Power Allocation in MISO Interference Channels with Stochastic CSIT

Weiqiang Xu, *Senior Member, IEEE,* Ali Tajer, *Member, IEEE,* Xiaodong Wang, *Fellow, IEEE,* Saleh Alshomrani

***Abstract*—This paper considers multiuser interference channels in which the transmitters have imperfect channel state information (CSI) where CSI perturbations are modeled stochastically. Transmitters are assumed to be equipped with multiple antennas serving single-antenna receivers. Transmitters use predesigned discrete codebooks for beamforming directions and dynamically (based on the available CSI) select the best set of beamformers from the given codebook. The objective is to perform optimal power allocation to different users while certain quality of service (QoS) guarantees are ensured for the users. Imposed by stochastic CSI uncertainties, guarantees provided for the QoS measures have a stochastic nature too. The primary focus is placed on the interference channels for which two power allocation problems are considered. The first problem minimizes power consumption subject to serving users at certain data rates and the second problem considers max-min rate allocation subject to given power budgets for the transmitters. The core step in formalizing these problems in mathematically tractable forms relies on using *Bernstein* approximation, which approximates and convexifies the non-convex stochastic guarantees by conservative convex and deterministic counterparts. For solving this resulting convex and deterministic optimization problem, a specialized version of the long-step logarithmic barrier cutting plane (LL-BCP) algorithm is used. Effectiveness of the proposed solutions and comparisons with other existing methods are assessed via extensive simulation results.**

***Index Terms*—Multiple-input single-output (MISO), interference channel, broadcast channel, power control, probability-constrained optimization, Bernstein approximation.**

I. Introduction

INTERFERENCE channels are among the most important building blocks of the modern multiuser communication networks. These channels emphasize the disruptive effects of interference that independent transmitter-receiver links impose on one another and is the fundamental model for studying interference management, e.g, inter-cell interference management in multi-cell multiuser networks [1]. Motivated by the significance of these multiuser channels their fundamental limits have been the subject of extensive studies. Exploiting the spatial selectivity of the channels and directional signal steering enabled by multi-antenna transmitters is known to be an effective approach to focus the energy of a transmitter towards its intended receiver(s) and harness its disruptive effects on the non-intended receivers.

Optimal directional transmission (beamforming) strongly hinges on the availability of *perfect* channel state information (CSI) at the transmitters sites. This requirement is, however, hard to meet in practice where there are multiple reasons, including channel estimation errors, delayed estimation finite-rate feedback channels, and quantization errors, causing perturbations to the CSI. It is shown that treating the imperfect CSI as perfect CSI and applying the optimal solutions for perfect CSI settings lead to rapid performance degradation as CSI perturbation level increases [2]. Driven by this observation and the premise that acquiring CSI is via only *imperfectly* there is a growing interest and body of literature studying the effects of CSI imperfections on designing the transmission strategies, which we review next.

Depending on the nature of CSI uncertainties there exist, broadly, two approaches coping with such uncertainties. In one approach the CSI uncertainties are assumed to be primarily shaped by the quantization errors introduced by finite-rate feedback from the receivers to the transmitters. In this approach channel uncertainties are modeled as bounded values confined within known hyper-spherical regions and the objective is to provide *worst-case* guarantees for the performance of the network. More specifically, these approaches search in the channel states uncertainty regions and identify the channel states that yield the weakest performance for the network and declare it as the worst-case performance, which in turn serves as a guaranteed performance level for the network. Such worst-guarantees are studied for power optimization in broadcast channel in [3], for beamforming in broadcast channels in [4]–[6], for beamforming in interference channels in [7], and for beamforming in multi-cell downlink transmission in [8], [9]. The worst-case-based optimization approaches provide robustness against CSI imperfections. Nevertheless, the actual worst case may occur with a very slim chance. Hence, the worst-case approach may be overly pessimistic and therefore, may lead to unnecessary performance degradation. The resulting optimization problem sometimes does not even have a feasible

Manuscript submitted August 29, 2013; revised November 20, 2013; accepted December 11, 2013. The associate editor coordinating the review of this paper and approving it for publication was D. Niyato.

This work was presented in part at the 2013 IEEE Globecom, December 2013, Atlanta, GA.

This work was supported in part by the National Science Foundation of China (NSFC) under grants 61374020, 61302076, 61272311, and 61329101, by the Key Project of Chinese Ministry of Education under grant 212066, by the Zhejiang Provincial Natural Science Foundation of China under grants Y12F020196, LQ13F010008, and by the Deanship of Scientific Research (DSR) , King Abdulaziz University (KAU), under grant 15-15-1432 HiCi.

W. Xu is with the School of Information Science & Technology, Zhejiang Sci-Tech University, Hangzhou, 310018, P. R. China (e-mail: wqxu@zstu.edu.cn).

A. Tajer is with the Electrical & Computer Engineering Department, Wayne State University, Detroit, MI, 48202, USA (e-mail: tajer@wayne.edu).

X. Wang is with the Electrical Engineering Department, Columbia University, New York, NY, 10027, USA. He is also affiliated with King Abdulaziz University, Jeddah, Saudi Arabia (e-mail: wangx@ee.columbia.edu).

S. Alshomrani is with the Faculty of Computing & Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia (e-mail: sshomrani@kau.edu.sa).

Digital Object Identifier 10.1109/TWC.2014.012814.131600

solution and even if the problem is feasible, the resource utilization is inefficient as most system resources must be dedicated to provide guarantees for the worst-case scenarios.

In order to provide less conservative solutions in favor of improved performance, in the second approach the CSI uncertainties are modeled stochastically that can potentially lie within bounded or unbounded regions. In this approach, instead of deterministic worst-case guarantees, probabilistic guarantees are provided for network performance. In these approaches the probabilistic guarantees are often approximated and convexified by deterministic ones and often exhibit improved average performance compared with the worst-case solutions, which controlling the likelihood of the worst-case performance. The relevant literature on stochastic imperfect CSI is mostly focused on single-user channels and multiuser broadcast and multiple access channels. Effects of channel uncertainties on the capacity of single-user multi-antenna systems is analyzed in [10] where lower and upper bounds on the mutual function are derived. Designing robust transceivers in single-users multi-antenna channels is studied in [11]. Receiver design in multiple access channels and its pertinent space-time code design are considered in [12] and [13], respectively, and transmission precoding in broadcast channels is studied in [10]–[12], [14]–[18]. Specifically, non-linear precoding (Tomlinson-Harashima) is treated in [16] where the stochastic quality of service (QoS) constraints are converted to deterministic second-order cone constraints via Schur-Complement properties. Power control under linear precoding (beamforming) in broadcast channels is studied in [17] where Vysochanskii-Petunin inequality is used to transform the problem into a deterministic convex optimization problem, and finally, beamforming under stochastic CSI uncertainties is studied in [14], [15] where the stochastic QoS guarantees are converted to conservative deterministic ones via a Bernstein-type *inequality* [19].

This paper focuses on interference channels, in which multiple independent multi-antenna transmitters wish to communicate with their respective single-antenna receivers. We assume that the multi-antenna transmitters employ some fixed beamformers to perform directional transmission and treat two closely related optimization problems. In the first problem the objective is to minimize the total transmission power subject to outage constraints, and in the second problem the goal is to maximize the achievable rate margins under the power constraint. We make use Bernstein *approximation* technique [20], [21], which is a recent advance in the field of chance-constrained programming that provides conservative deterministic, convex, and computationally tractable approximations for a computationally intractable probabilistic constraints.

It is noteworthy that the Bernstein approximation approach is different from the Bernstein-type inequality approach of [14], [15], [19] used for beamforming design in broadcast channels. Specifically, Bernstein-type inequality of [19] is a concentration-based inequality that upper bounds the likelihood that a quadratic function of a set of *Gaussian* random variables deviates from the concentration point. Bernstein approximation method of [20], on the other hand, builds approximates for the chance constraints under the assumption that the constraints are affine in perturbations and the entries of the perturbations are independent. Advantages of Bernstein approximation over Bernstein inequality are demonstrated via simulation results in Section IV, respectively. Finally, we remark that Bernstein approximation method was first deployed in wireless communication for treating chance-constrained resource allocation in adaptive adaptive orthogonal frequency division multiple access (OFDMA) systems [21] and we use the technique for applying on power allocation in multi-input single-output (MISO) interference channels.

The remainder of this paper is organized as follows. The chance-constrained power optimization problems in interference channels is formalized in Section II. In Section III, we offer the solutions to the stochastic optimization problems based on the Berstein approximation technique and the long-step logarithmic barrier cutting plane (LLBCP) algorithm. Section IV presents the simulation results and finally, the concluding remarks are provided in Section V.

II. Preliminaries

A. System Model

We consider a MISO interference channel with K transmitters and K receivers. Each transmitter employs M transmit antennas and each receiver is equipped with a single receive antenna. We assume that all receivers treat co-channel interference as noise, i.e., they make no attempt to decode the interference. Assuming a narrowband channel model, the received signal at receiver $i \in \{1, \dots, K\}$ is given by

$$y_i \;=\; \boldsymbol{h}_{ii}\boldsymbol{x}_i \;+\; \sum_{j\neq i}\boldsymbol{h}_{ij}\boldsymbol{x}_j \;+\; n_i \;, \tag{1}$$

where $\boldsymbol{x}_i \in \mathbb{C}^{M\times 1}$ is the transmitted signal vector by the i-th transmitter, $\boldsymbol{h}_{ij} \in \mathbb{C}^{1\times M}$ is the channel vector linking the j-th transmitter and the i-th receiver, and $n_i \sim \mathcal{N}_{\mathbb{C}}(0, \eta^2)$ is the additive complex Gaussian channel noise. Transmitters employ beamforming to perform directional transmission and we denote the beamforming vector of the i-th transmitter by $\boldsymbol{w}_i \in \mathbb{C}^{M\times 1}$. Therefore, for the transmission vector of the i-the transmitter we have $\boldsymbol{x}_i = \boldsymbol{w}_i s_i$, where $s_i \in \mathbb{C}$ denotes the complex data symbol intended for the i-th receiver. We assume that the receivers can acquire *perfect* CSI while the transmitters have access to only *imperfect* CSI. More specifically, we assume that channel $\boldsymbol{h}_{ij}$ can be decomposed into

$$\boldsymbol{h}_{ij} \;\triangleq\; \hat{\boldsymbol{h}}_{ij} \;+\; \boldsymbol{\delta}_{ij} \;, \tag{2}$$

where $\hat{\boldsymbol{h}}_{ij}$ denotes the imperfect estimate of the channel vector $\boldsymbol{h}_{ij}$, known to transmitter i, and $\boldsymbol{\delta}_{ij}$ denotes the unknown part of $\boldsymbol{h}_{ij}$ with covariance matrix $\boldsymbol{C}_{ij} \triangleq \mathbb{E}[\boldsymbol{\delta}_{ij}^H \boldsymbol{\delta}_{ij}]$. Driven by the practical needs of multiuser systems we assume that beamforming directions are selected from pre-designed beamforming codebooks [3], [17], e.g., discrete Fourier transform (DFT) based codebooks in Long Term Evolution (LTE) systems. Hence, we assume that the beamforming vector $\boldsymbol{w}_i$ can be decomposed as

$$\boldsymbol{w}_i \;\triangleq\; \sqrt{p_i}\,\boldsymbol{g}_i \;, \tag{3}$$

where p_i and $\boldsymbol{g}_i$ denote the power and direction of $\boldsymbol{w}_i$, i.e., $p_i = \|\boldsymbol{w}_i\|^2$ and $\|\boldsymbol{g}_i\|^2 = 1$. The beamforming directions $\{\boldsymbol{g}_i\}$

are based on known the channel estimates $\{\boldsymbol{h}_i\}$. On the other hand, the design of the transmission powers $\{p_i\}$ is much more sophisticated and it depends not only on the channel estimates, but also on the channel perturbation statistic. In this paper, we assume that the beam directions $\{\boldsymbol{g}_i\}$ are selected based on $\{\boldsymbol{h}_i\}$ and are fixed, and we focus on the power allocation problem, i.e., finding the optimal choices of $\{p_i\}$. Finally, we assume that the receivers employ single-use decoders where the interferers at each receiver are treated as Gaussian noise. Hence, the signal to noise plus interference ratio (SINR) of the i-th user is

$$\mathsf{SINR}_i = \frac{p_i|\boldsymbol{h}_{ii}\boldsymbol{g}_i|^2}{\eta^2 + \sum_{j\neq i} p_j|\boldsymbol{h}_{ij}\boldsymbol{g}_j|^2} \ , \tag{4}$$

and consequently, the rate sustained by the channel linking the i-the transmitter to the i-th receiver is

$$R_i \ \triangleq \ \log(1+\mathsf{SINR}_i) \ . \tag{5}$$

B. Problem Formulations

We next formulate the chance-constrained power allocation problems. Our goal is to optimize power allocation, while, in parallel, stochastic constraints on outage events are satisfied. We consider two chance-constrained optimization problems as follows.

1) Rate-constrained Power Optimization: The first problem seeks to minimize the average transmit power subject to the rate constraints. Specifically, given the desired rate r_i and outage probability ε_i for the i-th transmitter and receiver pair, we aim to minimize the average transmit power while meeting the rate outage constraints of all users, i.e.,

$$\mathcal{P}_1(\boldsymbol{r},\boldsymbol{\varepsilon}) \ \triangleq \ \left\{ \begin{array}{ll} \min_{\{p_i\}} & \sum_{i=1}^K p_i \\ \text{s.t.} & \mathbb{P}(R_i \ \leq \ r_i) \ \leq \ \varepsilon_i \ , \quad \forall i \\ & 0 \ \leq \ p_i \ , \quad \forall i \end{array}\right. , \tag{6}$$

where $\boldsymbol{r} \triangleq [r_1,\dots,r_K]$, $\boldsymbol{\varepsilon} \triangleq [\varepsilon_1,\dots,\varepsilon_K]$, and $\mathbb{P}(A)$ denotes the probability of the event A. The design parameter ε_i ensures that receiver i is served with rate R_i no less than r_i at least $(1-\varepsilon_i)\times 100\%$ of the time.

2) Max-Min Rate Optimization: The second problem seeks to maximize the minimum rate among all users, subject to rate outage constraints, and individual transmit power constraints. This problem can be cast as

$$\mathcal{P}_2(\bar{\boldsymbol{p}},\boldsymbol{\varepsilon}) \ \triangleq \ \left\{ \begin{array}{ll} \max_{\{p_i\}} & \min_i \ r_i \\ \text{s.t.} & \mathbb{P}(R_i \ \leq \ r_i) \ \leq \ \varepsilon_i \ , \quad \forall i \\ & 0 \ \leq \ p_i \ \leq \ \bar{p}_i \ , \quad \forall i \end{array}\right. , \tag{7}$$

where $\bar{\boldsymbol{p}} \triangleq [\bar{p}_1,\dots,\bar{p}_K]$ captures the individual power constraints. In problems $\mathcal{P}_1(\boldsymbol{r},\boldsymbol{\varepsilon})$ and $\mathcal{P}_2(\bar{\boldsymbol{p}},\boldsymbol{\varepsilon})$ given in (6) and (7), respectively, the probabilistic constraints make the optimization intractable. To circumvent the above hurdles, we make use of the Bernstein approximation technique [20], [21] to convert the probabilistic constraints to convex deterministic constraints. Next we briefly introduce the Bernstein approximation.

3) Bernstein Approximation: Consider the following optimization problem:

$$\mathcal{P} \ \triangleq \ \left\{ \begin{array}{ll} \min_{\boldsymbol{x}\in\mathcal{X}} & f(\boldsymbol{x}) \\ \text{s.t.} & \mathbb{P}(F(\boldsymbol{x},\boldsymbol{\zeta}) \ \geq \ 0) \ \leq \ \varepsilon \end{array}\right. , \tag{8}$$

where $\boldsymbol{\zeta}$ is a random vector supported on $\Xi \subseteq \mathbb{R}^d$, $\boldsymbol{x} \in \mathbb{R}^n$ and $F: \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}$ is a function of $\boldsymbol{z} \in \mathbb{R}^n$ and $\boldsymbol{\zeta} \in \mathbb{R}^d$. Also define $\psi: \mathbb{R} \to \mathbb{R}$ as a non-negative, non-decreasing, and convex function satisfying $\psi(z) \geq \psi(0) = 1$ for all $x \in \mathbb{R}_+$. Then, as discussed in details in [20] and reviewed in [21], the chance-constrained optimization problem $\mathcal{P}$ can be approximated *conservatively* by the following deterministic optimization problem:

$$\hat{\mathcal{P}} \ \triangleq \ \left\{ \begin{array}{ll} \min_{\boldsymbol{x}\in\mathcal{X}} & f(\boldsymbol{x}) \\ \text{s.t.} & \inf_t\{\hat{\Psi}(\boldsymbol{x},t) - t\log\varepsilon\} \ \leq 0 \\ & 0 \leq t \end{array}\right. , \tag{9}$$

where

$$\hat{\Psi}(\boldsymbol{x},t) \ \triangleq \ t \ \log\mathbb{E}\{\psi[t^{-1}F(\boldsymbol{x},\boldsymbol{\zeta})]\} \ . \tag{10}$$

Moreover, when $F(\cdot,\boldsymbol{\zeta})$ is convex in its first argument for a given $\boldsymbol{\zeta}$, by following the same footsteps as in [20] it can be readily shown that

$$\inf_t\{\hat{\Psi}(\boldsymbol{x},t) - t\log\varepsilon\} \tag{11}$$

is also convex. As an important special of the optimization problem $\mathcal{P}$, when the components of the random vector $\boldsymbol{\zeta} \triangleq [\zeta_1,\dots,\zeta_d]$ are independent, $F(\boldsymbol{z},\boldsymbol{\zeta})$ is affine in $\boldsymbol{\zeta}$, i.e.,

$$F(\boldsymbol{x},\boldsymbol{\zeta}) = f_0(\boldsymbol{x}) + \sum_{j=1}^d \zeta_j f_j(\boldsymbol{x}) \ , \tag{12}$$

where the functions $f_j(\boldsymbol{x}), j = 0,1,...,d$, are well defined and convex on $\mathbb{R}^n$, and we set $\psi(z) \triangleq \exp(z)$, then the function $\hat{\Psi}(\boldsymbol{x},\boldsymbol{\zeta})$ defined in (10) is convex and we have

$$\hat{\Psi}(\boldsymbol{x},t) \ = \ f_0(\boldsymbol{x}) + \sum_{j=1}^d t\log\mathbb{E}\{\exp[t^{-1}\zeta_j f_j(\boldsymbol{x})]\} \tag{13}$$

Hence, by using the Bernstein approximation method of [20], the optimization problem $\mathcal{P}$ with stochastic constrains can be converted to the optimization problem $\hat{\mathcal{P}}$ with convex and deterministic constraints.

III. Power Optimization in Interference Channels

In this section, we apply the Bernstein approximation to obtain the convex approximations to the probabilistic constraints in problems (6) and (7) and then solve the resulting convex problems using the long-step logarithmic barrier cutting plan (LLBCP) algorithm.

A. Rate-constrained Power Optimization

The major difficulty in the robust power optimization design is to convert the probabilistic constraint into a deterministic one. To that end we apply the Bernstein approximation to obtain the counterpart deterministic approximation in the following proposition.

Proposition 1: The stochastic rate-constrained power optimization problem $\mathcal{P}_1(\boldsymbol{r},\boldsymbol{\varepsilon})$ defined in (6) can be approximated conservatively by the deterministic optimization problem $\hat{\mathcal{P}}_1(\boldsymbol{r},\boldsymbol{\varepsilon})$ defined as

$$\hat{\mathcal{P}}_1(\boldsymbol{r},\boldsymbol{\varepsilon}) \triangleq \left\{ \begin{array}{lll} \min_{\{p_i\}} & \sum_{i=1}^{K} p_i & \\ \text{s.t.} & \inf_{t_i>\rho_i}\{\hat{\Psi}_i(\boldsymbol{p},t_i) - t_i \log \varepsilon_i\} \ \leq 0, \forall i & . \\ & 0 \ \leq \ p_i \ , \quad \forall i & \end{array} \right. \tag{14}$$

where we have defined

$$\hat{\Psi}(\boldsymbol{p},t_i) \triangleq \alpha_i \eta^2 \ + \left(\sum_{j\neq i} \frac{\alpha_i p_j |\hat{\boldsymbol{h}}_{ij}\boldsymbol{g}_j|^2}{1 - t_i^{-1}\alpha_i p_j \sigma_{ij}^2} - \frac{p_i |\hat{\boldsymbol{h}}_{ii}\boldsymbol{g}_i|^2}{1 - t_i^{-1} p_i \sigma_{ii}^2} \right) - t_i \left(\log\left(1 - t_i^{-1}\alpha_i p_j \sigma_{ij}^2\right) - \log\left(1 - t_i^{-1} p_i \sigma_{ii}^2\right) \right) \tag{15}$$

$$\alpha_i \triangleq 2^{r_i} - 1 \tag{16}$$

$$\rho_i \triangleq \max\left\{ p_i \sigma_{ii}^2 \ , \ \max_{j\neq i} \alpha_i p_j \sigma_{ij}^2 \right\} \tag{17}$$

$$\sigma_{ij}^2 \ \triangleq \ \|\boldsymbol{C}_{ij}^{1/2} \ \boldsymbol{g}_j\|^2 \ . \tag{18}$$

Proof: See Appendix A. ■

B. Max-Min Rate Optimization

We next consider the max-min rate optimization problem as formalized in (7). Since it is difficult to verify directly whether problem (7) is convex, we use the similar method in [9] to solve (7). Specifically, by introducing a slack variable $a > 0$, the epigraph form of the robust max-min rate optimization problem with individual power constraints (7) is given by

$$\mathcal{P}_2(\bar{\boldsymbol{p}},\boldsymbol{\varepsilon}) \ = \left\{ \begin{array}{lll} \max_{\{p_i\},a} & a & \\ \text{s.t.} & \mathbb{P}(R_i \ \leq \ a) \ \leq \ \varepsilon_i \ , \quad \forall i & , \\ & 0 \ \leq \ p_i \ \leq \ \bar{p}_i \ , \quad \forall i & \end{array} \right. \tag{19}$$

We demonstrate that solving $\mathcal{S}(\boldsymbol{p})$ can be facilitated via solving a power optimization problem defined as

$$\mathcal{S}(\bar{\boldsymbol{p}},\boldsymbol{\varepsilon},a) \triangleq \left\{ \begin{array}{lll} \min_{\{p_i\},a} & b & \\ \text{s.t.} & \mathbb{P}(R_i \ \leq \ a) \ \leq \ \varepsilon_i \ , \quad \forall i & , \\ & 0 \ \leq \ p_i \ \leq \ bp_i \ , \quad \forall i & \end{array} \right. \tag{20}$$

which can be solved using the similar method for solving the chance-constrained power minimization problem given in (6). The connection between $\mathcal{P}_2(\bar{\boldsymbol{p}},\boldsymbol{\varepsilon})$ and $\mathcal{S}(\bar{\boldsymbol{p}},\boldsymbol{\varepsilon},a)$ is given by the following Proposition

Proposition 2: Problem $\mathcal{S}(\bar{\boldsymbol{p}},\boldsymbol{\varepsilon},a)$ is strictly increasing and continuous in a at any feasible solution and it is related to $\mathcal{P}_2(\bar{\boldsymbol{p}},\boldsymbol{\varepsilon})$ via

$$\mathcal{S}(\bar{\boldsymbol{p}},\boldsymbol{\varepsilon},\mathcal{P}_2(\bar{\boldsymbol{p}},\boldsymbol{\varepsilon})) \ = \ 1 \ . \tag{21}$$

Proof: See Appendix B. ■

Since $\mathcal{S}(\bar{\boldsymbol{p}},\boldsymbol{\varepsilon},a)$ is strictly increasing and continuous in a, there exists a unique a^* satisfying $\mathcal{S}(\bar{\boldsymbol{p}},\boldsymbol{\varepsilon},a^*) = 1$. It follows from Proposition 2 that solving $\mathcal{P}_2(\bar{\boldsymbol{p}},\boldsymbol{\varepsilon})$ boils down to finding a^* that satisfies $\mathcal{S}(\bar{\boldsymbol{p}},\boldsymbol{\varepsilon},a^*) = 1$. Due to monotonicity and continuity of $\mathcal{S}(\bar{\boldsymbol{p}},\boldsymbol{\varepsilon},a)$, a^* can be obtained by a bi-section search over a. The iterative bi-section search for given a, involves solving $\mathcal{S}(\bar{\boldsymbol{p}},\boldsymbol{\varepsilon},a)$, which in turn, can be solved using the chance-constrained power optimization problem $\mathcal{P}_1(\boldsymbol{r},\boldsymbol{\varepsilon})$.

C. Long-step Logarithmic Barrier Cutting Plane (LLBCP) Algorithm

It is well-known that the convex optimization problem, as usual, can easily achieve the global optimum, such as in [22]. However, the constraint of the approximate deterministic convex optimization in (14) itself is a convex optimization problem. It is not in a standard form for which computationally efficient methods already exist and cannot be solved by straightforward application of the know convex optimization methods. In this Section we propose to use the long-step logarithmic barrier cutting plane (LLBCP) algorithm to solve it.

The detailed characteristics and development of the LLBCP algorithm can be found found in [23] [24]. Here we outline the basic ideas of this method. Suppose that we would like to find a solution $\boldsymbol{p}$ that is feasible for (14) and is a reasonable vicinity of the optimal solution, i.e., $\|\boldsymbol{p} - \boldsymbol{p}^*\| < \epsilon$[1], for some optimal solution $\boldsymbol{p}^*$ to (14) and error tolerance parameter $\epsilon > 0$ [2]. It is an iterative algorithm and at the beginning of each iteration, the feasible set, if exists, is contained in a bounded polytope. Then, a trial point is generated by constructing the analytic center inside the bounded polytope. If a slack variable has become large, and the associated variational quantity is small, that hyperplane is deemed *unimportant*, and then dropped. If the slack variables stay small, we test whether or not the trial point belongs to the feasible set. If this trial point is not feasible, a hyperplane through the trial point is introduced to cut off the violated constraint(s), so that the remaining polytope contains the feasible set. When the trial point is feasible but not optimal, by updating the lower bound on the optimal objective function value of problem (29) and reducing the barrier parameter(as defined in (24)), the new optimality constraint(s) is generated to update the polytope. We can then proceed to the next iteration with the new polytope until the termination condition is satisfied. Assuming that there exists a set of feasible solutions to (14), the iterative LLBCP algorithm terminates if one of the following three conditions is met:

1) **Termination 1:** The number of hyperplanes exceeds a certain level, so that the volume of the current polytope would be too small to contain a small enough ball.
2) **Termination 2:** The smallest slack is smaller than a certain number, so that the polytope would be too narrow to contain a small enough ball.
3) **Termination 3:** The duality gap is enough small, so that the algorithm may be terminated with optimality.

Given these termination rules, it is shown in [23] [24] that the LLBCP algorithm is guaranteed to terminate with a solution $\boldsymbol{p}$ that is feasible for (14) and satisfies $\|\boldsymbol{p} - \boldsymbol{p}^*\| < \epsilon$ for

[1] $\|\cdot\|$ denotes Euclidean norm operator.

[2] It is assumed that there exist the set of feasible solutions to (14) and a problem dependent constant ϵ such that the set of optimal solutions is guaranteed to be contained in the K dimensional hypercube of half-width $1/\epsilon$ and to contains a full dimensional ball of radius ϵ. Also it suffices to find a solution within an accuracy ϵ.

some optimal solution $\boldsymbol{p}^*$ after at most $\mathcal{O}(n(\log_2(1/\epsilon))^2)$ iterations, where n is the number of variables. Now, we begin to specialize the generic LLBCP algorithm to solve the power allocation problem at hand. The algorithm is initialized with the following relaxed power optimization problem.

$$\mathcal{Q}_0 \triangleq \begin{cases} \min_{\boldsymbol{p}} & \mathbf{1}^T \cdot \boldsymbol{p} \\ \text{s.t.} & \boldsymbol{p} \succeq -\frac{1}{\epsilon} \cdot \mathbf{1} \\ & \boldsymbol{p} \preceq \frac{1}{\epsilon} \cdot \mathbf{1} \\ & \boldsymbol{p} \succeq \mathbf{0} \\ & \mathbf{1}^T \cdot \boldsymbol{p} \geq \frac{1}{\epsilon}\sqrt{K} \end{cases} \tag{22}$$

where $\mathbf{1}$ and $\mathbf{0}$ denote the $K \times 1$ vector of all ones, and the $K \times 1$ vector of zeros, respectively, and $\preceq$, $\succeq$ represent component-wise inequalities. The first and second K hyperplanes[3] included in this initial relaxation will be referred to as the box hyperplanes, and serve to ensure that the polytope we have at any iteration is always bounded. The third K hyperplanes ensure $p_k \geq 0, \forall k$. It is easy to find the the first K hyperplanes are redundant, but they are retained as their inclusion facilitates solving problem $\mathcal{Q}_0$. The final K constraints give a initial bound $\frac{1}{\epsilon}\sqrt{K}$ on the optimal value of $\boldsymbol{p}$, which does not affect the feasible region of the initial relaxation. When the algorithm obtains better lower bounds on the optimal value, the right hand side of this lower bound constraint is updated. After the initialization given by $\mathcal{Q}_0$ in (22), the algorithm continues iteratively, where at the i-th iteration, the following optimization problem is solved.

$$\mathcal{Q}_i \triangleq \begin{cases} \min_{\boldsymbol{p}} & \mathbf{1}^T \cdot \boldsymbol{p} \\ \text{s.t.} & \boldsymbol{p} \succeq \mathbf{0} \\ & \boldsymbol{p} \preceq \frac{1}{\epsilon} \cdot \mathbf{1} \\ & \mathbf{1}^T \cdot \boldsymbol{p} \geq l_i \\ & \boldsymbol{A}^i \cdot \boldsymbol{p} \geq \boldsymbol{c}^i \end{cases} \tag{23}$$

where $\boldsymbol{A}^i \cdot \boldsymbol{p} \geq \boldsymbol{c}^i$ is the hyperplanes in the current relaxation and l_i is some lower bound on the optimal objective function value in the current relaxation.

Based on these the problems $\mathcal{Q}_i$, as the main ingredients of the LLBCP algorithm, in Fig. 1 we provide a detailed flow chart of the LLBCP algorithm with the step-by-step descriptions articulated in Algorithm 1. In the remainder of this subsection the key components of the LLBCP algorithm are elaborated in order to explain the dynamics of updating $\boldsymbol{A}^i \cdot \boldsymbol{p} \geq \boldsymbol{c}^i$ and l_i throughout algorithm iterations.

1) *Finding μ-center (Line 11):* In each iteration, we need to generate a trial point inside the polytope. To this end, we define $\mu > 0$ as a barrier parameter and define the corresponding logarithmic barrier function as

$$f(\boldsymbol{p}, \mu) \triangleq \frac{\mathbf{1}^T \cdot \boldsymbol{p}}{\mu} - \sum_n \ln(s_n) , \tag{24}$$

where we have defined

$$s_n \triangleq \boldsymbol{a}_n^T \boldsymbol{p} - c_n , \tag{25}$$

where $\boldsymbol{a}_n^T$ and c_n are the the n-th rows of $\boldsymbol{A}^i$ and $\boldsymbol{c}^i$, respectively. For a given value of μ, we define $\boldsymbol{p}(\mu)$ as the unique minimizer of this barrier function. We refer this unique point as the μ-center and has the property that an approximate the μ-center is sufficient to serve as a trial point. An approximate μ-center for the $(i+1)$-th iteration can be obtained from an approximate μ-center for the i-th iteration by applying $\mathcal{O}(1)$ Newton steps [25].

2) *Dropping unimportant constraints (lines 16-20):* Follows the same footsteps as in [21], [24].

3) *Feasibility test:* Follows the same footsteps as in [21], [24].

4) *Cutting off the violated constraints (lines 31-34):* Follows the same footsteps as in [21], [24]. If the trial point $\boldsymbol{p}_i$ is infeasible, then a hyperplane is generated at $\boldsymbol{p}$ as follows:

5) *Updating lower bound and reducing barrier parameter (lines 40-44):* If the point $\boldsymbol{p}_i$ is feasible but not optimal, the lower bound $l = \mathbf{1}^T \cdot \boldsymbol{p} - 1.25N\mu$ to the optimal objective function value of problem (14) is updated (Lines 35-36), and the value of the barrier parameter μ is reduced (Line 37). Notice that according to the definition of (24), for a fixed value of $\mu > 0$, it is desirable to minimize $f(\boldsymbol{p}, \mu)$, leading to a balance between the objective function and centrality. When we need to drive the objective function value down, we just reduce the value of the barrier parameter μ, leading to increasing emphasis on the objective function . When μ is driven to zero, we have the convergence to an optimal solution.

D. Related Literature

1) Cutting Plane Methods: The related method of short-step cutting plane method (SSCP) to design a slow adaptive OFDMA scheme in wireless system was first used in [21]. The main advantage of the proposed LLBCP approach compared to approach of [20] pertains to its lower computational complexity [20], primarily due to the linearity of the associated objective function. Specifically, while the method of [20] can be directly applied to the problem at hand, our objective function, as opposed to [20], is linear, which when exploited judiciously can bring about improvement in computational complexity.

For a linear objective function, LLBCP algorithm incorporates the linear objective function into a barrier function explicitly (defined in (24)), and iteratively reduces the value of a barrier parameter μ, leading to increasing emphasis on the objective function, and much greater progress in objective function. For a linear objective function, the μ-center can always be solved in $\mathcal{O}(1)$ Newton steps. Moreover, it avoids having to increase the number of constraints whenever we need to drive the objective function value down. Thus, the computation time at each iteration decreases. Moreover, for a linear objective function, the LLBCP algorithm maintains primal and dual variables and it allows for early termination when the suboptimality is deemed to be within allowable limits. This is in contrast to problems with non-linear objective function, in which it is not easy to maintain the primal and dual variables and whenever infeasibility is encountered, with non-linear objectives the algorithm has to be used until the current iterate again becomes feasible. Thus, although the LLBCP algorithm has the same *order* of complexity as [21], in

[3]Hyperplane and constraint are used exchangeably

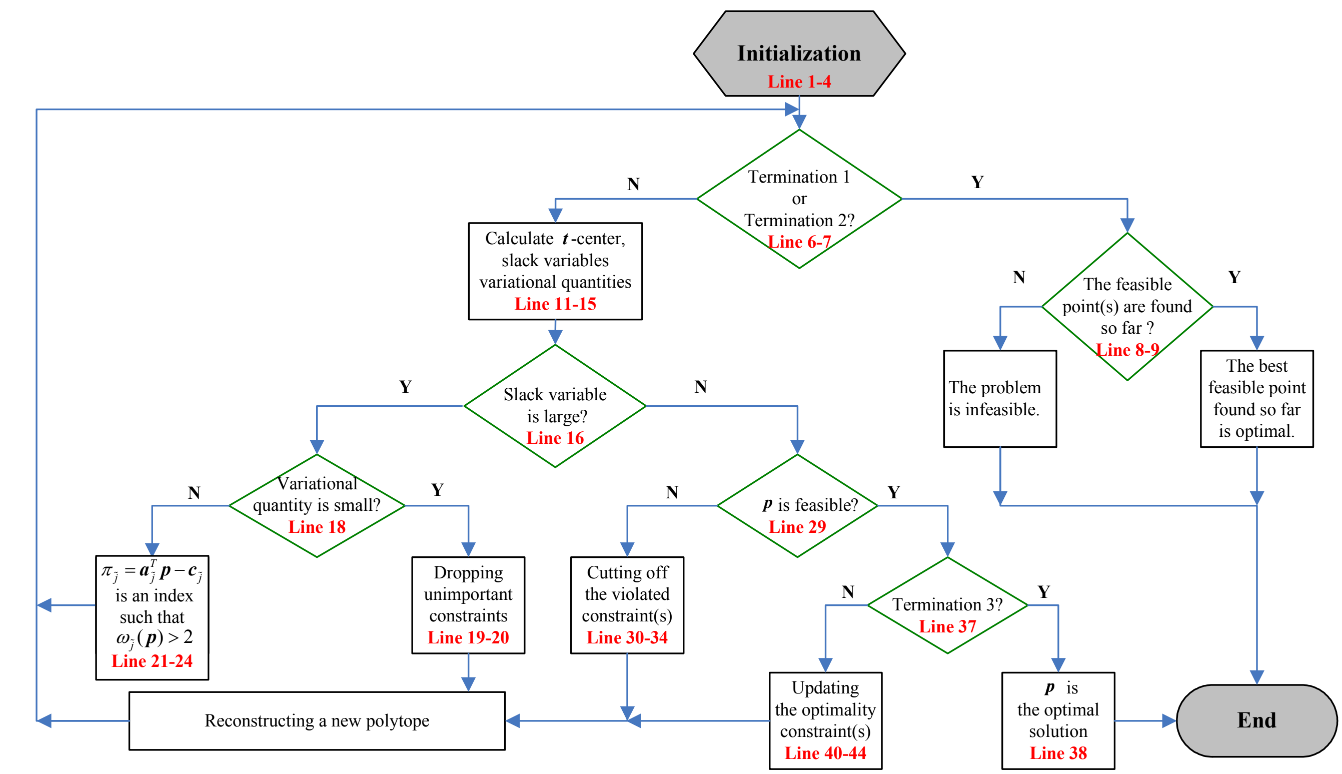

Fig. 1. Flow chart of the LLBCP algorithm.

practice it is computationally more efficient when the objective function is linear. The simulation results verify the advantages of LLBCP algorithm, including higher convergence speed, and more efficient computation at each iteration. We also remark that when the objective function is not linear, LLBCP algorithm is not more efficient than SSCP, since LLBCP has to solve an additional nonlinear optimization problem to find the μ-center, which can not be found in $\mathcal{O}(1)$ Newton steps. Otherwise, in SSCP, the analytic center (AC) can always be found in $\mathcal{O}(1)$ Newton steps.

In terms of the details of the algorithms, the barrier function $f(\boldsymbol{p}, \mu)$ defined in 24, which is the basis for designing the decisions in each iteration, are also different in LLBCP and the method of [21]. Specifically, LLBCP modifies the potential function used in [21] by introducing barrier parameter μ and adjusting the potential function via a shift proportional to barrier parameter. As explained in details in [24], inclusion of the barrier parameter leads to a balance between the objective function and centrality and gradually the barrier parameter is driven to zero, leading to increasing emphasis on the objective function and convergence to an optimal solution. Overall, by capitalizing on linearity of the objective function, the inclusion of the barrier parameter leads to less computational complexity at each iteration of the algorithm. We have tried to put this as the focal distinction of LLBCP from [21] and have summarized the entire section on LLBCP by skipping the and citing the details that are available in [24].

2) Bernstein Inequality Approach: As will be shown in the simulation results, compared with the Bernstein Inequality approach of [14], Bernstein approximation followed by the LLBCP algorithm provides a tighter solution to the power and rate optimization problems. This gain, however, is viable at the expense of higher computational complexity. Specifically, the simulation results show that our proposed approach is about 2(average) times slower than the approach of [14]. This mainly stems from the fact that our proposed Bernstein approximation approach has higher computational complexity due to the first constraint sets in the optimization problem (14), while the approach of [14] can reformulate the stochastic constraints into the deterministic convex constraints. This observation shows that there exists a tradeoff between complexity and performance across these two approaches.

We would to remark that besides better performance, another advantage of the Bernstein *approximation* approach [20] used in this paper is that it is more general in the sense that it applies to CSI perturbations with distributions other than Gaussian or any other perturbation model where channel uncertainty regions are convex compact sets which might not be even known precisely. The Bernstein-type inequality of [19], however, deals with the case of stochastic processes that involve quadratic forms of Gaussian variables.

IV. Simulation Results

In this section, we present extensive simulation results to illustrate the performance of proposed Bernstein approximation approach and LLBCP algorithm for solving the chance-constrained power optimization in MISO interference channels. Also, we provide comparisons with the existing literature,

Algorithm 1 : LLBCP Algorithm

initialization
Set $\epsilon > 0$, $\mu = \frac{1}{\epsilon}$, Set $\pi_k = \frac{1}{\epsilon}, k = 1, \ldots, 2K, \pi_{2K+1} = \frac{1}{\epsilon}\sqrt{K}$.
According to (29), set $\boldsymbol{A} = [\boldsymbol{I}_K \ \ -\boldsymbol{I}_K \ \ \mathbf{1}]^T$, $\boldsymbol{p} = \mathbf{0}_K$,
$\boldsymbol{c} = [\mathbf{0}_K \ \ -\frac{1}{\epsilon}\mathbf{1}^T \ \ -\frac{1}{\epsilon}\sqrt{K}]^T$, and $\boldsymbol{s} = \boldsymbol{A}\boldsymbol{p} - \boldsymbol{c}$.
repeat
  if $\min_k(\boldsymbol{a}_k^T\boldsymbol{p} - c_k) < 10^{-5}\epsilon^3/[2K^{1.5}\log_2(1/\epsilon)]$ or
    $N \geq 4093K\log_2(1/\epsilon)$ **then**
    STOP: the best feasible point found so far is optimal.
      Otherwise, no feasible point is found.
  else
    Find a new approximate μ-center $\boldsymbol{p}(\mu)$ as the minimizer of (24)
    Calculate $\omega_k(\boldsymbol{p}) = \frac{\boldsymbol{a}_k^T\boldsymbol{p} - c_k}{\pi_k}, \forall k$.
    Set $\omega_k(\boldsymbol{p}) = 1$, if k indexes the lowerbound constraint that get added in **Subcase 2.2**.
    If $\omega_{\tilde{j}}(\boldsymbol{p}) > 2$, calculate $\varpi_{\tilde{j}}(\boldsymbol{p}) \triangleq \frac{\boldsymbol{a}_{\tilde{j}}^T(\nabla^2 f(\boldsymbol{p},\mu))^{-1}\boldsymbol{a}_j}{s_{\tilde{j}}^2}$.
    if $\max_k(\omega_k(\boldsymbol{p})) > 2$ **then**
      Case 1 :
      if for some $\tilde{j}$, we have $\varpi_{\tilde{j}}(\boldsymbol{p}) < 0.04$ **then**
        Subcase 1.1:
        Drop the hyperplane $\boldsymbol{a}_{\tilde{j}}$;
      else
        Subcase 1.2:
        Reset $\pi_{\tilde{j}} = \boldsymbol{a}_{\tilde{j}}^T\boldsymbol{p} - c_{\tilde{j}}$, where $\tilde{j}$ be an index such that $\omega_{\tilde{j}}(\boldsymbol{p}) > 2$.
      end if
    end if
    if $\max_k(\omega_k(\boldsymbol{p})) \leq 2$ **then**
      Case 2 :
      if $\boldsymbol{p}$ is not feasible in the problem (14) **then**
        Subcase 2.1:
        For $\tilde{k} \in \tilde{\boldsymbol{K}}$, generate hyperplane(s) according to
        $\frac{\boldsymbol{\nabla}_{i,\tilde{k}}^T}{\|\boldsymbol{\nabla}_{i,\tilde{k}}\|} \cdot \boldsymbol{p} \leq \frac{\boldsymbol{\nabla}_{i,\tilde{k}}^T}{\|\boldsymbol{\nabla}_{i,\tilde{k}}\|} \cdot \boldsymbol{p}_i, \quad \tilde{k} \in \tilde{\boldsymbol{K}}$
        Set $\pi_{N+\tilde{k}} = \frac{\boldsymbol{\nabla}_{i,\tilde{k}}^T}{\|\boldsymbol{\nabla}_{i,\tilde{k}}\|} \cdot \boldsymbol{p}_i - \frac{\boldsymbol{\nabla}_{i,\tilde{k}}^T}{\|\boldsymbol{\nabla}_{i,\tilde{k}}\|} \cdot \boldsymbol{p}$.
        Set $N \leftarrow N + |\tilde{\boldsymbol{K}}|$
      else
        Subcase 2.2:
        if $1.25N\mu < \epsilon$ **then**
          $\boldsymbol{p}_i$ is the optimal solution, and STOP.
        else
          set the lower bound $l = \mathbf{1}^T\boldsymbol{p} - 1.25N\mu$ on optimal objective function of (29).
          Let l_{prev} denote previous lower bound. If $l_{prev} < l$, replace $\mathbf{1}^T\boldsymbol{p} \geq l_{prev}$ by $\mathbf{1}^T\boldsymbol{p} \geq l$.
          Set $\mu \leftarrow \theta\mu$, where $\theta \in (0.5, 1)$.
        end if
      end if
    end if
  end if
until STOP

especially with the application of the Bernstein-type inequality approach of [14], which has been developed for beamforming design in broadcast channels, on the power optimization problem in interference channels studied in this paper. For this purpose the beamforming directions are selected from a pre-designed codebook that based on their distance to the channel pseudoinverse and we use our proposed method to obtain the optimal powers to be allocated to the beamformers. Throughout the section and in plots we refer to the approach of [14] by "Bern-Ineq". This section is concluded by providing simulation results on power optimization on broadcast channels and compare it with the relevant literature.

Throughout the simulations for the interference channel we assume a $K = 4$ user interference channels with 4-antenna transmitters and single-antenna receivers. The channel coefficients are generated as i.i.d. complex Gaussian random variables with zero mean and unit variance. We apply a simple method to ensure that the direct links are stronger than the interference links [26] and set the strengthened direct channel links by a multiplicative constant 5. For simplicity, we assume that the covariance matrices of the channel uncertainty vectors are $\boldsymbol{C}_{ij} = \mathbb{E}[\boldsymbol{\delta}_{ij}^H\boldsymbol{\delta}_{ij}] = \sigma^2\boldsymbol{I}$, and consider cases of $\sigma^2 = 0.001$ and $\sigma^2 = 0.002$ in the simulations. The noise variances of all receivers are set to $\eta^2 = 0.01$, and the outage probabilities are set $\boldsymbol{\varepsilon} = \varepsilon \cdot \mathbf{1}$ and consider cases of $\varepsilon = 0.05$ and $\varepsilon = 0.1$ in the simulations. Moreover, we assume that all receivers have the identical target rates $\boldsymbol{r} = r \cdot \mathbf{1}$, or equivalently, identical target SINR levels $\boldsymbol{\alpha} = \alpha \cdot \mathbf{1}$. As the rate and SINR values are uniquely interchangeable according to (5) all the simulations are provided for SINR values. Finally, the Monte-Carlo simulations are performed over 2000 sets randomly generated channels.

A. *Rate-constrained Power Optimization*

In this section we focus on the rate-constrained power optimization problem $\mathcal{P}_1(\boldsymbol{r}, \boldsymbol{\varepsilon})$ Fig. 2 demonstrates the feasibility frequency of the problem versus target SINR and Fig. 3 evaluates the average transmission power of required by our proposed methods and by Bern-Ineq [14]. Both plots conform in that our proposed method demonstrates improved performance compared with Bern-Ineq [14] by achieving higher feasibility rate and lower average transmission power under different choices for parameters σ^2 and ε. We also implement the bisection technique for our proposed methods and Bern-Ineq [14]. From the simulation result shown in Fig. 4, we can see that the bisection technique can reduce the conservativeness of our proposed methods and Bern-Ineq, and our proposed methods still has a bit better performance. However, the computation time of our proposed methods is about 1.2-3.8 times (average about 2.2 times) of the computation time of Bern-Ineq [14].

From Fig. 2-4, we can see the impact of the channel uncertainty σ^2 and outage probability requirement ε as follows,

- When the channel uncertainty σ^2 increases, it takes more power to meet the rate (SINR) outage requirement. For a fixed channel uncertainty level σ^2, as the target SINR value α increases, it becomes exceedingly difficult to meet the outage requirement; and moreover, the transmit power increases drastically, and even making the optimization problem infeasible, and thus feasibility rate decreases.
- It is seen that as the outage requirement ε becomes more stringent, i.e., when ε becomes smaller, it takes more power to meet the SINR outage requirement, the maximum feasible SINR value α becomes smaller, and feasibility rate decreases.

In order to highlight the advantages of using the information about the uncertainty model, Fig. 2 and Fig. 3 also depict the feasibility rate and the average transmit power rate corresponding to a setting where the uncertainty model information

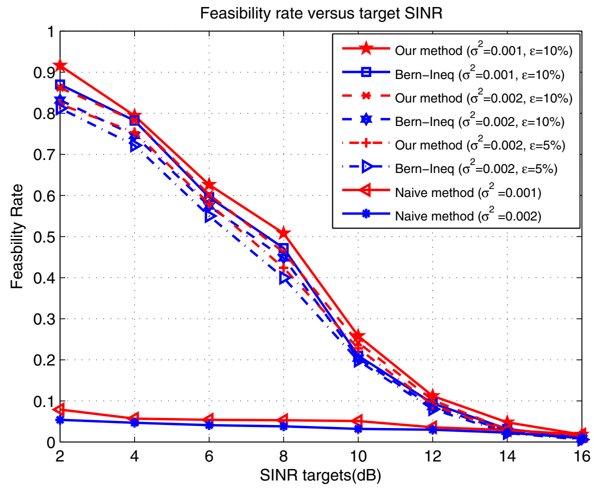

Fig. 2. Feasibility rate versus target SINR in a MISO interference channel.

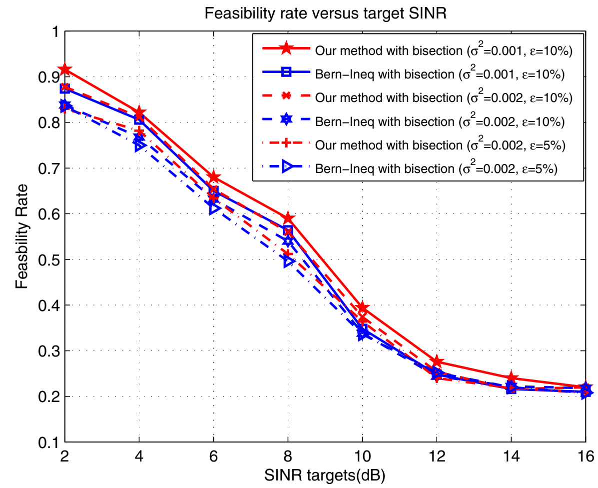

Fig. 4. Feasibility rate versus target SINR in a MISO interference channel (with bisection).

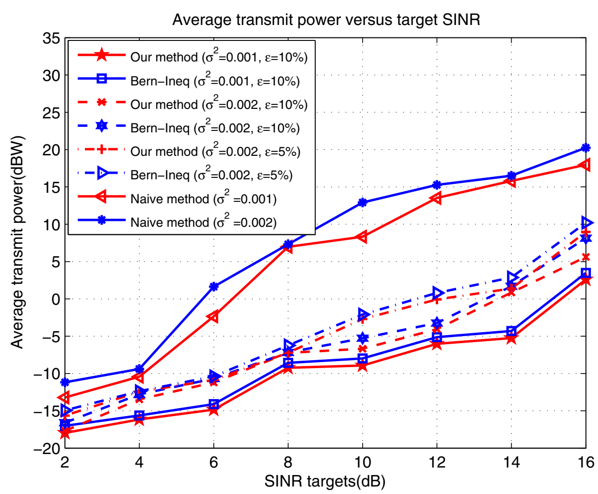

Fig. 3. Average transmit power versus target SINR in a MISO interference channel.

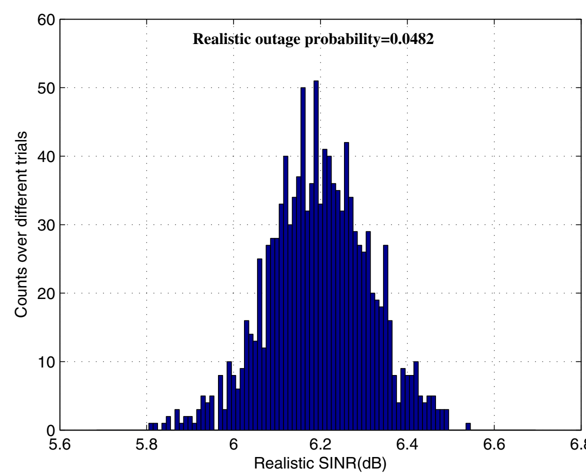

Fig. 5. Histograms of realistic SINR of our proposed method.

is not used. As indicated by both figures, not utilizing this information incurs a considerable penalty, i.e., reduction in feasibility rate and increase in transmit power.

In order to gain more insight into the behavior of these two methods, we verify the actual outage probability by plotting the histograms of the realistic SINR in Fig. 5 and Fig. 6. The system parameters are $\alpha = 6$dB, $\sigma^2 = 0.002$ and $\varepsilon = 0.05$. The simulation results show that the actual outage probability in both methods is smaller than the target outage probability, corroborating that both methods provide conservative approximations to the original probabilistic constraints. However, from Figs 5 and 6 it is observed that the realistic outage probability yielded by our approach is closer to the target value $\varepsilon = 0.05$, indicating that it is less conservative.

Next, we provide simulation results for assessing the performance of the LLBCP algorithm and the relevant comparisons with the SSCP algorithm used in [21]. Specifically, Fig. 7 demonstrates the convergence of LLBCP and SSCP algorithms for a feasible channel realization, when $\alpha = 6$dB, $\sigma^2 = 0.002$, and $\varepsilon = 0.05$. According to Fig. 7, LLBCP algorithm exhibits a higher convergence speed than SSCP algorithm. It is also observed that LLBCP algorithm is more computationally efficient due to the less number of constraints at each iteration as discussed in Section III-C. The simulations results confirm the advantages of LLBCP algorithm. Detailed discussions on convergence is available in [24].

B. Max-min Rate Optimization

We next consider the max-min rate optimization problems with individual power constraints in MISO interference channels. We assume that the transmitters have the same maximum allowable individual transmit power $\bar{\boldsymbol{p}} = \bar{p}\cdot\mathbf{1}$, where $\bar{p}$ denotes the total transmission power of each transmitter.

Fig. 8 shows the maximum achievable SINR versus the maximum allowable transmission power for different outage probability ε values and for different channel uncertainty σ^2 levels. Note that for a given allowable transmission power, the maximum achievable SINR decreases as the channel uncertainty σ^2 increases, or as the outage probability ε decreases.

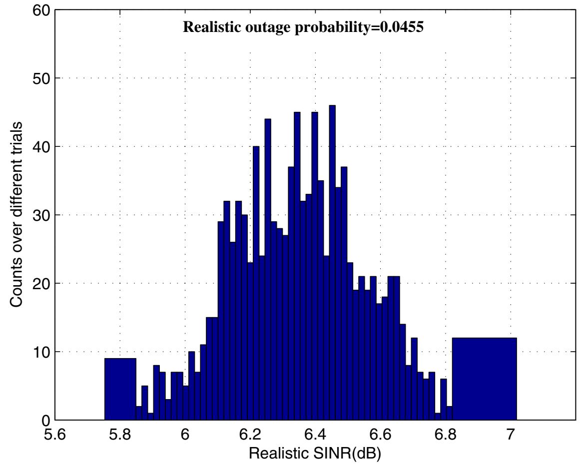

Fig. 6. Histograms of realistic SINR of Bern-Ineq [14].

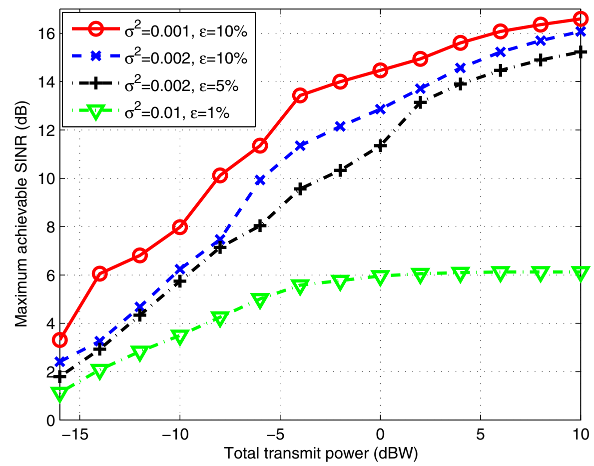

Fig. 8. Maximum achievable SINR versus maximum allowable individual transmit power.

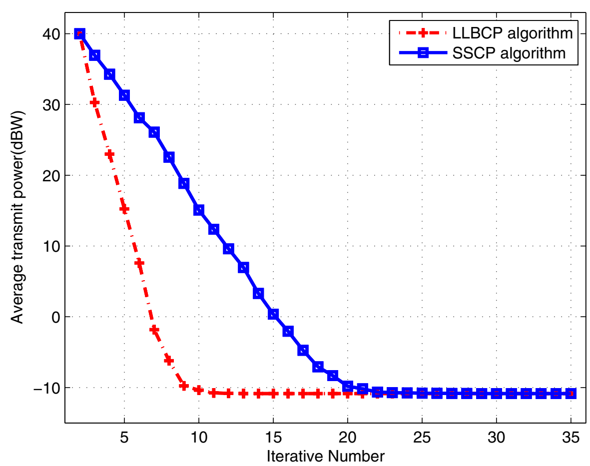

Fig. 7. Convergence of LLBCP algorithm and the cutting plane algorithm used in [21].

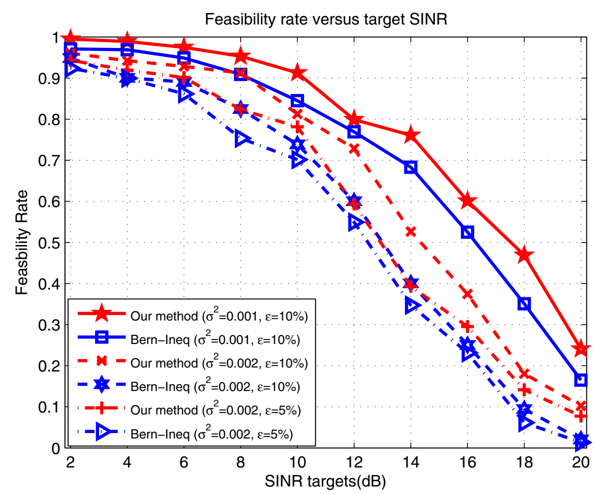

Fig. 9. Feasibility rate versus target SINRin the MISO broadcast channel.

It is noteworthy that for the channel uncertainty $\sigma^2 = 0.01$, the maximum achievable SINR is always smaller than 6.2dB, even if the maximum allowable total transmit increases from 2dBW to 10dBW.

C. Rate-constrained Power Optimization in Broadcast Channels

In this section we evaluate the application of the proposed Bernstein approximation method for power optimization in MISO broadcast channels. We assume a broadcast channel of one 4-antenna base-station serving $K = 4$ single-antenna users. Channel coefficients are distributed according to i.i.d. complex Gaussian random variables with zero mean and unit variance. The performance of feasibility rate versus target SINR, and transmission power versus target SINR are shown in Fig. 9 and Fig. 10, respectively. The results in Fig. 9 and Fig. 10 imply that our proposed method is amenable to support a wider range of target SINRs under the different parameter settings of σ^2 and ε.

We also verify the actual outage probability at $\alpha = 12$dB, $\sigma^2 = 0.002$ and $\varepsilon = 0.05$. From the histograms of the realistic SINR shown in Fig. 11 and Fig. 12, we can see that the outage probability of both methods is smaller than the target outage probability, and Bern-Ineq [14] provides a slightly more conservative approximation to the original probabilistic constraint.

V. Conclusions

We have treated the problems of robust power allocation in MISO interference channels with stochastic and imperfect transmitter-side CSI. The multi-antenna transmitters are assumed to employ beamforming directions based on predesigned beamforming codebooks. For MISO interference channels, we have considered the minimizing the transmission power and maximizing the rate of the weakest users in the network subject to constraints on the rate outage probabilities. The key contribution is to employ the Bernstein approximation to conservatively transform the probabilistic constraints into deterministic and convex ones, and consequently convert the

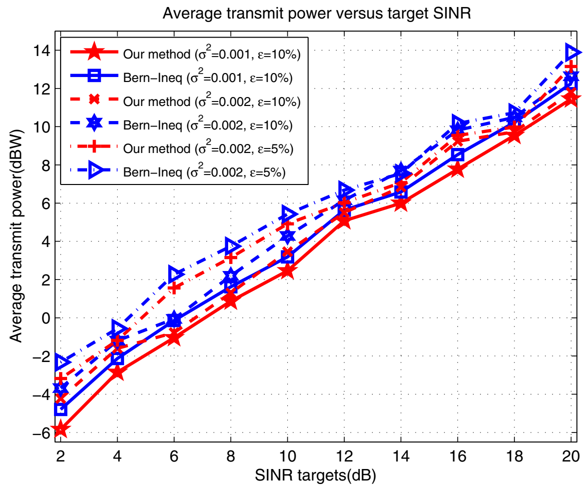

Fig. 10. Average transmit power versus target SINR in the MISO broadcast channel.

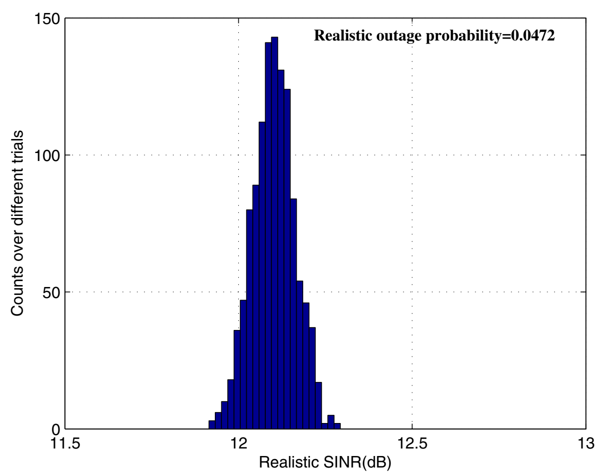

Fig. 11. Histograms of realistic SINR of our proposed method.

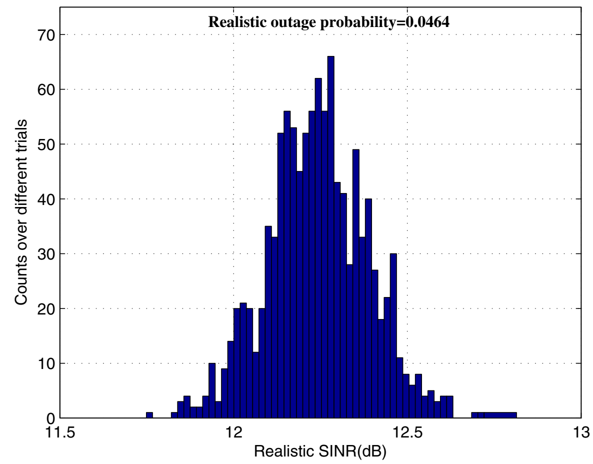

Fig. 12. Histograms of realistic SINR of Bern-Ineq [14].

original stochastic optimization problems into convex optimization problems. We also propose to propose to use the long-step logarithmic barrier cutting plane (LLBCP) algorithm to solve the deterministic problems. Extensive simulation results establish the gains of the proposed approach over the existing literature.

Appendix A
Proof of Proposition 1

Given the definitions of SINR_i and R_i given in (4) and (5), respectively, the constraint $R_i \leq r_i$ can be equivalently state as

$$F_i(\boldsymbol{p}, \boldsymbol{\zeta}_i) \ \geq \ 0 \ , \tag{26}$$

where

$$F_i(\boldsymbol{p}, \boldsymbol{\zeta}_i) \ \triangleq \ \alpha_i \eta^2 + \alpha_i \sum_{j \neq i} p_j \zeta_{ij} - p_i \zeta_{ii} \ , \tag{27}$$

which is in the affine form of (12) and also

$$\zeta_{ij} \ \triangleq \ |\boldsymbol{h}_{ij} \boldsymbol{g}_j|^2 = |\hat{\boldsymbol{h}}_{ij} \boldsymbol{g}_j + \boldsymbol{\delta}_{ij} \boldsymbol{g}_j|^2 \ , \tag{28}$$

Hence, based on the discussions in Section II-B3 and Equations (8) and (9) the rate-constrained power optimization problem $\mathcal{P}(\boldsymbol{r}, \boldsymbol{\varepsilon})$ can be approximated by

$$\hat{\mathcal{P}}_1(\boldsymbol{r}, \boldsymbol{\varepsilon}) \triangleq \begin{cases} \min_{\{p_i\}} & \sum_{i=1}^K p_i \\ \text{s.t.} & \inf_{t_i} \{\hat{\Psi}_i(\boldsymbol{p}, t_i) - t_i \log \varepsilon_i\} \ \leq 0, \forall i \\ & 0 \ \leq \ p_i \ , \quad \forall i \\ & 0 \ \leq \ t_i \ , \quad \forall i \end{cases} \tag{29}$$

where

$$\begin{aligned} &\hat{\Psi}(\boldsymbol{p}, t_i) \\ &\quad = t_i \log \mathbb{E}\{\exp[t_i^{-1} F_i(\boldsymbol{p}, \boldsymbol{\zeta}_i)]\} \\ &\quad \overset{(27)}{=} t_i \log \mathbb{E}\left[\exp\left(t_i^{-1}\alpha_i\eta^2 + t_i^{-1}\alpha_i \sum_{j\neq i} p_j \zeta_{ij} - t_i^{-1} p_i \zeta_{ii}\right)\right] \\ &\quad = \alpha_i \eta^2 + t_i \sum_{j \neq i} \log \mathbb{E}\left[\exp\left(t_i^{-1}\alpha_i p_j \zeta_{ij}\right)\right] \\ &\quad + t_i \log \mathbb{E}\left[\exp\left(-t_i^{-1} p_i \zeta_{ii}\right)\right] \ . \end{aligned} \tag{30}$$

In order to further simplify (30) we next find the distributions of the random variables ζ_{ij}. This this end, by recalling that the covariance matrix of $\boldsymbol{\delta}_{ij}$ is denoted by $\boldsymbol{C}_{ij}$ for the random vector $\boldsymbol{s}_{ij} \triangleq \boldsymbol{\delta}_{ij} \boldsymbol{C}_{ij}^{-1/2}$ we have $\boldsymbol{s}_{ij} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{I})$, and consequently,

$$\boldsymbol{\delta}_{ij} \boldsymbol{g}_j = \boldsymbol{s}_{ij} \ \boldsymbol{C}_{ij}^{1/2} \ \boldsymbol{g}_j \ \sim \mathcal{N}_{\mathbb{C}}(0, \sigma_{ij}^2) \ , \tag{31}$$

where

$$\sigma_{ij}^2 \triangleq \|\boldsymbol{C}_{ij}^{1/2} \ \boldsymbol{g}_j\|^2 \ .$$

Therefore, corresponding to ζ_{ij} defined in (28) we have

$$\frac{2\zeta_{ij}}{\sigma_{ij}^2} \ \sim \ \chi_2^2 \left(\frac{2|\hat{\boldsymbol{h}}_{ij} \boldsymbol{g}_j|^2}{\sigma_{ij}^2}\right) \ . \tag{32}$$

By recalling the fact that for $X \sim \chi_2^2(\lambda)$, we have

$$\mathbb{E}\{\exp(\ell X)\} = \frac{\exp\left(\frac{\lambda \ell}{1-2\ell}\right)}{(1-2\ell)} \ , \quad \text{with} \ \ \ell < \frac{1}{2} \ . \tag{33}$$

we find that

$$\mathbb{E}\left[\exp\left(\ell\zeta_{ij}\right)\right] = \mathbb{E}\left[\exp\left(\frac{\ell\sigma_{ij}^2}{2}\cdot\frac{2\zeta_{ij}}{\sigma_{ij}^2}\right)\right] \overset{(32)-(33)}{=} \frac{\exp\left(\frac{\ell|\hat{\boldsymbol{h}}_{ij}\boldsymbol{g}_j|^2}{1-\ell\sigma_{ij}^2}\right)}{1-\ell\sigma_{ij}^2}\ , \quad \ell\sigma_{ij}^2 < 1\ . \quad (34)$$

Hence, based on (30) and (34) we find that

$$\hat{\Psi}(\boldsymbol{p}, t_i) = \alpha_i\eta^2 + \left(\sum_{j\neq i}\frac{\alpha_i p_j|\hat{\boldsymbol{h}}_{ij}\boldsymbol{g}_j|^2}{1-t_i^{-1}\alpha_i p_j\sigma_{ij}^2} - \frac{p_i|\hat{\boldsymbol{h}}_{ii}\boldsymbol{g}_i|^2}{1-t_i^{-1}p_i\sigma_{ii}^2}\right) - t_i\left(\log\left(1-t_i^{-1}\alpha_i p_j\sigma_{ij}^2\right) - \log\left(1-t_i^{-1}p_i\sigma_{ii}^2\right)\right)\ . \quad (35)$$

Note that satisfying the constraints $\ell\sigma_{ij}^2 < \frac{1}{2}$ necessitate that

$$t_i^{-1}\alpha_i p_j\sigma_{ij}^2 \ < \ 1 \quad \text{and} \quad t_i^{-1}p_i\sigma_{ii}^2 \ < \ 1 \qquad (36)$$

or equivalently

$$t_i \ > \ \rho_i \triangleq \max\left\{p_i\sigma_{ii}^2\ ,\ \max_{j\neq i}\alpha_i p_j\sigma_{ij}^2\right\}\ . \qquad (37)$$

Hence, the rate-constraint power optimization problem $\mathcal{P}(\boldsymbol{r}, \boldsymbol{\varepsilon})$ can be approximated by $\hat{\mathcal{P}}_1(\boldsymbol{r}, \boldsymbol{\varepsilon})$ as follows:

$$\begin{array}{ll} \min_{\{p_i\}} & \sum_{i=1}^K p_i \\ \text{s.t.} & \inf_{t_i>\rho_i}\{\hat{\Psi}_i(\boldsymbol{p}, t_i) - t_i\log\varepsilon_i\} \ \leq 0\ ,\ \forall i \\ & 0 \ \leq \ p_i\ , \quad \forall i \end{array} \qquad (38)$$

where $\hat{\Psi}_i(\boldsymbol{p}, t_i)$ is defined in (35).

Appendix B
Proof of Proposition 2

The proof follows a similar line of argument as in [9]. Let us denote the set of powers obtained from solving $\mathcal{P}_2(\bar{\boldsymbol{p}}, \boldsymbol{\varepsilon})$ by $\bar{\boldsymbol{p}}^*$ and their corresponding smallest rates by $\boldsymbol{r}^*$. From the definition of $\mathcal{P}_2(\bar{\boldsymbol{p}}, \boldsymbol{\varepsilon})$ we have $p_i \leq \bar{p}_i, \forall i$ and $\min_i r_i^* = \mathcal{P}_2(\bar{\boldsymbol{p}}, \boldsymbol{\varepsilon})$ concluding that for all i we have $r_i^* \geq \mathcal{P}_2(\bar{\boldsymbol{p}}, \boldsymbol{\varepsilon})$. As a result from the definition of $\mathcal{S}(\bar{\boldsymbol{p}}, \boldsymbol{\varepsilon}, a)$ we find that for the choice of $\boldsymbol{p}^*$ the choice of $b = 1$ is achievable for $\mathcal{S}(\bar{\boldsymbol{p}}, \boldsymbol{\varepsilon}, \mathcal{P}_2(\bar{\boldsymbol{p}}, \boldsymbol{\varepsilon}))$ and therefore $\mathcal{S}(\bar{\boldsymbol{p}}, \boldsymbol{\varepsilon}, \mathcal{P}_2(\bar{\boldsymbol{p}}, \boldsymbol{\varepsilon})) \leq 1$.

Next we show that $\mathcal{S}(\bar{\boldsymbol{p}}, \boldsymbol{\varepsilon}, \mathcal{P}_2(\bar{\boldsymbol{p}}, \boldsymbol{\varepsilon}))$ cannot be less than one. Let us denote the set of powers obtained by solving $\mathcal{S}(\bar{\boldsymbol{p}}, \boldsymbol{\varepsilon}, \mathcal{P}_2(\bar{\boldsymbol{p}}, \boldsymbol{\varepsilon}))$ by $\boldsymbol{p}^{**}$ with corresponding rates $\boldsymbol{r}^{**}$. From the definition of $\mathcal{S}(\bar{\boldsymbol{p}}, \boldsymbol{\varepsilon}, \mathcal{P}_2(\bar{\boldsymbol{p}}, \boldsymbol{\varepsilon}))$ we clearly have $r_i^{**} \geq \mathcal{P}_2(\bar{\boldsymbol{p}}, \boldsymbol{\varepsilon})$ for all i. If $\mathcal{S}(\bar{\boldsymbol{p}}, \boldsymbol{\varepsilon}, \mathcal{P}_2(\bar{\boldsymbol{p}}, \boldsymbol{\varepsilon})) < 1$, i.e., if $\max_i \frac{p_i^{**}}{\bar{p}_i} = c < 1$, then we define the set of powers $\hat{p}_i = p_i^{**}/c$. Set of powers $\{\hat{p}_i\}$ clearly satisfy the power constraints and we have their corresponding SINRs satisfying

$$\hat{\mathsf{SINR}}_i = \frac{\hat{p}_i\left|\boldsymbol{h}_{ii}\boldsymbol{g}_i\right|^2}{\eta^2 + \sum\limits_{j\neq i}\hat{p}_j\left|\boldsymbol{h}_{ij}\boldsymbol{g}_j\right|^2} = \frac{\frac{p_i^{**}}{c}\left|\boldsymbol{h}_{ii}\boldsymbol{g}_i\right|^2}{\eta^2 + \sum\limits_{j\neq i}\frac{p_j^{**}}{c}\left|\boldsymbol{h}_{ij}\boldsymbol{g}_j\right|^2} \quad (39)$$

$$= \frac{p_i^{**}\left|\boldsymbol{h}_{ii}\boldsymbol{g}_i\right|^2}{c\eta^2 + \sum\limits_{j\neq i}p_j^{**}\left|\boldsymbol{h}_{ij}\boldsymbol{g}_j\right|^2} > \frac{p_i^{**}\left|\boldsymbol{h}_{ii}\boldsymbol{g}_i\right|^2}{\eta^2 + \sum\limits_{j\neq i}p_j^{**}\left|\boldsymbol{h}_{ij}\boldsymbol{g}_j\right|^2} \quad (40)$$

Since $c < 1$ we have $\hat{\mathsf{SINR}}_i > \mathsf{SINR}_i^{**}$. Therefore, we have found a set of powers which satisfy the power constraints and yet yield a strictly larger smallest SINR compared to what the powers $\boldsymbol{p}^{**}$ obtain. This contradicts the optimality of $\boldsymbol{p}^{**}$ and therefore $\mathcal{S}(\bar{\boldsymbol{p}}, \boldsymbol{\varepsilon}, \mathcal{P}_2(\bar{\boldsymbol{p}}, \boldsymbol{\varepsilon})) = 1$. The strict monotonicity and continuity of $\mathcal{S}(\bar{\boldsymbol{p}}, \boldsymbol{\varepsilon}, \mathcal{P}_2(\bar{\boldsymbol{p}}, a)$ in a, at any strictly feasible region, follows from a similar line of argument.

References

[1] W. Xu and X. Wang, "Pricing-based distributed downlink beamforming in multi-cell OFDMA networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 9, pp. 1605–1613, Oct. 2012.

[2] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.

[3] M. Payaro, A. Pascual-Iserte, and M. A. Lagunas, "Robust power allocation designs for multiuser and multiantenna downlink communication systems through convex optimization," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1390–1401, Sept. 2007.

[4] J. Wang and M. Payaro, "On the robustness of transmit beamformings," *IEEE Trans. Signal Process.*, vol. 58, no. 11, pp. 5933–5938, Nov. 2010.

[5] 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.

[6] M. B. Shenouda and T. N. Davidson, "Convex conic formulations of robust downlink precoder designs with quality of service constraints," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 4, pp. 714–724, Dec. 2007.

[7] A. Tajer, N. Prasad, and X. Wang, "Robust transceiver design for the multi-user interference channel," in *Proc. 2010 IEEE International Conf. Commun.*

[8] ——, "Robust beamforming for multi-cell downlink transmission," in *Proc. 2010 IEEE International Symp. Inf. Theory.*

[9] ——, "Robust linear precoder design for multi-cell downlink transmission," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 235–251, Jan. 2011.

[10] M. Ding and S. D. Blostein, "Maximum mutual information design for MIMO systems with imperfect channel knowledge," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 4793–4801, Oct. 2010.

[11] X. Zhang, D. P. Palomar, and B. Ottersten, "Statistically robust design of linear MIMO transceivers," *IEEE Trans. Signal Process.*, vol. 564, no. 8, pp. 3678–3689, Aug. 2008.

[12] G. Zheng, K.-K. Wong, and T.-S. Ng, "Energy-efficient multiuser SIMO: achieving probabilistic robustness with Gaussian channel uncertainty," *IEEE Trans. Commun.*, vol. 57, no. 7, pp. 1866–1878, July 2009.

[13] Y. Rong, S. A. Vorobyov, and A. B. Gershman, "Robust linear receivers for multi-access space-time block-coded MIMO systems: a probabilistically constrained approach," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1560–1570, Aug. 2006.

[14] K.-Y. Wang, T.-H. Chang, W.-K. Ma, A. M.-C. So, and C.-Y. Chi, "Probabilistic SINR constrained robust transmit beamforming: a Bernstein-type inequality based conservative approach," in *Proc. 2011 IEEE International Conf. Speech Signal Process.*, pp. 3080–3083.

[15] 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." Available: arXiv:1108.0982.

[16] M. B. Shenouda and T. N. Davidson, "Outage-based designs for multi-user transceivers," in *Proc. 2009 IEEE International Conf. Speech Signal Process.*, vol. 1.

[17] N. Vucic and H. Boche, "A tractable method for chance-constrained power control in downlink multiuser MISO systems with channel uncertainty," *IEEE Signal Process. Lett.*, vol. 16, no. 5, pp. 346–349, May 2009.

[18] B. Chalise, S. Shahbazpanahi, A. Czylwik, and A. Gershman, "Robust downlink beamforming based on outage probability specifications," *IEEE Trans. Wireless Commun.*, vol. 6, no. 10, pp. 3498–3505, Oct. 2007.

[19] I. Bechar, "A Bernstein-type inequality for stochastic processes of quadratic forms of Gaussian variables." Available: arXiv:0909.3595v1.

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

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

[22] J. Chen, W. Xu, S. He, Y. Sun, P. Thulasiraman, and X. Shen, "Utility-based asynchronous flow control algorithm for wireless sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 7, pp. 1116–1126, July 2010.

[23] J. E. Mitchell and S. Ramaswamy, "A long-step cutting plane algorithm for linear and convex programmings," *Annals Operations Research*, vol. 99, pp. 95–122, Apr. 2000.

[24] J. E. Mitchell, "Polynomial interior point cutting plane methods," *Optimization Methods Software*, vol. 18, no. 5, pp. 507–534, Oct. 2003.

[25] J.-L. Goffin and J.-P. Vial, "Convex nondifferentiable optimization: a survey focussed on the analytic center cutting plane method," *Optimization Methods Software*, vol. 17, no. 5, pp. 805–867, 2002.

[26] S. Timotheou, I. Krikidis, and B. Ottersten., "MISO interference channel with QoS and RF energy harvesting constraints," in *2013 IEEE International Conf. Commun.*

Weiqiang Xu (M'09-SM'14) received his M.Sc. degree in Communications and Information System from Southwest Jiao-Tong University, China, and his Ph.D. degree in Control Science and Engineering from Zhejiang University, China, in 2003 and 2006, respectively. He also was a postdoctor at Zhejiang University, and visiting scholar at Columbia University. He is currently a professor with the School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou, China. He is currently a committee membership of the Youth Committee of China Institute of Communications, and the Committee of Chinese Institute of Electronics on Information Theory. His research interests include multi-cell networks, Ad Hoc networks, wireless sensor networks, wireless optical networks, congestion control, and smart grid, etc. He has served as publication chair for Chinacom 2014, a TPC member for IEEE ICC 2013-2014, IEEE Globecom 2012-2013, WCSP 2013, IWCMC 2009, 2010, 2014, PMSN 2009, IHMSC 2009-2013, CCCC 2012-2013. He has also served as a peer reviewer for a variety of IEEE journals and conferences.

Ali Tajer [S'05-M'10] received the B.Sc. and M.Sc. degrees in Electrical Engineering from Sharif University of Technology, and the M.A. degree in Statistics and the Ph.D. degree in Electrical Engineering from Columbia University. During 2010-2012 he was a Postdoctoral Research Associate at Princeton University and an Adjunct Assistant Professor at Columbia University. Since Fall 2012 he has been an Assistant Professor of Electrical and Computer Engineering at Wayne State University. His research interests lie in the general areas of applied statistics, network information theory, and energy systems.

Xiaodong Wang (S'97-M'98-SM'04-F'08) received the Ph.D. degree in Electrical Engineering from Princeton University. He is a Professor of Electrical Engineering at Columbia University in New York. Dr. Wang's research interests fall in the general areas of computing, signal processing and communications, and has published extensively in these areas. Among his publications is a book entitled "Wireless Communication Systems: Advanced Techniques for Signal Reception", published by Prentice Hall in 2003. His current research interests include wireless communications, statistical signal processing, and genomic signal processing. Dr. Wang received the 1999 NSF CAREER Award, the 2001 IEEE Communications Society and Information Theory Society Joint Paper Award, and the 2011 IEEE Communication Society Award for Outstanding Paper on New Communication Topics. He has served as an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON SIGNAL PROCESSING, and the IEEE TRANSACTIONS ON INFORMATION THEORY. He is a Fellow of the IEEE and listed as an ISI Highly-cited Author.

Saleh Alshomrani is an Associate Professor of Information Systems Department at King Abdulaziz University. He is also serving now as the Vice-Dean of Faculty of Computing and Information Technology, North Jeddah Campus at King Abdulaziz University. He earned his Bachelor degree in Computer Science (BSc) from King Abdulaziz University, Saudi Arabia 1997. He received his Master degree in Computer Science from Ohio University, USA 2001. He earned his Ph.D. in Computer Science from Kent State University 2008, Ohio, USA. His research areas include: Web Distributed Systems and Internet Computing, Data Mining, and Algorithms.