

# The Minimization of Piecewise Functions: Pseudo Stationarity

**Ying Cui**

*Dept. of Industrial and Systems Engineering, University of Minnesota, Minneapolis, U.S.A.  
yingcui@umn.edu*

**Junyi Liu**

*Department of Industrial Engineering, Tsinghua University, Beijing, China  
junyiliu@mail.tsinghua.edu.cn*

**Jong-Shi Pang**

*The Daniel J. Epstein Department of Industrial and Systems Engineering,  
University of Southern California, Los Angeles, U.S.A.  
jongship@usc.edu*

*Dedicated to Roger J-B Wets on the occasion of his 85th birthday.*

Received: February 2, 2022

Accepted: July 4, 2022

There are many significant applied contexts that require the solution of discontinuous optimization problems in finite dimensions. Yet these problems are very difficult, both computationally and analytically. With the functions being discontinuous and a minimizer (local or global) of the problems, even if it exists, being impossible to verifiably compute, a foremost question is what kind of “stationary solutions” one can expect to obtain; these solutions provide promising candidates for minimizers; i.e., their defining conditions are necessary for optimality. Motivated by recent results on sparse optimization, we introduce in this paper such a kind of solution, termed “pseudo B- (for Bouligand) stationary solution”, for a broad class of discontinuous optimization problems with objective and constraint defined by indicator functions of the positive real axis composite with functions that are possibly nonsmooth. We present two approaches for computing such a solution. One approach is based on lifting the problem to a higher dimension via the epigraphical formulation of the indicator functions; this requires the addition of some auxiliary variables. The other approach is based on certain continuous (albeit not necessarily differentiable) piecewise approximations of the indicator functions and the convergence to a pseudo B-stationary solution of the original problem is established. The conditions for convergence are discussed and illustrated by an example.

*Keywords:* Piecewise functions, stationarity, epigraphical formulation, nonifier approximation.

*2010 Mathematics Subject Classification:* 90C26, 90C59.

## 1. Introduction

There are many significant applied contexts that require the solution of discontinuous optimization problems in finite dimensions. Yet these problems are very difficult, both computationally and analytically. Like all nonconvex problems, the

The work of the third author was based on research supported by the U.S. Air Force Office of Sponsored Research under grants FA9550-18-1-0382 and FA9550-22-1-0045.

computational task of a minimizer, local or global, is prohibitively challenging, if not impossible. Thus a realistic goal is to try to compute a solution that is a promising candidate for a minimizer, in particular, one that satisfies some necessary conditions of a local minimizer, i.e., a stationary solution of some sort. There is a very long tradition of investigation of such conditions, starting from the classical one of a zero objective gradient for an unconstrained differentiable optimization problem to the use of some advanced subdifferentials from variational analysis [49] to obtain a set-inclusion problem. With the functions involved being discontinuous, an insightful understanding, constructive characterization, and profitable employment of the latter subdifferentials all are not easy tasks, especially when there are constraints that are embedded in the objective function taking on infinite value.

As an alternative, the idea of approximating the discontinuous functions by smooth, say continuously differentiable, functions appears promising. One of the first smoothing methods for optimization of discontinuous functions was proposed in [27]. In general, such a method involves smoothing a locally integrable function by mollifier approximations using integration. A sequence of the resulting smoothed problems is then solved and their stationary solutions are used to define a stationarity concept for the original problem. Studied comprehensively in the seminal paper [18] under the framework of minimizing an extended-valued function, this approach leads to the definition of “mollifier subgradients” that serve as the target of a computational resolution for a discontinuous optimization problem. See [34, 35] for the applications of mollifier induced smoothing functions, called “averaged function” in [18, Definition 3.1], to nonlinear and semi-infinite programming, respectively. The paper [7] gives a survey of smoothing methods for nonconvex, nonsmooth optimization problems with the underlying functions being continuous. It is important to point out while the theory in [18, 7] is quite general, the applied problems surveyed in [7] are all of the kind of a univariate nonsmooth function composite with a smooth function; for these composite functions, smoothing is applied to the univariate component while the composition is maintained. Thus it is fair to say that to date, the practical aspect of smoothing is restricted to this class of composite functions where the convolutional operation requires only the evaluation of integrals of scalar functions of one real variable; the approach becomes most effective when the resulting averaged functions are explicitly available for computational purposes with the convolution operation staying in the background for general analysis.

Supported by diverse source problems, this paper studies a broad class of discontinuous optimization problems with objective and constraints defined by piecewise functions modeled as the products of nonconvex nonsmooth functions and the indicator functions of the positive/nonnegative real axis, called *Heaviside functions*. Per a referee, these functions first appeared in the work of Guillaume Libri [39, 40]. Oliver Heaviside (1850–1925) popularized these functions in his pioneering work on differential equations for the study of electromagnetic waves [29]. The “open” Heaviside function is the indicator of the open interval  $(0, \infty)$ . For our purpose, we include the indicator of the closed interval  $[0, \infty)$  also as a Heaviside function. With the advance of nonconvex nonsmooth optimization as documented in [12], our goal in approximating these problems is to least disrupt the nonsmoothness and nonconvexity of the given functions and do so only when needed, for instance in designing

practical computational methods. Thus, unlike the previous work [3, 16, 17, 18, 27] that smooth the discontinuous Heaviside function, the approximating problems developed in the present paper remain nonsmooth and nonconvex but are solvable by various surrogation methods; see [12, Chapter 7]. In addition to this departure from the smooth approximations, we adopt the approach to define an approximation-independent and subdifferential-free stationarity condition, called *pseudo B- (for Bouligand) stationarity*, that is necessarily satisfied by a local minimizer of the given discontinuous optimization problem. The definition is motivated by a recent study [25] of the  $\ell_0$ -optimization problem that lies at the center of sparse optimization in statistical estimation [28]. The well-known  $\ell_0$ -function is defined by:

$$|t|_0 \triangleq \begin{cases} 1 & \text{if } t \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

We describe two constructive approaches for computing/approximating such a pseudo B-stationary solution. Omitting the details that are left for a follow-up algorithmic study, we emphasize that these approaches can be computationally implemented in practice by difference-of-convex programming based algorithms [37, 42, 44, 45, 47] when the involved functions are of this kind, and more generally, by surrogation methods [12, Chapter 7] for broad classes of nonsmooth nonconvex optimization problems.

The rest of the paper is organized as follows. The next section begins with the formal definition of the problem to be studied followed by some preliminary remarks. Section 3 presents a host of discontinuous piecewise functions arising from various optimization contexts that are unified by our central problem. A main result in Section 4 identifies a principal sign condition that plays a central role throughout the paper; this result connects our problem with two related problems in the literature in terms of their global minima. As a remedy to the computational intractability of these global minima, Section 5 defines a pseudo B-stationarity concept that is amenable to computation. Two constructive approaches for computing such a stationary solution is presented in Sections 6 and 8, respectively. In particular, the approximation approach described in the last section can be traced back to a pioneering paper by Roger Wets and his collaborators that we expand in Section 7 and to whom we dedicate our work.

## 2. Problem definition and preliminary discussion

Consider the following piecewise optimization problem:

$$\begin{aligned} \underset{x \in X}{\text{minimize}} \quad & \Phi(x) \triangleq c(x) + \sum_{k=1}^K \varphi_k(x) \mathbf{1}_{(0,\infty)}(g_k(x)) \\ \text{subject to} \quad & \sum_{\ell=1}^L \phi_\ell(x) \mathbf{1}_{(0,\infty)}(h_\ell(x)) \leq b, \quad \text{called the functional constraint,} \end{aligned} \tag{1}$$

where  $\mathbf{1}_{(0,\infty)}$  is the “open” Heaviside function given by

$$\mathbf{1}_{(0,\infty)}(s) \triangleq \begin{cases} 1 & \text{if } s \in (0, \infty) \\ 0 & \text{otherwise} \end{cases}$$

and the following holds [the polyhedrality of  $X$  is not needed in several results; it is stated as a blanket assumption primarily to avoid the use of advanced constraint qualifications in the context of tangent cones]:

**Blanket assumption:**  $X$  is a polyhedron contained in the open subset  $\mathcal{O}$  of  $\mathbb{R}^n$  and  $c$ ,  $\{\varphi_k, g_k\}_{k=1}^K$ , and  $\{\phi_\ell, h_\ell\}_{\ell=1}^L$  are B-differentiable (where B is for Bouligand) functions from  $\mathcal{O} \rightarrow \mathbb{R}$ .  $\square$

By definition, a function  $\psi: \mathcal{O} \rightarrow \mathbb{R}$  is *B-differentiable* [12, Definition 4.1.1] at  $\bar{x} \in \mathcal{O}$  if  $\psi$  is locally Lipschitz continuous at  $\bar{x}$  (i.e., Lipschitz continuous in an open neighborhood of  $\bar{x}$ ) and directionally differentiable there; i.e., the elementary one-sided directional derivative

$$\psi'(\bar{x}; v) \triangleq \lim_{\tau \downarrow 0} \frac{\psi(\bar{x} + \tau v) - \psi(\bar{x})}{\tau}$$

exists for all  $v \in \mathbb{R}^n$ . We make several immediate remarks about the formulation (1); foremost is to note that the functional constraint adds considerable challenges to this problem, without which the analysis simplifies somewhat. Other remarks are as follows:

- While the open Heaviside function  $\mathbf{1}_{(0, \infty)}(\bullet)$  is lower semicontinuous on the real line, a product such as  $\varphi_k(\bullet) \mathbf{1}_{(0, \infty)}(g_k(\bullet))$  may not be lower semicontinuous at  $\bar{x} \in g_k^{-1}(0)$  unless  $\varphi_k$  is nonnegative there. For the problem (1), this sign condition on the pairs of functions  $\{\varphi_k, g_k\}_{k=1}^K$  and  $\{\phi_\ell, h_\ell\}_{\ell=1}^L$  will persist throughout the paper. This restriction is responsible for the closedness of the feasible region and for the existence of minimizers of the problem, and thus for their relaxations, such as that of a pseudo B-stationary point to be defined later.
- The well-known  $\ell_0$ -function  $|\bullet|_0$  in sparsity estimation [28] can be written as:

$$|s|_0 = \mathbf{1}_{(0, \infty)}(s) + \mathbf{1}_{(0, \infty)}(-s) = \mathbf{1}_{(0, \infty)}(|s|). \quad (2)$$

This is a prominent applied instance of the open Heaviside function and provides much motivation for the theory developed in this paper. Conversely, the developed theory also provides a deeper understanding of the sparse optimization problem as a result of (2); namely, any approximation of the open Heaviside function can immediately be specialized to the  $\ell_0$ -function.

- A term  $\psi(x) \mathbf{1}_{[0, \infty)}(f(x))$  involving the “closed” Heaviside function

$$\mathbf{1}_{[0, \infty)}(s) \triangleq \begin{cases} 1 & \text{if } s \in [0, \infty) \\ 0 & \text{otherwise} \end{cases}$$

can be written as:  $\psi(x) \mathbf{1}_{[0, \infty)}(f(x)) = \psi(x) - \psi(x) \mathbf{1}_{(0, \infty)}(-f(x))$ .

So the formulation (1) encompasses products of this kind that involves the closed Heaviside function. Subsequently, conditions imposed on the functions in (1) easily translate to corresponding conditions for functions involving the closed Heaviside function via the above relation.

In general, a locally Lipschitz function does not need to be directionally differentiable; however, for a univariate function  $f: (-\ell, u) \rightarrow \mathbb{R}$  defined on an interval, if

$f$  is locally Lipschitz and monotone (i.e., nondecreasing or nonincreasing), then the one-sided derivatives:

$$f'(t; \pm 1) \triangleq \lim_{\tau \downarrow 0} \frac{f(t \pm \tau) - f(t)}{\tau}, \quad t \in (-\ell, u) \quad (3)$$

exist with signs determined by the monotonicity. For an integer  $N > 0$ , we let  $[N] \triangleq \{1, \dots, N\}$ .

### 3. Some source problems

The product functions in (1) are discontinuous piecewise functions; they include several interesting special cases, which we highlight below. These cases illustrate the versatility of the Heaviside functions in modeling a host of discontinuous functions in diverse contexts.

- *Cost-efficient variable selection:* This is an extension of the sparsity optimization problem in statistical estimation [28] in which there is a cost associated the collection of data in the modeling process [5, 56, 57] which we may formulate as either a soft penalty embedded in the objective function:

$$\underset{x \in X \subseteq \mathbb{R}^n}{\text{minimize}} \quad e(x) + \lambda \sum_{i=1}^n c_i |x_i|_0, \quad (4)$$

or as a budget-type constraint:

$$\underset{x \in X \subseteq \mathbb{R}^n}{\text{minimize}} \quad e(x) \quad \text{subject to} \quad \sum_{i=1}^n c_i |x_i|_0 \leq b, \quad (5)$$

where  $e(x)$  is a loss function,  $\lambda > 0$  is a given parameter, and the coefficients  $c_i$  and the right-hand constant  $b$  are positive. In addition to the application in modern-day sparse optimization, the term  $c_i |x_i|_0$  is a classical modeling device in operations research applications as a set-up cost of an activity. Namely, a cost is incurred when there is a nonzero level of the  $i$ th activity and zero otherwise. The functional constraint expresses the available budget for the set-up or variable-selection costs applicable in both the data or operational contexts. A cost associated with the level of the activity is included in either the objective term  $e(x)$  or the constraint  $X$ . One mathematical feature of the summation term in both formulations is worth noting: namely, all the coefficients associated with the  $\ell_0$ -functions are positive, thus, satisfying the sign condition mentioned before. These problems suggest that the multiplicative functions  $\varphi_k(x)$  and  $\phi_\ell(x)$  in the general problem (1) may be interpreted as costs in some applied models.

- *Piecewise functions on complementary regions:* Consider a simple example of a piecewise function which may or may not be continuous:

$$\Psi(x) = \begin{cases} \psi_1(x) & \text{if } a \leq f(x) \leq b \\ \psi_2(x) & \text{if } f(x) < a \\ \psi_3(x) & \text{if } f(x) > b, \end{cases} \quad (6)$$

for some scalars  $a$  and  $b$  satisfying  $-\infty \leq a < b \leq \infty$ .

We then have

$$\begin{aligned}\Psi(x) &= \psi_1(x) \mathbf{1}_{[0,\infty)}(\min(b - f(x), f(x) - a)) \\ &\quad + \psi_2(x) \mathbf{1}_{(0,\infty)}(a - f(x)) + \psi_3(x) \mathbf{1}_{(0,\infty)}(f(x) - b) \\ &= \psi_1(x) - \psi_1(x) \mathbf{1}_{(0,\infty)}(\max(f(x) - b, a - f(x))) \\ &\quad + \psi_2(x) \mathbf{1}_{(0,\infty)}(a - f(x)) + \psi_3(x) \mathbf{1}_{(0,\infty)}(f(x) - b),\end{aligned}$$

whose validity is regardless of the continuity of  $\Psi$  on the sets  $f^{-1}(a)$  and  $f^{-1}(b)$ . An interesting application of this class of piecewise function concerns the constrained optimization with (soft) penalty. Specifically, let's say that we wish to minimize a function  $\psi_1(x)$  when the constraint  $f(x) \geq 0$  is satisfied and there is a penalty  $\psi_2(x)$  when the constraint is not satisfied. This problem can be formulated as minimizing the combined objective:  $\psi_1(x) \mathbf{1}_{[0,\infty)}(f(x)) + \psi_2(x) \mathbf{1}_{(0,\infty)}(-f(x))$ . A special case of (6) is when the three component functions  $\psi_i(x)$  are constants. More generally, a quantized function is a discontinuous step function that can very easily be described by a generalization of (6) with arbitrary (finite) number of mutually disjoint regions (intervals in the case of a univariate variable) within each of which the overall function is a constant. The references [41, 54, 55] have employed such quantized minimization problems for the training of deep neural networks.

- *Best constraint selection:* In contrast to the best variable selection problems (4) or (5), and yet so far at best minimally studied in the literature, the best constraint selection is a generalization of best variable selection and can be modeled using the summation:  $\sum_{k=1}^K \mathbf{1}_{[0,\infty)}(f_k(x))$ , where the family  $\{f_k(x) \geq 0\}_{k=1}^K$  consists of the constraints to be selected. A related problem is when there is a cost  $c_k(x)$  associated with the constraint  $f_k(x) \geq 0$  being disrupted. This problem can be formulated as minimizing the weighted sum:  $\sum_{k=1}^K c_k(x) \mathbf{1}_{(0,\infty)}(-f_k(x))$  or imposing a constraint defined by the sum. Our work offers a pathway for the potential applications of best constraint selection in instances like these.

- *On-off constraints by indicator variables:* In the literature such as [6, 30], such a constraint is  $f_k(x) \geq 0$  if  $y_k = 1$  where  $y_k \in \{0, 1\}$  is a binary variable that appears only in the constraint and for the sole purpose of turning on the constraint  $f_k(x) \geq 0$ . Equivalently, such an on-off constraint is equivalent to:  $f_k(x) \geq 0$  if  $y_k > 0$  with  $y_k$  restricted to be a continuous variable in the interval  $[0, 1]$ . In turn, the latter constraint is equivalent to  $y_k f_k(x) \geq 0$  with  $y_k \in [0, 1]$ , and instead of the constraint, a term  $\mathbf{1}_{[0,\infty)}(y_k f_k(x))$  can be added to the objective as part of the overall function to be minimized.

- *Binary classifications by the sign function:* In binary classification, it is customary to use the sign function to separate two classes. For instance, given a binary number  $\sigma = \pm 1$  denoting classes A and B, respectively and with  $f(x)$  as the feature-dependent classification function, we classify the outcome from this function as A if  $f(x) > 0$  and B if  $f(x) < 0$ . This classification scheme can be modeled by the composite indicator function:  $\mathbf{1}_{(0,\infty)}(-\sigma f(x))$ , which aims to count the number of misclassifications among the observational data. See [47] for a recent application of this formulation in the context of individualized decision making under uncertainty for medical treatment. Classification with margin is an extension that allows for minor errors; it classifies the outcome as A if  $f(x) \geq \varepsilon$  and B if  $f(x) \leq -\varepsilon$ , where

$\varepsilon > 0$  is a small margin. This can be modeled by  $\mathbf{1}_{(0,\infty)}(\varepsilon - \sigma f(x))$ , which again counts the number of misclassifications.

- *Products of indicators:* It is clear that

$$(\mathbf{1}_{[0,\infty)}(f(x))) (\mathbf{1}_{[0,\infty)}(g(x))) = \mathbf{1}_{[0,\infty)}(\min(f(x), g(x)));$$

thus products of closed Heaviside functions can be combined into a single Heaviside function of the same kind via the use of the pointwise minimum operator. Less clear is the product of a closed and an open Heaviside function, say

$$(\mathbf{1}_{[0,\infty)}(f(x))) (\mathbf{1}_{(0,\infty)}(g(x))).$$

Nevertheless, we have

$$\begin{aligned} (\mathbf{1}_{[0,\infty)}(f(x))) (\mathbf{1}_{(0,\infty)}(g(x))) &= (\mathbf{1}_{(0,\infty)}(g(x))) [1 - (\mathbf{1}_{(0,\infty)}(-f(x)))] \\ &= \mathbf{1}_{(0,\infty)}(g(x)) - \mathbf{1}_{(0,\infty)}(\min(-f(x), g(x))), \end{aligned}$$

which is the difference of two open Heaviside functions. An example where a product of an open and a closed Heaviside function may occur is a modification of the piecewise function (6): say  $\psi(x) = \psi_1(x)$  if  $a \leq f(x) < b$ . We see that

$$\begin{aligned} \psi(x) &= \psi_1(x) (\mathbf{1}_{[0,\infty)}(f(x) - a)) (\mathbf{1}_{(0,\infty)}(b - f(x))) \\ &= \psi_1(x) [\mathbf{1}_{(0,\infty)}(b - f(x)) - \mathbf{1}_{(0,\infty)}(a - f(x))]. \end{aligned}$$

- *Probabilistic functions and conditional expectations:* The use of the Heaviside functions in equivalent formulations of probabilistic functions is well known; see [11] for a recent comprehensive study of a nonconvex nonsmooth approach for chance-constrained stochastic programs that is built on this fundamental formulation. The simplest example is the probabilistic function  $\mathbb{P}_{\tilde{z}}(\Psi(x, \tilde{z}) \geq 0)$ , for some bivariate function  $\Psi(x, z)$ . In the cited reference, we have considered a broad class of such functions defined as the difference of two convex functions each being the pointwise maximum of finitely many convex smooth functions. An abstraction of such a difference-of-convex function  $\Psi(\bullet, z)$  is a piecewise function of which the following bivariate extension of (6) is an example:

$$\Psi(x, z) \triangleq \begin{cases} \psi_1(x, z) & \text{if } a \leq f(x, z) \leq b \\ \psi_2(x, z) & \text{if } f(x, z) < a \\ \psi_3(x, z) & \text{if } f(x, z) > b. \end{cases}$$

Appendix B of the paper [10] contains many piecewise functions of the above kind arising from individualized decision-making under partial identification. Such a piecewise random functional gives rise to an expectation of piecewise functions such as  $\mathbb{E}_{\tilde{z}}[\Psi(x, \tilde{z})]$  with  $\Psi$  given above.

Additionally, consider the conditional expectation:

$$\begin{aligned} \mathbb{E}_{\tilde{z}}[\phi(x, \tilde{z}) \mid f(x, \tilde{z}) \leq 0] &\triangleq \frac{\mathbb{E}_{\tilde{z}}[\phi(x, \tilde{z}) \mathbf{1}_{(-\infty, 0]}(f(x, \tilde{z}))]}{\mathbb{P}_{\tilde{z}}(f(x, \tilde{z}) \leq 0)} \\ &= \frac{\mathbb{E}_{\tilde{z}}[\phi(x, \tilde{z}) \mathbf{1}_{(-\infty, 0]}(f(x, \tilde{z}))]}{\mathbb{E}_{\tilde{z}}[\mathbf{1}_{(-\infty, 0]}(f(x, \tilde{z}))]}, \end{aligned}$$

which again involves the closed Heaviside function. Conditional expectations have applications in modeling low probability–high consequence accidents [51] and in the semi-supervised structured classification problem [59]. The understanding and treatment of the deterministic problem (1) provides useful insights for studying stochastic programs with composite indicator functions such as the optimization with chance constraints involving discontinuous piecewise functions and conditional expectation functions that lead to fractional expectation functions, where the Heaviside functions are prominently present.

#### 4. Equivalent formulations: Global minimizers

Before discussing stationarity solutions and local minimizers, we introduce classes of the problem (1) within which a (globally) optimal solution exists and there is an equivalence of the problem with several lifted formulations employing additional variables and special constraints. Equivalence means that there is a one-to-one correspondence between the optimal solutions of the problems and their optimal objective values are equal. The purpose of this section is twofold: (a) to show that the problem (1) has an optimal solution and is related to two well-studied problems in the literature under some sign restrictions on the family of functions  $\{\varphi_k, g_k\}_{k=1}^K$  and  $\{\phi_\ell, h_\ell\}_{\ell=1}^L$ , and (b) to support the sign conditions as a reasonable assumption to be imposed in subsequent sections. Since these equivalent formulations are for global minimizers whose computation is practically prohibitive if not impossible, and since they are not the focus in the later sections, we present the result below without the Heaviside constraint; i.e., for the following problem only:

$$\underset{x \in X}{\text{minimize}} \quad \Phi(x) \triangleq c(x) + \sum_{k=1}^K \varphi_k(x) \mathbf{1}_{(0,\infty)}(g_k(x)). \quad (7)$$

In the proof, we let  $(\bullet)_{\min}$  denote the minimum objective value of the referenced problem;  $(t)_{\pm} \triangleq \max(\pm t, 0)$  be the nonnegative and nonpositive part of a scalar  $t$ ; and let  $\perp$  denote the perpendicularity notation, which in the present context means the complementary slackness between the involved expressions.

**Proposition 4.1.** *Let  $X$  be a compact set. Suppose that the functions  $c$  and  $\{\varphi_k, g_k\}_{k=1}^K$  are continuous. The following two statements hold:*

(A) *Under the following sign restriction:*

- *for every  $k \in [K]$ , the function  $\varphi_k$  is nonnegative on the set  $X \cap g_k^{-1}(0)$ , the problem (7) has an optimal solution; moreover, it is equivalent to:*

— **MPCC-1:**

$$\underset{x \in X; s}{\text{minimize}} \quad \Phi_{\text{MPCC1}}(x, s) \triangleq c(x) + \sum_{k=1}^K \varphi_k(x) s_k \quad (8)$$

$$\text{subject to} \quad 0 \leq s_k \perp (g_k(x))_- \geq 0, \quad \forall k \in [K] \quad (9)$$

$$\text{and} \quad 0 \leq 1 - s_k \perp (g_k(x))_+ \geq 0, \quad \forall k \in [K]. \quad (10)$$

(B) *Under the following strengthened sign restriction:*

- *for every  $k \in [K]$ , the function  $\varphi_k$  is nonnegative on the set  $X \cap g_k^{-1}(-\infty, 0]$ ,*

the problem (7) is equivalent to either one of the following two problems:

— **MPCC-2:**

$$\begin{aligned}
 & \underset{x \in X; s}{\text{minimize}} && \Phi_{\text{MPCC2}}(x, s) \triangleq c(x) + \sum_{k=1}^K \varphi_k(x) s_k \\
 & \text{subject to} && 0 \leq 1 - s_k \perp (g_k(x))_+ \geq 0, \quad \forall k \in [K] \\
 & \text{and} && s \in [0, 1]^K \quad (\text{continuous variables});
 \end{aligned} \tag{11}$$

(note the absence of the constraint  $s_k(g_k(x))_- = 0$  compared to (8));

— **on-off constraints:**

$$\begin{aligned}
 & \underset{x \in X; z}{\text{minimize}} && \Phi_{\text{on/off}}(x, z) \triangleq c(x) + \sum_{k=1}^K \varphi_k(x) (1 - z_k) \\
 & \text{subject to} && g_k(x) \leq 0 \quad \text{if } z_k = 1, \quad \forall k \in [K] \\
 & \text{and} && z \in \{0, 1\}^K \quad (\text{binary variables}).
 \end{aligned} \tag{12}$$

**Proof.** We claim that the objective  $\Phi$  is lower semicontinuous on  $X$  by showing that each product  $\varphi_k(x) \mathbf{1}_{(0, \infty)}(g_k(x))$  is lower semicontinuous on  $X$  under the sign assumption on the element functions  $\{\varphi_k\}_{k=1}^K$ . For this purpose, it suffices to show if  $\varphi$  and  $g$  are two continuous functions such that  $\varphi$  is nonnegative on  $X \cap g^{-1}(0)$ , then the level set:

$$L(\alpha) \triangleq \{x \in X \mid \varphi(x) \mathbf{1}_{(0, \infty)}(g(x)) \leq \alpha\}$$

is closed for all scalars  $\alpha \in \mathbb{R}$ . Let  $\{x^\nu\}$  be a given sequence in  $L(\alpha)$  converging to a limit  $\bar{x} \in X$ . There are 2 cases to consider:

- $\alpha \geq 0$ : We must have for every  $\nu$ , either  $g(x^\nu) \leq 0$  or  $[g(x^\nu) > 0 \text{ and } \varphi(x^\nu) \leq \alpha]$ . Then the limit  $\bar{x}$  satisfies: either  $g(\bar{x}) \leq 0$  or  $[g(\bar{x}) > 0 \text{ and } \varphi(\bar{x}) \leq \alpha]$ . This shows that  $\bar{x} \in L(\alpha)$ , without requiring the sign restriction on  $\varphi$ .
- $\alpha < 0$ : We must have  $[g(x^\nu) > 0 \text{ and } \varphi(x^\nu) \leq \alpha]$ . By the sign assumption on  $\varphi$ , the limit  $\bar{x}$  must satisfy:  $g(\bar{x}) > 0$  and  $\varphi(\bar{x}) \leq \alpha$ ; so  $\bar{x} \in \mathcal{L}(\alpha)$ , completing the proof of the closedness of  $L(\alpha)$ .

Consequently, the problem (7) is a minimization problem of a lower semicontinuous function on a compact set; thus it has an optimal solution. Omitting the proof of equivalence with the problem (8), we directly show the equivalence of the two problems (7) and (11). Let  $x \in X$  be arbitrary. Then the pair  $(x, s)$  is feasible to (11), where  $s_k \triangleq \mathbf{1}_{(0, \infty)}(g_k(x))$  for all  $k \in [K]$ . Hence the minimum objective value of (7) is no less than that of (11). Conversely, if  $(x, s)$  is feasible to (11), then  $\Phi_{\text{MPCC2}}(x, s) \geq \Phi(x)$  by the sign restriction of  $\varphi_k$  on  $X \cap g_k^{-1}(-\infty, 0]$ . Therefore, equality holds and an optimal solution of one problem readily yields an optimal solution of the other. To show the equivalence of (7) and (12), let  $x \in X$  be arbitrary. Then the pair  $(x, z)$  is feasible to (12), where  $1 - z_k \triangleq \mathbf{1}_{(0, \infty)}(g_k(x))$  for all  $k \in [K]$ . Hence  $(7)_{\min} \geq (12)_{\min}$ . Conversely, if  $(x, z)$  is feasible to (12), then  $\Phi_{\text{on/off}}(x, z) \geq \Phi(x)$ . Hence equality holds and there is a one-to-to correspondence between the optimal solutions of these two problems.  $\square$

Being instances of a mathematical programming problem with complementarity constraints (MPCC), the formulations (8) and (11) are similar to the equivalent formulations of an  $\ell_0$ -minimization problem used in [22]. These two MPCC's (8) and (11) are special instances of a mathematical program with vanishing constraints for which there is an extensive literature; see [1, 14, 32, 33] which contain many more references. With an emphasis toward more general constraints, these references have paid significant attention to constraint qualifications [24] and their consequences for stationarity conditions and regularization methods. In contrast, our study bypasses such complementarity constraint qualifications and aims to analyze a kind of stationary solutions defined directly on the discontinuous piecewise problem (1).

The formulation (12) is one with on-off constraints described by indicator variables [6, 30]. The method of proof of the equivalence of the two problems (11) and (12) is fairly elementary; the same methodology is also discussed and used, e.g., in [36, 43].

The principal purpose of presenting Proposition 4.1 is to highlight the two sign assumptions in (A) and (B). They motivate similar assumptions to be imposed later. The sign conditions can be constructively verified (if desired) when the functions  $\varphi_k$  are convex and  $g_k$  are affine (for (A)) or convex (for (B)). In general, these assumptions can be formulated equivalently as the optimum objective values of the minimization problems being nonnegative:

$$\underset{x \in X}{\text{minimize}} \varphi_k(x) \text{ subject to } g_k(x) = 0 \mid \underset{x \in X}{\text{minimize}} \varphi_k(x) \text{ subject to } g_k(x) \leq 0,$$

respectively. These are convex programs under stated stipulations of  $g_k$  and  $\varphi_k$ . Of course, there are applied situations where these sign conditions hold easily.

## 5. Pseudo Bouligand stationarity

With the discontinuity of the (open) Heaviside function it is not easy to derive transparent necessary conditions for a local minimizer of the problem (1), let alone computing it. Although there are various one-sided directional derivatives (such as those of the Dini kind), and more generally, the subderivatives in modern variational analysis [49] that one may apply to the objective function  $\Phi(x)$ , these derivatives are defined for general functions; in particular, their specializations to functions such as  $\Phi(x)$  in (1) do not immediately yield useful insights about the problem without carefully unwrapping the details of the derivatives. Instead, our approach herein is based on elementary one-sided directional derivatives and basic optimization theory. One immediate benefit of our approach is that it handles constraints with the Heaviside functions at much ease, unlike the variational approach which requires the constraints to be converted to extended-valued functions embedded in the objective.

Another important point to make is that we aim to connect the theory with computations; namely, we want to ensure two goals of the defined solution concepts: (i) they are computationally achievable, and (ii) they are not over-relaxed. As the background for the subsequent Definition 5.2, we provide a result [25, Proposition 3] that was stated for a global minimizer in the reference, which we broaden to allow for a local minimizer.

**Proposition 5.1.** *Let  $X \subseteq \mathbb{R}^n$ ,  $\theta: \mathcal{O} \supseteq X \rightarrow \mathbb{R}$ , and  $\gamma > 0$  be given. If  $\bar{x}$  is a global (local) minimizer of the problem:*

$$\underset{x \in X}{\text{minimize}} \quad \theta(x) + \gamma \sum_{i=1}^n |x_i|_0, \quad (13)$$

*then  $\bar{x}$  is a global (local, respectively) minimizer of the problem:*

$$\underset{x \in X}{\text{minimize}} \quad \theta(x) \quad \text{subject to} \quad x_{\mathcal{A}_0(\bar{x})} = 0, \quad (14)$$

where  $\mathcal{A}_0(\bar{x}) \triangleq \{i \in [n] \mid \bar{x}_i = 0\}$ . □

Restricted to problems with  $\theta$  convex, a vector  $\bar{x}$  that is a global minimizer of (14) is termed a “pseudo-minimizer” of (13) in the reference. It turns out that such minimizers are the points of attraction of sequences of stationary solutions of the “folded concave” approximations of the  $\ell_0$ -problem; see [25, Proposition 5]. The main idea of the problem (14) to deal with the discontinuous function  $|\bullet|_0$  is to transfer, or “pull down” the discontinuity of the objective to a constraint. We plan to apply this idea to the problem (1); nevertheless the resulting “pulled-down” extension of (14) will have nonconvex objectives and constraints. This necessitates us to recall the definition of a Bouligand stationary (B-stationary in short) solution of a B-differentiable program. For an abstract optimization problem:

$$\underset{x \in \hat{X}}{\text{minimize}} \quad \theta(x),$$

where  $\hat{X}$  is a closed set in  $\mathbb{R}^n$  and  $\theta: \mathcal{O} \rightarrow \mathbb{R}$  is a B-differentiable function, a vector  $\bar{x} \in \hat{X}$  is a Bouligand stationary solution of  $\theta$  on  $\hat{X}$  [12, Definition 6.1.1] if  $\theta'(\bar{x}; v) \geq 0$  for all  $v \in \mathcal{T}(\hat{X}; \bar{x})$ , where

$$\mathcal{T}(\hat{X}; \bar{x}) \triangleq \left\{ v \in \mathbb{R}^n \left| \begin{array}{l} \exists \{x^\nu\} \subset \hat{X} \text{ converging to } \bar{x} \text{ and} \\ \{\tau_\nu\} \downarrow 0 \text{ such that } v = \lim_{\nu \rightarrow \infty} \frac{x^\nu - \bar{x}}{\tau_\nu} \end{array} \right. \right\}$$

is the tangent cone of  $\hat{X}$  at  $\bar{x}$ . If  $\hat{X}$  is additionally convex, B-stationarity reduces to d-stationary (“d” for directional), which is:  $\theta'(\bar{x}; x - \bar{x}) \geq 0$  for all  $x \in \hat{X}$ .

For the problem (1), we define the following six index sets corresponding to a given  $\bar{x} \in X$ :

$$\begin{aligned} \mathcal{K}_>(\bar{x}) &\triangleq \{k \in [K] \mid g_k(\bar{x}) > 0\}; & \mathcal{K}_=(\bar{x}) &\triangleq \{k \in [K] \mid g_k(\bar{x}) = 0\} \\ \mathcal{K}_<(\bar{x}) &\triangleq \{k \in [K] \mid g_k(\bar{x}) < 0\}; & \mathcal{L}_>(\bar{x}) &\triangleq \{\ell \in [L] \mid h_\ell(\bar{x}) > 0\} \\ \mathcal{L}_=(\bar{x}) &\triangleq \{\ell \in [L] \mid h_\ell(\bar{x}) = 0\}; & \mathcal{L}_<(\bar{x}) &\triangleq \{\ell \in [L] \mid h_\ell(\bar{x}) < 0\}. \end{aligned}$$

With these index sets, we define the “pulled-down”, or “pseudo stationarity” problem at  $\bar{x}$ :

$$\begin{aligned}
& \underset{x}{\text{minimize}} \quad \Phi(x; \bar{x}) \triangleq c(x) + \sum_{k \in \mathcal{K}_>(\bar{x})} \varphi_k(x) \\
& \text{subject to} \quad \left. \begin{aligned}
& x \in X; \quad \sum_{\ell \in \mathcal{L}_>(\bar{x})} \phi_\ell(x) \leq b \\
& g_k(x) \leq 0 \quad \forall k \in \mathcal{K}_=(\bar{x}) \cup \mathcal{K}_<(\bar{x}) \triangleq \mathcal{K}_\leq(\bar{x}) \\
& g_k(x) \geq 0 \quad \forall k \in \mathcal{K}_>(\bar{x}) \\
& h_\ell(x) \leq 0 \quad \forall \ell \in \mathcal{L}_=(\bar{x}) \cup \mathcal{L}_<(\bar{x}) \triangleq \mathcal{L}_\leq(\bar{x}) \\
& \text{and} \quad h_\ell(x) \geq 0 \quad \forall \ell \in \mathcal{L}_>(\bar{x})
\end{aligned} \right\} \begin{array}{l} \text{feasible} \\ \text{set} \\ \text{denoted} \\ S_{\text{ps}}(\bar{x}). \end{array} \quad (15)
\end{aligned}$$

Quite different from the three problems in Proposition 4.1, the above problem is obtained by pulling the indicator functions out of the objective function and the functional constraint and imposing constraints restricting the variable  $x$  according to those defined by the functions  $\{g_k, h_\ell\}$  and satisfied by the vector  $\bar{x}$  on hand. Specialized to the  $\ell_0$ -optimization problem (13), the above pulled-down problem reduces to (14). Clearly, if  $\bar{x}$  is a globally optimal solution to (1) and provided that  $\varphi_k$  is nonnegative on  $X \cap g_k^{-1}(0)$  for  $k \in \mathcal{K}_>(\bar{x})$  and  $\phi_\ell$  is nonnegative on  $X \cap h_\ell^{-1}(0)$  for  $\ell \in \mathcal{L}_>(\bar{x})$ , then  $\bar{x}$  is globally optimal for (15); however, the converse is generally not true because  $S_{\text{ps}}(\bar{x})$  is at best only a subset of the feasible set of (1). This failed equivalence is not surprising for the main reason that (15) is defined at a given  $\bar{x}$ ; this definition is for the purpose of answering the question: what is a necessary condition for a given feasible vector of (1) to be its local minimizer?

The following definition specifies a pseudo solution of (1) as a fixed point of the self-defined “locmin” or “B-stationarity” mapping. Note that with  $\bar{x}$  given, problem (15) is a standard nonlinear program; thus, a B-stationary point of the problem is a vector  $\hat{x} \in S_{\text{ps}}(\bar{x})$  such that  $\Phi(\bullet; \bar{x})'(\hat{x}; v) \geq 0$  for all  $v \in \mathcal{T}(S_{\text{ps}}(\bar{x}); \hat{x})$ .

**Definition 5.2.** A vector  $\bar{x} \in X$  is said to be a

- *pseudo B-stationary point* of (1) if  $\bar{x}$  is a B-stationary point of the pulled-down problem (15).
- *pseudo local minimizer* of (1) if  $\bar{x}$  is a local minimizer of the pulled-down problem (15). □

Besides its dictionary meaning of being not real, the adjective “pseudo” carries the hidden meaning that the concepts are defined in terms of an auxiliary problem induced by the candidate solution on hand. The following result shows that the above pseudo conditions are necessary for a local minimizer of (1), thus providing a first step in computing a promising candidate for a local minimum of the problem.

**Proposition 5.3.** Let  $c$ ,  $\{\varphi_k, g_k\}_{k=1}^K$ , and  $\{\phi_\ell, h_\ell\}_{\ell=1}^L$  be B-differentiable functions on the open set  $\mathcal{O}$  containing the closed set  $X$ . Among the following three statements for a vector  $\bar{x} \in X$ , it holds that (a)  $\Rightarrow$  (b)  $\Rightarrow$  (c):

- (a)  $\bar{x}$  is a local minimizer of (1).
- (b)  $\bar{x}$  is a pseudo local minimizer of (1).
- (c)  $\bar{x}$  is a pseudo B-stationary point of (1).

**Proof.** It suffices to show that the local minimizer of problem (1) must be a pseudo local minimizer. By the continuity of the functions  $\{g_k\}_{k=1}^K$ , there exists a neighborhood  $\mathcal{N}$  of  $\bar{x}$  such that for all  $x \in \mathcal{N}$ , it holds that:  $g_k(x) < 0$  for all  $k \in \mathcal{K}_<(\bar{x})$  and  $g_k(x) > 0$  for all  $k \in \mathcal{K}_>(\bar{x})$ ; and similarly for the  $\{\phi_\ell, h_\ell\}$ -functions. We may restrict the neighborhood  $\mathcal{N}$  so that  $\bar{x}$  is a minimizer of  $\Phi$  on  $X \cap \mathcal{N}$ . Hence if  $x \in \mathcal{N}$  is feasible to (15), then  $x$  is also feasible to (1) because

$$\sum_{\ell=1}^L \phi_\ell(x) \mathbf{1}_{(0,\infty)}(h_\ell(x)) = \sum_{\ell \in \mathcal{L}_>(\bar{x})} \phi_\ell(x).$$

Thus we have  $\Phi(x; \bar{x}) = \Phi(x) \geq \Phi(\bar{x}) = \Phi(\bar{x}; \bar{x})$  showing that  $\bar{x}$  is a local minimizer of the problem (15); thus (b) holds.  $\square$

An important point of Proposition 5.3 is that the two pseudo conditions are necessary for a local minimizer of (1). With the two approaches described in the two later sections and with the aid of a host of existing algorithms for practically implementing the approaches (as mentioned in the closing of the Introduction), the computation of a pseudo B-stationary solution can be accomplished by iterative algorithms for large classes of functions. This is in contrast to the computation of a pseudo local minimizer which in general is a daunting, if not impossible task. It is therefore natural to ask whether there are nonconvex nondifferentiable classes of problems for which a pseudo B-stationary point must be pseudo locally minimizing. It turns out that the answer is affirmative based on the convex-like property defined as follows; see [11, Section 4.1]. A function  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  is said to be *convex-like* near a vector  $\bar{x}$  if there exists a neighborhood  $\mathcal{N}$  of  $\bar{x}$  such that

$$f(x) \geq f(\bar{x}) + f'(\bar{x}; x - \bar{x}), \quad \forall x \in \mathcal{N}.$$

A very broad class of convex-like functions consists of the following 3-layer composite functions:

$$f(x) \triangleq \varphi \circ \theta \circ \psi(x), \quad (16)$$

where  $\varphi: \mathbb{R} \rightarrow \mathbb{R}$  is piecewise affine and nondecreasing;  $\theta: \mathbb{R} \rightarrow \mathbb{R}$  is convex, and  $\psi: \mathbb{R}^n \rightarrow \mathbb{R}$  is piecewise affine; see [11, Lemma 10]. In particular, piecewise affine functions and convex functions are convex-like near any point. Under the convexity-like conditions, we aim to specialize [11, Proposition 9] to the problem (15). For this purpose, let  $\bar{x} \in S_{\text{ps}}(\bar{x})$ . We then have the inclusion:

$$\mathcal{T}(S_{\text{ps}}(\bar{x}); \bar{x}) \subseteq \left\{ v \in \mathcal{T}(X; \bar{x}) \left| \begin{array}{l} \sum_{\ell \in \mathcal{L}_>(\bar{x})} \phi'_\ell(\bar{x}; v) (\leq 0) \\ g'_k(\bar{x}; v) \leq 0, \quad \forall k \in \mathcal{K}_=(\bar{x}) \\ h'_\ell(\bar{x}; v) \leq 0, \quad \forall \ell \in \mathcal{L}_=(\bar{x}) \end{array} \right. \right\} \triangleq \mathcal{L}(S_{\text{ps}}(\bar{x}); \bar{x}), \quad (17)$$

where the notation  $(\leq 0)$  means that this constraint is vacuous if the functional constraint holds as a strict inequality at  $\bar{x}$ ; i.e., if

$$\sum_{\ell=1}^L \phi_\ell(\bar{x}) \mathbf{1}_{(0,\infty)}(h_\ell(\bar{x})) = \sum_{\ell \in \mathcal{L}_>(\bar{x})} \phi_\ell(\bar{x}) < b.$$

The equality of the left- and right-hand cones in (17) is the Abadie constraint qualification (ACQ) for the set  $S_{\text{ps}}(\bar{x})$  at the member vector  $\bar{x}$ . Sufficient conditions for this CQ to hold are known; in particular, either one of the following two sets of conditions yields the ACQ (see e.g. [44, Section 4.1]):

- (piecewise polyhedrality) all the functions  $\{g_k\}_{k \in \mathcal{K}_=(\bar{x})}$  and  $\{h_\ell\}_{\ell \in \mathcal{L}_=(\bar{x})}$ , (and also  $\{\phi_\ell\}_{\ell \in \mathcal{L}_>(\bar{x})}$  if the functional constraint is binding at  $\bar{x}$ ) are piecewise affine; or
- (directional Slater) there exists a vector  $v \in \mathcal{T}(X; \bar{x})$  such that

$$\begin{aligned} \sum_{\ell \in \mathcal{L}_>(\bar{x})} \phi'_\ell(\bar{x}; v) &< 0 \\ g'_k(\bar{x}; v) &< 0, \quad \forall k \in \mathcal{K}_=(\bar{x}) \\ h'_\ell(\bar{x}; v) &< 0, \quad \forall \ell \in \mathcal{L}_=(\bar{x}), \end{aligned}$$

where  $(< 0)$  has the same meaning as  $(\leq 0)$  when the functional constraint is not binding at  $\bar{x}$ . We have the following result which follows readily from [11, Proposition 9].

**Proposition 5.4.** *Under the blanket assumption of (1), suppose that  $c$ ,  $\{g_k\}_{k \in \mathcal{K}_=(\bar{x})}$ ,  $\{\varphi_k\}_{k \in \mathcal{K}_>(\bar{x})}$ ,  $\{h_\ell\}_{\ell \in \mathcal{L}_=(\bar{x})}$ , and  $\{\phi_\ell\}_{\ell \in \mathcal{L}_>(\bar{x})}$ , are all convex-like near  $\bar{x}$ , which is a pseudo B-stationary solution of (1). If the ACQ holds for the set  $S_{\text{ps}}(\bar{x})$  at  $\bar{x}$ , then  $\bar{x}$  is a pseudo local minimizer of (1).  $\square$*

In terms of the larger cone  $\mathcal{L}(S_{\text{ps}}(\bar{x}); \bar{x})$ , it follows that a feasible vector  $\bar{x}$  of (1) is a pseudo B-stationary solution if the following implication holds:

$$v \in \mathcal{L}(S_{\text{ps}}(\bar{x}); \bar{x}) \Rightarrow \Phi(\bullet; \bar{x})'(\bar{x}; v) \geq 0; \quad (18)$$

or equivalently, if 
$$0 \in \underset{v \in \mathcal{L}(S_{\text{ps}}(\bar{x}); \bar{x})}{\operatorname{argmin}} \Phi(\bullet; \bar{x})'(\bar{x}; v).$$

When the functions  $c$ ,  $\{g_k\}_{k \in \mathcal{K}_=(\bar{x})}$ ,  $\{\varphi_k\}_{k \in \mathcal{K}_>(\bar{x})}$ ,  $\{h_\ell\}_{\ell \in \mathcal{L}_=(\bar{x})}$ , and  $\{\phi_\ell\}_{\ell \in \mathcal{L}_>(\bar{x})}$  are of the difference-of-convex (dc) kind, then the latter minimization problem in  $v$  (with  $\bar{x}$  given) is a dc constrained dc program that has been studied in [12, Section 6.4]. Under the ACQ for the set  $S_{\text{ps}}(\bar{x})$  at  $\bar{x}$ , the implication (18) is necessary and sufficient for pseudo B-stationarity. The upshot of this discussion is that with (or without) the ACQ, a feasible vector can in principle be checked for pseudo B-stationarity by methods existed in the literature; nevertheless, these methods can not be applied to compute a pseudo B-stationary solution. This is the task in the rest of the paper.

The pseudo B-stationarity definition can be phrased in an equivalent way, which when strengthened, yields a sufficient condition of a local minimizer of the problem (1). The latter condition is related to the formulations (8), (11) and (12) albeit with some obvious differences. First, the auxiliary variables  $s$  and  $z$  are pulled outside of these formulations; thereby eliminating the products in the objective functions of these problems. More interestingly, the result below highlights the difference between the necessary conditions (in Proposition 5.3) and the sufficient conditions for a local minimum. For the former, the condition is the *existence* of (binary) “multipliers”; for the latter, the condition is “for all” such multipliers.

**Proposition 5.5.** *Let  $c$ ,  $\{\varphi_k, g_k\}_{k=1}^K$ , and  $\{\phi_\ell, h_\ell\}_{\ell=1}^L$  be  $B$ -differentiable functions on the open set  $\mathcal{O}$  containing the closed set  $X$ . Let  $\bar{x} \in X$  be given. The following three statements hold.*

(a) *A necessary condition for  $\bar{x}$  to be a local minimizer of (1) is that there exist (finite) families of binary multipliers  $\{\xi_k\}_{k \in \mathcal{K}_=(\bar{x})} \subset \{0, 1\}^{|\mathcal{K}_=(\bar{x})|}$  and  $\{\mu_\ell\}_{\ell \in \mathcal{L}_=(\bar{x})} \subset \{0, 1\}^{|\mathcal{L}_=(\bar{x})|}$  such that  $\bar{x}$  is a local minimizer of the problem:*

$$\begin{aligned} & \underset{x}{\text{minimize}} \Phi(x; \bar{x}) \triangleq c(x) + \sum_{k \in \mathcal{K}_>(\bar{x})} \varphi_k(x) \\ & \text{subject to } \left. \begin{aligned} & x \in X; \quad \sum_{\ell \in \mathcal{L}_>(\bar{x})} \phi_\ell(x) \leq b \\ & (1 - \xi_k)g_k(x) \leq 0 \quad \forall k \in \mathcal{K}_=(\bar{x}) \\ & g_k(x) \leq 0 \quad \forall k \in \mathcal{K}_<(\bar{x}) \\ & g_k(x) \geq 0 \quad \forall k \in \mathcal{K}_>(\bar{x}) \\ & (1 - \mu_\ell)h_\ell(x) \leq 0 \quad \forall \ell \in \mathcal{L}_=(\bar{x}) \\ & h_\ell(x) \leq 0 \quad \forall \ell \in \mathcal{L}_<(\bar{x}) \\ & \text{and } h_\ell(x) \geq 0 \quad \forall \ell \in \mathcal{L}_>(\bar{x}) \end{aligned} \right\} \begin{aligned} & \text{denoted as } \widehat{S}_{ps}(\bar{x}; \xi, \mu), \\ & \text{same as } S_{ps}(\bar{x}) \text{ except} \\ & \text{for the constraints} \\ & \text{indexed by} \\ & \mathcal{K}_=(\bar{x}) \times \mathcal{L}_=(\bar{x}). \end{aligned} \end{aligned} \quad (19)$$

(b) *Conversely, suppose that in a neighborhood of  $\bar{x}$ , the products  $\varphi_k [g_k]_+$  and  $\phi_\ell [h_\ell]_+$  are nonnegative for all pairs  $(k, \ell)$  in  $\mathcal{K}_=(\bar{x}) \times \mathcal{L}_=(\bar{x})$ . Then a sufficient condition for  $\bar{x}$  to be a local minimizer of (1) is that for all  $\{\xi_k\}_{k \in \mathcal{K}_=(\bar{x})} \subset \{0, 1\}^{|\mathcal{K}_=(\bar{x})|}$  and  $\{\mu_\ell\}_{\ell \in \mathcal{L}_=(\bar{x})} \subset \{0, 1\}^{|\mathcal{L}_=(\bar{x})|}$ ,  $\bar{x}$  is a local minimizer of (19).*

(c) *Alternatively, suppose that  $\varphi_k(\bar{x})$  and  $\phi_\ell(\bar{x})$  are nonnegative for all  $(k, \ell)$  in  $\mathcal{K}_=(\bar{x}) \times \mathcal{L}_=(\bar{x})$ . Then a sufficient condition for  $\bar{x}$  to be a local minimizer of (1) is that for all  $\{\xi_k\}_{k \in \mathcal{K}_=(\bar{x})} \subset \{0, 1\}^{|\mathcal{K}_=(\bar{x})|}$  and  $\{\mu_\ell\}_{\ell \in \mathcal{L}_=(\bar{x})} \subset \{0, 1\}^{|\mathcal{L}_=(\bar{x})|}$ ,  $\bar{x}$  is a local minimizer of*

$$\begin{aligned} & \underset{x}{\text{minimize}} \Phi_{\geq}^{\xi}(x; \bar{x}) \triangleq c(x) + \sum_{k \in \mathcal{K}_>(\bar{x})} \varphi_k(x) + \underbrace{\sum_{k \in \mathcal{K}_=(\bar{x})} \xi_k \varphi_k(x)}_{\text{extra term with multiplier}} \\ & \text{subject to } \left. \begin{aligned} & x \in X; \quad \sum_{\ell \in \mathcal{L}_>(\bar{x})} \phi_\ell(x) + \underbrace{\sum_{\ell \in \mathcal{L}_=(\bar{x})} \mu_\ell \phi_\ell(x)}_{\text{extra term with multiplier}} \leq b \\ & (1 - \xi_k)g_k(x) \leq 0 \quad \forall k \in \mathcal{K}_=(\bar{x}) \\ & g_k(x) \leq 0 \quad \forall k \in \mathcal{K}_<(\bar{x}) \\ & g_k(x) \geq 0 \quad \forall k \in \mathcal{K}_>(\bar{x}) \\ & (1 - \mu_\ell)h_\ell(x) \leq 0 \quad \forall \ell \in \mathcal{L}_=(\bar{x}) \\ & h_\ell(x) \leq 0 \quad \forall \ell \in \mathcal{L}_<(\bar{x}) \\ & \text{and } h_\ell(x) \geq 0 \quad \forall \ell \in \mathcal{L}_>(\bar{x}) \end{aligned} \right\}. \end{aligned} \quad (20)$$

**Proof.** The necessary condition in (a) is clear because we can let  $\xi_k = 0 = \mu_\ell$  for all  $(k, \ell) \in \mathcal{K}_=(\bar{x}) \times \mathcal{L}_=(\bar{x})$ . For the sufficiency, suppose that  $\bar{x}$  is a local minimizer of (19) for all tuples  $\{\xi_k\}_{k \in \mathcal{K}_=(\bar{x})}$  and  $\{\mu_\ell\}_{\ell \in \mathcal{L}_=(\bar{x})}$  as stated. We first note that  $\bar{x}$  must be feasible to (1) because

$$\sum_{\ell=1}^L \phi_\ell(\bar{x}) \mathbf{1}_{(0,\infty)}(h_\ell(\bar{x})) = \sum_{\ell \in \mathcal{L}_>(\bar{x})} \phi_\ell(\bar{x}).$$

Since the family of problems (19) is finite, there exists a neighborhood  $\mathcal{N}$  of  $\bar{x}$  such that for all  $\{\xi_k\}_{k \in \mathcal{K}_=(\bar{x})} \subset \{0,1\}^{|\mathcal{K}_=(\bar{x})|}$  and  $\{\mu_\ell\}_{\ell \in \mathcal{L}_=(\bar{x})} \subset \{0,1\}^{|\mathcal{L}_=(\bar{x})|}$ , if  $x \in \mathcal{N}$  is feasible to (19), then  $\Phi(x; \bar{x}) \geq \Phi(\bar{x}; \bar{x}) = \Phi(\bar{x})$ . Without loss of generality, we may assume that this neighborhood is such that for all  $x \in \mathcal{N}$ , it holds that  $\varphi_k(x) [g_k(x)]_+ \geq 0$  and  $\phi_\ell(x) [h_\ell(x)]_+ \geq 0$  for all  $(k, \ell) \in \mathcal{K}_=(\bar{x}) \times \mathcal{L}_=(\bar{x})$ ;  $g_k(x) < 0$  for all  $k \in \mathcal{K}_<(\bar{x})$  and  $g_k(x) > 0$  for all  $k \in \mathcal{K}_>(\bar{x})$ ; and similarly for the  $\{\phi_\ell, h_\ell\}$ -functions. Let  $x \in \mathcal{N}$  be feasible to (1). For each pair  $(k, \ell) \in \mathcal{K}_=(\bar{x}) \times \mathcal{L}_=(\bar{x})$ , let  $\xi_k \triangleq \mathbf{1}_{(0,\infty)}(g_k(x))$  and  $\mu_\ell \triangleq \mathbf{1}_{(0,\infty)}(h_\ell(x))$ . To show that  $x$  is feasible to (19) for this pair  $(\xi, \mu)$ , it suffices to verify the functional constraint and  $(1 - \xi_k) g_k(x) \leq 0$  and  $(1 - \mu_\ell) h_\ell(x) \geq 0$  for all  $(k, \ell) \in \mathcal{K}_=(\bar{x}) \times \mathcal{L}_=(\bar{x})$ . The latter is clear. For the former, we have

$$\begin{aligned} b &\geq \sum_{\ell=1}^L \phi_\ell(x) \mathbf{1}_{(0,\infty)}(h_\ell(x)) = \sum_{\ell \in \mathcal{L}_>(\bar{x})} \phi_\ell(x) \mathbf{1}_{(0,\infty)}(h_\ell(x)) + \sum_{\ell \in \mathcal{L}_=(\bar{x})} \phi_\ell(x) \mathbf{1}_{(0,\infty)}(h_\ell(x)) \\ &\geq \sum_{\ell \in \mathcal{L}_>(\bar{x})} \phi_\ell(x) \mathbf{1}_{(0,\infty)}(h_\ell(x)) \quad \text{because } \phi_\ell(x) [h_\ell(x)]_+ \geq 0 \text{ for } \ell \in \mathcal{L}_=(\bar{x}) \\ &= \sum_{\ell \in \mathcal{L}_>(\bar{x})} \phi_\ell(x). \end{aligned}$$

Hence it follows that

$$\begin{aligned} \Phi(x) &= c(x) + \sum_{k \in \mathcal{K}_>(\bar{x})} \varphi_k(x) \mathbf{1}_{(0,\infty)}(g_k(x)) + \sum_{k \in \mathcal{K}_=(\bar{x})} \varphi_k(x) \mathbf{1}_{(0,\infty)}(g_k(x)) \\ &\geq \Phi(x, \bar{x}) \quad \text{because } \varphi_k(x) [g_k(x)]_+ \geq 0 \text{ for } k \in \mathcal{K}_=(\bar{x}) \\ &\geq \Phi(\bar{x}; \bar{x}) = \Phi(\bar{x}), \end{aligned}$$

as desired. Finally, the proof of statement (c) is very similar to that of (b). Omitting the details, we simply note that  $\bar{x}$ , being feasible to (20), must be feasible to (19), and thus to (1).  $\square$

**Remark 5.6.** To more precisely connect the pull-down problem (15) with the problem (19) for various choices of the binary pairs  $(\xi, \mu)$ , we note that the given  $\bar{x}$  is a local minimizer of the former if and only if  $\bar{x}$  is a local minimizer of the latter for some  $(\xi, \mu)$ . In view of statement (b) in Proposition 5.5, which involves *for all*  $(\xi, \mu)$ , it is natural to ask the following question: suppose  $\bar{x}$  is a local minimizer of (15), are there sufficient conditions that will ensure  $\bar{x}$  to be a local minimizer of (1)? Since local optimality involves a neighborhood of  $\bar{x}$  wherein the signs of  $\varphi_k(x) [g_k(x)]_+$  and  $\phi_\ell(x) [h_\ell(x)]_+$  are relevant, an answer to the question would essentially reduce to statement (b) in Proposition 5.5.

Incidentally, the difference between the parts (b) and (c) in Proposition 5.5 is in their respective assumptions on the functions  $(\varphi_k, g_k)$  and  $(\phi_\ell, h_\ell)$  for elements  $(k, \ell) \in \mathcal{K}_=(\bar{x}) \times \mathcal{L}_=(\bar{x})$  and the two resulting problems (19) and (20).  $\square$

## 6. The epigraphical approach

Guided by the sign assumption in case (A) of Proposition 4.1, which has two parts for the problem (1),

- for every  $k \in [K]$ , each  $\varphi_k$  is nonnegative on the set  $X \cap g_k^{-1}(0)$ ; and
- for every  $\ell \in [L]$ , each  $\phi_\ell$  is nonnegative on the set  $X \cap h_\ell^{-1}(0)$ ;

we present in this section a constructive approach for computing a pseudo B-stationary point of (1). The approach is based on an epigraphical formulation of the problem that lifts it to a higher dimension. We recall that the epigraph of a function  $f$  on the closed set  $S \subseteq \mathcal{O}$  is the set

$$\text{epi}(f; S) \triangleq \{(t, x) \in \mathbb{R} \times S \mid t \geq f(x)\}.$$

In what follows, we describe the tangent cone of the epigraph of a discontinuous product function  $\psi \mathbf{1}_{(0, \infty)}(f)$ ; the description not only provides insights for the demonstration of the epigraphical approach to successfully accomplish the computational goal, but is also of independent interest as it relates to some existing results in the literature. A basic result of this kind for a locally Lipschitz function is available from [8, Theorem 2.4.9 part (a)]; but it is not applicable to the discontinuous Heaviside function. An advanced result for an arbitrary function, and for the indicator function in particular, can be found in [49, Theorem 8.2] which is based on the notion of subderivatives. Rather than going through the calculation of the latter derivatives for the product function of interest, we present an elementary derivation that exposes the epigraph of  $\psi \mathbf{1}_{(0, \infty)}(f)$  as the union of two closed sets and highlights the consequence of the sign condition of  $\psi$  on the zero set of  $f$ .

**Proposition 6.1.** *Let  $\psi$  and  $f$  be B-differentiable functions on the open set  $\mathcal{O}$  that contains a closed set  $S$ . Suppose that  $\psi$  is nonnegative on  $S \cap f^{-1}(0)$ . Let  $\pi(x) \triangleq \psi(x) \mathbf{1}_{(0, \infty)}(f(x))$ . Then,*

$$\begin{aligned} \text{epi}(\pi; S) = & \underbrace{\{(t, x) \in \mathbb{R} \times S \mid t \geq \psi(x), f(x) \geq 0\}}_{\text{denoted } E_1} \\ & \cup \underbrace{(\mathbb{R}_+ \times (S \cap f^{-1}(-\infty, 0]))}_{\text{denoted } E_2}. \end{aligned} \quad (21)$$

Thus for any pair  $(t, x) \in \text{epi}(\pi; S)$ ,

$$\mathcal{T}(\text{epi}(\pi; S); (t, x)) = \mathcal{T}(E_1; (t, x)) \cup \mathcal{T}(E_2; (t, x)), \quad (22)$$

where  $\mathcal{T}(E_i; (t, x)) \triangleq \emptyset$  if  $(t, x) \notin E_i$  for  $i = 1, 2$ . Moreover, the following three statements (a), (b), and (c) hold:

- (a)  $(t, x) \in \text{epi}(\pi; S)$  if and only if  $x \in S$  and

$$\underbrace{\min(\max(\psi(x) - t, -f(x)), \max(f(x), -t))}_{dc \text{ in } (x, t) \text{ if } \psi \text{ and } f \text{ are dc}} \leq 0; \quad (23)$$

The function on the left side is piecewise affine in  $(x, t)$  if  $\psi$  and  $f$  are piecewise affine.

(b) For a pair  $(\bar{t}, \bar{x}) \in \text{epi}(\pi; S)$  with  $\bar{t} = \pi(\bar{x})$ , it holds that

- if  $f(\bar{x}) > 0$ , then  $\mathcal{T}(\text{epi}(\pi; S); (\bar{t}, \bar{x})) = \{(dt, v) \in \mathbb{R} \times \mathcal{T}(S; \bar{x}) \mid dt \geq \psi'(\bar{x}; v)\}$ ;
- if  $f(\bar{x}) < 0$ , then  $\mathcal{T}(\text{epi}(\pi; S); (\bar{t}, \bar{x})) = \mathbb{R}_+ \times \mathcal{T}(S; \bar{x})$ ;
- if  $f(\bar{x}) = 0 < \psi(\bar{x})$ , then

$$\mathcal{T}(\text{epi}(\pi; S); (\bar{t}, \bar{x})) \subseteq \{(dt, v) \in \mathbb{R}_+ \times \mathcal{T}(S; \bar{x}) \mid f'(\bar{x}; v) \leq 0\} \quad (24)$$

with equality holding if the set  $\{x \in S \mid f(x) \leq 0\}$  satisfies the ACQ at  $\bar{x}$ ;

- if  $f(\bar{x}) = 0 = \psi(\bar{x})$ , then

$$\begin{aligned} \mathcal{T}(\text{epi}(\pi; S); (\bar{t}, \bar{x})) \subseteq \{(dt, v) \in \mathbb{R} \times \mathcal{T}(S; \bar{x}) \mid dt \geq \psi'(\bar{x}; v), f'(\bar{x}; v) \geq 0\} \\ \bigcup \{(dt, v) \in \mathbb{R}_+ \times \mathcal{T}(S; \bar{x}) \mid f'(\bar{x}; v) \leq 0\}; \end{aligned} \quad (25)$$

moreover, if  $\psi'(\bar{x}; \bullet)$  is nonnegative on  $\mathcal{T}(S; \bar{x}) \cap (f'(\bar{x}; \bullet))^{-1}(0)$ , then the right-hand union in (25) is equal to  $\text{epi}(\psi'(\bar{x}; \bullet) \mathbf{1}_{(0, \infty)}(f'(\bar{x}; \bullet)); \mathcal{T}(S; \bar{x}))$ ; lastly, if the two sets  $\bar{S}_+ \triangleq \{x \in S \mid f(x) \geq 0\}$  and  $\bar{S}_- \triangleq \{x \in S \mid f(x) \leq 0\}$  satisfy the ACQ at  $\bar{x}$ , then

$$\mathcal{T}(\text{epi}(\pi; S); (\bar{t}, \bar{x})) = \text{epi}(\psi'(\bar{x}; \bullet) \mathbf{1}_{(0, \infty)}(f'(\bar{x}; \bullet)); \mathcal{T}(S; \bar{x})). \quad (26)$$

(c) For a pair  $(\bar{t}, \bar{x}) \in \text{epi}(\pi; S)$  with  $\bar{t} > \pi(\bar{x})$ , it holds that

- if  $f(\bar{x}) \neq 0$  or  $f(\bar{x}) = 0 = \psi(\bar{x})$ , then  $\mathcal{T}(\text{epi}(\pi; S); (\bar{t}, \bar{x})) = \mathbb{R} \times \mathcal{T}(S; \bar{x})$ ;
- if  $f(\bar{x}) = 0 < \psi(\bar{x}) \neq \bar{t}$ , then

$$\mathcal{T}(\text{epi}(\pi; S); (\bar{t}, \bar{x})) \begin{cases} = \mathbb{R} \times \mathcal{T}(S; \bar{x}) & \text{if } \bar{t} > \psi(\bar{x}) \\ \subseteq \mathbb{R} \times (\mathcal{T}(S; \bar{x}) \cap (f'(\bar{x}; \bullet))^{-1}(-\infty, 0]) & \text{if } \bar{t} < \psi(\bar{x}); \end{cases}$$

- if  $f(\bar{x}) = 0 < \psi(\bar{x}) = \bar{t}$ , then

$$\begin{aligned} \mathcal{T}(\text{epi}(\pi; S); (\bar{t}, \bar{x})) \subseteq \{(dt, v) \in \mathbb{R} \times \mathcal{T}(S; \bar{x}) \mid dt \geq \psi'(\bar{x}; v), f'(\bar{x}; v) \geq 0\} \\ \bigcup \{(dt, v) \in \mathbb{R} \times \mathcal{T}(S; \bar{x}) \mid f'(\bar{x}; v) \leq 0\}. \end{aligned}$$

**Proof.** By definition of the epigraph, we have

$$\text{epi}(\pi; S) = \{(t, x) \in \mathbb{R} \times S \mid t \geq \psi(x), f(x) > 0\} \cup (\mathbb{R}_+ \times (S \cap f^{-1}(-\infty, 0]));$$

thus  $\text{epi}(\pi; S) \subseteq E_1 \cup E_2$ . Conversely, let  $(t, x) \in E_1$  be such that  $f(x) = 0$ . Then  $t \geq \psi(x) \geq 0$ ; hence  $(t, x) \in E_2 \subseteq \text{epi}(\pi; S)$ . Thus, (21) holds. With the definition that  $\mathcal{T}(E_i; (t, x)) \triangleq \emptyset$  if  $(t, x) \notin E_i$  for  $i = 1, 2$ , the equality (22) is clear. Statement (a) holds because of the following equivalence:

$$\begin{aligned} [t \geq \psi(x) \text{ and } f(x) \geq 0] \text{ or } [t \geq 0 \text{ and } f(x) \leq 0] \\ \Leftrightarrow [\max(\psi(x) - t, -f(x)) \leq 0] \text{ or } [\max(f(x), -t) \leq 0] \\ \Leftrightarrow \min(\max(\psi(x) - t, -f(x)), \max(f(x), -t)) \leq 0. \end{aligned}$$

To prove (b), suppose  $f(\bar{x}) > 0$ . Then  $f(x) > 0$  for all  $x$  sufficiently close to  $\bar{x}$ . In this case, the equality  $\mathcal{T}(\text{epi}(\pi; S); (\bar{t}, \bar{x})) = \{(dt, v) \in \mathbb{R} \times \mathcal{T}(S; \bar{x}) \mid dt \geq \psi'(\bar{x}; v)\}$  can be proved as follows. The inclusion  $\subseteq$  is straightforward to prove.

For the “ $\supseteq$ ” inclusion, let  $(dt, v)$  be such that  $v \in \mathcal{T}(S; \bar{x})$  and  $dt \geq \psi'(\bar{x}; v)$ . Let  $v = \lim_{\nu \rightarrow \infty} (x^\nu - \bar{x})/\tau_\nu$  for some sequence  $\{x^\nu\} \subset S$  converging to  $\bar{x}$  and some sequence of scalars  $\{\tau_\nu\} \downarrow 0$ . Suppose  $dt > \psi'(\bar{x}; v)$ . Then  $\bar{t} + \tau_\nu dt > \psi(x^\nu)$  for all  $\nu$  sufficiently large. Thus  $(\bar{t} + \tau_\nu dt, x^\nu) \in \text{epi}(\pi; S)$  for all such  $\nu$ . This shows that  $(dt, v) \in \mathcal{T}(\text{epi}(\pi; S); (\bar{t}, \bar{x}))$  in this case. Suppose that  $dt = \psi'(\bar{x}; v)$ . Then for every  $\varepsilon > 0$ ,  $(dt + \varepsilon, v) \in \mathcal{T}(\text{epi}(\pi; S); (\bar{t}, \bar{x}))$ . Since the tangent cone is closed, we deduce  $(dt, v) \in \mathcal{T}(\text{epi}(\pi; S); (\bar{t}, \bar{x}))$ , completing the proof of the claimed equality of the two cones.

The second case where  $f(\bar{x}) < 0$  can be similarly argued. For the third case where  $f(\bar{x}) = 0 < \psi(\bar{x})$ , we have  $\bar{t} = 0$  and  $(\bar{t}, \bar{x}) \in E_2 \setminus E_1$ . So the desired inclusion (24) and the equality under the stated CQ both hold easily.

Consider the last case where  $f(\bar{x}) = 0 = \psi(\bar{x})$  so that  $(\bar{t}, \bar{x}) \in E_1 \cap E_2$ . It is easy to prove that  $\mathcal{T}(E_i; (\bar{t}, \bar{x}))$  for  $i = 1, 2$  is contained in the two sets on the right-hand side of (25), respectively. If  $\psi'(\bar{x}; \bullet)$  is nonnegative on  $\mathcal{T}(S; \bar{x}) \cap (f'(\bar{x}; \bullet))^{-1}(0)$ , the claim that the right-hand union in (25) is equal to  $\text{epi}(\psi'(\bar{x}; \bullet) \mathbf{1}_{(0, \infty)}(f'(\bar{x}; \bullet); \mathcal{T}(S; \bar{x})))$  can be proved similarly to that of the equality (21). Finally, under the two additional ACQs, the equality (26) holds by the respective representation of  $\mathcal{T}(E_i; (\bar{t}, \bar{x}))$ . The proof of statement (c) is similar and omitted.  $\square$

Summarizing the various cases in the above proposition, we conclude that there are three “constraint qualifications” (in place if needed) to be satisfied by a pair  $(\bar{t}, \bar{x}) \in \text{epi}(\pi; S)$  in order for the tangent cone  $\mathcal{T}(\text{epi}(\pi; S); (\bar{t}, \bar{x}))$  to have an exact representation:

- (i)  $\psi'(\bar{x}; \bullet)$  is nonnegative on  $\mathcal{T}(S; \bar{x}) \cap (f'(\bar{x}; \bullet))^{-1}(0)$ ;
- (ii) the set  $\bar{S}_+ \triangleq \{x \in S \mid f(x) \geq 0\}$  satisfies the ACQ at  $\bar{x}$ ; and
- (iii) the set  $\bar{S}_- \triangleq \{x \in S \mid f(x) \leq 0\}$  satisfies the ACQ at  $\bar{x}$ .

Moreover, refining the above analysis, it is easy to derive sufficient conditions for the set

$$\text{epi}(\pi; S) = \{(t, x) \in \mathbb{R} \times S \mid (23) \text{ holds}\}$$

to satisfy the ACQ at the pair  $(\pi(\bar{x}), \bar{x})$ .

### 6.1. The penalized epigraphical formulation

In addition to the challenges associated with the Heaviside functions, the functional constraint also complicates the analysis and the design of computational algorithms for the problem (1). There are two ways to address this constraint: one is a direct treatment as a hard constraint; the other is by a soft penalty with the goal of recovering the satisfaction of the constraint and achieving the stationarity of the problem. In this section and Section 8, we adopt the latter penalty approach as it offers a unified treatment with the composite Heaviside functions occurring in the objective only. Throughout this subsection, the functions  $\varphi_k$  and  $\phi_\ell$  are nonnegative on  $X \cap g_k^{-1}(0)$  and  $X \cap h_\ell^{-1}(0)$ , respectively.

As a first step toward a computationally tractable formulation of (1), we make the substitutions

$$t_k = \varphi_k(x) \mathbf{1}_{(0,\infty)}(g_k(x)) \quad \text{and} \quad s_\ell = \phi_\ell(x) \mathbf{1}_{(0,\infty)}(h_\ell(x)), \quad \forall (k, \ell) \in [K] \times [L], \quad (27)$$

and relax these definitional equalities to inequalities to obtain the epigraphical constraints:

$$(t_k, x) \in \text{epi}(\pi_k^\varphi; X) \quad \text{and} \quad (s_\ell, x) \in \text{epi}(\pi_\ell^\phi; X),$$

where  $\pi_k^\varphi(x) \triangleq \varphi_k(x) \mathbf{1}_{(0,\infty)}(g_k(x))$  and  $\pi_\ell^\phi(x) \triangleq \phi_\ell(x) \mathbf{1}_{(0,\infty)}(h_\ell(x))$ .

We also relax the functional constraint by penalizing it in the objective. These maneuvers lead to the following penalty problem defined for a given parameter  $\lambda > 0$ :

$$\begin{aligned} \underset{x \in X; t; s}{\text{minimize}} \quad & \Phi_\lambda(x, t, s) \triangleq \underbrace{c(x) + \sum_{k=1}^K t_k}_{\Phi(x) \text{ in epi-form}} + \lambda \underbrace{\max \left( \sum_{\ell=1}^L s_\ell - b, 0 \right)}_{\text{constraint residual fnc. in epi-form}} \end{aligned} \quad (28)$$

**subject to**  $(t_k, x) \in \text{epi}(\pi_k^\varphi; X)$  and  $(s_\ell, x) \in \text{epi}(\pi_\ell^\phi; X)$ ,  $\forall (k, \ell) \in [K] \times [L]$ .

We remark that the  $s$ -variables and the corresponding constraints  $(s_\ell, x) \in \text{epi}(\pi_\ell^\phi; X)$  are not needed for the problem (7) where the functional constraint is absent.

By part (a) of Proposition 6.1, problem (28) is equivalent to the following problem with the epigraphical constraints exposed:

$$\underset{x \in X; t; s}{\text{minimize}} \quad \Phi_\lambda(x, t, s) \triangleq c(x) + \sum_{k=1}^K t_k + \lambda \max \left( \sum_{\ell=1}^L s_\ell - b, 0 \right) \quad (29)$$

**subject to**  $\min(\max(\varphi_k(x) - t_k, -g_k(x)), \max(g_k(x), -t_k)) \leq 0, \quad \forall k \in [K]$

**and**  $\min(\max(\phi_\ell(x) - s_\ell, -h_\ell(x)), \max(h_\ell(x), -s_\ell)) \leq 0, \quad \forall \ell \in [L]$ .

The latter formulation is the computational workhorse to derive a B-stationary solution of (28) via the epigraphical approach. In what follows, we show that under a suitable directional conditions on the functions  $\phi_\ell$  in the functional constraint, a finite value of  $\lambda$  exists such that a B-stationary solution of (28) is a B-stationary solution of (15), hence a pseudo B-stationary solution of (1). The demonstration is carried out in 2 steps, with the first step being the recovery of the equalities of the  $t$ -variables in (27). This step is accomplished by the following lemma.

**Lemma 6.2.** *Let  $\theta: \mathcal{O}_x \times \mathcal{O}_y \rightarrow \mathbb{R}$  be a B-differentiable bivariate variable on the open set  $\mathcal{O}_x \times \mathcal{O}_y$  containing the closed set  $\widehat{X} \subseteq \mathbb{R}^{n+m}$ . If  $(\bar{x}, \bar{y})$  is a B-stationary point of  $\theta$  on  $\widehat{X}$ , then  $\bar{x}$  is a B-stationary point of  $\theta(\bullet, \bar{y})$  on  $\widehat{X}(\bar{y}) \triangleq \{x \mid (x, \bar{y}) \in \widehat{X}\}$ . A similar statement holds for the  $y$ -variable.*

**Proof.** Indeed, let  $v \in \mathcal{T}(\widehat{X}(\bar{y}); \bar{x})$  be given such that  $v = \lim_{\nu \rightarrow \infty} (x^\nu - \bar{x})/\tau_\nu$  for some sequence  $\{x^\nu\}$  in  $\widehat{X}(\bar{y})$  converging to  $\bar{x}$  and some sequence  $\{\tau_\nu\} \downarrow 0$ .

It then follows that  $(v, 0) \in \mathcal{T}(\widehat{X}; (\bar{x}, \bar{y}))$  and

$$\theta(\bullet, \bar{y})'(\bar{x}; v) = \lim_{\nu \rightarrow \infty} \frac{\theta(x^\nu, \bar{y}) - \theta(\bar{x}, \bar{y})}{\tau_\nu} = \theta'((\bar{x}, \bar{y}); (v, 0)) \geq 0,$$

where the two equalities hold by the B-differentiability of  $\theta$ .  $\square$

Let 
$$\{ \bar{x}; \{ \bar{t}_k \}_{k=1}^K; \{ \bar{s}_\ell \}_{\ell=1}^L \} \quad (30)$$

be a B-stationary tuple of (28). In consequence of Lemma 6.2, it follows that the tuple  $\{ \{ \bar{t}_k \}_{k=1}^K; \{ \bar{s}_\ell \}_{\ell=1}^L \}$  is a B-stationary solution of the problem:

$$\begin{aligned} & \underset{t, s}{\text{minimize}} && \sum_{k=1}^K t_k + \lambda \max \left( \sum_{\ell=1}^L s_\ell - b, 0 \right) \\ & \text{subject to} && t_k \geq \varphi_k(\bar{x}) \mathbf{1}_{(0, \infty)}(g_k(\bar{x})), \quad \forall k \in [K] \\ & \text{and} && s_\ell \geq \phi_\ell(\bar{x}) \mathbf{1}_{(0, \infty)}(h_\ell(\bar{x})), \quad \forall \ell \in [L]. \end{aligned}$$

This is a trivial convex piecewise linear program in the  $(t, s)$ -variables with lower-bound constraints only. It can easily be seen that  $\bar{t}_k = \pi_k^\varphi(\bar{x})$  for all  $k \in [K]$ , but similar equalities are not guaranteed for the  $s$ -variables. In fact, there are two possibilities:

- (i)  $\sum_{\ell=1}^L \bar{s}_\ell \leq b$  and there exists  $\bar{\ell} \in [L]$  such that  $\bar{s}_{\bar{\ell}} > \pi_{\bar{\ell}}^\phi(\bar{x})$ ; or
- (ii)  $\bar{s}_\ell = \pi_\ell^\phi(\bar{x})$  for all  $\ell \in [L]$ .

The following analysis addresses both cases. In case (i), we obtain  $\sum_{\ell=1}^L \pi_\ell^\phi(\bar{x}) < b$ , thus in particular  $\bar{x}$  is feasible to (1). We will return to complete this case in the main Theorem 6.4. For now, we consider the second case where both sets of equalities in (27) hold at  $\bar{x}$ . We next show that the vector  $\bar{x}$  in the B-stationary tuple (30) must be a B-stationary solution of the following problem defined with respect to a given  $\bar{x} \in X$  and without the auxiliary variables  $(t, s)$ :

$$\begin{aligned} & \underset{x \in X}{\text{minimize}} && \widehat{\Phi}_\lambda(x; \bar{x}) \triangleq c(x) + \sum_{k \in \mathcal{K}_>(\bar{x})} \varphi_k(x) + \lambda \max \left( \sum_{\ell \in \mathcal{L}_>(\bar{x})} \phi_\ell(x) - b, 0 \right) \\ & \text{subject to} && \left. \begin{aligned} g_k(x) &\leq 0 & \forall k \in \mathcal{K}_\leq(\bar{x}) \\ g_k(x) &\geq 0 & \forall k \in \mathcal{K}_>(\bar{x}) \\ h_\ell(x) &\leq 0 & \forall \ell \in \mathcal{L}_\leq(\bar{x}) \\ h_\ell(x) &\geq 0 & \forall \ell \in \mathcal{L}_>(\bar{x}) \end{aligned} \right\} && \begin{aligned} & \text{including } x \in X, \text{ the set} \\ & \text{of these constraints is} \\ & \text{denoted by } \widehat{S}_{\text{ps}}(\bar{x}) \text{ and} \\ & \text{is equal to } S_{\text{ps}}(\bar{x}) \text{ without} \\ & \text{the functional constraint} \end{aligned} \\ & \text{and} && \end{aligned} \quad (31)$$

Let  $dx$  be a tangent vector of  $\widehat{S}_{\text{ps}}(\bar{x})$  at  $\bar{x}$ . There exist sequences  $\{x^\nu\} \rightarrow \bar{x}$  and  $\{\tau_\nu\} \downarrow 0$  such that  $x^\nu \in \widehat{S}_{\text{ps}}(\bar{x})$  for all  $\nu$  and  $dx = \lim_{\nu \rightarrow \infty} (x^\nu - \bar{x})/\tau_\nu$ .

Define 
$$dt_k \triangleq \begin{cases} \varphi'_k(\bar{x}; dx) & \text{if } k \in \mathcal{K}_>(\bar{x}) \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad ds_\ell \triangleq \begin{cases} \phi'_\ell(\bar{x}; dx) & \text{if } \ell \in \mathcal{L}_>(\bar{x}) \\ 0 & \text{otherwise.} \end{cases}$$

We claim that  $(dt_k, dx) \in \mathcal{T}(\text{epi}(\pi_k^\varphi; X); (\bar{t}_k, \bar{x}))$

and  $(ds_\ell, dx) \in \mathcal{T}(\text{epi}(\pi_\ell^\phi; X); (\bar{s}_\ell, \bar{x})), \quad \forall (k, \ell) \in [K] \times [L].$

This is obviously true for all  $k \in \mathcal{K}_>(\bar{x}) \cup \mathcal{K}_<(\bar{x})$  and  $\ell \in \mathcal{L}_>(\bar{x}) \cup \mathcal{L}_<(\bar{x})$ , by the first two subcases of Proposition 6.1(b). For an index  $k \in \mathcal{K}_=(\bar{x})$  we have  $g_k(x^\nu) \leq 0$  by (31). With  $t_k^\nu \triangleq \pi_k^\varphi(x^\nu) = 0$ , recalling that we have  $\bar{t}_k = \pi_k^\varphi(\bar{x}) = 0$ , we obtain  $dt_k = 0 = \lim_{\nu \rightarrow \infty} (t_k^\nu - \bar{t}_k) / \tau_\nu$ . Therefore,  $(dt_k, dx)$  belongs to  $\mathcal{T}(\text{epi}(\pi_k^\varphi; X); (\bar{t}_k, \bar{x}))$  for all  $k \in [K]$ . Similarly, we can prove  $(ds_\ell, dx) \in \mathcal{T}(\text{epi}(\pi_\ell^\phi; X); (\bar{s}_\ell, \bar{x}))$  for all  $\ell \in [L]$ . Furthermore, it is easy to see that  $\Phi'_\lambda((\bar{x}, \bar{t}, \bar{s}); (dx, dt, ds)) = \widehat{\Phi}_\lambda(\bullet; \bar{x})'(\bar{x}, dx)$ . This completes the proof that  $\bar{x}$  is a B-stationary solution of (31).

So far, we have not imposed any constraint qualifications; for the last step in concluding that  $\bar{x}$  is B-stationary solution of (15), we need the last condition imposed in Theorem 6.4 below. This condition is the one stated in [12, Theorem 9.2.1] specialized to the problem (31); we refer to this reference for a brief history of the condition in the theory of exact penalization. For the problem (7) this extra condition is not needed; therefore we obtain a constructive approach for obtaining a pseudo B-stationary solution of this problem, which has no constraint involving the Heaviside functions, under only the sign conditions on the  $\varphi_k$  functions (and the blanket B-differentiability assumption of the problem). For the sake of clarity in the last step of the proof of Theorem 6.4, we state the following simple lemma but omit its proof.

**Lemma 6.3.** *Let  $f: \mathcal{O} \rightarrow \mathbb{R}$  be B-differentiable at  $\bar{x}$ . Let  $f_{\max}(x) \triangleq \max(f(x), 0)$ .*

*Then*  $f'_{\max}(\bar{x}; v) \leq \max(f'(\bar{x}; v), 0), \quad \forall v \in \mathbb{R}^n.$

**Theorem 6.4.** *Under the blanket assumption of problem (1), assume that  $\varphi_k$  and  $\phi_\ell$  are nonnegative on  $X \cap g_k^{-1}(0)$  and  $X \cap h_\ell^{-1}(0)$ , respectively, for all  $(k, \ell) \in [K] \times [L]$ , and that  $c$  and each  $\varphi_k$  are Lipschitz continuous on  $X$  with Lipschitz constants  $\text{Lip}_c$  and  $\text{Lip}_\varphi$ , respectively. If  $\{\bar{x}, \bar{t}, \bar{s}\}$  is a B-stationary tuple of (28) corresponding to a  $\lambda$  satisfying*

$$\lambda > \text{Lip}_c + K \text{Lip}_\varphi, \quad (32)$$

*then  $\bar{x}$  is a pseudo B-stationary solution of problem (1), provided that there exists a vector  $\bar{v} \in \mathcal{T}(\widehat{S}_{\text{ps}}(\bar{x}); \bar{x})$  with unit length satisfying  $\sum_{\ell \in \mathcal{L}_>(\bar{x})} \phi'_\ell(\bar{x}; \bar{v}) \leq -1$ .*

**Proof.** Continuing the above analysis of the case where (27) holds at  $\bar{x}$ , we suppose that  $\bar{x}$  fails the functional constraint; i.e.,  $\sum_{\ell=1}^L \pi_\ell^\phi(\bar{x}) > b$ . Then we have

$$\begin{aligned} 0 &\leq \widehat{\Phi}_\lambda(\bullet; \bar{x})'(\bar{x}, \bar{v}) = c'(\bar{x}; \bar{v}) + \sum_{k \in \mathcal{K}_>(\bar{x})} \varphi'_k(\bar{x}; \bar{v}) + \lambda \sum_{\ell \in \mathcal{L}_>(\bar{x})} \phi'_\ell(\bar{x}; \bar{v}) \\ &\leq (\text{Lip}_c + K \text{Lip}_\varphi) - \lambda. \end{aligned}$$

This contradicts the condition (32). Hence  $\bar{x}$  is feasible to (1). To show that  $\bar{x}$  is B-stationary for (15), there are two cases to consider:

$$\sum_{\ell \in \mathcal{L}_>(\bar{x})} \phi_\ell(\bar{x}) = b \quad \text{or} \quad \sum_{\ell \in \mathcal{L}_>(\bar{x})} \phi_\ell(\bar{x}) < b.$$

We consider the left case first. Let  $v \in \mathcal{T}(S_{\text{ps}}(\bar{x}); \bar{x})$ . Then  $v \in \mathcal{T}(\widehat{S}_{\text{ps}}(\bar{x}); \bar{x})$  and  $\sum_{\ell \in \mathcal{L}_{>}(\bar{x})} \phi'_\ell(\bar{x}; v) \leq 0$ . We have

$$0 \leq \widehat{\Phi}_\lambda(\bullet; \bar{x})'(\bar{x}, v) = c'(\bar{x}; v) + \sum_{k \in \mathcal{K}_{>}(\bar{x})} \varphi'_k(\bar{x}; v), \quad (33)$$

establishing the B-stationarity of  $\bar{x}$  for the problem (15). In the right case where  $\sum_{\ell \in \mathcal{L}_{>}(\bar{x})} \phi_\ell(\bar{x}) < b$ , we have  $\mathcal{T}(S_{\text{ps}}(\bar{x}); \bar{x}) = \mathcal{T}(\widehat{S}_{\text{ps}}(\bar{x}); \bar{x})$  and the same expression (33) also holds.

What remains to be proved is the case where the B-stationary tuple (30) of (28) is such that

$$\sum_{\ell \in \mathcal{L}_{>}(\bar{x})} \phi_\ell(\bar{x}) = \sum_{\ell=1}^L \pi_\ell^\phi(\bar{x}) < \sum_{\ell=1}^L \bar{s}_\ell \leq b$$

and  $\bar{s}_{\bar{\ell}} > \pi_{\bar{\ell}}^\phi(\bar{x})$  for at least one  $\bar{\ell}$ . Let  $dx \in \mathcal{T}(S_{\text{ps}}(\bar{x}); \bar{x})$  be arbitrary. There exist sequences  $\{x^\nu\} \rightarrow \bar{x}$  and  $\{\tau_\nu\} \downarrow 0$  such that  $x^\nu \in S_{\text{ps}}(\bar{x}) \subseteq \widehat{S}_{\text{ps}}(\bar{x})$  for all  $\nu$  and  $dx = \lim_{\nu \rightarrow \infty} (x^\nu - \bar{x})/\tau_\nu$ . As we have already shown,  $(dt_k, dx) \in \mathcal{T}(\text{epi}(\pi_k^\phi; X); (\bar{t}_k, \bar{x}))$  for all  $k \in [K]$ , where

$$dt_k \triangleq \begin{cases} \varphi'_k(\bar{x}; dx) & \text{if } k \in \mathcal{K}_{>}(\bar{x}) \\ 0 & \text{otherwise.} \end{cases}$$

For an arbitrary scalar  $M > 0$ , define a vector  $ds^M$  as follows:

$$ds_\ell^M \triangleq \begin{cases} \phi'_\ell(\bar{x}; dx) & \text{if } \ell \in \mathcal{L}_{>}(\bar{x}) \text{ and } \bar{s}_\ell = \pi_\ell^\phi(\bar{x}), \quad \text{index set denoted } \mathcal{L}_{>}^=(\bar{x}) \\ -M & \text{if } \bar{s}_\ell > \pi_\ell^\phi(\bar{x}) \\ 0 & \text{otherwise.} \end{cases}$$

As before, we have  $(ds_\ell^M, dx) \in \mathcal{T}(\text{epi}(\pi_\ell^\phi; X); (\bar{s}_\ell, \bar{x}))$  for all  $\ell \in [L]$  except possibly when  $\bar{s}_\ell > \pi_\ell^\phi(\bar{x})$ . We show that the latter exception can be removed. Indeed, with  $M$  fixed but arbitrary, for an index  $\ell$  of the latter kind, there are 2 cases to consider:

- (i)  $h_\ell(\bar{x}) \neq 0$ : then  $(-M, dx) \in \mathcal{T}(\text{epi}(\pi_\ell^\phi; X); (\bar{s}_\ell, \bar{x}))$  by part (c) of Proposition 6.1.
- (ii)  $h_\ell(\bar{x}) = 0$ : then  $\bar{s}_\ell > \pi_\ell^\phi(\bar{x}) = 0 \geq h_\ell(x^\nu)$ , where the last inequality holds because  $x^\nu \in S_{\text{ps}}(\bar{x})$ . Thus,  $\bar{s}_\ell - \tau_\nu M > \pi_\ell^\phi(x^\nu) = 0$  for all  $\nu$  sufficiently large and  $(-M, dx) \in \mathcal{T}(\text{epi}(\pi_\ell^\phi; X); (\bar{s}_\ell, \bar{x}))$  follows.

We have thus completed the proof that  $(ds_\ell^M, dx) \in \mathcal{T}(\text{epi}(\pi_\ell^\phi; X); (\bar{s}_\ell, \bar{x}))$  for all  $\ell \in [L]$ . We have

$$\sum_{\ell=1}^L ds_\ell^M = \sum_{\ell \in \mathcal{L}_{>}^=(\bar{x})} \phi'_\ell(\bar{x}; dx) - M \underbrace{|\{\ell : \bar{s}_\ell > \pi_\ell^\phi(\bar{x})\}|}_{\geq 1}.$$

By choosing  $M$  sufficiently large, we have  $\sum_{\ell=1}^L ds_\ell^M \leq 0$ .

By Lemma 6.3, it follows that

$$\begin{aligned}
 0 &\leq \Phi_\lambda((\bar{x}, \bar{t}, \bar{s}); (dx, dt, ds^M)) \\
 &\leq c'(\bar{x}; dx) + \sum_{k \in \mathcal{K}_>(\bar{x})} \varphi'_k(\bar{x}; dx) + \lambda \max \left( \sum_{\ell}^L ds_\ell^M, 0 \right) \\
 &= c'(\bar{x}; dx) + \sum_{k \in \mathcal{K}_>(\bar{x})} \varphi'_k(\bar{x}; dx),
 \end{aligned}$$

showing that  $\bar{x}$  is a B-stationary solution of (15), as desired.  $\square$

## 7. A digression: Approximations of the (open) Heaviside function

The lifted formulation (28), or its computational workhorse (29), requires the auxiliary variables  $t$  and  $s$  but has the advantage of leading directly to a pseudo B-stationary point of the original problem (1) under some mild conditions as stated in Theorem 6.4. Of independent interest, the next approach is based on approximations of the open Heaviside function and leads to approximated problems in the  $x$ -variable only. In this section, we digress from the discussion of the problem (1) and present a focussed discussion on two approaches to construct such approximations and show how they are related to each other. One construction is based on truncation [11] and the other is based on the classical work of [18] on mollifiers, or smoothing [7].

**Definition 7.1.** We say that the bivariate functions  $\theta: \mathbb{R} \times \mathbb{R}_{++} \rightarrow [0, 1]$  *p*-approximate (“p” for pointwise) the (open) Heaviside function  $\mathbf{1}_{(0, \infty)}$  if there exist endpoint functions  $\underline{\theta}$  and  $\bar{\theta}: \mathbb{R}_{++} \rightarrow \mathbb{R}_+$  satisfying

$$(A0) \quad \lim_{\delta \downarrow 0} \underline{\theta}(\delta) = \lim_{\delta \downarrow 0} \bar{\theta}(\delta) = 0;$$

$$(A1) \quad \lim_{\delta \downarrow 0} \theta(t, \delta) = \mathbf{1}_{(0, \infty)}(t) \text{ for all } t \in \mathbb{R}; \text{ [(A1) allows us to define } \theta(t, 0) \triangleq \mathbf{1}_{(0, \infty)}(t), \text{ thereby extending the domain of definition of } \theta \text{ to } \mathbb{R} \times \mathbb{R}_+]; \text{ and}$$

$$(A2) \quad \theta(t, \delta) = \begin{cases} 0 & \forall t \leq -\underline{\theta}(\delta) \\ 1 & \forall t \geq \bar{\theta}(\delta) \end{cases} \text{ for all } \delta > 0.$$

For subsequent purposes, we further stipulate that for all  $\delta > 0$ ,

$$(A3) \quad \text{the function } \theta(\bullet, \delta) \text{ is B-differentiable on } [-\underline{\theta}(\delta), \bar{\theta}(\delta)] \text{ with the directional derivatives satisfying}$$

$$\theta(\bullet, \delta)'(t; 1) \geq 0 \geq \theta(\bullet, \delta)'(t; -1), \quad \forall t \in [-\underline{\theta}(\delta), \bar{\theta}(\delta)]. \quad (34)$$

By (A2), condition (A3) implies that  $\theta(\bullet, \delta)'(t; 1) \geq 0 \geq \theta(\bullet, \delta)'(t; -1)$  for all  $t \in \mathbb{R}$ . Clearly, if  $\theta$  p-approximates the Heaviside function  $\mathbf{1}_{(0, \infty)}$ , then so does the function  $\psi(t, \delta) \triangleq \theta(\frac{t}{m(\delta)}, \delta)$  with the variable  $t$  being scaled by the positive univariate function  $m: \mathbb{R}_{++} \rightarrow \mathbb{R}_{++}$  satisfying  $\limsup_{\delta \downarrow 0} m(\delta) < \infty$ . In the next two subsections, we present two ways to obtain p-approximations of the Heaviside function satisfying the B-differentiability requirement.

### 7.1. Truncation derived approximations

Let  $\widehat{\theta}: \mathbb{R} \times \mathbb{R}_{++} \rightarrow \mathbb{R}$  be such that there exist end-point functions  $\underline{\theta}$  and  $\bar{\theta}: \mathbb{R}_{++} \rightarrow \mathbb{R}_+$  satisfying (A0) and

$$(T1) \quad \widehat{\theta}(t, \delta) \begin{cases} \leq 0 & \forall t \leq -\underline{\theta}(\delta) \\ \geq 1 & \forall t \geq \bar{\theta}(\delta) \end{cases} \text{ for all } \delta > 0;$$

$$(T2) \quad \lim_{\delta \downarrow 0} \widehat{\theta}(0, \delta) = 0;$$

$$(T3) \quad \widehat{\theta}(\bullet, \delta) \text{ is B-differentiable on an open interval containing } [-\underline{\theta}(\delta), \bar{\theta}(\delta)] \text{ with}$$

$$\widehat{\theta}(\bullet, \delta)'(t; 1) \geq 0 \geq \widehat{\theta}(\bullet, \delta)'(t; -1), \quad \forall t \in [-\underline{\theta}(\delta), \bar{\theta}(\delta)].$$

To obtain the p-approximation function  $\theta$  from  $\widehat{\theta}$ , let

$$\begin{aligned} T_{[0,1]}(t) &\triangleq \min \{ \max(t, 0), 1 \} = \max \{ \min(t, 1), 0 \} \\ &= \max(t, 0) - \max(t - 1, 0), \quad t \in \mathbb{R} \end{aligned}$$

be the truncation operator to the range  $[0, 1]$  and define the composite function:

$$\theta_{\text{tr}}(t, \delta) \triangleq T_{[0,1]}(\widehat{\theta}(t, \delta)), \quad (t, \delta) \in \mathbb{R} \times \mathbb{R}_{++}. \quad (35)$$

We formally state that the truncated function  $\theta$  is a p-approximation of the Heaviside function in the first part of the next proposition. The truncation function was used in [11] as a unification scheme for many approximations of chance constraints in stochastic programs; see the cited reference for many prior works on the latter subject. Part (b) of the result provides easy sufficient conditions for (T1) and (T3) to hold; this part is the bridge between the truncation approach discussed herein and the nonifier approach to be discussed in the next subsection. Part (c) presents a broad family of approximating functions by truncation that generalize the perspective functions studied extensively in convex analysis [4, 9, 31, 48] and employed extensively for integer programs in recent years; see e.g. [26]. The approximating functions in the last part of the proposition yield the family of folded concave approximations of the  $\ell_0$  function [12, Subsection 3.1.4]; these include the smoothly clipped absolute deviation SCAD function [20]; the minimax concave penalty MCP function [58, 13]; the capped  $\ell_1$ -function [38, Section 5], and others. As shown in [2], all these functions are of the difference-of-convex kind and not differentiable at the origin as they are approximations of the  $\ell_0$ -function that is discontinuous there.

**Proposition 7.2.** *The following statements hold:*

- (a) *If  $\widehat{\theta}$  satisfies (T1), (T2), and (T3), then its truncation function  $\theta_{\text{tr}}$  p-approximates the Heaviside function  $\mathbf{1}_{(0, \infty)}$  and satisfies (A3).*
- (b) *If  $\widehat{\theta}(\bullet, \delta)$  is nondecreasing and satisfies:*
  - (T1')  $\widehat{\theta}(-\underline{\theta}(\delta), \delta) = 0$  and  $\widehat{\theta}(\bar{\theta}(\delta), \delta) = 1$  for all  $\delta > 0$ ,*then (T1) holds; if additionally  $\widehat{\theta}(\bullet, \delta)$  is locally Lipschitz, then (T3) holds.*

- (c) If  $\psi: \mathbb{R} \rightarrow \mathbb{R}$  is a nondecreasing  $B$ -differentiable function satisfying  $\psi(0) = 0$  and  $\psi(1) = 1$ , then

$$\theta(t, \delta) \triangleq T_{[0,1]} \left( \psi \left( q(\delta) + \frac{t}{m(\delta)} \right) \right)$$

with  $q: \mathbb{R}_{++} \rightarrow [0, 1]$  and  $m: \mathbb{R}_{++} \rightarrow \mathbb{R}_{++}$  satisfying

$$\lim_{\delta \downarrow 0} q(\delta) = 0 = \lim_{\delta \downarrow 0} m(\delta) \quad (36)$$

$p$ -approximates  $\mathbf{1}_{(0,\infty)}$  and satisfies (A3).

- (d) If  $\theta_1(t, \delta)$  and  $\theta_2(t, \delta)$  both  $p$ -approximate the open Heaviside function  $\mathbf{1}_{(0,\infty)}$ , then their sum  $\theta_1(t, \delta) + \theta_2(-t, \delta)$   $p$ -approximates the  $\ell_0$ -function  $|t|_0$ .

**Proof.** For (a), we need to show that (T1) and (T3) imply (A1) and (A2) and that (T3) implies (A3). Clearly, (T2) implies (A1) for  $t = 0$ . For a  $t_* > 0$ , we have  $t_* \geq \bar{\theta}(\delta)$  for all  $\delta > 0$  sufficiently small. Thus,  $\theta_{\text{tr}}(t_*, \delta) = 1 = \mathbf{1}_{(0,\infty)}(t_*)$  by (T1) and the definition of truncation, which is (A1) at  $t_*$ . Similarly, (A1) also holds for  $t_* < 0$ . Thus the pointwise convergence condition (A1) holds for all  $t \in \mathbb{R}$ . Clearly (A2) holds by truncation. For (A3), we have

$$\theta_{\text{tr}}(\bullet, \delta)'(t; \pm 1) = T'_{[0,1]}(\hat{\theta}(t, \delta); \hat{\theta}(\bullet, \delta)'(t; \pm 1)).$$

Since  $T'_{[0,1]}(t; 1) \geq 0 \geq T'_{[0,1]}(t; -1)$  for all  $t \in \mathbb{R}$ , we readily obtain (A3) from (T3).

If  $\hat{\theta}$  is nondecreasing, then (T1') clearly implies (T1); moreover, this function is directionally differentiable and by the definition (3), we have

$$\hat{\theta}(\bullet, \delta)'(t; 1) \geq 0 \geq \hat{\theta}(\bullet, \delta)'(t; -1) \quad \text{for all } t;$$

statement (b) thus holds. For the proof of statement (c), define  $\underline{\theta}(\delta) \triangleq -m(\delta)q(\delta)$  and  $\bar{\theta}(\delta) \triangleq m(\delta)(1 - q(\delta))$ . By the limit on the function  $m$  in (36), we obtain  $\lim_{\delta \downarrow 0} \bar{\theta}(\delta) = 0 = \lim_{\delta \downarrow 0} \underline{\theta}(\delta)$ . Finally, with

$$\hat{\theta}(t, \delta) \triangleq \psi \left( q(\delta) + \frac{t}{m(\delta)} \right),$$

condition (T2) holds also by the limits on  $q(\delta)$  in (36). The last statement (d) is obvious by (2).  $\square$

To illustrate the function  $\theta$  in part (c) of the above proposition and the role of condition (T2) consider a common approximation of the Heaviside function  $\mathbf{1}_{(0,\infty)}$  by the truncated hinge loss functions [53, 47]:

$$T_h(t, \delta) \triangleq \frac{1}{2\delta} [\max(t + \delta, 0) - \max(t - \delta, 0)] \quad \text{for } \delta > 0.$$

Notice that  $T_h(0, \delta) = \frac{1}{2}$  for all  $\delta$ ; thus this approximation function fails condition (T2) and as a result does not “recover” the Heaviside function as  $\delta \downarrow 0$ . When this function  $T_h(\bullet, \delta)$  is employed in sampled discretization of a probability function, as in the cited references:

$$\mathbb{P}_{\tilde{z}}(f(x, \tilde{z}) > 0) = \mathbb{E}_{\tilde{z}}[\mathbf{1}_{(0,\infty)}(f(x, \tilde{z}))] \approx \frac{1}{N} \sum_{s=1}^N T_h(f(x, z^s), \delta),$$

where  $\{z^s\}_{s=1}^N$  is a sample batch of size  $N$  of the random variable  $\tilde{z}$ , the gap between the indicator function and its approximation at the origin is less important because

of the common assumption that  $\mathbb{P}_{\tilde{z}}(f(x, \tilde{z}) = 0) = 0$ . Nevertheless, in a deterministic context, such a gap can be significant because the set  $f^{-1}(0)$  is typically of most interest for the composite function  $\mathbf{1}_{(0, \infty)}(f(x))$ . Part of the reason for the gap is due to the symmetry of the function  $T_h(\bullet, \delta)$  with respect to  $t = 0$ . To recover the Heaviside function exactly as  $\delta \downarrow 0$ , we may consider the following modified hinge loss function:

$$\tilde{T}_h(t, \delta) \triangleq \min \left\{ \max \left( \frac{t}{\delta + \sqrt{\delta}} + \frac{\sqrt{\delta}}{1 + \sqrt{\delta}}, 0 \right), 1 \right\}, \quad \text{for } \mathbf{1}_{(0, \infty)}(t), \quad (37)$$

which is derived from  $\psi$  being the identity function,

$$q(\delta) = \frac{\sqrt{\delta}}{1 + \sqrt{\delta}} \quad \text{and} \quad m(\delta) = \delta + \sqrt{\delta},$$

both satisfying the limits (36).

## 7.2. Nonifiers induced approximations

We present the other approach for deriving approximations for the Heaviside function based on “averaged functions” as defined in [18, Definition 3.1].

**Definition 7.3.** Given a locally integrable function  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  and a family of bounded mollifiers  $\{\psi(\bullet, \delta): \mathbb{R}^n \rightarrow \mathbb{R}_+; \delta \in \mathbb{R}_+\}$  that satisfy

$$\int_{\mathbb{R}^n} \psi(z, \delta) dz = 1, \\ \text{supp } \psi(\bullet, \delta) \triangleq \{z \in \mathbb{R}^n \mid \psi(z, \delta) > 0\} \subseteq \rho_\delta \mathbb{B}, \text{ for some } \{\rho_\delta\} \downarrow 0 \text{ as } \delta \downarrow 0,$$

where  $\mathbb{B}$  is a unit Euclidean ball in  $\mathbb{R}^n$ , the associated family  $\{f_\psi(\bullet, \delta)\}_{\delta \geq 0}$  of *averaged functions* is given by

$$f_\psi(x, \delta) \triangleq \int_{\mathbb{R}^n} f(x - z) \psi(z, \delta) dz = \int_{\mathbb{R}^n} f(z) \psi(x - z, \delta) dz.$$

As noted in [18],  $\psi(\bullet, \delta)$  is a probability density function whose support tends to zero as  $\delta \downarrow 0$ ; moreover,  $\psi(\bullet, \delta)$  does not need to be continuous. Focus of the references [18, 7] has been on mollifiers  $\psi(\bullet, \delta)$  that lead to smooth (i.e., continuously differentiable) averaged functions  $f_\psi(\bullet, \delta)$ . Two sets of conditions ensure the latter property: (i)  $f$  is continuous with certain special choices of the family  $\{\psi(\bullet, \delta)\}$  (see [18, Proposition 3.11]), and (ii) the mollifiers  $\psi(\bullet, \delta)$  are continuously differentiable (see [18, Proposition 3.9]). Since we are most interested in nonsmooth averaged functions, we coin the term “nonifier” for  $\psi$  with the intention that the induced averaged functions  $f_\psi(\bullet, \delta)$  are not necessarily differentiable. In what follows, we derive these averaged functions for the Heaviside function and show how they are related to those obtained from the previous truncation approach.

To begin, we take a bivariate function  $\psi: \mathbb{R} \times \mathbb{R}_{++} \rightarrow \mathbb{R}_+$  such that  $\psi(\bullet, \delta)$  is integrable on  $\mathbb{R}$  for every  $\delta > 0$  with  $\int_{-\infty}^{\infty} \psi(t, \delta) dt = 1$  and there exist end-point functions  $\underline{\psi}: \mathbb{R}_{++} \rightarrow \mathbb{R}_+$  and  $\bar{\psi}: \mathbb{R}_{++} \rightarrow \mathbb{R}_+$  satisfying condition (A0), i.e.,  $\lim_{\delta \downarrow 0} \underline{\psi}(\delta) = 0 = \lim_{\delta \downarrow 0} \bar{\psi}(\delta)$ , such that  $\text{supp}(\psi(\bullet, \delta)) \subseteq [-\underline{\psi}(\delta), \bar{\psi}(\delta)]$ .

It then follows that

$$\begin{aligned}\theta_\psi(t, \delta) &= \int_0^\infty \psi(t-s, \delta) ds = \int_{-\infty}^t \psi(s, \delta) ds, \quad \forall (t, \delta) \in \mathbb{R} \times \mathbb{R}_{++} \\ &= \int_{-\underline{\psi}(\delta)}^{\min(t, \bar{\psi}(\delta))} \psi(s, \delta) ds = \begin{cases} 1 & \text{if } t \geq \bar{\psi}(\delta) \\ \int_{-\underline{\psi}(\delta)}^t \psi(s, \delta) ds & \text{if } -\underline{\psi}(\delta) \leq t \leq \bar{\psi}(\delta) \\ 0 & \text{if } t \leq -\underline{\psi}(\delta). \end{cases}\end{aligned}$$

Defining the cumulative distribution function:

$$\widehat{\theta}_\psi(t, \delta) \triangleq \int_{-\underline{\psi}(\delta)}^t \psi(s, \delta) ds, \quad \forall (t, \delta) \in \mathbb{R} \times \mathbb{R}_{++}, \quad (38)$$

we see that  $\widehat{\theta}_\psi(\bullet, \delta)$  is equal to its own truncation; i.e.,

$$\widehat{\theta}_\psi(t, \delta) = T_{[0,1]}(\widehat{\theta}_\psi(t, \delta)) = \theta_\psi(t, \delta);$$

moreover,  $\widehat{\theta}_\psi(\bullet, \delta)$  is nondecreasing (because  $\psi(\bullet, \delta)$  is nonnegative) and satisfies condition (T1'). The following lemma pertains to conditions (T2) and (T3) for the function  $\widehat{\theta}_\psi(\bullet, \delta)$ ; no continuity of  $\psi(\bullet, \delta)$  is needed.

**Lemma 7.4.** *The following two statements hold:*

- If the one-sided limits  $\psi(t\pm; \delta) \triangleq \lim_{\tau \downarrow 0} \psi(t \pm \tau, \delta)$  exist, then  $\widehat{\theta}_\psi(\bullet, \delta)$  is B-differentiable at  $t$  with

$$\widehat{\theta}_\psi(\bullet, \delta)'(t; \pm 1) = \pm \psi(t\pm; \delta), \quad \text{respectively}; \quad (39)$$

thus condition (T3) (and (A3)) hold for  $\widehat{\theta}_\psi(\bullet, \delta)$ ; i.e.,

$$\widehat{\theta}_\psi(\bullet, \delta)'(t; 1) \geq 0 \geq \widehat{\theta}_\psi(\bullet, \delta)'(t; -1).$$

- If  $\limsup_{\delta \downarrow 0} \left\{ \left[ \underline{\psi}(\delta) + \bar{\psi}(\delta) \right] \sup_{t \in \mathbb{R}} \psi(t, \delta) \right\} < \infty$  and  $\lim_{\delta \downarrow 0} \frac{\psi(\delta)}{\bar{\psi}(\delta)} = 0$ , then  $\lim_{\delta \downarrow 0} \widehat{\theta}_\psi(0, \delta) = 0$ . Thus condition (T2) holds for  $\widehat{\theta}_\psi(\bullet, \delta)$ .

**Proof.** To show the locally Lipschitz continuity of  $\widehat{\theta}_\psi(\bullet, \delta)$  near  $t$ , let  $\varepsilon$  and  $\bar{\tau}$  be positive scalars such that

$$\tau \in [0, \bar{\tau}] \Rightarrow |\psi(t \pm \tau, \delta) - \psi(t\pm, \delta)| \leq \varepsilon.$$

Let  $t_1$  and  $t_2$  be two scalars in  $[-\bar{\tau}, \bar{\tau}]$ . Consider first the case where both  $t_1$  and  $t_2$  are on the same side of  $t$ . We can write  $t_1 = t + \tau_1$  and  $t_2 = t + \tau_2$ , with  $\bar{\tau} \geq \tau_1 \geq \tau_2 \geq 0$ .

$$\widehat{\theta}_\psi(t_1, \delta) - \widehat{\theta}_\psi(t_2, \delta) = \int_{t_2}^{t_1} \psi(s, \delta) ds = \int_{t_2}^{t_1} [\psi(s, \delta) - \psi(t+, \delta)] ds + \psi(t+, \delta) (t_1 - t_2).$$

Hence, 
$$|\widehat{\theta}_\psi(t_1, \delta) - \widehat{\theta}_\psi(t_2, \delta)| \leq (\varepsilon + \psi(t+, \delta)) (t_1 - t_2).$$

If  $t_1$  and  $t_2$  are on opposite sides of  $t$ , then can write  $t_1 = t + \tau_1$  and  $t_2 = t - \tau_2$  with  $\tau_1$  and  $\tau_2$  both in the interval  $[0, \bar{\tau}]$ . It follows that

$$\begin{aligned}\widehat{\theta}_\psi(t_1, \delta) - \widehat{\theta}_\psi(t_2, \delta) &= \int_{t_2}^t \psi(s, \delta) ds + \int_t^{t_1} \psi(s, \delta) ds \\ &= \int_{t_2}^t [\psi(s, \delta) - \psi(t-, \delta)] ds + \psi(t-, \delta) (t - t_2) \\ &\quad + \int_t^{t_1} [\psi(s, \delta) - \psi(t+, \delta)] ds + \psi(t+, \delta) (t_1 - t).\end{aligned}$$

Hence,

$$\begin{aligned}|\widehat{\theta}_\psi(t_1, \delta) - \widehat{\theta}_\psi(t_2, \delta)| &\leq (\varepsilon + \psi(t-, \delta)) (t - t_2) + (\varepsilon + \psi(t+, \delta)) (t_1 - t) \\ &\leq [\varepsilon + \max(\psi(t-, \delta), \psi(t+, \delta))] [(t - t_2) + (t_1 - t)] \\ &= [\varepsilon + \max(\psi(t-, \delta), \psi(t+, \delta))] (t_1 - t_2).\end{aligned}$$

The locally Lipschitz continuity of  $\theta_\psi(\bullet, \delta)$  near  $t$  follows. Since  $\theta_\psi(\bullet, \delta)$  is nondecreasing, the B-differentiability of  $\theta_\psi(\bullet, \delta)$  at  $t$  follows. For the two limits in (39), we prove only for the negative direction; i.e.,  $\widehat{\theta}_\psi(\bullet, \delta)'(t; -1) = -\psi(t-, \delta)$ , as the proof for the plus direction is similar (and a little more straightforward). We have

$$\begin{aligned}\widehat{\theta}_\psi(\bullet, \delta)'(t; -1) + \psi(t-, \delta) &= \lim_{\tau \downarrow 0} \frac{\widehat{\theta}_\psi(t - \tau, \delta) - \widehat{\theta}_\psi(t, \delta) + \tau \psi(t-, \delta)}{\tau} \\ &= \lim_{\tau \downarrow 0} \frac{\int_t^{t-\tau} [\psi(s, \delta) - \psi(t-, \delta)] ds}{\tau} \leq \lim_{\tau \downarrow 0} \sup_{s \in [t-\tau, t]} |\psi(s, \delta) - \psi(t-, \delta)| = 0.\end{aligned}$$

For the second statement, we have

$$\begin{aligned}\widehat{\theta}_\psi(0, \delta) &= \int_{-\underline{\psi}(\delta)}^0 \psi(s, \delta) ds \leq \underline{\psi}(\delta) \sup_{s \in \mathbb{R}} \psi(s, \delta) \\ &\leq \frac{\underline{\psi}(\delta)}{\underline{\psi}(\delta) + \overline{\psi}(\delta)} [\underline{\psi}(\delta) + \overline{\psi}(\delta)] \sup_{t \in \mathbb{R}} \psi(t, \delta) = \frac{\underline{\psi}(\delta)/\overline{\psi}(\delta)}{1 + \underline{\psi}(\delta)/\overline{\psi}(\delta)} [\underline{\psi}(\delta) + \overline{\psi}(\delta)] \sup_{t \in \mathbb{R}} \psi(t, \delta).\end{aligned}$$

Thus  $\lim_{\delta \downarrow 0} \widehat{\theta}_\psi(0, \delta) = 0$  as desired.  $\square$

In summary, starting from a nonifier  $\psi(\bullet, \delta)$  satisfying the conditions in Lemma 7.4, the cumulative distribution function  $\widehat{\theta}_\psi(\bullet, \delta)$  yields a nondecreasing p-approximation function of the open Heaviside function satisfying conditions (A1), (A2), and (A3).

We next consider the reverse; i.e., we are given a bivariate function  $\widehat{\theta}(t, \delta)$  with  $\widehat{\theta}(\bullet, \delta)$  satisfying the conditions in part (b) of Proposition 7.2 and also (T2). Since  $\widehat{\theta}(\bullet, \delta)$  is B-differentiable, it is almost everywhere differentiable. Let  $\psi(\bullet, \delta)$  be any integrable function such that  $\psi(t, \delta) = \widehat{\theta}(\bullet, \delta)'(t)$  for almost all  $t \in [-\underline{\theta}(\delta), \bar{\theta}(\delta)]$ .

Defining  $\psi(\bullet, \delta)$  to be zero outside the latter interval, we deduce that  $\psi(\bullet, \delta)$  is a well-defined nonifier; moreover

$$\widehat{\theta}(t, \delta) = \int_{-\underline{\theta}(\delta)}^t \widehat{\theta}(\bullet, \delta)'(s) ds = \int_{-\underline{\theta}(\delta)}^t \psi(s, \delta) ds = \widehat{\theta}_\psi(t, \delta).$$

Combining the two parts of the analysis, we conclude that the p-approximating functions for the open Heaviside function obtained from the truncation approach coincide with those from the nonifier approach satisfying some mild properties.

**Example 7.5.** The truncated hinge loss function

$$T_h(t, \delta/2) = \frac{1}{\delta} [\max(t + \delta/2, 0) - \max(t - \delta/2, 0)]$$

is the averaged function derived from the “symmetric” one-dimensional Steklov mollifier [18, Definition 3.8]:

$$\psi(t, \delta) = \begin{cases} 1/\delta & \text{if } |t| \leq \delta/2 \\ 0 & \text{otherwise} \end{cases},$$

whose support is the interval  $[-\delta/2, \delta/2]$ . As mentioned before, the truncated hinge loss  $T_h(\bullet, \delta/2)$  fails to p-approximate the open Heaviside function because it violates condition (A1). More interesting is the fact there do not exist nonifiers for this function that satisfy the conditions in Lemma 7.4. Nevertheless, by considering the asymmetric Steklov function:

$$\psi_a(t, \delta) = \begin{cases} \frac{1}{\underline{\psi}(\delta) + \overline{\psi}(\delta)} & \text{if } t \in [-\underline{\psi}(\delta), \overline{\psi}(\delta)] \\ 0 & \text{otherwise} \end{cases}$$

with the end bounds  $\underline{\psi}(\delta)$  and  $\overline{\psi}(\delta)$  satisfying the conditions in Lemma 7.4, we can derive a host of modified truncated hinge loss functions, such as (37), that p-approximate the open Heaviside function.  $\square$

## 8. Convergence to pseudo B-stationary solutions

We consider the approximation of a pseudo B-stationary solution of the problem (1) by a combination of penalization of the functional constraint and approximation of the composite Heaviside functions:

$$\begin{aligned} \underset{x \in X}{\text{minimize}} \quad & \widehat{\Phi}_\lambda(x, \delta) \triangleq c(x) + \underbrace{\sum_{k=1}^K \varphi_k(x) \theta_k^\varphi(g_k(x), \delta)}_{\text{denoted } \varphi(x, \delta)} \\ & + \lambda \max \left( \underbrace{\sum_{\ell=1}^L \phi_\ell(x) \theta_\ell^\phi(h_\ell(x), \delta) - b}_{\text{denoted } \phi(x, \delta)}, 0 \right), \end{aligned} \quad (40)$$

where for each pair  $(k, \ell) \in [K] \times [L]$ ,  $\theta_k^\varphi(\bullet, \delta)$  and  $\theta_\ell^\phi(\bullet, \delta)$  are p-approximations of the Heaviside function  $\mathbf{1}_{(0, \infty)}$  with support in the interval  $[-\underline{\theta}_k^\varphi(\delta), \bar{\theta}_k^\varphi(\delta)]$  and  $[-\underline{\theta}_\ell^\phi(\delta), \bar{\theta}_\ell^\phi(\delta)]$ , respectively, that shrink to zero when  $\delta \downarrow 0$ ; in particular, (A1), (A2), and (A3) in Definition 7.1 are satisfied by these approximating functions. With the functions  $c$  and each  $\varphi_k$  being Lipschitz continuous on  $X$  with Lipschitz constants  $\text{Lip}_c$  and  $\text{Lip}_\varphi$ , respectively, we take  $\lambda$  satisfying (32).

One of the conditions that we will impose in the analysis is a sign condition, labelled (C2) below, on the directional derivatives of the functions  $g_k$  and  $h_\ell$ . The lemma below shows that this will hold if these are convex piecewise affine functions.

**Lemma 8.1.** *Let  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  be a convex piecewise affine function. For every vector  $\bar{x}$ , there exists a neighborhood  $\mathcal{N}$  of  $\bar{x}$  such that for all  $v \in \mathbb{R}^n$ ,*

$$f'(\bar{x}; v) \leq 0 \Rightarrow f'(x; v) \leq 0, \quad \forall x \in \mathcal{N}.$$

**Proof.** Since  $f$  is convex piecewise affine, it can be written as the pointwise maximum of finitely many affine functions [12, Proposition 4.4.6]; i.e.,

$$f(x) = \max_{1 \leq i \leq I} ((a^i)^\top x + b_i), \quad \forall x \in \mathbb{R}^n,$$

for some positive integer  $I$ ,  $n$ -vectors  $\{a^i\}_{i=1}^I$ , and scalars  $\{b_i\}_{i=1}^I$ . For any  $x$ , let

$$\mathcal{A}(x) \triangleq \{i \in [I] \mid f(x) = (a^i)^\top x + b_i\}$$

be the maximizing index set of the affine pieces of  $f$ . It then follows that for the given  $\bar{x}$ , there exists a neighborhood  $\mathcal{N}$  such that  $\mathcal{A}(x) \subseteq \mathcal{A}(\bar{x})$  for all  $x \in \mathcal{N}$ . Since

$$f'(x; v) = \max_{i \in \mathcal{A}(x)} ((a^i)^\top v + b_i), \quad \forall (x, v) \in \mathbb{R}^n \times \mathbb{R}^n,$$

the desired conclusion of the lemma follows readily.  $\square$

Let  $\{\delta_\nu\}$  be a sequence of positive scalars converging to zero. Let  $\{x^\nu\}$  be a corresponding sequence of d-stationary points of  $\widehat{\Phi}_\lambda(\bullet, \delta_\nu)$  on  $X$  so that we have  $\widehat{\Phi}_\lambda(\bullet, \delta_\nu)'(x^\nu; v) \geq 0$  for all  $v \in \mathcal{T}(X; x^\nu)$ . Suppose that  $\{x^\nu\}$  converges to the limit  $x^*$ . In what follows, we show that  $x^*$  is a pseudo B-stationary point of (1) by verifying two things under the condition (32) on  $\lambda$ : (a)  $x^*$  is feasible to (1), and (b) the implication (18), which we restate below in terms of the vector  $x^*$  on hand:

$$v \in \mathcal{L}(S_{\text{ps}}(x^*); x^*) \Rightarrow \Phi(\bullet; x^*)'(x^*; v) \geq 0. \quad (41)$$

Letting  $\sigma_k(x; v) \triangleq \text{sgn}(g'_k(x; v))$  with  $\text{sgn}(0)$  defined to be zero, we write the directional derivative of  $\varphi(\bullet, \delta_\nu)$  at  $x^\nu$  as the sum of 3 terms:

$$\begin{aligned} & \varphi(\bullet, \delta_\nu)'(x^\nu; v) \\ &= \sum_{k=1}^K [\varphi'_k(x^\nu; v) \theta_k^\varphi(g_k(x^\nu), \delta_\nu) + \varphi_k(x^\nu) |g'_k(x^\nu; v)| (\theta_k^\varphi(\bullet, \delta_\nu))'(g_k(x^\nu); \sigma_k(x^\nu; v))] \\ &= T_{>}^\varphi(x^\nu; v) + T_{=}^\varphi(x^\nu; v) + T_{<}^\varphi(x^\nu; v), \quad \text{where} \end{aligned}$$

$$\begin{aligned}
T_{>}^{\varphi}(x^{\nu}; v) &= \sum_{k \in \mathcal{K}_{>}(x^*)} \left[ \varphi'_k(x^{\nu}; v) \theta_k^{\varphi}(g_k(x^{\nu}), \delta_{\nu}) \right. \\
&\quad \left. + \varphi_k(x^{\nu}) |g'_k(x^{\nu}; v)| (\theta_k^{\varphi}(\bullet, \delta_{\nu}))'(g_k(x^{\nu}); \sigma_k(x^{\nu}; v)) \right] \\
T_{=}^{\varphi}(x^{\nu}; v) &= \sum_{k \in \mathcal{K}_{=}(x^*)} \left[ \varphi'_k(x^{\nu}; v) \theta_k^{\varphi}(g_k(x^{\nu}), \delta_{\nu}) \right. \\
&\quad \left. + \varphi_k(x^{\nu}) |g'_k(x^{\nu}; v)| (\theta_k^{\varphi}(\bullet, \delta_{\nu}))'(g_k(x^{\nu}); \sigma_k(x^{\nu}; v)) \right] \\
T_{<}^{\varphi}(x^{\nu}; v) &= \sum_{k \in \mathcal{K}_{<}(x^*)} \left[ \varphi'_k(x^{\nu}; v) \theta_k^{\varphi}(g_k(x^{\nu}), \delta_{\nu}) \right. \\
&\quad \left. + \varphi_k(x^{\nu}) |g'_k(x^{\nu}; v)| (\theta_k^{\varphi}(\bullet, \delta_{\nu}))'(g_k(x^{\nu}); \sigma_k(x^{\nu}; v)) \right].
\end{aligned}$$

Consider the sum  $T_{>}^{\varphi}(x^{\nu}; v)$ . For each  $k \in \mathcal{K}_{>}(x^*)$ , we have  $g_k(x^*) > 0$ . Hence  $g_k(x^{\nu}) > \bar{\theta}_k^{\varphi}(\delta_{\nu})$  for all  $\nu$  sufficiently large; this yields, by (A1):

$$\theta_k(g_k(x^{\nu}), \delta_{\nu}) = 1 \quad \text{and} \quad \theta'_k(\bullet, \delta_{\nu})(g_k(x^{\nu}); \sigma_k(x^{\nu}; v)) = 0.$$

Thus, 
$$T_{>}^{\varphi}(x^{\nu}; v) = \sum_{k \in \mathcal{K}_{>}(x^*)} \phi'_k(x^{\nu}; v).$$

Similarly, we can show that  $T_{<}^{\varphi}(x^{\nu}; v) = 0$  for all  $\nu$  sufficiently large. The rest of the proof is divided into three parts.

• To analyze the term  $T_{=}^{\varphi}(x^{\nu}; v)$  and the corresponding term in  $\phi(\bullet, \delta_{\nu})'(x^*; v)$ , we need assumptions (C1) and (C2):

(C1) There exists an open neighborhood  $\mathcal{N}_*$  such that on  $X \cap \mathcal{N}_*$ , the functions  $\{\varphi_k\}_{k \in \mathcal{K}_{=}(x^*)}$  and  $\{\phi_{\ell}\}_{\ell \in \mathcal{L}_{=}(x^*)}$  are nonnegative; this is a pointwise sign condition related to those in Proposition 4.1, demanding in particular the nonnegativity of these functions in a region around  $x^*$  and possibly outside the respective sets  $X \cap g_k^{-1}(-\infty, 0]$  and  $X \cap h_{\ell}^{-1}(\infty, 0]$ .

(C2) The implications

$$\left\{ \begin{array}{l} g'_k(x^*; v) \leq 0 \quad \forall k \in \mathcal{K}_{=}(x^*) \\ v \in \mathcal{T}(X; x^*) \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} g'_k(x; v) \leq 0 \text{ for all } k \in \mathcal{K}_{=}(x^*) \\ \text{and all } x \in \mathcal{N}_* \end{array} \right\} \quad (42)$$

$$\left\{ \begin{array}{l} h'_{\ell}(x^*; v) \leq 0 \quad \forall \ell \in \mathcal{L}_{=}(x^*) \\ v \in \mathcal{T}(X; x^*) \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} h'_{\ell}(x; v) \leq 0 \text{ for all } \ell \in \mathcal{L}_{=}(x^*) \\ \text{and all } x \in \mathcal{N}_* \end{array} \right\} \quad (43)$$

are motivated by Lemma 8.1 which provides sufficient conditions for them to hold.

Under (C1) and (C2), we deduce that if  $v$  satisfies the left-hand condition in (42), then by (34) applied to  $\theta_k^{\varphi}$ , we obtain, for all  $\nu$  sufficiently large and all  $k \in \mathcal{K}_{=}(x^*)$ ,

$$\varphi_k(x^{\nu}) |g'_k(x^{\nu}; v)| \theta_k^{\varphi}(\bullet, \delta_{\nu})(g_k(x^{\nu}); \sigma_k(x^{\nu}; v)) \leq 0$$

Hence, for such  $\nu$ , we have

$$T_{k;=}^{\varphi}(x^{\nu}; v) \leq \varphi'_k(x^{\nu}; v) \theta_k^{\varphi}(g_k(x^{\nu}), \delta_{\nu}).$$

Consequently, we deduce, for all  $\nu$  sufficiently large,

$$g'_k(x^*; v) \leq 0 \quad \forall k \in \mathcal{K}_{=}(x^*) \quad \text{and} \quad v \in \mathcal{T}(X; x^*) \quad (44)$$

$$\Rightarrow \varphi(\bullet, \delta_{\nu})'(x^{\nu}, v) \leq \sum_{k \in \mathcal{K}_{>}(x^*)} \varphi'_k(x^{\nu}; v) + \sum_{k \in \mathcal{K}_{=}(x^*)} \varphi'_k(x^{\nu}; v) \theta_k^{\varphi}(g_k(x^{\nu}), \delta_{\nu}).$$

Similarly, we also have, for all  $\nu$  sufficiently large,

$$h'_\ell(x^*; v) \leq 0 \quad \forall \ell \in \mathcal{L}_=(x^*) \quad \text{and} \quad v \in \mathcal{T}(X; x^*) \quad (45)$$

$$\Rightarrow \phi(\bullet, \delta_\nu)'(x^\nu, v) \leq \sum_{\ell \in \mathcal{L}_>(x^*)} \phi'_\ell(x^\nu; v) + \sum_{\ell \in \mathcal{L}_=(x^*)} \phi'_\ell(x^\nu; v) \theta_k^\phi(h_\ell(x^\nu), \delta_\nu).$$

• To establish the feasibility of  $x^*$  for the problem (1), we postulate three more assumptions: as we will see below, assumption (C3) ensures the objective recovery of the sequence  $\{x^\nu\}$ ; (C4) is the same as that in Theorem 6.4 and (C5) is the Clarke regularity on several key functions.

(C3) For all  $k \in \mathcal{K}_=(x^*)$  and  $\ell \in \mathcal{L}_=(x^*)$ , the limits

$$\lim_{\nu \rightarrow \infty} \theta_k^\varphi(g_k(x^\nu), \delta_\nu) = 0 \quad \text{and} \quad \lim_{\nu \rightarrow \infty} \theta_\ell^\phi(h_\ell(x^\nu), \delta_\nu) = 0 \quad (46)$$

are *functional consistency* requirements of the sequence  $\{x^\nu\}$  in the following sense. These limits ensure that for all pairs  $(k, \ell)$ ,

$$\lim_{\nu \rightarrow \infty} \theta_k^\varphi(g_k(x^\nu), \delta_\nu) = \mathbf{1}_{(0, \infty)}(g_k(x^*)) \quad \text{and} \quad \lim_{\nu \rightarrow \infty} \theta_\ell^\phi(h_\ell(x^\nu), \delta_\nu) = \mathbf{1}_{(0, \infty)}(h_\ell(x^*)).$$

Together with the same limits for  $k \in \mathcal{K}_>(x^*) \cup \mathcal{K}_<(x^*)$  and  $\ell \in \mathcal{L}_>(x^*) \cup \mathcal{L}_<(x^*)$ , we deduce in particular that

$$\lim_{\nu \rightarrow \infty} \left[ c(x^\nu) + \sum_{k=1}^K \varphi_k(x^*) \theta_k^\varphi(g_k(x^\nu), \delta_\nu) \right] = \Phi(x^*),$$

which is reasonable to postulate in order for  $x^*$  to be a stationarity point of some kind for the objective function of the original problem (1). If the approximation function  $\theta_k$  is chosen as in part (c) of Proposition 7.2, i.e. if

$$\theta_k(t, \delta) \triangleq T_{[0,1]} \left( \psi_k \left( q_k(\delta) + \frac{t}{m_k(\delta)} \right) \right)$$

where the functions  $\psi_k$ ,  $q_k$ , and  $m_k$  are as specified in the proposition, then the limit (46) holds if  $\lim_{\nu \rightarrow \infty} g_k(x^\nu)/m_k(\delta_\nu) = 0$ . Admittedly, this is a condition on the sequence of iterates  $\{x^\nu\}$  relative to the sequence of parameters  $\{\delta_\nu\}$  that needs to be addressed from the source of the iterates, e.g., as generated by an iterative algorithm. Regrettably, the design of such algorithms is beyond the scope of the present work but will be the focus of a subsequent computational study which will be guided by this high-level background result. The limit on the sequence  $\{\theta_\ell^\phi(h_\ell(x^\nu), \delta_\nu)\}$  has to do with the satisfaction of the functional constraint and facilitates the demonstration of the desired pseudo B-stationarity of the limit  $x^*$ ; this will become clear in what follows. Subsequently, we will establish a weaker stationarity property of the limit  $x^*$  without (C3).

(C4) For any  $x \in X$  such that we have  $\sum_{\ell \in \mathcal{L}_>(x)} \phi_\ell(x) > b$ , there exists a vector  $\bar{v} \in \mathcal{T}(\widehat{S}_{\text{ps}}(x); x)$  with unit length satisfying:  $\sum_{\ell \in \mathcal{L}_>(x)} \phi'_\ell(x; \bar{v}) \leq -1$ ; and

(C5) the functions  $c$ ,  $\{\varphi_k\}_{k \in \mathcal{K}_>(x^*)}$ , and  $\{\phi_\ell\}_{\ell \in \mathcal{L}_>(x^*)}$  are Clarke regular [8] at  $x^*$ .

In the context of a B-differentiable function  $f: \mathcal{O} \rightarrow \mathbb{R}$  (which these functions are), Clarke regularity at a vector  $\bar{x} \in \mathcal{O}$  means that for all sequences  $\{z^\nu\}$  converging to  $\bar{x}$ , it holds that

$$\limsup_{\nu \rightarrow \infty} f'(z^\nu; v) \leq f'(\bar{x}; v), \quad \forall v \in \mathbb{R}^n.$$

Condition (C4) requires that for any  $x \in X$  that is infeasible to problem (1), the constraint function  $\sum_{\ell \in \mathcal{L}_{>}(x)} \phi_\ell(\bullet)$  has a descent direction at  $x$ . This condition is in the same spirit as a classical one since the early days of the theory of exact penalty methods [46] for the recovery of feasibility (and hence stationarity) in a penalized problem; see also [12, Proposition 9.2.2 (a)]. As in problems without the Heaviside function, condition (C4) can be related to the concept of weak-sharp minima [23] properly extended. We give a brief discussion of the connection in an Appendix. The Clarke regularity in (C5) holds in particular if the considered functions are the composition of a convex function with a smooth mapping [49, Exercise 10.25].

Continuing the analysis, assume for the sake of contradiction that

$$\sum_{\ell=1}^L \phi_\ell(x^*) \mathbf{1}_{(0,\infty)}(h_\ell(x^*)) > b.$$

Then for all  $\nu$  sufficiently large we have

$$\phi(x^\nu, \delta_\nu) = \sum_{\ell=1}^L \phi_\ell(x^\nu) \theta_\ell^\varphi(h_\ell(x^\nu), \delta_\nu) - b > 0.$$

Hence, for the vector  $\bar{v}$  in (C4), we have

$$\begin{aligned} 0 &\leq \widehat{\Phi}_\lambda(\bullet, \delta_\nu)'(x^\nu; \bar{v}) = c'(x^\nu; \bar{v}) + \varphi(\bullet, \delta_\nu)'(x^\nu; \bar{v}) + \lambda \phi(\bullet, \delta_\nu)'(x^\nu; \bar{v}) \\ &\leq c'(x^\nu; \bar{v}) + \sum_{k \in \mathcal{K}_{>}(x^*)} \varphi'_k(x^\nu; \bar{v}) + \sum_{k \in \mathcal{K}_{=}(x^*)} \varphi'_k(x^\nu; \bar{v}) \theta_k^\varphi(g_k(x^\nu), \delta_\nu) \\ &\quad + \lambda \left[ \sum_{\ell \in \mathcal{L}_{>}(x^*)} \phi'_\ell(x^\nu; \bar{v}) + \sum_{\ell \in \mathcal{L}_{=}(x^*)} \phi'_\ell(x^\nu; \bar{v}) \theta_k^\phi(h_\ell(x^\nu), \delta_\nu) \right] \quad (47) \\ &\leq (\text{Lip}_c + K \text{Lip}_\varphi) + \lambda \left[ \sum_{\ell \in \mathcal{L}_{>}(x^*)} \phi'_\ell(x^\nu; \bar{v}) + \text{Lip}_\phi \sum_{\ell \in \mathcal{L}_{=}(x^*)} \theta_k^\phi(h_\ell(x^\nu), \delta_\nu) \right], \end{aligned}$$

where  $\text{Lip}_\phi$  is a local Lipschitz constant of  $\phi_\ell$  near  $x^*$ . Taking the limit  $\nu \rightarrow \infty$  and using (46), we obtain, by the Clarke regularity of the functions  $\{\phi_\ell\}_{\ell \in \mathcal{L}_{>}(x^*)}$  at  $x^*$ ,

$$\begin{aligned} 0 &\leq \text{Lip}_c + K \text{Lip}_\varphi + \lambda \sum_{\ell \in \mathcal{L}_{>}(x^*)} \limsup_{\nu \rightarrow \infty} \phi'_\ell(x^\nu; \bar{v}) \\ &\leq \text{Lip}_c + K \text{Lip}_\varphi + \lambda \sum_{\ell \in \mathcal{L}_{>}(x^*)} \phi'_\ell(x^*; \bar{v}) \leq \text{Lip}_c + K \text{Lip}_\varphi - \lambda, \end{aligned}$$

This contradiction completes the feasibility proof of  $x^*$  for (1).

• Finally, to complete the proof of the desired implication (41), suppose that  $x^*$  satisfies the functional constraint as an equality; i.e.,

$$\sum_{\ell=1}^L \phi_\ell(x^*) \mathbf{1}_{(0,\infty)}(h_\ell(x^*)) = b.$$

Let  $v \in \mathcal{L}(S_{\text{ps}}(x^*); x^*)$  be arbitrary. By Lemma 6.3 we have

$$\begin{aligned}
 \limsup_{\nu \rightarrow \infty} [\max(\phi(\bullet, \delta_\nu), 0)]'(x^\nu; v) &\leq \limsup_{\nu \rightarrow \infty} \max(\phi(\bullet, \delta_\nu)'(x^\nu; v), 0), \\
 &\leq \limsup_{\nu \rightarrow \infty} \max \left( \sum_{\ell \in \mathcal{L}_>(x^*)} \phi'_\ell(x^\nu; v) + \sum_{\ell \in \mathcal{L}_=(x^*)} \phi'_\ell(x^\nu; v) \theta_\ell^\phi(h_\ell(x^\nu), \delta_\nu), 0 \right) \text{ by (45)} \\
 &\leq \max \left( \limsup_{\nu \rightarrow \infty} \left[ \sum_{\ell \in \mathcal{L}_>(x^*)} \phi'_\ell(x^\nu; v) + \text{Lip}_\phi \sum_{\ell \in \mathcal{L}_=(x^*)} \theta_\ell^\phi(h_\ell(x^\nu), \delta_\nu) \right], 0 \right) \\
 &\leq \max \left( \sum_{\ell \in \mathcal{L}_>(x^*)} \phi'_\ell(x^*; v), 0 \right), \text{ by Clarke regularity and (C3)} \\
 &= 0, \text{ because } \sum_{\ell \in \mathcal{L}_>(x^*)} \phi'_\ell(x^*; v) \leq 0 \text{ is a stipulation on } v.
 \end{aligned}$$

In the other case where  $\sum_{\ell=1}^L \phi_\ell(x^*) \mathbf{1}_{(0, \infty)}(h_\ell(x^*)) < b$ , we must have  $\phi(x^\nu, \delta_\nu) < b$  for all  $\nu$  sufficiently large. Thus,  $[\max(\phi(\bullet, \delta_\nu), 0)]'(x^\nu; v) = 0$  for all such  $\nu$ . Consequently, in either case, it follows that

$$\begin{aligned}
 0 &\leq \limsup_{\nu \rightarrow \infty} [c'(x^\nu; v) + \varphi(\bullet, \delta_\nu)'(x^\nu; v)] \\
 &\leq c'(x^*; v) + \sum_{k \in \mathcal{K}_>(x^*)} \varphi'_k(x^*; v), \text{ by (44), (C3), and Clarke regularity,}
 \end{aligned}$$

establishing the right-hand side of (41). We have thus proved the following main result of this section.

**Theorem 8.2.** *Under the blanket assumption of problem (1), let  $c$  and each  $\varphi_k$  be Lipschitz continuous on  $X$  with Lipschitz constants  $\text{Lip}_c$  and  $\text{Lip}_\varphi$ , respectively. Let  $\{\delta_\nu\}$  be a sequence of positive scalars converging to zero and for each  $\nu$ , let  $x^\nu$  be a  $d$ -stationary point of  $\hat{\Phi}_\lambda(\bullet, \delta_\nu)$  on  $X$  with  $\lambda$  satisfying (32). Suppose  $\lim_{\nu \rightarrow \infty} x^\nu = x^*$ . Under the assumptions (C1)–(C5) it holds that  $x^*$  is a pseudo B-stationary solution of (1).  $\square$*

Without (C3), the desired pseudo B-stationarity of the limit  $x^*$  as defined in Definition 5.2 is in jeopardy. By strengthening (C4) and (C5), we can still establish a weak pseudo B-stationarity property of  $x^*$ . The two strengthened conditions are:

(C4') There exists a vector  $\bar{v} \in \mathcal{T}(\hat{S}_{\text{ps}}(x^*); x^*)$  with unit length satisfying:

$$\sum_{\ell \in \mathcal{L}_>(\bar{x})} \phi'_\ell(x^*; \bar{v}) + \sum_{\ell \in \mathcal{L}_=(\bar{x})} \max(\phi'_\ell(x^*; \bar{v}), 0) \leq -1. \quad (48)$$

While more demanding than (C4), this strengthened condition is still in the spirit of the common requirements in the theory of exact penalization (see conditions (a) and (b) in [12, Theorem 9.2.1]).

(C5') In addition to those in (C5), the functions  $\{\varphi_k\}_{k \in \mathcal{K}_=(x^*)}$ , and  $\{\phi_\ell\}_{\ell \in \mathcal{L}_=(x^*)}$  are also Clarke regular at  $x^*$ .

**Proposition 8.3.** Under the blanket assumption of problem (1), let  $c$  and each  $\varphi_k$  be Lipschitz continuous on  $X$  with Lipschitz constants  $Lip_c$  and  $Lip_{\varphi}$ , respectively. Let  $\{\delta_\nu\}$  be a sequence of positive scalars converging to zero and for each  $\nu$ , let  $x^\nu$  be a  $d$ -stationary point of  $\widehat{\Phi}_\lambda(\bullet, \delta_\nu)$  on  $X$  with  $\lambda$  satisfying (32). Suppose  $\lim_{\nu \rightarrow \infty} x^\nu = x^*$ . Under the assumptions (C1), (C2), (C4'), and (C5'),  $x^*$  is feasible to (1) and there exist scalars  $\{\xi_k^*\}_{k \in \mathcal{K}=(x^*)} \cup \{\mu_\ell^*\}_{\ell \in \mathcal{L}=(x^*)} \subset [0, 1]$  such that  $x^*$  is a  $B$ -stationary solution of:

$$\begin{aligned} & \underset{x \in X}{\text{minimize}} \quad \Phi_{\geq}^*(x; x^*) \triangleq c(x) + \sum_{k \in \mathcal{K}_{>}(x^*)} \varphi_k(x) + \underbrace{\sum_{k \in \mathcal{K}=(x^*)} \xi_k^* \varphi_k(x)}_{\text{extra term}} \\ & \text{subject to} \quad \left. \begin{aligned} & \sum_{\ell \in \mathcal{L}_{>}(x^*)} \phi_\ell(x) + \underbrace{\sum_{\ell \in \mathcal{L}=(x^*)} \mu_\ell^* \phi_\ell(x)}_{\text{extra term}} \leq b \\ & g_k(x) \leq 0 \quad \forall k \in \mathcal{K}_{\leq}(x^*) \\ & g_k(x) \geq 0 \quad \forall k \in \mathcal{K}_{>}(x^*) \\ & h_\ell(x) \leq 0 \quad \forall \ell \in \mathcal{L}_{\leq}(x^*) \\ & h_\ell(x) \geq 0 \quad \forall \ell \in \mathcal{L}_{>}(x^*) \end{aligned} \right\}. \end{aligned} \quad (49)$$

and

**Proof.** Belonging to the interval  $[0, 1]$ , the sequences

$$\{\theta_k^\varphi(g_k(x^\nu), \delta_\nu)\}_{k \in \mathcal{K}=(x^*)} \quad \text{and} \quad \{\theta_\ell^\phi(h_\ell(x^\nu), \delta_\nu)\}_{\ell \in \mathcal{L}=(x^*)}$$

have accumulation points, say  $\{\xi_k^*\}_{k \in \mathcal{K}=(x^*)}$  and  $\{\mu_\ell^*\}_{\ell \in \mathcal{L}=(x^*)}$ , which we may assume, without loss of generality are limits of the displayed sequences, respectively. We claim that  $x^*$  satisfies the functional constraint in (49). Assume otherwise. Then

$$\sum_{\ell \in \mathcal{L}_{>}(x^*)} \phi_\ell(x^*) + \sum_{\ell \in \mathcal{L}=(x^*)} \mu_\ell^* \phi_\ell(x^*) > b.$$

Since  $\lim_{\nu \rightarrow \infty} \phi(x^\nu, \delta_\nu) = \sum_{\ell \in \mathcal{L}_{>}(x^*)} \phi_\ell(x^*) + \sum_{\ell \in \mathcal{L}=(x^*)} \mu_\ell^* \phi_\ell(x^*)$

it follows that  $\phi(x^\nu, \delta_\nu) > b$  for all  $\nu$  sufficiently large. Continuing (47), we have

$$\begin{aligned} 0 & \leq \widehat{\Phi}_\lambda(\bullet, \delta_\nu)'(x^\nu; \bar{v}) \\ & \leq c'(x^\nu; \bar{v}) + \sum_{k \in \mathcal{K}_{>}(x^*)} \varphi'_k(x^\nu; \bar{v}) + \sum_{k \in \mathcal{K}=(x^*)} \varphi'_k(x^\nu; \bar{v}) \theta_k^\varphi(g_k(x^\nu), \delta_\nu) \\ & \quad + \lambda \left[ \sum_{\ell \in \mathcal{L}_{>}(x^*)} \phi'_\ell(x^\nu; \bar{v}) + \sum_{\ell \in \mathcal{L}=(x^*)} \phi'_\ell(x^\nu; \bar{v}) \theta_\ell^\phi(h_\ell(x^\nu), \delta_\nu) \right] \\ & \leq (Lip_c + K Lip_\varphi) + \lambda \left[ \sum_{\ell \in \mathcal{L}_{>}(x^*)} \phi'_\ell(x^\nu; \bar{v}) + \sum_{\ell \in \mathcal{L}=(x^*)} \max(\phi'_\ell(x^\nu; \bar{v}), 0) \right] \end{aligned}$$

because both  $\theta_k^\varphi(g_k(x^\nu), \delta_\nu)$  and  $\theta_\ell^\phi(h_\ell(x^\nu), \delta_\nu)$  are in  $[0, 1]$ .

The strengthened condition (48) and the Clarke regularity condition (C5') then yield a contradiction by letting  $\nu \rightarrow \infty$ . This shows that

$$\sum_{\ell \in \mathcal{L}_{>}(x^*)} \phi_{\ell}(x^*) + \sum_{\ell \in \mathcal{L}_{=}(x^*)} \mu_{\ell}^* \phi_{\ell}(x^*) \leq b.$$

Since  $\mu_{\ell}^* \phi_{\ell}(x^*) \geq 0$  for all  $\ell \in \mathcal{L}_{>}(x^*)$ , we have

$$\sum_{\ell \in \mathcal{L}_{>}(x^*)} \phi_{\ell}(x^*) + \sum_{\ell \in \mathcal{L}_{=}(x^*)} \mu_{\ell}^* \phi_{\ell}(x^*) \geq \sum_{\ell=1}^L \phi_{\ell}(x^*) \mathbf{1}_{(0,\infty)}(h_{\ell}(x^*)),$$

thus  $x^*$  is feasible to (1). Finally, the proof that  $x^*$  is a B-stationary solution of (49) is similar to that in the last part of the proof of Theorem 8.2. The details are not repeated.  $\square$

### 8.1. An illustrative example

We use the following slight modification of the function (6) to illustrate the sign assumptions in Theorems 6.4 and 8.2 on the simplified problem:

$$\underset{x \in X}{\text{minimize}} \Psi(x), \quad \text{where } \Psi(x) = \begin{cases} \psi_1(x) & \text{if } a \leq f(x) < b \\ \psi_2(x) & \text{if } f(x) < a \\ \psi_3(x) & \text{if } f(x) \geq b. \end{cases} \quad (50)$$

Similar to the previous derivation, we can write

$$\Psi(x) = \psi_3(x) + (\psi_1(x) - \psi_3(x)) \mathbf{1}_{(0,\infty)}(b - f(x)) + (\psi_2(x) - \psi_1(x)) \mathbf{1}_{(0,\infty)}(a - f(x)).$$

Given a vector  $\bar{x} \in X$ , there are five pseudo stationarity problems (15) at  $\bar{x}$  for the problem  $\underset{x \in X}{\text{minimize}} \Psi(x)$  depending on the value of  $f(\bar{x})$ ; these problems are

- [  $\underset{x \in X}{\text{minimize}} \psi_1(x)$  ] if  $f(\bar{x}) \in (a, b)$ ;
- [  $\underset{x \in X}{\text{minimize}} \psi_2(x)$  ] if  $f(\bar{x}) \in (-\infty, a)$ ;
- [  $\underset{x \in X}{\text{minimize}} \psi_3(x)$  ] if  $f(\bar{x}) \in (b, \infty)$ ;
- [  $\underset{x \in X}{\text{minimize}} \psi_1(x)$  subject to  $f(x) \geq a$  ] if  $f(\bar{x}) = a$ ;
- [  $\underset{x \in X}{\text{minimize}} \psi_3(x)$  subject to  $f(x) \geq b$  ] if  $f(\bar{x}) = b$ .

Clearly, not all stationary solutions of the above five problems are local minimizers of  $\Psi$  on  $X$ ; nevertheless, they provide candidate minimizers. The sign conditions in Theorem 6.4 require that:

- $[f(x) = a, x \in X] \Rightarrow \psi_2(x) \geq \psi_1(x)$ ; and
- $[f(x) = b, x \in X] \Rightarrow \psi_1(x) \geq \psi_3(x)$ .

Under these conditions, which basically stipulate that the function  $\Psi$  can not rise up at a point of discontinuity, the epi-hypographical approach will produce a pseudo B-stationary solution of  $\Psi$  on  $X$ . In contrast, the pointwise sign stipulations at  $x^*$  in condition (i) of Theorem 8.2 require that

- $f(x^*) = a \Rightarrow \psi_2(x) \geq \psi_1(x)$  in a neighborhood of  $x^*$ , and
- $f(x^*) = b \Rightarrow \psi_1(x) \geq \psi_3(x)$  in a neighborhood of  $x^*$ .

Condition (42) becomes:

- there exists a neighborhood  $\mathcal{N}$  of  $\bar{x}$  such that

$$\left. \begin{array}{l} f'(x^*; v) \geq 0 \\ v \in \mathcal{T}(X; x^*) \end{array} \right\} \Rightarrow f'(x; v) \geq 0 \quad \forall x \in \mathcal{N}.$$

Along with the objective consistency condition of the iterates, the above sign stipulations provide sufficient conditions for  $x^*$  to be a pseudo B-stationary solution of  $\Psi$  on  $X$ .

## 9. Concluding remarks

At the completion of the paper, the authors are grateful to receive a preprint [50] in which the author introduces two concepts of consistent approximations in composite optimization; see Definition 2.2 therein. The setting of this reference is the minimization of an objective that is the sum of an extended-valued constraint indicator function and a composite function  $h \circ F(x)$  where  $h$  is a convex extended-valued function and  $F$  is a vector function whose components are locally Lipschitz continuous functions. While in principle, the consistency concepts defined therein can be extended to any optimization problem without bother of its structure, the challenge is twofold:

- how to define the approximations, and
- establishing the consistency of the derived approximations.

Although the problem (40) is an approximation of (1), it is rather doubtful if the theory in the reference could be applied to our context. For one thing, a product  $\psi \mathbf{1}_{(0, \infty)}(f)$  is very different from the composite family  $h \circ F$  with  $h$  and  $F$  as stated. As one can see, our analysis makes extensive use of the product form and properties of the approximations of the Heavisde functions, in particular requiring assumptions that are akin to such a structure. Hopefully, our work will provide a motivation to extend the theory of consistent approximations to broader composite classes of discontinuous functions.

**Acknowledgements.** The authors are grateful to the referees for their insightful comments that have helped to improve the presentation of the manuscript.

## A. Appendix: (C4) and weak sharp minima

Consider the optimization problem

$$\underset{x \in S}{\text{minimize}} \quad f(x), \tag{51}$$

where  $f$  is a continuous function bounded below on the closed set  $S$ . Let

$$f_{\min} \triangleq \underset{x \in S}{\text{minimum}} \quad f(x) \quad \text{and} \quad \mathcal{F}_{\min} \triangleq \underset{x \in S}{\text{argmin}} \quad f(x).$$

The problem (51), or the pair  $(f, S)$ , is said to have weak sharp minima if there exists a constant  $\eta > 0$  such that

$$f(x) - f_{\min} \geq \eta^{-1} \text{dist}(x, \mathcal{F}_{\min}) \quad \forall x \in S.$$

The definition of weak sharp minima was introduced in the Ph.D. Thesis of Ferris [23]. Extensive discussion of this property and its role in optimization can be found in [19, Section 6.5]. The proposition below contains 3 statements. Statement (a) asserts the existence of weak sharp minima for (51). Statement (b) is known as Takahashi condition [52] in nonlinear analysis; it implies in particular the existence of a global minimizer to the optimization problem in question. Statement (c) is the key to connecting condition (C4) to weak sharp minima. Proof of the implication (b) is by the renowned Ekeland's variational principle [15]; proof of the proposition can be found in [19, Section 6.5]; see also [12, Section 8.5.4].

**Proposition A.1.** *Let  $S$  be a closed set in  $\mathbb{R}^n$  and  $f$  be a continuous real-valued function defined and bounded below on  $S$ . Let  $f_{\inf}$  denote the infimum value of  $f$  on  $S$ . The two statements (a) and (b) are equivalent and either statement is implied by statement (c):*

(a) *The pair  $(f, S)$  has weak sharp minima; i.e.,  $\mathcal{F}_{\min} \neq \emptyset$  and for some constant  $c > 0$  we have*

$$f(x) - f_{\min} \geq c \text{dist}(x, \mathcal{F}_{\min}), \quad \forall x \in S. \quad (52)$$

(b) *There exists a constant  $c > 0$  such that for each  $x \in S$  with  $f(x) > f_{\inf}$ , there exists a vector  $y \in S$  distinct from  $x$  such that*

$$f(y) + c \|x - y\| \leq f(x). \quad (53)$$

(c) *Suppose that  $f$  is  $B$ -differentiable on  $S$  and there exists a scalar  $\delta > 0$  such that for every  $x \in S$  with  $f(x) > f_{\inf}$ , a vector  $d \in \mathcal{T}(S; x)$  with unit (Euclidean) length exists satisfying  $f'(x; d) \leq -\delta$ .*

*Moreover, if  $f$  is a convex function and  $S$  is a convex set, then (b) implies (c); so all three statements (a), (b), and (c) are equivalent.  $\square$*

To apply Proposition A.1 to condition (C4), we rewrite the various index sets in  $\widehat{S}_{\text{ps}}(x)$  more generally. To be more specifically, for any two triplets of index sets  $\mathcal{K} \triangleq \{\mathcal{K}_{>}, \mathcal{K}_{=}, \mathcal{K}_{<}\}$  and  $\mathcal{L} \triangleq \{\mathcal{L}_{>}, \mathcal{L}_{=}, \mathcal{L}_{<}\}$  with the former partitioning  $\{1, \dots, K\}$  and the latter partitioning  $\{1, \dots, L\}$ , define the closed set

$$\mathcal{S}(\mathcal{K}, \mathcal{L}) \triangleq \left\{ x \in X \left| \begin{array}{l} g_k(x) \leq 0 \quad \forall k \in \mathcal{K}_{<} \cup \mathcal{K}_{=} \\ g_k(x) \geq 0 \quad \forall k \in \mathcal{K}_{>} \\ h_\ell(x) \leq 0 \quad \forall \ell \in \mathcal{L}_{<} \cup \mathcal{L}_{=} \\ h_\ell(x) \geq 0 \quad \forall \ell \in \mathcal{L}_{>} \end{array} \right. \right\}.$$

and consider the optimization problem:

$$\underset{x \in \mathcal{S}(\mathcal{K}, \mathcal{L})}{\text{minimize}} \quad f_{\mathcal{K}, \mathcal{L}}(x) \triangleq \max \left( \sum_{\ell \in \mathcal{L}_{>}} \phi_\ell(x) - b, 0 \right).$$

Assume that each function  $\phi_\ell$  is bounded below on  $X$ . Then a very loose sufficient condition for (C4) to hold is that for all pairs  $(\mathcal{K}, \mathcal{L})$ , the minimum value of  $f_{\mathcal{K}, \mathcal{L}}$  on  $\mathcal{S}(\mathcal{K}, \mathcal{L})$  is zero and the pair  $(f_{\mathcal{K}, \mathcal{L}}, \mathcal{S}(\mathcal{K}, \mathcal{L}))$  satisfies condition (c) of Proposition A.1 with  $\delta = 1$ . We believe that it may be possible to tighten this sufficient condition and derive a result similar to the proposition for condition (C4); such details are beyond the scope of this work and best left for a separate investigation.

## References

- [1] W. Achziger, C. Kanzow: *Mathematical programs with vanishing constraints: Optimality conditions and constraint qualifications*, Math. Programming 114/1 (2008) 69–99.
- [2] M. Ahn, J. S. Pang, J. Xin: *Difference-of-convex learning: Directional stationarity, optimality, and sparsity*, SIAM J. Optimization 27/3 (2017) 1637–1665.
- [3] V. D. Batukhtin: *On solving discontinuous extremal problems*, J. Optimization Theory Appl. 77/3 (1993) 575–589.
- [4] H. H. Bauschke, P. L. Combettes: *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, New York (2011).
- [5] W. Bian, X. Chen: *A smoothing proximal gradient algorithm for nonsmooth convex regression with cardinality penalty*, SIAM J. Numer. Analysis 58 (2020) 858–883.
- [6] P. Bonami, A. Lodi, A. Tramontani, S. Wiese: *On mathematical programming with indicator constraints*, Math. Programming, Ser. B 151 (2015) 191–223.
- [7] X. Chen: *Smoothing methods for nonsmooth, nonconvex minimization*, Math. Programming, Ser. B 134 (2012) 71–99.
- [8] F. H. Clarke: *Optimization and Nonsmooth Analysis*, Classics in Applied Mathematics 5, SIAM, Philadelphia (1990).
- [9] P. L. Combettes: *Perspective functions: Properties, constructions, and examples*, Set-Valued Var. Analysis 26 (2018) 247–264.
- [10] Y. Cui: *Individualized decision making under partial identification: Three perspectives, two optimality results, and one paradox*, Harvard Data Sci. Review 3.3 (2021).
- [11] Y. Cui, J. Liu, J. S. Pang: *Nonconvex and nonsmooth approaches for affine chance constrained stochastic programs*, J. Set-Valued Var. Analysis 30 (2022) 1149–1211.
- [12] Y. Cui, J. S. Pang: *Modern Nonconvex and Nondifferentiable Optimization*, SIAM Publications, Philadelphia (2021).
- [13] H. Dong, K. Chen, J. Linderoth: *Regularization vs. relaxation: A conic optimization perspective of statistical variable selection*, arXiv: 1510.06083 (2015).
- [14] J. P. Dussault, M. Haddoub, T. Migot: *Mathematical programs with vanishing constraints: Constraint qualifications, their applications, and a new regularization method*, Optimization 68/2-3 (2019) 509–538.
- [15] I. Ekeland: *On the variational principle*, J. Math. Analysis Appl. 47 (1974) 324–354.
- [16] Y. Ermoliev, V. Norkin: *On constrained discontinuous optimization*, in: *Stochastic Programming Methods and Technical Applications*, Springer, Berlin (1998), 128–144.
- [17] Y. Ermoliev, V. Norkin: *Stochastic optimization of risk functions via parametric smoothing*, in: *Dynamic Stochastic Optimization*, Springer, Berlin (2004) 225–247.
- [18] Y. M. Ermoliev, V. I. Norkin, R. J.-B. Wets: *The minimization of semicontinuous functions: Mollifier subgradients*, SIAM J. Control Optimization 3 (1995) 149–167.

- [19] F. Facchinei, J. S. Pang: *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Springer Ser. Oper. Research, Springer, New York (2003).
- [20] J. Fan, R. Li: *Variable selection via nonconcave penalized likelihood and its oracle properties*, J. Amer. Statist. Assoc. 96/456 (2001) 1348–1360.
- [21] J. Fan, L. Xue, H. Zou: *Strong oracle optimality of folded concave penalized estimation*, Ann. Statistics 42/3 (2014) 819–849.
- [22] M. Feng, J. E. Mitchell, X. Shen, A. Wächter: *Complementarity formulations of  $\ell_0$ -norm optimization problems*, Pacific J. Optimization 14/2 (2018) 273–305.
- [23] M. C. Ferris: *Weak Sharp Minima and Exact Penalty Functions*, Ph. D. Thesis, Department of Computer Sciences, University of Wisconsin, Madison (1988).
- [24] M. Fukushima, J. S. Pang: *Complementarity constraint qualifications and simplified B-stationarity conditions for mathematical programs with equilibrium constraints*, Comp. Optimization Appl. 13 (1998) 111–136.
- [25] A. Gomez, Z. He, J. S. Pang: *Linear-step solvability of some folded concave and singly-parametric sparse optimization problems*, Math. Program. B 198/2 (2023) 1339–1380.
- [26] O. Günlük, J. Linderoth: *Perspective reformulations of mixed integer nonlinear programs with indicator variables*, Math. Programming, Ser. B 124 (2010) 183–205.
- [27] A. M. Gupal, V. I. Norkin: *Algorithm for the minimization of discontinuous functions*, Cybernetics 13 (1977) 220–223.
- [28] T. Hastie, R. Tibshirani, M. Wainwright: *Statistical Learning with Sparsity: The Lasso and Generalizations*, Monographs on Statistics and Applied Probability 143, CRC Press, Boca Raton (2015).
- [29] O. Heaviside: *Electromagnetic waves, the propagation of potential, and the electromagnetic effects of a moving charge*, The Electrician (1888).
- [30] H. Hijazi, P. Bonami, G. Cornéjols, A. Ouorou: *Mixed-integer nonlinear programs featuring “on/off” constraints*, Comp. Optimization Appl. 52 (2012) 537–558.
- [31] J. B. Hiriart-Urruty, C. Lemaréchal: *Convex Analysis and Minimization Algorithms*, Springer, New York (1993).
- [32] T. Hoheisel, C. Kanzow, A. Schwartz: *Mathematical programs with vanishing constraints: A new regularization approach with strong convergence properties*, Optimization 61/6 (2012) 619–636.
- [33] A. F. Izmailov, M. V. Solodov: *Mathematical programs with vanishing constraints: Optimality conditions, sensitivity, and a relaxation method*, J. Optimization Theory Appl. 142/3 (2009) 501–532.
- [34] H. Th. Jongen, O. Stein: *Smoothing by mollifiers. I: Semi-infinite optimization*, J. Global Optimization 41 (2008) 319–334.
- [35] H. Th. Jongen, O. Stein: *Smoothing by mollifiers. II: Nonlinear optimization*, J. Global Optimization 41 (2008) 335–350.
- [36] A. I. Kibzun, A. V. Naumov, V. I. Norkin: *On reducing a quantile optimization problem with discrete distribution to a mixed integer programming problem*, Automation Remote Control 74/6 (2013) 951–967.
- [37] H. A. Le Thi, D. T. Pham: *The DC programming and DCA revised with DC models of real world nonconvex optimization problems*, Ann. Oper. Research 133 (2015) 25–46.
- [38] H. A. Le Thi, T. Pham Dinh, X. T. Vo: *DC approximation approaches for sparse optimization*, Eur. J. Oper. Research 244/1 (2015) 26–46.
- [39] G. Libri: *Note sur les valeurs de la fonction  $0^{0^x}$* , J. Reine Angew. Math. 6 (1830) 67–72.

- [40] G. Libri: *Mémoire sur les fonctions discontinues*, J. Reine Angew. Math. 10 (1833) 303–316.
- [41] Z. Long, P. Yin, J. Xin: *Learning quantized neural nets by coarse gradient method for nonlinear classification*, Res. Math. Sciences 8/48 (2021), 19 pages.
- [42] Z. Lu, Z. Zhou, Z. Sun: *Enhanced proximal DC algorithms with extrapolation for a class of structured nonsmooth DC minimization*, Math. Program. 176/12 (2019) 369–401.
- [43] B. Pagnoncelli, S. Ahmed, A. Shapiro: *Sample average approximation method for chance-constrained programming: Theory and applications*, J. Optimization Theory Appl. 142 (2009) 399–416.
- [44] J. S. Pang, M. Razaviyayn, A. Alvarado: *Computing B-stationary points of non-smooth DC programs*, Math. Oper. Res. 42/1 (2017) 95–118.
- [45] T. Pham Dinh, H. A. Le Thi: *Convex analysis approach to DC programming: Theory, algorithm and applications*, Acta Math. Vietnamica 22/1 (1997) 289–355.
- [46] G. D. Pillo, F. Facchinei: *Exact penalty functions for nondifferentiable programming problems*, in: *Nonsmooth Optimization and Related Topics*, Springer, Boston (1989) 89–107.
- [47] Z. Qi, Y. Cui, Y. Liu, J. S. Pang: *Estimation of individualized decision rules based on an optimized covariate-dependent equivalent of random outcomes*, SIAM J. Optimization 29/3 (2019) 2337–2362.
- [48] R. T. Rockafellar: *Convex Analysis*, Princeton University Press, Princeton (1970).
- [49] R. T. Rockafellar, R. J.-B. Wets: *Variational Analysis*, 3rd printing, Grundlehren der mathematischen Wissenschaften 317, Springer, Berlin (2009).
- [50] J. Royset: *Consistent approximations in composite optimization*, Math. Programming A 201/1-2 (2023) 339–372.
- [51] H. D. Sherali, L. D. Brizendine, T. S. Glickman, S. Subramanian: *Low probability high consequence considerations in routing hazardous material shipments*, Transportation Science 31/3 (1997) 237–251.
- [52] W. Takahashi: *Existence theorems generalizing fixed point theorems for multivalued mappings*, in: *Fixed Point Theory and Applications*, M. A. Théra and J. B. Baillon (eds.), Pitman Research Notes in Mathematics 252, Longman Scientific & Technical, Harlow (1991) 397–406.
- [53] Y. Wu, Y. Liu: *Robust truncated hinge loss support vector machines*, J. Amer. Statist. Assoc. 102 (2007) 974–983.
- [54] J. Xin: *Training quantized deep neural networks and applications with blended coarse gradient descent*, SIAM News 52/4 (2019).
- [55] P. Yin, S. Zhang, Y. Qi, J. Xin: *Quantization and training of low bit-width convolutionary neural networks for object detection* J. Comput. Math. 37/3 (2019) 349–359.
- [56] G. Yu, H. Fu, Y. Liu: *High-dimensional cost-constrained regression via non-convex optimization*, Technometrics 64/1 (2022) 52–64.
- [57] L. H. Yue: *Cost-efficient variable selection using branching lars*, Electronic Thesis and Dissertation Repository 43 (2010), <https://ir.lib.uwo.ca/etd/43/>.
- [58] C. Zhang: *Nearly unbiased variable selection under minimax concave penalty*, Ann. Statistics 38/2 (2010) 894–942.
- [59] V. W. Zheng, K. C. C. Chang: *Regularizing structured classifier with conditional probabilistic constraints for semi-supervised learning*, in: CIKM 2016, Proc. 2016 ACM Conf. Information and Knowledge Management (2016) 1029–1038.