Chance-Constrained Optimization of OFDMA Cognitive Radio Uplinks

Nasim Yahya Soltani, Seung-Jun Kim, *Senior Member, IEEE,* and Georgios B. Giannakis, *Fellow, IEEE*

***Abstract*—A resource allocation task for the uplink of OFDMA-based cognitive radio (CR) systems is considered. The weighted sum-rate is maximized over subcarrier assignment as well as over power loading per CR user, while protecting primary user (PU) systems. However, due to the lack of explicit support from PU systems, the channels from CR users to the PU may not be accurately acquired. Thus, the PU interference constraint is posed as a chance constraint, for which conservative convex approximation is employed for tractability. In particular, to mitigate the combinatorial complexity incurred for optimal subcarrier assignment, a separable structure is pursued, and the dual decomposition method is adopted to obtain near-optimal solutions. Numerical tests verify that the proposed algorithms yield higher weighted sum-rate at lower computational complexity than a benchmark algorithm.**

***Index Terms*—Cognitive radios, resource allocation, orthogonal frequency-division multiple access, channel uncertainty, chance-constrained programming.**

I. INTRODUCTION

COGNITIVE radios (CRs) aim to mitigate the scarcity of spectral resources by allowing opportunistic use of the bands licensed to primary user (PU) systems. In an *overlay* scenario, CRs target to transmit over unused parts of the spectrum (called white space), which can be identified via, e.g., spectrum sensing. In the *underlay* scenario, which is the setting of interest in the present work, CRs share the same band with the PUs by carefully controlling the interference caused to the PU system. In this case, the channel gain estimates between the CR transmitters and the PU receivers are needed [1].

Since CRs do not receive explicit support from the PUs, acquiring accurate channel estimates is often challenging. Therefore, much research effort has been devoted to ensure that the interference constraint is effected robustly against channel uncertainty [2], [3], [4], [5]. Statistical knowledge of the channel uncertainties was assumed in [2], [3], [4]. A bounded region (e.g., an ellipsoid) in which the uncertainty is confined was postulated in [5]. Multiple-antenna CRs with channel uncertainty were treated in [6], [7].

Orthogonal frequency division multiple access (OFDMA) is a natural candidate for CR systems due to its flexibility in controlling spectrum usage. A major resource allocation (RA) task for OFDMA radios is to allocate subcarriers to different users, and to load each subcarrier with proper power levels. Extensive research results are available on this topic including [8] and [9]. A heuristic algorithm based on the multi-dimensional knapsack problem was proposed for OFDM CRs in [8]. A weighted sum-rate maximization problem was considered for OFDMA CRs and near-optimal algorithms were proposed in [9]. However, very few works have addressed the RA problem for OFDMA-based CRs under channel uncertainty. Moreover, most existing approaches focus on the downlink scenario, which are not readily extendible to the uplink [3], [10].

CR interference constraints under channel uncertainty can be cast as chance constraints. However, chance constraints are typically more difficult to handle than their deterministic counterparts, as they may be either nonconvex, or tough to verify as being convex. Moreover, it is sometimes difficult to express these constraints in closed form. In such cases, convex approximation of chance constraints is of practical merit, as was exemplified in [11], where the Bernstein method was first employed for a chance-constrained RA problem for wireless systems.

The present paper addresses the RA task for OFDMA uplink CRs with uncertain CR-to-PU channels. A weighted sum-rate maximization problem is formulated under a probabilistic interference constraint and maximum transmit-power constraints for the CR users. The Bernstein method is adopted to approximate the probabilistic constraint by a convex constraint. Even after the approximation, the overall problem is still nonconvex due to the combinatorial assignment of users to each subcarrier. By employing appropriate bounds, the approximation emerging from the interference constraint can be further made *separable* across subcarriers. This opens the door to the dual decomposition approach, which leads to a near-optimal and computationally efficient solution [12].

A dual decomposition approach was taken in [3] for OFDMA *downlink* CR problem with channel uncertainty. Different challenges are associated with the downlink and the uplink scenarios; see. Fig. 1. In the downlink setup, the CR base station (BS) is the sole transmitter, and the interference channel (depicted as the dashed arrow in Fig. 1) between the CR BS and the PU does not depend on the CR user assignment decision on each subcarrier. In the uplink, however, the user assignment per subcarrier completely alters the associated interference channel toward the PU. In [3], the PU protection constraint was modeled as a second-order cone constraint involving transmit-powers in different subcarriers,

Manuscript received January 31, 2012; revised July 6, October 15, and November 26, 2012; accepted November 26, 2012. The associate editor coordinating the review of this paper and approving it for publication was G. Yue.

The authors are with the Dept. of ECE, University of Minnesota, 200 Union Street SE, Minneapolis, MN 55455, U.S.A. (e-mail: {yahya015, seungjun, georgios}@umn.edu).

Part of the work was presented at the *Intl. Conf. on Acoust., Speech and Signal Proc.*, Kyoto, Japan, March 2012. This work was supported by QNRF grant NPRP 09-341-2-128.

Digital Object Identifier 10.1109/TWC.2013.011713.120155

which was then tightly approximated by a polyhedral constraint for tractability. This technique cannot be applied to the uplink because the relevant constraint depends not only on the transmit-powers but also on the user assignment.

The rest of the paper is organized as follows. The problem is formulated in Sec. II, and Bernstein's approximation technique tailored for chance constraints is outlined in Sec. III. The RA algorithm is developed in Sec. IV. Numerical tests are presented in Sec. V, followed by conclusions in Sec. VI.

A note on notations: Vector quantities are denoted as bold letters, and the sets as calligraphic upper-case letters. Superscript $^{(n)}$ as in $p^{(n)}$ denotes the quantities related to the n-th subcarrier, while CR user indices are represented using subscripts as in $g_k^{(n)}$. Transposition is denoted by ^{T} and $\Pr\{\cdot\}$ represents the probability.

II. Problem Statement

Consider the uplink mode of a network comprising K CR users communicating with their BS using OFDMA over N subcarriers. The instantaneous channel gain $\tilde{g}_k^{(n)}$ between CR user $k \in \mathcal{K} \triangleq \{1, 2, \ldots, K\}$ and the CR BS on subcarrier $n \in \mathcal{N} \triangleq \{1, 2, \ldots, N\}$ is assumed to be acquired accurately via conventional channel estimation techniques. It is further assumed that during the spectrum sensing phase, the presence of an active PU has been detected. It is straightforward to handle multiple PUs in the framework to be proposed. In order to limit the interference inflicted to the PU, the channels from the CR users to the PU must be known. Let $g_k^{(n)}$ denote the channel gain from the k-th CR to the PU receiver on subcarrier n. Due to lack of cooperation from the PU system, it is difficult to estimate $g_k^{(n)}$ precisely. To capture this uncertainty, $g_k^{(n)}$ is modeled as a random variable.

A relevant RA problem is to maximize the weighted sum of all CR throughputs under the transmit-power constraints (one per CR), and the PU interference constraint. Let $p^{(n)}$ denote the transmit-power loaded on subcarrier n, where $0 \leq p^{(n)} \leq P_{\max}^{(n)}$. Let $\mathbf{p}$ and $\mathbf{P}_{\max}$ be the vectorized versions of $\{p^{(n)}\}$ and $\{P_{\max}^{(n)}\}$, respectively. Also, let $k(n) \in \mathcal{K}$ represent the index of the user served on subcarrier n, and define $\mathbf{k} \triangleq [k(1), \ldots, k(N)]^T$. With w_k denoting the positive weight for user $k \in \mathcal{K}$, the following chance-constrained optimization problem is of interest.

$$\text{(P1)} \quad \max_{0 \preceq \mathbf{p} \preceq \mathbf{P}_{\max}, \mathbf{k} \in \mathcal{K}^N} \sum_{n \in \mathcal{N}} w_{k(n)} \log\left(1 + \tilde{g}_{k(n)}^{(n)} p^{(n)}\right) \tag{1}$$

$$\text{subject to} \quad \sum_{n \in \mathcal{N}: k(n) = k} p^{(n)} \leq P_{k,\max}, \quad k \in \mathcal{K} \tag{2}$$

$$\Pr\left\{\sum_{n \in \mathcal{N}} g_{k(n)}^{(n)} p^{(n)} < I_{\max}\right\} \geq 1 - \epsilon \tag{3}$$

where (2) is the per-CR power constraint, and (3) enforces that the interference power at the PU stays below $I_{\max}$ with probability no less than $1 - \epsilon$ with $\epsilon \in (0, 1)$ denoting the desired upper-bound on the probability that the interference threshold is exceeded. In case of more than one PUs present, multiple PU protection constraints analogous to (3) can be imposed. The widely accepted constraint is to control the "interference temperature" such that the interference experienced by the PU system is below a certain threshold on the order of the background noise and interference.

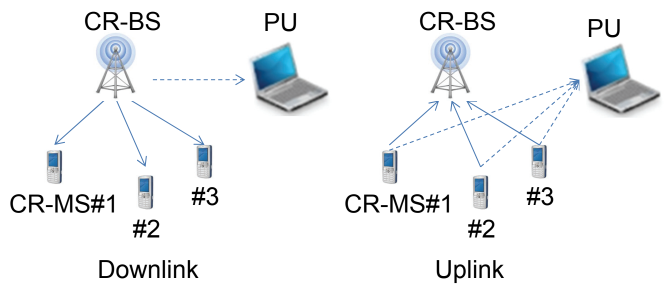


Fig. 1. Downlink and uplink modes.

The feasible set of (3) can be either convex or nonconvex, depending on the distribution of $g_k^{(n)}$ [13]. For example, $\Pr\left\{\mathbf{a}^T\mathbf{u} < \mathbf{b}\right\} \geq 1 - \epsilon$ is convex for $\epsilon < 1/2$, if $[\mathbf{a}^T \ \mathbf{b}^T]^T$ has a symmetric logarithmically concave density [13]. However, even if (3) is convex, it may not be straightforward to express it in closed form, rendering the optimization problem intractable. Moreover, the overall problem would be still nonconvex due to the combinatorial search over $\mathbf{k}$ needed for the subcarrier assignment to CR users.

In the following, a convex approximation of (3) is advocated. In particular, the approximation is made conservative in the sense that the approximate constraint implies the original constraint (3). This will be achieved by using the Bernstein method [13], [14]. By judiciously choosing the form of the approximation, one can also ensure that the approximate constraint is separable in n. Then, an efficient near-optimal solution will be obtained via dual decomposition [12], [15].

III. Approximation of Chance Constraints

A. Bernstein Approximation

A useful class of approximation techniques for chance constraints known as Bernstein approximations is briefly reviewed in the present context [13], [14]. Consider a chance constraint of the form

$$\Pr\left\{f_0(\mathbf{p}) + \sum_{n=1}^{N} \zeta_n f_n(\mathbf{p}) < 0\right\} \geq 1 - \epsilon \tag{4}$$

where $\mathbf{p}$ is a deterministic parameter vector, and $\{\zeta_n\}$ are random variables with marginal distributions denoted as $\{\pi_n\}$. Suppose that one desires to meet this constraint for a given family of $\{\zeta_n\}$ distributions, under the following assumptions.

as1) $\{f_n(\mathbf{p})\}$ are affine in $\mathbf{p}$ for $n = 0, 1, \ldots, N$;
as2) $\{\zeta_n\}$ are independent of each other; and
as3) $\{\pi_n\}$ have a common bounded support of $[-1, 1]$; that is, $-1 \leq \zeta_n \leq 1$ for all $n = 1, \ldots, N$.

Under these assumptions, the following constraint constitutes a conservative substitute and thus implies (4)

$$\inf_{\rho > 0} \left[f_0(\mathbf{p}) + \rho \sum_{n=1}^{N} \Omega_n\left(\rho^{-1} f_n(\mathbf{p})\right) + \rho \log\left(\frac{1}{\epsilon}\right)\right] \leq 0 \tag{5}$$

where $\Omega_n(y) \triangleq \max_{\pi_n} \log\left(\int \exp(xy)d\pi_n(x)\right)$. Moreover, it is guaranteed that (5) is convex [13], [14]. The approximation is useful when $\{\Omega_n(y)\}$ can be evaluated efficiently. In general, one can consider an upper-bound for $\Omega_n(y)$ given by

$$\Omega_n(y) \leq \max\{\mu_n^- y, \mu_n^+ y\} + \frac{\sigma_n^2}{2}y^2, \quad n = 1, \ldots, N \tag{6}$$

where μ_n^-, μ_n^+ with $-1 \leq \mu_n^- \leq \mu_n^+ \leq 1$ and $\sigma_n \geq 0$ are constants that depend on the given families of probability distributions. Some examples are given in [14, Table 1], where the useful prior knowledge includes the support, unimodality (with respect to the center of the support), and symmetry of the distribution, as well as the ranges of the first- and the second-order moments. Using more prior knowledge leads to tighter approximation. Replacing $\Omega_n(\cdot)$ in (5) with this upper-bound, and invoking the arithmetic-geometric inequality, yields

$$f_0(\mathbf{p}) + \sum_{n=1}^{N} \max\{\mu_n^- f_n(\mathbf{p}), \mu_n^+ f_n(\mathbf{p})\} + \sqrt{2\log\frac{1}{\epsilon}}\left(\sum_{n=1}^{N}\sigma_n^2 f_n(\mathbf{p})^2\right)^{\frac{1}{2}} \leq 0 \tag{7}$$

as a convex conservative surrogate for (4).

Suppose now that the distributions of $g_k^{(n)}$ have bounded supports $[a_k^{(n)}, b_k^{(n)}]$.[1] The case with unbounded supports will be treated in Sec. III-B. Introduce constants $\alpha_k^{(n)} \triangleq \frac{1}{2}(b_k^{(n)} - a_k^{(n)}) \neq 0$ and $\beta_k^{(n)} \triangleq \frac{1}{2}(b_k^{(n)} + a_k^{(n)})$ to normalize the supports to $[-1, 1]$ per as3); that is,

$$\zeta_k^{(n)} \triangleq \frac{g_k^{(n)} - \beta_k^{(n)}}{\alpha_k^{(n)}} \in [-1, 1]. \tag{8}$$

Then, letting $f_0(\mathbf{p}) = -I_{\max} + \sum_{n=1}^{N} \beta_{k(n)}^{(n)} p^{(n)}$ and $f_n(\mathbf{p}) = \alpha_{k(n)}^{(n)} p^{(n)}$ for $n \in \mathcal{N}$, it follows that (4) is equivalent to (3). Thus, substituting these into (7), and noting that $p^{(n)} \geq 0$, one obtains

$$-I_{\max} + \sum_{n=1}^{N}\beta_{k(n)}^{(n)}p^{(n)} + \sum_{n=1}^{N}\mu_{k(n)}^{(n)+}\alpha_{k(n)}^{(n)}p^{(n)} + \sqrt{2\log\frac{1}{\epsilon}}\left(\sum_{n=1}^{N}\left(\sigma_{k(n)}^{(n)}\alpha_{k(n)}^{(n)}p^{(n)}\right)^2\right)^{\frac{1}{2}} \leq 0. \tag{9}$$

The overall RA problem corresponding to (P1) with (3) replaced by (9) is still nonconvex due to the combinatorial search for the optimal $\mathbf{k}$ over $\mathcal{K}^N$, which is a nonconvex set. In fact, as the variables $p^{(n)}$ are coupled nonlinearly through the last term in (9), the search complexity grows rapidly as N increases. To mitigate these issues, we consider adopting the dual decomposition method, which is applicable when the problem has a separable structure; that is, after relaxing coupling constraints through Lagrange relaxation, the problem must decompose into subproblems that can be solved independently. To bring about such a separable structure, we further approximate (9) by noting that the last term in (9) involves the ℓ_2-norm of the vector $[\sigma_{k(1)}^{(1)}\alpha_{k(1)}^{(1)}p^{(1)}, \ldots, \sigma_{k(N)}^{(N)}\alpha_{k(N)}^{(N)}p^{(N)}]$, and that $\|\mathbf{x}\|_2 \leq \sqrt{N}\|\mathbf{x}\|_\infty$ for any $\mathbf{x} \in \mathbb{R}^N$. Thus, the constraint becomes

$$\sum_{n=1}^{N}\gamma_{k(n)}^{(n)}p^{(n)} + \sqrt{2N\log\frac{1}{\epsilon}}\max_{n\in\mathcal{N}}\sigma_{k(n)}^{(n)}\alpha_{k(n)}^{(n)}p^{(n)} \leq I_{\max} \tag{10}$$

where $\gamma_{k(n)}^{(n)} \triangleq \mu_{k(n)}^{(n)+}\alpha_{k(n)}^{(n)} + \beta_{k(n)}^{(n)}$.

Alternatively, one can appeal to the fact that $\|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1$ to obtain yet another substitute for (3) as

$$\sum_{n=1}^{N}\gamma_{k(n)}^{(n)}p^{(n)} + \sqrt{2\log\frac{1}{\epsilon}}\sum_{n=1}^{N}|\sigma_{k(n)}^{(n)}\alpha_{k(n)}^{(n)}p^{(n)}| \leq I_{\max}. \tag{11}$$

Both (10) and (11) are amenable to dual decomposition, as will be discussed in Sec. IV. In the sequel, (P1) with (3) replaced by (9), (10) and (11) will be referred to as the ℓ_2-, ℓ_∞-, and ℓ_1-*approximate problems*, respectively.

Even when the channel distribution is known, if the corresponding chance constraint is nonconvex or hard to express analytically, it may be reasonable to apply Bernstein approximation for tractability, as long as conditions as1)–as3) are met, which include boundedness of the support of the distribution. The case of distributions with an unbounded support is discussed next.

B. Extensions to Channels with Unbounded Support

In the preceding discussion, Bernstein approximations were applied to bounded channel gains. While this may be reasonable considering the finite dynamic ranges of the A/D converters in the radios, presuming too large a range for the uncertain parameters might lead to a very loose approximation of the chance constraint. An alternative approach is developed here when the channel distributions are known (as opposed to the preceding case where the *family* of possible distributions was assumed to be known.)

Upon defining $I \triangleq \sum_n g_{k(n)}^{(n)}p^{(n)}$, it is possible to express $\Pr\{I < I_{\max}\}$ in (3) as

$$\Pr\{I < I_{\max}|\mathbf{a} \leq \mathbf{g} \leq \mathbf{b}\}\Pr\{\mathbf{a} \leq \mathbf{g} \leq \mathbf{b}\} + \Pr\{I < I_{\max}|\mathbf{g} < \mathbf{a} \text{ or } \mathbf{g} > \mathbf{b}\}\Pr\{\mathbf{g} < \mathbf{a} \text{ or } \mathbf{g} > \mathbf{b}\} \tag{12}$$

where $\mathbf{g} \triangleq [g_{k(1)}^{(1)}, ..., g_{k(N)}^{(N)}]^T$, $\mathbf{a} \triangleq [a_{k(1)}^{(1)}, ..., a_{k(N)}^{(N)}]^T$ and $\mathbf{b} \triangleq [b_{k(1)}^{(1)}, ..., b_{k(N)}^{(N)}]^T$ are appropriate constants determined such that

$$\delta \triangleq \Pr\{\mathbf{a} \leq \mathbf{g} \leq \mathbf{b}\} \in (1-\epsilon, 1). \tag{13}$$

Then, noting that the second term in (12) is no larger than $(1-\delta)$, and thus neglecting this term, (3) can be approximated conservatively as

$$\Pr\{I < I_{\max}\ |\mathbf{a} \leq \mathbf{g} \leq \mathbf{b}\} \geq \frac{1-\epsilon}{\Pr\{\mathbf{a} \leq \mathbf{g} \leq \mathbf{b}\}} = \frac{1-\epsilon}{\delta} \triangleq 1-\epsilon' \tag{14}$$

which can now be approximated by (10) or (11) with ϵ replaced by ϵ'. For concreteness, consider the cases with and without channel estimation over Rayleigh fading channels.

[1] In practice, the channel estimation techniques often yield not only the channel gain estimates, but also the associated confidence interval, from which one can effectively bound the channel gains.

1) Without Channel Estimation: Suppose first that channel estimation is not attempted. Then, one has only the prior knowledge that the probability density function (*p.d.f.*) of $g_k^{(n)}$ is exponential with mean $\bar{g}_k^{(n)}$ as

$$f_{g_k^{(n)}}(x) = \frac{1}{\bar{g}_k^{(n)}} \exp\left(-\frac{x}{\bar{g}_k^{(n)}}\right). \tag{15}$$

It then follows from the independence assumption that

$$\Pr\{\mathbf{a} \le \mathbf{g} \le \mathbf{b}\} = \prod_{n=1}^{N} \Pr\left\{a_{k(n)}^{(n)} \le g_{k(n)}^{(n)} \le b_{k(n)}^{(n)}\right\} = \prod_{n=1}^{N}\left[\exp\left(-\frac{a_{k(n)}^{(n)}}{\bar{g}_{k(n)}^{(n)}}\right) - \exp\left(-\frac{b_{k(n)}^{(n)}}{\bar{g}_{k(n)}^{(n)}}\right)\right] = \delta. \tag{16}$$

Since the peak of the *p.d.f.* of $g_k^{(n)}$ is located at the origin, it is natural to choose the lower-bound as $\mathbf{a} = \mathbf{0}$. To determine $\mathbf{b}$, $\Pr\left\{0 \le g_{k(n)}^{(n)} \le b_{k(n)}^{(n)}\right\}$ is enforced to be constant across n. Then, $\mathbf{b}$ is obtained as

$$b_{k(n)}^{(n)} = \bar{g}_{k(n)}^{(n)} \log \frac{1}{1-\delta^{\frac{1}{N}}}, \quad n \in \mathcal{N}. \tag{17}$$

2) With Channel Estimation: Consider now that OFDM channel estimation is performed using the minimum mean-square error (MMSE) estimator under Rayleigh fading [16]. The complex channel coefficient vector of the k-th user, $\mathbf{h}_k = \{h_k^{(n)}\}_{n=1}^N$, can be modeled as

$$\mathbf{h}_k = \hat{\mathbf{h}}_k + \tilde{\mathbf{h}}_k \tag{18}$$

where $\hat{\mathbf{h}}_k$ is the channel estimate, and $\tilde{\mathbf{h}}_k$ zero-mean Gaussian-distributed estimation error. The covariance matrix of $\tilde{\mathbf{h}}_k$ is given by [16]

$$\mathbf{\Sigma}_{\tilde{\mathbf{h}}_k} = \frac{N\sigma_h^2}{L+1}\mathbf{F}_L\mathbf{F}_L^H \tag{19}$$

where $(L+1)$ is the number of paths in the fading channel, σ_h^2 is a parameter that depends on the system parameters including Doppler spread and the noise variance at the receiver, and $\mathbf{F}_L$ is the $N \times (L+1)$ DFT matrix with (n,ℓ)-th entry given by $\frac{1}{\sqrt{N}}e^{-j\frac{2\pi}{N}(n-1)(\ell-1)}$. Therefore, $\mathbf{h}_k$ is distributed as circularly symmetric complex Gaussian with mean $\hat{\mathbf{h}}_k$ and covariance $\mathbf{\Sigma}_{\tilde{\mathbf{h}}_k}$. It is assumed that $\mathbf{h}_k$ is independent across k.

Upon defining $\hat{g}_k^{(n)} \triangleq \left|\hat{h}_k^{(n)}\right|^2$, the random variable $g_k^{(n)} = \left|h_k^{(n)}\right|^2$ is non-central Chi-square with two degrees of freedom and non-centrality parameter $\frac{2\hat{g}_k^{(n)}}{\sigma_h^2}$, with its *p.d.f.* given by

$$f_{g_k^{(n)}}(x) = \frac{1}{\sigma_h^2}\exp\left(-\frac{\hat{g}_k^{(n)}+x}{\sigma_h^2}\right) I_0\left(\frac{2}{\sigma_h^2}\sqrt{\hat{g}_k^{(n)}x}\right). \tag{20}$$

To apply Bernstein's approximation, it is further assumed here that $\{h_k^{(n)}\}$ are independent across n. This essentially holds true when the number of subchannels is close to the number of paths. However, even if this is not the case, the numerical tests in Sec. 5 will verify that the interference constraints are approximated conservatively. Proceeding under the assumption, one obtains

$$\Pr\{\mathbf{a} \le \mathbf{g} \le \mathbf{b}\} = \prod_{n=1}^{N} \Pr\left\{a_{k(n)}^{(n)} \le g_{k(n)}^{(n)} \le b_{k(n)}^{(n)}\right\} = \prod_{n=1}^{N}\left[Q\left(\sqrt{\frac{2\hat{g}_{k(n)}^{(n)}}{\sigma_h^2}}, \sqrt{\frac{2a_{k(n)}^{(n)}}{\sigma_h^2}}\right) - Q\left(\sqrt{\frac{2\hat{g}_{k(n)}^{(n)}}{\sigma_h^2}}, \sqrt{\frac{2b_{k(n)}^{(n)}}{\sigma_h^2}}\right)\right] = \delta \tag{21}$$

where $Q(\cdot,\cdot)$ is Marcum's Q-function. Similar to the case without channel estimation, we choose to set the factors in the product in (21) equal across n. Since the mode of the distribution of $g_{k(n)}^{(n)}$ is tightly upper-bounded by $\hat{g}_{k(n)}^{(n)}$ [17], it is natural to set $a_{k(n)}^{(n)} = \hat{g}_{k(n)}^{(n)} - \check{g}_{k(n)}^{(n)}$ and $b_{k(n)}^{(n)} = \hat{g}_{k(n)}^{(n)} + \check{g}_{k(n)}^{(n)}$ for $0 \le \check{g}_{k(n)}^{(n)} \le \hat{g}_{k(n)}^{(n)}$ such that

$$Q\left(\sqrt{\frac{2\hat{g}_{k(n)}^{(n)}}{\sigma_h^2}}, \sqrt{\frac{2(\hat{g}_{k(n)}^{(n)} - \check{g}_{k(n)}^{(n)})}{\sigma_h^2}}\right) - Q\left(\sqrt{\frac{2\hat{g}_{k(n)}^{(n)}}{\sigma_h^2}}, \sqrt{\frac{2(\hat{g}_{k(n)}^{(n)} + \check{g}_{k(n)}^{(n)})}{\sigma_h^2}}\right) = \delta^{\frac{1}{N}} \tag{22}$$

is satisfied. If such $\check{g}_{k(n)}^{(n)}$ does not exist due to small $\frac{2\hat{g}_{k(n)}^{(n)}}{\sigma_h^2}$ (or, equivalently, large channel uncertainty), it is prudent to choose $a_{k(n)}^{(n)} = 0$ and find $b_{k(n)}^{(n)}$ from

$$Q\left(\sqrt{\frac{2\hat{g}_{k(n)}^{(n)}}{\sigma_h^2}}, \sqrt{\frac{2b_{k(n)}^{(n)}}{\sigma_h^2}}\right) = 1 - \delta^{\frac{1}{N}}. \tag{23}$$

It is worth noting that even if the channel *p.d.f.*'s are known to be either exponential or non-central Chi-square, it is not straightforward to express the *p.d.f.* of I in closed form [18], [19], which underlines the usefulness of the proposed approach.

Remark 1: It can be checked that when the channel estimation error vanishes, the approximated chance-constraints (9)–(11) fall back to the desired deterministic interference constraint, namely, $\sum_{n\in\mathcal{N}} \hat{g}_{k(n)}^{(n)} p^{(n)} \le I_{\max}$. To see this, first note that $\check{g}_{k(n)}^{(n)} \to 0$ as $\sigma_h^2 \to 0$ from (22). Thus, $\alpha_{k(n)}^{(n)}$ and $\beta_{k(n)}^{(n)}$ approach 0 and $\hat{g}_{k(n)}^{(n)}$, respectively. Substituting these into (9)–(11) yields the desired result.

IV. Resource Allocation Algorithms

The OFDMA RA problems with separable structure can be tackled efficiently in the dual domain. In this approach, the overall problem is divided into multiple smaller per-subcarrier subproblems, which can be solved separately, coordinated by the range multipliers. Moreover, it can be shown that the duality gap vanishes as the number of subcarriers increases. This approach has been widely applied to the RA problems for multi-carrier systems [12], [15], [20].

The ℓ_1-approximate problem is clearly separable in n. As for the ℓ_∞-approximate problem, by introducing auxiliary variables $\mathbf{u} \triangleq [u_1, \ldots, u_N]^T$, the following separable constraints can be shown to be equivalent to (10).

$$\sum_{n=1}^{N} \gamma_{k(n)}^{(n)} p^{(n)} + \sqrt{2\log\frac{1}{\epsilon}} \sum_{n=1}^{N} u_n \leq I_{\max} \quad (24)$$

$$\sqrt{N}\sigma_{k(n)}^{(n)} \alpha_{k(n)}^{(n)} p^{(n)} \leq \sum_{n'=1}^{N} u_{n'}, \quad n \in \mathcal{N} \quad (25)$$

This can be seen from the following argument. Suppose that (25) is slack for all $n \in \mathcal{N}$ at the optimum. Then, the sum $\sum_{n'=1}^{N} u_{n'}$ can be decreased by small amount without any penalty in the objective. However, this makes (24) less tight, leading to net increase in the objective (unless the power constraint (2) is already tight), which contradicts the initial assumption of being at the optimum.

We continue the derivation of the RA algorithm using (24) and (25). The case of the ℓ_1-approximation will be discussed briefly in Sec. IV-B.

A. Algorithm for the ℓ_∞-Approximate Problem

Introducing dual variables $\boldsymbol{\lambda} \triangleq [\lambda_1, \lambda_2, \ldots, \lambda_N]^T \succeq \mathbf{0}$, $\boldsymbol{\mu} \triangleq [\mu_1, \mu_2, \ldots, \mu_K]^T \succeq \mathbf{0}$ and $\nu \geq 0$ to relax (25), (2), and (24), respectively, one can write the Lagrangian as

$$\begin{aligned} L(\mathbf{p}, \mathbf{u}; \boldsymbol{\lambda}, \boldsymbol{\mu}, \nu) = & \sum_{n \in \mathcal{N}} w_{k(n)} \left[\log\left(1 + \tilde{g}_k^{(n)} p^{(n)}\right) \right. \\ & - \left(\nu\gamma_{k(n)}^{(n)} + \mu_{k(n)} + \lambda_n \sqrt{N}\sigma_{k(n)}^{(n)} \alpha_{k(n)}^{(n)}\right) p^{(n)} \\ & \left. + \left(\sum_{n' \in \mathcal{N}} \lambda_{n'} - \nu\sqrt{2\log\frac{1}{\epsilon}}\right) u_n \right] + \nu I_{\max} \\ & + \sum_{k \in \mathcal{K}} \mu_k P_{k,\max} \end{aligned} \quad (26)$$

Therefore, the dual function is

$$\begin{aligned} D(\boldsymbol{\lambda}, \boldsymbol{\mu}) = & \sup_{\mathbf{0} \preceq \mathbf{p} \preceq \mathbf{P}_{\max}, \mathbf{u}, \mathbf{k} \in \mathcal{K}^N} L(\mathbf{p}, \mathbf{u}; \boldsymbol{\lambda}, \boldsymbol{\mu}, \nu) \\ = & \sup_{\substack{\mathbf{0} \preceq \mathbf{p} \preceq \mathbf{P}_{\max} \\ \mathbf{k} \in \mathcal{K}^N}} \sum_{n \in \mathcal{N}} L_n(p^{(n)}, k(n)) + \nu I_{\max} \\ & + \sum_{k \in \mathcal{K}} \mu_k P_{k,\max} \end{aligned} \quad (27)$$

where

$$L_n(p^{(n)}, k) \triangleq w_k \log\left(1 + \tilde{g}_k^{(n)} p^{(n)}\right) - t_k^{(n)} p^{(n)} \quad (28)$$

$$t_k^{(n)} \triangleq \nu\gamma_k^{(n)} + \mu_k + \lambda_n \sqrt{N}\sigma_k^{(n)} \alpha_k^{(n)} \quad (29)$$

and $\nu := \left(2\log\frac{1}{\epsilon}\right)^{-\frac{1}{2}} \sum_{n' \in \mathcal{N}} \lambda_{n'}$. The dual problem is thus

$$\inf_{\boldsymbol{\lambda} \succeq \mathbf{0}, \boldsymbol{\mu} \succeq \mathbf{0}} D(\boldsymbol{\lambda}, \boldsymbol{\mu}) \ . \quad (30)$$

It can be seen from (27) that the optimization can be decoupled to per-tone problems given by

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

If $k(n) = k$, the optimal power loading $p^{*(n)}[k]$ can be shown to be

$$p^{*(n)}[k] = \left[\frac{w_k}{t_k^{(n)}} - \frac{1}{\tilde{g}_k^{(n)}}\right]_0^{P_{\max}^{(n)}}, \quad n \in \mathcal{N} \quad (32)$$

where $[\cdot]_a^b \triangleq \min\{\max\{0, a\}, b\}$. The optimal user allocation $\mathbf{k}^*$ is then given by

$$k^*(n) \in \arg\max_{k \in \mathcal{K}} L_n(p^{*(n)}[k], k), \quad n \in \mathcal{N} \quad (33)$$

and the optimal power loading by

$$p^{*(n)} = p^{*(n)}[k^*(n)], \quad n \in \mathcal{N} \ . \quad (34)$$

The dual problem (30) can be solved using, e.g., the subgradient method, or the ellipsoid method, which require the subgradient of $D(\cdot)$ w.r.t. $[\boldsymbol{\mu}^T \boldsymbol{\lambda}^T]^T$. One such subgradient is

$$-\begin{bmatrix} \sum_{n \in \mathcal{N}: k(n)=1} p^{(n)} - P_{1,\max} \\ \vdots \\ \sum_{n \in \mathcal{N}: k(n)=K} p^{(n)} - P_{K,\max} \\ \frac{-I_{\max} + \sum_{n \in \mathcal{N}} \gamma_{k(n)}^{(n)} p^{(n)}}{\sqrt{2\log\frac{1}{\epsilon}}} + \sqrt{N}\sigma_{k(n)}^{(1)} \alpha_{k(n)}^{(1)} p^{(1)} \\ \vdots \\ \frac{-I_{\max} + \sum_{n \in \mathcal{N}} \gamma_{k(n)}^{(n)} p^{(n)}}{\sqrt{2\log\frac{1}{\epsilon}}} + \sqrt{N}\sigma_{k(N)}^{(N)} \alpha_{k(N)}^{(N)} p^{(N)} \end{bmatrix}. \quad (35)$$

The overall RA algorithm based on the ellipsoid method is described in the following.

Algorithm 1:

1: Initialize $\mathbf{\Sigma}$ and $\boldsymbol{\theta} \triangleq [\boldsymbol{\mu}^T \ \boldsymbol{\lambda}^T]^T$;
and set $V = N + K$ for the ℓ_∞-approximation.
[Alternatively, $\boldsymbol{\theta} \triangleq [\boldsymbol{\mu}^T \ \nu]^T$, and set $V = K + 1$ for the ℓ_1-approximation].
Set tolerance τ
2: Repeat
3: If $\boldsymbol{\theta} < 0$ for some entries $i \in \mathcal{I}$
set $\mathbf{d} = \sum_{i \in \mathcal{I}} \mathbf{e}_i$ ($\mathbf{e}_i$ is the i-th canonical basis)
4: Otherwise:
5: Find $\mathbf{k}^*$ and $\mathbf{p}^*$ from (32)–(33)
[Alternatively use (36)–(37)]
6: Set $\mathbf{d}$ as (35) [Alternatively, use (38)]
7: If $\sqrt{\mathbf{d}^T \mathbf{\Sigma} \mathbf{d}} < \tau$, stop.
8: Perform the ellipsoid update:
9: $\mathbf{d} \leftarrow \mathbf{d}/\sqrt{\mathbf{d}^T \mathbf{\Sigma} \mathbf{d}}$
10: $\boldsymbol{\theta} \leftarrow \boldsymbol{\theta} - \mathbf{\Sigma}\mathbf{d}/(V+1)$
11: $\mathbf{\Sigma} \leftarrow \frac{V^2}{V^2-1}\left(\mathbf{\Sigma} - \frac{2}{V+1}\mathbf{\Sigma}\mathbf{d}\mathbf{d}^T\mathbf{\Sigma}\right)$

B. Algorithm for the ℓ_1-Approximate Problem

The ℓ_1-approximate problem can be similarly solved by the dual method. Introduce $(K+1)$ dual variables $\boldsymbol{\mu}$ and ν to relax (2) and (11), respectively, and let $s_k^{(n)} \triangleq \nu\left(\gamma_k^{(n)} + \sqrt{2\log\frac{1}{\epsilon}}\sigma_k^{(n)}\alpha_k^{(n)}\right) + \mu_k$. Then, one can obtain the optimal power loading as

$$p^{*(n)}[k] = \left[\frac{w_k}{s_k^{(n)}} - \frac{1}{\tilde{g}_k^{(n)}}\right]_0^{P_{\max}^{(n)}}, \quad n \in \mathcal{N} \tag{36}$$

and the optimal user allocation $\mathbf{k}^*$ as

$$k^*(n) \in \arg\max_{k\in\mathcal{K}} \log(1+\tilde{g}_k^{(n)}p^{(n)}) - s_k^{(n)}p^{*(n)}[k], \quad n \in \mathcal{N}. \tag{37}$$

A subgradient of the dual function w.r.t. $[\boldsymbol{\mu}^T \ \nu]^T$ is given by

$$-\begin{bmatrix} \sum_{n\in\mathcal{N}:k(n)=1} p^{(n)} - P_{1,\max} \\ \vdots \\ \sum_{n\in\mathcal{N}:k(n)=K} p^{(n)} - P_{K,\max} \\ -I_{\max} + \sum_{n\in\mathcal{N}} \left(\gamma_{k(n)}^{(n)} + \sqrt{2\log\frac{1}{\epsilon}}\sigma_{k(n)}^{(n)}\alpha_{k(n)}^{(n)}\right)p^{(n)} \end{bmatrix}. \tag{38}$$

The overall algorithm is implemented by following the alternative steps in the brackets in Algorithm 1 presented in Sec. IV-A.

Remark 2: A discussion on the complexity order of the proposed algorithms is in order. One must consider the total number of operations per iteration multiplied by the number of iterations required for convergence. The number of iterations needed to obtain an ϵ-optimal solution using the ellipsoid method with V variables is $\mathcal{O}\left(V^2 \log\frac{1}{\epsilon}\right)$ [21]. In the algorithm for the ℓ_∞-approximation problem, the number of variables is $(N+K)$. At each iteration, (32) and (33) need to be performed KN times, in addition to $(N+K)^2$ operations needed for the ellipsoid update. The overall complexity order is thus $\mathcal{O}((N+K)^4)$. On the other hand, the algorithm for the ℓ_1-approximate problem involves $(K+1)$ dual variables, leading to a complexity order of $\mathcal{O}(K^3N+K^4)$.

C. Suboptimal Algorithm

To obtain a performance benchmark, a suboptimal RA algorithm based on alternating maximization is derived. A suboptimal algorithm is a reasonable alternative used in a number of works such as [22]. Note that the algorithm still aims to solve the ℓ_2-, ℓ_∞-, and ℓ_1-approximate problems that result from applying the Bernstein method. The algorithm involves the following steps.

1) Initialize the power loading; e.g., set $p^{(n)} := \min\{\frac{1}{N}\min_k P_{k,\max}, P_{\max}^{(n)}\}$, $n \in \mathcal{N}$.
2) For each subcarrier $n \in \mathcal{N}$, find the user k^* that maximizes $w_k \log(1+\tilde{g}_k^{(n)}p^{(n)})$ and allocate subcarrier n to that user, i.e., $k(n) := k^*$
3) For fixed $\mathbf{k}$ obtained from step (2), optimize over $\mathbf{p}$ by solving the convex problem, using e.g., the interior-point method or the subgradient method
4) Repeat steps (2)–(3) until convergence.

It can be shown easily that the objective function does not decrease per iteration, thus guaranteeing convergence.

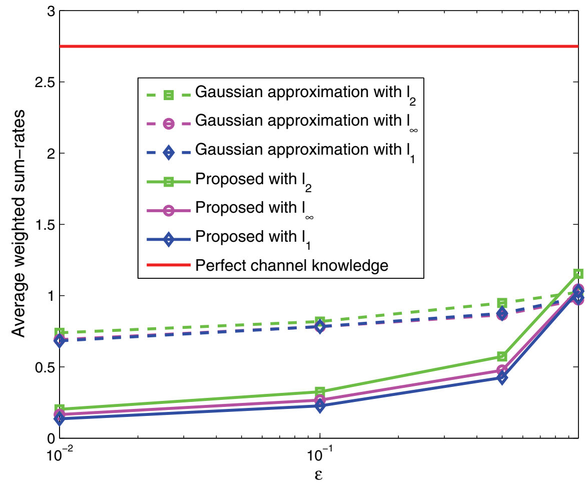


Fig. 2. Average weighted sum-rates without channel estimation.

D. Algorithms Based on the Central Limit Theorem

A yet another set of algorithms are derived here to approximately solve (1)–(3) without resorting to Bernstein's approximation when the mean $\hat{g}_k^{(n)}$ and the variance $(v_k^{(n)})^2$ of the channel gains $g_k^{(n)}$ are known. They will be useful in Sec. V to characterize the performance degradation due to the conservatism introduced by Bernstein's approximation. The idea is to apply the central limit theorem to the interference power $\sum_{n\in\mathcal{N}} g_{k(n)}^{(n)}p^{(n)}$ in (3) and approximate it as a Gaussian random variable with mean $\sum_{n\in\mathcal{N}} \hat{g}_{k(n)}^{(n)}p^{(n)}$ and variance $\sum_{n\in\mathcal{N}}(v_{k(n)}^{(n)}p^{(n)})^2$ (again assuming independence). Then, (3) can be expressed as

$$-I_{\max} + \sum_{n\in\mathcal{N}} \hat{g}_{k(n)}^{(n)}p^{(n)} + Q^{-1}(\epsilon)\left(\sum_{n\in\mathcal{N}}(v_{k(n)}^{(n)}p^{(n)})^2\right)^{\frac{1}{2}} \le 0 \tag{39}$$

Note the resemblance of (39) and (9). In particular, they both involve the ℓ_2-norm of vector $[v_{k(1)}^{(1)}p^{(1)}, \ldots, v_{k(N)}^{(N)}p^{(N)}]$. Thus, one can readily derive separable surrogates based on the ℓ_∞- and the ℓ_1-norms, given by [cf. (10) and (11)]

$$\sum_{n\in\mathcal{N}} \hat{g}_{k(n)}^{(n)}p^{(n)} + Q^{-1}(\epsilon)\max_{n\in\mathcal{N}} v_{k(n)}^{(n)}p^{(n)} \le I_{\max} \tag{40}$$

$$\sum_{n\in\mathcal{N}} \hat{g}_{k(n)}^{(n)}p^{(n)} + Q^{-1}(\epsilon)\sum_{n\in\mathcal{N}} |v_{k(n)}^{(n)}p^{(n)}| \le I_{\max} \tag{41}$$

respectively.

V. Numerical Tests

The proposed RA algorithms were tested via numerical experiments. The wideband channels $\tilde{g}_k^{(n)}$ and $g_k^{(n)}$ were simulated as 4-path Rayleigh fading channels. The pathloss exponent was set to $\alpha = 2$. Since the pathloss exponent in practice is usually larger than 2, this choice corresponds to a worst-case scenario in terms of PU interference. The parameters $\mu_k^{(n)+}$, $\mu_k^{(n)-}$ and $\sigma_k^{(n)}$ for the Bernstein approximations were chosen from [14, Table 1] using the known first-

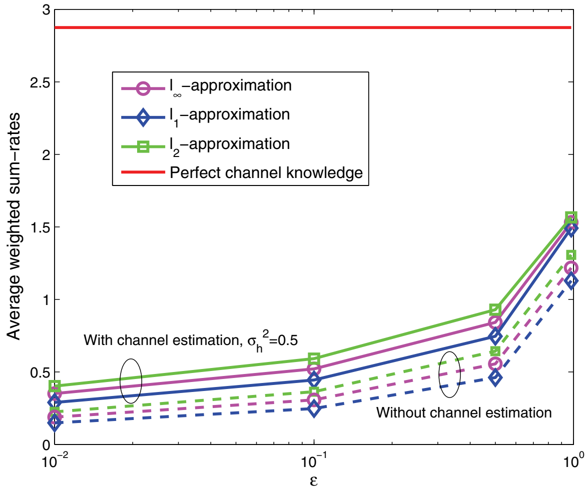


Fig. 3. Average weighted sum-rates with and without channel estimation.

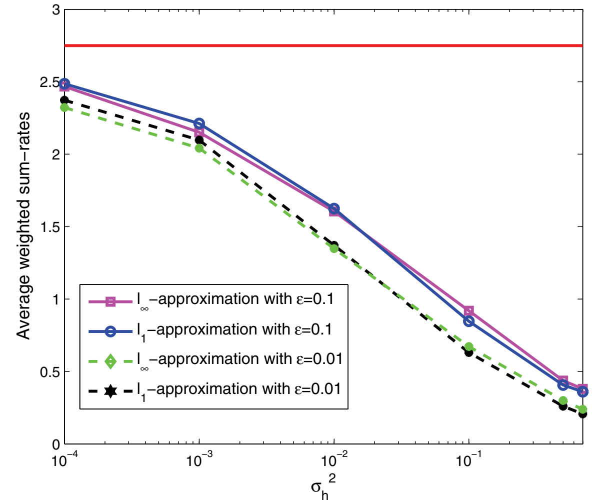


Fig. 4. Average weighted sum-rates versus σ_h^2.

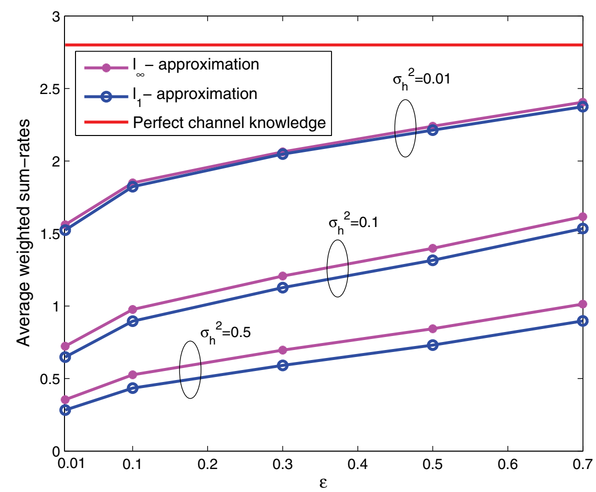


Fig. 5. Average weighted sum-rates versus ϵ.

and second-order moments of the truncated channel gains. Specifically, the values of $\mu_k^{(n)+}$ and $\mu_k^{(n)-}$ are set as (cf. (8))

$$\mu_k^{(n)+} = \mu_k^{(n)-} = E\left\{ \frac{g_k^{(n)} - \beta_k^{(n)}}{\alpha_k^{(n)}} \middle| a_k^{(n)} \leq g_k^{(n)} \leq b_k^{(n)} \right\}. \quad (42)$$

Then, upon defining

$$\eta_k^{(n)} \triangleq E\left\{ \left(\frac{g_k^{(n)} - \beta_k^{(n)}}{\alpha_k^{(n)}} \right)^2 \middle| a_k^{(n)} \leq g_k^{(n)} \leq b_k^{(n)} \right\}, \text{ and}$$

$$q_{\mu,\eta}(t) \triangleq \begin{cases} \log\left(\frac{(1-\mu)^2 \exp \frac{t(\mu-\eta^2)}{1-\mu} + (\eta^2-\mu^2)\exp(t)}{1-2\mu+\eta^2} \right), & \text{if } t \geq 0 \\ \log\left(\frac{(1+\mu)^2 \exp \frac{t(\mu+\eta^2)}{1+\mu} + (\eta^2-\mu^2)\exp(-t)}{1+2\mu+\eta^2} \right), & \text{otherwise} \end{cases} \quad (43)$$

the values of $\sigma_k^{(n)}$ are obtained from solving

$$\sigma_k^{(n)} = \min\left\{ c \geq 0 : q_{\mu_k^{(n)+}, \eta_k^{(n)}}(t) \leq \mu_k^{(n)+} t + \frac{c^2 t^2}{2}, \forall t \right\}. \quad (44)$$

Unless stated otherwise, the value of δ was set to $1-0.5\epsilon$, two CR users were experimented with $w_1 = 0.2$ and $w_2 = 0.8$, and the results were averaged over 80 realizations of $\tilde{g}_k^{(n)}$. The CR users were located equidistant from the CR BS throughout the tests.

Fig. 2 depicts the average weighted sum-rates for different values of ϵ without channel estimation for both Bernstein approximation-based and Gaussian approximation-based algorithms. The CR users are located equidistant from the PU. The solid line without markers is for the case when $\{g_k^{(n)}\}$ are perfectly known. In this case, the chance constraint (3) boils down to a deterministic constraint $\sum_{n\in\mathcal{N}} g_{k(n)}^{(n)} p^{(n)} < I_{\max}$, and thus the sum-rates do not depend on ϵ. The solid and dashed lines with markers correspond to the Bernstein approximation-based and the Gaussian approximation-based algorithms, respectively. The curves marked with squares, circles and diamonds were obtained by solving the ℓ_2-, ℓ_∞-, and ℓ_1-approximate problems, respectively. Note that in the case of the ℓ_2-approximate problem, exhaustive search over $\mathbf{k}$ was performed, as the dual decomposition technique could not be applied. It can be observed that the average weighted sum-rates increase as ϵ increases, since larger ϵ renders the chance constraint more lenient. Also, it can be seen that the curves corresponding to the ℓ_∞- or the ℓ_1-approximate problems are very close to the ones from the ℓ_2-approximate problem, underlining the usefulness of the dual decomposition-based low-complexity solutions. Interestingly, the ℓ_∞-approximation seems to yield a slightly better performance than the ℓ_1-approximation with Bernstein approximation, while the performance differences are quite negligible in the Gaussian approximation case. However, as discussed in Remark 2, the ℓ_1-approximation incurs lower computational complexity than the ℓ_∞-approximation.

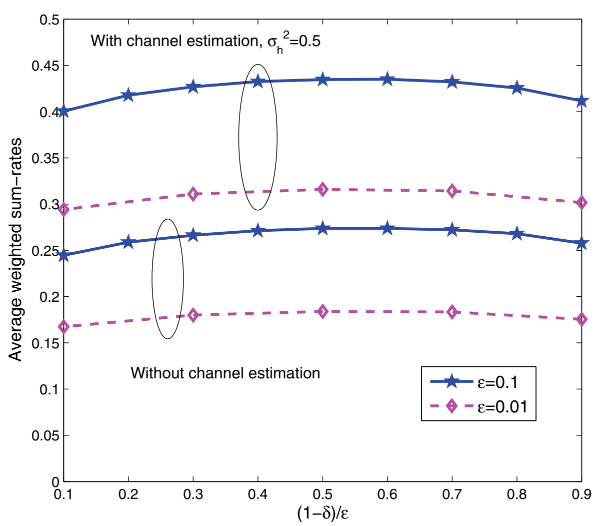
There is a large gap between the performance based on perfect channel knowledge and the robust RA performance. As can be seen, the Gaussian approximation yields much im-

TABLE I
AVERAGE RUN TIMES.

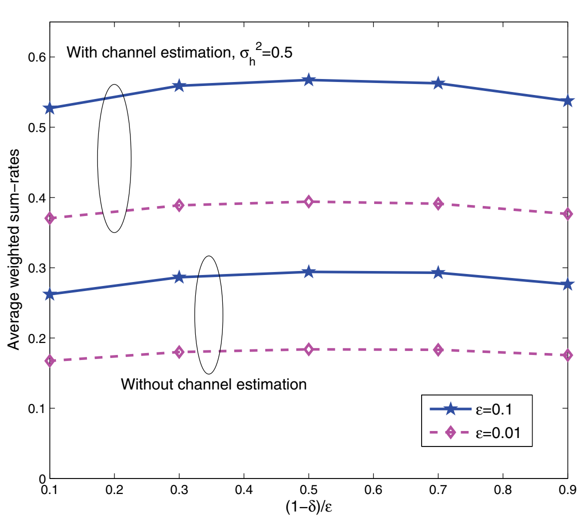
ϵ	0.01	0.1	0.5
Suboptimal algorithm	7.4782 sec	6.0122 sec	6.1220 sec
Dual method with ℓ_∞-approximation	3.7597 sec	3.6138 sec	3.3216 sec
Dual method with ℓ_1-approximation	0.7312 sec	0.7102 sec	0.7060 sec

TABLE II
SIMULATED VALUES OF $\Pr\{I < I_{\max}\}$ FOR BERNSTEIN AND THE GAUSSIAN APPROXIMATIONS.

$1-\epsilon$	0.9	0.5	0.3
Bernstein approx. with ℓ_2 without channel est.	0.9995	0.9600	0.8925
Bernstein approx. with ℓ_∞ without channel est.	1	0.9720	0.9526
Bernstein approx. with ℓ_1 without channel est.	0.9988	0.9600	0.9201
Bernstein approx. with ℓ_2 with channel est. ($\sigma_h^2 = 0.5$)	0.9967	0.9155	0.8234
Bernstein approx. with ℓ_∞ with channel est. ($\sigma_h^2 = 0.5$)	0.9998	0.9662	0.9226
Bernstein approx. with ℓ_1 with channel est. ($\sigma_h^2 = 0.5$)	0.9997	0.9655	0.8988
Gaussian approx. with ℓ_2 without channel est.	0.7785	0.6981	0.6718
Gaussian approx. with ℓ_∞ without channel est.	0.8310	0.7266	0.6982
Gaussian approx. with ℓ_1 without channel est.	0.8575	0.7602	0.7217



(a) CR users equidistant from the PU



(b) CR user 1 twice farther from the PU than CR user 2

Fig. 6. Average weighted sum-rates versus δ.

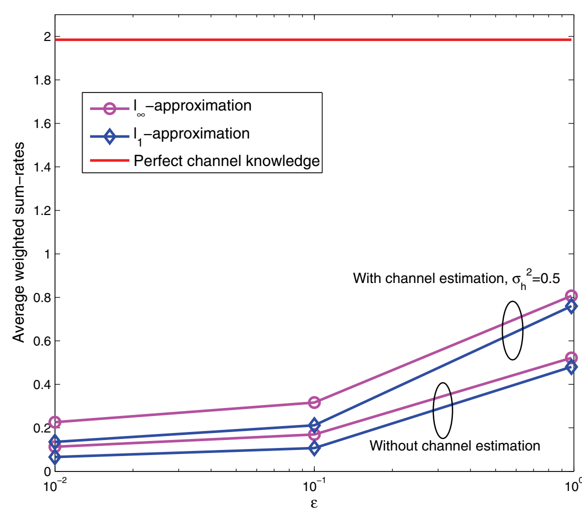


Fig. 7. Average weighted sum-rates for $K = 6$ and $N = 16$.

provement in the weighted sum-rates, compared to Bernstein approximation. However, even the Gaussian approximation can achieve only a small portion of what can be achieved under perfect channel knowledge. This illustrates that the major portion in the performance degradation may be attributed to the uncertainty in the channel, rather than the conservatism in Bernstein approximation.

The said gap can be alleviated by employing channel estimation, as shown in Fig. 3 for $\sigma_h^2 = 0.5$. The figure depicts the case where user 1 is twice farther from the PU than user 2. Among the curves obtained with channel estimation, the performance-complexity trade-offs similar to what appeared in the case without channel estimation are observed.

Fig. 4 shows the average weighted sum-rate performance as σ_h^2 is varied when both CR users are in the same distance from the PU. Two sets of curves corresponding to $\epsilon = 0.1$ and $\epsilon = 0.01$ are presented. It is seen that as the channel estimation accuracy improves (smaller σ_h^2), the performance of the RA algorithms also increases. As was discussed in Remark 1, the

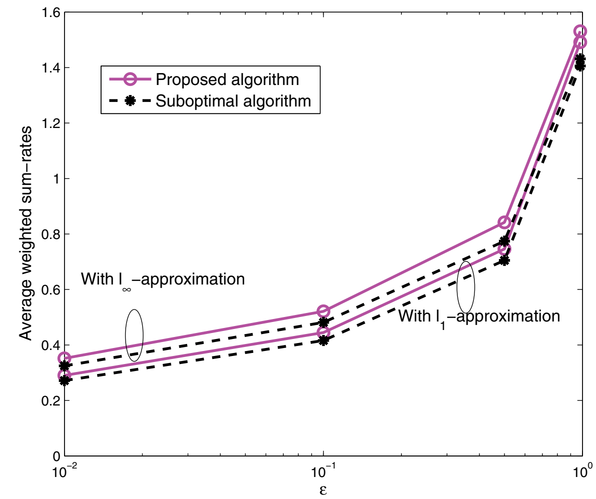


Fig. 8. Comparison with suboptimal algorithm.

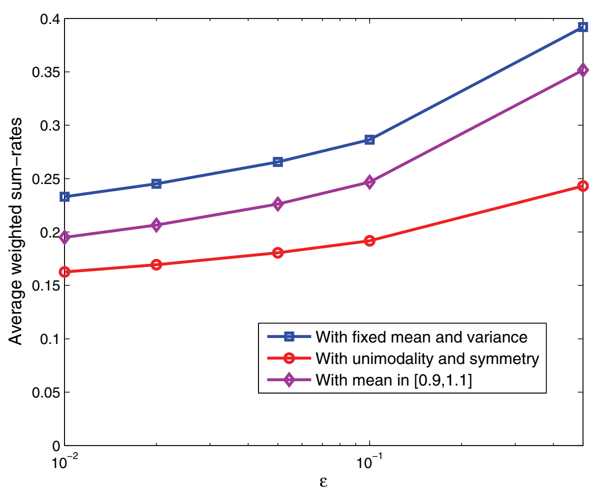


Fig. 9. The case of unknown p.d.f.

gap will eventually close as σ_h^2 vanishes. Similar trends are observed in Fig. 5, where the sum-rate performance versus ϵ is plotted with different values of σ_h^2, all with user 1 located twice farther away from the PU than user 2 is.

The sensitivity of the weighted sum-rate performance to the choice of δ is examined in Fig. 6 for $\epsilon = 0.01$ and 0.1 for two different network topologies. It is seen that the performance is maximized at about $\delta = 1 - 0.5\epsilon$ and is quite robust to the choice of δ. Interestingly, such an observation seems to hold regardless of the experimented network topologies and whether channel estimation is performed.

We also tested the case with $K = 6$ CR users with $\{w_k\}$ set to $[0.1, 0.2, 0.3, 0.2, 0.1, 0.1]$ using $N = 16$ subcarriers. Fig. 7 plots the weighted sum-rates averaged over 40 realizations of $\tilde{g}_k^{(n)}$. Due to the prohibitive complexity of exhaustive search required for the ℓ_2-approximate problem, only the results for the ℓ_∞- and the ℓ_1-approximate problems are reported. The overall trend remains unchanged from the two-user case.

To highlight the effectiveness of the dual decomposition-based approach, the performance of the suboptimal algorithm is compared to the proposed algorithms under channel estimation with $\sigma_h^2 = 0.5$ in Fig. 8 for the two-user case with CR user 1 twice farther from the PU than CR user 2.

The suboptimal algorithm is seen to be inferior to the proposed dual method-based algorithm both for the ℓ_∞- and the ℓ_1-approximate problems, although the gaps are small. On the other hand, the algorithms differ much in terms of computational complexity. Table I presents the average run times for the algorithms, where the experiments were done using a 2.13-GHz Intel CPU with 2 GB of RAM. It can be seen that the suboptimal algorithm takes much longer than the proposed algorithms until convergence. In fact, the suboptimal algorithm spent a lot of time to solve the convex subproblems, while the number of iterations were actually small.

To confirm that the interference constraint (3) is satisfied by the proposed methods, actual values of $\Pr\{I < I_{\max}\}$ are listed in Table II for the case with CR users equidistant from the PU, for different target values $1 - \epsilon$. It is clearly seen that the interference constraint is enforced conservatively with Bernstein approximation. (Although the achieved values of $\Pr\{I < I_{\max}\}$ are often quite off from the prescribed target $(1-\epsilon)$, this should not be interpreted as severe suboptimality of the proposed algorithms, because the performance must be ultimately gauged in terms of the weighted sum rates; see Fig. 2 and the associated discussion in this section.) However, it is seen that the Gaussian approximation approach does not always yield a solution feasible for the chance constraint, especially when ϵ is small. Therefore, the modest performance degradation as was seen in Fig. 2 may be thought of as the price to pay for guaranteed feasibility while maintaining tractability and accommodating a large class of distributions for uncertain parameters.

To illustrate the performance obtained when the channel *p.d.f.* is bounded but not known precisely, the average weighted sum-rate curves are plotted in Fig. 9 when the CR-to-PU channel gains belong to a bounded interval $[0, 7.38]$ for both CR users. The curve with circle markers represents the case where the support as well as the unimodality and the symmetry information are used. The curve with the diamonds corresponds to the case where the information that the mean lies in $[0.9, 1.1]$ was used in addition to the support information. The curve with the squares depicts the case of using the known mean and variance in addition to the support information. It can be seen that the performance is improved by employing more prior knowledge on the distribution.

VI. Conclusions

Weighted sum-rate maximization of an OFDMA CR uplink was considered, where the power loading and the user assignment over individual subcarriers are performed while ensuring that the interference power experienced at the PU location is less than a prescribed threshold. Since the channel gains between CR transmitters and PU receivers often cannot be estimated accurately, the PU interference constraint was cast as a chance constraint. As the resulting optimization problem is intractable, two layers of approximations were introduced. First, a convex conservative surrogate of the chance constraint was employed using Bernstein approximation, to bypass the need to analytically represent the chance constraint, even without precise knowledge of the distribution of uncertain channel gains. Secondly, due to the combinatorial complexity of searching for the optimal user assignment, approximation involving the ℓ_1- or the ℓ_∞-norms were employed so that the OFDMA RA problems possess separable structures, and can be tackled in the dual domain. Although such approximations indeed introduce conservatism, this is a side-effect often shared by a broad class of robust optimization approaches, and arguably constitutes the price paid to obtain guaranteed feasible solutions to chance-constrained OFDMA CR RA problems at an affordable complexity. Algorithms based on the dual decomposition method were developed for the cases with and without channel estimation. Numerical tests showed that the proposed algorithms outperformed a benchmark suboptimal algorithm in terms of both weighted sum-rate performance and computational complexity. It was also shown that the performance degradation due to the approximation introduced for enforcing the separability is rather insignificant.

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] 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.

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

[4] 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.

[5] 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.

[6] E. A. Gharavol, Y.-C. Liang, and K. Mouthaan, "Robust linear transceiver design in MIMO ad hoc cognitive radio networks with imperfect channel state information," *IEEE Trans. Wireless Commun.*, vol. 10, no. 5, pp. 1448–1457, May 2011.

[7] T. Al-Khasib, M. B. Shenouda, and L. Lampe, "Dynamic spectrum management for multiple-antenna cognitive radio systems: design with imperfect CSI," *IEEE Trans. Wireless Commun.*, vol. 10, no. 9, pp. 2850–2859, Sept. 2011.

[8] 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.

[9] 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.

[10] 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.

[11] W. W.-L. Li, Y. J. Zhang, A. M.-C. So, and M. Z. Win, "Slow adaptive OFDMA systems through chance constrained programming," *IEEE Trans. Signal Process.*, vol. 58, no. 7, pp. 3858–3869, July 2010.

[12] 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.

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

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

[15] 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.

[16] 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.

[17] P. K. Sen, "The mean-median-mode inequality and noncentral chi square distributions," *Sankhyā: The Indian J. Statistics*, vol. 51, pp. 106–114, 1989.

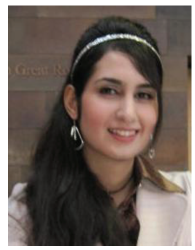
[18] A. Castaño-Martínez and F. López-Blázquez, "Distribution of a sum of weighted central chi-square variables," *Commun. Statistics - Theory and Methods*, vol. 34, no. 3, pp. 515–524, 2005.

[19] ——, "Distribution of a sum of weighted noncentral chi-square variables," *Test*, vol. 14, no. 2, pp. 397–415, 2005.

[20] 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.

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

[22] J. Huang, V. G. Subramanian, R. Agrawal, and R. A. Berry, "Downlink scheduling and resource allocation for OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 1, pp. 288–296, Jan. 2009.



Nasim Yahya Soltani 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 include optimal resource allocation for wireless communications, cognitive radios, robust optimization, array processing and smart power grids.

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. Since 2008, he has been with the Department of Electrical and Computer Engineering at the University of Minnesota, where he is currently a Research Assistant Professor. His research interests lie in applying signal processing and optimization techniques to various domains including wireless communication and networking, smart power grids, bio and social networks.

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 340 journal papers, 560 conference papers, 20 book chapters, two edited books and two research monographs (h-index 100). Current research focuses on compressive sensing, cognitive radios, cross-layer designs, wireless sensors, social and power grid 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.