SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

July 2026, Vol. 69, Iss. 7, 172304:1–172304:16
https://doi.org/10.1007/s11432-025-4740-0

On covert transmission of distributed antenna systems against heterogeneous adversaries

Weiyu CHEN[1], Haiyang DING[2], Shilian WANG[1*], Junshan LUO[1,3] & Fengkui GONG[4]

[1]*College of Electronic Science and Technology, National University of Defense Technology, Changsha 410073, China*
[2]*Engineering University of Information Support Force, Wuhan 430019, China*
[3]*The Sixty-Third Research Institute, National University of Defense Technology, Nanjing 210007, China*
[4]*State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an 710071, China*

Received 15 June 2025/Revised 7 October 2025/Accepted 28 December 2025/Published online 8 June 2026

Abstract This paper investigates the covert transmission from multiple ground nodes to a full-duplex unmanned aerial vehicle (UAV) under the surveillance of three types of wardens. The transmit nodes constitute a distributed antenna system to form sharp and flexible beams. The UAV receiver generates jamming signals while receiving to confuse the wardens. The wardens are classified as ground wardens, near UAV wardens, and far UAV wardens based on their positions; different channel models and covert communication strategies apply to different warden types. Analytical expressions are derived to evaluate the wardens' detection performance. Furthermore, the benefits of increasing the number of collaborative transmit nodes are revealed via asymptotic analyses. The collaborative transmit nodes' transmit powers, phases, and the receiver's jamming power are jointly optimized to minimize the receiver's outage probability while satisfying the covertness constraints for all the wardens. The formulated optimization problem is subtly decomposed by using the monotonicity properties of the Marcum Q function and is tackled by alternating optimization, with one subproblem having a closed-form optimal solution and the other solved with the semidefinite relaxation method. Numerical results show significant performance gains of the proposed algorithm over the existing strategy.

Keywords distributed communication systems, low-probability-of-detection communications, physical layer security, collaborative beamforming, covertness performance

Citation Chen W Y, Ding H Y, Wang S L, et al. On covert transmission of distributed antenna systems against heterogeneous adversaries. Sci China Inf Sci, 2026, 69(7): 172304, https://doi.org/10.1007/s11432-025-4740-0

1 Introduction

Distributed antenna systems (DASs) have recently gained prominence due to their inherent advantages [1], such as no single point of failure, on-demand deployment, high scalability, and ability extension of hardware/resource-limited nodes. Civil [2, 3] and military [4, 5] prototypes have been reported. A typical civil application of DASs is wireless sensor networks, where the main benefits of DASs are increased sensor node life time and extended communication range [1]. A typical military application of DASs is long-range communications at the squad level, where the main benefits of DASs are reduced size, weight and power and the elimination of single points of failure [4].

In many practical scenarios, wireless transmission must be stealthy. Typical examples include hiding nodes in military applications and hiding secret activities under surveillance. To these ends, covert communication has gained considerable research interest [6]. In the typical model of covert communication, transmitter Alice aims to transmit information to receiver Bob without being discovered by warden Willie, in the sense of a high probability of error for Willie's detector. The existence and variations of background noise and interference play an important role; they make Willie uncertain about whether his observations contain target communication signals. The most basic and inherent uncertainty stems from the variations in the instantaneous value of the noise at Willie [7]. Furthermore, Willie's uncertainty about noise variance [8], transmission time [9], interference power [10], etc., can be exploited to enhance the covert data rate.

Another effective way to enhance the covert data rate is to equip the transmit node with multiple antennas. Specifically, by using the beamforming techniques, the signals from different transmit antennas can be constructively superimposed at Bob, whereas the signals arriving at Willie are generally randomly superimposed or even can be destructively superimposed in particular scenarios. In this field, the authors in [11] investigated the beamforming

* Corresponding author (email: wangsl@nudt.edu.cn)

Table 1 Comparisons of the research on covert DASs.

Ref.	Scenario	Channel model	Beamforming strategy	Source of wardens' uncertainty	Main contributions
[16]	2D	Rayleigh	EGT	Friendly jammers	Comparisons of DASs and CASs
[17]	2D	LoS	EGT	Noise	Analyses of performance limit
[18]	2D	LoS	Phase align and power optimization	Friendly jammers	Deployment of jammers and transmitters
This work	3D	Three models for three types of wardens	Joint optimization of phase and power	Full-duplex receiver	Performance analyses and beamforming designs

designs for three scenarios in terms of the channel state information (CSI) available at Alice, wherein the received signal-to-noise ratio was maximized under the covertness constraint. It was revealed that the transmit power is not constrained when full CSI of Willie is available, and the maximum ratio transmission (MRT) strategy is optimal in Rayleigh fading channels when Alice only knows Willie's statistical CSI. The work in [12,13] further incorporated the full-duplex technique, where Bob generates interference while receiving to create interference uncertainty at Willie. The data beamformer and the interference beamformer were jointly optimized. The authors in [14] investigated the employment of the frequency diverse array to generate a distance-and-angle-dependent beampattern for covert transmission. Very recently, the authors in [15] incorporated the fluid antenna technique, where Alice adjusts the positions of antennas to enhance covertness.

The aforementioned research is on centralized antenna systems (CASs). Although both CASs and DASs have been substantially examined in terms of physical layer security [16], only a few studies have considered covert DASs [16–18], which are summarized in Table 1. Specifically, the work in [16] showed that MRT-based CASs outperform equal gain transmission (EGT)-based DASs in Rayleigh fading channels with randomly located wardens and interferers. The authors in [17] proved that in line-of-sight (LoS) channels, if the number of collaborative transmit nodes is large enough (of order of magnitude $\mathcal{O}(\sqrt{n})$), the limit covert and reliable data rate can be $\mathcal{O}(n)$ bits per n channel uses. The work in [18] studied the joint deployment of unmanned aerial vehicle (UAV) transmitters and friendly UAV jammers for covert collaborative transmission.

Despite the pioneering contributions made by [16–18], the investigation of covert DASs is still nascent. As summarized in Table 1, existing studies simply adopted the EGT beamforming strategy or merely optimized the power allocation. Note that the performance of DASs considerably depends on the adopted beamforming strategy. Jointly designing transmit power, phase shifts, and other key system parameters may fully exploit the potential of covert DASs, which has not yet been investigated to the knowledge of the authors. In addition, as seen from Table 1, existing studies merely considered a single type of channel model under two-dimensional (2D) scenarios. However, in practical 3D scenarios, DASs may simultaneously face heterogeneous adversaries with diverse geographical relationships and channel characteristics. The performance analyses and parameter designs of covert DASs under such unfriendly scenarios remain unsolved problems. To elucidate the aforementioned problems, this paper studies covert DASs against three types of wardens, named ground Willie, near UAV Willie, and far UAV Willie. Theoretical analyses are conducted to reveal the potential of the systems, and the key system parameters are jointly designed. The main contributions are summarized as follows.

(1) For the covert collaborative transmission from multiple ground transmit nodes to a remote full-duplex UAV receiver Bob, the diversity of warden types is considered in terms of the geographical relationships and channel characteristics. Analytical expressions of the probabilities that wardens' detection performance is worse than a preset tolerance are derived to evaluate the covertness. Furthermore, the potential of collaborative beamforming for covert communications is revealed. Specifically, on the basis of our constructed beamforming strategies, it is proven that the outage probability and the covertness can be simultaneously improved to any desired level by increasing the number of transmit nodes.

(2) The transmit power and phase shifts at the transmit nodes, as well as Bob's interference power, are jointly optimized to minimize the outage probability, on the premise of satisfying the covertness constraints for all three types of wardens. The formulated nonconvex problem is solved by exploiting the monotonicity properties of the Marcum Q function and by adopting the alternating optimization method. Particularly, the closed-form optimal solution to the power optimization subproblem is obtained. Simulation results validate the theoretical results and demonstrate the superior performance of the proposed algorithm over the benchmarks.

The main symbols and notations used in this paper are listed in Appendix G.

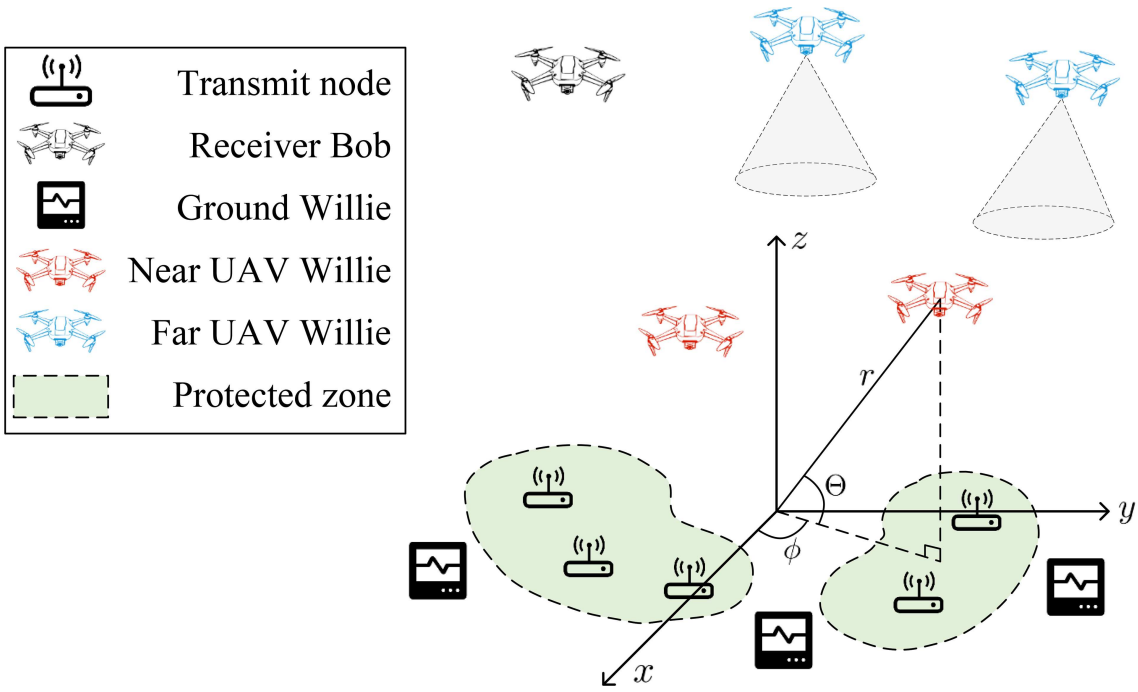

Figure 1 (Color online) Illustration of the considered covert DASs.

2 System model

As shown in Figure 1, in the considered system model, in order to covertly transmit information to a distant UAV (Bob) under the surveillance of three types of wardens, namely ground Willie, near UAV Willie, and far UAV Willie, N_a single-antenna ground transmit nodes collaboratively form the beam. For conciseness, we refer to the whole of the transmit nodes as Alice. Each ground Willie is located on the ground, and is equipped with an omnidirectional antenna so as to detect the communication behaviors around him. Similarly, each near UAV Willie is assumed to hover around Alice's area and is equipped with an omnidirectional antenna for detection. In contrast, the distances from Alice to Bob/far UAV Willies are assumed to be much larger than the distances between different transmit nodes, which means that Bob and far UAV Willies are in the far field of the array formed by the transmit nodes. To create interference uncertainty at the wardens, Bob uses an antenna to generate interference and meanwhile uses another antenna for reception [12,13]. From a conservative perspective, each far UAV Willie is assumed to be equipped with a directional antenna, whose main lobe covers Alice's area. For tractability, it is assumed that the wardens are non-colluding and all the nodes are static during one transmission [16].

Quasi-static flat fading is assumed for all the channels. Specifically, the ground-to-ground channels are assumed to follow the Rayleigh fading. The air-to-air channels are considered pure LoS channels. The ground-to-air channels are modeled by the Rician fading [19,20], with different path loss exponents $\alpha(\Theta)$ and different Rician K-factors $K(\Theta)$ for different elevation angles Θ [21]. Perfect synchronization is assumed among the transmit nodes [16,17], which can be done by cables [16] or by existing wireless synchronization techniques[1)] [1]. As a result, Bob's received signals can be written as

$$y_b(n) = \frac{1}{\sqrt{d_b^{\alpha(\Theta_b)}}} \boldsymbol{h}_b^{\mathrm{T}} \boldsymbol{w} x(n) + n_b(n). \tag{1}$$

Herein, $x(n) \sim \mathcal{CN}(0,1)$ denotes the normalized base-band symbol, where $n = 1, 2, \ldots, N$ is the index of channel uses. We focus on the scenarios where the number of channel uses in one transmission N is sufficiently large [16–18]. $n_b(n) \sim \mathcal{CN}\left(0, P_b\rho + \sigma_b^2\right)$ denotes the combination of residual loop interference and noise, with P_b being Bob's interference power and $0 < \rho < 1$ being the self-interference coefficient [22,23]. Both Alice and Willie are uninformed about the specific value of P_b, which is fixed during one transmission but is randomly chosen from the uniform distribution $\mathrm{U}(0, P_{\max})$ to confuse Willie [10]. d_b and Θ_b represent the distance and the elevation angle from Alice's

1) The covertness of the wireless synchronization process is an interesting research topic but does not fit the focus of this work. Note that the distances between the transmit nodes are much shorter than that from Alice to Bob, which reduces the difficulty of covert synchronization. In addition, Bob can broadcast training sequences periodically to help with synchronization.

geometric center to Bob, respectively. The large-scale path losses from all the transmit nodes to Bob are regarded as equal [24], because Bob is in the far field of Alice. $\boldsymbol{w} \triangleq \left[\sqrt{P_1}e^{\iota\varphi_1}, \sqrt{P_2}e^{\iota\varphi_2}, \ldots, \sqrt{P_{N_a}}e^{\iota\varphi_{N_a}}\right]^{\mathrm{T}}$ denotes the transmission weights of Alice, where P_i and φ_i are the transmit power and the phase shift of the i-th transmit node ($i = 1, 2, \ldots, N_a$), respectively. The normalized channel coefficients from Alice to Bob $\boldsymbol{h}_b \in \mathbb{C}^{N_a \times 1}$ are modeled as

$$\boldsymbol{h}_b = \sqrt{\frac{K\left(\Theta_b\right)}{K\left(\Theta_b\right)+1}}e^{\iota\psi_b}\bar{\boldsymbol{h}}_b + \sqrt{\frac{1}{K\left(\Theta_b\right)+1}}\hat{\boldsymbol{h}}_b. \tag{2}$$

The first and the second terms of the right-hide side (RHS) of (2) are the LoS and the non-LoS components, respectively, with $\hat{\boldsymbol{h}}_b$ following $\mathcal{CN}\left(\boldsymbol{0}, \boldsymbol{I}_{N_a}\right)$. Without loss of generality, let Alice's geometric center be the origin of coordinates, as shown in Figure 1. Then, ψ_b is the phase shift due to the propagation from the origin to Bob, which follows distribution $\mathrm{U}\left(0, 2\pi\right)$ [17]. Steering vector $\bar{\boldsymbol{h}}_b$ can be given by [25]

$$\bar{\boldsymbol{h}}_b \triangleq \left[e^{\iota\frac{2\pi}{\lambda}r_1\cos\left(\Theta_b\right)\cos\left(\phi_b-\phi_1\right)}, e^{\iota\frac{2\pi}{\lambda}r_2\cos\left(\Theta_b\right)\cos\left(\phi_b-\phi_2\right)}, \ldots, e^{\iota\frac{2\pi}{\lambda}r_{Na}\cos\left(\Theta_b\right)\cos\left(\phi_b-\phi_{N_a}\right)}\right]^{\mathrm{T}}, \tag{3}$$

where λ denotes the wavelength, r_i stands for the distance from the origin to the i-th transmit node, whereas ϕ_b and ϕ_i are the azimuth angles of Bob and the i-th transmit node, respectively.

Let the null hypothesis H_0 represent that Alice keeps silence, and let the alternate hypothesis H_1 denote that Alice transmits information. Conservatively, we follow the common assumption that Willies know the time framework of the possible transmission [26]. The detection at ground Willies can be modeled as

$$\begin{aligned}
H_0 &: y_{g,j}\left(n\right) = \sqrt{\frac{P_b}{d_{g,b,j}^{\alpha\left(\Theta_{g,b,j}\right)}}}h_{g,b,j}v\left(n\right) + n_{g,j}\left(n\right), \\
H_1 &: y_{g,j}\left(n\right) = \sum_{i=1}^{N_a}\sqrt{\frac{P_i}{d_{g,i,j}^{\alpha(0)}}}h_{g,i,j}e^{\iota\varphi_i}x\left(n\right) + \sqrt{\frac{P_b}{d_{g,b,j}^{\alpha\left(\Theta_{g,b,j}\right)}}}h_{g,b,j}v\left(n\right) + n_{g,j}\left(n\right),
\end{aligned} \tag{4}$$

where $j = 1, 2, \ldots, J_g$ and J_g denote the index and the number of ground Willies, respectively. $v\left(n\right) \sim \mathcal{CN}\left(0, 1\right)$ and $n_{g,j}\left(n\right) \sim \mathcal{CN}\left(0, \sigma_w^2\right)$ are the normalized interference and the noise, respectively. $d_{g,i,j}$ and $d_{g,b,j}$ are the distances from the i-th transmit node and Bob to the j-th ground Willie, respectively. $\Theta_{g,b,j}$ is the elevation angle from the j-th ground Willie to Bob. $h_{g,b,j} = \sqrt{\frac{K\left(\Theta_{g,b,j}\right)}{K\left(\Theta_{g,b,j}\right)+1}}e^{\iota\psi_{g,b,j}} + \sqrt{\frac{1}{K\left(\Theta_{g,b,j}\right)+1}}\hat{h}_{g,b,j}$ and $h_{g,i,j} \sim \mathcal{CN}\left(0, 1\right)$ are the normalized channel coefficients from Bob and the i-th transmit node to the j-th ground Willie, respectively, with $\psi_{g,b,j}$ following $\mathrm{U}\left(0, 2\pi\right)$ and $\hat{h}_{g,b,j}$ following $\mathcal{CN}\left(0, 1\right)$, as before.

Similarly, the detection of near UAV Willies is modeled as

$$\begin{aligned}
H_0 &: y_{c,j}\left(n\right) = \sqrt{\frac{P_b}{d_{c,b,j}^2}}e^{\iota\psi_{c,b,j}}v\left(n\right) + n_{c,j}\left(n\right), \\
H_1 &: y_{c,j}\left(n\right) = \sum_{i=1}^{N_a}\sqrt{\frac{P_i}{d_{c,i,j}^{\alpha\left(\Theta_{c,i,j}\right)}}}h_{c,i,j}e^{\iota\varphi_i}x\left(n\right) + \sqrt{\frac{P_b}{d_{c,b,j}^2}}e^{\iota\psi_{c,b,j}}v\left(n\right) + n_{c,j}\left(n\right),
\end{aligned} \tag{5}$$

where $j = 1, 2, \ldots, J_c$ and J_c denote the index and the number of near UAV Willies, respectively. $d_{c,i,j}$ and $d_{c,b,j}$ are the distances from the i-th transmit node and Bob to the j-th near UAV Willie, respectively. $\Theta_{c,i,j}$ denotes the elevation angle from the i-th transmit node to the j-th near UAV Willie. $n_{c,j}\left(n\right) \sim \mathcal{CN}\left(0, \sigma_w^2\right)$ denotes the noise. The ground-to-air channel coefficient from the i-th transmit node to the j-th near UAV Willie $h_{c,i,j}$ follows the same model as $h_{g,b,j}$, whereas $\psi_{c,b,j} \sim \mathrm{U}\left(0, 2\pi\right)$ is the phase shift due to the propagation from Bob to the j-th near UAV Willie.

As the far UAV Willies are in the far field of Alice, the channels from Alice to the far UAV Willies are modeled based on the steering vector. As a result, the detection of far UAV Willies can be modeled as

$$\begin{aligned}
H_0 &: y_{f,j}\left(n\right) = \sqrt{w_f}\sqrt{\frac{P_b}{d_{f,b,j}^2}}e^{\iota\psi_{f,b,j}}v\left(n\right) + n_{f,j}\left(n\right), \\
H_1 &: y_{f,j}\left(n\right) = \frac{\sqrt{W_f}}{\sqrt{d_{f,j}^{\alpha\left(\Theta_{f,j}\right)}}}\boldsymbol{h}_{f,j}^{\mathrm{T}}\boldsymbol{w}x\left(n\right) + \sqrt{w_f}\sqrt{\frac{P_b}{d_{f,b,j}^2}}e^{\iota\psi_{f,b,j}}v\left(n\right) + n_{f,j}\left(n\right),
\end{aligned} \tag{6}$$

where $j = 1, 2, \ldots, J_f$ denotes the index of far UAV Willies. Herein, W_f and w_f represent the far UAV Willies' antenna gain towards Alice and that towards Bob[2)], respectively. Without loss of generality, W_f and w_f are assumed as the same for different far UAV Willies for conciseness. $d_{f,j}$ and $\Theta_{f,j}$ represent the distance and the elevation angle from Alice's geometric center to the j-th far UAV Willie, respectively. As before, $d_{f,b,j}$ and $\psi_{f,b,j} \sim \mathrm{U}(0, 2\pi)$ represent the distance from Bob to the j-th far UAV Willie and the corresponding phase shift, respectively, whereas $n_{f,j}(n) \sim \mathcal{CN}\left(0, \sigma_w^2\right)$ denotes the noise. Similar to $\boldsymbol{h}_b$, herein the normalized channel coefficients from Alice to the j-th far UAV Willie $\boldsymbol{h}_{f,j}$ are modeled as $\boldsymbol{h}_{f,j} = \sqrt{\frac{K(\Theta_{f,j})}{K(\Theta_{f,j})+1}} e^{\iota\psi_{f,j}} \bar{\boldsymbol{h}}_{f,j} + \sqrt{\frac{1}{K(\Theta_{f,j})+1}} \hat{\boldsymbol{h}}_{f,j}$. Herein, $\bar{\boldsymbol{h}}_{f,j}$ is the steering vector and can also be written as (3), with Θ_b and ϕ_b replaced by $\Theta_{f,j}$ and the j-th far UAV Willie's azimuth angle $\phi_{f,j}$, respectively. $\psi_{f,j}$ and $\hat{\boldsymbol{h}}_{f,j}$ follow $\mathrm{U}(0, 2\pi)$ and $\mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_{N_a})$, respectively, as before.

Alice is assumed to know the instantaneous CSI of Alice-Bob channels [16], because Bob can periodically transmit training sequences. However, Alice does not know the CSI of Alice-Willie channels, because the wardens can be completely passive. Alternatively, it is assumed that Alice knows the positions of near UAV Willies and far UAV Willies. This can be done completely passively based on visible light [27], infrared [28], acoustics [29], passive radar [30], or communication/control signals [31]. In addition, additional cooperative nodes that do not need to hide themselves can play the role of radar/reconnaissance to obtain and then broadcast the positions. Conservatively, the number and the positions of ground Willies are considered unknown, due to the obstacles and/or camouflage. Alternatively, the concept of protected zone [32] is adopted. Specifically, each transmit node knows a lower bound of its minimum possible distance to the nearest ground Willie from it, denoted by $d_{g,i,\min}$ $(i = 1, 2, \ldots, N_a)$.

3 Performance analyses

In this section, Bob's decoding performance and wardens' detection performance are analyzed, based on which the benefits of increasing the number of collaborative transmit nodes are analyzed.

3.1 Decoding performance at Bob

As per equation (1), the received signal-to-interference-plus-noise ratio (SINR) at Bob can be written as

$$\gamma_b = \frac{\left|\boldsymbol{h}_b^{\mathrm{T}} \boldsymbol{w}\right|^2}{d_b^{\alpha(\Theta_b)} \left(P_b \rho + \sigma_b^2\right)}. \tag{7}$$

Define the outage probability at Bob as $\delta \triangleq \Pr\left\{\gamma_b < 2^R - 1\right\}$, where R denotes the transmit rate. Then, because P_b follows $\mathrm{U}(0, P_{\max})$, with given $\boldsymbol{h}_b$, the outage probability can be given by

$$\delta(\boldsymbol{h}_b) = \min\left\{1, \left[1 - \frac{\frac{\left|\boldsymbol{h}_b^{\mathrm{T}} \boldsymbol{w}\right|^2}{d_b^{\alpha(\Theta_b)}(2^R - 1)} - \sigma_b^2}{P_{\max}\rho}\right]^+\right\}, \tag{8}$$

where $[x]^+$ denotes $\max\{x, 0\}$.

3.2 Detection performance at the wardens

Wardens' detection performance is evaluated based on the sum of false alarm probability P_{FA} and miss detection probability P_{MD}[3)]. Let $w \in \{g, c, f\}$ represent the types of the wardens. Then, it follows from (4)–(6) that each Willie's reception can be uniformly expressed as

$$\begin{aligned} H_0 &: y_{w,j}(n) = \sqrt{P_b} g_{w,j} v(n) + n_{w,j}(n), \\ H_1 &: y_{w,j}(n) = h_{w,j} x(n) + \sqrt{P_b} g_{w,j} v(n) + n_{w,j}(n), \end{aligned} \tag{9}$$

2) Because the far UAV Willies are in the far field, for each far UAV Willie, his antenna gains towards different transmit nodes are regarded as the same. To ensure covertness, Alice only needs to know an upper bound of W_f and a lower bound of w_f, which can be conservatively determined (e.g., according to the antenna aperture and antenna pattern models based on public information, intelligence, or reconnaissance).

3) In general, the wardens do not know the prior probabilities of the hypotheses (i.e., $\Pr\{H_0\}$ and $\Pr\{H_1\}$). Because $\Pr\{H_0\} P_{\mathrm{FA}} + \Pr\{H_1\} P_{\mathrm{MD}} \geqslant \min\{\Pr\{H_0\}, \Pr\{H_1\}\} (P_{\mathrm{FA}} + P_{\mathrm{MD}})$, regardless of the values of the prior probabilities, Willie's detection performance can be restricted by using $P_{\mathrm{FA}} + P_{\mathrm{MD}}$ as the metric [7, 10].

Herein, coefficients $h_{w,j}$ and $g_{w,j}$ represent different expressions for different types of Willies. For example, for ground Willie, we have $h_{w,j} \triangleq \sum_{i=1}^{N_a} \sqrt{\frac{P_i}{d_{g,i,j}^{\alpha(0)}}} h_{g,i,j} e^{\iota\varphi_i}$ and $g_{w,j} \triangleq \sqrt{\frac{1}{d_{g,b,j}^{\alpha\left(\Theta_{g,b,j}\right)}}} h_{g,b,j}$. From a conservative perspective, it is assumed that each Willie perfectly knows his noise variance, the distribution of P_b, coefficients $h_{w,j}$ and $g_{w,j}$[4]. Next, it follows from [10, Lemma 2] that the optimal detector at each Willie is a radiometer, which can be expressed as

$$\frac{\sum_{n=1}^{N} \left|y_{w,j}(n)\right|^2}{N} \underset{\mathcal{D}_0}{\overset{\mathcal{D}_1}{\gtrless}} t_{w,j}. \tag{10}$$

Herein, $t_{w,j}$ stands for the threshold awaiting to be determined. Notations $\mathcal{D}_0$ and $\mathcal{D}_1$ denote that Willie regards H_0 and H_1 as true, respectively. For conciseness, indices w and j are omitted in what follows.

Based on the central limit theorem [8], the sum of error probabilities of detector (10) is

$$P_{\mathrm{FA}} + P_{\mathrm{MD}} \approx \begin{cases} 0, & P_b |g|^2 + \sigma_w^2 \leqslant t \leqslant |h|^2 + P_b |g|^2 + \sigma_w^2 \triangleq \xi(P_b, t), \\ 1, & \text{otherwise.} \end{cases} \tag{11}$$

By taking Willie's interference uncertainty (the randomness of P_b) into account, Willie's detection performance should be evaluated by the expectation of $\xi(P_b, t)$ with respect to P_b, denoted by $E\{\xi(P_b, t)\}$.

Lemma 1. When $|h|^2 > P_{\max} |g|^2$, arbitrary value within $\left[P_{\max} |g|^2 + \sigma_w^2, |h|^2 + \sigma_w^2\right]$ is Willie's optimal threshold t^*. Otherwise, arbitrary value within $\left[|h|^2 + \sigma_w^2, P_{\max} |g|^2 + \sigma_w^2\right]$ is the optimal threshold. For both cases, Willie's optimal detection performance can be written as

$$\xi^* \triangleq \min_t E\{\xi(P_b, t)\} = \left[1 - \frac{|h|^2}{P_{\max} |g|^2}\right]^+. \tag{12}$$

Proof. Please refer to Appendix A.

It follows from (12) that Alice cannot ensure perfect covertness (i.e., $\xi^* = 1$) unless she keeps silence. Suppose that Alice expects Willie's detection performance satisfying $\xi^* \geqslant 1 - \epsilon$, or equivalently $|h|^2 \leqslant \epsilon P_{\max} |g|^2$, where $0 < \epsilon < 1$ denotes the tolerance. Then, as h is unknown and random from Alice's perspective, Alice cannot ensure $|h|^2 \leqslant \epsilon P_{\max} |g|^2$, either. Therefore, we define covert probability

$$P_{\mathrm{covert}} = \Pr\left\{|h|^2 \leqslant \epsilon P_{\max} |g|^2\right\} \tag{13}$$

as the metric for evaluating Willie's detection performance from Alice's perspective.

Proposition 1. For the j-th ground Willie, the covert probability can be determined by

$$P_{\mathrm{covert}} = 1 - \frac{\frac{\beta}{\varrho} e^{\frac{\beta}{\varrho}}}{K\left(\Theta_{g,b,j}\right) e^{K\left(\Theta_{g,b,j}\right)}}, \tag{14}$$

where we define $\beta \triangleq K\left(\Theta_{g,b,j}\right) \mathcal{K}$, $\varrho \triangleq \frac{\epsilon P_{\max}}{\sum_{i=1}^{N_a} \frac{P_i}{d_{g,i,j}^{\alpha(0)}}} + \mathcal{K}$, and $\mathcal{K} \triangleq d_{g,b,j}^{\alpha\left(\Theta_{g,b,j}\right)} \left(K\left(\Theta_{g,b,j}\right) + 1\right)$.

Proof. Please refer to Appendix B.

When the number of the transmit nodes is large enough, by utilizing the central limit theorem, it can be proven that Eq. (14) serves as a well approximation for the scenarios in which the ground-to-ground channels follow other distributions. This extends the applicability of the analyses and algorithm presented in what follows. Furthermore, recall that the number and the positions of ground Willies are unknown from Alice and Bob's perspective, which means that J_g, $d_{g,i,j}$, $d_{g,b,j}$, and $\Theta_{g,b,j}$ are unknown. As a result, Eq. (14) cannot be directly used for further analyses and parameter optimization. To address this, it is noted that due to the path loss and obstacles, we only

4) For each of these involved parameters, Willie may suffer from different levels of uncertainty, depending on the specific application scenarios. Thus, there are numerous specific situations to be discussed. Alternatively, it is assumed that Willie only suffers from the uncertainty about the random jamming power. This assumption is conservative and is universally applicable from Alice's perspective [33]. Specifically, because the resultant covert communication performance serves as a lower bound of the performance when the wardens know less, the performance analyses and optimization based on this assumption can ensure covertness for any scenarios where Willie suffers from more uncertainties.

need to consider the ground Willies surrounding Alice, whose detection performance serves as an upper bound of that of the ground Willies far away from Alice. As a result, the unknown parameters $d_{g,b,j}$ and $\Theta_{g,b,j}$ can be approximated by d_b and Θ_b, respectively, since Bob is in the far field of Alice. On the other hand, it is noted that Eq. (14) monotonically increases with $d_{g,i,j}$. Thus, we can replace $d_{g,i,j}$ by $d_{g,i,\min}$ to lower bound the covert probability, with the help of the protected zone. As a result, for ground Willies, we have

$$P_{\text{covert}} \geqslant 1 - \frac{\frac{\beta'}{\varrho'} e^{\frac{\beta'}{\varrho'}}}{K\left(\Theta_b\right) e^{K(\Theta_b)}}, \tag{15}$$

where we define $\beta' \triangleq K\left(\Theta_b\right)\mathcal{K}'$, $\varrho' \triangleq \frac{\epsilon P_{\max}}{\sum_{i=1}^{N_a} \frac{P_i}{d_{g,i,\min}^{\alpha(0)}}} + \mathcal{K}'$, and $\mathcal{K}' \triangleq d_b^{\alpha(\Theta_b)}\left(K\left(\Theta_b\right)+1\right)$.

Proposition 2. For the j-th near UAV Willie, the accurate covert probability is

$$P_{\text{covert}} = \frac{1}{2}\int_0^{\frac{\epsilon P_{\max}}{d_{c,b,j}^2}} \int_0^{\infty} J_0\left(v\sqrt{x}\right) \exp\left(-\frac{v^2\sigma^2}{2}\right) \left[\prod_{i=1}^{N_a} J_0\left(V_i v\right)\right] v dv dx, \tag{16}$$

where we define $V_i \triangleq \sqrt{\frac{P_i}{d_{c,i,j}^{\alpha(\Theta_{c,i,j})}}}\sqrt{\frac{K(\Theta_{c,i,j})}{K(\Theta_{c,i,j})+1}}$ and $2\sigma^2 \triangleq \sum_{i=1}^{N_a} \frac{P_i}{d_{c,i,j}^{\alpha(\Theta_{c,i,j})}(K(\Theta_{c,i,j})+1)}$. Furthermore, the covert probability can be well approximated by

$$P_{\text{covert}} \approx 1 - \exp\left(-\frac{\epsilon P_{\max}}{d_{c,b,j}^2 \sum_{i=1}^{N_a} \frac{P_i}{d_{c,i,j}^{\alpha(\Theta_{c,i,j})}}}\right). \tag{17}$$

Proof. Please refer to Appendix C.

Proposition 3. For the j-th far UAV Willie, the covert probability can be determined by

$$P_{\text{covert}} = 1 - Q_1\left(\sqrt{2K\left(\Theta_{f,j}\right)\left|\bar{\boldsymbol{h}}_{f,j}^{\mathrm{T}}\frac{\boldsymbol{w}}{||\boldsymbol{w}||}\right|^2}, \sqrt{\frac{2\epsilon P_{\max}\left(K\left(\Theta_{f,j}\right)+1\right)w_f d_{f,j}^{\alpha(\Theta_{f,j})}}{W_f d_{f,b,j}^2 ||\boldsymbol{w}||^2}}\right), \tag{18}$$

where $Q_1\left(a,b\right) = \int_b^{\infty} x \exp\left(-\frac{x^2+a^2}{2}\right) I_0\left(ax\right) dx$ is the Marcum Q function.

Proof. Please refer to Appendix D.

Remark 1. Different principles apply for the countermeasures against different types of wardens. Specifically, as per [34], $Q_1\left(a,b\right)$ monotonically increases with a for all $a \geqslant 0$ and $b > 0$. In addition, it monotonically decreases with b for all $a, b \geqslant 0$. Thus, it follows from (18) that Alice should reduce $\left|\bar{\boldsymbol{h}}_{f,j}^{\mathrm{T}}\frac{\boldsymbol{w}}{||\boldsymbol{w}||}\right|$ and $||\boldsymbol{w}||$ for confusing the far UAV Willies, which means that she should construct null lobes towards far UAV Willies and reduce the total transmit power. In comparison, for near UAV Willies, it follows from (17) that Alice should adjust each transmit node's transmit power according to their respective large-scale path losses from the near UAV Willies. For ground Willies, it follows from (15) that Alice should adjust each transmit node's transmit power according to the sizes of their respective protected zones. In conclusion, it is not straightforward to conceive a single transmission strategy under the hierarchical surveillance of the three types of wardens with multiple wardens of each type. In Section 4, the joint optimization of key system parameters will be presented to solve this problem.

3.3 The benefits of increasing the number of collaborative transmit nodes

We firstly demonstrate the benefits of increasing N_a in the absence of far UAV Willies, by focusing on the large-N_a region. For such, a specific collaborative beamforming strategy is constructed as

$$\sqrt{P_i}e^{\iota\varphi_i} = \frac{\sqrt{P_{gc}}}{N_a}\frac{\left(\left[\boldsymbol{h}_b\right]_i\right)^*}{\left|\left[\boldsymbol{h}_b\right]_i\right|}, \forall i = 1, 2, \ldots, N_a. \tag{19}$$

Herein, $P_{gc} \triangleq \frac{d_b^{\alpha(\Theta_b)}\left(P_{\max}\rho+\sigma_b^2\right)\left(2^R-1\right)+\Delta}{{C_{gc}}^2}$ is a constant, where $\Delta > 0$ can be an arbitrary positive constant, whereas $C_{gc} \triangleq \sqrt{\frac{\pi}{K(\Theta_b)+1}}\frac{e^{-\frac{K(\Theta_b)}{2}}}{2} \times \left[\left(1+K\left(\Theta_b\right)\right)I_0\left(\frac{K(\Theta_b)}{2}\right) + K\left(\Theta_b\right)I_1\left(\frac{K(\Theta_b)}{2}\right)\right]$. Based on this beamforming strategy,

it follows from (7) that Bob's received SINR is given by

$$\gamma_b = \frac{P_{gc}}{d_b^{\alpha(\Theta_b)}\left(P_b\rho+\sigma_b^2\right)}\left(\frac{\sum_{i=1}^{N_a}\left|[\boldsymbol{h}_b]_i\right|}{N_a}\right)^2. \tag{20}$$

Note that $|[\boldsymbol{h}_b]_i|$ independently and identically follows the Rice distribution for different i. Thus, as $N_a\to\infty$, it follows from the law of large numbers that

$$\frac{\sum_{i=1}^{N_a}\left|[\boldsymbol{h}_b]_i\right|}{N_a} \overset{\text{p1}}{=} E\left\{\left|[\boldsymbol{h}_b]_i\right|\right\} = C_{gc}. \tag{21}$$

Combining the foregoing results, we can conclude that Bob's decoding requirement $\gamma_b > 2^R-1$ is ensured in the large-N_a region.

Next, we turn to consider the covert performance. It follows from (19) that

$$\varrho' = \frac{\epsilon P_{\max}}{\frac{P_{gc}}{N_a^2}\sum_{i=1}^{N_a}\frac{1}{d_{g,i,\min}^{\alpha(0)}}} + \mathcal{K}' \geqslant N_a\frac{\epsilon P_{\max}\min_{i=1,2,\ldots,N_a}\left\{d_{g,i,\min}^{\alpha(0)}\right\}}{P_{gc}} + \mathcal{K}'. \tag{22}$$

As $N_a\to\infty$, it follows from (22) that $\varrho'\to\infty$. As a result, the RHS of (15) approaches one. Similarly, one can readily show that the RHS of (17) approaches one as $N_a\to\infty$ by noting that

$$\sum_{i=1}^{N_a}\frac{P_i}{d_{c,i,j}^{\alpha(\Theta_{c,i,j})}} = \frac{P_{gc}}{N_a^2}\sum_{i=1}^{N_a}\frac{1}{d_{c,i,j}^{\alpha(\Theta_{c,i,j})}} \leqslant \frac{P_{gc}}{N_a\min_{i=1,2,\ldots,N_a}\left\{d_{c,i,j}^{\alpha(\Theta_{c,i,j})}\right\}}. \tag{23}$$

In conclusion, by adopting the collaborative beamforming strategy in (19) and by increasing N_a, the covert performance for ground Willies and near UAV Willies can be arbitrarily enhanced and meanwhile Bob's decoding threshold is satisfied.

Now we turn to the scenarios with far UAV Willies only, which is much more complicated due to the involvement of the steering vectors. To obtain tractable results, hereafter and until the end of this section, it is additionally assumed that the transmit nodes are located uniformly in a disc centered at the origin with radius R, whereas the rest of the paper is applicable regardless of the specific distribution of the transmit nodes. Then, a collaborative beamforming strategy is constructed as

$$\boldsymbol{w} = \frac{\sqrt{P_f}}{\sqrt{N_a\left(1-\boldsymbol{g}^{\mathrm{H}}\boldsymbol{E}^{-1}\boldsymbol{g}\right)}}\left(\bar{\boldsymbol{h}}_b - \boldsymbol{A}\boldsymbol{E}^{-1}\boldsymbol{g}\right)^*, \tag{24}$$

where we define $\boldsymbol{A}\triangleq\left[\bar{\boldsymbol{h}}_{f,1}\ \bar{\boldsymbol{h}}_{f,2}\ \cdots\ \bar{\boldsymbol{h}}_{f,J_f}\right]$, whereas constant $P_f>0$ will be elaborated in what follows. Herein, matrix $\boldsymbol{E}\in\mathbb{R}^{J_f\times J_f}$ and vector $\boldsymbol{g}\in\mathbb{R}^{J_f\times 1}$ are defined, respectively, as

$$[\boldsymbol{E}]_{m,n}\triangleq\begin{cases}\frac{2}{\alpha_{m,n}}J_1\left(\alpha_{m,n}\right), & m\neq n,\\ 1, & m=n,\end{cases} \tag{25}$$

$$[\boldsymbol{g}]_m\triangleq\frac{2}{\alpha_m}J_1\left(\alpha_m\right), \tag{26}$$

where $\alpha_{m,n}$ and α_m are defined as

$$\alpha_{m,n}\triangleq\frac{2\pi R}{\lambda}\sqrt{\left(\cos\left(\Theta_{f,n}\right)\cos\left(\phi_{f,n}\right)-\cos\left(\Theta_{f,m}\right)\cos\left(\phi_{f,m}\right)\right)^2+\left(\cos\left(\Theta_{f,n}\right)\sin\left(\phi_{f,n}\right)-\cos\left(\Theta_{f,m}\right)\sin\left(\phi_{f,m}\right)\right)^2}, \tag{27}$$

$$\alpha_m\triangleq\frac{2\pi R}{\lambda}\sqrt{\left(\cos\left(\Theta_b\right)\cos\left(\phi_b\right)-\cos\left(\Theta_{f,m}\right)\cos\left(\phi_{f,m}\right)\right)^2+\left(\cos\left(\Theta_b\right)\sin\left(\phi_b\right)-\cos\left(\Theta_{f,m}\right)\sin\left(\phi_{f,m}\right)\right)^2}. \tag{28}$$

Next, by generalizing the derivations in [24] from 2D scenarios to 3D scenarios, one can show that

$$\lim_{N_a\to\infty}\left|\bar{\boldsymbol{h}}_{f,j}^{\mathrm{T}}\boldsymbol{w}\right|^2 \overset{\text{p1}}{=} E\left\{\left|\bar{\boldsymbol{h}}_{f,j}^{\mathrm{T}}\boldsymbol{w}\right|^2\right\} = P_f, j=1,2,\ldots,J_f, \tag{29}$$

$$\lim_{N_a\to\infty}\left|\bar{\boldsymbol{h}}_b^{\mathrm{T}}\boldsymbol{w}\right|^2\overset{\mathrm{p1}}{=}E\left\{\left|\bar{\boldsymbol{h}}_b^{\mathrm{T}}\boldsymbol{w}\right|^2\right\}=N_aP_f\left(1-\boldsymbol{g}^{\mathrm{T}}\boldsymbol{E}^{-1}\boldsymbol{g}\right),\tag{30}$$

$$\lim_{N_a\to\infty}||\boldsymbol{w}||^2\overset{\mathrm{p1}}{=}E\left\{||\boldsymbol{w}||^2\right\}=P_f.\tag{31}$$

As a result, for the far UAV Willies, it follows from (18) that

$$\lim_{N_a\to\infty}P_{\mathrm{covert}}\overset{\mathrm{p1}}{=}1-Q_1\left(\sqrt{2K\left(\Theta_{f,j}\right)},\sqrt{\frac{2\epsilon P_{\max}\left(K\left(\Theta_{f,j}\right)+1\right)w_fd_{f,j}^{\alpha(\Theta_{f,j})}}{W_fd_{f,b,j}^2P_f}}\right).\tag{32}$$

Since $\lim_{b\to\infty}Q_1(a,b)=0$ for $a\geqslant 0$, it follows from (32) that, in the large-N_a region, based on the constructed collaborative beamforming strategy (24), the covert probability can be arbitrarily close to one by adopting a small enough P_f, whereas N_a has no impacts on the covert probability.

Next, we consider Bob's outage probability. Since $P_b\leqslant P_{\max}$, we have

$$\begin{aligned}\delta&\leqslant\Pr\left\{\frac{\left|\boldsymbol{h}_b^{\mathrm{T}}\boldsymbol{w}\right|^2}{d_b^{\alpha(\Theta_b)}\left(P_{\max}\rho+\sigma_b^2\right)}<2^R-1\right\}\triangleq\bar{\delta}\\&=1-Q_1\left(\sqrt{2K\left(\Theta_b\right)\left|\bar{\boldsymbol{h}}_b^{\mathrm{T}}\frac{\boldsymbol{w}}{||\boldsymbol{w}||}\right|^2},\sqrt{\frac{2\left(K\left(\Theta_b\right)+1\right)d_b^{\alpha(\Theta_b)}\left(P_{\max}\rho+\sigma_b^2\right)\left(2^R-1\right)}{||\boldsymbol{w}||^2}}\right),\end{aligned}\tag{33}$$

where the second step can be proven by following a similar derivation process as that in Appendix D. Then, it follows from (30), (31), and (33) that

$$\lim_{N_a\to\infty}\bar{\delta}\overset{\mathrm{p1}}{=}1-\lim_{N_a\to\infty}Q_1\left(\sqrt{2K\left(\Theta_b\right)N_a\left(1-\boldsymbol{g}^{\mathrm{T}}\boldsymbol{E}^{-1}\boldsymbol{g}\right)},\sqrt{\frac{2\left(K\left(\Theta_b\right)+1\right)d_b^{\alpha(\Theta_b)}\left(P_{\max}\rho+\sigma_b^2\right)\left(2^R-1\right)}{P_f}}\right).\tag{34}$$

Note that herein $\boldsymbol{g}^{\mathrm{T}}\boldsymbol{E}^{-1}\boldsymbol{g}$ is a constant determined by the azimuth angles and the elevation angles of far UAV Willies and Bob, as per (25)–(28).

Thus, we can conclude that $\lim_{N_a\to\infty}\delta\to 0$ for arbitrarily given P_f. Combining this conclusion with (32), it is concluded that, in the large-N_a region, based on the constructed collaborative beamforming strategy (24), an arbitrarily given covert probability requirement against far UAV Willies can be satisfied and meanwhile Bob's decoding performance can be arbitrarily enhanced, by adopting a small enough P_f and then by increasing N_a.

Until now, by focusing on the scenarios with only one or two warden types and by focusing on the large-N_a region, the benefits of increasing N_a have been analyzed to demonstrate the potential of collaborative beamforming for covert communications. In what follows, key system parameters are jointly designed, where all three warden types are jointly considered and meanwhile the whole N_a region is considered.

4 Parameter optimization against hierarchical surveillance

4.1 Problem formulation and conversion

This section aims to minimize Bob's outage probability while satisfying the covert constraint $P_{\mathrm{covert}}\geqslant 1-\chi$ for all the wardens, where $0<\chi<1$ reflects the required level of covertness. To begin with, define $\hat{P}_i$ and $\hat{P}_{\max}$ as the transmit power limits of the i-th transmit node and Bob, respectively. According to expressions (8), (15), and propositions 2 and 3, the optimization problem is formulated as

$$\begin{aligned}\mathrm{P0}:\ &\max_{\boldsymbol{w},P_{\max}}\ \frac{\left[\frac{\left|\boldsymbol{h}_b^{\mathrm{T}}\boldsymbol{w}\right|^2}{d_b^{\alpha(\Theta_b)}\left(2^R-1\right)}-\sigma_b^2\right]^+}{P_{\max}\rho}\\&\text{s.t.}\ \ \text{(a)}\ 0\leqslant P_i\leqslant\hat{P}_i,i=1,2,\ldots,N_a,\\&\qquad\text{(b)}\ 0\leqslant P_{\max}\leqslant\hat{P}_{\max},\end{aligned}$$

$$\text{(c)}\ \frac{\beta'}{\frac{\epsilon P_{\max}}{\sum_{i=1}^{N_a}\frac{P_i}{d_{g,i,\min}^{\alpha(0)}}}+\mathcal{K}'}\exp\left(\frac{\beta'}{\frac{\epsilon P_{\max}}{\sum_{i=1}^{N_a}\frac{P_i}{d_{g,i,\min}^{\alpha(0)}}}+\mathcal{K}'}\right)\leqslant \chi K\left(\Theta_b\right)e^{K(\Theta_b)},$$

$$\text{(d)}\ \exp\left(-\frac{\epsilon P_{\max}}{d_{c,b,j}^2\sum_{i=1}^{N_a}\frac{P_i}{d_{c,i,j}^{\alpha(\Theta_{c,i,j})}}}\right)\leqslant \chi, j=1,2,\ldots,J_c,$$

$$\text{(e)}\ Q_1\left(\sqrt{2K\left(\Theta_{f,j}\right)\left|\bar{\boldsymbol{h}}_{f,j}^{\mathrm{T}}\frac{\boldsymbol{w}}{||\boldsymbol{w}||}\right|^2},\sqrt{\frac{2\epsilon P_{\max}\left(K\left(\Theta_{f,j}\right)+1\right)w_f d_{f,j}^{\alpha(\Theta_{f,j})}}{W_f d_{f,b,j}^2||\boldsymbol{w}||^2}}\right)\leqslant \chi, j=1,2,\ldots,J_f. \tag{35}$$

To simplify P0, note that constraint (c) can be rewritten as $P_{\max}\geqslant\frac{\left(\frac{K(\Theta_b)}{x_g^*}-1\right)d_b^{\alpha(\Theta_b)}(K(\Theta_b)+1)}{\epsilon}\sum_{i=1}^{N_a}\frac{P_i}{d_{g,i,\min}^{\alpha(0)}}$, where x_g^* is the only solution to equation $xe^x=\chi K\left(\Theta_b\right)e^{K(\Theta_b)}$ with respect to x. Note that x_g^* can be efficiently obtained by using the bisection method and inequality $x_g^*<K\left(\Theta_b\right)$ definitely holds. Furthermore, one can readily rewrite constraint (d) as $P_{\max}\geqslant\frac{\ln\left(\frac{1}{\chi}\right)d_{c,b,j}^2}{\epsilon}\sum_{i=1}^{N_a}\frac{P_i}{d_{c,i,j}^{\alpha(\Theta_{c,i,j})}}, j=1,2,\ldots,J_c$.

Unfortunately, constraint (e) involves the Marcum Q function and thus is difficult to address. To address this, it is noted that in constraint (e), optimization variable $\boldsymbol{w}$ appears in the forms of $\frac{\boldsymbol{w}}{||\boldsymbol{w}||}$ and $||\boldsymbol{w}||^2$. Particularly, $\frac{\boldsymbol{w}}{||\boldsymbol{w}||}$ only appears in the first component of the Marcum Q function, whereas $||\boldsymbol{w}||^2$ as well as optimization variable $P_{\max}$ only appear in the second component. This motivates us to decompose optimization variable $\boldsymbol{w}$ into optimization variables $\boldsymbol{w}_d\triangleq\frac{\boldsymbol{w}}{||\boldsymbol{w}||}\triangleq[w_{d,1},w_{d,2},\ldots,w_{d,N_a}]^{\mathrm{T}}$ and $P_{\text{total}}\triangleq||\boldsymbol{w}||^2$, named collaborative beamforming direction and total transmit power, respectively. Corresponding, we have $\boldsymbol{w}=\sqrt{P_{\text{total}}}\boldsymbol{w}_d$ and $P_i=P_{\text{total}}\left|w_{d,i}\right|^2$. Then, P0 can be rewritten as

$$\begin{aligned}
\text{P1}:\ &\max_{\boldsymbol{w}_d,P_{\max},P_{\text{total}}}\ \frac{\left[\frac{P_{\text{total}}\left|\boldsymbol{h}_b^{\mathrm{T}}\boldsymbol{w}_d\right|^2}{d_b^{\alpha(\Theta_b)}(2^R-1)}-\sigma_b^2\right]^+}{P_{\max}\rho}\\
\text{s.t.}\ &\text{(a)}\ 0\leqslant P_{\text{total}}\left|w_{d,i}\right|^2\leqslant\hat{P}_i, i=1,2,\ldots,N_a,\\
&\text{(b)}\ 0\leqslant P_{\max}\leqslant\hat{P}_{\max},\\
&\text{(c)}\ P_{\max}\geqslant P_{\text{total}}\frac{\left(\frac{K(\Theta_b)}{x_g^*}-1\right)d_b^{\alpha(\Theta_b)}\left(K\left(\Theta_b\right)+1\right)}{\epsilon}\sum_{i=1}^{N_a}\frac{\left|w_{d,i}\right|^2}{d_{g,i,\min}^{\alpha(0)}},\\
&\text{(d)}\ P_{\max}\geqslant P_{\text{total}}\frac{\ln\left(\frac{1}{\chi}\right)d_{c,b,j}^2}{\epsilon}\sum_{i=1}^{N_a}\frac{\left|w_{d,i}\right|^2}{d_{c,i,j}^{\alpha(\Theta_{c,i,j})}}, j=1,2,\ldots,J_c,\\
&\text{(e)}\ Q_1\left(\sqrt{2K\left(\Theta_{f,j}\right)\left|\bar{\boldsymbol{h}}_{f,j}^{\mathrm{T}}\boldsymbol{w}_d\right|^2},\sqrt{\frac{2\epsilon P_{\max}\left(K\left(\Theta_{f,j}\right)+1\right)w_f d_{f,j}^{\alpha(\Theta_{f,j})}}{W_f d_{f,b,j}^2 P_{\text{total}}}}\right)\leqslant\chi, j=1,2,\ldots,J_f,\\
&\text{(f)}\ ||\boldsymbol{w}_d||^2=1.
\end{aligned}\tag{36}$$

Now, the two components of the Marcum Q function in constraint (e) do not involve the same optimization variables. Specifically, $\boldsymbol{w}_d$ only appears in the first component, whereas $P_{\max}$ and P_{total} only appear in the second component. This facilitates us to optimize $\boldsymbol{w}_d$ and $\{P_{\max},P_{\text{total}}\}$ alternately, with the help of the monotonicity properties of the Marcum Q function, which is elaborated in what follows.

4.2 Optimize $\{P_{\max},P_{\text{total}}\}$ with given $\boldsymbol{w}_d$

It is noted that the Marcum Q function $Q_1\left(a,b\right)$ is strictly decreasing in b for all $a,b\geqslant 0$ [34]. As a result, with given $\boldsymbol{w}_d$, constraint (e) of P1 can be rewritten as

$$\frac{2\epsilon P_{\max}\left(K\left(\Theta_{f,j}\right)+1\right)w_f d_{f,j}^{\alpha(\Theta_{f,j})}}{W_f d_{f,b,j}^2 P_{\text{total}}}\geqslant x_{P,j}^*, j=1,2,\ldots,J_f, \tag{37}$$

where $x_{P,j}^*$ is the only solution to equation

$$Q_1\left(\sqrt{2K\left(\Theta_{f,j}\right)\left|\bar{\boldsymbol{h}}_{f,j}^{\mathrm{T}}\boldsymbol{w}_d\right|^2},\sqrt{x}\right)=\chi \tag{38}$$

with respect to x and can be determined efficiently by the bisection method. Furthermore, we define

$$\mathrm{MAX}_c \triangleq \max_{j=1,2,\ldots,J_c}\left\{\frac{\ln\left(\frac{1}{\chi}\right)d_{c,b,j}^2}{\epsilon}\sum_{i=1}^{N_a}\frac{|w_{d,i}|^2}{d_{c,i,j}^{\alpha(\Theta_{c,i,j})}}\right\}, \tag{39}$$

$$\mathrm{MAX}_f \triangleq \max_{j=1,2,\ldots,J_f}\left\{\frac{W_f d_{f,b,j}^2 x_{P,j}^*}{2\epsilon\left(K\left(\Theta_{f,j}\right)+1\right)w_f d_{f,j}^{\alpha(\Theta_{f,j})}}\right\}, \tag{40}$$

$$\mathrm{MAX}_w \triangleq \max\left\{\mathrm{MAX}_c,\mathrm{MAX}_f,\frac{\left(\frac{K(\Theta_b)}{x_g^*}-1\right)d_b^{\alpha(\Theta_b)}\left(K\left(\Theta_b\right)+1\right)}{\epsilon}\sum_{i=1}^{N_a}\frac{|w_{d,i}|^2}{d_{g,i,\min}^{\alpha(0)}}\right\}. \tag{41}$$

Then, with given $\boldsymbol{w}_d$, constraints (b)–(e) of P1 can be combined as $P_{\mathrm{total}}\mathrm{MAX}_w \leqslant P_{\max} \leqslant \hat{P}_{\max}$. As a result, the subproblem of optimizing $\{P_{\max}, P_{\mathrm{total}}\}$ can be written as

$$\begin{aligned}\mathrm{SubP1}: \max_{P_{\max},P_{\mathrm{total}}}\ & \frac{\left[\frac{P_{\mathrm{total}}\left|\boldsymbol{h}_b^{\mathrm{T}}\boldsymbol{w}_d\right|^2}{d_b^{\alpha(\Theta_b)}\left(2^R-1\right)}-\sigma_b^2\right]^+}{P_{\max}\rho}\\ \text{s.t.}\ \ & \text{(a)}\ 0\leqslant P_{\mathrm{total}}\leqslant \mathrm{MIN}_P,\\ & \text{(b)}\ P_{\mathrm{total}}\mathrm{MAX}_w\leqslant P_{\max}\leqslant \hat{P}_{\max},\end{aligned} \tag{42}$$

where we define $\mathrm{MIN}_P \triangleq \min_{i=1,2,\ldots,N_a}\frac{\hat{P}_i}{|w_{d,i}|^2}$. Note that problem SubP1 definitely has feasible solution $\{P_{\max}, P_{\mathrm{total}}\} = \{0,0\}$.

Next, because the objective function of SubP1 is nonincreasing in $P_{\max}$, for arbitrarily given P_{total}, the optimal $P_{\max}$ is

$$P_{\max}^* = P_{\mathrm{total}}\mathrm{MAX}_w. \tag{43}$$

Inserting (43) into (42), we can arrive at

$$\begin{aligned}\mathrm{SubP1(a)}: \max_{P_{\mathrm{total}}}\ & \left[\frac{\left|\boldsymbol{h}_b^{\mathrm{T}}\boldsymbol{w}_d\right|^2}{\rho\mathrm{MAX}_w d_b^{\alpha(\Theta_b)}\left(2^R-1\right)}-\frac{\sigma_b^2}{\rho\mathrm{MAX}_w P_{\mathrm{total}}}\right]^+\\ \text{s.t.}\ \ & \text{(a)}\ 0\leqslant P_{\mathrm{total}}\leqslant \mathrm{MIN}_P,\\ & \text{(b)}\ P_{\mathrm{total}}\mathrm{MAX}_w\leqslant \hat{P}_{\max}.\end{aligned} \tag{44}$$

Because the objective function of SubP1(a) is nondecreasing in P_{total}, the optimal P_{total} is

$$P_{\mathrm{total}}^* = \begin{cases}\mathrm{MIN}_P, & \mathrm{MIN}_P\mathrm{MAX}_w\leqslant \hat{P}_{\max},\\ \frac{\hat{P}_{\max}}{\mathrm{MAX}_w}, & \text{otherwise.}\end{cases} \tag{45}$$

Remark 2. A larger signal power P_{total} enhances transmission reliability but requires a larger interference power $P_{\max}$ to ensure covertness, which in turn leads to larger self-interference at Bob. Thus, it is not straightforward to know the optimal setup of P_{total}. Eqs. (43) and (45) show that, for any given self-interference coefficient ρ and beamforming direction $\boldsymbol{w}_d$, P_{total} should be as large as possible, on the premise that $P_{\max}$ is set as the corresponding minimum required value that satisfies the covert constraints.

4.3 Optimize $\boldsymbol{w}_d$ with given $\{\boldsymbol{P}_{\max}, \boldsymbol{P}_{\text{total}}\}$

It is noted that the Marcum Q function $Q_1(a,b)$ is strictly increasing in a for all $a \geqslant 0$ and $b > 0$ [34]. As a result, with given $\{P_{\max}, P_{\text{total}}\}$, constraint (e) of P1 can be rewritten as $\left|\bar{\boldsymbol{h}}_{f,j}^{\mathrm{T}} \boldsymbol{w}_d\right|^2 \leqslant x_{d,j}^*, j = 1,2,\ldots,J_f$, where $x_{d,j}^*$ is the only solution to equation

$$Q_1\left(\sqrt{2K\left(\Theta_{f,j}\right)x}, \sqrt{\frac{2\epsilon P_{\max}\left(K\left(\Theta_{f,j}\right)+1\right)w_f d_{f,j}^{\alpha(\Theta_{f,j})}}{W_f d_{f,b,j}^2 P_{\text{total}}}}\right) = \chi$$

with respect to x and can be determined efficiently by the bisection method. Furthermore, note that in P1, variable $\boldsymbol{w}_d$ appears in the form of either $\left|\boldsymbol{h}^{\mathrm{T}}\boldsymbol{w}_d\right|^2$, $||\boldsymbol{w}_d||^2$, or $|w_{d,i}|^2$. This motivates us to use the semidefinite relaxation (SDR) technique to solve the subproblem of optimizing $\boldsymbol{w}_d$. Specifically, we define $\mathbf{W} \triangleq \boldsymbol{w}_d\boldsymbol{w}_d^{\mathrm{H}}$, $\mathbf{H}_b \triangleq \left(\boldsymbol{h}_b\boldsymbol{h}_b^{\mathrm{H}}\right)^{\mathrm{T}}$, and $\bar{\mathbf{H}}_{f,j} \triangleq \left(\bar{\boldsymbol{h}}_{f,j}\bar{\boldsymbol{h}}_{f,j}^{\mathrm{H}}\right)^{\mathrm{T}}$. Then, we have $\left|\boldsymbol{h}_b^{\mathrm{T}}\boldsymbol{w}_d\right|^2 = \boldsymbol{w}_d^{\mathrm{H}}\left(\boldsymbol{h}_b\boldsymbol{h}_b^{\mathrm{H}}\right)^{\mathrm{T}}\boldsymbol{w}_d = \mathrm{tr}\left(\mathbf{H}_b\mathbf{W}\right)$ and $\left|\bar{\boldsymbol{h}}_{f,j}^{\mathrm{T}}\boldsymbol{w}_d\right|^2 = \mathrm{tr}\left(\bar{\mathbf{H}}_{f,j}\mathbf{W}\right)$. Also, we have $|w_{d,i}|^2 \triangleq [\mathbf{W}]_{i,i}$. As a result, after relaxing the rank-one constraint, the subproblem can be formulated as

$$\begin{aligned}
\text{SubP2}: &\max_{\mathbf{W}} \mathrm{tr}\left(\mathbf{H}_b\mathbf{W}\right)\\
&\text{s.t. (a) } 0 \leqslant P_{\text{total}}[\mathbf{W}]_{i,i} \leqslant \hat{P}_i, i = 1,2,\ldots,N_a,\\
&\quad\text{(b) } P_{\max} \geqslant P_{\text{total}}\frac{\left(\frac{K(\Theta_b)}{x_g^*}-1\right)d_b^{\alpha(\Theta_b)}\left(K\left(\Theta_b\right)+1\right)}{\epsilon}\sum_{i=1}^{N_a}\frac{[\mathbf{W}]_{i,i}}{d_{g,i,\min}^{\alpha(0)}},\\
&\quad\text{(c) } P_{\max} \geqslant P_{\text{total}}\frac{\ln\left(\frac{1}{\chi}\right)d_{c,b,j}^2}{\epsilon}\sum_{i=1}^{N_a}\frac{[\mathbf{W}]_{i,i}}{d_{c,i,j}^{\alpha(\Theta_{c,i,j})}}, j = 1,2,\ldots,J_c,\\
&\quad\text{(d) } \mathrm{tr}\left(\bar{\mathbf{H}}_{f,j}\mathbf{W}\right) \leqslant x_{d,j}^*, j = 1,2,\ldots,J_f, \quad \text{(e) } \mathrm{tr}\left(\mathbf{W}\right) = 1, \quad \text{(f) } \mathbf{W} \succcurlyeq 0,
\end{aligned} \tag{46}$$

which is a convex optimization problem and can be efficiently solved by the CVX tool.

The solution $\mathbf{W}^*$ obtained by solving SubP2 is generally not rank-one. The Gaussian randomization method [35] is invoked to recover $\boldsymbol{w}_d$ from $\mathbf{W}^*$. Specifically, multiple candidate collaborative beamforming directions $\boldsymbol{w}_r = \mathbf{U}\boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{r}/\left\|\mathbf{U}\boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{r}\right\|$ are generated based on eigenvalue decomposition $\mathbf{W}^* = \mathbf{U}\boldsymbol{\Sigma}\mathbf{U}^{\mathrm{H}}$, where random vector $\boldsymbol{r}$ follows $\mathcal{CN}\left(\mathbf{0}, \boldsymbol{I}_{N_a}\right)$. Next, note that with given $\{P_{\max}, P_{\text{total}}\}$, a specific candidate direction $\boldsymbol{w}_r$ does not necessarily satisfy constraints (a)–(d) of SubP2. Also, it is noted that (a) it may happen that $\{\boldsymbol{w}_r, P_{\max}, P_{\text{total}}\}$ merely slightly violate the constraints but lead to good performance; (b) it may happen that $\{\boldsymbol{w}_r, P_{\max}, P_{\text{total}}\}$ lead to bad performance but $\boldsymbol{w}_r$ is still a potential collaborative beamforming direction because there is a large margin for the constraints. Therefore, to fully utilize all the obtained $\boldsymbol{w}_r$, the alternating optimization of $\{P_{\max}, P_{\text{total}}\}$ is incorporated into the selection of $\boldsymbol{w}_d$. Specifically, for each $\boldsymbol{w}_r$, Eqs. (43) and (45) are applied to determine the corresponding optimal $\{P_{\max}, P_{\text{total}}\}$, denoted by $\{P_{\max}^*(\boldsymbol{w}_r), P_{\text{total}}^*(\boldsymbol{w}_r)\}$. Then, the $\{\boldsymbol{w}_r, P_{\max}^*(\boldsymbol{w}_r), P_{\text{total}}^*(\boldsymbol{w}_r)\}$ that maximizes the objective function of P1 is selected as the solution. Note that each $\{\boldsymbol{w}_r, P_{\max}^*(\boldsymbol{w}_r), P_{\text{total}}^*(\boldsymbol{w}_r)\}$ is definitely a feasible solution, as per the analyses in Subsection 4.2.

4.4 Overall algorithm and complexity analyses

Before alternating optimization, an initial $\boldsymbol{w}_d$ is required. Recall that P1 is definitely feasible since we can set $\{P_{\max}, P_{\text{total}}\} = \{0,0\}$. Therefore, the initial $\boldsymbol{w}_d$ indeed can be randomly generated. Alternatively, in order to start the alternating optimization with a good initial solution, the initial $\boldsymbol{w}_d$ is generated by

$$\boldsymbol{w}_{d,\text{initial}} = \begin{cases} \frac{\boldsymbol{h}_b^*}{||\boldsymbol{h}_b||}, & N_a \leqslant J_f,\\ \frac{(\boldsymbol{I}-\boldsymbol{P}_{\boldsymbol{A}})(\boldsymbol{h}_b)^*}{||(\boldsymbol{I}-\boldsymbol{P}_{\boldsymbol{A}})(\boldsymbol{h}_b)^*||}, & \text{otherwise}, \end{cases} \tag{47}$$

where we define $\boldsymbol{P}_{\boldsymbol{A}} \triangleq \boldsymbol{A}^*\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}^*\right)^{-1}\boldsymbol{A}^{\mathrm{T}}$, with $\boldsymbol{A}$ defined as in (24). The explanation of the above proposed initial $\boldsymbol{w}_d$ is as follows. To begin with, note that it is difficult to optimize $\boldsymbol{w}_d$ before $\{P_{\max}, P_{\text{total}}\}$ have been determined. Alternatively, we resort to closed-form expression (47), which requires few computational resources and meanwhile

does not rely on the values of $\{P_{\max}, P_{\text{total}}\}$. To be more specific, the second case of the RHS of (47) is the optimal solution to [24]

$$\begin{aligned}&\text{P-Initial}: \max_{\boldsymbol{w}_d} \left|\boldsymbol{h}_b^{\mathrm{T}}\boldsymbol{w}_d\right|^2\\ &\quad\text{s.t.}\quad \text{(a) } \bar{\boldsymbol{h}}_{f,j}^{\mathrm{T}}\boldsymbol{w}_d = 0, j = 1,2,\ldots,J_f, \qquad \text{(b) } ||\boldsymbol{w}_d||^2 = 1.\end{aligned} \tag{48}$$

P-Initial maximizes Bob's received power on the premise of nulling the signals leaked to the far UAV Willies via the LoS links, whereas ground Willies and near UAV Willies are ignored. Note that the phases of the elements of $\boldsymbol{w}_d$ do not influence the covert constraints for ground Willies or near UAV Willies, whereas the covert constraints for far UAV Willies are influenced by both the phases and amplitudes of $\boldsymbol{w}_d$. Thus, the latter deserves priorities when $\boldsymbol{w}_d$ is initialized. In contrast, the first case of the RHS of (47) is the MRT strategy, which merely considers enhancing the transmission reliability and ignores the covertness. The first case is adopted when $N_a \leqslant J_f$ because P-Initial is infeasible when $N_a \leqslant J_f$.

Finally, the whole algorithm for solving P0 is summarized in Algorithm 1, where N_G and μ denote the Gaussian randomization times and the convergence tolerance, respectively, whereas $\delta\left(\boldsymbol{w}_d, P_{\max}, P_{\text{total}}\right)$ denotes the outage probability determined by solution $\{\boldsymbol{w}_d, P_{\max}, P_{\text{total}}\}$. Note that step 11 ensures the convergence of Algorithm 1 since the outage probability is lower bounded by zero [36–38].

Algorithm 1 Outage probability minimization algorithm.

1: Initialize iteration index of alternating optimization $t = 0$;
2: Initialize $\boldsymbol{w}_d^{(t)} = \boldsymbol{w}_{d,\text{initial}}$ according to (47); compute the corresponding $x_{P,j}^*$ by the bisection method as per (38);
3: Determine $\left\{P_{\max}^{(t)}, P_{\text{total}}^{(t)}\right\}$ as per (43) and (45) based on $\boldsymbol{w}_d = \boldsymbol{w}_{d,\text{initial}}$;
4: **repeat**
5: $t = t + 1$;
6: Generate $\mathbf{W}^*$ by solving SubP2 based on $\{P_{\max}, P_{\text{total}}\} = \left\{P_{\max}^{(t-1)}, P_{\text{total}}^{(t-1)}\right\}$;
7: Conduct eigenvalue decomposition $\mathbf{W}^* = \mathbf{U}\boldsymbol{\Sigma}\mathbf{U}^{\mathrm{H}}$;
8: Generate N_G candidate collaborative beamforming directions $\boldsymbol{w}_r = \mathbf{U}\boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{r}/\left|\left|\mathbf{U}\boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{r}\right|\right|$; compute the corresponding $x_{P,j}^*$ by the bisection method as per (38);
9: Determine $\{P_{\max}^*(\boldsymbol{w}_r), P_{\text{total}}^*(\boldsymbol{w}_r)\}$ as per (43) and (45) based on each $\boldsymbol{w}_r$;
10: Select the candidate solution $\{\boldsymbol{w}_r, P_{\max}^*(\boldsymbol{w}_r), P_{\text{total}}^*(\boldsymbol{w}_r)\}$ that maximizes the objective function of P1 as the solution of the t-th iteration $\left\{\boldsymbol{w}_d^{(t)}, P_{\max}^{(t)}, P_{\text{total}}^{(t)}\right\}$;
11: **until** $\delta\left(\boldsymbol{w}_d^{(t-1)}, P_{\max}^{(t-1)}, P_{\text{total}}^{(t-1)}\right) - \delta\left(\boldsymbol{w}_d^{(t)}, P_{\max}^{(t)}, P_{\text{total}}^{(t)}\right) < \mu\delta\left(\boldsymbol{w}_d^{(t-1)}, P_{\max}^{(t-1)}, P_{\text{total}}^{(t-1)}\right)$;
12: **return** $\left\{\boldsymbol{w}_d^{(t)}, P_{\max}^{(t)}, P_{\text{total}}^{(t)}\right\}$.

The computational complexity of Algorithm 1 is mainly due to steps 8–10, because the Gaussian randomization technique generally requires numerous realizations to obtain a relatively good solution [35]. Furthermore, note that the execution of the bisection method accounts for the major computational complexity among steps 8–10, because the Marcum Q function is basically a double integral, and because the remaining calculations have closed-form expressions and thus result in much smaller computational complexity. As a result, the computational complexity of Algorithm 1 can be given by $\mathcal{O}\left(TN_GJ_f\log\left(\bar{x}_P/\mu_P\right)\right)$, where T denotes the iteration times of the alternating optimization, whereas $[0, \bar{x}_P]$ and μ_P represent the range and the solution accuracy of the bisection method, respectively.

5 Numerical results and discussion

In this section, representative numerical results are presented to demonstrate the performance of the proposed algorithm. Without loss of generality, two near UAV Willies and two far UAV Willies are considered. The locations of the near UAV Willies are set as $(20, -20, 200)$ and $(-20, 20, 200)$ in meters. The locations of the far UAV Willies are set as $(100, -100, 1000)$ and $(-100, 100, 1000)$ in meters. Bob's location is set as $(100, 100, 1000)$ in meters. The locations of the transmit nodes are uniformly and randomly generated within a square with side length of 20 m centered at $(0, 0, 0)$; the specific setup is listed in Appendix E. Unless otherwise specified, the remaining parameters are set as $d_{g,i,\min} = 30$ m[5)], $\lambda = 0.1$ m, $W_f = 100$, $w_f = 0.01$, $N_G = 1000$, $\hat{P}_i = 1$ W, $\hat{P}_{\max} = 10$ W, $R = 1$ bit/s/Hz, $\epsilon = 0.01$, $\chi = 0.005$, $N_a = 10$, $\rho = -90$ dB [12], and $\sigma_b^2 = -110$ dBm. One important criterion of the parameter setup is that three types of covert constraints should be balanced, in the sense that none of them is much more stringent or loose than the others.

5) We do not need to specify the number and the coordinates of ground Willies because of the protected zone.

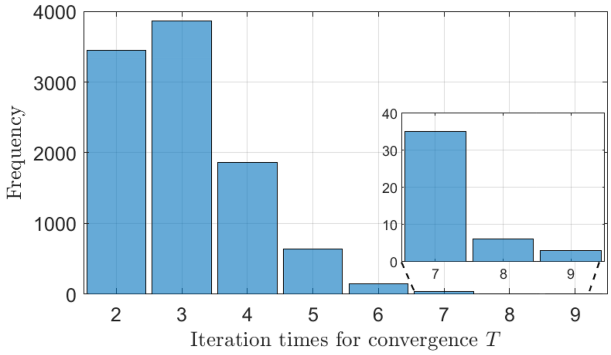

Figure 2 (Color online) Distribution of the required number of iterations for convergence.

For the path loss exponent and the Rician K-factor, we adopt the same models as those in [21]. Specifically, the path loss exponent is determined by $\alpha(\Theta) = a_1 P_{\mathrm{LoS}}(\Theta) + b_1$, where $P_{\mathrm{LoS}}(\Theta) = \frac{1}{1+a_2 e^{-b_2\Theta}}$ denotes the probability of LoS. Herein, a_1, a_2, b_1, and b_2 are the parameters determined by environment characteristics. Note that $a_1 \approx \alpha\left(\frac{\pi}{2}\right) - \alpha(0)$ and $b_1 \approx \alpha(0)$, because $P_{\mathrm{LoS}}(0) \approx 0$ and $P_{\mathrm{LoS}}\left(\frac{\pi}{2}\right) \approx 1$. Therefore, the values of a_1 and b_1 can be set by setting $\alpha\left(\frac{\pi}{2}\right)$ and $\alpha(0)$. The Rician K-factor is modeled as $K(\Theta) = a_3 e^{b_3\Theta}$, where parameters a_3 and b_3 can be set according to $a_3 = K(0)$ and $b_3 = \frac{2}{\pi}\ln\left(K\left(\frac{\pi}{2}\right)/K(0)\right)$. Without loss of generality, in the simulations, we set $\alpha\left(\frac{\pi}{2}\right) = 2$, $\alpha(0) = 3.5$, $a_2 = 10$, $b_2 = 5$, $K(0) = 5$ dB, and $K\left(\frac{\pi}{2}\right) = 15$ dB [21].

Figure 2 shows the convergence behaviors of the proposed algorithm. The results are based on 10^4 times random realizations of wireless channels. As the figure shows, the required number of iterations T is stable and acceptable, which demonstrates the efficiency of the proposed algorithm.

Figure 3 shows the performance of the proposed algorithm under different levels of covertness constraint in terms of various χ. Herein, Benchmark 1 is the EGT strategy adopted in [16,17], in which the collaborative beamforming direction is determined by $w_{d,i} = \frac{1}{\sqrt{N_a}} \frac{\left([\boldsymbol{h}_b]_i\right)^*}{\left|[\boldsymbol{h}_b]_i\right|}$ $(i = 1, 2, ..., N_a)$. Benchmark 2 is the initial solution of the proposed algorithm, in which the collaborative beamforming direction is determined by $\boldsymbol{w}_d = \frac{(\boldsymbol{I}-\boldsymbol{P}_{\boldsymbol{A}})(\boldsymbol{h}_b)^*}{||(\boldsymbol{I}-\boldsymbol{P}_{\boldsymbol{A}})(\boldsymbol{h}_b)^*||}$ (i.e., the optimal solution of P-Initial). For fair comparison and to satisfy the covertness constraints, in Benchmarks 1 and 2, Bob's interference power and Alice's total transmit power are set as their optimal values according to (43) and (45), respectively. As can be observed, when the number of collaborative transmit nodes increases, the performance of the proposed algorithm is dramatically improved. Note that this performance gain comes from at least three aspects. First, increasing N_a allows Alice to generate narrower and more flexible beams, which is critical for countering far UAV Willies. Second, increasing N_a allows Alice to adjust the spatial distribution of the signal energy more flexibly. Specifically, a transmit node that is relatively far from all the ground Willies and near UAV Willies can adopt a relatively large transmit power. Third, both ground Willies' and near UAV Willies' received signals from different transmit nodes are randomly superimposed, whereas Bob's received signals can be constructively superimposed by properly setting $\boldsymbol{w}_d$. Thus, a larger N_a leads to a larger gap of the received power between Bob and ground/near UAV Willies. Moreover, the performance of Benchmark 1 is observed to be much worse than that of the proposed algorithm. This finding indicates that merely optimizing Bob's reception while ignoring Willies' reception cannot fully exploit the potential of collaborative beamforming for covert communication. In contrast, Benchmark 2's performance is close to but is still worse than that of the proposed algorithm. This performance gap arises mainly because Benchmark 2 only considers far UAV Willies when determining $\boldsymbol{w}_d$, which indicates the importance of considering heterogeneous wardens in 3D scenarios. Note that another drawback of Benchmark 2 is that it is feasible only when $N_a > J_f$. However, Benchmark 2 can be adopted as a low-complexity solution of problem P0 when $N_a > J_f$.

Figure 4 shows the effects of the limit of Bob's interference power (i.e., $\hat{P}_{\max}$) on the performance of the proposed algorithm. It is first noted from problem P0 that increasing $\hat{P}_{\max}$ can only enhance and will not degrade the performance, since increasing $\hat{P}_{\max}$ extends the feasible set. However, as observed from the figure, increasing $\hat{P}_{\max}$ has negligible effects on the performance. This finding can be explained by revisiting optimal solutions (43), (45), and the objective function of SubP1(a). Specifically, expressions (43) and (45) indicate that at least one of the optimal values of P_{total} and $P_{\max}$ equals the corresponding upper limits MIN_P and $\hat{P}_{\max}$. For covert communication, the interference power should be large to conceal the communication behaviors, whereas Alice's transmit power is generally much smaller, so that generally it is $P_{\max}$ that equals the corresponding upper limit $\hat{P}_{\max}$. In this case, we

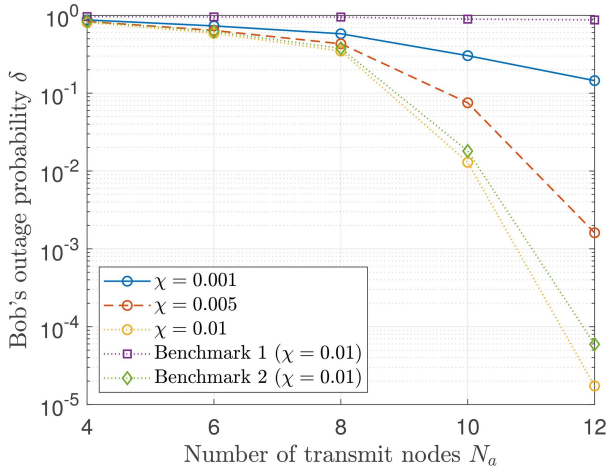

Figure 3 (Color online) Effects of the number of collaborative transmit nodes.

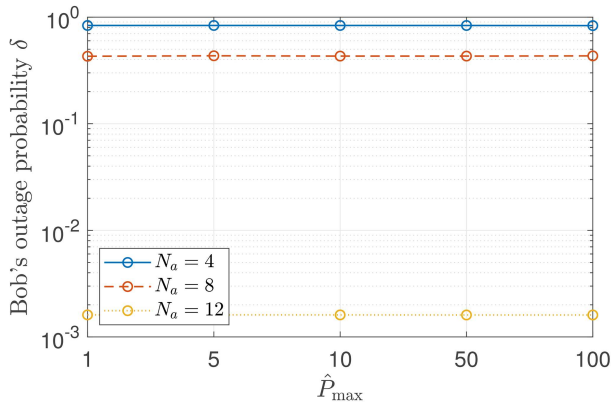

Figure 4 (Color online) Effects of the limit of Bob's interference power.

have $P^*_{\text{total}} = \frac{\hat{P}_{\max}}{\text{MAX}_w}$ so that the objective function of SubP1(a) equals $\left[\frac{|\boldsymbol{h}_b^{\text{T}}\boldsymbol{w}_d|^2}{\rho \text{MAX}_w d_b^{\alpha(\Theta_b)}(2^R-1)} - \frac{\sigma_b^2}{\rho\hat{P}_{\max}}\right]^+$. As observed from this expression, when the self-interference at Bob is much larger than his received power (i.e., $\rho\hat{P}_{\max} \gg \sigma_b^2$), which commonly holds, increasing $\hat{P}_{\max}$ brings negligible improvement to the performance. To sum up, although Bob's interference plays the role of Willie's uncertainty source, its marginal benefit is limited, as increasing the interference simultaneously degrades Bob's received SINR.

6 Concluding remarks

In this paper, the covert transmission from a ground DAS to a remote full-duplex UAV receiver was investigated. The expressions of the detection performance of ground Willies, near UAV Willies, and far UAV Willies were derived, based on which the potential of collaborative transmission was analyzed. Furthermore, the interference power, collaborative beamforming direction, and transmit power were jointly optimized to minimize the receiver's outage probability on the premise of satisfying the covert constraints for the three types of wardens. Numerical results demonstrated significant performance gains of the proposed algorithm over the EGT benchmark, validated the key role of the number of collaborative transmit nodes, and demonstrated the importance of considering the heterogeneity of wardens.

An important future direction is to further analytically reveal the effects of the distribution of the transmit nodes and to optimize the deployment of the transmit nodes. Additionally, the covert performance and algorithm designs merit study under more general channel models in future work.

Acknowledgements This work was supported in part by National Natural Science Foundation of China (Grant Nos. 62171445, 62371457, 62201590) and Postdoctoral Fellowship Program and China Postdoctoral Science Foundation (Grant No. BX20240471).

Supporting information Appendixes A–G. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Jayaprakasam S, Rahim S K A, Leow C Y. Distributed and collaborative beamforming in wireless sensor networks: classifications, trends, and research directions. IEEE Commun Surv Tutorials, 2017, 19: 2092–2116
2 Quitin F, Rahman M M U, Mudumbai R, et al. A scalable architecture for distributed transmit beamforming with commodity radios: design and proof of concept. IEEE Trans Wireless Commun, 2013, 12: 1418–1428
3 Wentz M, Chowdhury K R. Intra-network synchronization and retrodirective distributed transmit beamforming with UAVs. IEEE Trans Veh Technol, 2024, 73: 2017–2031
4 Scherber D, Bidigare P, O'Donnell R, et al. Coherent distributed techniques for tactical radio networks: enabling long range communications with reduced size, weight, power and cost. In: Proceedings of IEEE Military Communications Conference, San Diego, 2013. 655–660
5 Kramarev D, Ahmad I, Layton K, et al. Event-triggered synchronization for mobile distributed transmit beamforming. In: Proceedings of IEEE Military Communications Conference, Norfolk, 2019. 343–348
6 Chen X, An J, Xiong Z, et al. Covert communications: a comprehensive survey. IEEE Commun Surv Tutorials, 2023, 25: 1173–1198
7 Bash B A, Goeckel D, Towsley D. Limits of reliable communication with low probability of detection on AWGN channels. IEEE J Sel Areas Commun, 2013, 31: 1921–1930
8 He B, Yan S, Zhou X, et al. On covert communication with noise uncertainty. IEEE Commun Lett, 2017, 21: 941–944

9 Bash B A, Goeckel D, Towsley D. Covert communication gains from adversary's ignorance of transmission time. IEEE Trans Wireless Commun, 2016, 15: 8394–8405
10 Sobers T V, Bash B A, Guha S, et al. Covert communication in the presence of an uninformed jammer. IEEE Trans Wireless Commun, 2017, 16: 6193–6206
11 Yu X C, Luo Y, Chen W. Covert communication with beamforming over MISO channels in the finite blocklength regime. Sci China Inf Sci, 2021, 64: 192303
12 Chen X, Sun W, Xing C, et al. Multi-antenna covert communication via full-duplex jamming against a warden with uncertain locations. IEEE Trans Wireless Commun, 2021, 20: 5467–5480
13 Wang C, Li Z, Ng D W K Covert rate optimization of millimeter wave full-duplex communications. IEEE Trans Wireless Commun, 2022, 21: 2844–2861
14 Zhang Y, Zhang Y, Wang J, et al. Distance-angle beamforming for covert communications via frequency diverse array: toward two-dimensional covertness. IEEE Trans Wireless Commun, 2023, 22: 8559–8574
15 Yao J, Xin L, Wu T, et al. FAS for secure and covert communications. IEEE Internet Things J, 2025, 12: 18414–18418
16 Zheng T X, Wang H M, Ng D W K, et al. Multi-antenna covert communications in random wireless networks. IEEE Trans Wireless Commun, 2019, 18: 1974–1987
17 Wang J Q, Yu P X, Xiao S, et al. Achieving positive covert rate in distributed antenna system. In: Proceedings of IEEE GlobeCom Workshops, Rio de Janeiro, Brazil, 2022. 625–630
18 Chen W, Ding H, Wang S, et al. Jammer-aided covert collaborative AAV communications against directional beam. IEEE Trans Commun, 2025, 73: 12413–12429
19 Zeng Y, Wu Q, Zhang R. Accessing from the sky: a tutorial on UAV communications for 5G and beyond. Proc IEEE, 2019, 107: 2327–2375
20 Khawaja W, Guvenc I, Matolak D W, et al. A survey of air-to-ground propagation channel modeling for unmanned aerial vehicles. IEEE Commun Surv Tutorials, 2019, 21: 2361–2391
21 Azari M M, Rosas F, Chen K C, et al. Ultra reliable UAV communication using altitude and cooperation diversity. IEEE Trans Commun, 2018, 66: 330–344
22 Shu F, Xu T, Hu J, et al. Delay-constrained covert communications with a full-duplex receiver. IEEE Wireless Commun Lett, 2019, 8: 813–816
23 Li B, Zhao S, Zhang R, et al. Full-duplex UAV relaying for multiple user pairs. IEEE Internet Things J, 2021, 8: 4657–4667
24 Zarifi K, Affes S, Ghrayeb A. Collaborative null-steering beamforming for uniformly distributed wireless sensor networks. IEEE Trans Signal Process, 2010, 58: 1889–1903
25 Ochiai H, Mitran P, Poor H V, et al. Collaborative beamforming for distributed wireless ad hoc sensor networks. IEEE Trans Signal Process, 2005, 53: 4110–4124
26 Chen W Y, Ding H Y, Wang S L, et al. Beamforming design for covert broadcast communication with hidden adversary. Sci China Inf Sci, 2024, 67: 162304
27 Vrba M, Saska M. Marker-less micro aerial vehicle detection and localization using convolutional neural networks. IEEE Robot Autom Lett, 2020, 5: 2459–2466
28 Zhao B, Wang C, Fu Q, et al. A novel pattern for infrared small target detection with generative adversarial network. IEEE Trans Geosci Remote Sens, 2021, 59: 4481–4492
29 Sedunov A, Haddad D, Salloum H, et al. Stevens drone detection acoustic system and experiments in acoustics UAV tracking. In: Proceedings of IEEE International Symposium on Technologies for Homeland Security, Woburn, 2019. 1–7
30 Knoedler B, Steffes C, Koch W. Detecting and tracking a small UAV in GSM passive radar using track-before-detect. In: Proceedings of IEEE Radar Conference, Florence, 2020. 1–6
31 Nie W, Han Z C, Li Y, et al. UAV detection and localization based on multi-dimensional signal features. IEEE Sens J, 2022, 22: 5150–5162
32 Ma R, Yang W, Tao L, et al. Covert communications with randomly distributed wardens in the finite blocklength regime. IEEE Trans Veh Technol, 2022, 71: 533–544
33 Forouzesh M, Samsami Khodadad F, Azmi P, et al. Simultaneous secure and covert transmissions against two attacks under practical assumptions. IEEE Internet Things J, 2023, 10: 10160–10171
34 Sun Y, Baricz Á, Zhou S. On the monotonicity, log-concavity, and tight bounds of the generalized Marcum and Nuttall Q-functions. IEEE Trans Inform Theor, 2010, 56: 1166–1186
35 Luo Z, Ma W, So A, et al. Semidefinite relaxation of quadratic optimization problems. IEEE Signal Process Mag, 2010, 27: 20–34
36 Xu H, Wong K-K, Zhu Y X, et al. FAS-assisted federated learning over wireless communication systems. Sci China Inf Sci, 2025, 68: 170304
37 Guo Y F, Luo J S, Wang F G, et al. Dual-end fluid antennas for robust anti-jamming in low-altitude air-ground communications. 2025. ArXiv:2509.02260
38 Luo J S, Wang S L, He B X. Large-scale aerial reconfigurable intelligent surface-aided robust anti-jamming transmission. 2025. ArXiv:2509.10280