Commun. Optim. Theory 2027 (2027) 16 https://doi.org/10.23952/cot.2027.16

MATHRES Communications in Optimization Theory
Available online at http://cot.mathres.org

FAST INERTIAL ALGORITHMS WITH A CORRECTION TERM FOR STRUCTURED MINIMIZATION PROBLEMS

ANDRÉ WENG-LAW*, PAUL-EMILE MAINGÉ

Université des Antilles, Campus de Schoelcher, 97233, Schoelcher, Martinique (F.W.I.), France

Abstract. In this paper, we develop and study a new class of rapidly convergent proximal-like algorithms CRIPGA for solving structured non-smooth and convex minimization problems in a real Hilbert space framework. It can be regarded as a discrete counterpart of some recently discussed continuous Newton-like dynamical system with remarkable convergence properties. This new numerical scheme incorporates a correction term, besides the typical momentum term (arising in acceleration techniques introduced by Nesterov). As a special case of the parameters, we recover the very recently investigated method RIGPA (with no occurrence of correction term), for which interesting convergence properties were first established in the context of non-smooth minimization. A main specificity of both these latter processes is that they generate sequences that not only converge (towards equilibria) with some property of fast decaying subgradients (well-known as an uncommon efficient stability phenomenon), but also that minimize the function values at the rate $o(n^{-2})$ (as for the most famous related methods). However, our algorithm provides us some additional degree of freedom (through the correction term) that considerably influences its convergence properties. Numerical experiments are furthermore performed so as to illustrate the efficiency of CRIPGA compared with both classical and recently studied methods.

Keywords. Composite minimization; Fast first-order method; Inertial-type algorithm; Nesterov-type algorithm; Optimal gradient method.

2020 Mathematics Subject Classification. 90C25, 90C06, 65F22.

1. INTRODUCTION

The purpose of this work is to propose and discuss a numerical approach to structured and (possibly) non-smooth minimization problems, by means of rapidly convergent proximal methods, in the context of real Hilbert spaces. Throughout this paper, H is a real Hilbert space endowed with inner product and induced norm denoted by $\langle .,. \rangle$ and $\| \, . \, \|$, respectively. Our attention here will be focused on solving the structured convex programming problem:

$$\min\{\Theta(x) := f(x) + g(x) : \quad x \in H\}, \tag{1.1}$$

*Corresponding author.
E-mail address: Andre.Wenglaw@gmail.com (A. Weng-Law).
Received January 6, 2025; Accepted March 20, 2025.

©2027 Communications in Optimization Theory

where f and g satisfy the conditions below:

$$g : H \to (-\infty, \infty] \text{ is convex, proper and lower semi-continuous,} \tag{1.2}$$

$$f : H \to (-\infty, \infty) \text{ is convex and continuously differentiable, with } L\text{-Lipschitz continuous gradient } \nabla f, \tag{1.3}$$

$$S := \operatorname{argmin}_H \Theta \neq \emptyset. \tag{1.4}$$

This problem has been intensively studied in the last decades due to its applications in numerous real-life fields, such as for instance, image processing, computer vision, machine learning or signal processing. It also encompasses other important issues such as variational inequalities. As an example developed in this paper, the considered problem occurs in image reconstruction with $g(x) = \|x-b\|_1$ (b being the input image to be processed (see, e.g., [12, 28])), which acts as a data fidelity term.

Remark 1.1. The proximal mapping of g with index $\mu > 0$ is defined for $x \in H$ by $\operatorname{prox}_{\mu g}(x) := \operatorname{argmin}_{y \in H} \left(g(y) + (2\mu)^{-1}\|x-y\|^2 \right)$. This is nothing but the resolvent operator $J_{\mu \partial g} := (I + \mu \partial g)^{-1}$, where $\partial g : H \to 2^H$ is the Fenchel subdifferential of g. It is well-known to be single-valued and everywhere defined (see, e.g., [9, 13, 17]).

In order to approximate some solution to (1.1)-(1.4), we consider the sequences $\{x_n, y_n\} \subset H$ generated by the following Corrected and Relaxed Inertial Proximal Gradient Algorithm (CRIPGA, for short).

General form of CRIPGA :

▷ **Initialization**: take $\{z_{-1}, x_0, y_0\} \subset H$.

▷ **Updates** : given $\{z_{n-1}, x_n, y_n\} \subset H$, we compute the updates by

$$z_n = x_n - \theta_n(y_n - x_n) - \eta_n(x_n - z_{n-1}), \tag{1.5}$$

$$x_{n+1} = J_{\mu \partial g}(z_n - \mu \nabla f(z_n)), \tag{1.6}$$

$$y_{n+1} = (1-\kappa) y_n + \kappa x_n, \tag{1.7}$$

where $\{\kappa, \mu\} \subset (0, \infty)$, $(\eta_n) \subset [0,1)$, while $(\theta_n) \subset (0, \infty)$ is defined, for some additional parameters $e_* > 0$ and $(\nu_n) \subset (0, \infty)$, by :

$$\theta_n := \tfrac{\nu_n - (1-\kappa)\nu_{n+1}}{e_* + \nu_{n+1}}. \tag{1.8}$$

The parameters ν_n and η_n are intended (for some constant $\varepsilon \in [0,1)$) to be such that

$$\tfrac{\nu_{n+1}}{\nu_n} \to 1, \quad \nu_n \to \infty \quad \text{and} \quad \eta_n \to \varepsilon \quad (\text{as } n \to \infty). \tag{1.9}$$

The form of this method was recently suggested by Maingé-Weng-Law [21] as a discrete counterpart of a continuous dynamical system (see Section 2.2). The simple form of CRIPGA generalizes some interesting convergence features of Newton-like methods for convex minimization.

Remark 1.2. Let us stress that CRIPGA (1.5)-(1.8) can be equivalently written as the algorithm CRIPGA-B (1.10)-(1.13) below whose formulation is more familiar to the optimization community (see Proposition 2.1).

CRIPGA-B (as an alternative form of CRIPGA):
$\triangleright$ For $n \geq 1$, given $\{z_{n-2}, y_{n-1}, y_n\}$, we compute the updates by

$$z_{n-1} = y_n + \alpha_{n-1}(y_n - y_{n-1}) + \eta_n(z_{n-2} - y_{n-1}), \tag{1.10}$$

$$x_n = \mathrm{prox}_{\mu g}(z_{n-1} - \mu \nabla f(z_{n-1})), \tag{1.11}$$

$$y_{n+1} = (1-\kappa)y_n + \kappa x_n, \tag{1.12}$$

where the momentum coefficient $(\alpha_n) \subset (0, \kappa^{-1})$ is defined by

$$\alpha_n = \kappa^{-1}(1 + \theta_n - \eta_n) - 1. \tag{1.13}$$

It can be seen that the algorithm CRIPGA-B includes a relaxation term $(1-\kappa)y_n$, an extrapolation parameter α_n along with a correction term $\eta_n(z_{n-2} - y_{n-1})$. Pioneering approaches to accelerated proximal methods only combining relaxation and inertia were discussed for instance by Attouch-Cabot [2] $\big($using varying parameters (κ_n) and (μ_n) instead of the constant values κ and $\mu\big)$ and Iutzeler-Hendrickx [15] (relative to averaged operators). It can be noticed that, as in these previous schemes, the extrapolation coefficient α_n usually belongs to $[0,1]$, while, for CRIPGA-B, it satisfies $\alpha_n \to (1-\varepsilon)\kappa^{-1}$ $\big($so that, when $\kappa \in (0,1)$, α_n can be greater than 1 under the assumptions in (1.9)$\big)$. Let us mention that a similar correction term as in CRIPGA-B was incorporated in proximal-like methods $\big($even relative to more general structured inclusion problems (see Labarre [16], Maingé-Weng-Law [22], Tran-Dinh [31])$\big)$ in which the correction coefficient satisfies $\eta_n \to 1$, while for CRIPGA-B, η_n is set to be such that $\eta_n \to \varepsilon \in [0,1)$ (as $n \to \infty$).

A new class of algorithms given by (1.5)-(1.8) is proposed for solving non-smooth convex structured minimization problems. This new class recovers the recently studied method RIGPA (see Maingé-Labarre [20]) when $\eta_n = 0$, for which interesting convergence properties were stated even in the context of non-smooth optimization. Convergence results are established for (1.5)-(1.8) in the setting of either general parameters (Theorem 5.1) or particular ones (Corollary 5.2) with $\nu_n = n + \nu_0$ (for some positive constant ν_0). Our approach relies on Lyapunov properties of some energy-like sequence $E_n(s,q)$ (defined in (3.10) for $s > 0$ and $q \in S$) associated with our process. Specifically, as for the methods Accelerated Forward-Backward (AFB - see Attouch-Peypouquet [6]) and RIGPA, we prove under appropriate conditions $\big($including the ranges of parameters $\kappa \in (0,2)$ and $e_* > 2\kappa^{-1}\big)$ that (1.5)-(1.8) also generates (weakly) convergent sequences (x_n) that achieve a complexity result of $o(n^{-2})$ for minimizing the function values. Moreover, by exploiting the first-order formulation of (1.5)-(1.8), we successfully set out (even in the non-smooth setting) the existence of subgradients $\xi_n^* \in \partial\Theta(x_n)$ that obey $\sum_n n^2 \|\xi_n^*\|^2 < \infty$. This estimate corresponds to the stability phenomenon known as the property of fast decaying subgradients obtained for RIGPA (see also, Shi-Du-Jordan-Su [29], Attouch-Chbani-Fadili-Riahi [3]).

Our new class of algorithms provides us with an additional degree of freedom (compared with RIGPA) through the correction coefficient η_n that considerably influences the convergence of the method particularly in the range $\varepsilon \in (0,1)$.

An outline of this paper is as follows. In Section 2, we show connections between CRIPGA (1.5)-(1.8) and the state of the art, but we also prove the equivalency between CRIPGA and CRIPGA-B (1.10)-(1.13). In Section 3, an important inequality is stated in view of a Lyapunov analysis of the method. In Section 4, we then conduct a preliminary convergence analysis of the

proposed algorithm. Section 5 is devoted to the main estimates and convergence results related to CRIPGA. Numerical experiments are performed in Section 6 so as to illustrate the efficiency of the considered method.

Remark 1.3. From now on, for the sake of legibility, we use the following notations:
- Given any sequence (u_n), we denote $\dot{u}_n = u_n - u_{n-1}$ (but we sometimes write $[u_n]_\circ = u_n - u_{n-1}$ instead) and $\ddot{u}_n = \dot{u}_n - \dot{u}_{n-1}$.
- For any real t, we denote $(t)_+ = \max\{0,t\}$.

2. Connection of the Considered Method with the State of the Art and Continuous Dynamical Systems

2.1. **Some reminder regarding existing algorithms.** Let us underline that the proposed algorithm CRIPGA (1.5)-(1.8) can be regarded as an enhanced version of other proximal-type algorithms studied in the literature.

Indeed, in the special instance where $\kappa = 1$ and $\nu_n = \eta_n = 0$, CRIPGA reduces to the well-known Forward-Backward Algorithm (see, e.g. Lions-Mercier [18], Passty [27]), which considers the sequence (x_n) generated by the following process

$$x_{n+1} = \mathrm{prox}_{\mu_n g}(x_n - \mu_n \nabla f(x_n)). \tag{2.1}$$

For a function f such that ∇f is L-Lipschitz continuous, and for indexes $(\mu_n) \subset (0, L^{-1})$, (2.1) was shown to generate weakly convergent sequences (x_n) that satisfy the (sub-linear) rate $\Theta(x_n) - \inf_H \Theta = O(n^{-1})$ (see [10]).

In the special case when $\kappa = 1$, $\eta_n = 0$, $\nu_n = n-1$ and $e_* = \alpha - 1$, CRIPGA reduces to the algorithm below proposed by Chambolle-Dossal [11] (also see Attouch-Peypouquet [5])

$$z_n = x_n + \tfrac{n-1}{n+\alpha-1}(x_n - x_{n-1}), \qquad x_{n+1} = \mathrm{prox}_{\mu g}(z_n - \mu \nabla f(z_n)), \tag{2.2}$$

where $\mu \in (0, L^{-1})$ and $\alpha > 0$. For $\alpha > 3$, the authors successfully proved (see [5]) that (2.2) generates (weakly) convergent sequences (x_n) that minimize the function values with a convergence rate $\Theta(x_n) - \inf_H \Theta = o(n^{-2})$. This algorithm can be seen as an improved variant of FISTA (*Fast Iterative Soft Thresholding Algorithm*) discussed by Beck-Teboulle [8] (also see [32]). Recall that FISTA, which incorporates the acceleration techniques introduced by Güler [14] and Nesterov [23, 24, 25], was shown to generate iterates (x_n) that satisfy the convergence rate $\Theta(x_n) - \inf_H \Theta = O(n^{-2})$ but whose convergence has not been proved.

In absence of correction term (namely $\eta_n = 0$), CRIPGA (through its formulation CRIPGA-B) reduces to the algorithm RIGPA investigated by Maingé-Labarre [20] for which the property of fast vanishing subgradients was established (also see, e.g., Attouch-Peypouquet-Redont [7], Shi-Du-Jordan-Su [29]), besides the convergence rate $\Theta(x_n) - \inf_H \Theta = o(n^{-2})$.

It is worthwhile mentioning that a closely related corrected and relaxed inertial forward-backward algorithm was recently investigated by Maingé-Weng-Law in [22] relative to the more general framework of structured monotone inclusions (see also Tran-Dinh [31] for a similar approach). It consists of a modification of the classical forward-backward algorithm, that incorporates a correction term besides the classical momentum term, combined with a constant relaxation factor and a preconditioning process. The authors successfully proved the weak convergence to equilibria of the iterates (x_n) with worst-case convergence rates of $o(n^{-1})$ for both the discrete velocity and the fixed point residual.

2.2. **Links with continuous dynamical systems.** It is interesting to recall that the proposed method (1.5)-(1.8) can be regarded as a discrete counterpart of the following second-order differential system (see Maingé-Weng-Law [21] (Paragraph 5.2)) :

$$\xi(t) \in \partial f(x(t)), \tag{2.3}$$

$$\big(x(\cdot)+\sigma(\cdot)\xi(\cdot)\big)^{(2)}(t)+\alpha(t)\dot{x}(t)+\beta(t)\big(\sigma(\cdot)\xi(\cdot)\big)^{(1)}(t)+b(t)\sigma(t)\xi(t)=0, \tag{2.4}$$

where $\{\alpha(\cdot),\beta(\cdot),b(\cdot),\sigma(\cdot)\}$ are conveniently chosen positive functions, with $\alpha(\cdot)$, $\beta(\cdot)$, and $b(\cdot)$ expressed in terms of some auxiliary parameters $\theta(\cdot)$ and $\omega(\cdot)$ (see [21, Paragraph 5.2]).

This system was inspired by the Minty representation of maximally monotone operators and the approach due to Attouch-Chbani-Fadili-Riahi [3]. The term $\big(x(\cdot)+\sigma(\cdot)\xi(\cdot)\big)^{(2)}(t)$ acts as a singular perturbation of the possibly degenerated classical Newton continuous dynamical system in which a time scaling parameter $\sigma(\cdot)$ is incorporated. In addition, system (2.3)-(2.4) embeds some geometric damping (through the terms $\dot{\xi}(\cdot)$), coupled with an isotropic damping carried by the coefficient $\alpha(\cdot)$ that can be intended to vanish asymptotically. This system was considered with a particular choice of parameters $\{\theta(\cdot),\sigma(\cdot),\omega(\cdot)\}$ taken such that

$$\theta(t)=\tfrac{\kappa v(t)-\dot{v}(t)}{v(t)+e_*},$$
$$\sigma(t)=\sigma_0 e^{\delta\int_0^t \frac{1}{v(s)+e_*}ds},$$
$$\omega(t)=\Big(\kappa-\tfrac{\dot{v}(t)}{v(t)}\Big)\vartheta(t)-\tfrac{\delta}{v(t)+e_*},$$

where δ is a nonnegative constant, $\{e_*,\sigma_0\}$ are positive constants and $\{v(\cdot),\vartheta(\cdot)\}$ are positive mappings of class C^1. This model has been thoroughly studied in [21] through a first-order reformulation of (2.3) as the following equivalent system:

$$\xi(t)\in\partial f(x(t)),$$
$$\dot{x}(t)+\sigma(t)\dot{\xi}(t)+\theta(t)\,(y(t)-x(t))+\big(\dot{\sigma}(t)+\sigma(t)\omega(t)\big)\xi(t)=0,$$
$$\dot{y}(t)+\kappa(y(t)-x(t))=0.$$

It was shown to produce dynamics $(x,\xi):[0,\infty)\to H^2$ that satisfy the weak convergence of x toward equilibrium together with the strong convergence to zero of both ξ and $\dot{\xi}$, with fast decaying properties. The authors also established remarkable convergence rates (including linear rate) with appropriate settings of parameters. Therefore, it can reasonably be conjectured that the proposed algorithm can inherit some of those interesting convergence properties.

2.3. **Equivalency between CRIPGA and CRIPGA-B.** The following result shows us that the algorithms CRIPGA (1.5)-(1.8) and CRIPGA-B (1.10)-(1.13) are equivalent.

Proposition 2.1. *We have the following two statements:*

(i1) *If $\{x_n,z_n,y_n\}\subset H$ is generated, from $\{z_{-1},x_0,y_0\}$, by (1.5)-(1.8), then, for $n\geq 1$, $\{x_n,z_n,y_n\}$ satisfies (1.10)-(1.13).*

(i2) *If $\{x_n,z_n,y_n\}\subset H$ is generated, from $\{z_{-1},x_0,y_0\}$, by (1.10)-(1.13), then, for $n\geq 1$, $\{x_n,z_n,y_n\}$ satisfies (1.5)-(1.8).*

Proof. For the sake of simplicity we denote $\dot{y}_n=y_n-y_{n-1}$ and $u_n=y_n-x_n$. Let us prove **(i1)**. Observe that, for $n\geq 1$, (1.5) gives us

$$z_{n-1}=x_{n-1}-\theta_{n-1}u_{n-1}+\eta_{n-1}u_{n-1}-\eta_{n-1}(y_{n-1}-z_{n-2}).$$

Note also that x_{n-1} can be simply rewritten as $x_{n-1} = y_n - u_{n-1} - \dot{y}_n$, which by the previous formulation of z_{n-1} yields

$$z_{n-1} = y_n - (1 + \theta_{n-1} - \eta_{n-1})u_{n-1} - \dot{y}_n - \eta_{n-1}(y_{n-1} - z_{n-2}).$$

Hence, by $u_{n-1} = -\kappa^{-1}\dot{y}_n$ (from (1.7)), we obtain

$$\begin{aligned} z_{n-1} &= y_n + \big(\kappa^{-1}(1 + \theta_{n-1} - \eta_{n-1}) - 1\big)\dot{y}_n - \eta_{n-1}(y_{n-1} - z_{n-2}) \\ &= y_n + \alpha_{n-1}\dot{y}_n - \eta_{n-1}(y_{n-1} - z_{n-2}), \end{aligned}$$

where $\alpha_n = \kappa^{-1}(1 + \theta_n - \eta_n) - 1$. This amounts to (1.10). Equalities (1.11) and (1.12) are obvious from (1.6) and (1.7), respectively.

Let us prove **(i2)**. For $n \geq 0$, by (1.12), we have $\dot{y}_{n+1} = \kappa(x_n - y_n)$, that is, (1.7), while the definition of x_n in (1.11) gives us $x_{n+1} = J_{\mu\partial g}(z_n - \mu\nabla f(z_n))$, that is, (1.6). Moreover, by definition of z_n in (1.10) and noticing that y_{n+1} can be simply rewritten as $y_{n+1} = x_n + u_n + \dot{y}_{n+1}$, we have

$$z_n = y_{n+1} + \alpha_n\dot{y}_{n+1} + \eta_n(z_{n-1} - y_n) = x_n + u_n + (1 + \alpha_n)\dot{y}_{n+1} - \eta_n(x_{n-1} - z_{n-1}) - \eta_n u_n.$$

Then, by $\dot{y}_{n+1} = -\kappa u_n$ (from (1.12)), we obtain

$$z_n + \eta_n(x_n - z_{n-1}) = x_n + u_n + (1 + \alpha_n)(-\kappa u_n) - \eta_n u_n = x_n + \big(1 - \kappa(1 + \alpha_n) - \eta_n\big)u_n.$$

Thus, noticing that $\theta_n = -1 + \kappa(1 + \alpha_n) + \eta_n$, we deduce (1.5) □

3. Introducing a General Framework for the Study of the Method

In this section, we set out general conditions on the parameters that guarantee the existence of a discrete-like energy sequence in connection with the sequences $\{x_n, y_n, z_n\}$ generated by the proposed algorithm (1.5)-(1.8). This approach will be crucial for a Lyapunov analysis of the model. As recurrent quantities involved in our computations, we introduce the sequences $(\rho_n)_{n\geq 0}$ and $(\tau_n)_{n\geq 0}$ defined by

$$\rho_n = 1 - (1 - \kappa)\tfrac{\nu_{n+1}}{\nu_n}, \quad \tau_n = e_* + \nu_{n+1}. \tag{3.1}$$

Our methodology requires that there exists some integer N such that :

$$(\nu_n)_{n\geq N} \text{ is non-decreasing,} \tag{3.2}$$

$$(1 - \kappa)\tfrac{\nu_{n+1}}{\nu_n} < 1 \quad \text{for } n \geq N. \tag{3.3}$$

We also suppose for some $\lambda \in [0,1)$ and $(\beta_n)_{n\geq N} \subset [0,\infty)$ that, for $n \geq N$:

$$\eta_n \leq 1 - e_*\rho_n\tau_n^{-1}, \tag{3.4}$$

$$\eta_{n+1} \leq \lambda\,\tau_n\tau_{n+1}^{-1}, \tag{3.5}$$

$$[\rho_n^{-1}\tau_n^2(1 - \eta_n)]_\circ \leq (1 - \lambda)\beta_n\tau_n. \tag{3.6}$$

As a last condition on η_n we assume for some $\varepsilon \in [0,1)$ that :

$$\lim_{n\to\infty} \eta_n = \varepsilon. \tag{3.7}$$

Remark 3.1. Observe that condition (3.3) ensures that $(\rho_n)_{n\geq N}$ (given by (3.1)) is a positive sequence.

3.1. **A general result for a Lyapunov analysis.** Let us recall that L denotes a Lipschitz constant of the gradient of f (from condition (1.3)) and that S stands for the set of minimizers of the objective $\Theta := f+g$. As a well-known trick for evaluating sub-gradient values in proximal processes relative to structured optimization (see, for instance, [2, 6]), we introduce the operator $M_\mu(.)$ defined for $z \in H$ by

$$M_\mu(z) = \mu^{-1}\Big(z - J_{\mu\partial g}\big(z-\mu\nabla f(z)\big)\Big). \tag{3.8}$$

We recall that $M_\mu(z)=0$ is equivalent to $z\in S$.

In connection with the sequences $\{x_n, y_n, z_n\}$ produced by (1.5)-(1.8), under conditions (1.2)-(1.4) and (3.2)-(3.6), we introduce the following sequences $(T_n(s))_{n\geq 0}$ and $(E_n(s,q))_{n\geq 0}$ defined for $(s,q)\in[0,\infty)\times S$, $\delta\in[0,1]$ and for $n\geq 0$ by

$$\begin{aligned} T_n(s) &= \tfrac{1}{2}(1-\kappa)\rho_n\nu_n\nu_{n+1}\Big\|y_n - x_n + \Big(\tfrac{1}{\theta_n} - \tfrac{e_*-s}{\nu_n\rho_n}\Big)\dot{x}_{n+1}\Big\|^2 \\ &\quad + \tfrac{1}{2}(e_*-s)\big((e_*+2\nu_{n+1})\rho_n^{-1} + s(\rho_n^{-1}-1)\big)\|\dot{x}_{n+1}\|^2, \end{aligned} \tag{3.9}$$

$$\begin{aligned} E_n(s,q) &= \tfrac{1}{2}\|s(q-x_n)+\nu_n(y_n-x_n)\|^2 + \tfrac{1}{2}s(e_*-s)\|x_n-q\|^2 \\ &\quad + s\mu\eta_n(e_*+\nu_{n+1})\langle M_\mu(z_{n-1}), x_n-q\rangle \\ &\quad + \mu(e_*+\nu_n)^2\big(\rho_{n-1}^{-1}(1-\eta_{n-1}) - \tfrac{s\lambda}{e_*+\nu_n}\big)\big(\Theta(x_n)-\min\Theta\big) \\ &\quad + \tfrac{1}{2}\eta_n\rho_n^{-1}(e_*+\nu_{n+1})^2\big(\delta(1-\eta_n)+\mu L\big)\|\mu M_\mu(z_{n-1})\|^2. \end{aligned} \tag{3.10}$$

We focus here on establishing Lyapunov properties for $E_n(s,q)$. The form of this energy functional is well-known in the literature when $\lambda=0$ (see, for instance, [1, 4, 11, 30]).

For the sake of legibility, we introduce the set of conditions below :

$$(\mathbf{CP}):\left[\begin{array}{l} \text{Let (1.2)-(1.4) hold (for some } L\geq 0\text{), let } \{\kappa, e_*, \mu, \nu_n\}\subset(0,\infty) \text{ and } \eta_n\in[0,1) \text{ for} \\ n \text{ large enough.} \\ \text{Let (3.2)-(3.6) hold for some } (N,\lambda)\in \mathbb{N}\times[0,1) \text{ and } (\beta_n)\subset[0,\infty). \end{array}\right.$$

An important preliminary result regarding the discrete-like energy sequence $(E_n(s,q))$ is given below.

Lemma 3.2. *Suppose that the set of conditions* **(CP)** *is satisfied. Assume in addition that $\{x_n,y_n,z_n\}\subset H$ are generated by (1.5)-(1.8). Then, for any $\delta\in[0,1]$, $(s,q)\in[0,e_*]\times S$ and $n\geq N$, we have*

$$\begin{aligned} &\dot{E}_{n+1}(s,q) + (1-\lambda)\mu\big(s-\beta_n\big)(e_*+\nu_{n+1})\big(\Theta(x_n)-\min\Theta\big) \\ &\quad + \tfrac{1}{2}(1-\delta)\mu^2\rho_n^{-1}\tau_n^2\|M_\mu(z_n)-\eta_n M_\mu(z_{n-1})\|^2 \\ &\quad + \tfrac{1}{2}\mu^2(e_*+\nu_{n+1})^2\varpi_{1,n}\|M_\mu(z_n)\|^2 \leq -T_n(s), \end{aligned} \tag{3.11}$$

where $T_n(s)$ and $E_n(s,q)$ are given by (3.9) and (3.10), respectively, while $\varpi_{1,n}$ is defined by

$$\varpi_{1,n} = \rho_n^{-1}\big(\delta(1-\eta_n)-\mu L\big) - \eta_{n+1}\rho_{n+1}^{-1}\big(\tfrac{e_*+\nu_{n+2}}{e_*+\nu_{n+1}}\big)^2\big(\delta(1-\eta_{n+1})+\mu L\big). \tag{3.12}$$

This result will be proved in the next section.

3.2. **Proof of Lemma 3.2.** In order to prove Lemma 3.2, we first show that (1.5)-(1.8) fits into an abstract model for which estimations were established in [20, Proposition 3.2]. Next, we state other preliminary results regarding the objective values and the considered parameters.

Finally, estimations on our numerical method will be derived as a special case of the abstract model.

3.2.1. *A key equality for an abstract model.* Estimates on the sequences produced by (1.5)-(1.8) can be obtained through the wider framework of sequences $\{x_n, y_n, d_n\} \subset H$ verifying

$$\theta_n = \frac{\nu_n - (1-\kappa)\nu_{n+1}}{(e_* + \nu_{n+1})}, \qquad \dot{x}_{n+1} + d_n + \theta_n(y_n - x_n) = 0, \qquad \dot{y}_{n+1} = -\kappa(y_n - x_n), \tag{3.13}$$

where $\{\kappa, e_*, \nu_n\}$ are positive parameters. To this aim, we recall the next proposition that was established in [20, Proposition 3.2].

Proposition 3.3. *Let $N \in \mathbb{N}$, $\{e_*, \kappa, \nu_n\} \subset (0,\infty)$, let $\{x_n, y_n, d_n\} \subset H$, and suppose for $n \geq N$ that (3.3) and (3.13) hold. Then, for $(s,q) \in [0,\infty) \times H$ and for $n \geq N$, we have*

$$\begin{aligned} &\dot{G}_{n+1}(s,q) + \tfrac{1}{2}\rho_n^{-1}(e_* + \nu_{n+1})^2 \|\dot{x}_{n+1} + \theta_n(y_n - x_n)\|^2 + T_n(s) \\ &\quad + s(e_* + \nu_{n+1})\langle d_n, x_{n+1} - q\rangle + \left(\rho_n^{-1} - s\tfrac{1}{e_* + \nu_{n+1}}\right)(e_* + \nu_{n+1})^2 \langle d_n, \dot{x}_{n+1}\rangle = 0, \end{aligned} \tag{3.14}$$

where ρ_n and $T_n(s)$ are given by (3.1) and (3.9), respectively, while $G_n(s,q)$ is defined by

$$G_n(s,q) = \tfrac{1}{2}\|s(q - x_n) + \nu_n(y_n - x_n)\|^2 + \tfrac{1}{2}s(e_* - s)\|x_n - q\|^2. \tag{3.15}$$

3.2.2. *Link between the method and the abstract model.* The next results are intended to prove that (1.5)-(1.8) fits into the abstract model (3.13) with some particular sequence (d_n). We begin by giving some elementary facts that are set up as classical results regarding the (proximal) formulation (1.6). The proof is given below for the sake of completeness.

Proposition 3.4. *For any sequences $\{x_n, y_n, z_n\}$ generated by the scheme (1.5)-(1.8) with parameters $\{\kappa, e_*, \mu, \nu_n\} \subset (0,\infty)$ and $(\eta_n) \subset [0,1)$, we have, for $n \geq 0$,*

$$z_n - x_{n+1} = \mu M_\mu(z_n), \tag{3.16}$$

$$M_\mu(z_n) \in \partial g(x_{n+1}) + \nabla f(z_n). \tag{3.17}$$

Proof. Using the definition of M_μ (given in (3.8)) together with (1.6) yields

$$\mu M_\mu(z_n) = z_n - J_{\mu\partial g}\big(z_n - \mu\nabla f(z_n)\big) = z_n - x_{n+1}, \text{ that is, (3.16).}$$

Furthermore, (1.6) writes $x_{n+1} = (I + \mu\partial g)^{-1}\big(z_n - \mu\nabla f(z_n)\big)$. Thus, we get $z_n - \mu\nabla f(z_n) - x_{n+1} \in \mu\partial g(x_{n+1})$, or alternatively $z_n - x_{n+1} \in \mu\left(\partial g(x_{n+1}) + \nabla f(z_n)\right)$, which in light of (3.16) leads us to (3.17). □

Now, we are able to claim the following result.

Proposition 3.5. *For any $\{x_n, y_n, z_n\} \subset H$ generated by (1.5)-(1.8) with parameters $\{\kappa, e_*, \mu, \nu_n\} \subset (0,\infty)$ and $(\eta_n) \subset [0,1)$, we have, for $n \geq 1$,*

$$\dot{x}_{n+1} + \theta_n(y_n - x_n) + \mu M_\mu(z_n) - \mu\eta_n M_\mu(z_{n-1}) = 0. \tag{3.18}$$

Proof. For $n \geq 1$, we have $x_{n+1} - z_n = -\mu M_\mu(z_n)$ (according to (3.16)), hence replacing z_n in the left side of this last equality with its formulation in (1.5) (namely $z_n = x_n - \theta_n(y_n - x_n) + \eta_n(z_{n-1} - x_n)$) yields

$$\dot{x}_{n+1} + \theta_n(y_n - x_n) - \eta_n(z_{n-1} - x_n) = -\mu M_\mu(z_n),$$

which, by $z_{n-1} - x_n = \mu M_\mu(z_{n-1})$ (using again (3.16)), amounts to (3.18). □

Remark 3.6. As a straightforward consequence of Proposition 3.5, we infer that the discrete model (1.5)-(1.8) fits into the special instance of the abstract model (3.13) with the particular term d_n defined by

$$d_n = \mu M_\mu(z_n) - \mu\eta_n M_\mu(z_{n-1}). \tag{3.19}$$

3.2.3. *Intermediate estimates from the model and the convex part of the problem.* For the sake of simplicity and convenience of the reader (regarding our forthcoming computations), we introduce the following notations:

$$\xi_n := \mu M_\mu(z_{n-1}), \quad \tau_n := e_* + \nu_{n+1}, \quad c_n := \rho_n^{-1}\tau_n^2, \quad \bar{\vartheta}_n(s) = 1 - s\tfrac{\rho_n}{\tau_n}, \tag{3.20}$$

$$U_n := \langle \xi_n, x_n - q\rangle \quad \text{and} \quad W_n(s) = \rho_n^{-1}\tau_n^2\left(1 - \eta_n - s\lambda\,\tfrac{\rho_n}{\tau_n}\right). \tag{3.21}$$

Several useful estimates and inequalities are established, through the following parts **(A1)**, **(A2)** and **(A3)** :

(A1) An intermediate equality from the general model. To get estimations, we begin by applying Proposition 3.3 to the special framework of sequences $\{x_n, y_n\}$ generated by (1.5)-(1.8).

Proposition 3.7. *For any $\{x_n, y_n, z_n\} \subset H$ generated by (1.5)-(1.8) with parameters $\{\kappa, e_*, \mu, \nu_n\} \subset (0,\infty)$ and $(\eta_n) \subset [0,1)$, we have, for $n \geq 0$,*

$$\begin{aligned}[G_{n+1}(s,q) + s\eta_{n+1}\tau_{n+1}U_{n+1}]_\circ + \tfrac{1}{2}c_n\|\dot{x}_{n+1} + \theta_n(y_n - x_n)\|^2 \\ + R_n + \eta_n c_n\langle\dot{\xi}_{n+1}, \dot{x}_{n+1}\rangle = -T_n(s),\end{aligned} \tag{3.22}$$

where R_n is defined by

$$R_n := \big(s\tau_n - \eta_{n+1}s\tau_{n+1}\big)U_{n+1} + c_n(\bar{\vartheta}_n(s) - \eta_n)\langle\xi_{n+1}, \dot{x}_{n+1}\rangle. \tag{3.23}$$

Proof. As afore-mentioned through Remark 3.6, we know that numerical scheme (1.5)-(1.8) fits into abstract model (3.13), along with $d_n = \xi_{n+1} - \eta_n\xi_n$. Then, given $(s,q) \in [0, e_*] \times S$, by Proposition 3.3 and noticing that $(\rho_n)_{n\geq N}$ is a positive sequence (from Remark 3.1), we get (for $n \geq N$)

$$\dot{G}_{n+1}(s,q) + \tfrac{1}{2}c_n\|\dot{x}_{n+1} + \theta_n(y_n - x_n)\|^2 + T_n(s) + Q_n = 0, \tag{3.24}$$

where Q_n is defined by

$$Q_n := s\tau_n\langle d_n, x_{n+1} - q\rangle + c_n\bar{\vartheta}_n(s)\langle d_n, \dot{x}_{n+1}\rangle. \tag{3.25}$$

Let us proceed by estimating the quantity Q_n defined above. For simplification purposes, we also set $P_n := \langle\dot{\xi}_{n+1}, \dot{x}_{n+1}\rangle$ and $Z_n = \langle\xi_{n+1}, \dot{x}_{n+1}\rangle$. On the one hand, by $d_n := \xi_{n+1} - \eta_n\xi_n$ we have $\langle d_n, x_{n+1} - q\rangle = \langle\xi_{n+1}, x_{n+1} - q\rangle - \eta_n\langle\xi_n, x_{n+1} - q\rangle$, namely

$$\begin{aligned}\langle d_n, x_{n+1} - q\rangle &= U_{n+1} - \eta_n\big(\langle\xi_n, x_n - q\rangle - \langle\dot{\xi}_{n+1}, \dot{x}_{n+1}\rangle + \langle\xi_{n+1}, \dot{x}_{n+1}\rangle\big) \\ &= U_{n+1} - \eta_n U_n + \eta_n P_n - \eta_n Z_n.\end{aligned}$$

On the other hand, using again the above formulation of d_n we have

$$\begin{aligned}\langle d_n, \dot{x}_{n+1}\rangle &= \langle\xi_{n+1}, \dot{x}_{n+1}\rangle - \eta_n\langle\xi_n, \dot{x}_{n+1}\rangle \\ &= Z_n - \eta_n\big(\langle\xi_n - \xi_{n+1}, \dot{x}_{n+1}\rangle + \langle\xi_{n+1}, \dot{x}_{n+1}\rangle\big) = (1 - \eta_n)Z_n + \eta_n P_n.\end{aligned}$$

So, by the definition of Q_n (given in (3.25)) along with the previous two equalities, we infer that

$$\begin{aligned}Q_n &= s\tau_n\big(U_{n+1} - \eta_n U_n + \eta_n P_n - \eta_n Z_n\big) + c_n\bar{\vartheta}_n(s)\big((1 - \eta_n)Z_n + \eta_n P_n\big) \\ &= s\tau_n U_{n+1} - \eta_n s\tau_n U_n + t_{1,n}Z_n + t_{2,n}P_n,\end{aligned} \tag{3.26}$$

where $t_{1,n} := (1-\eta_n)c_n\bar{\vartheta}_n(s) - s\eta_n\tau_n$ and $t_{2,n} := \eta_n\big(s\tau_n + c_n\bar{\vartheta}_n(s)\big)$. Let us simplify the formulation of the parameters $t_{1,n}$ and $t_{2,n}$. Concerning $t_{1,n}$, we easily notice that

$$s\tau_n + c_n\bar{\vartheta}_n(s) = c_n \text{ (according to the definition of } \bar{\vartheta}_n(s) \text{ given by (3.20))}, \tag{3.27}$$

hence, by definition of $t_{1,n}$, we obtain

$$t_{1,n} = c_n\bar{\vartheta}_n(s) - \eta_n\big(c_n\bar{\vartheta}_n(s) + s\tau_n\big) = c_n\big(\bar{\vartheta}_n(s) - \eta_n\big). \tag{3.28}$$

Concerning $t_{2,n}$, by its definition and using again equality (3.27) we simply get

$$t_{2,n} = \eta_n\big(s\tau_n + c_n\bar{\vartheta}_n(s)\big) = \eta_n c_n. \tag{3.29}$$

As a result, from (3.26) in light of (3.28) and (3.29), we deduce

$$\begin{aligned} Q_n &= s\tau_n U_{n+1} - \eta_n s\tau_n U_n + c_n(\bar{\vartheta}_n(s) - \eta_n)Z_n + \eta_n c_n P_n \\ &= \eta_{n+1}s\tau_{n+1}U_{n+1} - \eta_n s\tau_n U_n + \big(s\tau_n - \eta_{n+1}s\tau_{n+1}\big)U_{n+1} \\ &\quad + c_n(\bar{\vartheta}_n(s) - \eta_n)Z_n + \eta_n c_n P_n. \end{aligned} \tag{3.30}$$

Therefore, by (3.24), together with (3.30) and denoting $E_{1,n}(s,q) := G_n(s,q) + s\eta_n\tau_n U_n$, we obtain

$$\begin{aligned} &-T_n(s) - \tfrac{1}{2}c_n\|\dot{x}_{n+1} + \theta_n(y_n - x_n)\|^2 \\ &\quad = \dot{G}_{n+1}(s,q) + Q_n \\ &\quad = \dot{E}_{1,n+1}(s,q) + \big(s\tau_n - \eta_{n+1}s\tau_{n+1}\big)U_{n+1} + c_n(\bar{\vartheta}_n(s) - \eta_n)Z_n + \eta_n c_n P_n. \end{aligned}$$

This leads us to (3.22)-(3.23). □

(A2) Additional estimates from the convex properties of the objective. As a classical result used in optimization theory, we recall some important inequality (relative to the objective function Θ arising in (1.1) and the operator M_μ introduced in (3.8)) through the next proposition. Its proof can be found in [6, 20] and is given here for the sake of completeness.

Proposition 3.8. *Let μ be any positive constant and suppose that conditions (1.2)-(1.4) hold. Then for any $(x,z) \in H^2$, we have*

$$\Theta\big(z - \mu M_\mu(z)\big) \le \Theta(x) + \langle M_\mu(z), z - x\rangle - \mu\left(1 - \frac{L\mu}{2}\right)\|M_\mu(z)\|^2. \tag{3.31}$$

Proof. Given $(x,z) \in H^2$, we set $y = z - \mu M_\mu(z)$ and $v = y - z$. Applying the Taylor expansion formula yields

$$f(y) - f(z) = \langle \nabla f(z), y - z\rangle + \int_0^1 \langle y - z, \nabla f(z + hv) - \nabla f(z)\rangle dh.$$

Hence, by L-Lipschitz continuity of ∇f and noticing that $\|y - z\| = \|v\| = \mu\|M_\mu(z)\|$, we readily deduce

$$f(y) - f(z) \le \langle \nabla f(z), y - z\rangle + \frac{1}{2}L\mu^2\|M_\mu(z)\|^2.$$

Furthermore using the convex inequality simply gives us $f(z) - f(x) \le \langle \nabla f(z), z - x\rangle$. So adding the previous two inequalities, we obtain

$$f(y) - f(x) \le \langle \nabla f(z), y - x\rangle + \tfrac{1}{2}L\mu^2\|M_\mu(z)\|^2. \tag{3.32}$$

Now, we recall that $M_\mu(z) = \mu^{-1}\Big(z - \mathrm{prox}_{\mu g}(z - \mu\nabla f(z))\Big)$. It is a classical matter to see that this formulation, in light of $y = z - \mu M_\mu(z)$, can be rewritten as $M_\mu(z) - \nabla f(z) \in \partial g(y)$. Then

setting $\zeta := M_\mu(z) - \nabla f(z)$ $\big($hence $\zeta \in \partial g(y)\big)$ and again using the convex inequality we clearly deduce that $g(y) - g(x) \leq \langle \zeta, y - x\rangle$, together with $y - x = z - x - \mu M_\mu(z)$. As a consequence, we also obtain

$$g(y) - g(x) \leq \langle M_\mu(z), y - x\rangle - \langle \nabla f(z), y - x\rangle = \langle M_\mu(z), z - x\rangle - \mu\|M_\mu(z)\|^2 - \langle \nabla f(z), y - x\rangle. \tag{3.33}$$

Adding (3.32) and (3.33) amounts immediately to the desired inequality. □

As straightforward consequences of the above proposition, we deduce the next useful estimations.

Proposition 3.9. *Let $q \in H$. Suppose that (1.2)-(1.4) hold and that $\{x_n, z_n\}$ are generated by (1.5)-(1.8) with $\{\kappa, e_*, \mu, \nu_n\} \subset (0,\infty)$ and $(\eta_n) \subset [0,\infty)$. Then, for $n \geq 1$, we have*

$$\langle M_\mu(z_n), x_{n+1} - q\rangle \geq \Theta(x_{n+1}) - \Theta(q) - \tfrac{1}{2}L\mu^2\|M_\mu(z_n)\|^2, \tag{3.34}$$

$$\langle M_\mu(z_n), \dot{x}_{n+1}\rangle \geq \Theta(x_{n+1}) - \Theta(x_n) - \tfrac{1}{2}L\mu^2\|M_\mu(z_n)\|^2, \tag{3.35}$$

$$\langle M_\mu(z_n) - M_\mu(z_{n-1}), \dot{x}_{n+1}\rangle \geq -\tfrac{1}{2}L\mu^2\big(\|M_\mu(z_n)\|^2 + \|M_\mu(z_{n-1})\|^2\big). \tag{3.36}$$

Proof. Let us denote $\xi_n := \mu M_\mu(z_{n-1})$. By applying (3.31) to the respective couples $(z = z_n, x = q)$ and $(z = z_n, x = x_n)$, we get

$$\Theta(z_n - \xi_{n+1}) \leq \Theta(q) + \mu^{-1}\langle \xi_{n+1}, z_n - q\rangle - \mu^{-1}\left(1 - \tfrac{L\mu}{2}\right)\|\xi_{n+1}\|^2, \tag{3.37}$$

$$\Theta(z_n - \xi_{n+1}) \leq \Theta(x_n) + \mu^{-1}\langle \xi_{n+1}, z_n - x_n\rangle - \mu^{-1}\left(1 - \tfrac{L\mu}{2}\right)\|\xi_{n+1}\|^2. \tag{3.38}$$

These inequalities together with $x_{n+1} = z_n - \xi_{n+1}$ (from (3.16)) can be rewritten as

$$\Theta(x_{n+1}) \leq \Theta(q) + \mu^{-1}\langle \xi_{n+1}, x_{n+1} - q\rangle + \tfrac{1}{2}L\|\xi_{n+1}\|^2,$$
$$\Theta(x_{n+1}) \leq \Theta(x_n) + \mu^{-1}\langle \xi_{n+1}, x_{n+1} - x_n\rangle + \tfrac{1}{2}L\|\xi_{n+1}\|^2.$$

It immediately follows (3.34) and (3.35), respectively. It remains to prove item (3.36). Using (3.31) for $z = z_{n-1}$ and $x = x_{n+1}$, while noticing that $x_n = z_{n-1} - \xi_n$, we get

$$\Theta(x_n) \leq \Theta(x_{n+1}) + \mu^{-1}\langle \xi_n, z_{n-1} - x_{n+1}\rangle - \mu^{-1}\left(1 - \tfrac{L\mu}{2}\right)\|\xi_n\|^2,$$

hence, by $z_{n-1} = x_n + \xi_n$, we obtain $-\mu^{-1}\langle \xi_n, \dot{x}_{n+1}\rangle \geq \Theta(x_n) - \Theta(x_{n+1}) - \frac{L}{2}\|\xi_n\|^2$, which can be alternatively written as

$$-\langle M_\mu(z_{n-1}), \dot{x}_{n+1}\rangle \geq -\big(\Theta(x_{n+1}) - \Theta(x_n)\big) - \frac{1}{2}L\mu^2\|M_\mu(z_{n-1})\|^2. \tag{3.39}$$

Consequently, summing inequalities (3.35) and (3.39) amounts to (3.36). □

Remark 3.10. Two helpful observations are given below:

(r1) It is of importance to notice that $\{z_n, x_n, y_n\}$ generated by (1.5)-(1.8) verify

$$x_{n+1} - z_n = -\mu M_\mu(z_n), \quad (\text{for } n \geq 0). \tag{3.40}$$

(r2) It can be checked (see, for instance, [2, Lemma B1]), when ∇f is L-Lipschitz continuous (hence ∇f is L^{-1}-cocoercive) that for $\mu \in (0, 2L^{-1})$, the operator M_μ is σ-cocoercive with $\sigma = \mu\left(1 - \frac{\mu L}{4}\right)$.

(A3) Some classical inequalities. Basic but useful inequalities are also set up through the following proposition.

Proposition 3.11. *For any $(a,b) \in H^2$ and $t \in [0,\infty)$ we have*

$$\|a+tb\|^2 \leq (1+t)\|a\|^2 + t(1+t)\|b\|^2, \tag{3.41}$$
$$\|a-tb\|^2 \geq (1-t)\|a\|^2 - t(1-t)\|b\|^2. \tag{3.42}$$

Proof. Indeed, we classically have

$$\|a+tb\|^2 = \|a\|^2 + t^2\|b\|^2 + 2t\langle a,b\rangle,$$
$$\|a-tb\|^2 = \|a\|^2 + t^2\|b\|^2 - 2t\langle a,b\rangle,$$

In addition, Young's inequality gives us $2\langle a,b\rangle \leq \|a\|^2 + \|b\|^2$. Consequently, we readily deduce

$$\|a+tb\|^2 \leq \|a\|^2 + t^2\|b\|^2 + t\|a\|^2 + t\|b\|^2 = (1+t)\|a\|^2 + (t+t^2)\|b\|^2,$$
$$\|a-tb\|^2 \geq \|a\|^2 + t^2\|b\|^2 - t\|a\|^2 - t\|b\|^2 = (1-t)\|a\|^2 + (t^2-t)\|b\|^2,$$

which leads to (3.41) and (3.42). □

3.2.4. *Proving Lemma 3.2.* We keep on using the notations in (3.20). According to Proposition 3.7 we know that, for $\{x_n, y_n, z_n\}$ generated by (1.5)-(1.8), and for $n \geq 0$, we have

$$\dot{E}_{1,n+1}(s,q) + \tfrac{1}{2}c_n\|\dot{x}_{n+1} + \theta_n(y_n - x_n)\|^2 + R_n + \eta_n c_n\langle\dot{\xi}_{n+1}, \dot{x}_{n+1}\rangle = -T_n(s), \tag{3.43}$$

where $E_{1,n} := G_n + s\eta_n\tau_n U_n$, and R_n is defined by

$$R_n := \big(s\tau_n - \eta_{n+1}s\tau_{n+1}\big)U_{n+1} + c_n\big(\bar{\vartheta}_n(s) - \eta_n\big)\langle\xi_{n+1}, \dot{x}_{n+1}\rangle. \tag{3.44}$$

For simplifications, given $s \in [0, e_*]$ and $\lambda \in [0,1)$, we set

$$a_n(s) := s\big(\tau_n - \tau_{n+1}\eta_{n+1}\big), \tag{3.45}$$
$$b_n(s) := c_n(\bar{\vartheta}_n(s) - \eta_n) = c_n\left(1 - s\frac{\rho_n}{\tau_n} - \eta_n\right). \tag{3.46}$$

Observe that $W_n(s)$ (introduced in (3.21)) can be rewritten as

$$W_n(s) = \rho_n^{-1}\tau_n^2\left(1 - \eta_n - s\lambda\frac{\rho_n}{\tau_n}\right) = c_n(1-\eta_n) - \lambda s\tau_n = (1-\lambda)s\tau_n + b_n(s), \tag{3.47}$$

which, recalling that $\dot{\tau}_n \geq 0$ (from condition (3.2)) entails

$$\dot{W}_n(s) = [c_n(1-\eta_n)]_\circ - \lambda s\dot{\tau}_n \leq [c_n(1-\eta_n)]_\circ. \tag{3.48}$$

Let us recall that the assumption (3.4)-(3.6) on parameters writes :

$$1 - e_*\rho_n\tau_n^{-1} - \eta_n \geq 0 \qquad \big(\Rightarrow b_n(s) \geq b_n(e_*) \geq 0\big), \tag{3.49}$$
$$\eta_{n+1} \leq \lambda\,\tau_n\tau_{n+1}^{-1} \qquad \big(\Rightarrow a_n(s) \geq s(1-\lambda)\tau_n\big), \tag{3.50}$$
$$[c_n(1-\eta_n)]_\circ \leq (1-\lambda)\beta_n\tau_n \quad \big(\Rightarrow \dot{W}_n(s) \leq (1-\lambda)\beta_n\tau_n \text{ (from (3.48))}\big). \tag{3.51}$$

Clearly, the above conditions ensure that $a_n(s)$ and $b_n(s)$ are nonnegative sequences. The rest of the proof can be divided into the following steps **(b1)** and **(b2)**:

(b1) Set $\bar{\Theta} := \Theta - \min\Theta$ and let us prove that the quantity R_n (given by (3.44)) satisfies

$$\mu^{-1}R_n \geq [W_n(s)\bar{\Theta}(x_{n+1})]_\circ + \big((1-\lambda)s\tau_n - \dot{W}_n(s)\big)\bar{\Theta}(x_n) - \big(a_n(s) + b_n(s)\big)\big(\tfrac{1}{2}L\|\xi_{n+1}\|^2\big). \tag{3.52}$$

Indeed, by definition of R_n (given by (3.44)), we have

$$\mu^{-1}R_n = a_n(s)\langle M_\mu(z_n), x_{n+1}-q\rangle + b_n(s)\langle M_\mu(z_n), \dot{x}_{n+1}\rangle. \tag{3.53}$$

In addition, by inequalities (3.34) and (3.35) of Proposition 3.9, we have

$$\langle M_\mu(z_n), x_{n+1}-q\rangle \geq \bar{\Theta}(x_{n+1}) - \tfrac{1}{2}L\|\xi_{n+1}\|^2, \tag{3.54}$$

$$\langle M_\mu(z_n), \dot{x}_{n+1}\rangle \geq \bar{\Theta}(x_{n+1}) - \bar{\Theta}(x_n) - \tfrac{1}{2}L\|\xi_{n+1}\|^2. \tag{3.55}$$

Then, for $n \geq N$, combining (3.53), (3.54) and (3.55), in light of $a_n(s) \geq 0$ and $b_n(s) \geq 0$, yields

$$\mu^{-1}R_n \geq a_n(s)\bar{\Theta}(x_{n+1}) + b_n(s)\big(\bar{\Theta}(x_{n+1}) - \bar{\Theta}(x_n)\big) - \big(a_n(s)+b_n(s)\big)\left(\frac{1}{2}L\|\xi_{n+1}\|^2\right). \tag{3.56}$$

Consequently, by $a_n(s) \geq s(1-\lambda)\tau_n$ (from (3.50)) and $W_n(s) = s(1-\lambda)\tau_n + b_n(s)$ (from (3.47)), we obtain

$$\begin{aligned}
&a_n(s)\bar{\Theta}(x_{n+1}) + b_n(s)\big(\bar{\Theta}(x_{n+1}) - \bar{\Theta}(x_n)\big)\\
&\quad\geq s(1-\lambda)\tau_n\bar{\Theta}(x_{n+1}) + b_n(s)\big(\bar{\Theta}(x_{n+1}) - \bar{\Theta}(x_n)\big)\\
&\quad= s(1-\lambda)\tau_n\bar{\Theta}(x_{n+1}) - s(1-\lambda)\tau_n\big(\bar{\Theta}(x_{n+1}) - \bar{\Theta}(x_n)\big)\\
&\qquad+\big(s(1-\lambda)\tau_n + b_n(s)\big)\big(\bar{\Theta}(x_{n+1}) - \bar{\Theta}(x_n)\big)\\
&\quad= s(1-\lambda)\tau_n\bar{\Theta}(x_n) + W_n(s)\big(\bar{\Theta}(x_{n+1}) - \bar{\Theta}(x_n)\big)\\
&\quad= \big[W_n(s)\bar{\Theta}(x_{n+1})\big]_\circ + \big(s(1-\lambda)\tau_n - \dot{W}_n(s)\big)\bar{\Theta}(x_n).
\end{aligned} \tag{3.57}$$

Then (3.56) in light of (3.57) amounts to (3.52).

(b2) Finally, we are in position to establish (3.11). It is easily noticed from (3.18) and $d_n := \xi_{n+1} - \eta_n\xi_n$ that $\|\dot{x}_{n+1} + \theta_n(y_n - x_n)\|^2 = \|d_n\|^2$, hence equality (3.22) can be reduced to

$$-T_n(s) = \dot{E}_{1,n+1}(s,q) + R_n + \tfrac{1}{2}c_n\|d_n\|^2 + \eta_n c_n\langle\dot{\xi}_{n+1}, \dot{x}_{n+1}\rangle, \tag{3.58}$$

where R_n is defined by (3.44). In addition, setting $\bar{\Theta} := \Theta - \min\Theta$, by inequality (3.52) together with $\dot{W}_n(s) \leq (1-\lambda)\beta_n\tau_n$ (from (3.51)) we readily have

$$R_n \geq \mu\big[W_n(s)\bar{\Theta}(x_{n+1})\big]_\circ + \mu(1-\lambda)\tau_n\big(s-\beta_n\big)\bar{\Theta}(x_n) - \tfrac{1}{2}\mu L\big(a_n(s)+b_n(s)\big)\|\xi_{n+1}\|^2. \tag{3.59}$$

Thus, denoting $E_{2,n} := E_{1,n} + \mu W_{n-1}(s)\overline{\Theta}(x_n)$, by (3.58)-(3.59), we deduce

$$\begin{aligned}
&-T_n(s) - \eta_n c_n\langle\dot{\xi}_{n+1}, \dot{x}_{n+1}\rangle\\
&\quad= \dot{E}_{1,n+1}(s,q) + R_n + \tfrac{1}{2}c_n\|d_n\|^2\\
&\quad\geq \dot{E}_{1,n+1}(s,q) + \mu\big[W_n(s)\bar{\Theta}(x_{n+1})\big]_\circ + \mu(1-\lambda)\tau_n\big(s-\beta_n\big)\bar{\Theta}(x_n)\\
&\qquad-\tfrac{1}{2}\mu L\big(a_n(s)+b_n(s)\big)\|\xi_{n+1}\|^2 + \tfrac{1}{2}c_n\|d_n\|^2\\
&\quad= \dot{E}_{2,n+1}(s,q) + \mu(1-\lambda)\tau_n\big(s-\beta_n\big)\bar{\Theta}(x_n) + \tfrac{1}{2}c_n\|d_n\|^2 - \tfrac{1}{2}\mu L\big(a_n(s)+b_n(s)\big)\|\xi_{n+1}\|^2.
\end{aligned} \tag{3.60}$$

It can be also noticed that

$$a_n(s) + b_n(s) = c_n - s\tau_{n+1}\eta_{n+1} - c_n\eta_n = (1-\eta_n)c_n - s\tau_{n+1}\eta_{n+1} \leq (1-\eta_n)c_n, \tag{3.61}$$

which by (3.60) amounts to

$$\begin{aligned}
-T_n(s) - \eta_n c_n\langle\dot{\xi}_{n+1}, \dot{x}_{n+1}\rangle \geq \dot{E}_{2,n+1}(s,q) + \mu(1-\lambda)\tau_n\big(s-\beta_n\big)\bar{\Theta}(x_n)\\
+\tfrac{1}{2}c_n\|d_n\|^2 - \tfrac{1}{2}\mu L(1-\eta_n)c_n\|\xi_{n+1}\|^2.
\end{aligned} \tag{3.62}$$

We estimate the third term in the right side of the previous inequality. Let us recall that $\eta_n \in [0,1)$. Then, from $d_n = \xi_{n+1} - \eta_n\xi_n$ (according to the definition of d_n) and using Proposition 3.11 we simply get

$$\|d_n\|^2 = \|\xi_{n+1} - \eta_n\xi_n\|^2 \geq (1-\eta_n)\|\xi_{n+1}\|^2 - \eta_n(1-\eta_n)\|\xi_n\|^2. \tag{3.63}$$

Moreover, by (3.36) (from Proposition 3.9), we have

$$\big\langle \dot{\xi}_{n+1}, \dot{x}_{n+1} \big\rangle = \mu \big\langle M_\mu(z_n) - M_\mu(z_{n-1}), \dot{x}_{n+1} \big\rangle \geq -\frac{1}{2}\mu L \bigg(\|\xi_{n+1}\|^2 + \|\xi_n\|^2 \bigg). \tag{3.64}$$

Hence, taking any $\delta \in [0,1]$ and denoting $\bar{\delta} = 1-\delta$, by combining (3.62), (3.63) and (3.64) we obtain

$$\begin{aligned}
&-T_n(s) - \dot{E}_{2,n+1}(s,q) - \mu(1-\lambda)\tau_n\big(s-\beta_n\big)\bar{\Theta}(x_n) \\
&\quad \geq \tfrac{1}{2}c_n\bar{\delta}\|d_n\|^2 + \tfrac{1}{2}\delta c_n\|d_n\|^2 - \tfrac{1}{2}\mu L(1-\eta_n)c_n\|\xi_{n+1}\|^2 + \eta_n c_n\langle \dot{\xi}_{n+1}, \dot{x}_{n+1}\rangle \\
&\quad \geq \tfrac{1}{2}c_n\bar{\delta}\|d_n\|^2 + \tfrac{1}{2}\delta c_n\big((1-\eta_n)\|\xi_{n+1}\|^2 - \eta_n(1-\eta_n)\|\xi_n\|^2\big) \\
&\qquad\quad -\tfrac{1}{2}\mu L(1-\eta_n)c_n\|\xi_{n+1}\|^2 - \tfrac{1}{2}\eta_n c_n\big(\mu L\|\xi_{n+1}\|^2 + \mu L\|\xi_n\|^2\big) \\
&\quad = \tfrac{1}{2}c_n\bar{\delta}\|d_n\|^2 + \tfrac{1}{2}c_n\big(\delta(1-\eta_n) - (1-\eta_n)\mu L - \eta_n\mu L\big)\|\xi_{n+1}\|^2 \\
&\qquad\quad -\tfrac{1}{2}c_n\big(\delta\eta_n(1-\eta_n) + \eta_n\mu L\big)\|\xi_n\|^2 \\
&\quad = \tfrac{1}{2}c_n\bar{\delta}\|d_n\|^2 + \phi_n\|\xi_{n+1}\|^2 - \psi_n\|\xi_n\|^2,
\end{aligned} \tag{3.65}$$

where ϕ_n and ψ_n are defined by

$$\phi_n := \tfrac{1}{2}c_n\big(\delta(1-\eta_n) - \mu L\big), \tag{3.66}$$

$$\psi_n := \tfrac{1}{2}c_n\eta_n\big(\delta(1-\eta_n) + \mu L\big). \tag{3.67}$$

Thus, noticing that $E_n := E_{2,n} + \psi_n\|\xi_n\|^2$ (according to (3.10)), (3.65) leads us to

$$\begin{aligned}
&-T_n(s) - \mu(1-\lambda)\tau_n\big(s-\beta_n\big)\bar{\Theta}(x_n) \\
&\quad \geq \dot{E}_{2,n+1}(s,q) + \phi_n\|\xi_{n+1}\|^2 - \psi_n\|\xi_n\|^2 + \tfrac{1}{2}c_n\bar{\delta}\|d_n\|^2 \\
&\quad = \dot{E}_{2,n+1}(s,q) + \psi_{n+1}\|\xi_{n+1}\|^2 - \psi_n\|\xi_n\|^2 + \big(\phi_n - \psi_{n+1}\big)\|\xi_{n+1}\|^2 + \tfrac{1}{2}c_n\bar{\delta}\|d_n\|^2 \\
&\quad = \dot{E}_{n+1}(s,q) + \big(\phi_n - \psi_{n+1}\big)\|\xi_{n+1}\|^2 + \tfrac{1}{2}c_n\bar{\delta}\|d_n\|^2.
\end{aligned} \tag{3.68}$$

Let us now estimate the coefficient $\phi_n - \psi_{n+1}$ occurring in the right side of the above inequality. Remembering that $c_n = \rho_n^{-1}\tau_n^2$ $\big($hence $\frac{c_{n+1}}{c_n} = \frac{\rho_n\tau_{n+1}^2}{\rho_{n+1}\tau_n^2}\big)$, by (3.66)-(3.67), we readily obtain

$$\begin{aligned}
\phi_n - \psi_{n+1} &= \tfrac{1}{2}c_n\big(\delta(1-\eta_n) - \mu L\big) - \tfrac{1}{2}c_{n+1}\eta_{n+1}\big(\delta(1-\eta_{n+1}) + \mu L\big) \\
&= \tfrac{1}{2}c_n\bigg(\big(\delta(1-\eta_n) - \mu L\big) - \eta_{n+1}\tfrac{c_{n+1}}{c_n}\big(\delta(1-\eta_{n+1}) + \mu L\big)\bigg) \\
&= \tfrac{1}{2}c_n\bigg(\big(\delta(1-\eta_n) - \mu L\big) - \eta_{n+1}\tfrac{\rho_n\tau_{n+1}^2}{\rho_{n+1}\tau_n^2}\big(\delta(1-\eta_{n+1}) + \mu L\big)\bigg).
\end{aligned} \tag{3.69}$$

Combining (3.68) with the above inequality (3.69) leads to the desired result (3.11), which completes the proof. □

4. Preliminary Convergence Analysis in a General Framework

In this section, we conduct a Lyapunov analysis by means of Lemma 3.2 based upon the discrete-like energy sequence (E_n) associated with the method (1.5)-(1.8). This methodology will allow us to derive convergence rates and estimations regarding the sequences generated by the proposed algorithm.

4.1. **Intermediate results for a convergence analysis.** We state in this section important rates and estimations that stand as an intermediary step towards our main results. For convenience of the reader, given $\delta \in [0,1]$ and $\lambda \in [0,1)$ being the constant arising in conditions (3.5)-(3.6), we recall or introduce the definitions of the parameters $\varpi_{i,n}$ and $\zeta_{i,n}(s)$ (for $i = 1,2$) :

$$\varpi_{1,n} = \rho_n^{-1}\big(\delta(1-\eta_n) - \mu L\big) - \eta_{n+1}\rho_{n+1}^{-1}\big(\tfrac{\tau_{n+1}}{\tau_n}\big)^2\big(\delta(1-\eta_{n+1}) + \mu L\big), \quad (4.1)$$

$$\varpi_{2,n} = 1 - \delta + (1-\kappa)\big(1 + \tfrac{\dot{\nu}_{n+1}}{\nu_n}\big), \quad (4.2)$$

$$\zeta_{1,n}(s) = \rho_{n-1}^{-1}(1-\eta_{n-1}) - \tfrac{s\lambda}{\tau_{n-1}}, \quad (4.3)$$

$$\zeta_{2,n}(s) = \rho_n^{-1}\big(\delta(1-\eta_n) + \mu L\big) - s\mu L\tau_n^{-1}. \quad (4.4)$$

The next proposition deals with a preliminary set of estimations regarding the algorithm under consideration within a general setting of parameters.

Lemma 4.1. *Suppose that the set of conditions* **(CP)** *is satisfied. Assume furthermore for some $\delta \in (0,1)$ that, for $n \geq N$:*

$$e_* - \beta_n \geq 0, \quad (4.5)$$

$$\varpi_{i,n} \geq 0, \quad \zeta_{i,n}(e_*) \geq 0 \quad \textit{(for } i = 1,2\textit{)}. \quad (4.6)$$

Then, for any sequences $\{x_n, y_n, z_n\} \subset H$ generated by (1.5)-(1.8) and for $q \in S$, $(E_n(e_*,q))_{n\geq N}$ is non-increasing and convergent. Furthermore, the following estimates are reached :

$$e_*(q - x_n) + \nu_n(y_n - x_n) \textit{ is bounded}, \quad (4.7)$$

$$\big(e_*\eta_n\nu_n + \zeta_{1,n}(e_*)\nu_n^2\big)\big(\Theta(x_n) - \min\Theta\big) \leq \mu^{-1}E_N(e_*,q), \quad (4.8)$$

$$\zeta_{2,n}(e_*)\nu_n^2\|\mu M_\mu(z_{n-1})\|^2 \leq 2E_N(e_*,q), \quad (4.9)$$

$$(1-\lambda)\textstyle\sum_{n\geq N}(e_* - \beta_n)\,\nu_n\big(\Theta(x_n) - \min\Theta\big) \leq \mu^{-1}E_N(e_*,q), \quad (4.10)$$

$$\textstyle\sum_{n\geq N}\rho_n^{-1}\varpi_{2,n}\nu_n^2\|M_\mu(z_n) - \eta_n M_\mu(z_{n-1})\|^2 \leq 2\mu^{-2}E_N(e_*,q), \quad (4.11)$$

$$\textstyle\sum_{n\geq N}\varpi_{1,n}\nu_n^2\|M_\mu(z_n)\|^2 \leq 2\mu^{-2}E_N(e_*,q), \quad (4.12)$$

$$\textstyle\sum_{n\geq N}\rho_n^{-1}\varpi_{2,n}\nu_n^2\|\dot{x}_{n+1} + \theta_n(y_n - x_n)\|^2 \leq 2E_N(e_*,q). \quad (4.13)$$

4.1.1. *Proof of Lemma 4.1.* The proof can be divided into the following steps (a) and (b):

(a) Estimation for the energy sequence. The next result will be useful to set up sufficient conditions on the parameters that ensure the nonnegativity of $(E_n(e_*,q))_{n\geq N}$.

Proposition 4.2. *Suppose that the set of conditions* **(CP)** *is satisfied. Assume in addition that $\{x_n, y_n, z_n\} \subset H$ are generated by (1.5)-(1.8). Then, for any $\delta \in [0,1]$, $(s,q) \in [0,e_*] \times S$ and $n \geq N$,*

$$\begin{aligned} E_n(s,q) \geq\ & \tfrac{1}{2}\|s(q-x_n) + \nu_n(y_n - x_n)\|^2 + \tfrac{1}{2}s(e_* - s)\|x_n - q\|^2 \\ & + \mu\big(s\eta_n\tau_n + \zeta_{1,n}(e_*)\tau_{n-1}^2\big)\big(\Theta(x_n) - \min\Theta\big) + \tfrac{1}{2}\eta_n\tau_n^2\zeta_{2,n}(e_*)\|\xi_n\|^2, \end{aligned} \quad (4.14)$$

where $\zeta_{1,n}(s)$ and $\zeta_{2,n}(s)$ are given by (4.3) and (4.4), respectively.

Proof. Set $\tau_n := e_* + \nu_{n+1}$, $\xi_n = \mu M_\mu(z_{n-1})$ and $\bar{\Theta} := \Theta - \min\Theta$. Using the definitions of $E_n(.,.)$ (given by (3.10)) and $G_n(.,.)$ (given by (3.15)) entails

$$\begin{aligned} E_n(s,q) = G_n(s,q) + s\eta_n\tau_n\langle\xi_n, x_n - q\rangle + \mu\tau_{n-1}^2\zeta_{1,n}(s)\bar{\Theta}(x_n) \\ + \tfrac{1}{2}\eta_n\rho_n^{-1}\tau_n^2\big(\delta(1-\eta_n) + \mu L\big)\|\xi_n\|^2. \end{aligned} \tag{4.15}$$

It can be noticed that $\zeta_{1,n}(s) \geq \zeta_{1,n}(e_*)$ $\big($since $s \in [0, e_*]\big)$, while (3.34) gives us

$$\mu^{-1}\langle\xi_n, x_n - q\rangle \geq \bar{\Theta}(x_n) - \frac{1}{2}L\|\xi_n\|^2.$$

So, (4.15) in light of the previous two arguments and $\eta_n \geq 0$ entails

$$\begin{aligned} &E_n(s,q) - G_n(s,q) \\ &\quad = s\mu\eta_n\tau_n\mu^{-1}\langle\xi_n, x_n - q\rangle + \mu\zeta_{1,n}(s)\tau_{n-1}^2\bar{\Theta}(x_n) + \tfrac{1}{2}\eta_n\rho_n^{-1}\tau_n^2\big(\delta(1-\eta_n) + \mu L\big)\|\xi_n\|^2 \\ &\quad \geq s\mu\eta_n\tau_n\big(\bar{\Theta}(x_n) - \tfrac{1}{2}L\|\xi_n\|^2\big) + \mu\zeta_{1,n}(e_*)\tau_{n-1}^2\bar{\Theta}(x_n) + \tfrac{1}{2}\eta_n\rho_n^{-1}\tau_n^2\big(\delta(1-\eta_n) + \mu L\big)\|\xi_n\|^2 \\ &\quad = \mu\big(s\eta_n\tau_n + \zeta_{1,n}(e_*)\tau_{n-1}^2\big)\bar{\Theta}(x_n) + \tfrac{1}{2}\eta_n\tau_n^2\left(\rho_n^{-1}\big(\delta(1-\eta_n) + \mu L\big) - s\mu L\tau_n^{-1}\right)\|\xi_n\|^2. \end{aligned}$$

This leads us to the desired inequality. □

(b) Proving Lemma 4.1. The proof is mainly based upon Lemma 3.2. For simplification purposes we denote $\tau_n := e_* + \nu_{n+1}$, $\xi_{n+1} := \mu M_\mu(z_n)$ and $d_n := \xi_{n+1} - \eta_n\xi_n$. Clearly, noticing that $\theta_n = \frac{\nu_n}{\tau_n}\rho_n$ and $\frac{\nu_{n+1}}{\nu_n} = 1 + \frac{\dot{\nu}_{n+1}}{\nu_n}$ readily gives us

$$\frac{\rho_n\nu_n\nu_{n+1}}{\theta_n^2} = \rho_n\nu_n\nu_{n+1}\left(\frac{\tau_n}{\nu_n\rho_n}\right)^2 = \rho_n^{-1}\tau_n^2\left(\frac{\nu_{n+1}}{\nu_n}\right) = \rho_n^{-1}\tau_n^2\left(1 + \frac{\dot{\nu}_{n+1}}{\nu_n}\right).$$

Then, by $T_n(e_*) = \frac{1}{2}(1-\kappa)\frac{\rho_n\nu_n\nu_{n+1}}{\theta_n^2}\|\theta_n(y_n - x_n) + \dot{x}_{n+1}\|^2$ (from the definition of T_n), along with $\dot{x}_{n+1} + \theta_n(y_n - x_n) = -d_n$ (from (3.18)), we readily obtain

$$T_n(e_*) = \frac{1}{2}(1-\kappa)\rho_n^{-1}\tau_n^2(1 + \frac{\dot{\nu}_{n+1}}{\nu_n})\|d_n\|^2. \tag{4.16}$$

Hence, for $n \geq N$, by Lemma 3.2 with $s = e_*$ together with the equality (3.18) while using the notation $\bar{\delta} = 1 - \delta$ and setting $\Gamma_n := \dot{E}_{n+1}(e_*, q) + (1-\lambda)\mu\big(e_* - \beta_n\big)\tau_n\bar{\Theta}(x_n)$ where $\bar{\Theta} := \Theta - \min\Theta$, we infer that

$$\Gamma_n + \frac{1}{2}\bar{\delta}c_n\|d_n\|^2 + \frac{1}{2}(e_* + \nu_{n+1})^2\varpi_{1,n}\|\xi_{n+1}\|^2 \leq -T_n(e_*), \tag{4.17}$$

which by (4.16) yields

$$\begin{aligned} -\Gamma_n &\geq \tfrac{1}{2}\bar{\delta}c_n\|d_n\|^2 + T_n(e_*) + \tfrac{1}{2}\tau_n^2\varpi_{1,n}\|\xi_{n+1}\|^2 \\ &= \tfrac{1}{2}\bar{\delta}c_n\|d_n\|^2 + \tfrac{1}{2}(1-\kappa)\rho_n^{-1}\tau_n^2\left(1 + \tfrac{\dot{\nu}_{n+1}}{\nu_n}\right)\|d_n\|^2 + \tfrac{1}{2}\tau_n^2\varpi_{1,n}\|\xi_{n+1}\|^2 \\ &= \tfrac{1}{2}\rho_n^{-1}\tau_n^2\varpi_{2,n}\|d_n\|^2 + \tfrac{1}{2}\tau_n^2\varpi_{1,n}\|\xi_{n+1}\|^2, \end{aligned} \tag{4.18}$$

where $\varpi_{1,n}$ and $\varpi_{2,n}$ are defined by (4.1) and (4.2), respectively. Clearly, inequality (4.18) in light of $e_* - \beta_n \geq 0$ (from (4.5)), $\varpi_{1,n} \geq 0$ and $\varpi_{2,n} \geq 0$ (from (4.6)) implies that the sequence $\big(E_n(e_*, q)\big)_{n\geq N}$ is non-increasing, which, by Proposition 4.2 entails that for any $n \geq N$

$$\begin{aligned} E_N(e_*, q) \geq E_n(e_*, q) \geq \tfrac{1}{2}\|e_*(q - x_n) + \nu_n(y_n - x_n)\|^2 \\ + \mu\big(e_*\eta_n\tau_n + \zeta_{1,n}(e_*)\tau_{n-1}^2\big)\bar{\Theta}(x_n) + \tfrac{1}{2}\tau_n^2\zeta_{2,n}(e_*)\|\xi_n\|^2, \end{aligned} \tag{4.19}$$

where $\zeta_{1,n}(.)$ and $\zeta_{2,n}(.)$ are defined in (4.3) and (4.4), respectively. So, by $\zeta_{1,n}(e_*) \geq 0$ and $\zeta_{2,n}(e_*) \geq 0$ (under condition (4.6)), we obviously deduce from (4.19) that the sequence

$(E_n(e_*,q))_{n\geq N}$ is nonnegative. Therefore, $(E_n(e_*,q))_{n\geq N}$, as a non-increasing and nonnegative sequence, is convergent. It is also immediately seen that the three terms in the right side of (4.19) are nonnegative, hence each of them is bounded by $E_N(e_*,q)$ which leads us to (4.7), (4.8) and (4.9) (in light of $\nu_n \leq \tau_n$). Furthermore, summing inequality (4.18) between N and $n-1 > N$ classically yields

$$\begin{array}{l} E_n(e_*,q) + (1-\lambda)\mu \sum_{k=N}^{n-1} (e_* - \beta_k)\, \tau_k \bar{\Theta}(x_k) \\ \qquad + \frac{1}{2}\sum_{k=N}^{n-1} \rho_k^{-1} \tau_k^2 \varpi_{2,k} \|d_k\|^2 + \frac{1}{2}\sum_{k=N}^{n-1} \tau_k^2 \varpi_{1,k} \|\xi_{k+1}\|^2 \leq E_N(e_*,q). \end{array} \tag{4.20}$$

Noticing that the four terms in the left side of the previous inequality are nonnegative, we deduce that each of them is bounded by $E_N(e_*,q)$ which amounts to items (4.10), (4.11) and (4.12) (in light of $\nu_n \leq \tau_n$). Item (4.13) follows straightforwardly from (4.11) and (3.18). □

4.2. **Convergence analysis.** A key result of this section is given by the next lemma.

Lemma 4.3. *Suppose that the set of conditions* **(CP)** *is satisfied. Assume furthermore that* $\{x_n, y_n, z_n\} \subset H$ *are generated by (1.5)-(1.8) with positive parameters satisfying the following conditions:*

$$(\dot{\nu}_n)_{n\geq N} \subset [\bar{\nu}_m^*, \bar{\nu}_M^*] \qquad \textit{(with constants } \bar{\nu}_m^* > 0 \textit{ and } \bar{\nu}_M^* > 0\textit{)}, \tag{4.21}$$

$$e_* > \sup_{n\geq N} \beta_n, \tag{4.22}$$

$$\inf_{n\geq N} \zeta_{i,n}(e_*) \geq \bar{\zeta}_i \textit{ (for } i = 1,2 \textit{ and for some positive constants } \bar{\zeta}_i\textit{)}, \tag{4.23}$$

$$\varpi_{1,n} \geq 0, \quad \inf_{n\geq N} \varpi_{2,n} \geq \bar{\varpi}_2 \textit{ (for some positive constants } \bar{\varpi}_2\textit{)}, \tag{4.24}$$

where $\zeta_{1,n}$, $\zeta_{2,n}$, $\varpi_{1,n}$ *and* $\varpi_{2,n}$ *are defined by (4.1)-(4.4). Then, for* $q \in S$, *the energy-like sequence* $(E_n(e_*,q))_{n\geq N}$ *(given in (3.10)) is non-increasing. In addition,* (x_n) *and* (y_n) *converge weakly to some element* $\bar{z} \in S$ *and the following properties are reached :*

$$\sum_n n\,(\Theta(x_n) - \min\Theta) < \infty, \quad \Theta(x_n) - \min\Theta = o(n^{-2}), \tag{4.25}$$

$$\sum_n n\|\dot{x}_n\|^2 < \infty, \quad \|\dot{x}_n\| = o(n^{-1}), \tag{4.26}$$

$$\sum_n n\|\dot{y}_n\|^2 < \infty, \quad \|\dot{y}_n\| = o(n^{-1}), \tag{4.27}$$

$$\exists \xi_n^* \in \partial\Theta(x_n) \quad s.t. \quad \sum_n n^2 \|\xi_n^*\|^2 < \infty. \tag{4.28}$$

4.2.1. *Proof of Lemma 4.3.* The proof can be divided into the next steps **(A1)**, **(A2)**, **(A3)** :

(A1) A useful inequality. The next result was implicitly established in [20] and its proof is given for the sake of completeness.

Proposition 4.4. *Let (1.2)-(1.4) hold (for some* $L \geq 0$*), let* $\{\kappa, e_*, \mu, \nu_n\} \subset (0,\infty)$ *and* $(\eta_n) \subset [0,1)$. *Then, for any sequences* $\{x_n, y_n, z_n\} \subset H$ *generated by (1.5)-(1.8), there exists* $\xi_n^* \in \partial\Theta(x_n)$ *such that*

$$\|\xi_{n+1}^*\|^2 \leq (1+L\mu)^2 \|M_\mu(z_n)\|^2. \tag{4.29}$$

Proof. In light of (3.16), we have $z_n - x_{n+1} = \mu\left(x_{n+1}^* + \nabla f(z_n)\right)$ (for some $x_{n+1}^* \in \partial g(x_{n+1})$), while the triangular inequality together with the L-Lipschitz continuity of ∇f (from condition (1.3)) gives us

$$\|x_{n+1}^* + \nabla f(x_{n+1})\| \leq \|x_{n+1}^* + \nabla f(z_n)\| + L\|z_n - x_{n+1}\|.$$

Combining the last two results, we obtain $\|x^*_{n+1}+\nabla f(x_{n+1})\| \leq (1+L\mu)\,\|x^*_{n+1}+\nabla f(z_n)\|$, which, in light of $M_\mu(z_n) = x^*_{n+1}+\nabla f(z_n)$ (from (3.17)), leads us to

$$\|x^*_{n+1}+\nabla f(x_{n+1})\|^2 \leq (1+L\mu)^2\,\|M_\mu(z_n)\|^2.$$

Therefore, denoting $\xi^*_n = x^*_n+\nabla f(x_n)$ $\big($hence $\xi^*_n \in \partial\Theta(x_n)\big)$ gives us (4.29). □

(A2) Asymptotic properties of the parameters. Let us make the following remark.

Remark 4.5. Some basic properties regarding the parameters ν_n, ρ_n, and θ_n are established. We begin by underlining that (4.21) obviously ensures that, as $n\to\infty$

$$\nu_n \to \infty, \quad \nu_n = O(n) \quad \text{and} \quad \nu_n^{-1} = O(n^{-1}). \tag{4.30}$$

In addition, by $\bar{\nu}^*_m \nu_n^{-1} \leq \nu_{n+1}\nu_n^{-1} - 1 \leq \bar{\nu}^*_M \nu_n^{-1}$ for $n \geq N$ $\big($according to (4.21) and the obvious equality $\dot{\nu}_{n+1}\nu_n^{-1} = \nu_{n+1}\nu_n^{-1} - 1\big)$, and recalling that $\rho_n = 1-(1-\kappa)\nu_{n+1}\nu_n^{-1}$ (from (3.1)) and that $\theta_n = \nu_n\rho_n(e_*+\nu_{n+1})^{-1}$ (from (1.8)), we easily deduce that

$$\lim_{n\to\infty} \tfrac{\nu_{n+1}}{\nu_n} = 1, \quad \lim_{n\to\infty}\rho_n = \kappa, \quad \lim_{n\to\infty}\theta_n = \kappa. \tag{4.31}$$

(A3) Proving Lemma 4.3. The proof will be divided into the following steps **(b0)** to **(b3)** :

(b0) Let us show that the following intermediate results are reached:

$$\textstyle\sum_n n^2\|M_\mu(z_n) - \eta_n M_\mu(z_{n-1})\|^2 < \infty, \tag{4.32}$$
$$\textstyle\sum_n n^2\|\dot{x}_{n+1} + \theta_n(y_n - x_n)\|^2 < \infty, \tag{4.33}$$
$$\textstyle\sum_n n\bar{\Theta}(x_n) < \infty, \tag{4.34}$$
$$\textstyle\sum_n n^2\|M_\mu(z_n)\|^2 < \infty, \tag{4.35}$$
$$\exists \xi^*_n \in \partial\Theta(x_n) \text{ s.t. } \textstyle\sum_{n\geq N} n^2\|\xi^*_n\|^2 < \infty. \tag{4.36}$$

Indeed, Lemma 4.1 gives us the set of estimates (4.10)-(4.13), which in light of $\nu_n^{-1} = O(1/n)$ (from (4.30)), $\lim_{n\to\infty}\rho_n = \kappa$ (from (4.31)), $e_* > \sup_{n\geq N}\beta_n$ (from condition (4.22)), $\varpi_{1,n} \geq 0$ and $\varpi_{2,n} \geq \bar{\varpi}_2 > 0$ (from condition (4.24)) immediately gives us (4.32)-(4.35). The last item (4.36) is a straightforward consequence of (4.35) and Proposition 4.4.

(b1) Now, we denote $u_n = y_n - x_n$ and $\bar{\beta} = \sup_{j\geq N}\beta_j$ $\big($hence $\bar{\beta} < e_*$ (from (4.22))$\big)$, and we show under assumptions (4.21) that the properties below hold:

$$(E_n(s,q))_n \text{ is convergent, for any } (s,q) \in [\bar{\beta}, e_*)\times S, \tag{4.37}$$
$$\textstyle\sum_n n\|\dot{x}_n\|^2 < \infty, \tag{4.38}$$
$$\textstyle\sum_n n\|u_n\|^2 < \infty, \tag{4.39}$$
$$\lim_{n\to\infty} n^2\big(\Theta(x_n) - \min\Theta\big) = \lim_{n\to\infty} n\|u_n\| = \lim_{n\to\infty} n\|M_\mu(z_{n-1})\| = 0. \tag{4.40}$$

Indeed, taking $(s,q) \in (0,e_*]\times S$, and using Lemma 3.2 in light of (4.21) together with (3.18), we deduce that, for $n \geq N$ $\big(N$ arising in the set of conditions (3.2)-(3.6)$\big)$

$$\begin{aligned}
&\dot{E}_{n+1}(s,q) + (1-\lambda)\mu\,(s-\beta_n)\,\tau_n\bar{\Theta}(x_n) \\
&\qquad + \tfrac{1}{2}\tau_n^2\varpi_{1,n}\|\mu M_\mu(z_n)\|^2 + \tfrac{1}{2}(1-\delta)\rho_n^{-1}\tau_n^2\|\dot{x}_{n+1}+\theta_n u_n\|^2 \\
&\leq -T_n(s) \leq \tfrac{1}{2}(\kappa-1)_+\rho_n\nu_n\nu_{n+1}\|u_n + \Big(\tfrac{1}{\theta_n} - \tfrac{e_*-s}{\nu_n\rho_n}\Big)\dot{x}_{n+1}\|^2 \\
&\qquad\qquad - \tfrac{1}{2}(e_*-s)\big((e_*+2\nu_{n+1})\rho_n^{-1} + s(\rho_n^{-1}-1)\big)\,\|\dot{x}_{n+1}\|^2,
\end{aligned} \tag{4.41}$$

where $\bar{\Theta} := \Theta - \min\Theta$, $T_n(s)$ is given by (3.9) and $\varpi_{1,n}$ is defined in (4.1). Let us estimate the first term in the right side of (4.41). Recalling that $\theta_n = \frac{\nu_n\rho_n}{\tau_n}$ $\big($hence $\frac{\rho_n\nu_n}{\theta_n^2} = \rho_n^{-1}\nu_n^{-1}\tau_n^2\big)$, and using the Young inequality, we classically obtain

$$\begin{aligned}\rho_n\nu_n\nu_{n+1}\|u_n + \Big(\tfrac{1}{\theta_n} - \tfrac{e_*-s}{\nu_n\rho_n}\Big)\dot{x}_{n+1}\|^2 &= \big(\tfrac{\rho_n\nu_n}{\theta_n^2}\big)\nu_{n+1}\|\theta_n u_n + \dot{x}_{n+1} - \theta_n\tfrac{e_*-s}{\nu_n\rho_n}\dot{x}_{n+1}\|^2\\ &\le 2\big(\tfrac{\rho_n\nu_n}{\theta_n^2}\big)\nu_{n+1}\|\dot{x}_{n+1}+\theta_n u_n\|^2 + 2(e_*-s)^2\rho_n^{-1}\tfrac{\nu_{n+1}}{\nu_n}\|\dot{x}_{n+1}\|^2\\ &= 2\tfrac{\nu_{n+1}}{\nu_n}\rho_n^{-1}\tau_n^2\|\dot{x}_{n+1}+\theta_n u_n\|^2 + 2(e_*-s)^2\rho_n^{-1}\tfrac{\nu_{n+1}}{\nu_n}\|\dot{x}_{n+1}\|^2.\end{aligned}$$

Hence, inequality (4.41) entails that

$$\begin{aligned}&\dot{E}_{n+1}(s,q) + (1-\lambda)\mu\,(s-\beta_n)\,\tau_n\bar{\Theta}(x_n) + \tfrac{1}{2}\tau_n^2\varpi_{1,n}\|\mu M_\mu(z_n)\|^2\\ &\quad\le (\kappa-1)_+\tfrac{\nu_{n+1}}{\nu_n}\rho_n^{-1}\tau_n^2\|\dot{x}_{n+1}+\theta_n u_n\|^2\\ &\qquad -\tfrac{1}{2}(e_*-s)\Big((e_*+2\nu_{n+1})\rho_n^{-1} + s(\rho_n^{-1}-1) - 2(\kappa-1)_+(e_*-s)\rho_n^{-1}\tfrac{\nu_{n+1}}{\nu_n}\Big)\|\dot{x}_{n+1}\|^2\\ &\quad= l_n\|\dot{x}_{n+1}+\theta_n u_n\|^2 - \tfrac{1}{2}(e_*-s)h_n(s)\|\dot{x}_{n+1}\|^2,\end{aligned} \tag{4.42}$$

where l_n and h_n $\big($noticing that $\rho_n^{-1} = 1 + \frac{(1-\kappa)\nu_{n+1}}{\nu_n\rho_n}$ $\big)$ are defined by

$$l_n = (\kappa-1)_+\tfrac{\nu_{n+1}}{\nu_n}\rho_n^{-1}\tau_n^2, \tag{4.43}$$

$$h_n(s) = (e_*+2\nu_{n+1})\Big(1 + \tfrac{(1-\kappa)\nu_{n+1}}{\nu_n\rho_n}\Big) + s\tfrac{(1-\kappa)\nu_{n+1}}{\nu_n\rho_n} - 2(e_*-s)(\kappa-1)_+\rho_n^{-1}\tfrac{\nu_{n+1}}{\nu_n}. \tag{4.44}$$

Clearly, by the above definitions of l_n and $h_n(.)$, along with $\nu_n\to\infty$, $\frac{\nu_{n+1}}{\nu_n}\to 1$ (from (4.31)), $\rho_n\to\kappa$ and $\nu_n = O(n)$ (from (4.30)), we easily deduce that for $s\in(0,e_*]$ (as $n\to\infty$)

$$h_n(s) \sim \tfrac{2}{\kappa}\nu_{n+1} \quad \text{(hence } h_n(s) = O(n)\text{)}, \tag{4.45}$$

$$l_n = O(n^2). \tag{4.46}$$

Consider now $\bar{\beta} = \sup_{j\ge N}\beta_j$ (introduced in (4.37)). Since $\bar{\beta} < e_*$ (by (4.22)), we can choose $s_0\in[\bar{\beta},e_*)$. It is readily seen by $\sum_n n^2\|\dot{x}_{n+1}+\theta_n u_n\|^2<\infty$ (from (4.33)) along with $l_n = O(n^2)$ (from (4.46)) that

$$\textstyle\sum_n l_n\|\dot{x}_{n+1}+\theta_n u_n\|^2 < \infty. \tag{4.47}$$

Furthermore, (4.45) implies that for n large enough, $h_n(s)>0$, which by (4.42) yields

$$\dot{E}_{n+1}(s,q) + (1-\lambda)\mu\,(s-\beta_n)\,\tau_n\bar{\Theta}(x_n) + \frac{1}{2}\tau_n^2\varpi_{1,n}\|\mu M_\mu(z_n)\|^2 \le l_n\|\dot{x}_{n+1}+\theta_n u_n\|^2. \tag{4.48}$$

Therefore, using inequality (4.48) with $s = s_0$ together with (4.47), while recalling that $s_0 - \bar{\beta}\ge 0$ and $E_{n+1}(s_0,q)\ge 0$ $\big($from Proposition 4.2 in light of $\zeta_{1,n}(e_*)\ge 0$ and $\zeta_{2,n}(e_*)\ge 0$ under condition (4.6)$\big)$, classically yields that $(E_n(s_0,q))_n$ is convergent, which proves (4.37). Furthermore, summing (4.42) between N and $n-1>N$, in light of $E_{n+1}(s_0,q)\ge 0$ yields

$$\tfrac{1}{2}(e_*-s_0)\textstyle\sum_{k=N}^n h_k(s_0)\|\dot{x}_{k+1}\|^2 \le E_N(s_0,q) + \sum_{k=N}^n l_k\|\dot{x}_{k+1}+\theta_k u_k\|^2. \tag{4.49}$$

Then, using (4.47) and $h_n(s) = O(n)$ (from (4.45)) leads us to $\sum_n n\|\dot{x}_n\|^2<\infty$, namely (4.38). To prove (4.39), we notice that using the Young inequality entails

$$n\|\theta_n u_n\|^2 = n\|\theta_n u_n + \dot{x}_{n+1} - \dot{x}_{n+1}\|^2 \le 2n\|\theta_n u_n + \dot{x}_{n+1}\|^2 + 2n\|\dot{x}_{n+1}\|^2. \tag{4.50}$$

Thus, by $\sum_n n\|\theta_n u_n+\dot{x}_{n+1}\|^2<\infty$ (from (4.33)) and $\sum_n n\|\dot{x}_{n+1}\|^2<\infty$ (from (4.38)), we get $\sum_n n\|\theta_n u_n\|^2<\infty$. Recalling that $\lim_{n\to\infty}\theta_n = \kappa$ (from (4.31)) we deduce that

$$\textstyle\sum_n n\|u_n\|^2<\infty \quad \text{(that is (4.39))}. \tag{4.51}$$

It remains to prove (4.40). From the definition of $E_n(s,q)$ given by (3.10) with $s=0$, we have for $n \geq N$

$$\begin{aligned} E_n(0,q) = \tfrac{1}{2}\|\nu_n u_n\|^2 + \mu\tau_{n-1}^2\rho_{n-1}^{-1}(1-\eta_{n-1})\overline{\Theta}(x_n) \\ +\tfrac{1}{2}\eta_n\rho_n^{-1}\tau_n^2\left(\delta(1-\eta_n)+\mu L\right)\|\mu M_\mu(z_{n-1})\|^2. \end{aligned} \tag{4.52}$$

Moreover, passing to the limit as $s \to 0^+$ in inequality (4.42) yields

$$E_{n+1}(0,q) - E_n(0,q) \leq l_n\|\dot{x}_{n+1}+\theta_n u_n\|^2 + (1-\lambda)\mu\bar{\beta}\tau_n\bar{\Theta}(x_n) - \frac{1}{2}e_* h_n(0)\|\dot{x}_{n+1}\|^2. \tag{4.53}$$

This, by $h_n(0) \sim \frac{2}{\kappa}\nu_{n+1}$ (as $n\to\infty$) (from 4.45)), $\sum_n \tau_n\bar{\Theta}(x_n) < \infty$ (from (4.34)) and $\tau_n = O(n)$ (from (4.30)), along with $\sum_n l_n\|\dot{x}_{n+1}+\theta_n u_n\|^2 < \infty$ (from (4.47)), classically entails that

$$(E_n(0,q)) \text{ is convergent as } n \to \infty. \tag{4.54}$$

Next, multiplying equality (4.52) by n^{-1} yields

$$\begin{aligned} n^{-1}E_n(0,q) = \tfrac{1}{2}n^{-1}\|\nu_n u_n\|^2 + \mu n^{-1}\tau_{n-1}^2\rho_{n-1}^{-1}(1-\eta_{n-1})\overline{\Theta}(x_n) \\ +\tfrac{1}{2}n^{-1}\eta_n\rho_n^{-1}\tau_n^2\left(\delta(1-\eta_n)+\mu L\right)\|\mu M_\mu(z_{n-1})\|^2. \end{aligned} \tag{4.55}$$

We will now prove that $\sum_n n^{-1}E_n(0,q) < \infty$ by estimating each term in the right side of the above inequality. From $\nu_n = O(n)$ and $\sum_n n\|u_n\|^2 < \infty$ (from (4.39)), we get

$$\sum_n n^{-1}\|\nu_n u_n\|^2 = \sum_n n^{-1}\nu_n^2\|u_n\|^2 < \infty. \tag{4.56}$$

Furthermore, by $\tau_n = O(n)$, $\eta_n \in [0,1)$ and $\rho_n \sim \kappa$ as $n\to\infty$ (from (4.31)) we deduce that $n^{-1}\tau_{n-1}^2\rho_{n-1}^{-1}(1-\eta_{n-1}) = O(n)$, which together with $\sum_n n\bar{\Theta}(x_n) < \infty$ (from (4.34)) ensures that

$$\sum_n n^{-1}\tau_{n-1}^2\rho_{n-1}^{-1}(1-\eta_{n-1})\bar{\Theta}(x_n) < \infty. \tag{4.57}$$

In addition, again using (4.31) entails that $n^{-1}\rho_n^{-1}\tau_n^2 = O(n)$, which by $\sum_n n^2\|M_\mu(z_n)\|^2 < \infty$ (from (4.35)) readily yields

$$\sum_n n^{-1}\eta_n\rho_n^{-1}\tau_n^2\big(\delta(1-\eta_n)+\mu L\big)\|M_\mu(z_{n-1})\|^2 < \infty. \tag{4.58}$$

Combining inequality (4.55) with the previous results (4.56), (4.57) and (4.58) yields that $\sum_n n^{-1}E_n(0,q) < \infty$. As a consequence, recalling that $E_n(0,q)$ is convergent (as $n\to\infty$) (from (4.54)) and noticing that $\sum_n n^{-1} = \infty$, we deduce that $\lim_{n\to\infty} E_n(0,q) = 0$. Now, using Proposition 4.2 with $s=0$ entails

$$E_n(0,q) \geq \frac{1}{2}\|\nu_n u_n\|^2 + \mu\,\zeta_{1,n}(e_*)\tau_{n-1}^2\bar{\Theta}(x_n) + \frac{1}{2}\eta_n\zeta_{2,n}(e_*)\tau_n^2\|\xi_n\|^2. \tag{4.59}$$

Note that the three terms in the right side of the previous inequality are nonnegative. Therefore, each of them converges to zero (as $n\to\infty$) (since $\lim_{n\to\infty} E_n(0,q) = 0$). This, combined with $\tau_n = O(n)$ and $\rho_n \to \kappa$ (as $n\to\infty$) (from (4.30) and (4.31)) together with the conditions $\inf_n \zeta_{i,n} \geq \bar{\zeta}_i > 0$ (for $i=1,2$) (from (4.23)), immediately leads us to (4.40).

(b2) Now we are in position to prove the estimates of Lemma 4.3. The first parts of items (4.25) and (4.26) are nothing but (4.34) and (4.38), respectively, while (4.28) is exactly (4.36). Furthermore, inequality (4.39) along with $\dot{y}_{n+1} = -\kappa u_n$ (according to (1.7)) leads us to $\sum_n n\|\dot{y}_n\|^2 < \infty$ (namely the first estimate of item (4.27)). Moreover, (4.40) gives us $\lim_{n\to\infty} n^2\bar{\Theta}(x_n) = 0$ (that is, the second estimate of (4.25)). We also have $\lim_{n\to\infty} n\|u_n\| = 0$ (again from (4.40)), which together with $\dot{y}_{n+1} = -\kappa u_n$ (from (1.7)) yields $\|\dot{y}_n\| = o(n^{-1})$ (that

is the second estimate of (4.27)). Finally, we recall that, as $n \to \infty$, $n\|\dot{x}_{n+1} + \theta_n u_n\| \to 0$ (according to (4.33)), while $n\|u_n\| \to 0$ (from (4.40)). It readily follows from the boundedness of (θ_n) that $n\|\dot{x}_{n+1}\| \to 0$ (as $n \to \infty$), which proves the second estimate of item (4.26).

(b3) It remains to prove that any sequences (x_n) and (y_n) generated by (1.5)-(1.8) converge weakly (as $n \to \infty$) to some element of S. This can be shown by means of the classical Opial's lemma [26]. From the definition of $(E_n(s,q))_{n\geq 0}$ given in (3.10) we have

$$\begin{aligned} E_n(s,q) = {} & \tfrac{1}{2}\|s(q-x_n) + \nu_n(y_n - x_n)\|^2 + \tfrac{1}{2}s(e_* - s)\|x_n - q\|^2 \\ & + s\mu\eta_n\tau_n\langle M_\mu(z_{n-1}), x_n - q\rangle + \mu\tau_{n-1}^2\big(\rho_{n-1}^{-1}(1-\eta_{n-1}) - \tfrac{s\lambda}{\tau_{n-1}}\big)\bar{\Theta}(x_n) \\ & + \tfrac{1}{2}\eta_n\rho_n^{-1}\tau_n^2\big(\delta(1-\eta_n) + \mu L\big)\|\mu M_\mu(z_{n-1})\|^2, \end{aligned} \tag{4.60}$$

where $\bar{\Theta} := \Theta - \min\Theta$. Let us estimate separately the third, fourth and fifth terms in the right side of inequality (4.60). To that end, we first note that, by (4.40) along with $\nu_n = O(n)$, $\tau_n = O(n)$, $\eta_n \in [0,1)$ and $\rho_n \sim \kappa$ (as $n \to \infty$) (from (4.31)), we have

$$\lim_{n\to\infty} \nu_n\|y_n - x_n\| = 0, \tag{4.61}$$

$$\lim_{n\to\infty} \tau_n^2\rho_{n-1}^{-1}\bar{\Theta}(x_n) = \lim_{n\to\infty} \tau_n^2\|M_\mu(z_{n-1})\|^2 = 0. \tag{4.62}$$

By (4.7) we also know that $e_*(q - x_n) + \nu_n(y_n - x_n)$ is bounded, which in light of (4.61) entails that (x_n) is also bounded. Concerning the third term in the right side of (4.60), the Cauchy-Schwarz inequality yields

$\eta_n\tau_n|\langle M_\mu(z_{n-1}), x_n - q\rangle| \leq \eta_n\tau_n\|M_\mu(z_{n-1})\|\|x_n - q\|$.

So, by $\lim_{n\to\infty}\tau_n\|M_\mu(z_{n-1})\| = 0$ (from (4.62)) and boundedness of (x_n) and (η_n), we get

$\eta_n\tau_n\langle M_\mu(z_{n-1}), x_n - q\rangle \to 0$ as $n \to \infty$.

Concerning the fourth and fifth terms in the right side of (4.60), we observe that the quantities $\rho_{n-1}^{-1}(1-\eta_{n-1}) - \frac{s\lambda}{\tau_{n-1}}$ and $\eta_n\rho_n^{-1}\big(\delta(1-\eta_n) + \mu L\big)$ are bounded. Hence, by (4.62), we readily see that the last two terms in the right side of (4.60) tend to 0 (as $n \to \infty$). We now take $s \in [\bar{\beta}, e_*)$. By (4.37) we know that $(E_n(s,q))$ is convergent as $n \to \infty$, Therefore, passing to the limit in (4.60) as $n \to \infty$, in light of the previous arguments, immediately yields that

$\lim_{n\to\infty} E_n(s,q) = \Big(\frac{s^2}{2} + \frac{1}{2}s(e_* - s)\Big)\lim_{n\to\infty}\|x_n - q\|^2$.

Thus we deduce that $\|x_n - q\|$ is convergent. Furthermore, it can be also noticed that (4.25) ensures that $\bar{\Theta}(x_n) \to 0$ as $n \to \infty$. So, given any subsequence (x_{n_k}) of (x_n) such that (x_{n_k}) converges weakly to some $x_* \in H$, by invoking the weak lower semi-continuity of Θ we obtain

$$0 = \liminf_{k\to\infty}(\Theta(x_{n_k}) - \Theta(q)) \geq (\Theta(x_*) - \Theta(q)) \geq 0. \tag{4.63}$$

This entails that $\Theta(x_*) = \Theta(q)$, namely $x_* \in S$. Therefore, any weak-cluster point of (x_n) belongs to S. Applying Opial's lemma [26] we conclude that the whole sequence (x_n) converges weakly to some element $\bar{z} \in S$. So does (y_n) in light of (4.61). This ends the proof. □

5. Main Convergence Results and Estimates

5.1. **The general case of correction parameters.** We are now in position to claim the main result of this section.

Theorem 5.1. *Let (1.2)-(1.4) hold (for some $L \geq 0$), let $\{\kappa, e_*, \mu, \nu_n\} \subset (0,\infty)$ be such that $\kappa \in (0,2)$ and $(\eta_n) \subset [0,1)$ satisfy (3.2)-(3.7) for some $(N,\lambda) \in \mathbb{N} \times [0,1)$ and $(\beta_n) \subset [0,\infty)$. Assume furthermore that $\{x_n, y_n, z_n\} \subset H$ are generated by (1.5)-(1.8) with parameters satisfying the following conditions:*

$$(\dot{\nu}_n)_{n\geq N} \subset [\bar{\nu}_m^*, \bar{\nu}_M^*] \qquad \text{(with constants } \bar{\nu}_m^* > 0 \text{ and } \bar{\nu}_M^* > 0\text{)}, \tag{5.1}$$

$$e_* > \sup_{n\geq N} \beta_n, \tag{5.2}$$

$$\mu L < \frac{(1-\varepsilon)^2}{1+\varepsilon} \min\{1, 2-\kappa\}, \tag{5.3}$$

Then, (x_n) and (y_n) converge weakly to some element $\bar{z} \in S$ and the following properties are reached:

$$\sum_n n\,(\Theta(x_n) - \min\Theta) < \infty, \quad \Theta(x_n) - \min\Theta = o(n^{-2}), \tag{5.4}$$

$$\sum_n n\|\dot{x}_n\|^2 < \infty,\ \|\dot{x}_n\| = o(n^{-1}), \tag{5.5}$$

$$\sum_n n\|\dot{y}_n\|^2 < \infty,\ \|\dot{y}_n\| = o(n^{-1}), \tag{5.6}$$

$$\exists \xi_n^* \in \partial\Theta(x_n) \quad s.t. \quad \sum_n n^2\|\xi_n^*\|^2 < \infty. \tag{5.7}$$

Proof. Indeed, observe that condition (5.3) can be reformulated as $\mu L \frac{1+\varepsilon}{(1-\varepsilon)^2} < \min\{1, 2-\kappa\}$. Hence, we can choose $\delta \in (0,1) \cap (0, 2-\kappa)$ such that

$$\mu L \tfrac{1+\varepsilon}{(1-\varepsilon)^2} < \delta < \min\{1, 2-\kappa\}. \tag{5.8}$$

Let us now consider the parameters introduced in (4.1) (involved in Lemma 4.1) along with this particular value of δ. For convenience of the reader, we recall that these parameters write

$$\varpi_{1,n} = \rho_n^{-1}\big(\delta(1-\eta_n) - \mu L\big) - \eta_{n+1}\rho_{n+1}^{-1}\big(\tfrac{\tau_{n+1}}{\tau_n}\big)^2\big(\delta(1-\eta_{n+1}) + \mu L\big), \tag{5.9}$$

$$\varpi_{2,n} = \bar{\delta} + (1-\kappa)\big(1 + \tfrac{\dot{\nu}_{n+1}}{\nu_n}\big), \tag{5.10}$$

$$\zeta_{1,n}(e_*) = \rho_{n-1}^{-1}(1-\eta_{n-1}) - \tfrac{e_*\lambda}{\tau_{n-1}}, \qquad \zeta_{2,n}(e_*) = \rho_n^{-1}\big(\delta(1-\eta_n) + \mu L\big) - e_*\mu L\tau_n^{-1}. \tag{5.11}$$

Passing to the limit as $n \to \infty$ in the above equalities, in light of (4.30) and (4.31), readily yields

$$\lim_{n\to\infty} \varpi_{1,n} = \kappa^{-1}\left(\delta(1-\varepsilon) - \mu L - \varepsilon\big(\delta(1-\varepsilon) + \mu L\big)\right) = \kappa^{-1}\big(\delta(1-\varepsilon)^2 - (1+\varepsilon)\mu L\big), \tag{5.12}$$

$$\lim_{n\to\infty} \varpi_{2,n} = 1 - \delta + 1 - \kappa = 2 - \delta - \kappa, \tag{5.13}$$

$$\lim_{n\to\infty} \zeta_{1,n}(e_*) = \kappa^{-1}(1-\varepsilon), \qquad \lim_{n\to\infty} \zeta_{2,n}(e_*) = \kappa^{-1}\big(\delta(1-\varepsilon) + \mu L\big). \tag{5.14}$$

It is obviously checked from (5.8) that the above limits are positive, which readily implies the existence of an integer N_* such that $n > N_*$ yields the following results

$$\inf_{n\geq N_*} \zeta_{i,n}(e_*) \geq \bar{\zeta}_i \text{ (for } i = 1,2 \text{ and for some positive constants } \bar{\zeta}_i), \tag{5.15}$$

$$\varpi_{1,n} \geq 0, \quad \inf_{n\geq N_*} \varpi_{2,n} \geq \bar{\varpi}_2 \text{ (for some positive constants } \bar{\varpi}_2). \tag{5.16}$$

Consequently, the desired conclusion immediately follows from Lemma 4.3. □

5.2. **Convergence results in particular cases of the correction parameter.** In this section, we discuss through the following corollary the convergence of (1.5)-(1.8) relative to special instances of parameters η_n and ν_n.

Corollary 5.2. *Let (1.2)-(1.4) hold (for some $L \geq 0$), let $\{\kappa, e_*, \nu_0, \mu\}$ be positive parameters such that $\kappa \in (0,2)$, and let $s_1 \in [0, e_*)$ and $\varepsilon \in [0,1)$. Assume in addition that*

$$e_* > 2\kappa^{-1}. \tag{5.17}$$

Let $\{x_n, y_n, z_n\} \subset H$ be generated by (1.5)-(1.8) with parameters (ν_n) and (η_n) defined by

$$\nu_n := n + \nu_0, \tag{5.18}$$

$$\eta_n := \varepsilon \left(1 - \tfrac{s_1}{e_* + \nu_{n+1}} \left(1 - (1-\kappa)\tfrac{\nu_{n+1}}{\nu_n}\right)\right). \tag{5.19}$$

Suppose furthermore that the following condition holds :

$$\mu L < \frac{(1-\varepsilon)^2}{1+\varepsilon} \min\{1, 2-\kappa\}. \tag{5.20}$$

Then, (x_n) and (y_n) converge weakly to some element $\bar{z} \in S$ and the properties (5.4)-(5.7) are reached.

Remark 5.3. Note that the above corollary covers the special case when η_n is constant (by setting $s_1 = 0$).

Proof. The results are derived from Theorem 5.1. Indeed, let us prove that the assumptions of this theorem hold. It is obvious from the definition of (ν_n) and (η_n) given in (5.18)-(5.19) that the conditions (3.2) and (3.3) are verified and that we have

$$\lim_{n\to\infty} \eta_n = \varepsilon \quad \text{(that is (3.7))}. \tag{5.21}$$

Concerning now conditions (3.4)-(3.6), we set for simplification $\rho_n = 1 - (1-\kappa)\frac{\nu_{n+1}}{\nu_n}$ and $\tau_n = e_* + \nu_{n+1}$, and we introduce the sequence (β_n) defined for $n \geq 0$ by

$$\beta_n = \frac{1}{1-\varepsilon}\left((1-\eta_{n-1})\left(\frac{2}{\rho_n} + \frac{(\kappa-1)_+}{\rho_n\rho_{n-1}\nu_n\nu_{n-1}}\tau_n\right) + \frac{\varepsilon s_1 \rho_n^{-1}}{\tau_{n-1}}\left(\rho_n + (\kappa-1)\frac{\tau_n}{\nu_n\nu_{n-1}}\right)\right). \tag{5.22}$$

We easily deduce from (4.30), (4.31) and (5.21) that

$$\lim_{n\to\infty} \left(\eta_n + e_*\rho_n\tau_n^{-1}\right) = \varepsilon < 1, \tag{5.23}$$

which guarantees that condition (3.4) is verified. Furthermore, by (5.18) and (5.21) we readily infer that $\lim_{n\to\infty} \eta_{n+1}\frac{\tau_{n+1}}{\tau_n} = \varepsilon$, which ensures that condition (3.5) holds for some $\lambda \in (\varepsilon, 1)$.

Let us now prove that (3.6) is satisfied for the sequence (β_n) defined by (5.22). We have

$$[\rho_n^{-1}\tau_n^2(1-\eta_n)]_\circ = (1-\eta_{n-1})[\rho_n^{-1}\tau_n^2]_\circ + \rho_n^{-1}\tau_n^2\dot{\eta}_n. \tag{5.24}$$

By an easy computation while noticing that $\tau_{n-1} = \tau_n - 1$ we get

$$\begin{aligned}[\rho_n^{-1}\tau_n^2]_\circ &= \tfrac{\tau_n^2}{\rho_n} - \tfrac{\tau_{n-1}^2}{\rho_{n-1}} = \tfrac{1}{\rho_n}\left(\tau_n^2 - \tau_{n-1}^2\right) + \left(\tfrac{1}{\rho_n} - \tfrac{1}{\rho_{n-1}}\right)\tau_{n-1}^2 \\ &= \tfrac{1}{\rho_n}(2\tau_n - 1) + \left(\tfrac{\rho_{n-1}-\rho_n}{\rho_n\rho_{n-1}}\right)\tau_{n-1}^2 \leq \tfrac{2}{\rho_n}\tau_n + \left(\tfrac{\rho_{n-1}-\rho_n}{\rho_n\rho_{n-1}}\right)\tau_{n-1}^2.\end{aligned} \tag{5.25}$$

In addition, we readily have

$$\frac{\rho_{n-1}-\rho_n}{\rho_n\rho_{n-1}} = \frac{(\kappa-1)}{\rho_n\rho_{n-1}}\left(\frac{\nu_n}{\nu_{n-1}} - \frac{\nu_{n+1}}{\nu_n}\right) = \frac{\kappa-1}{\rho_n\rho_{n-1}\nu_n\nu_{n-1}} \leq \frac{(\kappa-1)_+}{\rho_n\rho_{n-1}\nu_n\nu_{n-1}} \tag{5.26}$$

Hence, by combining (5.25) and (5.26) in light of $\tau_n \geq \tau_{n-1}$ (from (3.2)) , we obtain

$$[\rho_n^{-1}\tau_n^2]_\circ \leq \left(\frac{2}{\rho_n} + \frac{(\kappa-1)_+}{\rho_n\rho_{n-1}\nu_n\nu_{n-1}}\tau_n\right)\tau_n. \tag{5.27}$$

Furthermore, we can see that η_n can be rewritten as $\eta_n = \varepsilon\left(1-\frac{s_1}{\tau_n}\rho_n\right)$. Therefore, as $\nu_{n+1} = \nu_n + 1$, $\nu_{n-1} = \nu_n - 1$ and $\tau_{n+1} = \tau_n + 1$ we have

$$\begin{aligned}
\dot{\eta}_n &= -\varepsilon s_1\left(\tfrac{\rho_n}{\tau_n} - \tfrac{\rho_{n-1}}{\tau_{n-1}}\right) = -\tfrac{\varepsilon s_1}{\tau_n\tau_{n-1}}\left(\rho_n(\tau_n-1) - \rho_{n-1}\tau_n\right) \\
&= -\tfrac{\varepsilon s_1}{\tau_n\tau_{n-1}}\left(\tau_n(\rho_n-\rho_{n-1}) - \rho_n\right) = -\tfrac{\varepsilon s_1}{\tau_n\tau_{n-1}}\left((1-\kappa)\tau_n\left(\tfrac{\nu_n}{\nu_{n-1}} - \tfrac{\nu_{n+1}}{\nu_n}\right) - \rho_n\right) \\
&= -\tfrac{\varepsilon s_1}{\tau_n\tau_{n-1}}\left((1-\kappa)\tau_n\tfrac{\nu_n^2-(\nu_n+1)(\nu_n-1)}{\nu_n\nu_{n-1}} - \rho_n\right) = \tfrac{\varepsilon s_1}{\tau_n\tau_{n-1}}\left(\rho_n + (\kappa-1)\tfrac{\tau_n}{\nu_n\nu_{n-1}}\right).
\end{aligned} \tag{5.28}$$

Equality (5.24) together with (5.27) and (5.28) entails

$$\begin{aligned}
[\rho_n^{-1}\tau_n^2(1-\eta_n)]_\circ &= [\rho_n^{-1}\tau_n^2]_\circ(1-\eta_{n-1}) + \rho_n^{-1}\tau_n^2\dot{\eta}_n \\
&\leq \left(\tfrac{2}{\rho_n} + \tfrac{(\kappa-1)_+}{\rho_n\rho_{n-1}\nu_n\nu_{n-1}}\tau_n\right)\tau_n(1-\eta_{n-1}) + \tfrac{\varepsilon s_1\rho_n^{-1}\tau_n}{\tau_{n-1}}\left(\rho_n + (\kappa-1)\tfrac{\tau_n}{\nu_n\nu_{n-1}}\right). \\
&= (1-\varepsilon)\tfrac{1}{1-\varepsilon}\left(\left(\tfrac{2}{\rho_n} + \tfrac{(\kappa-1)_+}{\rho_n\rho_{n-1}\nu_n\nu_{n-1}}\tau_n\right)(1-\eta_{n-1}) + \tfrac{\varepsilon s_1\rho_n^{-1}}{\tau_{n-1}}\left(\rho_n + (\kappa-1)\tfrac{\tau_n}{\nu_n\nu_{n-1}}\right)\right)\tau_n \\
&= (1-\varepsilon)\beta_n\tau_n.
\end{aligned} \tag{5.29}$$

Furthermore, an easy computation using the definition of β_n (given in (5.22)) together with (4.30), (4.31) and (5.21) gives us $\beta_n \to 2\kappa^{-1}$ as $n \to \infty$, which guarantees that for n large enough, $\beta_n \geq 0$. Inequality (5.29) along with the previous observation ensures that condition (3.6) is satisfied.

Let us now prove that the set of conditions (5.1)-(5.3) holds. Condition (5.1) is a straightforward consequence of $\nu_n = n + \nu_0$. Finally, by $\beta_n \to 2\kappa^{-1}$ as $n \to \infty$, along with $e_* > 2\kappa^{-1}$ (from assumption (5.17)), we infer that (5.2) is satisfied. Therefore, the desired result immediately follows from Theorem 5.1. □

6. Numerical Experiments

In this section, we conduct some numerical experiments by implementing algorithm (1.5)-(1.8) so as to solve different problems in various Banach spaces. Our aim is to illustrate the behaviour of the proposed method as well as the influence of some parameters on its performance. We also intend to compare the converging rates with that of recently studied methods. In the sequel we denote by $\|.\|_1$ and $\|.\|_2$ the L_1-norm and the L_2-norm of $\mathbb{R}^d$, respectively.

6.1. **Application to a machine learning problem.** We start by addressing a classical machine learning problem, which consists in performing a binary classification of a diabetes dataset (available at http://networkrepository.com/pima-indians-diabetes.php), by solving the following minimization problem:

$$\min_{\mathbb{R}^d}\{\Theta(x) := f(x) + g(x)\}, \tag{6.1}$$

where $g(x) = \lambda\|x\|_1$ (for some $\lambda > 0$) and $f(x) = \frac{1}{N}\sum_{i=1}^N f_i(x)$, for some given integer N and functions $f_i(x) = \log\left(1 + e^{-b_i\langle x, a_i\rangle}\right)$ with training label data $(a_i, b_i) \in \mathbb{R}^d \times \{+1, -1\}$, $i = 1, ..., N$. This problem has been numerically investigated with the algorithm RIGPA in [20]. It can be checked that a Lipschitz constant of ∇f_i is given by $L_i = \|a_i\|^2/4$. So a Lipschitz constant for the function f is $\overline{L} = \frac{1}{N}\sum_{i=1}^N L_i$. For computational reasons, the data vectors a_i were

normalized, so that $\overline{L} = 1/4$. The other parameters are $d = 8$, $N = 768$ and $\lambda = 1/768$ while s_1 (which has to be in $[0, e_*)$ for convergence of the method) is chosen as $\frac{e_*}{2}$. The starting points are $x_0 = y_0 = (0,0,0,0,0,0,0,0)$. On all the figures, $\Theta(x_n) - \min\Theta$ is set to 10^{-13} whenever it is less than 10^{-13}.

6.1.1. *Influence of parameters κ and e_*.* Figure 1 illustrates the influence of κ on the trajectories for various fixed values of e_*. It can be seen that increasing the value of κ considerably improves the convergence, while higher values of e_* seem to reduce the oscillations visible on the trajectories.

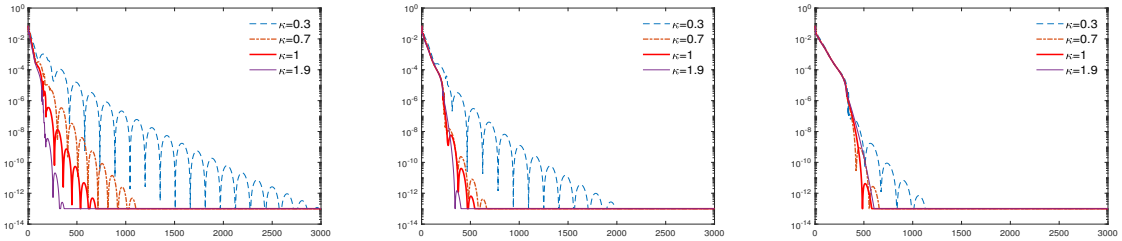


FIGURE 1. Evolution of $\Theta(x_n) - \min\Theta$ with $\varepsilon = 0.014$ for several values of κ for CRIPGA, for $e_* = 7$ (left), $e_* = 15$ (center) and $e_* = 30$ (right). The other parameters are $\mu = 0.35$, $\nu_n = n + \nu_0$ with $\nu_0 = 144$ and $s_1 = \frac{e_*}{2}$.

6.1.2. *Influence of parameter ε - comparison with RIGPA.* Figure 2 illustrates the influence of parameter ε on the trajectories, for various fixed values of e_*. It appears that the correction term ε has a damping effect on the oscillations generated by CRIPGA, especially for low values of e_*. Increasing e_* leads to a significant reduction of these oscillations, therefore reducing the incidence of the correction term. Finally, the best results in terms of convergence rate are obtained for $e_* = 5$ and ε between 0.01 and 0.015 (it is worthwhile remembering that for $\varepsilon = 0$, CRIPGA reduces to RIGPA which has been studied in [20]).

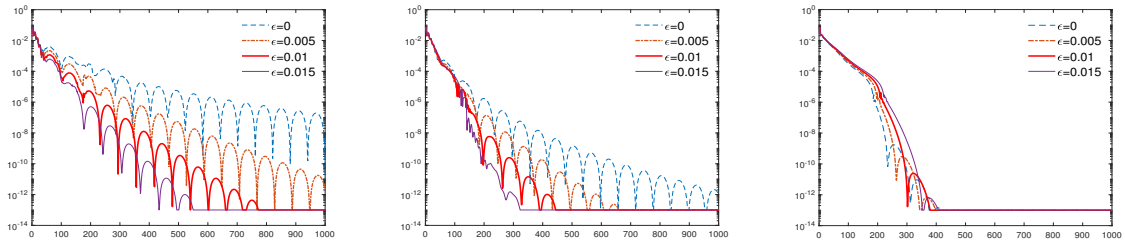


FIGURE 2. Evolution of $\Theta(x_n) - \min\Theta$ for CRIPGA for $\kappa = 1.9$ and various values of parameter ε, with $e_* = 2$ (left), $e_* = 5$ (center) and $e_* = 7$ (right). The other parameters are unchanged from Figure 1.

6.2. **Application to image processing.** We now consider the classical inverse problem of image deblurring (deconvolution) and denoising.

6.2.1. *The minimization problem.* We consider the case of grayscale images, represented as elements of $\mathbb{R}^{MN}$ (namely the set of $M \times N$ matrices with real coefficients), and we aim at computing a restored image u from an altered image $b \in \mathbb{R}^{MN}$. A classical approach consists in computing a solution of the total variation denoising and deconvolution model (see, e.g., [12, 19]) :

$$\min_{u \in \mathbb{R}^{MN}} \Theta(u) := \|\nabla u\|_1 + \lambda \|\ell u - b\|_2, \tag{6.2}$$

where $\ell : \mathbb{R}^{MN} \to \mathbb{R}^{MN}$ is a known blurring operator, ∇ is the discrete gradient operator on $\mathbb{R}^{MN}$ (∇ is a linear function from $\mathbb{R}^{MN}$ to $\mathbb{R}^{2MN}$) and λ is a trade-off parameter between data fidelity and smoothness. $\|\nabla(.)\|_1$ is the classical total variation operator introduced by Rudin-Osher-Fatemi in their seminal paper [28] and defined by

$$\|\nabla u\|_1 = \textstyle\sum_{i,j} \|(\nabla u)_{i,j}\|_2 \qquad \text{(since at any pixel } (i,j),\ (\nabla u)_{i,j} \text{ is a } \mathbb{R}^2\text{-vector).} \tag{6.3}$$

Minimizing the total variation is known to reduce the noise while preserving edges. This problem has been considered in [22] through a saddle point reformulation. However, it could hardly be solved with our method since the non-smooth function $\|\nabla(.)\|_1$ does not have an easy-to-compute proximal mapping.

To overcome this difficulty, we are led to consider a variant of the problem, which also presents operational advantages. Indeed, Total Variation methods for image denoising tend to artificially create flat areas in the restored image. This is called the staircasing problem (see [12] for more details on that subject). A classical technique to avoid this unwanted effect is to replace the L^1 norm in (6.3) by the Huber-norm defined for some positive (and small) parameter α by $J_\alpha(\nabla u) = \sum_{i,j} \psi_\alpha\left(\|(\nabla u)_{i,j}\|_2\right)$,

$$\text{where } \psi_\alpha\ (\mathbb{R} \to \mathbb{R}) \text{ is defined by } \psi_\alpha(x) = \begin{cases} |x| & \text{if } |x| > \alpha \\ \frac{|x|^2}{2\alpha} + \frac{\alpha}{2} & \text{otherwise} \end{cases}$$

This variant has also the advantage of smoothing the total variation term, which otherwise is not differentiable in any image having pixels with a null discrete gradient. The other modification of model (6.2) that we will consider consists in replacing the L^2 norm by the L^1 norm in the data term. This variant offers some advantages since it is contrast invariant and it is well known to be more effective against noise containing strong outliers, like for instance, salt and pepper noise. Thus, the problem we propose to solve becomes

$$\min_{u \in \mathbb{R}^{MN}} \Theta(u) := J_\alpha(\nabla u) + \lambda \|\ell u - b\|_1. \tag{6.4}$$

It can be easily checked that the total variation term $J_\alpha(\nabla u)$ is convex and continuously differentiable, with $\frac{\sqrt{8}}{\alpha}$-Lipschitz continuous gradient, while the data term $\lambda\|\ell u - b\|_1$ is convex, proper and lower semi-continuous, but not smooth. Therefore, (6.4) can be cast into problem (1.1) by setting $f(u) = J_\alpha(\nabla u)$ and $g(u) = \lambda\|\ell u - b\|_1$.

6.2.2. *Simulations for CRIPGA.* We aim at illustrating the performance of CRIPGA to solve problem (6.4), and highlighting the influence of the parameters involved in our algorithm. The experiments concern a denoising case ($\ell = I_H$), with the parameter $\lambda = 1.15$. The results are shown on Figure 3.

Next, we focus on the influence on the convergence to zero of $\bar{\Theta}(x_n) = \Theta(x_n) - \min\Theta$ of parameters $\{e_*, \nu_0\}$ (Figure 4) and $\{\kappa, \varepsilon\}$ (Figure 5). The influence of ε is again visible in



FIGURE 3. Application to total variation based image denoising with ℓ_1 fitting term. Original image (left), noisy input image (center), restored image (right) using $\lambda = 1.15$.

this example. In particular, it can be seen on Figure 5 that CRIPGA with adequate values of ε outperforms RIGPA (which corresponds to $\varepsilon = 0$). Moreover, the convergence is faster for low values of e_*, ν_0 between 50 and 100, and ε close to 0.01. Concerning κ, condition (5.20) enforces an implicit relation between κ and μ, which acts as a time step. Therefore, when illustrating the influence of κ, it is relevant to also adjust μ to a value close to its upper bound allowed by (5.20). Then, the best performance seems to be reached for κ around 1.

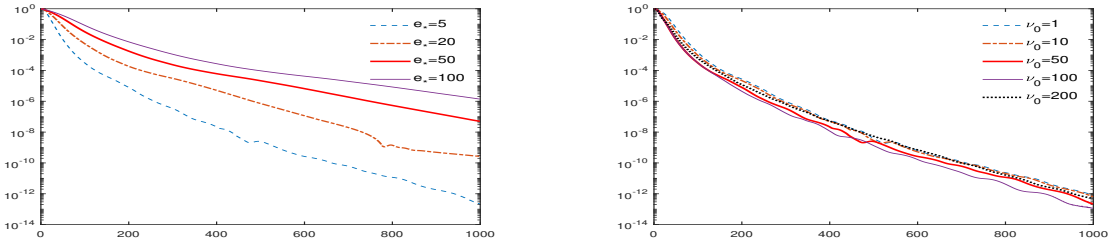


FIGURE 4. Evolution of $(\Theta(x_n) - \min\Theta)/(\Theta(x_0) - \min\Theta)$ for CRIPGA with $\nu_n = n + \nu_0$, for $\nu_0 = 50$ and several values of e_* (left figure) and for $e_* = 5$ and different values of ν_0 (right figure). The other parameters are $\kappa = 1$, $\mu = 4.8e-4$, $s_1 = 2.5$ and $\varepsilon = 0.01$.

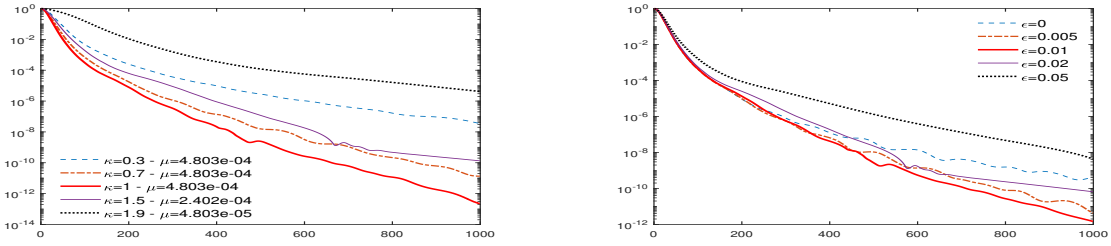


FIGURE 5. Evolution of $(\Theta(x_n) - \min\Theta)/(\Theta(x_0) - \min\Theta)$ for CRIPGA with $\nu_n = n + \nu_0$, for $\varepsilon = 0.01$ and several values of κ and μ (left figure) and different values of ε along with $\kappa = 1$ and $\mu = 4.25e-4$ (right figure). The other parameters are $e_* = 5$, $s_1 = 2.5$ and $\nu_0 = 50$.

6.3. **Comparing CRIPGA with some other well-suited algorithms.** In this section, we compare the performances of our proposed algorithm with that of other classical methods, namely the Forward-Backward Algorithm (FBA) and FISTA. On Figure 6, we compare the trajectories

of the iterates (x_n) given by CRIPGA, FBA and FISTA, through the convergence to zero of the value $(\Theta(x_n) - \min\Theta)/(\Theta(x_0) - \min\Theta)$ (Θ being the objective function in (6.4)). It is clearly observed that CRIPGA outperforms the other algorithms on the considered example.

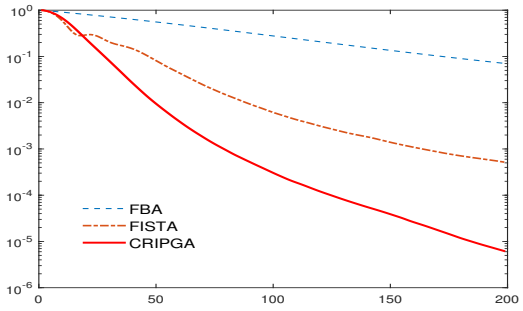


FIGURE 6. Profiles (in semi-logarithmic scale) of the quantity $\bar{\Theta}(x_n)/|\bar{\Theta}(x_0)|$ given by CRIPGA (with $e* = 5$, $s_1 = 2.5$, $\nu_n = n + 50$, $\kappa = 1$, $\varepsilon = 0.01$ and $\mu = 4.8e-04$), FBA (with $\mu = 4.95e-04$), and FISTA (with $\alpha = 3.1$ and $\mu = 4.95e-04$).

REFERENCES

[1] V. Apidopoulos, J.F. Aujol, C. Dossal, Convergence rate of inertial forward-backward algorithms beyond Nesterov's rule, Math. Program. 180 (2020), 137-156.
[2] H. Attouch, A. Cabot, Convergence of a relaxed inertial forward-backward algorithm for structured monotone inclusions, Appl. Math. Optim. 80 (2019), 547-598.
[3] H. Attouch, Z. Chbani, J. Fadili, H. Riahi, First-order optimization algorithms via inertial systems with Hessian driven damping, Math. Program. 193 (2022), 113–155.
[4] H. Attouch, Z. Chbani, J. Peypouquet, P. Redont, Fast convergence of inertial dynamics and algorithms with asymptotic vanishing viscosity, Math. Program. 168 (2018), 123–175.
[5] H. Attouch, J. Peypouquet, The rate of convergence of Nesterov's accelerated forward-backward method is actually faster than $1/k^2$, SIAM J. Optim. 26 (2016), 1824-1834.
[6] H. Attouch, J. Peypouquet, Convergence of inertial dynamics and proximal algorithms governed by maximal monotone operators, Math. Program. 174 (2019), 391–432.
[7] H. Attouch, J. Peypouquet, P. Redont, Fast convex optimization via inertial dynamics with Hessian driven damping, J. Differential Equations 261 (2016), 5734-5783.
[8] A. Beck, M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, SIAM J. Imaging Sci. 2 (2009), 183-202.
[9] H. Brezis, Opérateurs maximaux monotones, Math. Stud. 5, North-Holland, Amsterdam, 1973.
[10] H. Brezis, P.L. Lions: Produits infinis de résolvantes, Isr. J. Math. 29 (1978), 329-345.
[11] A. Chambolle, C. Dossal, On the convergence of the iterates of Fista, J. Optim. Theory Appl. 166 (2015), 968-982.
[12] A. Chambolle, T. Pock, A first-order primal-dual algorithm for convex problems with applications to imaging, J. Math. Imaging Vision 40 (2011), 120–145.
[13] O. Güler, On the convergence of the proximal point algorithm for convex minimization, SIAM J. Control Optim. 29 (1991), 403-419.
[14] O. Güler, New proximal point algorithms for convex minimization, SIAM J. Optim. 2 (1992), 649-664.
[15] F. Iutzeler, J.M. Hendrickx, A generic online acceleration scheme for optimization algorithms via relaxation and inertia, Optim. Methods Softw. 34 (2019), 383–405.
[16] F. Labarre, Approche numérique de problèmes d'optimisation et applications, PhD Thesis, Université des Antilles, 2021.

[17] B. Lemaire, The proximal algorithm, in: New methods in optimization and their industrial uses, J.P. Penot (ed), Internat. Ser. Numer. Math. vol. 87, pp. 73-87, Birkhauser, Basel, 1989.
[18] P.L. Lions, B. Mercier, Splitting algorithms for the sum of two nonlinear operators, SIAM J. Numer. Anal. 16 (1979), 964-979.
[19] D.A. Lorenz, T. Pock, An inertial forward-backward algorithm for monotone inclusions, J. Math. Imaging Vision 51 (2015), 311-325.
[20] P.E. Maingé, F. Labarre, Accelerated methods with fastly vanishing subgradients for structured non-smooth minimization, Numer. Algorithms 90 (2022), 99-136.
[21] P.E. Maingé, A. Weng-Law, Fast continuous dynamics inside the graph of subdifferentials of nonsmooth convex functions, Appl. Math. Optim. 89 (2024), Paper No. 1.
[22] P.E. Maingé, A. Weng-Law, Accelerated forward–backward algorithms for structured monotone inclusions, Comput. Optim. Appl. 88 (2024), 1-49.
[23] Y. Nesterov, A method of solving a convex programming problem with convergence rate O(1/k2), Soviet Math. Doklady 27 (1983), 372-376.
[24] Y. Nesterov, A method for unconstrained convex minimization problem with the rate of convergence O(1/k2), Dokl. Akad. Nauk. USSR 269 (1983), 543–547.
[25] Y. Nesterov, Gradient methods for minimizing composite objective function, Math. Program. 140 (2013), 125-161.
[26] Z. Opial, Weak convergence of the sequence of successive approximations for nonexpansive mappings, Bull. Amer. Math. Soc. 73 (1967), 591-597.
[27] G.B. Passty, Ergodic convergence to a zero of the sum of monotone operators in Hilbert space, J. Math. Anal. Appl. 72 (1979), 383-390.
[28] L. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, Physica D, 60 (1992), 259–268.
[29] B. Shi, S.S. Du, M.I. Jordan, W.J. Su, Understanding the acceleration phenomenon via high-resolution differential equations, Math. Program. 195 (2022), 79-148.
[30] W. Su, S. Boyd, E.J. Candès, A differential equation for modeling Nesterov's accelerated gradient method: theory and insights, J. Mach. Learn. Res. 17 (2016), 1-43.
[31] Q. Tran-Dinh, The connection between Nesterov's accelerated methods and Halpern fixed-point iterations, arXiv:2203.04869, 2022.
[32] S. Villa, S. Salzo, L. Baldassarre, A. Verri, Accelerated and inexact forward-backward algorithms, SIAM J. Optim. 23 (2013), 1607–1633.