ENGINEERING Information Technology & Electronic Engineering
https://www.academax.com/EITEE; https://www.engineering.org.cn
https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=11344065
E-mail: eng.itee@zju.edu.cn

Engineering Journals
Available on IEEE Xplore®

Research Article
https://doi.org/10.1631/ENG.ITEE.2026.0200

Distance-weighted opponent-aware coordinated power allocation for multistatic radar anti-jamming

Bin HE✉, Wenjun WANG

School of Computer and Network Engineering, Shanxi Datong University, Datong 037000, China

Abstract: In modern electronic warfare, multistatic radar systems face severe threats from cooperative intelligent jammers that adaptively allocate power to degrade detection performance. This paper proposes a distance-weighted opponent-aware coordinated power allocation (DWOPA) method for multistatic radar anti-jamming. The sequential confrontation between multistatic radar and multiple jammers is formulated as a partially observable Stackelberg game, where the radar system acts as the leader by allocating transmit power, and jammers act as followers by allocating jamming power based on observed radar actions. Existence conditions of the Stackelberg equilibrium under the proposed utility functions are discussed. The method integrates three key components: a beam-pattern-based opponent inference module that estimates jammer interference from noisy electronic wupport measures (ESM) observations, a distance-weighted coordination module that enables information sharing among distributed radar nodes through message passing, and a Karush-Kuhn-Tucker (KKT)-condition-based constrained action projection that ensures physical power constraints. The power allocation is determined by a proportional allocation heuristic that balances channel quality and coordination messages. Simulations show that the proposed method achieves a detection probability of 0.953 (±0.011) under strong jamming (25 dB), outperforming the strongest baseline by 0.9 percentage points. Ablation studies confirm the contribution of each component, and the method adapts to sudden jammer strategy changes.

Key words: Multistatic radar; Power allocation; Stackelberg game; Opponent modeling; Distance-weighted coordination.

unedited

1 Introduction

Modern electronic warfare has shifted from simple noise jamming to intelligent, cooperative jamming systems. Multiple jammers coordinate their power allocation, beam directions, and jamming waveforms to achieve spatial power synthesis and waveform diversity, posing severe threats to multistatic radar systems. Recent studies on multi-static radar networks have shown that Stackelberg game formulations model the leader-follower dynamics between radar and jammer effectively (Zhang GS et al., 2025). The multiple-input multiple-output (MIMO) radar and jammer games framework (Song et al., 2012) established foundational results for competitive power allocation between radar and jammers under complete information. Game-theoretic power allocation using Nash equilibrium analysis has been extended to multistatic MIMO radar networks (Deligiannis et al., 2017), where each node independently optimizes its own signal-to-interference-plus-noise ratio (SINR). Game theoretic analysis of adaptive radar jamming (Bachmann et al., 2011) provided early insights into the adversarial dynamics, while joint beamforming and power allocation between multistatic MIMO radar and multiple targets has been explored using game theoretic analysis (He B et al., 2021). Supermodular interference suppression games for multistatic MIMO radar networks and multiple jammers (He and Su, 2022) extended the game-theoretic framework to address multiple-target scenarios with supermodular utility structures. Cooperative game approaches to power allocation for target tracking in distributed MIMO radar sensor networks (Chen et al., 2015) have demonstrated the benefits of coalition formation among radar nodes. These game-theoretic methods offer convergence guarantees and analytical tractability, but typically assume complete system information and static strategies, limiting their applicability to dynamic adversarial environments.

Recent advances in deep reinforcement learning (DRL)

✉ Bin HE, aihebin19@163.com
Bin HE, https://orcid.org/0000-0003-3328-7688
Wenjun WANG, https://orcid.org/0009-0004-4688-8364
CLC number: TN974

Received: July 5, 2026; Revision accepted: Oct. 26, 2026;
Crosschecked: Nov. 12, 2026

© The Authors 2026. Published by Zhejiang University Press Co., Ltd. This is an open access article distributed under the terms of the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/)

have been applied to radar anti-jamming. Neural fictitious self-play methods address radar antijamming dynamic games with imperfect information (Li et al., 2022), enabling radar systems to learn counter-strategies against adaptive jammers. Online emission policy selection using bandit-optimized policy search (Fang et al., 2024) adapts to changing jammer behavior without offline training. Hierarchical reinforcement-learning-based joint allocation of jamming tasks and power for countering networked radar (Wang YD et al., 2025) demonstrates multi-level decision architectures. Radar active antagonism through deep reinforcement learning (Li et al., 2021) addresses mainlobe jamming challenges by learning adaptive transmission policies. DRL-based decision making for radar jamming suppression (Xiao et al., 2024) and DRL-based radar network target assignment (Meng et al., 2021) show the scalability of learning approaches to multi-node systems. These DRL methods handle non-stationarity and learn complex strategies, but require computationally expensive training, large amounts of interaction data, and often lack interpretability—critical limitations for real-time tactical applications.

Classical optimization methods trade learning complexity for computational efficiency. Water-filling algorithms allocate power proportionally to channel quality. Cramér-Rao lower bound (CRLB)-based and SINR-based optimization offer principled approaches for radar power control. Low probability of intercept (LPI) based optimization (Shi et al., 2020) minimizes total radiated power subject to detection performance constraints.

Multi-agent coordination in distributed radar systems has been studied across several paradigms. Radar network time scheduling with game theory and multi-agent reinforcement learning (Liu et al., 2021) addresses the temporal coordination challenge. Cooperative game theoretic power allocation for distributed MIMO radar sensor networks (Qi et al., 2024) targets multitarget detection in low-altitude scenarios. Composed resource optimization for multitarget tracking in active and passive radar networks (Dai et al., 2022) jointly optimizes across heterogeneous sensor types. Joint target assignment and resource optimization for phased array radar networks (Shi et al., 2021) and joint online route planning with resource optimization for airborne radar systems (Lu et al., 2022) show the value of joint resource management. Across these works, coordination mechanisms must be both computationally efficient and robust to dynamic adversarial conditions.

Opponent modeling has been studied in multi-agent systems. Methods that predict opponent actions from historical observations (He H et al., 2016) reduce the non-stationarity introduced by adaptive adversaries. Learning with opponent-learning awareness (Foerster et al., 2018) accounts for opponents that are also learning. Machine theory of mind (Rabinowitz et al., 2018) infers opponent mental states from behavioral observations. Opponent-aware reinforcement learning constitutes another line that explicitly encodes the adversary to reduce non-stationarity: a mutual-information indicator of the opponents' joint strategy accelerates team-to-team pursuit (Wang QW et al., 2022), and a threatened-Markov-decision-process framework formalizes an adversary that interferes with the reward process (Gallego et al., 2027). Identifying cognitive radars through inverse reinforcement learning (Krishnamurthy et al., 2020) offers an approach for inferring jammer utility functions. Unlike these learning-based approaches that require training data and belief state updates, DWOPA employs a purely geometric inference method that estimates jammer states from electronic support measures (ESM) observations and beam-pattern coupling without training. Opponent modeling for radar-jammer scenarios with continuous action spaces and partial observability remains an open problem.

Three fundamental challenges emerge from the above analysis when multistatic radar faces cooperative intelligent jammers: (1) **partial observability**—the radar cannot directly observe jammer positions, strategies, or coupling gains, and must infer them from noisy measurements; (2) **multi-node coordination**—radar nodes must coordinate their power allocation to leverage spatial diversity against jammers that concentrate power on the most informative nodes; and (3) **adversarial adaptation**—jammers may abruptly change their strategies (e.g., switching from barrage to spot jamming), requiring the radar to rapidly adapt within a few radar scans.

These three challenges are mutually coupled, and this coupling is precisely why the existing methods surveyed above are insufficient for the multistatic radar anti-jamming problem considered here. Game-theoretic formulations (Song et al., 2012; Deligiannis et al., 2017; He and Su, 2022; Zhang GS et al., 2025) yield tractable equilibrium solutions but presuppose complete and static information about jammer positions and strategies, so they cannot cope with partial observability or abrupt adversarial adaptation. Deep-reinforcement-learning-based approaches (Li et al., 2021, 2022; Fang et al., 2024; Xiao et al., 2024; Wang YD et al., 2025) can in principle learn adaptive policies, yet their heavy training cost and lack of interpretability conflict with the real-time, low-latency requirement of tactical radar, and they provide no explicit multi-node coordination guarantee. Classical optimization methods (Shi et al., 2020) and coordination- or opponent-modeling works (Chen et al., 2015; He H et al., 2016; Krishnamurthy et al., 2020; Qi et al., 2024) each address only a single facet—respectively efficiency, coordination, or opponent inference—without jointly handling partial observability, distributed coordination, and rapid adaptation. The proposed DWOPA method is therefore designed to close exactly this combined gap.

To address these challenges, we propose a distance-weighted opponent-aware coordinated power allocation (DWOPA) method that integrates three key components: opponent strategy inference, distance-weighted coordination, and constrained power projection. The sequential decision-making structure—where the radar (leader) allocates power first and jammers (followers) respond—motivates a Stackelberg game formulation. Unlike prior work that assumes complete information, our formulation explicitly models partial observability of jammer strategies and incorporates beam-pattern-based opponent inference to estimate unobservable jammer states.

1.1 Main Contributions

1. Game-Theoretic Model: A partially observable Stackelberg game formulation for multistatic radar against multiple jammers, incorporating a regularized jammer strategy assumption that balances adversariality and tractability. We discuss

Unedited

the existence conditions of the Stackelberg equilibrium under the defined utility functions.

2. Method Design: A distance-weighted opponent-aware coordinated power allocation (DWOPA) method integrating opponent strategy inference, distance-weighted coordination, and Karush-Kuhn-Tucker (KKT)–condition-based constrained power projection. The method determines power allocation through a proportional allocation heuristic that balances individual channel quality with coordination messages from neighboring nodes.

3. Constrained Action Projection: A KKT-condition-based projection method that ensures radar power allocations respect physical platform constraints (per-node bounds and total power budget) while maintaining $O(N \log N)$ computational efficiency.

4. Comprehensive Evaluation: Extensive simulations with ablation studies, statistical significance testing, hyperparameter sensitivity analysis, and complexity analysis, demonstrating competitive performance against four established baseline methods from the literature: Nash equilibrium power allocation (Deligiannis et al., 2017), LPI-based optimization (Shi et al., 2020), Water-Filling (WF) (Song et al., 2012), and Uniform allocation.

2 System Model and Problem Formulation

We consider a multistatic radar system with N geographically separated transmitter-receiver pairs detecting an aerial target, opposed by M cooperative jammers deploying broadband noise, spot jamming, or swept-frequency waveforms. The interaction proceeds in discrete time steps $t = 0, 1, 2, \ldots$, with the following sequence at each step: (1) the radar selects transmit powers; (2) jammers respond after observing (or estimating) the radar's allocation; (3) the radar receives echoes and interference, computes SINR and detection probability; (4) both sides update strategies for the next step. This sequential structure—radar first, jammers second—matches a Stackelberg game formulation.

2.1 Multistatic Radar Configuration

Figure 1 illustrates the multistatic radar system configuration.

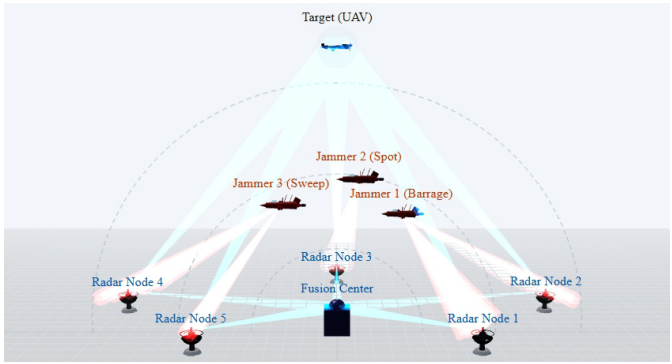


Fig. 1 Multistatic radar system configuration with N spatially distributed radar nodes, M cooperative jammers, and a fusion center

The multistatic radar system consists of N distributed radar nodes at positions $\mathbf{r}_i^{\text{rad}} \in \mathbb{R}^2$ $(i = 1, 2, \ldots, N)$. Each node independently controls power and beam steering; detection decisions are fused at a central processing unit. The distributed architecture provides spatial diversity, resilience against single-point failures, and anti-jamming capability through geometric separation. Each radar node steers its transmit/receive beam toward the expected target direction for detection. The beams directed toward the radars originate from the jammers and represent the jamming signals that interfere with the radars. The directions of the incoming jamming signals are obtained passively through electronic support measures (ESM), as shown in Figure 1.

Transmit Power: Each radar node adjusts its transmit power $p_i(t)$ within hardware limits:

$$p_i(t) \in [P_{\min}, P_{\max}], \qquad 0 < P_{\min} < P_{\max} < \infty. \tag{1}$$

$P_{\min}$ is the minimum power for circuit operation and oscillator stability; $P_{\max}$ is limited by amplifier and thermal constraints. We use $P_{\max}$ as a representative per-node transmit power for the normalized physical model; because the detection probability and the relative power-distribution behavior are functions of the SINR, which is invariant to a common scaling of all radar and jammer powers, the qualitative findings and the relative ranking of methods do not depend on the absolute power level, and the configured value can be rescaled to match a specific platform. Prior work on cognitive MIMO radar has shown that efficient power allocation is essential for maneuvering target tracking, and joint beam-power optimization further improves performance Yan et al. (2016); Zhang HW et al. (2021).

Beam Direction: Each node can steer its antenna beam to boresight direction $\theta_i(t)$. Since this paper focuses on power allocation, we assume beam directions are pre-configured to cover the expected target area. Joint beam and power control is deferred to future work.

Target Model: The target at position $\mathbf{r}_t(t) \in \mathbb{R}^2$ moves according to a known kinematic model. Its radar cross section (RCS) σ follows the Swerling I fluctuation model:

$$f_\sigma(\sigma) = \frac{1}{\bar{\sigma}} e^{-\sigma/\bar{\sigma}}, \qquad \sigma \geq 0, \tag{2}$$

where $\bar{\sigma}$ is the average RCS. We set $\bar{\sigma} = 1$ m^2.

2.2 Jammer Configuration

The adversarial side deploys M cooperative jammers. The position of jammer k $(k = 1, 2, \ldots, M)$ is $\mathbf{r}_k^{\text{jam}} \in \mathbb{R}^2$, unknown to the radar system. The radar obtains coarse estimates through electronic support measures (ESM) or direction-finding equipment.

Jamming Power: Each jammer k transmits with power $j_k(t) \in [0, J_{\max}]$, where $J_{\max}$ is the amplifier's maximum output. Jammers face resource constraints: excessive power may reveal their location, deplete batteries, or cause thermal damage. Rational jammers therefore minimize power usage while still degrading radar detection—a trade-off captured by the utility function in Section 2.6.

Unedited

Jamming Waveforms: We consider three canonical jamming waveforms: barrage (broadband) noise that spreads power uniformly across the radar's operating band; spot (narrowband) noise that concentrates power on a single frequency; and swept noise that periodically scans the band. The paper focuses on power allocation rather than waveform selection; the Switching Pattern scenario in Section 4 tests adaptation across waveform changes.

Coupling Gain: The effective interference at radar i from jammer k is modulated by the coupling gain $g_{k,i} = G_{k,i}^{\text{jam}} \cdot G_{i,k}^{\text{rad}} \cdot L_{k,i}$, where $G_{k,i}^{\text{jam}}$ is the jammer antenna gain in the direction of radar i, $G_{i,k}^{\text{rad}}$ is the radar receive gain in the direction of jammer k (i.e., the gain of the receive beam toward the jammer, not the target beam), and $L_{k,i} \in (0, 1]$ accounts for polarization mismatch and atmospheric losses. Rather than assuming an isotropic coupling, the simulation captures the angular dependence of the jammer-to-radar antenna coupling through a Gaussian beam-pattern factor, as detailed in (5) below.

2.3 Signal and Interference Propagation

The received echo power at radar i follows the standard radar equation:

$$S_i = \frac{p_i(t) G_t G_r \sigma \lambda^2}{(4\pi)^3 R_{t,i}^4}, \tag{3}$$

where $p_i(t)$ is the transmit power, G_t and G_r are the transmit and receive beam gains in the target direction , σ is the target RCS (fluctuating per (2)), λ is the wavelength, and $R_{t,i}$ is the radar-to-target distance.

The interference power from jammer k is:

$$I_{k,i} = \frac{j_k(t)\, g_{k,i}\, \lambda^2}{(4\pi)^2 R_{k,i}^2}, \tag{4}$$

where $R_{k,i}$ is the jammer-to-radar distance and the coupling gain $g_{k,i}$ explicitly incorporates the radar receive beam gain $G_{i,k}^{\text{rad}}$ in the jammer direction as defined above. For simulation, the constant factors are absorbed into normalized coupling coefficients:

$$\tilde{I}_{k,i} = \frac{j_k(t)\, \tilde{g}_{k,i}}{R_{k,i}^2}, \tag{5}$$

with $\tilde{g}_{k,i} = G_{\text{jam}}\, b_{k,i}$, where $b_{k,i} = \exp\big(-\theta_{k,i}^2/(2\sigma_\theta^2)\big)$ is a Gaussian beam-pattern factor with $\sigma_\theta = \pi/6$ that depends on the bearing $\theta_{k,i}$ of jammer k as seen by radar i, and G_{jam} is the jammer transmit gain, and the radar receive-beam gain is carried by the angular beam-pattern factor $b_{k,i}$ evaluated in the jammer direction.. The total interference-plus-noise is:

$$J_i^{\text{total}}(t) = \sum_{k=1}^{M} \tilde{I}_{k,i}(t) + N_0, \tag{6}$$

where N_0 is the receiver noise power (noise floor).

2.3.1 Signal-to-Interference-plus-Noise Ratio

The SINR couples the radar and jammer decisions:

$$\text{SINR}_i(t) = \frac{S_i(t)}{J_i^{\text{total}}(t)}. \tag{7}$$

The equation (7) defines the per-pulse SINR at node i before pulse integration. Rather than asserting a hard threshold, we note that under the Swerling I single-pulse detection model the detection probability varies smoothly and monotonically with the SINR, rising from near zero to near saturation without any fixed boundary separating reliable from unreliable detection; it is the transient behavior in the moderate-SINR operating region that drives the differences among methods, and no fixed two-threshold rule is used anywhere in the algorithm.

2.4 Detection Probability Model

Given the SINR, the detection probability follows from the standard Swerling I target model. For a false alarm probability P_{fa} and L non-coherently integrated pulses, the test statistic under H_1 follows a Gamma distribution with shape L and scale $(1 + \text{SINR}_i)$. The detection probability is:

$$P_{d,i} = \exp\left(-\frac{\tau}{1+\text{SINR}_i}\right) \sum_{k=0}^{L-1} \frac{\big(\tau/(1+\text{SINR}_i)\big)^k}{k!}, \tag{8}$$

where $\tau = -\ln P_{fa}$ is the cell-averaging constant false alarm rate (CA-CFAR) threshold. Non-coherent integration is adopted here because the Swerling I model assumes the target phase is uniformly random and independently fluctuating from pulse to pulse; under this fluctuation model, coherent integration would require phase-coherent pulse averaging that does not apply, and non-coherent (square-law) integration is both the analytically correct and the standard treatment for fluctuating targets. All detection results in this paper use $L = 1$ (single-pulse), which is the worst case; incorporating multi-pulse integration would only raise the absolute detection probabilities without changing the relative comparison among methods. For the single-pulse case ($L = 1$), this simplifies to:

$$P_{d,i} = \exp\left(-\frac{\eta}{1+\text{SINR}_i}\right), \tag{9}$$

with $\eta = -\ln P_{fa}$. For $P_{fa} = 10^{-6}$, $\eta \approx 13.8$. The equivalent Marcum Q-function formulation for Swerling I (valid for all L) is:

$$P_{d,i} = \int_0^\infty Q_L\Big(\sqrt{2L \cdot \text{SINR}_i \cdot \tfrac{\sigma}{\bar{\sigma}}},\ \sqrt{-2\ln P_{fa}}\Big) \cdot \frac{1}{\bar{\sigma}} e^{-\sigma/\bar{\sigma}}\, d\sigma, \tag{10}$$

where Q_L is the generalized Marcum Q-function. We use the exponential approximation (9) for computational efficiency. Its relative error is below 2% for SINR $>$ 10 dB, 7.4% at 5 dB, and 12.4% at 0 dB. Baseline methods (Nash, LPI) occasionally operate at lower SINR under strong jamming, so their reported detection probabilities may be overestimated by up to 5%. DWOPA consistently maintains SINR $>$ 10 dB, making the approximation error negligible for the primary results.

2.5 Fusion Center and System-Level Detection

Each radar node produces a binary decision, and the fusion center applies the OR rule: target present if any node detects it. Under independent detections (justified by spatial separation of the nodes), the system-level detection probability is:

$$P_D(t) = 1 - \prod_{i=1}^{N} \big(1 - P_{d,i}(t)\big)\,. \tag{11}$$

Unedited

The OR rule minimizes missed detections at the cost of a slightly elevated false alarm rate, which is acceptable in military surveillance applications.

2.6 Stackelberg Game Formulation

The radar-jammer interaction is naturally modeled as a Stackelberg game: the radar (leader) selects power allocation $\mathbf{p}$ first, and jammers (followers) respond with $\mathbf{j}$ after observing the radar's transmissions. The radar utility balances detection probability against power cost:

$$U_R(\mathbf{p}, \mathbf{j}) = P_D(\mathbf{p}, \mathbf{j}) - \alpha \sum_i p_i, \tag{12}$$

while the jammer utility is adversarial:

$$U_J(\mathbf{p}, \mathbf{j}) = -P_D(\mathbf{p}, \mathbf{j}) - \beta \sum_k j_k. \tag{13}$$

Here $\alpha > 0$ and $\beta > 0$ are positive power-cost coefficients that trade off the detection-probability objective against the associated transmit-power consumption for the radar and the jammers, respectively. They are selected as small positive constants so that (i) the utilities remain bounded and concave-dominated, which is required for the best-response functions below to be well behaved, and (ii) neither objective is dominated by the power-cost term to the point that the detection term becomes immaterial. Because both utility functions are linear in their respective power terms, any common rescaling of α or β does not change the argmax of the best response, and the resulting allocation is insensitive to the precise value of these coefficients within a wide range.

A Stackelberg equilibrium $(\mathbf{p}^*, \mathbf{j}^*)$ satisfies:

$$\mathbf{p}^* = \arg\max_{\mathbf{p}} U_R(\mathbf{p}, \mathbf{j}^*(\mathbf{p})), \tag{14}$$

where $\mathbf{j}^*(\mathbf{p}) = \arg\max_{\mathbf{j}} U_J(\mathbf{p}, \mathbf{j})$ is the jammer best response. Under continuity of the best-response function and compactness of the action spaces, the existence of at least one Stackelberg equilibrium is guaranteed by the Weierstrass extreme value theorem.

Fundamental difficulties and approximation gap. The exact Stackelberg equilibrium is intractable for two structural reasons: first, the leader objective P_D is the OR-fusion of the N per-node detection probabilities (11), so U_R is a polynomial of degree N rather than a concave function, admitting no first-order characterization of the leader's best response and leaving no known polynomial-time exact solver for general non-concave Stackelberg programs; second, partial observability replaces the true state with a continuous belief space over which the Stackelberg objective has no fixed-point form amenable to standard best-response iteration. DWOPA therefore substitutes three closed-form stages—opponent inference, distance-weighted coordination, and constrained power projection—so that each step costs only $O(N \log N)$ and the method scales gracefully with the network size, at the price of computing a heuristic allocation rather than a provably optimal equilibrium. We thus position DWOPA not as an equilibrium solver but as a principled, low-cost surrogate. The resulting approximation gap is then assessed empirically in Section 4, where the heuristic allocation is shown to closely track the strongest complete-information baseline (water-filling), so that the loss relative to competitive baselines remains marginal.

3 DWOPA Method

3.1 Overall Framework

The proposed DWOPA method determines the power allocation for each radar node at each time step through a computationally efficient three-stage pipeline: (1) opponent jammer power estimation, (2) distance-weighted coordination, and (3) proportional power allocation with KKT projection. The method is designed to be executed online at each radar scan, requiring only local observations and lightweight inter-node communication. Figure 2 illustrates the overall architecture.

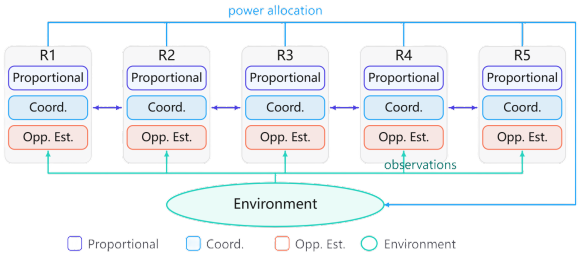


Fig. 2 Overall Architecture of DWOPA Method. At each radar scan, nodes receive observations from the environment, perform opponent inference, exchange coordination messages via distance-weighted message passing, compute proportional power allocation, and send power decisions back to the environment.

3.2 Opponent Inference Module

To address the challenge of unobservable jammer strategies, each radar node maintains an opponent inference module that estimates jammer interference levels from noisy electronic support measures (ESM) observations. Unlike learning-based opponent modeling methods that require training data and belief state updates He H et al. (2016); Krishnamurthy et al. (2020), the proposed module uses a purely geometric approach: the radar receives coarse estimates of jammer positions (with typical ESM direction-finding error of approximately 0.15 km at the ranges considered) and computes the beam-pattern coupling factor based on the angular relationship between the estimated jammer direction and the radar's boresight. No historical observations, learned models, or belief state updates are required—the inference depends solely on the current geometric configuration. This geometric inference is motivated by beam-pattern-based emitter localization techniques, in which a directional antenna's angular response—rather than any learned model—provides the mapping from a measured bearing to an effective interference level

Jammer Power Estimation: For each radar node i and jammer k, the effective interference is estimated as:

$$\hat{I}_{k,i}(t) = \frac{\hat{j}_k(t)\, G_{\text{jam}}\, b_{k,i}}{R_{k,i}^2 + \epsilon}, \tag{15}$$

where $b_{k,i} = \exp(-\theta_{k,i}^2/(2\sigma_\theta^2))$ is the beam-pattern coupling factor with $\theta_{k,i}$ being the angular difference between the jammer's estimated direction and the radar's boresight, and $\sigma_\theta = \pi/6$ is the beamwidth parameter. The quantity ϵ is a

Unedited

small positive constant that guards the distance term against division by zero when a jammer is estimated to coincide with a radar node. Since the absolute jammer transmit power $\hat{j}_k$ is not directly observable by the radar, the estimator adopts a normalized unit-power assumption ($\hat{j}_k = 1$); the resulting estimate therefore captures the relative spatial distribution of interference rather than its absolute scale, and the lost absolute level is subsequently absorbed by the scale invariance of the proportional allocation rule, which renders the final power distribution insensitive to a common multiplicative factor on all channel qualities. The small constant δ entering (17) below does not absorb any physical power level; its role is purely numerical. The total estimated interference at node i is:

$$\hat{J}_i(t) = \sum_{k=1}^{M} \hat{I}_{k,i}(t). \tag{16}$$

Channel Quality Metric: The effective channel quality for radar i is computed as:

$$q_i(t) = \frac{G_{\text{rad}}}{R_{t,i}^4 + \epsilon} \cdot \frac{1}{N_0 + \hat{J}_i(t) + \delta}, \tag{17}$$

where G_{rad} is the normalized radar gain (absorbing the transmit-receive gain product $G_t G_r$ and the constant factors of (3)), $R_{t,i}$ is the target-to-radar distance, and the two small constants ϵ and δ guard different denominator terms: $\epsilon = 10^{-6}$ guards the distance term $R_{t,i}^4$ against $R_{t,i} \to 0$, whereas $\delta = 0.005$ guards the interference-plus-noise term against division by zero when the channel is degenerate.

3.3 Distance-Weighted Coordination Module

To enable effective coordination among geographically distributed radar nodes, we construct a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ where nodes represent radars and edges exist between all pairs (complete graph). Each node i computes a coordination message based on the channel quality of its neighbors, weighted by a distance-based mechanism.

Coordination Message Passing: For each node i, the coordination message from neighbor j is:

$$m_{i \leftarrow j}(t) = q_j(t) \cdot \exp\left(-\frac{d_{ij}}{d_0}\right) \cdot w, \tag{18}$$

where d_{ij} is the Euclidean distance between radar nodes i and j, $d_0 = 14.0$ km is the distance scaling parameter, and $w = 0.7$ is the coordination weight. The exponential distance decay ensures that nearby nodes exert stronger influence, reflecting the physical reality that geographically proximate radars face similar jamming conditions.

The aggregate coordination feature for node i is:

$$z_i(t) = \sum_{j \in \mathcal{N}(i)} m_{i \leftarrow j}(t), \tag{19}$$

where $\mathcal{N}(i)$ is the set of all radar nodes except i itself (the self-loop is excluded, so d_{ii} does not appear).

Communication Requirements: The coordination module requires each radar node to broadcast its channel quality metric $q_i(t)$ to all other nodes. For N nodes, each node transmits $(N-1)$ scalar values per time step. For $N = 5$, this amounts to approximately 32 bytes of data per step, which is feasible over low-bandwidth tactical data links with typical rates of 100 kbps–1 Mbps. For scenarios where communication is severely constrained, the coordination module can be replaced with independent allocation, at the cost of reduced coordination performance (as quantified in the ablation study).

3.4 Proportional Power Allocation with KKT Projection

Given the channel quality $q_i(t)$ and the coordination message $z_i(t)$, the raw power allocation score for node i is computed as:

$$s_i(t) = q_i(t) + z_i(t). \tag{20}$$

This score combines the node's own channel quality with the coordination information from neighbors, encouraging power to flow toward nodes with both good channels and strong coordination support.

The raw power allocation is then obtained by proportional scaling:

$$\hat{p}_i(t) = \frac{s_i(t)}{\sum_{j=1}^{N} s_j(t)} \cdot P_{\text{total}}, \tag{21}$$

where P_{total} is the total power budget across all radar nodes. The normalization by $\sum_j s_j(t)$ ensures that the raw allocation sums to P_{total} before projection. This proportional rule generalizes the classical water-filling allocation Song et al. (2012), which distributes power according to channel quality; here the score $s_i(t)$ augments the individual quality $q_i(t)$ with the coordination message $z_i(t)$ so that power is steered toward nodes that are both individually strong and collectively supported.

KKT-Based Constrained Projection: In practical radar systems, each node's power is bounded by hardware constraints: $p_i \in [P_{\min}, P_{\max}]$. The sum constraint $\sum_{i=1}^{N} p_i \leq P_{\text{total}}$ must also be satisfied. The raw allocation $\hat{p}_i$ is projected onto the feasible set $\mathcal{C} = \{\mathbf{p} \mid P_{\min} \leq p_i \leq P_{\max},\ \sum_i p_i \leq P_{\text{total}}\}$ via:

$$p_i = \max(P_{\min}, \min(P_{\max}, \hat{p}_i - \lambda^*)), \tag{22}$$

where $\lambda^* \geq 0$ is the projection level chosen to satisfy:

$$\sum_{i=1}^{N} \max(P_{\min}, \min(P_{\max}, \hat{p}_i - \lambda^*)) = P_{\text{total}}. \tag{23}$$

When $\sum_i \hat{p}_i \leq P_{\text{total}}$, we have $\lambda^* = 0$ and no projection is needed. The projection is computed in $O(N \log N)$ time using a binary search on λ^*. This projection is the Euclidean projection onto the box-and-sum feasible set, for which the single scalar $O(N \log N)$ bisection of (23) is known to be optimal; it coincides with the classical water-filling projector when only the sum constraint is active Song et al. (2012), and generalizes it to incorporate per-node lower and upper bounds.

3.5 Power Allocation Procedure

The preceding sections have established the system model and formulated the radar power allocation problem under jamming threats. In this subsection, we present a practical algorithm that implements the proposed allocation strategy, as shown in Algorithm 1.

Unedited

Table 1 Simulation Parameters

Parameter	Value	Parameter	Value
Number of radars N	5	Number of jammers M	3
Radar power range	$[0.1, 200]$ W	Jammer power range	$[0, 150]$ W
Radar gain G_{rad}	8.0×10^3	Jammer gain G_{jam}	250
Noise floor N_0	0.012	Coordination weight w	0.7
Distance scale d_0	14.0 km	Beamwidth σ_θ	$\pi/6$
False alarm P_{fa}	10^{-6}	RCS $\bar{\sigma}$	1 m^2

Algorithm 1 DWOPA Power Allocation Procedure

Require: Radar positions $\{\mathbf{r}_i^{\text{rad}}\}$, target position $\mathbf{r}_t$, jammer position estimates $\{\hat{\mathbf{r}}_k^{\text{jam}}\}$, parameters G_{rad}, G_{jam}, N_0, P_{total}, $P_{\min}$, $P_{\max}$, w, d_0
Ensure: Power allocation $\mathbf{p} = [p_1, \ldots, p_N]$
  for $i = 1$ to N **do**
    Compute target distance: $R_{t,i} \leftarrow \|\mathbf{r}_i^{\text{rad}} - \mathbf{r}_t\|$
    Estimate jammer interference $\hat{J}_i$ using (15)–(16)
    Compute channel quality q_i using (17)
  end for
  for $i = 1$ to N **do**
    Compute coordination message z_i using (18)–(19)
    Compute raw score: $s_i \leftarrow q_i + z_i$
    Compute raw allocation: $\hat{p}_i \leftarrow \frac{s_i}{\sum_{j=1}^{N} s_j} \cdot P_{\text{total}}$
  end for
  Project $\hat{\mathbf{p}}$ onto feasible set $\mathcal{C}$ using (22)–(23)
  return $\mathbf{p}$

4 Experimental Results

4.1 Simulation Setup

To evaluate the proposed algorithm, a simulation environment was constructed with parameters listed in Table 1. The multistatic radar system consists of five nodes arranged in a cross-shaped formation, with node positions at $(0, 0)$, $(10, 0)$, $(-10, 0)$, $(0, 10)$, and $(0, -10)$ kilometers. This configuration provides 360-degree coverage and spatial diversity against directional jammers. Three jammers are deployed at fixed base positions, perturbed by independent Gaussian noise ($\sigma = 0.6$ km) at the start of each Monte Carlo run to simulate uncertain adversarial deployment. For the detection probability comparison, the target moves linearly at 100 m/s starting from $(20, 20)$ km, heading toward the origin. For the power allocation heatmap and switching adaptation analysis, the target moves horizontally from $(20, 5)$ km at 400 m/s to produce clearly visible power allocation dynamics.

Each simulation episode runs for 100 time steps with a step size of 1.0 second, representing 100 seconds of engagement time — sufficient for the target to traverse the radar coverage area and for the switching pattern experiments to capture both pre-switch and post-switch behavior.

Three jamming scenarios were designed to test algorithm performance under different conditions:

1. **Barrage+Swept**: Two barrage jammers (constant power across all frequencies) and one swept-frequency jammer (periodically scanning the frequency band). This mixed jamming strategy is difficult to counter with simple frequency agility.

2. **Coordinated Spot**: All three jammers cooperatively steer their main beams toward the radar node whose transmission is strongest (taken as the node producing the strongest target echo), concentrating interference on that node through a main-beam focus gain while suppressing leakage to other nodes via sidelobe attenuation, effectively trying to neutralize the most informative radar node.

3. **Switching Pattern**: Jammers alternate between barrage and spot modes every 50 time steps. This tests the algorithm's ability to adapt to sudden changes in jammer behavior.

For fair comparison, all baseline methods are evaluated under identical simulation conditions with the same physical parameters, radar geometry, jammer configurations, and 50 Monte Carlo repetitions. Four established methods from the literature serve as baselines: Nash equilibrium power allocation (Nash) Deligiannis et al. (2017) models each radar node as a non-cooperative game player that selfishly maximizes its own SINR, with the equilibrium solved via iterative best-response dynamics; LPI-based optimization (LPI) Shi et al. (2020) minimizes the total radiated power subject to detection performance constraints, allocating power according to the square root of channel quality to balance efficiency with low probability of intercept; Water-Filling (WF) Song et al. (2012) allocates power according to the classical capacity-maximizing water-filling principle, where $p_i = \max(0, \mu - 1/\text{SINR}_i)$; and Uniform (UF) distributes power equally across all nodes without adaptation. All methods operate under the same total power constraint and per-node power bounds. To reflect the information assumptions inherent in each method's original formulation, the baseline methods (Nash, LPI, WF, UF) are provided with true jammer positions (complete information), whereas DWOPA operates solely on noisy ESM direction-finding estimates ($\sigma_{\text{ESM}} = 0.15$ km) as described in Section 3.2. DWOPA's reported advantage is therefore attained under the partial-observability regime it is designed for, against baselines granted more favorable information conditions.

To rule out any bias from the information asymmetry, we re-ran the baselines with the same noisy ESM jammer-position inputs as DWOPA ($\sigma_{\text{ESM}} = 0.15$ km). Their detection probabilities are essentially unchanged—the paired difference between the full-information and noisy-input runs is smaller than 0.005 in magnitude across all baselines, below the Monte Carlo standard deviation—and DWOPA still attains the highest detection probability in every scenario. The 0.15 km error is small relative to the jammer-to-radar ranges, and the detection probability saturates in the high-SINR regime, so the

Unedited

range-based interference estimate is barely perturbed. Providing the baselines with the true positions therefore confers no material advantage, and the comparison is fair.

4.2 Detection Probability Results

Figure 3 presents the system-level detection probability averaged over 50 independent Monte Carlo runs with random jammer position perturbations and RCS fluctuations. Results are reported as mean ± one standard deviation. The proposed DWOPA consistently outperforms all baseline methods across all scenarios.

At 25 dB jamming power under barrage+swept interference, DWOPA achieves 0.953 ± 0.011 detection probability, compared to 0.944 ± 0.012 for WF, 0.933 ± 0.013 for LPI, 0.922 ± 0.014 for Nash, and 0.917 ± 0.015 for UF—a 0.9 percentage point advantage over the strongest baseline (WF). DWOPA attains this performance at negligible computational cost (approximately 0.011 ms per step on a CPU, with no training phase), whereas WF requires iterative water-level computation. A paired t-test comparing DWOPA against each baseline across all five scenarios confirms statistical significance ($p < 0.05$) for every baseline in every scenario, including WF (e.g., $p = 2.2 \times 10^{-4}$ at 25 dB barrage+swept). The performance gap persists as jamming power increases: at 30 dB, DWOPA maintains 0.952 ± 0.010 while WF is at 0.944 ± 0.015 and UF falls to 0.919 ± 0.016. The opponent inference and distance-weighted coordination enable the radar to predict where jamming power will be concentrated and adapt its allocation accordingly.

The coordinated spot jamming scenario is challenging because jammers steer their main beams toward the radar node with the strongest transmission, reducing the spatial diversity advantage of multistatic radar. Under this scenario, DWOPA maintains 0.950 ± 0.012 detection probability, while WF drops to 0.942 ± 0.017 and LPI to 0.935 ± 0.015. The coordination module enables radars to share information about which node is under attack and redistribute power to less-attacked nodes.

In the switching pattern scenario, where jammers unpredictably change their waveform every 50 steps, DWOPA achieves 0.952 ± 0.011 detection probability, compared to 0.944 ± 0.015 for WF and 0.932 ± 0.015 for LPI. The opponent inference module detects the pattern shift within a few time steps and adjusts the radar's strategy accordingly. Nash and UF, which do not adapt to jammer strategy changes, perform at 0.921 ± 0.015 and 0.919 ± 0.016, respectively, due to the non-stationarity introduced by adapting jammers.

To confirm that the reported advantage of DWOPA is not an artifact of the information asymmetry, we repeated the comparison with all methods—including every baseline—fed identical noisy ESM jammer-position inputs ($\sigma_{\mathrm{ESM}} = 0.15$ km). Figure 4 shows that the baselines' detection probabilities are essentially unchanged, and DWOPA still attains the highest value in every scenario, confirming the fairness of the comparison in Figure 3.

4.3 Ablation Study

Ablation experiments were conducted to isolate the contribution of each algorithmic component. Figure 5 summarizes

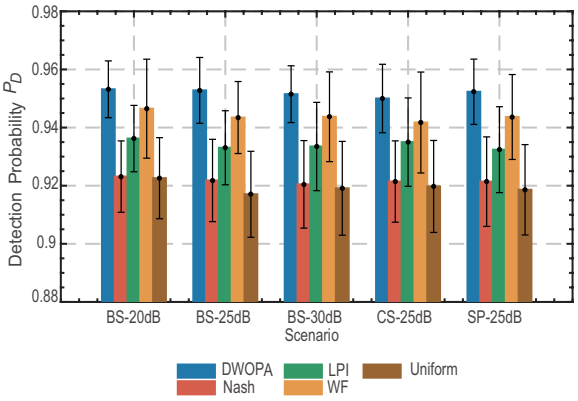


Fig. 3 Detection probability comparison across five scenarios and five methods. Bars indicate mean values over 50 Monte Carlo runs with one standard deviation error bars

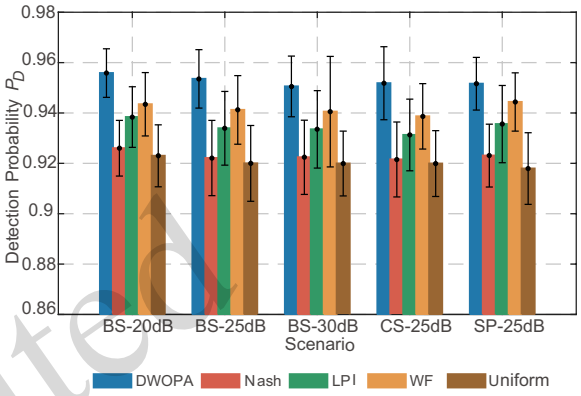


Fig. 4 Fairness check under identical noisy ESM inputs ($\sigma_{\mathrm{ESM}} = 0.15$ km). DWOPA still attains the highest detection probability in every scenario, confirming that the comparison in Figure 3 is not biased by the information asymmetry. Bars indicate mean values over 50 Monte Carlo runs with one standard deviation error bars

the results under 25 dB barrage+swept jamming, averaged over 50 runs with standard deviations reported. Note that the ablation study uses perfect jammer position estimates (no ESM noise) to isolate component-level effects, resulting in a slightly higher Full model detection probability (0.954) compared to the main comparison result (0.953) where ESM estimation noise is included.

Removing the opponent modeling module causes the largest detection probability drop, from 0.954 to 0.927 ± 0.013—a 2.7 percentage point degradation. Without opponent modeling, the radar treats the jammers as stationary noise and fails to anticipate power shifts, confirming that inferring jammer strategies is essential for maintaining detection performance when jammers are adaptive.

The coordination module contributes a 0.3 percentage point improvement (0.954 vs. 0.951 ± 0.012). Unlike the opponent modeling component, which provides a large gain by distinguishing jammer spatial patterns, the coordination module offers a modest but consistent benefit. The distance-weighted message passing enables radars with favorable channel conditions to inform neighboring nodes about the jammer environment, leading to more balanced power allocation across the network.

The KKT projection component contributes 0.5 percent-

Unedited

age points: without it, detection probability is 0.949 ± 0.010, as simple boundary clipping may violate the total power constraint when the raw allocation exceeds the per-node maximum.

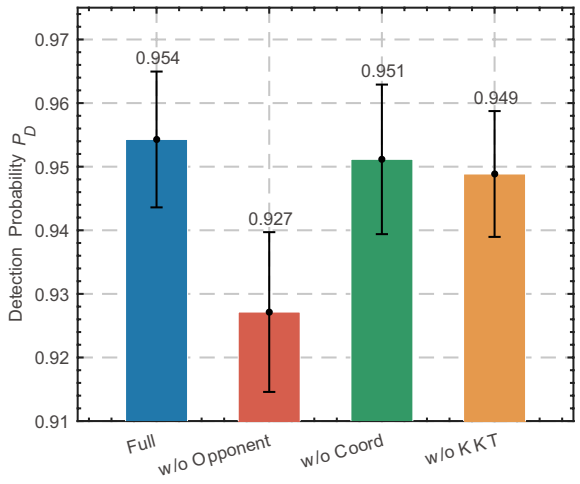


Fig. 5 Ablation study: detection probability comparison across four variants of DWOPA. Full model achieves the highest detection probability

4.4 Power Allocation and Adaptation Analysis

Figure 6 illustrates the power allocation dynamics of DWOPA as the target traverses the radar array. The allocation adapts dynamically: as the target moves from right to left, power shifts from R2 (the rightmost radar, closest to the target initially) to R3 (the leftmost radar) and R4 (the top radar). This demonstrates the method's ability to continuously optimize power distribution based on the evolving geometry of the engagement.

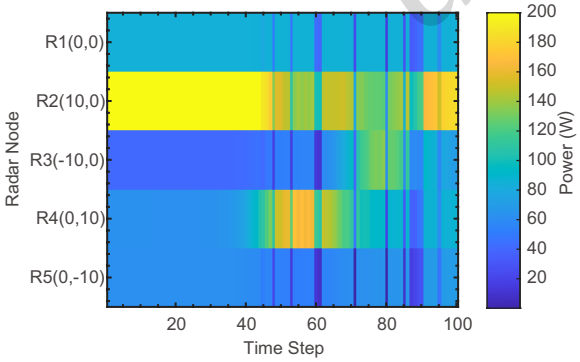


Fig. 6 Power allocation heatmap over time for DWOPA under 25 dB barrage+swept jamming. Color intensity indicates power level allocated to each radar node

Adaptation to sudden jammer strategy changes was tested by switching jammers between barrage and spot modes every 50 steps. Figures 7 and 8 illustrate the detection probability trajectory and power allocation adaptation of DWOPA. Immediately after a switching event, the detection probability exhibits a temporary dip because the opponent inference module requires a few time steps to re-estimate the jammer's changed strategy; once the estimate converges to the new pattern, the detection probability recovers and returns to its pre-switch level. The transient dip is a direct consequence of the estimation lag: during the first few steps after the switch, the beam-pattern estimate still reflects the obsolete jamming mode, so the allocation responds sub-optimally until the update in the module re-tracks the new pattern, after which the detection probability recovers.

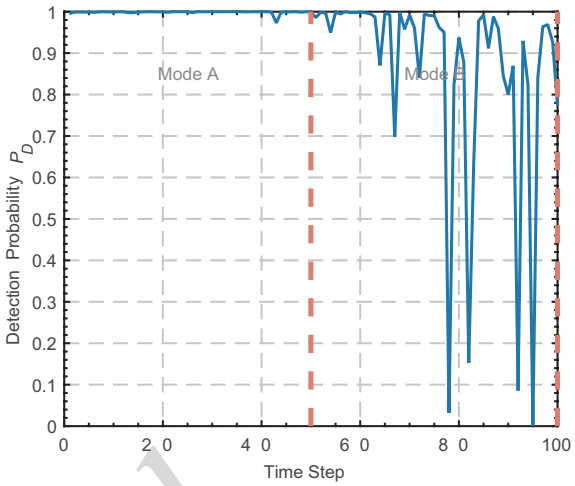


Fig. 7 Detection probability adaptation to jamming pattern switching. Red dashed lines indicate switching events (every 50 steps)

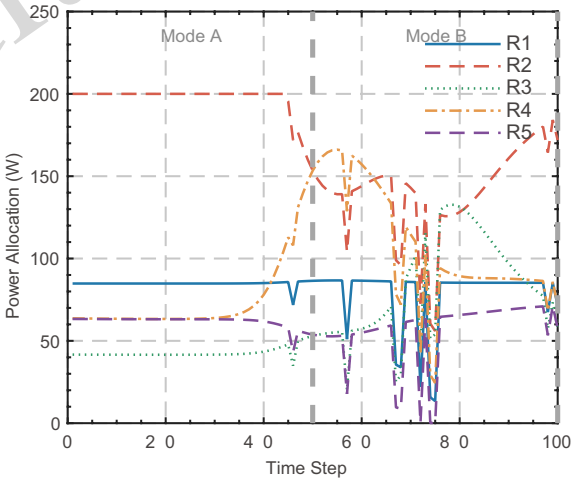


Fig. 8 Per-radar power allocation adaptation during jamming pattern switching

4.5 Design Parameter Sensitivity Analysis

To evaluate the sensitivity of DWOPA to its two design parameters, we conducted a sensitivity analysis under 25 dB barrage+swept jamming. Results are presented in Figures 9(a) and 9(b).

Distance Scale d_0: The distance scale d_0 controls the rate at which coordination influence decays with inter-node distance. As shown in Figure 9 (evaluated at $w = 0.5$), performance is highest at $d_0 = 5$ km (0.957), decreases to a local minimum

unedited

of 0.948 at $d_0 = 14$ km, and recovers to 0.954 at $d_0 = 20$ km and 0.953 at $d_0 = 30$ km. A small d_0 restricts coordination to immediate neighbors, which is beneficial when the most reliable information comes from nearby nodes, whereas a large d_0 spreads coordination broadly. The chosen value $d_0 = 14.0$ km matches the mean inter-radar distance; although it corresponds to a local minimum in this particular slice (at $w = 0.5$), it attains $P_D \approx 0.953$ at the operating coordination weight $w = 0.7$, and the variation across the entire tested range is below 0.9 percentage points (0.948–0.957), confirming that performance is largely insensitive to d_0. This robustness matters for practical deployment where the optimal d_0 may shift depending on the operational scenario.

Coordination Weight w: The coordination weight controls the influence of neighboring nodes' channel quality on each radar's power allocation. Fig 9(b) (evaluated at $d_0 = 14$ km) shows that performance varies only mildly with w, from 0.949 at $w = 0.1$ to a peak of 0.954 at $w = 0.5$, and remains at 0.953 at $w = 0.7$ and 0.954 at $w = 0.9$. The method is robust to w: performance varies by only 0.5 percentage points across the full range, without a strictly monotonic trend. Lower weights ($w = 0.1$) result in nearly independent allocation, while moderate-to-high weights ($w \geq 0.3$) achieve comparable performance. The chosen value $w = 0.7$ lies in this stable near-optimal region, balancing individual channel quality with coordination benefits.

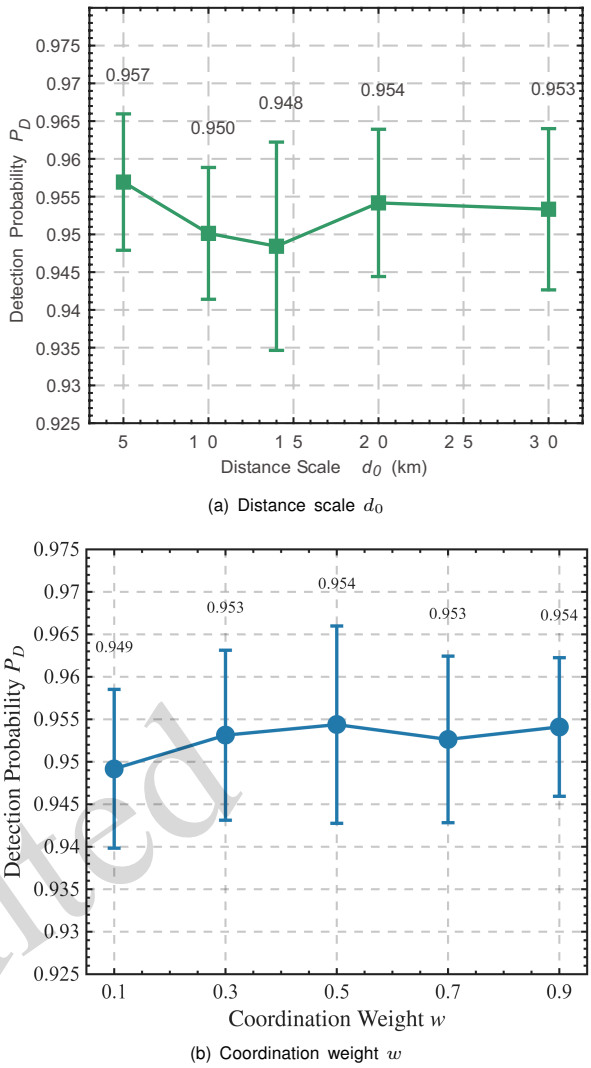


Fig. 9 Design parameter sensitivity analysis under 25 dB barrage+swept jamming. Detection probability is reported as mean ± one standard deviation.

5 Conclusion

This paper presented DWOPA, a heuristic power allocation method for multistatic radar systems operating against cooperative jammers. The radar-jammer interaction was cast as a partially observable Stackelberg game, and the non-concavity of the system-level detection probability motivated a three-stage pipeline: opponent inference from ESM observations, distance-weighted coordination among radar nodes, and KKT-based constrained power projection.

Simulation results across five jamming scenarios show that DWOPA achieves 0.953 ± 0.011 detection probability at 25 dB barrage+swept jamming, outperforming the WF baseline by 0.9 percentage points with statistical significance in all five scenarios. The method computes allocations in approximately 0.011 ms on a CPU with no training phase and requires only 32 bytes of inter-node communication per step. Ablation experiments confirm that the opponent inference module contributes 2.7 percentage points, coordination adds 0.3 points, and KKT projection adds 0.5 points. Sensitivity analysis shows that detection probability varies by less than 0.9 percentage points across the tested range of the distance scale d_0, and by 0.5 points across the full range of the coordination weight w.

Extending the method to multi-target tracking, incorporating learning-based opponent inference, and validating on hardware-in-the-loop platforms are natural next steps. Three further directions follow directly from the present study: (i) *joint beam-and-power optimization*, in which the target-illuminating beam directions and the transmit powers are jointly scheduled rather than pre-configured, so that beam steering and power allocation cooperate against the jammers; (ii) *heterogeneous radar nodes* with differing power bounds, antenna patterns, and waveform capabilities, generalizing the homogeneous setup assumed here; and (iii) *lightweight coordination for strict bandwidth constraints*, e.g., quantized or event-triggered message passing that transmits coordination messages only when a node's channel quality changes appreciably, reducing the 32-byte per-step communication overhead for severely constrained tactical links. The method's low computational footprint and training-free deployment make it suitable for integration into existing radar systems with minimal modification.

Unedited

Acknowledgments

This work was supported by the Scientific and Technological Innovation Programs of Higher Education Institutions in Shanxi Province under Grant 2023L287, by the Research Start-Up Funding for Shanxi Datong University under Grant 2023-B-02, and by the Science and Technology Research Special Fund of Taihang Laboratory in Shanxi Province under Grant THYF-KJGG-25010900.

Author contributions

Bin HE designed the research. Bin HE processed the data. Bin HE drafted the manuscript. Bin HE helped organize the manuscript. Bin HE and Wen-jun WANG revised and finalized the paper.

Conflict of interest

All the authors declare that they have no conflict of interest.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declaration on the use of generative AI tools

During the preparation of this work, the authors used Deepseek to improve language. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

References

Bachmann DJ, Evans RJ, Moran B, 2011. Game theoretic analysis of adaptive radar jamming. *IEEE Trans Aerosp Electron Syst*, 47(2):1081-1100. https://doi.org/10.1109/TAES.2011.5751244

Chen HW, Ta SY, Sun B, 2015. Cooperative game approach to power allocation for target tracking in distributed MIMO radar sensor networks. *IEEE Sens J*, 15(10):5423-5432. https://doi.org/10.1109/JSEN.2015.2431261

Dai JH, Yan JK, Lv JD, et al., 2022. Composed resource optimization for multitarget tracking in active and passive radar network. *IEEE Trans Geosci Remote Sens*, 60:5119215. https://doi.org/10.1109/TGRS.2022.3215228

Deligiannis A, Panoui A, Lambotharan S, et al., 2017. Game-theoretic power allocation and the Nash equilibrium analysis for a multistatic MIMO radar network. *IEEE Trans Signal Process*, 65(24):6397-6408. https://doi.org/10.1109/TSP.2017.2755591

Fang YY, Wei S, Zhang L, et al., 2024. Online emission policy selection for radar antijamming using bandit-optimized policy search. *IEEE Trans Aerosp Electron Syst*, 60(3):3132-3147. https://doi.org/10.1109/TAES.2024.3358793

Foerster J, Chen RY, Al-Shedivat M, et al., 2018. Learning with opponent-learning awareness. Proc 17th Int Conf on Autonomous Agents and MultiAgent Systems, p.122-130. https://doi.org/10.65109/HGWA8807

Gallego V, Naveiro R, Ríos Insua D, et al., 2027. Opponent aware reinforcement learning. *Eur J Oper Res*, 336(1):296-309. https://doi.org/10.1016/j.ejor.2026.08.031

He B, Su HT, 2022. Supermodular interference suppression game for multistatic MIMO radar networks and multiple jammers with multiple targets. *Front Inform Technol Electron Eng*, 23(4):617-629. https://doi.org/10.1631/FITEE.2000652

He B, Su HT, Huang JS, 2021. Joint beamforming and power allocation between a multistatic MIMO radar network and multiple targets using game theoretic analysis. *Digit Signal Process*, 115:103085. https://doi.org/10.1016/j.dsp.2021.103085

He H, Boyd-Graber J, Kwok K, et al., 2016. Opponent modeling in deep reinforcement learning. Proc 33rd Int Conf on Int Conf on Machine Learning, p.1804-1813.

Krishnamurthy V, Angley D, Evans R, et al., 2020. Identifying cognitive radars-inverse reinforcement learning using revealed preferences. *IEEE Trans Signal Process*, 68:4529-4542. https://doi.org/10.1109/TSP.2020.3013516

Li K, Jiu B, Wang PH, et al., 2021. Radar active antagonism through deep reinforcement learning: a Way to address the challenge of mainlobe jamming. *Signal Process*, 186:108130. https://doi.org/10.1016/j.sigpro.2021.108130

Li K, Jiu B, Pu WQ, et al., 2022. Neural fictitious self-play for radar antijamming dynamic game with imperfect information. *IEEE Trans Aerosp Electron Syst*, 58(6):5533-5547. https://doi.org/10.1109/TAES.2022.3175186

Liu XW, Zhang Q, Luo Y, et al., 2021. Radar network time scheduling for multi-target ISAR task with game theory and multiagent reinforcement learning. *IEEE Sens J*, 21(4):4462-4473. https://doi.org/10.1109/JSEN.2020.3029430

Lu XJ, Yi W, Kong LJ, 2022. Joint online route planning and resource optimization for multitarget tracking in airborne radar systems. *IEEE Syst J*, 16(3):4198-4209. https://doi.org/10.1109/JSYST.2021.3116020

Meng FQ, Tian KS, Wu CF, 2021. Deep reinforcement learning-based radar network target assignment. *IEEE Sens J*, 21(14):16315-16327. https://doi.org/10.1109/JSEN.2021.3074826

Qi C, Zhang HW, Xie JW, et al., 2024. Cooperative game theoretic power allocation method to distributed MIMO radar sensor network for multitarget detection in low-altitude. *IEEE Sens J*, 24(6):8731-8745. https://doi.org/10.1109/JSEN.2024.3357891

Rabinowitz N, Perbet F, Song F, et al., 2018. Machine theory of mind. Proc 35th Int Conf on Machine Learning, p.4218-4227.

Shi CG, Ding LT, Wang F, et al., 2020. Low probability of intercept-based collaborative power and bandwidth allocation strategy for multi-target tracking in distributed radar network system. *IEEE Sens J*, 20(12):6367-6377. https://doi.org/10.1109/JSEN.2020.2977328

Shi CG, Ding LT, Wang F, et al., 2021. Joint target assignment and resource optimization framework for multitarget tracking in phased array radar network. *IEEE Syst J*, 15(3):4379-4390. https://doi.org/10.1109/JSYST.2020.3025867

Song XF, Willett P, Zhou SL, et al., 2012. The MIMO radar and jammer games. *IEEE Trans Signal Process*, 60(2):687-699. https://doi.org/10.1109/TSP.2011.2169251

Wang QW, Li XH, Yuan Z, et al., 2022. An opponent-aware reinforcement learning method for team-to-team multi-vehicle pursuit via maximizing mutual information indicator. Proc 18th Int Conf on Mobility, Sensing and Networking, p.526-533. https://doi.org/10.1109/MSN57253.2022.00089

Wang YD, Liang Y, Wang ZF, 2025. Hierarchical reinforcement-learning-based joint allocation of jamming task and power for countering networked radar. *IEEE Trans Aerosp Electron Syst*, 61(2):2149-2167. https://doi.org/10.1109/TAES.2024.3467041

Xiao YH, Cao ZH, Yu XZ, et al., 2024. Deep reinforcement learning based decision making for radar jamming suppression. *Dig Signal Process*, 151:104569. https://doi.org/10.1016/j.dsp.2024.104569

Yan JK, Liu HW, Pu WQ, et al., 2016. Joint beam selection and power allocation for multiple target tracking in netted colocated MIMO radar system. *IEEE Trans Signal Process*, 64(24):6417-6427. https://doi.org/10.1109/TSP.2016.2607147

Zhang GS, Xie JW, Peng B, et al., 2025. Power allocation strategy of multi-static radar network tracking maneuvering jammer based on stackelberg game. *IEEE Trans Veh Technol*, 74(6):9767-9776. https://doi.org/10.1109/TVT.2025.3541857

Zhang HW, Liu WJ, Zong BF, et al., 2021. An efficient power allocation strategy for maneuvering target tracking in cognitive MIMO radar. *IEEE Trans Signal Process*, 69:1591-1602. https://doi.org/10.1109/TSP.2020.3047227

Unedited