Resource Allocation for OFDMA Cognitive Radios Under Channel Uncertainty

Seung-Jun Kim, Nasim Yahya Soltani, and Georgios B. Giannakis

Abstract—A resource allocation problem for a cognitive radio base station (CR-BS) communicating with multiple CR mobile stations (MSs) is considered in the downlink, which relies on orthogonal frequency-division multiple access (OFDMA). To protect the incumbent primary user (PU) system operating over the same frequency band, the interference inflicted to the PU receiver must be regulated. Since the channel gain estimates from the CR-BS to the PU receiver are typically uncertain in practice, a robust interference power constraint is advocated. Although the latter translates to a second-order cone constraint, the overall optimization problem for joint transmit-power and subcarrier allocation is non-convex, and coupled over all subcarriers in general. To circumvent the resulting computational hurdle, a tight polyhedral approximation of the second-order cone is employed. Practical finite-alphabet constellations are adopted and total weighted achievable rate is maximized. It is shown that a near-optimal solution can be obtained based on the Lagrangian dual.

Index Terms—Cognitive radios, resource allocation, OFDMA, channel uncertainty, optimization.

I. INTRODUCTION

DUE to increasing demand for high-speed wireless data access, efficient utilization of the limited frequency spectrum has become crucial. The cognitive radio (CR) technology puts forth a promising proposal to mitigate the scarcity of spectral resources by allowing unlicensed transmitters to use underutilized licensed bands in an opportunistic fashion. To avoid interfering licensed primary user (PU) systems and capture spectrum opportunities efficiently, CRs must adapt to the spectrum occupancy dynamically and intelligently.

There are a number of coexistence models for the CR concept [1]. In the so-called spectrum overlay scenario (also termed *interweave* in [2]), CRs detect frequency bands unoccupied by PUs through spectrum sensing, and exploit those "spectrum holes" using dynamic resource allocation. Under a spectrum underlay model, CRs may transmit concurrently with PUs sharing the same bands, provided that the interference experienced by the PUs is controlled to a tolerable level, say, comparable to that of the ambient noise and interference.

In order to utilize a broad swath of spectrum flexibly, CR systems often employ multi-channel architectures such as

orthogonal frequency division multiple access (OFDMA). Key resource allocation (RA) issues in OFDMA radios include optimal assignment of subcarriers to individual users, and optimal power loading across the subcarriers. Extensive research has been carried out on this important topic, with recent applications to CR scenarios. A heuristic algorithm based on the multi-dimensional knapsack problem was proposed for OFDM CRs in [3]. A weighted sum-rate maximization problem has been considered in [4].

Channel estimation in CR systems is challenging, since the CRs gain access to the medium only intermittently, and the incumbent PU system often does not explicitly support channel estimation for CR systems. In addition, CRs might not have prior knowledge of PU signal characteristics, and thus are forced to resort to less efficient channel estimation techniques. Nevertheless, PU transmissions must be strictly protected from the interference due to CRs. Therefore, it is widely recognized that RA for CRs must account for channel uncertainty.

There are a couple of approaches that are typically considered for optimization with uncertain parameters [5]. When the uncertain parameters are regarded as random, statistical knowledge such as their mean and covariance, or their distribution is assumed, which leads to chance-constrained formulations. A deterministic alternative is to adopt a robust optimization framework, where a bounded uncertainty region is postulated. In certain cases, the two approaches can be shown to be intimately related [6].

Transmit-beamformers were designed for multi-antenna CRs to maximize the signal-to-noise-plus-interference-ratio (SINR) subject to probabilistic interference constraints in [7]. A worst-case formulation was employed in [8] for beamformer design, where an ellipsoidal channel uncertainty region was assumed. Probabilistic interference constraints were considered for a CR power control problem in [9], where uncertainty in composite fading channels comprising shadowing and Nakagami fading was accounted for. RA problems for generic OFDMA systems with channel uncertainty have also been investigated extensively [10], [11].

On the other hand, not many works have addressed the RA problems for OFDMA-based CRs under channel uncertainty. A heuristic RA algorithm for OFDMA CR systems has been reported in [12], where the sum rate was maximized under power, interference, and average bit error rate constraints. However, the algorithm accounted for uncertainties only in the channels between the CR-BS and the CR-MSs, but not between the CR-BS and the PU. CR-PU channel uncertainty was considered in an OFDM setup by [13], but the subcarrier

assignment issue for OFDMA was not addressed. Both channel estimation and spectrum sensing errors were taken into account for OFDMA CRs in [14] under a spectrum overlay model, in which the bands occupied by the PUs are completely avoided.

The present paper considers an OFDMA CR network operating in a spectrum underlay set-up. A weighted sum rate maximization problem is formulated for a CR base station (BS) that transmits to a set of CR mobile stations (MSs), while respecting a strict interference constraint to protect PUs when the channel estimate between the CR-BS and the PU receiver contains uncertainty. It turns out that the interference constraint can be converted to a second-order cone constraint, which is convex. However, the overall optimization problem is still non-convex in general due to the combinatorial search necessary for subcarrier assignment. It is well known that when an OFDMA RA problem has a separable structure, the duality gap vanishes as the number of subcarriers grows large [15], [16]. To induce such a separable structure, a tight linear approximation for the interference constraint is introduced at a modest (polynomial) increase in the problem dimension. The approximated problem can then be readily solved using the dual decomposition method, which leads to a near-optimal, computationally efficient algorithm.

Different from our conference precursor [17], in which the user rates were modeled by logarithmic functions of transmit-powers, justified under ideal Gaussian coding, practical finite-alphabet constellations are adopted in this work. The optimization objective is a weighted sum of mutual information for individual CR-MSs. As the rate does not grow without bound even when the transmit-power is increased, a mercury/water-filling step is introduced rather than the conventional water-filling step [18]. The dual problem is solved using the relationship between the minimum mean-square error (MMSE) and the derivative of mutual information [19]. Moreover, a chance-constrained formulation was pursued in [17], while a robust optimization framework is adopted here, with the close relationship between the two delineated. A suboptimal benchmark algorithm is also considered, which aims at a locally optimal solution. A time complexity analysis for the proposed algorithm is provided as well.

In our companion paper [20], an OFDMA *uplink* problem was tackled, which entails a significantly different set of challenges than the *downlink* case treated in this paper. While conservative approximations of probabilistic interference constraints with strictly positive gaps were employed in [20], the present work pursues an arbitrarily tight approximation of robust interference constraints.

Although important components of the proposed algorithm, such as the Lagrange relaxation technique, robust optimization, and rate maximization for practical constellations, have received much attention in the literature, it must be emphasized that bringing these together for OFDMA CR RA under channel uncertainty incurs significant and novel challenges. On a high level, our contribution can be put in the following perspective. It is widely accepted that the Lagrange relaxation technique is a powerful tool for OFDMA RA. It is also generally recognized that robust optimization is an important framework to address RA problems with uncertain data. The

critical question is whether it is possible to marry these two to tackle the challenging OFDMA RA problem for CRs. The present work offers a positive answer to this intriguing query, with associated trade-offs carefully examined. Moreover, the resulting contribution may have impact beyond the particular RA task treated here, as it provides a means to efficiently tackle a class of large-scale robust mixed-integer problems involving second-order cone constraints.

The rest of the paper is organized as follows. Sec. II formulates the problem, and Sec. III briefs a polyhedral approximation technique for second-order cones. The proposed RA algorithm is derived in Sec. IV, along with complexity analysis as well as a sub-optimal benchmark algorithm. Results of numerical tests are presented in Sec. V, followed by conclusions in Sec. VI.

II. PROBLEM STATEMENT

Consider a CR-BS allocating resources to K CR-MSs (users) employing OFDMA using N subcarriers. The channel gain $h_k^{(n)} \geq 0$ between the CR-BS and the k-th CR-MS for each $k \in \mathcal{K} := \{1, 2, \ldots, K\}$, and each subcarrier $n \in \{1, 2, \ldots, N\}$ is assumed to have been acquired accurately via conventional channel estimation techniques. Suppose also that during the sensing phase, the CR-BS has detected the presence of an active PU. Let $p^{(n)}$ denote the transmit-power loaded on subcarrier n, and $P_{\max}^{(n)}$ the maximum transmit-power allowed on subcarrier n. Vectors $\mathbf{p}$ and $\mathbf{P}_{\max}$ collect $\{p^{(n)}\}$ and $\{P_{\max}^{(n)}\}$, respectively. Also, let $k(n) \in \mathcal{K}$ represent the index of the user served on subcarrier n, and define $\mathbf{k} := [k(1), \ldots, k(N)]^T$.

Rather than employing ideal Gaussian signaling, practical discrete constellations are adopted. Thus, the received signal for user k on subcarrier n is modeled as

$$y_k^{(n)} = \sqrt{h_k^{(n)} p^{(n)}} x_k^{(n)} + v_k^{(n)} \tag{1}$$

where $x_k^{(n)}$ is the unit-average-power input signal on subcarrier n drawn from a practical constellation of user k, where the inputs for different users and subcarriers are assumed to be independent. Random variables $\{v_k^{(n)}\}$ represent independent and identically distributed (i.i.d.) complex Gaussian noise with mean 0 and variance 1. The mutual information for user k on subcarrier n is denoted as $\mathcal{I}_k^{(n)}(\gamma) := I\left(x_k^{(n)}; \sqrt{\gamma} x_k^{(n)} + v_k^{(n)}\right)$, where γ represents the received signal power-to-noise power ratio (SNR). Let w_k denote the positive weight of user $k \in \mathcal{K}$. Then, a relevant RA objective is to maximize a weighted sum of mutual information $\sum_{n=1}^N w_{k(n)} \mathcal{I}_{k(n)}^{(n)}(h_k^{(n)} p^{(n)})$ of the CR system, while adhering to a transmit-power constraint, and a PU interference constraint.

To model the PU interference constraint, let $g^{(n)} \geq 0$ denote the channel gain from the CR-BS to the PU receiver on subcarrier n. Unlike the channels between CR transceivers, it is difficult to estimate $\{g^{(n)}\}$ precisely, due to the lack of cooperation from the PU system. Thus, $\{g^{(n)}\}$ are assumed uncertain. Typically, uncertainty in the parameters of an optimization problem is captured either deterministically or statistically [6]. Under the deterministic approach, bounded

uncertainty sets of the parameters are assumed, and the worst-case solution immunized to all elements in the uncertainty set is sought. When the parameters are viewed as random, chance constraints are constructed from the distributions of the parameters. In this work, a deterministic robust optimization approach is adopted, which can be also shown to be intimately related to the chance-constrained approach under special circumstances.

Specifically, it is assumed that the uncertainty in vector $\mathbf{g} := [g^{(1)}\ g^{(2)}\ \ldots\ g^{(N)}]^T$ can be captured by an ellipsoidal uncertainty region given by

$$\mathcal{G} := \left\{ \bar{\mathbf{g}} + \Delta\mathbf{g} : \Delta\mathbf{g}^T \mathbf{C}_g^{-1} \Delta\mathbf{g} \leq \Omega^2 \right\} \quad (2)$$

where $\bar{\mathbf{g}}$ is the nominal value of $\mathbf{g}$, $\Delta\mathbf{g}$ the deviation from the nominal value, $\mathbf{C}_g$ a symmetric positive definite matrix, and $\Omega \geq 0$ a given constant. Then, a robust interference constraint guarantees tolerable interference to the PU system for all possible $\mathbf{g} \in \mathcal{G}$; that is,

$$\sum_{n=1}^{N} g^{(n)} p^{(n)} \leq I_{\max} \text{ for all } \mathbf{g} \in \mathcal{G}. \quad (3)$$

One can easily confirm that (3) is equivalent to[1]

$$\mathbf{p} \in \mathcal{C} := \left\{ \mathbf{p} \,\middle|\, I_{\max} - \bar{\mathbf{g}}^T \mathbf{p} \geq \Omega \sqrt{\mathbf{p}^T \mathbf{C}_g \mathbf{p}} \right\} \quad (4)$$

which is a second-order cone constraint [6].

Alternatively, one can arrive at (4) by assuming that $\mathbf{g}$ is jointly Gaussian with mean $\bar{\mathbf{g}}$ and covariance $\mathbf{C}_g$, and considering the chance constraint given by

$$\Pr\left\{ \sum_{n=1}^{N} g^{(n)} p^{(n)} > I_{\max} \right\} \leq \epsilon \quad (5)$$

which enforces the interference afflicted to the PU not to exceed $I_{\max}$ with probability at least $1 - \epsilon$. Constraint (5) is equivalent to (4) with $\Omega = Q^{-1}(\epsilon)$, where $Q(\cdot)$ is the standard Gaussian tail function. Even when $\mathbf{g}$ is not jointly Gaussian (for example, $g^{(n)}$ follows an exponential distribution under Rayleigh fading), one may appeal to the fact that N is often quite large in modern OFDMA systems, and invoke the central limit theorem to reach the same conclusion [13].

Yet another connection to the chance-constrained problem can be made by assuming that $\{g^{(n)}\}$ are random and independent of one another, and have finite support such that $g^{(n)} - \bar{g}^{(n)} \in [-\hat{g}^{(n)}, \hat{g}^{(n)}]$ for all n. Independence is a reasonable assumption when the number of subcarriers is on par with the number of multiple channel paths. Then, a conservative approximation of (5) is again given by (4) with $\mathbf{C}_g$ being a diagonal matrix whose n-th diagonal entry equals to $(\hat{g}^{(n)})^2$, and $\Omega = \sqrt{2 \log\left(\frac{1}{\epsilon}\right)}$ [5].

The upshot is that the proposed framework can embrace a rather broad range of modeling perspectives, based on either statistical or deterministic approaches. Thus, our formulation

and solution methodology can readily accommodate practical system requirements and deployment scenarios. In any case, the following optimization problem is of interest:

$$\text{(P1)} \quad \max_{0 \preceq \mathbf{p} \preceq \mathbf{P}_{\max}, \mathbf{k} \in \mathcal{K}^N} \sum_{n=1}^{N} w_{k(n)} \mathcal{I}_{k(n)}^{(n)}(h_k^{(n)} p^{(n)}) \quad (6)$$

$$\text{subject to} \quad \sum_{n=1}^{N} p^{(n)} \leq P_{\max} \quad (7)$$

$$\bar{\mathbf{g}}^T \mathbf{p} - I_{\max} + \Omega\sqrt{\mathbf{p}^T \mathbf{C}_g \mathbf{p}} \leq 0. \quad (8)$$

Note that (P1) is convex provided $K = 1$, i.e., for a single-user system. In the general case of multiple users with $K > 1$, (P1) is non-convex due to the combinatorial assignment of users on each subcarrier.

When an OFDMA RA problem has a separable structure, in which the Lagrangian dual can be decomposed into per-subcarrier subproblems, it can be shown that the duality gap vanishes as the number of the subcarriers increases [21], [15]. This implies that the problem can be solved near-optimally in polynomial time for sufficiently large N. However, due to the coupling of $p^{(n)}$ belonging to different subcarriers n in (8), (P1) does not possess this separable structure. The novel approach here is to enforce such a structure by approximating (8) by a set of linear constraints. To this end, a general result that approximates second-order cone constraints by polyhedral constraints is employed, which is outlined next.

III. POLYHEDRAL APPROXIMATION OF SECOND-ORDER CONES

Motivated by the availability of efficient large-scale linear program (LP) solvers, a second-order cone program was tackled via an approximate LP in [22]. To this end, it was shown that the second-order Lorentz cone

$$\mathcal{L}^N := \left\{ (y_0, \mathbf{y}) \in \mathbb{R} \times \mathbb{R}^N : ||\mathbf{y}|| \leq y_0 \right\} \quad (9)$$

admits a polyhedral approximation of accuracy δ, comprising variables and constraints whose number is polynomial in N and $\log(1/\delta)$; see also [23] for a refinement of this result. Here, we briefly recap the idea behind this approximation.

Direct construction of a polyhedral approximation circumscribing $\mathcal{L}^N$ in the $(N+1)$-dimensional space is bound to have its number of facets growing exponentially with the dimension N, which implies that it needs an exponentially growing number of linear inequalities to define. The key idea is to reduce the number of inequalities by lifting the polyhedron to a higher-dimensional space by introducing additional variables, and considering its projection onto the $(N+1)$-dimensional subset. Since a projection of a higher-dimensional polyhedral set can significantly multiply the number of facets, this approach yields a relaxation that is "efficient" in the sense that it is very tight, yet it is defined using a relatively small number of constraints and extra variables.

The first step is to decompose the $(N+1)$-dimensional Lorentz cone $\mathcal{L}^N$ to a number of 3-dimensional Lorentz cones using the "tower of variables" concept. Suppose for simplicity that $N = 2^d$ for some integer d. Then, by introducing a vector of $N/2$ new variables $\boldsymbol{\rho}^{(1)} := \left[\rho_1^{(1)}, \rho_2^{(1)}, \ldots, \rho_{\frac{N}{2}}^{(1)} \right]^T$, where

---

[1]One can first note that (3) is equivalent to $\bar{\mathbf{g}}^T \mathbf{p} + \max_{\Delta\mathbf{g}:\Delta\mathbf{g}^T \mathbf{C}_g^{-1} \Delta\mathbf{g} \leq \Omega^2} \Delta\mathbf{g}^T \mathbf{p} \leq I_{\max}$. Upon defining $\Delta\tilde{\mathbf{g}} := \mathbf{C}_g^{-\frac{1}{2}} \Delta\mathbf{g}$, and invoking Cauchy-Schwarz's inequality as $\Delta\mathbf{g}^T \mathbf{p} = \Delta\tilde{\mathbf{g}}^T \mathbf{C}_g^{\frac{1}{2}} \mathbf{p} \leq ||\Delta\tilde{\mathbf{g}}||_2 ||\mathbf{C}_g^{\frac{1}{2}} \mathbf{p}||_2 \leq \Omega\sqrt{\mathbf{p}^T \mathbf{C}_g \mathbf{p}}$, one obtains the desired result.

the superscript $^{(\ell)}$ in $\boldsymbol{\rho}^{(\ell)}$ denotes the ℓ-th layer of the "tower," the cone $\mathcal{L}^N$ can be equivalently written as

$$\mathcal{L}^N = \left\{ (y_0, \mathbf{y}) \in \mathbb{R} \times \mathbb{R}^N : \exists \boldsymbol{\rho}^{(1)} \in \mathbb{R}^{\frac{N}{2}}, \sum_{n=1}^{N} \rho_i^{(1)2} \leq y_0^2, \right.$$

$$\left. y_{2i-1}^2 + y_{2i}^2 \leq \rho_i^{(1)2}, \; i = 1, 2, \ldots, \frac{N}{2} \right\} \quad (10)$$

$$= \left\{ (y_0, \mathbf{y}) \in \mathbb{R} \times \mathbb{R}^N : \exists \boldsymbol{\rho}^{(1)} \in \mathbb{R}^{\frac{N}{2}}, (y_0, \boldsymbol{\rho}^{(1)}) \in \mathcal{L}^{\frac{N}{2}}, \right.$$

$$\left. (\rho_i^{(1)}, y_{2i-1}, y_{2i}) \in \mathcal{L}^2, \; i = 1, 2, \ldots, \frac{N}{2} \right\}. \quad (11)$$

Decomposing further the $\left(\frac{N}{2} + 1\right)$-dimensional Lorentz cone in (11) by applying the idea repeatedly, one obtains eventually $(N - 1)$ three-dimensional second-order cone constraints via $(N - 2)$ new variables $\boldsymbol{\rho}^{(\ell)} := \left[\rho_1^{(\ell)}, \ldots, \rho_{\frac{N}{2^\ell}}^{(\ell)}\right]^T$, $\ell = 1, 2, \ldots, d - 1$.

The remaining task is to approximate $\mathcal{L}^2$ using a polynomial number of variables and constraints. Consider a polyhedral δ-relaxation Π_δ^N of $\mathcal{L}^N$ with $\delta > 0$ in the sense that

$$\mathcal{L}^N \subset \Pi_\delta^N \subset \mathcal{L}_\delta^N := \left\{ (y_0, \mathbf{y}) \in \mathbb{R} \times \mathbb{R}^N : ||\mathbf{y}|| \leq (1 + \delta) y_0 \right\}. \quad (12)$$

Then, for an integer q with

$$\delta = \frac{1}{\cos\left(\frac{\pi}{2^q}\right)} - 1 \quad (13)$$

it can be shown that the set of points $(y_0, y_1, y_2) = (\alpha_{q+1}, \alpha_0, \beta_0)$ satisfying the following set of linear constraints is a δ-relaxation of $\mathcal{L}^2$ [22], [23]:

$$\alpha_{i+1} = \alpha_i \cos\left(\frac{\pi}{2^i}\right) + \beta_i \sin\left(\frac{\pi}{2^i}\right), \quad i = 0, 1, \ldots, q \quad (14)$$

$$\beta_{i+1} \geq \left| \beta_i \cos\left(\frac{\pi}{2^i}\right) - \alpha_i \sin\left(\frac{\pi}{2^i}\right) \right|, \quad i = 0, \ldots, q - 1 \quad (15)$$

where $\boldsymbol{\alpha} := [\alpha_1, \alpha_2, \ldots, \alpha_q]^T$ and $\boldsymbol{\beta} := [\beta_1, \ldots, \beta_q]^T$ are extra variables introduced to "lift" the approximation to a higher-dimensional space. Thus, $2q$ extra variables have been introduced to form $(q + 1)$ equality constraints and $2q$ inequality constraints. One can further reduce the number of variables and constraints by eliminating $\boldsymbol{\alpha}$ and β_q using the equalities (14). The resulting set of linear constraints contains $(q-1)$ extra variables and only $2q$ linear inequality constraints.

Overall, using δ_ℓ-relaxations for the cones in the ℓ-th layer, a polyhedral δ-relaxation of $\mathcal{L}^N$ can be obtained as

$$\Pi_\delta^N := \left\{ (y_0, \mathbf{y}) = (\rho_1^{(d)}, \boldsymbol{\rho}^{(0)}) \in \mathbb{R} \times \mathbb{R}^N : \exists \boldsymbol{\rho}^{(1)}, \ldots, \boldsymbol{\rho}^{(d-1)}, \right.$$

$$\left. (\rho_i^{(\ell)}, \rho_{2i-1}^{(\ell-1)}, \rho_{2i}^{(\ell-1)}) \in \Pi_{\delta_\ell}^2, \; i = 1, \ldots, \frac{N}{2^\ell}, \ell = 1, \ldots, d \right\} \quad (16)$$

where $\delta = \prod_{\ell=1}^d (1 + \delta_\ell) - 1$ holds. Thus, the overall approximation introduces $v(N) := (N - 2) + \sum_{\ell=1}^d (q_\ell - 1)\frac{N}{2^\ell}$ extra variables, and $c(N) := \sum_{\ell=1}^d \frac{N \cdot q_\ell}{2^{\ell-1}}$ inequality constraints. Given the overall accuracy requirement δ, $\{\delta_\ell\}$ were optimized in [23], and can be obtained by plugging in $q = q_\ell = \lceil \frac{\ell+1}{2} \rceil - \lfloor \log_4\left(\frac{16}{9}\pi^{-2}\log(1 + \delta)\right) \rfloor$ to (13).

IV. RESOURCE ALLOCATION ALGORITHMS

A. Algorithm Based On Lagrangian Dual

In order to obtain a feasible solution to the original robust RA problem, the set of linear constraints approximating (4) must be tighter than the original constraint (4). Thus, for a given small positive constant δ, consider a tightened constraint set given by

$$\mathcal{C}' := \left\{ \mathbf{p} \left| \frac{I_{\max} - \bar{\mathbf{g}}^T \mathbf{p}}{(1 + \delta)\Omega} \geq \sqrt{\mathbf{p}^T \mathbf{C}_g \mathbf{p}} \right. \right\}. \quad (17)$$

Then, a δ-relaxed lifted polyhedral approximation of the form

$$\mathcal{C}_\delta' := \left\{ \mathbf{p} \left| \exists \mathbf{q} \in \mathbb{R}^{n_q}, \mathbf{A}\mathbf{p} + \mathbf{B}\mathbf{q} \preceq \mathbf{b} \right. \right\} \quad (18)$$

exists, where $\mathbf{A} \in \mathbb{R}^{n_c \times N}$, $\mathbf{B} \in \mathbb{R}^{n_c \times n_q}$, and $\mathbf{b} \in \mathbb{R}^{n_c}$ are obtained from the procedure outlined in Sec. III, and $\mathbf{q} \in \mathbb{R}^{n_q}$ is the vector of additional variables, with $n_c = c(N)$ and $n_q = v(N)$. From (12), it can be seen that

$$\mathcal{C}' \subset \mathcal{C}_\delta' \subset \mathcal{C} \quad (19)$$

holds. Therefore, the following optimization problem is a conservative surrogate for (P1):

$$\text{(P2)} \quad \max_{\mathbf{p}, \mathbf{q}, \mathbf{k} \in \mathcal{K}^N} \sum_{n=1}^N w_{k(n)} \mathcal{I}_{k(n)}^{(n)} (h_{k(n)}^{(n)} p^{(n)}) \quad (20)$$

$$\text{subject to} \sum_{n=1}^N p^{(n)} \leq P_{\max} \quad (21)$$

$$0 \leq p^{(n)} \leq P_{\max}^{(n)}, \quad n = 1, 2, \ldots, N \quad (22)$$

$$\mathbf{A}\mathbf{p} + \mathbf{B}\mathbf{q} \preceq \mathbf{b}. \quad (23)$$

Problem (P2) is again non-convex. However, it can be shown that the duality gap vanishes asymptotically as $N \to \infty$ [16], [24], [25]. A precise statement and its proof can be found in the Appendix. Therefore, (P2) can be solved efficiently using the dual method. Introducing dual variables $\lambda \geq 0$ and $\boldsymbol{\mu} := [\mu_1, \mu_2, \ldots, \mu_{n_c}]^T \succeq \mathbf{0}$, the (partial) Lagrangian is

$$L(\mathbf{p}, \mathbf{q}) = \sum_{n=1}^N w_{k(n)} \mathcal{I}_{k(n)}^{(n)} (h_{k(n)}^{(n)} p^{(n)}) - \lambda \left(\sum_{n=1}^N p^{(n)} - P_{\max} \right)$$

$$- \boldsymbol{\mu}^T (\mathbf{A}\mathbf{p} + \mathbf{B}\mathbf{q} - \mathbf{b}) \quad (24)$$

$$= \sum_{n=1}^N \left\{ w_{k(n)} \mathcal{I}_{k(n)}^{(n)} (h_{k(n)}^{(n)} p^{(n)}) - (\lambda + \boldsymbol{\mu}^T \mathbf{A}(:, n)) p^{(n)} \right\}$$

$$+ \lambda P_{\max} - \boldsymbol{\mu}^T \mathbf{B}\mathbf{q} + \boldsymbol{\mu}^T \mathbf{b} \quad (25)$$

where $\mathbf{A}(:, n)$ denotes the n-th column of matrix $\mathbf{A}$. Upon defining

$$L_n(p^{(n)}, k) := w_k \mathcal{I}_k^{(n)} (h_k^{(n)} p^{(n)}) - (\lambda + \boldsymbol{\mu}^T \mathbf{A}(:, n)) p^{(n)} \quad (26)$$

the dual function is given by

$$D(\lambda, \boldsymbol{\mu}) = \sup_{\mathbf{0} \preceq \mathbf{p} \preceq \mathbf{P}_{\max}, \mathbf{q}, \mathbf{k} \in \mathcal{K}^N} L(\mathbf{p}, \mathbf{q}) \qquad (27)$$

$$= \begin{cases} \sup\limits_{\mathbf{0} \preceq \mathbf{p} \preceq \mathbf{P}_{\max}, \mathbf{k} \in \mathcal{K}^N} \sum\limits_{n=1}^{N} L_n(p^{(n)}, k(n)) + \lambda P_{\max} + \boldsymbol{\mu}^T \mathbf{b}, \\ \hspace{4cm} \text{if } \mathbf{B}^T \boldsymbol{\mu} = \mathbf{0} \\ \infty, \hspace{3cm} \text{otherwise} \end{cases}$$

$$(28)$$

and thus the dual optimization problem boils down to

$$\inf_{\lambda, \boldsymbol{\mu}} D(\lambda, \boldsymbol{\mu}) \qquad (29)$$

$$\text{subject to } \lambda \geq 0, \ \boldsymbol{\mu} \succeq \mathbf{0}, \ \mathbf{B}^T \boldsymbol{\mu} = \mathbf{0}. \qquad (30)$$

It is interesting to note that the auxiliary variables $\mathbf{q}$ introduced for the lifted polyhedral relaxation do not need to be determined in order to obtain the dual function. Moreover, it is immediate that the optimization in (28) can be decomposed into individual tones, thanks to the transformation of the nonlinear cone constraint into multiple linear ones, so that the interference contributions due to individual $p^{(n)}$ can be separately assessed.

Specifically, for each subcarrier $n \in \{1, 2, \ldots, N\}$

$$\max_{0 \leq p^{(n)} \leq P_{\max}^{(n)}, k(n) \in \mathcal{K}} L_n(p^{(n)}, k(n)) \qquad (31)$$

needs to be solved. It is revealing to interpret (26) from an economic perspective, as balancing the payoff due to an achievable rate and the cost incurred for causing interference, where $\boldsymbol{\mu}$ contains the prices associated with different facets of the polyhedral model. If $k(n) = k \in \mathcal{K}$, the optimal power loading can be obtained by taking the derivative of $L_n(p^{(n)}, k)$ w.r.t. $p^{(n)}$, and setting it to zero. To this end, one needs the derivative of the mutual information $\mathcal{I}_k^{(n)}$. Since it is not straightforward to obtain the mutual information of an arbitrary input distribution in closed form and compute its derivative, a fundamental relationship between the derivative of mutual information and the nonlinear minimum mean-square error (MMSE) can be utilized [19].

Specifically, let $\mathsf{MMSE}_k^{(n)}(h_k^{(n)} p^{(n)})$ denote the MMSE for estimating $x_k^{(n)}$ given $y_k^{(n)}$ based on model (1). Then, it is known that

$$\frac{d}{dp^{(n)}} \mathcal{I}_k^{(n)}(h_k^{(n)} p^{(n)}) = h_k^{(n)} \mathsf{MMSE}_k^{(n)}(h_k^{(n)} p^{(n)}). \qquad (32)$$

Using (32), it can be shown that the optimal power loading $p^{*(n)}[k]$ is given by

$$p^{*(n)}[k] = \begin{cases} P_{\max}^{(n)}, \quad \text{if } \ \lambda + \boldsymbol{\mu}^T \mathbf{A}(:, n) < 0 \\ \left[\frac{1}{h_k^{(n)}} \mathsf{MMSE}_k^{(n)-1} \left(\frac{\lambda + \boldsymbol{\mu}^T \mathbf{A}(:, n)}{w_k h_k^{(n)}} \right) \right]_0^{P_{\max}^{(n)}}, \\ \hspace{2cm} \text{if } 0 \leq \frac{\lambda + \boldsymbol{\mu}^T \mathbf{A}(:, n)}{w_k h_k^{(n)}} \leq 1 \\ 0, \quad \text{otherwise}, \end{cases}$$

$$n = 1, 2, \ldots, N \quad (33)$$

TABLE I
ALGORITHM FOR SOLVING (P2).

1: Initialize $\boldsymbol{\Sigma}$ and $\boldsymbol{\nu} := [\lambda, \boldsymbol{\theta}^T]^T$. Set tolerance τ
2: Repeat
3:       If $\lambda < 0$, set $\mathbf{d} = -\mathbf{i}_1$ (first canonical basis)
4:       Or, if $\mathbf{Z}(i, :)\boldsymbol{\theta} < 0$ for some $i \in \{1, 2, \ldots, n_c\}$,
set $\mathbf{d} = [0 \ - \mathbf{Z}(i, :)]^T$
5:       Otherwise:
6:         Find $\mathbf{k}^*$ and $\mathbf{p}^*$ from (34)–(35)
7:         Set $\mathbf{d}$ equal to (36)
8:       If $\sqrt{\mathbf{d}^T \boldsymbol{\Sigma} \mathbf{d}} < \tau$, stop
9:       Perform the ellipsoid update:
10:      $\mathbf{d} \leftarrow \mathbf{d}/\sqrt{\mathbf{d}^T \boldsymbol{\Sigma} \mathbf{d}}$
11:      $\boldsymbol{\nu} \leftarrow \boldsymbol{\nu} - \boldsymbol{\Sigma} \mathbf{d}/(n_\theta + 2)$
12:      $\boldsymbol{\Sigma} \leftarrow \frac{(n_\theta+1)^2}{(n_\theta+1)^2-1} \left(\boldsymbol{\Sigma} - \frac{2}{n_\theta+2} \boldsymbol{\Sigma} \mathbf{d}\mathbf{d}^T \boldsymbol{\Sigma} \right)$

where $\mathsf{MMSE}_k^{(n)-1}(\cdot)$ is the inverse function of $\mathsf{MMSE}_k^{(n)}(\gamma)$, and $[\cdot]_a^b := \min\{\max\{0, a\}, b\}$. Thus, the optimal user allocation $\mathbf{k}^*$ and power loading $\mathbf{p}^*$ are given, respectively, by

$$k^*(n) \in \arg \max_{k \in \mathcal{K}} L_n(p^{*(n)}[k], k), \quad n = 1, 2, \ldots, N \quad (34)$$

$$p^{*(n)} = p^{*(n)}[k^*(n)], \quad n = 1, 2, \ldots, N. \qquad (35)$$

The optimal solution of (29)–(30) can be obtained via iterative optimization methods for non-differentiable objectives, such as the subgradient method or the ellipsoid method. To ensure $\mathbf{B}^T \boldsymbol{\mu} = \mathbf{0}$ as required in (30), $\boldsymbol{\mu}$ is parametrized by $\boldsymbol{\theta} \in \mathbb{R}^{n_\theta}$ as $\boldsymbol{\mu} = \mathbf{Z}\boldsymbol{\theta}$, where the columns of $\mathbf{Z}$ constitute the basis vectors of the null space of $\mathbf{B}^T$. It can be shown that

$$- \begin{bmatrix} \sum_{n=1}^{N} p^{*(n)} - P_{\max} \\ \mathbf{Z}^T (\mathbf{A}\mathbf{p}^* - \mathbf{b}) \end{bmatrix} \qquad (36)$$

is a subgradient of $D(\lambda, \mathbf{Z}\boldsymbol{\theta})$ w.r.t. $[\lambda \ \boldsymbol{\theta}^T]^T$. The overall procedure for solving (P2) using the ellipsoid method is presented in Table I, where $\mathbf{Z}(i, :)$ is the i-th row of $\mathbf{Z}$ for $i = 1, 2, \ldots, n_c$.

It is of interest to examine the complexity of the algorithm in Table I. The overall complexity can be characterized by multiplying the complexity per iteration by the number of iterations needed. The number of iterations needed for the ellipsoid method to converge grows as the square of the number of optimization variables [26]. The number of variables in the proposed algorithm equals $(n_\theta + 1)$, where $n_\theta = n_c - \mathrm{rank}(\mathbf{B}^T)$. Note that $\mathrm{rank}(\mathbf{B}^T) \leq \min(n_c, n_q) = n_q$ since $v(N) < c(N)$. On the other hand, it can be shown that the values of n_q and n_c both grow as $\mathcal{O}(N \log \frac{1}{\delta})$ for $\delta \leq \frac{1}{2}$ [23]. Therefore, the value of n_θ also grows at most as $\mathcal{O}(N \log \frac{1}{\delta})$. It can thus be deduced that the required number of iterations for our algorithm grows as $\mathcal{O}(N^2 (\log \frac{1}{\delta})^2)$.

It turns out that the number of operations per iteration grows as $\mathcal{O}(N^2 (\log \frac{1}{\delta})^2)$ as well, leading to the overall complexity order of $\mathcal{O}(N^4 (\log \frac{1}{\delta})^4)$. The per-iteration complexity can be verified as follows. In line 4 of Table I, $\mathcal{O}(n_c n_\theta)$ operations are needed. In line 6, $\mathcal{O}(n_c N)$ operations are necessary to compute the powers on all subcarriers using (35), where the dominant component is the calculation of $\boldsymbol{\mu}^T \mathbf{A}$. Once the values of $\boldsymbol{\mu}^T \mathbf{A}$ are stored, $L_n(p^{*(n)}[k], k)$ must be evaluated KN times to find the maximizing k per (34). In line 7, $\mathcal{O}(n_c N + n_c n_\theta)$ operations are necessary. In lines 8–12, the dominant calculation is the computation of $\boldsymbol{\Sigma} \mathbf{d}$, which
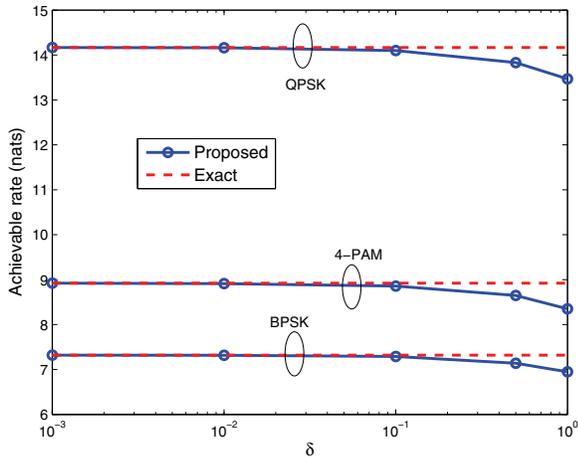
Fig. 1.   Achievable rates for a single-user case.

Fig. 2.   Number of constraints and extra variables.

takes $\mathcal{O}(n_\theta^2)$ operations. Therefore, the overall per-iteration complexity with large N and small δ grows as $\mathcal{O}(N^2(\log \frac{1}{\delta})^2)$.

B. Suboptimal Algorithm

To benchmark performance of the proposed near-optimal scheme, a simple suboptimal algorithm based on alternating minimization is also considered. Note that with the subcarrier allocation $\mathbf{k}$ fixed, the optimization problem (P1) with regard to only $\mathbf{p}$ is a convex problem. Also, with $\mathbf{p}$ fixed, solving for $\mathbf{k}$ is straightforward since the problem naturally decouples across subcarriers, and boils down to choosing on each subcarrier n the user that yields the maximum contribution to the objective. Iterating these two steps will monotonically increase the objective, which can be stopped when no further increase is made. It is noted that this algorithm can be thought of as an extension of the algorithm in [13] to the OFDMA case, since the latter can be employed as a subroutine (after appropriate adaptation to cope with practical constellations). The procedure can be described in pseudocode as follows.

- Step 1: Initialize $\mathbf{p}$, e.g., set $p^{(n)} = P_{\max}/N$ for $n = 1, 2, \ldots, N$.
- Step 2: Set $k(n) = \arg\max_k w_k \mathcal{I}_k^{(n)}(h_k^{(n)} p^{(n)})$ for $n = 1, 2, \ldots, N$.
- Step 3: With $\mathbf{k}$ fixed, solve (P1) only over $\mathbf{p}$ using convex optimization techniques (or by using an adapted version of the algorithm in [13].)
- Step 4: If the objective is not increased, stop; otherwise, go to Step 2.

V. NUMERICAL TESTS

The proposed RA algorithm is verified via numerical tests. A Rayleigh-faded, 4-path channel is simulated. The pathloss exponent is set to $\alpha = 2$. Assuming Gaussian-distributed $\mathbf{g}$, we set $\Omega = Q^{-1}(\epsilon)$ for small $\epsilon > 0$. The covariance of the channel estimation error follows the model in [27]. Unless stated otherwise, the values of $P_{\max} = 10^2$ and $I_{\max} = 1$ were used with unit-power channel coefficients and additive noise.
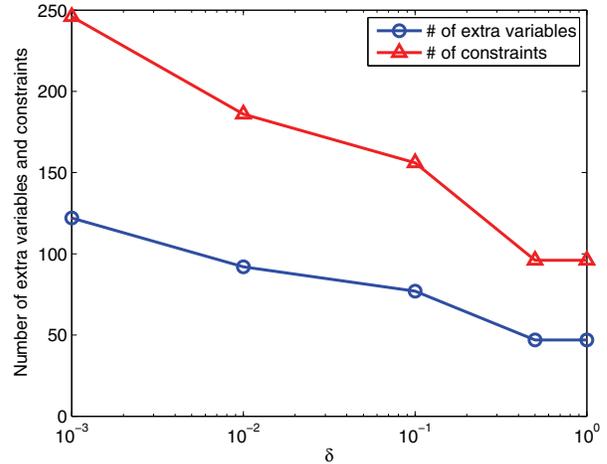
In order to validate the polyhedral approximation of the second-order cone constraint, the single-user case is first examined. Recall that when $K = 1$, the original problem (P1) is convex, which can be easily solved using generic convex optimization software for optimal power allocation across subcarriers. The solid lines with circle markers in Fig. 1 represent the sum rates in nats obtained by solving (P2) with $\epsilon = 0.1$ for various values of δ, where $N = 16$ subcarriers were used. Recall that when δ is small, the polyhedral approximation for the interference constraint is tight. Results using three different practical constellations are shown, namely, QPSK, 4-PAM, and BPSK constellations. Also shown in dashed lines are the optimal sum rates obtained by solving (P1) directly. It can be seen that the rate performances from (P2) are indistinguishable from those of (P1) when $\delta \leq 0.1$.

As was discussed in Sec. III, decreasing δ incurs increase in the number of optimization variables and constraints. Fig. 2 depicts the number of extra variables introduced for the polyhedral approximation as well as the number of constraints, as δ is varied. It can be observed that as δ is decreased by orders of magnitude, the corresponding growth in the problem complexity is rather mild.

In the multi-user case with $K > 1$, (P1) becomes hard due to the combinatorial search for optimal subcarrier assignment. Exhaustive search would require enumerating K^N different assignments, and solving a (convex) power allocation problem per (P1) for each of the assignments. Fig. 3 shows the case of $K = 2$ users, with user 2 located four times farther than user 1 from the CR-BS. Equal weights of $w_1 = w_2 = 0.5$ were used. User 1 employed a BPSK constellation and user 2 QPSK, and the number of subcarriers $N = 8$. Two sets of curves corresponding to $\epsilon = 0.1$ and $\epsilon = 0.01$ are presented in Fig. 3 in solid and dashed lines, respectively. In each set, the line with circle markers represents the weighted sum rate obtained using the proposed algorithm, and the line without markers corresponds to the results from exhaustive search. In addition, the results from the suboptimal algorithm in Sec. IV-B are depicted as the lines with triangle markers, and the rate achievable under perfect channel state information on
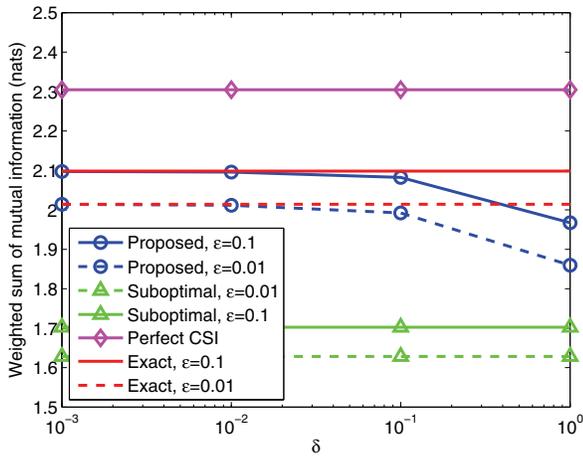
Fig. 3.   Weighted sum rates when $K = 2$.
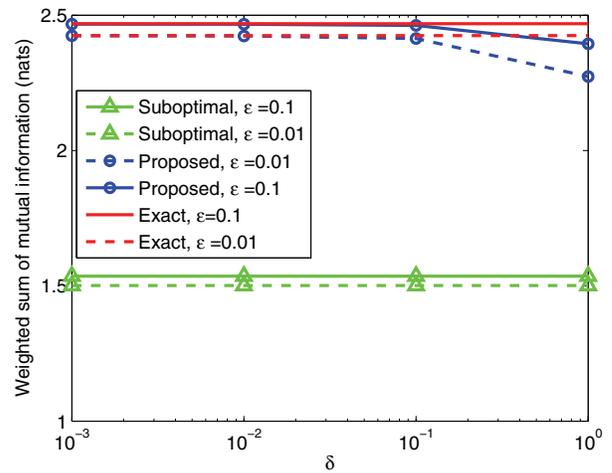
Fig. 5.   Weighted sum rates when $K = 3$.
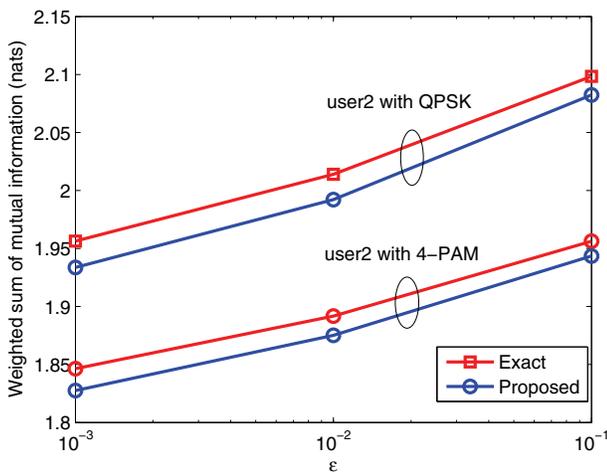
Fig. 4.   Weighted sum rates versus ϵ.
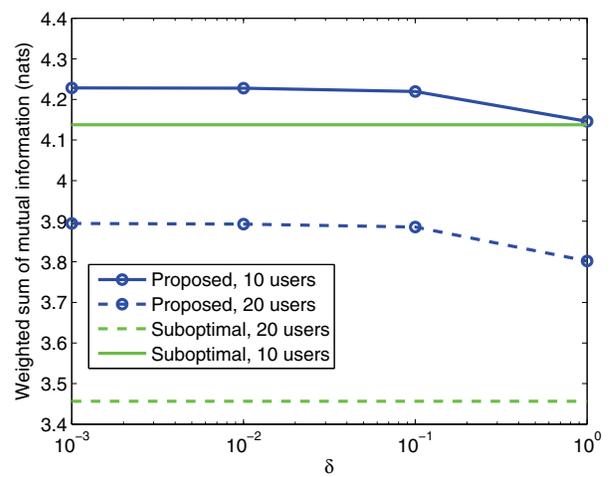
Fig. 6.   Weighted sum rates when $K = 10$ and $K = 20$.

the interfering channel with diamond markers. It can be seen that the proposed algorithm approaches the performance of exhaustive search with sizeable improvement relative to that of the suboptimal heuristic, although the loss due to channel uncertainty is evident.

Fig. 4 depicts the weighted sum rates for the same set-up as used for Fig. 3, but with user 2 employing QPSK or 4-PAM modulations. The value of $\delta = 0.1$ was used and the values of ϵ were varied. The lines with square markers and the ones with circle markers correspond to the results of exhaustive search and the proposed method, respectively. It can be verified that the proposed method can achieve near-optimal performance at a moderate value of δ. Also, as expected, the rate performance improves as ϵ grows larger, as this amounts to more lenient interference constraints.

The case with $K = 3$ CR users were also tested and shown in Fig. 5 for $\epsilon = 0.1$ and $\epsilon = 0.01$. The distances of users 2 and 3 from the CR-BS were four times that of user 1, and the input constellations for user 1, 2, and 3 were set to BPSK, 4-PAM, and QPSK, respectively. Equal weights were used. The value of $I_{\max}$ was set to 10 and $N = 8$ subcarriers (or subbands) were employed. The results from the proposed and the suboptimal algorithms as well as the exhaustive search are

presented. The overall trend is quite similar to the two-user case.

The proposed algorithm was tested for larger values of K. Fig. 6 depicts the cases with $K = 10$ and $K = 20$, where $N = 32$ and $\epsilon = 0.1$ were used. Due to prohibitive computational complexity associated with exhaustive search, only the results from the proposed and the suboptimal methods are plotted. Again, it can be seen that the achieved weighted sum rates saturate with $\delta \leq 0.1$, implying that they have reached the optimal values.

To get a rough idea on the practical implementation complexity of the algorithms, actual run times of the proposed and the benchmark algorithms are presented in Table II for the cases with $N = 8$ and $N = 32$ using $\tau = n_\theta \times 10^{-3}$. A 2.3 GHz Intel processor with 6 GB of RAM was used. Since the convex optimization subroutine called by the suboptimal algorithm was implmented in a native code, while the proposed algorithm was run in Matlab, which is interpreter-based, the suboptimal algorithm may be at a slight advantage. Nonetheless, it can be seen that our algorithm is quite competitive.

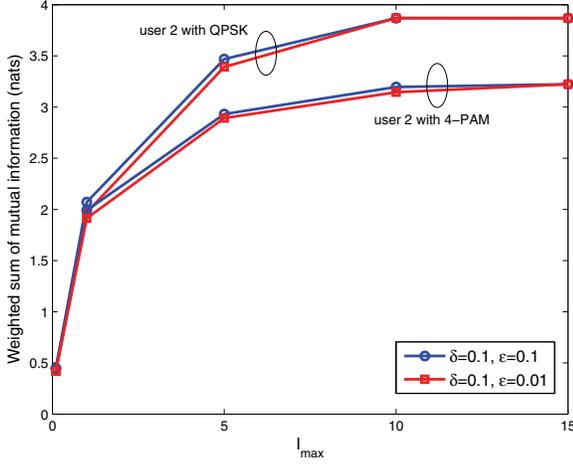To assess the sensitivity of the rate performance to the

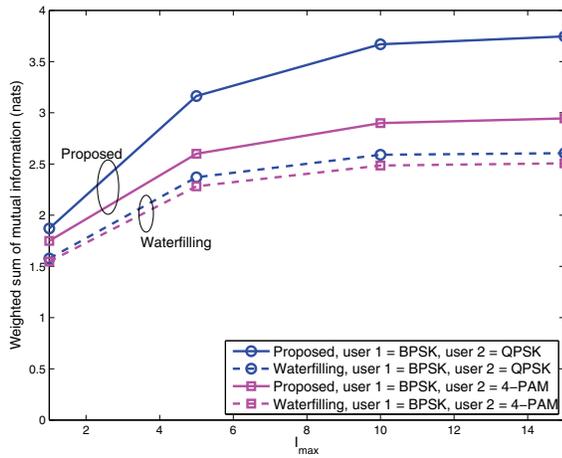Fig. 7. Weighted sum rates versus I_max.

Fig. 8. Performance comparison to algorithm with water-filling.

PU interference constraint, the weighted sum rates are plotted against the value of I_max in Fig. 7 for $\delta = 0.1$ with $N = 8$ and $K = 2$. As I_max is increased, the CR can allocate more power to serve the CR users, yielding a higher total rate. Once I_max becomes large enough, the total power constraint becomes the bottleneck, which explains the leveling of the curves at high I_max in Fig. 7. On the other hand, when I_max is very small, the use of higher order modulation does not seem to be warranted.

To characterize the performance gain of specifically accounting for practical modulation constellations, the rates achievable using the power and subcarrier allocation based on traditional Gaussian coding are compared to the rates from the proposed design. The same set-up as in Fig. 3 is used with $K = 2$ and $\epsilon = 0.01$. The solid lines in Fig. 8 represent the case where user 2 uses QPSK modulation, while the dashed lines the case with 4-PAM. The circle markers signify the proposed method in which the mercury/water-filling is used, whereas the square markers correspond to the method with the water-filling. It can be seen that regardless of the constellation, the proposed formulation yields superior performance.

TABLE II
RUN TIMES OF RA ALGORITHMS IN SECONDS.

		$\delta = 0.01$	$\delta = 0.1$	$\delta = 1$
$N = 32$	proposed	758	473	336
$K = 10$	suboptimal		180.	
$N = 8$	proposed	18.1	14.2	12.2
$K = 3$	suboptimal		2.74	

VI. CONCLUSIONS

A weighted sum rate maximization problem was formulated for a CR system employing OFDMA. Due to the uncertainty present in the CR-to-PU channel, a robust interference constraint with an ellipsoidal uncertainty set was imposed to protect the PU system, which is equivalent to a second-order cone constraint. When only one CR user is served by the BS, the optimization problem is convex. When multiple users are present, due to the combinatorial search for optimal subcarrier assignment, the problem is non-convex and hard to solve. In fact, since the second-order cone constraint lacks separable structure, the optimization variables are coupled across all subcarriers. Therefore, a polyhedral approximation was introduced to break the coupling. The dual method could then be employed to decompose the overall problem into per-subcarrier sub-problems, which can be easily solved. Since the complexity of the polyhedral approximation scales modestly (polynomially) in the number of subcarriers, the overall algorithm can efficiently find the near-optimal power loading and subcarrier assignment. Numerical tests verified the efficiency of the novel algorithm.

APPENDIX

APPENDIX: DUALITY GAP OF (P2)

Here, a claim related to the duality gap of (P2) is proved. To this end, we will invoke [21, Proposition 5.26], which considers problems of the following form:

$$\text{(P3)} \quad \min \sum_{i=1}^{I} f_i(\mathbf{x}_i) \tag{37}$$

$$\text{subject to } \mathbf{x}_i \in \mathcal{X}_i \tag{38}$$

$$\sum_{i=1}^{I} \mathbf{h}_i(\mathbf{x}_i) \leq \boldsymbol{\gamma} \tag{39}$$

where $\mathcal{X}_i \subset \mathbb{R}^{d_i}$, $\boldsymbol{\gamma} \in \mathbb{R}^m$, $f_i : \mathcal{X}_i \to \mathbb{R}$ and $\mathbf{h}_i : \mathcal{X}_i \to \mathbb{R}^m$. Then, it was shown in [21] that under conditions

c1) There exists at least one feasible solution to (P3).
c2) For each i, the set $\{\mathbf{x}_i, \mathbf{h}_i(\mathbf{x}_i), f_i(\mathbf{x}_i) | \mathbf{x}_i \in \mathcal{X}_i\}$ is compact.
c3) For each i, given any $\tilde{\mathbf{x}}_i \in \mathrm{conv}(\mathcal{X}_i)$, there exists $\mathbf{x}_i \in \mathcal{X}_i$ such that $\mathbf{h}_i(\mathbf{x}_i) \leq \tilde{\mathbf{h}}_i(\tilde{\mathbf{x}}_i)$, where $\mathrm{conv}(\cdot)$ denotes the convex hull operator, and

$$\tilde{\mathbf{h}}_i(\tilde{\mathbf{x}}) := \inf \left\{ \sum_{j=1}^{d_i+1} \pi^j \mathbf{h}_i(\mathbf{x}^j) \middle| \tilde{\mathbf{x}} = \sum_{j=1}^{d_i+1} \pi^j \mathbf{x}^j, \ \mathbf{x}^j \in \mathcal{X}_i, \right.$$

$$\left. \sum_{j=1}^{d_i+1} \pi^j = 1, \ \pi^j \geq 0 \right\} \tag{40}$$

the duality gap can be upper-bounded by $(m+1)\max_{i\in\{1,2,\dots,I\}}\rho_i$, where

$$\rho_i \le \sup\{f_i(\mathbf{x}_i)|\mathbf{x}_i \in \mathcal{X}_i\} - \inf\{f_i(\mathbf{x}_i)|\mathbf{x}_i \in \mathcal{X}_i\}. \quad (41)$$

Thus, as long as $\max_i \rho_i$ grows sublinearly in i, the duality gap of (P3) normalized by I vanishes as $I \to \infty$. In other words, provided that the optimal objective of (P3) is proportional in I, the fraction of the duality gap to the objective goes to zero as I grows.

This result is applicable to (P2). For this, set $I = N + 1$, and let $\mathbf{x}_n := (k(n), p^{(n)})$ for $n = 1, 2, \dots, N$, and $\mathbf{x}_{N+1} := \mathbf{q}$. Then, $\mathcal{X}_n := \mathcal{K} \times [0, P_{\max}^{(n)}]$ for $n = 1, \dots, N$. Without loss of generality, $\mathbf{q}$ can be confined to a convex and compact set $\mathcal{Q}$, and let $\mathcal{X}_{N+1} := \mathcal{Q}$. Furthermore, let $f_n(\mathbf{x}_n) := w_{k(n)}\mathcal{I}_{k(n)}^{(n)}(h_{k(n)}^{(n)}p^{(n)})$ for $n = 1, \dots, N$, $f_{N+1}(\mathbf{x}_{N+1}) = 0$, $\mathbf{h}_n(\mathbf{x}_n) := [1, \mathbf{A}(:, n)^T]^T p^{(n)}$ for $n = 1, \dots, N$, $\mathbf{h}_{N+1}(\mathbf{x}_{N+1}) := [\mathbf{0}^T, \mathbf{B}^T]^T \mathbf{q}$, and $\boldsymbol{\gamma} = [P_{\max}, \mathbf{b}^T]^T$. Then, conditions c1) and c2) are readily satisfied. It is also trivial to verify that c3) is met, since $\mathbf{h}_n(\mathbf{x}_n)$ for $n = 1, \dots, N$, is constant with respect to $k(n)$, and also linear in $p^{(n)}$. Therefore, a duality gap over-estimate of (P2) is given by

$$(n_c + 2) \max_{n\in\{1,2,\dots,N\}} \max_{k\in\mathcal{K}} w_k \mathcal{I}_k^{(n)}(h_k^{(n)}P_{\max}^{(n)}) \quad (42)$$

which is related to the highest-order constellation employed by the OFDMA users, and is thus finite. Therefore, the ratio of the duality gap of (P2) to the optimal objective approaches zero as N grows large.

REFERENCES

[1] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access: signal processing, networking, and regulatory policy," *IEEE Signal Process. Mag.*, vol. 4, no. 3, pp. 79–89, May 2007.

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

[3] Y. Zhang and C. Leung, "Resource allocation in an OFDM-based cognitive radio system," *IEEE Trans. Commun.*, vol. 57, no. 7, pp. 1928–1931, July 2009.

[4] S. M. Almalfouh and G. L. Stüber, "Interference-aware radio resource allocation in OFDMA-based cognitive radio networks," *IEEE Trans. Veh. Technol.*, vol. 60, no. 4, pp. 1699–1713, May 2011.

[5] A. Nemirovski and A. Shapiro, "Convex approximations of chance constrained programs," *SIAM J. Optim.*, vol. 17, no. 4, pp. 969–996, 2006.

[6] A. Ben-Tal and A. Nemirovski, "Selected topics in robust convex optimization," *Math. Prog., Ser. B*, vol. 112, pp. 125–158, 2008.

[7] Y. J. Zhang and A. M.-C. So, "Optimal spectrum sharing in MIMO cognitive radio networks via semidefinite programming," *IEEE J. Sel. Topics Commun.*, vol. 29, no. 2, pp. 362–373, Feb. 2011.

[8] G. Zheng, K.-K. Wong, and B. Ottersten, "Robust cognitive beamforming with bounded channel uncertainties," *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4871–1881, Dec. 2009.

[9] E. Dall'Anese, S.-J. Kim, G. B. Giannakis, and S. Pupolin, "Power control for cognitive radio networks under channel uncertainty," *IEEE Trans. Wireless Commun.*, vol. 10, no. 10, pp. 3541–3551, Dec. 2011.

[10] I. C. Wong and B. L. Evans, "Resource allocation in the OFDMA downlink with imperfect channel knowledge," *IEEE Trans. Wireless Commun.*, vol. 57, no. 1, pp. 232–241, Jan. 2009.

[11] R. Aggarwal, M. Assaad, C. E. Koksal, and P. Schniter, "Joint scheduling and resource allocation in the OFDMA downlink: utility maximization under imperfect channel-state information," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5589–5604, Nov. 2011.

[12] D. Huang, Z. Shen, C. Miao, and C. Leung, "Resource allocation in MU-OFDM cognitive radio systems with partial channel state information," *EURASIP J. Wireless Commun. Netw.*, Apr. 2010.

[13] K. Son, B. C. Jung, S. Chong, and D. K. Sung, "Power allocation for OFDM-based cognitive radio systems under outage constraints," in *Proc. 2010 Intl. Conf. Commun.*.

[14] R. Wang, V. K. N. Lau, L. Lv, and B. Chen, "Joint cross-layer scheduling and spectrum sensing for OFDMA cognitive radio systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 5, pp. 2410–2416, May 2009.

[15] K. Seong, M. Mohseni, and J. M. Cioffi, "Optimal resource allocation for OFDMA downlink systems," in *Proc. 2006 Intl. Symp. on Info. Theory*, pp. 1394–1398.

[16] Z.-Q. Luo and S. Zhang, "Dynamic spectrum management: complexity and duality," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 57–73, Feb. 2008.

[17] S.-J. Kim, N. Y. Soltani, and G. B. Giannakis, "Resource allocation for OFDMA cognitive radios under channel uncertainty," in *Proc. 2011 Intl. Conf. on Acoustics, Speech and Signal Processing*, pp. 3188–3191.

[18] A. Lozano, A. M. Tulino, and S. Verdú, "Optimum power allocation for multiuser OFDM with arbitrary signal constellations," *IEEE Trans. Commun.*, vol. 56, no. 5, pp. 828–837, May 2008.

[19] D. Guo, S. Shamai, and S. Verdú, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1261–1283, Apr. 2005.

[20] N. Y. Soltani, S.-J. Kim, and G. B. Giannakis, "Chance-constrained optimization of uplink parameters for OFDMA cognitive radios," in *Proc. 2012 Intl. Conf. on Acoustics, Speech and Signal Processing*.

[21] D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*. Athena Scientific, 1996.

[22] A. Ben-Tal and A. Nemirovski, "On polyhedral approximations of the second-order cone," *Math. Oper. Res.*, vol. 26, no. 2, pp. 193–205, May 2001.

[23] F. Glineur, "Computational experiments with a linear approximation of second-order cone optimization," Service de Mathématique et de Recherche Opérationnelle, Faculté Polytechnique de Mons, Mons, Belgium, Tech. Rep. 0001, Nov. 2000.

[24] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, and T. Bostoen, "Optimal multiuser spectrum balancing for digital subscriber lines," *IEEE Trans. Wireless Commun.*, vol. 54, no. 5, pp. 922–933, May 2006.

[25] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization for multicarrier systems," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1310–1322, July 2006.

[26] S. Boyd, "Ellipsoid method," http://www.stanford.edu/class/ee364b/lectures/ellipsoid_method_notes.pdf, 2008.

[27] Y. Yao and G. B. Giannakis, "Rate-maximizing power allocation in OFDM based on partial channel knowledge," *IEEE Trans. Wireless Commun.*, vol. 4, no. 3, pp. 1073–1083, May 2005.

Seung-Jun Kim (SM'12) received his B.S. and M.S. degrees from Seoul National University in Seoul, Korea in 1996 and 1998, respectively, and his Ph.D. from the University of California at Santa Barbara in 2005, all in electrical engineering. From 2005 to 2008, he worked for NEC Laboratories America in Princeton, New Jersey, as a research staff member. He is currently with the Digital Technology Center at the University of Minnesota, where he is a Research Associate. He is also affiliated with the Department of Electrical and Computer Engineering at the University of Minnesota, where he is a Research Assistant Professor. His research interests lie signal processing, optimization, and machine learning with applications to wireless communication and networking, smart power grids, bio and social networks.

Nasim Yahya Soltani (S'13) received the B.Sc. degree in electrical engineering from the University of Tehran, Tehran, Iran, in 2003 and the M.Sc. degree in electrical engineering from Iran University of Science and Technology, Tehran, Iran, in 2006. Since September 2010, she has been working toward the Ph.D. degree with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis. Her research interests lie in the areas of statistical signal processing, robust optimization, machine learning, network science, wireless communications and smart power grids.

G. B. Giannakis (Fellow'97) received his Diploma in Electrical Engr. from the Ntl. Tech. Univ. of Athens, Greece, 1981. From 1982 to 1986 he was with the Univ. of Southern California (USC), where he received his MSc. in Electrical Engineering, 1983, MSc. in Mathematics, 1986, and Ph.D. in Electrical Engr., 1986. Since 1999 he has been a professor with the Univ. of Minnesota, where he now holds an ADC Chair in Wireless Telecommunications in the ECE Department, and serves as director of the Digital Technology Center.

His general interests span the areas of communications, networking and statistical signal processing - subjects on which he has published more than 350 journal papers, 580 conference papers, 20 book chapters, two edited books and two research monographs (h-index 104). Current research focuses on sparsity and big data analytics, wireless cognitive radios, mobile ad hoc networks, renewable energy, power grid, gene-regulatory, and social networks. He is the (co-) inventor of 21 patents issued, and the (co-) recipient of 8 best paper awards from the IEEE Signal Processing (SP) and Communications Societies, including the G. Marconi Prize Paper Award in Wireless Communications. He also received Technical Achievement Awards from the SP Society (2000), from EURASIP (2005), a Young Faculty Teaching Award, and the G. W. Taylor Award for Distinguished Research from the University of Minnesota. He is a Fellow of EURASIP, and has served the IEEE in a number of posts, including that of a Distinguished Lecturer for the IEEE-SP Society.