

Wireless Surveillance in a MIMO System with Spoofing Relay and UAV-enabled Eavesdropper*

Delin Guo^{†‡}, Hui Ding[†], Lan Tang[†], *Member, IEEE*, Xinggan Zhang[†], Ying-Chang Liang[§], *Fellow, IEEE*

Abstract—This paper investigates a proactive eavesdropping scenario, where a legitimate unmanned aerial vehicle (UAV)-enabled eavesdropper aims to eavesdrop on multiple suspicious communication links with the help of a spoofing relay. Each suspicious link comprises of a mobile source and a fixed destination. The full-duplex spoofing relay receives suspicious information and amplify-and-forwards (AF) it to the eavesdropper and suspicious destinations. To maximize the long-run eavesdropping rate, we jointly optimize the precoder at the spoofing relay, the receive combiner at the eavesdropper, and the UAV's moving trajectory. Considering the difficulty of joint optimization, we decompose the optimization into two parts, where a semi-definite relaxation (SDR)-based approach and a low-complexity solution are proposed to optimize the radio actions (the precoder and receive combiner), while the moving trajectory of UAV is learned by reinforcement learning (RL). Numerical simulations demonstrate the superiority of proposed method over other benchmarks in various experimental setups.

Index Terms—Proactive eavesdropping, trajectory planning, spoofing relay, reinforcement learning (RL), unmanned aerial vehicle (UAV).

I. INTRODUCTION

In order to ensure public safety, certain authorized parties strongly require the development of effective methods for eavesdropping on suspicious wireless communication links. In light of this context, wireless surveillance has garnered significant research attention in recent years [1].

The most covert approach is passive eavesdropping, which deploys a radio-silent eavesdropper to monitor suspicious information. However, this method is ineffective when the wiretap channel is worse than the suspicious channel. To enhance eavesdropping efficiency, [2] and [3] proposed a proactive eavesdropping scheme based on jamming. In this scheme, a full-duplex (FD) eavesdropper sends jamming signals to the suspicious destination to induce the suspicious source to reduce its transmission rate. This enables the eavesdropper to reliably decode the suspicious information. However, when the eavesdropping channel is better than the suspicious channel, it is desired to spoof the suspicious source to increase its transmission rate so that more suspicious information can be overheard by the eavesdropper. Therefore, [4] proposed a more general proactive eavesdropping scheme by utilizing a spoofing relay. Specifically, the relay amplifies and forwards received signals to the suspicious destination. By designing the transmit precoder at the relay, it becomes possible to adjust (enhance or reduce) both wiretap link and suspicious link capacities in favor of improving eavesdropping performance. In [5], both jammer and spoofing relay were adopted for improving eavesdropping efficiency.

All the aforementioned studies have focused on scenarios involving fixed nodes. However, due to its inherent flexibility, Unmanned

Aerial Vehicles (UAVs) have recently emerged as a promising tool in wireless communications [6, 7]. Moreover, the utilization of UAV-enabled eavesdroppers has garnered significant research attention [8–12]. In [8], trajectories of multiple radio-silent UAV eavesdroppers were optimized to enhance eavesdropping rates. [9] jointly optimized receiving vectors and trajectories of a radio-silent UAV eavesdropper to maximize the eavesdropping rate. In [10], jamming power and position of a jamming-only UAV were optimized to improve the eavesdropping rate at a radio-silent receiver. [11] employed a jamming-enabled UAV eavesdropper to overhear communications between multiple sources and a destination. In [12], multiple UAV eavesdroppers were deployed to eavesdrop on multiple suspicious links by jointly optimizing their jamming powers and trajectories.

In this paper, we focus on wireless surveillance implemented by a UAV eavesdropper. To our knowledge, existing research on UAV eavesdroppers has primarily considered passive eavesdropping (e.g., [8, 9]) or proactive eavesdropping via jamming (e.g., [10–12]). However, UAV eavesdroppers possess the ability to fly flexibly to appropriate positions for eavesdropping (e.g., close to suspicious sources) and typically benefit from high-quality line-of-sight (LoS) channels. Consequently, the capacity of the eavesdropping channel is likely to surpass that of the suspicious channel. In such scenarios, it is necessary to induce the suspicious source to enhance its transmission rate via a spoofing relay. Motivated by this, we investigate a novel wireless surveillance scenario wherein a radio-silent UAV eavesdropper aims to eavesdrop on multiple communication suspicious links with assistance from a full-duplex spoofing relay. Each suspicious link comprises a mobile source and a fixed destination. The relay receives information from suspicious sources and amplify-and-forwards (AF) it to both the eavesdropper and suspicious destinations. To maximize the long-run eavesdropping rate, we jointly optimize the precoder at the spoofing relay, the moving trajectory, and the receiving vector of the UAV eavesdropper. The main contributions are summarized as follows:

1) We investigate a novel wireless surveillance scheme that utilizes a spoofing relay to assist in the eavesdropping of the UAV eavesdropper. The radio action (i.e., the precoder of a relay and the receive vector at a UAV) and the moving action (i.e., UAV's trajectory) are jointly optimized to improve the long-run eavesdropping rate.

2) Considering the Markov property and the high non-convexity of the investigated problem, we propose a hierarchical approach. Specifically, a semi-definite relaxation (SDR)-based method is proposed to solve the radio action in each time-step, and given this SDR-based optimizer, the UAV's moving policy is learned by the reinforcement learning (RL) technique. In addition, a low-complexity solution to the radio action is provided.

3) Extensive experimental results demonstrate that our method significantly outperforms other comparison schemes.

Notations: $\|\cdot\|$, $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^\dagger$, $\text{tr}(\cdot)$, $\text{Vec}(\cdot)$ and $\text{rank}(\cdot)$ denote the Frobenius norm, transpose, Hermitian, conjugate, trace, vectorization and rank of a matrix, respectively. $\mathbf{I}$ is the identity matrix, $\otimes$ denotes Kronecker product, $\mathbb{R}/\mathbb{C}$ denotes the set of real/complex numbers, and $\text{diag}(\cdot, \dots, \cdot)$ denotes the block-wise diagonalization of matrices.

II. SYSTEM MODEL

As depicted in Fig. 1, we consider a wireless surveillance system comprising of a UAV eavesdropper E equipped with N_{rE} receive antennas, a spoofing relay R equipped with N_{tR} transmit antennas and N_{rR} receive antennas, and K suspicious communication links. Each suspicious link consists of a mobile source S_k with N_{tS_k} transmit antennas, and a single-antenna destination D_k ($1 \leq k \leq K$). For example, multiple suspects move around sensitive areas to gather crucial information such as military equipment and buildings, and

Copyright (c) 2024 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

*This work was supported in part by the National Natural Science Foundation of China under grants 62072229, and in part by State Key Laboratory of Air Traffic Management System, NO. SKLATM202302.

[†]Delin Guo, Lan Tang and Xinggan Zhang are with the School of Electronic Science and Engineering, Nanjing University, Nanjing 210093, P. R. China. (e-mail: tanglan@nju.edu.cn). (Corresponding author: Lan Tang.)

[‡]Delin Guo and Hui Ding are with the State Key Laboratory of Air Traffic Management System, Nanjing 210023, China.

[§]Ying-Chang Liang is with the Center for Intelligent Networking and Communications (CINC), University of Electronic Science and Technology of China (UESTC), Chengdu 611731, China. (e-mail: liangyc@ieee.org).

transmit the collected information back to the fixed access points. The K suspicious links are assumed to operate in frequency division multiple access (FDMA) mode with equal bandwidth. The UAV eavesdropper E can dynamically adjust its position for enhanced eavesdropping capabilities while the relay R amplify-and-forwards signals transmitted by the mobile sources S_k s towards both E and their respective destinations D_k s, thereby adapting channel capacities for both eavesdropping and suspicious links. The legitimate nodes (i.e., E and R) aim to collaboratively monitor all the suspicious links by optimizing the transmit precoder at R, the trajectory and the receive vector of E.

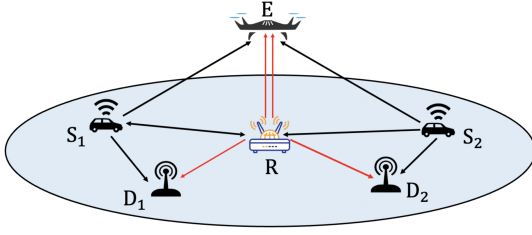


Fig. 1: System model with $K = 2$.

The channel coefficient/vector/matrix from the node X to Y is denoted as $h_{XY}/\mathbf{h}_{XY}/\mathbf{H}_{XY}$. The channel is modeled as $\mathbf{H}_{XY} = \frac{\beta_0}{d_{XY}^\alpha} \tilde{\mathbf{H}}_{XY}$, where β_0 represents the reference power at a distance of 1 meter, d_{XY} denotes the distance between nodes X and Y, α represents the path loss exponent, and $\tilde{\mathbf{H}}_{XY}$ refers to the small-scale fading matrix with elements following Rayleigh/Rician distribution in ground-to-ground (G2G)/air-to-ground (A2G) scenarios. The time is slotted as $t = 1, 2, \dots$, with length ΔT , and it is assumed that channels and node locations are quasi-static within each time step. The time-division duplex (TDD) mode is adopted, so the channels satisfy the reciprocal property.

In the time step t , the signal in the k th frequency band received by R is given by¹ (The index t is omitted here for brevity)

$$\mathbf{y}_R^k = \mathbf{H}_{S_k R} \mathbf{f}^k x_k + \mathbf{n}_R^k, \quad (1)$$

where x_k denotes the transmit signal of S_k with power $P_k = \mathbb{E}\{|x_k|^2\}$, $\mathbf{f}^k \in \mathbb{C}^{N_{tS_k} \times 1}$ denotes the beamformer of S_k , and $\mathbf{n}_R^k \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_{N_{tR}})$ is the noise at R. The signal forwarded by R is

$$\mathbf{x}_R^k = \mathbf{W}_k \mathbf{y}_R^k = \mathbf{W}_k (\mathbf{H}_{S_k R} \mathbf{f}^k x_k + \mathbf{n}_R^k), \quad (2)$$

where $\mathbf{W}_k \in \mathbb{C}^{N_{tR} \times N_{tR}}$ is the precoding matrix that needs to be optimized. According to (2), the transmit power of R is $P_R = P_k \|\mathbf{W}_k \mathbf{H}_{S_k R} \mathbf{f}^k\|^2 + \sigma^2 \|\mathbf{W}_k\|^2$. Similar to [4], it is assumed that the AF delay can be negligible compared to the symbol duration. Therefore, the signal received at D_k is

$$y_{D_k} = (\mathbf{h}_{S_k D_k}^H + \mathbf{h}_{RD_k}^H \mathbf{W}_k \mathbf{H}_{S_k R}) \mathbf{f}^k x_k + \mathbf{h}_{RD_k}^H \mathbf{W}_k \mathbf{n}_R^k + n_{D_k}, \quad (3)$$

where $n_{D_k} \sim \mathcal{CN}(0, \sigma^2)$ is the noise at R. We assume that all the suspicious links are not aware of being eavesdropped, so $\mathbf{f}^k$ is set to $\mathbf{h}_{S_k D_k} / \|\mathbf{h}_{S_k D_k}\|$. Similarly, in the k th frequency band, the signal received by E is given by

$$\mathbf{y}_E^k = (\mathbf{H}_{S_k E} + \mathbf{H}_{RE} \mathbf{W}_k \mathbf{H}_{S_k R}) \mathbf{f}^k x_k + \mathbf{H}_{RE} \mathbf{W}_k \mathbf{n}_R^k + \mathbf{n}_E^k. \quad (4)$$

¹The self-interference of R is assumed to be eliminated similar to [2, 3].

In the k th frequency band, the signal to interference and noise ratio (SINRs) at E and D_k are

$$\gamma_E^k = \frac{P_k |\mathbf{u}_k^H (\mathbf{H}_{S_k E} + \mathbf{H}_{RE} \mathbf{W}_k \mathbf{H}_{S_k R}) \mathbf{f}^k|^2}{\sigma^2 (1 + \|\mathbf{u}_k^H \mathbf{H}_{RE} \mathbf{W}_k\|^2)}, \quad (5)$$

$$\gamma_{D_k}^k = \frac{P_k |(\mathbf{h}_{S_k D_k}^H + \mathbf{h}_{RD_k}^H \mathbf{W}_k \mathbf{H}_{S_k R}) \mathbf{f}^k|^2}{\sigma^2 (1 + \|\mathbf{h}_{RD_k}^H \mathbf{W}_k\|^2)},$$

respectively, where $\mathbf{u}_k^H$ ($\|\mathbf{u}_k^H\| = 1$) is the receive vector of E in the k th frequency band.

III. PROBLEM FORMULATION

In this paper, our objective is to maximize the long-term eavesdropping rate by jointly optimizing the radio actions $\{\mathbf{W}_k(t), \mathbf{u}_k(t)\}_{k=1}^K$ and the velocity $\mathbf{v}_t$ of UAV E across all time steps. We denote the k th group of channel information at time step t as $h_t^k = (\mathbf{h}_{S_k D_k}(t), \mathbf{h}_{RD_k}(t), \mathbf{H}_{S_k R}(t), \mathbf{H}_{S_k E}(t), \mathbf{H}_{RE}(t))$. The k th radio action is represented as $\mathbf{a}_t^{r,k} = (\mathbf{W}_k(t), \mathbf{u}_k(t))$, and $\mathbf{a}_t^r = (a_t^{r,1}, \dots, a_t^{r,K})$ denotes the joint radio actions.

With the above representations, the problem is formulated as

$$\begin{aligned} \max_{\{\mathbf{a}_t^r, \mathbf{v}_t\}_{t=1}^T} \quad & \sum_{t=1}^T \sum_{k=1}^K \log_2(1 + \gamma_{D_k}(h_t^k, \mathbf{a}_t^{r,k})) \\ \text{s.t.} \quad & \text{C1: } \gamma_E^k(h_t^k, \mathbf{a}_t^{r,k}) \geq \gamma_{D_k}(h_t^k, \mathbf{a}_t^{r,k}), \forall k, t \\ & \text{C2: } P_R^k(h_t^k, \mathbf{a}_t^{r,k}) \leq P_{\max}, \forall k, t \\ & \text{C3: } \|\mathbf{v}_t\| \leq V_{\max}, \forall t \\ & \text{C4: } \mathbf{q}_{t+1}^E = \mathbf{q}_t^E + \mathbf{v}_t \Delta T, \forall t, \end{aligned} \quad (6)$$

where $\mathbf{q}_t^E$ denotes the position of E at the time step t . In the problem (6), C1 is to ensure successful eavesdropping. That is, the signals received by E can be reliably decoded only when C1 is satisfied [2]. C2 is the peak power constraint. C3 is the velocity constraints of E. C4 characterizes the motion characteristics of E.

A. MDP representation

The Markov Decision Process (MDP) can be represented by a quadruple $\Gamma = \langle S, \mathcal{A}, P, r \rangle$, where S denotes the state space, $\mathcal{A}$ represents the action space, P signifies the transition probability (typically unknown), and r stands for the reward function. Subsequently, we provide detailed definitions of state, action, and reward.

State: Similar to [2, 4], we assume the global channel information is available to the eavesdropping party. At the time step t , the global state (information) is defined as $s_t = (\{h_t^k\}_{k=1}^K, \{\mathbf{q}_t^X\}_{X \in \mathcal{T}})$, where $\mathbf{q}_t^X$ represents the position of the node X and $\mathcal{T} = \{S_k, D_k\}_{k=1}^K, R, E\}$. Let $\mathcal{S}$ denote the space of s_t .

Action: At the time-step t , the joint action refers to $\mathbf{a}_t = (\mathbf{a}_t^r, \mathbf{v}_t)$. For any k and t , define the radio action space $\mathcal{A}_k^r(s_t)$ as the intersection of C1 and C2. Define the moving action space $\mathcal{A}_v$ as the set satisfying C3. Based on the above representations, feasible solutions satisfying C1–C3 must be in the joint action space $\mathcal{A}(s_t) = \mathcal{A}_r(s_t) \times \mathcal{A}_v$, where $\mathcal{A}_r(s_t) = \prod_{k=1}^K \mathcal{A}_k^r(s_t)$.

Note that C4 is not included in a action space, because it is a part of the system dynamics and should be incorporated into the system transition probability $P(s_{t+1}|s_t, \mathbf{a}_t)$ rather than the action space.

Reward: With the state-action pair $(s_t, \mathbf{a}_t)$, the reward function is defined as

$$r(s_t, \mathbf{a}_t) = \sum_{k=1}^K \log_2(1 + \gamma_{D_k}(h_t^k, \mathbf{a}_t^{r,k})) \cdot \mathbb{I}(h_t^k, \mathbf{a}_t^{r,k}), \quad (7)$$

where $\mathbb{I}(h_t^k, \mathbf{a}_t^{r,k}) = \mathbb{I}\{\gamma_E^k(h_t^k, \mathbf{a}_t^{r,k}) \geq \gamma_{D_k}(h_t^k, \mathbf{a}_t^{r,k})\}$, and $\mathbb{I}\{x\}$ is the boolean indicator function which is 1 if x is true and 0 otherwise.

Problem Reformulation: Define the policy as $\pi : \mathcal{S} \rightarrow \mathcal{P}(\mathcal{A}(\mathcal{S}))$, i.e., $a_t \sim \pi(\cdot|s_t) \in \mathcal{A}(s_t)$, where $\mathcal{P}(\mathcal{A}(\mathcal{S}))$ is the set of probability measures on $\mathcal{A}(\mathcal{S})$. The objective of an MDP is to optimize the policy π to maximize the long-run return $\eta(\pi) = \mathbb{E}\{\sum_{t=0}^{T-1} \gamma^t r(s_t, \mathbf{a}_t) | \pi\}$, where $\mathbf{a}_t \sim \pi(\mathbf{a}_t|s_t)$, $s_{t+1} \sim P(s_{t+1}|s_t, \mathbf{a}_t)$, $\forall t \geq 0$, and $\gamma \in (0, 1)$ (typically close to 1) represents the discount factor. Therefore, the problem (6) is re-formulated as

$$\max_{\pi} \eta(\pi). \quad (8)$$

IV. SOLUTION TO (8)

In the previous section, we formulate the original problem in (6) as a standard MDP problem (8), which could be solved by RL algorithms in principle. However, due to the presence of constraints C1 and C2, it becomes impractical to design a policy π (typically represented by neural networks) that can select actions $\mathbf{a}_t$ from the action space $\mathcal{A}(s_t)$. Therefore, the problem (8) can hardly be solved by applying solely the RL technique.

Based on the quasi-static assumption of the considered system, it could be observed from (6) that: 1) $\mathbf{v}_t$ only influences the subsequent state s_{t+1} and has no relevance to the current reward $r(s_t, \mathbf{a}_t)$, 2) the radio action $\mathbf{a}_t^r$ only determines the current reward and has no impact on s_{t+1} , 3) The action space for movement $\mathcal{A}_v$ remains fixed, enabling easy representation in a RL policy.

Motivated by the above observations, we propose a hierarchical approach to address (8). Firstly, we employ conventional optimization methods to maximize the reward at each time step t . For consistency in notation, we denote the radio-action solver (not necessarily the optimal) as $\pi^r : \mathcal{S} \rightarrow \mathcal{P}(\mathcal{A}_r(\mathcal{S}))$. Secondly, given π^r , we optimize the moving-action policy $\pi^v : \mathcal{S} \rightarrow \mathcal{P}(\mathcal{A}_v)$ using the RL technique.

Let $\hat{\pi}^r$ denote the optimal solver of radio actions, i.e., $\sum_{\mathbf{a}^r \in \mathcal{A}_r(s)} \hat{\pi}^r(\mathbf{a}^r|s) r(s, \mathbf{a}^r) \geq \sum_{\mathbf{a}^r \in \mathcal{A}_r(s)} \pi^r(\mathbf{a}^r|s) r(s, \mathbf{a}^r)$, $\forall s, \forall \pi^r$. Denote the optimal moving action policy under the given $\hat{\pi}^r$ as $\hat{\pi}^v = \arg \max_{\pi^v} \eta(\pi^v, \hat{\pi}^r)$. It has been proved in [12] that the joint policy $\hat{\pi} = \hat{\pi}^v \times \hat{\pi}^r$ is optimal, i.e., $\eta(\hat{\pi}) \geq \eta(\pi), \forall \pi$. The above discussion indicates that sequentially optimizing π^r and π^v does not impair the global optimality. In the following two subsections, we will provide the solutions to radio actions and moving policy.

A. Solution to Radio-action

In each time step t , the global problem in (8) (or (6)) can be decomposed into K sub-problems, where the k th sub-problem is

$$\max_{\mathbf{u}_k, \mathbf{W}_k} \gamma_{D_k}, \text{ s.t.: } \gamma_E^k \geq \gamma_{D_k}, P_R^k \leq P_{\max}. \quad (9)$$

Obviously, the problem (9) is non-convex so that it is hard to obtain the global optimal solution. Therefore, we propose a sub-optimal solution by employing the block coordinate descent (BCD) approach. That is, $\mathbf{u}_k$ is fixed when optimizing $\mathbf{W}_k$ and vice versa.

1) Optimize $\mathbf{W}_k$ with Given $\mathbf{u}_k$: Denote $\mathbf{w}_k = \text{Vec}(\mathbf{W}_k)$. With a given $\mathbf{u}_k$, the problem (9) can be transformed into

$$\begin{aligned} \max_{\mathbf{w}_k} & \frac{|\mathbf{h}_{S_k D_k}^H \mathbf{f}^k + \mathbf{z}_2^H \mathbf{w}_k|^2}{1 + \mathbf{w}_k^H \mathbf{Z}_4 \mathbf{w}_k} \\ \text{s.t.:} & \frac{|\mathbf{u}_k^H \mathbf{H}_{S_k E} \mathbf{f}^k + \mathbf{z}_1^H \mathbf{w}_k|^2}{1 + \mathbf{w}_k^H \mathbf{Z}_3 \mathbf{w}_k} \geq \frac{|\mathbf{h}_{S_k D_k}^H \mathbf{f}^k + \mathbf{z}_2^H \mathbf{w}_k|^2}{1 + \mathbf{w}_k^H \mathbf{Z}_4 \mathbf{w}_k} \\ & P_k \mathbf{w}_k^H \mathbf{Z}_5 \mathbf{w}_k + \sigma^2 \mathbf{w}_k^H \mathbf{w}_k \leq P_{\max}, \end{aligned} \quad (10)$$

where $\mathbf{z}_1 = ((\mathbf{H}_{S_k R} \mathbf{f}^k)^T \otimes (\mathbf{u}_k^H \mathbf{H}_{RE}^k))^H$, $\mathbf{z}_2 = ((\mathbf{H}_{S_k R} \mathbf{f}^k)^T \otimes \mathbf{h}_{RD_k}^H)^H$, $\mathbf{Z}_3 = \mathbf{I}_{N_{rR}} \otimes (\mathbf{H}_{RE}^k)^H \mathbf{u}_k \mathbf{u}_k^H \mathbf{H}_{RE}^k$, $\mathbf{Z}_4 = \mathbf{I}_{N_{rR}} \otimes \mathbf{h}_{RD_k} \mathbf{h}_{RD_k}^H$, and $\mathbf{Z}_5 = (\mathbf{H}_{S_k R} \mathbf{f}^k (\mathbf{H}_{S_k R} \mathbf{f}^k)^H)^T \otimes \mathbf{I}_{N_{tR}}$. The index k is omitted here for brevity.

The problem (10) is still difficult to solve due to the 2-norm terms in the denominators. Therefore, we further define $\tilde{\mathbf{w}}_k = \begin{bmatrix} e^{jx} \\ \mathbf{w}_k \end{bmatrix}$,

$\tilde{\mathbf{z}}_1 = \begin{bmatrix} \mathbf{u}_k^H \mathbf{H}_{S_k E} \mathbf{f}^k \\ \mathbf{z}_1 \end{bmatrix}$ and $\tilde{\mathbf{z}}_2 = \begin{bmatrix} \mathbf{h}_{S_k D_k}^H \mathbf{f}^k \\ \mathbf{z}_2 \end{bmatrix}$. In addition, denote $\tilde{\mathbf{Z}}_1 = \tilde{\mathbf{z}}_1 \tilde{\mathbf{z}}_1^H$, $\tilde{\mathbf{Z}}_2 = \tilde{\mathbf{z}}_2 \tilde{\mathbf{z}}_2^H$, $\tilde{\mathbf{Z}}_3 = \text{diag}(0, \mathbf{Z}_3)$, $\tilde{\mathbf{Z}}_4 = \text{diag}(0, \mathbf{Z}_4)$, and $\tilde{\mathbf{Z}}_5 = \text{diag}(0, \mathbf{Z}_5)$. Then, the problem (10) can be equivalently represented as

$$\begin{aligned} \max_{\tilde{\mathbf{w}}_k} & \frac{\tilde{\mathbf{w}}_k^H \tilde{\mathbf{Z}}_2 \tilde{\mathbf{w}}_k}{1 + \tilde{\mathbf{w}}_k^H \tilde{\mathbf{Z}}_4 \tilde{\mathbf{w}}_k} \\ \text{s.t.:} & \frac{\tilde{\mathbf{w}}_k^H \tilde{\mathbf{Z}}_1 \tilde{\mathbf{w}}_k}{1 + \tilde{\mathbf{w}}_k^H \tilde{\mathbf{Z}}_3 \tilde{\mathbf{w}}_k} \geq \frac{\tilde{\mathbf{w}}_k^H \tilde{\mathbf{Z}}_2 \tilde{\mathbf{w}}_k}{1 + \tilde{\mathbf{w}}_k^H \tilde{\mathbf{Z}}_4 \tilde{\mathbf{w}}_k} \\ & P_k \tilde{\mathbf{w}}_k^H \tilde{\mathbf{Z}}_5 \tilde{\mathbf{w}}_k + \sigma^2 \tilde{\mathbf{w}}_k^H \tilde{\mathbf{w}}_k - \sigma^2 \leq P_{\max}, \\ & |\tilde{\mathbf{w}}_k[1]| = |e^{jx}| = 1. \end{aligned} \quad (11)$$

Using the fact that $\tilde{\mathbf{w}}_k^H \tilde{\mathbf{Z}}_1 \tilde{\mathbf{w}}_k = \text{tr}(\tilde{\mathbf{Z}}_1 \tilde{\mathbf{w}}_k \tilde{\mathbf{w}}_k^H)$ and defining $\mathbf{X}_k = \tilde{\mathbf{w}}_k \tilde{\mathbf{w}}_k^H$, the problem (11) can be re-expressed as

$$\begin{aligned} \max_{\mathbf{X}_k \succeq 0} & \frac{\text{tr}(\tilde{\mathbf{Z}}_2 \mathbf{X}_k)}{1 + \text{tr}(\tilde{\mathbf{Z}}_4 \mathbf{X}_k)} \\ \text{s.t.:} & \frac{\text{tr}(\tilde{\mathbf{Z}}_1 \mathbf{X}_k)}{1 + \text{tr}(\tilde{\mathbf{Z}}_3 \mathbf{X}_k)} \geq \frac{\text{tr}(\tilde{\mathbf{Z}}_2 \mathbf{X}_k)}{1 + \text{tr}(\tilde{\mathbf{Z}}_4 \mathbf{X}_k)} \\ & \text{tr}((P_k \tilde{\mathbf{Z}}_5 + \mathbf{I}) \mathbf{X}_k) \leq P_{\max} + \sigma^2, \\ & \mathbf{X}_k[1, 1] = 1, \text{rank}(\mathbf{X}_k) = 1. \end{aligned} \quad (12)$$

By introducing an auxiliary variable $\beta = \frac{\text{tr}(\tilde{\mathbf{Z}}_1 \mathbf{X}_k)}{1 + \text{tr}(\tilde{\mathbf{Z}}_3 \mathbf{X}_k)}$, the following problem is obtained:

$$\begin{aligned} \max_{\beta, \mathbf{X}_k \succeq 0} & \beta \\ \text{s.t.:} & \frac{\text{tr}(\tilde{\mathbf{Z}}_1 \mathbf{X}_k)}{1 + \text{tr}(\tilde{\mathbf{Z}}_3 \mathbf{X}_k)} \geq \beta, \frac{\text{tr}(\tilde{\mathbf{Z}}_2 \mathbf{X}_k)}{1 + \text{tr}(\tilde{\mathbf{Z}}_4 \mathbf{X}_k)} = \beta \\ & \text{tr}((P_k \tilde{\mathbf{Z}}_5 + \mathbf{I}) \mathbf{X}_k) \leq P_{\max} + \sigma^2 \\ & \mathbf{X}_k[1, 1] = 1, \text{rank}(\mathbf{X}_k) = 1. \end{aligned} \quad (13)$$

The above problem can be transformed into convex one by fixing β and relaxing the rank-one constraint,

$$\begin{aligned} \min_{\mathbf{X}_k \succeq 0} & \text{tr}((P_k \tilde{\mathbf{Z}}_5 + \mathbf{I}) \mathbf{X}_k) \\ \text{s.t.:} & \text{tr}((\beta \tilde{\mathbf{Z}}_4 - \tilde{\mathbf{Z}}_2) \mathbf{X}_k) + \beta \leq 0 \\ & \text{tr}((\beta \tilde{\mathbf{Z}}_3 - \tilde{\mathbf{Z}}_1) \mathbf{X}_k) + \beta = 0 \\ & \mathbf{X}_k[1, 1] - 1 = 0. \end{aligned} \quad (14)$$

The problem (14) is a standard semi-definite programming (SDP) problem, which can be solved by some standard convex optimization tools. The optimal β can be searched linearly in an outer iteration.

It should be noted that the optimal solution $\mathbf{X}_k^*$ of problem (14) may not necessarily be rank-one, which is a requirement for recovering $\tilde{\mathbf{w}}_k$. In such cases, we employ eigenvalue decomposition (ED) methods to obtain an approximate rank-one solution [13]. Additionally, the following proposition presents a sufficient condition for $\text{rank}(\mathbf{X}_k^*) = 1$.

Proposition 1: If $N_{rR} = 1$, then the rank of the optimal solution to the problem (14) is one, i.e., $\text{rank}(\mathbf{X}_k^*) = 1$.

Proof: We provide the proof in Appendix A.

2) Optimize $\mathbf{u}_k$ for given $\mathbf{W}_k$: Given $\mathbf{W}_k$, the problem (9) is degenerated as $\max_{\mathbf{u}_k} \gamma_E^k$, s.t.: $\|\mathbf{u}_k\| = 1$. Rewrite γ_E^k as $\frac{\mathbf{u}_k^H \mathbf{Q}_1^k \mathbf{u}_k}{\mathbf{u}_k^H \mathbf{Q}_2^k \mathbf{u}_k}$, where $\mathbf{Q}_1^k = (\mathbf{H}_{S_k E} \mathbf{f}^k + \mathbf{H}_{RE}^k \mathbf{W}_k \mathbf{H}_{S_k R} \mathbf{f}^k)(\mathbf{H}_{S_k E} \mathbf{f}^k + \mathbf{H}_{RE}^k \mathbf{W}_k \mathbf{H}_{S_k R} \mathbf{f}^k)^H$ and $\mathbf{Q}_2^k = \mathbf{I}_{N_{rE}} + (\mathbf{H}_{RE}^k \mathbf{W}_k)(\mathbf{H}_{RE}^k \mathbf{W}_k)^H$. The above problem is a general Rayleigh quotient problem. Therefore, the optimal $\mathbf{u}_k$ is the eigenvector corresponding to the largest eigenvalue of $(\mathbf{Q}_2^k)^{-1} \mathbf{Q}_1^k$.

Low-complexity scheme: The proposed SDR-based method demonstrates ideal performance, as shown in the simulation section. However, this approach requires iterative solving the SDP problem (14), resulting in highly computational complexity. Therefore, we also

$$\gamma_E^k(x, y) = \frac{|\mathbf{h}_{RE}^H \mathbf{C} \mathbf{h}_{S_k R}|^2 x^2 + 2|\mathbf{h}_{S_k E}| |\mathbf{h}_{RE}^H \mathbf{C} \mathbf{h}_{S_k R}| \cos(\phi_{S_k E} - \phi_1 - y)x + |\mathbf{h}_{S_k E}|^2}{1 + \|\mathbf{h}_{RE}^H \mathbf{C}\|^2 x^2} \quad (16)$$

$$\gamma_{D_k}(x, y) = \frac{|\mathbf{h}_{RD_k}^H \mathbf{C} \mathbf{h}_{S_k R}|^2 x^2 + 2|\mathbf{h}_{S_k D_k}| |\mathbf{h}_{RD_k}^H \mathbf{C} \mathbf{h}_{S_k R}| \cos(\phi_{S_k D_k} - \phi_2 - y)x + |\mathbf{h}_{S_k D_k}|^2}{1 + \|\mathbf{h}_{RD_k}^H \mathbf{C}\|^2 x^2} \quad (17)$$

provide a heuristic low-complexity scheme. Here, we denote $\mathbf{h}_{S_k E} = \mathbf{u}_k^H \mathbf{H}_{S_k E} \mathbf{f}^k$, $\mathbf{h}_{RE}^k = \mathbf{u}_k^H \mathbf{H}_{RE}^k$, $\mathbf{h}_{S_k R} = \mathbf{H}_{S_k R} \mathbf{f}^k$, and $\mathbf{h}_{S_k D_k} = \mathbf{h}_{S_k D_k}^H \mathbf{f}^k$ for simplicity.

According to Theorem 1 in [4], the optimal solution of the problem, $\max_{\mathbf{W}_k} \gamma_E^k$, s.t.: $P_R \leq P_{\max}$, has a closed-form solution $\mu_1^* e^{j\angle \mathbf{h}_{S_k E}} \mathbf{h}_{RE}^k \mathbf{h}_{S_k R}^H$, which can be regarded as the maximum ratio transmission (MRT) precoder for maximizing γ_E^k . Similarly, the MRT precoder for maximizing γ_{D_k} is $\mu_2^* e^{j\angle \mathbf{h}_{S_k D_k}} \mathbf{h}_{RD_k} \mathbf{h}_{S_k R}^H$. Motivated by the two MRT precoders, we heuristically set $\mathbf{W}_k = \nu \mathbf{h}_{RE}^k \mathbf{h}_{S_k R}^H + \kappa \mathbf{h}_{RD_k} \mathbf{h}_{S_k R}^H$. To minimize computational complexity, we only consider two extreme cases: $\nu = 0$ or $\kappa = 0$. The optimal $\mathbf{W}_k$ is selected from $\{\nu \mathbf{h}_{RE}^k \mathbf{h}_{S_k R}^H, \kappa \mathbf{h}_{RD_k} \mathbf{h}_{S_k R}^H\}$. When $\mathbf{W}_k = \nu \mathbf{h}_{RE}^k \mathbf{h}_{S_k R}^H$, the problem (9) with a given $\mathbf{u}_k$ is

$$\begin{aligned} \max_{x, y} \quad & \gamma_{D_k}(x, y), \\ \text{s.t.:} \quad & \gamma_E^k(x, y) \geq \gamma_{D_k}(x, y), \\ & P_R^k(x) \leq P_{\max}, \end{aligned} \quad (15)$$

where x and y are amplitude and phase of ν , respectively, $P_R^k(x) = (P_k \|\mathbf{C} \mathbf{h}_{S_k R}\|^2 + \sigma^2 \|\mathbf{C}\|^2) x^2$ and $\mathbf{C} = \mathbf{h}_{RE}^k \mathbf{h}_{S_k R}^H$. The full expressions of $\gamma_E^k(x, y)$ and $\gamma_{D_k}(x, y)$ are given at the top of this page, where $\phi_{S_k E} = \angle \mathbf{h}_{S_k E}$, $\phi_1 = \angle \mathbf{h}_{RE}^k \mathbf{C} \mathbf{h}_{S_k R}$, $\phi_{S_k D_k} = \angle \mathbf{h}_{S_k D_k}$, and $\phi_2 = \angle \mathbf{h}_{RD_k} \mathbf{C} \mathbf{h}_{S_k R}$. The problem (15) is easy to solve. For given y , the optimal x can be found by analyzing the feasible region of x and the monotonicity of the objective function in the region, and the optimal y can be found by linear search. The complexity of solving x is $\mathcal{O}(1)$, resulting in an overall complexity of only $\mathcal{O}(1/\epsilon)$, where ϵ is the relative accuracy for searching y .

B. Solution to Moving policy

Denote the proposed optimizer above for the radio action as $\bar{\pi}^r$. Next, with the given $\bar{\pi}^r$, we optimize the moving policy π^v using RL, i.e., $\max_{\pi^v} \eta(\pi^v, \bar{\pi}^r)$. The overall algorithm for optimizing π^v is presented in Algorithm 1. In RL, the policy π^v is parameterized as π_θ^v with parameters θ by using the artificial neural network (NN). In each episode, the eavesdropping party interacts with the environment using the policy $(\pi^v, \bar{\pi}^r)$, and stores the collected data $(s_t, a_t^v, r_t, s_{t+1})$ into the memory buffer $\mathcal{D}$. Then for each training step, the policy is updated by Soft Actor-Critic (SAC), an advanced RL algorithm (i.e., SACUpdate function). Since SAC has been well-established in [14] and we do not detail them here due to limited space.

V. NUMERICAL RESULTS

In this section, we present simulation results with discussion and analysis. Unless otherwise specified, we set $K = 4$, $N_{tR} = 2$, $N_{rR} = 1$, $N_{tS_k} = 2$, $N_{rE} = 1$, and $V_{\max} = 10m/s$. The time-step length $\Delta T = 1s$. The path loss exponent α for A2G and G2G links are 2.2 and 2.4, respectively. $\beta_0/\sigma^2 = 10^6$. $P_k = 20$ dBm and $P_{\max} = 20$ dBm. Cartesian coordinate is adopted to describe the locations of nodes. R is fixed at (0, 0, 0) (in meter), the initial location of E is (0, 300, 100), D_k s are randomly scattered within a range of [75, 150] meters from R, and each S_k moves within 125 meters around the corresponding D_k . In addition, we assume that E flies at a fixed height of 100 meters. As for the hyperparameters of SAC, the critic

Algorithm 1: Optimization of π^v

```

1 Initialize the moving policy  $\pi_\theta^v$  with parameters  $\theta$ ;
2 Initialize the memory buffer  $\mathcal{D}$ ;
3 for episode = 1, 2, ... do
4   Initialize  $s_1$ ;
5   for  $t = 1, 2, \dots, T$  do
6     Select  $\mathbf{a}_t = (a_t^v, a_t^r)$  using  $(\pi^v, \bar{\pi}^r)$ ;
7     Execute  $\mathbf{a}_t$  and receive  $r_t = r(s_t, \mathbf{a}_t)$ ;
8     System moves to the next state  $s_{t+1}$ ;
9     Store the transition  $(s_t, a_t^v, r_t, s_{t+1})$  into  $\mathcal{D}$ ;
10     $\pi_\theta^v \leftarrow \text{SACUpdate}(\mathcal{D})$ ;
11  end
12 end

```

and actor are represented by two-layer fully-connected networks with ReLU activation, and each hidden layer has 64 neurons. The smooth factor is 5×10^{-3} and the learning rate is 3×10^{-4} . In the following comparison schemes, the moving actions of E are learned by SAC, while the radio actions are determined with different methods.

1) Adj- γ_{D_k} [4]: For the case $N_{tR} > N_{rE}$, let $\mathbf{W}_k = \hat{\mathbf{V}}_k \hat{\mathbf{W}}_k$, where $\hat{\mathbf{V}}_k$ is composed of the right singular vectors corresponding to zero eigenvalues of $\mathbf{H}_{RE}^k$. Therefore, $\gamma_E^k = P_k |\mathbf{u}_k^H \mathbf{H}_{S_k E} \mathbf{f}^k|^2 / \sigma^2$ with $\mathbf{u}_k = \frac{\mathbf{H}_{S_k E} \mathbf{f}^k}{\|\mathbf{H}_{S_k E} \mathbf{f}^k\|}$. Then we optimize the problem $\max_{\mathbf{W}_k} \gamma_{D_k}$, s.t.: $\gamma_E^k \geq \gamma_{D_k}$, $P_R^k \leq P_{\max}$, which can be solved by using the method in the section III-C of [4]. Noting that γ_E^k is constant in this scheme and only γ_{D_k} needs to be adjusted.

2) Jamming: In this scheme, the relay R is replaced by a jammer. The jamming signals are assumed to be canceled at E by information sharing. Therefore, $\gamma_E^k = P_k |\mathbf{u}_k^H \mathbf{H}_{S_k E} \mathbf{f}^k|^2 / \sigma^2$ and $\gamma_{D_k} = (P_k |\mathbf{h}_{S_k D_k}^H \mathbf{f}^k|^2) / (\sigma^2 + P_{jam}^k \|\mathbf{h}_{RD_k}\|^2)$, where P_{jam}^k is the jamming power. The radio action for each k is obtained by solving $\max_{\mathbf{u}_k^H, P_{jam}^k} \gamma_{D_k}$, s.t.: $\gamma_E^k \geq \gamma_{D_k}$, $P_{jam}^k \leq P_{\max}$.

3) Max- γ_E [4]: Maximize γ_E^k without taking into account any effects on γ_{D_k} . That is, $\max_{\mathbf{u}_k, \mathbf{W}_k} \gamma_E^k$, s.t.: $P_R^k \leq P_{\max}$, which can be solved by using the theorem 1 in [4].

4) Passive: The relay R is removed and only the trajectory of E is optimized.

In the following simulation results, each point of the eavesdropping rate (i.e., $\mathbb{E}_t \{r(s_t, \mathbf{a}_t)\}$) is obtained by averaging at least 10^5 Monte Carlo simulations using the well-trained policy.

Fig. 2(a) shows the eavesdropping rate versus different maximum transmit power $P_{\max}$, where $N_{tR} = 4$ and $N_{rE} = 2$. In terms of the eavesdropping rate, our proposed SDR-based scheme significantly outperforms all other baselines, thereby validating the effectiveness of our method. The low-complexity scheme exhibits comparable performance to the SDR-based method, particularly when $P_{\max} \leq 20$ dBm, striking a good trade-off between performance and complexity. Furthermore, it can be observed that the curve of the Jamming scheme plateaus when $P_{\max} \geq 30$ dBm due to limitations imposed by both suspicious and eavesdropping channel capacities. For the Adj- γ_{D_k} scheme, its eavesdropping rate surpasses that of the Jamming scheme because R can induce suspicious sources to enhance transmission

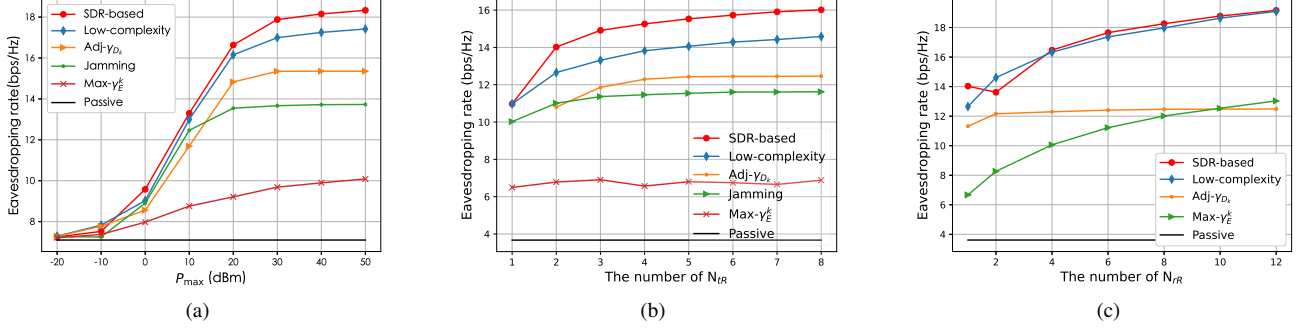


Fig. 2: The comparisons between the SDR-based scheme and baseline methods versus (a) different maximum transmit power $P_{\max}$, (b) different numbers of transmit antennas N_{tR} , and (c) different numbers of receive antennas N_{rR} .

rates when eavesdropping channels are favorable. However, the performance of Adj- γ_{D_k} is constrained by eavesdropping channel capacity, so it also reaches a steady state when $P_{\max} \geq 30$ dBm. In contrast, both SDR-based and low-complexity schemes continue to improve their performance beyond this threshold since relay R can simultaneously adjust capacities for both suspicious and eavesdropping channels.

Fig. 2(b) illustrates the impact of varying N_{tR} on the eavesdropping rate. It is worth noting that the point corresponding to $N_{tR} = 1$ is absent in the Adj- γ_{D_k} curve due to the requirement of $N_{tR} > N_{rE}$. The overall trend observed in Fig. 2(b) exhibits similarity with those in Fig. 2(a). In contrast to increasing $P_{\max}$, employing more transmit antennas in Max- γ_E yields little benefit and even introduces fluctuations in the eavesdropping rate.

Fig. 2(c) illustrates the eavesdropping rate versus varying N_{rR} . In this comparison, the performance gap between the low-complexity and SDR-based schemes is negligible. Interestingly, a slight decrease is observed in the SDR-based method when $N_{rR} = 2$. These results can be attributed to the fact that for $N_{rR} \geq 2$, the solution to the problem (14) may not necessarily be rank-one, leading to some loss due to rank-one approximation based on eigenvalue decomposition. Besides, since the eavesdropping rate of Adj- γ_{D_k} is limited by the capacity of the eavesdropping channel, it reaches a saturation point when $N_{rR} \geq 3$, while Max- γ_E scheme surpasses it when $N_{rR} \geq 10$.

Moving trajectory. Fig. 3 shows the trajectories of E for different locations of S_k, D_k , which may provide some intuitions. The solid line represents the trajectory of UAV E, while the dotted lines depict the trajectories of S_k s. Each suspicious source S_k moves around its respective destination (indicated by a dot). It is seen that UAV E strategically moves towards the area with higher concentrations of suspicious links and circles around in this area. As illustrated in Figs. 3(b), the UAV E tactically sacrifices eavesdropping on S_4 to achieve an overall higher eavesdropping rate.

Robustness to estimation errors. For the eavesdropping party, obtaining accurate $\{\mathbf{h}_{S_k D_k}\}_{k=1}^K$ is challenging, and a more plausible assumption is that only channels with estimation error can be obtained. We model the inaccurate channel as $\hat{\mathbf{h}}_{S_k D_k} = \sqrt{1 - \delta} \mathbf{h}_{S_k D_k} + \sqrt{\delta} \Phi$, where $\delta \in [0, 1]$ and Φ represents Gaussian estimation error. We solve the problem (9) with $\hat{\mathbf{h}}_{S_k D_k}$ and $\hat{\mathbf{f}}_k = \hat{\mathbf{h}}_{S_k D_k} / \|\hat{\mathbf{h}}_{S_k D_k}\|$, and then calculate the eavesdropping rate with the solution and accurate $\{\mathbf{h}_{S_k D_k}\}_{k=1}^K$. The experiments are conducted with $N_{tR} = 4$ and $N_{rE} = 2$. As shown in Fig. 4, the original SDR-based method (represented by the curve ‘Original’) exhibits high sensitivity to estimation errors. Even a small increase in δ leads to a significant decrease in the eavesdropping rate. This sensitivity can be

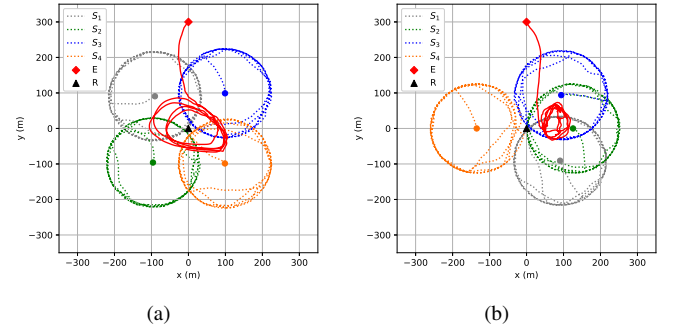


Fig. 3: The moving trajectories of E for different locations of (S_k, D_k) .

reduced by adding some redundancy in the constraint condition. Specifically, we can introduce an offset factor into the eavesdropping constraint, i.e., $\gamma_E^k \geq \zeta \gamma_{D_k}$, where $\zeta \geq 1$ is the offset factor. It is obvious that ζ should increase with the increase of δ . In our simulations, we set $\zeta = [1.1, 1.2, 1.35, 1.5, 2.0, 2.5, 3.0]$ for $\delta = [0.0001, 0.001, 0.01, 0.1, 0.2, 0.3, 0.4]$, respectively. In addition, it is found in experiments that the eavesdropping rate is insensitive to variations of ζ if ζ is within a reasonable range, so it does not need to delicately search the optimal ζ . As depicted by the curve ‘Modified’, although there is still an inevitable decrease in the eavesdropping rate with the increasing of δ , such a modification significantly enhances robustness of our proposed method against estimation errors.

VI. CONCLUSION

In this study, we employed a spoofing relay and a UAV eavesdropper to collaboratively eavesdrop on multiple suspicious links, each consisting of a mobile source and a stationary destination. To optimize the long-term eavesdropping rate, we proposed a hierarchical optimization method. Specifically, an SDR-based approach was introduced to determine the radio actions, while the trajectory of the UAV eavesdropper was learned using SAC. Additionally, we proposed a low-complexity solution for radio actions. Experimental results demonstrated that compared to the commonly used jamming-based eavesdropping scheme, the relay-based scheme increased the eavesdropping rate by over 20%. Moreover, the proposed low-complexity method achieved an excellent trade-off between performance and complexity.

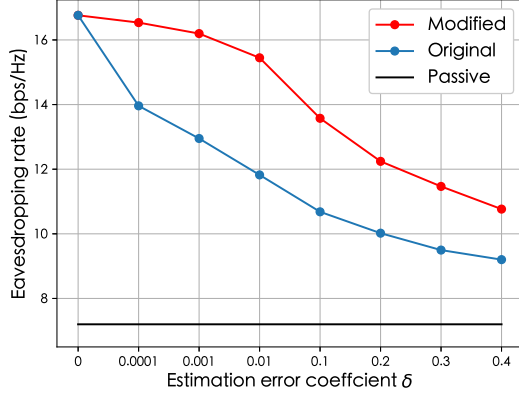


Fig. 4: Eavesdropping rate versus different δ .

APPENDIX A PROOF OF PROPOSITION 1

We first construct the Lagrangian function of the problem (14), which is given by

$$\mathcal{L}(\mathbf{X}_k, \mu_1, \mu_2, \mu_3) = \text{tr}(\mathbf{A}\mathbf{X}_k) + \mu_1\beta + \mu_2\beta - \mu_3, \quad (18)$$

where $\mathbf{A} = \mathbf{I} + P_k\tilde{\mathbf{Z}}_5 + \mu_1(\beta\tilde{\mathbf{Z}}_3 - \tilde{\mathbf{Z}}_1) + \mu_2(\beta\tilde{\mathbf{Z}}_4 - \tilde{\mathbf{Z}}_2) + \mu_3\mathbf{E}$, and $\mathbf{E}$ is the all-zero matrix except for $\mathbf{E}[1, 1] = 1$. To make sure the dual function $\inf_{\mathbf{X}_k \succeq 0} \mathcal{L}(\mathbf{X}_k, \mu_1, \mu_2, \mu_3)$ is lower-bounded, it must have $\mathbf{A} \succeq 0$ and $\text{tr}(\mathbf{A}\mathbf{X}_k^*) = 0$. Therefore, the dual problem is

$$\max_{\mu} \mathbf{c}^T \mu, \text{ s.t. : } \mathbf{A} \succeq 0, \mu_1 \geq 0, \quad (19)$$

where $\mathbf{c} = [\beta \ \beta \ -1]^T$ and $\mu = [\mu_1 \ \mu_2 \ \mu_3]^T$. Since $\text{tr}(\mathbf{A}\mathbf{X}_k^*) = 0$, it can be verified that $\text{rank}(\mathbf{A}) + \text{rank}(\mathbf{X}_k^*) = N_{rR}N_{tR} + 1$. Therefore, if we can determine the rank of $\mathbf{A}^* = \mathbf{I} + P_k\tilde{\mathbf{Z}}_5 + \mu_1^*(\beta\tilde{\mathbf{Z}}_3 - \tilde{\mathbf{Z}}_1) + \mu_2^*(\beta\tilde{\mathbf{Z}}_4 - \tilde{\mathbf{Z}}_2) + \mu_3^*\mathbf{E}$, where μ^* is the optimal solution of (19), then the rank of $\mathbf{X}_k^*$ is obtained.

We focus on the case with $N_{rR} = 1$ here. For the case of $N_{rR} = 1$, we denote $h_{s_kE} = \mathbf{u}_k^H \mathbf{H}_{s_kE} \mathbf{f}_k^k$, $h_{s_kD_k} = \mathbf{h}_{s_kD_k}^H \mathbf{f}_k^k$, and $h_{s_kR} = \mathbf{H}_{s_kR} \mathbf{f}_k^k$ for brevity. Based on definitions of $\tilde{\mathbf{Z}}_i$ in the text, we can derive $\beta\tilde{\mathbf{Z}}_3 - \tilde{\mathbf{Z}}_1 = a\mathbf{y}_1\mathbf{y}_1^H + b|h_{s_kE}|^2\mathbf{E}$, $\beta\tilde{\mathbf{Z}}_4 - \tilde{\mathbf{Z}}_2 = a\mathbf{y}_2\mathbf{y}_2^H + b|h_{s_kD_k}|^2\mathbf{E}$, and $\tilde{\mathbf{Z}}_5 = |h_{s_kR}|^2\mathbf{I} - |h_{s_kR}|^2\mathbf{E}$, where $\mathbf{y}_1 = [-h_{s_kE}/a \ \mathbf{z}_1^T]^T$, $\mathbf{y}_2 = [-h_{s_kD_k}/a \ \mathbf{z}_2^T]^T$, $a = \beta/|h_{s_kR}|^2 - 1$, and $b = -(1 + 1/a)$. Letting $\mathbf{Y}_1 = \mathbf{y}_1\mathbf{y}_1^H$ and $\mathbf{Y}_2 = \mathbf{y}_2\mathbf{y}_2^H$, we have $\mathbf{A} = q\mathbf{I} + d_1\mathbf{Y}_1 + d_2\mathbf{Y}_2 + d_3\mathbf{E}$, where $q = 1 + P_k|h_{s_kR}|^2$, $d_1 = a\mu_1$, $d_2 = a\mu_2$, and $d_3 = b|h_{s_kE}|^2\mu_1 + b|h_{s_kD_k}|^2\mu_2 + \mu_3$.² Denote $\mathbf{d} = [d_1 \ d_2 \ d_3]^T = \mathbf{G}\mu$. Obviously, $\mathbf{G}$ is invertible. Therefore, the dual problem (19) can be equivalently transformed as

$$\max_{\mathbf{d}} \mathbf{c}^T \mathbf{G}^{-1} \mathbf{d}, \text{ s.t. : } \mathbf{A} \succeq 0, d_1/a \geq 0. \quad (20)$$

Let (d_1^*, d_2^*, d_3^*) denote the optimal solution of the problem (20). To ensure that $\mathbf{X}_k$ has a non-zero solution, $\mathbf{A}^* = q\mathbf{I} + d_1^*\mathbf{Y}_1 + d_2^*\mathbf{Y}_2 + d_3^*\mathbf{E}$ must be singular. Since $\mathbf{A}^*$ is full-rank when $(d_1^*, d_2^*, d_3^*) \geq 0$, it means that at least one of (d_1^*, d_2^*, d_3^*) is less than 0. Without loss of generality, we assume $d_3^* < 0$ while make no assumption about d_1^* and d_2^* . We first prove by contradiction that $\mathbf{B}^* = q\mathbf{I} + d_1^*\mathbf{Y}_1 + d_2^*\mathbf{Y}_2$ must be full rank. Assume $\mathbf{B}^*$ is singular, and let $\mathbf{g}$ denote the eigenvector corresponding to zero eigenvalue of $\mathbf{B}^*$, i.e., $\mathbf{g}^H \mathbf{B}^* \mathbf{g} = 0$. The following relationship holds

$$\mathbf{g}^H \mathbf{A}^* \mathbf{g} = 0 + d_3^* \mathbf{g}^H \mathbf{E} \mathbf{g} \geq \lambda_{\min}(\mathbf{A}^*) = 0, \quad (21)$$

²Indeed, $d_3 = b|h_{s_kE}|^2\mu_1 + b|h_{s_kD_k}|^2\mu_2 + \mu_3 - P_k|h_{s_kR}|^2$. We have absorbed the constant term $-P_k|h_{s_kR}|^2$ into μ_3 here.

where $\lambda_{\min}(\cdot)$ indicates the minimum eigenvalue of a matrix. Since $\mathbf{E}$ is positive semi-definite (PSD), there must be $\mathbf{g}^H \mathbf{E} \mathbf{g} = 0$ and consequently $\mathbf{g}[1] = 0$. However, in the stochastic channel model, it can be proved $\Pr(\mathbf{g}[1] = 0) = 0$. Due to the limited spaces, we only provide the proof sketch of $\Pr(\mathbf{g}[1] = 0) = 0$. By utilizing the conditions, 1) $\mathbf{g}[1] = 0$, 2) $\mathbf{g} \in \text{span}\{\mathbf{y}_1, \mathbf{y}_2\}$, and 3) $\mathbf{B}^* \mathbf{g} = 0$, it can be obtained that $d_1^* = -q/(\|\mathbf{z}_1\|^2 - \frac{h_{s_kE}}{h_{s_kD_k}} \mathbf{z}_1^H \mathbf{z}_2)$.

Because d_1^* is a real number, $\frac{h_{s_kE}}{h_{s_kD_k}} \mathbf{z}_1^H \mathbf{z}_2$ is also necessarily a real number, which is impossible in the stochastic channel model. Consequently, it is obtained that $\mathbf{B}^*$ is full rank. Thereby, $N_{tR} \geq \text{rank}(\mathbf{A}^*) = \text{rank}(\mathbf{B}^* + d_3^*\mathbf{E}) \geq \text{rank}(\mathbf{B}^*) - 1 = N_{tR}$. Recall that $\text{rank}(\mathbf{A}) + \text{rank}(\mathbf{X}_k^*) = N_{tR} + 1$, and therefore $\text{rank}(\mathbf{X}_k^*) = 1$ holds. This completes the proof. $\square$

REFERENCES

- [1] J. Xu, L. Duan, and R. Zhang, "Surveillance and intervention of infrastructure-free mobile communications: A new wireless security paradigm," *IEEE Wireless Communications*, vol. 24, no. 4, pp. 152–159, 2017.
- [2] J. Xu, L. Duan, and et.al, "Proactive eavesdropping via cognitive jamming in fading channels," *IEEE Transactions on Wireless Communications*, vol. 16, no. 5, pp. 2790 – 2806, 2017.
- [3] J. Xu, L. Duan, and R. Zhang, "Proactive eavesdropping via jamming for rate maximization over rayleigh fading channels," *IEEE Wireless Communications Letters*, vol. 1, no. 5, pp. 80 – 83, 2016.
- [4] Y. Zeng and R. Zhang, "Wireless information surveillance via proactive eavesdropping with spoofing relay," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 8, pp. 1449 – 1461, 2017.
- [5] J. Moon, H. Lee, C. Song, S. Lee, and I. Lee, "Proactive eavesdropping with full-duplex relay and cooperative jamming," *IEEE Transactions on Wireless Communications*, vol. 17, no. 10, pp. 6707 – 6719, 2018.
- [6] B. Duo, J. Luo, Y. Li, H. Hu, and Z. Wang, "Joint trajectory and power optimization for securing uav communications against active eavesdropping," *China Communications*, vol. 18, no. 1, pp. 88–99, 2021.
- [7] C. Liao, K. Xu, H. Zhu, X. Xia, Q. Su, and N. Sha, "Secure transmission in satellite-uav integrated system against eavesdropping and jamming: A two-level stackelberg game model," *China Communications*, vol. 19, no. 7, pp. 53–66, 2022.
- [8] H. Huang, A. V. Savkin, and W. Ni, "Navigation of a uav team for collaborative eavesdropping on multiple ground transmitters," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 10, pp. 10450–10460, 2021.
- [9] J. Wang, P. Zhang, L. Tang, Y. Bai, and L. Yang, "Intelligent passive eavesdropping in massive mimo-ofdm systems via reinforcement learning," *IEEE Wireless Communications Letters*, vol. 11, no. 6, pp. 1248–1252, 2022.
- [10] G. Hu, Y. Cai, and Y. Cai, "Joint optimization of position and jamming power for uav-aided proactive eavesdropping over multiple suspicious communication links," *IEEE Wireless Communications Letters*, vol. 9, no. 12, pp. 2093–2097, 2020.
- [11] G. Hu, Z. Li, J. Si, K. Xu, D. Xu, Y. Cai, and N. Al-Dhahir, "Maxmin fairness for uav-enabled proactive eavesdropping with jamming over distributed transmit beamforming-based suspicious communications," *IEEE Transactions on Communications*, pp. 1–1, 2023.
- [12] D. Guo, L. Tang, X. Zhang, and Y.-C. Liang, "Joint optimization of trajectory and jamming power for multiple uav-aided proactive eavesdropping," *IEEE Transactions on Mobile Computing*, pp. 1–17, 2023.
- [13] Z.-q. Luo, W.-k. Ma, A. M.-c. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 20–34, 2010.
- [14] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor," in *International conference on machine learning*. PMLR, 2018, pp. 1861–1870.