Cooperative Anti-jamming for Infrastructure-less Wireless Networks with Stochastic Relaying

Liyang Zhang, Zhangyu Guan, and Tommaso Melodia
Department of Electrical Engineering
University at Buffalo, The State University of New York
Email: {liyangzh, zguan2, tmelodia}@buffalo.edu

***Abstract*—Denial-of-service (DoS) attacks launched by malicious jammers can pose significant threats to infrastructure-less wireless networks where a centralized controller may not be available. While significant recent research efforts have dealt with such attacks and several possible countermeasures have been proposed, little attention has been paid to the idea of *cooperative anti-jamming*.**

Inspired by this observation, we propose and study a cooperative anti-jamming scheme designed to enhance the quality of links degraded by jammers. To achieve this objective, users are allowed to cooperate at two levels. First, they cooperate to optimally regulate their channel access probabilities so that jammed users gain a higher share of channel utilization. Second, users leverage multiple-input single-output cooperative communication techniques to enhance the throughput of jammed links. We formulate the problem of optimal cooperative anti-jamming as a distributed pricing-based optimization problem and propose a best response algorithm to solve it in a distributed way. Simulations demonstrate that the proposed algorithm achieves considerable gains (compared to traditional noncooperative anti-jamming) especially under heavy traffic or high jamming power. Furthermore, by comparing the proposed algorithm with a provably-optimal centralized algorithm, we show that it achieves close-to-global optimality under moderate traffic load.

I. Introduction

Wireless networks are known to be vulnerable to denial-of-service (DoS) attacks, mostly as a consequence of their broadcast nature [1]–[3]. By radiating high-power radio-frequency signals, a malicious adversary can easily generate interference that may degrade the perceived signal-to-interference-plus-noise-ratio (SINR) and therefore the achievable link throughput of legitimate users. The situation may be exacerbated in infrastructureless wireless networks with no centralized entity to coordinate the transmission strategies of different users.

As a natural consequence, the problem of developing effective countermeasures to jamming attacks has attracted significant attention. Existing anti jamming techniques can be broadly classified into two main categories [2], i.e., techniques based on *frequency hopping* and techniques based on *optimal resource allocation*.

Frequency hopping spread spectrum (FHSS) has long been used to provide anti-jamming capabilities in wireless communications. By quickly shifting from one frequency carrier to another, FHSS allows legitimate users to actively avoid jamming attacks. However effective, FHSS has several shortcomings. First, it relies heavily on a pre-defined secret pattern. Therefore, it may not be suitable for ad hoc networks where it is difficult to share a common secret between transmitters and receivers; or for cognitive radio networks where the availability of spectrum holes may follow a random pattern [4]. Second, FHSS requires significantly more spectrum resources than single-carrier transmission strategies. Third, FHSS assumes that jammers can only jam one or a subset (but not all) of the available channels at the same time, while in some scenarios, it may be possible for a jammer to launch more powerful attacks by generating broadband interference on all the available channels. In recent years, several *adaptive* frequency hopping strategies have been proposed to address the first shortcoming, including uncoordinated frequency hopping (UFH) [5] and message-driven frequency hopping (MDFH) [6]. The communication efficiency of UFH was analyzed theoretically in [7], and practical algorithms were proposed in [8], [9]; the anti-jamming properties of MDFH were analyzed in [10], [11]. However, in scenarios with scarce spectrum resources, anti-jamming techniques that can utilize the spectrum resource more flexibly and efficiently are needed. Direct Sequence Spread Spectrum (DSSS) [12] is another effective technique against jamming. Compared to FHSS, DSSS makes it harder for the jammer to detect legitimate transmissions. However, DSSS is a purely physical layer technique; and it cannot make the network invulnerable to jamming especially when transmission power is limited [3].

Another common, more flexible, approach is to rely on *adaptive optimal resource allocation* techniques. Unlike frequency hopping, where transmission resources (transmission power and channel access probability, among others) are allocated dynamically on one channel only, in optimal resource allocation techniques a user typically tries to maximize its own information-theoretic capacity by allocating resources on several different channels, with a potential for increased diversity compared to frequency-hopping techniques. For example, in OFDM systems iterative water-filling algorithms can be used to maximize the achievable capacity of legitimate users in the presence of jammers [13]. Since each user independently and selfishly selects its optimal transmission strategy, these approaches are often analyzed using tools from non-cooperative game theory. An extensive literature has emerged using this approach [13]–[17]. In [18], optimal resource allocation and frequency-hopping are jointly analyzed, and an algorithm to select the best strategy is proposed.

In this paper, we attempt to give a positive answer to the following question: can we leverage additional degrees of diversity to provide enhanced anti-jamming capability? We observe that all the techniques discussed above take advantage

This work is based upon material supported by the National Science Foundation under grant CNS-1218717.

of *frequency diversity* exclusively. While frequency diversity is certainly an effective technique, it is not the only degree of freedom that an anti-jamming scheme can be built upon. Specifically, the *cooperative diversity* dimension has been underexplored in the context of anti-jamming techniques. Cooperative techniques can be jointly leveraged at the network, MAC, and physical layers to provide effective countermeasures against jamming. While in commercial networks it is natural for different terminals to operate selfishly and in a non-cooperative fashion, in sensor networks and tactical military networks, which are often managed by a single entity, cooperative behaviors can be more easily implemented.

Inspired by this idea, we propose and study a cooperative anti-jamming scheme designed to optimize the *fairness-constrained network throughput* in the presence of jammers. The proposed algorithm jointly optimizes the channel access probabilities and cooperative relaying probabilities of legitimate users. Legitimate users cooperate at two levels. At the medium access control layer, a cooperative channel access scheme is proposed where the channel access probabilities of different users are optimally regulated so that users degraded severely by jammers have an increased share of air time. In this way, users with good links "trade" capacity with those with jammed links. The second step is to extend the cooperation from MAC to physical layer. It is well known that, by using cooperative relays, virtual multiple-input-single-output transmission links can be formed to increase the link capacity. In the proposed scheme, users able to enhance the link capacity of another user through cooperative transmission cooperate as relays with a certain probability. Our new distributed algorithm jointly optimizes these two levels of cooperations, with significant gains in terms of achievable network throughput.

To summarize, we make the following contributions:

1) We propose the first *cooperative anti-jamming* scheme that jointly optimizes the cooperative behavior of nodes at the MAC and physical layers. To the best of our knowledge, ours is also the first anti-jamming scheme based on a virtual multiple-input-single-output (MISO) variant of cooperative communications;
2) We formulate the *optimal cooperative anti-jamming* problem as an optimization problem with the objective of maximizing the fairness-constrained network throughput. We design a *distributed* solution algorithm based on dynamic pricing that is guaranteed to converge even if the socially optimal problem is not convex.
3) We design a *provably-optimal centralized algorithm* based on the branch and bound framework and convex relaxation techniques. The algorithm provides a performance benchmark for any distributed algorithm designed to solve similar problems.
4) We compare the performance of the cooperative distributed algorithm with a non-cooperative distributed algorithm and with the optimal centralized algorithm. We show that the cooperative algorithm achieves near optimality under light and moderate traffic, and provides considerable gains compared to non-cooperative strategies.

The rest of the paper is organized as follows. Section II presents the system model and problem statement, while Section III derives a model of the utility for each legitimate user. In Section IV, we present and analyze the distributed solution algorithm of cooperative anti-jamming. The centralized algorithm is proposed in Section V. Some practical issues are discussed in Section VI. In Section VII, we analyze the performance of the proposed algorithms. Finally, we draw conclusions in Section VIII.

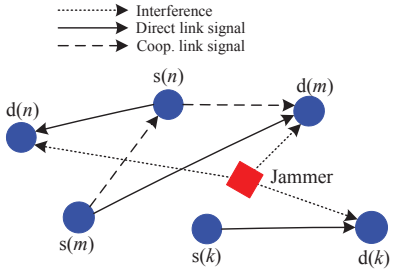


Fig. 1: An example topology with a jammer.

Notation: $\mathrm{P}(\cdot)$ represents the probability that an event occurs; $|\cdot|$ represents the cardinality of a set; $(\cdot)^{\mathrm{T}}$ represents the transpose of a vector or matrix; $\mathrm{E}(\cdot)$ is the expectation of a variable.

II. System Model

As illustrated in Fig. 1, we consider a wireless ad hoc network composed of a set $\mathcal{N}$ of legitimate users each consisting of a source-destination pair of nodes (also referred to as a *session*). Denote the source and destination nodes of each session $n \in \mathcal{N}$ as $\mathrm{s}(n)$ and $\mathrm{d}(n)$, respectively. Each $\mathrm{s}(n)$ generates data at an average rate R_n $\mathrm{bit/s}$ and then transmits the generated data to its intended receiver over a portion of the available spectrum, which is assumed to be divided into a set $\mathcal{F}$ of frequency-orthogonal-channels.

Jamming Model. We assume that there is one jammer node constrained by a limited power budget that attempts to degrade the throughput of the legitimate users by generating interference on the available channels. The model can be easily extended to the case of multiple jammers. If we denote by $\boldsymbol{p}_{\mathrm{J}} = (p_{\mathrm{J}}^f)_{f\in\mathcal{F}}$ the jammer power allocation profile, with p_{J}^f being the power allocated on channel f, we have

$$\mathbf{1}^{\mathrm{T}}\boldsymbol{p}_{\mathrm{J}} \le p_{\mathrm{J}}^{\max}, \tag{1}$$

where $p_{\mathrm{J}}^{\max}$ is the maximum power of the jammer, and $\mathbf{1}$ represents an $1 \times |\mathcal{N}|$ vector of '1' elements.

In the existing literature, jammers are typically categorized as constant jammers, deceptive jammers, random jammers, or reactive jammers [2]. We consider the more sophisticated reactive jammer model, i.e., the jammer adaptively adjusts its jamming strategy according to the channel states and other information such as the strategies of legitimate users. While an "omniscient" jammer can implement an optimal jamming strategy, it is usually not realistic for a jammer to know the transmission strategies of all legitimate users. We will therefore concentrate on a jammer with "moderate" abilities. Specifically, we assume that the jammer is only aware of the traffic on different channels at its own location. Therefore, the jamming strategy, i.e., its relative power allocation on different channels, is proportional to the sensed signal strength on each channel.

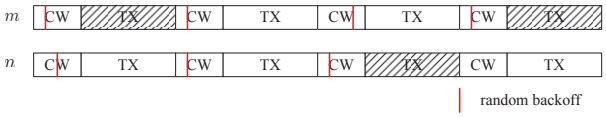


Fig. 2: Illustration of the multichannel slotted CSMA MAC.

MAC Layer Cooperation Model. We focus on a scenarios where legitimate users have more limited capabilities than the jammer. Specifically, unlike the jammer, which is able to transmit on multiple channels simultaneously, legitimate users can only use one channel at a time. The cooperative anti-jamming strategy is based on joint control of functionalities at the MAC and physical layer. At the MAC layer, users regulate their channel access probabilities to give higher opportunities to transmit to nodes that are being jammed.

We consider a stochastic MAC protocol that enables channel access cooperation in an uncoordinated network. For modeling purposes, we focus on a slotted multichannel CSMA-based protocol. However, the proposed scheme can be extended to other MAC protocols that enable stochastic channel access, with appropriate modifications in the mathematical formulation.

In the considered MAC protocol, the transmission time is divided into a set of consecutive time slots and all nodes are synchronized. At each time slot, a user either chooses a channel to compete for, or it serves as a cooperative relay. If the node chooses channel $f \in \mathcal{F}$ to compete for, it will first sense the channel at the beginning of the timeslot. If the channel is available, the node sets a random backoff and starts counting down, like in traditional CSMA. The first node counting to 0 wins the competition for the channel. We assume that the contention window size is sufficiently large so that the probability of collision is negligible. An example of the MAC protocol is illustrated in 2. We assume that the maximum contention window is set to the same value for all users on each channel. Therefore, the backoff strategy results in equal channel access probability for each contender.

With such MAC protocol, a user can regulate its channel access probability by simply adjusting the channel sensing probability on different channels. We let $q_n^f, f \in \mathcal{F}$ denote the channel sensing probabilities of user n on channel f. Since with non-zero probability node n may delay its own transmission and serve as relay for other nodes, we have $\sum_{f\in\mathcal{F}} q_n^f \leq 1$.

Physical Layer Cooperation Model. Physical-layer cooperation is obtained through relaying [19] [20]. Instead of transmitting its own traffic, a user can act as a relay and cooperatively transmit a packet on behalf of another user. Cooperative transmission is typically achieved by dividing the available transmission time into two phases: in the first phase, the transmitter broadcasts the message to both the destination and the relay; in the second phase, the relay forwards the received message to the destination, which then combines the two copies of the message and decodes. We focus on the decode-and-forward (DF) variant of cooperative communications, under which the relay node forwards the packet only when the information received from the source node can be successfully decoded. The analysis in this paper can be extended to other forwarding strategies, e.g., amplify-and-forward (AF).

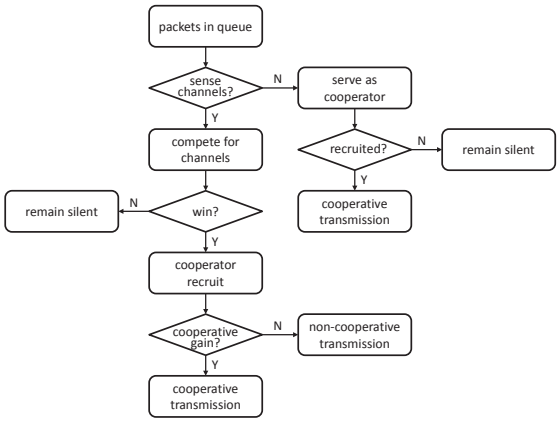


Fig. 3: Behavior of a legitimate user

To cope with the dynamic nature of the jammer, we consider a dynamic relay selection strategy to let the users form virtual MISO links. At each time slot, a user that chooses not to sense any channels will act as a relay for another user if there is a positive cooperative gain[1].

Strategy of Legitimate Users. The strategy of a legitimate user can be illustrated through the flowchart in Fig. 3. When a user is backlogged, it selects its channel sensing probability for each channel. If it chooses not to sense any channel, it serves as a potential cooperator for other legitimate users. Intuitively, the cooperative relaying probability - which will be derived formally in Section III - is a function of the channel sensing probability q_n^f, of the strategy of the jammer $\boldsymbol{p}_\mathrm{J}$, and of the network topology. Users that choose to sense the same channel compete for channel access by setting random backoffs, and the winner has the privilege to transmit.

The factors determining cooperations on the 2 layers, i.e., channel access probability and cooperative relaying probability, are both functions of channel access probability, for a given jammer strategy and network topology. Therefore, we can simply use the channel access probability as the strategy of a legitimate user, and denote it as $\boldsymbol{q}_n = (q_n^f)_{f\in\widetilde{\mathcal{F}}}$ with $\widetilde{\mathcal{F}} \triangleq \mathcal{F} \cup \{0\}$, where q_n^f indicates the sensing probability of channel f, and q_n^0 denotes the probability that n does not sense any channel. Then, we have

$$q_n^f > 0,\ \forall n \in \mathcal{N},\ \forall f \in \widetilde{\mathcal{F}} \tag{2}$$

$$q_n^f \leq 1,\ \forall n \in \mathcal{N},\ \forall f \in \widetilde{\mathcal{F}} \tag{3}$$

$$\mathbf{1}^{\mathrm{T}}\boldsymbol{q}_n = 1,\ \forall n \in \mathcal{N}. \tag{4}$$

We further denote by $\boldsymbol{q} = (\boldsymbol{q}_n)_{n\in\mathcal{N}}$ the sensing probability profile of all users in $\mathcal{N}$, and by $\boldsymbol{q}_{-n} = (\boldsymbol{q}_m)_{m\in\mathcal{N}/n}$ the profile of all users except for n.

Problem Statement. Our objective is to maximize the total utility of all legitimate users, which represents the fairness-constrained network throughput and will be defined formally in Section III, by choosing the optimal sensing probability profile for each user, for any given strategy of the jammer.

[1]Cooperative transmission does not always outperform direct transmission. The cooperative gain depends on the strategy of the jammer, the network topology and the instantaneous channel states - see Section III for details.

III. Legitimate User Utility

We consider the *expected capacity of a legitimate user* $n \in \mathcal{N}$, expressed as

$$C_n(\boldsymbol{q},\ \boldsymbol{p}_{\rm J}) = \sum_{f\in\mathcal{F}} q_n^f \rho_n^f(\boldsymbol{q},\ \boldsymbol{p}_{\rm J}) C_n^f(\boldsymbol{q},\ \boldsymbol{p}_{\rm J}), \tag{5}$$

where q_n^f is the probability that user n senses channel f, $\rho_n^f(\boldsymbol{q},\ \boldsymbol{p}_{\rm J})$ represents the probability that user n is able to successfully access the channel, and $C_n^f(\boldsymbol{q},\ \boldsymbol{p}_{\rm J})$ is the achievable capacity on that channel (through either direct transmission or by using a cooperative relay), for a given sensing probability profile $\boldsymbol{q}$ and jamming power profile $\boldsymbol{p}_{\rm J}$.

Channel Access Probability. MAC-layer cooperation is achieved through stochastic channel access. According to the slotted multichannel CSMA protocol described in Section II, user $n \in \mathcal{N}$ is able to successfully access channel $f \in \mathcal{N}$ if i) the channel is sensed to be idle at session n's source node $\mathrm{s}(n)$; and ii) session n wins the channel access competition. If we let $\widetilde{\rho}_n^f(\boldsymbol{p}_{\rm J})$ indicate the probability that channel f is idle and $\widehat{\rho}_n^f(\boldsymbol{q})$ the probability that session n wins the competition, the channel access probability $\rho_n^f(\boldsymbol{q},\ \boldsymbol{p}_{\rm J})$ in (5) can be expressed as

$$\rho_n^f(\boldsymbol{q},\ \boldsymbol{p}_{\rm J}) = \widetilde{\rho}_n^f(\boldsymbol{p}_{\rm J})\widehat{\rho}_n^f(\boldsymbol{q}). \tag{6}$$

If we let $p_{\rm th}$ represent the power threshold below which a channel is sensed idle, then $\widetilde{\rho}_n^f(\boldsymbol{p}_{\rm J})$ can be defined as

$$\widetilde{\rho}_n^f(\boldsymbol{p}_{\rm J}) \triangleq \mathrm{P}\left(p_{\rm J}^f H_{{\rm Js}(n)} \cdot (h^f_{{\rm Js}(n)})^2 + (\sigma^f_{{\rm s}(n)})^2 \le p_{\rm th}\right), \tag{7}$$

with $H_{{\rm Js}(n)},\ h^f_{{\rm Js}(n)}$ capturing path loss and fading of the link between the jammer and session n's source node $\mathrm{s}(n)$ on channel f, respectively, and $(\sigma^f_{{\rm s}(n)})^2$ being the noise power. For $h^f_{{\rm Js}(n)}$ Rayleigh distributed with fading factor $\Omega^f_{{\rm s}(n)}$, $\widetilde{\rho}_n^f(\boldsymbol{p}_{\rm J})$ in (6) can be written as

$$\widetilde{\rho}_n^f(\boldsymbol{p}_{\rm J}) = \int_0^{x_{\max}} 1 - e^{-x^2/\Omega^f_{{\rm s}(n)}}\ \mathrm{d}x, \tag{8}$$

with $x_{\max}$ calculated from (7) as

$$x_{\max} = \sqrt{(p_{\rm th} - (\sigma^f_{{\rm s}(n)})^2)/(p_{\rm J}^f H_{{\rm Js}(n)})}. \tag{9}$$

We now need to derive the probability that a user $n \in \mathcal{N}$ wins the medium access competition after sensing the channel $f \in \mathcal{F}$ to be idle. Denoting $\mathcal{N}_n^f \subset \mathcal{N}/n$ as the set of nodes competing with user n on channel f, the winning probability for user n can be written as $\frac{1}{1+|\mathcal{N}_n^f|}$, where $|\mathcal{N}_n^f|$ is the number of nodes in $\mathcal{N}_n^f$. Since each potential competing user $m \in \mathcal{N}/n$ joins the access competition with probability $q_m^f\widetilde{\rho}_m^f(\boldsymbol{p}_{\rm J})$, the cardinality of $\mathcal{N}_n^f$, i.e., $|\mathcal{N}_n^f|$, can be proven to be Poisson distributed with mean [21]

$$\mathrm{E}(|\mathcal{N}_n^f|) = \sum_{m\in\mathcal{N}/n} q_m^f \widetilde{\rho}_m^f(\boldsymbol{p}_{\rm J}). \tag{10}$$

Then, the overall probability of winning a medium access competition for user n, i.e, $\widehat{\rho}_n^f(\boldsymbol{q})$ in (6), can be expressed as

$$\widehat{\rho}_n^f(\boldsymbol{q}) = \sum_{k=0}^{|\mathcal{N}|-1} \frac{1}{1+k} \cdot \frac{\left(\mathrm{E}(|\mathcal{N}_n^f|)\right)^k e^{-\mathrm{E}(|\mathcal{N}_n^f|)}}{k!}. \tag{11}$$

Expected Capacity. Suppose that user $n \in \mathcal{N}$ has won the competition to access channel f. We can then derive the expected capacity achievable through either direct transmission or using a cooperative relay, i.e., $C_n^f(\boldsymbol{q},\ \boldsymbol{p}_{\rm J})$ in (5).

If direct transmission is used by n, the capacity is simple to derive. Denote the direct link capacity by $C^{\rm dir}_{n,f}(\boldsymbol{p}_{\rm J})$. Then, we have

$$C^{\rm dir}_{n,f}(\boldsymbol{p}_{\rm J}) = B\log\left(1+\gamma^{\rm s2d}_{n,f}(\boldsymbol{p}_{\rm J})\right), \tag{12}$$

where B is the bandwidth of each channel, and

$$\gamma^{\rm s2d}_{n,f}(\boldsymbol{p}_{\rm J}) \triangleq \frac{p_n H_n \cdot (h_n^f)^2}{(\delta^f_{{\rm d}(n)})^2 + p_{\rm J} H_{{\rm Jd}(n)} \cdot (h_{{\rm Jd}(n)})^2} \tag{13}$$

where p_n is the transmission power of user n; H_n and h_n^f are the path loss and fading, respectively; $(\delta^f_{{\rm d}(n)})^2$ is the noise power at the destination of user n denoted by $\mathrm{d}(n)$ on channel f. The expected capacity achievable with a direct link, denoted by $\widehat{C}^{\rm dir}_{n,f}(\boldsymbol{p}_{\rm J})$, can be computed by averaging over all possible channel fading outcomes of the links between $\mathrm{s}(n)$ and $\mathrm{d}(n)$, and the jammer and $\mathrm{d}(n)$, i.e.,

$$\widehat{C}^{\rm dir}_{n,f}(\boldsymbol{p}_{\rm J}) = \int_0^\infty\int_0^\infty C^{\rm dir}_{n,f}(\boldsymbol{p}_{\rm J}) \cdot \mathrm{P}(h_n^f = x_1)\mathrm{P}(h_{{\rm Jd}(n)} = x_2)\ \mathrm{d}x_1\ \mathrm{d}x_2. \tag{14}$$

As discussed in Section II, each source node $m \in \mathcal{N}/n$ serves as a potential relay with probability q_m^0. Therefore, with a certain probability, user n will receive cooperation assistance by one of the potential cooperators. Suppose user n chooses $\mathrm{s}(m)$ as the relay, then, the resulting cooperative capacity denoted by $C^{\rm cop}_{nm,f}(\boldsymbol{p}_{\rm J})$ can be expressed as [19]

$$C^{\rm cop}_{nm,f}(\boldsymbol{p}_{\rm J}) = \frac{B}{2}\log\left(1+\min\left(\gamma^{\rm s2r}_{nm,f}, \gamma^{\rm s2d}_{n,f} + \gamma^{\rm r2d}_{mn,f}\right)\right), \tag{15}$$

where $\gamma^{\rm s2r}_{nm,f} = \gamma^{\rm s2r}_{nm,f}(\boldsymbol{p}_{\rm J})$ and $\gamma^{\rm r2d}_{mn,f} = \gamma^{\rm r2d}_{mn,f}(\boldsymbol{p}_{\rm J})$ represents the SINR (defined as in (13)) of the link from source to relay, and from relay to destination, respectively.

Note, from (12) and (15), that the cooperative capacity $C^{\rm cop}_{nm,f}(\boldsymbol{p}_{\rm J})$ can be higher or lower than the direct capacity (because of the $\frac{1}{2}$ coefficient in (15)). If we define the following indicator function

$$\mathrm{I}(x,y) \triangleq \begin{cases} 1, & \text{if } x > y \\ 0, & \text{otherwise,} \end{cases} \tag{16}$$

then, the expected capacity achievable through cooperative communication (assuming that cooperative transmission outperforms direct transmission) can be defined as

$$\widehat{C}^{\rm cop}_{nm,f}(\boldsymbol{p}_{\rm J}) \triangleq \mathrm{E}\left(C^{\rm cop}_{nm,f}(\boldsymbol{p}_{\rm J}) | \mathrm{I}\left(C^{\rm cop}_{nm,f}(\boldsymbol{p}_{\rm J}), C^{\rm dir}_{n,f}(\boldsymbol{p}_{\rm J})\right) = 1\right) \tag{17}$$

The expected capacity achieved through cooperative communication in (17) can be computed by averaging over all possible channel fading outcomes of the links, for each channel $f \in \mathcal{F}$:

1) $h^f_{{\rm Jd}(n)}$: from jammer to $\mathrm{d}(n)$;
2) $h^f_{{\rm Js}(m)}$: from jammer to $\mathrm{s}(n)$;
3) h_n^f: from $\mathrm{s}(n)$ to $\mathrm{d}(n)$;
4) h^f_{nm}: from $\mathrm{s}(n)$ to $\mathrm{s}(m)$;
5) $\hat{h}^f_{mn}$: from $\mathrm{s}(m)$ to $\mathrm{d}(n)$.

Therefore, we have

$$\widehat{C}^{\text{cop}}_{nm,f}(\boldsymbol{p}_{\text{J}}) = \int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty C^{\text{cop}}_{nm,f}(\boldsymbol{p}_{\text{J}}) \cdot \text{I}\left(C^{\text{cop}}_{nm,f}(\boldsymbol{p}_{\text{J}}), C^{\text{dir}}_{n,f}(\boldsymbol{p}_{\text{J}})\right) \text{P}(h^f_{\text{Jd}(n)} = x_1) \cdot \text{P}(h^f_{\text{Js}(m)} = x_2)\text{P}(h^f_n = x_3) \cdot \text{P}(h^f_{nm} = x_4)\text{P}(\hat{h}^f_{mn} = x_5) \, \text{d}x_1 \, \text{d}x_2 \, \text{d}x_3 \, \text{d}x_4 \, \text{d}x_5. \quad (18)$$

The resulting probability that user n achieves a capacity gain through cooperative relaying can then be represented as

$$\phi^f_{nm}(\boldsymbol{p}_{\text{J}}) = q^0_m \int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty \text{I}\left(C^{\text{cop}}_{nm,f}(\boldsymbol{p}_{\text{J}}), C^{\text{dir}}_{n,f}(\boldsymbol{p}_{\text{J}})\right) \cdot \text{P}(h^f_{\text{Jd}(n)} = x_1) \cdot \text{P}(h^f_{\text{Js}(m)} = x_2)\text{P}(h^f_n = x_3) \cdot \text{P}(h^f_{nm} = x_4)\text{P}(\hat{h}^f_{mn} = x_5) \, \text{d}x_1 \, \text{d}x_2 \, \text{d}x_3 \, \text{d}x_4 \, \text{d}x_5, \quad (19)$$

and the corresponding sum probability can be written as

$$\phi^f_n(\boldsymbol{p}_{\text{J}}) = \sum_{m \in \mathcal{N}/n} \phi^f_{nm}(\boldsymbol{p}_{\text{J}}). \quad (20)$$

Finally, the expected capacity achievable by user n over channel f can be expressed as

$$C^f_n(\boldsymbol{q},\ \boldsymbol{p}_{\text{J}}) = \sum_{m \in \mathcal{N}/n} q^0_m \widehat{C}^{\text{cop}}_{nm,f}(\boldsymbol{p}_{\text{J}}) + \sum_{m \in \mathcal{N}/n} \left(1 - \phi^f_n(\boldsymbol{p}_{\text{J}})\right) \widehat{C}^{\text{dir}}_{n,f}(\boldsymbol{p}_{\text{J}}). \quad (21)$$

Note that (21) is exact when the probability that more than one cooperator participates in cooperative communication is very low. Otherwise, the capacity expression will be obtained as a sum of the expected cooperative capacities contributed by different cooperators. However, we can show through experimental results that in most cases this assumption is true. Readers are referred to the Appendix for details.

Social Problem Statement. So far, we have derived the expected capacity of each user $n \in \mathcal{N}$. If we consider a proportional fairness criterion, then the utility of each user can be defined as

$$U_n(\boldsymbol{q},\ \boldsymbol{p}_{\text{J}}) \triangleq \log(C_n(\boldsymbol{q},\ \boldsymbol{p}_{\text{J}})), \quad (22)$$

and the ideal objective is to maximize the sum utility of all users, i.e.,

$$\begin{array}{rl} \text{Given} & \boldsymbol{p}_{\text{J}} \\ \underset{\boldsymbol{q} \in (0,1]^{|\mathcal{N}|}}{\text{Maximize}} & U(\boldsymbol{q}, \boldsymbol{p}_{\text{J}}) = \sum\limits_{n \in \mathcal{N}} U_n(\boldsymbol{q},\ \boldsymbol{p}_{\text{J}}) \\ \text{subject to} & (2),\ (3),\ (4). \end{array} \quad (23)$$

However, this objective is not easily achievable with distributed control. In fact, the optimization problem is non-convex and the utility expressions in (2)-(22) are rather complex. Moreover, the non-convexity also implies that only suboptimal solutions can be computed in polynomial time even with centralized algorithms. Since we would like to design distributed solutions with low complexity, we follow here a different approach design a pricing-based distributed solution algorithm with provable convergence to a stationary point of the social problem.

IV. Distributed Solution Algorithm

The distributed solution algorithm is designed based on the recent framework results in [22], with the objective to achieve a stationary solution point of the social problem (23). Specifically, we design an iterative best-response algorithm based on a pricing mechanism. At each iteration, each session n maximizes its own utility minus a pricing term that acts as a penalty imposed to each session for being too aggressive in choosing its own strategy and thus "hurting" other sessions. The challenge in applying the framework is to design the pricing term of each iteration so that the distributed algorithm converges to a "good" stationary point (if more than one exists) of the social problem in (23). Since we are designing algorithm for legitimate users for which the strategy of the jammer, i.e., $\boldsymbol{p}_{\text{J}}$ is a given parameter, we will neglect it from the utility function for simplicity in this section and Section V.

We denote by $\boldsymbol{q}^\nu$ the sensing probability profile of iteration ν (with $\nu = 1, 2, \dots$). The pricing term for session $n \in \mathcal{N}$, denoted as $\Gamma_n(\boldsymbol{q}_n, \boldsymbol{q}^\nu_{-n})$, can be written as

$$\Gamma_n(\boldsymbol{q}_n, \boldsymbol{q}^\nu_{-n}) \triangleq (\boldsymbol{q}_n)^{\text{T}} \left(\Gamma^f_n(q^f_n, \boldsymbol{q}^\nu_{-n})\right)_{f \in \widetilde{\mathcal{F}}} - \frac{\tau_n}{2}\|\boldsymbol{q}_n - \boldsymbol{q}^\nu_n\|^2, \quad (24)$$

where

$$\Gamma^f_n(q^f_n, \boldsymbol{q}^\nu_{-n}) \triangleq \sum_{m \in \mathcal{N}/n} \frac{\partial U_m(\boldsymbol{q}^\nu)}{\partial q^f_n} \quad (25)$$

represents the marginal decrease of the sum-utility of the other sessions due to a variation of session n's sensing probability associated with channel f. Here $-\frac{\tau_n}{2}\|\boldsymbol{q}_n - \boldsymbol{q}^\nu_n\|^2$ is a proximal regulation with constant τ_n, whose value needs to be chosen properly to guarantee strong concavity of the resulting penalized utility function, and at the same time to prevent each session n from being too conservative in changing its sensing probability profile. To discuss the convergence, we first introduce Lemma 1.

Lemma 1: Given the sensing probability profiles of all other users $\boldsymbol{q}_{-n}$, the utility function $U_n(\boldsymbol{q}_n,\ \boldsymbol{q}_{-n})$ defined in (22) is strongly concave with respect to $\boldsymbol{q}_n$.

Proof: Since $C_n(\boldsymbol{q})$ is a linear function of $\boldsymbol{q}_n$ given $\boldsymbol{q}_{-n}$ and $C_n(\boldsymbol{q}) > 0$, $U_n(\boldsymbol{q}_n, \boldsymbol{q}_{-n})$ in (22) is concave over $\boldsymbol{q}_n$. Therefore all we need to prove is that the second derivative $\nabla^2_{\boldsymbol{q}_n} U_n(\boldsymbol{q}_n, \boldsymbol{q}_{-n})$ is bounded for $\forall \boldsymbol{q}_{-n} \in \Phi_{-n} \triangleq (\Phi_m)_{m \in \mathcal{N}/n}$ with

$$\Phi_m \triangleq \{\boldsymbol{q}_m | \text{constraints} : (2), (3), (4)\}. \quad (26)$$

The second derivative of $U(\boldsymbol{q}_n, \boldsymbol{q}_{-n})$ with respect to $\boldsymbol{q}_n$ can be written as

$$\nabla^2_{\boldsymbol{q}_n} U(\boldsymbol{q}_n, \boldsymbol{q}_{-n}) = \frac{1}{C_n(\boldsymbol{q}_n, \boldsymbol{q}_{-n})} \nabla^2_{\boldsymbol{q}_n} C_n(\boldsymbol{q}_n, \boldsymbol{q}_{-n}) - \frac{1}{C^2_n(\boldsymbol{q}_n, \boldsymbol{q}_{-n})} \nabla_{\boldsymbol{q}_n} C_n(\boldsymbol{q}_n, \boldsymbol{q}_{-n}).$$

It can be verified that both $\nabla_{\boldsymbol{q}_n} C_n(\boldsymbol{q}_n, \boldsymbol{q}_{-n})$ and $\nabla^2_{\boldsymbol{q}_n} C_n(\boldsymbol{q}_n, \boldsymbol{q}_{-n})$ are bounded for closed Φ_n. Since we

let $q_n^f > 0$ for $\forall n \in \mathcal{N}, f \in \mathcal{F}$, $\frac{1}{C_n(\boldsymbol{q}_n, \boldsymbol{q}_{-n})}$ and $\frac{1}{C_n^2(\boldsymbol{q}_n, \boldsymbol{q}_{-n})}$ are also bounded. Hence, $\nabla^2_{\boldsymbol{q}_n} U(\boldsymbol{q}_n, \boldsymbol{q}_{-n})$ is bounded. ■

Because Lemma 1 guarantees strong concavicity, we can set $\tau_n = 0$. The formal description of the algorithm is given in Algorithm 1, where the penalized version of utility function $U_n(\boldsymbol{q}_n, \boldsymbol{q}_{-n})$, denoted by $\widetilde{U}_n(\boldsymbol{q}_n,\ \boldsymbol{q}_{-n})$, is defined as

$$\widetilde{U}_n(\boldsymbol{q}_n,\ \boldsymbol{q}_{-n}) \triangleq U_n(\boldsymbol{q}_n,\ \boldsymbol{q}_{-n}) + \Gamma_n(\boldsymbol{q}_n, \boldsymbol{q}_{-n}^{\nu}, 0), \tag{27}$$

with $\Gamma_n(\boldsymbol{q}_n, \boldsymbol{q}_{-n}^{\nu}, 0)$ defined in (24). The convergence properties of Algorithm 1 are given in Theorem 1 below, where ζ is a parameter to guarantee the convergence of the algorithm.

Algorithm 1: Pricing Jacobi Algorithm

Data : $\{\zeta^{\nu}\} > 0$; Set $\nu = 0$.
(S.1) : If $\boldsymbol{q}^{\nu}$ satisfies a suitable termination criterion: STOP;
(S.2) : For all $n \in \mathcal{N}$, compute

$$\widehat{\boldsymbol{q}}_n(\boldsymbol{q}^{\nu}) \triangleq \underset{\boldsymbol{q}_n \in \Phi_n}{\arg\max}\ \widetilde{U}_n(\boldsymbol{q}_n,\ \boldsymbol{q}_{-n}) \tag{28}$$

(S.3) : Set $\boldsymbol{q}_n^{\nu+1} = \boldsymbol{q}_n^{\nu} + \zeta^{\nu}(\widehat{\boldsymbol{q}}_n(\boldsymbol{q}^{\nu}) - \boldsymbol{q}_n^{\nu})$.
(S.4) : $\nu \leftarrow \nu + 1$ and go to (S.1).

Theorem 1 (Convergence Condition): Given the social problem (23), suppose that $\{\zeta^{\nu}\}$ is chosen so that

$$\zeta^{\nu} \in (0, 1],\ \zeta^{\nu} \to 0,\ \text{and}\ \sum_{\nu} \zeta^{\nu} = +\infty. \tag{29}$$

Then, either Algorithm 1 converges in a finite number of iterations to a stationary solution of (23), or every limit point of the sequence $\{\boldsymbol{\zeta}^{\nu}\}$ (at least one of such points exists) is a stationary solution of (23). Moreover, no such point is a local minimum of the social function.

Proof: Based on the Descent Lemma in [23], it can be proven that the algorithm always converges to a feasible solution point of the social problem in (23). Then, together with Lemma 1, it can be further proven that each such solution point is also stationary for the social problem. An example of sequence η^{ν} satisfying conditions (29) in Theorem 1 is [22]:

$$\zeta^{\nu} = \frac{\zeta^{\nu-1} + \alpha(\nu)}{1 + \beta(\nu)},\ \nu = 1, \ldots, \tag{30}$$

where $\alpha(\nu) = \alpha$ and $\beta(\nu) = \nu\beta$ with $\alpha, \beta \in (0, 1)$ and $\alpha \leq \beta$. ■

V. Centralized Solution Algorithm

We now present a centralized solution algorithm to solve the social problem (23) to provide a performance benchmark for the distributed solution algorithm proposed in Section IV.

Objective. Denote U^* as the global optimum of the social problem, and $\varepsilon \in (0, 1]$ as a predefined optimality precision, then the objective of the algorithm is to obtain an ε-optimal solution $\boldsymbol{q}$ satisfying

$$U(\boldsymbol{q}) \geq \varepsilon U^*. \tag{31}$$

Here, the optimality precision ε can be set as close to 1 as we wish at the price of computational complexity.

Denote $\mathrm{UP}_{\mathrm{glb}}$ as a global upper bound, and $\mathrm{LR}_{\mathrm{glb}}$ as a global lower bound on the sum-utility $U(\boldsymbol{q})$ in (23), then it must be

$$\mathrm{LR}_{\mathrm{glb}} \leq U^* \leq \mathrm{UP}_{\mathrm{glb}}. \tag{32}$$

Then, the algorithm searches for the ε-optimal solution by iteratively updating $\mathrm{UP}_{\mathrm{glb}}$ and $\mathrm{LR}_{\mathrm{glb}}$ so that, the two bounds get closer and closer to each other, until

$$\mathrm{LR}_{\mathrm{glb}} \geq \varepsilon \cdot \mathrm{UP}_{\mathrm{glb}}. \tag{33}$$

We implement the above iteration based on a combination of the *branch-and-bound* framework and convex relaxations [24].

Algorithmic Framework. We solve a series of subproblems of the original social problem (23), obtained by partitioning its domain into a set of subdomains. Denote $\Phi_{\mathcal{N}} = \prod_{n \in \mathcal{N}} \Phi_n$ as the joint domain of all the users in $\mathcal{N}$ with Φ_n defined in (26); and $\Phi = \{\Phi_{\mathcal{N}}^i, i = 0, 1, 2, \ldots\}$ as the set of subdomains, with i denoting the subdomain index, $\Phi_{\mathcal{N}}^i = \Phi_{\mathcal{N}}$ for $i = 0$, and $\Phi_{\mathcal{N}}^i \subset \Phi_{\mathcal{N}}$ for the others. For each subproblem $\Phi_{\mathcal{N}}^i$, denote the local upper and lower bounds on sum-utility $U(\boldsymbol{q})$ by $\mathrm{UP}(\Phi_{\mathcal{N}}^i)$ and $\mathrm{LR}(\Phi_{\mathcal{N}}^i)$, respectively. Then, the global upper bound $\mathrm{UP}_{\mathrm{glb}}$, and lower bound $\mathrm{LR}_{\mathrm{glb}}$ are updated as follows.

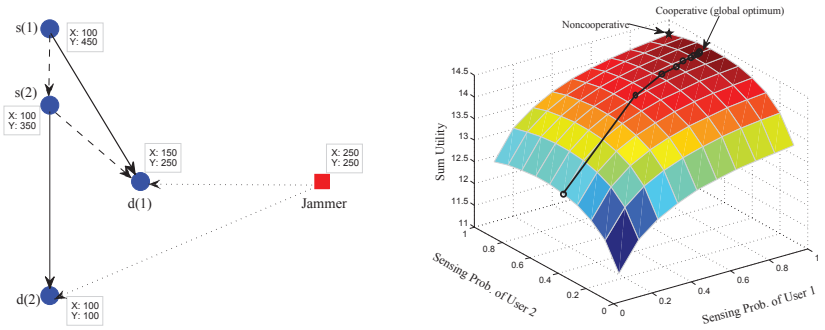
$$\mathrm{UP}_{\mathrm{glb}} = \max_{i=0,1,\ldots} \{\mathrm{UP}(\Phi_{\mathcal{N}}^i)\} \tag{34}$$

$$\mathrm{LR}_{\mathrm{glb}} = \max_{i=0,1,\ldots} \{\mathrm{LR}(\Phi_{\mathcal{N}}^i)\}. \tag{35}$$

The algorithm then checks how close the obtained global bounds are to each other. If the termination criterion (33) is satisfied, the algorithm terminates and sets the ε-optimal solution as $U(\boldsymbol{q}) = \mathrm{LR}_{\mathrm{glb}}$, and sets $\boldsymbol{q}$ accordingly; otherwise, the algorithm chooses one subdomain from Φ, partitions it into two smaller subdomains, then calculates the local upper and lower bounds for them each, and again updates $\mathrm{UP}_{\mathrm{glb}}$ and $\mathrm{LR}_{\mathrm{glb}}$. The above procedure is repeated until the gap between $\mathrm{UP}_{\mathrm{glb}}$ and $\mathrm{LR}_{\mathrm{glb}}$ converges to 0, and hence [according to (32)] converges to the global optimum U^*.

Convex Relaxation. In the above iterations, for a given $\Phi_{\mathcal{N}}^i$, the corresponding local upper bound $\mathrm{UP}(\Phi_{\mathcal{N}}^i)$ needs to be easy to compute. To this end, we rely on convex relaxation, i.e., we relax the original nonlinear nonconvex problem into a convex one that is easy to solve using standard convex programming techniques. We call the solution obtained by solving the relaxed optimization problem *relaxed solution*. Since the relaxed solution is also a feasible solution, we compute the sum throughput based on (22), and set the local lower bound $\mathrm{LR}(\Phi_{\mathcal{N}}^i)$ to the resulting solve.

To relax the objective function in (23) to be convex, we only need to relax the individual utility function of each user. Different approaches can be used (see [24] for details of possible relaxation techniques). Here, we adopt a simple but effective relaxation method based on the observations that $U_n(\boldsymbol{q})$ is a monotonically decreasing function with respect to q_m^f for any $f \in \mathcal{F}$. For given $\Phi_{\mathcal{N}}^i$, denoting the range of q_m^f as $[q_{m,f}^{\mathrm{L}}\ q_{m,f}^{\mathrm{U}}]$, $U_n\left(\boldsymbol{q}_n, (q_{m,f}^{L})_{m \in \mathcal{N}/n}^{f \in \mathcal{F}}\right)$ provides an upper bound on $U_n(\boldsymbol{q}_n, \boldsymbol{q}_{-n})$. By deriving the first and second derivatives of $U_n(\boldsymbol{q}_n, \boldsymbol{q}_{-n})$ with respect to $\boldsymbol{q}_n$, it can be seen that $U_n(\boldsymbol{q}_n)$ is a concave function whose global optimum can be easily computed, e.g., by using standard interior-point methods [25].



(a) (b)

Fig. 4: (a) Toy topology. X, Y: x and y coordinates, respectively; (b) Convergence of the distributed algorithm (to the global optimum in this example)

Variable Partition. We select the subdomain $\Phi^i_{\mathcal{N}}$ with the highest local upper bound from Φ for partition, i.e.,

$$i = \arg\max_i \mathrm{UP}(\Phi^i_{\mathcal{N}}). \tag{36}$$

The selected subproblem is then partitioned into two new subproblems by partitioning one of its variables, i.e., $\{q^f_n,\ n \in \mathcal{N}, f \in \mathcal{F}\}$. We select the variable with the largest range and partition it from the half, i.e., selecting q_{n^*,f^*} that satisfies

$$\{n^*, f^*\} = \arg\max_{n\in\mathcal{N}, f\in\mathcal{F}} (q^{\mathrm{U}}_{n,f} - q^{\mathrm{L}}_{n,f}) \tag{37}$$

and partition it as

$$q^{\mathrm{M}}_{n^*,f^*} = \frac{q^{\mathrm{U}}_{n^*,f^*} + q^{\mathrm{L}}_{n^*,f^*}}{2}, \tag{38}$$

which results in two new subproblems with domains of $[q^{\mathrm{L}}_{n^*,f^*}\ q^{\mathrm{M}}_{n^*,f^*}]$ and $[q^{\mathrm{M}}_{n^*,f^*}\ q^{\mathrm{U}}_{n^*,f^*}]$, respectively.

VI. Practical Considerations

Information Acquisition. Player strategies (for both legitimate users and jammer) are coupled. Therefore, some information exchange is required. Note that a node only needs to know the strategies of nodes within its contention range. Therefore, each node needs to periodically broadcast its strategy to the nodes within contention range. Legitimate users can exchange information through a control channel. Note that there is no information exchange between users and jammer. However, as described in Section II, the jammer does not have to know the strategies of the users. Instead, it adjusts its strategy based on traffic sensed on different channels.

Legitimate users, instead, must base their strategies on the strategy of the jammer because the jamming power allocation appears in the utility expression of legitimate users. We observe however that the jamming power only appears in the expression of the interference component in (13). Since a CSMA protocol at the MAC layer is considered, there is in principle no mutual interference between different users. Therefore, a user can estimate the jamming power by averaging over the receiver-side interference over time. Let us consider a node n as an example. The interference plus noise on channel f at $\mathrm{d}(n)$ is $IN_n = p^f_{\mathrm{J}} H_{\mathrm{Jd}(n)} \cdot (h_{\mathrm{Jd}(n)})^2 + (\delta^f_{\mathrm{d}(n)})^2$, which is a random variable for a given strategy of the jammer. The randomness lies in the noise, $(\delta^f_{\mathrm{d}(n)})^2$ and channel fading $h_{\mathrm{Jd}(n)}$. If the noise statistics are available, then user n can easily obtain the average interference value. Since channel fading is Rayleigh distributed, the interference, as a scalar of the square of the channel fading, is exponentially distributed. When a sufficient number of samples is available, user n can estimate the distribution of the interference, and calculate the expected capacity.

Cooperator Recruitment. The expected cooperative capacity in (21) is obtained by summing up the contributions of every potential relay, so an important prerequisite for this equation to approximate well the real cooperative capacity is that the probability that 2 or more relays with cooperative gain exist simultaneously is negligible.

Fortunately, in the scenario we are focusing on, i.e., a network with moderate or heavy traffic load and a powerful jammer, it is reasonable to assume that most of the legitimate users are affected by the jammer simultaneously, and thus the probability that relays exist with positive cooperative gain is not high in most cases (see Appendix). Besides, even if a relay is able to enhance another user's link through cooperation, it will only act as relay with a certain, typically small, probability (only when it chooses not to sense any channel). In most cases, then, the probability that multiple potential relays exist for a user in the same time slot is very low, and the approximation in (21) is quite accurate.

Based on this observation, we can just consider a simple relay selection rule, i.e., choose the available relay with maximum cooperative gain. In most cases, this rule will result in 0 or 1 available relays. Only in very rare cases, there will be 2 or more candidates. For scenarios in which the jammer is not very powerful and many cooperation opportunities exist, the cooperative capacity in (21) becomes an upper bound on the real value because of the overlap in cooperation probabilities.

VII. Performance Evaluation

System Setup. The topology is generated randomly. Specifically, all nodes are located in a $500\,\mathrm{m} \times 500\,\mathrm{m}$ area, with the distance between the transmitter-receiver pairs generated uniformly between 250 and $350\mathrm{m}$. The location of the jammer is fixed to the position $(250, 250)$. In most experiments, we assume that there are 2 different channels. The Rayleigh fading coefficients of the channels are set to different values generated uniformly from $[\frac{1}{2}, \frac{2}{3}]$. We set the path loss factor to 4.

Without loss of generality, the power of legitimate users is set to $1\,\mathrm{W}$, while the average noise power is set to $1 \times 10^{-10}\,\mathrm{W}$. The power of the jammer is set to different values in different experiments, but generally it is much higher than that

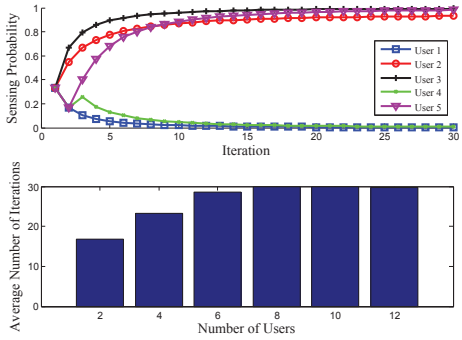


Fig. 5: Convergence of the distributed solution algorithm.

of the users to better highlight the anti-jamming performance. The jammer allocates its power budget to different channels proportionally to the sensed signal strength on each of them. The bandwidth of each channel is set to 20 kHz.

Case Study. We first show a toy example to gain insights on the convergence and optimality of the algorithm. We consider a 2-user-1-channel scenario. We vary the sensing probability of both users from 0 to 1 and calculate the expected utility for every possible tuple. We set the Rayleigh fading factors to 0.5. The locations of the nodes are as shown in Fig. 4(a).

We observe from Fig. 4 that user 2 is able to cooperate with user 1. We verify that the algorithm converges to $(0.95, 0.71)$ (the black line in the figure shows the convergence path). We can also easily verify that the convergence point is the global optimum. The non-cooperative optimum (each user senses the channel with probability 1 since there is only 1 channel in this case) is compared with our algorithm. Specifically, the total utility of the non-cooperative algorithm is 14.18 , while the total utility of the distributed algorithm is 14.53. The gain is moderate because the traffic in this case is fairly light (2 users competing for 1 channel).

Convergence Analysis. We now evaluate the convergence speed of the proposed algorithm. We set the number of users to 10, and the number of channels to 2. The power of the jammer is 10 W. The result is shown in Fig. 5(a). We only plot the strategy updates of users 1 to 5 on channel 1 for readability. We observe that the strategies converge quite quickly, i.e., within 20 iterations. The average convergence speed of our algorithm is also shown in Fig. 5(b). We vary the number of users from 2 to 10 in steps of 2, with the same settings for other parameters. For a fixed number of users, we randomly generate 20 topologies and calculate the average convergence speed. The algorithm is considered to have converged if the element-wise absolute difference of two consecutive iterations is no larger than 0.005. We observe that the distributed algorithm converges within about 30 iterations in all cases.

Utility Comparison. Finally, we compare the utility of our distributed algorithm vs. the frequency hopping algorithm and the centralized algorithm. In the frequency hopping algorithm, users select the best instantaneously available channel. Since we aim at maximizing the sum-log capacity, to make the comparison fair, we consider a frequency hopping algorithm designed to maximize the same objective function. The scenario is the same as described above. We vary the number of users between 2 and 18 in steps of 2, as shown in Fig. 6.

In all considered scenarios there are gains for our algorithm, up to 19.6%. We observe that, since the utility represents

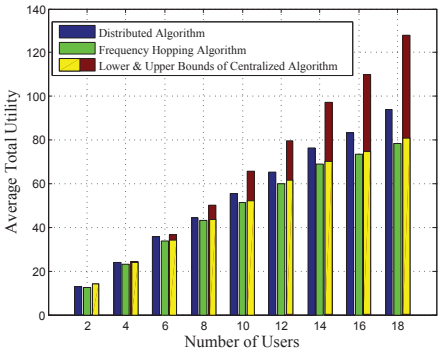


Fig. 6: Utility vs number of users.

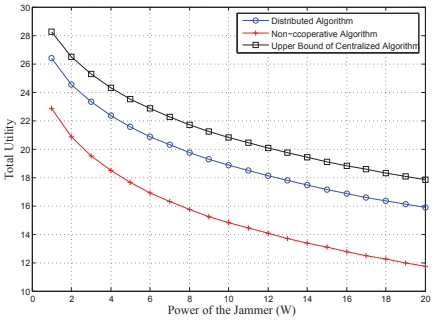


Fig. 7: Utility vs jamming power.

the logarithm of capacity, a gain of 19.6% is considerable. Compared with the centralized algorithm, when the number of users is small, our algorithm achieves utility very close to the upper bound of the centralized algorithm. To be specific, when the number of users is below 12, we achieve more than 80% of the upper bound of the centralized optimal value. In these cases, the lower bounds and the upper bounds of the centralized algorithm converge to one point, so the upper bound is actually the optimal point and our distributed algorithm has near-optimal utility performance. When the number of users is large, there are gaps between our algorithm and the upper bound. However, in these cases, the branch-and-bound-based algorithm fails to converge within the maximum number of iterations we set.[2] Therefore, the upper bound does not necessarily represent the actual optimal value. So, for these two cases, we are unable to make any conclusive statement about the global optimality.

Besides the density of users, the jamming power is also an important factor affecting the performance of the anti-jamming algorithm. To analyze this, we fix the number of users to 6. We vary the power of the jammer from 1 to 20 W. To better illustrate the impact of the jamming power, we fix the topology and the Rayleigh fading factors for each jamming power. The results are shown in Fig. 7. It can be observed that as the power of the jammer increases, the utility of both algorithms decreases. Our algorithm always outperforms the non-cooperative algorithm, with a gain up to 34.9%. An important observation is that the gain increases when the power of the jammer increases. This implies that the more severe jamming the nodes experience, the better performance our distributed algorithm obtains. Compared to the centralized

[2]Although theoretically the algorithm will eventually converge, there is no guarantee of convergence speed. In our experiment, the maximum number of iterations is set to 30000.

algorithm, our algorithm achieves approximately 90% of the upper bound, so we still have near-optimal utility.

VIII. Conclusions

We proposed and designed a cooperative anti-jamming scheme by introducing the notion of cooperative diversity into anti-jamming. There are two levels of cooperations. At the medium access control layer, a cooperative channel access scheme is proposed where the channel access probabilities of different users are optimally regulated so that users degraded severely by jammers have an increased share of air time; at the physical layer, users able to enhance the link capacity of another user through cooperative transmission cooperate as relays with a certain probability. We designed a pricing-based distributed algorithm to jointly optimize these two levels of cooperations. We proved that the algorithm always converges, even if the centralized optimization problem cannot be proven to be convex. Compared to non-cooperative algorithms, our algorithm achieves considerable gains. By comparing it with a branch-and-bound based centralized algorithm, we also showed that the proposed distributed algorithm achieves almost-global optimality in most cases. The gain is shown to increase with increasing network traffic and with jamming power. Our results also demonstrate significant cooperative gains when a network is experiencing very low throughput.

References

[1] W. Xu, W. Trappe, Y. Zhang, and T. Wood, "The feasibility of launching and detecting jamming attacks in wireless networks," in *Proc. of ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, NY, USA, May 2005.

[2] W. Xu, K. Ma, W. Trappe, and Y. Zhang, "Jamming sensor networks: attack and defense strategies," *IEEE Trans. on Networking*, vol. 20, no. 3, pp. 41–47, May 2006.

[3] A. Mpitziopoulos, D. Gavalas, C. Konstantopoulos, and G. Pantziou, "A survey on jamming attacks and countermeasures in WSNs," *IEEE Communications Surveys Tutorials*, vol. 11, no. 4, pp. 42–56, Fourth Quarter 2009.

[4] L. Ding, T. Melodia, S. Batalama, J. Matyjas, and M. Medley, "Cross-layer Routing and Dynamic Spectrum Allocation in Cognitive Radio Ad Hoc Networks," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 4, pp. 1969–1979, May 2010.

[5] M. Strasser, S. Capkun, C. Popper, and M. Cagalj, "Jamming-resistant key establishment using uncoordinated frequency hopping," in *Proc. of IEEE Symposium on Security and Privacy*, Oakland, CA, USA, May 2008.

[6] Q. Ling and T. Li, "Message-driven frequency hopping: Design and analysis," *IEEE Transactions on Wireless Communications*, vol. 8, no. 4, pp. 1773–1782, April 2009.

[7] C. Li, H. Dai, L. Xiao, and P. Ning, "Communication efficiency of anti-jamming broadcast in large-scale multi-channel wireless networks," *IEEE Transactions on Signal Processing*, vol. 60, no. 10, pp. 5281–5292, Oct. 2012.

[8] Q. Wang, P. Xu, K. Ren, and X. Li, "Delay-bounded adaptive UFH-based anti-jamming wireless communication," in *Proc. of IEEE International Conference on Computer Communications (INFOCOM)*, Shanghai, China, April 2011.

[9] K. Xu, Q. Wang, and K. Ren, "Joint UFH and power control for effective wireless anti-jamming communication," in *Proc. of IEEE International Conference on Computer Communications (INFOCOM)*, Orlando, FL, USA, March 2012.

[10] L. Zhang, H. Wang, and T. Li, "Anti-jamming message-driven frequency hopping part I: System design," *IEEE Transactions on Wireless Communications*, vol. 12, no. 1, pp. 70–79, Jan. 2013.

[11] L. Zhang and T. Li, "Anti-jamming message-driven frequency hopping part II: Capacity analysis under disguised jamming," *IEEE Transactions on Wireless Communications*, vol. 12, no. 1, pp. 80–88, Jan. 2013.

[12] L. Ding, K. Gao, T. Melodia, S. Batalama, D. Pados, and J. Matyjas, "All-spectrum Cognitive Networking through Jointly Optimal Distributed Channelization and Routing," *IEEE Transactions on Wireless Communications*, vol. 12, no. 11, pp. 5394–5405, Nov. 2013.

[13] R. Gohary, Y. Huang, Z.-Q. Luo, and J.-S. Pang, "A generalized iterative water-filling algorithm for distributed power control in the presence of a jammer," in *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Taipei Taiwan, April 2009.

[14] Y. Sagduyu, R. Berry, and A. Ephremides, "MAC games for distributed wireless network security with incomplete information of selfish and malicious user types," in *Proc. of International Conference on Game Theory for Networks (GameNets)*, Istanbul, Turkey, May 2009.

[15] Y. E. Sagduyu, R. A. Berryt, and A. Ephremidesi, "Wireless jamming attacks under dynamic traffic uncertainty," in *Proc. of International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt)*, Avignon, France, June 2010.

[16] B. Wang, Y. Wu, K. J. R. Liu, and T. Clancy, "An anti-jamming stochastic game for cognitive radio networks," *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 29, no. 4, pp. 877–889, April 2011.

[17] Y. Wu, B. Wang, K. Liu, and T. Clancy, "Anti-jamming games in multi-channel cognitive radio networks," *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 30, no. 1, pp. 4–15, Jan. 2012.

[18] X. Li, Y. Zhu, and B. Li, "Optimal anti-jamming strategy in sensor networks," in *Proc. of IEEE International Conference on Communications (ICC)*, Ottawa, Canada, June 2012.

[19] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative Diversity in Wireless Networks: Efficient Protocols and Outage Behavior," *IEEE Trans. on Info. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[20] Z. Guan, T. Melodia, D. Yuan, and D. A. Pados, "Distributed spectrum management and relay selection in interference-limited cooperative wireless networks," in *Proc. of ACM Intl. Conf. on Mobile Computing and Networking (MobiCom)*, Las Vegas, Nevada, USA, Sep. 2011.

[21] K. Butler and M. Stephens, "The distribution of a sum of binomial random variables," Tech. Rep. 467, Stanford University, Dept. of Statistics, 28 April 1993.

[22] G. Scutari, F. Facchinei, P. Song, D. Palomar, and J.-S. Pang, "Decomposition by Partial Linearization: Parallel Optimization of Multi-Agent Systems," *IEEE Trans. on Signal Processing*, vol. 62, no. 3, pp. 641–656, 2014.

[23] D. Bertsekas, *Nonlinear Programming*. Belmont, MA, USA: Athena Scientific, 2th Ed., 1999.

[24] H. D. Sherali and W. P. Adams, *A Reformulation-Linearization Technique for Solving Discrete and Continuous Nonconvex Problems*. Boston: MA: Kluwer Academic, 1999.

[25] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

Appendix

Here, we verify the assumption that cooperation opportunities are sparse. We consider a network with moderate traffic, i.e., with 10 sessions and 4 channels. All other settings are the same as in Section VII. The experiment is repeated with multiple randomly generated topologies. Figure 8 shows the average probability that a user can be assisted by a cooperative relay with positive gain, i.e., the probability that a cooperation opportunity exists. For readability, only selected channels are shown. We observe that the assumption that opportunities for cooperation are sparse is verified. In fact, for most users on most channels, the average probability that another user can provide cooperative gain is fairly low (below 0.1). Moreover, since the probability of cooperative transmission is the product of the probability shown in the figure and the probability that the cooperator does not sense any channel, the resulting cooperation probability is even lower.

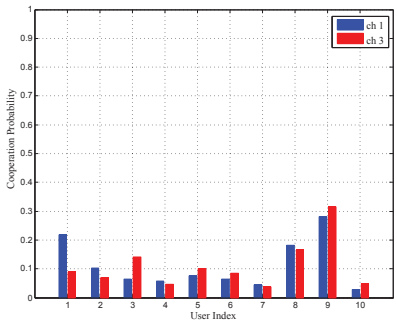


Fig. 8: Probability of positive cooperative gains.