

Erratum: Coping Uncertainty in Coexistence via Exploitation of Interference Threshold Violation

Shaoran Li

Virginia Tech

2040 Torgersen Hall,

Blacksburg, VA, 24061

shaoran@vt.edu

Published in Proc. MobiHoc 2019, pp. 71-80, Catania, Italy, July 2-5, 2019

There is a missing condition in Lemma 1, which is given as follows in Section 4:

- For any distribution of ξ_j (denote $f(u)$ as its probability density function) and a given interval $[a, b]$, $ab \leq 0$, there exists a discrete RV that ...

Which should be:

- For any distribution of ξ_j (denote $f(u)$ as its probability density function) and a given interval $[a, b]$, $ab \leq 0$, **s.t.** $a \leq \frac{\int_a^b u^2 f(u) du}{\int_a^b u f(u) du} \leq 0$, there exists a discrete RV that ...

This missing condition is always satisfied in the following derivations of Property 1 and Lemma 2. Therefore, the theoretical results in the published paper still holds.

Coping Uncertainty in Coexistence via Exploitation of Interference Threshold Violation

Shaoran Li[†] Yan Huang[†] Chengzhang Li[†] Brian A. Jalaian[‡] Y. Thomas Hou[†] Wenjing Lou[†]

[†] Virginia Polytechnic Institute and State University, Blacksburg, VA

[‡] U.S. Army Research Laboratory, Adelphi, MD

ABSTRACT

In underlay coexistence, secondary users (SUs) attempt to keep their interference to the primary users (PUs) under a threshold. Due to the absence of cooperation from the PUs, there exists much uncertainty at the SUs in terms of channel state information (CSI). An effective approach to cope such uncertainty is to introduce occasional interference threshold violation by the SUs, as long as such occasional violation can be tolerated by the PUs. This paper exploits this idea through a chance constrained programming (CCP) formulation, where the knowledge of uncertain CSI is limited to only the first and second order statistics rather than its complete distribution information. Our main contribution is the introduction of a novel and powerful technique, called *Exact Conic Reformulation* (ECR), to reformulate the intractable chance constraints. ECR guarantees an equivalent reformulation for linear chance constraints into deterministic conic constraints and does not suffer from the limitations associated with the state-of-the-art approach – Bernstein Approximation. Simulation results confirm that ECR offers significant performance improvement over Bernstein Approximation in uncorrelated channels and a competitive solution in correlated channels (where Bernstein Approximation is no longer applicable).

CCS CONCEPTS

• **Networks** → **Network resources allocation**; **Mobile networks**; *Network performance analysis*; Cognitive radios;

KEYWORDS

Underlay coexistence, Uncertainty, Scheduling

1 INTRODUCTION

Underlay coexistence is a key technique to improve spectrum efficiency by allowing simultaneous transmission of primary and secondary users (PUs and SUs) on the same spectrum [15]. The SUs must carefully control their transmission powers so that their

interference to each PU is under a threshold. An important feature (benefit) of underlay is that it does not require any cooperation (involving any hardware/software change) from the PUs to achieve coexistence, as the burden of successful coexistence with the PUs solely rests upon the SUs. Such feature is especially attractive for incremental deployment of new secondary networks over existing communication infrastructure, often referred to as primary networks. Due to this benefit, underlay coexistence has attracted many active efforts from the research community (see, e.g., [2, 10, 21, 23, 29]).

However, such benefits pose significant challenge for the SUs. Due to the absence of cooperation (feedback) from the PUs, accurate estimation of Channel State Information (CSI) is impossible. With such uncertainty in CSI, how to ensure the SUs limit their interference to the PUs under a threshold is a challenging problem. On the other hand, in many situations, we notice that occasional violations of interference threshold are not fatal to the PUs. First, to certain extent, the inherent channel coding is capable of recovering original transmitted symbols in the presence of interference [7]. Second, for applications such as video streaming and audio calls, human perception is quite tolerable to occasional errors (distortions) and there are numerous techniques to mitigate their impacts [28, 34].

Existing approaches to address CSI uncertainty can be classified into three categories: *stochastic programming*, *worst-case optimization* and *Chance Constrained Programming (CCP)*. Under stochastic programming, Random Variables (RVs) such as channel gains are assumed to have known distributions. For example, in [8], the wireless channel is assumed to have log-normal shadowing and Nakagami small-scale fading while in [32], it is assumed to have Rayleigh fading. However, in reality, many channels do not follow these simplified models and a blind assumption of these models could lead to misleading results (either overly optimistic or conservative). Even if we had accurate probability distributions for the RVs, the corresponding optimization problem could be extremely complicated, depending on the structure of the distributions.

Under worst-case optimization, the uncertainties are assumed to have some (known) upper and lower bounds and the constraints are enforced using the worst cases to achieve robustness. For example, in [37], the authors studied cognitive beamforming with a bounded ellipsoid for RVs (channel gains). In [30], the authors relaxed the interference constraint in underlay scenario to a linear constraint by defining a maximum estimation error. It is well known that such worst-case optimization is usually conservative with overly pessimistic performance. Further, many channel models are either

ACM acknowledges that this contribution was authored or co-authored by an employee, or contractor of the national government. As such, the Government retains a nonexclusive, royalty-free right to publish or reproduce this article, or to allow others to do so, for Government purposes only. Permission to make digital or hard copies for personal or classroom use is granted. Copies must bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. To copy otherwise, distribute, republish, or post, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

Mobihoc '19, July 2–5, 2019, Catania, Italy

© 2019 Association for Computing Machinery.

ACM ISBN 978-1-4503-6764-6/19/07...\$15.00

<https://doi.org/10.1145/3323679.3326505>

unbounded (e.g. Rayleigh fading) or an accurate estimation of the bounded set is difficult.

The third approach, *chance constrained programming (CCP)* [6], is a relatively new approach to address uncertainty in spectrum sharing [22, 24, 26, 31, 33]. In contrast to stochastic programming and worst-case optimization, CCP can be applied with any available knowledge of the unknown RVs, such as estimated mean, covariance and symmetricity, etc. To address uncertainty, CCP allows certain constraints to be violated and employs a control parameter called *risk level* to keep the violation probability below a limit. In this way, CCP explores a unique trade-off between performance objective and occasional constraint violations.

However, a major challenge in CCP is that chance constraints are usually mathematically intractable. A critical step in solving CCP is, therefore, to reformulate (substitute) the chance constraints with deterministic constraints and by doing so, to convert the CCP into a tractable optimization problem. The most primitive methods date back to Chebyshev and Markov inequalities, both of which introduce high relaxation errors [12, 17]. The state-of-the-art approach to perform this substitution (see, e.g., [22, 24, 26, 31, 33]) is the so-called *Bernstein Approximation* [27]. It performs such substitution by treating each RV separately (assuming they are independent and bounded) and solving an additional optimization problem for each RV to obtain the parameters used in the derived deterministic constraints. However, we find that there are a number of serious limitations with Bernstein Approximation. First, Bernstein Approximation explicitly requires that the RVs to be independent from each other. But this assumption does not always hold as correlations among uncertain RVs (e.g., CSI of sub-channels) are common and should be considered. Second, the performance of Bernstein Approximation depends heavily on the knowledge of the boundaries of uncertain RVs [27], which is hard to obtain in many cases. Finally, due to its generic nature, Bernstein Approximation does not explore the unique structure of linear CCP. As a consequence, its result tends to be rather conservative, as shown in our simulation results in Section 6.

In this paper, we study an underlay coexistence scenario where the PUs do not offer feedback to the SUs. Our goal is to maximize spectrum efficiency of picocells while keeping SUs' occasional interference threshold violation within a small probability. This scenario, in its simpler form (with one PU), was studied in [22, 31] following the Bernstein Approximation. In this paper, we introduce a novel technique called *Exact Conic Reformulation (ECR)* to address the underlying CCP. The proposed ECR allows us to handle more practical and general problem settings and to achieve better performance when compared to Bernstein Approximation. The main contributions of this paper are summarized as follows:

- To address channel uncertainty in underlay coexistence, we employ CCP but only rely on the first and second order statistics of the uncertain channel gains, which can be readily estimated and are quite accurate. By allowing occasional violation of interference threshold and keeping such violation under a target probability, we are able to exploit an optimal trade-off between spectrum efficiency and interference to the PUs.
- To reformulate the intractable chance constraints, we introduce ECR to offer mathematically exact conic reformulation and overcome the key limitations in the state-of-the-art approach (Bernstein Approximation). To the best of our knowledge, this is the first paper that has successfully addressed the limitations of Bernstein Approximation when it is used to study CCP problems in wireless networking and spectrum sharing in particular.
- We show that our solution (predicated on ECR) achieves higher spectrum efficiency when channel gains are independent and Bernstein Approximation is applicable. Specifically, our solution outperforms Bernstein Approximation by up to 60% (30% on average) higher spectrum efficiency. In the correlated scenario where Bernstein Approximation is no longer applicable, our proposed approach can still guarantee the violation probability while maximizing spectrum efficiency for the SUs.
- Our proposed approach is able to reap the full benefits of CCP in both general and practical settings thanks to our novel ECR technique. Through extensive simulations, we demonstrate the effectiveness of our approach under different settings of interference thresholds and channel models.

We organize this paper as follows. In Section 2, we introduce the system model and in Section 3, we formulate our problem. In Section 4, we present the novel ECR technique for CCP. In Section 5, we present the solution to the equivalent (reformulated) deterministic optimization problem. In Section 6, we present simulation results. Section 7 concludes the paper.

2 SYSTEM MODEL

Consider several picocells residing within a macrocell as shown in Fig. 1. An example of such scenario is that each picocell is installed as a set-up box inside a residential unit [4, 9]. Users connected with the macro base station (BS) are called PUs while users connected to the pico BSs are called SUs. We assume each picocell can use only a fraction of the spectrum allocated to the macrocell. To avoid the inter-cell interference between neighboring picocells, we assume adjacent picocells use different frequency bands (as shown in different colors of footprint in Fig. 1). This scheme is known as "fractional frequency reuse" in the literature [5, 11, 20].

In the underlay coexistence paradigm [15], the PUs are unaware of the presence of the SUs. The SUs take the sole responsibility of keeping their transmissions not to disrupt the normal operation of nearby PUs. Since the uplink problem (transmission from multiple SUs to the pico BS in a picocell) is harder than the downlink problem (only the pico BS in a picocell is transmitting), we focus on the (harder) uplink problem in this paper.

To keep the interference from the SUs to the PUs under control, each SU performs channel sensing before transmission. During channel sensing, a SU estimates both the channel conditions (for its own transmission¹) as well as those of nearby active PUs based on known signals (e.g., pilots) and channel reciprocity property [36]. Then the pico BSs will collect these CSI from the SUs through a dedicated control channel and find optimal solution for scheduling

¹For instance, a SU may send pilots to its connected pico BS for uplink channel training.

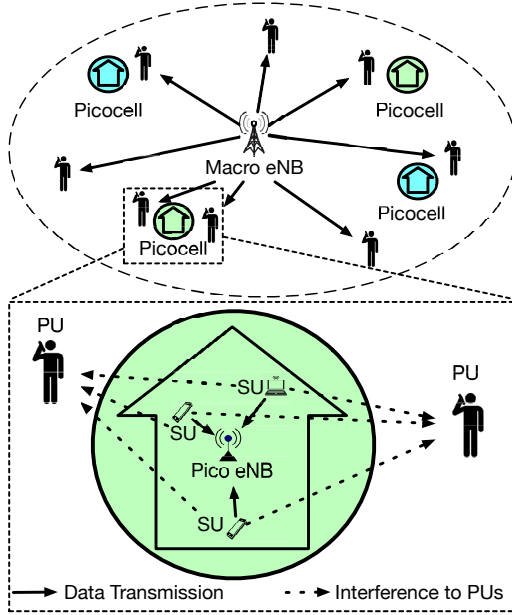


Figure 1: Network topology: multiple picocells within a macrocell.

(in spectrum and/or time) and power control. The goal is to maximize spectrum efficiency while keeping the aggregate interference from the SUs to each nearby PU below a threshold (see Fig. 1). The optimal solution for scheduling and power control will be sent to the SUs by the pico BSs and then the SUs can execute their uplink transmission based on this solution. Since neighboring picocells operate independently on non-overlapping frequency bands, we only need to study our problem in one picocell.

Consider one picocell (the lower portion of Fig. 1) with several nearby PUs. To control the aggregate interference to each PU, the CSI from the SUs to each PU is needed. Since there is no feedback from the PUs to the SUs, the SUs can only estimate CSI to the PUs unilaterally based on known signals from the PUs (e.g., pilot signal to the macro BS) and channel reciprocity property. As a result, channel gains from the SUs to the PUs, the key component for controlling the transmission powers of the SUs, can be characterized as RVs at best, rather than deterministic values. To differentiate different PUs, a SU can exploit the orthogonality in pilots as well as location techniques based on existing spectrum sensing algorithms [13, 35].

In our setting, we assume the PUs can tolerate occasional threshold violation as long as the probability of such violation is small. For practical purpose, such occasional violation is tolerable, as discussed in Section 1. As we shall see in the next section, such tolerance can be formulated as chance constraints under CCP.

3 MATHEMATICAL MODELING AND FORMULATION

We are interested in maximizing spectrum efficiency for the SUs in a picocell while keeping their violation of interference threshold to each nearby PU under a target small probability. Denote N as

the number of SUs in the picocell and J as the number of nearby PUs. Suppose the transmission bandwidth in the picocell is further divided into M sub-channels. Following cellular terminology, we call each sub-channel over one transmission time interval (TTI) as a resource block (RB). Due to multipath, channel gains vary over time and differ among different sub-channels (with perhaps some level of correlation).

For each TTI, a scheduling algorithm needs to allocate the available RBs among the SUs for uplink transmission. A popular scheduling objective is to achieve long-term proportional fair (PF) among SUs' throughput [25]. This is equivalent to maximizing a weighted sum of throughput in each TTI, with the weight of each SU being updated at the beginning of each TTI based on their long-term data rates. This is equivalent to assuming that the weights are given for the current TTI and we need to maximize the weighted sum rate for all SUs in the picocell.

Denote x_{iB}^m as a binary variable to indicate whether SU i will transmit to the pico BS on RB m , i.e.,

$$x_{iB}^m = \begin{cases} 1 & \text{if SU } i \text{ will transmit to the pico BS on RB } m, \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

Under single user Orthogonal Frequency-Division Multiple Access (OFDMA), each RB can be assigned to at most one SU. We have

$$\sum_{i \in N} x_{iB}^m \leq 1 \quad (m \in M), \quad (2)$$

where M is the set $\{1, 2, \dots, M\}$.

Denote p_{iB}^m as the transmission power from SU i to the pico BS on RB m . Denote P_{iB}^{\max} as the maximum power when SU i transmits to the pico BS over all RBs. Then we have

$$0 \leq p_{iB}^m \leq x_{iB}^m P_{iB}^{\max} \quad (i \in N, m \in M), \quad (3)$$

and

$$\sum_{m \in M} p_{iB}^m \leq P_{iB}^{\max} \quad (i \in N), \quad (4)$$

where $N = \{1, 2, \dots, N\}$. Constraints (4) represent the internal power control due to the SUs' equipment.

Assume each RB occupies the same bandwidth, which we normalize to 1 unit. Denote c_{iB}^m as SU i 's normalized capacity to the pico BS on RB m (w.r.t. normalized RB bandwidth). Then we have:

$$c_{iB}^m = \log_2(1 + h_{iB}^m p_{iB}^m) \quad (i \in N, m \in M), \quad (5)$$

where h_{iB}^m is the overall channel gain of SU i toward the pico BS on RB m , including both interference from the macro BS and thermal noise at the pico BS.

Denote g_{ij}^m as the channel gain from SU i to PU j on RB m and I_j as the interference threshold for PU j ($j \in \mathcal{J}$). Under CCP, the aggregate interference from the SUs to PU j is allowed to occasionally violate I_j but must be below a target (small) probability. This behavior, in its complementary form, can be formulated by the following chance constraints:

$$\mathbb{P} \left\{ \sum_{i \in N} \sum_{m \in M} g_{ij}^m p_{iB}^m \leq I_j \right\} \geq 1 - \epsilon_j \quad (j \in \mathcal{J}), \quad (6)$$

where $\mathcal{J} = \{1, 2, \dots, J\}$. $\mathbb{P}\{\cdot\}$ denotes the probability function and ϵ_j is called *risk level*. Note that proper power control at the SUs is the key to meet these chance constraints. This risk level

ϵ_j could vary over a wide range (e.g. 0.01 to 0.5) depending on the application of PU j . A higher ϵ_j means a larger tolerance to violation of interference threshold (and corresponding to a larger optimization space) and hence higher spectrum efficiency.

Per our earlier discussion, in (6), channel gains g_{ij}^m 's are modeled as RVs with unknown distributions. In this paper, we assume their mean and covariance can be obtained via online estimation. Specifically, whenever the SUs overhear the signals transmitted by PU j , the SUs can estimate the channel condition in current TTI based on channel reciprocity. But for those TTIs that PU j is silent, the channel state information becomes quickly outdated. However, the estimated mean and covariance are relatively time-invariant and remain valid. It is reasonable to assume such statistics are up-to-date at the SUs through continuous tracking of the mean and covariance over time. Thus, it is more prudent and practical to use the first and second order statistics (mean and covariance) when modeling RVs g_{ij}^m 's for our problem.

Denote w_i as the weight of SU i in current TTI. Then our problem can be formulated as follows:

$$\begin{aligned}
 \text{(P1)} \quad & \max_{x_{iB}^m, p_{iB}^m} \sum_{i \in \mathcal{N}} \sum_{m \in \mathcal{M}} w_i c_{iB}^m \\
 \text{s.t.} \quad & \text{RB allocations (2)} \\
 & \text{Transmission powers (3)} \\
 & \text{Internal power control (4)} \\
 & \text{Calculations of capacity (5)} \\
 & \text{External power control (6)} \\
 & x_{iB}^m \in \{0, 1\}, p_{iB}^m \geq 0
 \end{aligned}$$

Clearly, the main challenge in this optimization problem lies in chance constraints (6). Although we have the first and second order statistics of g_{ij}^m 's, we do not have the knowledge of their distributions.² For the same first and second order statistics, there are an infinite number of corresponding distributions. Since it is impossible to enumerate all possible distributions for constraints (6), P1 is intractable.

4 A NOVEL REFORMULATION OF CHANCE CONSTRAINTS

In this section, we present a novel approach to perform *exact reformulation* of (6) as *second-order cones*. By *Exact Conic Reformulation* (ECR), we mean that the newly derived deterministic constraints (with tractable conic formulations) are mathematically equivalent to the original chance constraints. We show that for the worst-case distribution, the supremum of threshold violation probability is exactly the risk level ϵ_j , while for all other distributions, the threshold violation probability is smaller than ϵ_j [3]. In addition, the derived deterministic constraints belong to second order cones which are guaranteed to be convex.

Since the J chance constraints in (6) are independent from each other, we can perform ECR for each chance constraint w.r.t. PU j .

²Even if we had knowledge of the exact distributions, it remains unclear if problem P1 could be solved. This would heavily depend on the underlying probability density functions.

For ease of exposition, we rewrite the j -th constraint in (6) as

$$\mathbb{P} \left\{ \mathbf{g}_j^T \mathbf{p} > I_j \right\} \leq \epsilon_j, \quad (7)$$

where superscript T denotes transposition. \mathbf{p} is a $MN \times 1$ column vector, i.e.,

$$\mathbf{p} = [p_{1B}^1, \dots, p_{1B}^M, p_{2B}^1, \dots, p_{2B}^M, \dots, p_{NB}^1, \dots, p_{NB}^M]^T, \quad (8)$$

which represents MN transmission powers from the SUs (over all RBs). \mathbf{g}_j is a $MN \times 1$ random column vector given as

$$\mathbf{g}_j = [g_{1j}^1, \dots, g_{1j}^M, g_{2j}^1, \dots, g_{2j}^M, \dots, g_{Nj}^1, \dots, g_{Nj}^M]^T, \quad (9)$$

which represents MN random channel gains from the SUs (over all RBs) to PU j .

Denote $\bar{\mathbf{g}}_j$ (a $MN \times 1$ column vector) and \mathbf{R}_j (a $MN \times MN$ matrix) as the mean and covariance matrix of \mathbf{g}_j , i.e., $\mathbf{g}_j \sim (\bar{\mathbf{g}}_j, \mathbf{R}_j)$. Since constraint (7) is satisfied for \mathbf{g}_j under all distributions with $\mathbf{g}_j \sim (\bar{\mathbf{g}}_j, \mathbf{R}_j)$, we have

$$\sup_{\mathbf{g}_j \sim (\bar{\mathbf{g}}_j, \mathbf{R}_j)} \mathbb{P} \left\{ \mathbf{g}_j^T \mathbf{p} > I_j \right\} \leq \epsilon_j, \quad (10)$$

where "sup" is taken over all distributions for \mathbf{g}_j with mean $\bar{\mathbf{g}}_j$ and covariance \mathbf{R}_j .

Denote ξ_j as a scalar RV defined as $\xi_j = \mathbf{g}_j^T \mathbf{p} - \bar{\mathbf{g}}_j^T \mathbf{p}$. It is easy to see that ξ_j has mean 0 and variance $\mathbf{p}^T \mathbf{R}_j \mathbf{p}$, i.e., $\xi_j \sim (0, \mathbf{p}^T \mathbf{R}_j \mathbf{p})$. For ease of exposition, denote ϕ_j as $\phi_j = I_j - \bar{\mathbf{g}}_j^T \mathbf{p}$. With ξ_j and ϕ_j , we can rewrite (7) as following

$$\sup_{\xi_j \sim (0, \mathbf{p}^T \mathbf{R}_j \mathbf{p})} \mathbb{P} \left\{ \xi_j > \phi_j \right\} \leq \epsilon_j. \quad (11)$$

To perform an exact reformulation of chance constraint (11), we need to derive a closed form expression for the supremum of violation probability where ξ_j can take any form of distribution with mean 0 and variance $\mathbf{p}^T \mathbf{R}_j \mathbf{p}$. Then we can upper bound this closed form expression by ϵ_j . Since this closed form expression has no randomness, we have a deterministic constraint on the decision variables in \mathbf{p} .

To derive a closed form expression for the "sup" in (11), we need to find the worst-case distribution of ξ_j that maximizes the violation probability $\mathbb{P}\{\xi_j > \phi_j\}$. This is not a trivial problem, as the worst-case distribution of ξ_j may take any form. As a start, we present the following lemma to shrink the searching space of all forms of distributions.

LEMMA 1. *For any distribution of ξ_j (denote $f(u)$ as its probability density function) and a given interval $[a, b]$, $ab \leq 0$, there exists a discrete RV that has two elements at 0 and $c \in [a, b]$, $c \neq 0$ with probabilities \mathbb{P}_0 and \mathbb{P}_c respectively, such that³*

$$\int_a^b f(u) du = \mathbb{P}_c + \mathbb{P}_0 \quad (\text{Probability}), \quad (12a)$$

$$\int_a^b u f(u) du = c \mathbb{P}_c \quad (\text{Mean}), \quad (12b)$$

$$\int_a^b u^2 f(u) du = c^2 \mathbb{P}_c \quad (\text{Variance}). \quad (12c)$$

³Here we use the notation of a closed interval, but a or b can be infinite.

Proof Sketch. Obviously, $\mathbb{P}_c, \mathbb{P}_0$ and c can be solved based on the three equations in (12). Specifically, c and \mathbb{P}_c can be solved based on (12b) and (12c). Then \mathbb{P}_0 can be obtained from (12a) and \mathbb{P}_c . It is easy to verify that \mathbb{P}_0 and \mathbb{P}_c are valid probabilities (i.e. $\mathbb{P}_0 \geq 0, \mathbb{P}_c \geq 0, \mathbb{P}_0 + \mathbb{P}_c \leq 1$). Further, we can show that $c \in [a, b], c \neq 0$ based on the relation between mean $c\mathbb{P}_c$ and variance $c^2\mathbb{P}_c$.

Lemma 1 states that this newly constructed discrete RV preserves probability, mean and variance in $[a, b]$ with no influence on other intervals since $0 \in [a, b], c \in [a, b]$. Based on Lemma 1, we are able to explore some properties associated with the worst-case distribution, as stated in Property 1.

PROPERTY 1. (Worst-case Distribution of ξ_j) *The worst-case distribution of ξ_j has the following properties:*

- (1) If $\phi_j < 0$, then $\mathbb{P}\{\xi_j \leq \phi_j\} = 0$,
- (2) If $\phi_j \geq 0$, then $\mathbb{P}\{\xi_j > \Phi_j\} = 0$, for all $\Phi_j > \phi_j$.

PROOF. Our proof is based on contradictions.

(1) When $\phi_j < 0$, suppose ξ'_j is a worst-case distribution but $\mathbb{P}\{\xi'_j \leq \phi_j\} > 0$, then the violation probability $\mathbb{P}\{\xi'_j > \phi_j\} = 1 - \mathbb{P}\{\xi'_j \leq \phi_j\}$ must be smaller than 1. However, consider the following discrete distribution of ξ_j :

$$\mathbb{P}\left\{\xi_j = \frac{\phi_j}{2}\right\} = \frac{4\mathbf{p}^T \mathbf{R}_j \mathbf{p}}{4\mathbf{p}^T \mathbf{R}_j \mathbf{p} + \phi_j^2}, \quad (13)$$

$$\mathbb{P}\left\{\xi_j = -\frac{2\mathbf{p}^T \mathbf{R}_j \mathbf{p}}{\phi_j}\right\} = \frac{\phi_j^2}{4\mathbf{p}^T \mathbf{R}_j \mathbf{p} + \phi_j^2}.$$

Obviously, the above distribution is valid and the violation probability $\mathbb{P}\{\xi_j > \phi_j\} = 1$. It is larger than the violation probability from ξ'_j , which contradicts the assumption that ξ'_j is the worst-case distribution. Therefore, any worst-case distribution of ξ_j has probability 0 in interval $(-\infty, \phi_j]$, i.e., $\mathbb{P}\{\xi_j \leq \phi_j\} = 0$.

(2) When $\phi_j \geq 0$, suppose ξ'_j is a worst-case distribution but has an positive element at $b' > \phi_j$ with probability $\mathbb{P}_{b'}$ ⁴. We can construct another distribution that has an element at b with probability \mathbb{P}_b such that

$$b' > b > \phi_j, \mathbb{P}_b > \mathbb{P}_{b'}. \quad (14)$$

Based on Lemma 1, such distribution always exists with necessary changes in interval $(-\infty, \phi_j]$ to maintain the first and second order statistics. Consequently, the violation probability becomes higher, which contradicts the assumption that ξ'_j is the worst-case distribution. Thus, the worst-case distribution of ξ_j when $\phi_j \geq 0$ only has one element that is sufficiently close to ϕ_j as all of the possibilities of ξ_j in interval $(\phi_j, +\infty)$ are pushed to ϕ_j , i.e., $\mathbb{P}\{\xi_j > \Phi_j\} = 0$, for all $\Phi_j > \phi_j$. \square

Under Property 1, there may still exist many forms for the worst-case distribution, i.e., we may have many worst case distributions. However, these worst-case distributions all share the same closed form expression for the supremum of violation probability. Moreover, Lemma 1 shows that for the purpose of deriving closed form expression of the supremum in (11), we only need to consider a discrete distribution with a small number of elements. Therefore, we have the following result.

⁴This conclusion also holds for continuous RV by discretization.

LEMMA 2. *A closed form expression for the supremum of violation probability is given by:*

$$\sup_{\xi_j \sim (0, \mathbf{p}^T \mathbf{R}_j \mathbf{p})} \mathbb{P}\{\xi_j > \phi_j\} = \begin{cases} 1 & \phi_j < 0, \\ \frac{\mathbf{p}^T \mathbf{R}_j \mathbf{p}}{\phi_j^2 + \mathbf{p}^T \mathbf{R}_j \mathbf{p}} & \phi_j \geq 0. \end{cases} \quad (15)$$

PROOF. Consider the following two cases.

Case 1. If $\phi_j < 0$, the distribution in (13) already achieves the supremum at 1, which is a trivial case.

Case 2. If $\phi_j \geq 0$, based on Property 1, any worst-case distribution only has one element in interval $(\phi_j, +\infty)$ at b approaching ϕ_j with probability \mathbb{P}_b . In terms of the other interval $(-\infty, \phi_j]$, it can be characterized by two discrete elements at $a < 0$ and 0 with probabilities \mathbb{P}_a and \mathbb{P}_0 respectively based on Lemma 1. This conversion from a worst-case distribution to a three-element discrete distribution preserves the violation probability $\mathbb{P}\{\xi_j > \phi_j\}$, which is calculated as

$$\mathbb{P}\{\xi_j > \phi_j\} = \mathbb{P}_b. \quad (16)$$

We can derive the closed form supremum by the following optimization problem

$$\begin{aligned} \sup_{\mathbb{P}_a, \mathbb{P}_b, \mathbb{P}_0, a, b} \quad & \mathbb{P}_b \\ \text{s.t.} \quad & \mathbb{P}_a + \mathbb{P}_0 + \mathbb{P}_b = 1 \\ & a\mathbb{P}_a + b\mathbb{P}_b = 0 \\ & a^2\mathbb{P}_a + b^2\mathbb{P}_b = \mathbf{p}^T \mathbf{R}_j \mathbf{p} \\ & \mathbb{P}_a, \mathbb{P}_0, \mathbb{P}_b \geq 0, a \leq 0, b > \phi_j \end{aligned} \quad (17)$$

The optimization problem in (17) is easy to solve since there are only five decision variables. The optimal objective (supremum of violation probability) is $\frac{\mathbf{p}^T \mathbf{R}_j \mathbf{p}}{\phi_j^2 + \mathbf{p}^T \mathbf{R}_j \mathbf{p}}$.

Combining the above two cases, we have the results in (15). \square

We now need to ensure the closed form expression for the supremum of violation probability is upper bounded by ϵ_j as in (11). Based on Lemma 2, since $1 > \epsilon_j$, the case with $\phi_j < 0$ is infeasible. So we only need to consider the case when $\phi_j \geq 0$. That is, chance constraint (11) can be replaced by the following two constraints.

$$\phi_j \geq 0 \quad (18a)$$

$$\frac{\mathbf{p}^T \mathbf{R}_j \mathbf{p}}{\phi_j^2 + \mathbf{p}^T \mathbf{R}_j \mathbf{p}} \leq \epsilon_j \quad (18b)$$

Rewrite (18b) as

$$\frac{1 - \epsilon_j}{\epsilon_j} \cdot \mathbf{p}^T \mathbf{R}_j \mathbf{p} \leq \phi_j^2. \quad (19)$$

Taking the square root of both sides in (19) and considering (18a), we have

$$\sqrt{\frac{1 - \epsilon_j}{\epsilon_j}} \cdot \sqrt{\mathbf{p}^T \mathbf{R}_j \mathbf{p}} \leq \phi_j. \quad (20)$$

Note that (20) has already considered the non-negativity of ϕ_j in (18a). Substituting $\phi_j = I_j - \mathbf{g}_j^T \mathbf{p}$ into (20), we have the following main result.

THEOREM 1. (ECR) *With respect to the decision variables in \mathbf{p} , chance constraints (6) are equivalent to the following second order cones*

$$\sqrt{\frac{1-\epsilon_j}{\epsilon_j}} \sqrt{\mathbf{p}^T \mathbf{R}_j \mathbf{p} + \bar{\mathbf{g}}_j^T \mathbf{p}} \leq I_j \quad (j \in \mathcal{J}). \quad (21)$$

This is our *Exact Conic Reformulation* for chance constraints (6) where "exact" means that the supremum of violation probability equals to ϵ_j . Constraints (21) are deterministic constraints since the random interference channel gains g_{ij}^m 's are eliminated and the mean $\bar{\mathbf{g}}_j$ and covariance matrix \mathbf{R}_j are given constants.

Compared with state-of-the-art approach (Bernstein Approximation [22, 24, 26, 27, 31, 33]), ECR has no requirements for the RVs in terms of their independence and boundaries. Therefore, it is more general than Bernstein Approximation for linear chance constraints.

Replacing (6) by (21) in P1, we have a deterministic maximization problem stated as

$$\begin{aligned} \text{(P2)} \quad & \max_{x_{iB}^m, p_{iB}^m} \sum_{i \in \mathcal{N}} \sum_{m \in \mathcal{M}} w_i c_{iB}^m \\ \text{s.t.} \quad & \text{Constraints (2) – (5)} \\ & \text{External power control from ECR (21)} \\ & x_{iB}^m \in \{0, 1\}, p_{iB}^m \geq 0 \end{aligned}$$

5 SOLVING THE DETERMINISTIC OPTIMIZATION PROBLEM

P2 is a Mixed-Integer Non-Linear Program (MINLP), which is NP-hard in general. The main difficulties reside in the two nonlinear terms in constraints (5) and (21). In this section, we show how to linearize them.

5.1 Logarithm Functions

In (5), we have logarithm functions in calculations of capacity. We propose to employ one convex hull to relax each logarithm term $\log_2(1 + h_{iB}^m p_{iB}^m)$ [18]. Since we have a maximization problem, we only need to consider the series of linear constraints to upper bound the convex hull.

For each log term $\log_2(1 + h_{iB}^m p_{iB}^m)$, we break the interval for p_{iB}^m (i.e., $[0, p_{iB}^{\max}]$) into K equal-length sub-intervals, each with length $\frac{p_{iB}^{\max}}{K}$. Then we can upper bound the log function $\log_2(1 + h_{iB}^m p_{iB}^m)$ with $K + 1$ linear functions as follows:

$$c_{iB}^m \leq \frac{1}{\ln 2} \cdot \left(\frac{K h_{iB}^m p_{iB}^m}{K + k h_{iB}^m p_{iB}^{\max}} + \ln \left(1 + \frac{k h_{iB}^m p_{iB}^{\max}}{K} \right) - \frac{k h_{iB}^m p_{iB}^{\max}}{K + k h_{iB}^m p_{iB}^{\max}} \right) \quad (22)$$

$(k = 0, 1, \dots, K, i \in \mathcal{N}, m \in \mathcal{M})$

Constraints (22) are linear with c_{iB}^m and p_{iB}^m as variables. Clearly, the larger the K , the tighter the linear relaxation. In our numerical results in Section 6, we set $K = 50$ and the relaxation errors are already smaller than 0.1%.

5.2 Second Order Cones

Constraints (21) from ECR are second order cones. Since \mathbf{R}_j is the covariance matrix of \mathbf{g}_j , it is guaranteed to be positive semi-definite and symmetric. To relax the non-linear terms $\mathbf{p}^T \mathbf{R}_j \mathbf{p}$, we introduce a constant matrix \mathbf{V}_j as the square root of \mathbf{R}_j , i.e., $\mathbf{R}_j = \mathbf{V}_j^T \mathbf{V}_j$. \mathbf{V}_j can easily be calculated based on Cholesky decomposition.

With \mathbf{V}_j , we can rewrite constraints (21) as

$$\sqrt{\frac{1-\epsilon_j}{\epsilon_j}} \sqrt{(\mathbf{V}_j \mathbf{p})^T (\mathbf{V}_j \mathbf{p}) + \bar{\mathbf{g}}_j^T \mathbf{p}} \leq I_j \quad (j \in \mathcal{J}). \quad (23)$$

For any feasible \mathbf{p} , denote r_j (an integer) as an upper bound on the number of non-zero elements in the column vector $\mathbf{V}_j \mathbf{p}$. Since \mathbf{p} is an $MN \times 1$ vector, the maximum value for r_j is MN . For a tighter relaxation, we propose to set $r_j (\leq MN)$ based on the sparse structures of \mathbf{V}_j and \mathbf{p} in our problem.

Recall \mathbf{p} (i.e., $[p_{1B}^1, \dots, p_{1B}^M, p_{2B}^1, \dots, p_{2B}^M, \dots, p_{NB}^1, \dots, p_{NB}^M]^T$) has at most M non-zero elements because each RB can only be allocated to at most one SU. Further, \mathbf{V}_j , as the square root of the covariance matrix \mathbf{R}_j , is an $MN \times MN$ matrix. Since the channel gains from different SUs to PU j are usually independent, \mathbf{V}_j is a block diagonal matrix with the i -th block (denoted as \mathbf{V}_{ij}) corresponding to SU i ($i = 1, \dots, N$), i.e.,

$$\mathbf{V}_j = \begin{bmatrix} \mathbf{V}_{1j} & & & \\ & \mathbf{V}_{2j} & & \\ & & \dots & \\ & & & \mathbf{V}_{Nj} \end{bmatrix}. \quad (24)$$

Moreover, for SU i , the correlation between two RBs decreases as they are further apart. Define L_j as the maximum subcarrier spacing that has correlations, meaning that each RB is correlated with at most $2L_j$ neighboring RBs. Then, each block \mathbf{V}_{ij} is a $M \times M$ band matrix [16] that has the following form:

$$\mathbf{V}_{ij} = \begin{bmatrix} v_{ij}^{11} & \dots & v_{ij}^{1(L_j+1)} & & & \\ v_{ij}^{21} & v_{ij}^{22} & \dots & v_{ij}^{2(L_j+2)} & & \\ \dots & \dots & \dots & \dots & \dots & \\ & & & & v_{ij}^{M(M-L_j)} & \dots & v_{ij}^{MM} \end{bmatrix}. \quad (25)$$

Based on the sparsity of \mathbf{V}_j and \mathbf{p} , we can set r_j as the maximum non-zero elements in $\mathbf{V}_j \mathbf{p}$ before solving \mathbf{p} . Specifically, we consider two cases to calculate r_j :

(i) When $M \geq 2L_j + 1$, denote r_j^m as the number of non-zero elements in $\mathbf{V}_j \mathbf{p}$ generated by p_{iB}^m if $p_{iB}^m > 0$. r_j^m is calculated (based on m) by

$$r_j^m = \begin{cases} m + L_j & \text{if } m \leq L_j \\ (2L_j + 1) & \text{if } L_j + 1 \leq m \leq M - L_j \\ M - m + L_j + 1 & \text{if } m \geq M - L_j + 1 \end{cases}. \quad (26)$$

The three cases in (26) corresponds to the RBs in the left, middle and right part of the picocell's spectrum.

Since only one transmission power p_{iB}^m in $\{p_{1B}^m, p_{2B}^m, \dots, p_{NB}^m\}$ (the transmission powers on RB m for all SUs) is non-zero, we can

set r_j as

$$r_j = \sum_{m \in \mathcal{M}} r_j^m = M(2L_j + 1) - L_j(L_j + 1) \quad (27)$$

(ii) When $M < 2L_j + 1$, r_j can be set to $r_j = \min\{MN, M(2L_j + 1) - L_j(L_j + 1)\}$ since MN is always an upper bound of r_j .

Combining both case, we set r_j as

$$r_j = \min\{MN, M(2L_j + 1) - L_j(L_j + 1)\} \quad (j \in \mathcal{J}). \quad (28)$$

Denote $\|\mathbf{V}_j \mathbf{p}\|_\infty$ as the infinity norm of $\mathbf{V}_j \mathbf{p}$. Based on the fact that $\sqrt{(\mathbf{V}_j \mathbf{p})^T (\mathbf{V}_j \mathbf{p})} \leq \sqrt{r_j} \|\mathbf{V}_j \mathbf{p}\|_\infty$, we have the following relaxation of constraints (23):

$$\sqrt{\frac{1 - \epsilon_j}{\epsilon_j}} \sqrt{r_j} \|\mathbf{V}_j \mathbf{p}\|_\infty + \bar{\mathbf{g}}_j^T \mathbf{p} \leq I_j \quad (j \in \mathcal{J}). \quad (29)$$

Constraints (29) are linear with the decision variables in \mathbf{p} .

With the above linear relaxation, we have the following relaxed optimization problem.

$$\begin{aligned} \text{(P3)} \quad & \max_{x_{iB}^m, p_{iB}^m} \sum_{i \in \mathcal{N}} \sum_{m \in \mathcal{M}} w_i c_{iB}^m \\ \text{s.t.} \quad & \text{Constraints (2) - (4)} \\ & \text{Linearly relaxed capacity constraints (22)} \\ & \text{Linearly relaxed external power control (29)} \\ & x_{iB}^m \in \{0, 1\}, p_{iB}^m \geq 0 \end{aligned}$$

P3 belongs to Mixed Integer Linear Programming (MILP), which can be solved by commercial solvers such as CPLEX. For our problem size, the solution can be obtained on the order of second. For real time implementation of MILP, one can employ a recent breakthrough in real-time optimization based on GPU platform [19].

6 SIMULATION RESULTS

In this section, we conduct simulations to evaluate our proposed solution predicated on ECR. For performance evaluation, we mainly focus on spectrum efficiency (our objective value) and threshold violation probability. Our numerical study covers general and practical settings in consideration of different risk levels, interference thresholds and channel models.

6.1 Simulation Settings

For all topologies in the simulation study, we set the distance between the macro BS and the pico BS to 400 meters. The radius of the picocell is set to 40 meters. The transmission power of macro BS on each RB is set to 46 dBm and the thermal noise on each RB is assumed to be 1×10^{-7} mW. We assume each SU has a maximum transmission power of 20 dBm across all RBs.

The channels are generated based on ITU path loss model and small-scale fading. We consider two types of path loss models [1] as following:

(i) The path loss for the macro BS follows the ITU free-space path loss model as:

$$PL(\text{dB}) = 128.1 + 37.6 \times \log_{10}(d_{\text{macro}}), \quad (30)$$

where d_{macro} is the distance from the macro BS to the pico BS (in kilometers).

(ii) The path loss from the SUs to nearby PUs and the pico BS follows the ITU indoor path loss model given as

$$PL(\text{dB}) = 38 + 30 \times \log_{10}(d_{\text{pico}}), \quad (31)$$

where d_{pico} is the distance from a SU to a nearby PU or the pico BS (in m).

For small-scale fading, we consider Rayleigh, Rician, and Nakagami fading as specified per discussion below.

6.2 The Case of Independent Channels with Rayleigh Fading

We consider two types of network topology where the SUs are randomly distributed in the picocell or closer to the PUs (a stressful scenario). The channels are generated independently with Rayleigh fading as the small-scale fading.

For comparison, we also include the results from Bernstein Approximation and worst-case optimization in the same figure. Worst-case optimization uses the upper bounds of channel gains to remove uncertainty and consequently the chance constraints (6) becomes deterministic linear constraints. As for Bernstein Approximation, since our channel model is unbounded, we employ the same truncation method proposed in [22, 31] when the channel is Rayleigh fading. Then the reformulated MINLP is solved directly by CPLEX without any further relaxation.

In each specific topology, we perform 200 simulation runs and the results presented in this section are the average objective values for P1 based on the feasible solution obtained from its relaxed problem P3. For each simulation, we generate 10000 samples of the channel gains from each SU to each PU. The first and second order statistics from the 10000 samples are used in our solution. The 10000 samples are also used when calculating the actual threshold violation probability after the solution for each simulation is obtained. All of the optimization problems are solved by CPLEX using Branch & Cut for mixed-integer solution and the relative gap is less than 1%.

6.2.1 Randomly distributed SUs inside the picocell. We test two settings with one PU and three PUs separately. Six SUs are randomly distributed in the picocell. Assuming the pico BS is at the origin, the coordinates of SUs are $(-10.67, -19.01)$, $(32.47, -11.77)$, $(-0.23, 8.90)$, $(12.04, -5.03)$, $(28.89, -10.00)$ and $(-22.66, 5.34)$. The corresponding weights of SUs are 0.22, 0.09, 0.09, 0.23, 0.19 and 0.18. The coordinates and interference threshold I_j for each PU are given per discussion below. The number of RBs is set to 12 and remains the same for the whole simulation study.

(i) *The setting with one PU.* We first consider the case with one PU located at $(50, 0)$ with interference threshold $I = 3 \times 10^{-7}$ mW. Fig. 2 shows the performance of our solution as a function of risk level ϵ and interference threshold I . In Fig. 2(a), we find that the objective value of our solution monotonically increases with the risk level ϵ . In particular, we achieve 2.92 bps/Hz with a risk level of $\epsilon = 0.01$ and 5.09 bps/Hz with $\epsilon = 0.5$.

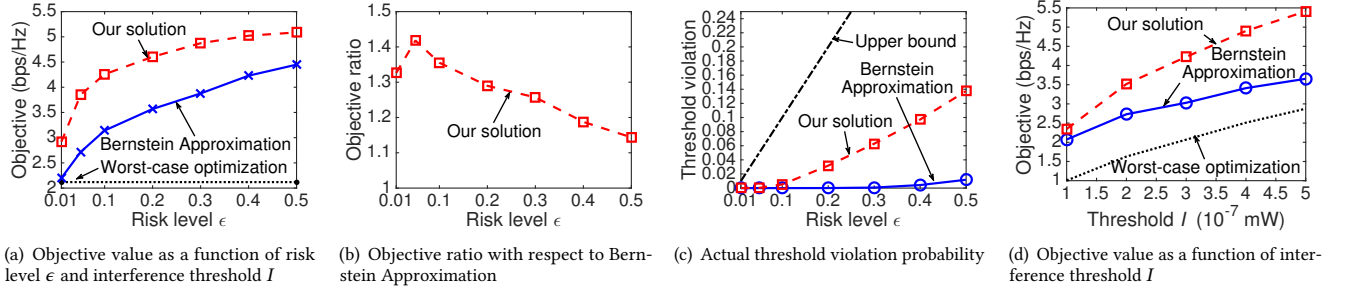


Figure 2: Performance of our solution as a function of risk level ϵ and interference threshold I with randomly distributed SUs

In Fig. 2(a), the objective value of worst-case optimization stays the same since it does not involve any risk level. Clearly, the performance of worst-case optimization is overly pessimistic due to zero tolerance of interference threshold violation.

The performance of our solution against Bernstein Approximation is shown in both Fig. 2(a) and Fig. 2(b). In particular, in Fig. 2(a), with a small risk level $\epsilon = 0.01$ (or 1%), the performance of Bernstein Approximation drops to 2.20 bps/Hz (almost the same with the one from worst-case optimization) while our solution achieves 2.92 bps/Hz. In Fig. 2(b), we plot the ratio of the objective value from our proposed solution over Bernstein Approximation. Our solution can achieve between 15% to 42% improvement over Bernstein Approximation.

Fig. 2(c) offers an in depth study of our solution and Bernstein Approximation in terms of actual threshold violation probability (i.e., percentage of samples where the interference threshold is actually violated). As shown in Fig. 2(c), the actual threshold violation probability from Bernstein Approximation stays below 0.02 even though the risk level is 0.5, which is unnecessarily conservative and lose significant benefits of CCP. In comparison, our solution violates the interference threshold with probability 0.14. Thus, our solution can better achieve the desired benefits while keeping the violation probability within the risk level.

Here the gap between actual threshold violation probability and risk level ϵ is because the channel model (ITU path loss and Rayleigh fading) is not the worst-case distribution since it does not have the properties in Property 1. Based on the discussion in Section 4, the violation probability does not achieve the supremum and consequently, a gap exists between the actual threshold violation probability and the risk level ϵ .

We also alter the interference threshold I from 1×10^{-7} mW to 5×10^{-7} mW while keeping the risk level $\epsilon = 0.1$. The results are shown in Fig. 2(d). As expected, the achievable objectives under all three solutions increase with higher interference threshold I . But our solution remains the best among three. The relative improvement from our our solution over Bernstein is from 13% to 50%. We omit the detailed discussion in this case study since the results are similar.

(ii) *The setting with three PUs* We test the setting with three PUs, located at $(50, 0)$, $(-45, 5)$ and $(-25, -35)$. In general, we can assign each PU with different interference thresholds I_j and different risk levels ϵ_j . Without loss of generality, we choose the interference thresholds for 3 PUs as 3×10^{-7} mW, 4×10^{-7} mW and 2×10^{-7} mW

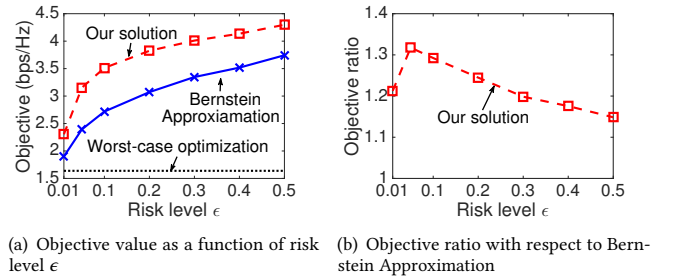


Figure 3: Performance of our solution as a function of risk level ϵ with three PUs

respectively. The risk levels ϵ_j of three PUs are the same ranging from 0.01 to 0.5. This configuration is sufficient enough to validate our proposed solution. The results are summarized in Fig. 3.

As is shown in Fig. 3, our solution outperforms the one with Bernstein Approximation with up to 30% improvement (when $\epsilon = 0.05$). Our solution achieves 4.3 bps/Hz while the one with Bernstein Approximation only obtains 3.5 bps/Hz when $\epsilon = 0.5$. We also note that the objective value is less than that of single PU scenario in Fig. 2(a). This is reasonable since the intersection of three second order cones inevitably shrinks the optimization space. Similar results with Fig. 2 (actual threshold violation probability and objective value with different thresholds I_j) are obtained under this setting so we omit these results.

6.2.2 Stressful scenario when SUs are closer to the PU. To show the robustness of our solution and the conservativeness of Bernstein Approximation, we generate a more stressful network topology where all the SUs in the picocell are closer to the PU. The coordinates of six SUs are $(20.79, -21.49)$, $(17.86, -24.49)$, $(32.16, 14.79)$, $(30.65, 5.08)$, $(24.90, -15.41)$ and $(34.21, -5.52)$. Their corresponding weights are 0.19, 0.13, 0.26, 0.12, 0.20 and 0.10. We consider one PU located at $(50, 0)$ with interference threshold $I = 3 \times 10^{-7}$ mW. Under this circumstance, the channel gains from the SUs to the PU are larger and thus strict power control on the SUs should be exercised. Fig. 4 shows the results under this circumstance.

Since the SUs are closer to the PU, they must further lower their transmission powers and consequently, the spectrum efficiency is lower compared with that from randomly distributed SUs. In

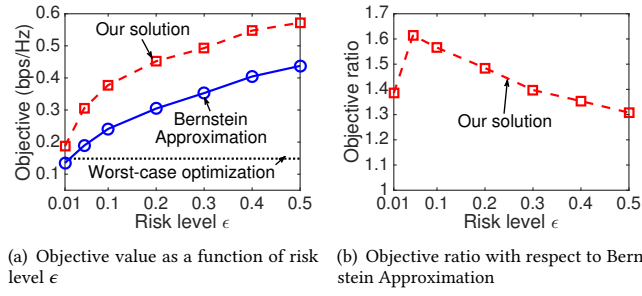


Figure 4: Performance of our solution as a function of risk level ϵ in stressful scenario

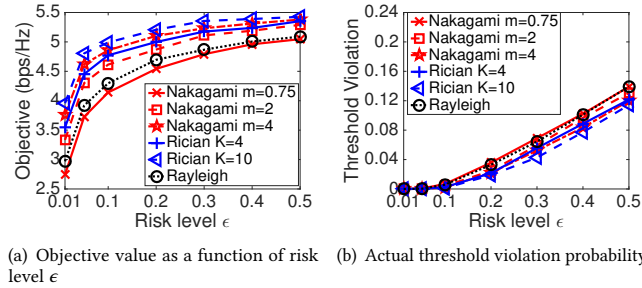


Figure 5: Performance of our solution as a function of risk level ϵ with different channel models

Fig. 4(a), we find that Bernstein Approximation has even lower spectrum efficiency compared with worst-case optimization when $\epsilon = 0.01$ due to its conservativeness. On the contrary, our solution obtains the highest spectrum efficiency for all risk levels. In fact, our solution has 60% improvement compared with Bernstein Approximation when ϵ is 0.05 as shown in Fig. 4(b).

6.3 The Case of Independent Channels with Rician Fading and Nakagami Fading

To show that our solution is applicable for all channel distributions, we test our solution with other wireless channel models using the topology in Section 6.2.1 with one PU located at (50, 0). We employ the method described in Section 6.2 and our solution only requires the knowledge of mean and covariance regardless of the channel distributions. To be specific, we employ the small-scale fading from the SUs to the PU with Nakagami fading and Rician fading in comparison with Rayleigh fading described above. The channels from the SUs to the pico BS are still based on Rayleigh fading. Since the channel models are unbounded and no truncation method for Bernstein Approximation regarding these channel models was proposed, we will only focus on our solution. The results are shown in Fig. 5.

To understand the results in Fig. 5, we need to back up some knowledge of wireless channel models. Nakagami fading is a general model for wireless channel and employs a parameter called "Nakagami- m " to adjust the seriousness of fading [14]. Both Rayleigh

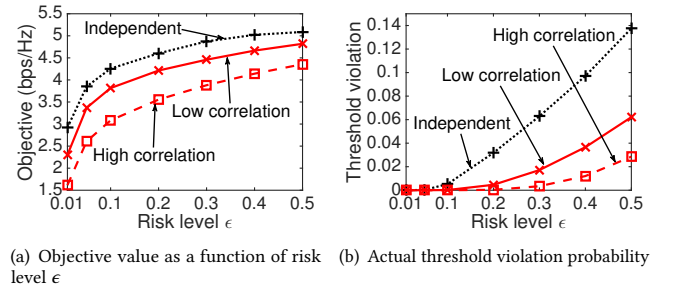


Figure 6: Performance of our solution as a function of risk level ϵ in correlated channels

fading and Rician fading can be characterized as special cases of Nakagami fading and we can calculate the "Nakagami- m " parameter accordingly. The six channel models we employed in Fig. 5 have the Nakagami- m parameters as 0.75, 2, 4, 2.78, 5.76, 1. It is shown in Fig. 5 that a smaller Nakagami- m leads to a lower spectrum efficiency even though the channels from the SUs to the pico BS are the same. This can be explained from the reformulated constraints (21) from ECR. Given the same mean, a small "Nakagami- m " generates a higher covariance and consequently we have a smaller optimization space that leads to a worse performance.

It is clearly verified in Fig 5(b) that our solution meets the threshold violation requirement (smaller than risk level ϵ) in all of the simulated channel models and thus our solution is applicable though we don't have the exact distribution knowledge. Moreover, we see the gap between risk level ϵ and actual threshold violation still exists because none of the tested channel models is a worst-case distribution. All of these results are consistent with our discussion of ECR in Section 4.

6.4 The Case of Correlated Channels with Rayleigh Fading

Since Bernstein Approximation explicitly requires independent channels and is no longer applicable under correlated channels, we will only show the results from our solution.

The topology and number of RBs are the same as those from Section 6.3. The channels are based on ITU pass loss model and Rayleigh fading and we also perform 200 simulation runs using the same method in Section 6.2. We consider both low correlation and high correlation settings. For the setting with low correlation, the correlation coefficient between adjacent RBs is set to 0.5 (i.e., $L = 1, r = 34$). For the setting with high correlation case, the correlation coefficient among all RBs from the same SU is set to 0.5 (i.e., $L = 11, r = 72$). We set the interference threshold $I = 3 \times 10^{-7}$ mW. Our results are depicted in Fig. 6.

As shown in Fig. 6(a), for the same risk level, when correlation increases, the performance of our objective (spectrum efficiency) decreases. Fig. 6(b) shows the actual threshold violation probability as a function of risk level ϵ . Here, we see that when the channels have high correlation, our solution tends to be conservative with lower actual threshold violation probability. This is because the original optimization space is smaller and our chosen of r_j contributes to

more relaxation errors as well. To the best of our knowledge, our solution to CCP is the first work that shows results with correlated channels in underlay coexistence.

7 CONCLUSIONS

In this paper, we studied underlay coexistence where occasional violation of interference threshold by SUs is allowed. We formulated the problem as a CCP and use a risk level to control the threshold violation probability. Our formulation only requires the first and second order statistics of the uncertain channel gains. For the CCP problem, we introduce a novel technique called *Exact Conic Reformulation* (ECR) that transforms the intractable chance constraints into equivalent second order cones. We show that the proposed ECR overcomes the limitations of the state-of-the-art Bernstein Approximation in terms of conservative performance and assumptions for random variables. Through extensive performance evaluation, we show that: (i) for uncorrelated channels, our proposed solution outperforms Bernstein Approximation by up-to 60 % (30% on average) in spectrum efficiency; (ii) for correlated channels, our proposed solution offers a competitive solution while Bernstein Approximation is no longer applicable.

8 ACKNOWLEDGMENTS

This research was supported in part by U.S. Army Research Laboratory under Cooperative Agreement No. W911NF1820293. The work of Y.T. Hou was also supported in part by NSF under grant 1642873.

REFERENCES

- [1] 3GPP. *3GPP TR 36.931: Radio Frequency (RF) requirements for LTE Pico Node B*, June 2018. Version 15.0.0. Available: <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2589>.
- [2] 3GPP. *3GPP TR 36.932: Scenarios and requirements for small cell enhancements for E-UTRA and E-UTRAN*, July 2018. Version 15.0.0. Available: <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2590>.
- [3] CALAFIORE, G. C., AND EL GHAOU, L. On distributionally robust chance-constrained linear programs. *Journal of Optimization Theory and Applications* 130, 1 (2006), 1–22.
- [4] CHANDRASEKHAR, V., ANDREWS, J. G., AND GATHERER, A. Femtocell networks: A survey. *IEEE Communications Magazine* 46, 9 (2008), 59–67.
- [5] CHANG, H.-B., AND RUBIN, I. Optimal downlink and uplink fractional frequency reuse in cellular wireless networks. *IEEE Transactions Vehicular Technology* 65, 4 (2016), 2295–2308.
- [6] CHARNES, A., AND COOPER, W. W. Chance-constrained programming. *Management Science* 6, 1 (1959), 73–79.
- [7] COSTELLO, D. J., AND FORNEY, G. D. Channel coding: The road to channel capacity. *Proceedings of the IEEE* 95, 6 (2007), 1150–1177.
- [8] DALL’ANESE, E., KIM, S.-J., GIANNAKIS, G. B., AND PUPOLIN, S. Power control for cognitive radio networks under channel uncertainty. *IEEE Transactions on Wireless Communications* 10, 10 (2011), 3541–3551.
- [9] DAMNJANOVIC, A., MONTOJO, J., WEI, Y., JI, T., LUO, T., VAJAPAYAM, M., YOO, T., SONG, O., AND MALLADI, D. A survey on 3GPP heterogeneous networks. *IEEE Wireless communications* 18, 3 (2011), 10–21.
- [10] DOPPLER, K., RINNE, M., WIJTING, C., RIBEIRO, C. B., AND HUGL, K. Device-to-device communication as an underlay to LTE-advanced networks. *IEEE Communications Magazine* 47, 12 (2009), 42–49.
- [11] ELSHERIF, A. R., CHEN, W.-P., ITO, A., AND DING, Z. Adaptive resource allocation for interference management in small cell networks. *IEEE Transactions on Communications* 63, 6 (2015), 2107–2125.
- [12] FELLER, W. *An introduction to probability theory and its applications*, vol. 2. John Wiley & Sons, 2008.
- [13] FURTADO, A., IRIO, L., OLIVEIRA, R., BERNARDO, L., AND DINIS, R. Spectrum sensing performance in cognitive radio networks with multiple primary users. *IEEE Transactions on Vehicular Technology* 65, 3 (2016), 1564–1574.
- [14] GOLDSMITH, A. *Wireless communications*. Cambridge University Press, 2005.
- [15] GOLDSMITH, A. J., JAFAR, S. A., MARIC, I., AND SRINIVASA, S. Breaking spectrum gridlock with cognitive radios: An information theoretic perspective. *Proceedings of the IEEE* 97, 97 (2009), 894–914.
- [16] GOLUB, G. H., AND VAN LOAN, C. F. *Matrix computations*, vol. 3. JHU press, 2012.
- [17] GUT, A. *Probability: A graduate course*, vol. 75. Springer Science & Business Media, 2013.
- [18] HOU, Y. T., SHI, Y., AND SHERALI, H. D. *Applied optimization methods for wireless networks*. Cambridge University Press, 2014.
- [19] HUANG, Y., LI, S., HOU, Y. T., AND LOU, W. GPF: A GPU-based design to achieve ~100 μ s scheduling for 5G NR. In *Proceedings of MobiCom* (2018), ACM, pp. 207–222.
- [20] JIN, F., ZHANG, R., AND HANZO, L. Fractional frequency reuse aided twin-layer femtocell networks: Analysis, design and optimization. *IEEE Transactions on Communications* 61, 5 (2013), 2074–2085.
- [21] KIM, S.-J., AND GIANNAKIS, G. B. Optimal resource allocation for MIMO ad hoc cognitive radio networks. *IEEE Transactions on Information Theory* 57, 5 (2011), 3117–3131.
- [22] KIM, S.-J., SOLTANI, N. Y., AND GIANNAKIS, G. B. Resource allocation for OFDMA cognitive radios under channel uncertainty. *IEEE Transactions on wireless communications* 12, 7 (2013), 3578–3587.
- [23] LE, L. B., AND HOSSAIN, E. Resource allocation for spectrum underlay in cognitive radio networks. *IEEE Transactions on Wireless Communications* 7, 12 (2008), 5306–5315.
- [24] LE, T. A., VIEN, Q.-T., NGUYEN, H. X., NG, D. W. K., AND SCHOBER, R. Robust chance-constrained optimization for power-efficient and secure SWIPT systems. *IEEE Transactions on Green Communications and Networking* 1, 3 (2017), 333–346.
- [25] LEE, S.-B., PEKIANAKIS, I., MEYERSON, A., XU, S., AND LU, S. Proportional fair frequency-domain packet scheduling for 3GPP LTE uplink. In *Proceedings of INFOCOM* (2009), IEEE, pp. 2611–2615.
- [26] LI, W. W.-L., ZHANG, Y. J., SO, A. M.-C., AND WIN, M. Z. Slow adaptive OFDMA systems through chance constrained programming. *IEEE Transactions on Signal Processing* 58, 7 (2010), 3858–3869.
- [27] NEMIROVSKI, A., AND SHAPIRO, A. Convex approximations of chance constrained programs. *SIAM Journal on Optimization* 17, 4 (2006), 969–996.
- [28] PAINTER, T., AND SPANIAS, A. Perceptual coding of digital audio. *Proceedings of the IEEE* 88, 4 (2000), 451–515.
- [29] PAN, C., WANG, J., ZHANG, W., DU, B., AND CHEN, M. Power minimization in multi-band multi-antenna cognitive radio networks. *IEEE Transactions on Wireless Communications* 13, 9 (2014), 5056–5069.
- [30] PARSAEFARD, S., AND SHARAFAT, A. R. Robust worst-case interference control in underlay cognitive radio networks. *IEEE Transactions on Vehicular Technology* 61, 8 (2012), 3731–3745.
- [31] SOLTANI, N. Y., KIM, S.-J., AND GIANNAKIS, G. B. Chance-constrained optimization of OFDMA cognitive radio uplinks. *IEEE Transactions on Wireless Communications* 12, 3 (2013), 1098–1107.
- [32] SURAWEERA, H. A., SMITH, P. J., AND SHAFI, M. Capacity limits and performance analysis of cognitive radio with imperfect channel knowledge. *IEEE Transactions on Vehicular Technology* 59, 4 (2010), 1811–1822.
- [33] WANG, S., SHI, W., AND WANG, C. Energy-efficient resource management in ofdm-based cognitive radio networks under channel uncertainty. *IEEE Transactions on Communications* 63, 9 (2015), 3092–3102.
- [34] WANG, Y., AND ZHU, Q.-F. Error control and concealment for video communication: A review. *Proceedings of the IEEE* 86, 5 (1998), 974–997.
- [35] WEI, L., AND TIRKKONEN, O. Spectrum sensing in the presence of multiple primary users. *IEEE Transactions on Communications* 60, 5 (2012), 1268–1277.
- [36] YUCEK, T., AND ARSLAN, H. A survey of spectrum sensing algorithms for cognitive radio applications. *IEEE Communications Surveys & Tutorials* 11, 1 (2009), 116–130.
- [37] ZHANG, L., LIANG, Y.-C., XIN, Y., AND POOR, H. V. Robust cognitive beamforming with partial channel state information. *IEEE Transactions on Wireless Communications* 8, 8 (2009), 4143–4153.