



A subgradient projection method for quasiconvex minimization

Juan Choque¹ · Felipe Lara² · Raúl T. Marcavillaca³

Received: 6 March 2024 / Accepted: 24 August 2024
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2024

Abstract

In this paper, a subgradient projection method for quasiconvex minimization problems is provided. By employing strong subdifferentials, it is proved that the generated sequence of the proposed algorithm converges to the solution of the minimization problem of a proper, lower semicontinuous, and strongly quasiconvex function (in the sense of Polyak in Soviet Math 7:72–75, 1966), under the same assumptions as those required for convex functions with the convex subdifferentials. Furthermore, a quasi-linear convergence rate of the iterates, extending similar results for the general quasiconvex case, is also provided.

Keywords Subgradient methods · First-order methods · Nonconvex optimization · Generalized convexity · Quasiconvexity

Mathematics Subject Classification 90C26 · 90C30 · 49M37

1 Introduction

Let K be a closed and convex set in \mathbb{R}^n and $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$ be a proper function such that $K \subseteq \text{dom } h$. We consider the constraint minimization problem

In loving memory of Juan Choque.

✉ Felipe Lara
felipelaraobrequ@gmail.com; flarao@uta.cl
https://felipelara.cl

Juan Choque
jucachof@gmail.com

Raúl T. Marcavillaca
raultm.rt@gmail.com; rtintaya@dim.uchile.cl

- 1 Programa de Magíster en Ciencias con Mención en Matemática, Facultad de Ciencias, Universidad de Tarapacá, Arica, Chile
- 2 Instituto de Alta Investigación (IAI), Universidad de Tarapacá, Arica, Chile
- 3 Center of Mathematical Modelling (CMM), Universidad de Chile, Santiago, Chile

given by:

$$\min_{x \in K} h(x). \quad (\text{COP})$$

If h is differentiable, then one of the most important iterative methods for finding a solution of problem (COP) is the gradient method, which has been well studied due to its important properties and applications in several fields of the mathematical sciences, engineering and economics among others.

If h is not differentiable, but lower semicontinuous (lsc henceforth) and convex, then we may apply the subgradient algorithm: Given $z^k \in \text{int dom } h \subseteq K$, we take

$$z^{k+1} = z^k - \alpha_k \xi^k, \quad (1.1)$$

where $\xi^k \in \partial h(z^k)$ is the convex subdifferential (see (2.7) below) and $\{\alpha_k\}_k$ is a positive sequence of parameters. If $\xi^k = 0$, then z^k is a solution of problem (COP), if not, then we take $k = k + 1$ and we repeat (1.1).

In contrast to the gradient method, the subgradient method (with the convex subdifferential) does not decrease at each step even for convex functions (see [2, Example 8.3] and also [20]). However, the (convex) subgradient method converges to the solution of the problem for convex functions (see [2, Theorem 8.17] for instance), but also for classes of quasiconvex functions with generalized subdifferentials (see [6, 7, 9, 13, 14, 22] and references therein).

In this paper, and by using new existence results and generalized subdifferentials, we provide a subgradient projection method for finding the unique solution of the minimization problem of a proper, lsc and strongly quasiconvex function. The subdifferential that we use is the strong subdifferential (see [12]) which has been especially introduced for studying strongly quasiconvex functions. We prove the convergence of the generated sequence to the solution of problem (COP) for a proper, lsc and strongly quasiconvex function under the same assumptions than the convex case with the convex subdifferential. Recall that the class of strongly quasiconvex functions were introduced in the sixties, it is the natural extension of the strongly convex functions and its definition goes back to the famous paper on existence theorems in extremum problems of Polyak [18].

The structure of the paper is as follows: In Sect. 2 we present notation, preliminaries and the basic results on generalized convexity and generalized subdifferentials. In Sect. 3, we present new theoretical results for the strong subdifferential which will be useful for the convergence analysis of the algorithm, then we present our subgradient method and the basic assumptions for its implementation. Furthermore, we provide the convergence analysis for the proposed algorithm in which we prove the convergence of the generated sequence to the unique solution of problem (COP) under a proper, lsc, and strongly quasiconvexity assumption with a quasi-linear convergence rate. Finally, theoretical examples showing the advantages of our approach are discussed in the conclusions.

2 Preliminaries and basic definitions

The inner product in \mathbb{R}^n and the Euclidean norm are denoted by $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$, respectively. Given a convex and closed set K , the projection of x on K is denoted by $P_K(x)$, and the indicator function on K by ι_K . The set $]0, +\infty[$ is denoted by \mathbb{R}_{++} . The unit sphere centered at the origin is denoted by \mathbb{B} .

For a closed and convex set $K \subseteq \mathbb{R}^n$, the metric projection satisfies that

$$v = P_K(u) \iff \langle u - v, w - v \rangle \leq 0, \forall w \in K. \tag{2.1}$$

Given any extended-valued function $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, the effective domain of h is defined by $\text{dom } h := \{x \in \mathbb{R}^n : h(x) < +\infty\}$. It is said that h is proper if $\text{dom } h$ is nonempty and $h(x) > -\infty$ for all $x \in \mathbb{R}^n$. The notion of properness is important when dealing with minimization problems.

It is indicated by $\text{epi } h := \{(x, t) \in \mathbb{R}^n \times \mathbb{R} : h(x) \leq t\}$ the epigraph of h , by $S_\lambda(h) := \{x \in \mathbb{R}^n : h(x) \leq \lambda\}$ the sublevel set of h at the height $\lambda \in \mathbb{R}$ and by $\text{argmin}_{\mathbb{R}^n} h$ the set of all minimal points of h . A function h is lower semicontinuous at $\bar{x} \in \mathbb{R}^n$ if for any sequence $\{x_k\}_k \in \mathbb{R}^n$ with $x_k \rightarrow \bar{x}$, $h(\bar{x}) \leq \liminf_{k \rightarrow +\infty} h(x_k)$. Furthermore, the current convention $\sup_{\emptyset} h := -\infty$ and $\inf_{\emptyset} h := +\infty$ is adopted.

A function h with a convex domain is said to be

(a) convex if, given any $x, y \in \text{dom } h$, then

$$h(\lambda x + (1 - \lambda)y) \leq \lambda h(x) + (1 - \lambda)h(y), \forall \lambda \in [0, 1]; \tag{2.2}$$

(b) semistrictly quasiconvex if, given any $x, y \in \text{dom } h$, with $h(x) \neq h(y)$, then

$$h(\lambda x + (1 - \lambda)y) < \max\{h(x), h(y)\}, \forall \lambda \in]0, 1[; \tag{2.3}$$

(c) quasiconvex if, given any $x, y \in \text{dom } h$, then

$$h(\lambda x + (1 - \lambda)y) \leq \max\{h(x), h(y)\}, \forall \lambda \in [0, 1]. \tag{2.4}$$

It is said that h is strictly convex (resp. strictly quasiconvex) if the inequality in (2.2) (resp. (2.4)) is strict whenever $x \neq y$.

Every (strictly) convex function is (strictly) quasiconvex and semistrictly quasiconvex, and every semistrictly quasiconvex and lsc function is quasiconvex. The continuous function $h : \mathbb{R} \rightarrow \mathbb{R}$ given by $h(x) := \min\{|x|, 1\}$ is quasiconvex without being semistrictly quasiconvex. Furthermore, recall that a function h is convex iff $\text{epi } h$ is a convex set and that is quasiconvex iff the sublevel sets $S_\lambda(h)$ are convex sets for all $\lambda \in \mathbb{R}$.

A function h with a convex domain is said to be (see [1, 2, 11, 18]):

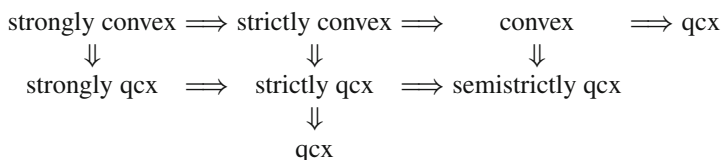
(a) strongly convex on $\text{dom } h$ if there exists $\gamma \in]0, +\infty[$ such that for all $x, y \in \text{dom } h$ and all $\lambda \in [0, 1]$, we have

$$h(\lambda y + (1 - \lambda)x) \leq \lambda h(y) + (1 - \lambda)h(x) - \lambda(1 - \lambda)\frac{\gamma}{2}\|x - y\|^2, \tag{2.5}$$

(b) strongly quasiconvex on $\text{dom } h$ if there exists $\gamma \in]0, +\infty[$ such that for all $x, y \in \text{dom } h$ and all $\lambda \in [0, 1]$, we have

$$h(\lambda y + (1 - \lambda)x) \leq \max\{h(y), h(x)\} - \lambda(1 - \lambda) \frac{\gamma}{2} \|x - y\|^2. \tag{2.6}$$

In these cases, it is said that h is strongly convex (resp. quasiconvex) with modulus $\gamma > 0$. Note that every strongly convex function is strongly quasiconvex, and every strongly quasiconvex function is strictly quasiconvex. The Euclidean norm $h_1(x) = \|x\|$ is strongly quasiconvex without being strongly convex on any bounded convex set $K \subseteq \mathbb{R}^n$ (see [11, Theorem 2]) and the function $h_2(x) = x^3$ is strictly quasiconvex without being strongly quasiconvex on \mathbb{R} . Summarizing (quasiconvex is denoted by qcx):



If in addition the function is lsc, then all the previous notions are quasiconvex.

Remark 2.1 There is no relationship between convex and strongly quasiconvex functions. Indeed, the function $h : \mathbb{R}^n \rightarrow \mathbb{R}$ given by $h(x) = \sqrt{\|x\|}$ is strongly quasiconvex on any bounded and convex set K in \mathbb{R}^n without being convex ([15, Theorem 17]), while the function $h(x) \equiv 1$ is convex without being strongly quasiconvex. However, strongly convex functions are both convex and strongly quasiconvex.

The following existence result is the starting point of our research.

Lemma 2.1 ([15, Corollary 3]) *Let $K \subseteq \mathbb{R}^n$ be a closed and convex set and $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be a proper, lsc, and strongly quasiconvex function on $K \subseteq \text{dom } h$ with modulus $\gamma > 0$. Then, $\text{argmin}_K h$ is a singleton.*

Given a proper function $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, the convex subdifferential of h at $\bar{x} \in \text{dom } h$ is defined by

$$\partial h(\bar{x}) := \{\xi \in \mathbb{R}^n : h(y) \geq h(\bar{x}) + \langle \xi, y - \bar{x} \rangle, \forall y \in \mathbb{R}^n\}, \tag{2.7}$$

and by $\partial h(x) = \emptyset$ if $x \notin \text{dom } h$.

The convex subdifferential is an outstanding tool in continuous optimization. It is especially useful when the function is proper, lsc and convex, but when the function is nonconvex, the convex subdifferential is no longer useful. For this reason, several authors have introduced different generalized subdifferentials for dealing with different classes of nonconvex functions (see [16, 17, 19]).

For the case of quasiconvex and strongly quasiconvex functions, we recall the following notion of generalized subdifferential from [12].

The strong subdifferential of h at $\bar{x} \in \text{dom } h \cap K$ is defined by

$$\begin{aligned} \partial_{\beta,\gamma}^K h(\bar{x}) := \{ \xi \in \mathbb{R}^n : \max\{h(y), h(\bar{x})\} \geq h(\bar{x}) + \frac{\lambda}{\beta} \langle \xi, y - \bar{x} \rangle \\ + \frac{\lambda}{2} \left(\gamma - \frac{\lambda}{\beta} - \lambda\gamma \right) \|y - \bar{x}\|^2, \forall y \in K, \forall \lambda \in [0, 1] \}. \end{aligned} \tag{2.8}$$

Clearly, $\partial_{\beta,\gamma}^K$ is a closed and convex set. Furthermore, we also have.

Lemma 2.2 ([12, Proposition 7(d)]) *Let $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be a proper function, $\beta > 0$, $\gamma \geq 0$, $K \subseteq \mathbb{R}^n$, and $\bar{x} \in \text{dom } h \cap K$. Then $\partial_{\beta,\gamma}^K h(\bar{x})$ is compact for every $\bar{x} \in \text{int}(\text{dom } h \cap K)$.*

We recall some interesting and useful properties of the strong subdifferentials. Before that, we first recall that \bar{x} is an (α, K) -strong minimum point of h if there exists $\alpha > 0$ such that

$$h(y) \geq h(\bar{x}) + \alpha \|y - \bar{x}\|^2, \forall y \in K \setminus \{\bar{x}\}, \tag{2.9}$$

Clearly, every (α, K) -strong minimum point is a strict minimum point and every strict minimum point is a global minimum point, but both reverse statements do not hold.

Lemma 2.3 ([12, Theorem 24(a)]) *Let $K \subseteq \mathbb{R}^n$ be a closed and convex set, $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be a proper function such that $K \subseteq \text{dom } h$, $\beta > 0$, $\gamma > 0$ and $\bar{x} \in K$. Then, $0 \in \partial_{\beta,\gamma}^K h(\bar{x})$ iff \bar{x} is a $\left(\frac{\gamma^2\beta}{8(1+\gamma\beta)}, K\right)$ -strong minimum point of h .*

The strong subdifferential is nonempty for bigger classes of quasiconvex functions as we can see below.

Lemma 2.4 ([12, Corollary 38(a)]) *Let $K \subseteq \mathbb{R}^n$ be a closed and convex set, $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be a proper and lsc function such that $K \subseteq \text{dom } h$, and $\beta > 0$. If h is strongly quasiconvex on K with modulus $\gamma > 0$, then $\partial_{\beta,\gamma}^K h(\bar{x}) \neq \emptyset$ for every $\bar{x} \in K$.*

Another useful property is the following.

Lemma 2.5 ([12, Proposition 40]) *Let $K \subseteq \mathbb{R}^n$ be a closed and convex set, $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be a proper and lsc function such that $K \subseteq \text{dom } h$, $\beta > 0$ and $\bar{x} \in K$. If h is strongly quasiconvex on K with modulus $\gamma > 0$, then*

$$y \in K \cap S_{h(\bar{x})}(h) \implies \langle \xi, y - \bar{x} \rangle \leq -\frac{\beta\gamma}{2} \|y - \bar{x}\|^2, \forall \xi \in \partial_{\beta,\gamma}^K h(\bar{x}). \tag{2.10}$$

We also recall the notion of quasi-Fejér convergence and its properties. A detailed survey on this subject can be found in [5].

Definition 2.1 Let $C \subseteq \mathbb{R}^n$ and $\{x^k\}_k \subseteq \mathbb{R}^n$ be any sequence in \mathbb{R}^n . Then,

- (i) The sequence $\{x^k\}_k \subseteq \mathbb{R}^n$ is called quasi-Fejér related to set $C \subseteq \mathbb{R}^n$ if for every $z \in C$ there exists a sequence $\{\epsilon_k\}_k \subseteq \mathbb{R}_+$ such that

$$\|x^{k+1} - z\|^2 \leq \|x^k - z\|^2 + \epsilon_k,$$

with $\sum_{k=1}^{+\infty} \epsilon_k < +\infty$.

- (ii) The sequence $\{x^k\}_k \subseteq \mathbb{R}^n$ converges to $\bar{x} \in \mathbb{R}^n$ quasi R -linearly, if there exists a constant $0 < \theta < 1$, $k_0 \in \mathbb{N}$, and $\{\epsilon_k\}_k \subseteq \mathbb{R}_+$ (which does not depend on x_k) such that

$$\|x^{k+1} - \bar{x}\|^2 \leq \theta \|x^k - \bar{x}\|^2 + \epsilon_k, \forall k \geq k_0$$

with $\sum_{k=1}^{+\infty} \epsilon_k < +\infty$.

To end this section, we state the main properties of the quasi-Fejér sequences that will be needed in our analysis.

Lemma 2.6 ([4, Theorem 1]) *Let $\{x^k\}_k \subseteq \mathbb{R}^n$ be quasi-Fejér convergent to a nonempty set $C \subseteq \mathbb{R}^n$. Then, the following assertions hold:*

- (a) $\{x^k\}_k$ is bounded;
 (b) If a cluster point x^* of $\{x^k\}_k$ belongs to C , then $\lim_{k \rightarrow +\infty} x^k = x^*$.

For a further study on nonsmooth analysis and generalized convexity we refer to [11, 12, 15–19].

3 Subgradient methods for quasiconvex minimization

Throughout the paper, we assume the following assumption on h :

- (A) h is a proper, lsc and strongly quasiconvex function on $K \subseteq \text{dom } h$ with modulus $\gamma > 0$.

3.1 Properties for the strong subdifferential

Before introducing our algorithm, we show some useful theoretical results.

Proposition 3.1 *Let $K \subseteq \mathbb{R}^n$ be a closed and convex set, $h : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ be a proper function such that $K \subseteq \text{dom } h$, $\beta > 0$, $\gamma > 0$ and $z \in K$. Given $\alpha > 0$ and $\xi \in \partial_{\beta, \gamma}^K h(z)$, set*

$$x^+ := P_K(z - \alpha\xi). \quad (3.1)$$

Then the following assertions hold:

- (a) $\|x^+ - z\| \leq \alpha \|\xi\|$.

(b) For every $x \in K$, we have

$$\|x^+ - x\|^2 \leq \|z - x\|^2 + \alpha^2 \|\xi\|^2 + 2\alpha \langle \xi, x - z \rangle. \tag{3.2}$$

Proof (a): It follows from (3.1) and (2.1) that,

$$\langle z - \alpha\xi - x^+, y - x^+ \rangle \leq 0, \forall y \in K,$$

taking $y = z$, then using the Cauchy–Schwarz inequality, we deduce

$$\|z - x^+\|^2 \leq \alpha \langle \xi, z - x^+ \rangle \leq \alpha \|\xi\| \|z - x^+\|$$

(b): Let $x \in K$. Using the nonexpansivity of the projection operator, we have

$$\begin{aligned} \|x^+ - x\|^2 &= \|P_K(z - \alpha\xi) - P_K(x)\|^2 \\ &\leq \|z - \alpha\xi - x\|^2 \\ &= \|z - x\|^2 + \alpha^2 \|\xi\|^2 + 2\alpha \langle \xi, x - z \rangle, \end{aligned}$$

which proves (3.2). □

By using the structure of the strong subdifferential, we obtain the following.

Proposition 3.2 *Let $K \subseteq \mathbb{R}^n$ be a closed and convex set, $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be a proper function such that $K \subseteq \text{dom } h$, $\beta > 0$, $\gamma > 0$ and $z \in K$. Given $\xi \in \partial_{\beta, \gamma}^K h(z)$, $\alpha > 0$ and $x^+ := P_K(z - \alpha\xi)$. Then, for every $x \in K$, we have*

(a) If $h(x) > h(z)$, then

$$\|x^+ - x\|^2 \leq \|z - x\|^2 + \alpha^2 \|\xi\|^2 + \frac{2\alpha(1 + \gamma\beta)}{\gamma} (h(x) - h(z)). \tag{3.3}$$

(b) If $h(x) \leq h(z)$, then

$$\|x^+ - x\|^2 \leq (1 - \alpha\gamma\beta) \|z - x\|^2 + \alpha^2 \|\xi\|^2. \tag{3.4}$$

Proof Since $\xi \in \partial_{\beta, \gamma}^K h(z)$, we have

$$\begin{aligned} \max\{h(y), h(z)\} - h(z) &\geq \frac{\lambda}{\beta} \langle \xi, y - z \rangle + \frac{\lambda}{2} \left(\gamma - \frac{\lambda}{\beta} - \lambda\gamma \right) \|y - z\|^2, \\ &\forall y \in K, \forall \lambda \in [0, 1]. \end{aligned} \tag{3.5}$$

Taking $y = x$ into (3.5), we separate in two cases:

(a) If $h(x) > h(z)$, then $\max\{h(x), h(z)\} - h(z) = h(x) - h(z)$. Hence (3.5) becomes

$$\langle \xi, x - z \rangle \leq -\frac{\beta}{2} \left(\gamma - \frac{\lambda}{\beta} - \lambda\gamma \right) \|x - z\|^2 + \frac{\beta}{\lambda} (h(x) - h(z)).$$

Replacing this in (3.2), we obtain

$$\|x^+ - x\|^2 \leq \left(1 - \alpha\beta\left(\gamma - \frac{\lambda}{\beta} - \lambda\gamma\right)\right) \|z - x\|^2 + \alpha^2 \|\xi\|^2 + 2\frac{\alpha\beta}{\lambda} (h(x) - h(z)).$$

Choosing $0 < \lambda = \frac{\gamma\beta}{1+\gamma\beta} < 1$ and taking into account $h(x) > h(z)$, we have

$$\|x^+ - x\|^2 \leq \|z - x\|^2 + \alpha^2 \|\xi\|^2 + \frac{2\alpha(1 + \gamma\beta)}{\gamma} (h(x) - h(z)).$$

(b) If $h(x) \leq h(z)$, then $\max\{h(x), h(z)\} - h(z) = 0$. By a similar argument to the used at item (a), we obtain

$$\|x^+ - x\|^2 \leq \left(1 - \alpha\beta\left(\gamma - \frac{\lambda}{\beta} - \lambda\gamma\right)\right) \|z - x\|^2 + \alpha^2 \|\xi\|^2.$$

Taking $\lambda \rightarrow 0^+$, we conclude

$$\|x^+ - x\|^2 \leq (1 - \alpha\gamma\beta) \|z - x\|^2 + \alpha^2 \|\xi\|^2,$$

which completes the proof. □

An interesting theoretical result for continuous strongly quasiconvex functions is given below.

Proposition 3.3 *Let $K \subseteq \mathbb{R}^n$ be a closed and convex set and $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be a proper, continuous and strongly quasiconvex function on $K \subseteq \text{dom } h$ with modulus $\gamma > 0$. Let $\{x^k\}_k \subseteq K$ and $\{\xi^k\}_k \subseteq \mathbb{R}^n$ be such that $\xi^k \in \partial_{\beta,\gamma}^K h(x^k)$ for all $k \in \mathbb{N}$. If $\xi^k \rightarrow 0$ and $x^k \rightarrow \bar{x}$, then $0 \in \partial_{\beta,\gamma}^K h(\bar{x})$. Hence, the point \bar{x} is a global minimizer.*

Proof Suppose the contrary that $0 \notin \partial_{\beta,\gamma}^K h(\bar{x})$, then by [12, Corollary 37] there exists $\widehat{x} \in K$ such that

$$h(\widehat{x}) < h(\bar{x}) + \frac{\gamma^2\beta}{8(1 + \gamma\beta)} \|\widehat{x} - \bar{x}\|^2. \tag{3.6}$$

Furthermore, since $\xi^k \in \partial_{\beta,\gamma}^K h(x^k)$ for every $k \in \mathbb{N}$, we have

$$\begin{aligned} \max\{h(y), h(x^k)\} &\geq h(x^k) + \frac{\lambda}{\beta} \langle \xi^k, y - x^k \rangle + \frac{\lambda}{2} \left(\gamma - \frac{\lambda}{\beta} - \lambda\gamma\right) \|y - x^k\|^2, \\ &\forall y \in K, \forall \lambda \in [0, 1]. \end{aligned}$$

By taking $y = \widehat{x}$,

$$\begin{aligned} \max\{h(\widehat{x}), h(x^k)\} &\geq h(x^k) + \frac{\lambda}{\beta} \langle \xi^k, \widehat{x} - x^k \rangle + \frac{\lambda}{2} \left(\gamma - \frac{\lambda}{\beta} - \lambda\gamma\right) \|\widehat{x} - x^k\|^2, \\ &\forall \lambda \in [0, 1]. \end{aligned} \tag{3.7}$$

Taking $\lim_{k \rightarrow +\infty}$ in (3.7), and since $\xi^k \rightarrow 0$ and $x^k \rightarrow \bar{x}$, we have

$$\begin{aligned} \max\{h(\hat{x}), h(\bar{x})\} &\geq h(\bar{x}) + \frac{\lambda}{2} \left(\gamma - \frac{\lambda}{\beta} - \lambda\gamma \right) \|\hat{x} - \bar{x}\|^2, \quad \forall \lambda \in [0, 1] \\ &\geq h(\bar{x}) + \frac{\gamma^2\beta}{8(1 + \gamma\beta)} \|\hat{x} - \bar{x}\|^2, \end{aligned} \tag{3.8}$$

where the second inequality yields from the fact that the real function $[0, 1] \ni \lambda \mapsto \frac{\lambda}{2} \left(\gamma - \frac{\lambda}{\beta} - \lambda\gamma \right)$ reaches its minimum at $\lambda = \frac{\gamma\beta}{2(1+\gamma\beta)}$ and its value is $\frac{\gamma^2\beta}{8(1+\gamma\beta)}$. Then, (3.8) contradicts (3.6). Therefore, $0 \in \partial_{\beta,\gamma}^K h(\bar{x})$ and the proof is complete. \square

Remark 3.1 A similar result to Proposition 3.3 may be found in [6, Proposition 7] for Plastria’s subdifferential (see [17]) and quasiconvex Lipschitz continuous functions. The previous result is expected to be useful for implementing inexact proximal point methods for strongly quasiconvex functions using strong subdifferentials.

We finish this subsection with the following result in the continuous case, which shows that the subgradients of points in a given compact set contained in the interior of the domain are always bounded as in the convex case for the convex subdifferential.

Proposition 3.4 *Let $K \subseteq \mathbb{R}^n$ be a closed and convex set and $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be a proper, continuous and strongly quasiconvex function on $K \subseteq \text{int dom } h$ with modulus $\gamma > 0$. If K is compact, then $Y := \bigcup_{x \in K} \partial_{\beta,\gamma}^K h(x)$ is nonempty and bounded.*

Proof We follow the reasoning line of [2, Theorem 3.16]. Since h is proper, continuous and strongly quasiconvex, $Y \neq \emptyset$ by Lemma 2.4. Now, suppose for the contrary that Y is unbounded. Then there exist sequences $\{x^k\}_k \subseteq K$ and $\xi^k \in \partial_{\beta,\gamma}^K h(x^k)$ such that $\|\xi^k\| \rightarrow +\infty$ as $k \rightarrow +\infty$.

Since K is compact, $(\text{int dom } h)^c$ is closed and $K \cap (\text{int dom } h)^c = \emptyset$. Then there exists $\varepsilon > 0$ such that

$$\|x - y\| \geq \varepsilon, \quad \forall x \in K, \quad \forall y \in (\text{int dom } h)^c. \tag{3.9}$$

On the other hand, since $\xi^k \in \partial_{\beta,\gamma}^K h(x^k)$, we have

$$\max\{h(y), h(x^k)\} - h(x^k) \geq \frac{\lambda}{\beta} \langle \xi^k, y - x^k \rangle + \frac{\lambda}{2} \left(\gamma - \frac{\lambda}{\beta} - \lambda\gamma \right) \|y - x^k\|^2.$$

Take $y = x^k + \frac{\varepsilon}{2} \frac{\xi^k}{\|\xi^k\|} (\in \text{int dom } h)$. Then,

$$\begin{aligned} \max \left\{ h \left(x^k + \frac{\varepsilon}{2} \frac{\xi^k}{\|\xi^k\|} \right), h(x^k) \right\} - h(x^k) &\geq \frac{\lambda}{\beta} \frac{\varepsilon}{2} \|\xi^k\| + \frac{\lambda}{2} \frac{\varepsilon^2}{4} \left(\gamma - \frac{\lambda}{\beta} - \lambda\gamma \right) \\ &= \frac{\lambda\varepsilon}{2} \left(\frac{\|\xi^k\|}{\beta} + \frac{\varepsilon}{4} \left(\gamma - \frac{\lambda}{\beta} - \lambda\gamma \right) \right). \end{aligned}$$

Taking $0 < \lambda = \frac{\gamma\beta}{1+\gamma\beta} < 1$, we have

$$\max \left\{ h \left(x^k + \frac{\varepsilon}{2} \frac{\xi^k}{\|\xi^k\|} \right), h(x^k) \right\} - h(x^k) \geq \frac{\gamma\varepsilon}{2(1+\gamma\beta)} \|\xi^k\|. \tag{3.10}$$

We claim that the left-hand side in (3.10) is bounded. Indeed, suppose for the contrary that there exist subsequences $\{x^{k_l}\}_l$ and $\left\{ \frac{\xi^{k_l}}{\|\xi^{k_l}\|} \right\}_l$ such that

$$h \left(x^{k_l} + \frac{\varepsilon}{2} \frac{\xi^{k_l}}{\|\xi^{k_l}\|} \right) - h(x^{k_l}) \rightarrow +\infty \text{ as } l \rightarrow +\infty. \tag{3.11}$$

Now, without loss of generality, we assume that

$$\max \left\{ h \left(x^{k_l} + \frac{\varepsilon}{2} \frac{\xi^{k_l}}{\|\xi^{k_l}\|} \right), h(x^{k_l}) \right\} = h \left(x^{k_l} + \frac{\varepsilon}{2} \frac{\xi^{k_l}}{\|\xi^{k_l}\|} \right), \forall l \in \mathbb{N}.$$

Since $\{x^{k_l}\}_l$ and $\left\{ \frac{\xi^{k_l}}{\|\xi^{k_l}\|} \right\}_l$ are bounded sequences, there exist sequences $\{x^{k_{l_j}}\}_j$ and $\left\{ \frac{\xi^{k_{l_j}}}{\|\xi^{k_{l_j}}\|} \right\}_j$ such that $x^{k_{l_j}} \rightarrow \widehat{x}$ and $\frac{\xi^{k_{l_j}}}{\|\xi^{k_{l_j}}\|} \rightarrow \widehat{\xi}$ as $j \rightarrow +\infty$. Finally, since $\{x^{k_{l_j}}\}_j$, $\{x^{k_{l_j}} + \frac{\varepsilon}{2} \frac{\xi^{k_{l_j}}}{\|\xi^{k_{l_j}}\|}\}_j$ and $\widehat{x} + \frac{\varepsilon}{2}\widehat{\xi}$ belongs to $\text{int dom } h$, it follows from the continuity of h that

$$h \left(x^{k_{l_j}} + \frac{\varepsilon}{2} \frac{\xi^{k_{l_j}}}{\|\xi^{k_{l_j}}\|} \right) - h(x^{k_{l_j}}) \longrightarrow h \left(\widehat{x} + \frac{\varepsilon}{2}\widehat{\xi} \right) - h(\widehat{x}), \text{ as } j \rightarrow +\infty,$$

which contradicts (3.11). Therefore, the left-hand side in (3.10) is bounded, i.e., the right-hand side in (3.10) is bounded too, which contradicts the fact that $\|\xi^k\| \rightarrow +\infty$.

Therefore, Y is bounded, and the proof is complete □

3.2 The algorithm

In order to present our algorithm, we first consider the following compatibility conditions:

- (C1) $K \subseteq \text{int dom } h$.
- (C2) There exists $M > 0$ such that $\|\xi\| \leq M$ for all $\xi \in \partial_{\beta,\gamma}^K h(x)$ and all $x \in K$.
- (C3) The sequence $\{\alpha_k\}_k \subseteq]0, +\infty[$ is such that $0 < \alpha_k < \frac{1}{\gamma\beta}$ for every $k \in \mathbb{N}$, and

$$\sum_{k=0}^{\infty} \alpha_k = +\infty, \quad \sum_{k=1}^{\infty} \alpha_k^2 < +\infty. \tag{3.12}$$

We analyze the previous assumptions below.

Remark 3.2 (i) Assumptions (Ci) with $i = 1, 2, 3$ are usual assumptions for subgradient algorithms. Indeed, (C1) and (C2) are actually used for convex functions with the convex subdifferential (see Assumption 8.7(C) and Assumption 8.12 in [2], respectively).

(ii) As in the convex case, we know by Proposition 3.4 that if K is compact, then assumption (C3) holds for proper, continuous and strongly quasiconvex functions.

The conceptual algorithm to find a solution of problem (COP) is given below.

Algorithm 1 Subgradient Method for Strongly Quasiconvex Functions

Step 0. Take $x^0 \in K$ and set $k := 0$.

Step 1. Given $x^k \in K$, take $\xi^k \in \partial_{\beta,\gamma}^K h(x^k)$.

If $\xi^k = 0$, then STOP, x^k is a solution of problem (COP). Otherwise, go to Step 2.

Step 2. Compute

$$x^{k+1} = P_K \left(x^k - \alpha_k \xi^k \right), \tag{3.13}$$

take $k = k + 1$ and go to Step 1.

Remark 3.3 (i) Note that under assumption (A), Algorithm 1 is well-defined. Indeed, for every $k \in \mathbb{N}$, the subgradient $\xi^k \in \partial_{\beta,\gamma}^K h(x^k)$ exists by Lemma 2.4. Furthermore, under assumption (C1) and [12, Proposition 7(d)], we have for every iterate x^k that $\partial_{\beta,\gamma}^K h(x^k)$ is bounded.

(ii) If Algorithm 1 stops, then $0 \in \partial_{\beta,\gamma}^K h(x^k)$, thus $x^k \in \operatorname{argmin}_K h$ by Lemma 2.3, i.e., x^k is solution to problem (COP).

(iii) The stopping criteria “ $0 \in \partial_{\beta,\gamma}^K h(x^k)$ ” can be replaced by “ $x^{k+1} = x^k$ ”. Indeed, if $x^{k+1} = x^k$, then $x^k = P_K(x^k - \alpha_k \xi^k)$, i.e., relation (2.1) yields

$$\langle -\alpha_k \xi^k, y - x^k \rangle \leq 0, \forall y \in K \iff \langle \xi^k, y - x^k \rangle \geq 0, \forall y \in K.$$

Since $\xi^k \in \partial_{\beta,\gamma}^K h(x^k)$ and $\langle \xi^k, y - x^k \rangle \geq 0$ for all $y \in K$, we have

$$\max\{h(y), h(x^k)\} \geq h(x^k) + \frac{\lambda}{2} \left(\gamma - \frac{\lambda}{\beta} - \lambda\gamma \right) \|y - x^k\|^2, \forall y \in K, \forall \lambda \in [0, 1].$$

Taking $0 < \lambda < \frac{\gamma\beta}{1+\gamma\beta} < 1$, we have

$$\max\{h(y), h(x^k)\} > h(x^k), \forall y \in K \setminus \{x^k\},$$

i.e., $x^k \in \operatorname{argmin}_K h$. Therefore, if $x^{k+1} = x^k$, then $x^k \in \operatorname{argmin}_K h$.

(iv) The projection step (3.13) is equivalent to the following regularized minimization problem

$$x^{k+1} = \operatorname{argmin} \left\{ \langle \alpha_k \xi^k, y \rangle + \frac{1}{2} \|y - x^k\|^2 : y \in K \right\}. \tag{3.14}$$

Indeed, (3.14) is equivalent to

$$x^{k+1} = \operatorname{argmin} \left\{ \langle \alpha_k \xi^k, y \rangle + \frac{1}{2} \|y - x^k\|^2 + \iota_K(y) : y \in \mathbb{R}^n \right\}.$$

Hence, using the optimality condition, we have

$$\begin{aligned} 0 \in \alpha_k \xi^k + x^{k+1} - x^k + N_K(x^{k+1}) &\iff (I + N_K)x^{k+1} = x^k - \alpha_k \xi^k \\ &\iff x^{k+1} = (I + N_K)^{-1}(x^k - \alpha_k \xi^k) \iff x^{k+1} = P_K(x^k - \alpha_k \xi^k), \end{aligned}$$

where the last equivalence holds due to [1, Example 23.4].

3.3 Convergence analysis

In order to prove the convergence of the sequence generated for Algorithm 1 under assumption (A), we consider the following set:

$$\Omega = \{x \in K : h(x) \leq h(x^k), \forall k \in \mathbb{N}\}. \quad (3.15)$$

Note that under assumption (A), $\operatorname{argmin}_K h$ is a singleton by [15, Corollary 3], thus $\Omega \neq \emptyset$.

The following corollary is an easy consequence of Proposition 3.2 which is a key for establishing the convergence analysis of subgradient methods.

Corollary 3.1 *Let $K \subseteq \mathbb{R}^n$ be a closed and convex set, $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be a function such that assumption (A) holds, $\{\alpha_k\}_k$ be a sequence of positive numbers, $\{x^k\}_k$ and $\{\xi^k\}_k$ be the sequences generated by Algorithm 1 and $\tilde{x} \in \Omega$. Then, for every $k \in \mathbb{N}$, we have*

$$\|x^{k+1} - \tilde{x}\|^2 \leq \|x^k - \tilde{x}\|^2 + \alpha_k^2 \|\xi^k\|^2 + 2\alpha_k \langle \xi^k, \tilde{x} - x^k \rangle, \quad (3.16)$$

$$\|x^{k+1} - \tilde{x}\|^2 \leq (1 - \gamma\beta\alpha_k) \|x^k - \tilde{x}\|^2 + \alpha_k^2 \|\xi^k\|^2. \quad (3.17)$$

Proof It follows immediately from relation (3.2) and Proposition 3.2(b) with $x^+ = x^{k+1}$, $z = x^k$, $\xi = \xi^k$, $x = \tilde{x}$ and $\alpha = \alpha_k$. \square

We also have the following.

Corollary 3.2 *Let $K \subseteq \mathbb{R}^n$ be a closed and convex set, $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be a function such that assumption (A) holds, $\{\alpha_k\}_k$ be a sequence of positive numbers, and $\{x^k\}_k$ and $\{\xi^k\}_k$ be the sequences generated by Algorithm 1. Then, for every $k \in \mathbb{N}$, we have*

$$\|x^{k+1} - x^k\| \leq \alpha_k \|\xi^k\|. \quad (3.18)$$

Proof It follows immediately from Proposition 3.1(a) with $x^+ = x^{k+1}$, $z = x^k$, $\xi = \xi^k$ and $\alpha = \alpha_k$. \square

Another result that will be useful for the convergence of the algorithm is the following:

Proposition 3.5 *Let $K \subseteq \mathbb{R}^n$ be a closed and convex set, $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be a function such that assumption (A) holds, $\{\alpha_k\}_k$ be a sequence of positive numbers and $\{x^k\}_k$ and $\{\xi^k\}_k$ be the sequences generated by Algorithm 1. If assumptions (Ci) with $i = 1, 2, 3$ holds, then*

(a) *for every $\tilde{x} \in \Omega$ we have*

$$\lim_{k \rightarrow \infty} \langle \xi^k, x^k - \tilde{x} \rangle = 0; \tag{3.19}$$

(b) $\sum_{k=0}^{\infty} \|x^{k+1} - x^k\|^2 < +\infty$. *In particular, $\lim_{k \rightarrow +\infty} \|x^{k+1} - x^k\| = 0$.*

Proof (a): Since $\xi^k \in \partial_{\beta, \gamma}^K h(x^k)$ for every $k \in \mathbb{N}$ and $\tilde{x} \in \Omega$, we have by Lemma 2.5 that

$$\langle \xi^k, x^k - \tilde{x} \rangle \geq \frac{\beta\gamma}{2} \|x^k - \tilde{x}\|^2 \geq 0.$$

Furthermore, it follows from relation (3.16) that

$$\sum_{k=0}^N \alpha_k \langle \xi^k, x^k - \tilde{x} \rangle \leq \frac{1}{2} \|x^0 - \tilde{x}\|^2 - \frac{1}{2} \|x^{N+1} - \tilde{x}\|^2 + \frac{M^2}{2} \sum_{k=0}^N \alpha_k^2.$$

Then, $\sum_{k=0}^{\infty} \alpha_k \langle \xi^k, x^k - \tilde{x} \rangle < +\infty$. Finally, by (C3), $\sum_{k=0}^{\infty} \alpha_k = +\infty$, thus $\lim_{k \rightarrow \infty} \langle \xi^k, x^k - \tilde{x} \rangle = 0$.

(b): It follows directly by summing up in (3.18) (after squaring both sides of the equality) taking into account (Ci) with $i = 1, 3$. □

Our main result, which shows that the sequences $\{x^k\}_k$, generated by Algorithm 1, converge to the optimal solution of problem (COP) under assumption (A), is given below.

Theorem 3.1 *Let $K \subseteq \mathbb{R}^n$ be a closed and convex set, $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be a function such that assumption (A) holds, $\{\alpha_k\}_k$ be a sequence of positive numbers and $\{x^k\}_k$ and $\{\xi^k\}_k$ be the sequences generated by Algorithm 1. Suppose that Assumption (Ci) with $i = 1, 2, 3$ holds. Then $\{x^k\}_k$ converges to $\{\bar{x}\} = \operatorname{argmin}_K h$. If moreover, h is continuous, $\lim_{k \rightarrow +\infty} h(x^k) = h(\bar{x}) = \min_K h$.*

Proof By assumption (C2), we have $\|\xi^k\| \leq M$ for every $k \in \mathbb{N}$. Hence, we have from (3.17) that for every $\tilde{x} \in \Omega$,

$$\|x^{k+1} - \tilde{x}\|^2 \leq (1 - \gamma\beta\alpha_k) \|x^k - \tilde{x}\|^2 + \alpha_k^2 \|\xi^k\|^2 \leq \|x^k - \tilde{x}\|^2 + M^2\alpha_k^2. \tag{3.20}$$

Furthermore, since $\sum_{k=0}^{\infty} \alpha_k^2 < +\infty$ by (C3), it follows from Lemma 2.6(a) that $\{x^k\}_k$ is quasi-Fejér and, as a consequence, $\{x^k\}_k$ is bounded, i.e., it has cluster points.

Let $\widehat{x} \in K$ a cluster point of $\{x^k\}_k$. Then there exists a subsequence $\{x^{k_j}\}_j \subseteq \{x^k\}_k$ such that $x^{k_j} \rightarrow \widehat{x}$ as $j \rightarrow +\infty$.

Since for every k , $\xi^k \in \partial_{\beta, \gamma}^K h(x^k)$, then for every $j \in \mathbb{N}$,

$$\begin{aligned} \max\{h(y), h(x^{k_j})\} &\geq h(x^{k_j}) + \frac{\lambda}{2} \left(\gamma - \frac{\lambda}{\beta} - \lambda\gamma \right) \|x^{k_j} - y\|^2 \\ &\quad + \frac{\lambda}{\beta} \langle \xi^{k_j}, y - x^{k_j} \rangle, \quad \forall y \in K. \end{aligned}$$

Take $y = \bar{x}$ (where $\{\bar{x}\} = \operatorname{argmin}_K h \subseteq \Omega$). Then

$$\langle \xi^{k_j}, \bar{x} - x^{k_j} \rangle \leq -\frac{\beta}{2} \left(\gamma - \frac{\lambda}{\beta} - \gamma\lambda \right) \|\bar{x} - x^{k_j}\|^2, \quad \forall \lambda \in]0, 1[,$$

which is equivalent to (taking $0 < \lambda < \frac{\gamma\beta}{1+\gamma\beta}$)

$$0 \leq \frac{\beta}{2} \left(\gamma - \frac{\lambda}{\beta} - \gamma\lambda \right) \|\bar{x} - x^{k_j}\|^2 \leq \langle \xi^{k_j}, x^{k_j} - \bar{x} \rangle. \tag{3.21}$$

Since $x^{k_j} \rightarrow \widehat{x}$ as $j \rightarrow +\infty$ and $\lim_{j \rightarrow \infty} \langle \xi^{k_j}, x^{k_j} - \bar{x} \rangle = 0$ by Proposition 3.5(i), taking the limit as $j \rightarrow +\infty$ in (3.21) we conclude that $\|\bar{x} - \widehat{x}\|^2 = 0$, i.e., $\bar{x} = \widehat{x}$. Hence, every cluster point of the sequence $\{x^k\}_k$ is a minimum of problem (COP), and since h is strongly quasiconvex by (A), $\operatorname{argmin}_K h$ is a singleton, hence the whole sequence $\{x^k\}_k$ converges to $\{\bar{x}\} = \operatorname{argmin}_K h$ by Lemma 2.6(b), and the proof is complete. The last statement of this theorem follows immediately from the fact that $\lim_{k \rightarrow +\infty} x^k = \bar{x}$, continuity of h and Lemma 2.1. \square

Next, we show a quasi-linear convergence rate of the iterates.

Corollary 3.3 *Let $K \subseteq \mathbb{R}^n$ be a closed and convex set, $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be a function such that assumption (A) holds, $\{\alpha_k\}_k$ be a sequence of positive numbers and $\{x^k\}_k$ and $\{\xi^k\}_k$ be the sequences generated by Algorithm 1. Suppose that Assumption (Ci) with $i = 1, 2, 3$ holds. Then, for $\{\bar{x}\} = \operatorname{argmin}_K h$, the sequence $\{x^k\}_k$ converges quasi linearly to \bar{x} , and*

$$\|x^k - \bar{x}\|^2 \leq \prod_{j=0}^{k-1} (1 - \gamma\beta\alpha_j) \|x^0 - \bar{x}\|^2 + \epsilon_k, \tag{3.22}$$

with $\lim_{k \rightarrow +\infty} \epsilon_k = 0$.

Proof From (3.17) and assumption (C2) it follows that for any $k \in \mathbb{N}$,

$$\|x^{k+1} - \bar{x}\|^2 \leq (1 - \gamma\beta\alpha_k) \|x^k - \bar{x}\|^2 + M^2\alpha_k^2.$$

Since $1 - \gamma\beta\alpha_k < 1$ and $\sum_k^\infty \alpha_k^2 < +\infty$ by assumption (C3), we show the first statement. Moreover, we obtain recursively

$$\begin{aligned} \|x^{k+1} - \bar{x}\|^2 &\leq (1 - \gamma\beta\alpha_k) \left[(1 - \gamma\beta\alpha_{k-1}) \|x^k - \bar{x}\|^2 + M^2\alpha_{k-1}^2 \right] + M^2\alpha_k^2 \\ &\vdots \\ &\leq \prod_{j=0}^k (1 - \gamma\beta\alpha_j) \|x^0 - \bar{x}\|^2 + M^2 \sum_{j=0}^k \alpha_j^2, \end{aligned}$$

which implies (3.22) with $\epsilon_k = \sum_{j=0}^{k-1} \alpha_{j-1}^2$. □

We finish the paper with the following remark.

Remark 3.4 In [6], the authors uses Plastria’s subdifferential [17] for quasiconvex Lipschitz continuous functions, which is completely different from our case with strongly quasiconvex functions with strong subdifferential because there is no relationship between strongly quasiconvex functions and quasiconvex Lipschitz continuous functions. Indeed, the function $h(x) = \sqrt{|x|}$ is strongly quasiconvex on any ball centered at the origin without being Lipschitz continuous while the function $h(x) = 0$ is quasiconvex and Lipschitz continuous without being strongly quasiconvex.

4 Conclusions

The main difference between Algorithm 1 with other subgradient methods based on generalized subdifferentials for quasiconvex functions (see [6, 8, 13, 14]) is that we use the strong subdifferential (see relation (2.8)). This theoretical difference which has many algorithmic consequences as, for instance, the strong subdifferential is compact at every point of the interior of the effective domain of the function (Lemma 2.2) while other generalized subdifferentials are unbounded (see [6, 8, 13, 14]). Furthermore, in [6, 8, 13, 14], the subgradient is taken from those unbounded subdifferentials and it needs to be normalized, thus the subgradient that they used belongs to the intersection with the unit ball in \mathbb{R}^n . In our case, Algorithm 1 can take any vector from the strong subdifferential and since the strong subdifferential was motivated by strongly quasiconvex functions, its geometrical structure will provide much more valuable information.

We added two examples to show the difference between the strong subdifferential and the intersection of the GP subdifferential and the unit ball. In these examples, we can see the advantages of using the strong subdifferential instead of the GP subdifferential.

Let $h : \mathbb{R}^n \rightarrow \mathbb{R}$ be a quasiconvex function and $x \in \mathbb{R}^n$. Then the Greenberg–Pierskalla (GP henceforth) of h at x is defined by

$$\partial^{GP} h(x) := \{\xi \in \mathbb{R}^n : \langle \xi, y - x \rangle < 0, \forall y \in S_{h(x)}^<(h)\},$$

while all the other subdifferentials used in [13] are variants of the previous one. Clearly, ∂^{GP} is unbounded (as well as its variants of [13] and the ones used in [6, 8, 14]).

Now, let us consider the following examples:

- (i) Let $K = [-1, 1]$ and $h : \mathbb{R} \rightarrow \mathbb{R}$ be given by $h(x) = \sqrt{|x|}$. h is strongly quasiconvex on K with modulus $\gamma = \frac{1}{2} > 0$. Take $x = 0$ and $\beta = 1$. Then, by [12, Remark 20] we have

$$\partial_{1, \frac{1}{2}}^{[-1, 1]} h(0) = \left[-\frac{3}{2}, \frac{3}{2}\right],$$

while $\partial^{GP} h(0) = \mathbb{R}$ and $\partial^{GP} h(0) \cap B(0, 1) = \{-1, 1\}$

- (ii) Let $h : \mathbb{R} \rightarrow \mathbb{R}$ be given by

$$h(x) = \begin{cases} 0, & \text{if } x = 0, \\ -\frac{1}{x}, & \text{if } 0 < x \leq 1, \\ +\infty, & \text{otherwise.} \end{cases}$$

Note that h is strongly quasiconvex with modulus $\gamma_h = 1$ (see [10]). Take $\bar{x} = 0$ and $\beta = 1$. Then by [12, Remark 6(ii)] we have

$$\partial_{1, 1}^{[-1, 1]} h(0) = \left] -\infty, -\frac{1}{2} \right],$$

while $\partial^{GP} h(0) = \left] -\infty, 0 \right[$ and $\partial^{GP} h(0) \cap B(0, 1) = \{-1\}$.

We observe that the strong subdifferential is smaller than the GP subdifferential and its variants for this class of functions. For instance, in case (i), where the function is continuous, the GP subdifferential is the entire space, which is not particularly informative, while the strong subdifferential is compact. In the more challenging case (ii), where the function is not even lower semicontinuous, both subdifferentials are unbounded. However, the strong subdifferential is smaller and has a positive distance from zero, which could prevent the algorithm from terminating prematurely due to approximation errors.

We hope this research offers new insights into subgradient-type methods for a broader range of quasiconvex functions. This includes nonlinear projected subgradient methods [3], conditional subgradient methods [7], incremental subgradient methods [8], and other approaches involving strongly quasiconvex functions and strong subdifferential.

Acknowledgements We dedicated our work to the first author of this manuscript, our formed student, colleague and friend MSc. Juan Choque, who passed away on August 13, 2024, after the submission of the file to the journal.

Author contributions All authors contributed equally to the study conception, design and implementation of the paper.

Funding This research was partially supported by postgraduate scholarship Apoyo Institucional para Magíster Académicos de la Universidad de Tarapacá, Arica, Chile (Choque), by ANID-Chile under project

Fondecyt Regular 1241040 (Lara) and by BASAL fund FB210005 for center of excellence from ANID-Chile (Marcavillaca).

Availability of supporting data No data sets were generated during the current study.

Declarations

Conflict of interest The authors declare no competing interests.

References

1. Bauschke, H.H., Combettes, P.L.: *Convex Analysis and Monotone Operators Theory in Hilbert Spaces*. CMS Books in Mathematics, 2nd edn. Springer, Berlin (2017)
2. Beck, A.: *First Order Methods in Optimization*. MOS-SIAM, Series on Optimization. SIAM, Philadelphia (2017)
3. Beck, A., Teboulle, M.: Mirror descent and nonlinear projected subgradient methods for convex optimization. *Oper. Res. Lett.* **31**, 167–175 (2003)
4. Burachik, R.S., Graña Drummond, L.M., Iusem, A., Svaiter, B.F.: Full convergence of the steepest descent method with inexact line search. *Optimization* **32**, 137–146 (1995)
5. Combettes, P.L.: Quasi-Fejerian analysis of some optimization algorithms. In: *Inherently parallel algorithms in feasibility and optimization and their applications* (Haifa, 2000), volume 8. *Stud. Comput. Math.* **2**, 115–152 (2001)
6. Cruz Neto, J.X., Lopes, J.O., Tragliai, M.V.: Algorithms for quasiconvex minimization. *Optimization* **60**, 1105–1117 (2011)
7. Hu, Y., Li, J., Yu, C.: Convergence rates of subgradient methods for quasi-convex optimization problems. *Comp. Optim. Appl.* **77**, 183–212 (2020)
8. Hu, Y., Wai Yu, C.K., Yang, X.: Incremental quasi-subgradient methods for minimizing the sum of quasi-convex functions. *J. Global Optim.* **75**, 1003–1028 (2019)
9. Hu, Y., Yang, X., Sim, C.K.: Inexact subgradient methods for quasi-convex optimization problems. *Eur. J. Oper. Res.* **240**, 315–327 (2015)
10. Jovanović, M.: On strong quasiconvex functions and boundedness of level sets. *Optimization* **20**, 163–165 (1989)
11. Jovanović, M.: A note on strongly convex and quasiconvex functions. *Math. Notes* **60**, 584–585 (1996)
12. Kabgani, A., Lara, F.: Strong subdifferentials: theory and applications in nonconvex optimization. *J. Global Optim.* **84**, 349–368 (2022)
13. Kiwiel, K.C.: Convergence and efficiency of subgradient methods for quasiconvex minimization. *Math. Prog. Ser. A* **99**, 1–25 (2001)
14. Konnov, I.V.: On convergence properties of a subgradient method. *Optim. Methods Softw.* **18**(1), 53–62 (2003)
15. Lara, F.: On strongly quasiconvex functions: existence results and proximal point algorithms. *J. Optim. Theory Appl.* **192**, 891–911 (2022)
16. Penot, J.P.: Are generalized derivatives useful for generalized convex functions? In: Crouzeix, J.P., et al. (eds.) *Generalized Convexity, Generalized Monotonicity*, pp. 3–60. Kluwer, Philadelphia (1998)
17. Plastria, F.: Lower subdifferentiable functions and their minimization by cutting planes. *J. Optim. Theory Appl.* **46**, 37–53 (1985)
18. Polyak, B.T.: Existence theorems and convergence of minimizing sequences in extremum problems with restrictions. *Soviet Math.* **7**, 72–75 (1966)
19. Rockafellar, R.T.: Generalized directional derivatives and subgradients of nonconvex functions. *Can. J. Math.* **32**, 257–280 (1980)
20. Vandenberghe, L., *Optimization Methods for Large-Scale Systems*. EE236C lecture notes, UCLA (2016)
21. Yen, L.H., Mu, L.D.: A parallel subgradient projection algorithm for quasiconvex equilibrium problems under the intersection of convex sets. *Optimization* **71**, 4447–4462 (2022)
22. Yu, C., Hu, Y., Yang, X., Choy, S.K.: Abstract convergence theorem for quasi-convex optimization 507 problems with applications. *Optimization* **68**, 1289–1304 (2018)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.