To Transmit or Not to Transmit? Distributed Queueing Games in Infrastructureless Wireless Networks

Zhangyu Guan, *Member, IEEE*, Tommaso Melodia, *Member, IEEE*, and Gesualdo Scutari, *Senior Member, IEEE*

***Abstract*—We study distributed queueing games in interference-limited wireless networks. We formulate the throughput maximization problem via distributed selection of users' transmission thresholds as a Nash Equilibrium Problem (NEP). We first focus on the solution analysis of the NEP and derive sufficient conditions for the existence and uniqueness of a Nash Equilibrium (NE). Then, we develop a general best-response-based algorithmic framework wherein the users can explicitly choose the degree of desired cooperation and signaling, converging to different types of solutions, namely: 1) a NE of the NEP when there is no cooperation among users and 2) a stationary point of the Network Utility Maximization (NUM) problem associated with the NEP, when some cooperation among the users in the form of (pricing) message passing is allowed. Finally, as a benchmark, we design a globally optimal but centralized solution method for the nonconvex NUM problem. Our experiments show that in many scenarios the sum-throughput at the NE of the NEP is very close to the global optimum of the NUM problem, which validates our noncooperative and distributed approach. When the gap of the NE from the global optimality is non negligible (e.g., in the presence of "high" coupling among users), exploiting cooperation among the users in the form of pricing enhances the system performance.**

***Index Terms*—Decentralized control, interference limited wireless networks, queueing games, successive convex approximation..**

I. Introduction

QUEUEING theory [2], [3] has served as a fundamental analytical tool to model and understand the behavior of computer networks and other complex systems. A common assumption underlying queuing theory is that arrivals are the product of exogenous factors.[1] Therefore, the traffic arriving at a queue does not depend on interactions between different users in the network. Typical approaches rely then on describing inputs as a stochastic process representing either the number of arrivals during a time interval or the time interval between successive arrivals. As a consequence, most existing queueing theoretical results *observe and model* the performance of a system without providing tools to design cross-layer algorithms and control strategies that *bring queuing dynamics in the decision process*.

On the other hand, in recent years, distributed optimization and game theoretic techniques have been extensively applied to model and design interactions of decision-makers in distributed networking problems [5]–[10]. However, usually existing (game theoretic) approaches ignore the queueing dynamics in networked systems. This is not without a reason. In fact, in distributed wireless networks (e.g., mobile ad hoc, vehicular, sensor networks) with no predetermined access scheme, users typically compete to access the available spectrum thus creating interference to each other. The resulting (time-varying, frequency- and location-dependent) coupling in the transmission strategies makes the formulation and analysis of the distributed queueing resource allocation problem a complicated and, to the best of our knowledge, unexplored design.

This paper aims at bridging this gap. We consider a queueing interference-limited wireless network wherein users share a common portion of the wireless spectrum (divided into frequency-orthogonal channels). When a new packet arrives, each user decides whether to enqueue the packet, or to transmit it to its destination. Clearly, the decision of each user affects all others. If a node decides to transmit, it generates interference to other users, thus potentially reducing their packet transmission success rate. On the other hand, if the user enqueues the packet, it increases its queuing delay and therefore the probability of dropping packets because of violating a maximum delay deadline. The question that arises is then *how to design distributed algorithms to decide whether to enqueue or transmit in order to achieve high network throughput*.

Main contributions. In this paper, we attack the distributed optimization of users transmission thresholds in queueing networks from two different but complementary perspectives, namely: a user-oriented noncooperative optimization and a holistic-based cooperative design. More specifically, in the first approach, we formulate the aforementioned optimization as a Nash Equilibrium Problem (NEP) wherein the users

Manuscript received March 24, 2014; revised December 23, 2014; accepted February 08, 2015; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor L. Andrew. This work was supported in part by the National Science Foundation under Grant CNS-1218717 and Grant CAREER-1254739, and by the Air Force Research Laboratory under Contract FA8750-14-1-0074. The work of Z. Guan was supported in part by the NSFCunder Grant 61101120. A preliminary shorter version of this paper appeared in the *Proceedings of the IEEE International Conference on Communications*, 2013.

Z. Guan is with the Department of Electrical and Computer Engineering, Northeastern University, Boston, MA 02115 USA, and also with the Department of Electrical Engineering, State University of New York (SUNY) at Buffalo, Buffalo, NY 14260 USA (e-mail: zguan@ece.neu.edu).

T. Melodia is with the Department of Electrical and Computer Engineering, Northeastern University, Boston, MA 02115 USA (e-mail: melodia@ece.neu.edu).

G. Scutari is with the Department of Electrical Engineering, State University of New York (SUNY) at Buffalo, Buffalo, NY 14260 USA (e-mail: gesualdo@buffalo.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNET.2015.2412116

[1]In this work we do not consider closed queueing networks, where packets do not enter the network from outside or depart the network ([4, ch. 7]).

(the players of the game) choose the optimal transmission threshold to maximize their own throughput: Each user decides to transmit if its "best" instantaneous frequency channel is above the optimal threshold value, and enqueue otherwise. The optimal threshold policy of each user depends on the choices of the other users through the multi-user interference. Building on the variational inequality (VI) framework [11], [12] we study the existence and uniqueness of a Nash equilibrium (NE) of the NEP. We then focus on distributed (possibly asynchronous) best-response algorithms solving the game, and derive sufficient conditions guaranteeing their convergence to the NE. Convergence conditions have an intuitive interpretation with significant practical consequences: the algorithm always converges to an equilibrium whenever: 1) the network is highly congested (high values of offered traffic); 2) applications are delay insensitive (i.e., high maximum queueing delay); and 3) the level of interference in the network is not "too high."

The second method we propose consists in formulating the threshold optimization problem as a Network Utility Maximization (NUM), where one maximizes the sum-throughput of the users with respect to the transmission thresholds. To cope with the nonconvexity of the NUM formulation, we build on recent successive convex approximation techniques [13], and propose a distributed pricing-based algorithm that converges to a stationary solution of the NUM while requiring only limited signaling (in the form of message passing) among the users. This second approach and algorithm is more suitable in "collaborative" contexts, where users are willing to exchange some information (albeit limited) in favor of better performance.

Overall, the two algorithms above complement each other in that they can be interpreted as two instances of a unified algorithmic framework: the users solve distributively a sequence of (strongly) convex problems whose objective function is chosen according to the desired level of cooperation and signaling, converging consequently to different types of solutions, namely: 1) a NE of the NEP or 2) a local optimal solution of the NUM. Given the two proposed approaches, noncooperative (game theoretical formulation) versus cooperative (NUM optimization via pricing), the following questions arise naturally: how does the sum-throughput corresponding to the solutions in 1) and 2) compare? How good are those solutions with respect to the globally optimal one of the NUM?

To address these questions, we develop a globally optimal but centralized solution methods for the NUM formulation; the scheme is based on a combination of the branch and bound and convex relaxation techniques. We then use it to benchmark the performance of the proposed noncooperative and cooperative distributed algorithms. Our experiments show that the two approaches, game theoretical and NUM formulations complement each other well: the NE leads to a sum-throughput that is very close to the globally optimal solution of the NUM problem, in many practical scenarios. The performance gap, when present (e.g., in high interference regimes), can be reduced allowing some cooperation among the users in the form of pricing.

In summary the main contributions of the paper are given here.

- We formulate the distributed selection problem of transmission thresholds in a queuing wireless network as NEP and propose distributed (possibly asynchronous) algorithms converging to the NE of the NEP.
- We consider the NUM problem associated to the NEP formulation and design a pricing-based distributed algorithm converging to its local optimal solutions.
- We develop a global solution method for the NUM and benchmark the performance of the proposed cooperative and noncooperative solution methods.

The remainder of this paper is organized as follows. Section II reviews the related work; the system model is described in Section III. Section IV introduces the queueing NEP along with its analysis and design of distributed solution algorithms. The NUM formulation is introduced in Section V, along with our pricing-based and global solution methods. Numerical results are presented in Section VI, where we compare the performance of all the algorithms introduced in the paper. Finally, Section VII draws some conclusions.

II. Related Work

There exists a sizable body of research on queueing-based wireless networks, including queueing in: 1) cellular networks [14], [15]; 2) wireless sensor and *ad hoc* networks [16]–[18]; and 3) cognitive radio networks [19]–[22]. However, to the date, the distributed optimization of users' decision variables (e.g., transmit power, rate allocation, or precoding matrices) over queueing networks based only on local state information remains a difficult and open problem. Good examples of efforts in this direction are [16], [23]–[31] (see also references therein). In these works, the authors designed distributed link scheduling or power control algorithms considering: 1) contention-based interference models [28], [29]; 2) deterministic interference models [16], [24]; or 3) static channel/interference models [25]. Partially distributed queueing schemes for networks with centralized controllers were considered in [30], [31].

Noncooperative game theory has been widely used in the literature to devise distributed decision strategies in wireless networks [5]–[12], [32]–[34]. For example, a game theoretical approach for the linear precoding design of interfering users in a wireless ad-hoc network has been proposed in [9], [10], [32]; [33] extends the design to cognitive radio networks; and in [11], [12], building on the advanced theory of variational inequalities (VIs), the authors developed a unified (asynchronous) best-response-based algorithmic framework for the solution of arbitrary NEPs along with many applications in communications and networking. None of these contributions however has considered queuing dynamics. Game theory has been also applied to wireless networks for queueing analysis [19], [35]–[38]. For example, the work in [35] formulated a multi-class queueing game to study differentiated services; in [36], the authors studied throughput–delay tradeoffs by formulating a M/M/1 queueing game; a dynamic pricing game for uplink wireless random access was introduced in [37]; and [38] formulated the problem of energy management in "sparse" distributed aloha networks as a stochastic evolutionary game.

Differently from the aforementioned works, in this paper we consider: 1) a decentralized wireless networks with arbitrarily user density; 2) a more realistic physical interference model (based on stochastic geometry theory) that captures explicitly the fluctuations of the wireless channel; and 3) *distributed* design of queueing-based interference-limited wireless networks in the aforementioned general setting. To the best of our knowledge, this is the first work merging game-theoretic results with queueing, stochastic interference effects, and distributed control in wireless networks.

III. System Model

We consider a set $\mathcal{N}$ of traffic sessions sharing the same spectrum. For each session, say $n \in \mathcal{N}$, a source-destination pair is identified. Each destination is assumed to be reachable via one-hop by its source node. The spectrum available for communications is divided into a set of $\mathcal{F}$ frequency channels, whereas the transmission time is divided into consecutive time slots. In each time slot, each backlogged source node can either transmit to its destination by selecting a frequency within $\mathcal{F}$ or be silent and enqueue any new packets in its buffer. The proposed model is sufficiently general to subsume several wireless networks, such as: 1) multi-hop wireless ad-hoc networks wherein transmissions are scheduled by a given link scheduling algorithm; 2) interfering multi-cell cellular systems wherein concurrent transmissions scheduled by the base stations (BSs) occur; and 3) device-to-device (D2D) communications in the envisioned 5G cellular networks [39], just to name a few.

We consider a threshold-based transmission policy, according to which each user $n \in \mathcal{N}$: 1) selects the frequency channel with the best channel gain, say f^*, and 2) transmits a packet to its destination over f^* if the corresponding channel gain is higher than a threshold β_n, and does not transmit otherwise. Each user dynamically (and selfishly) adjusts its β_n in order to maximize its expected throughput. High values of β_n lead to low transmission error rates (due to the good quality of the channels), but also reduces opportunities for a node to transmit. This may result in a large queueing delay, and consequently in a high probability that a packet exceed the maximum queueing delay and then be dropped at the receiver side. Conversely, low threshold values are expected to lead to a high transmission error rate, but low packet dropping rate. We show in the next section how to optimize the thresholds β_n's to exploit the optimal tradeoff between transmission and queueing.[2]

Channel model. Let h_n^f denote the channel transfer gain at frequency $f \in \mathcal{F}$ of user $n \in \mathcal{N}$ between its source and destination nodes; h_n^f can be written as $h_n^f = \hat{h}_n \tilde{h}_n^f$, where $\hat{h}_n$ represents the square root of path loss and $\tilde{h}_n^f$ is the channel fading coefficient. Considering a nonsingular path loss model,[3] then we have $\hat{h}_n = \sqrt{(1+l_n^\alpha)^{-1}}$ where l_n is the distance [m] between the transmitter and the receiver of user n, and α represents the path loss factor. We assume a block fading channel model, where each $\tilde{h}_{nm}^f$ is Rayleigh distributed (the model can also be extended to other fading channels, e.g., Rician and Nakagami), with parameter Ω, i.e.,

$$\mathrm{Pb}\left(\tilde{h}_{nm}^f = x\right) = \frac{2x}{\Omega} e^{-\frac{x^2}{\Omega}},\ f \in \mathcal{F}, \quad \forall n, m \in \mathcal{N}. \tag{1}$$

Threshold policy. Similar to [41], [42], we adopt a threshold-based policy. Let $\beta_n > 0$ be the channel fading threshold for user $n \in \mathcal{N}$ and let f^* be the best frequency channel, i.e., $f^* = \arg\max_{f \in \mathcal{F}} \hat{h}_n \tilde{h}_n^f$. Then, according to the threshold policy, user n chooses to transmit a packet over channel f^* if $\tilde{h}_n^{f^*} \geq \beta_n$ and to queue otherwise. Then, according to (1), the probability that $\tilde{h}_n^f$ is lower than the threshold β_n can be expressed as

$$\mathrm{Pb}(\tilde{h}_n^f < \beta_n) = \int_0^{\beta_n} \frac{2x}{\Omega} e^{-\frac{x^2}{\Omega}} dx = 1 - e^{-\frac{\beta_n^2}{\Omega}}. \tag{2}$$

The probability that user $n \in \mathcal{N}$ transmits a packet during a time slot, denoted as $\phi(\beta_n) = 1 - \mathrm{Pb}(\tilde{h}_n^{f^*} < \beta_n)$, is then

$$\phi(\beta_n) = 1 - \left(1 - e^{-\frac{\beta_n^2}{\Omega}}\right)^{|\mathcal{F}|} \tag{3}$$

where $|\mathcal{F}|$ represents the cardinality of the set $\mathcal{F}$.

Queueing model. Let ν_n denote the number of time slots that it takes for node $n \in \mathcal{N}$ to transmit a packet. The probability density function (pdf) of ν_n is

$$\mathrm{Pb}(\nu_n = k) = (1 - \phi(\beta_n))^{k-1} \phi(\beta_n). \tag{4}$$

The expected value of ν_n, denoted by $\mathrm{E}[\nu_n]$, can be represented as $\mathrm{E}[\nu_n] = 1/\phi(\beta_n)$.

To keep our analysis tractable, we approximate the pdf in (4) using an exponential distribution having the same first-order moment of the original pdf. Denoting by $\mu_n(\beta_n)$ the parameter of the exponential distribution function associated with user $n \in \mathcal{N}$, the pdf approximating (4) is

$$\mathrm{Pb}(\nu_n = k) \approx \mu_n(\beta_n) e^{-\mu_n(\beta_n)k} \tag{5}$$

where $\mu_n(\beta_n) = \phi(\beta_n)$, with $\phi(\beta_n)$ defined in (3). The accuracy of the approximated pdf is verified in Fig. 1, where the numerical results show that (5) fits very well the original pdf in (4).

Assuming that the packet arrival rate at each node $n \in \mathcal{N}$ follows a Poisson distribution with average rate λ_n, the queue at node $n \in \mathcal{N}$ can be modeled as an M/M/1 queue, based on the approximation in (5). Denoting by T_n^{th} the maximum queuing delay for user $n \in \mathcal{N}$, the packet loss probability of user n caused by exceeding the maximum queueing delay T_n^{th}, denoted by $P_n^{\mathrm{dly}}(\beta_n)$, is

$$P_n^{\mathrm{dly}}(\beta_n) \triangleq \mathrm{Pb}(T_n > T_n^{\mathrm{th}}) = e^{-\left(\frac{\mu_n(\beta_n)}{T_{\mathrm{slt}}} - \lambda_n\right) T_n^{\mathrm{th}}} \tag{6}$$

where T_{slt} is the time duration of a time slot. Since $P_n^{\mathrm{dly}}(\beta_n)$ in (6) can not be greater than 1, it must be

$$\frac{\mu_n(\beta_n)}{T_{\mathrm{slt}}} - \lambda_n \geq 0 \tag{7}$$

[2]Other physical- and MAC-layer functions can also be taken into consideration in the queueing game, including power control, Automatic Repeat-reQuest (ARQ), among others. This will result in queueing games with multiple strategy variables for each user. In this work, we focus on optimizing the transmission threshold to keep the theoretical modeling and analysis tractable.

[3]In the nonsingular path loss model, the path loss tends to one as propagation distance tends to zero, whereas in singular path model, the path loss tends to infinity [40].

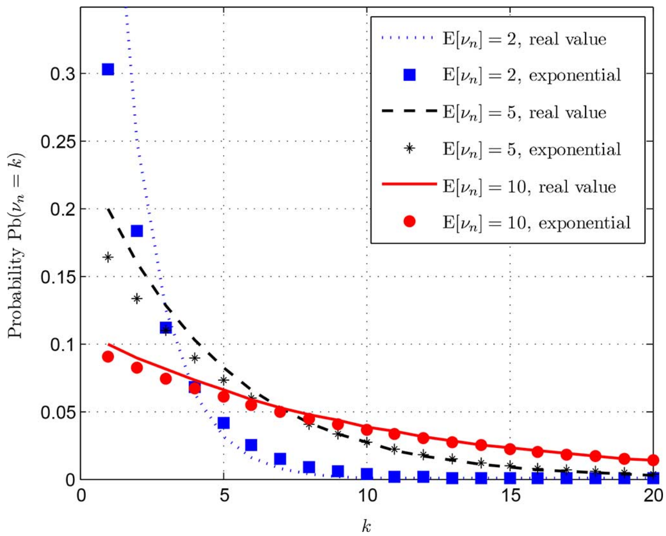


Fig. 1. Approximate the pdf of ν_n in (4) using exponential distribution function.

with $\mu_n(\beta_n) = \phi(\beta_n)$ defined in (3). This leads to the following upper bound on β_n:

$$\beta_n \leq \beta_n^{\mathrm{U}}, \qquad n \in \mathcal{N} \tag{8}$$

with

$$\beta_n^{\mathrm{U}} = \sqrt{-\ln\left[1 - (1 - T_{\mathrm{slt}}\lambda_n)^{\frac{1}{|\mathcal{F}|}}\right]\Omega}. \tag{9}$$

Interference model. If multiple users select the same frequency channel, they will interfere with each other. Letting γ_{th} be the threshold on the signal-to-noise-plus-interference ratio (SINR) γ_n above which a packet can be correctly decoded at a receiver, then the probability that a transmission error occurs for user $n \in \mathcal{N}$ is

$$\mathrm{Pb}(\gamma_n < \gamma_{\mathrm{th}}) = \mathrm{Pb}\left(\frac{P_n \hat{h}_n^2 \left(\tilde{h}_n^f\right)^2}{\left(\sigma_n^f\right)^2 + I_n^f(\boldsymbol{\beta}_{-n})} < \gamma_{\mathrm{th}}\right) \tag{10}$$

where P_n is the transmission power of user $n \in \mathcal{N}$, $(\sigma_n^f)^2$ is the power of the Gaussian noise, and $I_n^f(\boldsymbol{\beta}_{-n})$ represents the interference at the destination node of user n on the carrier $f \in \mathcal{F}$, which depends on the transmission thresholds of all the other users, i.e., $\boldsymbol{\beta}_{-n} \triangleq (\beta_m)_{m \in \mathcal{N}/n}$. In (10), the channel fading $\tilde{h}_n^f$ follows a Rayleigh distribution [cf. (1)] and takes values in $[\beta_n + \infty)$, where β_n is the transmission threshold of user n. The interference $I_n^f(\boldsymbol{\beta}_{-n})$ in (10) can be then expressed as

$$I_n^f(\boldsymbol{\beta}_{-n}) = \sum_{m \in \mathcal{N}/n} P_m \hat{h}_{mn}^2 (\tilde{h}_{mn}^f)^2 \alpha_m^f(\beta_m) \tag{11}$$

where $\hat{h}_{mn}^2$ and $(\tilde{h}_{mn}^f)^2$ represent the path loss and the square of channel fading between the source node $m \in \mathcal{N}$ and the destination node $n \in \mathcal{N}$, respectively; the value of $\alpha_m^f(\beta_m)$ is equal to one if user m transmits, and zero otherwise. Therefore, the aggregate interference measured at the receiver node of user n depends on: 1) the *locations of the source nodes* of all other sessions; 2) *whether each source transmits or not*; and 3) *which channel* each source uses for transmission.

We model the distribution of $I_n^f(\boldsymbol{\beta}_{-n})$ in (11) following a classical stochastic geometry approach [43]. More specifically, the node distribution in the network is assumed to follow a bidimensional Poisson Point Process (PPP). This is a well-accepted model of static ad-hoc networks with random deployment as well as networks with moving users. In this setting, the pdf of $I_n^f(\boldsymbol{\beta}_{-n})$ in (11) is the Gamma distribution function $\zeta_n(x)$, i.e.,

$$\begin{aligned}\zeta_n(x) &\triangleq \mathrm{Pb}\left(I_n^f(\boldsymbol{\beta}_{-n}) = x\right)\\ &= x^{k_n(\boldsymbol{\beta}_{-n})-1} \frac{e^{-x/\theta_n(\boldsymbol{\beta}_{-n})}}{\Gamma\left(k_n(\boldsymbol{\beta}_{-n})\right) \theta_n^{k_n(\boldsymbol{\beta}_{-n})}(\boldsymbol{\beta}_{-n})}\end{aligned} \tag{12}$$

where $k_n(\boldsymbol{\beta}_{-n})$ and $\theta_n(\boldsymbol{\beta}_{-n})$ are the shaping parameters of the Gamma function, and both are functions of $\boldsymbol{\beta}_{-n}$ (see Appendix A for the explicit functional expressions), and $\Gamma(k_n)$ in (12) is given by

$$\Gamma(k_n(\boldsymbol{\beta}_{-n})) = \int_0^\infty x^{k_n(\boldsymbol{\beta}_{-n})-1} e^{-x} dx. \tag{13}$$

Let $\vartheta_n(x, \boldsymbol{\beta}_{-n})$ be the complementary cumulative distribution function of $I_n^f(\boldsymbol{\beta}_{-n})$. Then, according to (12), $\vartheta_n(x, \boldsymbol{\beta}_{-n})$ can be written as

$$\begin{aligned}\vartheta_n(x, \boldsymbol{\beta}_{-n}) &\triangleq \mathrm{Pb}\left(I_n^f(\boldsymbol{\beta}_{-n}) > x\right)\\ &= 1 - \frac{\varphi\left(k_n(\boldsymbol{\beta}_{-n}), \frac{x}{\theta_n(\boldsymbol{\beta}_{-n})}\right)}{\Gamma\left(k_n(\boldsymbol{\beta}_{-n})\right)}\end{aligned} \tag{14}$$

where $\Gamma\left(k_n(\boldsymbol{\beta}_{-n})\right)$ is defined in (13), and $\varphi\left(k_n(\boldsymbol{\beta}_{-n}), \frac{x}{\theta_n(\boldsymbol{\beta}_{-n})}\right)$ is the incomplete gamma function given by

$$\varphi\left(k_n(\boldsymbol{\beta}_{-n}), \frac{x}{\theta_n(\boldsymbol{\beta}_{-n})}\right) = \int_0^{\frac{x}{\theta_n(\boldsymbol{\beta}_{-n})}} s^{k_n(\boldsymbol{\beta}_{-n})-1} e^{-s} ds. \tag{15}$$

Under the above assumptions, the packet loss probability is given by

$$\begin{aligned}P_n^{\mathrm{err}}(\boldsymbol{\beta}) &\triangleq \mathrm{Pb}(\gamma_n < \gamma_{\mathrm{th}})\\ &= \int_{\beta_n}^\infty \frac{2x}{\Omega} e^{-\frac{x^2}{\Omega}} \vartheta_n\left(\frac{P_n \hat{h}_n^2}{\gamma_{\mathrm{th}}} x^2 - (\sigma_n^f)^2, \boldsymbol{\beta}_{-n}\right) dx\end{aligned} \tag{16}$$

with $\vartheta_n(\cdot,\cdot)$ defined in (14), and $\boldsymbol{\beta} \triangleq (\beta_n)_{n \in \mathcal{N}}$.

Expected throughput. Using the expressions of the packet loss probability $P_n^{\mathrm{dly}}(\beta_n)$ and the transmission error rate $P_n^{\mathrm{err}}(\boldsymbol{\beta})$ of user n as given in (6) and (16), respectively, we can now introduce the overall loss rate $P_n^{\mathrm{los}}(\boldsymbol{\beta})$ and the expected throughput $R_n(\boldsymbol{\beta})$ of each user n, given by

$$P_n^{\mathrm{los}}(\boldsymbol{\beta}) = P_n^{\mathrm{dly}}(\beta_n) + \left[1 - P_n^{\mathrm{dly}}(\beta_n)\right] P_n^{\mathrm{err}}(\boldsymbol{\beta}) \tag{17}$$

$$R_n(\boldsymbol{\beta}) = \lambda_n \left(1 - P_n^{\mathrm{dly}}(\beta_n) - \left[1 - P_n^{\mathrm{dly}}(\beta_n)\right] P_n^{\mathrm{err}}(\boldsymbol{\beta})\right) \tag{18}$$

where λ_n is the average incoming packet rate of user n. Note that $R_n(\boldsymbol{\beta})$ is a function of $\boldsymbol{\beta}$ through $P_n^{\rm dly}(\beta_n)$ and $P_n^{\rm err}(\boldsymbol{\beta})$.

We now approximate the throughput $R_n(\boldsymbol{\beta})$ in (18) by neglecting the second-order term $P_n^{\rm dly}(\beta_n)P_n^{\rm err}(\boldsymbol{\beta})$, which is acceptable when the overall packet loss rate is low or moderate. For example, if $P_n^{\rm dly}(\beta_n) = 0.1$ and $P_n^{\rm err}(\boldsymbol{\beta}) = 0.1$ (which are still very high values), the overall packet loss rate is 0.19 and the approximation is 0.20. The resulting approximation error is only 0.01. The approximation error is 0.1 when $P_n^{\rm dly}(\beta_n)$ and $P_n^{\rm err}(\boldsymbol{\beta})$ take a value of 0.2. Based on the approximation, the throughput of user $n \in \mathcal{N}$ can be simplified as

$$R_n(\boldsymbol{\beta}) = \lambda_n \left[1 - P_n^{\rm dly}(\beta_n) - P_n^{\rm err}(\boldsymbol{\beta})\right]. \qquad (19)$$

IV. Distributed Design via Game Theory

Aiming at finding distributed low-complex algorithms for the network design, in this section, we cast the joint optimization of the channel thresholds into a game theoretical formulation, where the users are able to self-enforce the negotiated agreements on use of the available spectrum without the intervention of a centralized authority. More specifically, we consider a NEP wherein players are the users and the payoff function of each user is the achievable throughput $R_n(\boldsymbol{\beta})$; each player $n \in \mathcal{N}$ competes against the others by choosing the channel threshold β_n that maximizes $R_n(\boldsymbol{\beta})$, i.e.,

$$\mathcal{G}: \begin{array}{ll} \max\limits_{\beta_n} & R_n(\beta_n, \boldsymbol{\beta}_{-n}) \\ \text{s.t.} & \beta_n^{\rm L} \le \beta_n \le \beta_n^{\rm U} \end{array} \qquad \forall n \in \mathcal{N} \qquad (20)$$

where $\beta_n^{\rm U}$ is the upper bound of β_n as given in (9), $\beta_n^{\rm L}$ is a given lower bound (we discuss shortly how to choose this lower bound), and $\Phi = \prod_{n\in\mathcal{N}} \Phi_n$ with $\Phi_n = [\beta_n^{\rm L}\ \beta_n^{\rm U}]$ denotes the joint strategy set of the game. We will refer to the game (20) as $\mathcal{G}$.

Definition 1 (Nash Equilibrium): A channel fading threshold vector $\boldsymbol{\beta}^* = (\beta_n^*)_{n\in\mathcal{N}}$ is a NE of $\mathcal{G}$ if

$$\beta_n^* = \mathop{\arg\max}_{\beta_n \in [\beta_n^{\rm L}\ \beta_n^{\rm U}]} R_n\left(\beta_n, \boldsymbol{\beta}_{-n}^*\right), \quad \forall n \in \mathcal{N}. \qquad (21)$$

In words, at a NE of the game no user has incentive to modify its channel fading threshold, given the optimal threshold values of the other users. Note that a NEP may not have a NE, even if the single user optimization problems have a unique solution. The solution analysis of the game $\mathcal{G}$ is given in the following theorem.

Theorem 1 (Existence of NE): Given the NEP $\mathcal{G}$, suppose that $\beta_n^{\rm L} \ge \frac{\sqrt{2}}{2}\Omega$ for all $n \in \mathcal{N}$, where Ω is the Rayleigh fading factor. Then, the NEP has a NE.

Proof: The core of the proof is to show that the payoff function of each user, i.e., $R(\boldsymbol{\beta})$, is concave with respect to its own transmission threshold β_n, see Appendix B. ■

According to Theorem 1, a NE always exists if $\beta_n^{\rm L} \ge \frac{\sqrt{2}}{2}\Omega$ for all $n \in \mathcal{N}$. Quite interestingly, this condition has a practical interpretation/implication. Based on the Rayleigh fading distribution (1), in a single-channel network, user n transmits a packet in a given time-slot with probability around 0.6 if it chooses a threshold $\beta_n = \beta_n^{\rm L}$. This probability increases to 0.975 when there are four frequency channels. In practical wireless networks (e.g., IEEE 802.11 based wireless networks [44]), the available frequency channels is much larger than 4, implying that the corresponding user's transmission probability is very close to 1, which is equivalent to setting $\beta_n^{\rm L}$ to zero. Therefore, in the following and without loss of generality we set $\beta_n^{\rm L} = \frac{\sqrt{2}}{2}\Omega$ for all $n \in \mathcal{N}$.

A. Distributed Best-Response Algorithms

We focus now on distributed algorithms to compute a NE of the game $\mathcal{G}$. We consider the class of best-response Jacobi iterative schemes: at each iteration, all the users solve *in parallel* their own optimization problems (20) (given the strategy profile of the others at the previous iteration). Extensions to *totally asynchronous* (in the sense of [10], [11]) are also considered. The formal description of the algorithm is given in Algorithm 1, and its convergence properties are stated in Theorem 2.

Algorithm 1: Jacobi Best-response Algorithm

Data: $\boldsymbol{\beta}^0 \in \Phi$. Set $\kappa = 0$.

(S.1) : If $\boldsymbol{\beta}^\kappa$ satisfies a suitable termination criterion: STOP;

(S.2) : For all $n \in \mathcal{N}$, compute in parallel

$$\beta_n^{\kappa+1} \in \mathop{\arg\max}_{\beta_n\in\Phi_n} R_n\left(\beta_n, \boldsymbol{\beta}_{-n}^\kappa\right). \qquad (22)$$

(S.3) : $\kappa \leftarrow \kappa + 1$ and go to (S.1).

Theorem 2: Given the game $\mathcal{G}$ in the setting of Theorem 1, suppose that the following conditions are satisfied:

$$\min_{\boldsymbol{\beta}\in\Phi} \frac{\partial^2 P_n^{\rm los}(\boldsymbol{\beta})}{\partial \beta_n^2} > \sum_{m\in\mathcal{N}/n} \max_{\boldsymbol{\beta}\in\Phi} \frac{\partial^2 P_m^{\rm los}(\boldsymbol{\beta})}{\partial\beta_m \partial\beta_n}, \quad \forall n \in \mathcal{N}. \qquad (23)$$

Then, the sequence generated by Algorithm 1 converges to the unique NE of the game.

Proof: The theorem can be proven by constructing an auxiliary $|\mathcal{N}| \times |\mathcal{N}|$ matrix and verifying its P-matrix property based on the recent results in ([11], Proposition 3). See Appendix C for details. ■

On the convergence condition. Interestingly, convergence conditions (23) have an intuitive interpretation, as described next. The explicit computation of the derivatives in (23) (see Appendix B) leads to the following inequality

$$\frac{\partial^2 P_n^{\rm los}(\boldsymbol{\beta})}{\partial\beta_n^2} \ge \frac{\partial B(\boldsymbol{\beta})}{\partial\beta_n} \ge |\mathcal{F}| \frac{T_n^{\rm th}}{T_{\rm slt}} e^{\lambda_n T_n^{\rm th}} c_n, \quad \forall n \in \mathcal{N} \qquad (24)$$

where $B(\boldsymbol{\beta})$ is defined in (45) (cf. Appendix B) and $c_n > 0$ is a positive constant whose explicit expression is irrelevant for our discussion; and $\frac{\partial^2 P_m^{\rm los}(\boldsymbol{\beta})}{\partial\beta_m\partial\beta_n}$ is upper bounded for all $m, n \in \mathcal{N}$. Based on (24), sufficient conditions in (23) are satisfied in the following scenarios.

Congestion-dominated wireless network. This setting corresponds to large incoming average packet rates λ_n. Indeed, the left-hand side of (23) monotonically increases with λ_n, implying that condition (23) tends to be satisfied in networks with large λ_n, where the users tend to choose low transmission thresholds β_n (they transmit more often to avoid high packet loss rate due to exceeding the the maximum queueing delay $T_n^{\rm th}$).

Loosely coupled interference. This corresponds to sparse wireless networks (e.g., due to low density of nodes or large number of available channels). In this setting, the right-hand side of (23) tends to be "very small" (cf. Appendix B), and as a result the convergence condition is satisfied.

Delay-insensitive wireless networks. In this setting, we have "large" maximum tolerable queueing delays $T_n^{\rm th}$. Indeed, when $T_n^{\rm th}$s are "large", the left-hand side of (23) tends to be "large" too (cf. Appendix B), and consequently this makes the convergence conditions more likely to be satisfied. In this scenario, each user will choose to enqueue its data more often and transmit only when its channel quality is very good, resulting in low generated interference.

Implementation issue. To compute the optimal solution of its optimization problem, every user needs to estimate the parameters $k_n(\boldsymbol{\beta}_{-n})$ and $\theta_n(\boldsymbol{\beta}_{-n})$, which are both functions on $\boldsymbol{\beta}_{-n}$. Interestingly, this can be done locally without any exchange of information among the users. Indeed, each user n only needs to measure at his receiver the interference experienced in a number of consecutive time slots and use it to estimate the mean and the variance of the interference. Given these estimates, he can then use (43) and (44) given in Appendix A to obtain the desired expressions of $k_n(\boldsymbol{\beta}_{-n})$ and $\theta_n(\boldsymbol{\beta}_{-n})$.

Asynchronous updates. Note that the convergence conditions (23) do not depend on the specific updating order performed by the users. In fact, one can prove (see [12]) that these conditions guarantee also convergence of best-response algorithms under *totally asynchronous* updates of the player strategies (in the sense of [12]). In such asynchronous schemes, some users may change their strategies more frequently than the others (e.g., at random times) and they may even use an outdated information of the other users strategies, without affecting the convergence of the algorithm. It turns out that instances of the aforementioned asynchronous framework are robust against missing or outdated updates of the users. This feature strongly relaxes the constraints on the network synchronization; which makes the proposed class of algorithms truly appealing in many practical scenarios.

V. NUM-Based Design

Here, we focus on a complementary approach to the noncooperative NEP formulation; we cast the system design into a NUM problem wherein the users cooperate to maximize the sum-throughput in the network. More formally, we have the following:

$$\begin{array}{lll} \displaystyle\max_{\boldsymbol{\beta}\triangleq(\beta_n)_{n\in\mathcal{N}}} & R(\boldsymbol{\beta}) \triangleq \displaystyle\sum_{n\in\mathcal{N}} R_n(\boldsymbol{\beta}) & \\ \text{s.t.} & \beta_n^{\rm L} \leq \beta_n \leq \beta_n^{\rm U} & \forall n \in \mathcal{N}. \end{array} \tag{25}$$

The above optimization problem is nonconvex. In what follows, we will exploit the structure of (25) and, building on some recent SCA techniques introduced in [13], [45], we develop a fast (almost) distributed pricing-based algorithm converging to a local optimal solution of (25) (cf. Section V-A).[4] As benchmark, we then focus on global solution but centralized schemes for (25) (cf. Section V-B).

A. SCA Pricing-Based Algorithm

Traditionally nonconvex optimization problems in the form (25) have been tackled using gradient-based algorithms (when the goal is to compute local optimal solutions via low-complexity methods), which solve a sequence of convex problems where the social function $R(\boldsymbol{\beta})$ is replaced by its first (or second) order Taylor approximation. Because of that, however, those methods suffer from slow convergence. A faster still parallel algorithm can be obtained building on the idea introduced recently in [13]: since each $R_n(\boldsymbol{\beta}_n, \boldsymbol{\beta}_{-n})$ is concave in $\boldsymbol{\beta}_n$ for any given $\boldsymbol{\beta}_{-n}$ (in the setting of Theorem 1), one can convexify only the nonconcave part in $R(\boldsymbol{\beta})$, which is $\sum_{j\in\mathcal{N}\setminus\{n\}} R_j(\boldsymbol{\beta})$, and solve the sequence of resulting optimization problems, one for each user. Since such a procedure preserves some structure of the sum-utility function, it is expected to be faster than standard gradient-based algorithms while keeping the same parallel and distributed nature, a fact that is supported by our numerical experiments (see Section VI). We provide a formal description of the proposed SCA algorithm next.

Given the users' strategy profile $\boldsymbol{\beta}$, let us define for each user n the pricing quantity $\pi_n(\boldsymbol{\beta})$ as

$$\pi_n(\boldsymbol{\beta}) = \sum_{m\in\mathcal{N}/n} \nabla_{\beta_n} R_m(\boldsymbol{\beta}) \tag{26}$$

which represents the linearization of the nonconcave part of $R(\boldsymbol{\beta})$ with respect to $\boldsymbol{\beta}_n$. We can then introduce the following best-response dynamic for each user n: given the users' strategy profile $\boldsymbol{\beta}^\nu$ at iteration $\nu \geq 0$, let

$$\begin{aligned} \widehat{\beta}_n(\boldsymbol{\beta}^\nu) \triangleq \underset{\beta_n\in[\beta_n^{\rm L},\beta_n^{\rm U}]}{\arg\max} \Big\{ & R_n(\beta_n, \boldsymbol{\beta}_{-n}^\nu) - \pi_n(\boldsymbol{\beta}^\nu)\left(\beta_n - \beta_n^\nu\right) \\ & - \frac{\tau_n}{2}\left(\beta_n - \beta_n^\nu\right)^2 \Big\}, \end{aligned} \tag{27}$$

where τ_n is any arbitrary positive constant. Note that under the setting of Theorem 1, each optimization problem (27) is strongly convex and thus $\widehat{\beta}_n(\boldsymbol{\beta}^\nu)$, the unique solution of (27), is well defined.

We are now ready to introduce the proposed pricing-based algorithm, which roughly speaking consists in solving in parallel the subproblems (27), starting from a feasible $\boldsymbol{\beta}^0$. The formal description of the algorithm is given in Algorithm 2, where in Step 3 we also allow the use of a memory in the update of the threshold variables; the convergence properties of the algorithm are stated in Theorem 3, whose proof follows from Theorem 1 and ([13], Theorem 1) and thus is omitted.

Algorithm 2: Pricing Jacobi Algorithm

Data: $\boldsymbol{\tau} \triangleq (\tau_n)_{n\in\mathcal{N}} > \mathbf{0}, \{\eta^\nu\} > 0$. Set $\nu = 0$.

[4]Different SCA schemes have also been proposed to solve nonconvex optimization problems, including [46]–[50]. The schemes in these papers focus on utility functions that are convex in some optimization variables and concave in the others, and hence can not be applied to our problem where the sum-rate utility function does not have such a convex-concave structure.

(S.1) : If $\boldsymbol{\beta}^{\nu}$ satisfies a suitable termination criterion: STOP;

(S.2) : For all $n \in \mathcal{N}$, compute in parallel $\widehat{\beta}_n(\boldsymbol{\beta}^{\nu})$ [cf. (27)];

(S.3) : Set $\beta_n^{\nu+1} = \beta_n^{\nu} + \eta^{\nu}\left(\widehat{\beta}_n(\boldsymbol{\beta}^{\nu}) - \beta_n^{\nu}\right)$, for all $n \in \mathcal{N}$;

(S.4) : $\nu \leftarrow \nu + 1$ and go to (S.1).

Theorem 3: Given the NUM problem (25) under the setting of Theorem 1, suppose that the step-size sequence $\{\eta^{\nu}\}$ is chosen so that

$$\eta^{\nu} \in (0,1], \eta^{\nu} \to 0 \text{ and } \sum_{\nu} \eta^{\nu} = +\infty. \tag{28}$$

Then, either Algorithm 2 converges in a finite number of iterations to a stationary solution of the social problem, or every limit point of the sequence $\{\boldsymbol{\beta}^{\nu}\}$ (at least one of such points exists) is a stationary solution of the social problem. Moreover, none of such points is a local minimum of the social function. □

On Algorithm 2. The algorithm implements in a distributed way a pricing mechanism; each user n maximizes iteratively its own rate minus a pricing term that measures somehow the marginal increase of the sum-utility of the other users due to a variation of the strategy of user n. Roughly speaking, the pricing works like a punishment imposed to each user for being too aggressive in choosing its own strategy and thus "hurting" the other users. In fact, the presence of this pricing is what drives the system toward a stationary point of the NUM problem rather than a NE of the game $\mathcal{G}$ (which happens instead when $\pi_n = 0$ for all n, see Algorithm 1). Differently from Algorithm 1, Algorithm 2 is however not incentive compatible, in the sense that the users need to reach an agreement in following the best-response dynamics (27); therefore it has to be imposed as network protocol. Moreover, some signaling among the users is required to compute locally at each iterations the prices $\pi_n(\boldsymbol{\beta}^{\nu})$. However, Algorithm 2 convergences under consistently milder conditions on the network parameters that those required by Algorithm 1 (Theorem 2 versus Theorem 3), at the price of more (albeit limited) signaling.

As a final remark, note that conditions on the choice of the step-size sequence $\{\eta^{\nu}\}$ as in (28) are relatively weak; for instance all the step-size rules using in diminishing gradient-like schemes can be used here. In our experiments we observed that two effective rules are [13]: given $\eta^0 = 1$, we have

$$\text{Rule1} : \eta^{\nu} = \eta^{\nu-1}(1 - \epsilon\,\eta^{\nu-1}), \qquad \nu = 1, \ldots \tag{29}$$

$$\text{Rule2} : \eta^{\nu} = \frac{\eta^{\nu-1} + \beta_1}{1 + \beta_2 \nu}, \qquad \nu = 1, \ldots \tag{30}$$

where $\epsilon \in (0,1)$ and $\beta_1, \beta_2 \in (0,1)$ are given constants such that $\beta_1 \le \beta_2$.

B. Global Solution Centralized Method

Here, we address the issue of quantifying how good are the solutions obtained by Algorithms 1 and 2 in terms of sum-throughput. To provide an answer to this question, we develop a global solution method for the NUM problem (25) that is based on a combination of the *Branch-and-Bound* (BB) framework and *convex relaxation* techniques [51].

Algorithm design. Define $R^{\star}$ as the globally optimal sum-throughput objective value of (25), i.e.,

$$R^{\star} \triangleq \max_{\boldsymbol{\beta} \in \Phi} R(\boldsymbol{\beta}) \tag{31}$$

and denote $\boldsymbol{\beta}^{\star} \in \Phi$ as the optimal transmission profile corresponding to $R^{\star}$. Then, given the accuracy $\epsilon \in [0,1]$, the proposed algorithm seeks an ϵ-optimal solution, i.e., a feasible profile $\overline{\boldsymbol{\beta}}$ such that $R(\overline{\boldsymbol{\beta}}) \ge \epsilon R^{\star}$.

Let Φ^0 be the power set of the feasible set Φ of (25). The proposed algorithm generates a partition $\mathcal{P}_{\Phi} \triangleq \{\Phi^{i+1} \subseteq \Phi^i \subseteq \cdots \subseteq \Phi^0, i = 1, 2, \ldots\}$ of Φ^0, where Φ^i is obtained using the subroutine vrb_pat (a suitable variable partition procedure) described later on. For any Φ^i, we compute local upper $\mathrm{UP}(\Phi^i)$ and lower $\mathrm{LR}(\Phi^i)$ bounds of the sum-throughput $R(\boldsymbol{\beta})$ over Φ^i, solving convex subproblems; $\mathrm{UP}(\Phi^i)$ is obtained solving a convex optimization problem in cvx_rlx, whereas $\mathrm{LR}(\Phi^i)$ is computed as described in lcl_sch. Then, a global upper $\mathrm{UP}_{\mathrm{glb}}(\Phi^0)$ and lower $\mathrm{LR}_{\mathrm{glb}}(\Phi^0)$ bound over Φ^0 can be maintained as

$$\mathrm{UP}_{\mathrm{glb}}(\Phi^0) \triangleq \max_{\Phi^i \in \mathcal{P}_{\Phi}} \mathrm{UP}(\Phi^i) \tag{32}$$

$$\mathrm{LR}_{\mathrm{glb}}(\Phi^0) \triangleq \max_{\Phi^i \in \mathcal{P}_{\Phi}} \mathrm{LR}(\Phi^i) \tag{33}$$

respectively. A formal description of the algorithm is given in Algorithm 3, where subroutines cvx_rlx, lcl_sch and vrb_pat are described below.

Algorithm 3: Globally Optimal Solution Algorithm

Data: $\epsilon \in [0,1], \mathcal{P}_{\Phi} = \Phi^0$; set $\Phi^{\mathrm{sel}} = \Phi^0, \mathrm{UP}_{\mathrm{glb}} = \infty$, and $\mathrm{LR}_{\mathrm{glb}} = 0$.

(S.1) : Calculate $\mathrm{UP}(\Phi^{\mathrm{sel}})$ using cvx_rlx;

Calculate $\mathrm{LR}(\Phi^{\mathrm{sel}})$ using lcl_sch;

Set $\mathrm{UP}_{\mathrm{glb}} = \mathrm{UP}(\Phi^{\mathrm{sel}}), \mathrm{LR}_{\mathrm{glb}} = \mathrm{LR}(\Phi^{\mathrm{sel}})$.

(S.2) : if $\mathrm{LR}_{\mathrm{glb}} \ge \epsilon \mathrm{UP}_{\mathrm{glb}}$, STOP;

(S.3) : Select $\Phi^{\mathrm{sel}} \in \mathcal{P}_{\Phi}$ and partition it into $\Phi^{\mathrm{sel},1}$ and $\Phi^{\mathrm{sel},2}$ using vrb_pat.

(S.4) : For $j = 1, 2$,

Calculate $\mathrm{UP}(\Phi^{\mathrm{sel},j})$ using cvx_rlx;

Calculate $\mathrm{LR}(\Phi^{\mathrm{sel},j})$ using lcl_sch;

Set $\mathrm{LR}_{\mathrm{glb}} = \max(\mathrm{LR}_{\mathrm{glb}}, \mathrm{LR}(\Phi^{\mathrm{sel},j}))$;

if $\mathrm{UP}(\Phi^{\mathrm{sel},j}) \ge \mathrm{LR}_{\mathrm{glb}}$, set

$\mathcal{P}_{\Phi} \leftarrow \mathcal{P}_{\Phi} \bigcup \Phi^{\mathrm{sel},j}$.

(S.5) : Set $\mathrm{UP}_{\mathrm{glb}} = \max_{\Phi \in \mathcal{P}_{\Phi}} (\mathrm{UP}(\Phi))$; go to (S.2).

Convex Relaxation (cvx_rlx):

$$\begin{aligned}&\text{Input}:\Phi^{\mathrm{sel},j}\\ &\text{Output}:\mathrm{UP}(\Phi^{\mathrm{sel},j})\triangleq\max_{\boldsymbol{\beta}\in\Phi^{\mathrm{sel},j}}\sum_{n\in\mathcal{N}}\overline{R}_n(\boldsymbol{\beta})\\ &=\sum_{n\in\mathcal{N}}\overline{R}_n(\boldsymbol{\beta})\lambda_n\left(1-P_n^{\mathrm{dly}}(\beta_n)\right.\\ &\left.-\left[1-P_n^{\mathrm{dly}}\left(\underline{\beta}_n^{\mathrm{sel},j}\right)\right]P_n^{\mathrm{err}}\left(\overline{\boldsymbol{\beta}}_{-n}^{\mathrm{sel},j}\right)\right)\end{aligned}\tag{34}$$

where $\overline{\boldsymbol{\beta}}_{-n}^{\mathrm{sel},j}\triangleq(\overline{\beta}_m^{\mathrm{sel},j})_{m\in\mathcal{N}/n}$ with $[\underline{\beta}_n^{\mathrm{sel},j},\ \overline{\beta}_n^{\mathrm{sel},j}]$ being the range of transmission threshold β_n given current sub-domain $\Phi^{\mathrm{sel},j}$.

Remark: The rationale behind the relaxation is as follows. To relax the objective function in (25) to be convex, we only need to relax the individual utility function of each user. Based on the proof of Theorem 1, it can be proven that $P_n^{\mathrm{dly}}(\beta_n)$ (which represents the packet loss rate of user $n\in\mathcal{N}$ caused by exceeding the maximum delay time T_n^{th}) is a convex function of β_n. Therefore, we only need to relax the term $[1-P_n^{\mathrm{dly}}(\beta_n)]P_n^{\mathrm{err}}(\boldsymbol{\beta})$ in (18) to be convex. For this purpose, we adopt a simple but effective relaxation method, observing that i) $P_n^{\mathrm{dly}}(\beta_n)$ monotonically increases with β_n, and ii) $P_n^{\mathrm{err}}(\boldsymbol{\beta})$ monotonically decreases with $\boldsymbol{\beta}_{-n}$. Then, given $\Phi^{\mathrm{sel},j}$, $\mathrm{UP}(\Phi^{\mathrm{sel},j})$ obtained by solving the convex optimization in (34) provides a suitable relaxation of the original sum-throughput objective function. Denote the solution of the relaxed maximization problem (34) as $\boldsymbol{\beta}^{\mathrm{sel},j}\triangleq(\beta_n^{\mathrm{sel},j})_{n\in\mathcal{N}}$.

Local Search (lcl_sch):

$$\begin{aligned}&\text{Input}:\boldsymbol{\beta}^{\mathrm{sel},j}\\ &\text{Output}:\mathrm{LR}(\Phi^{\mathrm{sel},j})\triangleq R(\boldsymbol{\beta}^{\mathrm{sel},j}).\end{aligned}\tag{35}$$

Variable Partition (vrb_pat):

$$\begin{aligned}&\text{Input}:\Phi^{\mathrm{sel}}\triangleq\left\{(\beta_n^{i^\star})_{n\in\mathcal{N}}\right\}\ \text{with}\\ &i^\star=\arg\max_{\Phi_i\in\mathcal{P}_\Phi}\ \mathrm{UP}(\Phi_i)\end{aligned}\tag{36}$$

$$\begin{aligned}\text{Output}:\Phi^{\mathrm{sel},1}\triangleq\Bigg\{&\left(\beta_n^{\mathrm{sel},1}\right)_{n\in\mathcal{N}}\Bigg|\,\underline{\beta}_n^{i^\star}\le\beta_n^{\mathrm{sel},1}\\ &\le\overline{\beta}_n^{i^\star},\quad\forall n\in\mathcal{N}/n^\star,\underline{\beta}_{n^\star}^{i^\star}\le\beta_{n^\star}^{\mathrm{sel},1}\le\frac{\underline{\beta}_{n^\star}^{i^\star}+\overline{\beta}_{n^\star}^{i^\star}}{2}\Bigg\}\end{aligned}\tag{37}$$

$$\begin{aligned}\Phi^{\mathrm{sel},2}\triangleq\Bigg\{&\left(\beta_n^{\mathrm{sel},2}\right)_{n\in\mathcal{N}}\Bigg|\,\underline{\beta}_n^{i^\star}\le\beta_n^{\mathrm{sel},2}\\ &\le\overline{\beta}_n^{i^\star},\quad\forall n\in\mathcal{N}/n^\star,\frac{\underline{\beta}_{n^\star}^{i^\star}+\overline{\beta}_{n^\star}^{i^\star}}{2}\le\beta_{n^\star}^{\mathrm{sel},2}\le\overline{\beta}_{n^\star}^{i^\star}\Bigg\}\end{aligned}\tag{38}$$

with

$$n^\star=\arg\max_{n\in\mathcal{N}}\left(\overline{\beta}_n^{i^\star}-\underline{\beta}_n^{i^\star}\right).\tag{39}$$

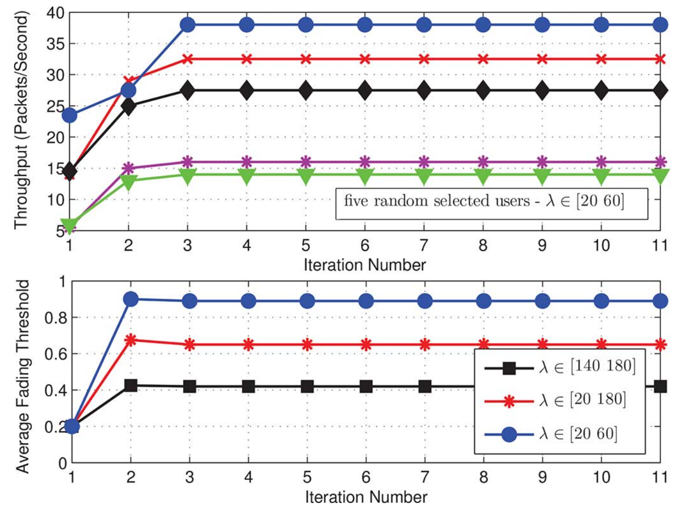


Fig. 2. Convergence of Algorithm 1. Achievable throughput (top) and average Rayleigh fading threshold (bottom) versus iterations.

Remark: Here, we select sub-domain $\Phi^{\mathrm{sel}}\in\mathcal{P}_\Phi$ corresponding to the largest local upper bound, and then partition the variable in it with the largest range from the middle; $[\underline{\beta}_n^{i^\star}\ \overline{\beta}_n^{i^\star}]$ represents the range of variable $\beta_n^{i^\star}$ for each $n\in\mathcal{N}$.

Example. In the first iteration with the initial partition $\mathcal{P}_\Phi=\{\Phi^0\}$, the two global bounds are set to $\mathrm{UP}_{\mathrm{glb}}=\mathrm{UP}(\Phi^0)$ and $\mathrm{LR}_{\mathrm{glb}}=\mathrm{LR}(\Phi^0)$, respectively. The algorithm then partitions Φ^0 into two sub-domains $\Phi^{0,j}$ with $j=1,2$, and for them each the algorithm calculates two local bounds $\mathrm{UP}(\Phi^{0,j})$ and $\mathrm{LR}(\Phi^{0,j})$. Then, if $\mathrm{UP}(\Phi^{0,j})<\mathrm{LR}_{\mathrm{glb}}$, this implies that it is impossible for the globally optimal transmission profile $\boldsymbol{\beta}^\star$ to be located in $\Phi^{0,j}$, and hence $\Phi^{0,j}$ can be removed from the partition $\mathcal{P}_\Phi$ and hence $\Phi^{0,j}$ will not be partitioned any more in the following iterations. Finally, the algorithm updates the global upper and lower bounds as

$$\mathrm{UP}_{\mathrm{glb}}=\max_{j=1,2}\left\{\mathrm{UP}(\Phi^{0,j})\right\},\tag{40}$$

$$\mathrm{LR}_{\mathrm{glb}}=\max_{j=1,2}\left\{\mathrm{LR}(\Phi^{0,j})\right\}\tag{41}$$

respectively.

Remark: The complexity of Algorithm 3, in terms of the total number of iterations required to compute an ϵ-optimal solution, depends on the number of users, maximum transmission power of each user, and also on the optimality precision. We observed through experiments that Algorithm 3 converges very fast (within a couple of seconds of runtime on Dell Optiplex 9020M with Intel Core(TM) i7-4785T CPU@2.20 GHz and 16.0 GB RAM) for wireless networks with less than 10 users; while for large networks, e.g., with more than 50 users, the algorithm converges quite slowly. The situation can be exacerbated if the optimality precision is set to a value very close to 1 (e.g., 0.999). In the worst case, algorithms designed based on the branch and bound framework end up with exhaustive search, i.e., examine all possible combinations of transmission strategies [52]. The optimality precision ϵ can be adjusted to achieve a tradeoff between solution optimality and computational complexity.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

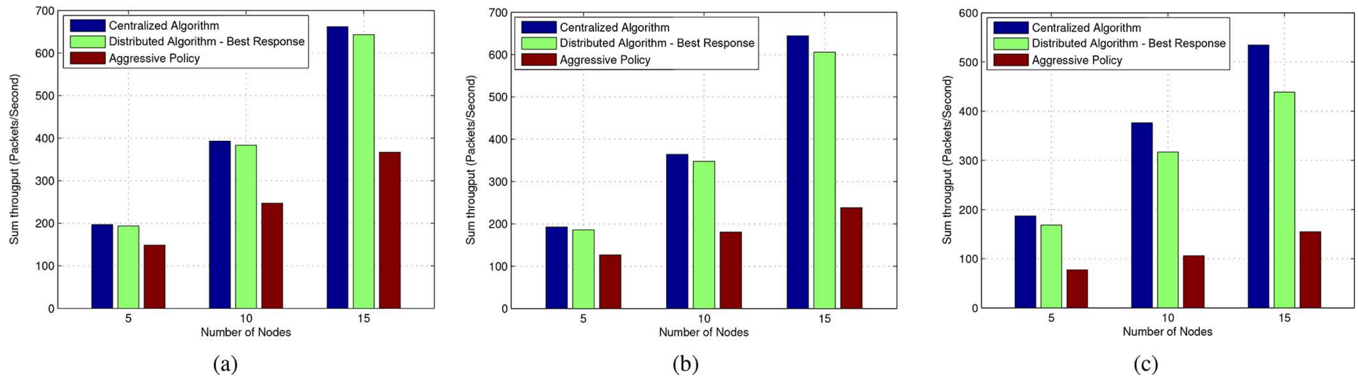


Fig. 3. Sum throughput versus the number of users with low level of the offered traffic load: $\lambda \in [20,\ 60]$, and with (a) ten, (b) five, and (c) two frequencies.

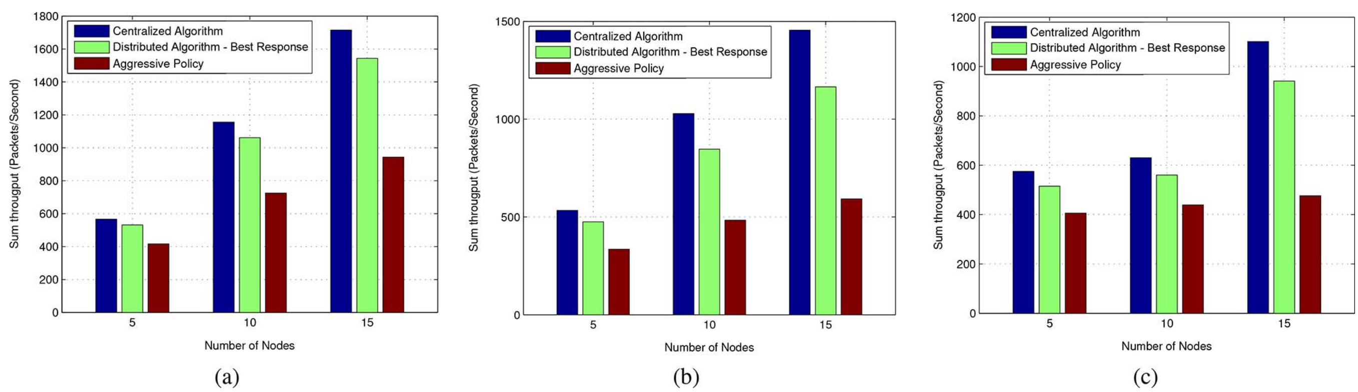


Fig. 4. Sum throughput versus the number of users with moderate level of the offered traffic load: $\lambda \in [20,\ 180]$, and with (a) ten, (b) five, and (c) two frequencies.

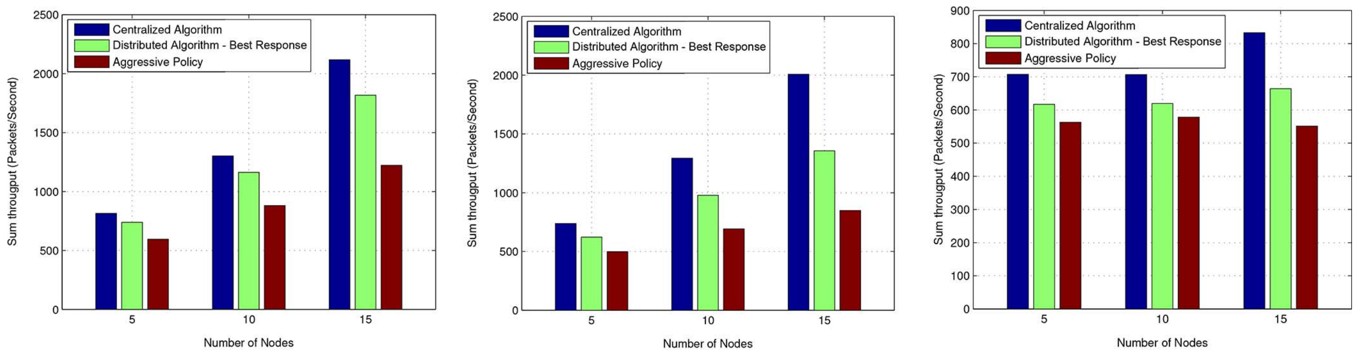
VI. Performance Evaluation

System setup. We consider a communication area with size of 200×200 m^2. The area is divided into 100 square blocks, each having a size of 20×20 m^2. Users are randomly placed in each block according to a Poisson distribution with different values of the density parameter μ. Each transmitter communicates with its intended receiver that is located in an arbitrary direction with a uniformly distributed random distance between 50 and 100 meters. The path loss factor between any two nodes is set to $\alpha = 3$, and the Rayleigh fading factor is set to $\Omega = 0.5$. The number of frequency channels is varied between 2 and 10. The SINR threshold $\gamma_{\rm th}$ to successfully decode a packet is set to a typical value of 10. The duration of each time slot is set to 5 ms for all users, leading to a maximum number of user's transmit packets equal to 200 per second. The average incoming packet rate, i.e., λ_n for user $n \in \mathcal{N}$, is varied from 20 to 180 with step of 20 to inject low, moderate and high traffic loads into the network. The maximum delay time is set to a uniformly distributed random number between 60 and 100 ms.

We first study the convergence performance of the proposed best-response based distributed algorithm (Algorithm 1), and then evaluate its performance in terms of the sum throughput by comparing it to the global optimum obtained using the centralized algorithm developed in Section V-B. In the centralized algorithm, the optimality precision ϵ is set to 95%, i.e., the centralized algorithm achieves greater than or equal to 95% of the global optimum. A simple "aggressive policy" is also tested to provide a performance benchmark. Based on the aggressive policy, each user has a fixed small transmission threshold of $\beta_n = \beta_n^{\rm L}$ (which is set to $\frac{\sqrt{2}}{2}\Omega$ as discussed in Section IV-A). Consequently, in each time slot each user will choose to transmit a packet as long as there exists one or more channels with coefficients greater than $\frac{\sqrt{2}}{2}\Omega$, which corresponds a high (hence aggressive) transmission probability. All results are averaged over 20 independent channel and topology realizations.

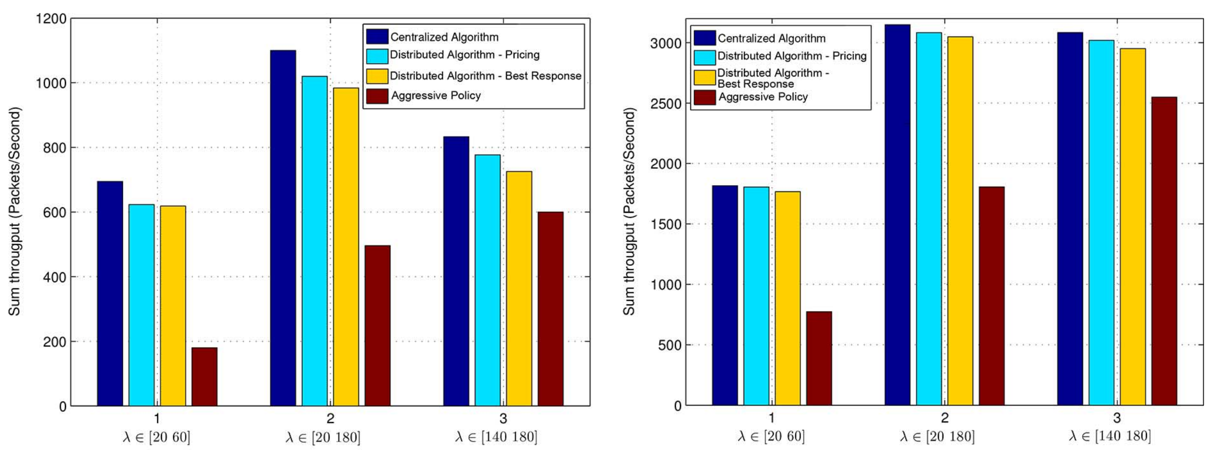
Convergence performance. Convergence of Algorithm 1 is shown in Fig. 2 for a network with 20 users. In the figure, we plot the throughput of five randomly selected users and the average fading threshold of all users versus the iteration number. The average incoming packet rates λ are randomly selected from the interval [20, 60]. We observed that the the proposed algorithm is very fast; in our tests, the desired accuracy is reached in four-five iterations. This validates our theoretical convergence analysis.

The algorithm also converges in the case of moderate load traffics with $\lambda \in [20, 180]$ and in the case of high load traffics with $\lambda \in [140, 180]$. The average Rayleigh fading thresholds corresponding to the above three cases are given in the bottom of Fig. 2. We observe that, in the case of low traffic load, at optimality users transmit using a high fading threshold, while they select a low threshold in the case of high load traffic, which is in agreement with our analysis in Section IV-A (i.e., in a high-congestion network, users choices converge to a low fading threshold in favor of more transmission opportunities and hence lower packet dropping rate caused by violations of the maximum delay constraint).



(a) (b) (c)

Fig. 5. Sum throughput versus the number of users with high level of the offered traffic load: $\lambda \in [140,\ 180]$, and with (a) ten, (b) five, and (c) two frequencies.



(a) (b)

Fig. 6. Sum throughput versus the level of offered traffic loads in high interference networks where there are (a) 15 and (b) 50 users, respectively, and two frequencies.

Throughput comparison. Comparison results of the three transmission schemes in terms of sum throughput are given in Figs. 3–6 for different numbers of frequencies (namely: 10, 5, and 2, respectively) users (namely: 5, 10, 15, and 50, respectively).

In Fig. 3, the average packet rate of each user is selected from [20, 60], which results in a network with low traffic load. We can see that Algorithm 1 achieves good sum throughput performance. For example, Fig. 3(b) shows that an 98.3% of the global optimum can be achieved in the case of five users and ten frequencies, and at least 82% of the optimum can be achieved with 15 users and 2 frequencies. It can also be seen that the aggressive policy achieves comparable (but lower) throughput only in the case of five users with ten and five frequencies, while it achieves a sum throughput much lower than the global optimum and the proposed best-response algorithm in all other cases. Good performance of Algorithm 1 are also observed in networks with moderate and high traffic loads, as shown in Figs. 4 and 5, respectively. For instance, in the worst case scenario corresponding to high traffic load and five frequencies, Algorithm 1 still achieves near 70% of the global optimum, which is almost the double of what is achievable by the aggressive policy. Additionally, in Fig. 5(c) where there are two channels and high traffic loads, the performance of the best-response distributed algorithm and the aggressive policy are close to each other. This confirms the observation in Fig. 2, that is, with the best-response algorithm users choose a lower channel fading threshold in networks with higher traffic loads.

The comparison of Figs. 3(c) and 4(c) shows that the sum throughput can be increased by injecting more data into a low-load network, whereas from Figs. 4(c) and 5(c) one infers that the throughput decreases when the network load is already high. Moreover, by comparing Figs. 3(a)–5(b) and (c), one can see that the gap between the throughput achievable by the NE and that of the global optimal solution is non negligible because of the large number of users, less available channels or high traffic load.

Fig. 6 shows that this performance loss can be partially compensated by allowing some degree of cooperation among users as implemented in Algorithm 2. More specifically, in the figure, we plot the sum throughput achievable by Algorithm 2 for different traffic loads. In fact, one can see that in case of high traffic load, Algorithm 2 achieves more than 7% of throughput improvement with respect to Algorithm 1, whereas the gain reduces to 3.5% in case of moderate load, and it is negligible when the traffic load is low. As expected, these results confirm that local cooperation (through pricing) is desirable in high interference networks [see Fig. 6(a) and (b)].

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

VII. Conclusion

In this paper, we studied the maximization of the sum-throughput in interference-limited queueing wireless networks via distributed selections of users' transmission thresholds. We formulated the optimization problem using a game theoretical approach (noncooperative design) and a more classical nonlinear programming approach (holistic design). We developed a general best-response algorithmic framework wherein the users can explicitly choose the level of cooperation (in the form of message passing) and consequently converge to: 1) a NE of the NEP, if no signaling exchange is allowed or 2) to a local optimal solution of the (nonconvex) NUM problem, if some coordination in the form of pricing exchange is performed among the users. Finally we developed a global solution method for the NUM problem and numerically compared the performance of the cooperative and noncooperative solutions. In many practical scenarios, a NE (noncooperative solution) yields sum-throughput very close to the global optimum. In case of "high" interference (i.e., coupling) among users, the performance loss can be partially filled by allowing some cooperation among the users and relying on the proposed pricing-based algorithm. We are currently implementing the proposed distributed algorithms on a USRP2/GNU radio testbed. As future research subjects, we will take power and rate control into consideration in the queueing game, and extend the queueing game to multi-hop wireless networks.

Appendix A
Model Parameter Estimation

We first derive the first and the second order moment of $I_n^f(\boldsymbol{\beta}_{-n})$. Using $\widetilde{\mathrm{E}}[I_n^f(\boldsymbol{\beta}_{-n})]$ to denote the first order moment of $I_n^f(\boldsymbol{\beta}_{-n})$, according to (11), $\widetilde{\mathrm{E}}[I_n^f(\boldsymbol{\beta}_{-n})]$ can be represented as

$$\begin{aligned}\widetilde{\mathrm{E}}\left[I_n^f(\boldsymbol{\beta}_{-n})\right] &= \sum_{m\in\mathcal{N}/n} P_m \mathrm{E}\left[h_{mn}^2(\tilde{h}_{mn}^f)^2\alpha_m^f(\beta_m)\right] \\ &= \sum_{m\in\mathcal{N}/n} P_m \hat{h}_{mn}^2 \int_{\beta_m}^{\infty} \frac{2x^3}{\Omega} \\ &\quad \times e^{-\frac{x^2}{\Omega}} dx \frac{1}{|\mathcal{F}|}\left(1-(1-e^{-\frac{\beta_m^2}{\Omega}})^{|\mathcal{F}|}\right) \quad (42)\end{aligned}$$

where $\phi(\beta_m)$ represents the probability that user $m \in \mathcal{N}/n$ transmits a packet in a time slot as defined in (3). The second order moment of $I_n^f(\boldsymbol{\beta}_{-n})$, denoted as $\widetilde{\mathrm{D}}[I_n^f(\boldsymbol{\beta}_{-n})]$, can also be represented similarly.

According to (12), the first and second order moments can also be represented as $\widetilde{\mathrm{E}}[I_n^f(\boldsymbol{\beta}_{-n})] = k_n(\boldsymbol{\beta}_{-n})\theta_n(\boldsymbol{\beta}_{-n})$, and $\widetilde{\mathrm{D}}[I_n^f(\boldsymbol{\beta}_{-n})] = k_n(\boldsymbol{\beta}_{-n})\theta_n^2(\boldsymbol{\beta}_{-n})$, respectively. Then, we can calculate $\theta_n(\boldsymbol{\beta}_{-n})$ and $k_n(\boldsymbol{\beta}_{-n})$ as

$$\theta_n(\boldsymbol{\beta}_{-n}) = \frac{\widetilde{\mathrm{D}}\left[I_n^f(\boldsymbol{\beta}_{-n})\right]}{\widetilde{\mathrm{E}}\left[I_n^f(\boldsymbol{\beta}_{-n})\right]} \quad (43)$$

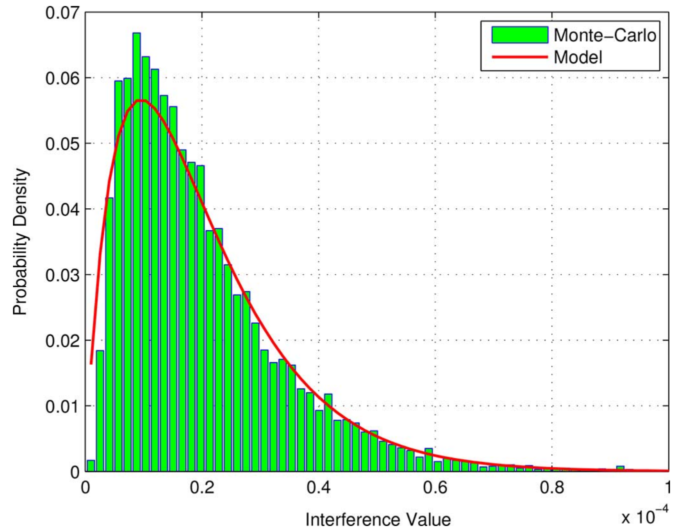


Fig. 7. Verification of interference model.

$$k_n(\boldsymbol{\beta}_{-n}) = \frac{\left(\widetilde{\mathrm{E}}\left[I_n^f(\boldsymbol{\beta}_{-n})\right]\right)^2}{\widetilde{\mathrm{D}}\left[I_n^f(\boldsymbol{\beta}_{-n})\right]} \quad (44)$$

respectively.

In Fig. 7, the interference model is verified by considering a wireless network with an area of $1000 \times 1000\ \mathrm{m}^2$ and 100 users. Results of Monte-Carlo are obtained by averaging over 10000 simulations. We can see the the interference model based on Gamma distribution function fits the real interference distribution very well.

Appendix B
Proof of Theorem 1

Invoking standard results on NEPs (see, e.g., [53]), the existence of a NE is guaranteed if: i) the strategy set of each player is convex and compact; and ii) the payoff function of each players is a continuous function on $\boldsymbol{\beta}$ and concave on β_n, for any given feasible $\boldsymbol{\beta}_{-n}$. We only need to prove the concavity of each function $R(\bullet, \boldsymbol{\beta}_{-n})$ on $[\beta_n^{\mathrm{L}}, \beta_n^{\mathrm{U}}]$, when $\beta_n^{\mathrm{U}} \geq \frac{\sqrt{2}}{2}\Omega$. This can be checked by computing the first and second derivatives of $P_n^{\mathrm{los}}(\boldsymbol{\beta})$ with respect to β_n, as given next. The first derive of $P_n^{\mathrm{los}}(\boldsymbol{\beta})$ with respect to β_n is

$$\begin{aligned}&\frac{\partial P_n^{\mathrm{los}}(\boldsymbol{\beta})}{\partial \beta_n} \\ &= \underbrace{-\frac{2\beta_n}{\Omega} e^{-\frac{\beta_n^2}{\Omega}} \vartheta_n\left(\frac{P_n \hat{h}_n^2}{\gamma_{\mathrm{th}}}\beta_n^2 - \left(\sigma_n^f\right)^2, \boldsymbol{\beta}_{-n}\right)}_{A(\boldsymbol{\beta})} \\ &\quad + \underbrace{\frac{2|\mathcal{F}|\beta_n T_n^{\mathrm{th}}}{T_{\mathrm{slt}}\Omega} e^{-(\frac{\mu_n(\beta_n)}{T_{\mathrm{slt}}} - \lambda_n)T_n^{\mathrm{th}}}\left(1 - e^{-\frac{\beta_n^2}{\Omega}}\right)^{|\mathcal{F}|-1}}_{B(\boldsymbol{\beta})}.\end{aligned} \quad (45)$$

$$\frac{\partial A(\boldsymbol{\beta})}{\partial \beta_n} = \underbrace{\frac{2}{\Omega} e^{-\frac{\beta_n^2}{\Omega}} \left(-1 + \frac{2\beta_n^2}{\Omega}\right) \vartheta_n \left(\frac{P_n \hat{h}_n^2}{\gamma_{\rm th}} \beta_n^2 - \left(\sigma_n^f\right)^2, \boldsymbol{\beta}_{-n}\right)}_{A_1} + \underbrace{\left(-\frac{2\beta_n}{\Omega} e^{-\frac{\beta_n^2}{\Omega}}\right) \frac{\partial}{\partial \beta_n} \vartheta_n \left(\frac{P_n \hat{h}_n^2}{\gamma_{\rm th}} \beta_n^2 - \left(\sigma_n^f\right)^2, \boldsymbol{\beta}_{-n}\right)}_{A_2} \quad (47)$$

$$\frac{\partial B(\boldsymbol{\beta})}{\partial \beta_n} = |\mathcal{F}| \frac{T_n^{\rm th}}{T_{\rm slt}} e^{\lambda_n T_n^{\rm th}} \left\{ a_n + \left(\underbrace{|\mathcal{F}| \frac{T_n^{\rm th}}{T_{\rm slt}} e^{-\frac{\beta_n^2}{\Omega}} \left(1 - e^{-\frac{\beta_n^2}{\Omega}}\right)^{|\mathcal{F}|-1}}_{B_1} - 1 \right) \left(\frac{2\beta_n}{\Omega}\right)^2 \right\} \quad (48)$$

The second derivative of $P_n^{\rm los}(\boldsymbol{\beta})$ with respect to β_n can be written as

$$\frac{\partial^2 P_n^{\rm los}(\boldsymbol{\beta})}{\partial \beta_n^2} = \frac{\partial A(\boldsymbol{\beta})}{\partial \beta_n} + \frac{\partial B(\boldsymbol{\beta})}{\partial \beta_n} \quad (46)$$

with $\partial A(\boldsymbol{\beta})/\partial \beta_n$ and $\partial B(\boldsymbol{\beta})/\partial \beta_n$ are given by (47) and (48), shown at the top of the page, where a_n is a positive constant whose explicit expression is not relevant and thus is omitted. Note that, if $\beta_n \geq \frac{\sqrt{2}}{2}\Omega$, $A_1 \geq 0$ and $A_2 > 0$, the latter due to $\frac{\partial \vartheta_n(\cdot,\cdot)}{\partial \beta_n} < 0$. Therefore, we have $\partial A(\boldsymbol{\beta})/\partial \beta_n > 0$. Also, when $|\mathcal{F}|(T_n^{\rm th}/T_{\rm slt}) \gg 1$ (i.e., delay-tolerant traffic or large number of sub-channels), we have $B_1 > 1$, implying $\partial B(\boldsymbol{\beta})/\partial \beta_n > 0$.

Appendix C
Proof of Theorem 2

We hinge on recent results in ([11], Proposition 3). Given $\mathrm{R} \triangleq (\nabla_{\beta_n} R_n)_{n\in\mathcal{N}}$, let us introduce the $|\mathcal{N}| \times |\mathcal{N}|$ matrix $\Upsilon_{\rm R}$ defined as

$$[\Upsilon_{\rm R}]_{mn} \triangleq \begin{cases} \omega_m^{\min}, & \text{if } m = n, \\ -\tau_{mn}^{\max}, & \text{otherwise} \end{cases}$$

where $\omega_m^{\min} \triangleq \inf_{\boldsymbol{\beta}\in\Phi} \lambda_{\rm least}(J_{\beta_m} \mathrm{R}_m(\boldsymbol{\beta}))$ and $\tau_{mn}^{\max} \triangleq \sup_{\boldsymbol{\beta}\in\Phi} \|J_{\beta_m} \mathrm{R}_n(\boldsymbol{\beta})\|$ with $\lambda_{\rm least}(\boldsymbol{A})$ denoting the least eigenvalue of $\boldsymbol{A}$. Then, invoking ([11], Proposition 3) one can show that Algorithm 1 converges to the unique NE of $\mathcal{G}$ if $\Upsilon_{\rm R}$ is a P-matrix. It turns our that condition (23) is sufficient for $\Upsilon_{\rm R}$ to be P (see also [12] for more details).

References

[1] Z. Guan, T. Melodia, and G. Scutari, “Distributed queuing games in interference-limited wireless networks,” in *Proc. IEEE Int. Conf. Commun.*, Budapest, Hungary, Jun. 2013, pp. 403–408.
[2] L. Kleinrock, *Queueing Systems*. New York, NY, USA: Wiley, 1975.
[3] D. Bertsekas and R. Gallager, *Data Networks*. Upper Saddle River, NJ, USA: Prentice-Hall, 2000.
[4] G. Bolch, S. Greiner, H. de Meer, and K. Trivedi, *Queueing Networks and Markov Chains*. New York, NY, USA: Wiley, 1998.
[5] Z. Guan, T. Melodia, D. Yuan, and D. Pados, “Distributed spectrum management and relay selection in interference-limited cooperative wireless networks,” in *Proc. ACM Int. Conf. Mobile Computing and Networking*, Las Vegas, NV, USA, Sep. 2011, pp. 229–240.
[6] R. Rahman, T. Vinko, D. Hales, J. Pouwelse, and H. Sips, “Design space analysis for modeling incentives in distributed systems,” in *Proc. ACM Special Interest Group on Data Commun.*, Toronto, ON, Canada, Aug. 2011, pp. 182–193.
[7] G. S. Kasbekar and S. Sarkar, “Spectrum pricing games with bandwidth uncertainty and spatial reuse in cognitive radio networks,” in *Proc. ACM Int. Symp. Mobile Ad Hoc Networking and Computing*, Chicargo, Illinois, USA, Sep. 2010, pp. 251–260.
[8] D. Yang, X. Fang, and G. Xue, “Truthful auction for cooperative communications,” in *Proc. ACM Int. Symp. Mobile Ad Hoc Networking and Computing*, Paris, France, May 2011, pp. 1–10.
[9] G. Scutari, D. P. Palomar, and S. Barbarossa, “Optimal linear precoding strategies for wideband noncooperative systems based on game theory—Part I & II: Nash equilibria and distributed algorithms,” *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 1230–1267, Mar. 2008.
[10] G. Scutari, D. P. Palomar, and S. Barbarossa, “Asynchronous iterative water-filling for Gaussian frequency-selective interference channels,” *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 2868–2878, Jul. 2008.
[11] G. Scutari, D. P. Palomar, F. Facchinei, and J.-S. Pang, , A. Rantzer and R. Johansson, Eds., “Monotone games for cognitive radio systems,” in *Distributed Decision-Making and Control*, ser. Lecture Notes in Control and Information Sciences Series. New York, NY, USA: Springer-Verlag, 2011.
[12] G. Scutari, F. Facchinei, J.-S. Pang, and D. Palomar, “Real and complex monotone communication games,” *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4197–4231, Jul. 2014.
[13] G. Scutari, F. Facchinei, P. Song, D. Palomar, and J.-S. Pang, “Decomposition by partial linearization: Parallel optimization of multi-agent systems,” *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 641–656, Feb. 2014.
[14] F. Ren and C. Lin, “Modeling and improving TCP performance over cellular link with variable bandwidth,” *IEEE Trans. Mobile Computing*, vol. 10, no. 8, pp. 1057–1070, Aug. 2011.
[15] H. Huang and V. K. N. Lau, “Delay-optimal user scheduling and intercell interference management in cellular network via distributive stochastic learning,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 12, pp. 3790–3797, Dec. 2010.
[16] L. Le, E. Modiano, and N. B. Shroff, “Longest-queue-first scheduling under SINR interference model,” in *Proc. ACM Int. Symp. Mobile Ad Hoc Networking and Computing*, Chicago, IL, USA, Sep. 2010, pp. 1–10.
[17] R. Urgaonkar and M. J. Neely, “Network capacity region and minimum energy function for a delay-tolerant mobile ad hoc network,” *IEEE/ACM Trans. Netw.*, vol. 19, no. 4, pp. 1137–1150, Aug. 2011.
[18] J. Chen and V. K. N. Lau, “Delay analysis of max-weight queue algorithm for time-varying wireless ad hoc networks control theoretical approach,” *IEEE Trans. Signal Process.*, vol. 61, no. 1, pp. 99–108, Jan. 2013.
[19] H. Li and Z. Han, “Socially optimal queuing control in cognitive radio networks subject to service interruptions: To queue or not to queue?,” *IEEE Trans. Wireless Commun.*, vol. 10, no. 5, pp. 1656–1666, May 2011.
[20] S. Wang, J. Zhang, and L. Tong, “Delay analysis for cognitive radio networks with random access: A fluid queue view,” in *Proc. IEEE Int. Conf. Computer Commun.*, San Diego, CA, USA, Mar. 2010, pp. 1–9.
[21] J. Wang, A. Huang, L. Cai, and W. Wang, “On the queue dynamics of multiuser multichannel cognitive radio networks,” *IEEE Trans. Veh. Technol.*, vol. 62, no. 3, pp. 1314–1328, Mar. 2013.
[22] C. T. Do, N. H. Tran, C. S. Hong, S. Lee, J.-J. Lee, and W. Lee, “A lightweight algorithm for probability-based spectrum decision scheme in multiple channels cognitive radio networks,” *IEEE Commun. Lett.*, vol. 17, no. 3, pp. 509–512, Mar. 2013.

[23] O. Tickoo and B. Sikdar, “Modeling queueing and channel access delay in unsaturated IEEE 802.11 random access MAC based wireless networks,” *IEEE/ACM Trans. Netw.*, vol. 16, no. 4, pp. 878–891, Aug. 2008.

[24] J.-G. Choi, C. Joo, J. Zhang, and N. B. Shroff, “Distributed link scheduling under SINR model in multihop wireless networks,” *IEEE/ACM Trans. Netw.*, vol. 22, no. 4, pp. 1204–1217, Aug. 2014.

[25] Y. Xi and E. M. Yeh, “Throughput optimal distributed power control of stochastic wireless networks,” *IEEE/ACM Trans. Netw.*, vol. 18, no. 4, pp. 1054–1066, Aug. 2010.

[26] M. J. Hossain, M.-S. Alouini, and V. Bhargava, “Channel access delay and buffer distribution of two-user opportunistic scheduling schemes in wireless networks,” *IEEE Trans. Wireless Commun.*, vol. 58, no. 7, pp. 2077–2087, July 2010.

[27] L. Ying, R. Srikant, D. Towsley, and S. Liu, “Cluster-based back-pressure routing algorithm,” *IEEE/ACM Trans. Netw.*, vol. 19, no. 6, pp. 1773–1786, Dec. 2011.

[28] J. Ghaderi and R. Srikant, “The impact of access probabilities on the delay performance of Q-CSMA algorithms in wireless networks,” *IEEE/ACM Trans. Netw.*, vol. 21, no. 4, pp. 1063–1075, Aug. 2013.

[29] D. Xue and E. Ekici, “Queuing behavior of queue-backlog-based random access scheduling in the many-channel regime,” in *Proc. Int. Symp. Workshops on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, Tsukuba Science City, Japan, May 2013, pp. 476–483.

[30] M. J. Neely, “A lyapunov optimization approach to repeated stochastic games,” in *Proc. Allerton Conf. Commun. Control, and Computing*, Monticello, IL, USA, Oct. 2013, pp. 1–8.

[31] L. Zheng and L. Cai, “A distributed demand response control strategy using Lyapunov optimization,” *IEEE Trans. on Smart Grid*, vol. 5, no. 4, pp. 2075–2083, July 2014.

[32] G. Scutari, D. P. Palomar, F. Facchinei, and J.-S. Pang, “Convex optimization, game theory, and variational inequality theory in multiuser communication systems,” *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 35–49, May 2010.

[33] J.-S. Pang, G. Scutari, D. P. Palomar, and F. Facchinei, “Design of cognitive radio systems under temperature-interference constraints: A variational inequality approach,” *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3251–3271, June 2010.

[34] K.-J. Park, C. Hou, T. Başar, and H. Kim, “Noncooperative carrier sense game in wireless networks,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 10, pp. 5280–5289, Oct. 2009.

[35] P. Dube and R. Jain, “Queueing game models for differentiated services,” in *Proc. Int. Conf. Game Theory for Networks*, Istanbul, Turkey, May 2009, pp. 523–532.

[36] Y. Gai, H. Liu, and B. Krishnamachari, “A packet dropping-based incentive mechanism for M/M/1 queues with selfish users,” in *Proc. IEEE Int. Conf. Comput. Commun.*, Shanghai, China, Apr. 2011, pp. 1–9.

[37] Y. Sarikaya, T. Alpcan, and O. Ercetin, “Dynamic pricing and queue stability in wireless random access games,” *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 4, pp. 140–150, Apr. 2012.

[38] E. Altman and Y. Hayel, “A stochastic evolutionary game approach to energy management in a distributed aloha network,” in *Proc. IEEE Int. Conf. Comput. Commun.*, Phoenix, AZ, USA, Apr. 2008, pp. 1–9.

[39] 4G Americas, “4G Americas recommendations on 5G requirements and solutions,” Oct. 2014 [Online]. Available: http://www.4gamericas.org

[40] T. S. Rappaport, *Wireless Communications*. New York, NY, USA: Prentice-Hall PTR, 1996.

[41] A. A. El-Sherif and K. J. R. Liu, “Cooperation in random access networks: Protocol design and performance analysis,” *IEEE J. Sel. Areas Commun.*, vol. 30, no. 9, pp. 1694–1702, Oct. 2012.

[42] A. Fanous and A. Ephremides, “Stable throughput in a cognitive wireless network,” *IEEE J. Sel. Areas Commun.*, vol. 31, no. 3, pp. 523–533, Mar. 2013.

[43] F. Baccelli and B. Blaszczyszyn, *Stochastic Geometry and Wireless Networks, Part I: Theory*. Norwell, MA, USA: Now, 2009.

[44] IEEE 802.11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications). (2007 Revision) Jun. 2007 [Online]. Available: doi:10.1109/IEEESTD.2007.373646, IEEE-SA

[45] G. Scutari, D. P. Palomar, F. Facchinei, and J.-S. Pang, “Distributed dynamic pricing for MIMO interfering multiuser systems: A unified approach,” in *Proc. Int. Conf. Network Games, Control and Optimization*, Paris, France, Oct. 12–14, 2011, pp. 1–5.

[46] A. L. Yuille and A. Rangarajan, “The concave-convex procedure,” *Neural Computation*, vol. 15, no. 4, pp. 915–936, Apr. 2003.

[47] B. K. Sriperumbudur and G. R. G. Lanckriet, “On the convergence of the concave-convex procedure,” in *Proc. NIPS Workshop Optimization for Machine Learning*, Whistler, British Columbia, Canada, Dec. 2009, pp. 1–9.

[48] Y. Cheng and M. Pesavento, “Joint optimization of source power allocation and distributed relay beamforming in multiuser peer-to-peer relay networks,” *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 2962–2973, Jun. 2012.

[49] S. Boyd, “Sequential convex programming,” *Lecture Note* 2013 [Online]. Available: http://www.stanford.edu/class/ee364b/lectures/seq_slides.pdf

[50] M. Hast, K. Astrom, B. Bernhardsson, and S. Boyd, “PID design by convex-concave optimization,” in *Proc. Eur. Control Conf.*, Zurich, Switzerland, Jul. 2013, pp. 4460–4645.

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

[52] S. Boyd and J. Mattingley, “Branch and bound methods,” Stanford Univ., Notes for EE364b, Mar. 2007 [Online]. Available: http://stanford.edu/class/ee364b/lectures/bb_slides.pdf

[53] J. B. Rosen, “Existence and uniqueness of equilibrium points for n-person games,” *Econometrica*, vol. 33, no. 3, pp. 520–534, Jul. 1965.

Zhangyu Guan (M'11) received the Ph.D. degree in communication and information systems from Shandong University, Jinan, China, in 2010.

He is currently a Postdoctoral Research Associate with the Department of Electrical and Computer Engineering, Northeastern University, Boston, MA, USA. He was a visiting Ph.D. student with the Department of Electrical Engineering, The State University of New York (SUNY) at Buffalo, Buffalo, NY, USA, from 2009 to 2010; he was a Lecturer with Shandong University from 2011 to 2014, and he was a Postdoctoral Research Associate with the Department of Electrical Engineering at SUNY Buffalo from November 2012 to August 2015. His current research interests are in cognitive radio and software-defined networking, wireless multimedia sensor networks, and underwater networks.

Dr. Guan has served as a TPC member for IEEE INFOCOM 2016, IEEE GLOBECOM 2015, IEEE ICNC 2012–2015, and IEEE VTC 2011-Fall and 2015-Fall, among others, and served as a reviewer for the IEEE TRANSACTIONS ON COMMUNICATIONS, the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, and the IEEE COMMUNICATIONS SURVEYS AND TUTORIALS, among others.

Tommaso Melodia (M'07) received the Ph.D. degree in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2007.

He is an Associate Professor with the Department of Electrical and Computer Engineering, Northeastern University, Boston, MA, USA. His research has been supported by the National Science Foundation, Air Force Research Laboratory, and the Office of Naval Research, among others. His current research interests are in modeling, optimization, and experimental evaluation of networked communication systems, with applications to ultrasonic intra-body networks, cognitive and cooperative networks, multimedia sensor networks, and underwater networks.

Prof. Melodia was a recipient of the National Science Foundation CAREER Award and coauthored a paper that was recognized as the ISI Fast Breaking Paper in the field of Computer Science for February 2009 and of an ACM WUWNet 2013 Best Paper Award. He was the Technical Program Committee Vice Chair for IEEE Globecom 2013 and the Technical Program Committee Vice Chair for Information Systems for IEEE INFOCOM 2013. He serves on the editorial boards of the IEEE TRANSACTIONS ON MOBILE COMPUTING, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON MULTIMEDIA, and *Computer Networks*.

Gesualdo Scutari (S'05–M'06–SM'11) received the Electrical Engineering and Ph.D. degrees (both with honors) from the University of Rome La Sapienza, Rome, Italy, in 2001 and 2005, respectively.

He is an Assistant Professor with the Department of Electrical Engineering, The State University of New York (SUNY) at Buffalo, Buffalo, NY, USA. He had previously held several research appointments, namely, at University of California at Berkeley, Berkeley, CA, USA, Hong Kong University of Science and Technology, Hong Kong, University of Rome, La Sapienza, Rome, Italy, and University of Illinois at Urbana-Champaign, Urbana, IL, USA. His primary research interests focus on theoretical and algorithmic issues related to continuous optimization, equilibrium programming and their applications to signal processing, communications and networking, big data analytics, and distributed decisions.

Dr. Scutari is an associate editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING, and he served as an associate editor of IEEE SIGNAL PROCESSING LETTERS. He serves on the IEEE Signal Processing Society Technical Committee on Signal Processing for Communications (SPCOM). He was the recipient of the 2006 Best Student Paper Award at the International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2006, the 2013 National Science Foundation Faculty Early Career Development (CAREER) Award, and the 2013 UB Young Investigator Award.