SOLUTION UNIQUENESS OF CONVEX PIECEWISE AFFINE FUNCTIONS BASED OPTIMIZATION WITH APPLICATIONS TO CONSTRAINED ℓ_1 MINIMIZATION

Seyedahmad Mousavi and Jinglai Shen[*]

Abstract. In this paper, we study the solution uniqueness of an individual feasible vector of a class of convex optimization problems involving convex piecewise affine functions and subject to general polyhedral constraints. This class of problems incorporates many important polyhedral constrained ℓ_1 recovery problems arising from sparse optimization, such as basis pursuit, LASSO, and basis pursuit denoising, as well as polyhedral gauge recovery. By leveraging the max-formulation of convex piecewise affine functions and convex analysis tools, we develop dual variables based necessary and sufficient uniqueness conditions *via* simple and yet unifying approaches; these conditions are applied to a wide range of ℓ_1 minimization problems under possible polyhedral constraints. An effective linear program based scheme is proposed to verify solution uniqueness conditions. The results obtained in this paper not only recover the known solution uniqueness conditions in the literature by removing restrictive assumptions but also yield new uniqueness conditions for much broader constrained ℓ_1-minimization problems.

1. Introduction

The ℓ_1-norm minimization, or simply ℓ_1 minimization, is a convex relaxation of ℓ_0-(pseudo)norm based sparse optimization, and it has received surging interest in diverse areas, such as compressed sensing, signal and image processing, machine learning, and high dimensional statistics and data analytics. Unlike the ℓ_p-norm with $p > 1$, the ℓ_1-norm is *not* strictly convex [20], and this yields many interesting issues in solution uniqueness which are critical to algorithm development and analysis. In addition to various important sufficient conditions for global and uniform solution uniqueness (or the so-called uniform recovery conditions) [8, 9, 23], necessary and sufficient conditions for solution uniqueness of an (arbitrary) individual vector are also established, *e.g.,* ([8], Sect. 4.3) and [10, 28, 29, 31], which are closely related to non-uniform recovery conditions in the sparse signal recovery literature [4, 8, 29].

It is worth mentioning that there are many different, nonetheless equivalent, solution uniqueness conditions for an individual vector. We are particularly interested in those conditions expressed in terms of dual variables or the so-called dual certificate conditions [9]. This is because dual variables usually have a smaller size in sparse

optimization. For example, the size of dual variables associated with a measurement matrix is the number of rows of this matrix, which is much smaller than the size of primal variables, *i.e.,* the number of columns of such a matrix. Therefore, solution uniqueness conditions in dual variables are numerically favorable. From an optimization point of view, such conditions are nontrivial and often require convex analysis tools to develop them. Moreover, it is desired that uniqueness conditions are explicitly dependent on problem parameters, *e.g.,* the measurement matrix and the measurement vector. Recent solution uniqueness results of this kind include [10, 28–31]. In particular, the paper ([30], Thm. 2.10) establishes solution uniqueness conditions for the standard ℓ_1 minimization, and the papers [28, 29] develop solution uniqueness conditions for several important ℓ_1 minimization problems and their variations, *e.g.,* basis pursuit (BP), the least absolute shrinkage and selection operator (LASSO), and basis pursuit denoising (BPDN). The recent paper [10] gives another proof of the uniqueness conditions of basis pursuit established in [28] and clarifies geometric meanings of these conditions with extensions to polyhedral gauge recovery. Motivated by constrained sparse signal recovery [7, 11, 27], solution uniqueness of basis pursuit under the nonnegative constraint is studied in [31]. Moreover, solution uniqueness conditions are recently developed for the standard ℓ_1 minimization over a general polyhedral set using the concept of restricted range space property [32, 33]. They can also be established *via* the uniqueness conditions of linear programs [13]; see Remark 4.12 for details. However, solution uniqueness of many other ℓ_1 minimization problems under general polyhedral constraints has not been fully addressed, despite various polyhedral constraints in applications, *e.g.,* the monotone cone constraint in order statistics, and the polyhedral constraint in the Dantzig selector [5] (*cf.* Sect. 3.5).

Inspired by the lack of solution uniqueness conditions under general polyhedral constraints and the fact that the ℓ_1-norm is a special convex piecewise affine (PA) function, we study a broad class of convex optimization problems involving convex PA functions and subject to general linear inequality constraints, and we develop necessary and sufficient solution uniqueness conditions for an individual feasible vector. This general framework incorporates many important ℓ_1 minimization problems under possible inequality constraints, such as BP, LASSO, BPDN, and polyhedral gauge recovery. Different from the techniques developed in a similar framework in [10], we exploit the max-formulation of a convex PA function (*cf.* Sect. 2). The max-formulation leads to much simpler, yet unifying and systematic, approaches to establish solution uniqueness conditions for a wide range of problems; see Remark 3.5 for comparison. These approaches not only recover all the known solution uniqueness conditions in the literature by removing restrictive assumptions but also shed light on new solution uniqueness conditions of much broader constrained ℓ_1 minimization problems, *e.g.,* the basis pursuit and sparse fused LASSO under general linear inequality constraints, and the Dantzig selector; see Section 5 for examples and details.

The rest of the paper is organized as follows. In Section 2, we introduce convex PA functions and discuss their properties. Section 3 develops solution uniqueness conditions for four convex optimization problems involving convex PA functions and subject to general linear inequality constrains, *i.e.,* basis pursuit-like problem, LASSO-like problem, and two basis pursuit denoising-like problems. By applying these results, Section 4 addresses solution existence and uniqueness of general ℓ_1 minimization problems. In Section 5, concrete uniqueness conditions are established for ℓ_1 minimization and compared with related results in the literature. Section 6 provides a simple and effective linear program-based scheme for verifying solution uniqueness conditions. Finally, conclusions are made in Section 7.

Notation. Let A be an $m \times N$ real matrix. For any index set $\mathcal{S} \subseteq \{1, \ldots, N\}$, let $|\mathcal{S}|$ denote the cardinality of $\mathcal{S}$, $\mathcal{S}^c$ denote the complement of $\mathcal{S}$, and $A_{\bullet \mathcal{S}}$ be the matrix formed by the columns of A indexed by elements of $\mathcal{S}$. Similarly, for an index set $\alpha \subseteq \{1, \ldots, m\}$, $A_{\alpha \bullet}$ is the matrix formed by the rows of A indexed by elements of α. For a given matrix A, $R(A)$ and $N(A)$ denote the range and null space of A, respectively. Denote by $\mathcal{N}_{\mathcal{C}}(x)$ the normal cone of a closed convex set $\mathcal{C}$ at $x \in \mathcal{C}$, and by int and ri the interior and the relative interior of a set, respectively. Besides, denote by $\mathbf{1}$ the column vectors of ones. In addition, $\mathbb{R}_+^N$ and $\mathbb{R}_{++}^N$ denote the nonnegative and positive orthants of $\mathbb{R}^N$, respectively. For a vector $z = (z_1, \ldots, z_n)^T$ whose each $z_i \neq 0$, we define $\operatorname{sgn}(z) := (z_1/|z_1|, \ldots, z_n/|z_n|)^T$.

2. Preliminary: Convex Piecewise Affine Functions

A real-valued continuous function $g : \mathbb{R}^N \to \mathbb{R}$ is piecewise affine (PA) if there exists a finite family of real-valued affine functions $\{g_i\}_{i=1}^{\ell}$ such that $g(x) \in \{g_i(x)\}_{i=1}^{\ell}$ for each $x \in \mathbb{R}^N$ [17]. A special class of continuous PA functions is continuous piecewise linear (PL) functions, for which each g_i is a linear function. A continuous PA function is globally Lipschitz, and we call it a *Lipschitz* PA function without loss of generality; see [18, 21, 22] for more geometric properties of these functions. A real-valued Lipischitz PA function can be described by the min-max formulation [17]. Furthermore, a convex (Lipschitz) PA function $g : \mathbb{R}^N \to \mathbb{R}$ (whose effective domain is $\mathbb{R}^N$) attains the max-formulation ([15], Sect. 19) or ([3], Prop. 2.3.5). Specifically, there exists a finite family of $(p_i, \gamma_i) \in \mathbb{R}^N \times \mathbb{R}, i = 1, \ldots, \ell$ such that

$$g(x) = \max_{i=1,\ldots,\ell} \left(p_i^T x + \gamma_i \right). \tag{2.1}$$

Similarly, a convex PL function attains the above max-formulation with all $\gamma_i = 0$. For a given $x \in \mathbb{R}^N$, define the index set $\mathcal{I} := \{i \in \{1, \ldots, \ell\} \,|\, p_i^T x + \gamma_i = g(x)\}$. Letting conv denote the convex hull of a set, the subdifferential of $g(x)$ at this x is then given by ([2], Prop. B.25)

$$\partial g(x) = \mathrm{conv}\left(\bigcup_{i \in \mathcal{I}} \partial(p_i^T x + \gamma_i) \right) = \mathrm{conv}\Big(\{p_i \,|\, i \in \mathcal{I}\} \Big). \tag{2.2}$$

The normal cone of $\{x \,|\, g(x) \leq 0\}$ at x^* is $\mathrm{cone}(\partial g(x^*))$ [16], where cone denotes the conic hull of a set. The following lemma presents additional properties of convex PA functions.

Lemma 2.1. *The following hold:*

(i) *The (real-valued) function $g : \mathbb{R}^N \to \mathbb{R}$ is a convex PA function if and only if its epigraph is a convex polyhedron in $\mathbb{R}^N \times \mathbb{R}$;*

(ii) *Let $f : \mathbb{R}^m \to \mathbb{R}$ be a convex PA function, and $h : \mathbb{R}^N \to \mathbb{R}^m$ be an affine function. Then $f \circ h$ is a convex PA function on $\mathbb{R}^N$;*

(iii) *Let $\{g_1, \ldots, g_r\}$ be a finite family of convex PA functions on $\mathbb{R}^N$. Then $\sum_{i=1}^{r} \lambda_i \cdot g_i(x)$ with $\lambda_i \geq 0$ is a convex PA function.*

Proof. Statement (i) follows from a similar proof for ([3], Prop. 2.3.5) by restricting the effective domain of g to $\mathbb{R}^N$, and statements (ii) and (iii) are trivial. □

Remark 2.2. A slightly more general class of convex PA functions is considered in ([15], Sect. 19) and [3]. Such a function, which is called the *polyhedral convex function* coined by R. T. Rockafellar, is defined as an *extended* real-valued function whose epigraph is a polyhedron in $\mathbb{R}^N \times \mathbb{R}$. It can be described by the sum of a real-valued convex PA function and the indicator function of a polyhedron, namely,

$$\widehat{g}(x) = \underbrace{\max_{i=1,\ldots,\ell} \left(h_i^T x + \beta_i \right)}_{:=g(x)} + \delta_{\mathcal{P}}(x),$$

where g is a real-valued convex PA function, $\mathcal{P} = \{x \,|\, Cx \geq d\}$ is a polyhedron in $\mathbb{R}^N$, and $\delta_{\mathcal{P}}$ is the indicator function of $\mathcal{P}$, *i.e.*, $\delta_{\mathcal{P}}(x) = 0$ if $x \in \mathcal{P}$, and $\delta_{\mathcal{P}}(x) = +\infty$ otherwise. See [3, 10, 15] for more discussions. However, in all the optimization problems to be considered in this paper, the polyhedron $\mathcal{P}$ corresponding to the indicator function in the function $\widehat{g}$ can be formulated as an additional linear inequality constraint, and thus be removed from $\widehat{g}(x)$. For example, the optimization problem: $\min \widehat{g}(x)$ subject to $Ax = y$ is equivalent to: $\min g(x)$ subject to $x \in \mathcal{P}$ and $Ax = y$. For this reason, we consider real-valued convex PA functions, or simply convex PA functions, throughout this paper.

Convex PA functions represent a broad class of nonsmooth convex functions in numerous applications, and we give several examples as follows. A (real-valued) *polyhedral gauge* is a convex PA function satisfying the following conditions: it is nonnegative, positively homogeneous of degree one, and vanishes at the origin [10, 15]. Since a (real-valued) convex function is continuous on $\mathbb{R}^N$, it must vanish at the origin if it is positively homogeneous of degree one, since for some $z \in \mathbb{R}^N$, $g(0) = \lim_{\lambda \downarrow 0} g(\lambda \cdot z) = \left(\lim_{\lambda \downarrow 0} \lambda \right) \cdot g(z) = 0$. Hence, a convex PA function is a polyhedral gauge if it is nonnegative and positively homogeneous of degree one. The following lemma shows that a polyhedral gauge is a convex PL function.

Lemma 2.3. *The function $g : \mathbb{R}^N \to \mathbb{R}$ is a polyhedral gauge if and only if there are finitely many $p_1, \ldots, p_\ell \in \mathbb{R}^N$ such that $g(x) = \max(p_1^T x, \ldots, p_\ell^T x, 0), \forall x \in \mathbb{R}^N$.*

Proof. The "if" part is trivial since the convex PA function $g(x) = \max(p_1^T x, \ldots, p_\ell^T x, 0)$ is nonnegative and positively homogeneous of degree one. We show the "only if" part as follows. Suppose $g : \mathbb{R}^N \to \mathbb{R}$ is a polyhedral gauge. Since g is a convex PA function, it attains the max-formulation and its domain attains a polyhedral subdivision of $\mathbb{R}^N$ [17]. Specifically, there are finitely many $(p_i, \gamma_i) \in \mathbb{R}^N \times \mathbb{R}$ and polyhedra $\mathcal{X}_i$, where $i = 1, \ldots, \ell$, such that $g(x) = \max_{i=1,\ldots,\ell}(p_i^T x + \gamma_i)$, and for each i, $g(x) = p_i^T x + \gamma_i$ for all $x \in \mathcal{X}_i$ ([6], Prop. 4.2.1). Here, $\Xi := \{\mathcal{X}_i\}_{i=1,\ldots,\ell}$ is a polyhedral subdivision of $\mathbb{R}^N$, *i.e.,* $\cup_{i=1}^\ell \mathcal{X}_i = \mathbb{R}^N$, each $\mathcal{X}_i$ has nonempty interior, and the intersection of any two polyhedra in Ξ is either empty or a common proper face of both polyhedra; see [6, 17–19] for more details. For any fixed $i \in \{1, \ldots, \ell\}$, let z be in the interior of $\mathcal{X}_i$. Therefore, $g(z) = p_i^T z + \gamma_i$, and for all $\lambda \in \mathbb{R}$ sufficiently close to 1, we have $\lambda \cdot z \in \mathcal{X}_i$ so that $g(\lambda \cdot z) = p_i^T (\lambda \cdot z) + \gamma_i$. Furthermore, since g is positively homogeneous of degree one, $g(\lambda \cdot z) = \lambda \cdot g(z)$ such that $\lambda \cdot p_i^T z + \gamma_i = \lambda \cdot p_i^T z + \lambda \cdot \gamma_i$ for all λ sufficiently close to 1. This shows that $\gamma_i = 0$ for each i. Therefore, $g(x) = \max(p_1^T x, \ldots, p_\ell^T x)$. Finally, since g is nonnegative, we have $g(x) = \max(g(x), 0)$ for all x. This shows that $g(x) = \max(p_1^T x, \ldots, p_\ell^T x, 0)$ for all $x \in \mathbb{R}^N$. $\qquad\square$

We mention a particular class of polyhedral gauges arising from applications as follows. Such a polyhedral gauge $g(x) = \max(p_1^T x, \ldots, p_\ell^T x, 0)$ with $p_i \neq 0, \forall i = 1, \ldots, \ell$ satisfies the following condition: for each nonzero p_i, there exists p_j such that $p_j = \beta_{j,i} \cdot p_i$ for some constant $\beta_{j,i} < 0$, where $\beta_{j,i}$ depends on p_i and p_j. We call such the polyhedral gauge *sign-symmetric*. Note that for each $x \in \mathbb{R}^N$,

$$g(x) = \max \Big\{ \max \big\{ \max(p_i^T x, p_j^T x) \mid i = 1, \ldots, \ell, \ p_j = \beta_{j,i} \cdot p_i, \ \beta_{j,i} < 0 \big\}, 0 \Big\}.$$

Since $\max(p_i^T x, p_j^T x) = \max(p_i^T x, \beta_{j,i} p_i^T x) \geq 0$ for any x, we have $g(x) = \max \big\{ \max(p_i^T x, p_j^T x) \mid i = 1, \ldots, \ell, \ p_i = \beta_{i,j} p_j, \ \beta_{i,j} < 0 \big\} = \max(p_1^T x, \ldots, p_\ell^T x)$. In other words, the zero term can be dropped in the max-formulation of a sign-symmetric polyhedral gauge. Examples of sign-symmetric polyhedral gauges include $\|Ex\|_1$ and $\|Ex\|_\infty$ for a matrix $E \in \mathbb{R}^{q \times N}$; see Section 4.2 for the max-formulation of $\|Ex\|_1$. Obviously, not every polyhedral gauge is sign-symmetric, *e.g.,* $\max(p^T x, 0)$ for some vector $p \neq 0$.

3. Unique optimal solution to a class of convex optimization problems involving convex PA functions

In this section, we develop dual variables based explicit conditions for unique optimal solutions to four convex optimization problems involving convex PA functions, which are motivated by basis pursuit (BP), LASSO, and basis pursuit denoising (BPDN) problems subject to possible polyhedral constraints. For each of these optimization problems, we assume that an optimal solution exists. A detailed study of solution existence requires different techniques and argument other than those for convex PA functions and uniqueness analysis. To avoid being off track from the main theme of the paper, we postpone the discussions of the solution existence issue to Section 4.1.

Among the four convex optimization problems treated in this section, three of them are involved with two functions: the first function, denoted by f, pertains to the cost due to measurement or approximation errors, while the second function corresponds to sparsity related penalty or objective value, which is usually a convex

PA function denoted by g. In the literature of statistics and decision theory, the first function is called a loss function. We consider the class of smooth (*i.e.*, C^1) and strictly convex loss functions through Sections 3.1–3.4, and study the class of convex PA loss functions in Section 3.5. A typical example of loss functions in the first class is the ℓ_2-loss $f(\cdot) = \|\cdot\|_2^2$, whereas examples of the second class are the ℓ_1-loss $\|\cdot\|_1$, the max-loss $\|\cdot\|_\infty$, and the hinge loss.

Throughout this section, let $g : \mathbb{R}^N \to \mathbb{R}$ be a convex PA function whose max-formulation is given in (2.1), $A \in \mathbb{R}^{m \times N}$, $y \in \mathbb{R}^m$, and $\mathcal{P} := \{x \in \mathbb{R}^N \mid Cx \geq d\}$ be a nonempty polyhedron, where $C \in \mathbb{R}^{\bar{m} \times N}$ and $d \in \mathbb{R}^{\bar{m}}$. For a given $x^* \in \mathbb{R}^N$ satisfying $Cx^* \geq d$, define the index sets

$$\alpha := \left\{i \in \{1, \ldots, \bar{m}\} \mid (Cx^* - d)_i = 0\right\}, \qquad \mathcal{I} := \left\{j \in \{1, \ldots, \ell\} \mid p_j^T x^* + \gamma_j = g(x^*)\right\}, \qquad (3.1)$$

and define the following matrix:

$$W := \begin{bmatrix} p_{i_1}^T \\ \vdots \\ p_{i_{|\mathcal{I}|}}^T \end{bmatrix}_{i_k \in \mathcal{I}} \in \mathbb{R}^{|\mathcal{I}| \times N}, \qquad (3.2)$$

where without loss of generality, we assume that for each $i_k \in \mathcal{I}$, p_{i_k} is *not* a convex combination of the other p_{i_j}'s with $i_j \in \mathcal{I}$. In light of (2.2), the columns of W^T are generators of the convex hull that forms the subdifferential $\partial g(x^*)$. Hence, finding the matrix W is equivalent to finding convex hull generators of $\partial g(x^*)$. This observation will be exploited to establish the matrix W; see Lemma 4.5 and Section 4.

3.1. Unique optimal solution to the basis pursuit-like problem

Consider the following convex optimization problem motivated by the basis pursuit (BP) subject to a linear inequality constraint:

$$\min_{x \in \mathbb{R}^N} \ g(x) \quad \text{subject to} \quad Ax = y, \quad Cx \geq d. \qquad (3.3)$$

We assume that this problem has an optimal solution. For a given feasible point $x^* \in \mathbb{R}^N$ of (3.3), *i.e.*, $Ax^* = y$ and $Cx^* \geq d$, recall the definitions of $\alpha, \mathcal{I}$, and W in (3.1) and (3.2).

Lemma 3.1. *Let $A \in \mathbb{R}^{m \times N}$ and $H \in \mathbb{R}^{r \times N}$ be given. Then $\{u \in \mathbb{R}^N \mid Au = 0, \ Hu \geq 0\} = \{0\}$ if and only if the following two conditions hold:*

(i) $\{u \in \mathbb{R}^N \mid Au = 0, \ Hu = 0\} = \{0\}$ and
(ii) there exist $z \in \mathbb{R}^m$ and $z' \in \mathbb{R}^r_{++}$ such that $A^T z = H^T z'$.

Proof. Consider the following linear program:

$$(LP): \quad \max_{u \in \mathbb{R}^N} \ \mathbf{1}^T Hu, \quad \text{subject to} \quad Au = 0, \quad Hu \geq 0. \qquad (3.4)$$

We claim that $\{u \in \mathbb{R}^N \mid Au = 0, \ Hu \geq 0\} = \{0\}$ if and only if the following hold:

(i') condition (i) holds, *i.e.*, $\{u \in \mathbb{R}^N \mid Au = 0, \ Hu = 0\} = \{0\}$ and
(ii') the linear program (LP) attains the zero optimal value (and the unique optimal solution $u^* = 0$).

To show the "if" part of this claim, suppose (i') and (ii') hold but there exists $u' \neq 0$ such that $Au' = 0$ and $Hu' \geq 0$. It follows from (i') that $Hu' \neq 0$. Therefore, $\mathbf{1}^T Hu' > 0$, a contradiction to (ii'). Conversely, suppose $\{u \in \mathbb{R}^N \mid Au = 0, \ Hu \geq 0\} = \{0\}$ holds. Clearly, it implies condition (i'). Furthermore, the feasible set of the (LP) is the singleton set $\{0\}$ such that (ii') holds. Consequently, the claim holds.

The dual problem of (LP) is given by:

$$\min_{(v,w)\in\mathbb{R}^m\times\mathbb{R}^r} 0, \qquad \text{subject to} \quad A^T v - H^T w = H^T \mathbf{1}, \quad w \geq 0.$$

In view of the strong duality theorem of linear program, the dual problem attains an optimal solution (v^*, w^*) such that $A^T v^* = H^T(\mathbf{1} + w^*)$ and $w^* \geq 0$. By suitable positive scaling, we deduce that there exist z and $z' > 0$ such that $A^T z = H^T z'$, which yields condition (ii). Since condition (ii) is also sufficient for the feasibility, and thus solvability, of the dual problem, it follows from the weak duality of linear program that $\mathbf{1}^T H u \leq 0$ for any feasible u of (LP). Since $Hu \geq 0$ for any feasible u of (LP), we must have $\mathbf{1}^T H u = 0$, which leads to condition (ii'). Therefore, conditions (ii) and (ii') are equivalent. In view of the claim proven above, the lemma holds. $\square$

Theorem 3.2. *Let $x^* \in \mathbb{R}^N$ be a feasible point of the optimization problem (3.3). Then x^* is the unique minimizer of the problem (3.3) if and only if the following two conditions hold:*

(i) $\{v \in \mathbb{R}^N \,|\, Av = 0,\ C_{\alpha\bullet}v = 0,\ Wv = 0\} = \{0\}$ *and*
(ii) there exist $z \in \mathbb{R}^m$, $z' \in \mathbb{R}^{|\alpha|}_{++}$, and $z'' \in \mathbb{R}^{|\mathcal{I}|}_{++}$ such that $A^T z - C_{\alpha\bullet}^T z' + W^T z'' = 0$.

Moreover, condition (ii) is equivalent to the following condition:

(iii) there exist $w \in \mathbb{R}^m$, $w' \in \mathbb{R}^{|\alpha|}_{++}$, and $w'' \in \mathbb{R}^{|\mathcal{I}|}$ with $0 < w'' < \mathbf{1}$ and $\mathbf{1}^T w'' = 1$ such that $A^T w - C_{\alpha\bullet}^T w' + W^T w'' = 0$.

Proof. Clearly, x^* is a unique (global) minimizer of the convex optimization problem (3.3) if and only if x^* is a local unique minimizer of (3.3). It is easy to see that for all x sufficiently close to x^*, $g(x) = g(x^*) + \max_{i\in\mathcal{I}}\left(p_i^T(x - x^*)\right)$. In other words, $g(x) - g(x^*)$ is piecewise linear (and convex) in $(x - x^*)$ for all x sufficiently close to x^*. By this observation, we deduce that x^* is the unique minimizer of (3.3) if and only if $v^* = 0$ is the unique minimizer of the following convex optimization problem:

$$\min_{v\in\mathbb{R}^N} \left(\max_{i\in\mathcal{I}} p_i^T v\right), \quad \text{subject to} \quad Av = 0, \quad C_{\alpha\bullet}v \geq 0. \tag{3.5}$$

Furthermore, it is easy to verify that $v^* = 0$ is the unique minimizer of (3.5) if and only if

$$\left\{ v \in \mathbb{R}^N \,\big|\, Av = 0,\ C_{\alpha\bullet}v \geq 0,\ \max_{i\in\mathcal{I}} p_i^T v \leq 0 \right\} = \{0\}.$$

In light of the definition of the matrix W given in (3.2), we see that $\max_{i\in\mathcal{I}} p_i^T v \leq 0$ is equivalent to $Wv \leq 0$. Hence, $v^* = 0$ is the unique minimizer of (3.5) if and only if $\{v \in \mathbb{R}^N \,|\, Av = 0,\ C_{\alpha\bullet}v \geq 0,\ Wv \leq 0\} = \{0\}$. By setting $H = \begin{bmatrix} C_{\alpha\bullet} \\ -W \end{bmatrix}$, we deduce *via* Lemma 3.1 that $\{v \in \mathbb{R}^N \,|\, Av = 0,\ C_{\alpha\bullet}v \geq 0,\ Wv \leq 0\} = \{0\}$ if and only if conditions (i) and (ii) hold. This leads to the desired result.

We finally show the equivalence of conditions (ii) and (iii). Clearly, condition (iii) implies condition (ii). Conversely, suppose there exist $z \in \mathbb{R}^m$, $z' \in \mathbb{R}^{|\alpha|}_{++}$, and $z'' \in \mathbb{R}^{|\mathcal{I}|}_{++}$ such that $A^T z - C_{\alpha\bullet}^T z' + W^T z'' = 0$. Note that $\mathbf{1}^T z'' > 0$. Therefore, letting

$$w = \frac{z}{\mathbf{1}^T z''}, \qquad w' = \frac{z'}{\mathbf{1}^T z''}, \qquad w'' = \frac{z''}{\mathbf{1}^T z''},$$

we obtain condition (iii) with the above w, w' and w''. Hence, conditions (ii) and (iii) are equivalent. $\square$

3.1.1. Comparison with related results in the literature

The paper [10] studies a problem similar to (3.3) but without the linear inequality constraint. For comparison, we apply these tools to the problem (3.3). Define $\mathcal{S} := \{x \,|\, Ax = y\}$ and $\mathcal{P} := \{x \,|\, Cx \geq d\}$. Note that the problem (3.3) is equivalent to the unconstrained problem: $\min_{x \in \mathbb{R}^N} J(x)$, where $J(x) := g(x) + \delta_\mathcal{S}(x) + \delta_\mathcal{P}(x)$, and δ is the indicator function defined in Remark 2.2. Since $\mathcal{S}$ and $\mathcal{P}$ are polyhedral, x^* is the unique minimizer of (3.3) if and only if $0 \in \mathrm{int}\big(\partial J(x^*)\big)$ ([10], Lem. 3.2), where $\partial J(x^*) = \partial g(x^*) + \mathcal{N}_\mathcal{S}(x^*) + \mathcal{N}_\mathcal{P}(x^*)$. Since $\mathcal{N}_\mathcal{S}(x^*) = R(A^T)$ and $\mathcal{N}_\mathcal{P}(x^*) = \{C_{\alpha\bullet}^T z \,|\, z \leq 0\}$, the condition $0 \in \mathrm{int}\big(\partial J(x^*)\big)$ is further equivalent to the following two conditions ([10], Prop. 4.2):

(a) $0 \in \mathrm{ri}(\partial g(x^*)) + R(A^T) + \mathrm{ri}(\{C_{\alpha\bullet}^T z \,|\, z \leq 0\})$ and
(b) $\mathrm{aff}(\partial g(x^*) + R(A^T) + \{C_{\alpha\bullet}^T z \,|\, z \leq 0\}) = \mathbb{R}^N$, where $\mathrm{aff}(\cdot)$ denotes the affine hull of a set.

In what follows, we show that conditions (a) and (b) are equivalent to conditions (i) and (iii) of Theorem 3.2. To achieve this goal, we first present a lemma which gives an explicit characterization of relative interiors of a polytope and a polyhedral cone.

Lemma 3.3. *Let $\mathcal{C} = conv(a_1, \ldots, a_k)$ and $\mathcal{K} = cone(b_1, \ldots, b_\ell)$, where $a_1, \ldots, a_k \in \mathbb{R}^n$ and $b_1, \ldots, b_\ell \in \mathbb{R}^n$. Then the relative interiors of $\mathcal{C}$ and $\mathcal{K}$ are*

$$ri\,\mathcal{C} = \Big\{ \sum_{i=1}^k \lambda_i a_i \,\Big|\, \sum_{i=1}^k \lambda_i = 1, \; 0 < \lambda_i < 1, \, \forall\, i = 1, \ldots, k \Big\}, \quad ri\,\mathcal{K} = \Big\{ \sum_{j=1}^\ell \mu_j b_j \,\Big|\, \mu_j > 0, \; \forall\, j = 1, \ldots, \ell \Big\}.$$

Proof. To establish the relative interior of $\mathcal{C}$, we note that $\mathcal{C} = \mathrm{conv}(\mathcal{C}_1 \cup \mathcal{C}_2 \cup \cdots \cup \mathcal{C}_k)$, where each $\mathcal{C}_i := \{a_i\}$ is a convex singleton set. Hence, $\mathrm{ri}\,\mathcal{C}_i = \{a_i\}$ for each i. It follows from ([15], Thm. 6.9) that $\mathrm{ri}\,\mathcal{C} = \Big\{ \sum_{i=1}^k \lambda_i \cdot \mathrm{ri}\,\mathcal{C}_i \,\big|\, \sum_{i=1}^k \lambda_i = 1, \; 0 < \lambda_i < 1, \, \forall\, i = 1, \ldots, k \Big\}$, which leads to the desired result for $\mathrm{ri}\,\mathcal{C}$. The relative interior of $\mathcal{K}$ also follows by positive scaling. $\qquad\square$

Proposition 3.4. *Conditions (a) and (b) for the optimization problem (3.3) are equivalent to conditions (i) and (iii) of Theorem 3.2.*

Proof. By Lemma 3.3 and the subdifferential of g at x^* given in (2.2), we have

$$\mathrm{ri}\big(\partial g(x^*)\big) = \Big\{ \sum_{i \in \mathcal{I}} \lambda_i \cdot p_i \,\Big|\, \sum_{i \in \mathcal{I}} \lambda_i = 1, \; 0 < \lambda_i < 1, \, \forall\, i \in \mathcal{I} \Big\}, \quad \mathrm{ri}\big(\{C_{\alpha\bullet}^T z \,|\, z \leq 0\}\big) = \{C_{\alpha\bullet}^T z \,|\, z < 0\}.$$

Therefore, condition (a) holds if and only if there exist $z, z' > 0$, and z'' with $0 < z'' < \mathbf{1}$ with $\mathbf{1}^T z'' = 1$ such that $A^T z - C_{\alpha\bullet}^T z' + W^T z'' = 0$, which is exactly condition (iii) of Theorem 3.2.

Moreover, $\mathrm{aff}(\partial g(x^*) + R(A^T) + \{C_{\alpha\bullet}^T z \,|\, z \leq 0\}) = \mathrm{aff}(\partial g(x^*)) + \mathrm{aff}(R(A^T)) + \mathrm{aff}(\{C_{\alpha\bullet}^T z \,|\, z \leq 0\})$, where, in view of $\partial g(x^*) = \mathrm{conv}(p_{i_1}, p_{i_2}, \ldots, p_{i_{|\mathcal{I}|}})$, we have

$$\mathrm{aff}(\partial g(x^*)) = p_{i_1} + \mathrm{span}\big(p_{i_2} - p_{i_1}, \ldots, p_{i_{|\mathcal{I}|}} - p_{i_1}\big), \quad \mathrm{aff}(R(A^T)) = R(A^T), \quad \mathrm{aff}(\{C_{\alpha\bullet}^T z \,|\, z \leq 0\}) = R(C_{\alpha\bullet}^T).$$

Therefore, condition (b) is equivalent to

$$\underbrace{\mathrm{span}\big(p_{i_2} - p_{i_1}, \ldots, p_{i_{|\mathcal{I}|}} - p_{i_1}\big)}_{:=\mathcal{V}} + R(A^T) + R(C_{\alpha\bullet}^T) = \mathbb{R}^N,$$

which is further equivalent to condition (b'): $\mathcal{V}^\perp \cap N(A) \cap N(C_{\alpha\bullet}) = \{0\}$, where $\mathcal{V}^\perp = \{v \in \mathbb{R}^N \,|\, p_{i_1}^T v = p_{i_2}^T v = \cdots = p_{i_{|\mathcal{I}|}}^T v\}$. Obviously, condition (b') implies condition (i) of Theorem 3.2. We show next that if

conditions (i) and (iii) holds, then condition (b') holds. It follows from condition (iii) that for any $v \in \mathbb{R}^N$, $v^T A z - v^T C_{\alpha\bullet}^T z' + v^T W^T z'' = 0$ for some $z, z' > 0$, and z'' with $0 < z'' < \mathbf{1}$ with $\mathbf{1}^T z'' = 1$. Therefore, for any $v \in \mathcal{V}^\perp \cap N(A) \cap N(C_{\alpha\bullet})$, *i.e.*, $Av = 0, C_{\alpha\bullet}v = 0$, and $p_{i_1}^T v = p_{i_2}^T v = \cdots = p_{i_{|\mathcal{I}|}}^T v$, we obtain $(Wv)^T z'' = 0$. By virtue of the expression of the matrix W in (3.2), we further have $0 = (Wv)^T z'' = (p_{i_1}^T v \cdot \mathbf{1})^T z'' = (p_{i_1}^T v) \cdot \mathbf{1}^T z'' = p_{i_1}^T v$. This shows that $Wv = 0$. Along with $Av = 0$ and $C_{\alpha\bullet}v = 0$, we see *via* condition (i) of Theorem 3.2 that $v = 0$ such that condition (b') holds. Consequently, conditions (a) and (b) hold if and only if conditions (i) and (iii) of Theorem 3.2 hold.   □

Remark 3.5. Proposition 3.4 shows that the techniques developed in [10] can be used to derive the exactly same solution uniqueness conditions given in Theorem 3.2. However, when establishing explicit uniqueness conditions in terms of problem parameters, the paper [10] considers a particular class of convex PA functions, *i.e.*, polyhedral gauges, and employs the inner representation of the unit sublevel set of a polyhedral gauge to obtain (equivalent) uniqueness conditions in a different form. Instead, the present paper gives a much simpler approach to derive the explicit uniqueness conditions in Theorem 3.2 for a general convex PA function *via* its max-formulation, which can be easily applied to any specific convex PA function. For example, by leveraging Lemma 2.3 and Theorem 3.2, explicit uniqueness conditions can be readily obtained for a polyhedral gauge. Furthermore, the proposed approach can be exploited for other relevant problems as shown in the subsequent subsections, and thus provides a simple, albeit unifying, framework for a broad class of problems. Nevertheless, conditions (a) and (b) derived in [10] give better geometric interpretation of the conditions obtained in Theorem 3.2.

3.2. Unique optimal solution to the LASSO-like problem

Letting $f : \mathbb{R}^m \to \mathbb{R}$ be a C^1 strictly convex function, we consider the following convex optimization problem motivated by the constrained LASSO:

$$\min_{x \in \mathbb{R}^N} \ f(Ax - y) + g(x) \quad \text{subject to} \quad Cx \geq d. \tag{3.6}$$

We assume that this optimization problem has an optimal solution. To characterize a unique optimal solution to (3.6), we first present some preliminary results as follows.

Being an extension of ([28], Lem. 4.1), the following lemma can be shown *via* an elementary argument in convex analysis; its proof is thus omitted.

Lemma 3.6. *Let $f : \mathbb{R}^m \to \mathbb{R}$ be a strictly convex function, and $h : \mathbb{R}^N \to \mathbb{R}$ be a convex function. If $f(Ax - y) + h(x)$ is constant on a convex set $\mathcal{S} \subseteq \mathbb{R}^N$, then $Ax = Az$ and $h(x) = h(z)$ for all $x, z \in \mathcal{S}$.*

In light of Lemma 3.6, we obtain the following proposition which generalizes ([28], Thm. 2.1) using a similar argument. To be self-contained, we present its proof as follows.

Proposition 3.7. *Let $f : \mathbb{R}^m \to \mathbb{R}$ be a strictly convex function, $h : \mathbb{R}^N \to \mathbb{R}$ be a convex function, and $\mathcal{C}$ be a convex set in $\mathbb{R}^N$ such that the following optimization problem has a minimizer $x^* \in \mathbb{R}^N$:*

$$(P_0): \qquad \min_{x \in \mathbb{R}^N} \ f(Ax - y) + h(x) \quad \text{subject to} \quad x \in \mathcal{C}.$$

Then x^* is the unique minimizer of (P_0) if and only if x^* is the unique minimizer of the following problem:

$$(P_1): \qquad \min_{x \in \mathbb{R}^N} \ h(x) \quad \text{subject to} \quad Ax = Ax^*, \text{ and } x \in \mathcal{C}.$$

Proof. Let $\mathcal{S}_0$ be the solution set of (P_0). It is easy to see that $\mathcal{S}_0$ is convex on which $f(Ax - y) + h(x)$ is constant. By Lemma 3.6, we have $Ax = Ax^*$ and $h(x) = h(x^*)$ for any $x \in \mathcal{S}_0 \subseteq \mathcal{C}$. To show the "if" part, suppose that x^* is the unique minimizer of (P_1) but there exists $z \in \mathcal{S}_0$ with $z \neq x^*$. It follows from the previous result that $h(z) = h(x^*)$, $Ax = Ax^*$, and $z \in \mathcal{C}$, contradicting the solution uniqueness of (P_1). Conversely, for the "only if"

part, we first show that x^* is a minimizer of (P_1). Suppose not, *i.e.,* there exists $z \in \mathcal{C}$ with $Az = Ax^*$ such that $h(z) < h(x^*)$. Then we have $f(Az - y) + h(z) < f(Ax^* - y) + h(x^*)$. This implies that x^* is not a minimizer of (P_0), contradiction. The solution uniqueness of (P_1) follows directly from that of (P_0) and the result given at the beginning of the proof. □

Theorem 3.8. *Let $x^* \in \mathbb{R}^N$ be a feasible point of the problem (3.6), where $f : \mathbb{R}^m \to \mathbb{R}$ is a C^1 strictly convex function. Then x^* is the unique minimizer of (3.6) if and only if all the following conditions hold:*

   (i) $\{v \in \mathbb{R}^N \,|\, Av = 0, \, C_{\alpha\bullet}v = 0, \, Wv = 0\} = \{0\}$;
   (ii) there exist $z \in \mathbb{R}^m$, $z' \in \mathbb{R}^{|\alpha|}_{++}$, and $z'' \in \mathbb{R}^{|\mathcal{I}|}$ with $0 < z'' < \mathbf{1}$ and $\mathbf{1}^T z'' = 1$ such that $A^T z - C^T_{\alpha\bullet}z' + W^T z'' = 0$;
   (iii) there exist $w \in \mathbb{R}^{|\alpha|}_+$ and $w' \in \mathbb{R}^{|\mathcal{I}|}$ with $0 \leq w' \leq \mathbf{1}$ and $\mathbf{1}^T w' = 1$ such that $A^T \nabla f(Ax^* - y) - C^T_{\alpha\bullet}w + W^T w' = 0$.

Proof. We first show that x^* is a minimizer of (3.6) if and only if condition (iii) holds. Recall that $\mathcal{P} := \{x \,|\, Cx \geq d\}$. Since $f(Ax - y) + g(x)$ is a real-valued convex function on $\mathbb{R}^N$ and $\mathcal{P}$ is a closed convex set, it follows from ([16], Thm. 3.33) that x^* is a minimizer of (3.6) if and only if $0 \in \partial f(Ax^* - y) + \partial g(x^*) + \mathcal{N}_{\mathcal{P}}(x^*)$. In light of

$$\partial f(Ax^* - y) = \{A^T \nabla f(Ax^* - y)\}, \quad \partial g(x^*) = \mathrm{conv}(p_{i_1}, p_{i_2}, \ldots, p_{i_{|\mathcal{I}|}}), \quad \mathcal{N}_{\mathcal{P}}(x^*) = \{C^T_{\alpha\bullet}u \,|\, u \leq 0\},$$

we see that x^* is a minimizer of (3.6) if and only if condition (iii) holds.

   Applying Proposition 3.7 with $\mathcal{C} := \{x \in \mathbb{R}^N \,|\, Cx \geq d\}$ and $h(x) = g(x)$, we deduce that a minimizer x^* is the unique minimizer of (3.6) if and only if it is the unique minimizer of the following problem in the form of (3.3):

$$(P_2): \qquad \min_{x \in \mathbb{R}^N} \; g(x) \quad \text{subject to} \quad Ax = Ax^*, \text{ and } Cx \geq d.$$

Clearly, x^* is a feasible point of (P_2). Hence, by Theorem 3.2, x^* is the unique minimizer of (P_2) if and only if conditions (i) and (ii) hold. This completes the proof. □

3.3. Unique optimal solution to the basis pursuit denoising I-like problem

   Letting $f : \mathbb{R}^m \to \mathbb{R}$ be a C^1 strictly convex function and $\varepsilon \in \mathbb{R}$, we consider the following convex optimization problem motivated by the BPDN-I problem with an additional linear inequality constraint:

$$\min_{x \in \mathbb{R}^N} \; g(x) \quad \text{subject to} \quad f(Ax - y) \leq \varepsilon, \text{ and } Cx \geq d. \tag{3.7}$$

We assume that this problem has an optimal solution. Note that this problem is different from that treated in [10], since the constraints are no longer polyhedral in general. Moreover, the papers [28, 29] consider a problem similar to (3.7) with $g(x) = \|Ex\|_1$ or $g(x) = \|x\|_1$ but without the linear inequality constraint $Cx \geq d$, and they show that its solution uniqueness can be reduced to that of a relevant basis pursuit problem. However, this reduction does not hold for (3.7) due to the presence of the general linear inequality constraint; see Section 5.1 for a counterexample. This calls for new techniques to handle (3.7).

Theorem 3.9. *Let $x^* \in \mathbb{R}^N$ be a feasible point of the problem (3.7).*

   C.1 Suppose $f(Ax^* - y) < \varepsilon$. Then x^* is the unique minimizer of (3.7) if and only if $\{v \in \mathbb{R}^N \,|\, C_{\alpha\bullet}v = 0, \, Wv = 0\} = \{0\}$ and there exist $z \in \mathbb{R}^{|\alpha|}_{++}$ and $z' \in \mathbb{R}^{|\mathcal{I}|}$ with $0 < z' < \mathbf{1}$ and $\mathbf{1}^T z' = 1$ such that $C^T_{\alpha\bullet}z = W^T z'$.
   C.2 Suppose $f(Ax^* - y) = \varepsilon$. Then x^* is the unique minimizer of (3.7) if and only if the following hold:
   (2.i) $\{v \in \mathbb{R}^N \,|\, Av = 0, \, C_{\alpha\bullet}v = 0, \, Wv = 0\} = \{0\}$;

(2.ii) *there exist $z \in \mathbb{R}^m$, $z' \in \mathbb{R}_{++}^{|\alpha|}$, and $z'' \in \mathbb{R}^{|\mathcal{I}|}$ with $0 < z'' < \mathbf{1}$ and $\mathbf{1}^T z'' = 1$ such that $A^T z - C_{\alpha\bullet}^T z' + W^T z'' = 0$;*

(2.iii) *if $\mathcal{K} := \{v \in \mathbb{R}^N \mid (\nabla f(Ax^* - y))^T Av < 0, \ C_{\alpha\bullet} v \geq 0\}$ is nonempty, then there exist $w \in \mathbb{R}_+^{|\alpha|}$ and $w' \in \mathbb{R}_+^{|\mathcal{I}|}$ such that $A^T \nabla f(Ax^* - y) - C_{\alpha\bullet}^T w + W^T w' = 0$.*

Remark 3.10. We give several remarks on the conditions in Theorem 3.9 before presenting its proof.

(a) Note that in C.2, if the cone $\mathcal{K}$ defined in condition (2.iii) is empty, then x^* is the unique minimizer if and only if conditions (2.i) and (2.ii) hold.

(b) The cone $\mathcal{K}$ is nonempty if and only if there is no $u \geq 0$ such that $A^T \nabla f(Ax^* - y) = C_{\alpha\bullet}^T u$. Geometrically, it means that $A^T \nabla f(Ax^* - y)$ is not in the dual cone of $\{v \mid C_{\alpha\bullet} v \geq 0\}$, which equals the normal cone of the polyhedron $\mathcal{P} := \{x \mid Cx \geq d\}$ at x^*. This condition provides a constraint qualification for the optimality condition shown in (2.iii).

(c) In view of remark (b), we see that if $\mathcal{K}$ is nonempty, then a nonnegative w' given in condition (2.iii) must be nonzero. Hence, condition (2.iii) can be equivalently written as: if $\mathcal{K}$ is nonempty, then there exist a positive real number θ, $\widetilde{w} \in \mathbb{R}_+^{|\alpha|}$, and $\widetilde{w}' \in \mathbb{R}^{|\mathcal{I}|}$ with $0 \leq \widetilde{w}' \leq \mathbf{1}$ and $\mathbf{1}^T \widetilde{w}' = 1$ such that $\theta \cdot A^T \nabla f(Ax^* - y) - C_{\alpha\bullet}^T \widetilde{w} + W^T \widetilde{w}' = 0$.

Proof. The proof is divided into the following two parts:

Case C.1: $f(Ax^* - y) < \varepsilon$. Due to the continuity of f, it is clear that x^* is a unique minimizer of (3.7) if and only if it is a unique (local) minimizer of the following problem on a small neighborhood of x^*:

$$\min_{x \in \mathbb{R}^N} g(x), \quad \text{subject to} \quad Cx \geq d.$$

By applying Theorem 3.2 with $A = 0$ and $y = 0$ to the above problem, we obtain the desired result.

Case C.2: $f(Ax^* - y) = \varepsilon$. Define the function $r(Av) := f(Ax^* - y + Av) - f(Ax^* - y) - (\nabla f(Ax^* - y))^T Av$ for $v \in \mathbb{R}^N$. Since f is strictly convex, we see that $r(Av) \geq 0$ for all v, and $r(Av) = 0$ if and only if $Av = 0$. Furthermore, since f is C^1, we have

$$\lim_{0 \neq Av \to 0} \frac{r(Av)}{\|Av\|} = 0. \tag{3.8}$$

For notational simplicity, we define $q := A^T \nabla f(Ax^* - y) \in \mathbb{R}^N$. Note that if $Av = 0$, so is $q^T v$.

Define the positively homogeneous function $\widetilde{g}(v) := \max_{i \in \mathcal{I}} p_i^T v$. By virtue of the piecewise linear structure of $g(x) - g(x^*)$ for all x sufficiently close to x^*, x^* is the unique minimizer of (3.7) if and only if $v^* = 0$ is a unique local minimizer of the following problem:

$$\min_{v \in \mathbb{R}^N} \widetilde{g}(v) \quad \text{subject to} \quad q^T v + r(Av) \leq 0, \quad C_{\alpha\bullet} v \geq 0. \tag{3.9}$$

We claim that $v^* = 0$ is the unique local minimizer of (3.9) if and only if the following hold:

(i') $u^* = 0$ is the unique minimizer of the problem

$$\min_{u \in \mathbb{R}^N} \widetilde{g}(u) \quad \text{subject to} \quad Au = 0, \quad C_{\alpha\bullet} u \geq 0 \quad \text{and}$$

(ii') if the cone $\mathcal{K} := \{u \mid q^T u < 0, \ C_{\alpha\bullet} u \geq 0\}$ is nonempty, then $\widetilde{g}(u) > 0$ for all $u \in \mathcal{K}$.

To show this claim, we first prove the "if" part. Let $\mathcal{U}$ be a neighborhood of $v^* = 0$ such that $\widetilde{g}(v) = g(x^* + v) - g(x^*)$ and $C_{\alpha^c\bullet}(x^* + v) > d_{\alpha^c}$ for all $v \in \mathcal{U}$. For any $0 \neq v \in \mathcal{U}$ with $q^T v + r(Av) \leq 0$ and $C_{\alpha\bullet} v \geq 0$, we consider two cases: $Av = 0$, and $Av \neq 0$. For the former case, we have $q^T v + r(Av) = 0$. By condition (i'), we

have $\widetilde{g}(v) > \widetilde{g}(0) = 0$. For the latter case, since $Av \neq 0$, we have $r(Av) > 0$ so that $q^T v < 0$. By condition (ii'), we also have $\widetilde{g}(v) > 0$. Therefore, $v^* = 0$ is the unique local minimizer of (3.9). We next prove the "only if" part. Suppose $v^* = 0$ is the unique local minimizer of (3.9). For any $u \neq 0$ with $Au = 0$ and $C_{\alpha\bullet} u \geq 0$, we have that for all sufficiently small $\beta > 0$, $q^T \beta u + r(\beta Au) = 0$ such that βu is a nonzero local feasible point of (3.9). This implies that $\widetilde{g}(\beta u) > 0$. By the positive homogeneity of $\widetilde{g}$, we see that $\widetilde{g}(u) > 0$ for all $u \neq 0$ with $Au = 0$ and $C_{\alpha\bullet} u \geq 0$. This leads to condition (i'). Furthermore, for any $u \in \mathcal{K}$, we deduce *via* $q^T u < 0$ that $u \neq 0$ and $Au \neq 0$ (recalling that $[Au = 0] \Rightarrow [q^T u = 0]$). It follows from (3.8) that for all sufficiently small $\beta > 0$,

$$\frac{q^T \beta u + r(A\beta u)}{\|\beta u\|} = \frac{q^T u}{\|u\|} + \frac{r(\beta Au)}{\|\beta Au\|} \cdot \frac{\|Au\|}{\|u\|} < 0.$$

Therefore, $q^T \beta u + r(A\beta u) < 0$ for all small $\beta > 0$. Hence, βu is a nonzero local feasible point of (3.9) so that $\widetilde{g}(\beta u) > 0$. We thus obtain condition (ii') *via* the positive homogeneity of $\widetilde{g}$ again. This completes the proof of the claim.

We finally show that conditions (i') and (ii') are equivalent to conditions (2.i), (2.ii), and (2.iii) stated in the theorem. Clearly, in light of Theorem 3.2, condition (i') is equivalent to conditions (2.i) and (2.ii). Moreover, when $\mathcal{K}$ is nonempty, condition (ii') is equivalent to the inconsistency of the following inequality system in u:

$$q^T u < 0, \quad C_{\alpha\bullet} u \geq 0, \quad \max_{i \in \mathcal{I}} p_i^T u \leq 0.$$

In view of the expression of the matrix W in (3.2), the above inequality system is equivalent to the following linear inequality system:

$$\text{(I)}: \quad q^T u < 0, \quad C_{\alpha\bullet} u \geq 0, \quad Wu \leq 0.$$

By the Motzkin's Transposition Theorem, system (I) has no solution if and only if there exists $z = (z_1, z_2, z_3)$ with $0 < z_1 \in \mathbb{R}$ and $(z_2, z_3) \geq 0$ such that $-z_1 \cdot q + C_{\alpha\bullet}^T z_2 - W^T z_3 = 0$. The latter condition is equivalent to the existence of $(w, w') \geq 0$ such that $q - C_{\alpha\bullet}^T w + W^T w' = 0$. This shows the equivalence of conditions (ii') and (2.iii). □

3.4. Unique optimal solution to the basis pursuit denoising II-like problem

Let $f : \mathbb{R}^m \to \mathbb{R}$ be a C^1 strictly convex function. For each $i = 1, \ldots, r$, $g_i : \mathbb{R}^N \to \mathbb{R}$ is a convex PA function whose max-formulation is $g_i(x) = \max_{s=1,\ldots,\ell_i} (p_{i,s}^T x + \gamma_{i,s})$, where each $(p_{i,s}, \gamma_{i,s}) \in \mathbb{R}^N \times \mathbb{R}$. Consider the following convex optimization problem motivated by the constrained BPDN-II problem:

$$\min_{x \in \mathbb{R}^N} \ f(Ax - y) \quad \text{subject to} \quad g_1(x) \leq \eta_1, \ \ldots, \ g_r(x) \leq \eta_r, \ \text{and} \quad Cx \geq d, \tag{3.10}$$

where $\eta_1, \ldots, \eta_r$ are real numbers. We assume that this problem has an optimal solution. This optimization problem allows multiple convex PA function defined constraints, which appear in applications, *e.g.,* the sparse fused LASSO [26]; see Section 5.3 for details.

A problem similar to (3.10) is treated in [28] with one inequality constraint $\|x\|_1 \leq \eta_1$ but without the polyhedral constraint $Cx \geq d$. Under a restrictive assumption on η_1, it is shown in [28] that its solution uniqueness is reduced to that of a related basis pursuit problem. However, this reduction fails for (3.10) due to the presence of the general polyhedral constraint; see Section 5.2 for more elaboration.

We introduce more notation first. For a given feasible point $x^* \in \mathbb{R}^N$, define the index set $\mathcal{J} := \{i \in \{1, \ldots, r\} \,|\, g_i(x^*) = \eta_i\}$, which corresponds to the active constraints defined by g_i's at x^*. For each $i \in \mathcal{J}$,

define the index set $\mathcal{I}_i := \{s \in \{1, \ldots, \ell_i\} \mid p_{i,s}^T x^* + \gamma_{i,s} = g_i(x^*)\}$, and the matrix

$$
W_i := \begin{bmatrix} p_{i,s_1}^T \\ \vdots \\ p_{i,s_{|\mathcal{I}_i|}}^T \end{bmatrix}_{s_k \in \mathcal{I}_i} \in \mathbb{R}^{|\mathcal{I}_i| \times N}. \tag{3.11}
$$

Theorem 3.11. *Let $x^* \in \mathbb{R}^N$ be a feasible point of the problem (3.10). Then x^* is the unique minimizer of (3.10) if and only if the following hold:*

- *(i) $\{v \in \mathbb{R}^N \mid Av = 0, \ C_{\alpha\bullet}v = 0, \ W_i v = 0, \ \forall i \in \mathcal{J}\} = \{0\}$;*
- *(ii) there exist $w \in \mathbb{R}^m$, $w' \in \mathbb{R}_{++}^{|\alpha|}$, and $w_i'' \in \mathbb{R}^{|\mathcal{I}_i|}$ with $0 < w_i'' < \mathbf{1}$ and $\mathbf{1}^T w_i'' = 1$ for each $i \in \mathcal{J}$ such that $A^T w - C_{\alpha\bullet}^T w' + \sum_{i \in \mathcal{J}} W_i^T w_i'' = 0$;*
- *(iii) there exist $\widetilde{z} \in \mathbb{R}_+^{|\alpha|}$ and $\widetilde{z}_i' \in \mathbb{R}_+^{|\mathcal{I}_i|}$ for each $i \in \mathcal{J}$ such that $A^T \nabla f(Ax^* - y) - C_{\alpha\bullet}^T \widetilde{z} + \sum_{i \in \mathcal{J}} W_i^T \widetilde{z}_i' = 0$. Moreover, condition (iii) is equivalent to the following condition:*
- *(iv) there exist $z \in \mathbb{R}_+^{|\alpha|}$, $\theta_i \in \mathbb{R}_+$, $z_i' \in \mathbb{R}^{|\mathcal{I}_i|}$ with $0 \le z_i' \le \mathbf{1}$ and $\mathbf{1}^T z_i' = 1$ for each $i \in \mathcal{J}$ such that*

$$
A^T \nabla f(Ax^* - y) - C_{\alpha\bullet}^T z + \sum_{i \in \mathcal{J}} \theta_i \cdot W_i^T z_i' = 0.
$$

Proof. We first show that x^* is a minimizer of the problem (3.10) if and only if condition (iii) holds. Note that for each $g_i(x) = \max_{s=1,\ldots,\ell_i}(p_{i,s}^T x + \gamma_{i,s})$, the constraint $g_i(x) \le \eta_i$ is equivalent to the linear inequality constraint $p_{i,s}^T x \le \eta_i - \gamma_{i,s}$ for all s. Hence, the problem (3.10) has a polyhedral constraint. In view of the definitions of $C_{\alpha\bullet}$ and W_i for each $i \in \mathcal{J}$, it is easy to see, *e.g.*, via ([16], Thm. 3.33), that x^* is a minimizer if and only if there exist $\widetilde{z} \in \mathbb{R}_+^{|\alpha|}$ and $\widetilde{z}_i' \in \mathbb{R}_+^{|\mathcal{I}_i|}$ for each $i \in \mathcal{J}$ such that $A^T \nabla f(Ax^* - y) - C_{\alpha\bullet}^T \widetilde{z} + \sum_{i \in \mathcal{J}} W_i^T \widetilde{z}_i' = 0$, which is condition (iii). To show the equivalence of conditions (iii) and (iv), we first observe that (iv) implies (iii). Conversely, suppose (iii) holds. It suffices to show that for each $\widetilde{z}_i' \in \mathbb{R}_+^{|\mathcal{I}_i|}$, there exist $\theta_i \in \mathbb{R}_+$ and $z_i' \in \mathbb{R}^{|\mathcal{I}_i|}$ with $0 \le z_i' \le \mathbf{1}$ and $\mathbf{1}^T z_i' = 1$ such that $\widetilde{z}_i' = \theta_i \cdot z_i'$. This result is trivial when $\widetilde{z}_i' = 0$. When $0 \ne \widetilde{z}_i' \ge 0$, we choose $\theta_i := \mathbf{1}^T \widetilde{z}_i > 0$ and $z_i' := \widetilde{z}_i'/\theta_i$, which leads to the desired result.

Suppose x^* is a minimizer of the problem (3.10) or equivalently x^* satisfies condition (iii). For each $i \in \mathcal{J}$, let $\widetilde{g}_i(v) := \max_{s \in \mathcal{I}_i} p_{i,s}^T v$. Let $\mathcal{U}$ be a convex neighborhood of x^* such that for all $x \in \mathcal{U}$, $g_i(x) - g_i(x^*) = \widetilde{g}_i(x - x^*)$ for each $i \in \mathcal{J}$, $g_i(x) < \eta_i$ for each $i \in \mathcal{J}^c$, and $C_{\alpha^c \bullet} x > d_{\alpha^c}$. Then x^* is the unique minimizer of (3.10) if and only if it is a unique local minimizer of the following problem on $\mathcal{U}$:

$$
\min_{x \in \mathbb{R}^N} \ f(Ax - y) + h(x), \quad \text{subject to} \quad x \in \mathcal{U}, \ \ \widetilde{g}_i(x - x^*) \le 0, \forall i \in \mathcal{J}, \quad \text{and} \ \ C_{\alpha\bullet}(x - x^*) \ge 0,
$$

where h is the zero function, and $\widetilde{g}_i(x - x^*)$ is convex in x for each $i \in \mathcal{J}$. Applying Proposition 3.7 to the above problem with the convex set $\mathcal{C} := \mathcal{U} \cap \{x \mid C_{\alpha\bullet}(x - x^*) \ge 0, \ \widetilde{g}_i(x - x^*) \le 0, \ \forall i \in \mathcal{J}\}$, we see that x^* is the unique minimizer of (3.10) if and only if it is the unique minimizer of the following problem:

$$
\min_{x \in \mathbb{R}^N} h(x), \quad \text{subject to} \quad Ax = Ax^*, \ x \in \mathcal{U}, \ \widetilde{g}_i(x - x^*) \le 0, \forall i \in \mathcal{J}, \quad \text{and} \ \ C_{\alpha\bullet}(x - x^*) \ge 0,
$$

It is equivalent to the equation $\{v \mid Av = 0, \ C_{\alpha\bullet}v \ge 0, \ \widetilde{g}_i(v) \le 0, \ \forall i \in \mathcal{J}\} = \{0\}$ in view of the positive homogeneity of $\widetilde{g}_i$. Since $\widetilde{g}_i(v) \le 0$ is equivalent to $W_i v \le 0$ where W_i is defined in (3.11), this equation holds

if and only if $\{v \,|\, Av = 0, \ C_{\alpha\bullet}v \geq 0, \ W_i v \leq 0, \ \forall\, i \in \mathcal{J}\} = \{0\}$. Using Lemma 3.1 with

$$H = \begin{bmatrix} C_{\alpha\bullet} \\ -W_{i_1} \\ \vdots \\ -W_{i_{|\mathcal{J}|}} \end{bmatrix}_{i_k \in \mathcal{J}}$$

and a similar argument for Theorem 3.2, we deduce that $\{v \,|\, Av = 0, \ C_{\alpha\bullet}v \geq 0, \ W_i v \leq 0, \ \forall\, i \in \mathcal{J}\} = \{0\}$ holds if and only if conditions (i) and (ii) hold. This completes the proof. $\qquad\square$

3.5. Extensions to convex PA loss functions

In this subsection, we extend the results in Sections 3.2–3.4 to a convex PA loss function $f : \mathbb{R}^m \to \mathbb{R}$. It follows from (ii) of Lemma 2.1 that $f(Ax - y)$ is a convex PA function on $\mathbb{R}^N$. By virtue of this property, we show below that under this f and a convex PA function g, each of the LASSO-like problem (3.6), the BPDN-I-like problem (3.7), and the BPDN-II-like problem (3.10) can be formulated as the BP-like problem (3.3) with a new convex PA objective function or suitable polyhedral constraints.

(a) The LASSO-like problem (3.6). Define $g_\star(x) := f(Ax - y) + g(x)$. Since both $f(Ax - y)$ and $g(x)$ are convex PA functions on $\mathbb{R}^N$, so is $g_\star$ in view of (iii) of Lemma 2.1. This leads to the BP-like problem (3.3) with the objective function $g_\star$ and without the equality constraint.

(b) The BPDN-I-like problem (3.7). Since $f(Ax - y)$ is a convex PA function on $\mathbb{R}^N$, the constraint set $\{x \,|\, f(Ax - y) \leq \varepsilon\}$ is polyhedral as shown in the proof of Theorem 3.11. Hence, the problem (3.7) can be formulated as the BP-like problem (3.3) with a new polyhedral constraint.

(c) The BPDN-II-like problem (3.10). Based on the argument for the above two cases, the problem (3.10) is also transferred to the BP-like problem (3.3).

Consequently, when f is a convex PA function, the solution uniqueness of the above three problems can be determined *via* Theorem 3.2 for a given x^*.

As an example, we consider the Dantzig selector which has gained tremendous interest in high-dimensional statistics [5]: $\min_{x \in \mathbb{R}^N} \|x\|_1$ subject to $\|A^T(Ax - y)\|_\infty \leq \varepsilon$. Let $g(x) := \|x\|_1, \forall\, x \in \mathbb{R}^N$, and $f(z) := \|A^T z\|_\infty, \forall\, z \in \mathbb{R}^m$, which are both convex PA functions. Hence, the Dantzig selector can be treated as the BPDN-I-like problem (3.7) with the convex PA loss function f and the objective function g.

4. Solution existence and uniqueness of ℓ_1-norm based constrained optimization problems

Since the ℓ_1-norm is a sign-symmetric polyhedral gauge and thus a convex PL function, we apply the general results developed in Section 3 to establish solution uniqueness conditions for several important and representative ℓ_1 minimization problems, possibly subject to linear inequality constraints.

4.1. Solution existence of ℓ_1-norm based constrained optimization problems

Solution existence is a fundamental issue for ℓ_1-norm based optimization problems. For the BP-like problem, it depends on the convex PA function g, whereas for the LASSO-like and two BPDN-like problems, it depends on the function f additionally. In this subsection, we first establish some general solution existence results, and then apply them to several problems of interest with $g(x) = \|Ex\|_1$ and $f(\cdot) = \|\cdot\|^s$, which find various applications in ℓ_1 minimization. We start from certain preliminary results.

Lemma 4.1. *Let $J : \mathbb{R}^\ell \to \mathbb{R}$ be a coercive and lower semi-continuous function that is bounded below, i.e., $\inf_{u \in \mathbb{R}^\ell} J(u) > -\infty$. Let a matrix $H \in \mathbb{R}^{\ell \times N}$ and a set $\mathcal{C} \subseteq \mathbb{R}^N$ be such that $H\mathcal{C}$ is a closed set in $\mathbb{R}^\ell$. Then for any $u' \in \mathbb{R}^\ell$, the minimization problem $\inf_{x \in \mathcal{C}} J(Hx + u')$ attains an optimal solution.*

Proof. Define the set $\mathcal{W} := H\mathcal{C} + \{u'\}$ in $\mathbb{R}^\ell$ for an arbitrary u'. Since $H\mathcal{C}$ is closed, so is $\mathcal{W}$. Consider the optimization problem $(P) : \inf_{u \in \mathcal{W}} J(u)$. Since J is coercive, lower semi-continuous, and bounded below and $\mathcal{W}$ is closed, it follows from a standard argument that (P) has a minimizer $u_* \in \mathcal{W}$. Therefore, there exists $x_* \in \mathcal{C}$ such that $Hx_* + u' = u_*$. Clearly, x_* is an optimal solution to the original problem. $\qquad\square$

Corollary 4.2. *Let $A \in \mathbb{R}^{m \times N}$, $F \in \mathbb{R}^{p \times N}$, and $H := \begin{bmatrix} A \\ F \end{bmatrix} \in \mathbb{R}^{(m+p) \times N}$. Let $J_1 : \mathbb{R}^m \to \mathbb{R}$ and $J_2 : \mathbb{R}^p \to \mathbb{R}$ be two coercive and lower semi-continuous functions that are bounded below. Suppose $\mathcal{C} \subseteq \mathbb{R}^N$ is such that $H\mathcal{C}$ is a closed set in $\mathbb{R}^{m+p}$. Then for any given $y \in \mathbb{R}^m$, the following problem attains a minimizer:*

$$\inf_{x \in \mathcal{C}} \quad J_1(Ax - y) + J_2(Fx). \tag{4.1}$$

Proof. For any $z = (z_\alpha, z_\beta) \in \mathbb{R}^{m+p}$ with $z_\alpha \in \mathbb{R}^m$ and $z_\beta \in \mathbb{R}^p$, define the function $J(z) := J_1(z_\alpha) + J_2(z_\beta)$. Clearly, J is coercive, lower semi-continuous, and bounded below on $\mathbb{R}^{m+p}$. For any given y, define $z' := (-y, 0) \in \mathbb{R}^{m+p}$. Hence, $J_1(Ax - y) + J_2(Fx) = J(Hx + z')$ for any $x \in \mathbb{R}^N$. Consequently, the minimization problem in (4.1) can be equivalently written as $\inf_{x \in \mathcal{C}} J(Hx + z')$. Since $H\mathcal{C}$ is closed, it follows from Lemma 4.1 that the minimization problem in (4.1) attains an optimal solution. $\qquad\square$

By exploiting the above results, we obtain the following solution existence results for some general minimization problems motivated by the basis pursuit, LASSO, and basis pursuit denoising problems.

Theorem 4.3. *Let $A \in \mathbb{R}^{m \times N}, C \in \mathbb{R}^{p \times N}, y \in \mathbb{R}^m, d \in \mathbb{R}^p$, and $E \in \mathbb{R}^{k \times N}$ be given, and suppose the functions $J_1 : \mathbb{R}^m \to \mathbb{R}$ and $J_2 : \mathbb{R}^k \to \mathbb{R}$ are coercive, bounded below, and lower semi-continuous. Then each of the following minimization problems attains an optimal solution as long as it is feasible:*

$$(P_1): \quad \min_{x \in \mathbb{R}^N} \quad J_2(Ex) \quad \text{subject to} \quad Ax = y, \quad \text{and} \quad Cx \geq d;$$

$$(P_2): \quad \min_{x \in \mathbb{R}^N} \quad J_1(Ax - y) + J_2(Ex) \quad \text{subject to} \quad Cx \geq d;$$

$$(P_3): \quad \min_{x \in \mathbb{R}^N} \quad J_1(Ax - y) \quad \text{subject to} \quad \|E_1 x\|_1 \leq \eta_1, \ \ldots, \ \|E_r x\|_1 \leq \eta_r, \ \text{and} \ Cx \geq d,$$

where $E_i \in \mathbb{R}^{k_i \times N}$ and $\eta_i \geq 0$ for each $i = 1, \ldots, r$ in (P_3). Moreover, if $J_1 : \mathbb{R}^m \to \mathbb{R}$ is coercive and lower semi-continuous, and $J_2 : \mathbb{R}^k \to \mathbb{R}$ satisfies the conditions specified above, then the following problem attains an optimal solution as long as it is feasible:

$$(P_4): \quad \min_{x \in \mathbb{R}^N} \quad J_2(Ex) \quad \text{subject to} \quad J_1(Ax - y) \leq \varepsilon, \quad \text{and} \quad Cx \geq d, \quad \text{where} \ \ \varepsilon \in \mathbb{R}.$$

Proof. (i) Consider the problem (P_1) first. Define the (nonempty) feasible set $\mathcal{C}_1 := \{x \mid Ax = y, \ Cx \geq d\}$. Since $\mathcal{C}_1$ is a convex polyhedron, we deduce *via* Minkowski-Wyel Decomposition Theorem that $E\mathcal{C}_1$ is also a convex polyhedron and thus closed. Applying Lemma 4.1 to $J(\cdot) = J_2(\cdot)$, $H = E$, $u' = 0$, and $\mathcal{C} = \mathcal{C}_1$, we conclude that this problem attains a minimizer.

(ii) We then consider the problem (P_2). Clearly, the (nonempty) feasible set $\mathcal{C}_2 := \{x \mid Cx \geq d\}$ is a convex polyhedron. Let $H := \begin{bmatrix} A \\ E \end{bmatrix} \in \mathbb{R}^{(m+k) \times N}$. Hence, $H\mathcal{C}_2$ is closed. It follows from Corollary 4.2 directly that a minimizer exists.

(iii) We next consider the problem (P_3). As indicated in the proof of Theorem 3.11, since each $g_i(x) = \|E_i x\|_1$ is a convex PL function, the (nonempty) feasible set $\mathcal{C}_3 := \{x \mid g_1(x) \leq \eta_1, \ \ldots, \ g_r(x) \leq \eta_r, \ \text{and} \ Cx \geq d\}$ is a polyhedron. Therefore, $A\mathcal{C}_3$ is closed. By letting $J(\cdot) = J_1(\cdot)$, $H = A$, and $u' = -y$, and $\mathcal{C} = \mathcal{C}_3$, the desired result follows readily from Lemma 4.1.

(iv) Lastly, we consider the problem (P_4). Let the (nonempty) set $\mathcal{D} := \{x \in \mathbb{R}^N \mid J_1(Ax - y) \leq \varepsilon\}$, and define $\mathcal{W} := R(A^T) \cap \mathcal{D}$. We claim that $\mathcal{D} = \mathcal{W} + N(A)$. It is straightforward to show that $\mathcal{D} \supseteq \mathcal{W} + N(A)$. For the converse, consider an arbitrary $x \in \mathcal{D}$. Note that there exist unique vectors $u \in R(A^T)$ and $v \in N(A)$ such that $x = u + v$. Since $Ax = Au$, we have $u \in \mathcal{D}$. Therefore, $u \in \mathcal{W}$ so that $x \in \mathcal{W} + N(A)$. This completes the proof of the claim.

We next show that $\mathcal{W}$ is a compact set. Toward this end, we note that since $J_1(\cdot)$ is lower semi-continuous, $J_1(Ax - y)$ is also lower semi-continuous in x. By observing that $\mathcal{D}$ is the sub-level set of a lower semi-continuous function, we deduce that $\mathcal{D}$ is closed. Sine $R(A^T)$ is also closed, so is $\mathcal{W}$. We show next that $\mathcal{W}$ is bounded. Since $J_1(\cdot)$ is coercive, we see *via* the definition of the set $\mathcal{D}$ that $A\mathcal{D}$ is bounded. Suppose, by contradiction, that $\mathcal{W}$ is unbounded. Then there exists a sequence (x_n) in $\mathcal{W} := R(A^T) \cap \mathcal{D}$ such that $(\|x_n\|) \to \infty$. Without loss of generality, we assume that $(x_n/\|x_n\|)$ converges to z_* with $\|z_*\| = 1$. Since (Ax_n) is in $A\mathcal{D}$, it is thus bounded so that $(Ax_n/\|x_n\|) \to 0$. This implies that $Az_* = 0$ or equivalently $z_* \in N(A)$. Furthermore, since $(x_n/\|x_n\|)$ is a convergent sequence in the closed set $R(A^T)$, we have $z_* \in R(A^T)$. In view of $N(A) \cap R(A^T) = \{0\}$, we have $z_* = 0$, a contradiction. Hence, $\mathcal{W}$ is bounded and thus compact.

Since $\mathcal{D} = \mathcal{W} + N(A)$, we have $E\mathcal{D} = E\mathcal{W} + EN(A)$. Note that $E\mathcal{W}$ is compact, and that $EN(A)$ is a subspace and thus closed. Consequently, $E\mathcal{D}$ is closed. Since the (nonempty) feasible set $\mathcal{C}_4 = \mathcal{D} \cap \mathcal{P}$, where $\mathcal{P} := \{x \mid Cx \geq d\}$, we have $E\mathcal{C}_4 = E\mathcal{D} \cap E\mathcal{P}$. As both $E\mathcal{D}$ and $E\mathcal{P}$ are closed, so is $E\mathcal{C}_4$. It follows from the similar argument as before that (P_4) attains an optimal solution. □

We apply the above theorem to several representative ℓ_1 minimization problems.

Corollary 4.4. *Let $A \in \mathbb{R}^{m \times N}, C \in \mathbb{R}^{p \times N}, y \in \mathbb{R}^m$, and $d \in \mathbb{R}^p$ be given, $g(x) = \|Ex\|_1$ for some $E \in \mathbb{R}^{k \times N}$, and $f(u) = \|u\|^s, \forall u \in \mathbb{R}^m$ where $\|\cdot\|$ is a norm on $\mathbb{R}^m$ and $s > 0$. Then each of the following minimization problems attains an optimal solution as long as it is feasible:*

$$\text{BP-like problem}: \quad \min_{x \in \mathbb{R}^N} \; g(x) \quad \text{subject to} \quad Ax = y, \quad \text{and} \;\; Cx \geq d;$$

$$\text{LASSO-like problem}: \quad \min_{x \in \mathbb{R}^N} \; f(Ax - y) + g(x) \quad \text{subject to} \quad Cx \geq d;$$

$$\text{BPDN-I like problem}: \quad \min_{x \in \mathbb{R}^N} \; g(x) \quad \text{subject to} \quad f(Ax - y) \leq \varepsilon, \quad \text{and} \;\; Cx \geq d, \;\; \text{where} \;\; \varepsilon > 0;$$

$$\text{BPDN-II-like problem}: \quad \min_{x \in \mathbb{R}^N} \; f(Ax - y) \quad \text{subject to} \quad \|E_1 x\|_1 \leq \eta_1, \ldots, \|E_r x\|_1 \leq \eta_r, \quad \text{and} \;\; Cx \geq d,$$

where $E_i \in \mathbb{R}^{k_i \times N}$ and $\eta_i \geq 0$ for each $i = 1, \ldots, r$ in the last problem.

Proof. It is a direct consequence of Theorem 4.3 by noting that $J_1(\cdot) = \|\cdot\|^s$ with $s > 0$ in f and $J_2(\cdot) = \|\cdot\|_1$ in g are coercive, continuous (thus lower semi-continuous), and bounded below by zero. □

4.2. Properties of ℓ_1-norm based convex PA functions

In order to apply the general results developed in Section 3 to an ℓ_1-norm based convex PA function, it is crucial to find the matrix W defined in (3.2) associated with this function for a given vector. Toward this end, we first establish this matrix for the ℓ_1-norm. Note that the max-formulation of the ℓ_1-norm on $\mathbb{R}^k$ is given by $g(z) := \|z\|_1 = \max_{1,\ldots,2^k} p_i^T z, \forall z \in \mathbb{R}^k$, where each

$$p_i \in \left\{ \left(\pm 1, \pm 1, \ldots, \pm 1 \right)^T \right\} \subset \mathbb{R}^k. \tag{4.2}$$

For a given $z^* \in \mathbb{R}^k$, let $\mathcal{S}$ be the support of z^* and $\mathcal{S}^c$ be its complement. Further, define the index set $\mathcal{I} := \{i \in \{1, \ldots, 2^k\} \mid p_i^T z^* = \|z^*\|_1\}$, and $b := \text{sgn}(z_{\mathcal{S}}^*) \in \mathbb{R}^{|\mathcal{S}|}$. Here $|\mathcal{I}| = 2^{|\mathcal{S}^c|}$. Using the definitions of $\mathcal{S}$ and $\mathcal{S}^c$, we can decompose $g(z)$ as the sum of two ℓ_1-norms on $\mathbb{R}^{|\mathcal{S}|}$ and $\mathbb{R}^{|\mathcal{S}^c|}$ respectively, *i.e.*, $g(z) = \|z\|_1 = \|z_{\mathcal{S}}\|_1 + \|z_{\mathcal{S}^c}\|_1, \forall z \in \mathbb{R}^k$. For notational purpose, define $g_{\mathcal{S}}(z_{\mathcal{S}}) := \|z_{\mathcal{S}}\|_1$ and $g_{\mathcal{S}^c}(z_{\mathcal{S}^c}) := \|z_{\mathcal{S}^c}\|_1$.

Hence the subdifferentials $\partial g_{\mathcal{S}}(z_{\mathcal{S}}^*) = \{b\}$, and $\partial g_{\mathcal{S}^c}(z_{\mathcal{S}^c}^*) = \partial g_{\mathcal{S}^c}(0) = \{u \in \mathbb{R}^{|\mathcal{S}^c|} \mid \|u\|_\infty \leq 1\}$. By the comment after equation (3.2), it is easy to verify that the matrix defined in (3.2) associated with $\|z\|_1$ at z^* is given by $\widehat{W} = \begin{bmatrix} \widehat{W}_{\bullet\mathcal{S}} & \widehat{W}_{\bullet\mathcal{S}^c} \end{bmatrix} \in \mathbb{R}^{|\mathcal{I}| \times k}$, where $\widehat{W}_{\bullet\mathcal{S}} = \mathbf{1} \cdot b^T$, and each row of $\widehat{W}_{\bullet\mathcal{S}^c}$ is of the form $(\pm 1, \ldots, \pm 1) \in \mathbb{R}^{|\mathcal{S}^c|}$.

For example, if $|\mathcal{S}^c| = 2$, $\widehat{W}_{\bullet\mathcal{S}^c} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ -1 & 1 \\ -1 & -1 \end{bmatrix}$. We collect several properties of $\widehat{W}_{\bullet\mathcal{S}^c}$ in the following lemma.

These properties will be used for the latter development; see the proofs of Lemma 4.6 and Propositions 4.7 and 4.8.

Lemma 4.5. *For the given $z^* \in \mathbb{R}^k$, the matrix $\widehat{W}_{\bullet\mathcal{S}^c} \in \mathbb{R}^{|\mathcal{I}| \times |\mathcal{S}^c|}$ defined above satisfies:*

(i) *the columns of $\widehat{W}_{\bullet\mathcal{S}^c}$ are linearly independent;*

(ii) *for any row $\widehat{W}_{i\mathcal{S}^c}$, there is another row $\widehat{W}_{j\mathcal{S}^c}$ with $i \neq j$ such that $\widehat{W}_{j\mathcal{S}^c} = -\widehat{W}_{i\mathcal{S}^c}$;*

(iii) $\mathrm{conv}\{\widehat{W}_{i\mathcal{S}^c}^T \mid i = 1, \ldots, |\mathcal{I}|\} = \partial g_{\mathcal{S}^c}(0) = \{u \in \mathbb{R}^{|\mathcal{S}^c|} \mid \|u\|_\infty \leq 1\}$, *and*

$$\left\{ \sum_{i=1}^{|\mathcal{I}|} \lambda_i \cdot \widehat{W}_{i\mathcal{S}^c}^T \,\Big|\, \sum_{i=1}^{|\mathcal{I}|} \lambda_i = 1, \ \ \lambda_i > 0, \ \forall\, i = 1, \ldots, |\mathcal{I}| \right\} = \left\{ u \in \mathbb{R}^{|\mathcal{S}^c|} \mid \|u\|_\infty < 1 \right\}.$$

Proof. Statements (i) and (ii) are trivial. To show the first part of statement (iii), it follows from the comment after equation (3.2) that the columns of $\widehat{W}_{\bullet\mathcal{S}^c}^T$ are convex hull generators (or vertices/extreme points) of $\partial g_{\mathcal{S}^c}(0)$, which is the closed unit ball with respect to the infinity-norm $\|\cdot\|_\infty$. Lastly, we deduce from Lemma 3.3 and the first part of (iii) that

$$\left\{ \sum_{i=1}^{|\mathcal{I}|} \lambda_i \cdot \widehat{W}_{i\mathcal{S}^c}^T \,\Big|\, \sum_{i=1}^{|\mathcal{I}|} \lambda_i = 1, \ \ \lambda_i > 0, \ \forall\, i = 1, \ldots, |\mathcal{I}| \right\}$$

$$= \mathrm{ri}\Big(\mathrm{conv}\{\widehat{W}_{i\mathcal{S}^c}^T \mid i = 1, \ldots, |\mathcal{I}|\} \Big) = \mathrm{ri}\Big(\{u \in \mathbb{R}^{|\mathcal{S}^c|} \mid \|u\|_\infty \leq 1\} \Big)$$

$$= \mathrm{int}\Big(\{u \in \mathbb{R}^{|\mathcal{S}^c|} \mid \|u\|_\infty \leq 1\} \Big) = \left\{ u \in \mathbb{R}^{|\mathcal{S}^c|} \mid \|u\|_\infty < 1 \right\},$$

where the second-to-last equation follows from the fact that the unit closed ball with respect to the infinity-norm $\|\cdot\|_\infty$ has nonempty interior. $\square$

Motivated by generalized ℓ_1 minimization, we consider a sign-symmetric polyhedral gauge of the form $g(x) = \|Ex\|_1$ for a (nonzero) matrix $E \in \mathbb{R}^{k \times N}$. Many ℓ_1-norm based convex PL functions arising from applications can be represented by this form, *e.g.,* ℓ_1-trend filtering [12], sparse fused LASSO [26], and generalized LASSO [24]; see Sections 4.3.1 and 5.3 for more discussions and examples. For a given $x^* \in \mathbb{R}^N$, let $\mathcal{S}$ denote the support of Ex^*, *i.e.,* $\mathcal{S} = \{i \in \{1, \ldots, k\} \mid (Ex^*)_i \neq 0\}$, and $\mathcal{S}^c$ be its complement. Further, define $\widehat{b} := \mathrm{sgn}((Ex^*)_{\mathcal{S}})$ and the index set $\mathcal{I} := \{i \in \{1, \ldots, 2^k\} \mid p_i^T Ex^* = \|Ex^*\|_1\}$, where p_i's are defined in (4.2) for the max-formulation of the ℓ_1-norm. Here $|\mathcal{I}| = 2^{|\mathcal{S}^c|}$. In light of the comment after equation (3.2), we obtain the matrix $\widehat{W}$ defined in (3.2) associated with $\|\cdot\|_1$ at Ex^* as $\widehat{W} = \begin{bmatrix} \widehat{W}_{\bullet\mathcal{S}} & \widehat{W}_{\bullet\mathcal{S}^c} \end{bmatrix} \in \mathbb{R}^{|\mathcal{I}| \times k}$, where $\widehat{W}_{\bullet\mathcal{S}} = \mathbf{1} \cdot \widehat{b}^T$ and $\widehat{W}_{\bullet\mathcal{S}^c} \in \mathbb{R}^{|\mathcal{I}| \times |\mathcal{S}^c|}$ is the matrix whose columns of its transpose form the vertices of the closed unit ball in $\mathbb{R}^{|\mathcal{S}^c|}$ with respect to $\|\cdot\|_\infty$. Note that the matrix $\widehat{W}_{\bullet\mathcal{S}^c}$ satisfies the conditions given in Lemma 4.5. In view of $\partial g(x^*) = E^T \partial \|\cdot\|_1(Ex^*)$, we see that the matrix W associated with the function g at x^* is

$$W = \widehat{W} \cdot E = \begin{bmatrix} \widehat{W}_{\bullet\mathcal{S}} & \widehat{W}_{\bullet\mathcal{S}^c} \end{bmatrix} \begin{bmatrix} E_{\mathcal{S}\bullet} \\ E_{\mathcal{S}^c\bullet} \end{bmatrix} = \mathbf{1} \cdot b^T + \widehat{W}_{\bullet\mathcal{S}^c} E_{\mathcal{S}^c\bullet} \in \mathbb{R}^{|\mathcal{I}| \times N}, \quad b := E_{\mathcal{S}\bullet}^T \widehat{b} \in \mathbb{R}^N. \quad (4.3)$$

By virtue of these results, we obtain the following lemma which characterizes the null space of W.

Lemma 4.6. *Let the matrix W be defined in (4.3) for the function $g(x) = \|Ex\|_1$ at x^*. For a given $v \in \mathbb{R}^N$, $Wv = 0$ if and only if $b^T v = 0$ and $E_{\mathcal{S}^c \bullet} v = 0$.*

Proof. The "if" part is trivial, and we show the "only if" part only. Suppose $Wv = 0$. Let $\alpha := b^T v \in \mathbb{R}$ and $u := E_{\mathcal{S}^c \bullet} v \in \mathbb{R}^{|\mathcal{S}^c|}$. Hence, we have $Wv = \alpha \cdot \mathbf{1} + \widehat{W}_{\bullet \mathcal{S}^c} u = 0$. It follows from (ii) of Lemma 4.5 that the vector $W_{\bullet \mathcal{S}^c} u$ has two elements of the same absolute value (which is possibly zero) but opposite signs. This shows that α is zero, *i.e.,* $b^T v = 0$. This further implies that $W_{\bullet \mathcal{S}^c} u = 0$. Since the columns of $W_{\bullet \mathcal{S}^c}$ are linearly independent (*cf.* (i) of Lem. 4.5), we obtain $u = 0$ or equivalently $E_{\mathcal{S}^c \bullet} v = 0$. $\square$

4.3. Solution uniqueness of convex optimization problems involving $\|Ex\|_1$

Through this subsection, we let $g(x) = \|Ex\|_1, \forall\, x \in \mathbb{R}^N$ for a (nonzero) matrix $E \in \mathbb{R}^{k \times N}$, and let $\mathcal{P} := \{x \in \mathbb{R}^N \,|\, Cx \ge d\}$ be a nonempty polyhedron where $C \in \mathbb{R}^{p \times N}$ and $d \in \mathbb{R}^p$. Furthermore, for a given x^*, recall the definitions of the index sets α and $\mathcal{S}$ in Section 4.2, and the definitions of the matrix W and the vector b in (4.3). We first consider the BP-like problem (3.3) involving the function g.

Proposition 4.7. *Let $g(x) = \|Ex\|_1$, and x^* be a feasible point of the optimization problem (3.3). Then x^* is the unique minimizer if and only if the following conditions hold:*

(a) the matrix $\begin{bmatrix} A \\ C_{\alpha \bullet} \\ E_{\mathcal{S}^c \bullet} \end{bmatrix}$ has full column rank and

(b) there exist $u \in \mathbb{R}^m$, $u' \in \mathbb{R}^{|\alpha|}_{++}$, and $u'' \in \mathbb{R}^{|\mathcal{S}^c|}$ with $\|u''\|_\infty < 1$ such that $A^T u + C^T_{\alpha \bullet} u' - E^T_{\mathcal{S}^c \bullet} u'' = b$.

Proof. By Theorem 3.2, it suffices to show that conditions (i) and (iii) of Theorem 3.2 hold if and only if conditions (a) and (b) hold, where we recall that (i) $\{v \,|\, Av = 0,\ C_{\alpha \bullet} v = 0,\ Wv = 0\} = \{0\}$, and (iii) there exist $w \in \mathbb{R}^m$, $w' \in \mathbb{R}^{|\alpha|}_{++}$, and $w'' \in \mathbb{R}^{|\mathcal{I}|}$ with $0 < w'' < \mathbf{1}$ and $\mathbf{1}^T w'' = 1$ such that $A^T w - C^T_{\alpha \bullet} w' + W^T w'' = 0$.

"Only if". Suppose conditions (i) and (iii) of Theorem 3.2 hold with suitable w, w', and w'' satisfying the specified conditions. In view of the expression of W given in (4.3), we have $W^T w'' = b \cdot \mathbf{1}^T w'' + E^T_{\mathcal{S}^c \bullet} \widehat{W}^T_{\bullet \mathcal{S}^c} w'' = b + E^T_{\mathcal{S}^c \bullet} u''$, where $u'' := \widehat{W}^T_{\bullet \mathcal{S}^c} w''$ and we use $\mathbf{1}^T w'' = 1$. By the second part of statement (iii) Lemma 4.5, we obtain $\|u''\|_\infty < 1$. Hence, letting $u = -w$ and $u' = w' > 0$, we have $A^T u + C^T_{\alpha \bullet} u' - E^T_{\mathcal{S}^c \bullet} u'' = b$. This yields condition (b). Moreover, it follows from condition (i) and Lemma 4.6 that

$$\{v \,|\, Av = 0,\ C_{\alpha \bullet} v = 0,\ b^T v = 0,\ E_{\mathcal{S}^c \bullet} v = 0\} = \{0\}. \tag{4.4}$$

We claim that equation (4.4) implies condition (a). Suppose, in contrast, that (a) fails under (4.4), *i.e.,* there exists $v \ne 0$ such that $Av = 0, C_{\alpha \bullet} v = 0$, and $E_{\mathcal{S}^c \bullet} v = 0$. In view of condition (b), we have

$$v^T b = v^T \big(A^T u + C^T_{\alpha \bullet} u' - E^T_{\mathcal{S}^c \bullet} u''\big) = 0.$$

This gives rise to a contradiction to (4.4). Hence, condition (a) holds.

"If". Suppose conditions (a) and (b) hold. Note that condition (a) implies that $\{v \,|\, Av = 0,\ C_{\alpha \bullet} v = 0,\ b^T v = 0,\ E_{\mathcal{S}^c \bullet} v = 0\} = \{0\}$. By Lemma 4.6, we have $\{v \,|\, Av = 0,\ C_{\alpha \bullet} v = 0,\ Wv = 0\} = \{0\}$, which is condition (i) of Theorem 3.2. Furthermore, we deduce from condition (b) that there exist $u \in \mathbb{R}^m$, $u' \in \mathbb{R}^{|\alpha|}_{++}$, and $u'' \in \mathbb{R}^{|\mathcal{S}^c|}$ with $\|u''\|_\infty < 1$ such that $A^T u + C^T_{\alpha \bullet} u' - E^T_{\mathcal{S}^c \bullet} u'' = b$. By letting $w = -u$ and $w' = u'$, we have $A^T w - C^T_{\alpha \bullet} w' + b + E^T_{\mathcal{S}^c \bullet} u'' = 0$. Since $\|u''\|_\infty < 1$, we deduce *via* the second part of (iii) of Lemma 4.5 that there exists $w'' \in \mathbb{R}^{|\mathcal{I}|}$ with $0 < w'' < \mathbf{1}$ and $\mathbf{1}^T w'' = 1$ such that $u'' = W^T_{\bullet \mathcal{S}^c} w''$. Therefore, $b + E^T_{\mathcal{S}^c \bullet} u'' = \big(b \cdot \mathbf{1}^T + E^T_{\mathcal{S}^c \bullet} W^T_{\bullet \mathcal{S}^c}\big) w'' = W^T w''$, where the second equation follows from (4.3). This gives rise to condition (ii) of Theorem 3.2. $\square$

The necessary and sufficient conditions for unique optimal solutions to the LASSO-like problem (3.6) are given in the following proposition.

Proposition 4.8. *Let $g(x) = \|Ex\|_1$, $f : \mathbb{R}^m \to \mathbb{R}$ be a C^1 strictly convex function, and $x^* \in \mathbb{R}^N$ be a feasible point of the problem (3.6). Then x^* is the unique minimizer of (3.6) if and only if conditions (a) and (b) of Proposition 4.7 and the following condition hold:*

(c) *there exist $\widetilde{u} \in \mathbb{R}_+^{|\alpha|}$ and $\widetilde{u}' \in \mathbb{R}^{|\mathcal{S}^c|}$ with $\|\widetilde{u}'\|_\infty \leq 1$ such that $A^T \nabla f(Ax^* - y) - C_{\alpha\bullet}^T \widetilde{u} + b + E_{\mathcal{S}^c\bullet}^T \widetilde{u}' = 0$.*

Proof. In light of Theorem 3.8 and Proposition 4.7, we only need to show that condition (iii) of Theorem 3.8 is equivalent to condition (c) of this proposition. Using (iii) of Lemma 4.5, we deduce that $\|\widetilde{u}'\|_\infty \leq 1$ for some $\widetilde{u}' \in \mathbb{R}^{|\mathcal{S}^c|}$ if and only if there exists $w' \in \mathbb{R}^{|\mathcal{I}|}$ with $0 \leq w' \leq \mathbf{1}$ and $\mathbf{1}^T w' = 1$ such that $\widetilde{u}' = \widehat{W}_{\bullet\mathcal{S}^c}^T w'$. Applying this result and the similar argument in the proof of Proposition 4.7, we conclude that condition (iii) of Theorem 3.8 is equivalent to condition (c) of the proposition. $\square$

The following proposition pertains to the BPDN-I-like problem (3.7); condition (2.c) given below follows from statement (c) of Remark 3.10. Its proof is rather straightforward and thus omitted.

Proposition 4.9. *Let $g(x) = \|Ex\|_1$, $f : \mathbb{R}^m \to \mathbb{R}$ be a C^1 strictly convex function, and $x^* \in \mathbb{R}^N$ be a feasible point of the problem (3.7).*

C.1 *Suppose $f(Ax^* - y) < \varepsilon$. Then x^* is the unique minimizer of (3.7) if and only if $\begin{bmatrix} C_{\alpha\bullet} \\ E_{\mathcal{S}^c\bullet} \end{bmatrix}$ has full column rank, and there exist $u \in \mathbb{R}_{++}^{|\alpha|}$ and $u' \in \mathbb{R}^{|\mathcal{S}^c|}$ with $\|u'\|_\infty < 1$ such that $C_{\alpha\bullet}^T u = b + E_{\mathcal{S}^c\bullet}^T u'$.*

C.2 *Suppose $f(Ax^* - y) = \varepsilon$. Then x^* is the unique minimizer of (3.7) if and only if conditions (a) and (b) of Proposition 4.7 and the following condition hold:*

(2.c) *If $\mathcal{K} := \{v \in \mathbb{R}^N \mid (\nabla f(Ax^* - y))^T Av < 0, \ C_{\alpha\bullet}v \geq 0\}$ is nonempty, then there exist a positive real number θ, $\widetilde{u} \in \mathbb{R}_+^{|\alpha|}$, and $\widetilde{u}' \in \mathbb{R}^{|\mathcal{S}^c|}$ with $\|\widetilde{u}'\|_\infty \leq 1$ such that $\theta \cdot A^T \nabla f(Ax^* - y) - C_{\alpha\bullet}^T \widetilde{u} + b + E_{\mathcal{S}^c\bullet}^T \widetilde{u}' = 0$.*

The next result characterizes solution uniqueness of the following BPDN-II-like problem:

$$\min_{x \in \mathbb{R}^N} \ f(Ax - y) \quad \text{subject to} \quad g(x) \leq \eta, \ \text{and} \ Cx \geq d. \tag{4.5}$$

This problem is a special case of the problem (3.10) with $r = 1$, $g(x) = \|Ex\|_1$, and a constant $\eta > 0$.

Proposition 4.10. *Let $g(x) = \|Ex\|_1$, $f : \mathbb{R}^m \to \mathbb{R}$ be a C^1 strictly convex function, and $x^* \in \mathbb{R}^N$ be a feasible point of the problem (4.5).*

C.1 *Suppose $g(x^*) < \eta$. Then x^* is the unique minimizer of (4.5) if and only if the matrix $\begin{bmatrix} A \\ C_{\alpha\bullet} \end{bmatrix}$ has full column rank, and there exist $w \in \mathbb{R}^m$, $w' \in \mathbb{R}_{++}^{|\alpha|}$ and $u \in \mathbb{R}_+^{|\alpha|}$ such that $A^T w = C_{\alpha\bullet}^T w'$ and $A^T \nabla f(Ax^* - y) = C_{\alpha\bullet}^T u$;*

C.2 *Suppose $g(x^*) = \eta$. Then x^* is the unique minimizer of (4.5) if and only if conditions (a) and (b) of Proposition 4.7 and the following condition hold:*

(2.c) *there exist $\widetilde{u} \in \mathbb{R}_+^{|\alpha|}$, $\mu \in \mathbb{R}_+$, and $\widetilde{u}' \in \mathbb{R}^{|\mathcal{S}^c|}$ with $\|\widetilde{u}'\|_\infty \leq 1$ such that $A^T \nabla f(Ax^* - y) - C_{\alpha\bullet}^T \widetilde{u} + \mu \cdot \left(b + E_{\mathcal{S}^c\bullet}^T \widetilde{u}' \right) = 0$.*

Proof. The proof for the case C.1 follows directly from Theorem 3.11 by setting the index $\mathcal{J} = \emptyset$. For the case C.2, it suffices to show that condition (iv) of Theorem 3.11 is equivalent to condition (2.c) of this proposition. For this purpose, it follows from the expression of the matrix W in (4.3), (iii) of Lemma 4.5, and a similar argument for Proposition 4.7 that (a) for any $z' \geq 0$ with $\mathbf{1}^T z' = 1$, there exists $\widetilde{u}' \in \mathbb{R}^{|\mathcal{S}^c|}$ with $\|\widetilde{u}'\|_\infty \leq 1$ such

that $W^T z' = b + E_{\mathcal{S}^c \bullet}^T \widetilde{u}'$; and (b) for any $\widetilde{u}' \in \mathbb{R}^{|\mathcal{S}^c|}$ with $\|\widetilde{u}'\|_\infty \le 1$, there exists $z' \ge 0$ with $\mathbf{1}^T z' = 1$ such that $W^T z' = b + E_{\mathcal{S}^c \bullet}^T \widetilde{u}'$. Therefore, condition (iv) of Theorem 3.11 is equivalent to (2.c) of this proposition. □

4.3.1. Extension to a sum of ℓ_1-norm based convex PA functions

The results developed for $\|Ex\|_1$ can be extended to a wide range of ℓ_1-norm based convex PA functions; see Section 4.4 for the ℓ_1-norm. In this subsection, we consider a convex PL function of the following form which appears in such applications as the sparse fused LASSO [14, 26] (*cf.* Sect. 5.3):

$$g(x) = \sum_{i=1}^{q} \|F_i x\|_1, \qquad \forall\, x \in \mathbb{R}^N, \tag{4.6}$$

where each matrix $F_i \in \mathbb{R}^{k_i \times N}$, $i = 1, \ldots q$. Letting $k := k_1 + \cdots + k_q$, define the augmented matrix $F :=$ $\begin{bmatrix} F_1 \\ \vdots \\ F_q \end{bmatrix} \in \mathbb{R}^{k \times N}$. Then $g(x) = \|Fx\|_1, \forall\, x \in \mathbb{R}^N$. Consequently, all the results developed for $\|Ex\|_1$ hold for $g(x) = \|Fx\|_1$. Particularly, for a given $x^* \in \mathbb{R}^N$, let $\mathcal{S}_i$ be the support of $F_i x^*$ for each i, and $\mathcal{S}$ be the support of Fx^*. Then $F_{\mathcal{S}\bullet} = \begin{bmatrix} F_{\mathcal{S}_1 \bullet} \\ \vdots \\ F_{\mathcal{S}_q \bullet} \end{bmatrix}$. Similarly, $F_{\mathcal{S}^c \bullet} = \begin{bmatrix} F_{\mathcal{S}_1^c \bullet} \\ \vdots \\ F_{\mathcal{S}_q^c \bullet} \end{bmatrix}$. Further, let $\widehat{b} := \mathrm{sgn}((Fx^*)_{\mathcal{S}})$, and $b := F_{\mathcal{S}\bullet}^T \widehat{b}$. Therefore, Propositions 4.7–4.10 are readily extended to the function g in (4.6) using these matrices and vectors.

4.4. Solution uniqueness of convex optimization problems involving the ℓ_1-norm

Through this subsection, we consider the case where g is the ℓ_1-norm, *i.e.*, $g(x) = \|x\|_1, \forall\, x \in \mathbb{R}^N$. By setting E as the $N \times N$ identity matrix and applying the results in Section 4.3 to the ℓ_1-norm, we see that for a given $x^* \in \mathbb{R}^N$, the index set $\mathcal{S}$ is the support of x^*, the vector $\widehat{b} = \mathrm{sgn}(x_{\mathcal{S}}^*) \in \mathbb{R}^{|\mathcal{S}|}$, the vector $b = E_{\mathcal{S}\bullet}^T \widehat{b} = (b_{\mathcal{S}}, b_{\mathcal{S}^c}) = (\widehat{b}, 0) \in \mathbb{R}^N$, $E_{\mathcal{S}^c \bullet} = [0\ I_{\mathcal{S}^c \mathcal{S}^c}]$, and $E_{\mathcal{S}^c \bullet} v = v_{\mathcal{S}^c}$.

Corollary 4.11. *Let $g(x) = \|x\|_1$, and x^* be a feasible point of the optimization problem (3.3). Then x^* is the unique minimizer if and only if the following conditions hold:*

(i) the matrix $\begin{bmatrix} A_{\bullet \mathcal{S}} \\ C_{\alpha \mathcal{S}} \end{bmatrix}$ has full column rank and

(ii) there exist $u \in \mathbb{R}^m$ and $u' \in \mathbb{R}_{++}^{|\alpha|}$ such that $A_{\bullet \mathcal{S}}^T u + C_{\alpha \mathcal{S}}^T u' = \widehat{b}$ and $\left\| A_{\bullet \mathcal{S}^c}^T u + C_{\alpha \mathcal{S}^c}^T u' \right\|_\infty < 1$.

Proof. We apply Proposition 4.7 to this case with $E = I_N$. Since $E_{\mathcal{S}^c \bullet} = [0\ I_{\mathcal{S}^c \mathcal{S}^c}]$, we see that

$$\begin{bmatrix} A \\ C_{\alpha \bullet} \\ E_{\mathcal{S}^c \bullet} \end{bmatrix} = \begin{bmatrix} A_{\bullet \mathcal{S}} & A_{\bullet \mathcal{S}^c} \\ C_{\alpha \mathcal{S}} & C_{\alpha \mathcal{S}^c} \\ 0 & I_{\mathcal{S}^c \mathcal{S}^c} \end{bmatrix}.$$

Since $I_{\mathcal{S}^c \mathcal{S}^c}$ is invertible, condition (a) of Proposition 4.7 holds if and only if condition (i) holds, *i.e.*, $\begin{bmatrix} A_{\bullet \mathcal{S}} \\ C_{\alpha \mathcal{S}} \end{bmatrix}$ has full column rank. Using $E_{\mathcal{S}^c \bullet} = [0\ I_{\mathcal{S}^c \mathcal{S}^c}]$ and $b = (b_{\mathcal{S}}, b_{\mathcal{S}^c})$ with $b_{\mathcal{S}} = \widehat{b}$ and $b_{\mathcal{S}^c} = 0$, we deduce that condition (b) of Proposition 4.7 holds if and only if there exist $u \in \mathbb{R}^m$, $u' \in \mathbb{R}_{++}^{|\alpha|}$, and $u'' \in \mathbb{R}^{|\mathcal{S}^c|}$ with $\|u''\|_\infty < 1$ such that

$$A_{\bullet \mathcal{S}}^T u + C_{\alpha \mathcal{S}}^T u' = \widehat{b}, \quad \text{and} \quad A_{\bullet \mathcal{S}^c}^T u + C_{\alpha \mathcal{S}^c}^T u' = u''.$$

The latter equation given above is equivalent to $\left\|A_{\bullet S^c}^T u + C_{\alpha S^c}^T u'\right\|_\infty < 1$. This completes the proof. $\square$

Remark 4.12. The polyhedron constrained ℓ_1 minimization problem considered in Corollary 4.11 can be formulated as the following equivalent linear program:

$$\min_{(x,v)\in\mathbb{R}^N\times\mathbb{R}^N} \mathbf{1}^T v \qquad \text{subject to} \quad v \geq 0, \quad -v \leq x \leq v, \quad Ax = y, \quad Cx \geq d,$$

where $x \in \mathbb{R}^N$ is the original decision variable, and $v \in \mathbb{R}^N$ is the slack variable used for transforming the ℓ_1-norm into linear inequalities. Hence, the solution uniqueness conditions for linear programs, *e.g.*, ([13], Thm. 2), can be employed to obtain uniqueness conditions which are equivalent to those in Corollary 4.11. However, the linear program based conditions are described by large-size linear inequality systems due to the presence of the slack variable v, and they often require the knowledge of optimizers of the associated dual problem. Therefore, they are more expensive to verify numerically than those in Corollary 4.11. Furthermore, the conditions in Corollary 4.11 demonstrate better geometric meanings and yield simpler conditions for specific polyhedral constraints, in comparison with the linear program based uniqueness conditions.

The necessary and sufficient conditions for unique optimal solutions to the LASSO-like and the BPDN-I/II-like problems are presented below. Their proofs are omitted since they follow directly from Propositions 4.8–4.10 and particular structure associated with the ℓ_1-norm shown in Corollary 4.11.

Corollary 4.13. *Let $g(x) = \|x\|_1$, $f : \mathbb{R}^m \to \mathbb{R}$ be a C^1 strictly convex function, and $x^* \in \mathbb{R}^N$ be a feasible point of the problem (3.6). Then x^* is the unique minimizer of (3.6) if and only if conditions (i)–(ii) of Corollary 4.11 and the following condition hold: there exists $\widetilde{u} \in \mathbb{R}_+^{|\alpha|}$ such that $A_{\bullet S}^T \nabla f(Ax^* - y) - C_{\alpha S}^T \widetilde{u} + \widehat{b} = 0$, and $\left\|A_{\bullet S^c}^T \nabla f(Ax^* - y) - C_{\alpha S^c}^T \widetilde{u}\right\|_\infty \leq 1$.*

The next result characterizes a unique optimal solution to the BPDN-I-like problem (3.7).

Corollary 4.14. *Let $g(x) = \|x\|_1$, $f : \mathbb{R}^m \to \mathbb{R}$ be a C^1 strictly convex function, and $x^* \in \mathbb{R}^N$ be a feasible point of the problem (3.7).*

C.1 *Suppose $f(Ax^* - y) < \varepsilon$. Then x^* is the unique minimizer of (3.7) if and only if $C_{\alpha S}$ has full column rank and there exists $u \in \mathbb{R}_{++}^{|\alpha|}$ such that $C_{\alpha S}^T u = \widehat{b}$ and $\|C_{\alpha S^c}^T u\|_\infty < 1$.*

C.2 *Suppose $f(Ax^* - y) = \varepsilon$. Then x^* is the unique minimizer of (3.7) if and only if conditions (i)-(ii) of Corollary 4.11 and the following condition hold:*

(2.iii) *if $\mathcal{K} := \{v \in \mathbb{R}^N \mid \left(\nabla f(Ax^* - y)\right)^T Av < 0, \ C_{\alpha\bullet}v \geq 0\}$ is nonempty, then there exist a positive real number θ and $\widetilde{u} \in \mathbb{R}_+^{|\alpha|}$ such that $\theta \cdot A_{\bullet S}^T \nabla f(Ax^* - y) - C_{\alpha S}^T \widetilde{u} + \widehat{b} = 0$, and $\|\theta \cdot A_{\bullet S^c}^T \nabla f(Ax^* - y) - C_{\alpha S^c}^T \widetilde{u}\|_\infty \leq 1$.*

The last result of this subsection pertains to the BPDN-II-like problem defined below, which is a special case of the problem (3.10) with $r = 1$, $g(x) = \|x\|_1$, and a positive real number $\eta > 0$:

$$\min_{x\in\mathbb{R}^N} \ f(Ax - y) \qquad \text{subject to} \quad g(x) \leq \eta, \ \text{and} \ \ Cx \geq d. \tag{4.7}$$

Corollary 4.15. *Let $g(x) = \|x\|_1$, $f : \mathbb{R}^m \to \mathbb{R}$ be a C^1 strictly convex function, and $x^* \in \mathbb{R}^N$ be a feasible point of the problem (4.7).*

C.1 *Suppose $g(x^*) < \eta$. Then x^* is the unique minimizer of (4.7) if and only if the associated conditions given in C.1 of Proposition 4.10 hold.*

C.2 *Suppose $g(x^*) = \eta$. Then x^* is the unique minimizer of (4.7) if and only if conditions (i)–(ii) of Corollary 4.11 and the following condition hold:*

(2.iii) *there exist* $\widetilde{u} \in \mathbb{R}_+^{|\alpha|}$ *and* $\mu \in \mathbb{R}_+$ *such that* $A_{\bullet\mathcal{S}}^T \nabla f(Ax^* - y) - C_{\alpha\mathcal{S}}^T \widetilde{u} + \mu \cdot \widehat{b} = 0$, *and* $\|A_{\bullet\mathcal{S}^c}^T \nabla f(Ax^* - y) - C_{\alpha\mathcal{S}}^T \widetilde{u}\|_\infty \leq \mu$.

5. Applications to ℓ_1-norm recovery problems: examples and comparison with related results

In this section, we apply the results developed in the previous section to specific ℓ_1-norm recovery problems. We show that the general framework established in this paper not only recovers all the known results without imposing restrictive assumptions but also leads to many new results, *e.g.*, the sparse fused LASSO subject to polyhedral constraints (*cf.* Cor. 5.4), basis pursuit subject to the monotone cone constraint and the Dantzig selector (*cf.* Sect. 5.4). Besides, we compare our results with the existing work and demonstrate the broad applicability and efficiency of the general results of this paper.

By setting $C = 0$ and $d = 0$ in Propositions 4.7–4.10 and Corollaries 4.11–4.15, we obtain solution uniqueness conditions for ℓ_1-norm optimization problems without a linear inequality constraint for either $g(x) = \|Ex\|_1$ or $g(x) = \|x\|_1$. These results give rise to the same uniqueness conditions recently developed for $g(x) = \|Ex\|_1$ in [29] and $g(x) = \|x\|_1$ in [28] respectively. Also see Section 3.1.1 for a detailed comparison with the results on basis-pursuit-like problems in [10].

5.1. Applications to basis pursuit denoising I and comparison with related results

Let $f : \mathbb{R}^m \to \mathbb{R}$ be a C^1 strictly convex function, and $g(x) = \|Ex\|_1$ or $g(x) = \|x\|_1$. For the BPDN-I-like problem (3.7) without a linear inequality constraint $Cx \geq d$, it is shown in the two papers [28, 29] that $Ax - y$ is constant on the solution set $\mathcal{X}$ and that $f(Ax - y) = \varepsilon$ for all $x \in \mathcal{X}$ if $0 \notin \mathcal{X}$. By a similar argument for ([28], Lem. 4.2(3)), one can show that if a linear inequality constraint is involved but $0 \in \mathcal{P} := \{x \,|\, Cx \geq d\}$ (or equivalently $d \leq 0$), then the same results hold; particularly, $Ax - y$ is also constant on the solution set. This case is especially interesting since $\mathcal{P}$ is often a polyhedral cone in applications. Therefore, the case C.1 in Proposition 4.9 and Corollary 4.14 can be ignored in these scenarios. Nonetheless, when a general linear inequality constraint is considered, the case C.1 is needed, since $Ax - y$ is *not* always constant on the solution set as demonstrated by the following example.

Example 5.1. Consider the following problem on $\mathbb{R}^2$:

$$\min_{x=(x_1, x_2)\in\mathbb{R}^2} \|x\|_1 \quad \text{subject to} \quad \frac{x_1^2}{9} + \frac{x_2^2}{16} \leq 1 \ \text{ and } \ x_1 + x_2 \geq 2.$$

This problem is a special case of the BPDN-I-like problem in (3.7), where $g(x) = \|x\|_1$, $\varepsilon = 1$, $f(Ax - y) = \|Ax\|_2^2$ with $A = \text{diag}(1/3, 1/4)$ and $y = 0$, $C = [1\ 1]$, and $d = 2$. It is easy to show *via* a geometric argument that the solution set is $\{x = (x_1, x_2) \,|\, x_1 + x_2 = 2 \ \text{and} \ x \geq 0\}$ on which $Ax - y$ is varying.

5.2. Applications to basis pursuit denoising II and comparison with related results

We discuss the results on the BPDN-II-like problem subject to one ℓ_1-norm based constraint given in (4.5) or (4.7) and compare them with the previous results. The paper [28] studies this problem of the following form without the linear inequality constraint $Cx \geq d$:

$$\min_{x\in\mathbb{R}^N} \ f(Ax - y) \qquad \text{subject to} \qquad \|x\|_1 \leq \tau, \tag{5.1}$$

where f is a strictly convex function, and the constant τ is assumed to satisfy $0 < \tau \leq \inf\{\|x\|_1 \,|\, x \in \text{argmin}_{z\in\mathbb{R}^N} f(Az - y)\}$ ([28], Ass. 2.3). It is claimed in ([28], Lem. 4.2(4)) that under this assumption on τ, $\|x\|_1 = \tau$ on the (nonempty) solution set of (5.1), which is a key step to derive the solution uniqueness

conditions in [28]. However, the proof for this claim given in ([28], Lem. 4.2(4)) is invalid. In what follows, we provide a remedy proof in a general setting.

Lemma 5.2. *Let $h : \mathbb{R}^N \to \mathbb{R}$ be a convex function and $g : \mathbb{R}^N \to \mathbb{R}$ be a continuous function such that $\min_{x \in \mathbb{R}^N} h(x)$ has a nonempty solution set $\mathcal{H}$, and $\gamma := \inf\{g(x) \mid x \in \mathcal{H}\} > -\infty$. For a given $\tau \in \mathbb{R}$ with $\tau \leq \gamma$, suppose the following optimization problem attains a nonempty solution set $\mathcal{H}_P$:*

$$(P) : \quad \min_{x \in \mathbb{R}^N} h(x) \quad subject \ to \quad g(x) \leq \tau.$$

Then $g(x) = \tau$ for any $x \in \mathcal{H}_P$. In particular, if $h(x) = f(Ax - y)$ and $g(x) = \|x\|_1$, where f is strictly convex, then $\|x\|_1 = \tau$ on $\mathcal{H}_P$.

Proof. We prove this lemma by contradiction. Suppose there exists an optimal solution $\widehat{x} \in \mathcal{H}_P$ such that $g(\widehat{x}) < \tau$. Then $g(\widehat{x}) < \gamma := \inf\{g(x) \mid x \in \mathcal{H}\}$, which implies that $\widehat{x} \notin \mathcal{H}$. For a fixed $x' \in \mathcal{H}$, we thus have $h(\widehat{x}) > h(x')$. It follows from the continuity of g at $\widehat{x}$ and $g(\widehat{x}) < \tau$ that there exists a sufficiently small $\bar{\lambda} \in (0, 1)$ such that $g\big((1 - \bar{\lambda})\widehat{x} + \bar{\lambda}x'\big) = g\big(\widehat{x} + \bar{\lambda}(x' - \widehat{x})\big) < \tau$. Hence, $z := (1 - \bar{\lambda})\widehat{x} + \bar{\lambda}x'$ is feasible to the optimization problem (P). In view of $h(\widehat{x}) > h(x')$, $\bar{\lambda} > 0$, and the convexity of h, we further have

$$h(z) = h\big((1 - \bar{\lambda})\widehat{x} + \bar{\lambda}x'\big) \leq (1 - \bar{\lambda})h(\widehat{x}) + \bar{\lambda}h(x') < h(\widehat{x}).$$

This shows that $\widehat{x}$ is not an optimal solution to (P), contradiction. Therefore, $g(x) = \tau, \forall\, x \in \mathcal{H}_P$. □

Compared with the results for the problem (5.1) developed in [28], the present paper establishes the solution uniqueness conditions not only without imposing a restriction on the parameter τ but also taking a general linear inequality constraint as well as multiple convex PA functions based constraints into account; see Theorem 3.11, Proposition 4.10, and Corollary 4.15. This generalization is especially important because, as shown in the following example, the claim that $\|x\|_1$ is constant on the solution set fails when a linear inequality constraint is imposed.

Example 5.3. Consider the problem in $\mathbb{R}^2$:

$$\min_{x=(x_1,x_2) \in \mathbb{R}^2} \quad (x_1 + x_2 - 2)^2 \qquad subject \ to \quad \|x\|_1 \leq 1 \ \ and \ \ -x_1 - x_2 \geq 0.$$

This problem is a special case of (4.7), where $f(\cdot) = |\cdot|^2$, $A = [1\ \ 1]$, $y = 2$, $\eta = 1$, $C = [-1\ \ -1]$, and $d = 0$. It is noticed that the solution set of $\min_{x \in \mathbb{R}^2}\ (x_1 + x_2 - 2)^2$ is the line in $\mathbb{R}^2$ defined by $x_1 + x_2 = 2$, and $\inf\{\|x\|_1 \mid x_1 + x_2 = 2\} = 2$. Therefore, the bound $\eta = 1$ satisfies the specified assumption. However, by a simple geometric argument, we deduce that the solution set is $\{x = (x_1, x_2) \mid x_1 + x_2 = 0,\ -1/2 \leq x_1 \leq 1/2\}$. Clearly, $\|x\|_1$ is not constant on this set.

5.3. Applications to multiple ℓ_1-norm based convex PA functions

Multiple ℓ_1-norm based functions are involved in several sparse optimization problems arising from statistics, image processing, and machine learning. One notable example is the so-called sparse fused LASSO [14, 26] which takes the following form with the positive penalty parameters λ_1 and λ_2:

$$\min_{x \in \mathbb{R}^N} \|Ax - y\|_2^2 + \lambda_1 \cdot \|x\|_1 + \lambda_2 \cdot \|D_1 x\|_1, \tag{5.2}$$

where $A \in \mathbb{R}^{m \times N}$ and $y \in \mathbb{R}^m$ are given, and D_1 is the first-order difference matrix given by

$$D_1 := \begin{bmatrix} -1 & 1 & & & \\ & -1 & 1 & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \end{bmatrix} \in \mathbb{R}^{(N-1) \times N}. \tag{5.3}$$

Here $\|D_1 x\|_1$ characterizes the total variation of x. Another version of the sparse fused LASSO is

$$\min_{x \in \mathbb{R}^N} \|Ax - y\|_2^2 \quad \text{subject to} \quad \|x\|_1 \leq \eta_1, \quad \text{and} \quad \|D_1 x\|_1 \leq \eta_2, \tag{5.4}$$

where $\eta_1, \eta_2 > 0$. These two problems are closely related to the generalized LASSO [24, 25].

It is observed *via* the discussions in Section 4.3.1 that the sparse fused LASSO in (5.2) can be formulated as: $\min_{x \in \mathbb{R}^N} f(Ax - y) + g(x)$, where $f(\cdot) = \|\cdot\|_2^2$, and $g(x) = \|Ex\|_1$ with $E = \begin{bmatrix} \lambda_1 \cdot I_N \\ \lambda_2 \cdot D_1 \end{bmatrix} \in \mathbb{R}^{(2N-1) \times N}$. Therefore, its solution uniqueness is determined by Theorem 3.8 or Proposition 4.8. Furthermore, we observe that the problem (5.4) can be treated as the BPDN-II-like problem (3.10) subject to two ℓ_1-norm based constraints, and its solution uniqueness conditions follow from Theorem 3.11. These observations allow us to easily incorporate linear inequality constraints into the two sparse fused LASSO models. For illustration, we show the solution uniqueness conditions below for the sparse fused LASSO in (5.2) subject to the nonnegative constraint, *i.e.,* $x \in \mathbb{R}_+^N = \{x \mid Cx \geq d\}$ with $C = I_N$ and $d = 0$.

Corollary 5.4. *Consider the sparse fused LASSO in (5.2) subject to the nonnegative constraint. For a feasible point x^* and the matrix E defined above, let $\mathcal{S}$ be the support of Ex^*, and $b := E_{\mathcal{S}\bullet}^T \widehat{b}$ with $\widehat{b} := sgn((Ex^*)_{\mathcal{S}})$. Then x^* is the unique minimizer if and only if the following conditions hold:*

(i) $\begin{bmatrix} A_{\bullet \mathcal{S}_I} \\ (D_1)_{\mathcal{S}_D^c \mathcal{S}_I} \end{bmatrix}$ *has full column rank, where $\mathcal{S}_I$ is the support of x^*, and $\mathcal{S}_D$ is the support of $D_1 x^*$;*

(ii) there exist $u \in \mathbb{R}^m$, $u' \in \mathbb{R}_{++}^{|\mathcal{S}_I^c|}$, and $u'' \in \mathbb{R}^{|\mathcal{S}^c|}$ with $\|u''\|_\infty < 1$ such that $\begin{pmatrix} A_{\bullet \mathcal{S}_I}^T u \\ A_{\bullet \mathcal{S}_I^c}^T u + u' \end{pmatrix} - E_{\mathcal{S}^c \bullet}^T u'' = b$;

(iii) there exist $\widetilde{u} \in \mathbb{R}_+^{|\mathcal{S}_I^c|}$ and $\widetilde{u}' \in \mathbb{R}^{|\mathcal{S}^c|}$ with $\|\widetilde{u}'\|_\infty \leq 1$ such that $2A^T(Ax^* - y) - I_{\mathcal{S}_I^c \bullet}^T \widetilde{u} + b + E_{\mathcal{S}^c \bullet}^T \widetilde{u}' = 0$.

Proof. Based on the definitions of E, $\mathcal{S}$, $\mathcal{S}_I$, and $\mathcal{S}_D$ as well as $C = I_N$, we have $E_{\mathcal{S}^c \bullet} = \begin{bmatrix} \lambda_1 I_{\mathcal{S}_I^c \bullet} \\ \lambda_2 \cdot (D_1)_{\mathcal{S}_D^c \bullet} \end{bmatrix}$,

$\alpha = \mathcal{S}_I^c$, and $C_{\alpha \bullet} = I_{\mathcal{S}_I^c \bullet}$. Hence, we have $\begin{bmatrix} A \\ C_{\alpha \bullet} \\ E_{\mathcal{S}^c \bullet} \end{bmatrix} = \begin{bmatrix} A_{\bullet \mathcal{S}_I} & A_{\bullet \mathcal{S}_I^c} \\ 0 & I_{\mathcal{S}_I^c \mathcal{S}_I^c} \\ 0 & \lambda_1 I_{\mathcal{S}_I^c \mathcal{S}_I^c} \\ \lambda_2 (D_1)_{\mathcal{S}_D^c \mathcal{S}_I} & \lambda_2 (D_1)_{\mathcal{S}_D^c \mathcal{S}_I^c} \end{bmatrix}$. In view of this result,

Proposition 4.8, and condition (a) of Proposition 4.7, it can be verified *via* a straightforward calculation that the corollary holds. $\square$

5.4. Applications to ℓ_1-norm recovery subject to linear inequality constraints and comparison with related results

The general framework developed in Sections 3 and 4 provides considerable flexibility to incorporate various linear inequality constraints into recovery problems. We shows this advantage *via* several examples below.

An important linear inequality constraint that has received considerable attention in applications is the nonnegative constraint [7, 11, 27, 31], *i.e.,* $x \in \mathbb{R}_+^N$. As shown in Section 5.3, we have $C = I_N$ and $d = 0$ so that $\mathcal{P} = \{x \in \mathbb{R}^N \mid Cx \geq d\} = \mathbb{R}_+^N$. We first consider the basis pursuit with $g(x) = \|x\|_1$, and recover the solution uniqueness conditions established in ([31], Thm. 2.7).

Corollary 5.5. *Let $C = I_N$ and $d = 0$ so that $\mathcal{P} = \mathbb{R}_+^N$. Then a feasible point x^* is the unique minimizer of (3.3) with $\mathcal{S}$ being the support of x^* if and only if the following conditions hold:*

(i) *the columns of $A_{\bullet\mathcal{S}}$ are linearly independent and*
(ii) *there exists $u \in \mathbb{R}^m$ such that $A_{\bullet\mathcal{S}}^T u = \mathbf{1}$ and $A_{\bullet\mathcal{S}^c}^T u < \mathbf{1}$.*

Proof. Note that for a given $x^* \in \mathbb{R}_+^N$, we have $\widehat{b} = (\mathrm{sgn}(x_{\mathcal{S}}^*)) = \mathbf{1} \in \mathbb{R}^{|\mathcal{S}|}$, $\alpha = \mathcal{S}^c$, $C_{\alpha\mathcal{S}} = 0$ and $C_{\alpha\mathcal{S}^c} = I_{\mathcal{S}^c\mathcal{S}^c}$. It thus follows from Corollary 4.11 that x^* is the unique minimizer if and only if (i) $A_{\bullet\mathcal{S}}$ has full column rank, and (ii) there exist $u \in \mathbb{R}^m$ and $u' \in \mathbb{R}_{++}^{|\mathcal{S}^c|}$ such that $A_{\bullet\mathcal{S}}^T u = \widehat{b} = \mathbf{1}$ and $\|A_{\bullet\mathcal{S}^c}^T u + u'\|_\infty < 1$. Hence, it suffices to show that $(A^T u)_{\mathcal{S}^c} < \mathbf{1}$ if and only if $\|A_{\bullet\mathcal{S}^c}^T u + u'\|_\infty < 1$ for some $u' > 0$. The "if" part is trivial. To show the "only if" part, suppose for each $i \in \mathcal{S}^c$, $v_i := (A^T u)_i < 1$. Hence, either $-1 < v_i < 1$ or $v_i \le -1$. For the former case, choose $u_i' > 0$ sufficiently small so that $|v_i + u_i'| < 1$. For the latter case, choose s satisfying $0 \le \frac{-(1+v_i)}{1-v_i} < s < 1$. Then $u_i' := s \cdot (1 - v_i) > 0$ is such that $|v_i + u_i'| < 1$. This thus completes the proof. $\square$

The solution uniqueness conditions for the LASSO-like and the BPDN-I/II-like problems under the nonnegative constraint are given below. Their proofs follow directly from Corollaries 4.13–4.15 and the fact that $h \ge -\mathbf{1}$ if and only if there exists $u \ge 0$ such that $\|h - u\|_\infty \le 1$.

Corollary 5.6. *Let $g(x) = \|x\|_1$, and $f : \mathbb{R}^m \to \mathbb{R}$ be a C^1 strictly convex function, and x^* be a feasible point of the problem (3.6). Then x^* is the unique minimizer of (3.6) if and only if conditions (i) and (ii) of Corollary 5.5 and the following condition hold: $A_{\bullet\mathcal{S}}^T \nabla f(Ax^* - y) = -\mathbf{1}$, and $A_{\bullet\mathcal{S}^c}^T \nabla f(Ax^* - y) \ge -\mathbf{1}$.*

The next result characterizes a nonzero unique optimal solution to the BPDN-I-like problem (3.7). The case $f(Ax^* - y) < \varepsilon$ is ignored due to the fact that $0 \in \mathbb{R}_+^N$ and the discussions in Section 5.1.

Corollary 5.7. *Let $g(x) = \|x\|_1$, $f : \mathbb{R}^m \to \mathbb{R}$ be a C^1 strictly convex function, and x^* be a nonzero feasible point of the problem (3.7). Then x^* is the unique minimizer of (3.7) if and only if conditions (i) and (ii) of Corollary 5.5 and the following condition hold: if $\mathcal{K} := \{v \in \mathbb{R}^N \,|\, \left(\nabla f(Ax^* - y)\right)^T Av < 0, \; v_{\mathcal{S}^c} \ge 0\}$ is nonempty, then there exists a positive real number θ such that $\theta \cdot A_{\bullet\mathcal{S}}^T \nabla f(Ax^* - y) = -\mathbf{1}$, and $\theta \cdot A_{\bullet\mathcal{S}^c}^T \nabla f(Ax^* - y) \ge -\mathbf{1}$.*

We then consider the BPDN-II-like problem (4.7) subject to the nonnegative constraint.

Corollary 5.8. *Let $g(x) = \|x\|_1$, $f : \mathbb{R}^m \to \mathbb{R}$ be a C^1 strictly convex function, and x^* be a feasible point of the problem (4.7).*

C.1 *Suppose $g(x^*) < \eta$. Then x^* is the unique minimizer of (4.7) if and only if the associated conditions given in C.1 of Proposition 4.10 hold.*
C.2 *Suppose $g(x^*) = \eta$. Then x^* is the unique minimizer of (4.7) if and only if conditions (i) and (ii) of Corollary 5.5 and the following condition hold: there exists $\mu \in \mathbb{R}_+$ such that $A_{\bullet\mathcal{S}}^T \nabla f(Ax^* - y) = -\mu \cdot \mathbf{1}$, and $A_{\bullet\mathcal{S}^c}^T \nabla f(Ax^* - y) \ge -\mu \cdot \mathbf{1}$.*

We next consider the linear inequality constraint with $C = D_1$ and $d = 0$, where D_1 is the first-order difference matrix defined in (5.3). In other words, the variable x is subject to the monotone increasing constraint which appears in such applications as order statistics. For the purpose of illustration, we consider the BP-like problem (3.3) with $g(\cdot) = \|\cdot\|_1$ for a feasible x^*. Since the elements of x^* are monotonically increasing, we can write it as

$$x^* = \Big(\underbrace{x_1^*, \ldots, x_{k^-}^*}_{<0}, 0, \ldots, 0, \underbrace{x_{k^+}^*, \ldots, x_N^*}_{>0} \Big)^T \in \mathbb{R}^N,$$

where $x_{k^-}^*$ is the last negative element and $x_{k^+}^*$ is the first positive element, both from the left. Define the index sets $\mathcal{S}_- := \{1, \ldots, k^-\}$, and $\mathcal{S}_+ := \{k^+, \ldots, N\}$. Then the support of x^* is $\mathcal{S} = \mathcal{S}_- \cup \mathcal{S}_+$, and $\widehat{b}_{\mathcal{S}_-} = -\mathbf{1}$ and $\widehat{b}_{\mathcal{S}_+} = \mathbf{1}$. Further, the index set of active constraints is $\alpha = \alpha_- \cup \alpha_0 \cup \alpha_+$, where α_- and α_+ are the index sets

of active constraints associated with $(x_1^*, \ldots, x_{k^-}^*, 0)^T \in \mathbb{R}^{k^-+1}$ and $(0, x_{k^+}^*, \ldots, x_N^*)^T \in \mathbb{R}^{N-k^++2}$ respectively, and $\alpha_0 = \{k^- + 1, k^- + 2, \ldots, k^+ - 2\}$. Note that $\mathcal{S}^c = \{k^- + 1, k^- + 2, \ldots, k^+ - 1\}$, $\alpha_- \subset \mathcal{S}_-$, $\alpha_+ \subset \mathcal{S}_+$, and $(D_1)_{\alpha_0 \mathcal{S}_-} = 0$, $(D_1)_{\alpha_+ \mathcal{S}_-} = 0$, $(D_1)_{\alpha_0 \mathcal{S}_+} = 0$, $(D_1)_{\alpha_- \mathcal{S}_+} = 0$. Furthermore, defining $\alpha_- + 1 := \{i + 1 \,|\, i \in \alpha_-\}$ and $\alpha_+ + 1 := \{j + 1 \,|\, j \in \alpha_+\}$, we let $\overline{\alpha}_- := \alpha_- \cup (\alpha_- + 1) \subset \mathcal{S}_-$, and $\overline{\alpha}_+ := \alpha_+ \cup (\alpha_+ + 1) \subset \mathcal{S}_+$. Since the null spaces of $(D_1)_{\alpha_- \overline{\alpha}_-}$ and $(D_1)_{\alpha_+ \overline{\alpha}_+}$ are spanned by $\mathbf{1}$ respectively, and $(D_1)_{\alpha_- (\mathcal{S}_- \setminus \overline{\alpha}_-)} = 0$, $(D_1)_{\alpha_+ (\mathcal{S}_+ \setminus \overline{\alpha}_+)} = 0$, we have

$$
\begin{bmatrix} A_{\bullet \mathcal{S}} \\ (D_1)_{\alpha \mathcal{S}} \end{bmatrix} = \begin{bmatrix} A_{\bullet \mathcal{S}_-} & A_{\bullet \mathcal{S}_+} \\ (D_1)_{\alpha_- \mathcal{S}_-} & 0 \\ 0 & 0 \\ 0 & (D_1)_{\alpha_+ \mathcal{S}_+} \end{bmatrix} = \begin{bmatrix} A_{\bullet \overline{\alpha}_-} & A_{\bullet(\mathcal{S}_- \setminus \overline{\alpha}_-)} & A_{\bullet \overline{\alpha}_+} & A_{\bullet(\mathcal{S}_+ \setminus \overline{\alpha}_+)} \\ (D_1)_{\alpha_- \overline{\alpha}} & (D_1)_{\alpha_-(\mathcal{S}_- \setminus \overline{\alpha}_-)} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & (D_1)_{\alpha_+ \overline{\alpha}_+} & (D_1)_{\alpha_+(\mathcal{S}_+ \setminus \overline{\alpha}_+)} \end{bmatrix}.
$$

Therefore, condition (i) of Corollary 4.11 holds, *i.e.*, $\begin{bmatrix} A_{\bullet \mathcal{S}} \\ (D_1)_{\alpha \mathcal{S}} \end{bmatrix}$ has full column rank, if and only if

$$
\begin{bmatrix} A_{\bullet(\mathcal{S}_- \setminus \overline{\alpha}_-)} & A_{\bullet(\mathcal{S}_+ \setminus \overline{\alpha}_+)} & (A_{\bullet \overline{\alpha}_-} \mathbf{1} + A_{\bullet \overline{\alpha}_+} \mathbf{1}) \end{bmatrix}
$$

has full column rank. Moreover, in light of the above development, we see that condition (ii) of Corollary 4.11 is equivalent to the existence of $u \in \mathbb{R}^m$ and $(u'_-, u'_0, u'_+) > 0$ such that $A_{\bullet \mathcal{S}_-}^T u + [(D_1)_{\alpha_- \mathcal{S}_-}]^T u'_- = -\mathbf{1}$, $A_{\bullet \mathcal{S}_+}^T u + [(D_1)_{\alpha_+ \mathcal{S}_+}]^T u'_+ = \mathbf{1}$, and $\|A_{\bullet \mathcal{S}^c}^T u + [(D_1)_{\alpha_0 \mathcal{S}^c}]^T u'_0\|_\infty < 1$, where the last inequality follows from the fact that $(D_1)_{\alpha_- \mathcal{S}^c} = 0$ and $(D_1)_{\alpha_+ \mathcal{S}^c} = 0$.

Lastly, we consider the Dantzig selector [5] (*cf.* Sect. 3.5). As shown in Section 3.5, this problem can be formulated as the basis pursuit subject to a polyhedral constraint, *i.e.*, $\min_{x \in \mathbb{R}^N} \|x\|_1$ subject to $-\varepsilon \cdot \mathbf{1} \le A^T A x - A^T y \le \varepsilon \cdot \mathbf{1}$, where $\varepsilon > 0$ is given. For a feasible vector x^*, let $\mathcal{S}$ be its support, and α_+ and α_- be the active index sets of the constraints $A^T A x \ge A^T y - \varepsilon \cdot \mathbf{1}$ and $-A^T A x \ge -(A^T y + \varepsilon \cdot \mathbf{1})$ at x^*, respectively. Clearly, $\alpha_+ \cap \alpha_- = \emptyset$. It thus follows from Corollary 4.11 that x^* is the unique optimizer of the Dantzig selector if and only if (i) $\begin{bmatrix} A_{\bullet \alpha_+}^T \\ -A_{\bullet \alpha_-}^T \end{bmatrix} A_{\bullet \mathcal{S}}$ has full column rank, and (ii) there exists $u' \in \mathbb{R}_{++}^{|\alpha_+| + |\alpha_-|}$ such that $A_{\bullet \mathcal{S}}^T \begin{bmatrix} A_{\bullet \alpha_+} & -A_{\bullet \alpha_-} \end{bmatrix} u' = \mathrm{sgn}(x_{\mathcal{S}}^*)$ and $\|A_{\bullet \mathcal{S}^c}^T \begin{bmatrix} A_{\bullet \alpha_+} & -A_{\bullet \alpha_-} \end{bmatrix} u'\|_\infty < 1$. Note that condition (i) holds if and only if $A_{\bullet \mathcal{S}}$ has full column rank and $N(\begin{bmatrix} A_{\bullet \alpha_+}^T \\ A_{\bullet \alpha_-}^T \end{bmatrix}) \cap R(A_{\bullet \mathcal{S}}) = \{0\}$.

6. Numerical verification of the solution uniqueness conditions

In this section, we discuss numerical verification of the solution uniqueness conditions developed in the previous sections. It is observed that each set of solution uniqueness criteria involving a convex PA function and a C^1 strictly convex loss function established in Sections 3 and 4 consists of the following conditions: (a) the full column rank condition for a matrix; (b) the consistency of a linear inequality system with at least one strict inequality; and/or (c) the consistency of another linear inequality system with non-strict inequalities. The first two conditions characterize solution uniqueness, while the last condition pertains to solution optimality. Numerically, the first condition can be determined *via* standard linear algebraic tools, and the last condition can be checked *via* the feasibility test of a suitable linear program. To effectively verify the conditions involving strict inequalities, we show in the following lemma that the verification of such conditions can be formulated as a linear program.

Lemma 6.1. *Let $\widehat{z} \in \mathbb{R}^n$, $F \in \mathbb{R}^{n \times m}$, $G \in \mathbb{R}^{n \times r}$, and $H \in \mathbb{R}^{n \times s}$ be given. Then the linear inequality system $(I): \widehat{z} + Fw + Gw' + Hw'' = 0$, $w \in \mathbb{R}^m$, $w' \in \mathbb{R}_+^r$, $w'' \in \mathbb{R}_{++}^s$ has a solution if and only if the following linear program is solvable and attains a positive optimal value:*

$$
\max_{(w, w', w'', \varepsilon) \in \mathbb{R}^m \times \mathbb{R}^r \times \mathbb{R}^s \times \mathbb{R}} \varepsilon \quad \text{subject to} \quad \widehat{z} + Fw + Gw' + Hw'' = 0, \quad w' \ge 0, \quad w'' \ge \varepsilon \cdot \mathbf{1}, \quad \varepsilon \le 1. \tag{6.1}
$$

Proof. To show the "if" part, let $(w_*, w'_*, w''_*, \varepsilon_*) \in \mathbb{R}^m \times \mathbb{R}^r_+ \times \mathbb{R}^s \times \mathbb{R}$ be an optimal solution to the linear program (6.1) with $\varepsilon_* > 0$. Hence, we have $w''_* \geq \varepsilon_* \cdot \mathbf{1} > 0$. This shows that the system (I) has a solution. For the "only if" part, suppose there exists a triple $(\widetilde{w}, \widetilde{w}', \widetilde{w}'') \in \mathbb{R}^m \times \mathbb{R}^r_+ \times \mathbb{R}^s_{++}$ such that $\widehat{z} + F\widetilde{w} + G\widetilde{w}' + H\widetilde{w}'' = 0$. Since $\widetilde{w}'' > 0$, there exists a real number $\theta \in (0, 1]$ such that $\widetilde{w}'' \geq \theta \cdot \mathbf{1}$. Therefore, the linear program (6.1) is feasible. Furthermore, since its objective function is bounded above by 1 (and bounded below by θ) on the feasible set, the linear program (6.1) attains an optimal solution and its optimal value $\varepsilon_* \geq \theta > 0$. This yields the desired result. $\qquad\square$

In what follows, we apply Lemma 6.1 to check various uniqueness conditions involving strict inequalities. Particularly, we show how to formulate these conditions in the requested form in Lemma 6.1.

1) Condition (ii) of Theorem 3.2, *i.e.*, there exist $z \in \mathbb{R}^m$, $z' \in \mathbb{R}^{|\alpha|}_{++}$, and $z'' \in \mathbb{R}^{|\mathcal{I}|}_{++}$ such that $A^T z - C^T_{\alpha \bullet} z' + W^T z'' = 0$. Letting $\widehat{z} = 0$, $F = A^T$, $G = 0$, and $H = \begin{bmatrix} -C^T_{\alpha \bullet} & W^T \end{bmatrix}$, Lemma 6.1 can be applied.

2) Conditions for (C.1) of Theorem 3.9, which is equivalent to the existence of $z \in \mathbb{R}^{|\alpha|}_{++}$ and $z' \in \mathbb{R}^{|\mathcal{I}|}_{++}$ such that $C^T_{\alpha \bullet} z = W^T z'$. Letting $\widehat{z} = 0$, $F = 0$, $G = 0$, and $H = \begin{bmatrix} -C^T_{\alpha \bullet} & W^T \end{bmatrix}$, we use Lemma 6.1.

3) Condition (b) of Proposition 4.7, *i.e.*, there exist $u \in \mathbb{R}^m$, $u' \in \mathbb{R}^{|\alpha|}_{++}$, and $u'' \in \mathbb{R}^{|\mathcal{S}^c|}$ with $\|u''\|_\infty < 1$ such that $A^T u + C^T_{\alpha \bullet} u' - E^T_{\mathcal{S}^c \bullet} u'' = b$. This condition is equivalent to the consistency of the following linear inequality system in (u, u', u'', v, w):

$$A^T u + C^T_{\alpha \bullet} u' - E^T_{\mathcal{S}^c \bullet} u'' = b, \quad u'' + v = \mathbf{1}, \quad u'' - w = -\mathbf{1}, \quad (u', v, w) > 0.$$

Suitable $\widehat{z}, F, G$, and H can be easily found from the above system in order to make use of Lemma 6.1.

4) Condition (ii) of Corollary 4.11, *i.e.*, there exist $u \in \mathbb{R}^m$ and $u' \in \mathbb{R}^{|\alpha|}_{++}$ such that $A^T_{\bullet \mathcal{S}} u + C^T_{\alpha \mathcal{S}} u' = \widehat{b}$ and $\|A^T_{\bullet \mathcal{S}^c} u + C^T_{\alpha \mathcal{S}^c} u'\|_\infty < 1$. It is equivalent to the consistency of the linear inequality system in (u, u', v, w): $A^T_{\bullet \mathcal{S}} u + C^T_{\alpha \mathcal{S}} u' = \widehat{b}$, $A^T_{\bullet \mathcal{S}^c} u + C^T_{\alpha \mathcal{S}^c} u' + v = \mathbf{1}$, $A^T_{\bullet \mathcal{S}^c} u + C^T_{\alpha \mathcal{S}^c} u' - w = -\mathbf{1}$, and $(u', v, w) > 0$.

5) Condition (ii) of Corollary 5.5, *i.e.*, there exists $u \in \mathbb{R}^m$ such that $(A^T u)_{\mathcal{S}} = \mathbf{1}$ and $(A^T u)_{\mathcal{S}^c} < \mathbf{1}$. This condition is equivalent to the consistency of the linear inequality system in (u, v): $A^T_{\bullet \mathcal{S}} u = \mathbf{1}$, $A^T_{\bullet \mathcal{S}^c} u + v = \mathbf{1}$, and $v > 0$. This paves the way to exploit Lemma 6.1.

Remark 6.2. We briefly discuss the complexity of the linear program based verification for ℓ_1 minimization. To facilitate the discussion, we focus on the case without linear inequality constraints, *i.e.*, $C = 0$ and $d = 0$. In the most general scenario given in case (1) above, the complexity of the associated linear program depends on the size of $\mathcal{I}$ defined in (3.1), which corresponds to the active index set of the max-formulation of a convex PA function g. Indeed, the number of variables for case 1) is $m + |\mathcal{I}|$. Note that in ℓ_1 minimization problems such as basis pursuit, we usually have $m \ll N$ and $|\mathcal{S}| \ll N$ for a sparse vector x^*, where $\mathcal{S}$ is the support of x^*. Hence, $|\mathcal{S}^c| \approx N$ and $|\mathcal{I}| = 2^{|\mathcal{S}^c|} \approx 2^N$, leading to a high computational cost. Instead, using the specialized conditions given in cases 3)–5) and Lemma 6.1, we obtain linear programs whose numbers of variables are small multiples of N. This yields the complexity of $O(N^3 / \ln N)$ based on state-of-art linear programming techniques [1]. In comparison with the verification schemes developed in [28, 29] for basis pursuit without polyhedral constraints, we provide a simple, systematic, and yet effective verification scheme applicable to broader uniqueness conditions.

7. Conclusions

This paper studies the solution uniqueness of a class of convex optimization problems involving convex PA functions and subject to general polyhedral constraints. By exploiting the max-formulation of convex PA functions and convex analysis techniques, we develop simpler and unifying approaches to derive dual variable based solution uniqueness conditions for an individual vector. These results are applied to a variety of ℓ_1 minimization problems subject to possible polyhedral constraints. An effective scheme is also proposed to verify the obtained uniqueness conditions. The uniqueness conditions developed in the current paper assume the exact knowledge of a minimizer. As a future research topic, we will consider various robustness issues and their implications when a minimizer is unknown.

References

[1] K.M. Anstreicher, Linear programming in $O(\frac{n^3}{\ln n}L)$ operations. *SIAM J. Optim.* **9** (1999) 803–812.

[2] D.P. Bertsekas, Nonlinear Programming. 2nd edn. Athena Scientific, Belmont, MA (1999).

[3] D.P. Bertsekas, Convex Optimization Theory. Athena Scientific, Belmont, MA (2009).

[4] E. J. Candes and Y. Plan, A probabilistic and RIPless theory of compressed sensing. *IEEE Trans. Inf. Theory* **57** (2011) 7235–7254.

[5] E. Candes and T. Tao, The Dantzig selector: statistical estimation when p is much larger than n. *Ann. Stat.* **35** (2007) 2313–2351.

[6] F. Facchinei and J.-S. Pang, Finite-Dimensional Variational Inequalities and Complementarity Problems. Spring-Verlag, New York (2003).

[7] S. Foucart and D. Koslicki, Sparse recovery by means of nonnegative least squares. *IEEE Signal Process. Lett.* **21** (2014) 498–502.

[8] S. Foucart and H. Rauhut, A Mathematical Introduction to Compressive Sensing. Birkhäuser, Basel, Switzerland (2013).

[9] J.J. Fuchs, On sparse representations in arbitrary redundant bases. *IEEE Trans. Inf. Theory* **50** (2004) 1341–1344.

[10] J.C. Gilbert, On the solution uniqueness characterization in the L1 norm and polyhedral gauge recovery. *J. Optim. Theory Appl.* **172** (2017) 70–101.

[11] Y. Itoh, M.F. Duarte and M. Parente. Perfect recovery conditions for non-negative sparse modeling. *IEEE Trans. Signal Process.* **65** (2017) 69–80.

[12] S.J. Kim, K. Koh, S. Boyd and D. Gorinevsky, ℓ_1 trend filtering. *SIAM Rev.* **51** (2009) 339–360.

[13] O.L. Mangasarian, Uniqueness of solution in linear programming. *Linear Algebra Appl.* **25** (1979) 151–162.

[14] A. Rinaldo, Properties and refinements of the fused lasso. *Ann. Stat.* **37** (2009) 2922–2952.

[15] R.T. Rockafellar, Convex Analysis. Princeton University Press (1970).

[16] A.P. Ruszczynski, Nonlinear Optimization. Princeton University Press (2006).

[17] S. Scholtes, *Introduction to Piecewise Differentiable Equations*. Habilitation thesis, Institut für Statistik und Mathematische Wirtschaftstheorie, Universität Karlsruhe (1994).

[18] J. Shen, Robust non-Zenoness of piecewise affine systems with applications to linear complementarity systems. *SIAM J. Optim.* **24** (2014) 2023–2056.

[19] J. Shen and T.M. Lebair, Shape restricted smoothing splines via constrained optimal control and nonsmooth Newton's methods. *Automatica* **53** (2015) 216–224.

[20] J. Shen and S. Mousavi, Least sparsity of p-norm based optimization problems with $p > 1$. *SIAM J. Optim.* **28** (2018) 2721–2751.

[21] J. Shen and X. Wang, Estimation of monotone functions via P-splines: a constrained dynamical optimization approach. *SIAM J. Control Optim.* **49** (2011) 646–671.

[22] J. Shen, L. Han and J.-S. Pang, Switching and stability properties of conewise linear systems. *ESAIM: COCV* **16** (2010) 764–793.

[23] R.J. Tibshirani, The Lasso problem and uniqueness. *Electron. J. Stat.* **7** (2013) 1456–1490.

[24] R. J. Tibshirani, *The Solution Path of the Generalized LASSO*. Ph.D. thesis, Stanford University (2011).

[25] R. Tibshirani and J. Taylor, Degrees of freedom in Lasso problems. *Ann. Stat.* **40** (2012) 1198–1232.

[26] R. Tibshirani, M. Saunders, S. Rosset, J. Zhu and K. Knight, Sparsity and smoothness via the fused Lasso. *J. R. Stat. Soc.: B (Statistical Methodology)* **67** (2005) 91–108.

[27] M. Wang, W. Xu and A. Tang, A unique "nonnegative" solution to an underdetermined system: from vectors to matrices. *IEEE Trans. Signal Process.* **59** (2011) 1007–1016.

[28] H. Zhang, W. Yin and L. Cheng, Necessary and sufficient conditions of solution uniqueness in 1-norm minimization. *J. Optim. Theory Appl.* **164** (2015) 109–122.

[29] H. Zhang, M. Yan and W. Yin, One condition for solution uniqueness and robustness of both ℓ_1-synthesis and ℓ_1-analysis minimizations. *Adv. Comput. Math.* **42** (2016) 1381–1399.

[30] Y.-B. Zhao, RSP-based analysis for sparsest and least ℓ_1-norm solutions to underdetermined linear systems. *IEEE Trans. Signal Process.* **61** (2013) 5777–5788.

[31] Y.-B. Zhao, Equivalence and strong equivalence between the sparsest and least ℓ_1-norm nonnegative solutions of linear systems and their applications. *J. Oper. Res. Soc. China* **2** (2014) 171–193.

[32] Y.-B. Zhao, Sparse Optimization Theory and Methods. CRC Press, Taylor & Francis, NY (2018).

[33] Y.-B. Zhao and C. Xu, 1-bit compressive sensing: reformulation and RRSP-based sign recovery theory. *Sci. China Math.* **59** (2016) 2049–2074.