

Joint Power and Admission Control: Non-Convex L_q Approximation and An Effective Polynomial Time Deflation Approach

Ya-Feng Liu, *Member, IEEE*, Yu-Hong Dai, and Shiqian Ma

Abstract—In an interference limited network, joint power and admission control (JPAC) aims at supporting a maximum number of links at their specified signal-to-interference-plus-noise ratio (SINR) targets while using minimum total transmission power. Various convex approximation deflation approaches have been developed for the JPAC problem. In this paper, we propose an effective polynomial time *non-convex* approximation deflation approach for solving the problem. The approach is based on the non-convex ℓ_q ($0 < q < 1$) approximation of an equivalent sparse ℓ_0 reformulation of the JPAC problem. We show that, for any instance of the JPAC problem, there exists a $\bar{q} \in (0, 1)$ such that it can be exactly solved by solving its ℓ_q approximation problem with any $q \in (0, \bar{q}]$. We also show that finding the global solution of the ℓ_q approximation problem is NP-hard. Then, we propose a potential reduction interior-point algorithm, which can return an ϵ -KKT solution of the NP-hard ℓ_q approximation problem in polynomial time. The returned solution can be used to check the simultaneous supportability of all links in the network and to guide an iterative link removal procedure, resulting in the polynomial time non-convex approximation deflation approach for the JPAC problem. Numerical simulations show that the proposed approach outperforms the existing convex approximation approaches in terms of the number of supported links and the total transmission power, particularly exhibiting a quite good performance in selecting which subset of links to support.

Index Terms—Admission control, complexity, non-convex approximation, potential reduction algorithm, power control, sparse optimization.

Manuscript received November 01, 2013; revised October 01, 2014 and January 29, 2015; accepted April 14, 2015. Date of publication April 30, 2015; date of current version June 08, 2015. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Shuguang Cui. The first author is partially supported by the National Natural Science Foundation, Grants 11301516 and 11331012. The second author is partially supported by the China National Funds for Distinguished Young Scientists, Grant 1125107, and the Key Project of Chinese National Programs for Fundamental Research and Development, Grant 2015CB856000, and the CAS Program for Cross & Cooperative Team of the Science & Technology Innovation. S. Ma is partially supported by a Direct Grant of the Chinese University of Hong Kong (Project ID: 4055016) and the Hong Kong Research Grants Council General Research Fund Early Career Scheme (Project ID: CUHK 439513). Part of this work was presented in the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Vancouver, Canada, May 26–31, 2013 [25]. (Corresponding Author: Ya-Feng Liu.)

Y.-F. Liu and Y.-H. Dai are 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; dyh@lsec.cc.ac.cn).

S. Ma is with the Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong (e-mail: sqma@se.cuhk.edu.hk).

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

Digital Object Identifier 10.1109/TSP.2015.2428224

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 more than two decades [1]–[31]. The goal of JPAC 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 interference limited network cannot be simultaneously supported. JPAC can not only determine which interfering links must be turned off and rescheduled along orthogonal resource dimensions (such as time, space, or frequency slots), but also alleviate the difficulties of the convergence of stand-alone power control algorithms. For example, a longstanding issue associated with the Foschini-Miljanic algorithm [5] 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.

A. Related Work

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. Another globally optimal algorithm, which is based on the branch and bound strategy, is given in [10]. Theoretically, the problem is shown to be NP-hard to solve (to global optimality) and to approximate (to constant factor of global optimality) [1], [2], [4]. In recent years, various convex approximation based heuristics algorithms [1]–[11], [16]–[31] have been proposed for the problem, since convex optimization problems, such as linear program (LP), second-order cone program (SOCP), and semidefinite program (SDP), are relatively easy to solve¹.

Assuming perfect channel state information (CSI), [1] proposed the so-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 uses its solution to guide the removal of interfering links. The removal procedure is repeated until all the remaining links in the network are simultaneously supportable. In [2], the JPAC problem

¹For any convex optimization problem and any $\epsilon > 0$, the ellipsoid algorithm can find an ϵ -optimal solution (i.e., a feasible solution whose objective value is within ϵ from being globally optimal) with a complexity that is polynomial in the problem dimension and $\log(1/\epsilon)$ [32].

is shown to be equivalent to a sparse ℓ_0 minimization problem and then its ℓ_1 convex relaxation is used to derive an LP, which is different from the one in [1]. Again, the solution to the derived LP is used to guide an iterative link removal procedure (deflation), leading to an efficient new linear programming deflation (NLPD) algorithm. Another convex approximation based heuristics algorithm is proposed in [3]. Assuming the same SINR target for each link, the link that results in the largest increase in the achievable SINR is removed at each iteration until all the remaining links in the network are simultaneously supportable. To determine the removed link, a large number of extreme eigenvalue problems² need to be solved at each iteration, making the removal procedure computationally expensive. To reduce the computational complexity, the above idea is approximately implemented in the Algorithm II-B [3]. Similar convex approximation deflation ideas were used in [19], [30] to solve the joint beamforming and admission control problem for the cellular downlink network, where at each iteration an SDP needs to be solved to determine the link to be removed.

Under the imperfect CSI assumption, JPAC has been studied in [1], [28], [29]. In [1], the authors considered the worst-case robust JPAC problem with bounded channel estimation errors. The key there is that the relaxed LP with bounded uncertainty can be equivalently rewritten as an 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. [29] studied the JPAC problem under the assumption of the channel distribution information (CDI), and formulated the JPAC problem as 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 difficulty of the chance SINR constraint, [29] employed the sample (scenario) approximation scheme to convert the chance constraints into finitely many simple linear constraints. Then, the sample approximation of the chance SINR constrained JPAC problem is reformulated as a group sparse minimization problem and approximated by an SOCP. The solution of the SOCP approximation problem can be used to check the simultaneous supportability of all links in the network and to guide an iterative link removal procedure.

B. Our Contribution

This paper considers the JPAC problem under the perfect CSI assumption. We remark that similar techniques can be used for the case where the CSI is not perfectly known. As mentioned above, most existing algorithms for JPAC are based on (successive) convex approximations. The main contribution of this paper is to propose an effective polynomial time non-convex approximation deflation approach for solving the JPAC problem. To our knowledge, this is the first approach that solves the JPAC problem by (successive) non-convex approximations. The key idea is to approximate the sparse ℓ_0 minimization reformulation of the JPAC problem by the non-convex ℓ_q minimization problem with $q \in (0, 1)$ instead of the convex ℓ_1 minimization

problem as in [1], [2], and to design a polynomial time algorithm for computing an ϵ -KKT solution (its definition will be given later) of the non-convex ℓ_q minimization problem for any given $\epsilon > 0$. The main results of this paper are summarized as follows.

- We show that the non-convex ℓ_q minimization approximation problem shares the same solution with the ℓ_0 minimization problem if $q \in (0, \bar{q}]$, where \bar{q} is some value in $(0, 1)$. We also give an example of the JPAC problem, showing that the solution to its non-convex ℓ_q minimization approximation problem with any $q \in (0, 1)$ solves the original problem while its convex ℓ_1 minimization approximation problem fails to do so. We therefore show that the ℓ_q minimization problem with $q \in (0, 1)$ approximates the ℓ_0 minimization JPAC problem better than the ℓ_1 minimization problem.
- We show that, for any $q \in (0, 1)$, the ℓ_q minimization approximation problem is NP-hard. The proof is based on a polynomial time transformation from the partition problem. The complexity result suggests that there is no polynomial time algorithm which can solve the ℓ_q minimization approximation problem to global optimality (unless $P = NP$).
- We reformulate the ℓ_q minimization approximation problem and develop a potential reduction interior-point algorithm for solving its equivalent reformulation. We show that, for any given $\epsilon > 0$, the potential reduction algorithm can return an ϵ -KKT solution of the reformulated problem in polynomial time. The obtained ϵ -KKT solution can be used to check the simultaneous supportability of all links in the network and to guide an iterative link removal procedure, resulting in the polynomial time non-convex approximation deflation approach for the JPAC problem. Simulation results show that the proposed approach significantly outperforms the existing convex approximation deflation algorithms [1]–[3].

C. Notations

We adopt the following notations in this paper. 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 notations $\max\{\mathbf{x}\}$, $[\mathbf{x}]_k$, $\|\mathbf{x}\|_\infty$, and $\|\mathbf{x}\|_q^q := \sum_k |[\mathbf{x}]_k|^q$ ($0 \leq q \leq 1$) stand for its maximum entry, its k -th entry, its infinity norm, and its ℓ_q norm³, respectively. In particular, when $q = 0$, $\|\mathbf{x}\|_0$ stands for the number of nonzero entries in \mathbf{x} . For any subset $\mathcal{I} \subseteq \mathcal{K}$, we use $\mathbf{A}_{\mathcal{I}}$ to denote the matrix formed by the rows of \mathbf{A} indexed by \mathcal{I} . For a vector \mathbf{x} , the notation $\mathbf{x}_{\mathcal{I}}$ is similarly defined. Moreover, for any $\mathcal{J} \subseteq \mathcal{K}$, the notation $\mathbf{A}_{\mathcal{I}, \mathcal{J}}$ will denote the submatrix of \mathbf{A} obtained by taking the rows and columns of \mathbf{A} indexed by \mathcal{I} and \mathcal{J} respectively. The spectral radius and the infinity norm of a matrix \mathbf{A} are denoted by $\rho(\mathbf{A})$ and $\|\mathbf{A}\|_\infty$, respectively. Finally, we use \mathbf{e} to represent the vector with all components being one and \mathbf{I} to represent the identity matrix of an appropriate size, respectively.

³Strictly speaking, $\|\mathbf{x}\|_q^q$ with $0 \leq q < 1$ is not a norm, since it does not satisfy the triangle inequality. However, we still call it ℓ_q norm for convenience in this paper.

²Extreme eigenvalue problems are SDP representable.

II. SYSTEM MODEL, SPARSE FORMULATION, AND NLPD ALGORITHM

Consider a K -link (a link corresponds to a transmitter-receiver pair) 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 $k, j \in \mathcal{K} := \{1, 2, \dots, K\}$. 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

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

To some extent, the JPAC problem can be formulated as a two-stage optimization problem. The first stage maximizes the number of admitted links:

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

The optimal solution \mathcal{S}_0 of problem (1), which may not be unique, is called *maximum admissible set*. The second stage minimizes the total transmission power required to support the admitted links in \mathcal{S}_0 :

$$\begin{aligned} \min_{\{p_k\}_{k \in \mathcal{S}_0}} \quad & \sum_{k \in \mathcal{S}_0} p_k \\ \text{s.t.} \quad & \text{SINR}_k \geq \gamma_k, \quad k \in \mathcal{S}_0, \\ & 0 \leq p_k \leq \bar{p}_k, \quad k \in \mathcal{S}_0. \end{aligned} \quad (2)$$

Due to the special choice of \mathcal{S}_0 , power control problem (2) is feasible and can be efficiently solved by the Foschini-Miljanic algorithm [5].

The two-stage JPAC problem (1) and (2) is reformulated as a single-stage sparse ℓ_0 minimization problem in [2], which is based on a normalized channel. Next, we first introduce the channel normalization and then the sparse formulation of the JPAC problem. Denote the normalized power allocation vector by $\mathbf{x} = (x_1, x_2, \dots, x_K)^T$ with $x_k = p_k / \bar{p}_k$, and the normalized noise vector by $\mathbf{b} = (b_1, b_2, \dots, b_K)^T$ with $b_k = (\gamma_k \eta_k) / (g_{k,k} \bar{p}_k) > 0$. The normalized channel matrix is denoted by $\mathbf{A} \in \mathbb{R}^{K \times K}$ with the (k, j) -th entry

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

The JPAC problem can be reformulated as a single-stage sparse ℓ_0 minimization problem as follows

$$\begin{aligned} \min_{\mathbf{x}} \quad & \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_0 + \alpha \bar{\mathbf{p}}^T \mathbf{x} \\ \text{s.t.} \quad & \mathbf{0} \leq \mathbf{x} \leq \mathbf{e}, \end{aligned} \quad (3)$$

where α is a parameter satisfying

$$0 < \alpha < \alpha_1 := 1/\mathbf{e}^T \bar{\mathbf{p}}. \quad (4)$$

Notice that the formulation (3) is capable of finding the maximum admissible set with minimum total transmission power

and hence is superior to the two-stage formulation (1) and (2) in case of multiple maximum admissible sets.

The basic idea of the NLPD algorithm in [2] is to update the power and check whether all links can be supported. If not, drop one link from the network and update the power again. This process is repeated until all the remaining links are supported. More specifically, the NLPD algorithm checks whether all links in the network can be simultaneously supported by solving the ℓ_1 convex approximation of the ℓ_0 minimization problem (3)

$$\begin{aligned} \min_{\mathbf{x}} \quad & \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_1 + \alpha \bar{\mathbf{p}}^T \mathbf{x} \\ \text{s.t.} \quad & \mathbf{0} \leq \mathbf{x} \leq \mathbf{e}, \end{aligned} \quad (5)$$

which is equivalent to the following LP (see Theorem 2 in [2])

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{e}^T (\mathbf{b} - \mathbf{A}\mathbf{x}) + \alpha \bar{\mathbf{p}}^T \mathbf{x} \\ \text{s.t.} \quad & \mathbf{b} - \mathbf{A}\mathbf{x} \geq \mathbf{0}, \\ & \mathbf{0} \leq \mathbf{x} \leq \mathbf{e}. \end{aligned} \quad (6)$$

If all links in the network cannot be simultaneously supported, the NLPD algorithm drops the link

$$k_0 = \arg \max_{k \in \mathcal{K}} \left\{ \sum_{j \neq k} \left(|a_{k,j}| [\mathbf{b} - \mathbf{A}\mathbf{x}]_j + |a_{j,k}| [\mathbf{b} - \mathbf{A}\mathbf{x}]_k \right) \right\}. \quad (7)$$

To accelerate the deflation process, an easy-to-check necessary condition

$$(\boldsymbol{\mu}^+)^T \mathbf{e} - (\boldsymbol{\mu}^- + \mathbf{e})^T \mathbf{b} \geq 0 \quad (8)$$

for all links in the network to be simultaneously supported is also derived in [2], where $\boldsymbol{\mu}^+ = \max\{\boldsymbol{\mu}, \mathbf{0}\}$, $\boldsymbol{\mu}^- = \max\{-\boldsymbol{\mu}, \mathbf{0}\}$, and $\boldsymbol{\mu} = \mathbf{A}^T \mathbf{e}$. The necessary condition allows to iteratively remove strong interfering links from the network. In particular, the link

$$k_0 = \arg \max_{k \in \mathcal{K}} \left\{ \sum_{j \neq k} |a_{k,j}| + \sum_{j \neq k} |a_{j,k}| + b_k \right\} \quad (9)$$

is iteratively removed in the NLPD algorithm until (8) becomes true.

The complete description of the NLPD algorithm is given as follows.

Algorithm 1: The NLPD Algorithm

- Step 1. Initialization: Input data $(\mathbf{A}, \mathbf{b}, \bar{\mathbf{p}})$.
 - Step 2. Preprocessing: Remove link k_0 iteratively according to (9) until condition (8) holds true.
 - Step 3. Power control: Solve problem (5); 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 (7), set $\mathcal{K} = \mathcal{K} \setminus \{k_0\}$, and go to **Step 3**.
 - Step 5. Postprocessing: Check the removed links for possible admission.
-

III. A NON-CONVEX L_q APPROXIMATION DEFLATION APPROACH FOR JPAC

In this section, we develop a polynomial time non-convex ℓ_q ($0 < q < 1$) approximation deflation algorithm for the JPAC problem. The motivation for developing such a non-convex approximation deflation algorithm is that the ℓ_q minimization problem with $q \in (0, 1)$ should perform better than the ℓ_1 minimization problem in approximating the ℓ_0 minimization problem (3) and thus the deflation algorithm based on non-convex ℓ_q approximations should have a better performance than the NLPD algorithm, which is based on convex ℓ_1 approximations.

In the following, we first analyze exact recovery of the ℓ_q minimization approximation in solving the ℓ_0 minimization problem in Section III-A. Then, we prove in Section III-B that the ℓ_q minimization problem with any $q \in (0, 1)$ is NP-hard. In Section III-C, we develop a polynomial time interior-point algorithm for approximately solving the ℓ_q minimization problem. The approximate solution of the ℓ_q minimization problem can be used to check the simultaneous supportability of all links in the network and to guide an iterative link removal procedure, thus resulting in the polynomial time non-convex approximation deflation approach for the JPAC problem in Section III-D.

A. Exact Recovery of Non-Convex L_q Approximation

The sparse ℓ_0 minimization problem (3) is *successively* approximated by the ℓ_1 minimization problem (5) in the NLPD algorithm. Intuitively, the ℓ_q minimization problem with $0 < q < 1$,

$$\begin{aligned} \min_{\mathbf{x}} \quad & \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_q^q + \alpha \bar{\mathbf{p}}^T \mathbf{x} \\ \text{s.t.} \quad & \mathbf{0} \leq \mathbf{x} \leq \mathbf{e} \end{aligned} \quad (10)$$

should approximate (3) “better” than (5). To provide such an evidence, we give the following lemma.

Lemma 1: For any $q \in [0, 1]$, problem (10) is equivalent to

$$\begin{aligned} \min_{\mathbf{x}} \quad & \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_q^q + \alpha \bar{\mathbf{p}}^T \mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} \leq \mathbf{b}, \\ & \mathbf{0} \leq \mathbf{x} \leq \mathbf{e}. \end{aligned} \quad (11)$$

The above lemma can be verified in a similar way as the proof of Theorem 2 in [2] and a detailed proof is provided in [33, Section I]. Based on Lemma 1, we can show the following result (see Appendix A for its proof).

Theorem 1: For any given instance of problem (3), there exists $\bar{q} > 0$ (depending on $\mathbf{A}, \mathbf{b}, \alpha, \bar{\mathbf{p}}$) such that when $q \in (0, \bar{q}]$, any global solution to problem (10) is one of the global solutions to problem (3).

Theorem 1 states that the ℓ_q minimization problem (10) shares the same solution with the ℓ_0 minimization problem (3) if the parameter q (depending on $\mathbf{A}, \mathbf{b}, \alpha, \bar{\mathbf{p}}$) is chosen to be sufficiently small. In general, the ℓ_1 minimization problem (5) does not enjoy this exact recovery property, which is in sharp contrast to the results in [34]–[36].

It is shown in [34] that the problem of minimizing $\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_1$ is equivalent to the problem of minimizing $\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_0$ with high

probability if the vector $\mathbf{A}\mathbf{x} - \mathbf{b}$ at the true solution \mathbf{x}^* is sparse, where $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $m > n$, and if the entries of the matrix \mathbf{A} are independent and identically distributed (i.i.d.) Gaussian. The reason why the ℓ_1 minimization problem (10) fails to recover the solution of problem (3) is that the two assumptions required in [34] do not hold true for problem (3). Specifically, the vector $\mathbf{A}\mathbf{x} - \mathbf{b}$ may not be sparse even at the optimal power allocation vector \mathbf{x}^* . This depends on whether the (normalized) channel is strongly interfered or not. More importantly, the matrix \mathbf{A} in (3) is a square matrix and has a special structure; i.e., all diagonal entries are one and all non-diagonal entries are non-positive.

Next we give an example to illustrate the advantage of the use of ℓ_q norm with $0 < q < 1$ over the use of ℓ_1 norm to approximate problem (3). Suppose $\mathbf{A}, \mathbf{b}, \bar{\mathbf{p}}$ in (3) are given as follows:

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}, \quad \mathbf{b} = 0.5\mathbf{e}, \quad \bar{\mathbf{p}} = \mathbf{e}.$$

It can be shown that the optimal solution to the sparse optimization problem (3) is

$$\mathbf{x}^* = (0.5, 0.5, 0)^T$$

if the parameter α is chosen satisfying $0 < \alpha < 1/3$ (cf. (4)). We can also obtain the solutions to problems (5) and (10).

- By writing the KKT optimality conditions, we can check that $\mathbf{x} = 0$ is the unique global minimizer of problem (5) with any $\alpha \geq 0$.
- Lemma 1 implies that problem (10) is equivalent to

$$\begin{aligned} \min_{x_1, x_2, x_3} \quad & (0.5 - x_1 + x_3)^q + (0.5 - x_2 + x_3)^q \\ & + (0.5 + x_1 + x_2 - x_3)^q + \alpha(x_1 + x_2 + x_3) \\ \text{s.t.} \quad & 0.5 - x_1 + x_3 \geq 0, \\ & 0.5 - x_2 + x_3 \geq 0, \\ & 0.5 + x_1 + x_2 - x_3 \geq 0, \\ & 0 \leq x_1, x_2, x_3 \leq 1. \end{aligned} \quad (12)$$

For any given $q \in (0, 1)$, define

$$\bar{\alpha}_q := \min \{1 + (0.5)^q, 2^q\} - (1.5)^q > 0. \quad (13)$$

It can be checked (although tedious) that, as long as α in problem (12) is chosen such that $0 < \alpha \leq \bar{\alpha}_q$, the unique global minimizer of problem (12) is \mathbf{x}^* ; see [33, Section II].

We remark that, for a given instance of problem (3), it is generally not easy to determine \bar{q} in Theorem 1. However, our simulation results in Section IV-A show that it is generally not very small for small networks. In practice, we could set the parameter q in problem (10) to be a constant in $(0, 1)$ (more on the choice of the parameter q will be discussed in Section IV). Therefore, the solution to problem (10) might not be able to solve the ℓ_0 minimization problem (3). This is the reason why we do not just use the ℓ_q minimization (10) to approximate problem (3), but instead employ a deflation technique to successively approximate problem (3) in our proposed algorithm below.

B. Complexity Analysis of L_q Minimization (10)

Roughly speaking, convex optimization problems are relatively easy to solve, while non-convex optimization problems are difficult to solve. However, not all non-convex problems are computationally intractable since the lack of convexity may be due to an inappropriate formulation. In fact, many non-convex optimization problems admit a convex reformulation; see [37]–[44] for some examples. Therefore, convexity is useful but unreliable to evaluate the computational intractability of an optimization problem. A more robust tool is the computational complexity theory [45], [46].

In this subsection, we show that problem (10) is NP-hard for any given $q \in (0, 1)$. The NP-hardness proof is based on a polynomial time transformation from the partition problem: given a set of N positive integers s_1, s_2, \dots, s_N , determine whether there exists a subset \mathcal{S} of $\{1, 2, \dots, N\}$ such that

$$\sum_{n \in \mathcal{S}} s_n = \sum_{n \notin \mathcal{S}} s_n = \frac{1}{2} \sum_{n=1}^N s_n.$$

The partition problem is known to be NP-complete [45].

Theorem 2: For any given $0 < q < 1$, the ℓ_q minimization problem (10) is NP-hard.

The proof of Theorem 2 is relegated to Appendix B. Theorem 2 suggests that there is no efficient algorithm which can solve problem (10) to global optimality in polynomial time (unless $P = NP$), and finding an approximate solution for it is more realistic in practice.

We remark that some related problems

$$\begin{aligned} \min_{\mathbf{x}} \quad & \|\mathbf{x}\|_q^q \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \end{aligned} \quad (14)$$

and

$$\min_{\mathbf{x}} \quad \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_p^p + \lambda \|\mathbf{x}\|_q^q \quad (15)$$

are shown to be NP-hard in [47] and [48], where $p \in [1, \infty)$ and $q \in (0, 1)$. However, these NP-hardness results cannot imply our result in Theorem 2, since the considered problems are different. A key difference between our problem (10) and problems (14) and (15) is that problem (10) (equivalent to (11)) tries to find an \mathbf{x} such that the number of positive entries of the vector $\mathbf{b} - \mathbf{A}\mathbf{x}$ is as small as possible while problems (14) and (15) try to find a solution such that the number of nonzero entries of \mathbf{x} is as small as possible. Moreover, the matrix \mathbf{A} and the vector \mathbf{b} in (10) have special structures, i.e., all diagonal entries of \mathbf{A} are one, all non-diagonal entries of \mathbf{A} are non-positive, and all entries of \mathbf{b} are positive. This restriction on \mathbf{A} and \mathbf{b} makes it more technical and intricate to show the NP-hardness of problem (10) compared to show that of problems (14) and (15) for general \mathbf{A} and \mathbf{b} in [47] and [48]. The NP-hardness of a general composite ℓ_q minimization problem is also shown in [49]. Again, this result does not imply our result in Theorem 2, which is due to the special structure \mathbf{A} and \mathbf{b} in our problem (10).

C. A Polynomial Time Potential Reduction Algorithm for Problem (10)

In this subsection, we develop a polynomial time potential reduction interior-point algorithm for solving problem (10).

Based on Lemma 1, by introducing slack variables, we see that problem (10) can be equivalently formulated as

$$\begin{aligned} \min_{\mathbf{w}} \quad & f(\mathbf{w}) := \tilde{\mathbf{c}}^T \mathbf{w}_1 + \|\mathbf{w}_2\|_q^q \\ \text{s.t.} \quad & \tilde{\mathbf{A}}\mathbf{w} = \tilde{\mathbf{b}}, \\ & \mathbf{w} \geq \mathbf{0}, \end{aligned} \quad (16)$$

where

$$\begin{aligned} \tilde{\mathbf{A}} &= \begin{pmatrix} \mathbf{A} & \mathbf{I} & \mathbf{0} \\ \mathbf{I} & \mathbf{0} & \mathbf{I} \end{pmatrix} \in \mathbb{R}^{2K \times 3K}, \quad \tilde{\mathbf{b}} = \begin{pmatrix} \mathbf{b} \\ \mathbf{e} \end{pmatrix} \in \mathbb{R}^{2K}, \\ \tilde{\mathbf{c}} &= \alpha \bar{\mathbf{p}} \in \mathbb{R}^K, \quad \mathbf{w} = \begin{pmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \\ \mathbf{w}_3 \end{pmatrix} \in \mathbb{R}^{3K}. \end{aligned}$$

We extend the potential reduction algorithm in [47], [50], [51] to solve problem (16) to obtain one of its ϵ -KKT points (the definition of the ϵ -KKT point shall be given later). It can be shown that the potential reduction interior-point algorithm returns an ϵ -KKT point of problem (16) in polynomial time.

Before going into the details, we first give a high level preview of the proposed algorithm. The two basic ingredients of the potential reduction interior-point algorithm is the potential function (cf. (20)) and the update rule (cf. (25)). The potential function measures the progress of the algorithm, and the update rule guides to compute the next iterate based on the current one. More specifically, the next iterate is chosen as the feasible point that achieves the maximum potential reduction. The algorithm is terminated either when the potential function is below some threshold (cf. (21)) or when the potential reduction (cf. (26)) is smaller than a constant. In the former case, the algorithm returns an ϵ -optimal solution of problem (16), and in the latter case an ϵ -KKT point of problem (16). Moreover, the polynomial time convergence of the algorithm can be guaranteed by showing that the value of the potential function is decreased by at least a constant at each iteration and the potential function is bounded above and below by some thresholds.

Definition of ϵ -KKT Point: Since $f(\mathbf{w})$ is not differentiable when some entries of \mathbf{w}_2 are zero, the common definitions of the KKT point do not apply to problem (16). We thus define a weaker concept, the so-called ϵ -KKT point (or ϵ -KKT solution) for problem (16), in a similar way as done in [47], [50], [52]. Suppose that \mathbf{w}^* is a local minimizer of problem (16), and define $\mathcal{S} = \{j \mid [\mathbf{w}_2^*]_j > 0\}$. Then $\mathbf{w}_{\mathcal{S}}^* = (\mathbf{w}_1^*; [\mathbf{w}_2^*]_{\mathcal{S}}; \mathbf{w}_3^*)$ should be a local minimizer of problem

$$\begin{aligned} \min_{\mathbf{w}_{\mathcal{S}}} \quad & \tilde{\mathbf{c}}^T \mathbf{w}_1 + \|[\mathbf{w}_2]_{\mathcal{S}}\|_q^q \\ \text{s.t.} \quad & \tilde{\mathbf{A}}_{\mathcal{S}} \mathbf{w}_{\mathcal{S}} = \tilde{\mathbf{b}}, \\ & \mathbf{w}_{\mathcal{S}} \geq \mathbf{0}, \end{aligned} \quad (17)$$

where $\tilde{\mathbf{A}}_{\mathcal{S}} = \begin{pmatrix} \mathbf{A} & \mathbf{I}_{\mathcal{S}}^T & \mathbf{0} \\ \mathbf{I} & \mathbf{0} & \mathbf{I} \end{pmatrix} \in \mathbb{R}^{(2K) \times (2K + |\mathcal{S}|)}$. The KKT condition of problem (17) is: there exists a Lagrange multiplier vector $\boldsymbol{\lambda}^* \in \mathbb{R}^{2K}$ such that

$$\nabla_{\mathbf{w}_{\mathcal{S}}} f(\mathbf{w}^*) - \tilde{\mathbf{A}}_{\mathcal{S}}^T \boldsymbol{\lambda}^* \geq \mathbf{0} \quad (18)$$

and

$$(\nabla_{\mathbf{w}_{\mathcal{S}}} f(\mathbf{w}^*) - \tilde{\mathbf{A}}_{\mathcal{S}}^T \boldsymbol{\lambda}^*)^T \mathbf{w}_{\mathcal{S}}^* = 0.$$

Notice that given $\boldsymbol{\lambda}^*$, if $[\mathbf{w}^*]_n = 0$ for $n - K \notin \mathcal{S}$, we have

$$q[\mathbf{w}^*]_n^q - \left[\tilde{\mathbf{A}}^T \boldsymbol{\lambda}^* \right]_n [\mathbf{w}^*]_n = 0.$$

Therefore, the ϵ -KKT point of problem (16) can be defined as follows.

Definition 1: \mathbf{w}^* is called an ϵ -KKT point of problem (16) if

- (a) it is feasible;
- (b) there exists $\boldsymbol{\lambda}^*$ such that (18) holds true; and
- (c) the complementarity gap

$$\frac{\sum_{n=1}^{3K} \left(q[\mathbf{w}^*]_n^q - \left[\tilde{\mathbf{A}}^T \boldsymbol{\lambda}^* \right]_n [\mathbf{w}^*]_n \right)}{\bar{f} - \underline{f}} \leq \epsilon, \quad (19)$$

where \bar{f} and \underline{f} are upper and lower bounds on the objective value of problem (16), respectively.

In addition, \mathbf{w}^* is called an ϵ -optimal solution to problem (16) if $f(\mathbf{w}^*) \leq \epsilon$.

It is worthwhile remarking that if $\epsilon = 0$ in (19), the above definition reduces to the definition of the KKT point of problem (17). For simplicity, we set $\underline{f} = 0$ in (19) in this paper, since the objective function of problem (16) is always nonnegative.

Potential Function: For any given strictly feasible \mathbf{w} , define the following potential function

$$\phi(\mathbf{w}) = \rho \log(f(\mathbf{w})) - \sum_{k=1}^{3K} \log([\mathbf{w}]_k), \quad (20)$$

where ρ is a parameter to be specified later.

Lemma 2: Let $\epsilon > 0$ and $\rho > \frac{K}{q}$ be fixed. Suppose that \mathbf{w} is strictly feasible and satisfies

$$\phi(\mathbf{w}) \leq \underline{\phi} := \left(\rho - \frac{K}{q} \right) \log(\epsilon) + \frac{K}{q} \log(K) + K \log(4). \quad (21)$$

Then \mathbf{w} is an ϵ -optimal solution to problem (16).

Lemma 2 actually gives a lower bound of the potential function $\phi(\mathbf{w})$. Its proof can be found in Appendix C. In the following, we provide an upper bound for $\phi(\mathbf{w})$. Let

$$\mathbf{w}^0 = \begin{pmatrix} \frac{\min\{\mathbf{b}, \mathbf{e}\}}{2} \\ \mathbf{b} - \frac{\mathbf{A} \min\{\mathbf{b}, \mathbf{e}\}}{2} \\ \mathbf{e} - \frac{\min\{\mathbf{b}, \mathbf{e}\}}{2} \end{pmatrix}.$$

It can be verified that \mathbf{w}^0 in the above is an interior point of problem (16) by the use of the special structures of \mathbf{A} and \mathbf{b} . Since the potential function values are decreasing at each iteration \mathbf{w}^t of the potential reduction algorithm (see ‘‘Update Rule’’ further ahead), it follows that

$$\phi(\mathbf{w}^t) \leq \phi(\mathbf{w}^0), \quad \forall t \geq 0. \quad (22)$$

Update Rule: Consider one iteration update from \mathbf{w} to \mathbf{w}^+ by minimizing the potential reduction $\phi(\mathbf{w}^+) - \phi(\mathbf{w})$. Suppose that $\mathbf{w}^+ = \mathbf{w} + \mathbf{d} > \mathbf{0}$, where \mathbf{d} satisfies $\tilde{\mathbf{A}}\mathbf{d} = \mathbf{0}$. From the concavity of $\log(f(\mathbf{w}))$, we have

$$\log(f(\mathbf{w}^+)) - \log(f(\mathbf{w})) \leq \frac{1}{f(\mathbf{w})} \nabla f(\mathbf{w})^T \mathbf{d}. \quad (23)$$

On the other hand, we have the following standard lemma [53, Theorem 9.5].

Lemma 3: Let $\mathbf{W} = \text{Diag}(\mathbf{w})$. Suppose that $\|\mathbf{W}^{-1}\mathbf{d}\| \leq \beta < 1$. Then, we have

$$-\sum_{k=1}^{3K} \log([\mathbf{w}^+]_k) + \sum_{k=1}^{3K} \log([\mathbf{w}]_k) \leq -\mathbf{e}^T \mathbf{W}^{-1} \mathbf{d} + \frac{\beta^2}{2(1-\beta)}.$$

It is worthwhile remarking that if $\|\mathbf{W}^{-1}\mathbf{d}\| \leq \beta < 1$, then $\mathbf{w}^+ = \mathbf{w} + \mathbf{d} > \mathbf{0}$. By combining (23) and Lemma 3, we have

$$\phi(\mathbf{w}^+) - \phi(\mathbf{w}) \leq \left(\frac{\rho}{f(\mathbf{w})} \nabla f(\mathbf{w})^T \mathbf{W} - \mathbf{e}^T \right) \mathbf{W}^{-1} \mathbf{d} + \frac{\beta^2}{2(1-\beta)}. \quad (24)$$

Let $\tilde{\mathbf{d}} = \mathbf{W}^{-1}\mathbf{d}$. To achieve the maximum potential reduction, one can solve the following problem

$$\begin{aligned} \min_{\tilde{\mathbf{d}}} \quad & \mathbf{v}^T \tilde{\mathbf{d}} \\ \text{s.t.} \quad & \tilde{\mathbf{A}} \mathbf{W} \tilde{\mathbf{d}} = \mathbf{0}, \\ & \|\tilde{\mathbf{d}}\|^2 \leq \beta^2, \end{aligned} \quad (25)$$

where

$$\mathbf{v} = \frac{\rho}{f(\mathbf{w})} \mathbf{W} \nabla f(\mathbf{w}) - \mathbf{e}.$$

Problem (25) is simply a projection problem. The minimal value of problem (25) is

$$-\beta \|g(\mathbf{w})\|, \quad (26)$$

and the solution to problem (25) is $\tilde{\mathbf{d}} = \frac{\beta}{\|g(\mathbf{w})\|} g(\mathbf{w})$, where

$$g(\mathbf{w}) = \mathbf{e} - \frac{\rho}{f(\mathbf{w})} \mathbf{W} \left(\nabla f(\mathbf{w}) - \tilde{\mathbf{A}}^T \boldsymbol{\lambda} \right),$$

and

$$\boldsymbol{\lambda} = \left(\tilde{\mathbf{A}} \mathbf{W}^2 \tilde{\mathbf{A}}^T \right)^{-1} \tilde{\mathbf{A}} \mathbf{W} \left(\mathbf{W} \nabla f(\mathbf{w}) - \frac{f(\mathbf{w})}{\rho} \mathbf{e} \right).$$

To establish the polynomial time complexity[32] of the above potential reduction interior-point algorithm for solving problem (16), we make the following assumption.

Assumption 1: There exist $r \in (0, 1)$ and $R \in (1, \infty)$ such that the inputs of problem (16) (equivalent to problem (10)) satisfy $\mathbf{b} \geq r\mathbf{e}$ and $\max\{\|\mathbf{A}\|_\infty, \|\mathbf{b}\|_\infty, \|\alpha \bar{\mathbf{p}}\|_\infty\} \leq R$.

Theorem 3: Suppose Assumption 1 holds true. Then the potential reduction interior-point algorithm returns an ϵ -KKT point or an ϵ -optimal solution of problem (16) in no more than $O\left(\left(\frac{K}{\min\{\epsilon, q\}}\right) \log\left(\frac{K^{q+1}R}{\epsilon r}\right)\right)$ iterations.

The proof of Theorem 3 can be found in Appendix D. Note that for fixed $\epsilon > 0$ and $q \in (0, 1)$, the iteration complexity bound in Theorem 3 polynomially depends on the size and the bit length (number of bits) of the input data of problem (16). This shows the polynomial time complexity of the potential reduction interior-point algorithm for solving problem (16).

Since one projection problem in the form of (25) needs to be solved at each iteration of the potential reduction algorithm, and the complexity of solving problem (25) is $O(K^3)$, we immediately have the following corollary.

Corollary 1: Suppose Assumption 1 holds true. Then the potential reduction interior-point algorithm returns an ϵ -KKT point or ϵ -optimal solution of problem (16) (equivalent to problem (10)) in no more than $O\left(\left(\frac{K^4}{\min\{\epsilon, q\}}\right) \log\left(\frac{K^{q+1}R}{\epsilon r}\right)\right)$ operations.

One may ask why we restrict ourselves to use interior-point algorithms for solving problem (16) (equivalent to problem (10)). The reasons are the following. First, the objective function of problem (16) is *differentiable* in the interior feasible region. Moreover, we are actually interested in finding a feasible \mathbf{w} such that \mathbf{w}_2 is as sparse as possible; if we start from a \mathbf{w} , some entries of \mathbf{w}_2 are already zero, then it is very hard to make it nonzero. In contrast, if we start from an interior point, the interior-point algorithm may generate a sequence of interior points that bypasses solutions with the wrong zero supporting set and converges to the true one. This is exactly the idea of the interior-point algorithm developed in [50] for the non-convex quadratic programming.

To further improve the solution quality, we propose to run the above potential reduction algorithm multiple times to solve problem (16) and pick the best one among (potentially different) returned ϵ -KKT solutions, where at each time the potential reduction algorithm is initialized with a randomly generated interior point. Thanks to the special structure of \mathbf{A} and \mathbf{b} , it can be verified that the random point

$$\mathbf{w}(\boldsymbol{\xi}) = \begin{pmatrix} \boldsymbol{\xi} \circ \min\{\mathbf{b}, \mathbf{e}\} \\ \mathbf{b} - \mathbf{A}(\boldsymbol{\xi} \circ \min\{\mathbf{b}, \mathbf{e}\}) \\ \mathbf{e} - \boldsymbol{\xi} \circ \min\{\mathbf{b}, \mathbf{e}\} \end{pmatrix} \quad (27)$$

is an interior point to problem (16) with probability one, where each entry of $\boldsymbol{\xi}$ obeys the uniform distribution in the interval $[0, 1]$ and \circ is the Hadamard product operator.

D. A Polynomial Time Non-Convex Approximation Deflation Approach

The proposed ℓ_q minimization deflation (LQMD) algorithm, based on successive ℓ_q minimization approximations, is given as follows. The key difference between the LQMD algorithm and the NLPD algorithm lies in the power control step (i.e., **Step 3**), albeit the framework of the two algorithms are the same.

Algorithm 2: The LQMD Algorithm

- Step 1. Initialization: Input data $(\mathbf{A}, \mathbf{b}, \bar{\mathbf{p}})$, $q \in (0, 1)$, and positive integer N .
 - Step 2. Preprocessing: Remove link k_0 iteratively according to (9) until condition (8) holds true.
 - Step 3. Power control: Compute the parameter α by (28) and run the potential reduction algorithm with N randomly generated initial points (27) to solve problem (10); 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 (7), set $\mathcal{K} = \mathcal{K} \setminus \{k_0\}$, and go to **Step 3**.
 - Step 5. Postprocessing: Check the removed links for possible admission.
-

Two remarks on the LQMD algorithm are in order. First, the parameter α in ℓ_q minimization (10) is computed by

$$\alpha = \begin{cases} c_1 \alpha_1, & \text{if } \rho(\mathbf{I} - \mathbf{A}) \geq 1; \\ \min\{c_2 \alpha_1, c_3 \alpha_2\}, & \text{if } \rho(\mathbf{I} - \mathbf{A}) < 1, \end{cases} \quad (28)$$

where $c_3 > c_2$, $0 < c_1, c_2 < 1$ are three constants. In the above, α_1 is determined by the equivalence between problem (3) and the joint problem (1) and (2) (cf. (4)), and α_2 is determined by the so-called ‘‘Never-Over-Removal’’ property (cf. [2, (22)]). Second, the LQMD has a polynomial time worst-case complexity under Assumption 1, which is

$$O\left(\left(\frac{NK^5}{\min\{\epsilon, q\}}\right) \log\left(\frac{K^{q+1}R}{\epsilon r}\right)\right). \quad (29)$$

This is because that at most K links will be dropped and the complexity of dropping one link needs solving problem (10) N times in the LQMD algorithm. Combining this with Corollary 1, we immediately obtain the complexity result in (29).

IV. NUMERICAL SIMULATIONS

In this section, we carry out two sets of numerical experiments to illustrate the effectiveness of the non-convex ℓ_q approximation (10) and the LQMD algorithm (Algorithm 2), respectively. We employ the number of supported links and the total transmission power as the comparison metrics to compare different approximations and algorithms. In our simulations, the parameters in (28) are set to be $c_1 = c_2 = 0.2$ and $c_3 = 4$.

We generate the same channel parameters as in [1] in our simulations; i.e., each transmitter’s location obeys the uniform distribution over a 2 Km \times 2 Km square and the location of its corresponding receiver is uniformly generated in a disc with radius 400 m; channel gains are given by

$$g_{k,j} = \frac{1}{d_{k,j}^4}, \quad \forall k, j \in \mathcal{K}, \quad (30)$$

where $d_{k,j}$ is the Euclidean distance from the link of transmitter j to the link of receiver k . Each link’s SINR target is set to be $\gamma_k = 2$ dB ($\forall k \in \mathcal{K}$) and the noise power is set to be $\eta_k = -90$ dBm ($\forall k \in \mathcal{K}$). The power budget of the link of transmitter k is

$$\bar{p}_k = 2p_k^{\min}, \quad \forall k \in \mathcal{K}, \quad (31)$$

where p_k^{\min} is the minimum power needed for link k to meet its SINR requirement in the absence of any interference from other links.

Algorithm 3: A Heuristic Algorithm for Computing \bar{q}

- Step 1. Input data $(\mathbf{A}, \mathbf{b}, \bar{\mathbf{p}})$, positive integer N , and nonempty set \mathcal{Q} .
 - Step 2. Compute the parameter α by (28). Use the enumeration method to solve problem (3) and denote the global solution by \mathbf{x}^* .
 - Step 3. If \mathcal{Q} is empty, set $\bar{q} = 0$, return failure and terminate the algorithm; else pick the largest $q \in \mathcal{Q}$ and run the potential reduction algorithm with N randomly generated initial points (27) to solve problem (10). Denote the best point by \mathbf{x}_q^N .
 - Step 4. If $\mathbf{x}_q^N = \mathbf{x}^*$, set $\bar{q} = q$, return success and terminate the algorithm; else set $\mathcal{Q} = \mathcal{Q} \setminus \{q\}$ and go to **Step 3**.
-

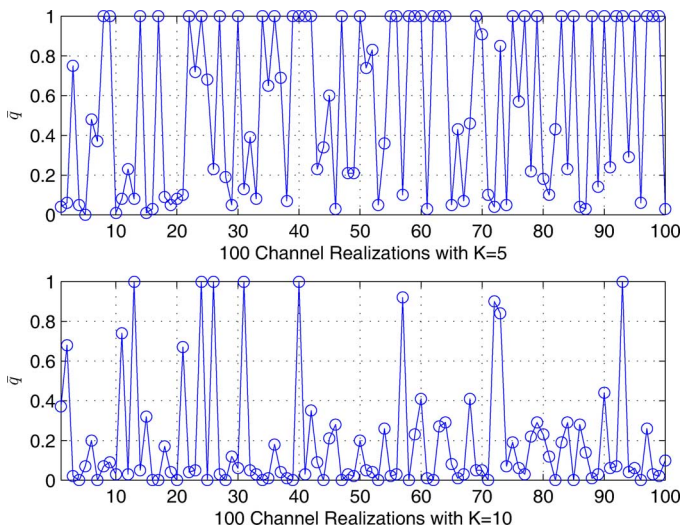


Fig. 1. The computed \bar{q} of 100 channel realizations with $K = 5$ and $K = 10$.

A. Non-Convex L_q Versus Convex L_1 Approximations

In this subsection, we do numerical simulations in two small networks where there are $K = 5$ and $K = 10$ links, respectively.

We first test how small q needs to be for the ℓ_q minimization problem (10) to exactly recover the global solution of problem (3). We propose Algorithm 3 to heuristically compute \bar{q} in Theorem 1. In Algorithm 3, the parameters N and Q are set to be

$$N = 100, Q = \{0.01, 0.02, \dots, 1\}. \quad (32)$$

Fig. 1 depicts the computed \bar{q} of 100 random channel realizations where $K = 5$ and $K = 10$. It shows that the parameter \bar{q} required in Theorem 1 is heavily problem-dependent. Fortunately, it is generally not very small, i.e., the average \bar{q} of the 100 random channel realizations for the two networks where $K = 5$ and $K = 10$ are 0.5364 and 0.1951, respectively. Fig. 1 also suggests that a smaller \bar{q} tends to be required as the number of total links in the network becomes large.

Among the above 100 random channel realizations, the ℓ_q minimization problem (10), with the parameter q being judiciously chosen by Algorithm 3, successfully finds the global solution of problem (3) 99 times and 82 times for the two networks where $K = 5$ and $K = 10$, respectively. This is consistent with our analysis in Theorem 1. The simulation results also show good performance of the potential reduction algorithm with multiple random initializations in finding the global solution of problem (10). As we can see, there are some instances that problem (10) fails to find the global solution of problem (3). The possible reasons are: a) the required \bar{q} in Theorem 1 for these instances might be less than 0.01, which is the smallest value we test in our simulations (cf. Q in (32)); and/or b) the potential reduction algorithm (even with 100 random initializations) does not find the global solution of problem (10), which is NP-hard as shown in Theorem 2.

We now compare the performance of the ℓ_q minimization problem (10) with q being fixed to be 0.1 and the ℓ_1 minimization problem (5) in approximating the ℓ_0 minimization problem (3). The corresponding ℓ_q minimization problem (10) is solved

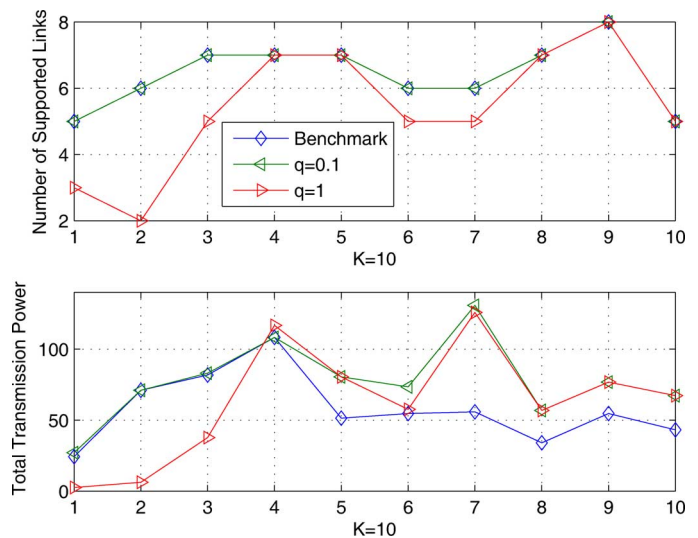


Fig. 2. Comparison of $\ell_{0.1}$ and ℓ_1 for 10 random channel realizations for the network where $K = 10$.

TABLE I
STATISTICS OF $L_{0.1}$ AND L_1 APPROXIMATIONS FOR 100
RANDOM CHANNEL REALIZATIONS

	$K = 5$			$K = 10$		
	Links	Power	Percentage	Links	Power	Percentage
Benchmark	4.05	43.71	100%	6.46	60.57	100%
$\ell_{0.1}$	4.05	51.36	69%	6.32	67.86	26%
ℓ_1	3.86	49.90	47%	5.28	51.94	3%

again by running the potential reduction algorithm with 100 random initializations and the corresponding ℓ_1 minimization problem (5) is solved by using the simplex method to solve its equivalent LP reformulation (6). The global solution obtained by “brute-force” enumeration is used as benchmark. Simulation results are summarized in Table I and Fig. 2. Table I reports the performance comparison of $\ell_{0.1}$ and ℓ_1 approximations in terms of average number of supported links, average total transmission power, and percentage of finding the global solution of problem (3) in the 100 random channel realizations. Fig. 2 illustrates the number of supported links and the total transmission power of 10 random channel realizations for the network where $K = 10$.

Table I shows that the $\ell_{0.1}$ minimization approximation significantly outperforms the ℓ_1 minimization approximation in terms of the number of supported links. It can be observed from Table I that the $\ell_{0.1}$ minimization problem can find the maximum admissible set for all of the 100 random channel realizations and successfully finds the maximum admissible set with minimum total transmission power with a percentage of 69% for the case $K = 5$. For the case $K = 10$, the percentage of the $\ell_{0.1}$ minimization problem for finding the ‘optimal’ maximum admissible set decreases to 26%, but still it performs much better than the ℓ_1 minimization problem in the sense that it supports 1.04 more links than ℓ_1 minimization in average; see Table I. Fig. 2 shows that the $\ell_{0.1}$ minimization problem finds the maximum admissible set for all these 10 channel realizations, and successfully finds the ‘optimal’ one for the first 4 channel realizations. Therefore, the ℓ_q minimization problem (10) indeed

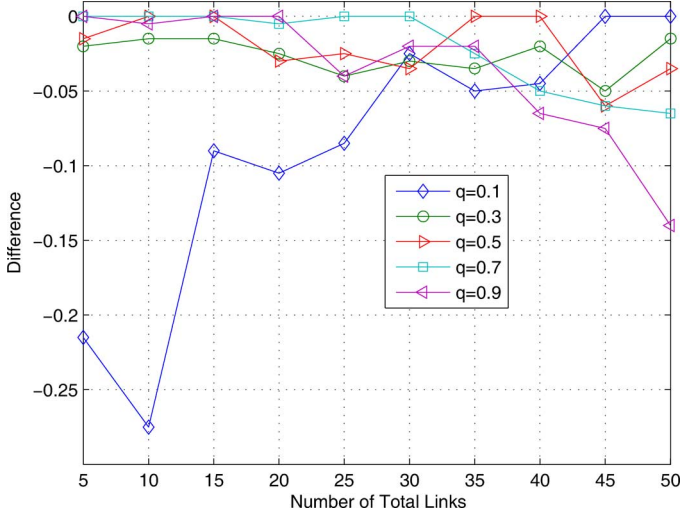


Fig. 3. The difference of average number of supported links by the LQMD algorithm with different q and the one by the LQMD algorithm with best $q \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$ versus the number of total links.

exhibits a significantly better capability in approximating the ℓ_0 minimization problem (3) compared to the ℓ_1 minimization problem (5).

B. Effectiveness of LQMD

We present some numerical simulation results to evaluate the effectiveness of the proposed LQMD algorithm in this part. We set $N = 5$ in the LQMD algorithm. All figures in this subsection are obtained by averaging over 200 Monte-Carlo runs.

We first test whether the performance of the LQMD algorithm is sensitive to the choice of the parameter q . Figs. 3 and 4 plot the performance comparison of the LQMD algorithm with different choices of the parameter q . More specifically, for each fixed K and $q \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$, we use the LQMD algorithm to solve 200 randomly generated JPAC problems, and denote the average number of supported links by $N(K, q)$. Each point in Fig. 3 denotes

$$D(K, q) := N(K, q) - \max_{q \in \{0.1, 0.3, 0.5, 0.7, 0.9\}} N(K, q).$$

It can be observed from Figs. 3 and 4 that the performance of the LQMD algorithm is somehow sensitive to the choice of the parameter q , which is mainly due to our implementation of the LQMD algorithm. Theoretically, the parameter q should be chosen as small as possible according to Theorem 1 (if the ℓ_q approximation problem can be solved to global optimality). However, as shown in Theorem 2, finding the global solution of the ℓ_q minimization problem is NP-hard for any $q \in (0, 1)$. In our implementation of the LQMD algorithm, we run the potential reduction algorithm with 5 random initializations to solve the ℓ_q minimization problem at affordable complexity. As can be seen from Fig. 3, the performance of the LQMD algorithm in terms of the number of supported links with $q = 0.1$ is not as good as expected. The main reason for this is because the ℓ_q minimization problem in the LQMD algorithm is solved by the potential reduction algorithm where the number of random initializations is set to be 5 and in this case the potential reduction algorithm will

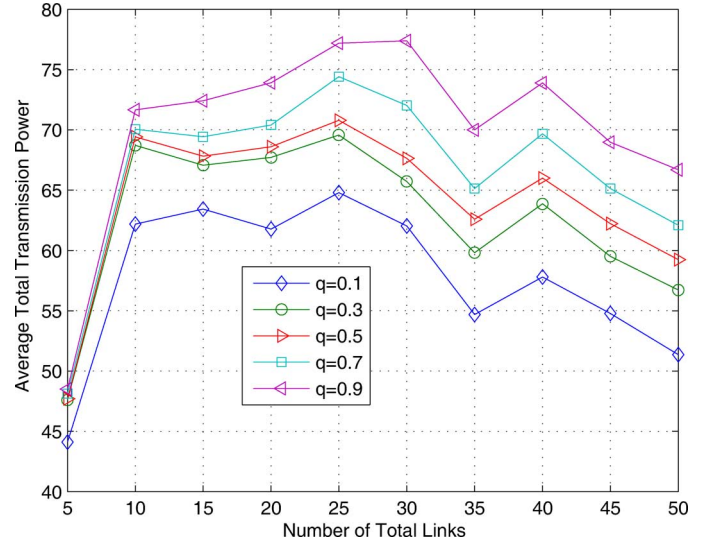


Fig. 4. Average total transmission power by the LQMD algorithm with different q versus the number of total links.

get stuck at a local minimizer of the ℓ_q minimization problem with a higher probability for a small q compared to a large q . One may set the number of random initializations to be very large, then the ℓ_q minimization problem with $q = 0.1$ will be solved to global optimality with a high probability and thus the corresponding LQMD algorithm will enjoy a good performance. However, this will lead to excessively high computational costs and make the corresponding LQMD algorithm impractical. On the other hand, a large q is also not suitable for the LQMD algorithm, since ℓ_q minimization with a large q cannot approximate the ℓ_0 minimization problem as good as the one with a small q . This can be clearly seen from Figs. 3 and 4, where the performance of the LQMD algorithm with $q = 0.7$ and $q = 0.9$ in terms of the number of supported links gradually deteriorate as the number of total links increases and the corresponding total transmission power is larger than that of $q = 0.1, 0.3, 0.5$ in the whole range.

The above simulation results and discussions provide us useful insights into the choice of the parameter q , i.e., both small and large q are not suitable for the LQMD algorithm due to either the practical implementation issue or the theoretical approximation issue and a median q is preferred in the LQMD algorithm in terms of leveraging the implementation issue and enjoying a relatively good approximation property. Figs. 3 and 4 suggest that $q = 0.3$ and $q = 0.5$ are the best ones in terms of the number of supported links and the total transmission power. We thus set $q = 0.5$ in all of the following simulations.

We now compare the performance of the proposed LQMD algorithm with that of the LPD algorithm in [1], the NLPD algorithm in [2], and the Algorithm II-B in [3], since all of them have been reported to have close-to-optimal performance in terms of the number of supported links. Figs. 5 – 6 plot the performance comparison of aforementioned various admission and power control algorithms. Fig. 5 shows that the proposed LQMD algorithm and the NLPD algorithm can support more links than the other two algorithms (the LPD algorithm and the Algorithm II-B) over the whole range of the tested number of total links.

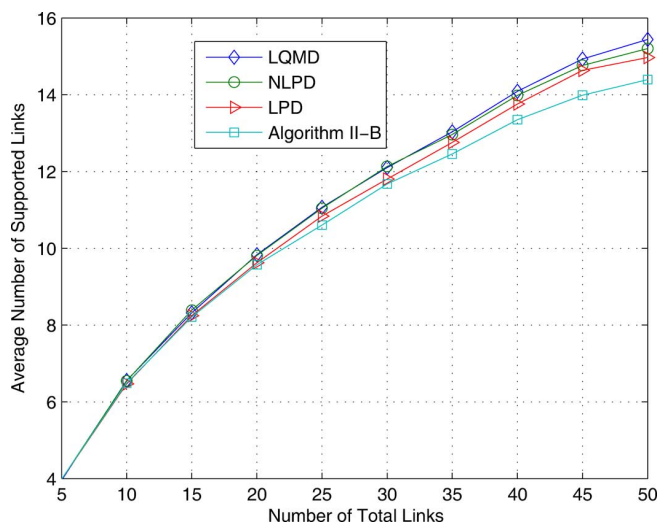


Fig. 5. Average number of supported links versus the number of total links.

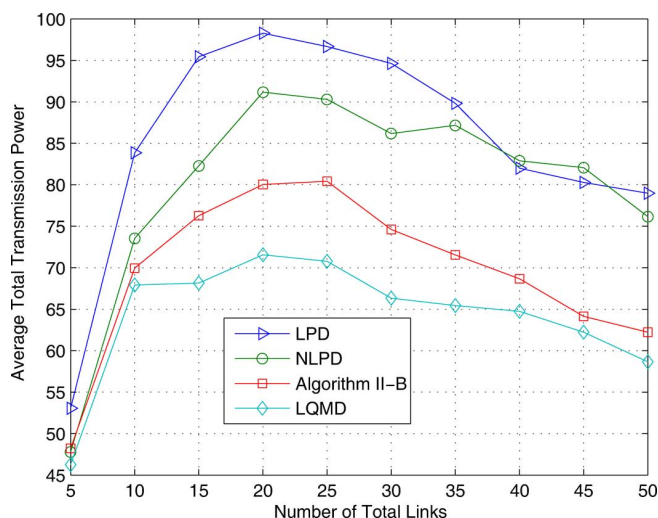


Fig. 6. Average total transmission power versus the number of total links.

Figs. 5 and 6 show that, compared to the LPD algorithm and the Algorithm II-B, the proposed LQMD algorithm can support more links with much less total transmission power.

Next, we focus on the performance comparison of the LQMD algorithm and the NLPD algorithm, since these two algorithms outperform the other two in terms of the number of supported links. The comparison results are presented in Figs. 7 and 8. The vertical axis “Win Ratio” in Fig. 7 shows the ratio of the number that the LQMD algorithm (the NLPD algorithm) winning the NLPD algorithm (the LQMD algorithm) to the total run number 200. Given an instance of the JPAC problem, the LQMD algorithm is said to win the NLPD algorithm if the former can support strictly more links than the latter for this instance. In a similar fashion, we can define that the NLPD algorithm wins the LQMD algorithm. It can be observed from Fig. 7 that the win ratio of the two algorithms are almost the same when the number of total links is less than or equal to 30,⁴ but the proposed LQMD algorithm wins the NLPD algorithm with a higher ratio when the number of total links is greater than 30. As depicted in Fig. 7, when there are $K = 50$ links in the network,

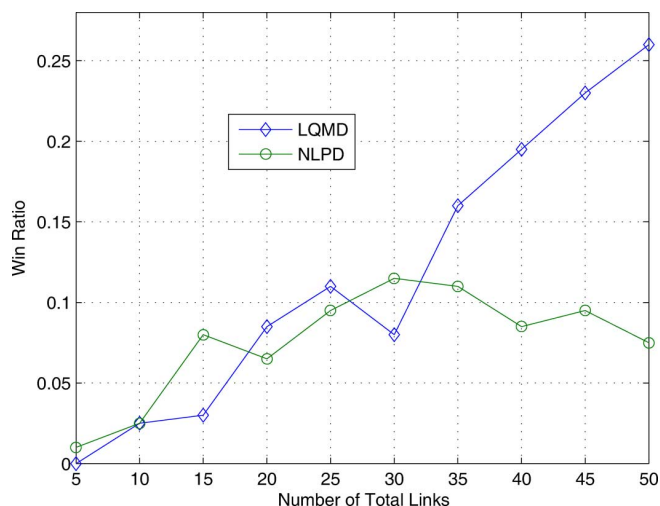


Fig. 7. Win ratio comparison of LQMD and NLPD versus the number of total links.

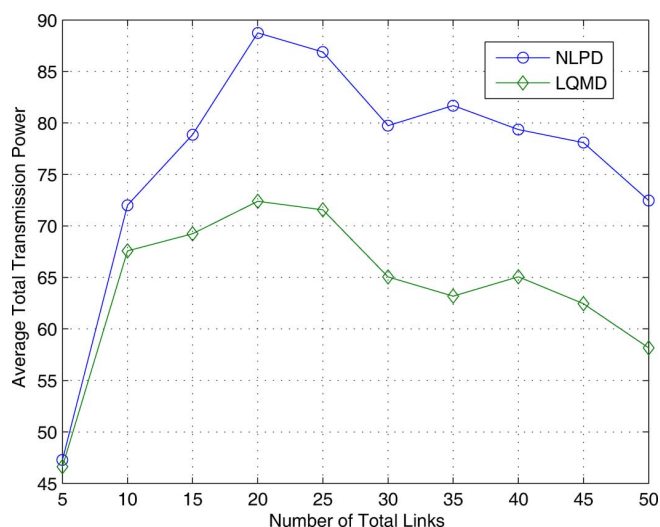


Fig. 8. Average total transmission power comparison of NLPD and LQMD when the two algorithms find the admissible set with same cardinality versus the number of total links.

the LQMD algorithm wins the NLPD algorithm 52 times, while the NLPD algorithm wins the LQMD algorithm only 15 times (among the total 200 runs). The two algorithms find the admissible set with same cardinality for the remaining 133 times. However, this does not mean that the two algorithms find the same admissible set in these cases. Fig. 8 plots the average total transmission power when the two algorithms can support the same number of links, which demonstrates that the LQMD algorithm is able to select a “better” subset of links to support, and can use much less total transmission power to support the same number of links (compared to the NLPD algorithm). As the number of total links in the network increases, the LQMD algorithm saves more power. In a nutshell, the LQMD algorithm

⁴In fact, it is impossible for the LQMD algorithm to achieve a large margin of the number of supported links over the NLPD algorithm for small networks, since it has been shown in [2] that the NLPD algorithm can achieve more than 98% of global optimality (by “brute force” enumeration) in terms of the number of supported links when $K \leq 18$.

TABLE II
RATIOS OF AVERAGE NUMBER OF SUPPORTED LINKS AND TOTAL TRANSMISSION POWER IN SETUP1 TO THAT IN SETUP2

Ratio \ K	5	10	15	20	25	30	35	40	45	50
Number of Supported Links	0.9875	1.0085	0.9946	0.9919	1.0189	1.0017	0.9992	0.9975	1.0000	0.9834
Total Transmission Power	3.6163	4.0054	4.0821	4.1320	3.9742	3.9698	3.9674	4.2082	4.2756	3.9979

exhibits a substantially better performance (than the NLPD algorithm) in selecting which subset of links to support, and thus yields a much better total transmission power performance.

Finally, we test the performance of the LQMD algorithm in setups with different levels of interference. For convenience, we call the former simulation setup as Setup1. We decrease the distances between all transmitters and receivers by a factor of 0.707 in Setup1 and call the obtained setup as Setup2. Notice that all the direct-link and cross-link channel gains and thus interference levels in Setup2 is $(1/0.707)^4 \approx 4$ times larger than that of Setup1 according to (30).

Table II summarizes the ratio of the average number of supported links and total transmission power in Setup1 to that in Setup2. It can be observed from it that the average number of supported links in both setups are roughly equal to each other, but the average total transmission power in Setup1 is approximately 4 times as large as that in Setup2. This is because when the simulation setup is switched from Setup1 to Setup2, all the channel gains increase by a factor of 4 and power budgets of all links decrease by a factor of 4 (cf. (31)). 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 (3) remains unchanged. However, it brings a benefit of a 75% reduction in the total transmission power, which is consistent with our engineering practice. Since the channel parameters in Setup1 and Setup2 are independently randomly generated, the ratios of average number of supported links and total transmission power in Setup1 to that in Setup2 in Table II are approximately (but not exactly) 1 and 4.

V. CONCLUSIONS

In this paper, we have proposed a polynomial time non-convex approximation deflation approach for the NP-hard joint power and admission control (JPAC) problem. Different from the existing convex approximation approaches, the proposed one solves the JPAC problem by successive non-convex ℓ_q ($0 < q < 1$) minimization approximations. We have shown exact recovery of the ℓ_q minimization problem, i.e., any global solution to the ℓ_q minimization problem is one of the global solutions to the JPAC problem as long as the parameter q is chosen to be sufficiently small. We have also developed a polynomial time potential reduction interior-point algorithm for solving the ℓ_q minimization problem, which makes the proposed deflation approach enjoy a polynomial time worst-case complexity. Numerical simulations demonstrate that the proposed approach is very effective, exhibiting a significantly better performance in selecting which subset of links to support compared to the existing convex approximation approaches.

APPENDIX A

PROOF OF THEOREM 1

To prove Theorem 1, we first introduce the following lemma.

Lemma 4 ([2]): Suppose \mathcal{S} is an admissible set⁵ of problem (3) and $\mathcal{S}^c = \mathcal{K} \setminus \mathcal{S}$ is its complement. Then the following statements hold.

- (1) For any admissible set \mathcal{S} , $\mathbf{A}_{\mathcal{S},\mathcal{S}}$ is invertible and $(\mathbf{A}_{\mathcal{S},\mathcal{S}})^{-1} \geq \mathbf{0}$.
- (2) For any admissible set \mathcal{S} and any feasible \mathbf{x} satisfying $[\mathbf{A}\mathbf{x} - \mathbf{b}]_{\mathcal{S}} \geq \mathbf{0}$,

$$\mathbf{x}_{\mathcal{S}} \geq (\mathbf{A}_{\mathcal{S},\mathcal{S}})^{-1} \mathbf{b}_{\mathcal{S}} > \mathbf{0}. \quad (33)$$

- (3) Let $\bar{\mathcal{S}}$ be the optimal maximum admissible set of problem (3). Then \mathbf{x}^* with

$$\mathbf{x}_{\bar{\mathcal{S}}}^* = (\mathbf{A}_{\bar{\mathcal{S}},\bar{\mathcal{S}}})^{-1} \mathbf{b}_{\bar{\mathcal{S}}}, \mathbf{x}_{\bar{\mathcal{S}}^c}^* = \mathbf{0} \quad (34)$$

is the solution to problem (3).

We are now ready to prove Theorem 1. We first prove that the theorem is true under the assumptions that the solution of problem (3) is unique and the maximum admissible set of problem (3) is also unique. Then, we remove these two assumptions and prove that the theorem remains true.

Case 1: Assume \mathbf{x}^* in (34) is the unique global minimizer of problem (3) and $\bar{\mathcal{S}}$ is the corresponding maximum admissible set, which is also unique. Next, we show \mathbf{x}^* is the unique global minimizer of problem (10). By Lemma 1, it is equivalent to show \mathbf{x}^* is the unique global minimizer of problem (11). We divide the proof into two parts.

Part A: In this part, we show that when q is sufficiently small, \mathbf{x}^* is the unique global minimizer of problem

$$\begin{aligned} \min_{\mathbf{x}} \quad & \|[\mathbf{b} - \mathbf{A}\mathbf{x}]_{\bar{\mathcal{S}}^c}\|_q^q + \alpha \bar{\mathbf{p}}^T \mathbf{x} \\ \text{s.t.} \quad & [\mathbf{b} - \mathbf{A}\mathbf{x}]_{\bar{\mathcal{S}}} = \mathbf{0}, \\ & [\mathbf{b} - \mathbf{A}\mathbf{x}]_{\bar{\mathcal{S}}^c} > \mathbf{0}, \\ & \mathbf{0} \leq \mathbf{x} \leq \mathbf{e}. \end{aligned} \quad (35)$$

Consider the following problem

$$\begin{aligned} \min_{\mathbf{x}} \quad & \min \{[\mathbf{b} - \mathbf{A}\mathbf{x}]_{\bar{\mathcal{S}}^c}\} \\ \text{s.t.} \quad & [\mathbf{b} - \mathbf{A}\mathbf{x}]_{\bar{\mathcal{S}}} = \mathbf{0}, \\ & [\mathbf{b} - \mathbf{A}\mathbf{x}]_{\bar{\mathcal{S}}^c} \geq \mathbf{0}, \\ & \mathbf{0} \leq \mathbf{x} \leq \mathbf{e}. \end{aligned} \quad (36)$$

We claim that the optimal value of problem (36) is greater than or equal to $\underline{\delta} := \delta(\mathbf{A}, \mathbf{b}) > 0$. Otherwise, there must exist a feasible point \mathbf{x} of problem (36) such that

$$[\mathbf{b} - \mathbf{A}\mathbf{x}]_j = 0, \quad \forall j \in \bar{\mathcal{S}} \cup k,$$

where $k \in \bar{\mathcal{S}}^c$. This contradicts the fact that $\bar{\mathcal{S}}$ is the maximum admissible set.

⁵The subset \mathcal{S} is called admissible if there exists a feasible \mathbf{x} such that $[\mathbf{A}\mathbf{x} - \mathbf{b}]_{\mathcal{S}} \geq \mathbf{0}$.

Suppose $\mathbf{g}(\mathbf{x})$ is the gradient of the objective function in (35). Then

$$\begin{aligned}\mathbf{g}(\mathbf{x}) &= \alpha \bar{\mathbf{p}} - q (\mathbf{A}_{\bar{\mathcal{S}}^c, \bar{\mathcal{S}}}, \mathbf{A}_{\bar{\mathcal{S}}^c, \bar{\mathcal{S}}^c})^T [\mathbf{b} - \mathbf{A}\mathbf{x}]_{\bar{\mathcal{S}}^c}^{q-1} \\ &\geq \alpha \bar{\mathbf{p}} - q \underline{\delta}^{q-1} (|\mathbf{A}_{\bar{\mathcal{S}}^c, \bar{\mathcal{S}}}|, |\mathbf{A}_{\bar{\mathcal{S}}^c, \bar{\mathcal{S}}^c}|)^T \mathbf{e},\end{aligned}$$

where $|\mathbf{A}|$ denotes the entry-wise absolute value of the matrix \mathbf{A} . If q is sufficiently small (say, $q \leq \bar{q}_1(\mathbf{A}, \mathbf{b}, \alpha, \bar{\mathbf{p}})$, where $\bar{q}_1(\mathbf{A}, \mathbf{b}, \alpha, \bar{\mathbf{p}})$ is a positive number such that for any $q \in (0, \bar{q}_1(\mathbf{A}, \mathbf{b}, \alpha, \bar{\mathbf{p}})]$, we have $\alpha \bar{\mathbf{p}} - q \underline{\delta}^{q-1} (|\mathbf{A}_{\bar{\mathcal{S}}^c, \bar{\mathcal{S}}}|, |\mathbf{A}_{\bar{\mathcal{S}}^c, \bar{\mathcal{S}}^c}|)^T \mathbf{e} > \mathbf{0}$), then the gradient $\mathbf{g}(\mathbf{x})$ of the objective function in (35) is component-wise positive at any feasible point. In addition, to guarantee $[\mathbf{b} - \mathbf{A}\mathbf{x}]_{\bar{\mathcal{S}}} = \mathbf{0}$, we need $\mathbf{x}_{\bar{\mathcal{S}}} \geq \mathbf{x}_{\bar{\mathcal{S}}}^* = \mathbf{A}_{\bar{\mathcal{S}}, \bar{\mathcal{S}}}^{-1} \mathbf{b}_{\bar{\mathcal{S}}}$ (cf. (33) and (34)). Therefore, \mathbf{x}^* is the unique global minimizer of problem (35), since for any feasible \mathbf{x} , we have $\mathbf{x}_{\bar{\mathcal{S}}} \geq \mathbf{x}_{\bar{\mathcal{S}}}^*$ and $\mathbf{x}_{\bar{\mathcal{S}}^c} \geq \mathbf{0}$. *Part B:* In this part, we show that when q is sufficiently small, \mathbf{x}^* is the unique global minimizer of problem (11). To show this, it suffices to show that, for any given admissible set $\mathcal{S} \subset \mathcal{K}$ with $|\mathcal{S}| < |\bar{\mathcal{S}}|$, the minimum value of problem

$$\begin{aligned}\min_{\mathbf{x}} \quad & \|[\mathbf{b} - \mathbf{A}\mathbf{x}]_{\mathcal{S}^c}\|_q^q + \alpha \bar{\mathbf{p}}^T \mathbf{x} \\ \text{s.t.} \quad & [\mathbf{b} - \mathbf{A}\mathbf{x}]_{\mathcal{S}} = 0, \\ & [\mathbf{b} - \mathbf{A}\mathbf{x}]_{\mathcal{S}^c} > 0, \\ & \mathbf{0} \leq \mathbf{x} \leq \mathbf{e},\end{aligned}\quad (37)$$

is greater than the one of problem (35). Without loss of generality, suppose the solution $\mathbf{x}^*(\mathcal{S})$ of problem (37) is attainable. Otherwise, there must exist $k \in \mathcal{S}^c$ such that $[\mathbf{b} - \mathbf{A}\mathbf{x}]_k \rightarrow 0$ at the optimal point. In this case, we consider problem (37) with \mathcal{S} replaced by $\mathcal{S} \cup k$. According to the assumption that the maximum admissible set of problem (3) is unique, we still have $|\mathcal{S} \cup k| < |\bar{\mathcal{S}}|$ unless $\mathcal{S} \cup k = \bar{\mathcal{S}}$.

Suppose $\mathbf{x}^*(\mathcal{S})$ is achievable, there must exist $\delta_{\mathcal{S}} > 0$ such that

$$[\mathbf{b} - \mathbf{A}\mathbf{x}^*(\mathcal{S})]_j \geq \delta_{\mathcal{S}} > 0, \quad \forall j \in \mathcal{S}^c, \quad (38)$$

where $\delta_{\mathcal{S}}$ depends on $\mathbf{A}, \mathbf{b}, \alpha, \bar{\mathbf{p}}, \mathcal{S}$. Since the number of admissible sets \mathcal{S} , with which the solution $\mathbf{x}^*(\mathcal{S})$ of problem (37) is achievable, is finite, then

$$\delta := \min \left\{ \min_{\mathcal{S} \subset \mathcal{K}} \{\delta_{\mathcal{S}}\}, \min \{\mathbf{b}/2\} \right\} > 0. \quad (39)$$

Here, δ only depends on $\mathbf{A}, \mathbf{b}, \alpha$ and $\bar{\mathbf{p}}$. Define

$$\Delta := \max_{\mathbf{0} \leq \mathbf{x} \leq \mathbf{e}, \mathbf{b} - \mathbf{A}\mathbf{x} \geq \mathbf{0}} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_{\infty}. \quad (40)$$

Let $\bar{q}_2(\mathbf{A}, \mathbf{b}, \alpha, \bar{\mathbf{p}})$ be a positive number such that for any $q \in (0, \bar{q}_2(\mathbf{A}, \mathbf{b}, \alpha, \bar{\mathbf{p}})]$, we have

$$(K - |\bar{\mathcal{S}}| + 1) \delta^q \geq (K - |\bar{\mathcal{S}}|) \Delta^q + \alpha \bar{\mathbf{p}}^T \mathbf{x}^*. \quad (41)$$

Therefore, if

$$0 < q \leq \bar{q} := \min \{\bar{q}_1(\mathbf{A}, \mathbf{b}, \alpha, \bar{\mathbf{p}}), \bar{q}_2(\mathbf{A}, \mathbf{b}, \alpha, \bar{\mathbf{p}})\}, \quad (42)$$

for any admissible \mathcal{S} with $|\mathcal{S}| < |\bar{\mathcal{S}}|$:

- if $\mathbf{x}^*(\mathcal{S}) \neq \mathbf{0}$, there holds

$$\begin{aligned}\|[\mathbf{b} - \mathbf{A}\mathbf{x}^*(\mathcal{S})]\|_q^q + \alpha \bar{\mathbf{p}}^T \mathbf{x}^*(\mathcal{S}) &> (K - |\mathcal{S}|) \delta^q\end{aligned}\quad (43)$$

$$\geq (K - |\bar{\mathcal{S}}| + 1) \delta^q \quad (44)$$

$$\geq (K - |\bar{\mathcal{S}}|) \Delta^q + \alpha \bar{\mathbf{p}}^T \mathbf{x}^* \quad (45)$$

$$\geq \|\mathbf{b} - \mathbf{A}\mathbf{x}^*\|_q^q + \alpha \bar{\mathbf{p}}^T \mathbf{x}^*, \quad (46)$$

where (43) is due to (38), (39), and $\mathbf{x}^*(\mathcal{S}) \neq \mathbf{0}$, (44) is due to the fact $|\mathcal{S}| < |\bar{\mathcal{S}}|$, (45) is due to (41) and (42), and (46) is by the definition of Δ (cf. (40));

- if $\mathbf{x}^*(\mathcal{S}) = \mathbf{0}$, then

$$\begin{aligned}\|\mathbf{b} - \mathbf{A}\mathbf{x}^*(\mathcal{S})\|_q^q + \alpha \bar{\mathbf{p}}^T \mathbf{x}^*(\mathcal{S}) &= \|\mathbf{b}\|_q^q > K \delta^q \geq \|\mathbf{b} - \mathbf{A}\mathbf{x}^*\|_q^q + \alpha \bar{\mathbf{p}}^T \mathbf{x}^*,\end{aligned}$$

where the first strict inequality is due to (39), and the second inequality can be obtained in a similar fashion as in the case of $\mathbf{x}^*(\mathcal{S}) \neq \mathbf{0}$.

Case II: Consider the case when problem (3) has multiple maximum admissible sets, but its solution \mathbf{x}^* remains unique. Then, for any feasible set $\mathcal{S} \neq \bar{\mathcal{S}}$ such that $|\mathcal{S}| = |\bar{\mathcal{S}}|$, we have

$$\|\mathbf{A}\mathbf{x}^*(\mathcal{S}) - \mathbf{b}\|_0 + \alpha \bar{\mathbf{p}}^T \mathbf{x}^*(\mathcal{S}) > \|\mathbf{A}\mathbf{x}^* - \mathbf{b}\|_0 + \alpha \bar{\mathbf{p}}^T \mathbf{x}^*,$$

where $\mathbf{x}^*(\mathcal{S})$ is the solution to problem (37). The above strict inequality is because \mathbf{x}^* is the unique solution to problem (3). Therefore, there exists $\bar{q}(\mathbf{A}, \mathbf{b}, \alpha, \bar{\mathbf{p}}, \mathcal{S}) > 0$ such that for all $q \in (0, \bar{q}(\mathbf{A}, \mathbf{b}, \alpha, \bar{\mathbf{p}}, \mathcal{S})]$, there holds

$$\|\mathbf{A}\mathbf{x}^*(\mathcal{S}) - \mathbf{b}\|_q^q + \alpha \bar{\mathbf{p}}^T \mathbf{x}^*(\mathcal{S}) > \|\mathbf{A}\mathbf{x}^* - \mathbf{b}\|_q^q + \alpha \bar{\mathbf{p}}^T \mathbf{x}^*.$$

Since problem (3) has at most $\binom{K}{|\bar{\mathcal{S}}|}$ maximum admissible sets, we can take the minimum among $\{\bar{q}(\mathbf{A}, \mathbf{b}, \alpha, \bar{\mathbf{p}}, \mathcal{S})\}$, and obtain a $\bar{q}(\mathbf{A}, \mathbf{b}, \alpha, \bar{\mathbf{p}}) > 0$ such that when $q \in (0, \bar{q}(\mathbf{A}, \mathbf{b}, \alpha, \bar{\mathbf{p}})]$, it has

$$\begin{aligned}\|\mathbf{A}\mathbf{x}^*(\mathcal{S}) - \mathbf{b}\|_q^q + \alpha \bar{\mathbf{p}}^T \mathbf{x}^*(\mathcal{S}) &> \|\mathbf{A}\mathbf{x}^* - \mathbf{b}\|_q^q + \alpha \bar{\mathbf{p}}^T \mathbf{x}^*, \quad \forall \mathcal{S} \neq \mathcal{S}^*, |\mathcal{S}| = |\mathcal{S}^*|.\end{aligned}$$

This, together with Case I, implies that when q is sufficiently small, \mathbf{x}^* is the unique global minimizer of problem (11) under the assumption that the solution \mathbf{x}^* of problem (3) is unique.

Case III: The remaining case is that when problem (3) has multiple solutions. Without loss of generality, we assume that there are two different solutions \mathbf{x}_1^* and \mathbf{x}_2^* . Then, the choice of α (cf. (4)) immediately implies

$$\|\mathbf{A}\mathbf{x}_1^* - \mathbf{b}\|_0 = \|\mathbf{A}\mathbf{x}_2^* - \mathbf{b}\|_0, \quad \bar{\mathbf{p}}^T \mathbf{x}_1^* = \bar{\mathbf{p}}^T \mathbf{x}_2^*.$$

If there exists an bijective mapping π from $\{1, 2, \dots, K\}$ to $\{1, 2, \dots, K\}$ such that $[\mathbf{A}\mathbf{x}_1^* - \mathbf{b}]_j = [\mathbf{A}\mathbf{x}_2^* - \mathbf{b}]_{\pi(j)}$ for all $j \in \mathcal{K}$, then both \mathbf{x}_1^* and \mathbf{x}_2^* are global minimizers of problem (11). Otherwise, we can find $\bar{q}(\mathbf{A}, \mathbf{b}, \alpha, \bar{\mathbf{p}}) > 0$ such that when $q \in (0, \bar{q}(\mathbf{A}, \mathbf{b}, \alpha, \bar{\mathbf{p}})]$, we have either

$$\|\mathbf{A}\mathbf{x}_1^* - \mathbf{b}\|_q^q + \alpha \bar{\mathbf{p}}^T \mathbf{x}_1^* < \|\mathbf{A}\mathbf{x}_2^* - \mathbf{b}\|_q^q + \alpha \bar{\mathbf{p}}^T \mathbf{x}_2^*$$

or

$$\|\mathbf{A}\mathbf{x}_1^* - \mathbf{b}\|_q^q + \alpha \bar{\mathbf{p}}^T \mathbf{x}_1^* > \|\mathbf{A}\mathbf{x}_2^* - \mathbf{b}\|_q^q + \alpha \bar{\mathbf{p}}^T \mathbf{x}_2^*.$$

Combining the above with Cases I and II, we know that \mathbf{x}_1^* (or \mathbf{x}_2^*) is the global minimizer of problem (11). This completes the proof of Theorem 1.

APPENDIX B PROOF OF THEOREM 2

Given an instance of the partition problem with s_1, s_2, \dots, s_N , define $S = \sum_{n=1}^N s_n$. Next, we construct an instance of problem (10), where

- $K = 2N + 2$;
- all entries of $\bar{\mathbf{p}} \in \mathbb{R}^{2N+2}$ are set to be 1;
- the first $2N$ entries of $\mathbf{b} \in \mathbb{R}^{2N+2}$ are set to be 1, and the last two entries of \mathbf{b} are set to be 0.5;
- all diagonal entries of $\mathbf{A} \in \mathbb{R}^{(2N+2) \times (2N+2)}$ are one, and non-diagonal entries of \mathbf{A} are
 - for $k = 2n - 1$, $n = 1, 2, \dots, N$, set $a_{k,j} = 0$ except $a_{k,k+1} = -1$;
 - for $k = 2n$, $n = 1, 2, \dots, N$, set $a_{k,j} = 0$ except $a_{k,k-1} = -1$;
 - for $k = 2N + 1$, set $a_{k,j} = 0$ except $a_{k,2n-1} = -\frac{s_n}{S}$ for $n = 1, 2, \dots, N$;
 - for $k = 2N + 2$, set $a_{k,j} = 0$ except $a_{k,2n} = -\frac{s_n}{S}$ for $n = 1, 2, \dots, N$; and
- the parameter α satisfies

$$0 < \alpha \leq \frac{2 - 2^q}{2}. \quad (47)$$

Then the constructed instance of problem (10) becomes problem (48).

$$\begin{aligned} \min_{x_1, \dots, x_{2N+2}} & \sum_{n=1}^N F(x_{2n-1}, x_{2n}) \\ & + \sum_{i=0}^1 H(x_{1+i}, x_{3+i}, \dots, x_{2N+1+i}) \\ \text{s.t.} & \quad 0 \leq x_n \leq 1, \quad n = 1, 2, \dots, 2N + 2, \end{aligned} \quad (48)$$

where

$$F(x_{2n-1}, x_{2n}) = |1 + x_{2n} - x_{2n-1}|^q + |1 + x_{2n-1} - x_{2n}|^q + \alpha(x_{2n-1} + x_{2n}), \quad n = 1, 2, \dots, N,$$

$$\begin{aligned} H(x_{1+i}, x_{3+i}, \dots, x_{2N+1+i}) \\ = \left| 0.5 + \frac{1}{S} \sum_{n=1}^N s_n x_{2n-1+i} - x_{2N+1+i} \right|^q \\ + \alpha x_{2N+1+i}, \quad i = 0, 1. \end{aligned}$$

Notice that $0 \leq x_{2n-1}, x_{2n} \leq 1$ for all $n = 1, 2, \dots, N$, it follows that

$$F(x_{2n-1}, x_{2n}) = (1 + x_{2n} - x_{2n-1})^q + (1 + x_{2n-1} - x_{2n})^q + \alpha(x_{2n-1} + x_{2n}), \quad n = 1, 2, \dots, N. \quad (49)$$

Next, we claim that the partition problem has a “yes” answer if and only if the optimal value of problem (48) is less than or equal to $2^q N + \alpha(N + 2)$. We prove the “if” and “only if” directions separately.

Let us first prove the “only if” direction. Suppose the partition problem has a “yes” answer and let \mathcal{S} be the subset of $\{1, 2, \dots, N\}$ such that

$$\sum_{n \in \mathcal{S}} s_n = S/2. \quad (50)$$

We show that there exists a feasible power allocation vector $\{x_n\}_{n=1}^{2N+2}$ such that the optimal value of problem (48) is less than or equal to $2^q N + \alpha(N + 2)$. In particular, let

$$\hat{x}_{2n-1} = \begin{cases} 1, & \text{if } n \in \mathcal{S}; \\ 0, & \text{if } n \notin \mathcal{S}, \end{cases} \quad \hat{x}_{2n} = 1 - \hat{x}_{2n-1}, \quad n = 1, 2, \dots, N, \quad (51)$$

and

$$\hat{x}_{2N+1} = \hat{x}_{2N+2} = 1.$$

It is simple to check

$$\begin{aligned} F(\hat{x}_{2n-1}, \hat{x}_{2n}) & = 2^q + \alpha, \quad n = 1, 2, \dots, N, \\ H(\hat{x}_1, \hat{x}_3, \dots, \hat{x}_{2N+1}) & = \left| 0.5 + \frac{1}{S} \sum_{n=1}^N s_n \hat{x}_{2n-1} - \hat{x}_{2N+1} \right|^q + \alpha \hat{x}_{2N+1} \\ & = \left| 0.5 + \frac{1}{S} \sum_{n \in \mathcal{S}} s_n - 1 \right|^q + \alpha \quad (\text{from (51)}) \\ & = \alpha, \quad (\text{from (50)}) \end{aligned}$$

$$\begin{aligned} H(\hat{x}_2, \hat{x}_4, \dots, \hat{x}_{2N+2}) & = \alpha. \end{aligned}$$

Thus, we have

$$\begin{aligned} \sum_{n=1}^N F(\hat{x}_{2n-1}, \hat{x}_{2n}) + \sum_{i=0}^1 H(\hat{x}_{1+i}, \hat{x}_{3+i}, \dots, \hat{x}_{2N+1+i}) \\ = N2^q + (N + 2)\alpha, \end{aligned}$$

which implies that the optimal value of problem (48) is less than or equal to $N2^q + (N + 2)\alpha$.

To show the “if” direction, suppose that the optimal solution of problem (48) is less than or equal to $N2^q + (N + 2)\alpha$. Consider a relaxation of problem (48) by dropping the constraints $x_{2N+1} \leq 1$ and $x_{2N+2} \leq 1$:

$$\begin{aligned} \min_{x_1, \dots, x_{2N+2}} & \sum_{n=1}^N F(x_{2n-1}, x_{2n}) \\ & + \sum_{i=0}^1 H(x_{1+i}, x_{3+i}, \dots, x_{2N+1+i}) \\ \text{s.t.} & \quad 0 \leq x_n \leq 1, \quad n = 1, 2, \dots, 2N, \\ & \quad x_{2N+1} \geq 0, \quad x_{2N+2} \geq 0. \end{aligned} \quad (52)$$

Clearly, the optimal value of problem (52) is less than or equal to the optimal value of problem (48). The relaxed problem (52) can be equivalently rewritten as

$$\begin{aligned} \min_{x_1, \dots, x_{2N+2}} & \sum_{n=1}^N F(x_{2n-1}, x_{2n}) \\ & + \sum_{i=0}^1 \hat{H}(x_{1+i}, x_{3+i}, \dots, x_{2N-1+i}) \\ \text{s.t.} & \quad 0 \leq x_n \leq 1, \quad n = 1, 2, \dots, 2N, \end{aligned}$$

where, for $i = 0, 1$, $\hat{H}(x_{1+i}, x_{3+i}, \dots, x_{2N-1+i})$ is defined as the optimal value of the problem

$$\begin{aligned} \min_{x_{2N+1+i}} \quad & H(x_{1+i}, x_{3+i}, \dots, x_{2N-1+i}, x_{2N+1+i}) \\ \text{s.t.} \quad & x_{2N+1+i} \geq 0. \end{aligned} \quad (53)$$

Since problem (53) is an univariate optimization problem and α satisfies (47), we can verify that, for any $x_{1+i}, x_{3+i}, \dots, x_{2N+1+i} \geq 0$, there holds

$$\begin{aligned} H(x_{1+i}, x_{3+i}, \dots, x_{2N+1+i}) &\geq H(x_{1+i}, x_{3+i}, \dots, \hat{x}_{2N+1+i}) \\ &= \frac{\alpha}{2} + \frac{\alpha}{S} \sum_{n=1}^N s_n x_{2n-1+i}, \end{aligned} \quad (54)$$

where

$$\hat{x}_{2N+1+i} = 0.5 + \frac{1}{S} \sum_{n=1}^N s_n x_{2n-1+i}, \quad i = 0, 1. \quad (55)$$

By the definition of $\hat{H}(x_{1+i}, x_{3+i}, \dots, x_{2N-1+i})$ (cf. (53)), we have

$$\begin{aligned} \hat{H}(x_{1+i}, x_{3+i}, \dots, x_{2N-1+i}) &= \frac{\alpha}{2} + \frac{\alpha}{S} \sum_{n=1}^N s_n x_{2n-1+i}, \\ & \quad i = 0, 1. \end{aligned}$$

As a result, problem (52) can be decomposed into N subproblems

$$\begin{aligned} \min_{x_{2n-1}, x_{2n}} \quad & \hat{F}(x_{2n-1}, x_{2n}) := F(x_{2n-1}, x_{2n}) \\ & + \frac{\alpha s_n}{S} (x_{2n-1} + x_{2n}) \\ \text{s.t.} \quad & 0 \leq x_{2n-1}, x_{2n} \leq 1. \end{aligned} \quad (56)$$

We know from (49) that $\hat{F}(x_{2n-1}, x_{2n})$ in (56) is strictly concave with respect to x_{2n-1} and x_{2n} in $[0, 1] \times [0, 1]$. Since the minimum of a strictly concave function is always attained at a vertex [54], we immediately obtain that the optimal solution of (56) must be $(0, 0)$, $(0, 1)$, $(1, 0)$, or $(1, 1)$. It is easy to see that

$$\begin{aligned} \hat{F}(0, 1) &= \hat{F}(1, 0) = 2^q + \alpha \left(1 + \frac{s_n}{S}\right), \\ \hat{F}(0, 0) &= 2, \quad \hat{F}(1, 1) = 2 + 2\alpha \left(1 + \frac{s_n}{S}\right). \end{aligned}$$

This, together with the facts $0 < s_n/S < 1$ and $\alpha \leq \frac{2-2^q}{2}$ (cf. (47)), shows the optimal solution of (56) is

$$(\hat{x}_{2n-1}, \hat{x}_{2n}) = (0, 1) \text{ or } (1, 0). \quad (57)$$

Now, we can use (55) and (57) to conclude that the optimal value of problem (52) is

$$\sum_{n=1}^N \left(2^q + \alpha + \frac{\alpha s_n}{S}\right) + \alpha = N2^q + (N+2)\alpha.$$

Since the optimal value of problem (48) is less than or equal to $N2^q + (N+2)\alpha$ (the assumption of the ‘‘if’’ direction), it follows from (55) that

$$\hat{x}_{2N+1} = 0.5 + \frac{1}{S} \sum_{n=1}^N s_n \hat{x}_{2n-1} \leq 1,$$

$$\hat{x}_{2N+2} = 0.5 + \frac{1}{S} \sum_{n=1}^N s_n \hat{x}_{2n} \leq 1.$$

Combining this with (57) yields

$$\sum_{n=1}^N s_n \hat{x}_{2n-1} = \sum_{n \in \mathcal{S}} s_n = \frac{S}{2}, \quad \sum_{n=1}^N s_n \hat{x}_{2n} = \sum_{n \notin \mathcal{S}} s_n = \frac{S}{2},$$

where $\mathcal{S} = \{n \mid \hat{x}_{2n-1} = 1\}$. Therefore, there exists a subset \mathcal{S} such that (50) holds true, which shows that the partition problem has a ‘‘yes’’ answer.

Finally, this transformation can be finished in polynomial time. Since the partition problem is NP-complete, we conclude that problem (10) is NP-hard.

APPENDIX C PROOF OF LEMMA 2

Since \mathbf{w} is feasible, it follows that

$$\begin{aligned} \sum_{k=1}^K \log([\mathbf{w}_1]_k [\mathbf{w}_3]_k) &= \sum_{k=1}^K \log([\mathbf{w}_1]_k (1 - [\mathbf{w}_1]_k)) \\ &\leq -K \log(4), \end{aligned} \quad (58)$$

and

$$\frac{K}{q} \log(\|\mathbf{w}_2\|_q^q) - \sum_{k=1}^K \log([\mathbf{w}_2]_k) \geq \frac{K}{q} \log K, \quad (59)$$

where (59) comes from

$$\frac{\|\mathbf{w}_2\|_q^q}{K} = \frac{\sum_{k=1}^K [\mathbf{w}_2]_k^q}{K} \geq (\prod_{k=1}^K [\mathbf{w}_2]_k^q)^{1/K}.$$

For any strictly feasible \mathbf{w} , by the definition of $\phi(\mathbf{w})$ (cf. (20)), we obtain

$$\begin{aligned} \phi(\mathbf{w}) &= \rho \log(f(\mathbf{w})) - \sum_{k=1}^K \log([\mathbf{w}_1]_k) \\ &\quad - \sum_{k=1}^K \log([\mathbf{w}_2]_k) - \sum_{k=1}^K \log([\mathbf{w}_3]_k) \\ &\geq \left(\rho - \frac{K}{q}\right) \log(f(\mathbf{w})) \\ &\quad + \left(\frac{K}{q} \log(\|\mathbf{w}_2\|_q^q) - \sum_{k=1}^K \log([\mathbf{w}_2]_k)\right) \\ &\quad - \sum_{k=1}^K \log([\mathbf{w}_1]_k [\mathbf{w}_3]_k) \\ &\geq \left(\rho - \frac{K}{q}\right) \log(f(\mathbf{w})) + \frac{K}{q} \log(K) + K \log(4), \end{aligned} \quad (60)$$

where the first inequality is because $\tilde{\mathbf{c}}^T \mathbf{w}_1 \geq 0$ for any feasible \mathbf{w} , and the second is by (58) and (59). Therefore, if (21) holds, we must have $f(\mathbf{w}) \leq \epsilon$, which shows that \mathbf{w} is an ϵ -optimal solution.

APPENDIX D PROOF OF THEOREM 3

We first show that the potential reduction interior-point algorithm can return an ϵ -KKT point or an ϵ -optimal solution of problem (16) (equivalent to problem (10)). Towards this end, we consider the following two cases.

- If $\|g(\mathbf{w})\| > 1$, then we know

$$\phi(\mathbf{w}^+) - \phi(\mathbf{w}) < -\beta + \frac{\beta^2}{2(1-\beta)}.$$

The potential function value is reduced by a constant $2 - \sqrt{3}$ if we set $\beta = 1 - \frac{\sqrt{3}}{3} < 1$. If this case would hold for

$$O\left(\left(\rho - \frac{K}{q}\right) \log\left(\frac{1}{\epsilon}\right)\right) \quad (61)$$

iterations (cf. (21) and (22)), we would obtain an ϵ -optimal solution of (16).

- If $\|g(\mathbf{w})\| \leq 1$, then, from the definition of $g(\mathbf{w})$, we must have

$$0 \leq \frac{\rho}{f(\mathbf{w})} \mathbf{W}(\nabla f(\mathbf{w}) - \tilde{\mathbf{A}}^T \boldsymbol{\lambda}) \leq 2\mathbf{e}.$$

In other words,

$$\left[\nabla f(\mathbf{w}) - \tilde{\mathbf{A}}^T \boldsymbol{\lambda} \right]_k \geq 0, \quad \frac{\mathbf{w}_k \left[\nabla f(\mathbf{w}) - \tilde{\mathbf{A}}^T \boldsymbol{\lambda} \right]_k}{f(\mathbf{w})} \leq \frac{2}{\rho}, \quad \forall k.$$

By choosing

$$\rho \geq \frac{6K}{\epsilon}, \quad (62)$$

we have

$$\frac{\mathbf{w}^T (\nabla f(\mathbf{w}) - \tilde{\mathbf{A}}^T \boldsymbol{\lambda})}{f(\mathbf{w})} \leq \epsilon.$$

Therefore,

$$\begin{aligned} & \frac{\sum_{n=1}^{3K} \left(q[\mathbf{w}_n^q - [\tilde{\mathbf{A}}^T \boldsymbol{\lambda}]_n [\mathbf{w}_n] \right)}{\bar{f} - \underline{f}} \\ &= \frac{\mathbf{w}^T (\nabla f(\mathbf{w}) - \tilde{\mathbf{A}}^T \boldsymbol{\lambda})}{\bar{f} - \underline{f}} \leq \frac{\mathbf{w}^T (\nabla f(\mathbf{w}) - \tilde{\mathbf{A}}^T \boldsymbol{\lambda})}{f(\mathbf{w})} \leq \epsilon. \end{aligned}$$

Recalling Definition 1, we know that \mathbf{w} is an ϵ -KKT point of problem (16).

Next we establish the iteration complexity of the potential reduction algorithm. To do so, it suffices to bound $\phi(\mathbf{w}^0)$. By Assumption 1 and the special structure of \mathbf{A} , it can be easily shown that

$$\phi(\mathbf{w}^0) \leq \bar{\phi} := \rho \log(KR + K^{q+1}R^q) + 3K \log\left(\frac{2}{r}\right).$$

From the above, (21), (61), and (62), we know that the algorithm terminates within at most

$$\frac{\bar{\phi} - \phi}{2 - \sqrt{3}} = O\left(\frac{K}{\min\{\epsilon, q\}} \log\left(\frac{K^{q+1}R}{\epsilon r}\right)\right) \quad (63)$$

iterations. The proof of Theorem 3 is completed.

ACKNOWLEDGMENT

The authors would like to thank Prof. Z.-Q. (Tom) Luo of University of Minnesota, Prof. S. Zhang of University of Minnesota, and Dr. Q. Li of Beijing Institute of Technology for their useful discussions. The authors also would like to thank the anonymous reviewers for their useful comments on simulations.

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] 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.
- [3] 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.
- [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] 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.
- [6] 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.
- [7] 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, Sep. 1995.
- [8] S. A. Grandhi, J. Zander, and R. Yates, "Constrained power control," *Wireless Personal Commun.*, vol. 1, no. 4, pp. 257–270, Dec. 1994.
- [9] 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.
- [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. Adv. Wireless Commun. (SPAWC)*, Jun. 2010, pp. 1–5.
- [11] 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.
- [12] J. Zander, "Distributed cochannel interference control in cellular radio systems," *IEEE Trans. Veh. Technol.*, vol. 41, no. 3, pp. 305–311, Aug. 1992.
- [13] T.-H. Lee, J.-C. Lin, and Y.-T. Su, "Downlink power control algorithms for cellular radio systems," *IEEE Trans. Veh. Technol.*, vol. 44, no. 1, pp. 89–94, Feb. 1995.
- [14] T.-H. Lee and J.-C. Lin, "A fully distributed power control algorithm for cellular mobile systems," *IEEE J. Sel. Areas Commun.*, vol. 14, no. 4, pp. 692–697, May 1996.
- [15] J.-T. Wang and T.-H. Lee, "Non-reinitialized fully distributed power control algorithm," *IEEE Commun. Lett.*, vol. 3, no. 12, pp. 329–331, Dec. 1999.
- [16] A. Behzad, I. Rubin, and P. Chakravarty, "Optimum integrated link scheduling and power control for ad hoc wireless networks," in *Proc. IEEE Int. Conf. Wireless Mobile Comput., Netw., Commun.*, Aug. 2005, vol. 3, pp. 275–283.
- [17] 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.
- [18] 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.
- [19] 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, Syst., Comput.*, Nov. 2012, pp. 559–563.
- [20] T. Holliday, A. Goldsmith, P. Glynn, and N. Bambos, "Distributed power and admission control for time varying wireless networks," in *Proc. Global Telecommun. Conf. (GLOBECOM)*, Nov. 2004, pp. 768–774.
- [21] Z. Marantz, P. Orenstein, and D. J. Goodman, "Admission control for maximal throughput in power limited CDMA systems," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Mar. 2005, vol. 2, pp. 695–700.
- [22] 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.
- [23] S. Stanczak, M. Kaliszan, and N. Bambos, "Admission control for power-controlled wireless networks under general interference functions," in *Proc. 42nd Asilomar Conf. Signals, Syst., Comput.*, Oct. 2008, pp. 900–904.
- [24] M. Kaliszan, S. Stanczak, and N. Bambos, "Admission control for autonomous wireless links with power constraints," in *Proc. IEEE ICASSP*, Mar. 2010, pp. 3138–3141.
- [25] Y.-F. Liu and Y.-H. Dai, "Joint power and admission control via p norm minimization deflation," in *Proc. IEEE ICASSP*, May 2013, pp. 4789–4793.
- [26] K. Hamdi, W. Zhang, and K. B. Letaief, "Joint beamforming and scheduling in cognitive radio networks," in *Proc. GLOBECOM*, Nov. 2007, pp. 2977–2981.
- [27] Y.-F. Liu, "An efficient distributed joint power and admission control algorithm," in *Proc. 31th Chin. Control Conf.*, Jul. 2012, pp. 5508–5512.
- [28] Y.-F. Liu and E. Song, "Distributionally robust joint power and admission control via SOCP deflation," in *Proc. SPAWC*, Jun. 2013, pp. 11–15.

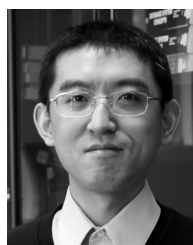
- [29] Y.-F. Liu, E. Song, and M. Hong, "Sample approximation-based deflation approaches for chance SINR constrained joint power and admission control," *IEEE Trans. Signal Process.*, submitted for publication.
- [30] E. Matakani, N. D. Sidiropoulos, Z.-Q. Luo, and L. Tassiulas, "Convex approximation techniques for joint multiuser downlink beamforming and admission control," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2682–2693, Jul. 2008.
- [31] M. H. Ahmed, "Call admission control in wireless networks: A comprehensive survey," *IEEE Commun. Surveys*, vol. 7, no. 1, pp. 50–69, 2005.
- [32] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization*, ser. SIAM-MPS Series on Optimization. Philadelphia, PA, USA: SIAM, 2001.
- [33] Y.-F. Liu, Y.-H. Dai, and S. Ma, "A companion technical report of "Joint power and admission control: Non-convex ℓ_q approximation and an effective polynomial time deflation approach"," *Acad. Math. Syst. Sci., Chin. Acad. Sci.*, 2013.
- [34] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.
- [35] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [36] Y. Zhang, "Theory of compressive sensing via ℓ_1 -minimization: A non-RIP analysis and extensions," *J. Oper. Res. Soc. China*, vol. 1, no. 1, pp. 79–105, Mar. 2013.
- [37] Z.-Q. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 57–73, Feb. 2008.
- [38] A. Wiesel, Y. C. Eldar, and S. S. Shitz, "Linear precoding via conic programming for fixed MIMO receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 161–176, Jan. 2006.
- [39] H. Dahrouj and W. Yu, "Coordinated beamforming for the multicell multi-antenna wireless system," *IEEE Trans. Wireless Commun.*, vol. 9, no. 5, pp. 1748–1759, May 2010.
- [40] Y.-F. Liu, Y.-H. Dai, and Z.-Q. Luo, "Coordinated beamforming for MISO interference channel: Complexity analysis and efficient algorithms," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1142–1157, Mar. 2011.
- [41] Y.-F. Liu, M. Hong, and Y.-H. Dai, "Max-min fairness linear transceiver design problem for a multi-user SIMO interference channel is polynomial time solvable," *IEEE Signal Process. Lett.*, vol. 20, no. 1, pp. 27–30, Jan. 2013.
- [42] A. Beck and M. Teboulle, "A convex optimization approach for minimizing the ratio of indefinite quadratic functions over an ellipsoid," *Math. Program., Ser. A*, vol. 118, pp. 13–35, 2009.
- [43] Y.-F. Liu and Y.-H. Dai, "On the complexity of joint subcarrier and power allocation for multi-user OFDMA systems," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 583–596, Feb. 2014.
- [44] Y.-F. Liu, "Complexity analysis of joint subcarrier and power allocation for the cellular downlink OFDMA system," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 661–664, Dec. 2014.
- [45] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*. San Francisco, CA, USA: Freeman, 1979.
- [46] C. H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1998.
- [47] D. Ge, X. Jiang, and Y. Ye, "A note on the complexity of L_p minimization," *Math. Prog.*, vol. 129, pp. 285–299, 2011.
- [48] X. Chen, D. Ge, Z. Wang, and Y. Ye, "Complexity of unconstrained l_2 - l_p minimization," *Math. Prog.*, vol. 143, pp. 371–383, 2014.
- [49] Y.-F. Liu, S. Ma, Y.-H. Dai, and S. Zhang, "An SSQP framework for a class of composite L_q minimization over polyhedron," *Math. Prog.*, submitted for publication.
- [50] Y. Ye, "On the complexity of approximating a KKT point of quadratic programming," *Math. Prog.*, vol. 80, pp. 195–211, 1998.
- [51] S. Ji, K.-F. Sze, Z. Zhou, A. M.-C. So, and Y. Ye, "Beyond convex relaxation: A polynomial-time non-convex optimization approach to network localization," in *Proc. IEEE INFOCOM*, 2013, pp. 2499–2507.
- [52] S. A. Vavasis, "Polynomial time weak approximation algorithms for quadratic programming," in *Complexity in Numerical Optimization*, C. A. Floudas and P. M. Pardalos, Eds. Hackensack, NJ, USA: World Scientific, 1993.
- [53] D. Bertsimas and J. Tsitsiklis, *Introduction to Linear Optimization*. Belmont, MA, USA: Athena Scientific, 1997.
- [54] H. L. Hoffman, "A method for globally minimizing concave functions over convex sets," *Math. Prog.*, vol. 20, no. 1, pp. 22–32, Dec. 1981.



Ya-Feng Liu (M'12) received the B.Sc. degree in applied mathematics in 2007 from Xidian University, Xi'an, China, and the Ph.D. degree in computational mathematics in 2012 from the Chinese Academy of Sciences (CAS), Beijing, China. During his Ph.D. study, he was supported by the Academy of Mathematics and Systems Science, CAS, to visit Professor Zhi-Quan (Tom) Luo at the University of Minnesota (Twins Cities) from February 2011 to February 2012. After his graduation, he joined the Institute of Computational Mathematics and Scientific/Engineering Computing, Beijing, China, in July 2012, where he is currently an Assistant Professor. His main research interests are nonlinear optimization and its applications to wireless communications, signal processing, and machine learning.



Yu-Hong Dai received the B.Sc. degree in applied mathematics from the Beijing Institute of Technology, Beijing, China, in 1992. He then studied nonlinear optimization in the Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences, and received the Ph.D. degree in nonlinear programming in 1997. After his graduation, he worked in the Academy of Mathematics and Systems Science (AMSS), Chinese Academy of Sciences, Beijing, China, and became a Full Professor in 2006. He is currently also Vice-Director of the Center for Optimization and Applications (COA) of AMSS. His research interests include nonlinear optimization and various optimization applications.



Shiqian Ma received the B.Sc. degree from Peking University in 2003, the M.S. degree from Chinese Academy of Sciences in 2006 and the Ph.D. degree in Industrial Engineering and Operations Research from Columbia University in 2011. He then spent one and half years in the Institute for Mathematics and Its Applications at University of Minnesota as an NSF postdoctoral fellow. Dr. Ma joined the Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong in December 2012. His current research interests include theory and algorithms for large-scale optimization and its applications in big data analytics, statistics, machine learning, signal processing and image processing.