Distributed Optimization of Channel Access Strategies in Reactive Cognitive Networks

Marco Levorato, *Member, IEEE*, Pradeep Chathuranga Weeraddana, *Member, IEEE*, and Carlo Fischione, *Member, IEEE*

***Abstract*— In reactive cognitive networks, the channel access and the transmission decisions of the cognitive terminals have a long-term effect on the network dynamics. When multiple cognitive terminals coexist, the optimization and implementation of their strategy is challenging and may require considerable coordination overhead. In this paper, such challenge is addressed by a novel framework for the distributed optimization of transmission and channel access strategies. The objective of the cognitive terminals is to find the optimal *action* distribution depending on the current network state. To reduce the coordination overhead, in the proposed framework the cognitive terminals distributively coordinate the policy, whereas the action in each individual time slot is independently selected by the terminals. The optimization of the transmission and channel access strategy is performed iteratively by using the alternate convex optimization technique, where at each iteration a cognitive terminal is selected to optimize its own action distribution while assuming fixed those of the other cognitive terminals. For a traditional primary–secondary user network configuration, numerical results show that the proposed algorithm converges to a stable solution in a small number of iterations, and a limited performance loss with respect to the perfect coordinated case.**

***Index Terms*— Cognitive networks, distributed optimization, Markov decision processes.**

I. Introduction

COGNITIVE radios open new opportunities and communication paradigms to increase the efficiency of wireless spectrum usage [1]–[3]. This area has attracted a considerable research effort, with particular attention devoted to the study of cognitive devices operating in the same bandwidth as primary wireless terminals. The goal of the cognitive devices, generally referred to as *secondary users*, is to maximize their performance while bounding the performance degradation caused to the non-cognitive primary devices, which are the legitimate owners of the radio resource. The underlay and interweave approaches [4] were proposed to optimize secondary users' transmissions in this scenario. In the former approach the secondary users are allowed to operate concurrently with the primary users as long as the interference at the receivers of the primary users' signals remains below a predefined threshold [5], [6]. In the interweave paradigm, the secondary users opportunistically identify and exploit time and frequency slots left unused by the primary users for their transmissions (the so called *white space* approach) [3], [7].

Most prior work on the interweave approach proposes frameworks where the primary users generate an idle/busy pattern over multiple radio channels. The activity of the primary users is described by a stochastic process with a binary state space whose statistics are assumed independent of the secondary users' activity. Due to this assumption, these frameworks cannot capture the complex interactions between the primary and the secondary users occurring when transmission and networking protocols are implemented. In fact, the activity of the secondary users may trigger a *reaction* of the primary users and impact the distribution of the future network's state. For instance, interference due to secondary user's transmissions may force the primary user to enter a backoff period if a channel sensing-based protocol is used to regulate primary users' access to the wireless resource [8], or may trigger a retransmission if the primary users implement Automatic Retransmission reQuest (ARQ) [9]. Therefore, the assumption of independence between the temporal evolution of the primary users' network and the activity of the secondary users limits the range of applications of the primary-secondary user framework to those network scenarios where the primary users do not implement any form of control. In [10] and [11], a framework, namely *reactive primary users*, was introduced to capture these interactions. The secondary users optimize their transmission strategy accounting for the impact of the generated interference on the temporal evolution of the state of the network. Importantly, the reactive framework also allows for a more accurate estimation of the performance degradation caused to the primary users. Primary-secondary users' interactions were further studied in [12]–[16].

One of the most important features of the reactive cognitive framework is the introduction of the notion of *network state*. By capturing the internal state of the network nodes and their protocols (see Sec. II and V for definitions and a detailed case study), the reactive framework opens to the advancement of the traditional primary-secondary user network setting to more general cognitive network scenarios where smart terminals optimize their transmission and channel access strategies to achieve the desired Quality of Service (QoS). Finite State Machines (FSM) can be defined to capture the

Manuscript received October 20, 2015; revised April 11, 2016 and August 8, 2016; accepted August 12, 2016. Date of publication August 24, 2016; date of current version October 14, 2016. The associate editor coordinating the review of this paper and approving it for publication was P. Mitran.

M. Levorato is with the Donald Bren School of Information and Computer Sciences, University of California at Irvine, Irvine, CA 92617 USA (e-mail: levorato@uci.edu).

P. C. Weeraddana is with the Sri Lanka Institute of Information Technology, Malabe 10115, Sri Lanka (e-mail: chathuranga.we@sliit.lk).

C. Fischione is with the ACCESS Center, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden (e-mail: carlofi@kth.se).

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/TCOMM.2016.2602207

dynamics of the internal state of the wireless terminals and the implemented network mechanisms (*e.g.*, packet buffering, packet retransmission, cooperative transmissions), as well as to track environmental variables (*e.g.*, channel state). The strategy of the cognitive terminals influences the statistics of FSM's state transition, and aims at steering the temporal evolution of the state on optimal trajectories.

Prior work based on this framework considers network scenarios focused on the primary-secondary user classification and with one smart secondary user. Herein, we propose a framework targeting a network setting with multiple smart terminals sharing the same channel resource. The objective of the smart terminals is to maximize some QoS metric (*e.g.*, aggregate throughput) under constraints on other QoS metrics (*e.g.*, packet failure). The reactive cognitive framework is used to formulate control policies defined on the network state, and to model the impact of transmission/interference resulting from the individual terminals on the statistics of the future network state. By removing the traditional primary-secondary user classification, we open our framework to novel network scenarios focused on the coexistence of smart, adaptive, terminals in modern cognitive networks. However, compared to the single smart secondary user, the presence of multiple such users poses major analytical and optimization challenges.

In the literature, the problem of distributively estimating the network state has been extensively studied [17], [18]. However, relatively few studies exist focusing on the problem of optimizing the operations of cognitive networks with multiple cognitive terminals. In [19], the authors propose a distributed resource management technique for such scenario. A learning approach to the same problem is presented in [20] and [21]. An information theoretic study is provided in [22]. In this paper, we propose an optimization strategy within the reactive cognitive framework specifically addressing a scenario with multiple secondary users. To make the implementation of the transmission strategy feasible, the secondary users coordinate their transmission policy, but, then, independently select their individual action (*e.g.*, transmission, idleness) in each time slot. In fact, action coordination would require an excessive amount of overhead to distributively select a global action, that is, the vector of all the actions, from the overall policy. To overcome this problem, we propose an iterative optimization strategy.

The iterative optimization is based on the Regularized Alternate Convex Optimization Approach (RACOA), where in a round-robin fashion, each individual cognitive terminal solves a local optimization problem and, then, transmits the result to the next cognitive terminal until a shared policy is found. Convergence to a coordinate-wise minimum of the objective function with linear convergence rate is demonstrated. The policy computation phase is repeated if the statistics of the stochastic process modeling the network's dynamics change. However, each phase requires the exchange of a small number of control packets containing action distributions, and no further coordination is needed between policy computation phases, where transmission actions are selected independently by the terminals. Note that coordination phases followed by implementation phases are explicitly or implicitly used in related papers focusing on different scenarios, such as [23] and [24]. The proposed technique can be used to optimize the transmission strategy of the secondary users in traditional primary-secondary user network where the primary users implement networking protocols. In order to exemplify this application of the framework, a case study is considered where multiple secondary users distributively optimize their transmission strategy to maximize their aggregate throughput under a constraint on the performance loss caused to a primary user implementing ARQ. Numerical results show a limited performance loss incurred by the distributed uncoordinated policy with respect to the ideal global optimum and fully coordinated policy. *However, we remark that the proposed framework can be used in more general cognitive network scenarios.*

The rest of the paper is organized as follows. Section II describes the considered network and defines the optimization problem. Section III presents the distributed optimization framework. Section IV discusses the convergence properties of the proposed algorithm. In Section V, numerical results are presented to assess the performance of the proposed framework for a case study network with a primary user implementing ARQ and multiple secondary users coordinating their transmission strategy. Section VI concludes the paper.

II. Problem Definition

First, we state the proposed algorithm for a general network optimization problem. A specific network scenario is instantiated in Sec. V. A network with N cognitive terminals sharing the wireless resource is considered. As observed in [25] and [26], the state of networking and communication protocols, as well as of the physical environment, is described by a finite collection of variables describing features of the network state. The individual variables track quantities and counters such as the number of packets in the buffers, the number of retransmissions of the packets being served, and the quality of the channel to the receivers,[1] the backoff countdown and transmission parameters. An exemplar of this construction is provided in [8], where Distributed Coordination Function (DCF) is modeled as a bi-dimensional Markov chain tracking the backoff counter and the retransmission index of the packet being served. Therefore, we model the operations of the network by a Finite State Machine (FSM), where the state of the FSM refers to the internal state of networking mechanisms. Slotted time is assumed, where the slots are indexed with $t = 0, 1, \ldots$.

Denote the state space of the FSM and the state at time t by $\mathcal{S}$ and $S(t) \in \mathcal{S}$, respectively. The state $S(t) \in \mathcal{S}$ is defined as the vector $\big(X_1(t), \ldots, X_M(t)\big)$, where the variables $X_m(t)$ describe features of the network state and M is the number of features. Due to the intrinsic randomness of *events* in the network (*e.g.*, packet arrival, packet decoding outcome), the state sequence $\mathbf{S} = (S(0), S(1), \ldots)$ is stochastic. Herein, we model the statistics of the temporal evolution of the network

[1]Continuous variables are assumed to be quantized to achieve an accurate approximation. See [27] for an example of wireless channel quantization strategy matched to the retransmission protocol.

state as a homogeneous Markov process. We remark that this is a widely accepted assumption in works addressing the modeling, analysis and control of wireless networks. In fact, most of the technical contributions in white space-based cognitive networks model the dynamics of the network as a binary Markov chain.

The cognitive terminals make transmission and channel access decisions based on the state of the network. We, then, define $U(t) \in \mathcal{U}$ as set of the actions of the cognitive terminals at time t. The action variable $U(t)$ consists of the vector $U(t) = \{U_1(t), \ldots, U_N(t)\}$, where $U_j(t) \in \mathcal{U}_j$, $j = 1, 2, \ldots, N$, is a random variable associated with the action of cognitive terminal j. For instance, the action space $\mathcal{U}$ can be a collection of N binary variables associated with transmission/idleness of the wireless terminals. However, the action space can include variables controlling transmission parameters and cooperative behaviors. Then, the statistics of $\mathbf{S}$ are completely defined by the transition probabilities

$$p(s'|s, u) = \Pr(S(t+1) = s'|S(t) = s, U(t) = u), \quad (1)$$

and initial state distribution $p_0(s)$, where $p_t(s) = \Pr(S(t) = s)$, where $\Pr(\cdot)$ denotes the probability of an event.

The objective of the cognitive terminals is to define a control sequence $U = \{U(0), U(1), \ldots\}$ maximizing some QoS metric measuring their performance (*e.g.*, aggregate throughput) under a constraint on some other QoS metric (*e.g.*, packet delivery delay). We remark that the cognitive terminals need to account for the interaction between their control sequence and the statistics of the Markov process determining the state trajectory of the FSM. QoS metrics such as throughput, packet failure probability and service time can be expressed as the time average of additive cost functions [28]. Define the transition cost functions $\omega(s'|s, u)$ and $\phi(s'|s, u)$, $s, s' \in \mathcal{S}$, $u \in \mathcal{U}$, namely the cost to move from s to s' as an effect of the action u taken by the terminals, where the first cost function corresponds to a QoS metric, and the second cost function corresponds to another QoS metric. The average costs conditioned on state s and action u are

$$\omega(s, u) = \sum_{s' \in \mathcal{S}} p(s'|s, u)\omega(s'|s, u), \quad (2)$$

$$\phi(s, u) = \sum_{s' \in \mathcal{S}} p(s'|s, u)\phi(s'|s, u). \quad (3)$$

An example of such cost functions is provided later in this paper. However, we remark that a wide array of performance metrics can be measured using the modeling structure adopted in this paper [28]. Then, the problem of optimizing the cognitive terminals control sequence is formulated as the constrained Markov Decision Process [29]

$$U^* = \arg\min \overline{\omega}(U) \text{ s.t. } \overline{\phi}(U) \leq C, \quad (4)$$

where $\overline{\omega}(\mu)$ and $\overline{\phi}(\mu)$ are the time averages

$$\overline{\omega}(U) = \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} E_U\left[\omega(S(t), U(t)) \middle| S(0) = s_0\right], \quad (5)$$

$$\overline{\phi}(U) = \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} E_U\left[\phi(S(t), U(t)) \middle| S(0) = s_0\right], \quad (6)$$

and $E_U[\cdot]$ denotes expectation conditioned on the policy U.

If the Markov process is unichain, that is, only one recurrent class exists, then at least one optimal policy is a randomized past-independent policy [30], that is, the optimal policy lies in the set of randomized maps $\mu : \mathcal{S} \times \mathcal{U} \to [0, 1]$. Without loss of optimality, then, we focus on this class of policies and define $\mu(s, u)$ as the probability that the action $u \in \mathcal{U}$ is selected given that the process is in state $s \in \mathcal{S}$, that is

$$\mu(u|s) = \Pr(U(t) = u|S(t) = s). \quad (7)$$

For any policy μ, the state sequence $S = (S(0), S(1), \ldots)$ is a Markov process. We remark the difference between action and policy. Action refers to the individual slot transmission decision, whereas the policy is the distribution of actions over the state space.

The following Linear Program (LP) can be used to find the optimal action distribution conditioned on the network state [30]. The optimization variables $z_{s,u}$, $\forall s, u$, are the steady-state joint probability that the Markov process is in state s and action u is selected, that is $z_{s,u} = \Pr(S(t) = s, U(t) = u)$.

$$z^* = \arg\min_{z} \sum_{s \in \mathcal{S}} \sum_{u \in \mathcal{U}} z_{s,u}\omega(s, u), \quad (8a)$$

$$\text{s.t. } \sum_{s \in \mathcal{S}} \sum_{u \in \mathcal{U}} z_{s,u}\phi(s, u) \leq C_t, \quad (8b)$$

$$\sum_{s \in \mathcal{S}} \sum_{u \in \mathcal{U}} z_{s,u} = 1, \quad (8c)$$

$$\sum_{u' \in \mathcal{U}} z_{s',u'} = \sum_{s \in \mathcal{S}} \sum_{u \in \mathcal{U}} z_{s,u}\ p(s'|s, u), \quad \forall s', \quad (8d)$$

with $z_{s,u} \geq 0, \forall s, u$, and where $z = \{z^*_{s,u}\}_{s \in \mathcal{S}, u \in \mathcal{U}}$. The equality constraints (8c) and (8d) enforce the variables $z_{s,u}$ to be a valid steady-state for the transition probability set. The optimal randomized past-independent policy μ^* obtained by solving the above LP, then, is $\mu^*(u|s) = z_{s,u} / \sum_{w \in \mathcal{U}} z_{s,w}$.

Assumption 1: Problem (8) has at least one feasible solution. ◀.

In the next section, we propose a distributed approach to the solution of problem (8).

III. UNCOORDINATED-ACTION FRAMEWORK

The formulation of the optimization problem and the found policy presume the existence of a coordinator. In fact, not only the LP in Eq. (8) needs to be solved centrally by one of the terminals or a control unit, but the action in each time slot needs to be centrally selected and then broadcast to the all cognitive terminals in the very short lapse of time between state identification and action deployment. These drawbacks limit the practical implementation of reactive cognitive networks in real-world networks. In this section, we develop a new framework where the optimization problem can be solved either centrally or distributively, an whose solution does not presume coordination of actions in each time slot. Instead, we assume that the cognitive terminals only coordinate the policy, i.e., they do not coordinate the actions selected in each slot.

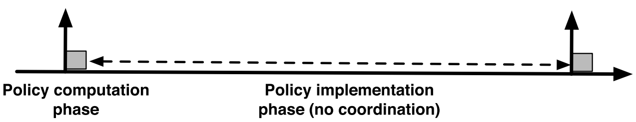

Fig. 1. In the proposed framework, the cognitive terminals periodically coordinate to compute the optimal policy and, then, independently select transmission actions in the individual slots.

Fig. 1 illustrates the operations of the cognitive terminals in the proposed framework: the cognitive terminals periodically coordinate to compute the optimal action distribution and, then, independently select transmission actions in the individual slots. Note that the action distribution needs to be recomputed only if the statistics of the stochastic process capturing the network's dynamics change. As shown later in the paper, the policy computation phase only requires the exchange of a small number of control packets. We refer to the policies where the secondary users coordinate, and do not coordinate actions, in each slot as *coordinated-action* and *uncoordinated-action* policy, respectively.

A. Uncoordinated-Action Policy

The *key property* exploited when designing the uncoordinated-action policy is the conditional *independence* of actions $(u_1, \ldots, u_N)$ of secondary users when they are in *state* $S = s$. In other words, we require the cognitive terminals to independently select the actions. This allows us to split the decision variable $z_{s,u}$ of problem (8) as follows:

$$z_{s,u} = \Pr\left(S = s,\ U = (u_1, \ldots, u_N)\right) \tag{9}$$

$$= \Pr(S = s) \Pr\left(U = (u_1, \ldots, u_N) \mid S = s\right) \tag{10}$$

$$= \Pr(S = s) \prod_{j=1}^{N} \Pr\left(U_i = u_i \mid S = s\right) = \prod_{j=1}^{N} z_j^{s,u_j}, \tag{11}$$

where the equalities above follow directly from basic probability theory and

$$z_j^{s,u_j} = y_j^s \Pr\left(U_i = u_i \mid S = s\right), \quad j = 1, \ldots, N \tag{12}$$

for some positive y_j^s, $j = 1, \ldots, N$ that conforms to

$$\prod_{k=1}^{N} y_k^s = \Pr(S = s). \tag{13}$$

Let us next see, how the independence property, consequently the splitting of $z_{s,u}$ can be used to decompose the *coordinated-action* policy $\mu(u \mid s)$. Recall that $\mu(u \mid s) = z_{s,u} / \sum_{w \in \mathcal{U}} z_{s,w}$, where $\{z_{s,u}\}_{s \in \mathcal{S}, u \in \mathcal{U}}$ is a feasible point of (8). With some straightforward manipulations, together with the conditional independence property of actions, the equality in Eq. (14), as shown at the bottom of this page, can be derived [see Appendix I], where $\mu_j(u_j \mid s)$ is the individual cognitive terminal's strategies, which we define as the *uncoordinated-action policy*. In particular, as shown in Appendix I, we can compute $\mu_j(u_j \mid s)$ as follows:

$$\mu_j(u_j \mid s) = \frac{z_j^{s,u_j}}{\sum_{w_k \in \mathcal{U}_j} z_j^{s,w_k}}. \tag{15}$$

Given the intrinsic difficulty of a coordinated-action policy, in the following we focus on determining an uncoordinated-action policy, which entails distributed optimization.

B. Distributed Optimization Problem

From (15), we see that the computation of the uncoordinated-action policies $\{\mu_j(u_j \mid s)\}_{s \in \mathcal{S}, u_j \in \mathcal{U}_j}$ of jth secondary user requires the knowledge of $\{z_j^{s,u_j}\}_{s \in \mathcal{S}, u_j \in \mathcal{U}_j}$. Thus, we simply use the splitting of $z_{s,u}$, given by $z_{s,u} = \prod_{j=1}^{N} z_j^{s,u_j}$, for all $s \in \mathcal{S}$ and $u \in \mathcal{U}_1 \times \cdots \times \mathcal{U}_N$ [compare with (11)]. Accordingly, the objective function $\sum_{s \in \mathcal{S}} \sum_{u \in \mathcal{U}} z_{s,u} \omega(s,u)$ of problem (8) becomes

$$\sum_{s \in \mathcal{S}} \sum_{u_1 \in \mathcal{U}_1} \cdots \sum_{u_N \in \mathcal{U}_N} \prod_{j=1}^{N} z_j^{s,u_j} \omega(s,u). \tag{16}$$

To simplify the notation, we denote by z_j^s the vector of probability values associated with state s and secondary user j, i.e., $z_j^s = (z_j^{s,u_j})_{u_j \in \mathcal{U}_j}$. Moreover, we denote by z_j the vector of probability values associated with secondary user j, i.e., $z_j = (z_j^s)_{s \in \mathcal{S}}$. Finally, we let the decision variables of the optimization, which recall are probability values, and precisely $z = (z_j)_{j=1,\ldots,N}$ is the steady-state state-action probabilities. It follows that optimization problem (4) can be equivalently expressed as

$$z = \arg\min_{z} \sum_{s \in \mathcal{S}} \sum_{u_1 \in \mathcal{U}_1} \cdots \sum_{u_N \in \mathcal{U}_N} \omega(s,u) \prod_{j=1}^{N} z_j^{s,u_j} \tag{17a}$$

$$\text{s.t. } \frac{1}{C_t} \sum_{s \in \mathcal{S}} \sum_{u_1 \in \mathcal{U}_1} \cdots \sum_{u_N \in \mathcal{U}_N} \phi(s,u) \prod_{j=1}^{N} z_j^{s,u_j} \leq 1 \tag{17b}$$

$$\sum_{s \in \mathcal{S}} \sum_{u_1 \in \mathcal{U}_1} \cdots \sum_{u_N \in \mathcal{U}_N} \prod_{j=1}^{N} z_j^{s,u_j} = 1 \tag{17c}$$

$$\frac{\sum_{s' \in \mathcal{S}} \sum_{y_1 \in \mathcal{U}_1} \cdots \sum_{y_N \in \mathcal{U}_N} p(s'|s,y) \prod_{j=1}^{N} z_j^{s',y_j}}{\sum_{u_1 \in \mathcal{U}_1} \cdots \sum_{u_N \in \mathcal{U}_N} \prod_{j=1}^{N} z_{s,u_j}^{j}} = 1, \tag{17d}$$

$$0 \leq z_j^{s,u_j} \leq 1, s \in \mathcal{S}, j = 1, \ldots, N, u_j \in \mathcal{U}_j, \tag{17e}$$

$$\mu(u \mid s) = \underbrace{\frac{\Pr\left(U_1 = u_1 \mid S = s\right)}{\sum_{w_1 \in \mathcal{U}_1} \Pr\left(U_1 = w_1 \mid S = s\right)}}_{\mu_1(u_1 \mid s)} \times \cdots \times \underbrace{\frac{\Pr\left(U_N = u_N \mid S = s\right)}{\sum_{w_N \in \mathcal{U}_N} \Pr\left(U_N = w_N \mid S = s\right)}}_{\mu_N(u_N \mid s)} = \prod_{j=1}^{N} \mu_j(u_j \mid s). \tag{14}$$

$$\omega_j^s(\bar{z}_j, u_j, \bar{u}_j) = \sum_{u_1 \in \mathcal{U}_1} \cdots \sum_{u_{j-1} \in \mathcal{U}_{j-1}} \sum_{u_{j+1} \in \mathcal{U}_{j+1}} \cdots \sum_{u_N \in \mathcal{U}_N} \omega(s, u_j, \bar{u}_j) \prod_{n \neq j} z_n^{s,u_n}, \tag{20}$$

$$\phi_j^s(\bar{z}_j, u_j, \bar{u}_j) = \sum_{u_1 \in \mathcal{U}_1} \cdots \sum_{u_{j-1} \in \mathcal{U}_{j-1}} \sum_{u_{j+1} \in \mathcal{U}_{j+1}} \cdots \sum_{u_N \in \mathcal{U}_N} \phi(s, u_j, \bar{u}_j) \prod_{n \neq j} z_n^{s,u_n}, \tag{21}$$

$$Z_j^s(\bar{z}_j, \bar{u}_j) = \sum_{u_1 \in \mathcal{U}_1} \cdots \sum_{u_{j-1} \in \mathcal{U}_{j-1}} \sum_{u_{j+1} \in \mathcal{U}_{j+1}} \cdots \sum_{u_N \in \mathcal{U}_N} \prod_{n \neq j} z_n^{s,u_n}, \tag{22}$$

$$P_j^{s's}(\bar{z}_j, u_j, \bar{u}_j) = \sum_{u_1 \in \mathcal{U}_1} \cdots \sum_{u_{j-1} \in \mathcal{U}_{j-1}} \sum_{u_{j+1} \in \mathcal{U}_{j+1}} \cdots \sum_{u_N \in \mathcal{U}_N} p(s'|s, u_j, \bar{u}_j) \prod_{n \neq j} z_n^{s,u_n}. \tag{23}$$

where the optimization vector is z. One of the classic method for decentralizing optimization problems is the decomposition technique [31]. However the specification of the new variables gives an optimization problem which does not necessarily posses structural properties (e.g., *separability*) favorable for classic distributed optimization [31, Sec. III]. Therefore, decomposition techniques do not apply for efficiently distributing the solution method for Problem (17). Nevertheless, in the sequel we propose a method inspired from alternating convex optimization techniques for decentralizing the solution method for the problem. The penalty is an increase in the required message exchanges compared to that of classic decomposition techniques. We remark that the optimization procedure can be performed centrally if needed.

C. Alternating Convex Optimization Approach

Based on optimization problem (17), we derive an algorithm for the distributed computation of the vector of steady-state state-action probabilities z. We denote by u_{-j} the vector obtained by eliminating the component j of u, i.e.,

$$u_{-j} = (u_1, \ldots, u_{j-1}, u_{j+1}, \ldots, u_N). \tag{18}$$

We also denote the overall action u with (u_j, u_{-j}) to distinguish the individual cognitive terminal's action u_j From the rest of the cognitive terminals' actions. Coherently, we use $\omega(s, u) \triangleq \omega(s, u_j, u_{-j})$ and $p(s'|s, u) \triangleq p(s'|s, u_j, u_{-j})$ for all $j = 1, \ldots, N$. In a similar manner, we denote by z_{-j} the vector obtained by eliminating z_j from z, that is,

$$z_{-j} = (z_1, \ldots, z_{j-1}, z_{j+1}, \ldots, z_N). \tag{19}$$

To formally establish the algorithm, it is also useful to define the functions in Eq. (20), (21), (22) and (23), as shown at the top of this page.

The key idea of the algorithm is that the optimization is performed with respect to different subsets of variables in the vector z, while the others are fixed. This strategy is based on the alternating convex optimization techniques [32]. For the reactive cognitive network considered herein, each cognitive terminal j optimizes with respect to the associated set of variables z_j, while considering z_{-j} fixed. Then, the optimization token is passed to another cognitive terminal that optimizes with respect to its own action distribution with all the others fixed and so on. Each cognitive terminal performs the optimization step multiple times, as every iteration is a descent step. The algorithm is executed until there is no significant modification in the optimization variables, namely until there is not a relevant policy modification. Interestingly, the optimization problem in each iteration is a Linear Program that can be solved efficiently by the individual cognitive terminal. For technical reasons, that will be discussed in Section IV, a regularized alternate convex optimization approach (RCOA) is considered, where instead we consider the problem

$$\begin{aligned} &\text{minimize } f(z_1, z_{-1}) + \textstyle\sum_{j=1}^N \rho ||z_j - y_j||_2^2 && \text{(24a)}\\ &\text{subject to } \texttt{constraint (17b)} && \text{(24b)}\\ &\qquad \texttt{constraint (17c)} && \text{(24c)}\\ &\qquad \texttt{constraint (17d)} && \text{(24d)}\\ &\qquad \texttt{constraint (17e)} && \text{(24e)} \end{aligned}$$

where $\rho > 0$ is called the *regularization coefficient*,

$$f(z_1, z_{-1}) = \sum_{s \in \mathcal{S}} \sum_{u_1 \in \mathcal{U}_1} \cdots \sum_{u_N \in \mathcal{U}_N} \omega(s, u) \prod_{j=1}^N z_j^{s,u_j}, \tag{25}$$

and the optimization variables are $z = (z_1, z_{-1}) = (z_1, \ldots, z_N)$ and $y = (y_1, \cdots, y_N)$.

Note that Problem (24) is identical to Problem (17a) except that (24) has additional optimization variables y and its objective function contains an additional quadratic term. The two problems are related in the sense that for every solution $z^\star$ of Problem (17a), $(z^\star, z^\star)$ is a solution of (24), and for every solution $(z^\star, y^\star)$ of (24), $z^\star$ is a solution for Problem (17a). Application of alternating convex optimization to Problem (24) yields an algorithm, where the optimization problem in each iteration becomes a quadratic program (QP), which is still solvable efficiently by the individual cognitive terminal.

The steady-state state action distribution vector evolves along iterations, where we let $z(k)$ be such a vector after the k-th iteration. The algorithm can be summarized as follows in Algorithm 1.

Step 1) initializes RACOA. Note that the initialization can be done in an fully decentralized manner, so that each cognitive terminal can select the value of z_j according to some probability distribution independently of the other cognitive terminals. However, we assume that the value of ρ to be used is initially agreed among the cognitive terminals. An important question is the value of the regularization coefficient ρ to be used. If the original problem is convex, convexity properties (e.g., use of Lagrange multipliers and zero duality results) can be exploited when characterizing ρ,

Algorithm 1 Regularized Alternate Convex Optimization Approach (RACOA)

1) Given a feasible point $z(0) = (z_1(0), \ldots, z_N(0))$, regularization index $\rho > 0$, and an accuracy level ε. Set cognitive terminal index $j = 1$, iteration index $k = 1$, and objective value $f(k) = \infty$.
2) For fixed state-action distribution $z_{-j}(k) = (z_1(k), \ldots, z_{j-1}(k), z_{j+1}(k-1), \ldots, z_N(k-1))$, the cognitive terminal j solves the QP

$$\min_{z_j} \ f_j^k(z_j) \triangleq \sum_{s\in\mathcal{S}} \sum_{u_j\in\mathcal{U}_j} \omega_j^s(z_{-j}(k), u_j, u_{-j}) z_j^{s,u_j} + \rho||z_j - z_j(k-1)||_2^2 \tag{26a}$$

$$\text{s.t.}\ \frac{1}{C_t} \sum_{s\in\mathcal{S}} \sum_{u_j\in\mathcal{U}_j} \omega_j^s(z_{-j}(k), u_j, u_{-j}) z_j^{s,u_j} \le 1 \tag{26b}$$

$$\sum_{s\in\mathcal{S}} \sum_{u_j\in\mathcal{U}_j} Z_j^s(z_{-j}(k), u_{-j}) z_j^{s,u_j} = 1 \tag{26c}$$

$$\frac{\sum_{s'\in\mathcal{S}} \sum_{u_j\in\mathcal{U}_j} P_j^{s's}(z_{-j}(k), u_j, u_{-j}) z_j^{s',u_j}}{\sum_{u_j\in\mathcal{U}_j} Z_j^s(z_{-j}(k), u_{-j}) z_j^{s,u_j}} = 1, \quad s\in\mathcal{S} \tag{26d}$$

$$0 \le z_j^{s,u_j} \le 1,\ s \in \mathcal{S},\ u_j \in \mathcal{U}_j \tag{26e}$$

where the optimization variable is only z_j and not the overall state-action vector distribution z. Denote the solution of problem (26) by $z_j^\star$. Set $z_j(k) = z_j^\star$.
3) If $j = N$, set $k = k + 1$ and go to step 4. Otherwise, set $j = j + 1$ and go to step 2).
4) Stopping criterion: If $|f_N^k(z_N^\star) - f(k-1)| < \varepsilon$, STOP. Otherwise, set $f(k) = f_N^k(z_N^\star)$, $z(k) = z(k-1)$, $j = 1$ and go to step 2.

e.g., [33, Th. 4.2]. However, when the problem becomes nonconvex, the mechanisms used in the convex settings do not apply anymore. However, the choice of ρ is justified in the following sense: ρ is directly linked with the curvature of the objective function of problem (26). Therefore, an undesired ill-conditioning can be forced upon the QP (26) routines if a large ρ is selected., and a common practice is to choose a moderate value of ρ, possibly based on preliminary experiments, see [34, p. 123].

In Step 2), the secondary user 1 solves the QP (26) to compute z_1. It is recommended to constraint components of z_{-j} to be sufficiently small (see Step 1) so that QP (26) is feasible (in particular the constraint (26b) for the cognitive terminal 1).[2] In step 3), the QP (26) is sequentially solved among other cognitive terminals in a round robin fashion. It is worth pointing out that after the QP is solved by the cognitive terminal 1, all the remaining alternating optimizations of QP (26), remain feasible. Step 4) checks for a stopping criterion, and repeats the round of optimization if the criterion is not satisfied. If $\rho > 0$, the RACOA terminates for any nonzero accuracy level specified by ε, as we will show next. We observe that the token passing optimization procedure may be subject to failure if a "token" message is lost. However, the token passing mechanisms used here is more flexible than that traditionally used in wired and wireless networks. In fact, the optimization order is not strictly rigid, and in case of lost token message the current terminal could perform another optimization round and, then, advertise a new policy. Additionally, whereas in token passing-based channel access the time needed to rebuild the ring directly corresponds to wasted transmission opportunities, in the considered scenario it would correspond to a suboptimal policy used until termination.

[2]Alternatively, the cognitive terminal 1 can scale z_{-j} appropriately and communicate the scaling factor to the next cognitive terminal.

IV. Convergence Properties of RACOA

In this section we investigate the convergence properties of our proposed algorithm RACOA. Some definitions and terminologies, which are technically necessary to characterize the convergence properties, are given in Appendix II.

A. Global Convergence and Convergent Point

Proposition 2: Let $f(k)$ be computed according to step (4) of RACOA. Then for every $\varepsilon > 0$, there is an integer K such that $|f(k) - f(k-1)| < \varepsilon$ if $k \ge K$. □

Proof: The objective function value of problem (24) is bounded below, continuous over all its feasible z, and $\rho > 0$, from [35, Lemma 2.2], we get

$$f(0) - f(n) \ge \sum_{k=1}^{n} \rho||z(k) - z(k-1)||_2^2, \tag{27}$$

where $z(k) = (z_1(k), \ldots, z_N(k))$. Note that both $f(0)$ and $f(n) = f_N^k(z_N(n))$ are bounded below, and therefore the partial sum $\sum_{k=1}^n \rho||z(k) - z(k-1)||_2^2 \triangleq y_n$ is bounded above for all $n \in \mathbb{N}$. Thus, [36, Th. 3.24] ensures that $\{y_n\}_{n\in\mathbb{N}}$ converges. From [36, Th. 3.23], we conclude that

$$||z(k) - z(k-1)||_2^2 \to 0. \tag{28}$$

Here we emphasize the importance of the use of equivalent regularized Problem (24) instead of (17a), which in turn ensure the convergence of $\{z(k) - z(k-1)\}_{k\in\mathbb{N}}$. As a result we have

$$\forall\epsilon > 0, \exists m \in \mathbb{N}, \quad \text{s. t. } k{\ge}m \Rightarrow |z_N(k) - z_N(k-1)| < \epsilon. \tag{29}$$

It is not difficult to show that the objective function $f_j^k(z_j)$ of problem (26) is Lipschitz continuous on $\mathcal{X}_j$ with some constant L_j, where $\mathcal{X}_j$ is the feasible set of problem (26). Therefore, we have

$$k \ge m \Rightarrow |f_N^k(z_N(k)) - f_N^{k-1}(z_N(k-1))| < L_N\epsilon. \tag{30}$$

Moreover, note that $f(k) = f_N^k(z_N(k))$ and $f(k-1) = f_N^{k-1}(z_N(k-1))$ [compare with step (4) of *Algorithm 1*]. This together with (30) ensures the convergence, since ϵ was arbitrary. ■

Note that (28) ensures $\lim_{k\to\infty}(z(k) - z(k-1)) = 0$, though there is no guarantees that the sequence $\{z(k)\}_{k\in\mathbb{N}}$

will converge. However, since $\{z(k)\}_{k\in\mathbb{N}}$ is a bounded sequence in $\mathbb{R}^N$, it always has convergent subsequences [36, Th. 3.6-(b)]. Let us next formally present the characteristics of these convergence points.

Proposition 3: Let $\{z(k_n)\}_{n\in\mathbb{N}}$ be a subsequence of $\{z(k)\}_{k\in\mathbb{N}}$ computed from RACOA with $\lim_{n\to\infty} z(k_n) = \bar{z}:=(\bar{z}_j, \bar{z}_{-j})$. The limit point $\bar{z}$ is a *coordinatewise minimum* (see Definition 6, Appendix II) of the objective function of problem (17a) □

Proof: We start by introducing some compact notations for clarity. Without loss of generality, we denote by $\{z(k)\}_{k\in\mathbb{N}}$ the convergent subsequence $\{z(k_n)\}_{n\in\mathbb{N}}$. Recall that the objective function (26a) is linear in z_j and has the form $a_j(z_{-j}(k))^\mathsf{T} z_j$, where $a_j(z_{-j}(k))$ is a column vector of appropriate length that can be constructed in a straight forward manner [compare with (20)]. Moreover, let the constraints (26b), (26c), and (26d) be expressed as $b_j(z_{-j}(k))^\mathsf{T} z_j \leq 1$, $c_j(z_{-j}(k))^\mathsf{T} z_j = 1$, and $D_j(z_{-j}(k))z_j = \mathbf{0}$, respectively, where $b_j(z_{-j}(k))$, $c_j(z_{-j}(k))$ are a column vector of appropriate length and $D_j(z_{-j}(k))$ is a matrix of appropriate size. Note that for all $j \in \{1, \ldots, N_s\}$, the vector valued functions $b_j(z_{-j})$, $c_j(z_{-j})$, and the matrix valued function $D_j(z_{-j})$ are continuous functions of z_{-j} over $\mathbb{R}^{m_{-j}}$, where m_{-j} is the number of components of z_{-j}. Moreover, note that the constraint (26e) has the compact form $\mathbf{0} \leq z_j \leq \mathbf{1}$. Therefore, the feasible set of problem (26) is given by $\mathcal{X}_j(z_{-j}(k))$, where

$$\mathcal{X}_j(z_{-j}) = \{z_j \mid b_j(z_{-j})^\mathsf{T} z_j \leq 1,\ c_j(z_{-j})^\mathsf{T} z_j = 1, \\ D_j(z_{-j})z_j = \mathbf{0}, \mathbf{0} \leq z_j \leq \mathbf{1}\},\ \ j \in \{1, \ldots, N\}. \quad (31)$$

From Lemma 11 (see Appendix III), we have

$$\lim_{k\to\infty} \mathcal{X}_j(z_{-j}(k)) = \mathcal{X}_j(\bar{z}_{-j}), \quad j \in \{1, \ldots, N\}. \quad (32)$$

We refer the reader to Appendix II, Definition 8 for formal set limit definitions. This combined with [35, Th. 2.3] yields the result, which we outline next for the completeness.

The set convergence (31) ensures that $\forall y_j \in \mathcal{X}_j(\bar{z}_{-j})$, $\forall j \in \{1, \ldots, N\}$, there exist a sequence $\{y_j(k) \in \mathcal{X}_j(z_{-j}(k))\}_{k\in\mathbb{N}}$ such that $\lim_{k\in\mathbb{N}} y_j(k) = y_j$, see Appendix II, Definition 8. Moreover, by noting that $z_j(k)$ is the minimizer of function $f_j^k(z_j)$ [see (26a)], we have $\forall j \in \{1, \ldots, N\}$

$$a_j(z_{-j}(k))^\mathsf{T} z_j(k) + \rho||z_j(k) - z_j(k-1)||_2^2 \\ \leq a_j(z_{-j}(k))^\mathsf{T} y_j(k) + \rho||y_j(k) - z_j(k-1)||_2^2, \quad (33)$$

$y_j(k) \in \mathcal{X}_j(z_{-j}(k))$. Now let $k \longrightarrow \infty$ to yield

$$a_j(\bar{z}_{-j})^\mathsf{T} \bar{z}_j \leq a_j(\bar{z}_{-j})^\mathsf{T} y_j + \rho||y_j - \bar{z}_j||_2^2, y_j \in \mathcal{X}_j(\bar{z}_{-j}). \quad (34)$$

From the inequality above we get that $\bar{z}_j$ is the minimizer of $a_j(\bar{z}_{-j})^\mathsf{T} y_j + \rho||y_j - \bar{z}_j||_2^2$ over $y_j \in \mathcal{X}_j(\bar{z}_{-j})$. This combined with that $\mathcal{X}_j(\bar{z}_{-j})$ is *convex*, we have, $a_j(\bar{z}_{-j})^\mathsf{T}(y_j - \bar{z}_j) \geq 0 \forall y_j \in \mathcal{X}_j(\bar{z}_{-j})$, which in turn ensures that

$$\bar{z}_j = \underset{y_j \in \mathcal{X}_j(\bar{z}_{-j})}{\arg\min}\ a_j(\bar{z}_{-j})^\mathsf{T} y_j. \quad (35)$$

Finally, we note that

$$f(\bar{z}_1, \ldots, \bar{z}_N) = a_j(\bar{z}_{-j})^\mathsf{T} \bar{z}_j,\ \ j \in \{1, \ldots, N\}, \quad (36)$$

where f is the objective function of problem (17a). This combined with (35) yields the required result concluding the proof. ■

Proposition 3 characterizes the properties of the subsequence limits of the sequence $\{z(k)\}_{k\in\mathbb{N}}$ generated by RACOA. The proposition claims that the subsequence limits are optimal in the sense that, just with independent minimizations performed at the secondary users, the overall cost incurred by the secondary users are not further improved. We note there is no guarantee that $z(k)$ given at the end of RACOA is a limit point. The termination of RACOA is based on the criterion $|f(k) - f(k-1) < \varepsilon$, see step 4 of the algorithm. Therefore, when RACOA terminates, it simply settles on the *current* $z(k)$ values, which is not necessarily a limit point. However, via extensive numerical evaluations, we empirically observed that the termination point $z(k)$ from RACOA resembles properties of a limit point. This suggests the relevance of the assertions of Proposition 2 for RACOA in practice.

It is worth pointing out that $\forall j \in \{1, \ldots, N\}$, the subsequence limits of the sequence $\{(z_j(k), z_{-j}(k))\}_{k\in\mathbb{N}}$, denoted $\bar{z}$ is a feasible point for problem (17a). This follows immediately from that the feasible set $\mathcal{X}$ of problem problem (17a) is compact and the feasible set of problem (26) is a compact subset of $\mathcal{X}$.

Note that the algorithm RACOA has considered a regularized objective function, where the regularization coefficient ρ is *positive*. Therefore, the optimization problem that is to be solved at every iteration of RACOA is a QP. A choice of $\rho = 0$ corresponds to an algorithm, where the optimization problem that is to be solved at every iteration of the algorithm is an LP. Such an algorithm may be preferred to RACOA, because solving an LP is computationally less intensive compared to solving a QP. However, such a computational gain is achieved at the cost of certain convergence properties. In particular, only the convergence of the resulting algorithm is proved and characteristics of the convergence point is not. Therefore, if one wants to trade-off convergent point characteristics with computational burden, a choice of $\rho = 0$ is suggestive ensures that the RACOA terminates. Let us now establish the convergence properties of the algorithm RACOA with a choice of $\rho = 0$.

Proposition 4: Let $f(k)$ be computed according to Step (4) of RACOA with the exception that $\rho = 0$ at Step (1). Then for every $\varepsilon > 0$, there is an integer K such that $|f(k)-f(k-1)| < \varepsilon$ if $k \geq N$. □

Proof: When $\rho = 0$ in the RACOA, we note that the sequence $\{f(k)\}_{k\in\mathbb{N}}$ is *monotonically decreasing* (compare with [36, Deinition 3.13]). Moreover, the sequence is *bounded below*. Therefore, from [36, Th. 3.14], we conclude that sequence $\{f(k)\}$ converges. Hence the sequence $\{f(k)\}$ is a Cauchy sequence [36, Th. 3.11]. More specifically, for every $\varepsilon > 0$, there is an integer ℓ such that $|f(k) - f(l)| < \varepsilon$ if $k \geq \ell$ and $l \geq \ell$, (compare with [36, Deinition 3.8]). This means $|f(k) - f(k-1)| < \varepsilon$ as required at the Step 4 of RACOA if $k \geq \ell + 1 = K$. ■

B. Convergence Rate

Characterizing the global convergence rate in the case of RACOA (see Proposition 3) is challenging.

Reference [35, Sec. 2.3] derived global convergence rate results for alternating optimization problems under certain conditions. However, these assertions (see [35, Th. 2.8]) are not applied for RACOA, because the feasible set of Problem (17a) is compact. In the sequel, we show *q-linear convergence* (see Definition 9, Appendix II) of RACOA, which is a direct consequence of [37, Th. 2]. It is worth pointing that the aforementioned q-linear convergence is associated to local convergence properties of RACOA, which holds if iterates get near a *local minimizer* (see Definition 10, Appendix II) of Problem (24).

Proposition 5: Let $\bar{z}$ be a local minimum of Problem (24) for which the Hessian of f evaluated at $\bar{z}$ is positive definite. Moreover let $\mathcal{X}$ denote the feasible region of Problem (24). Then $\exists\ \epsilon > 0$ such that for any $z(0) \in \mathcal{N}(\bar{z}, \epsilon) = \{z|\ \ ||z-\bar{z}||_2 < \epsilon, z \in \mathcal{X}\}$, the corresponding RACOA iteration sequence $\{z(k)\}_{k\in\mathbb{N}}$ converges q-linearly to $\bar{z}$. □

Proof: Note that the objective function f of Problem (24) is twice differentiable. In addition, in every iteration step (2) of RACOA yields a unique minimizer, because the associated objective function [see (26a)] is strictly convex. Application of [37, Th. 2] completes the proof. ■

V. Case Study and Numerical Results

In this section, we illustrate the proposed optimization framework with a case of relevant practical interest. For such a case, we then present numerical results. In order to show the application of the proposed framework to the traditional primary-secondary user scenario, we consider a network with one primary user implementing finite ARQ and N_s secondary users. In this case, the primary user's packet service consists of consecutive retransmissions that terminate with packet delivery or packet drop after F unsuccessful transmissions. After packet service is terminated, the primary user remains idle until a new packet arrives and the service of a new packet starts. The probability that a new packet service initiates in an otherwise idle slot is α. This model corresponds to a buffer with size 1 packets, where the probability that one or more packets arrive in each slot equal to α. We remark that the proposed approach can be seamlessly applied to a non-binary buffer model. An extensive discussion on the optimal policy for a case with non-binary buffer can be found in [25] and [28]. The secondary users are assumed backlogged, and in every slot t each individual secondary user makes a binary transmission/idleness decision, where $U_j(t) = 0$ and $U_j(t) = 1$ correspond to secondary user j being idle and transmitting in slot t, respectively. We remark that the proposed framework can be applied to general networks modeled as a Markov process, without any assumption regarding the number of primary users. However, a limitation could be imposed by optimization complexity and time if many primary users are active, due to state space explosion.

The state space of the primary user is $\mathcal{S} = \{0, 1, \ldots, F\}$, where $S(t) = 0$ corresponds to an idle slot and $S(t) = f$ to the f–th transmission of the packet being served in slot t. From state 0, the Markov chain tracking the primary user's state moves to 1 and remains in 0 with probability α and $1 - \alpha$, respectively. From state $S(t) = f$, $0 < f < F$, if packet decoding fails at the primary receiver, then the Markov chain moves to $S(t) = f + 1$. If packet decoding succeed, then the process moves either to state 0 or 1 depending on packet arrival. If the current state is $S(t) = F$, then packet service terminates both in case of successful or unsuccessful packet decoding, and the process returns either to state 0 or 1 depending on packet arrival (probability $1 - \alpha$ and α, respectively). Fig. 2 depicts the state space and the state transitions with non-zero probability for the considered case.

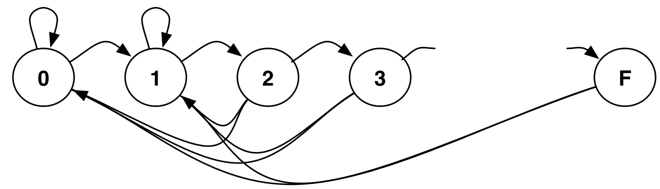

Fig. 2. State space of the primary user: 0 and $f > 0$ correspond to idle state and f–th retransmission of the packet currently being served. Arrows correspond to state transitions with non-zero probability.

For the sake of clear and readable results, we assume a symmetric network, where the packet decoding probability at the primary and secondary receivers in slot t is only a function of the number of active transmitters, determined by the primary user's state $S(t)$ and secondary user's action $U(t)$. In state $s \in \mathcal{S}$ and action vector $u = \{u_1, \ldots, u_2\}$, the number of active transmitter is $A(s, u) = \mathbb{1}(s > 0) + \sum_{j=1,\ldots,N_s} u_j$, where $\mathbb{1}(\cdot)$ is the indicator function. We, then, define as $\theta_p(A(s, u))$ and $\theta_s(A(s, u))$ the failure probability of the primary and secondary users, respectively, conditioned on the number of active transmitters.

We measure the performance of the primary user in terms of average throughput, and we define the cost function in Eq. (2)

$$\omega(s, u) = \begin{cases} \theta_p(A(s, u)) & \text{if } s > 0 \\ 1 & \text{otherwise.} \end{cases} \tag{37}$$

The performance of the secondary users is measured in terms of average aggregate throughput expressed in delivered packets per time slot. Then, we define the cost function in Eq. (3) as $\phi(s, u) = \sum_{j=1,\ldots,N_s} \phi_j(s, u)$, where

$$\phi_j(s, u) = \begin{cases} \theta_s(A(s, u)) & \text{if } u_j > 0 \\ 1 & \text{otherwise.} \end{cases} \tag{38}$$

A. Physical Layer

In order to address the wide interest on multiantenna networks, we consider nodes equipped with multiple antennas [38], [39]. The Layered Space-Time MUltiuser Detector (LASTMUD) [40] is used at the receiving nodes to provide resilience to interference by means of interference cancellation.

Assume that the receivers are equipped with arrays composed of $N_{\rm A}$ antennas. Assuming that $N_{\rm Tx}$ nodes with indices $\{1, 2, \ldots, N_{\rm Tx}\}$ are transmitting (each using 1 transmission antenna), let us define the column vector $\mathbf{s}'(k) = [s'_1(k), \ldots, s'_{N_{\rm Tx}}(k)]^T$ whose entry $s'_i(k)$ is the symbol transmitted from user i at time kT, where T is the symbol period

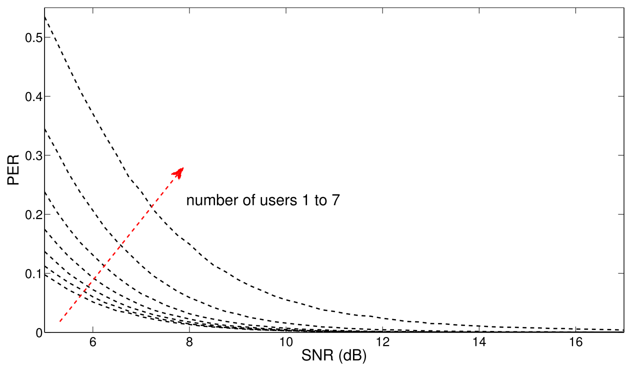

Fig. 3. Packet Error Rate (PER) as a function of the SNR. The antenna array at the receiver is composed of 4 antennas.

and ^{T} denotes the transpose operation. The power of the transmitted data signal from each transmitter i is $\sigma^2_{s'}(i) = \mathrm{E}[|s'_i(k)|^2] = P_{\mathrm{Tx}}$. For the sake of a simpler notation, we omit in the following the time index k in all signals.

The receivers use all the N_{A} antennas, and the column vector of the N_A received samples can be written as $\mathbf{r} = \mathbf{H}\mathbf{s}' + \nu'$, where ν' is the column noise vector of length N_{A}, and $\mathbf{H}$ is the $N_{\mathrm{A}} \times N_{\mathrm{Tx}}$ channel matrix whose entry $\mathbf{H}_{g,m}$ represents the complex baseband channel gain between the transmitter m and the g-th receive antenna.

In order to extract a sufficient statistics for detection, the receive node multiplies the vector of the received samples by a matrix matched to the channel. By defining the $N_{\mathrm{A}} \times N_{\mathrm{A}}$ matrix $\mathbf{R} = \mathbf{H}^H\mathbf{H}$, where ^{H} is the Hermitian operator, the obtained vector is $\mathbf{z} = \mathbf{H}^H\mathbf{r} = \mathbf{R}\mathbf{s} + \nu$, where $\nu = \mathbf{H}^H\nu'$.

The LASTMUD receiver performs the detection of the streams in stages. At each stage the stream with the highest signal to noise plus interference ratio (SNIR) is detected, and its contribution is removed from the vector $\mathbf{z}$ before the next stage [41], [42]. Further details on the LASTMUD receiver can be found in [40].

Fig. 3 depicts the Packet Error Rate (PER) of the LASTMUD receiver for a 4-element antenna array as a function of the Signal to Noise Ratio of the received signals. The the Gaussian approximation proposed in [43] is used to capture the residual noise in interference cancellation process. The packet size is set to 1000 bits.

B. Performance

In the following, the SNR of the received signals at the primary users is set to 15 dB, and the arrival rate is $\alpha = 0.8$. The quadrature coefficient used in the simulations is $\rho = 0.01$. Results not included here show that the throughput of the primary and cognitive terminals present minor variations as ρ is changed.

Fig. 4 depicts the average aggregate throughput achieved by the secondary users as a function of the SNR at the secondary users' receivers. Dashed lines correspond to the centralized solution obtained assuming that there is an ideal solver capable to compute in centralized manner the solution to optimization problem Eq. (8a), referred to as bound in the plots. Solid lines correspond to the solution of the proposed RACOA algorithm averaged over different initializations. The aggregate throughput is shown for a number of secondary users N_s equal to 2, 4 and 6. The number of antennas at the receivers is set to $N_A = 2$.

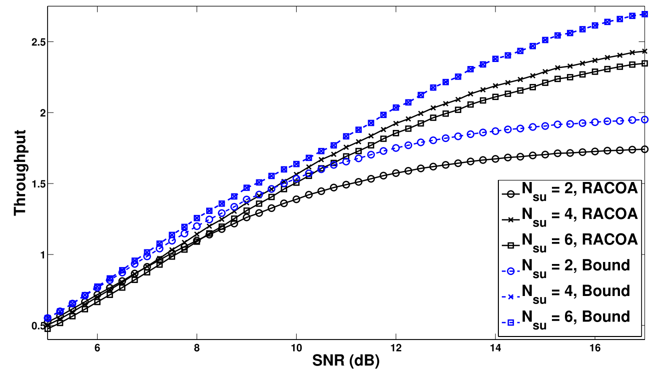

Fig. 4. Average aggregate throughput of the secondary users as a function of the SNR at the secondary users' receivers. The SNR at the primary user's receiver is set to 15 dB. The antenna array at the receiver is composed of 4 antennas.

The aggregate throughput monotonically increases as the SNR increases at the secondary users receivers. In fact, as shown in Fig. 3, a larger SNR implies that a larger number of transmitters can simultaneously access the channel achieving the same PER. It can be observed that there is a significant gap between the throughput with 2 and 4 secondary users, whereas the upper bound throughput for 4 and 6 overlap. For a given SNR regime and number of antennas, there is a specific number of transmitting users that maximizes the average number of packets delivered in a slot. The fully-coordinated implementation of the optimal policy can directly control the number of transmitters to match that number. In the proposed scheme, actions are determined by the individual users, and the overall number of transmitting users in a slot is controlled by a state-dependent distribution where sub-optimal global actions have positive probability. It can be seen that the degradation introduced by sub-optimal actions is more pronounced if the number of users present in the system exceeds the number of transmitters whose simultaneous transmission would maximize the average number of packets delivered in one slot. Finally, low SNR regimes further increase the degradation, as the number of transmitting users which maximizes the average number of delivered packets is smaller. We remark that the considerable communication overhead necessary to coordinate action in each slot would heavily penalize the performance of the centralized approach. Herein, we do not consider that overhead when computing the performance of the centralized approach to provide a comparison on the effectiveness of the transmission policy.

Fig. 5 shows the same plot where the secondary receivers' antenna arrays are composed of 4 antennas. In this case the receivers can successfully decouple a larger number of transmissions and a larger number of users corresponds to a larger throughput both in the optimal coordinated policy and in the proposed frameworks where secondary users do not coordinate actions. We emphasize that the performance of the proposed distributed solution is a function of the physical layer, as well as of the network setting.

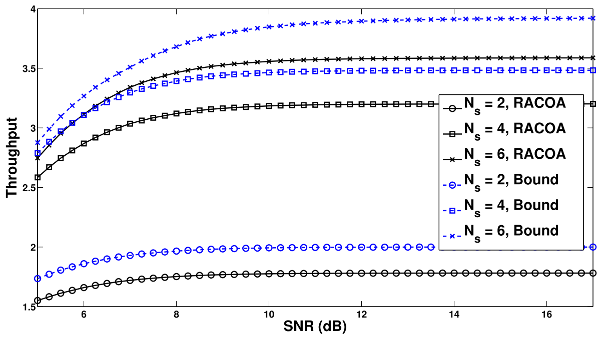

Fig. 5. Average aggregate throughput of the secondary users as a function of the SNR at the secondary users' receivers. The SNR at the primary user's receiver is set to 15 dB. The antenna array at the receiver is composed of 4 antennas.

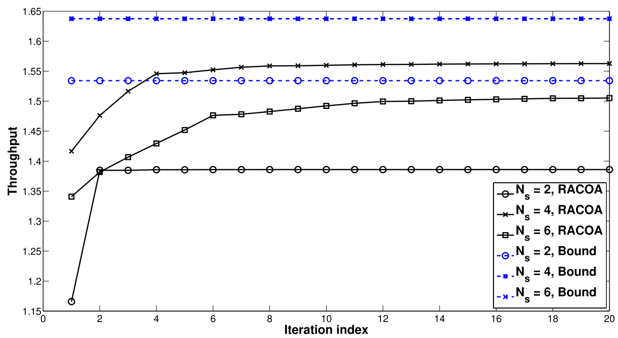

Fig. 6. Average aggregate throughput of the secondary users as a function of the iteration index. The SNR at the primary user's and secondary users' receivers is set to 15 dB. The antenna array at the receiver is composed of 4 antennas.

Fig. 6 shows the average throughput as a function of the iteration index for a setting with 2 antennas at the receivers, and where the the SNR at the receivers is set to 15 dB (primary user's receiver) and 10 dB (secondary users' receivers). The plots are averaged over different initialization points. It can be observed that the plot corresponding to 2 secondary users converges in a very small number of iterations to a stable solution, whereas convergence requires a larger number of iterations when a larger number of secondary users are present in the system. Intuitively, a larger number of secondary users results into a larger number of iterations to run the optimization problem a given number of times at each users.

VI. Conclusions

In this paper, we presented a novel framework for the distributed optimization of transmission and channel access strategies in cognitive networks. The objective of the cognitive terminals is to find the optimal *action* distribution conditioned on the current network state. In the proposed framework, the smart users distributively compute the policy, but do not coordinate actions in each slot in order to avoid a communication overhead. The algorithm is based on alternate convex optimization, where the smart users sequentially improve the current global solution. A traditional primary-secondary user scenario was considered to assess the performance of the proposed framework. Results show that the gap between the optimal centralized solution and the proposed distributed solution is a function of the number of smart users, as well as on the characteristics of the physical layer. Furthermore, the proposed algorithm presents fast converge to a stable solution. We remark that the framework presented in this paper finds application in any smart network where the policy of the cognitive terminals is a function of the network state.

Appendix I
Decomposition of the Coordinated-Action Policy

Recall that $\mu(u \mid s) = z_{s,u}/\sum_{w\in\mathcal{U}} z_{s,w}$, where $\{z_{s,u}\}_{s\in\mathcal{S},u\in\mathcal{U}}$ is a feasible point of (8). This, together with the conditional independence of actions, yields Eq. (1), as shown at the top of the next page, where where the equalities (2)-(3), as shown at the top of the next page, follow from straightforward manipulations, (4), as shown at the top of the next page, follows from (12), and (5), as shown at the top of the next page, is obtained by cancelling the common factors. The final equality follows from the definition of $\mu_j(u_j \mid s)$, the *uncoordinated-action* policy of the secondary user j. From (3)-(5), we note that

$$\mu_j(u_j \mid s) = \frac{z_j^{s,u_j}}{\sum_{w_k\in\mathcal{U}_j} z_j^{s,w_k}}. \tag{I.1}$$

Appendix II
Preliminary Definitions

Definition 6 (Coordinatewise Minimum of a Function): Let f be a scalar valued function with $\mathtt{dom}(f) \subseteq \mathbb{R}^n$ and $(x_1,\ldots,x_m)$ be a partition of $x\in\mathbb{R}^n$ such that $x_i\in\mathbb{R}^{n_i}$ and $\sum_{i=1}^m n_i = n$. Moreover, let $y_i = (0,\ldots,x_i,\ldots,0)\in\mathbb{R}^n$. We call a point $z\in\mathtt{dom}(f)$ a coordinatewise minimum of f with respect to the coordinates $1,\ldots,m$ if

$$f(z+y_i) \geq f(z), \quad \forall y_i\in\mathbb{R}^n \text{ with } (z+y_i)\in\mathtt{dom}(f)$$

for all $i\in\{1,\ldots,m\}$. ◀

Definition 7 (Distance of a Point From a Set): Let $\mathcal{X}$ be a subset of $\mathbb{R}^n$ and x denote a point in $\mathbb{R}^n$. We denote by

$$\mathtt{dis}(x,\mathcal{X}) := \begin{cases} \inf_{z\in\mathcal{X}} \|x-z\|_2 & \mathcal{X}\neq\emptyset \\ \infty & \text{otherwise} \end{cases}$$

the distance from x to $\mathcal{X}$. ◀

Definitions 6 and 7 are self-explanatory. For the completeness and the clarity of our subsequent convergence proofs, let us now formally express the convergence of sequences of sets, where we stick to the notations and the definitions of [44, Ch. 4].[3]

To this end we first denote by $\mathcal{N}_\infty$ the subsequences of $\mathbb{N}$ containing all k sufficiently large, i.e., $\mathcal{N}_\infty := \{\mathcal{N}\subseteq\mathbb{N} \mid \mathbb{N}\setminus\mathcal{N} \text{ finite}\}$. Moreover, we denote by $\mathcal{N}_\infty^\#$ all the subsequences of $\mathbb{N}$, i.e., $\mathcal{N}_\infty^\# := \{\mathcal{N}\subseteq\mathbb{N} \mid \mathcal{N} \text{ infinite}\}$.

Definition 8 (Limit of a Sequence of Subsets): Let $\{\mathcal{X}(k)\}_{k\in\mathbb{N}}$ be a sequence of subsets of $\mathbb{R}^n$. We call the set

$$\begin{aligned}
&\liminf_{k\to\infty} \mathcal{X}(k) \\
&:= \left\{x \,\middle|\, \limsup_{k\to\infty}\ \mathtt{dis}(x,\mathcal{X}(k)) = 0\right\} \\
&= \left\{x \,\middle|\, \exists\mathcal{N}\in\mathcal{N}_\infty, \exists x(k)\in\mathcal{X}(k)(k\in\mathcal{N}) \text{ with } x(k)\xrightarrow[\mathcal{N}]{} x\right\}
\end{aligned}$$

[3] We refer the reader to [44, Ch. 4] and [45, Ch. 1] for details.

$$\mu(u \mid s) = \frac{z_{s,u}}{\sum_{w\in\mathcal{U}} z_{s,w}} = \frac{z_{s,u}}{\sum_{w_1\in\mathcal{U}_1}\sum_{w_2\in\mathcal{U}_2}\cdots\sum_{w_N\in\mathcal{U}_N} z_{s,w}} \tag{1}$$

$$= \frac{\prod_{j=1}^{N} z_j^{s,u_j}}{\sum_{w_1\in\mathcal{U}_1}\sum_{w_2\in\mathcal{U}_2}\cdots\sum_{w_N\in\mathcal{U}_N} \prod_{j=1}^{N} z_j^{s,w_j}} \tag{2}$$

$$= \frac{z_1^{s,u_1}}{\sum_{w_1\in\mathcal{U}_1} z_1^{s,w_1}} \times \frac{z_2^{s,u_2}}{\sum_{w_2\in\mathcal{U}_2} z_2^{s,w_2}} \times \cdots \times \frac{z_N^{s,u_N}}{\sum_{w_N\in\mathcal{U}_N} z_N^{s,w_N}} \tag{3}$$

$$= \frac{y_s^1 \Pr\left(U_1 = u_1 \mid S = s\right)}{y_s^1 \sum_{w_1\in\mathcal{U}_1} \Pr\left(U_1 = w_1 \mid S = s\right)} \times \cdots \times \frac{y_s^N \Pr\left(U_N = u_N \mid S = s\right)}{y_s^N \sum_{w_N\in\mathcal{U}_N} \Pr\left(U_N = w_N \mid S = s\right)} \tag{4}$$

$$= \underbrace{\frac{\Pr\left(U_1 = u_1 \mid S = s\right)}{\sum_{w_1\in\mathcal{U}_1} \Pr\left(U_1 = w_1 \mid S = s\right)}}_{\mu_1(u_1 \mid s)} \times \cdots \times \underbrace{\frac{\Pr\left(U_N = u_N \mid S = s\right)}{\sum_{w_N\in\mathcal{U}_N} \Pr\left(U_N = w_N \mid S = s\right)}}_{\mu_N(u_N \mid s)} = \prod_{j=1}^{N} \mu_j(u_j \mid s), \tag{5}$$

the *inner limit* of the sequence $\{\mathcal{X}(k)\}_{k\in\mathbb{N}}$ and the set

$$\begin{aligned}\limsup_{k\to\infty} \mathcal{X}(k) &:= \left\{x \,\middle|\, \liminf_{k\to\infty}\ \mathtt{dis}(x, \mathcal{X}(k)) = 0\right\} \\ &= \left\{x \,\middle|\, \exists\mathcal{N} \in \mathcal{N}_\infty^\#, \exists x(k) \in \mathcal{X}(k)(k \in \mathcal{N}) \text{ with } x(k) \underset{\mathcal{N}}{\longrightarrow} x\right\}\end{aligned}$$

the *outer limit* of the sequence $\{\mathcal{X}(k)\}_{k\in\mathbb{N}}$, where $x(k) \underset{\mathcal{N}}{\longrightarrow} x$ designates the convergence of sequence $\{x(k)\}_{k\in\mathcal{N}}$ to x. The limit of the sequence $\{\mathcal{X}(k)\}_{k\in\mathbb{N}}$ exists if the outer and inner limit sets are equal, i.e.,

$$\lim_{k\to\infty} \mathcal{X}(k) := \limsup_{k\to\infty} \mathcal{X}(k) = \liminf_{k\to\infty} \mathcal{X}(k).$$

◀

Definition 8 is an extending concepts of *limits* and *limit points* of sequences of elements to sequences of sets. Technically, the inner limit is the set of limits of sequences $\{x(k) \in \mathcal{X}(k)\}_{k\in\mathbb{N}}$ and the outer limit is the set of limit points of sequences $\{x(k) \in \mathcal{X}(k)\}_{k\in\mathbb{N}}$, which follows immediately from the definition, compare with [45, Proposition 1.1.2].

Definition 9 (q-Linear Convergence [37]): We say a sequence $\{z(k)\}_{k\in\mathbb{N}}$ converges q-linearly to $\bar{z}$ if and only if $\exists n_0 \geq 0$, $\exists \beta \in [0, 1)$ such that $\forall k \geq n_0$ $||z(k+1) - \bar{z}||_2 \leq \beta||z(k) - \bar{z}||_2$.

Definition 10 (Local Minimum): Let $\mathcal{X}$ denote the feasible region of a constrained optimization problem P and f denote the objective function. A vector $x^\star \in \mathcal{X}$ is a local minimum of f over the set $\mathcal{X}$, if $\exists\ \epsilon > 0$ such that $f(x^\star) \leq f(x)$ for all $x \in \mathcal{X}$ with $||x - x^\star||_2 < \epsilon$.

Appendix III
Lemma 11

In this appendix, we present an intermediate result required for the proof of Proposition 3, see Section IV. Recall that, without loss of generality, we denote by $\{z(k)\}_{k\in\mathbb{N}}$ the convergent subsequence $\{z(k_n)\}_{n\in\mathbb{N}}$ to improve the notational clarity.

Lemma 11: Let $j \in \{1, \ldots, N\}$, $(\bar{z}_1, \ldots, \bar{z}_N)$ be the limit point of $\{(z_1(k), \ldots, z_j(k), z_{j+1}(k-1), \ldots, z_N(k-1))\}_{n\in\mathbb{N}}$, where $z_j(k)$ is the steady-state action frequencies computed by jth secondary user at outer loop k of RACOA. Then $\mathcal{X}_j(z_{-j}(k)) \underset{\mathbb{N}}{\longrightarrow} \mathcal{X}_j(\bar{z}_{-j})$. □

Proof: We first prove that $\lim_{k\in\mathbb{N}} \mathtt{dis}(x, \mathcal{X}_j(z_{-j}(k))) = \mathtt{dis}(x, \mathcal{X}_j(\bar{z}_{-j}))$ for all $x \in \mathbb{R}^{m_j}$, where m_j is the number of components of z_j. Then the result follows directly from [44, Corollary 4.7].

Since the problem (26) is feasible for all RACOA iterations [compare with Assumption 1] and $\mathcal{X}_j(z_{-j}(k))$ is compact, $\mathtt{dis}(x, \mathcal{X}_j(z_{-j}(k)))$ is the optimal value of the following optimization problem:

$$\min_{z_j}\ (1/2)||z_j - x||_2^2 \tag{III.1a}$$

$$\text{s.t. } b_j(z_{-j}(k))^\mathsf{T} z_j \leq 1 \tag{III.1b}$$

$$c_j(z_{-j}(k))^\mathsf{T} z_j = 1 \tag{III.1c}$$

$$D_j(z_{-j}(k)) z_j = \mathbf{0} \tag{III.1d}$$

$$\mathbf{0} \leq z_j \leq \mathbf{1}. \tag{III.1e}$$

Let us denote by $g(\lambda, z_{-j}(k))$ the dual function of problem (III.1a), where λ is the vector of all Lagrange multipliers associated with the constraints (III.1b)-(III.1e). Strong duality holds for problem (III.1a), and thus

$$\mathtt{dis}(x, \mathcal{X}_j(z_{-j}(k))) = \max_{\lambda\in\Pi} g(\lambda, z_{-j}(k)), \tag{III.2}$$

where Π is the *closed* feasible set of the dual problem. Similarly, we can show that

$$\mathtt{dis}(x, \mathcal{X}_j(\bar{z})) = \max_{\lambda\in\Pi} g(\lambda, \bar{z}_{-j}). \tag{III.3}$$

Note that arguments that maximize the right hand sides of (III.2) and (III.3) *exist*, see [46, Proposition 5.2.2]. From the minimizer of the associated Lagrangian function over z_j, which can be computed in closed-form, one can readily conclude that $g(\lambda, z_{-j}(k))$ is continuous in $(\lambda, z_{-j}(k))$. Therefore, we have $\lim_{k\in\mathbb{N}} g(\lambda, z_{-j}(k)) = g(\lambda, \bar{z}_{-j})$ for all $\lambda \in \Pi$. More specifically, $g(\lambda, z_{-j}(k)) \longrightarrow g(\lambda, \bar{z}_{-j})$ *uniformly*, see [36, Definition 7.7]. The uniform convergence ensures the desired result, i.e., $\lim_{k\in\mathbb{N}} \mathtt{dis}(x, \mathcal{X}_j(z_{-j}(k))) = \mathtt{dis}(x, \mathcal{X}_j(\bar{z}))$, see Lemma 12. This concludes the proof. ■

Lemma 12: Let $\{g_k\}_{k\in\mathbb{N}}$ be a sequence of *concave* functions that converges *uniformly* on Π to a function g.

Moreover, suppose for all k and for all $\lambda \in \Pi$, $g_k(\lambda)$ is bounded above and the set $\arg\max_{\lambda\in\Pi} g_k(\lambda)$ is nonempty. Then $\lim_{k\in\mathbb{N}} \sup_{\lambda\in\Pi} g_k(\lambda) = \sup_{\lambda\in\Pi} g(\lambda)$. □

Proof: Let λ_k denote the maximizer of $g_k(\lambda)$ over Π, i.e., $g_k(\lambda_k) = \sup_{\lambda\in\Pi} g_k(\lambda)$. We start by showing that the sequence $\{g_k(\lambda_k)\}_{k\in\mathbb{N}}$ converges, i.e., it has *only one* limit point.

Suppose to the contrary that $\{g_k(\lambda_k)\}_{k\in\mathbb{N}}$ has at least two distinct limit points denoted $\acute{g}$ and $\grave{g}$, with $|\acute{g} - \grave{g}| = \delta > 0$. By the properties of limit points we have

$$\forall\epsilon > 0, \quad \forall N \in \mathbb{N}, \exists n \geq N \text{ such that } |g_n(\lambda_n) - \grave{g}| \leq \epsilon, \tag{III.4}$$

$$\forall\epsilon > 0, \quad \forall N \in \mathbb{N}, \exists m \geq N \text{ such that } |g_m(\lambda_m) - \acute{g}| \leq \epsilon. \tag{III.5}$$

Moreover, without loss of generality we can assume that $g_m(\lambda_m) \geq g_n(\lambda_n)$. By combining (III.4)-(III.5), we get $\forall\epsilon \in (0, \delta), \forall N \in \mathbb{N}, \exists n, m \geq N$ such that $|g_n(\lambda_n) - \grave{g}| \leq \epsilon/2$ and $|g_m(\lambda_m) - \acute{g}| \leq \epsilon/2$, which implies

$$\begin{aligned}
\epsilon &\geq |(g_m(\lambda_m) - g_n(\lambda_n)) + (\grave{g} - \acute{g})| &\text{(III.6)}\\
&\geq |\grave{g} - \acute{g}| - |g_m(\lambda_m) - g_n(\lambda_n)| &\text{(III.7)}\\
&\geq |\grave{g} - \acute{g}| - |g_m(\lambda_m) - g_n(\lambda_m)| &\text{(III.8)}\\
&= \delta - |g_m(\lambda_m) - g_n(\lambda_m)|, &\text{(III.9)}
\end{aligned}$$

where (III.6) and (III.7) follow from the triangular inequality, and (III.8) follows from that $g_m(\lambda_m) \geq g_n(\lambda_n)$ and $g_n(\lambda_n) \geq g_n(\lambda_m)$ [λ_n is the maximizer of g_n]. Therefore, from (III.9) and by making ϵ arbitrarily small, we have $\forall N \in \mathbb{N}, \exists n, m \geq N$ such that $|g_m(\lambda_m) - g_n(\lambda_m)| \geq \delta$. This contradicts the fact that $|g_m(\lambda_m) - g_n(\lambda_m)|$ can be made arbitrarily small by choosing n, m sufficiently large [36, Th. 7.8]. Therefore, $\{g_k(\lambda_k)\}_{k\in\mathbb{N}}$ has *only one* limit point. Let us denote by $\hat{g}$ the associated limit, i.e., $\lim_{k\in\mathbb{N}} g_k(\lambda_k) = \hat{g}$. To complete the proof, we finally show that $\hat{g} = \sup_{\lambda\in\Pi} g(\lambda)$.

Because g_k is concave and λ_k is the maximizer of $g_k(\lambda)$ over Π, it is not difficult to see that

$$g_k(\lambda_k) \geq g_k(\lambda), \quad \forall\lambda \in \Pi. \tag{III.10}$$

Now let $k \longrightarrow \infty$ to yield

$$\hat{g} \geq g(\lambda), \quad \forall\lambda \in \Pi. \tag{III.11}$$

Moreover, the definition of $\hat{g}$ claims

$$\forall\epsilon > 0, \exists L \in \mathbb{N}, \quad \forall i, i \geq L \Longrightarrow |g_i(\lambda_i) - \hat{g}| \leq \epsilon. \tag{III.12}$$

Furthermore, the definition of uniform convergence [36, Definition 7.7], claims

$$\forall\epsilon > 0, \exists M \in \mathbb{N}, \quad \forall i, i \geq M \Longrightarrow |g_i(\lambda) - g(\lambda)| \leq \epsilon, \tag{III.13}$$

for all $\lambda \in \Pi$. Let $N = \max\{L, M\}$. By combining (III.12)-(III.13), we have $\forall\epsilon > 0, \exists N \in \mathbb{N}, \forall i, i \geq N \Longrightarrow |g_i(\lambda_i) - \hat{g}| \leq \epsilon/2$ and $|g_i(\lambda_i) - g(\lambda_i)| \leq \epsilon/2$, which implies

$$\begin{aligned}
\epsilon &\geq |(g_i(\lambda_i) - \hat{g}) + (g(\lambda_i) - g_i(\lambda_i))| &\text{(III.14)}\\
&= |g(\lambda_i) - \hat{g}|, &\text{(III.15)}
\end{aligned}$$

where (III.14) follows from the triangular inequality. As $\epsilon > 0$ was chosen arbitrarily, we conclude that $g(\lambda) = \hat{g}$ for some $\lambda \in \Pi$. This combined with (III.11) yields $\hat{g} = \sup_{\lambda\in\Pi} g(\lambda)$, which concludes the proof. ■

References

[1] J. Mitola, "Cognitive radio: An integrated agent architecture for software defined radio," Ph.D. dissertation, Royal Inst. Technol., Stockholm, Sweden, 2000.

[2] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.

[3] S. Geirhofer, L. Tong, and B. M. Sadler, "Dynamic spectrum access in the time domain: Modeling and exploiting white space," *IEEE Commun. Mag.*, vol. 45, no. 5, pp. 66–72, May 2007.

[4] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proc. IEEE*, vol. 97, no. 5, pp. 894–914, Apr. 2009.

[5] W. Zhang and U. Mitra, "A spectrum-shaping perspective on cognitive radio: Uncoded primary transmission case," in *Proc. IEEE ISIT*, Toronto, ON, Canada, Jul. 2008, pp. 1338–1342.

[6] Y. Xing, C. N. Mathur, M. A. Haleem, R. Chandramouli, and K. P. Subbalakshmi, "Dynamic spectrum access with QoS and interference temperature constraints," *IEEE Trans. Mobile Comput.*, vol. 6, no. 4, pp. 423–433, Apr. 2007.

[7] Y. Chen, Q. Zhao, and A. Swami, "Joint design and separation principle for opportunistic spectrum access in the presence of sensing errors," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 2053–2071, May 2008.

[8] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 3, pp. 535–547, Mar. 2000.

[9] H. O. Burton and D. D. Sullivan, "Errors and error control," *Proc. IEEE*, vol. 60, no. 11, pp. 1293–1301, Nov. 1972.

[10] M. Levorato, U. Mitra, and M. Zorzi, "Cognitive interference management in retransmission-based wireless networks," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 3023–3046, May 2012.

[11] M. Levorato, U. Mitra, and M. Zorzi, "On optimal control of wireless networks with multiuser detection, hybrid ARQ and distortion constraints," in *Proc. IEEE Int. Conf. Comput. Commun. (IEEE INFOCOM)*, Rio de Janeiro, Brazil, Apr. 2009, pp. 1926–1934.

[12] N. Michelusi, P. Popovski, O. Simeone, M. Levorato, and M. Zorzi, "Cognitive access policies under a primary ARQ process via forward-backward interference cancellation," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 11, pp. 2374–2386, Nov. 2013.

[13] N. Michelusi, P. Popovski, and M. Zorzi, "Cognitive access policies under a primary ARQ process via chain decoding," in *Proc. IEEE Inf. Theory Appl. Workshop (ITA)*, Feb. 2013, pp. 1–8.

[14] F. E. Lapiccirella, X. Liu, and Z. Ding, "Distributed control of multiple cognitive radio overlay for primary queue stability," *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 112–122, Jan. 2013.

[15] F. E. Lapiccirella, Z. Ding, and X. Liu, "Cognitive spectrum access control based on intrinsic primary ARQ information," in *Proc. IEEE Int. Conf. Commun. (IEEE ICC)*, May 2010, pp. 1–5.

[16] I. Burago and M. Levorato, "Network estimation in cognition-empowered wireless networks," *IEEE Trans. Cogn. Commun. Netw.*, vol. 1, no. 2, pp. 244–256, Jun. 2015.

[17] R. Chen, J.-M. Park, and K. Bian, "Robust distributed spectrum sensing in cognitive radio networks," in *Proc. IEEE 27th Conf. Comput. Commun. (IEEE INFOCOM)*, pp. 2190–2200, Apr. 2008.

[18] Z. Fanzi, C. Li, and Z. Tian, "Distributed compressive spectrum sensing in cooperative multihop cognitive networks," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 1, pp. 37–48, Feb. 2011.

[19] H.-P. Shiang and M. van der Schaar, "Distributed resource management in multihop cognitive radio networks for delay-sensitive transmission," *IEEE Trans. Veh. Technol.*, vol. 58, no. 2, pp. 941–953, Feb. 2009.

[20] M. van der Schaar and F. Fu, "Spectrum access games and strategic learning in cognitive radio networks for delay-critical applications," *Proc. IEEE*, vol. 97, no. 4, pp. 720–740, Apr. 2009.

[21] F. Fu and M. van der Schaar, "Stochastic game formulation for cognitive radio networks," in *Proc. 3rd IEEE Symp. New Frontiers Dyn. Spectr. Access Netw. (IEEE DySPAN)*, Oct. 2008, pp. 1–5.

[22] B. Maham, P. Popovski, X. Zhou, and A. Hjorungnes, "Cognitive multiple access network with outage margin in the primary system," *IEEE Trans. Wireless Commun.*, vol. 10, no. 10, pp. 3343–3353, Oct. 2011.

[23] O. Simeone, I. Stanojev, S. Savazzi, Y. Bar-Ness, U. Spagnolini, and R. Pickholtz, "Spectrum leasing to cooperating secondary ad hoc networks," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 203–213, Jan. 2008.

[24] Y. Xu, L. Wang, C. Fischione, and V. Fodor, "Distributed spectrum leasing via vertical cooperation in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1588–1601, Feb. 2016.

[25] M. Levorato, S. Firouzabadi, and A. Goldsmith, "A learning framework for cognitive interference networks with partial and noisy observations," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3101–3111, Sep. 2012.

[26] M. Levorato, U. Mitra, and A. Goldsmith, "Structure-based learning in wireless networks via sparse approximation," *EURASIP J. Wireless Commun. Netw.*, vol. 2012, no. 1, pp. 1–15, 2012.

[27] L. Badia, M. Levorato, and M. Zorzi, "A channel representation method for the study of hybrid retransmission-based error control," *IEEE Trans. Commun.*, vol. 57, no. 7, pp. 1959–1971, Jul. 2009.

[28] M. Levorato, D. O'Neill, A. Goldsmith, and U. Mitra, "Optimization of ARQ protocols in interference networks with QoS constraints," in *Proc. IEEE Int. Conf. Commun. (IEEE ICC)*, Jun. 2011, pp. 1–6.

[29] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, vol. 2, 2nd ed. Belmont, MA, USA: Athena Scientific, 2001.

[30] K. W. Ross, "Randomized and past-dependent policies for Markov decision processes with multiple constraints," *Oper. Res.*, vol. 37, no. 3, pp. 474–477, May/Jun. 1989.

[31] D. P. Palomar and M. Chiang, "A tutorial on decomposition methods for network utility maximization," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1439–1451, Aug. 2006.

[32] S. Boyd. (2007). *Sequential Convex Programming*. [Online]. Available: http://www.stanford.edu/class/ee364b/lectures/seq_slides.pdf

[33] M. P. Friedlander and P. Tseng, "Exact regularization of convex programs," *SIAM J. Optim.*, vol. 18, no. 4, pp. 1326–1350, 2007.

[34] D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*. Belmont, MA, USA: Athena Scientific, 1996.

[35] Y. Xu and W. Yin, "A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion," *SIAM J. Imag. Sci.*, vol. 6, no. 3, pp. 1758–1789, Sep. 2013.

[36] W. Rudin, *Principles of Mathematical Analysis*, 3rd ed. New York, NY, USA: McGraw-Hill, 1976.

[37] J. C. Bezdek and R. J. Hathaway, "Convergence of alternating optimization," *Neural, Parallel Sci. Comput.*, vol. 11, no. 4, pp. 351–368, 2003.

[38] K. Sundaresan, R. Sivakumar, M. A. Ingram, and T.-Y. Chang, "Medium access control in ad hoc networks with MIMO links: Optimization considerations and algorithms," *IEEE Trans. Mobile Comput.*, vol. 3, no. 4, pp. 350–365, Oct./Dec. 2004.

[39] P. Casari, M. Levorato, and M. Zorzi, "MAC/PHY cross-layer design of MIMO ad hoc networks with layered multiuser detection," *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4596–4607, Nov. 2008.

[40] S. Sfar, R. D. Murch, and K. B. Letaief, "Layered space-time multiuser detection over wireless uplink systems," *IEEE Trans. Wireless Commun.*, vol. 2, no. 4, pp. 653–668, Jul. 2003.

[41] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. IEEE ISSSE*, Pisa, Italy, Sep./Oct. 1998, pp. 295–300.

[42] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, vol. 1, no. 2, pp. 41–59, 1996.

[43] M. Levorato, S. Tomasin, P. Casari, and M. Zorzi, "Physical layer approximations for cross-layer performance analysis in MIMO-BLAST ad hoc networks," *IEEE Trans. Wireless Commun.*, vol. 6, no. 12, pp. 4390–4400, Dec. 2007.

[44] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis* (A Series of Comprehensive Studies in Mathematics), vol. 317, A. Chenciner, J. Coates, and S. Varadhan, Eds., 3rd ed. Heidelberg, Germany: Springer-Verlag, 2009.

[45] J.-P. Aubin and H. Frankowska, *Set-Valued Analysis*, vol. 3. Boston, MA, USA: Birkhäuser, 1990.

[46] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, MA, USA: Athena Scientific, 1999.

Marco Levorato (S'06–M'09) received the B.S. and M.S. degrees in electrical engineering *summa cum laude* from the University of Ferrara, Italy, in 2005 and 2003, respectively, and the Ph.D. degree in electrical engineering from the University of Padova, Italy, in 2009. He is currently an Assistant Professor in Computer Science with the University of California at Irvine. He has post-doctoral appointments with Stanford University, the University of Southern California, and the Royal Institute of Technology in Stockholm, Sweden. He is a recipient of a best paper award from the IEEE Globecom 2012, the UC Hellman Foundation Award, and has been twice nominated for the Best Young Researcher Award, Department of Information Engineering, University of Padova, Italy.

Pradeep Chathuranga Weeraddana (S'08–M'11) received the M.Eng. degree Telecommunication from the School of Engineering and Technology, Asian Institute of Technology, Thailand, in 2007, and the Ph.D. degree from the University of Oulu, Finland, in 2011. He was a Post-Doctoral Researcher with the Department of Automatic Control, School of Electrical Engineering and the ACCESS Linnaeus Center, KTH Royal Institute of Technology, Stockholm, Sweden, from 2012 to 2014. He is currently a Senior Lecturer with the Sri Lanka Institute of Information Technology, Malabe, Sri Lanka. His research interests include application of optimization techniques in various application domains, such as signal processing, wireless communications, and smart grids. He is also interested in application domains of computational topology for data analysis.

Carlo Fischione (M'05) received the Laurea degree in electronic engineering and the Ph.D. degree in electrical and information engineering from the University of L'Aquila, Italy, in 2001 and 2005, respectively. He is currently a Tenured Associate Professor with the School of Electrical Engineering, KTH Royal Institute of Technology, and the ACCESS Linnaeus Center, Stockholm, Sweden. He was with the Massachusetts Institute of Technology, Cambridge, MA, USA, in 2015, as a Visiting Professor, an Associate with Harvard University, Cambridge, MA, in 2015, a Visiting Scholar with the University of California at Berkeley, CA, USA, from 2004 to 2005, as a Research Associate from 2007 to 2008, and an Research Associate with the Royal Institute of Technology, Stockholm, Sweden from 2005 to 2007. He has co-authored over 100 publications, including a book, book chapters, international journals and conferences, and international patents. He received or co-received a number of awards, including the best paper award from the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS in 2007, the best paper awards at the IEEE International Conference on Mobile Ad-hoc and Sensor System 05 and 09 IEEE MASS 2005 and IEEE MASS 2009, the best paper award of the IEEE Sweden VT-COM-IT Chapter in 2014, the Best Business Idea awards from VentureCup East Sweden in 2010 and from Stockholm Innovation and Growth Life Science in Sweden in 2014, the Ferdinando Filauro Award from the University of LAquila, Italy, in 2003, the Higher Education Award from Abruzzo Region Government, Italy in 2004, the Junior Research Award from Swedish Research Council in 2007, the Silver Ear of Wheat Award in history from the Municipality of Tornimparte, Italy, in 2012. He is an Associate Editor of *Automatica* (Elsevier). He has chaired or served as a technical member of program committees of several international conferences. He is serving as a referee for technical journals. He is an Ordinary Member of the Academy of History Deputazione Abruzzese di Storia Patria.