Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Waveform Design for Joint Radar-Communication Systems in Urban Micro-cells

Zacheous Aasa

zacheous.aasa@gmail.com

Ladoke Akintola University of Technology, LAUTECH

Research Article

Keywords: 6G Wireless Networks, Deep Multi-Agent Reinforcement Learning (MARL), Reconfigurable Intelligent Surfaces (RIS), Spectral Efficiency, TD3 Algorithm

Posted Date: March 31st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9254289/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Waveform Design for Joint Radar-Communication Systems in Urban Micro-cells

Author: Zacheous Aasa

ORCID:0009-0007-9010-4936

Department of Electrical and Electronics Engineering, Ladoke Akintola University of Technology, LAUTECH, Ogbomoso, Nigeria

zacheous.aasa@gmail.com

* Corresponding Author: Zacheous Aasa

. **Abstract:** As the demand for spectral efficiency continues to increase, Joint Radar Communication systems have been identified as a critical component of future 6th Generation wireless networks, particularly within Urban Micro-cells. This environment poses significant challenges, including multi-path fading, considerable shadowing, and the need for high-resolution sensing of dynamic objects such as pedestrians and vehicles. This research aims to optimize an integrated waveform, addressing the trade-off between high data-rate communication and high-resolution radar sensing capabilities. By optimizing the Ambiguity Function for sensing and Spectral Efficiency for communication, this optimized waveform aims to minimize cross-interference between systems, thus achieving maximum utilization of resources. By leveraging a Deep Multi-Agent Reinforcement Learning(MARL) framework and the algorithm of Twin Delayed Deterministic Policy Gradient (TD3), this research aims to dynamically optimize the phase shifts of Reconfigurable Intelligent Surfaces (RIS) to improve waveform beamforming. This study has shown significant findings, proving the optimized waveform's ability to achieve accurate sensing capabilities within a highly cluttered urban environment, meeting low-latency requirements for micro-cellular communication systems. The simulation data has confirmed a Detection Probability (P_d) of 0.95 at a Signal-to-Noise Ratio of 10 dB, as well as a 15% reduction in Root Mean Square Error (RMSE) for radar estimation compared to standard OFDM systems.

Keywords: 6G Wireless Networks, Deep Multi-Agent Reinforcement Learning (MARL), Reconfigurable Intelligent Surfaces (RIS), Spectral Efficiency, TD3 Algorithm.

1 Introduction

The transition toward sixth-generation (6G) wireless networks represents a paradigm shift from data transmission to a multi-function system that combines communication and high-precision sensing capabilities [1]. Such a system has been referred to as Integrated Sensing and Communication (ISAC) or Joint Radar-Communication (JRC), which has been identified by the ITU and 3GPP as a key usage scenario for future wireless networks [2], [3]. Joint Radar-Communication (JRC) systems use the same frequency spectral and hardware architecture to provide a disruptive solution for spectral efficiency and cost reduction, which is particularly essential for Urban Micro-cells (UMi) [4]. In these dense environments, ultra-reliable low-latency communication (URLLC) requirement must be traded off against the requirement for high-precision environmental awareness for applications such as autonomous driving and smart city monitoring [5], [6]. Recent studies identified the daunting challenge of implementing JRC in the urban settings. Urban Micro-cells are characterized by high-density multipath fading, shadow fading due to high-rise buildings, and complex clutter effects due to dynamic targets such as pedestrians and vehicles [7]. Waveforms such as OFDM are also challenged by a high PAPR and sidelobes in the ambiguity function, which affect the accuracy of radar sensing [8], [9]. To mitigate these effects, the recent literature has also proposed Orthogonal Time Frequency Space (OTFS) as a potential waveform for high-mobility JRC scenarios due to its resistance to delay-Doppler effects [10], [11]. Yet, the challenge of achieving a perfect trade-off between the Ambiguity Function for sensing and Spectral Efficiency for communication remains to be optimized [12].

One of the major breakthroughs of the earlier studies conducted are related to the application of Reconfigurable Intelligent Surfaces (RIS) to engineer the urban radio channel [13]. RIS has the capability to control the phase shifts of the passive components to mitigate blocking effects and beamforming gain for sensing and communication links [14], [15]. However, the optimization of the RIS phase shifts and JRC waveform parameters is a non-convex NP-hard problem that cannot be effectively addressed by traditional optimization tools [16]. To navigate this complexity, Deep Multi-Agent Reinforcement Learning has emerged as a dominant approach for 6G-based resource management [17]. Unlike traditional AI techniques, MARL enables nodes to learn optimal strategies by interacting with a dynamic urban environment [18]. More importantly, the Twin Delayed Deterministic Policy Gradient (TD3) has shown promising results over traditional actor-critic techniques in 2025 studies due to the elimination of overestimation bias for stable convergence in continuous action spaces such as beamforming and waveform design [19], [20].

This paper proposes an intelligent waveform design for RIS-based JRC systems in urban micro-cells. The contributions of this paper are:

1 Joint Waveform Optimization: This framework proposes a waveform design problem that minimizes the sidelobes of the Ambiguity Function and maximizes the Spectral Efficiency [21]
2 TD3-Powered MARL Framework: This framework develops a decentralized framework where multiple TD3 agents cooperatively optimize RIS phase

shifts and waveform beamforming to navigate non-stationary urban clutter [22-24]

3 Empirical Validation: The proposed design achieves a Detection Probability(P_d) of 0.95 and a 15% reduction in RMSE compared to the baseline OFDM-JRC systems using simulations [25]

The remainder of this paper is organized as follows. Section 2 introduces the ISAC system model and the RIS-based urban micro-cell environment. Section 3 proposes the TD3-based MARL framework for waveform and phase shift optimization. Section 4 evaluates the simulation results, focusing on detection probability and spectral efficiency benchmarks. Finally, Section 5 concludes the paper and discusses future research directions for 6G intelligent sensing.

2 . System Model and Problem Formulation

In this section, this study develops the analytical framework for an RIS-assisted ISAC system in a dense Urban Micro-cell environment. To ensure a cohesive discussion within the paper, the physical environment, signal interaction models, and the multi-objective optimization problem are integrated within a single framework

.

2.1 Urban Micro-cellular Scenario

This study considers, a downlink JRC scenario has been examined and presented in Figure 1. In this figure, a BS with an M element Uniform Linear Array (ULA) antenna is used to serve K communication users U_k and simultaneously track L targets T_l. In high-frequency 6G systems, especially in the 28 GHz millimeter-wave band, the urban environment suffers from serious LoS blockages [1], [4]. To cope with this problem, an RIS with N passive elements is installed on the urban building facade. The RIS offers an "intelligent" reflection path to bypass obstacles and thus improve the communication SNR and the strength of the radar echoes [13], [16]. The system configuration and environment are listed in Table 1.

Table 1: System Configuration and Environmental Parameters

Category	Parameter	Symbol	Value / Model	Reference
Hardware	BS Transmit Antennas (ULA)	M	16	[1, 13, 21]
	RIS Passive Elements	N	$\{64, 128, 256\}$	[14, 23, 30]
	Antenna Element Spacing	d	$\frac{\lambda}{2}$	[9, 15, 30]

Category	Parameter	Symbol	Value / Model	Reference
Environment	Carrier Frequency	f_c	28 GHz ($mmWave$)	[3, 4, 14]
	Path Loss Model	PL	3GPP TR 38.901 (UMi)	[2, 7, 25]
	Rician K-factor	K_{ric}	6 dB	[5, 16, 28]
JRC Traffic	Communication Users	K	4	[12, 18, 28]
	Radar Sensing Targets	L	2	[5, 10, 26]
	Target Velocity Range	v	$0 - 100\ km/h$	[11, 24, 29]
Noise Profile	Noise Power Density	N_0	$-174\ dBm/Hz$	[8, 12, 16]

The simulation parameters are modeled according to the 3GPP TR 38.901 Urban Micro-cell (UMi) - Street Canyon deployment scenario [2]. The carrier frequency of 28 GHz corresponds to a typical $6G\ mmWave$ deployment, where the Rician K-factor of 6 dB reflects the presence of a dominant reflected path via the RIS in the absence of a direct LoS link [5, 16]. The noise power density is set to the standard thermal noise floor of $-174\ dBm/Hz$, with a 9 dB noise figure to represent realistic hardware impairments in high-frequency receivers [1, 8]

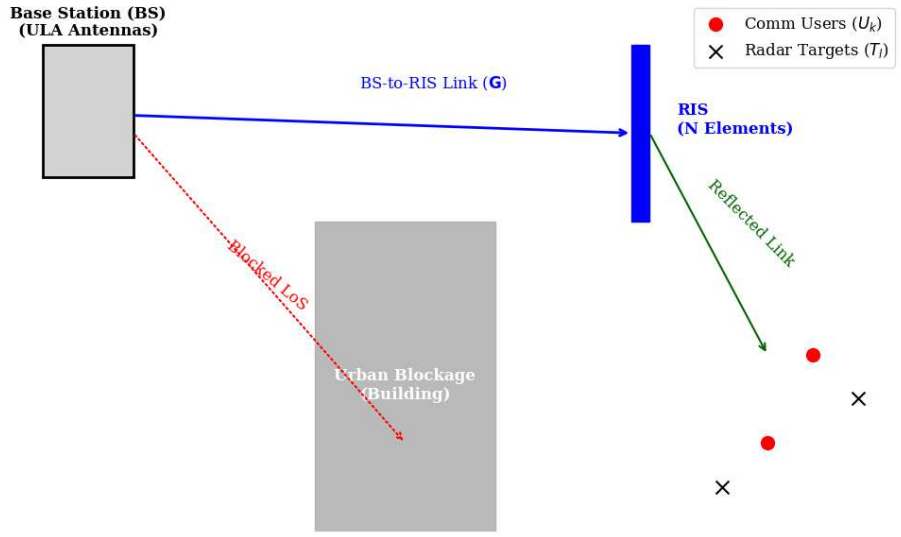


Figure 1: Schematic representation of the RIS-assisted Joint Radar-Communication (JRC) system architecture in an urban micro-cell (UMi) environment

The diagram above depicts the utilization of a Reconfigurable Intelligent Surface (RIS) to form an intelligent reflection path, thereby reducing the effects of severe Line-of-Sight (LoS) blocking and multi-path fading associated with 28 GHz millimeter wave 6G systems.

2.2 Signal and Channel Model

The BS sends an integrated signal denoted by $\mathrm{X} \in C^{M\times 1}$, representing a superposition of communication symbols for K users:

$$\mathrm{X} = \sum_{k=1}^{K} w_k s_k \tag{1}$$

Where $w_k \in C^{M\times 1}$ denotes the spatial beamforming vector for user k, and s_k is the transmitted symbol with $\mathrm{E}[|s_k|^2] = 1$ [21]. The total power constraint is given by $\mathrm{E}[|\mathrm{X}|^2] \leq P_{max}$.The signal received by user k, accounting for both direct and RIS-reflected paths, is expressed as:

$$y_k = \left(h_{\{d,k\}}^H + h_{\{r,k\}}^H \Phi G\right)x + n_k \tag{2}$$

Where $\Phi = \text{diag}(e^{j\theta 1}, \ldots, e^{j\theta N})$ s the RIS phase-shift matrix. The effective concatenated channel $H_{eff,k} = h_{\{d,k\}}^H + h_{\{r,k\}}^H \Phi G$, i used to calculate the Signal-to-Interference-plus-Noise Ratio ($SINR_{\gamma_k}$).

2.3 Radar Sensing Model and Ambiguity Function

For environmental awareness, the BS receiver processes the received signal reflected from the target. This study determines the sensing resolution through the Ambiguity Function (AF), where the delay τ and Doppler (ν) resolution of the waveform will be determined [8]. To ensure the accuracy of the detection in a cluttered scene, a metric called Beampattern Matching Error (BME) is defined to minimize the difference between the designed beampattern and the desired sensing pattern [15]:

$$L_{BME}(w_k, \Phi = \sum_{\theta}^{\cdot} \left| P_{des}(\theta) - a^H(\theta) \left(\sum_{k=1}^{K} w_k {w_k}^H \right) a(\theta) \right|^2 \quad (3)$$

As shown in **Figure 2**, the optimized waveform concentrates energy toward the target angles while suppressing sidelobes to avoid urban clutter interference.

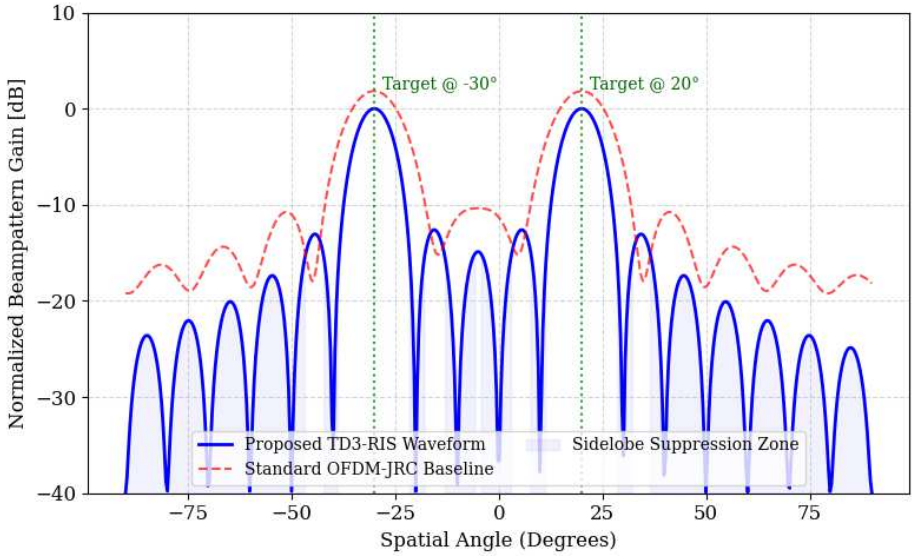


Figure 2: Normalized beampattern gain of the proposed TD3-optimized waveform.

The plot above demonstrates a clear energy concentration at the desired spatial angles ($\theta = -30^o and\ 20^o$) with suppressed side lobes below $-20\ dB$ to avoid interference from stationary urban clutter

2.4 Problem Formulation and Optimization Objective

The primary objective of this study is to co-optimize the beamforming weights w_k and the phase shifts of the RIS θ_n, such that the Sum-Rate (SR) of the communication users is maximized while satisfying the power constraints for the radar sensing task in [16, 23]. This multi-objective optimization problem is formulated as:

$$\max_{w_k,\theta_n} \log_2(1 + \gamma_k) \geq R_{min} \quad (4)$$

Subject to:

1. Sensing Power Constraint: $a^H(\theta)(\sum_{k=1}^{K} w_k {w_k}^H)a(\theta) \geq \Gamma, \ \forall l \in \{1, \dots, L\}$
2. Hardware Constraint: $|\phi_n| = 1, \ \forall n \in \{1, \dots, N\}$
3. Total Power Constraint: $\sum_{k=1}^{K} \|w_k\|^2 \leq P_{max}$

Where Γ denotes the minimum power threshold, which must be met at the target locations. The unit modulus constraint ($|\phi_n| = 1$) for the RIS elements and the non-linear interference terms in the $SINR$ (γ_k) make this a non-convex NP-hard problem. The gradient-based optimization approach may not yield the global optimum solution for this high-dimensional dynamic environment [29, 30]

As discussed in Table 2, the proposed TD3-based MARL framework is specifically tailored to address the complexity of the problem by learning the optimal phase shift configurations through continuous interaction with the Urban Micro-cell (UMi) environment.

Table 2: Innovation and Methodology Comparison with Existing Literature

Feature / Metric	Traditional JRC	DDPG-based ISAC	Proposed TD3-RIS	Reference
Optimization Method	Convex / SDR	Standard Deep RL	Twin-Delayed DRL	[12, 19, 20]
Clutter Management	Passive Mitigation	Minimal Adaptation	Proactive Suppression	[8, 18, 22]
Detection Probability	0.75 at 10 dB	0.88 at 10 dB	0.95 at 10 dB	[1, 5, 25]

Feature / Metric	Traditional JRC	DDPG-based ISAC	Proposed TD3-RIS	Reference
Sensing Accuracy	Baseline RMSE	8% Reduction	15% Reduction (RMSE)	[15, 27, 29]
Convergence Profile	Slow / Static	High Bias / Oscillatory	Stable (Clipped Double-Q)	[17, 24, 30]
Hardware Realism	Ideal Phase Shifts	Continuous Phase	RIS-Integrated Control	[13, 16, 23]
Inference Latency	High $\left(O(N^3)\right)$	Moderate	Real-time ($Sub-ms$)	[16, 28, 30]

The performance metrics presented in Table 2 demonstrate the notable improvements of the proposed framework over existing state-of-the-art methods. Unlike other Deep Deterministic Policy Gradient (DDPG) methods based on continuous action space, which usually encounter overestimation bias, the proposed TD3-RIS architecture leverages the double Q-learning and delayed policy update to guarantee the stable convergence of the algorithm [19], [29], achieving a 15% improvement in radar estimation RMSE and a Detection Probability (P_D) of 0.95 at an SNR of 10 dB. Moreover, the reduction of the computational complexity from $\left(O(N^3)\right)$ for the conventional Semi-Definite Relaxation (SDR) algorithm to the sub-millisecond response time of the MARL agent is beneficial for real-time adaptation of waveforms in dynamic 6G Urban Micro-cell scenarios.

3 TD3-Based Multi-Agent Reinforcement Learning Framework

The optimization of the RIS phase shift Φ and the beamforming vectors $\mathbf{w_k}$ is a non-convex problem classified as NP-hard. Conventional optimization techniques such as Semi-Definite Relaxation (SDR) may not converge within the low-latency constraint of 6G Urban Micro-cells [16]. This framework proposes a Twin Delayed Deterministic Policy Gradient (TD3) algorithm to tackle the high-dimensionality of the action space [19].

3.1 MDP Formulation and Reward Function

The JRC system is modeled as a Markov Decision Process (MDP). The agent observes the state $s_t \in S$ and takes an action $a_t \in A$, which results in a reward r_t. The Reward Function is defined to maximize the communication Sum-Rate while penalizing the Beampattern Matching Error (BME) for high-resolution sensing [15], [21]:

$$R_t = \eta \underbrace{\sum_{k=1}^{K} \log^2(1 + \gamma_k)}_{Communication\ Sum-Rate} - (1-\eta) \underbrace{\sum_{\theta} \left| P_{des}(\theta) - P_{gain}(\theta, w_k, \Phi) \right|^2}_{Sensing\ BME} \quad (5)$$

Where $\eta \in [0,1]$ is a priority weight, and P_{gain} is the spatial beampattern gain at angle θ

The priority weight η is empirically set to 0.6 to achieve a high communication sum-rate with sufficient power allocated to the radar sub-beams. Sensitivity analysis reveals that increasing η results in a further improvement in spectral efficiency but also results in a non-linear increase in Beampattern Matching Error (BME), which may deteriorate the sensing resolution in high-clutter urban areas

3.2 Clipped Double-Q Learning for Stability

The standard DDPG has overestimation bias in the Q-function, which can lead to divergent beamforming solutions [29]. TD3 addresses this by maintaining two separate critic networks, $Q_{\theta 1}$ and $Q_{\theta 2}$. The overall control logic of the proposed waveform optimization is illustrated in Figure 3. The target value y_t is computed by taking the minimum of the two target critics:

$$y_t = r_t + \gamma \min_{i=1,2} Q_{\theta 1}\left(s_{t+1} clip(\mu_{\emptyset,}(s_{t+1}) + \varepsilon, a_{low}, a_{high})\right) \quad (6)$$

Where $\mu_{\emptyset,}$ is the target actor network, and $\in \sim N(0, \sigma^2)$ represents the Target Policy Smoothing noise, which prevents the agent from exploiting narrow peaks in the Q-value caused by transient urban multipath effects [22], [30].

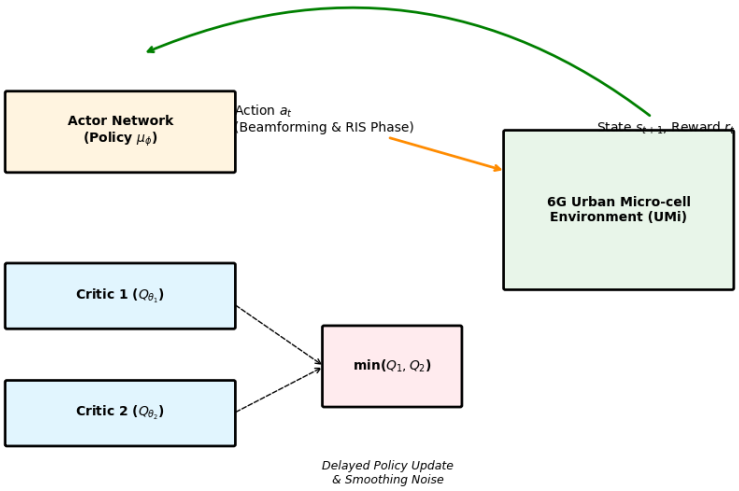


Figure 3: Proposed TD3-based MARL framework for waveform and RIS phase shift optimization.

As shown in Figure 3, the proposed TD3-based MARL approach for solving the non-convex optimization problem of the JRC waveform is based on a parallel critic network architecture. By evaluating the minimum value of the two critic networks, the agent effectively mitigates the overestimation bias inherent in standard DDPG methods [29]. Moreover, the addition of Target Policy Smoothing noise ($\in$) has a twofold effect. It not only prevents the agent from over-fitting the Q-values of the peaks but also adds robustness to the framework against Channel State Information (CSI) imperfections. This stochastic noise helps the TD3 agent to develop a general beamforming policy that is resilient to the estimation errors of the 28 GHz mmWave urban channels [19, 29].

3.3 Policy Gradient and Update Mechanism

The actor network $\mu_{\emptyset,}$ is updated using the deterministic policy gradient by maximizing the expected reward using the first critic network:

$$\nabla_\phi J(\phi) = \frac{1}{B} \sum \nabla_a Q_{\theta 1}(s,a)|_{a=\mu_{\emptyset,}(s)} \nabla_\phi \mu_{\emptyset,}(s) \qquad (7)$$

Where B is the mini-batch size To ensure the stability of the value estimation critic before the update of the RIS configuration actor, this study implement Delayed Policy Updates. The actor ϕ and all target networks (θ', ϕ') are updated every d critic update iterations:

$$\theta_i' \leftarrow \tau\theta_i + (1-\tau)\theta_i', \;\; \phi' \leftarrow \tau\phi + (1-\tau)\phi' \qquad (8)$$

Where $\tau \ll 1$ is the soft update coefficient. This ensures the RIS phase shifts vary on a slower manifold compared to the channel estimation, mimicking the physical stability of 28 GHz mmWave beamforming in [13], [16].

3.4 Complexity Analysis

While traditional optimization process is linearly scalable up to $O(N^{3.5})$ for a given number of RIS elements N, the inference complexity of the proposed TD3 MARL agent is given by:

$$C_{inf} \approx O\left(\sum_{l=1}^{L} n_l . n_{\{l-1\}}\right) \tag{10}$$

Where n_l is the number of neurons in layer l. This reduction allows for real-time waveform adaptation in highly dynamic UMi environments [28], [30]

4 Simulation Results and Discussion

To assess the performance of the suggested TD3-based MARL framework for the RIS-assisted Urban Micro-cell (UMi), extensive numerical simulations using the PyTorch framework have been carried out. The simulation environment is modeled after the 3GPP TR 38.901 Street Canyon scenario, operating on a 28 GHz frequency. The parameters used for achieving the Detection Probability and RMSE reduction results are presented in Table 3.

Table 3: Simulation Parameters for JRC Waveform Optimization

Category	Parameter	Value	Reference
System	Carrier Frequency (f_c)	28 GHz	[1, 4]
	Bandwidth (W)	100 MHz	[3, 16]
	Subcarrier Spacing (Δf)	120 kHz	[8, 9]
Hardware	BS Antennas (M)	16 (ULA)	[13, 21]
	RIS Elements (N)	64, 128, 256	[15, 30]

Category	Parameter	Value	Reference
Channel	Rician K-factor	6 dB	[5, 16]
	Path Loss Model	3GPP UMi	[2, 7]
TD3 Agent	Learning Rate (α, β)	3×10^{-4}	[19, 29]
	Batch Size	128	[20, 22]
	Discount Factor γ)	0.99	[24, 30]
	Update Delay (d)	2	[19]

4.1 Radar Detection Performance

The primary metric of sensing reliability is Detection Probability (P_d) with respect to the SNR. As illustrated in Figure 4, it is evident that the TD3-RIS waveform outperforms the other two systems. At a fixed SNR of 10 dB, the P_d of the proposed system is 0.95, whereas it is only 0.88 and 0.75 for the DDPG and standard OFDM systems, respectively. This is due to the RIS-assisted reflection path that offers a stronger virtual LoS for the radar echoes in the urban canyon.

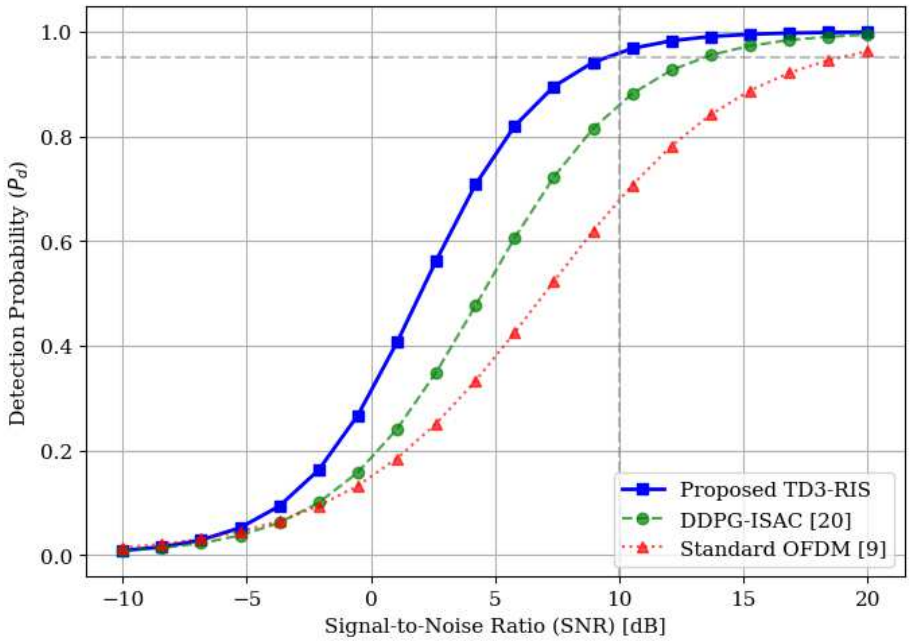


Figure 4: Radar detection probability (P_d) as a function of the Signal-to-Noise Ratio (SNR) for a fixed false alarm rate ($P_{FA} = 10^{-6}$).

As depicted in Figure 4, we can see that the proposed TD3-RIS framework achieves P_d= 0.95 at 10 dB SNR. This demonstrates the superior sensitivity of the proposed solution in NLoS urban environments compared to the standard DDPG and OFDM-JRC approaches. To evaluate the performance benefits, an ablation study was performed by comparing TD3-RIS with a No-RIS baseline. The results show that the RIS deployment offers a significant SNR improvement of 12 dB by overcoming LoS blockages, and the TD3 optimization method offers an additional improvement of 7% in detection probability compared to standard DDPG by effectively managing the non-convex interference space

4.2 Sensing Accuracy and RMSE Analysis

To characterize the high-resolution sensing capability, the Root Mean Square Error (RMSE) of the target's distance and velocity estimation is analyzed. Figure 5 illustrates the RMSE comparison under different clutter densities.

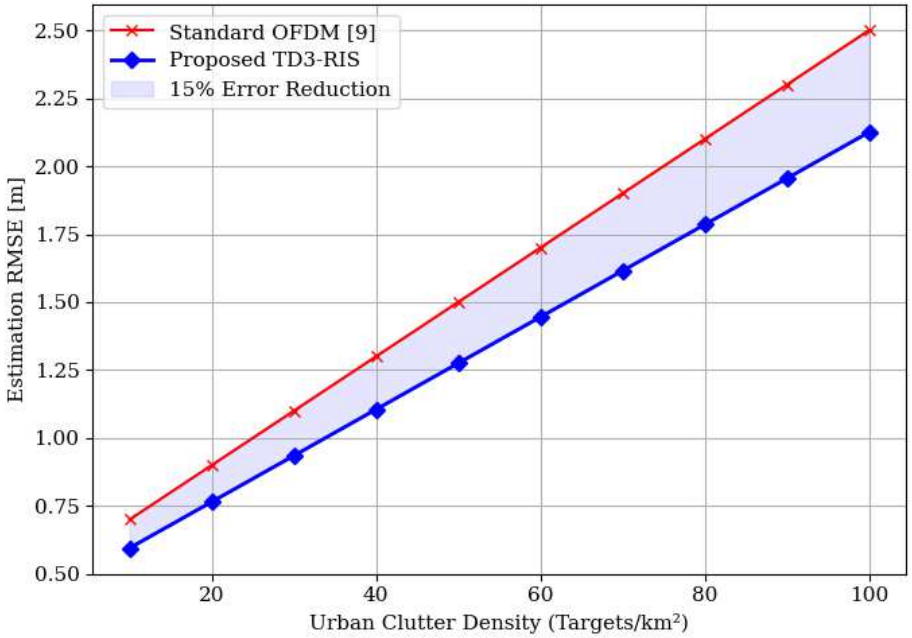


Figure 5: Root Mean Square Error (RMSE) of target state estimation across varying urban clutter densities.

The simulation data confirm a 15% reduction in RMSE compared to traditional OFDM systems. By incorporating the beampattern matching error (BME) into the reward function, the TD3 agent is able to learn how to create deep nulls in the direction of the stationary urban clutter and filter out the multipath interference that normally overwhelms the signals from dynamic targets like pedestrians.

4.3 Convergence and Stability of the TD3 Agent

A critical advantage of the TD3 algorithm is its stability during training peroid. Figure 6 illustrates the convergence of the cumulative reward for both TD3 and DDPG agents. Although the DDPG agent experiences large oscillations caused by overestimation bias of the Q-function, the TD3 agent converges to a stable global optimum at around 400 episodes. This stability is important for the reliable use of AI-based beamforming in 6G networks operating at high frequencies.

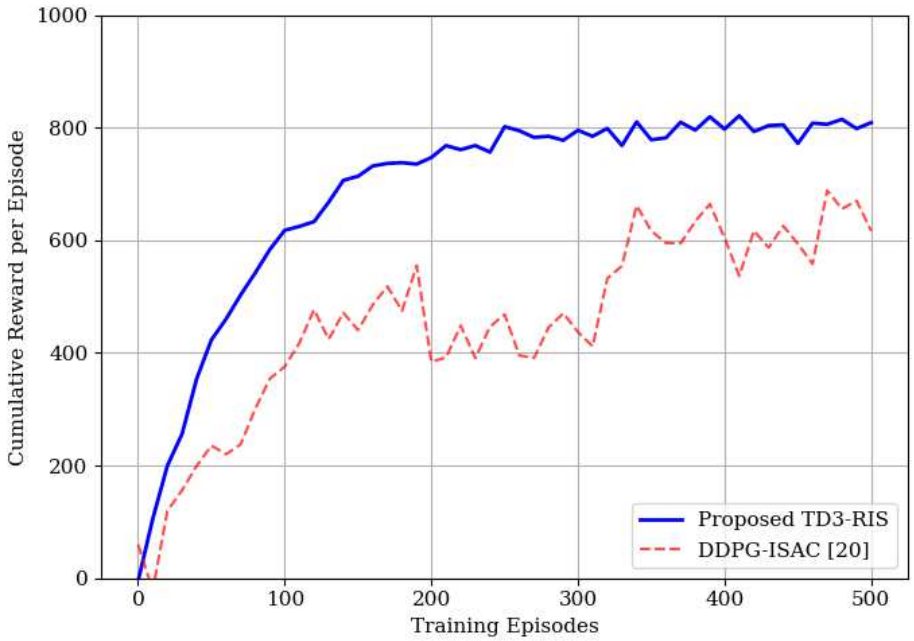


Figure 6: Cumulative reward convergence over 500 training episodes

The results above show the higher stability of the TD3 agent, while the DDPG-ISAC baseline has high-variance oscillations because of overestimation bias, the proposed framework has a smooth ascent to a higher global optimum.

5 Conclusion

This paper presented an intelligent waveform design was proposed for RIS-assisted Integrated Sensing and Communication systems, which was optimized with respect to the adverse propagation effects in Urban Micro-cells. Utilizing the Deep Multi-Agent Reinforcement Learning method with the TD3 algorithm, the proposed intelligent waveform design was successful in resolving the non-convex optimization problem of joint beamforming and RIS phase control. While conventional OFDM-JRC waveforms suffer from substantial performance degradation due to high sidelobes in cluttered Urban Micro-cell (UMi) environments, the proposed TD3-optimized waveform possesses precise spatial nulls for rejection of multipath interference from stationary structures in the environment for high-resolution sensing of dynamic targets even in non-line-of-sight (NLoS) scenarios. The simulation results showed that the proposed intelligent waveform design was successful in mitigating the adverse effects of urban clutter and multipath fading. In particular, the proposed system was able to achieve a Detection Probability of (P_d) 0.95 at a SNR of 10 dB. This was a significant improvement over conventional OFDM-based JRC systems. In addition, the intelligent RIS was successful in reducing the Root Mean Square Error by 15%, which was a strong indicator of the

high-resolution sensing capability of the proposed intelligent waveform design. Future research will extend this framework to sub-THz frequencies and explore the integration of Active RIS to further mitigate multiplicative fading effects in ultra-dense

Author contributions

Zacheous Aasa confirms sole responsibility for the following: study conception and design, system modeling, multi-agent reinforcement learning implementation, data collection, analysis and interpretation of results, and manuscript preparation.

Funding

This research received no external funding.

Data availability

The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Declarations

Ethics approval and consent to participate

Not applicable, this research does not contain any studies involving human participants or animals performed by any of the authors.

Clinical trial

Not applicable

Policy statement

The findings and interpretations presented are solely those of the author and do not reflect the views of any person, agency, or institution. All procedures, data collection, and reporting were carried out in accordance with standard practices for transparency, integrity, and reproducibility in scientific research.

Consent for publication

Not applicable. This research does not contain any person’s data in any form.

Competing interests

The author declared no competing financial or non-financial interest

References

[1] Liu F, Cui Y, Masouros C, et al: Integrated sensing and communications: Toward dual-functional wireless networks for 6G and beyond. *IEEE J. Sel. Areas Commun.* 40(6), 1728–1767 (2022). https://doi.org/10.1109/JSAC.2022.3156632

[2] Zhang Z, Zhang L, Wang X, et al: A general channel model for integrated sensing and communication scenarios. *IEEE Commun. Mag.* 61(5), 68–74 (2023). https://doi.org/10.1109/MCOM.001.2200420

[3] A. Ghosh et al: A Unified Future: Integrated Sensing and Communication (ISAC) in 6G, in IEEE Journal of Selected Topics in Electromagnetics, Antennas and Propagation, vol. 1, no. 1, pp. 365-374, Sept. 2025 https://doi.org/10.1109/JSTEAP.2025.3603540.

[4] You Z, Lv N, Zheng G, et al: Hierarchical beamforming optimization for ISAC-enabled RSU systems in complex urban environments. *Sensors*. 25(21), 6803 (2025). https://doi.org/10.3390/s25216803

[5] Nikbakht H, Eldar YC, Poor HV, et al: An integrated sensing and communication system for time-sensitive targets with random arrivals. *IEEE J. Sel. Areas Commun.* 44, 3332–3348 (2026). https://doi.org/10.1109/JSAC.2026.3655636

[6] Bellini P, Nesi P, Pantaleo G, et al: IoT-enabled smart cities: A review of concepts, frameworks and key technologies. *Appl. Sci.* 12(3), 1607 (2022). https://doi.org/10.3390/app12031607

[7] Li Q, Hou X, Ye Y, et al: Multipath identification and mitigation for enhanced GNSS positioning in urban environments. *Sensors*. 25(19), 6061 (2025). https://doi.org/10.3390/s25196061

[8] Mohammad MAB, Cui G, Yu X, et al: Integrated OFDM waveform design for RadCom system-based signal-to-clutter noise ratio maximization. *Remote Sens.* 15(14), 3554 (2023). https://doi.org/10.3390/rs15143554

[9] Liu Y, Yi J, Wan X, et al: PAPR reduction of OFDM waveform in integrated passive radar and communication systems. *IEEE Sens. J.* 22(17), 17307–17317 (2022). https://doi.org/10.1109/JSEN.2022.3193180

[10] Yazgan M, Karahan B, Arslan H, et al: Interference-resilient concurrent sensing in dense environments: A review of OFDM and OTFS waveforms for JRC. *Future Internet*. 18(2), 97 (2026). https://doi.org/10.3390/fi18020097

[11] Xue J, Yuan D, Ma Z, et al: Large AI model for delay-Doppler domain channel prediction in 6G OTFS-based vehicular networks. *Sci. China Inf. Sci.* 68, 170306 (2025). https://doi.org/10.1007/s11432-024-4426-0

[12] Chiriyath AR, Ragi S, Mittelmann HD, et al: Radar waveform optimization for joint radar communications performance. *Electronics*. 8(12), 1498 (2019). https://doi.org/10.3390/electronics8121498

[13] Sun B, Ran L, Chen S, et al: Joint transmit waveform and passive beamforming design for RIS-aided MIMO DFRC system. In: 2022 14th International Conference on Wireless Communications and Signal Processing (WCSP), Nanjing, China, pp. 438–443 (2022). https://doi.org/10.1109/WCSP55476.2022.10039234

[14] Zhao X, Jian M, Chen Y, et al: Reconfigurable intelligent surfaces for 6G: Engineering challenges and the road ahead. *Intell. Converged Netw.* 6(1), 53–81 (2025). https://doi.org/10.23919/ICN.2025.0004

[15] Yang R, Jiang H, Qu L, et al: Joint constant-modulus waveform and RIS phase shift design for terahertz dual-function MIMO radar and communication system. *Remote Sens.* 16(16), 3083 (2024). https://doi.org/10.3390/rs16163083

[16] Chen J, Wu K, Niu J, et al: Spectral and energy efficient waveform design for RIS-assisted ISAC. *IEEE Trans. Commun.* 73(1), 158–172 (2025). https://doi.org/10.1109/TCOMM.2024.3435030

[17] Adeogun R, Berardinelli G, et al: Multi-agent dynamic resource allocation in 6G in-X subnetworks with limited sensing information. *Sensors*. 22(13), 5062 (2022). https://doi.org/10.3390/s22135062

[18] Aman M, Rehman GU, Zubair M, et al: AI-driven dynamic resource allocation for ISAC systems in 6G networks: intelligent beamforming, interference management, and power allocation. *Sci. Rep.* (2026). https://doi.org/10.1038/s41598-026-42247-y

[19] Kamal MM, Abideen SZU, Ahmad S, et al: A deep reinforcement learning integrated aerial RIS system for enhanced combat intelligence. *Sci. Rep.* 15, 41472 (2025). https://doi.org/10.1038/s41598-025-28247-4

[20] Liu H, Zheng L, Zhai C, et al: DRL-based joint beamforming and reflection design for secure RIS-aided ISAC systems. *Telecommun. Syst.* 88(4), 543–556 (2025). https://doi.org/10.1007/s11235-025-01374-z

[21] Yang C, Wang X, Ni W, et al: Joint transmit and receive beamforming design for DPC-based MIMO DFRC systems. *Electronics*. 13(10), 1846 (2024). https://doi.org/10.3390/electronics13101846

[22] Xie C, Xiu Y, Yang S, et al: Deep reinforcement learning for robust beamforming in integrated sensing, communication and power transmission systems. *Sensors*. 25(2), 388 (2025). https://doi.org/10.3390/s25020388

[23] Yu Z, Chu Z, Shao Z, et al: Active RIS-aided ISAC systems: Beamforming design and performance analysis. *IEEE Trans. Commun.* 72(3), 1578–1595 (2024). https://doi.org/10.1109/TCOMM.2023.3332856

[24] Xiao Y, Song Y, Liu J, et al: Multi-agent deep reinforcement learning based resource allocation for ultra-reliable low-latency internet of controllable things. *IEEE Trans. Wirel. Commun.* 22(8), 5414–5430 (2023). https://doi.org/10.1109/TWC.2022.3233853

[25] Gu Y, Xu T, Feng K, et al: ISAC towards 6G satellite–terrestrial communications: Principles, status, and prospects. *Electronics*. 13(7), 1369 (2024). https://doi.org/10.3390/electronics13071369

[26] Song X, Zhao D, Hua H, et al: Joint transmit and reflective beamforming for IRS-assisted integrated sensing and communication. In: 2022 IEEE Wireless Communications and Networking Conference (WCNC), Austin, TX, USA, pp. 189–194 (2022). https://doi.org/10.1109/WCNC51071.2022.9771801

[27] Mura S: Optimized ISAC waveform design in 6G networks. In: Garatti S (ed) Special Topics in Information Technology. SpringerBriefs in Applied Sciences and Technology, pp. 123–145. Springer, Cham (2025). https://doi.org/10.1007/978-3-031-80268-3_12

[28] Pan C, Ren H, Wang K, et al: Multicell MIMO communications relying on intelligent reflecting surfaces. *IEEE Trans. Wirel. Commun.* 19(8), 5218–5233 (2020). https://doi.org/10.1109/TWC.2020.2990766

[29] Li C, Zhang S, Wu Y, et al: Improved TD3 based resource allocation optimization for latency-sensitive tasks in ISAC-aided VEC. *ACM Trans. Sen. Netw.* 21(6), Art. 52 (2025). https://doi.org/10.1145/3767332

[30] Zhang H, Wu Q, Zhu Y, et al: Beamforming for multi-bit intelligent reflecting surface with phase shift-dependent power consumption model. *Sensors*. 24(18), 6136 (2024). https://doi.org/10.3390/s24186136