Resource Allocation for Joint Interference Management and Security Enhancement in Cellular-Connected Internet-of-Drones Networks

Md. Zoheb Hassan, Georges Kaddoum, Senior Member, IEEE, and Ouassima Akhrif, Senior Member, IEEE

Abstract—Internet-of-drones (IoD) systems require enhanced data transmission security and efficient interference management to accommodate the rapidly growing drone-based and rate-intensive applications. This paper develops a novel resource allocation scheme to jointly manage interference and enhance the physical layer security of cellular-connected IoD networks in the presence of a multi-band eavesdropping drone. Our envisioned cellular-connected IoD network has multiple full-duplex cellular base stations (CBSs), where each CBS reserves an orthogonal cellular radio resource block (RRB) for the aerial communication links. To efficiently utilize cellular RRBs, each CBS is connected to a cluster of data transmitting drones using uplink non-orthogonal multiple access (NOMA) scheme. In addition, all the CBSs simultaneously transmit artificial noise signals to weaken the eavesdropper links. A joint optimization problem, considering the transmit power allocation and clustering of the legitimate drones, and the jamming power allocation of the CBSs, is formulated to maximize the worst-case average sum-secrecy-rate of the network. The joint optimization problem is decomposed into drone-clustering and power allocation sub-problems to obtain an efficient solution. A multi-agent reinforcement-learning framework is devised to solve the drone-clustering sub-problem. Meanwhile, the transmit and jamming power allocation sub-problem is solved by employing fractional programming, successive convex approximation, and alternating optimization techniques. By iteratively solving these two sub-problems, a convergent resource allocation algorithm, namely, security and interference management with reinforcement-learning and NOMA (SIREN), is proposed. The superiority of SIREN over several benchmark schemes is verified via extensive simulations.

Index Terms—Internet-of-drones, interference management, physical layer security, resource allocation.

I. INTRODUCTION

Cellular-connected Internet-of-drones (IoD) deploys drones as the flying Internet-of-things devices, and facilitates transmission of the data collected by the drones to the ground data centers via cellular radio resource blocks (RRBs) [1], [2]. Thanks to the seamless coverage and increased bandwidth of mobile-networks, using cellular RRBs improves both range and capacity of the aerial data transmission links [3]. However, the performance of cellular-connected IoD networks suffers from both inter-drone interference and drone-to-ground interference experienced at neighboring cellular base stations (CBSs) [4]. Besides, due to the broadcasting nature of wireless transmissions and the presence of dominant line-of-sight (LOS) links, the confidential data in IoD systems can be readily eavesdropped by ground and aerial adversaries. This work seeks to address these challenges by proposing a joint interference management and secure data transmission mechanism for cellular-connected IoD networks.

A. Related Works

1) Interference Management in Drone Networks: Transmit power allocation schemes that minimize the overall energy consumption in IoD systems were investigated in [5], [6]. However, the studies in [5], [6] did not consider any interference in the system. Due to the presence of strong LOS air-to-ground (A2G) communication links, the drones in cellular-connected IoD networks can generate strong interference at the neighboring CBSs. The mitigation of the interference between aerial and cellular links over shared RRBs have been studied in several works. In [7], using a cognitive radio framework, the drone’s transmit power control was exploited to maximize the data rate of the aerial communication links while keeping a low interference level at the cellular users. In [8], the joint optimization of the drones’ trajectory, transmit power, and CBS association was investigated to manage interference in cellular-connected drone networks. However, the studies conducted in [7], [8] ignored the inter-drone interference. In [9], the joint transmit power control and sub-channel allocation problems were investigated for a cellular-connected multi-drone network using both licensed and unlicensed channels. However, this study considered orthogonal scheduling of the drones over the licensed channels, and treated the interference over the unlicensed channels as-a-noise. Essentially, the inter-drone interference management scheme of [9] is also sub-optimal.

2) Physical Layer Security in Drone Networks: Conventional cryptographic methods enable secured data transmission using encryption keys. However, due to the limited on-board energy and computation capability, and high-mobility of the drones, the generation and distribution of encryption keys is challenging in aerial communication systems. Leveraging wiretap channel coding, physical layer security (PLS) can safeguard the transmitted data without the need for exchanging
complex encryption keys. As a result, PLS requires much less complexity than cryptographic methods, which makes PLS attractive for aerial communications [10]. Besides, a multi-layer security mechanism can be constructed by augmenting PLS with upper layer’s existing security protocols. The key performance metric of PLS is the secrecy-rate (SR), which quantifies an effective transmission rate of the confidential message without any information leakage at the eavesdropping nodes. A positive SR is guaranteed as long as the legitimate link has a higher channel capacity than the eavesdropping link. Hence, the SR can be enhanced by appropriately designing the resource allocations. The existing literature optimized several factors to enhance the SR of aerial communication systems, including the drones’ transmit power, the trajectory followed by the drones, and the sub-channel assignment between the drones and the ground users [11]–[13].

The SR of aerial communications can be further enhanced using cooperative jamming, where artificial noise (AN) is simultaneously transmitted with the confidential messages. The purpose of AN is to artificially increase interference at the eavesdroppers without affecting the legitimate links. In [14], dual-drone based cooperative jamming, where one drone transmits confidential messages and another drone simultaneously transmits AN signals to reduce the eavesdropping capability of the ground adversaries, was developed. In [15], a power-splitting method was developed to split the transmit power of the legitimate drone between the confidential messages and AN signals. The studies in [11], [12], [14], [15] assumed accurate estimation of the eavesdroppers’ locations. In [16]–[18], resource allocation schemes were developed to maximize the worst-case SR while considering bounded location error models for the eavesdroppers. In [19], a probabilistic location error model for the eavesdroppers was utilized to maximize the outage-constrained average SR via robust optimization of the drones’ trajectory, and the transmit and jamming powers. Nonetheless, most of the these studies ignored the interference by scheduling the available drone-to-ground links over orthogonal channels. Therefore, despite improving the PLS of drone networks, these studies fail to efficiently utilize the radio resources.

3) Interference Management With PLS: Non-orthogonal multiple access (NOMA) is a powerful technique that can enhance the spectral efficiency of multi-user networks [20]. In NOMA, signals from multiple users are multiplexed over the same RRB using different transmit power levels at the transmitter and decoded at the receiving node using successive interference cancellation (SIC). It is noted that the data transmission process in a cellular-connected IoD network is an uplink multiple access scheme. Hence, NOMA is promising to concurrently schedule multiple drones to transmit data over the same cellular RRB. By efficiently exploiting interference in the network, NOMA improves the achievable SR as well [21].

Recently, NOMA-enabled secured aerial communications has received considerable attention. In [22], NOMA was exploited to transmit the secured and public users’ messages from an aerial BS such that the secured message cannot be eavesdropped by public users. In [23], NOMA was studied to efficiently allocate the transmit power of an aerial BS between a pair of security-demanding and rate-demanding ground users without leakage of secured information. In [24], both NOMA and cooperative jamming were integrated to enhance the PLS of an aerial communication system. However, the studies in [22]–[24] considered single drone-based communications. We emphasize that the existing literature lacks works on the joint management of interference and PLS for multi-drone cellular-connected IoD networks considering both inter-drone and drone-to-ground interference into account.

B. Motivation, Applications, and Contributions

1) Motivation: The aim of this work is to develop a resource allocation scheme for cellular-connected IoD networks. In particular, we envision an IoD network in which multiple legitimate drones concurrently transmit confidential data to a set of full-duplex (FD) CBSs using cellular RRBs and power-domain NOMA in the presence of a multi-band eavesdropping drone. The motivation for using NOMA and FD technology is explained as follows. We specifically consider a dense IoD network in which there are more legitimate drones than available cellular RRBs. To efficiently utilize the cellular RRBs, a drone-clustering scheme is proposed so that the drones within a cluster can concurrently transmit over the same RRB. Nevertheless, drone-clustering inevitably introduces co-channel inter-drone interference that reduces the capacity of the legitimate data transmission links. To combat this interference, power-domain NOMA is employed in the drone-clusters. Meanwhile, cooperative jamming is used to make the drone-to-ground communications resilient to eavesdropping. We take inspiration from [16], [17] and implement the cooperative jamming by incorporating FD technology at the CBSs so that the CBSs can simultaneously receive data from legitimate drones and transmit AN signals to weaken eavesdropping links. Note that FD technology can be readily deployed at the CBSs by combining analog and digital self-interference (SI) cancellation techniques [25]. To the best of our knowledge, this is the first work that jointly employs drone-clustering, NOMA-enabled data transmission, and FD CBS-aided cooperative jamming to improve both the capacity and the security of multi-drone cellular-connected IoD networks.

2) Applications: IoD networks are promising for both civilian and military surveillance and monitoring applications [26]. In these types of applications, drones usually generate high-definition videos and images that are processed at the ground control station. As a result, these applications require high capacity drone-to-ground communications and beyond-visual-line-of-sight (BVLOS) drone operations for extended coverage. Existing point-to-point drone-to-ground communications rely mainly on the unlicensed 2.4 GHz industrial, scientific, and medical (ISM) frequency band, which supports only visual-line-of-sight drone operations and low data rate. In contrast, cellular networks provide reliable command and control links for drone operations in the BVLOS range, a large bandwidth, and extended coverage [27]. Thus, unlike the ISM band, cellular networks are well positioned to support a variety of IoD network applications. However, the cellular-connected IoD networks must be resilient to untrusted drones in the airspace, especially for security-critical applications. In this context, our proposed system model not only enhances the capacity of cellular-connected IoD networks by efficiently managing co-channel interference but also prevents unauthorized drones from eavesdropping on confidential information. Our proposed system model can therefore be used in various IoD network applications, namely, border surveillance, military search and rescue operations, and critical infrastructure monitoring that require both high-capacity and secured drone-to-ground data transmissions [26].

3) Contributions: In this work, a novel resource allocation scheme is proposed to enhance the overall SR of the cellular-connected IoD networks and efficiently control both co-channel
inter-drone interference and drone-to-CBS interference. The specific contributions of this work are summarized as follows.

1) The worst-case average sum-SR (WC-ASR) of the network is maximized subject to target SR constraints for the legitimate drones and maximum acceptable interference constraints at the CBSs. Towards this objective, an optimization problem is formulated by jointly considering the transmit power allocation and clustering of the legitimate drones, and the jamming power allocation of the CBSs. The presented joint optimization problem is NP-hard, making it inherently difficult to find a global optimal solution. A two-level optimization approach is proposed to design a low-complexity resource allocation scheme. More specifically, the joint optimization problem is decomposed into two sub-problems, i.e., drone-clustering and power allocation sub-problems. A sub-optimal yet efficient solution is obtained by iteratively solving these two sub-problems.

2) The drone-clustering sub-problem is formulated as a mixed-strategy repeated game and solved using a multi-agent reinforcement learning (RL) framework. Meanwhile, the transmit and jamming power allocation sub-problem is solved by employing fractional programming, successive convex approximation (SCA), and alternating optimization techniques. Near-optimality of the derived solutions to both sub-problems is proved.

3) A resource allocation algorithm of polynomial computational complexity, namely, security and interference management with reinforcement learning and NOMA (SIREN), is proposed. SIREN converges to a local-optimal solution of the presented joint optimization problem with a polynomial computational complexity. Through extensive simulations, the superiority of SIREN over several interference-aware power allocation, drone-clustering, and RL empowered methods is demonstrated.

The rest of this paper is organized as follows. In Section II, the system model and problem formulation are presented. The sub-problems’ solutions are provided in Sections III and IV. Section V presents an overview of the properties of the SIREN algorithm. Finally, simulation results and concluding remarks are provided in Sections VI and VII, respectively.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. System Model

1) System Overview: We consider a cellular-connected IoD network, as shown in Fig. 1, with $M$ legitimate drones, $K$ FD CBSs, and a multi-band eavesdropping drone that passively overhears the data transmitted by the legitimate drones over different RRBs. The CBSs are connected to a centralized network controller, that coordinates the overall resource allocation procedure. The geographical region is divided into $M$ different sensing zones, each of which is assigned to a single drone. In each sensing zone, the assigned drone sequentially visits $N$ stop-over points before returning to its initial location [5]. To facilitate this, the mission duration is divided into $N$ equal time-slots (TSs). In the $n$-th TS, where $n \in \{1, 2, \ldots, N\}$, the drones fly from the $(n-1)$-th stop-over point to the $n$-th stop-over point in their assigned sensing zones, collect data (e.g., images and videos) while statically hovering at the $n$-th stop-over point, and send the collected data to the CBSs using cellular RRBs. We emphasize that the stop-over points and the transition routes between the consecutive stop-over points need to be carefully designed to avoid drones colliding. Note that the proposed framework jointly optimizes the drone-to-CBS associations as well as the drones’ transmit and CBSs’ jamming power allocations at each TS. As a result, it is computationally intractable to optimize the drones’ stop-over points and transition routes in the existing framework. Meanwhile, our proposed optimization framework requires only the drones’ stop-over points. Therefore, for the analytical tractability, we consider that both the drones’ stop-over points and their transition routes in the sensing zones are predefined to avoid nearby drones colliding. Collision-free path optimization frameworks for a system of multiple drones are extensively investigated in the literature [28], [29] and can be leveraged to design the drones’ transition routes of our system model. We assume that the position of the eavesdropping drone can also differ from a TS to another.

Let $\mathcal{M} = \{1, 2, \ldots, M\}$ be the set of legitimate drones; $\mathcal{K} = \{1, 2, \ldots, K\}$ be the set of CBSs; $\mathbf{q}_{m,n} = \{x_{m,n}, y_{m,n}\}$ be the 2D location of the $n$-th stop-over point of the $m$-th drone; $\mathbf{v}^{(n)}_k$ be the 2D location of the eavesdropping drone at the $n$-th TS; $\mathbf{s}_k = \{x_k, y_k\}$ be the 2D coordinate of the $k$-th CBS; $H_k$ be the height of each CBS; $H_D$ and $H_c$ be the fixed altitudes of the legitimate and eavesdropping drones, respectively. Both legitimate and eavesdropping drones are equipped with a single omnidirectional antenna, whereas each CBS is equipped with two antennas to support FD transmission. Each CBS employs an orthogonal cellular RRB for collecting data from the associated drones. Such a cellular RRB is also reused in a set of predefined CBSs for uplink cellular communications, leading to an enhanced spectrum resource utilization [8].

2) Interference Management: The system model considered experiences both co-channel inter-drone and drone-to-ground interference. A hybrid multiple access scheme is adopted to mitigate the co-channel inter-drone interference. In particular, the data-transmitting drones are partitioned into a maximum of $K$ non-overlapping drone-clusters. Each drone-cluster is associated with only one CBS, and each CBS supports only one drone-cluster. Therefore, different drone-clusters utilize orthogonal RRBs for data transmission, and as a result, the inter-cluster interference is eliminated. To overcome the intra-cluster interference, uplink NOMA with suitable transmit power allocation is employed in each drone-cluster. To further alleviate the intra-cluster interference and utilize the A2G channel’s diversity, dynamic drone-clustering is introduced by allowing a given drone to be associated with different CBSs at different stop-over points.

Fig. 1. Cellular-connected IoD network with two FD CBSs, two drone-clusters, two RRBs, and one eavesdropping drone (drone-to-CBS interference and SI links are not shown for clarity).
Because the same RRB is reused at multiple CBSs for uplink aerial and cellular communications, the drones associated with a particular CBS can generate severe co-channel interference at neighboring CBSs [8]. Uplink power control is an effective methodology to prevent drones from generating harmful interference [27]. Let us consider that the RRB reserved for aerial communications in the \( k \)-th CBS is reused for uplink cellular communications by a set of CBSs, denoted by \( \mathcal{L}_k \). Essentially, the drones associated with the \( k \)-th CBS interfere with all the CBSs of the set \( \mathcal{L}_k \). Hence, the transmit power of the drones associated with the \( k \)-th CBS is controlled so that the resultant co-channel interference at each CBS in the set \( \mathcal{L}_k \) remains below an acceptable interference threshold.

3) Cooperative Jamming: To enhance the PLS, cooperative jamming is employed. In particular, leveraging its FD communication capability, each CBS simultaneously receives data from the associated drones and transmits AN signals over the RRB reserved for aerial communications. The transmitted AN signals reduce the channel capacity of the eavesdropping links at the cost of SI at the CBSs. We consider the CBSs to be equipped with state-of-the-art SI cancellation schemes [25] and thus affected only by residual SI. In particular, both the transmitted AN signal and distortions in the RF chain contribute to SI. RF distortions cannot be canceled out due to their stochastic behavior and equate to the residual SI.

4) Assumptions A1: The eavesdropping drone’s position is imperfectly known at the network controller [16]–[18]. Let \( \hat{\mathbf{d}}^{(n)}_k \) be the estimated 2D position of the eavesdropping drone at the \( n \)-th TS, and \( \mathbf{v}^{(n)}_k \) be the estimated velocity at the \( n \)-th TS, \( \mathbf{v}^{(n)}_k \in \Theta \triangleq \{||\hat{\mathbf{d}}^{(n)}_k - \mathbf{v}^{(n)}_k|| \leq \chi\} \), where \( \chi \) is the maximum estimation error and \( || \cdot || \) is the Euclidean norm. A2: The parameters of the A2G, ground-to-air (G2A), and air-to-air (A2A) channel models are accurately available at the network controller. A3: Due to shadowing and path loss, the uplink cellular users create negligible interference at the neighboring CBSs, and hence interference from the cellular users to the aerial links is not considered.

Remark 1: In the system considered, the eavesdropping drone can position itself in the vicinity of a legitimate drone. In the event that this happens and the legitimate drone is far from its associated CBS, the jamming signal might be weak and the eavesdropping drone might have a higher probability of intercepting the signal transmitted by a legitimate drone. This situation can be effectively tackled by exploiting co-channel inter-drone interference so that the deleterious impact of co-channel interference is minimized at the CBS and maximized at the eavesdropping drone. Note that the co-channel inter-drone interference of the proposed system jointly depends on the drone-clustering and transmit power allocation variables. Therefore, by optimizing the drone-clusters and the transmit power of the clustered drones, the SR can be improved even when the eavesdropping drone is far from the CBS and receives a weak jamming signal from it. To further enhance the SR, the collaboration among the drones and CBSs can be leveraged. In particular, the transmit power of each drone in a cluster can be optimally split between the confidential messages and AN signals such that the received jamming power at the eavesdropping drone is higher even when the eavesdropping drone is far from the CBS. In addition, multiple CBSs in the network can collaboratively transmit the weighted AN signals such that these AN signals coherently combine at the eavesdropping drone and maximize the received jamming power. These collaboration-based strategies for counteracting eavesdropping attacks are beyond the scope of this work and will be investigated in a future work.

B. Channel Model and Secrecy-Rate Expression

1) Channel Model: The A2G, G2A, and A2A communication channels are usually dominated by the LOS links [7], [8]. The A2G channel gain between the \( m \)-th legitimate drone and \( k \)-th CBS in the \( n \)-th TS follows the free space path loss model, given by \( h_{m,k}^{(n)} = \frac{\nu_{m,k}^{(n)}}{|q_{m,n} - \mathbf{r}^{(n)}_k|^2 + (H_d - H_e)} \), where \( \beta \) denotes the channel gain at the reference distance of 1 m. Similarly, the A2G channel gain between the \( k \)-th CBS and the eavesdropping drone in the \( n \)-th TS is expressed as \( f_{e,k}^{(n)} = \frac{\beta_2^{(n)}}{|q_{m,n} - \mathbf{v}_k^{(n)}|^2 + (H_d - H_e)} \)\(^2\). The A2A channel gain between the \( m \)-th legitimate and the eavesdropping drones in the \( n \)-th TS is expressed as \( g_{m,e}^{(n)} = \frac{\beta_2^{(n)}}{|q_{m,n} - \mathbf{v}_e^{(n)}|^2 + (H_d - H_e)} \)\(^2\), where \( \beta_2 \) denotes the channel gain at the reference distance of 1 m.

2) Secrecy-Rate Expression: In what follows, we provide the SR expression of the \( m \)-th legitimate drone in the \( n \)-th TS. Without loss of generality, we assume that in the \( n \)-th TS, the \( m \)-th legitimate drone is associated with the \( k \)-th CBS. Moreover, the set of drones associated with the \( k \)-th CBS in the \( n \)-th TS is denoted as \( \mathcal{S}_k^{(n)} \). Let \( P_{m,k}^{(n)} \) be the transmit power of the \( m \)-th drone in the \( n \)-th TS. The received signals at the \( k \)-th CBS and eavesdropping drone are expressed as

\[
y_k^{(n)} = \sum_{m=1}^{\mathcal{S}_k^{(n)}} P_{m,k}^{(n)} s_m + f_d c_0 + n_a,
\]

and

\[
y_e^{(n)} = \sum_{m=1}^{\mathcal{S}_k^{(n)}} P_{m,k}^{(n)} s_m + \sqrt{Q_{k}^{(n)}} f_{e,k}^{(n)} z_{AN} + n_a,
\]

respectively, where \( | \cdot | \) denotes the cardinality of a set. Here, \( s_m \) is the transmitted symbol from the \( m \)-th drone with \( |s_m|^2 = 1 \), \( f_d \) is the SI channel gain and \( c_0 \) is the residual SI at the CBS with \( c_0 \sim CN(0,\sigma^2) \) being the RF chain’s distortion, \( Q_{k}^{(n)} \) is the jamming power of the \( k \)-th CBS in the \( n \)-th TS, \( \epsilon \in (0,1) \) is SI cancellation coefficient, \( z_{AN} \sim CN(0,1) \) is the normalized AN signal, and \( n_a \sim CN(0,\sigma^2) \) is the additive white Gaussian noise.

Without loss of generality, the channel gains of the drones associated with the \( k \)-th CBS are sorted as \( h_{1,k}^{(n)} \geq h_{2,k}^{(n)} \geq \cdots \geq h_{\mathcal{S}_k^{(n)},k}^{(n)} \). The CBS applies SIC to decode the received signals following the decreasing order of the drones’ channel gains. However, our presented optimization framework can be applied following other decoding orders as well. Moreover, we consider a practical SIC where each drone’s signal is subjected to residual interference from the previously decoded signals. Based on [30, eq. (3)], the channel capacity of the data transmission link between the \( m \)-th legitimate drone and \( k \)-th CBS in the \( n \)-th TS is expressed as

\[
R_{m,k}^{(n)} = W \log_2 \left( 1 + \frac{P_{m,k}^{(n)} s_m^{(n)} P_{m,k}^{(n)} h_{m,k}^{(n)}}{I_{SIC} + I_{Sig} + e Q_{k}^{(n)} |f_d|^2 + \sigma^2} \right),
\]

where \( W \) is the bandwidth of the RRB reserved for aerial communications; \( I_{SIC} = \epsilon_c \sum_{l=1}^{\mathcal{S}_k^{(n)}} P_{l,k}^{(n)} h_{l,k}^{(n)} \) and \( I_{Sig} = \sum_{l=m+1}^{\mathcal{S}_k^{(n)}} P_{l,k}^{(n)} h_{l,k}^{(n)} \), \( \epsilon_c \in [0,1] \) is the SIC error coefficient, where \( \epsilon_c = 0 \) and \( \epsilon_c = 1 \) denote the ideal SIC and a scenario without any SIC, respectively.
In practice, it is more challenging to implement SIC-based decoding in a drone compared to a CBS. Similar to [22], we consider that the eavesdropping drone attempts to decode the transmitted signals of each drone-cluster by treating the co-channel inter-drone interference as a noise. The channel capacity of the eavesdropping link between the $m$-th legitimate and eavesdropping drones in the $n$-th TS is expressed as
\[
\hat{R}_{m}^{(n)} = W \log_{2} \left( 1 + \frac{P_{m,k}^{(n)} g_{m,e}}{\sum_{l=1, l \neq m}^{\left| S_{l}^{(n)} \right|} P_{l,k}^{(n)} g_{l,e} + Q_{k}^{(n)} f_{e,k} + \sigma^{2}} \right).
\]

The achievable SR of the $m$-th legitimate drone in the $n$-th TS is obtained as
\[
C_{m,k}^{(n)} = \left[ R_{m,k}^{(n)} - \max_{\nu_{e,k}^{(n)}} \hat{R}_{m}^{(n)} \right]^{+},
\]
where $[x]^{+} = \max(x, 0)$. Due to the uncertainty about the eavesdropping drone’s location, it is challenging to compute (5). To address this issue, we derive a tractable lower bound for (5).

In particular, we consider the eavesdropping drone’s location that results in the worst-case lower-bound of the achievable SR. Leveraging [16, eq. (5)] and [18, eq. (20)], we obtain an upper bound of the eavesdropping link’s channel capacity as
\[
\hat{R}_{m}^{(n)} = W \log_{2} \left( 1 + \frac{P_{m,k}^{(n)} g_{m,e}}{\sum_{l=1, l \neq m}^{\left| S_{l}^{(n)} \right|} P_{l,k}^{(n)} g_{l,e} + Q_{k}^{(n)} f_{e,k} + \sigma^{2}} \right),
\]
and where
\[
\hat{g}_{m,e}^{(n)} = \max_{\nu_{e,k}^{(n)}} g_{m,e}^{(n)} = \frac{\beta_{1}}{\left| \gamma_{m,n} - P_{m,k}^{(n)} \right| + \left| \chi_{m,n} \right| + (H_{d} - H_{e})},
\]
\[
\hat{g}_{m,k}^{(n)} = \min_{\nu_{e,k}^{(n)}} g_{m,k}^{(n)} = \frac{\beta_{2}}{\left| \gamma_{m,n} - P_{m,k}^{(n)} \right| + \left| \chi_{m,n} \right| + (H_{d} - H_{e})},
\]
\[
\hat{f}_{e,k}^{(n)} = \min_{\nu_{e,k}^{(n)}} f_{e,k}^{(n)} = \frac{\beta_{3}}{\left| \gamma_{m,n} - P_{m,k}^{(n)} \right| + \left| \chi_{m,n} \right| + (H_{d} - H_{e})}.
\]
Therefore, the worst-case SR of the $m$-th legitimate drone in the $n$-th TS is expressed as
\[
\hat{C}_{m,k}^{(n)} = \left[ R_{m,k}^{(n)} - \hat{R}_{m}^{(n)} \right]^{+}.
\]
Since (7) is analytically tractable, the ensuing problem formulation and analysis are developed using (7) instead of (5).

C. Problem Formulation

A drone-cluster is a set of the drones associated with the same CBS in a given TS. We denote $S_{m,k}^{(n)} = \{S_{1}^{(n)}, S_{2}^{(n)}, \ldots, S_{K}^{(n)}\}$ as the set of all drone-clusters in the $n$-th TS. Let $\alpha_{m,k}^{(n)} \in \{0, 1\}$ be a decision variable where $\alpha_{m,k}^{(n)} = 1$ if the $m$-th legitimate drone is associated with the $k$-th CBS in the $n$-th TS, otherwise $\alpha_{m,k}^{(n)} = 0$. Therefore, we obtain $S_{m,k}^{(n)} = \{m \in M | \alpha_{m,k}^{(n)} = 1\}, \forall k \in K$.

The optimization problem for maximizing the WC-ASR of the network is formulated as $P_{0}$, shown at the bottom of this page in (8). In $P_{0}$, $P_{\text{max}}$ denotes the maximum instantaneous transmit power of each legitimate drone; constraint $C_{1}$ ensures that the average SR of each legitimate drone is greater than or equal to a target SR, $R_{\text{t}}$; constraint $C_{2}$ provides a maximum average transmit power limit, $P_{\text{avg}}$, for each legitimate drone; constraint $C_{3}$ provides a maximum jamming power limit, $Q_{\text{max}}$, for each CBS; constraint $C_{4}$ mandates that the maximum interference caused by the drone-clusters to the neighboring CBSs(s) is not more than an acceptable interference level, $I_{th}$; constraint $C_{5}$ ensures that the drone-clusters are non-overlapping.

$P_{0}$ is a mixed-integer non-linear optimization problem. More specifically, the special instances of $P_{0}$, which are obtained by fixing either the power allocation or drone-clustering variables, are NP-complete. Consequently, $P_{0}$ is NP-hard and a global optimal solution to $P_{0}$ is computationally intractable. To address this issue, we propose a two-level iterative optimization approach. In particular, we decompose $P_{0}$ into the upper-level and lower-level sub-problems as follows.

**Upper-level sub-problem:** The upper-level sub-problem optimizes the drone-clusters in each TS. The essence of drone-clustering is to associate each data transmitting drone with a suitable CBS so that the sum-SR of the network is maximized. Based on such a fact, the upper-level sub-problem for the $n$-th TS, $\forall n = 1, 2, 3, \ldots, N$, is formulated as
\[
\text{P1:} \quad \alpha_{m,k}^{(n)} = \arg \max \left\{ \alpha_{m,k}^{(n)} \right\} \quad \alpha_{m,k}^{(n)} \in \{0, 1\}, \forall m \in M, \forall k \in K,
\]
\[
\text{s.t.} \quad C_{i}^{(n)} = \begin{cases} C_{m,k}^{(n)} + C_{\text{sys}}^{(n)}, & m = i \in M, \forall k \in K, \forall n \in \{1, 2, \ldots, N\}, \forall \alpha_{m,k}^{(n)} \in \{0, 1\} \end{cases}.
\]

**Lower-level sub-problem:** For a given set of drone-clusters, the lower-level sub-problem jointly optimizes the transmit power of the drones and the jamming power of the CBSs using the following optimization problem.

\[
\text{P2:} \quad \max_{0 \leq \{P_{m,k}^{(n)}\} \leq P_{\text{max}}}, \quad \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \max_{\alpha_{m,k}^{(n)}} \sum_{m=1}^{M} \sum_{k=1}^{K} \alpha_{m,k}^{(n)} C_{m,k}^{(n)}
\]
\[
\text{s.t.} \quad C_{i}^{(n)} \geq 0 \quad \alpha_{m,k}^{(n)} \in \{0, 1\}, \forall m \in M, \forall k \in K, \forall n \in \{1, 2, \ldots, N\}, \forall k \in K
\]

A converged sub-optimal solution to $P_{0}$ is obtained by iteratively solving $P_{1}$ and $P_{2}$. We first solve $P_{2}$ in Section III for any given set of drone-clusters. Thereafter, by leveraging the
solution of P2, P1 is solved in Section IV. Finally, the overall solution to P0 is presented in Section V.

III. POWER ALLOCATION SUB-PROBLEM’S SOLUTION

A. Proposed Solution Approach

Although P2 is non-convex, it can be solved in dual-domain without a notable loss of optimality, especially for large numbers of TSs. We therefore develop a dual-domain solution to P2. We first express the (partial) Lagrangian function of P2 as (11), shown at the bottom of the page.

In (11), \( \lambda, \rho, \) and \( \mu \) are the non-negative Lagrangian multipliers for the constraints C1, C2, and C4, respectively; \( p = [P^{(n)}_{m,k}]_{m \in M, k \in K, 1 \leq n \leq N}, q = [Q^{(n)}_{k}]_{k \in K, 1 \leq n \leq N}, \lambda = [\lambda_{m}]_{m \in M}, \rho = [\rho_{m}]_{m \in M}, \) and \( \mu = [\mu_{n,k}]_{k \in K, 1 \leq n \leq N} \). The dual problem of P2 is expressed as

\[
P3 : \min_{\lambda \geq 0, \rho \geq 0, \mu \geq 0} G(\lambda, \rho, \mu), \tag{12}
\]

where \( G(\lambda, \rho, \mu) \) is the Lagrangian dual function, defined as

\[
\begin{align*}
G(\lambda, \rho, \mu) &= \max_{p, q} \mathcal{L}(p, q, \lambda, \rho, \mu) \\
\text{s.t. } 0 \leq P^{(n)}_{m,k} \leq P_{\max}, 0 \leq Q^{(n)}_{k} \leq Q_{\max}, \forall m, k, n. \tag{13}
\end{align*}
\]

An optimal solution to (13) is required to solve the dual problem of P2. Evidently, (13) is separable per TS and decomposed into \( N \) optimization problems. The optimization problem for the \( n \)-th TS, where \( n \in \{1, 2, \ldots, N\} \), is expressed as

\[
P4 : \max_{p, q} \sum_{k=1}^{K} \sum_{m=1}^{M} (1 + \lambda_{m}) \alpha^{(n)}_{m,k} \tilde{r}^{(n)}_{m,k} - \sum_{k=1}^{K} \sum_{m=1}^{M} \left( \rho_{m} + \mu_{n,k} \max_{l \in L_{k}} \tilde{h}^{(n)}_{m,l} \right) \alpha^{(n)}_{m,k} P^{(n)}_{m,k} \tag{14}
\]

\[
\text{s.t. } 0 \leq P^{(n)}_{m,k} \leq P_{\max}, 0 \leq Q^{(n)}_{k} \leq Q_{\max}, \forall m, k. \tag{16}
\]

The optimal solution to P4 only depends on the drone-clusters in the current TS. Therefore, P2 can be optimally solved by repeating the following two steps in each TS: (i) solving P4 optimally for the given Lagrangian multipliers and (ii) updating the Lagrangian multipliers by minimizing the function \( G(\lambda, \rho, \mu) \) for the updated solution to P4.

B. Solution to P4 for the Given Lagrangian Multipliers

Using [12, Lemma 1], we can justify that the operator \([\cdot]^{+}\) can be omitted from the objective function of P4 without loss of optimality. Hence, P4 can be equivalently expressed as

\[
P5 : \max_{\lambda \geq 0, \rho \geq 0, \mu \geq 0} \mathcal{F}_{0}(p, q) \triangleq \sum_{k=1}^{K} \sum_{m=1}^{M} (1 + \lambda_{m}) \left( R^{(n)}_{m,k} - \tilde{R}^{(u,p,n)}_{m,k} \right) - \sum_{k=1}^{K} \sum_{m=1}^{M} \left( \rho_{m} + \mu_{n,k} \max_{l \in L_{k}} \tilde{h}^{(n)}_{m,l} \right) P^{(n)}_{m,k} \tag{17}
\]

\[
\text{s.t. } 0 \leq P^{(n)}_{m,k} \leq P_{\max}, 0 \leq Q^{(n)}_{k} \leq Q_{\max}, \forall m, k. \tag{19}
\]

Here, \( \{S^{(n)}_{1}, S^{(n)}_{2}, \ldots, S^{(n)}_{K} \} \) can be any feasible sets of drone-clusters in the \( n \)-th TS. Since the CBSs are associated with non-overlapping drone-clusters, P5 is separable per CBS. The optimization problem for the \( k \)-th CBS, \( \forall k \in K \), is written as

\[
\max_{\lambda \geq 0, \rho \geq 0, \mu \geq 0} \mathcal{F}_{k}(p, q) \triangleq \sum_{m \in S^{(n)}_{k}} (1 + \lambda_{m}) \left( R^{(n)}_{m,k} - \tilde{R}^{(u,p,n)}_{m,k} \right) - \sum_{m \in S^{(n)}_{k}} \left( \rho_{m} + \mu_{n,k} \max_{l \in L_{k}} \tilde{h}^{(n)}_{m,l} \right) P^{(n)}_{m,k} \tag{20}
\]

\[
\text{s.t. } 0 \leq P^{(n)}_{m,k} \leq P_{\max}, 0 \leq Q^{(n)}_{k} \leq Q_{\max}, \forall m \in S^{(n)}_{k}. \tag{21}
\]

(16) is a non-convex optimization problem, and thus solving (16) is non-trivial. To this end, we develop a low-complexity algorithm to near-optimally solve (16) by alternately updating the transmit and jamming power allocations. The detailed solution is provided as follows.

1) Jamming Power Allocation: Let us assume that the transmit power allocation of the legitimate drones is known. We first express the objective function of (16) as a difference-of-concave (DC) functions. To this end, using a change of variable \( Q^{(n)}_{k} = \frac{1}{\gamma_{k}} \), \( \forall k \), we obtain \( \mathcal{F}_{k}(p, q) = \mathcal{F}_{k}^{(1)}(p, u) \), where \( u = [u_{k}]_{k \in K} \), with

\[
\mathcal{F}_{k}^{(1)}(p, u) = \sum_{m \in S^{(n)}_{k}} \omega_{m} \left( 1 + \gamma^{(1)}_{m,k,n} \right) - \left( 1 + \gamma^{(2)}_{m,k,n} \right) \tag{22}
\]

\[
- \sum_{m \in S^{(n)}_{k}} \left( \rho_{m} + \mu_{n,k} \max_{l \in L_{k}} \tilde{h}^{(n)}_{m,l} \right) P^{(n)}_{m,k} \tag{23}
\]

In (17), \( \omega_{m} = \frac{W(1+\lambda_{m})}{\ln 2}, \gamma^{(1)}_{m,k,n} = \frac{u_{k} P^{(n)}_{m,k} \tilde{h}^{(n)}_{m,k} \gamma_{k} - \epsilon_{k}}{\sum_{l \in L_{k}} \tilde{h}^{(n)}_{m,l} \epsilon_{l}}, \) and \( \gamma^{(2)}_{m,k,n} = \frac{u_{k} P^{(n)}_{m,k} \tilde{h}^{(n)}_{m,k} \gamma_{k} + \epsilon_{k}}{\sum_{l \in L_{k}} \tilde{h}^{(n)}_{m,l} \epsilon_{l}} \). Therefore, for a given \( p \), we can equivalently express (16) as

\[
\max_{u \geq 0} \mathcal{F}_{k}^{(1)}(p, u), \text{s.t. } u_{k} \geq \frac{1}{Q_{\max}}, \forall k. \tag{24}
\]
(18) is a standard DC programming problem. We employ the SCA technique to solve (18). In particular, the first-order Taylor approximation is a global over-estimator of the concave function. Using such a fact, a lower bound to the objective function of (18) is obtained as $\mathcal{F}_{k}^{(1)}(p, u) \geq \mathcal{F}_{k}^{(2)}(p, u, \bar{u})$, where $\bar{u} = [\pi_{k}]_{k \in K}$ is a vector of the given feasible solution to (18) and where

$$
\mathcal{F}_{k}^{(2)}(p, u, \bar{u}) = \sum_{m \in S_{k}} \omega_{m} \left( \ln \left( 1 + \gamma_{m,k,n}^{(2)}(p, u) \right) - \ln \left( 1 + \gamma_{m,k,n}^{(2)}(p, u, \bar{u}) \right) \right) - \sum_{m \in S_{k}} \omega_{m} (u_{k} - \bar{u}_{k}) \Delta_{m}^{(1)}(\bar{u}_{k})
$$

and

$$
\Delta_{m}^{(1)}(\bar{u}_{k}) = \frac{\gamma_{m,k,n}^{(2)}(p, u, \bar{u})}{1 + \gamma_{m,k,n}^{(2)}(p, u, \bar{u})} \times \frac{1}{\sqrt{\omega_{m} \Delta_{m}^{(1)}(\bar{u}_{k})}}.
$$

An efficient solution to (18) is obtained by iteratively solving the following sequence of convex optimization problems.

$$
u^{(t)} = \arg \max_{\nu \geq 0} \mathcal{F}_{k}^{(2)}(p, u, \nu), \quad \text{s.t.} \quad u_{k} \geq \frac{1}{Q_{\max}}, \quad \forall k, \forall t
$$

where $u^{(t)}$ denotes the optimal solution for the $t$-th instance of (20), $\forall t \geq 1$. The optimal solution to (20) is obtained in the following proposition.

**Proposition 1:** An optimal solution to (20) is obtained at the limit point of the sequence $\{u_{k,j}^{(t+1)}\}$, where $1 \leq j \leq J_{\max}$, $\forall t \geq 1$, and $\forall k \in K$. Considering $u_{k,0}^{(t)}$ as a given initial point, each point of this sequence is iteratively computed as $u_{k,j+1}^{(t+1)} = f(u_{k,j+1}^{(t)})$, where

$$f(u_{k,j+1}^{(t)}) = \left[ \frac{\sum_{m \in S_{k}} \omega_{m} \gamma_{m,k,n}^{(2)}(p, u_{k,j+1}^{(t)})}{\sum_{m \in S_{k}} \omega_{m} \Delta_{m}^{(1)}(u_{k,j}^{(t)})} \right]_{\nu_{\max}},$$

with $[\cdot]_{\nu_{\max}}$ denoting the projection in the interval $[1/\nu_{\max}, \infty)$.

**Proof:** Due to space limitations, the proof has been moved to the extended version of the paper [36, Appendix A].

The jamming power allocation of the CBSs is optimized using the following iterative procedure. At first, the jamming power of the $k$-th CBS is initialized to some feasible value $Q_{k}$, and we set $u_{k}^{(0)} = (Q_{k})^{-1}, \forall k \in K$. Thereafter, at the $t$-th iteration, where $t \in \left\{1, 2, \ldots, T_{\max}\right\}$, $u_{k}^{(t)}$ is updated according to the limit point of the sequence presented in (21), i.e., $u_{k}^{(t)} \triangleq u_{k,j_{t}^{(t-1)}}$. The converged jamming power allocation of the $k$-th CBS is obtained as $Q_{k}^{(n)} = \frac{1}{\nu_{\max}}, \forall k \in K$.

2) **Transmit Power Allocation:** For a known jamming power allocation, using [31, Proposition 2], we can express (16) as

$$\max_{\nu \geq 0} W_{1} + W_{2} - \psi(p), \quad \text{s.t.} \quad P_{m,k}^{(n)} \in [0, P_{\max}], \forall m \in S_{k}^{(n)},$$

where $a = [a_{m}]_{m \in S_{k}^{(n)}}$ is a vector of auxiliary variables; $W_{1}$, $W_{2}$, and $\psi(p)$ are defined as

$$W_{1} = \sum_{m \in S_{k}^{(n)}} \omega_{m} \left( 1 + a_{m} \right),$$

$$W_{2} = \frac{\omega_{m}(1 + a_{m}) \gamma_{m,k,n}^{(1)}}{1 + \gamma_{m,k,n}^{(1)}} - \omega_{m} \left( 1 + \gamma_{m,k,n}^{(2)} \right),$$

and

$$\psi(p) = \sum_{m \in S_{k}^{(n)}} \left( \rho_{m} + \mu_{n,k} \times \max_{l \in L_{k}} h_{m,l}^{(n)} \right) P_{m,k}^{(n)}.$$

respectively. In (24), $\gamma_{m,k,n}^{(1)} = \frac{p_{m,k}^{(n)}}{\mu_{n,k} \times l_{m,k}^{(n)}}$ and

$$\gamma_{m,k,n}^{(2)} = \frac{\sum_{l=1}^{L_{k}} \omega_{m} h_{m,l}^{(n)} + \epsilon_{Q_{k}^{(n)}} (\omega_{m} h_{m,k}^{(n)} + \gamma_{m,k,n}^{(1)})}{\sum_{l=1}^{L_{k}} \omega_{m} h_{m,l}^{(n)} + \gamma_{m,k,n}^{(1)}} + \sigma_{Q_{k}^{(n)}}^{2},$$

(22) is decomposed into outer and inner optimization problems as

$$\max_{\nu \geq 0} W_{1} + W_{2},$$

and

$$\max_{p, \nu \geq 0} W_{3} - \psi(p), \quad \text{s.t.} \quad P_{m,k}^{(n)} \in [0, P_{\max}], \forall m \in S_{k}^{(n)},$$

where $W_{3}$ is defined in (29), shown at the bottom of this page. In (29), $\eta = [\eta_{m}]_{m \in S_{k}^{(n)}}$ is a vector of auxiliary variables for transforming the fractions into quadratic functions. For a fixed $\eta$, (29) is a non-convex function of the transmit power $p$. The SCA technique is utilized to solve (29) for a fixed $\eta$. A concave lower bound of $W_{3}$, with respect to $p$, is presented in (30), shown at the bottom of the next page. In (30), $p$ is a given feasible transmit power allocation vector. For a fixed $\eta$, an approximate solution to (28) is obtained by iteratively solving the following optimization problem,

$$\max_{p, \eta \geq 0} W_{3}^{(p, \eta)} - \psi(p), \quad \text{s.t.} \quad P_{m,k}^{(n)} \in [0, P_{\max}], \forall m \in S_{k}^{(n)},$$

and updating $\eta$ according to the solution obtained from (31).

It is noted that (26), (28), and (31) are the (unconstrained) convex optimization problems with respect to the variables $a$, $\eta$, and $p$, respectively. Therefore, the optimal solutions to these optimization problems can be directly obtained at the stationary
points of the objective functions. As a result, the optimal values of \( \{a_m\}, \{\eta_m\}, \) and \( \{p_m\}, \forall m \in S^{(n)}_k \), are obtained as

\[
a_m = \frac{\hat{z}_{m,k,n}^{(1)}(p,q)}{\eta_m}, \quad \eta_m = \sqrt{\omega_m(1 + a_m)P_{m,k}^{(n)}h_{m,k}^{(n)}} P_{m,k}^{(n)}h_{m,k}^{(n)} + I_{SIC} + I_{Sig} + \varepsilon Q_k^{(n)}|f_d|^2 + \sigma^2}, \quad (32)
\]

and

\[
P_{m,k}^{(n)} = \left[ \begin{array}{c} \omega_m(1 + a_m)h_{m,k}^{(n)} \eta_m \\ h_{m,k}^{(n)} \eta_m + \omega_m(\Delta_m^{(2)}(p,q) + D_{m,k}^{(n)}) \end{array} \right], \quad (34)
\]

respectively. In (34), \( \eta_m = \eta_m^2 + \varepsilon \sum_{l=1}^{m-1} \eta_l^2 + \sum_{l=m}^{N'} \eta_l^2 \), and \( D_{m,k} = \rho_m + \mu_k \times \max_{\ell < k} h_{m,\ell}^n \). For a known jamming power at the CBSs, a converged transmit power allocation is obtained by alternatively updating the variables \( a, \eta, \) and \( p \) using (32), (33), and (34), respectively. Such a fact is further confirmed by Proposition 2 of Section V.

3) Algorithm Development: The overall steps to solve P4 are summarized in Algorithm 1. Next, we analyze the computational complexity of Algorithm 1. In particular, Algorithm 1 iteratively executes two inner loops, where inner loop-I and inner loop-II update the jamming and transmit power allocations, respectively. The total number of iterations of both inner loops is \( T_{max} \times I_{max} \), and the total number of iterations of the outer loop is \( I_{max} \). The computational complexity associated to inner loop-I is \( O(T_{max} I_{max} K) \). Meanwhile, the computational complexity associated to inner loop-II is \( O(T_{max} I_{max} \sum_{k=1}^{K} |S_k^{(n)}|) = O(T_{max} I_{max} M) \). Therefore, the required computational complexity for executing a single iteration of the outer loop is \( O(T_{max} I_{max} (K + M)) \). The overall computational complexity of Algorithm 1 is obtained as \( O(\Delta_{max} (K + M)) \) where \( \Delta_{max} = T_{max} I_{max} M \).

C. Optimal Lagrangian Multipliers

We employ the well-known sub-gradient method to minimize \( G(\lambda, \rho, \mu) \) and find the optimal Lagrangian multipliers. The update equations of \( \lambda, \rho, \) and \( \mu \) are expressed as

\[
\lambda_m \leftarrow \lambda_m - \lambda^{(1)} \left( \frac{1}{N'} \sum_{n=1}^{N'} \sum_{k=1}^{K} a_m (\frac{P_{m,k}^{(n)}h_{m,k}^{(n)}}{P_{m,k}^{(n)}h_{m,k}^{(n)} + I_{SIC} + I_{Sig} + \varepsilon |f_d|^2 Q_k^{(n)} + \sigma^2 - \omega_m (1 + \hat{z}_{m,k,n}^{(2)}(p,q)) - \omega_m \ln(1 + \hat{z}_{m,k,n}^{(2)}(p,q)) \right)^2 \right) + \rho_m + \chi^{(2)} \left( \frac{1}{N'} \sum_{n=1}^{N'} \sum_{k=1}^{K} a_m (\frac{P_{m,k}^{(n)}h_{m,k}^{(n)}}{P_{m,k}^{(n)}h_{m,k}^{(n)} + I_{SIC} + I_{Sig} + \varepsilon |f_d|^2 Q_k^{(n)} + \sigma^2 - \omega_m (1 + \hat{z}_{m,k,n}^{(2)}(p,q)) - \omega_m \ln(1 + \hat{z}_{m,k,n}^{(2)}(p,q)) \right)^2 \right), \quad \forall m, \quad (35)
\]

\[
\rho_m \leftarrow \rho_m + \chi^{(2)} \left( \frac{1}{N'} \sum_{n=1}^{N'} \sum_{k=1}^{K} a_m (P_{m,k}^{(n)}h_{m,k}^{(n)} - P_{avg}) \right)^2 \right) + \eta_m, \quad (36)
\]

\[
\mu_{n,k} \leftarrow \mu_{n,k} + \chi^{(2)} \left( \max_{\ell < k} \sum_{m=1}^{M} \omega_m (P_{m,k}^{(n)}h_{m,k}^{(n)} - P_{avg}) \right)^2 \right) + \eta_m, \quad (37)
\]

Algorithm 1: Transmit and Jamming Power Allocation Algorithm for Solving P4 at the \( n \)-th TS.

1: **Input:** Drone-cluster sets \( \{S_1^{(n)}, S_2^{(n)}, \ldots, S_K^{(n)}\} \);
2:  Lagrangian multipliers \( \lambda, \rho, \) and \( \mu; \) maximum loop iterations \( I_{max}, J_{max}, \) and \( T_{max} \)
3: **Initialize:** Transmit power vector \( p \) and jamming power vector \( q; \) outer loop iteration index \( i = 1; \)
4: **repeat** (Start of outer loop)
5:  **for** Each \( k \in \{1, 2, \ldots, K\} \) **do**
6:  Initialize \( u_{k}^{(0)} = \frac{1}{Q_k^{(n)}} \), and iteration index \( t = 0; \)
7:  **while** \( t < T_{max} \) **do**
8:  Update \( u_{k}^{(t+1)} \) by iteratively executing (21), \( \forall j \in \{1, 2, \ldots, J_{max}\}. \)
9:  Assign \( u_{k}^{(t+1)} = u_{k}^{(t+1)} \)
10: **end**
11: **while** \( t < T_{max} \) **do**
12: **for** \( J = 1 : J_{max} \) **do**
13:  Update \( P_{m,k}^{(n)} \) using (34), \( \forall m \in S_k^{(n)} \)
14:  **end**
15: **end**
16: **end**
17: **end**
18: **end while** (End of inner loop-II)
19: **end for**
20: Increase outer loop iteration index \( i = i + 1. \)
21: **until** \( i > I_{max} \) **(End of outer loop)**
22: **Output:** Transmit power allocation \( \{P_{m,k}^{(n)}\} \) and jamming power allocation \( \{Q_{k}^{(n)}\}, \forall m \in S_k^{(n)} \) and \( \forall k \in K. \)

\[
W_s(p,q) = \sum_{m \in S_k^{(n)}} \omega_m (P_{m,k}^{(n)} - P_{m,k}^{(n)}) \times \left( \begin{array}{c} \hat{z}_{m,k,n}^{(2)}(p,q) \\ \frac{\hat{z}_{m,k,n}^{(2)}(p,q)}{P_{m,k}^{(n)}(1 + \hat{z}_{m,k,n}^{(2)}(p,q))} - \sum_{l=1}^{M} \frac{\hat{z}_{l,k,n}^{(2)}(p,q)}{P_{l,k}^{(n)}(1 + \hat{z}_{l,k,n}^{(2)}(p,q))} \times \frac{\hat{z}_{m,k,n}^{(2)}(p,q)}{\hat{z}_{l,k,n}^{(2)}(p,q)} \end{array} \right), \quad (30)
\]
respectively. Here, \( \xi(1) \), \( \xi(2) \), and \( \xi(3) \) are the positive and square-summable step-sizes. It is noted that C1 and C2 are the constraints on the average SR and average transmit power over the TSs, respectively, and C4 is an instantaneous interference threshold constraint. Hence, both \( \lambda \) and \( \rho \) are updated only at the end of each TS based on the average SR and average transmit power, respectively. Conversely, \( \mu \) is iteratively updated along with the instantaneous transmit power of the drones.

IV. DIALOGUE CLUSTERING SUB-PROBLEM’S SOLUTION

P1 is a combinatorial optimization problem of exponential complexity \( O(K^M) \). Thus, finding an optimal solution to P1 is prohibitive for large-scale systems. We apply the mixed-strategy repeated game framework [32] to solve P1, and the motivation is explained as follows. In particular, a prior knowledge of the achievable secrecy-rate (SR) of different drone-to-CBS associations is required. However, the achievable SR with different CBSs is influenced by the drone-to-CBS associations and the resultant inter-drone interference. Essentially, both drone-to-CBS association decisions and the resultant SR from the CBSs depend on each other. To solve this dilemma, it is imperative to better understand the achievable SR of different drone-to-CBS associations by exploring different drone-clusters. A mixed-strategy repeated game is used to fulfill this requirement. In a mixed-strategy repeated game, the players randomize their actions for each round of the game using certain probability mass functions (PMFs) and learn the achievable utility of different actions via exploration. Unlike in a one-shot game, the players in a mixed-strategy repeated game use some of the initial iterations to learn the behavior of other players and utility, and make more judicious decisions in later rounds [33]. Eventually, each player can select the most suitable strategy without having any prior knowledge of the other players’ strategies and utility of different actions. By exploiting a mixed-strategy repeated game with drone-clustering as the action, the network controller of our system model can learn the achievable SR of each drone with different CBSs and select the most suitable CBS for each drone. Thus, a mixed-strategy repeated game is well-suited to solve the drone-clustering problem considered in a computationally efficient manner.

A. Mixed-Strategy Repeated Game Formulation

The drone-clusters in the \( n \)-th TS, where \( n = 1, \ldots, N \), are obtained from the solution of a mixed-strategy repeated game \( \psi(n) = (\mathcal{M}, \{A_m^{(n)}\}_m \in \mathcal{M}, \{r_m^{(n)}\}_m \in \mathcal{M}) \). Here, the set of legitimate drones \( \mathcal{M} \) is the set of players. Each player \( m \in \mathcal{M} \) has a set of actions \( A_m^{(n)} = \{a_{m,1}^{(n)}, a_{m,2}^{(n)}, \ldots, a_{m,K}^{(n)}\} \), where taking action \( a_{m,k}^{(n)} \) implies the selection of the \( k \)-th CBS for the \( m \)-th player in the \( n \)-th TS. The joint action space of the players is denoted as \( \mathcal{A}^{(n)} = A_1^{(n)} \times A_2^{(n)} \times \cdots \times A_M^{(n)} \). At each round of the game, the \( m \)-th player selects an action from \( A_m^{(n)} \), \( \forall m \in \mathcal{M} \), leading to a joint action \( \pi \in \mathcal{A}(n) \). The utility function of the \( m \)-th player for the joint action \( \pi \in \mathcal{A}(n) \) is denoted as \( v_m^{(n)}(\pi) = \frac{C_{\pi m}(n)}{K_{\pi m}(n)} \), where \( C_{\pi m}(n) \) is the sum-SR of the network and it is defined in the objective function of P1. At each game round, based on the joint action of the players, a set of non-overlapping drone-clusters is formed. Thereafter, Algorithm 1 is executed for these drone-clusters, and near-optimal transmit and jamming power allocations are derived. By plugging such power allocations to the objective function of P1, the utility function of the players is determined.

For each player, there is a PMF over its action space, known as the mixed-strategy profile. Let \( \pi_m^{(n)} = \{\pi_{m,1}^{(n)}, \pi_{m,2}^{(n)}, \ldots, \pi_{m,K}^{(n)}\} \) be the mixed-strategy profile of the \( m \)-th player, where \( \pi_{m,k}^{(n)} \in [0, 1] \) is the \( m \)-th player’s probability of selecting the \( k \)-th CBS at the \( n \)-th TS, and \( \sum_{k=1}^{K} \pi_{m,k}^{(n)} = 1 \). The joint mixed-strategy profile of the players is denoted by \( \pi^{(n)} = \{\pi_1^{(n)}, \pi_2^{(n)}, \ldots, \pi_M^{(n)}\} \).

Due to the randomized actions of the players, the payoff of a joint mixed-strategy profile corresponds to the expected utility of the players. The expected utility of the \( m \)-th player in the \( n \)-th TS, \( \forall m \in \mathcal{M} \) and \( n = 1, 2, \ldots, N \), is obtained as

\[
\pi_m^{(n)}(\pi_m^{(n)}, \pi_{-m}^{(n)}) = \sum_{\pi \in \mathcal{A}(n)} \left( \prod_{j=1}^{M} \left( \sum_{k=1}^{K} \pi_{j,k}^{(n)} \right) \right)^{r_{m,k}^{(n)}}(\pi),
\]

where \( \pi_{-m}^{(n)} \) is the joint mixed-strategy profile of all the players except the \( m \)-th player, and \( \pi_{j,k}^{(n)}(\pi) = 1 \) if \( a_{m,n}^{(n)} \in \pi \), otherwise \( \pi_{j,k}^{(n)}(\pi) = 0 \).

From (38), the expected utility of a given player simultaneously depends on its own strategy and the strategies of the opponent players. Essentially, each player’s optimal strategy is a function of the opponent players’ strategies. We consider that the players are rational and choose the best-response strategies. The best-response strategy of the \( m \)-th player is defined as \( BR_m(\pi_m^{(n)}) \triangleq \arg \max_{\pi_{-m}^{(n)}} \pi_m^{(n)}(\pi_m^{(n)}, \pi_{-m}^{(n)}) \). The Nash-equilibrium (NE) solution of the game \( \psi(n) \) is formally defined as follows [34, Def. 3. 10].

**Definition 1:** A joint mixed-strategy profile \( \hat{\pi}^{(n)} = \{\hat{\pi}_1^{(n)}, \hat{\pi}_2^{(n)}, \ldots, \hat{\pi}_M^{(n)}\} \) is NE if, for each player \( m \in \mathcal{M} \), \( \hat{\pi}_m^{(n)} = BR_m(\hat{\pi}_{-m}^{(n)}) \). In particular, in the NE, the following condition is satisfied for each player \( m \in \mathcal{M} \).

\[
\pi_m^{(n)}(\hat{\pi}_m^{(n)}, \hat{\pi}_{-m}^{(n)}) \geq \pi_m^{(n)}(\hat{\pi}_m^{(n)}, \tilde{\pi}_{-m}^{(n)}) \quad \forall \tilde{\pi}_{-m}^{(n)}. \quad (39)
\]

In the NE, all the players jointly maximize the network sum-SR. Therefore, the NE solution of \( \psi(n) \) is sought to solve P1. To identify the NE solution of \( \psi(n) \), we first define the smooth best response (SBR) strategy of the \( m \)-th player, \( \forall m \in \mathcal{M} \), as follows.

**Definition 2:** For the opponents’ fixed joint strategy, \( \pi_{-m}^{(n)} \), the SBR strategy of the \( m \)-th player is defined as \( \beta_m^{(n)}(\pi_{-m}^{(n)}) \triangleq \{\beta_{m,1}^{(n)}(\pi_{-m}^{(n)}), \beta_{m,2}^{(n)}(\pi_{-m}^{(n)}), \ldots, \beta_{m,K}^{(n)}(\pi_{-m}^{(n)})\} \) [35, eq. (2.2)]. Here,

\[
\beta_{m,k}^{(n)}(\pi_{-m}^{(n)}) = \frac{\exp\left(\kappa_m^{(n)} r_m^{(n)}(k, \pi_{-m}^{(n)})\right)}{\sum_{k'=1}^{K} \exp\left(\kappa_m^{(n)} r_m^{(n)}(k', \pi_{-m}^{(n)})\right)}, \quad (40)
\]

where \( r_m^{(n)}(k, \pi_{-m}^{(n)}) \) is the expected utility of the \( m \)-th player obtained from the \( k \)-th individual action given that all other players select actions using the joint strategy \( \pi_{-m}^{(n)} \); \( \kappa_m > 0 \) is a positive parameter.

Let us consider that \( \hat{\pi}^{(n)} = \{\hat{\pi}_1^{(n)}, \hat{\pi}_2^{(n)}, \ldots, \hat{\pi}_M^{(n)}\} \) is a stable joint mixed-strategy profile for the players, where \( \hat{\pi}_m^{(n)} = \beta_m^{(n)}(\pi_{-m}^{(n)}) \), \( \forall m \in \mathcal{M} \). It is confirmed by Proposition 3 of Section V that \( \hat{\pi}^{(n)} \) can converge to the \( \varepsilon \)-equilibrium, where no player can improve its expected utility more than \( \varepsilon \) by unilaterally deviating from the converged strategy. More specifically,
the ε-equilibrium can be made sufficiently close to the NE when ε is small. Consequently, \( \pi^{(n)} \) can approach the NE solution of the game \( \psi^{(n)} \).

### B. Multi-Agent RL Framework to Learn the NE Solution

To compute the NE solution of the game \( \psi^{(n)} \) using (40), the network controller requires prior information about the utility and strategy models. In this case, the utility model refers to the expected utility from different joint actions of the players, and the strategy model refers to the distribution of the joint actions of the players. Since prior knowledge about both the utility and strategy models is unavailable, it is non-trivial to compute the NE solution of the game. To circumvent this issue, we exploit a multi-agent RL framework that can learn the strategy and utility models from the agents’ interactions with the environment. Essentially, the multi-agent RL framework is adopted to learn the NE solution of \( \psi^{(n)} \) without any prior statistical information.

\( \psi^{(n)} \) is modeled as a stateless multi-agent RL problem consisting of \( M \) virtual agents, each of which represents a drone. The action profile of the virtual agents is given by \( \pi = \{ \pi_m \in \mathbb{R}^n \}_m \). The instantaneous reward function of the \( m \)-th virtual agent, \( \forall m \in \mathcal{M} \), is denoted by \( r^{(n)}_m \). Both \( A_m \) and \( r^{(n)}_m \), \( \forall m \in \mathcal{M} \), are defined at the beginning of Section IV. A. The RL framework’s environment is modeled by the proposed cellular-connected IoD system. In each iterative game round (i.e., each episode of the RL problem), the virtual agents receive rewards for interacting with the environment and select new actions based on the rewards they receive. An interaction between the virtual agents and the environment is illustrated in Fig. 2. To facilitate the virtual agents’ interaction with the environment, we introduce the actor and critic RL processes as follows.

**Critic RL process:** Let \( V^{(n)}_{m,k}[l] \) be the value function for the \( k \)-th action of the \( m \)-th agent during the \( l \)-th round of the game. The value functions estimate the expected utility of the available actions. Let us denote \( V^{(n)}_{m,k} = \{ V^{(n)}_{m,1,k}, \ldots, V^{(n)}_{m,m,k} \} \). The critic RL process updates the value functions, \( \forall m \in \mathcal{M} \) and \( k = 1, 2, \ldots, K \), as

\[
V^{(n)}_{m,k}[l+1] = V^{(n)}_{m,k}[l] + \nu^{(l)}[l] \frac{r^{(n)}_{m} - V^{(n)}_{m,k}[l]}{\pi^{(n)}_{m,k}[l]},
\]

where \( \nu^{(l)}[l] \) is the critic’s learning step-size at the \( l \)-th round; \( \pi^{(n)}_{m,k}[l] \in (0, 1) \) is the actor’s estimated probability of selecting the \( k \)-th action at the \( l \)-th round; \( \nu^{(l)}[l] = 1 \) if the \( k \)-th action is selected for the \( m \)-th agent during the \( l \)-th round, otherwise \( \nu^{(l)}[l] = 0 \).

**Actor RL process:** Let \( \pi^{(n)}_{m}[l] = \{ \pi^{(n)}_{m,1}[l], \pi^{(n)}_{m,2}[l], \ldots, \pi^{(n)}_{m,K}[l] \} \) be the estimated mixed-strategy profile of the \( m \)-th agent during the \( l \)-th round of the game. The actor RL process updates the estimation of the mixed-strategy profile, \( \forall m \in \mathcal{M} \) and \( k = 1, 2, \ldots, K \), as

\[
\pi^{(n)}_{m,k}[l+1] = \pi^{(n)}_{m,k}[l] + \nu^{(2)}[l] \left( \beta^{(n)}_{m,k} \left( V^{(n)}_{m}[l] \right) - \pi^{(n)}_{m,k}[l] \right),
\]

where \( \nu^{(2)}[l] \) is the actor’s learning step-size at the \( l \)-th round.

In (41) and (42), the value function and mixed-strategy profile are initialized as \( \nu^{(n)}_{m}[0] = [0, 0, \ldots, 0] \) and \( \pi^{(n)}_{m}[0] = [\frac{1}{K}, \frac{1}{K}, \ldots, \frac{1}{K}] \), respectively, \( \forall m \in \mathcal{M} \). For the convergence of (41) and (42), the learning step-sizes should satisfy the following conditions: (i) \( \sum_{l=1}^{\infty} \nu^{(l)}[l] = \infty \), \( \sum_{l=1}^{\infty} \nu^{(l)}[l]^2 < \infty \), (ii) \( \sum_{l=1}^{\infty} \nu^{(2)}[l] = \infty \), \( \sum_{l=1}^{\infty} \nu^{(2)}[l]^2 < \infty \), and (iii) \( \lim_{l \to \infty} \nu^{(2)}[l] = 0 \) [35].

The proposed multi-agent RL framework consists of the following three steps. The **action selection step**, in which the virtual agent, \( \forall m \in \mathcal{M} \), samples an action from the \( A_m \) set using current mixed-strategy profiles; the **reward calculation step**, in which the drone-clusters are determined based on the actions selected by the virtual agents, Algorithm 1 is executed for these drone-clusters, and the virtual agents’ rewards are calculated using the derived power allocations; the **actor-critic update step**, in which the virtual agents’ value functions and mixed-strategy profiles are updated by plugging the calculated rewards into (41) and (42). The network controller learns the NE solution of \( \psi^{(n)} \) from the aforementioned three steps being iteratively repeated at each round of the game.

**Remark 2:** The parameter \( \kappa_m \) in (40) controls the trade-off between exploration and exploitation of the \( m \)-th player, \( \forall m \in \mathcal{M} \). It is evident that \( \lim_{\kappa_m \to 0} \beta^{(n)}_{m,k}(\mathbf{r}_m) = \frac{1}{K}, \quad \forall k = 1, 2, \ldots, K \). Therefore, when \( \{ \kappa_m \} \) is small, the mixed-strategy profiles converge to a uniform PMF, and the agents randomly explore different actions with equal probability. In contrast, one can readily show that \( \lim_{\kappa_m \to \infty} \beta^{(n)}_{m,k}(\mathbf{r}_m) = 1 \) and \( \lim_{\kappa_m \to \infty} \beta^{(n)}_{m,k \neq k}(\mathbf{r}_m) = 0 \), where \( k^* = \arg \max_{k \in \mathcal{K}} T^{(n)}(k, \pi^{(n)}_{m}) \). Thus, the best action for each agent is selected in a deterministic manner when \( \{ \kappa_m \} \) is asymptotically large. To learn the NE mixed-strategy profiles, it is imperative to make the agents explore available actions during the initial search, and exploit the best possible actions during the later search [35]. In this context, the value of \( \kappa_m, \forall m \in \mathcal{M} \), is gradually increased at each round of the game.

### V. OVERALL RESOURCE ALLOCATION ALGORITHM

#### A. Overview of the Proposed SIREN Algorithm

In Algorithm 2, the overall steps of SIREN are provided to solve the joint optimization problem P0. SIREN is implemented in a centralized manner. In each TS, SIREN provides a set of non-overlapping drone-clusters, the transmit power allocation of the drones, and the jamming power allocation of the CBSs using the following procedure. Each TS is divided into three intervals, i.e., namely, the transition interval, the scheduling interval, and the data transmission interval. In the transition interval, the drones move from their current stop-over point to the next stop-over point in their respective sensing zones and collect data. In the scheduling interval, Steps 4–12 of SIREN are
Algorithm 2: Security and Interference Management with REinforcement-learning and NOMA (SIREN) Algorithm.

1: Input: Maximum number of iterations, \( L_{\text{max}} \); coordinates of the stop-over points of the drones; number of TSs, \( N \).
2: Initialize: Lagrangian multipliers \( \lambda \) and \( \rho \).
3: for Each \( n = 1, 2, \ldots , N \) TS do
4: Initialize: \( \nu^{(m)}_0 \) and \( \pi^{(m)}_0 \), parameters \( k^{(m)} \), \( \forall m \in M \); iteration index \( l = 0 \); Lagrangian multiplier \( \mu \).
5: repeat
6: For each \( m \in M \) drone, sample a CBS using the mixed-strategy profile \( \pi^{(m)}_l \). Determine a set of drone-clusters, \( \{ S^{(1)}_1, S^{(2)}_2, \ldots , S^{(K)}_K \} \), based on the sampled drone-to-CBS associations.
7: Update the transmit and jamming power allocation by plugging the updated drone-clusters to Algorithm 1.
8: Determine the utility function (i.e., network sum-rate) for the updated drone-clusters and power allocations.
9: Update the value functions and mixed-strategy profiles using (41) and (42), respectively.
10: Update the parameter \( k^{(m)} = a^l - 1 \), where \( a > 1 \).
11: Update the Lagrangian multiplier \( \mu \) using (37); \( l = l + 1 \).
12: until \( l > L_{\text{max}} \).
13: Obtain the converged drone-clusters by associating each drone with the CBS having maximum value function. Obtain the final transmit and jamming power allocation.
14: Update the Lagrangian multipliers \( \lambda \) and \( \rho \) using (35) and (36), respectively.
15: end for
16: Output: Drone-clusters, transmit power allocation of the drones, and jamming power allocation of the CBSs for \( n = 1, 2, \ldots , N \) TSs.

Iteratively executed in the centralized network controller until the maximum number of iterations is reached. Following this, the network controller determines a set of non-overlapping drone-clusters at Step 13 of SIREN, and broadcasts the drone-to-CBS association matrix to all the CBSs in the network. Subsequently, each CBS determines the jamming power and the associated drones’ transmit power by executing Algorithm 1. Finally, the CBSs inform the associated drones about their transmit power allocation over reliable control channels. The network controller also updates the Lagrangian multipliers, \( \lambda \) and \( \rho \), as per Step 14 of SIREN. In the transmission interval, the drones transmit their collected data to the associated CBSs using the scheduled transmit power, and the CBSs also simultaneously transmit AN signals using the scheduled jamming power. The aforementioned resource allocation procedure is repeated in each TS.

Remark 3: From an implementation point of view, the computational delay required to execute SIREN needs to be orders of magnitude smaller than the duration of each TS. Note that SIREN requires a finite number of iterations for convergence as confirmed from our simulation results (Fig. 8(a) and (b)). Furthermore, the computational intensive part of SIREN, i.e., Steps 4–12, are implemented in a centralized network controller, which is usually hosted on an edge cloud platform and hence has powerful hardware for fast computation. For instance, SIREN’s execution can be accelerated by implementing Steps 4-12 using multiple parallel and commercial off-the-shelf (COTS) GPUs running in parallel. SIREN can therefore be rapidly executed by leveraging the centralized network controller’s fast computation capability.

B. Properties of the Proposed SIREN Algorithm

1) Convergence and Optimality: The following proposition capitalizes on the development of Algorithm 1, and confirms SIREN’s effectiveness at solving the transmit and jamming power allocation sub-problem.

Proposition 2: For a given set of drone-clusters, SIREN converges to a near-optimal solution to P2.
Proof: Due to space limitations, the proof has been moved to the extended version of the paper [36, Appendix B].

The following proposition capitalizes on the actor-critic multi-agent RL framework developed in Section IV.B., and confirms SIREN’s effectiveness at solving the drone-clustering sub-problem.

Proposition 3: For a sufficiently large number of iterations, i.e., \( L_{\text{max}} \to \infty \), SIREN converges to the NE solution of the drone-clustering game \( \psi^{(n)} \), \( \forall n = 1, 2, \ldots , N \).
Proof: Due to space limitations, the proof has been moved to the extended version of this paper [36, Appendix C].

We leverage these two propositions to confirm the local optimality of SIREN as follows.

Proposition 4: For a sufficiently large number of iterations, i.e., \( L_{\text{max}} \to \infty \), SIREN converges to a local optimal solution of the joint optimization problem P0.
Proof: Due to space limitations, the proof has been moved to the extended version of the paper [36, Appendix D].

2) Computational Complexity: We first determine the required computational complexity for SIREN in each TS. The computational complexity of SIREN in each TS is dominated by Steps 7, 9, 10, 11, 13 and 14. In particular, the required computational complexity for executing Step 7 is \( O(\Delta_{\text{max}}(K + M)) \). Meanwhile, for each drone, a total of \( (K + 1) \) computations are required to update the value functions and mixed-strategy profiles. Therefore, the required computational complexity of Step 9 is \( O(MK) \). Finally, Steps 10 and 11 of SIREN require \( O(M) \) and \( O(K) \) computations, respectively. Recall that the aforementioned steps are repeated for a total of \( L_{\text{max}} \) number of times in each TS, where \( L_{\text{max}} \gg 1 \). Conversely, Steps 13 and 14, whose required computational complexity is \( O(\Delta_{\text{max}}(K + M)) \) and \( O(M) \), respectively, are executed only once in each TS. Therefore, the required computational complexity for SIREN in each TS is approximated as \( O(L_{\text{max}}(MK + \Delta_{\text{max}}(K + M))) \). As a result, the overall computational complexity of SIREN is obtained as \( O(NL_{\text{max}}(MK + \Delta_{\text{max}}(K + M))) \). Evidently, SIREN requires a polynomial computational complexity to obtain a converged solution to P0.

3) Signaling Overhead: In each TS, SIREN requires the following information exchanges. First, at the end of scheduling interval of a TS, the network controller broadcasts the drone-to-CBS association matrix to all the CBSs in the network. This requires a total of \( MK \) information exchanges between the network controller and CBSs. Second, at the end of a TS, each CBS forwards the transmit power allocation information to its associated drones over reliable control channels. This requires a total of \( M \) information exchanges between the CBSs and drones. Consequently, a total of \( NM(K + 1) \) information exchanges are required by the SIREN algorithm.
TABLE I
SIMULATION PARAMETERS

| Parameter                          | Value |
|------------------------------------|-------|
| Jamming power limit, $Q_{\text{max}}$ | 10 watt |
| Maximum transmit power limit, $P_{\text{max}}$ | 5 watt |
| Drone’s average transmit power, $P_{\text{avg}}$ | 3 watt |
| SIC error coefficient, $\epsilon_s$ | $10^{-5}$ |
| SI cancellation coefficient of the CBSs, $\epsilon$ | $-90$ dBm |
| Acceptable interference level at the CBSs, $I_{\text{th}}$ | $-80$ dBm |
| Target SR, $R_t$ | 5 Mbps |
| AWGN variance, $\sigma^2$ | $-110$ dBm |

VI. NUMERICAL RESULTS

A. Simulation Setting

We consider a $1500$ m $\times$ $1500$ m geographical area, where six CBSs are placed at the vertices of a $600$ m $\times$ $600$ hexagonal region. We simulate five different system configurations consisting of $M = 5$, $10$, $15$, $20$ and $25$ data transmitting drones. Leveraging the circle packing tool [37], the entire geographical area of each system configuration is divided into $M$ non-overlapping sensing zones. Each sensing zone consists of $N = 30$ uniformly distributed stop-over points. For each system configuration, we repeat $25$ different simulation trials with independently generated stop-over points at the sensing zones, and present the average WC-ASR over all the independent simulation trials.

A circular flight trajectory is considered for the eavesdropping drone. The altitudes of the legitimate and eavesdropping drones are $H_L = 80$ m and $H_E = 50$ m, respectively. The height of each CBS is $H_b = 10$ m. The estimation error of the eavesdropping drone’s $2$D positions is $\chi = 10$ m, and the A2G/G2A and A2A channels’ gains at the reference distance are $\beta_1 = 10^{-5}$ and $\beta_2 = 10^{-4}$, respectively [17]. The SI channel gain, $f_d$, follows Rayleigh distribution. Each CBS employs an orthogonal RRB of bandwidth $W = 10$ MHz for aerial communications. Each CBS also shares its aerial RRB with all the other CBSs in the network for uplink cellular communications. The remaining simulation parameters, summarized in Table I, are selected based on [6], [7], [17]. Finally, to execute the proposed algorithms, we set the parameters $\nu^{(1)}(l) = \frac{1}{(1+a)^l}$, $\nu^{(2)}(l) = \frac{1}{(1+a)^l}$, $a = 1.1$, $J_{\text{max}} = 15$, $T_{\text{max}} = 100$, $I_{\text{max}} = 30$, and $I_{\text{max}} = 50$. Note that to avoid an infeasible solution to P0 for all the simulation settings considered, relatively large average and peak transmit power limits are considered for the legitimate drones. This is consistent with the existing literature [5], [6]. Nevertheless, by appropriately adjusting the Lagrangian multiplier, $\{\mu_{n,k}\}$, our proposed optimization framework confines the drones’ converged transmit power to a certain value that satisfies interference constraint $C4$ and prevents the drones from creating harmful interference at the terrestrial CBSs.

B. Performance Comparison With the Interference-Aware Benchmark Power Allocation Schemes

In this sub-section, we demonstrate the efficacy of SIREN in maximizing the achievable WC-ASR of the network compared to the following three interference-aware power allocation schemes.

1) Iterative Function Evaluation (IFE) Based Power Control: In the IFE scheme, the jamming and transmit power allocations are alternately updated, where the CBSs’ jamming power allocation is determined using Steps 4–9 of Algorithm 1 and the drones’ transmit power allocation is determined by satisfying the first order optimality condition of (16) with respect to $p$ [38].

2) Greedy Binary Power Control (GBPC): In GBPC, the jamming power of the CBSs and the transmit power of the legitimate drones are set to either the maximum or minimum values. To implement GBPC, the minimum and maximum values of the jamming power are set to $2$ W and $10$ W, respectively; the minimum and maximum values of the transmit power are set to $0.5$ W and $5$ W, respectively; Algorithm 1 of [39] is executed to alternately update the jamming and transmit power allocations at each drone-cluster until the maximum number of iterations is reached.

3) Fractional Power Control with Fixed Jamming Power (FPC-FJ): In FPC-FJ, the jamming power of each CBS is set to a fixed value of $2$ W, and the transmit power of the legitimate drones is optimized using Algorithm 3 of [31].

In Fig. 3, we compare the achievable WC-ASR of SIREN and the aforementioned three interference-aware power allocation schemes by varying the numbers of drones in the network. To make a fair comparison, the entire framework of SIREN is used for each benchmark power allocation scheme, while modifying only Step 7 of Algorithm 2. Fig. 3 shows that the IFE scheme provides a higher WC-ASR for small numbers of drones. Conversely, SIREN outperforms the IFE scheme for large numbers of drones. In particular, the inter-drone interference at the drone-clusters becomes substantial as the number of drones increases. Meanwhile, SIREN optimizes the drones’ transmit power by leveraging the quadratic transformation of fractional programming problems, which is highly effective for alleviating interference in the network [31]. Thanks to the efficient management of inter-drone interference, SIREN achieves an improved WC-ASR for large numbers of drones. It is observed from Fig. 3 that SIREN achieves $42.98\%$ higher WC-ASR than the IFE scheme for a network of $25$ drones. Fig. 3 shows that SIREN also outperforms both GBPC and FPC-FJ schemes for both small and large numbers of drones. In particular, since GBPC switches between the given maximum and minimum powers, the resultant jamming and transmit powers can be far from the optimal values. Meanwhile, FPC-FJ only takes the interference links into account, and ignores both eavesdropping and jamming links. Intuitively, both GBPC and FPC-FJ are sub-optimal, and thereby SIREN achieves substantial performance gains over both GBPC.
and FPC-FJ, especially for large numbers of drones. Fig. 3(b) depicts that for a network of 25 drones, SIREN achieves 1.99 times and 3.96 times more WC-ASR than the GBPC and FPC-FJ methods, respectively. We conclude that SIREN is interference resilient and particularly advantageous for dense IoD networks.

C. Performance Comparison With the Quasi-Exhaustive Search and Matching Empowered Drone-Clustering Methods

In Fig. 4, we compare the achievable WC-ASR of SIREN and a quasi-exhaustive search empowered drone-clustering method for various numbers of drones. In the quasi-exhaustive search method, the network controller determines (almost) all the possible combinations of drone-clusters, calculates the transmit and jamming power allocations for each combination of drone-clusters, and selects the drone-cluster combination that provides the largest WC-ASR for the network. By repeating these three steps $\Theta$ times at each TS, where $\Theta > 1$ is a predefined number, the network controller obtains near-optimal drone-clusters and the corresponding power allocation. Intuitively, the quasi-exhaustive search method can obtain the optimal WC-ASR, and thus outperforms SIREN. The overall computational complexity of the quasi-exhaustive search method is given by $O(N\Theta \Delta_{\text{max}}(M^2K + K^2M))$. It is clearly much greater when there is a large number of drones in the network. Conversely, SIREN’s required computational complexity increases linearly as the number of drones increases, and consequently, SIREN is scalable. Despite a significant reduction in the computational complexity, SIREN experiences only a small performance loss compared to the quasi-exhaustive search method. Fig. 4(a) shows that SIREN exhibits 6.9% smaller WC-ASR than the quasi-exhaustive search method when there are 25 drones in the network. Evidently, SIREN strikes a suitable balance between optimal performance and required computational complexity.

In Fig. 4, we also plot the achievable WC-ASR of a greedy matching empowered drone-clustering method for various numbers of drones. Here, we consider drone-clustering as a two-sided matching problem where each drone aims to be associated with the nearest CBS, and each CBS can be associated with a maximum of $N_d$ nearest drones. In particular, we set $N_d = 5$. This matching problem is near-optimally solved using a greedy algorithm. Note that the greedy matching method avoids iterating between the drone-clustering and power allocation phases, and thus requires a low computational complexity. However, due to the distance based drone-to-CBS associations, the drone-clusters generated by the greedy matching method can exhibit large inter-drone interference. As a result, the achievable WC-ASR of the greedy matching method is significantly reduced as the number of drones increases. Fig. 4 depicts that SIREN achieves 9.18 times more WC-ASR than the greedy matching empowered drone-clustering method when there are 25 drones in the network. Accordingly, SIREN is significantly more efficient than the greedy matching empowered drone-clustering method.

D. Performance Comparison With Benchmark RL Empowered Drone-Clustering Schemes

In this sub-section, the superiority of SIREN over two different RL empowered drone-clustering schemes is revealed. In the first benchmark scheme, the drone-clustering sub-problem P1 is solved by leveraging the stochastic learning automata (SLA) framework [40]. In the SLA empowered scheme, the network controller determines the resource allocation for each TS by iteratively carrying out the following three steps: (i) Update the mixed-strategy profiles of the drones by applying [40, eqs. (4), (5)]; (ii) determine a set of drone-clusters using the updated mixed-strategy profiles; (iii) calculate the transmit and jamming power allocations by executing Algorithm 1 for the updated drone-clusters. Fig. 5(a) compares the achievable WC-ASR of SIREN and SLA empowered scheme for various numbers of drones and six CBSs in the network. In the SLA empowered scheme, at each TS, the mixed-strategy profiles of the drones converge to stable PMFs. However, unlike SIREN, SLA does not provide any guarantee on the convergence of the mixed-strategy profiles to the SBR strategies. We emphasize that the SBR strategies obtain a pure-strategy NE solution, and thus near-optimally solve the sub-problem P1. Intuitively, compared to SIREN, the drone-clusters obtained by the SLA empowered method are sub-optimal. Accordingly, SIREN offers notable performance improvements over the SLA empowered method for both small and large numbers of drones. For instance, Fig. 5(a) illustrates that for a network of 25 drones, SIREN obtains 20% higher WC-ASR than the SLA empowered method.

In the second benchmark scheme, the drone-clustering sub-problem is solved by applying the multi-armed bandit (MAB) learning with upper confidence bound (UCB) policy [29]. Particularly, the MAB learning problem has a set of bandits, where each bandit has a set of arms to play at each game round. Here,
the goal is to learn suitable strategies for the bandits to play the arms so that the expected reward is maximized. To address the exploration-exploitation dilemma of the bandits, the UCB policy is utilized to select the arms. Such MAB-UCB method can be utilized to solve the sub-problem P1 by considering the drones as bandits and the CBSs as arms. More specifically, to obtain resource allocations using the MAB-UCB method, the network controller iteratively repeats the following two steps in each TS: (i) Select the CBSs for the drones by applying the UCB method [29, eq. (24)] and update the drone-clusters, and (ii) calculate the transmit and jamming power allocations by executing Algorithm 1 for the updated drone-clusters. In Fig. 5(a), we compare the achievable WC-ASR of SIREN and the MAB-UCB method for different numbers of drones. We can readily demonstrate that the MAB-UCB method generates stable drone-clusters. However, such drone-clusters are not necessarily the NE solution of sub-problem P1. Accordingly, the MAB-UCB method exhibits noticeable performance loss compared to the proposed SIREN algorithm. For instance, Fig. 5(a) illustrates that for a network of 25 drones, SIREN obtains 34.71% higher WC-ASR than the MAB-UCB method.

Fig. 5(b) compares the achievable WC-ASR of SIREN, SLA, and MAB-UCB methods for three CBSs and different numbers of drones in the network. Due to the lower diversity offered by A2G links, the achievable WC-ASR of all three schemes in Fig. 5(b) is reduced compared to Fig. 5(a). Nevertheless, SIREN outperforms both SLA and MAB-UCB methods for a small number of CBSs as well. For instance, Fig. 5(b) shows that SIREN obtains 8.25% and 38.09% higher WC-ASR than the SLA and MAB-UCB methods, respectively, with 5 drones in the network. Overall, SIREN is superior than both SLA and MAB-UCB empowered drone-clustering schemes.

E. Impact of SI Cancellation and SIC Error Coefficients

Fig. 6 presents the achievable WC-ASR of SIREN for different values of the SI cancellation coefficient, \(\epsilon\), in a network of 20 drones and six CBSs. When the values of \(\epsilon\) increase, the residual SI at the CBSs caused by the cooperative jamming increases and the channel capacity of the legitimate links strictly decreases. Hence, as observed from Fig. 6, the achievable WC-ASR of SIREN also decreases with the increase in \(\epsilon\). Recall that SIREN adapts the jamming power of the CBSs while taking the residual SI into account. To show the efficacy of SIREN over the fixed jamming power allocation strategy, the achievable WC-ASR of two SI agnostic schemes is plotted in Fig. 6. In the first SI agnostic scheme, denoted as “Max-JAM,” the CBSs always transmit the AN signals using the maximum jamming power. In the second SI agnostic scheme, denoted as “Low-JAM,” the CBSs always transmit the AN signals using a small jamming power. The fixed jamming power of the CBSs in the Low-JAM scheme is set to 0.1 W. When the values of \(\epsilon\) decrease, the residual SI at the CBSs is reduced and SIREN enables the CBSs to transmit the maximum jamming power. Fig. 6 thus shows that for sufficiently small values of \(\epsilon\), both SIREN and the Max-JAM scheme achieve almost the same WC-ASR. However, due to the severe residual SI at the CBSs, the achievable WC-ASR of the Max-JAM scheme is significantly reduced in the large \(\epsilon\) regime. Fig. 6 depicts that SIREN achieves 94.92 times higher WC-ASR than the Max-JAM scheme when \(\epsilon = -50\) dBm. Meanwhile, thanks to the small jamming power, the Low-JAM scheme exhibits small residual SI at the CBSs. In consequence, the Low-JAM scheme obtains higher WC-ASR than the Max-JAM scheme when \(\epsilon\) is large. Conversely, the Low-JAM scheme obtains higher WC-ASR than the Low-JAM scheme in the small \(\epsilon\) regime. The proposed SIREN, however, outperforms the Low-JAM scheme for both small and large values of \(\epsilon\). For instance, Fig. 6 shows that SIREN achieves 4.48 times higher WC-ASR than the Low-JAM scheme when \(\epsilon = 50\) dBm. In essence, thanks to the SI awareness, SIREN is more efficient than the fixed jamming power allocation schemes.

Fig. 7 illustrates SIREN’s achievable WC-ASR by varying the number of drones and the SIC error coefficients in a network of six CBSs. As expected, for a given number of drones, SIREN achieves the largest WC-ASR for ideal SIC at the CBSs, whereas an increase in SIC error coefficients reduces the achievable WC-ASR. Fig. 7 also shows that the multi-user diversity gain is reduced for large SIC error coefficients. This observation is intuitive since the resultant co-channel inter-drones interference at the CBSs is considerably larger for large drone-clusters and large SIC error coefficients. As a result, increasing numbers of drones provides a small increment in achievable WC-ASR. Fig. 7 shows that as the number of drones increases from 5 to 25, the achievable multi-user diversity gain of practical SIC is reduced by 20.27%, 64.81%, and 74.06% for \(\epsilon_c = 10^{-5}\), \(\epsilon_c = 10^{-3}\), and \(\epsilon_c = 10^{-2.5}\), respectively, with respect to ideal SIC. We conclude that the capability of performing accurate SIC at the CBSs notably impacts the proposed algorithm’s PLS performance.
Fig. 8. Convergence performance of SIREN.

**F. Convergence of SIREN**

Fig. 8(a) illustrates the evolution of the CBS selection probability of a single drone in a network of 20 drones and six CBSs with respect to the iterations of SIREN, $L_{\text{max}}$. In Fig. 8(a), at the beginning, the CBS selection probability of the drone in question is $\{0.1, 0.1, 0.1, 0.1, 0.1, 0.1\}$, i.e., the probability of associating this drone with a certain CBS is the same for all the CBSs in the network. As $L_{\text{max}}$ increases, the network controller’s knowledge of the drone’s achievable SR with different CBSs improves, and it is then able to select the most suitable CBS for the drone. Consequently, the PMF used to select the CBS for the drone becomes deterministic. Fig. 8(a) shows that the PMF used to select a CBS for the drone in question almost converges to $\{0, 0, 0, 0, 0, 0\}$ for large values of $L_{\text{max}}$, i.e., the drone is associated with the second CBS in the network. Likewise, we can show that the CBS selection PMFs used for other drones also become deterministic. This implies that drone-clusters become stable for large values of $L_{\text{max}}$. Moreover, when the drone-clusters become stable, consistent from *Proposition 2*, the transmit and jamming power allocations also become invariant. This fact leads to the convergence of the overall resource allocation. To confirm this, we plot the network’s achievable WC-ASR in Fig. 8(b) for various iterations of SIREN. We observe that WC-ASR increases with $L_{\text{max}}$. This is due to the fact, that as $L_{\text{max}}$ increases, the network controller can make more informed decisions about achieving a high network WC-ASR. Fig. 8(b) shows that the WC-ASR becomes almost stable after 35 iterations of the SIREN algorithm for both small and large numbers of drones. We therefore conclude that the convergence of SIREN to a stable resource allocation is guaranteed.

**VII. CONCLUSION**

We investigated a resource optimization scheme to manage interference and enhance the PLS of multi-drone cellular-connected IoT networks. A joint optimization problem was formulated to maximize the network’s WC-ASR while considering clustering of the legitimate drones, NOMA-enabled data transmission, and cooperative jamming from the FD CBSs. A two-level iterative optimization approach was devised to address the computational intractability of the joint optimization problem. A multi-agent RL framework was employed to optimize the drone-clusters, and the transmit power of the legitimate drones and jamming power of the CBSs were optimized by leveraging fractional programming, SCA, and alternating optimization techniques. A centralized and convergent resource allocation algorithm, entitled SIREN, was proposed to solve the joint optimization problem. Simulation results revealed that the proposed SIREN algorithm attains a small performance loss compared to the quasi-exhaustive search empowered method with a significantly reduced computational complexity. Simulation results also confirmed that in dense IoD networks, the proposed SIREN algorithm manages inter-drone interference more efficiently than the existing interference-aware power allocation schemes, and achieves a higher WC-ASR than the benchmark RL empowered drone-clustering schemes.

**ACKNOWLEDGMENT**

This article was supported in part by funding from the Innovation for Defence Excellence and Security (IDEaS) program from the Department of National Defence (DND). Any opinions and conclusions in this work are strictly those of the author(s) and do not reflect the views, positions, or policies of-and are not endorsed by-IDEaS, DND, or the Government of Canada.

**REFERENCES**

[1] H. Zhang, L. Song, Z. Han, and H. V. Poor, “Cooperation techniques for a cellular internet of unmanned aerial vehicles,” *IEEE Wireless Commun.*, vol. 26, no. 5, pp. 167–173, Oct. 2019.

[2] M. Z. Hassan, G. Kaddoum, and O. Akhrif, “Interference management in cellular-connected Internet-of-drones network with drone clustering and uplink rate-splitting multiple access,” *IEEE Internet Things J.*, early access, Feb. 17, 2022, doi: 10.1109/JIOT.2022.3152382.

[3] U. Challita, A. Ferdowsi, M. Chen, and W. Saad, “Machine learning for wireless connectivity and security of cellular-connected UAVs,” *IEEE Wireless Commun.*, vol. 26, no. 1, pp. 28–35, Feb. 2019.

[4] W. Mei and R. Zhang, “Aerial-ground interference mitigation for cellular-connected UAV,” *IEEE Wireless Commun. Mag.*, vol. 28, no. 1, pp. 167–173, Feb. 2021.

[5] J. Yao and N. Ansari, “QoS-aware power control in internet of drones for data collection service,” *IEEE Trans. Veh. Technol.*, vol. 68, no. 7, pp. 6649–6656, Jul. 2019.

[6] J. Yao and N. Ansari, “Wireless power and energy harvesting control in IoD by deep reinforcement learning,” *IEEE Trans. Green Commun. Netw.*, vol. 5, no. 2, pp. 980–989, Jun. 2021.

[7] Y. Huang, W. Mei, J. Xu, L. Qiu, and R. Zhang, “Cognitive UAV communication via joint maneuver and power control,” *IEEE Trans. Commun.*, vol. 67, no. 11, pp. 7872–7888, Nov. 2019.

[8] W. Mei, Q. Wu, and R. Zhang, “Cellular-connected UAV: Uplink association, power control and interference coordination,” *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5380–5393, Nov. 2019.

[9] A. S. Matar and X. Shen, “Joint subchannel allocation and power control in licensed and unlicensed spectrum for multi-cell UAV-cellular network,” *IEEE J. Sel. Areas Commun.*, vol. 39, no. 11, pp. 3542–3554, Nov. 2021.

[10] B. Li, Z. Fei, Y. Zhang, and M. Guizani, “Secure UAV communication networks over 5G,” *IEEE Wireless Commun.*, vol. 26, no. 5, pp. 114–120, Oct. 2019.

[11] J. Yao and N. Ansari, “Secure federated learning by power control for internet of drones,” *IEEE Trans. Cogn. Commun. Netw.*, vol. 7, no. 4, pp. 1021–1031, Dec. 2021.

[12] G. Zhang, Q. Wu, M. Cui, and R. Zhang, “Securing UAV communications via joint trajectory and power control,” *IEEE Trans. Wireless Commun.*, vol. 18, no. 2, pp. 1376–1389, Feb. 2019.

[13] Y. Cai, Z. Wei, R. Li, D. W. K. Ng, and J. Yuan, “Joint trajectory and resource allocation design for energy-efficient secure UAV communication systems,” *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 4536–4553, Jul. 2020.

[14] W. Wang et al., “Robust 3D-trajectory and time switching optimization for dual-UAV-enabled secure communications,” *IEEE J. Sel. Areas Commun.*, vol. 39, no. 11, pp. 3334–3347, Nov. 2021.

[15] K. Xu, M. -M. Zhao, Y. Cai, and L. Hanzo, “Low-complexity joint power allocation and trajectory design for UAV-enabled secure communications with power splitting,” *IEEE Trans. Commun.*, vol. 69, no. 3, pp. 1896–1911, Mar. 2021.
IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, VOL. 71, NO. 12, DECEMBER 2022

C. Zhong, J. Yao, and J. Xu, “Secure UAV communication with cooperative jamming and trajectory control,” IEEE Commun. Lett., vol. 23, no. 2, pp. 286–289, Feb. 2019.

W. Wang et al., “Energy-constrained UAV-assisted secure communications with position optimization and cooperative jamming,” IEEE Trans. Commun., vol. 68, no. 7, pp. 4476–4489, Jul. 2020.

Y. Zhou et al., “Secure communications for UAV-enabled mobile edge computing systems,” IEEE Trans. Commun., vol. 68, no. 1, pp. 376–388, Jan. 2020.

Y. Zhou, F. Zhou, H. Zhou, D. W. K. Ng, and R. Q. Hu, “Robust trajectory and transmit power optimization for secure UAV-enabled cognitive radio networks,” IEEE Trans. Commun., vol. 68, no. 7, pp. 4022–4034, Jul. 2020.

Y. Liu, Z. Qin, Y. Cai, Y. Gao, G. Y. Li, and A. Nallanathan, “UAV communications based on non-orthogonal multiple access,” IEEE Wireless Commun., vol. 26, no. 1, pp. 52–57, Feb. 2019.

X. Sun, D. W. K. Ng, Z. Ding, Y. Xu, and Z. Zhong, “Physical layer security in UAV systems: Challenges and opportunities,” IEEE Wireless Commun., vol. 26, no. 5, pp. 40–47, Oct. 2019.

N. Zhao et al., “Security enhancement for NOMA-UAV networks,” IEEE Trans. Veh. Technol., vol. 69, no. 4, pp. 3994–4005, Apr. 2020.

H. M. Wang and X. Zhang, “UAV secure downlink NOMA transmissions: A secure users oriented perspective,” IEEE Trans. Commun., vol. 68, no. 9, pp. 5732–5746, Sep. 2020.

Y. Li, H. Zhang, and K. Long, “Joint resource, trajectory, and artificial noise optimization in secure driven 3D UAVs with NOMA and imperfect CSI,” IEEE J. Sel. Areas Commun., vol. 39, no. 11, pp. 3363–3377, Nov. 2021.

K. Singh, K. Wang, S. Biswas, Z. Ding, F. A. Khan, and T. Ratnarajah, “Resource optimization in full duplex non-orthogonal multiple access systems,” IEEE Trans. Wireless Commun., vol. 18, no. 9, pp. 4312–4325, Sep. 2019.

H. Shakhatreh et al., “Unmanned aerial vehicles (UAVs): A survey on civil applications and key research challenges,” IEEE Access, vol. 7, pp. 48572–48634, Apr. 2019.

X. Lin et al., “The sky is not the limit: LTE for unmanned aerial vehicles,” IEEE Commun. Mag., vol. 56, no. 4, pp. 204–210, Apr. 2018.

Q. S. Jabbal, M. Aliuzzaman, A. Baz, H. Allakami, and J. Ben-Othman, “Dispatch of UAVs for urban vehicular networks: A deep reinforcement learning approach,” IEEE Trans. Veh. Technol., vol. 70, no. 12, pp. 13174–13189, Dec. 2021.

A. H. Arani, M. M. Azari, P. Hu, Y. Zhu, H. Yankomeroglu, and S. Safavii-Naeini, “Reinforcement learning for energy-efficient trajectory design of UAVs,” IEEE Internet Things J., vol. 9, no. 11, pp. 9060–9070, Jun. 2022.

A. Celik, M. Tsai, R. M. Radaydeh, F. S. Al-Qahtani, and M. Alouini, “Distributed cluster formation and power-bandwidth allocation for imperfect NOMA in DL-HetNets,” IEEE Trans. Commun., vol. 67, no. 2, pp. 1677–1692, Feb. 2019.

K. Shen and W. Yu, “Fractional programming for communication systems with power control and beamforming,” IEEE Trans. Signal Process., vol. 66, no. 10, pp. 2616–2630, May 2018.

S. Perkins, P. Mertikopoulos, and D. S. Leslie, “Mixed-strategy learning with continuous action sets,” IEEE Trans. Autom. Control, vol. 62, no. 1, pp. 379–384, Jan. 2017.

L. Busoni, R. Babuska, and B. De Schutter, “A comprehensive survey of multigame reinforcement learning,” IEEE Trans. Syst. Man Cybern.-part C Appl. Rev., vol. 38, no. 2, pp. 156–172, Mar. 2008.

Z. Han et al., “Game Theory in Wireless and Communication Networks: Theory, Models and Applications, Cambridge, U.K.: Cambridge Univ. Press, 2011.

D. S. Leslie and J. E. Collins, “Individual Q-learning in normal form games,” SIAM J. Control Optim., vol. 44, no. 2, pp. 495–514, Aug. 2005.

M. Z. Hassan et al., “Resource allocation for joint security and interference management in NOMA enabled and cellular-connected Internet-of-Things systems,” TechRxiv Preprint, 2002. [Online]. Available: https://doi.org/10.36227/techrxiv.19141907.v2

[Online]. Available: https://doi.org/10.36227/techrxiv.19141907.v4

H. Dahrouj et al., “Power spectrum optimization for interference mitigation via regretive function evaluation,”EURASIP J. Wireless Commun. Netw., vol. 2012, pp. 1–14.

A. Gjendemsjo, D. Gesbert, G. E. Oien, and S. G. Kiani, “Binary power control for sum rate maximization over multiple interfering links,”IEEE Trans. Wireless Commun., vol. 7, no. 8, pp. 3164–3173, Aug. 2008.

P. Vrancea, K. Verbeek, and A. Nowe, “Decentralized learning in Markov games,”IEEE Trans. Syst. Man Cybern., vol. 38, no. 4, pp. 976–981, Aug. 2008.

R. D. Yates, “A framework for uplink power control in cellular radio systems,” IEEE J. Sel. Areas Commun., vol. 13, no. 7, pp. 1341–1347, Sep. 1995.

Md. Zoheb Hassan received the Ph.D. degree from The University of British Columbia, Vancouver, BC, Canada, in 2019. He was a Postdoctoral Research Fellow with the School of Engineering, École de Technologie Supérieure, Montreal, QC, Canada, The University of British Columbia, Kelowna, BC, Canada. Since then, he has been a Research Assistant Professor with Wireless@VT, Bradley Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA, USA. His research interests include resource optimization for wireless communications, signal processing for interference management and detection, and application of machine learning in physical layer communications. Dr. Hassan was or is currently a Member of the Technical Program Committee of IEEE ICC 2019–2023, IEEE GLOBECOM 2021–2022, and IEEE VTC 2022-Fall. He was the recipient of the Natural Science and Engineering Research Council Postdoctoral Fellowship of Canada, in 2021.

Georges Kaddoum (Senior Member IEEE) received the bachelor’s degree in electrical engineering from the École Nationale Supérieure de Techniques Avancées (ENSTA Bretagne), Brest, France, and the Ph.D. degree in telecommunications and signal processing (circuits, systems, and signal processing) from the Université de Bretagne Occidentale and Telecom Bretagne (ENSBT), Brest, France, in 2005, and the Ph.D. degree (Hons.) in signal processing and telecommunications from the National Institute of Applied Sciences (INSA), University of Toulouse, Toulouse, France, in 2009. He is currently an Associate Professor and Tier 2 Canada Research Chair with the École de Technologie Supérieure (ETS), Université du Québec, Québec City, QC, Canada. Since 2010, he has been a Scientific Consultant in the field of space and wireless telecommunications with several US and Canadian companies. He has authored or coauthored more than 200 journals and conference papers and has two pending patents. His current research interests include mobile communication systems, modulations, security, and space communications, and navigation. Dr. Kaddoum was the recipient of the Best Papers Awards at the 2014 IEEE International Conference on Wireless and Mobile Computing, Networking, Communications (WIMOB), with three coauthors, and at the 2017 IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC), with four coauthors, and ETS Research Chair award in physical-layer security for wireless networks, in 2014. He was also the recipient of IEEE Transactions on Communications Exemplary Reviewer Award, for the year 2015, 2017, 2019, Research Excellence Award, Université du Québec, in 2018, and Research Excellence Award from the ETS in recognition of his outstanding research outcomes, in 2019. He is currently an Associate Editor for IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, and IEEE COMMUNICATIONS LETTERS.

Ouassima Akhrif (Senior Member IEEE) received the M.Sc. and Ph.D. degrees as a Fulbright Scholar in electrical engineering and control systems from the University of Maryland, College Park, College Park, MD, USA, in 1987 and 1989, respectively. She was an Assistant Professor for one year with the Systems Engineering Department, Case Western Reserve University, Cleveland, OH, USA. She moved to Montréal, Canada, and joined, in 1992, the “École de Technologie Supérieure (ETS),” Université du Québec, Québec City, QC, Canada, where she is currently a Full Professor with the Electrical Engineering Department. She is a Member of GREPCI (Groupe de Recherche en Génie Industriel et Electronique de Puissance), a research group that she directed from 2004 to 2009. Her research interests include bifurcation analysis, nonlinear geometric control, nonlinear adaptive control and their applications in electric drives, power systems, renewable energy integration, autopilot design, and flight control systems.