Robust MIMO Cognitive Radio Systems Under Interference Temperature Constraints

Yang Yang, *Student Member, IEEE,* Gesualdo Scutari, *Senior Member, IEEE,*
Peiran Song, *Student Member, IEEE,* and Daniel P. Palomar, *Fellow, IEEE*

***Abstract*—Cognitive Radio (CR) systems are built on the coexistence of primary users (PUs) and secondary users (SUs), the latter being allowed to share spectral resources with the PUs but under strict interference limitations. However, such limitations may easily be violated by SUs if perfect SU-to-PU channel state information (CSI) is not available at the secondary transmitters, which always happens in practice. In this paper, we propose a distributed design of MIMO CR networks under global interference temperature constraints that is robust (in the worst-case sense) against SU-to-PU channel uncertainties. More specifically, we consider two alternative formulations that are complementary to each other in terms of signaling and system performance, namely: a game-theoretical design and a social-oriented optimization. To study and solve the proposed formulations we hinge on the new theory of finite-dimensional variational inequalities (VI) in the complex domain and a novel parallel decomposition technique for nonconvex sum-utility problems with coupling constraints, respectively. A major contribution of this paper is to devise a new class of distributed best-response algorithms with provable convergence. The algorithms differ in computational complexity, convergence speed, communication overhead, and achievable performance; they are thus applicable to a variety of CR scenarios, either cooperative or non-cooperative, which allow the SUs to explore the trade-off between signaling and performance.**

***Index Terms*—Cognitive Radio System, Game Theory, MIMO System, Successive Convex Approximation, Complex Variational Inequalities, Worst-Case Robust Design.**

I. Introduction

THE INCREASING demand of wireless service calls for flexible and efficient usage of the scarce radio spectrum, which is however underutilized by current fixed spectrum assignment policies. Cognitive Radio (CR) has presented a promising solution to this problem by allowing intelligent cognitive nodes to access the licensed bandwidth [1]. As a widely accepted concept [2, 3], a CR system is built on a hierarchical structure, composed of primary users (PU), the legacy spectrum holders, and secondary users (SU), the unlicensed users who are allowed to access the spectrum, provided that they do not induce too much interference at the primary receivers.

The design of secondary CR systems has been addressed in a number of works, under different assumptions, e.g., [4–14]. Current formulations mainly follow two complementary design philosophies, namely: a "system-oriented" optimization (also termed network utility maximization (NUM)) [4–8], and a "user-oriented" distributed design, mostly based on game theory [9–13]. The quality of service (QoS) of the PUs is protected by imposing interference constraints to the SUs; either local interference constraints (i.e. at the level of each SU) [9, 10, 12] or global interference constraints (i.e., on the overall interference generated by all the SUs) [4, 5, 7, 11, 13] have been adopted. Figure 1 shows an example of the importance of using global rather than local interference constraints. In this figure, we compare the interference power received by a PU along different angular directions generated by the SUs in the presence of local and global interference constraints; as a benchmark, we also plot the interference experienced by the PU in the absence of any constraints on the SUs' transmissions (e.g., as in classic iterative waterfilling algorithm (IWFA) [15]). The analysis of the figure shows that local interference constraints might be too conservative, whereas global interference constraints lead to a more flexible design and thus efficient use of the available resources.

Global interference constraints are introduced in [4, 6, 11] for CR SISO, SIMO and MISO systems. In the context of MIMO systems, the analysis is mainly limited to local interference constraints [9, 10, 12, 16], with the exception of [5], where the authors focused on the sum-rate maximization of the SUs under global interference constraints, and proposed a primal-based decomposition algorithm, whose convergence however has been observed only numerically. Moreover, in the above works, the CR system is designed under the premise of *perfect* SU-to-PU channel state information (CSI), which is however not a realistic assumption, due to, e.g., inaccurate or limited CSI at the secondary transmitters and lack of full co-operation between SUs and PUs. It is therefore of paramount importance to take the channel uncertainties explicitly into account in the system design.

Capitalizing on a norm-bounded channel uncertainty model, worst-case robust (centralized and distributed) designs for MIMO system have been proposed in [6, 7, 12, 17–20], under *local* interference constraints. To the best of our knowledge, the only paper dealing with (worst-case) *robust global* interference constraints is [7] (appeared after the initial submission

Manuscript received November 18, 2012; revised April 25, 2013. The work of Y. Yang and D. P. Palomar is supported by the Hong Kong RGC 617911. The work of G. Scutari and P. Song is supported by the National Science Foundation Grant No. CNS-1218717.

Y. Yang and D. P. Palomar are with Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology (e-mail: {yangyang, palomar}@ust.hk).

G. Scutari and P. Song are with Department of Electrical Engineering, University at Buffalo, State University of New York, Buffalo, NY, USA (e-mail: {gesualdo, peiranso}@buffalo.edu).

Digital Object Identifier 10.1109/JSAC.2013.131131.

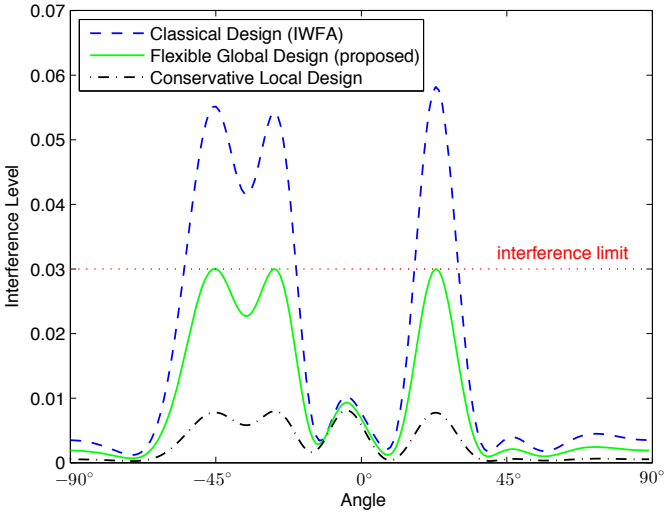


Fig. 1. Typical interference profile at a PU as a function of the angle caused by the SUs in a MIMO CR system, under local (black curve), global (green curve), and no (blue curve) interference constraints.

of this work), where the minimization of the sum-Mean-Square-Error (MSE) in MIMO systems is formulated. While promising, the algorithmic approach in [7] is however not supported by any theoretical study, leaving open the issue of the convergence of the algorithms therein. Moreover, algorithms in [7] are sequential schemes: only one SU at a time is allowed to update his transmit covariance matrix; a fact that in large scale networks may lead to excessive communication overhead and slow convergence.

The present paper considers an underlaying MIMO CR system, composed of an arbitrary number of PUs and SUs, the latter modeled as vector Gaussian interference channels (IC); inaccuracies in the SU-to-PU CSI are captured using a norm-bounded uncertainty model, and the QoS of the PU is protected by (worst-case) *robust global* interference constraints. Our interest in this setting is to develop *distributed* design solutions. To explore the trade-off between SUs' performance and (PU-to-SU and SU-to-SU) signaling, we follow two alternative design philosophies, namely: a game-theoretical and a social-oriented design. A description of these two approaches along with the main contributions of the paper is given next.

Game-Theoretical Formulation. We formulate the system design as a Nash Equilibrium Problem (NEP), where each SU competes against the others to maximize his own transmission rate subject to robust global interference constraints as well as power constraints. Global interference constraints introduce a challenge in the system design since they couple the SUs: how to enforce such constraints without requiring a centralized optimization? We address this issue by introducing a pricing mechanism in the game, through a penalization in the SUs' objective functions. The prices need to be chosen so that the robust interference constraints are satisfied at any solution of the game and a clearing condition holds; they are thus additional variables to be determined.

This formulation faces many challenges: i) the non-differentiability of the SUs' objective functions; ii) the semi-infinite attribute of the robust interference constraints; iii) the lack of boundedness of the price variables and the presence of global constraints (in the form of a clearing condition); and iv) the complex matrix nature of the optimization variables. To cope with these issues, we hinge on the theory of complex variational inequalities (VI), recently developed in [21]. Capitalizing on [21], we provide a satisfactory solution analysis of the proposed NEP (in terms of conditions for the existence and uniqueness of a solution), devise distributed asynchronous best-response algorithms, and study their convergence. We remark that the proposed framework based on complex VIs is expected to be broadly applicable to other complex NEP formulations with global constraints.

The developed solution methods are suitable for a distributed implementation, provided that a limited signaling is exchanged between PUs and SUs. There are CR scenarios wherein such a SU-to-PU signaling is either inaccurate or impossible at all. Building on consensus algorithms, we discuss how to implement our schemes when no SU-to-PU signaling is allowed. Finally, we corroborate the proposed game-theoretical model via numerical results, showing that our schemes based on robust *global* interference constraints outperform (in terms of achievable sum-rate) current decentralized state-of-the-art designs based on robust *local* interference constraints [12].

Social-oriented Design. It is well-known that a Nash Equilibrium (NE) may not be Pareto efficient. It is then important to quantify the performance loss in using game-theoretical solutions. To do that, we consider the social counterpart of the proposed NEP, that is the sum-rate maximization problem subject to the same robust global interference constraints. Since the problem is NP-hard, our goal is to develop distributed algorithms converging to stationary (possibly locally optimal) solutions. Here, the challenging issue is the nonconvexity of the sum-rate function, which prevents the application of standard primal/dual decomposition techniques [22, 23]. Indeed, under a non-zero duality gap, the convergence of primal/dual schemes is in jeopardy [5, 7]. By exploiting the successive convex approximation (SCA) framework proposed in [24], a second major contribution of this paper is to devise a first class of distributed primal and dual decomposition algorithms with *provable convergence.* In the proposed schemes, the SUs solve *in parallel* a sequence of convex optimization problems, converging to a stationary solution of the robust sum-rate maximization problem with coupling constraints. To be implemented, however, such algorithms require more PU-to-SU and/or SU-to-SU signaling than that of game-theoretical schemes. Because of that, in some scenarios they might not be implementable, making the game-theoretical schemes the only feasible and valuable choice. Numerical results show that stationary solutions yield higher sum-rates than those at the NE; for instance, in the simulated scenarios, we experience a performance gap of no more than 10%, which happens when the interference limits become the dominant constraints (i.e., much smaller than the power budget).

Overall, the paper presents two complementary system designs together with distributed solution methods, which differ in complexity, convergence speed, SU-to-SU and PU-to-SU communication overhead, and sum-rate performance; they are thus applicable to a variety of CR scenarios, either cooperative or non-cooperative, allowing the SUs to explore the trade-off between signaling and performance.

Notations: $\|\mathbf{x}\|$ is the Euclidean norm of $\mathbf{x}$ and $\|\mathbf{X}\|$ is the Frobenius norm of $\mathbf{X}$. We use $\mathbf{x}$ (or $\mathbf{X}$) as a compact notation for $\mathbf{x}_1,\ldots,\mathbf{x}_I$ (or $\mathbf{X}_1,\ldots,\mathbf{X}_I$): $\mathbf{x}=(\mathbf{x}_i)_{i=1}^I$ (or $\mathbf{X}=(\mathbf{X}_i)_{i=1}^I$). $\lambda_{\min}(\mathbf{X})$ and $\lambda_{\max}(\mathbf{X})$) denotes the smallest and largest eigenvalue of $\mathbf{X}$, respectively. $\sigma_{\max}(\mathbf{X})$ is the largest singular value of $\mathbf{X}$. $\Re(\bullet)$ is the real operator. $\langle \mathbf{x},\mathbf{y}\rangle \triangleq \mathbf{x}^T\mathbf{y}$ is the inner product of $\mathbf{x}$ and $\mathbf{y}$, and $\langle \mathbf{X},\mathbf{Y}\rangle \triangleq \Re(\mathrm{tr}(\mathbf{X}^H\mathbf{Y}))$ is the inner product between complex matrices $\mathbf{X}$ and $\mathbf{Y}$.

The rest of the paper is organized as follows. Section II introduces the system model and outlines the two complementary robust formulations. Section III deals with the robust game-theoretical formulation, and addresses the solution analysis and design of distributed algorithms along with their convergence properties; some practical implementation issues are also discussed. The robust sum-rate maximization is studied in IV, where a new class of primal/dual decomposition algorithms with provable convergence is introduced. A numerical comparison of the two approaches is given in Section V, and conclusions are drawn in Section VI.

II. System Model and Problem Formulation

In this section we introduce the system model (Section II-A) and outline the two robust design problems: a game-theoretical formulation (Section II-B) and a NUM-based optimization (Section II-C).

A. System model

Consider a hierarchical MIMO CR system composed of P PUs sharing the licensed spectrum with I SUs, modeled as a MIMO Gaussian IC. Each SU i is equipped with n_{T_i} and n_{R_i} transmit and receive antennas, respectively, and PUs may have multiple antennas. Let $\mathbf{H}_{ji}$ (resp. $\mathbf{G}_{pi}$) be the cross-channel function from SU i to SU j (resp. PU p). Under basic information theoretical assumptions, the transmission rate of SU i is

$$r_i(\mathbf{Q}_i,\mathbf{Q}_{-i}) \triangleq \log\det\left(\mathbf{I}+\mathbf{H}_{ii}^H\mathbf{R}_i(\mathbf{Q}_{-i})^{-1}\mathbf{H}_{ii}\mathbf{Q}_i\right) \quad (1)$$

where $\mathbf{Q}_i$ is the transmit covariance matrix of SU i, $\mathbf{Q}_{-i}\triangleq(\mathbf{Q}_j)_{j\neq i}$, $\mathbf{R}_i(\mathbf{Q}_{-i})\triangleq\mathbf{R}_{n_i}+\sum_{j\neq i}\mathbf{H}_{ij}\mathbf{Q}_j\mathbf{H}_{ij}^H$ with $\mathbf{R}_{n_i}\succ\mathbf{0}$ being the covariance matrix of the noise plus the interference generated by the active PUs. The feasible set of SU i is

$$\mathcal{Q}_i \triangleq \left\{\mathbf{Q}_i\succeq\mathbf{0}:\ \begin{array}{l}\mathrm{tr}(\mathbf{Q}_i)\le P_i^{\mathrm{tot}},\ \lambda_{\max}(\mathbf{Q}_i)\le P_i^{\mathrm{peak}}\\ [\mathbf{Q}_i]_{qq}\le P_{iq}^{\mathrm{ant}},\ q=1,\ldots,n_{T_i}\end{array}\right\}, \quad (2)$$

where P_i^{tot} is the total transmit power in units of energy per transmission; in (2), we have also included peak average and per-antenna power constraints, P_i^{peak} is the spatial peak average power threshold, and P_{iq}^{ant} is the average power constraint for the q-th antenna of SU i.

Interference constraints. According to the CR paradigm, SUs are allowed to transmit provided that they do not induce "too much" interference against the primary receivers. In this paper we consider global interference constraints, in the general form described next. Denoting by $\mathbf{B}_{pi}$ the effective channel from the secondary transmitter i to the primary receiver p (this includes the actual cross-channel $\mathbf{G}_{pi}$ and the filtering/beamforming performed by the primary receiver), the interference covariance matrix at PU is $\sum_{i=1}^I\mathbf{B}_{pi}\mathbf{Q}_i\mathbf{B}_{pi}^H$, leading to the following metric of the overall interference experienced by PU p: $\sum_{i=1}^I\mathrm{tr}\left(\mathbf{B}_{pi}\mathbf{Q}_i\mathbf{B}_{pi}^H\right)$. Specific examples for $\mathbf{B}_{pi}$ include:

- $\mathbf{B}_{pi}=\mathbf{G}_{pi}$: the interference metric $\sum_{i=1}^I\mathrm{tr}\left(\mathbf{G}_{pi}\mathbf{Q}_i\mathbf{G}_{pi}^H\right)$ reduces to the aggregate interference generated by all SUs at the primary receiver p [5, 7];
- $\mathbf{B}_{pi}=\mathbf{a}_p^H(\theta)\mathbf{G}_{pi}$, with $\mathbf{a}_p(\theta)$ being the spatial steering vector of angle θ [25]: the interference metric becomes $\sum_{i=1}^I\mathbf{a}_p^H(\theta)\mathbf{G}_{pi}\mathbf{Q}_i\mathbf{G}_{pi}^H\mathbf{a}_p(\theta)$, which measures the interference strength along a specific direction θ.
- $\mathbf{B}_{pi}=\mathbf{w}_p^H\mathbf{G}_{pi}$, with $\mathbf{w}_p$ being the receive beamformer of PU p: the interference metric reduces to $\sum_{i=1}^I\mathbf{w}_p^H\mathbf{G}_{pi}\mathbf{Q}_i\mathbf{G}_{pi}^H\mathbf{w}_p$, which is the interference perceived at the primary receiver p.

With a slight abuse of terminology, we will use the words "channel" and "effective channel" interchangeably. An estimate of the effective channel $\mathbf{B}_{pi}$ can be obtained at the secondary transmitters using standard signal processing techniques. A typical situation is when PUs adopt a time-division duplex (TDD) strategy, and the SUs have a priori knowledge of the PUs' pilot symbols [6]. Note that, in a TDD mode, the primary receiver p uses $\mathbf{w}_p$ for both receive and transmit beamformers, so $\mathbf{B}_{pi}=\mathbf{w}_p^H\mathbf{G}_{pi}$ can be estimated by sensing the pilots from the PU receivers as well.

In typical CR scenarios, however, the SU-to-PU (effective) channels $\mathbf{B}_{pi}$ are difficult to estimate accurately. To model inaccurate or limited SU-to-PU CSI, we express each channel matrix $\mathbf{B}_{pi}$ as [12, 17–20]:

$$\mathbf{B}_{pi}\triangleq\widehat{\mathbf{B}}_{pi}+\Delta\mathbf{B}_{pi}, \quad (3)$$

where $\widehat{\mathbf{B}}_{pi}$ is the estimated channel at the secondary transmitters, and $\Delta\mathbf{B}_{pi}$ quantifies the estimation error which takes values from the so-called uncertainty region $\mathcal{U}_{pi}$:

$$\mathcal{U}_{pi}\triangleq\left\{\Delta\mathbf{B}_{pi}\ :\ \mathrm{tr}\left(\Delta\mathbf{B}_{pi}\mathbf{T}_{pi}\Delta\mathbf{B}_{pi}^H\right)\le\varepsilon_{pi}^2\right\} \quad (4)$$

where $\varepsilon_{pi}>0$ reflects the amount of uncertainty associated with $\widehat{\mathbf{B}}_{pi}$, and $\mathbf{T}_{pi}$ is a given positive definite (weight) matrix. Based on (4), a (worst-case) robust global interference constraint can be written as [26, 27]: for each PU $p=1,\ldots,P$,

$$\sum\nolimits_{i=1}^I\phi_{pi}(\mathbf{Q}_i)\le I_p^{\max}, \quad (5a)$$

where

$$\phi_{pi}(\mathbf{Q}_i)\triangleq \max_{\Delta\mathbf{B}_{pi}\in\mathcal{U}_{pi}}\left\{\mathrm{tr}\left((\widehat{\mathbf{B}}_{pi}+\Delta\mathbf{B}_{pi})\mathbf{Q}_i(\widehat{\mathbf{B}}_{pi}+\Delta\mathbf{B}_{pi})^H\right)\right\} \quad (5b)$$

represents the worst-case interference generated by SU i to PU p, and $I_p^{\max}$ is the maximum level of interference tolerable by PU p. Note that $\phi_{pi}(\mathbf{Q}_i)$ is a convex but non-differentiable function of the covariance matrix $\mathbf{Q}_i$. Constraint (5) is essentially semi-infinite because it has to be satisfied for every $\Delta\mathbf{B}_{pi}$ in the compact set $\mathcal{U}_{pi}$.

Given the local constraint set $\mathcal{Q}_i$ in (2) and the robust global interference constraints (5), we define the joint feasible set $\mathcal{Q}$ of all SUs as

$$\mathcal{Q}\triangleq\left\{(\mathbf{Q}_i)_{i=1}^I:\ \begin{array}{l}\mathbf{Q}_i\in\mathcal{Q}_i,\ i=1,\ldots,I\\ \sum_{i=1}^I\phi_{pi}(\mathbf{Q}_i)\le I_p^{\max},\ p=1,\ldots,P\end{array}\right\}, \quad (6)$$

and we assume without loss of generality that $\mathbf{I}^{\max} = (I_p^{\max})_{p=1}^P > \mathbf{0}$ (this implies that $\mathcal{Q}$ has a non-empty interior).

We remark that other interference constraints can also be included and treated using the same methodology we are going to introduce; examples are null constraints $\mathbf{U}_i^H \mathbf{Q}_i = \mathbf{0}$, and worst-case peak average interference constraint $\sum_{i=1}^I \max_{\Delta\mathbf{B}_{pi}\in\mathcal{U}_{pi}} \lambda_{\max}((\widehat{\mathbf{B}}_{pi}+\Delta\mathbf{B}_{pi})\mathbf{Q}_i(\widehat{\mathbf{B}}_{pi}+\Delta\mathbf{B}_{pi})^H)$.

B. Game-theoretical formulation

The first system design we propose is based on a robust game-theoretical formulation: the SUs are modeled as players of a NEP aiming at maximizing their own transmission rate (1) subject to the power and robust global interference constraints (2) and (5). To keep the system design as decentralized as possible, the robust global interference constraints are enforced by introducing a pricing term in the objective function of each SU: there is one price μ_p associated with each of the global interference constraints; let us denote by $\boldsymbol{\mu} \triangleq (\mu_p)_{p=1}^P$ the vector of all prices. Stated in mathematical terms, we have the following pricing-based NEP: anticipating the rival strategies $\mathbf{Q}_{-i}$ and the price $\boldsymbol{\mu}$, each SU $i = 1, \ldots, I$ solves

$$\begin{array}{ll} \underset{\mathbf{Q}_i}{\text{maximize}} & r_i(\mathbf{Q}_i, \mathbf{Q}_{-i}) - \sum_{p=1}^P \mu_p \cdot \phi_{pi}(\mathbf{Q}_i) \\ \text{subject to} & \mathbf{Q}_i \in \mathcal{Q}_i. \end{array} \tag{7a}$$

The game is completed with the robust global constraints to be satisfied by the price vector $\boldsymbol{\mu}$:

$$0 \le \mu_p \perp I_p^{\max} - \sum\nolimits_{i=1}^I \phi_{pi}(\mathbf{Q}_i) \ge 0, \quad p = 1, \ldots, P, \tag{7b}$$

where $a \perp b$ means $a \cdot b = 0$. Condition (7b) says that the robust global interference constraints must be satisfied together with nonnegative pricing; in addition, they imply that if a constraint is satisfied with strict inequality, the corresponding price must be zero (the SUs are not punished if they already satisfy the interference requirements).

Related Works. The proposed NEP differs from previous game-theoretical formulations [11, 12, 28, 29] in several aspects, namely: i) our design incorporates *robust global* interference temperature constraints (in [28, 29] there are no global constraints, and in [11] the global constraints are nonrobust, whereas in [12] there are only local robust interference constraints); ii) SU's objective functions are nondifferentiable; iii) the pricing term is nonlinear in the optimization variables (due to the worst-case nature); and iv) players' optimization variables are complex matrices (in [11, 28, 29] the optimization variables are real vectors, whereas in [12] the problem can be conveniently written using real matrices). Because of these issues, the NEP (7) is much more involved and cannot be studied by a direct application of existing results in [11, 12]. Section III addresses these technical issues and provides a comprehensive treatment of the NEP in terms of solution analysis and algorithms.

C. Social-oriented design

A complementary formulation to the game-theoretical design is the classical NUM problem, which is

$$\begin{array}{ll} \underset{\mathbf{Q}_1,\ldots,\mathbf{Q}_I}{\text{maximize}} & U(\mathbf{Q}) \triangleq \sum_{i=1}^I r_i(\mathbf{Q}_i, \mathbf{Q}_{-i}) \\ \text{subject to} & \mathbf{Q}_i \in \mathcal{Q}_i, \quad i = 1, \ldots, I, \\ & \sum_{i=1}^I \phi_{pi}(\mathbf{Q}_i) \le I_p^{\max}, \quad p = 1, \ldots, P. \end{array} \tag{8}$$

Related Works. Instances of (8) have been widely studied in the literature; relevant works dealing with local and/or global interference constraints but under *perfect* SU-to-PU CSI are [5, 8, 30]. In [7], the authors focused on the minimization of the (nonconvex) sum-MSE subject to robust global interference constraints, and proposed centralized and distributed schemes aiming at reaching stationary solutions of the problem. Convergence of the algorithms in [7] however remains an open and challenging problem. In this paper, for the first time, we propose a class of distributed *convergent* algorithms to stationary solutions of (8). Interestingly, our algorithms and convergence analysis apply also to the problems studied in [7]. This framework is developed in Section IV.

III. Game-Theoretical Design

The first step of our analysis is to get rid of the nondifferentiability of the SU's utility function by rewriting (7) in an equivalent but more convenient form. More specifically, introducing the real slack variables $\mathbf{t}_i \triangleq (t_{pi})_{p=1}^P \ge \mathbf{0}$ and the augmented (convex) feasible set

$$\widetilde{\mathcal{Q}}_i \triangleq \{(\mathbf{Q}_i, \mathbf{t}_i) : \mathbf{Q}_i \in \mathcal{Q}_i,\ \phi_{pi}(\mathbf{Q}_i) - t_{pi} \le 0,\ \forall\, p\}, \tag{9}$$

the NEP (7a) can be rewritten as: for SU $i = 1, \ldots, I$,

$$\mathcal{G}_{\boldsymbol{\mu}}: \quad \begin{array}{ll} \underset{\mathbf{Q}_i,\mathbf{t}_i}{\text{maximize}} & r_i(\mathbf{Q}_i, \mathbf{Q}_{-i}) - \langle \boldsymbol{\mu}, \mathbf{t}_i \rangle \\ \text{subject to} & (\mathbf{Q}_i, \mathbf{t}_i) \in \widetilde{\mathcal{Q}}_i, \end{array} \tag{10a}$$

along with the complementary condition

$$0 \le \mu_p \perp I_p^{\max} - \sum\nolimits_{i=1}^I t_{pi} \ge 0, \quad p = 1, \ldots, P. \tag{10b}$$

For future convenience, we will refer to the NEP in (10a) with fixed (exogenous) price vector $\boldsymbol{\mu} \ge \mathbf{0}$ as $\mathcal{G}_{\boldsymbol{\mu}}$. With a slight abuse of notation, we will denote the overall game $\mathcal{G}_{\boldsymbol{\mu}}$ in (10a) plus the complementary condition (10b) by $\mathcal{G}$. We focus next on the solution analysis of $\mathcal{G}$ as well as the design of distributed algorithms.

A. Solution analysis: connection to complex VIs

The definition of NE for a game with price equilibrium conditions as $\mathcal{G}$ is the natural generalization of the same concept introduced for standard NEPs (with no global constraints) and is given next.

Definition 1 (NE of $\mathcal{G}$)**.** A Nash Equilibrium of the game $\mathcal{G}$ in (10) is a strategy-price tuple $(\mathbf{Q}^\star, \mathbf{t}^\star, \boldsymbol{\mu}^\star)$ such that

$$(\mathbf{Q}_i^\star, \mathbf{t}_i^\star) \in \underset{(\mathbf{Q}_i,\mathbf{t}_i)\in\widetilde{\mathcal{Q}}_i}{\arg\max} \left\{ r_i(\mathbf{Q}_i, \mathbf{Q}_{-i}^\star) - \langle \boldsymbol{\mu}^\star, \mathbf{t}_i \rangle \right\},\ i = 1, \ldots, I,$$

and

$$0 \le \mu_p^\star \perp I_p^{\max} - \sum\nolimits_{i=1}^I t_{pi}^\star \ge 0, \quad p = 1, \ldots, P.$$

Note that even though each SU's optimization problem is convex, the NEP may not have a NE; moreover standard game-theoretical existence results cannot be applied because of the unboundedness of the price vector (the prices cannot be normalized due to the lack of homogeneity in the players' optimization problem). To cope with these issues, we hinge on the advanced framework of finite-dimensional VI [31]. Standard VIs are defined in the real domain [31], but the players' optimization variables in the game $\mathcal{G}$ are complex matrices. Of course, one approach would be to rewrite the SUs' optimization problems in terms of real and imaginary parts of the original complex variables. This is however awkward because it "destroys" the structure of the optimization problem, and leads to final conditions that cannot be easily written in terms of the original complex setup. It seems instead more convenient to work directly in the complex domain. This naturally calls for the definition of the VI problem in the domain of complex matrices; *complex* VIs and their connection with standard *complex* NEP have been introduced and studied in [21]. Capitalizing on [21], in the present paper we extend (some of) those results to complex NEPs with *pricing and global constraints*. We introduce next the definition of complex VI; then we establish the connection with the game $\mathcal{G}$ and study its main properties.

Definition 2 (Complex VI [21])**.** Given a convex and closed set $\mathcal{X} \in \mathbb{C}^{n\times m}$ and a complex-valued matrix function $\mathbf{F}^{\mathbb{C}}(\mathbf{X}) : \mathcal{X} \ni \mathbf{X} \to \mathbb{C}^{n\times m}$, the complex VI problem, denoted by VI$(\mathcal{X}, \mathbf{F}^{\mathbb{C}})$, consists in finding a point $\mathbf{X}^\star \in \mathcal{X}$ such that

$$\left\langle \mathbf{X} - \mathbf{X}^\star, \mathbf{F}^{\mathbb{C}}(\mathbf{X}^\star)\right\rangle \geq 0, \quad \forall\, \mathbf{X} \in \mathcal{X}, \tag{11}$$

where $\langle \mathbf{A}, \mathbf{B}\rangle \triangleq \Re\left(\mathrm{tr}\left(\mathbf{A}^H\mathbf{B}\right)\right)$.

If $\mathbf{X}$ and $\mathbf{F}^{\mathbb{C}}(\mathbf{X})$ are partitioned according to $\mathbf{X} = (\mathbf{X}_i)_{i=1}^I$ and $\mathbf{F}^{\mathbb{C}}(\mathbf{X}) = (\mathbf{F}_i^{\mathbb{C}}(\mathbf{X}))_{i=1}^I$ [such that $\left\langle \mathbf{X}_i, \mathbf{F}_i^{\mathbb{C}}(\mathbf{X})\right\rangle$ is defined for all $i = 1, \ldots, I$], the inner product in (11) is intended as

$$\sum\nolimits_{i=1}^I \left\langle \mathbf{X}_i - \mathbf{X}_i^\star, \mathbf{F}_i^{\mathbb{C}}(\mathbf{X}^\star)\right\rangle \geq 0.$$

To establish the connection between a suitably defined complex VI and the game $\mathcal{G}$, let us recall the joint feasible set $\mathcal{Q}$ defined in (6) and introduce the complex-valued matrix mapping $\mathbf{F}^{\mathbb{C}}(\mathbf{Q}) = \left(\mathbf{F}_i^{\mathbb{C}}(\mathbf{Q})\right)_{i=1}^I$ with

$$\begin{aligned}\mathbf{F}_i^{\mathbb{C}}(\mathbf{Q}) &= -\nabla_{\mathbf{Q}_i^*} r_i(\mathbf{Q}_i, \mathbf{Q}_{-i}) \\ &= -\mathbf{H}_{ii}^H\Big(\mathbf{R}_{n_i} + \sum\nolimits_{j=1}^I \mathbf{H}_{ij}\mathbf{Q}_j\mathbf{H}_{ij}^H\Big)^{-1}\mathbf{H}_{ii},\end{aligned} \tag{12}$$

where $\nabla_{\mathbf{Q}_i^*} r_i(\mathbf{Q}_i, \mathbf{Q}_{-i})$ denotes the gradient of $r_i(\mathbf{Q}_i, \mathbf{Q}_{-i})$ with respect to (w.r.t.) $\mathbf{Q}_i^*$ (the conjugate of $\mathbf{Q}_i$); see, e.g., [21, 32]. Then we have the following.

Proposition 3. *The complex game $\mathcal{G}$ in (10) is equivalent to the complex VI$(\mathcal{Q}, \mathbf{F}^{\mathbb{C}})$. The equivalence is in the following sense. If $\mathbf{Q}^{VI} = (\mathbf{Q}_i^{VI})_{i=1}^I$ is a solution of the VI$(\mathcal{Q}, \mathbf{F}^{\mathbb{C}})$, then there exists a $\boldsymbol{\mu}^{VI} \triangleq (\mu_p^{VI})_{p=1}^P$−the multiplier of the VI associated with the robust global interference constraint (5)−such that $(\mathbf{Q}^{VI}, \mathbf{t}^{VI}, \boldsymbol{\mu}^{VI})$ is a NE of $\mathcal{G}$, with $\mathbf{t}^{VI} \triangleq (t_{pi}^{VI})_{p,i}$ and $t_{pi}^{VI} = \phi_{pi}(\mathbf{Q}_i^{VI})$. Conversely, if $(\mathbf{Q}^{NE}, \mathbf{t}^{NE}, \boldsymbol{\mu}^{NE})$ is a NE of $\mathcal{G}$, then $t_{pi}^{NE} = \phi_{pi}(\mathbf{Q}_i^{NE})$ and $\mathbf{Q}^{NE}$ is a solution of the VI$(\mathcal{Q}, \mathbf{F}^{\mathbb{C}})$ with $\boldsymbol{\mu}^{NE}$ being the multiplier associated with the robust global interference constraint (5).*

The proof of the proposition comes from the equivalence of the Karush-Kuhn-Tucker (KKT) conditions of $\mathcal{G}$ and those of the VI$(\mathcal{Q}, \mathbf{F}^{\mathbb{C}})$. Note that $\mathcal{Q}$ in (6) has a non-empty interior, so Slater's constraint qualification is satisfied, see [31, Section 3.2]. We omit the details due to space limit.

The equivalence of the game $\mathcal{G}$ with the VI$(\mathcal{Q}, \mathbf{F}^{\mathbb{C}})$ paves the way to the study of existence and uniqueness of a NE, capitalizing on the solution analysis of complex VIs [21]. We introduce the $I \times I$ matrix $\boldsymbol{\Upsilon}$:

$$[\boldsymbol{\Upsilon}]_{ij} = \begin{cases} 1, & \text{if } i = j, \\ -\rho\left(\mathbf{H}_{ii}^{\dagger H}\mathbf{H}_{ij}^H\mathbf{H}_{ij}\mathbf{H}_{ii}^\dagger\right)\cdot \mathtt{INNR}_{ij}, & \text{if } i \neq j, \end{cases}$$

where $\mathbf{H}_{ii}$ is assumed to be full column rank, $\mathbf{A}^\dagger$ is the Moore-Penrose pseudoinverse of $\mathbf{A}$, $\rho(\mathbf{A})$ is the spectral radius of $\mathbf{A}$, and $\mathtt{INNR}_{ij}$ is defined as

$$\mathtt{INNR}_{ij} \triangleq \frac{\rho\left(\mathbf{R}_{n_i} + \sum_{i=1}^I P_i^{\mathrm{tot}}\mathbf{H}_{ij}\mathbf{H}_{ij}^H\right)}{\lambda_{\min}\left(\mathbf{R}_{n_i}\right)}, \tag{13}$$

the main existence/uniqueness result for the NE of $\mathcal{G}$ is the following.

Proposition 4. *The VI$(\mathcal{Q}, \mathbf{F}^{\mathbb{C}})$ always has a solution; therefore the game $\mathcal{G}$ has a NE. Moreover, if $\boldsymbol{\Upsilon} \succ \mathbf{0}$, the VI$(\mathcal{Q}, \mathbf{F}^{\mathbb{C}})$ has a unique solution; therefore the $(\mathbf{Q}, \mathbf{t})$-tuple of the NE of $\mathcal{G}$ is unique.*

Remark 5 (On the uniqueness conditions). A sufficient condition for $\boldsymbol{\Upsilon}$ to be positive definite is: for all $i = 1, \ldots, I$,

$$\begin{aligned}&\tfrac{1}{2}\textstyle\sum_{j\neq i}\left\{\rho\left(\mathbf{H}_{ii}^{\dagger H}\mathbf{H}_{ij}^H\mathbf{H}_{ij}\mathbf{H}_{ii}^\dagger\right)\cdot \mathtt{INNR}_{ij}\right\} \\ &+\tfrac{1}{2}\textstyle\sum_{j\neq i}\left\{\rho\left(\mathbf{H}_{jj}^{\dagger H}\mathbf{H}_{ji}^H\mathbf{H}_{ji}\mathbf{H}_{jj}^\dagger\right)\cdot \mathtt{INNR}_{ji}\right\} < 1.\end{aligned} \tag{14}$$

The condition is the standard diagonal dominance of (the symmetric part of) $\boldsymbol{\Upsilon}$. Uniqueness conditions above have an interesting physical interpretation: the NE is unique if the interference among the SUs is "sufficiently small". For instance, the first sum in (14) can be interpreted as a constraint on the maximum amount of interference that each secondary receiver can tolerate [$\rho(\mathbf{H}_{ii}^{\dagger H}\mathbf{H}_{ij}^H\mathbf{H}_{ij}\mathbf{H}_{ii}^\dagger)\cdot \mathtt{INNR}_{ij}$ is the maximum amount of interference that SU i can experience due to the transmissions of SU j], whereas the second sum in (14) imposes an upper bound on the maximum interference that each secondary transmitter can generate. Note that the above conditions are independent of the price vector $\boldsymbol{\mu}$.

B. Distributed algorithms

The challenging goal of this subsection is to devise a decentralized mechanism solving the game $\mathcal{G}$. While useful for the solution analysis, the VI reformulation of $\mathcal{G}$ (cf. Proposition 3) does not help much in obtaining distributed algorithms. This is due to the fact that the feasible set $\mathcal{Q}$ of the VI [cf. (6)] does not have a Cartesian structure [cf. (5)], leaving coupled the strategies of the players; therefore, standard solution methods for the VI$(\mathcal{Q}, \mathbf{F}^{\mathbb{C}})$ as in [31] would lead to centralized schemes, which are not applicable in our CR scenario. To deal with this issue, we start rewriting the game $\mathcal{G}$ as a *Nonlinear Complementarity Problem* (NCP).

A Nonlinear Complementarity Reformulation. Let us focus preliminarily on the NEP $\mathcal{G}_{\boldsymbol{\mu}}$ with exogenous (fixed) price $\boldsymbol{\mu} \geq \mathbf{0}$ [cf. (10a)] and assume that $\mathcal{G}_{\boldsymbol{\mu}}$ has a unique NE for any given $\boldsymbol{\mu} \geq \mathbf{0}$, denoted by $(\mathbf{Q}^\star(\boldsymbol{\mu}), \mathbf{t}^\star(\boldsymbol{\mu}))$ (we derive conditions for the uniqueness in Proposition 6), where we have made explicit the dependence on $\boldsymbol{\mu}$. Under this condition, let us introduce the mapping:

$$\mathbf{M}(\boldsymbol{\mu}) : \mathbb{R}_+^P \ni \boldsymbol{\mu} \mapsto \mathbf{M}(\boldsymbol{\mu}) \triangleq \left(I_p^{\max} - \sum\nolimits_{i=1}^{I} t_{pi}^\star(\boldsymbol{\mu})\right)_{p=1}^{P}, \tag{15}$$

which measures the violation of the robust global interference constraints at the NE $(\mathbf{Q}^\star(\boldsymbol{\mu}), \mathbf{t}^\star(\boldsymbol{\mu}))$. It turns out that if $\boldsymbol{\mu}$ is such that the complementarity condition (10b) is satisfied, that is $\mathbf{0} \leq \boldsymbol{\mu} \perp \mathbf{M}(\boldsymbol{\mu}) \geq \mathbf{0}$, then $(\mathbf{Q}^\star(\boldsymbol{\mu}), \mathbf{t}^\star(\boldsymbol{\mu}), \boldsymbol{\mu})$ is a NE of $\mathcal{G}$. This is made formal in the following proposition.

Proposition 6. *Given the game $\mathcal{G}$, suppose that $\boldsymbol{\Upsilon}$ is a P-matrix. Then the following hold.*

(a): *$\mathcal{G}_{\boldsymbol{\mu}}$ has a unique NE, denoted by $(\mathbf{Q}^\star(\boldsymbol{\mu}), \mathbf{t}^\star(\boldsymbol{\mu}))$, for any given $\boldsymbol{\mu} \geq \mathbf{0}$;*

(b): *$\mathcal{G}$ is equivalent to the nonlinear complementary problem*

$$\text{NCP}(\mathbf{M}): \quad \mathbf{0} \leq \boldsymbol{\mu} \perp \mathbf{M}(\boldsymbol{\mu}) \geq \mathbf{0}. \tag{16}$$

The equivalence is in the following sense: the NCP($\mathbf{M}$) *must have a solution and, for any such solution $\boldsymbol{\mu}^{NCP}$, the tuple $(\mathbf{Q}^\star(\boldsymbol{\mu}^{NCP}), \mathbf{t}^\star(\boldsymbol{\mu}^{NCP}), \boldsymbol{\mu}^{NCP})$ is a NE of $\mathcal{G}$; conversely, if $(\mathbf{Q}^{NE}, \mathbf{t}^{NE}, \boldsymbol{\mu}^{NE})$ is a NE of $\mathcal{G}$, then $\boldsymbol{\mu}^{NE}$ is a solution of the* NCP($\mathbf{M}$) *with $(\mathbf{Q}^\star(\boldsymbol{\mu}^{NE}), \mathbf{t}^\star(\boldsymbol{\mu}^{NE})) = (\mathbf{Q}^{NE}, \mathbf{t}^{NE})$;*

(c): *If the condition on $\boldsymbol{\Upsilon}$ is strengthened to $\boldsymbol{\Upsilon} \succ \mathbf{0}$, then for any solution $\boldsymbol{\mu}^{NCP}$ of* NCP($\mathbf{M}$) *we have $(\mathbf{Q}^\star(\boldsymbol{\mu}^{NCP}), \mathbf{t}^\star(\boldsymbol{\mu}^{NCP})) = (\mathbf{Q}^{NE}, \mathbf{t}^{NE})$, where $(\mathbf{Q}^{NE}, \mathbf{t}^{NE})$ is the unique $(\mathbf{Q}, \mathbf{t})$-component of the NE of $\mathcal{G}$.*

The difference between Proposition 6 (b) and (c) is that, under the more stringent condition in (c), all the solutions $\boldsymbol{\mu}^{\text{NCP}}$ of the NCP($\mathbf{M}$) yield equilibrium pairs $(\mathbf{Q}^\star(\boldsymbol{\mu}^{\text{NCP}}), \mathbf{t}^\star(\boldsymbol{\mu}^{\text{NCP}}))$ of the game $\mathcal{G}$ having the *same* $(\mathbf{Q}, \mathbf{t})$-component. In Proposition 6 we invoked the P property of the matrix $\boldsymbol{\Upsilon}$; we refer to [33] for a detailed treatment of P matrices. Here we recall only that a matrix $\boldsymbol{\Upsilon}$ is P if every principal minor of $\boldsymbol{\Upsilon}$ is positive. Note that every positive definite matrix is a P matrix, but the reverse does not hold (unless the matrix is symmetric). A sufficient condition for $\boldsymbol{\Upsilon}$ to be P is that either the first or the second sum in (14) is less than $1/2$; see Remark 5 for a physical interpretation of these conditions.

In the setting of Proposition 6, one can compute the NE of $\mathcal{G}$ by solving the NCP($\mathbf{M}$), which offers the possibility of devising distributed algorithms (the feasible set of the NCP is the nonnegative orthant and thus has a Cartesian structure), whose convergence can be studied using well-known results from the theory of VIs. An example is the *Projection Algorithm with variable step-size* [31, Chapter 12] applied to the NCP($\mathbf{M}$) and formally described in Algorithm 1, where in (17) the symbol $[\bullet]^+$ denotes the (componentwise) Euclidean projection onto the nonnegative orthant.

Convergence conditions of Algorithm 1 are given in the next theorem, whose proof follows from [31, Th. 12.1.8] and the co-coercivity properties of the mapping $\mathbf{M}(\boldsymbol{\mu})$ on $\mathbb{R}_+^P$ (proved in Appendix A); the co-coercivity constant $c_{\text{coc}} > 0$ in Theorem 7 is given by

Algorithm 1: Projection Algorithm with Variable Step-size for the NCP($\mathbf{M}$) in (16)

Data: $\boldsymbol{\mu}^0 \geq \mathbf{0}$ and $\{\gamma^\nu\}_{\nu=0}^{\infty} > 0$. Set $\nu = 0$.
(S.1): If $\boldsymbol{\mu}^\nu$ satisfies a suitable termination criterion: STOP.
(S.2): Compute the unique NE of $\mathcal{G}_{\boldsymbol{\mu}^\nu}$ in (10a).
(S.3): Update the price vector according to

$$\boldsymbol{\mu}^{\nu+1} = \left[\boldsymbol{\mu}^\nu - \gamma^\nu \mathbf{M}(\boldsymbol{\mu}^\nu)\right]^+. \tag{17}$$

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

$$c_{\text{coc}} = \frac{c_{\text{sm}}}{\sum_{k=1}^{K}\left(\sum_{i=1}^{I} L_{\phi,pi}\right)^2}, \tag{18}$$

where $c_{\text{sm}} = \frac{1}{2}\lambda_{\min}(\boldsymbol{\Upsilon} + \boldsymbol{\Upsilon}^T)$ and $L_{\phi,pi}$ is the Lipschitz constant of $\phi_{pi}(\mathbf{Q}_i)$, whose existence is well justified [34, Th. 10.4] and given by

$$L_{\phi,pi} = \frac{(P_i^{\text{tot}} + d(\mathcal{Q}_i))(\varepsilon_{pi}/\lambda_{\min}(\mathbf{T}_{pi}) + d(\mathcal{Q}_i))\sigma_{\max}^2(\widehat{\mathbf{B}}_{pi})}{d(\mathcal{Q}_i)^2}, \tag{19}$$

where $d(\mathcal{Q}_i)$ is the diameter of the smallest ball containing the compact set $\mathcal{Q}_i$.

Theorem 7. *Suppose that $\boldsymbol{\Upsilon} \succ \mathbf{0}$ and $\{\gamma^\nu\}_{\nu=0}^{\infty}$ is chosen so that $0 < \inf_\nu \gamma^\nu \leq \gamma^\nu \leq \sup_\nu \gamma^\nu < 2c_{\text{coc}}$. Then the sequence $\{\boldsymbol{\mu}^\nu\}$ generated by Algorithm 1 converges to a solution of the* NCP($\mathbf{M}$) *in (16).*

Remark 8 (On Algorithm 1). Algorithm 1 provides an update rule for the price vector, guaranteeing convergence to a solution of the NCP($\mathbf{M}$) and thus to a NE of the game $\mathcal{G}$ (cf. Proposition 6). It is essentially a double-loop scheme. In the inner loop (Step 2), given the price vector $\boldsymbol{\mu}^\nu \geq \mathbf{0}$, the SUs compute the unique NE of $\mathcal{G}_{\boldsymbol{\mu}^\nu}$; once a NE is reached, the price vector $\boldsymbol{\mu}^\nu$ is updated according to (17) in the outer loop (Step 3); then the game $\mathcal{G}_{\boldsymbol{\mu}}$ is played again, but now with $\boldsymbol{\mu} = \boldsymbol{\mu}^{\nu+1}$. The outer loop is terminated (Step 1) when the price change in two consecutive iterations is smaller than a (given) positive constant ϵ, e.g., $\left\|\boldsymbol{\mu}^{\nu+1} - \boldsymbol{\mu}^\nu\right\| \leq \epsilon$. The implementation of Algorithm 1 requires the computation of the solution of the NE of $\mathcal{G}_{\boldsymbol{\mu}^\nu}$ in (10a). This can be done using best-response schemes wherein the SUs, according to a given scheduling (e.g., simultaneously or sequentially), solve their optimization problems (10a), given $\boldsymbol{\mu} = \boldsymbol{\mu}^\nu$. The interesting result is that conditions in Theorem 7 are sufficient also for the convergence of this inner loop, under *simultaneous, sequential, or asynchronous* updates from the SUs [21, 35, 36]. An instance of Step 2 when a Jacobi best-response iterate is used to compute a NE of $\mathcal{G}_{\boldsymbol{\mu}^\nu}$ is given in Algorithm 2.

C. Discussion on the implementation

We discuss now some implementation issues related to the proposed algorithms. As already observed in Remark 8, Algorithm 1 is composed of two major steps: the computation of the NE of the inner game $\mathcal{G}_{\boldsymbol{\mu}}$ (Step 2) and the update of the price vector (Step 3). Then, some natural questions are: i) Who will perform the update of the price vector in Step

Algorithm 2: Algorithm 1 with inner Jacobi best-response iterates.

The steps of the algorithm are the same as those of Algorithm 1, except for Step 2, which is modified as follows. Given $(\mathbf{Q}_i^0, \mathbf{t}_i^0) \in \widetilde{\mathcal{Q}}_i$, for all $i = 1, \ldots, I$; $\boldsymbol{\mu}^\nu$; $\delta > 0$; set $n = 0$,

(S.2a): For each $i = 1, \ldots, I$, compute $(\mathbf{Q}_i^{n+1}, \mathbf{t}_i^{n+1})$ as

$$\left(\mathbf{Q}_i^{n+1}, \mathbf{t}_i^{n+1}\right) \in \underset{(\mathbf{Q}_i, \mathbf{t}_i) \in \widetilde{\mathcal{Q}}_i}{\arg\max}\ r_i(\mathbf{Q}_i, \mathbf{Q}_{-i}^n) - \left\langle \boldsymbol{\mu}^\nu, \mathbf{t}_i \right\rangle \quad (20)$$

(S.2b): If $\left\| \sum_{i=1}^I \mathbf{t}_i^{n+1} - \sum_{i=1}^I \mathbf{t}_i^n \right\| \leq \delta$, then STOP; otherwise set $n \leftarrow n+1$ and go to (S.2a).

3; ii) How to implement this update in a (fairly) distributed way? iii) What is the convergence time and PU-to-SU/SU-to-SU communication overhead of the two steps? iv) How can the SUs compute efficiently their best-response in Step 2? We discuss next different protocols to address these issues, each of them characterized by a different level PU-to-SU and SU-to-SU signaling and complexity.

1) PU-to-SU/SU-to-SU signaling and communication overhead in Step 3: Depending on the debate paradigm assumed for the CR network, *spectrum leasing* versus *common* model, the price update in Step 3 can be performed either by the PUs or by the SUs, as detailed next.

CR spectrum leasing model. In this setting, an interaction between the PUs and SUs is allowed. It is thus natural that the price update is performed by the PUs. This however requires some explicit signaling between PUs and SUs: at each iteration ν, each SU i needs to communicate his (scalar) worst-case interference value t_{pi}^ν to PU p, so that the PUs can evaluate the function $\mathbf{M}(\boldsymbol{\mu}^\nu)$, update the price vector $\boldsymbol{\mu}$ according to (17), and broadcast the new value $\boldsymbol{\mu}^{\nu+1}$. Note that the computation of the projection onto the nonnegative orthant in (17) is a component-wise operation, implying that each PU can update locally his own price performing a computationally inexpensive operation. We remark that the signaling and communication overhead of this scheme is lower than that of state-of-the-art algorithms dealing with (robust) global interference constraints, e.g., [5, 7], see Section IV-C.

CR common model. In this scenario the PUs are oblivious of the presence of the SUs (e.g., the PUs are legacy systems), thus behaving as if no secondary activities were present. Therefore, the update of the price vector in Step 3 needs to be performed by the SUs themselves. This can be done in different ways, at the cost however of some signaling among the SUs and computational complexity. A centralized approach would be to elect a "sink" node in the secondary network that collects the required information from the other SUs (using the network control channel), updates and broadcasts the price vector $\boldsymbol{\mu}$.

An alternative method less demanding in terms of network signaling is to estimate the violation function $\mathbf{M}(\boldsymbol{\mu})$ locally at each secondary node by running consensus algorithms, which requires the interaction only between nearby nodes. Consensus schemes have become popular over the last decade as a practical scheme for the in-network distributed calculation of general functions of the node values (see, e.g., [37, 38]); here, we suggest to use the *finite-time distributed* convergence linear scheme proposed in [39]. The main advantage of this scheme w.r.t. the more classical consensus/gossip algorithms whose convergence is only asymptotic (i.e., exact consensus is not reached in a finite number of times) is that, at no extra signaling, each node can immediately calculate the consensus value after observing the evolution of its own value over a finite number of iterations (specifically, upper bounded by the size of the network). Therefore, starting from the local information $\mathbf{t}_i^\star(\boldsymbol{\mu})$ (which is available at node i) and running the algorithm in [39, Figures 1-2], each node i will converge to the estimate of the sums $\sum_{i=1}^I t_{pi}^\star(\boldsymbol{\mu})$ for all $p = 1, \ldots, P$ [and thus the violation function $\mathbf{M}(\boldsymbol{\mu})$].

The communication cost incurred by this protocol can be characterized as follows. Modeling the secondary network as a directed graph, the consensus algorithm described in [39] converges in at most $\tau^{\max} \leq I - \min_i |\mathcal{N}_i| + 1$ iterations (a tighter bound can be found in [39]), where $|\mathcal{N}_i|$ is the number of neighbors of SU i (the nodes that interfere with node i), also termed in-degree of node i. According to the consensus algorithm, in each step, each SU i transmits a scalar value on each outgoing edge (the outgoing edges from each node i link the nodes associated with the SUs who receive interference from SU i); since there are at most $\tau^{\max} - 1$ runs, each SU i will in principle have to transmit $\tau^{\max} \cdot \mathtt{deg}_i^{\text{out}}$ messages, where $\mathtt{deg}_i^{\text{out}}$ is the out-degree of node i (i.e., the number of SUs having SU i as interferer). Thanks to the broadcast nature of the wireless channel, however, a single transmission of each SU i will be equivalent to communicating a message along each of $\mathtt{deg}_i^{\text{out}}$ outgoing edges, and thus each node would only have to transmit $\tau^{\max}$ messages. Summing over all nodes in the network, there will be $I \cdot \tau^{\max}$ overall messages that have to be transmitted to run the consensus protocol.

Finally we mention that there are some scenarios where the PUs cannot communicate with the SUs but the primary receivers have a fixed geographical location. Then it might be possible to install some monitoring devices close to each primary receiver having the functionality of cross-channel estimation as well as price computation and broadcast. Note that the uncertainty region (4) and the proposed robust optimization model fits well into this situation, in which the cross-channel between the monitoring devices and the SUs can only represent an estimate of the real PU-to-SU channel.

2) On Step 2 and the computation of the best-response: A natural question dealing with Step 2 of Algorithm 1 is how to compute the NE of the inner game $\mathcal{G}_{\boldsymbol{\mu}}$, given the price $\boldsymbol{\mu} \geq \mathbf{0}$. One can use best-response algorithms (e.g., Algorithm 2), as already discussed in Remark 8. Note that once the price $\boldsymbol{\mu}$ is available at the secondary transmitters, the implementation of such algorithms is distributed. Indeed, to compute his best-response, each SU only needs to locally measure the covariance matrix of the interference plus noise. Moreover the *asynchronous* implementation of best-response algorithms relaxes the requirements on the network synchronization: some SUs are allowed to update their strategies more frequently than others and even use outdated interference measurements.

The last issue left to discuss is the computation of the SU's best-response (20), given the price vector $\boldsymbol{\mu}$ and the covariance matrix $\mathbf{R}_{-i}(\mathbf{Q}_{-i})$. Since the maximum-value function $\phi_{pi}(\mathbf{Q}_i)$ in the feasible set makes the computation quite difficult, we reformulate it as a linear matrix inequality (LMI)

based on the so-called *S-procedure* [18]: $\phi_{pi}(\mathbf{Q}_i) \leq t_{pi}$ if and only if there is a η_{pi} such that

$$\eta_{pi} \geq 0 \quad \text{and} \quad \mathbf{S}_{pi}(\mathbf{Q}_i) \succeq \mathbf{0}, \tag{21a}$$

where

$$\mathbf{S}_{pi}(\mathbf{Q}_i) \triangleq \begin{bmatrix} (\eta_{pi}\mathbf{T}_{pi} - \mathbf{Q}_i)^T \otimes \mathbf{I} & -\mathrm{vec}(\widehat{\mathbf{B}}_{pi}\mathbf{Q}_i) \\ -\mathrm{vec}(\widehat{\mathbf{B}}_{pi}\mathbf{Q}_i)^H & t_{pi} - \eta_{pi}\varepsilon_{pi}^2 \\ & -\mathrm{tr}(\widehat{\mathbf{B}}_{pi}\mathbf{Q}_i\widehat{\mathbf{B}}_{pi}^H) \end{bmatrix}. \tag{21b}$$

Based on (21), each optimization problem (20) can be rewritten in the following equivalent form:

$$\begin{array}{ll} \underset{\mathbf{Q}_i, \mathbf{t}_i, \boldsymbol{\eta}_i}{\text{maximize}} & r_i(\mathbf{Q}_i, \mathbf{Q}_{-i}) - \langle \boldsymbol{\mu}, \mathbf{t}_i \rangle \\ \text{subject to} & \mathbf{Q}_i \in \mathcal{Q}_i,\ \boldsymbol{\eta}_i \geq \mathbf{0}, \\ & \mathbf{S}_{pi}(\mathbf{Q}_i) \succeq \mathbf{0}, \quad p = 1, \ldots, P, \end{array} \tag{22}$$

with $\boldsymbol{\eta}_i \triangleq (\eta_{pi})_{i=1}^I$ being additional slack variables. This is a tractable convex optimization problem, whose solution can be obtained from many solvers, e.g., `SeDuMi` [41].

IV. Social-oriented Sum-Rate Maximization

The design of distributed solution methods for the sum-rate maximization problem (8) confronts three major challenges, namely: i) the *nonconvexity* of the sum-rate function; ii) the *lack of separability* of the sum-rate function (the rate of each SU depends also on the covariance matrices of the others); and iii) the presence of *coupling* (*semi-infinite*) robust constraints, and they make the optimization problem arduous to manage. In this section, we propose for this problem a first class of *distributed* algorithms with *provable convergence*.

Specifically, we follow a two-step procedure. First, we cope with the nonconvexity and nonseparability of the sum-rate function capitalizing on successive convex approximation techniques recently proposed in [24]. More specifically, the original nonconvex problem (8) is replaced by a suitably defined *sequence* of convex subproblems, each of them having a *separable strongly convex* function, but still with coupling interference constraints. The unique solution of any of such subproblems can be efficiently computed via centralized algorithms; inexact solutions are also allowed without affecting the overall convergence. Then, building on standard primal and dual decomposition techniques, we propose two efficient methods to solve the convex subproblems in a distributed way.

A. Centralized SCA-based algorithm

We start the discussion from an informal description of the proposed solution method that sheds light on the core idea of the novel decomposition technique and establishes the connection with classical gradient-based ascent schemes.

A classical approach to solve the nonconvex problem (8) would be using some well-known gradient-based ascent schemes. A simple way to generate a feasible ascent direction is for example using the conditional gradient method [42, Sec. 2.2.2]: given the current iterate $\mathbf{Q}^\nu = (\mathbf{Q}_i^\nu)_{i=1}^I$, the next feasible covariance matrix $\mathbf{Q}^{\nu+1}$ is given by

$$\mathbf{Q}^{\nu+1} = \mathbf{Q}^\nu + \gamma^\nu \left(\bar{\mathbf{Q}}^\nu - \mathbf{Q}^\nu\right) \tag{23}$$

where $\bar{\mathbf{Q}}^\nu$ is a solution to the following convex problem:

$$\bar{\mathbf{Q}}^\nu \in \underset{\mathbf{Q} \in \mathcal{Q}}{\arg\max} \Big\{ \sum\nolimits_{i=1}^I \left\langle \nabla_{\mathbf{Q}_i^*} U(\mathbf{Q}^\nu), \mathbf{Q}_i - \mathbf{Q}_i^\nu \right\rangle \Big\}, \tag{24}$$

$\gamma^\nu \in (0, 1]$ is the step-size of the algorithm, $\mathcal{Q}$ is defined in (6), and $\nabla_{\mathbf{Q}_i^*} U(\mathbf{Q})$ denotes the gradient of $U(\mathbf{Q})$ w.r.t. $\mathbf{Q}^*$.

Looking at (24), one infers that conditional gradient methods are based on solving a sequence of convex problems, obtained by linearizing the *whole* utility function $U(\mathbf{Q})$ around $\mathbf{Q}^\nu$. This however does not exploit the structure of $U(\mathbf{Q})$. Since each transmission rate function $r_i(\mathbf{Q}_i, \mathbf{Q}_{-i})$ in $U(\mathbf{Q})$ is a concave function in $\mathbf{Q}_i$, a better exploitation of this "partial" concavity is to replace each linear term $\left\langle \nabla_{\mathbf{Q}_i^*} U(\mathbf{Q}^\nu), \mathbf{Q}_i - \bar{\mathbf{Q}}_i^\nu \right\rangle$ in (24) with a concave approximation of $U(\mathbf{Q})$ at $\mathbf{Q}^\nu$ that leaves $r_i(\mathbf{Q}_i, \mathbf{Q}_{-i})$ untouched and linearizes the other (nonconcave) terms. A natural choice for this approximation is then: given $\mathbf{Q}^\nu = (\mathbf{Q}_i^\nu)_{i=1}^I$,

$$\begin{aligned} \widetilde{U}_i(\mathbf{Q}_i; \mathbf{Q}^\nu) \triangleq\ & r_i(\mathbf{Q}_i, \mathbf{Q}_{-i}^\nu) + \langle \boldsymbol{\Pi}_i(\mathbf{Q}^\nu), \mathbf{Q}_i - \mathbf{Q}_i^\nu \rangle \\ & - \frac{\tau_i}{2} \left\| \mathbf{Q}_i - \mathbf{Q}_i^\nu \right\|^2, \end{aligned} \tag{25}$$

where $\boldsymbol{\Pi}_i(\mathbf{Q}^\nu)$ is the linearization of the terms in $U(\mathbf{Q})$ that are not concave in $\mathbf{Q}_i$, that is

$$\boldsymbol{\Pi}_i(\mathbf{Q}^\nu) \triangleq \sum\nolimits_{j \neq i} \nabla_{\mathbf{Q}_i^*} r_j(\mathbf{Q}^\nu) = \sum\nolimits_{j \neq i} \mathbf{H}_{ji}^H \widetilde{\mathbf{R}}_j(\mathbf{Q}^\nu) \mathbf{H}_{ji}, \tag{26}$$

where $\widetilde{\mathbf{R}}_j(\mathbf{Q}) \triangleq \left(\mathbf{R}_j(\mathbf{Q}_{-j}) + \mathbf{H}_{jj}^H \mathbf{Q}_j \mathbf{H}_{jj}\right)^{-1} - \mathbf{R}_j(\mathbf{Q}_{-j})^{-1}$, and the last (quadratic) term in (25), with $\tau_i > 0$, is a proximal regularization, whose numerical benefits are well-known [35]. Therefore, $\widetilde{U}_i(\mathbf{Q}; \mathbf{Q}^\nu)$ is a strongly concave function in $\mathbf{Q}_i$ obtained by preserving the original concave part $r_i(\mathbf{Q}_i, \mathbf{Q}_{-i})$ in $U(\mathbf{Q})$, while linearizing the nonconcave part $\sum_{j \neq i} r_j(\mathbf{Q})$. We denote by $c_\tau > 0$ the constant of uniformly strong concavity of $\widetilde{U}(\bullet; \mathbf{Q}^\nu)$ on $\mathcal{Q}$, for any $\mathbf{Q}^\nu \in \mathcal{Q}$, i.e., c_τ is the smallest positive scalar such that

$$\begin{aligned} & \left\langle \mathbf{Q}^1 - \mathbf{Q}^2, \nabla_{\mathbf{Q}^*} \widetilde{U}(\mathbf{Q}^1; \mathbf{Q}) - \nabla_{\mathbf{Q}^*} \widetilde{U}(\mathbf{Q}^2; \mathbf{Q}) \right\rangle \\ & \leq -c_\tau \left\| \mathbf{Q}^1 - \mathbf{Q}^2 \right\|, \quad \forall \mathbf{Q}^1, \mathbf{Q}^2, \mathbf{Q} \in \mathcal{Q}. \end{aligned} \tag{27}$$

Note that $c_\tau \geq \min_{i=1,\ldots,I}\{\tau_i\}$, and the equality is reached when all channels $\{\mathbf{H}_{ii}\}_{i=1}^I$ are column rank deficient.

Associated with each $\widetilde{U}_i(\mathbf{Q}; \mathbf{Q}^\nu)$ we can define the following "best response" map that resembles (24): given $\mathbf{Q}^\nu$,

$$\begin{aligned} & \mathcal{Q} \ni \mathbf{Q}^\nu \mapsto \hat{\mathbf{Q}}(\mathbf{Q}^\nu) \triangleq \\ & \qquad \underset{\mathbf{Q} \in \mathcal{Q}}{\arg\max} \left\{ \widetilde{U}(\mathbf{Q}; \mathbf{Q}^{(n)}) \triangleq \sum\nolimits_{i=1}^I \widetilde{U}_i(\mathbf{Q}_i; \mathbf{Q}^\nu) \right\}. \end{aligned} \tag{28}$$

Note that $\hat{\mathbf{Q}}(\mathbf{Q}^\nu)$ is always well-defined, since the problem in (28) is strongly convex and it thus has a unique solution.

According to the above discussion, the proposed candidate search direction at point $\mathbf{Q}^\nu$ in (23) becomes $\hat{\mathbf{Q}}(\mathbf{Q}^\nu) - \mathbf{Q}^\nu$, which leads to the SCA-based algorithm formally described in Algorithm 3. The challenging question now is whether such a direction is still an ascent direction for the sum-rate function $U(\mathbf{Q})$ in (8) at $\mathbf{Q}^\nu$ (as $\bar{\mathbf{Q}}^\nu - \mathbf{Q}^\nu$ is in conditional gradient methods), and how to choose the free parameters τ_i's and γ^ν's in order to guarantee convergence to a stationary solution of the nonconvex problem (8). This is addressed in Theorem 9.

Algorithm 3: Exact Centralized Robust Sum-Rate Maximization Algorithm

Data: $\mathbf{Q}_i^0 \in \mathcal{Q}_i$ for $i = 1, \ldots, I$; $\{\gamma^\nu\} > 0$, $(\tau_i)_{i=1}^I > 0$; set $\nu = 0$;
(S.1): If $\mathbf{Q}^\nu$ satisfies a termination criterion: STOP.
(S.2): Compute the best-response $\hat{\mathbf{Q}}(\mathbf{Q}^\nu)$ solving (28);
(S.3): Set $\mathbf{Q}^{\nu+1} = \mathbf{Q}^\nu + \gamma^\nu(\hat{\mathbf{Q}}(\mathbf{Q}^\nu) - \mathbf{Q}^\nu)$.
(S.4): $\nu \leftarrow \nu + 1$ and go back to (S.1).

Theorem 9. *Given the social problem (8), suppose that: i) $(\tau_i)_{i=1}^I > \mathbf{0}$[1], and ii) the sequence $\{\gamma^\nu\}_{\nu=1}^\infty$ is chosen so that*

$$\gamma^\nu \in (0,1], \quad \gamma^\nu \to 0, \quad \textit{and} \quad \sum\nolimits_\nu \gamma^\nu = +\infty. \tag{29}$$

Then, either Algorithm 3 converges in a finite number of iterations to a stationary point of (8), or every limit point of the sequence $\{\mathbf{Q}^\nu\}$ (at least one such a point exists) is a stationary point of (8). Moreover, none of such points is a local minimum of (8).

Proof: See Appendix B. ■

Remark 10 (On Algorithm 3). The algorithm implements a novel convergent SCA decomposition: a stationary solution of (8) is computed by solving a sequence of strongly convex optimization problems in the form of (28). A suitable termination criterion in Step 1 is $|U(\mathbf{Q}^n) - U(\mathbf{Q}^{n-1})| < \epsilon$, where ϵ is the prescribed accuracy. The algorithm is expected to perform better than classical gradient-based schemes (at least in terms of convergence speed) at the cost of no extra signaling, because the structure of the objective functions is better preserved. We remark that this new decomposition is the nontrivial generalization to nonconvex sum-utility problems with *coupling* (convex) constraints of the technique proposed in [24] for nonconvex problems with *local* constraints only.

Convergence of Algorithm 3 is guaranteed if the step-size sequence satisfies the diminishing step-size rule (29). An example of such a rule is: given $\gamma^0 = 1$,

$$\gamma^{\nu+1} = \gamma^\nu (1 - \alpha\gamma^\nu), \quad \nu = 1, \ldots, \tag{30}$$

where $\alpha \in (0,1)$ is a given constant; see [24] for other alternative rules. We remark that a *constant* step-size $\gamma^\nu = \gamma$ can also be used; convergence is guaranteed if γ is "sufficiently small". We omit the details because of space limit.

Discussion on the implementation. To be implemented, Algorithm 3 needs a CR fusion center, collecting SU-to-SU CSI $(\mathbf{H}_{ij})_{i,j}$, SU-to-PU imperfect estimates $(\widehat{\mathbf{B}}_{ij})_{i,j}$, and the confident intervals $(\epsilon_{pi})_{p,i}$ [cf. (4)]. Once this information is available, problem (28) can be solved in a centralized fashion (Step 3 of Algorithm 3), and the solution of (8) is then broadcast to the SUs. We can reduce the computational complexity of Algorithm 3 by allowing *inexact* computations of the best-response $\hat{\mathbf{Q}}(\mathbf{Q}^\nu)$ in (28). The inexact version of Algorithm 3 is given in Algorithm 4, and its convergence conditions are stated in Theorem 11.

[1]If the channel $\mathbf{H}_{ii}$ is full column rank, one can also choose $\tau_i = 0$.

Algorithm 4: Inexact Centralized Robust Sum-Rate Maximization Algorithm

Data: $\mathbf{Q}_i^0 \in \mathcal{Q}_i$ for $i = 1, \ldots, I$; $\{\gamma^\nu\} > 0$, $(\tau_i)_{i=1}^I > 0$; set $\nu = 0$;
(S.1): If $\mathbf{Q}^\nu$ satisfies a termination criterion: STOP.
(S.2): Find $\mathbf{Z}^\nu$ such that $\left\|\mathbf{Z}^\nu - \hat{\mathbf{Q}}(\mathbf{Q}^\nu)\right\| \le \epsilon^\nu$;
(S.3): Set $\mathbf{Q}^{\nu+1} = \mathbf{Q}^\nu + \gamma^\nu(\mathbf{Z}^\nu - \mathbf{Q}^\nu)$.
(S.4): $\nu \leftarrow \nu + 1$ and go back to (S.1).

Theorem 11. *Let $\{\mathbf{Q}^\nu\}$ be the sequence generated by Algorithm 4, under the setting of Theorem 9. Suppose that the sequences $\{\gamma^\nu\}$ and $\{\epsilon^\nu\}$ satisfy the following conditions: i) $\gamma^\nu \in (0,1]$; ii) $\gamma^\nu \to 0$; iii) $\sum_\nu \gamma^\nu = +\infty$; iv) $\sum_\nu (\gamma^\nu)^2 < +\infty$; and v) $\sum_\nu \epsilon_i^\nu \gamma^\nu < +\infty$ for all $i = 1, \ldots, I$. Then, either Algorithm 4 converges in a finite number of iterations to a stationary point of (8), or every limit point of $\{\mathbf{Q}^\nu\}$ (at least one such a point exists) is a stationary point of (8).*

Proof: See Appendix B. ■

In Algorithm 4, each subproblem (28) is solved within the accuracy ϵ_i^ν (cf. Step 2). As expected, in the presence of errors, convergence is guaranteed if the sequence of approximated subproblems are solved within an increasing accuracy. Note that, in addition to requiring $\epsilon_i^\nu \to 0$, condition v) of Theorem 11 imposes also a constraint on the rate ϵ_i^ν goes to zero, which depends on the decreasing rate of $\{\gamma^\nu\}$. An example of error sequence satisfying condition v) is $\epsilon_i^\nu \le \beta_i \gamma^\nu$, where β_i is any finite positive constant. Besides, the step-size rule (30) satisfies the summability condition iv).

To alleviate the heavy communication overhead associated with the (centralized) implementation of Algorithms 3 and 4, it is desirable to obtain a distributed version of these schemes. Interestingly, the proposed decomposition technique lends itself to a distributed optimization procedure that can be implemented in an on-line fashion. Indeed, the additive structure in both the approximation function $\widetilde{U}(\mathbf{Q}; \mathbf{Q}^{(n)})$ and the coupling constraints [see (5)] is suitable for a parallel decomposition of each subproblem (28) across the SUs; this can be obtained using standard primal or dual decomposition techniques, while preserving the convergence of the algorithms. Note that this is possible because each of the subproblems (28) is convex (e.g., under Slater's constraint qualification, we have strong duality). This is a major departure from previous works in the literature using SCA-based techniques (see, e.g., [5, 7]), where the lack of convexity and strong duality does not guarantee the convergence of the developed solution schemes. The proposed primal and dual distributed implementations of Step 2 of Algorithm 3 (and Algorithm 4) are described in the next two subsections.

B. Distributed dual-decomposition-based algorithms

The subproblem (28) can be solved in a distributed manner if the coupling constraints are relaxed into the Lagrangian (see, e.g., [23]). We rewrite (28) as

$$\begin{array}{ll} \underset{(\mathbf{Q}_i, \mathbf{t}_i)_{i=1}^I}{\text{maximize}} & \sum_{i=1}^I \widetilde{U}_i(\mathbf{Q}; \mathbf{Q}^\nu) \\ \text{subject to} & (\mathbf{Q}_i, \mathbf{t}_i) \in \widetilde{\mathcal{Q}}_i, \quad i = 1, \ldots, I, \\ & \sum_{i=1}^I t_{pi} \le I_p^{\max}, \quad p = 1, \ldots, P, \end{array} \tag{31}$$

with $\widetilde{\mathcal{Q}}_i$ defined in (9). The dual problem of (31) is

$$\underset{\boldsymbol{\mu}\geq\mathbf{0}}{\text{minimize}} \quad D(\boldsymbol{\mu};\mathbf{Q}^\nu), \tag{32}$$

where

$$D(\boldsymbol{\mu};\mathbf{Q}^\nu) \triangleq \max_{(\mathbf{Q}_i,\mathbf{t}_i)\in\widetilde{\mathcal{Q}}_i} \left\{\sum_{i=1}^{I}\left(\widetilde{U}_i(\mathbf{Q}_i;\mathbf{Q}^\nu)-\langle\boldsymbol{\mu},\mathbf{t}_i\rangle\right)\right\}+\langle\boldsymbol{\mu},\mathbf{I}^{\max}\rangle \tag{33}$$

and $\mathbf{I}^{\max}\triangleq(I_p^{\max})_{p=1}^P$. Note that strong duality holds for (31) [and thus (28)]. We can then focus on the dual formulation (32), which is suitable for a distributed implementation because the maximization in (33) can be performed in parallel by the SUs while the optimization problem for each SU is a tractable convex optimization problem [the semi-infinite constraint $\phi_{pi}(\mathbf{Q}_i)\leq t_i$ can be reformulated as a LMI (21)].

The next lemma summarizes some desirable properties of the dual function $D(\boldsymbol{\mu};\mathbf{Q}^\nu)$, where $(\mathbf{Q}^\star(\boldsymbol{\mu};\mathbf{Q}^\nu),\ \mathbf{t}^\star(\boldsymbol{\mu};\mathbf{Q}^\nu))$ denotes the unique solution of the inner maximization in (33), for a given $\boldsymbol{\mu}\geq\mathbf{0}$ and $\mathbf{Q}^\nu$. The proof of the lemma follows the same line of analysis as in Appendix A, and thus is omitted.

Lemma 12. *The dual function $D(\boldsymbol{\mu};\mathbf{Q}^\nu)$ is differentiable on $\mathbb{R}_+^P$, with gradient*

$$\nabla_{\boldsymbol{\mu}}D(\boldsymbol{\mu};\mathbf{Q}^\nu)=\sum\nolimits_{i=1}^{I}\mathbf{t}_i^\star(\boldsymbol{\mu};\mathbf{Q}^\nu)-\mathbf{I}^{\max}.$$

The gradient $\nabla_{\boldsymbol{\mu}}D(\boldsymbol{\mu};\mathbf{Q}^\nu)$ is Lipschitz continuous on $\mathbb{R}_+^P$ with Lipschitz constant $L_{\nabla D}$, given by

$$L_{\nabla D}=\frac{c_\tau}{\sum_{p=1}^{P}\left(\sum_{i=1}^{I}L_{\phi,pi}\right)^2} \tag{34}$$

with $L_{\phi,pi}$ defined in (19) and c_τ defined in (27).

The dual problem can be solved, e.g., using well-known gradient algorithms; an instance is given in Algorithm 5, whose convergence is stated in Theorem 13. The proof of the theorem follows from Lemma 12 and standard convergence results of gradient projection algorithms [43].

Algorithm 5: Dual-based Distributed Implementation of Step 2 of Algorithm 3

Data: $\boldsymbol{\mu}^0\geq\mathbf{0}$, $\mathbf{Q}^\nu$, $\{s^n\}_{n=0}^\infty$; set $n=0$.
(S.2a): If $\boldsymbol{\mu}^n$ satisfies a suitable termination criterion: STOP.
(S.2b): The SUs solve in parallel the following optimization problems: for all $i=1,\ldots,I$,

$$\underset{(\mathbf{Q}_i,\mathbf{t}_i)\in\widetilde{\mathcal{Q}}_i}{\text{maximize}} \quad \widetilde{U}_i(\mathbf{Q}_i;\mathbf{Q}^\nu)-\langle\boldsymbol{\mu}^n,\mathbf{t}_i\rangle \tag{35}$$

(S.2c): Update $\boldsymbol{\mu}$ according to

$$\mu_p^{n+1}=\left[\mu_p^n+s^n\left(\sum\nolimits_{i=1}^{I}t_{pi}^\star\left(\boldsymbol{\mu}^n;\mathbf{Q}^\nu\right)-I_p^{\max}\right)\right]^+,\ \forall\, p. \tag{36}$$

(S.2d): $n\leftarrow n+1$ and go back to (S.2a).

Theorem 13. *Given the dual problem (32), suppose that the step-size sequence $\{s^n\}_{n=0}^\infty$ satisfies one of the following two conditions:*

$$\begin{aligned}&(a):\ 0<s^n=s<2/L_{\nabla D},\\ &(b):\ s^n\to 0,\quad \textstyle\sum_n s^n=\infty,\quad \sum_n\left(s^n\right)^2<\infty.\end{aligned} \tag{37}$$

Then, the sequence $\{\boldsymbol{\mu}^n\}$ generated by Algorithm 5 converges to a solution of (32).

Remark 14 (On Algorithm 5). The price update in (36) is computational inexpensive and can be performed in parallel among PUs. Note that a step-size rule satisfying condition (b) in Theorem 13 is the one in (30), but with an arbitrary $s^0>0$ and any $\alpha\in(0,\,1/s^0)$. When there is only one PU (i.e., $P=1$), the gradient projection update in (36) can be replaced by a bisection search, which in general converges quite fast. When there are multiple PUs, the potential slow convergence of the gradient update can be alleviated using accelerated gradient-based update; we omit the details because of space limit, see, e.g., [44]. A last remark is to observe that the size of the dual problem is equal to the number P of the PUs, but independent of the number of SUs. This makes Algorithm 5 scalable with the number of SUs.

Algorithm 5 can also be used to implement (Step 2 of) Algorithm 4 in a distributed way. To do that, the only issue to discuss is how to choose the termination criterion in Step 2a of Algorithm 5 so that the desired accuracy ϵ^ν is reached in Step 2 of Algorithm 4, that is $\left\|\mathbf{Z}^\nu-\hat{\mathbf{Q}}(\mathbf{Q}^\nu)\right\|\leq\epsilon^\nu$. Recalling the definition of $\hat{\mathbf{Q}}(\mathbf{Q}^\nu)$ in (28) and using the saddle-point connection $\hat{\mathbf{Q}}(\mathbf{Q}^\nu)=\mathbf{Q}^\star(\boldsymbol{\mu}^\star;\mathbf{Q}^\nu)$ [$\boldsymbol{\mu}^\star$ is an optimal solution of (32)], it is not difficult to show that (see Appendix A for a similar line of analysis):

$$\left\|\mathbf{Q}^\star(\boldsymbol{\mu})-\hat{\mathbf{Q}}(\mathbf{Q}^\nu)\right\|\leq\sqrt{\frac{L_{\nabla D}}{c_\tau}}\left\|\boldsymbol{\mu}-\boldsymbol{\mu}^\star\right\|. \tag{38}$$

It turns out that terminating Step 2 of Algorithm 5 when the following accuracy in the price variable $\boldsymbol{\mu}$ is reached

$$\left\|\boldsymbol{\mu}^n-\boldsymbol{\mu}^\star\right\|\leq\sqrt{\frac{c_\tau}{L_{\nabla D}}\sum\nolimits_{i=1}^{I}\left(\epsilon_i^\nu\right)^2} \tag{39}$$

guarantees that $\left\|\mathbf{Z}^\nu-\hat{\mathbf{Q}}(\mathbf{Q}^\nu)\right\|\leq\epsilon^\nu$ is satisfied. Interestingly, under the linear independence constraint qualification (LICQ) of (31) at $\hat{\mathbf{Q}}(\mathbf{Q}^\nu)$ [31, Sec. 3.2], i.e., the gradient of the *active* robust global interference constraints are linearly independent at $\hat{\mathbf{Q}}(\mathbf{Q}^\nu)$, the dual optimal variable $\boldsymbol{\mu}^\star$ is unique, and condition (39) can be forced in a distributed way by using classical error bound results in convex analysis; see, e.g., Propositions 6.2.1 and 6.3.3 of [31]. We omit further details because of space limit.

Remark 15 (On the LICQ of (31)). When the channels $\widehat{\mathbf{B}}_{pi}$ are known with no error, $\phi_{pi}(\mathbf{Q}_i)$ is differentiable and the LICQ of (31) holds if the following matrix is full row rank:

$$\left[\operatorname{vec}\left(\mathbf{I}\otimes(\widehat{\mathbf{B}}_{p1}^H\widehat{\mathbf{B}}_{p1})\right)\ \ldots\ \operatorname{vec}\left(\mathbf{I}\otimes(\widehat{\mathbf{B}}_{pI}^H\widehat{\mathbf{B}}_{pI})\right)\right],\quad p\in\mathcal{I}^\nu,$$

where $\mathcal{I}^\nu\triangleq\left\{p:\ \sum_{i=1}^{I}\phi_{pi}(\hat{\mathbf{Q}}_i(\mathbf{Q}^\nu))=I_p^{\max}\right\}$, i.e., $\mathcal{I}^\nu$ is the set of active global interference constraints. This can be satisfied almost surely if $P\leq\sum_i n_{T_i}^2$ (since the channels $\widehat{\mathbf{B}}_{pi}$ are independent), which is generally true in practice (the number of SUs is much larger than the number of PUs). When the channels are imperfect and $\phi_{pi}(\mathbf{Q}_i)$ is non-differentiable, one can work on the alternative (differentiable) formulation by rewriting the non-differentiable semi-infinite constraint $\phi_{pi}(\mathbf{Q}_i)\leq t_{pi}$ as a linear generalized inequality constraint (21), whose LICQ can be formulated based on the line of

analysis given in [45, Sec. 5.3]; we omit the details because of space limit.

Discussion on the implementation. Combining Algorithm 5 with Algorithm 3 or Algorithm 4, we obtain a distributed *double-loop* scheme converging to a stationary solution of the nonconvex social problem (8). The inner loop deals with the update of the price variables (Algorithm 5): given $\mathbf{Q}^\nu = (\mathbf{Q}_i^\nu)_{i=1}^I$, $(\mathbf{\Pi}_i(\mathbf{Q}^\nu))_{i=1}^I$, and $\boldsymbol{\mu}^n = (\mu_p^n)_{p=1}^P$, at each iteration n, the SUs solve in *parallel* their strongly convex optimization problems (35), resulting in the optimal solution $(\mathbf{Q}_i^\star(\boldsymbol{\mu}^n;\mathbf{Q}^\nu), \mathbf{t}_i^\star(\boldsymbol{\mu}^n;\mathbf{Q}^\nu))_{i=1}^I$. Note that based on (21) (and similar to (22)), (35) can be transformed into a tractable form amenable for many solvers like `SeDuMi` [41]. Once the new values $\sum_{i=1}^I t_{pi}^\star(\boldsymbol{\mu}^n;\mathbf{Q}^\nu)$ are available, the new price μ_p^{n+1} can be computed by an inexpensive projection onto $\mathbb{R}_+$. In the outer loop, which happens once the inner loop price sequence $\{\boldsymbol{\mu}^n\}$ has reached convergence (within the desired accuracy), one just updates the covariance matrices $\mathbf{Q}_i^\nu$ according to $\mathbf{Q}_i^{\nu+1} = \mathbf{Q}_i^\star(\boldsymbol{\mu}^\infty;\mathbf{Q}^\nu)$, where $\boldsymbol{\mu}^\infty$ is the limit point of $\{\boldsymbol{\mu}^n\}$. Overall, the algorithm can be interpreted as a single-scale scheme: the prices are kept being updated according to (36) [which requires the SUs keep solving their optimization problems (35)], and "from time to time" [more precisely when the accuracy in the inner loop is reached, e.g., (39) is satisfied] the objective functions in the SUs optimization problems (35) are changed by updating the pricing matrices $(\mathbf{\Pi}_i(\mathbf{Q}^\nu))_{i=1}^I$ and the regularization terms $(\|\mathbf{Q}_i - \mathbf{Q}_i^\nu\|^2)_{i=1}^I$.

The natural candidates for updating the prices (the inner loop) are the PUs, after receiving the worst-case interference values $t_{pi}^\star(\boldsymbol{\mu}^n;\mathbf{Q}^\nu)$ from the SUs. The SUs take care implicitly of the outer loop solving in parallel the convex optimization problems (35). In CR scenarios where the PU cannot participate in the updating process, the SUs can perform also the operations in the inner loop, at the cost of more signaling, e.g., using consensus algorithms; the implementation solutions proposed in Section III-C apply also to Algorithm 5.

As far as the communication overhead is concerned, the same remarks we have made for the algorithms in Section III-C apply here. The only difference is that now, in order to solve (35), each SU i needs to exchange some extra signaling to estimate the (pricing) matrix $\mathbf{\Pi}_i(\mathbf{Q})$, which can be obtained from the neighboring links via local message passing. This leads to an extra $\mathcal{O}(I^2 n_R)$ amount of message exchange per-iteration w.r.t. the game-theoretical solution methods in Section III-B, where n_R is the number of receive antennas (assumed here equal for all the receivers). This extra communication overhead however comes in favor of better performance; we numerically compare the proposed game-theoretical and social-oriented algorithms (in terms of achievable sum-rate) in Section V, sheding light on the achievable trade-off between performance and signaling.

As a final remark note that, since Algorithm 5 is based on relaxation of the coupling interference constraints into the Lagrangian, it may happen that these constraints are violated by the transmit covariance matrices during the intermediate iterations. This issue can be alleviated in practice by choosing a "large" $\boldsymbol{\mu}^0$ as the initial price. In the next subsection, we propose an alternative distributed scheme which does not suffer from this issue; we cope with the coupling interference constraints using a primal-based decomposition.

C. Distributed primal-decomposition-based algorithms

In this section we provide a primal-based decomposition for problem (28) in Step 2 of Algorithm 3. This scheme is suitable for an on-line implementation, because the global interference constraints are satisfied during all network operations.

We start rewriting (28) into the following equivalent form by introducing the slack variables $(\kappa_{pi})_{p,i}$:

$$\begin{array}{ll} \underset{\{\mathbf{Q}_i,\mathbf{t}_i,\boldsymbol{\kappa}_i\}}{\text{maximize}} & \sum_{i=1}^I \widetilde{U}_i(\mathbf{Q}_i;\mathbf{Q}^\nu) \\ \text{subject to} & (\mathbf{Q}_i,\mathbf{t}_i) \in \widetilde{\mathcal{Q}}_i, \quad i=1,\ldots,I, \\ & t_{pi} \le \kappa_{pi}, \quad i=1,\ldots,I,\ p=1,\ldots,P, \\ & \sum_{i=1}^I \kappa_{pi} \le I_p^{\max}, \quad p=1,\ldots,P, \end{array} \tag{40}$$

where κ_{pi} can be interpreted as the interference budget assigned by PU p to SU i. When $\boldsymbol{\kappa}$ is fixed[2], (40) can be decoupled across the SUs, each of them solving

$$\begin{array}{ll} \underset{\mathbf{Q}_i}{\text{maximize}} & \widetilde{U}_i(\mathbf{Q}_i;\mathbf{Q}^\nu) \\ \text{subject to} & (\mathbf{Q}_i,\mathbf{t}_i) \in \widetilde{\mathcal{Q}}_i, \\ & t_{pi} \overset{\lambda_{pi}(\kappa_{pi};\mathbf{Q}^\nu)}{\le} \kappa_{pi}, \quad p=1,\ldots,P, \end{array} \tag{41}$$

where $\lambda_{pi}(\kappa_{pi};\mathbf{Q}^\nu)$ is the dual variable associated with the inequality constraint $t_{pi} \le \kappa_{pi}$. It is easy to verify that strong duality holds for (41) for any $\kappa_{pi} \ge 0$, so the existence of $\lambda_{pi}(\kappa_{pi};\mathbf{Q}^\nu)$ is guaranteed [46, Sec. 9.1.3, Th. 4], but $\lambda_{pi}(\kappa_{pi};\mathbf{Q}^\nu)$ is generally not unique (e.g., this happens when $\kappa_{pi}=0$). We recall that (41) is a tractable convex optimization problem because the semi-infinite constraint $\phi_{pi}(\mathbf{Q}_i) \le t_{pi}$ can be reformulated as a LMI constraint, see (21).

Given $\boldsymbol{\kappa}_i$, the unique solution of (41) is denoted as $\mathbf{Q}_i^\star(\boldsymbol{\kappa}_i;\mathbf{Q}^\nu)$. Then the remaining issue is to find the optimal interference budget $\boldsymbol{\kappa}^\star$ such that the sum-rate in (40) is maximized. This is equivalent to the following so-called *master* (convex) problem [23, 47]:

$$\begin{array}{ll} \underset{\boldsymbol{\kappa}}{\text{maximize}} & g(\boldsymbol{\kappa};\mathbf{Q}^\nu) \triangleq \sum_{i=1}^I \widetilde{U}_i\left(\mathbf{Q}_i^\star(\boldsymbol{\kappa}_i;\mathbf{Q}^\nu)\right) \\ \text{subject to} & \sum_{i=1}^I \kappa_{pi} \le I_p^{\max}, \quad p=1,\ldots,P. \end{array} \tag{42}$$

Due to the non-uniqueness of $\lambda_{pi}(\kappa_{pi};\mathbf{Q}^\nu)$, the objective function in (42) is non-differentiable; problem (42) can be solved by subgradient projection methods [23], and a subgradient of $g(\boldsymbol{\kappa};\mathbf{Q}^\nu)$ at $\boldsymbol{\kappa}=\boldsymbol{\kappa}^n$ is

$$\partial_{\kappa_{pi}} g(\boldsymbol{\kappa}^n;\mathbf{Q}^\nu) = \lambda_{pi}(\kappa_{pi}^n;\mathbf{Q}^\nu), \quad \forall\, p,\, i.$$

The subgradient projection method solving (42) is summarized in Algorithm 6, whose convergence properties are given in Theorem 16; the proof follows from standard convergence results of subgradient algorithms [43, Prop. 8.2.6]. Note that the operator $\Pi_{\mathcal{S}_p}(\mathbf{x})$ in Step 2c) of (43) denotes the Euclidean projection of $\mathbf{x}$ onto the simplex $\mathcal{S}_p \triangleq \left\{\mathbf{x} : \mathbf{x} \ge \mathbf{0}, \mathbf{1}^T\mathbf{x} \le I_p^{\max}\right\}$ ($\mathbf{1}$ is a column vector with all entries equal to 1).

[2]With a slight abuse of notation, we will use the same symbol $\boldsymbol{\kappa}$ to denote two different rearrangements of the components κ_{pi}, namely: i) $\boldsymbol{\kappa} = (\boldsymbol{\kappa}_i)_{i=1}^I$, where by $\boldsymbol{\kappa}_i$ we will mean $\boldsymbol{\kappa}_i \triangleq (\kappa_{pi})_{p=1}^P$; and ii) $\boldsymbol{\kappa} = (\boldsymbol{\kappa}_p)_{p=1}^P$, where by $\boldsymbol{\kappa}_p$ we will mean $\boldsymbol{\kappa}_p \triangleq (\kappa_{pi})_{i=1}^I$.

Algorithm 6 : Primal-based Distributed Implementation of Step 2 of Algorithm 3

Data: $\boldsymbol{\kappa}^0 \geq 0$, $\mathbf{Q}^\nu$, $\{s^n\}_{n=0}^\infty$; set $n=0$.
(S.2a): If $\boldsymbol{\kappa}^n$ satisfies a suitable termination criterion: STOP.
(S.2b): The SUs solve in parallel the following optimization problems: for all $i=1,\ldots,I$,

$$\begin{aligned} \underset{\mathbf{Q}_i,\mathbf{t}_i}{\text{maximize}} \quad & \widetilde{U}_i(\mathbf{Q}_i;\mathbf{Q}^\nu) \\ \text{subject to} \quad & (\mathbf{Q}_i,\mathbf{t}_i)\in\widetilde{\mathcal{Q}}_i, \\ & t_{pi} \overset{\lambda_{pi}(\kappa_{pi}^n;\mathbf{Q}^\nu)}{\leq} \kappa_{pi}^n, \quad p=1,\ldots,P, \end{aligned}$$

(S.2c): Given $\boldsymbol{\lambda}_p(\boldsymbol{\kappa}_p^n;\mathbf{Q}^\nu) \triangleq (\lambda_{pi}(\kappa_{pi}^n;\mathbf{Q}^\nu))_{i=1}^I$, update $\boldsymbol{\kappa}_p \triangleq (\kappa_{pi})_{i=1}^I$ according to

$$\boldsymbol{\kappa}_p^{n+1} = \Pi_{\mathcal{S}_p}\left(\boldsymbol{\kappa}_p^n + s^n\boldsymbol{\lambda}_p(\boldsymbol{\kappa}_p^n;\mathbf{Q}^\nu)\right),\ \forall\, p=1,\ldots,P, \quad (43)$$

(S.2d): $n\leftarrow n+1$ and go back to (S.2a).

Theorem 16. *Suppose that the step-size sequence $\{s^n\}_{n=0}^\infty$ satisfies condition (b) in Theorem 13 [cf. (37)]. Then the sequence $\{\boldsymbol{\kappa}^n\}$ generated by Algorithm 6 converges to an optimal solution of (42).*

Remark 17 (On Algorithm 6). The projection onto the simplex $\mathcal{S}_p$ in (43) has a closed-form solution (up to an unknown scalar which can be found by bisection) [47, Lemma 1], and thus can be performed efficiently. The projection also lends the primal decomposition method suitable for an on-line implementation: the robust global interference constraints are always satisfied at each immediate iteration of Algorithm 6.

Discussion on the implementation. Combining Algorithm 6 with Algorithm 3, we obtain a distributed *double-loop* scheme converging to a stationary solution of the nonconvex social problem (8). The inner loop deals with the update of the interference budget variables (Algorithm 6): given $\mathbf{Q}^\nu=(\mathbf{Q}_i^\nu)_{i=1}^I$, $(\boldsymbol{\Pi}_i(\mathbf{Q}^\nu))_{i=1}^I$, and $\boldsymbol{\kappa}^n$, at each iteration n, the SUs solve in *parallel* their strongly convex optimization problems (41), resulting in the optimal solution $(\mathbf{Q}_i^\star(\boldsymbol{\kappa}_i^n;\mathbf{Q}^\nu),\mathbf{t}_i^\star(\boldsymbol{\kappa}_i^n;\mathbf{Q}^\nu))$ and Lagrange multipliers $\boldsymbol{\lambda}_i=(\lambda_{pi}(\kappa_{pi}^n;\mathbf{Q}^\nu))_{p=1}^P$. Then each SU i passes $\lambda_{pi}^*(\kappa_{pi}^n;\mathbf{Q}^\nu)$ to PU p, who will compute the new interference budget $\boldsymbol{\kappa}_p^{n+1}=(\kappa_{pi}^{n+1})_{i=1}^I$ by performing an inexpensive projection onto a simplex [cf. Remark 17]. After the inner loop converges, one just updates in the outer loop the covariance matrices $\mathbf{Q}_i$ according to $\mathbf{Q}_i^{\nu+1}=\mathbf{Q}_i^\star(\boldsymbol{\kappa}_i^\infty;\mathbf{Q}^\nu)$, where $\boldsymbol{\kappa}^\infty$ is the limit point of $\{\boldsymbol{\kappa}^n\}$. Similar to the dual-based scheme (Algorithm 5), the primal-based implementation can also be interpreted as a single-scale scheme with the only difference that the PUs keep updating the interference budgets $\boldsymbol{\kappa}$ rather than the prices $\boldsymbol{\mu}$; the discussion on implementation issues of Algorithm 5 applies thus also to Algorithm 6, see Section IV-B.

As a final remark, note that in the primal-based scheme, the interference budget κ_{pi} set by PU p for SU i may be *different* from other SUs. This is a major difference with the dual-based implementation in Section IV-B and the game-theoretical schemes in Section III, where each PU p sets an *equal* price μ_p to all SUs. As a result, the PU-to-SU signaling necessary to implement Algorithm 6 is heavier than the one required for Algorithm 5. Interestingly, the signaling in Algorithm 6 is at the same level of state-of-the-art dealing with (robust) interference constraints [5, 7]. Moreover, due to the larger dimension of $\boldsymbol{\kappa}$ (which is $P\cdot I$) and lack of differentiability of $g(\boldsymbol{\kappa};\mathbf{Q}^\nu)$ in the master problem (42), Algorithm 6 is expected to be slower than Algorithm 5 (which solves a P-dimensional differentiable problem using a gradient projection method). These are the prices to pay for having the interference constraints satisfied at each (intermediate) iteration of the Algorithm 6. A numerical comparison of Algorithms 5 and 6 is given in Section V.

V. Numerical Results

In this section we run some numerical tests to show the benefits provided by the proposed framework. We consider a cellular system composed of one base station, the PU (i.e., $P=1$, so $I_p^{\max}$ and $I^{\max}$ are used interchangeably), and multiple secondary links (whose number varies from one figure to another). The PU is equipped with 4 receive antennas, while the SUs have 4 transmit and 4 receive antennas. We assume equal power budget $P_i^{\text{tot}}=P^{\text{tot}}$ and white Gaussian noise with variance $\sigma_i^2=\sigma^2$ for all the SUs; the SNR of each SU is then $\texttt{snr}\triangleq P^{\text{tot}}/\sigma^2$. The distance between each secondary transmitter and receiver is set to one (so that the strengths of direct channels $\{\mathbf{H}_{ii}\}_{i=1}^I$ are comparable), whereas we consider different (normalized) inter-pair distances between the secondary transmitters and receiver, which provides a simple way to control the coupling among the SUs. The (normalized) path loss between each secondary transmitter and the PU is given by $d^{-\alpha}$, where d is the relative distance between the secondary transmitter and the base station; we set $d=10$ and $\alpha=3$. We consider spherical uncertainty regions, i.e., $\mathbf{T}_{pi}=\mathbf{I}$, and $\epsilon_{pi}=0.1\cdot\|\widehat{\mathbf{G}}_{pi}\|$.

A. Robust system design and worst-case interference

In this experiment, we compare the proposed robust global design [Algorithm 1] (termed as "robust-global") with the nonrobust global design [Algorithm 1 with $\varepsilon_{pi}=0$ in (4)] (termed as "nonrobust-global"). As a benchmark, we have also simulated the classical iterative waterfilling algorithm (IWFA) [15], which does not consider any interference constraints. The secondary network is composed of 4 SUs, each with $\texttt{snr}=10$ dB, and the global interference limit is set to be $I^{\max}=0.03$. In Figure 2, we plot the worst-case interference $\sum_{i=1}^I t_{pi}^\star(\boldsymbol{\mu}^\nu)$ obtained by Algorithm 1 versus the iteration index ν. The worst-case interference in the nonrobust-global version of the algorithm is estimated using the largest interference observed by randomly generating 1000 channel realizations in the uncertainty region.

Figure 2 shows that the proposed "robust-global" based on Algorithm 1 (marker of triangle) converges quite fast (in a few iterations) while satisfying the global interference temperature constraints. On the contrary, nonrobust designs such as the global nonrobust version of Algorithm 1 ("nonrobust-global", marker of triangle) and the IWFA (marker of plus) result in a violation of the interference constraint, and a possible degradation in the licensed PUs' performance. This consolidates the necessity of a robust system design to deal with channel uncertainties.

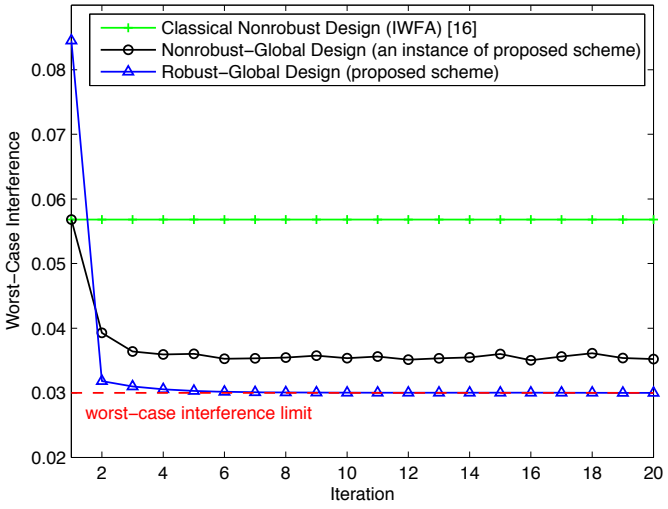


Fig. 2. Worst-case interference $\sum_{i=1}^{I} t_{pi}^{\star}(\boldsymbol{\mu}^{\nu})$ versus iteration number ν in Algorithm 1 for robust and nonrobust formulation.

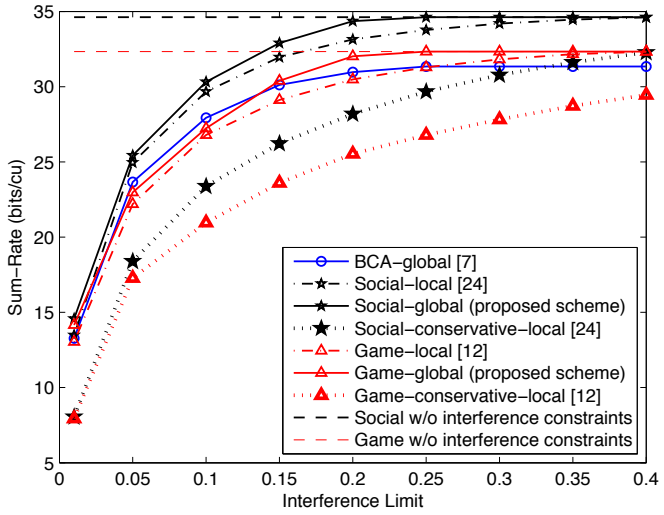


Fig. 3. Sum-rate of all SUs $\sum_{i=1}^{I} r_i(\mathbf{Q})$ at a NE/socially optimal solution versus interference limit $I^{\max}$.

B. Game-theoretical versus social-oriented designs

In this example, we compare the achievable sum-rate by the two proposed alternative designs: the game-theoretical and the social-oriented ones; we also contrast our schemes with state-of-the-art algorithms in literature for similar problems [7, 12, 24]. We consider a secondary network composed of 8 SUs, each with `snr=5` dB, and compare the following algorithms in Figure 3: i) Algorithm 1 (termed as "game-global"); ii) Algorithm 3 (termed as "social-global"); iii) the state-of-the-art block coordinate ascent algorithm (termed as "BCA-global") proposed in [7], minimizing the sum-MSE subject to robust global interference constraints; iv) the algorithm proposed in [12] (termed as "game-local"), solving the rate-maximization game subject to robust local interference constraints; and v) the "social-local" where robust local interference constraints $\phi_{pi}(\mathbf{Q}_i) \leq I_p^{\max}/I$ are considered and the sum-rate can be calculated using the methodology of [24]. Note that "social-local" is a particular instance of (28) by replacing the robust global interference constraints $\sum_{i=1}^{I} \phi_{pi}(\mathbf{Q}_i) \leq I_p^{\max}$ with robust local interference constraints $\phi_{pi}(\mathbf{Q}_i) \leq I_p^{\max}/I$. We also consider the so-called "game-conservative-local" and "social-conservative-local" instances, and they correspond to the scenario in which there are only local interference constraints but the number of SUs I is unknown; in such a case, a conservative estimate on the number of SUs $\bar{I}(= 20 \geq I = 8)$ is used, resulting in conservative robust local interference constraints $\phi_{pi}(\mathbf{Q}_i) \leq I_p^{\max}/\bar{I} \leq I_p^{\max}/I$.

Global vs local interference constraints: Figure 3 shows that both game-theoretical and social-oriented designs under *global* interference constraints ("game-global" and "social-global") outperform those based on *local* interference constraints ("game-local" and "social-local"). This is because global interference constraints provide more flexibility than local constraints to the SUs in satisfying the PUs' requirements. The performance gap between these two approaches becomes even more significant when the exact number of the SUs is not known; indeed in such a case, $\bar{I}$ may be much larger than the real I, leading to even more restrictive local interference constraints $\phi_{pi}(\mathbf{Q}_i) \leq I_p^{\max}/\bar{I} \ll I_p^{\max}/I$. This issue is clear comparing the sum-rate performance achievable by the "game-local" and "social-local" designs with "game-conservative-local" and "social-conservative-local", respectively. Of course the better performance using global interference constraints comes at the cost of extra (albeit limited) signaling among the SUs (see Section IV-B and Section IV-C for a detailed discussion on this).

Comparison with state-of-the-art algorithms: Figure 3 clearly shows that the proposed algorithm "social-global" outperforms the "BCA-global" [7] over all interference limit; the sum-rate gain of our scheme goes from 7.5% to 10% over the simulated range of the interference constraints. We also remark that "BCA-global" is of a Gauss-Seidel type (i.e., the updates of the SUs' strategies occur sequentially), which may incur a large delay when the number of SUs is large. In contrast, the proposed algorithm is based on simultaneous update among the SUs, which makes it much more scalable.

How good is the NE? The last issue to address is quantifying the quality of the NE in terms of achievable sum-rate. From Figure 3, one can see that, as expected, the social-oriented design outperforms the game-theoretical design, but the gap is limited up to 10%, which happens when the interference constraints are the dominant constraints. We recall that better performance is at the cost of more signaling. Figure 3 thus provides an indication on the signaling/performance tradeoff.

C. Convergence of robust sum-rate maximization algorithm

In this experiment, we show the convergence of the exact centralized robust sum-rate maximization algorithm [Algorithm 3]. We consider a CR network composed of 4 (dotted curves) or 8 SUs (solid curves), each SU with three `snr`: 0 dB (marker of square), 10 dB (marker of triangle), and 20 dB (marker of circle). In Figure 4 we plot the sum-rate of the SUs $\sum_{i=1}^{I} r_i(\mathbf{Q}^{\nu})$ versus the iteration index ν; the initial point is set to $\mathbf{Q}_i^0 = \mathbf{0}$ for all $i = 1, \ldots, I$ and stepsize $\gamma^{\nu+1} = \gamma^{\nu}(1 - 10^{-5}\gamma^{\nu})$ with $\gamma^0 = 1$. Figure 4 clearly shows the (asymptotic) ascent property of the proposed algorithm: the sum-rate is increased after each of the simultaneous update of

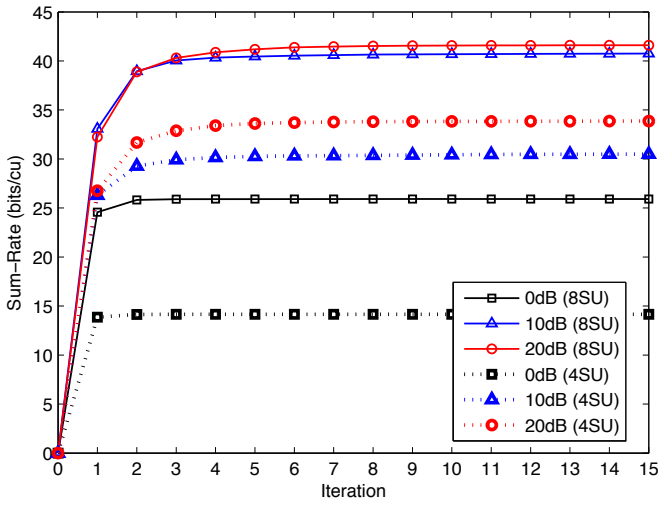


Fig. 4. Convergence of Algorithm 3: Sum-rate of all SUs $\sum_{i=1}^{I} r_i(\mathbf{Q}^\nu)$ versus iteration number ν with $\mathbf{Q}_i^0 = \mathbf{0}$ for all $i = 1, \ldots, I$.

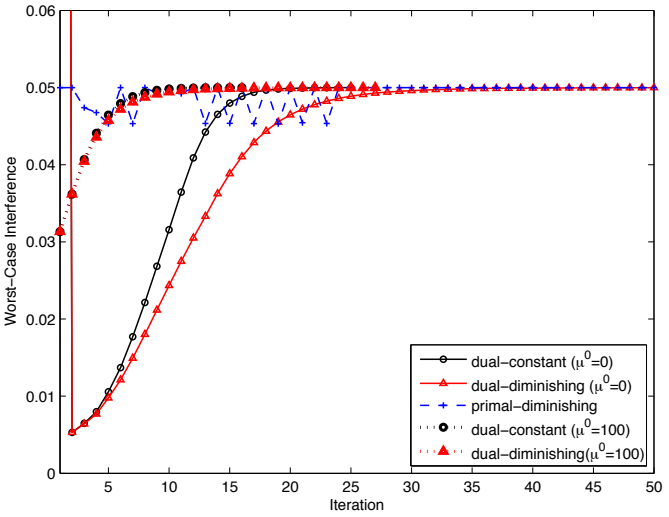


Fig. 5. Worst-case aggregate interference at the PU: $\sum_{i=1}^{I} t_{pi}^{\star}(\boldsymbol{\mu}^n; \mathbf{Q}^\nu)$ in the dual-based Algorithm 5 and $\sum_{i=1}^{I} \kappa_{pi}^n(\mathbf{Q}^\nu)$ in the primal-based Algorithm 6 versus iteration number n for a given $\mathbf{Q}^\nu$.

all SUs. The convergence speed is reasonable fast (less than 10 iterations). More importantly, in the simulated scenarios, it seems that the convergence speed is not affected much by the number of SUs (compare the curves with the same marker), which makes the algorithm scalable and applicable in large-scale CR networks.

D. Distributed implementation: convergence of primal and dual decomposition schemes

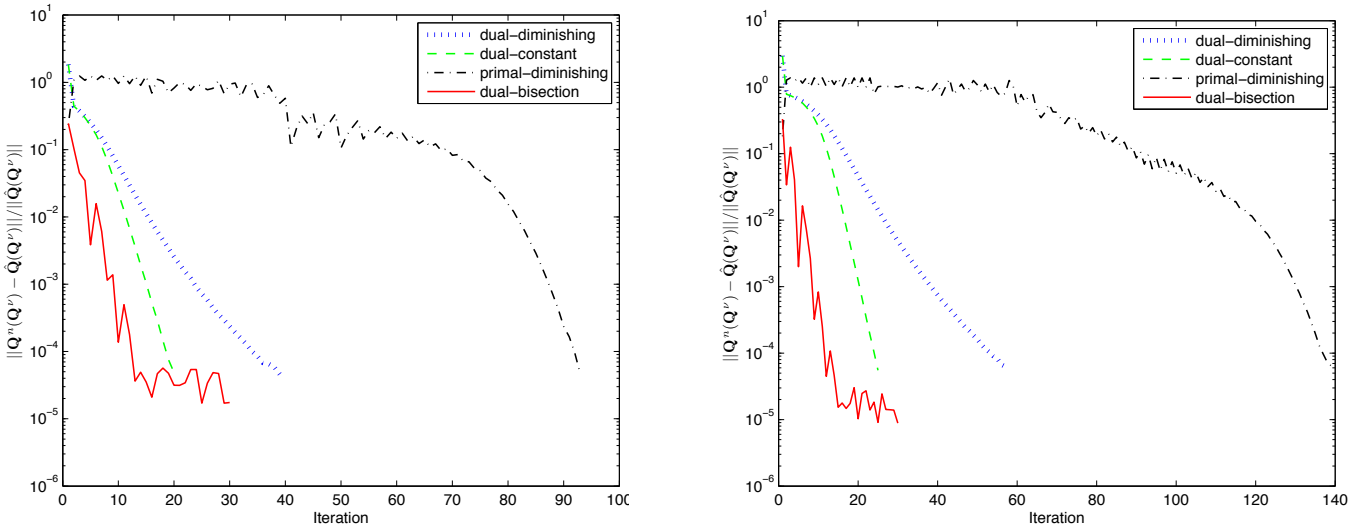
In this experiment, we compare the convergence of dual decomposition [cf. Algorithm 5] and primal decomposition scheme [cf. Algorithm 6] under different stepsize rules, and we plot in Figure 6 the metric $\left\|\mathbf{Q}^n(\mathbf{Q}^\nu) - \mathbf{Q}^{\nu+1}\right\| / \left\|\mathbf{Q}^{\nu+1}\right\|$ versus iteration number n for a given $\mathbf{Q}^\nu$, where $\mathbf{Q}^n(\mathbf{Q}^\nu) = \mathbf{Q}^\star(\boldsymbol{\mu}^n; \mathbf{Q}^\nu)$ and $\mathbf{Q}^n(\mathbf{Q}^\nu) = \mathbf{Q}^\star(\boldsymbol{\kappa}^n; \mathbf{Q}^\nu)$ in the context of dual decomposition and primal decomposition, respectively; $\mathbf{Q}^{\nu+1} = \mathbf{Q}^\star(\boldsymbol{\mu}^\infty; \mathbf{Q}^\nu) = \mathbf{Q}^\star(\boldsymbol{\kappa}^\infty; \mathbf{Q}^\nu)$ and it can be obtained by applying the centralized solver `SeDuMi` to (28). In particular, "dual-constant", "dual-diminishing", and "dual-bisection" refers to dual decomposition method with a constant stepsize rule ($s^n = 500$), a diminishing stepsize rule ($s^{n+1} = s^n(1 - 10^{-4}s^n)$ with $s^0 = 500$), and the bisection method, respectively. "Primal-diminishing" refers to primal decomposition method with a diminishing stepsize rule ($s^{n+1} = s^n(1 - 0.5s^n)$ with $s^0 = 1$).

Figure 6 shows that dual decomposition schemes (under different step-size rules) always converge faster than the primal decomposition schemes. This is because the dual problem is differentiable with a Lipschitz continuous gradient and thus it can be solved by gradient projection method with a constant step-size, and the dual variable has a smaller dimension than the interference budget variable. More specifically, 1) "dual-constant" performs better than "dual-diminishing" because the step-size is constant in "dual-constant" which is however diminishing in "dual-diminishing"; 2) under the same step-size rule, "dual-diminishing" performs better than "primal-diminishing" because the dimension of dual variable $\boldsymbol{\mu}$ in (31) is P while the dimension of interference budget $\boldsymbol{\kappa}$ in (42) is $P \cdot I$; 3) By comparing Figure 6 (a) and (b), we see that the convergence speed of dual decomposition is independent of the number of SUs, but the convergence of primal decomposition is slower when the number of SUs is doubled; 4) When there is only one PU, the dual variable is a single scalar and it can be updated by bisection method, which converges quite fast.

E. Worst-case interference of primal and dual decomposition

Even though dual schemes are faster than primal ones, they do not guarantee that the interference constraints are always satisfied while the algorithm in running, whereas primal algorithms do. This is shown by the experiment in Figure 5, where we plot the worst-case aggregate interference at the PU in each iteration of the dual and primal decomposition. In particular, we plot the metric $\sum_{i=1}^{I} t_{pi}^{\star}(\boldsymbol{\mu}^n; \mathbf{Q}^\nu)$ as generated by the dual-based Algorithm 5 and $\sum_{i=1}^{I} \kappa_{pi}^n(\mathbf{Q}^\nu)$ by the primal-based Algorithm 6 versus the iteration index n, for a given $\mathbf{Q}^\nu$. The number of SUs is 8, each with a `snr` =10 dB, and the interference limit is $I^{\max} = 0.05$. The stepsize parameters are as same as those in Figure 6.

Figure 5 shows that, as expected, the worst-case aggregate interference generated in each iteration by the primal decomposition (blue dashed curve with plus) is always below the interference limit. This is not necessarily true for dual decomposition, because the robust global interference constraints are relaxed into the Lagrangian function to trade for parallel computation among the SUs. It can be observed that the violation of the interference constraints is more likely when the initial Lagrange multiplier μ^0 is small, e.g., $\mu^0 = 0$ (see solid curves); this is because the penalty for interference is small. However, as we have already remarked in Section IV-B, this problem can be alleviated if one starts with a large initial Lagrange multiplier μ^0, corresponding to a large penalty for interference. Indeed, the dotted curve shows that the worst-case aggregate interference with a large μ^0(=100) is always below the interference limit.



(a) Number of SUs: 4

(b) Number of SUs: 8

Fig. 6. Convergence of dual-based Algorithm 5 and primal-based Algorithm 6: $\|\mathbf{Q}^n(\mathbf{Q}^\nu) - \hat{\mathbf{Q}}(\mathbf{Q}^\nu)\|/\|\hat{\mathbf{Q}}(\mathbf{Q}^\nu)\|$ versus iteration number n for a given $\mathbf{Q}^\nu$, where $\mathbf{Q}^n(\mathbf{Q}^\nu) \triangleq \mathbf{Q}^\star(\boldsymbol{\mu}^n;\mathbf{Q}^\nu)$ in Algorithm 5 and $\mathbf{Q}^n(\mathbf{Q}^\nu) \triangleq \mathbf{Q}^\star(\boldsymbol{\kappa}^n;\mathbf{Q}^\nu)$ in Algorithm 6, and $\hat{\mathbf{Q}}(\mathbf{Q}^\nu)$ is defined in (28).

VI. Conclusions

In this paper, we have considered the worst-case robust and decentralized designs for CR systems with multiple primary and secondary users over MIMO IC. We have proposed two alternative approaches, namely: a game-theoretical and a social-oriented approach. In the game-theoretical approach, the CR network design is formulated as a pricing game with robust global interference constraints in the form of a price clearing condition. Building on the theory of complex VIs, we have shown that the game always has an NE, and have proposed distributed algorithms converging (under technical conditions) to a NE of the game. To deal with the potential inefficiencies of NE, we have considered the more classical social-oriented optimization, where the nonconvex sum-rate of the SUs is maximized subject to robust global interference constraints. Building on the framework of successive convex approximation, we have proposed for the first time a class of distributed SCA-based algorithms, based on primal/dual decomposition, with provable convergence to stationary solutions. Overall, the two complementary system designs differ in PU-to-SU and/or SU-to-SU signaling, complexity, convergence speed, and sum-rate performance, and are thus applicable to a variety of CR scenarios, either cooperative or non-cooperative, offering a tool to explore the trade-off between signaling and performance.

Appendix A
The Co-coercivity of $\mathbf{M}(\boldsymbol{\mu})$

The co-coercivity of $\mathbf{M}(\boldsymbol{\mu})$ is stated in the following lemma.

Lemma 18. *If $\boldsymbol{\Upsilon} \succ \mathbf{0}$, there exists a positive constant $c_{\rm coc}$ given by (18) such that for any $\boldsymbol{\mu}^1,\ \boldsymbol{\mu}^2 \geq \mathbf{0}$:*

$$\left\langle \mathbf{M}(\boldsymbol{\mu}^1) - \mathbf{M}(\boldsymbol{\mu}^2), \boldsymbol{\mu}^1 - \boldsymbol{\mu}^2 \right\rangle \geq c_{\rm coc} \left\| \mathbf{M}(\boldsymbol{\mu}^1) - \mathbf{M}(\boldsymbol{\mu}^2) \right\|^2.$$

Proof: Let $(\mathbf{Q}^\star(\boldsymbol{\mu}), \mathbf{t}^\star(\boldsymbol{\mu}))$ be the unique NE of $\mathcal{G}_{\boldsymbol{\mu}}$, for a given $\boldsymbol{\mu} \geq \mathbf{0}$. Then, it follows from the variational principle [21, Lemma 24] that: for all $i = 1, \ldots, I$,

$$\begin{aligned} &\left\langle \mathbf{Q}_i^\star(\boldsymbol{\mu}^2) - \mathbf{Q}_i^\star(\boldsymbol{\mu}^1), \mathbf{F}_i^{\mathbb{C}}\left(\mathbf{Q}^\star(\boldsymbol{\mu}^1)\right)\right\rangle \\ &\qquad + \left\langle \mathbf{t}_i^\star(\boldsymbol{\mu}^2) - \mathbf{t}_i^\star(\boldsymbol{\mu}^1), \boldsymbol{\mu}^1 \right\rangle \geq 0, \end{aligned} \tag{44a}$$

$$\begin{aligned} &\left\langle \mathbf{Q}_i^\star(\boldsymbol{\mu}^1) - \mathbf{Q}_i^\star(\boldsymbol{\mu}^2), \mathbf{F}_i^{\mathbb{C}}\left(\mathbf{Q}^\star(\boldsymbol{\mu}^2)\right)\right\rangle \\ &\qquad + \left\langle \mathbf{t}_i^\star(\boldsymbol{\mu}^1) - \mathbf{t}_i^\star(\boldsymbol{\mu}^2), \boldsymbol{\mu}^2 \right\rangle \geq 0, \end{aligned} \tag{44b}$$

with $\mathbf{F}_i^{\mathbb{C}}(\mathbf{Q})$ defined in (12). Adding (44a) and (44b) we have

$$\begin{aligned} &\left\langle \mathbf{Q}_i^\star(\boldsymbol{\mu}^1) - \mathbf{Q}_i^\star(\boldsymbol{\mu}^2), \mathbf{F}_i^{\mathbb{C}}\left(\mathbf{Q}^\star(\boldsymbol{\mu}^1)\right) - \mathbf{F}_i^{\mathbb{C}}\left(\mathbf{Q}^\star(\boldsymbol{\mu}^2)\right)\right\rangle \\ &\leq \left\langle -\mathbf{t}_i^\star(\boldsymbol{\mu}^1) + \mathbf{t}_i^\star(\boldsymbol{\mu}^2), \boldsymbol{\mu}^1 - \boldsymbol{\mu}^2 \right\rangle. \end{aligned} \tag{45}$$

Summing (45) over i and recalling the definition of $\mathbf{M}(\boldsymbol{\mu})$, we obtain:

$$\begin{aligned} &\left\langle \mathbf{M}(\boldsymbol{\mu}^1) - \mathbf{M}(\boldsymbol{\mu}^2), \boldsymbol{\mu}^1 - \boldsymbol{\mu}^2 \right\rangle \\ &= \left\langle -\sum\nolimits_{i=1}^I \mathbf{t}_i^\star\left(\boldsymbol{\mu}^1\right) + \sum\nolimits_{i=1}^I \mathbf{t}_i^\star\left(\boldsymbol{\mu}^2\right), \boldsymbol{\mu}^1 - \boldsymbol{\mu}^2 \right\rangle \\ &\geq \left\langle \mathbf{Q}^\star(\boldsymbol{\mu}^1) - \mathbf{Q}^\star(\boldsymbol{\mu}^2), \mathbf{F}^{\mathbb{C}}\left(\mathbf{Q}^\star(\boldsymbol{\mu}^1)\right) - \mathbf{F}^{\mathbb{C}}\left(\mathbf{Q}^\star(\boldsymbol{\mu}^2)\right)\right\rangle \end{aligned} \tag{46}$$

It is shown in [21, Prop. 38] that if $\mathbf{H}_{ii}$ is full column rank for all i and $\boldsymbol{\Upsilon} \succ \mathbf{0}$, $\mathbf{F}^{\mathbb{C}}(\mathbf{Q})$ is strongly monotone [21, Prop. 29] on $\mathcal{Q}$: there exists a positive constant $c_{\rm sm}$ [cf. (18)] such that

$$\begin{aligned} &\left\langle \mathbf{Q}^\star(\boldsymbol{\mu}^1) - \mathbf{Q}^\star(\boldsymbol{\mu}^2), \mathbf{F}^{\mathbb{C}}\left(\mathbf{Q}^\star(\boldsymbol{\mu}^1)\right) - \mathbf{F}^{\mathbb{C}}\left(\mathbf{Q}^\star(\boldsymbol{\mu}^2)\right)\right\rangle \\ &\geq c_{\rm sm} \left\| \mathbf{Q}^\star(\boldsymbol{\mu}^1) - \mathbf{Q}^\star(\boldsymbol{\mu}^2) \right\|^2. \end{aligned} \tag{47}$$

On the other hand, it follows from [34, Th. 10.4] that the functions $\phi_{pi}(\mathbf{Q}_i)$ are Lipschitz continuous, i.e., there exist positive constants $L_{\phi,pi}$ [cf. (19)] such that

$$\begin{aligned} &\left\| \mathbf{M}(\boldsymbol{\mu}^1) - \mathbf{M}(\boldsymbol{\mu}^2) \right\|^2 \\ &= \sum\nolimits_{p=1}^P \left\| \sum\nolimits_{i=1}^I \left[\phi_{pi}\left(\mathbf{Q}_i^\star(\boldsymbol{\mu}^1)\right) - \phi_{pi}\left(\mathbf{Q}_i^\star(\boldsymbol{\mu}^2)\right)\right] \right\|^2 \\ &\leq \left[\sum\nolimits_{p=1}^P \left(\sum\nolimits_{i=1}^I L_{\phi,pi} \right)^2 \right] \left\| \mathbf{Q}^\star(\boldsymbol{\mu}^1) - \mathbf{Q}^\star(\boldsymbol{\mu}^2) \right\|^2, \end{aligned} \tag{48}$$

where in (48) we have also used the fact that $\left\|\mathbf{Q}_i^\star(\boldsymbol{\mu}^1)-\mathbf{Q}_i^\star(\boldsymbol{\mu}^2)\right\| \le \left\|\mathbf{Q}^\star(\boldsymbol{\mu}^1)-\mathbf{Q}^\star(\boldsymbol{\mu}^2)\right\|$.

Combining (46)-(48), we obtain the desired result

$$\left\langle \mathbf{M}(\boldsymbol{\mu}^1)-\mathbf{M}(\boldsymbol{\mu}^2), \boldsymbol{\mu}^1-\boldsymbol{\mu}^2\right\rangle \ge \frac{c_{\text{sm}}}{\sum_{p=1}^P\left(\sum_{i=1}^I L_{\phi,pi}\right)^2}\left\|\mathbf{M}(\boldsymbol{\mu}^1)-\mathbf{M}(\boldsymbol{\mu}^2)\right\|^2,$$

which completes the proof. ∎

APPENDIX B
PROOF OF THEOREM 9 AND THEOREM 11

A key result to prove convergence of the SCA-based algorithm is to show that the proposed new feasible direction $\hat{\mathbf{Q}}(\mathbf{Q})-\mathbf{Q}$ is an ascent direction of $U(\mathbf{Q})$ (cf. (8)) at $\mathbf{Q}$. This is proved in Proposition 19 below along with some other properties of the best-response map $\hat{\mathbf{Q}}(\mathbf{Q})$, instrumental to prove Theorems 9 and 11.

Proposition 19. *The map $\mathcal{Q}\ni\mathbf{Q}\mapsto\hat{\mathbf{Q}}(\mathbf{Q})\in\mathcal{Q}$ defined in (28) has the following properties:*

(a): *$\hat{\mathbf{Q}}(\mathbf{Q})$ is Lipschitz continuous on $\mathcal{Q}$, i.e., there exists a positive constant $L_{\hat{\mathbf{Q}}}$ such that for any $\mathbf{Q}^1,\mathbf{Q}^2\in\mathcal{Q}$, we have $\left\|\hat{\mathbf{Q}}(\mathbf{Q}^1)-\hat{\mathbf{Q}}(\mathbf{Q}^2)\right\|\le L_{\hat{\mathbf{Q}}}\left\|\mathbf{Q}^1-\mathbf{Q}^2\right\|$;*

(b): *For any $\mathbf{Q}\in\mathcal{Q}$, $\hat{\mathbf{Q}}(\mathbf{Q})-\mathbf{Q}$ is an ascent direction of $U(\mathbf{Q})$ at $\mathbf{Q}$ such that*

$$\left\langle\hat{\mathbf{Q}}(\mathbf{Q})-\mathbf{Q},\nabla_{\mathbf{Q}^*}U(\mathbf{Q})\right\rangle\ge c_\tau\left\|\hat{\mathbf{Q}}(\mathbf{Q})-\mathbf{Q}\right\|^2,$$

where $c_\tau>0$ is defined in (27);

(c): *The set of fixed points of the map $\hat{\mathbf{Q}}(\mathbf{Q})$ coincides with the set of stationary points of (8); therefore the map $\hat{\mathbf{Q}}(\mathbf{Q})$ has a fixed point;*

(d): *$\hat{\mathbf{Q}}(\mathbf{Q})-\mathbf{Q}$ is bounded on $\mathcal{Q}$: there exists a finite positive constant B such that $\left\|\hat{\mathbf{Q}}(\mathbf{Q})-\mathbf{Q}\right\|\le B$, for any $\mathbf{Q}\in\mathcal{Q}$.*

Proof: We only prove (a) and (b) due to page limit.

(a): In view of the compactness of the joint feasible set $\mathcal{Q}$ defined in (6), the function $U(\mathbf{Q})$ has a bounded Hessian and $\nabla_{\mathbf{Q}^*}U(\mathbf{Q})$ is thus Lipschitz continuous on $\mathcal{Q}$. It is also easy to verify that $\nabla_{\mathbf{Q}^*}\widetilde{U}(\mathbf{Q};\bullet)$ is uniformly Lipschitz continuous on $\mathcal{Q}$; we denote by $\tilde{L}_\nabla$ its Lipschitz constant.

By definition, $\hat{\mathbf{Q}}(\mathbf{Q}^1)$ satisfies the minimum principle of (28) [21, Lemma 24]:

$$\left\langle\mathbf{Q}-\hat{\mathbf{Q}}(\mathbf{Q}^1),\nabla_{\mathbf{Q}^*}\widetilde{U}(\hat{\mathbf{Q}}(\mathbf{Q}^1);\mathbf{Q}^1)\right\rangle\le 0,\quad\forall\,\mathbf{Q}\in\mathcal{Q};\tag{49a}$$

likewise, so does $\hat{\mathbf{Q}}(\mathbf{Q}^2)$:

$$\left\langle\mathbf{Q}-\hat{\mathbf{Q}}(\mathbf{Q}^2),\nabla_{\mathbf{Q}^*}\widetilde{U}(\hat{\mathbf{Q}}(\mathbf{Q}^2);\mathbf{Q}^2)\right\rangle\le 0,\quad\forall\,\mathbf{Q}\in\mathcal{Q}.\tag{49b}$$

Setting $\mathbf{Q}=\hat{\mathbf{Q}}(\mathbf{Q}^2)$ in (49a) and $\mathbf{Q}=\hat{\mathbf{Q}}(\mathbf{Q}^1)$ in (49b), and summing the resulting inequalities we obtain: denoting for short $\hat{\mathbf{Q}}(\mathbf{Q}^1)$ by $\hat{\mathbf{Q}}^1$ and $\hat{\mathbf{Q}}(\mathbf{Q}^2)$ by $\hat{\mathbf{Q}}^2$,

$$\begin{aligned}0&\ge\left\langle\hat{\mathbf{Q}}^1-\hat{\mathbf{Q}}^2,\nabla_{\mathbf{Q}^*}\widetilde{U}(\hat{\mathbf{Q}}^2;\mathbf{Q}^2)-\nabla_{\mathbf{Q}^*}\widetilde{U}(\hat{\mathbf{Q}}^1;\mathbf{Q}^1)\right\rangle\\&=\left\langle\hat{\mathbf{Q}}^1-\hat{\mathbf{Q}}^2,\nabla_{\mathbf{Q}^*}\widetilde{U}(\hat{\mathbf{Q}}^2;\mathbf{Q}^2)-\nabla_{\mathbf{Q}^*}\widetilde{U}(\hat{\mathbf{Q}}^2;\mathbf{Q}^1)\right\rangle\\&\quad+\left\langle\hat{\mathbf{Q}}^1-\hat{\mathbf{Q}}^2,\nabla_{\mathbf{Q}^*}\widetilde{U}(\hat{\mathbf{Q}}^2;\mathbf{Q}^1)-\nabla_{\mathbf{Q}^*}\widetilde{U}(\hat{\mathbf{Q}}^1;\mathbf{Q}^1)\right\rangle.\end{aligned}\tag{50}$$

We first apply the Cauchy-Schwarz inequality on the first term in (50):

$$\begin{aligned}&\left\langle\hat{\mathbf{Q}}^1-\hat{\mathbf{Q}}^2,\nabla_{\mathbf{Q}^*}\widetilde{U}(\hat{\mathbf{Q}}^2;\mathbf{Q}^2)-\nabla_{\mathbf{Q}^*}\widetilde{U}(\hat{\mathbf{Q}}^2;\mathbf{Q}^1)\right\rangle\\&\ge-\left\|\hat{\mathbf{Q}}^1-\hat{\mathbf{Q}}^2\right\|\left\|\nabla_{\mathbf{Q}^*}\widetilde{U}(\hat{\mathbf{Q}}^2;\mathbf{Q}^2)-\nabla_{\mathbf{Q}^*}\widetilde{U}(\hat{\mathbf{Q}}^2;\mathbf{Q}^1)\right\|\\&\ge-\tilde{L}_\nabla\left\|\hat{\mathbf{Q}}^1-\hat{\mathbf{Q}}^2\right\|\left\|\mathbf{Q}^1-\mathbf{Q}^2\right\|,\end{aligned}\tag{51}$$

where the last inequality comes from the Lipschitz property of $\nabla_{\mathbf{Q}^*}\widetilde{U}(\mathbf{Q};\bullet)$ (with constant $\tilde{L}_\nabla$). Furthermore, we infer from strong concavity of $\widetilde{U}(\bullet;\mathbf{Q})$ (cf. (27)) the following inequality for the second term in (50):

$$\begin{aligned}&\left\langle\hat{\mathbf{Q}}^1-\hat{\mathbf{Q}}^2,\nabla_{\mathbf{Q}^*}\widetilde{U}(\hat{\mathbf{Q}}^1;\mathbf{Q}^1)-\nabla_{\mathbf{Q}^*}\widetilde{U}(\hat{\mathbf{Q}}^2;\mathbf{Q}^1)\right\rangle\\&\le-c_\tau\left\|\hat{\mathbf{Q}}^1-\hat{\mathbf{Q}}^2\right\|^2\end{aligned}\tag{52}$$

Combining (51) and (52), we get the desired Lipschitz property of $\hat{\mathbf{Q}}(\mathbf{Q})$: for any $\mathbf{Q}^1$ and $\mathbf{Q}^2\in\mathcal{Q}$,

$$\left\|\hat{\mathbf{Q}}^1-\hat{\mathbf{Q}}^2\right\|\le c_\tau^{-1}\tilde{L}_\nabla\left\|\mathbf{Q}^1-\mathbf{Q}^2\right\|.$$

(b): Since $\hat{\mathbf{Q}}(\mathbf{Q}^\nu)$ is the (unique) optimal solution of the convex problem (28), it follows from the minimum principle [21, Lemma 24] that (recall $\hat{\mathbf{Q}}^\nu$ stands for $\hat{\mathbf{Q}}(\mathbf{Q}^\nu)$)

$$\left\langle\mathbf{Q}^\nu-\hat{\mathbf{Q}}^\nu,\nabla_{\mathbf{Q}^*}\widetilde{U}(\hat{\mathbf{Q}}^\nu;\mathbf{Q}^\nu)\right\rangle\le 0,$$

which further indicates that

$$\begin{aligned}0&\ge\left\langle\mathbf{Q}^\nu-\hat{\mathbf{Q}}^\nu,\nabla_{\mathbf{Q}^*}\widetilde{U}(\hat{\mathbf{Q}}^\nu;\mathbf{Q}^\nu)\right\rangle\\&=\left\langle\mathbf{Q}^\nu-\hat{\mathbf{Q}}^\nu,\nabla_{\mathbf{Q}^*}\widetilde{U}(\mathbf{Q}^\nu;\mathbf{Q}^\nu)\right\rangle\\&\quad+\left\langle\mathbf{Q}^\nu-\hat{\mathbf{Q}}^\nu,\nabla_{\mathbf{Q}^*}\widetilde{U}(\hat{\mathbf{Q}}^\nu;\mathbf{Q}^\nu)-\nabla_{\mathbf{Q}^*}\widetilde{U}(\mathbf{Q}^\nu;\mathbf{Q}^\nu))\right\rangle\\&\ge\left\langle\mathbf{Q}^\nu-\hat{\mathbf{Q}}^\nu,\nabla_{\mathbf{Q}^*}U(\mathbf{Q}^\nu)\right\rangle+c_\tau\left\|\hat{\mathbf{Q}}^\nu-\mathbf{Q}^\nu\right\|^2.\end{aligned}$$

Rearranging the terms, we get the inequality in **(b)**. ∎

Proof of Theorem 9 and Theorem 11: The proof is based on Proposition 19 and follows the same line of [24, Th. 2 and 4]; we omit the details because of space limit.

REFERENCES

[1] J. Mitola and G. Maguire, "Cognitive radio: making software radios more personal," *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13–18, 1999.

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

[3] A. Goldsmith, S. 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, May 2009.

[4] L. Zhang, Y.-C. Liang, and Y. Xin, "Joint Beamforming and Power Allocation for Multiple Access Channels in Cognitive Radio Networks," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 38–51, Jan. 2008.

[5] S.-J. Kim and G. B. Giannakis, "Optimal Resource Allocation for MIMO Ad Hoc Cognitive Radio Networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 3117–3131, May 2011.

[6] K. Phan, S. Vorobyov, N. Sidiropoulos, and C. Tellambura, "Spectrum Sharing in Wireless Networks via QoS-Aware Secondary Multicast Beamforming," *IEEE Trans. Signal Process.*, vol. 57, no. 6, pp. 2323–2335, Jun. 2009.

[7] Y. Zhang, E. Dall'Anese, and G. B. Giannakis, "Distributed Optimal Beamformers for Cognitive Radios Robust to Channel Uncertainties," *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6495–6508, 2012.

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

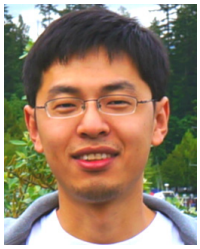
[9] G. Scutari, D. P. Palomar, and S. Barbarossa, "Cognitive MIMO radio," *IEEE Signal Process. Mag.*, vol. 25, no. 6, pp. 46–59, Nov. 2008.

[10] G. Scutari and D. P. Palomar, "MIMO Cognitive Radio: A Game Theoretical Approach," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 761–780, Feb. 2010.

[11] 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, Jun. 2010.

[12] J. Wang, G. Scutari, and D. P. Palomar, "Robust MIMO Cognitive Radio Via Game Theory," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1183–1201, Mar. 2011.

[13] G. Scutari, D. Palomar, F. Facchinei, and J.-S. Pang, "Monotone Games for Cognitive Radio Systems," in *Distributed Decision Making and Control*, R. Johansson and A. Rantzer, Eds. Springer, 2012.
[14] S. Hu, Y.-D. Yao, and Z. Yang, "Cognitive medium access control protocols for secondary users sharing a common channel with time division multiple access primary users," *Wireless Communications and Mobile Computing*, 2012.
[15] G. Scutari, D. P. Palomar, and S. Barbarossa, "The MIMO Iterative Waterfilling Algorithm," *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 1917–1935, May 2009.
[16] R. Zhang, Y.-C. Liang, and S. Cui, "Dynamic resource allocation in cognitive radio networks," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 102–114, 2010.
[17] Y. Eldar and N. Merhav, "A Competitive Minimax Approach to Robust Estimation of Random Parameters," *IEEE Trans. Signal Process.*, vol. 52, no. 7, pp. 1931–1946, Jul. 2004.
[18] J. Wang and D. P. Palomar, "Worst-Case Robust MIMO Transmission With Imperfect Channel Knowledge," *IEEE Trans. Signal Process.*, vol. 57, no. 8, pp. 3086–3100, Aug. 2009.
[19] G. Zheng, K.-K. Wong, and B. Ottersten, "Robust Cognitive Beamforming With Bounded Channel Uncertainties," *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4871–4881, Dec. 2009.
[20] L. Zhang, Y.-C. Liang, Y. Xin, and H. Poor, "Robust cognitive beamforming with partial channel state information," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 4143–4153, Aug. 2009.
[21] G. Scutari, F. Facchinei, J.-S. Pang, and D. P. Palomar, "Real and Complex Monotone Communication Games," Nov. 2012, submitted to *IEEE Trans. Inf. Theory*. [Online]. Available: http://arxiv.org/abs/1212.6235
[22] S. Low and D. Lapsely, "Optimization flow control. I. Basic algorithm and convergence," *IEEE/ACM Trans. Netw.*, vol. 7, no. 6, pp. 861–874, 1999.
[23] 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.
[24] G. Scutari, F. Facchinei, P. Song, D. P. Palomar, and J.-S. Pang, "Decomposition by Partial Linearization: Parallel Optimization of Multi-Agent Systems," Feb. 2013, submitted to *IEEE Trans. Signal Process.*. [Online]. Available: http://arxiv.org/abs/1302.0756
[25] D. H. Johnson and D. E. Dudgeon, *Array signal processing: concepts and techniques*. Englewood Cliffs, N.J.: Prentice Hall, 1993.
[26] S. Vorobyov, A. Gershman, and Z.-Q. Luo, "Robust adaptive beamforming using worst-case performance optimization: a solution to the signal mismatch problem," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 313–324, Feb. 2003.
[27] A. Gershman, N. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten, "Convex optimization-based beamforming," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 62–75, 2010.
[28] W. Yu, G. Ginis, and J. Cioffi, "Distributed multiuser power control for digital subscriber lines," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 5, pp. 1105–1115, Jun. 2002.
[29] R. Etkin, A. Parekh, and D. N. C. Tse, "Spectrum sharing for unlicensed bands," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 3, pp. 517–528, Apr. 2007.
[30] M. Chiang, C. W. Tan, D. P. Palomar, D. Oneill, and D. Julian, "Power Control By Geometric Programming," *IEEE Transactions on Wireless Communications*, vol. 6, no. 7, pp. 2640–2651, Jul. 2007.
[31] F. Facchinei and J.-S. Pang, *Finite-dimensional variational inequalities and complementarity problems*. Springer-Verlag, 2003.
[32] A. Hjørungnes, *Complex-valued matrix derivatives with applications in signal processing and communications*. Cambridge University Press.
[33] R. Cottle, J.-S. Pang, and R. E. Stone, *The linear complementarity problem*. Academic Press, 1992.
[34] R. T. Rockafellar, *Convex analysis*. Princeton University Press, 1997.
[35] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and distributed computation: Numerical methods*. Prentice Hall, 1989.
[36] 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.
[37] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proc. IEEE*, vol. 95, no. 1, pp. 215–223, Jan. 2007.
[38] G. Scutari, S. Barbarossa, and L. Pescosolido, "Distributed decision through self-synchronizing sensor networks in the presence of propagation delays and asymmetric channels," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1667–1684, April 2008.
[39] S. Sundaram and C. N. Hadjicostis, "Distributed function calculation and consensus using linear iterative strategies," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 4, pp. 650–660, May 2008.
[40] O. L. Mangasarian and L. McLinden, "Simple bounds for solutions of monotone complementarity problems and convex programs," *Mathematical Programming*, vol. 32, no. 1, pp. 32–40, May 1985.
[41] J. F. Sturm, "Using SeDuMi 1.02: A Matlab toolbox for optimization over symmetric cones," *Optimization Methods and Software*, vol. 11, no. 1-4, pp. 625–653, Jan. 1999.
[42] D. P. Bertsekas, *Nonlinear programming*. Athena Scientific, 1999.
[43] D. P. Bertsekas, A. Nedic, and A. E. Ozdaglar, *Convex Analysis and Optimization*. Athena Scientific, 2003.
[44] Y. Nesterov, *Introductory Lectures on Convex Optimization: A Basic Course*. Springer, 2004.
[45] J. F. Bonnans and A. Shapiro, *Perturbation analysis of optimization problems*. Springer, 2000.
[46] B. Polyak, *Introduction to optimization*. Optimization Software, 1987.
[47] D. P. Palomar, "Convex primal decomposition for multicarrier linear MIMO transceivers," *IEEE Trans. Signal Process.*, vol. 53, no. 12, pp. 4661–4674, Dec. 2005.



Yang Yang (S'09) received the B.S. degree in Information Engineering from Southeast University, Nanjing, China, in 2009. Since then, he has been working towards the Ph.D. degree at Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology. He held a visiting research appointment at the Institut für Nachrichtentechnik, Technische Universität Darmstadt, Darmstadt, Germany, from Jul. 2011 to Aug. 2012 and Feb. 2013 to May 2013, respectively.

His research interests are in distributed solution methods in convex optimization, nonlinear programming, and game theory, with applications in communication networks, signal processing, and financial engineering.

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 in the Department of Electrical Engineering at State University of New York (SUNY) at Buffalo, Buffalo, NY. He had previously held several research appointments, namely, at the University of California at Berkeley, Berkeley, CA; Hong Kong University of Science and Technology, Hong Kong; University of Rome, "La Sapienza"; University of Illinois at Urbana-Champaign, Urbana, IL. He has participated in many research projects funded by the European Commission and the Hong Kong Government.

His primary research interests include applications of nonlinear optimization theory, game theory, and variational inequality theory to signal processing, communications and networking; sensor networks; and distributed decisions. Dr. Scutari is an Associate Editor of IEEE TRANSACTIONS ON SIGNAL PROCESSING, and he has served as an Associated Editor of the IEEE SIGNAL PROCESSING LETTERS. He serves on the IEEE Signal Processing Society Technical Committee on Signal Processing for Communications (SPCOM). He received the 2006 Best Student Paper Award at the International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2006, and the 2013 NSF Faculty Early Career Development (CAREER) award.

Peiran Song (S'10) received the B.S. degree in information and computing science and the M.S. degree in computational mathematics form Beijing Jiaotong University, Beijing, China, in 2008 and 2010, respectively. She is currently a Ph.D. student in the Department of Electrical Engineering at the State University of New York at Buffalo, NY.

Her research interests include applications of convex optimization theory, cooperative communications and distributed designs.

Daniel P. Palomar (S'99-M'03-SM'08-F'12) received the Electrical Engineering and Ph.D. degrees from the Technical University of Catalonia (UPC), Barcelona, Spain, in 1998 and 2003, respectively.

He is an Associate Professor in the Department of Electronic and Computer Engineering at the Hong Kong University of Science and Technology (HKUST), Hong Kong, which he joined in 2006. Since 2013 he is a Fellow of the Institute for Advance Study (IAS) at HKUST. He had previously held several research appointments, namely, at King's College London (KCL), London, UK; Technical University of Catalonia (UPC), Barcelona; Stanford University, Stanford, CA; Telecommunications Technological Center of Catalonia (CTTC), Barcelona; Royal Institute of Technology (KTH), Stockholm, Sweden; University of Rome "La Sapienza", Rome, Italy; and Princeton University, Princeton, NJ. His current research interests include applications of convex optimization theory, game theory, and variational inequality theory to financial systems and communication systems.

Dr. Palomar is an IEEE Fellow, a recipient of a 2004/06 Fulbright Research Fellowship, the 2004 Young Author Best Paper Award by the IEEE Signal Processing Society, the 2002/03 best Ph.D. prize in Information Technologies and Communications by the Technical University of Catalonia (UPC), the 2002/03 Rosina Ribalta first prize for the Best Doctoral Thesis in Information Technologies and Communications by the Epson Foundation, and the 2004 prize for the best Doctoral Thesis in Advanced Mobile Communications by the Vodafone Foundation and COIT.

He serves as an Associate Editor of IEEE Transactions on Information Theory, and has been an Associate Editor of IEEE Transactions on Signal Processing, a Guest Editor of the IEEE Signal Processing Magazine 2010 Special Issue on "Convex Optimization for Signal Processing," the IEEE Journal on Selected Areas in Communications 2008 Special Issue on "Game Theory in Communication Systems," and the IEEE Journal on Selected Areas in Communications 2007 Special Issue on "Optimization of MIMO Transceivers for Realistic Communication Networks."