A primal-dual dynamical approach to structured convex minimization problems

Radu Ioan Boţ [*]   Ernö Robert Csetnek [†]   Szilárd Csaba László [‡]

April 30, 2020

Abstract. In this paper we propose a primal-dual dynamical approach to the minimization of a structured convex function consisting of a smooth term, a nonsmooth term, and the composition of another nonsmooth term with a linear continuous operator. In this scope we introduce a dynamical system for which we prove that its trajectories asymptotically converge to a saddle point of the Lagrangian of the underlying convex minimization problem as time tends to infinity. In addition, we provide rates for both the violation of the feasibility condition by the ergodic trajectories and the convergence of the objective function along these ergodic trajectories to its minimal value. Explicit time discretization of the dynamical system results in a numerical algorithm which is a combination of the linearized proximal method of multipliers and the proximal ADMM algorithm.

Keywords. structured convex minimization, dynamical system, proximal ADMM algorithm, primal-dual algorithm

AMS Subject Classification. 37N40, 49N15, 90C25, 90C46

1 Introduction and preliminaries

For $\mathcal{H}$ and $\mathcal{G}$ real Hilbert spaces, we consider the convex minimization problem

$$\inf_{x \in \mathcal{H}} f(x) + h(x) + g(Ax), \tag{1}$$

where $f : \mathcal{H} \longrightarrow \overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ and $g : \mathcal{G} \longrightarrow \overline{\mathbb{R}}$ are proper, convex and lower semicontinuous functions, $h : \mathcal{H} \longrightarrow \mathbb{R}$ is a convex and Fréchet differentiable function with L_h-Lipschitz continuous gradient ($L_h \geq 0$), i.e. $\|\nabla h(x) - \nabla h(y)\| \leq L_h\|x - y\|$ for every $x, y \in \mathcal{H}$, and $A : \mathcal{H} \longrightarrow \mathcal{G}$ is a continuous linear operator.

Problem (1) can be rewritten as

$$\inf_{\substack{(x,z) \in \mathcal{H} \times \mathcal{G} \\ Ax - z = 0}} f(x) + h(x) + g(z). \tag{2}$$

Obviously, $x^* \in \mathcal{H}$ is an optimal solution of (1) if and only if $(x^*, z^*) \in \mathcal{H} \times \mathcal{G}$ is an optimal solution of (2) and $Ax^* = z^*$.

---

[*]University of Vienna, Faculty of Mathematics, Oskar-Morgenstern-Platz 1, A-1090 Vienna, Austria, email: radu.bot@univie.ac.at. Research partially supported by FWF (Austrian Science Fund), project I 2419-N32.

[†]University of Vienna, Faculty of Mathematics, Oskar-Morgenstern-Platz 1, A-1090 Vienna, Austria, email: er-noe.robert.csetnek@univie.ac.at. Research supported by FWF (Austrian Science Fund), project P 29809-N32.

[‡]Technical University of Cluj-Napoca, Department of Mathematics, Memorandumului 28, Cluj-Napoca, Romania, email: szilard.laszlo@math.utcluj.ro. This work was supported by a grant of Ministry of Research and Innovation, CNCS - UEFISCDI, project number PN-III-P1-1.1-TE-2016-0266, within PNCDI III.

Based on this reformulation of problem (1) we define its Lagrangian

$$l : \mathcal{H} \times \mathcal{G} \times \mathcal{G} \longrightarrow \overline{\mathbb{R}}, \; l(x, z, y) = f(x) + h(x) + g(z) + \langle y, Ax - z \rangle.$$

An element $(x^*, z^*, y^*) \in \mathcal{H} \times \mathcal{G} \times \mathcal{G}$ is said to be a saddle point of the Lagrangian l, if

$$l(x^*, z^*, y) \leq l(x^*, z^*, y^*) \leq l(x, z, y^*), \; \forall (x, z, y) \in \mathcal{H} \times \mathcal{G} \times \mathcal{G}.$$

It is known that $(x^*, z^*, y^*) \in \mathcal{H} \times \mathcal{G} \times \mathcal{G}$ is a saddle point of l if and only if x^* is an optimal solution of (1), $Ax^* = z^*$, and y^* is an optimal solution of the Fenchel dual to problem (1), which reads

$$\sup_{y \in \mathcal{G}} \left(-(f^* \square h^*)(-A^* y) - g^*(y) \right). \tag{3}$$

In this situation the optimal objective values of (1) and (3) coincide.

In the formulation of (3),

$$f^* : \mathcal{H} \to \overline{\mathbb{R}}, \; f^*(u) = \sup_{x \in \mathcal{H}} (\langle u, x \rangle - f(x)), \quad h^* : \mathcal{H} \to \overline{\mathbb{R}}, \; h^*(u) = \sup_{x \in \mathcal{H}} (\langle u, x \rangle - h(x)),$$

and

$$g^* : \mathcal{G} \to \overline{\mathbb{R}}, \; g^*(y) = \sup_{z \in \mathcal{G}} (\langle y, z \rangle - g(z)),$$

denote the conjugate functions of f, h and g, respectively, and $A^* : \mathcal{G} \to \mathcal{H}$ denotes the adjoint operator of A. The infimal convolution $f^* \square h^* : \mathcal{H} \to \overline{\mathbb{R}}$ of the functions f^* and h^* is defined by

$$(f^* \square h^*)(x) = \inf_{y \in \mathcal{H}} (f^*(y) + h^*(x - y)).$$

It is also known that $(x^*, z^*, y^*) \in \mathcal{H} \times \mathcal{G} \times \mathcal{G}$ is a saddle point of the Lagrangian l if and only if it is a solution of the following system of primal-dual optimality conditions

$$\begin{cases} 0 \in \partial f(x) + \nabla h(x) + A^* y \\ Ax = z, \, Ax \in \partial g^*(y). \end{cases}$$

We recall that the convex subdifferential of the function $f : \mathcal{H} \to \overline{\mathbb{R}}$ at $x \in \mathcal{H}$ is defined by $\partial f(x) = \{u \in \mathcal{H} : f(x') - f(x) \geq \langle u, x' - x \rangle \; \forall x' \in \mathcal{H}\}$, for $f(x) \in \mathbb{R}$, and by $\partial f(x) = \emptyset$, otherwise.

A saddle point of the Lagrangian l exists whenever the primal problem (1) has an optimal solution and the so-called Attouch-Brézis regularity condition

$$0 \in \operatorname{sqri}(\operatorname{dom} g - A(\operatorname{dom} f))$$

holds. Here,

$$\operatorname{sqri} Q := \{x \in Q : \cup_{\lambda > 0} \lambda (Q - x) \text{ is a closed linear subspace of } \mathcal{G}\}$$

denotes the strong quasi-relative interior of a set $Q \subseteq \mathcal{G}$. We refer the reader to [9, 11, 28] for more insights into the world of regularity conditions and convex duality theory.

Let $S_+(\mathcal{H})$ denote the family of continuous linear operators $U : \mathcal{H} \longrightarrow \mathcal{H}$ which are self-adjoint and positive semidefinite. For $U \in S_+(\mathcal{H})$ we introduce the following seminorm on $\mathcal{H}$:

$$\|x\|_U^2 = \langle x, Ux \rangle \; \forall x \in \mathcal{H}.$$

This introduces on $S_+(\mathcal{H})$ the following partial ordering: for $U_1, U_2 \in S_+(\mathcal{H})$

$$U_1 \succcurlyeq U_2 \Leftrightarrow \|x\|_{U_1}^2 \geq \|x\|_{U_2}^2 \; \forall x \in \mathcal{H}.$$

For $\alpha > 0$ fixed, let be
$$P_\alpha(\mathcal{H}) = \{U \in S_+(\mathcal{H}) : U \succcurlyeq \alpha I\},$$
where $I : \mathcal{H} \longrightarrow \mathcal{H}$, $I(x) = x$, denotes the identity operator on $\mathcal{H}$.

The subject of our investigations in this paper will be the following dynamical system, for which we will show that it asymptotically approaches the set of solutions of the primal-dual pair of optimization problems (1)-(3)

$$
\begin{cases}
\dot{x}(t) + x(t) \in (\partial f + cA^*A + M_1(t))^{-1} (M_1(t)x(t) + cA^*z(t) - A^*y(t) - \nabla h(x(t))) \\[2mm]
\dot{z}(t) + z(t) \in (\partial g + cI + M_2(t))^{-1} (M_2(t)z(t) + cA(\gamma\dot{x}(t) + x(t)) + y(t)) \\[2mm]
\dot{y}(t) = cA(x(t) + \dot{x}(t)) - c(z(t) + \dot{z}(t)) \\[2mm]
x(0) = x^0 \in \mathcal{H}, \ z(0) = z^0 \in \mathcal{G}, \ y(0) = y^0 \in \mathcal{G},
\end{cases}
\tag{4}
$$

where $c > 0$, $\gamma \in [0,1]$, and $M_1 : [0, +\infty) \longrightarrow S_+(\mathcal{H})$ and $M_2 : [0, +\infty) \longrightarrow S_+(\mathcal{G})$ are such that

$$(Cstrong) \quad \text{there exists } \alpha > 0 \text{ such that } cA^*A + M_1(t) \in P_\alpha(\mathcal{H}) \quad \forall t \in [0, +\infty).$$

One of the motivation for the study of this dynamical system comes from the fact that, as we will see in Remark 1, it provides through explicit time discretization a numerical algorithm which is a combination of the linearized proximal method of multipliers and the proximal ADMM algorithm.

In the next section we will show the existence and uniqueness of strong global solutions for the dynamical system (4) in the framework of the Cauchy-Lipschitz Theorem. In Section 3 we will prove some technical results, which will play an important role in the asymptotic analyis. In Section 4 we will investigate the asymptotic behaviour of the trajectories as the time tends to infinity. By carrying out a Lyapunov analysis and by relying on the continuous variant of the Opial Lemma, we are able to prove that the trajectories generated by (4) asymptotically convergence to a saddle point of the Lagrangian l. Furthermore, we provide convergence rates of $\mathcal{O}(\frac{1}{t})$ for the violation of the feasibility condition by ergodic trajectories and the convergence of the objective function along these ergodic trajectories to its minimal value.

The approach of optimization problems by dynamical systems has a long tradition. Crandall and Pazy considered in [20] dynamical systems governed by subdifferential operators (and more general by maximally monotone operators) in Hilbert spaces, addressed questions like the existence and uniqueness of solution trajectories, and related the latter to the theory of semi-groups of nonlinear contractions. Brézis [14] studied the asymptotic behaviour of the trajectories for dynamical systems governed by convex subdifferentials, and Bruck carried out in [15] a similar analysis for maximally monotone operators. Dynamical systems defined via resolvent/proximal evaluations of the governing operators have enjoyed much attention in the last years, as they result by explicit time discretization in relaxed versions of standard numerical algorithms, with high flexibility and good numerical performances. Abbas and Attouch introduced in [1] a forward-backward dynamical system, by extending to more general optimization problems an approach proposed by Antipin in [5] and Bolte in [10] on a gradient-projected dynamical system associated to the minimization of a smooth convex function over a convex closed set. Implicit dynamical systems were considered also in [13] in the context of monotone inclusion problems. A dynamical system of forward-backward-forward type was considered in [7], while, a dynamical system of Douglas-Rachford type was recently introduced in [21].

It is important to notice that the approaches mentioned above have been introduced in connection with the study of "simple" monotone inclusion and convex minimization problems. They rely on straightforward splitting strategies and cannot be efficiently used when addressing structured

minimization problems, like (1), which need to be addressed from a primal and a dual perspective, thus, require for tools and techniques from the convex duality theory. The dynamical approach we introduce and investigate in this paper is, to our knowledge, the first meant to address structured convex minimization problems in the spirit of the full splitting paradigm.

Remark 1. The first inclusion in (4) can be equivalently written as

$$0 \in \partial f(\dot{x}(t)+x(t))+cA^*A(\dot{x}(t)+x(t))+M_1(t)\dot{x}(t)-(cA^*z(t)-A^*y(t)-\nabla h(x(t))) \ \forall t \in [0,+\infty), \ (5)$$

while the second one as

$$0 \in \partial g(\dot{z}(t) + z(t)) + c(\dot{z}(t) + z(t)) - cA(\gamma\dot{x}(t) + x(t)) - y(t) + M_2(t)\dot{z}(t) \ \forall t \in [0,+\infty). \quad (6)$$

The explicit discretization of (5) with respect to the time variable t and constant step equal to 1 yields the iterative scheme

$$0 \in \frac{1}{c}\partial f(x^{k+1}) + A^*Ax^{k+1} + \frac{M_1^k}{c}(x^{k+1} - x^k) - A^*z^k + \frac{A^*}{c}y^k + \frac{1}{c}\nabla h(x^k) \ \forall k \geq 0.$$

By convex subdifferential calculus, one can easily see that this can be for every $k \geq 0$ equivalently written as

$$0 \in \partial \left(f(x) + \langle x - x^k, \nabla h(x^k) \rangle + \frac{c}{2} \left\| Ax - z^k + \frac{y^k}{c} \right\|^2 + \frac{1}{2}\|x - x^k\|_{M_1^k}^2 \right) \bigg|_{x=x^{k+1}}$$

and, further, as

$$x^{k+1} \in \underset{x\in\mathcal{H}}{\operatorname{argmin}} \left(f(x) + \langle x - x^k, \nabla h(x^k) \rangle + \frac{c}{2} \left\| Ax - z^k + \frac{y^k}{c} \right\|^2 + \frac{1}{2}\|x - x^k\|_{M_1^k}^2 \right).$$

Similarly, (6) leads for every $k \geq 0$ to

$$0 \in \partial \left(g(z) + \frac{c}{2} \left\| A(\gamma x^{k+1} + (1-\gamma)x^k) - z + \frac{y^k}{c} \right\|^2 + \frac{1}{2}\|z - z^k\|_{M_2^k}^2 \right) \bigg|_{z=z^{k+1}},$$

which is nothing else than

$$z^{k+1} \in \underset{z\in\mathcal{G}}{\operatorname{argmin}} \left(g(z) + \frac{c}{2} \left\| A(\gamma x^{k+1} + (1-\gamma)x^k) - z + \frac{y^k}{c} \right\|^2 + \frac{1}{2}\|z - z^k\|_{M_2^k}^2 \right).$$

Here, $(M_1^k)_{k\geq 0}$ and $(M_2^k)_{k\geq 0}$ are two operator sequences in $S_+(\mathcal{H})$ and $S_+(\mathcal{G})$, respectively.

Thus the dynamical system (4) leads through explicit time discretization to a numerical algorithm, which, for a starting point $(x^0, z^0, y^0) \in \mathcal{H} \times \mathcal{G} \times \mathcal{G}$, generates a sequence $(x^k, z^k, y^k)_{k\geq 0}$ for every $k \geq 0$ as follows

$$\begin{cases} x^{k+1} \in \underset{x\in\mathcal{H}}{\operatorname{argmin}} \left(f(x) + \langle x - x^k, \nabla h(x^k) \rangle + \frac{c}{2} \left\| Ax - z^k + \frac{y^k}{c} \right\|^2 + \frac{1}{2}\|x - x^k\|_{M_1^k}^2 \right) \\[2ex] z^{k+1} \in \underset{z\in\mathcal{G}}{\operatorname{argmin}} \left(g(z) + \frac{c}{2} \left\| A(\gamma x^{k+1} + (1-\gamma)x^k) - z + \frac{y^k}{c} \right\|^2 + \frac{1}{2}\|z - z^k\|_{M_2^k}^2 \right) \\[2ex] y^{k+1} = y^k + c(Ax^{k+1} - z^{k+1}). \end{cases} \quad (7)$$

4

The algorithm (7) is a combination of the linearized proximal method of multipliers and the proximal ADMM algorithm.

Indeed, in the case when $\gamma = 1$, (7) becomes the proximal ADMM algorithm with variable metrics from [8] (see, also, [12]). If, in addition, $h = 0$ and the operator sequences $(M_1^k)_{k \geq 0}$ and $(M_2^k)_{k \geq 0}$ are constant, then (7) becomes the proximal ADMM algorithm investigated in [25, Section 3.2] (see, also, [23]). It is known that the proximal ADMM algorithm can be seen as a generalization of the full splitting primal-dual algorithms of Chambolle-Pock (see [16]) and Condat-Vu (see [19,27]).

On the other hand, in the case when $\gamma = 0$, (7) becomes an extension of the linearized proximal method of multipliers of Chen-Teboulle (see [17], [25, Algorithm 1]).

In the following remark we provide a particular choice for the maps M_1 and M_2, which transforms (4) into a dynamical system of primal-dual type formulated in the spirit of the full splitting paradigm.

Remark 2. For every $t \in [0, +\infty)$, define

$$M_1(t) = \frac{1}{\tau(t)} I - cA^*A \text{ and } M_2(t) = 0,$$

where $\tau(t) > 0$ is such that $c\tau(t)\|A\|^2 \leq 1$.

Let $t \in [0, +\infty)$ be fixed. In this particular setting, (5) is equivalent to

$$\left(\frac{1}{\tau(t)} I - cA^*A \right) x(t) + cA^*z(t) - A^*y(t) - \nabla h(x(t)) \in \frac{1}{\tau(t)} \dot{x}(t) + \frac{1}{\tau(t)} x(t) + \partial f(\dot{x}(t) + x(t))$$

and further to

$$\dot{x}(t) + x(t) = (I + \tau(t)\partial f)^{-1} \left((I - c\tau(t)A^*A)x(t) + c\tau(t)A^*z(t) - \tau(t)A^*y(t) - \tau(t)\nabla h(x(t)) \right).$$

In other words,

$$\dot{x}(t) + x(t) = \mathrm{prox}_{\tau(t)f} \left((I - c\tau(t)A^*A)x(t) + c\tau(t)A^*z(t) - \tau(t)A^*y(t) - \tau(t)\nabla h(x(t)) \right),$$

where

$$\mathrm{prox}_\kappa : \mathcal{H} \to \mathcal{H}, \quad \mathrm{prox}_\kappa(x) = \underset{y \in \mathcal{H}}{\mathrm{argmin}} \left\{ \kappa(y) + \frac{1}{2} \|x - y\|^2 \right\} = (I + \partial\kappa)^{-1}(x),$$

denotes the proximal point operator of a proper, convex and lower semicontinuous function $\kappa : \mathcal{H} \to \overline{\mathbb{R}}$.

On the other hand, relation (6) is equivalent to

$$\dot{y}(t) + y(t) + c(\gamma - 1)A\dot{x}(t) \in \partial g(\dot{z}(t) + z(t)),$$

hence,

$$\dot{z}(t) + z(t) \in \partial g^*(\dot{y}(t) + y(t) + c(\gamma - 1)A\dot{x}(t)).$$

This is further equivalent to

$$A(\gamma\dot{x}(t) + x(t)) + \frac{1}{c}y(t) \in \frac{1}{c}\dot{y}(t) + \frac{1}{c}y(t) + (\gamma - 1)A\dot{x}(t) + \partial g^*(\dot{y}(t) + y(t) + c(\gamma - 1)A\dot{x}(t))$$

and further to

$$\dot{y}(t) + y(t) + c(\gamma - 1)A\dot{x}(t) = (I + c\partial g^*)^{-1}(cA(\gamma\dot{x}(t) + x(t)) + y(t)).$$

In other words,

$$\dot{y}(t) + y(t) + c(\gamma - 1)A\dot{x}(t) = \mathrm{prox}_{cg^*} \left(cA(\gamma\dot{x}(t) + x(t)) + y(t) \right).$$

5

Consequently, in this particular setting, the dynamical system (4) can be equivalently written as

$$
\begin{cases}
\dot{x}(t) + x(t) = \mathrm{prox}_{\tau(t)f}\left((I - c\tau(t)A^*A)x(t) + c\tau(t)A^*z(t) - \tau(t)A^*y(t) - \tau(t)\nabla h(x(t))\right) \\[2mm]
\dot{y}(t) + y(t) + c(\gamma - 1)A\dot{x}(t) = \mathrm{prox}_{cg^*}\left(cA(\gamma\dot{x}(t) + x(t)) + y(t)\right) \\[2mm]
\dot{y}(t) = cA(x(t) + \dot{x}(t)) - c(z(t) + \dot{z}(t)) \\[2mm]
x(0) = x^0 \in \mathcal{H},\ z(0) = z^0 \in \mathcal{G},\ y(0) = y^0 \in \mathcal{G}.
\end{cases}
\tag{8}
$$

Let us also mention that when $h = 0$ and $\gamma = 1$ the dynamical system (8) reads

$$
\begin{cases}
\dot{x}(t) + x(t) = \mathrm{prox}_{\tau(t)f}\left(x(t) - \tau(t)A^*(y(t) + cAx(t) - cz(t))\right) \\[2mm]
\dot{y}(t) + y(t) = \mathrm{prox}_{cg^*}\left(y(t) + cA(\dot{x}(t) + x(t))\right) \\[2mm]
\dot{y}(t) = cA(x(t) + \dot{x}(t)) - c(z(t) + \dot{z}(t)) \\[2mm]
x(0) = x^0 \in \mathcal{H},\ z(0) = z^0 \in \mathcal{G},\ y(0) = y^0 \in \mathcal{G}.
\end{cases}
\tag{9}
$$

The explicit time discretization of (9) leads to a numerical algorithm, which, for a starting point $(x^0, z^0, y^0) \in \mathcal{H} \times \mathcal{G} \times \mathcal{G}$, generates the sequence $(x^k, z^k, y^k)_{k \geq 0}$ for every $k \geq 0$ as follows

$$
\begin{cases}
x^{k+1} = \mathrm{prox}_{\tau_k f}\left(x^k - \tau_k A^*(y^k + cAx^k - cz^k)\right) \\[2mm]
y^{k+1} = \mathrm{prox}_{cg^*}(y^k + cAx^{k+1}) \\[2mm]
y^{k+1} = y^k + c(Ax^{k+1} - z^{k+1}).
\end{cases}
\tag{10}
$$

By substituting in the first equation of (10) the term $cAx^k - cz^k$ by $y^k - y^{k-1}$, which is allowed according to the last equation, one can easily see that (10) is equivalent to the following numerical algorithm, which, for a starting point $(x^0, y^0, y^{-1}) \in \mathcal{H} \times \mathcal{G} \times \mathcal{G}, y^0 = y^{-1}$, generates the sequence $(x^k, y^k)_{k \geq 0}$ for every $k \geq 0$ as follows

$$
\begin{cases}
x^{k+1} = \mathrm{prox}_{\tau_k f}\left(x^k - \tau_k A^*(2y^k - y^{k-1})\right) \\[2mm]
y^{k+1} = \mathrm{prox}_{cg^*}(y^k + cAx^{k+1}).
\end{cases}
\tag{11}
$$

For $\tau_k = \tau > 0$ for every $k \geq 0$, (11) is nothing else than the primal-dual algorithm proposed by Chambolle and Pock in [16].

Remark 3. The maps $M_1 : [0, +\infty) \longrightarrow S_+(\mathcal{H})$ and $M_2 : [0, +\infty) \longrightarrow S_+(\mathcal{G})$ can be seen as inducing a variable renorming of the underlying Hilbert space and, as seen in the remark above, allow the use of variable step sizes. In addition, they might provide favourable settings for the derivation of convergence rates for function values along the trajectories, as it is the case for the discrete time counterpart of (4) (see [12, Section 3]).

Example 1. We will illustrate the way in which the parameters γ, c and $\tau(t), t \in [0, +\infty)$ may influence the asymptotic convergence of the primal and dual trajectories via numerical experiments. In this scope, we considered the following primal optimization problem

$$
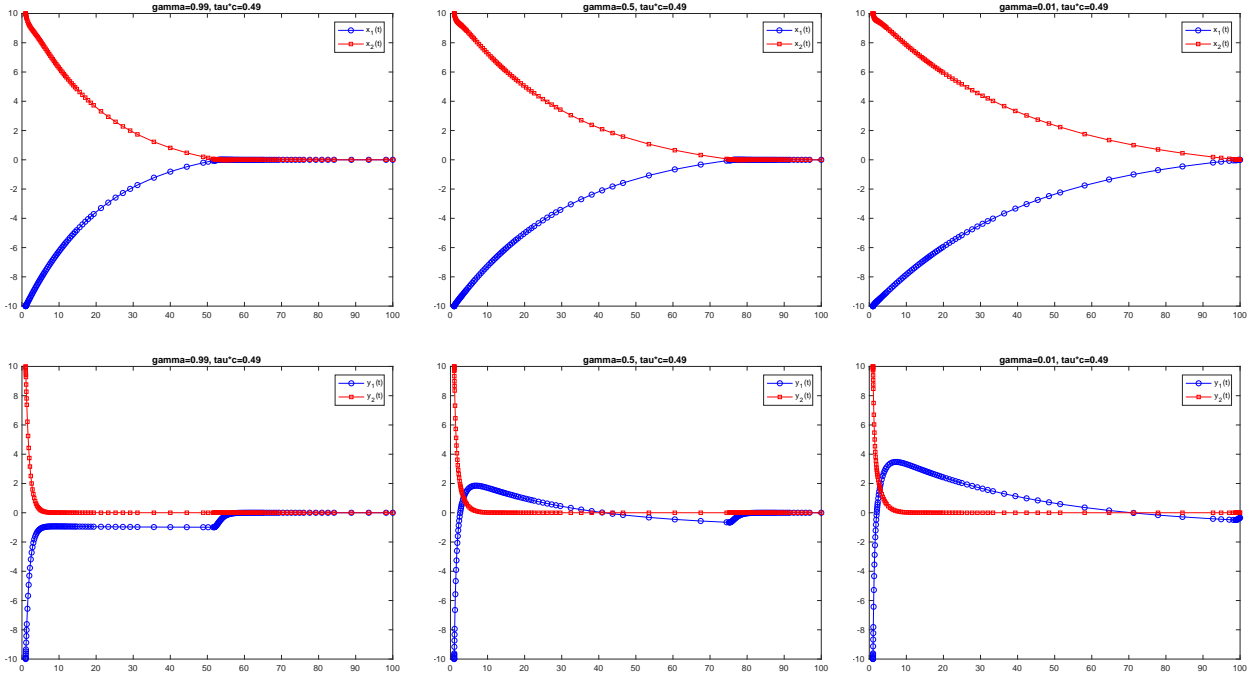\inf_{(x_1, x_2) \in \mathbb{R}^2} \frac{1}{2}(x_1^2 + x_2^2) + |x_1 - x_2| + |x_1 + x_2|,
\tag{12}
$$

Figure 1: First row: the primal trajectory $x(t)$ approaching the primal optimal solution $(0,0)$ for $\tau c = 0.49$ and starting point $x^0 = (-10, 10)$. Second row: the dual trajectory $y(t)$ approaching a dual optimal solution for $\tau c = 0.49$ and starting point $y^0 = (-10, 10)$.

which is in fact problem (1) written in the following particular setting: $\mathcal{H} = \mathcal{G} = \mathbb{R}^2$, $f, g, h : \mathbb{R}^2 \to \mathbb{R}$, $f(x) = \frac{1}{2}\|x\|_2^2$, $g(x) = \|x\|_1$, $h(x) = 0$, for every $x \in \mathbb{R}^2$, and $A : \mathbb{R}^2 \to \mathbb{R}^2$, $A(x_1, x_2) = (x_1 - x_2, x_1 + x_2)$. Then $\overline{x} = (0,0)$ is the unique optimal solution of (12) and that

$$\sup_{\|(y_1, y_2)\|_\infty \leq 1} -y_1^2 - y_2^2 \tag{13}$$

is the Fenchel dual problem of (12). Thus $\overline{y} = (0,0)$ is the unique dual optimal solution.

We considered the dynamical system (8) attached to the primal-dual pair (12)-(13) with starting points $x^0 = (-10, 10)$, $z^0 = Ax^0 = (-20, 0)$ and $y^0 = (-10, 10)$ in the case when $\tau(t) = \tau > 0$ for every $t \in [0, +\infty)$ is a constant function. In order to solve the resulting dynamical system we used the Matlab function `ode15s` and, to this end, we reformulated it as

$$\begin{cases} \dot{U}(t) = \Gamma(U(t)) \\ U(0) = (x^0, y^0, z^0), \end{cases}$$

where

$$U(t) = (x(t), y(t), z(t)) \in \mathcal{H} \times \mathcal{G} \times \mathcal{G}$$

and

$$\Gamma : \mathcal{H} \times \mathcal{G} \times \mathcal{G} \to \mathcal{H} \times \mathcal{G} \times \mathcal{G}, \ \Gamma(u_1, u_2, u_3) = (u_4, u_5, u_6),$$

is defined as

$$\begin{cases} u_4 = \operatorname{prox}_{\tau f}\left(u_1 - \tau A^*(u_2 + cAu_1 - cu_3)\right) - u_1 \\[2mm] u_5 = \operatorname{prox}_{cg^*}\left(u_2 + cA(\gamma u_4 + u_1)\right) - u_2 - c(\gamma - 1)Au_4 \\[2mm] u_6 = A(u_1 + u_4) - u_3 - \frac{1}{c}u_5. \end{cases}$$
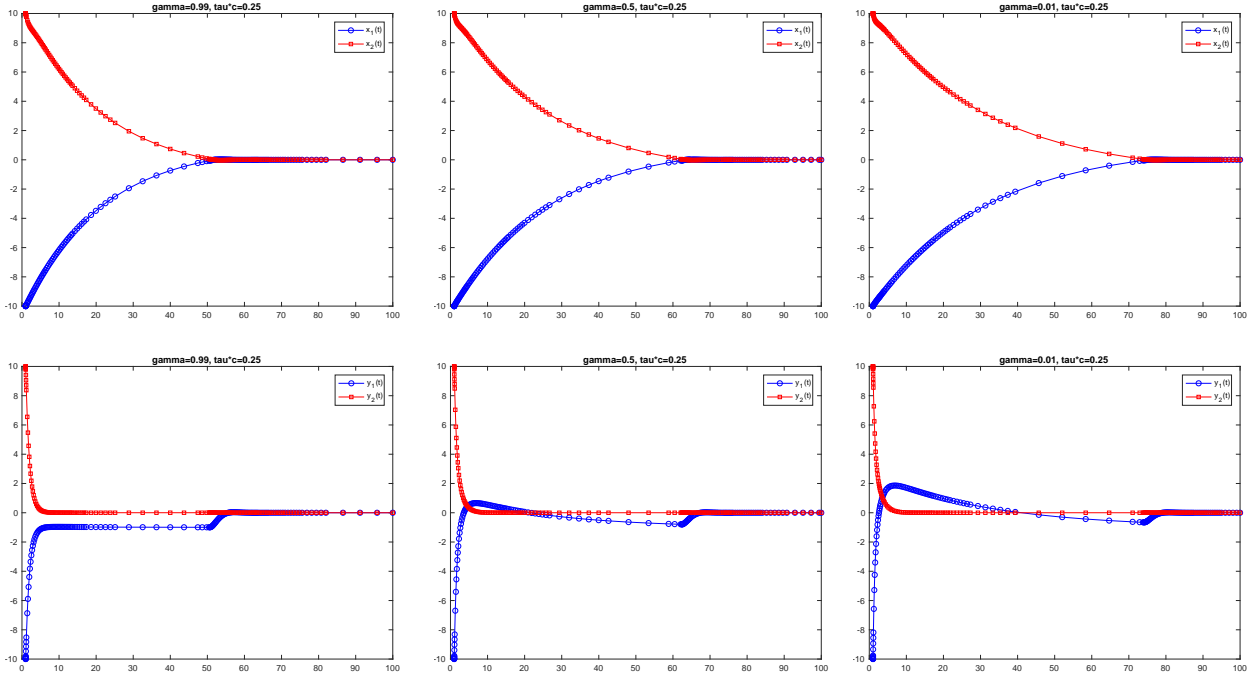
7

Figure 2: First row: the primal trajectory $x(t)$ approaching the primal optimal solution $(0,0)$ for $\tau c = 0.25$ and starting point $x^0 = (-10, 10)$. Second row: the dual trajectory $y(t)$ approaching a dual optimal solution for $\tau c = 0.25$ and starting point $y^0 = (-10, 10)$.

Notice that

$$A^*(x_1, x_2) = (x_1 + x_2, -x_1 + x_2) \quad \forall (x_1, x_2) \in \mathbb{R}^2,$$

$$\text{prox}_{\tau f}(x) = \frac{1}{1 + \tau} x \quad \forall x \in \mathbb{R}^2,$$

$$\text{prox}_{cg^*}(y) = \text{proj}_{[-1,1]^2}(y) \quad \forall y \in \mathbb{R}^2,$$

where proj_Q denotes the projection operator on a convex and closed set $Q \subseteq \mathcal{H}$.

According to Theorem 12, the asymptotic convergence of the trajectories can be guaranteed when $\tau c \|A\|^2 \leq 1$. Since $\|A\| = \sqrt{2}$, we considered for $\tau c \in (0, \frac{1}{2})$ three different choices, namely, $\tau c = 0.49, 0.25$ and 0.01. The primal and the dual trajectories generated by the dynamical system for each of these three choices are represented in the figures 1, 2 and 3, respectively. The first row of each figure represents the primal trajectories $x(t)$ for $\gamma = 0.99, 0.5$ and 0.01, while the second row represents the dual trajectories $y(t)$ for the same choices of the parameter γ.

One can see that the parameter γ has a strong influence on the asymptotic behaviour of the primal and dual trajectories. Namely, in all three figures, thus independently of the choice of the parameters τ and c, the primal and the dual trajectories converge faster to the corresponding primal and dual solutions, respectively, for larger values of γ, namely, when γ is closer to 1. As γ is the coefficient of the derivative in the second inclusion of the dynamical system, it can be seen as a constant which characterizes its level of implicitness. In particular, more implicitness promotes a better asymptotic convergence. On the other hand, we notice that the smaller the values of τc are, the smaller is the influence of γ on the asymptotic convergence of the trajectories.

Notations. The following two functions will play an important role in particular in the forthcoming analysis

$$F : [0, +\infty) \times \mathcal{H} \longrightarrow \overline{\mathbb{R}}, \quad F(t, x) = f(x) + \frac{c}{2}(\|Ax\|^2 - \|x\|^2) + \frac{1}{2}\|x\|_{M_1(t)}^2,$$
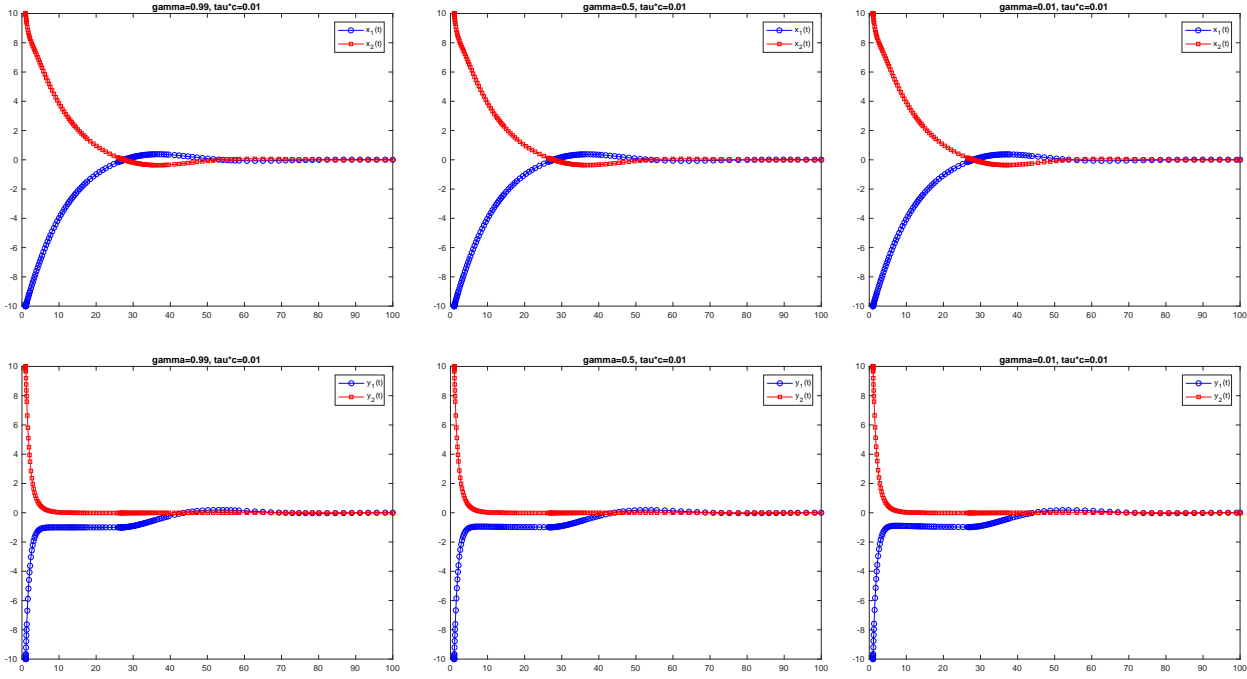
8

Figure 3: First row: the primal trajectory $x(t)$ approaching the primal optimal solution $(0,0)$ for $\tau c = 0.1$ and starting point $x^0 = (-10, 10)$. Second row: the dual trajectory $y(t)$ approaching a dual optimal solution for $\tau c = 0.1$ and starting point $y^0 = (-10, 10)$.

and

$$G : [0, +\infty) \times \mathcal{G} \longrightarrow \overline{\mathbb{R}}, \quad G(t, z) = g(z) + \frac{1}{2} \|z\|^2_{M_2(t)}.$$

With these two notations, the dynamical system (4) can be rewritten as

$$
\begin{cases}
\dot{x}(t) + x(t) \in \underset{x \in \mathcal{H}}{\operatorname{argmin}} \left(F(t, x) + \frac{c}{2} \left\| x - \left(\frac{1}{c} M_1(t) x(t) + A^* z(t) - \frac{A^*}{c} y(t) - \frac{1}{c} \nabla h(x(t)) \right) \right\|^2 \right) \\[2ex]
\dot{z}(t) + z(t) = \underset{z \in \mathcal{G}}{\operatorname{argmin}} \left(G(t, z) + \frac{c}{2} \left\| z - \left(\frac{1}{c} M_2(t) z(t) + A(\gamma \dot{x}(t) + x(t)) + \frac{1}{c} y(t) \right) \right\|^2 \right) \\[2ex]
\dot{y}(t) = cA(x(t) + \dot{x}(t)) - c(z(t) + \dot{z}(t)) \\[2ex]
x(0) = x^0 \in \mathcal{H}, \ y(0) = y^0 \in \mathcal{G}, \ z(0) = z^0 \in \mathcal{G}.
\end{cases}
\tag{14}
$$

Let $t \in [0, +\infty)$ be fixed. The function $G(t, \cdot)$ is proper, convex and lower semicontinuous, hence $z \to G(t, z) + \frac{c}{2} \|z - v\|^2$ is proper, strongly convex and lower semicontinuous for every $v \in \mathcal{G}$. This allows us to use the sign equal in the second relation of (14). On the other hand, a sufficient condition which guarantees that the function $x \mapsto F(x, t) + \frac{c}{2} \left\| x - \left(\frac{1}{c} M_1(t) x(t) + A^* z(t) - \frac{A^*}{c} y(t) - \frac{1}{c} \nabla h(x(t)) \right) \right\|^2$, which is proper and lower semicontinuous, is strongly convex is that there exists $\alpha(t) > 0$ such that $cA^*A + M_1(t) \in P_{\alpha(t)}(\mathcal{H})$. This actually ensures that $x \to F(t, x) + \frac{c}{2} \|x - u\|^2$ is proper, strongly convex and lower semicontinuous for every $u \in \mathcal{H}$.

This means that if the assumption

$(Cweak)$ for every $t \in [0, +\infty)$ there exists $\alpha(t) > 0$ such that $cA^*A + M_1(t) \in P_{\alpha(t)}(\mathcal{H})$

holds, then we can use also in the first relation of (14) the sign equal. It is easy to see, that, if $(Cweak)$ holds, then $\partial f + cA^*A + M_1(t)$ is $\alpha(t)$-strongly monotone for every $t \in [0, +\infty)$. In

9

other words, for every $t \in [0, +\infty)$, all $u, v \in \mathcal{H}$ and all $u^* \in (\partial f + cA^*A + M_1(t))(u), x^* \in (\partial f + cA^*A + M_1(t))(x)$ we have

$$\langle u^* - x^*, u - x \rangle \geq \alpha(t) \|u - x\|^2.$$

Notice that, since $A^*A \in S_+(\mathcal{H})$ and $M_1(t) \in S_+(\mathcal{H})$ for every $t \in [0, +\infty)$, $(Cweak)$ is fulfilled, if

$$\text{for every } t \in [0, +\infty) \text{ there exists } \alpha(t) > 0 \text{ such that } M_1(t) \in P_{\alpha(t)}(\mathcal{H}) \tag{15}$$

or, if

$$\text{there exists } \alpha > 0 \text{ such that } A^*A \in P_\alpha(\mathcal{H}). \tag{16}$$

Notice also that, if $\mathcal{H}$ is a finite dimensional Hilbert space, then (16), which is independent of t, is nothing else than A^*A is positively definite or, equivalently, A is injective.

Let $S = \{x \in \mathcal{H} : \|x\| = 1\}$ be the unit sphere of $\mathcal{H}$. Assumption $(Cweak)$ is fulfilled if and only if $\inf_{x \in S} \|x\|^2_{cA^*A + M_1(t)} > 0$ for every $t \in [0, +\infty)$. In this case we can take $\alpha(t) := \inf_{x \in S} \|x\|^2_{cA^*A + M_1(t)}$ for every $t \in [0, +\infty)$.

Obviously, if $(Cstrong)$ holds, then $(Cweak)$ holds with $\alpha(t) := \alpha > 0$ for every $t \in [0, +\infty)$.

2 Existence and uniqueness of the trajectories

In this section we will investigate the existence and uniqueness of the trajectories generated by (4). We start by recalling the definition of a locally absolutely continuous map.

Definition 1. *A function* $x : [0, +\infty) \to \mathcal{H}$ *is said to be locally absolutely continuous, if it is absolutely continuous on every interval* $[0, T]$, $T > 0$; *that is, for every* $T > 0$ *there exists an integrable function* $y : [0, T] \to \mathcal{H}$ *such that*

$$x(t) = x(0) + \int_0^t y(s)ds \quad \forall t \in [0, T].$$

Remark 4. (a) Every absolutely continuous function is differentiable almost everywhere, its derivative coincides with its distributional derivative almost everywhere and one can recover the function from its derivative $\dot{x} = y$ by the above integration formula.

(b) Let be $T > 0$ and $x : [0, T] \to \mathcal{H}$ an absolutely continuous function. This is equivalent to (see [2, 6]): for every $\varepsilon > 0$ there exists $\eta > 0$ such that for any finite family of intervals $I_k = (a_k, b_k) \subseteq [0, T]$ the following property holds:

for any subfamily of disjoint intervals I_j with $\sum_j |b_j - a_j| < \eta$ it holds $\sum_j \|x(b_j) - x(a_j)\| < \varepsilon$.

From this characterization it is easy to see that, if $B : \mathcal{H} \to \mathcal{H}$ is L-Lipschitz continuous with $L \geq 0$, then the function $z = B \circ x$ is absolutely continuous, too. This means that z is differentiable almost everywhere and $\|\dot{z}(\cdot)\| \leq L \|\dot{x}(\cdot)\|$ holds almost everywhere.

The following definition specifies which type of solutions we consider in the analysis of the dynamical system (4).

Definition 2. *Let* $(x^0, z^0, y^0) \in \mathcal{H} \times \mathcal{G} \times \mathcal{G}$, $c > 0$, $\gamma \in [0, 1]$, *and* $M_1 : [0, +\infty) \to \mathcal{S}_+(\mathcal{H})$ *and* $M_2 : [0, +\infty) \to \mathcal{S}_+(\mathcal{G})$. *We say that the function* $(x, z, y) : [0, +\infty) \longrightarrow \mathcal{H} \times \mathcal{G} \times \mathcal{G}$ *is a strong global solutions of* (4), *if the following properties are satisfied:*

(i) *the functions* x, z, y *are locally absolutely continuous;*

(ii) for almost every $t \in [0, +\infty)$

$$\dot{x}(t) + x(t) \in (\partial f + cA^*A + M_1(t))^{-1} \left(M_1(t)x(t) + cA^*z(t) - A^*y(t) - \nabla h(x(t)) \right),$$
$$\dot{z}(t) + z(t) \in (\partial g + c\,\mathrm{Id} + M_2(t))^{-1} \left(M_2(t)z(t) + cA(\gamma \dot{x}(t) + x(t)) + y(t) \right)$$
$$\dot{y}(t) = cA(x(t) + \dot{x}(t)) - c(z(t) + \dot{z}(t));$$

(iii) $x(0) = x^0$, $z(0) = z^0$, and $y(0) = y^0$.

The following results will be useful in the proof of the existence and uniqueness theorem.

Lemma 1. *Assume that $(Cweak)$ holds. Then, for every fixed $t \in [0, +\infty)$, the operator*

$$S_t : \mathcal{H} \longrightarrow \mathcal{H}, \quad S_t(u) = \operatorname*{argmin}_{x \in \mathcal{H}} \left(F(t, x) + \frac{c}{2} \|x - u\|^2 \right),$$

is Lipschitz continuous.

Proof. Let $t \in [0, +\infty)$ be fixed and $u, v \in \mathcal{H}$. By subdifferential calculus we obtain that

$$cu \in \partial f(S_t(u)) + \left(cA^*A + M_1(t) \right)(S_t(u))$$

and

$$cv \in \partial f(S_t(v)) + \left(cA^*A + M_1(t) \right)(S_t(v)).$$

Using that, due to $(Cweak)$, $\partial f + cA^*A + M_1(t)$ is $\alpha(t)$-strongly monotone, we get

$$\alpha(t)\|S_t u - S_t v\|^2 \le c \langle u - v, S_t(u) - S_t(v) \rangle.$$

By the Cauchy-Schwartz inequality we obtain

$$\|S_t u - S_t v\| \le \frac{c}{\alpha(t)} \|u - v\|,$$

which shows that S_t is Lipschitz continuous with constant $\frac{c}{\alpha(t)}$. $\qquad\square$

Now we are going to prove another technical result which will be used in the proof of the main theorem of this section.

Lemma 2. *Assume that $(Cweak)$ holds. Let be $(x, z, y) \in \mathcal{H} \times \mathcal{G} \times \mathcal{G}$ and the maps $R_{(x,z,y)} : [0, +\infty) \longrightarrow \mathcal{H}$,*

$$R_{(x,z,y)}(t) = \operatorname*{argmin}_{u \in \mathcal{H}} \left(F(t, u) + \frac{c}{2} \left\| u - \left(\frac{1}{c} M_1(t)x + A^*z - \frac{1}{c} A^*y - \frac{1}{c} \nabla h(x) \right) \right\|^2 \right) - x,$$

and $Q_{(x,z,y)} : [0, +\infty) \longrightarrow \mathcal{G}$,

$$Q_{(x,z,y)}(t) = \operatorname*{argmin}_{v \in \mathcal{G}} \left(G(t, v) + \frac{c}{2} \left\| v - \left(\frac{1}{c} M_2(t)z + A\left(\gamma R_{(x,z,y)}(t) + x \right) + \frac{1}{c} y \right) \right\|^2 \right) - z.$$

Then the following statements are true for every $t, r \in [0, +\infty)$:

(i) $\|R_{(x,z,y)}(t) - R_{(x,z,y)}(r)\| \le \dfrac{\|R_{(x,z,y)}(r)\|}{\alpha(t)} \|M_1(t) - M_1(r)\|$;

(ii) $\|Q_{(x,z,y)}(t) - Q_{(x,z,y)}(r)\| \le \dfrac{\|Q_{(x,z,y)}(r)\|}{c} \|M_2(t) - M_2(r)\| + \dfrac{\gamma\|A\|\|R_{(x,z,y)}(r)\|}{\alpha(t)} \|M_1(t) - M_1(r)\|$.

11

Proof. Let $t, r \in [0, +\infty)$ be fixed.

(i) From the definition of $R_{(x,z,y)}$ one has

$$M_1(t)x + cA^*z - A^*y - \nabla h(x) \in \partial f(R_{(x,z,y)}(t) + x) + \big(cA^*A + M_1(t)\big)(R_{(x,z,y)}(t) + x)$$

and

$$M_1(r)x + cA^*z - A^*y - \nabla h(x) \in \partial f(R_{(x,z,y)}(r) + x) + \big(cA^*A + M_1(r)\big)(R_{(x,z,y)}(r) + x),$$

which is equivalent to

$$M_1(t)(R_{(x,z,y)}(r) + x) - M_1(r)(R_{(x,z,y)}(r)) + cA^*z - A^*y - \nabla h(x) \in$$
$$\partial f(R_{(x,z,y)}(r) + x) + \big(cA^*A + M_1(t)\big)(R_{(x,z,y)}(r) + x).$$

Using again that $\partial f + cA^*A + M_1(t)$ is $\alpha(t)$-strongly monotone for every $t \in [0, +\infty)$, we obtain

$$\langle M_1(t)(R_{(x,z,y)}(r)) - M_1(r)(R_{(x,z,y)}(r)), R_{(x,z,y)}(r) - R_{(x,z,y)}(t) \rangle \geq \alpha(t)\|R_{(x,z,y)}(r) - R_{(x,z,y)}(t)\|^2.$$

The conclusion follows via the Cauchy-Schwarz inequality.

(ii) From the definition of $Q_{(x,z,y)}$ one has

$$M_2(t)z + cA(\gamma R_{(x,z,y)}(t) + x) + y \in \partial g(Q_{(x,z,y)}(t) + z) + \big(M_2(t) + cI\big)(Q_{(x,z,y)}(t) + z)$$

and

$$M_2(r)z + cA(\gamma R_{(x,z,y)}(r) + x) + y \in \partial g(Q_{(x,z,y)}(r) + z) + \big(M_2(r) + cI\big)(Q_{(x,z,y)}(r) + z),$$

which is equivalent to

$$-M_2(r)(Q_{(x,z,y)}(r)) + M_2(t)(Q_{(x,z,y)}(r) + z) + cA(\gamma R_{(x,z,y)}(r) + x) + y \in$$
$$\partial g(Q_{(x,z,y)}(r) + z) + \big(M_2(t) + cI\big)(Q_{(x,z,y)}(r) + z).$$

Using that $\partial g + M_2(t) + cI$ is $c-$strongly monotone, we obtain

$$\langle M_2(t)(Q_{(x,z,y)}(r)) - M_2(r)(Q_{(x,z,y)}(r)) + c\gamma A(R_{(x,z,y)}(r) - R_{(x,z,y)}(t)), Q_{(x,z,y)}(r) - Q_{(x,z,y)}(t) \rangle \geq$$
$$c\|Q_{(x,z,y)}(r) - Q_{(x,z,y)}(t)\|^2.$$

From the Cauchy-Schwarz inequality and (i) it follows

$$\|Q_{(x,z,y)}(r) - Q_{(x,z,y)}(t)\| \leq \frac{\|Q_{(x,z,y)}(r)\|}{c}\|M_2(t) - M_2(r)\| + \frac{\gamma\|A\|\|R_{(x,z,y)}(r)\|}{\alpha(t)}\|M_1(t) - M_1(r)\|.$$

$\square$

Now we can prove existence and uniqueness of a strong global solution of (4) under $(CStrong)$. To this end we will first reformulate (14) as a particular first order dynamical system in a suitably chosen product space (see also [4]). Subsequently we will make use of the Cauchy-Lipschitz-Picard Theorem for absolutely continues trajectories (see, for example, [24, Proposition 6.2.1], [26, Theorem 54]). Notice that under $(CStrong)$ the operator S_t in Lemma 1 is Lipschitz continuous with constant $\frac{c}{\alpha}$ for every $t \in [0, +\infty)$.

Theorem 3. *Assume that $(Cstrong)$ holds, and $M_1 \in L^1_{loc}([0, +\infty), \mathcal{H})$ and $M_2 \in L^1_{loc}([0, +\infty), \mathcal{G})$, namely,*

$$t \longrightarrow \|M_1(t)\| \text{ and } t \longrightarrow \|M_2(t)\|$$

are integrable on $[0, T]$ for every $T > 0$. Then, for every starting points $(x^0, z^0, y^0) \in \mathcal{H} \times \mathcal{G} \times \mathcal{G}$, the dynamical system (4) has a unique strong global solution $(x, z, y) : [0, +\infty) \longrightarrow \mathcal{H} \times \mathcal{G} \times \mathcal{G}$.

Proof. Denoting $U(t) = (x(t), z(t), y(t))$, the dynamical system (4) can be rewritten as

$$\begin{cases} \dot{U}(t) = \Gamma(t, U(t)) \\ U(0) = (x^0, z^0, y^0), \end{cases} \tag{17}$$

where

$$\Gamma : [0, +\infty) \times \mathcal{H} \times \mathcal{G} \times \mathcal{G} \longrightarrow \mathcal{H} \times \mathcal{G} \times \mathcal{G}, \quad \Gamma(t, x, z, y) = (u, v, w),$$

is defined as

$$u = u(t, x, z, y) = \underset{a \in \mathcal{H}}{\text{argmin}} \left(F(t, a) + \frac{c}{2} \left\| a - \left(\frac{1}{c} M_1(t)x + A^* z - \frac{1}{c} A^* y - \frac{1}{c} \nabla h(x) \right) \right\|^2 \right) - x$$

$$v = v(t, x, z, y) = \underset{b \in \mathcal{G}}{\text{argmin}} \left(G(t, b) + \frac{c}{2} \left\| b - \left(\frac{1}{c} M_2(t)z + A(\gamma u + x) + \frac{1}{c} y \right) \right\|^2 \right) - z$$

$$= \text{prox}_{\frac{1}{c} G(t, \cdot)} \left(\frac{1}{c} M_2(t)z + A(\gamma u + x) + \frac{1}{c} y \right) - z,$$

$$w = w(t, x, z, y) = cA(u + x) - c(v + z).$$

The existence and uniqueness of a strong global solution follows according to the Cauchy-Lipschitz-Picard Theorem, if we show: (1) that $\Gamma(t, \cdot, \cdot, \cdot)$ is $L(t)$-Lipschitz continuous for every $t \in [0, +\infty)$ and the Lipschitz constant as a function of time has the property that $L(\cdot) \in L^1_{loc}([0, +\infty), \mathbb{R})$; (2) that $\Gamma(\cdot, x, z, y) \in L^1_{loc}([0, +\infty), \mathcal{H} \times \mathcal{G} \times \mathcal{G})$ for every $(x, z, y) \in \mathcal{H} \times \mathcal{G} \times \mathcal{G}$.

(1) Let $t \in [0, +\infty)$ be fixed and consider $(x, z, y), (\overline{x}, \overline{z}, \overline{y}) \in \mathcal{H} \times \mathcal{G} \times \mathcal{G}$. We have

$$\|\Gamma(t, x, z, y) - \Gamma(t, \overline{x}, \overline{z}, \overline{y})\| = \sqrt{\|u - \overline{u}\|^2 + \|v - \overline{v}\|^2 + \|w - \overline{w}\|^2},$$

where (see Lemma 1)

$$u - \overline{u} = \underset{a \in \mathcal{H}}{\text{argmin}} \left(F(t, a) + \frac{c}{2} \left\| a - \left(\frac{1}{c} M_1(t)x + A^* z - \frac{1}{c} A^* y - \frac{1}{c} \nabla h(x) \right) \right\|^2 \right)$$

$$- \underset{a \in \mathcal{H}}{\text{argmin}} \left(F(t, a) + \frac{c}{2} \left\| a - \left(\frac{1}{c} M_1(t)\overline{x} + A^* \overline{z} - \frac{1}{c} A^* \overline{y} - \frac{1}{c} \nabla h(\overline{x}) \right) \right\|^2 \right) + \overline{x} - x$$

$$= S_t \left(\frac{1}{c} M_1(t)x + A^* z - \frac{1}{c} A^* y - \frac{1}{c} \nabla h(x) \right) - S_t \left(\frac{1}{c} M_1(t)\overline{x} + A^* \overline{z} - \frac{1}{c} A^* \overline{y} - \frac{1}{c} \nabla h(\overline{x}) \right)$$

$$+ \overline{x} - x.$$

Hence,

$$\|u - \overline{u}\|^2 \le 2 \left\| S_t \left(\frac{M_1(t)}{c} x + A^* z - \frac{A^*}{c} y - \frac{1}{c} \nabla h(x) \right) - S_t \left(\frac{M_1(t)}{c} \overline{x} + A^* \overline{z} - \frac{A^*}{c} \overline{y} - \frac{1}{c} \nabla h(\overline{x}) \right) \right\|^2$$

$$+ 2\|\overline{x} - x\|^2.$$

Using Lemma 1 and taking into account that $(Cstrong)$ is fulfilled, which means that the Lipschitz constant of the operator S_t is $\frac{c}{\alpha}$, it follows

$$\|u - \overline{u}\|^2 \le 2 \frac{c^2}{\alpha^2} \left\| \frac{1}{c} M_1(t)(x - \overline{x}) + A^*(z - \overline{z}) - \frac{1}{c} A^*(y - \overline{y}) - \frac{1}{c} (\nabla h(x) - \nabla h(\overline{x})) \right\|^2 + 2\|\overline{x} - x\|^2$$

$$\le 2 \frac{c^2}{\alpha^2} \left(4 \frac{\|M_1(t)\|^2}{c^2} \|x - \overline{x}\|^2 + 4\|A\|^2 \|z - \overline{z}\|^2 + 4 \frac{\|A\|^2}{c^2} \|y - \overline{y}\|^2 + \frac{4}{c^2} \|\nabla h(x) - \nabla h(\overline{x})\|^2 \right)$$

$$+ 2\|x - \overline{x}\|^2$$

$$\le 2 \left(\frac{4\|M_1(t)\|^2 + 4L_h^2}{\alpha^2} + 1 \right) \|x - \overline{x}\|^2 + \frac{8c^2}{\alpha^2} \|A\|^2 \|z - \overline{z}\|^2 + \frac{8}{\alpha^2} \|A\|^2 \|y - \overline{y}\|^2.$$

13

By taking into account the nonexpansiveness of the proximal operator and that $\gamma \in [0,1]$, it also follows

$$
\begin{aligned}
\|v - \overline{v}\|^2 &\leq 2 \left\| \text{prox}_{\frac{1}{c}G(t,\cdot)}\left(\frac{1}{c}M_2(t)z + A(\gamma u + x) + \frac{1}{c}y\right) - \text{prox}_{\frac{1}{c}G(t,\cdot)}\left(\frac{1}{c}M_2(t)\overline{z} + A(\gamma\overline{u} + \overline{x}) + \frac{1}{c}\overline{y}\right) \right\|^2 \\
&\quad + 2\|z - \overline{z}\|^2 \\
&\leq 2\left\| \frac{1}{c}M_2(t)(z - \overline{z}) + A(\gamma(u - \overline{u}) + x - \overline{x}) + \frac{1}{c}(y - \overline{y}) \right\|^2 + 2\|z - \overline{z}\|^2 \\
&\leq \frac{8\|M_2(t)\|^2}{c^2}\|z - \overline{z}\|^2 + 8\gamma^2\|A\|^2\|u - \overline{u}\|^2 + 8\|A\|^2\|x - \overline{x}\|^2 + \frac{8}{c^2}\|y - \overline{y}\|^2 + 2\|z - \overline{z}\|^2 \\
&\leq 8\|A\|^2\left(\frac{8\|M_1(t)\|^2 + 8L_h^2}{\alpha^2} + 3\right)\|x - \overline{x}\|^2 + \left(\frac{64c^2}{\alpha^2}\|A\|^4 + \frac{8\|M_2(t)\|^2}{c^2} + 2\right)\|z - \overline{z}\|^2 \\
&\quad + \left(\frac{64}{\alpha^2}\|A\|^4 + \frac{8}{c^2}\right)\|y - \overline{y}\|^2.
\end{aligned}
$$

Finally,

$$
\begin{aligned}
\|w - \overline{w}\|^2 &= \|cA(u - \overline{u} + x - \overline{x}) - c(v - \overline{v} + z - \overline{z})\|^2 \\
&\leq 4c^2\|A\|^2\|u - \overline{u}\|^2 + 4c^2\|A\|^2\|x - \overline{x}\|^2 + 4c^2\|v - \overline{v}\|^2 + 4c^2\|z - \overline{z}\|^2 \\
&\leq 36c^2\|A\|^2\left(\frac{8\|M_1(t)\|^2 + 8L_h^2}{\alpha^2} + 3\right)\|x - \overline{x}\|^2 \\
&\quad + 4\left(\frac{72c^4}{\alpha^2}\|A\|^4 + 8\|M_2(t)\|^2 + 3c^2\right)\|z - \overline{z}\|^2 + 32\left(\frac{9c^2}{\alpha^2}\|A\|^4 + 1\right)\|y - \overline{y}\|^2.
\end{aligned}
$$

Consequently,

$$
\begin{aligned}
\|\Gamma(t,x,z,y) - \Gamma(t,\overline{x},\overline{z},\overline{y})\| &\leq \sqrt{L_1(t)\|x - \overline{x}\|^2 + L_2(t)\|z - \overline{z}\|^2 + L_3(t)\|y - \overline{y}\|^2}\| \\
&\leq \sqrt{L_1(t) + L_2(t) + L_3(t)}\sqrt{\|x - \overline{x}\|^2 + \|z - \overline{z}\|^2 + \|y - \overline{y}\|^2} \\
&= L(t)\|(x,z,y) - (\overline{x},\overline{z},\overline{y})\|,
\end{aligned}
$$

where

$$
L(t) = \sqrt{L_1(t) + L_2(t) + L_3(t)}
$$

and

$$
\begin{aligned}
L_1(t) &= 2\left(\frac{4\|M_1(t)\|^2 + 4L_h^2}{\alpha^2} + 1\right) + 8\|A\|^2\left(\frac{8\|M_1(t)\|^2 + 8L_h^2}{\alpha^2} + 3\right) \\
&\quad + 36c^2\|A\|^2\left(\frac{8\|M_1(t)\|^2 + 8L_h^2}{\alpha^2} + 3\right), \\
L_2(t) &= \frac{8c^2}{\alpha^2}\|A\|^2 + \left(\frac{64c^2}{\alpha^2}\|A\|^4 + \frac{8\|M_2(t)\|^2}{c^2} + 2\right) + 4\left(\frac{72c^4}{\alpha^2}\|A\|^4 + 8\|M_2(t)\|^2 + 3c^2\right), \\
L_3(t) &= \frac{8}{\alpha^2}\|A\|^2 + \left(\frac{64}{\alpha^2}\|A\|^4 + \frac{8}{c^2}\right) + 32\left(\frac{9c^2}{\alpha^2}\|A\|^4 + 1\right),
\end{aligned}
$$

which means that $\Gamma(t,\cdot,\cdot,\cdot)$ is $L(t)$-Lipschitz continuous. Since $M_1 \in L^1_{loc}([0,+\infty),\mathcal{H})$ and $M_2 \in L^1_{loc}([0,+\infty),\mathcal{G})$, it is obvious that $L(\cdot) \in L^1_{loc}([0,+\infty),\mathbb{R})$.

(2) Now we will show that $\Gamma(\cdot,x,z,y) \in L^1_{loc}([0,+\infty),\mathcal{H}\times\mathcal{G}\times\mathcal{G})$ for every $(x,z,y) \in \mathcal{H}\times\mathcal{G}\times\mathcal{G}$. Let $(x,z,y) \in \mathcal{H}\times\mathcal{G}\times\mathcal{G}$ be fixed and $T > 0$. We have

$$
\int_0^T \|\Gamma(t,x,z,y)\|dt = \int_0^T \sqrt{\|u(t,x,z,y)\|^2 + \|v(t,x,z,y)\|^2 + \|w(t,x,z,y)\|^2}dt.
$$

14

By Lemma 2 and taking into account that $\alpha(t) = \alpha > 0$ for every $t \in [0, +\infty)$ and $\gamma \in [0, 1]$, we have for every $t \in [0, +\infty)$ that

$$\|u(t, x, z, y)\|^2 \leq 2\|u(t, x, z, y) - u(0, x, z, y)\|^2 + 2\|u(0, x, z, y)\|^2$$
$$\leq \frac{2\|u(0, x, z, y)\|^2}{\alpha^2}\|M_1(t) - M_1(0)\|^2 + 2\|u(0, x, z, y)\|^2,$$

$$\|v(t, x, z, y)\|^2 \leq 2\|v(t, x, z, y) - v(0, x, z, y)\|^2 + 2\|v(0, x, z, y)\|^2$$
$$\leq \frac{4\|v(0, x, z, y)\|^2}{c^2}\|M_2(t) - M_2(0)\|^2 + \frac{4\|A\|^2\|u(0, x, z, y)\|^2}{\alpha^2}\|M_1(t) - M_1(0)\|^2$$
$$+ 2\|v(0, x, z, y)\|^2$$

and

$$\|w(t, x, z, y)\|^2 = c^2\|(Au(t, x, z, y) + x) - (v(t, x, z, y) + z)\|^2$$
$$\leq 3c^2(\|A\|^2\|u(t, x, z, y)\|^2 + \|v(t, x, z, y)\|^2 + \|Ax - z\|^2)$$
$$\leq \frac{18c^2\|A\|^2\|u(0, x, z, y)\|^2}{\alpha^2}\|M_1(t) - M_1(0)\|^2 + 12\|v(0, x, z, y)\|^2\|M_2(t) - M_2(0)\|^2$$
$$+ 3c^2\big(2\|A\|^2\|u(0, x, z, y)\|^2 + 2\|v(0, x, z, y)\|^2 + \|x - z\|^2\big).$$

Since $M_1 \in L^1_{loc}([0, +\infty), \mathcal{H})$ and $M_2 \in L^1_{loc}([0, +\infty), \mathcal{G})$, it follows that the integral

$$\int_0^T \|\Gamma(t, x, z, y)\| dt$$

exists and it is finite, in other words, $\Gamma(\cdot, x, z, y) \in L^1_{loc}([0, +\infty), \mathcal{H} \times \mathcal{G} \times \mathcal{G})$.

Consequently, the dynamical system (17) has a unique locally absolutely continuous solution, which means that the dynamical system (14) has a unique strong global solution. $\qquad\square$

3 Some technical results

In this section we will prove some technical results which will be useful in the asymptotic analysis of the dynamical system (4). We endow the real linear space $\mathcal{L}(\mathcal{H}) := \{A : \mathcal{H} \longrightarrow \mathcal{H} : A \text{ is linear and continuous}\}$ with the norm

$$\|A\| = \sup_{\|x\| \leq 1} \|Ax\|.$$

If $A \in \mathcal{L}(\mathcal{H})$ is self-adjoint, then it holds (see [29, Lemma 3.2.4 iv)])

$$\|A\| = \sup_{\|x\| \leq 1} |\langle Ax, x \rangle|.$$

Definition 3. *We say that the map $M : [0, +\infty) \longrightarrow \mathcal{L}(\mathcal{H})$, $t \longrightarrow M(t)$, is derivable at $t_0 \in [0, +\infty)$, if the limit*

$$\lim_{h \longrightarrow 0} \frac{M(t_0 + h) - M(t_0)}{h}$$

taken with respect to the norm topology of $\mathcal{L}(\mathcal{H})$ exists. In this case we denote by $\dot{M}(t_0) \in \mathcal{L}(\mathcal{H})$ the value of this limit.

If $\dot{M}(t_0)$ exists, for $t_0 \in [0, +\infty)$, then one can easily see that

$$\dot{M}(t_0)x = \lim_{h \longrightarrow 0} \frac{M(t_0 + h)x - M(t_0)x}{h} \quad \text{for every } x \in \mathcal{H}.$$

According to Remark 4, if M is locally absolutely continuous then $\dot{M}(t)$ exists for almost every $t \in [0, +\infty)$.

Assume now that $M(t) \in \mathcal{L}(\mathcal{H})$ is self-adjoint for every $t \in [0, +\infty)$ and that it is derivable at $t_0 \in [0, +\infty)$. For all $x, u \in \mathcal{H}$ we have

$$\langle \dot{M}(t_0)x, u \rangle = \left\langle \lim_{h \longrightarrow 0} \frac{M(t_0 + h)x - M(t_0)x}{h}, u \right\rangle = \lim_{h \longrightarrow 0} \left\langle \frac{M(t_0 + h)x - M(t_0)x}{h}, u \right\rangle$$

$$= \lim_{h \longrightarrow 0} \left\langle x, \frac{M(t_0 + h)u - M(t_0)u}{h} \right\rangle = \langle x, \dot{M}(t_0)u \rangle,$$

which shows that $\dot{M}(t_0)$ is also self-adjoint.

Lemma 4. *Let $M : [0, +\infty) \longrightarrow \mathcal{L}(\mathcal{H})$, $t \longrightarrow M(t)$, be derivable at $t_0 \in [0, +\infty)$, and let the maps $x, y : [0, +\infty) \longrightarrow \mathcal{H}$ be also derivable at t_0. Then the real function $t \longrightarrow \langle M(t)x(t), y(t) \rangle$ is derivable at t_0 and one has*

$$\frac{d}{dt}\langle M(t)x(t), y(t) \rangle\big|_{t=t_0} = \langle \dot{M}(t_0)x(t_0), y(t_0) \rangle + \langle M(t_0)\dot{x}(t_0), y(t_0) \rangle + \langle M(t_0)x(t_0), \dot{y}(t_0) \rangle.$$

Proof. We have

$$\frac{d}{dt}M(t)x(t)\big|_{t=t_0} = \lim_{h \longrightarrow 0} \frac{M(t_0 + h)x(t_0 + h) - M(t_0)x(t_0)}{h}$$

$$= \lim_{h \longrightarrow 0} M(t_0 + h)\left(\frac{x(t_0 + h) - x(t_0)}{h}\right) + \lim_{h \longrightarrow 0} \frac{M(t_0 + h)x(t_0) - M(t_0)x(t_0)}{h}$$

$$= M(t_0)\dot{x}(t_0) + \dot{M}(t_0)x(t_0).$$

The derivation formula of the scalar product leads to the desired conclusion

$$\frac{d}{dt}\langle M(t)x(t), y(t) \rangle\big|_{t=t_0} = \left\langle \frac{d}{dt}M(t)x(t)\big|_{t=t_0}, y(t_0) \right\rangle + \langle M(t_0)x(t_0), \dot{y}(t_0) \rangle$$

$$= \langle M(t_0)\dot{x}(t_0), y(t_0) \rangle + \langle \dot{M}(t_0)x(t_0), y(t_0) \rangle + \langle M(t_0)x(t_0), \dot{y}(t_0) \rangle.$$

$\square$

The main result of this section follows.

Lemma 5. *Assume that (Cstrong) holds and that the maps $M_1 : [0, +\infty) \longrightarrow S_+(\mathcal{H})$ and $M_2 : [0, +\infty) \longrightarrow S_+(\mathcal{G})$ are locally absolutely continuous. For a given starting point $(x^0, z^0, y^0) \in \mathcal{H} \times \mathcal{G} \times \mathcal{G}$, let $(x, z, y) : [0, +\infty) \longrightarrow \mathcal{H} \times \mathcal{G} \times \mathcal{G}$ be the unique strong global solution of the dynamical system (4). Then*

$$t \longrightarrow (\dot{x}(t), \dot{z}(t), \dot{y}(t))$$

is locally absolutely continuous, hence $(\ddot{x}(t), \ddot{z}(t), \ddot{y}(t))$ exists for almost every $t \in [0, +\infty)$.

In addition, if $\sup_{t \geq 0} \|M_1(t)\| < +\infty$ and $\sup_{t \geq 0} \|M_2(t)\| < +\infty$, then there exists $L > 0$ such that

$$\|\ddot{x}(t)\| + \|\ddot{z}(t)\| + \|\ddot{y}(t)\| \leq L\big(\|\dot{x}(t)\| + \|\dot{z}(t)\| + \|\dot{y}(t)\| + \|\dot{M}_1(t)\|\|\dot{x}(t)\| + \|\dot{M}_2(t)\|\|\dot{z}(t)\|\big)$$

for almost every $t \in [0, +\infty)$.

16

Proof. Let $T > 0$ be fixed. We will use the same notations as in the proof of Theorem 3. Let $t, r \in [0, T]$ be fixed. We have

$$\|\dot{U}(t) - \dot{U}(r)\| = \|\Gamma(t, U(t)) - \Gamma(r, U(r))\| \leq \|\Gamma(t, U(t)) - \Gamma(t, U(r))\| + \|\Gamma(t, U(r)) - \Gamma(r, U(r))\|$$
$$\leq \|u(t, x(t), z(t), y(t)) - u(t, x(r), z(r), y(r))\| + \|v(t, x(t), z(t), y(t)) - v(t, x(r), z(r), y(r))\|$$
$$+ \|w(t, x(t), z(t), y(t)) - w(t, x(r), z(r), y(r))\|$$
$$+ \|u(t, x(r), z(r), y(r)) - u(r, x(r), z(r), y(r))\| + \|v(t, x(r), z(r), y(r)) - v(r, x(r), z(r), y(r))\|$$
$$+ \|w(t, x(r), z(r), y(r)) - w(r, x(r), z(r), y(r))\|.$$

Since

$$u(t, x(t), z(t), y(t)) - u(t, x(r), z(r), y(r)) = S_t \left(\frac{1}{c} M_1(t) x(t) + A^* z(t) - \frac{1}{c} A^* y(t) - \frac{1}{c} \nabla h(x(t)) \right)$$
$$- S_t \left(\frac{1}{c} M_1(t) x(r) + A^* z(r) - \frac{1}{c} A^* y(r) - \frac{1}{c} \nabla h(x(r)) \right)$$
$$- x(t) + x(r),$$

according to Lemma 1, we get

$$\|u(t, x(t), z(t), y(t)) - u(t, x(r), z(r), y(r))\|$$
$$\leq \left(\frac{\|M_1(t)\|}{\alpha} + \frac{L_h}{\alpha} + 1 \right) \|x(t) - x(r)\| + \frac{c}{\alpha} \|A\| \|z(t) - z(r)\| + \frac{\|A\|}{\alpha} \|y(t) - y(r)\|.$$

Since $t \longrightarrow \|M_1(t)\|$ is bounded on $[0, T]$, there exists $L_1 := L_1(T) > 0$ such that

$$\|u(t, x(t), z(t), y(t)) - u(t, x(r), z(r), y(r))\| \leq L_1(\|x(t) - x(r)\| + \|z(t) - z(r)\| + \|y(t) - y(r)\|). \quad (18)$$

Similarly, since

$$v(t, x(t), z(t), y(t)) - v(t, x(r), z(r), y(r))$$
$$= \text{prox}_{\frac{1}{c} G(t, \cdot)} \left(\frac{1}{c} M_2(t) z(t) + A(\gamma u(t, x(t), z(t), y(t)) + x(t)) + \frac{1}{c} y(t) \right)$$
$$- \text{prox}_{\frac{1}{c} G(t, \cdot)} \left(\frac{1}{c} M_2(t) z(r) + A(\gamma u(t, x(r), z(r), y(r)) + x(r)) + \frac{1}{c} y(r) \right) - z(t) + z(r),$$

by the nonexpansiveness of the proximal operator we get

$$\|v(t, x(t), z(t), y(t)) - v(t, x(r), z(r), y(r))\|$$
$$\leq \left(\frac{\|M_2(t)\|}{c} + 1 \right) \|z(t) - z(r)\| + \|A\| \|x(t) - x(r)\| + \frac{1}{c} \|y(t) - y(r)\|$$
$$+ \gamma \|A\| \|u(t, x(t), z(t), y(t)) - u(t, x(r), z(r), y(r))\|.$$

Since $t \longrightarrow \|M_2(t)\|$ is bounded on $[0, T]$, by taking into consideration (18), one can easily see that there exists $L_2 := L_2(T) > 0$ such that

$$\|v(t, x(t), z(t), y(t)) - v(t, x(r), z(r), y(r))\| \leq L_2(\|x(t) - x(r)\| + \|z(t) - z(r)\| + \|y(t) - y(r)\|). \quad (19)$$

Further, by using (18) and (19), we get

$$\|w(t, x(t), z(t), y(t)) - w(t, x(r), z(r), y(r))\|$$
$$\leq c \|A(u(t, x(t), z(t), y(t)) - u(t, x(r), z(r), y(r)) + x(t) - x(r))\|$$
$$+ c \|v(t, x(t), z(t), y(t)) - v(t, x(r), z(r), y(r)) + z(t) - z(r)\|$$
$$\leq c(\|A\| L_1 + \|A\| + L_2) \|x(t) - x(r)\| + c(\|A\| L_1 + L_2 + 1) \|z(t) - z(r)\|$$
$$+ c(\|A\| L_1 + L_2) \|y(t) - y(r)\|.$$

17

Hence, there exists $L_3 := c(\|A\|L_1 + \|A\| + L_2 + 1) > 0$ such that

$$\|w(t, x(t), z(t), y(t)) - w(t, x(r), z(r), y(r))\| \leq L_3(\|x(t) - x(r)\| + \|z(t) - z(r)\| + \|y(t) - y(r)\|). \quad (20)$$

Using now Lemma 2 (i), we get

$$\|u(t, x(r), z(r), y(r)) - u(r, x(r), z(r), y(r))\| = \|R_{(x(r), z(r), y(r))}(t) - R_{(x(r), z(r), y(r))}(r)\| \atop \leq \frac{\|R_{(x(r), z(r), y(r))}(r)\|}{\alpha} \|M_1(t) - M_1(r)\|. \quad (21)$$

Since $r \mapsto S_r$ and ∇h are Lipschitz continuous and x, z, y and M_1 are absolutely continuous, the map

$$r \mapsto R_{(x(r), z(r), y(r))}(r) = S_r\left(\frac{1}{c}M_1(r)x(r) + A^*z(r) - \frac{1}{c}A^*y(r) - \frac{1}{c}\nabla h(x(r))\right) - x(r)$$

is bounded on $[0, T]$. Consequently, there exists $L_4 := L_4(T) > 0$ such that

$$\|u(t, x(r), z(r), y(r)) - u(r, x(r), z(r), y(r))\| \leq L_4\|M_1(t) - M_1(r)\|. \quad (22)$$

Similarly, using this time Lemma 2 (ii), we get

$$\|v(t, x(r), z(r), y(r)) - v(r, x(r), z(r), y(r))\| = \|Q_{(x(r), z(r), y(r))}(t) - Q_{(x(r), z(r), y(r))}(r)\| \atop \leq \frac{\|A\|\|R_{(x(r), z(r), y(r))}(r)\|}{\alpha}\|M_1(t) - M_1(r)\| + \frac{\|Q_{(x(r), z(r), y(r))}(r)\|}{c}\|M_2(t) - M_2(r)\|. \quad (23)$$

Since the proximal operator is nonexpansive and x, z, y and M_2 are absolutely continuous, the map

$$r \mapsto Q_{(x(r), z(r), y(r))}(r) = \text{prox}_{\frac{1}{c}G(r, \cdot)}\left(\frac{1}{c}M_2(r)z(r) + A(\gamma u(r, x(r), z(r), y(r)) + x(r)) + \frac{1}{c}y(r)\right) - z(r)$$

is bounded on $[0, T]$. Consequently, there exists $L_5 := L_5(T) > 0$ such that

$$\|v(t, x(r), z(r), y(r)) - v(r, x(r), z(r), y(r))\| \leq L_5(\|M_1(t) - M_1(r)\| + \|M_2(t) - M_2(r)\|). \quad (24)$$

Further, by using (22) and (24), we get

$$\|w(t, x(r), z(r), y(r)) - w(r, x(r), z(r), y(r))\|$$
$$\leq c\|A(u(t, x(r), z(r), y(r)) - u(r, x(r), z(r), y(r)))\| + c\|v(t, x(r), z(r), y(r)) - v(r, x(r), z(r), y(r))\|$$
$$\leq c(\|A\|L_4 + L_5)\|M_1(t) - M_1(r)\| + cL_5\|M_2(t) - M_2(r)\|$$

Consequently, there exists $L_6 := c(\|A\|L_4 + L_5) > 0$ such that

$$\|w(t, x(r), z(r), y(r)) - w(r, x(r), z(r), y(r))\| \leq L_6(\|M_1(t) - M_1(r)\| + \|M_2(t) - M_2(r)\|). \quad (25)$$

Summing the relations (18)-(25) we obtain that there exists $L_7 > 0$ such that

$$\|\dot{U}(t) - \dot{U}(r)\|$$
$$\leq L_7(\|x(t) - x(r)\| + \|z(t) - z(r)\| + \|y(t) - y(r)\| + \|M_1(t) - M_1(r)\| + \|M_2(t) - M_2(r)\|).$$

Let be $\epsilon > 0$. Since the maps x, z, y, M_1 and M_2 are absolutely continuous on $[0, T]$, there exists $\eta > 0$ such that for any finite family of intervals $I_k = (a_k, b_k) \subseteq [0, T]$ such that for any subfamily of disjoint intervals I_j with $\sum_j |b_j - a_j| < \eta$ it holds

$$\sum_j \|x(b_j) - x(a_j)\| < \frac{\varepsilon}{5L_7}, \sum_j \|z(b_j) - z(a_j)\| < \frac{\varepsilon}{5L_7}, \sum_j \|y(b_j) - y(a_j)\| < \frac{\varepsilon}{5L_7},$$

$$\sum_j \|M_1(b_j) - M_1(a_j)\| < \frac{\varepsilon}{5L_7} \quad \text{and} \quad \sum_j \|M_2(b_j) - M_2(a_j)\| < \frac{\varepsilon}{5L_7}.$$

Consequently,

$$\sum_j \|\dot{U}(b_j) - \dot{U}(a_j)\| < \varepsilon,$$

hence $\dot{U}(\cdot) = (\dot{x}(\cdot), \dot{z}(\cdot), \dot{y}(\cdot))$ is absolutely continuous on $[0, T]$. This proves that the second order derivatives $\ddot{x}, \ddot{z}, \ddot{y}$ exist almost everywhere on $[0, +\infty)$.

We come now to the proof of the second statement and assume to this end that $\sup_{t \geq 0} \|M_1(t)\| < +\infty$ and $\sup_{t \geq 0} \|M_2(t)\| < +\infty$. Under these assumption, L_1, L_2 and L_3 appearing in (18), (19) and (20), respectively, can be taken as being global constants, that is, (18)-(20) hold for every $t, r \in [0, +\infty)$.

Since $R_{(x(r), z(r), y(r))}(r) = \dot{x}(r)$ and $Q_{(x(r), z(r), y(r))}(r) = \dot{z}(r)$ for every $r \in [0, +\infty)$, from (21) and (23) we get

$$\|u(t, x(r), z(r), y(r)) - u(r, x(r), z(r), y(r))\| \leq \frac{\|\dot{x}(r)\|}{\alpha} \|M_1(t) - M_1(r)\|$$

and, respectively,

$$\|v(t, x(r), z(r), y(r)) - v(r, x(r), z(r), y(r))\| \leq \frac{\|A\| \|\dot{x}(r)\|}{\alpha} \|M_1(t) - M_1(r)\| + \frac{\|\dot{z}(r)\|}{c} \|M_2(t) - M_2(r)\|$$

for every $t, r \in [0, +\infty)$. Consequently,

$$\|w(t, x(r), z(r), y(r)) - w(r, x(r), z(r), y(r))\|$$
$$\leq c\|A\| \|u(t, x(r), z(r), y(r)) - u(r, x(r), z(r), y(r))\| + c\|v(t, x(r), z(r), y(r)) - v(r, x(r), z(r), y(r))\|$$
$$\leq \frac{2c\|A\|}{\alpha} \|\dot{x}(r)\| \|M_1(t) - M_1(r)\| + \|\dot{z}(r)\| \|M_2(t) - M_2(r)\|$$

for every $t, r \in [0, +\infty)$. This shows that there exists $L > 0$ such that

$$\|\dot{U}(t) - \dot{U}(r)\| \leq$$
$$\frac{L}{\sqrt{3}} (\|x(t) - x(r)\| + \|z(t) - z(r)\| + \|y(t) - y(r)\| + \|\dot{x}(r)\| \|M_1(t) - M_1(r)\| + \|\dot{z}(r)\| \|M_2(t) - M_2(r)\|)$$

for every $t, r \in [0, +\infty)$.

Now we fix $r \in [0, +\infty)$ at which the second derivative of the trajectories exist and take in the above inequality $t = r + h$ for some $h > 0$. This yields

$$\|\dot{x}(r+h) - \dot{x}(r)\| + \|\dot{z}(r+h) - \dot{z}(r)\| + \|\dot{y}(r+h) - \dot{y}(r)\|) \leq \sqrt{3} \|\dot{U}(r+h) - \dot{U}(r)\|$$
$$\leq L(\|x(r+h) - x(r)\| + \|z(r+h) - z(r)\| + \|y(r+h) - y(r)\|)$$
$$+ L(\|\dot{x}(r)\| \|M_1(r+h) - M_1(r)\| + \|\dot{z}(r)\| \|M_2(r+h) - M_2(r)\|).$$

After dividing in the above inequality by h and letting $h \longrightarrow 0$, we obtain

$$\|\ddot{x}(r)\| + \|\ddot{z}(r)\| + \|\ddot{y}(r)\| \leq L(\|\dot{x}(r)\| + \|\dot{z}(r)\| + \|\dot{y}(r)\| + \|\dot{x}(r)\| \|\dot{M_1}(r)\| + \|\dot{z}(r)\| \|\dot{M_2}(r)\|).$$

This inequality holds for almost every $r \in [0, +\infty)$. $\qquad \square$

4 Asymptotic analysis

In this section we will address the asymptotic behaviour of the trajectories generated by the dynamical system (4). At the beginning we will recall two results which will play a central role in the asymptotic analysis (see [2, Lemma 5.1] and [2, Lemma 5.2], respectively).

Lemma 6. *Suppose that $A : [0, +\infty) \to \mathbb{R}$ is locally absolutely continuous and bounded from below and that there exists $B \in L^1([0, +\infty), \mathbb{R})$ such that for almost every $t \in [0, +\infty)$*

$$\frac{d}{dt} A(t) \leq B(t).$$

Then there exists $\lim_{t \to +\infty} A(t) \in \mathbb{R}$.

Lemma 7. *If $1 \leq p < \infty$, $1 \leq r \leq \infty$, $A : [0, +\infty) \to [0, +\infty)$ is locally absolutely continuous, $A \in L^p([0, +\infty), \mathbb{R})$, $B : [0, +\infty) \to \mathbb{R}$, $B \in L^r([0, +\infty), \mathbb{R})$ and for almost every $t \in [0, +\infty)$*

$$\frac{d}{dt} A(t) \leq B(t),$$

then $\lim_{t \to +\infty} A(t) = 0$.

The first result which we prove in this section is a continuous version of the Opial Lemma formulated in the setting of variable metrics (see [18, Theorem 3.3] for its discrete counterpart).

Lemma 8. *Let $\mathcal{C} \subseteq \mathcal{H}$ be a nonempty set and $x : [0, +\infty) \to \mathcal{H}$ a continuous map. Let $M : [0, +\infty) \longrightarrow S_+(\mathcal{H})$ be such that $M(t_1) \succcurlyeq M(t_2)$ for every $t_1, t_2 \in [0, +\infty)$ with $t_1 \leq t_2$ and there exists $\alpha > 0$ with $M(t) \in P_\alpha(\mathcal{H})$ for every $t \in [0, +\infty)$. If the following two conditions are fulliled*
  (i) the limit $\lim_{t \to +\infty} \|x(t) - z\|_{M(t)}$ exists for every $z \in \mathcal{C}$;
  (ii) every weak sequential cluster point of $x(t), t \in [0, +\infty)$, belongs to $\mathcal{C}$;
then there exists $x_\infty \in \mathcal{C}$ such that $x(t), t \in [0, +\infty)$, converges weakly to x_∞ as $t \to +\infty$.

Proof. Since $\mathcal{C} \neq \emptyset$ and $M(t) \in P_\alpha(\mathcal{H})$, by (i) we have that x is bounded, hence it possesses at least one weak sequential cluster point, which belongs to $\mathcal{C}$. We show that x has exactly one weak sequential cluster point.

Indeed, let x_1, x_2 two weak sequential cluster points of x. For our claim it is enough to show that $x_1 = x_2$. Obviously $x_1, x_2 \in \mathcal{C}$ and there exist the sequences $(t_n^1)_{n \geq 0}, (t_n^2)_{n \geq 0} \subseteq [0, +\infty)$ with $\lim_{n \to +\infty} t_n^1 = +\infty$ and $\lim_{n \to +\infty} t_n^2 = +\infty$ such that $(x(t_n^1))_{n \geq 0}$ converges weakly to x_1 and $(x(t_n^2))_{n \geq 0}$ converges weakly to x_2 as $n \to +\infty$.

Further, since $M(t_1) \succcurlyeq M(t_2)$ for every for every $t_1, t_2 \in [0, +\infty)$ with $t_1 \leq t_2$ and $M(t) \in P_\alpha(\mathcal{H})$ for every $t \in [0, +\infty)$, it follows that for every $z \in \mathcal{H}$ the function

$$[0, +\infty) \to [0, +\infty), t \longrightarrow \|z\|_{M(t)}^2,$$

is decreasing and is bounded from below, hence there exists

$$\lim_{t \to +\infty} \|z\|_{M(t)}^2 \in \mathbb{R}. \tag{26}$$

Since $x_1, x_2 \in \mathcal{C}$, we have that the limits $\lim_{t \to +\infty} \|x(t) - x_1\|_{M(t)}^2$ and $\lim_{t \to +\infty} \|x(t) - x_2\|_{M(t)}^2$ exist. Further, since

$$-\langle x(t), M(t)(x_1 - x_2)\rangle = \frac{1}{2} \left(\|x(t) - x_1\|_{M(t)}^2 - \|x(t) - x_2\|_{M(t)}^2 - \|x_1\|_{M(t)}^2 + \|x_2\|_{M(t)}^2 \right)$$

holds for every $t \in [0, +\infty)$, the limit

$$\lambda := \lim_{t \longrightarrow +\infty} \langle x(t), M(t)(x_1 - x_2) \rangle \in \mathbb{R} \tag{27}$$

exists.

Next we show that the limits

$$\lim_{t \longrightarrow +\infty} M(t)z \tag{28}$$

exists for every $z \in \mathcal{H}$. To this end we fix $z \in \mathcal{H}$. We will actually show that

$$\lim_{s,t \to +\infty} \|M(t)z - M(s)z\| = 0$$

and the conclusion will follow by the Cauchy criterion.

For $U \in S_+(\mathcal{H})$ we have by the generalized Cauchy-Schwarz inequality that

$$|\langle Ux, z \rangle| \le \|x\|_U \|z\|_U \text{ for every } x, z \in H.$$

Hence, for $t, s \in [0, +\infty)$ with $t \le s$ we have $M(t) - M(s) \in S_+(\mathcal{H})$, therefore

$$
\begin{aligned}
\|(M(t) - M(s))z\|^2 &= \langle (M(t) - M(s))z, M(t) - M(s))z \rangle \\
&\le \|z\|_{(M(t)-M(s))} \|(M(t) - M(s))z\|_{(M(t)-M(s))} \\
&= \|z\|_{(M(t)-M(s))} \left(\langle (M(t) - M(s))^2 z, (M(t) - M(s))z \rangle \right)^{\frac{1}{2}} \\
&\le \|z\|_{(M(t)-M(s))} \|M(t) - M(s)\|^{\frac{3}{2}} \|z\|.
\end{aligned}
$$

Since $M(0) \succcurlyeq M(t)$, we have that

$$\|M(t)\| = \sup_{\|x\|=1} \langle M(t)x, x \rangle \le \sup_{\|x\|=1} \langle M(0)x, x \rangle \le \|M(0)\|$$

for every $t \in [0, +\infty)$. This shows that $\|M(t) - M(s)\|, t, s \in [0, +\infty)$, is bounded. This, together with the fact that $\lim_{s,t \to +\infty} \|z\|^2_{(M(t)-M(s))} = 0$, which follows from (26), implies

$$\|(M(t) - M(s))z\| \longrightarrow 0 \text{ as } s, t \longrightarrow +\infty.$$

This proves (28). For every $z \in \mathcal{H}$ let us denote by $Mz := \lim_{t \longrightarrow +\infty} M(t)z$. Since $M(t) \in P_\alpha(\mathcal{H})$ for every $t \in [0, +\infty)$, it holds

$$M \in P_\alpha(\mathcal{H}).$$

Since $(x(t_n^1))_{n \ge 0}$ converges weakly to x_1 and $(x(t_n^2))_{n \ge 0}$ converges weakly to x_2 as $n \to +\infty$ and

$$M(t_n^1)(x_2 - x_1) \longrightarrow M(x_2 - x_1) \text{ and } M(t_n^2)(x_2 - x_1) \longrightarrow M(x_2 - x_1) \text{ as } n \longrightarrow +\infty,$$

passing to the limit in (27) we get

$$\lim_{n \longrightarrow +\infty} \langle x(t_n^1), M(t_n^1)(x_2 - x_1) \rangle = \langle x_1, M(x_2 - x_1) \rangle = \lambda$$

and

$$\lim_{n \longrightarrow +\infty} \langle x(t_n^2), M(t_n^2)(x_2 - x_1) \rangle = \langle x_2, M(x_2 - x_1) \rangle = \lambda.$$

In conclusion,

$$0 = \langle x_2, M(x_2 - x_1) \rangle - \langle x_1, M(x_2 - x_1) \rangle = \|x_2 - x_1\|_M^2 \ge \alpha \|x_2 - x_1\|^2,$$

which shows that $x_1 = x_2$. $\qquad\square$

Remark 5. If a map $M : [0, +\infty) \longrightarrow S_+(\mathcal{H})$ satisfies $M(t_1) \succcurlyeq M(t_2)$ for every $t_1, t_2 \in [0, +\infty)$ with $t_1 \leq t_2$ we say that M is monotonically decreasing. If M is monotonically decreasing and locally absolutely continuous, then $\dot{M}(t)$ exists and $\langle \dot{M}(t)x, x \rangle \leq 0$ for almost every $t \in [0, +\infty)$.

The following result is an adaptation of a result from [3] to our setting.

Proposition 9. *(see [3, Proposition 2.4]) In the setting of the optimization problem* (1), *let* $(a_n, a_n^*)_{n \geq 0}$ *be a sequence in the graph of* $\partial(f + h)$ *and* $(b_n, b_n^*)_{n \geq 0}$ *a sequence in the graph of* ∂g. *Suppose that* a_n *converges weakly to* $\overline{x} \in \mathcal{H}$, b_n^* *converges weakly to* $\overline{v} \in \mathcal{G}$, $a_n^* + A^* b_n^* \longrightarrow 0$, *and* $Aa_n - b_n \longrightarrow 0$ *as* $n \to +\infty$. *Then*

$$\langle a_n, a_n^* \rangle + \langle b_n, b_n^* \rangle \longrightarrow 0 \ as \ n \to +\infty$$

and

$$\overline{v} \in \partial g(A\overline{x}), \ -A^*\overline{v} \in \partial f(\overline{x}) + \nabla h(\overline{x}).$$

The theorem which states the asymptotic convergence of the trajectories generated by the dynamical system (4) to a saddle point of the Lagrangian of the problem (1) follows.

Theorem 10. *In the setting of the optimization problem* (1), *assume that the set of saddle points of the Lagrangian* l *is nonempty, the maps*

$$[0, +\infty) \to S_+(\mathcal{H}), t \mapsto M_1(t), \ and \ [0, +\infty) \to S_+(\mathcal{G}), t \mapsto M_2(t),$$

are locally absolutely continuous and monotonically decreasing,

$$M_1(t) + \frac{c(1-\gamma)}{4}A^*A - \frac{L_h}{4}I \in S_+(\mathcal{H}) \quad \forall t \in [0, +\infty),$$

and

$$\sup_{t \geq 0} \|\dot{M}_1(t)\| < +\infty \ and \ \sup_{t \geq 0} \|\dot{M}_2(t)\| < +\infty.$$

For an arbitrary starting point $(x^0, z^0, y^0) \in \mathcal{H} \times \mathcal{G} \times \mathcal{G}$, *let* $(x, z, y) : [0, +\infty) \longrightarrow \mathcal{H} \times \mathcal{G} \times \mathcal{G}$ *be the unique strong global solution of the dynamical system* (4). *If one of the following conditions holds:*

(I) *there exists* $\alpha > 0$ *such that* $M_1(t) + \frac{c(1-\gamma)}{4}A^*A - \frac{L_h}{4}I \in P_\alpha(\mathcal{H})$ *for every* $t \in [0, +\infty)$;

(II) $\gamma \in [0, 1)$ *and there exists* $\alpha > 0$ *such that* $A^*A \in P_\alpha(\mathcal{H})$;

then the trajectory $(x(t), z(t), y(t))$ *converges weakly to a saddle point of* l *as* $t \longrightarrow +\infty$.

Proof. The proof of the theorem relies on Lemma 8. An important step in the proof will be the derivation of two inequalities of Lyapunov type, namely, (37), in the case when $L_h \neq 0$, and (43), in the case when $L_h = 0$. Let $(x^*, z^*, y^*) \in \mathcal{H} \times \mathcal{G} \times \mathcal{G}$ be a saddle point of the Lagrangian l. Then

$$\begin{cases} 0 \in \partial f(x^*) + \nabla h(x^*) + A^*y^* \\ Ax^* = z^*, Ax^* \in \partial g^*(y^*). \end{cases}$$

According to (5) we have for almost every $t \in [0, +\infty)$

$$-cA^*A(\dot{x}(t) + x(t)) - M_1(t)\dot{x}(t) + cA^*z(t) - A^*y(t) - \nabla h(x(t)) \in \partial f(\dot{x}(t) + x(t)), \quad (29)$$

which yields, by taking into account the monotonicity of ∂f,

$$\langle -cA^*A(\dot{x}(t)+x(t)) - M_1(t)\dot{x}(t) + cA^*z(t) - A^*(y(t)-y^*) - (\nabla h(x(t)) - \nabla h(x^*)), \dot{x}(t)+x(t)-x^* \rangle \geq 0. \quad (30)$$

22

Similarly, according to (6) we have for almost every $t \in [0, +\infty)$

$$-c(\dot{z}(t) + z(t)) + cA(\gamma\dot{x}(t) + x(t)) - M_2(t)\dot{z}(t) + y(t) \in \partial g(\dot{z}(t) + z(t)), \tag{31}$$

which yields, by taking into account the monotonicity of ∂g,

$$\langle -c(\dot{z}(t) + z(t)) + cA(\gamma\dot{x}(t) + x(t)) - M_2(t)\dot{z}(t) + (y(t) - y^*), \dot{z}(t) + z(t) - Ax^* \rangle \geq 0. \tag{32}$$

By using the last equation of (4) we obtain for almost every $t \in [0, +\infty)$

$$
\langle -A^*(y(t) - y^*), \dot{x}(t) + x(t) - x^* \rangle + \langle y(t) - y^*, \dot{z}(t) + z(t) - Ax^* \rangle
$$
$$
= -\langle y(t) - y^*, A(\dot{x}(t) + x(t)) - Ax^* - (\dot{z}(t) + z(t)) + Ax^* \rangle = -\frac{1}{c}\langle y(t) - y^*, \dot{y}(t) \rangle \tag{33}
$$
$$
= -\frac{1}{2c}\frac{d}{dt}\|y(t) - y^*\|^2.
$$

Assume that $L_h > 0$. By using the Baillon-Haddad Theorem we have for almost every $t \in [0, +\infty)$

$$
\langle -(\nabla h(x(t)) - \nabla h(x^*)), \dot{x}(t) + x(t) - x^* \rangle
$$
$$
= -\langle \nabla h(x(t)) - \nabla h(x^*), x(t) - x^* \rangle - \langle \nabla h(x(t)) - \nabla h(x^*), \dot{x}(t) \rangle
$$
$$
\leq -\frac{1}{L_h}\|\nabla h(x(t)) - \nabla h(x^*)\|^2 - \langle \nabla h(x(t)) - \nabla h(x^*), \dot{x}(t) \rangle \tag{34}
$$
$$
= -\frac{1}{L_h}\left(\left\| \nabla h(x(t)) - \nabla h(x^*) + \frac{L_h}{2}\dot{x}(t) \right\|^2 - \frac{L_h^2}{4}\|\dot{x}(t)\|^2 \right).
$$

By summing (30) and (32) and by taking into account (33) and (34) we obtain for almost every $t \in [0, +\infty)$

$$
0 \leq \langle -cA^*A(\dot{x}(t) + x(t)) - M_1(t)\dot{x}(t) + cA^*z(t), \dot{x}(t) + x(t) - x^* \rangle
$$
$$
+ \langle -c(\dot{z}(t) + z(t)) + cAx(t) + c\gamma A\dot{x}(t) - M_2(t)\dot{z}(t), \dot{z}(t) + z(t) - Ax^* \rangle \tag{35}
$$
$$
-\frac{1}{2c}\frac{d}{dt}\|y(t) - y^*\|^2 - \frac{1}{L_h}\left(\left\| \nabla h(x(t)) - \nabla h(x^*) + \frac{L_h}{2}\dot{x}(t) \right\|^2 - \frac{L_h^2}{4}\|\dot{x}(t)\|^2 \right).
$$

We have for almost every $t \in [0, +\infty)$

$$
\langle -cA^*A(\dot{x}(t) + x(t)) + cA^*z(t), \dot{x}(t) + x(t) - x^* \rangle
$$
$$
+ \langle -c(\dot{z}(t) + z(t)) + cAx(t) + c\gamma A\dot{x}(t), \dot{z}(t) + z(t) - Ax^* \rangle
$$
$$
= -\frac{1}{c}\|\dot{y}(t)\|^2 + \langle -c\dot{z}(t), A(\dot{x}(t) + x(t) - x^*) \rangle + \langle (\gamma - 1)cA\dot{x}(t), \dot{z}(t) + z(t) - Ax^* \rangle
$$
$$
= -\frac{1}{c}\|\dot{y}(t)\|^2 + \left\langle -c\dot{z}(t), \frac{1}{c}\dot{y}(t) + \dot{z}(t) + z(t) - Ax^* \right\rangle
$$
$$
+ \left\langle (\gamma - 1)cA\dot{x}(t), A(\dot{x}(t) + x(t)) - \frac{1}{c}\dot{y}(t) - Ax^* \right\rangle
$$
$$
= -\frac{1}{c}\|\dot{y}(t)\|^2 - c\|\dot{z}(t)\|^2 + (\gamma - 1)c\|A\dot{x}(t)\|^2 - \langle \dot{z}(t), \dot{y}(t) \rangle + (1 - \gamma)\langle A\dot{x}(t), \dot{y}(t) \rangle
$$
$$
+ \frac{c(\gamma - 1)}{2}\frac{d}{dt}\left(\|Ax(t) - Ax^*\|^2 \right) - \frac{c}{2}\frac{d}{dt}\left(\|z(t) - Ax^*\|^2 \right).
$$

Since

$$
\langle \dot{z}(t), \dot{y}(t) \rangle = \left\| \frac{\sqrt{3c}}{2}\dot{z}(t) + \frac{1}{\sqrt{3c}}\dot{y}(t) \right\|^2 - \frac{3c}{4}\|\dot{z}(t)\|^2 - \frac{1}{3c}\|\dot{y}(t)\|^2,
$$

23

and

$$\langle A\dot{x}(t), \dot{y}(t) \rangle = -\left\|\frac{\sqrt{3c}}{2}A\dot{x}(t) - \frac{1}{\sqrt{3c}}\dot{y}(t)\right\|^2 + \frac{3c}{4}\|A\dot{x}(t)\|^2 + \frac{1}{3c}\|\dot{y}(t)\|^2,$$

we obtain from above that for almost every $t \in [0, +\infty)$ it holds

$$
\begin{aligned}
&\langle -cA^*A(\dot{x}(t) + x(t)) + cA^*z(t), \dot{x}(t) + x(t) - x^* \rangle \\
&\quad + \langle -c(\dot{z}(t) + z(t)) + cAx(t) + c\gamma A\dot{x}(t), \dot{z}(t) + z(t) - Ax^* \rangle \\
&= -\frac{\gamma+1}{3c}\|\dot{y}(t)\|^2 - \frac{c}{4}\|\dot{z}(t)\|^2 - \frac{(1-\gamma)c}{4}\|A\dot{x}(t)\|^2 - \left\|\frac{\sqrt{3c}}{2}\dot{z}(t) + \frac{1}{\sqrt{3c}}\dot{y}(t)\right\|^2 \\
&\quad - (1-\gamma)\left\|\frac{\sqrt{3c}}{2}A\dot{x}(t) - \frac{1}{\sqrt{3c}}\dot{y}(t)\right\|^2 + \frac{c(\gamma-1)}{2}\frac{d}{dt}\left(\|Ax(t) - Ax^*\|^2\right) - \frac{c}{2}\frac{d}{dt}\left(\|z(t) - Ax^*\|^2\right).
\end{aligned}
\tag{36}
$$

By using Lemma 4 we observe that for almost every $t \in [0, +\infty)$ it holds

$$
\begin{aligned}
\langle -M_1(t)\dot{x}(t), \dot{x}(t) + x(t) - x^* \rangle &= -\|\dot{x}(t)\|^2_{M_1(t)} - \langle M_1(t)\dot{x}(t), x(t) - x^* \rangle \\
&= -\|\dot{x}(t)\|^2_{M_1(t)} + \frac{1}{2}\langle \dot{M}_1(t)(x(t) - x^*), x(t) - x^* \rangle \\
&\quad - \frac{1}{2}\frac{d}{dt}\|x(t) - x^*\|^2_{M_1(t)}
\end{aligned}
$$

and

$$
\begin{aligned}
\langle -M_2(t)\dot{z}(t), \dot{z}(t) + z(t) - Ax^* \rangle &= -\|\dot{z}(t)\|^2_{M_2(t)} - \langle M_2(t)\dot{z}(t), z(t) - Ax^* \rangle \\
&= -\|\dot{z}(t)\|^2_{M_2(t)} + \frac{1}{2}\langle \dot{M}_2(t)(z(t) - Ax^*), z(t) - Ax^* \rangle \\
&\quad - \frac{1}{2}\frac{d}{dt}\|z(t) - Ax^*\|^2_{M_2(t)}.
\end{aligned}
$$

By plugging the last two identities and (36) into (35), we obtain for almost every $t \in [0, +\infty)$

$$
\begin{aligned}
0 \leq & -\frac{\gamma+1}{3c}\|\dot{y}(t)\|^2 - \frac{c}{4}\|\dot{z}(t)\|^2 - \frac{(1-\gamma)c}{4}\|A\dot{x}(t)\|^2 - \left\|\frac{\sqrt{3c}}{2}\dot{z}(t) + \frac{1}{\sqrt{3c}}\dot{y}(t)\right\|^2 \\
& - (1-\gamma)\left\|\frac{\sqrt{3c}}{2}A\dot{x}(t) - \frac{1}{\sqrt{3c}}\dot{y}(t)\right\|^2 + \frac{c(\gamma-1)}{2}\frac{d}{dt}\left(\|Ax(t) - Ax^*\|^2\right) - \frac{c}{2}\frac{d}{dt}\left(\|z(t) - Ax^*\|^2\right) \\
& - \|\dot{x}(t)\|^2_{M_1(t)} + \frac{1}{2}\langle \dot{M}_1(t)(x(t) - x^*), x(t) - x^* \rangle - \frac{1}{2}\frac{d}{dt}\|x(t) - x^*\|^2_{M_1(t)} \\
& - \|\dot{z}(t)\|^2_{M_2(t)} + \frac{1}{2}\langle \dot{M}_2(t)(z(t) - Ax^*), z(t) - Ax^* \rangle - \frac{1}{2}\frac{d}{dt}\|z(t) - Ax^*\|^2_{M_2(t)} \\
& - \frac{1}{2c}\frac{d}{dt}\|y(t) - y^*\|^2 - \frac{1}{L_h}\left(\left\|\nabla h(x(t)) - \nabla h(x^*) + \frac{L_h}{2}\dot{x}(t)\right\|^2 - \frac{L_h^2}{4}\|\dot{x}(t)\|^2\right).
\end{aligned}
$$

According to Remark 5,

$$\langle \dot{M}_1(t)(x(t) - x^*), x(t) - x^* \rangle \leq 0 \text{ and } \langle \dot{M}_2(t)(z(t) - Ax^*), z(t) - Ax^* \rangle \leq 0$$

24

for almost every $t \in [0, +\infty)$. This means that for almost every $t \in [0, +\infty)$ we have

$$
\frac{1}{2}\frac{d}{dt}\left(\|x(t) - x^*\|^2_{M_1(t)+c(1-\gamma)A^*A} + \|z(t) - Ax^*\|^2_{M_2(t)+cI} + \frac{1}{c}\|y(t) - y^*\|^2\right) +
$$

$$
\|\dot{x}(t)\|^2_{M_1(t)+\frac{(1-\gamma)c}{4}A^*A-\frac{L_h}{4}I} + \|\dot{z}(t)\|^2_{M_2(t)+\frac{c}{4}I} + \frac{\gamma+1}{3c}\|\dot{y}(t)\|^2 +
$$

$$
\left\|\frac{\sqrt{3c}}{2}\dot{z}(t) + \frac{1}{\sqrt{3c}}\dot{y}(t)\right\|^2 + (1-\gamma)\left\|\frac{\sqrt{3c}}{2}A\dot{x}(t) - \frac{1}{\sqrt{3c}}\dot{y}(t)\right\|^2 + \tag{37}
$$

$$
\frac{1}{L_h}\left\|\nabla h(x(t)) - \nabla h(x^*) + \frac{L_h}{2}\dot{x}(t)\right\|^2 \leq 0.
$$

From Lemma 6 we have

$$
\lim_{t \to +\infty}\left(\|x(t) - x^*\|^2_{M_1(t)+c(1-\gamma)A^*A} + \|z(t) - Ax^*\|^2_{M_2(t)+cI} + \frac{1}{c}\|y(t) - y^*\|^2\right) \in \mathbb{R}. \tag{38}
$$

Let be $T > 0$. By integrating (37) on the interval $[0, T]$ we obtain

$$
\frac{1}{2}\left(\|x(T) - x^*\|^2_{M_1(T)+c(1-\gamma)A^*A} + \|z(T) - z^*\|^2_{M_2(T)+cI} + \frac{1}{c}\|y(T) - y^*\|^2\right) +
$$

$$
\int_0^T \|\dot{x}(t)\|^2_{M_1(t)+\frac{(1-\gamma)c}{4}A^*A-\frac{L_h}{4}I}dt + \int_0^T \|\dot{z}(t)\|^2_{M_2(t)+\frac{c}{4}I}dt + \frac{\gamma+1}{3c}\int_0^T \|\dot{y}(t)\|^2dt +
$$

$$
\int_0^T \left\|\frac{\sqrt{3c}}{2}\dot{z}(t) + \frac{1}{\sqrt{3c}}\dot{y}(t)\right\|^2 dt + (1-\gamma)\int_0^T \left\|\frac{\sqrt{3c}}{2}A\dot{x}(t) - \frac{1}{\sqrt{3c}}\dot{y}(t)\right\|^2 dt +
$$

$$
\frac{1}{L_h}\int_0^T \left\|\nabla h(x(t)) - \nabla h(x^*) + \frac{L_h}{2}\dot{x}(t)\right\|^2 dt \leq
$$

$$
\frac{1}{2}\left(\|x_0 - x^*\|^2_{M_1(0)+c(1-\gamma)A^*A} + \|z_0 - z^*\|^2_{M_2(0)+cI} + \frac{1}{c}\|y_0 - y^*\|^2\right).
$$

Letting T converge to $+\infty$ we find

$$
\|\dot{x}(\cdot)\|^2_{M_1(\cdot)+\frac{(1-\gamma)c}{4}A^*A-\frac{L_h}{4}I} \in L^1([0, +\infty), \mathbb{R}), \tag{39}
$$

$$
\|\dot{z}(\cdot)\|^2_{M_2(\cdot)+\frac{c}{4}I} \in L^1([0, +\infty), \mathbb{R}), \dot{y}(\cdot) \in L^2([0, +\infty), \mathcal{G}), \tag{40}
$$

$$
\frac{\sqrt{3c}}{2}\dot{z}(\cdot) + \frac{1}{\sqrt{3c}}\dot{y}(\cdot) \in L^2([0, +\infty), \mathcal{G}), (1-\gamma)\left(\frac{\sqrt{3c}}{2}A\dot{x}(\cdot) - \frac{1}{\sqrt{3c}}\dot{y}(\cdot)\right) \in L^2([0, +\infty), \mathcal{G}), \tag{41}
$$

and, consequently,

$$
\dot{z}(\cdot), (1-\gamma)A\dot{x}(\cdot) \in L^2([0, +\infty), \mathcal{G}). \tag{42}
$$

In the case when $L_h = 0$, which corresponds to the situation when h is an affine-continuous function, instead of (37) we obtain that for almost every $t \in [0, +\infty)$

$$
\frac{1}{2}\frac{d}{dt}\left(\|x(t) - x^*\|^2_{M_1(t)+c(1-\gamma)A^*A} + \|z(t) - Ax^*\|^2_{M_2(t)+cI} + \frac{1}{c}\|y(t) - y^*\|^2\right) +
$$

$$
\|\dot{x}(t)\|^2_{M_1(t)+\frac{(1-\gamma)c}{4}A^*A} + \|\dot{z}(t)\|^2_{M_2(t)+\frac{c}{4}I} + \frac{\gamma+1}{3c}\|\dot{y}(t)\|^2 + \tag{43}
$$

$$
\left\|\frac{\sqrt{3c}}{2}\dot{z}(t) + \frac{1}{\sqrt{3c}}\dot{y}(t)\right\|^2 + (1-\gamma)\left\|\frac{\sqrt{3c}}{2}A\dot{x}(t) - \frac{1}{\sqrt{3c}}\dot{y}(t)\right\|^2 \leq 0.
$$

25

By arguing as above, we obtain also in this case (38), (39)-(41) and (42).

Further, we have that $\dot{x}(\cdot) \in L^2([0,+\infty),\mathcal{H})$. Indeed, in case (I), when we assume that there exists $\alpha > 0$ such that $M_1(t) + \frac{(1-\gamma)c}{4}A^*A - \frac{L_h}{4}I \in P_\alpha(\mathcal{H})$ for every $t \in [0,+\infty)$, then this yields automatically. In case (II), from $(1-\gamma)A\dot{x}(\cdot) \in L^2([0,+\infty),\mathcal{G})$ and $\gamma \in [0,1)$, we have

$$A\dot{x}(\cdot) \in L^2([0,+\infty),\mathcal{G}).$$

But, since $A^*A \in P_\alpha(H)$, it yields $\|A\dot{x}(t)\|^2 \geq \alpha\|\dot{x}(t)\|^2$ for almost every $t \in [0,+\infty)$, which means that also in this case

$$\dot{x}(\cdot) \in L^2([0,+\infty),\mathcal{H}).$$

According to Lemma 5, this yields

$$\ddot{x}(\cdot) \in L^2([0,+\infty),\mathcal{H}) \text{ and } \ddot{z}(\cdot), \ddot{y}(\cdot) \in L^2([0,+\infty),\mathcal{G}).$$

Consequently, for almost every $t \in [0,+\infty)$ it holds

$$\frac{d}{dt}\|\dot{x}(t)\|^2 = 2\langle \ddot{x}(t), \dot{x}(t)\rangle \leq \left(\|\ddot{x}(t)\|^2 + \|\dot{x}(t)\|^2\right)$$

and the right-hand side is a function in $L^1([0,+\infty),\mathbb{R})$. Hence, according to Lemma 7,

$$\lim_{t\longrightarrow+\infty}\dot{x}(t) = 0.$$

Similarly, we obtain that

$$\lim_{t\longrightarrow+\infty}\dot{z}(t) = 0 \text{ and } \lim_{t\longrightarrow+\infty}\dot{y}(t) = 0.$$

We will close the proof of the theorem by showing that the asymptotic convergence of the trajectory follows from Lemma 8. One can easily notice that (38) is nothing else but condition (i) of this lemma when applied in the product space for the trajectory

$$[0,+\infty) \mapsto \mathcal{H} \times \mathcal{G} \times \mathcal{G}, \quad t \longrightarrow (x(t), z(t), y(t)),$$

the monotonically decreasing map

$$W : [0,+\infty) \mapsto \mathcal{H} \times \mathcal{G} \times \mathcal{G}, \quad W(t) = \left(M_1(t) + c(1-\gamma)A^*A, M_2(t) + cI, \frac{1}{c}I\right)$$

and the set $\mathcal{C}$ taken as the set of saddle points of the Lagrangian l.

Next we will show that also condition (ii) in Lemma 8 is fulfilled, namely, that every weak sequential cluster point of the trajectory $(x(t), z(t), y(t)), t \in [0,+\infty)$, is a saddle point of the Langrangian l.

Let $(\overline{x}, \overline{z}, \overline{y})$ be such a weak sequentially cluster point. This means that there exists a sequence $(s_n)_{n\geq 0}$ with $s_n \longrightarrow +\infty$ such that $(x(s_n), z(s_n), y(s_n))$ converges to $(\overline{x}, \overline{z}, \overline{y})$ as $n \longrightarrow +\infty$ in the weak topology of $\mathcal{H} \times \mathcal{G} \times \mathcal{G}$.

From (29) and (31) we get for every $n \geq 0$

$$-cA^*A(\dot{x}(s_n) + x(s_n)) - M_1(s_n)\dot{x}(s_n) + cA^*z(s_n) - A^*y(s_n) - \nabla h(x(s_n)) \in \partial f(\dot{x}(s_n) + x(s_n))$$

and

$$-c(\dot{z}(s_n) + z(s_n)) + cA(\gamma\dot{x}(s_n) + x(s_n)) - M_2(s_n)\dot{z}(s_n) + y(s_n) \in \partial g(\dot{z}(s_n) + z(s_n)),$$

26

respectively. For every $n \geq 0$, let

$$a_n^* := -cA^*A(\dot{x}(s_n) + x(s_n)) - M_1(s_n)\dot{x}(s_n) + cA^*z(s_n) - A^*y(s_n) - \nabla h(x(s_n)) + \nabla h(\dot{x}(s_n) + x(s_n))$$

and

$$a_n := \dot{x}(s_n) + x(s_n).$$

Hence, $(a_n, a_n^*)_{n \geq 0} \subseteq \operatorname{Gr} \partial(f + h)$. Similarly, for every $n \geq 0$, let

$$b_n^* := -c(\dot{z}(s_n) + z(s_n)) + cA(\gamma\dot{x}(s_n) + x(s_n)) - M_2(s_n)\dot{z}(s_n) + y(s_n)$$

and

$$b_n := \dot{z}(s_n) + z(s_n).$$

Hence, $(b_n, b_n^*)_{n \geq 0} \subseteq \operatorname{Gr} \partial g$.

Since $\lim_{t \to +\infty} \dot{x}(t) = 0$, $\lim_{t \to +\infty} \dot{z}(t) = 0$ and $\lim_{t \to +\infty} \dot{y}(t) = 0$ it follows that $(a_n)_{n \geq 0}$ converges weakly to $\overline{x}$ as $n \to \infty$. Furthermore, since $(M_2(s_n))_{n \geq 0}$ is bounded, and

$$b_n^* = c(\gamma - 1)A\dot{x}(s_n) + \dot{y}(s_n) - M_2(s_n)\dot{z}(s_n) + y(s_n) \ \forall n \geq 0,$$

it follows that $(b_n^*)_{n \geq 0}$ converges weakly to $\overline{y}$ as $n \to \infty$.

From (14) we have

$$Aa_n - b_n = \frac{1}{c}\dot{y}(s_n) \longrightarrow 0 \ (n \to +\infty),$$

which implies that $A\overline{x} = \overline{z}$. On the other hand, since ∇h is Lipschitz continuous, we have

$$\nabla h(\dot{x}(s_n) + x(s_n)) - \nabla h(x(s_n)) \longrightarrow 0 \ (n \to +\infty),$$

hence

$$\begin{aligned}
&\lim_{n \to +\infty} (a_n^* + A^*b_n^*) \\
&= \lim_{n \to +\infty} (c(\gamma - 1)A^*A\dot{x}(s_n) - cA^*\dot{z}(s_n) - A^*M_2(s_n)\dot{z}(s_n) - M_1(s_n)\dot{x}(s_n)) \\
&\quad + \lim_{n \to +\infty} (\nabla h(\dot{x}(s_n) + x(s_n)) - \nabla h(x(s_n))) \\
&= 0.
\end{aligned}$$

Thus, according to Proposition 9, we have

$$-A^*\overline{y} - \nabla h(\overline{x}) \in \partial f(\overline{x}) \text{ and } \overline{y} \in \partial g(A\overline{x}).$$

Consequently, $(\overline{x}, \overline{z}, \overline{y})$ is a saddle point of l.

The conclusion of the theorem follows from Lemma 8. $\qquad\square$

Next we will address two particular cases of the dynamical system (4). We consider first the case when $M_1(t) = M_2(t) = 0$ for every $t \in [0, +\infty)$, thus, the system (4) becomes

$$\begin{cases}
\dot{x}(t) + x(t) \in \operatorname*{argmin}_{x \in \mathcal{H}} \left(f(x) + \langle x, \nabla h(x(t)) \rangle + \frac{c}{2} \left\| Ax - z(t) + \frac{1}{c}y(t) \right\|^2 \right) \\[2mm]
\dot{z}(t) + z(t) = \operatorname*{argmin}_{x \in \mathcal{G}} \left(g(x) + \frac{c}{2} \left\| x - \left(A(\gamma\dot{x}(t) + x(t)) + \frac{1}{c}y(t) \right) \right\|^2 \right) \\[2mm]
\dot{y}(t) = cA(x(t) + \dot{x}(t)) - c(z(t) + \dot{z}(t)) \\[2mm]
x(0) = x^0 \in \mathcal{H}, \ z(0) = z^0 \in \mathcal{G}, \ y(0) = y^0 \in \mathcal{G},
\end{cases} \tag{44}$$

where $c > 0$ and $\gamma \in [0, 1]$. The dynamical system (44) can be seen as the continuous counterpart of the classical ADMM algorithm. The corresponding convergence result follows as a particular case of Theorem 10.

Theorem 11. *In the setting of the optimization problem* (1), *assume that the set of saddle points of the Lagrangian l is nonempty, $\gamma \in [0, 1)$ and that there exists $\alpha > 0$ such that $A^*A - \frac{L_h}{c(1-\gamma)} I \in P_\alpha(\mathcal{H})$. For an arbitrary starting point $(x^0, z^0, y^0) \in \mathcal{H} \times \mathcal{G} \times \mathcal{G}$, let $(x, z, y) : [0, +\infty) \longrightarrow \mathcal{H} \times \mathcal{G} \times \mathcal{G}$ be the unique strong global solution of the dynamical system* (44). *Then the trajectory $(x(t), z(t), y(t))$ converges weakly to a saddle point of l as $t \longrightarrow +\infty$.*

Next we consider the setting from Remark 2 with $M_1(t) = \frac{1}{\tau(t)} I - cA^*A$ and $M_2(t) = 0$, where $\tau(t)$ is such that $c\tau(t)\|A\|^2 \leq 1$, for every $t \in [0, +\infty)$. The resulting dynamical system is the primal-dual system (8). The corresponding convergence result follows again as a particular case of Theorem 10.

Theorem 12. *In the setting of the optimization problem* (1), *assume that the set of saddle points of the Lagrangian l is nonempty, the map $\tau : [0, +\infty) \to (0, +\infty)$ is locally absolutely continuous and monotonically increasing with*

$$c\tau(t)\|A\|^2 \leq 1 \ \text{and} \ \frac{4 - \tau(t)L_h}{4\tau(t)} I - \frac{c(3+\gamma)}{4} A^*A \in S_+(\mathcal{H}) \quad \forall t \in [0, +\infty),$$

and $\sup_{t \geq 0} \frac{\tau'(t)}{\tau^2(t)} < +\infty$. For an arbitrary starting point $(x^0, z^0, y^0) \in \mathcal{H} \times \mathcal{G} \times \mathcal{G}$, let $(x, z, y) : [0, +\infty) \longrightarrow \mathcal{H} \times \mathcal{G} \times \mathcal{G}$ be the unique strong global solution of the dynamical system (8). *If one of the following assumptions holds:*

(I) there exists $\alpha > 0$ such that $\frac{4 - \tau(t)L_h}{4\tau(t)} I - \frac{c(3+\gamma)}{4} A^*A \in P_\alpha(\mathcal{H})$ for every $t \in [0, +\infty)$;

(II) $\gamma \in [0, 1)$ and there exists $\alpha > 0$ such that $A^*A \in P_\alpha(\mathcal{H})$;

then the trajectory $(x(t), z(t), y(t))$ converges weakly to a saddle point of l as $t \longrightarrow +\infty$.

Remark 6. Let be $t \in [0, +\infty)$. Notice that the condition $\frac{4 - \tau(t)L_h}{4\tau(t)} I - \frac{c(3+\gamma)}{4} A^*A \in S_+(\mathcal{H})$ is fulfilled if and only if

$$\tau(t) \left(\frac{L_h}{4} + \frac{c(3+\gamma)}{4} \|A\|^2 \right) \leq 1.$$

On the other hand, the condition $\frac{4 - \tau(t)L_h}{4\tau(t)} I - \frac{c(3+\gamma)}{4} A^*A \in P_\alpha(\mathcal{H})$ holds, for $\alpha > 0$, if and only if

$$\tau(t) \left(\alpha + \frac{L_h}{4} + \frac{c(3+\gamma)}{4} \|A\|^2 \right) \leq 1.$$

For the last result of this paper we go back to the general dynamical system (4) and provide convergence rates for the violation of the feasibility condition by ergodic trajectories and the convergence of the objective function along these ergodic trajectories to its minimal value. The result can be seen as the continuous counterpart of a convergence rate result proved for the ADMM algorithm in [22, Theorem 4.3].

Theorem 13. *In the setting of the optimization problem* (1), *assume that the set of saddle points of the Lagrangian l is nonempty, the maps*

$$[0, +\infty) \to S_+(\mathcal{H}), t \mapsto M_1(t), \ \text{and} \ [0, +\infty) \to S_+(\mathcal{G}), t \mapsto M_2(t),$$

28

are locally absolutely continuous and monotonically decreasing,

$$M_1(t) + \frac{c(1-\gamma)}{4}A^*A - \frac{L_h}{2}I \in S_+(\mathcal{H}) \quad \forall t \in [0,+\infty),$$

$$\sup_{t\geq 0}\|\dot{M}_1(t)\| < +\infty \ \text{and} \ \sup_{t\geq 0}\|\dot{M}_2(t)\| < +\infty$$

and that one of the following conditions holds:

(I) there exists $\alpha > 0$ such that $M_1(t) + \frac{c(1-\gamma)}{4}A^*A - \frac{L_h}{4}I \in P_\alpha(\mathcal{H})$ for every $t \in [0,+\infty)$;

(II) $\gamma \in [0,1)$ and there exists $\alpha > 0$ such that $A^*A \in P_\alpha(\mathcal{H})$;

For an arbitrary starting point $(x^0, z^0, y^0) \in \mathcal{H} \times \mathcal{G} \times \mathcal{G}$, let $(x,z,y) : [0,+\infty) \longrightarrow \mathcal{H} \times \mathcal{G} \times \mathcal{G}$ be the unique strong global solution of the dynamical system (4). Consider further for every $t \in (0,+\infty)$ the ergodic trajectories

$$\tilde{x}(t) = \frac{1}{t}\int_0^t (\dot{x}(s) + x(s))ds$$

and

$$\tilde{z}(t) = \frac{1}{t}\int_0^t (\dot{z}(s) + z(s))ds.$$

Then there exists $K \geq 0$ such that for every $t \in (0,+\infty)$

$$\|A\tilde{x}(t) - \tilde{z}(t)\| \leq \frac{K}{t}.$$

In addition, for every $\overline{x} \in \mathcal{H}$ and every $t \in (0,+\infty)$ such that $(\tilde{x}(t),\tilde{z}(t)) \in \mathrm{dom}\,f \times \mathrm{dom}\,g$, one has

$$\Big((f+h)(\tilde{x}(t)) + g(\tilde{z}(t))\Big) - \Big((f+h)(\overline{x}) + g(A\overline{x})\Big) \leq \frac{\|(x^0,z^0,y^0) - (\overline{x}, A\overline{x}, 0)\|_{W(0)}^2}{2t},$$

where

$$W(t) = \left(M_1(t) + c(1-\gamma)A^*A, M_2(t) + cI, \frac{1}{c}I\right).$$

Proof. Let $\overline{x} \in \mathcal{H}$ be fixed. By using (5), that is

$$-cA^*A(\dot{x}(t) + x(t)) - M_1(t)\dot{x}(t) + cA^*z(t) - A^*y(t) - \nabla h(x(t)) \in \partial f(\dot{x}(t) + x(t)),$$

it yields

$$f(\dot{x}(t)+x(t)) - f(\overline{x}) \leq \langle cA^*A(\dot{x}(t)+x(t)) + M_1(t)\dot{x}(t) - cA^*z(t) + A^*y(t) + \nabla h(x(t)), \overline{x} - (\dot{x}(t)+x(t))\rangle$$

$$(45)$$

for almost every $t \in [0,+\infty)$. Similarly, by using (31), that is

$$-c(\dot{z}(t) + z(t)) + cA(\gamma\dot{x}(t) + x(t)) - M_2(t)\dot{z}(t) + y(t) \in \partial g(\dot{z}(t) + z(t)),$$

it yields

$$g(\dot{z}(t)+z(t)) - g(A\overline{x}) \leq \langle c(\dot{z}(t)+z(t)) - cA(\gamma\dot{x}(t)+x(t)) + M_2(t)\dot{z}(t) - y(t), A\overline{x} - (\dot{z}(t)+z(t))\rangle. \quad (46)$$

for almost every $t \in [0,+\infty)$. Further, by using the convexity of h and the Descent Lemma we obtain for almost every $t \in [0,+\infty)$

$$h(\overline{x}) - h(\dot{x}(t) + x(t)) - \langle \nabla h(x(t)), \overline{x} - (\dot{x}(t) + x(t))\rangle \geq$$

$$h(x(t)) + \langle \nabla h(x(t)), \overline{x} - x(t)\rangle - h(\dot{x}(t) + x(t)) - \langle \nabla h(x(t)), \overline{x} - (\dot{x}(t) + x(t))\rangle = \quad (47)$$

$$h(x(t)) - h(\dot{x}(t) + x(t)) + \langle \nabla h(x(t)), \dot{x}(t)\rangle \geq -\frac{L_h}{2}\|\dot{x}(t)\|^2.$$

Adding (45) and (47) we obtain for almost every $t \in [0, +\infty)$

$$(f+h)(\dot{x}(t) + x(t)) - (f+h)(\overline{x}) \leq$$
$$\langle cA^*A(\dot{x}(t) + x(t)) + M_1(t)\dot{x}(t) - cA^*z(t) + A^*y(t), \overline{x} - (\dot{x}(t) + x(t)) \rangle + \tfrac{L_h}{2}\|\dot{x}(t)\|^2. \tag{48}$$

We recall the following four identities from the proof of Theorem 10 (here were actually replace x^* with $\overline{x}$ and y^* by 0)

$$\langle -A^*y(t), \dot{x}(t) + x(t) - \overline{x} \rangle + \langle y(t), \dot{z}(t) + z(t) - A\overline{x} \rangle$$
$$= -\langle y(t), A(\dot{x}(t) + x(t)) - A\overline{x} - (\dot{z}(t) + z(t)) + A\overline{x} \rangle = \frac{1}{c}\langle y(t), \dot{y}(t) \rangle$$
$$= -\frac{1}{2c}\frac{d}{dt}\|y(t)\|^2,$$

which corresponds to (33),

$$\langle -cA^*A(\dot{x}(t) + x(t)) + cA^*z(t), \dot{x}(t) + x(t) - \overline{x} \rangle$$
$$+ \langle -c(\dot{z}(t) + z(t)) + cAx(t) + c\gamma A\dot{x}(t), \dot{z}(t) + z(t) - A\overline{x} \rangle$$
$$= -\frac{\gamma+1}{3c}\|\dot{y}(t)\|^2 - \frac{c}{4}\|\dot{z}(t)\|^2 - \frac{(1-\gamma)c}{4}\|A\dot{x}(t)\|^2 - \left\|\frac{\sqrt{3c}}{2}\dot{z}(t) + \frac{1}{\sqrt{3c}}\dot{y}(t)\right\|^2$$
$$- (1-\gamma)\left\|\frac{\sqrt{3c}}{2}A\dot{x}(t) - \frac{1}{\sqrt{3c}}\dot{y}(t)\right\|^2 + \frac{c(\gamma-1)}{2}\frac{d}{dt}\left(\|Ax(t) - A\overline{x}\|^2\right) - \frac{c}{2}\frac{d}{dt}\left(\|z(t) - A\overline{x}\|^2\right),$$

which corresponds to (36), and

$$\langle -M_1(t)\dot{x}(t), \dot{x}(t) + x(t) - \overline{x} \rangle = -\|\dot{x}(t)\|^2_{M_1(t)} - \langle M_1(t)\dot{x}(t), x(t) - \overline{x} \rangle$$
$$= -\|\dot{x}(t)\|^2_{M_1(t)} + \frac{1}{2}\langle \dot{M}_1(t)(x(t) - \overline{x}), x(t) - \overline{x} \rangle$$
$$- \frac{1}{2}\frac{d}{dt}\|x(t) - \overline{x}\|^2_{M_1(t)}$$

and

$$\langle -M_2(t)\dot{z}(t), \dot{z}(t) + z(t) - A\overline{x} \rangle = -\|\dot{z}(t)\|^2_{M_2(t)} - \langle M_2(t)\dot{z}(t), z(t) - A\overline{x} \rangle$$
$$= -\|\dot{z}(t)\|^2_{M_2(t)} + \frac{1}{2}\langle \dot{M}_2(t)(z(t) - A\overline{x}), z(t) - A\overline{x} \rangle$$
$$- \frac{1}{2}\frac{d}{dt}\|z(t) - A\overline{x}\|^2_{M_2(t)},$$

which all hold for for almost every $t \in [0, +\infty)$. By adding the four identities, (48) and (46), we obtain for almost every $t \in [0, +\infty)$

$$\left((f+h)(\dot{x}(t) + x(t)) + g(\dot{z}(t) + z(t))\right) - \left((f+h)(\overline{x}) + g(A\overline{x})\right) \leq$$
$$-\frac{\gamma+1}{3c}\|\dot{y}(t)\|^2 - \frac{c}{4}\|\dot{z}(t)\|^2 - \frac{(1-\gamma)c}{4}\|A\dot{x}(t)\|^2$$
$$-\left\|\frac{\sqrt{3c}}{2}\dot{z}(t) + \frac{1}{\sqrt{3c}}\dot{y}(t)\right\|^2 - (1-\gamma)\left\|\frac{\sqrt{3c}}{2}A\dot{x}(t) - \frac{1}{\sqrt{3c}}\dot{y}(t)\right\|^2$$
$$-\frac{c(1-\gamma)}{2}\frac{d}{dt}\left(\|Ax(t) - A\overline{x}\|^2\right) - \frac{c}{2}\frac{d}{dt}\left(\|z(t) - A\overline{x}\|^2\right) + \frac{L_h}{2}\|\dot{x}(t)\|^2 - \frac{1}{2c}\frac{d}{dt}\|y(t)\|^2$$
$$-\|\dot{x}(t)\|^2_{M_1(t)} + \frac{1}{2}\langle \dot{M}_1(t)(x(t) - \overline{x}), x(t) - \overline{x} \rangle - \frac{1}{2}\frac{d}{dt}\|x(t) - \overline{x}\|^2_{M_1(t)}$$
$$-\|\dot{z}(t)\|^2_{M_2(t)} + \frac{1}{2}\langle \dot{M}_2(t)(z(t) - A\overline{x}), z(t) - A\overline{x} \rangle - \frac{1}{2}\frac{d}{dt}\|z(t) - A\overline{x}\|^2_{M_2(t)}.$$

By neglecting the negative terms (here we use also that $M_1(t) + \frac{c(1-\gamma)}{4}A^*A - \frac{L_h}{2}I \in S_+(\mathcal{H})$), we obtain for almost every $t \in [0, +\infty)$

$$\Big((f+h)(\dot{x}(t) + x(t)) + g(\dot{z}(t) + z(t))\Big) - \Big((f+h)(\overline{x}) + g(A\overline{x})\Big) \le$$

$$-\frac{1}{2}\frac{d}{dt}\left(\|x(t) - \overline{x}\|^2_{M_1(t)+c(1-\gamma)A^*A} + \|z(t) - A\overline{x}\|^2_{M_2(t)+cI} + \frac{1}{c}\|y(t)\|^2\right) = \tag{49}$$

$$-\frac{1}{2}\frac{d}{dt}\|(x(t), z(t), y(t)) - (\overline{x}, A\overline{x}, 0)\|^2_{W(t)},$$

where

$$W(t) = \left(M_1(t) + c(1-\gamma)A^*A, M_2(t) + cI, \frac{1}{c}I\right).$$

For $\tilde{x}(t) = \frac{1}{t}\int_0^t(\dot{x}(s) + x(s))ds$ and $\tilde{z}(t) = \frac{1}{t}\int_0^t(\dot{z}(s) + z(s))ds$, it holds

$$A\tilde{x}(t) - \tilde{z}(t) = \frac{1}{t}\int_0^t A(\dot{x}(s) + x(s)) - (\dot{z}(s) + z(s))ds = \frac{1}{ct}\int_0^t \dot{y}(s)ds = \frac{y(t) - y_0}{ct} \quad \forall t \in (0, +\infty).$$

From Theorem 10 it follows that the trajectory $(x(t), z(t), y(t)) \to (x_\infty, z_\infty, y_\infty), t \in [0, +\infty)$, converges weakly to a saddle point of l as $t \to +\infty$. This means that $y(t), t \in [0, +\infty)$, it is bounded, thus there exists $K \ge 0$ such that

$$\|A\tilde{x}(t) - \tilde{z}(t)\| \le \frac{K}{t} \quad \forall t \in (0, +\infty).$$

Let $t \in (0, +\infty)$ be such that $(\tilde{x}(t), \tilde{z}(t)) \in \operatorname{dom} f \times \operatorname{dom} g$. By Jensen's inequality in the integral form we have for every $t \in (0, +\infty)$

$$(f+h)(\tilde{x}(t)) = (f+h)\left(\frac{1}{t}\int_0^t(\dot{x}(s) + x(s))ds\right) \le \frac{1}{t}\int_0^t(f+h)(\dot{x}(s) + x(s))ds$$

and

$$g(\tilde{z}(t)) = g\left(\frac{1}{t}\int_0^t(\dot{z}(s) + z(s))ds\right) \le \frac{1}{t}\int_0^t g(\dot{z}(s) + z(s))ds,$$

which, combined with (49), yields

$$(f+h)(\tilde{x}(t)) + g(\tilde{z}(t)) \le$$

$$\frac{1}{t}\int_0^t\Big((f+h)(\dot{x}(s) + x(s)) + g(\dot{z}(s) + z(s))\Big)ds \le$$

$$\frac{1}{t}\int_0^t\left(\Big((f+h)(\overline{x}) + g(A\overline{x})\Big) - \frac{1}{2}\frac{d}{ds}\|(x(s), z(s), y(s)) - (\overline{x}, A\overline{x}, 0)\|^2_{W(s)}\right)ds =$$

$$(f+h)(\overline{x}) + g(A\overline{x}) - \frac{1}{2t}\left(\|(x(t), z(t), y(t)) - (\overline{x}, A\overline{x}, 0)\|^2_{W(t)} - \|(x(0), z(0), y(0)) - (\overline{x}, A\overline{x}, 0)\|^2_{W(0)}\right) \le$$

$$(f+h)(\overline{x}) + g(A\overline{x}) + \frac{\|(x^0, z^0, y^0) - (\overline{x}, A\overline{x}, 0)\|^2_{W(0)}}{2t}.$$

Hence,

$$\Big((f+h)(\tilde{x}(t)) + g(\tilde{z}(t))\Big) - \Big((f+h)(\overline{x}) + g(A\overline{x})\Big) \le \frac{\|(x^0, z^0, y^0) - (\overline{x}, A\overline{x}, 0)\|^2_{W(0)}}{2t}.$$

$\square$

References

[1] B. Abbas, H. Attouch, *Dynamical systems and forward-backward algorithms associated with the sum of a convex subdifferential and a monotone cocoercive operator*, Optimization 64(10), 2223–2252, 2015

[2] B. Abbas, H. Attouch, B.F. Svaiter, *Newton-like dynamics and forward-backward methods for structured monotone inclusions in Hilbert spaces*, Journal of Optimization Theory and its Applications 161(2), 331–360, 2014

[3] A. Alotaibi, P. L. Combettes and N. Shahzad, *Solving coupled composite monotone inclusions by successive Fejér approximations of their Kuhn-Tucker set*, SIAM Journal on Optimization 24(4), 2076–2095, 2014

[4] F. Alvarez, H. Attouch, J. Bolte, P. Redont, *A second-order gradient-like dissipative dynamical system with Hessian-driven damping. Application to optimization and mechanics*, Journal de Mathématiques Pures et Appliquées (9) 81(8), 747–779, 2002

[5] A.S. Antipin, *Minimization of convex functions on convex sets by means of differential equations*, (Russian) Differentsial'nye Uravneniya 30(9), 1475–1486, 1994; translation in Differential Equations 30(9), 1365–1375, 1994

[6] H. Attouch, B.F. Svaiter, *A continuous dynamical Newton-like approach to solving monotone inclusions*, SIAM Journal on Control and Optimization 49(2), 574–598, 2011

[7] S. Banert, R.I. Boţ, *A forward-backward-forward differential equation and its asymptotic properties*, Journal of Convex Analysis 25(2), 371–388, 2018

[8] S. Banert, R.I. Boţ, E.R. Csetnek, *Fixing and extending some recent results on the ADMM algorithm*, Numerical Algorithms, DOI: 10.1007/s11075-020-00934-5, 2020

[9] H.H. Bauschke, P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, CMS Books in Mathematics, Springer, New York, 2011

[10] J. Bolte, *Continuous gradient projection method in Hilbert spaces*, Journal of Optimization Theory and its Applications 119(2), 235–259, 2003

[11] R.I. Boţ, *Conjugate Duality in Convex Optimization*, Lecture Notes in Economics and Mathematical Systems, Vol. 637, Springer, Berlin Heidelberg, 2010

[12] R.I. Boţ, E.R. Csetnek, *ADMM for monotone operators: convergence analysis and rates*, Advances in Computational Mathematics 45(1), 327-359, 2019

[13] R.I. Boţ, E.R. Csetnek, *A dynamical system associated with the fixed points set of a nonexpansive operator*, Journal of Dynamics and Differential Equations 29(1), 155–168, 2017

[14] H. Brézis, *Opérateurs Maximaux Monotones et Semi-Groupes de Contractions Dans les Espaces de Hilbert*, North-Holland Mathematics Studies No. 5, Notas de Matemática (50), North-Holland/Elsevier, New York, 1973

[15] R.E. Bruck, Jr., *Asymptotic convergence of nonlinear contraction semigroups in Hilbert space*, Journal of Functional Analysis 18, 15–26, 1975

[16] A. Chambolle, T. Pock, *A first-order primal-dual algorithm for convex problems with applications to imaging*, Journal of Mathematical Imaging and Vision 40(1), 120–145, 2011

[17] G. Chen, M. Teboulle, *A proximal-based decomposition method for convex minimization problems*, Mathematical Programming 64, 81–101, 1994

[18] P.L. Combettes, B.C. Vu, *Variable metric quasi-Fejér monotonicity*, Nonlinear Analysis 78, 17-31, 2013

[19] L. Condat, *A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms*, Journal of Optimization Theory and Applications 158(2), 460–479, 2013

[20] M.G. Crandall, A. Pazy, *Semi-groups of nonlinear contractions and dissipative sets*, Journal of Functional Analysis 3, 376-418, 1969

[21] E.R. Csetnek, Y. Malitsky, M.K. Tam, *Shadow Douglas-Rachford splitting for monotone inclusions*, Applied Mathematics and Optimization 80, 665–678, 2019

[22] Y. Cui, X. Li, D. Sun, K.-C. Toh, *On the convergence properties of a majorized alternating direction method of multipliers for linearly constrained convex optimization problems with coupled objective functions*, Journal of Optimization Theory and Applications 169, 1013–1041, 2016

[23] M. Fazel, T.K. Pong, D. Sun, P. Tseng, *Hankel matrix rank minimization with applications in system identification and realization*, SIAM Journal on Matrix Analysis and Applications 34, 946–977, 2013

[24] A. Haraux, *Systèmes Dynamiques Dissipatifs et Applications*, Recherches en Mathématiques Appliquées 17, Masson, Paris, 1991

[25] R. Shefi, M. Teboulle, *Rate of convergence analysis of decomposition methods based on the proximal method of multipliers for convex optimization*, SIAM Journal on Optimization 24(1), 269–297, 2014

[26] E.D. Sontag, *Mathematical Control Theory. Deterministic Finite-Dimensional Systems*, Texts in Applied Mathematics 6, Springer-Verlag, New York, 1998

[27] B.C. Vũ, *A splitting algorithm for dual monotone inclusions involving cocoercive operators*, Advances in Computational Mathematics 38(3), 667–681, 2013

[28] C. Zălinescu, *Convex Analysis in General Vector Spaces*, World Scientific, Singapore, 2002

[29] R. Zimmer, *Essential Results of Functional Analysis*, The University of Chicago Press, Chicago, 1990