J. Nonlinear Var. Anal. 10 (2026), No. 5, pp. 965-989
Available online at http://jnva.biemdas.com
https://doi.org/10.23952/jnva.10.2026.5.06

PRECONDITIONED TIKHONOV REGULARISED MONOTONE DYNAMICAL SYSTEMS AND APPLICATIONS TO PRIMAL-DUAL ALGORITHM IN CONVEX OPTIMIZATION

FOUAD BATTAHI, ZAKI CHBANI, HASSAN RIAHI*

*Laboratory of Mathematics, Modeling and Automatic Systems, Cadi Ayyad University,
Faculty of Sciences Semlalia, 40000, Marrakesh Morocco*

Abstract. The primary objective of this paper is to achieve strong convergence toward a zero of a maximally monotone operator in a Hilbert space. To this end, we first propose a continuous dynamic system that builds upon the framework presented by Boţ and Nguyen (2025), while remaining distinct from it. Specifically, we apply a positive linear operator to the velocity term $\dot{x}(t)$. This modification enables the adaptation of a suitable proximal algorithm when applying the proposed system to the context of the composite minimization problem $f(x)+g(Ax)$, where f and g are convex functions and A is a linear operator, and then to convex minimization problems under linear constraints, via a primal-dual algorithm by Chambolle and Pock (2011) for saddle points of the associated Lagrangian. The continuous model is analyzed for a single-valued operator M, allowing us to illustrate -through an appropriate discretization- the corresponding proximal point algorithm for a set-valued operator and to provide a consistent proof for the suitable convergence rates. The algorithmic contribution of this work is particularly significant, as it establishes strong convergence to the minimum norm of zeros of M without requiring additional conditions on its maximal monotonicity, a result that remains absent from more recent literature. We then provide strong convergence to the minimum norm solution, and also the same rate of convergence for values and constraints for the composite minimization problem and convex minimization under linear constraints.

Keywords. Composite convex minimization; First-order dynamical system; Strong convergence; Rate of convergence; Vanishing Tikhonov regularization.

1. INTRODUCTION

In this paper, our primary objective was to quickly achieve a solution to the following convex minimization problem:

$$\min_{x\in\mathscr{X}} f(x)+g(Ax), \qquad (\mathscr{P})$$

*Corresponding author.
E-mail address: h-riahi@uca.ac.ma (H. Riahi)
Received 30 November 2025; Accepted 4 April 2026; Published online 10 July 2026.

where

$$\begin{cases} * \quad \mathscr{X} \text{ and } \mathscr{Y} \text{ are real Hilbert spaces,} \\ * \quad f:\mathscr{X}\to\mathbb{R}\cup\{+\infty\} \text{ and } g:\mathscr{Y}\to\mathbb{R}\cup\{+\infty\} \text{ are proper, closed and convex functions ,} \\ * \quad A:\mathscr{X}\to\mathscr{Y} \text{ is a linear continuous operator .} \end{cases} \tag{$\mathbf{H}_0$}$$

So, we assume that the (closed and convex) set of the optimal solutions S of ($\mathscr{P}$) is non-empty. Using the first-order optimality conditions and a qualification condition on the maximality of the sum (see [1, 2]), we first ensure that

$$\begin{aligned} \bar{x}\in S \quad &\Longleftrightarrow \quad 0\in\partial(f+g\circ A)(\bar{x})=\partial f(\bar{x})+A^*\partial g(A\bar{x}) \\ &\Longleftrightarrow \quad \exists\bar{\lambda}\in\mathscr{Y}^* \text{ such that } -A^*\bar{\lambda}\in\partial f(\bar{x}) \text{ and } A\bar{x}\in\partial g^*(\bar{\lambda}), \end{aligned} \tag{1.1}$$

where $g^*:\mathscr{Y}\to\mathbb{R}\cup\{+\infty\}$ is the Fenchel conjugate of g defined as

$$g^*(z)=\sup_{y\in\mathscr{Y}}(\langle z,y\rangle-g(y)).$$

Without any qualification condition, the above inverse implication remains true, since we still have the sum of the subdifferentials is included in the subdifferential of the sum. The primal-dual conditions in (1.1) are equivalent to finding a saddle point of the associated Lagrangian $\mathscr{L}:\mathscr{X}\times\mathscr{Y}\to\mathbb{R}$ defined by

$$\mathscr{L}(x,\lambda)=f(x)+\langle\lambda,Ax\rangle-g^*(\lambda). \tag{1.2}$$

This means (see [3]) that $\bar{x}\in\mathscr{X}$ is an optimal solution of ($\mathscr{P}$) if and only if there exists a corresponding Lagrange multiplier $\bar{\lambda}\in\mathscr{Y}$ such that $(\bar{x},\bar{\lambda})$ is a saddle point of $\mathscr{L}$, that is,

$$\mathscr{L}(\bar{x},\lambda)\le\mathscr{L}(\bar{x},\bar{\lambda})\le\mathscr{L}(x,\bar{\lambda}) \quad \forall(x,\lambda)\in\mathscr{X}\times\mathscr{Y}.$$

This justifies that the set $\mathscr{S}$ of saddle points for $\mathscr{L}$ is nonempty.

Let us further assume that the Lagrangian $\mathscr{L}$ is $\mathscr{C}^1$. The necessary and sufficient primal-dual optimality conditions to attain a saddle point of $\mathscr{L}$ are read as follows:

$$(\bar{x},\bar{\lambda})\in\mathscr{S}\Leftrightarrow\begin{cases}\nabla_x\mathscr{L}(\bar{x},\bar{\lambda})=0\\ \nabla_\lambda\mathscr{L}(\bar{x},\bar{\lambda})=0\end{cases}\Leftrightarrow\begin{cases}\nabla f(\bar{x})+A^*\bar{\lambda} &=0\\ A\bar{x}-\nabla g^*(\bar{\lambda}) &=0.\end{cases} \tag{1.3}$$

Moreover, by setting $M:\mathscr{H}:=\mathscr{X}\times\mathscr{Y}\to\mathscr{H}$, defined by $M(x,\lambda)=(\nabla_x\mathscr{L}(x,\lambda),-\nabla_\lambda\mathscr{L}(x,\lambda))$ for all $(x,\lambda)\in\mathscr{H}$, the saddle points of the Lagrangian $\mathscr{L}$ are characterized by finding the zeros of the operator M. Therefore, to deal with problem ($\mathscr{P}$), we now turn to the following problem:

$$\text{Find } u\in\mathscr{H} \text{ such that } Mu=0 \tag{$\mathscr{M}$}$$

where, due to Rockafellar [4, 5], the conditions in ($\mathbf{H}_0$) ensure that $M:\mathscr{H}\to\mathscr{H}$ is maximally monotone. Much more, this problem includes, as special cases, variational inequalities and non-smooth convex optimization problems. Noticing also that the zeros of M, which are the saddle points of L, are characterized by $\mathscr{S}=S_f\times L_f$, where S_f is the set of optimal solutions of ($\mathscr{P}$) and L_f is the set of associated Lagrange multipliers.

Thus, we start from dealing with the problem of finding a zero of a maximally monotone operator M defined on a real Hilbert space $\mathscr{H}$, and then we apply our results to the saddle points of the associate Lagrangian $\mathscr{L}$ in order to reach a minimum of a convex function under linear constraints. In order to tackle these problems in a more finely tuned and less demanding way, we use the links between the algorithms and dissipative continuous dynamical systems with Tikhonov regularization, as well as their asymptotic analysis by Lyapunov methods. As we can see, our algorithms can be derived from the time discretization of the associated dynamic systems. To express this discretization relationship for saddle points, we try to come up with a less costly and more efficient proximal algorithm via the related monotone operators. To resolve ($\mathscr{P}$), in the case where f and g^* are not necessarily differentiable, we start in reverse with the following discrete algorithm:

$$\begin{cases} x_{k+1}-x_k+\partial f(x_{k+1})+A^*\lambda_k+\varepsilon_{k+1}x_{k+1} & \ni 0, \\ -2Ax_{k+1}+Ax_k+\lambda_{k+1}-\lambda_k+\partial g^*(\lambda_{k+1})+\varepsilon_{k+1}\lambda_{k+1} & \ni 0, \end{cases} \tag{SP_k}$$

which, by setting $\sigma_k=\frac{1}{1+\varepsilon_{k+1}}$, can be written as

$$\begin{cases} x_{k+1}=[I_{\mathscr{X}}+\sigma_k\partial f]^{-1}\left(\sigma_k(x_k-A^*\lambda_k)\right), \\ \lambda_{k+1}=(I_{\mathscr{Y}}+\sigma_k\partial g^*)^{-1}\left(\sigma_k(\lambda_k+2Ax_{k+1}-Ax_k)\right). \end{cases} \tag{1.4}$$

This algorithm is more suitable for solving problem ($\mathscr{P}$), because in the first part of this iteration it uses the proximal of the objective function f in a direct way to calculate x_{k+1}, then in the second part, the calculation of λ_{k+1} is deduced directly from x_k,λ_k and x_{k+1}.

First, we note that algorithm (SP_k) is related to the well-known Chambolle-Pock method [6] by exchanging the updates for λ_{k+1} and x_{k+1}. In fact, we extend the study to general Hilbert spaces, then add a Tikhonov regularization term $\varepsilon_{k+1}I_{\mathscr{X}\times\mathscr{Y}}$ which allows us to achieve strong convergence of the iterates to a selected solution of ($\mathscr{P}$) . Return to (SP_k) and rewrite it as

$$\begin{cases} x_{k+1}-x_k+\partial_x\mathscr{L}(x_{k+1},\lambda_{k+1})-A^*(\lambda_{k+1}-\lambda_k)+\varepsilon_{k+1}x_{k+1}\ni 0, \\ -A(x_{k+1}-x_k)+\lambda_{k+1}-\lambda_k-\partial_\lambda\mathscr{L}(x_{k+1},\lambda_{k+1})+\varepsilon_{k+1}\lambda_{k+1}\ni 0. \end{cases} \tag{1.5}$$

By setting $u_k=(x_k,\lambda_k)$ and for $u=(x,\lambda), Mu=(\partial_x\mathscr{L}(x,\lambda),-\partial_\lambda\mathscr{L}(x,\lambda))$, in view of (1.5), ($SP_k$) is equivalent to

$$C(u_{k+1}-u_k)+Mu_{k+1}+\varepsilon_{k+1}u_{k+1}\ni 0, \tag{S_k}$$

where $C=\begin{pmatrix} I_{\mathscr{X}} & -A^* \\ -A & I_{\mathscr{Y}}\end{pmatrix}$ is the symmetric preconditioner operator. This last reformulation can be seen as a backward discretization associated to the continuous dynamics

$$C\dot{u}(t)+Mu(t)+\varepsilon(t)u(t)\ni 0. \tag{S_0}$$

We recall that, for the study of ($\mathscr{M}$), Boţ and Nguyen [7] considered the dynamics

$$\dot{u}(t)+Mu(t)+\varepsilon(t)u(t)=0, u(t_0)=u_0\in\mathscr{H}, \tag{1.6}$$

which, by implicit discretization, gives

$$(u_{k+1}-u_k)+Mu_{k+1}+\varepsilon_{k+1}u_{k+1}=0. \tag{1.7}$$

Since $M(x,\lambda)=(\nabla_x\mathscr{L}(x,\lambda),-\nabla_\lambda\mathscr{L}(x,\lambda))$, we deduce the continuous dynamics

$$\begin{cases} \dot{x}(t)+\nabla_x\mathscr{L}(x(t),\lambda(t))+\varepsilon(t)x(t), \\ \dot{\lambda}(t)-\nabla_\lambda\mathscr{L}(x(t),\lambda(t))+\varepsilon(t)\lambda(t), \end{cases} \tag{SP_0}$$

and the associate implicit discretization

$$\begin{cases} x_{k+1}=[I_{\mathscr{Y}}+\sigma_k\nabla f]^{-1}\left(\sigma_k(x_k-A^*\lambda_{k+1})\right), \\ \lambda_{k+1}=[I_{\mathscr{Y}}+\sigma_k\nabla g^*]^{-1}\left(\sigma_k(\lambda_k+Ax_{k+1})\right) \end{cases} \tag{1.8}$$

As we notice, this algorithm requires at each iteration a resolution of a system in (x_{k+1},λ_{k+1}), which is more expensive to evaluate. So, to avoid this difficulty, the introduction of the operator C in the system (S_0), allows us to evaluate in (SP_k) directly x_{k+1} as a function of (x_k,λ_k) (see (1.4)), and then, λ_{k+1} is deduced from (x_k,λ_k) and x_{k+1}.

We now turn to a brief history of convergence study of backward discretization and continuous aspect of the first-order system $0\in\dot{u}(t)+Mu(t)$. The backward methods for reaching a zero of a maximally monotone operator were pioneered by Martinet [8] who proposed the proximal algorithm for minimizing a convex function, then Rockafellar in [9] extended this proximal method to maximally monotone operators and proved the weak convergence of this algorithm. As explained by Güler, this proximal algorithm doesn't converge strongly in an infinite-dimensional Hilbert space for a general maximally monotone operator. In order to justify the interest of the proximal point method in accelerating its convergence rate, there were many interesting contributions to both theoretical and practical interest (see, for example, [4, 5, 7, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20]).

The continuous aspect began with Bruck's result [21], who showed that in a real Hilbert space $\mathscr{X}$, the trajectories of the system $0\in\dot{x}(t)+\partial f(x(t))$ converge weakly to a minimum of f. Baillon and Brézis [22] generalized this result to maximal monotone operators, and immediately afterward Baillon [23] provided an example in which the solutions of the first system converge weakly, but not strongly. We refer to [24, 25], for more details on the convergence for discrete and continuous dynamical systems governed by maximal monotone operators. In [26], Attouch and Cominetti coupled the dynamic steepest descent method and a Tikhonov regularization term

$$\dot{x}(t)+\partial f(x(t))+\varepsilon(t)x(t)\ni 0. \tag{1.9}$$

The striking point of their analysis is the strong convergence of the trajectory $x(t)$ when the regularization parameter $\varepsilon(t)$ tends to zero with a sufficiently slow rate of convergence $\varepsilon\notin L^1(\mathbb{R}_+,\mathbb{R})$. Then the strong limit is the minimum norm element of agrmin f. However, if $\varepsilon(t)=0$ we can only expect a weak convergence of the induced trajectory $x(t)$. Cominetti et al. [27] considered, for a maximally monotone operator on a Hilbert space, the dynamics

$$\dot{u}(t)+Mu(t)+\varepsilon(t)u(t)=0, u\left(t_0\right)=u_0\in\mathscr{H}. \tag{1.10}$$

They established a strong convergence of $u(t)$ to the minimum norm element in $M^{-1}0$ whenever $\varepsilon(t)\searrow 0$ as $t\nearrow\infty$, $\int_0^\infty\varepsilon(t)dt=\infty$ and either $\lim_{t\to+\infty}\frac{\dot{\varepsilon}(t)}{\varepsilon(t)^2}=0$ ([27, Proposition 5]), or $\int_0^\infty\dot{\varepsilon}(t)dt<\infty$ ([27, Theorem 9]). More recently, Boţ and Nguyen [7] have established the same result for the dynamic (1.10) by omitting the last two conditions involving the derivative of $\varepsilon(t)$.

In order to solve ($\mathscr{P}$), the authors in [28] considered the following preconditioned dynamic system:

$$D(t)\dot{u}(t)+Mu(t)+\varepsilon(t)u(t)\ni 0, \tag{1.11}$$

where

$$u=(x,\lambda),\; Mu=\left(\nabla_x\mathscr{L}(x,\lambda),-\nabla_\lambda\mathscr{L}(x,\lambda)\right),\; D(t)=\begin{pmatrix}\alpha(t)I_{\mathscr{X}} & \beta(t)A^*\\ \gamma(t)A & \delta(t)I_{\mathscr{Y}}\end{pmatrix}$$

and $\alpha,\beta,\gamma,\delta$ are scalar functions. This system generalizes the Luo's [29] preconditioned triangular one, where the author assume $\beta(t)=0,\gamma(t)=1$ and f is L-smooth and μ-convex with

$\mu \geq 0$. In (1.11), system (1.10) requires a time-varying preconditioner that acts on the time derivatives. In the case of a linear constrained optimization problem, i.e., $g = \iota_b$, they proved in the asymptotically antisymmetric case, i.e.,

$$\alpha(t) = \frac{\alpha_0}{\delta_0}\delta(t) = \frac{\alpha_0}{\beta_0}\beta(t)\exp\left(-\int_0^t \frac{ds}{\beta(s)}\right), \alpha_0, \beta_0, \delta_0 > \text{ and } \gamma(t) = -\beta(t),$$

that the constraint $\|Ax - b\|$ and the values $|f(x(t)) - f(\bar{x})|$ decrease exponentially. Furthermore, they concluded that each weak limit point of the subsequence of trajectories is a primal-dual solution. If the objective function is strongly convex, they obtained strong convergence of the primal trajectory to a solution of the linear constrained optimization problem. Nevertheless, referring to Example 5.2, when comparing the convergence rates for a quadratic minimization problem under linear constraints similar to that in [28, Example 8.2], we observe that the convergence rate of the values and the runtime are more efficient for (3.14) than for (1.11). This can be explained by the time-dependent factors included in the operator $D(t)$, which may require more time and effort to solve the system (1.11).

Inspired by recent works in [7, 28, 30], the first contribution of this paper is devoted to the asymptotic analysis as $t \to +\infty$ for solutions of the continuous dynamical system with vanishing Tikhonov regularization (S_0). Under the suitable proposed conditions $\lim_{t\to+\infty} \varepsilon(t) = 0$ and $\lim_{t\to+\infty} \int_{t_0}^t \varepsilon(s)ds = +\infty$ on the Tikhonov terms $\varepsilon(t)$, we show (see Theorem 3.2) that the solutions of the regularized differential system (S_0) strongly converges toward the element of least norm within the solution set $\mathscr{S}$. Moreover, we provide fast asymptotic decay rate estimates for convergence of $Mu(t)$ and $\dot{u}(t)$ to the origine. We therefore, slightly extend [7, Corollaries 2.10, 2.11] by adding the linear operator C in the system (S_0). For situations where $M(x,\lambda) = (\nabla_x \mathscr{L}(x,\lambda), -\nabla_\lambda \mathscr{L}(x,\lambda))$, we present, as application of Theorem 3.2, for solutions of the system (SP_0) the same convergence rates for the operator $(\nabla_x \mathscr{L}(x,\lambda), -\nabla_\lambda \mathscr{L}(x,\lambda))$ and $(\dot{x}(t), \dot{\lambda}(t))$ (see Theorem 3.3). This theorem also provides the same rate of convergence for values and constraints. Section 4, the main contribution of this paper, is devoted to an implicit discretization of the continuous systems (S_0) for monotone inclusions and (SP_0) for linear constrained convex minimization problems. We propose the appropriate proximal algorithms (1.7) and (1.8) for solving ($\mathscr{M}$) and ($\mathscr{P}$), respectively. We then give strong convergence results and suitable convergence rates for values and constraints.

2. Notations and Preliminaries

Let $\mathscr{H}$ be a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and corresponding norm $\|\cdot\|^2$. Suppose that C is a continuous, self-adjoint, and positive-definite linear operator on $\mathscr{H}$, i.e., there exist $\alpha > 0$ and $\beta > 0$ such that

$$\beta\|u\|^2 \leq \langle Cu, u\rangle \leq \alpha\|u\|^2, \ \forall u \in \mathscr{H}. \tag{2.1}$$

This ensures that the norm associated with C defined by $\|u\|_C = \sqrt{\langle Cu, u\rangle}$ is equivalent to the norm $\|\cdot\|$. On the other hand, the fact that C is self-adjoint allows us to define the inner product associated with C by

$$\langle u, y\rangle_C = \langle u, Cy\rangle = \langle Cu, y\rangle = \langle y, u\rangle_C.$$

Note also that (2.1) is equivalent to

$$\frac{1}{\alpha}\|u\|^2 \leq \|u\|^2_{C^{-1}} \leq \frac{1}{\beta}\|u\|^2, \ \forall u \in \mathscr{H}. \tag{2.2}$$

For a set-valued operator $M:\mathscr{H}\to 2^{\mathscr{H}}$, we denote by $\mathrm{dom}(M)=\{u\in\mathscr{H}: Mu\neq\emptyset\}$, $M^{-1}0=\{u\in\mathscr{H}: 0\in Mu\}$. We identify the following forms

$$(u,\xi)\in M \iff \xi\in Mu \iff u\in M^{-1}\xi \iff (\xi,u)\in M^{-1}.$$

M is called monotone if $\left\langle \xi-\zeta\ ,\ u-v\right\rangle\geq 0$ for all $(u,\xi),(v,\zeta)\in M$, and it is maximally monotone provided that it is monotone and there is no other monotone operator containing M in $\mathscr{H}\times\mathscr{H}$, i.e., M is maximal in the family of monotone subsets of $\mathscr{H}\times\mathscr{H}$, ordered by inclusion. As examples of maximally monotone operators, we cite [31] for the subdifferential of proper lower semicontinuous convex functions and [4, Corollary 1] for closed proper convex-concave saddle functions. The resolvent of the operator M is defined as $J_M=(I_{\mathscr{H}}+M)^{-1}$. If M is maximally monotone, then the resolvent J_M is an every where defined, nonexpansive operator (and thus single-valued). For more details on monotone operators, we refer the reader to [32].

3. Continuous Asymptotic Aspects

In this section, for a single valued monotone operator $M:\mathscr{H}\to\mathscr{H}$, we consider a continuous aspect which is the system (S_0) to solve the equation $Mu=0$, and saddle points problem (1.3). Subsequently, by temporal discretization of the dynamics, we adapt a class of proximal inertial algorithms with fast convergence properties.

3.1. **Reaching a zero of the monotone operator M.** The following theorem is similar to [7, Theorem 2.9]; however, its proof requires an adaptation to the form of system (S_0). In order to justify the majorizations associated with the operator C, we next detail its proof. For $t\in[t_0,+\infty[$, define $\gamma(t)=\exp\left(\frac{1}{\alpha}\int_{t_0}^{t}\varepsilon(s)ds\right)$, and the energy function

$$E(t):=\frac{1}{2}\|Mu(t)+\varepsilon(t)u(t)\|^2_{C^{-1}}. \tag{3.1}$$

We first remark that

$$E(t)=\frac{1}{2}\|C\dot{u}(t)\|^2_{C^{-1}}=\frac{1}{2}\left\langle C\dot{u}(t),C\dot{u}(t)\right\rangle_{C^{-1}}=\frac{1}{2}\left\langle C\dot{u}(t),\dot{u}(t)\right\rangle=\frac{1}{2}\|\dot{u}(t)\|^2_{C}. \tag{3.2}$$

Theorem 3.1. *Let $u:[t_0;+\infty[\to\mathscr{H}$ be a solution of (S_0) and assume that $\int_{t_0}^{+\infty}\varepsilon(t)dt=+\infty$. Then*

(1) *the trajectory $u(t)$ is bounded;*
(2) *it holds the estimation: for each $t>t_0$,*

$$E(t)\leqslant\frac{E(t_0)\gamma(t_0)}{\gamma(t)}+\frac{R}{2\gamma(t)}\int_{t_0}^{t}\frac{(\dot{\varepsilon}(s))^2}{\varepsilon(s)}\gamma(s)ds.$$

Proof. (1) Let x^* be an element of $M^{-1}0$ and consider $w(t):=\frac{1}{2}\|u(t)-x^*\|^2_C$. Then ($S_0$) ensures

$$\dot{w}(t)=\left\langle C\dot{u}(t),u(t)-x^*\right\rangle=-\left\langle Mu(t),u(t)-x^*\right\rangle-\varepsilon(t)\left\langle u(t),u(t)-x^*\right\rangle.$$

Since $Mx^*=0$ and M is monotone, we have

$$\left\langle Mu(t),u(t)-x^*\right\rangle=\left\langle Mu(t)-Mx^*,u(t)-x^*\right\rangle\geqslant 0,$$

which gives

$$\begin{aligned}
\dot{w}(t) &\leqslant -\varepsilon(t)\langle u(t), u(t)-x^*\rangle \\
&= -\varepsilon(t)\|u(t)-x^*\|^2 - \varepsilon(t)\langle x^*, u(t)-x^*\rangle . \\
&\leqslant -\frac{\varepsilon(t)}{2}\|u(t)-x^*\|^2 + \frac{\varepsilon(t)}{2}\|x^*\|^2 \\
&\leqslant -\frac{\varepsilon(t)}{2\alpha}\|u(t)-x^*\|_C^2 + \frac{\varepsilon(t)}{2}\|x^*\|^2 \\
&= -\frac{\varepsilon(t)}{\alpha}w(t) + \frac{\varepsilon(t)}{2}\|x^*\|^2.
\end{aligned}$$

We deduce for almost all $t > t_0$,

$$\dot{w}(t) + \frac{\varepsilon(t)}{\alpha}w(t) \leqslant \frac{\varepsilon(t)}{2}\|x^*\|^2.$$

Multiplying by $\gamma(t)$, we obtain

$$\frac{d}{dt}(\gamma(t)w(t)) \leq \frac{\|x^*\|^2}{2}\varepsilon(t)\gamma(t) = \frac{\alpha\|x^*\|^2}{2}\dot{\gamma}(t)$$

Integrating on $[t_0, t[$, we obtain

$$\gamma(t)w(t) - \gamma(s)w(s) \leq \frac{\|x^*\|^2}{2}\int_s^t \gamma(\tau)\varepsilon(\tau)d\tau = \alpha(\gamma(t)-\gamma(s)).$$

If $t > s \geq t_0$, $\gamma(t)(w(t)-\alpha) \leq \gamma(s)(w(s)-\alpha) \leq \gamma(t_0)(w(t_0)-\alpha)$. Thus, for $t > t_0$, $\gamma(t)(w(t)-\alpha) \leq \gamma(t_0)(w(t_0)-\alpha)$,

$$w(t) \leq \frac{w(t_0)\gamma(t_0)}{\gamma(t)} + \frac{\|x^*\|^2}{2\gamma(t)}\int_{t_0}^t \gamma(s)\varepsilon(s)ds.$$

Since $\alpha\dot{\gamma}(t) = \varepsilon(t)\gamma(t)$, we have $\int_{t_0}^t \gamma(s)\varepsilon(s)ds = \alpha(\gamma(t)-\gamma(t_0))$. Then

$$w(t) \leq \frac{w(t_0)\gamma(t_0)}{\gamma(t)} + \frac{\alpha\|x^*\|^2}{2},$$

which gives that $w(t) := \frac{1}{2}\|u(t)-x^*\|_C^2$ is bounded since $\lim_{t\to+\infty}\gamma(t) = +\infty$. Thus $u(t)$ is bounded.

(2) Differentiating the energy function $E(t)$ defined by (3.1), we have

$$\begin{aligned}
\dot{E}(t) &= \left\langle \frac{d}{dt}(Mu(t)+\varepsilon(t)u(t)), \, Mu(t)+\varepsilon(t)u(t) \right\rangle_{C^{-1}} \\
&= \left\langle \frac{d}{dt}(Mu(t)) + \dot{\varepsilon}(t)u(t) + \varepsilon(t)\dot{u}(t), \, \underbrace{Mu(t)+\varepsilon(t)u(t)}_{-C\dot{u}(t)} \right\rangle_{C^{-1}} \\
&= \left\langle \frac{d}{dt}(Mu(t)) + \dot{\varepsilon}(t)u(t) + \varepsilon(t)\dot{u}(t), \, -\dot{u}(t) \right\rangle \\
&= -\left\langle \frac{d}{dt}(Mu(t)), \, \dot{u}(t) \right\rangle - \dot{\varepsilon}(t)\langle \dot{u}(t), u(t)\rangle - \varepsilon(t)\|\dot{u}\|^2.
\end{aligned}$$

Since $-\left\langle \frac{d}{dt}(Mu(t)), \dot{u}(t)\right\rangle \leqslant 0$ (this is due to the monotonicity of M), then

$$\begin{aligned}\dot{E}(t) &\leqslant -\dot{\varepsilon}(t)\langle \dot{u}(t), u(t)\rangle - \varepsilon(t)\|\dot{u}\|^2\\ &\leqslant \frac{\dot{\varepsilon}(t)^2}{2\varepsilon(t)}\|u(t)\|^2 - \frac{\varepsilon(t)}{2}\|\dot{u}(t)\|^2\\ &\leqslant \frac{\dot{\varepsilon}(t)^2}{2\varepsilon(t)}\|u(t)\|^2 - \frac{\varepsilon(t)}{2\alpha}\|\dot{u}(t)\|_C^2\\ &= \frac{\dot{\varepsilon}(t)^2}{2\varepsilon(t)}\|u(t)\|^2 - \frac{\varepsilon(t)}{\alpha}E(t),\end{aligned}$$

where the last inequality is obtained from (3.2). We then obtain

$$\dot{E}(t) + \frac{\varepsilon(t)}{\alpha}E(t) \leqslant \frac{(\dot{\varepsilon}(t))^2}{2\varepsilon(t)}\|u(t)\|^2.$$

Using the boundedness of $(u(t))$, there exists $R > 0$ such that

$$\dot{E}(t) + \frac{\varepsilon(t)}{\alpha}E(t) \leq \frac{R}{2}\frac{\dot{\varepsilon}(t)^2}{\varepsilon(t)}.$$

Multiplying by $\gamma(t)$ and integrating on $[t_0, t[$, we obtain

$$E(t) \leqslant \frac{E(t_0)\gamma(t_0)}{\gamma(t)} + \frac{R}{2\gamma(t)}\int_{t_0}^{t}\frac{\dot{\varepsilon}(s)^2}{\varepsilon(s)}\gamma(s)ds \tag{3.3}$$

□

To illustrate the convergence results, let us consider the specific case of $\varepsilon(t) = \frac{\alpha\delta}{t}$, where the parameters α, δ are positive.

Theorem 3.2. *Let $x : [t_0; +\infty[\to \mathscr{H}$ be the unique solution of the following system:*

$$C\dot{u}(t) + Mu(t) + \frac{\alpha\delta}{t}u(t) = 0,$$

Then, $u(t)$ converges strongly to the minimum norm solution $\bar{x} \in M^{-1}(0)$ and $\lim_{t\to+\infty} Mu(t) = \lim_{t\to+\infty}\dot{u}(t) = 0$, with the following estimations:

(1)

$$\|Mu(t)\| = \mathscr{O}\left(\frac{1}{t^{\delta/2}} + \frac{1}{t}\right), \textit{ if } \delta \neq 2; \tag{3.4}$$

$$\|Mu(t)\| = \mathscr{O}\left(\frac{\sqrt{\ln(t)}}{t}\right), \textit{ if } \delta = 2. \tag{3.5}$$

(2)

$$\|\dot{u}(t)\| = \mathscr{O}\left(\frac{1}{t^{\delta/2}} + \frac{1}{t}\right), \textit{ if } \delta \neq 2; \tag{3.6}$$

$$\|\dot{u}(t)\| = \mathscr{O}\left(\frac{\sqrt{\ln(t)}}{t}\right), \textit{ if } \delta = 2. \tag{3.7}$$

Proof. (1) Since $\varepsilon(t) := \frac{\alpha\delta}{t}$, we have, for $t > t_0$,

$$\gamma(t) = \exp\left(\frac{1}{\alpha}\int_{t_0}^{t}\varepsilon(s)ds\right) = \exp\left(\ln\left(\frac{t}{t_0}\right)^{\delta}\right) = \left(\frac{t}{t_0}\right)^{\delta},$$

and $\gamma(t_0) = 1$. Also $\dot{\varepsilon}(t) = -\frac{\alpha\delta}{t^2}$. Then (3.3) becomes

$$E(t) \leqslant \frac{t_0^{\delta}E(t_0)\gamma(t_0)}{t^{\delta}} + \frac{R}{2t^{\delta}}\int_{t_0}^{t}\frac{\alpha\delta}{s^{3-\delta}}ds$$

• If $\delta \neq 2$, then

$$\int_{t_0}^{t}\frac{\alpha\delta}{s^{3-\delta}}ds = \frac{\alpha\delta t^{\delta-2}}{\delta-2} - \frac{\alpha\delta t_0^{\delta-2}}{\delta-2}.$$

Thus we obtain

$$E(t) \leqslant \frac{t_0^{\delta}E(t_0)}{t^{\delta}} + \frac{\alpha\delta R}{2(\alpha-2)t^2}. \tag{3.8}$$

Moreover, according to (2.2), we have

$$\|Mu(t)+\varepsilon(t)u(t)\|^2 \leq \alpha\|Mu(t)+\varepsilon(t)u(t)\|_{C^{-1}}^2 = 2\alpha E(t). \tag{3.9}$$

According to (3.8), we get

$$\|Mu(t)+\varepsilon(t)u(t)\| \leq \frac{\sqrt{2\alpha E(t_0)}}{t^{\frac{\delta}{2}}} + \sqrt{\frac{\alpha^2\delta R}{\alpha-2}}\frac{1}{t}. \tag{3.10}$$

On the other hand, we have

$$\|Mu(t)\| \leq \|Mu(t)+\varepsilon(t)u(t)\| + \varepsilon(t)\|u(t)\|. \tag{3.11}$$

Combining the above inequality with (3.10), we easily arrive at (3.4).

• If $\delta = 2$, then

$$E(t) \leqslant \frac{t_0^2E(t_0)}{t^2} + \frac{Rt_0^2}{2t^2}\int_{t_0}^{t}\frac{2\alpha}{s}ds \leqslant \frac{t_0^2E(t_0)}{t^2} + \frac{\alpha Rt_0^2\ln(t)}{t^2}. \tag{3.12}$$

Combining the above inequality with (3.9), we have

$$\|Mu(t)+\varepsilon(t)u(t)\| \leq \frac{t_0}{t}\left(\sqrt{2\alpha E(t_0)} + \alpha\sqrt{2R\ln(t)}\right),$$

and then (3.11) ensures (3.5).

(2) From (2.1) and (3.2), we have $\|\dot{u}(t)\|^2 \leq \frac{1}{\beta}\|\dot{u}(t)\|_C^2 = \frac{2}{\beta}E(t)$. Then (3.8) ensures (3.6) if $\delta \neq 2$, and (3.12) ensures (3.7) if $\delta = 2$.

(3) The strong convergence of $\{u(t)\}$ towards the minimal norm solution $\bar{x} \in M^{-1}0$ was justified by [27, Proposition 6] since our settings do not affect the validity. So, we only need to justify that every weak cluster point of $\{u(t)\}$, for $t \to \infty$, belongs to $M^{-1}0$. Let us consider a weak cluster point x^* of $\{u(t)\}$. Then there exists $t_k \to \infty$ such that $\{u(t_k)\}$ converges weakly to x^*. Since $\{u(t)\}$ is bounded and the sequences $\{\varepsilon(t_k)\}, \{\|\dot{u}(t_k)\|^2\}$ converge to zero, as $k \to +\infty$, then $\{-C\dot{u}(t_k) - \varepsilon(t_k)u(t_k)\}$ strongly converges to the origine of $\mathscr{H}$. Let us tend k towards infinity in $-C\dot{u}(t_k) - \varepsilon(t_k)u(t_k) = Mu(t_k)$. We deduce from weak-strong sequential closedness of the graph of each maximally monotone operator, that is, $0 = Mu^*$. Thus $u^* \in M^{-1}0$. □

Remark that $\varepsilon(t)=\frac{\alpha\delta}{t}$ does not satisfy the condition $\lim_{t\to+\infty}\frac{\dot{\varepsilon}(t)}{\varepsilon(t)^2}=0$ imposed in [27, Proposition 5] to deduce strong convergence.

3.2. **Application to a composite convex minimization.** As we already explained that in the introduction composite minimization problem ($\mathscr{P}$) is equivalent to a zero of the maximally monotone operator $(x,\lambda)\mapsto(\nabla_x\mathscr{L}(x,\lambda),-\nabla_\lambda\mathscr{L}(x,\lambda))$. To use Theorem 3.2, we slightly modify the form of operator C. This ensures that regularity condition (2.1) is verified on C by only assuming A to be positive.

Lemma 3.1. *Let A be positive. Then, for $\tau\in(0,1/\|A\|)$, the linear operator $C=\begin{pmatrix} I_{\mathscr{X}} & -\tau A^* \\ -\tau A & I_{\mathscr{Y}}\end{pmatrix}$ is positive-definite and satisfies, for $\eta=\|A\|$,*

$$(1-\tau\eta)\|(x,\lambda)\|^2\le\|(x,\lambda)\|_C^2\le(1+\tau\eta)\|(x,\lambda)\|^2\quad\forall(x,\lambda)\in\mathscr{X}\times\mathscr{Y}.$$

Proof. Note that

$$\begin{aligned}\|(x,\lambda)\|_C^2&=\left\langle C\begin{pmatrix}x\\ \lambda\end{pmatrix},\begin{pmatrix}x\\ \lambda\end{pmatrix}\right\rangle=\left\langle\begin{pmatrix}x-\tau A^*\lambda\\ -\tau Ax+\lambda\end{pmatrix},\begin{pmatrix}x\\ \lambda\end{pmatrix}\right\rangle\\ &=\|x\|^2-2\tau\langle Ax,\lambda\rangle+\|\lambda\|^2\\ &\ge\|x\|^2-2\tau\eta\|x\|\|\lambda\|+\|\lambda\|^2 \qquad\qquad (\text{for } \eta=\|A\|)\\ &\ge\|x\|^2-\tau\eta(\|x\|^2+\|\lambda\|^2)+\|\lambda\|^2\\ &=(1-\tau\eta)\|(x,\lambda)\|^2.\end{aligned}$$

Since $1-\tau\eta>0$, we deduce that C is positive-definite and $(1-\tau\eta)\|(x,\lambda)\|^2\le\|(x,\lambda)\|_C^2$. Similarly, we have the right-hand inequality

$$\|(x,\lambda)\|_C^2=\left\langle C\begin{pmatrix}x\\ \lambda\end{pmatrix},\begin{pmatrix}x\\ \lambda\end{pmatrix}\right\rangle\le(1+\tau\eta)\|(x,\lambda)\|^2.$$

□

In order to use Theorem 3.2 to study the asymptotic behavior of the trajectories of the associated differential equation, we consider $M(x,\lambda)=\tau(\nabla_x\mathscr{L}(x,\lambda),-\nabla_\lambda\mathscr{L}(x,\lambda))$. Thus system (S) takes the form C. Now, we use Theorem 3.2 to study the asymptotic behavior of the trajectories of the associate differential equation. $M(x,\lambda)=\tau(\nabla_x\mathscr{L}(x,\lambda),-\nabla_\lambda\mathscr{L}(x,\lambda))$

$$C(\dot{u}(t))+Mu(t)+\varepsilon(t)u(t)=0,\ \text{where } u(t):=(x(t),\lambda(t)),$$

which means

$$\begin{cases}\dot{x}(t)-\tau A^*\dot{\lambda}(t)+(\nabla f(x(t))+A^*\lambda(t))+\varepsilon(t)x(t)=0_{\mathscr{X}},\\ \dot{\lambda}(t)-\ \tau A\dot{x}(t)-(Ax(t)-\nabla g^*(\lambda(t)))\ +\varepsilon(t)\lambda(t)=0_{\mathscr{Y}}.\end{cases}\tag{SP_0}$$

According to Theorem 3.2, we confirm the following convergence results.

Theorem 3.3. *Let $(x(t),\lambda(t))$ be a trajectory solution of (SP_0). Let $\phi(\delta)=\min\left(\frac{\delta}{2},1\right)$. Then, the following statements are true:*

(1) *The trajectory solution* $(x(t),\lambda(t))$ *converges strongly to the saddle point* $(\bar{x},\bar{\lambda})$ *of the Lagrangian* $\mathscr{L}$*; moreover,* $\bar{x}$ *(resp.* $\bar{\lambda}$*) is the projection of the origin in* $\mathscr{X}$ *(resp.* $\mathscr{Y}$*) on* S_f *the set of optimal solutions of (*$\mathscr{P}$*) (resp.* L_f *the set of associate Lagrange multipliers).*

(2) *For* $\delta \neq 2$ *and* t *large enough,*

$$i)\quad \left\|\nabla_x \mathscr{L}(x(t),\lambda(t))\right\| = \mathscr{O}\left(\frac{1}{t^{\phi(\delta)}}\right); \qquad ii)\quad \left\|\nabla_\lambda \mathscr{L}(x(t),\lambda(t))\right\| = \mathscr{O}\left(\frac{1}{t^{\phi(\delta)}}\right);$$

$$iii)\quad \left\|(\dot{x}(t),\dot{\lambda}(t))\right\| = \mathscr{O}\left(\frac{1}{t^{\phi(\delta)}}\right); \qquad iv)\quad \mathscr{L}\big(x(t),\bar{\lambda}\big) - \mathscr{L}\big(\bar{x},\lambda(t)\big) = \mathscr{O}\left(\frac{1}{t^{\phi(\delta)}}\right);$$

$$v)\quad \|Ax(t) - \nabla g^*(\lambda(t))\| = \mathscr{O}\left(\frac{1}{t^{\phi(\delta)}}\right);$$

$$vi)\quad f(x(t)) + g^*(\lambda(t)) - \big(f(\bar{x}) + g^*(\bar{\lambda})\big) \;+\; \langle Ax(t),\bar{\lambda}\rangle - \langle A\bar{x},\lambda(t)\rangle = \mathscr{O}\left(\frac{1}{t^{\phi(\delta)}}\right).$$

(3) *For* $\delta = 2$ *and* t *large enough,*

$$i)\quad \left\|\nabla_x \mathscr{L}(x(t),\lambda(t))\right\| = \mathscr{O}\left(\frac{\sqrt{\ln(t)}}{t}\right); \qquad ii)\quad \left\|\nabla_\lambda \mathscr{L}(x(t),\lambda(t))\right\| = \mathscr{O}\left(\frac{\sqrt{\ln(t)}}{t}\right);$$

$$iii)\quad \left\|(\dot{x}(t),\dot{\lambda}(t))\right\| = \mathscr{O}\left(\frac{\sqrt{\ln(t)}}{t}\right); \qquad iv)\quad \mathscr{L}(x(t),\bar{\lambda}) - \mathscr{L}(\bar{x},\lambda(t)) = \mathscr{O}\left(\frac{\sqrt{\ln(t)}}{t}\right);$$

$$v)\quad \|Ax(t) - \nabla g^*(\lambda(t))\| = \mathscr{O}\left(\frac{\sqrt{\ln(t)}}{t}\right);$$

$$vi)\quad f(x(t)) + g^*(\lambda(t)) - \big(f(\bar{x}) + g^*(\bar{\lambda})\big) \;+\; \langle Ax(t),\bar{\lambda}\rangle - \langle A\bar{x},\lambda(t)\rangle = \mathscr{O}\left(\frac{\sqrt{\ln(t)}}{t^{\phi(\delta)}}\right).$$

Proof. Since the monotone operator $M(x,\lambda)$ satisfies all the conditions of Theorem 3.2, $\mathscr{S} = S_f \times L_f$ and the saddle points of the Lagrangian $\mathscr{L}$ are nothing other than the zeros of the operator M, we automatically obtain assertion (1) and the first two assertion. Also, we have

$$\|M(x(t),\lambda(t))\| = \mathscr{O}\left(\frac{\sqrt{\ln(t)}}{t}\right) \text{ if } \delta = 2 \text{ and } \|M(x(t),\lambda(t))\| = \mathscr{O}\left(\frac{1}{t^{\phi(\alpha)}}\right) \text{ if } \delta \neq 2. \quad (3.13)$$

To prove assertions (2) and (3), we primarily go back to the convex-concavity of $\mathscr{L}$ to have $M(x(t),\lambda(t)),(x(t),\lambda(t)) = (\nabla_x \mathscr{L}(x(t),\lambda(t)), -\nabla_\lambda \mathscr{L}(x(t),\lambda(t)))$, which ensures

$$\mathscr{L}\big(x(t),\lambda(t)\big) - \mathscr{L}\big(\bar{x},\lambda(t)\big) \leq \big\langle \nabla_x \mathscr{L}\big(x(t),\lambda(t)\big)\,,\, x(t) - \bar{x}\big\rangle$$
$$-\mathscr{L}\big(x(t),\lambda(t)\big) + \mathscr{L}\big(x(t),\bar{\lambda}(t)\big) \leq \big\langle -\nabla_\lambda \mathscr{L}(x(t),\lambda(t))\,,\, \lambda(t)\bar{\lambda}\big\rangle.$$

Summing these two inequalities and using $(\bar{x},\bar{\lambda}) \in \mathscr{S}$, we obtain

$$\begin{aligned} 0 \leqslant \mathscr{L}(x(t),\bar{\lambda}) - \mathscr{L}(\bar{x},\lambda(t)) &\leq \big\langle M(x(t),\lambda(t)),(x(t),\lambda(t)) - (\bar{x},\bar{\lambda})\big\rangle \\ &\leq \|\mathrm{M}(x(t),\lambda(t))\| \cdot \|(x(t),\lambda(t)) - (\bar{x},\bar{\lambda})\|. \end{aligned}$$

Since $(x(t),\lambda(t))$ is bounded, by using (3.13), we get the convergence rate of values for the Lagrangian $\mathscr{L}$. Using (3.13), we conclude

$$\|Ax(t) - \nabla g^*(\lambda(t))\| = \|\nabla_\lambda \mathscr{L}(x(t),\lambda(t))\| \leqslant \|M(x(t),\lambda(t))\|.$$

Also, we have

$$f(x(t))-f(\bar{x})=\mathscr{L}(x(t),\bar{\lambda})-\mathscr{L}(\bar{x},\lambda(t))-\langle Ax(t),\bar{\lambda}\rangle+\langle A\bar{x},\lambda\rangle-g^*(\lambda)+g^*(\bar{\lambda}).$$

Then, the convergence rate of values for the Lagrangian, ensures the desired estimations. So, according to (3.13) and Theorem 3.2, we obtain the remaining assertion. □

3.3. **Application to a linearly-constrained minimization problem.** When supposing $g=\mathbf{1}_{\{b\}}$, where $b\in\mathscr{X}$, problem ($\mathscr{P}$) reduces to the linearly-constrained minimization problem:

$$\min_{x\in\mathscr{X}} f(x) \quad \text{subject to} \quad Ax-b=0. \qquad (\mathscr{P}_c)$$

Note that the conjugate function g^* is always convex and lower semicontinuous, and its differentiability specifically requires the strict convexity of g, which is not satisfied for $g=\mathbf{1}_{\{b\}}$. This requires us to return to Theorem 3.2 to derive the related convergence assertions. The associate Lagrangian for ($\mathscr{P}_c$) is defined by $\mathscr{L}(x,\lambda)=f(x)+\langle\lambda,Ax-b\rangle$. Let $M:\mathscr{H}:=\mathscr{X}\times\mathscr{Y}\to\mathscr{H}$ be defined by

$$M(x,\lambda)=(\nabla_x\mathscr{L}(x,\lambda)\ ,\ -\nabla_\lambda\mathscr{L}(x,\lambda))=(\nabla f(x)+A^*\lambda\ ,\ b-Ax)$$

for all $(x,\lambda)\in\mathscr{H}$. Then the saddle points of the Lagrangian $\mathscr{L}$ are characterized by finding the zeros of the operator M, and system (S_0) becomes

$$\begin{cases}\dot{x}(t)-\tau A^*\dot{\lambda}(t)+(\nabla f(x(t))+A^*\lambda(t))+\varepsilon(t)x(t)=0_{\mathscr{X}},\\ \dot{\lambda}(t)-\ \tau A\dot{x}(t)\ -\ (Ax(t)-b)\ +\ \varepsilon(t)\lambda(t)=0_{\mathscr{Y}}.\end{cases} \tag{3.14}$$

Theorem 3.4. *Let $(x(t),\lambda(t))$ be a trajectory solution of (3.14). Then, in addition to the assertions in Theorem 3.3, we have the following convergence rates for the values and the constraints:*

(1) *For $\delta\neq 2$, for t large enough, we have*

$$i)\quad \|Ax(t)-b\|=\mathscr{O}\left(\frac{1}{t^{\phi(\delta)}}\right);\quad ii)\quad |f(x(t))-f(\bar{x})|=\mathscr{O}\left(\frac{1}{t^{\phi(\delta)}}\right).$$

(2) *For $\delta=2$, for t large enough, we have*

$$i)\quad \|Ax(t)-b\|=\mathscr{O}\left(\frac{\sqrt{\ln(t)}}{t}\right);\quad ii)\quad |f(x(t))-f(\bar{x})|=\mathscr{O}\left(\frac{\sqrt{\ln(t)}}{t}\right).$$

Proof. Since the monotone operator $M(x,\lambda)$ satisfies all the conditions of Theorem 3.2, we get all the assertions in Theorem 3.3 similarly.

For the convergence rates of the constraints, we only use (3.13) and

$$\|Ax(t)-b\|=\|\nabla_\lambda\mathscr{L}(x(t),\lambda(t))\|\leqslant\|M(x(t),\lambda(t))\|.$$

Also, for the convergence rates of the values, we use

$$\begin{aligned} f(x(t))-f(\bar{x}) &= \mathscr{L}(x(t),\bar{\lambda})-\mathscr{L}(\bar{x},\lambda(t))-\left\langle\bar{\lambda},Ax(t)-b\right\rangle\\ &\leqslant \mathscr{L}(x(t),\bar{\lambda})-\mathscr{L}(\bar{x},\lambda(t))+\|\bar{\lambda}\|\|Ax(t)-b\|\\ &\leqslant \mathscr{L}(x(t),\bar{\lambda})-\mathscr{L}(\bar{x},\lambda(t))+\|\bar{\lambda}\|\|M(x(t),\lambda(t))\|.\end{aligned}$$

□

4. The Associate Implicit Algorithms

4.1. **Proximal algorithms to attain zeros of monotone operators.** Here, we suppose $M : \mathscr{H} \to 2^{\mathscr{H}}$ a set-valued maximally monotone operator whose domaine is nonempty. As a backward discrete version of (S_0), consider the proximal algorithm:

$$\begin{cases} C(u_{k+1}-u_k)+\xi_{k+1}+\varepsilon_{k+1}u_{k+1}=0, \\ \xi_{k+1}\in Mu_{k+1} \text{ and } u_0\in\mathscr{H}. \end{cases} \tag{S_k}$$

Set $\bar{u} := \mathrm{proj}_{M^{-1}(0)}(0)$. As the first result, we show that the iterations $\{u_k\}$, generated by (S_k), converge strongly to $\bar{u}$, provides that each weak cluster point of $\{u_k\}$ belongs to S.

Proposition 4.1. *Suppose that $\{u_k\}$ is bounded. Then the following assertions are equivalent:*

(a) *Each weak cluster point of $\{u_k\}$ belongs to $S = M^{-1}(0)$.*
(b) $\liminf_{k\to+\infty} \|u_k\| \geq \|\bar{u}\|$.
(c) *$u_k \to \bar{u}$ strongly as $k \to +\infty$.*

Proof. The implication $(c) \Rightarrow (a)$ is immediate, while $(a) \Rightarrow (b)$ follows from the weak lower-semicontinuity of the norm. For the last implication $(b) \Rightarrow (c)$, we consider $s_k := \frac{1}{2}\|u_k-\bar{u}\|_C^2$. Since $\|\cdot\|_C$ defines a norm in $\mathscr{H}$, then

$$\begin{aligned} \|u_{k+1}-\bar{u}\|_C^2-\|u_k-\bar{u}\|_C^2 &= -\|u_k-u_{k+1}\|_C^2-2\langle u_k-u_{k+1},u_{k+1}-\bar{u}\rangle_C \\ &\leq 2\langle u_{k+1}-u_k,u_{k+1}-\bar{u}\rangle_C. \end{aligned}$$

Thus

$$s_{k+1}-s_k \leqslant \langle C(u_{k+1}-u_k),u_{k+1}-\bar{u}\rangle. \tag{4.1}$$

Using iteration (S_k), we have

$$\begin{aligned} \langle C(u_{k+1}-u_k),u_{k+1}-\bar{u}\rangle &= -\langle \xi_{k+1}+\varepsilon_{k+1}u_{k+1},u_{k+1}-\bar{u}\rangle \\ &= -\langle \xi_{k+1},u_{k+1}-\bar{u}\rangle-\varepsilon_{k+1}\langle u_{k+1},u_{k+1}-\bar{u}\rangle. \end{aligned}$$

By monotonicity of M, $\xi_{k+1}\in Mu_{k+1}$ and $0\in M\bar{u}$ ensure $\langle -\xi_{k+1},u_{k+1}-\bar{u}\rangle \leqslant 0$. Thus

$$\langle C(u_{k+1}-u_k),u_{k+1}-\bar{u}\rangle \leq -\varepsilon_{k+1}\langle u_{k+1},u_{k+1}-\bar{u}\rangle.$$

Coming back to (4.1) and (2.1), we see that

$$\begin{aligned} s_{k+1}-s_k &\leq -\varepsilon_{k+1}\langle u_{k+1}\ ,\ u_{k+1}-\bar{u}\rangle \\ &\leq \frac{\varepsilon_{k+1}}{2}\left[\|\bar{u}\|^2-\|u_{k+1}\|^2-\|u_{k+1}-\bar{u}\|^2\right] \\ &\leq \frac{\varepsilon_{k+1}}{2}\left[\|\bar{u}\|^2-\|u_{k+1}\|^2-\alpha^{-1}\|u_{k+1}-\bar{u}\|_C^2\right], \end{aligned}$$

which gives

$$(s_{k+1}-s_k)+\alpha^{-1}\varepsilon_{k+1}s_{k+1} \leqslant \alpha^{-1}\varepsilon_{k+1}\frac{\alpha}{2}\left[\|\bar{u}\|^2-\|u_{k+1}\|^2\right].$$

Setting $h_k := \frac{\alpha}{2}\left[\|\bar{u}\|^2-\|u_k\|^2\right]$ and $\xi_k := \alpha^{-1}\varepsilon_{k+1}$ in Lemma A.3, we obtain that $\{s_k\}$ is bounded, so is $\{u_k\}$ and

$$\limsup_{k\to+\infty}\|u_k-\bar{u}\|_C^2 = 2\limsup_{k\to+\infty}s_k \leq 2\limsup_{k\to+\infty}h_k = \alpha\limsup_{k\to+\infty}\left[\|\bar{u}\|^2-\|u_k\|^2\right].$$

By (b), we have $\limsup_{k\to+\infty}\left[\|\bar{u}\|^2-\|u_k\|^2\right] \leq 0$, so $\|u_k-\bar{u}\|_C \to 0$. This ends the proof. ☐

Consider now the energy sequence:

$$E_k := \frac{1}{2}\|u_k - u_{k-1}\|_C^2 = \frac{1}{2}\|C(u_k - u_{k-1})\|_{C^{-1}}^2 = \frac{1}{2}\|\xi_k + \varepsilon_k u_k\|_{C^{-1}}^2.$$

We then have the following iterative sequential control.

Theorem 4.1. *Let $\{u_k\}$ be the sequence generated by (S_k), and suppose that $\{\varepsilon_k\}$ satisfies*

$$\sum_{k=k_0}^{+\infty} \frac{\beta\varepsilon_k}{1+\beta\varepsilon_k} = +\infty \quad \textit{and} \quad \lim_{k\to+\infty} \frac{(\varepsilon_{k+1}-\varepsilon_k)^2}{\beta{\varepsilon_{k+1}}^2} = 0. \tag{4.2}$$

Then the following statements hold:

(1) *the generated sequence $\{u_k\}$ is bounded;*
(2) *it holds the following estimation:*

$$E_{k+1} \leq \frac{1}{1+\beta\varepsilon_{k+1}}E_k + \frac{(\varepsilon_{k+1}-\varepsilon_k)^2}{\varepsilon_{k+1}(1+\beta\varepsilon_{k+1})}R_0; \tag{4.3}$$

(3) *the sequence $\{M^o u_k\}$, generated for each k by the minimum norm element of Mu_k, converges strongly to the origin in $\mathscr{H}$;*
(4) *$\{u_k\}$ converges strongly to the minimum norm solution $\bar{u}$ in $M^{-1}0$.*

Proof. (1) Fix $u \in \mathscr{S}$ and define the nonnegative sequence $\{w_k\}$ whose general term is $w_k := \frac{1}{2}\|u_k - u\|_C^2$. Then, we have

$$\langle C(u_{k+1}-u_k), u_{k+1}-u\rangle = \frac{1}{2}(\|u_{k+1}-u_k\|_C^2 + \|u_{k+1}-u\|_C^2 - \|u_k - u\|_C^2).$$

Thus,

$$w_{k+1} - w_k = \langle C(u_{k+1}-u_k), u_{k+1}-u\rangle - \frac{1}{2}\|u_{k+1}-u_k\|_C^2. \tag{4.4}$$

Since $C(u_{k+1}-u_k) = -Mu_{k+1} - \varepsilon_{k+1}u_{k+1}$, then (4.4) becomes

$$w_{k+1} - w_k = -\langle \xi_{k+1}, u_{k+1}-u\rangle - \varepsilon_{k+1}\langle u_{k+1}, u_{k+1}-u\rangle - \frac{1}{2}\|u_{k+1}-u_k\|_C^2,$$

which by monotonicity of M and $u \in \mathscr{S}$ implies $\langle \xi_{k+1}, u_{k+1}-u\rangle \geq 0$. Thus

$$w_{k+1} - w_k \leq -\varepsilon_{k+1}\langle u_{k+1}, u_{k+1}-u\rangle - \frac{1}{2}\|u_{k+1}-u_k\|_C^2. \tag{4.5}$$

Note that $-\varepsilon_{k+1}\langle u_{k+1}, u_{k+1}-u\rangle = -\varepsilon_k\|u_{k+1}-u\|^2 + \varepsilon_k\langle u, u_{k+1}-u\rangle$. Using the fact that

$$\varepsilon_k\langle u, u_{k+1}-u\rangle \leq \frac{\varepsilon_k}{2}\|u_{k+1}-u\|^2 + \frac{\varepsilon_k}{2}\|u\|^2,$$

we obtain

$$-\varepsilon_{k+1}\langle u_{k+1}, u_{k+1}-u\rangle \leq -\frac{\varepsilon_k}{2}\|u_{k+1}-u\|^2 + \frac{\varepsilon_k}{2}\|u\|^2.$$

Combining the last inequality with (4.5), we have

$$w_{k+1} - w_k \leq -\frac{\varepsilon_k}{2}\|u_{k+1}-u\|^2 + \frac{\varepsilon_k}{2}\|u\|^2.$$

According to (2.1), we have

$$w_{k+1} - w_k \leq -\frac{\varepsilon_k}{2\alpha}\|u_{k+1}-u\|_C^2 + \frac{\varepsilon_k}{2}\|u\|^2 = -\frac{\varepsilon_k}{\alpha}w_{k+1} + \frac{\varepsilon_k}{2}\|u\|^2.$$

Then, setting $\alpha_k = \frac{\varepsilon_k}{\alpha+\varepsilon_k} \in]0,1[$, this leads to

$$w_{k+1} \leq (1-\alpha_{k+1})w_k + \alpha_{k+1}\frac{\alpha\|u\|^2}{2}.$$

Basing on the above relation and reasoning by induction, we prove that

$$w_k \leq \max\left(w_{k_0}, \frac{\alpha}{2}\|u\|^2\right), \ \forall k \geq k_0.$$

This leads to the existence of a nonnegative real constant R such that $\|u_k\|_C \leq R$. According to (2.1), we conclude that $\{u_k\}$ is bounded.

(2) Let us remind that, for all $a, b \in \mathscr{H}$, $\frac{1}{2}\left(\|b\|^2_{C^{-1}} - \|a\|^2_{C^{-1}}\right) \leq \langle b-a, C^{-1}b\rangle$, which obtains

$$E_{k+1} - E_k \leqslant \langle \xi_{k+1} - \xi_k + \varepsilon_{k+1}u_{k+1} - \varepsilon_k u_k \ , \ C^{-1}(\xi_{k+1} + \varepsilon_{k+1}u_{k+1})\rangle.$$

Using the iteration $\xi_{k+1} + \varepsilon_{k+1}u_{k+1} = -C(u_{k+1} - u_k)$, we have

$$E_{k+1} - E_k \leqslant -\langle \xi_{k+1} - \xi_k, u_{k+1} - u_k\rangle - \langle \varepsilon_{k+1}u_{k+1} - \varepsilon_k u_k, u_{k+1} - u_k\rangle.$$

Since M is monotone, we deduce

$$\begin{aligned}
E_{k+1} - E_k &\leqslant -\langle \varepsilon_{k+1}u_{k+1} - \varepsilon_k u_k, u_{k+1} - u_k\rangle \\
&= -\varepsilon_{k+1}\|u_{k+1} - u_k\|^2 - (\varepsilon_{k+1} - \varepsilon_k)\langle u_k, u_{k+1} - u_k\rangle \\
&\leqslant -\varepsilon_{k+1}\|u_{k+1} - u_k\|^2 + \frac{(\varepsilon_{k+1} - \varepsilon_k)^2}{2\varepsilon_{k+1}}\|u_k\|^2 + \frac{\varepsilon_{k+1}}{2}\|u_{k+1} - u_k\|^2 \\
&= \frac{(\varepsilon_{k+1} - \varepsilon_k)^2}{2\varepsilon_{k+1}}\|u_k\|^2 - \frac{\varepsilon_{k+1}}{2}\|u_{k+1} - u_k\|^2.
\end{aligned}$$

Using (2.1) , (2.2), and the boundedness of $\{u_k\}$, we arrive at

$$E_{k+1} - E_k \leqslant \frac{(\varepsilon_{k+1} - \varepsilon_k)^2}{2\varepsilon_{k+1}}R_0 - \underbrace{\frac{\beta\varepsilon_{k+1}}{2}\|u_{k+1} - u_k\|^2_{C^{-1}}}_{\beta\varepsilon_{k+1}E_{k+1}},$$

where R_0 is a positive constant. Rearranging the previous inequality, we achieve (4.3).

(3) Setting

$$q_k := \frac{1}{1+\beta\varepsilon_{k+1}} \text{ and } \beta_k := \frac{(\varepsilon_{k+1} - \varepsilon_k)^2}{\varepsilon_{k+1}(1+\beta\varepsilon_{k+1})}R_0,$$

we have

$$1 - q_k = \frac{\beta\varepsilon_{k+1}}{1+\beta\varepsilon_{k+1}} \text{ and } \frac{\beta_k}{1-q_k} = \frac{(\varepsilon_{k+1} - \varepsilon_k)^2}{\beta\varepsilon_{k+1}{}^2}.$$

According to (4.2), sequences $\{q_k\}$ and $\{\beta_k\}$ satisfy all the conditions in Lemma A.1. Then, we conclude that $E_k \to 0$ as $k \to +\infty$. Using again (2.1), we have $\|\xi_k + \varepsilon_k u_k\| \to 0$ as $k \to +\infty$. Then, the boundedness of $\{u_k\}$ and $\varepsilon_k \to 0$ ensure $\lim\limits_{k\to+\infty}\|\xi_k\| = 0$. Therefore, $\lim_{k\to+\infty}\|M^o u_k\| = 0$.

(4) To prove strong convergence of $\{u_k\}$ to the minimum norm solution $\bar{u} \in M^{-1}0$, we appeal Proposition 4.1. So, we only need to justify that every weak cluster point of $\{u_k\}$, for $k \to \infty$, belongs to $M^{-1}0$. Let us consider $\nu_k \nearrow +\infty$ such that $\{u_{\nu_k}\}$ converges weakly to u^*. Using (3), we see that sequence $\{y_{\nu_k} := M^o u_{\nu_k} \in Mu_{\nu_k}\}$ strongly converges to the origin in $\mathscr{H}$. Knowing that the graph of a maximally monotone operator is weak×strong sequentially closed in $\mathscr{H} \times \mathscr{H}$ (see [32, Proposition 20.38]), we ensure that $0 = Mu^*$. This ends the proof. □

4.2. **The case : $\varepsilon_k = \frac{d}{\beta(k-d-1)}, d > 0$.**

Theorem 4.2. *Let $\{u_k\}$ be the sequence generated by (SP_0), where $\varepsilon_k = \frac{d}{\beta(k-d-1)}$ for some $d > 0$. Then,*

(1) $\lim_{k\to+\infty} \|M^o u_k\| = 0$;
(2) *there exist $p \in (1,2)$ and $R_0 > 0$ such that*
  i) *if* $d > 2$, *then* $\|M^o u_k\| = \mathscr{O}\left(\sqrt{E_k}\right) = \mathscr{O}\left(k^{-1}\right)$;
  ii) *if* $d = 2$, *then* $\|M^o u_k\| = \mathscr{O}\left(\sqrt{E_k}\right) = \mathscr{O}\left(k^{-1}(\ln k)^{1/2}\right)$;
  iii) *if* $0 < d < 2$, *then* $\|M^o u_k\| = \mathscr{O}\left(\sqrt{E_k}\right) = \mathscr{O}\left(k^{-d/2}\right)$.

Proof. (1) Since $\varepsilon_k = \frac{d}{\beta(k-d-1)}$, a simple calculation ensures that

$$\frac{1}{1+\beta\varepsilon_{k+1}} = 1 - \frac{d}{k} \quad \text{and} \quad \frac{(\varepsilon_{k+1}-\varepsilon_k)^2}{\varepsilon_{k+1}(1+\beta\varepsilon_{k+1})} = \frac{d}{\beta k^3(1-(d+1)k^{-2})^2}$$

Using $1 - \frac{d+1}{k^2} \geq \frac{1}{2}$, for $k \geq k_0$ large enough, and (4.3) in Theorem 4.1, we obtain the existence of $k_1 > k_0$ such that

$$E_{k+1} \leq \left(1 - \frac{d}{k}\right) E_k + \frac{4dR_0}{\beta k^3} \quad \forall k \geq k_1.$$

Using Lemma A.2 with $p = 2$, we conclude that

- if $d > 2$, then $E_k = \mathscr{O}\left(k^{-2}\right)$;
- if $d = 2$, then $E_k = \mathscr{O}\left(k^{-2}\ln(k)\right)$;
- if $d < 2$, then $E_k = \mathscr{O}\left(k^{-d}\right)$.

Also, we have

$$\begin{aligned} \|M^o u_k\| &\leq \|\xi_k\| \leq \|Mu_k + \varepsilon_k u_k\| + \varepsilon_k \|u_k\| \\ &\leq \frac{1}{\beta} \|\xi_k + \varepsilon_k u_k\|_C + \varepsilon_k \|u_k\| \\ &= \frac{\sqrt{2}}{\beta}\sqrt{E_k} + \frac{d}{\beta(k-d-1)} \|u_k\|. \end{aligned}$$

Since $\{u_k\}$ is bounded and the above rates of convergence for $\{E_k\}$, we deduce all the assertions in this theorem. □

4.3. **Adequate proximal algorithm for composite convex minimization.** In this section, we apply the setting of the next subsection to composite convex optimization problems. Moreover, we attain convergence rates for different concepts of values for Lagrangian saddle functions and the constraints. For $f : \mathscr{X} \to \mathbb{R} \cup \{+\infty\}$, $g : \mathscr{Y} \to \mathbb{R} \cup \{+\infty\}$ and $A : \mathscr{X} \to \mathscr{Y}$, consider the composite convex minimization problem

$$\min_{x\in\mathscr{X}} f(x) + g(Ax), \tag{$\mathscr{P}$}$$

Using the conditions (**H**$_0$), composite minimization problem ($\mathscr{P}$) can be equivalently reformulated as the Lagrangian function $\mathscr{L}(x,\lambda) = f(x) + \langle \lambda, Ax\rangle - g^*(\lambda)$ as the saddle point problem

$\mathscr{L}(\bar{x},\lambda) \le \mathscr{L}(\bar{x},\bar{\lambda}) \le \mathscr{L}(x,\bar{\lambda})$, which is equivalent to the corresponding optimality conditions

$$\begin{cases} \partial_x \mathscr{L}(\bar{x},\bar{\lambda}) \ni 0, \\ \partial_\lambda \mathscr{L}(\bar{x},\bar{\lambda}) \ni 0, \end{cases} \iff \begin{cases} \partial f(\bar{x}) + A^*\bar{\lambda} \ni 0, \\ A\bar{x} - \partial g^*(\bar{\lambda}) \ni 0. \end{cases}$$

We associate to ($\mathscr{P}$) the continuous dynamic system (S_0), where, for $u(t) = (x(t),\lambda(t))$, $Mu := \tau(\partial_x \mathscr{L}(x,\lambda), -\partial_\lambda \mathscr{L}(x,\lambda))$ for every $u = (x,\lambda) \in \mathscr{H} = \mathscr{X} \times \mathscr{Y}$, and $C = \begin{pmatrix} I_{\mathscr{X}} & -\tau A^* \\ -\tau A & I_{\mathscr{Y}} \end{pmatrix}$, where τ satisfying $0 < \tau < 1/\|A\|$ (see Lemma 3.1). We note that

$$M^o(x,\lambda) = \tau\big(\partial_x^o \mathscr{L}(x,\lambda), -\partial_\lambda \mathscr{L}^o(x,\lambda)\big),$$

where $\partial_x^o \mathscr{L}(x,\lambda)$ and $\partial_\lambda^o \mathscr{L}(x,\lambda)$ are the minmum norm elements of $\partial \mathscr{L}(\cdot,\lambda)(x)$ and $\partial \mathscr{L}(x,\cdot)(\lambda)$, respectively.

The corresponding optimization algorithm associated to (S_k) satisfies the following scheme

$$\begin{cases} x_{k+1} - x_k - \tau A^*(\lambda_{k+1} - \lambda_k) + \tau \partial_x \mathscr{L}(x_{k+1},\lambda_{k+1}) + \varepsilon_{k+1} x_{k+1} \ni 0, \\ \lambda_{k+1} - \lambda_k - \tau A(x_{k+1} - x_k) - \tau \partial_\lambda \mathscr{L}(x_{k+1},\lambda_{k+1}) + \varepsilon_{k+1} \lambda_{k+1} \ni 0. \end{cases} \tag{SP_k}$$

This last reformulation can be seen as a discretization associated to continuous dynamic system (SP_0).

According to Theorem 4.1, we derive the following first convergence result.

Theorem 4.3. *Let $\{(x_k,\lambda_k)\}$ be the sequence generated by (SP_k) and suppose that the regularization sequence ε_k satisfies (4.2). Then,*

(1) *$\{(x_k,\lambda_k)\}$ converges strongly to $(\bar{x},\bar{\lambda})$, where $\bar{x}$ is the projection of the origin on S_f and $\bar{\lambda}$ is the projection of the origin on the set of associated Lagrange multipliers;*
(2) *$\{(\partial_x^o \mathscr{L}(x_k,\lambda_k), -\partial_\lambda^o \mathscr{L}(x_k,\lambda_k))\}$ converges strongly to the origin in $\mathscr{X} \times \mathscr{Y}$.*

Proof. These follow immediately from the assertions (3) and (4) of Theorem 4.1. We only point out that the projection of the origin onto $M^{-1}0 = S_f \times L_f$ is nothing else than the pair $(\bar{x},\bar{\lambda})$, where $\bar{x}$ and $\bar{\lambda}$ are respectively the projections of the origin onto S_f and L_f, the set of all Lagrange multipliers associated to ($\mathscr{P}$). $\square$

Now, we choose $\varepsilon_k = \frac{d}{(1-\eta)(k-d-1)}$, where $d > 0$. Then, based on Theorem 4.2, we obtain, in addition to the strong convergences predicted in the previous theorem, the following rates of convergence.

Theorem 4.4. *Let $\{(x_k,\lambda_k)\}$ be the sequence generated by (SP_k) and suppose that $\{\varepsilon_k\}$ satisfies (4.2). Then, the following rates of convergence hold:*

- *If $d > 2$, then, for t large enough,*
  (i) $\|\big(\partial_x^o \mathscr{L}(x_k,\lambda_k), -\partial_\lambda^o \mathscr{L}(x_k,\lambda_k)\big)\| = \mathscr{O}\left(k^{-1}\right)$,
  (ii) $\mathscr{L}(x_k,\bar{\lambda}) - \mathscr{L}(\bar{x},\lambda_k) = \mathscr{O}\left(k^{-1}\right)$,
  (iii) $\|Ax_k - \partial^o g^*(\lambda_k)\| = \mathscr{O}\left(k^{-1}\right)$,
  (iv) $f(x_k) + g^*(\lambda_k) - \big(f(\bar{x}) + g^*(\bar{\lambda})\big) + \langle Ax_k, \bar{\lambda}\rangle - \langle A\bar{x}, \lambda_k\rangle = \mathscr{O}\left(k^{-1}\right)$.
- *If $d = 2$, then, for t large enough,*
  (i) $\|\big(\partial_x^o \mathscr{L}(x_k,\lambda_k), -\partial_\lambda^o \mathscr{L}(x_k,\lambda_k)\big)\| = \mathscr{O}\left(k^{-1}(\ln k)^{1/2}\right)$,
  (ii) $\mathscr{L}(x_k,\bar{\lambda}) - \mathscr{L}(\bar{x},\lambda_k)| = \mathscr{O}\left(k^{-1/2}(\ln k)^{1/4}\right)$,

(iii) $\|Ax_k - \partial^o g^*(\lambda_k)\| = \mathscr{O}\left(k^{-1}(\ln k)^{1/2}\right)$,

(iv) $f(x_k) + g^*(\lambda_k) - \left(f(\bar{x}) + g^*(\bar{\lambda})\right) + \langle Ax_k, \bar{\lambda}\rangle - \langle A\bar{x}, \lambda_k\rangle = \mathscr{O}\left(k^{-1}(\ln k)^{1/2}\right)$.

- *If* $0 < d < 2$*, then, for t large enough,*
  (i) $\|\left(\partial_x^o \mathscr{L}(x_k,\lambda_k), -\partial_\lambda^o \mathscr{L}(x_k,\lambda_k)\right)\| = \mathscr{O}\left(k^{-d/2}\right)$,
  (ii) $\mathscr{L}(x_k,\bar{\lambda}) - \mathscr{L}(\bar{x},\lambda_k) = \mathscr{O}\left(k^{-d/2}\right)$,
  (iii) $\|Ax_k - \partial^o g^*(\lambda_k)\| = \mathscr{O}\left(k^{-d/2}\right)$,
  (iv) $f(x_k) + g^*(\lambda_k) - \left(f(\bar{x}) + g^*(\bar{\lambda})\right) + \langle Ax_k, \bar{\lambda}\rangle - \langle A\bar{x}, \lambda_k\rangle = \mathscr{O}\left(k^{-d/2}\right)$.

Proof. Observe that $M^o(x_k,\lambda_k) \in (\partial_x \mathscr{L}(x_k,\lambda_k), -\partial_\lambda \mathscr{L}(x_k,\lambda_k))$. Then, for every $k \geq k_0$, we derive from inequalities in the proof of Theorem 3.3 that

$$0 \leqslant \mathscr{L}(x_k,\bar{\lambda}) - \mathscr{L}(\bar{x},\lambda_k) \leqslant \frac{1}{\tau}\|M^o(x_k,\lambda_k)\| \cdot \|(x_k,\lambda_k) - (\bar{x},\bar{\lambda})\|,$$

$$\|Ax_k - \partial^o g^*(\lambda_k)\| \leqslant \frac{1}{\tau}\|M^o(x_k,\lambda_k)\|,$$

and

$$f(x_k) - f(\bar{x}) = \mathscr{L}(x_k,\bar{\lambda}) - \mathscr{L}(\bar{x},\lambda_k) - \langle Ax_k,\bar{\lambda}\rangle + \langle A\bar{x},\lambda_k\rangle - g^*(\lambda_k) + g^*(\bar{\lambda}).$$

Since the generated sequence $\{(x_k,\lambda_k)\}$ by (SP_k) is bounded, we derive from Theorem 4.2 the desired rates of convergence immeately. □

4.4. **Case of linearly-constrained minimization problem.** We return to problem ($\mathscr{P}_c$), with f not necessarily differentiable, and rewrite it for some $\tau > 0$ as

$$\min_{x\in\mathscr{X}} \tau f(x) \quad \text{subject to} \quad \tau Ax = \tau b.$$

If we take $\tau \in (0, 1/\|A\|)$ (see Lemma 3.1), then (SP_k) becomes

$$\begin{cases} x_{k+1} - x_k \;+\; \tau A^*\lambda_k + \tau\partial f(x_{k+1}) \;+\; \varepsilon_{k+1}x_{k+1} \ni 0, \\ \lambda_{k+1} - \lambda_k - 2\tau Ax_{k+1} + \tau Ax_k + \tau b + \varepsilon_{k+1}\lambda_{k+1} = 0. \end{cases}$$

which is equivalent, for $\sigma_k := \frac{1}{1+\varepsilon_{k+1}}$, to the following Forward-Backward algorithm

$$\begin{cases} x_{k+1} = \operatorname{prox}_{\tau\sigma_k f}\left(\sigma_k(x_k - \tau A^*\lambda_k)\right), \\ \lambda_{k+1} = \sigma_k\left(\lambda_k + 2\tau Ax_{k+1} - \tau(Ax_k + b)\right). \end{cases} \tag{4.6}$$

Theorem 4.5. *Let* $\{(x_k,\lambda_k)\}$ *be the sequence generated by* (4.6) *and suppose that* $\{\varepsilon_k\}$ *satisfies* (4.2)*. Then, in addition to the Lagrangian estimates of Theorem 4.4, we have the following convergence rates for values and constraints:*

- *If* $d > 2$*, then*
  (i) $\|Ax_k - b\| = \mathscr{O}\left(k^{-1}\right)$,
  (ii) $|f(x_k) - f(\bar{x})| = \mathscr{O}\left(k^{-1}\right)$.
- *If* $d = 2$*, then*
  (i) $\|Ax_k - b\| = \mathscr{O}\left(k^{-1}(\ln k)^{1/2}\right)$,
  (ii) $|f(x_k) - f(\bar{x})| = \mathscr{O}\left(k^{-1}(\ln k)^{1/2}\right)$.

- *If* $0 < d < 2$, *then*
  (i) $\|Ax_k - b\| = \mathscr{O}\left(k^{-d/2}\right)$,
  (ii) $|f(x_k) - f(\bar{x})| = \mathscr{O}\left(k^{-d/2}\right)$.

Proof. Recall that $M^o(x_k, \lambda_k) \in (\partial_x \mathscr{L}(x_k, \lambda_k), -\partial_\lambda \mathscr{L}(x_k, \lambda_k))$. Then, for every $k \geq k_0$, we derive from inequalities in the proof of Theorem 3.3 that

$$\begin{aligned}
&0 \leqslant \mathscr{L}(x_k, \bar{\lambda}) - \mathscr{L}(\bar{x}, \lambda_k) \leqslant \|M^o(x_k, \lambda_k)\| \cdot \|(x_k, \lambda_k) - (\bar{x}, \bar{\lambda})\|, \\
&\tau\|Ax_k - b\| \leqslant \|M^o(x_k, \lambda_k)\|, \\
&f(x_k) - f(\bar{x}) \leqslant \mathscr{L}(x_k, \bar{\lambda}) - \mathscr{L}(\bar{x}, \lambda_k) + \|\bar{\lambda}\|\|M^o(x_k, \lambda_k)\| \\
&\qquad \leqslant \Big(\|(x_k, \lambda_k) - (\bar{x}, \bar{\lambda})\| + \|\bar{\lambda}\|\Big)\|M^o(x_k, \lambda_k)\|.
\end{aligned}$$

Since the generated sequence $\{(x_k, \lambda_k)\}$ by (4.6) is bounded, the desired rates of convergence follow from Theorem 4.2 immediately. ☐

5. Examples for Comparison

In this section, we consider numerical examples for the trajectories generated by dynamical system (3.14). Here, the comparison examples are solved numerically with the Runge-Kutta adaptive method in Scilab 6.1.0, and all the numerical tests are run on a Mac Book Pro 2.8 GHz Intel Core i7.

Example 5.1 (Nonstrictly convex function)**.** Consider the differentiable and nonstrictly convex function f defined on $\mathbb{R}^3$ by $f(x) = (x_1 + x_2)^2 + x_3^2$ and the constraint-set:

$$C := \left\{ x \in \mathbb{R}^3 : Ax - b = \frac{1}{10}\begin{pmatrix} x_1 + x_2 + x_3 \\ -x_1 - x_2 + x_3 \end{pmatrix} \right\}.$$

We have $\min_C f = 0$ and $\operatorname{argmin}_C f = \mathbb{R}(1, -1, 0)$. For selected cases for the parameter δ, in Figure 1 we plot the time-dependent evolution of the convergence rate to zero for the values $f(x(t)) - \min f$, the velocity $\|\dot{x}(t)\|$ and the constraints $\|Ax(t) - b\|$, as well as, for the pathways towards the minimum norm solution x^*. First, we obtain an approximate optimal solution $x = [-3,9 \times 10^{-6} \quad 4 \times 10^{-6} \quad 6 \times 10^{-7}]$ and the associated multiplier $\lambda = [7,2 \times 10^{-6} \quad 4,9\ 10^{-6}]$ when $\alpha = 1.2$ and $\delta = 2$. Then, for $\alpha = 1,2$ and $\delta = 50$, the solution $x = [3,28 \times 10^{-115} \quad 3,28 \times 10^{-115} \quad 9,67 \times 10^{-116}]$ and $\lambda = [4,67 \times 10^{-114} \quad -3,48 \times 10^{-114}]$ is much more accurate. The corresponding results in Figure 1 justify a much faster convergence when the parameter δ increases significantly.

Example 5.2 (Quadratic minimization problem subject to linear constraints)**.** In this example, we consider the problem $\min f(x)$ under constraint $Ax = 0$, where $f(x) = \frac{1}{2}\langle Bx, x\rangle$ and $A \in \mathbb{R}^{10\times 20}$ is a random matrix and $B \in \mathbb{R}^{20\times 20}$ is a random positive semidefinite matrix with non-empty kernel (10% of the eigenvalues equal to 0). This example is adapted from [28, Example 8.2] in order to compare the convergence rate of values for a similar constrained minimization problem. To solve the linearly-constrained quadratic minimization problem ($\mathscr{P}_c$), Apidopoulos et al. propose the differential system

$$\begin{pmatrix} \beta e^{-t/\beta}\mathbf{I}_{20} & 0,1A^* \\ -0,1A & \beta e^{-t/\beta}\mathbf{I}_{10} \end{pmatrix}\begin{pmatrix} \dot{x}(t) \\ \dot{\lambda}(t) \end{pmatrix} + \begin{pmatrix} \nabla f(x(t)) + A^*\lambda(t) \\ b - Ax(t) \end{pmatrix} = 0, \tag{5.1}$$

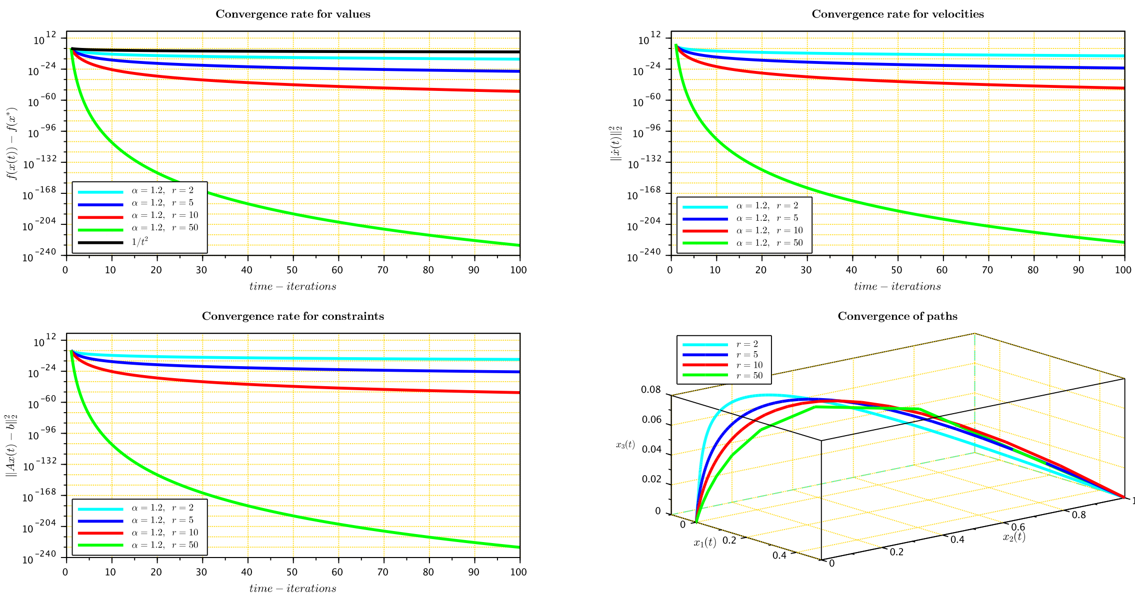

FIGURE 1. Here, for the convex minimization problem $\min_{x\in C} f(x) = (x_1 + x_2)^2 + x_3^2$ where $C = \{x : x_1 + x_2 + x_3 = 0, x_1 + x_2 - x_3 = 0\}$, we schematize the convergence rate to zero for values, the velocity and the constrains, as well as, for the pathways towards the minimum norm solution $x^* = (0,0,0)$. In this numerical experiments we consider the starting points $x_0 = (1/2,1,0)$ and $\lambda_0 = (1,1)$.

where $\mathbf{I}_m$ represents identity operator on $\mathbb{R}^m$. This system generalizes the following Luo's [29] preconditioned triangular one

$$\begin{pmatrix} \alpha e^{-t}\mathbf{I}_{20} & 0 \\ \alpha\gamma A & \gamma e^{-t}\mathbf{I}_{10} \end{pmatrix} \begin{pmatrix} \dot{x}(t) \\ \dot{\lambda}(t) \end{pmatrix} + \begin{pmatrix} \nabla f(x(t)) + A^*\lambda(t) \\ b - Ax(t) \end{pmatrix} = 0. \tag{5.2}$$

In Figure 2, we plot the time-dependent evolution of the convergence rate to zero for values. We note that the convergence rate of the values and the runtime are better for (3.14), when compared to those for (5.1). This may be due to the exponential factors in (5.1), which could require more computation time at each iteration over time. Indeed, at each time iteration, solving (5.1) requires solving the system of linear equations involving the exponential factors, which significantly increases the number of iterations. Whereas in (3.14) the square matrix C is constant, and therefore its inverse only occurs once at the beginning of the resolution of this system. Hereinafter, we specify the resolution times of the systems (3.14), (5.1) and (5.2) in Table 1 for selected value of α, γ, β and r.

Example 5.3 (Proximal algorithm for linear constrained convex minimization)**.** Considering the minimization of the nondifferentiable convex function $f(x) = \frac{1}{2}(x_1 + x_2)^2 + |x_3|$ under the

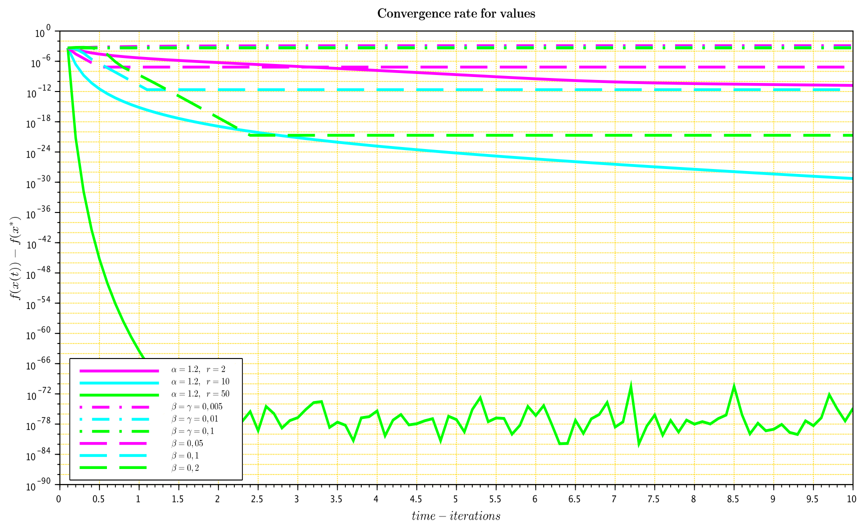

FIGURE 2. For the quadratic minimization problem $f(x) = \frac{1}{2}\langle Bx, x\rangle$ subject to linear constraints $Ax = 0$, where $A \in \mathbb{R}^{10\times 20}$ is a random matrix and $B \in \mathbb{R}^{20\times 20}$ is a random positive semidefinite matrix with non-empty kernel (10% of the eigenvalues equal to 0), we compare values convergence rates for the systems (3.14) and (5.1) by starting from $(x(0), \lambda(0)) = \frac{1}{20}\mathbf{I}(30)$, where $\mathbf{I}(n)$ denotes the 30-dimensional vector whose coordinates are all equal to one.

TABLE 1. In Exemple 5.2, the choice of parameters α, β, γ and r for the systems (3.14), (5.1) and (5.2) are simulated in the following table.

System	Data	Execution time
(5.2)	$\alpha = \gamma = 0,005$	1,03 s
(5.2)	$\alpha = \gamma = 0,001$	0,99 s
(5.2)	$\alpha = \gamma = 0,05$	1,01 s
(5.1)	$\beta = 0,05$	808,86 s
(5.1)	$\beta = 0,1$	754,65 s
(5.1)	$\beta = 0,2$	665,70 s
(3.14)	$r = 2$	1,80 s
(3.14)	$r = 10$	2,30 s
(3.14)	$r = 50$	3,31 s

linear constraints $x_1 + x_2 + x_3 = 0$ and $x_1 + x_2 - x_3 = 0$, we adjust the algorithm (4.6) by setting:

$$\begin{cases} x_{1,k+1} = \frac{\sigma_k}{2\sigma_k+1}\left[(\sigma_k+1)x_{1,k} - \sigma_k x2(n) - \frac{1}{\delta}(\lambda_{1,k} - \lambda_{2,k})\right], \\ x_{2,k+1} = \frac{\sigma_k}{2\sigma_k+1}\left[(\sigma_k+1)x2(n) - \sigma_k x_{1,k} - \frac{1}{\delta}(\lambda_{1,k} - \lambda_{2,k})\right], \\ x_{3,k+1} = \sigma_k\left[\max\left(\left|x_{3,k} - \frac{1}{\delta}(\lambda_{1,k} + \lambda_{2,k})\right| - 1\right)\right].sgn\left(x_{3,k} - \frac{1}{\delta}(\lambda_{1,k} + \lambda_{2,k})\right), \\ \lambda_{1,k+1} = \sigma_k\left[\lambda_{1,k} + \frac{2}{\delta}(x_{1,k+1} + x_{2,k+1} + x_{3,k+1}) - \frac{1}{\delta}(x_{1,k} + x2(n) + x3(n))\right], \\ \lambda_{2,k+1} = \sigma_k\left[\lambda_{2,k} - \frac{2}{\delta}(x_{1,k+1} + x_{2,k+1} - x_{3,k+1}) + \frac{1}{\delta}(x_{1,k} + x_{2,k} - x_{3,k})\right], \end{cases} \tag{5.3}$$

where $\sigma_k = \frac{(1-\eta)k}{(1-\eta)k+d}$ and $sgn(x) = 1$ if $x \geq 0$, $\quad sgn(x) = -1$ if $x < 0$.

Figure 3 compares the convergence rates for the algorithm (5.3) when the initial point is

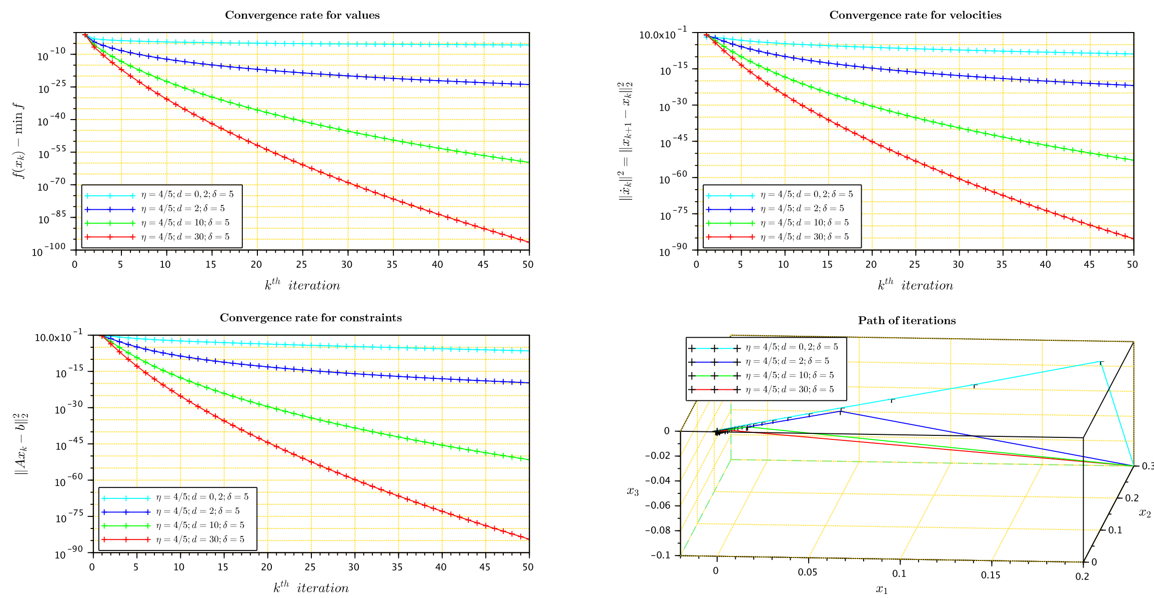

FIGURE 3. Here, we implement the algorithm described in (5.3). Varying the parameter d, we notice a variation in the rate of convergence for the values, velocities and constraints. We observe that the associated proximal algorithm achieved convergence rates similar to the ones obtained in Example 5.1 for the continuous case.

$(x_0, \lambda_0) = (0.2, 0.3, -0.1, -2, 1)$, the data d varies from 0 to 30 and $\eta = 4/5, \delta = 5$ remain fixed. We note that the parameter d plays a crucial role in improving the convergence rates to zero for values, velocities, and constraints. The parameter d thus contributes even more to the improvement of these convergence rates than the suggested assertions in Theorem 4.5.

6. CONCLUSION AND COMMENTS

Our aim in this paper is to find a zero of a maximally monotone operator in a Hilbert space. To this end, we first proposed a continuous aspect that differs little from that proposed by Boţ and Nguyen [7]. We apply a positive linear operator to the factor $\dot{x}(t)$ in order to adapt an appropriate proximal algorithm when applying the proposed system to a composite convex minimization problem and to minimization under linear constraints via a saddle point of an associated Lagrangian. The continuous aspect is treated for a single-valued operator M, in order to be able to illustrate via a suitable discretization the appropriate proximal algorithm and then to perform a similar proof for any convergence rate. The proximal algorithm developed in this paper is more interesting, since strong convergence to a Zero of M with no additional condition on the maximal monotonicity of M does not appear in later works. As mentioned earlier, the study of the case that M is the subdifferential of a convex lsc function has been carefully treated in [26]. We also note that it is the Tykhonov's penalization term $\varepsilon_{k+1} u_{k+1}$ that enables us to achieve this strong convergence, to the minimum norm element in $M^{-1}0$, of u_k generated by the algorithm

$0 \in C(u_{k+1}-u_k)+Mu_{k+1}+\varepsilon_{k+1}u_{k+1}$. In the algorithm $0 \in C(u_{k+1}-u_k)+Mu_{k+1}+\varepsilon_{k+1}u_{k+1}$, it is the Tykhonov perturbation $\varepsilon_{k+1}u_{k+1}$ that enables us to achieve this strong convergence to the minimum norm element in $M^{-1}0$. Furthermore, in solving a minimization problem under linear constraints, we observe that the associated algorithm (1.4) through the introduction of the operator C proposes at each iteration to directly evaluate x_{k+1} based on (x_k,λ_k), and then to deduce λ_{k+1} from (x_k,λ_k) and x_{k+1}. This avoids, as can be seen when implicitly discretizing the system proposed in [7], the need to directly evaluate (x_{k+1},λ_{k+1}) as a fixed point of an appropriate proximal operator, see (1.8). Also, notice that system (S_0) can be seen as a dynamic system associated with the two-level hierarchical problem: $\min \frac{1}{2}\|u\|^2_{\mathscr{H}}$ under constrained $u \in M^{-1}0$, which is equivalent by optimality condition to find $\bar{u} \in \mathscr{H}$ such that

$$\begin{aligned} 0 \in \partial\left(\frac{1}{2}\|\cdot\|^2_{\mathscr{H}}+\delta_{M^{-1}0}\right)\bar{u} \quad &\iff \quad 0 \in \bar{u}+\mathscr{N}_{M^{-1}0}\bar{u} \\ &\iff \quad \bar{u} \in M^{-1}0, \text{ and } \langle v-\bar{u}\,,\,\bar{u}\rangle \geq 0, \forall v \in M^{-1}0 \\ &\iff \quad \bar{u} \in M^{-1}0, \text{ and } \|\bar{u}\|_{\mathscr{H}} \leq \|\bar{v}\|_{\mathscr{H}}, \forall v \in M^{-1}0 \\ &\iff \quad \bar{u} \text{ is the minimal norm element of } M^{-1}0. \end{aligned}$$

So, to reach another solution on $M^{-1}0$, it is possible to choose another differentiable objective function φ, for which the attained solution would be the minimum of φ in $M^{-1}0$.

In this article, when dealing with convex minimization under linear constraints and the composite minimization problem, we restrict ourselves to the case where the operator C, in the associate dynamic system, is symmetric. A detailed study of the different types of systems with symmetric, antisymmetric, and triangular, as well as time-dependent and constant operators C, could be the subject of a future article. We would compare these different types of systems from both theoretical and numerical perspectives, as methods for achieving faster convergence rates. Another more attractive direction is to pass from the first-order system (S_0), to a second-order system in order to obtain an acceleration rate of order t^{-2}. This is the subject of a current work.

Appendix A. Convergence for Nonnegative Real Sequences

Lemma A.1. [33, Lemma 2.2.3] *Let (y_k) be a real sequence satisfying*

$$y_{k+1} \leq (1-\alpha_k)y_k+\beta_k, \text{ with } \beta_k \geq 0,\ 0<\alpha_k \leq 1,\ \sum_{k=k_0}^{\infty}\alpha_k=\infty \text{ and } \frac{\beta_k}{\alpha_k} \to 0.$$

Then, $\limsup_{k\to+\infty} y_k \leq 0$. *In particular, if $y_k \geq 0$, then $y_k \to 0$.*

Lemma A.2. [34, Lemma 1] *Let (y_k) be a nonnegative real sequence satisfying*

$$y_{k+1} \leq \left(1-\frac{d}{k}\right)y_k+\frac{c}{k^{p+1}} \text{ with } c>0, d>0, p>0.$$

(i) *If $d \neq p$, then $y_k = \mathscr{O}\left(k^{-\min(p,d)}\right)$;*
(ii) *if $d=p$, then $y_k = \mathscr{O}\left(k^{-p}\ln k\right)$.*

Lemma A.3. *Let $\{\theta_k\}$, $\{h_k\}$ and $\{\xi_k\}$ be three real sequences, such that $\{h_k\}$ is bounded and $\xi_k > -1, \forall k \geq k_0$. Suppose that, for every $k \geq k_0$,*

$$\theta_{k+1} - \theta_k + \xi_k \theta_{k+1} \leq \xi_k h_{k+1} \text{ and } \sum_{k=k_0}^{+\infty} \xi_k = +\infty.$$

Then $\{\theta_k\}$ is bounded and we have $\limsup_{k\to+\infty} \theta_k \leq \limsup_{k\to+\infty} h_k$.

Proof. Setting $\phi_n := \sup_k \{h_k : k \geq n\}$, we have, for every $k \geq n$, $\theta_{k+1} - \theta_k + \xi_k(\theta_{k+1} - \phi_n) \leq 0$. By setting $E_k = \theta_k - \phi_n$ for all $k \geq n$, the last inequality becomes $(1+\xi_k)E_{k+1} - E_k \leq 0$. Multiplying by $\Pi_k := \prod_{j=k_0}^{k-1}(1+\xi_j)$, we have

$$(1+\xi_k)\Pi_k E_{k+1} - \Pi_k E_k = \Pi_{k+1}E_{k+1} - \Pi_k E_k \leq 0.$$

Summation from n to k, we obtain

$$E_{k+1} \leq \frac{\Pi_n E_n}{\Pi_{k+1}}.$$

In view of $\sum_{k=k_0}^{+\infty} \xi_k = +\infty$, we have $\lim_{k\to+\infty} \Pi_k = +\infty$, which implies

$$\limsup_{k\to+\infty}(\theta_k - \phi_n) = \limsup_{k\to+\infty} E_k \leq \lim_{k\to+\infty} \frac{\Pi_n E_n}{\Pi_k} = 0.$$

Letting once more $n \to +\infty$, we deduce that $\limsup_{k\to+\infty} \theta_k \leq \lim_{n\to+\infty} \phi_n = \limsup_{k\to+\infty} h_k$, which completes the proof. □

REFERENCES

[1] R. T. Rockafellar, On the maximality of sums of nonlinear monotone operators, Trans. Amer. Math. Soc. 149 (1970), 75–88.

[2] H. Attouch, H. Riahi, M. Théra, Somme ponctuelle d'opérateurs maximaux monotones, Serdica Math. J. 22 (1996), 267–292.

[3] I. Ekeland, R. Témam, Convex Analysis and Variational Problems, Classics in Applied Mathematics, SIAM, Philadelphia, 1999.

[4] R.T. Rockafellar, Monotone operators associated with saddle-functions and minimax problems, Proc. Symp. Pure Math. Amer. Math. Sot. Part 1, 18 (1970), 241–250.

[5] R.T. Rockafellar, Augmented Lagrangians and applications of the proximal point algorithm in convex programming, Math. Oper. Res. 1 (1976), 97–116.

[6] A. Chambolle, T. Pock, A first-order primal-dual algorithm for convex problems with applications to imaging, J. Math. Imaging Vis. 40 (2011), 120–145.

[7] R.I. Boţ, D.K. Nguyen, Tikhonov regularization of monotone operator flows not only ensures strong convergence of the trajectories but also speeds up the vanishing of the residuals, J. Complexity 94 (2026) 102029.

[8] B. Martinet, Régularisation d'inéquations variationnelles par approximations successives, Rev. Française d'Inform, Recherche Oper. 4 (1970), 154–159.

[9] R.T. Rockafellar, Monotone operators and the proximal point algorithm, SIAM J. Contr. Opter. 14 (1976), 877–898.

[10] P.L. Lions, B. Mercier, Splitting algorithms for the sum of two nonlinear operators, SIAM J. Numerical Anal. 16 (1979), 964–979.

[11] J. Eckstein, D.P. Bertsekas, On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators, Math. Program. 55 (1992), 293–318.

[12] G.H. Chen, R.T. Rockafellar, Convergence rates in forward-backward splitting, SIAM J. Optim. 7 (1997), 421–444.

[13] H. Raguet, J. Fadili, G. Peyré, A generalized forward-backward splitting, SIAM J. Imaging Sci. 6 (2013), 1199–1226.

[14] R. Boţ, E. Csetnek, C. Hendrich, Inertial Douglas-Rachford splitting for monotone inclusion problems, Appl. Math. Comput. 256 (2015), 472–487.

[15] E.K. Ryu, S. Boyd, A primer on monotone operator methods (survey), Appl. Comput. Math. 15 (2016), 3–43.

[16] Y. Nesterov, Lectures on Convex Optimization, Springer, 2018. https://doi.org/10.1007/978-3-319-91578-4

[17] D. Kim, Accelerated proximal point method for maximally monotone operators, Math. Program. 190 (2021), 57–87.

[18] D. Medhi, C.D. Ha, Generalized proximal point algorithm for convex optimization, J. Optim. Theory Appl. 88 (1996), 475–488.

[19] L. Leuştean, A. Nicolae, A. Sipoş, An abstract proximal point algorithm, J. Global Optim. 72 (2018), 553–577.

[20] A.C. Bagy, Z. Chbani, H. Riahi, On the strong convergence of an inertial proximal algorithm with a time scale, Hessian-driven damping, and a Tikhonov regularization term, J. Appl. Numer. Optim. 6 (2024), 249–269.

[21] R.E. Bruck, Asymptotic convergence of nonlinear contraction semigroups in Hilbert space, J. Funct. Anal., **18**(1) (1975), 15–26.

[22] J.B. Baillon,H. Brézis, Une remarque sur le comportement asymptotique des semigroupes non linéaires, Houston J. Math. 2 (1976), 5–7.

[23] J.B. Baillon, Un exemple concernant le comportement asymptotique de la solution du problème $du/dt + \partial\Phi(u) \ni 0$, J. Funct. Anal. 28 (1978), 369–376.

[24] F. Alvarez, H. Attouch, An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping, Set-Valued Anal. 9 (2001), 3–11.

[25] J. Peypouquet, S. Sorin, Evolution equations for maximal monotone operators: Asymptotic analysis in continuous and discrete iime, J. Convex Anal. 17 (2010), 1113–1163.

[26] H. Attouch, R. Cominetti, A dynamical approach to convex minimization coupling approximation with the steepest descent method, J. Differential Equations 128 (1996), 519–540.

[27] R. Cominetti, J. Peypouquet, S. Sorin, Strong asymptotic convergence of evolution equations governed by maximally monotone operators with Tikhonov regularization, J. Differential Equations 245 (2008), 3753–3763.

[28] V. Apidopoulos, C. Molinari, J. Peypouquet, S. Villa, Preconditioned primal-dual dynamics in convex optimization: Non-ergodic convergence rates, Nonlinear Anal. 60 (2026), 101674.

[29] H. Luo, A primal-dual flow for affine constrained convex optimization, ESAIM Control Optim. Calc. Var. 28 (2022), 33.

[30] F. Battahi, Z. Chbani, S.K. Niederländer, H. Riahi, Asymptotic behavior of the Arrow-Hurwicz differential system with Tikhonov regularization, Preprint, arXiv:2411.17656 [math.OC] (2025).

[31] R.T. Rockafellar, On the maximal monotonicity of subdifferential mappings, Pacific J. Math. 33 (1970), 209–216.

[32] H. Bauschke, P. L. Combettes, Convex Analysis and Monotone Operator Theory in Hilbert spaces, CMS Books in Mathematics, Springer, 2011.

[33] B. Polyak, Introduction to optimization, New York: Optimization Software, Inc. 1987.

[34] K.L. Chung, On a Stochastic Approximation Method, Ann. Math. Statistics, 25 (1954), 463–483.