



## Mathematics of Operations Research

Publication details, including instructions for authors and subscription information:  
<http://pubsonline.informs.org>

### On the Minimization Over Sparse Symmetric Sets: Projections, Optimality Conditions, and Algorithms

Amir Beck, Nadav Hallak

To cite this article:

Amir Beck, Nadav Hallak (2015) On the Minimization Over Sparse Symmetric Sets: Projections, Optimality Conditions, and Algorithms. *Mathematics of Operations Research*

Published online in Articles in Advance 23 Sep 2015

<http://dx.doi.org/10.1287/moor.2015.0722>

Full terms and conditions of use: <http://pubsonline.informs.org/page/terms-and-conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact [permissions@informs.org](mailto:permissions@informs.org).

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2015, INFORMS

Please scroll down for article—it is on subsequent pages



INFORMS is the largest professional society in the world for professionals in the fields of operations research, management science, and analytics.

For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

# On the Minimization Over Sparse Symmetric Sets: Projections, Optimality Conditions, and Algorithms

Amir Beck, Nadav Hallak

Faculty of Industrial Engineering and Management, Technion–Israel Institute of Technology, Haifa 3200003, Israel  
 {becka@ie.technion.ac.il, ndvhllk@campus.technion.ac.il}

We consider the problem of minimizing a general continuously differentiable function over symmetric sets under sparsity constraints. These type of problems are generally hard to solve because the sparsity constraint induces a combinatorial constraint into the problem, rendering the feasible set to be nonconvex. We begin with a study of the properties of the orthogonal projection operator onto sparse symmetric sets. Based on this study, we derive efficient methods for computing sparse projections under various symmetry assumptions. We then introduce and study three types of optimality conditions: basic feasibility,  $L$ -stationarity, and coordinatewise optimality. A hierarchy between the optimality conditions is established by using the results derived on the orthogonal projection operator. Methods for generating points satisfying the various optimality conditions are presented, analyzed, and finally tested on specific applications.

*Keywords:* sparse optimization; optimality conditions; block-type methods

*MSC2000 subject classification:* Primary: 90C26; secondary: 90C46

*ORMS subject classification:* Primary: nonlinear programming; secondary: algorithms

*History:* Received June 16, 2014; revised December 7, 2014. Published online in *Articles in Advance*.

## 1. Introduction.

**1.1. Problem formulation.** Sparse optimization problems have been a major research topic across different disciplines in recent years, such as compressed sensing. Since the sparsity constraint induces a combinatorial constraint into the problem, it is generally hard to reach an optimal solution efficiently, even if the objective function is convex or if no other constraints are imposed. Residing on the border between continuous and discrete optimization, methods addressing the problem appear in the literature of both fields. Yet, the theory of optimality conditions in the sparse optimization literature is lacking. One of the objectives of this paper will be to rectify this situation. We discuss the following sparsity-constrained minimization problem:

$$(P) \quad \begin{aligned} \min \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{x} \in C_s \cap B, \end{aligned} \tag{1}$$

where the set  $C_s$  comprises all vectors with at most  $s$  nonzero elements:

$$C_s = \{\mathbf{x} \in \mathbb{R}^n: \|\mathbf{x}\|_0 \leq s\}, \tag{2}$$

where  $s \in \{1, 2, \dots, n\}$  and  $\|\cdot\|_0$  is the so-called  $l_0$  norm, which counts the number of nonzero elements in the vector:

$$\|\mathbf{x}\|_0 \equiv \#\{i: x_i \neq 0\}.$$

The following standing assumptions are made from now on in the paper.

[A]  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  is a lower-bounded continuously differentiable function.

[B]  $B$  is a closed and convex set.

In some cases, which will be explicitly noted, we will extend Assumption A to:

[A+]  $f \in C_{L(f)}^{1,1}$ , meaning that  $f$  has a Lipschitz-continuous gradient with constant  $L(f)$  (see §2.2).

When additional assumptions will be imposed, they will be explicitly stated.

**1.2. Literature review.** The sparse optimization literature is dominated by compressed sensing-oriented papers, which mostly focus on the problem of recovering a sparse signal  $\mathbf{x}$  with a sampling matrix  $\mathbf{A}$  and a measurements vector  $\mathbf{b}$ , see, for example, the comprehensive reviews Bruckstein et al. [9], Davenport et al. [14], Duarte and Eldar [16], and Tropp and Wright [30]. Exact recovery properties are known when the sampling matrix is assumed to have some properties such as the restricted isometry property (RIP) (Candes and Tao [11]) or conditions based on the mutual coherence (David and Elad [15]), see also Cartis and Thompson [12] for a different type of condition warranting exact recovery.

We roughly distinguish between two types of methods: those who relax the sparsity constraint and those who do not. Relaxation methods usually involve the  $l_1$  norm, such as the famous basis pursuit (BP) (Chen et al. [13]), the Dantzing selector (Candes and Tao [10]), or regularization techniques such as  $l_1$  regularization (Tropp and Wright [30], Beck and Teboulle [4]). Our main interest is in methods, which address the sparsity-constrained problem directly. These usually attempt to solve the minimization of the function  $\|\mathbf{Ax} - \mathbf{b}\|^2$  subject to the sparsity constraint. Two such methods are the IHT (Blumensath and Davies [5, 6], Blumensath [7]) and CoSaMP (Needell and Tropp [24]); see also Lu [19] and Lu and Zhang [20] for further extensions and analysis of IHT-type methods, as well as Attouch et al. [1] for an analysis of the convergence under semialgebraic assumptions on the feasible set and the objective function. For a general objective function, the Gradient Support Pursuit (GraSP) method was introduced and studied in Bahmani et al. [2].

As was already noted in this paper, we are interested in deriving necessary optimality conditions for the problem of minimizing a continuously differentiable function over  $B \cap C_s$ . Such a study was carried out in Beck and Eldar [3] on the problem with  $B = \mathbb{R}^n$ , where several necessary optimality conditions related to the classical notions of stationarity and coordinatewise optimality were presented and analyzed. It was shown in Beck and Eldar [3] that coordinatewise optimality conditions are more restrictive (that is, stronger) than those, which are based on the notion of stationarity.

This hierarchy of the optimality conditions also implied a hierarchy between algorithms. In particular, the IHT method was shown to be inferior to coordinate descent-type methods.

The results of Beck and Eldar [3] are limited to the case where only a sparsity constraint is present, and the natural question, which is the main motivation for this paper, is whether we can generalize the results when additional constraints are imposed. We will answer this question affirmatively when certain symmetry assumptions (in addition to convexity and closedness) will be imposed on the set  $B$ . An important aspect of the theory on optimality conditions that will be developed is that it is naturally accompanied with appropriate algorithms that converge to the devised optimality conditions. Our insights on the optimality conditions will assist in qualifying the derived algorithms. In addition, the derivation of the conditions is made possible due to a development of a unified theory encompassing properties and algorithms related to the computation of the orthogonal projection operator onto sets with various symmetry properties.

**1.3. Paper layout.** Mathematical preliminaries that are required for the analysis of constrained sparse problems are defined and studied in §2. Section 3 is devoted to the study of properties of the orthogonal projection operator over symmetric sets. The two main results in this context are the monotonicity lemma and the order preservation property. Based on the results on the orthogonal projection operator onto sparse symmetric sets, we develop, in §4, a unified theory for efficiently computing sparse projections onto sparse sets. The theory on sparse orthogonal projections is the basis for the development of the stationarity-based optimality conditions presented and analyzed in §5. We continue with defining coordinatewise optimality conditions in §6, and explore their relation to the stationary-based ones, concluding with a hierarchy between all of the devised conditions and several results on the representation of the conditions. Section 7 presents several methods that are guaranteed to converge to the conceived optimality conditions, and finally, §8 illustrates the validity of the theoretical hierarchy by two sets of numerical experiments on problems over the sparse unit simplex—one on randomly generated data and the other on the sparse index tracking problem with real sampled data.

**1.4. Notation.** The complement of a set  $A$  is denoted by  $A^c$ . Matrices and vectors are denoted by boldface letters. The  $n$ -length vector of all zeros is denoted by  $\mathbf{0}_n$  and the  $n$ -length vector of all ones is denoted by  $\mathbf{1}_n$ . When the dimensions are clear from the context, we will frequently omit the subscripts and just write  $\mathbf{0}$  and  $\mathbf{1}$ . The vector  $\mathbf{e}_i$  has 1 in the  $i$ -th component and zeros elsewhere. For a vector  $\mathbf{x} \in \mathbb{R}^n$ , we define  $[\mathbf{x}]_+$  and  $|\mathbf{x}|$  to be the vectors whose  $i$ -th component is  $[x_i]_+ = \max\{x_i, 0\}$  and  $|x_i|$ , respectively. For any  $p \geq 1$ , the  $l_p$  ball in the space  $\mathbb{R}^n$  is denoted by

$$B_p^n[\mathbf{0}, 1] = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_p \leq 1\}.$$

The  $n$  dimensional unit simplex is given by

$$\Delta_n = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{1}^T \mathbf{x} = 1, \mathbf{x} \geq \mathbf{0}\},$$

and the unit-sum set is the set

$$\Delta'_n = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{1}^T \mathbf{x} = 1\}.$$

The sign vector of a given  $\mathbf{x} \in \mathbb{R}^n$  is denoted by  $\text{sign}(\mathbf{x})$ , and its  $i$ -th component is

$$\text{sign}(\mathbf{x})_i \equiv \begin{cases} 1 & x_i \geq 0, \\ -1 & x_i < 0. \end{cases}$$

For two vectors  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ , the Hadmard product is defined by  $\mathbf{x} \circ \mathbf{y} \equiv (x_i y_i)_{i=1}^n$ . Given a set  $S \subseteq \mathbb{R}^n$  and a vector  $\mathbf{x} \in \mathbb{R}^n$ , the orthogonal projection of  $\mathbf{x}$  onto  $S$  is defined as the set

$$P_S(\mathbf{x}) = \text{argmin}\{\|\mathbf{y} - \mathbf{x}\|^2: \mathbf{y} \in S\},$$

where here and elsewhere in the paper  $\|\cdot\|$  denotes the  $l_2$  norm on  $\mathbb{R}^n$ . If  $S$  is closed, the set  $P_S(\mathbf{x})$  is nonempty, and if, in addition,  $S$  is also convex, then  $P_S(\mathbf{x})$  is a singleton and we associate  $P_S(\mathbf{x})$  with the vector that it comprises. The gradient of a given function  $h: \mathbb{R}^n \rightarrow \mathbb{R}$  at a point  $\mathbf{x} \in \mathbb{R}^n$ , given that it exists, is denoted by  $\nabla h(\mathbf{x})$ . The  $i$ th partial derivative is denoted by  $\nabla_i h(\mathbf{x})$  and given two different indices  $i, j \in \{1, 2, \dots, n\}, i \neq j$ , the vector  $\nabla_{i,j} h(\mathbf{x})$  is the two-dimensional column vector  $(\nabla_i h(\mathbf{x}), \nabla_j h(\mathbf{x}))^T$ .

## 2. Mathematical preliminaries.

**2.1. Stationarity in smooth problems over convex sets.** We begin by recalling the notion of stationarity in smooth problems over closed and convex sets. Consider the problem

$$\min\{h(\mathbf{x}): \mathbf{x} \in C\}, \tag{3}$$

where  $h: \mathbb{R}^n \rightarrow \mathbb{R}$  is a continuously differentiable function and  $C \subseteq \mathbb{R}^n$  is a closed and convex set. A vector  $\mathbf{x}^*$  is called a *stationary point* of (3) if

$$\nabla h(\mathbf{x}^*)^T (\mathbf{x} - \mathbf{x}^*) \geq 0 \quad \text{for any } \mathbf{x} \in C. \tag{4}$$

This necessary optimality condition means that there are no feasible descent directions at  $\mathbf{x}^*$ . It is well known that the condition can be rewritten as

$$\mathbf{x}^* = P_C \left( \mathbf{x}^* - \frac{1}{L} \nabla h(\mathbf{x}^*) \right) \tag{5}$$

for some  $L > 0$ . Interestingly, although condition (5) is expressed in terms of the parameter  $L$ , it is actually independent of  $L$  by its equivalence to condition (4). For many special cases of the set  $C$ , there are more explicit expressions of the stationarity condition that are easier to handle. We recall in Table 1 some of the examples that will be used later on in the paper.

When the objective function  $h$  is convex, then stationarity is a necessary *and sufficient* condition for optimality.

TABLE 1. Explicit stationarity conditions for simple sets.

Feasible set	Explicit stationarity condition
$\mathbb{R}^n$	$\nabla f(\mathbf{x}^*) = \mathbf{0}$
$\mathbb{R}_+^n$	$\frac{\partial f}{\partial x_i}(\mathbf{x}^*) \begin{cases} = 0, & x_i^* > 0, \\ \geq 0, & x_i^* = 0. \end{cases}$
$\Delta_n$	$\exists \mu \in \mathbb{R}: \frac{\partial f}{\partial x_i}(\mathbf{x}^*) \begin{cases} = \mu, & x_i > 0, \\ \geq \mu, & x_i = 0. \end{cases}$
$\Delta'_n$	$\exists \mu \in \mathbb{R}: \frac{\partial f}{\partial x_i}(\mathbf{x}^*) = \mu, i = 1, 2, \dots, n.$
$B_2^n[\mathbf{0}, 1]$	$\nabla f(\mathbf{x}^*) = \mathbf{0}$ or $\ \mathbf{x}^*\  = 1$ and $\exists \lambda \leq 0: \nabla f(\mathbf{x}^*) = \lambda \mathbf{x}^*$ .
$[\ell, u]^n$ ( $\ell < u$ )	$\frac{\partial f}{\partial x_i}(\mathbf{x}^*) \begin{cases} = 0, & \ell < x_i < u, \\ \geq 0, & x_i = \ell, \\ \leq 0, & x_i = u. \end{cases}$

**2.2. The class of  $C_L^{1,1}$  functions.** A function  $h: \mathbb{R}^n \rightarrow \mathbb{R}$  is said to belong to  $C_L^{1,1}$  if it is continuously differentiable and its gradient is Lipschitz continuous with parameter  $L > 0$ , meaning that

$$\|\nabla h(\mathbf{x}) - \nabla h(\mathbf{y})\| \leq L\|\mathbf{x} - \mathbf{y}\| \quad \text{for all } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

An important property of  $C_L^{1,1}$  functions is described in the well-known descent lemma.

LEMMA 2.1 (DESCENT LEMMA). *Suppose that  $h \in C_{L(h)}^{1,1}$ . Then, for any  $\mathbf{x}, \mathbf{d} \in \mathbb{R}^n$  and  $L \geq L(h)$ , the following inequality is satisfied:*

$$h(\mathbf{x} + \mathbf{d}) \leq h(\mathbf{x}) + \nabla h(\mathbf{x})^T \mathbf{d} + \frac{L}{2} \|\mathbf{d}\|^2.$$

We will also be interested in a more refined version of the descent lemma, called *the block descent lemma* in which the perturbation vector  $\mathbf{d}$  has at most two nonzero components. For that, we will define the block Lipschitz constant. Let  $h \in C_{L(h)}^{1,1}$ . Then, for any  $i \neq j$ , there exists a constant  $L_{i,j}(h)$  for which

$$\|\nabla_{i,j} h(\mathbf{x}) - \nabla_{i,j} h(\mathbf{x} + \mathbf{d})\| \leq L_{i,j}(h) \|\mathbf{d}\| \quad (6)$$

for any  $\mathbf{x} \in \mathbb{R}^n$  and any  $\mathbf{d} \in \mathbb{R}^n$ , which has at most two nonzero components. Here,  $\nabla_{i,j} h(\mathbf{x})$  denotes a vector of length-2 whose elements are the  $i$ -th and  $j$ -th elements of  $\nabla h(\mathbf{x})$ . The block Lipschitz constants are defined as

$$L_2(h) \equiv \max_{i \neq j} L_{i,j}(h).$$

Clearly,  $L_2(h) \leq L(h)$  and, in general, the block Lipschitz constant  $L_2(h)$  can be much smaller than the global Lipschitz constant  $L(h)$ . The block Lipschitz constant is used in a “block” version of the descent lemma.

LEMMA 2.2 (BLOCK DESCENT LEMMA). *Suppose that  $h \in C_L^{1,1}$  and that  $L \geq L_2(h)$ . Then,*

$$h(\mathbf{x} + \mathbf{d}) \leq h(\mathbf{x}) + \nabla h(\mathbf{x})^T \mathbf{d} + \frac{L}{2} \|\mathbf{d}\|^2$$

for any vector  $\mathbf{d} \in \mathbb{R}^n$  with at most two nonzero components.

**2.3. Supports, super supports, and restriction on index sets.** The support set of a vector  $\mathbf{x} \in \mathbb{R}^n$  is denoted by

$$I_1(\mathbf{x}) \equiv \{i \in \{1, \dots, n\}: x_i \neq 0\}.$$

The off-support set of a vector  $\mathbf{x} \in \mathbb{R}^n$  is denoted by

$$I_0(\mathbf{x}) \equiv \{i \in \{1, \dots, n\}: x_i = 0\}.$$

Of course,  $I_0(\mathbf{x})$  is the complement of  $I_1(\mathbf{x})$ . A vector is said to have a *full support* if  $\|\mathbf{x}\|_0 = s$ , and to have an *incomplete support* if  $\|\mathbf{x}\|_0 < s$ . A set  $T$  is a *super support* of a vector  $\mathbf{y} \in C_s \cap B$  if  $I_1(\mathbf{y}) \subseteq T$  and  $|T| = s$ . Of course, if  $\mathbf{y}$  has a full support, then the only super support set is the support set itself. However, if  $\mathbf{y}$  does not have a full support, then there are  $\binom{n - \|\mathbf{y}\|_0}{s - \|\mathbf{y}\|_0}$  possible super supports. For example, if  $s = 3$ ,  $n = 5$ , and  $\mathbf{y} = (-3, 4, 0, 0, 0)^T$ . Then, the three super supports of  $\mathbf{y}$  are

$$\{1, 2, 3\}, \quad \{1, 2, 4\}, \quad \{1, 2, 5\}.$$

Given a vector  $\mathbf{x} \in \mathbb{R}^n$ , the vector composed of the components of  $\mathbf{x}$  whose indices are in a given subset  $T \subseteq \{1, \dots, n\}$  is denoted by  $\mathbf{x}_T \in \mathbb{R}^{|T|}$ , the matrix  $\mathbf{U}_T$  denotes the submatrix of the  $n \times n$  identity matrix  $\mathbf{I}_n$  constructed from the columns corresponding to the index set  $T$ . In this notation,  $\mathbf{x}_T = \mathbf{U}_T^T \mathbf{x}$  (note that the superscript stands for the transpose operation). In addition, if  $T$  is a super support of a vector  $\mathbf{x} \in \mathbb{R}^n$ , then  $\mathbf{x} = \mathbf{U}_T \mathbf{x}_T$ . We use the notation

$$B_T = \{\mathbf{x} \in \mathbb{R}^{|T|}: \mathbf{U}_T \mathbf{x} \in B\},$$

and the set  $B_T$  will be called “the restriction of  $B$  to  $T$ .” For example, if

$$B = \{(x_1, x_2, x_3, x_4)^T: x_1 + 2x_2 + 3x_3 + 4x_4 = 1\},$$

then

$$B_{\{1,2\}} = \{(x_1, x_2)^T: x_1 + 2x_2 = 1\}, \quad B_{\{2,4\}} = \{(x_2, x_4)^T: 2x_2 + 4x_4 = 1\}.$$

In a similar manner, given a continuously differentiable function  $f$ , we denote “the restriction of the vector  $\nabla f(\mathbf{x})$  to  $T$ ” by  $\nabla_T f(\mathbf{x}) = \mathbf{U}_T^T \nabla f(\mathbf{x})$ . For example, if  $f(\mathbf{x}) = x_1 x_2 + x_2^2 + x_3^3$  and  $T = \{1, 3\}$ , then

$$\nabla_T f(\mathbf{x}) = \begin{pmatrix} x_2 \\ 3x_3^2 \end{pmatrix}.$$

TABLE 2. Simple sets and their symmetry properties.

Set	Description	Type-1	Nonneg. type-1	Type-2
$\mathbb{R}^n$	Entire space	✓		✓
$\mathbb{R}_+^n$	Nonnegative orthant	✓	✓	
$\Delta_n$	Unit simplex	✓	✓	
$\Delta_n$	Unit sum	✓		
$B_p[0, 1]$ ( $p \geq 1$ )	$p$ -ball	✓		✓
$[\ell, u]^n$ ( $\ell < u$ )	Box	✓		

**3. Projection onto symmetric sets.** In this section, we present two types of set symmetries that will be discussed in the paper, and study some key properties related to the orthogonal projection operator onto these types of sets. We will show later on, that the derived properties of the orthogonal projection operator will be important in the derivation and study of various optimality conditions and algorithms for sparsity-constrained problems.

**3.1. Type-1 and type-2 symmetries.** The permutation group of the set of indices  $\{1, \dots, n\}$  will be denoted by  $\Sigma_n$ , and for a given vector  $\mathbf{x} \in \mathbb{R}^n$  and a permutation  $\sigma \in \Sigma_n$ , the vector  $\mathbf{x}^\sigma$  is the vector defined by

$$(\mathbf{x}^\sigma)_i = x_{\sigma(i)},$$

that is, the vector that is a reordering of  $\mathbf{x}$  according to  $\sigma$ . For example, if  $\mathbf{x} = (4, 5, 6)^T$ , and  $\sigma$  is the permutation given by

$$\sigma(1) = 3, \quad \sigma(2) = 2, \quad \sigma(3) = 1,$$

then

$$\mathbf{x}^\sigma = (6, 5, 4)^T.$$

In this example,  $\sigma$  reordered the elements of  $\mathbf{x}$  in a nonascending order. Because such permutations will be important in our analysis, we will define them formally.

**DEFINITION 3.1 (SORTING PERMUTATIONS).** Let  $\mathbf{x} \in \mathbb{R}^n$ . Then, a permutation that sorts the elements of  $\mathbf{x}$  in a nonascending order will be called a *sorting permutation*. The set of all the sorting permutations of  $\mathbf{x}$  is a subset of  $\Sigma_n$  and will be denoted by  $\tilde{\Sigma}(\mathbf{x})$ . Explicitly,  $\sigma \in \tilde{\Sigma}(\mathbf{x})$  if and only if

$$x_{\sigma(1)} \geq x_{\sigma(2)} \geq \dots \geq x_{\sigma(n-1)} \geq x_{\sigma(n)}.$$

**DEFINITION 3.2 (SWAP PERMUTATIONS).** The swap permutation<sup>1</sup>  $\tau_{i,j} \in \Sigma_n$ , which swaps between two indices  $i, j \in \{1, 2, \dots, n\}$  is defined for any  $i \neq j$  by

$$\tau_{i,j}(l) = \begin{cases} l & l \neq i, j, \\ i & l = j, \\ j & l = i. \end{cases}$$

We will consider the following classes of sets, and most of the analysis in this paper will be based on assuming that the underlying set belongs to at least one of these classes.

**DEFINITION 3.3 (TYPE-1 SYMMETRIC SETS).** Let  $D \subseteq \mathbb{R}^n$ . Then,  $D$  is a *type-1 symmetric set* if for any vector  $\mathbf{x} \in D$  and  $\sigma \in \Sigma_n$ , we have  $\mathbf{x}^\sigma \in D$ .

**DEFINITION 3.4 (NONNEGATIVE SETS).** A set  $D \subseteq \mathbb{R}^n$  is *nonnegative* if  $\mathbf{x} \geq \mathbf{0}$  for any  $\mathbf{x} \in D$ .

**DEFINITION 3.5 (TYPE-2 SYMMETRIC SETS).** Let  $D \subseteq \mathbb{R}^n$  be a type-1 symmetric set. Then,  $D$  is a *type-2 symmetric set* if for any  $\mathbf{x} \in D$ ,  $\sigma \in \Sigma_n$ , and  $\mathbf{y} \in \{-1, 1\}^n$ , the vector  $\mathbf{x} \circ \mathbf{y} = (x_i y_i)_{i=1}^n$  is in  $D$ .

Type-1 and type-2 symmetric sets appear quite often like feasibility sets in optimization problems. Some frequently appearing examples, as well as their affiliation to the different symmetry types are summarized in Table 2.

<sup>1</sup> Swap permutations are also called “transpositions” in the literature.

**3.2. Basic properties of projections on symmetric sets.** Our objective is to show that for symmetric type-1 sets, the orthogonal projection operator satisfies an order preservation property. This property will be the basis for the detection of a super support in the problem of projecting onto  $C_s \cap B$ .

We introduce the following lemma, describing an important monotonicity property associated with projections onto type-1 symmetric sets that will play an important role in the derivation of the order preservation property.

**LEMMA 3.1 (SYMMETRIC PROJECTION MONOTONICITY LEMMA).** *Let  $D$  be a symmetric type-1 set. Let  $\mathbf{x} \in \mathbb{R}^n$  and  $\mathbf{y} \in P_D(\mathbf{x})$ . Then, for any permutation  $\sigma \in \Sigma_n$ , it holds that*

$$\mathbf{y}^T(\mathbf{x} - \mathbf{x}^\sigma) \geq 0. \quad (7)$$

In particular,

$$(y_i - y_j)(x_i - x_j) \geq 0 \quad (8)$$

for any  $i, j \in \{1, 2, \dots, n\}$ .

**PROOF.** Since  $\mathbf{y} \in P_D(\mathbf{x})$  and  $D$  is type-1 symmetric, for any permutation  $\sigma \in \Sigma_n$ , we have

$$\|\mathbf{x} - \mathbf{y}\|^2 \leq \|\mathbf{x} - \mathbf{y}^{\sigma^{-1}}\|^2, \quad (9)$$

where  $\sigma^{-1}$  denotes the inverse permutation of  $\sigma$ . The inequality (9) is equivalent to (omitting the term  $\|\mathbf{x}\|^2$ )

$$\|\mathbf{y}\|^2 - 2\mathbf{y}^T\mathbf{x} \leq \|\mathbf{y}^{\sigma^{-1}}\|^2 - 2(\mathbf{y}^{\sigma^{-1}})^T\mathbf{x}.$$

Hence, using the fact that  $\|\mathbf{y}\|^2 = \|\mathbf{y}^{\sigma^{-1}}\|^2$ , it follows that

$$\mathbf{y}^T\mathbf{x} - (\mathbf{y}^{\sigma^{-1}})^T\mathbf{x} \geq 0.$$

Finally, by the obvious identity  $(\mathbf{a}^{\sigma^{-1}})^T\mathbf{b} = \mathbf{a}^T\mathbf{b}^\sigma$  for any  $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ , we obtain that

$$\mathbf{y}^T(\mathbf{x} - \mathbf{x}^\sigma) \geq 0.$$

Plugging the swap permutation  $\sigma = \tau_{i,j}$  into the latter inequality yields

$$(y_i - y_j)(x_i - x_j) \geq 0. \quad \square \quad (10)$$

**DEFINITION 3.6.** Let  $\mathbf{x} \in \mathbb{R}^n$ . Then, a permutation  $\sigma \in \Sigma_n$  will be called a *value-preserving permutation* of  $\mathbf{x}$  if  $\mathbf{x} = \mathbf{x}^\sigma$ . The set of all of value-preserving permutations of  $\mathbf{x}$  is denoted by  $\Sigma^v(\mathbf{x})$ .

**EXAMPLE 3.1.** If  $\mathbf{x} = (2, 1, 2, 1)^T$ , then

$$\Sigma^v(\mathbf{x}) = \left\{ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix} \right\},$$

where a permutation  $\sigma \in \Sigma_4$  is denoted here as

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ \sigma(1) & \sigma(2) & \sigma(3) & \sigma(4) \end{pmatrix}.$$

The next lemma shows that if  $\mathbf{y} \in P_D(\mathbf{x})$  for some type-1 symmetric set  $D$ , then  $\mathbf{y}^\sigma \in P_D(\mathbf{x})$  for each value preservation permutation  $\sigma$  of  $\mathbf{x}$ .

**LEMMA 3.2.** *Let  $D$  be a type-1 symmetric set,  $\mathbf{x} \in \mathbb{R}^n$  and  $\mathbf{y} \in P_D(\mathbf{x})$ . Then, for any  $\sigma \in \Sigma^v(\mathbf{x})$ , it holds that  $\mathbf{y}^\sigma \in P_D(\mathbf{x})$ .*

**PROOF.** Suppose that  $\mathbf{y} \in P_D(\mathbf{x})$ . Then, by the symmetry of  $D$ ,  $\mathbf{y}^\sigma \in D$ . Since minimizing the function  $f(\mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|^2$  is equivalent to minimizing the function  $g(\mathbf{y}) = \|\mathbf{y}\|^2 - 2\mathbf{y}^T\mathbf{x}$ , all we need to show is that  $g(\mathbf{y}) = g(\mathbf{y}^\sigma)$ , and indeed,

$$g(\mathbf{y}^\sigma) = \|\mathbf{y}^\sigma\|^2 - 2(\mathbf{y}^\sigma)^T\mathbf{x} = \|\mathbf{y}\|^2 - 2(\mathbf{y}^\sigma)^T\mathbf{x}^\sigma = \|\mathbf{y}\|^2 - 2\mathbf{y}^T\mathbf{x} = g(\mathbf{y}),$$

where we used the facts that  $\|\mathbf{y}^\sigma\| = \|\mathbf{y}\|$ ,  $\mathbf{x}^\sigma = \mathbf{x}$  and  $(\mathbf{y}^\sigma)^T\mathbf{x}^\sigma = \mathbf{y}^T\mathbf{x}$ .  $\square$

We now turn to prove a key property of the projection set  $P_D(\mathbf{x})$  in the case of type-1 symmetric sets—there always exists a vector in the projection set whose components are in the same order as  $\mathbf{x}$ .

**THEOREM 3.1 (ORDER PRESERVATION PROPERTY).** *Let  $D$  be a type-1 symmetric set,  $\mathbf{x} \in \mathbb{R}^n$  and  $\sigma \in \tilde{\Sigma}(\mathbf{x})$ . Suppose that  $P_D(\mathbf{x})$  is nonempty. Then, there exists  $\mathbf{y} \in P_D(\mathbf{x})$  such that  $\sigma \in \tilde{\Sigma}(\mathbf{y})$ .*

**PROOF.** Let  $\mathbf{y} \in P_D(\mathbf{x})$ , and suppose that there exist indices  $i_1 < i_2$  such that  $y_{\sigma(i_1)} < y_{\sigma(i_2)}$  (otherwise, the proof is complete). By Lemma 3.1,

$$(y_{\sigma(i_1)} - y_{\sigma(i_2)})(x_{\sigma(i_1)} - x_{\sigma(i_2)}) \geq 0,$$

which in turn, implies that  $x_{\sigma(i_1)} \leq x_{\sigma(i_2)}$ . Since  $\sigma \in \tilde{\Sigma}(\mathbf{x})$ , it follows that  $x_{\sigma(i_1)} \geq x_{\sigma(i_2)}$ , and hence  $x_{\sigma(i_1)} = x_{\sigma(i_2)}$ . Therefore, the swap permutation  $\hat{\sigma} = \tau_{\sigma(i_1), \sigma(i_2)}$  in which  $\sigma(i_1)$  is swapped with  $\sigma(i_2)$ , is a value-preserving permutation of  $\mathbf{x}$ . By Lemma 3.2,  $\mathbf{y}^{\hat{\sigma}} \in P_D(\mathbf{x})$ , and we set  $\mathbf{y} \leftarrow \mathbf{y}^{\hat{\sigma}}$ . This procedure can be repeated as long as there are indices  $i < j$ , which violate the order ( $y_{\sigma(i)} < y_{\sigma(j)}$ ). Since at each iteration of the procedure, the number of pairs of indices which violate the order is strictly reduced, the process is finite and ends with a vector  $\mathbf{y}$  for which  $\sigma \in \tilde{\Sigma}(\mathbf{y})$ .  $\square$

**EXAMPLE 3.2.** Consider the type-1 symmetric set  $D = \Delta'_4 \cap C_2$  (for a definition of  $C_s$ , see (2)), and consider the problem of finding  $P_D(\mathbf{x})$ , where  $\mathbf{x} = (-4, 3, 1, -4)^T$ . A sorting permutation of  $\mathbf{x}$  is

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix}.$$

This is, of course, not the only sorting permutation. A simple computation shows that the projection set of  $\mathbf{x}$  is

$$P_{\Delta'_4 \cap C_2}(\mathbf{x}) = \{\mathbf{v}_1 = (-4, 3, 0, 0)^T, \mathbf{v}_2 = (0, 3, 0, -4)^T\}.$$

Since  $\Delta'_4 \cap C_2$  is a type-1 symmetric set, then by Theorem 3.1, it follows that  $\sigma$  is a sorting permutation of at least one of the projection vectors, and indeed,  $\sigma \in \tilde{\Sigma}(\mathbf{v}_2)$ .

**3.3. From type-2 to nonnegative type-1 symmetric sets.** The analysis of type-2 symmetric sets will be frequently done by using the following lemma that connects the projection operator on the two types of sets. In many cases, the analysis of type-1 symmetric sets along with Lemma 3.3 will immediately imply the corresponding results for nonnegative type-1 symmetric sets.

**LEMMA 3.3.** *Let  $D \subseteq \mathbb{R}^n$  be a type-2 symmetric set. Then, the relation*

$$\tilde{\mathbf{y}} \in P_D(\mathbf{x})$$

*holds if and only if*

$$\text{sign}(\mathbf{x}) \circ \tilde{\mathbf{y}} \in P_{D \cap \mathbb{R}_+^n}(|\mathbf{x}|).$$

**PROOF.** Denote  $\mathbf{E} = \text{diag}(\text{sign}(\mathbf{x}))$ . We have for any  $\mathbf{y} \in \mathbb{R}^n$ ,

$$\|\mathbf{y} - \mathbf{x}\|^2 = \|\mathbf{E}\mathbf{y} - \mathbf{E}\mathbf{x}\|^2 = \|\mathbf{E}\mathbf{y} - |\mathbf{x}|\|^2.$$

Hence the projection problem

$$\min_{\mathbf{y}} \{\|\mathbf{y} - \mathbf{x}\|^2 : \mathbf{y} \in D\} \tag{11}$$

is the same as

$$\min_{\mathbf{y}} \{\|\mathbf{E}\mathbf{y} - |\mathbf{x}|\|^2 : \mathbf{y} \in D\}. \tag{12}$$

By the type-2 symmetry property of  $D$ , we conclude that (11) is equivalent to

$$\min_{\mathbf{y}} \{\|\mathbf{E}\mathbf{y} - |\mathbf{x}|\|^2 : \mathbf{E}\mathbf{y} \in D\},$$

and by making the change of variables  $\mathbf{z} = \mathbf{E}\mathbf{y}$ , we arrive at the following equivalent formulation:

$$\min_{\mathbf{z}} \{\|\mathbf{z} - |\mathbf{x}|\|^2 : \mathbf{z} \in D\}. \tag{13}$$

Now, all the optimal solutions of (13) must be nonnegative. Otherwise, if  $\mathbf{z}$  is an optimal solution for which  $z_i < 0$  for some  $i$ , then the vector  $\tilde{\mathbf{z}} \in C_s \cap B$  defined by

$$\tilde{z}_j = \begin{cases} z_j, & j \neq i, \\ -z_i, & j = i, \end{cases}$$

has a smaller objective function value, which is a contradiction to the assumed optimality of  $\mathbf{z}$  w.r.t. (13). Therefore, we can add (redundant) nonnegativity constraints, and conclude that  $\tilde{\mathbf{y}}$  is an optimal solution of (11) if and only if  $\mathbf{E}\tilde{\mathbf{y}}$  is an optimal solution of

$$\min_{\mathbf{z}} \{ \|\mathbf{z} - |\mathbf{x}|\|^2 : \mathbf{z} \in D \cap \mathbb{R}_+^n \}, \quad (14)$$

that is, if and only if  $\text{sign}(\mathbf{x}) \circ \tilde{\mathbf{y}} = \mathbf{E}\tilde{\mathbf{y}} \in P_{D \cap \mathbb{R}_+^n}(|\mathbf{x}|)$ .  $\square$

The following is a direct consequence of the latter lemma.

**COROLLARY 3.1.** *Let  $D$  be a type-2 symmetric set. If  $\mathbf{y} \in P_D(\mathbf{x})$ , then  $|\mathbf{y}| \in P_{D \cap \mathbb{R}_+^n}(|\mathbf{x}|)$ .*

**PROOF.** By Lemma 3.3, it follows that  $\text{sign}(\mathbf{x}) \circ \mathbf{y} \in P_{D \cap \mathbb{R}_+^n}(|\mathbf{x}|)$ . Since  $D \cap \mathbb{R}_+^n$  is obviously a nonnegative set, it follows that  $\text{sign}(\mathbf{x}) \circ \mathbf{y} = |\text{sign}(\mathbf{x}) \circ \mathbf{y}| = |\mathbf{y}|$ , from which the desired relation  $|\mathbf{y}| \in P_{D \cap \mathbb{R}_+^n}(|\mathbf{x}|)$  follows.  $\square$

#### 4. Sparse projection over symmetric sets.

**4.1. The problem.** So far, we have studied several properties of projection onto symmetric sets. Building on the results of the previous section, this section studies the properties of the orthogonal projection operator onto the intersection of a symmetric closed and convex set  $B$  and the set of  $s$ -sparse vectors. As a by-product of this theoretical study, we show how the operator can be efficiently computed under various symmetry properties. Later on, the derived properties will play a key role in characterizing various optimality conditions of problem (P) (see §5).

The following is the exact problem we consider:

*The sparse projection problem*

Given a closed and convex set  $B$ , and a vector  $\mathbf{x} \in \mathbb{R}^n$ , find an element in the orthogonal projection set of  $\mathbf{x}$  onto  $B \cap C_s$ :

$$P_{C_s \cap B}(\mathbf{x}) = \text{argmin} \{ \|\mathbf{z} - \mathbf{x}\|^2 : \mathbf{z} \in B \cap C_s \}. \quad (15)$$

We will refer to  $P_{C_s \cap B}$  as “the  $s$ -sparse projection set onto  $B$ ,” and an element of the latter set is called “an  $s$ -sparse projection vector onto  $B$ ,” or just “a sparse projection vector.” By the closedness of  $B \cap C_s$ , it follows that  $P_{C_s \cap B}(\mathbf{x})$  is a nonempty set. However,  $C_s \cap B$  is nonconvex, and hence the set  $P_{C_s \cap B}(\mathbf{x})$  is not always a singleton. For example, when  $B = \mathbb{R}^n$ ,  $P_{C_s \cap B}(\mathbf{x}) = P_{C_s}(\mathbf{x})$  is comprised of all vectors consisting of the  $s$  components of  $\mathbf{x}$  with the largest absolute values, and with zeros elsewhere. In general, there can be more than one choice to the  $s$  largest components in absolute value, and each of these choices gives rise to another vector in the set  $P_{C_s}(\mathbf{x})$ . For example,

$$P_{C_2}((2, 1, 1)^T) = \{(2, 1, 0)^T, (2, 0, 1)^T\}.$$

Finding the set  $P_{C_s \cap B}(\mathbf{x})$ , or even just a vector in the set, is, in general, a difficult task since the corresponding optimization problem is nonconvex. However, we will show that under several symmetry properties, finding such a vector is a tractable mission. We would like to stress that, as shown in the sequel, the orthogonal projection plays a key role in the sparse optimization problem. There are very few examples in the literature of computations of sparse projections. One such paper is Kyrillidis et al. [18], which presents an algorithm for finding sparse projections onto the unit-simplex and unit-sum sets, and also covers the case where the sum of variables is not necessarily one. Another paper in which the sparse projection operator was mentioned is Luss and Teboulle [21], where as part of a study of an algorithm for sparse PCA, the authors computed the sparse projection operator onto the  $l_2$ -norm ball. In Bolte et al. [8, Proposition 4], a formula for the sparse projection operator onto the nonnegative orthant was derived. We will present a unified theory that will enable us to compute sparse projections onto closed and convex and symmetric sets. The mentioned examples will be special cases of the general theory.

We begin with a simple lemma, which shows that the  $s$ -sparse projection onto  $B$  can be done in two phases: finding a super support at a first stage, and then finding the projection onto the restriction of  $B$  to the super support.

**LEMMA 4.1.** *Let  $\mathbf{x} \in \mathbb{R}^n$ . Suppose that  $\mathbf{y} \in P_{C_s \cap B}(\mathbf{x})$ . Then, for any super support set  $T$  of  $\mathbf{y}$ , it holds that*

$$\mathbf{y}_T = P_{B_T}(\mathbf{x}_T).$$

PROOF. Suppose on the contrary that  $\mathbf{y} \in P_{C_s \cap B}(\mathbf{x})$ , but  $\mathbf{y}_T \neq P_{B_T}(\mathbf{x}_T)$ . The relation  $\mathbf{y} \in P_{C_s \cap B}(\mathbf{x})$  explicitly means that  $\mathbf{y}$  is a solution of the minimization problem

$$\min_{\mathbf{z} \in C_s \cap B} \|\mathbf{z} - \mathbf{x}\|^2. \quad (16)$$

The objective function in the latter minimization problem, evaluated at  $\mathbf{z} = \mathbf{y}$ , can be decomposed into two terms: one comprised of the elements in  $T^c$ , which are all nonsupport elements, and the other comprised of elements corresponding to the set  $T$ :

$$\|\mathbf{y} - \mathbf{x}\|^2 = \|\mathbf{y}_T - \mathbf{x}_T\|^2 + \|\mathbf{x}_{T^c}\|^2.$$

Denote by  $\mathbf{u} \neq \mathbf{y}$ , the vector for which it holds that  $\mathbf{u}_T = P_{B_T}(\mathbf{x}_T)$  and  $\mathbf{u}_{T^c} = \mathbf{0}$ . Then,  $\mathbf{u}_T \neq \mathbf{y}_T$ , and by the uniqueness of the orthogonal projection operator onto closed and convex sets, it follows that  $\|\mathbf{u}_T - \mathbf{x}_T\|^2 < \|\mathbf{y}_T - \mathbf{x}_T\|^2$ , and hence

$$\|\mathbf{y} - \mathbf{x}\|^2 = \|\mathbf{y}_T - \mathbf{x}_T\|^2 + \|\mathbf{x}_{T^c}\|^2 > \|\mathbf{u}_T - \mathbf{x}_T\|^2 + \|\mathbf{x}_{T^c}\|^2 = \|\mathbf{u} - \mathbf{x}\|^2,$$

thus contradicting the optimality of  $\mathbf{y}$  w.r.t. (16).  $\square$

The latter lemma implies that if a super support  $T$  of a sparse projection vector is known, then the vector  $\mathbf{v} = P_{B_T}(\mathbf{x})$  induces a sparse projection vector in the sense that  $\mathbf{U}_T \mathbf{v} \in P_{C_s \cap B}(\mathbf{x})$ . Therefore, an inefficient method to find a vector in  $P_{C_s \cap B}(\mathbf{x})$  is to go over all the potential  $\binom{n}{s}$  super supports  $T \subseteq \{1, 2, \dots, n\}$ ,  $|T| = s$ , compute the corresponding projections  $P_{B_T}(\mathbf{x}_T)$ , and finally, choose the support corresponding to the minimal distance  $\|\mathbf{x} - \mathbf{U}_T P_{B_T}(\mathbf{x}_T)\|$ .

**4.2. Sparse projections onto type-1 symmetric sets.** It is important to note when  $B$  is nonnegative/type-1 symmetric/type-2 symmetric, then so is the set  $D = C_s \cap B$ , so all the results of §3 can be employed. Using these results, we will show that under general symmetry properties, a super support set of a sparse projection vector can be evaluated efficiently in advance, without the need of the exhaustive search procedure described above. We will show how the order preservation property (Theorem 3.1) leads to the insight that there are at most  $s + 1$  possible support sets. For that, we first define the set  $S_{[j_1, j_2]}^\sigma$  as the set of indices from  $\sigma(j_1)$  to  $\sigma(j_2)$ .

DEFINITION 4.1. For any permutation  $\sigma \in \Sigma_n$ , the set  $S_{[j_1, j_2]}^\sigma$  is defined as

$$S_{[j_1, j_2]}^\sigma = \begin{cases} \{\sigma(j_1), \sigma(j_1 + 1), \dots, \sigma(j_2)\}, & 0 < j_1 \leq j_2 \leq n, \\ \emptyset, & \text{otherwise.} \end{cases}$$

Now, the  $s$ -sparse projection onto a symmetric type-1 set can be evaluated efficiently in  $s + 1$  steps, by searching for a super support among  $s + 1$  possibilities, stated in the following theorem.

THEOREM 4.1 (SYMMETRIC TYPE-1 PROJECTION THEOREM). *Let  $B$  be a closed and convex type-1 symmetric set,  $\mathbf{x} \in \mathbb{R}^n$ , and  $\sigma \in \tilde{\Sigma}(\mathbf{x})$ . Then, there exists  $\mathbf{y} \in P_{C_s \cap B}(\mathbf{x})$  for which*

$$I_1(\mathbf{y}) \subseteq S_{[1, k]}^\sigma \cup S_{[n+k-(s-1), n]}^\sigma \quad (17)$$

for some  $k \in \{0, \dots, s\}$ .

PROOF. Note that by the closedness of  $C_s \cap B$ , it follows that  $P_{C_s \cap B}(\mathbf{x})$  is nonempty, and hence by the order preservation property (Theorem 3.1), there exists  $\mathbf{y} \in P_{C_s \cap B}(\mathbf{x})$  such that  $\sigma \in \tilde{\Sigma}(\mathbf{y})$ . Define

$$i_{\max} = \begin{cases} 0, & \mathbf{y} \leq \mathbf{0}, \\ \max\{i: y_{\sigma(i)} > 0\}, & \text{otherwise,} \end{cases}$$

$$i_{\min} = \begin{cases} n + 1, & \mathbf{y} \geq \mathbf{0}, \\ \min\{i: y_{\sigma(i)} < 0\}, & \text{otherwise.} \end{cases}$$

Evidently,

$$I_1(\mathbf{y}) = S_{[1, i_{\max}]}^\sigma \cup S_{[i_{\min}, n]}^\sigma.$$

Since the number of nonzero elements in  $\mathbf{y}$  is at most  $s$ , we have that

$$i_{\max} + (n - i_{\min} + 1) \leq s. \quad (18)$$

Define  $k = i_{\max}$ , then by (18), it follows that  $i_{\min} \geq n + k - (s - 1)$ , and hence

$$S_{[i_{\min}, n]}^{\sigma} \subseteq S_{[n+k-(s-1), n]}^{\sigma},$$

and we conclude (recalling that  $k = i_{\max}$ ) that

$$I_1(\mathbf{y}) = S_{[1, i_{\max}]}^{\sigma} \cup S_{[i_{\min}, n]}^{\sigma} \subseteq S_{[1, k]}^{\sigma} \cup S_{[n+k-(s-1), n]}^{\sigma},$$

establishing the desired result.  $\square$

The type-1 symmetric projection theorem readily implies that a super support of a sparse projection vector can be found by going over  $s + 1$  possible super support sets, computing the orthogonal projections over the restriction of  $B$  to the corresponding index set, and choosing the vector associated with the minimal distance. This method is described in details below.

---

**Algorithm 1** (Projection onto a type-1 symmetric set).

**Input:**  $\mathbf{x} \in \mathbb{R}^n$ .

**Output:**  $\mathbf{u} \in P_{C_s \cap B}(\mathbf{x})$ .

1. Find  $\sigma \in \tilde{\Sigma}(\mathbf{x})$ .
  2. For any  $k = s, s - 1, \dots, 0$  do:
    - (a) Set  $T_k = S_{[1, k]}^{\sigma} \cup S_{[n+k-(s-1), n]}^{\sigma}$ .
    - (b) Compute  $\mathbf{g}_k = P_{B_{T_k}}(\mathbf{x}_{T_k})$  and define  $\mathbf{z}^k = \mathbf{U}_{T_k} \mathbf{g}_k$ .
  3. Return  $\mathbf{u} = \operatorname{argmin}\{\|\mathbf{z} - \mathbf{x}\|^2: \mathbf{z} \in \{\mathbf{z}^k: k = s, s - 1, \dots, 0\}\}$ .
- 

REMARK 4.1. Note that it is not really necessary to compute a sorting permutation  $\sigma \in \tilde{\Sigma}(\mathbf{x})$ , as described in step 1, and it is actually enough to be able to compute the sets  $T_k$ , which can be done in linear time.

**4.3. Sparse projection onto nonnegative type-1 symmetric sets.** For nonnegative type-1 symmetric sets, a super support of a vector in  $P_{C_s \cap B}(\mathbf{x})$  can be found instantly: a set containing the indices corresponding to the  $s$  largest values of  $\mathbf{x}$ .

**THEOREM 4.2 (NONNEGATIVE TYPE-1 SYMMETRIC PROJECTION THEOREM).** *Let  $B$  be a closed and convex nonnegative type-1 symmetric set. Let  $\mathbf{x} \in \mathbb{R}^n$  and  $\sigma \in \tilde{\Sigma}(\mathbf{x})$ . Then, there exists  $\mathbf{y} \in P_{C_s \cap B}(\mathbf{x})$  for which  $S_{[1, s]}^{\sigma}$  is a super support.*

**PROOF.** Since  $P_{C_s \cap B}(\mathbf{x})$  is nonempty, it follows by the order preservation property (Theorem 3.1) that there exists  $\mathbf{y} \in P_{C_s \cap B}(\mathbf{x})$  such that  $\sigma \in \tilde{\Sigma}(\mathbf{y})$ . Therefore, since  $\mathbf{y} \geq \mathbf{0}$  and  $|I_1(\mathbf{y})| \leq s$ , it follows that  $I_1(\mathbf{y}) \subseteq S_{[1, s]}^{\sigma}$ .  $\square$

Using Theorem 4.2, we can write explicitly the algorithm for finding a sparse projection vector onto a closed and convex nonnegative type-1 symmetric set.

---

**Algorithm 2** (Projection onto a nonnegative type-1 symmetric set).

**Input:**  $\mathbf{x} \in \mathbb{R}^n$ .

**Output:**  $\mathbf{u} \in P_{B \cap C_s}(\mathbf{x})$ .

1. Compute  $T = S_{[1, s]}^{\sigma}$  for  $\sigma \in \tilde{\Sigma}(\mathbf{x})$ .
  2. Return  $\mathbf{u} = \mathbf{U}_T P_{B_T}(\mathbf{x}_T)$ .
- 

**4.4. Sparse projection onto type-2 symmetric sets.** When the underlying set  $B$  is a type-2 symmetric set, a super support of a sparse projection vector onto  $B$  can also be instantly detected—the set of indices corresponding to  $s$  indices with the largest *absolute value*.

**THEOREM 4.3 (TYPE-2 SYMMETRIC PROJECTION THEOREM).** *Let  $B$  be a closed and convex type-2 symmetric set and for  $\mathbf{x} \in \mathbb{R}^n$ , let  $\sigma \in \tilde{\Sigma}(|\mathbf{x}|)$ . Then, there exists  $\mathbf{y} \in P_{C_s \cap B}(\mathbf{x})$  for which  $S_{[1, s]}^{\sigma}$  is a super support.*

**PROOF.** Since  $B \cap \mathbb{R}_+^n$  is obviously a nonnegative type-1 symmetric set, it follows by Theorem 4.2 that there exists  $\mathbf{z} \in P_{C_s \cap B \cap \mathbb{R}_+^n}(|\mathbf{x}|)$  such that  $I_1(\mathbf{z}) \subseteq S_{[1, s]}^{\sigma}$ . By Lemma 3.3, it follows that  $\operatorname{sign}(\mathbf{x}) \circ \mathbf{z} \in P_{C_s \cap B}(\mathbf{x})$ , and hence taking  $\mathbf{y} = \operatorname{sign}(\mathbf{x}) \circ \mathbf{z}$ , we obtain that  $\mathbf{y} \in P_{C_s \cap B}(\mathbf{x})$  and that  $I_1(\mathbf{y}) = I_1(\mathbf{z}) \subseteq S_{[1, s]}^{\sigma}$ .  $\square$

The corresponding algorithm for computing a sparse projection vector onto a type-2 symmetric set is stated explicitly below.

**Algorithm 3** (Projection onto a type-2 symmetric set).

**Input:**  $\mathbf{x} \in \mathbb{R}^n$ .

**Output:**  $\mathbf{u} \in P_{C_s \cap B}(\mathbf{x})$ .

1. Compute  $T = S_{[1, s]}^\sigma$  for  $\sigma \in \tilde{\Sigma}(|\mathbf{x}|)$ .
2. Return  $\mathbf{u} = \mathbf{U}_T P_{B_T}(\mathbf{x}_T)$ .

**4.5. Unifying the analysis.** Since we will investigate nonnegative type-1 and type-2 symmetric sets, we would like to unify the analysis of the two settings as much as possible. This is done by defining the following *symmetry function*  $p: \mathbb{R}^n \rightarrow \mathbb{R}^n$ :

$$p(\mathbf{x}) \equiv \begin{cases} \mathbf{x}, & B \text{ is nonnegative type-1,} \\ |\mathbf{x}|, & B \text{ is type-2 symmetric.} \end{cases} \quad (19)$$

Using the definition of the symmetry function, we can combine Theorems 4.2 and 4.3, to obtain the following result.

**THEOREM 4.4 (UNIFIED SYMMETRIC PROJECTION THEOREM).** *Let  $B \subseteq \mathbb{R}^n$  be a closed and convex set, which is additionally either a nonnegative type-1 or type-2 symmetric set. Let  $\mathbf{x} \in \mathbb{R}^n$  and  $\sigma \in \tilde{\Sigma}(p(\mathbf{x}))$ , where  $p(\cdot)$  is defined in (19). Then, there exists  $\mathbf{y} \in P_{C_s \cap B}(\mathbf{x})$  for which  $S_{[1, s]}^\sigma$  is a super support.*

**4.6. Examples.** Using the results obtained so far, we can now go back to the symmetric sets considered in §3.1 and write in details how to compute a sparse projection vector onto each of them. This is done in Table 3.

Some comments related to Table 3 are in order:

- When there are  $s + 1$  candidates for the sparse projection vector  $\mathbf{z}_k = \mathbf{U}_{T_k} P_{B_{T_k}}(\mathbf{x}_{T_k})$ ,  $k = 0, 1, \dots, s$ , the sparse projection vector is chosen to be the one for which  $\|\mathbf{z}_k - \mathbf{x}\|^2$  is minimal.
- Performing the orthogonal projection onto the unit-simplex  $\Delta_s$  amounts to finding a root of a one-dimensional strictly decreasing function (see, e.g., Kiwiel [17]).
- The projection of a vector  $\mathbf{y} \in \mathbb{R}^s$  onto the unit-sum set  $\Delta'_s$  is given by

$$P_{\Delta'_s}(\mathbf{y}) = \mathbf{y} + \lambda \mathbf{1}_s,$$

where  $\lambda = (1 - \mathbf{1}_s^T \mathbf{y})/s$ .

- The projection onto the unit  $l_p$ -ball can be done via a one-dimensional root finding procedure.
- For general  $\ell, u$ , the set  $[\ell, u]^n$  is a type-1 symmetric set, and hence  $s + 1$  possible super sets should be explored. However, if  $\ell \geq 0$ , then the set is, in addition, nonnegative and hence a super support set is  $T = S_{[1, s]}^\sigma$ , where  $\sigma \in \tilde{\Sigma}(\mathbf{x})$ . If  $\ell = -u$ , then the set is a type-2 symmetric set and hence a super support set is  $T = S_{[1, s]}^\sigma$ , where  $\sigma \in \tilde{\Sigma}(|\mathbf{x}|)$ .

We also note that in some cases, the search of the correct super support among the  $s + 1$  possibilities can be done more efficiently than just performing  $s + 1$  projections on the restrictions of the underlying set  $B$ . As an illustration, let us consider the sparse projection onto the unit-sum set. A naive implementation of the procedure will consist of the computation of the  $s + 1$  vectors

$$\mathbf{v}_k = \mathbf{U}_{T_k}(\mathbf{x}_{T_k} + \lambda_k \mathbf{1}_s),$$

TABLE 3. Super supports for sparse projection onto simple sets.

$B$	Candidates for sparse projection vectors	Super support set(s)	Restriction of $B$ on the support
$\mathbb{R}^n$	$\mathbf{U}_T \mathbf{x}_T$	$T = S_{[1, s]}^\sigma$ , $\sigma \in \tilde{\Sigma}( \mathbf{x} )$	$B_T = \mathbb{R}^s$
$\mathbb{R}_+^n$	$\mathbf{U}_T [\mathbf{x}_T]_+$	$T = S_{[1, s]}^\sigma$ , $\sigma \in \tilde{\Sigma}(\mathbf{x})$	$B_T = \mathbb{R}_+^s$
$\Delta_n$	$\mathbf{U}_T P_{B_T}(\mathbf{x}_T)$	$T = S_{[1, s]}^\sigma$ , $\sigma \in \tilde{\Sigma}(\mathbf{x})$	$B_T = \Delta_s$
$\Delta'_n$	$\mathbf{U}_{T_k} P_{B_{T_k}}(\mathbf{x}_{T_k})$	$T_k = S_{[1, k]}^\sigma \cup S_{[n+k-(s-1), n]}^\sigma$ $k = 0, 1, \dots, s$ , $\sigma \in \tilde{\Sigma}(\mathbf{x})$	$B_{T_k} = \Delta'_s$
$B_p^n[\mathbf{0}, 1]$ ( $p \geq 1$ )	$\mathbf{U}_T P_{B_T}(\mathbf{x}_T)$	$T = S_{[1, s]}^\sigma$ , $\sigma \in \tilde{\Sigma}( \mathbf{x} )$	$B_T = B_p^s[\mathbf{0}, 1]$
$[\ell, u]^n$ ( $\ell < u$ )	$\mathbf{U}_{T_k} P_{B_{T_k}}(\mathbf{x}_{T_k})$	$T_k = S_{[1, k]}^\sigma \cup S_{[n+k-(s-1), n]}^\sigma$ $k = 0, 1, \dots, s$ , $\sigma \in \tilde{\Sigma}(\mathbf{x})$	$B_{T_k} = [\ell, u]^s$

Downloaded from informs.org by [132.68.247.241] on 27 October 2015, at 05:10. For personal use only, all rights reserved.

where

$$\lambda_k = \frac{1 - \mathbf{1}_s^T \mathbf{x}_{T_k}}{s},$$

and  $T_k = S_{[1, k]}^\sigma \cup S_{[n+k-(s-1), n]}^\sigma$ ,  $k = 0, 1, \dots, s$ . The sparse projection vector is  $\mathbf{v}_k$  with  $k$  chosen to minimize the expression

$$f_k \equiv \|\mathbf{x} - \mathbf{v}_k\|^2 = s\lambda_k^2 + \|\mathbf{x}_{T_k^c}\|^2.$$

Since the evaluation of  $f_k$  requires  $O(n)$  operations, the described process needs  $O(ns)$  operations. We can decrease the number of operations by noting that  $\lambda_k$  and  $f_k$  can be computed recursively. Indeed, for any  $k = 1, 2, \dots, s$ ,

$$\lambda_k = \frac{1}{s} \left( 1 - \sum_{j=1}^k x_{\sigma(j)} - \sum_{j=n+k+1-s}^n x_{\sigma(j)} \right).$$

Thus

$$\lambda_k = \frac{1}{s} \left( 1 - x_{\sigma(k)} + x_{\sigma(n+k-s)} - \sum_{j=1}^{k-1} x_{\sigma(j)} - \sum_{j=n+k-s}^n x_{\sigma(j)} \right) = \frac{1}{s} (-x_{\sigma(k)} + x_{\sigma(n+k-s)}) + \lambda_{k-1}.$$

Therefore the recurrence relation for the sequence  $\{\lambda_k\}$  is

$$\lambda_{k-1} = \lambda_k + \frac{1}{s} (x_{\sigma(k)} - x_{\sigma(n+k-s)}).$$

A similar argument shows how to compute  $f_{k-1}$  out of  $f_k$ :

$$\begin{aligned} f_{k-1} &= s\lambda_{k-1}^2 + \sum_{j=k}^{n+k-1+s} x_{\sigma(j)}^2 \\ &= s\lambda_k^2 + \sum_{j=k+1}^{n+k-s} x_{\sigma(j)}^2 + x_{\sigma(k)}^2 - x_{\sigma(n+k-s)}^2 + s(\lambda_{k-1}^2 - \lambda_k^2) \\ &= f_k + s(\lambda_{k-1}^2 - \lambda_k^2) + x_{\sigma(k)}^2 - x_{\sigma(n+k-s)}^2. \end{aligned}$$

The recurrence-based algorithm for computing a sparse projection vector on the unit-sum set is described below.

**Algorithm 4** (Projection onto the sparse unit-sum set)

**Input:**  $\mathbf{x} \in \mathbb{R}^n$ .

**Output:**  $\mathbf{u} \in P_{C_s \cap \Delta_n}(\mathbf{x})$ .

1. Compute  $S_{[1, s]}^\sigma$  and  $S_{[n-(s-1), n]}^\sigma$  for  $\sigma \in \tilde{\Sigma}(\mathbf{x})$ .
2. Sort the elements of  $\mathbf{x}$  corresponding to the indices  $S_{[1, s]}^\sigma \cup S_{[n-(s-1), n]}^\sigma$ .
3. Set  $\lambda_s = \frac{1}{s} \left( 1 - \sum_{j=1}^s x_{\sigma(j)} \right)$ .
4.  $f_s = s\lambda_s^2 + \sum_{j=s+1}^n x_{\sigma(j)}^2$ .
5. For any  $k = s, s-1, \dots, 1$  compute the following:
  - (a)  $\lambda_{k-1} = \lambda_k + \frac{1}{s} (x_{\sigma(k)} - x_{\sigma(n+k-s+1)})$ .
  - (b)  $f_{k-1} = f_k + s(\lambda_{k-1}^2 - \lambda_k^2) + x_{\sigma(k)}^2 - x_{\sigma(n+k-s)}^2$ .
6.  $m \in \underset{k=0, 1, \dots, s}{\operatorname{argmin}} \{f_k\}$
7. Return

$$u_i = \begin{cases} x_i + \lambda_m, & i \in S_{[1, m]}^\sigma \cup S_{[n+m-(s-1), n]}^\sigma, \\ 0, & \text{else.} \end{cases}$$

Overall, the number of operations required in the above implementation is  $o(n + s \log(s))$  for the initial computation and sorting of  $S_{[1, s]}^\sigma \cup S_{[n-(s-1), n]}^\sigma$ , and  $s \cdot O(1)$  operations for the evaluation of the  $s$  values  $f_0, f_1, \dots, f_s$ , resulting with a total (reduced) complexity of  $O(n + s \log(s))$ .

**5. Optimality conditions I: Stationarity-based conditions.** So far, we have studied properties and computational methods of orthogonal projections onto symmetric sets and sparse symmetric sets (§§3 and 4, respectively). The importance of this study is twofold: first, some algorithms that attempt to find an optimal solution of problem (1) actually require the computation of the orthogonal projection (e.g., the IHT method—see §8). Second, the orthogonal projection plays a key role in the development of necessary optimality conditions for sparsity-constrained problems. In particular, in this section, two of the optimality conditions that we will consider—basic feasibility (BF) and  $L$ -stationarity—heavily rely on orthogonal projections.

**5.1. Basic feasibility (BF).** The most elementary stationarity-based optimality condition is *basic feasibility*. For problem (1), loosely speaking, it states that a basic feasible point is a point that satisfies the first-order optimality conditions over any possible super support set.

**DEFINITION 5.1 (BASIC FEASIBILITY).** A vector  $\mathbf{x} \in C_S \cap B$  is called a *basic feasible (BF) point* of (P) if for any super support set  $S$  of  $\mathbf{x}$ , it holds that for some  $L > 0$ :

$$\mathbf{x}_S = P_{B_S} \left( \mathbf{x}_S - \frac{1}{L} \nabla_S f(\mathbf{x}) \right). \quad (20)$$

**REMARK 5.1.** We note the following:

(a) If  $|I_1(\mathbf{x})| = s$ , then the only super support set is the support itself, and hence basic feasibility is the same as the condition

$$\mathbf{x}_{I_1(\mathbf{x})} = P_{B_{I_1(\mathbf{x})}} \left( \mathbf{x}_{I_1(\mathbf{x})} - \frac{1}{L} \nabla_{I_1(\mathbf{x})} f(\mathbf{x}) \right). \quad (21)$$

The above condition is always satisfied for BF points (but is not sufficient when the support is incomplete).

(b) The BF condition is equivalent to the condition that for any super support set  $S$  of  $\mathbf{x}$ ,  $\mathbf{x}_S$  is a stationary point of the convex-constrained problem

$$\min \{f(\mathbf{U}_S \mathbf{d}) : \mathbf{d} \in B_S\}.$$

The stationarity condition (20) is actually independent of  $L$ , although it is expressed in terms of  $L$ , and it can also be written alternatively as (see §2.1)

$$\langle \nabla_S f(\mathbf{x}), \mathbf{y}_S - \mathbf{x}_S \rangle \geq 0 \quad \text{for any } \mathbf{y} \in B \text{ s.t. } I_1(\mathbf{y}) \subseteq S.$$

The fact that BF is a necessary optimality condition is shown next.

**THEOREM 5.1.** *Let  $\mathbf{x}^*$  be an optimal solution of problem (1). Then,  $\mathbf{x}^*$  is a basic feasible point of (P).*

**PROOF.** Let  $S$  be a super support set of  $\mathbf{x}^*$ . Since  $\mathbf{x}^*$  is an optimal solution of (P), we have, in particular, that  $\mathbf{x}_S^*$  is an optimal solution of

$$\min_{\mathbf{d} \in B_S} f(\mathbf{U}_S \mathbf{d}), \quad (22)$$

and thus,  $\mathbf{x}_S^*$  is a stationary point of (22). We conclude that for any  $L > 0$ ,

$$\mathbf{x}_S^* = P_{B_S} \left( \mathbf{x}_S^* - \frac{1}{L} \mathbf{U}_S^T \nabla f(\mathbf{U}_S \mathbf{x}_S^*) \right).$$

By noting the equality  $\mathbf{U}_S^T \nabla f(\mathbf{U}_S \mathbf{x}_S^*) = \nabla_S f(\mathbf{x}^*)$ , we conclude that  $\mathbf{x}^*$  is a basic feasible point.  $\square$

Note that when the support of  $\mathbf{x}$  is not full, verifying whether it is a basic feasible point requires in principle checking the condition (20) for each of the  $\binom{n - \|\mathbf{x}\|_0}{s - \|\mathbf{x}\|_0}$  choices of the super support set. We will see in §5.2.3 that when the underlying set  $B$  is either a nonnegative type-1 symmetric set or a type-2 symmetric set, there are simple ways to verify that a point without a full support is a basic feasible point by checking that condition (20) holds for a specific super support set. In the meantime, using the explicit stationarity conditions for the examples that appear in Table 1, we can write the BF conditions for the full support case. For all these examples, when

Downloaded from informs.org by [132.68.247.241] on 27 October 2015, at 05:10. For personal use only, all rights reserved.

the support is not full, the condition for BF is exactly the stationarity condition for the problem of minimizing  $f$  over  $B$  (without the sparsity constraint).

$B$	BF conditions (full support)
$\mathbb{R}^n$	$\nabla_{I_1(\mathbf{x}^*)}f(\mathbf{x}^*) = \mathbf{0}$
$\mathbb{R}_+^n$	$\nabla_{I_1(\mathbf{x}^*)}f(\mathbf{x}^*) = \mathbf{0}$
$\Delta_n$	$\exists \mu \in \mathbb{R}: \nabla_i f(\mathbf{x}^*) = \mu, i \in I_1(\mathbf{x}^*)$
$\Delta'_n$	$\exists \mu \in \mathbb{R}: \nabla_i f(\mathbf{x}^*) = \mu, i \in I_1(\mathbf{x}^*)$
$B_2^n[\mathbf{0}, 1]$	$\nabla_{I_1(\mathbf{x}^*)}f(\mathbf{x}^*) = \mathbf{0}$ or $\ \mathbf{x}^*\  = 1$ and $\exists \lambda \leq 0: \nabla_{I_1(\mathbf{x}^*)}f(\mathbf{x}^*) = \lambda \mathbf{x}_{I_1(\mathbf{x}^*)}^*$
$[\ell, u]^n$ ( $\ell < u$ )	$\frac{\partial f}{\partial x_i}(\mathbf{x}^*) \begin{cases} = 0, & \ell < x_i < u, \\ \geq 0, & x_i = \ell, \\ \leq 0, & x_i = u_i, \end{cases} \quad i \in I_1(\mathbf{x}^*).$

We end this section with a general property of basic feasible points that will be useful later on. This property holds when  $B$  is a type-2 symmetric set.

LEMMA 5.1. *Suppose that  $B$  is a type-2 symmetric set, and let  $\mathbf{x}^*$  be a basic feasible solution of (P). Then,*

$$x_i^* \nabla_i f(\mathbf{x}^*) \leq 0 \quad \text{for all } i \in I_1(\mathbf{x}^*).$$

PROOF. Since  $\mathbf{x}^*$  satisfies (21) with  $\mathbf{x} = \mathbf{x}^*$ , it is a stationary point of the problem

$$\min_{\mathbf{d} \in B_S} f(\mathbf{U}_S \mathbf{d}),$$

where  $S = I_1(\mathbf{x}^*)$ . Thus

$$\mathbf{U}_S^T \nabla f(\mathbf{U}_S \mathbf{x}_S^*)^T (\mathbf{y}_S - \mathbf{x}_S^*) \geq 0 \quad \text{for any } \mathbf{y} \in B \text{ s.t. } I_1(\mathbf{y}) \subseteq S. \tag{23}$$

Since  $B$  is type-2 symmetric, it follows that the vector  $\tilde{\mathbf{x}}$  defined by

$$\tilde{x}_j = \begin{cases} x_j^*, & j \neq i, \\ -x_i^*, & j = i. \end{cases}$$

is in  $B$  and obviously  $I_1(\tilde{\mathbf{x}}) = I_1(\mathbf{x}^*)$ . Plugging  $\mathbf{y} = \tilde{\mathbf{x}}$  into (23) yields the inequality  $\nabla_i f(\mathbf{x}^*) x_i^* \leq 0$ .  $\square$

## 5.2. $L$ -stationarity.

**5.2.1. Definition and basic properties.** As was already noted, BF is a notion related to stationarity over a restriction of  $B$  to super support sets of the vector. It does not say anything about the ‘‘optimality’’ of the support, and in that respect, it is a rather weak condition. A stronger condition is the  $L$ -stationarity condition that we introduce now.

DEFINITION 5.2 ( $L$ -STATIONARITY). Let  $L > 0$ . A vector  $\mathbf{x} \in C_S \cap B$  is an  $L$ -stationary point of (P) if

$$\mathbf{x} \in P_{C_S \cap B} \left( \mathbf{x} - \frac{1}{L} \nabla f(\mathbf{x}) \right).$$

The fact that  $L$ -stationarity is a more restrictive condition than BF is now stated and proved.

LEMMA 5.2. *Let  $\mathbf{x}^* \in C_S \cap B$  be an  $L$ -stationary point of (P). Then,  $\mathbf{x}^*$  is a basic feasible point of (P).*

PROOF. Let  $S$  be a super support of  $\mathbf{x}$ . Plugging  $T = S$ ,  $\mathbf{x} = \mathbf{x}^* - (1/L)\nabla f(\mathbf{x}^*)$  and  $\mathbf{y} = \mathbf{x}^*$  in Lemma 4.1, implies that

$$\mathbf{x}_S^* = P_{B_S} \left( \mathbf{x}_S^* - \frac{1}{L} \nabla_S f(\mathbf{x}^*) \right)$$

showing the required result.  $\square$

Downloaded from informs.org by [132.68.247.241] on 27 October 2015, at 05:10. For personal use only, all rights reserved.

We have thus shown that any  $L$ -stationary point is a basic feasible point. We continue to show that  $L$ -stationarity is a more restrictive condition as  $L$  becomes smaller. For that, we use the following trivial fact.

LEMMA 5.3. *Let  $L > 0$ . Then,  $\mathbf{x} \in C_s \cap B$  is an  $L$ -stationary point of (P) if and only if*

$$\mathbf{x} \in \operatorname{argmin}_{\mathbf{y} \in C_s \cap B} \left\{ h_L(\mathbf{y}, \mathbf{x}) \equiv \nabla f(\mathbf{x})^T \mathbf{y} + \frac{L}{2} \|\mathbf{y} - \mathbf{x}\|^2 \right\}. \quad (24)$$

PROOF. Note that

$$\left\| \mathbf{y} - \left( \mathbf{x} - \frac{1}{L} \nabla f(\mathbf{x}) \right) \right\|^2 = \left( \frac{1}{L^2} \|\nabla f(\mathbf{x})\|^2 - \frac{2}{L} \nabla f(\mathbf{x})^T \mathbf{x} \right) + \frac{2}{L} \left[ \nabla f(\mathbf{x})^T \mathbf{y} + \frac{L}{2} \|\mathbf{y} - \mathbf{x}\|^2 \right].$$

Therefore  $\mathbf{x}$  is an  $L$ -stationary point, meaning that  $\mathbf{x}$  satisfies

$$\mathbf{x} \in \operatorname{argmin}_{\mathbf{y} \in C_s \cap B} \left\| \mathbf{y} - \left( \mathbf{x} - \frac{1}{L} \nabla f(\mathbf{x}) \right) \right\|^2$$

if and only if (24) holds.  $\square$

THEOREM 5.2. *Suppose that  $L_1 \geq L_2 \geq 0$ . Then, if  $\mathbf{x} \in C_s \cap B$  is an  $L_2$ -stationary point, then it is also an  $L_1$ -stationary point.*

PROOF. Since  $\mathbf{x}$  is an  $L_2$ -stationary point, then by Lemma 5.3, it satisfies

$$\mathbf{x} \in \operatorname{argmin}_{\mathbf{y} \in C_s \cap B} h_{L_2}(\mathbf{y}, \mathbf{x}),$$

which means that  $h_{L_2}(\mathbf{y}, \mathbf{x}) \geq h_{L_2}(\mathbf{x}, \mathbf{x})$  for any  $\mathbf{y} \in C_s \cap B$ , and hence for any such  $\mathbf{y}$ :

$$h_{L_1}(\mathbf{y}, \mathbf{x}) = h_{L_2}(\mathbf{y}, \mathbf{x}) + \frac{L_1 - L_2}{2} \|\mathbf{y} - \mathbf{x}\|^2 \geq h_{L_2}(\mathbf{x}, \mathbf{x}) + 0 = h_{L_1}(\mathbf{x}, \mathbf{x})$$

showing that  $\mathbf{x} \in \operatorname{argmin}_{\mathbf{y} \in C_s \cap B} h_{L_1}(\mathbf{y}, \mathbf{x})$ , which by Lemma 5.3, implies that  $\mathbf{x}$  is an  $L_1$ -stationary point.  $\square$

Another result that can be proved without any symmetry assumptions on  $B$  is that if  $\nabla f$  has a Lipschitz constant  $L(f)$ , then  $L$ -stationarity is a necessary optimality condition for any  $L > L(f)$ . Later on, in §3, we will show how the result can be improved, that is, can be shown for smaller values of  $L$  when symmetry conditions are assumed.

THEOREM 5.3. *Suppose that Assumption [A+] is satisfied. Then, if  $\mathbf{x}^*$  is an optimal solution of (P), then it is an  $L$ -stationary point for any  $L > L(f)$ .*

PROOF. Let  $\mathbf{x}^*$  be an optimal solution of (P), and let  $L$  satisfy  $L > L(f)$ . By the descent lemma (Lemma 2.1), it follows that for any  $\mathbf{y} \in C_s \cap B$ , the following inequality holds:

$$f(\mathbf{y}) \leq f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^T (\mathbf{y} - \mathbf{x}^*) + \frac{L}{2} \|\mathbf{y} - \mathbf{x}^*\|^2,$$

which by the fact that  $f(\mathbf{y}) \geq f(\mathbf{x}^*)$ , implies that for any  $\mathbf{y} \in C_s \cap B$ :

$$g(\mathbf{y}) \equiv \nabla f(\mathbf{x}^*)^T (\mathbf{y} - \mathbf{x}^*) + \frac{L}{2} \|\mathbf{y} - \mathbf{x}^*\|^2 \geq 0.$$

Therefore  $g(\mathbf{y}) \geq g(\mathbf{x}^*)$  for any  $\mathbf{y} \in C_s \cap B$ , and hence

$$\mathbf{x}^* \in \operatorname{argmin}_{\mathbf{y} \in C_s \cap B} \left\{ \nabla f(\mathbf{x}^*)^T (\mathbf{y} - \mathbf{x}^*) + \frac{L}{2} \|\mathbf{y} - \mathbf{x}^*\|^2 \right\} = \operatorname{argmin}_{\mathbf{y} \in C_s \cap B} h_L(\mathbf{y}, \mathbf{x}^*).$$

Finally, by Lemma 5.3, we conclude that  $\mathbf{x}^*$  is an  $L$ -stationary point of (P).  $\square$

**5.2.2. L-stationarity under symmetry assumptions.** We now continue to write more explicit conditions for L-stationarity under the assumption that the underlying set is either nonnegative type-1 or type-2 symmetric.

**THEOREM 5.4 (L-STATIONARITY CHARACTERIZATION).** *Let  $B$  be either a nonnegative type-1 or a type-2 symmetric set. A vector  $\mathbf{x}^* \in C_s \cap B$  is an L-stationary point of  $(P)$  if and only if*

$$\mathbf{x}^* \text{ is a BF point and } p(Lx_i^* - \nabla_i f(\mathbf{x}^*)) \geq p(-\nabla_j f(\mathbf{x}^*)) \text{ for any } i \in I_1(\mathbf{x}^*) \text{ and } j \in I_0(\mathbf{x}^*),$$

where  $p(\cdot)$  is given by

$$p(\mathbf{x}) \equiv \begin{cases} \mathbf{x} & B \text{ is nonnegative type-1,} \\ |\mathbf{x}| & B \text{ is type-2 symmetric.} \end{cases}$$

**PROOF.** Assume that  $\mathbf{x}^* \in C_s \cap B$  is an L-stationary point, that is,  $\mathbf{x}^* \in P_{C_s \cap B}(\mathbf{x}^* - (1/L)\nabla f(\mathbf{x}^*))$ . By Lemma 5.2, it is also a basic feasible point. In addition,

$$p(\mathbf{x}^*) \in P_{C_s \cap B \cap \mathbb{R}_+^n} \left( p \left( \mathbf{x}^* - \frac{1}{L} \nabla f(\mathbf{x}^*) \right) \right). \tag{25}$$

Relation (25) is valid since when  $p(\mathbf{x}^*) \equiv \mathbf{x}^*$ , it is actually a tautology, and when  $p(\mathbf{x}^*) = |\mathbf{x}^*|$ , it holds by Corollary 3.1. Now, let  $i \in I_1(\mathbf{x}^*)$  and  $j \in I_0(\mathbf{x}^*)$ . By (25) and Lemma 3.1, we have

$$\left( p \left( x_i^* - \frac{1}{L} \nabla_i f(\mathbf{x}^*) \right) - p \left( x_j^* - \frac{1}{L} \nabla_j f(\mathbf{x}^*) \right) \right) (p(x_i^*) - p(x_j^*)) \geq 0,$$

which by the fact that  $p(x_i^*) > p(x_j^*)$  ( $i \in I_1(\mathbf{x}^*), j \in I_0(\mathbf{x}^*)$ ), implies that

$$p \left( x_i^* - \frac{1}{L} \nabla_i f(\mathbf{x}^*) \right) \geq p \left( x_j^* - \frac{1}{L} \nabla_j f(\mathbf{x}^*) \right).$$

Hence, since  $x_j^* = 0$ , we have that  $p(Lx_i^* - \nabla_i f(\mathbf{x}^*)) \geq p(-\nabla_j f(\mathbf{x}^*))$ .

To prove the reverse direction, assume that  $\mathbf{x}^*$  is a basic feasible point, and that for any  $i \in I_1(\mathbf{x}^*)$  and  $j \in I_0(\mathbf{x}^*)$ , the inequality

$$p(Lx_i^* - \nabla_i f(\mathbf{x}^*)) \geq p(-\nabla_j f(\mathbf{x}^*)) \tag{26}$$

holds. Since  $x_j^* = 0$  ( $j \in I_0(\mathbf{x}^*)$ ), we have that

$$p \left( x_i^* - \frac{1}{L} \nabla_i f(\mathbf{x}^*) \right) \geq p \left( x_j^* - \frac{1}{L} \nabla_j f(\mathbf{x}^*) \right)$$

for any  $i \in I_1(\mathbf{x}^*)$  and  $j \in I_0(\mathbf{x}^*)$ . Therefore, there exists a sorting permutation

$$\sigma \in \tilde{\Sigma} \left( p \left[ \mathbf{x}^* - \frac{1}{L} \nabla f(\mathbf{x}^*) \right] \right)$$

for which  $I_1(\mathbf{x}^*) \subseteq S_{[1,s]}^\sigma$ . By Theorem 4.4, there exists  $\mathbf{z} \in P_{C_s \cap B}(\mathbf{x}^* - (1/L)\nabla f(\mathbf{x}^*))$  such that  $I_1(\mathbf{z}) \subseteq S_{[1,s]}^\sigma$ . Since  $\mathbf{z} \in P_{C_s \cap B}(\mathbf{x}^* - (1/L)\nabla f(\mathbf{x}^*))$ , by Lemma 4.1,

$$\mathbf{z}_{S_{[1,s]}^\sigma} = P_{B_{S_{[1,s]}^\sigma}} \left( p \left[ \mathbf{x}_{S_{[1,s]}^\sigma}^* - \frac{1}{L} \nabla_{S_{[1,s]}^\sigma} f(\mathbf{x}^*) \right] \right).$$

Since  $\mathbf{x}^*$  is a basic feasible point of  $(P)$ , and  $S_{[1,s]}^\sigma$  is a super support of  $\mathbf{x}^*$ , then

$$\mathbf{x}_{S_{[1,s]}^\sigma}^* = P_{B_{S_{[1,s]}^\sigma}} \left( p \left[ \mathbf{x}_{S_{[1,s]}^\sigma}^* - \frac{1}{L} \nabla_{S_{[1,s]}^\sigma} f(\mathbf{x}^*) \right] \right).$$

Therefore, by the uniqueness of the projection operator onto closed and convex sets (in particular here, on  $B_{S_{[1,s]}^\sigma}$ ), and since  $I_1(\mathbf{x}^*) \cup I_1(\mathbf{z}) \subseteq S_{[1,s]}^\sigma$ , it follows that  $\mathbf{z} = \mathbf{x}^*$ , which is the desired result.  $\square$

**REMARK 5.2.** When  $B = \mathbb{R}^n$ , it can be shown that the conditions of Theorem 5.4 reduce to

$$|\nabla_i f(\mathbf{x}^*)| \begin{cases} \leq LM_s(\mathbf{x}^*), & \text{if } i \in I_0(\mathbf{x}^*), \\ = 0, & \text{if } i \in I_1(\mathbf{x}^*), \end{cases}$$

where  $M_s(\mathbf{x}^*)$  is the  $s$ -th largest absolute value component in  $\mathbf{x}^*$ . This result was established in Beck and Eldar [3] for general objective functions, and in Blumensath and Davies [6] for the case of a least squares objective function.

Downloaded from informs.org by [132.68.247.241] on 27 October 2015, at 05:10. For personal use only, all rights reserved.

**5.2.3. Characterization of BF points with incomplete support.** In Lemma 5.2, we showed that an  $L$ -stationary point is necessarily a basic feasible point. We will now show that the reverse implication also holds if the point does not have a full support and  $B$  is either a nonnegative type-1 or a type-2 symmetric set. Therefore, checking BF of a point in  $C_s \cap B$  is equivalent to checking  $L$ -stationarity, and in any case, there is no need to go over all the possible super support sets. Before proving the result, we will establish a technical lemma.

LEMMA 5.4. *Let  $B$  be either a nonnegative type-1 or a type-2 symmetric set and let  $\mathbf{x} \in C_s \cap B$  satisfy  $|I_1(\mathbf{x})| < s$ . Let  $S$  be a super support set of  $\mathbf{x}$ , and assume that for some  $L > 0$ ,*

$$\mathbf{x}_S = P_{B_S} \left( \mathbf{x}_S - \frac{1}{L} \nabla_S f(\mathbf{x}) \right). \quad (27)$$

Then,

$$p \left( x_i - \frac{1}{L} \nabla_i f(\mathbf{x}) \right) \geq p \left( x_j - \frac{1}{L} \nabla_j f(\mathbf{x}) \right), \quad \text{for any } i \in I_1(\mathbf{x}), \quad j \in S \cap I_0(\mathbf{x}).$$

PROOF. Since  $\mathbf{x}$  satisfies (27), it is a stationary point of the problem

$$\min_{\mathbf{d} \in B_S} f(\mathbf{U}_S \mathbf{d})$$

. Thus

$$\mathbf{U}_S^T \nabla f(\mathbf{U}_S \mathbf{x}_S)^T (\mathbf{y}_S - \mathbf{x}_S) \geq 0, \quad \text{for any } \mathbf{y} \in B \text{ s.t. } I_1(\mathbf{y}) \subseteq S,$$

which is the same as

$$\nabla_S f(\mathbf{x})^T (\mathbf{y}_S - \mathbf{x}_S) \geq 0, \quad \text{for any } \mathbf{y} \in B \text{ s.t. } I_1(\mathbf{y}) \subseteq S. \quad (28)$$

Let  $i \in I_1(\mathbf{x})$  and  $j \in I_0(\mathbf{x}) \cap S$ . Since  $B$  is type-1 symmetric, we have that  $\mathbf{x}^{\tau_{i,j}} \in B$ , where  $\tau_{i,j}$  is the swap permutation of the indices  $i, j$ . In addition, since  $i, j \in S$ , it follows that  $I_1(\mathbf{x}^{\tau_{i,j}}) \subseteq S$ , and hence we can plug  $\mathbf{y} = \mathbf{x}^{\tau_{i,j}}$  into the inequality in (28) and obtain

$$x_i (\nabla_j f(\mathbf{x}) - \nabla_i f(\mathbf{x})) \geq 0. \quad (29)$$

At this point, we split the analysis between the nonnegative type-1 and type-2 settings.

- $B$  is nonnegative type-1 symmetric. Since  $i \in I_1(\mathbf{x})$ , we have  $x_i > 0$ , and hence by (29),

$$\nabla_j f(\mathbf{x}) \geq \nabla_i f(\mathbf{x}), \quad (30)$$

which combined with the fact  $x_i > 0$  ( $i \in I_1(\mathbf{x})$ ) and  $x_j = 0$  ( $j \in I_0(\mathbf{x})$ ) implies the following relation:

$$x_i - \frac{1}{L} \nabla_i f(\mathbf{x}) > x_j - \frac{1}{L} \nabla_j f(\mathbf{x}).$$

- $B$  is type-2 symmetric. By Lemma 5.1, it follows that

$$\nabla_i f(\mathbf{x}) x_i \leq 0. \quad (31)$$

Plugging  $\mathbf{y} = -\mathbf{x}^{\tau_{i,j}}$  into inequality (28) results in

$$-x_i (\nabla_j f(\mathbf{x}) + \nabla_i f(\mathbf{x})) \geq 0. \quad (32)$$

Multiplying inequalities (29) and (32), implies that the following holds:

$$|\nabla_i f(\mathbf{x})| \geq |\nabla_j f(\mathbf{x})|, \quad (33)$$

which combined with (31), yields

$$\left| x_i - \frac{1}{L} \nabla_i f(\mathbf{x}) \right| \geq \frac{1}{L} |\nabla_j f(\mathbf{x})|.$$

Hence, using (33) and the fact that  $x_j = 0$ , we obtain that

$$\left| x_i - \frac{1}{L} \nabla_i f(\mathbf{x}) \right| \geq \left| x_j - \frac{1}{L} \nabla_j f(\mathbf{x}) \right|.$$

Combining the two cases, we conclude that

$$p\left(x_i - \frac{1}{L}\nabla_i f(\mathbf{x})\right) \geq p\left(x_j - \frac{1}{L}\nabla_j f(\mathbf{x})\right) \quad \text{for any } i \in I_1(\mathbf{x}), j \in S \cap I_0(\mathbf{x}),$$

which is the desired result.  $\square$

We are now ready to prove that BF and  $L$ -stationarity are equivalent when the vector has an incomplete support.

**THEOREM 5.5.** *Let  $B$  be either a nonnegative type-1 or type-2 symmetric set and  $\mathbf{x} \in C_s \cap B$  such that  $|I_1(\mathbf{x})| < s$ . Then, the following claims are equivalent:*

- (a)  $\mathbf{x}$  is a basic feasible point of  $(P)$ .
- (b)  $\mathbf{x}$  is an  $L$ -stationary point of  $(P)$  over  $C_s \cap B$  for any  $L > 0$ .

**PROOF.** The implication (b)  $\Rightarrow$  (a) follows by Lemma 5.2. To prove the reverse implication, assume that  $\mathbf{x}$  is a basic feasible point. Let  $i \in I_1(\mathbf{x})$ ,  $j \in I_0(\mathbf{x})$ , and let  $S$  be a super support set of  $\mathbf{x}$  for which  $j \in S$  (such a set exists since  $|I_1(\mathbf{x})| < s$ ). Then, employing Lemma 5.4, it follows that

$$p\left(x_i - \frac{1}{L}\nabla_i f(\mathbf{x})\right) \geq p\left(x_j - \frac{1}{L}\nabla_j f(\mathbf{x})\right).$$

Since the above inequality is valid for all  $i \in I_1(\mathbf{x})$ ,  $j \in I_0(\mathbf{x})$ , we conclude that there exists a sorting permutation

$$\sigma \in \tilde{\Sigma}\left(p\left[\mathbf{x} - \frac{1}{L}\nabla f(\mathbf{x})\right]\right)$$

for which  $I_1(\mathbf{x}) \subseteq S_{[1,s]}^\sigma$ . By Theorem 4.4, there exists

$$\mathbf{z} \in P_{C_s \cap B}\left(\mathbf{x} - \frac{1}{L}\nabla f(\mathbf{x})\right) \quad (34)$$

such that  $I_1(\mathbf{z}) \subseteq S_{[1,s]}^\sigma$ . Consequently,

$$I_1(\mathbf{x}) \cup I_1(\mathbf{z}) \subseteq S_{[1,s]}^\sigma. \quad (35)$$

Since  $\mathbf{z} \in P_{C_s \cap B}(\mathbf{x} - (1/L)\nabla f(\mathbf{x}))$  and  $I_1(\mathbf{z}) \subseteq S_{[1,s]}^\sigma$ , by Lemma 4.1,

$$\mathbf{z}_{S_{[1,s]}^\sigma} = P_{B_{S_{[1,s]}^\sigma}}\left(\mathbf{x}_{S_{[1,s]}^\sigma} - \frac{1}{L}\nabla_{S_{[1,s]}^\sigma} f(\mathbf{x})\right).$$

Since  $\mathbf{x}$  is a basic feasible point of  $(P)$ , and  $I_1(\mathbf{x}) \subseteq S_{[1,s]}^\sigma$ , then

$$\mathbf{x}_{S_{[1,s]}^\sigma} = P_{B_{S_{[1,s]}^\sigma}}\left(\mathbf{x}_{S_{[1,s]}^\sigma} - \frac{1}{L}\nabla_{S_{[1,s]}^\sigma} f(\mathbf{x})\right).$$

Finally, we have from the uniqueness of the projection operator onto closed and convex sets, that  $\mathbf{x}_{S_{[1,s]}^\sigma} = \mathbf{z}_{S_{[1,s]}^\sigma}$ , which combined with (35), implies that  $\mathbf{x} = \mathbf{z}$ , and hence (see (34)) that

$$\mathbf{x} \in P_{C_s \cap B}\left(\mathbf{x} - \frac{1}{L}\nabla f(\mathbf{x})\right). \quad \square$$

Theorem 5.5 states that when the support is incomplete, BF is equivalent to  $L$ -stationarity (for any  $L > 0$ ). We will now prove a different characterization of basic feasible points with an incomplete support that will be extremely useful in the algorithmic part. Particularly, we show that in order to check BF of a vector with an incomplete support, only one super support set should be checked to satisfy (20).

**THEOREM 5.6.** *Suppose that  $B$  is either a nonnegative type-1 or a type-2 symmetric set. Let  $\mathbf{x} \in C_s \cap B$  and  $\sigma \in \tilde{\Sigma}(-p(-\nabla f(\mathbf{x})))$ . Let  $i \in \{1, \dots, n+1\}$  be such that  $|S_{[i,n]}^\sigma \cup I_1(\mathbf{x})| = s$ , and let  $T = I_1(\mathbf{x}) \cup S_{[i,n]}^\sigma$ . If*

$$\mathbf{x}_T = P_{B_T}\left(\mathbf{x}_T - \frac{1}{L}\nabla_T f(\mathbf{x})\right), \quad (36)$$

then  $\mathbf{x}$  is a basic feasible point of  $(P)$ .

PROOF. If  $|I_1(\mathbf{x})| = s$ , then  $i = n + 1$ . That is,  $S_{[i,n]}^\sigma = \emptyset$ ,  $T = I_1(\mathbf{x})$  and

$$\mathbf{x}_{I_1(\mathbf{x})} = P_{B_{I_1(\mathbf{x})}} \left( \mathbf{x}_{I_1(\mathbf{x})} - \frac{1}{L} \nabla_{I_1(\mathbf{x})} f(\mathbf{x}) \right),$$

which is exactly the condition for BF for points with a full support, see Remark 5.1(a). If  $|I_1(\mathbf{x})| < s$ , then  $i \leq n$ . That is, the set  $S_{[i,n]}^\sigma$  is nonempty and  $T = I_1(\mathbf{x}) \cup S_{[i,n]}^\sigma$ . Since (36) holds, we have by Lemma 5.4 that

$$p \left( x_i - \frac{1}{L} \nabla_i f(\mathbf{x}) \right) \geq p \left( x_j - \frac{1}{L} \nabla_j f(\mathbf{x}) \right), \quad \text{for any } i \in I_1(\mathbf{x}), j \in T \cap I_0(\mathbf{x}). \quad (37)$$

In addition, by the definition of  $T$ , for any  $j \in T \cap I_0(\mathbf{x})$  and  $k \in I_0(\mathbf{x}) \setminus T$ , we have (recalling that  $x_j = x_k = 0$ ):

$$p \left( x_j - \frac{1}{L} \nabla_j f(\mathbf{x}) \right) \geq p \left( x_k - \frac{1}{L} \nabla_k f(\mathbf{x}) \right),$$

which combined with (37) implies that

$$p \left( x_i - \frac{1}{L} \nabla_i f(\mathbf{x}) \right) \geq p \left( x_j - \frac{1}{L} \nabla_j f(\mathbf{x}) \right), \quad \text{for any } i \in I_1(\mathbf{x}), j \in I_0(\mathbf{x}).$$

Therefore, by Theorem 5.4, it follows that  $\mathbf{x}$  is an  $L$ -stationary point, which readily implies that it is a basic feasible point.  $\square$

REMARK 5.3. We note that the approach for deriving optimality conditions in this paper is a departure from the classical derivation of optimality conditions for nonconvex programming that usually use characterizations based on normal cones and subdifferentials (Rockafellar [26], Rockafellar and Wets [27], Bolte et al. [8]). Recently, in the work of Pan et al. [25], explicit expressions for sparsity-constrained problems were derived by using normal cones that correspond to either Clarke or Bouligand tangent cones. It was shown in Pan et al. [25] that  $L$ -stationarity is a more restrictive condition than the derived optimality conditions. In fact, when the only constraint is the sparsity constraint, the conditions obtained in Pan et al. [25] are equivalent to BF (see Pan et al. [25, Tables 1 and 2]).

**6. Optimality conditions II: Coordinatewise-based conditions.** So far, we presented and studied optimality conditions, which are based on stationarity notions. For example, BF is a type of stationarity on the underlying set restricted to the support, and  $L$ -stationarity is a natural extension of the standard stationarity condition for smooth problems on convex sets. In any case, these types of conditions are related to standard optimality conditions in continuous optimization. Another type of optimality conditions that will be the subject of this section are coordinatewise optimality conditions. Loosely speaking, points satisfying such conditions are required to have a function value, which is better (or no worse) than function values of a certain set of points whose support is only slightly different from the current support. These types of conditions are of a more combinatorial flavor, and we will see that they are, in fact, superior to the stationarity-based conditions.

**6.1. Simple coordinatewise (simple-CW) optimality.** We begin by defining an unrestrictive optimality condition and proving its optimality.

DEFINITION 6.1 (SIMPLE-CW OPTIMALITY IN NONNEGATIVE TYPE-1 OR TYPE-2 SYMMETRIC SETS). Suppose that  $B$  is either a nonnegative type-1 or type-2 symmetric set. Let  $\mathbf{x} \in C_s \cap B$  be a basic feasible point of (P). Let

$$i \in \operatorname{argmin}_{\ell \in D(\mathbf{x})} \{ p(-\nabla_\ell f(\mathbf{x})) \}, \quad \text{with } D(\mathbf{x}) = \operatorname{argmin}_{k \in I_1(\mathbf{x})} p(x_k) \quad (38)$$

$$j \in \operatorname{argmin}_{\ell \in I_0(\mathbf{x})} \{ -p(-\nabla_\ell f(\mathbf{x})) \}, \quad (39)$$

where  $p(\cdot)$  is defined in (19). Then,  $\mathbf{x}$  is a *simple-CW point* of (P) if

$$f(\mathbf{x}) \leq \begin{cases} \min \{ f(\mathbf{x} - x_i \mathbf{e}_i + x_i \mathbf{e}_j), f(\mathbf{x} - x_i \mathbf{e}_i - x_i \mathbf{e}_j) \}, & B \text{ type-2,} \\ f(\mathbf{x} - x_i \mathbf{e}_i + x_i \mathbf{e}_j), & B \text{ nonneg. type-1.} \end{cases}$$

Note that there might be several choices for the indices  $i$  and  $j$  in the above definition. We will make the convention that in such a case, we will always pick the smallest index among the possible choices. In addition, note that the condition is independent of any Lipschitz constant.

When  $B$  is a type-2 (nonnegative type-1) symmetric set, then the variable entering the support corresponds to the maximal absolute value partial derivative (minimal partial derivative), whereas the choice of the variable leaving the support is made in two stages: first, we consider all the indices corresponding to the variables with minimal absolute value (value) among the support indices, and from this set of indices, we pick an index corresponding to the minimal absolute value (maximal value) of the partial derivative.

Obviously, simple-CW optimality is a necessary optimality condition.

**LEMMA 6.1.** *Let  $B$  be either nonnegative type-1 set or a type-2 symmetric set, and let  $\mathbf{x}^*$  be an optimal solution of problem (1). Then,  $\mathbf{x}^*$  is a simple-CW point of (P).*

Simple-CW optimality seems like a rather unrestrictive condition, and yet, it turns out that when Assumption [A+] holds, it is better (i.e., more restrictive) than  $L_2(f)$ -stationarity. For that, we first need to show the following result.

**LEMMA 6.2.** *Let  $B$  be either a nonnegative type-1 or a type-2 symmetric set, and let  $\mathbf{x} \in C_s \cap B$  be a basic feasible solution of (P). Assume that  $i$  is an index chosen according to (38). Then, for any  $L > 0$ ,*

$$i \in \operatorname{argmin}_{\ell \in I_1(\mathbf{x})} p(Lx_\ell - \nabla_\ell f(\mathbf{x})).$$

**PROOF.** Let  $L > 0$ . Since  $\mathbf{x}$  is a BF point of (P), it follows that

$$\mathbf{x}_{I_1(\mathbf{x})} = P_{B_{I_1(\mathbf{x})}} \left( \mathbf{x}_{I_1(\mathbf{x})} - \frac{1}{L} \nabla_{I_1(\mathbf{x})} f(\mathbf{x}) \right),$$

and hence, by Corollary 3.1 with  $D = B_{I_1(\mathbf{x})}$  (for nonnegative type-1 sets, the implication is trivial),

$$p(\mathbf{x}_{I_1(\mathbf{x})}) = P_{B_{I_1(\mathbf{x})} \cap \mathbb{R}_+^{|I_1(\mathbf{x})|}} \left( p \left[ \mathbf{x}_{I_1(\mathbf{x})} - \frac{1}{L} \nabla_{I_1(\mathbf{x})} f(\mathbf{x}) \right] \right).$$

Since  $B_{I_1(\mathbf{x})} \cap \mathbb{R}_+^{|I_1(\mathbf{x})|}$  is a nonnegative type-1 symmetric set, it follows by Theorem 3.1, and the uniqueness of the projection onto closed and convex sets, that there is a permutation that sorts  $\{p(x_k)\}_{k \in I_1(\mathbf{x})}$  and  $\{p(Lx_k - \nabla_k f(\mathbf{x}))\}_{k \in I_1(\mathbf{x})}$ . Consequently, there exists an index minimizing  $p(Lx_k - \nabla_k f(\mathbf{x}))$  among the indices minimizing  $p(x_k)$ . Hence,

$$\min_{k \in I_1(\mathbf{x})} p(Lx_k - \nabla_k f(\mathbf{x})) = \min_{k \in D(\mathbf{x})} p(Lx_k - \nabla_k f(\mathbf{x})). \quad (40)$$

Obviously,  $p(x_k)$  has the same value for all  $k \in D(\mathbf{x})$ , which we will denote by  $p^*$ . In addition, since  $p(x) \equiv x$  when  $B$  is type-1 symmetric and  $x_k \nabla_k f(\mathbf{x}) \leq 0$  (by Lemma 5.1),  $p(x) = |x|$  when  $B$  is type-2 symmetric, it follows that for any  $k \in D(\mathbf{x})$

$$p(Lx_k - \nabla_k f(\mathbf{x})) = Lp^* + p(-\nabla_k f(\mathbf{x})),$$

which combined with (40) implies that the index  $i$ , which is the index corresponding to the minimal value of  $p(-\nabla_k f(\mathbf{x}))$  over  $D(\mathbf{x})$ , also corresponds to the minimal value of  $p(Lx_k - \nabla_k f(\mathbf{x}))$  over  $k \in I_1(\mathbf{x})$ .  $\square$

We can now show that under Assumption [A+], any simple-CW point is also an  $L_2(f)$ -stationary point. In a sense, this is a rather surprising result since the simple-CW condition only checks that the point has a better value than one or two points, which are slightly different than the current point.

**THEOREM 6.1 (SIMPLE-CW  $\Rightarrow$   $L_2(f)$ -STATIONARITY).** *Let  $f$  satisfy Assumption [A+] and let  $B$  be either a nonnegative type-1 or a type-2 symmetric set. Then, any simple-CW point of (P) is an  $L$ -stationary point of (P) for any  $L \geq L_2(f)$ .*

**PROOF.** Let  $\mathbf{x} \in C_s \cap B$  be a simple-CW point, and let  $L \geq L_2(f)$ . By definition,  $\mathbf{x}$  is a BF-point, and therefore if  $|I_1(\mathbf{x})| < s$ , then by Theorem 5.5,  $\mathbf{x}$  is an  $L$ -stationarity point of (P) (actually, for any  $L$ ), and the result is proven.

Suppose then that  $|I_1(\mathbf{x})| = s$ , then by the definition of simple-CW optimality and Lemma 6.2, for

$$j \in \operatorname{argmin}_{\ell \in I_0(\mathbf{x})} \{-p(-\nabla_\ell f(\mathbf{x}))\} \quad \text{and} \quad i \in \operatorname{argmin}_{\ell \in I_1(\mathbf{x})} \{p(Lx_\ell - \nabla_\ell f(\mathbf{x}))\}, \quad (41)$$

it holds that

$$f(\mathbf{x}) \leq f(\mathbf{x} + x_i(\mathbf{e}_j - \mathbf{e}_i)). \quad (42)$$

By the block descent lemma (Lemma 2.2), for any  $\mathbf{y} \in \mathbb{R}^n$  satisfying  $\|\mathbf{y} - \mathbf{x}\|_0 \leq 2$ ,

$$f(\mathbf{y}) - f(\mathbf{x}) \leq \nabla f(\mathbf{x})^T (\mathbf{y} - \mathbf{x}) + \frac{L}{2} \|\mathbf{y} - \mathbf{x}\|^2. \quad (43)$$

In particular, for  $\mathbf{y} = \mathbf{x}^{i,j} = \mathbf{x} - x_i \mathbf{e}_i + x_j \mathbf{e}_j$ , (43) reduces to

$$\nabla_j f(\mathbf{x})(x_j - x_i) + \nabla_i f(\mathbf{x})(x_j - x_i) + L\|x_i - x_j\|^2 \geq f(\mathbf{y}) - f(\mathbf{x}).$$

Using (42) and the fact that  $x_j = 0$ , we obtain

$$x_i(\nabla_j f(\mathbf{x}) - \nabla_i f(\mathbf{x}) + Lx_i) \geq 0. \quad (44)$$

We now split the analysis between the two types of symmetries.

- If  $B$  is nonnegative type-1 symmetric, then using the fact that  $i \in I_1(\mathbf{x})$ , we have

$$Lx_i - \nabla_i f(\mathbf{x}) \geq -\nabla_j f(\mathbf{x}).$$

- If  $B$  is type-2 symmetric, then using the same argument as above on the condition

$$f(\mathbf{x}) \leq f(\mathbf{x} - x_i \mathbf{e}_i - x_j \mathbf{e}_j),$$

yields the inequality

$$x_i(-\nabla_j f(\mathbf{x}) - \nabla_i f(\mathbf{x}) + Lx_i) \geq 0.$$

Multiplying the latter inequality with (44), and using the fact that  $i \in I_1(\mathbf{x})$ , implies

$$(Lx_i - \nabla_i f(\mathbf{x}))^2 \geq \nabla_j f(\mathbf{x})^2,$$

and consequently

$$|Lx_i - \nabla_i f(\mathbf{x})| \geq |\nabla_j f(\mathbf{x})|.$$

We have thus shown that in both cases

$$p(Lx_i - \nabla_i f(\mathbf{x})) \geq p(-\nabla_j f(\mathbf{x})).$$

By the definitions of the indices  $i$  and  $j$  given in (42), it follows that

$$p(Lx_\ell - \nabla_\ell f(\mathbf{x})) \geq p(-\nabla_m f(\mathbf{x}))$$

for any  $\ell \in I_1(\mathbf{x})$  and  $m \in I_0(\mathbf{x})$ , which by Theorem 5.4, implies that  $\mathbf{x}$  is an  $L$ -stationarity point.  $\square$

A direct consequence of Theorem 6.1 is that when  $B$  is either a nonnegative-type 1 or a type-2 symmetric set,  $L_2(f)$ -stationarity is a necessary optimality condition. This is a stronger result than the one shown in Theorem 5.3, where  $L$ -stationarity was shown for  $L > L(f)$  (without any symmetry assumption).

**THEOREM 6.2.** *Suppose that Assumption [A+] is satisfied. Then, if  $\mathbf{x}^*$  is an optimal solution of (P), then it is an  $L$ -stationary point for any  $L \geq L_2(f)$ .*

**6.2. Zero-CW and full-CW optimal points.** We will now present two additional CW-type optimality conditions whose validation requires the ability to minimize the objective function over a given super support. This assumption holds, for example, when the objective function is convex, but there are other scenarios when it is possible, such as the case when  $B$  is the  $l_2$  unit ball and the objective function is a (possibly nonconvex) quadratic function. The minimization in this case reduces to the solution of a trust region subproblem (More and Sorensen [23]). The zero-CW optimality condition is similar to simple-CW optimality, but requires the point to be of a smaller or equal value than the points with a super support set constructed by swapping the indices  $i$  and  $j$  defined in (38) and (39), and adding (if the support is not full) indices corresponding to the largest values of  $p(-\nabla_j f(\mathbf{x}))$ .

DEFINITION 6.2 (ZERO-CW OPTIMAL POINT). Let  $\mathbf{x}$  be a basic feasible point of (P), and assume that  $\sigma \in \tilde{\Sigma}(-p(-\nabla f(\mathbf{x})))$ . Let  $k \in \{1, 2, \dots, n\}$  satisfy

$$|(S_{[k,n]}^\sigma \cup I_1(\mathbf{x}) \cup \{j\}) \setminus \{i\}| = s,$$

where  $i$  and  $j$  are defined in (38) and (39), respectively. Let

$$T = (S_{[k,n]}^\sigma \cup I_1(\mathbf{x}) \cup \{j\}) \setminus \{i\}.$$

Then,  $\mathbf{x}$  is called a *zero-CW optimal point* if

$$f(\mathbf{x}) \leq \min\{f(\mathbf{y}) : \mathbf{y} \in B, I_1(\mathbf{y}) \subseteq T\}.$$

Similarly, a *full-CW optimal point* is a point in which any possible swap between support and nonsupport indices does not result with a better (i.e., smaller) function value.

DEFINITION 6.3 (FULL-CW OPTIMAL POINT). Let  $\mathbf{x}$  be a basic feasible point of (P), and assume that  $\sigma \in \tilde{\Sigma}(-p(-\nabla f(\mathbf{x})))$ . For each  $i \in I_1(\mathbf{x})$ ,  $j \in I_0(\mathbf{x})$ , define  $k^{i,j}$  as an index satisfying

$$|(S_{[k^{i,j},n]}^\sigma \cup I_1(\mathbf{x}) \cup \{j\}) \setminus \{i\}| = s.$$

Let

$$T_{i,j} = (S_{[k^{i,j},n]}^\sigma \cup I_1(\mathbf{x}) \cup \{j\}) \setminus \{i\}.$$

Then,  $\mathbf{x}$  is called a *full-CW optimal point* if for any  $i \in I_1(\mathbf{x})$ ,  $j \in I_0(\mathbf{x})$ , it holds that

$$f(\mathbf{x}) \leq \min\{f(\mathbf{y}) : \mathbf{y} \in B, I_1(\mathbf{y}) \subseteq T_{i,j}\}.$$

We will now state formally some obvious facts on zero- and full-CW optimal points. First, by their definition, zero- and full-CW optimal points are also simple optimal points, and consequently, by Theorem 6.1, also  $L_2(f)$ -stationary points.

THEOREM 6.3. *Let  $B$  be either a nonnegative type-1 or a type-2 symmetric set. If  $\mathbf{x}$  is a zero- or full-CW optimal point, then it is also a simple-CW optimal point, and consequently, they are also  $L_2(f)$ -stationary points.*

Another quite obvious fact is that zero and full-CW optimality are necessary optimality conditions for problem (P). This fact does not require any symmetry assumption.

THEOREM 6.4. *Let  $\mathbf{x}$  be an optimal solution of problem (P). Then, it is a zero-CW as well as full-CW optimal point.*

**6.3. Examples.** The following examples demonstrate the hierarchy between the various optimality conditions.

EXAMPLE 6.1. Consider the problem

$$\max\{f(x_1, x_2, x_3) \equiv d_1 x_1^2 + d_2 x_2^2 + d_3 x_3^2; -1 \leq x_1, x_2, x_3 \leq -1, \|\mathbf{x}\|_0 \leq 2\},$$

where  $d_1 > d_2 > d_3 > 0$ . It is not difficult to show that there are actually 12 BF points:  $(\pm 1, \pm 1, 0)$ ,  $(0, \pm 1, \pm 1)$ ,  $(\pm 1, 0, \pm 1)$ . In this example  $L(f) = L_2(f) = 2d_1$ . By Theorem 5.4, a BF point whose off support consists of the index  $i_1 \in \{1, 2, 3\}$  is an  $L(f)$ -stationary point if and only if (recalling that we are actually minimizing  $-f$ ):

$$|L(f)x_i + \nabla_i f(\mathbf{x})| \geq |\nabla_i f(\mathbf{x})|, \quad \text{for all } i \in \{1, 2, 3\} \setminus \{i_1\}.$$

This condition is satisfied for all BF points since in this example  $\nabla_i f(\mathbf{x}) = 2d_i x_{i_1} = 0$ . Therefore, all 12 BF points are actually  $L(f)$ -stationary points. On the other hand, it is to see that only the four optimal points  $(1, 1, 0)$ ,  $(-1, 1, 0)$ ,  $(1, -1, 0)$ ,  $(-1, -1, 0)$  are simple-CW points, and hence also zero and full-CW points.

TABLE 4. Hierarchy between optimality condition in problem (45).

Support	{1, 2}	{1, 3}	{1, 4}	{2, 3}	{2, 4}	{3, 4}	6
Values	0.003 0.997 0 0	0.003 0 0.997 0	0.002 0 0 0.998	0 0.910 0 0			Total
BF	✓	✓	✓	✓			4
$L(f)$ -stationary			✓	✓			2
Simple CW			✓	✓			2
Zero CW			✓				1
Full CW			✓				1

EXAMPLE 6.2. Consider the following 2-sparse least squares problem over the unit  $\ell_1$ -norm ball:

$$\min \left\{ \left\| \begin{pmatrix} 1,000 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0.01 & 1 \end{pmatrix} \mathbf{x} - \begin{pmatrix} 3 \\ 1 \\ 9 \end{pmatrix} \right\|_2^2 : \mathbf{x} \in C_2 \cap B_1^4[\mathbf{0}, 1] \right\}. \quad (45)$$

Table 4 depicts the different BF points (with three digits of accuracy) per corresponding super support, with a check mark indicating which conditions are satisfied by each point. The blank cells for the super supports  $\{2, 4\}$ ,  $\{3, 4\}$  indicate that there is no point satisfying an optimality condition having one of these super supports. This situation arises when a stationary point of the restricted problem over a specific super support has an incomplete support. For example, the stationary point corresponding to the restriction of the problem to the super support  $\{2, 4\}$  is  $(0, 0, 0, 1)^T$ . By definition, for the former point to be a BF point, it must be stationary over any super support. Although it is stationary for  $\{2, 4\}$ , it is not for  $\{1, 4\}$ , implying that there is no BF point with the super support  $\{2, 4\}$ . In this example, the two BF points  $\mathbf{v} = (0.002, 0, 0, 0.998)^T$ ,  $\mathbf{w} = (0, 0.910, 0.09, 0)^T$  are the only  $L(f)$ -stationary points as well as simple-CW points. This means that the nonoptimal point  $\mathbf{w}$  satisfies both optimality conditions. However, the only point that is either a zero of full-CW point is the optimal point  $\mathbf{v}$ .

**7. Algorithms.** The hierarchy between the optimality conditions established in the previous section suggests that there is also a hierarchy of algorithms with respect to the quality of the points to which they are guaranteed to converge. Thus, for example, methods that are guaranteed to converge to  $L(f)$ -stationary points are worse than algorithms that are guaranteed to converge to simple-CW points in the sense that they produce points that satisfy relatively weak optimality conditions. Later on, in §8, we will show by empirical experiments that methods, which are guaranteed to converge to points satisfying strong optimality conditions, tend to outperform methods that are only guaranteed to converge points satisfying weaker optimality conditions. The algorithms that we will present here require the ability to minimize the function over a given super support set. Thus, given a set  $T \subseteq \{1, 2, \dots, n\}$  such that  $|T| = s$ , we will assume that the following problem is solvable:

$$v_T \equiv \min \{f(\mathbf{x}) : \mathbf{x} \in B, I_1(\mathbf{x}) \subseteq T\}. \quad (46)$$

The above problem can be solved, for example, when the objective function is convex, or when  $f$  is a (possibly nonconvex) quadratic function and  $B$  is an  $l_2$  ball (More and Sorensen [23]). A solution  $\mathbf{x}$  of problem (46) is called a *support optimal point*. It is important to note that since there are  $\binom{n}{s}$  possible choices for the set  $T$ , there is only a finite number of values  $v_T$ .

**7.1. BFS search.** The first algorithm that we need to define is an algorithm that, given a specific feasible point, seeks a basic feasible point with a lower or equal function value. The method receives a feasible point ( $\mathbf{x} \in C_s \cap B$ ) as an input. If the support of  $\mathbf{x}$  is full, the method finds the minimizer of the objective function over its support. If the support of  $\mathbf{x}$  is incomplete, the method adds the indices corresponding to the largest elements in  $p(-\nabla f(\mathbf{x}))$  to the support set, and optimizes the function over the constructed index set. If the support of the obtained vector is full, the process is terminated; otherwise, it continues until either a full support is reached or no decrease in the function value is achieved.

Downloaded from informs.org by [132.68.247.241] on 27 October 2015, at 05:10. For personal use only, all rights reserved.

**Algorithm 5** (Basic feasible search (BFS)).

**Initialization:**  $\mathbf{x}^0 \in C_s \cap B$ ,  $k = 0$ .

**Output:**  $\mathbf{u} \in C_s \cap B$ , which is a basic feasible point.

1. Repeat
  - (a) let  $\sigma \in \tilde{\Sigma}(-p(-\nabla f(\mathbf{x}^k)))$
  - (b) set  $i \in \{1, \dots, n+1\}$  such that  $|S_{[i,n]}^\sigma \cup I_1(\mathbf{x}^k)| = s$
  - (c) set  $T_k = I_1(\mathbf{x}^k) \cup S_{[i,n]}^\sigma$
  - (d) take  $\mathbf{x}^{k+1} \in \operatorname{argmin}\{f(\mathbf{y}): \mathbf{y} \in B, I_1(\mathbf{y}) \subseteq T_k\}$
  - (e)  $k \leftarrow k+1$
 Until  $f(\mathbf{x}^{k-1}) \leq f(\mathbf{x}^k)$
2. Set  $\mathbf{u} = \mathbf{x}^{k-1}$ .

The process is obviously finite since the sequence of objective function values is strictly decreasing:  $v_{T_1} > v_{T_2} > \dots$ , and there is a finite number of possible values  $v_T$ . We only need to show that the output  $\mathbf{u} = \mathbf{x}^{k-1}$  is indeed a basic feasible point ( $k$  being the last index in the process).

LEMMA 7.1. *Let  $B$  be either a nonnegative type-1 or a type-2 symmetric set. Let  $\mathbf{u} = \mathbf{x}^{k-1}$  be the output of the BFS procedure. Then,  $\mathbf{u}$  is a basic feasible point.*

PROOF. By the definition of the method, the point  $\mathbf{z}$  defined as

$$\mathbf{z} \in \operatorname{argmin}\{f(\mathbf{y}): \mathbf{y} \in B, I_1(\mathbf{y}) \subseteq T_{k-1}\}$$

satisfies  $f(\mathbf{x}^{k-1}) \leq f(\mathbf{z})$ . Since  $\mathbf{x}^{k-1} \in B$  and  $I_1(\mathbf{x}^{k-1}) \subseteq T_{k-1}$ , it follows that

$$\mathbf{x}^{k-1} \in \operatorname{argmin}\{f(\mathbf{y}): \mathbf{y} \in B, I_1(\mathbf{y}) \subseteq T_{k-1}\},$$

which is the same as

$$\mathbf{x}_{T_{k-1}}^{k-1} \in \operatorname{argmin}\{f(\mathbf{U}_{T_{k-1}} \mathbf{w}): \mathbf{w} \in B_{T_{k-1}}\}.$$

Hence  $\mathbf{x}_{T_{k-1}}^{k-1}$  is a stationary point of the above problem, and therefore for any  $L > 0$ ,

$$\mathbf{x}_{T_{k-1}}^{k-1} = P_{B_{T_{k-1}}}\left(\mathbf{x}_{T_{k-1}}^{k-1} - \frac{1}{L} \nabla_{T_{k-1}} f(\mathbf{x}^{k-1})\right),$$

which by Theorem 5.6 implies that  $\mathbf{x}^{k-1}$  is a basic feasible point.  $\square$

We also note that all the methods that will be discussed in the sequel move from one basic feasible point, which is also support optimal, using the BFS procedure, while maintaining a strictly decreasing sequence of function values. Therefore, since there is a finite number of possible function values for the generated sequence (as it is contained in the set  $\{v_T: T \subseteq \{1, 2, \dots, n\}, |T| = s\}$ ), the algorithms are always finite.

**7.2. The zero-CW search method.** The zero-CW search method, as its name suggests, finds a zero-CW point by swapping at each iteration the support and nonsupport indices  $i$  and  $j$  defined in (38) and (39). If this swap induces a better basic feasible point in terms of the objective function, then the procedure proceeds with the new point; otherwise, the process terminates with a zero-CW point.

**Algorithm 6** (Zero-CW search method (ZCWS)).

**Initialization:**  $\mathbf{x}^0 \in C_s \cap B$  a basic feasible point,  $k = 0$ .

**Output:**  $\mathbf{u} \in C_s \cap B$ , which is a zero-CW point.

**General Step** ( $k = 0, 1, 2, \dots$ )

1.  $D(\mathbf{x}^k) = \operatorname{argmin}_{\ell \in I_1(\mathbf{x}^k)} p(x_\ell^k)$ .
2.  $i \in \operatorname{argmin}_{\ell \in D(\mathbf{x}^k)} \{p(-\nabla_\ell f(\mathbf{x}^k))\}$ .
3.  $j \in \operatorname{argmin}_{\ell \in I_0(\mathbf{x}^k)} \{-p(-\nabla_\ell f(\mathbf{x}^k))\}$ .
4. Let  $\sigma \in \tilde{\Sigma}(-p(-\nabla f(\mathbf{x}^k)))$  and let  $\ell$  be such that  $|(S_{[\ell,n]}^\sigma \cup I_1(\mathbf{x}^k) \cup \{j\}) \setminus \{i\}| = s$ .
5. Define  $T_k = (S_{[\ell,n]}^\sigma \cup I_1(\mathbf{x}^k) \cup \{j\}) \setminus \{i\}$ .

6. Set  $\mathbf{x} \in \operatorname{argmin}\{f(\mathbf{y}): \mathbf{y} \in B, I_1(\mathbf{y}) \subseteq T_k\}$ .
7.  $\mathbf{x}^{k+1} = \text{BFS}(\mathbf{x})$
8. If  $f(\mathbf{x}^k) \leq f(\mathbf{x}^{k+1})$ , then STOP and the output is  $\mathbf{u} = \mathbf{x}^k$ . Otherwise,  $k \leftarrow k + 1$  and go back to step 1.

Since the termination of the ZCWS method is exactly the validity of the zero-CW optimality conditions, it follows that the procedure produces a zero-CW point.

**7.3. The full-CW search method.** The full-CW search method is a scheme aimed at finding full-CW points. The process first executes the ZCWS method and then proceeds to check if the point is a full-CW point by examining the objective function’s value for any possible swaps between indices in the support with indices in the off support. If there is a swap which induces a reduction in the function value, then the procedure continues with that point. Otherwise, a full-CW point is returned.

**Algorithm 7** (Full-CW search method).

**Initialization:**  $\mathbf{x}^0 \in C_s \cap B$ —a basic feasible point,  $k = 0$ .

**Output:**  $\mathbf{u} \in C_s \cap B$ , which is a full-CW point.

**General Step** ( $k = 0, 1, 2, \dots$ )

1.  $\mathbf{w}^k = \text{ZCWS}(\mathbf{x}^k)$ .
2. Let  $\sigma \in \tilde{\Sigma}(-p(-\nabla f(\mathbf{x}^k)))$ , and for any  $i \in I_1(\mathbf{w}^k)$ ,  $j \in I_0(\mathbf{w}^k)$ , let  $\ell^{i,j}$  be such that  $|(S_{[\ell^{i,j}, n]}^\sigma \cup I_1(\mathbf{w}^k) \cup \{j\}) \setminus \{i\}| = s$
3. Define  $T_k^{i,j} = (S_{[\ell^{i,j}, n]}^\sigma \cup I_1(\mathbf{w}^k) \cup \{j\}) \setminus \{i\}$ .
4. Take  $\mathbf{z}^{i,j} \in \operatorname{argmin}\{f(\mathbf{y}): \mathbf{y} \in B, I_1(\mathbf{y}) \subseteq T_k^{i,j}\}$ .
5. Set  $(i_0, j_0) \in \operatorname{argmin}\{f(\mathbf{z}^{i,j}): i \in I_1(\mathbf{w}^k), j \in I_0(\mathbf{w}^k)\}$ .
6. Define  $\mathbf{x}^{k+1} = \text{BFS}(\mathbf{z}^{i_0, j_0})$ .
7. If  $f(\mathbf{x}^k) \leq f(\mathbf{x}^{k+1})$ , then STOP and the output is  $\mathbf{u} = \mathbf{x}^k$ . Otherwise,  $k \leftarrow k + 1$  and go back to step 1.

The full-CW search method obviously find a full-CW point in a finite number of steps.

**8. Numerical experiments and applications.** The objective of the simulations is to demonstrate how the hierarchy between the optimality conditions can be observed in concrete problems. We will examine two applications. The first is sparse index tracking (Takeda et al. [29]), and the second is the compressed sensing problem of retrieving a sparse signal from an underdetermined system of equations. We will compare four methods:

1. The zero-CW search method, which attains zero-CW points.
2. The full-CW search method, which attains full-CW points.
3. The IHT method, which attains  $L$ -stationary points (for  $L > L(f)$ ).
4. The TGA method, which greedily builds a super support set, and is a generalization of the greedy algorithm defined in Takeda et al. [29].

We will use the acronyms given in Table 5.

**Algorithm 8** (Iterative hard thresholding (IHT)).

**Initialization:**  $\mathbf{x}^0 \in \mathbb{R}^n$ ,  $k = 0$ .

**Output:**  $\mathbf{u} \in C_s \cap B$ .

1. Repeat
  - (a)  $\mathbf{x}^{k+1} \in P_{C_s \cap B}(\mathbf{x}^k - (1/L)\nabla f(\mathbf{x}^k))$ .
  - (b)  $k \leftarrow k + 1$ .
 Until  $\|\mathbf{x}^{k-1} - \mathbf{x}^k\| \leq \epsilon$ .
2. Return  $\mathbf{u} = \mathbf{x}^k$ .

TABLE 5. The acronyms of the four compared methods.

Acronym	Method
ZCWS	The zero-CW search method (Algorithm 6)
FCWS	The full-CW search method (Algorithm 7)
IHT	The IHT method (Algorithm 8)
TGA	The greedy support pursuit method (Algorithm 9)

**Algorithm 9** (TGA).**Initialization:**  $\mathbf{x} = \mathbf{0}_n$ ,  $S = \emptyset$ .**Output:**  $\mathbf{x} \in C_s \cap B$ .

1. While  $|S| < s$ , do:
  - (a)  $(j, \mathbf{x}) \in \underset{(\ell \in S^c, \mathbf{z} \in B)}{\operatorname{argmin}} \{f(\mathbf{z}): I_1(\mathbf{z}) \subseteq S \cup \{\ell\}\}$
  - (b) Set  $S \leftarrow S \cup \{j\}$ .
2. Return  $\mathbf{x}$ .

The TGA algorithm starts from the zeros vector, and greedily adds indices until the super support contains exactly  $s$  indices. Several remarks should be taken into account:

- Once an index enters the suggested super support set, it does not leave the set.
- The method is terminated once the super support is full, even if the support of the actual solution is not.
- The method always starts from the same point (zeros vector).
- The output might not satisfy any of the optimality conditions.
- If  $B = \mathbb{R}^n$ , then this method is exactly the orthogonal matching pursuit (OMP) method (Mallat and Zhang [22]), and when  $B = \Delta'_n$ , the method is the same as the greedy method introduced in Takeda et al. [29].

**8.1. Sparse index tracking.** The index tracking problem is the problem of tracking an index using a set of assets. Mathematically, it is defined as a minimization problem of a least squares term  $\|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2$  where  $\mathbf{A}$  is the sample matrix,  $\mathbf{b}$  is the so-called index vector, and the optimization is made over a set  $C$ , of all admissible vectors. The problem was addressed in Takeda et al. [29], with  $C$  being the sparse unit-sum set. The authors in Takeda et al. [29] offered to use sparsity as a tool for controlling the trade-off between the performance of the tracker and the robustness of the model (overfitting).

We also note that the similar problem of portfolio optimization was presented in Kyrillidis et al. [18] as an application. We will test the sparse index tracking problem with no “short” allowed. The problem formulation is

$$\begin{aligned} \min \quad & \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 \\ \text{s.t.} \quad & \mathbf{x} \in C_s \cap \Delta_n, \end{aligned}$$

where the matrix  $\mathbf{A}$  contains the daily returns of stocks traded in the New York Stock Exchange, and the vector  $\mathbf{b}$  consists of the daily returns of the S&P 500 index.

We created 180 random sets of 54 stocks and 72 days of trade, and tested three levels of sparsity: 9, 18, 27. For each random problem, we executed each method from the output of the TGA method with  $s = 1$ . Then, we ran each method (except for the TGA) from the output of each of the other methods and counted the number of times the results were improved, which is the number of times these methods did not reach better optimality than their worst-case theoretical guarantees. For example, if the output of the IHT method was improved by the ZCWS method in all its runs, then we can conclude that the outputs of the IHT methods were always  $L$ -stationary points that are not zero-CW point.

The results are summarized in Table 6. Each cell in the table indicates the number of times the output of the algorithm in the first column improved the algorithm in the second column with the corresponding sparsity level.

Several conclusions can be deduced from the results:

1. The hierarchy between the optimality conditions is (unsurprisingly) validated.
2. The IHT method never reached a zero-CW or full-CW point.

TABLE 6. Number of improvements by sparsity level.

Improver	Improved	$s = 9$	$s = 18$	$s = 27$	Total
ZCWS	FCWS	0	0	0	0
	IHT	60	60	60	180
	TGA	9	56	50	115
FCWS	ZCWS	33	11	17	61
	IHT	60	60	60	180
	TGA	15	56	51	122
IHT	ZCWS	0	0	0	0
	FCWS	0	0	0	0
	TGA	3	56	50	109

TABLE 7. Number of improvements by sparsity level.

Improver	Improved	$s = 9$	$s = 18$	$s = 27$	Total
ZCWS	FCWS	0	0	0	0
	IHT	60	60	60	180
	TGA	15	60	60	135
FCWS	ZCWS	0	6	14	20
	IHT	60	60	60	180
	TGA	15	60	60	135
IHT	ZCWS	0	0	0	0
	FCWS	0	0	0	0
	TGA	11	57	60	128

- The ZCWS reached a full-CW point in two-thirds of the instances (66%).
- The TGA was improved in most of the instances by all the methods when the sparsity level was greater than 9.

We note that, in practice, it is quite rare to obtain a BF point with an incomplete support. In the index tracking problem, however, there were a few cases where the algorithms did example BF points with an incomplete support.

**8.2. Compressed sensing with signals from the unit simplex.** We will now consider randomly generated compressed sensing problems of the form

$$\begin{aligned} \min \quad & \|\mathbf{Ax} - \mathbf{b}\|^2 \\ \text{s.t.} \quad & \mathbf{x} \in C_s \cap \Delta_n. \end{aligned}$$

We generated 180 random problems with a matrix  $\mathbf{A} \in \mathbb{R}^{63 \times 91}$  whose components are independently generated from a standard normal distribution. Then, we generated a sparse signal  $\mathbf{x}_{\text{true}}$  with a sparsity level  $s \in \{9, 18, 27\}$  from the unit-simplex set (see Smith and Tromble [28]). The vector  $\mathbf{b}$  was then chosen as  $\mathbf{b} = \mathbf{Ax}_{\text{true}} + \mathbf{n}$ , where the components of  $\mathbf{n}$  were generated by a normal distribution with zero mean and standard deviation  $\sigma = 0.6$ . The results are summarized in Table 7.

The conclusions that can be drawn from this set of results are almost the same as the ones seen in the previous set of experiments: The ZCWS methods reached a full-CW point in many instances (89% of the instances), IHT is strictly dominated by the CW-type methods, and the TGA methods were almost always improved by all the methods when the sparsity level is greater than 9.

**Acknowledgment.** The research of Amir Beck was partially supported by the Israel Science Foundation [Grant ISF 253/12].

## References

- Attouch H, Bolte J, Svaiter BF (2013) Convergence of descent methods for semi-algebraic and tame problems: Proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods. *Math. Programming Ser. A* 137(1–2):91–129.
- Bahmani S, Raj B, Boufounos PT (2013) Greedy sparsity-constrained optimization. *J. Machine Learn. Res.* 14(1):807–841.
- Beck A, Eldar YC (2013) Sparsity constrained nonlinear optimization: Optimality conditions and algorithms. *SIAM J. Optim.* 23(3):1480–1509.
- Beck A, Teboulle M (2009) A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.* 2(1):183–202.
- Blumensath T, Davies ME (2008) Iterative thresholding for sparse approximations. *J. Fourier Anal. Appl.* 14(5):629–654.
- Blumensath T, Davies ME (2009) Iterative hard thresholding for compressed sensing. *Appl. Comput. Harmonic Anal.* 27(3):265–274.
- Blumensath T, Davies ME (2009) Normalised iterative hard thresholding; guaranteed stability and performance. *IEEE J. Selected Topics Signal Processing* 4(2):298–309.
- Bolte J, Sabach S, Teboulle M (2014) Proximal alternating linearized minimization for nonconvex and nonsmooth problems. *Math. Programming Ser. A* 146(1–2):459–494.
- Bruckstein AM, Donoho DL, Elad M (2009) From sparse solutions of systems of equations to sparse modeling of signals and images. *SIAM Rev.* 51(1):34–81.
- Candes E, Tao T (2007) The Dantzig selector: Statistical estimation when  $p$  is much larger than  $n$ . *Ann. Statist.* 35(6):2313–2351.
- Candes EJ, Tao T (2005) Decoding by linear programming. *IEEE Trans. Inform. Theory* 51(12):4203–4215.
- Cartis C, Thompson A (2013) A new and improved quantitative recovery analysis for iterative hard thresholding algorithms in compressed sensing. Technical Report NA-13/20, Mathematical Institute, University of Oxford, Oxford, UK).

- [13] Chen S, Donoho D, Saunders M (1998) Atomic decomposition by basis pursuit. *SIAM J. Scientific Comput.* 20(1):33–61.
- [14] Davenport MA, Duarte MF, Eldar YC, Kutyniok G (2011) Introduction to compressed sensing. Preprint:1–68.
- [15] David LD, Elad M (2003) Optimally sparse representation in general (nonorthogonal) dictionaries via L1 minimization. *Proc. Natl. Acad. Sci.* 100(5):2197–2202.
- [16] Duarte MF, Eldar YC (2011) Structured compressed sensing: From theory to applications. *IEEE Trans. Signal Processing* 59(9):4053–4085.
- [17] Kiwiel KC (2007) On linear-time algorithms for the continuous quadratic knapsack problem. *J. Optim. Theory Appl.* 134(3):549–554. [English.]
- [18] Kyrillidis A, Becker S, Cevher V, Koch C (2013) Sparse projections onto the simplex. Preprint, [arXiv:1206.1529](https://arxiv.org/abs/1206.1529) 28.
- [19] Lu Z (2013) Iterative hard thresholding methods for  $l_0$  regularized convex cone programming. *Math. Programming Ser. A* 147(1–2):125–154.
- [20] Lu Z, Zhang Y (2013) Sparse approximation via penalty decomposition methods. *SIAM J. Optim.* 23(4):2448–2478.
- [21] Luss R, Teboulle M (2013) Conditional gradient algorithms for rank-one matrix approximations with a sparsity constraint. *SIAM Rev.* 55(1):65–98.
- [22] Mallat S, Zhang Z (1993) Matching pursuits with time-frequency dictionaries. *IEEE Trans. Signal Processing* 41(12):3397–3415.
- [23] More J, Sorensen D (1983) Computing a trust region step. *SIAM J. Sci. Statist. Comput.* 4(3):553–572.
- [24] Needell D, Tropp JA (2009) CoSaMP: Iterative signal recovery from incomplete and inaccurate samples. *Appl. Comput. Harmonic Anal.* 26(3):301–321.
- [25] Pan L, Xiu N, Zhou S Gradient support projection algorithm for affine feasibility problem with sparsity and nonnegativity. Preprint, <http://arxiv-web3.library.cornell.edu/pdf/1406.7178v1.pdf>.
- [26] Rockafellar RT (1981) *The Theory of Subgradients and Its Applications to Problems of Optimization*, R&E, Vol. 1 (Heldermann Verlag, Berlin).
- [27] Rockafellar RT, Wets RJ-B (1998) *Variational analysis*. Grundlehren der mathematischen Wissenschaften (Fundamental Principles of Mathematical Sciences), Vol. 317 (Springer, Berlin).
- [28] Smith NA, Tromble RW (2004) Sampling uniformly from the unit simplex. Technical report, Johns Hopkins University, Baltimore.
- [29] Takeda A, Niranjan M, Gotoh J, Kawahara Y (2012) Simultaneous pursuit of out-of-sample performance and sparsity in index tracking portfolios. *Comput. Management Sci.* 10(1):21–49.
- [30] Tropp JA, Wright SJ (2010) Computational methods for sparse solution of linear inverse problems. *Proc. IEEE* 98(6):948–958.