Computationally Efficient Robust Beamforming for SINR Balancing in Multicell Downlink With Applications to Large Antenna Array Systems

Muhammad Fainan Hanif, Le-Nam Tran, *Member, IEEE*, Antti Tölli, *Senior Member, IEEE*, and Markku Juntti, *Senior Member, IEEE*

***Abstract*—We address the problem of the downlink beamformer design for signal-to-interference-plus-noise ratio balancing in a multiuser multicell environment with imperfectly estimated channels at base stations. We first present a semidefinite program (SDP)-based approximate solution to the problem. Then, as our main contribution, by exploiting some properties of the robust counterpart of the optimization problem, we arrive at a second-order cone program (SOCP)-based approximation of the balancing problem. The advantages of the proposed SOCP-based design are twofold. First, it greatly reduces the computational complexity compared to the SDP-based method. Second, it applies to a wide range of uncertainty models. As a case study, we investigate the performance of proposed formulations when the base station is equipped with a massive antenna array. Numerical experiments are carried out to confirm that the proposed robust designs achieve favorable results in scenarios of practical interest.**

***Index Terms*—SINR balancing, massive MIMO, very large-scale antenna arrays, reduced complexity, interference channel, multicell beamforming.**

I. Introduction

In practical wireless systems, it is virtually impossible to provide an error-free estimate of channel state information (CSI) to the transmitter. Although beamforming is very attractive from implementation and performance perspective, its applicability is reduced due to its sensitivity to channel estimation errors which may arise as a consequence of pilot contamination in multicell systems [1], quantization effects due to digital processing [2] etc. Motivated by this dilemma, various studies have been conducted to design 'uncertainty immune' precoders, see, e.g., [3]–[6] and references therein. The key tool common to all the studies is the application of various important results from the robust optimization theory [7], [8]. For any optimization problem, the design of robust counterpart can potentially suffer from two major difficulties, namely, (i) hurdles in obtaining tractable representation of the robust counterpart of the original program thereby compelling to employ various approximations, and, (ii) once a tractable formulation is obtained an increase in the complexity of the robust counterpart is seen as compared to the original problem. This pattern is common to most robust designs pertaining to signal processing and communication applications in the literature.

The significance of (ii) in designing uncertainty immune precoders is further enhanced when some of the parameters involved in the system setup take very large values. In this context, the recently envisaged large scale massive multiple-input multiple-output (MIMO) systems [9] can be considered. Indeed, such large-scale antenna arrays promise increased link reliability, better spectral efficiency and low power consumption at the cost of manyfold increase in the number of transmitter antennas compared with the traditional multiple-antenna systems. For instance, values of the order of hundreds of base station antennas have been proposed in [9]. The sub-optimality of traditional precoding methods like zero-forcing, block-diagonalization is now well understood [10]. Any algorithm that is, for example, based on traditional mathematical programming is likely to outperform the heuristic approaches of zero-forcing etc. It is also pertinent to point out that the mathematical analysis of the present paper can be easily leveraged to the case of maximizing weighted sum rates based on the development presented in [10]. The traditional approaches of introducing robustness in the the precoder design can end up in a semidefinite program (SDP). The complexity of an SDP is highly sensitive to the precoder size (more details on this appear in Section III-B), and hence the SDP-based solutions can either incur appreciable computational cost or in certain circumstances the digital resources may not be sufficient to cater for the memory requirements of an SDP-based solution. On the other hand, second-order cone programs (SOCPs) are much more computationally efficient (again the details appear in Section III-B), and can certainly provide a viable alternative to designing algorithms for very large-scale antenna arrays. This motivates arriving at robust SOCP formulations for optimizing certain performance metric in modern communication systems.

In this paper, we study the problem of signal-to-interference-plus-noise ratio (SINR) balancing in multicell multiple-input single-output (MISO) downlink or interfering broadcast channel with a realistic assumption of imperfect CSI. We focus on

Manuscript received June 29, 2013; revised November 4, 2013 and February 5, 2014; accepted April 23, 2014. Date of publication April 29, 2014; date of current version June 18, 2014. This work was supported by the Academy of Finland and the Finnish Funding Agency for Technology and Innovation. The editor coordinating the review of this paper and approving it for publication was A. Ghrayeb.

The authors are with the Department of Communications Engineering and Centre for Wireless Communications, University of Oulu, 90014 Oulu, Finland (e-mail: mfh21@uclive.ac.nz; ltran@ee.oulu.fi; atolli@ee.oulu.fi; markku.juntti@ee.oulu.fi).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCOMM.2014.2320913

0090-6778 © 2014 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

centralized base station (BS) control. The worst case design philosophy that is commonly employed in the existing literature is considered. We first show that the robust counterpart can be relaxed to an SDP, and, thus, can be (suboptimally) solved in conjunction with a bisection search. It appears that the SDP-based formulations provide a general solution to the robust design in many works, e.g., [4]–[6]. However, the SDP-based methods may rely on a rank relaxation scheme, which is in general a suboptimal technique. Furthermore, the SDP-based approaches generally result in computationally expensive tractable robust counterparts. As our main contribution, we propose a robust design which is merely based on solving SOCPs, i.e., the proposed method does not represent much increase in complexity in comparison to the original version of the SINR balancing problem. This is accomplished by exploiting various properties of the constraints in the robust counterpart of the balancing problem. In particular, we avoid formulating the beamformer design by projecting it to the space of semidefinite matrices which normally results in a rank constrained SDP. More importantly, the proposed SOCP-based design can be used in a wide range of uncertainty models. We notice that the SDP-based design formulations commonly used in literature [4]–[6] are only applicable to the cases where the channel errors lie in an ellipsoid. As mentioned above, we also compare and contrast the SDP and SOCP solutions, particulary from the computational cost perspective, when the number of base station antennas is very large [9]. Finally, through numerical investigations, we show that the proposed SOCP-based solution offers comparable performance to the approach in [6] and the SDP-based method when same uncertainty set (a ball) is used to represent channel perturbations.

The rest of the paper is organized as follows. Section II presents problem formulation, a solution for perfect CSI, and modeling of the balancing problem with imperfect CSI. Section III discusses in detail various solutions with imperfect CSI along with a comparative discussion about their properties. Finally, Sections IV and V describe numerical experiments and conclusions, respectively.

Notation: We use bold lowercase letters to express vectors and bold uppercase letters to represent matrices. $(.)^{\mathrm{H}}$, $(.)^{\mathrm{T}}$ and $\mathrm{Tr}(.)$ represent the Hermitian, transpose and the trace operators, respectively. $\mathbb{C}^{a\times b}$ and $\mathbb{R}^{p\times q}$ represent the space of complex and real matrices (vectors) of dimensions given as superscripts, respectively. $|\mathcal{M}|$ denotes the cardinality of set $\mathcal{M}$. $[\mathbf{p}]_k$ represents the kth component of vector $\mathbf{p}$. $|c|$ and $\Re(c)$ represent the absolute value and the real part of a complex number c, respectively. $\mathbf{I}_T$ denotes a $T\times T$ identity matrix. Finally, $\|.\|_2$ represents the l_2 norm.

II. System Model

Consider a system of B coordinated BSs. Each BS is equipped with T transmit antennas and each user with a single receive antenna. A sketch of the system model is presented in Fig. 1. Interference originating outside the coordinated system is omitted. We assume that the indices of the users served by a BS b are contained in $\mathcal{U}_b$ so that the number of users it serves, K_b, are given by the cardinality of this set i.e., $|\mathcal{U}_b| = K_b$. The

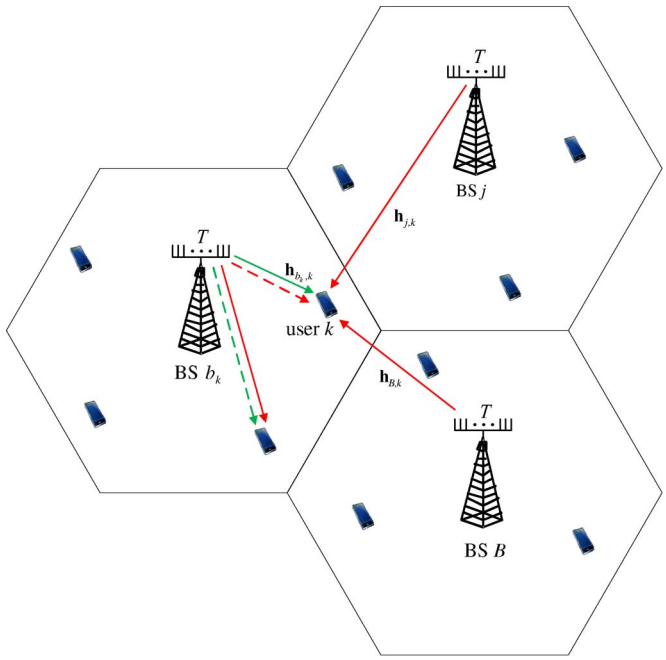


Fig. 1. System model of a multicell MISO downlink channel. Green lines represent desired signals, while red ones denote interference. The serving base station for user k is denoted by b_k.

total number of users in the whole system is K i.e., $\sum_b K_b = K$. The serving BS for the kth user is denoted by b_k. Hence, the signal received by the kth user is

$$y_{b_k,k} = \mathbf{h}_{b_k,k}\mathbf{x}_k + \sum_{i=1,i\neq k}^{K} \mathbf{h}_{b_i,k}\mathbf{x}_i + n_k \tag{1}$$

where $\mathbf{h}_{b_k,k} \in \mathbb{C}^{1\times T}$ is the channel (row) vector from BS b_k to user k, $\mathbf{x}_k \in \mathbb{C}^{T\times 1}$ is the transmitted signal vector from the BS b_k to user k and $n_k \sim \mathcal{CN}(0,\sigma^2)$ represents circularly symmetric zero mean complex Gaussian noise with variance σ^2. The transmitted signal vector is defined as $\mathbf{x}_k = \mathbf{m}_k d_k$, where $\mathbf{m}_k \in \mathbb{C}^{T\times 1}$ is the unnormalized beamforming vector and d_k is the normalized complex data symbol. The total power transmitted by BS b is $\sum_{k\in\mathcal{U}_b} \mathrm{Tr}(\mathrm{E}[\mathbf{x}_k\mathbf{x}_k^{\mathrm{H}}]) = \sum_{k\in\mathcal{U}_b} \|\mathbf{m}_k\|_2^2$. The SINR at user k served by BS b_k is

$$\gamma_{b_k,k} = \frac{|\mathbf{h}_{b_k,k}\mathbf{m}_k|^2}{\sigma^2 + \sum\limits_{i\in\mathcal{U}_{b_k}\setminus k} |\mathbf{h}_{b_k,k}\mathbf{m}_i|^2 + \sum\limits_{b=1,b\neq b_k}^{B}\sum\limits_{i\in\mathcal{U}_b} |\mathbf{h}_{b,k}\mathbf{m}_i|^2} \tag{2}$$

where the interference power in the denominator is divided into intra- and inter-cell interference components. We would like to remark that the level of coordination assumed in our framework only considers sharing channel information, and not the data of the users. Such an information is collected by several BSs and passed to a centralized controller connected to them via fast microwave or fiber optic links. Nonetheless, it is straightforward to extend our results to the case when infinite capacity backhaul link is assumed to exist between several BSs, so that a user is potentially served by more than one BSs.

A. Problem Statement and Solution for Perfect CSI

For the case of perfect CSI, the maximin SINR balancing can be cast as

$$\underset{\mathbf{m}_k:\sum_{k\in\mathcal{U}_b}\|\mathbf{m}_k\|_2^2\le P_b,\forall b}{\text{maximize}} \quad \min_k \alpha_{b_k,k}\gamma_{b_k,k} \tag{3}$$

where $\alpha_{b_k,k}$ are positive weighting factors. Using (2), we can equivalently reformulate (3) as (see the formulation in (4) at the bottom of the page) where $\mathbf{M}_b = [\mathbf{m}_{\mathcal{U}_b(1)}, \ldots, \mathbf{m}_{\mathcal{U}_b(|\mathcal{U}_b|)}]$ includes the precoders of all users being served in the bth cell and the operation $\mathrm{vec}(.)$ vectorizes the argument matrix by stacking columns. Furthermore, we can still find an optimal solution of (4) even if $\mathbf{h}_{b_k,k}\mathbf{m}_k$, for all k, is forced to be real [3], [11], [12]. In this way, the constraints in (4) represent second-order cone (SOC) constraints for fixed t. Therefore, the original problem can be solved as a series of SOCP feasibility problems using bisection search [11], [13].

B. Modeling of Imperfect CSI

In real systems it is impossible to achieve perfect transmitter CSI due to several reasons mentioned in, e.g., [5], [6]. Hence, robust designs dealing with channel errors are of practical importance. For large scale antenna array systems, time division duplex (TDD) techniques have been envisaged [14] for channel estimation purposes. However, even the TDD cannot guarantee an errorless estimation of channels. For instance, the errors at the receiver during channel estimation propagate to the transmitter even if reciprocity is exploited to obtain the channel estimate at the transmitter. Therefore, the focus of this work is not on estimating channels. Instead, without focusing on a particular estimation procedure, we aim to optimize system performance in the presence of unavoidable channel errors. We consider the channel uncertainty model in which the true channel vectors are of the form

$$\mathbf{h}_{b,k} = \hat{\mathbf{h}}_{b,k} + \sum_{i=1}^{l_{b,k}} \boldsymbol{\delta}_{b,k}^i [\mathbf{v}_{b,k}]_i = \hat{\mathbf{h}}_{b,k} + \mathbf{v}_{b,k}\mathbf{A}_{b,k},\ \forall b,k \tag{5}$$

where $\hat{\mathbf{h}}_{b,k}$ represents the nominal (known) value of the channels, $l_{b,k} \in \{1,2,\ldots,T\}$, the vectors $\boldsymbol{\delta}_{b,k}^i$ (channel perturbation directions) form the rows of $\mathbf{A}_{b,k} \in \mathbb{C}^{l_{b,k}\times T}$ and $\mathbf{v}_{b,k}\mathbf{A}_{b,k}$ gives the error vector in the downlink channel from BS b to user k [7], [15]. We denote by $\mathcal{S}$ the uncertainty set that includes all channel error row vectors $\mathbf{v}_{b,k}$. As seen in (5) the above model assumes that the uncertainty vector affects the data in an affine manner. This philosophy has been widely used, e.g., in [5], [6] etc. In addition to affecting the true channels in an affine manner, the error vectors are also constrained to lie in an uncertainty set $\mathcal{S}_{b,k}$ as

$$\mathcal{S}_{b,k} = \{\mathbf{v}_{b,k} : \|\mathbf{v}_{b,k}\| \le \rho_{b,k},\ \forall b,k\} \tag{6}$$

where $\|\cdot\|$ is an appropriate norm specified by the parameter $\rho_{b,k}$, and is chosen based on how one wishes to model channel uncertainties. Normally, modeling channel errors by exploiting their statistical nature is prone to numerous difficulties. To name a few, it requires information about the statistics of the error vectors which is mostly not available on account of myriad of phenomenon involved in the channel estimation process. Then even if some information about the statistics of the error vectors is exploitable, a part from simple linear constraints contaminated with Gaussian errors, it is virtually impossible to arrive at exact tractable versions of *stochastic constraints*. Motivated by the dilemma, the uncertainty in channels is modeled by norm-bounded sets (6). With such modeling, it does not remain necessary to know information about the, say, probability law that the uncertainty vectors follow. Further to this, as we will see in the discussion to follow, norm-bounded uncertainty sets model various real world scenarios very well. One more advantage of such modeling is that in several cases of interest, the norm-bounded uncertainty models permit either exact tractable formulations or good approximations [15], [16]. Some studies related to zero-forcing like solutions in the presence of errors have appeared in the literature [9], [14], [17]. We will see in the sections to follow that our method is much more flexible and versatile in terms of handling various types and/or combinations of uncertainty sets.

C. Worst Case Design Formulation

We will concentrate on the worst case robust optimization approach of [8], [15] that has been traditionally used in the existing literature for different problems [4]–[6]. The worst case approach amounts to satisfying the constraints for all possible channel vectors. Hence, the robust counterpart of (3) is written as

$$\begin{aligned}
&\max_{\mathbf{m}_k,t}. \quad t \\
&\text{s.t.} \quad \begin{cases} \frac{\alpha_{b_k,k}}{t}|\mathbf{h}_{b_k,k}\mathbf{m}_k|^2 \ge \sum\limits_{i\in\mathcal{U}_{b_k}\setminus k} |\mathbf{h}_{b_k,k}\mathbf{m}_i|^2 \\ \quad + \sum\limits_{b=1,b\ne b_k}^{B}\sum\limits_{i\in\mathcal{U}_b} |\mathbf{h}_{b,k}\mathbf{m}_i|^2 + \sigma^2,\ \forall k, \\ \quad \forall\{\mathbf{v}_{b,k}\mathbf{A}_{b,k} : \mathbf{v}_{b,k}\in\mathcal{S}_{b,k}\} \end{cases} \\
&\quad \sum_{k\in\mathcal{U}_b}\|\mathbf{m}_k\|_2^2 \le P_b, \forall b.
\end{aligned} \tag{7}$$

$$\begin{aligned}
&\underset{\mathbf{m}_k,t}{\text{maximize}} \quad t \\
&\text{subject to} \quad \left\{ \left\|(\mathbf{h}_{1,k}\mathbf{M}_1 \ \cdots \ \mathbf{h}_{B,k}\mathbf{M}_B \ \sigma)^{\mathrm{H}}\right\|_2 \le \sqrt{1+\tfrac{\alpha_{b_k,k}}{t}}|\mathbf{h}_{b_k,k}\mathbf{m}_k|,\ \forall k, \quad \|\mathrm{vec}(\mathbf{M}_b)\|_2 \le \sqrt{P_b}, \qquad \forall b \right.
\end{aligned} \tag{4}$$

We note that the formulation in (7) is intrinsically intractable owing to its semi-infinite nature i.e., finite optimization variables and infinite constraints.

Remark 1: It is worth mentioning here that for the case of receivers equipped with multiple antennas (and hence the possibility of transmitting multiple data streams), an option could be to employ receiver combining matrix and study the balancing problem on per stream basis. The problem is then exactly similar to our case for a given receiver processing matrix per user. Here, we stress that even in the presence of perfect CSI, joint transmitter-receiver processing matrix design is a difficult nonconvex problem.

III. Robust Beamformer Designs for Imperfect CSI

A. SDP-Based Robust Design

The solution to problem (7) significantly depends on the type of the uncertainty set. In the context of SDP design, the most commonly considered uncertainty model is the one where error vectors $\mathbf{v}_{b,k}$ are bounded in a Euclidean ball of radius $\rho_{b,k}$, i.e., the norm involved in $\mathcal{S}_{b,k}$ [cf. (6)] is simply an l_2 norm. Based on the affine model in (5), $\mathbf{h}_{b,k}$ is assumed to lie in an ellipsoid centered at $\hat{\mathbf{h}}_{b,k}$, which is characterized by $\rho_{b,k}$ and $\mathbf{A}_{b,k}$. This type of channel error model is commonly known as ellipsoidal uncertainty model in the related literature. For practical channel estimation schemes it is known that the channel estimation error follows Gaussian distribution [18]. Most of the probability content of multi-dimensional Gaussian density is localized in its certain region. This clearly motivates modeling the error using an ellipsoid. Further, when vector quantization is used at the receiver, quantization errors can also be approximated by ellipsoids [19]. For this specific uncertainty set, we will show that the robust counterpart in (7) can be solved using SDP approximations. We note that in (7) the uncertain part of $\mathbf{h}_{b_k,k}$ varies in the same set on both sides of the inequality. It is well known that this renders the problem intractable [4], [8], [16]. However, we will see that, after a suitable relaxation, this constraint can be written in a tractable form. To this end, we define $\mathbf{P}_k = \mathbf{m}_k\mathbf{m}_k^{\mathrm{H}}$, $\mathbf{Q}_b = \sum_{k\in\mathcal{U}_b}\mathbf{P}_k$ and $\mathbf{W}_k = (\alpha_{b_k,k}/t)\mathbf{P}_k - \sum_{i\in\mathcal{U}_{b_k}\setminus k}\mathbf{P}_i$. Then by using slack variables we observe that the constraints involving perturbed channels can be cast into tractable linear matrix inequalities (LMIs) using the so called $\mathcal{S}$-Lemma or $\mathcal{S}$-Procedure [8], [20]. After some manipulations, (7) can be equivalently rewritten as (8), shown at the bottom of the page. Therefore, after ignoring the nonconvex rank constraints, bisection search over t can be used to obtain covariance matrices. However, we cannot guarantee optimality (see [21]) of the proposed solution if we obtain the rank of precoding matrices greater than unity. Similar rank relaxation approach was also adopted in the recent works, e.g., [5]. We may need some randomization procedure [22] to extract the beamformer $\mathbf{m}_k$ if $\mathrm{rank}(\mathbf{P}_k) > 1$. Nonetheless, randomization trick may not always be useful [5].

B. SOCP-Based Robust Scheme

The robust counterpart of any optimization problem can potentially pose two issues, one related to its tractability, and the other related to its complexity. Very often, the worst case principle leads to an intractable problem, since, as noted previously, the robust counterpart is an optimization problem over an infinite set of constraints. Furthermore, commonly employed approximation schemes usually increase the complexity of the original problem by one degree, i.e., a linear program becomes an SOCP and an SOCP transforms to an SDP. In what follows, we propose a robust design which is merely based on iteratively solving SOCPs, i.e., we attempt to minimize the complexity of the robust version of the balancing problem. Interestingly enough, we note that the SOCP-based scheme can also encompass a wide variety of uncertainty sets. In order to emphasize the capability of the SOCP scheme to handle a variety of uncertainty sets, we will not specify any particular norm to represent the uncertainty set.

We will arrive at a reduced complexity tractable robust scheme by incorporating uncertainty and exploiting the structure of the SOC constraints in (4). To start with, we consider a relaxation of (7), which is written as (9), shown at the bottom of the following page, where $C_k = \sqrt{1+(\alpha_{b_k,k}/t)}$, $\mathbf{m}_k \in \mathbb{C}^{T\times 1}$, $z_{b,k}\in\mathbb{R}$, $\mathbf{M}_b = [\mathbf{m}_{\mathcal{U}_b(1)},\ldots,\mathbf{m}_{\mathcal{U}_b(|\mathcal{U}_b|)}]$ are optimization variables and we have removed the absolute value, and only consider the real part of the left side of the constraints in (9a).

$$\underset{\mathbf{P}_k,t_{b,k},\lambda_{b,k},\tau_k}{\text{maximize}} \quad t$$

$$\text{s.t.} \quad \begin{pmatrix} \mathbf{A}_{b_k,k}\mathbf{W}_k\mathbf{A}_{b_k,k}^{\mathrm{H}} + \lambda_{b_k,k}\mathbf{I} & \mathbf{A}_{b_k,k}\mathbf{W}_k\hat{\mathbf{h}}_{b_k,k}^{\mathrm{H}} \\ \hat{\mathbf{h}}_{b_k,k}\mathbf{W}_k\mathbf{A}_{b_k,k}^{\mathrm{H}} & \hat{\mathbf{h}}_{b_k,k}\mathbf{W}_k\hat{\mathbf{h}}_{b_k,k}^{\mathrm{H}} - \tau_k - \lambda_{b_k,k}\rho_{b_k,k}^2 \end{pmatrix} \succeq 0, \forall\, k \tag{8a}$$

$$\begin{pmatrix} -\mathbf{A}_{b,k}\mathbf{Q}_b\mathbf{A}_{b,k}^{\mathrm{H}} + \lambda_{b,k}\mathbf{I} & -\mathbf{A}_{b,k}\mathbf{Q}_b\hat{\mathbf{h}}_{b,k}^{\mathrm{H}} \\ -\hat{\mathbf{h}}_{b,k}\mathbf{Q}_b\mathbf{A}_{b,k}^{\mathrm{H}} & -\hat{\mathbf{h}}_{b,k}\mathbf{Q}_b\hat{\mathbf{h}}_{b,k}^{\mathrm{H}} + t_{b,k} - \lambda_{b,k}\rho_{b,k}^2 \end{pmatrix} \succeq 0\ \forall\, k, b\neq b_k \tag{8b}$$

$$\sum_{k\in\mathcal{U}_b}\mathrm{trace}(\mathbf{P}_k) \le P_b, \forall\, b, \quad \sum_{b=1,b\neq b_k}^{B} t_{b,k} + \sigma^2 \le \tau_k, \tau_k \ge 0\ \forall\, k \tag{8c}$$

$$\begin{cases} \lambda_{b,k} \ge 0\ \forall\, b,k, t_{b,k} \ge 0\ \forall\, k, b\neq b_k, \\ \mathbf{P}_k \succeq 0, \mathrm{rank}(\mathbf{P}_k) = 1\ \forall\, k. \end{cases} \tag{8d}$$

Unlike the non-robust version of the problem (4), we cannot force the imaginary part of $[\hat{\mathbf{h}}_{b_k,k}+\sum_{i=1}^{l_{b_k,k}}\boldsymbol{\delta}^i_{b_k,k}[\mathbf{v}_{b_k,k}]_i]\mathbf{m}_k$ to zero for all channel error realizations. Since for a complex number x, $|x|\geq\Re(x)$, a feasible point for (9a) is also feasible for the exact robust counterpart given in (7). That is to say, to arrive at a reduced complexity approach, we consider a conservative approximation of the exact robust counterpart of the problem. For notational simplicity, when clear from context, we avoid mentioning the real operator $\Re(\cdot)$ explicitly from this point onwards.

Now we make a key manipulation by substituting $\mathbf{v}_{b,k}=\boldsymbol{\theta}_{b,k}-\boldsymbol{\phi}_{b,k}$ for all b, k, in (9b) and $\Re(\mathbf{v}_{b_k,k})=\Re(\boldsymbol{\theta}_{b_k,k})-\Re(\boldsymbol{\phi}_{b_k,k})$ such that $\Re(\boldsymbol{\theta}_{b_k,k})\geq 0$ and $\Re(\boldsymbol{\phi}_{b_k,k})\geq 0$ in (9a). After this we obtain an approximation of (9)

$$\begin{aligned}\text{maximize}\quad & t\\ \text{s.t.}\quad & C_k\Bigg(\Re\left(\hat{\mathbf{h}}_{b_k,k}\mathbf{m}_k\right)+\sum_{i=1}^{l_{b_k,k}}\Re\left(\boldsymbol{\delta}^i_{b_k,k}\mathbf{m}_k\right)\\ &\qquad\times\left(\Re\left([\boldsymbol{\theta}_{b_k,k}]_i\right)-\Re\left([\boldsymbol{\phi}_{b_k,k}]_i\right)\right)\Bigg)\\ &\quad-\left\|(z_{1,k}\dots z_{B,k}\sigma)^{\mathrm{T}}\right\|_2\geq 0,\forall k,\\ &\quad\forall\boldsymbol{\theta}_{b_k,k},\boldsymbol{\phi}_{b_k,k}:\|\Re(\boldsymbol{\theta}_{b_k,k})+\Re(\boldsymbol{\phi}_{b_k,k})\|\leq\rho'_{b_k,k}\end{aligned}\tag{10a}$$

$$\begin{aligned}& z_{b,k}-\left\|\mathbf{M}_b^{\mathrm{H}}\left[\hat{\mathbf{h}}_{b,k}+\sum_{i=1}^{l_{b,k}}\boldsymbol{\delta}^i_{b,k}\right.\right.\\ &\qquad\left.\left.\times\left([\boldsymbol{\theta}_{b,k}]_i-[\boldsymbol{\phi}_{b,k}]_i\right)\right]^{\mathrm{H}}\right\|_2\\ &\geq 0,\forall b,k,\forall\boldsymbol{\theta}_{b,k},\boldsymbol{\phi}_{b,k}:\left\||\boldsymbol{\theta}_{b,k}|+|\boldsymbol{\phi}_{b,k}|\right\|\leq\rho'_{b,k},\end{aligned}\tag{10b}$$

$$z_{b,k}\geq 0,\forall b,k,\quad \|\mathrm{vec}(\mathbf{M}_b)\|_2\leq\sqrt{P_b},\qquad\forall b,\tag{10c}$$

where $[\boldsymbol{\theta}_{b_k,k}]_i,[\boldsymbol{\phi}_{b_k,k}]_i$ denote the ith components of $\boldsymbol{\theta}_{b_k,k}$, $\boldsymbol{\phi}_{b_k,k}$ for all b, k, respectively, and for a vector $\mathbf{y}$ the symbol $|\mathbf{y}|$ represents that $[|\mathbf{y}|]_i=|[\mathbf{y}]_i|$ for all i. Another change introduced in (10) is that for all b, k we have replaced $\rho_{b_k,k}$ with $\rho'_{b_k,k}$. The motivation for this variation of the uncertainty set parameter $\rho_{b_k,k}$ becomes clear as we outline the fact that splitting the uncertainty vector $\mathbf{v}_{b_k,k}$ into a difference of two vectors and manipulating the left side of uncertainty sets as done in (10) transforms the problem into *nearly* a *safe approximation* of its original version while also rendering it *tractability*.

Remark 2: As noted earlier, it appears difficult, if not impossible, to cast the worst case robust counterpart (7) into its exact equivalent tractable formulation. For example, in the first approach based on SDP formulation, we had to drop the unit rank constraints to arrive at a tractable representation. Naturally, to arrive at an SOCP representation of the problem, we have to resort to an approximation of the original feasible set i.e.,

$$\begin{aligned}\mathcal{O}_{orig}=\{&\text{Optimization variables in (7) such that}\\ &\text{all constraints in (7) are satisfied}\}\end{aligned}\tag{11}$$

with its *tractable subset* that may also include some additional analysis variables.

In our case, it should be a feasible set for an SOC problem. By doing so we can ensure that a solution for the approximation is definitely feasible for the original optimization program, and thus promises *safety* in a sense that we do not violate the original constraints. By just considering the real part of the left side of constraints in (9a), we follow this strategy. However, alone this move can be rather conservative. By including additional terms, this conservatism can be compensated to some extent. Particularly, the sum of terms in (10a) is an approximation for an exact equivalent of real of a product of two complex numbers.[1] Hence, a decrease in conservatism is expected. Although no guarantee on the safety can be promised analytically, our numerical results reveal that this neutralization of the conservative real approximation is almost safe from the perspective defined above.

Next, let us first focus on the set of constraints in (10b) and rewrite it in a form similar to the one presented in (9b)

$$\begin{aligned}z_{b,k}-&\left\|\mathbf{M}_b^{\mathrm{H}}\left[\hat{\mathbf{h}}_{b,k}+\sum_{i=1}^{l_{b,k}}\boldsymbol{\delta}^i_{b,k}[\mathbf{v}_{b,k}]_i\right]^{\mathrm{H}}\right\|_2\\ &\geq 0,\forall b,k,\forall\mathbf{v}_{b,k}:\|\mathbf{v}_{b,k}\|\leq\rho'_{b,k}.\end{aligned}\tag{12}$$

[1]Recall that if $c_1=a_1+jb_1$, $c_2=a_2+jb_2$, then $\Re(c_1c_2)=\Re(c_1)\Re(c_2)-\Im(c_1)\Im(c_2)$.

$$\begin{aligned}&\text{maximize } t\\ &\text{s.t.}\begin{cases}C_k\Re\left(\left[\hat{\mathbf{h}}_{b_k,k}+\sum_{i=1}^{l_{b_k,k}}\boldsymbol{\delta}^i_{b_k,k}[\mathbf{v}_{b_k,k}]_i\right]\mathbf{m}_k\right)\\ \quad-\left\|(z_{1,k}\dots z_{B,k}\sigma)^{\mathrm{T}}\right\|_2\geq 0,\forall\, k,\\ \forall\,\mathbf{v}_{b_k,k}:\|\mathbf{v}_{b_k,k}\|\leq\rho_{b_k,k}\end{cases}\end{aligned}\tag{9a}$$

$$\begin{cases}z_{b,k}-\left\|\mathbf{M}_b^{\mathrm{H}}\left[\hat{\mathbf{h}}_{b,k}+\sum_{i=1}^{l_{b,k}}\boldsymbol{\delta}^i_{b,k}[\mathbf{v}_{b,k}]_i\right]^{\mathrm{H}}\right\|_2\geq 0,\ \forall\, b,k,\\ \forall\,\mathbf{v}_{b,k}:\|\mathbf{v}_{b,k}\|\leq\rho_{b,k},\end{cases}\tag{9b}$$

$$z_{b,k}\geq 0,\forall\, b,k,\quad\|\mathrm{vec}(\mathbf{M}_b)\|_2\leq\sqrt{P_b},\qquad\forall\, b,\tag{9c}$$

It is worthy making an important observation now.

Proposition 1: A set of optimization variables satisfies the constraints in (10b) if and only if it satisfies the set of constraints in (12).

Proof: The proof of this statement is deferred to Appendix A. ■

The main goal of the development so far is to approximate (10a) and (10b) by SOC constraints so that the resulting robust counterparts in (10a)–(10c) can be cast as an SOCP for fixed t. Since (10a) and (10b) have the same form, it is sufficient to concentrate on tackling the more difficult set of constraints in (10b). We use the concavity of the negative norm to bound (10b) from below as

$$z_{b,k} - \left\|\mathbf{M}_b^{\mathrm{H}}\hat{\mathbf{h}}_{b,k}^{\mathrm{H}}\right\|_2 - \left\|\mathbf{M}_b^{\mathrm{H}}\left[\sum_{i=1}^{l_{b,k}}\boldsymbol{\delta}_{b,k}^i\left([\boldsymbol{\theta}_{b,k}]_i - [\boldsymbol{\phi}_{b,k}]_i\right)\right]^{\mathrm{H}}\right\|_2 \geq 0,$$
$$\forall b,k,\forall \boldsymbol{\theta}_{b,k},\boldsymbol{\phi}_{b,k} : \||\boldsymbol{\theta}_{b,k}| + |\boldsymbol{\phi}_{b,k}|\| \leq \rho'_{b,k}. \quad (13)$$

Again using the concavity argument, the left side of constraints in (13) can be further lower bounded as

$$\begin{aligned} z_{b,k} &- \left\|\mathbf{M}_b^{\mathrm{H}}\hat{\mathbf{h}}_{b,k}^{\mathrm{H}}\right\|_2 \\ &+ \sum_{i=1}^{l_{b,k}}\Big[-\left\|\left\{\boldsymbol{\delta}_{b,k}^i\mathbf{M}_b\right\}^{\mathrm{H}}[\boldsymbol{\theta}_{b,k}]_i\right\|_2 \\ &\qquad - \left\|\left\{\boldsymbol{\delta}_{b,k}^i\mathbf{M}_b\right\}^{\mathrm{H}}\left(-[\boldsymbol{\phi}_{b,k}]_i\right)\right\|_2\Big] \\ &\geq 0, \forall b,k\forall\boldsymbol{\theta}_{b,k},\boldsymbol{\phi}_{b,k} : \||\boldsymbol{\theta}_{b,k}| + |\boldsymbol{\phi}_{b,k}|\| \leq \rho'_{b,k}. \quad (14) \end{aligned}$$

Reading the inequalities from (14) backwards, and recalling the equivalence of (10b) and (12), we observe that a solution of (14) is also feasible for (12) or (9b). We note that while we have to bound the the norm terms as above, we do not need to use any such bound in case of (10a) as it is a simple linear constraint and can be readily cast into the desired form.

Remark 3: Before presenting a tractable formulation of the constraints in (14), we again note that the optimal solution of the proposed SOCP relaxation ideally should also be feasible for the original worst case robust counterpart in (7). Therefore, if $\mathcal{O}_{socp}$ represents the feasible set for the SOCP relaxation, then $\mathcal{O}_{socp} \subseteq \mathcal{O}_{orig}$ should hold. This will imply both safety and tractability for the proposed SOCP approximation. Although the transformations that will lead to an SOCP formulation for (14) ensure both these factors, the same, on account of the reasons mentioned above, cannot be observed for the constraint in (10a). Nevertheless, we will see in Section IV that the relation $\mathcal{O}_{socp} \subseteq \mathcal{O}_{orig}$ almost remains valid at least for the cases considered. Being nearly a subset of the original problem, the proposed approximation can be rather conservative, as also noted in [16], [23]. To provide additional flexibility in this regard, and as noted above, we make $\rho_{b,k}$ a design parameter and replace it by $\rho'_{b,k}$ in (10). With the introduction of this maneuver, we may be able to improve the achieved objective, albeit this may come at the cost of degradation in achieving it for given realizations of channel errors as we probe in the results section.

Let us define

$$\begin{cases} f_1(\mathbf{M}_b, z_{b,k}, \hat{\mathbf{h}}_{b,k}) \triangleq z_{b,k} - \|\mathbf{M}_b^{\mathrm{H}}\hat{\mathbf{h}}_{b,k}^{\mathrm{H}}\|_2, \\ f_2(\mathbf{M}_b, \boldsymbol{\delta}_{b,k}^i) \triangleq -\left\|\{\boldsymbol{\delta}_{b,k}^i\mathbf{M}_b\}^{\mathrm{H}}\right\|_2. \end{cases} \quad (15)$$

We note that $f_2(\mathbf{M}_b, \boldsymbol{\delta}_{b,k}^i) = f_2(\mathbf{M}_b, -\boldsymbol{\delta}_{b,k}^i)$. With the above definitions and using the fact that $\|k\mathbf{m}\|_2 = |k|\|\mathbf{m}\|_2$, the constraints in (14), for all b, k, can be equivalently written as

$$\begin{aligned} f_1(\mathbf{M}_b, z_{b,k}, \hat{\mathbf{h}}_{b,k}) + \min_{\||\boldsymbol{\theta}_{b,k}|+|\boldsymbol{\phi}_{b,k}|\| \leq \rho'_{b,k}} &\sum_{i=1}^{l_{b,k}}\Big\{f_2(\mathbf{M}_b, \boldsymbol{\delta}_{b,k}^i)|[\boldsymbol{\theta}_{b,k}]_i| \\ &+ f_2(\mathbf{M}_b, -\boldsymbol{\delta}_{b,k}^i)|[\boldsymbol{\phi}_{b,k}]_i|\Big\} \geq 0 \quad (16) \end{aligned}$$

where we emphasize that the first term is independent of $\boldsymbol{\phi}_{b,k}$ and $\boldsymbol{\theta}_{b,k}$, and hence the constrained minimization appears across the second term only. The constraint in (16) can be cast into tractable form using [16, Theorem 1], which is stated as:

Theorem 1: Working in the real domain, given a function $f(\mathbf{x}, \mathbf{U})$ that is concave in data $\mathbf{U}$ for all given $\mathbf{x}$ and scales linearly with the data, we consider a constraint of the following form:

$$\begin{aligned} \min_{\mathbf{u}_1,\mathbf{u}_2 \geq 0 : \|\mathbf{u}_1+\mathbf{u}_2\| \leq \omega} & f(\mathbf{x}, \mathbf{U}^n) \\ &+ \sum_j \Big[f(\mathbf{x}, \boldsymbol{\delta}^j)[\mathbf{u}_1]_j + f(\mathbf{x}, -\boldsymbol{\delta}^j)[\mathbf{u}_2]_j\Big] \geq 0 \quad (17) \end{aligned}$$

where $\mathbf{U}^n$ is the nominal part of the data, $\boldsymbol{\delta}^j$ is a vector representing perturbation direction in the jth component of the data and $\mathbf{u}_1$ and $\mathbf{u}_2$ are real vectors of appropriate dimensions and the norm in (17) satisfies the property [16, Eq. 6]

$$\|\mathbf{u}\| = \||\mathbf{u}|\| \quad (18)$$

where $|\mathbf{u}| = (|u_1|, \ldots, |u_d|)$. Notice that first term in (17) is constant with respect to $\mathbf{u}_1$ and $\mathbf{u}_1$, and the minimization is over the second term only. The constraint (17) admits an equivalent representation of the form $f(\mathbf{x}, \mathbf{U}^n) \geq \omega\|\boldsymbol{\gamma}\|^\star$, where $[\boldsymbol{\gamma}]_j = \max\{-f(\mathbf{x}, \boldsymbol{\delta}^j), -f(\mathbf{x}, -\boldsymbol{\delta}^j)\} \geq 0$ and $\|\boldsymbol{\gamma}\|^\star \triangleq \max_{\|\mathbf{s}\|\leq 1}\mathbf{s}^{\mathrm{T}}\boldsymbol{\gamma}$ is the dual norm of $\boldsymbol{\gamma}$.

Proof: The proof of the theorem is available in [16, Theorem 1]. However, for the sake of completeness and for demonstrating its applicability on (16) it is relegated to the Appendix B. ■

It should be emphasized that the norm in (17) can be arbitrary, as long as it satisfies (18), meaning that the proposed SOCP-based scheme presented next is applicable to a wide variety and combinations of norms and thus uncertainty sets. For the special case of l_2 norm, the norms remain l_2 because of the self dual property of the l_2 norm. Following similar steps used to tackle the constraints in (10b), we can easily see that the uncertain constraints in (10a) can be cast in a form that is amenable to applying Theorem 1. Specifically, notice that the constraints in (10a) can be straightforwardly written in a

functional form as required by Theorem 1 i.e., as the following minimization over $\boldsymbol{\theta}_{b_k,k}$, $\boldsymbol{\phi}_{b_k,k}$

$$\begin{aligned}\min_{\left\|\Re(\boldsymbol{\theta}_{b_k,k})+\Re\left(\boldsymbol{\phi}_{b_k,k}\right)\right\|\leq\rho'_{b_k,k}} & \\ \times\Bigg(\Re\left(\hat{\mathbf{h}}_{b_k,k}\mathbf{m}_k\right)+\sum_{i=1}^{l_{b_k,k}}\Re\left(\boldsymbol{\delta}^i_{b_k,k}\mathbf{m}_k\right)\Re\left([\boldsymbol{\theta}_{b_k,k}]_i\right) & \\ +\Re\left(-\boldsymbol{\delta}^i_{b_k,k}\mathbf{m}_k\right)\Re\left(\left[\phi_{b_k,k}\right]_i\right)\Bigg) & \\ \geq\left\|(z_{1,k}\ldots z_{B,k}\sigma)^{\mathrm{T}}\right\|_2,\forall k. & \end{aligned} \tag{19}$$

In addition, we remark that the equivalent formulation in Theorem 1 i.e., $f(\mathbf{x},\mathbf{U}^n)\geq\omega\|\boldsymbol{\gamma}\|^\star$ can also be written as a system of following constraints:

$$\begin{cases} f(\mathbf{x},\mathbf{U}^n)\geq\omega\eta,\|\boldsymbol{\eta}\|^\star\leq\eta,\ f(\mathbf{x},\boldsymbol{\delta}^i)+[\boldsymbol{\eta}]_i\geq 0,\ \forall\, i,\\ f(\mathbf{x},-\boldsymbol{\delta}^i)+[\boldsymbol{\eta}]_i\geq 0,\ \forall\, i\end{cases} \tag{20}$$

where η and $\boldsymbol{\eta}$ are auxiliary analysis variables. In the formulations to follow, we will adopt this convention. We remark that while obtaining the equivalent formulation of (10a) using this method, we follow the same strategy of approximating the modulus of a complex number with its real part. As mentioned earlier, we avoid showing this operation explicitly. The performance of the proposed approximation is tested in Section IV. However, some important observations should be re-stressed at this point. Although the constraints in (10a) are linear, the solution of this approximation does not necessarily imply (9a). This differs from the previous scenario where the conversion of the constraints in (9b)–(10b) is safe. Therefore, obtaining a safe, tractable and least possible conservative version of (9a) is left as an open question for future research.

With the aid of Theorem 1, the approximate robust counterpart of the original problem can be written in the following tractable form:

$$\text{maximize } t$$

$$\text{s.t. } C_k\hat{\mathbf{h}}_{b_k,k}\mathbf{m}_k-\left\|(z_{1,k}\ \cdots\ z_{B,k}\ \sigma)^{\mathrm{T}}\right\|_2 \geq\rho'_{b_k,k}L_{b_k,k},\forall k \tag{21a}$$

$$C_k\boldsymbol{\delta}^q_{b_k,k}\mathbf{m}_k+[\mathbf{q}_{b_k,k}]_q\geq 0,\forall k,q=1,\ldots,l_{b_k,k} \tag{21b}$$

$$\begin{cases}-C_k\boldsymbol{\delta}^q_{b_k,k}\mathbf{m}_k+[\mathbf{q}_{b_k,k}]_q\geq 0,\forall k,q=1,\ldots,l_{b_k,k},\\ \|\mathbf{q}_{b_k,k}\|\leq L_{b_k,k},\forall k\end{cases} \tag{21c}$$

$$z_{b,k}-\|\mathbf{M}_b^{\mathrm{H}}\hat{\mathbf{h}}_{b,k}^{\mathrm{H}}\|_2\geq\rho'_{b,k}\nu_{b,k}, \tag{21d}$$

$$\begin{cases}-\left\|\{\boldsymbol{\delta}^i_{b,k}\mathbf{M}_b\}^{\mathrm{H}}\right\|_2+[\boldsymbol{\mu}_{b,k}]_i\geq 0,\\ \quad i=1\ldots l_{b,k},\forall b,k,\\ \|\boldsymbol{\mu}_{b,k}\|\leq\nu_{b,k},\forall b,k\end{cases} \tag{21e}$$

$$\|\mathrm{vec}(\mathbf{M}_b)\|_2\leq\sqrt{P_b},\qquad\forall b \tag{21f}$$

where $\mathbf{m}_k\in\mathbb{C}^{T\times 1}$, $\mathbf{M}_b=[\mathbf{m}_{\mathcal{U}_b(1)},\ldots,\mathbf{m}_{\mathcal{U}_b(|\mathcal{U}_b|)}]$, $z_{b,k}\in\mathbb{R}$, $L_{b_k,k}\in\mathbb{R}$, $\nu_{b,k}\in\mathbb{R}$, $\mathbf{q}_{b_k,k}\in\mathbb{C}^{l_{b_k,k}}$, $\boldsymbol{\mu}_{b,k}\in\mathbb{C}^{l_{b_k,k}}$ are optimization variables. The above optimization problem represents a tractable approximation, in the form of SOCP in conjunction with bisection search, of the robust counterpart of the problem under consideration. In the following, we provide some remarks regarding the tractability and reduced complexity of the proposed SOCP-based robust design.

Tractability: We emphasize that while the SDP-based solution is only applicable to ellipsoidal uncertainty models, the SOCP-based approach is flexible enough to deal with other types of uncertainty sets. For example, in certain situations, the errors in each of the individual terms of the channel vector are bounded i.e., $|[\mathbf{v}_{b,k}]_i|\leq\xi_{b,k}$ for all b, k, i. This amounts to saying that $\|\mathbf{v}_{b,k}\|_\infty\leq\xi_{b,k}$. In fact, in practical systems where each entry of $\mathbf{h}_{b,k}$ is quantized independently at the receiver and fed back to the corresponding transmitters, the interval uncertainty model is more appropriate [24]. Clearly, this uncertainty model can be easily handled with the above approach since the dual of the l_∞ norm is well known [8]. In other situations, it may happen that the entries of the uncertainty vector are symmetrically random and bounded. In such scenarios it is well known that the perturbation set can be represented as the intersection of the l_2 and l_∞ norms of $\mathbf{v}_{b,k}$ [25]. For this uncertainty model, it may be difficult, if not impossible, to straightforwardly use worst case design philosophy. Hence, the SDP-based method is not applicable and problem (7) appears to be intractable. However, the SOCP-based approach admits tractability in the approximate solution of the robust counterpart using the dual of the $l_2\cap l_\infty$ norm [25].

Complexity Reduction: The SOCP-based robust design also offers a great reduction in computational complexity compared to the SDP-based method. In what follows, we give a complexity comparison of the SDP- and SOCP-based solutions for the special case where $\mathbf{A}_{b,k}=\mathbf{I}_T$ for all b, k, which is commonly considered in the related works.[2] First, let us focus on the equivalent representation obtained using Theorem 1 and explore it by considering any robust equivalent constraint (without loss of generality) from (21e). We note that under this setting each entry of a channel can be written as $[\mathbf{h}_{b,k}]_m=[\hat{\mathbf{h}}_{b,k}]_m+[\boldsymbol{\delta}_{b,k}]_m[\mathbf{v}_{b,k}]_m, 1\leq m\leq T$, where $\mathbf{v}_{b,k}$ belongs to the uncertainty set defined in (6). The vector $\boldsymbol{\gamma}_{b,k}$ corresponding to the equivalent formulation of Theorem 1 becomes

$$\boldsymbol{\gamma}^{\mathrm{T}}_{b,k}=\left[\left|\left[\mathbf{m}_{\mathcal{U}_b(1)}\right]_m[\boldsymbol{\delta}_{b,k}]_m\right|,\ldots,\left|\left[\mathbf{m}_{\mathcal{U}_b(|\mathcal{U}_b|)}\right]_m[\boldsymbol{\delta}_{b,k}]_m\right|\right],\ \forall\, b,k. \tag{22}$$

With this type of $\boldsymbol{\gamma}_{b,k}$ it has been shown in [16] that for ellipsoidal uncertainty set, instead of having multiple additional constraints of the type mentioned above, we can stack all corresponding variables into one SOC constraint, $\|\boldsymbol{\mu}_{b,k}\|_2\leq\nu_{b,k}$, and one variable $\nu_{b,k}$ for all b, k. Similarly, the constraints involving user (b_k,k), shown in (21b) and (21c), can be greatly simplified.

To provide a complexity comparison, we base our discussion on the simplification noted above and focus on an arbitrary bisection step. According to (21a)–(21f), the number of real optimization variables per bisection iteration of the SOCP-based robust design is $4TBK+2BK+2KT+K$. More

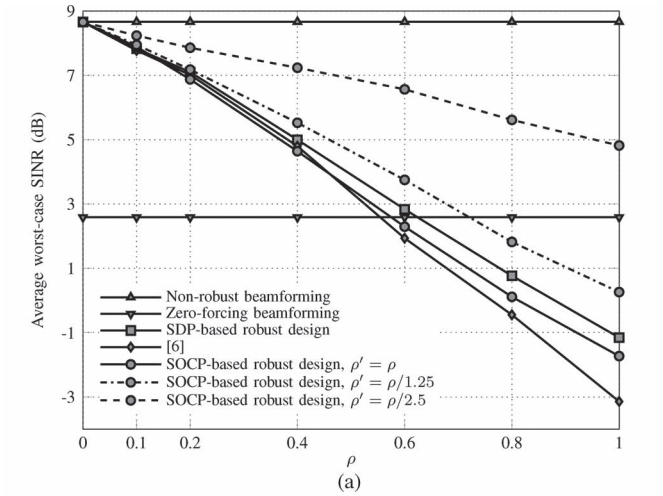
[2]The same arguments in this part also apply to case where the entries of channel vectors undergo independent perturbations.

specifically, there are BK constraints of real dimension $(2TK_b + 1)$ that occur thrice including the power constraint. Again using the above mentioned simplification, we obtain two constraints of real dimensions $B + 2$ and $T + 1$ that are K in number. Combining all these the worst case per iteration complexity of the SOCP approach approximates as $\mathcal{O}(K_b(TBK)^3)$ [26], [27]. The per iteration complexity of the SDP-based method is found to be $\mathcal{O}((KBT)^6)$ [26], [28], which is clearly higher than the SOCP counterpart. Further to this, based on [26], [27], the worst case estimate of the number of iterations needed to arrive at a numerically acceptable value of the SOCP-based design is $\mathcal{O}(\sqrt{KB})$. As similar calculation reveals that such an estimate for the SDP-based method results in a higher value of $\mathcal{O}(\sqrt{KTB})$ on account of its dependence on the size of the matrix inequalities. A more detailed exploration that compares run times of the proposed approaches with different solvers is given in Section IV.

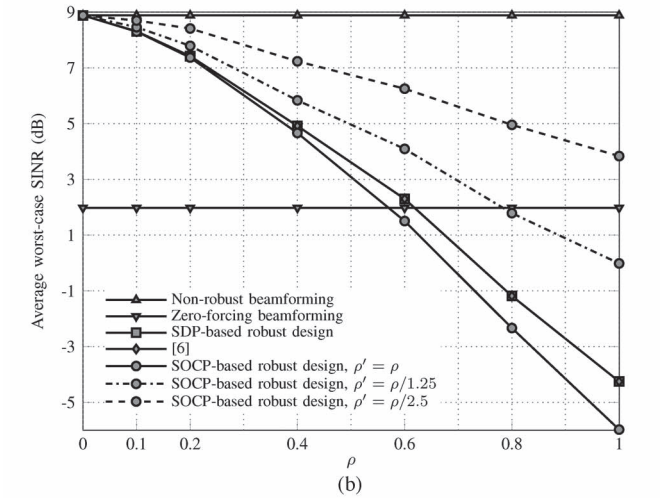
IV. Numerical Results

In order to compare the performance of the proposed approaches we report results of numerical simulations in this section. For all simulation setups, we consider a system of two cells $(B = 2)$, while the number of users per cell is mentioned for individual numerical experiments. The channel vector from the BS b to user k is given by $\mathbf{h}_{b,k} = \sqrt{\kappa_{b,k}}\tilde{\mathbf{h}}_{b,k}$ where $\kappa_{b,k}$ represents both the path loss and the shadow fading and $\tilde{\mathbf{h}}_{b,k}$ follows $\mathcal{CN}(0, \mathbf{I})$. In Figs. 2–4 to follow, we only consider the small-scale fading (i.e., $\kappa_{b,k} = 1$ for all b and k). These setups can be considered to correspond to the worst-case scenario where all users are at the cell edge. A more realistic channel model where large-scale fading is taken into account is investigated in Fig. 6 for a massive MISO system. All noise variances are taken as unity and the transmit power is normalized with respect to the noise variance. For the sake of simplicity, but without compromising generality, we take $\alpha_{b_k,k} = 1$ for all k, $\mathbf{A}_{b,k} = \mathbf{I}_T$, and $\rho_{b,k} = \rho$ for all b, k. Unless otherwise mentioned, the error vectors are assumed to lie in a hypersphere of radius ρ. We evaluate the performance of the three approaches in terms of the worst-case SINR [i.e., the objective obtained at the end of the bisection procedure when solving (8) and (21)] and the probability of exceeding the worst-case SINR which is referred to as PE from now on. For the simulation setup considered in this paper, unless otherwise mentioned, the SDP-based approach is numerically found to produce precoding matrices close to rank-1 matrices.

Fig. 2 plots the average worst-case SINRs (over 200 realizations of the nominal channels $\hat{\mathbf{h}}_{b,k}$) versus the radius of the uncertainty sets, ρ, for all approaches. In this simulation setup, we only consider the small-scale fading (i.e., $\kappa_{b,k} = 1$ for all b and k). We consider the cases when the number of users in each cell is 2 and 1, i.e., there are $K = 4$ and $K = 2$ users, respectively. It is seen in Fig. 2(a) that the SOCP-based solution gives the worst-case SINR close to that of the SDP-based approach and slightly higher than that of [6] when $\rho' = \rho$ especially for larger ρ. It can be seen from Fig. 2 that the worst-case SINR of the SOCP-based solution is improved when we take $\rho' = \rho/2.5$, but this implies reduced PE as indicated in



(a)



(b)

	ρ	0.1	0.2	0.4	0.6	0.8	1
$B = 2, K = 4$ (two users per cell)	PE non-robust	0	0	0	0	0	0
	PE zero-forcing	0	0	0	0	0	0
	PE SDP	1	1	1	1	1	1
	PE [6]	1	1	1	1	1	1
	PE ($\rho' = \rho$)	1	1	1	1	1	1
	PE ($\rho' = \rho/1.25$)	0.99	1	0.99	0.99	0.98	0.99
	PE ($\rho' = \rho/2.5$)	0.81	0.87	0.82	0.75	0.80	0.73
$B = 2, K = 2$ (one user per cell)	PE non-robust	0	0	0	0	0	0
	PE zero-forcing	0	0	0	0	0	0
	PE SDP	1	1	1	1	0.99	0.98
	PE [6]	1	1	1	1	1	1
	PE ($\rho' = \rho$)	0.98	0.99	0.99	0.98	0.98	0.98
	PE ($\rho' = \rho/1.25$)	0.94	0.93	0.94	0.93	0.93	0.93
	PE ($\rho' = \rho/2.5$)	0.81	0.85	0.84	0.79	0.74	0.72

(c)

Fig. 2. Average worst-case SINR versus ρ for different approaches where channel uncertainties are bounded by l_2-norm. The value of power for both BSs is taken as 5 dB. The number of transmit antennas at each BS is $T = 8$. (a) Average worst-case SINR versus ρ. The total number of users is $K = 4$ (2 users per base station). (b) Average worst-case SINR versus ρ. The total number of users is $K = 2$ (1 users per base station). (c) Variation of PE with ρ for the proposed SOCP-based robust design.

Fig. 2(c). The values of PE given in Fig. 2(c) are obtained with 10^6 realization of channel errors that are uniformly distributed in the ball of ρ using the toolbox of [29]. However, in the case of $K = 2$ users shown in Fig. 2(b), compared to the optimal result in [6], the proposed SOCP design is rather conservative for large ρ in the $\rho = \rho'$ case. We also see that in both Fig. 2(a) and (b), consistent with the observation made in [6], for the given number of antennas the zero-forcing beamforming yields lower SINR compared to the non-robust design. Remarkably, unlike

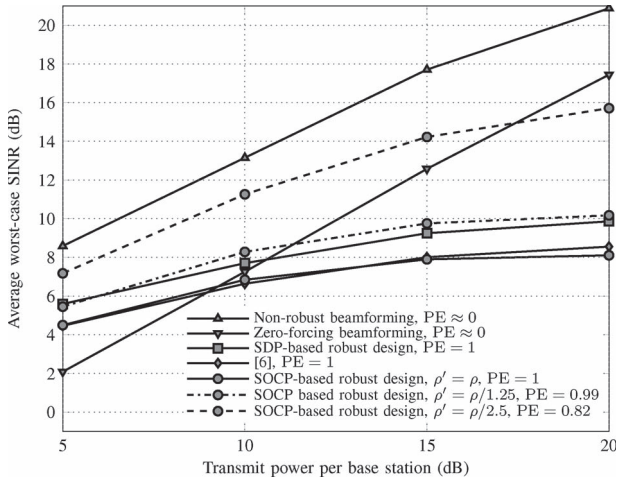
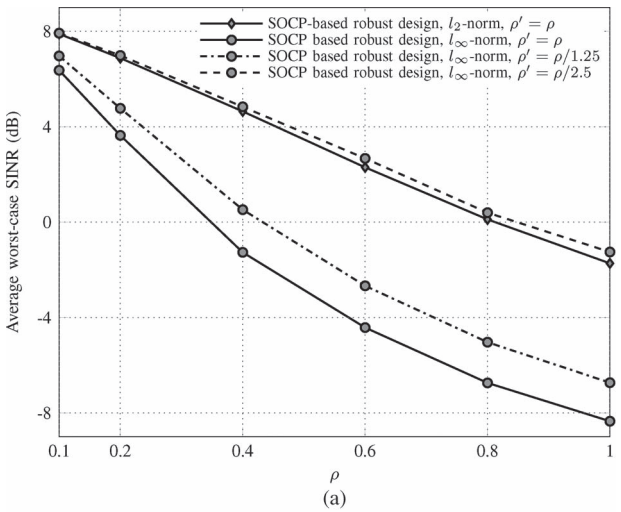


Fig. 3. Comparison of the worst-case SINR of non-robust, SOCP-, SDP-based design, and the approach of [6] as a function of the transmit power of the BSs for $\rho = 0.4$. The performances of the SOCP-based method are shown for three cases where $\rho' = \rho$, $\rho' = \rho/1.25$, and $\rho' = \rho/2.5$. We take $T = 8$ and $K = 4$ (2 users per base station).



(a)

ρ	0.1	0.2	0.4	0.6	0.8	1
PE ($\rho' = \rho$)	1	1	1	1	0.96	0.98
PE ($\rho' = \rho/1.25$)	1	1	0.99	0.99	0.95	0.97
PE ($\rho' = \rho/2$)	1	0.97	0.96	0.95	0.95	0.91
PE ($\rho' = \rho/2.5$)	0.8	0.88	0.81	0.81	0.81	0.80

(b)

Fig. 4. Variation of average worst-case SINR versus ρ for l_∞-norm (i.e., "box" uncertainty). The value of power for both BSs has been taken as 5 dB. The number of transmit antennas at each BS is $T = 8$. The total number of users is $K = 4$. (a) Variation of average worst-case SINR versus ρ for l_∞-norm. (b) Variation of PE with ρ for l_∞-norm for the proposed SOCP-based robust approach.

the case presented in Fig. 2(a), the SDP approach in Fig. 2(b) did not produce precoding matrices for both users that are numerically close to rank 1. Hence, in the Fig. 2(c), the PE metric shows a degraded value in this case for large ρ. As expected, the non-robust approach delivers the maximum SINR, and similar to the zero-forcing strategy, virtually zero PE in all cases.

Although the SDP formulations can offer better worst-case SINR, they are not practically useful for large-scale antenna systems especially from the complexity perspective. Another disadvantage of the SDP approach is its inability to handle various uncertainty sets and limited choice of solvers compared to those for SOCP-based solutions. The flexibility in choosing a solver is important because a general purpose convex programing solver may not be efficient for all problems. We compare the simulation time of the SDP and SOCP-based robust designs using YALMIP [30] with two widely used conic programming solvers (SeDuMi [31] and SDPT3 [32]). Note that the proposed SOCP-based method allows us to make use of GUROBI [33] as a solver as well which is claimed to be very efficient for detecting feasibility of large-scale SOCPs. For the robust SOCP-based design, we use the simplified representation in (22). In Table I, we show the average run time (in seconds) of all robust approaches as a function of the number of transmit antennas, T, for solving the corresponding optimization problem. The bisection procedure terminates if feasibility is detected and the relative difference ϵ of the objectives between two bisection steps is less than or equal to 10^{-2}. The lower threshold of the bisection algorithm is set to 0, while the upper one is equal to the balanced SINR obtained from the non-robust design. The codes are executed on a 64-bit desktop that supports 8 Gbyte RAM and Intel CORE i7. For both solvers, it can be clearly seen that the SOCP-based design requires a lower run time and the difference is considerable as the number of transmit antennas T increases. This observation matches with the theory presented in the subsection on reduced complexity in Section III-B. Moreover, we notice that the SDP-based methods are not capable of producing a solution when $T \geq 50$ due to lack of memory (denoted by a cross mark "$\times$" in Table I). When $T \geq 300$, SeDuMi and SDPT3 are not suitable solvers for the SOCP-based method since they are not able to produce a solution even after several hours (denoted by "$-$" in Table I). We have observed that GUROBI is the most efficient solver for the SOCP-based method in particular for large-scale antenna array systems.

In Fig. 3, the average worst-case balanced SINR (again over 200 realizations of the nominal channels $\hat{\mathbf{h}}_{b,k}$) is plotted with the transmit power per BS, P, for different approaches. We note that the SOCP-based approach performs nearly as good as the SDP one. The reduced minimum SINR of [6] can be probably attributed to the fact that it is completely a conservative approximation of the robust counterpart. Recall that, in our proposed SOCP-based design, we can control the degree of conservatism of the design by finding proper value of ρ'. The values of PE for three approaches are also provided in Fig. 3. Further, being oblivious to channel error vectors, the non robust design delivers the best worst-case SINR. Nonetheless, as expected and seen previously, this comes at the cost of unacceptably low PE, i.e., PE ≈ 0. We have also included a performance curve for the traditional zero-forcing strategy for setting shown in the figure. It is seen, as reportedly previously in [6] as well, that zero forcing does not yield the best SINR. In addition, its PE is also practically zero. It is found that the SOCP-based method gives PE $= 0.99$ for $\rho' = \rho/1.25$, while the approach of [6] and the SDP-based solution both produce PE $= 1.0$. The value of PE for the SOCP-based design is reduced to 0.82 as we set $\rho' = \rho/2.5$. Interestingly, this decrease in PE is accompanied by a corresponding increase in the worst-case SINR, thereby providing a tradeoff between the two parameters. We note that

TABLE I
AVERAGE RUN TIME (IN SECONDS) VERSUS THE NUMBER OF TRANSMIT ANTENNAS, T, AT EACH BS FOR THE ROBUST DESIGNS. THE NUMBER OF BSS IS $B = 2$, EACH SERVING 10 USERS. THE BISECTION PROCEDURE TERMINATES WHEN THE DIFFERENCE BETWEEN THE OBJECTIVE VALUES OF TWO BISECTION STEPS, $\epsilon \leq 10^{-2}$

Antennas	8	12	16	50	100	200	300	400	500
SDP-based design (SDPT3) (sec)	96.48	477.55	5620.3397	×	×	×	×	×	×
SDP-based design (SeDuMi) (sec)	31.44	162.68	684.57	×	×	×	×	×	×
[6] (SDPT3) (sec)	88.34	130.83	240.33	×	×	×	×	×	×
[6] (SeDuMi) (sec)	48.01	61.09	156.78	×	×	×	×	×	×
SOCP-based design (SDPT3) (sec)	1.63	4.04	4.07	28.09	99.23	285.74	–	–	–
SOCP-based design (SeDuMi) (sec)	0.66	1.09	1.23	21.92	51.31	149.68	–	–	–
SOCP-based design (GUROBI) (sec)	1.08	1.95	2.14	12.02	23.66	44.91	47.37	67.46	90.72

the trend of values of PE is observed to be typical for the range of transmit power considered in Fig. 3.

In Fig. 4 we evaluate performance of robust beamforming for SINR balancing where the errors in elements of channel vectors are bounded within a (multi-dimensional) box of size ρ, i.e., $|[\mathbf{v}_{b,k}]_i| \leq \rho$ for all i. This is equivalent to saying that $\|\mathbf{v}_{b,k}\|_\infty \leq \rho$. For this case, we note that the SDP formulations and the approximations used in [6] are not applicable. The curves in Fig. 4 have been obtained by noting the fact that the dual of l_∞-norm is l_1-norm. It is seen that, with box uncertainty (Fig. 4), the worst-case SINR is lower than that in the case of ellipsoidal uncertainty for the same ρ. This can be explained as follows. We note that for a vector $\mathbf{v}$, $\|\mathbf{v}\|_1 \geq \|\mathbf{v}\|_2 \geq \|\mathbf{v}\|_\infty$, which means that for the same ρ, the l_∞-norm defines a smaller feasible set in (21) compared to the l_2-norm. Thus the worst-case SINR for the l_∞-norm uncertainty is lower than that of the l_2-norm. However, when errors are uniformly distributed in a box, we note in the table given in Fig. 4(b) a slight degradation in PE when $\rho' = \rho$. This stems from the fact that the proposed approach is not guaranteed to be safe as discussed earlier in the paper.

Now we move on to exploring the performance of our algorithm when the number of base station antennas is made very large. To do so, in spirit of the work done in [14], we first obtain a genie-aided interference free upper bound to the achievable SINR with our approach. We assume that each user receives an interference free direct signal corrupted with an error vector present in ball of radius $\rho = 0.1$. By making all interference zero in (7), we obtain an average of worst-case SINR over several hundred channel realizations and compare the performance against the proposed scheme as reported in Fig. 5. The difference between the achievable SINR and the genie-sided SNR is enhanced at higher transmit powers (for all values of T shown) owing to the predominant effect of interference on the achievable SINR at these transmit power values. Therefore, interestingly enough, the results reveal that the large-scale antenna array gain is overwhelmed by the interference at larger powers. More importantly, it can also be clearly seen from the results in Fig. 5 that with an increase in the number of transmit antennas, in general, the gap between the average SINR achieved by our algorithm, and that of the upper bound SNR decreases. This shows that our solution performs well and the expected massive antenna array effect is retained even when we take care of the channel uncertainties.

Finally, in Fig. 6(a) we investigate how the balanced SINR of all users scales with the number of transmit antennas. In particular, we consider a system of two cells, each serving

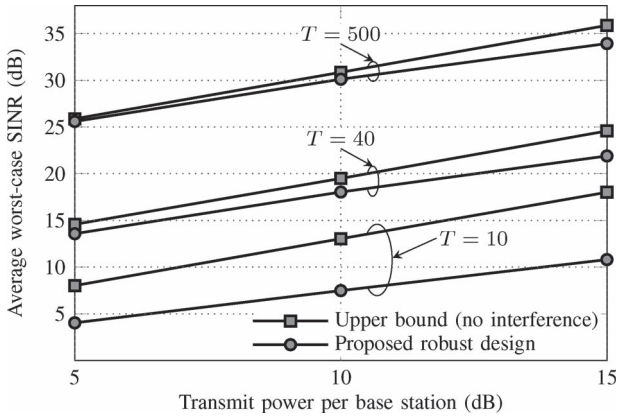


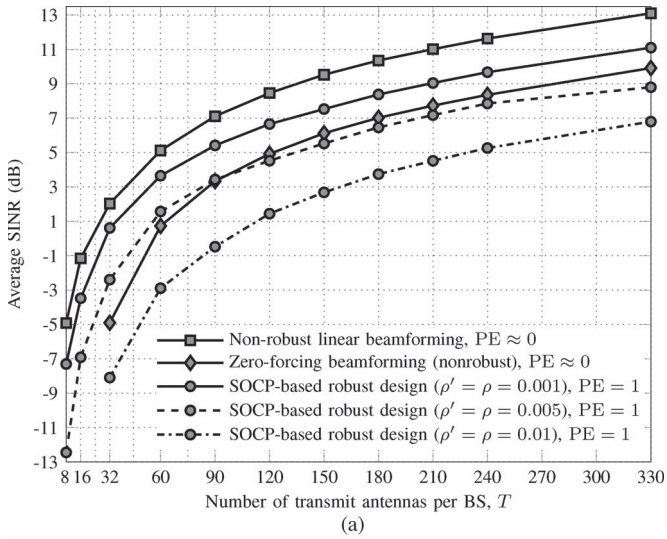
Fig. 5. Comparison of the worst-case SINR of proposed robust SOCP-based design and an upper bound (obtained by ignoring all interference) as a function of the transmit power of the BSs for three values of transmit antennas: $T = 10$, $T = 40$, and $T = 500$. We take $K = 8$ (4 users per base station) and the channel errors are assumed to be in a ball of radius $\rho = 0.1$.

10 users. The users are uniformly distributed in the cell and are not allowed to be closer to the BS by more than $d_0 = 100$ meters [34]. We also assume that the cell diameter (to a vertex) is 1000 meters. The large-scale fading coefficient is modeled as $\kappa_{b,k} = \beta_{b,k}(d_{b,k}/d_0)^{-\nu}$ where $\beta_{b,k}$ accounts for shadow fading assumed to follow a log-normal distribution with standard deviation σ_{shadow}, ν is the path loss exponent, and $d_{b,k}$ is the distance between BS b and user k. In Fig. 6(a), we choose $\sigma_{\text{shadow}} = 8$ dB and $\nu = 3.8$ as in [34]. The performance of zero-forcing beamforming (ZF-BF) scheme is also included for comparison. In ZF-BF, multiuser interference for each user is forced to zero, i.e., $\mathbf{h}_{b_i,j}\mathbf{m}_i = 0$ for all $j \neq i$ [35], [36]. In this way, problem (4) is simplified as

$$\underset{(t,\mathbf{m}_k)\in\mathcal{F}}{\text{maximize}} \quad t \tag{23}$$

where $\mathcal{F} \triangleq \{(t,\mathbf{m}_k) | \mathbf{h}_{b_i,j}\mathbf{m}_i = 0, \quad \forall\, i \neq j, \|\mathbf{M}_b\|_2 \leq \sqrt{P_b}, \forall\, b, \sigma(t/\alpha_{b_k,k}) \leq \Re(\mathbf{h}_{b_k,k}\mathbf{m}_k), \ \forall\, k\}$ which is convex. Consequently, problem (23) is jointly convex in t and $\mathbf{m}_k$. Using the null-space technique as devised in [35], [36], we can further remove the ZF constraints $\mathbf{h}_{b_i,j}\mathbf{m}_i = 0, \ \forall\, i \neq j$ in $\mathcal{F}$ without loss of optimality as follows. Let $\bar{\mathbf{M}}_k = [\mathbf{h}_{b_1,k}^{\mathrm{T}}\, \mathbf{h}_{b_2,k}^{\mathrm{T}} \cdots \mathbf{h}_{b_{k-1},k}^{\mathrm{T}}\, \mathbf{h}_{b_{k+1},k}^{\mathrm{T}} \cdots \mathbf{h}_{b_K,k}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^{(K-1)\times T}$ and $\mathbf{G}_k \in \mathbb{C}^{T\times(T-K+1)}$ be a matrix of orthogonal columns that span the null space of $\bar{\mathbf{M}}_k$.[3] Then, to satisfy the ZF constraints,

[3]For the ZF-BF scheme to be feasible, i.e., $\dim(\mathbf{G}_k) > 0$, we must have $T \geq K$.



(a)

	Zero-forcing beamforming ($\rho = 0.1$)		Proposed SOCP ($\rho' = \rho = 0.1$)	
T/K	SINR (dB)	PE	SINR (dB)	PE
100/2	9.1033	0.015	7.0280	0.99
100/5	3.9776	5.2×10^{-4}	2.4009	0.98
200/2	12.4326	0.12	11.5282	1
200/10	5.4046	10^{-5}	3.9289	0.99
300/3	9.90	0.042	8.58	1
300/12	6.2221	10^{-6}	4.1877	0.99

(b)

Fig. 6. Variation of average balanced SINR versus the number of transmit antennas per BS, T. For robust designs, channel uncertainties are bounded in a ball of radius ρ. The value of the normalized power for both BSs is taken as 20 dB. (a) Variation of T versus average SINR (dB). The total number of users is $K = 20$ (10 users per cell). (b) Performance comparison of zero-forcing beamforming and the proposed SOCP method with large-scale antenna setup.

we can write $\mathbf{m}_k = \mathbf{G}_k \tilde{\mathbf{m}}_k$. The problem (23) is thus further equivalent to

$$\underset{(t,\tilde{\mathbf{m}}_k)\in\tilde{\mathcal{F}}}{\text{maximize}} \quad t \tag{24}$$

where $\tilde{\mathbf{h}}_{b_k,k} \triangleq \mathbf{h}_{b_k,k}\mathbf{G}_k$ and $\tilde{\mathbf{M}}_b \triangleq [\tilde{\mathbf{m}}_{U_b(1)}, \ldots, \tilde{\mathbf{m}}_{U_b(|U_b|)}]$ and $\tilde{\mathcal{F}} \triangleq \{\tilde{\mathbf{m}}_k \,|\, \|(\tilde{\mathbf{M}}_b)\|_2 \le \sqrt{P_b}\ \forall\, b, \sigma(t/\alpha_{b_k,k}) \le \Re(\tilde{\mathbf{h}}_{b_k,k}\tilde{\mathbf{m}}_k),\ \forall\, k\}$. The advantage of the ZF-BF scheme is that we can avoid the bisection procedure that must be carried out to solve (4) for optimal linear beamforming. However, the performance of ZF-BF is still far away from that of the non-robust design as shown in Fig. 6(a). For example, when $T = 120$, a gap of about 4 dB is observed between ZF-BF and general linear beamforming. A similar observation has also been made in [6]. At the same time, as expected, both non-robust and ZF-BF solutions fail to deliver appreciable PEs when ρ is not zero. It is seen that when $\rho = \rho' = 0.005$, the ZF-BF solution surpasses the SOCP approach beyond $T = 90$ antennas. Hence, if we assume that the channels are perfectly estimated, the ZF-BF solution provides a plausible option for massive MIMO case. Of course, in the presence of errors, the benefit of achieving high SINR is no longer available as it cannot be achieved $(PE \approx 0)$ for a non zero ρ. It is also seen in Fig. 6(a) that the robust SOCP design scales noticeably with the number of transmit antennas for the given number of users. A similar conclusion also applies to the robust designs when ρ is taken as 0.001, 0.005 and 0.01. We note that these values of ρ are comparable to the average channel gains of $\mathbf{h}_{b,k}$ taking into account the effect of shadowing and path loss. Therefore, owing to the conservative nature of robust designs it is not possible to achieve nontrivial SINRs for higher values of ρ. In the last table shown in Fig. 6(b), we study the effect of the variation of the ratio T/K on the balanced SINRs and the PEs achieved by the SOCP design and the ZF-BF solution. The values taken for the ratio are similar to the typical ones used in the literature for large-scale antenna systems, see for example, [37]. Our investigations confirm the expected result that although the ZF-BF solution can produce comparatively large balanced SINR, the SOCP design more or less guarantees that its SINR is achieved with a high probability.

V. Conclusion

We have studied the design of beamformers that balance the SINR of users in a multicell downlink system in the presence of channel uncertainties. Norm bounded channel uncertainty model is used. As a first approach to solving the problem in this scenario, we present an $\mathcal{S}$-lemma based approximate solution in which the beamformers are obtained by solving an SDP in conjunction with bisection search. Later, by exploiting various properties of the functions involved in the problem, we present a solution in which robust beamformers are solutions to an SOCP-based formulation. We show that in addition to being capable of handling different uncertainty sets, the SOCP-based approximation exhibits a much reduced complexity solution. We have tested the performance of proposed approaches for the recently conceived massive antenna systems, and have determined that the SOCP approach outperforms the SDP-based solution from computational cost perspective. Finally, we have also shown that the reduced complexity SOCP-based approach yields a balanced SINR and the probability of achieving it which is comparable with the SDP approach.

Appendix A
Proof of Proposition 1

For a given b, k, let us assume that $z_{b,k}$ and $\mathbf{M}_b$ are infeasible in (10b), i.e., there exist $\boldsymbol{\theta}_{b,k}$, $\boldsymbol{\phi}_{b,k}$ and $\||\boldsymbol{\theta}_{b,k}| + |\boldsymbol{\phi}_{b,k}|\| \le \rho'_{b,k}$ such that

$$z_{b,k} - \left\| \mathbf{M}_b^{\mathrm{H}} \left[\hat{\mathbf{h}}_{b,k} + \sum_{i=1}^{l_{bq,k}} \boldsymbol{\delta}_{b,k}^i \left([\boldsymbol{\theta}_{b,k}]_i - [\boldsymbol{\phi}_{b,k}]_i \right) \right]^{\mathrm{H}} \right\|_2 < 0. \tag{25}$$

Let $[\mathbf{v}_{b,k}]_i = [\boldsymbol{\theta}_{b,k}]_i - [\boldsymbol{\phi}_{b,k}]_i$ for all i. Thus it is easy to see that $|[\mathbf{v}_{b,k}]_i| \le |[\boldsymbol{\theta}_{b,k}]_i| + |[\boldsymbol{\phi}_{b,k}]_i|$. Therefore, we obtain $\|\mathbf{v}_{b,k}\| \le \||\boldsymbol{\theta}_{b,k}| + |\boldsymbol{\phi}_{b,k}|\| \le \rho'_{b,k}$, and hence (12) is also infeasible. Next, we assume conversely that for a given b, k, $z_{b,k}$ and $\mathbf{M}_b$ are infeasible in (12), i.e.,

$$z_{b,k} - \left\| \mathbf{M}_b^{\mathrm{H}} \left[\hat{\mathbf{h}}_{b,k} + \sum_{i=1}^{l_{b,k}} \boldsymbol{\delta}_{b,k}^i [\mathbf{v}_{b,k}]_i \right]^{\mathrm{H}} \right\|_2 < 0 \tag{26}$$

for certain $\mathbf{v}_{b,k}$ such that $\|\mathbf{v}_{b,k}\| \le \rho'_{b,k}$. Let $[\boldsymbol{\theta}_{b,k}]_i = (1 - \vartheta_{b,k})[\mathbf{v}_{b,k}]_i$ and $[\boldsymbol{\phi}_{b,k}]_i = -\vartheta_{b,k}[\mathbf{v}_{b,k}]_i$, where $\vartheta_{b,k} \in [0, 1]$.

$$f_1(\mathbf{M}_b, z_{b,k}, \hat{\mathbf{h}}_{b,k}) \geq -\min_{\||\boldsymbol{\theta}_{b,k}|+|\boldsymbol{\phi}_{b,k}|\|\leq\rho'_{b,k}} \sum_{i=1}^{l_{b,k}} \left\{ f_2\left(\mathbf{M}_b, \boldsymbol{\delta}^i_{b,k}\right) |[\boldsymbol{\theta}_{b,k}]_i| + f_2\left(\mathbf{M}_b, -\boldsymbol{\delta}^i_{b,k}\right) |[\boldsymbol{\phi}_{b,k}]_i| \right\} \tag{29a}$$

$$= \max_{\||\boldsymbol{\theta}_{b,k}|+|\boldsymbol{\phi}_{b,k}|\|\leq\rho'_{b,k}} \sum_{i=1}^{l_{b,k}} \left\{ -f_2\left(\mathbf{M}_b, \boldsymbol{\delta}^i_{b,k}\right) |[\boldsymbol{\theta}_{b,k}]_i| - f_2(\mathbf{M}_b, -\boldsymbol{\delta}^i_{b,k}) |[\boldsymbol{\phi}_{b,k}]_i| \right\} \tag{29b}$$

$$= \max_{\|\mathbf{v}_{3b,3k}\|\leq\rho'_{b,k}} \sum_{i=1}^{l_{b,k}} \left\{ \max\left(-f_2\left(\mathbf{M}_b, \boldsymbol{\delta}^i_{b,k}\right), -f_2\left(\mathbf{M}_b, -\boldsymbol{\delta}^i_{b,k}\right), 0\right) \left|[\mathbf{v}_{3b,3k}]_i\right| \right\} \tag{29c}$$

With this substitution, it is seen that $[\mathbf{v}_{b,k}]_i = [\boldsymbol{\theta}_{b,k}]_i - [\boldsymbol{\phi}_{b,k}]_i$. Similarly, these substitutions imply $|[\boldsymbol{\theta}_{b,k}]_i| + |[\boldsymbol{\phi}_{b,k}]_i| = |(1-\vartheta_{b,k})||[\mathbf{v}_{b,k}]_i| + |-\vartheta_{b,k}||[\mathbf{v}_{b,k}]_i| = |[\mathbf{v}_{b,k}]_i|$, and, thus, $\||\boldsymbol{\theta}_{b,k}| + |\boldsymbol{\phi}_{b,k}|\| = \|\mathbf{v}_{b,k}\| \leq \rho'_{b,k}$. Therefore, the variables $z_{b,k}$ and $\mathbf{M}_b$ are infeasible in (10b) as well. Hence, we conclude that the feasibility of the constraints in (10b) implies the feasibility of (12), and vice versa. Therefore, the proposition follows.

APPENDIX B
PROOF OF THEOREM 1

We will follow the notation in (16) and obtain a tractable version of this constraint by adapting the arguments developed in [16]. Let O_1 and O_2 be the optimal solutions of

$$\max \quad \mathbf{a}_1^{\mathrm{T}}\mathbf{v}_1 + \mathbf{a}_2^{\mathrm{T}}\mathbf{v}_2 \tag{27a}$$

$$\text{subject to} \quad \|\mathbf{v}_1 + \mathbf{v}_2\| \leq \vartheta \tag{27b}$$

$$\mathbf{v}_1 \geq 0, \mathbf{v}_2 \geq 0 \tag{27c}$$

and

$$\max \quad \sum_{i\in\mathcal{I}} \max\left\{[\mathbf{a}_1]_i, [\mathbf{a}_2]_i, 0\right\} [\mathbf{v}_3]_i \tag{28a}$$

$$\text{subject to} \quad \|\mathbf{v}_3\| \leq \vartheta \tag{28b}$$

respectively. It is shown in [16] that $O_1 = O_2$. Now using (16), we can obtain the set of relations (29) shown at the top of the page. In (29c) we have employed the result in optimization problem formulation (28), and have slightly changed the representation in (16) by not specifying any particular norm in the constraint set. Now recalling the definition of the dual norm $\vartheta\|\boldsymbol{\gamma}\|^\star \triangleq \max_{\|\mathbf{s}\|\leq\vartheta} \mathbf{s}^{\mathrm{T}}\boldsymbol{\gamma}$ of vector $\boldsymbol{\gamma}$, we obtain the result stated in Theorem 1 for the constraint set of interest.

ACKNOWLEDGMENT

The authors would like to thank Prof. Johan Löfberg of the Linköping University, Sweden for his helpful advice regarding the numerical results presented in the paper.

REFERENCES

[1] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.

[2] H. A. Suraweera, P. J. Smith, and M. Shafi, "Capacity limits and performance analysis of cognitive radio with imperfect channel knowledge," *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 1811–1822, May 2010.

[3] S. A. Vorobyov, A. B. Gershman, and Z.-Q. Luo, "Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 313–324, Feb. 2003.

[4] N. Vucic and H. Boche, "Robust QoS-constrained optimization of downlink multiuser MISO systems," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 714–725, Feb. 2009.

[5] G. Zheng, K. K. Wong, and B. Ottersten, "Robust cognitive beamforming with bounded channel uncertainties," *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4871–4881, Dec. 2009.

[6] A. Tajer, N. Prasad, and X. Wang, "Robust linear precoder design for multi-cell downlink transmission," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 235–251, Jan. 2011.

[7] D. Bertsimas, D. B. Brown, and C. Caramanis, "Theory and applications of robust optimization," *SIAM Rev.*, vol. 53, no. 3, pp. 464–501, 2011.

[8] A. B. Tal, L. E. Ghaoui, and A. Nemirovski, *Robust Optimization*. Princeton, NJ, USA: Princeton Univ. Press, Aug. 2009.

[9] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[10] L.-N. Tran, M. F. Hanif, A. Tolli, and M. Juntti, "Fast converging algorithm for weighted sum rate maximization in multicell MISO downlink," *IEEE Signal Process. Lett.*, vol. 19, no. 12, pp. 872–875, Dec. 2012.

[11] A. Tolli, M. Codreanu, and M. Juntti, "Linear multiuser MIMO transceiver design with quality of service and per-antenna power constraints," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3049–3055, Jul. 2008.

[12] A. Wiesel, Y. C. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 161–176, Jan. 2006.

[13] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U. K.: Cambridge Univ. Press, 2004.

[14] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.

[15] A. B. Tal and A. Nemirovski, "Robust convex optimization," *Math. Oper. Res.*, vol. 23, no. 4, pp. 769–805, Nov. 1998.

[16] D. Bertsimas and M. Sim, "Tractable approximations to robust conic optimization problems," *Math. Program.*, vol. 107, no. 1, pp. 5–36, Jun. 2006.

[17] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[18] T. Yoo and A. Goldsmith, "Capacity and power allocation for fading MIMO channels with channel estimation error," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2203–2214, May 2006.

[19] J. Zheng, E. Duni, and B. Rao, "Analysis of multiple-antenna systems with finite-rate feedback using high-resolution quantization theory," *IEEE Trans. Signal Process.*, vol. 55, no. 4, pp. 1461–1476, Apr. 2007.

[20] S. Boyd, L. E. Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in Systems and Control Theory*. Philadelphia, PA, USA: SIAM, 1994.

[21] E. Song, Q. Shi, M. Sanjabi, R. Sun, and Z.-Q. Luo, "Robust SINR-constrained MISO downlink beamforming: When is semidefinite programming relaxation tight?" in *Proc. IEEE ICASSP*, May 2011, pp. 3096–3099.

[22] Z.-Q. Luo, W. K. Ma, A. M. Cho, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.

[23] D. Bertsimas and D. B. Brown, "Constrained stochastic LQC: A tractable approach," *IEEE Trans. Autom. Control*, vol. 52, no. 10, pp. 1826–1841, Oct. 2007.
[24] M. Shenouda and T. Davidson, "Nonlinear and linear broadcasting with QoS requirements: Tractable approaches for bounded channel uncertainties," *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 1936–1947, May 2009.
[25] A. B. Tal and A. Nemirovski, "Robust solutions of linear programming problems contaminated with uncertain data," *Math. Program.*, vol. 88, no. 3, pp. 411–424, Sep. 2000.
[26] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Rev.*, vol. 38, no. 1, pp. 49–95, Mar. 1996.
[27] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, "Applications of second-order cone programming," *Linear Algebra Appl., Special Issue Linear Algebra Control, Signals Image Process.*, vol. 284, pp. 193–228, Nov. 1998.
[28] Z.-Q. Luo, T. N. Davidson, G. B. Giannakis, and K. M. Wong, "Transceiver optimization for block-based multiple access through ISI channels," *IEEE Trans. Signal Process.*, vol. 52, no. 4, pp. 1037–1052, Apr. 2004.
[29] A. Tremba *et al.*, "RACT: Randomized algorithms control toolbox for MATLAB," in *Proc. 17th World Congr. IFAC*, Seoul, Korea, 2008, pp. 390–395.
[30] J. Löfberg, "YALMIP: A toolbox for modeling and optimization in MATLAB," in *Proc. CACSD Conf.*, Taipei, Taiwan, 2004, pp. 284–289.
[31] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optim. Methods Softw.*, vol. 11, no. 1–4, pp. 625–653, Jan. 1999.
[32] K. C. Toha, M. J. Todd, and R. H. Tutuncuc, "SDPT3—A Matlab software package for semidefinite programming, version 1.3," *Optim. Methods Softw.*, vol. 11, no. 1–4, pp. 545–581, Jan. 1999.
[33] Gurobi Optimization, Gurobi Optimizer Version 5.1., Houston, TX, USA, Mar./Apr. 2013. [Online]. Available: http://www.gurobi.com/
[34] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta (2013, Apr.). Energy and spectral efficiency of very large multiuser MIMO systems. *IEEE Trans. Commun.* [Online]. *61(4)*, pp. 1436–1449. Available: http://arxiv.org/abs/1112.3810
[35] L.-N. Tran, M. Bengtsson, and B. Ottersten, "Iterative precoder design and user scheduling for block-diagonalized systems," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3726–3739, Jul. 2012.
[36] L.-N. Tran, M. Juntti, M. Bengtsson, and B. Ottersten, "Beamformer designs for MISO broadcast channels with zero-forcing dirty paper coding," *IEEE Trans. Wireless Commun.*, vol. 12, no. 3, pp. 1173–1185, Mar. 2013.
[37] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Feb. 2013.

Muhammad Fainan Hanif received the Ph.D. degree in electrical and electronic engineering from the University of Canterbury, Christchurch, New Zealand, in 2011.

Later, he was a Postdoctoral Researcher with the Centre for Wireless Communications, Department of Communications Engineering, University of Oulu, Oulu, Finland. His research interests include convex optimization applications in signal processing and modern communication systems.

Le-Nam Tran (M'09) received the B.S. degree in electrical engineering from Ho Chi Minh National University of Technology, Ho Chi Minh, Vietnam, in 2003, and the M.S and Ph.D. degrees in radio engineering from Kyung Hee University, Seoul, Korea, in 2006 and 2009, respectively.

In 2009, he joined the Department of Electrical Engineering, Kyung Hee University, as a lecturer. From September 2010 to July 2011, he was a Postdoctoral Fellow with the Signal Processing Laboratory, ACCESS Linnaeus Centre, KTH Royal Institute of Technology, Stockholm, Sweden. Since August 2011, he has been with the Centre for Wireless Communications and the Department of Communications Engineering, University of Oulu, Oulu, Finland. His current research interests include multiuser MIMO systems, energy efficient communications, and full duplex transmission. He was the recipient of the Best Paper Award from the Institute of Information Technology Advancement (IITA) in August 2005.

Antti Tölli (M'08–SM'14) received his Dr.Sc. (Tech.) degree in electrical engineering from the University of Oulu, Oulu, Finland, in June 2008.

Before joining the Department of Communication Engineering (DCE) and the Centre for Wireless Communications, University of Oulu, he worked five years for Nokia Networks, IP Mobility Networks Division, Espoo, Finland, as a Research Engineer and Project Manager both in Finland and Spain. Currently, he works as a Senior Research Fellow and Adjunct Professor with the DCE, University of Oulu. He served as a General Cochair of IEEE WDN in 2010 and 2011, as well as Finance Chair in IEEE Communication Theory Workshop 2011. His main research interests are in radio resource management and transceiver design for broadband wireless communications with special emphasis on interference management in multiuser MIMO cellular systems. He has authored about 80 papers and several patents in the areas of signal processing and wireless communications.

Markku Juntti (S'93–M'98–SM'04) received the M.Sc. (Tech.) and Dr.Sc. (Tech.) degrees in electrical engineering from University of Oulu, Oulu, Finland in 1993 and 1997, respectively.

He was with the University of Oulu in 1992–1998. In academic year 1994–95, he was a Visiting Scholar with Rice University, Houston, TX, USA. In 1999–2000, he was a Senior Specialist with Nokia Networks, Espoo, Finland. Since 2000, he has been a Professor of communications engineering with the University of Oulu, Department of Communication Engineering and Centre for Wireless Communications. His research interests include signal processing for wireless networks, as well as communication and information theory. He has authored or coauthored some 200 papers published in international journals and conference records, as well as in the book *WCDMA for UMTS* published by Wiley. He is also an Adjunct Professor with the Department of Electrical and Computer Engineering, Rice University, Houston, TX, USA.

Dr. Juntti is an Editor of IEEE TRANSACTIONS ON COMMUNICATIONS and was an Associate Editor for IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. He was a Secretary of the IEEE Communication Society Finland Chapter in 1996–97 and the Chairman for years 2000–2001. He has been Secretary of the Technical Program Committee of the 2001 IEEE International Conference on Communications (ICC'01), and the Cochair of the Technical Program Committee of 2004 Nordic Radio Symposium and 2006 IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2006). He is the General Chair of 2011 IEEE Communication Theory Workshop (CTW 2011).