Eng. Sci., 2025, **37**, 1817

Engineered Science
DOI: https://dx.doi.org/10.30919/es1817

Q-Learning Based Equivalent False Target Generation for Multi-point Source Jamming System

Zhihao Cai, Shiqi Xing,* Sinong Quan, Xinyuan Su, Junpeng Wang and Weize Meng

Abstract

In the multi-point source jamming (MPSJ) system, the antenna phase error, as an intrinsic issue, seriously affects the distribution of equivalent false targets (EFTs). To ensure that the position of EFTs remains within the constraint range in the presence of phase error, a multi-point source jamming system equivalent false target generation method is proposed in this paper. First, by establishing the spatial geometric relationship between the triplet antenna and the target, the distribution law of EFT in the MPSJ system is derived. Second, based on the feasibility of the MPSJ system EFT generation method, the MPSJ system EFT generation model is constructed by introducing the constraint function. Third, a dual dynamic inductive Q-learning (DDIQL) algorithm with a head-tail dynamic inductive strategy is proposed to solve this model, wherein initial optimization and constraint adjustment are emphasized. Experiments results demonstrate that when the phase error is 9°, the EFTs corresponding to the generated amplitude and phase parameters are all controllable in position. When the radius of the constraint range is reduced to the baseline length of the MPSJ system, the positions of EFTs corresponding to more than 10% of the amplitude and phase parameters are outside the constraint range.

Keywords: Equivalent false target generation; Multi-point source jamming system; Dual dynamic inductive Q-learning algorithm.

Received: 30 August 2025; Revised: 11 October 2025; Accepted: 15 October 2025.
Article type: Research article.

1. Introduction

Against the backdrop of the continuous advancement of modern radar technology, jamming techniques in electronic countermeasure systems are faced with increasingly complex challenges.[1-5] As an effective jamming method against monopulse radar, the cross-eye jamming has been subjected to extensive and in-depth research.[6] The core principle of cross-eye jamming is that the jammer emits one or more pairs of signals with the amplitude ratio close to 1 and the phase difference close to 180°, causing the monopulse radar to mistakenly point to a false target position, thereby enabling effective deflection of the monopulse radar to an area outside the line connecting the jamming sources. The development course of cross-eye jamming can be roughly divided into three stages: traditional two-point source cross-eye jamming, two-point source reverse cross-eye jamming, and multi-point source reverse cross-eye jamming.[7]

The cross-eye jamming technique was first inspired by the physical phenomenon of angle glint. Redmill *et al*.[8] conducted mathematical modeling and system design for traditional two-point source cross-eye jamming. However, due to the limitations of hardware conditions at that time, the 180° phase difference between the two jamming sources was difficult to be precisely controlled, resulting in poor jamming effects. For a subsequent period, the research on cross-eye jamming was relatively scarce.

To meet the stringent requirements of traditional two-point source cross-eye jamming for system parameters, Plessis *et al*. innovatively introduced the reverse antenna structure into the cross-eye jamming system, conducting rigorous mathematical derivations and experimental validations.[9-13] Through mathematical modelling, Ma *et al*.[14] elucidated the effect of target echo on the performance of cross-eye jamming. Liu *et al*.[15] primarily modelled the sum and difference channels of monopulse radar, concluding that cross-eye jamming distorts the direction pattern of the difference channel, thereby affecting radar angle measurement. Wang *et al*.[16] compared the impacts of two-point source reverse cross-eye and non-reverse cross-eye on the sum and difference channels of monopulse radar.

To tackle the challenge of low system parameter tolerance in two-point source reverse cross-eye jamming, Liu *et al*.[17]

College of Electronic Science and Technology, National University of Defense Technology, Changsha, 410073, China
**Email:* xingshiqi@nudt.edu.cn (S. Xing)

adopted multi-loop cross-eye jamming to ameliorate the harsh implementation conditions. By utilizing the median value of cross-eye gain, Yang *et al*.[18] analyzed the requirement of jamming-to-signal ratio for cross-eye jamming, indicating the conditions for achieving stable jamming. Under the condition of isolating target echo, Lee *et al*.[19] analyzed the parameter tolerance of reverse cross-eye in linear arrays. Liu *et al*.[20] proposed the rectangular array reverse cross-eye jamming technology, which demonstrates superior jamming range and angle compared with two-source reverse cross-eye jamming.

To overcome the limitations of complex multi-point source cross-eye jamming systems that are difficult to practicalize, Chen *et al*.[21] initiated from the principle of testing the angle measurement accuracy of monopulse radars in radio frequency hardware-in-the-loop simulation systems, proposing a multi-element vector synthesis jamming method. Based on the perspective of electromagnetic field vector synthesis, this method calculates the position of the equivalent radiation center (apparent false target) by means of the Poynting vector method, with its jamming range being a two-dimensional planar space-larger than the one-dimensional range of cross-eye jamming. Ning *et al*.[22] constructed a three-element combined jamming experimental model, and field test data verified the effectiveness of the jamming. Moreover, Ning *et al*. further established the "triplet antenna" as the basic unit structure for multi-vector synthetic interference. This structure adopts a triangular spatial distribution of three antennas (rather than the traditional one-dimensional linear layout). By establishing a mathematical model between the apparent position of the false target and the amplitude phase of the three antennas, theoretical derivation proves that this configuration can extend the originally linear interference range to a two-dimensional plane, thereby greatly enhancing the dimension and flexibility of the interference efficiency.

More importantly, the practical application of this technology still confronts the severe challenges long existing in cross-eye jamming: signal phase errors significantly affect the spatial position accuracy of synthetic false targets, leading to jamming effects deviating from expectations or even complete failure. In essence, the vector synthesis process of multi-point source signals is extremely sensitive to phase consistency, while phase inconsistency between antenna channels in actual systems is unavoidable. Therefore, instead of requiring precise positioning of false targets, it suffices that a series of false target positions fall within the constrained jamming range when considering the influence of phase errors.

To address the aforementioned challenges, this paper proposes a multi-point source jamming (MPSJ) system equivalent false target (EFT) generation method. The primary contributions are as follows: (1) Through the determination of spatial geometric relationships between triplet antenna and targets in MPSJ system, the distribution law of synthetic false targets is theoretically derived. (2) Based on contour map analysis of triplet antenna configurations, the feasibility of the proposed the MPSJ system EFT generation method is validated, and a constrained function is introduced to construct the MPSJ system EFT generation model. (3) Based on the Q-learning algorithm, through optimizing the initial values of the Q-table (head) and the reward function (tail), thus a dual dynamic inductive (DI) Q-learning algorithm with head-tail DI strategy is proposed. It is capable of deriving amplitude and phase parameters corresponding that the EFTs of the MPSJ system locate within the constraint range. Two different jamming cases are considered and the outcomes ensure that the generated EFTs remained position-controllable under a phase error of 9°, thereby validating the effectiveness of the proposed method.

2. Materials and methods

2.1 Spatial geometric configuration of MPSJ system

First, the layout of triplet antenna in the absence of target echo is determined.[23] As shown in Fig. 1(a), a spherical coordinate system is constructed with the monopulse radar P as the center. In this coordinate system, the azimuth angle is defined based on the XOY plane: taking P as the reference point, with the distance r from the triplet antenna plane to P as the radius, the direction is determined along a directed circular arc. The pitch angle is defined based on the XOZ plane: similarly taking P as the reference point, with the distance r from the plane to P as the radius, the direction is determined along a directed circular arc. The triplet antenna a, b, and c form a triangle, with their position coordinates being (φ_a,θ_a), (φ_b,θ_b), and (φ_c,θ_c), respectively. The baseline lengths between antennas a, b, and c are denoted as x_1, x_2, and x_3 in sequence. When target echo exists, the spatial position of the target is added to the spherical coordinate system of the MPSJ system, as shown in Fig. 1(b). The target O is located at $(\varphi_o,\theta_o)=(0,0)$.

In Fig. 1, the equivalent radiation center is the inferred position of the target radiation source obtained by analyzing the normal direction of the distorted equiphasic surface after the radar receives echoes from the triplet antenna. However, since no real radiation source exists at this position, the equivalent radiation center can essentially be regarded as a false target perceived by the radar, namely the equivalent false target of the MPSJ system. Additionally, given that the distance r between the monopulse radar and the target is much greater than the actual distance d_T between the EFT and the target, the azimuth component d_φ and pitch component d_θ of the actual distance between the EFT and the target can be calculated by Eq. (1).[24]

$$\begin{cases} d_\varphi = r\cos\theta_T \sin\varphi_T \approx r\varphi_T \\ d_\theta = r\sin\theta_T \approx r\theta_T \end{cases} \tag{1}$$

2.2 Distribution law of the equivalent false target

Subsequently, the coordinates of the triplet antenna EFT are derived based on the principle of multi-element vector synthesis (Eq. (2)). The average energy flux density at the

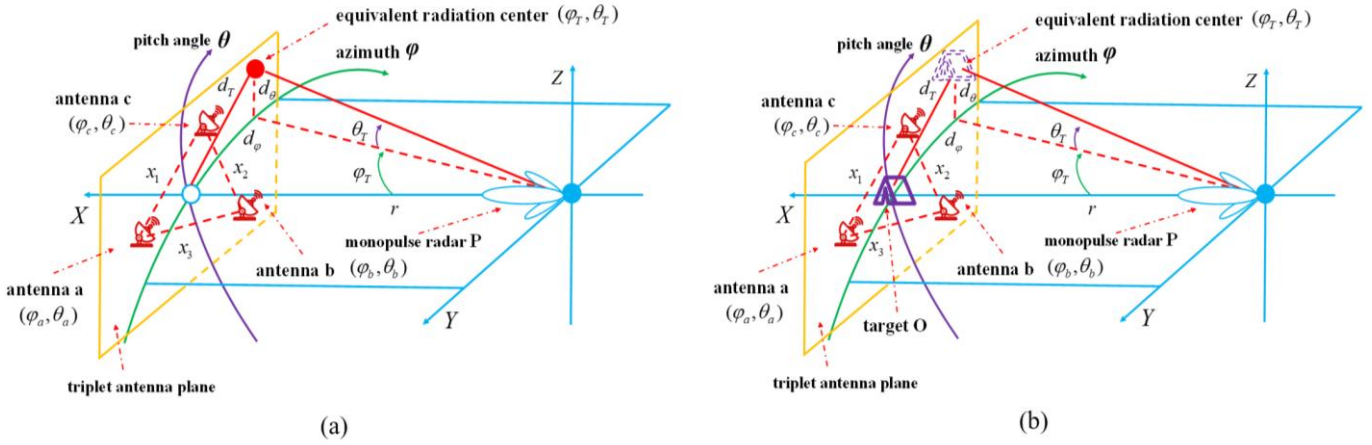

Fig. 1: The spherical coordinate system of MPSJ system. (a) No target echo case. (b) Target echo case.

$$\begin{cases} S_x = -[1+\alpha^2+\beta^2+2\alpha\cos\phi_{ab}+2\beta\cos\phi_{ac}+2\alpha\beta\cos(\phi_{ab}-\phi_{ac})]/(2\eta) \\ S_y = [\varphi_a+\alpha^2\varphi_b+\beta^2\varphi_c+\alpha(\varphi_a+\varphi_b)\cos\phi_{ab}+\beta(\varphi_a+\varphi_c)\cos\phi_{ac}+\alpha\beta(\varphi_b+\varphi_c)\cos(\phi_{ab}-\phi_{ac})]/(2\eta) \\ S_z = -[\theta_a(1+\alpha\cos\phi_{ab}+\beta\cos\phi_{ac})+\theta_b(\alpha^2+\alpha\cos\phi_{ab}+\alpha\beta\cos(\phi_{ab}-\phi_{ac})+\theta_c(\beta^2+\beta\cos\phi_{ac}+\alpha\beta\cos(\phi_{ab}-\phi_{ac}))]/(2\eta) \end{cases} \tag{2}$$

$$\begin{cases} S_x = -[1+\alpha^2+\beta^2+\gamma^2+2\alpha\cos\phi_{ab}+2\beta\cos\phi_{ac}+2\gamma\cos\phi_{ao}+2\alpha\beta\cos(\phi_{ab}-\phi_{ac})+2\alpha\gamma\cos(\phi_{ab}-\phi_{ao}) \\ \quad +2\beta\gamma\cos(\phi_{ac}-\phi_{ao})]/(2\eta) \\ S_y = [\varphi_a+\alpha^2\varphi_b+\beta^2\varphi_c+\alpha(\varphi_a+\varphi_b)\cos\phi_{ab}+\beta(\varphi_a+\varphi_c)\cos\phi_{ac}+\gamma(\varphi_a+\varphi_o)\cos\phi_{ao}+\alpha\beta(\varphi_b+\varphi_c)\cos(\phi_{ab}-\phi_{ac}) \\ \quad +\alpha\gamma(\varphi_b+\varphi_o)\cos(\phi_{ab}-\phi_{ao})+\beta\gamma(\varphi_c+\varphi_o)\cos(\phi_{ac}-\phi_{ao})]/(2\eta) \\ S_z = -[\theta_a(1+\alpha\cos\phi_{ab}+\beta\cos\phi_{ac}+\gamma\cos\phi_{ao})+\theta_b(\alpha^2+\alpha\cos\phi_{ab}+\alpha\beta\cos(\phi_{ab}-\phi_{ac}) \\ \quad +\alpha\gamma\cos(\phi_{ab}-\phi_{ao}))+\theta_c(\beta^2+\beta\cos\phi_{ac}+\alpha\beta\cos(\phi_{ab}-\phi_{ac})+\beta\gamma\cos(\phi_{ac}-\phi_{ao}))]/(2\eta) \end{cases} \tag{3}$$

$$\begin{bmatrix} \varphi_{Tm} \\ \theta_{Tm} \end{bmatrix} = \frac{\begin{bmatrix} \varphi_a & \varphi_b & \varphi_c \\ \theta_a & \theta_b & \theta_c \end{bmatrix} \begin{bmatrix} 1+\alpha\cos\phi_{ab}+\beta\cos\phi_{ac} \\ \alpha^2+\alpha\cos\phi_{ab}+\alpha\beta\cos(\phi_{ab}-\phi_{ac}) \\ \beta^2+\beta\cos\phi_{ac}+\alpha\beta\cos(\phi_{ab}-\phi_{ac}) \end{bmatrix}}{(1+\alpha^2+\beta^2+2\alpha\cos\phi_{ab}+2\beta\cos\phi_{ac}+2\alpha\beta\cos(\phi_{ab}-\phi_{ac})+1/\mathrm{JSR})} \tag{5}$$

center P of the sphere is given. For detailed derivation, please refer to the appendix.

where α and β are amplitude ratios: $\alpha = E_b / E_a$, $\beta = E_c / E_a$; ϕ_{ab} and ϕ_{ac} are phase difference: $\phi_{ab} = \phi_b - \phi_a$, $\phi_{ac} = \phi_c - \phi_a$. E_i denotes the amplitude of the signal arriving at the sphere center P, and ϕ_i represents the phase of the signal arriving at the sphere center P; η is the impedance in free space.

Similarly, in the presence of target echo, the average energy flux densities at the sphere center P in the three directions of vectors x, y, and z are in Eq. (3):

where γ is the amplitude ratio: $\gamma = E_o / E_a$; ϕ_{ao} is the phase difference $\phi_{ao} = \phi_o - \phi_a$.

Based on Eq. (4), the three directional components in the cartesian coordinate system are transformed back into the spherical coordinate system, thereby deriving the position coordinate (φ_T, θ_T) of the EFT.[25]

$$\begin{cases} \varphi_T = \arctan(-\dfrac{S_y}{S_x}) \approx -\dfrac{S_y}{S_x} \\ \theta_T = \arctan(-\dfrac{S_z}{\sqrt{S_x^2+S_y^2}}) \approx -\dfrac{S_z}{\sqrt{S_x^2+S_y^2}} \end{cases} \tag{4}$$

In practical application environments, the phase of target echo is unpredictable, so it is assumed that its phase follows a uniform distribution within the interval $[0, 2\pi]$. Based on this assumption, the coordinates of the EFT are not fixed but determined by the specific value of the target echo phase. According to the simulation analysis results in the relevant reference,[10] the median variation $(\varphi_{Tm}, \theta_{Tm})$ trend of the EFT is more stable relative to its maximum and minimum values, demonstrating high stability. Therefore, this paper selects the median of the EFT as the estimated value of its coordinates, specifically as given in Eq. (5):

where JSR is jamming-to-signal ratio: $\mathrm{JSR} = E_a^2 / E_o^2 = 1/\gamma^2$.

2.3 MPSJ system equivalent false target generation model

The positions of the above-mentioned MPSJ system EFTs are calculated based on theoretical formulas. However, in practical applications, due to the presence of errors, there will be certain deviations between the actual positions of the EFTs and their theoretical positions. In other words, when considering error factors, it is extremely challenging to achieve jamming on precise positions. But in actual jamming operations, what is usually concerned is not the specific position where the EFT appears, but whether this position is far away from the triplet antenna, whether it can confuse the enemy's radar, and thus achieve the purpose of protecting the target.

In practical applications, the effect of phase error is generally more significant than that of amplitude error. As a result, existing EFT generation methods have become difficult to apply to realize jamming at specific positions. Such methods must account for certain phase errors and find amplitude and phase parameters that can consistently locate the positions of EFTs within the constraint range.

To illustrate the feasibility of the MPSJ system EFT generation method, this section explains it by means of the parameter tolerance of the triplet antenna. The contour map of the angle factor can intuitively present the parameter tolerance situation of the triple antenna. The expression of the angle factor is as given in Eq. (6).[26]

$$AF_T = \left| \sqrt{\varphi_{\mathrm{Tm}}^2 + \theta_{\mathrm{Tm}}^2} \,/\, \max(\varphi_i, \theta_i) \right|, i = a, b, c, o \tag{6}$$

As shown in Fig. 2, Fig. 2(a) depicts the contour map of the triplet antenna angle factor when α =0dB, β =0dB in the absence of target echo; while Fig. 2(b) illustrates the contour map of the triplet antenna angle factor when α =0dB, β =0dB, and jamming-to-signal ratio (JSR) = 30 dB in the presence of target echo. It can be observed from the figures that taking the constraint range $[2d_{\max}, 5d_{\max}]$ (the closest distance $2d_{\max}$ and the farthest distance $5d_{\max}$ from the EFT position to the triplet antenna center, $d_{\max}$ is the maximum baseline length between triple) as an example, even in the presence of phase error, amplitude and phase parameters (represented by yellow dots in the figures) can be found such that the corresponding EFT positions still locate within the constraint range under the influence of current error. Therefore, the MPSJ system EFT generation method is feasible. For the convenience of subsequent elaboration, the following two cases are denoted by abbreviations respectively: the MPSJ system EFT generation model without target echo (denoted as case1); the MPSJ system EFT generation model with target echo (denoted as case2).

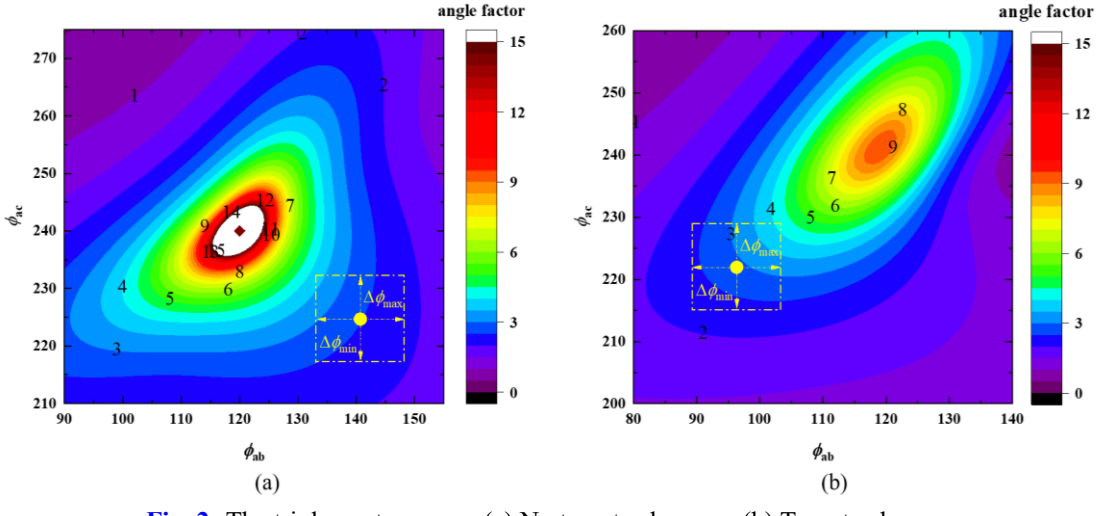

Fig. 2: The triple contour map. (a) No target echo case. (b) Target echo case.

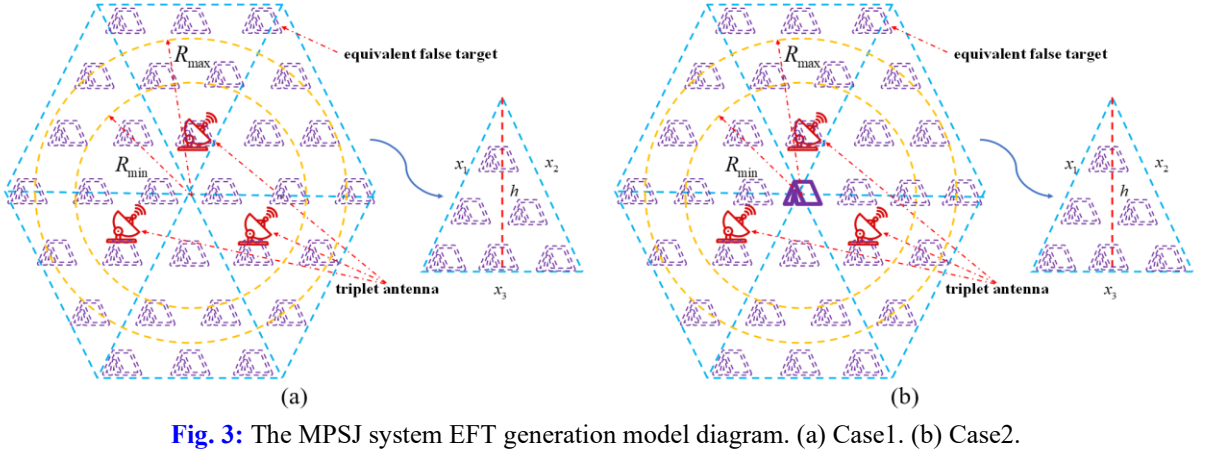

Fig. 3: The MPSJ system EFT generation model diagram. (a) Case1. (b) Case2.

To facilitate the construction of the MPSJ system EFT generation model, this section defines a constraint function, as specifically given in Eq. (7).

$$M_i(\boldsymbol{E}^{'},\boldsymbol{\phi}^{'})=\frac{1}{m}\sum_{k=1}^{m}\begin{cases}1, R_{\min}\le F(\boldsymbol{E}^{'},\boldsymbol{\phi}^{'})\le R_{\max}\\ 0, else\end{cases}, i=1,...,n \quad (7)$$

where $\boldsymbol{E}^{'}$ is a vector composed of antenna amplitude ratios, $\boldsymbol{E}^{'}=[\alpha+\Delta\alpha,\beta+\Delta\beta]$; $\boldsymbol{\phi}^{'}$ is a vector composed of antenna phase differences, $\boldsymbol{\phi}^{'}=[\phi_{ab}+\Delta\phi_{ab},\phi_{ac}+\Delta\phi_{ac}]$; $\Delta\alpha$ and $\Delta\beta$ are amplitude ratio errors respectively; $\Delta\phi_{ab}$ and $\Delta\phi_{ac}$ are phase difference errors respectively; $F(\boldsymbol{E}^{'},\boldsymbol{\phi}^{'})$ denotes substituting $\boldsymbol{E}^{'}$ and $\boldsymbol{\phi}^{'}$ into Eqs. (1), (3), (4) and (5) to calculate the distance from the EFT to the center of the triplet antenna or the target; $R_{\min}$ and $R_{\max}$ represent the closest and farthest distances from the constraint range to the center of the triplet antenna or the target respectively; m is the number of amplitude-phase combinations within the device error range; and n is the number of all amplitude-phase combinations.

From the above analysis, it is clear that in the absence of target echo, the MPSJ system EFT generation model for the MPSJ system takes a certain degree of phase error into account and is committed to finding amplitude and phase parameters to ensure that the position of the EFT can still locate within the established constraint range under the influence of the current error. Correspondingly, the MPSJ system EFT generation model without target echo is presented in Fig. 3(a). Similarly, the MPSJ system EFT generation model with target echo is shown in Fig. 3(b). The general mathematical model of the MPSJ system EFT generation models corresponding to the above two different cases is described in Eq. (8).

$$\max_{\boldsymbol{E}^{'},\boldsymbol{\phi}^{'}} M_i(\boldsymbol{E}^{'},\boldsymbol{\phi}^{'}), i=1,...,n$$
$$\text{s.t.}\begin{cases}\alpha_{\min}\le\alpha\le\alpha_{\max},\beta_{\min}\le\beta\le\beta_{\max}\\ \phi_{abmin}\le\phi_{ab}\le\phi_{abmax},\phi_{acmin}\le\phi_{ac}\le\phi_{acmax}\\ \Delta\alpha=0,\Delta\beta=0,n\ge1\\ \Delta\phi_{abmin}\le\Delta\phi_{ab}\le\Delta\phi_{abmax},\Delta\phi_{acmin}\le\Delta\phi_{ac}\le\Delta\phi_{acmax}\end{cases} \quad (8)$$

2.4 Q-Learning algorithm

The problem of finding optimal parameters in the aforementioned dynamic process necessitates continuous adjustment of current parameters upon receiving feedback, which precisely aligns with the concept of reinforcement learning that adjusts its actions through continuous interaction with the environment.

2.4.1 Basic Q-learning algorithm

The Q-learning (QL) algorithm, a type of reinforcement learning algorithm, primarily consists of elements including the agent, states, actions, rewards, and the environment.[27-29] The algorithm achieves self-learning through continuous information exchange between the agent and the environment. In the QL algorithm, the Q-table reflects the result of interactions between the agent and the environment, so the process of updating the Q-table essentially constitutes the interactive process between the agent and the environment.[30] When the agent is in the current state s, selecting an action δ, after being affected by the environment, it will transition to a new state $s^{'}$ and obtain a corresponding reward $r(s)$. After the agent updates the Q-table via the QL algorithm, if the value of $Q(s,\delta)$ decreases, this indicates that choosing the action at the current position is not the optimal selection. When the agent returns to the same position or state next time, it can avoid repeating the action. By continuously repeating this process, the agent interacts with the environment to accumulate massive data until the Q-table converges. The QL algorithm uses the obtained data to correct its action strategy, then continues to interact with the environment to acquire new data, which are further used to optimize the strategy.[31,32] After multiple iterations, the agent can ultimately obtain the optimal action.

The Q-table update formula is in Eq. (9):

$$Q(s,\delta)=(1-\lambda)*Q(s,\delta)+\lambda*(r(s)+\mu*(\max_{a}Q(s^{'},\delta^{'}))-Q(s,\delta)) \quad (9)$$

where the set of state is shown in Eq. (10):

$$S=\{s_1,...,s_i,...,s_u\}$$
$$s_i=F(\boldsymbol{E}^{'},\boldsymbol{\phi}^{'}) \quad (10)$$

The set of action is in Eq. (11):

$$A=\{\delta_1,...,\delta_i,...,\delta_v\}$$
$$\delta_i=[\Lambda\alpha,\Lambda\beta,\Lambda\phi_{ab},\Lambda\phi_{ac}] \quad (11)$$

where the reward $r(s)$ denotes the reward value in the current state s; λ is the learning rate, $\lambda\in[0,1]$. When $\lambda=0$, it indicates that the agent only learns from past states, while when $\lambda=1$, the agent can only learn from received information. μ is the discount factor, $\mu\in[0,1]$. When $\mu=1$, the agent can learn all future rewards, whereas when $\mu=0$, the agent can only accept immediate rewards at present. u and v denote the number of states and actions, respectively. $\Lambda\alpha$, $\Lambda\beta$, $\Lambda\phi_{ab}$ and $\Lambda\phi_{ac}$ denote the variation of α, β, ϕ_{ab} and ϕ_{ac} after the action is taken, respectively.

2.4.2 Double dynamic incentive Q-learning algorithm

In the basic QL algorithm, the Q-table is typically initialized to 0, leading to blind action selection during the early exploration phase.[33] By setting an inductive Q-table in the initial stage, the agent can be induced to learn toward optimal actions, accelerating the algorithm's search for optimal amplitude and phase parameters. Therefore, designing a Q-table with the maximum value at the center of the constraint range and gradually decreasing values toward the constraint boundary can induce the agent to approach the constraint range rapidly. This design leverages the gradient of Q-values

to guide the learning process, ensuring that the agent prioritizes parameter spaces within the constraint boundary while minimizing unnecessary exploration of suboptimal regions. The inductive Q-table is in Eq. (12):

$$Q=\begin{bmatrix} s_1 & s_1 & \cdots & s_1 \\ \vdots & \vdots & \vdots & \vdots \\ s_p & s_p & \cdots & s_p \\ \vdots & \vdots & \vdots & \vdots \\ s_{\text{mean}} & s_{\text{mean}} & \cdots & s_{\text{mean}} \\ \vdots & \vdots & \vdots & \vdots \\ s_q & s_q & s_q & s_q \\ \vdots & \vdots & \vdots & \vdots \\ s_u & s_u & s_u & s_u \end{bmatrix} \tag{12}$$

where s_p is the state corresponding to $R_{\min}$, s_q is the state corresponding to $R_{\max}$, and s_{mean} is the state corresponding to $(R_{\min}+R_{\max})/2$. The values corresponding to each state are using Eq. (13):

$$s_i=\begin{cases} 0.5*(i-1)/(p-1), 1\le i<p \\ (i-p)/(q-p)+0.5, p\le i<(p+q)/2 \\ 1-(i-p/2-q/2)/(q-p), (p+q)/2\le i<q \\ 0.5*(u-i)/(u-q), q\le i\le u \end{cases} \tag{13}$$

In the basic QL algorithm, the exploration strategy employs a greedy policy.[34] When the exploration factor ε approaches 1, the agent tends to attempt random actions. However, if the agent always acts randomly, it is not suitable for searching optimal parameters. When ε approaches 0, the agent tends to exploit the external environment and select the action with the maximum action-value function. In this case, the algorithm may fail to converge effectively. Therefore, the selection of ε is of vital importance. To address this issue, a dynamic exploration factor is adopted in Eq. (14).

$$\varepsilon=\exp(-\varepsilon_0*t) \tag{14}$$

where t is the number of episodes; $\varepsilon_0=1$. As the number of episodes increases, the proportion of the exploration factor gradually decreases, enabling the agent to make more use of the external environment in the next selection to improve the convergence speed.

The design of the reward function directly influences the learning direction of the agent, and an irrational reward function may lead to low exploration efficiency of the algorithm. To improve the convergence ability of the algorithm, it is necessary to set up a reasonable reward function according to the objectives. To better construct the dual-objective reward function, the constraint metric function is first defined. The constraint metric function reflects the proportion of EFT positions corresponding to amplitude and phase parameters within the constraint range after introducing phase errors in Eq. (15).

$$G=M(\boldsymbol{E}',\boldsymbol{\phi}')/m \tag{15}$$

Subsequently, a dual-objective reward function is constructed, which consists of two parts. The first part is the distance reward, namely the distance between the EFT position corresponding to the current amplitude and phase parameters and the boundary of the constraint range. The smaller the distance, the larger the reward value. The second part is the constraint reward, that is, the proportion of EFT positions corresponding to the current amplitude and phase parameters falling within the constraint range after considering phase errors. The higher the proportion, the higher the reward value. These two parts of rewards can not only accelerate the speed at which the algorithm converges to the optimal parameters but also ensure that the generated false target positions are all within the constraint range (Eq. (16)).

$$r(s)=\begin{cases} r^*-\omega*(1-G), R_{\min}\le s\le R_{\max} \\ -(s-s_{\text{mean}})^2, else \end{cases} \tag{16}$$

where r^* denotes the positive reward for amplitude and phase parameters within the constraint range, which is a positive number; ω is the balancing factor, balancing the distance reward and the constraint reward to guide the agent to take actions that achieve dual objectives.

3. Simulation experiment and result analysis

3.1 Simulation experimental environment and parameters

The simulation experiments are conducted on a Windows 11 (64-bit) operating system, with hardware platform parameters including an Intel(R) Core(TM) i9-13900K CPU, a main frequency of 3.00 GHz, and 16 GB of RAM. The computing environment is MATLAB R2024a.

The parameters involved in all the simulation experiments in this paper are shown in Table 1. The triplet antenna array is arranged in a regular triangle pattern. The states range from 0.1 $d_{\max}$ to 10 $d_{\max}$ at intervals of 0.1 $d_{\max}$. The action amplitude ratios have variation amounts of 0.1 dB and -0.1 dB, while the phase difference variation amounts are 1° and -1°, forming a total of 16 actions.

Table 1: The simulation experiment parameters.

Parameter	Value	Parameter	Value
$d_{\max}$	0.2km	Episodes	2000
Maximum step	1000	JSR	15.71dB
r^*	10	ω	10
$\alpha_{\min}$	-1.5dB	$\alpha_{\max}$	1.5dB
$\beta_{\min}$	-1.5dB	$\beta_{\max}$	1.5dB
ϕ_{abmin}	0°	ϕ_{abmax}	360°
ϕ_{acmin}	0°	ϕ_{acmax}	360°

3.2 Generation of equivalent false target

Without loss of generality, the phase error in this section's experiment is set to [−5°, 5°], while $R_{min} = 2d_{max}$ and $R_{max} = 5d_{max}$. The setting of phase error is consistent with the typical error range in actual situations, and the setting of constraint radius is also consistent with the general situation in actual scenarios. In actual scenarios, the constraint radius corresponds to 400m-1000m, which is a reasonable protection radius.[35-37] Additionally, within the error range, the position variation of the EFT is investigated with a step of 1°, which is sufficient to meet the requirements. Based on this, the precision requirements in the algorithm are specified: the initialization, iteration process, and output precision of optimal parameters are all set to 1°.

To enhance the algorithm's insensitivity to initial parameters, the parameters are randomly initialized during the training process. The specific iterative results are shown in the Fig. 4(a). And the x-axis represents the episode of the algorithm, and the y-axis is the reward obtained by the algorithm during the iteration process. The larger the reward, the better the algorithm's performance. It can be observed that the algorithm converges to the optimal value of 9.5 in both cases, with the difference lying in the number of episodes required for convergence, which is attributed to the varying complexity of the models under the two cases. The model with target echo is more complex than that without target echo.

Additionally, the specific results of the distances of synthetic false targets corresponding to amplitude-phase combinations within the phase error range are presented in Fig. 4(b), where points of two different colors represent EFT distance points under different conditions. Moreover, the positions of EFTs generated by the MPSJ system EFT generation method and those under phase error conditions are shown in Fig. 5 for both cases, with the red dashed line indicating the constraint range. It is evident that all EFT distances fall within the range [$2d_{max}$, $5d_{max}$].

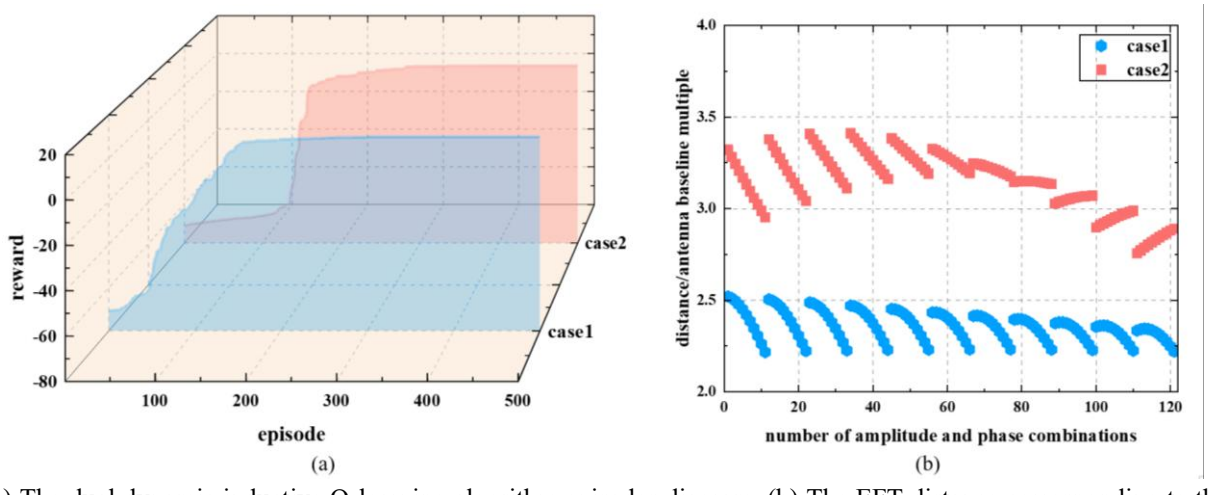

Fig. 4: (a) The dual dynamic inductive Q-learning algorithm episodes diagram. (b) The EFT distance corresponding to the solved amplitude and phase parameters in the case of phase error.

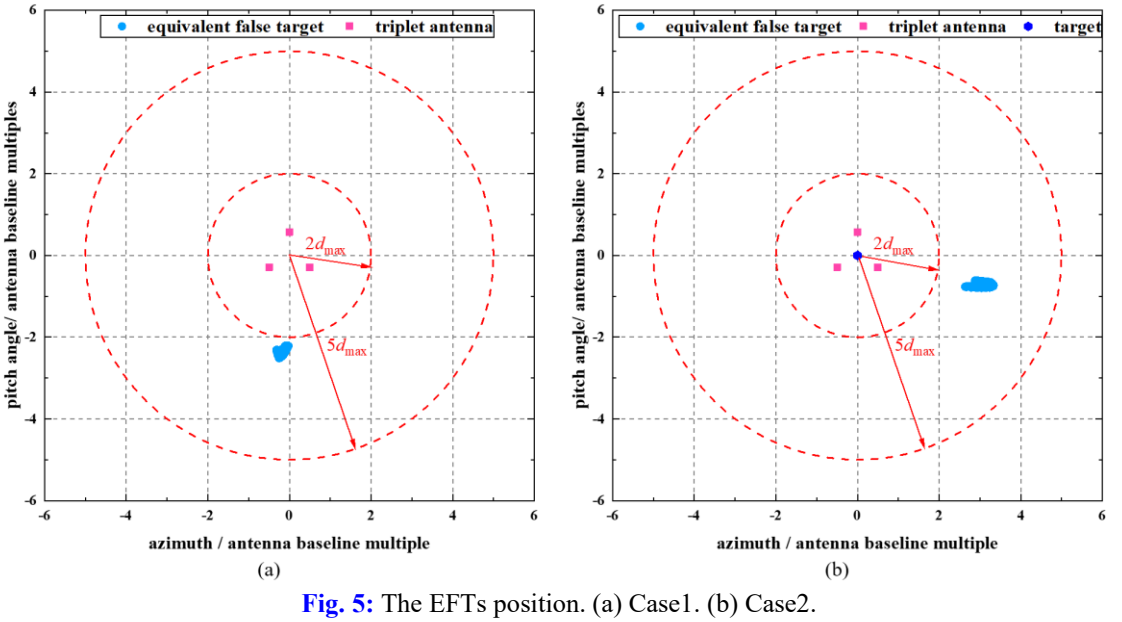

Fig. 5: The EFTs position. (a) Case1. (b) Case2.

3.3 The influence of model parameters

3.3.1 The influence of phase error

In the experiment of this section, the constraint range [$2d_{max}$, $5d_{max}$] is kept unchanged, and the influence of phase error on the MPSJ system EFT generation method is studied by gradually increasing the phase error with a step of 1°. The specific results are shown in Table 2 and Fig. 6.

According to the data in Table 2, within the phase error range of 9°, the positions of all EFTs corresponding to amplitude and phase parameters are within the constraint range. Moreover, as the phase error increases, the constraint metric gradually decreases. When the phase error reaches 12°, the constraint metric in both cases are less than 0.9, indicating that within the 12° phase error range, the EFT positions corresponding to more than 10% of the amplitude-phase combinations exceed the [$2d_{max}$, $5d_{max}$] range.

Table 2: The specific results of the impact of phase error on the MPSJ system.

	Phase error	Constraint metric		Phase error	Constraint metric
Case1	5°	1	Case2	5°	1
	6°	1		6°	1
	7°	1		7°	1
	8°	1		8°	1
	9°	1		9°	1
	10°	0.995		10°	0.993
	11°	0.964		11°	0.954
	12°	0.898		12°	0.877

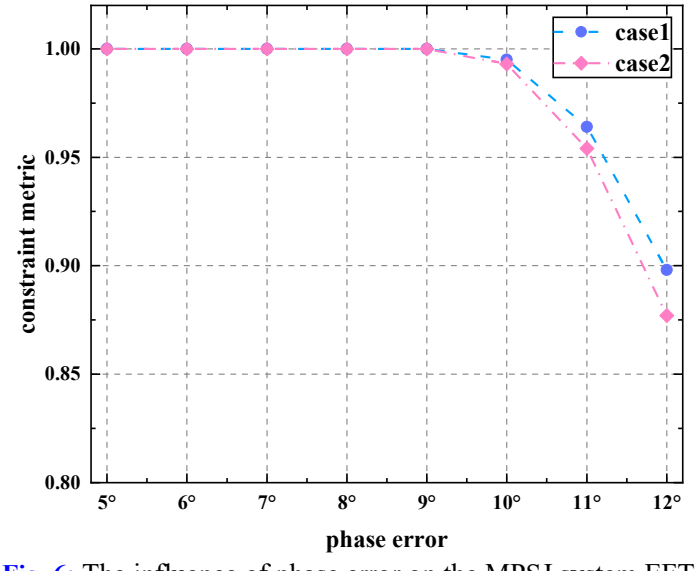

Fig. 6: The influence of phase error on the MPSJ system EFT generation method.

3.3.2 The influence of constraint range size

In this section, while keeping the phase error constant, the influence of the constraint range size on the MPSJ system EFT generation method is investigated by gradually reducing the constraint range at intervals of d_{max} . The specific results are presented in Table 3 and Fig. 7. In Fig. 7, R1, R2 and R3 represent [$2d_{max}$, $5d_{max}$], [$2d_{max}$, $4d_{max}$] and [$2d_{max}$, $3d_{max}$], respectively.

Based on the data in Table 3, within the phase error range of 5°-8°, the constraint metric remains unchanged as the constraint range decreases. Conversely, within the phase error range of 9°-12°, the constraint metric gradually declines with the reduction of the constraint range. This indicates that under larger phase errors, the size of the constraint range significantly affects the MPSJ system EFT generation method, thereby influencing the number of amplitude-phase combinations corresponding to EFTs within the constraint range. When the phase error exceeds 10° and the radius of the constraint range is reduced to the baseline length of the MPSJ system, the positions of EFTs corresponding to more than 10% of the amplitude and phase parameters fall outside the constraint range.

Overall, the MPSJ system EFT generation method is highly sensitive to phase error. Under a phase error of 10°, no combination of amplitude and phase parameters can ensure that the positions of EFTs remain within the constraint range under the influence of the current error. Additionally, the method is also affected by the size of the constraint range. Within the phase error range of 9°-12°, the constraint metric gradually decreases as the constraint range shrinks, indicating a reduction in the number of amplitude-phase combinations corresponding to EFTs within the constraint range.

4. Discussion

4.1 The impact of array arrangement

4.1.1 Array arrangement in the absence of target echo

This section demonstrates the distribution of EFTs in the absence of target echo through simulations, so as to explain the basis for choosing the regular triangle array arrangement. In the absence of target echo, according to Eq. (3) and Eq. (5), to maintain generality, the amplitude ratio of the triplet antenna is set to freely take values in the interval [-1.5dB, 1.5 dB] with a step of 0.1 dB, and the phase difference is set to freely take values in the interval [0, 360°] with a step of 1°. Through the above settings, the complete plane distribution diagram of the triplet antenna EFTs in the case of no target echo can be simulated, as shown in Fig. 8(a). The triplet antenna array configuration is an arbitrary triangular configuration.

As can be observed from Fig. 8(a), the distribution of EFTs in an arbitrary triangular array configuration exhibits a general hexagonal shape. Given that the constraint range defined in this paper is a circular ring region, the regular triangular array configuration is more suitable. The EFTs corresponding to the

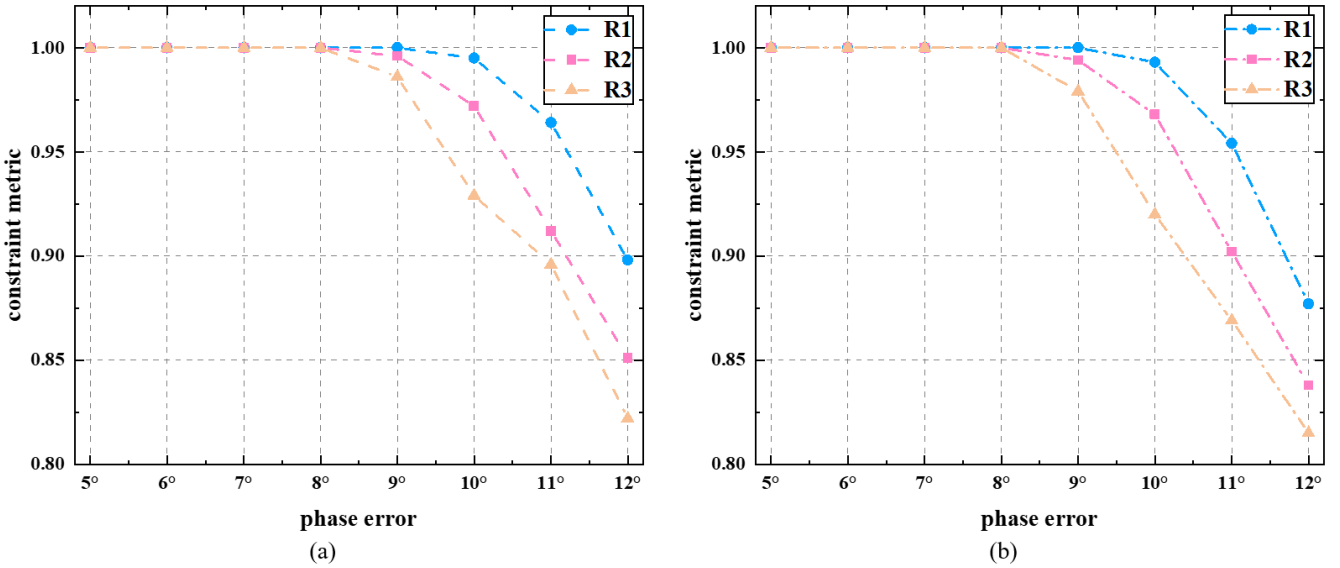

Fig. 7: The influence of constraint range size on the MPSJ system EFT generation method. (a) Case1. (b) Case2.

Table 3: The specific results of the influence of the size of the constraint range on the MPSJ system.

	Phase error	Constraint range	Constraint metric		Phase error	Constraint range	Constraint metric
Case1	5°	$[2d_{max}, 5d_{max}]$	1	Case2	5°	$[2d_{max}, 5d_{max}]$	1
	5°	$[2d_{max}, 4d_{max}]$	1		5°	$[2d_{max}, 4d_{max}]$	1
	5°	$[2d_{max}, 3d_{max}]$	1		5°	$[2d_{max}, 3d_{max}]$	1
	6°	$[2d_{max}, 5d_{max}]$	1		6°	$[2d_{max}, 5d_{max}]$	1
	6°	$[2d_{max}, 4d_{max}]$	1		6°	$[2d_{max}, 4d_{max}]$	1
	6°	$[2d_{max}, 3d_{max}]$	1		6°	$[2d_{max}, 3d_{max}]$	1
	7°	$[2d_{max}, 5d_{max}]$	1		7°	$[2d_{max}, 5d_{max}]$	1
	7°	$[2d_{max}, 4d_{max}]$	1		7°	$[2d_{max}, 4d_{max}]$	1
	7°	$[2d_{max}, 3d_{max}]$	1		7°	$[2d_{max}, 3d_{max}]$	1
	8°	$[2d_{max}, 5d_{max}]$	1		8°	$[2d_{max}, 5d_{max}]$	1
	8°	$[2d_{max}, 4d_{max}]$	1		8°	$[2d_{max}, 4d_{max}]$	1
	8°	$[2d_{max}, 3d_{max}]$	1		8°	$[2d_{max}, 3d_{max}]$	1
	9°	$[2d_{max}, 5d_{max}]$	1		9°	$[2d_{max}, 5d_{max}]$	1
	9°	$[2d_{max}, 4d_{max}]$	0.996		9°	$[2d_{max}, 4d_{max}]$	0.994
	9°	$[2d_{max}, 3d_{max}]$	0.986		9°	$[2d_{max}, 3d_{max}]$	0.979
	10°	$[2d_{max}, 5d_{max}]$	0.995		10°	$[2d_{max}, 5d_{max}]$	0.993
	10°	$[2d_{max}, 4d_{max}]$	0.972		10°	$[2d_{max}, 4d_{max}]$	0.968
	10°	$[2d_{max}, 3d_{max}]$	0.929		10°	$[2d_{max}, 3d_{max}]$	0.920
	11°	$[2d_{max}, 5d_{max}]$	0.964		11°	$[2d_{max}, 5d_{max}]$	0.954
	11°	$[2d_{max}, 4d_{max}]$	0.912		11°	$[2d_{max}, 4d_{max}]$	0.902
	11°	$[2d_{max}, 3d_{max}]$	0.896		11°	$[2d_{max}, 3d_{max}]$	0.869
	12°	$[2d_{max}, 5d_{max}]$	0.898		12°	$[2d_{max}, 5d_{max}]$	0.877
	12°	$[2d_{max}, 4d_{max}]$	0.851		12°	$[2d_{max}, 4d_{max}]$	0.838
	12°	$[2d_{max}, 3d_{max}]$	0.822		12°	$[2d_{max}, 3d_{max}]$	0.815

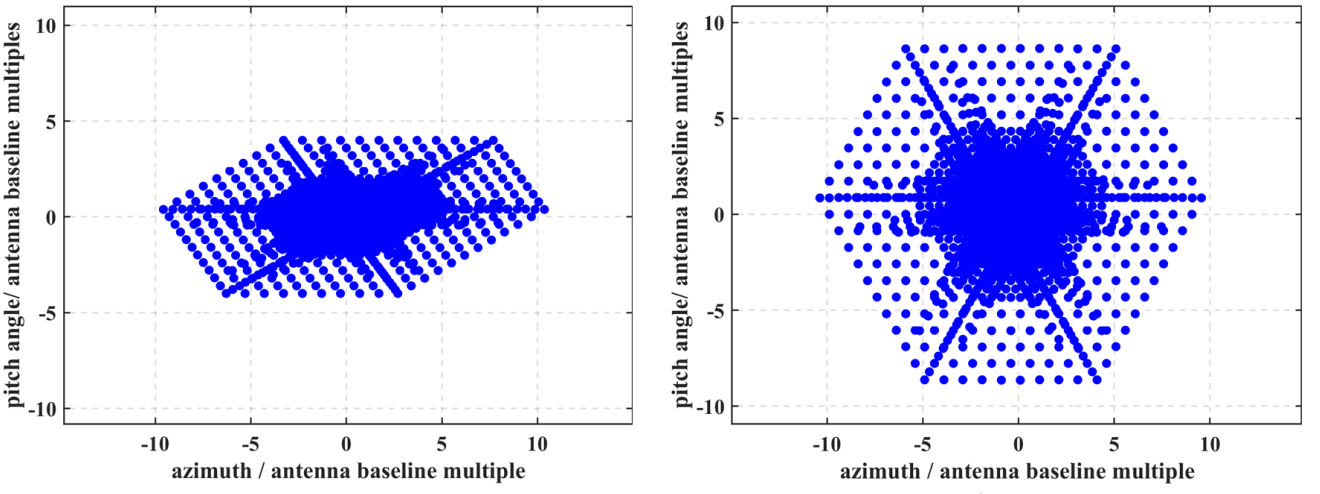

Fig. 8: The EFT distribution in the absence of target echo (a) arbitrary triangle (b) regular triangle.

regular triangular array are distributed in a regular hexagon, as illustrated in Fig. 8(b). This distribution pattern can conform to the set constraint range, thereby enabling the MPSJ system EFT generation method to more efficiently locate the optimal parameters that satisfy the conditions.

4.1.2 Array arrangement in the presence of target echo

In the presence of target echo, not only the triplet antenna array configuration but also the influence of JSR must be considered. This section analyzes the trend of the angular factor varying with JSR under different parameter conditions through simulations. As shown in Fig. 9, each curve in Fig. 9(a) represents the results of different parameters for β and ϕ_{ac} under the conditions of α = 0.5 dB and ϕ_{ab} = 115°; each curve in Fig. 9(b) represents the results of different parameters for α and ϕ_{ab} under the conditions of β = 0.5 dB and ϕ_{ac} = 235°.

The simulation results indicate that the angular factor of the triplet antenna exhibits a monotonically increasing trend as JSR increases. This phenomenon arises because, as JSR increases, the proportion of the signal emitted by the triplet antenna in the total echo gradually increases, thereby producing a stronger jamming effect on the radar. When the JSR reaches 10 dB, the amplitude of the total angular factor of the triplet antenna exceeds 1, indicating that the positions of EFTs generated by the triplet antenna have exceeded the internal range of the triangle. Additionally, when the JSR increases to 30 dB, the amplitude of the angular factor approaches saturation. Therefore, to ensure the effectiveness of triplet antenna jamming, the optimal range of JSR should be controlled between 10 dB and 30 dB.

In summary, the selection of JSR significantly impacts the MPSJ system EFT generation method. A smaller JSR fails to meet the expected constraint range in the presence of target echo, while a larger JSR results in energy waste. Therefore, to

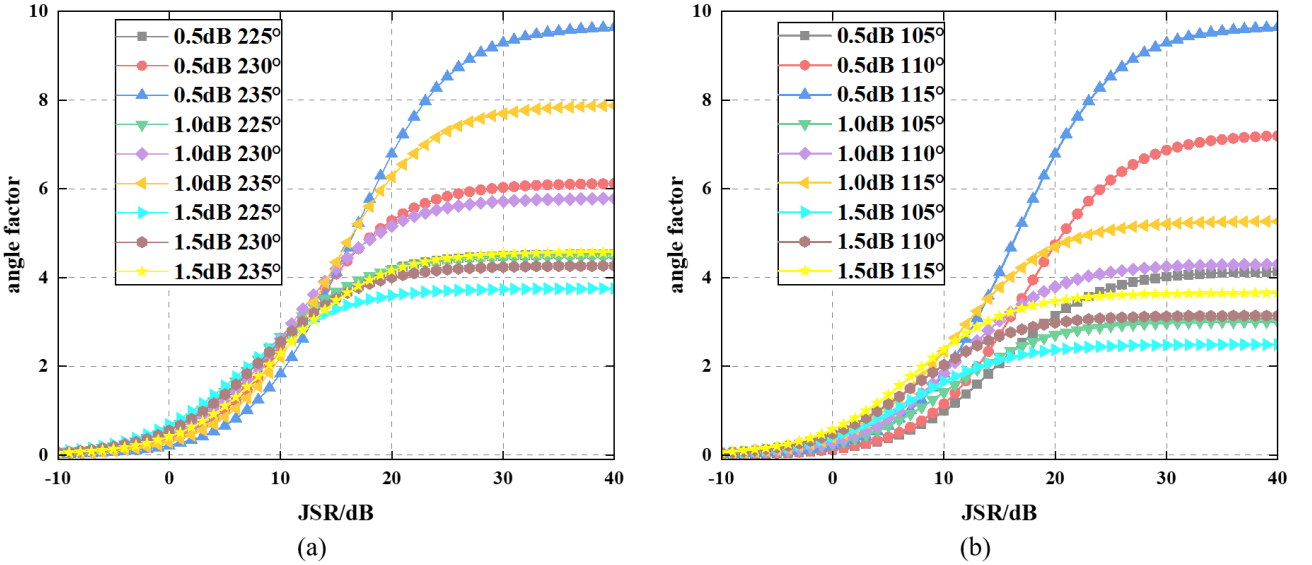

Fig. 9: The curves of angle factor with JSR under different parameters. (a) α = 0.5dB, ϕ_{ab} =115°. (b) β = 0.5dB, ϕ_{ac} =235°.

satisfy the constraint range settings in this paper, the JSR value should be set to 15.71 dB. Similarly, in the presence of target echo, the triplet antenna should adopt a regular triangular array configuration.

4.2 Performance of DDIQL algorithm

4.2.1 Convergence ability

This section compares the DDIQL algorithm with three classic reinforcement learning algorithms, namely QL, deep Q network (DQN),[38] 5-step-Sarsa,[39] and Sarsa.[40] Each experiment is independently conducted 100 times, and the convergence capability of each algorithm is evaluated through two indicators: the average episodes and the constraint metric. The episodes diagrams of each algorithm are shown in Fig. 10, and the specific experimental results are presented in Fig. 10.

The data in Table 4 demonstrates that the DDIQL algorithm has fewer average convergence times and higher constraint metric than the comparative algorithms in both scenarios, verifying the superior convergence performance of the DDIQL algorithm.

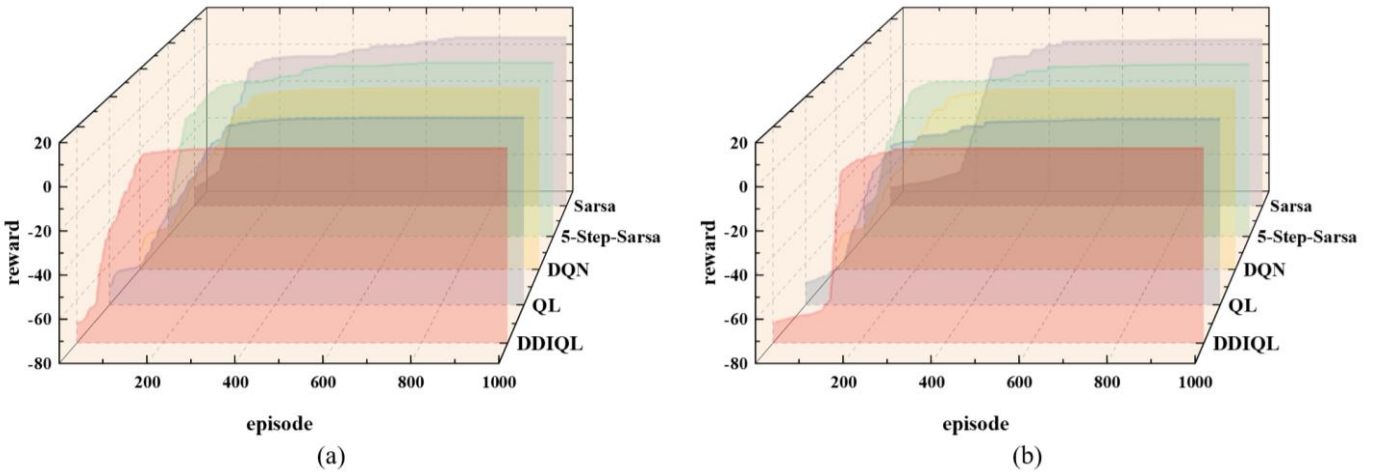

Fig. 10: The episodes diagrams of each algorithm in 2 cases. (a) Case1. (b) Case2.

Table 4: The specific results and constraint metrics of average episodes.

	Algorithm	Average episodes	Constraint metric		Algorithm	Average episodes	Constraint metric
Case1	Sarsa	700	0.427	Case2	Sarsa	900	0.412
	5-step-Sarsa	650	0.457		5-step-Sarsa	800	0.446
	DQN	630	0.612		DQN	750	0.596
	QL	600	0.518		QL	700	0.506
	DDIQL	300	1		DDIQL	400	1

Table 5: The comparison results of the effectiveness of the DDIQL algorithm improvement strategies.

	Inductive strategy	Dynamic exploration factor	Dual-objective reward function	Average episodes	Constraint metric
Case1	×	×	×	600	0.518
	×	×	√	400	1
	×	√	×	560	0.518
	×	√	√	360	1
	√	×	×	550	0.518
	√	×	√	350	1
	√	√	×	500	0.518
	√	√	√	300	1
Case2	×	×	×	700	0.506
	×	×	√	510	1
	×	√	×	665	0.506
	×	√	√	470	1
	√	×	×	660	0.506
	√	×	√	460	1
	√	√	×	610	0.506
	√	√	√	400	1

4.2.2 Effectiveness of the improvement strategy

This section conducts comparative experiments on the effectiveness of different improvement strategies for the DDIQL algorithm under two cases, with each experiment independently repeated 100 times. The effectiveness of different improvement strategies for the DDIQL algorithm is evaluated through two indicators: the average episodes and constraint metric of 100 experiments, with specific results shown in Table 5.

As can be seen from the data in Table 5, the inductive strategy can guide the algorithm to search toward optimal parameters in the early stage, thereby improving the convergence performance of the algorithm; the dynamic exploration factor can ensure that the algorithm no longer performs random searches in the later stage, thus accelerating the convergence speed; the dual-objective reward function can not only guide the algorithm to search toward optimal parameters but also enhance the constraint metric to ensure that the positions of generated false targets are all within the constraint range. In summary, the dual-objective reward function plays the most significant role in improving the convergence and constraint metric of the DDIQL algorithm.

4.3 Hyperparameter sensitivity analysis

In this section, the marginal effects of the learning rate, discount factor, and exploration factor on the algorithm performance (quantified by the episode and constraint metric) are systematically evaluated through single-factor sensitivity experiments. To eliminate interaction effects, the experiments adhere to the principle of "varying only one factor at a time while fixing the others to their verified optimal values"; each value point is independently repeated 10 times, with the mean value calculated to mitigate random fluctuations. All three hyperparameters are sampled at equal intervals of 0.05 within the range of 0.70-0.95, and the corresponding results are summarized in Table 6.

As observed in Table 6, when the learning rate and discount factor are simultaneously within the range of 0.70-0.75, the algorithm prioritizes immediate rewards at the expense of long-term stability: the constraint metric fails to reach 1, while the convergence speed is significantly enhanced. Once both parameters increase to the range of 0.80-0.95, the constraint metric maintains a high level, whereas the convergence speed decreases accordingly. In contrast, variations in the exploration factor within the range of 0.70-0.95 exert almost no impact on performance, which can be attributed to the built-in dynamic exploration mechanism of the algorithm that has effectively balanced short-term gains and long-term planning. Based on the above results, this study fixes the learning rate and discount factor at 0.80, and the exploration factor can take any value within the range of 0.70-0.95.

4.4 The selection of parameters for the reward function

This section illustrates the basis for selecting reward function parameters through the trends of average episodes and constraint metric varying with the distance constraint ratio under two scenarios, as shown in Table 7 and Fig. 11. The distance constraint ratio (DCR) is defined as in Eq. (17):

$$\mathrm{DCR} = 10 * \log 10(r^* / \omega) \tag{17}$$

According to the data in Table 7, the algorithm converges relatively slowly when the DCR is small. This is because a smaller distance reward provides weaker incentives for the agent to learn toward optimal parameters. Conversely, when the DCR is large, the constraint metric of the algorithm decreases, as an excessively large distance reward causes the agent to ignore the smaller constraint reward. Therefore, to balance fast convergence and high constraint metric, the magnitudes of the distance reward and constraint reward should be roughly equivalent.

Table 6: The influence of hyperparameters on algorithm performance.

	Range		Average episodes	Constraint metric		Average episodes	Constraint metric		Average episodes	Constraint metric
Case1	0.70	Learning rate	255	0.962	Discount factor	260	0.966	Exploration factor	300	1
	0.75		280	0.984		285	0.982		300	1
	0.80		300	1		300	1		300	1
	0.85		320	1		320	1		300	1
	0.90		345	1		340	1		300	1
	0.95		375	1		365	1		300	1
Case2	0.70		350	0.958		355	0.962		400	1
	0.75		375	0.976		380	0.978		400	1
	0.80		400	1		400	1		400	1
	0.85		430	1		425	1		400	1
	0.90		460	1		450	1		400	1
	0.95		485	1		480	1		400	1

Table 7: The trends of average episodes and constraint metrics.

	DCR/dB	Average episodes	Constraint metric		DCR/dB	Average episodes	Constraint metric
	-25	440	1		-25	550	1
	-20	390	1		-20	520	1
	-15	360	1		-15	480	1
	-10	330	1		-10	450	1
	-5	310	1		-5	420	1
Case1	0	300	1	Case2	0	400	1
	5	290	0.996		5	390	0.991
	10	270	0.983		10	380	0.976
	15	250	0.969		15	370	0.960
	20	230	0.952		20	350	0.932
	25	200	0.936		25	330	0.914

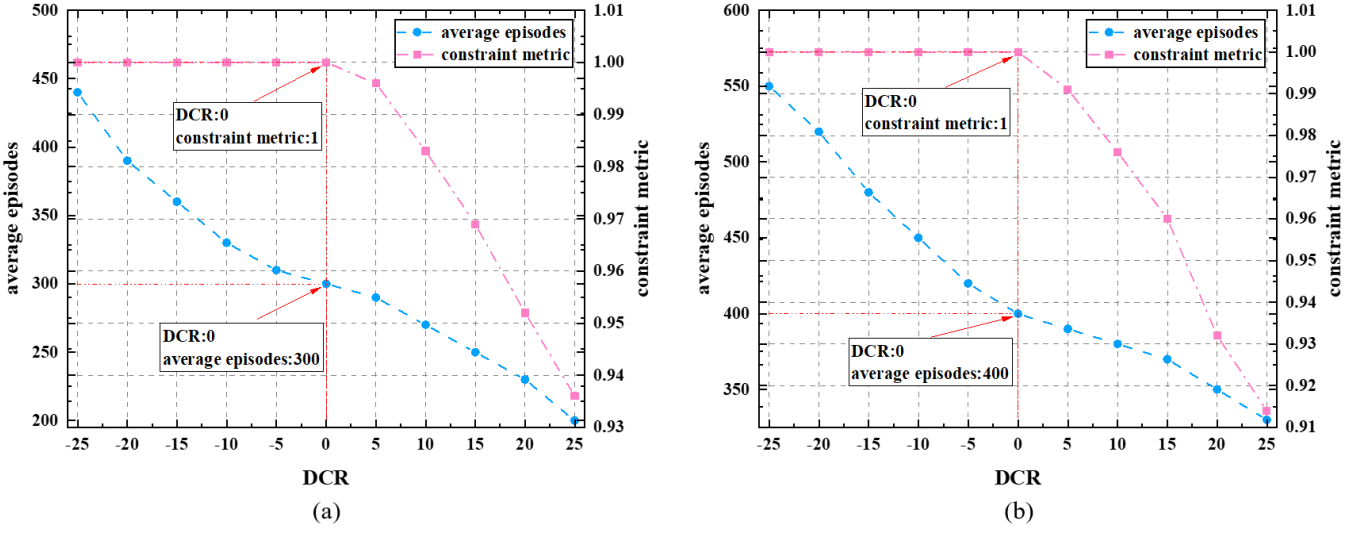

Fig. 11: The trends of average episodes and constraint metric in 2 cases. (a) Case1. (b) Case2.

5. Conclusion

To eliminate the influence of phase error in the multi-point source jamming (MPSJ) system, a Q-learning based MPSJ system equivalent false target (EFT) generation method is proposed. On the one hand, through determining the spatial geometric position relationship between the triplet antenna and the target in the MPSJ system, the distribution law of EFTs in the MPSJ system is derived. On this basis, the feasibility of the MPSJ system EFT generation method is analyzed according to the contour map of the triplet antenna, and a controllable EFT generation model with phase errors is constructed. On the other hand, by introducing a dual-objective reward function, a dynamic exploration factor, and an inductive strategy, a head-tail dynamic Inductive Q-learning algorithm is proposed to derive amplitude and phase parameters that keep the EFT positions locate within the constraint range. The influences of phase error and the size of the constraint range are discussed. Experimental results show that when the phase error is 9° and the radius of the constraint range is larger than the baseline length of the system, the positions of all EFTs to the generated amplitude and phase parameters are within the constraint range. Comparative experiments demonstrate that the proposed algorithm performs best in terms of convergence ability and constraint metric, among which the dual-objective reward function contributes the most. Furthermore, the effect of the dual-objective reward function is optimal when the magnitudes of the distance reward and the constraint reward are comparable. However, in actual scenarios, phase errors are dynamically changing. Our current DDIQL algorithm updates parameters based on batch processing of static errors, which may lead to response delays to dynamic error changes. Therefore, looking to the future, this research will continue to improve the proposed controllable EFT generation method to make it adaptable to scenarios with dynamic errors.

$$\begin{cases} S_x = -[1+\alpha^2+\beta^2+2\alpha\cos\phi_{ab}+2\beta\cos\phi_{ac}+2\alpha\beta\cos(\phi_{ab}-\phi_{ac})]/(2\eta) \\ S_y = [\varphi_a+\alpha^2\varphi_b+\beta^2\varphi_c+\alpha(\varphi_a+\varphi_b)\cos\phi_{ab}+\beta(\varphi_a+\varphi_c)\cos\phi_{ac}+\alpha\beta(\varphi_b+\varphi_c)\cos(\phi_{ab}-\phi_{ac})]/(2\eta) \\ S_z = -[\theta_a(1+\alpha\cos\phi_{ab}+\beta\cos\phi_{ac})+\theta_b(\alpha^2+\alpha\cos\phi_{ab}+\alpha\beta\cos(\phi_{ab}-\phi_{ac})+\theta_c(\beta^2+\beta\cos\phi_{ac}+\alpha\beta\cos(\phi_{ab}-\phi_{ac}))]/(2\eta) \end{cases} \tag{20}$$

$$\begin{cases} S_x = -[1+\alpha^2+\beta^2+\gamma^2+2\alpha\cos\phi_{ab}+2\beta\cos\phi_{ac}+2\gamma\cos\phi_{ao}+2\alpha\beta\cos(\phi_{ab}-\phi_{ac})+2\alpha\gamma\cos(\phi_{ab}-\phi_{ao}) \\ \quad +2\beta\gamma\cos(\phi_{ac}-\phi_{ao})]/(2\eta) \\ S_y = [\varphi_a+\alpha^2\varphi_b+\beta^2\varphi_c+\alpha(\varphi_a+\varphi_b)\cos\phi_{ab}+\beta(\varphi_a+\varphi_c)\cos\phi_{ac}+\gamma(\varphi_a+\varphi_o)\cos\phi_{ao}+\alpha\beta(\varphi_b+\varphi_c)\cos(\phi_{ab}-\phi_{ac}) \\ \quad +\alpha\gamma(\varphi_b+\varphi_o)\cos(\phi_{ab}-\phi_{ao})+\beta\gamma(\varphi_c+\varphi_o)\cos(\phi_{ac}-\phi_{ao})]/(2\eta) \\ S_z = -[\theta_a(1+\alpha\cos\phi_{ab}+\beta\cos\phi_{ac}+\gamma\cos\phi_{ao})+\theta_b(\alpha^2+\alpha\cos\phi_{ab}+\alpha\beta\cos(\phi_{ab}-\phi_{ac}) \\ \quad +\alpha\gamma\cos(\phi_{ab}-\phi_{ao}))+\theta_c(\beta^2+\beta\cos\phi_{ac}+\alpha\beta\cos(\phi_{ab}-\phi_{ac})+\beta\gamma\cos(\phi_{ac}-\phi_{ao}))]/(2\eta) \end{cases} \tag{21}$$

Acknowledgments

This work was supported by the National Natural Science Foundation of China (62471471).

Conflict of Interest

There is no conflict of interest.

Supporting Information

Not applicable.

CRediT Statement

Conceptualization, **Zhihao Cai** and **Shiqi Xing**; methodology, **Zhihao Cai** and **Shiqi Xing**; software, **Zhihao Cai**; validation, **Zhihao Cai**, **Sinong Quan**, and **Junpeng Wang**; formal analysis, **Zhihao Cai** and **Xinyuan Su**; investigation, **Zhihao Cai** and **Shiqi Xing**; resources, **Shiqi Xing**; data curation, **Zhihao Cai** and **Shiqi Xing**; writing — original draft preparation, **Zhihao Cai**; writing—review and editing, **Shiqi Xing** and **Sinong Quan**; visualization, **Zhihao Cai** and **Sinong Quan**; supervision, **Shiqi Xing**; project administration, **Shiqi Xing**; funding acquisition, **Sinong Quan**.

Appendix

The following is a detailed derivation of the coordinates of the triplet antenna EFT. Given the transformation relationship between the spherical coordinate system and the cartesian coordinate system, the vectors $\boldsymbol{\varphi}$ and $\boldsymbol{\theta}$ can be expressed by means of the unit vectors $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{z}$. On this basis, the electric field component and magnetic field component at the sphere center P can be expressed as Eq. (18).

$$\begin{cases} \boldsymbol{E}_i = E_i e^{j\phi_i}[-\sin\theta_i\cos\varphi_i\boldsymbol{x}+\sin\theta_i\sin\varphi_i\boldsymbol{y}+\cos\theta_i\boldsymbol{z}] \\ \boldsymbol{H}_i = \dfrac{E_i}{\eta}e^{j\phi_i}[\sin\varphi_i\boldsymbol{x}+\cos\varphi_i\boldsymbol{y}] \end{cases}, i=a,b,c,o \tag{18}$$

where E_i denotes the amplitude of the signal arriving at the sphere center P, and ϕ_i represents the phase of the signal arriving at the sphere center P; η is the impedance in free space.

Given that the distance between the radar and the triplet antenna plane is relatively far, the following approximate relationships can be obtained:

$$\begin{cases} \sin\theta_i \approx \theta_i, \cos\theta_i \approx 1 \\ \sin\varphi_i \approx \varphi_i, \cos\varphi_i \approx 1 \end{cases}, i=a,b,c,o \tag{19}$$

Combining Eqs. (18) and (19) and based on the calculation formula of the average Poynting vector, the average energy flux density at the sphere center P in the three directions of vectors $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{z}$ can be obtained as Eq. (20):

where α and β are amplitude ratios: $\alpha = E_b/E_a$, $\beta = E_c/E_a$; ϕ_{ab} and ϕ_{ac} are phase difference: $\phi_{ab}=\phi_b-\phi_a$, $\phi_{ac}=\phi_c-\phi_a$.

Similarly, in the presence of target echo, the average energy flux densities at the sphere center P in the three directions of vectors $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{z}$ are in Eq. (21):

where γ is the amplitude ratio: $\gamma = E_o/E_a$; ϕ_{ao} is the phase difference $\phi_{ao}=\phi_o-\phi_a$.

References

[1] J. Huang, J. Yin, M. An, Y. Li, Azimuth pointing calibration for rotating phased array radar based on ground clutter correlation, *IEEE Transactions on Geoscience and Remote Sensing*, 2024, **63**, 1000315, doi: 10.1109/TGRS.2024.3514310.

[2] W. D. Plessis, Cross-eye gain in multiloop retrodirective cross-eye jamming, *IEEE Transactions on Aerospace and Electronic Systems*, 2016, **52**, 875-882, doi: 10.1109/TAES.2016.140112.

[3] W. P. du Plessis, Path-length effects in multiloop retrodirective cross-eye jamming, *IEEE Antennas and Wireless Propagation Letters*, 2015, **15**, 626-629, doi: 10.1109/LAWP.2015.2465815.

[4] S. Song, Y. Dai, S. Sun, T. Jin, Efficient image reconstruction methods based on structured sparsity for short-range radar, *IEEE*

Transactions on Geoscience and Remote Sensing, 2024, **62**, 5212615, doi: 10.1109/TGRS.2024.3404626.

[5] L. Chen, J. Wang, X. Liu, D. Feng, G. Sun, A flexible range–Doppler modulation method for pulse-Doppler radar using phase-switched screen, *IEEE Transactions on Antennas and Propagation*, 2025, **73**, 6774-6787, doi: 10.1109/TAP.2025.3567479.

[6] J. Ma, D. Lu, J. Liu, L. Shi, A cross-eye jamming mitigation approach for polarized monopulse radars, *IEEE Transactions on Aerospace and Electronic Systems*, 2025, **61**, 5495-5498, doi: 10.1109/TAES.2024.3468294.

[7] T. Liu, X. Wei, Z. Liu, Overview of cross-eye jamming research, *Journal of Radars*, 2019, **8**, 140-153, doi: 10.12000/JR19013.

[8] Z. Cai, S. Xing, S. Quan, X. Su, J. Wang, Generation of equivalent false target for multiple triplet antenna system, *Results in Engineering*, 2025, 28, doi: 10.1016/j.rineng.2025.107433.

[9] W. P. du Plessis, J. W. Odendaal, J. Joubert, Extended analysis of retrodirective cross-eye jamming, *IEEE Transactions on Antennas and Propagation*, 2009, **57**, 2803-2806, doi: 10.1109/TAP.2009.2027353.

[10] W. P. DuPlessis, J. W. Odendaal, J. Joubert, Experimental simulation of retrodirective cross-eye jamming, *IEEE Transactions on Aerospace and Electronic Systems*, 2011, **47**, 734-740, doi: 10.1109/TAES.2011.5705704.

[11] W. P. DuPlessis, J. W. Odendaal, J. Joubert, Tolerance analysis of cross-eye jamming systems, *IEEE Transactions on Aerospace and Electronic Systems*, 2011, **47**, 740-745, doi: 10.1109/TAES.2011.5705705.

[12] W. P. Du Plessis, Platform skin return and retrodirective cross-eye jamming, *IEEE Transactions on Aerospace and Electronic Systems*, 2012, **48**, 490-501, doi: 10.1109/TAES.2012.6129650.

[13] W. P. Du Plessis, Limiting apparent target position in skin-return influenced cross-eye jamming, *IEEE Transactions on Aerospace and Electronic Systems*, 2013, **49**, 2097-2101, doi: 10.1109/TAES.2013.6558044.

[14] J. Ma, L. Shi, S. Xiao, X. Wang, Mitigation of cross-eye jamming using a dual-polarization array, *Journal of Systems Engineering and Electronics*, 2018, **29**, 491-498.

[15] S. Liu, C. Dong, J. Xu, G. Zhao, Y. Zhu, Analysis of rotating cross-eye jamming, *IEEE Antennas and Wireless Propagation Letters*, 2015, **14**, 939-942, doi: 10.1109/LAWP.2014.2387423.

[16] C. Wang, Z. He, J. Gong, Analysis of retro-directive cross-eye jamming for multiple elements, *Journal of Systems Engineering and Electronics*, 2017, **39**, 1457-1463, doi: 10.3969/j.issn.1001-506X.2017.07.05.

[17] T. Liu, D. Liao, X. Wei, L. Li, Performance analysis of multiple-element retrodirective cross-eye jamming based on linear array, *IEEE Transactions on Aerospace and Electronic Systems*, 2015, **51**, 1867-1876, doi: 10.1109/TAES.2015.140035.

[18] D. Yang, B. Liang, D. Zhao, Cross-eye gain distribution of multiple-element retrodirective cross-eye jamming, *Journal of Systems Engineering and Electronics*, 2018, **29**, 1170-1179, doi: 10.21629/JSEE.2018.06.06.

[19] J. H. Lee, J. Jo, H. K. Ryu, C. Lee, Experimental approach to estimate cross-eye gain for a nonretrodirective cross-eye jamming system, *IEEE Antennas and Wireless Propagation Letters*, 2022, **21**, 1120-1123, doi: 10.1109/LAWP.2022.3158683.

[20] W. Liu, J. Meng, L. Zhou, Interference modeling of rectangular array retrodirective cross-eye jamming method, *Journal of Systems Engineering and Electronics*, 2019, **41**, 2453-2459, doi: 10.3969/j.issn.1001-506X.2019.11.08.

[21] Z. Chen, Y. Ning, S. He, M. Yu, Experimental study on multiantenna interference technology, *IEEE Transactions on Aerospace and Electronic Systems*, 2023, **59**, 7878-7889, doi: 10.1109/TAES.2023.3297088.

[22] Y. Ning, M. Yu, Performance analysis of multiple antennas synthetic false target jamming, *IEEE Access*, 2022, **10**, 16178-16187.

[23] J. Chen, Q. Shi, Z. Huang, Q. Wang, N. Yuan, Performance analysis of multi-group three-tuple cross-eye jamming, *Journal of Systems Engineering and Electronics*, 2022, **33**, 80-90, doi: 10.23919/JSEE.2022.000009.

[24] R. Sui, J. Wang, G. Sun, Z. Xu, D. Feng, A dual-polarimetric high range resolution profile modulation method based on time-modulated APCM, *IEEE Transactions on Antennas and Propagation*, 2025, **73**, 1007-1017, doi: 10.1109/TAP.2025.3526901.

[25] Z. Cai, S. Xing, S. Quan, X. Su, J. Wang, A power-distribution joint optimization arrangement for multi-point source jamming system, *Results in Engineering*, 2025, **27**, doi: 10.1016/j.rineng.2025.106856.

[26] Z. Cai, S. Xing, X. Su, J. Wang, W. Meng, Z. Xiao, A joint optimization method for power and array of multi-point sources system, *Remote Sensing*, 2025, **17**, 2445, doi: 10.3390/rs17142445.

[27] K. Mishra, G. N. V. Rajareddy, U. Ghugar, G. S. Chhabra, A. H. Gandomi, A collaborative computation and offloading for compute-intensive and latency-sensitive dependency-aware tasks in dew-enabled vehicular fog computing: a federated deep Q-learning approach, *IEEE Transactions on Network and Service Management*, 2023, **20**, 4600-4614, 10.1109/TNSM.2023.3282795.

[28] R. B. Diddigi, C. Kamanchi, S. Bhatnagar, A generalized minimax Q-learning algorithm for two-player zero-sum stochastic games, *IEEE Transactions on Automatic Control*, 2022, **67**, 4816-4823, doi: 10.1109/TAC.2022.3159453.

[29] H. Saberi, C. Zhang, Z. Y. Dong, A multi-agent deep constrained Q-learning method for smart building energy management under uncertainties, *IEEE Transactions on Smart Grid*, 2024, **15**, 4649-4661, doi: 10.1109/TSG.2024.3386896.
[30] H. Jiang, G. Li, J. Xie, J. Yang, Action candidate driven clipped double Q-learning for discrete and continuous action tasks, *IEEE Transactions on Neural Networks and Learning Systems*, 2024, **35**, 5269-5279, doi: 10.1109/TNNLS.2022.3203024.
[31] Y. Shi, Z. Rong, Analysis of Q-learning like algorithms through evolutionary game dynamics, *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2022, **69**, 2463-2467, doi: 10.1109/TCSII.2022.3161655.
[32] J. Wang, S. Quan, S. Xing, Y. Li, H. Wu, W. Meng, PSO-based fine polarimetric decomposition for ship scattering characterization, *ISPRS Journal of Photogrammetry and Remote Sensing*, 2025, **220**, 18-31, doi: 10.1016/j.isprsjprs.2024.11.015.
[33] D. Wang, Y. Wang, M. Zhao, J. Qiao, Iterative Q-learning for model-free optimal control with adjustable convergence rate, *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2024, **71**, 2224-2228, doi: 10.1109/TCSII.2023.3339577.
[34] K. S. Shin, H. H. Choi, H. Lee, Knowledge transfer-based multiagent Q-learning for medium access in dense cellular networks, *IEEE Wireless Communications Letters*, 2022, **11**, 2542-2545, doi: 10.1109/LWC.2022.3207493.
[35] S. Song, Y. Dai, Y. Song, T. Jin, Efficient near-field radar microwave imaging based on joint constraints of low-rank and structured sparsity at low SNR, *IEEE Transactions on Microwave Theory and Techniques*, 2025, **73**, 2962-2977, doi: 10.1109/TMTT.2024.3479189.
[36] Z. Cai, S. Xing, S. Quan, X. Su, J. Wang, Generation of equivalent false target for multiple triplet antenna system, *Results in Engineering*, 2025, **28**, 107433, doi: 10.1016/j.rineng.2025.107433.
[37] W. P. du Plessis, Statistical skin-return results for retrodirective cross-eye jamming, *IEEE Transactions on Aerospace and Electronic Systems*, 2019, **55**, 2581-2591, doi: 10.1109/TAES.2019.2930168.
[38] J. Chen, S. Lv, T. Zhang, Y. Wang, Y. Wang, The semidouble DQN resource optimization strategy for UAV-aided networks: a case study, *IEEE Transactions on Aerospace and Electronic Systems*, 2025, **61**, 7852-7862, doi: 10.1109/TAES.2025.3541168.
[39] H. Mohammadi Rouzbahani, H. Karimipour, L. Lei, Optimizing resource swap functionality in IoE-based grids using approximate reasoning reward-based adjustable deep double Q-learning, *IEEE Transactions on Consumer Electronics*, 2023, **69**, 522-532, doi: 10.1109/TCE.2023.3279138.
[40] W. Fan, J. Xiong, Q-learning methods for LQR control of completely unknown discrete-time linear systems, *IEEE Transactions on Automation Science and Engineering*, 2024, **22**, 5933-5943, doi: 10.1109/TASE.2024.3434533.