Commun. Optim. Theory 2025 (2025) 12 https://doi.org/10.23952/cot.2025.12

MATHRES

Communications in Optimization Theory

Available online at http://cot.mathres.org

INEXACT QUANTIZED QUASI-SUBGRADIENT METHOD FOR QUASI-CONVEX OPTIMIZATION PROBLEMS

CARISA KWOK WAI YU[1,*], YAOHUA HU[2]

[1]Department of Mathematics, Statistics and Insurance,
The Hang Seng University of Hong Kong, Shatin, Hong Kong
[2]School of Mathematical Sciences, Shenzhen University, Shenzhen 518060, China

Dedicated to Professor Elijah Lucien Polak for his remarkable contributions to the fields of unconstrained and constrained optimization, nonsmooth analysis, particularly in minimax and semi-infinite programming, and optimal control.

Abstract. Motivated by applications in distributed optimization, in this paper, we consider the nondifferentiable quantized quasi-convex constrained optimization problem and propose a quantized approximate quasi-subgradient method (QAQSGM). Each iteration of the QAQSGM consists of an inexact subgradient iteration and a quantization operator successively. The inexactness stems from computation errors and noise, which come from practical considerations and applications. Assuming that the computational errors and noise are deterministic and bounded, we investigate the convergence analysis and study the effect of the inexactness on the QAQSGM when the constraint set is compact or the objective function has a set of generalized weak sharp minima.

Keywords. Ssubgradient method; Quasi-convex optimization; Quantization operator; Noise; Weak sharp minima.

2020 Mathematics Subject Classification. 90C26, 49J52, 65K05.

*Corresponding author.
E-mail address: carisayu@hsu.edu.hk (C.K.W. Yu), mayhhu@szu.edu.cn (Y. Hu).
Received December 20, 2023; Accepted March 13, 2024.

1. INTRODUCTION

Mathematical optimization provides a unified framework for a wide variety of application problems in many disciplines, in which we usually consider a general constrained optimization problem

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & x \in X, \end{array} \tag{1.1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a real-valued function, and $X \subseteq \mathbb{R}^n$ is a nonempty, closed and convex constraint set. Convex optimization plays a fundamental role in mathematical optimization. However, convexity can become a too restrictive assumption for many real-life problems encountered in economics, finance and management science. For the latter problems, quasi-convex optimization can provide a much more accurate representation of reality, while still sharing some desirable properties enjoyed by convex optimization problems. Hence there is a significant increase of interest in quasi-convex optimization; see [3, 10, 13, 28] and references therein.

Popular first-order algorithms for solving (convex or quasi-convex) optimization problems are the so-called *projected subgradient methods*. The classical projected subgradient method was originally introduced by Polyak [25] and Ermoliev [11] in the 1970s to solve the nondifferentiable convex optimization problem (1.1) (i.e., with f in (1.1) assumed to be convex) and has the following iterative formula:

$$x_{k+1} := \mathbb{P}_X(x_k - v g_k), \tag{1.2}$$

where $\mathbb{P}_X(\cdot)$ denotes the Euclidean projection onto X, g_k is a (convex) subgradient of f at x_k, and v is a positive stepsize. Over the past five decades, various features and applications/extensions of projected subgradient methods have been established for convex optimization problems [7, 14, 23, 27]. Projected subgradient methods have also been extended and developed to solve nondifferentiable quasi-convex optimization problems (i.e., (1.1) with f being quasi-convex); see [15, 16, 18, 22, 30] and references therein.

Motivated by practical reasons, approximate subgradient methods (also called ε-subgradient methods) are widely studied for convex optimization problems [1, 23, 24, 27] and extended to quasi-convex optimization problems [16, 17, 18, 20].

The distributed optimization problem in networks usually requires the data at each node and transmitted data to reach a quantization level (see, e.g., [4, 21, 26]).

Motivated by applications, in this paper, we consider the following nondifferentiable quantized quasi-convex constrained optimization problem:

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & x \in X \cap \Lambda, \end{array} \tag{1.3}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a quasi-convex and continuous function, X is a closed and convex set, and Λ is a quantization lattice defined by

$$\Lambda := \Delta \mathbb{Z}^n = \{(\lambda_1 \Delta, \lambda_2 \Delta, \cdots, \lambda_n \Delta) : \lambda_i \in \mathbb{Z}\}, \tag{1.4}$$

where $\Delta > 0$ is the given quantization scalar. We denote the optimal solution set and the optimal value of problem (1.3) respectively by X^* and f_*, and we assume that X^* is nonempty and compact.

Inspired by the idea in [18], we propose a quantized approximate quasi-subgradient method (QAQSGM), which applies a quantization operator after the subgradient iteration along the approximate quasi-subgradient, to solve problem (1.3), and investigate the influence of inexact terms and convergence behavior on the QAQSGM. Considering a generic inexact subgradient algorithm for problem (1.3) and assuming the inexact terms are deterministic and bounded, we establish convergence results in two cases: (i) X is compact and (ii) f satisfies the generalized weak sharp minima condition. In this paper, we only consider the constant stepsize rule and obtain the best constant stepsize by minimizing the tolerance in approaching the optimal value.

The paper is organized as follows. In section 2, we present the notations and preliminary results. In section 3, we propose the QAQSGM for solving the quantized quasi-convex optimization problem (1.3) and demonstrate convergence properties of the QAQSGM when X is compact or when f satisfies a generalized weak sharp minima condition.

2. Notations and preliminary results

The notations used in the present paper are standard; see, e.g., [6]. We consider the n-dimensional Euclidean space $\mathbb{R}^n$ with inner product $\langle\cdot,\cdot\rangle$ and norm $\|\cdot\|$. For $x \in \mathbb{R}^n$ and $r > 0$, we use $\mathbb{B}(x,r)$ to denote the closed ball centered at x with radius r, and use $\mathbb{S}$ to denote the unit sphere centered at the origin.

For $x \in \mathbb{R}^n$ and $Z \subseteq \mathbb{R}^n$, the Euclidean distance of x from Z and the Euclidean projection of x onto Z are respectively defined by

$$\operatorname{dist}(x,Z) := \min_{z\in Z} \|x-z\| \quad \text{and} \quad \mathbb{P}_Z(x) := \arg\min_{z\in Z} \|x-z\|.$$

A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be quasi-convex if for each $x, y \in \mathbb{R}^n$ and each $\alpha \in [0,1]$, the following inequality holds

$$f((1-\alpha)x+\alpha y) \le \max\{f(x), f(y)\}.$$

For $\alpha \in \mathbb{R}$, the sublevel sets of f are denoted by

$$\operatorname{lev}_{<\alpha} f := \{x \in \mathbb{R}^n : f(x) < \alpha\} \quad \text{and} \quad \operatorname{lev}_{\le\alpha} f := \{x \in \mathbb{R}^n : f(x) \le \alpha\}.$$

It is well-known that f is quasi-convex if and only if $\operatorname{lev}_{<\alpha} f$ (and/or $\operatorname{lev}_{\le\alpha} f$) is convex for each $\alpha \in \mathbb{R}$.

The convex subdifferential $\partial f(x) := \{g \in \mathbb{R}^n : f(y) \ge f(x) + \langle g, y-x\rangle, \forall y \in \mathbb{R}^n\}$ might be empty for the quasi-convex functions. Hence, the introduction of (nonempty) subdifferential of quasi-convex functions plays an important role in quasi-convex optimization. Several different types of subdifferentials of quasi-convex functions have been introduced in the literature, see [2, 12, 18, 22] and references therein. In particular, Kiwiel [22] and Hu et al. [18] introduced a quasi-subdifferential and applied this quasi-subgradient in their proposed subgradient methods; see, e.g., [18, 19, 22].

Definition 2.1. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a quasi-convex function and let $\varepsilon > 0$. The quasi-subdifferential and ε-quasi-subdifferential of f at $x \in \mathbb{R}^n$ are respectively defined by*

$$\partial^* f(x) := \left\{ g \in \mathbb{R}^n : \langle g, y - x \rangle \leq 0, \forall y \in \mathrm{lev}_{<f(x)} f \right\},$$

and

$$\partial^*_\varepsilon f(x) := \left\{ g \in \mathbb{R}^n : \langle g, y - x \rangle \leq 0, \forall y \in \mathrm{lev}_{<f(x)-\varepsilon} f \right\}.$$

Any vector $g \in \partial^* f(x)$ or $g \in \partial^*_\varepsilon f(x)$ is called a quasi-subgradient or an ε-quasi-subgradient of f at x, respectively.

3. Quantized Approximate Quasi-Subgradient Method

For a given scalar $\Delta > 0$, the quantization lattice Λ in (1.4) consists of points regularly spaced by Δ along each coordinate axis. We define the quantization operator $Q : \mathbb{R}^n \to X \cap \Lambda$ by

$$\mathbb{Q}(\cdot) = \mathbb{P}_{X\cap\Lambda}(\mathbb{P}_X(\cdot)), \tag{3.1}$$

which projects its argument first onto the constraint set X and then onto the nearest lattice point in $X \cap \Lambda$. Note that applying the quantization operator $\mathbb{Q}(\cdot)$ to $x \in \mathbb{R}^n$ is not equivalent to directly projecting x to the nearest point in $X \cap \Lambda$. In particular, when $x \notin X$, the nearest point to x in $X \cap \Lambda$ can be different from $\mathbb{Q}(x)$, and may result in a large error (see [26]).

By using the quantization operator, we propose a quantized approximate quasi-subgradient method (QAQSGM) to solve problem (1.3) as follows.

Quantized approximate quasi-subgradient method (QAQSGM)
Select the stepsize v, an error sequence $\{\varepsilon_k\}$ and a noise sequence $\{r_k\}$, start with an initial point $x_0 \in X$, and generate a sequence $\{x_k\} \in X$ via the iteration

$$x_{k+1} = \mathbb{Q}(x_k - v\tilde{g}_k), \tag{3.2}$$

where the direction $\tilde{g}_k$ is an approximate quasi-subgradient of the following form

$$\tilde{g}_k := g_k + r_k \quad \text{with } g_k \in \partial^*_{\varepsilon_k} f(x_k) \cap \mathbb{S}. \tag{3.3}$$

The convergence analysis is divided into two cases: (i) X is compact, and (ii) f satisfies the generalized weak sharp minima condition.

3.1. Convergence Analysis for a Compact X.

In this subsection, we investigate the convergence property of the QAQSGM when X is compact. Throughout this subsection, the following three assumptions are made.

Assumption 1. *The constraint set X is compact.*

Assumption 2. *f satisfies the Hölder condition of order $p > 0$ with modulus $\mu > 0$ on $\mathbb{R}^n$, that is,*

$$f(x) - f_* \leq L\,\mathrm{dist}^p(x, X^*) \quad \textit{for all } x \in \mathbb{R}^n. \tag{3.4}$$

Assumption 3. *The noise and errors are bounded, i.e., there exist some scalars R and $\varepsilon \geq 0$ such that*

$$\|r_k\| \leq R, \forall k \geq 0 \quad \text{and} \quad \limsup_{k\to\infty} \varepsilon_k = \varepsilon.$$

Since the constraint set X is compact, all iterates are bounded. Hence there exists some $d>0$ (such as the diameter of X) such that

$$\|x_k - x\| \le d \quad \text{for all } x \in X \text{ and } k \ge 0. \tag{3.5}$$

Moreover under the bounded noise assumption, it follows from (3.3) that approximate quasi-subgradients are uniformly bounded, i.e.,

$$\|\tilde{g}_k\| \le 1+R \quad \text{for all } k \ge 0. \tag{3.6}$$

We start with the following lemmas that describe a very important property of the quantization operator and show the basic inequality of the subgradient iteration respectively.

Lemma 3.1 ([22, Lemma 6]). *If* $\mathbb{B}(\bar{x},\bar{r}) \subset \mathrm{cl}(\mathrm{lev}_{<f(x_k)-\varepsilon_k} f)$ *for some* $\bar{x} \in \mathbb{R}^n$ *and* $\bar{r} \ge 0$, *then* $\langle g_k, x_k - \bar{x}\rangle \ge \bar{r}$.

Lemma 3.2. *If Assumption 2 holds and* $f(x_k) > f_* + L\bar{r}^p + \varepsilon_k$ *holds for some* $\bar{r} \ge 0$, *then* $\langle g_k, x_k - x^*\rangle \ge \bar{r}$ *for all* $x^* \in X^*$.

Proof. Given $x^* \in X^*$, by the Hölder condition of order p and the assumption given in the lemma, for all $x \in \mathbb{B}(x^*,\bar{r})$, we have

$$f(x) - f_* \le L\mathrm{dist}^p(x,X^*) \le L\bar{r}^p < f(x_k) - f_* - \varepsilon_k,$$

which implies $\mathbb{B}(x^*,\bar{r}) \subseteq \mathrm{cl}(\mathrm{lev}_{<f(x_k)-\varepsilon_k} f)$. Hence, the conclusion follows from Lemma 3.1. ■

Lemma 3.3. *For all* $x \in \mathbb{R}^n$ *and* $y \in X$, *we have* $\|\mathbb{Q}(x) - y\| \le \|x-y\| + \sqrt{n}\Delta$.

Proof. Due to the structure of quantization lattice (1.4), for all $x \in \mathbb{R}^n$ and $y \in X$, we obtain

$$\|\mathbb{Q}(x) - y\| \le \|\mathbb{Q}(x) - \mathbb{P}_X(x)\| + \|\mathbb{P}_X(x) - y\| \le \sqrt{n}\Delta + \|x-y\|$$

thanks to the nonexpansive property of the projection operator. ■

Lemma 3.4. *Let* $\{x_k\}$ *be a sequence generated by the QAQSGM, and suppose Assumptions 1 and 3 hold.*

Then for all $x \in X$ *and* $k \in \mathbb{N}$, *we have*

$$\|x_{k+1} - x\|^2 \le \|x_k - x\|^2 - 2v\Big(\langle g_k, x_k - x\rangle - d\left(\frac{\sqrt{n}\Delta}{v} + R\right) - \frac{(\sqrt{n}\Delta + v(1+R))^2}{2v}\Big).$$

Proof. By relations (3.2)-(3.3) and Lemma 3.3, we have for all $x \in X$ that

$$\begin{aligned} \|x_{k+1} - x\|^2 &\le (\|x_k - v\tilde{g}_k - x\| + \sqrt{n}\Delta)^2 \\ &= \|x_k - v\tilde{g}_k - x\|^2 + 2\sqrt{n}\Delta\|x_k - v\tilde{g}_k - x\| + n\Delta^2. \end{aligned} \tag{3.7}$$

Note by (3.5) and (3.6) that

$$\begin{aligned} \|x_k - v\tilde{g}_k - x\|^2 &= \|x_k - x\|^2 - 2v\langle g_k + r_k, x_k - x\rangle + v^2\|g_k + r_k\|^2 \\ &\le \|x_k - x\|^2 - 2v\langle g_k, x_k - x\rangle + 2vRd + v^2(1+R)^2, \end{aligned}$$

and

$$\|x_k - v\tilde{g}_k - x\| \le \|x_k - x\| + v\|\tilde{g}_k\| \le d + v(1+R).$$

Combining the above two inequalities, (3.7) is reduced to the conclusion of this lemma. ■

The convergence result of the QAQSGM is demonstrated as follows.

Theorem 3.1. *Let $\{x_k\}$ be a sequence generated by the QAQSGM, and suppose Assumptions 1-3 hold.*

Then

$$\liminf_{k\to\infty} f(x_k) \le f_* + L\Big(d\left(\frac{\sqrt{n}\Delta}{v} + R\right) + \frac{(\sqrt{n}\Delta + v(1+R))^2}{2v}\Big)^p + \varepsilon. \tag{3.8}$$

Proof. By contradiction, we assume that

$$\liminf_{k\to\infty} f(x_k) > f_* + L\Big(d\left(\frac{\sqrt{n}\Delta}{v} + R\right) + \frac{(\sqrt{n}\Delta + v(1+R))^2}{2v}\Big)^p + \varepsilon,$$

that is, by Assumption 3, there exist some $\delta > 0$ and $k_0 \in \mathbb{N}$ such that

$$f(x_k) > f_* + L\Big(d\left(\frac{\sqrt{n}\Delta}{v} + R\right) + \frac{(\sqrt{n}\Delta + v(1+R))^2}{2v} + \delta\Big)^p + \varepsilon_k \quad \text{for each } k \ge k_0.$$

Thus it follows from Lemma 3.2 that for all $x^* \in X^*$ and $k \ge k_0$ there holds

$$\langle g_k, x_k - x^*\rangle \ge d\left(\frac{\sqrt{n}\Delta}{v} + R\right) + \frac{(\sqrt{n}\Delta + v(1+R))^2}{2v} + \delta.$$

Therefore from Lemma 3.4 with $x^* \in X^*$, we obtain

$$\|x_{k+1} - x^*\|^2 < \|x_k - x^*\|^2 - 2v\delta < \cdots < \|x_{k_0} - x^*\|^2 - 2(k - k_0 + 1)v\delta,$$

which yields a contradiction for sufficiently large k. The proof is complete. ■

For a given quantization scalar Δ, the best constant stepsize can be obtained by minimizing the tolerance estimated in (3.8), that is

$$v^* = \frac{\sqrt{n\Delta^2 + 2d\sqrt{n}\Delta}}{1+R}.$$

It is observed that the tolerance given in (3.8) has a same expression as that in [18, Theorem 3.1] if the quantization operator is infinitely precise (i.e., Δ is sufficiently small).

3.2. **Convergence Analysis for f with Generalized Weak Sharp Minima.** In this subsection, we consider the case when f satisfies the generalized weak sharp minima condition. The concept of weak sharp minima was introduced by Burke and Ferris [9], and has been extensively studied and widely used to analyze the convergence rates of many optimization algorithms; see [5, 8, 14, 29] and references therein. The generalized weak sharp minima condition was introduced in [18] to investigate the inexact subgradient method for quasi-convex optimization. In particular, we introduce the following two assumptions.

Assumption 4. *The function f satisfies the generalized weak sharp minima condition over X, that is, there exist some scalars $\eta > 0$, $q \geq p$ and a function $g : \mathbb{R}_+ \to \mathbb{R}_+$, satisfying $g(\cdot) \geq p$, $\sup_{t \geq 0} g(t) = q$ and $\lim_{t \to \infty} g(t) = p$, such that*

$$f(x) - f_* \geq \eta \operatorname{dist}^{g(\operatorname{dist}(x,X^*))}(x, X^*) \quad \text{for all } x \in X. \tag{3.9}$$

Assumption 5. *$\{r_k\}$ is a low level noise sequence, e.g., $R + \frac{\sqrt{n}\Delta}{v} < \left(\frac{\eta}{L}\right)^{\frac{1}{p}}$.*

Recall that L and p are scalars given in Assumption 2, and R and ε are scalars given in Assumption 3. For each $\theta \geq 0$ and $t > 0$, we define a new function $K_\theta^t : \mathbb{R}_+ \to \mathbb{R}$ by

$$K_\theta^t(y) := L\Big(\frac{(\sqrt{n}\Delta + v(1+R))^2}{2v} + \left(\frac{\sqrt{n}\Delta}{v} + R\right)\left(\frac{y}{\eta}\right)^{\frac{1}{t}}\Big)^p + \varepsilon + \theta - y, \tag{3.10}$$

The maximum solution of the inequality $K_\theta^{g(\operatorname{dist}(x,X^*))}(y) \geq 0$ over X is defined by

$$y_\theta^* := \sup\left\{y : K_\theta^{g(\operatorname{dist}(x,X^*))}(y) \geq 0 \text{ for some } x \in X\right\}. \tag{3.11}$$

Assumption 4 says that $p \leq g(\operatorname{dist}(x,X^*)) \leq q$ for all $x \in X$. Hence, from (3.10), we have

$$K_\theta^{g(\operatorname{dist}(x,X^*))}(y) \leq \max\left\{K_\theta^p(y), K_\theta^q(y)\right\} \quad \text{for all } y \geq 0, x \in X.$$

For the sake of simplicity, we write

$$y_\theta^p := \sup\left\{y : K_\theta^p(y) \geq 0\right\} \quad \text{and} \quad y_\theta^q := \sup\left\{y : K_\theta^q(y) \geq 0\right\}. \tag{3.12}$$

Hence, by applying (3.11) and Assumption 4, y_θ^* can be rewritten as

$$y_\theta^* = \max\left\{y_\theta^p, y_\theta^q\right\}. \tag{3.13}$$

Since $K_\theta^{g(\operatorname{dist}(x,X^*))}(0) > 0$ and $K_\theta^{g(\operatorname{dist}(x,X^*))}(y)$ is continuous on variable y for all $x \in X$, then y_θ^* is positive. Following a line of analysis similar to [18, Lemma 4.1], we achieve that y_θ^* is finite and continuous on parameter θ under Assumptions 4-5.

Lemma 3.5. *Suppose that Assumptions 4-5 hold. Then the following statements hold:*

(i) *y_θ^* is finite for all $\theta \geq 0$,*

(ii) $\lim_{\theta \to 0_+} y_\theta^* = y_0^*$.

To proceed the convergence property, we start with the following basic inequality.

Lemma 3.6. *Let $\{x_k\}$ be a sequence generated by the QAQSGM, and suppose Assumption 3 holds. Then, for all $x \in X$ and $k \in \mathbb{N}$, we have*

$$\|x_{k+1} - x\|^2 \leq \|x_k - x\|^2 - 2v\Big(\langle g_k, x_k - x\rangle - \|x_k - x\|\left(\frac{\sqrt{n}\Delta}{v} + R\right) - \frac{(\sqrt{n}\Delta + v(1+R))^2}{2v}\Big).$$

Proof. Note by (3.6) that

$$\|x_k - v\tilde{g}_k - x\|^2 \leq \|x_k - x\|^2 - 2v\langle g_k, x_k - x\rangle + 2vR\|x_k - x\| + v^2(1+R)^2,$$

and

$$\|x_k - v\tilde{g}_k - x\| \leq \|x_k - x\| + v(1+R).$$

Combining these inequalities and (3.7), we obtain the conclusion of this lemma. ■

Below we consider the boundedness of the sequence $\{x_k\}$.

Lemma 3.7. *Let $\{x_k\}$ be a sequence generated by the QAQSGM, and suppose Assumptions 2-5 hold. Then, $\{x_k\}$ is bounded.*

Proof. Since $\limsup_{k\to\infty} \varepsilon_k = \varepsilon$, for any $\theta > 0$, there exists some $k_0 \in \mathbb{N}$ such that

$$\varepsilon_k < \varepsilon + \theta \quad \text{for each } k \geq k_0. \tag{3.14}$$

Define

$$T := \sup\left\{ t \in \mathbb{R}_+ : t^{g(t)} \leq \frac{y_\theta^*}{\eta} \right\}, \tag{3.15}$$

which is finite, since y_θ^* is finite (cf. Lemma 3.5(i)) and $\lim_{t\to\infty} t^{g(t)} = +\infty$. Next, we claim that the following inequality holds for all $i \geq k_0$:

$$\operatorname{dist}(x_i, X^*) \leq \max\{\operatorname{dist}(x_{k_0}, X^*), T + v(1+R) + \sqrt{n}\Delta\}. \tag{3.16}$$

It is obvious that the relation (3.16) holds if $i = k_0$. Proving by the induction, we assume the relation (3.16) holds for some $i = k$ $(\geq k_0)$ and consider the following two cases.

Case 1. If $f(x_k) \leq f_* + L\left(\frac{(\sqrt{n}\Delta + v(1+R))^2}{2v} + \left(\frac{\sqrt{n}\Delta}{v} + R\right)\left(\frac{f(x_k)-f_*}{\eta}\right)^{\frac{1}{g(\operatorname{dist}(x_k,X^*))}}\right)^p + \varepsilon_k$, together with (3.10) and (3.14), we have $K_\theta^{g(\operatorname{dist}(x_k,X^*))}(f(x_k) - f_*) \geq 0$. Hence from (3.11), we obtain $f(x_k) - f_* \leq y_\theta^*$, and thus by (3.9) that $\operatorname{dist}(x_k,X^*)^{g(\operatorname{dist}(x_k,X^*))} \leq \frac{y_\theta^*}{\eta}$. It follows from (3.15) that $\operatorname{dist}(x_k,X^*) < T$, and thus by Lemma 3.3 that

$$\operatorname{dist}(x_{k+1}, X^*) \leq \operatorname{dist}(x_k - v\tilde{g}_k, X^*) + \sqrt{n}\Delta < T + v(1+R) + \sqrt{n}\Delta;$$

that is, the relation (3.16) holds for $i = k+1$.

Case 2. If $f(x_k) > f_* + L\left(\frac{(\sqrt{n}\Delta + v(1+R))^2}{2v} + \left(\frac{\sqrt{n}\Delta}{v} + R\right)\left(\frac{f(x_k)-f_*}{\eta}\right)^{\frac{1}{g(\operatorname{dist}(x_k,X^*))}}\right)^p + \varepsilon_k$, then, it follows from Lemma 3.2 and (3.9)that

$$\langle g_k, x_k - x^*\rangle \geq \frac{(\sqrt{n}\Delta + v(1+R))^2}{2v} + \left(\frac{\sqrt{n}\Delta}{v} + R\right)\operatorname{dist}(x_k, X^*).$$

Hence, applying Lemma 3.6 with $x^* = \mathbb{P}_{X^*}(x_k)$, we obtain

$$\operatorname{dist}^2(x_{k+1}, X^*) \leq \|x_{k+1} - x^*\|^2 \leq \operatorname{dist}^2(x_k, X^*).$$

Hence, the relation (3.16) holds for $i = k+1$.

Therefore by induction, the relation (3.16) holds for all $i \geq k_0$. Since the right hand side of (3.16) is finite and X^* is compact, hence $\{x_k\}$ is bounded. The proof is complete. ■

From Lemma 3.7, $\{x_k\}$ is bounded and hence $\{f(x_k)\}$ is bounded from above by the Hölder condition. The upper bound on $\{f(x_k)\}$ is denoted by M in what follows.

Theorem 3.2. *Let $\{x_k\}$ be a sequence generated by the QAQSGM, and suppose Assumptions 2-5 hold. Then* $\liminf_{k\to\infty} f(x_k) \le f_* + y_0^*$.

Proof. The finiteness of y_0^* has been proved in Lemma 3.5(i). To prove the convergence property, we first show that

$$\liminf_{k\to\infty} f(x_k) < f_* + y_\theta^* \quad \text{for all } \theta > 0 \tag{3.17}$$

Proving by contradiction, we assume that the following inequality holds for some $\theta > 0$,

$$\liminf_{k\to\infty} f(x_k) \ge f_* + y_\theta^*.$$

Thus, there exist some $0 < \delta < \min\left\{\frac{\theta}{2}, y_\theta^*\right\}$ and $k_0 \in \mathbb{N}$ such that

$$f(x_k) > f_* + y_\theta^* - \delta \quad \text{and} \quad \varepsilon_k < \varepsilon + \frac{\theta}{2} \tag{3.18}$$

for all $k \ge k_0$, where the second inequality holds due to $\limsup_{k\to\infty} \varepsilon_k = \varepsilon$.

From (3.11) and (3.18), we obtain $f(x_k) - f_* + \delta > \sup\{y : K_\theta^{g(\mathrm{dist}(x_k,X^*))}(y) \ge 0\}$ and thus $K_\theta^{g(\mathrm{dist}(x_k,X^*))}(f(x_k) - f_* + \delta) < 0$. Hence we get by (3.10) and (3.18) that

$$\begin{aligned} f(x_k) &> f_* + L\left(\frac{(\sqrt{n}\Delta + v(1+R))^2}{2v} + \left(\frac{\sqrt{n}\Delta}{v} + R\right)\left(\frac{f(x_k) - f_* + \delta}{\eta}\right)^{\frac{1}{g(\mathrm{dist}(x_k,X^*))}}\right)^p + \varepsilon_k \\ &\ge f_* + L\left(\frac{(\sqrt{n}\Delta + v(1+R))^2}{2v} + \left(\frac{\sqrt{n}\Delta}{v} + R\right)\left(\frac{f(x_k) - f_*}{\eta}\right)^{\frac{1}{g(\mathrm{dist}(x_k,X^*))}} + \delta'\right)^p + \varepsilon_k, \end{aligned}$$

where the second inequality follows from the Taylor expansion with some positive scalar $\delta' = \left(\frac{\sqrt{n}\Delta}{v} + R\right)\min\left\{\frac{\delta}{\eta q}\left(\frac{y_\theta^*}{\eta}\right)^{\frac{1}{q}-1}, \frac{\delta}{\eta p}\left(\frac{M - f_*}{\eta}\right)^{\frac{1}{p}-1}\right\}$. Therefore, by using Lemmas 3.2 and 3.6, we obtain

$$\langle g_k, x_k - x^* \rangle \ge \frac{(\sqrt{n}\Delta + v(1+R))^2}{2v} + \left(\frac{\sqrt{n}\Delta}{v} + R\right)\mathrm{dist}(x_k, X^*) + \delta',$$

and hence

$$\mathrm{dist}^2(x_{k+1}, X^*) \le \mathrm{dist}^2(x_k, X^*) - 2v\delta' \le \cdots \le \mathrm{dist}^2(x_0, X^*) - 2(k - k_0 + 1)v\delta',$$

which yields a contradiction for sufficiently large k. Thus (3.17) is proved. Taking the limit on (3.17) as $\theta \to 0$ and applying Lemma 3.5(ii), we arrive at the conclusion. ∎

We now give an explicit expression for the tolerance in approaching f_* in Theorem 3.2 in a specific case of p and $g(t)$. By solving (3.12) and (3.13), we obtain the following corollary where the total error is given in an explicit expression.

Corollary 3.1. *Let $\{x_k\}$ be a sequence generated by the QAQSGM, and suppose Assumptions 2-5 hold with $g(t) \equiv p$ and $p = 1$. Then*

$$\liminf_{k\to\infty} f(x_k) \le f_* + \eta \frac{L(\sqrt{n}\Delta + v(1+R))^2 + 2v\varepsilon}{2(v\eta - vLR - L\sqrt{n}\Delta)}.$$

Proof. By the assumptions, since $g(t) \equiv p$ and $p = q = 1$, we have

$$K_0^p(y) = K_0^q(y) = L\Big(\frac{(\sqrt{n}\Delta + v(1+R))^2}{2v} + \left(\frac{\sqrt{n}\Delta}{v} + R\right)\frac{y}{\eta}\Big) + \varepsilon - y.$$

It is clear that $K_0^p(y)$ is linear and decreasing due to $R + \frac{\sqrt{n}\Delta}{v} < \frac{\eta}{L}$. Thus by (3.12), y_0^p is just the solution of $K_0^p(y) = 0$. Thus from (3.13), we have

$$y_0^* = y_0^p = \eta \frac{L(\sqrt{n}\Delta + v(1+R))^2 + 2v\varepsilon}{2(v\eta - vLR - L\sqrt{n}\Delta)}.$$

Hence by Theorem 3.2, we arrive at the conclusion. ■

The total error given in Corollary 3.1 has a same expression as that of [18, Corollary 4.1] if the quantization operator is infinitely precise (i.e., Δ is sufficiently small).

Remark 3.1. *It is worth mentioning that Zaslavski investigated the convergence properties of the inexact quasi-subgradient method in* [31, Chapter 5]*, in which the inexact terms consist of the inexact quasi-subgradient and the inexact projection. The QAQSGM can be understood of a type of inexact quasi-subgradient method, while the convergence analysis obtained in this paper is difference from the ones in* [31]*. In particular, the convergence complexity of inexact quasi-subgradient method was investigated in* [31, Theorem 5.4] *under the coercive assumption, while our Theorem 3.2 established the convergence property of the QAQSGM under the assumption of generalized weak sharp minima.*

Acknowledgements

The first author is supported in part by the Research Grants Council of the Hong Kong Special Administrative Region, China (UGC/FDS14/P02/21, UGC/FDS14/P07/22). The second author is supported in part by the National Natural Science Foundation of China (12222112, 12071306, 32170655), Project of Educational Commission of Guangdong Province (2023ZDZX1017), Shenzhen Science and Technology Program (RCJC20221008092753082), and Research Team Cultivation Program of Shenzhen University (2023QNT011).

REFERENCES

[1] A. Auslender, M. Teboulle, Interior Gradient and ε-Subgradient Descent Methods for Constrained Convex Minimization, Mathematics of Operations Research, 29 (2004), 1-26.

[2] D. Aussel, A. Danilidis, Normal characterization of the main classes of quasiconvex functions, Set-Valued Analysis, 8 (2000), 219-236.

[3] M. Avriel, W.E. Diewert, S. Schaible, I. Zang, Generalized Concavity, Plenum Press, New York, 1988.

[4] T.C. Aysal, M.J. Coates, M.G. Rabbat, Distributed average consensus with dithered quantization, IEEE Trnsaction on Signal Processing, 56 (2008), 4905-4918.
[5] J. Bao, C.K.W. Yu, J. Wang, Y. Hu, J.-C. Yao, Modified inexact Levenberg–Marquardt methods for solving nonlinear least squares problems, Computational Optimization and Applications, 74 (2019), 547-582.
[6] D.P. BErtsekas, Nonlinear Programming, Athena Scientific, Cambridge, 1999.
[7] D.P. Bertsekas, Convex Optimization Algorithms, Athena Scientific, Massachuesetts, 2015.
[8] J. Bolte, T.P. Nguyen, J. Peypouquet, B.W. Suter, From Error bounds to the complexity of first-order descent methods for convex fundctions, Mathematical Programming, 165 (2017), 471-507.
[9] J.V. Burke, M.C. Ferris, Weak sharp minima in mathematical programming, SIAM Journal on Control and Optimization, 31 (1993), 1340-1359.
[10] J.-P. Crouzeix, J.-E. Martinez-Legaz, M. Volle, Generalized Convexity, Generalized Monotonicity, Kluwer Academic Publishers, Dordrecht, 1998.
[11] Y.M. Ermoliev, Methods of solution of nonlinear extremal problems, Cybernetics and Systems Analysis, 2 (1966), 1-14.
[12] H.J. Greenberg, W.P. Pierskalla, Quasiconjugate functions and surrogate duality, Cahiers Centre Études Recherche Opertionnelle, 15 (1973), 437-448.
[13] N. Hadjisavvas, S. Komlósi, S. Schaible, Handbook of Generalized Convexity and Generalized Monotonicity, Springer-Verlag, New York, 2005.
[14] Y. Hu, C. Li, X. Yang, On convergence rates of linearized proximal algorithms for convex composite optimization with applications, SIAM Journal on Optimization, 26 (2016), 1207-1235.
[15] Y. Hu, G. Li, C.K.W. Yu, T.L. Yip, Quasi-convex feasibility problems: Subgradient methods and convergence rates, European Journal of Operational Research, 298 (2022), 45-58.
[16] Y. Hu, J. Li, C.K.W. Yu, Convergenece rates of subgradient methods for quasi-convex optimization problems, Computational Optimization and Applications, 77 (2020), 183-212.
[17] Y. Hu, Y. Liu, M. Li, The Effect of Deterministic Noise on Quasi-subgradient Method for Quasi-convex Feasibility Problems, Journal of Applied and Numerical Optimization, 2 (2020), 235-247.
[18] Y. Hu, X. Yang, C.-K. Sim, Inexact subgradient methods for quasi-convex optimization problems, European Journal of Operational Research, 240 (2015), 315-327.
[19] Y. Hu, C.K.W. Yu, C. Li, X. Yang, Conditional subgradient methods for constrained quasi-convex optimization problems, Journal of Nonlinear and Convex Analysis, 17 (2016), 2143-2158.
[20] Y. Hu, C.K.W. Yu, X. Yang, Incremental quasi-subgradient methods for minimizing the sum of quasi-convex functions, Journal of Global Optimization, 75 (2019), 1003-1028.
[21] A. Kashyap, T. Basar, R. Srikant, Quantized consensus, Automatica, 43 (2007), 1192-1203.
[22] K.C. Kiwiel, Convergence and efficiency of subgradient methods for quasiconvex minimization, Mathematical Programming, 90 (2001), 1-25.
[23] K.C. Kiwiel, Convergence of approximate and incremental subgradient methods for convex optimization, SIAM Journal on Optimization, 14 (2004), 807-840.
[24] A. Nedić, D.P. Bertsekas, The effect of deterministic noise in subgradient methods, Mathematical Programming, 125 (2010), 75-99.
[25] B.T. Polyak, A general method for solving extremum problems, Soviet Mathematics Doklady, 8 (1967), 593-597.
[26] M.G. Rabbat, R.D. Nowak, Quantized incremental algorithms for distributed optimization, IEEE Journal on Selected Areas in Communication, 23 (2005), 798-808.
[27] N.Z. Shor, Minimization Methods for Non-differentiable Functions, Springer-Verlag, New York, 1985.
[28] I.M. Stancu-Minasian, Fractional Programming, Kluwer Academic Publishers, Dordrecht, 1997.
[29] J. Wang, Y. Hu, C. Li, J.-C. Yao, Linear convergence of CQ algorithms and applications in gene regulatory network inference, Inverse Problems, 33 (2017), 055017.

[30] C.K.W. Yu, Y. Hu, X. Yang, S.K. Choy, Abstract convergence theorem for quasi-convex optimization problems with applications, Optimization, 68 (2019), 1289-1304.
[31] A.J. Zaslavski, The projected subgradient algorithm in convex optimization, Springer, Switzerland, 2020.