Distributed Cooperation Under Uncertainty in Drone-Based Wireless Networks: A Bayesian Coalitional Game

Vandana Mittal, Setareh Maghsudi, and Ekram Hossain, *Fellow, IEEE*

Abstract—**We study the resource sharing problem in a drone-based wireless network. We consider a distributed control setting under uncertainty (i.e. unavailability of full information). In particular, the drones cooperate in serving the users while pooling their spectrum and energy resources in the absence of prior knowledge about different system characteristics such as the amount of available power at the other drones. We cast the aforementioned problem as a Bayesian cooperative game in which the agents (drones) engage in a coalition formation process, where the goal is to maximize the overall transmission rate of the network. The drones update their beliefs using a novel technique that combines the maximum likelihood estimation with Kullback-Leibler divergence. We propose a decision-making strategy for repeated coalition formation that converges to a stable coalition structure. We analyze the performance of the proposed approach by both theoretical analysis and simulations.**

Index Terms—**Drone-based wireless network, Bayesian cooperative games, coalition formation, uncertainty, distributed resource sharing, Kullback-Leibler divergence**

I. INTRODUCTION

The current wireless networking paradigm makes a significant step towards building up the ultra-reliable, low latency, power and spectral efficient communication with the help of technologies such as millimeter wave and massive MIMO communications, seamless integration of licensed and unlicensed bands, as well as intelligent spectrum usage and management [1]. Nonetheless, the rapid increase of wireless traffic calls for further enhancement through the integration of aerial base stations (e.g. drones) in order to improve the coverage, capacity, and connectivity of existing terrestrial cellular networks. The potential of using drones as aerial base stations stems from their altitude flexibility and the possibility of establishing the line-of-sight (LoS) link towards ground users. Despite the huge potential, drone-based wireless communication systems face a variety of challenges including precise channel modeling, efficient users-to-drones association, interference management, trajectory optimization, and resource management and control, etc. [2].

In general, centralized management and control of a network requires the availability of the global network information at a central controller that performs network optimization.

Vandana Mittal and Ekram Hossain are with Department of Electrical and Computer Engineering, University of Manitoba (E-mail: mittalv@myumanitoba.ca, ekram.hossain@umanitoba.ca). Setareh Maghsudi is with the Department of Electrical Engineering and Computer Science, Technical University of Berlin, 10623 Berlin, Germany (E-mail: maghsudi@tu-berlin.de).

These methods, however, suffer from excessive overhead and computational cost. From this point of view, it is important to develop distributed solution methods that are robust to uncertainty and information shortage. One mathematical tool that enables rigorous analysis of multi-agent systems under uncertainty is the cooperative game theory. An example is [3] that investigates distributed task management in networked cyber-physical systems.

Based on the flying mechanism of drones, they are broadly classified into two categories as fixed-wing drones and rotary-wing drones [4], [5]. Fixed-wing drones usually can carry heavy payload and can travel at a high speed. However, they need to maintain continuous forward motion to remain aloft; Therefore, they are not suitable for stationary applications. Rotary-wing drones, despite having limited payload capacity and mobility, can either move in any direction or stay stationary in the air. Thus, such drones can be used as hovering drones at a certain location to ensure continuous coverage. However, they consume a significant amount of power to keep them hovering in the air all the time. A particular type of drone is selected depending on the application. In this paper, we consider the rotary-wing hovering drones, each placed at a fixed location to provide maximum coverage to its assigned users. We assume that the available power at each drone is time-varying. This variation in the available power in a drone is due to the fact that the power supply and energy dissipation at each drone can vary depending on its type, mode of operation, location during communication (and hence propagation condition), etc. As a specific example, a drone with energy harvesting capability can harvest energy from stochastic resources and then communicate with the ground users or base stations by using the harvested energy [4]. Since energy harvesting can be intermittent and uncertain, the amount of available energy will be statistically varying.

In the setting described above, we address the distributed resource sharing problem, where there is some uncertainty about network parameters such as the power availability in each drone. More specifically, to enhance network performance, the drones can make cooperative clusters. The drones within a cluster can share their available power and spectrum resource to optimize the service provided to the users. Given no prior information about the statistical characteristics of the available power, the goal is to find an optimal structure of drone clusters along with the best channel assignment and power allocation to each user in the cluster such that the overall transmission rate (social welfare) in the network is maximized.

To solve the described problem, we formulate it as a Bayesian Coalition Formation Game (BCFG), where the drones represent the players. Each drone, being uncertain about the available power, or the *type*, of others, forms some belief. Given the beliefs, the drones engage in a coalition formation process that aims at optimizing the service for the users assigned to the coalition members. By observing the outcome of its action and possibly some side-information, each drone then updates its belief. The process continues until convergence. The main contributions of the paper include the following:

- For a drone-based wireless network, we formulate the resource sharing problem under uncertainty as an optimization problem. The joint optimization problem includes (i) learning he *types* along with finding the optimal cluster configurations of the drones, (ii) user-channel assignment, and (iii) power allocation in each cluster. Since this is a combinatorial problem, we divide the optimization problem into three sub-problems. We cast the cooperation problem among drones as a Bayesian coalition formation game. We propose a distributed approach based on best-reply dynamics to obtain the optimal clusters of drones. In each cluster, the user-association and channel allocation problem is formulated as a bipartite matching problem. Finally, we use a water-filling algorithm for power allocation.

- For numerical analysis, we simulate a variety of network settings. We investigate the performance of our proposed solution, as well as the effect of uncertainty and cooperation, by implementing several benchmarking approaches. These include the baseline configuration in which the drones perform independently without cooperation, distributed coalition formation with full information, and the social optimal case. The results establish the superior performance of the proposed scheme in terms of improvement in the sum rate of the network and individual rate of the drones. We also show the effect of overlap in the distribution of the *types* on convergence.

- Concerning the methodology of Bayesian coalition formation, we propose a novel belief-updating method. More precisely, instead of updating the belief using Bayes' rule, our approach first estimates the parameters of the distribution (of *type*) and then finds the closeness of the estimated parameters with the given set of types followed by averaging. We also use local information to update the belief; that is, the drones share the information only inside the corresponding coalition, thereby reducing the feedback and signaling overhead.

Compared to the state-of-the-art, our approach offers the following advantages:

- It is more scalable since we assume limited use of a central controller (e.g. a software-defined controller) in the cooperation process or the availability of precise information about the critical variables.

- It can be used for networks where the users or drones leave/join the network dynamically. Upon detection of the changes in the network dynamics or type sets (e.g. by the controller), the algorithm can be triggered to execute using the updated information.

- The proposed coalition formation model and decision-making strategy do not depend on the subsequent resource allocation or the statistical characteristics (such as the distribution) of the available power at the drones. Therefore, it is highly adaptable to different systems beyond the drone networks.

The organization of the paper is as follows. **Section II** provides a brief overview on the state-of-the-art research. **Section III** describes the system model. We formulate the resource sharing problem in **Section IV**. In **Section V**, we model the formulated problem as a Bayesian coalition formation games, and **Section VI** includes the algorithmic solutions. **Section VII** presents the theoretical analysis. In **Section VIII**, we evaluate the performance numerically. **Section IX** concludes the paper.

II. RELATED WORK

Resource management and energy efficiency for drone-assisted cellular wireless networks have attracted significant attention of the research community. Reference [6] proposes an energy-efficient scheduling framework for cooperative drone communication. In [7], the authors propose an optimal resource allocation strategy for an energy-harvesting flying access point. Reference [8] studies the problem of flight time optimization and bandwidth allocation of the drones that serve ground users. The authors of [9] propose a resource allocation framework for cache-enabled drones that provide services to the ground users over unlicensed and licensed bands. Reference [10] studies optimal user and drone assignment for capacity improvement in drone-assisted heterogeneous wireless networks. In [11], the authors jointly optimize the number of active drones and their locations to maximize the coverage. In [12], the authors propose a framework based on stochastic game theory to optimize the performance of multi-drone networks by a joint selection of power levels, sub-channels, and users. The authors in [13] propose a novel method for the strategic placement of multiple drones along with base stations in a large scale network. Reference [14] studies the backhaul-aware optimal placement of drones and base stations to maximize the number of served users. Reference [15] develops a framework consisting of swarms of UAVs as flying relays for delay-intolerant and bandwidth-hungry applications. Reference [16] presents a method for resource allocation and trajectory planning for multiple UAVs that deliver data in vehicular networks. Reference [17] studies the positioning of nodes in a dynamic UAV swarm network with the goal of optimal throughput communication. In [18], the authors develop a framework for UAV-enabled wireless-powered Internet of Things. They maximize the sum throughput of the network by joint optimization of UAV placement, time allocation, and the UAV-device association. Reference [19] investigates joint optimization of UAV location planning, content placement, transmit beamforming, and user admission decisions. The objective is to maximize the number of served users while satisfying the minimum rate requirements, where the UAVs have limited storage capacity. Authors of [20] study

drone-mounted in-band full-duplex heterogeneous networks. They maximize the network's transmission performance by joint optimization of drone placement, power and bandwidth allocation, and user assignment.

Beyond drone networks, there is a large body of literature that study the resource allocation and user association problems in heterogeneous networks. For instance, [21] formulates a centralized framework to analyze and compare various user association-, resource allocation-, and interference management schemes. Similarly, the authors in [22] study joint optimization of user association, power allocation, and channel allocation in multi-cell multi-association OFDMA heterogeneous networks. They decompose the problem into two subproblems. The sub-problems are then solved alternatively to obtain the local optimal solution. Reference [23] develops a multi-agent reinforcement learning-based distributed solution.

From the methodology perspective, coalition formation game without uncertainty has been a popular tool to solve wireless communication problems. For example, in [24], the authors propose distributed cooperation among single antennas to form the virtual multi-antenna system to improve network performance. In [25], users collaborate for sharing the sub-channels in the cognitive LTE femtocells environment. Cooperation among femtocell access points by sharing the excess computational resources is proposed in [26]. The goal is cost- and delay reduction by avoiding unnecessary offloading to the remote cloud. References [27], [28], and [29] study the application of coalition formation game in wireless sensors networks. Specifically, [27] considers the network's lifetime expansion with the desired quality of service requirements. Reference [28] proposes to balance the energy efficiency and QoS provisioning for cooperation in a clustered wireless sensor network. Reference [30] shows the application coalition formation game for cooperative networks with simultaneous wireless information and power transfer. A few papers use coalition formation game with uncertainty to solve the resource allocation problem in wireless networks. For example, the authors of [31] use the Bayesian coalition game with nontransferable utility (NTU) for packet delivery among mobile nodes under uncertainty in node behavior. Reference [32] proposes utilizing coalition formation based on Bayesian reinforcement learning for distributed resource sharing in device-to-device (D2D) enabled heterogeneous cellular networks. Similarly, [33] presents a Bayesian overlapping coalition formation game for spectrum sharing between multiple co-located cellular networks and a set of D2D links. Reference [34] models the dynamics of coalition formation games for spectrum sharing in an interference channel. Reference [35] studies the cooperative task allocation problem among heterogeneous cyber-physical systems under uncertainty about the randomly arrived tasks and stochastic systems' types. The authors model the problem as a multi-state stochastic cooperative game with state uncertainty. Reference [36] solves the cooperative user association problem under uncertainty using an approach based on the exchange economy.

None of the aforementioned works investigate the problem of distributed cooperation among drones for resource sharing under uncertainty considering realistic network and channel propagation conditions.

III. SYSTEM MODEL AND ASSUMPTIONS

In this section, we describe the system model. **Table I** summarizes the most important variables that are frequently used in this paper.
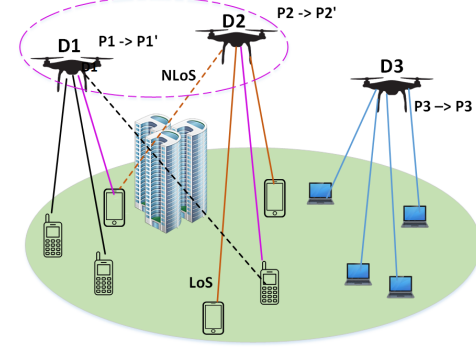

Figure 1. Network model: Drones $D1$ and $D2$ cooperate by forming a coalition to ensure the possibility of line-of-sight (LoS) transmission, thereby improving the transmission performance.

A. Network Model

As shown in **Fig. 1**, we consider a drone network with multiple hovering drones with fixed locations. Each drone shall serve a pre-defined set of users that are distributed uniformly over a wide geographical area. This happens, among others, when the drones belong to different service providers, or when the drones support the overload traffic from the terrestrial cellular system. We represent the set of drones, channels, and users respectively by $\mathcal{D} = \{1, \ldots, D\}$, $\mathcal{Q} = \{1, \ldots, Q\}$, and $\mathcal{N} = \{1, \ldots, N\}$. Each drone $d \in \mathcal{D}$ has access to Q_d orthogonal channels, each with bandwidth B. The drone allocates each channel only to one of its assigned user. Therefore, the number of users that can be served by a drone at each time instant is at most equal to the number of available channels. Primarily, each drone operates independently of the others, routing the traffic of its own users only. We refer to this configuration as the *baseline network*. In this configuration, a drone is placed at the centroid location of its users. We use a k-means clustering algorithm to find the centroid location for all the drones [37].

We consider that the available power at each drone is a random variable. The distribution of this random variable is referred to as the *type* of the drone. Every drone knows its own type; however, concerning other drones, it only knows the set of possible types, i.e., a set of possible distributions. The possible types depend on the environment and the location of each drone so that it can be inferred from the historical data. Each drone maintains a belief about the types of other drones.

B. Channel Model

Let $\mathcal{N}_d$ and $\mathcal{Q}_d$ denote the set of users and channels of the drone $d \in \mathcal{D}$. Moreover, let user $n \in \mathcal{N}_d$ and drone $d \in \mathcal{D}$ be located at $(x, y, 0)$ and (x_d, y_d, h_d), respectively. Then the

Table I
IMPORTANT NOTATIONS AND DEFINITIONS

Notation	Definition
$\mathcal{D}$, $\mathcal{N}$, and $\mathcal{Q}$	Set of drones, users, and channels
$\mathcal{D}^{C_k}$, $\mathcal{N}^{C_k}$, and $\mathcal{Q}^{C_k}$	Set of drones, users, and channels in coalition C_k
$\mathcal{W}$	Set of coalition structures for D drones with cardinality W
$w = \{C_1, \ldots, C_k, \ldots, C_l\}$	Coalition structure $w \in \mathcal{W}$ consisting of l coalitions
w^*	Nash stable coalition structure
$\mathbf{X}_w, \mathbf{Y}_w$	Set of channel and user- assignment matrices for coalition structure w
$\mathbf{p}^{C_k} = [p_1, \ldots, p_n, \ldots, p_{\mathcal{N}^{C_k}}]$	Power vector consisting of power of each user in coalition C_k
$\bar{L}_{dn,q}$	Average path loss of drone d towards user n over channel q
$R_{dn}(q, p_n)$	Transmission rate of user n from drone d over channel q with power p_n
x_{qn}, y_{dn}	Binary variables with value 1 if an assignment exists and 0 otherwise
P_d	Power of drone d
$P^{C_k} = \sum_{d=1}^{\mathcal{D}} P_d$	Total power of the coalition C_k
$\mathcal{T} = \{T_1, \ldots, T_M\}$	Type set consisting of M types
$t^d \in \mathcal{T}, P_d = \mathbb{E}(t^d)$	Type of drone d, power of drone (expected value of its type)
$\mathbf{T} \in \otimes_{d=1}^{D} \{T_1, T_2, \ldots, T_M\}$	Type space of D drones
$\mathbf{T}^{-d} \in \otimes_{x \in \mathcal{D} \backslash d} \{T_1, T_2, \ldots, T_M\}$	Type space of $D \backslash d$ drones
$\mathbf{t}^{-d} = [t_1^d, \ldots, t_j^d \ldots, t_{\mathcal{D} \backslash d}^d] \in \mathbf{T}^{-d}$	Belief space of drone d
$\mathbf{t}_{C_k} \in \mathbf{T}_{C_k} = \otimes_{d \in C_k} \mathcal{T}$	Type vector of drones in coalition C_k belonging to the type space $\mathbf{T}_{C_k}$
$\mathbf{t}_{C_k}^{-d} \in \mathbf{T}_{C_k}^{-d} = \otimes_{j \in \{C_k \backslash d\}} \mathcal{T}$	Type vector based on drone d about the members $j \in \{C_k \backslash d\}$ from type space $\mathbf{T}_{C_k}^{-d}$
$B(\mathbf{t}_{C_k}^{-d})$	Belief of drone d about the type vector $\mathbf{t}_{C_k}^{-d}$
$P(\mathbf{t}_{C_k}^{-d})$	Power of coalition C_k as a function of type vector $\mathbf{t}_{C_k}^{-d}$
$\mathcal{A}^{C_k}$	Set of coalition actions consisting of all possible channel associations and power allocations
$\bar{q}_d^d, \bar{q}_j^d$	Expected pay off of drone d and drone j based on drone d's belief about members of C_k
$\bar{q}_{total}$	Total expected payoff of the network
$\mathbf{W}, \rho_{w,w'}$	Transition matrix of the Markov chain of BCFG, Transition probability from w to w'
π_w	Formation probability for coalition structure w

path-loss of the downlink communication from drone d to user n over the channel q is given by [38]

$$L_{dn,q}[dB] = 20 \log \left(\frac{4\pi f_c d_{dn}(x,y)}{c} \right) + \zeta_{dn} + 10 \log(\Omega_{dn}), \tag{1}$$

where $d_{dn}(x,y) = \sqrt{(x_d - x)^2 + (y_d - y)^2 + (h_d)^2}$ is the distance between drone d and user n. Moreover, f_c and c are the carrier frequency and the speed of light, respectively. Also, ζ_{dn} is the average loss due to the free-space propagation, which depends on the environment. If the wireless link between drone d and user n is LoS, then $\zeta_{dn}^{\text{LoS}} = N(\mu_{\text{LoS}}, \sigma_{\text{LoS}}^2)$; otherwise, the non-line-of-sight (NLoS) link has a loss of $\zeta_{dn}^{\text{NLoS}} = N(\mu_{\text{NLoS}}, \sigma_{\text{NLoS}}^2)$. Therefore, ζ_{dn}^{LoS} and ζ_{dn}^{NLoS} are shadow fading with normal distribution in dB scale for LoS and NLoS links. The expected value and variance of the shadow fading for LoS and NLoS links are $(\mu_{\text{LoS}}, \sigma_{\text{LoS}}^2)$ and $(\mu_{\text{NLoS}}, \sigma_{\text{NLoS}}^2)$, respectively. The variance depends on the elevation angle and type of the environment [39], i.e.

$$\sigma_{\text{LoS}}(\theta_{dn}) = k_1 \exp(-k_2 \theta_{dn}) \tag{2}$$
$$\sigma_{\text{NLoS}}(\theta_{dn}) = g_1 \exp(-g_2 \theta_{dn}), \tag{3}$$

where $\theta_{dn} = \sin^{-1}(h_d/d_{dn})$ is the elevation angle between drone d and user n. The parameters k_1, k_2, g_1, and g_2 are constants that depend on the environment. Let Ω_{dn} be a random variable capturing the effects of the small-scale fading between drone d and user n with $\bar{\Omega}_{dn} = 1$. The random variable Ω_{dn} follows a non-central chi-square probability distribution, given by [39]

$$f_{\Omega_{dn}}(\rho) = \frac{(K+1)e^{-K}}{\bar{\Omega}_{dn}} e^{-\frac{(K+1)\rho}{\bar{\Omega}_{dn}}} I_0 \left(2\sqrt{\frac{(K+1)\rho}{\bar{\Omega}_{dn}}} \right). \tag{4}$$

In (4), $\rho \geq 0$, and $I_0(\cdot)$ is the zero-order modified Bessel function of the first kind. Moreover, K is the Rician factor that corresponds to the ratio of the power of the LoS component and the power of the multipath components. For $K = 0$, the Rician model reduces to a Rayleigh fading distribution. Hence, in general a small value of K represents that multipath component is dominant whereas a large value corresponds to a strong LoS between the drone and user. Thus, $\Omega_{dn}^{\text{LoS}} = f_{\Omega_{dn}}(\rho)$ and $\Omega_{dn}^{\text{NLoS}} = f_{\Omega_{dn}}(\rho)|_{K=0}$. The Rician factor K depends on some factors out of which the elevation angle θ between the drone and user plays the dominant role. We consider [39]

$$K = \psi(\theta) = ae^{b\theta}, \tag{5}$$

where a and b are some constants whose values depend on the system parameters and the environment. The constants a and b are determined as

$$a = k_0, \quad b = \frac{2}{\pi} \left(\ln \frac{k_{\frac{\pi}{2}}}{k_0} \right), \tag{6}$$

where k_0 and $k_{\frac{\pi}{2}}$ could be determined from measurements in a concrete scenario [39]. Let p_{dn}^{LoS} be the probability of existence of the LoS link between a drone d and user n, given by [38]

$$p_{dn}^{\text{LoS}} = \alpha \left(\frac{180}{\pi} \theta_{dn} - \theta_o \right)^{\gamma}, \tag{7}$$

where α and γ are constant values reflecting the environment impact. Moreover, θ_o is the minimum angle between a user and a drone. Then $p_{dn}^{\text{NLoS}} = 1 - p_{dn}^{\text{LoS}}$ is the probability of having a NLoS link. The average path loss from drone d to user n is

$$\bar{L}_{dn,q} = p_{dn}^{\text{LoS}} L_{dn,q}^{\text{LoS}} + p_{dn}^{\text{NLoS}} L_{dn,q}^{\text{NLoS}}. \tag{8}$$

The signal-to-interference-plus-noise ratio (SINR) from drone d to a user n over channel q yields

$$\text{SINR}_{dn,q} = \frac{p_n G / \bar{L}_{dn,q}}{B_w (N_0 + I_0)}, \tag{9}$$

where p_n is the average transmission power of drone d towards user n, N_0 and I_0 are noise power and interference power spectral density (e.g. due to co-channel interference from other networks), respectively, B_w is the channel bandwidth, and G is the antenna gain. The downlink transmission rate is then given by

$$R_{dn}(q, p_n) = B_w \log(1 + \text{SINR}_{dn,q}). \tag{10}$$

IV. RATE MAXIMIZATION IN DRONE NETWORKS: PROBLEM FORMULATION

A. Optimization Problem Formulation

Primarily, each drone serves its assigned users independently. However, this might not be optimal in several scenarios, as the users associated with one drone might have better channels (e.g. LoS link) to some other drone. Moreover, the amount of available power is not identical for all drones and also changes over time. Consequently, promoting cooperation among drones improves network performance significantly. For intelligent cooperation, the drones can be divided into clusters. Inside each cluster, the drones share their radio resources including spectrum and energy (e.g. energy sharing among drones can be possible through wireless power transfer [40]). They also cooperate in serving users, irrespective of the initial association. In brief, by shuffling users and pooling the resources, the drones in each cluster improve the aggregate performance in terms of the transmission rate. To formalize the problem, we note the following:

1) The drones are clustered into disjoint groups. Let $\mathcal{W}$ with cardinality W be the set of all possible partition structures. For any partition structure $w \in \mathcal{W}$, we have $w = \{C_1, \cdots, C_k, \cdots, C_l\}$, where $C_k \cap C_{k'} = \phi$ and $\bigcup_{k=1}^{l} C_k = \mathcal{D}$. Let $C_k \in w$ denote an arbitrary coalition. Moreover, $\mathcal{D}^{C_k}$ is the set of drones in C_k. Also, $\mathcal{N}^{C_k}$ and $\mathcal{Q}^{C_k}$ are the set of channels and users that the drones in C_k share.

2) Inside each cluster C_k, any user $n \in \mathcal{N}^{C_k}$ shall be served by one drone, say $d \in \mathcal{D}^{C_k}$. Each drone can serve multiple users. Matrix $\mathbf{Y}^{C_k}$ of dimension $D^{C_k} \times N^{C_k}$ represents the user assignment. Therefore, in coalition structure w with l coalitions, there exists l such assignment matrices. We gather these as $\mathbf{Y}_w = \{\mathbf{Y}^{C_1}, \ldots, \mathbf{Y}^{C_k}, \ldots, \mathbf{Y}^{C_l}\}$. The element y_{dn} of the matrix $\mathbf{Y}^{C_k}$ is defined as follows:

$$y_{dn} = \begin{cases} 1, & \text{if user } n \text{ and drone } d \text{ are associated;} \\ 0, & \text{otherwise.} \end{cases} \tag{11}$$

3) In addition to user assignment, each channel $q \in \mathcal{Q}^{C_k}$ is allocated to one user $n \in \mathcal{N}^{C_k}$. Matrix $\mathbf{X}^{C_k}$ with dimension $Q^{C_k} \times N^{C_k}$ denotes the channel assignment. The set of channel allocation matrices for a partition

structure w is $\mathbf{X}_w = \{\mathbf{X}^{C_1}, \ldots, \mathbf{X}^{C_k}, \ldots, \mathbf{X}^{C_l}\}$. Each element x_{qn} of the matrix $\mathbf{X}^{C_k}$ is defined as

$$x_{qn} = \begin{cases} 1, & \text{if channel } q \text{ is allocated to user } n; \\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

4) Inside each cluster, the drones optimize the power allocation. Let P_d denote the available power at drone d. The total power of the group C_k then yields $P^{C_k} = \sum_{d=1}^{\mathcal{D}^{C_k}} P_d$, where we neglect the loss that might occur due to energy sharing (e.g. via wireless power transfer). The power P^{C_k} is then allocated to the users $\mathcal{N}^{C_k}$. Let $\mathbf{p}^{C_k} = [p_1, \ldots, p_n, \ldots, p_{\mathcal{N}^{C_k}}]$ be the allocation vector, where $\sum_{n=1}^{\mathcal{N}^{C_k}} p_n \leq P^{C_k}$.

Let $R_{dn}(q, p_n)$ be the transmission rate provided by drone $d \in \mathcal{D}^{C_k}$ to user $n \in \mathcal{N}^{C_k}$ over channel $q \in \mathcal{Q}^{C_k}$ with power p_n. Based on the discussion above, the objective is to select the partition $w \in \mathcal{W}$, the user association matrices $\mathbf{Y}_w$, the channel allocation matrices $\mathbf{X}_w$, and the power allocation $\mathbf{p}^{C_k}$ in each cluster $C_k \in w$ to maximize the total transmission rate. Formally, for all $C_k \in w$, $n \in \mathcal{N}^{C_k}$, $q \in \mathcal{Q}^{C_k}$, $d \in \mathcal{D}^{C_k}$,

$$\underset{w, \mathbf{Y}_w, \mathbf{X}_w, \mathbf{p}^{C_k}}{\text{maximize}} \sum_{C_k \in w} \sum_{n \in \mathcal{N}^{C_k}} R_{dn}(q, p_n) x_{qn} y_{dn} \tag{13a}$$

$$\text{s.t.} \quad \sum_{d=1}^{\mathcal{D}^{C_k}} y_{dn} = 1, \quad y_{dn} \in \{0, 1\} \tag{13b}$$

$$\sum_{q=1}^{\mathcal{Q}^{C_k}} \sum_{n=1}^{\mathcal{N}^{C_k}} x_{qn} = 1, \quad x_{dn} \in \{0, 1\} \tag{13c}$$

$$\sum_{n=1}^{\mathcal{N}^{C_k}} p_n \leq P^{C_k}, \quad p_n \geq 0 \tag{13d}$$

$$\sum_{n=1}^{\mathcal{N}^{C_k}} R_{dn}(q, p_n) y_{dn} > R_d^{(B)}. \tag{13e}$$

Constraints (13b) and (13c), respectively, ensure that any user is associated to only one drone and is allocated only one channel. Moreover, (13d) guarantees positive allocated power for each user while the total allocated power does not exceed the available power. Constraint (13e) ensures that, for every drone d, the aggregate transmission rate of its associated users after cooperation, i.e. $\sum_{n=1}^{\mathcal{N}^{C_k}} R_{dn} y_{dn}$, is larger than the transmission rate provided by drone d before cooperation (baseline configuration), denoted by $R_d^{(B)}$.

However, solving the formulated optimization problem is not feasible due to the following reasons: (i) Given limited information and uncertainty in the network parameters, the objective function might not be known; (ii) Mixed integer non-linear program formulation and a combinatorial number of possible partitions (given by the Bell number) make the problem practically infeasible to solve. Therefore, we divide the optimization problem into three sub-problems: (i) clustering; 2) user-association and channel assignment; (iii) power control. Indeed, the drones first make coalition based

on their limited information, and inside each coalition, they optimize the user association as well as the channel- and power allocation.

B. Clustering/Coalition Formation

Centralized partitioning requires full information and it is computationally expensive. Hence we allow drones to form coalitions by local decision-making. When forming coalitions, each drone acts rationally to maximize its performance. Therefore, cooperation under uncertainty becomes challenging, as the optimal coalition and action are not known in advance. To address this challenge, we take advantage of Bayesian coalition formation games. Then w is a *coalition structure* and each group $C_k \in w$ in the partition is a *coalition*. In every coalition, the utility of each drone and the coalition depends on user- and channel assignment, as well as power allocation. *Coalition formation will be discussed in detail in Section V and Section VI.*
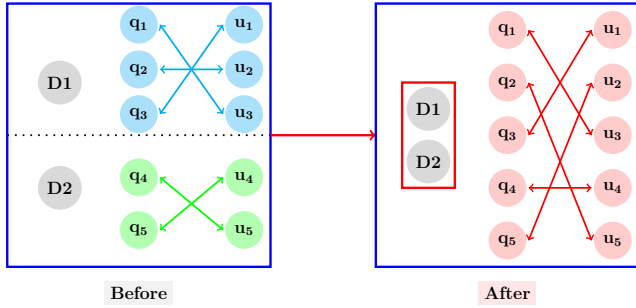

Figure 2. Example of the channel assignment. The left figure shows the baseline configuration when drones $D1$ and $D2$ perform independently. By forming the coalition $\{D1, D2\}$, the drones share their spectrum resources, thereby enabling more channel assignment choices.

C. User and Channel Assignment

For any coalition C_k, we formulate the user- and channel assignment problem as a weighted bipartite matching problem. We then use Hungarian algorithm [41] to solve the problem.

Definition 1 (Weighted Bipartite Matching): Let $G = (\mathcal{V}, \mathcal{E})$ be a weighted bipartite graph, where $\mathcal{V}$ consists of two set of vertices $\mathcal{V}_1$ and $\mathcal{V}_2$, such that $\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2$ and $\mathcal{V}_1 \cap \mathcal{V}_2 = \varnothing$. Moreover, $\mathcal{E} \subseteq \mathcal{V}_1 \times \mathcal{V}_2$ is the set of edges connecting the vertices. Let $e(i, j)$ denotes the edge between a vertex $i \in \mathcal{V}_1$ and $j \in \mathcal{V}_2$, and w_{ij} represents the weight of the edge $e(i, j)$. The weights are represented by the graph matrix $\mathbf{W} = [w_{ij}]$ of dimension $V_1 \times V_2$.

Matching: A matching is the subset of edges $\mathcal{M} \subseteq \mathcal{E}$ such that for $e, e' \in M$, there is no vertex v on which both the edges e and e' incident.

Maximum (minimum) matching: Let $W_{\mathcal{M}}$ represent the total weight of the selected edges in the matching $\mathcal{M}$. A matching $\mathcal{M}$ is maximum (minimum) if for any other matching $\mathcal{M}'$, we have $W_{\mathcal{M}} \geq W_{\mathcal{M}'}$ ($W_{\mathcal{M}} \leq W_{\mathcal{M}'}$).

Concerning the user- and channel assignment problem in any coalition C_k, $\mathcal{V}_1$ and $\mathcal{V}_2$ represent the set of channels $\mathcal{Q}^{C_k}$ and

the set of users $\mathcal{N}^{C_k}$, respectively. The weight of the an edge between a channel $q \in \mathcal{Q}^{C_k}$ and a user $n \in \mathcal{N}^{C_k}$, denoted by w_{qn}, is equal to the inverse average path-loss $1/\bar{L}_{dn,q}$. We then use the maximum bipartite matching for the assignment, i.e. we maximize the aggregate channel gains while associating users with drones/channels. For a given power allocation, such an assignment is equivalent to maximizing the overall transmission rate, as logarithmic function is monotone increasing. Note that the users are not explicitly assigned to the drones, rather implicitly by including fading- and shadowing effects in the overall channel gains of the users. **Fig. 2** shows an example of user- and channel allocation.

D. Power Control

After channel assignment, the drones perform power allocation, e.g. using water-filling algorithm [42], given constraint (13d). By the consensus among the drones in every coalition, any other power allocation strategy can be used without affecting the procedure of coalition formation.

V. BAYESIAN COALITION FORMATION GAME

A. Formulation of Bayesian Coalition Formation Game

We formulate the problem of distributed cooperation among drones under uncertainty as a Bayesian coalition formation game (BCFG) [43] with nontransferable utility (NTU). The game is the tuple $G : \langle \mathcal{D}, \mathcal{T}, B, \mathcal{A}, \vec{\mathbf{q}}_{d \in \mathcal{D}} \rangle$, where

- $\mathcal{D} = \{1, \ldots, D\}$ is the set of drones.
- $\mathcal{T}$ is the set of *types*. The available power at each drone follows a distribution belonging to the *type set* $\mathcal{T} = \{T_1, T_2, \ldots, T_M\}$, where M are the number of *types*. The space of drones' joint type is then $\mathbf{T} \in \otimes_{d=1}^{D}\{T_1, T_2, \ldots, T_M\}$. Every drone $d \in \mathcal{D}$ knows its own type $t^d \in \mathcal{T}$, i.e. the distribution of its available power, but not those others. For each drone d, the type set of other drones yields $\mathbf{T}^{-d} = \otimes_{x \in \mathcal{D}\backslash d}\{T_1, T_2, \ldots, T_M\}$. Note that the *amount* of available power of any drone d is the expected value of its type, i.e. $P_d = \mathbb{E}(t^d)$. The *type* of each drones remains fixed during the coalition formation process.
- B is the belief function. Let $\mathbf{t}^{-d} \in \mathbf{T}^{-d}$ be any vector from the set $\mathbf{T}^{-d}$ where, $\mathbf{t}^{-d} = [t_1^d, \ldots, t_j^d \ldots, t_{\mathcal{D}\backslash d}^d]$ and $t_j^d \in \mathcal{T}$ represents the type of the drone j according to the drone d. Therefore, $B(\mathbf{t}^{-d})$ is the joint belief of drone d about others having type profile $\mathbf{t}^{-d}$ as follows:

$$B(\mathbf{t}^{-d}) = \prod_{j \in \mathcal{D}\backslash d} \Pr[t_j^d], \qquad (14)$$

where $\Pr[t_j^d]$ is the probability of the drone d about drone j's type. Similarly, the function $B(\mathbf{t}_{C_k})$ indicates the marginal of B over any coalition C_k with members' types $\mathbf{t}_{C_k} = \{t^d\}_{d \in C_k} \in \mathbf{T}_{C_k} = \otimes_{d \in C_k}\mathcal{T}$. Moreover, $B(\mathbf{t}_{C_k}^{-d})$ indicates the joint belief of drone d about the coalition members $j \in \{C_k \backslash d\}$, where $\mathbf{t}_{C_k}^{-d} \in \mathbf{T}_{C_k}^{-d} = \otimes_{j \in \{C_k \backslash d\}}\mathcal{T}$. Further, $B^d(t^d)$ represents the drone d's belief about its own *type*, which assign probability 1 to its actual type and 0 to all others.

- $\mathcal{A}$ indicates the coalition actions. Indeed, each coalition C_k has a set of coalition actions, denoted by $\mathcal{A}^{C_k}$. The coalition actions $\mathcal{A}^{C_k}$ is the set of all solutions for joint user association, channel assignments, and power allocations in coalition C_k.

- Based on the drone d's belief, the available power of coalition C_k is $P(\mathbf{t}_{C_k}^{-d}) = \sum_{j=1}^{\mathcal{D}^{C_k}} P(t_j^d)$, where $P(t_j^d)$ is the power of the drone j based on d's belief t_j^d. The power is divided among users $\mathcal{N}^{C_k}$. Based on the *a priori* available channel information and the belief, we use weighted bipartite matching and the water-filling algorithm for channel selection and power allocation, respectively, as discussed in Section (IV-C) and Section (IV-D). Let $R_{dn}(\mathbf{t}_{C_k}^{-d}, a^{C_k})$ be the rate of the user n connected to the drone d in coalition C_k, calculated by drone d based on the type $\mathbf{t}_{C_k}^{-d}$ and action $a^{C_k} \in \mathcal{A}^{C_k}$. Let the selected channel and allocated power for user n be q and p_n, respectively, based on a^{C_k} and $\mathbf{t}_{C_k}^{-d}$. Thus, $R_{dn}(\mathbf{t}_{C_k}^{-d}, a^{C_k}) = R_{dn}(q, p_n)$. The total transmission rate provided by done d then yields $R_d(\mathbf{t}_{C_k}^{-d}, a^{C_k}) = \sum_{n=1}^{\mathcal{N}^{C_k}} R_{dn}(\mathbf{t}_{C_k}^{-d}, a^{C_k}) y_{dn}$. The expected payoff of drone d is then given by

$$\bar{q}_d^d(B, a^{C_k}) = \sum_{\mathbf{t}_{C_k}^{-d} \in \mathbf{T}_{C_k}^{-d}} B(\mathbf{t}_{C_k}^{-d}) R_d(\mathbf{t}_{C_k}^{-d}, a^{C_k}). \quad (15)$$

Finally, the maximum achievable expected payoff of a drone d in coalition C_k is

$$\bar{q}_d^d(B) = \max_{a^{C_k} \in \mathcal{A}^{C_k}} \bar{q}_d^d(B, a^{C_k}). \quad (16)$$

Each drone aims at maximizing its own expected payoff by joining the best coalition.

B. Belief Update Mechanism

At every iteration of coalition formation, the members of each coalition share the instantaneous information of the available power (type). For updating its belief, each drone performs the following: (i) Each drone uses the maximum likelihood estimation (MLE) method to estimate the parameters of the distribution of the type; (ii) It uses Kullback-Leibler divergence (KL divergence) method to compare the closeness of the estimated parameters with the given set of types; (iii) Finally, each drone updates its belief about the types of other drones based on averaging. Details follow.

1) The maximum likelihood estimation (MLE): In MLE, the objective function is the likelihood of the data $\mathbf{X}$ given the model. The goal is to find the parameter θ that maximizes the evaluation function (the likelihood). Formally,

$$\theta_{MLE} = \underset{\theta}{\operatorname{argmax}} \ p(\mathbf{X}|\theta). \quad (17)$$

Let $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n]$ be the dataset consisting of n samples of Gaussian process.[1] The *log likelihood* $\mathcal{LL}$ of the data is

given by

$$\mathcal{LL} = \log(\mathcal{N}(\mathbf{X}|\mu, \sigma^2)) = \sum_{n=1}^{N} \log(\mathcal{N}(\mathbf{x}_n|\mu, \sigma^2)$$

$$= \sum_{n=1}^{N} \log\left(\frac{1}{\sqrt{2\pi\sigma^2}} \exp^{\frac{1}{2}\left(\frac{(\mathbf{x}_n - \mu)^2}{\sigma^2}\right)}\right). \quad (18a)$$

The best μ and σ^2 are obtained by taking partial derivatives of above *log likelihood* function $\mathcal{LL}$ w.r.t. these parameters and equating them to zero. After solving the two equations, the estimated parameters are

$$\mu_{MLE} = \frac{1}{N}\sum_{n=1}^{N}\mathbf{x}_n; \quad \sigma_{MLE}^2 = \frac{1}{N}\sum_{n=1}^{N}(\mathbf{x}_n - \mu)^2. \quad (19)$$

2) Kullback-Leibler divergence (KL divergence): The KL divergence (also called relative entropy) is a measure of how one probability distribution is different from other. KL divergence of 0 indicates that the two distributions are identical and 1 indicates that they are completely different. Suppose p and q are the density of the normal random variables with mean $[\mu_1, \mu_2]$ and variance $[\sigma_1^2, \sigma_2^2]$. The KL distance from q to p is given by

$$KL_{pq} = \log\left(\frac{\sigma_2}{\sigma_1}\right) + \frac{\sigma_1^2 + (\mu_1 - \mu_2)^2}{2\sigma_2^2} - \frac{1}{2}. \quad (20)$$

3) Averaging: The drones repeat the coalition formation process at $t = 1, 2, \ldots$ until convergence. At every round, given the shared information, each drone re-estimates the parameters of the types' distributions (beliefs) using (18). Afterwards, it identifies the types using (20). It then calculates the average frequency of observation for each type, i.e. the number of times each type is observed so far over the total number of interactions.

VI. DISTRIBUTED ALGORITHM FOR BAYESIAN COALITION FORMATION GAME

In this section, we propose a distributed method for coalition formation under uncertainty that consists of main parts: (i) Initialization of beliefs over the types, (ii) Coalition formation process based on the beliefs; and (iii) Updating the belief based on the local observations. **Algorithm 1** and **Algorithm 2**, respectively, summarize the overall procedure and the coalition formation process.

A. Algorithm 1: Repeated coalition formation under uncertainty

For the few first iterations, the drones form the grand coalition, where each drone shares the information about its instantaneous value of the type with other drones. The drones utilize this information to initialize the beliefs about each others' type by the belief updating mechanism discussed in section V-B. If forming a grand coalition is not possible, the drones simply start from some arbitrary belief.

After initialization, at each round t, the controller draws a realization of some Bernoulli random variable X with parameter ϵ. If the outcome is 1, the agents form the grand coalition;

---

[1]Here, we assume that the available power of each drone follows a normal distribution with unknown parameters. Adapting the entire analysis to any other distribution is straightforward.

otherwise, they follow the distributed coalition formation process based on the *best reply dynamics* discussed in **Algorithm 2**. The result is some coalition structure $w \in \mathcal{W}$. Inside each coalition, the drones share the instantaneous information. Given this information, each member of coalition C_k updates the belief about its partners' types. Note that, forming the grand coalition with prob ϵ allows for enough information exchange so that the drones do not get stuck in forming some particular coalitions based on their beliefs. The process is repeated until there is no further change in the belief and a stable coalition structure is formed.

Algorithm 1 Repeated coalition formation under uncertainty

1: ***Initialization:*** Form the grand coalition for a few iterations. Initialize the belief B given the observations and based on the mechanism V-B.

2: **loop**

3:     A central controller draws a realization x of a Bernoulli random variable X with probability ϵ.

4:     **if** $X = 1$ **then**

5:       The drones form the grand coalition.

6:     **else**

7:       Drones $\mathcal{D}$ engage in the coalition formation process using **Algorithm 2**.

8:       **Algorithm 2** is run until Nash-stable/absorbing state w^* is reached (i.e., when no drone benefits by leaving its coalition and join another coalition in the same coalition structure.)

9:     **end if**

10:    Each drone broadcasts the instantaneous power information to the members of its coalition.

11:    Based on this local information, each drone updates the belief about its coalition members using the belief update mechanism in Section V-B .

12: **end loop**

B. Algorithm 2: Distributed Coalition Formation Algorithm

The dynamic coalition formation algorithm [44], [45] operates iteratively. First, the algorithm is initialized with any random feasible coalition structure. At every stage, a controller selects a drone d with probability $1/D$, which is called the *proposer*, to change its strategy.

The proposer d observes the current coalition structure and decides for one of the following: (i) staying in the current coalition, (ii) joining another coalition in the same coalition structure, or (iii) forming the singleton coalition. Since the drones are rational, the proposer d selects the coalition S that results in the maximum expected payoff while believing that the affected members agree. Ties are broken simply at random. Thus, based on its belief, the proposer calculates its expected payoff in every coalition. It switches its coalition provided that its payoff is strictly higher and the expected payoff of other drones in the coalition S where he joins, are at-least equal to their payoff before joining the proposer, i.e. they would agree that the proposer joins. The proposer then sends the request to the controller for joining its preferred coalition

S. Upon receiving the proposal to join, all the members of coalition S calculate their expected payoffs based on their beliefs conditioned that the proposer joins their coalition. They then compare the payoffs for before and after joining. Every drone in S agrees that d joins only if it does not result in a reduction in its expected payoff. Every drone in S sends its positive or negative response to the controller, and the proposal is accepted only if all agree. If not allowed to join, the proposer repeats the same procedure with the next best coalition. This procedure is the *best reply dynamics*.

The process continues until the drones achieve a stable coalition structure, also called the *absorbing state*/Nash-stable coalition structure. In an absorbing state, no drone has an incentive to change its strategy, given the prevailing coalition structure. There can be multiple stable coalition structures (i.e. multiple absorbing states). The achieved absorbing state depends on the random sequence of the proposers. The best reply dynamics converges to one of the absorbing states with probability one as time tends to infinity [44].

During the coalition formation process, the drones of each coalition share the information in a broadcast control channel. Therefore, no pairwise interaction is required. As such, the signaling overhead remains low. **Fig. 3** shows the overall solution approach.

Algorithm 2 Distributed coalition formation algorithm

1: Initialize $w(t)$.

2: **loop**

3:     At time t, the controller selects a *proposer* d uniformly at random. Let d belong to coalition C_k of w. Upon being selected, it can decide to join some other coalition $S \in w(t) \backslash C_k$, to remain in C_k, or to go singleton.

4:     Based on its belief B, d computes its expected payoff if joining any coalition S, $\bar{q}_d^d(S \cup \{d\})$, using (15).

5:     Every drone $k \in S$ computes its expected payoff based on its belief for the case d joins S as $\bar{q}_k^k(S \cup \{d\})$.

6:     Based on the computed expected payoffs, evaluate $C_k' = \arg\max_{S \in w(t) \backslash C_k} \bar{q}_d^d(S \cup \{d\})$, such that it holds $\bar{q}_d^d(S \cup \{d\}) > \bar{q}_d^d(C_k)$ and $\bar{q}_k^k(S \cup \{d\}) \geq \bar{q}_k^k(S)$.

7:     **if** C_k' is not empty **then**

8:       drone d leaves coalition C_k and joins C_k'. Then the new state is

9:       $w(t+1) = ((w(t) \backslash \{C_k\}) \backslash \{C_k'\}) \cup \{C_k' \cup \{d\}\} \cup \{C_k \backslash \{d\}\})$

10:    **else**

11:       $w(t+1) = w(t)$

12:    **end if**

13:    $t = t + 1$

14: **end loop**   when the stable/absorbing state w^* is reached.

VII. ANALYSIS OF THE BAYESIAN COALITION FORMATION GAME

To analyze the stability of the formulated Bayesian coalition formation game, we use the concept of Nash stability/ absorbing states. We also develop a discrete-time Markov chain for the coalition formation algorithm that provides the Nash-stable solution.
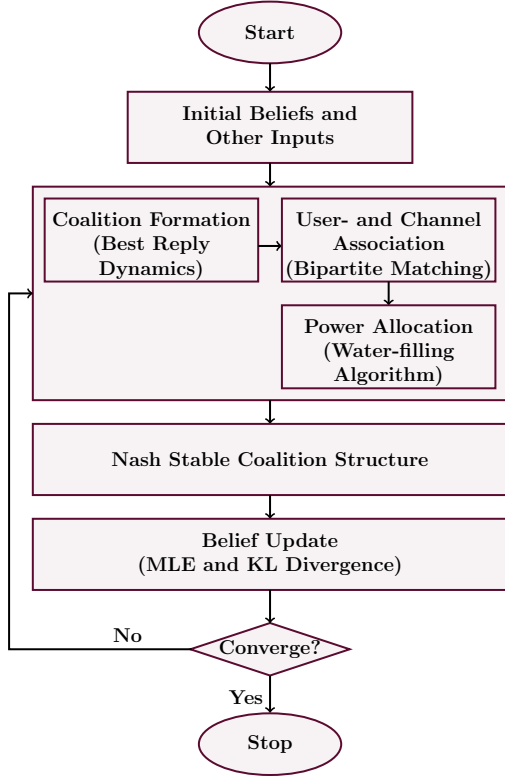
Figure 3. Distributed coalition formation under uncertainty and resource allocation.

A. Convergence Analysis

We start by stating the following definitions.

Definition 2: Let $\mathcal{W} = \{w_1, \ldots, w_m, \ldots, w_W\}$ be the set of states. A subset $F \subset \mathcal{W}$ is ergodic if, for $w \in F$ and $w' \notin F$, the transition probability $\rho_{\omega, \omega'} = 0$ and no any other non-empty subset has this property. The singleton ergodic states are called absorbing states. A state w is absorbing if $\rho_{w,w} = 1$. Once the coalition formation process reaches the ergodic state, it remains there forever.

Definition 3 (Nash-stability/Absorbing state [44] [45]): Let drone d belong to coalition $C_k(d)$ in state w. The state or a coalition structure $w = \{C_1, \ldots, C_k \ldots, C_l\}$ is an absorbing (Nash-stable) state if for all $d \in \mathcal{D}$, for any proposal $S \cup \{d\}$, either

- $\bar{q}_d^d(S \cup \{d\}) \leq \bar{q}_d^d(C_k)$, or
- $\exists j \in S$ such that $\bar{q}_j^j(S \cup \{d\}) < \bar{q}_j^j(S)$.

The condition guarantees that *either*

- The *proposer* d cannot improve its payoff by leaving its current coalition C_k to join coalition S. For the tie $\bar{q}_d^d(S \cup \{d\}) = \bar{q}_d^d(C_k)$, the drone remains in C_k as it was already in effect and accepted by all the members, *or*
- The proposal is denied by at-least one member of the coalition S, i.e. there is some agent $j \in S$ for which $q_j^j(S \cup \{d\}) < q_j^j(S)$. In this case, although the proposer has the incentive to join the coalition S, he cannot due to the denial.

Thus, no drone $d \in \mathcal{D}$ leaves its current coalition to join another coalition in the same coalition structure. Also, no drone has an incentive to leave it current coalition to act alone.

The following proposition describes the convergence behavior of **Algorithm 2** [44] [45].

Proposition 1: **Algorithm 2** converges to a Nash-stable solution or absorbing state [44].

Definition 4 (Preference): Let a drone d be the member of coalition C_k in a coalition structure w. Then $S \succ_d C_k$ implies the strong preference of d for coalition $S \in \{w \backslash C_k\}$ over coalition C_k. The preference $S \succ_d C_k$ is valid if the following two conditions are satisfied:

- The expected payoff ($\bar{q}_d^d$) of drone d is strictly higher when it joins coalition S than when it remains a part of coalition C_k, i.e. $\bar{q}_d^d(S \cup \{d\}) > \bar{q}_d^d(C_k)$.
- No member of coalition S believes that joining d reduces its payoff, i.e. $\bar{q}_j^j(S \cup \{d\}) \geq \bar{q}_j^j(S), \forall j \in S$.

Similarly, $S \succeq_d C_k$ represents the weak preference of d for S over C_k. The weak preference means that the expected payoff of drone d in coalition S is at least equal to the expected payoff in its current coalition, i.e. $\bar{q}_d^d(S \cup \{d\}) \geq \bar{q}_d^d(C_k)$. However, the conditions on the affected members remain same; that is, $\bar{q}_j^j(S \cup \{d\}) \geq \bar{q}_j^j(S), \forall j \in S$.

Definition 5 (Characteristic function): Let $\vec{\mathbf{q}}^{\mathcal{D}} = [\bar{q}_1(\mathcal{D}), \ldots, \bar{q}_d(\mathcal{D}), \ldots \bar{q}_D(\mathcal{D})]$ be the expected payoff vector of all drones in grand coalition. Moreover, $\vec{\mathbf{q}}^S = [\ldots, \bar{q}_d(\mathcal{D}), \ldots]$ denotes the expected payoff vector of the drones in any coalition $S \subset \mathcal{D}$. For coalition S, the characteristic function $V(\mathcal{D})$ is a set of feasible payoff vectors $\vec{\mathbf{x}}^S$ of length $|S|$, where

$$V(S) = \{\vec{\mathbf{x}}^S \in \mathbb{R}^S | \vec{\mathbf{x}}^S \leq \vec{\mathbf{q}}^S\}. \tag{21}$$

Definition 6 (Weak Bayesian core [31]): The weak Bayesian core of an NTU game is defined as

$$\mathcal{C} = \{\vec{\mathbf{q}}^{\mathcal{D}} \in V(\mathcal{D}) | \forall S \subseteq \mathcal{D}, \nexists \vec{\mathbf{q}}^S \in V(S) \text{ s.t.}$$
$$S \succeq_d \mathcal{D}, \forall d \in S\}. \tag{22}$$

In words, in weak Bayesian core, the payoff of every drone in the grand coalition is such that no drone leaves the grand coalition as it is not better off in any other coalition, i.e. there is no blocking coalition.

Definition 7 (Strong Bayesian core [31]): The strong Bayesian core of an NTU game is defined as

$$\mathcal{C} = \{\vec{\mathbf{q}}^{\mathcal{D}} \in V(\mathcal{D}) | \forall S \subseteq \mathcal{D}, \nexists \vec{\mathbf{q}}^S \in V(S) \text{ s.t.}$$
$$S \succeq_d \mathcal{D}, \forall d \in S \text{ and } S \succeq_j^d \mathcal{D}, \forall j \in S, j \neq d\}. \tag{23}$$

The definition indicates that there exists a payoff profile in the grand coalition such that no drone has an incentive to leave the grand coalition. Also, every drone, based on its information about the expected payoff of others, believes that other drones are not better off if it leaves the grand coalition. The strong Bayesian core is the subset of the weak Bayesian core. Moreover, the conventional core is a special case of Bayesian core when there is no uncertainty, i.e. the drones know each others' types.

Remark 1: The distributed coalition formation algorithm converges to a Nash stable solution but not necessarily to the Bayesian core. For the proposed drone network, there may exist a blocking coalition for the grand coalition due

to the individual rationality of the players/drones. However, Bayesian core exists if the grand coalition is the only stable coalition structure. If there are multiple Nash stable coalition structures and one of them is the grand coalition then the convergence to a particular Nash stable solution depends on the random sequence of proposers. The best case is to initialize the coalition structure with the grand coalition. If it is not stable, then the drones leave the grand coalition and form a stable coalition as a result of successive interactions.

B. Discrete-Time Markov Chain-Based Analysis of Bayesian Coalition Formation Game

We formulate a finite state Markov chain to find the solution of coalition formation and to analyze the stable coalition structures [44] [45]. The state space of the Markov chain is the set of all coalition structures, i.e. $\mathcal{W}$. The transition matrix of the Markov chain is then $\mathbf{W}$. Let drone d belong to coalition $C_k(d)$ in the state w at some time t. Moreover, $C'_k(d) = S \cup \{d\}$ is the coalition in state w' at the next iteration $t + 1$. Thus, the state $w' = ((w \backslash \{C_k\} \backslash \{S\}) \cup (\{C_k \backslash \{d\}\} \cup \{S \cup \{d\}\}))$. Transition from state w to w' depends on the decision of the proposer drone d at time t. The transition probability $\rho_{w,w'}$, from state w to w', is given by

$$\rho_{w,w'} = \frac{1}{D} \sum_{d \in \mathcal{D}} \varphi_d(w'|w), \qquad (24)$$

where $1/D$ is the probability that drone d becomes a proposer to change its strategy. The term φ_d is defined by best reply dynamics as follows:

- $\varphi_d = 1/k_{w'}^d$, where $k_{w'}^d$ are the potential maximizers of the expected payoff of proposer d at state w. The proposer joins one of these maximizers at random. The proposal to join the coalition S by the proposer is feasible only if every members of the coalition S accepts the proposal based on its private beliefs. Also, in the case $w \neq w'$, the $\varphi_d(w'|w)$ is non-zero for at most one d in $\sum_{d \in \mathcal{D}} \varphi_d(w'|w)$.
- $\varphi_d = 1$, if $C'_k(d) = S \cup \{d\}$ is the maximizer of the proposer d but the proposal is rejected by at least one member of the affected coalition S so that $w = w'$.
- $\varphi_d = 0$, otherwise.

The stable coalitions are the absorbing states of the finite Markov chain. The stationary probability vector is $\vec{\pi} = [\pi_1 \ldots \pi_w \ldots \pi_W]^T$ where π_w is the probability that the drones form the coalition structure w. Given that the transition matrix $\mathbf{W}$, the probability vector is the solution of

$$\vec{\pi}^T \mathbf{W} = \vec{\pi}^T, \text{ where } \vec{\pi}^T \vec{1} = 1. \qquad (25)$$

VIII. Numerical Results

We consider an outdoor urban environment. The users that are initially assigned to each drone are randomly positioned inside the service area of 4km × 4km. Each drone is placed at the centroid of assigned users at a height of 1 km above the ground. This configuration is the *baseline* configuration. For a comprehensive evaluation, we implement four different

Table II
THE PARAMETERS OF SIMULATED NETWORKS (EACH COLUMN REPRESENTS A SIMULATION SETTING)

Simulation Settings / Parameters	S1	S2	S3	S4
Number of Drones (D)	3	4	5	6
Number of Channels (Q)	9	12	15	18
Number of Users (N)	9	12	15	18
Users per Drone (N_d)	3	3	3	3
Channels per Drone (Q_d)	3	3	3	3

Table III
ENVIRONMENT PARAMETERS [46]

Environment	Parameters			
	α, γ	k_1, k_2	g_1, g_2	$\mu_{\text{LoS}}, \mu_{\text{NLoS}}$
Urban	0.6, 0.11	10.39, 0.05	29.06, 0.03	1, 20
Dense urban	0.36, 0.21	8.96, 0.04	35.97, 0.04	1.6, 23
High-rise urban	0.05, 0.61	7.37, 0.03	37.08, 0.03	2.3, 34

settings concerning several network parameters. **Table II** summarizes the scenarios.

The drone-based communication operates over 2 GHz carrier frequency (f_c = 2 GHz). The sum of noise and interference power spectral density $N_0 + I_0$ is -70 dBm/Hz, antenna gain G is 10 dB and bandwidth B_w is assumed to be 1 Hz. Without loss of generality, the bandwidth of each channel is 1 Hz. The minimum angle θ_o is 15°. The small scale fading parameters are $k_0, k_{\frac{\pi}{2}} = [3, 30]$ dB [39]. **Table III** summarizes the parameters for different simulated environments. For each network type, we generate 100 topologies. We then simulate the coalition formation procedure for each topology 30 times, as each topology can have multiple absorbing states, i.e. stable coalition structures. This means that each repetition of the coalition formation process at the same topology might converge into any of the possible absorbing states depending on the initial coalition structure and selection of the drones in the coalition formation process. It is also worth noting that for all topologies and simulation runs, the coalition formation process converges to a stable state in at most 50 iterations. Convergence in such a short time is in particular interesting since, in theory, the convergence follows asymptotically.

The proposed solution is compared with

1) Baseline configuration, in which the drones do not cooperate. For each drone and its assigned users, the channel assignment and power allocation follow by bipartite matching and water-fill algorithm, respectively.
2) The solution using distributed best reply dynamics with complete information. That is, the drones have the full knowledge of channel quality as well as available power profiles.
3) The socially optimal solution obtained using the exhaustive search, where the central controller has global information. The drones within each coalition share their channel and power. The coalition structure that has a maximum overall rate is selected as the optimal solution. Note that in this solution, achieving the highest aggregate performance has a higher priority compared to optimizing the performance for each drone.
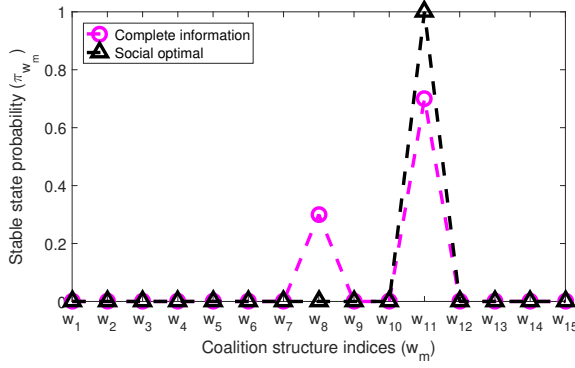
Figure 4. Stationary probability of a network consisting of 4 drones and 3 users connected to each drone (Type 1: $\mu = 12$, $\sigma = 3$; Type 2: $\mu = 18$, $\sigma = 3$).
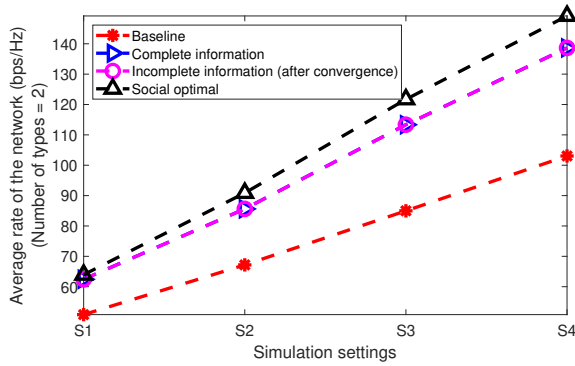

Figure 7. Convergence for each simulation setting shown in Table II for two types ($T_1 : \mu = 12, \sigma = 3$; $T_2 : \mu = 18, \sigma = 3$).

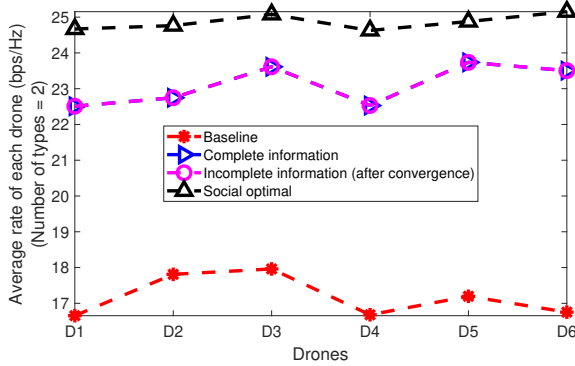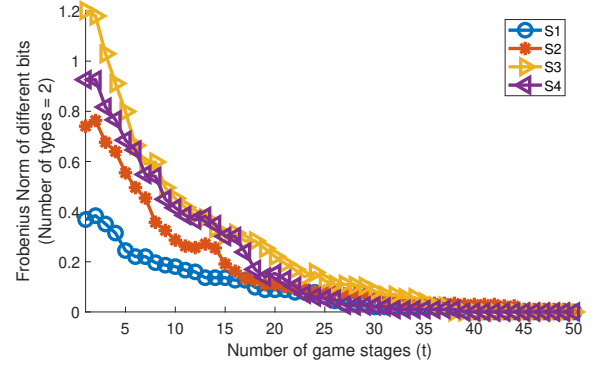Figure 5. Average transmission rate of the network for each simulation setting shown in Table II for two types ($T_1 : \mu = 12, \sigma = 3$; $T_2 : \mu = 18, \sigma = 3$).

Figure 6. Average transmission rate of each drone of the simulation setting with S4 (consists of 6 drones) for two types as shown in Table II ($T_1 : \mu = 12, \sigma = 3$; $T_2 : \mu = 18, \sigma = 3$).

For the distributed coalition formation process, for every topology, there are two ways to define the network's overall transmission rate:

- The maximum expected transmission rate for the network achieved by any stable coalition (the approach adopted in this work);
- The expected transmission rate of all the stable coalition structures. Formally,

$$\bar{q}_{total} = \sum_{d \in \mathcal{D}} \sum_{w=1}^{W} \pi_w \bar{q}_d^d(C_k^d), \quad \text{for} \quad C_k^d \in w. \quad (26)$$

We present the stationary probabilities obtained from the Markov model in **Fig. 4**. We show the stationary probabilities of an arbitrary topology from 100 random topologies of simulation setting $S4$ given in **Table II**. It consists of 4 drones with 2 types. Each drone has 3 channels and 3 users. The network of 4 drones has a total of 15 possible coalition structures $\{w_1, \ldots, w_{15}\}$, where $w_8 = \{\{d_1\}, \{d_2, d_3, d_4\}\}$ and $w_{11} = \{\{d_1, d_2, d_3\}, \{d_4\}\}$. For this specific topology, the best coalition structure is w_{11} that provides the maximum sum rate for all drones. Thus, if drones seek to maximize the social welfare, the probability of forming w_{11} is 1. However, the best-reply dynamics with complete information yields the coalition structures w_{11} and w_8 as absorbing states. Thus, w_{11} and w_8 are formed 70% and 30% of the time, respectively, if the drones seek individual rationality ($\pi_{w_{11}} = 0.7$, $\pi_{w_8} = 0.3$).

Fig. 5 shows the average total transmission rate of the network for four simulation settings given in **Table II**. We consider two *types* that are Gaussian random variables defined as $T_1 : \mu = 12, \sigma = 3$; $T_2 : \mu = 18, \sigma = 3$. We observe that the average transmission rate increases by moving from simulation setting $S1$ to $S4$. This happens since increasing the number of drones, users, and channels, improves the chances of experiencing good channel quality.

Fig. 6 shows the performance improvement for individual drones for the simulation setting $S4$.

Fig. 7 shows the learning of drones type over time. We quantify the progress of learning using the *Frobenius norm*, as described in the following. Each drone updates its belief about the type of other drones after each iteration of coalition formation based on the local information. For a network of D drones and M types, every drone $d \in \mathcal{D}$ maintains a matrix $\mathbf{T}$ of dimension $D \times D$ for each type, thus there are M such matrices. After every iteration, any element a_{ij} of the matrix for a given type is 1 if the drone i predicts that type for drone j and 0 otherwise. We do averaging after every iteration to obtain the probability of predicting the type by every drone for others in the network. For the evaluation of convergence, we convert the probability values to binary values. We then use the Frobenius norm to show the average number of bits that differ from true type bits. The Frobenius norm of a $m \times n$
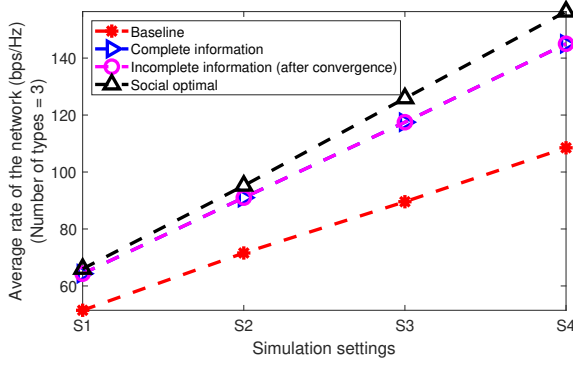
Figure 8. Average transmission rate of the network for each simulation setting shown in Table II for three types ($T_1 : \mu = 12, \sigma = 3$; $T_2 : \mu = 18, \sigma = 3$; $T_3 : \mu = 24, \sigma = 3$).
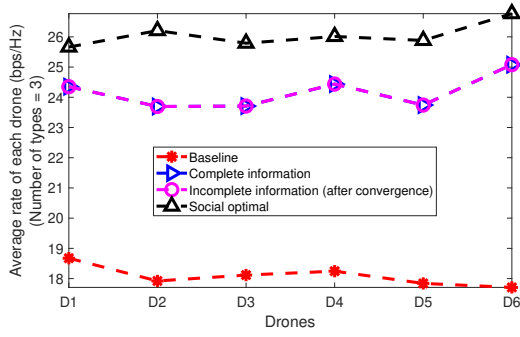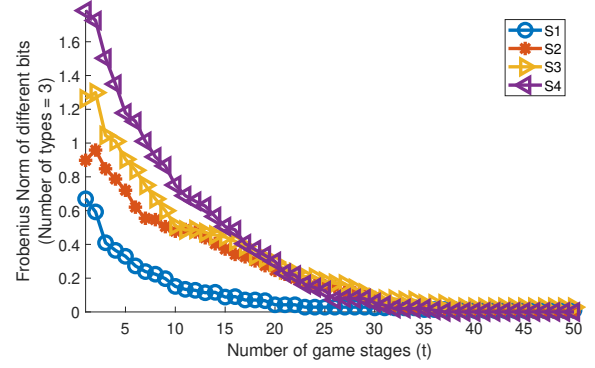

Figure 10. Convergence for each simulation setting shown in Table II for three types ($T_1 : \mu = 12, \sigma = 3$; $T_2 : \mu = 18, \sigma = 3$; $T_3 : \mu = 24, \sigma = 3$).

Figure 9. Average transmission rate of each drone of the simulation setting with S4 (consists of 6 drones) for three types as shown in Table II ($T_1 : \mu = 12, \sigma = 3$; $T_2 : \mu = 18, \sigma = 3$ $T_2 : \mu = 24, \sigma = 3$).
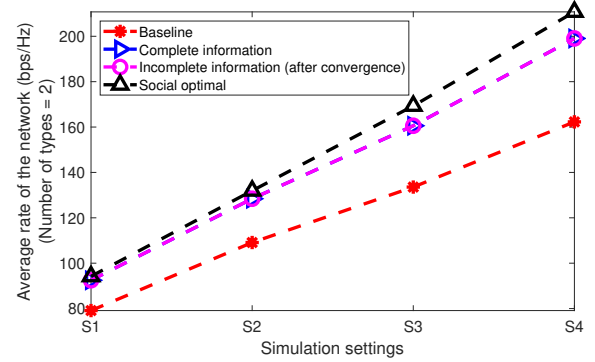

Figure 11. Average transmission rate of the network for each simulation setting shown in Table II, with the increase of the power level and two types ($T_1 : \mu = 120, \sigma = 5$; $T_2 : \mu = 180, \sigma = 5$).

matrix $\mathbf{A}$ is defined as

$$\left\| \mathbf{A}_F \right\| = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2}, \qquad (27)$$

where matrix $\mathbf{A}_F$ is difference of the matrix with true bits and the matrix obtained after belief update at every iteration. Frobenius norm 0 represents the convergence to true value.

Similarly, **Fig. 8**, **Fig. 9**, and **Fig. 10** show the performance for the same simulations setting given in Table II but for three Gaussian types as $T_1 : \mu = 12, \sigma = 3$; $T_2 : \mu = 18, \sigma = 3$; $T_3 : \mu = 24, \sigma = 3$. Here, the trend is similar to the scenario with two types.

In **Fig. 11**, we show the effect of increasing the power level. In this case, the Gaussian types are $T_1 : \mu = 120, \sigma = 5$; $T_2 : \mu = 180, \sigma = 5$. The figure shows that the network's transmission performance improves compared to the previous case of low power availability.

In **Fig. 12**, we show the comparison of the two different approaches of calculating the overall network transmission rate. We observe that by choosing the Nash-stable coalition structure having the maximum sum rate, the overall network's performance improves better compared to the approach of calculating the expected rate over all possible Nash-stable coalition structures, as given by (26).

Fig. 13 shows the network's performance for various environments. We observe that compared to the dense-urban

and urban environments, the performance of cooperation is better in high-rise urban environments. This is due to severe reflections, scattering, and shadowing in high-rise urban environment. In other words, leveraging cooperation among drones improves the network's performance substantially, as a high-quality LoS link towards a helping drone might be available to some users that have a low-quality channel towards the original serving drone.

So far, we have selected the types in a way that the corresponding probability distributions overlap only very slightly or not at all. However, it is worth noting that in case of highly overlapping type distributions, the accuracy and the convergence rate would degrade. **Fig. 14** illustrates the case where the types' distributions overlap approximately up to 60%.

As the final remark, we mention the following: The repeated coalition formation algorithms with incomplete and complete information yield greater average rates than the baseline configuration and lower than the social optimal case. In the social optimal configuration, the optimal coalition structure is the one that maximizes the overall rate of the network such that every drone provides a rate more than the baseline configuration. However, the socially optimal solution might not be a stable coalition structure and shall be enforced by an authority.
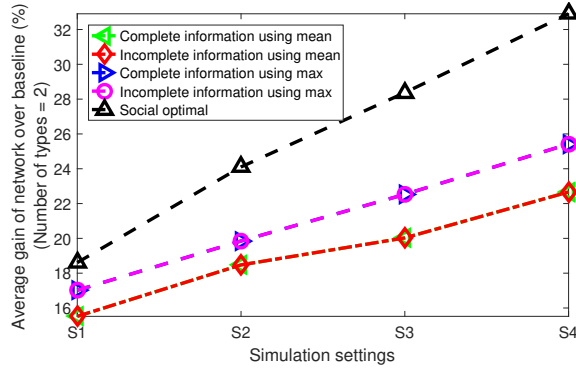
Figure 12. Average transmission gain for each simulation setting shown in Table II for two types ($T_1 : \mu = 120, \sigma = 5$; $T_2 : \mu = 180, \sigma = 5$).
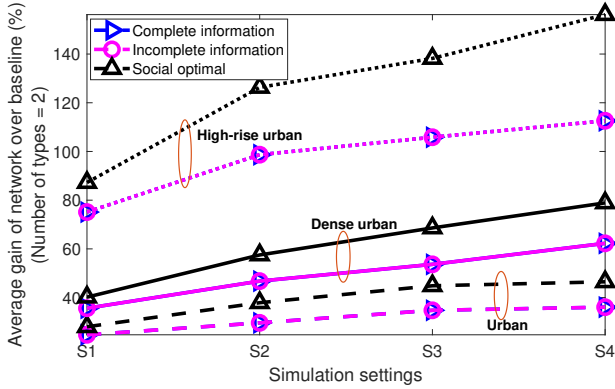

Figure 13. Average transmission gain for various environments for two types ($T_1 : \mu = 12, \sigma = 3$; $T_2 : \mu = 18, \sigma = 3$).

IX. Conclusion

For a drone-based wireless network, we have developed an approach for distributed cooperation among drones under uncertainty based on a Bayesian coalition formation process in order to maximize the overall transmission rate in the network. Given only limited information about the *type* (i.e. amount of available power) in other drones, coalitions of drones are formed, the spectrum and energy resources are pooled, and then the channel allocations to users are shuffled to improve network performance. We have evaluated the method through theoretical and numerical analysis. Theoretical analysis has shown the convergence to the stable coalition structure. Also, simulation results have showed performance improvement over baseline method. Future works include generalization to the scenario with state-dependent types with unknown states as well as the scenario with high-dimensional types.

References

[1] M. Giordani, M. Polese, M. Mezzavilla, S. Rangan, and M. Zorzi, "Toward 6g networks: Use cases and technologies," *IEEE Communications Magazine*, vol. 58, no. 3, pp. 55–61, 2020.

[2] S. Sekander, H. Tabassum, and E. Hossain, "Multi-tier drone architecture for 5g/b5g cellular networks: Challenges, trends, and prospects," *IEEE Communications Magazine*, vol. 56, no. 3, pp. 96–103, 2018.

[3] S. Maghsudi and E. Hossain, "Distributed user association in energy harvesting small cell networks: A probabilistic bandit model," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1549–1563, 2017.

[4] S. Sekander, H. Tabassum, and E. Hossain, "Statistical performance modeling of solar and wind-powered uav communications," *IEEE Transactions on Mobile Computing*, 2020.

[5] Y. Zeng, R. Zhang, and T. J. Lim, "Wireless communications with unmanned aerial vehicles: Opportunities and challenges," *IEEE Communications Magazine*, vol. 54, no. 5, pp. 36–42, 2016.

[6] T. X. Tran, A. Hajisami, and D. Pompili, "Cooperative hierarchical caching in 5G cloud radio access networks," *IEEE Network*, vol. 31, no. 4, pp. 35–41, 2017.

[7] E. T. Ceran, T. Erkilic, E. Uysal-Biyikoglu, T. Girici, and K. Leblebicioglu, "Optimal energy allocation policies for a high altitude flying wireless access point," *Transactions on Emerging Telecommunications Technologies*, vol. 28, no. 4, p. e3034, 2017.

[8] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Wireless communication using unmanned aerial vehicles (UAVs): Optimal transport theory for hover time optimization," *IEEE Transactions on Wireless Communications*, vol. 16, no. 12, pp. 8052–8066, 2017.

[9] M. Chen, W. Saad, and C. Yin, "Liquid state machine learning for resource allocation in a network of cache-enabled LTE-U UAVs," in *GLOBECOM 2017-2017 IEEE Global Communications Conference*. IEEE, 2017, pp. 1–6.

[10] V. Sharma, M. Bennis, and R. Kumar, "UAV-assisted heterogeneous networks for capacity enhancement," *IEEE Communications Letters*, vol. 20, no. 6, pp. 1207–1210, 2016.

[11] E. Kalantari, H. Yanikomeroglu, and A. Yongacoglu, "On the number and 3D placement of drone base stations in wireless cellular networks," in *2016 IEEE 84th Vehicular Technology Conference (VTC-Fall)*. IEEE, 2016, pp. 1–6.

[12] J. Cui, Y. Liu, and A. Nallanathan, "Multi-agent reinforcement learning-based resource allocation for UAV networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 2, pp. 729–743, 2019.

[13] F. Lagum, I. Bor-Yaliniz, and H. Yanikomeroglu, "Strategic densification with UAV-BSs in cellular networks," *IEEE Wireless Communications Letters*, vol. 7, no. 3, pp. 384–387, 2017.

[14] E. Kalantari, M. Z. Shakir, H. Yanikomeroglu, and A. Yongacoglu, "Backhaul-aware robust 3D drone placement in 5G+ wireless networks," in *2017 IEEE international conference on communications workshops (ICC workshops)*. IEEE, 2017, pp. 109–114.

[15] H. Ghazzai, M. B. Ghorbel, A. Kassler, and M. J. Hossain, "Trajectory optimization for cooperative dual-band uav swarms," in *2018 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2018, pp. 1–7.

[16] M. Samir, S. Sharafeddine, C. Assi, T. M. Nguyen, and A. Ghrayeb, "Trajectory planning and resource allocation of multiple uavs for data delivery in vehicular networks," *IEEE Networking Letters*, vol. 1, no. 3, pp. 107–110, 2019.

[17] A. M. Koushik, F. Hu, and S. Kumar, "Deep Q-learning-based node positioning for throughput-optimal communications in dynamic uav swarm network," *IEEE Transactions on Cognitive Communications and Networking*, vol. 5, no. 3, pp. 554–566, 2019.

[18] H.-T. Ye, X. Kang, J. Joung, and Y.-C. Liang, "Joint uplink and downlink 3d optimization of an uav swarm for wireless-powered nb-iot," in *2019 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2019, pp. 1–6.

[19] P. Dinh, T. M. Nguyen, S. Sharafeddine, and C. Assi, "Joint location and beamforming design for cooperative uavs with limited storage capacity," *IEEE Transactions on Communications*, vol. 67, no. 11, pp. 8112–8123, 2019.

[20] L. Zhang and N. Ansari, "A framework for 5g networks with in-band full-duplex enabled drone-mounted base-stations," *IEEE Wireless Communications*, vol. 26, no. 5, pp. 121–127, 2019.

[21] D. Fooladivanda and C. Rosenberg, "Joint resource allocation and user association for heterogeneous wireless cellular networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 1, pp. 248–257, 2012.

[22] F. Wang, W. Chen, H. Tang, and Q. Wu, "Joint optimization of user association, subchannel allocation, and power allocation in multi-cell multi-association OFDMA heterogeneous networks," *IEEE Transactions on Communications*, vol. 65, no. 6, pp. 2672–2684, 2017.

[23] N. Zhao, Y.-C. Liang, D. Niyato, Y. Pei, M. Wu, and Y. Jiang, "Deep reinforcement learning for user association and resource allocation in heterogeneous cellular networks," *IEEE Transactions on Wireless Communications*, vol. 18, no. 11, pp. 5141–5152, 2019.

[24] W. Saad, Z. Han, M. Debbah, and A. Hjorungnes, "A distributed coalition formation framework for fair user cooperation in wireless networks," *IEEE Transactions on wireless communications*, vol. 8, no. 9, pp. 4580–4593, 2009.
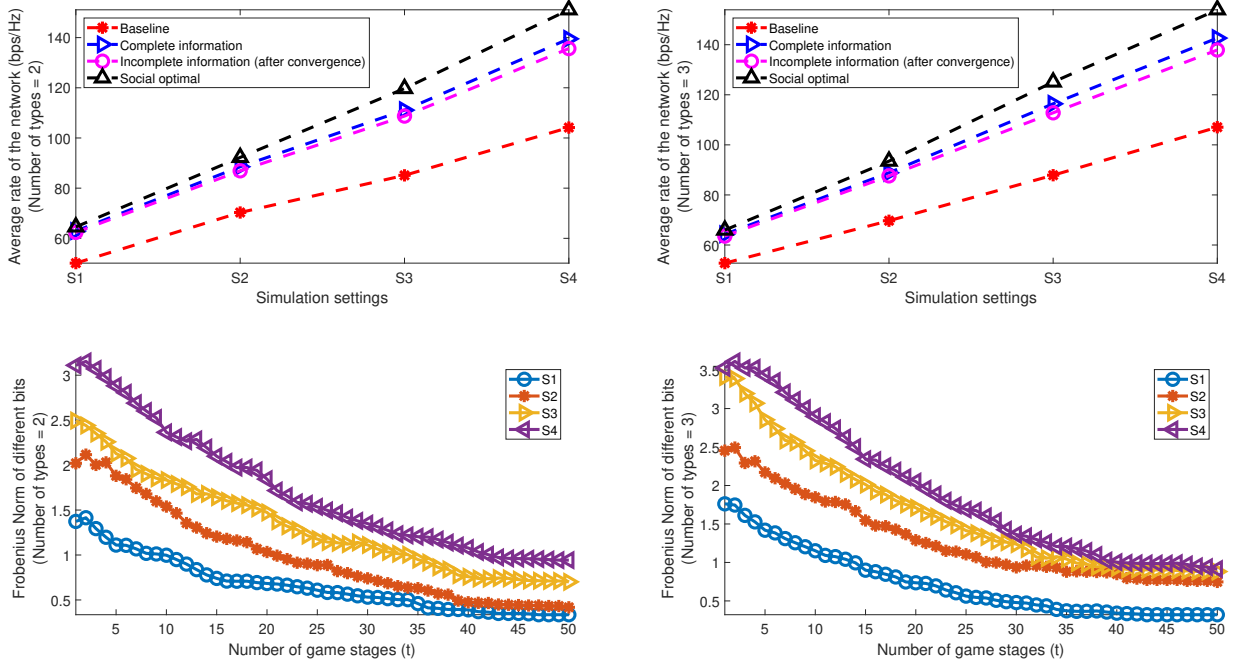
Figure 14. Average transmission rate of the network for each simulation setting shown in Table II. It shows the effect of increase in the overlap of the distributions. First case is for two types ($T_1 : \mu = 12, \sigma = 6$; $T_2 : \mu = 18, \sigma = 6$) and second case is for three types ($T_1 : \mu = 12, \sigma = 6$; $T_2 : \mu = 18, \sigma = 6$; $T_3 : \mu = 24, \sigma = 6$).

[25] O. N. Gharehshiran, A. Attar, and V. Krishnamurthy, "Collaborative sub-channel allocation in cognitive LTE femto-cells: A cooperative game-theoretic approach," *IEEE Transactions on Communications*, vol. 61, no. 1, pp. 325–334, 2012.

[26] S. S. Tanzil, O. N. Gharehshiran, and V. Krishnamurthy, "A distributed coalition game approach to femto-cloud formation," *IEEE Transactions on Cloud Computing*, vol. 7, no. 1, pp. 129–140, 2016.

[27] D. Wu, Y. Cai, and J. Wang, "A coalition formation framework for transmission scheme selection in wireless sensor networks," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 6, pp. 2620–2630, 2011.

[28] D. Wu, Y. Cai, L. Zhou, and J. Wang, "A cooperative communication scheme based on coalition formation game in clustered wireless sensor networks," *IEEE Transactions on wireless communications*, vol. 11, no. 3, pp. 1190–1200, 2012.

[29] O. N. Gharehshiran and V. Krishnamurthy, "Coalition formation for bearings-only localization in sensor networks—a cooperative game approach," *IEEE Transactions on Signal Processing*, vol. 58, no. 8, pp. 4322–4338, 2010.

[30] J. Ren, M. Xu, W. Chen, Z. Ding, and Z. Wang, "Coalition formation approaches for cooperative networks with SWIPT," *IEEE Access*, vol. 5, pp. 17 644–17 659, 2017.

[31] K. Akkarajitsakul, E. Hossain, and D. Niyato, "Coalition-based cooperative packet delivery under uncertainty: A dynamic Bayesian coalitional game," *IEEE Transactions on Mobile Computing*, vol. 12, no. 2, pp. 371–385, 2011.

[32] A. Asheralieva, "Bayesian reinforcement learning-based coalition formation for distributed resource sharing by device-to-device users in heterogeneous cellular networks," *IEEE Trans. Wirel. Commun*, vol. 16, pp. 5016–5032, 2017.

[33] Y. Xiao, K.-C. Chen, C. Yuen, Z. Han, and L. A. DaSilva, "A Bayesian overlapping coalition formation game for device-to-device spectrum sharing in cellular networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 7, pp. 4034–4051, 2015.

[34] Z. Khan, S. Glisic, L. A. DaSilva, and J. Lehtomaki, "Modeling the dynamics of coalition formation games for cooperative spectrum sharing in an interference channel," *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 3, no. 1, pp. 17–30, 2010.

[35] S. Maghsudi and M. van der Schaar, "Distributed task management in cyber-physical systems: How to cooperate under uncertainty?" *IEEE Transactions on Cognitive Communications and Networking*, vol. 5, no. 1, pp. 165–180, 2019.

[36] S. Maghsudi and E. Hossain, "Distributed user association in energy harvesting small cell networks: An exchange economy with uncertainty," *IEEE Transactions on Green Communications and Networking*, vol. 1, no. 3, pp. 294–308, 2017.

[37] J. Lyu, Y. Zeng, R. Zhang, and T. J. Lim, "Placement optimization of UAV-mounted mobile base stations," *IEEE Communications Letters*, vol. 21, no. 3, pp. 604–607, 2016.

[38] Q. Zhang, M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Machine learning for predictive on-demand deployment of UAVs for wireless communications," in *2018 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2018, pp. 1–6.

[39] M. M. Azari, F. Rosas, K.-C. Chen, and S. Pollin, "Optimal UAV positioning for terrestrial-aerial communication in presence of fading," in *2016 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2016, pp. 1–7.

[40] Z. Zhang, H. Pang, A. Georgiadis, and C. Cecati, "Wireless power transfer—an overview," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 2, pp. 1044–1058, 2019.

[41] H. W. Kuhn, "The Hungarian method for the assignment problem," *Naval research logistics quarterly*, vol. 2, no. 1-2, pp. 83–97, 1955.

[42] R. G. Gallager, *Information Theory and Reliable Communication*. Frederiksberg, Denmark: John Wiley & Sons, 1968.

[43] G. Chalkiadakis and C. Boutilier, "Sequentially optimal repeated coalition formation under uncertainty," *Autonomous Agents and Multi-Agent Systems*, vol. 24, no. 3, pp. 441–484, 2012.

[44] ——, "Bayesian reinforcement learning for coalition formation under uncertainty," in *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 3*. IEEE Computer Society, 2004, pp. 1090–1097.

[45] T. Arnold and U. Schwalbe, "Dynamic coalition formation and the core," *Journal of economic behavior & organization*, vol. 49, no. 3, pp. 363–380, 2002.

[46] A. Al-Hourani, S. Kandeepan, and A. Jamalipour, "Modeling air-to-ground path loss for low altitude platforms in urban environments," in *Global Communications Conference (GLOBECOM), 2014 IEEE*. IEEE, 2014, pp. 2898–2904.