q-networks (DQNs) [27], [28], which has great generalization performance and high training efficiency. However, the associated computation complexity is extremely high due to the curse of dimensionality.

In this paper, we propose a Wolpertinger policy-assisted reinforcement learning framework for optimization of antenna selection and quantized transmit beamforming vectors in an ISAC system for automotive radar. The primary goal is to concentrate the output power on the desired sensing and communication directions in a dynamic environment, while minimizing the interference to other automotive radars. In autonomous driving scenario, the radar target of interest, communication user and other automotive radars are highly dynamic rather than stationary. As a result, the automotive radar ISAC system is required to adjust its transmit parameters accordingly. With DRL, an efficient antenna selection, beamforming, and waveform constraints can be achieved in rapidly changing automotive scenarios. The introduction of double phase shifters ensures simultaneous transmission of radar perception and communication signals. Actor-critic networks are used to solve the curse of dimensionality caused by the large action space of deep reinforcement learning (DRL) techniques. The proposed ISAC automotive radar system not only improves the signal-to-noise ratio (SNR) of radar perception and communication, but also attenuates interference to other automotive radars on the road.

II. SYSTEM MODEL

We consider a multiple-input multiple-output (MIMO) frequency-modulated continuous-wave (FMCW) automotive radar ISAC system consisting of an N_t-element uniform linear transmit array and an N_r-element receive array. The transmit array is used for both the radar sensing and communication functions. The system block diagram of the integrated automotive radar sensing and communication system is shown in Fig. 1. In the radar sensing and communication mode, a single radio frequency (RF) chain is driving a phase shifter bank to adjust its transmission phase. Therefore, the analog precoder $\mathbf{f}_{\mathrm{RF}}$ has dimension $\mathbb{C}^{N_t \times 1}$. Each shifter bank has a pair of q bits of quantized phase shifters and a finite number of tunable phase values that are uniformly drawn from a quantized set $\mathcal{D}$ with 2^q possible phase values. The data streams are encoded in the slow-time domain using code symbol

Ω and all transmit antennas carry the same communication symbol at the same chirp.

At the n-th chirp, the transmit waveform vector for the N_t transmit antennas is

$$\mathbf{x}(n,t) = e^{j2\pi\left(f_c t + \frac{B}{2T}t^2\right)} e^{j2\pi\Omega_n} \mathbf{f}_{\mathrm{RF}}, \tag{1}$$

where f_c is the carrier frequency, B denotes the bandwidth, T is the pulse duration, and Ω_n is the n-th data symbol.

II-A. Antenna Selection

In order to reduce the power consumption of the radar system and the mutual coupling of the antennas, we design a reconfigurable sparse array by activating only a subset of available transmit antennas. Define a element selection matrix $\mathbf{S} = [\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_{N_t}]$ where the column vector $\mathbf{u}_i$ represents the state of the i-th antenna and its i-th entry is "1" if it is already activated, otherwise it should be "0". The aperture size of the transmit array determines the width of the transmit beam, so a larger aperture implies a narrower beamwidth. We fix the two antennas at both ends to maintain the same array aperture, and choose M other antennas in-between, rendering a total of $M+2$ antennas, so that the trace of $\mathbf{S}$, i.e., $\mathrm{tr}(\mathbf{S})$, is $M+2$. The first element in $\mathbf{u}_1$ and the last element in $\mathbf{u}_{N_t}$ are "1", and all other entries in $\mathbf{u}_1$ and $\mathbf{u}_{N_t}$ are "0".

II-B. Radar Transmit Beamforming

The radar transmit beampattern is given by

$$B(\theta) = \mathbf{a}_t^H(\theta)\mathbf{S}^H\mathbf{W}\mathbf{S}\mathbf{a}_t(\theta), \tag{2}$$

where the transmitter steering vector $\mathbf{a}_t(\theta)$ is

$$\mathbf{a}_t(\theta) = \left[1, e^{j\frac{2\pi}{\lambda}d\sin\theta}, \cdots, e^{j\frac{2\pi}{\lambda}(N_t-1)d\sin\theta}\right]^T,$$

with d denoting the interelement spacing of the transmit array, and $\mathbf{W} \in \mathbb{C}^{N_t \times N_t}$ is the beamforming weight matrix, expressed as

$$\mathbf{W} = \mathbb{E}\left[\mathbf{x}(n,t)\mathbf{x}^H(n,t)\right] = \mathbf{f}_{\mathrm{RF}}\mathbf{f}_{\mathrm{RF}}^H. \tag{3}$$

To perform the radar sensing function, the analog precoder $\mathbf{f}_{\mathrm{RF}}$ is designed to steer the mainlobe to the region of interest. $\mathbf{f}_{\mathrm{RF}}$ is controlled by the phase shifters, which can be replaced by the radar sensing beamformer $\mathbf{w}_r$, defined as

$$\mathbf{w}_r = \frac{1}{\sqrt{N_t}}\left[e^{j\phi_1}, e^{j\phi_2}, \cdots, e^{j\phi_{N_t}}\right]^T, \tag{4}$$

where $\phi_i \in \mathcal{D}$ for all $i \in \{1, \cdots, N_t\}$.

II-C. Communication Model

Assume that the communication receiver is equipped with an N_c-element array. The mmWave channels are considered to have limited scattering [29], i.e., the number of independent propagation paths L is smaller than N_t. Then, the downlink channel matrix $\mathbf{H} \in \mathbb{C}^{N_c \times N_t}$ is given by

$$\mathbf{H} = \sqrt{\frac{N_t N_c}{L}} \sum_{l=1}^{L} \beta_l \mathbf{b}_c(\theta_{cl})\mathbf{a}_t^H(\theta_{tl})\mathbf{S}^H, \tag{5}$$

where β_l is the l-th complex path gain. Here, $\mathbf{b}_c(\theta_{cl})$ and $\mathbf{a}_t(\theta_{tl})$ are communication system receive and transmit array steering vectors, respectively, with θ_{cl} and θ_{tl} being the angles of arrival and departure of the l-th path. The received signal at the communication receiver is

$$\mathbf{y}_c(n,t) = \sqrt{\rho} e^{j2\pi\left[f_c(t-\tau_c)+\frac{1B}{2T}(t-\tau_c)^2\right]} e^{j2\pi\Omega_n}\mathbf{H}\mathbf{f}_{\mathrm{RF}} + \mathbf{n}(n,t), \tag{6}$$

where ρ denotes the average received power and τ_c is the delay between the radar transmitter and the communication receiver. In the communication mode, the analog beamformer $\mathbf{f}_{\mathrm{RF}}$ is replaced with a beamformer $\mathbf{w}_c$, defined as

$$\mathbf{w}_c = \frac{1}{\sqrt{N_t}}\left[e^{j\Phi_1}, e^{j\Phi_2}, \cdots, e^{j\Phi_{N_t}}\right]^T, \tag{7}$$

where $\Phi_i \in \mathcal{D}$ for all $i \in \{1, \cdots, N_t\}$.

II-D. Beam Synthesis via Double Phase Shifters

The ISAC system shares the same transmit array. Therefore, in order to realize the two functions at the same time, the radar sensing beamformer $\mathbf{w}_r$ and the communication beamformer $\mathbf{w}_c$ need to be fused into a new beamformer $\mathbf{w}$. This beamformer can synthesize two main beams that focus the energy of the array in the directions of interest for each of the two functions.

The transmit array forms a single beam. By connecting a pair of phase shifters to each antenna we introduce more degrees of freedom to shape the beampattern [25]. The hybrid transmit beamformer $\mathbf{w}$ is given as

$$\begin{aligned}\mathbf{w} &= c_1\mathbf{w}_r + c_2\mathbf{w}_c \\ &= \frac{1}{\sqrt{N_t}}[c_1 e^{j\phi_1} + c_2 e^{j\Phi_1}, c_1 e^{j\phi_2} + c_2 e^{j\Phi_2}, \cdots, \\ &\quad c_1 e^{j\phi_{N_t}} + c_2 e^{j\Phi_{N_t}}]^T.\end{aligned} \tag{8}$$

where $c_1 \in [0,1]$ and $c_2 \in [0,1]$ are weighting factors that balance radar sensing and communication capabilities, respectively.

III. TRANSMIT BEAMFORMING DESIGN

In this section, we consider the design of a transmit beamformer $\mathbf{w}$ and an antenna selection matrix $\mathbf{S}$ that maintain certain power at the radar targets and communication destination, and at the same time generates low interference towards other directions and low peak sidelobe levels. The problem can be formulated as

$$\begin{aligned}\min_{\{\mathbf{w},\mathbf{S},\alpha_1,\alpha_2,\alpha_3\}} \quad & \alpha_1 + \alpha_2 + \alpha_3 \\ \text{s.t.} \quad & \left|\mathbf{w}^H\mathbf{S}\mathbf{a}(\theta_r)\right| = c_1^2, \\ & \left|\mathbf{w}^H\mathbf{S}\mathbf{a}(\theta_c)\right| = c_2^2, \\ & \left|\mathbf{w}^H\mathbf{S}\mathbf{a}(\theta_l)\right| \le \rho_1 + \alpha_1,\ \theta_l \in \Theta,\ l = 1,2,\cdots,L, \\ & \left|\mathbf{w}^H\mathbf{S}\mathbf{a}(\theta_i)\right| \le \rho_2 + \alpha_2, \\ & \mathbf{w} = c_1\mathbf{w}_r + c_2\mathbf{w}_c, \\ & \left|\theta_r - \hat{\theta_r}\right| \le \alpha_3, \\ & \mathrm{tr}(\mathbf{S}) = M+2,\end{aligned} \tag{9}$$

where $\theta_r, \theta_c, \theta_l, \theta_i$ denote the sensing direction, departure angle to the communication user, discretized angle in the sidelobe region, and direction of a target of no interest, respectively. The power allocated to radar sensing and communication is c_1^2 and c_2^2, respectively. The specific peak sidelobe level constraint is denoted by ρ_1, and the required interference attenuation level is ρ_2. The auxiliary variables α_1 and α_2 are used to relax the specific level constraints in peak sidelobe and interference attenuation, and their minimum values are 0. The parameter α_3 measures the deviation between ground truth θ_r and the actual main beam direction $\hat{\theta}$.

This is a combinatorial NP-hard optimization problem that is difficult to solve, without mentioning that the quantized phase has to be further taken into account. To alleviate the optimization difficulties caused by the size explosion, and considering the limited phase tunability characteristics of practical phase shifters, we adopt a framework based on deep reinforcement learning to dynamically activate or deactivate antennas and tune the phase of the activated antennas.

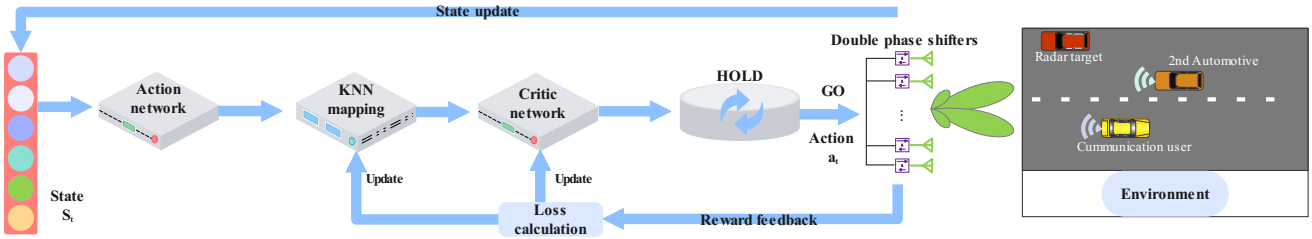


Fig. 1: The proposed automotive radar ISAC system diagram with action-critic network.

III-A. Deep Reinforcement Learning

The problem of sparse array beamforming is to find the optimal subarray set and the corresponding beamforming matrix policy, so reinforcement learning is a powerful tool. Reinforcement learning intelligently updates the mapping policy Γ by tracking the reward r_t and maps state s_t, action a_t, and reward r_t to the action-value function $Q^*(s_t, a_t)$ as $Q^*(s_t, a_t) = \max_\Gamma \mathbb{E}[r_t | s_t = s, a_t = a, \Gamma]$. DQN using neural networks maps above parameters. When the action dimension is large, however, it is difficult to search for the desired mapping policy using DQN reinforcement learning. Therefore, we adapt the Wolpertinger policy-based reinforcement learning framework and implement time-wise tractable training. Three basic elements, i.e., action network, k-nearest neighbor (KNN) map, and critic network, make up the Wolpertinger policy. The action network maps the input state to an output proto-action, and the KNN block is then used to map the proto-action to the feasible space.

All action-state pairs from the KNN mapping by $a_t = \arg \max_{a_t \in g_t(\hat{a}_t)} Q_{\theta^Q}(s_t, a_t)$ will be evaluated by the critic network. The loss function is given as $\text{Loss} = (y - Q_{\theta^Q}(s, a))^2$, where y is the output of the target network that is used to generate independent and identically distributed (i.i.d.) data to stabilize the training process. The deep deterministic policy gradient (DDPG) is used to train the networks [30].

III-B. Beamforming Design with DRL

The beamforming design is implemented using the actor-critic network.

1) **Action Space:** Selection of $M+2$ from N_t antennas with fixed antennas at both ends yields $Q = C^M_{N_t-2}$ possible solutions. Each antenna is connected with two q-bit quantized phase shifters. To obtain $\mathbf{w}$, we can simultaneously optimize the phase of the phase shifter, and the adjustment dimension of the phase is $\mathbb{R}^{Q \cdot 2^{q \times (M-2)}}$.

2) **State:** After taking an action from the action space, the state vector $\mathbf{s}$ is changed and consists of the transmit array phase shifter status. At the i-th iteration, the state $\mathbf{s}_i^T = [w_1, w_2, \cdots, w_{M+2}]_i$. One element phase change (activation or deactivation) implies taking an action from the action space.

3) **Hold and Go:** In the radar search mode, the range, Doppler, and angle of the target are estimated so that the region of interest (ROI) can be defined. In the hold stage, we only explore one set of phase shifters instead of two, and perform a pre-beamforming check before inputting two phase shifters with the desired phase. The two columns of the beamformer recorder matrix $\mathbf{W}_{rc} = [\mathbf{w}_r^T(\theta_r), \mathbf{w}_c^T(\theta_c)]$ form the respective beams in the target and communication receiver directions. The fused beamformer $\mathbf{w} = \mathbf{w}_r^T(\theta_r) + \mathbf{w}_c^T(\theta_c)$. Here, we assume that $c_1 = c_2 = 1$. The set of flags $\mathbf{f}_d$ is used to check whether the phase of double phase shifters should be changed. It consists of two flag bits. A one-time trigger flag bit f_{d1} is used to detect the dimension of matrix $\mathbf{W}_{rc}$ and, once the dimension satisfies the two columns, the holding phase is then ended and the external environment interaction is entered. Another flag f_{d2} is used to indicate whether $\mathbf{W}_{rc}$ changes and, if so, update the reward.

4) **Reward:** Assume that the ROI for radar sensing spans $[-\theta_{\text{ROI}}/2, \theta_{\text{ROI}}/2]$ and the 3-dB beamwidth is given by $\Delta_\theta = 2\arcsin(1.4\lambda/(\pi D))$. The region outside the first nulls of the mainlobe is defined as the sidelobe region. The difference between the mainlobe peak and the PSL is $\xi_i = \max(P_{\text{ROI},i}) - \max(\text{PSL}_i)$ at the i-th update, where $\max(P_{\text{ROI},i})$ and $\max(\text{PSL}_i)$ are the maximum main beam level and PSL at the i-th iteration, respectively. The main beam deviation is given by $d_r = |\theta_r - \hat{\theta}_r|$. The term ξ and d_r are used to guarantee the main beam to be steered to ROI while the PSL is minimized. The reward is given by

$$r_{ri} = \begin{cases} 1, & \text{if } \xi_i > \xi_{i-1} \text{ and } d_{ri} \le d_{ri-1}, \\ -1, & \text{if } \xi_i \le \xi_{i-1} \text{ and } d_{ri} > d_{ri-1}, \\ 0, & \text{other cases.} \end{cases} \tag{10}$$

For communication evaluation, the received gain is expressed as $g_c = |\mathbf{H}\mathbf{w}|^2$. Assuming the channel parameters are estimated, the reward in communication is given by

$$r_{ci} = \begin{cases} 1, & \text{if } g_{ci} > g_{ci-1} \\ 0, & \text{if } g_{ci} = g_{ci-1} \\ -1, & \text{if } g_{ci} < g_{ci-1}. \end{cases} \tag{11}$$

Algorithm 1 DRL-based automotive radar ISAC system

Initialize networks with corresponding parameters.
HOLD = TRUE, $f_{d1} = 0$; $f_{d2} = 0$.
Initialize $\xi_0 = 0$, $d_{r0} = 1$, $g_{c0} = 0$, $p_0 = 1$.
Initial sample a random beamforming vector $\mathbf{w}_{rc1}$ as initial state s_1 and record action a_1.
for $i = 1$ to T **do**
  Receive proto-action $\hat{a}_i$ from actor network.
  Action embedding $g(\hat{a}_i)$ through KNN mapping.
  while HOLD **do**
    Update $\mathbf{W}_{rc1}$.
    f_{d1} = column($\mathbf{W}_{rc}$).
    if f_{d1} == 2 **then**
      HOLD = FALSE.
      Execute action $\mathbf{w}_1$ passed from critic network.
      Calculate reward and update state $s_{i+1} = a_i$.
      Update ξ_1 , d_{r1} , g_{c1} and p_1.
    end if
  end while
  Update $\mathbf{W}_{rci}$.
  if $\mathbf{W}_{rci} \neq \mathbf{W}_{rci-1}$ **then**
    Execute action $\mathbf{w}_i$ passed from critic network.
    Calculate reward and update state $s_{i+1} = a_i$.
    Update ξ_i , d_{ri} , g_{ci} and p_i.
    Update all networks.
  end if
end for

This dynamic gain will be reported to the automotive radar by communication user through an uplink channel.

To avoid interference to the other potential automotive radar user, we desire that the synthesized beamformer has the ability to create a null toward its direction of departure θ_i. The attenuation level is given by $p = |\mathbf{w}^H \mathbf{a}(\theta_i)|$. The reward in interference attenuation at the i-th update is given by

$$r_{pi} = \begin{cases} 1, & \text{if} \quad p_i < p_{i-1} \\ 0, & \text{if} \quad p_i = p_{i-1} \\ -1, & \text{if} \quad p_i > p_{i-1}. \end{cases} \tag{12}$$

The final triple reward r_i at the i-th update can be expressed as

$$r_i = \lambda_1 r_{ri} + \lambda_2 r_{ci} + \lambda_3 r_{pi}, \tag{13}$$

where λ_1, λ_2, and λ_3 represent the respective weights trading off between the radar and communication functions, and interference attenuation. The pseudo code of DRL-based automotive ISAC using Wolpertinger policy is given by Algorithm 1.

IV. NUMERICAL RESULTS

We consider an FMCW radar system with 15 transmit antennas with d being half wavelength, and normalized spatial frequency of the half field of view of the array is set to 0.7, corresponding to $\theta_{\text{ROI}}/2 = 44.2^\circ$. The 3-dB beamwidth is $\Delta_f = 0.119$, corresponding to $\Delta_\theta = 6.81^\circ$. We select ten out of all transmit antennas to form a final transmit array, yielding $Q = 1,287$ possible solutions. Each antenna has two 3-bit quantization phase shifters. The dimension of the total quantized action space exceeds $10^{10} = 10$ billions. The hyper-parameters for model training is described in Table I. All networks are trained on a Lambda machine with an Intel Core$^{\text{TM}}$ i9-10920X CPU and four Nvidia Quadro RTX 6000 GPUs.

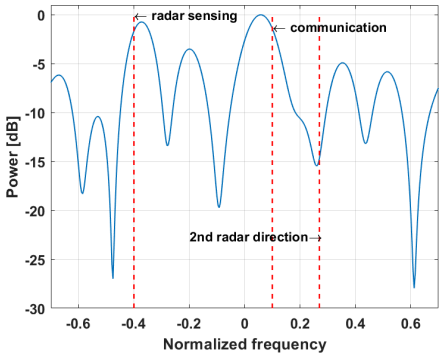
Table I: Hyper-parameters for training

Parameter	Value	
Models	Actor-Net	Critic-Net
Replay Buffer	4096	4096
Mini-batch	64	64
Learning rate	0.001	0.001
Decay	0.001	0.001

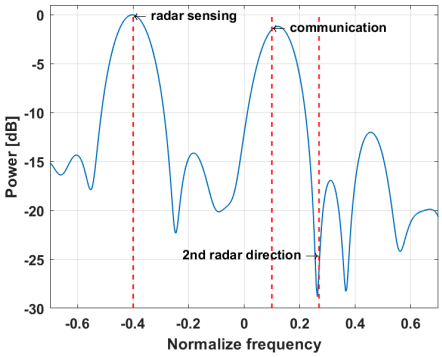
We assume that the radar tracking target is located at $f = -0.4$, corresponding to $\theta = -23.6^\circ$. The communicating user is located at $f = 0.1$, which corresponds to $\theta = 5.7^\circ$. A second radar direction is located at $f = 0.25$, corresponding to $\theta = 14.4^\circ$. As shown in Fig. 2, at the beginning of the iterative optimization, although two beams are formed in the directions of the radar target and the communication receiver, the sidelobe level in the direction of no interest is very high. After optimization, the sidelobe level is substantially reduced to effectively attenuate the interference to other automotive radar. Figs. 3 (a) and 3 (b) show the antenna selection processing. Fig. 4 gives the average reward during the training process. After around 80 epochs, the network can intelligently adjust the phases so that the main beam is steered to the ROI based on the current observation state.

V. CONCLUSIONS

We have proposed a reinforcement learning framework based on the Wolpertinger strategy to design an automotive radar ISAC system that intelligently adjusts the distribution of the antennas and the phases of quantized double phase shifters to steer its main beam to simultaneously track targets of interest and enhance communications ability and attenuate interference to other directions of no interest. The proposed method is suitable for extremely dimensional action spaces while avoiding exhausted action search. The feasibility of the proposed method has been verified by simulations.

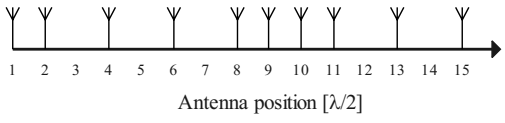


(a)

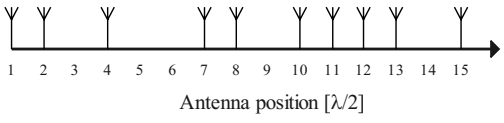


(b)

Fig. 2: (a) The transmit beamforming in the initial phase; (b) The transmit beamforming after optimization; Ground truth directions are indicated in red dash lines.



(a)



(b)

Fig. 3: (a) The transmit array configuration in the initial phase; (b) The transmit array configuration after optimization.

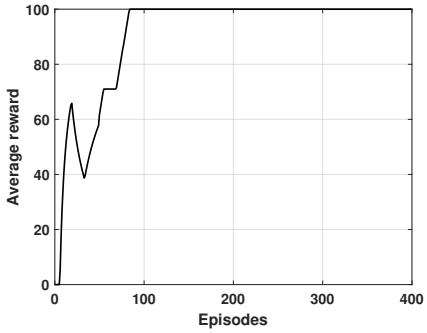


Fig. 4: The reward during the training.

VI. REFERENCES

[1] S. Sun, A. P. Petropulu, and H. V. Poor, "MIMO radar for advanced driver-assistance systems and autonomous driving: Advantages and challenges," *IEEE Signal Process. Mag.*, vol. 37, no. 4, pp. 98–117, 2020.

[2] S. Sun and Y. D. Zhang, "4D automotive radar sensing for autonomous vehicles: A sparsity-oriented approach," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 4, pp. 879–891, 2021.

[3] M. Markel, *Radar for Fully Autonomous Driving*. Boston, MA: Artech House, 2022.

[4] B. Li, A. P. Petropulu, and W. Trappe, "Optimum co-design for spectrum sharing between matrix completion based MIMO radars and a MIMO communication system," *IEEE Trans. Signal Process.*, vol. 64, no. 17, pp. 4562–4575, 2016.

[5] F. Liu, L. Zhou, C. Masouros, A. Li, W. Luo, and A. Petropulu, "Toward dual-functional radar-communication systems: Optimal waveform design," *IEEE Trans. Signal Process.*, vol. 66, no. 16, pp. 4264–4279, 2018.

[6] A. Hassanien, M. G. Amin, E. Aboutanios, and B. Himed, "Dual-function radar communication systems: A solution to the spectrum congestion problem," *IEEE Signal Process. Mag.*, vol. 36, no. 5, pp. 115–126, 2019.

[7] L. Zheng, M. Lops, Y. C. Eldar, and X. Wang, "Radar and communication coexistence: An overview: A review of recent methods," *IEEE Signal Process. Mag.*, vol. 36, no. 5, pp. 85–99, 2019.

[8] K. V. Mishra, M. R. B. Shankar, V. Koivunen, B. Ottersten, and S. A. Vorobyov, "Toward millimeter-wave joint radar communications: A signal processing perspective," *IEEE Signal Process. Mag.*, vol. 36, no. 5, pp. 100–114, 2019.

[9] F. Liu, C. Masouros, A. P. Petropulu, H. Griffiths, and L. Hanzo, "Joint radar and communication design: Applications, state-of-the-art, and the road ahead," *IEEE Trans. Commun.*, vol. 68, no. 6, pp. 3834–3862, 2020.

[10] L. G. de Oliveira, B. Nuss, M. B. Alabd, A. Diewald, M. Pauli, and T. Zwick, "Joint radar-communication systems: Modulation schemes and system design," *IEEE Trans. Microw. Theory Tech.*, vol. 70, no. 3, pp. 1521–1551, 2022.

[11] A. Hassanien, M. G. Amin, E. Aboutanios, and B. Himed, "Dual-function radar communication systems: A solution to the spectrum congestion problem," *IEEE Signal Processing Magazine*, vol. 36, no. 5, pp. 115–126, 2019.

[12] D. Ma, N. Shlezinger, T. Huang, Y. Liu, and Y. C. Eldar, "Joint radar-communication strategies for autonomous vehicles: Combining two key automotive technologies," *IEEE signal processing magazine*, vol. 37, no. 4, pp. 85–97, 2020.

[13] F. Liu, L. Zheng, Y. Cui, C. Masouros, A. P. Petropulu, H. Griffiths, and Y. C. Eldar, "Seventy years of radar and communications: The road from separation to integration," *arXiv preprint arXiv:2210.00446*, 2022.

[14] A. Ahmed, S. Zhang, and Y. D. Zhang, "Optimized sensor selection for joint radar-communication systems," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Barcelona, Spain, May 4–8, 2020, pp. 4682–4686.

[15] D. Ma, N. Shlezinger, T. Huang, Y. Liu, and Y. C. Eldar, "Joint radar-communication strategies for autonomous vehicles: Combining two key automotive technologies," *IEEE Signal Process. Mag.*, vol. 37, no. 4, pp. 85–97, 2020.

[16] ——, "FRaC: FMCW-based joint radar-communications system via index modulation," *IEEE J. Sel. Topics Signal Process.*, vol. 15, no. 6, pp. 1348–1364, 2021.

[17] A. Hassanien, M. G. Amin, Y. D. Zhang, and F. Ahmad, "Dual-function radar-communications: Information embedding using sidelobe control and waveform diversity," *IEEE Trans. Signal Process.*, vol. 64, no. 8, pp. 2168–2181, 2015.

[18] A. M. Elbir, K. V. Mishra, and S. Chatzinotas, "Terahertz-band joint ultra-massive MIMO radar-communications: Model-based and model-free hybrid beamforming," *IEEE J. Sel. Topics Signal Process.*, vol. 15, no. 6, pp. 1468–1483, 2021.

[19] A. M. Elbir and K. V. Mishra, "Joint antenna selection and hybrid beamformer design using unquantized and quantized deep learning networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 1677–1688, 2019.

[20] L. Xu, R. Zheng, and S. Sun, "A deep reinforcement learning approach for integrated automotive radar sensing and communication," in *IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, Trondheim, Norway, June 20-23, 2022, pp. 316–320.

[21] A. Hassanien, M. G. Amin, Y. D. Zhang, and F. Ahmad, "Phase-modulation based dual-function radar-communications," *IET Radar, Sonar & Navigation*, vol. 10, no. 8, pp. 1411–1421, 2016.

[22] X. Wang, A. Hassanien, and M. G. Amin, "Sparse transmit array design for dual-function radar communications by antenna selection," *Digital Signal Processing*, vol. 83, pp. 223–234, 2018.

[23] Z. Xu, F. Liu, and A. P. Petropulu, "Cramér-Rao bound and antenna selection optimization for dual radar-communication design," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Singapore, May 22-27, 2022, pp. 5168–5172.

[24] Texas Instruments Inc., "Design guide: TIDEP-01012 imaging radar using cascaded mmWave sensor reference design (REV. A)," [Available Online] https://www.ti.com/lit/ug/tiduen5a/tiduen5a.pdf, Mar. 2020.

[25] Z. Xu and A. P. Petropulu, "Phased array with improved beamforming capability via use of double phase shifters," in *IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, Trondheim, Norway, June 20-23, 2022, pp. 66–70.

[26] X. Yu, J. Zhang, and K. B. Letaief, "Doubling phase shifters for efficient hybrid precoder design in millimeter-wave communication systems," *Journal of Communications and Information Networks*, vol. 4, no. 2, pp. 51–67, 2019.

[27] A. M. Ahmed, A. A. Ahmad, S. Fortunati, A. Sezgin, M. S. Greco, and F. Gini, "A reinforcement learning based approach for multitarget detection in massive MIMO radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 57, no. 5, pp. 2622–2636, 2021.

[28] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[29] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 101–107, 2011.

[30] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015.