Metric Subregularity of Subdifferential and KL Property of Exponent 1/2

Shaohua Pan∗, Dongdong Zhang† and Yulan Liu‡

December 11, 2018

Abstract

For a proper lower semicontinuous function, we study the relations between the metric subregularity of its limiting subdifferential relative to the critical set and the KL property of exponent 1/2. When the function is convex, we establish the equivalence between them under a continuity condition. When the function is nonconvex, we show that the KL property of exponent 1/2 along with the quadratic growth on the critical set implies the metric subregularity of the subdifferential relative to the critical set; and if the function is primal-lower-nice, under an assumption on stationary values, the latter implies the former. These results provide a bridge for the two kinds of regularity and contribute to enriching each other.

Keywords: Metric subregularity; KL property of exponent 1/2; subdifferential

1 Introduction

For a proper lower semicontinuous (lsc) function, the metric subregularity of its (limiting) subdifferential at a (limiting) critical point for the origin states that for any point near the critical point, its distance to the critical set is upper bounded by the remoteness (i.e., distance to zero) of the subdifferential at this point. When the function is convex, Artacho and Geoffroy [1] provided an excellent characterization for this property in terms of the quadratic growth condition of the function, and this property plays a key role in the linear convergence analysis of algorithms; for example, the proximal point algorithm [2]. However, when the function is nonconvex, there is less knowledge on the metric subregularity of the subdifferential at critical points except that the metric subregularity of the subdifferential at a local minimizer implies the quadratic growth (see [3, 4]), and it is even unclear how to use the property to achieve the linear convergence of algorithms.

∗School of Mathematics, South China University of Technology, Guangzhou. shhpan@scut.edu.cn
†School of Mathematics, South China University of Technology. mathzdd@mail.scut.edu.cn
‡Corresponding author(ylliu@gdut.edu.cn), School of Applied Mathematics, Guangdong University of Technology, Guangzhou.
Recently, the Kurdyka-Łojasiewicz (KL) property of a proper lsc function is successfully used to analyze the global convergence of algorithms for nonconvex and nonsmooth optimization problems (see [5, 6, 7]). In particular, the KL property of exponent 1/2 is the crucial part in achieving the linear convergence of the corresponding algorithms. Then, it is natural to ask whether there is a link or not between the metric subregularity of the subdifferential at a critical point and the KL property of exponent 1/2 of the function at the reference point. This work aims at investigating this so as to build a bridge for the two kinds of regularity. For the convex function, we establish their equivalence. For the nonconvex function, we show that the KL property of exponent 1/2 along with the quadratic growth on the critical set implies the metric subregularity of the subdifferential relative to the critical set; and if the function is primal-lower-nice (pln), under an assumption on stationary values, the latter also implies the former.

For a class of structured nonconvex functions, Li and Pong [8] showed that under the assumption on stationary values, the Luo-Tseng error bound implies the KL property of exponent 1/2. Our result (see Proposition 3.4) shows that under this assumption the metric subregularity of the subdifferential relative to the critical set also implies the KL property of exponent 1/2. Then, it is interesting to know what is the link between the metric subregularity of the subdifferential relative to the critical set and the Luo-Tseng error bound. In Section 3.3, we provide a full exploration and extend the result of [9, Section 3] partially to the nonconvex case.

2 Notations and preliminaries

We denote \( X \) by a finite-dimensional vector space equipped with the inner product \( \langle \cdot, \cdot \rangle \) and its induced norm \( \| \cdot \| \). For an extended real-valued function \( f : X \to (-\infty, +\infty) \), we denote \( \text{dom} \ f = \{ x \in X : f(x) < \infty \} \) by its domain, and for given real numbers \( \alpha \) and \( \beta \), set \( [\alpha \leq f \leq \beta] := \{ x \in X : \alpha \leq f(x) \leq \beta \} \), and \( f \) is called proper if \( \text{dom} f \neq \emptyset \). For a proper \( f : X \to (-\infty, +\infty) \), we use \( x' \to x \) to signify \( x' \to x \) and \( f(x') \to f(x) \).

a given \( \mathbf{r} \in X \) and a constant \( \varepsilon > 0 \), \( B(\mathbf{r}, \varepsilon) \) denotes the closed ball centered at \( \mathbf{r} \) with radius \( \varepsilon \). For a proper convex function \( \phi : X \to (-\infty, +\infty) \), denote by \( \text{prox}_\phi \) its proximal operator. For a proper convex function \( \phi : X \to (-\infty, +\infty) \), denote by \( \text{prox}_\phi \) its proximal operator, and for a closed set \( S \subseteq X \), \( \Pi_S \) denotes the projection operator on \( S \), which is a multi-valued mapping when the set \( S \) is not convex.

2.1 Limiting subdifferential and metric subregularity

**Definition 2.1 ([10, Definition 8.3])** Consider a function \( f : X \to (-\infty, +\infty) \) and a point \( x \in \text{dom} \ f \). The regular subdifferential of \( f \) at \( x \) is defined by

\[
\hat{\partial}f(x) := \left\{ v \in X : \liminf_{x' \to x, \neq x} \frac{f(x') - f(x) - \langle v, x' - x \rangle}{\|x' - x\|} \geq 0 \right\},
\]

2
and the (limiting) subdifferential of $f$ at $x$ is defined by
\[ \partial f(x) := \left\{ v \in \mathbb{X} : \exists x^k \to x, v^k \to v \text{ with } v^k \in \hat{\partial} f(x^k) \text{ for each } k \right\}. \]

**Remark 2.1** (a) At each $x \in \text{dom} f$, it holds that $\hat{\partial} f(x) \subseteq \partial f(x)$, where the former is closed convex, and the latter is closed but generally not convex. When $f$ is convex, both $\hat{\partial} f(x)$ and $\partial f(x)$ reduce to the subdifferential of $f$ at $x$ in the sense of convex analysis.

(b) The point $\overline{x}$ at which $0 \in \partial f(\overline{x})$ is called a (limiting) critical point of $f$. We denote by $\text{crit} f$ the set of critical points (or the critical set). By [10, Theorem 10.1], the local minimizer of $f$ is necessarily the critical point of $f$.

**Definition 2.2** Let $F : \mathbb{X} \rightrightarrows \mathbb{X}$ be a multifunction. Consider an arbitrary $(\overline{x}, \overline{y}) \in \text{gph} F$. We say that $F$ is metrically subregular at $\overline{x}$ for $\overline{y}$ if there exist $\varepsilon > 0$ and $\kappa > 0$ such that for all $x \in B(\overline{x}, \varepsilon)$, $\text{dist}(x, F^{-1}(\overline{y})) \leq \kappa \text{dist}(\overline{y}, F(x))$.

Definition 2.2 has a little difference from the original one, and we here adopt an equivalent form by [11, Section 3H]. It is well-known that the metric subregularity of $F$ at $\overline{x}$ for $\overline{y} \in F(\overline{x})$ if and only if its inverse mapping $F^{-1}$ is calm at $\overline{x}$ for $\overline{y} \in F^{-1}(\overline{y})$. For the recent discussions on the criterion of calmness and metric subregularity, the reader may refer to [12, 13]. In this work, we focus on the metric subregularity of the subdifferential of a proper lsc function $f : \mathbb{X} \to (-\infty, +\infty]$, and say that $\partial f$ is metrically subregular relative to $\text{crit} f$ if it is metrically subregular at each $x \in \text{crit} f$ for the origin. Generally, it is not an easy task to check whether the subdifferential of a proper (even convex) function is metrically subregular or not at a critical point. Appendix A summarizes some convex functions whose subdifferentials are metrically subregular at each point of their graphs.

### 2.2 Kurdyka-Łojasiewicz property

**Definition 2.3** Let $f : \mathbb{X} \to (-\infty, +\infty]$ be a proper function. The function $f$ is said to have the Kurdyka-Łojasiewicz (KL) property at $\overline{x} \in \text{dom} \partial f$ if there exist $\eta \in (0, +\infty]$, a continuous concave function $\varphi : [0, \eta] \to \mathbb{R}_+$ satisfying the following properties

(i) $\varphi(0) = 0$ and $\varphi$ is continuously differentiable on $(0, \eta)$;

(ii) for all $s \in (0, \eta)$, $\varphi'(s) > 0$,

and a neighborhood $\mathcal{U}$ of $\overline{x}$ such that for all $x \in \mathcal{U} \cap [f(\overline{x}) < f < f(\overline{x}) + \eta]$,

$$\varphi'(f(x) - f(\overline{x})) \text{dist}(0, \partial f(x)) \geq 1.$$  

If the corresponding $\varphi$ can be chosen as $\varphi(s) = cs^{1/2}$ for some $c > 0$, then $f$ is said to have the KL property at $\overline{x}$ with an exponent of $1/2$. If $f$ has the KL property of exponent $1/2$ at each point of $\text{dom} \partial f$, then $f$ is called a KL function of exponent $1/2$.

**Remark 2.2** To argue that a proper $f$ is a KL function of exponent $1/2$, it suffices to check whether it has the KL property of $1/2$ at all critical points or not, since by [5, Lemma 2.1] it has the property at any noncritical point.
2.3 Primal-lower-nice functions

Definition 2.4 Let $h : \mathbb{X} \to (-\infty, +\infty]$ be a proper function and consider $x \in \text{dom} h$. The $h$ is said to be pln at $x$, if there exist $R > 0, c' > 0$ and $\varepsilon' > 0$ such that for all $x \in B(\mathbb{X}, \varepsilon')$, for all $r > R$ and all $v \in \partial h(x)$ with $\|v\| \leq c'r$,

$$h(y) \geq h(x) + \langle v, y - x \rangle - \frac{r}{2}\|y - x\|^2$$

for all $y \in B(\mathbb{X}, \varepsilon')$.

If $h$ is pln at each $x \in \text{dom} h$, then we say that $h$ is a pln function.

Remark 2.3 In the original definition (see [14]), the proximal subdifferential instead of the limiting subdifferential is used. However, by [15, Remark 1.5], for a pln function $h$, $\partial \text{P} h(x) = \hat{\partial} h(x) = \partial h(x) = \partial \text{C} h(x)$ for all $x \in \text{dom} h$, where $\partial \text{P} h$ and $\partial \text{C} h$ are the proximal and Clarke subdifferentials of $h$. That is, the definition of the pln property is independent of the involved subdifferential.

The pln function includes strongly amenable functions and semiconvex functions. For strongly amenable functions, the reader is invited to see [10, Section 10.F]. Next, we recall from [16] the formal definition of semiconvexity.

Definition 2.5 A proper function $h : \mathbb{X} \to (-\infty, +\infty]$ is called semiconvex of modulus $\gamma \geq 0$ if for any $\gamma' \geq \gamma$, the function $x \mapsto h(x) + \frac{1}{2}\gamma'\|x\|^2$ is convex.

Remark 2.4 (a) Definition 2.5 is equivalent to saying that for each $\gamma' \geq \gamma$,

$$h(x') \geq h(x) + \langle \xi, x' - x \rangle - \gamma'\|x' - x\|^2 \quad \forall x', x \in \mathbb{X} \text{ and } \forall \xi \in \partial h(x)$$

(b) A continuously differentiable function $h : \mathbb{X} \to \mathbb{R}$ with Lipschitz continuous gradient is necessarily semiconvex. Every twice continuously differentiable function with bounded second derivative is also semiconvex.

(c) By [17, Theorem 6] a locally Lipschitzian function $h : \mathbb{X} \to \mathbb{R}$ is lower-C$^2$ if and only if $\partial \text{C} h$ is strictly hypomonotone, that is, for all $x \in \mathbb{X}$,

$$\liminf_{x' \to x, y' \in \partial \text{C} h(x')} \frac{\langle y' - y'', x' - x'' \rangle}{\|x' - x''\|^2} > -\infty.$$ 

It is not difficult to verify that the minimization of a locally Lipschitzian lower-C$^2$ function over a compact convex set $\Omega \subseteq \mathbb{X}$ is equivalent to the unconstrained minimization of a certain semiconvex function.

3 Metric subregularity and KL property of exponent $1/2$

For a proper lsc $f : \mathbb{X} \to (-\infty, +\infty]$, we investigate the link between the metric subregularity of $\partial f$ relative to crit$f$ and the KL property of exponent 1/2.
3.1 The case that $f$ is convex

In this part we shall establish the equivalence between the metric subregularity of $\partial f$ relative to $\text{crit } f$ and the KL property of exponent $1/2$ of $f$ under a continuity condition.

This requires the following lemma that summarizes some favorable properties of the trajectories of the differential inclusion for a proper lsc convex function.

**Lemma 3.1** Let $h: \mathbb{X} \to (-\infty, +\infty]$ be a proper lsc convex function with $\text{crit } h \neq \emptyset$. For each $x \in \text{dom } h$, there is a unique absolutely continuous curve $\chi_x: [0, +\infty) \to \mathbb{X}$ such that

\[
\begin{cases}
\chi_x(t) \in -\partial h(\chi_x(t)) & \text{a.e. on } (0, +\infty), \\
\chi_x(0) = x.
\end{cases}
\]

(1)

Also, the curve $\chi_x$ (called a subgradient curve) has the following properties:

(i) For all $t > 0$, the right derivative $\frac{d}{dt} \chi_x(t^+)$ of $\chi_x$ is well defined and satisfies

\[
\frac{d}{dt} \chi_x(t^+) = -\partial^0 h(\chi_x(t)) := -\arg \min_{z \in \partial h(\chi_x(t))} \|z\|;
\]

(ii) $\frac{d}{dt} h(\chi_x(t^+)) = -\|\chi_x(t^+)\|^2$ for all $t > 0$;

(iii) for each $z \in \text{crit } h$, the function $t \mapsto \|\chi_x(t) - z\|$ decreases;

(iv) the function $t \mapsto h(\chi_x(t))$ is nonincreasing and $\lim_{t \to \infty} h(\chi_x(t)) = h^*$;

(v) $\chi_x(t)$ converges to some $\hat{x} \in \text{crit } h$ as $t \to \infty$.

Lemma 3.1 can be found in [18], and its proof is provided in [19] except for part (v), which was proved in [20]. Comparing with [16, Theorem 13], we see that the subgradient curve of a semiconvex function generally does not satisfy (iii) and (v).

**Theorem 3.1** Let $f: \mathbb{X} \to (-\infty, +\infty]$ be a proper lsc convex function. Consider the following statements:

(i) the multifunction $\partial f$ is metrically subregular relative to $\text{crit } f$;

(ii) for each $\overline{x} \in \text{crit } f$, there exist $\overline{\epsilon} > 0$ and $\nu > 0$ such that for all $x \in B(\overline{x}, \overline{\epsilon})$,

\[
f(x) - f(\overline{x}) \geq \nu \text{dist}^2(x, (\partial f)^{-1}(0));
\]

(2)

(iii) $f$ is a KL function with an exponent of $1/2$.

Then (i) $\iff$ (ii) $\implies$ (iii). If, in addition, $f$ is continuous on $\text{crit } f$, (ii) $\iff$ (iii).
Proof: The equivalence between (i) and (ii) follows by [1, Theorem 3.3]. We next establish the implication (ii) ⇒ (iii). Fix an arbitrary \( \varphi \in \text{crit} \, f \). Suppose that there exist \( \hat{\varepsilon} > 0 \) and \( \nu > 0 \) such that (2) holds for all \( x \in B(\varphi, \hat{\varepsilon}) \). Pick up an arbitrary \( x \in B(\varphi, \hat{\varepsilon}) \). If \( \partial f(x) = \emptyset \), the conclusion holds automatically. So, it suffices to consider the case \( \partial f(x) \neq \emptyset \). From the convexity of \( f \), for any \( z \in \text{crit} \, f \), it holds that

\[
f(z) \geq f(x) + \langle \xi, z - x \rangle \quad \text{for every} \ \xi \in \partial f(x).
\]

Then,

\[
f(x) - f(z) \leq \inf_{\xi \in \partial f(x)} ||\xi|| \inf_{z \in \text{crit} \, f} ||z - x|| \leq \text{dist}(x, (\partial f)^{-1}(0)) \inf_{\xi \in \partial f(x)} ||\xi||.
\]

By setting \( z = \varphi \) in the last inequality and using inequality (2), we obtain

\[
[f(x) - f(\varphi)]^{1/2} \leq \frac{1}{\sqrt{\nu}} \inf_{\xi \in \partial f(x)} ||\xi|| = \frac{1}{\sqrt{\nu}} \text{dist}(0, \partial f(x)).
\]

This shows that \( f \) has the KL property of exponent 1/2 at \( \varphi \). By the arbitrariness of \( \varphi \) in \( \text{crit} \, f \), the conclusion in (iii) follows.

Assume that \( f \) is continuous on \( \text{crit} \, f \). We argue that (iii) ⇒ (ii) holds. Fix an arbitrary \( \varphi \in \text{crit} \, f \). Since \( f \) is a KL function of exponent 1/2 and \( \varphi \in \text{dom} \partial f \), there exist \( \varepsilon_1 > 0, \eta > 0 \) and \( c > 0 \) such that for all \( x \in B(\varphi, \varepsilon_1) \cap [f(\varphi) < f < f(\varphi) + \eta] \),

\[
\text{dist}(0, \partial f(x)) \geq c \sqrt{f(x) - f(\varphi)}.
\]

Since \( f \) is continuous at \( \varphi \), there exists \( \varepsilon_2 > 0 \) such that for all \( x \in B(\varphi, \varepsilon_2) \),

\[
f(\varphi) \leq f(x) < f(\varphi) + \eta.
\]

Set \( \varepsilon = \min(\varepsilon_1, \varepsilon_2) \). The last two inequalities imply that for all \( x \in B(\varphi, \varepsilon) \setminus \text{crit} \, f \),

\[
\text{dist}(0, \partial f(x)) \geq c \sqrt{f(x) - f(\varphi)}.
\]

Since \( f \) is a proper lsc and convex function, by Lemma 3.1 for every \( x \in \text{dom} \, f \) there is a unique absolutely continuous subgradient curve \( \chi_x : [0, +\infty) \to \mathbb{R} \) that satisfies the differential inclusion (1). Define \( \omega(t) := \sqrt{f(\chi_x(t)) - f^*} \) for \( t \in [0, +\infty) \). Fix an arbitrary \( x \in B(\varphi, \varepsilon) \setminus \text{crit} \, f \) and consider the differential inclusion (1). By Lemma 3.1(iii) it follows that \( \|\chi_x(t) - \bar{x}\| \leq \|x - \bar{x}\| \leq \varepsilon \) for any \( t > 0 \). Fix an arbitrary \( T > 0 \). Then,

\[
\frac{d\omega(t)}{dt} = \frac{\left\langle \partial f(\chi_x(t)), \chi_x(t) \right\rangle}{2 \sqrt{f(\chi_x(t)) - f(\bar{x})}} = -\frac{\|\dot{\chi}_x(t)\|^2}{2 \sqrt{f(\chi_x(t)) - f(\bar{x})}} \leq \frac{-c}{2} \|\dot{\chi}_x(t)\| \quad \forall t \in (0, T),
\]

where the second equality is due to Lemma 3.1(ii), and the inequality is using (3) and the fact that \( \dot{\chi}_x(t) \in -\partial f(\chi_x(t)) \). The last inequality implies that

\[
\omega(T) - \omega(0) = \lim_{\varepsilon \to 0^+} \int_{\varepsilon}^{T} \frac{d\omega(t)}{dt} dt \leq \frac{-c}{2} \int_{0}^{T} \|\dot{\chi}_x(t)\| dt \leq \frac{-c}{2} \text{dist}(\chi_x(T), \chi_x(0))
\]
where the last inequality is since the length of the curve connecting any two points is at least as long as the Euclidean distance between them. Thus, together with the nonnegativity of $\omega(T)$, it is immediate to obtain
\[ \sqrt{f(x) - f(\overline{x})} = \omega(0) \geq \frac{c}{2} \text{dist}(\chi x(T), \chi x(0)). \]
Notice that the function $\chi x(\cdot)$ is continuous by [16, Theorem 14]. By taking the limiting $T \to \infty$ and using Lemma 3.1(v), it follows that
\[ \sqrt{f(x) - f(\overline{x})} = \omega(0) \geq \frac{c}{2} \text{dist}(x, \text{crit}f) = \frac{c}{2} \text{dist}(x, (\partial f)^{-1}(0)). \]
This, by the arbitrariness of $x \in B(\overline{x}, \epsilon) \setminus \text{crit}f$, implies that (ii) holds.}

**Remark 3.1** (a) Notice that $\partial f$ is maximally monotone. By [21, Theorem 3.1] or [22, Proposition 2.4], we know that the assertion (i) is also equivalent to the metric subregularity of the mapping $R_f(x) := x - \text{prox}_f(x)$ relative to $\text{crit}f$. Thus, one may add the metric subregularity of $R_f$ to the list of assertions.

(b) By Theorem 3.1, it is immediate to know that all the convex functions appearing in Lemma A.1-A.2 are the KL one of exponent 1/2.

### 3.2 The case that $f$ is nonconvex

We first show that under the quadratic growth condition in (2), the KL property of exponent 1/2 implies the metric subregularity of $\partial f$ relative to $\text{crit}f$.

**Proposition 3.1** Let $f : X \to (-\infty, +\infty)$ be a proper function. Suppose $f$ is continuous on $\text{crit}f$ and the quadratic growth (2) holds at each $\overline{x} \in \text{crit}f$. If $f$ is a KL function of exponent 1/2, $\partial f$ is metrically subregular relative to $\text{crit}f$.

**Proof:** Fix an arbitrary $\overline{x} \in \text{crit}f$. To prove that $\partial f$ is metrically subregular at $\overline{x}$ for the origin, it suffices to argue that there exist $\epsilon > 0$ and $\kappa > 0$ such that
\[ \text{dist}(x, (\partial f)^{-1}(0)) \leq \kappa \text{dist}(0, \partial f(x)) \quad \forall x \in B(\overline{x}, \epsilon). \quad (4) \]
Since $f$ has the KL property of exponent 1/2 at $\overline{x}$, there exist $\eta \in (0, +\infty)$, $\epsilon_1 > 0$ and $c > 0$ such that for all $x \in B(\overline{x}, \epsilon_1) \cap \{f(\overline{x}) < f(x) < f(\overline{x}) + \eta\}$,
\[ \text{dist}(0, \partial f(x)) \geq \frac{2}{c} \sqrt{f(x) - f(\overline{x})}. \quad (5) \]
Since $f$ is continuous at $\overline{x}$, there exists $\epsilon_2 > 0$ such that for all $x \in B(\overline{x}, \epsilon_2)$,
\[ f(x) \leq f(\overline{x}) + \eta/2. \quad (6) \]
Since the quadratic growth holds at $\overline{x}$, there exist $\hat{\epsilon} > 0$ and $\nu > 0$ such that
\[ f(x) \geq f(\overline{x}) + \nu \text{dist}^2(x, (\partial f)^{-1}(0)) \quad \forall x \in B(\overline{x}, \hat{\epsilon}). \quad (7) \]
Fix an arbitrary $x \in \mathbb{B}(\overline{x}, \varepsilon)$ with $\varepsilon = \min(\varepsilon_1, \varepsilon_2, \hat{\varepsilon})$. By (5) and (7), we have
\[
\text{dist}(0, \partial f(x)) \geq \frac{2\sqrt{n}}{c}\text{dist}(x, (\partial f)^{-1}(0))
\]
provided that $x \in [f(\overline{x}) < f < f(\overline{x}) + \eta]$. If $x \notin [f(\overline{x}) < f < f(\overline{x}) + \eta]$, by (6) we have $f(x) \leq f(\overline{x})$, and then dist$^2(x, (\partial f)^{-1}(0)) = 0$ follows from (7). Then, the last inequality still holds. Thus, the desired (4) holds with $\kappa = \frac{2\sqrt{n}}{c}$. \hfill \Box

**Remark 3.2 (a)** By [3, Remark 2.2(iii)], when $\overline{x}$ is a local minimum, the metric subregularity of $\partial f$ at $\overline{x}$ for the origin implies the quadratic growth at $\overline{x}$. This shows that the quadratic growth is necessary for the conclusion.

(b) The continuity of $f$ on $\text{crit} f$ is also necessary for the conclusion. Consider
\[
f(x) := \begin{cases} 
x^4 & \text{if } x \neq 0; \\
-1 & \text{if } x = 0
\end{cases}
\quad \text{for } x \in \mathbb{R}.
\]
An elementary calculation yields $\partial f(0) = \mathbb{R}$ and $\partial f(x) = \{4x^3\}$ if $x \neq 0$. Clearly, $f$ is not continuous on $\text{crit} f = \{0\}$. It is easy to check that $\partial f$ is not metrically subregular at $\overline{x} = 0$ although the quadratic growth condition (2) holds at $\overline{x}$ and $f$ is a KL function with an exponent of $1/2$.

Now it is unclear whether the converse conclusion of Proposition 3.1 holds or not for a general nonconvex $f$. For the pln function, we may achieve it under the following assumption concerning separation of stationary values:

**Assumption 3.1** For each $\overline{x} \in \text{crit} f$, there exists $\delta > 0$ such that $f(y) \leq f(\overline{x})$ for all $y \in \mathbb{B}(\overline{x}, \delta) \cap \text{crit} f$.

**Proposition 3.2** Let $f : \mathbb{X} \to (-\infty, +\infty]$ be a proper lsc pln function. Suppose that $f$ is locally Lipschitz continuous on $\text{crit} f$ and Assumption 3.1 holds. If $\partial f$ is metrically subregular relative to $\text{crit} f$, $f$ is a KL function of exponent $1/2$.

**Proof:** Fix an arbitrary $\overline{x} \in \text{crit} f$. Since $\partial f$ is metrically subregular at $\overline{x}$ for the origin, there exist $\varepsilon > 0$ and $\kappa > 0$ such that (4) holds. Since $f$ is pln at $\overline{x}$, by Definition 2.4 there exist $R > 0$, $\varepsilon' > 0$ and $\kappa' > 0$ such that for all $x \in \mathbb{B}(\overline{x}, \varepsilon')$, all $r > R$ and all $v \in \partial f(x)$ with $||v|| \leq \varepsilon'r$, one has
\[
f(y) \geq f(x) + \langle v, y - x \rangle - (r/2)||y - x||^2 \quad \text{for all } y \in \mathbb{B}(\overline{x}, \varepsilon'). \tag{8}
\]
Since $f$ is locally Lipschitz near $\overline{x}$, there exist $\delta' > 0$ and $L > 0$ such that for all $z, z' \in \mathbb{B}(\overline{x}, \delta')$, $|f(z) - f(z')| \leq L||z - z'||$. Take $\epsilon = \min(\delta, \varepsilon, \varepsilon', \delta')/4$ where $\delta$ is the constant in Assumption 3.1. Choose an arbitrary $\eta > 0$. Consider an arbitrary $x \in \mathbb{B}(\overline{x}, \epsilon) \cap [f(\overline{x}) < f < f(\overline{x}) + \eta]$. Since $\mathbb{B}(x, \delta'/2) \subseteq \mathbb{B}(\overline{x}, \delta')$, the function $f$ is locally
Lipschitz near \( x \). By [23, Proposition 2.1.2], we have \( \|\xi\| \leq L \) for all \( \xi \in \partial f(x) \subseteq \partial cf(x) \). Let \( u \in \Pi_{\text{crit}\ f}(x) \). Then,

\[
\|u - \bar{f}\| = \|u - x + x - \bar{f}\| \leq \|u - x\| + \|x - \bar{f}\| \leq 2\|x - \bar{f}\|.
\]

Take an arbitrary \( \xi \in \partial f(x) \). From the last inequality and (8), it follows that

\[
f(u) \geq f(x) + \langle \xi, y - x \rangle -(r/2)\|u-x\|^2 \quad \text{with} \quad r = \max(R+1, L/c).
\]

Notice that \( f(u) \leq f(\bar{f}) \) by Assumption 3.1. Then, it holds that

\[
f(x) - f(\bar{f}) \leq f(x) - f(u) \leq \inf_{\xi \in \partial f(u)} \|\xi\|\|x-u\| + (r/2)\|x-u\|^2
\]

\[
\leq \inf_{\xi \in \partial f(x)} \|\xi\|\|x-u\| + (r/2)\|x-u\|^2
\]

\[
\leq \text{dist}(0, \partial f(x))\text{dist}(x, (\partial f)^{-1}(0)) + (r/2)\text{dist}^2(x, (\partial f)^{-1}(0)).
\]

Since \( \text{dist}(x, (\partial f)^{-1}(0)) \leq \kappa\text{dist}(0, \partial f(x)) \) by (4), the last inequality implies

\[
f(x) - f(\bar{f}) \leq (\kappa + r\kappa^2/2)\text{dist}^2(0, \partial f(x)).
\]

Along with \( f(x) > f(\bar{f}) \), this shows that \( f \) has the KL property of exponent 1/2 at \( \bar{f} \). By the arbitrariness of \( \bar{f} \in \text{crit}\ f \), the desired result follows.

\[\square\]

**Remark 3.3** Assumption 3.1 and the local Lipschitz continuity of \( f \) on \( \text{crit}\ f \) are not necessary to the conclusion. For example, for \( f(x) := \text{sign}(|x|) \) for \( x \in \mathbb{R} \), one may check that \( f \) is pln with \( \text{crit}\ f = \mathbb{R} \), \( f \) is a KL function of exponent \( \frac{1}{2} \) and \( \partial f \) is metrically subregular relative to \( \text{crit}\ f \), but \( f \) is not locally Lipschitz continuous on \( \text{crit}\ f \) and Assumption 3.1 does not hold at \( \bar{f} = 0 \).

When the function \( f \) is semiconvex, Proposition 3.2 can be strengthened as follows.

**Proposition 3.3** Let \( f : \mathbb{X} \to (-\infty, +\infty] \) be a proper lsc semiconvex function of modulus \( \gamma \) and let \( f_\gamma(\cdot) := f(\cdot) + \frac{\gamma}{2}\|\cdot\|^2 \). Then, \( f \) is a KL function of exponent 1/2 under Assumption 3.1 and either of the following equivalent conditions:

(i) the mapping \( \partial f \) is metrically subregular relative to \( \text{crit}\ f \);

(ii) \( \mathcal{R}_{f_\gamma}(x) := x - \text{prox}_{f_\gamma}(x + \gamma x) \) is metrically subregular relative to \( \text{crit}\ f \).

**Proof:** Notice that \( \partial f_\gamma \) is maximally monotone since \( f_\gamma \) is closed and proper convex. By the definition of \( f_\gamma \) and [10, Exercise 8.8(c)], for all \( x \in \text{dom}\partial f, \partial f(x) = \partial f_\gamma(x) - \gamma x \). The equivalence between (i) and (ii) then follows from [21, Theorem 3.1] or [22, Proposition 2.4]. Since \( f \) is a proper lsc semiconvex function of modulus \( \gamma \), by Remark 2.4 for any \( z, y \in \mathbb{X} \) and all \( \xi \in \partial f(y) \),

\[
f(z) - f(y) \geq \langle \xi, z - y \rangle - \frac{\gamma}{2}\|z - y\|^2.
\]

Now suppose that the condition (i) holds. Following the same arguments as those after (9) for Proposition 3.2, one may get the desired result. \[\square\]
Remark 3.4 An example satisfying the assumption of Proposition 3.3 is the function
\[ f(x) = \|Ax - b\|_1 + \frac{1}{2\mu}\|x\|^2 + \sum_{i=1}^P \phi_\lambda(x_i), \tag{10} \]
where \( A \in \mathbb{R}^{n \times P} \) and \( b \in \mathbb{R}^n \) are the observation matrix and vector, respectively, \( \mu > 0 \) is the regularization parameter, and \( \phi_\lambda (\lambda > 0) \) is the SCAD function from [25]
\[ \phi_\lambda(t) = \begin{cases} \lambda |t| & \text{if } |t| \leq \lambda, \\ \frac{-t^2 + 2\lambda|t| - \lambda^2}{2(a - 1)} & \text{if } \lambda < |t| \leq a\lambda, \quad \text{with } a > 2 \\ \frac{1}{2}(a + 1)\lambda^2 & \text{if } |t| > a\lambda \end{cases} \tag{11} \]
or the MCP function from [26] which takes the following form for \( a > 0 \)
\[ \phi_\lambda(t) = \begin{cases} \lambda |t| - t^2/(2a) & \text{if } |t| \leq a\lambda, \\ \frac{1}{2a}\lambda^2 & \text{if } |t| > a\lambda. \end{cases} \tag{12} \]
One may check that \( f \) is semiconvex of modulus \( \gamma = \frac{1}{a - 1} - \frac{b}{b} \) for \( \phi_\lambda \) in (11) and \( \gamma = \frac{1}{a} - \frac{b}{b} \) for \( \phi_\lambda \) in (12). By invoking [10, Theorem 10.6], it is not hard to check that at each \( x \in \mathbb{R}^p \), \( \partial f(x) = A^T \partial\|\cdot\|_1(Ax - b) + \mu^{-1}x + \partial \phi_\lambda(x_1) \times \cdots \times \partial \phi_\lambda(x_p) \). This implies that \( \partial f \) is polyhedral, i.e., \( \text{gph} \partial f \) is the union of finite many polyhedral convex sets, and hence is metrically subregular at each \( \tau \in \text{crit } f \) for 0 by [24, Proposition 1]. By Proposition 3.3, \( f \) has the KL property of exponent 1/2 at those critical points satisfying Assumption 3.1. Similarly, when replacing \( \|Ax - b\|_1 \) with \( \|Ax - b\|_\infty \), \( f \) still has the KL property of exponent 1/2 at those critical points satisfying Assumption 3.1. It is worthwhile to point out that the KL property of exponent 1/2 can be derived by [8, Corollary 5.2].

3.3 The case that \( f \) is structured nonconvex
In this part we focus on the structured nonconvex \( f \) which takes the form of
\[ f(x) = g(x) + h(x), \tag{13} \]
where \( g: \mathbb{X} \to (-\infty, +\infty] \) is a proper lsc function with an open domain and is continuously differentiable on \( \text{dom } g \), and \( h: \mathbb{X} \to (-\infty, +\infty] \) is a proper lsc convex function with \( \text{dom } h \cap \text{dom } g \neq \emptyset \). Clearly, the composite \( \tilde{g}(\cdot) + \tilde{h}(\cdot) \), where \( \tilde{g} \) has the same property as \( g \) does and \( \tilde{h} : \mathbb{X} \to (-\infty, +\infty] \) is a proper lsc semiconvex function, can be rewritten as the form of (13). It is easy to check that \( f \) is pln but is not semiconvex unless the gradient of \( g \) is Lipschitz continuous. Now by using its special structure of \( f \) we can obtain the conclusion of Proposition 3.3 without the local Lipschitz continuity of \( f \) on \( \text{crit } f \).

Proposition 3.4 For the nonconvex \( f \) in (13), under Assumption 3.1, if \( \partial f \) is metrically subregular relative to \( \text{crit } f \), \( f \) is a KL function of exponent 1/2.
Proof: Fix an arbitrary $\overline{x} \in \text{crit } f$. Since $\partial f$ is metrically subregular at $\overline{x}$ for the origin, there exist $\varepsilon > 0$ and $\kappa > 0$ such that (4) holds. Since $\overline{x} \in \text{dom } g$ and $g$ is smooth on dom$g$, $\nabla g$ is Lipschitz continuous near $\overline{x}$, i.e., there exist $\delta' > 0$ and $L_g > 0$ such that $B(\overline{x}, \delta') \subset \text{dom } g$ and for any $y, z \in B(\overline{x}, \delta')$, $\|\nabla g(y) - \nabla g(z)\| \leq L_g \|y - z\|$. By the descent lemma, it follows that

$$g(y) \leq g(z) + \langle \nabla g(z), y - z \rangle + \frac{L_g}{2}\|y - z\|^2 \quad \forall y, z \in B(\overline{x}, \delta').$$

In addition, since $h$ is a proper lsc convex function, it holds that

$$h(z) - h(y) \geq \langle \zeta, z - y \rangle \quad \forall z, y \in \mathbb{X} \text{ and all } \zeta \in \partial h(y).$$

Take $\epsilon \in (0, \frac{1}{2}\min(\varepsilon, \delta', \delta))$ such that $B(\overline{x}, \epsilon) \subseteq \text{dom } g$, where $\delta > 0$ is the constant in Assumption 3.1. Fix an arbitrary $x \in B(\overline{x}, \epsilon/2)$ and let $u \in \Pi_{\text{crit } f}(x)$. Notice that $\|u - \overline{x}\| = \|u - x + x - \overline{x}\| \leq 2\|x - \overline{x}\|$. Take an arbitrary $\zeta \in \partial h(x)$. Then, together with the last two inequalities, it follows that

$$f(x) - f(\overline{x}) \leq f(x) - f(u) = g(x) - g(u) + h(x) - h(u)$$

$$\leq \langle \nabla g(u), x - u \rangle + \frac{L_g}{2}\|x - u\|^2 + \langle \zeta, x - u \rangle$$

$$= \langle \nabla g(u) - \nabla g(x), x - u \rangle + \frac{L_g}{2}\|x - u\|^2 + \langle \nabla g(x) + \zeta, x - u \rangle$$

$$\leq (3L_g/2)\|x - u\|^2 + \langle \nabla g(x) + \zeta, x - u \rangle$$

$$= (3L_g/2)\text{dist}^2(x, (\partial f)^{-1}(0)) + \langle \nabla g(x) + \zeta, x - u \rangle,$$

where the first inequality is due to $u \in B(\overline{x}, \delta) \cap \text{crit } f$ and Assumption 3.1. Taking the infimum over all possible $\zeta \in \partial h(x)$, by (4) and the last inequality,

$$f(x) - f(\overline{x}) \leq \frac{3L_g}{2}\text{dist}^2(x, (\partial f)^{-1}(0)) + \text{dist}(0, \partial f(x))\text{dist}(x, (\partial f)^{-1}(0))$$

$$\leq \left(\frac{3L_g}{2} + 1\right)\text{dist}^2(0, \partial f(x)).$$

By the arbitrariness of $\overline{x}$ in $\text{crit } f$, $f$ is a KL function of exponent $1/2$. \hfill \Box

Recently, Li and Pong [8] showed that under Assumption 3.1, the KL property of exponent $1/2$ is also implied by the Luo-Tseng error bound [27] which is stated as: “for any $\zeta \geq \inf_{z \in \mathbb{X}} f(z)$, there exist $\tilde{\varepsilon} > 0$ and $\tilde{\kappa} > 0$ such that

$$\text{dist}(x, (\partial f)^{-1}(0)) \leq \tilde{\kappa}\|\text{prox}_{h}(x - \nabla g(x)) - x\|$$

(14)

whenever $f(x) \leq \zeta$ and $\|\text{prox}_{h}(x - \nabla g(x)) - x\| < \tilde{\varepsilon}$. Then, it is natural to ask what is the link between the metric subregularity of $\partial f$ relative to $\text{crit } f$ and the Luo-Tseng error bound. For a class of structured convex $f$, the results of [9, Section 3] actually show that under the compactness of $\text{crit } f$, the former implies the latter, and we here extend it to the nonconvex case.
Lemma 3.2 For the nonconvex $f$ in (13), the following assertions hold.

(i) If the function $f$ is continuous on $\text{crit} f$, then the Luo-Tseng error bound implies that $\partial f$ is metrically subregular relative to $\text{crit} f$.

(ii) If the critical set $\text{crit} f$ is compact, the metric subregularity of $\partial f$ relative to $\text{crit} f$ implies the following error bound: there exist $\tau > 0$ and $\varpi > 0$ such that for all $z \in \text{dom} f$ with $\text{dist}(z, (\partial f)^{-1}(0)) \leq \tau$,

$$\text{dist}(z, (\partial f)^{-1}(0)) \leq \varpi \|\text{prox}_h(z - \nabla g(z)) - z\|, \quad (15)$$

which in turn implies that the Luo-Tseng error bound holds.

Proof: For convenience, we write $R(x) := \text{prox}_h(x - \nabla g(x)) - x$ for $x \in X$.

(i) Fix an arbitrary $\overline{x} \in \text{crit} f$. Since $g$ is continuously differentiable on $\text{dom} g$, $\nabla g$ is locally Lipschitz continuous at $\overline{x}$. Hence, there exist $\delta' > 0$ and $L > 0$ such that for all $z \in B(\overline{x}, \delta'), \|\nabla g(z) - \nabla g(\overline{x})\| \leq L\|z - \overline{x}\|$, and consequently

$$\|R(x)\| = \|R(x) - R(\overline{x})\| \leq (L + 2)\|x - \overline{x}\|,$$

where the equality is due to $R(\overline{x}) = 0$. Since the Luo-Tseng error bound holds, there exist $\tilde{\varepsilon} > 0$ and $\tilde{\kappa} > 0$ such that $\text{dist}(z, (\partial f)^{-1}(0)) \leq \tilde{\kappa}\|R(z)\|$ whenever $f(z) \leq f(\overline{x}) + 1$ and $\|R(z)\| < \tilde{\varepsilon}$. Since $f$ is continuous at $\overline{x}$, there exists $\delta'' > 0$ such that for all $z \in B(\overline{x}, \delta'')$, $f(z) \leq f(\overline{x}) + 1$. Set $\varepsilon = \min(\delta', \delta'', \frac{\tilde{\varepsilon}}{L + 1})$. Then, for any $x \in B(\overline{x}, \varepsilon)$, $f(x) \leq f(\overline{x}) + 1$ and $\|R(x)\| < \tilde{\varepsilon}$, and hence $\text{dist}(x, (\partial f)^{-1}(0)) \leq \tilde{\kappa}\|R(x)\|$, which by [8, Lemma 4.1] implies $\text{dist}(x, (\partial f)^{-1}(0)) \leq \tilde{\kappa}\text{dist}(0, \partial f(x))$. So, $\partial f$ is metrically subregular at $\overline{x}$ for $0$. By the arbitrariness of $\overline{x}$, the mapping $\partial f$ is metrically subregular relative to $\text{crit} f$.

(ii) Since $\text{crit} f \subseteq \text{dom} g$ and $\text{crit} f$ is compact, there exists $\delta' > 0$ such that $\{u \in X : \text{dist}(u, \text{crit} f) \leq \delta'\} \subseteq \text{dom} g$. By the smoothness of $g$ on $\text{dom} g$, there exists $L > 0$ such that for all $z, z' \in \{u \in X : \text{dist}(u, \text{crit} f) \leq \delta'\}$,

$$\|\nabla g(z) - \nabla g(z')\| \leq L\|z - z'\|.$$ 

In addition, since $\partial f$ is metrically subregular relative to the compact $\text{crit} f$, by following the similar arguments as those for [9, Proposition 2], there exist $\delta'' > 0$ and $\kappa'' > 0$ such that for all $x \in \text{crit} f + \delta'' B_X$,

$$\text{dist}(x, (\partial f)^{-1}(0)) \leq \kappa''\text{dist}(0, \partial f(x)). \quad (16)$$

Fix an arbitrary $z \in \text{dom} f$ with $\text{dist}(z, (\partial f)^{-1}(0)) \leq \min(\frac{\delta''}{L + 1}, \delta')$. Notice that

$$\text{dist}(z + R(z), \text{crit} f) = \text{dist}(z + R(z), (\partial f)^{-1}(0)) \leq \text{dist}(z, (\partial f)^{-1}(0)) + \|R(z)\|$$

$$= \text{dist}(z, (\partial f)^{-1}(0)) + \|R(z) - R(z^*)\|$$

$$\leq (L + 3)\text{dist}(z, (\partial f)^{-1}(0)) \leq \delta''.$$
where the inequality is using $\|\nabla g(z) - \nabla g(z^*)\| \leq L \|z - z^*\|$. By (16) we have
\[
\text{dist}(z + R(z), (\partial f)^{-1}(0)) \leq \kappa'' \text{dist}(0, \partial f(z + R(z)))
\]
\[
= \kappa'' \min_{\eta \in \partial h(\text{prox}_h(z - \nabla g(z))) \cap \partial h(z + R(z))} \|\eta + \nabla g(z + R(z))\|.
\]
Notice that $\nabla g(z) - \nabla g(z) \in \partial h(\text{prox}_h(z - \nabla g(z))) = \partial h(z + R(z))$. Then,
\[
\text{dist}(z + R(z), (\partial f)^{-1}(0)) \leq \kappa'' \|\nabla g(z) - \nabla g(z) + \nabla g(z + R(z))\|
\]
\[
\leq \kappa'' (1 + L) \|R(z)\|,
\]
which implies that $\text{dist}(z, (\partial f)^{-1}(0)) \leq [\kappa'' (1 + L) + 1] \|R(z)\|$. By the arbitrariness of $z$, (15) holds with $\tau = \min\left(\frac{\delta''}{3+L}, \delta'\right)$ and $\kappa' = [\kappa'' (1 + L) + 1]$. The last part follows by using the same arguments as those for [9, Proposition 3].

\section{Conclusions}

For a proper lsc function, we have disclosed the relation between the metric subregularity of its subdifferential relative to the critical set and its KL property of exponent $1/2$, and the link between the former and the Luo-Tseng error bound for a class structured nonconvex functions. The obtained results are helpful to excavate new classes of KL functions of exponent $1/2$.

\section*{Acknowledgements}

This work is supported by the National Natural Science Foundation of China under project No.11571120 and No.11701186. The first author would like to thank Professor Ting Kei Pong for pointing out a gap in the original Remark 3.4.

\section*{References}

[1] Aragón Artacho, F. J., Geoffroy, M. H.: Characterization of metric regularity of subdifferential. J. Convex Anal. 15, 365-380 (2008)

[2] Luque, F. J.: Asymptotic convergence analysis of the proximal point algorithm. SIAM J. Control Optim. 22, 277-293 (1984)

[3] Aragón Artacho, F. J., Geoffroy, M. H.: Metric subregularity of the convex subdifferential in Banach spaces. J. Nonlinear Convex Anal. 15, 35-47 (2014)

[4] Drusvyatskiy D., Ioffe, A. D.: Quadratic growth and critical point stability of semialgebraic functions. Math. Program. 3, 243-264 (2014)

[5] Attouch, H., Bolte, J., Redont, P., Soubeyran, A.: Proximal alternating minimization and projection methods for nonconvex problems: an approach based on the Kurdyka-Łojasiewicz inequality. Math. Oper. Res. 35, 438-457 (2010)
[6] Attouch, H., Bolte, J., Svaiter, B. F.: Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods. Math. Program. 137, 91-129 (2013)

[7] Bolte, J., Sabach, S., Teboulle, M.: Proximal alternating linearized minimization for nonconvex and nonsmooth problems. Math. Program. 146, 459-494 (2014).

[8] Li, G. Y., Pong, T. K.: Calculus of the exponent of Kurdyka-Łojasiewicz inequality and its applications to linear convergence of first-order methods. Found. Comput. Math. DOI 10.1007/s10208-017-9366-8.

[9] Zhou Z. R., So, A. M.-C.: A unified approach to error bounds for structured convex optimization problems. Math. Program. 165, 689-728 (2017)

[10] Rockafellar, R. T., Wets, R. J-B.: Variational Analysis, Springer, New York (1998)

[11] Dontchev A. L., Rockafellar, R. T.: Implicit Functions and Solution Mappings, Springer Monographs in Mathematics, New York (2009)

[12] Gfrerer, H.: On directional metric regularity, subregularity and optimality conditions for nonsmooth mathematical programs. Set-Valued Var. Anal. 21, 151-176 (2013)

[13] Henrion, H., Outrata, J. V.: On the calmness of a class of multifunctions. SIAM J. Optim. 13, 603-618 (2002)

[14] Poliquin, R. A.: Integration of subdifferentials of nonconvex functions. Nonlinear Anal. 17, 385-398 (1991)

[15] Marcellin, S., Thibault, L.: Evolution problems associated with primal lower nice functions. J. Convex Anal. 13, 385-421 (2006)

[16] Bolte, J., Daniilidis, A., Ley, O., Mazet, L.: Characterizations of Łojasiewicz inequalities: subgradient flows, talweg, convexity. Trans. Amer. Math. Soc. 362, 3319-3363 (2010)

[17] Rockafellar, R. T.: Favorable classes of Lipschitz-continuous functions in subgradient optimization. In: Nurminski, E. (eds.): Progress in Nondifferentiable Optimization, pp. 125-144. IIASA Collaborative Proceedings Series. International Institute of Applied Systems Analysis, Laxenburg, Austria (1982).

[18] Bolte, J., Nguyen, T. P., Peyouquet J., Suter, B. W.: From error bounds to the complexity of first-order descent methods for convex functions. Math. Program. 165, 471-507 (2017)

[19] Brézis, H.: Opérateurs maximaux monotones et semi-groupes de contractions dans les espace Hilbert. North-Holland Mathematics studies 5.
[20] Bruck, R. E.: Asymptotic convergence of nonlinear contraction semigroups in Hilbert space. J. Funct. Anal. 18, 15-26 (1975)

[21] Dong, Y. D.: An extension of Luque’s growth condition. Appl. Math. Lett. 22, 1390-1393 (2009)

[22] Cui, Y., Sun D. F., Toh, K. C.: On the asymptotic superlinear convergence of the augmented Lagrangian method for semidefinite programming with multiple solutions. arXiv:1610.00875, 2016.

[23] Clarke, F. H.: Nonsmooth Analysis and Optimization, Wiley, New York (1983)

[24] Robinson, S. M.: Some continuity properties of polyhedral multifunctions. Math. Program. Study 14, 206-214 (1981)

[25] Fan J. Q., Li, R. Z.: Variable selection via nonconcave penalized likelihood and its oracle properties. J. Amer. Statist. Assoc. 96, 1348-1360 (2001)

[26] Zhang, C. H.: Nearly unbiased variable selection under minimax concave penalty. Ann. Statist. 38, 894-942 (2010)

[27] Tseng, P., Yun, S.: A coordinate gradient descent method for nonsmooth separable minimization. Math. Program. 117, 387-423 (2009)

[28] Sun, J.: On Monotropic Piecewise Quadratic Programming. Ph.D Thesis, Department of Mathematics, University of Washington, Seattle (1986)

[29] Zhou, Z. R.: Error Bounds for Structured Convex Programming: Theory and Applications. PhD thesis, The Chinese University of Hong Kong (2015)

[30] Sun Y., Pan S. H., Bi, S. J.: Metric subregularity and/or calmness of the normal cone mapping to the $p$-order conic constraint system. Accepted by Optim. Lett., (2018)

[31] Liu, Y. L., Pan, S. H.: Computation of graphical derivatives of normal cone maps to a class of conic constraint sets. Set-Valued Var. Anal., Doi.org/10.1007/s11228-018-0494-3

[32] Cui, Y., Ding, C., Zhao, X. Y.: Quadratic growth conditions for convex matrix optimization problems associated with spectral functions. SIAM J. Optim. 27, 2332-2355 (2017)

A Some functions with metric subregular subdifferential

Lemma A.1 The subdifferentials of the following functions are metrically subregular at every point of its graph:

(i) the piecewise linear function;
(ii) the piecewise linear-quadratic convex function and its conjugate;

(iii) the vector $\ell_p$-norm with $p \in [1, 2] \cup \{+\infty\}$;

(iv) the indicator function on the $p$-order cone with $p \in [1, 2] \cup \{+\infty\}$;

(v) the indicator function on the $C^2$-cone reducible self-dual closed convex cone;

(vi) the spectral function $g(\lambda(X))$ for $X \in \mathbb{S}^n$ where $g: \mathbb{R}^n \rightarrow (-\infty, +\infty)$ is a proper lsc symmetric convex function and $\partial g$ is metrically subregular at any $(\overline{x}, \overline{y}) \in \operatorname{gph} \partial g$;

(vii) the spectral function $g(\sigma(X))$ for $X \in \mathbb{R}^{n_1 \times n_2}$ where $g: \mathbb{R}^n \rightarrow (-\infty, +\infty]$ is a proper lsc absolute symmetric convex function with $n = \min(n_1, n_2)$ and $\partial g$ is metrically subregular at any $(\overline{x}, \overline{y}) \in \operatorname{gph} \partial g$.

**Proof:** Part (i) follows by [24, Proposition 1]; part (ii) follows from [28] or [10, Proposition 12.30]; part (iii) is due to [29]; part (iv) can be found in [30]; part (v) is due to [31, Theorem 2.1]; and part (vi)-(vii) follows by [32, Proposition 15].

By following the arguments in [9, Section 3.3] and the proof of [22, Theorem 3.1], it is not difficult to verify that the following conclusion holds.

**Lemma A.2** Consider $f(x) = g(Ax) + \phi(x)$ with $\operatorname{crit} f \neq \emptyset$, where $g: \mathbb{Z} \rightarrow (-\infty, +\infty]$ is strongly convex with an open domain and is continuously differentiable on $\operatorname{dom} g$, and $\phi: \mathbb{X} \rightarrow (-\infty, +\infty]$ is a proper lsc convex function with $\operatorname{dom} \phi \subseteq A^{-1}(\operatorname{dom} g)$. Suppose that $A^{-1}\{Ax^*\} \cap \text{ri}[\partial \phi^*(-A^*\nabla g(Ax^*))] \neq \emptyset$ for some $x^* \in \operatorname{crit} f$ and $\partial \phi$ is metrically subregular at every $(x, y) \in \operatorname{gph} \partial \phi$. Then, $\partial f$ is metrically subregular relative to $\operatorname{crit} f$. 

16