

# Sample Approximation-Based Deflation Approaches for Chance SINR Constrained Joint Power and Admission Control

Ya-Feng Liu, Mingyi Hong, and Enbin Song

## Abstract

Consider the joint power and admission control (JPAC) problem for a multi-user single-input single-output (SISO) interference channel. Most existing works on JPAC assume the perfect instantaneous channel state information (CSI). In this paper, we consider the JPAC problem with the imperfect CSI, that is, we assume that only the channel distribution information (CDI) is available. We formulate the JPAC problem into a chance (probabilistic) constrained program, where each link's SINR outage probability is enforced to be less than or equal to a specified tolerance. To circumvent the computational difficulty of the chance SINR constraints, we propose to use the sample (scenario) approximation scheme to convert them into finitely many simple linear constraints. Furthermore, we reformulate the sample approximation of the chance SINR constrained JPAC problem as a composite group sparse minimization problem and then approximate it by a second-order cone program (SOCP). The solution of the SOCP approximation can be used to check the simultaneous supportability of all links in the network and to guide an iterative link removal procedure (the deflation approach). We exploit the special structure of

Part of this work has been presented in the 14th IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Darmstadt, Germany, June 16–19, 2013 [12].

Y.-F. Liu (corresponding author) is with the State Key Laboratory of Scientific and Engineering Computing, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, 100190, China (e-mail: yafliu@lsec.cc.ac.cn). Y.-F. Liu is supported by the National Natural Science Foundation of China under Grants 11301516, 11331012, and 11571221.

M. Hong is with the Department of Industrial and Manufacturing Systems Engineering (IMSE), Iowa State University, Ames, IA 50011, USA (email: mingyi@iastate.edu). M. Hong is supported in part by NSF under Grant CCF-1526078 and by AFOSR under grant 15RT0767.

E. Song is with the Department of Mathematics, Sichuan University, Chengdu, 610064, China (e-mail: e.b.song@163.com). E. Song is supported by the National Natural Science Foundation of China under Grant 61473197.

the SOCP approximation and custom-design an efficient algorithm for solving it. Finally, we illustrate the effectiveness and efficiency of the proposed sample approximation-based deflation approaches by simulations.

### **Index Terms**

Chance SINR constraint, group sparse, power and admission control, sample approximation.

## **I. INTRODUCTION**

Joint power and admission control (JPAC) has been recognized as an effective tool for interference management in cellular, ad hoc, and cognitive underlay wireless networks for two decades. Generally speaking, there are two kinds of JPAC: one is to support a maximum number of links at their specified signal to interference plus noise ratio (SINR) targets while using minimum total transmission power when all links in the network cannot be simultaneously supported [1]–[22], and the other is to determine whether a new arrival link can be admitted to the network while maintaining the SINR of all already admitted links above their required SINR levels [23]–[25]. This paper focuses on the former one, which not only determines the set of links that must be turned off and rescheduled (possibly along orthogonal resource dimensions such as time, space, or frequency slots), but also alleviates the difficulties of the convergence of stand-alone power control algorithms. For example, a longstanding issue associated with the Foschini-Miljanic algorithm [3] is that, it does not converge when the preselected SINR levels are infeasible. In this case, a JPAC approach must be adopted to determine which links to be removed.

The JPAC problem can be solved to global optimality by checking the simultaneous supportability of every subset of links. However, the computational complexity of this enumeration approach grows exponentially with the total number of links. Theoretically, the problem is known to be NP-hard to solve (to global optimality) and to approximate (to constant ratio of global optimality) [1], [4], [8], so various heuristic algorithms have been proposed [1]–[22]. In particular, the reference [1] proposed a convex approximation-based algorithm, called linear programming deflation (LPD) algorithm. Instead of solving the original NP-hard problem directly, the LPD algorithm solves an appropriate LP approximation of the original problem at each iteration and use its solution to guide the removal of interfering links. The removal procedure is terminated

if all the remaining links in the network are simultaneously supportable. The reference [8] developed another LP approximation-based new linear programming deflation (NLPD) algorithm for the JPAC problem. In [8], the JPAC problem is first equivalently reformulated as a sparse  $\ell_0$ -minimization problem and then its  $\ell_1$ -convex approximation is used to derive a LP, which is different from the one in [1]. Again, the solution to the derived LP can guide an iterative link removal procedure, and the removal procedure is terminated if all the remaining links in the network are simultaneously supportable. Similar ideas were also used in [13], [16], [20] to solve the joint beamforming and admission control problem for the cellular downlink network.

Most of the aforementioned works on the joint power/beamforming and admission control problem assume the perfect instantaneous channel state information (CSI) except [1], [4], [14], [22]. In [1], the authors also considered the worst-case robust JPAC problem with bounded channel estimation errors. The key in [1] is that the LP approximation with bounded uncertainty can be equivalently rewritten as a second-order cone program (SOCP). The overall approximation algorithm remains similar to LPD for the case of the perfect CSI, except that the SOCP formulation is used to carry out power control and its solution is used to check whether links are simultaneously supportable in the worst case. In [4], [14], [22], the authors employed the Foschini-Miljanic algorithm [3] or its variants to update the power and then use the updated power to guide the removals of links without assuming the perfect CSI (as long as the SINR can be measured at the receiver and feedbacked to the corresponding transmitter). The Foschini-Miljanic algorithm [3] can leverage the perfect CSI assumption when updating the power, but it does not take admission control into consideration compared to the disciplined convex approximation-based power control algorithms in [1], [8], [13], [16], [20]. This makes the JPAC algorithms where the power is updated by the Foschini-Miljanic algorithm suffer a significant performance loss in the number of supported links compared to those where the power is updated by the disciplined convex approximation-based power control algorithms.

The assumption of the perfect CSI generally does not hold true due to CSI estimation errors or limited CSI feedback in practice [26], [27]. Even though the instantaneous CSI can be perfectly available, dynamic JPAC in accordance with its variations would lead to excessively high computational and signaling costs. In this paper, we consider the chance (probabilistic or outage-based) SINR constrained JPAC problem, where each link's SINR outage probability must be kept below a given tolerance. Different from most of the aforementioned works on

JPAC where the perfect CSI is assumed, our new formulation only requires the availability of the channel distribution information (CDI). Due to the fact that the CDI can remain unchanged over a relatively long period of time, JPAC based on the CDI can therefore be performed on a relatively slow timescale (compared to fast fluctuations of instantaneous channel conditions), hence the overall computational cost and signaling overhead can be significantly reduced, which is particularly appealing from the network operator's perspective. Moreover, the chance SINR constrained JPAC formulation can maximize the number of *long-term* supported links by using minimum total transmission power, and at the same time guarantee that *short-term* SINR requirements are respected with high probability, which depends on the user-specified outage tolerance.

It is well-known that characterizing Quality-of-Service (QoS) constraints in terms of an outage probability can significantly improve practicality of the resource allocation algorithms including power control and beamforming design; see [27]–[34] and references therein. Therefore, the chance constrained programming methodology has been widely applied to wireless system designs in recent years. However, as far as we know, such methodology has not been used in the context of JPAC. This is largely due to the computational challenge of solving the chance SINR constrained JPAC problem. First, chance SINR constraints do not have closed-form expressions and are nonconvex in general. Second, even when the CSI is perfectly available, the JPAC problem is NP-hard to solve and to approximate [1], [4], [8].

This is the first work that formulates the chance SINR constrained JPAC problem and proposes efficient deflation approaches for solving it. The main contributions of this paper are twofold.

- *Novel Problem Formulation and Reformulation.* In this paper, we assume that only the CDI is available, which is different from most of the existing works on JPAC where the perfect CSI is assumed. We propose the first chance SINR constrained JPAC formulation, where each link's SINR outage probability is required to be less than or equal to a preselected tolerance. Furthermore, we approximate the chance SINR constraint via sampling [35], [36] and reformulate the sample approximation of the chance SINR constrained JPAC problem as a composite group sparse minimization problem.
- *Efficient Deflation Approaches.* We propose an efficient convex SOCP approximation (different from that in [1]) of the group sparse minimization reformulation. The solution of the SOCP approximation can be used to check the simultaneous supportability of all links

in the network and to guide an iterative link removal procedure (the deflation approach). Instead of relying on standard SOCP solvers to solve the derived SOCP, we exploit its special structure and custom-design an efficient algorithm for solving it. Note that the standard SOCP solvers cannot efficiently solve the SOCP approximation here because both the number of constraints and unknown variables in the SOCP approximation increase linearly with the number of samples, which is generally large in order to guarantee the approximation performance.

*Notation.* We denote the index set  $\{1, 2, \dots, K\}$  by  $\mathcal{K}$ . Lowercase boldface and uppercase boldface are used for vectors and matrices, respectively. For a given vector  $\mathbf{x}$ , the notation  $\max\{\mathbf{x}\}$ ,  $\min\{\mathbf{x}\}$ ,  $(\mathbf{x})_k$ , and  $\|\mathbf{x}\|_0$  stand for its maximum entry, its minimum entry, its  $k$ -th entry, and the indicator function of  $\mathbf{x}$  (i.e.,  $\|\mathbf{x}\|_0 = 0$  if  $\mathbf{x} = \mathbf{0}$  and  $\|\mathbf{x}\|_0 = 1$  otherwise), respectively. The expression  $\max\{\mathbf{x}_1, \mathbf{x}_2\}$  ( $\min\{\mathbf{x}_1, \mathbf{x}_2\}$ ) represents the component-wise maximum (minimum) of two vectors  $\mathbf{x}_1$  and  $\mathbf{x}_2$ . For any subset  $\mathcal{I} \subseteq \mathcal{K}$ ,  $\mathbf{A}_{\mathcal{I}}$  stands for the matrix formed by the rows of  $\mathbf{A}$  indexed by  $\mathcal{I}$ . We use  $(\mathbf{A}_1, \mathbf{A}_2)$  to denote the matrix formed by stacking matrices  $\mathbf{A}_1$  and  $\mathbf{A}_2$  by column and use  $(\mathbf{A}_1; \mathbf{A}_2)$  to denote the matrix formed by stacking  $\mathbf{A}_1$  and  $\mathbf{A}_2$  by row. Similar notation applies to stacking of vectors and scalars. Finally, we use  $\mathbf{e}$  to represent the vector with all components being one,  $\mathbf{I}$  the identity matrix, and  $\mathbf{E}_k$  the matrix with all entries being zero except its  $k$ -th column entries being one, respectively.

## II. REVIEW OF THE NLPD ALGORITHM

The algorithms developed for the chance SINR constrained JPAC problem in this paper are based on the NLPD algorithm [8] for the JPAC problem that assumes the perfect CSI. To streamline the presentation, we briefly review the NLPD algorithm in this section. The basic idea of the NLPD algorithm is to update the power and check whether all links can be simultaneously supported or not. If the answer is yes, then terminate the algorithm; else drop one link from the network and update the power again. The above process is repeated until all the remaining links can be simultaneously supported.

Specifically, consider a  $K$ -link (a link corresponds to a transmitter-receiver pair) single-input single-output interference channel with channel gains  $g_{k,j} \geq 0$  (from transmitter  $j$  to receiver  $k$ ), noise power  $\eta_k > 0$ , SINR target  $\gamma_k > 0$ , and power budget  $\bar{p}_k > 0$  for  $k, j \in \mathcal{K}$ . Denote the power allocation vector by  $\mathbf{p} = (p_1, p_2, \dots, p_K)^T$  and the power budget vector by  $\bar{\mathbf{p}} = (\bar{p}_1, \bar{p}_2, \dots, \bar{p}_K)^T$ .

Treating interference as noise, we can write the SINR at the  $k$ -th receiver as

$$\text{SINR}_k(\mathbf{p}) = \frac{g_{k,k}p_k}{\eta_k + \sum_{j \neq k} g_{k,j}p_j}, \quad \forall k \in \mathcal{K}. \quad (1)$$

Correspondingly, we introduce an equivalent normalized channel. In particular, we use

$$\mathbf{q} = (q_1, q_2, \dots, q_K)^T \quad (2)$$

with  $q_k = p_k/\bar{p}_k$  to denote the normalized power allocation vector, and use  $\mathbf{c} = (c_1, c_2, \dots, c_K)^T$  with  $c_k = (\gamma_k \eta_k)/(g_{k,k} \bar{p}_k) > 0$  to denote the normalized noise vector. We denote the normalized channel matrix by  $\mathbf{A} \in \mathbb{R}^{K \times K}$  with its  $(k, j)$ -th entry

$$a_{k,j} = \begin{cases} 1, & \text{if } k = j; \\ -\frac{\gamma_k g_{k,j} \bar{p}_j}{g_{k,k} \bar{p}_k}, & \text{if } k \neq j. \end{cases}$$

With these notation, it is simple to check that  $\text{SINR}_k(\mathbf{p}) \geq \gamma_k$  if and only if  $(\mathbf{A}\mathbf{q} - \mathbf{c})_k \geq 0$ .

Based on the Balancing Lemma [5], we reformulate the JPAC problem as a sparse optimization problem

$$\begin{aligned} \min_{\mathbf{q}} \quad & \sum_{k \in \mathcal{K}} \|(\mathbf{c} - \mathbf{A}\mathbf{q})_k\|_0 + \alpha \bar{\mathbf{p}}^T \mathbf{q} \\ \text{s.t.} \quad & \mathbf{0} \leq \mathbf{q} \leq \mathbf{e}. \end{aligned} \quad (3)$$

In the above,  $\alpha$  is a parameter and  $\mathbf{e}$  is the all-one vector of length  $K$ . For details on the choice of the parameter  $\alpha$ , we refer the readers to [8, Section III-B]. Since problem (3) is NP-hard [1], we further consider its  $\ell_1$ -convex approximation (which is equivalent to an LP; see [8])

$$\begin{aligned} \min_{\mathbf{q}} \quad & \sum_{k \in \mathcal{K}} |(\mathbf{c} - \mathbf{A}\mathbf{q})_k| + \alpha \bar{\mathbf{p}}^T \mathbf{q} \\ \text{s.t.} \quad & \mathbf{0} \leq \mathbf{q} \leq \mathbf{e}. \end{aligned} \quad (4)$$

By solving (4), we know whether all links in the network can be simultaneously supported or not. If not, we drop one link (mathematically, delete the corresponding row and column of  $\mathbf{A}$  and the corresponding entry of  $\bar{\mathbf{p}}$  and  $\mathbf{c}$ ) from the network according to some removal strategy, and solve a reduced problem (4) until all the remaining links are supported.

### III. PROBLEM FORMULATION

Consider the chance SINR constrained JPAC problem, where the channel gains  $\{g_{k,j}\}$  in the SINR expression (1) are random variables. In this paper, we assume the distribution of  $\{g_{k,j}\}$  is known. However, we do not assume any specific channel distribution, which is different from most of the existing works on outage probability constrained resource allocation for wireless systems [29], [30], [32]. We also assume that all coordinations and computations are carried out by a central controller who knows the CDI of all links. Since  $\{g_{k,j}\}$  in (1) are random variables, we need to redefine the concept of a supported link. We call link  $k$  is supported if its outage probability is below a specified tolerance  $\epsilon \in (0, 1)$ , i.e.,

$$\mathbb{P}(\text{SINR}_k(\mathbf{p}) \geq \gamma_k) \geq 1 - \epsilon, \quad (5)$$

where the probability is taken with respect to the random variables  $\{g_{k,j}\}$ .

The chance SINR constrained JPAC problem aims to maximize the number of supported links while using minimum total transmission power. Mathematically, the problem can be formulated as

$$\begin{aligned} \max_{\mathbf{p}, \mathcal{S}} \quad & |\mathcal{S}| - \alpha \mathbf{e}^T \mathbf{p} \\ \text{s.t.} \quad & \mathbb{P}(\text{SINR}_k(\mathbf{p}) \geq \gamma_k) \geq 1 - \epsilon, \quad k \in \mathcal{S} \subseteq \mathcal{K}, \\ & \mathbf{0} \leq \mathbf{p} \leq \bar{\mathbf{p}}. \end{aligned} \quad (6)$$

In the above,  $\mathcal{S}$  denotes the set of supported links and  $|\mathcal{S}|$  denotes its cardinality, i.e., the number of supported links; the parameter  $\alpha$  balances the relative importance of the two goals, i.e., maximizing the number of supported links (the first term  $|\mathcal{S}|$  in the objective) and minimizing the total transmission power (the second term  $\mathbf{e}^T \mathbf{p}$  in the objective).

To gain further understanding of formulation (6), we compare it with the following two-stage formulation. Specifically, the first stage maximizes the number of admitted links:

$$\begin{aligned} \max_{\mathbf{p}, \mathcal{S}} \quad & |\mathcal{S}| \\ \text{s.t.} \quad & \mathbb{P}(\text{SINR}_k(\mathbf{p}) \geq \gamma_k) \geq 1 - \epsilon, \quad k \in \mathcal{S} \subseteq \mathcal{K}, \\ & \mathbf{0} \leq \mathbf{p} \leq \bar{\mathbf{p}}. \end{aligned} \quad (7)$$

We use  $\mathcal{S}_0$  to denote the optimal solution for problem (7) and call it *the maximum admissible set*. Notice that the solution for (7) might not be unique. The second stage minimizes the total



transmission power required to support the admitted links:

$$\begin{aligned} \min_{\mathbf{p}} \quad & \mathbf{e}^T \mathbf{p} \\ \text{s.t.} \quad & \mathbb{P}(\text{SINR}_k(\mathbf{p}) \geq \gamma_k) \geq 1 - \epsilon, \quad k \in \mathcal{S}_0, \\ & \mathbf{0} \leq \mathbf{p} \leq \bar{\mathbf{p}}. \end{aligned} \tag{8}$$

Due to the choice of  $\mathcal{S}_0$ , power control problem (8) is feasible.

Although the above two-stage formulation (i.e., (7) and (8)) is intuitive and easy to understand, the formulation (6) is better in terms of modeling the JPAC problem; see the following Theorem 1. Theorem 1 can be shown by a similar argument as used in [8, Theorem 1] and a detailed proof is provided in Section I of [37].

*Theorem 1:* Suppose the parameter  $\alpha$  satisfies

$$0 < \alpha < \alpha_1 := 1/\mathbf{e}^T \bar{\mathbf{p}}. \tag{9}$$

Then the optimal value of problem (7) is  $M$  if and only if the optimal value of problem (6) lies in  $(M - 1, M)$ . Moreover, suppose  $(\mathcal{S}^*, \mathbf{p}^*)$  is the solution of problem (6). Then,  $\mathcal{S}^*$  is a maximum admissible set and  $\mathbf{e}^T \mathbf{p}^*$  is the minimum total transmission power to support any maximum admissible set.

Theorem 1 states that the single-stage formulation (6) with  $\alpha \in (0, \alpha_1)$  is equivalent to the two-stage formulation (7) and (8) in terms of finding the maximum admissible set. Moreover, it is capable of picking the maximum admissible set with minimum total transmission power from potentially multiple maximum admissible sets.

In the rest of this paper, we develop sample approximation-based deflation approaches for (approximately) solving the chance SINR constrained JPAC problem (6).

#### IV. SAMPLE APPROXIMATION AND REFORMULATION

In general, the chance SINR constrained optimization problem (6) is difficult to solve exactly, since it is difficult to obtain the closed-form expression of (5). In this section, we first approximate the computationally intractable chance SINR constraint via sampling, and then reformulate the sample approximation of problem (6) as a composite group sparse optimization problem. Three distinctive advantages of the sample approximation scheme in the context of approximating the chance SINR constraint (5) are as follows. First, it works for general channel distribution models and thus is distributionally robust. Second, the sample approximation technique significantly



simplifies problem (6) by replacing the difficult chance SINR constraint with finitely many simple linear constraints (depending on the sample size). Last but not the least, solving the sample approximation problem returns a solution to the original chance constrained problem with guaranteed performance [35], [36].

It is worthwhile remarking that safe tractable approximation [38], [39] is an alternative approach to the sample approximation approach to dealing with the chance constraint. The safe tractable approximation approach builds an analytic upper bound of the probability for the chance constraint to be violated. The advantage of this line of approach over the sample approximation approach is that solving the deterministic analytic upper bound will return a feasible solution to the chance constraint for sure. However, to build such an analytic upper bound, some strict conditions on structures of the function composed in the chance constraint and on the distribution of the random variables are required.

#### A. Sample Approximation

We handle the chance SINR constraint via sample approximations [35], [40]. Suppose  $\{g_{k,j}^n\}_{n=1}^N$  are  $N$  independent samples drawn according to the distribution of  $\{g_{k,j}\}$  by the central controller, we use

$$\text{SINR}_k^n(\mathbf{p}) := \frac{g_{k,k}^n p_k}{\eta_k + \sum_{j \neq k} g_{k,j}^n p_j} \geq \gamma_k, \quad n \in \mathcal{N} := \{1, 2, \dots, N\} \quad (10)$$

to approximate the chance SINR constraint (5). Since the samples are random variables, the power allocation vector  $\mathbf{p}$  satisfying the sampled SINR constraints (10) is also a random variable. Intuitively, if the sample size  $N$  is sufficiently large, then the power allocation vector  $\mathbf{p}$  satisfying (10) will satisfy the chance SINR constraint (5) with high probability.

The above intuition has been rigorously shown in [36, Theorem 1] and [28, Theorem 1]. It is shown that, if the sample size  $N$  satisfies

$$N \geq N^* := \left\lceil \frac{1}{\epsilon} \left( K - 1 + \ln \frac{1}{\delta} + \sqrt{2(K-1) \ln \frac{1}{\delta} + \ln^2 \frac{1}{\delta}} \right) \right\rceil \quad (11)$$

for any  $\delta \in (0, 1)$ , then any solution to the linear system

$$\text{SINR}_k^n(\mathbf{p}) \geq \gamma_k, \quad k \in \mathcal{K}, \quad n \in \mathcal{N} \quad (12)$$

will satisfy the chance SINR constraint (5) for all  $k \in \mathcal{K}$  with probability at least  $1 - \delta$ . In particular, if  $\delta$  is chosen to be a very small value, any solution to (12) will almost surely be feasible for the chance SINR constraints (5) for all  $k \in \mathcal{K}$ . Note that the number of samples needed will not increase significantly as  $\delta$  decreases, since  $N^*$  has only a logarithmic dependence on  $1/\delta$ . Although the dependence of  $N^*$  on  $\epsilon$  is  $N^* = O(1/\epsilon)$ , really small values of  $\epsilon$  are of no interest in the scenario considered in this paper.

The linear system (12) might have multiple solutions. The most interesting solution is the one that minimizes the total transmission power, i.e., the solution to the following problem

$$\begin{aligned} \min_{\mathbf{p}} \quad & \mathbf{e}^T \mathbf{p} \\ \text{s.t.} \quad & \text{SINR}_k^n(\mathbf{p}) \geq \gamma_k, \quad k \in \mathcal{K}, \quad n \in \mathcal{N}, \\ & \mathbf{0} \leq \mathbf{p} \leq \bar{\mathbf{p}}. \end{aligned} \tag{13}$$

Suppose  $\mathbf{p}$  is the solution to problem (13). Then, for each  $k \in \mathcal{K}$ , there must exist an index  $n_k \in \mathcal{N}$  such that

$$\text{SINR}_k^{n_k}(\mathbf{p}) = \gamma_k. \tag{14}$$

For simplicity, we will refer link  $k$  to be supported if all constraints in (10) are satisfied in the sequel.

### B. Sampled Channel Normalization

To facilitate the reformulation of the sample approximation of problem (6) and the development of efficient algorithms, we normalize the sampled channel parameters. To this end, we use

$$\mathbf{c}_k = \left( \frac{\gamma_k \eta_k}{g_{k,k}^1 \bar{p}_k}, \frac{\gamma_k \eta_k}{g_{k,k}^2 \bar{p}_k}, \dots, \frac{\gamma_k \eta_k}{g_{k,k}^N \bar{p}_k} \right)^T \in \mathbb{R}^{N \times 1}$$

to denote the normalized noise vector of link  $k$ . Define

$$a_{k,j}^n = \begin{cases} 1, & \text{if } k = j; \\ -\frac{\gamma_k g_{k,j}^n \bar{p}_j}{g_{k,k}^n \bar{p}_k}, & \text{if } k \neq j, \end{cases}$$

$$\mathbf{a}_k^n = (a_{k,1}^n, a_{k,2}^n, \dots, a_{k,K}^n) \in \mathbb{R}^{1 \times K}, \quad n \in \mathcal{N}, \quad k \in \mathcal{K},$$

and

$$\mathbf{A}_k = (\mathbf{a}_k^1; \mathbf{a}_k^2; \dots; \mathbf{a}_k^N) \in \mathbb{R}^{N \times K}, \quad k \in \mathcal{K}.$$

Notice that the entries of the  $k$ -th column of  $\mathbf{A}_k$  are one, and all the other entries are nonpositive. This special structure of  $\mathbf{A}_k$  ( $k \in \mathcal{K}$ ) will play an important role in the following algorithm design. Furthermore, we let

$$\mathbf{c} = (\mathbf{c}_1; \mathbf{c}_2; \dots; \mathbf{c}_K) \in \mathbb{R}^{NK \times 1} \text{ and } \mathbf{A} = (\mathbf{A}_1; \mathbf{A}_2; \dots; \mathbf{A}_K) \in \mathbb{R}^{NK \times K}.$$

With the above notation and (2), we can see that  $\text{SINR}_k^n(\mathbf{p}) \geq \gamma_k$  for all  $n \in \mathcal{N}$  if and only if  $\mathbf{A}_k \mathbf{q} \geq \mathbf{c}_k$ . Consequently, the sample approximation of problem (6) can be equivalently rewritten as

$$\begin{aligned} \max_{\mathbf{q}, \mathcal{S}} \quad & |\mathcal{S}| - \alpha \bar{\mathbf{p}}^T \mathbf{q} \\ \text{s.t.} \quad & \mathbf{A}_k \mathbf{q} - \mathbf{c}_k \geq 0, \quad k \in \mathcal{S} \subseteq \mathcal{K}, \\ & \mathbf{0} \leq \mathbf{q} \leq \mathbf{e}. \end{aligned} \tag{15}$$

### C. Composite Group Sparse Minimization Reformulation

By the definition of  $\|\cdot\|_0$ , the sampled JPAC problem (15) can be reformulated as the following composite group sparse optimization problem

$$\begin{aligned} \min_{\mathbf{q}} \quad & \sum_{k \in \mathcal{K}} \|\max\{\mathbf{c}_k - \mathbf{A}_k \mathbf{q}, \mathbf{0}\}\|_0 + \alpha \bar{\mathbf{p}}^T \mathbf{q} \\ \text{s.t.} \quad & \mathbf{0} \leq \mathbf{q} \leq \mathbf{e}. \end{aligned} \tag{16}$$

Problem (16) has the following property stated in Proposition 1, which is mainly due to the special structure of  $\mathbf{A}_k$ . The proof of Proposition 1 can be found in Appendix A.

*Proposition 1:* Suppose that  $\mathbf{q}^*$  is the solution to problem (16) and link  $k$  is supported at the point  $\mathbf{q}^*$  (i.e.,  $\mathbf{A}_k \mathbf{q}^* \geq \mathbf{c}_k$ ). Then there must exist an index  $n_k \in \mathcal{N}$  such that  $(\mathbf{c}_k - \mathbf{A}_k \mathbf{q}^*)_{n_k} = 0$ .

Proposition 1 implies that problem (16) can be viewed as an (nontrivial) extension of problem (3). In fact, we know from Proposition 1 that when  $N = 1$ , the solution of problem (16) satisfies  $\mathbf{c} - \mathbf{A} \mathbf{q}^* \geq \mathbf{0}$  (i.e.,  $(\mathbf{c} - \mathbf{A} \mathbf{q}^*)_k = 0$  for supported links and  $(\mathbf{c} - \mathbf{A} \mathbf{q}^*)_k > 0$  for unsupported links), and problem (16) reduces to problem (3). Since problem (3) is NP-hard to solve to global optimality and NP-hard to approximate to constant factor of global optimality [1], [4], [8], it follows that problem (16) is also NP-hard to solve and approximate.

A key difference between problems (16) and (3) lies in the max operator introduced in problem (16). In problem (3), if link  $k$  is supported, then  $\mathbf{c}_k - \mathbf{A}_k \mathbf{q}^*$  is a scalar and equals zero; while in problem (16), if link  $k$  is supported, then  $\mathbf{c}_k - \mathbf{A}_k \mathbf{q}^* \leq \mathbf{0}$  but not necessarily equal to zero.

Therefore, to correctly formulate the JPAC problem, we introduce a  $\max$  operation and put  $\max\{\mathbf{c}_k - \mathbf{A}_k \mathbf{q}, \mathbf{0}\}$  in  $\|\cdot\|_0$  instead of  $\mathbf{c}_k - \mathbf{A}_k \mathbf{q}$ ; see problem (16). Notice that in the sparse formulation it is desirable that a link is supported if and only if the corresponding  $\ell_0$ -quasi-norm is zero. To further illustrate this, we give the following example, where  $K = N = 2$ , and

$$\mathbf{A} = (\mathbf{A}_1; \mathbf{A}_2) = \begin{pmatrix} 1 & -0.2 \\ 1 & -0.5 \\ -0.3 & 1 \\ -0.5 & 1 \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} 0.5 \\ 0.5 \\ 0.5 \\ 0.5 \end{pmatrix}.$$

It can be checked that the only possible way to simultaneously support the two links  $\{1, 2\}$  is  $\mathbf{q}^* = \mathbf{e}$  and  $\max\{\mathbf{c}_1 - \mathbf{A}_1 \mathbf{q}^*, \mathbf{0}\} = \max\{\mathbf{c}_2 - \mathbf{A}_2 \mathbf{q}^*, \mathbf{0}\} = \mathbf{0}$  but  $\mathbf{c}_1 - \mathbf{A}_1 \mathbf{q}^* \neq \mathbf{0}$  and  $\mathbf{c}_2 - \mathbf{A}_2 \mathbf{q}^* \neq \mathbf{0}$ .

## V. EFFICIENT DEFLATION APPROACHES FOR THE SAMPLED JPAC PROBLEM

In this section, we develop efficient convex approximation-based deflation algorithms for solving the sampled JPAC problem (16). As can be seen, problem (16) has a discontinuous objective function due to the first term. However, it allows for an efficient convex approximation. We first approximate problem (16) by a convex problem, which is actually equivalent to an SOCP, and then design efficient algorithms for solving the approximation problem. The solution to the approximation problem can be used to check the simultaneous supportability of all links in the network and to guide an iterative link removal procedure (the deflation approach). We conclude this section with two convex approximation-based deflation algorithms for solving the sampled joint control problem (16).

### A. Convex Approximation

Recall that problem (16) aims to find a feasible  $\mathbf{q}$  such that the vector  $\mathbf{x} = (\mathbf{x}_1; \mathbf{x}_2; \dots; \mathbf{x}_K)$  is as sparse as possible in the group sense, where  $\mathbf{x}_k := \max\{\mathbf{c}_k - \mathbf{A}_k \mathbf{q}, \mathbf{0}\}$ . The nonsmooth mixed  $\ell_2/\ell_1$  norm,  $\sum_{k \in \mathcal{K}} \|\mathbf{x}_k\|_2$ , is shown in [41] to be quite effective in characterizing and inducing the *group sparsity* of the vector  $\mathbf{x}$ . To understand this, observe that  $\sum_{k \in \mathcal{K}} \|\mathbf{x}_k\|_2$ , the  $\ell_1$  norm of the vector  $(\|\mathbf{x}_1\|_2, \|\mathbf{x}_2\|_2, \dots, \|\mathbf{x}_K\|_2)^T$ , is a good approximation of its  $\ell_0$  norm, which is equal to the  $\ell_0$  norm of the vector  $(\|\mathbf{x}_1\|_0, \|\mathbf{x}_2\|_0, \dots, \|\mathbf{x}_K\|_0)^T$ . More discussions on using the mixed  $\ell_2/\ell_1$  norm to recover the group sparsity can be found in [41].

Motivated by the above discussion and the NP-hardness of problem (16), we consider the convex approximation of problem (16) as follows:

$$\begin{aligned} \min_{\mathbf{q}} \quad & f(\mathbf{q}) := \sum_{k \in \mathcal{K}} \|\max\{\mathbf{c}_k - \mathbf{A}_k \mathbf{q}, \mathbf{0}\}\|_2 + \alpha \bar{\mathbf{p}}^T \mathbf{q} \\ \text{s.t.} \quad & \mathbf{0} \leq \mathbf{q} \leq \mathbf{e}. \end{aligned} \quad (17)$$

The convexity of the objective function of problem (17) follows directly from [42, Section 3.2.5].

Compared to problem (16), the objective function of problem (17) is continuous in  $\mathbf{q}$ , but still nonsmooth. We give the subdifferential [43] of the function  $\|\max\{\mathbf{c}_k - \mathbf{A}_k \mathbf{q}, \mathbf{0}\}\|_2$  in Proposition 2, which is important in the following analysis and algorithm design. The proof of Proposition 2 is provided in Appendix B.

*Proposition 2:* Define  $h_k(\mathbf{q}) = \|\max\{\mathbf{c}_k - \mathbf{A}_k \mathbf{q}, \mathbf{0}\}\|_2$ . Suppose  $\mathbf{c}_k - \mathbf{A}_k \bar{\mathbf{q}} \leq \mathbf{0}$  and  $\mathcal{N}_k^- := \{n \mid (\mathbf{c}_k - \mathbf{A}_k \bar{\mathbf{q}})_n = 0\} \neq \emptyset$ , then

$$\partial h_k(\bar{\mathbf{q}}) = \left\{ - \sum_{n \in \mathcal{N}_k^-} s_n (\mathbf{a}_k^n)^T \mid s_n \geq 0, \sum_{n \in \mathcal{N}_k^-} s_n^2 \leq 1 \right\}.$$

In particular, if  $\mathcal{N}_k^- = \mathcal{N}$ , then  $\partial h_k(\bar{\mathbf{q}}) = \{-\mathbf{A}_k^T \mathbf{s} \mid \mathbf{s} \geq \mathbf{0}, \|\mathbf{s}\|_2 \leq 1\}$ . Further, if  $\mathcal{N}_k^+ := \{n \mid (\mathbf{c}_k - \mathbf{A}_k \bar{\mathbf{q}})_n > 0\} \neq \emptyset$ , then

$$\nabla h_k(\bar{\mathbf{q}}) = \frac{-\sum_{n \in \mathcal{N}_k^+} (\mathbf{c}_k - \mathbf{A}_k \bar{\mathbf{q}})_n (\mathbf{a}_k^n)^T}{\|\max\{\mathbf{c}_k - \mathbf{A}_k \bar{\mathbf{q}}, \mathbf{0}\}\|_2} = \frac{-\mathbf{A}_k^T \max\{\mathbf{c}_k - \mathbf{A}_k \bar{\mathbf{q}}, \mathbf{0}\}}{\|\max\{\mathbf{c}_k - \mathbf{A}_k \bar{\mathbf{q}}, \mathbf{0}\}\|_2}. \quad (18)$$

We now discuss the choice of the parameter  $\alpha$  in (17). The parameter  $\alpha$  in (17) should be chosen appropriately such that the following “Never-Over-Removal” property is satisfied: the solution of problem (17) should simultaneously support all links at their desired SINR targets with minimum total transmission power as long as all links in the network are simultaneously supportable. Otherwise, since the solution of (17) will be used to check the simultaneous supportability of all links and to guide the links’ removal, it may mislead us to remove the links unnecessarily. Notice that problem (17) with  $\alpha = 0$  indeed can simultaneously support all links as long as the links are simultaneously supportable but not necessarily with minimum total transmission power, i.e., the solution of problem (17) with  $\alpha = 0$  might not solve (13). Theorem 2 gives an interval of the parameter  $\alpha$  to guarantee the “Never-Over-Removal” property. The proof of Theorem 2 (see Appendix C) is mainly based on Proposition 2.

*Theorem 2:* Suppose there exists some vector  $\mathbf{q}$  such that  $\mathbf{0} \leq \mathbf{q} \leq \mathbf{e}$  and  $\mathbf{A}\mathbf{q} \geq \mathbf{c}$ . Then any solution of problem (17) with

$$0 < \alpha \leq \alpha_2 := \frac{\min \{\mathbf{c}\}}{K \max \{\bar{\mathbf{p}}\}} \quad (19)$$

can simultaneously support all links at their desired SINR targets with minimum total transmission power.

Combining (9) and (19), we propose to choose the parameter  $\alpha$  in (17) according to

$$\alpha = \min \{c_1\alpha_1, c_2\alpha_2\}, \quad (20)$$

where  $c_1, c_2 \in (0, 1)$  are two constants.

*Link Removal Strategy.* The solution of problem (17) can be used to guide the link removal process. In particular, by solving (17) with  $\alpha$  given in (20), we know whether all links in the network can be simultaneously supported by simply checking if its solution  $\bar{\mathbf{q}}$  satisfies  $\mathbf{A}\bar{\mathbf{q}} \geq \mathbf{c}$ . Furthermore, if all links in the network cannot be simultaneously supported, we need to remove at least one link from the network. In particular, picking the worst sampled channel index  $\bar{n}_k = \arg \max \{c_k - \mathbf{A}_k \bar{\mathbf{q}}\}$ , we remove the link with the largest interference plus noise footprint

$$k = \arg \max \left\{ \sum_{j \neq k} |a_{k,j}^{\bar{n}_k}| \bar{q}_j + \sum_{j \neq k} |a_{j,k}^{\bar{n}_j}| \bar{q}_k + c_k^{\bar{n}_k} \right\}. \quad (21)$$

In the next subsection, we design efficient algorithms to solve the convex but nonsmooth problem (17).

### B. Solution for Approximation Problem (17)

By introducing auxiliary variables  $\mathbf{x} = (\mathbf{x}_1; \mathbf{x}_2; \dots; \mathbf{x}_K)$  and  $\mathbf{t} = (t_1; t_2; \dots; t_K)$ , problem (17) can be transformed into the following SOCP

$$\begin{aligned} \min_{\mathbf{q}, \mathbf{x}, \mathbf{t}} \quad & \sum_{k \in \mathcal{K}} t_k + \alpha \bar{\mathbf{p}}^T \mathbf{q} \\ \text{s.t.} \quad & \|\mathbf{x}_k\| \leq t_k, \quad k \in \mathcal{K}, \\ & \mathbf{c} - \mathbf{A}\mathbf{q} \leq \mathbf{x}, \\ & \mathbf{0} \leq \mathbf{x}, \\ & \mathbf{0} \leq \mathbf{q} \leq \mathbf{e}, \end{aligned} \quad (22)$$

which can be solved by using the standard solver like CVX [44]. However, it is not an efficient way of solving problem (17) by solving its equivalent SOCP reformulation (22), since both

the number of constraints and the number of unknown variables of problem (22) are of order  $O(NK)$  while (11) suggests that the sample size  $N$  is generally very large.

Next, we develop a custom-design algorithm for problem (17) by first smoothing the problem and then applying the efficient projected alternate Barzilai-Borwein (PABB) algorithm [45], [46] to solve its smooth counterpart. More specifically, we smooth problem (17) by

$$\begin{aligned} \min_{\mathbf{q}} \quad & \tilde{f}(\mathbf{q}, \mu) = \sum_{k \in \mathcal{K}} \sqrt{\|\max\{\mathbf{c}_k - \mathbf{A}_k \mathbf{q}, \mathbf{0}\}\|_2^2 + \mu^2} + \alpha \bar{\mathbf{p}}^T \mathbf{q} \\ \text{s.t.} \quad & \mathbf{0} \leq \mathbf{q} \leq \mathbf{e}, \end{aligned} \quad (23)$$

where  $\mu > 0$  is the smoothing parameter. By (18) in Proposition 2, the objective function  $\tilde{f}(\mathbf{q}, \mu)$  of problem (23) is differentiable everywhere with respect to  $\mathbf{q}$  and its gradient is given by

$$\nabla \tilde{f}(\mathbf{q}, \mu) = \sum_{k \in \mathcal{K}} \frac{-\mathbf{A}_k^T \max\{\mathbf{c}_k - \mathbf{A}_k \mathbf{q}, \mathbf{0}\}}{\sqrt{\|\max\{\mathbf{c}_k - \mathbf{A}_k \mathbf{q}, \mathbf{0}\}\|_2^2 + \mu^2}} + \alpha \bar{\mathbf{p}}.$$

It can be shown that, as the parameter  $\mu$  tends to zero,  $\tilde{f}(\mathbf{q}, \mu)$  *uniformly* converges to  $f(\mathbf{q})$  in (17) and the solution of the smoothing problem (23) also converges to the one of problem (17); see Section II of [37]. Therefore, when the parameter  $\mu$  is very close to zero, the solution of problem (23) will be very close to the one of problem (17).

We apply the PABB algorithm [45], [46] to solve the smoothing problem (23). Three distinctive advantages of the PABB algorithm in the context of solving problem (23) are as follows. First, the box constraint is easy to project onto, and thus the PABB algorithm can be easily implemented to solve problem (23). Second, the PABB algorithm requires only the gradient information but not the high-order derivative information, which makes it suitable for solving large-scale optimization problem (23). Last but not least, the PABB algorithm enjoys a quite good numerical performance due to the use of the BB stepsize [46]. When using the PABB algorithm to solve problem (23), we employ the *continuation* technique [47], [48]. That is, to obtain an approximate solution of (17), we solve (23) with a series of gradually decreasing values for  $\mu$ , instead of using a tiny fixed  $\mu$ . The continuation technique can reasonably improve the computational efficiency. Solving problem (17) by the PABB algorithm (combined with the smoothing and continuation techniques) is much faster than solving its SOCP reformulation (22) by the standard SOCP solver. Simulation results will be given later in Section VI.



### C. Convex Approximation-Based Deflation Algorithms

The basic idea of the proposed convex approximation-based deflation algorithm for the sampled JPAC problem (16) is to solve the power control problem (17) and check whether all links can be supported or not; if not, remove a link from the network, and solve a reduced problem (17) again until all the remaining links are supported.

As in [8], to accelerate the deflation procedure (avoid solving too many optimization problems in the form of (17)), we can derive an easy-to-check necessary condition for all links in the network to be simultaneously supported. It is easy to verify that the condition

$$\boldsymbol{\mu}_+^T \mathbf{e} - (\boldsymbol{\mu}_-^T \mathbf{c}^{\max} + \mathbf{e}^T \mathbf{c}) \geq 0 \quad (24)$$

is necessary for all links to be simultaneously supported, where  $\boldsymbol{\mu}_+ = \max\{\boldsymbol{\mu}, \mathbf{0}\}$ ,  $\boldsymbol{\mu}_- = \max\{-\boldsymbol{\mu}, \mathbf{0}\}$ ,  $\boldsymbol{\mu} = \mathbf{A}^T \mathbf{e}$ , and  $\mathbf{c}^{\max} = (\max\{\mathbf{c}_1\}; \max\{\mathbf{c}_2\}; \dots; \max\{\mathbf{c}_K\})$ . If (24) is violated, we remove the link  $k_0$  according to

$$k_0 = \arg \max_{k \in \mathcal{K}} \left\{ \sum_{j \neq k} |\bar{a}_{k,j}| + \sum_{j \neq k} |\bar{a}_{j,k}| + \bar{c}_k \right\}, \quad (25)$$

which corresponds to applying the SMART rule [4] to the normalized sampled channel and substituting  $\mathbf{q} = \mathbf{e}$ . In (25),  $\bar{a}_{k,j}$  and  $\bar{c}_k$  are the averaged sample channel gain and noise, i.e.,

$$\bar{\mathbf{a}}_k = (\bar{a}_{k,1}, \bar{a}_{k,2}, \dots, \bar{a}_{k,K}) = \frac{\mathbf{e}^T \mathbf{A}_k}{N}, \quad \bar{c}_k = \frac{\mathbf{e}^T \mathbf{c}_k}{N}, \quad k \in \mathcal{K}.$$

The proposed convex approximation-based deflation algorithmic framework for problem (16) is described in Algorithm 1. It is worthwhile remarking the difference between the proposed Algorithm 1 and the NLPD algorithm in [8]. The first key difference is that Algorithm 1 is designed for solving the sample approximation of the chance SINR constrained JPAC problem (6) (i.e., problem (16)) while the NLPD algorithm is designed for solving the instantaneous SINR constrained JPAC problem (3). As discussed in Subsection IV-C, problem (16) includes problem (3) as a special case. The second key difference between the two algorithms lies in the power control step (i.e., **Step 3**). More specifically, at each iteration, the proposed Algorithm 1 solves the SOCP (17) to update the power while the NLPD algorithm solves the LP (4) to update the power. We also remark that the SOCP approximation (17) used in Algorithm 1 is

different from the one used in [1]. The two SOCP approximations take different forms and are derived from different perspectives.

**Algorithm 1: A Convex Approximation-Based Deflation Algorithmic Framework**

**Step 1.** Initialization: Input data  $(\mathbf{A}, \mathbf{c}, \bar{\mathbf{p}})$ .

**Step 2.** Preprocessing: Remove link  $k_0$  iteratively according to (25) until condition (24) holds true.

**Step 3.** Power control: Compute parameter  $\alpha$  by (20) and solve problem (17); check whether all links are supported: if yes, go to **Step 5**; else go to **Step 4**.

**Step 4.** Admission control: Remove link  $k_0$  according to (21), set  $\mathcal{K} = \mathcal{K} \setminus \{k_0\}$ , and go to **Step 3**.

**Step 5.** Postprocessing: Check the removed links for possible admission.

In the above framework, if the power control problem (17) is solved via solving its equivalent SOCP reformulation (22), we call the corresponding algorithm SOCP-D; while if problem (17) is solved via using the PABB algorithm to solve its smoothing counterpart (23), we call the corresponding algorithm PABB-D. The SOCP-D algorithm is of polynomial time complexity, i.e., it has a complexity of  $O(N^{3.5}K^{4.5})$ , since it needs to solve at most  $K$  SOCP problems (22) and solving one SOCP problem in the form of (22) requires  $O(N^{3.5}K^{3.5})$  operations [49, Page 423]. It is hard to analyze the complexity of the PABB-D algorithm. This is because global (linear) convergence rate of the PABB algorithm, when it is used to solve general nonlinear optimization problems, remains unknown [50]. The postprocessing step (Step 5) aims at admitting the links removed in the preprocessing and admission control steps [1], [8]. A specification of the postprocessing step can be found in Section III of [37].

## VI. NUMERICAL SIMULATIONS

To illustrate the effectiveness and efficiency of the two proposed convex approximation-based deflation algorithms (SOCP-D and PABB-D), we present some numerical simulation results in this section. The number of supported links, the total transmission power, and the execution CPU time are used as the metrics for comparing different algorithms.

*Simulation Setup:* As in [1], each transmitter's location obeys the uniform distribution over a  $D_1 \text{ Km} \times D_1 \text{ Km}$  square and the location of each receiver is uniformly generated in a disc with

center at its corresponding transmitter and radius  $D_2$  Km, excluding a radius of 10 m. Suppose that the channel coefficient  $h_{k,j}$  is generated from the Rician channel model [51], i.e.,

$$h_{k,j} = \left( \sqrt{\frac{\kappa}{\kappa+1}} + \sqrt{\frac{1}{\kappa+1}} \zeta_{k,j} \right) \frac{1}{d_{k,j}^2}, \quad \forall k, j \in \mathcal{K}, \quad (26)$$

where  $\zeta_{k,j}$  obeys the standard complex Gaussian distribution, i.e.,  $\zeta_{k,j} \sim \mathcal{CN}(0, 1)$ ,  $d_{k,j}$  is the Euclidean distance from the link of transmitter  $j$  to the link of receiver  $k$ , and  $\kappa$  is the ratio of the power in the line of sight (LOS) component to the power in the other (non-LOS) multipath components. For  $\kappa = 0$  we have Rayleigh fading and for  $\kappa = \infty$  we have no fading (i.e., a channel with no multipath and only a LOS component). The parameter  $\kappa$  therefore is a measure of the severity of the channel fading: a small  $\kappa$  implies severe fading and a large  $\kappa$  implies relatively mild fading. The channel gain  $\{g_{k,j}\}$  are set to be:

$$g_{k,j} = |h_{k,j}|^2 = \left| \sqrt{\frac{\kappa}{\kappa+1}} + \sqrt{\frac{1}{\kappa+1}} \zeta_{k,j} \right|^2 \frac{1}{d_{k,j}^4}, \quad k, j \in \mathcal{K}. \quad (27)$$

Each link's SINR target is set to be  $\gamma_k = 2$  dB ( $\forall k \in \mathcal{K}$ ), each link's noise power is set to be  $\eta_k = -90$  dB ( $\forall k \in \mathcal{K}$ ), and the power budget of the link of transmitter  $k$  is set to be

$$\bar{p}_k = b \underline{p}_k, \quad k \in \mathcal{K}, \quad (28)$$

where  $\underline{p}_k$  is the minimum power needed by link  $k$  to meet its SINR requirement in the absence of any interference from other links when  $\kappa = +\infty$  in (26).

*Benchmark:* When  $\kappa = +\infty$ , there is no uncertainty of channel gains, and the number of supported links in this case should be greater than or equal to the number of supported links under the same channel conditions except where  $\kappa < +\infty$ . In addition, if the number of supported links under these two cases are equal to each other, the total transmission power in the former channel condition should be less than the one in the latter channel condition. In fact, when  $\kappa = +\infty$ , the corresponding JPAC problem (16) reduces to problem (3), which can be solved efficiently by the NLPD algorithm in [8]. The solution given by the NLPD algorithm will be used as the benchmark to compare with the two proposed algorithms<sup>1</sup>, since the NLPD algorithm was

<sup>1</sup>We remark that this is the first paper that addresses the JPAC problem based on the CDI assumption without specifying any particular distribution, and there is no existing algorithms dealing with the same issue that we can compare the proposed algorithms with.

reported to have the close-to-global-optimal performance in terms of the number of supported links in [8].

*Choice of Parameters:* We set the parameters  $\epsilon$ ,  $\delta$ , and  $K$  in (11) to be 0.1, 0.05, and 10, respectively. We remark that  $K$  in equation (11) is the number of supported links but not the number of total links. Substituting these parameters in (11), we obtain  $N^* = 200$ , and we set  $N = 200$  in all of our simulations. Both of the parameters  $c_1$  and  $c_2$  in (20) are set to be 0.999. We do simulations in two different setups where  $(D_1, D_2) = (2, 0.4)$  and  $(D_1, D_2) = (1, 0.2)$ . For convenience, we call the former setup as Setup1 and the latter one as Setup2. Notice that Setup2 represents a dense network where the distance between the transmitters and receivers are closer (i.e., half of that of the Setup1). Under each setup, we test three different sets of parameters, where one is  $(\kappa, b) = (+\infty, 2)$ , one is  $(\kappa, b) = (100, 4)$ , and another one is  $(\kappa, b) = (10, 40)$ . Finally, we use CVX [44] to solve the SOCP problems in the SOCP-D algorithm.

*Simulation Results and Analysis:* Table I summarizes the statistics of the number of supported links of 200 Monte-Carlo runs of numerical experiments with different choices of simulation parameters. For instance, “664 = 2 \* 19 + 3 \* 98 + 4 \* 83” in the third column of Table I stands for that when  $(K, D_1, D_2, \kappa, b) = (4, 2, 0.4, +\infty, 2)$ , total 664 links are supported in these 200 Monte-Carlo runs, and amongst them, 2 links are supported 19 times, 3 links are supported 98 times, and 4 links are supported 83 times. Figs. 1, 2, and 3 are obtained by averaging over the 200 Monte-Carlo runs. They plot the average number of supported links, the average total transmission power, and the average execution CPU time of the proposed SOCP-D and PABB-D algorithms (for solving the sampled JPAC problem (16)) and the benchmark versus different number of total links in Setup1.

It can be seen from Fig. 1 that the number of supported links by the two proposed algorithms (for fading channels) is less than the benchmark (for deterministic channels). This shows that the uncertainty of channel gains could lead to a (significant) reduction in the number of supported links. This can also be clearly observed from Table I. For instance, when  $K = 4$  (see the first three lines of Table I) all links can be simultaneously supported 83 times when  $\kappa = +\infty$ , 78 times when  $\kappa = 100$ , and only 50 times when  $\kappa = 10$ . In fact, this is the reason why we associate different  $\kappa$  with different  $b$  in our simulations. We expect that a large  $b$  and thus large power budgets  $\bar{p}_k$  (cf. (28)) can compensate the performance degradation of the number of supported links caused by the large uncertainty of channel gains.

TABLE I  
STATISTICS OF THE NUMBER OF SUPPORTED LINKS OF 200 MONTE-CARLO RUNS.

Parameters ( $K, D_1, D_2, \kappa, b$ )	Algorithm	Statistics of the Number of Supported Links
(4, 2, 0.4, $+\infty$ , 2)	Benchmark	664=2*19+3*98+4*83
(4, 2, 0.4, 100, 4)	SOCP-D/PABB-D	659=2*19+3*103+4*78
(4, 2, 0.4, 10, 40)	SOCP-D/PABB-D	609=1*1+2*39+3*110+4*50
(12, 2, 0.4, $+\infty$ , 2)	Benchmark	1468=5*8+6*36+7*68+8*60+9*24+10*4
(12, 2, 0.4, 100, 4)	SOCP-D/PABB-D	1431=5*12+6*42+7*72+8*54+9*17+10*3
(12, 2, 0.4, 10, 40)	SOCP-D/PABB-D	1236=4*10+5*39+6*79+7*53+8*15+9*4
(20, 2, 0.4, $+\infty$ , 2)	Benchmark	1953=6*1+7*8+8*21+9*50+10*67+11*38+12*11+13*3+14*1
(20, 2, 0.4, 100, 4)	SOCP-D/PABB-D	1864=7*13+8*40+9*58+10*58+11*22+12*8+13*1
(20, 2, 0.4, 10, 40)	SOCP-D/PABB-D	1570=5*2+6*17+7*65+8*60+9*39+10*15+11*2
(28, 2, 0.4, $+\infty$ , 2)	Benchmark	2342=8*1+9*13+10*23+11*54+12*56+13*31+14*17+15*2+16*1+17*2
(28, 2, 0.4, 100, 4)	SOCP-D/PABB-D	2250=8*3+9*21+10*40+11*50+12*50+13*22+14*11+15*1+16*2
(28, 2, 0.4, 10, 40)	SOCP-D/PABB-D	1875=5*1+6*1+7*13+8*37+9*60+10*48+11*25+12*14+13*0+14*1
(4, 1, 0.2, $+\infty$ , 2)	Benchmark	639=2*26+3*109+4*65
(4, 1, 0.2, 100, 4)	SOCP-D/PABB-D	632=1*1+2*27+3*111+4*61
(4, 1, 0.2, 10, 40)	SOCP-D/PABB-D	589=1*2+2*44+3*117+4*37
(12, 1, 0.2, $+\infty$ , 2)	Benchmark	1443=4*3+5*4+6*38+7*85+8*45+9*22+10*3
(12, 1, 0.2, 100, 4)	SOCP-D/PABB-D	1403=4*3+5*11+6*45+7*82+8*41+9*16+10*2
(12, 1, 0.2, 10, 40)	SOCP-D/PABB-D	1214=3*2+4*8+5*56+6*64+7*52+8*14+9*4
(20, 1, 0.2, $+\infty$ , 2)	Benchmark	1942=6*1+7*3+8*29+9*49+10*69+11*37+12*11+13*1
(20, 1, 0.2, 100, 4)	SOCP-D/PABB-D	1882=6*1+7*10+8*32+9*65+10*58+11*23+12*11
(20, 1, 0.2, 10, 40)	SOCP-D/PABB-D	1577=5*4+6*18+7*54+8*59+9*55+10*6+11*4
(28, 1, 0.2, $+\infty$ , 2)	Benchmark	2333=8*2+9*13+10*34+11*41+12*50+13*37+14*18+15*4+16*1
(28, 1, 0.2, 100, 4)	SOCP-D/PABB-D	2236=7*1+8*5+9*22+10*36+11*57+12*38+13*27+14*13+15*1
(28, 1, 0.2, 10, 40)	SOCP-D/PABB-D	1857=6*5+7*16+8*38+9*51+10*51+11*28+12*9+13*2

Table I, Fig. 1, and Fig. 2 show that the two proposed algorithms always return the same solution to the sampled JPAC problem (16), i.e., supporting same number of links with same total transmission power. However, Fig. 3 shows that the PABB-D algorithm substantially outperforms the SOCP-D algorithm in terms of the average CPU time. This is not surprising, since in each power control step (i.e., solving the convex approximation problem (17)), the custom-design algorithm is used to carry out power control in the PABB-D algorithm while a general-purpose solver CVX is used to update power in the SOCP-D algorithm. Note that both the number of

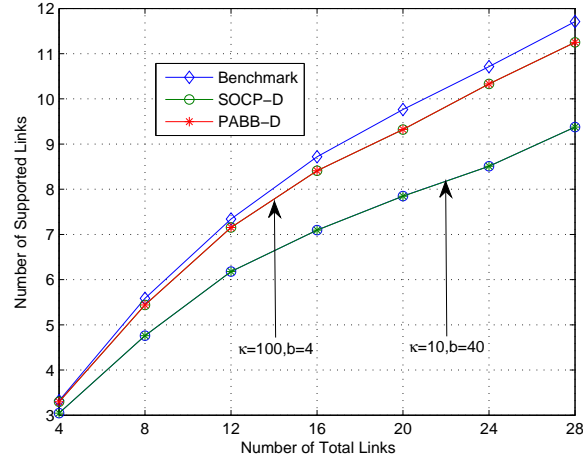


Fig. 1. Average number of supported links versus the number of total links in Setup1.

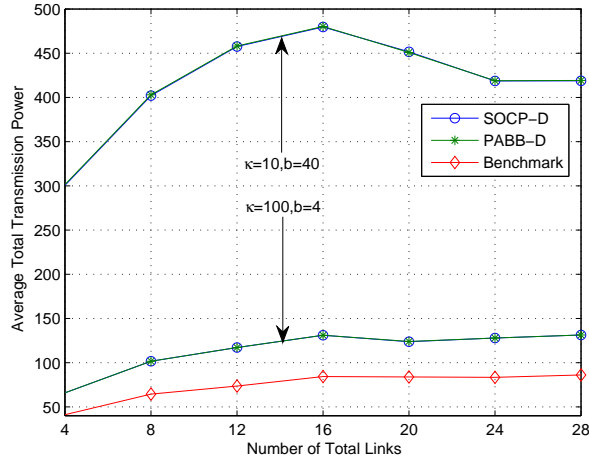


Fig. 2. Average total transmission power versus the number of total links in Setup1.

constraints and the number of unknown variables in problem (22) are of order  $O(NK)$  and (11) shows that the sample size  $N$  needs to be large to guarantee the approximation performance, which makes CVX unsuitable to be used to solve problem (17) via solving its equivalent SOCP reformulation (22).

By comparing the two sets of numerical experiments where  $(\kappa, b) = (100, 4)$  and  $(\kappa, b) = (10, 40)$ , it can be observed from Figs. 1 and 2 that more links can be supported with significantly less total transmission power in the former case than the latter case. This is because

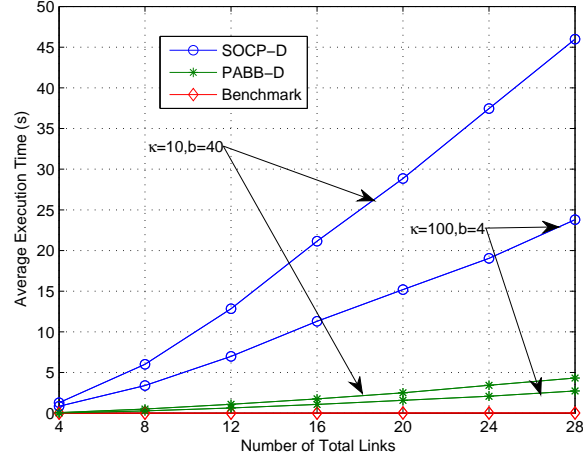


Fig. 3. Average execution time (in seconds) versus the number of total links in Setup1.

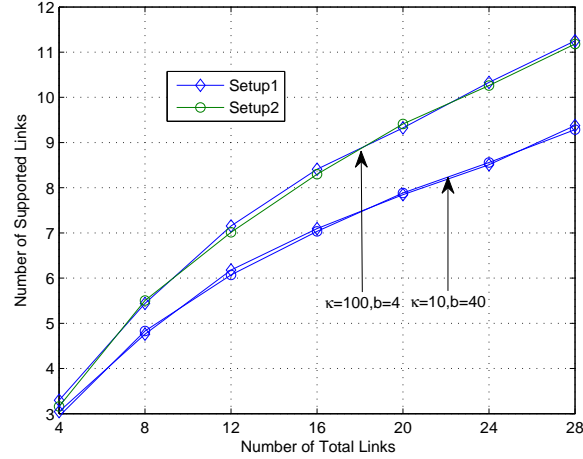


Fig. 4. Average number of supported links versus the number of total links in Setup1 and Setup2.

the uncertainty of the channel gains with  $\kappa = 100$  is generally much smaller than the one with  $\kappa = 10$ . We also point out that the execution CPU time of the two proposed deflation algorithms mainly depends on how many times the power control problem (17) is solved. In general, the larger number of links are supported, the smaller number of links are removed from the network and the smaller number of power control problems in form of (17) are solved. Therefore, the average CPU time of the proposed algorithms when  $(\kappa, b) = (10, 40)$  is larger than the one when  $(\kappa, b) = (100, 4)$ ; see Fig. 3.



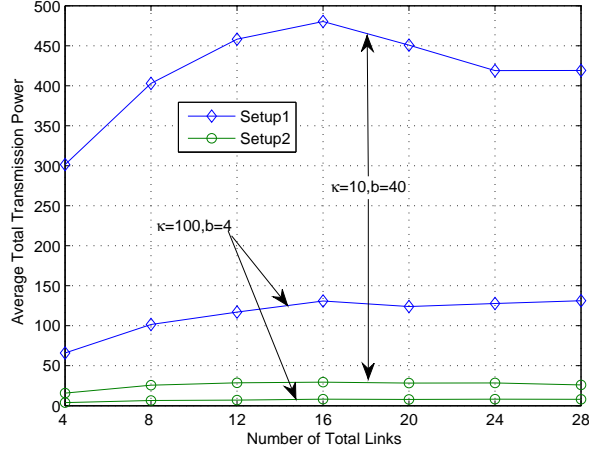


Fig. 5. Average total transmission power versus the number of total links in Setup1 and Setup2.

For conciseness, we do not present the comparison of the PABB-D and SOCP-D algorithms when both algorithms are used to solve the sampled JPAC problem (16) in Setup2, since the same observations as in Setup1 can be made in Setup2. Instead, we focus on the comparison of two different setups in the following. Since the PABB-D algorithm always returns the same solution as the SOCP-D algorithm but takes much less CPU time, we choose to use the PABB-D algorithm to solve the sampled JPAC problem (16) in the following.

TABLE II  
THE RATIO OF THE AVERAGE TOTAL TRANSMISSION POWER IN SETUP1 TO THAT IN SETUP2.

$(\kappa, b) \backslash K$	4	8	12	16	20	24	28
(100, 4)	17.1513	15.7554	16.5563	16.1834	15.9193	15.6577	16.4100
(10, 40)	19.1574	15.7159	16.0476	16.2672	15.9009	14.7517	16.0995

Figs. 4 and 5 report the average number of supported links and the average total transmission power (returned by the PABB-D algorithm for solving the sampled JPAC problem (16)) versus different number of total links in Setup1 and Setup2. It can be observed that the average number of supported links and the average execution time in both setups are roughly equal to each other, but the average total transmission power in Setup1 is approximately 16 times as large as that in

Setup2 (also see Table II). This is because when the setup is switched from Setup1 to Setup2 (with the random variables  $\zeta_{k,j}$  in (26) and (27) neglected for the time being), the corresponding distances  $d_{k,j}$  between the transmitters and receivers decrease by half. According to (27) and (28), all the channel gains (including both the direct-link and cross-link channel gains) increase and power budgets decrease by a factor of 16. As the channel gains are increased and power budgets are decreased by a same factor while the noise powers remain to be fixed, the number of supported links in problem (16) remains unchanged. However, it brings a benefit of a 93.75% (=15/16) reduction in the total transmission power, which is consistent with our engineering practice. Due to the effects of the random variables  $\zeta_{k,j}$  in (26) and (27), the ratio of the average total transmission power in Setup1 to that in Setup2 is approximately (but not exactly) 16.

## VII. CONCLUSIONS

In this paper, we have considered the chance SINR constrained JPAC problem, and have proposed two sample approximation-based deflation approaches for solving the problem. We first approximated the computationally intractable chance SINR constraints by sampling, and then reformulated the sampled JPAC problem as a composite group sparse minimization problem. Furthermore, we approximated the NP-hard group sparse minimization problem by a convex problem (equivalent to an SOCP) and used its solution to check the simultaneous supportability of all links and to guide an iterative link removal procedure (the deflation approach), resulting in two efficient deflation algorithms (SOCP-D and PABB-D).

The proposed approaches are particularly attractive for practical implementations for the following reasons. First, the two proposed approaches only require the CDI, which is more practical than most of the existing algorithms for JPAC where the perfect instantaneous CSI is required. Second, the two proposed approaches enjoy a low computational complexity. The SOCP-D approach has a polynomial time complexity. To further improve the computational efficiency, the special structure of the SOCP approximation problem is exploited, and an efficient algorithm, PABB-D, is custom designed for solving it. The PABB-D algorithm significantly outperforms the SOCP-D algorithm in terms of the execution CPU time. Finally, our simulation results show that the proposed approaches are very effective by using the NLPD algorithm as the benchmark.

## APPENDIX A

*Proof of Proposition 1:* We prove Proposition 1 by contraction. Assume that link  $k$  is supported in (16) but  $\mathbf{A}_k \mathbf{q}^* > \mathbf{c}_k$  holds true. Then we can construct a feasible point  $\hat{\mathbf{q}}$  satisfying

$$\sum_{k \in \mathcal{K}} \|\max\{\mathbf{c}_k - \mathbf{A}_k \hat{\mathbf{q}}, \mathbf{0}\}\|_0 + \alpha \bar{\mathbf{p}}^T \hat{\mathbf{q}} < \sum_{k \in \mathcal{K}} \|\max\{\mathbf{c}_k - \mathbf{A}_k \mathbf{q}^*, \mathbf{0}\}\|_0 + \alpha \bar{\mathbf{p}}^T \mathbf{q}^*. \quad (29)$$

Define  $\hat{\mathbf{q}} = (\hat{q}_1, \hat{q}_2, \dots, \hat{q}_K)^T$  with

$$\hat{q}_j = \begin{cases} \max\{(\mathbf{E}_k - \mathbf{A}_k) \mathbf{q}^* + \mathbf{c}_k\}, & \text{if } j = k; \\ q_j^*, & \text{if } j \neq k. \end{cases}$$

Recalling the definitions of  $\mathbf{E}_k$  and  $\mathbf{A}_k$ , we know  $\mathbf{E}_k - \mathbf{A}_k$  is a nonnegative matrix, and thus  $\hat{q}_k = \max\{(\mathbf{E}_k - \mathbf{A}_k) \mathbf{q}^* + \mathbf{c}_k\} > 0$ . Since

$$\mathbf{A}_k \mathbf{q}^* = \mathbf{E}_k \mathbf{q}^* - (\mathbf{E}_k - \mathbf{A}_k) \mathbf{q}^* = q_k^* \mathbf{e} - (\mathbf{E}_k - \mathbf{A}_k) \mathbf{q}^* > \mathbf{c}_k,$$

it follows from the definition of  $\hat{q}_k$  that

$$q_k^* > \max\{(\mathbf{E}_k - \mathbf{A}_k) \mathbf{q}^* + \mathbf{c}_k\} = \hat{q}_k.$$

Hence,  $\hat{\mathbf{q}}$  is feasible (i.e.,  $\mathbf{0} \leq \hat{\mathbf{q}} \leq \mathbf{q}^* \leq \mathbf{e}$  and the inequality  $\hat{\mathbf{q}} \leq \mathbf{q}^*$  holds true strictly for the  $k$ -th entry) and

$$\bar{\mathbf{p}}^T \hat{\mathbf{q}} < \bar{\mathbf{p}}^T \mathbf{q}^*. \quad (30)$$

Moreover, it follows from the definition of  $\hat{\mathbf{q}}$  that  $\mathbf{A}_k \hat{\mathbf{q}} \geq \mathbf{c}_k$ . For any  $j \neq k$ , if  $\mathbf{A}_j \mathbf{q}^* \geq \mathbf{c}_j$ , we have  $\mathbf{A}_j \hat{\mathbf{q}} - \mathbf{c}_j \geq \mathbf{A}_j \mathbf{q}^* - \mathbf{c}_j \geq \mathbf{0}$ , where the first inequality is due to the fact that all entries of  $\mathbf{A}_j$  except its  $j$ -th column are nonpositive and the fact  $\hat{\mathbf{q}} \leq \mathbf{q}^*$ . Consequently, there holds  $\mathcal{J}^* \subset \hat{\mathcal{J}}$ , where  $\mathcal{J}^* = \{j \mid \mathbf{A}_j \mathbf{q}^* \geq \mathbf{c}_j\}$  and  $\hat{\mathcal{J}} = \{j \mid \mathbf{A}_j \hat{\mathbf{q}} \geq \mathbf{c}_j\}$ . Thus, we have

$$\begin{aligned} \sum_{k \in \mathcal{K}} \|\max\{\mathbf{c}_k - \mathbf{A}_k \mathbf{q}^*, \mathbf{0}\}\|_0 &= \sum_{k \notin \mathcal{J}^*} \|\max\{\mathbf{c}_k - \mathbf{A}_k \mathbf{q}^*, \mathbf{0}\}\|_0 \\ &\geq \sum_{k \notin \hat{\mathcal{J}}} \|\max\{\mathbf{c}_k - \mathbf{A}_k \hat{\mathbf{q}}, \mathbf{0}\}\|_0 \\ &= \sum_{k \in \mathcal{K}} \|\max\{\mathbf{c}_k - \mathbf{A}_k \hat{\mathbf{q}}, \mathbf{0}\}\|_0. \end{aligned} \quad (31)$$

Combining (30) and (31) yields (29), which contradicts the optimality of  $\mathbf{q}^*$ . This completes the proof of Proposition 1.

## APPENDIX B

*Proof of Proposition 2:* To prove Proposition 2, we first consider the simple case where  $h(\mathbf{q}) = \|\max\{\mathbf{q}, \mathbf{0}\}\|_2$ .

*Lemma 1:* Suppose  $h(\mathbf{q}) = \|\max\{\mathbf{q}, \mathbf{0}\}\|_2$ . Then  $\partial h(\mathbf{0}) = \{\mathbf{s} \mid \mathbf{s} \geq \mathbf{0}, \|\mathbf{s}\|_2 \leq 1\}$ . If there exists  $i$  such that  $(\mathbf{q})_i > 0$ , then  $h(\mathbf{q})$  is differentiable and

$$\nabla h(\mathbf{q}) = \frac{\max\{\mathbf{q}, \mathbf{0}\}}{\|\max\{\mathbf{q}, \mathbf{0}\}\|_2}. \quad (32)$$

*Proof:* Define  $\mathcal{S} = \{\mathbf{s} \mid \mathbf{s} \geq \mathbf{0}, \|\mathbf{s}\|_2 \leq 1\}$ . We claim  $\partial h(\mathbf{0}) = \mathcal{S}$ . On one hand, taking any  $\mathbf{s} \in \mathcal{S}$ , we have that

$$h(\mathbf{q}) = \|\max\{\mathbf{q}, \mathbf{0}\}\|_2 \geq \mathbf{s}^T \max\{\mathbf{q}, \mathbf{0}\} \geq \mathbf{s}^T \mathbf{q} = h(\mathbf{0}) + \mathbf{s}^T (\mathbf{q} - \mathbf{0}), \quad \forall \mathbf{q},$$

where the first inequality is due to the Cauchy-Schwarz inequality and the fact  $\|\mathbf{s}\|_2 \leq 1$ , and the second inequality is due to  $\mathbf{s} \geq \mathbf{0}$ . This shows that  $\mathcal{S} \subset \partial h(\mathbf{0})$  according to the definition of  $\partial h(\mathbf{0})$  [43]. On the other hand, to show  $\partial h(\mathbf{0}) \subset \mathcal{S}$ , it suffices to show that any point  $\mathbf{s} \notin \mathcal{S}$  is not a subgradient of  $h(\mathbf{q})$  at point  $\mathbf{0}$ . In particular, if  $\|\mathbf{s}\|_2 > 1$ , then

$$h(\mathbf{q}) = \|\max\{\mathbf{q}, \mathbf{0}\}\|_2 \leq \|\mathbf{q}\|_2 = 1 < \|\mathbf{s}\|_2 = \mathbf{s}^T \mathbf{q} = h(\mathbf{0}) + \mathbf{s}^T (\mathbf{q} - \mathbf{0})$$

with  $\mathbf{q} = \mathbf{s}/\|\mathbf{s}\|_2$ . Thus, the subgradient  $\mathbf{s}$  of  $h(\mathbf{q})$  at point  $\mathbf{0}$  must satisfy  $\|\mathbf{s}\|_2 \leq 1$ . If  $\|\mathbf{s}\|_2 \leq 1$  but  $(\mathbf{s})_1 < 0$  (without loss of generality), we test the point  $\mathbf{q} = (-1, 0, \dots, 0)^T$ , and obtain

$$h(\mathbf{q}) = \|\max\{\mathbf{q}, \mathbf{0}\}\|_2 = 0 < -(\mathbf{s})_1 = \mathbf{s}^T \mathbf{q} = h(\mathbf{0}) + \mathbf{s}^T (\mathbf{q} - \mathbf{0}).$$

Consequently, the subgradient  $\mathbf{s}$  of  $h(\mathbf{q})$  at point  $\mathbf{0}$  must satisfy  $\|\mathbf{s}\|_2 \leq 1$  and  $\mathbf{s} \geq \mathbf{0}$ . Hence,  $\partial h(\mathbf{0}) = \{\mathbf{s} \mid \mathbf{s} \geq \mathbf{0}, \|\mathbf{s}\|_2 \leq 1\}$ .

Next, we show that  $h(\mathbf{q})$  is differentiable at the point  $\mathbf{q}$  which has at least one positive entry, and the corresponding gradient is given in (32). In fact, although the function  $\max\{q, 0\}$  is nondifferentiable at point  $q = 0$ , its square  $f(q) = (\max\{q, 0\})^2$  is differentiable everywhere; i.e.,

$$f'(q) = \begin{cases} 0, & \text{if } q \leq 0; \\ 2q, & \text{if } q > 0. \end{cases}$$

According to the composite rule of differentiation, we know that the gradient of  $h(\mathbf{q})$  is given by (32). This completes the proof of Lemma 1. ■

Equipping with Lemma 1, we now can prove Proposition 2. Without loss of generality, assume that  $\mathbf{c}_k - \mathbf{A}_k \bar{\mathbf{q}} = \mathbf{0}$ . Then, for any  $\mathbf{q}$  and any  $\mathbf{s}$  satisfying  $\|\mathbf{s}\|_2 \leq 1$  and  $\mathbf{s} \geq \mathbf{0}$ , we have

$$\begin{aligned} h_k(\mathbf{q}) &= \|\max\{\mathbf{c}_k - \mathbf{A}_k \mathbf{q}, \mathbf{0}\}\|_2 \geq \mathbf{s}^T (\mathbf{c}_k - \mathbf{A}_k \mathbf{q}) \quad (\text{from Lemma 1}) \\ &= h_k(\bar{\mathbf{q}}) + (-\mathbf{A}_k^T \mathbf{s})^T (\mathbf{q} - \bar{\mathbf{q}}), \end{aligned} \quad (33)$$

which shows that all vectors in  $\{-\mathbf{A}_k^T \mathbf{s} \mid \mathbf{s} \geq \mathbf{0}, \|\mathbf{s}\|_2 \leq 1\}$  are subgradients of  $h_k(\mathbf{q})$  at point  $\bar{\mathbf{q}}$ . In the same way as in the proof of Lemma 1, we can show that if  $\mathbf{s}$  does not satisfy  $\|\mathbf{s}\|_2 \leq 1$  and  $\mathbf{s} \geq \mathbf{0}$ , the inequality in (33) will violate for some special choice of  $\mathbf{q}$ . Hence,  $\partial h_k(\bar{\mathbf{q}}) = \{-\mathbf{A}_k^T \mathbf{s} \mid \mathbf{s} \geq \mathbf{0}, \|\mathbf{s}\|_2 \leq 1\}$ .

If  $\mathcal{N}_k^+ \neq \emptyset$ , we know from the composite rule of differentiation and Lemma 1 that  $h_k(\bar{\mathbf{q}})$  is differentiable and its gradient is given by

$$\nabla h_k(\bar{\mathbf{q}}) = \frac{-\sum_{n \in \mathcal{N}_k^+} (\mathbf{c}_k - \mathbf{A}_k \bar{\mathbf{q}})_n (\mathbf{a}_k^n)^T}{\|\max\{\mathbf{c}_k - \mathbf{A}_k \bar{\mathbf{q}}, \mathbf{0}\}\|_2} = \frac{-\mathbf{A}_k^T \max\{\mathbf{c}_k - \mathbf{A}_k \bar{\mathbf{q}}, \mathbf{0}\}}{\|\max\{\mathbf{c}_k - \mathbf{A}_k \bar{\mathbf{q}}, \mathbf{0}\}\|_2}.$$

This completes the proof of Proposition 2.

## APPENDIX C

*Proof of Theorem 2:* Suppose all links in the network can be simultaneously supported (i.e., there exists  $\mathbf{0} \leq \mathbf{q} \leq \mathbf{e}$  satisfying  $\mathbf{A}\mathbf{q} \geq \mathbf{c}$ ) and  $\bar{\mathbf{q}}$  is the solution to problem

$$\begin{aligned} \min_{\mathbf{q}} \quad & \bar{\mathbf{p}}^T \mathbf{q} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{q} - \mathbf{c} \geq \mathbf{0}, \\ & \mathbf{0} \leq \mathbf{q} \leq \mathbf{e}. \end{aligned}$$

To prove Theorem 2, it suffices to show that  $\bar{\mathbf{q}}$  is also the solution to problem (17) with  $\alpha \in (0, \alpha_2]$ . Moreover, to show  $\bar{\mathbf{q}}$  is the solution to problem (17), we only need to show that the subdifferential of the objective function of problem (17) at point  $\bar{\mathbf{q}}$  contains  $\mathbf{0}$  [43]. Next, we claim the latter is true.

We first characterize the subdifferential of the objective function of problem (17) at point  $\bar{\mathbf{q}}$ . It follows from (14) that there exists  $\mathcal{I} = \{n_1, n_2, \dots, n_K\}$  such that  $\bar{\mathbf{q}}$  is the solution to the following linear system

$$\mathbf{A}_{\mathcal{I}} \mathbf{q} := [\mathbf{a}_1^{n_1}; \mathbf{a}_2^{n_2}; \dots; \mathbf{a}_K^{n_K}] \mathbf{q} = (c_1^{n_1}; c_2^{n_2}; \dots; c_K^{n_K}) := \mathbf{c}_{\mathcal{I}}.$$

Recalling the definition of  $\mathbf{a}_k^n$  (see Subsection IV-B), we know that  $\mathbf{I} - \mathbf{A}_{\mathcal{I}}$  is a nonnegative matrix. Moreover, from [52, Theorem 1.15],  $\mathbf{A}_{\mathcal{I}}$  is nonsingular,  $\mathbf{A}_{\mathcal{I}}^{-1}$  is nonnegative, and

$$0 < \mathbf{e}^T \mathbf{A}_{\mathcal{I}}^{-1} \mathbf{c}_{\mathcal{I}} = \mathbf{e}^T \bar{\mathbf{q}} \leq K. \quad (34)$$

Define  $h_k(\mathbf{q}) = \|\max\{\mathbf{c}_k - \mathbf{A}_k \mathbf{q}, \mathbf{0}\}\|_2$  for  $k \in \mathcal{K}$ . It follows from [43, Theorem 23.8] that the subdifferential of the objective function of problem (17) at point  $\bar{\mathbf{q}}$  is given by

$$\left\{ \sum_{k \in \mathcal{K}} \mathbf{g}_k + \alpha \bar{\mathbf{p}} \mid \mathbf{g}_k \in \partial h_k(\bar{\mathbf{q}}), k \in \mathcal{K} \right\}.$$

According to Proposition 2,  $\partial h_k(\bar{\mathbf{q}})$  contains<sup>2</sup> all vectors in

$$\mathcal{S}_k = \left\{ -s_k (\mathbf{a}_k^{n_k})^T \mid 0 \leq s_k \leq 1 \right\}. \quad (35)$$

Therefore, all vectors in

$$\mathcal{S} = \left\{ -\mathbf{A}_{\mathcal{I}}^T \mathbf{s} + \alpha \bar{\mathbf{p}} \mid \mathbf{0} \leq \mathbf{s} \leq \mathbf{e} \right\} = \left\{ -\sum_{k \in \mathcal{K}} s_k (\mathbf{a}_k^{n_k})^T + \alpha \bar{\mathbf{p}} \mid 0 \leq s_k \leq 1, k \in \mathcal{K} \right\}$$

are subgradients of the objective function of problem (17) at point  $\bar{\mathbf{q}}$ .

If  $\mathbf{0} \in \mathcal{S}$ , the subdifferential of the objective function of problem (17) at point  $\bar{\mathbf{q}}$  contains  $\mathbf{0}$  [43], which completes the proof of Theorem 2. Next, we show  $\mathbf{0} \in \mathcal{S}$  is true. Consider the vector  $\mathbf{s} = \alpha \mathbf{A}_{\mathcal{I}}^{-T} \bar{\mathbf{p}}$ . It is a nonnegative vector (since  $\mathbf{A}_{\mathcal{I}}^{-1}$  is nonnegative), and each of its entries is less than or equal to 1 as long as  $\alpha \leq \alpha_2$ . This is because

$$\mathbf{e}^T \mathbf{s} = \mathbf{e}^T \alpha \mathbf{A}_{\mathcal{I}}^{-T} \bar{\mathbf{p}} \leq \alpha \max\{\bar{\mathbf{p}}\} \mathbf{e}^T \mathbf{A}_{\mathcal{I}}^{-T} \mathbf{e} \leq \alpha \frac{\max\{\bar{\mathbf{p}}\}}{\min\{\mathbf{c}_{\mathcal{I}}\}} \mathbf{c}_{\mathcal{I}}^T \mathbf{A}_{\mathcal{I}}^{-T} \mathbf{e} \stackrel{(a)}{\leq} \alpha \frac{\max\{\bar{\mathbf{p}}\}}{\min\{\mathbf{c}\}} K \leq 1,$$

where (a) is due to (34) and the fact  $\min\{\mathbf{c}\} \leq \min\{\mathbf{c}_{\mathcal{I}}\}$ . Substituting  $\mathbf{s} = \alpha \mathbf{A}_{\mathcal{I}}^{-T} \bar{\mathbf{p}}$  into  $\mathcal{S}$ , we obtain  $-\mathbf{A}_{\mathcal{I}}^T \mathbf{s} + \alpha \bar{\mathbf{p}} = -\mathbf{A}_{\mathcal{I}}^T (\alpha \mathbf{A}_{\mathcal{I}}^{-T} \bar{\mathbf{p}}) + \alpha \bar{\mathbf{p}} = \mathbf{0}$ . Thus,  $\mathbf{0} \in \mathcal{S}$ .

#### ACKNOWLEDGMENT

The authors would like to thank Professor Yu-Hong Dai of the Chinese Academy of Sciences for many useful discussions and the anonymous reviewers for their useful comments.

<sup>2</sup>The subdifferential of  $h_k(\mathbf{q})$  at point  $\bar{\mathbf{q}}$  is not necessarily equal to  $\mathcal{S}_k$  in (35). This is because that some other entries (except the  $n_k$ -th entry) of the vector  $\mathbf{c}_k - \mathbf{A}_k \bar{\mathbf{q}}$  might also be zero.

## REFERENCES

- [1] I. Mitliagkas, N. D. Sidiropoulos, and A. Swami, "Joint power and admission control for ad-hoc and cognitive underlay networks: Convex approximation and distributed implementation," *IEEE Trans. Wireless Commun.*, vol. 10, no. 12, pp. 4110–4121, Dec. 2011.
- [2] H. Mahdavi-Doost, M. Ebrahimi, and A. K. Khandani, "Characterization of SINR region for interfering links with constrained power," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2816–2828, Jun. 2010.
- [3] G. J. Foschini and Z. Miljanic, "A simple distributed autonomous power control algorithm and its convergence," *IEEE Trans. Veh. Technol.*, vol. 42, no. 4, pp. 641–646, Nov. 1993.
- [4] M. Andersin, Z. Rosberg, and J. Zander, "Gradual removals in cellular PCS with constrained power control and noise," *Wireless Netw.*, vol. 2, no. 1, pp. 27–43, Mar. 1996.
- [5] J. Zander, "Performance of optimum transmitter power control in cellular radio systems," *IEEE Trans. Veh. Technol.*, vol. 41, no. 1, pp. 57–62, Feb. 1992.
- [6] R. D. Yates, "A framework for uplink power control in cellular radio systems," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 7, pp. 1341–1347, Sept. 1995.
- [7] S. A. Grandhi, J. Zander, and R. Yates, "Constrained power control," *Wireless Personal Commun.*, vol. 1, no. 4, pp. 257–270, Dec. 1994.
- [8] Y.-F. Liu, Y.-H. Dai, and Z.-Q. Luo, "Joint power and admission control via linear programming deflation," *IEEE Trans. Signal Process.*, vol. 61, no. 6, pp. 1327–1338, Mar. 2013.
- [9] Y.-F. Liu, Y.-H. Dai, and S. Ma, "Joint power and admission control: Non-convex lq approximation and an effective polynomial time deflation approach," *IEEE Trans. Signal Process.*, vol. 63, no. 14, pp. 3641–3656, July 2015.
- [10] D. I. Evangelinakis, N. D. Sidiropoulos, and A. Swami, "Joint admission and power control using branch & bound and gradual admissions," in *Proc. Int. Workshop Signal Process. Advances Wireless Commun. (SPAWC)*, Jun. 2010, pp. 1–5.
- [11] Y.-F. Liu, "An efficient distributed joint power and admission control algorithm," in *Proc. 31th Chinese Control Conference (CCC)*, July, 2012, pp. 5508–5512.
- [12] Y.-F. Liu and E. Song, "Distributionally robust joint power and admission control via SOCP deflation," in *Proc. SPAWC*, June, 2013, pp. 11–15.
- [13] E. Matskani, N. D. Sidiropoulos, Z.-Q. Luo, and L. Tassiulas, "Convex approximation techniques for joint multiuser downlink beamforming and admission control," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2682–2693, Jul. 2008.
- [14] J.-T. Wang, "Admission control with distributed joint diversity and power control for wireless networks," *IEEE Trans. Veh. Technol.*, vol. 58, no. 1, pp. 409–419, Jan. 2009.
- [15] T. Elbatt and A. Ephremides, "Joint scheduling and power control for wireless ad hoc networks," *IEEE Trans. Wireless Commun.*, vol. 3, no. 1, pp. 74–85, Jan. 2004.
- [16] H.-T. Wai and W.-K. Ma, "A decentralized method for joint admission control and beamforming in coordinated multicell downlink," in *Proc. 46th Asilomar Conf. Signals, Systems, and Computers*, Nov. 2012, pp. 559–563.
- [17] Y.-F. Liu and Y.-H. Dai, "Joint power and admission control via  $p$  norm minimization deflation," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, May 2013, pp. 4789–4793.
- [18] W. S. Lai, T. H. Chang, and T. S. Lee, "Joint power and admission control for spectral and energy efficiency maximization in heterogeneous OFDMA networks," to appear in *IEEE Trans. Wireless Commun.*
- [19] X. Gong, S. A. Vorobyov, and C. Tellambura, "Joint bandwidth and power allocation with admission control in wireless multi-user networks with and without relaying," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1801–1813, Apr. 2011.



- [20] E. Matskani, N. D. Sidiropoulos, Z.-Q. Luo, and L. Tassiulas, "Efficient batch and adaptive approximation algorithms for joint multicast beamforming and admission control," *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4882–4894, Dec. 2009.
- [21] K. T. Phan, T. Le-Ngoc, S. A. Vorobyov, and C. Tellambura, "Power allocation in wireless multi-user relay networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 5, pp. 2535–2545, May, 2009.
- [22] M. Rasti, A. R. Sharafat, and J. Zander, "Pareto and energy-efficient distributed power control with feasibility check in wireless networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 245–255, Jan. 2011.
- [23] N. Bambos, S.-C. Chen, and G. J. Pottie, "Channel access algorithms with active link protection for wireless communication networks with power control," *IEEE Trans. Netw.*, vol. 8, no. 5, pp. 583–597, Oct. 2000.
- [24] M. Andersin, Z. Rosberg, and J. Zander, "Soft and safe admission control in cellular networks," *IEEE/ACM Trans. Netw.*, vol. 5, no. 2, pp. 255–265, Apr. 1997.
- [25] M. Xiao, N. B. Shroff, and E. K. P. Chong, "Distributed admission control for power-controlled cellular wireless systems," *IEEE/ACM Trans. Netw.*, vol. 9, no. 6, pp. 790–800, Dec. 2001.
- [26] D. J. Love, R. W. Heath, V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1341–1365, Oct. 2008.
- [27] K.-Y. Wang, N. Jacklin, Z. Ding, and C.-Y. Chi, "Robust MISO transmit optimization under outage-based QoS constraints in two-tier heterogeneous networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 4, pp. 1883–1897, Apr. 2013.
- [28] A. M.-C. So and Y.-J. Zhang, "Distributionally robust slow adaptive OFDMA with soft QoS via linear programming," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 5, pp. 947–958, 2013.
- [29] C. Fischione, M. Butussi, K. H. Johansson, and M. D'Angelo, "Power and rate control with outage constraints in CDMA wireless networks," *IEEE Trans. Commun.*, vol. 57, no. 8, pp. 2225–2229, Aug. 2009.
- [30] W.-C. Li, T.-H. Chang, C. Lin, and C.-Y. Chi, "Coordinated beamforming for multiuser MISO interference channel under rate outage constraints," *IEEE Trans. Signal Process.*, vol. 61, no. 5, pp. 1087–1103, Mar. 2013.
- [31] R. Sun, H. Baligh, and Z.-Q. Luo, "Long-term transmit point association for coordinated multipoint transmission by stochastic optimization," in *Proc. SPAWC*, June, 2013, pp. 91–95.
- [32] B. Khoshnevis and W. Yu, "Joint power control and beamforming codebook design for MISO channels under the outage criterion," *Journal of Franklin Institute Special Issue on Modelling and Simulation in Advanced Communications*, vol. 349, no. 1, pp. 140–161, Feb. 2012.
- [33] S. Ghosh, B. D. Rao, and J. R. Zeidler, "Outage-efficient strategies for multiuser MIMO networks with channel distribution information," *IEEE Trans. Signal Process.*, vol. 58, no. 12, pp. 6312–6324, Dec. 2010.
- [34] Y. Rong, S. A. Vorobyov, and A. B. Gershman, "Robust linear receivers for multiaccess space-time block-coded MIMO systems: A probabilistically constrained approach," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1560–1570, Aug. 2006.
- [35] G. Calafiore and M. C. Campi, "Uncertain convex programs: Randomized solutions and confidence levels," *Math. Prog.*, vol. 102, pp. 25–46, 2005.
- [36] M. C. Campi and S. Garatti, "The exact feasibility of randomized solutions of uncertain convex programs," *SIAM J. Optim.*, vol. 19, no. 3, pp. 1211–1230, 2008.
- [37] Y.-F. Liu, M. Hong, and E. Song, "A companion technical report of "Sample approximation-based deflation approaches for chance SINR constrained joint power and admission control"," *Academy of Mathematics and Systems Science, Chinese Academy of Sciences*, 2015. [Online]. Available: [http://lsec.cc.ac.cn/~yafliu/report\\_robust.pdf](http://lsec.cc.ac.cn/~yafliu/report_robust.pdf)

- [38] A. Nemirovski and A. Shapiro, “Convex approximations of chance constrained programs,” *SIAM J. Optim.*, vol. 17, no. 4, pp. 969–996, 2006.
- [39] A. M.-C. So, “Moment inequalities for sums of random matrices and their applications in optimization,” *Math. Prog.*, vol. 130, pp. 125–151, 2011.
- [40] J. R. Birge and F. Louveaux, *Introduction to Stochastic Programming*, 2nd ed. New York, U.S.A.: Springer, 2011.
- [41] M. Yuan and Y. Lin, “Model selection and estimation in regression with grouped variables,” *J. R. Stat. Soc. Ser. B*, vol. 68, no. 1, pp. 49–67, 2006.
- [42] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, U.S.A.: Cambridge University Press, 2004.
- [43] R. T. Rockafellar, *Convex Analysis*. Princeton, NJ, U.S.A.: Princeton University Press, 1970.
- [44] M. Grant, S. Boyd, and Y. Ye, *CVX: matlab software for disciplined convex programming*, 2006. [Online]. Available: <http://www.stanford.edu/~boyd/cvx/>.
- [45] Y.-H. Dai and R. Fletcher, “Projected Barzilai-Borwein methods for large-scale box-constrained quadratic programming,” *Numer. Math.*, vol. 100, no. 1, pp. 21–47, 2005.
- [46] J. Barzilai and J. M. Borwein, “Two-point step size gradient methods,” *IMA J. Numer. Anal.*, vol. 8, no. 1, pp. 141–148, 1988.
- [47] E. T. Hale, W. Yin, and Y. Zhang, “A numerical study of fixed-point continuation applied to compressed sensing,” *J. Comp. Math.*, vol. 28, no. 2, pp. 170–194, 2010.
- [48] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. New York, U.S.A.: Springer, 2006.
- [49] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization*. Philadelphia, U.S.A.: SIAM-MPS Series on Optimization, SIAM Publications, 2001.
- [50] Y.-H. Dai and L.-Z. Liao, “R-linear convergence of the Barzilai and Borwein gradient method,” *IMA J. Numer. Anal.*, vol. 22, pp. 1–10, 2002.
- [51] A. Goldsmith, *Wireless Communications*. New York, U.S.A.: Cambridge University Press, 2005.
- [52] Y. Saad, *Iterative Methods for Sparse Linear Systems*, 2nd ed. Philadelphia, U.S.A.: SIAM Publications, 2003.