Full Convergence of Regularized Methods for Unconstrained Optimization

Andrea Cristofari*

*Department of Civil Engineering and Computer Science Engineering
University of Rome "Tor Vergata"
Via del Politecnico, 1, 00133 Rome (Italy)
E-mail: andrea.cristofari@uniroma2.it

Abstract. Typically, the sequence of points generated by an optimization algorithm may have multiple limit points. Under convexity assumptions, however, (sub)gradient methods are known to generate a convergent sequence of points. In this paper, we extend the latter property to a broader class of algorithms. Specifically, we study unconstrained optimization methods that use local quadratic models regularized by a power $r \geq 3$ of the norm of the step. In particular, we focus on the case where only the objective function and its gradient are evaluated. Our analysis shows that, by a careful choice of the regularized model at every iteration, the whole sequence of points generated by this class of algorithms converges if the objective function is pseudoconvex. The result is achieved by employing appropriate matrices to ensure that the sequence of points is variable metric quasi-Fejér monotone.

Keywords. Regularized methods. Full convergence. Variable metric quasi-Fejér monotonicity.

1 Introduction

Let us consider the following unconstrained optimization problem:

$$\min_{x \in \mathbb{R}^n} f(x), \tag{1}$$

where the objective function $f \colon \mathbb{R}^n \to \mathbb{R}$ has a continuous gradient $\nabla f(x)$. Among the huge number of algorithms proposed in the literature to solve problem (1), a popular approach is represented by the class of *regularized methods*, which are characterized by the use of suitable models combined with a regularization term. In particular, here we consider *quadratic models* regularized by a power r of the norm of the step. Namely, to update a point x_k generated at iteration k, we use the following regularized model:

$$m_k(s) := f(x_k) + \nabla f(x_k)^T s + \frac{1}{2} s^T Q_k s + \frac{\sigma_k}{r} \|s\|^r, \tag{2}$$

with symmetric matrix $Q_k \in \mathbb{R}^{n \times n}$ and regularization parameter $\sigma_k \geq 0$, while, here and in the rest of the paper, $\|\cdot\|$ denotes the Euclidean norm. Then, a step s_k is computed as

an approximate minimizer of m_k in order to set the next iterate as $x_{k+1} = x_k + s_k$ if an appropriate objective decrease is achieved.

In this paper, we focus on the case where *only the objective function and its gradient are evaluated* at each iteration. In such a setting, convergence and complexity results of regularized methods can be found in [6, 7, 8, 9].

Our analysis shows that, when f is pseudoconvex and $r \geq 3$, *the whole sequence* $\{x_k\}$ *generated by the regularized methods converges* if an appropriate sequence of matrices $\{Q_k\}$ is employed. Specifically, $\{Q_k\}$ must be chosen so as to guarantee that $\{x_k\}$ is a *variable metric quasi-Fejér monotone* sequence.

In the literature, it is known that the whole sequence of points converges, under convexity assumptions, for (sub)gradient methods [1, 3, 5, 15, 18, 21]. Our result hence extends that property to a broader class of regularized methods. Interestingly, our analysis also shows that the convergence of the algorithm does not require the assumption of bounded level sets, which is typically needed to obtain the $\mathcal{O}(1/k)$ convergence rate of regularized methods using first-order information in the convex case [8].

Let us also remark that the matrix Q_k used in model (2) is not required to approximate the Hessian (which may not exist in our setting, since we only assume access to f and ∇f). This differentiates our approach from methods that explicitly employ Hessian approximations, e.g., via finite differences [14].

The rest of the paper is organized as follows. In Section 2, we review the notion of Fejér monotonicity and its extensions. In Section 3, we describe the class of regularized methods under analysis. In Section 4, we give the main result of the paper showing the convergence of the whole sequence of points generated by the considered class of algorithms. In Section 5, we provide an example of a regularized method that generates a sequence of points converging to a minimizer of the objective function. In Section 6, we describe an iterative procedure to compute appropriate matrices that satisfy the conditions required for the considered class of algorithms. In Section 7, we analyze the practical performance of the proposed algorithmic scheme under different choices for the sequence of matrices. Finally, some conclusions are drawn in Section 8.

2 Fejér Monotonicity and Extensions

Let us consider a sequence $\{x_k\} \subseteq \mathbb{R}^n$ generated by an optimization algorithm. A common tool to prove the convergence of $\{x_k\}$ comes from the notion of *Fejér monotonicity*, defined as follows.

Definition 1 (Fejér monotonicity)**.** *Given a non-empty set $\mathcal{F} \subseteq \mathbb{R}^n$, we say that $\{x_k\} \subseteq \mathbb{R}^n$ is a Fejér monotone sequence with respect to $\mathcal{F}$ if, for all $y \in \mathcal{F}$, it holds that*

$$\|x_{k+1} - y\|^2 \leq \|x_k - y\|^2 \quad \forall k \geq 0.$$

It is well known that, if $\{x_k\}$ is a Fejér monotone sequence with respect to $\mathcal{F}$ and $\{x_k\}$ has a limit point $x^* \in \mathcal{F}$, then the whole sequence $\{x_k\}$ converges to x^* (see, e.g., [3]).

In the literature, Fejér monotonicity has been used to show the convergence of some optimization algorithms under appropriate assumptions. In particular, when the objective

function f is convex with Lipschitz continuous gradient, the sequence $\{x_k\}$ generated by the gradient method is Fejér monotone if the stepsize is chosen within an appropriate interval, hence converging to a minimizer of f [18]. For a non-smooth convex objective function f, also the subgradient method produces a Fejér monotone sequence $\{x_k\}$ when using the Polyak stepsize, thus converging to a minimizer of f [21].

Some extensions of Fejér monotonicity have been proposed in the literature in order to guarantee the same convergence properties for a broader family of sequences. A first extension is represented by *quasi-Fejér monotonicity*, which is defined below.

Definition 2 (quasi-Fejér monotonicity)**.** *Given a non-empty set $\mathcal{F} \subseteq \mathbb{R}^n$, we say that $\{x_k\} \subseteq \mathbb{R}^n$ is a quasi-Fejér monotone sequence with respect to $\mathcal{F}$ if, for all $y \in \mathcal{F}$, it holds that*

$$\|x_{k+1} - y\|^2 \leq \|x_k - y\|^2 + \epsilon_k \quad \forall k \geq 0,$$

where $\{\epsilon_k\}$ is a sequence of scalars such that

$$\{\epsilon_k\} \subseteq [0, \infty), \quad \sum_{k=0}^{\infty} \epsilon_k < \infty.$$

Using quasi-Fejér monotonicity, it is still possible to show that, under convexity assumptions, the whole sequence generated by gradient methods converges to a minimizer of the objective function f using suitable line search techniques (which include the classical Armijo line search) [5, 15]. In the non-smooth case, also subgradient methods with a diminishing stepsize generate a quasi-Fejér monotone sequence, under convexity assumptions, hence converging to a minimizer of the objective function [1, 3]. Further, quasi-Fejér monotonicity has been used even in vector optimization to show the convergence of the whole sequence of points generated by the steepest descent method under convexity assumptions [13].

Additionally, a more general definition of quasi-Fejér monotonicity was given in [16] using a divergence measure in place of the quadratic norm appearing in Definition 2, whereas a comparison between different types of quasi-Fejér monotonicity and an analysis of their use in optimization can be found in [11] .

Going beyond Definition 2, a further extension has been proposed in [12] with the notion of *variable metric quasi-Fejér monotonicity* which employs variable metrics to measure the distance between points. Using that tool, weak and strong convergence results for sequences in a real Hilbert space were established in [12].

In the following, we report a simplified definition of variable metric quasi-Fejér monotonicity adapted from [12] to our purposes, using, here and in the rest of the paper, the notation $\|Q\|$ to denote the norm induced by the vector Euclidean norm for any matrix Q, while $\|\cdot\|_Q$ indicates the norm induced by a symmetric positive definite matrix $Q \in \mathbb{R}^{n \times n}$, that is, $\|v\|_Q = (v^T Q v)^{1/2}$ for all $v \in \mathbb{R}^n$.

Definition 3 (variable metric quasi-Fejér monotonicity)**.** *Given a non-empty set $\mathcal{F} \subseteq \mathbb{R}^n$ and a sequence of symmetric matrices $\{Q_k\} \subseteq \mathbb{R}^{n \times n}$ such that*

$$Q_k \succeq aI \quad \forall k \geq 0, \quad \textit{with} \quad a > 0, \tag{3}$$

$$\|Q_k\| \leq b \quad \forall k \geq 0, \quad \textit{with} \quad b < \infty, \tag{4}$$

we say that $\{x_k\} \subseteq \mathbb{R}^n$ *is a variable metric quasi-Fejér monotone sequence with respect to* $\mathcal{F}$ *relative to* $\{Q_k\}$ *if, for all* $y \in \mathcal{F}$*, it holds that*

$$\|x_{k+1} - y\|^2_{Q_{k+1}} \leq (1 + \psi_k)\|x_k - y\|^2_{Q_k} + \epsilon_k \quad \forall k \geq 0, \tag{5}$$

where $\{\psi_k\}$ *and* $\{\epsilon_k\}$ *are two sequences of scalars such that*

$$\{\psi_k\} \subseteq [0, \infty), \quad \sum_{k=0}^{\infty} \psi_k < \infty, \tag{6}$$

$$\{\epsilon_k\} \subseteq [0, \infty), \quad \sum_{k=0}^{\infty} \epsilon_k < \infty. \tag{7}$$

Variable metric quasi-Fejér monotonicity will be central to our analysis. Specifically, we will use the following result to ensure the convergence of a sequence of points.

Theorem 1. *Let* $\{x_k\}$ *be a variable metric quasi-Fejér monotone sequence with respect to a set* $\mathcal{F}$ *relative to a sequence of matrices* $\{Q_k\}$*, according to Definition 3. The following holds:*

(i) $\|x_k - y\| \leq R$ *for all* $k \geq 0$ *and all* $y \in \mathcal{F}$*, where* $R := \left(\frac{1}{a}\prod_{j=0}^{\infty}(1+\psi_j)\left(\|x_0 - y\|^2_{Q_0} + \sum_{i=0}^{\infty}\epsilon_i\right)\right)^{1/2} < \infty$ *with* a*,* $\{\psi_k\}$ *and* $\{\epsilon_k\}$ *given in Definition 3;*

(ii) $\{x_k\}$ *is bounded;*

(iii) if $x^* \in \mathcal{F}$ *is a limit point of* $\{x_k\}$*, then* $\{x_k\}$ *converges to* x^**.*

Proof. Proof Fix any $y \in \mathcal{F}$. Let a, b, $\{\psi_k\}$ and $\{\epsilon_k\}$ be as given in Definition 3. Since $\{\psi_k\}$ satisfies (6), we can define

$$1 \leq \zeta := \prod_{i=0}^{\infty}(1 + \psi_i) < \infty. \tag{8}$$

Applying the variable metric quasi-Fejér inequality (5) recursively from an index k to an index $t \leq k$, we have that

$$\|x_k - y\|^2_{Q_k} \leq \prod_{i=t}^{k-1}(1 + \psi_i)\|x_t - y\|^2_{Q_t} + \sum_{i=t}^{k-2}\prod_{j=i+1}^{k-1}(1 + \psi_j)\epsilon_i + \epsilon_{k-1} \quad \forall k \geq t \geq 0.$$

Using (8) together with the fact that $\{\psi_k\}$ and $\{\epsilon_k\}$ are sequences of non-negative real numbers from (6) and (7), respectively, it follows that

$$\|x_k - y\|^2_{Q_k} \leq \zeta\|x_t - y\|^2_{Q_t} + \zeta\sum_{i=t}^{k-2}\epsilon_i + \epsilon_{k-1} \leq \zeta\left(\|x_t - y\|^2_{Q_t} + \sum_{i=t}^{\infty}\epsilon_i\right) \quad \forall k \geq t \geq 0.$$

Since, from (3), we have $\|x_k - y\|^2_{Q_k} \geq a\|x_k - y\|^2$ for all $k \geq 0$, we get

$$\|x_k - y\|^2 \leq \frac{\zeta}{a}\left(\|x_t - y\|^2_{Q_t} + \sum_{i=t}^{\infty} \epsilon_i\right) < \infty \quad \forall k \geq t \geq 0, \tag{9}$$

where the last inequality follows from the fact that $\{\epsilon_k\}$ satisfies (7) together with the fact that ζ/a is finite from (8) and (3). Hence, items (i)–(ii) hold.

Now, let $x^* \in \mathcal{F}$ be a limit point of $\{x_k\}$, that is, there exists a subsequence $\{x_k\}_K \to x^*$, with $K \subseteq \{0, 1, \ldots\}$. Using (7), for any $\omega > 0$ there exists an index k_ω such that

$$\|x_{k_\omega} - x^*\|^2 \leq \frac{a\omega}{2b\zeta},$$
$$\sum_{i=k_\omega}^{\infty} \epsilon_i \leq \frac{a\omega}{2\zeta}.$$

Hence, using (9) with $y = x^*$ and $t = k_\omega$, for all $k \geq k_\omega$ we have that

$$\begin{aligned}
\|x_k - x^*\|^2 \leq \frac{\zeta}{a}\left(\|x_{k_\omega} - x^*\|^2_{Q_{k_\omega}} + \sum_{i=k_\omega}^{\infty} \epsilon_i\right) &\leq \frac{\zeta}{a}\left(\|Q_{k_\omega}\|\|x_{k_\omega} - x^*\|^2 + \sum_{i=k_\omega}^{\infty} \epsilon_i\right) \\
&\leq \frac{\zeta}{a}\left(b\|x_{k_\omega} - x^*\|^2 + \sum_{i=k_\omega}^{\infty} \epsilon_i\right) \\
&\leq \frac{\zeta}{a}\left(\frac{a\omega}{2\zeta} + \frac{a\omega}{2\zeta}\right) \\
&= \omega,
\end{aligned}$$

where we have used (4) in the third inequality. Since this is true for any arbitrary $\omega > 0$, we conclude that $\{x_k\}$ converges to x^*, thus proving item (iii). □

In the next lemma, we give necessary and sufficient conditions for $\{x_k\}$ to be a variable metric quasi-Fejér monotone sequence.

Lemma 1. *Let $\mathcal{F} \subseteq \mathbb{R}^n$ be a non-empty set and let $\{Q_k\} \subseteq \mathbb{R}^{n\times n}$ be a sequence of symmetric matrices satisfying (3)–(4). Then, $\{x_k\}$ is a variable metric quasi-Fejér monotone sequence with respect to $\mathcal{F}$ relative to $\{Q_k\}$, according to Definition 3, if and only if, for all $y \in \mathcal{F}$, it holds that*

$$\|x_{k+1} - y\|^2_{Q_{k+1}} \leq (1 + \psi_k)\|x_k - y\|^2_{Q_k} + \theta_k\|x_k - y\| + \epsilon_k \quad \forall k \geq 0, \tag{10}$$

where $\{\psi_k\}$ and $\{\epsilon_k\}$ satisfy (6) and (7), respectively, while $\{\theta_k\}$ is a sequence of scalars such that

$$\{\theta_k\} \subseteq [0, \infty), \quad \sum_{k=0}^{\infty} \theta_k < \infty. \tag{11}$$

Proof. Proof Take any $y \in \mathcal{F}$ and let us distinguish the necessary and sufficient implications separately.

- Necessity. If Definition 3 is satisfied, then (10)–(11) hold by using $\theta_k = 0$ for all $k \geq 0$.

- Sufficiency. Assume that (10)–(11) hold for a given $y \in \mathcal{F}$. We can partition the index set $\{0, 1, \ldots\}$ into two subsets K_1 and K_2 such that

$$\begin{aligned} k \in K_1 \quad &\Leftrightarrow \quad \|x_k - y\| < 1, \\ k \in K_2 \quad &\Leftrightarrow \quad \|x_k - y\| \geq 1. \end{aligned}$$

Observe that, in view of (3), we can write

$$\|x_k - y\|_{Q_k}^2 \geq a\|x_k - y\|^2 \geq a\|x_k - y\| \quad \forall k \in K_2.$$

Then, using (10), we get

$$\begin{aligned} \|x_{k+1} - y\|_{Q_{k+1}}^2 &\leq (1 + \psi_k)\|x_k - y\|_{Q_k}^2 + \theta_k + \epsilon_k \quad \forall k \in K_1, \\ \|x_{k+1} - y\|_{Q_{k+1}}^2 &\leq \left(1 + \psi_k + \frac{\theta_k}{a}\right)\|x_k - y\|_{Q_k}^2 + \epsilon_k \quad \forall k \in K_2. \end{aligned}$$

Since $a > 0$ and $\theta_k \geq 0$ for all $k \geq 0$, it follows that

$$\begin{aligned} \|x_{k+1} - y\|_{Q_{k+1}}^2 &\leq \left(1 + \psi_k + \frac{\theta_k}{a}\right)\|x_k - y\|_{Q_k}^2 + \theta_k + \epsilon_k \quad \forall k \in K_1, \\ \|x_{k+1} - y\|_{Q_{k+1}}^2 &\leq \left(1 + \psi_k + \frac{\theta_k}{a}\right)\|x_k - y\|_{Q_k}^2 + \theta_k + \epsilon_k \quad \forall k \in K_2. \end{aligned}$$

Namely, recalling that $K_1 \cup K_2 = \{0, 1, \ldots\}$, we have that

$$\|x_{k+1} - y\|_{Q_{k+1}}^2 \leq (1 + \tilde{\psi}_k)\|x_k - y\|_{Q_k}^2 + \tilde{\epsilon}_k \quad \forall k \geq 0,$$

where

$$\tilde{\psi}_k = \psi_k + \frac{\theta_k}{a} \quad \text{and} \quad \tilde{\epsilon}_k = \epsilon_k + \theta_k.$$

Using (6), (7) and (11), it follows that $\tilde{\psi}_k$ and $\tilde{\epsilon}_k$ are two sequences of scalars such that

$$\begin{aligned} \tilde{\psi}_k \subseteq [0, \infty), \quad \sum_{k=0}^{\infty} \tilde{\psi}_k < \infty, \\ \tilde{\epsilon}_k \subseteq [0, \infty), \quad \sum_{k=0}^{\infty} \tilde{\epsilon}_k < \infty. \end{aligned}$$

Hence, $\{x_k\}$ is a variable metric quasi-Fejér monotone sequence with respect to $\mathcal{F}$ relative to $\{Q_k\}$, according to Definition 3

□

3 Regularized Methods

In this section, we describe a simple framework for regularized methods which will be the subject of our analysis. The scheme we consider here is intentionally given *as general as possible* to include only the ingredients needed to ensure the convergence of the generated sequence $\{x_k\}$ to a point $x^* \in \mathbb{R}^n$. Other properties, such as the conditions under which x^* is a minimizer of f and the convergence rate, will be analyzed in the next section.

Given a regularization power $r \geq 3$, at every iteration k we consider the model m_k defined as in (2) by choosing appropriate regularization parameter $\sigma_k \geq 0$ and symmetric matrix $Q_k \in \mathbb{R}^{n \times n}$. For the sake of convenience, let us also define the quadratic part in the regularized model as

$$q_k(s) := f(x_k) + \nabla f(x_k)^T s + \frac{1}{2} s^T Q_k s, \tag{12}$$

so that, from (2) and (12), we can write

$$m_k(s) = q_k(s) + \frac{\sigma_k}{r} \|s\|^r.$$

In order to get the next point x_{k+1}, we compute s_k as an inexact minimizer of $m_k(s)$. In particular, here we require s_k to satisfy an approximate first-order condition, that is,

$$\|\nabla m_k(s_k)\| \leq \tau \|s_k\| \min\{\|s_k\|, 1\}, \tag{13}$$

with $\tau \geq 0$. Note that, in the considered setting where $r \geq 3$, the function $m_k(s)$ is coercive and has a global minimizer. Consequently, a point s_k satisfying (13) can be obtained in finite time by applying a descent method for unconstrained optimization.

Then, the next point x_{k+1} is obtained by means of an update rule ensuring that

$$x_{k+1} \neq x_k \quad \Rightarrow \quad f(x_k) - f(x_{k+1}) \geq \eta(m_k(0) - m_k(s_k)) > 0, \tag{14}$$

with $\eta > 0$. We see that, by the update rule (14), we can set $x_{k+1} = x_k + s_k$ only if the objective decrease given in (14) holds, otherwise we have to set $x_{k+1} = x_k$. Then, we say that k is a *successful iteration* if $x_{k+1} = x_k + s_k$, denoting the set of successful iterations by $\mathcal{S}$. Namely,

$$k \in \mathcal{S} \quad \Leftrightarrow \quad x_{k+1} = x_k + s_k. \tag{15}$$

The scheme is reported in Algorithm 1.

As mentioned at the beginning of this section, Algorithm 1 is a general framework in which a number of details are intentionally left unspecified. Namely, besides the choice of Q_k—which will be discussed in the next subsection—two main issues should be addressed: how to compute σ_k and how to guarantee the objective decrease condition (14).

Regarding the choice of σ_k, various strategies can be found in the literature. Common approaches make use of adaptive rules which dynamically update σ_k during the iterations in a trust-region fashion (see, e.g., [6]). Moreover, in Section 5, we will describe an algorithm where σ_k can assume any value in $[0, \sigma_{\max}]$ with an appropriate choice of the matrix Q_k under Lipschitz continuity of ∇f.

Algorithm 1 General scheme of a regularized method

given $r \in [3, \infty)$, $\tau \in [0, \infty)$, $\eta > 0$ and $\sigma_{\max} \in [0, \infty)$
choose $x_0 \in \mathbb{R}^n$
for $k = 0, 1, \ldots$ **do**
    compute $\sigma_k \in [0, \sigma_{\max}]$
    compute a symmetric matrix $Q_k \in \mathbb{R}^{n \times n}$
    compute $s_k \in \mathbb{R}^n$ satisfying (13)
    set either $x_{k+1} = x_k + s_k$ or $x_{k+1} = x_k$ in order to satisfy (14)
end for

To guarantee the objective decrease condition (14), usually $f(x_k+s_k)-f(x_k)$ is compared to the decrease achieved by either the regularized model m_k or the quadratic model q_k, in order to decide whether to set $x_{k+1} = x_k + s_k$ or $x_{k+1} = x_k$. In more detail, a first possibility [6, 9] is to compute

$$\rho_k = \frac{f(x_k) - f(x_k + s_k)}{m_k(0) - m_k(s_k)} \tag{16}$$

and set $x_{k+1} = x_k + s_k$ if $\rho_k \geq \eta$. By requiring s_k to be such that $m_k(0) - m_k(s_k) > 0$, it is straightforward to see that this update rule satisfies (14). A second possibility [4, 10] is to use

$$\rho_k = \frac{f(x_k) - f(x_k + s_k)}{q_k(0) - q_k(s_k)} \tag{17}$$

and set $x_{k+1} = x_k + s_k$ if $\rho_k \geq \eta$. By requiring s_k to be such that $q_k(0) - q_k(s_k) > 0$, also this update rule satisfies (14) since

$$q_k(0) - q_k(s_k) = m_k(0) - m_k(s_k) + \frac{\sigma_k}{r}\|s_k\|^r \geq m_k(0) - m_k(s_k).$$

Hence, by means of the general condition (14) in Algorithm 1, all the above update rules are included in our analysis.

Also note that Algorithm 1 encompasses the classical gradient method defined by the kth iteration $x_{k+1} = x_k - \alpha_k \nabla f(x_k)$, with the stepsize $\alpha_k > 0$ chosen by an Armijo-type line search such that $f(x_{k+1}) \leq f(x_k) - \gamma\alpha_k\|\nabla f(x_k)\|^2$, where $\gamma \in (0, 1)$ (see, e.g., [2]). Specifically, we can use $\tau = 0$ while setting $\sigma_k = 0$ and $Q_k = \alpha_k^{-1} I$ for every iteration k. It is straightforward to verify that, for every iteration k, such a parameter choice leads to $s_k = -\alpha_k \nabla f(x_k)$ and $m_k(0) - m_k(s_k) = (\alpha_k/2)\|\nabla f(x_k)\|^2$, implying that (14) is satisfied with $\eta = 2\gamma$. If ∇f is Lipschitz continuous with constant L, then we can also use, for every iteration k, a constant stepsize $\alpha_k = \alpha \in (0, 2(1-\gamma)/L]$ as it still guarantees the objective decrease required by the Armijo-type line search [2].

4 Analysis of Full Convergence

In this section, we will show that the whole sequence of points $\{x_k\}$ generated by Algorithm 1 converges under appropriate hypotheses.

The key element in our analysis is an adequate choice of $\{Q_k\}$, reported in Condition 1 below, to guarantee that $\{x_k\}$ is a variable metric quasi-Fejér monotone sequence, according to Definition 3, provided that f is pseudoconvex. This, in view of Theorem 1, will allow us to state the convergence of $\{x_k\}$.

Condition 1. *The sequence of matrices $\{Q_k\}$ used in Algorithm 1 is such that, for every iteration $k \geq 0$, it holds that*

$$Q_k \succeq aI, \tag{18a}$$

$$Q_{k+1} \preceq (1+\psi_k)Q_k, \tag{18b}$$

where $\{\psi_k\}$ satisfies (6) *and*

$$a > 2\tau. \tag{19}$$

From Condition 1, we see that (18a) and (19) impose that all the eigenvalues of Q_k are greater than 2τ, where τ is the parameter used in (13) to control the degree of approximation when computing s_k as an inexact minimizer of m_k. Moreover, (18b) relates two consecutive matrices and guarantees that

$$\|v\|^2_{Q_{k+1}} \leq (1+\psi_k)\|v\|^2_{Q_k} \quad \forall v \in \mathbb{R}^n. \tag{20}$$

Remark 1. *The easiest way to satisfy Condition 1 is choosing a matrix $Q \succ 2\tau I$ and setting $Q_k = Q$ for all $k \geq 0$. A more elaborate iterative procedure to compute $\{Q_k\}$ will be presented in Section 6.*

4.1 Preliminary Results

Let us provide some preliminary results which will be useful to get the main result in the next subsection. First note that, for every iteration k, we can express the gradient of the regularized model (2) as follows:

$$\nabla m_k(s) = \nabla f(x_k) + Q_k s + \sigma_k \|s\|^{r-2} s. \tag{21}$$

To begin with, in the following lemma we show that the sequence of matrices $\{Q_k\}$ is bounded.

Lemma 2. *Let $\{Q_k\}$ be the sequence of matrices used in Algorithm 1. If Condition 1 is satisfied, then $b \in \mathbb{R}$ exists such that*

$$\|Q_k\| \leq b \quad \forall k \geq 0.$$

Proof. Proof Fix any iteration $k \geq 1$. Since $Q_k \in \mathbb{R}^{n\times n}$ is symmetric and positive definite from Condition 1, we can define v_k as a unit-norm eigenvector of Q_k associated to the largest eigenvalue of Q_k. Namely,

$$v_k^T Q_k v_k = \|Q_k\|, \quad \|v_k\| = 1. \tag{22}$$

Using (22) and (20), we can write

$$\|Q_k\| = v_k^T Q_k v_k = \|v_k\|^2_{Q_k} \leq (1+\psi_{k-1})\|v_k\|^2_{Q_{k-1}} = (1+\psi_{k-1}) v_k^T Q_{k-1} v_k. \tag{23}$$

Since $\{\psi_k\}$ satisfies (6) in view of Condition 1, we can define

$$1 \leq \zeta := \prod_{i=0}^{\infty}(1+\psi_i) < \infty$$

and, applying (23) recursively, we get

$$\|Q_k\| \leq \prod_{i=0}^{k-1}(1+\psi_i) v_k^T Q_0 v_k \leq \zeta \|Q_0\|,$$

where, in the last inequality, we have used the fact that $\|v_k\| = 1$ from (22). Then, the desired result holds with $b = \zeta\|Q_0\|$. □

Now, we show how we can lower bound the decrease in the regularized model and in the objective function.

Lemma 3. *If Condition 1 is satisfied, then, for every iteration k of Algorithm 1, we have*

(i) $m_k(0) - m_k(s_k) \geq c\|s_k\|^2$,

(ii) $f(x_k) - f(x_{k+1}) \geq \eta c \|x_{k+1} - x_k\|^2$,

where

$$c := \Big(\frac{1}{2}a - \tau\Big) > 0.$$

Proof. Proof Fix any iteration $k \geq 0$. Recalling the expression of ∇m_k given in (21), we can write

$$\|\nabla m_k(s_k)\|\|s_k\| \geq \nabla m_k(s_k)^T s_k = \nabla f(x_k)^T s_k + s_k^T Q_k s_k + \sigma_k \|s_k\|^r. \tag{24}$$

Using the approximate optimality condition on s_k given (13), we also have

$$\|\nabla m_k(s_k)\| \leq \tau \|s_k\| \min\{\|s_k\|, 1\} \leq \tau \|s_k\|. \tag{25}$$

From (24)–(25), it follows that

$$-\nabla f(x_k)^T s_k \geq s_k^T Q_k s_k + \sigma_k \|s_k\|^r - \tau \|s_k\|^2. \tag{26}$$

Hence, recalling the expression of m_k given in (2), we get

$$\begin{aligned} m_k(0) - m_k(s_k) &= -\nabla f(x_k)^T s_k - \frac{1}{2} s_k^T Q_k s_k - \frac{\sigma_k}{r}\|s_k\|^r \\ &\geq \frac{1}{2} s_k^T Q_k s_k - \tau \|s_k\|^2 + \frac{\sigma_k (r-1)}{r}\|s_k\|^r \\ &\geq \Big(\frac{1}{2}a - \tau\Big)\|s_k\|^2 + \frac{\sigma_k (r-1)}{r}\|s_k\|^r, \end{aligned}$$

where we have used (26) in the first inequality and (18a) from Condition 1 in the second inequality. Since, from the instructions of Algorithm 1, we have $\sigma_k \geq 0$ and $r \geq 3$, then item (i) is proved.

Item (ii) follows from item (i) and the objective decrease condition (14). Finally, note that $c > 0$ follows from (19) in Condition 1. □

An immediate consequence of the above lemma is the monotonicity of $\{f(x_k)\}$.

Lemma 4. *Let $\{x_k\}$ be the sequence generated by Algorithm 1. If Condition 1 is satisfied, then $\{f(x_k)\}$ is monotonically non-increasing.*

Proof. Proof It follows from item (ii) of Lemma 3 recalling, from the instructions of Algorithm 1, that $\eta > 0$. □

When f is bounded from below, we show in the following two lemmas that $\{\|s_k\|^p\}_{k\in\mathcal{S}}$ is summable for any $p \geq 2$, and consequently, thanks to the approximate optimality condition on s_k given in (13), $\{\|\nabla m_k(s_k)\|\}_{k\in\mathcal{S}}$ is also summable.

Lemma 5. *Let $\{x_k\}$ be the sequence generated by Algorithm 1. If Condition 1 is satisfied and f is bounded from below, then there exists $T := \max_{k\in\mathcal{S}}\|s_k\| < \infty$ and*

$$\sum_{k\in\mathcal{S}}\|s_k\|^p < \infty \quad \forall p \geq 2.$$

Proof. Proof The result is straightforward if $\mathcal{S}$ is a finite set, so let us consider the case where $\mathcal{S}$ is an infinite set. From item (ii) of Lemma 3, we can write

$$f(x_0) - \inf_{x\in\mathbb{R}^n} f(x) \geq f(x_0) - f(x_h) = \sum_{k=0}^{h-1}(f(x_k) - f(x_{k+1})) \geq \eta c \sum_{k=0}^{h-1}\|x_{k+1} - x_k\|^2 \quad \forall h \geq 0.$$

Letting $h \to \infty$, we obtain

$$f(x_0) - \inf_{x\in\mathbb{R}^n} f(x) \geq \eta c \sum_{k=0}^{\infty}\|x_{k+1} - x_k\|^2.$$

Since $\inf_{x\in\mathbb{R}^n} f(x)$ is finite by assumption and recalling, from the instructions of Algorithm 1 and Lemma 3, that $\eta c > 0$, we get

$$\sum_{k=0}^{\infty}\|x_{k+1} - x_k\|^2 < \infty.$$

Hence, from the instructions of Algorithm 1 and the definition of $\mathcal{S}$ given in (15), it follows that

$$\sum_{k=0}^{\infty}\|x_{k+1} - x_k\|^2 = \sum_{k\in\mathcal{S}}\|x_{k+1} - x_k\|^2 = \sum_{k\in\mathcal{S}}\|s_k\|^2 < \infty.$$

This implies that $\{\|s_k\|\}_{k\in\mathcal{S}} \to 0$ and we can define $T < \infty$ as in the assertion, leading to

$$\sum_{k\in\mathcal{S}}\|s_k\|^p = \sum_{k\in\mathcal{S}}\|s_k\|^2\|s_k\|^{p-2} \leq T^{p-2}\sum_{k\in\mathcal{S}}\|s_k\|^2 < \infty \quad \forall p \geq 2.$$

□

Lemma 6. *Let $\{x_k\}$ be the sequence generated by Algorithm 1. If Condition 1 is satisfied and f is bounded from below, then*

$$\sum_{k\in\mathcal{S}}\|\nabla m_k(s_k)\| < \infty.$$

Proof. Proof The result is straightforward if $\mathcal{S}$ is a finite set, so let us consider the case where $\mathcal{S}$ is an infinite set. From Lemma 5, it follows that $\{\|s_k\|\}_{k\in\mathcal{S}} \to 0$. Then, using the approximate optimality condition on s_k given (13), an iteration $\bar{k} \in \mathcal{S}$ exists such that

$$\|\nabla m_k(s_k)\| \leq \tau\|s_k\|^2 \quad \forall k \geq \bar{k},\, k \in \mathcal{S}.$$

The desired result follows by invoking Lemma 5 again and noting that, from the instructions of Algorithm 1, we have $\tau < \infty$. □

Let us conclude this subsection of preliminary results by relating, for every iteration k, the optimality violation $\|\nabla f(x_k)\|$ to $\|s_k\|$.

Lemma 7. *Let $\{x_k\}$ be the sequence generated by Algorithm 1. If Condition 1 is satisfied, then*

$$\|\nabla f(x_k)\| \leq (\tau + b + \sigma_{max}T^{r-2})\|s_k\| \quad \forall k \in \mathcal{S},$$

where b and T are defined in Lemma 2 and Lemma 5, respectively.

Proof. Proof Fix any iteration $k \in \mathcal{S}$. Using the approximate optimality condition on s_k given (13) and recalling the expression of ∇m_k given in (21), we get

$$\begin{aligned}
\|\nabla f(x_k)\| &= \|\nabla m_k(s_k) - Q_k s_k - \sigma_k\|s_k\|^{r-2}s_k\| \\
&\leq \|\nabla m_k(s_k)\| + \|Q_k s_k\| + \|\sigma_k\|s_k\|^{r-2}s_k\| \\
&\leq \tau\|s_k\|\min\{\|s_k\|, 1\} + (\|Q_k\| + \sigma_k\|s_k\|^{r-2})\|s_k\| \\
&\leq (\tau + \|Q_k\| + \sigma_{\max}\|s_k\|^{r-2})\|s_k\|,
\end{aligned}$$

where, in the last inequality, we have used the fact that $\sigma_k \leq \sigma_{\max}$ from the instructions of Algorithm 1. The desired result is hence obtained by observing that $\|Q_k\|$ can be upper bounded by b in view of Lemma 2, while $\|s_k\|^{r-2}$ can be upper bounded by T^{r-2}in view of Lemma 5 and the fact that $r \geq 3$ from the instructions of Algorithm 1. □

4.2 Main Result

To establish the convergence of the whole sequence $\{x_k\}$ generated by Algorithm 1, we need to assume pseudoconvexity of the objective function f and the existence of a minimizer, as stated below.

Assumption 1. *The objective function f is pseudoconvex and has a global minimizer.*

We recall that pseudoconvexity of a continuously differentiable function f means that

$$\nabla f(x)^T(y-x) \geq 0 \quad \Rightarrow \quad f(y) \geq f(x) \quad \forall x, y \in \mathbb{R}^n,$$

or equivalently,

$$f(y) < f(x) \quad \Rightarrow \quad \nabla f(x)^T(y-x) < 0 \quad \forall x, y \in \mathbb{R}^n. \tag{27}$$

We also recall that a point x^* is a global minimizer of a pseudoconvex function f if and only if $\nabla f(x^*) = 0$.

We now show that, under Assumption 1 and Condition 1, the sequence $\{x_k\}$ generated by Algorithm 1 is variable metric quasi-Fejér monotone according to Definition 3.

Proposition 1. *Let $\{x_k\}$ be the sequence generated by Algorithm 1. If Assumption 1 holds and Condition 1 is satisfied, then $\{x_k\}$ is a variable metric quasi-Fejér monotone sequence with respect to $\mathcal{F}$ relative to $\{Q_k\}$, according to Definition 3, where*

$$\mathcal{F} := \Big\{ x \in \mathbb{R}^n \colon f(x) \leq \lim_{k\to\infty} f(x_k) \Big\}. \tag{28}$$

Proof. Proof First note that $\mathcal{F}$ is well defined and non-empty. This follows from Lemma 4 and Assumption 1, which ensure that $\{f(x_k)\}$ converges to a value greater than or equal to the minimum value of f.

Now, consider the sequences of scalars $\{\psi_k\}$, $\{\epsilon_k\}$ and $\{\theta_k\}$ defined as follow:

$$\{\psi_k\} \text{ from Condition 1, i.e, satisfying (6)}, \tag{29a}$$

$$\epsilon_k = \begin{cases} (1+\psi_k)\|s_k\|^2_{Q_k} & \text{if } k \in \mathcal{S}, \\ 0 & \text{otherwise}, \end{cases} \tag{29b}$$

$$\theta_k = \begin{cases} 2(1+\psi_k)(\|\nabla m_k(s_k)\| + \sigma_k\|s_k\|^{r-1}) & \text{if } k \in \mathcal{S}, \\ 0 & \text{otherwise.} \end{cases} \tag{29c}$$

To obtain the desired result, we want to show that all the hypothesis of Lemma 1 are satisfied, with $\mathcal{F}$ given in the assertion, by using $\{Q_k\}$ satisfying Condition 1 and $\{\psi_k\}$, $\{\epsilon_k\}$, $\{\theta_k\}$ defined as in (29). Namely, we want to show that

(i) $\{Q_k\}$ satisfies (3)–(4);

(ii) $\{\psi_k\}$, $\{\epsilon_k\}$ and $\{\theta_k\}$ satisfy (6), (7) and (11), respectively;

(iii) inequality (10) holds for all $y \in \mathcal{F}$.

Let us first observe that $\{Q_k\}$ satisfies (3)–(4) from (18a), (19) and Lemma 2. Moreover, recalling (29), the following holds:

- $\{\psi_k\}$ satisfies (6) by definition;
- From Lemma 2, we have

  $$\epsilon_k \leq \Big(1 + \sup_{k\in\mathcal{S}} \psi_k\Big)\|Q_k\|\|s_k\|^2 \leq \Big(1 + \sup_{k\in\mathcal{S}} \psi_k\Big) b\|s_k\|^2 \quad \forall k \in \mathcal{S}.$$

  Using Lemma 5 and the fact that $0 \leq \sup_{k\in\mathcal{S}} \psi_k < \infty$ since $\{\psi_k\}$ satisfies (6), it follows that $\{\epsilon_k\}$ satisfies (7);

- From the fact that $0 \leq \sup_{k\in\mathcal{S}} \sigma_k \leq \sigma_{\max} < \infty$ by the instructions of Algorithm 1, we have
$$\theta_k \leq 2\Big(1 + \sup_{k\in\mathcal{S}} \psi_k\Big)(\|\nabla m_k(s_k)\| + \sigma_{\max}\|s_k\|^{r-1}) \quad \forall k \in \mathcal{S}.$$
Using Lemma 5 with $p = r - 1 \geq 2$, Lemma 6 and the fact that $0 \leq \sup_{k\in\mathcal{S}} \psi_k < \infty$ since $\{\psi_k\}$ satisfies (6), it follows that $\{\theta_k\}$ satisfies (11).

What is left to show is that (10) holds for all $y \in \mathcal{F}$. Consider any iteration $k \geq 0$ and fix any $y \in \mathcal{F}$. Using (20), we have that
$$\|x_{k+1} - y\|^2_{Q_{k+1}} \leq (1 + \psi_k)\|x_{k+1} - y\|^2_{Q_k}. \tag{30}$$
We distinguish two cases depending whether $x_{k+1} = x_k$ or $x_{k+1} \neq x_k$.

(i) Assume that $x_{k+1} = x_k$. Using (30) together with the fact that $\{\theta_k\} \subseteq [0, \infty)$ and $\{\epsilon_k\} \subseteq [0, \infty)$, we can write
$$\|x_{k+1} - y\|^2_{Q_{k+1}} \leq (1 + \psi_k)\|x_k - y\|^2_{Q_k} \leq (1 + \psi_k)\|x_k - y\|^2_{Q_k} + \theta_k\|x_k - y\| + \epsilon_k.$$
Then, (10) holds.

(ii) Assume that $x_{k+1} \neq x_k$. From the instructions of Algorithm 1 and the definition of $\mathcal{S}$ given in (15), we have that
$$s_k = x_{k+1} - x_k \neq 0 \tag{31}$$
and
$$k \in \mathcal{S}. \tag{32}$$
Hence, using item (ii) of Lemma 3 and the fact that $\eta > 0$ from the instructions of Algorithm 1, we have that
$$f(x_k) > f(x_{k+1}) \geq f(y), \tag{33}$$
where the last inequality follows from the fact that $y \in \mathcal{F}$. Now, from simple calculations and using (31), we can write
$$\begin{aligned} \|x_{k+1} - y\|^2_{Q_k} &= \|x_k - y\|^2_{Q_k} + \|x_{k+1} - x_k\|^2_{Q_k} + 2(x_k - y)^T Q_k(x_{k+1} - x_k) \\ &= \|x_k - y\|^2_{Q_k} + \|s_k\|^2_{Q_k} + 2(x_k - y)^T Q_k s_k. \end{aligned} \tag{34}$$
In order to upper bound $2(x_k - y)^T Q_k s_k$ in (34), note that, recalling the expression of ∇m_k given in (21), we have
$$Q_k s_k = \nabla m_k(s_k) - \sigma_k\|s_k\|^{r-2}s_k - \nabla f(x_k).$$
Hence,
$$\begin{aligned} (x_k - y)^T Q_k s_k &= (\nabla m_k(s_k) - \sigma_k\|s_k\|^{r-2}s_k)^T(x_k - y) + \nabla f(x_k)^T(y - x_k) \\ &\leq (\|\nabla m_k(s_k)\| + \sigma_k\|s_k\|^{r-1})\|x_k - y\| + \nabla f(x_k)^T(y - x_k) \\ &\leq (\|\nabla m_k(s_k)\| + \sigma_k\|s_k\|^{r-1})\|x_k - y\|, \end{aligned}$$

where the last inequality, recalling (33), follows from the pseudoconvexity condition (27) of Assumption 1. Combining the above chain of inequalities with (34), we obtain

$$\|x_{k+1} - y\|_{Q_k}^2 \leq \|x_k - y\|_{Q_k}^2 + \|s_k\|_{Q_k}^2 + 2(\|\nabla m_k(s_k)\| + \sigma_k \|s_k\|^{r-1})\|x_k - y\|.$$

Multiplying all terms by $(1 + \psi_k)$ and using (30), we get

$$\|x_{k+1} - y\|_{Q_{k+1}}^2 \leq$$
$$(1 + \psi_k)\|x_k - y\|_{Q_k}^2 + (1 + \psi_k)\|s_k\|_{Q_k}^2 + 2(1 + \psi_k)(\|\nabla m_k(s_k)\| + \sigma_k \|s_k\|^{r-1})\|x_k - y\|.$$

Namely, using (29) and recalling (32), we have

$$\|x_{k+1} - y\|_{Q_{k+1}}^2 \leq (1 + \psi_k)\|x_k - y\|_{Q_k}^2 + \theta_k \|x_k - y\| + \epsilon_k,$$

thus satisfying (10).

□

Finally, by combining Proposition 1 and Theorem 1, we can easily prove the main result of the paper, that is, that the whole sequence $\{x_k\}$ generated by Algorithm 1 converges under Assumption 1 and Condition 1.

Theorem 2. *Let $\{x_k\}$ be the sequence generated by Algorithm 1. If Assumption 1 holds and Condition 1 is satisfied, then $x^* \in \mathbb{R}^n$ exists such that*

$$\lim_{k \to \infty} x_k = x^*.$$

Proof. Proof From Proposition 1, we have that $\{x_k\}$ is a variable metric quasi-Fejér monotone sequence with respect to $\mathcal{F}$ relative to $\{Q_k\}$, where $\mathcal{F}$ is defined as in (28). From point (ii) of Theorem 1, it follows that $\{x_k\}$ is bounded and hence has a limit point x^*. Then, using the monotonicity of $\{f(x_k)\}$ from Lemma 4 and the continuity of f, we get that $\{f(x_k)\} \to f(x^*)$, that is, $x^* \in \mathcal{F}$. The desired result finally follows from item (iii) of Theorem 1. □

5 An Example of a Convergent Algorithm

In Theorem 2, we have shown that the whole sequence $\{x_k\}$ generated by Algorithm 1 converges to a point $x^* \in \mathbb{R}^n$, provided that the objective function f is pseudoconvex and the sequence of matrices $\{Q_k\}$ satisfies Condition 1. Note that Theorem 2 by itself does not say anything about the nature of x^*, that is, whether or not x^* is a minimizer of f. This is due to the fact that, in Algorithm 1, we have considered a general framework where some details were intentionally left unspecified.

To guarantee that the whole sequence $\{x_k\}$ converges to *a minimizer* of a pseudoconvex function f, we need to refine Algorithm 1 in order to guarantee that

$$\lim_{k \to \infty} \|\nabla f(x_k)\| = 0. \tag{35}$$

The above limit, together with Theorem 2, will ensure that the whole sequence $\{x_k\}$ converges to x^* and the latter is a minimizer of f. In the next proposition, we show that this surely occurs when we have an infinite number of successful iterations.

Proposition 2. *Let $\{x_k\}$ be the sequence generated by Algorithm 1 and assume that $\mathcal{S}$ is an infinite set. If Assumption 1 holds and Condition 1 is satisfied, then*

$$\lim_{k\to\infty} x_k = x^*$$

such that x^* is a minimizer of f.

Proof. Proof Theorem 2 ensures that $x^* \in \mathbb{R}^n$ exists such that $\{x_k\} \to x^*$. The continuity of ∇f hence implies that $\{\nabla f(x_k)\} \to \nabla f(x^*)$. Therefore, using the fact that $\mathcal{S}$ is an infinite set, from Lemma 7 we can write

$$\|\nabla f(x^*)\| = \lim_{k\to\infty} \|\nabla f(x_k)\| = \lim_{k\in\mathcal{S}} \|\nabla f(x_k)\| \leq (\tau + b + \sigma_{\max} T^{r-2}) \lim_{k\in\mathcal{S}} \|s_k\|,$$

with $\tau + b + \sigma_{\max} T^{r-2} < \infty$. Since Lemma 5 implies that $\{\|s_k\|\}_{k\in\mathcal{S}}$ converges to 0, it follows that $\|\nabla f(x^*)\| = 0$. As f is pseudoconvex from Assumption 1, we conclude that x^* is a minimizer of f. □

In Algorithm 2, we provide an example of a scheme that uses all the features of Algorithm 1 and guarantees the convergence to a minimizer of a pseudoconvex function f by leveraging Proposition 2.

Algorithm 2 Example of a regularized method with convergence guarantees

given $r \in [3, \infty)$, $\tau \in [0, \infty)$ and $\sigma_{\max} \in [0, \infty)$
choose $x_0 \in \mathbb{R}^n$
for $k = 0, 1, \ldots$ **do**
    choose any $\sigma_k \in [0, \sigma_{\max}]$
    compute a symmetric matrix $Q_k \in \mathbb{R}^{n\times n}$
    compute $s_k \in \mathbb{R}^n$ satisfying (13)
    set $x_{k+1} = x_k + s_k$
end for

Note that, for every iteration k of Algorithm 2, we can choose σ_k to be any value in $[0, \sigma_{\max}]$ and set $x_{k+1} = x_k + s_k$ without checking the new objective value $f(x_{k+1})$. As shown below, the objective decrease condition (14) is guaranteed by assuming Lipschitz continuity of ∇f and strengthening the lower bound requirement on the parameter a appearing in Condition 1.

Assumption 2. *The gradient ∇f of the objective function f is Lipschitz continuous with constant $L > 0$, that is,*

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\| \quad \forall x, y \in \mathbb{R}^n.$$

Condition 2. *Under Assumption 2, the same requirements expressed in Condition 1 hold, but* (19) *is replaced with*

$$a \geq 2\tau + \frac{L}{(1-\eta)}, \tag{36}$$

where $\eta \in (0, 1)$ and L is the Lipschitz constant appearing in Assumption 2.

Now, under Assumptions 1–2 and Condition 2, we can show that the whole sequence $\{x_k\}$ generated by Algorithm 2 converges to a minimizer of a pseudoconvex function f.

Theorem 3. *Let $\{x_k\}$ be the sequence generated by Algorithm 2. If Assumptions 1–2 hold and Condition 2 is satisfied, then all iterations belong to $\mathcal{S}$ and*

$$\lim_{k\to\infty} x_k = x^*$$

such that x^* is a minimizer of f.

Proof. Proof Let us show that, under the assumptions and conditions stated, Algorithm 2 is a special case of Algorithm 1. Namely, we want to show that the objective decrease condition (14) is satisfied for every iteration k of Algorithm 2 with the same η appearing in Condition 2.

Proceeding by contradiction, assume that an iteration $k \geq 0$ exists such that (14) does not hold. Reasoning as in the proof of item (i) of Lemma 3, we still get

$$m_k(0) - m_k(s_k) \geq \Big(\frac{1}{2}a - \tau\Big)\|s_k\|^2. \tag{37}$$

Then, using (36) from Condition 2, it follows that $m_k(0) - m_k(s_k) > 0$ whenever $\|s_k\| > 0$. Hence, since $x_{k+1} = x_k + s_k$ from the instructions of Algorithm 2, the fact that (14) does not hold implies that

$$f(x_k) - f(x_k + s_k) = f(x_k) - f(x_{k+1}) < \eta(m_k(0) - m_k(s_k)). \tag{38}$$

By the Lipschitz continuity of ∇f from Assumption 2, using known results [18] we can write

$$\begin{aligned} f(x_k + s_k) &\leq f(x_k) + \nabla f(x_k)^T s_k + \frac{L}{2}\|s_k\|^2 \\ &\leq f(x_k) + \nabla f(x_k)^T s_k + \frac{L}{2}\|s_k\|^2 + \frac{1}{2}s_k^T Q_k s_k + \frac{\sigma_k}{r}\|s_k\|^r, \end{aligned}$$

where the last inequality follows from the fact that $\frac{1}{2}s_k^T Q_k s_k + \frac{\sigma_k}{r}\|s_k\|^r \geq 0$ in view of the instructions of Algorithm 2 together with Condition 2. Hence, recalling the expression of m_k, we get

$$f(x_k + s_k) \leq f(x_k) + \frac{L}{2}\|s_k\|^2 - (m_k(0) - m_k(s_k)). \tag{39}$$

Using (38) and (39), we get

$$\frac{L}{2}\|s_k\|^2 \geq f(x_k + s_k) - f(x_k) + (m_k(0) - m_k(s_k)) > (1-\eta)(m_k(0) - m_k(s_k)).$$

In view of (37), it follows that

$$\frac{L}{2}\|s_k\|^2 > (1-\eta)\Big(\frac{1}{2}a - \tau\Big)\|s_k\|^2,$$

thus contradicting (36) from Condition 2 and proving that Algorithm 2 is a special case of Algorithm 1. Moreover, since $x_{k+1} = x_k + s_k$ for all $k \geq 0$ from the instructions of Algorithm 2, then all iterations are successful according to (15), that is,

$$\mathcal{S} = \{0, 1, \ldots\}. \tag{40}$$

Therefore, observing that Condition 1 is satisfied as it is implied by Condition 2, it follows from Proposition 2 that $\{x_k\} \to x^*$ such that x^* is a minimizer of f. □

5.1 Convergence Rate in the Convex Case

Let us conclude the analysis of Algorithm 2 by showing that, when f is convex, not only does the whole sequence of points $\{x_k\}$ converge to a minimizer, but the objective error also converges to zero at a rate of the order of k^{-1}. Hence, we attain the same convergence rate for the objective error as other regularized methods that use first-order information in a convex setting [8].

Assumption 3. *The objective function f is convex and has a global minimizer.*

Proposition 3. *Let $\{x_k\}$ be the sequence generated by Algorithm 2. If Assumptions 2–3 hold and Condition 2 is satisfied, then*

$$f(x_k) - f^* \leq \frac{R^2(f(x_0) - f^*)}{R^2 + \nu k(f(x_0) - f^*)} \quad \forall k \geq 1,$$

where $f^* := \min_{x \in \mathbb{R}^n} f(x)$ and

$$\nu := \frac{\eta c}{(\tau + b + \sigma_{max} T^{r-2})^2} > 0,$$

while R, b, c and T are defined as in Theorem 1, Lemma 2, Lemma 3 and Lemma 5, respectively.

Proof. Proof Reasoning as in the proof of Theorem 3, we have that Algorithm 2 is a special case of Algorithm 1 where (40) holds. Now fix an iteration $k \geq 0$. Using item (ii) of Lemma 3 and the fact that $x_{k+1} = x_k + s_k$, we get

$$f(x_k) - f(x_{k+1}) \geq \eta c \|s_k\|^2.$$

Since $k \in \mathcal{S}$ as (40) holds, using Lemma 7 we obtain

$$f(x_k) - f(x_{k+1}) \geq \nu \|\nabla f(x_k)\|^2, \tag{41}$$

where ν is defined as in the statement.

Moreover, using the convexity of f from Assumption 3, we can write

$$f(x_k) - f^* \leq \nabla f(x_k)^T (x_k - x^*) \leq \|\nabla f(x_k)\| \|x_k - x^*\|. \tag{42}$$

Since Algorithm 2 is a special case of Algorithm 1, while Assumption 1 and Condition 1 are implied by Assumptions 3 and Condition 2, respectively, then we can use Proposition 1 to get that $\{x_k\}$ is a variable metric quasi-Fejér monotone sequence. Using item (i) of Theorem 1, it follows that $\|x_k - x^*\| \leq R$. Hence, from (42) we obtain

$$f(x_k) - f^* \leq R \|\nabla f(x_k)\|. \tag{43}$$

Now, denote

$$E_k := f(x_k) - f^*.$$

It follows from (41) and (43) that

$$E_k - E_{k+1} \geq \frac{\nu}{R^2} E_k^2.$$

Then,

$$\frac{1}{E_{k+1}} - \frac{1}{E_k} = \frac{E_k - E_{k+1}}{E_{k+1} E_k} \geq \frac{\nu E_k}{R^2 E_{k+1}} \geq \frac{\nu}{R^2}.$$

Applying the above chain of inequalities recursively, we get

$$\frac{1}{E_{k+1}} \geq \frac{1}{E_0} + \frac{\nu(k+1)}{R^2} = \frac{R^2 + E_0 \nu(k+1)}{E_0 R^2}.$$

Then, the desired result follows. □

Remark 2. *The $\mathcal{O}(1/k)$ convergence rate given in Proposition 3 does not require bounded level sets, thus showing an improvement over classical results of regularized methods that use first-order information in the convex case (cfr. [8, Theorem 2.5]).*

6 An iterative procedure for matrix computation

The requirements given in Condition 1 on the sequence of matrices $\{Q_k\}$ are rather general and can be satisfied in many ways. As highlighted in Remark 1, the easiest choice is setting Q_k to an appropriate constant matrix at any iteration k.

In this section, we describe a strategy based on an iterative update of Q_k. In what follows, given a symmetric matrix $M \in \mathbb{R}^{n \times n}$, we denote its lowest eigenvalue by $\lambda_{\min}(M)$.

Given $\{\psi_k\}$ and a satisfying (6) and (19), respectively, let $\{a_k\}$ be a sequence of scalars such that

$$a < a_{k+1} < a_k \quad \forall k \geq 0. \tag{44}$$

At iteration k, assume that we have a symmetric matrix $Q_k \in \mathbb{R}^{n \times n}$ such that

$$Q_k \succeq a_k I \tag{45}$$

and we have computed a symmetric matrix $\tilde{Q}_{k+1} \in \mathbb{R}^{n \times n}$. Now define

$$\Delta_k := \tilde{Q}_{k+1} - Q_k$$

and choose $\beta_k \geq 0$ to set

$$Q_{k+1} = Q_k + \beta_k \Delta_k = (1 - \beta_k) Q_k + \beta_k \tilde{Q}_{k+1}. \tag{46}$$

We will show that, choosing β_k sufficiently small, the resulting sequence $\{Q_k\}$ satisfies Condition 1 provided that $Q_0 \succ a_0 I$.

In particular, since (18a) holds by assumption in view of (44)–(45), what we need to show is that, using suitable values of β_k, (18b) holds and (45) is satisfied with k replaced by $k+1$ as well. If this is true, reasoning by induction we get that Condition 1 is satisfied provided that $Q_0 \succ a_0 I$. In the sequel, we will describe two possible strategies to choose β_k.

6.1 First option

At iteration k, assuming without loss of generality that $\|\Delta_k\| > 0$ (otherwise $Q_{k+1} = Q_k$), we can set

$$\beta_k \in \begin{cases} \left[0, \dfrac{\psi_k \lambda_{\min}(Q_k)}{\|\Delta_k\|}\right] & \text{if } \lambda_{\min}(\tilde{Q}_{k+1}) \geq \lambda_{\min}(Q_k), \\ \left[0, \min\left\{\dfrac{\psi_k \lambda_{\min}(Q_k)}{\|\Delta_k\|}, \dfrac{\lambda_{\min}(Q_k) - a_{k+1}}{\lambda_{\min}(Q_k) - \lambda_{\min}(\tilde{Q}_{k+1})}\right\}\right] & \text{if } \lambda_{\min}(\tilde{Q}_{k+1}) < \lambda_{\min}(Q_k). \end{cases} \tag{47}$$

As discussed above, to satisfy Condition 1 provided that $Q_0 \succ a_0 I$, we need to prove that (18b) holds and that (45) is satisfied with k replaced by $k+1$ as well. Indeed, at every iteration k we have the following:

- Since $\beta_k \leq \psi_k \lambda_{\min}(Q_k)/\|\Delta_k\|$ from (47), then

  $$\beta_k \Delta_k \preceq \psi_k Q_k.$$

  Adding Q_k to both terms and using (46), it follows that (18b) holds.

- Using (46) and the Weyl's inequality, we can write

  $$\lambda_{\min}(Q_{k+1}) \geq (1-\beta_k)\lambda_{\min}(Q_k) + \beta_k \lambda_{\min}(\tilde{Q}_{k+1}) = \beta_k(\lambda_{\min}(\tilde{Q}_{k+1}) - \lambda_{\min}(Q_k)) + \lambda_{\min}(Q_k). \tag{48}$$

  From (48) and the non-negativity of β_k expressed in (47), it follows that:

  - if $\lambda_{\min}(\tilde{Q}_{k+1}) \geq \lambda_{\min}(Q_k)$, then $\lambda_{\min}(Q_{k+1}) \geq \lambda_{\min}(Q_k)$, implying from (44)–(45) that

    $$\lambda_{\min}(Q_{k+1}) \geq a_{k+1} \quad \forall \beta_k \geq 0;$$

  - if $\lambda_{\min}(\tilde{Q}_{k+1}) < \lambda_{\min}(Q_k)$, then

    $$\lambda_{\min}(Q_{k+1}) \geq a_{k+1} \quad \forall \beta_k \leq \frac{\lambda_{\min}(Q_k) - a_{k+1}}{\lambda_{\min}(Q_k) - \lambda_{\min}(\tilde{Q}_{k+1})}.$$

  From the above two cases, we get that (47) guarantees $\lambda_{\min}(Q_{k+1}) \geq a_{k+1}$, that is, (45) is satisfied with k replaced by $k+1$ as well.

Note that, provided that $\psi_k > 0$, the upper bound on β_k obtained from (47) is positive in view of (44)–(45).

Finally, also note that the acceptable values of β_k obtained from (47) may be ≥ 1. This can be useful, for example, if one intends to use a quasi-Newton formula to compute $\tilde{Q}_{k+1}$ since we may set $Q_{k+1} = \tilde{Q}_{k+1}$ in such a case, hence recovering the standard quasi-Newton update.

6.2 Second option

The strategy described in the previous subsection requires to compute $\lambda_{\min}(Q_k)$ and $\lambda_{\min}(\tilde{Q}_{k+1})$ at iteration k. In order to reduce the computational burden associated with the lowest eigenvalue computation, here we describe an alternative strategy which, at iteration k, requires to compute $\lambda_{\min}(Q_k)$ only. This comes at the price of requiring

$$\tilde{Q}_{k+1} \succeq 0 \tag{49}$$

and, as to be discussed below, restricting the acceptable values of β_k.

In more detail, assuming again without loss of generality that $\|\Delta_k\| > 0$ (otherwise $Q_{k+1} = Q_k$), we can choose

$$0 \leq \beta_k \leq \min\left\{ \frac{\psi_k \lambda_{\min}(Q_k)}{\|\Delta_k\|}, 1 - \frac{a_{k+1}}{\lambda_{\min}(Q_k)} \right\}. \tag{50}$$

As discussed above, to satisfy Condition 1 provided that $Q_0 \succ a_0 I$, we need to prove that (18b) holds and that (45) is satisfied with k replaced by $k+1$ as well. Indeed, at every iteration k we have the following:

- Reasoning as in Subsection 6.1, since $\beta_k \leq \psi_k \lambda_{\min}(Q_k)/\|\Delta_k\|$ from (50), then

  $$\beta_k \Delta_k \preceq \psi_k Q_k.$$

  Adding Q_k to both terms and using (46), it follows that (18b) holds.

- Since $\beta_k \leq 1 - a_{k+1}/\lambda_{\min}(Q_k)$ from (50), and using the fact that $\lambda_{\min}(Q_k) > 0$ from (44)–(45), we have

  $$(1 - \beta_k)\lambda_{\min}(Q_k) \geq a_{k+1}. \tag{51}$$

  Moreover, reasoning as in Subsection 6.1, using (46) and the Weyl's inequality, we get (48). In particular, since $\beta_k \lambda_{\min}(\tilde{Q}_{k+1}) \geq 0$ from (49) and the non-negativity of β_k expressed in (50), from (48) we obtain

  $$\lambda_{\min}(Q_{k+1}) \geq (1 - \beta_k)\lambda_{\min}(Q_k). \tag{52}$$

  Combining (51) and (52), we get $\lambda_{\min}(Q_{k+1}) \geq a_{k+1}$, that is, (45) is satisfied with k replaced by $k+1$ as well.

Note that, from (44)–(45), we have

$$0 < a_{k+1} < \lambda_{\min}(Q_k).$$

Then, provided that $\psi_k > 0$, the upper bound on β_k obtained from (50) is positive, as for the strategy outlined in Subsection 6.1. However, here the acceptable values of β_k obtained from (50) are < 1. Namely, using (50) we can never set $Q_{k+1} = \tilde{Q}_{k+1}$, differently from the strategy described in Subsection 6.1.

7 Numerical examples

In this section, we analyze the practical performance of the proposed algorithmic framework using different strategies to compute a sequence of matrices $\{Q_k\}$ satisfying Condition 1. The experiments were run in Matlab R2025b on an Apple MacBook Pro with an Apple M1 Pro Chip and 16 GB RAM.

We consider the ARC algorithm [6], which is a special case of Algorithm 1 with $r = 3$. In particular, at each iteration k:

- s_k is computed to satisfy, together with condition (13) required by our scheme, also the Cauchy condition

$$m_k(s_k) \leq m_k(s_k^{\mathrm{C}}), \quad \text{where} \quad s_k^{\mathrm{C}} = -\alpha_k^{\mathrm{C}} \nabla f(x_k), \quad \alpha_k^{\mathrm{C}} \in \operatorname*{Argmin}_{\alpha \geq 0} m_k(-\alpha \nabla f(x_k));$$

- after computing ρ_k as in (16), we set

$$x_{k+1} = \begin{cases} x_k + s_k & \text{if } \rho_k \geq \eta_1, \\ x_k & \text{otherwise,} \end{cases}$$

and

$$\sigma_{k+1} = \begin{cases} \max\{\sigma_k/2, \sigma_{\min}\} & \text{if } \rho_k \geq \eta_2, \\ \sigma_k & \text{if } \eta_1 \leq \rho_k < \eta_2, \\ 2\sigma_k & \text{if } \rho_k < \eta_1, \end{cases}$$

with $\eta_1 = 0.1$, $\eta_2 = 0.9$, $\sigma_{\min} = 10^{-6}$ and $\sigma_0 = 1$.

We consider four versions of ARC, each using a different strategy to compute Q_k at iteration k among those analyzed in Sections 5–6:

- ARC_1 uses the strategy described in Section 5 with a constant matrix $Q_k = aI$, where a is such that (36) holds with equality.
- ARC_2 uses the strategy described in Section 6, where $\tilde{Q}_{k+1}$ is the BFGS update [20] and β_k is computed as described in Subsection 6.1. In particular, β_k is chosen as the minimum between 1 and the largest value obtained from (47) so as to obtain the standard BFGS update (i.e., $\beta_k = 1$) whenever possible.
- ARC_3 differs from ARC_2 in that β_k is computed as described in Subsection 6.2. In particular, β_k is chosen as the largest value obtained from (50) (which is surely < 1).
- ARC_4 differs from ARC_2 in that $\tilde{Q}_{k+1}$ is a diagonal matrix obtained from the diagonal part of the ordinary BFGS update [17], with Q_k diagonal as well. Note that, in this case, computing $\lambda_{\min}(Q_k)$ and $\lambda_{\min}(\tilde{Q}_{k+1})$ has a negligible cost.

Observe that ARC_1 is a special case of Algorithm 2 which trivially satisfies Condition 2 and, consequently, Condition 1, regardless of how $\{\psi_k\}$ is chosen. Moreover, in view of

Theorem 3, all iterations of ARC_1 are successful and $\{\sigma_k\}$ is therefore bounded from above by σ_0.

For the other variants, which implement the strategies described in Section 6, we use $\{\psi_k\} = \{10^8(k+1)^{-2}\}$ and $\{a_k\} = \{a+(k+1)^{-1}\}$, with $a = 2\tau+10^{-6}$. Moreover, σ_k remains below $\sigma_{\max} = 10^{15}$ in all our experiments. However, in case σ_k exceeds a pre-specified $\sigma_{\max}$, one may switch to ARC_1.

We consider the following two classes of pseudoconvex problems that admit multiple optimal solutions:

- *Binary classification with squared hinge loss.* Let $\{v_1, \dots, v_m\} \subseteq \mathbb{R}^n$, with $m = 20,000$ and $n = 1000$, be a set of vectors that can be divided into two equally sized groups by a hyperplane passing through the origin with margin 0.01. Using labels $y^1, \dots, y^m \in \{\pm 1\}$ assigned to the points, we want to compute a separating hyperplane by minimizing a squared hinge loss as in the following problem:

$$\min f(x) = \sum_{i=1}^{m} \max\{0, 1 - y^i (v^i)^T x\}^2.$$

  Note that f is convex and any separating hyperplane passing through the origin, after an appropriate rescaling, defines an optimal solution.

- *Linear feasibility with log-penalty.* Let $A \in \mathbb{R}^{m \times n}$ be the matrix of a linear system, with $m = 100$ and $n = 1000$. We want to compute a solution of $Ax = 0$ by minimizing a log-penalty loss as in the following problem:

$$\min f(x) = \log(1 + \|Ax\|^2).$$

  Note that f is pseudoconvex and any vector in the kernel of A is an optimal solution.

For both classes of problems, we run the considered algorithms on 10 randomly generated instances, and the average results are reported in Figure 1. We observe that ARC_1, which uses a constant matrix at all iterations, performs the worst on both problems. Among the other three versions, ARC_2 and ARC_3 yield very similar results and outperform ARC_4, especially on the first class of problems. Overall, these results seem to suggest that a BFGS-type approach can be beneficial for computing the sequence of matrices $\{Q_k\}$.

8 Conclusions

In this paper, we have established some convergence properties of regularized methods for unconstrained optimization, that is, algorithms that use quadratic models regularized by a power $r \geq 3$ of the norm of the step, focusing on the case where only the objective function and its gradient are evaluated. We have shown that, if the objective function is pseudoconvex and the regularized model is properly chosen at every iteration, then the whole sequence of points generated by this class of algorithms converges. Moreover, the convergence does not require bounded level sets.

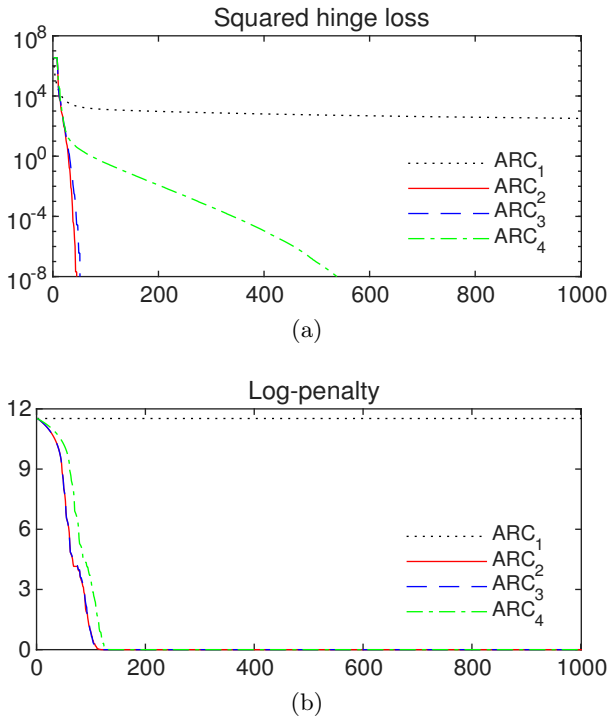

Figure 1: Objective error decrease on (a) binary classification problems using squared hinge loss and (b) linear feasibility problems with log-penalty. In the left plot, the y axis is in logarithmic scale.

The key to obtain the above result is choosing, at every iteration, the matrix of the quadratic model in such a way that the resulting sequence of points is variable metric quasi-Fejér monotone.

We finally observe that the convergence of the whole sequence of points was established in [6, 19] with $r = 3$ using second-order information. Specifically, the result was obtained in [6, Theorem 4.4] when the Hessian of f is continuous and there exists a subsequence of iterates converging to a point where the Hessian of f is non-singular; whereas the result was obtained in [19, Theorem 3] by employing exact minimizers of the regularized model built with the Hessian matrix of f, the latter assumed to be Lipschitz continuous and such that its smallest eigenvalue at the starting point is sufficiently large compared to the gradient norm. We see that our analysis differs mainly in that we do not require f to have second-order derivatives (and we allow for a regularization power $r \geq 3$), although we impose the additional pseudoconvexity condition on f. Moreover, compared to [6], our results do not need to assume specific properties at the limit, whereas, compared to [19], our results are independent of the starting point and use approximate minimizers of the regularized model.

References

[1] Y. I. Alber, A. N. Iusem, and M. V. Solodov. On the projected subgradient method for nonsmooth convex optimization in a Hilbert space. *Mathematical Programming*, 81: 23–35, 1998.

[2] D. P. Bertsekas. *Nonlinear programming*. Athena Scientific, Belmont, MA, 1999.

[3] D. P. Bertsekas. *Convex Optimization Algorithms*. Athena Scientific, 2015.

[4] E. G. Birgin, J. Gardenghi, J. M. Martínez, S. A. Santos, and P. L. Toint. Worst-case evaluation complexity for unconstrained nonlinear optimization using high-order regularized models. *Mathematical Programming*, 163:359–368, 2017.

[5] R. Burachik, L. Graña Drummond, A. N. Iusem, and B. F. Svaiter. Full convergence of the steepest descent method with inexact line searches. *Optimization*, 32(2):137–146, 1995.

[6] C. Cartis, N. I. Gould, and P. L. Toint. Adaptive cubic regularisation methods for unconstrained optimization. Part I: motivation, convergence and numerical results. *Mathematical Programming*, 127(2):245–295, 2011.

[7] C. Cartis, N. I. Gould, and P. L. Toint. Adaptive cubic regularisation methods for unconstrained optimization. Part II: worst-case function-and derivative-evaluation complexity. *Mathematical Programming*, 130(2):295–319, 2011.

[8] C. Cartis, N. I. Gould, and P. L. Toint. Evaluation complexity of adaptive cubic regularization methods for convex unconstrained optimization. *Optimization Methods and Software*, 27(2):197–219, 2012.

[9] C. Cartis, N. I. Gould, and P. L. Toint. Worst-case evaluation complexity of regularization methods for smooth unconstrained optimization using Hölder continuous gradients. *Optimization Methods and Software*, 32(6):1273–1298, 2017.

[10] C. Cartis, N. I. Gould, and P. L. Toint. Universal regularization methods: varying the power, the smoothness and the accuracy. *SIAM Journal on Optimization*, 29(1): 595–615, 2019.

[11] P. L. Combettes. Quasi-Fejérian analysis of some optimization algorithms. In *Studies in Computational Mathematics*, volume 8, pages 115–152. Elsevier, 2001.

[12] P. L. Combettes and B. C. Vũ. Variable metric quasi-Fejér monotonicity. *Nonlinear Analysis: Theory, Methods & Applications*, 78:17–31, 2013.

[13] L. G. Drummond and B. F. Svaiter. A steepest descent method for vector optimization. *Journal of computational and applied mathematics*, 175(2):395–414, 2005.

[14] G. N. Grapiglia, M. L. Gonçalves, and G. Silva. A cubic regularization of Newton's method with finite difference Hessian approximations. *Numerical Algorithms*, 90(2): 607–630, 2022.

[15] A. N. Iusem. On the convergence properties of the projected gradient method for convex optimization. *Computational & Applied Mathematics*, 22:37–52, 2003.

[16] A. N. Iusem, B. F. Svaiter, and M. Teboulle. Entropy-like proximal methods in convex programming. *Mathematics of Operations Research*, 19(4):790–814, 1994.

[17] D. Li, X. Wang, and J. Huang. Diagonal BFGS updates and applications to the limited memory BFGS method. *Computational Optimization and Applications*, 81(3):829–856, 2022.

[18] Y. Nesterov. *Introductory Lectures on Convex Optimization: A Basic Course*, volume 87. Springer Science & Business Media, 2013.

[19] Y. Nesterov and B. T. Polyak. Cubic regularization of Newton method and its global performance. *Mathematical Programming*, 108(1):177–205, 2006.

[20] J. Nocedal and S. J. Wright. *Numerical optimization*. Springer, 2006.

[21] B. T. Polyak. *Introduction to optimization*. Optimization Software, New York, 1987.