



# The sparsest solution of the union of finite polytopes via its nonconvex relaxation

Guowei You<sup>1</sup> · Zheng-Hai Huang<sup>2</sup> · Yong Wang<sup>2</sup>

Received: 2 November 2017 / Published online: 18 February 2019  
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

## Abstract

Sparse optimization problems have gained much attention since 2004. Many approaches have been developed, where nonconvex relaxation methods have been a hot topic in recent years. In this paper, we study a partially sparse optimization problem, which finds a partially sparsest solution of a union of finite polytopes. We discuss the relationship between its solution set and the solution set of its nonconvex relaxation. In details, by using geometrical properties of polytopes and properties of a family of well-defined nonconvex functions, we show that there exists a positive constant  $p^* \in (0, 1]$  such that for every  $p \in [0, p^*)$ , all optimal solutions to the nonconvex relaxation with the parameter  $p$  are also optimal solutions to the original sparse optimization problem. This provides a theoretical basis for solving the underlying problem via its nonconvex relaxation. Moreover, we show that the problem we concerned covers a wide range of problems so that several important sparse optimization problems are its subclasses. Finally, by an example we illustrate our theoretical findings.

**Keywords** Nonconvex relaxation · Sparse optimization · Sparse solution

**Mathematics Subject Classification** 90C25 · 93C41 · 65K05

---

This work was partially supported by the National Natural Science Foundation of China (Grant Nos. 11431002 and 11871051).

---

✉ Zheng-Hai Huang  
huangzhenghai@tju.edu.cn

Guowei You  
youziyou1017@126.com

Yong Wang  
wang\_yong@tju.edu.cn

<sup>1</sup> School of Mathematics and Statistics, Henan University of Science and Technology, Luoyang 471023, People's Republic of China

<sup>2</sup> School of Mathematics, Tianjin University, Tianjin 300072, People's Republic of China

## 1 Introduction

Compressive Sensing (CS) has been popularized for the past more than 10 years (see Candés et al. 2006; Candés and Tao 2005; Foucart and Rauhut 2013); its target is to reconstruct a sparse signal from an underdetermined system of linear equations. Mathematically, CS is to find a sparsest solution to a system of linear equations, which is NP-hard (see Natarajan 1995) because of the combinatorial feature of the problem. Many approaches have been developed for solving this class of problems, of which one of the most popular methods is the relaxation method, i.e., find a solution to the original problem via its relaxation problem.

So far, the convex relaxation method of CS has been developed well, in both theory and algorithm. However, it has been found from both theoretical and numerical aspects that it needs fewer measurements for exact recovery via the non-convex relaxation (such as  $\ell_p$ -minimization) than the convex relaxation (such as  $\ell_1$ -minimization) (Chartrand 2007; Lyu et al. 2013; Zhang et al. 2013). Thus, nonconvex relaxation methods, including  $\ell_p$ -minimization, have become a hot spot in recent years (see, for example, Chartrand 2007; Davies and Gribonval 2009; Foucart and Lai 2009; Gribonval and Nielson 2003; Lai and Wang 2011; Peng et al. 2015; Saab et al. 2008; Sun 2012; Wang et al. 2011; Xu et al. 2012; Zhang et al. 2013).

Recently, Fung and Mangasarian (2011) discussed the relationship between the  $\ell_0$  and  $\ell_p$  minimal solutions to the system of linear equalities and inequalities with box constraints; and they showed that there is a (problem dependent) constant smaller or equal to one, denoted by  $p^*$ , such that any optimal solution of  $\ell_p$ -minimization with  $p \in [0, p^*)$  also solves the concerned  $\ell_0$ -minimization. Such an important property has also been investigated for the system of linear equations (Peng et al. 2015), the linear complementarity problem (Chen and Xiang 2016) and the phaseless compressive sensing problem (You et al. 2017).

Inspired by the studies mentioned above, we investigate the partially sparsest solution of the union of finite polytopes. This problem covers a wide range which contains the problems discussed in Fung and Mangasarian (2011), Peng et al. (2015) and You et al. (2017) as special cases. We consider this problem via its nonconvex relaxation by using a class of nonconvex approximation functions which contains Schatten- $p$  quasi-norm as a special case. We show that there exists a constant  $p^* \in (0, 1]$  such that for any parameter  $p$  involving in the nonconvex approximation function satisfying  $p \in [0, p^*)$ , every optimal solution of the relaxation problem is also an optimal solution of the original problem, which is a generalization of those obtained in Fung and Mangasarian (2011), Peng et al. (2015) and You et al. (2017). We give an example to illustrate our theoretical result.

The paper is organized as follows. In Sect. 2, we state the problem we concerned and provide some basic concepts and results which are useful in later sections. In Sect. 3, we reformulate equivalently the original sparse optimization problem and its nonconvex relaxation problem into new forms by means of some results from polytope, then study the relationship between the solution sets of these two problems. In Sect. 4, we show that several classes of sparse optimization problems are subclasses of the

problem we concerned; and hence, the main result is valid for these problems. Some conclusions and comments are given in Sect. 5.

The conventions of the notations in this paper. Uppercase (greek) letters denote matrices, e.g.,  $A, \Phi$ ; boldface lowercase letters denote column vectors, e.g.,  $\mathbf{x}, \mathbf{b}$ ;  $\|\mathbf{x}\|_0$  means the number of nonzero components of vector  $\mathbf{x}$ ;  $\|\mathbf{x}\|_\infty$  is infinity norm defined by  $\max_i \{|x_i|\}$ ;  $\|\mathbf{x}\|_p$  is Schatten- $p$  quasi-norm of  $\mathbf{x}$  defined by  $(\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$ , called ' $\ell_p$ -norm of  $\mathbf{x}$ ', where  $0 < p < 1$ ;  $\mathbf{1}$  is a vector with each component being 1;  $\mathbf{0}$  is a zero-vector;  $|\mathbf{u}|$  is a vector with the  $i$ -th component being  $|u_i|$  for all  $i$ ;  $(\mathbf{x}, \mathbf{y})$  stands for the vector  $(\mathbf{x}^\top, \mathbf{y}^\top)^\top$ , which is a column vector;  $\mathbf{u} \leq \mathbf{v}$  signifies vector inequality which satisfies  $u_i \leq v_i$  for all  $i$ ;  $A^\top$  represents the transpose of matrix  $A$  and  $\mathbf{u}^\top$  the transpose of vector  $\mathbf{u}$ ; lowercase letters denote real scalars, e.g.,  $u_i, a, b, \dots$ , particularly,  $u_i$  stands for the  $i$ -th component of  $\mathbf{u}$ ;  $[n]$  is the set  $\{1, 2, \dots, n\}$  with  $n$  being a positive integer;  $\mathbb{R}_+$  is the set of non-negative real numbers;  $\mathbb{R}_+^n$  is the non-negative orthant  $\{(x_1, \dots, x_n)^\top | x_i \geq 0, i \in [n]\}$ ;  $\mathbb{R}^{m \times n}$  is the set of all real matrices with  $m$  rows and  $n$  columns;  $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_m)^\top \in \{-1, 1\}^m$  is an  $m$ -dimensional vector with each  $\epsilon_i$  being  $-1$  or  $1$ ;  $\mathbf{u} \circ \mathbf{v}$  is Hadamard product of vectors  $\mathbf{u}$  and  $\mathbf{v}$ ;  $\text{sign}(a)$  is sign function, e.g., if  $a > 0$ , then  $\text{sign}(a) = 1$ , else if  $a = 0$ , then  $\text{sign}(a) = 0$ , else  $\text{sign}(a) = -1$  if  $a < 0$ ;  $\text{sign}(\mathbf{u})$  is a sign vector with the  $i$ -th component being  $\text{sign}(u_i)$ ;  $\mathbb{I}, \mathbb{I}^c$  are two index sets contained in  $[n]$ , with  $\mathbb{I}^c = [n] \setminus \mathbb{I}$ ;  $\#\mathbb{I}$  is cardinality of index set  $\mathbb{I}$ , i.e., the number of elements of  $\mathbb{I}$ ;  $\Phi_{\mathbb{I}}$  denotes a submatrix of  $\Phi$  constructed from columns corresponding to indices of index set  $\mathbb{I}$ , and  $\mathbf{x}_{\mathbb{I}}$  a sub-vector of  $\mathbf{x}$  constructed from components corresponding to indices of index set  $\mathbb{I}$ ;  $\text{Conv}(\mathbb{T})$  is convex hull of the set  $\mathbb{T}$ .  $\text{Diag}(\mathbf{v})$  denotes a diagonal matrix generated by vector  $\mathbf{v}$ .

## 2 Preliminaries

In this paper, we study the following partially sparse optimization problem:

$$\min_{(\mathbf{x}, \mathbf{y})} \|\mathbf{y}\|_0 \quad \text{s.t.} \quad (\mathbf{x}, \mathbf{y}) \in \bigcup_{i \in \mathcal{I}} \mathbb{P}_i, \mathbf{1} \leq \mathbf{x} \leq \mathbf{u}, \mathbf{0} \leq \mathbf{y} \leq \mathbf{1}, \tag{1}$$

where  $\mathbf{x} \in \mathbb{R}^{n_1}, \mathbf{y} \in \mathbb{R}^{n_2}, (\mathbf{x}, \mathbf{y}) \in \mathbb{R}^N$  with  $N = n_1 + n_2$ , each  $\mathbb{P}_i$  is a polyhedron in  $\mathbb{R}^N$ ,  $\mathcal{I}$  is a finite index set, and  $\mathbf{1}, \mathbf{u} \in \mathbb{R}^{n_1}$  are two known vectors satisfying  $\mathbf{1} \leq \mathbf{u}$ .

To make sense of problem (1), we assume that the feasible set of problem (1) is nonempty throughout this paper.

We denote the feasible set of (1) by

$$\mathbb{T} := \bigcup_{i \in \mathcal{I}} \mathbb{T}_i \subseteq \mathbb{R}^N \tag{2}$$

where, for any  $i \in \mathcal{I}$ ,

$$\mathbb{T}_i := \left\{ (\mathbf{x}, \mathbf{y}) \in \mathbb{R}^N \mid (\mathbf{x}, \mathbf{y}) \in \mathbb{P}_i, \mathbf{1} \leq \mathbf{x} \leq \mathbf{u}, \mathbf{0} \leq \mathbf{y} \leq \mathbf{1} \right\}. \tag{3}$$

By the assumption that the feasible set of problem (1) is nonempty and the fact that a set is a polytope if and only if it is a bounded polyhedral set (see Grünbaum 2003, p. 32), we have the following result.

**Proposition 2.1** *The set  $\mathbb{T}$  defined by (2) is nonempty and it is a union of finitely many polytopes in  $\mathbb{R}^N$ .*

By Proposition 2.1, problem (1) is a sparse optimization problem over a union of finitely many polytopes. It should be noticed that  $\mathbb{T}$  is convex when  $\#\mathcal{I} = 1$  and not necessarily convex when  $\#\mathcal{I} > 1$ .

**Remark 2.1** Consider the problem

$$\min_{(\mathbf{x}, \mathbf{z})} \|\mathbf{z}\|_0 \quad \text{s.t.} \quad (\mathbf{x}, \mathbf{z}) \in \bigcup_{i \in \mathcal{I}} \tilde{\mathbb{P}}_i, \mathbf{1} \leq \mathbf{x} \leq \mathbf{u}, \mathbf{0} \leq \mathbf{z} \leq \mathbf{v}, \tag{4}$$

where  $\mathbf{z} \in \mathbb{R}^{n_2}$ ,  $\mathbf{v} \in \mathbb{R}^{n_2}$  with  $\mathbf{v} > \mathbf{0}$ , and every  $\tilde{\mathbb{P}}_i$  is a polyhedron in  $\mathbb{R}^N$ . Then, it is easy to see that problem (4) can be reformulated as a problem in the form of problem (1).

Problem (1) is a class of nonconvex optimization problems with wide range, which contains many popular problems as special cases (see some discussions in Sect. 4). In this paper, we investigate this class of problems via its nonconvex relaxation. For this purpose, we define the following set:

$$\mathcal{F}(t; p) := \{f(t; p) \mid t \in [0, \infty), f(t; p) \text{ satisfies properties (C1)–(C3)}\}, \tag{5}$$

where properties (C1)–(C3) are as follows: for every given number  $p \in (0, 1]$ ,

(C1)  $f(t; p)$  is strictly (increasing) concave function in  $t \in [0, \infty)$ ;

(C2)  $f(0; p) = 0$  and  $0 < f(t; p) \leq 1$  whenever  $t \in (0, 1]$ ;

(C3)  $\lim_{p \downarrow 0} f(t; p) = 1$  as  $t > 0$ .

Several functions  $f(t; p) \in \mathcal{F}(t; p)$  have been proposed in Gribonval and Nielsen (2007), Bradley and Mangasarian (1998) and Haddou and Migot (2015). For example,

$$(a) f_1(t; p) = \frac{t}{t + p}; \quad (b) f_2(t; p) = 1 - e^{-\frac{t}{p}}; \quad (c) f_3(t; p) = t^p.$$

**Remark 2.2** Let  $f(t; p) \in \mathcal{F}(t; p)$ . By properties (C2) and (C3), it is easy to see that  $\text{sign}(t) = \lim_{p \rightarrow 0} f(t; p)$  for any real number  $t \geq 0$ .

For any function  $f(t; p) \in \mathcal{F}(t; p)$ , we will consider the following relaxation problem of problem (1):

$$\min_{(\mathbf{x}, \mathbf{y})} \sum_{i=1}^{n_2} f(y_i; p) \quad \text{s.t.} \quad (\mathbf{x}, \mathbf{y}) \in \bigcup_{i \in \mathcal{I}} \mathbb{P}_i, \mathbf{1} \leq \mathbf{x} \leq \mathbf{u}, \mathbf{0} \leq \mathbf{y} \leq \mathbf{1}, \tag{6}$$

To further discuss the relationship between problems (1) and (6), we need the following definitions concerning monotonically nondecreasing vector-valued and monotonically nondecreasing real functions.

**Definition 2.1** (You et al. 2017, Definition 2.1) (i) Define a function  $\mathbf{f} : \mathbb{R}_+^n \rightarrow \mathbb{R}_+^n$  by

$$\mathbf{f}(\mathbf{x}) = (f(x_1), f(x_2), \dots, f(x_n))^\top, \tag{7}$$

where  $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ . The (vector-valued) function  $\mathbf{f}$  is said to be monotonically nondecreasing if for any two nonnegative vectors  $\mathbf{u}$  and  $\mathbf{v}$ ,  $\mathbf{u} \leq \mathbf{v}$  implies  $\mathbf{f}(\mathbf{u}) \leq \mathbf{f}(\mathbf{v})$ .

(ii) Let  $F : \mathbb{R}_+^n \rightarrow \mathbb{R}_+$ . The function  $F$  is said to be monotonically nondecreasing if for any two nonnegative vectors  $\mathbf{u}$  and  $\mathbf{v}$ ,  $\mathbf{u} \leq \mathbf{v}$  implies  $F(\mathbf{u}) \leq F(\mathbf{v})$ .

Two monotonically nondecreasing functions are given in the following example, which will be frequently used in our subsequent analysis.

**Example 2.1** (i) Suppose that  $\mathbf{f} : \mathbb{R}_+^n \rightarrow \mathbb{R}_+^n$  is defined by (7) with  $f(x_i) = \text{sign}(x_i)$  for any  $i \in [n]$ . We denote this function by  $\text{sign}(\cdot)$ . It is easy to see that  $\text{sign}(\cdot)$  is a monotonically nondecreasing function.

(ii) Suppose that  $\mathbf{f} : \mathbb{R}_+^n \rightarrow \mathbb{R}_+^n$  is defined by (7) with strictly increasing concave function  $f(x_i) = f(x_i; p)$  for any  $i \in [n]$  and  $p \in (0, 1]$ , where  $f \in \mathcal{F}(\cdot; p)$  which is defined in (5). It is easy to see that this function is monotonically nondecreasing.

For any given vectors  $\mathbf{u}$  and  $\mathbf{v}$  with  $\mathbf{0} \leq \mathbf{u} \leq \mathbf{v}$  and a monotonically nondecreasing function  $\mathbf{f}$ , we have

$$\mathbf{1}^\top \mathbf{f}(\mathbf{u}) = \sum_{i=1}^n f(u_i) \leq \sum_{i=1}^n f(v_i) = \mathbf{1}^\top \mathbf{f}(\mathbf{v}).$$

Therefore,  $\mathbf{1}^\top \mathbf{f}$  is monotonically nondecreasing.

### 3 Relationship between optimal solutions of problem (1) and problem (6)

In this section, we study the relationship between optimal solutions of problem (1) and problem (6).

Applying the set  $\mathbb{T}$  defined by (2), problem (1) can be simply rewritten as

$$\min_{(\mathbf{x}, \mathbf{y})} \|\mathbf{y}\|_0 \quad \text{s.t.} \quad (\mathbf{x}, \mathbf{y}) \in \mathbb{T}. \tag{8}$$

As is known that the objective function of (8) is nonconvex and the feasible set is usually nonconvex, problem (8) is NP-hard (Natarajan 1995). Similarly, problem (6) is simply rewritten as

$$\min_{(\mathbf{x}, \mathbf{y})} \mathbf{1}^\top \mathbf{f}(\mathbf{y}; p) \quad \text{s.t.} \quad (\mathbf{x}, \mathbf{y}) \in \mathbb{T}, \tag{9}$$

where  $\mathbf{f}(\mathbf{t}; p) = (f(t_1; p), f(t_2; p), \dots, f(t_{n_2}; p))^\top \in \mathbb{R}^{n_2}$  is a vector-valued function with  $f(t; p) \in \mathcal{F}(t; p)$ .

Denote the convex hull of the set  $\mathbb{T}$  defined in (2) by

$$\mathbb{S} := \text{Conv}(\mathbb{T}) = \text{Conv}\left(\bigcup_{i \in \mathcal{I}} \mathbb{T}_i\right). \tag{10}$$

Then, we have the following result.

**Proposition 3.1** *The set  $\mathbb{S}$  defined by (10) is a nonempty polytope contained in the box  $\prod_{i=1}^{n_1} [l_i, u_i] \times [0, 1]^{n_2} \subset \mathbb{R}^N$  with  $N = n_1 + n_2$ .*

**Proof** By Proposition 2.1, it follows that the set  $\mathbb{T}$  is nonempty and it is the union of finitely many polytopes in  $\mathbb{R}^N$ . Thus, the set  $\mathbb{S}$  is nonempty and is a convex hull of finitely many polytopes. By the fact that the convex hull of finitely many polytopes is a polytope (see Grünbaum 2003, p. 32), it follows that  $\mathbb{S}$  is a polytope. Furthermore, since  $\mathbb{T} \subset \prod_{i=1}^{n_1} [l_i, u_i] \times [0, 1]^{n_2}$  with the latter set being a (convex) box and  $\mathbb{S}$  is the smallest convex set containing  $\mathbb{T}$ , it is easy to see that  $\mathbb{S}$  is a subset of this box. Thus, the result of the proposition holds.  $\square$

For the set  $\mathbb{T}$  defined in (2), we call  $\mathbf{x} \in \mathbb{T}$  a *pseudo-extreme point* of  $\mathbb{T}$  if it is a vertex of some polytope  $\mathbb{T}_i$ , which implies that there are no two distinct points  $\mathbf{y}, \mathbf{z} \in \mathbb{T}_i$  and a constant  $\lambda \in (0, 1)$  such that  $\mathbf{x} = \lambda\mathbf{y} + (1 - \lambda)\mathbf{z}$ . In addition, an extreme point of  $\mathbb{S}$  is called a vertex of  $\mathbb{S}$  as usually defined. In the following, we use  $\text{vert}(\mathbb{S})$  to denote the set of all vertices of  $\mathbb{S}$  and  $\text{pext}(\mathbb{T})$  to denote the set of all pseudo-extreme points of  $\mathbb{T}$ .

Using the convex hull  $\mathbb{S}$ , problem (9) is relaxed as

$$\min_{(\mathbf{x}, \mathbf{y})} \mathbf{1}^\top \mathbf{f}(\mathbf{y}; p) \quad \text{s.t.} \quad (\mathbf{x}, \mathbf{y}) \in \mathbb{S}. \tag{11}$$

In what follows, we discuss the relationship between problem (9) and problem (11). We have the following result.

**Theorem 3.1** *Suppose that  $\mathbf{x}^* \in \text{vert}(\mathbb{S})$  is an optimal solution to problem (11), then  $\mathbf{x}^* \in \text{pext}(\mathbb{T})$  and it is an optimal solution to problem (9).*

**Proof** Notice that problem (11) and problem (9) are different mainly in their feasible sets, that is, one is  $\mathbb{T}$  and another is its convex hull  $\mathbb{S}$ . By polytope geometry, it follows that  $\mathbb{S} = \text{Conv}(\text{pext}(\mathbb{T}))$  (see Ziegler 1995, Theorem 2.15(3)). Furthermore, by the fact that if a polytope can be written as the convex hull of a finite point set, then the set contains all the vertices of the polytope (see Ziegler 1995, Proposition 2.2), we have that  $\text{pext}(\mathbb{T}) \supseteq \text{vert}(\mathbb{S})$ . Therefore, by You et al. (2017, Lemma 2.2), each vertex solution to problem (11) is a pseudo-extreme point which is an optimal solution to problem (9).  $\square$

In the following, if a pseudo-extreme point of  $\mathbb{T}$  is an optimal solution to problem (9), then we call it a *pseudo-extreme-point optimal solution to problem (9)*. The following addresses the existence of optimal solution to problem (9).

**Theorem 3.2** *For any given  $p \in (0, 1)$ , there exists at least one pseudo-extreme-point optimal solution to problem (9).*

**Proof** Denote  $f(\mathbf{x}, \mathbf{y}) := \mathbf{1}^\top \mathbf{f}(\mathbf{y}; p) = \sum_{i=1}^{n_2} f(y_i; p)$  with  $p \in (0, 1)$ , then  $-f$  is a convex function on  $\text{dom } f = \mathbb{R}^{n_1} \times \mathbb{R}_+^{n_2}$ . By Proposition 3.1, it follows that the set  $\mathbb{S}$  is nonempty and  $\mathbb{S} \subseteq \prod_{i=1}^{n_1} [l_i, u_i] \times [0, 1]^{n_2} \subset \text{dom } f$ . There are no half-lines and no lines in  $\mathbb{S}$ ; and  $-f$  is bounded above on  $\mathbb{S}$ . By Rockafellar (1970, Corollary 32.3.4), there exists a vertex optimal solution to problem (11), which is also a pseudo-extreme-point optimal solution to problem (9) by Theorem 3.1. In other words, for each real number  $p \in (0, 1)$ , there exists at least one pseudo-extreme-point optimal solution to problem (9).  $\square$

**Theorem 3.3** *Suppose  $\{p_k\}$  is an arbitrarily given sequence with  $0 < p_{k+1} < p_k < 1$  for any  $k = 1, 2, \dots$  and  $\lim_{k \rightarrow \infty} p_k = 0$ . Then,*

- (i) *there is a subsequence  $\{p_{i_k}\} \subseteq \{p_k\}$  such that some fixed pseudo-extreme point  $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$  of  $\mathbb{T}$  is an optimal solution to (9) with any  $p_{i_j} \in \{p_{i_k}\}$ ; and*
- (ii)  *$(\hat{\mathbf{x}}, \hat{\mathbf{y}})$  is an optimal solution to (8).*

**Proof** (i) For each  $p_i \in \{p_k\}$  with  $0 < p_k < 1$  for any  $k = 1, 2, \dots$  and  $\lim_{k \rightarrow \infty} p_k = 0$ , by Theorem 3.2 it follows that there exists a pseudo-extreme-point optimal solution  $(\hat{\mathbf{x}}(p_i), \hat{\mathbf{y}}(p_i))$  to problem (9) with  $p = p_i$ .

Since the number of pseudo-extreme points of the set  $\mathbb{T}$  is finite, it follows that the number of pseudo-extreme-point optimal solutions is finite; while the set  $\{p_k\}$  is infinite, so we conclude that there exists an infinite sub-sequence  $\{p_{i_k}\} \subset \{p_k\}$  such that for any  $p_{i_j} \in \{p_{i_k}\}$ , problem (9) with  $p = p_{i_j}$  has the same pseudo-extreme-point optimal solution  $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ . The first part of the theorem is proved.

(ii) We show that  $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$  is a partially sparsest solution of the original problem (1). Noting that  $\mathbf{y} \in \mathbb{R}^{n_2}$  and  $\mathbf{0} \leq \mathbf{y} \leq \mathbf{1}$ , we have

$$(\text{sign}(\mathbf{y}))_i = \text{sign}(y_i) = \begin{cases} 1, & \text{if } y_i > 0, \\ 0, & \text{if } y_i = 0 \end{cases}$$

and

$$\|\mathbf{y}\|_0 = \sum_{i=1}^{n_2} \text{sign}(y_i). \tag{12}$$

Since  $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$  is an optimal solution to problem (9) with  $p = p_{i_j}$  for any  $p_{i_j} \in \{p_{i_k}\}$ , we have  $(\hat{\mathbf{x}}, \hat{\mathbf{y}}) \in \mathbb{T}$ ; and hence,  $\mathbf{1} \geq \hat{\mathbf{y}} \geq \mathbf{0}$ . Furthermore,  $\|\hat{\mathbf{y}}\|_0 = \sum_{i=1}^{n_2} \text{sign}(\hat{y}_i)$  and

$$\mathbf{1}^\top \mathbf{f}(\hat{\mathbf{y}}; p_{i_j}) = \min_{(\mathbf{x}, \mathbf{y}) \in \mathbb{T}} \mathbf{1}^\top \mathbf{f}(\mathbf{y}; p_{i_j}). \tag{13}$$

For simplicity, we use  $p_j$  to replace  $p_{i_j}$ . Therefore,

$$\begin{aligned} \|\hat{\mathbf{y}}\|_0 &= \sum_{i=1}^{n_2} \text{sign}(\hat{y}_i) = \sum_{i=1}^{n_2} \lim_{j \rightarrow \infty} f(\hat{y}_i; p_j) = \lim_{j \rightarrow \infty} \sum_{i=1}^{n_2} f(\hat{y}_i; p_j) = \lim_{j \rightarrow \infty} \mathbf{1}^\top \mathbf{f}(\hat{\mathbf{y}}; p_j) \\ &= \lim_{j \rightarrow \infty} \min_{(\mathbf{x}, \mathbf{y}) \in \mathbb{T}} \mathbf{1}^\top \mathbf{f}(\mathbf{y}; p_j) = \lim_{j \rightarrow \infty} \min_{(\mathbf{x}, \mathbf{y}) \in \mathbb{T}} \sum_{i=1}^{n_2} f(y_i; p_j) \end{aligned}$$

$$\begin{aligned} &\leq \lim_{j \rightarrow \infty} \sum_{i=1}^{n_2} f(y_i; p_j) \text{ for all } (\mathbf{x}, \mathbf{y}) \in \mathbb{T} \\ &\leq \sum_{i=1}^{n_2} \text{sign}(y_i) = \|\mathbf{y}\|_0 \text{ for all } (\mathbf{x}, \mathbf{y}) \in \mathbb{T} \end{aligned} \tag{14}$$

where the first equality follows from (12), the fifth equality from (13), and the last equality from (12).

From (14) and the arbitrariness of  $(\mathbf{x}, \mathbf{y}) \in \mathbb{T}$ , we have

$$\|\hat{\mathbf{y}}\|_0 \leq \min_{(\mathbf{x}, \mathbf{y}) \in \mathbb{T}} \|\mathbf{y}\|_0,$$

which implies that  $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$  is an optimal solution to problem (8), i.e., (1). The second part is proved. □

**Theorem 3.4** *There exists a constant  $p^\sharp \in (0, 1]$  such that for any fixed  $p \in (0, p^\sharp)$ , there exists a pseudo-extreme-point optimal solution to problem (9), which is also an optimal solution to problem (8).*

**Proof** Suppose the statement is not true. Then for any fixed  $\tilde{p} \in (0, 1]$ , there exists a  $\bar{p} \in (0, \tilde{p})$  such that none of all pseudo-extreme-point optimal solutions to (9) with  $p = \bar{p}$  is the optimal solution to problem (8). In what follows, we will derive a contradiction.

Define a sequence  $\mathbb{P} := \{\tilde{p}_1, \tilde{p}_2, \dots\}$  with  $1 \geq \tilde{p}_i > \tilde{p}_{i+1}, i = 1, 2, \dots$  and  $\lim_{i \rightarrow +\infty} \tilde{p}_i = 0$ . By the above statement, for any  $\tilde{p}_i \in \mathbb{P}$ , there exists a  $\bar{p}_i > 0$  with  $\bar{p}_i < \tilde{p}_i$  such that none of all pseudo-extreme point optimal solutions to problem (9) with  $p = \bar{p}_i$  is the optimal solution to problem (8). Let  $\hat{p}_i = \min\{\tilde{p}_{i+1}, \bar{p}_i\}$ . Again by the same statement, there exists a  $\bar{p}_{i+1} > 0$  with  $\bar{p}_{i+1} < \hat{p}_i$  such that none of all pseudo-extreme-point optimal solutions to problem (9) with  $p = \bar{p}_{i+1}$  is the optimal solution to problem (8).

Noticing that  $\tilde{p}_i > \bar{p}_i \geq \hat{p}_i > \bar{p}_{i+1} > 0$  with  $i = 1, 2, \dots$ , we get a sequence  $\{\bar{p}_i\}$ , which satisfies that  $0 < \bar{p}_{i+1} < \bar{p}_i < 1$  and  $\lim_{i \rightarrow +\infty} \bar{p}_i = 0$ , such that none of all pseudo-extreme-point optimal solutions to problem (9) with any  $p \in \{\bar{p}_i\}$  is the optimal solution to problem (8). This contradicts the results of Theorem 3.3. The proof is complete. □

In the following, we will give two main results in this section.

**Theorem 3.5** *Suppose that  $p^\sharp \leq 1$  is given by Theorem 3.4. Then there exists a constant  $p^* \in (0, p^\sharp]$  such that for any given real number  $p$  with  $0 < p < p^*$ , every pseudo-extreme-point optimal solution to problem (9) is an optimal solution to problem (8).*

**Proof** Assume that the statement is not true. Then for an arbitrary real number  $\tilde{p} \in (0, p^\sharp] \subseteq (0, 1]$ , there is a real number  $\bar{p} \in (0, \tilde{p})$  such that there is a pseudo-extreme-point optimal solution  $(\bar{\mathbf{x}}(\bar{p}), \bar{\mathbf{y}}(\bar{p}))$  to problem (9) with  $p = \bar{p}$ , but it is not an optimal



solution to problem (8); meanwhile, by Theorem 3.4, there is another pseudo-extreme-point optimal solution  $(\hat{\mathbf{x}}(\bar{p}), \hat{\mathbf{y}}(\bar{p}))$  to problem (9) with  $p = \bar{p}$  that is also an optimal solution to problem (8). In what follows, we will deduce a contradiction.

Define a sequence  $\mathbb{P} := \{\tilde{p}_1, \tilde{p}_2, \dots\} \subseteq (0, p^\sharp]$  with  $1 \geq p^\sharp \geq \tilde{p}_i > \tilde{p}_{i+1} > 0, i = 1, 2, \dots$  and  $\lim_{i \rightarrow +\infty} \tilde{p}_i = 0$ . By the statement above, for any  $\tilde{p}_i \in \mathbb{P}$ , there exists  $\bar{p}_i > 0$  with  $\bar{p}_i < \tilde{p}_i$  such that there is a pseudo-extreme-point optimal solution  $(\bar{\mathbf{x}}(\bar{p}_i), \bar{\mathbf{y}}(\bar{p}_i))$  to problem (9) with  $p = \bar{p}_i$ , but it is not an optimal solution to problem (8); meanwhile, by Theorem 3.4, there is another pseudo-extreme-point optimal solution  $(\hat{\mathbf{x}}(\bar{p}_i), \hat{\mathbf{y}}(\bar{p}_i))$  to problem (9) with  $p = \bar{p}_i$  that is also an optimal solution to problem (8). Let  $\hat{p}_i = \min\{\tilde{p}_{i+1}, \bar{p}_i\}$ . Again by the same statement, there exists  $\bar{p}_{i+1} > 0$  with  $\bar{p}_{i+1} < \hat{p}_i$  such that there is a pseudo-extreme-point optimal solution  $(\bar{\mathbf{x}}(\bar{p}_{i+1}), \bar{\mathbf{y}}(\bar{p}_{i+1}))$  to problem (9) with  $p = \bar{p}_{i+1}$ , but it is not an optimal solution to problem (8); meanwhile, by Theorem 3.4, there is another pseudo-extreme-point optimal solution  $(\hat{\mathbf{x}}(\bar{p}_{i+1}), \hat{\mathbf{y}}(\bar{p}_{i+1}))$  to problem (9) with  $p = \bar{p}_{i+1}$  that is also an optimal solution to problem (8).

Noticing that  $\tilde{p}_i > \bar{p}_i \geq \hat{p}_i > \bar{p}_{i+1} > 0$  with  $i = 1, 2, \dots$ , we get a sequence  $\{\bar{p}_i\}$ , which satisfies that  $0 < \bar{p}_{i+1} < \bar{p}_i \leq 1$  and  $\lim_{i \rightarrow +\infty} \bar{p}_i = 0$ , such that there is a pseudo-extreme-point optimal solution  $(\bar{\mathbf{x}}(\bar{p}_i), \bar{\mathbf{y}}(\bar{p}_i))$  to problem (9) with any  $p \in \{\bar{p}_i\}$ , but it is not an optimal solution to problem (8); meanwhile, there is another pseudo-extreme-point optimal solution  $(\hat{\mathbf{x}}(\bar{p}_i), \hat{\mathbf{y}}(\bar{p}_i))$  to problem (9) with any  $p \in \{\bar{p}_i\}$  that is also an optimal solution to problem (8).

Since the number of pseudo-extreme points of  $\mathbb{T}$  is finite, it follows that the pseudo-extreme-point optimal solution set  $\{(\bar{\mathbf{x}}(\bar{p}_i), \bar{\mathbf{y}}(\bar{p}_i)), (\hat{\mathbf{x}}(\bar{p}_i), \hat{\mathbf{y}}(\bar{p}_i))\}$  is finite; while the sequence  $\{\bar{p}_i\}$  is infinite, we conclude that there must be a sub-sequence  $\{\bar{p}_{i_k}\} \subseteq \{\bar{p}_i\}$  such that there is a fixed pseudo-extreme point  $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$  that is an optimal solution to every problem (9) with  $p = \bar{p}_{i_k}$ , however,  $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$  is not an optimal solution to problem (8); meanwhile there is some fixed pseudo-extreme point  $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$  being an optimal solution to problem (9) with  $p = \bar{p}_{i_k}$  that is also an optimal solution to problem (8). For simplicity, we use  $\bar{p}_i$  instead of  $\bar{p}_{i_k}$ . Then,

$$\mathbf{1}^\top \mathbf{f}(\bar{\mathbf{y}}; \bar{p}_i) = \mathbf{1}^\top \mathbf{f}(\hat{\mathbf{y}}; \bar{p}_i) = \min_{(\mathbf{x}, \mathbf{y}) \in \mathbb{T}} \mathbf{1}^\top \mathbf{f}(\mathbf{y}; \bar{p}_i) \text{ with } 0 < \bar{p}_i < p^\sharp.$$

Let  $\|\hat{\mathbf{y}}\|_0 = s$ . From above we have

$$\|\bar{\mathbf{y}}\|_0 = \lim_{\bar{p}_i \downarrow 0} \mathbf{1}^\top \mathbf{f}(\bar{\mathbf{y}}; \bar{p}_i) = \lim_{\bar{p}_i \downarrow 0} \mathbf{1}^\top \mathbf{f}(\hat{\mathbf{y}}; \bar{p}_i) = \|\hat{\mathbf{y}}\|_0 = s,$$

which shows that  $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$  is also an optimal solution to problem (8). This contradicts to the afore-mentioned statement that  $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$  does not solve problem (8), which shows the assumption given in the beginning of the proof is false. The proof is complete.  $\square$

**Theorem 3.6** Consider sparse optimization problem (1) which is subject to the union of finite polyhedrons within a box  $\prod_{i=1}^{n_1} [l_i, u_i] \times \prod_{j=1}^{n_2} [0, 1]$ . There exists a constant  $p^* \in (0, 1]$  such that for any given real number  $p$  with  $0 < p < p^*$ , every optimal solution to nonconvex minimization problem (6) is an optimal solution to problem (1).

**Proof** Notice that problem (1) and problem (8) is equivalent, and problem (6) and problem (9) is equivalent. By virtue of Theorem 3.5, there exists a constant  $p^* \in (0, 1]$  such that for any given real number  $p$  with  $0 < p < p^*$ , every pseudo-extreme-point optimal solution to problem (6) is an optimal solution to problem (1).

To accomplish the proof, we divide the feasible set  $\mathbb{T}$  [defined by (2)] of problem (6) into some subsets. First, we observe that  $\mathbb{T}$  is composed of the following two subsets, one is concerned with all the (finite) pseudo-extreme points, the other is concerned with all the (feasible) non-pseudo-extreme points; these two subsets are denoted by  $\mathbb{T}_{pseudo}$  and  $\mathbb{T}_{non-pseudo}$ , respectively. Then,

$$\mathbb{T} = \mathbb{T}_{pseudo} \cup \mathbb{T}_{non-pseudo}. \tag{15}$$

Furthermore, we divide the two subsets in (15) into several smaller subsets, respectively. As to  $\mathbb{T}_{pseudo}$ , each pseudo-extreme point is either optimal solution to problem (6) or non-optimal feasible point; we denote the corresponding subsets by  $\mathbb{T}_{pseudo}^{opt}$  and  $\mathbb{T}_{pseudo}^{non-opt}$ , respectively, that is,

$$\mathbb{T}_{pseudo}^{opt} := \{(\hat{\mathbf{x}}_{opt}^i, \hat{\mathbf{y}}_{opt}^i)\}_{i \in I_1} \quad \text{and} \quad \mathbb{T}_{pseudo}^{non-opt} := \{(\hat{\mathbf{x}}_{non-opt}^j, \hat{\mathbf{y}}_{non-opt}^j)\}_{j \in I_2}, \tag{16}$$

where  $I_1$  and  $I_2$  are two finite index sets. Hence,

$$\mathbb{T}_{pseudo} = \mathbb{T}_{pseudo}^{opt} \cup \mathbb{T}_{pseudo}^{non-opt}. \tag{17}$$

As to  $\mathbb{T}_{non-pseudo}$ , composed of all the (feasible) non-pseudo-extreme points, we divide it specifically into three subsets:

$$\mathbb{T}_{non-pseudo} = \mathbb{T}_{non-pseudo}^{opt} \cup \mathbb{T}_{non-pseudo}^{non-opt} \cup \mathbb{T}_{non-pseudo}^{else}, \tag{18}$$

where  $\mathbb{T}_{non-pseudo}^{opt}$  stands for a feasible subset constructed from all the  $\mathbf{y}$ -blocks of all the pseudo-extreme point optimal solutions to problem (6):

$$\mathbb{T}_{non-pseudo}^{opt} := \bigcup_{i \in I_1} \{(\mathbf{x}, \hat{\mathbf{y}}_{opt}^i) : (\mathbf{x}, \hat{\mathbf{y}}_{opt}^i) \in \mathbb{T}, \mathbf{x} \neq \hat{\mathbf{x}}_{opt}^i\};$$

$\mathbb{T}_{non-pseudo}^{non-opt}$  stands for a feasible subset constructed from all the  $\mathbf{y}$ -blocks of all the non-optimal pseudo-extreme points of  $\mathbb{T}$ :

$$\mathbb{T}_{non-pseudo}^{non-opt} := \bigcup_{j \in I_2} \{(\mathbf{x}, \hat{\mathbf{y}}_{non-opt}^j) : (\mathbf{x}, \hat{\mathbf{y}}_{non-opt}^j) \in \mathbb{T}, \mathbf{x} \neq \hat{\mathbf{x}}_{non-opt}^j\};$$

and  $\mathbb{T}_{non-pseudo}^{else}$  stands for a feasible subset in which all the  $\mathbf{y}$ -blocks of all of its points are not from corresponding blocks of all the pseudo-extreme points of  $\mathbb{T}$ , that is,

$$\mathbb{T}_{non-pseudo}^{else} := \{(\mathbf{x}, \mathbf{y}) : (\mathbf{x}, \mathbf{y}) \in \mathbb{T}, \mathbf{y} \neq \hat{\mathbf{y}}_{opt}^i, \mathbf{y} \neq \hat{\mathbf{y}}_{non-opt}^j, i \in I_1, j \in I_2\}.$$

From (17) and (18), we observe that

$$\mathbb{T} = \mathbb{T}_{pseudo}^{opt} \cup \mathbb{T}_{pseudo}^{non-opt} \cup \mathbb{T}_{non-pseudo}^{opt} \cup \mathbb{T}_{non-pseudo}^{non-opt} \cup \mathbb{T}_{non-pseudo}^{else}. \tag{19}$$

Notice that if  $\mathbb{T}_{non-pseudo}^{opt}$  is nonempty, then all the points in  $\mathbb{T}_{non-pseudo}^{opt}$  are optimal solutions to problem (6), and furthermore, they are also optimal solutions to problem (1) since the optimal values of objective function of problem (6) is derived from all the  $\mathbf{y}$ -blocks of pseudo-extreme point optimal solutions. In addition, it is obvious that all the points in  $\mathbb{T}_{pseudo}^{non-opt}$  and  $\mathbb{T}_{non-pseudo}^{non-opt}$  are not optimal solutions to problem (6).

From Theorem 3.5, (19) and the discussions above, we observe that only the feasible subset  $\mathbb{T}_{non-pseudo}^{else}$  remains to be discussed. Suppose  $(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$  is any given (feasible) point in  $\mathbb{T}_{non-pseudo}^{else}$ . Since  $\tilde{\mathbf{y}} \neq \hat{\mathbf{y}}_{opt}^i, \tilde{\mathbf{y}} \neq \hat{\mathbf{y}}_{non-opt}^j, \forall i \in I_1, \forall j \in I_2$ , it is not a pseudo-extreme point. Therefore,  $(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$  can be represented as a convex combination of some (at least two) pseudo-extreme points, with at least two *distinct*  $\mathbf{y}$ -blocks. This can be written as

$$(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = \sum_{i \in \tilde{I}_1 \cup \tilde{I}_2} \lambda_i (\hat{\mathbf{x}}^i, \hat{\mathbf{y}}^i) = \left( \sum_{i \in \tilde{I}_1 \cup \tilde{I}_2} \lambda_i \hat{\mathbf{x}}^i, \sum_{i \in \tilde{I}_1 \cup \tilde{I}_2} \lambda_i \hat{\mathbf{y}}^i \right), \tag{20}$$

where  $\tilde{I}_1 \subseteq I_1, \tilde{I}_2 \subseteq I_2$ ; and we omit all the terms that have zero coefficients. Hence, each  $\lambda_i > 0$  and  $\sum_{i \in \tilde{I}_1 \cup \tilde{I}_2} \lambda_i = 1$ . Let

$$\mathbf{1}^\top \mathbf{f}(\hat{\mathbf{y}}^{i_0}; p) = \min_{i \in \tilde{I}_1 \cup \tilde{I}_2} \mathbf{1}^\top \mathbf{f}(\hat{\mathbf{y}}^i; p). \tag{21}$$

Notice that every  $(\hat{\mathbf{x}}^i, \hat{\mathbf{y}}^i)$  is a pseudo-extreme point of  $\mathbb{T}$  and that at least two pseudo-extreme points have *distinct*  $\mathbf{y}$ -blocks, we have

$$\begin{aligned} \mathbf{1}^\top \mathbf{f}(\tilde{\mathbf{y}}; p) &= \mathbf{1}^\top \mathbf{f} \left( \sum_{i \in \tilde{I}_1 \cup \tilde{I}_2} \lambda_i \hat{\mathbf{y}}^i; p \right) = \sum_{j=1}^{n_2} f \left( \sum_{i \in \tilde{I}_1 \cup \tilde{I}_2} \lambda_i \hat{\mathbf{y}}_j^i; p \right) \\ &> \sum_{j=1}^{n_2} \sum_{i \in \tilde{I}_1 \cup \tilde{I}_2} \lambda_i f(\hat{\mathbf{y}}_j^i; p) = \sum_{i \in \tilde{I}_1 \cup \tilde{I}_2} \lambda_i \sum_{j=1}^{n_2} f(\hat{\mathbf{y}}_j^i; p) \\ &= \sum_{i \in \tilde{I}_1 \cup \tilde{I}_2} \lambda_i \mathbf{1}^\top \mathbf{f}(\hat{\mathbf{y}}^i; p) \geq \sum_{i \in \tilde{I}_1 \cup \tilde{I}_2} \lambda_i \mathbf{1}^\top \mathbf{f}(\hat{\mathbf{y}}^{i_0}; p) \\ &= \mathbf{1}^\top \mathbf{f}(\hat{\mathbf{y}}^{i_0}; p) \sum_{i \in \tilde{I}_1 \cup \tilde{I}_2} \lambda_i = \mathbf{1}^\top \mathbf{f}(\hat{\mathbf{y}}^{i_0}; p) \\ &\geq \min_{(\mathbf{x}, \mathbf{y}) \in \mathbb{T}} \mathbf{1}^\top \mathbf{f}(\mathbf{y}; p), \end{aligned} \tag{22}$$

where the first strict inequality follows from the strict concavity of  $f(\cdot; p)$  on  $\mathbb{S}$ ; the second inequality from (21); and the last inequality is obvious.

From (22), we have  $\mathbf{1}^\top \mathbf{f}(\tilde{\mathbf{y}}; p) > \min_{(\mathbf{x}, \mathbf{y}) \in \mathbb{T}} \mathbf{1}^\top \mathbf{f}(\mathbf{y}; p)$ , which implies that point  $(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$  is not an optimal solution to problem (6). Since  $(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$  is arbitrary in  $\mathbb{T}_{non-pseudo}^{else}$ , we see that all the points in  $\mathbb{T}_{non-pseudo}^{else}$  are not optimal solutions to problem (6).

By all the discussions above, we derive that each optimal solution to problem (6) with  $0 < p < p^*$  must have a  $\mathbf{y}$ -block being the corresponding block of some pseudo-extreme point optimal solution to problem (6). This demonstrates that each optimal solution to problem (6) with  $0 < p < p^*$  must be an optimal solution to problem (1), since the optimal values of the objective functions of these both problems are all determined by the  $\mathbf{y}$ -block of each pseudo-extreme-point optimal solution of problem (6).

The proof is complete. □

### 4 Several important subclasses of problem (1)

Problem (1) covers a wide range of problems. In this section, we give several important special cases of it.

#### 4.1 Partially sparse solutions of equalities and inequalities

Consider the following partially sparse optimization problem

$$\min_{(\mathbf{x}, \mathbf{y})} \|\mathbf{y}\|_0 \quad \text{s.t.} \quad \mathbf{Ax} + \mathbf{By} = \mathbf{b}, \mathbf{Cx} + \mathbf{Dy} \leq \mathbf{c}, \mathbf{l} \leq \mathbf{x} \leq \mathbf{u}, \mathbf{0} \leq \mathbf{y} \leq \mathbf{1}, \quad (23)$$

where  $A \in \mathbb{R}^{m_1 \times n_1}$ ,  $B \in \mathbb{R}^{m_1 \times n_2}$ ,  $C \in \mathbb{R}^{m_2 \times n_1}$ ,  $D \in \mathbb{R}^{m_2 \times n_2}$ ,  $\mathbf{b} \in \mathbb{R}^{m_1}$ ,  $\mathbf{c} \in \mathbb{R}^{m_2}$ ,  $\mathbf{l}, \mathbf{u} \in \mathbb{R}^{n_1}$  satisfying  $\mathbf{l} \leq \mathbf{u}$ , and  $\mathbf{x} \in \mathbb{R}^{n_1}$ ,  $\mathbf{y} \in \mathbb{R}^{n_2}$ ,  $(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^N$  with  $N = n_1 + n_2$ .

We have the following result.

**Theorem 4.1** *Suppose that  $f(t; p) \in \mathcal{F}(t; p)$ . We consider sparse optimization problem (23) and its nonconvex relaxation problem*

$$\min_{(\mathbf{x}, \mathbf{y})} \sum_{i=1}^{n_2} f(y_i; p) \quad \text{s.t.} \quad \mathbf{Ax} + \mathbf{By} = \mathbf{b}, \mathbf{Cx} + \mathbf{Dy} \leq \mathbf{c}, \mathbf{l} \leq \mathbf{x} \leq \mathbf{u}, \mathbf{0} \leq \mathbf{y} \leq \mathbf{1}. \quad (24)$$

*It follows that there exists a constant  $p^* \in (0, 1]$  such that for an arbitrarily given real number  $p$  with  $0 < p < p^*$ , every optimal solution to nonconvex minimization problem (24) is an optimal solution to (23).*

**Proof** Let  $\mathbb{P} = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^N : \mathbf{Ax} + \mathbf{By} = \mathbf{b}, \mathbf{Cx} + \mathbf{Dy} \leq \mathbf{c}\}$ . Then, problem (23) can be written as

$$\min_{(\mathbf{x}, \mathbf{y})} \|\mathbf{y}\|_0 \quad \text{s.t.} \quad (\mathbf{x}, \mathbf{y}) \in \mathbb{P}, \mathbf{l} \leq \mathbf{x} \leq \mathbf{u}, \mathbf{0} \leq \mathbf{y} \leq \mathbf{1}, \quad (25)$$

It is obvious that  $\mathbb{P}$  is a polyhedral in  $\mathbb{R}^N$ . Thus, problem (25) is a special case of problem (1). By Theorem 3.6, the theorem holds immediately. □

In the following, we further consider two special cases of problem (23).

**Problem 1** We consider the problem of Compressive Sensing (see, e.g., Candés et al. 2006; Candés and Tao 2005; Foucart and Rauhut 2013):

$$\min_{\mathbf{z}} \|\mathbf{z}\|_0 \quad \text{s.t.} \quad \Phi \mathbf{z} = \mathbf{b}, \tag{26}$$

where  $\Phi \in \mathbb{R}^{m \times n}$  with  $m \ll n$  is of full row rank, i.e.,  $\text{rank}(\Phi) = m$ ,  $\mathbf{b} \in \mathbb{R}^m$  and  $\mathbf{z} \in \mathbb{R}^n$ . By You et al. (2017, Theorem 3.1), there is a number  $r_0 > 0$  such that problem (26) can be reformulated as

$$\min_{\mathbf{z}} \|\mathbf{z}\|_0 \quad \text{s.t.} \quad \Phi \mathbf{z} = \mathbf{b}, \|\mathbf{z}\|_\infty \leq r_0. \tag{27}$$

Let  $\mathbf{x} = \frac{1}{r_0} \mathbf{z}$  and  $A = r_0 \Phi$ , then problem (27) can be equivalently transformed to

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{s.t.} \quad A \mathbf{x} = \mathbf{b}, \|\mathbf{x}\|_\infty \leq 1. \tag{28}$$

Furthermore, we introduce a variable  $\mathbf{y} \in \mathbb{R}^n$  such that  $|\mathbf{x}| \leq \mathbf{y}$ ; and define two-block matrices  $C := (-E, E)^\top$  and  $D := (-E, -E)^\top$  where  $E \in \mathbb{R}^{n \times n}$  is the  $n \times n$  identity matrix, then it is easy to verify that  $|\mathbf{x}| \leq \mathbf{y}$  if and only if  $C \mathbf{x} + D \mathbf{y} \leq \mathbf{0}$ . Take  $B = \mathbf{0}$ , i.e.,  $B$  denotes an  $m$  by  $n$  zero matrix. Then, by You et al. (2017, Corollary 2.1), problem (28) can be written as

$$\min_{(\mathbf{x}, \mathbf{y})} \|\mathbf{y}\|_0 \quad \text{s.t.} \quad A \mathbf{x} + B \mathbf{y} = \mathbf{b}, C \mathbf{x} + D \mathbf{y} \leq \mathbf{0}, -\mathbf{1} \leq \mathbf{x} \leq \mathbf{1}, \mathbf{0} \leq \mathbf{y} \leq \mathbf{1}. \tag{29}$$

It is easy to see that this problem is a special case of problem (23). To relax problem (29), we assume  $f(t; p) \in \mathcal{F}(t; p)$  and consider the following problem:

$$\min_{(\mathbf{x}, \mathbf{y})} \mathbf{1}^\top \mathbf{f}(\mathbf{y}; p) \quad \text{s.t.} \quad A \mathbf{x} + B \mathbf{y} = \mathbf{b}, C \mathbf{x} + D \mathbf{y} \leq \mathbf{0}, -\mathbf{1} \leq \mathbf{x} \leq \mathbf{1}, \mathbf{0} \leq \mathbf{y} \leq \mathbf{1}, \tag{30}$$

where  $\mathbf{f}(\mathbf{y}; p) = (f(y_1; p), \dots, f(y_n; p))^\top$ . Therefore, from Theorem 4.1 it follows that there exists a constant  $p^* \in (0, 1]$  such that for an arbitrarily given real number  $p$  with  $0 < p < p^*$ , every optimal solution to nonconvex minimization problem (30) is an optimal solution to problem (29).

Noting that we have shown that problem (29) is equivalent to problem (26). By using a similar way we can obtain that problem (30) is equivalent to

$$\min_{\mathbf{z}} \mathbf{1}^\top \mathbf{f} \left( \frac{1}{r_0} |\mathbf{z}|; p \right) \quad \text{s.t.} \quad \Phi \mathbf{z} = \mathbf{b}, \|\mathbf{z}\|_\infty \leq r_0, \tag{31}$$

where  $r_0$  is given by (27). Thus, we can further obtain the following result.

**Theorem 4.2** *There exist a positive number  $r_0$  and a constant  $p^* \in (0, 1]$  such that for an arbitrarily given real number  $p$  with  $0 < p < p^*$ , every optimal solution to nonconvex minimization problem (31) is an optimal solution to problem (26).*

In particular, if we take  $f(t; p) = t^p \in \mathcal{F}(t; p)$ , we can further obtain that problem (31) is equivalent to

$$\min_{\mathbf{z}} \|\mathbf{z}\|_p^p \quad \text{s.t.} \quad \Phi \mathbf{z} = \mathbf{b}. \tag{32}$$

Thus, from Theorem 4.2 it follows that there exists a constant  $p^* \in (0, 1]$  such that for an arbitrarily given real number  $p$  with  $0 < p < p^*$ , every optimal solution to nonconvex minimization problem (32) is an optimal solution to problem (26). Such a result has been obtained in Peng et al. (2015).

**Problem 2** We consider sparsest solutions of linear equalities and inequalities (see Fung and Mangasarian 2011):

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{s.t.} \quad A\mathbf{x} = \mathbf{b}, F\mathbf{x} \leq \mathbf{d}, \|\mathbf{x}\|_\infty \leq 1, \tag{33}$$

where  $A \in \mathbb{R}^{m \times n}$ ,  $\mathbf{b} \in \mathbb{R}^m$ ,  $F \in \mathbb{R}^{k \times n}$ ,  $\mathbf{d} \in \mathbb{R}^k$ , and  $\mathbf{x} \in \mathbb{R}^n$ . We introduce a variable  $\mathbf{y} \in \mathbb{R}^n$  such that  $|\mathbf{x}| \leq \mathbf{y}$  and define  $\mathbf{c} := (\mathbf{d}^\top, \mathbf{0}^\top, \mathbf{0}^\top)^\top$ ,  $C := (F^\top, -E, E)^\top$  and  $D := (O^\top, -E, -E)^\top$ , where  $O$  is a  $k \times n$  zero matrix and  $E \in \mathbb{R}^{n \times n}$  is the  $n \times n$  identity matrix, then it is easy to verify that both  $F\mathbf{x} \leq \mathbf{d}$  and  $|\mathbf{x}| \leq \mathbf{y}$  hold if and only if  $C\mathbf{x} + D\mathbf{y} \leq \mathbf{c}$ . Let  $B = \mathbf{O}$ , i.e.,  $B$  denotes an  $m$  by  $n$  zero matrix. Then, by You et al. (2017, Corollary 2.1), problem (33) can be written as

$$\min_{(\mathbf{x}, \mathbf{y})} \|\mathbf{y}\|_0 \quad \text{s.t.} \quad A\mathbf{x} + B\mathbf{y} = \mathbf{b}, C\mathbf{x} + D\mathbf{y} \leq \mathbf{c}, -\mathbf{1} \leq \mathbf{x} \leq \mathbf{1}, \mathbf{0} \leq \mathbf{y} \leq \mathbf{1}. \tag{34}$$

It is easy to observe that this problem is a special case of problem (23).

To relax problem (34), we assume  $f(t; p) \in \mathcal{F}(t; p)$  and consider the nonconvex relaxation of problem (34) as follows

$$\min_{(\mathbf{x}, \mathbf{y})} \mathbf{1}^\top \mathbf{f}(\mathbf{y}; p) \quad \text{s.t.} \quad A\mathbf{x} + B\mathbf{y} = \mathbf{b}, C\mathbf{x} + D\mathbf{y} \leq \mathbf{c}, -\mathbf{1} \leq \mathbf{x} \leq \mathbf{1}, \mathbf{0} \leq \mathbf{y} \leq \mathbf{1}, \tag{35}$$

where  $\mathbf{f}(\mathbf{y}; p) = (f(y_1; p), \dots, f(y_n; p))^\top$ . Therefore, from Theorem 3.6 it follows that there exists a constant  $p^* \in (0, 1]$  such that for an arbitrarily given real number  $p$  with  $0 < p < p^*$ , every optimal solution to nonconvex minimization problem (35) is an optimal solution to problem (34).

By You et al. (2017, Corollary 2.1), we observe that problem (35) is equivalent to

$$\min_{\mathbf{x}} \mathbf{1}^\top \mathbf{f}(|\mathbf{x}|; p) \quad \text{s.t.} \quad A\mathbf{x} = \mathbf{b}, F\mathbf{x} \leq \mathbf{d}, \|\mathbf{x}\|_\infty \leq 1, \tag{36}$$

and notice that (34) is equivalent to (33). Therefore, we derive

**Theorem 4.3** *There exists a constant  $p^* \in (0, 1]$  such that for an arbitrarily given real number  $p$  with  $0 < p < p^*$ , every optimal solution to nonconvex relaxation problem (36) is an optimal solution to (33).*

In particular, if we take  $f(t; p) = t^p \in \mathcal{F}(t; p)$ , then the corresponding result was obtained in Fung and Mangasarian (2011, Proposition 1).

### 4.2 Partially sparsest solutions of absolute equalities and inequalities

Consider the following partially sparse optimization problem

$$\min_{(\mathbf{x}, \mathbf{y})} \|\mathbf{y}\|_0 \quad \text{s.t. } |\mathbf{Ax} + \mathbf{By}| = \mathbf{b}, \mathbf{Cx} + \mathbf{Dy} \leq \mathbf{c}, \mathbf{l} \leq \mathbf{x} \leq \mathbf{u}, \mathbf{0} \leq \mathbf{y} \leq \mathbf{1}, \quad (37)$$

where  $A \in \mathbb{R}^{m_1 \times n_1}, B \in \mathbb{R}^{m_1 \times n_2}, C \in \mathbb{R}^{m_2 \times n_1}, D \in \mathbb{R}^{m_2 \times n_2}, \mathbf{0} \leq \mathbf{b} \in \mathbb{R}^{m_1}, \mathbf{c} \in \mathbb{R}^{m_2}, \mathbf{l}, \mathbf{u} \in \mathbb{R}^{n_1}$  satisfying  $\mathbf{l} \leq \mathbf{u}$ , and  $(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^N = \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ .

We have the following result.

**Theorem 4.4** *Assume  $f(t; p) \in \mathcal{F}(t; p)$ . We consider sparse optimization problem (37) and its nonconvex relaxation problem*

$$\min_{(\mathbf{x}, \mathbf{y})} \sum_{i=1}^{n_2} f(y_i; p) \quad \text{s.t. } |\mathbf{Ax} + \mathbf{By}| = \mathbf{b}, \mathbf{Cx} + \mathbf{Dy} \leq \mathbf{c}, \mathbf{l} \leq \mathbf{x} \leq \mathbf{u}, \mathbf{0} \leq \mathbf{y} \leq \mathbf{1}. \quad (38)$$

It follows that there exists a constant  $p^* \in (0, 1]$  such that for an arbitrarily given real number  $p$  with  $0 < p < p^*$ , every optimal solution to nonconvex relaxation problem (38) is an optimal solution to (37).

**Proof** Let

$$\mathbb{T} := \left( \bigcup_{\epsilon \in \{-1, 1\}^{m_1}} \mathbb{P}_\epsilon \right) \cap \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^N : \mathbf{l} \leq \mathbf{x} \leq \mathbf{u}, \mathbf{0} \leq \mathbf{y} \leq \mathbf{1}\}, \quad (39)$$

where each  $\mathbb{P}_\epsilon := \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^N : \mathbf{Ax} + \mathbf{By} = \epsilon \circ \mathbf{b}, \mathbf{Cx} + \mathbf{Dy} \leq \mathbf{c}\}$  is obviously a polyhedral in  $\mathbb{R}^N$ . Hence, (37) can be rewritten as

$$\min_{(\mathbf{x}, \mathbf{y})} \|\mathbf{y}\|_0 \quad \text{s.t. } (\mathbf{x}, \mathbf{y}) \in \bigcup_{\epsilon \in \{-1, 1\}^{m_1}} \mathbb{P}_\epsilon, \mathbf{l} \leq \mathbf{x} \leq \mathbf{u}, \mathbf{0} \leq \mathbf{y} \leq \mathbf{1}. \quad (40)$$

Thus, problem (37) is a special case of problem (1). By Theorem 3.6, the theorem holds immediately.  $\square$

In the following, we further consider a special case of problem (37).

**Problem 3** We consider the problem of Phaseless Compressive Sensing (see, e.g., Voroninski and Xu 2016):

$$\min_{\mathbf{z}} \|\mathbf{z}\|_0 \quad \text{s.t. } |\Phi \mathbf{z}| = \mathbf{b}, \quad (41)$$

where  $\Phi \in \mathbb{R}^{m \times n}$  with  $m \ll n$  is of full row rank, i.e.,  $\text{rank}(\Phi) = m$ ,  $\mathbf{b} \in \mathbb{R}^m$  and  $\mathbf{z} \in \mathbb{R}^n$ . By You et al. (2017, Theorem 3.2), there is a number  $r_0 > 0$  such that problem (41) can be reformulated as

$$\min_{\mathbf{z}} \|\mathbf{z}\|_0 \quad \text{s.t.} \quad |\Phi\mathbf{z}| = \mathbf{b}, \|\mathbf{z}\|_\infty \leq r_0. \tag{42}$$

Let  $\mathbf{x} = \frac{1}{r_0}\mathbf{z}$  and  $A = r_0\Phi$ , then problem (42) can be equivalently transformed to

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{s.t.} \quad |\mathbf{Ax}| = \mathbf{b}, \|\mathbf{x}\|_\infty \leq 1. \tag{43}$$

Furthermore, we introduce a variable  $\mathbf{y} \in \mathbb{R}^n$  such that  $|\mathbf{x}| \leq \mathbf{y}$ ; and define two-block matrices  $C := (-E, E)^\top$  and  $D := (-E, -E)^\top$  where  $E \in \mathbb{R}^{n \times n}$  is the  $n \times n$  identity matrix, then it is easy to verify that  $|\mathbf{x}| \leq \mathbf{y}$  if and only if  $C\mathbf{x} + D\mathbf{y} \leq \mathbf{0}$ . Take  $B = \mathbf{0}$ , i.e.,  $B$  denotes an  $m$  by  $n$  zero matrix. Then, by You et al. (2017, Corollary 2.1), problem (28) can be written as

$$\min_{(\mathbf{x}, \mathbf{y})} \|\mathbf{y}\|_0 \quad \text{s.t.} \quad |\mathbf{Ax} + B\mathbf{y}| = \mathbf{b}, C\mathbf{x} + D\mathbf{y} \leq \mathbf{0}, -\mathbf{1} \leq \mathbf{x} \leq \mathbf{1}, \mathbf{0} \leq \mathbf{y} \leq \mathbf{1}. \tag{44}$$

It is easy to see that this problem is a special case of problem (37). To relax problem (44), we assume  $f(t; p) \in \mathcal{F}(t; p)$  and consider the following problem:

$$\min_{(\mathbf{x}, \mathbf{y})} \mathbf{1}^\top \mathbf{f}(\mathbf{y}; p) \quad \text{s.t.} \quad |\mathbf{Ax} + B\mathbf{y}| = \mathbf{b}, C\mathbf{x} + D\mathbf{y} \leq \mathbf{0}, -\mathbf{1} \leq \mathbf{x} \leq \mathbf{1}, \mathbf{0} \leq \mathbf{y} \leq \mathbf{1}, \tag{45}$$

where  $\mathbf{f}(\mathbf{y}; p) = (f(y_1; p), \dots, f(y_n; p))^\top$ . Therefore, from Theorem 4.4 it follows that there exists a constant  $p^* \in (0, 1]$  such that for an arbitrarily given real number  $p$  with  $0 < p < p^*$ , every optimal solution to nonconvex minimization problem (45) is an optimal solution to problem (44).

Noting that we have shown that problem (44) is equivalent to problem (41). By using a similar way we can obtain that problem (45) is equivalent to

$$\min_{\mathbf{z}} \mathbf{1}^\top \mathbf{f}\left(\frac{1}{r_0}|\mathbf{z}|; p\right) \quad \text{s.t.} \quad |\Phi\mathbf{z}| = \mathbf{b}, \|\mathbf{z}\|_\infty \leq r_0, \tag{46}$$

where  $r_0$  is given by (42).

Thus, we can further obtain the following result.

**Theorem 4.5** *There exist a positive number  $r_0$  and a constant  $p^* \in (0, 1]$  such that for an arbitrarily given real number  $p$  with  $0 < p < p^*$ , every optimal solution to nonconvex minimization problem (46) is an optimal solution to problem (41).*

In particular, if we take  $f(t; p) = t^p \in \mathcal{F}(t; p)$ , we can further obtain that problem (46) is equivalent to

$$\min_{\mathbf{z}} \|\mathbf{z}\|_p^p \quad \text{s.t.} \quad |\Phi\mathbf{z}| = \mathbf{b}. \tag{47}$$

Thus, from Theorem 4.5 it follows that there exists a constant  $p^* \in (0, 1]$  such that for an arbitrarily given real number  $p$  with  $0 < p < p^*$ , every optimal solution to nonconvex minimization problem (47) is an optimal solution to problem (41). Such a result has been obtained in You et al. (2017).



### 4.3 Sparsest solutions of absolute equalities and inequalities

In this subsection, we consider a class of problems that find the sparsest solutions defined as follows. Let  $\Phi \in \mathbb{R}^{m \times n}$ ,  $\mathbf{b} \in \mathbb{R}^m$ ,  $\mathbf{d} \in \mathbb{R}^s$ ,  $P \in \mathbb{R}^{s \times n}$ ,  $\mathbf{l} := (l_1, \dots, l_n)^\top$  and  $\mathbf{u} := (u_1, \dots, u_n)^\top$ . We consider

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{s.t.} \quad |\Phi\mathbf{x}| = \mathbf{b}, P\mathbf{x} \leq \mathbf{d}, \mathbf{l} \leq \mathbf{x} \leq \mathbf{u}. \tag{48}$$

We do not care two cases:  $\mathbf{0} < \mathbf{l} < \mathbf{u}$  or  $\mathbf{l} < \mathbf{u} < \mathbf{0}$ , since, if so, all the feasible points are solutions with objective  $\min \|\mathbf{x}\|_0 = n$ . For convenience, we only consider the above problem when  $\mathbf{l} < \mathbf{u}$  with any one of four special cases as follows: (i)  $\mathbf{l} = \mathbf{0}$  and  $\mathbf{u} > \mathbf{0}$ , (ii)  $\mathbf{l} \geq \mathbf{0}$  and  $\mathbf{u} > \mathbf{0}$ , (iii)  $\mathbf{l} < \mathbf{0}$  and  $\mathbf{u} \leq \mathbf{0}$  and (iv)  $\mathbf{l} < \mathbf{0} < \mathbf{u}$ .

**Proposition 4.1** *For any given case defined above, problem (48) can be equivalently transformed to a special case of problem (37).*

**Proof** We divide the proof into four parts.

(i)  $\mathbf{l} = \mathbf{0}$  and  $\mathbf{u} > \mathbf{0}$ . In this case, the problem (48) is

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{s.t.} \quad |\Phi\mathbf{x}| = \mathbf{b}, P\mathbf{x} \leq \mathbf{d}, \mathbf{0} \leq \mathbf{x} \leq \mathbf{u}. \tag{49}$$

Let  $\Lambda = \text{Diag}(\mathbf{u})$ ,  $\mathbf{y} = \Lambda^{-1}\mathbf{x}$ ,  $B = \Phi\Lambda$ ,  $D = P\Lambda$ . Then, (49) is written as

$$\min_{\mathbf{y}} \|\mathbf{y}\|_0 \quad \text{s.t.} \quad |B\mathbf{y}| = \mathbf{b}, D\mathbf{y} \leq \mathbf{d}, \mathbf{0} \leq \mathbf{y} \leq \mathbf{1},$$

which can be regarded as a special case of problem (37).

(ii)  $\mathbf{l} \geq \mathbf{0}$  and  $\mathbf{u} > \mathbf{0}$ . Define an index set  $\mathbb{I} := \{i \mid l_i = 0, i \in [n]\}$ . Then, we have  $\Phi\mathbf{x} = \Phi_{\mathbb{I}}\mathbf{x}_{\mathbb{I}} + \Phi_{\mathbb{I}^c}\mathbf{x}_{\mathbb{I}^c}$  and  $P\mathbf{x} = P_{\mathbb{I}}\mathbf{x}_{\mathbb{I}} + P_{\mathbb{I}^c}\mathbf{x}_{\mathbb{I}^c}$ , and hence, (48) can be rewritten as

$$\begin{aligned} \min_{\mathbf{x}} \|\mathbf{x}_{\mathbb{I}}\|_0 \\ \text{s.t.} \quad |\Phi_{\mathbb{I}}\mathbf{x}_{\mathbb{I}} + \Phi_{\mathbb{I}^c}\mathbf{x}_{\mathbb{I}^c}| = \mathbf{b}, P_{\mathbb{I}}\mathbf{x}_{\mathbb{I}} + P_{\mathbb{I}^c}\mathbf{x}_{\mathbb{I}^c} \leq \mathbf{d}, \mathbf{0} \leq \mathbf{x}_{\mathbb{I}} \leq \mathbf{u}_{\mathbb{I}}, \mathbf{l}_{\mathbb{I}^c} \leq \mathbf{x}_{\mathbb{I}^c} \leq \mathbf{u}_{\mathbb{I}^c}, \end{aligned} \tag{50}$$

which is a special case of problem (4). By Remark 2.1, (50) can be equivalently transformed to a special case of (1), more specifically, to (37).

(iii)  $\mathbf{l} < \mathbf{0}$  and  $\mathbf{u} \leq \mathbf{0}$ . In this case, let  $\bar{\mathbf{x}} = -\mathbf{x}$ ,  $\bar{\Phi} = -\Phi$ ,  $\bar{P} = -P$ ,  $\bar{\mathbf{l}} = -\mathbf{u}$  and  $\bar{\mathbf{u}} = -\mathbf{l}$ , then  $\bar{\mathbf{l}} \geq \mathbf{0}$  and  $\bar{\mathbf{u}} > \mathbf{0}$ , and hence, (48) can be rewritten as

$$\min \|\bar{\mathbf{x}}\|_0 \quad \text{s.t.} \quad |\bar{\Phi}\bar{\mathbf{x}}| = \mathbf{b}, \bar{P}\bar{\mathbf{x}} \leq \mathbf{d}, \bar{\mathbf{l}} \leq \bar{\mathbf{x}} \leq \bar{\mathbf{u}},$$

which is the case (ii).

(iv)  $\mathbf{l} < \mathbf{0} < \mathbf{u}$ . In this case, we introduce a variable  $\mathbf{z} \in \mathbb{R}^n$  such that  $|\mathbf{x}| \leq \mathbf{z}$ . Let  $\mathbf{v} = \max\{-\mathbf{l}, \mathbf{u}\}$ . Since that  $\text{sign}(\mathbf{z})$  is monotonically nondecreasing, noting that

objective function  $\|\mathbf{z}\|_0 = \mathbf{1}^\top \mathbf{sign}(\mathbf{z})$ , by You et al. (2017, Corollary 2.1), we have the following equivalent form:

$$\min_{(\mathbf{x}, \mathbf{z})} \|\mathbf{z}\|_0 \quad \text{s.t.} \quad |\Phi \mathbf{x}| = \mathbf{b}, P\mathbf{x} \leq \mathbf{d}, |\mathbf{x}| \leq \mathbf{z}, \mathbf{l} \leq \mathbf{x} \leq \mathbf{u} \text{ and } \mathbf{0} \leq \mathbf{z} \leq \mathbf{v}. \quad (51)$$

Let  $\Psi$  be a zero matrix,  $\Theta = (-E, E, P^\top)^\top$ ,  $\Gamma = (-E, E, O)^\top$  with  $E$  being an  $n$  by  $n$  Identity matrix and  $O$  being an  $n$  by  $s$  zero matrix, and  $\mathbf{c} = (\mathbf{0}^\top, \mathbf{0}^\top, \mathbf{d}^\top)^\top$ . Then, problem (51) can be rewritten as

$$\min_{(\mathbf{x}, \mathbf{z})} \|\mathbf{z}\|_0 \quad \text{s.t.} \quad |\Phi \mathbf{x} + \Psi \mathbf{z}| = \mathbf{b}, \Theta \mathbf{x} + \Gamma \mathbf{z} \leq \mathbf{c}, \mathbf{l} \leq \mathbf{x} \leq \mathbf{u}, \mathbf{0} \leq \mathbf{z} \leq \mathbf{v}, \quad (52)$$

By Remark 2.1, (52) can be equivalently transformed to a special case of (1), more specifically, to (37).

Combining parts (i)–(iv), we complete the proof. □

By Proposition 4.1, Theorem 4.4 and You et al. (2017, Corollary 2.1), with notation replacements if necessary, we can obtain the following results.

**Theorem 4.6** *Consider problem (48) with bounds  $\mathbf{l} < \mathbf{u}$  satisfying one of four special cases defined above. Let  $\mathbb{I} := \{i : l_i u_i \leq 0, i \in [n]\}$ ,  $\mathbf{v} := \max\{|\mathbf{l}_{\mathbb{I}}|, |\mathbf{u}_{\mathbb{I}}|\}$  and  $\Lambda := \text{Diag}(\mathbf{v})$ . Then, for the following nonconvex minimization problem*

$$\min_{\mathbf{x}} \mathbf{1}^\top \mathbf{f}(\Lambda^{-1}|\mathbf{x}_{\mathbb{I}}|; p) \quad \text{s.t.} \quad |\Phi \mathbf{x}| = \mathbf{b}, P\mathbf{x} \leq \mathbf{d}, \mathbf{l} \leq \mathbf{x} \leq \mathbf{u}, \quad (53)$$

where  $\mathbf{f}(\mathbf{y}; p) = (f(y_i; p)_{i \in \mathbb{I}})^\top$  with any given  $f(t; p) \in \mathcal{F}(t; p)$ , there exists a constant  $p^* \in (0, 1]$  such that for any given real number  $p$  with  $0 < p < p^*$ , every optimal solution to nonconvex minimization problem (53) is an optimal solution to problem (48).

Next, we take an example to illustrate the validity of Theorem 4.6.

**Example 4.1** We consider the 2-dimensional problem:

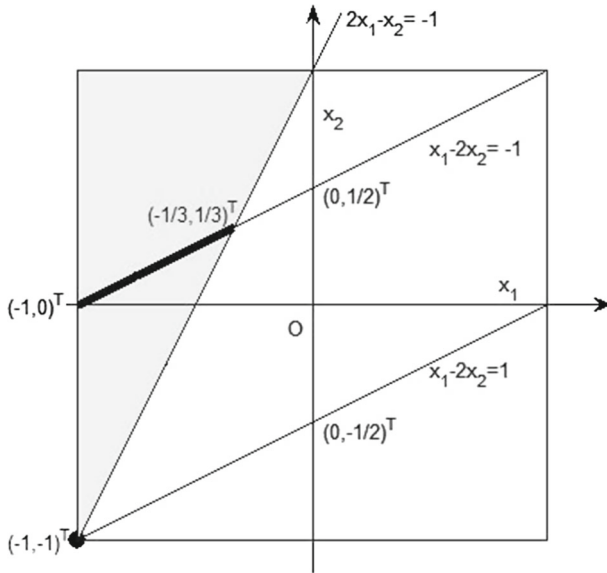
$$\min_{\mathbf{x} \in \mathbb{R}^2} \|\mathbf{x}\|_0 \quad \text{s.t.} \quad |x_1 - 2x_2| = 1, 2x_1 - x_2 \leq -1, \|\mathbf{x}\|_\infty \leq 1, \quad (54)$$

where  $\mathbf{x} = (x_1, x_2)^\top$ .

We apply two specific nonconvex functions and consider the corresponding nonconvex relaxation problems to confirm the result obtained in Theorem 4.6.

**Case 1.** We take  $f(t; p) = t^p \in \mathcal{F}(t; p)$ , and consider nonconvex relaxation problem

$$\min_{\mathbf{x} \in \mathbb{R}^2} \|\mathbf{x}\|_p^p \quad \text{s.t.} \quad |x_1 - 2x_2| = 1, 2x_1 - x_2 \leq -1, \|\mathbf{x}\|_\infty \leq 1. \quad (55)$$



**Fig. 1** All feasible points of problem (54): point  $(-1, -1)^T$  and points on the bold line segment, with the point  $(-1, 0)^T$  being the (sparsest) optimal solution to the original problem (54)

The equality constraint implies  $x_1 - 2x_2 = -1$  or  $x_1 - 2x_2 = 1$ . Accordingly, the feasible set (see Fig. 1), denoted by  $\mathbb{D}$ , can be written as  $\mathbb{D} = \mathbb{D}_1 \cup \mathbb{D}_2$ , where

$$\mathbb{D}_1 = \{(x_1, x_2)^T \mid x_1 - 2x_2 = -1, 2x_1 - x_2 \leq -1, -1 \leq x_i \leq 1, i = 1, 2\}$$

and

$$\mathbb{D}_2 = \{(x_1, x_2)^T \mid x_1 - 2x_2 = 1, 2x_1 - x_2 \leq -1, -1 \leq x_i \leq 1, i = 1, 2\}.$$

Note that the objective function  $\|\mathbf{x}\|_p^p = |x_1|^p + |x_2|^p$  with  $0 < p < 1$ . It follows that problem (55) is equivalent to

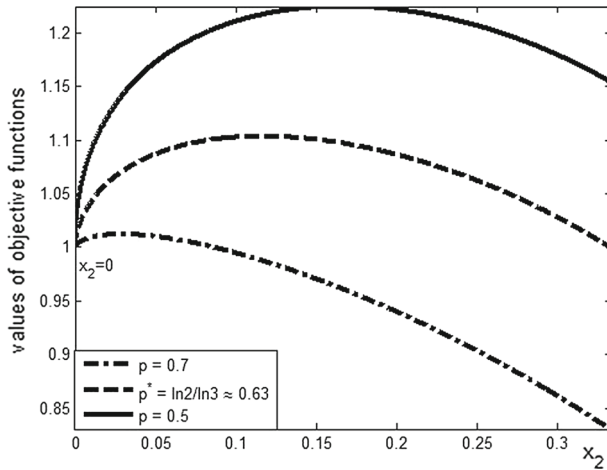
$$\min \left\{ \min_{\mathbf{x} \in \mathbb{D}_1} |x_1|^p + |x_2|^p, \min_{\mathbf{x} \in \mathbb{D}_2} |x_1|^p + |x_2|^p \right\}. \tag{56}$$

Simple computations implies that

$$\mathbb{D}_1 = \left\{ (2x_2 - 1, x_2)^T \mid 0 \leq x_2 \leq \frac{1}{3} \right\} \quad \text{and} \quad \mathbb{D}_2 = \left\{ (-1, -1)^T \right\}.$$

Therefore, (56) can be rewritten as

$$\min \left\{ \min_{0 \leq x_2 \leq \frac{1}{3}} |2x_2 - 1|^p + |x_2|^p, 2 \right\}$$



**Fig. 2** Objective functions of (58), being all concave over closed interval  $[0, \frac{1}{3}]$ ; among three specific  $p$  values,  $p^*$  is the constant that is claimed in our main results in Sect. 3

$$= \min \left\{ \min_{0 \leq x_2 \leq \frac{1}{3}} (1 - 2x_2)^p + x_2^p, 2 \right\}. \tag{57}$$

Consider the subproblem in (57):

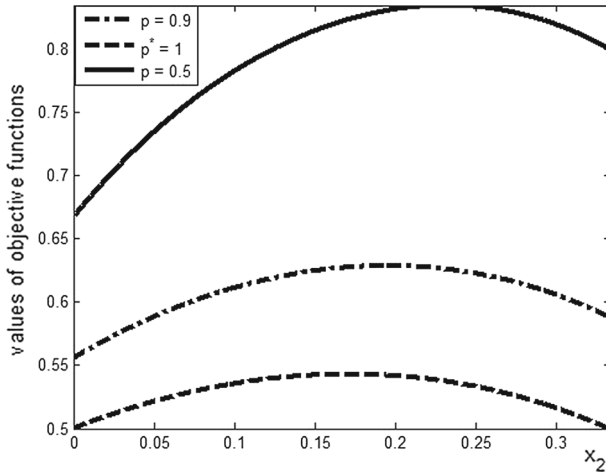
$$\min (1 - 2x_2)^p + x_2^p \quad \text{s.t.} \quad 0 \leq x_2 \leq \frac{1}{3}. \tag{58}$$

When  $0 < p < 1$ , problem (58) is a one-dimensional concave minimization problem over a closed interval (see Fig. 2); the corresponding objective values at the endpoints 0 and  $\frac{1}{3}$  are 1 and  $2\left(\frac{1}{3}\right)^p$ , respectively. Hence, (57) is equivalent to

$$\min \left\{ 1, 2\left(\frac{1}{3}\right)^p, 2 \right\} = \min \left\{ 1, 2\left(\frac{1}{3}\right)^p \right\}. \tag{59}$$

Note that when  $x_2 = 0$ , we have  $x_1 = 2x_2 - 1 = -1$ ; and when  $x_2 = \frac{1}{3}$ , we have  $x_1 = 2x_2 - 1 = -\frac{1}{3}$ ; that is, the candidate optimal points are  $(-1, 0)^T$  and  $(\frac{1}{3}, \frac{1}{3})^T$ . However, we desire the point  $(-1, 0)^T$  as an optimal solution to (55), meanwhile, as a (sparsest) optimal solution to the original problem (54). Hence we need the value of objective function of problem (55) at  $(-1, 0)^T$  is strictly smaller than that at  $(\frac{1}{3}, \frac{1}{3})^T$ . That is, we look forward to  $1 < 2\left(\frac{1}{3}\right)^p$ ; we derive  $0 < p < \frac{\ln 2}{\ln 3}$ . Then, we find  $p^* = \frac{\ln 2}{\ln 3}$ .

In sum, there exists  $p^* = \frac{\ln 2}{\ln 3}$  such that for any  $p$  with  $0 < p < p^*$ , an optimal solution to nonconvex relaxation problem (55) is  $(-1, 0)^T$ , which is also an optimal solution to the original problem (54). This confirm our results obtained in Sect. 3.



**Fig. 3** Objective functions of (63), being all concave over closed interval  $[0, \frac{1}{3}]$ ; among three specific  $p$  values,  $p^*$  is the constant that is claimed in our main results in Sect. 3

**Case 2.** We take  $f(t; p) = \frac{t}{t+p} \in \mathcal{F}(t; p)$ , and consider nonconvex relaxation problem

$$\min_{x \in \mathbb{R}^2} \left( \frac{|x_1|}{|x_1| + p} + \frac{|x_2|}{|x_2| + p} \right) \quad \text{s.t. } |x_1 - 2x_2| = 1, 2x_1 - x_2 \leq -1, \|x\|_\infty \leq 1. \tag{60}$$

The feasible set is same as Case 1 (see Fig. 1). Hence, problem (60) is equivalent to

$$\min \left\{ \min_{x \in \mathbb{D}_1} \left( \frac{|x_1|}{|x_1| + p} + \frac{|x_2|}{|x_2| + p} \right), \min_{x \in \mathbb{D}_2} \left( \frac{|x_1|}{|x_1| + p} + \frac{|x_2|}{|x_2| + p} \right) \right\}. \tag{61}$$

Therefore, (61) can be rewritten as

$$\begin{aligned} & \min \left\{ \min_{0 \leq x_2 \leq \frac{1}{3}} \frac{|2x_2 - 1|}{|2x_2 - 1| + p} + \frac{|x_2|}{|x_2| + p}, \frac{2}{1 + p} \right\} \\ & = \min \left\{ \min_{0 \leq x_2 \leq \frac{1}{3}} \frac{1 - 2x_2}{1 - 2x_2 + p} + \frac{x_2}{x_2 + p}, \frac{2}{1 + p} \right\}. \end{aligned} \tag{62}$$

Consider the subproblem in (62):

$$\min f(x_2) = \frac{1 - 2x_2}{1 - 2x_2 + p} + \frac{x_2}{x_2 + p} \quad \text{s.t. } 0 \leq x_2 \leq \frac{1}{3}. \tag{63}$$

When  $0 < p < 1$ , problem (63) is a one-dimensional concave minimization problem over a closed interval (see Fig. 3); the corresponding objective values at the endpoints 0 and  $\frac{1}{3}$  are  $\frac{1}{1+p}$  and  $\frac{2}{1+3p}$ , respectively. Hence, (62) is equivalent to

$$\min_{x_2 \in \{0, \frac{1}{3}\}} \left\{ f(0) = \frac{1}{1+p}, f\left(\frac{1}{3}\right) = \frac{2}{1+3p}, \frac{2}{1+p} \right\} = \min_{x_2 \in \{0, \frac{1}{3}\}} \left\{ \frac{1}{1+p}, \frac{2}{1+3p} \right\}. \quad (64)$$

Note that when  $x_2 = 0$ , we have  $x_1 = 2x_2 - 1 = -1$ ; and when  $x_2 = \frac{1}{3}$ , we have  $x_1 = 2x_2 - 1 = -\frac{1}{3}$ ; that is, the candidate optimal points are  $(-1, 0)^\top$  and  $(\frac{1}{3}, \frac{1}{3})^\top$ . However, we desire the point  $(-1, 0)^\top$  as an optimal solution to (55), meanwhile, as a (sparsest) optimal solution to the original problem (54). Hence we need the value of objective function of problem (55) at  $(-1, 0)^\top$  is strictly smaller than that at  $(\frac{1}{3}, \frac{1}{3})^\top$ . That is, we look forward to  $\frac{1}{1+p} < \frac{2}{1+3p}$ ; we derive  $0 < p < 1$ . Then, we find  $p^* = 1$ .

In sum, there exists  $p^* = 1$  such that for any  $p$  with  $0 < p < p^*$ , an optimal solution to nonconvex relaxation problem (60) is  $(-1, 0)^T$ , which is also an optimal solution to the original problem (54). This confirms our results obtained in Sect. 3.

**Remark 4.1** From the two cases discussed above, we not only confirmed the validness of one of our results, but also observed that for different types of nonconvex relaxation approaches, we derived different values of constant  $p^*$ , which are not too small to exceed the range of applications.

Additionally, if we solve the following  $\ell_1$ -norm minimization problem

$$\min_{\mathbf{x} \in \mathbb{R}^2} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad |x_1 - 2x_2| = 1, 2x_1 - x_2 \leq -1, \|\mathbf{x}\|_\infty \leq 1, \quad (65)$$

then we obtain a unique solution  $(-\frac{1}{3}, \frac{1}{3})^T$ , which is not one solution of the original problem (54). This implies that the solution of  $\ell_0$ -norm minimization problem (54) can not be obtained by solving the corresponding  $\ell_1$ -minimization (65).

## 5 Conclusions

In this paper, we studied a partially sparsest optimization problem, whose feasible set is a union of finite polytopes, via its nonconvex relaxation method in terms of a class of concave functions with a parameter  $p$ . We discussed the relationship between its optimal solution set and the optimal solution set of its nonconvex relaxation problem. We showed that there is a positive constant parameter  $p^* \leq 1$  such that any optimal solution to the nonconvex relaxation with any given parameter  $0 < p < p^*$  is also an optimal solution of the original sparse optimization problem. It is possible that an analytic expression of such a constant  $p^*$  could be derived. In addition, it is important to design an effective algorithm for solving the nonconvex relaxation. These are important issues to be studied in the future.

## References

Bradley PS, Mangasarian OL (1998) Feature selection via concave minimization and support vector machines. ICML 98:82–90

- Candés EJ, Romberg J, Tao T (2006) Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans Inf Theory* 52(2):489–509
- Candés EJ, Tao T (2005) Decoding by linear programming. *IEEE Trans Inf Theory* 51(12):4203–4215
- Chartrand R (2007) Exact reconstruction of sparse signals via nonconvex minimization. *IEEE Signal Process Lett* 14(10):707–710
- Chen X, Xiang S (2016) Sparse solution of linear complementarity problems. *Math Program* 159(1):539–556
- Davies ME, Gribonval R (2009) Restricted isometry constants where  $l_p$  sparse recovery can fail for  $0 < p < 1$ . *IEEE Trans Inf Theory* 55(5):2203–2214
- Foucart S, Lai MJ (2009) Sparsest solutions of underdetermined linear systems via  $l_q$ -minimization for  $0 < q \leq 1$ . *Appl Comput Harmon Anal* 26(3):395–407
- Foucart S, Rauhut H (2013) A mathematical introduction to compressive sensing. Birkhäuser, Basel
- Fung G, Mangasarian O (2011) Equivalence of minimal  $l_0$  and  $l_p$  norm solutions of linear equalities, inequalities and linear programs for sufficiently small  $p$ . *J Optim Theory Appl* 151:1–10
- Gribonval R, Nielsen M (2007) Highly sparse representations from dictionaries are unique and independent of the sparseness measure. *Appl Comput Harmon Anal* 22(3):335–355
- Gribonval R, Nielson M (2003) Sparse representations in unions of bases. *IEEE Trans Inf Theory* 49(12):3320–3325
- Grünbaum B (2003) Convex polytopes, 2nd edn. Springer, New York
- Haddou M, Migot T (2015) A smoothing method for sparse optimization over polyhedral sets. *Modelling, computation and optimization in information systems and management sciences*. Springer, Berlin, pp 369–379
- Lai MJ, Wang J (2011) An unconstrained  $l_p$  minimization with  $0 < q \leq 1$  for sparse solution of underdetermined linear systems. *SIAM J Optim* 21(1):82–101
- Lyu Q, Lin Z, She Y et al (2013) A comparison of typical  $l_p$  minimization algorithms. *Neurocomputing* 119:413–424
- Natarajan BK (1995) Sparse approximate solutions to linear systems. *SIAM J Comput* 24(2):227–234
- Peng J, Yue S, Li H (2015) NP/CMP equivalence: a phenomenon hidden among sparsity models  $l_0$  minimization and  $l_p$  minimization for information processing. *IEEE Trans Inf Theory* 61(7):4028–4033
- Rockafellar RT (1970) Convex analysis. Princeton University Press, Princeton
- Saab R, Chartrand R, Yilmaz O (2008) Stable sparse approximations via nonconvex optimization. In: *Proceedings of international conference on acoustics speech and signal processing*, pp 3885–3888
- Sun QY (2012) Recovery of sparsest signals via  $l_p$  minimization. *Appl Comput Harmon Anal* 32(3):329–341
- Voroninski V, Xu Z (2016) A strong restricted isometry property, with an application to phaseless compressed sensing. *Appl Comput Harmon Anal* 40(2):386–395
- Wang M, Xu WY, Tang A (2011) On the performance of sparse recovery via  $l_p$ -minimization ( $0 \leq p \leq 1$ ). *IEEE Trans Inf Theory* 57(11):7255–7278
- Xu ZB, Chang XY, Xu FM, Zhang H (2012)  $L_{1/2}$  regularization: a thresholding representation theory and a fast solver. *IEEE Trans Neural Netw Learn Syst* 23(7):1013–1027
- You G, Huang ZH, Wang Y (2017) A theoretical perspective of solving phaseless compressive sensing via its nonconvex relaxation. *Inf Sci* 415:254–268
- Zhang M, Huang ZH, Zhang Y (2013) Restricted  $p$ -isometry properties of nonconvex matrix recovery. *IEEE Trans Inf Theory* 59(7):4136–4323
- Ziegler GM (1995) Lectures on polytopes, Revised First Edition. Springer, New York