

Weakly Quasar-Convex Optimization

Abbas Khademi[✉] * Felipe Lara[✉] †

October 7, 2026

Abstract

This paper introduces the class of weakly quasar-convex functions, a notion that simultaneously generalizes weak convexity (and hence convexity) and quasar-convexity. We establish the theoretical foundation for this new class by developing its basic calculus, deriving characterizations, and systematically investigating its closure and structural properties. Building on these tools, we analyze the proximal point algorithm for minimizing a weakly quasar-convex function under local quadratic growth, proving Q -linear convergence of the iterates and R -linear convergence of the function values to a global minimizer, with iteration complexity $O(\ln(\varepsilon^{-1}))$; we also give a fully computable, minimizer-free stopping rule via the proximal residual with matching complexity. Finally, we construct a weakly quasar-convex function, satisfying all our standing assumptions, that is neither weakly convex, quasi-convex, star quasi-convex, nor quasar-convex for any modulus, showing that our convergence theory applies to instances outside every prior framework in the literature.

Keywords: Nonconvex optimization; Quasar-convexity; Moreau envelope; Proximal point method; Structured nonconvexity.

Mathematics Subject Classification: 90C26, 49J52, 49M37.

1 Introduction

Many of the most important problems in modern optimization give rise to highly nonconvex objectives, including deep neural network training [26], low-rank matrix recovery [25, 35], phase retrieval [28], robust statistics [23], quantum process tomography [34], and dynamical systems [10, 15]. Since arbitrary nonconvexity rules out general-purpose guarantees, attention has turned to structured nonconvexity, which is weaker than convexity yet strong enough to support rigorous analysis of stochastic gradient methods [37], momentum schemes [5, 22], and proximal point methods [13]. Identifying which structural assumptions yield which guarantees remains an active area of research [4]. This paper contributes to this effort by introducing a unified framework that bridges the prominent notions of weak convexity and quasar-convexity.

Identifying tractable problem classes between convexity and full nonconvexity is a central question in optimization. One prominent example is star-convexity, in which a global minimizer remains visible from every point of the domain [31]. A further relaxation, quasar-convexity (originally called weak quasi-convexity), was proposed in the study of linear dynamical systems [15]. For a quasar-convex function, the negative gradient always points toward a region of smaller objective value, and this observation has led to near-optimal algorithms for star-convex functions and their extensions [16]. In recent years, algorithmic developments based on quasar-convexity have attracted considerable attention [1, 11, 19, 30, 32, 40]; see the recent survey [18] for a broad taxonomy of quasar-convexity and related concepts.

Weak convexity occupies a different corner of this landscape. Dating back to early work in nonsmooth optimization [33], it is characterized by the fact that adding a sufficiently large quadratic penalty to the objective renders it convex. This permits a controlled amount of nonconvexity while preserving enough structure for proximal and subgradient methods, together with smoothing techniques for composite and stochastic problems, to retain rigorous convergence guarantees [7, 12, 21, 29, 38, 41].

Weak convexity and quasar-convexity describe different forms of structure in nonconvex optimization. Weak convexity controls nonconvexity through a uniform quadratic curvature bound, whereas quasar-convexity imposes a geometric condition relative to a global minimizer. These complementary viewpoints

*Department of Decision Sciences, HEC Montréal, Montréal, H3T 2A7, Québec, Canada.

E-mail: abbaskhademi92@gmail.com; abbas.khademi@hec.ca

†Instituto de Alta Investigación (IAI), Universidad de Tarapacá, Arica, Chile.

E-mail: felipelaraobreque@gmail.com; flarao@academicos.uta.cl. Web: felipelara.cl

motivate a common framework that accommodates both types of structure. We introduce *weakly quasar-convex* functions through a parametrized inequality anchored at a global minimizer. This class includes weakly convex functions and quasar-convex functions, while also admitting functions outside both classes. We investigate its structural properties and identify conditions under which it supports convergence of the proximal point algorithm.

Our contributions are both theoretical and algorithmic and can be summarized as follows. First, we introduce the class of weakly quasar-convex functions, defined via a parametrized inequality anchored at a global minimizer that simultaneously generalizes weak convexity and quasar-convexity, and we prove that this generalization is strict via explicit examples that fail to be weakly convex, quasi-convex, or quasar-convex. Second, we develop the basic calculus of this new class, together with first-order characterizations, and show that, under a local quadratic growth condition, weak quasar-convexity rules out spurious stationary points and implies a restricted secant inequality (also called star quasiconvexity in [20]). Furthermore, we prove that the Moreau envelope preserves weak quasar-convexity, with an explicit modulus formula, yielding coercivity and characterization results for the associated proximity operator. Finally, we analyze the proximal point algorithm under local quadratic growth, proving Q -linear convergence of the iterates and R -linear convergence of the function values to a global minimizer; we illustrate this theory on an example lying outside every prior framework in the literature, where no other existing convergence result applies.

The remainder of the paper is organized as follows. Section 2 recalls notation and preliminary background. Section 3 introduces weakly quasar-convex functions and develops their theoretical foundations, detailing their basic calculus, characterizations, preservation under the Moreau envelope, and the properties of the associated proximity operator. Section 4 establishes convergence of the proximal point algorithm under local quadratic growth. Section 5 concludes the paper.

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. We denote $\mathbb{R}_+ := [0, +\infty[$, $\mathbb{R}_{++} :=]0, +\infty[$, $\overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$, and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$.

Given any $x, y, z \in \mathbb{R}^n$ and any $\beta \in \mathbb{R}$, the following relations hold

$$\langle x - z, y - x \rangle = \frac{1}{2}\|z - y\|^2 - \frac{1}{2}\|x - z\|^2 - \frac{1}{2}\|y - x\|^2, \quad (1)$$

$$\|\beta x + (1 - \beta)y\|^2 = \beta\|x\|^2 + (1 - \beta)\|y\|^2 - \beta(1 - \beta)\|x - y\|^2. \quad (2)$$

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. The set $\text{epi}(h) := \{(x, t) \in \mathbb{R}^n \times \mathbb{R} : h(x) \leq t\}$ is the epigraph of h , and $\text{argmin}_{\mathbb{R}^n} h$ is 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 $\{x^k\}_k \subseteq \mathbb{R}^n$ with $x^k \rightarrow \bar{x}$, $h(\bar{x}) \leq \liminf_{k \rightarrow +\infty} h(x^k)$.

For $h : \mathbb{R}^n \rightarrow \mathbb{R}$, $\mathcal{K} \subseteq \mathbb{R}^n$, $\bar{x} \in \mathcal{K}$, and $\zeta > 0$, the ζ -level set of h over $\mathcal{K}$ at $\bar{x}$ is

$$\mathcal{L}_\zeta^h(\mathcal{K}, \bar{x}) := \{x \in \mathcal{K} : h(x) \leq h(\bar{x}) + \zeta\}.$$

A function h satisfies a μ -local Quadratic Growth (QG) condition with parameter $\zeta > 0$ if

$$h(x) - h(\bar{x}) \geq \frac{\mu}{2}\|x - \bar{x}\|^2, \quad \forall x \in \mathcal{L}_\zeta^h(\mathbb{R}^n, \bar{x}). \quad (3)$$

In the literature, there exist several interesting examples of functions that satisfy the (local) quadratic growth condition (e.g., [39, Example E.5]).

Let $h : \mathbb{R}^n \rightarrow \mathbb{R}$ be a differentiable function and let $\bar{x} \in \text{argmin}_{\mathbb{R}^n} h$. A point $x \in \mathbb{R}^n$ is called a spurious stationary point of h (with respect to $\bar{x}$) if $\nabla h(x) = 0$ but $x \neq \bar{x}$.

Let us recall the following generalized convexity notions, which have been introduced and developed in virtue of its interesting properties in machine learning and continuous optimization (see [14, 16, 31, 42]).

Let $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be a proper function with $\bar{x} \in \text{argmin}_{\mathbb{R}^n} h$. Then h is said to be:

- (i) (Strongly) star-convex [31] if there exists $\rho \geq 0$ such that for every $y \in \text{dom}(h)$ and every $\lambda \in [0, 1]$, we have

$$h(\lambda \bar{x} + (1 - \lambda)y) \leq \lambda h(\bar{x}) + (1 - \lambda)h(y) - \lambda(1 - \lambda)\frac{\rho}{2}\|y - \bar{x}\|^2.$$

(ii) (Strongly) quasar-convex [14, 16] if there exists $\gamma \in]0, 1]$ and $\rho \geq 0$ such that for every $y \in \text{dom}(h)$, we have

$$h(\lambda \bar{x} + (1 - \lambda)y) \leq \lambda \gamma h(\bar{x}) + (1 - \lambda \gamma)h(y) - \lambda \left(1 - \frac{\lambda}{2 - \gamma}\right) \frac{\gamma \rho}{2} \|y - \bar{x}\|^2.$$

Every (strongly) star-convex function is (strongly) quasar-convex, while the reverse statement does not hold in general (see [8, Example 8] with $0 < p < 1$).

Furthermore, it follows from the definition that $\text{argmin}_{\mathbb{R}^n} h$ is a singleton when h is strongly ($\rho > 0$) quasar-convex (resp. star-convex). Moreover, if $\gamma = 1$, then a quasar-convex function reduces to a strongly star-convex function ($\rho > 0$) and star-convex function ($\rho = 0$).

In the differentiable case, we have the following useful characterization for (strongly) quasar-convex functions. Let $\mathcal{K} \subseteq \mathbb{R}^n$ be a convex and open set and $h : \mathbb{R}^n \rightarrow \mathbb{R}$ be a differentiable function with $\text{argmin}_{\mathcal{K}} h \neq \emptyset$. Then h is (strongly) quasar-convex with modulus $\gamma \in]0, 1]$ and $\rho \geq 0$ on $\mathcal{K}$ if and only if (see [16, Lemma 10])

$$h(\bar{x}) \geq h(y) + \frac{1}{\gamma} \langle \nabla h(y), \bar{x} - y \rangle + \frac{\rho}{2} \|y - \bar{x}\|^2, \quad \forall y \in \mathcal{K}. \quad (4)$$

Given a nonempty, closed, and convex set $\mathcal{K} \subset \mathbb{R}^n$, the proximity operator on $\mathcal{K}$ of parameter $\beta > 0$ of a proper lower semicontinuous function $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ at $z \in \mathbb{R}^n$ is defined as the operator $\text{Prox}_{\beta h}(\mathcal{K}, \cdot) : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ for which

$$\text{Prox}_{\beta h}(\mathcal{K}, z) := \text{argmin}_{x \in \mathcal{K}} \left\{ h(x) + \frac{1}{2\beta} \|z - x\|^2 \right\}. \quad (5)$$

When $\mathcal{K} = \mathbb{R}^n$, we simply write $\text{Prox}_{\beta h}(z) := \text{Prox}_{\beta h}(\mathbb{R}^n, z)$. If h is proper, lower semicontinuous, and convex, then $\text{Prox}_{\beta h}$ turns out to be a single-valued operator (see, for instance, [2, Proposition 12.15]). Closely related is the *Moreau envelope* $h^\beta : \mathbb{R}^n \rightarrow \mathbb{R}$ of h with parameter $\beta > 0$, defined by

$$h^\beta(x) := \inf_{y \in \mathbb{R}^n} \left\{ h(y) + \frac{1}{2\beta} \|x - y\|^2 \right\}. \quad (6)$$

Let us recall some notions related to the Clarke directional derivative [6]. Let $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be proper and locally Lipschitz around $x \in \text{dom}(h)$. Then the Clarke directional derivative of h at x in direction $d \in \mathbb{R}^n$ is defined by

$$h^\circ(x; d) = \limsup_{\substack{y \rightarrow x \\ t \downarrow 0}} \frac{h(y + td) - h(y)}{t},$$

and the Clarke subdifferential of h at x is given by

$$\partial^C h(x) := \{v \in \mathbb{R}^n : \langle v, d \rangle \leq h^\circ(x; d), \quad \forall d \in \mathbb{R}^n\},$$

By convention, $\partial^C h(x) := \emptyset$ for $x \notin \text{dom}(h)$. Moreover,

$$h^\circ(x; d) := \max_{\xi \in \partial^C h(x)} \langle \xi, d \rangle, \quad \forall d \in \mathbb{R}^n. \quad (7)$$

A point $x \in \text{dom}(h)$ is called a Clarke critical point of h if $0 \in \partial^C h(x)$.

Definition 1. Let $\{x^k\}_k \subset \mathbb{R}^n$ be a sequence converging to $x^* \in \mathbb{R}^n$. It is said that:

(a) $\{x^k\}_k$ converges *Q-linearly* to x^* if there exists a constant $r \in]0, 1[$ such that for all sufficiently large k , we have

$$\|x^{k+1} - x^*\| \leq r \|x^k - x^*\|. \quad (8)$$

(b) $\{x^k\}_k$ converges *R-linearly* to x^* if there exists a constant $r \in]0, 1[$ and a constant $C > 0$ such that for all sufficiently large k , we have

$$\|x^k - x^*\| \leq C r^k. \quad (9)$$

Clearly, *Q-linear convergence implies R-linear convergence*, but the converse is not true in general. $\square$

3 Weak Quasar-Convexity

In this section, we introduce and develop the notions of weak quasar-convexity, a notion that extends weak convexity and quasar-convexity, simultaneously.

Definition 2 (Weakly Quasar-Convex). *A function $h : \mathbb{R}^n \rightarrow \mathbb{R}$ is said to be (ρ, γ) -weakly quasar-convex with respect to $\bar{x} \in \operatorname{argmin}_{\mathbb{R}^n} h$ if for every $x \in \mathbb{R}^n$ and every $\lambda \in [0, 1]$, we have*

$$h(\lambda\bar{x} + (1 - \lambda)x) \leq \gamma\lambda h(\bar{x}) + (1 - \gamma\lambda)h(x) + \frac{\gamma\rho}{2}\lambda\left(1 - \frac{\lambda}{2 - \gamma}\right)\|x - \bar{x}\|^2, \quad (10)$$

where $\gamma \in]0, 1]$ and $\rho \geq 0$. $\square$

Definition 2 can be extended to any proper function $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, with $\bar{x} \in \operatorname{argmin}_{\mathbb{R}^n} h$, once (10) is read under the usual conventions of extended arithmetic. In that setting, the inequality is automatically satisfied (and imposes no constraint) whenever $x \notin \operatorname{dom}(h)$, while for $x \in \operatorname{dom}(h)$ it forces $\lambda\bar{x} + (1 - \lambda)x \in \operatorname{dom}(h)$ for every $\lambda \in [0, 1]$; that is, weak quasar-convexity forces $\operatorname{dom}(h)$ to be star-shaped with respect to $\bar{x}$.

In particular, when $\gamma = 1$, Definition 2 reduces to ρ -weak star-convexity, that is,

$$h(\lambda\bar{x} + (1 - \lambda)x) \leq \lambda h(\bar{x}) + (1 - \lambda)h(x) + \frac{\rho}{2}\lambda(1 - \lambda)\|x - \bar{x}\|^2, \quad \forall x \in \mathbb{R}^n, \quad \forall \lambda \in [0, 1]. \quad (11)$$

These definitions generalize standard weak convexity by anchoring the relaxation to a specific minimizer $\bar{x}$ rather than requiring the property to hold for all pairs of points. When $\rho = 0$ and $\gamma = 1$, definitions recover standard star-convexity with respect to $\bar{x}$. When $\rho = 0$, Definition 2 connects to quasar-convexity concepts from recent optimization literature [8, 16, 18, 42]. Then the following relationship holds

$$\begin{array}{ccc} \text{Convex} & \xrightarrow{\star} & \text{Quasar-Convex} \\ \downarrow & & \downarrow \\ \text{Weakly Convex} & \xrightarrow{\star} & \text{Weakly Quasar-Convex} \end{array}$$

where “ $\star$ ” denotes that $\operatorname{argmin}_{\mathbb{R}^n} h \neq \emptyset$ is required.

All the reverse implications do not hold in general. Indeed, examples of quasar-convex and weakly convex functions which are not convex are well-known (see [16, 31]). It remains to present functions which is weakly quasar-convex without being neither quasar-convex nor weakly convex for any modulus.

To that end, we first present a function which is weakly quasar-convex, but it is not quasar-convex for any $\gamma \in]0, 1]$.

Example 3. Let $\gamma = \frac{1}{2}$, $\rho = \frac{5}{2}$, and define the function $h : \mathbb{R} \rightarrow \mathbb{R}$ by

$$h(x) = x^6 - 2x^4 + \frac{11}{10}x^2 = x^2 \left((x^2 - 1)^2 + \frac{1}{10} \right) \geq 0. \quad (12)$$

Clearly, $h(x) \geq 0$ for all $x \in \mathbb{R}$, $h(0) = 0$, and $h(x) \rightarrow +\infty$ as $|x| \rightarrow \infty$, hence $\bar{x} = 0$ is a global minimizer of h on $\mathbb{R}$.

Furthermore, h is $(\frac{5}{2}, \frac{1}{2})$ -weakly quasar-convex with respect to $\bar{x} = 0$. With $\bar{x} = 0$ and $h(\bar{x}) = 0$, the condition (10) becomes, for all $x \in \mathbb{R}$ and $\lambda \in [0, 1]$,

$$h((1 - \lambda)x) \leq \left(1 - \frac{\lambda}{2}\right) h(x) + \frac{5\lambda}{8} \left(1 - \frac{2\lambda}{3}\right) x^2, \quad (13)$$

where we used $\gamma = \frac{1}{2}$, $\rho = \frac{5}{2}$, $\frac{\gamma\rho}{2} = \frac{5}{8}$, and $\frac{\lambda}{2 - \gamma} = \frac{2\lambda}{3}$. Define the gap function

$$\Delta(\lambda, x) := \left(1 - \frac{\lambda}{2}\right) h(x) + \frac{5\lambda}{8} \left(1 - \frac{2\lambda}{3}\right) x^2 - h((1 - \lambda)x).$$

We must show $\Delta(\lambda, x) \geq 0$ for all $x \in \mathbb{R}$ and $\lambda \in [0, 1]$. Indeed, substituting $h(x) = x^6 - 2x^4 + \frac{11}{10}x^2$, we decompose

$$\Delta(\lambda, x) = A_6(\lambda)x^6 + A_4(\lambda)x^4 + A_2(\lambda)x^2,$$

where the coefficients for all $\lambda \in [0, 1]$ are

$$A_6(\lambda) := \left(1 - \frac{\lambda}{2}\right) - (1 - \lambda)^6 = -\lambda^6 + 6\lambda^5 - 15\lambda^4 + 20\lambda^3 - 15\lambda^2 + \frac{11}{2}\lambda \in \left[0, \frac{1}{2} + \frac{5}{12\sqrt[5]{12}}\right],$$

$$A_4(\lambda) := -2\left(1 - \frac{\lambda}{2}\right) + 2(1 - \lambda)^4 = 2\lambda^4 - 8\lambda^3 + 12\lambda^2 - 7\lambda \in \left[-\frac{11}{8}, 0\right],$$

$$A_2(\lambda) := \frac{11}{10}\left(1 - \frac{\lambda}{2}\right) + \frac{5\lambda}{8}\left(1 - \frac{2\lambda}{3}\right) - \frac{11}{10}(1 - \lambda)^2 = -\frac{91}{60}\lambda^2 + \frac{91}{40}\lambda \in \left[0, \frac{273}{320}\right].$$

Since $x^2 \geq 0$, we can view $\Delta(\lambda, x)$ as a cubic polynomial in $u = x^2 \geq 0$,

$$\Delta(\lambda, x) = u(A_6(\lambda)u^2 + A_4(\lambda)u + A_2(\lambda)).$$

To show $\Delta(\lambda, x) \geq 0$, it suffices to prove that the quadratic $Q(u) = A_6(\lambda)u^2 + A_4(\lambda)u + A_2(\lambda)$ is non-negative for all $u \geq 0$ and all $\lambda \in [0, 1]$. Since $A_6(\lambda) \geq 0$ and $A_4(\lambda) \leq 0$ for all $\lambda \in [0, 1]$, consider the quadratic

$$Q(u) = A_6(\lambda)u^2 + A_4(\lambda)u + A_2(\lambda), \quad \forall u \geq 0.$$

Case $\lambda = 0$. Here $A_6(\lambda) = A_4(\lambda) = A_2(\lambda) = 0$, so $Q \equiv 0$.

Case $\lambda \in]0, 1]$. Then $A_6(\lambda) > 0$, so Q is an upward-opening parabola with vertex at

$$u^* = -\frac{A_4(\lambda)}{2A_6(\lambda)} \geq 0,$$

because $A_4 \leq 0$. Thus the global minimum of Q over $\mathbb{R}$ already occurs at $u^* \geq 0$, and

$$Q(u) \geq 0 \quad \forall u \geq 0 \iff \min_{\mathbb{R}} Q = Q(u^*) = A_2(\lambda) - \frac{A_4(\lambda)^2}{4A_6(\lambda)} \geq 0 \iff A_4(\lambda)^2 \leq 4A_6(\lambda)A_2(\lambda). \quad (14)$$

Hence it suffices to prove $A_4(\lambda)^2 \leq 4A_6(\lambda)A_2(\lambda)$ for all $\lambda \in [0, 1]$. Direct expansion gives, for all λ , that $A_6(\lambda) = \lambda B_6(\lambda)$, $A_4(\lambda) = \lambda B_4(\lambda)$ and $A_2(\lambda) = \lambda B_2(\lambda)$, where

$$B_2(\lambda) = \frac{91}{120}(3 - 2\lambda), \quad B_4(\lambda) = -7 + 12\lambda - 8\lambda^2 + 2\lambda^3, \quad \text{and} \quad B_6(\lambda) = \frac{11}{2} - 15\lambda + 20\lambda^2 - 15\lambda^3 + 6\lambda^4 - \lambda^5.$$

For $\lambda = 0$, (14) is trivial. For $\lambda > 0$, (14) is equivalent to

$$B_4(\lambda)^2 \leq 4B_6(\lambda)B_2(\lambda), \quad \forall \lambda \in [0, 1]. \quad (15)$$

Expanding both sides,

$$\begin{aligned} B_4(\lambda)^2 &= 4\lambda^6 - 32\lambda^5 + 112\lambda^4 - 220\lambda^3 + 256\lambda^2 - 168\lambda + 49, \\ 4B_6(\lambda)B_2(\lambda) &= \frac{91}{15}\lambda^6 - \frac{91}{2}\lambda^5 + \frac{728}{5}\lambda^4 - \frac{1547}{6}\lambda^3 + 273\lambda^2 - \frac{2548}{15}\lambda + \frac{1001}{20}. \end{aligned}$$

Define

$$P(\lambda) := 60(4B_6(\lambda)B_2(\lambda) - B_4(\lambda)^2) = 124\lambda^6 - 810\lambda^5 + 2016\lambda^4 - 2270\lambda^3 + 1020\lambda^2 - 112\lambda + 63.$$

Since (15) is equivalent to $P(\lambda) \geq 0$, it remains to show that

$$P(\lambda) \geq 0, \quad \forall \lambda \in [0, 1]. \quad (16)$$

Write P in the Bernstein basis [9] of degree 6 on $[0, 1]$,

$$P(\lambda) = \sum_{k=0}^6 b_k \binom{6}{k} \lambda^k (1 - \lambda)^{6-k}.$$

The coefficients b_k are obtained from the power-basis coefficients a_i of P via the standard conversion

$$b_k = \sum_{i=0}^k \binom{k}{i} a_i, \quad a_0 = 63, \quad a_1 = -112, \quad a_2 = 1020, \quad a_3 = -2270, \quad a_4 = 2016, \quad a_5 = -810, \quad \text{and} \quad a_6 = 124.$$

Carrying out this computation gives $b_0 = 63$, $b_1 = \frac{133}{3}$, $b_2 = \frac{281}{3}$, $b_3 = \frac{195}{2}$, $b_4 = \frac{1151}{15}$, $b_5 = \frac{155}{3}$ and $b_6 = 31$. Since the Bernstein basis polynomials $\binom{6}{k} \lambda^k (1 - \lambda)^{6-k}$ are non-negative for $\lambda \in [0, 1]$, we conclude

$$P(\lambda) = \sum_{k=0}^6 b_k \binom{6}{k} \lambda^k (1 - \lambda)^{6-k} \geq 0, \quad \forall \lambda \in [0, 1],$$

which establishes (16).

On the other hand, h is not γ -quasar-convex at $\bar{x} = 0$ for any $\gamma \in]0, 1]$. Indeed, if h is γ -quasar-convex at $\bar{x} = 0$ (with $h(\bar{x}) = 0$), then for all $x \in \mathbb{R}$ and all $\lambda \in [0, 1]$, we should have $h((1 - \lambda)x) \leq (1 - \gamma\lambda)h(x)$. In particular, take $x = 1$ and $\lambda = \frac{1}{2}$. Then, $h(\frac{1}{2}) \leq (1 - \frac{\gamma}{2})h(1)$. Substituting $h(1) = \frac{32}{320}$ and $h(\frac{1}{2}) = \frac{53}{320}$, we get

$$\frac{53}{320} \leq (1 - \frac{\gamma}{2}) \frac{32}{320} \implies 53 \leq 32 - 16\gamma \implies \gamma \leq -\frac{21}{16},$$

a contradiction. Therefore, h is not quasar-convex for any $\gamma \in]0, 1]$. $\square$

The following proposition will be useful.

Proposition 4. *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a function with $0 \in \operatorname{argmin}_{\mathbb{R}} f$, $f(0) = 0$, and let $g : \mathbb{R}^n \rightarrow \mathbb{R}$ be a function such that $1 \leq g \leq M$ for some constant $M > 0$. Define,*

$$h(x) := \begin{cases} f(\|x\|) g\left(\frac{x}{\|x\|}\right), & x \in \mathbb{R}^n \setminus \{0\} \\ 0, & x = 0 \end{cases}$$

If f is (ρ, γ) -weakly quasar-convex with respect to 0, with $\rho \geq 0$ and $\gamma \in]0, 1]$, then h is $(M\rho, \gamma)$ -weakly quasar-convex with respect to 0.

Proof. First, we verify that $0 \in \operatorname{argmin}_{\mathbb{R}^n} h$. Since $f(0) = 0$ and $0 \in \operatorname{argmin}_{\mathbb{R}} f$, we have $f(t) \geq 0$ for all $t \geq 0$. Also, $g \geq 1 > 0$. Hence, for any $x \in \mathbb{R}^n$, we have

$$h(x) = f(\|x\|) g\left(\frac{x}{\|x\|}\right) \geq 0 = h(0),$$

with equality if and only if $f(\|x\|) = 0$. Thus, $0 \in \operatorname{argmin}_{\mathbb{R}^n} h$.

Now, fix $y \in \mathbb{R}^n$, $\lambda \in [0, 1]$, and let $\bar{x} = 0$. Since $h(0) = 0$ and $\bar{x} = 0$,

$$h(\lambda\bar{x} + (1 - \lambda)y) = h((1 - \lambda)y) = f((1 - \lambda)\|y\|) g\left(\frac{y}{\|y\|}\right).$$

Since f is (ρ, γ) -weakly quasar-convex with respect to 0, we have

$$f((1 - \lambda)\|y\|) \leq (1 - \gamma\lambda)f(\|y\|) + \frac{\gamma\rho}{2}\lambda\left(1 - \frac{\lambda}{2 - \gamma}\right)\|y\|^2.$$

Multiplying by $g(y/\|y\|) \geq 0$ and using $g \geq 1$, we obtain

$$\begin{aligned} h((1 - \lambda)y) &= f((1 - \lambda)\|y\|) g\left(\frac{y}{\|y\|}\right) \\ &\leq (1 - \gamma\lambda)f(\|y\|) g\left(\frac{y}{\|y\|}\right) + \frac{\gamma\rho}{2}\lambda\left(1 - \frac{\lambda}{2 - \gamma}\right)\|y\|^2 g\left(\frac{y}{\|y\|}\right) \\ &= (1 - \gamma\lambda)h(y) + \frac{\gamma\rho}{2}\lambda\left(1 - \frac{\lambda}{2 - \gamma}\right)\|y\|^2 g\left(\frac{y}{\|y\|}\right). \end{aligned}$$

Since $1 \leq g(y/\|y\|) \leq M$, we have $g\left(\frac{y}{\|y\|}\right) \leq M$. Thus,

$$h((1 - \lambda)y) \leq (1 - \gamma\lambda)h(y) + \frac{\gamma(M\rho)}{2}\lambda\left(1 - \frac{\lambda}{2 - \gamma}\right)\|y\|^2.$$

This is exactly the $(M\rho, \gamma)$ -weakly quasar-convex inequality for h with respect to 0. Therefore, h is $(M\rho, \gamma)$ -weakly quasar-convex with respect to 0. $\square$

Finally, we present the desired example.

Example 5. (Strictly Large) *Let $f : [-1, 1] \rightarrow \mathbb{R}$ be the function given by*

$$f(s) = 2 - \sqrt{|s|},$$

and g be the function of Example 3, i.e., $g(s) = s^2((s^2 - 1)^2 + \frac{1}{10}) \geq 0$.

Clearly, f is positive on $[-1, 1]$, but f is not weakly convex on $[-1, 1]$ for any modulus, since the one-sided derivatives at $s = 0$ are $\{\pm\infty\}$. Also, since $(s^2 - 1)^2 + \frac{1}{10} \geq \frac{1}{10} > 0$ for all s , we have $g(s) = 0$ if and only if $s = 0$, so 0 is the unique global minimizer of g on $\mathbb{R}$. By Example 3, g is $(\frac{5}{2}, \frac{1}{2})$ -weakly quasar-convex with respect to 0, and g is not γ -quasar-convex with respect to 0 for any $\gamma \in]0, 1]$. Set

$$\mathcal{A} := \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \geq 0, -x_1 \leq x_2 \leq x_1\},$$

and define the function $h : \mathcal{A} \rightarrow \mathbb{R}$ by

$$h(x_1, x_2) := \begin{cases} g(x_1) f\left(\frac{x_2}{x_1}\right), & (x_1, x_2) \neq (0, 0), \\ 0, & (x_1, x_2) = (0, 0). \end{cases} \quad (17)$$

Since $g \geq 0$ and $f > 0$ on $[-1, 1]$, we have $h \geq 0$ on $\mathcal{A}$. Moreover, for $(x_1, x_2) \in \mathcal{A}$ with $x_1 > 0$, $h(x_1, x_2) = g(x_1)f(x_2/x_1) > 0$ since $g(x_1) > 0$ and $f > 0$, and for $x_1 = 0$, the domain constraint of $\mathcal{A}$ forces $x_2 = 0$. Hence $\bar{x} = (0, 0)$ is the unique global minimizer of h on $\mathcal{A}$, i.e., $\{(0, 0)\} = \operatorname{argmin}_{\mathcal{A}} h$. Moreover, along the ray $x_2 = 0$, $h(t, 0) = f(0)g(t) = 2g(t) \rightarrow 0$ as $t \downarrow 0$, and more generally, since f is bounded on $[-1, 1]$ and $g(x_1) \rightarrow g(0) = 0$ as $x_1 \downarrow 0$,

$$h(x_1, x_2) = g(x_1) f\left(\frac{x_2}{x_1}\right) \rightarrow 0$$

uniformly over $x_2/x_1 \in [-1, 1]$ as $x_1 \downarrow 0$, thus h is continuous at the origin. We claim that h is weakly quasar-convex with respect to $(0, 0)$, but is neither weakly convex nor quasi-convex nor quasar-convex on $\mathcal{A}$ for any modulus. Indeed, let us prove this point by point:

(i) h is weakly quasar-convex with respect to $(0, 0)$: On any ray through the origin, the ratio x_2/x_1 is constant, say $x_2/x_1 = c \in [-1, 1]$, so restricted to that ray, $h(t, ct) = f(c) g(t)$ is a positive scalar multiple of g , with $f(c) \in [1, 2]$. Since g is $(\frac{5}{2}, \frac{1}{2})$ -weakly quasar-convex with respect to 0, multiplying (10) (with $g(0) = 0$) by the positive constant $f(c)$ preserves $\gamma = \frac{1}{2}$ while scaling ρ by $f(c)$, yielding

$$h((1 - \lambda)x_1, (1 - \lambda)x_2) \leq (1 - \frac{\lambda}{2}) h(x_1, x_2) + \frac{\frac{1}{2} \cdot f(c) \cdot \frac{5}{2}}{2} \lambda (1 - \frac{2\lambda}{3}) x_1^2.$$

Since $x = (x_1, cx_1)$, we have $\|x\|^2 = x_1^2(1 + c^2) \geq x_1^2$. Using $x_1^2 \leq \|x\|^2$ together with $f(c) \leq 2$, we get $f(c) \cdot \frac{5}{2} \leq 5$, which gives

$$h((1 - \lambda)x) \leq (1 - \frac{\lambda}{2}) h(x) + \frac{\frac{1}{2} \cdot 5}{2} \lambda (1 - \frac{2\lambda}{3}) \|x\|^2.$$

Thus, h is $(5, \frac{1}{2})$ -weakly quasar-convex with respect to $(0, 0)$ on $\mathcal{A}$.

(ii) h is not weakly convex on $\mathcal{A}$ for any modulus: Fix $\rho \geq 0$ and consider, for $s \in [0, 1]$, the points $(1, -s), (1, s) \in \mathcal{A}$ and their midpoint $\frac{1}{2}(1, -s) + \frac{1}{2}(1, s) = (1, 0)$. Since f is even, $h(1, -s) = h(1, s) = g(1)f(s)$, and $h(1, 0) = g(1)f(0)$. Weak convexity with modulus ρ would require

$$g(1)f(0) \leq g(1)f(s) + \frac{\rho}{2} \cdot \frac{1}{4}(2s)^2 = g(1)f(s) + \frac{\rho}{2}s^2,$$

i.e., using $f(0) - f(s) = \sqrt{s}$, we get $g(1)\sqrt{s} \leq \frac{\rho}{2}s^2$.

Since $g(1) = 1 - 2 + \frac{11}{10} = \frac{1}{10} > 0$, dividing by $\sqrt{s} > 0$ gives $\frac{1}{10} \leq \frac{\rho}{2}s^{3/2}$, which fails as $s \downarrow 0$ for any fixed $\rho \geq 0$. Hence, h is not ρ -weakly convex on $\mathcal{A}$ for any $\rho \geq 0$.

(iii) h is not quasi-convex, not star quasi-convex, and not quasar-convex on $\mathcal{A}$ for any modulus: It suffices to consider h restricted to the half-line $(t, 0)$, $t \geq 0$, where $h(t, 0) = f(0)g(t) = 2g(t)$. Using $g(0) = 0$, $g(1) = \frac{1}{10} = \frac{32}{320}$, and $g(\frac{1}{2}) = \frac{53}{320}$ (see Example 3), we get

$$h(0, 0) = 0, \quad h(1, 0) = \frac{1}{5} = \frac{32}{160}, \quad \text{and} \quad h(\frac{1}{2}, 0) = \frac{53}{160}.$$

Since $\frac{53}{160} > \frac{32}{160}$, the point $(\frac{1}{2}, 0)$, which lies on the segment joining $(0, 0)$ and $(1, 0)$, satisfies $h(\frac{1}{2}, 0) > \max\{h(0, 0), h(1, 0)\}$, which directly violates quasi-convexity and, in particular, star quasi-convexity with respect to $(0, 0)$ (see [20, Theorem 12]).

Furthermore, h is not γ -quasar-convex with respect to $(0, 0)$ for any $\gamma \in]0, 1]$. Indeed, taking $x = (1, 0)$ and $\lambda = \frac{1}{2}$ in the quasar-convexity inequality $h((1 - \lambda)x) \leq (1 - \gamma\lambda)h(x)$ reduces, after cancelling the common factor of 2 from $h(t, 0) = 2g(t)$, to $g(\frac{1}{2}) \leq (1 - \frac{\gamma}{2})g(1)$, i.e., $\frac{53}{320} \leq (1 - \frac{\gamma}{2})\frac{32}{320}$, which is exactly the inequality shown in Example 3 to force $\gamma \leq -\frac{21}{16}$, a contradiction.

Therefore, h is weakly quasar-convex on $\mathcal{A}$ with respect to $(0, 0)$ (indeed $(5, \frac{1}{2})$ -weakly quasar-convex), yet neither weakly convex, nor quasi-convex, nor star quasi-convex, nor quasar-convex, for any modulus. $\square$

3.1 Basic Properties

We now establish several fundamental properties of weakly quasar-convex functions.

Proposition 6. h is (ρ, γ) -weakly quasar-convex with respect to $\bar{x} \in \operatorname{argmin}_{\mathbb{R}^n} h$ (where $\rho \geq 0$, $\gamma \in]0, 1[$) if and only if for every $d \in \mathbb{R}^n$ with $\|d\| = 1$, the function

$$\varphi_d(t) := h(\bar{x} + td), \quad t \geq 0, \quad (18)$$

is (ρ, γ) -weakly quasar-convex with respect to $t = 0$ on $[0, +\infty[$, i.e., for every $t \geq 0$ and every $\lambda \in [0, 1]$, we have

$$\varphi_d((1 - \lambda)t) \leq \gamma \lambda \varphi_d(0) + (1 - \gamma \lambda) \varphi_d(t) + \frac{\gamma \rho}{2} \lambda \left(1 - \frac{\lambda}{2 - \gamma}\right) t^2. \quad (19)$$

Proof. ($\Rightarrow$) Fix d with $\|d\| = 1$. For any $t \geq 0$ and $\lambda \in [0, 1]$, define $x = \bar{x} + td$. Then,

$$\varphi_d((1 - \lambda)t) = h(\bar{x} + (1 - \lambda)td) = h(\lambda \bar{x} + (1 - \lambda)(\bar{x} + td)) = h(\lambda \bar{x} + (1 - \lambda)x).$$

Applying the (ρ, γ) -weak quasar-convexity of h at $\bar{x}$ and since $\|x - \bar{x}\| = \|td\| = t$, we get

$$\begin{aligned} h(\lambda \bar{x} + (1 - \lambda)x) &\leq \gamma \lambda h(\bar{x}) + (1 - \gamma \lambda) h(x) + \frac{\gamma \rho}{2} \lambda \left(1 - \frac{\lambda}{2 - \gamma}\right) \|x - \bar{x}\|^2 \\ &= \gamma \lambda \varphi_d(0) + (1 - \gamma \lambda) \varphi_d(t) + \frac{\gamma \rho}{2} \lambda \left(1 - \frac{\lambda}{2 - \gamma}\right) t^2, \end{aligned}$$

i.e., (19) holds.

($\Leftarrow$) Let $x \in \mathbb{R}^n$ be arbitrary. If $x = \bar{x}$, the weak quasar-convexity inequality holds trivially (both sides equal $h(\bar{x})$). Otherwise, define $d = \frac{x - \bar{x}}{\|x - \bar{x}\|}$ and $t = \|x - \bar{x}\| > 0$, hence $\|d\| = 1$ and $x = \bar{x} + td$. Then, for any $\lambda \in [0, 1]$,

$$h(\lambda \bar{x} + (1 - \lambda)x) = h(\bar{x} + (1 - \lambda)td) = \varphi_d((1 - \lambda)t).$$

Applying relation (19) to φ_d gives

$$\begin{aligned} \varphi_d((1 - \lambda)t) &\leq \gamma \lambda \varphi_d(0) + (1 - \gamma \lambda) \varphi_d(t) + \frac{\gamma \rho}{2} \lambda \left(1 - \frac{\lambda}{2 - \gamma}\right) t^2 \\ &= \gamma \lambda h(\bar{x}) + (1 - \gamma \lambda) h(x) + \frac{\gamma \rho}{2} \lambda \left(1 - \frac{\lambda}{2 - \gamma}\right) \|x - \bar{x}\|^2, \end{aligned}$$

i.e., h is (ρ, γ) -weakly quasar-convex at $\bar{x}$. $\square$

Proposition 7. Let $h_i : \mathbb{R}^n \rightarrow \mathbb{R}$ be (ρ_i, γ) -weakly quasar-convex with respect to $\bar{x} \in \operatorname{argmin}_{\mathbb{R}^n} h_i$ for every $i = 1, \dots, m$, where $\gamma \in]0, 1[$ is the same for all i . Then for every $\alpha_i \geq 0$, the function $h := \sum_{i=1}^m \alpha_i h_i$ is $(\sum_{i=1}^m \alpha_i \rho_i, \gamma)$ -weakly quasar-convex with respect to $\bar{x}$.

Proof. Since $\bar{x} \in \cap_i \operatorname{argmin}_{\mathbb{R}^n} h_i$ and $\alpha_i \geq 0$, it follows that $\bar{x} \in \operatorname{argmin}_{\mathbb{R}^n} h$. Then for every $x \in \mathbb{R}^n$ and every $\lambda \in [0, 1]$, we have

$$\begin{aligned} h(\lambda \bar{x} + (1 - \lambda)x) &= \sum_{i=1}^m \alpha_i h_i(\lambda \bar{x} + (1 - \lambda)x) \\ &\leq \sum_{i=1}^m \alpha_i \left(\gamma \lambda h_i(\bar{x}) + (1 - \gamma \lambda) h_i(x) + \frac{\gamma \rho_i}{2} \lambda \left(1 - \frac{\lambda}{2 - \gamma}\right) \|x - \bar{x}\|^2 \right) \\ &= \gamma \lambda \sum_{i=1}^m \alpha_i h_i(\bar{x}) + (1 - \gamma \lambda) \sum_{i=1}^m \alpha_i h_i(x) + \frac{\gamma}{2} \lambda \left(1 - \frac{\lambda}{2 - \gamma}\right) \|x - \bar{x}\|^2 \sum_{i=1}^m \alpha_i \rho_i \\ &= \gamma \lambda h(\bar{x}) + (1 - \gamma \lambda) h(x) + \frac{\gamma}{2} \left(\sum_{i=1}^m \alpha_i \rho_i \right) \lambda \left(1 - \frac{\lambda}{2 - \gamma}\right) \|x - \bar{x}\|^2. \end{aligned}$$

Therefore, h is (ρ, γ) -weakly quasar-convex at $\bar{x}$ with modulus $\rho = \sum_{i=1}^m \alpha_i \rho_i$ and $\gamma \in]0, 1[$. $\square$

The following calculus rule will be useful.

Proposition 8. Let $h : \mathbb{R}^n \rightarrow \mathbb{R}$ and let $\bar{x} \in \operatorname{argmin}_{\mathbb{R}^n} h$. For $\gamma \in]0, 1[$ and $\rho \geq 0$, define

$$g(x) := h(x) + \frac{\gamma \rho}{2(2 - \gamma)} \|x - \bar{x}\|^2. \quad (20)$$

Then h is (ρ, γ) -weakly quasar-convex with respect to $\bar{x}$ if and only if g is γ -quasar-convex with respect to $\bar{x}$.

Proof. Clearly, $\bar{x} \in \operatorname{argmin}_{\mathbb{R}^n} g$. Furthermore, throughout the proof, “ g is γ -quasar-convex with respect to $\bar{x}$ ” refers to the $\rho = 0$ instance of (10), namely

$$g(\lambda\bar{x} + (1-\lambda)x) \leq \gamma\lambda g(\bar{x}) + (1-\gamma\lambda)g(x), \quad \forall x \in \mathbb{R}^n, \quad \forall \lambda \in [0, 1]. \quad (21)$$

Set $c := \frac{\gamma\rho}{2(2-\gamma)}$. Since $\gamma \in]0, 1]$, we have $2-\gamma \geq 1 > 0$, so c is well defined and $c \geq 0$ because $\rho \geq 0$. Furthermore, the term $c\|x - \bar{x}\|^2$ is non-negative for every $x \in \mathbb{R}^n$, vanishing precisely at $x = \bar{x}$ when $\rho > 0$ (and identically, with $g = h$, when $\rho = 0$). Consequently,

$$g(x) - g(\bar{x}) = (h(x) - h(\bar{x})) + c\|x - \bar{x}\|^2 \geq 0, \quad \forall x \in \mathbb{R}^n,$$

since both summands are non-negative, the first because $\bar{x} \in \operatorname{argmin}_{\mathbb{R}^n} h$. It therefore follows that $\bar{x}$ is likewise a global minimizer of g over $\mathbb{R}^n$, thus the base point $\bar{x}$ remains valid for the shifted function g , and the assertion that g is γ -quasar-convex with respect to $\bar{x}$ is consequently well posed.

Fix $x \in \mathbb{R}^n$ and $\lambda \in [0, 1]$, and set $z_\lambda := \lambda\bar{x} + (1-\lambda)x$. Since $z_\lambda - \bar{x} = (1-\lambda)(x - \bar{x})$, we have $\|z_\lambda - \bar{x}\|^2 = (1-\lambda)^2\|x - \bar{x}\|^2$, thus

$$g(\bar{x}) = h(\bar{x}), \quad g(x) = h(x) + c\|x - \bar{x}\|^2, \quad \text{and} \quad g(z_\lambda) = h(z_\lambda) + c(1-\lambda)^2\|x - \bar{x}\|^2.$$

Substituting these three expressions in (21) and rearranging shows that $g(z_\lambda) \leq \gamma\lambda g(\bar{x}) + (1-\gamma\lambda)g(x)$ holds if and only if

$$h(z_\lambda) \leq \gamma\lambda h(\bar{x}) + (1-\gamma\lambda)h(x) + c\left((1-\gamma\lambda) - (1-\lambda)^2\right)\|x - \bar{x}\|^2.$$

A direct expansion gives $(1-\gamma\lambda) - (1-\lambda)^2 = \lambda(2-\gamma-\lambda)$, so that

$$c\left((1-\gamma\lambda) - (1-\lambda)^2\right) = \frac{\gamma\rho}{2(2-\gamma)}\lambda(2-\gamma-\lambda) = \frac{\gamma\rho}{2}\lambda\left(1 - \frac{\lambda}{2-\gamma}\right),$$

which is exactly the coefficient of $\|x - \bar{x}\|^2$ appearing in (10). Then, for every $x \in \mathbb{R}^n$ and every $\lambda \in [0, 1]$, the inequality $g(z_\lambda) \leq \gamma\lambda g(\bar{x}) + (1-\gamma\lambda)g(x)$ holds if and only if

$$h(z_\lambda) \leq \gamma\lambda h(\bar{x}) + (1-\gamma\lambda)h(x) + \frac{\gamma\rho}{2}\lambda\left(1 - \frac{\lambda}{2-\gamma}\right)\|x - \bar{x}\|^2,$$

that is, precisely when inequality (10) holds for h at the same pair (x, λ) . Since both notions of quasar-convexity require their defining inequality to hold for every $x \in \mathbb{R}^n$ and every $\lambda \in [0, 1]$, taking the conjunction over all such pairs on each side yields h is (ρ, γ) -weakly quasar-convex with respect to $\bar{x}$ if and only if g is γ -quasar-convex with respect to $\bar{x}$, which proves the proposition. $\square$

Now, we provide a sufficient condition for the preservation of the weak quasar-convexity under linear composition.

Proposition 9. *Let $h : \mathbb{R}^m \rightarrow \mathbb{R}$ be (ρ, γ) -weakly quasar-convex with respect to $\bar{y} \in \operatorname{argmin} h$, with $\rho \geq 0$ and $\gamma \in]0, 1]$. Let $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear operator satisfying $A\bar{x} = \bar{y}$, and assume $\bar{x} \in \operatorname{argmin}_{\mathbb{R}^n} g$ where $g(x) = h(Ax)$. Then g is $(\rho\|A\|_{\text{op}}^2, \gamma)$ -weakly quasar-convex with respect to $\bar{x}$, where $\|A\|_{\text{op}}$ denotes the operator norm of A .*

Proof. Since A is linear and $A\bar{x} = \bar{y}$, we have for every $x \in \mathbb{R}^n$ and every $\lambda \in [0, 1]$ that $A(\lambda\bar{x} + (1-\lambda)x) = \lambda A\bar{x} + (1-\lambda)Ax = \lambda\bar{y} + (1-\lambda)Ax$. Applying the (ρ, γ) -weak quasar-convexity of h at $\bar{y}$ with $y = Ax$ yields

$$\begin{aligned} g(\lambda\bar{x} + (1-\lambda)x) &= h(A(\lambda\bar{x} + (1-\lambda)x)) = h(\lambda\bar{y} + (1-\lambda)Ax) \\ &\leq \gamma\lambda h(\bar{y}) + (1-\gamma\lambda)h(Ax) + \frac{\gamma\rho}{2}\lambda\left(1 - \frac{\lambda}{2-\gamma}\right)\|Ax - \bar{y}\|^2. \end{aligned} \quad (22)$$

Since $\bar{y} = A\bar{x}$, we have $\|Ax - \bar{y}\| = \|A(x - \bar{x})\|$. By definition of the operator norm,

$$\|A(x - \bar{x})\| \leq \|A\|_{\text{op}}\|x - \bar{x}\| \implies \|Ax - \bar{y}\|^2 \leq \|A\|_{\text{op}}^2\|x - \bar{x}\|^2.$$

Substituting this into (22) and using $g(\bar{x}) = h(\bar{y})$ and $g(x) = h(Ax)$, we obtain

$$\begin{aligned} g(\lambda\bar{x} + (1-\lambda)x) &\leq \gamma\lambda g(\bar{x}) + (1-\gamma\lambda)g(x) + \frac{\gamma\rho}{2}\lambda\left(1 - \frac{\lambda}{2-\gamma}\right)\|A\|_{\text{op}}^2\|x - \bar{x}\|^2 \\ &= \gamma\lambda g(\bar{x}) + (1-\gamma\lambda)g(x) + \frac{\gamma(\rho\|A\|_{\text{op}}^2)}{2}\lambda\left(1 - \frac{\lambda}{2-\gamma}\right)\|x - \bar{x}\|^2. \end{aligned}$$

Hence g is $(\rho\|A\|_{\text{op}}^2, \gamma)$ -weak quasar-convex at $\bar{x}$. $\square$

Another interesting property of weakly quasar-convex functions, is the following extension of [3, Theorem 4.8] from the quasar-convex to the weakly quasar-convex case, (see also [24, Theorem 3]). Its similar proof is omitted.

Proposition 10. *Let $h_i : \mathbb{R}^{n_i} \rightarrow \mathbb{R}$, with $i \in \{1, \dots, m\}$ be functions, with $n_i \geq 0$ and $\sum_{i=1}^m n_i = n$, and $h : \mathbb{R}^n \rightarrow \mathbb{R}$ be a function such that*

$$h(x_1, \dots, x_m) = \sum_{i=1}^m h_i(x_i), \quad (23)$$

Let $\bar{x} = (\bar{x}_1, \dots, \bar{x}_m) \in \mathbb{R}^n$. Then, for any $\rho \geq 0$ and $\gamma \in]0, 1]$, h is (ρ, γ) -weakly quasar-convex with respect to $\bar{x} \in \operatorname{argmin}_{\mathbb{R}^n} h$ if and only if every h_i is (ρ, γ) -weakly quasar-convex with respect to $\bar{x}_i \in \operatorname{argmin}_{\mathbb{R}^{n_i}} h_i$ (with the same ρ and γ).

The proof of the following characterization follows the reasoning line from [1, Theorem 31] (which studies the strong case), as a consequence, it is omitted.

Proposition 11 (Clarke First-Order Characterization). *Assume h is a locally Lipschitz function with a convex domain. Then h is (ρ, γ) -weakly quasar-convex with respect to $\bar{x} \in \operatorname{argmin} h$, where $\rho \geq 0$ and $\gamma \in]0, 1]$, if and only if for every $x \neq \bar{x}$, we have*

$$h(\bar{x}) \geq h(x) + \frac{1}{\gamma} \langle v, \bar{x} - x \rangle - \frac{\rho}{2} \|x - \bar{x}\|^2, \quad \forall v \in \partial^C h(x). \quad (24)$$

Since a continuously differentiable function is locally Lipschitz, and its Clarke subdifferential reduces to the gradient (i.e., $\partial^C h(x) = \{\nabla h(x)\}$ [6]), Proposition 11 immediately yields the following smooth characterization.

Corollary 12 (First-Order Characterization). *Assume that h is continuously differentiable on a convex domain $\operatorname{dom}(h)$. Then h is (ρ, γ) -weakly quasar-convex with respect to $\bar{x} \in \operatorname{argmin} h$, where $\rho \geq 0$ and $\gamma \in]0, 1]$, if and only if*

$$h(\bar{x}) \geq h(x) + \frac{1}{\gamma} \langle \nabla h(x), \bar{x} - x \rangle - \frac{\rho}{2} \|x - \bar{x}\|^2.$$

An immediate structural consequence of combining the μ -local quadratic growth condition (3) with (ρ, γ) -weak quasar-convexity under the assumption $\mu > \rho$, is that the ζ -sublevel set $\mathcal{L}_\zeta^h(\mathbb{R}^n, \bar{x})$ is free of spurious stationary points, as we show next.

Proposition 13 (Absence of Spurious Stationary Points). *Let $h : \mathbb{R}^n \rightarrow \mathbb{R}$ be a continuously differentiable function that satisfies the μ -local quadratic growth condition (3) and is (ρ, γ) -weakly quasar-convex with respect to the global minimizer $\bar{x}$, with $\rho \geq 0$ and $\gamma \in]0, 1]$. If $\mu > \rho$, then h contains no spurious stationary points within the sublevel set $\mathcal{L}_\zeta^h(\mathbb{R}^n, \bar{x})$ for some $\zeta > 0$, that is, for every $x \in \mathcal{L}_\zeta^h(\mathbb{R}^n, \bar{x})$ with $x \neq \bar{x}$, we have $\nabla h(x) \neq 0$.*

Proof. Let $x \in \mathcal{L}_\zeta^h(\mathbb{R}^n, \bar{x})$. By the first-order characterization in Corollary 12, we have (the fundamental gradient bound)

$$\langle \nabla h(x), x - \bar{x} \rangle \geq \gamma(h(x) - h(\bar{x})) - \frac{\gamma\rho}{2} \|x - \bar{x}\|^2. \quad (25)$$

Since $x \in \mathcal{L}_\zeta^h(\mathbb{R}^n, \bar{x})$, the local quadratic growth condition (3) applies, giving $h(x) - h(\bar{x}) \geq \frac{\mu}{2} \|x - \bar{x}\|^2$. Because $\gamma > 0$, we can substitute this lower bound into (25) to obtain

$$\langle \nabla h(x), x - \bar{x} \rangle \geq \gamma \left(\frac{\mu}{2} \|x - \bar{x}\|^2 \right) - \frac{\gamma\rho}{2} \|x - \bar{x}\|^2 = \frac{\gamma(\mu - \rho)}{2} \|x - \bar{x}\|^2.$$

Now, suppose for the contrary that there exists a spurious stationary point $x \in \mathcal{L}_\zeta^h(\mathbb{R}^n, \bar{x})$ with $x \neq \bar{x}$. Hence, $\nabla h(x) = 0$. Also, since $\gamma > 0$ and $\mu > \rho$ by assumption, we get $\gamma(\mu - \rho) > 0$, and since $x \neq \bar{x}$, $\|x - \bar{x}\|^2 > 0$. Consequently,

$$0 \geq \frac{\gamma(\mu - \rho)}{2} \|x - \bar{x}\|^2 > 0,$$

a contradiction. Therefore, no such spurious stationary point can exist in $\mathcal{L}_\zeta^h(\mathbb{R}^n, \bar{x})$. $\square$

Remark 14 (Alternative Sufficient Conditions for Local QG). *Several other well-known regularity conditions guarantee the local quadratic growth condition (3). Notably, local strong (quasar-)convexity inherently enforces a quadratic lower bound, directly satisfying the QG condition. Furthermore, the Restricted Secant Inequality (RSI) ensures that the gradient points sufficiently toward the optimal set; as established in the literature, strong quasar-convexity and strong star quasi-convexity imply the RSI, which is fundamentally linked to and often used to derive quadratic growth [18, Properties 6.13, 6.14, 6.17, 6.18]. These conditions collectively form a hierarchy of regularity that relaxes classical strong convexity while preserving the favorable algorithmic guarantees provided by local QG.* $\square$

In the following result, we provide a sufficient condition for weakly quasar-convex functions to satisfies the RSI property locally around the minimizer.

Proposition 15. *Let $h : \mathbb{R}^n \rightarrow \mathbb{R}$ be a continuously differentiable and (ρ, γ) -weakly quasar-convex function with respect to $\bar{x} \in \operatorname{argmin} h$, with $\rho \geq 0$ and $\gamma \in]0, 1]$. If h satisfies the μ -local quadratic growth condition (3) on $\mathcal{L}_\zeta^h(\mathbb{R}^n, \bar{x})$ for some $\zeta > 0$ and $\mu > \rho$, then h satisfies RSI on $\mathcal{L}_\zeta^h(\mathbb{R}^n, \bar{x})$ with parameter $\nu = \frac{\gamma(\mu-\rho)}{2} > 0$, that is,*

$$\langle \nabla h(x), x - \bar{x} \rangle \geq \frac{\gamma(\mu-\rho)}{2} \|x - \bar{x}\|^2, \quad \forall x \in \mathcal{L}_\zeta^h(\mathbb{R}^n, \bar{x}). \quad (26)$$

Proof. By the first-order characterization of (ρ, γ) -weakly quasar-convexity (Corollary 12), we have

$$\langle \nabla h(x), x - \bar{x} \rangle \geq \gamma(h(x) - h(\bar{x})) - \frac{\gamma\rho}{2} \|x - \bar{x}\|^2, \quad \forall x \in \mathbb{R}^n.$$

Since the local quadratic growth condition (3) holds, there exists $\zeta > 0$ such that for all $x \in \mathcal{L}_\zeta^h(\mathbb{R}^n, \bar{x})$ we have $h(x) - h(\bar{x}) \geq \frac{\mu}{2} \|x - \bar{x}\|^2$. Using this in the previous inequality and since $\gamma > 0$, we obtain

$$\langle \nabla h(x), x - \bar{x} \rangle \geq \gamma \left(\frac{\mu}{2} \|x - \bar{x}\|^2 \right) - \frac{\gamma\rho}{2} \|x - \bar{x}\|^2 = \frac{\gamma(\mu-\rho)}{2} \|x - \bar{x}\|^2,$$

which completes the proof. $\square$

3.2 Moreau Envelope

We recall the following well-known property of the Moreau envelope (e.g., see [17]), and then show that this smoothing operation also preserves weak quasar-convexity.

Lemma 16. *Let h be proper lower semicontinuous function, and $\bar{x} \in \operatorname{argmin}_{\mathbb{R}^n} h$. Then, for every $\beta > 0$, h^β is finite-valued on $\mathbb{R}^n$, $h^\beta(\bar{x}) = h(\bar{x})$, and $\operatorname{argmin}_{\mathbb{R}^n} h^\beta = \operatorname{argmin}_{\mathbb{R}^n} h$.*

We now show that the Moreau envelope preserves the weak quasar-convexity of the original function.

Theorem 17 (Moreau Envelope Preserves Weak Quasar-Convexity). *Let $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be proper, lower semicontinuous, and (ρ, γ) -weakly quasar-convex with respect to $\bar{x} \in \operatorname{argmin}_{\mathbb{R}^n} h$, with $\rho \geq 0$ and $\gamma \in]0, 1]$. Let*

$$0 < \beta < \frac{2-\gamma}{\rho\gamma}, \quad (27)$$

with the convention that the right-hand side is $+\infty$ when $\rho = 0$. Then h^β is (ρ_β, γ) -weakly quasar-convex with respect to the same point $\bar{x} \in \operatorname{argmin}_{\mathbb{R}^n} h^\beta$, where

$$\rho_\beta = \frac{(2-\gamma)\rho}{(2-\gamma) - \gamma\rho\beta} \geq 0. \quad (28)$$

Proof. By Lemma 16 (applied with this particular $\beta > 0$), h^β is finite-valued, $h^\beta(\bar{x}) = h(\bar{x})$, and $\bar{x} \in \operatorname{argmin}_{\mathbb{R}^n} h^\beta$. It remains to prove the chord inequality for h^β with modulus ρ_β . Fix $x \in \mathbb{R}^n$, $\lambda \in [0, 1]$, and $\varepsilon > 0$. Choose $y^* \in \mathbb{R}^n$ with

$$h(y^*) + \frac{1}{2\beta} \|x - y^*\|^2 \leq h^\beta(x) + \varepsilon. \quad (29)$$

Set $z_\lambda := \lambda\bar{x} + (1-\lambda)x$ and $y_\lambda := \lambda\bar{x} + (1-\lambda)y^*$. Since y_λ is feasible for $h^\beta(z_\lambda)$ and $z_\lambda - y_\lambda = (1-\lambda)(x - y^*)$,

$$h^\beta(z_\lambda) \leq h(y_\lambda) + \frac{(1-\lambda)^2}{2\beta} \|x - y^*\|^2. \quad (30)$$

Applying (10) to the pair $(\bar{x}, y^*)$, we have

$$h(y_\lambda) \leq \gamma\lambda h(\bar{x}) + (1-\gamma\lambda)h(y^*) + \frac{\gamma\rho}{2}\lambda\left(1 - \frac{\lambda}{2-\gamma}\right)\|y^* - \bar{x}\|^2.$$

Substituting into (30) and using $h(\bar{x}) = h^\beta(\bar{x})$,

$$h^\beta(z_\lambda) \leq \gamma\lambda h^\beta(\bar{x}) + (1-\gamma\lambda)h(y^*) + \frac{\gamma\rho}{2}\lambda\left(1 - \frac{\lambda}{2-\gamma}\right)\|y^* - \bar{x}\|^2 + \frac{(1-\lambda)^2}{2\beta}\|x - y^*\|^2. \quad (31)$$

From (29), $h(y^*) \leq h^\beta(x) + \varepsilon - \frac{1}{2\beta}\|x - y^*\|^2$. Since $1 - \gamma\lambda \geq 0$, substitute into (31) we get

$$\begin{aligned} h^\beta(z_\lambda) &\leq \gamma\lambda h^\beta(\bar{x}) + (1 - \gamma\lambda)(h^\beta(x) + \varepsilon) + \frac{\gamma\rho}{2}\lambda\left(1 - \frac{\lambda}{2-\gamma}\right)\|y^* - \bar{x}\|^2 \\ &\quad + \frac{1}{2\beta}((1 - \lambda)^2 - (1 - \gamma\lambda))\|x - y^*\|^2. \end{aligned} \quad (32)$$

Since $(1 - \lambda)^2 - (1 - \gamma\lambda) = -\lambda(2 - \gamma - \lambda)$, $\lambda \in [0, 1]$ and $\gamma \in [0, 1]$, we have $2 - \gamma - \lambda \geq 1 - \gamma \geq 0$, this bracket is ≤ 0 . Let $c := \frac{\gamma\rho}{2-\gamma} \geq 0$, so $\frac{\gamma\rho}{2}(1 - \frac{\lambda}{2-\gamma}) = \frac{c}{2}(2 - \gamma - \lambda)$. Factor $\frac{\lambda(2-\gamma-\lambda)}{2}$ from the two quadratic terms in (32) gives

$$\frac{c}{2}(2 - \gamma - \lambda)\lambda\|y^* - \bar{x}\|^2 - \frac{\lambda(2-\gamma-\lambda)}{2\beta}\|x - y^*\|^2 = \frac{\lambda(2-\gamma-\lambda)}{2}\left(c\|y^* - \bar{x}\|^2 - \frac{1}{\beta}\|x - y^*\|^2\right).$$

Set $u := y^* - x$ and $v := x - \bar{x}$, so $y^* - \bar{x} = u + v$, and

$$c\|y^* - \bar{x}\|^2 - \frac{1}{\beta}\|x - y^*\|^2 = -\left(\frac{1}{\beta} - c\right)\|u\|^2 + 2c\langle u, v \rangle + c\|v\|^2.$$

Let $A := \frac{1}{\beta} - c$. The assumption (27) is equivalent to $c\beta < 1$. Indeed, if $\rho = 0$, then $c = 0$, and if $\rho > 0$, then $\beta < \frac{2-\gamma}{\gamma\rho}$ if and only if $\frac{\gamma\rho\beta}{2-\gamma} < 1$. Hence, $A = \frac{1}{\beta} - c = \frac{1-c\beta}{\beta} > 0$.

We can maximize this quadratic form over u by completing the square

$$-A\|u\|^2 + 2c\langle u, v \rangle + c\|v\|^2 = -A\left\|u - \frac{c}{A}v\right\|^2 + \frac{c^2}{A}\|v\|^2 + c\|v\|^2 \leq c\left(\frac{c+A}{A}\right)\|v\|^2.$$

Substituting $A = \frac{1}{\beta} - c$ back into the scalar coefficient yields $c\frac{(c+(\frac{1}{\beta}-c))}{(\frac{1}{\beta}-c)} = \frac{c}{1-c\beta}$. Then, we obtain the tight upper bound

$$c\|y^* - \bar{x}\|^2 - \frac{1}{\beta}\|x - y^*\|^2 \leq \frac{c}{1-c\beta}\|x - \bar{x}\|^2.$$

Hence,

$$\frac{\gamma\rho}{2}\lambda\left(1 - \frac{\lambda}{2-\gamma}\right)\|y^* - \bar{x}\|^2 - \frac{\lambda(2-\gamma-\lambda)}{2\beta}\|x - y^*\|^2 \leq \frac{\lambda(2-\gamma-\lambda)}{2} \cdot \frac{c}{1-c\beta}\|x - \bar{x}\|^2.$$

Substituting into (32) and using $\lambda(2 - \gamma - \lambda) = (2 - \gamma)\lambda(1 - \frac{\lambda}{2-\gamma})$, we have

$$\begin{aligned} h^\beta(z_\lambda) &\leq \gamma\lambda h^\beta(\bar{x}) + (1 - \gamma\lambda)h^\beta(x) + (1 - \gamma\lambda)\varepsilon + \frac{c(2-\gamma)}{2(1-c\beta)}\lambda\left(1 - \frac{\lambda}{2-\gamma}\right)\|x - \bar{x}\|^2 \\ &\stackrel{\varepsilon \downarrow 0}{\implies} h^\beta(z_\lambda) \leq \gamma\lambda h^\beta(\bar{x}) + (1 - \gamma\lambda)h^\beta(x) + \frac{c(2-\gamma)}{2(1-c\beta)}\lambda\left(1 - \frac{\lambda}{2-\gamma}\right)\|x - \bar{x}\|^2. \end{aligned}$$

Comparing with (10), this is the (ρ_β, γ) -weak quasar-convexity inequality for h^β , with

$$\frac{\gamma\rho_\beta}{2} = \frac{c(2-\gamma)}{2(1-c\beta)} \iff \rho_\beta = \frac{(2-\gamma)c}{\gamma(1-c\beta)} = \frac{\rho}{1-c\beta},$$

using $c = \frac{\gamma\rho}{2-\gamma}$. This is exactly (28). Since $x \in \mathbb{R}^n$ and $\lambda \in [0, 1]$ were arbitrary, the proof is complete. $\square$

Remark 18. Let $\rho \geq 0$ and $\gamma \in [0, 1]$, with the convention that both $\frac{1}{\rho}$ and $\frac{2-\gamma}{\gamma\rho}$ equal $+\infty$ when $\rho = 0$. Since $\gamma \leq 2 - \gamma$, it holds that

$$\frac{1}{\rho} \leq \frac{2-\gamma}{\gamma\rho},$$

with equality if and only if $\gamma = 1$ (or $\rho = 0$). Consequently, Theorem 17 applies in particular for every $0 < \beta < \frac{1}{\rho}$, and the same conclusion remains valid. In this range, the relation $c = \frac{\gamma\rho}{2-\gamma} \leq \rho$ yields the simpler (yet weaker) upper bound

$$\rho_\beta = \frac{\rho}{1-c\beta} \leq \frac{\rho}{1-\rho\beta}.$$

Condition (27) is precisely the requirement needed in the proof, namely $A = \frac{1}{\beta} - c > 0$. $\square$

3.3 On the Proximity Operator

In this subsection, we present the basic properties of the proximity operator of a weak quasar-convex function.

As a first result, we show a sufficient condition for the supercoercivity of the proximity operator of a weakly quasar-convex function.

Lemma 19. Let $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be proper with $\bar{x} \in \operatorname{argmin}_{\mathbb{R}^n} h$, and let $\beta > 0$. Then, for every $x \in \mathbb{R}^n$, the function $\Phi(y) := h(y) + \frac{1}{2\beta}\|x - y\|^2$ is 2-supercoercive; more precisely,

$$\Phi(y) \geq h(\bar{x}) + \frac{1}{2\beta}\|x - y\|^2 \quad \forall y \in \mathbb{R}^n, \quad \text{hence} \quad \liminf_{\|y\| \rightarrow \infty} \frac{\Phi(y)}{\|y\|^2} \geq \frac{1}{2\beta} > 0. \quad (33)$$

Proof. The first inequality follows from $h(y) \geq h(\bar{x})$. Dividing by $\|y\|^2$ and using $\|x - y\|^2/\|y\|^2 \rightarrow 1$ as $\|y\| \rightarrow \infty$ gives the second one. $\square$

Remark 20. Lemma 19 does not use weak quasar-convexity, nor any restriction on β : it only uses that $\bar{x}$ is a global minimizer of h . In particular, the proximal subproblems in Corollary 21 and Algorithm 1 are solvable for every $\beta > 0$, and the upper bound $\beta_k < 1/\rho$ in (37) can be dropped.

As a consequence of Lemma 19 and the lower semicontinuity of h .

Corollary 21. Let $\mathcal{K} \subseteq \mathbb{R}^n$ be a nonempty closed convex set and let $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be proper and lower semicontinuous, with $\mathcal{K} \cap \operatorname{dom}(h) \neq \emptyset$ and $\bar{x} \in \operatorname{argmin}_{\mathcal{K}} h$. Then, for every $\beta > 0$ and every $x \in \mathbb{R}^n$, the set $\operatorname{Prox}_{\beta h}(\mathcal{K}, x)$ is nonempty and compact. In particular, this holds for every (ρ, γ) -weakly quasar-convex function with respect to $\bar{x}$, without any restriction on β .

Proof. For every $y \in \mathcal{K}$ we have $h(y) \geq h(\bar{x})$, hence $h(y) + \frac{1}{2\beta}\|x - y\|^2 \geq h(\bar{x}) + \frac{1}{2\beta}\|x - y\|^2$, exactly as in Lemma 19. Thus $\Phi := h + \iota_{\mathcal{K}} + \frac{1}{2\beta}\|x - \cdot\|^2$ is proper, lower semicontinuous and coercive, so its set of minimizers is nonempty and compact. $\square$

Note that $\operatorname{Prox}_{\beta h}(\cdot, x)$ may be set-valued even for (strongly) quasar-convex functions as noted in [8, Remark 14].

Proposition 22. Let $\mathcal{K} \subseteq \mathbb{R}^n$ be a closed convex set and $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be a proper, lower semicontinuous and (ρ, γ) -weakly quasar-convex with respect to $\bar{x} \in \operatorname{argmin}_{\mathcal{K}} h$ ($\gamma \in]0, 1]$ and $\rho \geq 0$) on $\mathcal{K} \subseteq \operatorname{dom}(h)$. Let $z \in \mathcal{K}$ and $\beta > 0$. If $x^* \in \operatorname{Prox}_{\beta h}(\mathcal{K}, z)$, then

$$h(x^*) - h(\bar{x}) \leq \frac{1}{\gamma\beta} \langle x^* - z, \bar{x} - x^* \rangle + \frac{\rho}{2} \|x^* - \bar{x}\|^2. \quad (34)$$

Proof. Since $x^* \in \operatorname{Prox}_{\beta h}(\mathcal{K}, z)$, we have

$$h(x^*) + \frac{1}{2\beta}\|x^* - z\|^2 \leq h(x) + \frac{1}{2\beta}\|x - z\|^2, \quad \forall x \in \mathcal{K}. \quad (35)$$

Fix $\lambda \in]0, 1]$ and set $x = \lambda\bar{x} + (1 - \lambda)x^* \in \mathcal{K}$. Since h is (ρ, γ) -weakly quasar-convex,

$$\begin{aligned} h(x^*) + \frac{1}{2\beta}\|x^* - z\|^2 &\leq \gamma\lambda h(\bar{x}) + (1 - \gamma\lambda)h(x^*) + \frac{\gamma\rho}{2}\lambda \left(1 - \frac{\lambda}{2 - \gamma}\right) \|x^* - \bar{x}\|^2 + \frac{1}{2\beta}\|\lambda\bar{x} + (1 - \lambda)x^* - z\|^2 \\ &= \gamma\lambda h(\bar{x}) + (1 - \gamma\lambda)h(x^*) + \frac{\gamma\rho}{2}\lambda \left(1 - \frac{\lambda}{2 - \gamma}\right) \|x^* - \bar{x}\|^2 \\ &\quad + \frac{\lambda}{2\beta}\|\bar{x} - z\|^2 + \frac{1 - \lambda}{2\beta}\|x^* - z\|^2 - \frac{\lambda(1 - \lambda)}{2\beta}\|\bar{x} - x^*\|^2. \end{aligned}$$

That is,

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

Dividing by $\lambda > 0$ and using identity (1), we obtain

$$\begin{aligned} \gamma(h(x^*) - h(\bar{x})) &\leq \frac{1}{\beta} \langle x^* - z, \bar{x} - x^* \rangle + \frac{1}{2\beta}\|\bar{x} - x^*\|^2 + \frac{\gamma\rho}{2} \left(1 - \frac{\lambda}{2 - \gamma}\right) \|x^* - \bar{x}\|^2 - \frac{1 - \lambda}{2\beta}\|\bar{x} - x^*\|^2 \\ &= \frac{1}{\beta} \langle x^* - z, \bar{x} - x^* \rangle + \frac{\lambda}{2\beta}\|x^* - \bar{x}\|^2 + \frac{\gamma\rho}{2} \left(1 - \frac{\lambda}{2 - \gamma}\right) \|x^* - \bar{x}\|^2, \end{aligned}$$

and the result follows by taking $\lambda \downarrow 0$. $\square$

The following sufficient decrease property holds:

Proposition 23. Let $\mathcal{K} \subseteq \mathbb{R}^n$ be a closed convex set and $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be a proper, lower semicontinuous, and (ρ, γ) -weakly quasar-convex with respect to $\bar{x} \in \operatorname{argmin}_{\mathcal{K}} h$ ($\gamma \in]0, 1]$ and $\rho \geq 0$) on $\mathcal{K} \subseteq \operatorname{dom}(h)$. Then for every $z \in \mathcal{K}$, $\beta > 0$ and $x^* \in \operatorname{Prox}_{\beta h}(\mathcal{K}, z)$, we have

$$h(x^*) + \frac{1}{2\beta}\|x^* - z\|^2 \leq h(z). \quad (36)$$

Furthermore, if $0 < \gamma\beta\rho < 1$, then

$$\|x^* - \bar{x}\| \leq \frac{1}{\sqrt{1 - \gamma\beta\rho}} \|z - \bar{x}\|.$$

Proof. The inequality (36) holds for any proper functions (does not need convexity assumptions), and follows directly by taking $x = z$ in (35).

For the second part, from relation (34) and since $h(x^*) \geq h(\bar{x})$, we have

$$0 \leq \frac{1}{\gamma\beta} \langle x^* - z, \bar{x} - x^* \rangle + \frac{\rho}{2} \|x^* - \bar{x}\|^2.$$

Using identity (1) we obtain

$$\begin{aligned} 0 &\leq \frac{1}{2\gamma\beta} (\|\bar{x} - z\|^2 - \|x^* - z\|^2 - \|\bar{x} - x^*\|^2) + \frac{\rho}{2} \|x^* - \bar{x}\|^2, \\ \implies \left(\frac{1}{\gamma\beta} - \rho\right) \|x^* - \bar{x}\|^2 &\leq \frac{1}{\gamma\beta} (\|\bar{x} - z\|^2 - \|x^* - z\|^2). \end{aligned}$$

Since $\|x^* - z\|^2 \geq 0$, the right-hand side is bounded above by $\frac{1}{\gamma\beta} \|\bar{x} - z\|^2$. Under the assumption $0 < \gamma\beta\rho < 1$ (so that $\frac{1}{\gamma\beta} - \rho > 0$), we obtain

$$\|x^* - \bar{x}\|^2 \leq \frac{\frac{1}{\gamma\beta}}{\frac{1}{\gamma\beta} - \rho} \|\bar{x} - z\|^2 = \frac{1}{1 - \gamma\beta\rho} \|z - \bar{x}\|^2,$$

which completes the proof. $\square$

4 The Proximal Point Method and Convergence Analysis

In this part, we consider the following general optimization problem

$$\min_{x \in \mathbb{R}^n} h(x), \tag{P}$$

where h is a weakly quasar-convex function. Our purpose is to solve problem (P) via the proximal point method in its classical form (see [36]).

Algorithm 1 Proximal Point Method (PPM)

Input: Initial point $x^0 \in \mathbb{R}^n$, $\beta' > 0$, and a sequence of step sizes $\{\beta_k\}_k$ such that

$$\beta_k \geq \beta' > 0, \quad \forall k \in \mathbb{N}_0. \tag{37}$$

Iterate: For $k = 0, 1, 2, \dots$

$$x^{k+1} \leftarrow \text{Prox}_{\beta_k h}(x^k) = \operatorname{argmin}_{z \in \mathbb{R}^n} \left\{ h(z) + \frac{1}{2\beta_k} \|z - x^k\|^2 \right\}. \tag{38}$$

Until: Terminate when the stopping criteria are satisfied.

Return: x^{k+1} .

Corollary 21 establishes the existence of solutions to all proximal subproblems, rendering the iterative scheme well-defined.

Lemma 24 (Descent Property). *Let $h : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ be proper, lower semicontinuous, and (ρ, γ) -weakly quasar-convex with respect to $\bar{x} \in \operatorname{argmin}_{\mathbb{R}^n} h$, with $\rho \geq 0$ and $\gamma \in]0, 1]$. Then for every $\beta_k > 0$ and $x^{k+1} \in \text{Prox}_{\beta_k h}(x^k)$, we have*

$$h(x^{k+1}) + \frac{1}{2\beta_k} \|x^{k+1} - x^k\|^2 \leq h(x^k). \tag{39}$$

In particular, $\{h(x^k)\}_k$ is a non-increasing sequence.

Proof. Directly from (35) with $z = x^k$ and $x^* = x^{k+1}$. $\square$

The following useful relation holds.

Proposition 25. *Let $\mathcal{K} \subset \mathbb{R}^n$ be a closed convex set and $h : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ be a proper, lower semicontinuous and (ρ, γ) -weakly quasar-convex function with respect to $\bar{x} \in \operatorname{argmin}_{\mathcal{K}} h$ ($\gamma \in]0, 1]$ and $\rho \geq 0$) on $\mathcal{K} \subset \operatorname{dom}(h)$, $x^k \in \mathbb{R}^n$ and $\beta_k > 0$ for all $k \in \mathbb{N}_0$. If $x^{k+1} \in \text{Prox}_{\beta_k h}(\mathcal{K}, x^k)$, then*

$$h(x^{k+1}) - h(\bar{x}) \leq \frac{\rho}{2} \|x^{k+1} - \bar{x}\|^2 + \frac{1}{\gamma\beta_k} \langle x^{k+1} - x^k, \bar{x} - x^{k+1} \rangle. \tag{40}$$

Proof. Directly from Proposition 22 with $z = x^k$ and $x^* = x^{k+1}$. $\square$

In the next result, we establish the convergence of the sequence generated by Algorithm 1 to a global solution of the minimization problem.

Theorem 26 (Convergence of PPM under Local Quadratic Growth). *Let $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be proper, lower semicontinuous, (ρ, γ) -weakly quasar-convex with respect to $\bar{x} \in \operatorname{argmin}_{\mathbb{R}^n} h$, with $\rho \geq 0$ and $\gamma \in]0, 1]$, and that satisfies the μ -local quadratic growth condition (3). Assume that $\mu > \rho$. Let $\{x^k\}_k$ be generated by Algorithm 1 with step sizes $\{\beta_k\}_k$ satisfying (37). If $x^0 \in \mathcal{L}_\zeta^h(\mathbb{R}^n, \bar{x})$ for some $\zeta > 0$, then setting*

$$\tau := \frac{1}{1 + (\mu - \rho)\gamma\beta'} \in]0, 1[, \quad (41)$$

we have for every $k \in \mathbb{N}_0$ that:

- (a) $x^k \in \mathcal{L}_\zeta^h(\mathbb{R}^n, \bar{x})$.
- (b) $\|x^{k+1} - \bar{x}\|^2 \leq \tau \|x^k - \bar{x}\|^2 \leq \tau^{k+1} \|x^0 - \bar{x}\|^2$.
- (c) $h(x^k) \rightarrow h(\bar{x})$.

Proof. (a): It follows by Lemma 24 that

$$h(x^{k+1}) \leq h(x^k) + \frac{1}{2\beta_k} \|x^{k+1} - x^k\|^2 \leq h(x^k). \quad (42)$$

Hence $\{h(x^k)\}$ is non-increasing, and by induction $h(x^k) \leq h(x^0) \leq h(\bar{x}) + \zeta$, so $x^k \in \mathcal{L}_\zeta^h(\mathbb{R}^n, \bar{x})$ for every $k \in \mathbb{N}_0$.

(b): Applying Proposition 25 with $\mathcal{K} = \mathbb{R}^n$ and rearranging yields

$$2\langle x^{k+1} - x^k, x^{k+1} - \bar{x} \rangle \leq \rho\gamma\beta_k \|x^{k+1} - \bar{x}\|^2 + 2\gamma\beta_k (h(\bar{x}) - h(x^{k+1})). \quad (43)$$

Since $x^{k+1} \in \mathcal{L}_\zeta$ by part (a), condition (3) gives

$$h(x^{k+1}) - h(\bar{x}) \geq \frac{\mu}{2} \|x^{k+1} - \bar{x}\|^2, \quad (44)$$

so that $h(\bar{x}) - h(x^{k+1}) \leq -\frac{\mu}{2} \|x^{k+1} - \bar{x}\|^2$. Substituting into (43),

$$2\langle x^{k+1} - x^k, x^{k+1} - \bar{x} \rangle \leq -(\mu - \rho)\gamma\beta_k \|x^{k+1} - \bar{x}\|^2. \quad (45)$$

Using (1), we have

$$\begin{aligned} \|x^{k+1} - \bar{x}\|^2 - \|x^k - \bar{x}\|^2 + \|x^{k+1} - x^k\|^2 &\leq -(\mu - \rho)\gamma\beta_k \|x^{k+1} - \bar{x}\|^2, \\ \implies (1 + (\mu - \rho)\gamma\beta_k) \|x^{k+1} - \bar{x}\|^2 &\leq \|x^k - \bar{x}\|^2. \end{aligned} \quad (46)$$

Since $\mu > \rho$, $\gamma > 0$, and $\beta_k \geq \beta' > 0$, we have $1 + (\mu - \rho)\gamma\beta_k \geq 1 + (\mu - \rho)\gamma\beta' = \frac{1}{\tau} > 1$, where $\tau \in]0, 1[$ is defined in (41). Dividing (46) by $1 + (\mu - \rho)\gamma\beta_k$ yields

$$\|x^{k+1} - \bar{x}\|^2 \leq \frac{1}{1 + (\mu - \rho)\gamma\beta_k} \|x^k - \bar{x}\|^2 \leq \tau \|x^k - \bar{x}\|^2.$$

Applying this bound recursively, we obtain (b).

(c): Since $\bar{x}$ is a global minimizer of h , $h(x^{k+1}) \geq h(\bar{x})$ for all k . Furthermore, by applying Proposition 25 with $\mathcal{K} = \mathbb{R}^n$, we get

$$0 \leq h(x^{k+1}) - h(\bar{x}) \leq \frac{\rho}{2} \|x^{k+1} - \bar{x}\|^2 + \frac{1}{\gamma\beta_k} \langle x^{k+1} - x^k, \bar{x} - x^{k+1} \rangle. \quad (47)$$

Applying the Cauchy-Schwarz inequality to the inner product and using $\beta_k \geq \beta' > 0$, we get

$$h(x^{k+1}) - h(\bar{x}) \leq \frac{\rho}{2} \|x^{k+1} - \bar{x}\|^2 + \frac{1}{\gamma\beta'} \|x^{k+1} - x^k\| \cdot \|x^{k+1} - \bar{x}\|. \quad (48)$$

From part (b), we know that $\|x^k - \bar{x}\| \leq \tau^{k/2} \|x^0 - \bar{x}\| \rightarrow 0$. By the triangle inequality,

$$\|x^{k+1} - x^k\| \leq \|x^{k+1} - \bar{x}\| + \|x^k - \bar{x}\| \leq (\tau^{(k+1)/2} + \tau^{k/2}) \|x^0 - \bar{x}\| \rightarrow 0.$$

Hence both $\|x^{k+1} - \bar{x}\| \rightarrow 0$ and $\|x^{k+1} - x^k\| \rightarrow 0$, so the right-hand side of (48) tends to 0. Combining with (47) yields $h(x^k) \rightarrow h(\bar{x})$, which completes the proof. $\square$

In particular, we have the following convergence rates.

Corollary 27 (Convergence Rates). *Under the same assumptions as Theorem 26, the following hold:*

(a) *The sequence of iterates $\{x^k\}_k$ converges Q -linearly to $\bar{x}$ with Q -factor $\sqrt{\tau}$, i.e.,*

$$\|x^{k+1} - \bar{x}\| \leq \frac{1}{\sqrt{1 + (\mu - \rho)\gamma\beta'}} \|x^k - \bar{x}\|, \quad \forall k \in \mathbb{N}_0. \quad (49)$$

(b) *The sequence of function values $\{h(x^k)\}_k$ converges R -linearly to $h(\bar{x})$ with R -factor $\sqrt{\tau}$, i.e.,*

$$h(x^k) - h(\bar{x}) \leq C(\sqrt{\tau})^{k-1}, \quad \forall k \in \mathbb{N}, \quad (50)$$

where

$$C := \left(\frac{\rho}{2} + \frac{1}{\gamma\beta'} (1 + \sqrt{\tau}) \right) \|x^0 - \bar{x}\|^2. \quad (51)$$

Proof. (a): This follows directly from Theorem 26(b). The contraction factor τ comes from the inequality

$$\|x^{k+1} - \bar{x}\|^2 \leq \frac{1}{1 + (\mu - \rho)\gamma\beta_k} \|x^k - \bar{x}\|^2 \leq \tau \|x^k - \bar{x}\|^2,$$

where the last inequality uses $\beta_k \geq \beta'$ from (37), and the result follows by applying this recursively.

(b): From Proposition 22 with $\mathcal{K} = \mathbb{R}^n$, we have

$$h(x^{k+1}) - h(\bar{x}) \leq \frac{\rho}{2} \|x^{k+1} - \bar{x}\|^2 + \frac{1}{\gamma\beta_k} \langle x^{k+1} - x^k, \bar{x} - x^{k+1} \rangle.$$

Applying the Cauchy-Schwarz inequality and using $\beta_k \geq \beta' > 0$, we get

$$h(x^{k+1}) - h(\bar{x}) \leq \frac{\rho}{2} \|x^{k+1} - \bar{x}\|^2 + \frac{1}{\gamma\beta'} \|x^{k+1} - x^k\| \cdot \|x^{k+1} - \bar{x}\|.$$

Using part (a) and the triangle inequality,

$$\|x^{k+1} - x^k\| \leq \|x^{k+1} - \bar{x}\| + \|x^k - \bar{x}\| \leq \left(\tau^{(k+1)/2} + \tau^{k/2} \right) \|x^0 - \bar{x}\|.$$

Substituting and simplifying,

$$\begin{aligned} h(x^{k+1}) - h(\bar{x}) &\leq \left(\frac{\rho}{2} \tau^{k+1} + \frac{1}{\gamma\beta'} (\tau^{(k+1)/2} + \tau^{k/2}) \tau^{(k+1)/2} \right) \|x^0 - \bar{x}\|^2 \\ &\leq \left(\frac{\rho}{2} + \frac{1}{\gamma\beta'} (1 + \sqrt{\tau}) \right) \tau^{k/2} \|x^0 - \bar{x}\|^2 \\ &= C \tau^{k/2}, \end{aligned}$$

with C as in (51). Since this bounds $h(x^{k+1}) - h(\bar{x})$ for every $k \in \mathbb{N}_0$, reindexing $n := k + 1$ gives

$$h(x^n) - h(\bar{x}) \leq C \tau^{(n-1)/2} = C(\sqrt{\tau})^{n-1}, \quad \forall n \geq 1,$$

which is the claimed bound. $\square$

We finish this subsection with the following complexity analysis. The standard proof is omitted.

Corollary 28 (Iteration Complexity). *Under the same assumptions as Corollary 27, the following assertions hold:*

(a) *For any $\varepsilon > 0$, after at most*

$$K_{\text{iter}}(\varepsilon) = \left\lceil \frac{2}{-\ln(\tau)} \ln \left(\frac{\|x^0 - \bar{x}\|}{\varepsilon} \right) \right\rceil \quad (52)$$

iterations, we have $\|x^K - \bar{x}\| \leq \varepsilon$.

(b) *For any $\varepsilon > 0$, after at most*

$$K_{\text{func}}(\varepsilon) = \left\lceil 1 + \frac{2}{-\ln(\tau)} \ln \left(\frac{C}{\varepsilon} \right) \right\rceil \quad (53)$$

iterations, we have $h(x^K) - h(\bar{x}) \leq \varepsilon$, where C is defined in Corollary 27.

In both cases, the iteration complexity is $O\left(\frac{1}{-\ln(\tau)} \ln(1/\varepsilon)\right)$, which is the standard complexity for linearly convergent methods.

Remark 29. Theorem 26 extends [27, Theorem 4.3] from weakly convex to weakly quasar-convex functions. Furthermore, when $\gamma = 1$ (i.e., h is weakly star-convex), our convergence rate in Corollary 27 is exactly as good as [27, Theorem 4.3] ($\tau = (1 + (\mu - \rho)\beta')^{-1}$ in our case and $\tau = (1 + c_{\min}(\mu_q - \rho))^{-1}$ in their case). $\square$

4.1 Computable Complexity via the Proximal Residual

The complexity results of Corollaries 27 and 28 are stated in terms of $\|x^0 - \bar{x}\|$ or the constant C , both of which require knowledge of the (typically unknown) minimizer $\bar{x}$. In this subsection, we derive a complementary, fully computable stopping criterion based on the *proximal residual*, together with its own iteration complexity, and we show it is consistent with the sub-optimality complexity already established.

Definition 30 (Proximal Residual). *Let $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be proper, lower semicontinuous, and $\beta > 0$. For $x \in \mathbb{R}^n$ and $x^* \in \text{Prox}_{\beta h}(x)$, the proximal residual at x (associated with x^*) is defined by*

$$r_\beta(x) := \frac{1}{\beta} (x - x^*).$$

Unlike $\|x - \bar{x}\|$, the quantity $\|r_\beta(x)\|$ is observable during the execution of Algorithm 1, since along the PPM trajectory $x^{k+1} \in \text{Prox}_{\beta_k h}(x^k)$ we have simply $r_{\beta_k}(x^k) = \frac{1}{\beta_k}(x^k - x^{k+1})$, computable from two consecutive iterates without any knowledge of $\bar{x}$.

We first relate the residual to sub-optimality, under the local quadratic growth condition (3) already used in Corollary 27.

Proposition 31 (Proximal-PL-type Inequality). *Let $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be proper, lower semicontinuous, (ρ, γ) -weakly quasar-convex with respect to $\bar{x} \in \text{argmin}_{\mathbb{R}^n} h$, and satisfy the μ -quadratic growth condition (3) at $\bar{x}$ globally on $\mathbb{R}^n$ (i.e., for every $\zeta > 0$), with $\mu > \rho$. Let $\beta > 0$, $x \in \mathbb{R}^n$, and $x^* \in \text{Prox}_{\beta h}(x)$. Then,*

$$h(x^*) - h(\bar{x}) \leq \frac{2\mu}{\gamma^2(\mu-\rho)^2} \|r_\beta(x)\|^2. \quad (54)$$

Proof. Taking $z = x$ in Proposition 22, we obtain

$$h(x^*) - h(\bar{x}) \leq \frac{1}{\gamma\beta} \langle x^* - x, \bar{x} - x^* \rangle + \frac{\rho}{2} \|x^* - \bar{x}\|^2.$$

By definition, $x^* - x = -\beta r_\beta(x)$, so, applying the Cauchy-Schwarz inequality,

$$\frac{1}{\gamma\beta} \langle x^* - x, \bar{x} - x^* \rangle = -\frac{1}{\gamma} \langle r_\beta(x), \bar{x} - x^* \rangle \leq \frac{1}{\gamma} \|r_\beta(x)\| \|x^* - \bar{x}\|.$$

Combining the two displays,

$$h(x^*) - h(\bar{x}) \leq \frac{1}{\gamma} \|r_\beta(x)\| \|x^* - \bar{x}\| + \frac{\rho}{2} \|x^* - \bar{x}\|^2. \quad (55)$$

On the other hand, by the μ -local quadratic growth condition (3) applied at x^* , we have $h(x^*) - h(\bar{x}) \geq \frac{\mu}{2} \|x^* - \bar{x}\|^2$. Combining this lower bound with (55) gives

$$\frac{\mu}{2} \|x^* - \bar{x}\|^2 \leq \frac{1}{\gamma} \|r_\beta(x)\| \|x^* - \bar{x}\| + \frac{\rho}{2} \|x^* - \bar{x}\|^2.$$

If $x^* = \bar{x}$, then (54) holds trivially since $h(x^*) - h(\bar{x}) = 0$. Otherwise, dividing by $\|x^* - \bar{x}\| > 0$ and rearranging, using $\mu > \rho$, we get $\|x^* - \bar{x}\| \leq \frac{2}{\gamma(\mu-\rho)} \|r_\beta(x)\|$. Substituting this bound back into (55),

$$\begin{aligned} h(x^*) - h(\bar{x}) &\leq \frac{1}{\gamma} \|r_\beta(x)\| \frac{2}{\gamma(\mu-\rho)} \|r_\beta(x)\| + \frac{\rho}{2} \left(\frac{2}{\gamma(\mu-\rho)} \right)^2 \|r_\beta(x)\|^2 \\ &= \frac{2}{\gamma^2(\mu-\rho)} \|r_\beta(x)\|^2 + \frac{2\rho}{\gamma^2(\mu-\rho)^2} \|r_\beta(x)\|^2 \\ &= \frac{2\mu}{\gamma^2(\mu-\rho)^2} \|r_\beta(x)\|^2, \end{aligned}$$

which is (54). $\square$

We now show that the PPM iterates drive the residual to zero linearly, at the same rate $\sqrt{\tau}$ already established for the iterate distance and function values.

Corollary 32 (Residual Convergence Rate). *Under the same assumptions as Corollary 27, let $\{x^k\}_k$ be generated by Algorithm 1 with $\beta_k \equiv \beta'$ satisfying (37), and set*

$$r^k := r_{\beta'}(x^k) = \frac{1}{\beta'} (x^k - x^{k+1}).$$

Then,

$$\|r^k\| \leq \frac{1+\sqrt{\tau}}{\beta'} (\sqrt{\tau})^k \|x^0 - \bar{x}\|, \quad \forall k \in \mathbb{N}_0. \quad (56)$$

Proof. By the triangle inequality and Corollary 27(a),

$$\|r^k\| = \frac{1}{\beta'} \|x^k - x^{k+1}\| \leq \frac{1}{\beta'} \left(\|x^k - \bar{x}\| + \|x^{k+1} - \bar{x}\| \right) \leq \frac{1}{\beta'} (\tau^{k/2} + \tau^{(k+1)/2}) \|x^0 - \bar{x}\|,$$

which simplifies to (56). $\square$

Corollary 33 (Residual-Based Iteration Complexity). *Under the same assumptions as Corollary 27, let $\{x^k\}_k$ be generated by Algorithm 1 with $\beta_k \equiv \beta'$ satisfying (37). Then for any $\varepsilon > 0$, after at most*

$$K_{\text{res}}(\varepsilon) = \left\lceil \frac{2}{-\ln(\tau)} \ln \left(\frac{(1+\sqrt{\tau})\|x^0 - \bar{x}\|}{\beta' \varepsilon} \right) \right\rceil \quad (57)$$

iterations, we have $\|r^K\| \leq \varepsilon$; this criterion is computable online, without knowledge of $\bar{x}$, from $\|x^K - x^{K+1}\|$ alone. Moreover, if the assumptions of Proposition 31 hold, then $\|r^K\| \leq \varepsilon$ implies

$$h(x^{K+1}) - h(\bar{x}) \leq \frac{2\mu}{\gamma^2(\mu-\rho)^2} \varepsilon^2.$$

In both cases, the iteration complexity is $O\left(\frac{1}{-\ln(\tau)} \ln(1/\varepsilon)\right)$, matching the order already established in Corollary 28.

Proof. The bound (57) follows from (56), requiring $\frac{1+\sqrt{\tau}}{\beta'} \tau^{K/2} \|x^0 - \bar{x}\| \leq \varepsilon$, and solving for K exactly as in the derivation of Corollary 28. The observability of $\|r^K\| = \frac{1}{\beta'} \|x^K - x^{K+1}\|$ from consecutive iterates is immediate from Definition 30. The final sub-optimality bound follows directly from Proposition 31, applied at $x = x^K$, $x^* = x^{K+1}$. $\square$

Remark 34. Corollary 33 recovers a complexity of the same order $O\left(\frac{1}{-\ln(\tau)} \ln(1/\varepsilon)\right)$ as Corollary 28, confirming consistency between the two criteria, but with an essential practical advantage: the stopping rule $\|x^K - x^{K+1}\| \leq \beta' \varepsilon$ can be checked by the algorithm itself during execution, whereas the criteria of Corollary 28 require an a priori estimate of $\|x^0 - \bar{x}\|$ or of the constant C , both of which depend on the unknown minimizer $\bar{x}$. $\square$

The following lemma guarantees that both the weak quasar-convexity and the local quadratic growth condition are preserved under such a quadratic perturbation.

Lemma 35. *Let $\phi : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ be (ρ, γ) -weakly quasar-convex with respect to $\bar{x} \in \text{argmin}_{\mathbb{R}^n} \phi$, and suppose ϕ satisfies the μ -local Quadratic Growth (QG) condition at $\bar{x}$ with parameter $\zeta > 0$. Then, for any $L \geq 0$, the perturbed function $\psi(x) := \phi(x) + L\|x - \bar{x}\|^2$ is (ρ_L, γ) -weakly quasar-convex with respect to $\bar{x}$, and satisfies the μ_L -local QG condition at $\bar{x}$ with the same parameter ζ , where*

$$\rho_L = \max \left\{ 0, \rho - 2L \left(\frac{2}{\gamma} - 1 \right) \right\} \quad \text{and} \quad \mu_L = \mu + 2L.$$

Proof. First, observe that since $\bar{x} \in \text{argmin}_{\mathbb{R}^n} \phi$ and $L \geq 0$, we have $\psi(x) = \phi(x) + L\|x - \bar{x}\|^2 \geq \phi(\bar{x}) = \psi(\bar{x})$ for all $x \in \mathbb{R}^n$. This ensures that $\bar{x} \in \text{argmin}_{\mathbb{R}^n} \psi$, which is a strict requirement for the definition of weak quasar-convexity. We now prove the two properties separately. For any $x \in \mathcal{L}_\zeta^\psi(\mathbb{R}^n, \bar{x})$, we have

$$\phi(x) \leq \psi(x) \leq \psi(\bar{x}) + \zeta = \phi(\bar{x}) + \zeta,$$

which implies $x \in \mathcal{L}_\zeta^\phi(\mathbb{R}^n, \bar{x})$. Therefore, the μ -QG condition for ϕ applies to x . By the definition of ψ , we obtain

$$\begin{aligned} \psi(x) - \psi(\bar{x}) &= (\phi(x) + L\|x - \bar{x}\|^2) - (\phi(\bar{x}) + 0) = (\phi(x) - \phi(\bar{x})) + L\|x - \bar{x}\|^2 \\ &\geq \frac{\mu}{2} \|x - \bar{x}\|^2 + L\|x - \bar{x}\|^2 = \frac{\mu + 2L}{2} \|x - \bar{x}\|^2. \end{aligned}$$

Thus, ψ satisfies the μ_L -QG condition with $\mu_L = \mu + 2L$ on $\mathcal{L}_\zeta^\psi(\mathbb{R}^n, \bar{x})$.

Now, let $x_\lambda = \lambda \bar{x} + (1 - \lambda)x$ for $\lambda \in [0, 1]$. Note that $x_\lambda - \bar{x} = (1 - \lambda)(x - \bar{x})$, which implies $\|x_\lambda - \bar{x}\|^2 = (1 - \lambda)^2 \|x - \bar{x}\|^2$. Using the definition of ψ and the (ρ, γ) -weak quasar-convexity of ϕ , we obtain

$$\begin{aligned} \psi(x_\lambda) &= \phi(x_\lambda) + L\|x_\lambda - \bar{x}\|^2 \\ &\leq \gamma \lambda \phi(\bar{x}) + (1 - \gamma \lambda) \phi(x) + \frac{\gamma \rho}{2} \lambda \left(1 - \frac{\lambda}{2 - \gamma} \right) \|x - \bar{x}\|^2 + L(1 - \lambda)^2 \|x - \bar{x}\|^2. \end{aligned}$$

Substitute $\phi(\bar{x}) = \psi(\bar{x})$ and $\phi(x) = \psi(x) - L\|x - \bar{x}\|^2$ into the inequality

$$\begin{aligned}\psi(x_\lambda) &\leq \gamma\lambda\psi(\bar{x}) + (1 - \gamma\lambda)(\psi(x) - L\|x - \bar{x}\|^2) + \left[\frac{\gamma\rho}{2}\lambda\left(1 - \frac{\lambda}{2 - \gamma}\right) + L(1 - \lambda)^2\right]\|x - \bar{x}\|^2 \\ &= \gamma\lambda\psi(\bar{x}) + (1 - \gamma\lambda)\psi(x) + B(\lambda)\|x - \bar{x}\|^2,\end{aligned}$$

where $B(\lambda)$ is the coefficient of $\|x - \bar{x}\|^2$. We simplify $B(\lambda)$ by expanding the terms involving L

$$\begin{aligned}B(\lambda) &= \frac{\gamma\rho}{2}\lambda\left(1 - \frac{\lambda}{2 - \gamma}\right) + L(1 - 2\lambda + \lambda^2 - 1 + \gamma\lambda) \\ &= \frac{\gamma\rho}{2}\lambda\left(1 - \frac{\lambda}{2 - \gamma}\right) - L\lambda(2 - \gamma)\left(1 - \frac{\lambda}{2 - \gamma}\right) \\ &= \frac{\gamma}{2}\left[\rho - 2L\left(\frac{2}{\gamma} - 1\right)\right]\lambda\left(1 - \frac{\lambda}{2 - \gamma}\right).\end{aligned}$$

For $\lambda \in [0, 1]$ and $\gamma \in]0, 1]$, the term $\lambda\left(1 - \frac{\lambda}{2 - \gamma}\right)$ is non-negative because $1 - \frac{\lambda}{2 - \gamma} \geq 1 - \frac{1}{2 - \gamma} = \frac{1 - \gamma}{2 - \gamma} \geq 0$. Therefore, to satisfy the definition of (ρ_L, γ) -weak quasar-convexity, we require

$$B(\lambda) \leq \frac{\gamma\rho_L}{2}\lambda\left(1 - \frac{\lambda}{2 - \gamma}\right).$$

This holds if and only if $\rho_L \geq \rho - 2L\left(\frac{2}{\gamma} - 1\right)$. Since the weak quasar-convexity modulus must be non-negative by definition, we take $\rho_L = \max\left\{0, \rho - 2L\left(\frac{2}{\gamma} - 1\right)\right\}$, which completes the proof. $\square$

Remark 36. Lemma 35 shows why a naive quadratic perturbation in the function of Example 3 cannot by itself yield an instance with $\mu > \rho$ while preserving failure of quasiconvexity: as L grows, ρ_L is driven toward 0 exactly in the regime needed for $\mu_L > \rho_L$, at which point the perturbed function becomes genuinely (star-)convex. This motivated the alternative, piecewise construction of Example 37. $\square$

Finally, we present an example in which Theorem 26 applies, and no other result in the literature does. To that end, we first show the following.

Example 37 (A Radial Function with $\mu > \rho$). Let $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ be the even function defined, for $t \geq 0$, by

$$\varphi(t) := \begin{cases} t^2, & 0 \leq t \leq 1, \\ 1 + 2(t - 1) - \frac{1}{8}(t - 1)^2, & 1 \leq t \leq 13, \\ 7 - (t - 13) + (t - 13)^2, & t \geq 13, \end{cases} \quad \varphi(-t) := \varphi(t). \quad (58)$$

Clearly, φ is continuously differentiable on $\mathbb{R}$. Indeed, at $t = 1$, both pieces give value 1 and slope 2, while at $t = 13$, both pieces give value 7 and slope -1 , and φ is smooth and even at $t = 0$. We now record the main properties of φ .

- (i) Unique global minimizer. $\varphi(0) = 0$, and $\varphi(t) > 0$ for every $t \neq 0$: on $[0, 1]$, $\varphi(t) = t^2 > 0$ for $t > 0$; on $[1, 13]$, φ increases from $\varphi(1) = 1$ to a strict local maximum $\varphi(9) = 9$ (since $\varphi'(t) = 2 - \frac{1}{4}(t - 1)$ vanishes exactly at $t = 9$), then decreases to $\varphi(13) = 7$; on $[13, +\infty[$, φ decreases further to a local minimum $\varphi(13.5) = \frac{27}{4}$ (since $\varphi'(t) = -1 + 2(t - 13)$ vanishes at $t = 13.5$), then increases to $+\infty$. In particular $\varphi(t) \geq \frac{27}{4} > 0$ for all $t \geq 13$. Hence, $\{0\} = \operatorname{argmin}_{\mathbb{R}} \varphi$.
- (ii) Global weak convexity. On $]0, 1[$, $\varphi''(t) = 2$; on $]1, 13[$, $\varphi''(t) = -\frac{1}{4}$; on $]13, +\infty[$, $\varphi''(t) = 2$, with the corresponding formulas on the negative half-line by evenness. Since $\varphi + \frac{1}{8}|\cdot|^2$ is continuously differentiable and piecewise quadratic with nonnegative second derivative on each quadratic piece, its derivative is nondecreasing. Hence it is convex. Thus φ is $\frac{1}{4}$ -weakly convex on $\mathbb{R}$. In particular, it is $(\rho, 1)$ -weakly quasar-convex with respect to 0 for every $\rho \geq \frac{1}{4}$, by (10).
- (iii) Local quadratic growth. Since $\varphi(t) = t^2$ exactly on $[-1, 1]$, and $\varphi(t) \geq 1$ for every t with $|t| \geq 1$ (as φ increases on $[1, 9]$ from 1, decreases on $[9, 13]$ to 7, and is bounded below by $\frac{27}{4} > 1$ on $[13, +\infty[$), the sublevel set $\{t \in \mathbb{R} : \varphi(t) \leq \zeta\}$ is contained in $[-1, 1]$ for every $\zeta \in]0, 1[$, and on that set $\varphi(t) = t^2$ exactly. Hence φ satisfies the μ -local quadratic growth condition (3) at $\bar{x} = 0$ with $\mu = 2$, for every $\zeta \in]0, 1[$.

(iv) Failure of (star) quasi-convexity. The point $t = 9$ lies on the segment joining 0 and 13.5, and $\varphi(9) = 9 > \frac{27}{4} = \max\{\varphi(0), \varphi(13.5)\}$, which directly violates quasi-convexity and, in particular, star quasi-convexity with respect to 0.

(v) Failure of quasar-convexity, for every $\gamma \in]0, 1]$. Taking $x = 13.5$ and $\lambda = \frac{1}{3}$ (so that $(1 - \lambda)x = 9$), γ -quasar-convexity with respect to 0 would require $\varphi(9) \leq (1 - \frac{\gamma}{3}) \varphi(13.5)$. Since $\varphi(13.5) = \frac{27}{4}$ and $1 - \frac{\gamma}{3} \leq 1$ for every $\gamma \in]0, 1]$, we get $(1 - \frac{\gamma}{3}) \varphi(13.5) \leq \frac{27}{4} = 6.75 < 9 = \varphi(9)$, a contradiction for every $\gamma \in]0, 1]$. Hence φ is not γ -quasar-convex with respect to 0, for any $\gamma \in]0, 1]$. $\square$

Now we present the desired example.

Example 38 (An Application for PPM). Let $f(s) = 2 - \sqrt{|s|}$ be as in Example 5, and let φ be the function of Example 37. Divide $\mathbb{R}^2$ into the four closed regions delimited by the lines $x_2 = x_1$ and $x_2 = -x_1$, namely

$$\begin{aligned} \mathcal{A} &:= \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \geq 0, -x_1 \leq x_2 \leq x_1\}, & \mathcal{B} &:= \{(x_1, x_2) \in \mathbb{R}^2 : x_2 \geq 0, -x_2 \leq x_1 \leq x_2\}, \\ \mathcal{C} &:= \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \leq 0, x_1 \leq x_2 \leq -x_1\}, & \mathcal{D} &:= \{(x_1, x_2) \in \mathbb{R}^2 : x_2 \leq 0, x_2 \leq x_1 \leq -x_2\}. \end{aligned}$$

Define $H : \mathbb{R}^2 \rightarrow \mathbb{R}$ by $H(0, 0) := 0$, and for $(x_1, x_2) \neq (0, 0)$,

$$H(x_1, x_2) := \begin{cases} \varphi(x_1) f\left(\frac{x_2}{x_1}\right), & (x_1, x_2) \in \mathcal{A}, \\ \varphi(x_2) f\left(\frac{x_1}{x_2}\right), & (x_1, x_2) \in \mathcal{B}, \\ \varphi(-x_1) f\left(\frac{x_2}{-x_1}\right), & (x_1, x_2) \in \mathcal{C}, \\ \varphi(-x_2) f\left(\frac{x_1}{-x_2}\right), & (x_1, x_2) \in \mathcal{D}. \end{cases}$$

The four formulas agree on the shared boundary rays $x_2 = \pm x_1$, so H is well defined and continuous on $\mathbb{R}^2$.

Unique global minimizer. Since $\varphi(s) = 0$ if and only if $s = 0$ by Example 37(i) and $f > 0$ on $[-1, 1]$, the same argument as in Example 5 shows $H(x) = 0$ if and only if $x = (0, 0)$, thus $\{(0, 0)\} = \operatorname{argmin}_{\mathbb{R}^2} H$.

Local quadratic growth with $\mu_H = 1$. On $\mathcal{A}$, using $\varphi(t) = t^2$ for $t \in [0, 1]$ and $f \geq 1$, we have

$$H(x_1, x_2) = \varphi(x_1) f\left(\frac{x_2}{x_1}\right) \geq \varphi(x_1) \geq x_1^2 \geq \frac{1}{2} \|x\|^2, \quad \forall (x_1, x_2) \in \mathcal{A}, \quad 0 < x_1 \leq 1,$$

using $|x_2| \leq x_1$ on $\mathcal{A}$. The same bound holds trivially at $(0, 0)$. On the other hand, for every $(x_1, x_2) \in \mathcal{A}$ with $x_1 \geq 1$ we have $H(x_1, x_2) \geq \varphi(x_1) \geq 1$, by Example 37(iii) and $f \geq 1$. Hence, for every $\zeta \in]0, 1[$, the sublevel set $\{H \leq \zeta\} \cap \mathcal{A}$ is contained in $\{x_1 < 1\}$, where the bound above holds.

Weak quasar-convexity with $\rho_H = \frac{1}{2}$. On any ray through the origin, $x_2/x_1 = c \in [-1, 1]$ constant, so restricted to that ray, $H(t, ct) = f(c)\varphi(t)$ is a positive scalar multiple of φ , with $f(c) \in [1, 2]$. Since φ is $\frac{1}{4}$ -weakly convex on $\mathbb{R}$ by Example 37(ii), multiplying by $f(c) \leq 2$ gives that $t \mapsto f(c)\varphi(t)$ is $\frac{1}{2}$ -weakly convex, hence $(\frac{1}{2}, 1)$ -weakly quasar-convex with respect to $t = 0$, for every $c \in [-1, 1]$. As in Example 5(i), this shows H is $(\frac{1}{2}, 1)$ -weakly quasar-convex with respect to $(0, 0)$ on $\mathbb{R}^2$.

Not weakly convex for any modulus. Fix $\rho \geq 0$ and, for $s \in]0, 1]$, consider $(1, -s), (1, s) \in \mathbb{R}^2$ and their midpoint $(1, 0)$. Since f is even, $H(1, \pm s) = \varphi(1)f(s)$ and $H(1, 0) = \varphi(1)f(0)$. As in Example 5(ii), weak convexity with modulus ρ would require $\varphi(1)\sqrt{s} \leq \frac{\rho}{2}s^2$, using $\varphi(1) = 1 > 0$ and $f(0) - f(s) = \sqrt{s}$; this fails as $s \rightarrow 0^+$ for any fixed $\rho \geq 0$. Hence, H is not weakly convex on $\mathbb{R}^2$ for any modulus.

Neither (star) quasi-convexity nor γ -quasar-convexity for any $\gamma \in]0, 1]$. It suffices to consider H restricted to $(t, 0)$, $t \geq 0$, where $H(t, 0) = f(0)\varphi(t) = 2\varphi(t)$. Using Example 37(iv)–(v), we get

$$H(0, 0) = 0, \quad H(9, 0) = 18, \quad H(13.5, 0) = \frac{27}{2}.$$

Since $H(9, 0) = 18 > \frac{27}{2} = \max\{H(0, 0), H(13.5, 0)\}$, and $(9, 0)$ lies on the segment joining $(0, 0)$ and $(13.5, 0)$, this violates quasi-convexity and star quasi-convexity with respect to $(0, 0)$ (see [20, Theorem

12]). Furthermore, taking $x = (13.5, 0)$, $\lambda = \frac{1}{3}$, the same computation as in Example 37(v) (scaled by the common factor 2) shows $H(9, 0) = 18 > (1 - \frac{\gamma}{3}) \cdot \frac{27}{2}$ for every $\gamma \in]0, 1]$, so H is not γ -quasar-convex with respect to $(0, 0)$, for any $\gamma \in]0, 1]$.

Therefore, H is $(\frac{1}{2}, 1)$ -weakly quasar-convex with respect to its unique global minimizer $\bar{x} = (0, 0)$, and satisfies the 1-local quadratic growth condition (3) at $\bar{x}$ for every $\zeta \in]0, 1[$, with

$$\mu_H = 1 > \frac{1}{2} = \rho_H,$$

yet H is neither weakly convex, nor quasi-convex, nor star quasi-convex, nor quasar-convex, for any modulus. As a consequence, provided the initial point x^0 of the PPM satisfies $H(x^0) < 1$ (so that the trajectory remains within the region of validity $\zeta \in]0, 1[$), Theorem 26 applies to H , while no other result in the literature does. $\square$

5 Conclusion

We introduced the class of weakly quasar-convex functions, unifying weak convexity and quasar-convexity under a single (ρ, γ) -parametrized inequality anchored at a global minimizer. We developed its basic properties and showed that the class strictly contains the union of its two parent notions. We further proved that the Moreau envelope preserves weak quasar-convexity with an explicit modulus update, yielding coercivity and characterization results for the associated proximity operator. Building on this machinery, we established Q -linear convergence of the proximal point algorithm under a local quadratic growth condition, together with matching iteration complexity and a computable, minimizer-free stopping rule based on the proximal residual.

6 Declarations

6.1 Availability of supporting data

No data sets were generated during the current study.

6.2 Author Contributions

All authors contributed equally to the study conception, design and implementation and wrote and corrected the manuscript.

6.3 Conflicts of Interest

There are no conflicts of interest or competing interests related to this manuscript.

6.4 Funding

A. Khademi acknowledges financial support from IVADO and the Canada First Research Excellence Fund, and F. Lara from ANID-Chile through project Fondecyt Regular 1241040.

References

- [1] M. Ahookhosh, J. M. de Brito, A. Kabgani, F. Lara, and J. Yuan. Quasar-convex optimization: Fundamental properties and high-order proximal-point methods. *arXiv preprint arXiv:2604.26735*, 2026.
- [2] H. H. Bauschke and P. L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. CMS Books in Mathematics. Springer International Publishing, Cham, 2017.
- [3] A. Beck, F. Lara, and C. Vega. Block descent type methods for quasar-convex minimization. *Submitted*, 2025.
- [4] C. Cartis, N. I. Gould, and P. L. Toint. *Evaluation Complexity of Algorithms for Nonconvex Optimization: Theory, Computation and Perspectives*. SIAM, Philadelphia, PA, USA, 2022.

- [5] Y. Chen and Q. Yang. A distributed accelerated algorithm based on a unified momentum method. *Journal of Optimization Theory and Applications*, 203(3):2908–2953, 2024.
- [6] F. H. Clarke. *Optimization and Nonsmooth Analysis*. SIAM, Philadelphia, 1990.
- [7] D. Davis and B. Grimmer. Proximally guided stochastic subgradient method for nonsmooth, non-convex problems. *SIAM Journal on Optimization*, 29(3):1908–1930, 2019.
- [8] J. de Brito, F. Lara, and D. Liu. Extending linear convergence of the proximal point algorithm: The quasar-convex case. *arXiv preprint arXiv:2509.04375*, 2025.
- [9] R. T. Farouki. The bernstein polynomial basis: A centennial retrospective. *Computer Aided Geometric Design*, 29(6):379–419, 2012.
- [10] Z. Gao, X. Sun, and L. He. Inertial dynamical systems and accelerated algorithms with implicit hessian-driven damping for nonconvex optimization. *Journal of Optimization Theory and Applications*, 210(1):21, 2026.
- [11] E. Gbaguidi and J. Hermant. Accelerated stochastic zeroth-order quasar-convex optimization. *arXiv preprint arXiv:2607.19965*, 2026.
- [12] A. Goujon, S. Neumayer, and M. Unser. Learning weakly convex regularizers for convergent image-reconstruction algorithms. *SIAM Journal on Imaging Sciences*, 17(1):91–115, 2024.
- [13] S.-M. Grad, F. Lara, and R. Tintaya Marcavillaca. Relaxed-inertial proximal point algorithms for nonconvex equilibrium problems with applications. *Journal of Optimization Theory and Applications*, 203(3):2233–2262, 2024.
- [14] S. Guminov, A. Gasnikov, and I. Kurozov. Accelerated methods for weakly quasi-convex optimization problems. *Computational Management Science*, 20(1):367, 2023.
- [15] M. Hardt, T. Ma, and B. Recht. Gradient descent learns linear dynamical systems. *Journal of Machine Learning Research*, 19(29):1–44, 2018.
- [16] O. Hinder, A. Sidford, and N. Sohoni. Near-optimal methods for minimizing star-convex functions and beyond. *arXiv preprint arXiv:1906.11985*, 2023. Version 3.
- [17] A. Kabgani and M. Ahookhosh. ItsOPT: An inexact two-level smoothing framework for nonconvex optimization via high-order Moreau envelope. *SIAM Journal on Optimization*, 2026. To appear.
- [18] A. Khademi. The convexity zoo: A taxonomy of function classes in optimization. *Optimization*, pages 1–52, 2026.
- [19] A. Khademi and A. Silveti-Falls. Adaptive conditional gradient descent. *Optimization Online*, 2025.
- [20] P. Khanh and F. Lara. Star quasiconvexity: a unified approach for linear convergence of first-order methods beyond convexity. *arXiv preprint arXiv: 2510.24981*, 2025.
- [21] P. D. Khanh, B. S. Mordukhovich, V. T. Phat, and D. B. Tran. Inexact proximal methods for weakly convex functions. *Journal of Global Optimization*, 91(3):611–646, 2025.
- [22] P. D. Khanh, B. S. Mordukhovich, and D. B. Tran. Convergence of first-order algorithms with momentum from the perspective of an inexact gradient descent method. *Journal of Optimization Theory and Applications*, 208(1):40, 2026.
- [23] D. La Vecchia and H. Liu. E-ROBOT: a dimension-free method for robust statistics and machine learning via schrödinger bridge. *Statistics and Computing*, 36(4):157, 2026.
- [24] F. Lara. On Debreu-Koopmans theorem for additively decomposed quasiconvex functions with applications. *arXiv preprint arXiv: 2603.17088*, 2026.
- [25] X. Li and D. Wu. Low-rank matrix recovery via nonconvex optimization methods with application to errors-in-variables matrix regression. *Journal of Optimization Theory and Applications*, 205(3):42, 2025.

- [26] S. Liang, R. Sun, and R. Srikant. Revisiting landscape analysis in deep neural networks: Eliminating decreasing paths to infinity. *SIAM Journal on Optimization*, 32(4):2797–2827, 2022.
- [27] F.-Y. Liao, L. Ding, and Y. Zheng. Error bounds, PL condition, and quadratic growth for weakly convex functions, and linear convergences of proximal point methods. *arXiv preprint arXiv:2312.16775*, 2023.
- [28] Y. Liu. On solving phase retrieval problems via neural networks: a unified framework with provable guarantees. *Journal on Advances in Signal Processing*, 2026(1):15, 2026.
- [29] S. López-Rivera, P. Pérez-Aros, and E. Vilches. A projected variable smoothing for weakly convex optimization and supremum functions. *Journal of Optimization Theory and Applications*, 208(3):115, 2026.
- [30] N. Nandhan, A. Khademi, and A. Silveti-Falls. Boosted stochastic Frank-Wolfe for constrained nonconvex optimization. *Optimization Online*, 2026.
- [31] Y. Nesterov and B. Polyak. Cubic regularization of Newton method and its global performance. *Mathematical Programming*, 108(1):177–205, 2006.
- [32] N. H. Nguyen, T. Q. Trinh, and P. T. Vuong. The projected heavy-ball method for quasar-monotone variational inequalities and quasar-convex optimization. *Journal of Optimization Theory and Applications*, 208(2):83, 2026.
- [33] E. A. Nurminskii. The quasigradient method for the solving of the nonlinear programming problems. *Cybernetics*, 9:145–150, 1973.
- [34] D. A. Quiroga and A. Kyriallidis. Using non-convex optimization in quantum process tomography: Factored gradient descent is tough to beat. In *2023 IEEE International Conference on Rebooting Computing (ICRC)*, pages 1–10. IEEE, 2023.
- [35] M. Rahimi, S. Ghaderi, Y. Moreau, and M. Ahookhosh. Projected subgradient methods for paraconvex optimization: Application to robust low-rank matrix recovery. *Journal of Global Optimization*, pages 1–40, 2026.
- [36] R. T. Rockafellar. Monotone operators and proximal point algorithms. *SIAM J. Control Optim.*, pages 877–898, 1976.
- [37] A. Sadiev, A. Beznosikov, A. J. Almansoori, D. Kamzolov, R. Tappenden, and M. Takáč. Stochastic gradient methods with preconditioned updates. *Journal of Optimization Theory and Applications*, 201(2):471–489, 2024.
- [38] A. Silveti-Falls, C. Molinari, and Z. Woodstock. Frank-Wolfe with Moreau envelope smoothing for nonsmooth nonconvex problems. *arXiv preprint arXiv:2606.00853*, 2026.
- [39] S. Takahashi, S. Pokutta, and A. Takeda. Fast Frank-Wolfe algorithms with adaptive Bregman step-size for weakly convex functions. In *International Conference on Learning Representations*, volume 2026, pages 78614–78651, 2026.
- [40] O. S. Valdivia Viscarra and F. Stonyakin. On the convergence of first-order methods for quasar convex functions. In *International Conference on Optimization and Applications*, pages 88–103. Springer, 2025.
- [41] W. van Ackooij, F. Atenas, and C. Sagastizábal. Weak convexity and approximate subdifferentials. *Journal of Optimization Theory and Applications*, 203(2):1686–1709, 2024.
- [42] J.-K. Wang and A. Wibisono. Continuized acceleration for quasar convex functions in non-convex optimization. *arXiv preprint arXiv: 2302.07851*, 2023.