Distributionally Robust Relay Beamforming in Wireless Communications

Shimin Gong
Shenzhen Institutes of Advanced Technology
Chinese Academy of Sciences, China
sm.gong@siat.ac.cn

Sissi Xiaoxiao Wu
Dept. of Systems Engineering and Engineering Management
The Chinese University of Hong Kong, Hong Kong
xxwu@ee.cuhk.edu.hk

Anthony Man-Cho So
Dept. of Systems Engineering and Engineering Management
The Chinese University of Hong Kong, Hong Kong
manchoso@se.cuhk.edu.hk

Xiaoxia Huang
Shenzhen Institutes of Advanced Technology
Chinese Academy of Sciences, China
xx.huang@siat.ac.cn

ABSTRACT

We consider a wireless network with densely deployed user devices (e.g., a device-to-device or wireless sensor network) underlaying a cellular system, in which some user devices act as relays to facilitate data transmissions between a distant transceiver pair under imperfect channel information. Motivated by the observation that most of the channel distributions are unimodal, we formulate a novel distributionally robust beamforming problem, in which the random channel coefficient follows a class of unimodal distribution with known first- and second-order moments. Our design objective is to maximize the worst-case signal-to-noise ratio (SNR) at the dedicated user device while satisfying a probabilistic interference constraint at the cellular user equipment (CUE). Though such a unimodal distributionally robust (UDR) beamforming problem is non-convex, we show that an approximate solution can be computed efficiently using semidefinite programming. Our simulation results show that under mild conditions, the UDR model yields significant beamforming performance improvement over conventional robust models that merely rely on first- and second-order moments of the channel distribution.

CCS Concepts

•**Networks** → *Network performance analysis; Mobile ad hoc networks;*

Keywords

Relay beamforming; distributional uncertainty; robust optimization

MSWiM '16, November 13-17, 2016, Malta, Malta
© 2016 ACM. ISBN 978-1-4503-4502-6/16/11...$15.00
DOI: http://dx.doi.org/10.1145/2988287.2989137

1. INTRODUCTION

As wireless devices (e.g., smart phones, wearable devices, sensors, etc.) become ubiquitous, device-to-device (D2D) communication is a promising way to increase data rates, extend network coverage, and reduce energy consumptions in cellular networks. It is proposed primarily to offload data traffic between the cellular base station (CBS) and the cellular user equipment (CUE) within or beyond the cellular coverage [12, 17]. With the upsurge of social applications (e.g., interactive gaming, file sharing, video streaming, etc.), D2D communication has become especially advantageous for a local social network, as the access to local contents is more frequent than that to the core network via the CBS. In this paper, instead of emphasizing on the performance improvement of cellular systems, we focus on the D2D communications and aim to improve the quality-of-service (QoS) of a local D2D network within the cellular coverage.

The dense deployment of D2D user equipment (DUE) allows us to achieve spatial diversity by employing multiple DUE as collaborative relays. It has been shown in [4] that the use of relays can significantly improve network performance while causing insignificant increase in end-to-end delay. However, the dense DUE makes the spectrum usage more crowded. An economical and spectrally efficient solution is to let the DUE share the same spectrum with the cellular system [7, 9]. This requires an efficient power control strategy at the DUE relays to manage the interference among the DUE and CUE. By analyzing the interference to the CUE, a strategy for selecting different D2D operating modes is proposed in [23]. In [18], the power of each DUE relay is controlled separately while the interference to the CUE is controlled by selecting the number of relays. The authors in [8] considered an ad hoc D2D network, in which the interference to the cellular system is kept below a certain interference temperature. In general, to achieve optimal relay performance, one needs to design a distributed beamformer according to different relays' channel conditions [15, 24]. However, most of the aforementioned works rely on the unrealistic assumption that the channel information is precisely known. In practice, the channel information is usually unreliable due to quantization errors, processing delay, or lack of coordination between the CUE and numerous DUE. Thus, there has been a growing effort in modeling the channel uncertainty, so that proactive strategies can be developed to avoid sharp performance deterioration. Broadly speaking, there are three different types of channel uncertainty models. A *stochastic* (STO) model assumes that the channel estimate follows an explicit distribution, such as Rician or Rayleigh [11]. Under this model, we can formulate the QoS metric in a probabilistic manner and obtain a chance-constrained robust beamforming problem [16, 22]. The *worst-case robust* (WCR) model assumes that the channel error lies in a bounded convex set and gives rise to a max-min beamforming problem [5, 6]. A less conservative approach is to build the channel models based on partial distributional information that is easy to estimate with high accuracy, such as low-order moments of the channel error. The *distributionally robust* (DRO) model was recently employed to study robust beamforming by leveraging the first- and second-order moments of the channel distribution [10]. However, such an approach can still be too conservative, as the worst-case channel distribution is shown to be discrete and hardly observed in practice [21].

The above discussion motivates us to impose additional requirements on the shape or structure of the channel distribution. Empirically, we find that the channel distribution is typically smooth, symmetric, unimodal, or even similar to some known patterns [2]. Such structural information can be extracted and exploited to further mitigate the conservatism of the DRO model. In this work, we focus on the unimodal structure and propose the *unimodal distributionally robust* (UDR) model to improve the relays' beamformer design. Our goal is to maximize the DUE receiver's SNR subject to the CUE's probabilistic interference constraints. The main contributions of this paper are summarized as follows:

- *Unimodal distributionally robust model*: Motivated by the observation that most channel distributions are unimodal, we introduce a novel UDR model, which assumes that the random channel follows a class of unimodal distributions with known first- and second-order moments. By leveraging the generalized notion of unimodality [20], we parameterize a class of unimodal distributions by a positive scalar α, which allows us to derive an upper bound on the worst-case interference violation probability. To the best of our knowledge, our work is the first to adopt the UDR model in the relays' robust beamforming problem.

- *Heuristic beamforming algorithm:* As we shall see, different values of α correspond to different levels of conservatism in channel modeling, and the UDR model degenerates into the DRO model as α approaches infinity. This observation motivates us to design an iterative algorithm to tackle the relays' beamforming problem by alternately solving the well-known DRO-based beamforming problem and updating the beamformer based on a feasibility check on the UDR-based beamforming problem. Our simulation results show that the UDR model significantly improves the DUE's throughput performance.

The rest of this paper is organized as follows. We describe the system model and the channel uncertainty in Section 2. In Section 3, we propose the relays' robust beamforming problem and design the beamforming algorithm based on the proposed UDR model. Simulations and conclusions are given in Sections 4 and 5, respectively.

2. SYSTEM MODEL

We consider a D2D network with densely deployed user devices underlaying a downlink cellular system, in which each DUE transmitter (DTx), DUE relay, and DUE receiver (DRx) are equipped with a single antenna. We assume that a direct link between the DTx and DRx is not available due to the long transmission distance and limited power at the DTx, and the DUE's data transmission is assisted by a group of DUE relays, denoted by the set $\mathcal{N} = \{1, 2, \dots, N\}$. The relays' transmissions also introduce interference to the CUE. Without loss of generality, we consider one CUE in the system for simplicity. The system model is illustrated in Figure 1. The CUE will not be interrupted if the interference from the DUE relays is less than a prescribed threshold. The downlink transmissions from the CBS to the CUE may also introduce interference to the D2D communications. Herein, we assume that such interference is constant and treat it as background noise.

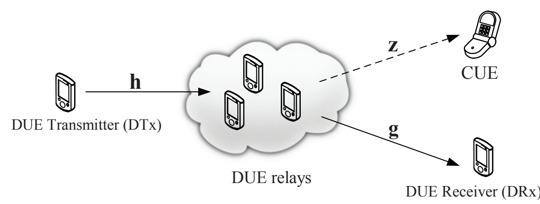

Figure 1: System model

Let $\mathbf{h} \triangleq [h_1, h_2, \dots, h_N]^T$ and $\mathbf{g} \triangleq [g_1, g_2, \dots, g_N]^T$ denote the channel coefficients from the DTx to the relays and from the relays to the DRx, respectively, and $\mathbf{z} \triangleq [z_1, z_2, \dots, z_N]^T$ denote the channel coefficients from the relays to the CUE. We assume that all channels are frequency-flat and block fading [1]; i.e., the channel coefficients remain constant during one data frame and may change independently during different data frames. We assume that there is a coordinator in the system that schedules one DTx to transmit at a time to avoid conflicts between different DUE [19]. Thus, we can focus on one DUE transceiver pair each time.

2.1 Nominal Beamforming Problem

In the system, the relays' information delivery follows a two-hop amplify-and-forward (AF) protocol. In the first hop, the DTx broadcasts a symbol s to the nearby relays. The received signal at relay n is $m_n = h_n s + \sigma_n$, where $\sigma_n \sim \mathbb{N}(0, 1)$ is the Gaussian noise with zero mean and unit variance. In the second hop, relay n forwards the signal m_n amplified by a weight w_n and thus $\mathbf{w} = [w_1, w_2, \dots, w_N]^T$ constitutes the relays' beamformer. The received signal at the DRx is

$$y = \sum_{n=1}^{N} g_n w_n h_n s + \sum_{n=1}^{N} g_n w_n \sigma_n + v_d,$$

where $v_d \sim \mathbb{N}(0, 1)$ is the noise at the DRx. Note that the first term of the received signal y contains the useful information while the second term is the noise signal forwarded by the relays. Assuming unit transmit power at the DTx (i.e., $\mathbb{E}\big[|s|^2\big] = 1$), the SNR at the DRx is given by

$$\gamma(\mathbf{w}) = \left|\sum_{n=1}^{N} w_n g_n h_n\right|^2 \Bigg/ \left(1 + \sum_{n=1}^{N} |w_n|^2 |g_n|^2\right). \quad (1)$$

We observe that it may not be optimal for all the relays to transmit at their peak power, as the relays also amplify and forward the noise signals. Let $\mathbf{a} = \mathbf{g} \circ \mathbf{h}$ be the component-wise product of the channels $\mathbf{g}$ and $\mathbf{h}$. Then, we can rewrite the DUE's SNR $\gamma(\mathbf{w})$ more compactly as $\gamma(\mathbf{w}) = \frac{\mathbf{w}^T \mathbf{A} \mathbf{w}}{1+\mathbf{w}^T \mathbf{B} \mathbf{w}}$, where $\mathbf{A} = \mathbf{a}\mathbf{a}^T$ and $\mathbf{B} = \mathrm{D}(\mathbf{g} \circ \mathbf{g})$ are positive semidefinite matrices. Here, $\mathrm{D}(\mathbf{g})$ is a diagonal matrix with the diagonal elements specified by $\mathbf{g}$.

To ensure harmonic coexistence with the cellular system, the DUE relays' transmit beamforming has to restrict the aggregate interference to the CUE. As the transmit power at relay n is $|w_n|^2(1+|h_n|^2)$, the aggregate interference received by the CUE can be expressed as $\phi(\mathbf{w}) = \sum_{n=1}^{N} |z_n|^2 |w_n|^2 (1+|h_n|^2)$. Let $\mathbf{\Lambda_w} = \mathrm{D}(\mathbf{k} \circ \mathbf{k})\mathrm{D}(\mathbf{w} \circ \mathbf{w})$, where $\mathbf{k} = [(1+h_n^2)^{1/2}]_{n\in\mathcal{N}}$ is the known information at the relays. Then, we have $\phi(\mathbf{w}) = \mathbf{z}^T \mathbf{\Lambda_w} \mathbf{z}$. Our target is to maximize the DUE's SNR by optimizing the relays' beamformer $\mathbf{w}$, subject to the CUE's interference constraint; i.e.,

$$\max_{\mathbf{w}} \left\{ \frac{\mathbf{w}^T \mathbf{A} \mathbf{w}}{1 + \mathbf{w}^T \mathbf{B} \mathbf{w}} : \mathbf{z}^T \mathbf{\Lambda_w} \mathbf{z} \leq \bar{\phi} \right\}, \quad (2)$$

where $\bar{\phi}$ is a prescribed threshold that represents the CUE's sensitivity to interference. We can also add individual or sum power budget constraints at the relays, but this would not affect our subsequent analysis. Thus, we only consider the CUE's interference constraint for simplicity. Our formulation can be easily extended to the setting where there are multiple CUE. Each CUE would then have its own interference constraint, thus ensuring protection for the most vulnerable CUE.

2.2 Channel Uncertainty Model

The solution to problem (2) requires exact knowledge of the channel coefficients. In this paper, we assume that $\mathbf{h}$ in the first hop is perfectly known by the relays' channel estimation; e.g., the DTx can broadcast a known pilot signal to facilitate the relays' channel estimation. However, due to limited or untimely responses from the DRx and CUE, the relays are unable to estimate $\mathbf{g}$ and $\mathbf{z}$ accurately. To model the uncertainties in $\mathbf{g}$ and $\mathbf{z}$, we assume that $\mathbf{g}$ and $\mathbf{z}$ are random variables following distributions $\mathbb{P}_\mathbf{g}$ and $\mathbb{P}_\mathbf{z}$, respectively. Since it is relatively easy to estimate the mean $\mathbf{u_g}$ and covariance $\mathbf{S_g}$ of the channel coefficient $\mathbf{g}$, we can define the distributional uncertainty set of $\mathbf{g}$ as

$$\mathbb{P}_\mathbf{g} \in \mathcal{P}(\mathbf{u_g}, \mathbf{S_g}) \subset \mathcal{P}_\infty, \quad (3)$$

where $\mathcal{P}(\mathbf{u_g}, \mathbf{S_g})$ is a set of distributions having the same moment statistics $(\mathbf{u_g}, \mathbf{S_g})$ and $\mathcal{P}_\infty$ is the set of all probability distributions. In the same vein, we can define the distributional uncertainty set of $\mathbf{z}$ as $\mathbb{P}_\mathbf{z} \in \mathcal{P}(\mathbf{u_z}, \mathbf{S_z})$, where $(\mathbf{u_z}, \mathbf{S_z})$ is the first- and second-order moments of the channel coefficient $\mathbf{z}$, which are known to the DUE relays. The distributional uncertainty set defined in (3) gives rise to a typical DRO model [10].

In this work, we aim to mitigate the conservatism of the DRO model by incorporating additional structural information that can be extracted from channel measurements. Specifically, we require $\mathbb{P}_\mathbf{g}$ or $\mathbb{P}_\mathbf{z}$ to be unimodal, which is a commonly observed property in the area of wireless communications and signal processing [13]. Intuitively, unimodality implies a single local maxima (referred to as the mode) in the density function. For *univariate* unimodal distributions, the density function is non-decreasing to the left of the mode and non-increasing to the right of the mode. In particular, Gaussian, Rayleigh, and Rician distributions are examples of univariate unimodal distribution. By incorporating unimodality in the DRO model, we can remove those hardly observed multi-modal distributions from the uncertainty set (3), thereby leading to more practical distributional uncertainty sets for $\mathbf{g}$ and $\mathbf{z}$. To fulfill this purpose, we first need an analytical characterization of unimodality for *multivariate* distributions. This can be achieved using the notion of α-unimodality [20]:

Definition 1. *Suppose that $\mathbb{P} \in \mathcal{P}_\infty$ has a continuous probability density function f on $\mathbb{R}^N$. We say that $\mathbb{P}$ is α-unimodal with mode $\mathbf{0}$ for some $\alpha > 0$ if $t^{N-\alpha} f(t\mathbf{x})$ is non-increasing in $t > 0$ for all $\mathbf{x} \neq \mathbf{0}$. The set of all α-unimodal distributions with mode $\mathbf{0}$ is denoted as $\mathcal{P}_\alpha$.*

When $\alpha = N$, α-unimodality implies that the density function is non-increasing along the rays emanating from its mode, which coincides with our intuition for univariate unimodal distributions. When $\alpha > N$, the density function f may increase along the rays emanating from its mode, but $t^{N-\alpha} f(t\mathbf{x})$ is still non-increasing in $t > 0$. The rate of increase of f is controlled by the parameter α. Hence, α-unimodality is a more general characterization of unimodal distributions. Now, we can define the unimodal distributionally robust (UDR) model for $\mathbf{g}$ and $\mathbf{z}$ as follows:

$$\mathbb{P}_\mathbf{g} \in \mathcal{P}_\mathbf{g}^\alpha \triangleq \mathcal{P}_\mathbf{g}(\mathbf{u_g}, \mathbf{S_g}) \bigcap \mathcal{P}_\alpha, \quad (4a)$$

$$\mathbb{P}_\mathbf{z} \in \mathcal{P}_\mathbf{z}^\alpha \triangleq \mathcal{P}_\mathbf{z}(\mathbf{u_z}, \mathbf{S_z}) \bigcap \mathcal{P}_\alpha. \quad (4b)$$

The parameterized sets $\mathcal{P}_\mathbf{g}^\alpha$ and $\mathcal{P}_\mathbf{z}^\alpha$ allow flexibility in modeling the uncertainties in $\mathbf{g}$ and $\mathbf{z}$ under different channel conditions. When α is equal to the number of DUE relays, most of the practically observed distributions will fall in the set $\mathcal{P}_\alpha$ [20]. When α approaches infinity, every distribution belongs to the set $\mathcal{P}_\infty$. In this case, the UDR model (4) degenerates into the DRO model (3).

3. ROBUST BEAMFORMING PROBLEM

Both the SNR performance $\gamma(\mathbf{w})$ and the aggregate interference $\phi(\mathbf{w})$ are functions of the channel conditions. Thus, they become stochastic when the channels $\mathbf{g}$ and $\mathbf{z}$ are drawn from $\mathbb{P}_\mathbf{g}$ and $\mathbb{P}_\mathbf{z}$, respectively. Considering the UDR model (4), we formulate the robust counterpart of (2) as follows:

$$\max_{\mathbf{w}} \quad \min_{\mathbb{P}_\mathbf{g} \in \mathcal{P}_\mathbf{g}^\alpha} \frac{\mathbb{E}_{\mathbb{P}_\mathbf{g}}[\mathbf{w}^T \mathbf{A} \mathbf{w}]}{1 + \mathbb{E}_{\mathbb{P}_\mathbf{g}}[\mathbf{w}^T \mathbf{B} \mathbf{w}]} \quad (5a)$$

$$s.t. \quad \max_{\mathbb{P}_\mathbf{z} \in \mathcal{P}_\mathbf{z}^\alpha} \mathbb{P}_\mathbf{z}\left(\mathbf{z}^T \mathbf{\Lambda_w} \mathbf{z} \geq \bar{\phi}\right) \leq \eta. \quad (5b)$$

Here, we optimize the relays' beamformer $\mathbf{w}$ to maximize the DUE's worst-case SNR while limiting the CUE's worst-case (with respect to all distributions with the given structural and moment information) interference violation probability below a prescribed probability threshold η.

Due to the non-convex probabilistic constraint (5b), the UDR beamforming problem (5) is generally difficult to solve optimally. To simplify it, we introduce the SNR target $\rho \geq 0$ and observe that (5a) is equivalent to maximizing ρ subject to the following QoS constraint:

$$\rho + \max_{\mathbb{P}_{\mathbf{g}} \in \mathcal{P}_{\mathbf{g}}^{\alpha}} \mathbb{E}_{\mathbb{P}_{\mathbf{g}}} \left[\mathbf{g}^T \left(\rho \mathrm{D}(\mathbf{w} \circ \mathbf{w}) - \mathbf{q}\mathbf{q}^T \right) \mathbf{g} \right] \leq 0, \quad (6)$$

where we denote $\mathbf{q} = \mathrm{D}(\mathbf{h})\mathbf{w}$ for simplicity. Note that the SNR constraint (6) only involves the second-order moment $\mathbf{S}_g$ and not the particular structure of the distribution $\mathbb{P}_{\mathbf{g}}$. Hence, we can simplify (6) as

$$\rho + \rho \mathbf{Tr}(\mathrm{D}(\mathbf{w} \circ \mathbf{w})\mathbf{S}_{\mathbf{g}}) - \mathbf{q}^T \mathbf{S}_{\mathbf{g}} \mathbf{q} \leq 0. \quad (7)$$

To simplify the probabilistic constraint (5b), let $e(\mathbf{z}) = \mathbf{1}(\mathbf{z}^T \mathbf{\Lambda}_{\mathbf{w}} \mathbf{z} \geq \bar{\phi})$, where $\mathbf{1}(\cdot)$ is the indicator function. Define

$$B_\alpha(\mathbf{u}_{\mathbf{z}}, \mathbf{S}_{\mathbf{z}}) \ = \max_{\mathbb{P}_{\mathbf{z}} \in \mathcal{P}_{\mathbf{z}}^{\alpha}} \mathbb{E}_{\mathbb{P}_{\mathbf{z}}} \left[e(\mathbf{z}) \right] \quad (8)$$

to be the worst-case interference violation probability on the LHS of (5b). Since the maximization in (8) is semi-infinite and the constraint $B_\alpha(\mathbf{u}_{\mathbf{z}}, \mathbf{S}_{\mathbf{z}}) \leq \eta$ has no known closed-form convex equivalence, we shall focus on deriving an upper bound on $B_\alpha(\mathbf{u}_{\mathbf{z}}, \mathbf{S}_{\mathbf{z}})$ in the sequel. To proceed, we first study the approximation of $B_\alpha(\mathbf{u}_{\mathbf{z}}, \mathbf{S}_{\mathbf{z}})$ in a special case of the UDR model; namely, the DRO model (which corresponds to α approaching infinity). After getting some insights from this special case, we then consider the UDR model in its full generality.

3.1 A Special Case of the UDR Model

When α approaches infinity, the UDR model degenerates into the DRO model. In this case, we have the following equivalence [25]:

$$B_\infty(\mathbf{u}_{\mathbf{z}}, \mathbf{S}_{\mathbf{z}}) \ = \ \min_{\mathbf{M}, \nu} \ \mathbf{Tr}(\mathbf{\Sigma}_{\mathbf{z}} \mathbf{M}) \quad (9a)$$

$$s.t. \ \ \mathbf{M} \succeq \begin{bmatrix} \nu \mathbf{\Lambda}_{\mathbf{w}} & \mathbf{0} \\ \mathbf{0} & 1 - \nu \bar{\phi} \end{bmatrix}, \quad (9b)$$

$$\mathbf{M} \succeq \mathbf{0}, \quad \nu \geq 0, \quad (9c)$$

where $\mathbf{M}$, ν are dual variables and $\mathbf{\Sigma}_{\mathbf{z}} = \begin{bmatrix} \mathbf{S}_{\mathbf{z}} & \mathbf{u}_{\mathbf{z}} \\ \mathbf{u}_{\mathbf{z}}^T & 1 \end{bmatrix}$ denotes the second-order moment matrix of the channel $\mathbf{z}$. For any fixed $\mathbf{w}$, (9b) is a linear matrix inequality. Hence, problem (9) is a semidefinite program (SDP) and provides a tractable reformulation of $B_\infty(\mathbf{u}_{\mathbf{z}}, \mathbf{S}_{\mathbf{z}})$. Upon replacing (5b) by (9), the optimal beamformer $\mathbf{w}^\star$ can be obtained as follows:

$$\max_{\rho \geq 0, \mathbf{w}} \ \rho \quad (10a)$$

$$s.t. \ \ \rho + \rho \mathbf{Tr}(\mathrm{D}(\mathbf{w} \circ \mathbf{w})\mathbf{S}_{\mathbf{g}}) - \mathbf{q}^T \mathbf{S}_{\mathbf{g}} \mathbf{q} \leq 0, \quad (10b)$$

$$\mathbf{Tr}(\mathbf{\Sigma}_{\mathbf{z}} \mathbf{M}) \leq \nu \eta, \quad (10c)$$

$$\mathbf{M} \succeq \begin{bmatrix} \mathrm{D}(\mathbf{k} \circ \mathbf{k})\mathrm{D}(\mathbf{w} \circ \mathbf{w}) & \mathbf{0} \\ \mathbf{0} & \nu - \bar{\phi} \end{bmatrix}, \quad (10d)$$

$$\mathbf{M} \succeq \mathbf{0}, \quad \nu \geq 0. \quad (10e)$$

Unfortunately, problem (10) is still non-convex, as the constraints (10b) and (10d) are quadratic in $\mathbf{w}$. To circumvent this difficulty, we apply the semidefinite relaxation (SDR) technique [14]. Specifically, by first introducing the rank-one matrix $\mathbf{W} = \mathbf{w}\mathbf{w}^T$ and then dropping the non-convex rank-one constraint on $\mathbf{W}$, we obtain the following SDR of problem (10):

$$(\mathrm{SUB}): \max_{\rho \geq 0, \mathbf{W} \succeq \mathbf{0}} \ \rho \quad (11a)$$

$$s.t. \quad \rho + \rho \mathbf{Tr}(\mathbf{\Delta}(\mathbf{W})\mathbf{S}_{\mathbf{g}}) - \mathbf{q}^T \mathbf{S}_{\mathbf{g}} \mathbf{q} \leq 0, \quad (11b)$$

$$\mathbf{Tr}(\mathbf{\Sigma}_{\mathbf{z}} \mathbf{M}) \leq \nu \eta, \quad (11c)$$

$$\mathbf{M} \succeq \begin{bmatrix} \mathrm{D}(\mathbf{k} \circ \mathbf{k})\mathbf{\Delta}(\mathbf{W}) & \mathbf{0} \\ \mathbf{0} & \nu - \bar{\phi} \end{bmatrix}, \quad (11d)$$

$$\mathbf{M} \succeq \mathbf{0}, \quad \nu \geq 0, \quad (11e)$$

where $\mathbf{\Delta}(\mathbf{W}) = \mathrm{D}(\mathbf{w} \circ \mathbf{w})$ is the diagonal matrix by setting all off-diagonal elements of $\mathbf{W}$ to zero.

Problem (SUB) can be solved optimally by a bisection method. Indeed, for a given target SNR ρ, it is clear that (11b) and (11c) are linear inequalities and (11d) is a linear matrix inequality. This implies that checking the feasibility of the constraints (11b)–(11e) is an SDP, which can be solved efficiently by the interior-point algorithms embedded in some well-known optimization toolbox, such as SeDuMi and CVX [3]. In particular, we can increase or decrease ρ depending on whether the constraints (11b)–(11e) are satisfied or not, until convergence is achieved. Now, suppose that the bisection method converges to the optimal beamforming matrix $\mathbf{W}^\star$. If $\mathbf{W}^\star$ happens to be a rank-one matrix, then the optimal beamformer $\mathbf{w}^\star$ to problem (10) can be extracted by eigen-decomposition. Otherwise, an approximate rank-one solution can be extracted from $\mathbf{W}^\star$ by, e.g., a Gaussian randomization method [14].

3.2 The General Case of the UDR Model

Now, let us focus on the general UDR model and derive an upper bound on $B_\alpha(\mathbf{u}_{\mathbf{z}}, \mathbf{S}_{\mathbf{z}})$ for finite values of α. By definition of α-unimodality, a finite α imposes non-trivial structure on the distribution $\mathbb{P}_{\mathbf{z}}$, and a smaller α implies more stringent structural requirement; i.e., $\mathcal{P}_{\alpha_1} \supset \mathcal{P}_{\alpha_2}$ for $\alpha_1 \geq \alpha_2 > 0$. Therefore, by tuning the value of α, we can construct a flexible robust model that captures the information uncertainty in different channel conditions. However, the lack of a concrete representation of $\mathcal{P}_{\mathbf{z}}^{\alpha}$ makes it very difficult to solve problem (8) in closed-form. A simplification of $\mathcal{P}_{\mathbf{z}}^{\alpha}$ relies on the construction of a special class of the α-unimodal distributions.

Definition 2. *For any $\alpha > 0$ and $\mathbf{x} \in \mathbb{R}^n$, the radial α-unimodal distribution supported on the line segment $[\mathbf{0}, \mathbf{x}] \subset \mathbb{R}^n$, denoted by $\delta_{[\mathbf{0},\mathbf{x}]}^{\alpha}(\cdot)$, is a distribution with the property that $\delta_{[\mathbf{0},\mathbf{x}]}^{\alpha}([\mathbf{0}, t\mathbf{x}]) = t^\alpha$ for $t \in [0, 1]$.*

Lemma 1 ([20]). *For any $\mathbb{P} \in \mathcal{P}_\alpha$, there exists a unique distribution $\mathbf{m} \in \mathcal{P}_\infty$ such that $\mathbb{P}(\cdot) = \int_{\mathbb{R}^n} \delta_{[\mathbf{0},\mathbf{x}]}^{\alpha}(\cdot)\mathbf{m}(d\mathbf{x})$.*

Lemma 1 maps any α-unimodal distribution $\mathbb{P}_{\mathbf{z}} \in \mathcal{P}_{\mathbf{z}}^{\alpha}$ to a general distribution $\mathbf{m} \in \mathcal{P}_\infty$ without the unimodality constraint. This mapping allows us to reformulate (8) as a conventional moment constrained problem:

$$B_\alpha(\mathbf{u}_{\mathbf{z}}, \mathbf{S}_{\mathbf{z}}) = \max_{\mathbf{m} \in \mathcal{P}_\infty} \ \mathbb{E}_{\mathbf{m}} \left[p_\alpha(\mathbf{x}) \right] \quad (12a)$$

$$s.t. \quad \mathbb{E}_{\mathbf{m}} \begin{bmatrix} \mathbf{x}\mathbf{x}^T & \mathbf{x} \\ \mathbf{x}^T & 1 \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{S}}_{\mathbf{z}} & \bar{\mathbf{u}}_{\mathbf{z}} \\ \bar{\mathbf{u}}_{\mathbf{z}}^T & 1 \end{bmatrix}, \quad (12b)$$

where $p_\alpha(\mathbf{x}) = \int_{\mathbb{R}^n} e(\mathbf{z}) \delta_{[\mathbf{0},\mathbf{x}]}^{\alpha}(d\mathbf{z})$ and the moment statistics $\bar{\mathbf{S}}_{\mathbf{z}}$ and $\bar{\mathbf{u}}_{\mathbf{z}}$ are given by $\frac{2+\alpha}{\alpha}\mathbf{S}_{\mathbf{z}}$ and $\frac{1+\alpha}{\alpha}\mathbf{u}_{\mathbf{z}}$, respectively. The equivalence between problems (8) and (12) is straightforward by applying Lemma 1 to (8). Such equivalence implies that once we find an $\mathbf{m}$ that is feasible for (12), we can construct an α-unimodal distribution $\mathbb{P}_{\mathbf{z}}$ that is feasible for (8). Thus, in the sequel, we shall focus on problem (12).

Although problem (12) has a similar structure to that in the DRO model, the Lagrangian method does not give a tractable dual form of (12). Instead, we resort to the following approximation:

PROPOSITION 1. *An upper bound on $B_\alpha(\mathbf{u}_{\mathbf{z}}, \mathbf{S}_{\mathbf{z}})$ is given by*

$$\max_{\mathbf{Y}\succeq\mathbf{0},\mathbf{y}\geq\mathbf{0},\lambda,\tau\geq0} \quad \lambda-\tau \tag{13a}$$

$$s.t. \quad \begin{bmatrix} \bar{\mathbf{S}}_{\mathbf{z}} & \bar{\mathbf{u}}_{\mathbf{z}} \\ \bar{\mathbf{u}}_{\mathbf{z}}^T & 1 \end{bmatrix} - \begin{bmatrix} \mathbf{Y} & \mathbf{y} \\ \mathbf{y}^T & \lambda \end{bmatrix} \succeq \mathbf{0}, \tag{13b}$$

$$\begin{bmatrix} \mathbf{Y} & \mathbf{y} \\ \mathbf{y}^T & \lambda \end{bmatrix} \succeq \mathbf{0}, \tag{13c}$$

$$\tau^2\big(\mathbf{Tr}(\mathbf{\Lambda_w Y})\big)^\alpha \geq \lambda^{\alpha+2}\bar{\phi}^\alpha. \tag{13d}$$

PROOF. We need to show that for any feasible solution $\mathbf{m}$ to problem (12), we can construct a feasible solution $(\mathbf{Y}, \mathbf{y}, \lambda, \tau)$ to problem (13) such that it achieves at least the same probability bound as $\mathbf{m}$. Towards that end, let $\mathbb{P}_{\mathbf{z}}$ be a feasible α-unimodal distribution for (8) and $\mathbf{m}$ denote the corresponding mixture distribution in (12). Let $\Xi \triangleq \{\mathbf{x}\,|\,\mathbf{x}^T\mathbf{\Lambda_w x} < \bar{\phi}\}$ and $\bar{\Xi} = \mathbb{R}^n \setminus \Xi$. Then, we can construct a feasible solution to problem (13) by the following rules:

$$\begin{bmatrix} \mathbf{Y} & \mathbf{y} \\ \mathbf{y}^T & \lambda \end{bmatrix} = \int_{\bar{\Xi}} \begin{bmatrix} \mathbf{x}\mathbf{x}^T & \mathbf{x} \\ \mathbf{x}^T & 1 \end{bmatrix} \mathbf{m}(d\mathbf{x}) \succeq \mathbf{0}, \tag{14a}$$

$$\tau = \int_{\bar{\Xi}} \big(1 - p_\alpha(\mathbf{x})\big)\,\mathbf{m}(d\mathbf{x}) \geq 0. \tag{14b}$$

Obviously, we have $\lambda \leq 1$ and

$$\begin{bmatrix} \bar{\mathbf{S}}_{\mathbf{z}} & \bar{\mathbf{u}}_{\mathbf{z}} \\ \bar{\mathbf{u}}_{\mathbf{z}}^T & 1 \end{bmatrix} = \int_{\mathbf{x}\in\Xi\cup\bar{\Xi}} \begin{bmatrix} \mathbf{x}\mathbf{x}^T & \mathbf{x} \\ \mathbf{x}^T & 1 \end{bmatrix} \mathbf{m}(d\mathbf{x}) \succeq \begin{bmatrix} \mathbf{Y} & \mathbf{y} \\ \mathbf{y}^T & \lambda \end{bmatrix}.$$

Note that $p_\alpha(\mathbf{x}) = \int_{\mathbb{R}^n} \mathbf{1}(\mathbf{z}^T\mathbf{\Lambda_w z} \geq \bar{\phi})\delta^\alpha_{[\mathbf{0},\mathbf{x}]}(d\mathbf{z})$. Moreover, by the construction of Ξ, we have

$$1 - p_\alpha(\mathbf{x}) = \begin{cases} \left(\frac{\bar{\phi}}{\mathbf{x}^T\mathbf{\Lambda_w x}}\right)^{\frac{\alpha}{2}} & \text{for } \mathbf{x} \in \bar{\Xi}, \\ 1 & \text{for } \mathbf{x} \in \Xi, \end{cases} \tag{15}$$

which implies $\tau \geq \left(\frac{\bar{\phi}}{\int_{\bar{\Xi}} \mathbf{x}^T\mathbf{\Lambda_w x}\,\mathbf{m}(d\mathbf{x})}\right)^{\frac{\alpha}{2}} \geq \lambda\left(\frac{\lambda\bar{\phi}}{\mathbf{Tr}(\mathbf{Y\Lambda_w})}\right)^{\frac{\alpha}{2}}$. The first inequality comes from the Jensen inequality and the second inequality is due to the fact that $\int_{\bar{\Xi}} \mathbf{x}\mathbf{x}^T\mathbf{m}(d\mathbf{x}) = \mathbf{Y}$ and $\lambda \leq 1$. The final step is to verify that the same objective can be achieved by the construction in (14): $\mathbb{E}_{\mathbf{m}}[p_\alpha(\mathbf{x})] = \int_{\bar{\Xi}} \mathbf{m}(d\mathbf{x}) - \int_{\bar{\Xi}} \big(1 - p_\alpha(\mathbf{x})\big)\mathbf{m}(d\mathbf{x}) = \lambda - \tau$. This completes the proof. □

Proposition 1 implies that we can find a safe approximation of (5b) by solving the upper bound problem (13). However, the nonlinear constraint in (13d) is non-convex and depends on the value of α. To further simplify it, we can introduce an auxiliary variable $s \geq 0$ such that

$$\big(\mathbf{Tr}(\mathbf{\Lambda_w Y})\big)^\alpha s^\alpha \geq \lambda^{2\alpha}\bar{\phi}^\alpha, \tag{16a}$$

$$\tau^2\lambda^{\alpha-2}s^{2^l-\alpha} \geq s^{2^l}, \tag{16b}$$

where $l = \lceil \log_2(\alpha) \rceil$ is the smallest integer such that $2^l - \alpha \geq 0$. By imposing mild conditions on α, we can derive an equivalent convex formulation of (16a)–(16b) as follows:

PROPOSITION 2. *Let $\alpha \geq 2$ be an integer. By introducing auxiliary variables $t_{2^k+m} \geq 0$, where $k = 0, 1, \ldots, l-1$, the constraints (16a)–(16b) are equivalent to*

$$\begin{bmatrix} \mathbf{Tr}(\mathbf{\Lambda_w Y}) & \lambda\bar{\phi}^{\frac{1}{2}} \\ \lambda\bar{\phi}^{\frac{1}{2}} & s \end{bmatrix} \succeq \mathbf{0}, \tag{17a}$$

$$\begin{bmatrix} t_{2^{k+1}+2m} & t_{2^k+m} \\ t_{2^k+m} & t_{2^{k+1}+2m+1} \end{bmatrix} \succeq \mathbf{0}, \tag{17b}$$

$$t_1 \geq s \tag{17c}$$

for $m = 0, 1, \ldots, 2^k - 1$ and $k = 0, 1, \ldots, l-1$, where

$$t_{2^l+m} = \begin{cases} \tau & \text{for } m = 0, 1, \\ \lambda & \text{for } m = 2, \ldots, \alpha - 1, \\ s & \text{for } m = \alpha, \ldots, 2^l - 1. \end{cases} \tag{18}$$

PROOF. It is easy to see that (17a) is equivalent to (16a). Thus, we only need to show the equivalence between (16b) and (17b)–(17c). By the construction (18) and the inequalities in (17b)–(17c), all power exponents in (16b) are nonnegative. Thus, we have

$$\tau^2\lambda^{\alpha-2}s^{2^l-\alpha} = \prod_{m=0}^{2^l-1} t_{2^l+m} \geq \left(\prod_{m=0}^{2^{l-1}-1} t_{2^{l-1}+m}\right)^2 \geq \cdots \geq t_1^{2^l} \geq s^{2^l}.$$

The converse is also true. For any (τ, λ, s) that is feasible for (16b), we can simply construct a set of values $\{t_{2^{l-1}+m}\}_{0\leq m\leq 2^{l-1}-1}$ such that $t_{2^l+2m}t_{2^l+2m+1} = t^2_{2^{l-1}+m}$ and

$$\tau^2\lambda^{\alpha-2}s^{2^l-\alpha} = \prod_{m=0}^{2^{l-1}-1} t^2_{2^{l-1}+m} \geq s^{2^l}.$$

The same construction applies to the second inequality and yields $\tau^2\lambda^{\alpha-2}s^{2^l-\alpha} = \cdots = t_1 \geq s$. □

Given a fixed beamformer $\mathbf{w}$, both constraints (17a) and (17b) are linear matrix inequalities. Hence, we can solve problem (13) efficiently by leveraging on the SDP reformulation (17).

3.3 UDR-based Beamforming Algorithm

In view of the approximation of $B_\alpha(\mathbf{u}_{\mathbf{z}}, \mathbf{S}_{\mathbf{z}})$, we can reformulate the general UDR-based beamforming problem as the following bi-level optimization problem:

$$(\text{GUB}) : \max_{\rho\geq0,\mathbf{w}} \; \rho \tag{19a}$$

$$s.t. \quad \rho + \rho\mathbf{Tr}(\mathrm{D}(\mathbf{w}\circ\mathbf{w})\mathbf{S}_{\mathbf{g}}) - \mathbf{q}^T\mathbf{S}_{\mathbf{g}}\mathbf{q} \leq 0, \tag{19b}$$

$$\eta \geq \max\{\lambda - \tau : (13b), (13c), (17a)\text{–}(17c)\}. \tag{19c}$$

On the upper-level of problem (GUB), we optimize the rank-one beamformer $\mathbf{w}$ to maximize the target SNR ρ. The constraint (19b) ensures the target SNR at the designated DUE and the constraint (19c) bounds the CUE's worst-case interference violation probability, which is the optimum of the lower-level optimization problem as in (13). In particular, similar to problem (SUB), we can again pinpoint the maximum target SNR $\rho^\star$ by checking the feasibility of (19b)–(19c) for a given ρ and applying a bisection method. However, the feasibility check now requires solving the lower-level optimization problem, which is non-convex due to the coupling between $\mathbf{w}$ and $\mathbf{Y}$ in the matrix inequality (17a). Nevertheless, note that problems (GUB) and (SUB) differ only in the interference constraint. By exploiting the connection between problems (GUB) and (SUB), we aim to design a simpler algorithm that can bypass checking the feasibility of (19b)–(19c) directly.

PROPOSITION 3. *The upper bound on $B_\alpha(\mathbf{u_z}, \mathbf{S_z})$ derived in problem (13) is an increasing function of α.*

PROOF. For any $\alpha \geq 2$ and $\mathbf{\Gamma} = (\mathbf{Y}, \mathbf{y}, \lambda, \tau)$ that is feasible for problem (13), let $u_b(\alpha, \mathbf{\Gamma}) = \lambda - \tau$ denote the objective value of $\mathbf{\Gamma}$. Now, consider a fixed $\alpha \geq 2$ and let $\mathbf{\Gamma}^\star = (\mathbf{Y}^\star, \mathbf{y}^\star, \lambda^\star, \tau^\star)$ be an optimal solution to problem (13). Note that the inequality constraint (13d) is satisfied as an equality at $\mathbf{\Gamma}^\star$; i.e., $\tau^\star = \lambda^\star \left(\frac{\lambda^\star \bar{\phi}}{\mathbf{Tr}(\mathbf{\Lambda_w Y}^\star)}\right)^{\frac{\alpha}{2}}$. Hence, we have $\frac{\mathbf{Tr}(\mathbf{\Lambda_w Y}^\star)}{\lambda^\star \bar{\phi}} = \left(\frac{\lambda^\star}{\tau^\star}\right)^{\frac{2}{\alpha}} \geq 1$, which implies that

$$\frac{\partial u_b(\alpha, \mathbf{\Gamma}^\star)}{\partial \alpha} = \frac{\lambda^\star}{2} \left(\frac{\lambda^\star \bar{\phi}}{\mathbf{Tr}(\mathbf{\Lambda_w Y}^\star)}\right)^{\frac{\alpha}{2}} \log\left(\frac{\mathbf{Tr}(\mathbf{\Lambda_w Y}^\star)}{\lambda^\star \bar{\phi}}\right) \geq 0.$$

In particular, for any $0 < \alpha_1 \leq \alpha_2$, we have $u_b(\alpha_1, \mathbf{\Gamma}_1^\star) \leq u_b(\alpha_2, \mathbf{\Gamma}_1^\star) \leq u_b(\alpha_2, \mathbf{\Gamma}_2^\star)$, where the first inequality is due to the monotonicity of $u_b(\alpha, \mathbf{\Gamma}_1^\star)$ with respect to α and the second inequality is due to the fact that $\mathbf{\Gamma}_2^\star$ is an optimal solution corresponding to α_2. This completes the proof. ☐

Proposition 3 implies that (SUB) produces the largest evaluation of $B_\alpha(\mathbf{u_z}, \mathbf{S_z})$ and thus it is a restricted version of (GUB); i.e., any beamformer $\mathbf{w}$ that is feasible for (SUB) is also feasible for (GUB). Moreover, since the added structural information in (GUB) mitigates the conservatism in channel estimation, (GUB) will achieve a better SNR target than that of (SUB) for the same beamformer $\mathbf{w}$. This observation motivates us to approximate (GUB) by iteratively checking the feasibility of (SUB). The basic idea is to start from a beamformer $\mathbf{w}$ that is feasible for (SUB). Then, we fix $\mathbf{w}$ and solve the lower-level problem (13). Once we obtain an upper bound on $B_\alpha(\mathbf{u_z}, \mathbf{S_z})$, we check the feasibility of (19c), which then facilitates the update of $\mathbf{w}$. The entire procedure is summarized in Algorithm 1.

Algorithm 1 UDR-Based Beamforming Algorithm

1: initialize $\tilde{\eta}_{\min} = \eta$, $\tilde{\eta}_{\max} = 1$, and $\tilde{\eta} = \eta$
2: **while** $|\tilde{\eta}_{\max} - \tilde{\eta}_{\min}| \geq \varepsilon$
3: find $(\rho, \mathbf{W})$ by solving (SUB) with $\tilde{\eta}$
4: extract rank-one beamformer $\mathbf{w}$ from $\mathbf{W}$
5: evaluate $B_\alpha(\mathbf{u_z}, \mathbf{S_z})$ by solving SDP (13)
6: **if** $B_\alpha(\mathbf{u_z}, \mathbf{S_z}) \leq \eta$
7: $\tilde{\eta}_{\min} \leftarrow \tilde{\eta}$
8: **else**
9: $\tilde{\eta}_{\max} \leftarrow \tilde{\eta}$
10: **end if**
11: $\tilde{\eta} \leftarrow (\tilde{\eta}_{\max} + \tilde{\eta}_{\min})/2$
12: **end while**
13: return convergent $(\rho, \mathbf{w})$

Given the CUE's probability threshold η (e.g., $\eta = 0.1$), we initialize $(\rho, \mathbf{W})$ in line 3 by solving (SUB) with the initial threshold $\tilde{\eta} = \eta$. The rank-one solution $\mathbf{w}$ can be extracted from $\mathbf{W}$ by the Gaussian randomization method [14]. Simulation results reveal that problem (SUB) always gives rank-one solutions and a similar observation has been reported in [10]. Next, fixing the beamformer $\mathbf{w}$, we check the feasibility of (19c) in problem (GUB) and update the probability threshold $\tilde{\eta}$ by the bisection method shown in lines 6–11. Then, we solve (SUB) again to update $(\rho, \mathbf{W})$. Note that Algorithm 1 always returns a feasible rank-one solution to (GRB), which provides a lower bound on the optimal value of problem (5) for any integer $\alpha \geq 2$.

4. NUMERICAL RESULTS

In this section, we demonstrate the efficacy of the proposed UDR beamforming design and compare it with some existing robust models. For simplicity, we consider 3 DUE relays collaboratively amplifying and forwarding the received signal to the DUE receiver. The noise at each relay and at the DUE has zero mean and unit variance. The channel coefficients $\mathbf{g}$ and $\mathbf{z}$ are random variables with known mean and covariance, but their exact distribution functions are unknown. We require the channel distributions to be α-unimodal. Since our safe approximation of (5b) is valid only when $\alpha \geq 2$ is an integer, we shall consider integer values of α in the simulations. By setting α to be the number of relays, the UDR model contains those commonly observed unimodal distributions in practice.

4.1 Comparing with the DRO Model [10]

In the general UDR model, an upper bound on $B_\alpha(\mathbf{u_z}, \mathbf{S_z})$ is obtained by solving problem (13) with a fixed beamformer $\mathbf{w}$. This leads to a safe approximation of the probabilistic constraint (5b). When α approaches infinity, the UDR model degenerates into the DRO model [10], which admits the equivalent SDP formulation (9) and gives an exact evaluation of $B_\alpha(\mathbf{u_z}, \mathbf{S_z})$. To compare the UDR and DRO models, we fix the same beamformer $\mathbf{w}$ in both problems (9) and (13) and then numerically evaluate the interference violation probability and the DUE's data rate. The comparison results are presented in Figure 2, where the DRO model is denoted by "$\alpha \to \infty$". We observe that a larger α implies less stringent structural requirement and a larger (thus more conservative) evaluation of $B_\alpha(\mathbf{u_z}, \mathbf{S_z})$, thus resulting in over-protection for the CUE and performance loss at the DUE. As a result, the DRO model significantly overestimates the interference violation probability; see Figure 2(a). Next, we set the probability threshold at $\eta = 0.2$ and show the data rate of the D2D communications in Figure 2(b). We see that although the general UDR model can only be solved by a heuristic method (i.e., Algorithm 1), it still provides the DUE a significantly higher data rate than the exactly solvable DRO model.

4.2 Comparing with the BER model [22]

The BER model assumes that the channel coefficients follow a Gaussian distribution. Under this model, a safe approximation of the constraint $B_\alpha(\mathbf{u_z}, \mathbf{S_z}) \leq \eta$ can be developed using the Bernstein-type inequality [22]. Since the Gaussian distribution is α-unimodal for any α greater than the dimension of the channel vector [20], the BER model demands more stringent structural information than that of the UDR model. To compare these two models, we first optimize the beamformer $\mathbf{w}$ in different models and then

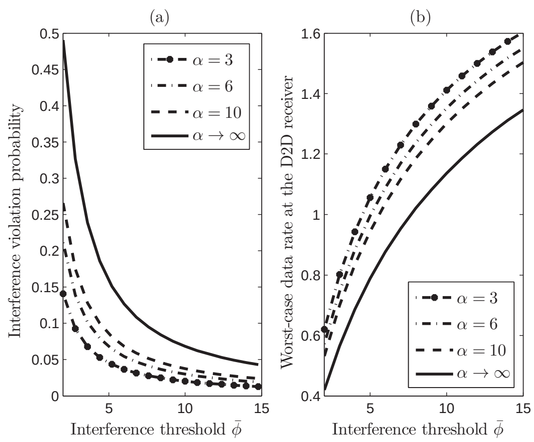

Figure 2: The UDR model outperforms the DRO model

check their throughput performance with randomized channel realizations of $\mathbf{g}$ and $\mathbf{z}$. We set the CUE's probability threshold at $\eta = 0.2$ and generate 10^6 realizations of the uncertain channel coefficients $\mathbf{g}$ and $\mathbf{z}$ according to their moment information in each simulation run. For each interference threshold $\bar{\phi}$, we record the CUE's interference violation probabilities in Table 1 and show the DUE's data rate in Figure 3. We observe that the interference violation probabilities corresponding to different models do not vary too much with the increase of $\bar{\phi}$. We also notice that the interference violation probability in the DRO model is much smaller than that of the BER and UDR models due to the lack of structural information. This explains the DRO model's conservative DUE throughput performance in Figure 3.

Table 1: Simulation results with different $\bar{\phi}$

$\bar{\phi}$	5	6	7	8	9	10
DRO	0.71%	0.65%	0.67%	0.65%	0.60%	0.66%
BER	2.01%	1.98%	2.09%	2.05%	2.00%	2.08%
UDR	6.43%	6.32%	6.46%	6.34%	6.31%	6.29%

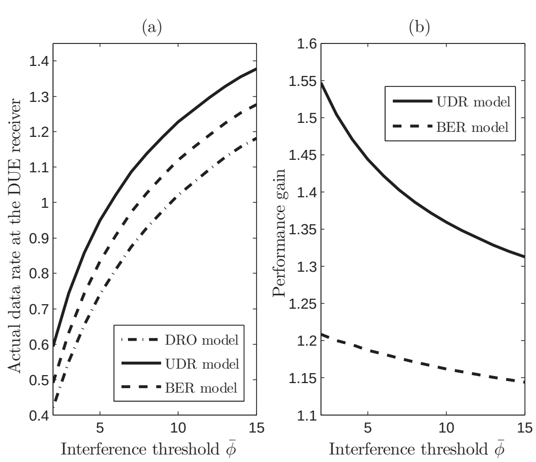

Figure 3: Throughput performance in three robust models

A counter-intuitive observation in Table 1 is that the UDR model has the highest interference violation probability and consequently a better throughput performance than that of the BER model; see Figure 3. Though the BER model demands more structural information than that of the UDR model, the looser approximation in the Bernstein-type inequality results in an over-estimation of the interference violation probability. Therefore, the robust beamformer design in the BER model is more conservative than that of the UDR model. In Figure 3(b), the performance gain of the UDR (or BER) model is defined as the expected ratio $\mathbb{E}[\rho_U/\rho_D]$ (or $\mathbb{E}[\rho_B/\rho_D]$), where ρ_D and ρ_U (or ρ_B) denote the optimal SNR targets of the DRO and UDR (or BER) models, respectively. We find that the UDR model improves the DUE's performance significantly by exploiting the additional structural information in channel estimation, especially when the CUE has a small interference threshold.

By tuning the value of α, the UDR model has the flexibility to model the channel uncertainty. To show this, we vary the value of α and record the actual performance of these robust models in Table 2. In the simulation, we set $\bar{\phi} = 6$ and $\eta = 0.2$. Note that the DRO and BER models do not take the unimodal structure into account. Thus, we record nearly constant interference violation probabilities in the first two rows of Table 2. By the same reason, the performance gain $\mathbb{E}[\rho_B/\rho_D]$ is also a constant for different α. However, when we reduce α and thus impose more stringent structural requirement on the distribution $\mathbb{P}_{\mathbf{z}}$, the UDR model achieves higher SNR for the DUE by pushing the actual interference violation probability closer to its target level η. Hence, we observe an increasing value of $\mathbb{E}[\rho_U/\rho_D]$ as α decreases. When α approaches infinity, the UDR model degenerates into the DRO model and the performance gain $\mathbb{E}[\rho_U/\rho_D]$ tends to unity. A practical choice of the parameter α is to relate it to the dimension of the channel coefficients. Specifically, by setting $\alpha = N$, the set $\mathcal{P}_{\mathbf{z}}^{\alpha}$ contains most of the commonly observed distributions in wireless systems, and the UDR model achieves a significant performance improvement (around 50%) over the DRO model; see Table 2.

Table 2: Simulation results with different α values

α	3	4	5	6	10
DRO	0.96%	0.94%	0.97%	1.06%	1.13%
BER	1.90%	1.92%	1.91%	1.99%	2.02%
UDR	7.56%	6.49%	5.98%	5.37%	3.96%
$\mathbb{E}[\rho_B/\rho_D]$	1.151	1.151	1.152	1.153	1.152
$\mathbb{E}[\rho_U/\rho_D]$	1.556	1.487	1.452	1.415	1.305

5. CONCLUSIONS

In this work, we studied DUE relays' beamforming in a cellular network under imperfect channel information, where our goal is to maximize the DUE's worst-case SNR subject to the CUE's probabilistic interference constraint. Based on the notion of α-unimodality, we proposed the UDR model that can significantly mitigate the conservatism of the conventional DRO and BER models in robust beamforming design. As a future work, we may incorporate more structural information into the channel uncertainty model and study the more involved setting where the channel information in both hops of a D2D relay network is uncertain.

6. ACKNOWLEDGMENTS

This work was supported jointly by the National Natural Science Foundation of China under Grant No. 61503368, 61601449, U1501255, and U1301256; the Guangdong Science and Technology Project under Grant No. 2015A010103009; the Shenzhen Science and Technology Project under grant No. CXZZ20150401152251212, JCYJ20150401150223648, and JCYJ20150401145529016.

7. REFERENCES

[1] F. Gao, R. Zhang, and Y.-C. Liang. Channel estimation for OFDM modulated two-way relay networks. *IEEE Trans. Signal Process.*, 57(11):4443–4455, November 2009.

[2] S. Gong, P. Wang, and L. Duan. Distributed power control with robust protection for PUs in cognitive radio networks. *IEEE Trans. Wireless Commun.*, 14(6):3247–3258, June 2015.

[3] M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming, version 2.1. http://cvxr.com/cvx, March 2014.

[4] M. Hasan and E. Hossain. Distributed resource allocation for relay-aided device-to-device communication: A message passing approach. *IEEE Trans. Wireless Commun.*, 13(11):6326–6341, November 2014.

[5] M. Hasan and E. Hossain. Distributed resource allocation for relay-aided device-to-device communication under channel uncertainties: A stable matching approach. *IEEE Trans. Wireless Commun.*, 63(10):3882–3897, October 2015.

[6] M. Hasan, E. Hossain, and D. I. Kim. Resource allocation under channel uncertainties for relay-aided device-to-device communication underlaying LTE-A cellular networks. *IEEE Trans. Wireless Commun.*, 13(4):2322–2338, April 2014.

[7] K. Jayasinghe, P. Jayasinghe, N. Rajatheva, and M. Latva-Aho. Linear precoder-decoder design of MIMO device-to-device communication underlaying cellular communication. *IEEE Trans. Wireless Commun.*, 62(12):4304–4319, December 2014.

[8] B. Kaufman, J. Lilleberg, and B. Aazhang. Spectrum sharing scheme between cellular users and ad-hoc device-to-device users. *IEEE Trans. Wireless Commun.*, 12(3):1038–1049, March 2013.

[9] P. Li, S. Guo, T. Miyazaki, and W. Zhuang. Fine-grained resource allocation for cooperative device-to-device communication in cellular networks. *IEEE Wireless Commun.*, 21(5):35–40, October 2014.

[10] Q. Li, A. M.-C. So, and W.-K. Ma. Distributionally robust chance-constrained transmit beamforming for multiuser MISO downlink. In *Proc. of IEEE ICASSP*, pages 3479–3483, May 2014.

[11] M. Lin, J. Ouyang, and W. P. Zhu. Joint beamforming and power control for device-to-device communications underlaying cellular networks. *IEEE J. Sel. Areas Commun.*, 34(1):138–150, January 2016.

[12] J. Liu, Y. Kawamoto, H. Nishiyama, N. Kato, and N. Kadowaki. Device-to-device communications achieve efficient load balancing in LTE-advanced networks. *IEEE Wireless Commun.*, 21(2):57–65, April 2014.

[13] S. Loyka, V. Kostina, and F. Gagnon. On convexity of error rates in digital communications. *IEEE Trans Inf. Theory*, 59(10):6501–6516, October 2013.

[14] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang. Semidefinite relaxation of quadratic optimization problems. *IEEE Signal Process. Mag.*, 27(3):20–34, May 2010.

[15] D. H. N. Nguyen, H. H. Nguyen, and T. T. Pham. Distributed beamforming in multiuser multi-relay networks with guaranteed QoS. In *Proc. of IEEE GLOBECOM*, November 2009.

[16] Y. Qin, M. Ding, M. Zhang, H. Yu, and H. Luo. Relaying robust beamforming for device-to-device communication with channel uncertainty. *IEEE Commun. Lett.*, 18(10):1859–1862, October 2014.

[17] Y. A. Sambo, M. Z. Shakir, K. A. Qaraqe, E. Serpedin, and M. A. Imran. Expanding cellular coverage via cell-edge deployment in heterogeneous networks: Spectral efficiency and backhaul power consumption perspectives. *IEEE Commun. Mag.*, 52(6):140–149, June 2014.

[18] S. H. Seyedmehdi and G. Boudreau. An efficient clustering algorithm for device-to-device assisted virtual MIMO. *IEEE Trans. Wireless Commun.*, 13(3):1334–1343, March 2014.

[19] M. N. Tehrani, M. Uysal, and H. Yanikomeroglu. Device-to-device communication in 5G cellular networks: Challenges, solutions, and future directions. *IEEE Commun. Mag.*, 52(5):86–92, May 2014.

[20] B. P. G. Van Parys, P. J. Goulart, and D. Kuhn. Generalized Gauss inequalities via semidefinite programming. *Mathematical Programming*, 156(1):271–302, 2016.

[21] L. Vandenberghe, S. Boyd, and K. Comanor. Generalized Chebyshev bounds via semidefinite programming. *SIAM Review*, 49(1):52–64, 2007.

[22] K.-Y. Wang, A. M.-C. So, T.-H. Chang, W.-K. Ma, and C.-Y. Chi. Outage constrained robust transmit optimization for multiuser MISO downlinks: Tractable approximations by conic optimization. *IEEE Trans. Signal Process.*, 62(21):5690–5705, November 2014.

[23] L. Wang, T. Peng, Y. Yang, and W. Wang. Interference constrained D2D communication with relay underlaying cellular networks. In *Proc. of IEEE VTC Fall*, September 2013.

[24] Y. Wang, B. Ji, Y. Huang, T. Ban, and L. Yang. Robust collaborative relay beamforming design for two-way relay systems with reciprocal CSI. *Wireless Networks*, 21(7):2209–2221, October 2015.

[25] S. Zymler, D. Kuhn, and B. Rustem. Distributionally robust joint chance constraints with second-order moment information. *Mathematical Programming*, 137(1-2):167–198, 2013.