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 allow occasional interference threshold violation by the SUs, as long as such violation can be tolerated by the PUs. This paper exploits this idea through a chance constrained programming (CCP) formulation, where the knowledge of 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, namely 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 have 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in uncorrelated channels and a competitive solution in correlated channels (where Bernstein Approximation is no longer applicable). As a result, ECR represents a new tool to fully exploit the benefits of CCP when addressing uncertainty problems in spectrum sharing.

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 [14]. The SUs must carefully control their transmission power so that their interference to the PUs 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 increment deployment of new secondary networks over existing PU 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., [1, 10, 18, 20, 27]).

However, such benefits pose significant challenge for the SUs. Due to the absence of cooperation (feedback) from the PUs, accurate channel estimation of 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, we notice that occasional violations of interference thresholds, in many situations, are not considered fatal to the PUs. First, to certain extent, the inherited 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 biological perception is rather tolerable to occasional errors (distortions) and there are numerous techniques to mitigate their impacts [26, 32].

The approaches to address channel uncertainty fall into three classes: 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], channel is assumed to have log-normal shadowing and Nakagami small-scale fading while in [30], channel is assumed to have Rayleigh fading. However, in reality, many channels do not follow these simplified models and an arbitrary use of these models could lead to misleading results (either overly optimistic or conservative). Even if we had an accurate probability distribution for the RVs, the corresponding optimization problem could be extremely complicated, depending on the structure of the distribution functions.

In 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 [35], the authors studied cognitive beamforming with a bounded ellipsoid for RVs (channel gains). In [28], 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 unbounded (e.g. Rayleigh fading) or an accurate estimation of the bounded set is difficult.

The third approach, called chance constrained programming (CCP) [6], is a relatively new approach to address uncertainty in spectrum sharing [19, 21, 24, 29, 31]. 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 employs a parameter called risk level to keep violation probability below a limit and explore 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 to solve CCP is, therefore, to reformulate (substitute) the chance constraint with a deterministic constraint and by doing so, to convert the CCP into a tractable optimization problem. The state-of-the-art approach to
perform this substitution (see, e.g., [19, 21, 24, 29, 31]) is the so-called Bernstein Approximation [25]. It performs such a 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 constraint. 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 uncertainty RVs (e.g., CSI among sub-channels) are common and should be considered. Second, the performance of Bernstein Approximation depends heavily on the knowledge of the boundaries of uncertainty RVs [25], which is hard to obtain in many cases. Finally, due to its generic nature, Bernstein Approximation does not explore the unique structure in linear CCP. As a consequence, its result tends to be rather conservative, which adversely limits the potential of CCP, 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 prior work [19, 29] following the approach of Bernstein Approximation. In this paper, we propose to employ a novel technique called Exact Conic Reformulation (ECR) to address the underlying CCP. The proposed ECR allows 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 rather 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 on 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. As such, it offers a new and effective technique to solve the class of CCP problems in wireless networks.
- 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 involving 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 a novel ECR technique for CCP. In Section 5, we present the solution to the the equivalent (reformulated) deterministic optimization problem. In Section 6, we present simulation results. Section 7 concludes the paper.

2 SYSTEM MODEL

Consider a picocell residing within a macrocell as shown in Fig. 1. An example of such a scenario is that a 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 is also known as “fractional frequency reuse” in the literature [5, 11, 17].

In the underlay coexistence paradigm [14], 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
SU to the pico BS) is harder than the downlink problem (involving only one transmitter), 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 transmission\(^1\)) as well as those of nearby active PUs based on known pilots signals or channel reciprocity property [34]. Then the pico BS will collect these CSI from the SUs through a dedicated control channel and find optimal solution for scheduling (in spectrum and/or time) and power control. The goal is to maximize spectrum efficient 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 BS and then the SUs can execute their uplink transmissions based on this solution. Since neighboring picocells operate 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 macrocell BS) and channel reciprocity property. As a result, channel gains, the key parameter for controlling the transmission power 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 [12, 33].

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 tolerable threshold violation is 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

Consider the single picocell in Fig. 1 (lower portion). We are interested in maximizing spectrum efficiency for the SUs while keeping their violation of interference threshold to a PU under a target small probability. Denote Number 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 gain varies over time and differs 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 [22]. 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

\[ \text{rates. This is equivalent to assuming weights are given for the current TTI and we need to maximize the weighted sum rate for all SUs in the picocell.} \]

Denote \( x_{iB}^m \) as a binary variable to indicate whether SU \( i \) transmits to picocell BS on RB \( m \), i.e.,

\[ x_{iB}^m = \begin{cases} 1 & \text{if SU } i \text{ transmits to picocell BS on RB } m, \\ 0 & \text{otherwise}. \end{cases} \quad (1) \]

Under single user OFDMA, each RB can be assigned to at most one SU. We have

\[ \sum_{i \in N} x_{iB}^m \leq 1 \quad (m \in M). \quad (2) \]

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 (4) represents inherited power control constraint due to the SU’s equipment.

Assume each RB occupies the same bandwidth, which we normalize to 1 unit. Denote \( e_{iB}^m \) as SU \( i \)’s normalized capacity to pico

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>( c_{iB}^m )</td>
<td>Capacity of SU ( i ) toward its pico BS on RB ( m )</td>
</tr>
<tr>
<td>( g_{ij}^m )</td>
<td>Channel gain from SU ( i ) to PU ( j ) on RB ( m )</td>
</tr>
<tr>
<td>( g_j^M )</td>
<td>A column vector of ( g_{ij}^m ): ( \begin{bmatrix} g_{ij}^1, g_{ij}^2, \cdots, g_{ij}^M \end{bmatrix}^T )</td>
</tr>
<tr>
<td>( \bar{g}_j )</td>
<td>Mean of channel gain vector ( g_j )</td>
</tr>
<tr>
<td>( R_j )</td>
<td>Covariance matrix of channel gain vector ( g_j )</td>
</tr>
<tr>
<td>( h_{iB}^m )</td>
<td>Channel gain from SU ( i ) to the pico BS on RB ( m )</td>
</tr>
<tr>
<td>( l_j )</td>
<td>Interference threshold for PU ( j )</td>
</tr>
<tr>
<td>( \epsilon_j )</td>
<td>Risk level (probability upper bound) of violating PU ( j )’s interference threshold ( l_j )</td>
</tr>
<tr>
<td>( J )</td>
<td>The set of integers equal and smaller than ( J ): ( {1, 2, 3, \cdots, J} )</td>
</tr>
<tr>
<td>( M )</td>
<td>Number of RBs for transmission in the pico cell</td>
</tr>
<tr>
<td>( M )</td>
<td>The set of integers equal and smaller than ( M ): ( {1, 2, 3, \cdots, M} )</td>
</tr>
<tr>
<td>( N )</td>
<td>Number of SUs in the picocell</td>
</tr>
<tr>
<td>( N )</td>
<td>The set of integers equal and smaller than ( N ): ( {1, 2, 3, \cdots, N} )</td>
</tr>
<tr>
<td>( p_{iB}^m )</td>
<td>Transmission power from SU ( i ) to pico BS on RB ( m )</td>
</tr>
<tr>
<td>( R )</td>
<td>A column vector of ( p_{iB}^m ): ( \begin{bmatrix} p_{iB}^1, p_{iB}^2, \cdots, p_{iB}^M \end{bmatrix} )</td>
</tr>
<tr>
<td>( P_{iB}^{\max} )</td>
<td>Maximum transmission power of SU ( i ) over all RBs</td>
</tr>
<tr>
<td>( w_i )</td>
<td>Weight of SU ( i ) in current TTI</td>
</tr>
</tbody>
</table>
BS on RB $m$ (w.r.t. normalized RB bandwidth). Then we have:

$$c_i^{m} = \log_2(1 + h_i^{m} P_i^{m}) \quad (i \in N, \ m \in M),$$

(5)

where $h_i^{m}$ is the overall channel gain of SU $i$ toward pico BS on RB

$m$, including both interference and thermal noise at 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 $j$ ($j \in J$). 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:

$$\{ j \in J \} .$$

(6)

where $\mathbb{P}\{ \cdot \}$ denotes the probability function and $\epsilon_j$ is called risk level. Note that proper power control at SUs is the key to meet this chance constraint. This risk level $\epsilon_j$ could vary over a wide range (e.g., 0.01 to 0.5) depending on the application of PU $j$. A higher $\epsilon_j$ 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 distribution. In this paper, we assume that 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 such 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:

$$\text{max}_{x_{B_i}^{m}, P_i^{m}} \sum_{i \in N} \sum_{m \in M} w_i c_i^{m}$$

$$s.t. \text{ Scheduling decisions (2)}$$

$$\text{Inherited power control (4)}$$

$$\text{Calculation of capacities (5)}$$

$$\text{External power control (6)}$$

$$x_{B_i}^{m} \in \{0, 1\}, P_i^{m} \geq 0$$

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 knowledge of their distributions.\(^2\) For the same first and second order statistics, there is an infinite number of corresponding distributions. Since it is impossible to

\(^2\)Even 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 function.

enumerate all possible distributions for constraints (6), problem P1 is intractable.

4 A NOVEL REFORMULATION OF CHANCE CONSTRAINT

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 a tractable conic formulation) 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 $\epsilon_j$, while for all other distributions, the threshold violation probability is smaller than $\epsilon_j$ [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 one another, we can perform ECR for each chance constraint w.r.t. PU $j$. For ease of exposition, we rewrite (6) in matrix form

$$\mathbb{P}\{ g_j^{T} p > I_j \} \leq \epsilon_j$$

(7)

where superscript $T$ denotes transposition, $p$ is a column vector

$$p = [P_1^{1B}, P_2^{1B}, \ldots, P_1^{M}, P_2^{M}, \ldots, P_N^{M}]^T,$$

(8)

which represents $MN$ transmission powers from the SUs (over all RBs), $g_j$ is a column vector

$$g_j = [g_{1j1}, g_{1j2}, \ldots, g_{1j1}, g_{1j2}, \ldots, g_{Nj}]^T.$$ (9)

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

Denote $\overline{g}_j$ and $R_j$ as the mean and covariance matrix of $g_j$, i.e., $g_j \sim (\overline{g}_j, R_j)$. Since constraint (7) is satisfied for $g_j$ under all distributions with $g_j \sim (\overline{g}_j, R_j)$, we have

$$\sup_{g_j \sim (\overline{g}_j, R_j)} \mathbb{P}\{ g_j^{T} p > I_j \} \leq \epsilon_j ,$$

(10)

where “sup” is taken over all distributions for $g_j$ with mean $\overline{g}_j$ and covariance $R_j$.

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

$$\sup_{\xi_j \sim (0, p^T R_p p)} \mathbb{P}\{ \xi_j > \phi_j \} \leq \epsilon_j .$$

(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 $\xi_j$ can take any form of distribution with mean 0 and variance $p^T R_p p$. Then we can upper bound this closed form expression by $\epsilon_j$. Since this closed form expression has no randomness, we have a deterministic constraint on the decision variables in $p$.

To derive a closed form expression for the “sup” in (11), we need to find the worst-case distribution of $\xi_j$ that maximizes the violation probability $\mathbb{P}\{ \xi_j > \phi_j \}$. This is not a trivial problem, as the worst-case distribution of $\xi_j$ 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 $\xi_j$ (denote $f(u)$ as its probability density function) and a given interval $[a, b]$, if $ab \leq 0$, there exists a discrete RV that has two elements at $0$ and $c \in [a, b]$, $c \neq 0$ with probabilities $P_a$ and $P_c$ respectively, such that

$$
\int_a^b f(u) du = P_c + P_0 \quad \text{(Probability)} , \\
\int_a^b u f(u) du = cP_c \quad \text{(Mean)} , \\
\int_a^b u^2 f(u) du = c^2P_c \quad \text{(Variance)} .
$$

**Proof.** Obviously, $c, P_0, P_c$ should be solved based on the three equations in (12). We can justify that the solution always exists and is a valid distribution with $c \in [a, b]$. The detailed proof is summarized in our technical report [23].

Lemma 1 states that this newly constructed discrete RV preserves probability, mean and variance in $[a, b]$ with no influence on other intervals since $0 \in [a, b]$. 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 $\xi_j$)** The worst-case distribution of $\xi_j$ has the following properties:

1. If $\phi_j < 0$, then $P\{\xi_j \leq \phi_j\} = 0$.
2. If $\phi_j \geq 0$, then $P\{\xi_j > \Phi_j\} = 0, \forall \Phi_j > \phi_j$.

**Proof.** Our proof is based on contradictions.

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

$$
\xi_j = \begin{cases} 
\frac{\phi_j}{2} & P_1 = \frac{4p^T R_j p}{4p^T R_j p + \phi_j^2}, \\
-\frac{2p^T R_j p}{\phi_j} & P_2 = \frac{\phi_j^2}{4p^T R_j p + \phi_j^2}.
\end{cases}
$$

The above distribution is valid since $0 \leq P_1, P_2 \leq 1$. We see that the violation probability $P\{\xi_j > \phi_j\} = P_1 + P_2 = 1$. It is large than the violation probability from $\xi_j'$, which contradicts the assumption that $\xi_j'$ is the worst-case distribution. Therefore, any worst-case distribution of $\xi_j$ has $0$ probability in interval $(-\infty, \phi_j]$, i.e., $P\{\xi_j \leq \phi_j\} = 0$.

2. When $\phi_j \geq 0$, suppose $\xi_j'$ is a worst-case distribution but has a positive element at $b' > \phi_j$ with probability $P_{b'}$. We can construct another distribution that has an element at $b$ with probability $P_b$ such that

$$
b' > b > \phi_j, P_b > P_{b'} .
$$

Based on Lemma 1, such distribution always exists with necessary changes in interval $(-\infty, \phi_j]$ to maintain the first and second order statistics (details are in our technical report [23]). Consequently, the violation probability becomes higher, which contradicts the assumption that $\xi_j'$ is the worst-case distribution. Therefore, the worst-case distribution of $\xi_j$ when $\phi_j \geq 0$ has one element that is sufficiently close to $\phi_j$ as all of the possibilities of $\xi_j$ in interval $(\phi_j, +\infty)$ are pushed to $\phi_j$, i.e., $P\{\xi_j > \Phi_j\} = 0, \forall \Phi_j > \phi_j$.

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.

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

$$
\sup_{\xi_j \in (0, p^T R_j p)} P\{\xi_j > \phi_j\} = \begin{cases} 
1 & \phi_j < 0 , \\
\frac{p^T R_j p}{\phi_j^2 + p^T R_j p} & \phi_j \geq 0.
\end{cases}
$$

**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 at $b$ approaching $\phi_j$ with probability $P_b$. In terms of the other interval $(-\infty, \phi_j]$, it can be characterized by two discrete elements at $a < 0$ and $0$ with probabilities $P_a$ and $P_0$ respectively based on Lemma 1. This conversion from a worst-case distribution to a three-element discrete distribution preserves the violation probability $P\{\xi_j > \phi_j\}$, which is calculated as

$$
P\{\xi_j > \phi_j\} = P_b .
$$

We can derive its supremum by the following optimization problem

$$
\sup_{P_a, P_b, a, b} P_b \\
\text{s.t.} \quad P_a + P_0 + P_b = 1 \\
aP_a + bP_b = 0 \\
a^2P_a + b^2P_b = p^T R_j p \\
P_a, P_0, P_b \geq 0, a \leq 0, b > \phi_j .
$$

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

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

We now need to ensure the closed from expression for the supremum of violation probability is upper bounded by $\epsilon_j$ 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.

\[
\phi_j \geq 0 \tag{18a}
\]

\[
\frac{p^T R_j p}{\phi_j^2 + p^T R_j p} \leq \epsilon_j \tag{18b}
\]

Rewrite (18b) as

\[
\sqrt{1 - \epsilon_j} \cdot \sqrt{p^T R_j p} \leq \phi_j^2 . \tag{19}
\]

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

\[
\sqrt{1 - \epsilon_j} \sqrt{p^T R_j p} \leq \phi_j . \tag{20}
\]

Note that (20) has already considered the non-negativity of \( \phi_j \) in (18a). Substituting \( \phi_j = I_j - g_j^T \) into (20), we have the following main result.

**Theorem 1. (ECR)** With respect to decision variables in \( p \), chance constraints (11) is equivalent to the following second order cones

\[
\sqrt{1 - \epsilon_j} \sqrt{p^T R_j p} + g_j^T p \leq I_j \quad (\text{for } j \in J) . \tag{21}
\]

This is our Exact Conic Reformulation for chance constraints (11) where "exact" refers to the supremum of violation probability equals to \( \epsilon_j \).

Compared with state-of-the-art approach (Bernstein Approximation [19, 24, 25, 29, 31]), 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) in P1 by (21), we have

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

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 function for capacity calculations. We propose to employ a convex hull to relax each logarithm term \( \log_2(1 + h_{iB}^m p_{iB}^m) \) [15]. 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., \([0, p_{iB}^{\text{max}}}]\) into \( K \) equal-length sub-intervals, each with length \( \frac{p_{iB}^{\text{max}}}{K} \). Then we can upper bound the log function \( \log_2(1 + h_{iB}^m p_{iB}^m) \) with \( K + 1 \) linear functions as follows:

\[
c_{iB}^m \leq \frac{1}{\ln 2} \left( \frac{kh_{iB}^m}{K} + \ln(1 + \frac{k}{i_{B}^{p_{iB}^{\text{max}}}}) - \frac{kp_{iB}^{\text{max}}}{K} \right) \quad (k = 0, 1, \cdots, K, i \in N, m \in M) \tag{22}
\]

Constraints (22) are linear with \( c_{iB} \) 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

ECR constraints (21) are second order cones. Since \( R_j \) is the covariance matrix of \( g_j \), it is guaranteed to be positive semi-definite and symmetric. To relax the non-linear terms \( p^T R_j p \), we introduce constant matrix \( V_j \) as the square root of \( R_j \), i.e., \( R_j = V_j^T V_j \). \( V_j \) can easily be calculated based on Cholesky decomposition.

With \( V_j \), we can rewrite constraints (21) as

\[
\sqrt{1 - \epsilon_j} \sqrt{(V_j p)^T (V_j p) + g_j^T p} \leq I_j . \tag{23}
\]

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

Recall \( p \) (i.e., \([p_{iB}^1, p_{iB}^2, \cdots, p_{iB}^M]^T\)) has at most \( M \) non-zero elements because each RB can be allocation to one SU. Further, \( V_j \), as the square root of the covariance matrix \( R_j \), is an \( MN \times MN \) matrix. Since the channel gains from different SUs to PU \( j \) are usually independent, \( V_j \) is a block diagonal matrix with the \( i \)th block (denoted as \( V_{ij} \)) corresponding to SU \( i \) (\( i = 1, \cdots, N \)), i.e.,

\[
V_j = \begin{bmatrix}
V_{1j} & V_{2j} & \cdots & V_{Nj}
\end{bmatrix} . \tag{24}
\]

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 an RB is correlated with at most \( 2L_j \) neighboring RBs. Then, each block \( V_{ij} \) is a band matrix in the following form:

\[
V_{ij} = \begin{bmatrix}
\psi_{11}^{ij} & \cdots & \psi_{1(L_j+1)}^{ij} \\
\psi_{21}^{ij} & \cdots & \psi_{2(L_j+2)}^{ij} \\
\vdots & \ddots & \vdots \\
\psi_{M(L-j)}^{ij} & \cdots & \psi_{MM}^{ij}
\end{bmatrix} . \tag{25}
\]

Based on the sparse properties of \( V_j \) and \( p \), we consider two cases to calculate \( r_j \): (i) When \( M \geq 2L_j + 1 \), it can be shown that we can set \( r_j = M(2L_j + 1) - L_j(L_j + 1) \) \[23\]. (ii) When \( M < 2L_j + 1 \), \( r_j \) can be set to \( r_j = \min\{MN, M(2L_j + 1) - L_j(L_j + 1)\} \) \[23\]. Combining both case, we set \( r_j \) as

\[
r_j = \min\{MN, M(2L_j + 1) - L_j(L_j + 1)\} . \tag{26}
\]
Denote $||V_jp||_\infty$ as the infinity norm of $V_jp$. Since $\sqrt{(V_jp)^T (V_jp)} \leq \sqrt{r_j}||V_jp||_\infty$, we have the following relaxation of constraints (23):

$$\sqrt{\frac{1 - \epsilon_j}{\epsilon_j}} \sqrt{r_j}||V_jp||_\infty + g_j^T p \leq I_j.$$  \hspace{1cm} (27)

Constraints (27) are linear with decision variables in $p$. With the above linear relaxation, we have the following relaxed optimization problem.

\[ \text{(P3)} \quad \max_{x_{iB}^{m},\beta_{iB}^{m}} \sum_{i \in N} \sum_{m \in M} w_i e_i^{m} \]
\[ \text{s.t.} \quad \text{Constraints (2) } - (4) \]
\[ \text{Linearly relaxed capacity constraints (22)} \]
\[ \text{Linearly relaxed power control constraints (27)} \]
\[ x_{iB}^{m} \in \{0, 1\}, p_{iB}^{m} \geq 0 \]

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 [16].

6 SIMULATION RESULTS

In this section, we conduct simulations to evaluate our proposed solution based 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 a picocell is set to 40 meters. The transmission power of macro BS between the macro BS and the pico BS to 400 meters. The radius of all topologies in the simulation study, we set the distance be-

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 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. Worse-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 [19, 29] 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 gain 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 & Bound 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 6 SUs at the same locations but with one PU and three PUs separately. The number of RBs are set to 12 and remains the same for the rest of this simulation study.

(i) The setting with one PU. Fig. 2 shows a network topology where the SUs are randomly distributed in the picocell with one nearby PU. The location and weight of each SU are given in Table 2. The interference threshold $I$ is set to $3 \times 10^{-7}$ mW. Fig. 3 shows the performance of our solution as a function of risk level $\epsilon$ and interference threshold $I$. In Fig. 3(a), we find that the objective value of our solution monotonically increases with the risk level. In particular, we achieve 2.92 bps/Hz with a risk of $\epsilon = 0.01$ and 5.09 bps/Hz with $\epsilon = 0.5$.

In Fig. 3(a), the objective value of worst-case optimization stays the same since it does not involve any risk. Clearly, the performance of worst-case optimization is overly pessimistic due to zero tolerance of interference threshold violation.

\begin{table}[h]
\centering
\caption{Coordinates and weight for each SU.}
\begin{tabular}{|c|c|c|}
\hline
SU & Coordinates & Weight \\
\hline
1 & (-10.67,-19.01) & 0.22 \\
2 & (32.47,-11.77) & 0.09 \\
3 & (-3.23,8.90) & 0.09 \\
4 & (12.04,-5.03) & 0.23 \\
5 & (28.89,-10.00) & 0.19 \\
6 & (-22.66,5.34) & 0.18 \\
\hline
\end{tabular}
\end{table}
The performance of our solution against Bernstein Approximation is shown in both Fig. 3(a) and Fig. 3(b). In particular, in Fig. 3(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. 3(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. 3(c) offers an in depth study of our solution and Bernstein Approximation in terms of actual threshold violation probability (i.e., percentage of instances where the interference threshold is actually violated). As shown in Fig. 3(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. On the other hand, 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 $\epsilon$ 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 $\epsilon$.

We also alter the interference threshold $I$ from $1 \times 10^{-7}$ mW to $5 \times 10^{-7}$ mW while keeping the risk level at $\epsilon = 0.1$. The results are shown in Fig. 3(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 solution over Bernstein is from 13% to 50%. Due to page limit, we omit the detailed discussion in this case study.

(ii) The Case of Three PUs

We test the setting with three PUs, as the topology is shown in Fig. 4 and Table 3. In general, we can assign each PU with different interference threshold $I_j$ and different risk level $\epsilon_j$. Without loss of generality, we choose different interference thresholds but the same risk level $\epsilon$ to three PUs. This configuration is sufficient enough to validate our proposed solution. The results are summarized in Fig. 5.

As is shown in Fig. 5, 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. 3(a). This is reasonable since the intersection of three second order cones inevitably shrinks the optimization space. Similar results with Fig. 3 (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 close to the PU, as shown in Fig. 6. Table 4 shows the coordinates and weights of each SU. 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. Again, we assume the interference threshold $I = 3 \times 10^{-7}$ mW at the PU.

![Figure 3: Performance of our solution as a function of risk level $\epsilon$ and interference threshold $I$](image1)

![Figure 4: Network topology with three PUs](image2)

![Figure 5: Performance of our solution as a function of risk level $\epsilon$ with three PUs](image3)
Table 4: Coordinates and weight for each SU

<table>
<thead>
<tr>
<th>SU</th>
<th>Coordinates</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(20.79, -21.49)</td>
<td>0.19</td>
</tr>
<tr>
<td>2</td>
<td>(17.86, -24.49)</td>
<td>0.13</td>
</tr>
<tr>
<td>3</td>
<td>(32.16, 14.79)</td>
<td>0.26</td>
</tr>
<tr>
<td>4</td>
<td>(30.65, 5.08)</td>
<td>0.12</td>
</tr>
<tr>
<td>5</td>
<td>(24.90, -15.41)</td>
<td>0.20</td>
</tr>
<tr>
<td>6</td>
<td>(34.21, -5.52)</td>
<td>0.10</td>
</tr>
</tbody>
</table>

Fig. 6: Network topology, stressful scenario

Fig. 7: Performance of our solution as a function of risk level $\epsilon$ in stressful scenario

Fig. 8: Performance of our solution as a function of risk level $\epsilon$ with different channel models

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.

We use the same topology in Fig. 2 with 12 RBs. The channels are based on 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. 9.

As shown in Fig. 9(a), for the same risk level, when correlation increases, the performance of our solution (spectrum efficiency) decreases. Fig. 9(b) shows the actual interference threshold violation probability as a function of risk level $\epsilon$. 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 contributes to more relaxation errors. To the best of our knowledge, our solution to CCP is the first work that shows results with correlated channels.

7 CONCLUSIONS

In this paper, we studied underlay coexistence where occasional violation of interference threshold by SUs can be tolerated. We formulated the problem as CCP where we use a risk level to control the violation probability by SUs. Our formulation only requires the first and second order statistics of the uncertain channel gains. For the CCP problem, we propose 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 is superior to the state-of-the-art Bernstein Approximation in terms of less conservativeness and no further assumptions for random variables. Through extensive numerical study, we show that our proposed solution (predicated on ECR) outperforms that based on Bernstein Approximation. Specifically, (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 not applicable.

REFERENCES


[34] Zhang, L., Liang, Y.-C., Kim, Y., and Poor, H. V. Robust cognitive beamforming with partial channel state information. IEEE Trans. on Wireless Communications 8, 8 (2009).