Second order forward-backward dynamical systems for monotone inclusion problems

Radu Ioan Boţ *   Ernö Robert Csetnek †

March 16, 2015

Abstract. We begin by considering second order dynamical systems of the from $\ddot{x}(t) + \Gamma(\dot{x}(t)) + \lambda(t)B(x(t)) = 0$, where $\Gamma : \mathcal{H} \to \mathcal{H}$ is an elliptic bounded self-adjoint linear operator defined on a real Hilbert space $\mathcal{H}$, $B : \mathcal{H} \to \mathcal{H}$ is a cocoercive operator and $\lambda : [0, +\infty) \to [0, +\infty)$ is a relaxation function depending on time. We show the existence and uniqueness of strong global solutions in the framework of the Cauchy-Lipschitz-Picard Theorem and prove weak convergence for the generated trajectories to a zero of the operator B, by using Lyapunov analysis combined with the celebrated Opial Lemma in its continuous version. The framework allows to address from similar perspectives second order dynamical systems associated with the problem of finding zeros of the sum of a maximally monotone operator and a cocoercive one. This captures as particular case the minimization of the sum of a nonsmooth convex function with a smooth convex one and allows us to recover and improve several results from the literature concerning the minimization of a convex smooth function subject to a convex closed set by means of second order dynamical systems. When considering the unconstrained minimization of a smooth convex function we prove a rate of $\mathcal{O}(1/t)$ for the convergence of the function value along the ergodic trajectory to its minimum value. A similar analysis is carried out also for second order dynamical systems having as first order term $\gamma(t)\dot{x}(t)$, where $\gamma : [0, +\infty) \to [0, +\infty)$ is a damping function depending on time.

Key Words. dynamical systems, Lyapunov analysis, monotone inclusions, convex optimization problems, continuous forward-backward method

AMS subject classification. 34G25, 47J25, 47H05, 90C25

1 Introduction and preliminaries

This paper is motivated by the heavy ball with friction dynamical system

$$\ddot{x} + \gamma\dot{x} + \nabla f(x) = 0, \tag{1}$$

which is a nonlinear oscillator with damping $\gamma > 0$ and potential $f : \mathcal{H} \to \mathbb{R}$, supposed to be a convex and differentiable function defined on the real Hilbert space $\mathcal{H}$. When

*University of Vienna, Faculty of Mathematics, Oskar-Morgenstern-Platz 1, A-1090 Vienna, Austria, email: radu.bot@univie.ac.at.

†University of Vienna, Faculty of Mathematics, Oskar-Morgenstern-Platz 1, A-1090 Vienna, Austria, email: ernoe.robert.csetnek@univie.ac.at. Research supported by FWF (Austrian Science Fund), Lise Meitner Programme, project M 1682-N25.

$\mathcal{H} = \mathbb{R}^2$, the system (1) is a simplified version of the differential system describing the motion of a heavy ball that keeps rolling over the graph of the function f under its own inertia until friction stops it at a critical point of f (see [11]).

The second order dynamical system (1) has been considered by several authors in the context of minimizing the function f, these investigations being either concerned with the convergence of the generated trajectories to a critical point of f or with the convergence of the function along the trajectories to its global minimum value (see [3, 7, 8, 11]). It is also worth to mention that the time discretization of the heavy ball with friction dynamical system leads to the so-called inertial-type algorithms, which are numerical schemes sharing the feature that the current iterate of the generated sequence is defined by making use of the previous two iterates (see, for instance, [3–5, 18, 21, 23]).

In order to approach the minimization of f over a nonempty, convex and closed set $C \subseteq \mathcal{H}$, the gradient-projection second order dynamical system

$$\ddot{x} + \gamma\dot{x} + x - P_C(x - \eta\nabla f(x)) = 0, \tag{2}$$

has been considered, where $P_C : \mathcal{H} \to C$ denotes the projection onto the set C and $\eta > 0$. Convergence statements for the trajectories to a global minimizer of f over C have been provided in [7, 8]. Furthermore, in [8], these investigations have been expanded to more general second order dynamical systems of the form

$$\ddot{x} + \gamma\dot{x} + x - Tx = 0, \tag{3}$$

where $T : \mathcal{H} \to \mathcal{H}$ is a nonexpansive operator. It has been shown that when $\gamma^2 > 2$ the trajectory of (8) converges weakly to an element in the fixed points set of T, provided it is nonempty.

In the first part of the present manuscript we treat the second order dynamical system

$$\ddot{x}(t) + \Gamma(\dot{x}(t)) + \lambda(t)B(x(t)) = 0, \tag{4}$$

where $\Gamma : \mathcal{H} \to \mathcal{H}$ is an elliptic bounded self-adjoint linear operator, $B : \mathcal{H} \to \mathcal{H}$ is a cocoercive operator and $\lambda : [0, +\infty) \to [0, +\infty)$ is a relaxation function in time. We notice that the presence of the elliptic operator induces an anisotropic damping and refer to [3], where a similar construction has been used fin the context of minimizing a convex and smooth function. The existence and uniqueness of strong global solutions for (4) is obtained by applying the classical Cauchy-Lipschitz-Picard Theorem (see [20, 27]). We show that under mild assumptions on the relaxation function the trajectory $x(t)$ converges weakly as $t \to +\infty$ to a zero of the operator B, provided it has a nonempty set of zeros. To this end we will use Lyapunov analysis combined with the continuous version of the Opial Lemma (see also [3, 7, 8], where similar techniques have been used).

Further, we approach the problem of finding a zero of the sum of a maximally monotone operator and a cocoercive one via a second order dynamical system formulated by making use of the resolvent of the set-valued operator, see (31). Dynamical systems of implicit type have been already considered in the literature in [1, 2, 9, 12, 14, 16, 17]. We specialize these investigations to the minimization of the sum of a nonsmooth convex function with a smooth convex function one, fact which allows us to recover and improve results given in [7,8] in the context of studying the dynamical system (2). Whenever B is the gradient of a smooth convex function we show that the latter converges along the ergodic trajectories generated by (4) to its minimum value with a rate of convergence of $\mathcal{O}(1/t)$.

We close the paper by showing that a similar analysis can be carried out when using as starting point dynamical systems of the form

$$\ddot{x}(t) + \gamma(t)\dot{x}(t) + \lambda(t)B(x(t)) = 0, \tag{5}$$

where the damping coefficient $\gamma : [0, +\infty) \to [0, +\infty)$ is a function depending on time.

Throughout this paper $\mathbb{N} = \{0, 1, 2, ...\}$ denotes the set of nonnegative integers and $\mathcal{H}$ a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and corresponding norm $\| \cdot \| = \sqrt{\langle \cdot, \cdot \rangle}$.

2 A dynamical system: existence and uniqueness of strong global solutions

This section is devoted to the study of existence and uniqueness of strong global solutions of a second order dynamical system governed by Lipschitz continuous operators.

Let $\Gamma : \mathcal{H} \to \mathcal{H}$ be an L_Γ-Lipschitz continuous operator (that is $L_\Gamma \geq 0$ and $\|\Gamma x - \Gamma y\| \leq L_\Gamma \|x - y\|$ for all $x, y \in \mathcal{H}$), $B : \mathcal{H} \to \mathcal{H}$ L_B-Lipschitz continuous operator, $\lambda : [0, +\infty) \to [0, +\infty)$ a Lebesgue measurable function, $u_0, v_0 \in \mathcal{H}$ and consider the dynamical system

$$\begin{cases} \ddot{x}(t) + \Gamma(\dot{x}(t)) + \lambda(t)B(x(t)) = 0 \\ x(0) = u_0, \dot{x}(0) = v_0. \end{cases} \tag{6}$$

As in [2, 12], we consider the following definition of an absolutely continuous function.

Definition 1 (see, for instance, [2, 12]) A function $x : [0, b] \to \mathcal{H}$ (where $b > 0$) is said to be absolutely continuous if one of the following equivalent properties holds:

(i) there exists an integrable function $y : [0, b] \to \mathcal{H}$ such that

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

(ii) x is continuous and its distributional derivative is Lebesgue integrable on $[0, b]$;

(iii) for every $\varepsilon > 0$, there exists $\eta > 0$ such that for any finite family of intervals $I_k = (a_k, b_k)$ we have the implication

$$\left(I_k \cap I_j = \emptyset \text{ and } \sum_k |b_k - a_k| < \eta \right) \Longrightarrow \sum_k \|x(b_k) - x(a_k)\| < \varepsilon.$$

Remark 1 (a) It follows from the definition that an 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 integration formula (i).

(b) If $x : [0, b] \to \mathcal{H}$ (where $b > 0$) is absolutely continuous and $B : \mathcal{H} \to \mathcal{H}$ is L-Lipschitz continuous (where $L \geq 0$), then the function $z = B \circ x$ is absolutely continuous, too. This can be easily seen by using the characterization of absolute continuity in Definition 1(iii). Moreover, z is almost everywhere differentiable and the inequality $\|\dot{z}(\cdot)\| \leq L\|\dot{x}(\cdot)\|$ holds almost everywhere.

Definition 2 We say that $x : [0,+\infty) \to \mathcal{H}$ is a strong global solution of (6) if the following properties are satisfied:

(i) $x, \dot{x} : [0,+\infty) \to \mathcal{H}$ are locally absolutely continuous, in other words, absolutely continuous on each interval $[0,b]$ for $0 < b < +\infty$;

(ii) $\ddot{x}(t) + \Gamma(\dot{x}(t)) + \lambda(t)B(x(t)) = 0$ for almost every $t \in [0,+\infty)$;

(iii) $x(0) = u_0$ and $\dot{x}(0) = v_0$.

For proving the existence and uniqueness of strong global solutions of (6) we use the Cauchy-Lipschitz-Picard Theorem for absolutely continues trajectories (see for example [20, Proposition 6.2.1], [27, Theorem 54]). The key observation here is that one ca rewrite (6) as a certain first order dynamical system in a product space (see also [6]).

Theorem 2 *Let $\Gamma : \mathcal{H} \to \mathcal{H}$ be an L_Γ-Lipschitz continuous operator, $B : \mathcal{H} \to \mathcal{H}$ a L_B-Lipschitz continuous operator and $\lambda : [0,+\infty) \to [0,+\infty)$ a Lebesgue measurable function such that $\lambda \in L^1_{\text{loc}}([0,+\infty))$ (that is $\lambda \in L^1([0,b])$ for every $0 < b < +\infty$). Then for each $u_0, v_0 \in \mathcal{H}$ there exists a unique strong global solution of the dynamical system* (6).

Proof. The system (6) can be equivalently written as a first order dynamical system in the phase space $\mathcal{H} \times \mathcal{H}$

$$\begin{cases} \dot{Y}(t) = F(t, Y(t)) \\ Y(0) = (u_0, v_0), \end{cases} \tag{7}$$

with

$$Y : [0,+\infty) \to \mathcal{H} \times \mathcal{H},\ Y(t) = (x(t), \dot{x}(t))$$

and

$$F : [0,+\infty) \times \mathcal{H} \times \mathcal{H} \to \mathcal{H} \times \mathcal{H},\ F(t,u,v) = (v, -\Gamma v - \lambda(t)Bu).$$

We endow $\mathcal{H} \times \mathcal{H}$ with scalar product $\langle (u,v), (\overline{u}, \overline{v}) \rangle_{\mathcal{H}\times\mathcal{H}} = \langle u, \overline{u} \rangle + \langle v, \overline{v} \rangle$ and corresponding norm $\|(u,v)\|_{\mathcal{H}\times\mathcal{H}} = \sqrt{\|u\|^2 + \|v\|^2}$.

(a) For arbitrary $u, \overline{u}, v, \overline{v} \in \mathcal{H}$, by using the Lipschitz continuity of the involved operators, we obtain

$$\begin{aligned} \|F(t,u,v) - F(t,\overline{u},\overline{v})\|_{\mathcal{H}\times\mathcal{H}} &= \sqrt{\|v - \overline{v}\|^2 + \|\Gamma\overline{v} - \Gamma v + \lambda(t)(B\overline{u} - Bu)\|^2} \\ &\le \sqrt{(1 + 2L_\Gamma^2)\|v - \overline{v}\|^2 + 2L_B^2\lambda^2(t)\|u - \overline{u}\|^2} \\ &\le \sqrt{1 + 2L_\Gamma^2 + 2L_B^2\lambda^2(t)}\|(u,\overline{u}) - (v,\overline{v})\|_{\mathcal{H}\times\mathcal{H}} \\ &\le (1 + \sqrt{2}L_\Gamma + \sqrt{2}L_B\lambda(t))\|(u,\overline{u}) - (v,\overline{v})\|_{\mathcal{H}\times\mathcal{H}}\ \forall t \ge 0. \end{aligned}$$

As $\lambda \in L^1_{\text{loc}}([0,+\infty))$, the Lipschitz constant of $F(t,\cdot,\cdot)$ is local integrable.

(b) Next we show that

$$\forall u, v \in \mathcal{H},\ \forall b > 0, \quad F(\cdot, u, v) \in L^1([0,b], \mathcal{H} \times \mathcal{H}). \tag{8}$$

For arbitrary $u, v \in \mathcal{H}$ and $b > 0$ it holds

$$\begin{aligned} \int_0^b \|F(t,u,v)\|_{\mathcal{H}\times\mathcal{H}} dt &= \int_0^b \sqrt{\|v\|^2 + \|\Gamma v + \lambda(t)Bu\|^2} dt \\ &\le \int_0^b \sqrt{\|v\|^2 + 2\|\Gamma v\|^2 + 2\lambda^2(t)\|Bu\|^2} dt \\ &\le \int_0^b \left(\sqrt{\|v\|^2 + 2\|\Gamma v\|^2} + \sqrt{2}\lambda(t)\|Bu\| \right) dt \end{aligned}$$

and from here (8) follows, by using the assumptions made on λ.

In the light of the statements (a) and (b), the existence and uniqueness of a strong global solution for (7) are consequences of the Cauchy-Lipschitz-Picard Theorem for first order dynamical systems (see, for example, [20, Proposition 6.2.1], [27, Theorem 54]). From here, due to the equivalence of (6) and (7), the conclusion follows. ■

3 Convergence of the trajectories

In this section we address the convergence properties of the trajectories generated by the dynamical system (6) by assuming that $B : \mathcal{H} \to \mathcal{H}$ is a β-*cocoercive* operator for $\beta > 0$, that is $\beta\|Bx - By\|^2 \leq \langle x - y, Bx - By\rangle$ for all $x, y \in \mathcal{H}$. To this end we will make use of the following well-known results, which can be interpreted as continuous versions of the quasi-Fejér monotonicity for sequences. For their proofs we refer the reader to [2, Lemma 5.1] and [2, Lemma 5.2], respectively.

Lemma 3 *Suppose that* $F : [0, +\infty) \to \mathbb{R}$ *is locally absolutely continuous and bounded below and that there exists* $G \in L^1([0, +\infty))$ *such that for almost every* $t \in [0, +\infty)$

$$\frac{d}{dt}F(t) \leq G(t).$$

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

Lemma 4 *If* $1 \leq p < \infty$, $1 \leq r \leq \infty$, $F : [0, +\infty) \to [0, +\infty)$ *is locally absolutely continuous,* $F \in L^p([0, +\infty))$, $G : [0, +\infty) \to \mathbb{R}$, $G \in L^r([0, +\infty))$ *and for almost every* $t \in [0, +\infty)$

$$\frac{d}{dt}F(t) \leq G(t),$$

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

The next result which we recall here is the continuous version of the Opial Lemma (see, for example, [2, Lemma 5.3], [1, Lemma 1.10]).

Lemma 5 *Let* $S \subseteq \mathcal{H}$ *be a nonempty set and* $x : [0, +\infty) \to \mathcal{H}$ *a given map. Assume that*

(i) for every $x^* \in S$, $\lim_{t\to+\infty} \|x(t) - x^*\|$ *exists;*

(ii) every weak sequential cluster point of the map x *belongs to* S.

Then there exists $x_\infty \in S$ *such that* $x(t)$ *converges weakly to* x_∞ *as* $t \to +\infty$.

In order to prove the convergence of the trajectories of (6), we make the following assumptions on the operator Γ and the relaxation function λ, respectively:

(A1) $\Gamma : \mathcal{H} \to \mathcal{H}$ is a bounded self-adjoint linear operator, assumed to be elliptic, that is, there exists $\gamma > 0$ such that $\langle \Gamma u, u\rangle \geq \gamma\|u\|^2$ for all $u \in \mathcal{H}$;

(A2) $\lambda : [0, +\infty) \to (0, +\infty)$ is locally absolutely continuous and there exists $\theta > 0$ such that for almost every $t \in [0, +\infty)$ we have

$$\dot{\lambda}(t) \geq 0 \text{ and } \lambda(t) \leq \frac{\beta\gamma^2}{1+\theta}. \tag{9}$$

Due to Definition 1 and Remark 1(a) $\dot{\lambda}(t)$ exists for almost every $t \geq 0$ and $\dot{\lambda}$ is Lebesgue integrable on each interval $[0, b]$ for $0 < b < +\infty$. If $\dot{\lambda}(t) \geq 0$ for almost every $t \geq 0$, then λ is monotonically increasing, thus, as λ is assumed to take only positive values, (A2) yields the existence of a lower bound $\underline{\lambda}$ such that for almost every $t \in [0, +\infty)$ one has

$$0 < \underline{\lambda} \leq \lambda(t) \leq \frac{\beta\gamma^2}{1+\theta}. \tag{10}$$

Theorem 6 *Let $B : \mathcal{H} \to \mathcal{H}$ be a β-cocoercive operator for $\beta > 0$ such that $\operatorname{zer} B := \{u \in \mathcal{H} : Bu = 0\} \neq \emptyset$, $\Gamma : \mathcal{H} \to \mathcal{H}$ be an operator fulfilling (A1), $\lambda : [0, +\infty) \to (0, +\infty)$ be a function fulfilling (A2) and $u_0, v_0 \in \mathcal{H}$. Let $x : [0, +\infty) \to \mathcal{H}$ be the unique strong global solution of (6). Then the following statements are true:*

(i) the trajectory x is bounded and $\dot{x}, \ddot{x}, Bx \in L^2([0, +\infty); \mathcal{H})$;

(ii) $\lim_{t\to+\infty} \dot{x}(t) = \lim_{t\to+\infty} \ddot{x}(t) = \lim_{t\to+\infty} B(x(t) = 0$;

(iii) $x(t)$ converges weakly to an element in $\operatorname{zer} B$ as $t \to +\infty$.

Proof. Notice that the existence and uniqueness of the trajectory x follows from Theorem 2, since B is $1/\beta$-Lipschitz continuous, Γ is $\|\Gamma\|$-Lipschitz continuous and (A2) ensures $\lambda(\cdot) \in L^1_{\text{loc}}([0, +\infty))$.

(i) Take an arbitrary $x^* \in \operatorname{zer} B$ and consider for every $t \in [0, +\infty)$ the function $h(t) = \frac{1}{2}\|x(t) - x^*\|^2$. We have $\dot{h}(t) = \langle x(t) - x^*, \dot{x}(t)\rangle$ and $\ddot{h}(t) = \|\dot{x}(t)\|^2 + \langle x(t) - x^*, \ddot{x}(t)\rangle$ for almost every $t \in [0, +\infty)$. Taking into account (6), we get for almost every $t \in [0, +\infty)$

$$\ddot{h}(t) + \gamma\dot{h}(t) + \lambda(t)\left\langle x(t) - x^*, B(x(t))\right\rangle + \left\langle x(t) - x^*, \Gamma(\dot{x}(t)) - \gamma\dot{x}(t)\right\rangle = \|\dot{x}(t)\|^2. \tag{11}$$

Now we introduce the function $p : [0, +\infty) \to \mathbb{R}$,

$$p(t) = \frac{1}{2}\left\langle (\Gamma - \gamma \operatorname{Id})(x(t) - x^*), x(t) - x^*\right\rangle, \tag{12}$$

where Id denotes the identity on $\mathcal{H}$. Due to (A1), as $\langle (\Gamma - \gamma \operatorname{Id}) u, u\rangle \geq 0$ for all $u \in \mathcal{H}$, it holds

$$p(t) \geq 0 \text{ for all } t \geq 0. \tag{13}$$

Moreover, $\dot{p}(t) = \langle (\Gamma - \gamma \operatorname{Id})(\dot{x}(t)), x(t) - x^*\rangle$, which combined with (11), the cocoercivity of B and the fact that $Bx^* = 0$ yields for almost every $t \in [0, +\infty)$

$$\ddot{h}(t) + \gamma\dot{h}(t) + \beta\lambda(t)\|B(x(t))\|^2 + \dot{p}(t) \leq \|\dot{x}(t)\|^2.$$

Taking into account (6) one obtains for almost every $t \in [0, +\infty)$

$$\ddot{h}(t) + \gamma\dot{h}(t) + \frac{\beta}{\lambda(t)}\|\ddot{x}(t) + \Gamma(\dot{x}(t))\|^2 + \dot{p}(t) \leq \|\dot{x}(t)\|^2,$$

hence

$$\ddot{h}(t) + \gamma\dot{h}(t) + \frac{\beta}{\lambda(t)}\|\ddot{x}(t)\|^2 + \frac{2\beta}{\lambda(t)}\left\langle \ddot{x}(t), \Gamma(\dot{x}(t))\right\rangle + \frac{\beta}{\lambda(t)}\|\Gamma(\dot{x}(t))\|^2 + \dot{p}(t) \leq \|\dot{x}(t)\|^2. \tag{14}$$

According to (A1) we have

$$\gamma\|u\| \leq \|\Gamma u\| \text{ for all } u \in \mathcal{H}, \tag{15}$$

which combined with (14) yields for almost every $t \in [0, +\infty)$

$$\ddot{h}(t) + \gamma\dot{h}(t) + \dot{p}(t) + \frac{\beta}{\lambda(t)}\frac{d}{dt}\big(\left\langle \dot{x}(t), \Gamma(\dot{x}(t))\right\rangle\big) + \left(\frac{\beta\gamma^2}{\lambda(t)} - 1\right)||\dot{x}(t)||^2 + \frac{\beta}{\lambda(t)}||\ddot{x}(t)||^2 \leq 0.$$

By taking into account that for almost every $t \in [0, +\infty)$

$$\begin{aligned}\frac{1}{\lambda(t)}\frac{d}{dt}\big(\left\langle \dot{x}(t), \Gamma(\dot{x}(t))\right\rangle\big) = & \frac{d}{dt}\left(\frac{1}{\lambda(t)}\left\langle \dot{x}(t), \Gamma(\dot{x}(t))\right\rangle\right) + \frac{\dot{\lambda}(t)}{\lambda^2(t)}\left\langle \dot{x}(t), \Gamma(\dot{x}(t))\right\rangle \\ \geq & \frac{d}{dt}\left(\frac{1}{\lambda(t)}\left\langle \dot{x}(t), \Gamma(\dot{x}(t))\right\rangle\right) + \gamma\frac{\dot{\lambda}(t)}{\lambda^2(t)}\|\dot{x}(t)\|^2, \end{aligned} \tag{16}$$

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

$$\begin{aligned}&\ddot{h}(t) + \gamma\dot{h}(t) + \dot{p}(t) + \\ &\beta\frac{d}{dt}\left(\frac{1}{\lambda(t)}\left\langle \dot{x}(t), \Gamma(\dot{x}(t))\right\rangle\right) + \left(\frac{\beta\gamma^2}{\lambda(t)} + \beta\gamma\frac{\dot{\lambda}(t)}{\lambda^2(t)} - 1\right)||\dot{x}(t)||^2 + \frac{\beta}{\lambda(t)}||\ddot{x}(t)||^2 \leq 0. \end{aligned} \tag{17}$$

By using now assumption (A2) we obtain that the following inequality holds for almost every $t \in [0, +\infty)$

$$\ddot{h}(t) + \gamma\dot{h}(t) + \dot{p}(t) + \beta\frac{d}{dt}\left(\frac{1}{\lambda(t)}\left\langle \dot{x}(t), \Gamma(\dot{x}(t))\right\rangle\right) + \theta||\dot{x}(t)||^2 + \frac{1+\theta}{\gamma^2}||\ddot{x}(t)||^2 \leq 0. \tag{18}$$

This implies that the function $t \mapsto \dot{h}(t) + \gamma h(t) + p(t) + \frac{\beta}{\lambda(t)}\left\langle \dot{x}(t), \Gamma(\dot{x}(t))\right\rangle$, which is locally absolutely continuous, is monotonically decreasing. Hence there exists a real number M such that for almost every $t \in [0, +\infty)$

$$\dot{h}(t) + \gamma h(t) + p(t) + \frac{\beta}{\lambda(t)}\left\langle \dot{x}(t), \Gamma(\dot{x}(t))\right\rangle \leq M, \tag{19}$$

which yields, together with (13) and (A2), that for almost every $t \in [0, +\infty)$

$$\dot{h}(t) + \gamma h(t) \leq M.$$

By multiplying this inequality with $e^{\gamma t}$ and then integrating from 0 to T, where $T > 0$, one easily obtains

$$h(T) \leq h(0)e^{-\gamma T} + \frac{M}{\gamma}(1 - e^{-\gamma T}),$$

thus

$$h \text{ is bounded} \tag{20}$$

and, consequently,

$$\text{the trajectory } x \text{ is bounded.} \tag{21}$$

On the other hand, from (19), by taking into account (13), (A1) and (A2), it follows that for almost every $t \in [0, +\infty)$

$$\dot{h}(t) + \frac{1+\theta}{\gamma}\|\dot{x}(t)\|^2 \leq M,$$

hence

$$\langle x(t)-x^*, \dot{x}(t)\rangle + \frac{1+\theta}{\gamma}\|\dot{x}(t)\|^2 \leq M.$$

This inequality, in combination with (21), yields

$$\dot{x} \text{ is bounded,} \tag{22}$$

which further implies that

$$\dot{h} \text{ is bounded.} \tag{23}$$

Integrating the inequality (18) we obtain that there exists a real number $N \in \mathbb{R}$ such that for almost every $t \in [0, +\infty)$

$$\dot{h}(t) + \gamma h(t) + p(t) + \frac{\beta}{\lambda(t)}\left\langle \dot{x}(t), \Gamma(\dot{x}(t))\right\rangle + \theta\int_0^t ||\dot{x}(s)||^2 ds + \frac{1+\theta}{\gamma^2}\int_0^t ||\ddot{x}(s)||^2 ds \leq N.$$

From here, via (23), (13) and (A1), we conclude that $\dot{x}(\cdot), \ddot{x}(\cdot) \in L^2([0,+\infty);\mathcal{H})$. Finally, from (6), (A1) and (A2) we deduce $Bx \in L^2([0,+\infty);\mathcal{H})$ and the proof of (i) is complete.

(ii) For almost every $t \in [0, +\infty)$ it holds

$$\frac{d}{dt}\left(\frac{1}{2}\|\dot{x}(t)\|^2\right) = \langle \dot{x}(t), \ddot{x}(t)\rangle \leq \frac{1}{2}\|\dot{x}(t)\|^2 + \frac{1}{2}\|\ddot{x}(t)\|^2$$

and Lemma 4 together with (i) lead to $\lim_{t\to+\infty}\dot{x}(t) = 0$.

Further, by taking into consideration Remark 1(b), for almost every $t \in [0, +\infty)$ we have

$$\frac{d}{dt}\left(\frac{1}{2}\|B(x(t))\|^2\right) = \left\langle B(x(t)), \frac{d}{dt}(Bx(t))\right\rangle \leq \frac{1}{2}\|B(x(t))\|^2 + \frac{1}{2\beta^2}\|\dot{x}(t)\|^2.$$

By using again Lemma 4 and (i) we get $\lim_{t\to+\infty} B(x(t)) = 0$, while the fact that $\lim_{t\to+\infty}\ddot{x}(t) = 0$ follows from (6), (A1) and (A2).

(iii) We are going to prove that both assumptions in Opial Lemma are fulfilled. The first one concerns the existence of $\lim_{t\to+\infty}\|x(t)-x^*\|$. As seen in the proof of part (i), the function $t \mapsto \dot{h}(t) + \gamma h(t) + p(t) + \frac{\beta}{\lambda(t)}\left\langle \dot{x}(t), \Gamma(\dot{x}(t))\right\rangle$ is monotonically decreasing, thus from (i), (ii), (13), (A1) and (A2) we deduce that $\lim_{t\to+\infty}(\gamma h(t) + p(t))$ exists and it is a real number. It remains to prove that $\lim_{t\to+\infty} p(t)$ exists and it is a real number and this will prove the first part of the Opial Lemma.

Indeed, from (18) we get that for almost every $t \in [0, +\infty)$

$$\dot{p}(t) \leq -\ddot{h}(t) - \gamma\dot{h}(t) - \beta\frac{d}{dt}\left(\frac{1}{\lambda(t)}\left\langle \dot{x}(t), \Gamma(\dot{x}(t))\right\rangle\right). \tag{24}$$

On the other hand, by (A1), for every $T \geq 0$ we have

$$\int_0^T \left[-\ddot{h}(t) - \gamma\dot{h}(t) - \beta\frac{d}{dt}\left(\frac{1}{\lambda(t)}\left\langle \dot{x}(t), \Gamma(\dot{x}(t))\right\rangle\right)\right] dt =$$

$$- \dot{h}(T) - \gamma h(T) - \frac{\beta}{\lambda(T)}\left\langle \dot{x}(T), \Gamma(\dot{x}(T))\right\rangle + \dot{h}(0) + \gamma h(0) + \frac{\beta}{\lambda(0)}\left\langle \dot{x}(0), \Gamma(\dot{x}(0))\right\rangle \leq$$

$$- \dot{h}(T) + \dot{h}(0) + \gamma h(0) + \frac{\beta}{\lambda(0)}\left\langle \dot{x}(0), \Gamma(\dot{x}(0))\right\rangle.$$

Since $\lim_{T\to+\infty}\dot{h}(T)=0$ (see (i) and (ii)), we deduce that the function $t\mapsto -\ddot{h}(t)-\gamma\dot{h}(t)-\beta\frac{d}{dt}\left(\frac{1}{\lambda(t)}\langle\dot{x}(t),\Gamma(\dot{x}(t))\rangle\right)$ is in $L^1([0,+\infty))$. From Lemma 3 it follows that there exists $\lim_{t\to+\infty}p(t)\in\mathbb{R}$.

We come now to the second assumption of the Opial Lemma. Let $\overline{x}$ be a weak sequential cluster point of x, that is, there exists a sequence $t_n\to+\infty$ (as $n\to+\infty$) such that $(x(t_n))_{n\in\mathbb{N}}$ converges weakly to $\overline{x}$. Since B is a maximally monotone operator (see for instance [13, Example 20.28]), its graph is sequentially closed with respect to the weak-strong topology of the product space $\mathcal{H}\times\mathcal{H}$. By using also that $\lim_{n\to+\infty}B(x(t_n))=0$, we conclude that $B\overline{x}=0$, hence $\overline{x}\in\operatorname{zer}B$ and the proof is complete. ■

A standard instance of a cocoercive operator defined on a real Hilbert spaces is the one that can be represented as $B=\operatorname{Id}-T$, where $T:\mathcal{H}\to\mathcal{H}$ is a *nonexpansive operator*, that is, a 1-Lipschitz continuous operator. As it easily follows from the nonexpansiveness of T, B is in this case 1/2-cocoercive. For this particular choice of the operator B, the dynamical system (6) becomes

$$\left\{\begin{array}{l}\ddot{x}(t)+\Gamma(\dot{x}(t))+\lambda(t)\big(x(t)-T(x(t))\big)=0\\ x(0)=u_0,\dot{x}(0)=v_0,\end{array}\right. \tag{25}$$

while assumption (A2) reads

(A3) $\lambda:[0,+\infty)\to(0,+\infty)$ is locally absolutely continuous and there exists $\theta>0$ such that for almost every $t\in[0,+\infty)$ we have

$$\dot{\lambda}(t)\geq 0 \text{ and } \lambda(t)\leq\frac{\gamma^2}{2(1+\theta)}. \tag{26}$$

Theorem 6 gives rise to the following result.

Corollary 7 *Let $T:\mathcal{H}\to\mathcal{H}$ be a nonexpansive operator such that* $\operatorname{Fix}T=\{u\in\mathcal{H}: Tu=u\}\neq\emptyset$, $\Gamma:\mathcal{H}\to\mathcal{H}$ *be an operator fulfilling* (A1), $\lambda:[0,+\infty)\to(0,+\infty)$ *be a function fulfilling* (A3) *and* $u_0,v_0\in\mathcal{H}$. *Let $x:[0,+\infty)\to\mathcal{H}$ be the unique strong global solution of* (25). *Then the following statements are true:*

(i) the trajectory x is bounded and $\dot{x},\ddot{x},(\operatorname{Id}-T)x\in L^2([0,+\infty);\mathcal{H})$;

(ii) $\lim_{t\to+\infty}\dot{x}(t)=\lim_{t\to+\infty}\ddot{x}(t)=\lim_{t\to+\infty}(\operatorname{Id}-T)(x(t))=0$;

(iii) $x(t)$ converges weakly to a point in $\operatorname{Fix}T$ *as $t\to+\infty$.*

Remark 8 In the particular case when $\Gamma=\gamma\operatorname{Id}$ for $\gamma>0$ and $\lambda(t)=1$ for all $t\in[0,+\infty)$ the dynamical system (25) becomes

$$\left\{\begin{array}{l}\ddot{x}(t)+\gamma\dot{x}(t)+x(t)-T(x(t))=0\\ x(0)=u_0,\dot{x}(0)=v_0.\end{array}\right. \tag{27}$$

The convergence of the trajectories generated by (27) has been studied in [8, Theorem 3.2] under the condition $\gamma^2>2$. In this case (A3) is obviously fulfilled for an arbitrary $0<\theta\leq(\gamma^2-2)/2$. However, different to [8], we allow in Corollary 7 an anisotropic damping through the use of the elliptic operator Γ and also a variable relaxation function λ depending on time (in [3] the anisotropic damping has been considered as well in the context of minimizing of a smooth convex function via second order dynamical systems).

We close the section by addressing an immediate consequence of the above corollary applied to second order dynamical systems governed by averaged operators. The operator $R : \mathcal{H} \to \mathcal{H}$ is said to be *α-averaged* for $\alpha \in (0,1)$, if there exists a nonexpansive operator $T : \mathcal{H} \to \mathcal{H}$ such that $R = (1-\alpha)\operatorname{Id} + \alpha T$. For $\alpha = \frac{1}{2}$ we obtain as an important representative of this class the *firmly nonexpansive* operators. For properties and insights concerning these families of operators we refer to the monograph [13].

We consider the dynamical system

$$\begin{cases} \ddot{x}(t) + \Gamma(\dot{x}(t)) + \lambda(t)\big(x(t) - R(x(t))\big) = 0 \\ x(0) = u_0, \dot{x}(0) = v_0 \end{cases} \tag{28}$$

and formulate the assumption

(A4) $\lambda : [0,+\infty) \to (0,+\infty)$ is locally absolutely continuous and there exists $\theta > 0$ such that for almost every $t \in [0,+\infty)$ we have

$$\dot{\lambda}(t) \geq 0 \text{ and } \lambda(t) \leq \frac{\gamma^2}{2\alpha(1+\theta)}. \tag{29}$$

Corollary 9 *Let $R : \mathcal{H} \to \mathcal{H}$ be an α-averaged operator for $\alpha \in (0,1)$ such that $\operatorname{Fix} R \neq \emptyset$, $\Gamma : \mathcal{H} \to \mathcal{H}$ be an operator fulfilling (A1), $\lambda : [0,+\infty) \to (0,+\infty)$ be a function fulfilling (A4) and $u_0, v_0 \in \mathcal{H}$. Let $x : [0,+\infty) \to \mathcal{H}$ be the unique strong global solution of (28). Then the following statements are true:*

(i) the trajectory x is bounded and $\dot{x}, \ddot{x}, (\operatorname{Id} - R)x \in L^2([0,+\infty);\mathcal{H})$;

(ii) $\lim_{t\to+\infty} \dot{x}(t) = \lim_{t\to+\infty} \ddot{x}(t) = \lim_{t\to+\infty} (\operatorname{Id} - R)(x(t)) = 0$;

(iii) $x(t)$ converges weakly to a point in $\operatorname{Fix} R$ as $t \to +\infty$.

Proof. Since R is α-averaged, there exists a nonexpansive operator $T : \mathcal{H} \to \mathcal{H}$ such that $R = (1-\alpha)\operatorname{Id} + \alpha T$. The conclusion is a direct consequence of Corollary 7, by taking into account that (28) is equivalent to

$$\begin{cases} \ddot{x}(t) + \Gamma(\dot{x}(t)) + \alpha\lambda(t)\big(x(t) - T(x(t))\big) = 0 \\ x(0) = u_0, \dot{x}(0) = v_0, \end{cases}$$

and $\operatorname{Fix} R = \operatorname{Fix} T$. ■

4 Forward-backward second order dynamical systems

In this section we address the monotone inclusion problem

$$\text{find } 0 \in A(x) + B(x),$$

where $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is a maximally monotone operator and $B : \mathcal{H} \to \mathcal{H}$ is a β-cocoercive operator for $\beta > 0$ via a second-order forward-backward dynamical system with anisotropic damping and variable relaxation parameter.

For readers convenience we recall at the beginning some standard notions and results in monotone operator theory which will be used in the following (see also [13,15,26]). For an arbitrary set-valued operator $A : \mathcal{H} \rightrightarrows \mathcal{H}$ we denote by $\operatorname{Gr} A = \{(x,u) \in \mathcal{H} \times \mathcal{H} : u \in Ax\}$ its graph. We use also the notation $\operatorname{zer} A = \{x \in \mathcal{H} : 0 \in Ax\}$ for the set of zeros of A. We say that A is monotone, if $\langle x-y, u-v\rangle \geq 0$ for all $(x,u),(y,v) \in \operatorname{Gr} A$. A monotone operator A is said to be maximally monotone, if there exists no proper monotone extension of the graph of A on $\mathcal{H}\times\mathcal{H}$. The resolvent of A, $J_A : \mathcal{H} \rightrightarrows \mathcal{H}$, is defined by $J_A = (\operatorname{Id} + A)^{-1}$. If A is maximally monotone, then $J_A : \mathcal{H} \to \mathcal{H}$ is single-valued and maximally monotone (see [13, Proposition 23.7 and Corollary 23.10]). For an arbitrary $\gamma > 0$ we have (see [13, Proposition 23.2])

$$p \in J_{\gamma A}x \text{ if and only if } (p, \gamma^{-1}(x-p)) \in \operatorname{Gr} A. \tag{30}$$

The operator A is said to be uniformly monotone if there exists an increasing function $\phi_A : [0,+\infty) \to [0,+\infty]$ that vanishes only at 0 and fulfills $\langle x-y, u-v\rangle \geq \phi_A(\|x-y\|)$ for every $(x,u) \in \operatorname{Gr} A$ and $(y,v) \in \operatorname{Gr} A$. A popular class of operators having this property is the one of the strongly monotone operators. We say that A is γ-strongly monotone for $\gamma > 0$, if $\langle x-y, u-v\rangle \geq \gamma\|x-y\|^2$ for all $(x,u),(y,v) \in \operatorname{Gr} A$.

For $\eta > 0$ we consider the dynamical system

$$\left\{ \begin{array}{l} \ddot{x}(t) + \Gamma(\dot{x}(t)) + \lambda(t)\left[x(t) - J_{\eta A}\Big(x(t) - \eta B(x(t))\Big)\right] = 0 \\ x(0) = u_0, \dot{x}(0) = v_0. \end{array} \right. \tag{31}$$

We formulate the following assumption, where $\delta := \min\{1, \beta/\eta\} + 1/2$:

(A5) $\lambda : [0,+\infty) \to (0,+\infty)$ is locally absolutely continuous and there exists $\theta > 0$ such that for almost every $t \in [0,+\infty)$ we have

$$\dot{\lambda}(t) \geq 0 \text{ and } \lambda(t) \leq \frac{\delta\gamma^2}{2(1+\theta)}. \tag{32}$$

Theorem 10 *Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be a maximally monotone operator and $B : \mathcal{H} \to \mathcal{H}$ be β-cocoercive operator for $\beta > 0$ such that $\operatorname{zer}(A+B) \neq \emptyset$. Let $\eta \in (0, 2\beta)$ and set $\delta := \min\{1, \beta/\eta\} + 1/2$. Let $\Gamma : \mathcal{H} \to \mathcal{H}$ be an operator fulfilling (A1), $\lambda : [0,+\infty) \to (0,+\infty)$ be a function fulfilling (A5), $u_0, v_0 \in \mathcal{H}$ and $x : [0,+\infty) \to \mathcal{H}$ be the unique strong global solution of (31). Then the following statements are true:*

(i) the trajectory x is bounded and $\dot{x}, \ddot{x}, \big(\operatorname{Id} - J_{\eta A} \circ (\operatorname{Id} - \eta B)\big)x \in L^2([0,+\infty); \mathcal{H})$;

(ii) $\lim_{t\to+\infty} \dot{x}(t) = \lim_{t\to+\infty} \ddot{x}(t) = \lim_{t\to+\infty} \big(\operatorname{Id} - J_{\eta A} \circ (\operatorname{Id} - \eta B)\big)(x(t)) = 0$;

(iii) $x(t)$ converges weakly to a point in $\operatorname{zer}(A+B)$ as $t \to +\infty$;

(iv) if $x^* \in \operatorname{zer}(A+B)$, then $B(x(\cdot)) - Bx^* \in L^2([0,+\infty); \mathcal{H})$, $\lim_{t\to+\infty} B(x(t)) = Bx^*$ and B is constant on $\operatorname{zer}(A+B)$;

(v) if A or B is uniformly monotone, then $x(t)$ converges strongly to the unique point in $\operatorname{zer}(A+B)$ as $t \to +\infty$.

Proof. (i)-(iii) It is immediate that the dynamical system (31) can be written in the form

$$\left\{ \begin{array}{l} \ddot{x}(t) + \Gamma(\dot{x}(t)) + \lambda(t)\big(x(t) - R(x(t))\big) = 0 \\ x(0) = u_0, \dot{x}(0) = v_0, \end{array} \right. \tag{33}$$

where $R = J_{\eta A} \circ (\operatorname{Id} - \eta B)$. According to [13, Corollary 23.8 and Remark 4.24(iii)], $J_{\eta A}$ is 1/2-cocoercive. Moreover, by [13, Proposition 4.33], $\operatorname{Id} - \eta B$ is $\eta/(2\beta)$-averaged. Combining this with [13, Proposition 4.32], we derive that R is $1/\delta$-averaged. The statements (i)-(iii) follow now from Corollary 9 by noticing that $\operatorname{Fix} R = \operatorname{zer}(A+B)$ (see [13, Proposition 25.1(iv)]).

(iv) The fact that B is constant on $\operatorname{zer}(A+B)$ follows from the cocoercivity of B and the monotonicity of A. A proof of this statement when A is the subdifferential of a proper, convex and lower semicontinuous function is given for instance in [1, Lemma 1.7].

Take an arbitrary $x^* \in \operatorname{zer}(A+B)$. From the definition of the resolvent we have for almost every $t \in [0, +\infty)$

$$-B(x(t)) - \frac{1}{\eta\lambda(t)}\ddot{x}(t) - \frac{1}{\eta\lambda(t)}\Gamma(\dot{x}(t)) \in A\left(\frac{1}{\lambda(t)}\ddot{x}(t) + \frac{1}{\lambda(t)}\Gamma(\dot{x}(t)) + x(t)\right), \tag{34}$$

which combined with $-Bx^* \in Ax^*$ and the monotonicity of A leads to

$$0 \leq \left\langle \frac{1}{\lambda(t)}\ddot{x}(t) + \frac{1}{\lambda(t)}\Gamma(\dot{x}(t)) + x(t) - x^*, -B(x(t)) + Bx^* - \frac{1}{\eta\lambda(t)}\ddot{x}(t) - \frac{1}{\eta\lambda(t)}\Gamma(\dot{x}(t)) \right\rangle. \tag{35}$$

After using the cocoercivity of B we obtain for almost every $t \in [0, +\infty)$

$$\begin{aligned}
\beta\|B(x(t)) - Bx^*\|^2 \leq& \left\langle \frac{1}{\lambda(t)}\ddot{x}(t) + \frac{1}{\lambda(t)}\Gamma(\dot{x}(t)), -B(x(t)) + Bx^* \right\rangle - \frac{1}{\eta\lambda^2(t)}\|\ddot{x}(t) + \Gamma(\dot{x}(t))\|^2 \\
&+ \left\langle x(t) - x^*, -\frac{1}{\eta\lambda(t)}\ddot{x}(t) - \frac{1}{\eta\lambda(t)}\Gamma(\dot{x}(t)) \right\rangle \\
\leq& \frac{1}{2\beta}\left\|\frac{1}{\lambda(t)}\ddot{x}(t) + \frac{1}{\lambda(t)}\Gamma(\dot{x}(t))\right\|^2 + \frac{\beta}{2}\|B(x(t)) - Bx^*\|^2 \\
&+ \left\langle x(t) - x^*, -\frac{1}{\eta\lambda(t)}\ddot{x}(t) - \frac{1}{\eta\lambda(t)}\Gamma(\dot{x}(t)) \right\rangle.
\end{aligned}$$

For evaluating the last term of the above inequality we use the functions $h : [0, +\infty) \to \mathbb{R}, h(t) = \frac{1}{2}\|x(t) - x^*\|^2$ and $p : [0, +\infty) \to \mathbb{R}$, $p(t) = \frac{1}{2}\langle(\Gamma - \gamma \operatorname{Id})(x(t) - x^*), x(t) - x^*\rangle$, already used in the proof of Theorem 6. For almost every $t \in [0, +\infty)$ we have

$$\langle x(t) - x^*, \ddot{x}(t)\rangle = \ddot{h}(t) - \|\dot{x}(t)\|^2$$

and

$$\dot{p}(t) = \langle x(t) - x^*, \Gamma(\dot{x}(t))\rangle - \gamma\langle x(t) - x^*, \dot{x}(t)\rangle = \langle x(t) - x^*, \Gamma(\dot{x}(t))\rangle - \gamma\dot{h}(t),$$

hence

$$\left\langle x(t) - x^*, -\frac{1}{\eta\lambda(t)}\ddot{x}(t) - \frac{1}{\eta\lambda(t)}\Gamma(\dot{x}(t)) \right\rangle = -\frac{1}{\eta\lambda(t)}\left(\ddot{h}(t) + \gamma\dot{h}(t) + \dot{p}(t) - \|\dot{x}(t)\|^2\right). \tag{36}$$

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

$$\begin{aligned}
\frac{\beta}{2}\|B(x(t)) - Bx^*\|^2 \leq& \frac{1}{2\beta}\left\|\frac{1}{\lambda(t)}\ddot{x}(t) + \frac{1}{\lambda(t)}\Gamma(\dot{x}(t))\right\|^2 \\
&- \frac{1}{\eta\lambda(t)}\left(\ddot{h}(t) + \gamma\dot{h}(t) + \dot{p}(t) - \|\dot{x}(t)\|^2\right).
\end{aligned} \tag{37}$$

By taking into account (A5) we obtain a lower bound $\underline{\lambda}$ such that for almost every $t \in [0, +\infty)$ one has

$$0 < \underline{\lambda} \leq \lambda(t) \leq \frac{\delta\gamma^2}{2(1+\theta)}.$$

By multiplying (37) with $\lambda(t)$ we obtain for almost every $t \in [0, +\infty)$ that

$$\frac{\beta\underline{\lambda}}{2}\|B(x(t)) - Bx^*\|^2 + \frac{1}{\eta}\left(\ddot{h}(t) + \gamma\dot{h}(t) + \dot{p}(t)\right) \leq \frac{1}{2\beta\underline{\lambda}}\|\ddot{x}(t) + \Gamma(\dot{x}(t))\|^2 + \frac{1}{\eta}\|\dot{x}(t)\|^2.$$

After integration we obtain that for every $T \in [0, +\infty)$

$$\frac{\beta\underline{\lambda}}{2}\int_0^T \|B(x(t)) - Bx^*\|^2 dt + \frac{1}{\eta}\left(\dot{h}(T) - \dot{h}(0) + \gamma h(T) - \gamma h(0) + p(T) - p(0)\right)$$
$$\leq \int_0^T \left(\frac{1}{2\beta\underline{\lambda}}\|\ddot{x}(t) + \Gamma(\dot{x}(t))\|^2 + \frac{1}{\eta}\|\dot{x}(t)\|^2\right) dt.$$

As $\dot{x}, \ddot{x} \in L^2([0, +\infty); \mathcal{H})$, $h(T) \geq 0, p(T) \geq 0$ for every $T \in [0, +\infty)$ and $\lim_{T\to+\infty} \dot{h}(T) = 0$, it follows that $B(x(\cdot)) - Bx^* \in L^2([0, +\infty); \mathcal{H})$.

Further, by taking into consideration Remark 1(b), we have

$$\frac{d}{dt}\left(\frac{1}{2}\|B(x(t)) - Bx^*\|^2\right) = \left\langle B(x(t)) - Bx^*, \frac{d}{dt}(Bx(t))\right\rangle$$
$$\leq \frac{1}{2}\|B(x(t)) - Bx^*\|^2 + \frac{1}{2\beta^2}\|\dot{x}(t)\|^2$$

and from here, in light of Lemma 4, it follows that $\lim_{t\to+\infty} B(x(t)) = Bx^*$.

(v) Let x^* be the unique element of $\operatorname{zer}(A + B)$. For the beginning we suppose that A is uniformly monotone with corresponding function $\phi_A : [0, +\infty) \to [0, +\infty]$, which is increasing and vanishes only at 0.

By similar arguments as in the proof of statement (iv), for almost every $t \in [0, +\infty)$ we have

$$\phi_A\left(\left\|\frac{1}{\lambda(t)}\ddot{x}(t) + \frac{1}{\lambda(t)}\Gamma(\dot{x}(t)) + x(t) - x^*\right\|\right) \leq$$
$$\left\langle \frac{1}{\lambda(t)}\ddot{x}(t) + \frac{1}{\lambda(t)}\Gamma(\dot{x}(t)) + x(t) - x^*, -B(x(t)) + Bx^* - \frac{1}{\eta\lambda(t)}\ddot{x}(t) - \frac{1}{\eta\lambda(t)}\Gamma(\dot{x}(t))\right\rangle,$$

which combined with the inequality

$$\langle x(t) - x^*, B(x(t)) - Bx^*\rangle \geq 0$$

yields

$$\phi_A\left(\left\|\frac{1}{\lambda(t)}\ddot{x}(t) + \frac{1}{\lambda(t)}\Gamma(\dot{x}(t)) + x(t) - x^*\right\|\right) \leq$$
$$\left\langle \frac{1}{\lambda(t)}\ddot{x}(t) + \frac{1}{\lambda(t)}\Gamma(\dot{x}(t)), -B(x(t)) + Bx^*\right\rangle - \frac{1}{\eta\lambda^2(t)}\|\ddot{x}(t) + \Gamma(\dot{x}(t))\|^2 +$$
$$\left\langle x(t) - x^*, -\frac{1}{\eta\lambda(t)}\ddot{x}(t) - \frac{1}{\eta\lambda(t)}\Gamma(\dot{x}(t))\right\rangle \leq$$
$$\left\langle \frac{1}{\lambda(t)}\ddot{x}(t) + \frac{1}{\lambda(t)}\Gamma(\dot{x}(t)), -B(x(t)) + Bx^*\right\rangle + \left\langle x(t) - x^*, -\frac{1}{\eta\lambda(t)}\ddot{x}(t) - \frac{1}{\eta\lambda(t)}\Gamma(\dot{x}(t))\right\rangle.$$

As λ is bounded by positive constants, by using (i)-(iv) it follows that the right-hand side of the last inequality converges to 0 as $t \to +\infty$. Hence

$$\lim_{t\to+\infty} \phi_A\left(\left\|\frac{1}{\lambda(t)}\ddot{x}(t) + \frac{1}{\lambda(t)}\Gamma(\dot{x}(t)) + x(t) - x^*\right\|\right) = 0$$

and the properties of the function ϕ_A allow to conclude that $\frac{1}{\lambda(t)}\ddot{x}(t)+\frac{1}{\lambda(t)}\Gamma(\dot{x}(t))+x(t)-x^*$ converges strongly to 0 as $t \to +\infty$. By using again the boundedness of λ and (ii) we obtain that $x(t)$ converges strongly to x^* as $t \to +\infty$.

Finally, suppose that B is uniformly monotone with corresponding function $\phi_B : [0,+\infty) \to [0,+\infty]$, which is increasing and vanishes only at 0. The conclusion follows by letting t in the inequality

$$\langle x(t) - x^*, B(x(t)) - Bx^*\rangle \geq \phi_B(\|x(t) - x^*\|)\ \forall t \in [0,+\infty)$$

converge to $+\infty$ and by using that x is bounded and $\lim_{t\to+\infty}(B(x(t) - Bx^*) = 0$. ■

Remark 11 We would like to emphasize the fact that the statements in Theorem 10 remain valid also for $\eta := 2\beta$. Indeed, in this case the cocoercivity of B implies that $\operatorname{Id} - \eta B$ is nonexpansive, hence the operator $R = J_{\eta A} \circ (\operatorname{Id} - \eta B)$ used in the proof is nonexpansive, too, and so the statements in (i)-(iii) follow from Corollary 7. Furthermore, the proof of the statements (iv) and (v) can be repeated also for $\eta = 2\beta$.

In the remaining of this section we turn our attention to optimization problems of the form

$$\min_{x\in\mathcal{H}} f(x) + g(x),$$

where $f : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ is a proper, convex and lower semicontinuous function and $g : \mathcal{H} \to \mathbb{R}$ is a convex and (Fréchet) differentiable function with $1/\beta$-Lipschitz continuous gradient for $\beta > 0$.

We recall some standard notations and facts in convex analysis. For a proper, convex and lower semicontinuous function $f : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$, its (convex) subdifferential at $x \in \mathcal{H}$ is defined as

$$\partial f(x) = \{u \in \mathcal{H} : f(y) \geq f(x) + \langle u, y - x\rangle\ \ \forall y \in \mathcal{H}\}.$$

When seen as a set-valued mapping, it is a maximally monotone operator (see [24]) and its resolvent is given by $J_{\eta\partial f} = \operatorname{prox}_{\eta f}$ (see [13]), where $\operatorname{prox}_{\eta f} : \mathcal{H} \to \mathcal{H}$,

$$\operatorname{prox}_{\eta f}(x) = \operatorname*{argmin}_{y\in\mathcal{H}} \left\{f(y) + \frac{1}{2\eta}\|y - x\|^2\right\}, \tag{38}$$

denotes the proximal point operator of f and $\eta > 0$. According to [13, Definition 10.5], f is said to be uniformly convex with modulus function $\phi : [0,+\infty) \to [0,+\infty]$, if ϕ is increasing, vanishes only at 0 and fulfills $f(\alpha x + (1-\alpha)y) + \alpha(1-\alpha)\phi(\|x - y\|) \leq \alpha f(x) + (1-\alpha)f(y)$ for all $\alpha \in (0,1)$ and $x, y \in \operatorname{dom} f := \{x \in \mathcal{H} : f(x) < +\infty\}$. Notice that if this inequality holds for $\phi = (\nu/2)|\cdot|^2$ for $\nu > 0$, then f is said to be ν-strongly convex.

In the following statement we approach the minimizers of $f+g$ via the second order dynamical system

$$\left\{\begin{array}{l} \ddot{x}(t)+\Gamma(\dot{x}(t))+\lambda(t)\Big[x(t)-\operatorname{prox}_{\eta f}\Big(x(t)-\eta\nabla g(x(t))\Big)\Big]=0\\ x(0)=u_0,\dot{x}(0)=v_0. \end{array}\right. \tag{39}$$

Corollary 12 *Let $f:\mathcal{H}\rightarrow\mathbb{R}\cup\{+\infty\}$ by a proper, convex and lower semicontinuous function and $g:\mathcal{H}\rightarrow\mathbb{R}$ be a convex and (Fréchet) differentiable function with $1/\beta$-Lipschitz continuous gradient for $\beta>0$ such that $\operatorname{argmin}_{x\in\mathcal{H}}\{f(x)+g(x)\}\neq\emptyset$. Let $\eta\in(0,2\beta]$ and set $\delta:=\min\{1,\beta/\eta\}+1/2$. Let $\Gamma:\mathcal{H}\rightarrow\mathcal{H}$ be an operator fulfilling* (A1)*, $\lambda:[0,+\infty)\rightarrow(0,+\infty)$ be a function fulfilling* (A5)*, $u_0,v_0\in\mathcal{H}$ and $x:[0,+\infty)\rightarrow\mathcal{H}$ be the unique strong global solution of* (39)*. Then the following statements are true:*

(i) the trajectory x is bounded and $\dot{x},\ddot{x},\big(\operatorname{Id}-\operatorname{prox}_{\eta f}\circ(\operatorname{Id}-\eta\nabla g)\big)x\in L^2([0,+\infty);\mathcal{H})$;

(ii) $\lim_{t\rightarrow+\infty}\dot{x}(t)=\lim_{t\rightarrow+\infty}\ddot{x}(t)=\lim_{t\rightarrow+\infty}\big(\operatorname{Id}-\operatorname{prox}_{\eta f}\circ(\operatorname{Id}-\eta\nabla g)\big)(x(t))=0$;

(iii) $x(t)$ converges weakly to a minimizer of $f+g$ as $t\rightarrow+\infty$;

(iv) if x^* is a minimizer of $f+g$, then $\nabla g(x(\cdot))-\nabla g(x^*)\in L^2([0,+\infty);\mathcal{H})$, $\lim_{t\rightarrow+\infty}\nabla g(x(t))=\nabla g(x^*)$ and ∇g is constant on $\operatorname{argmin}_{x\in\mathcal{H}}\{f(x)+g(x)\}$;

(v) if f or g is uniformly convex, then $x(t)$ converges strongly to the unique minimizer of $f+g$ as $t\rightarrow+\infty$.

Proof. The statements are direct consequences of the corresponding ones from Theorem 10 (see also Remark 11), by choosing $A:=\partial f$ and $B:=\nabla g$, by taking into account that

$$\operatorname{zer}(\partial f+\nabla g)=\operatorname*{argmin}_{x\in\mathcal{H}}\{f(x)+g(x)\}$$

and by making use of the Baillon-Haddad Theorem, which says that ∇g is $1/\beta$-Lipschitz if and only if ∇g is β-cocoercive (see [13, Corollary 18.16]). For statement (v) we also use the fact that if f is uniformly convex with modulus ϕ, then ∂f is uniformly monotone with modulus 2ϕ (see [13, Example 22.3(iii)]). ■

Remark 13 Consider again the setting in Remark 8, namely, when $\Gamma=\gamma\operatorname{Id}$ for $\gamma>0$ and $\lambda(t)=1$ for every $t\in[0,+\infty)$. Furthermore, for C a nonempty, convex, closed subset of $\mathcal{H}$, let $f=\delta_C$ be the indicator function of C, which is defined as being equal to 0 for $x\in C$ and to $+\infty$, else. The dynamical system (39) attached in this setting to the minimization of g over C becomes

$$\left\{\begin{array}{l} \ddot{x}(t)+\gamma\dot{x}(t)+x(t)-P_C\big(x(t)-\eta\nabla g(x(t))\big)=0\\ x(0)=u_0,\dot{x}(0)=v_0, \end{array}\right. \tag{40}$$

where P_C denotes the projection onto the set C.

The convergence of the trajectories of (40) has been studied in [8, Theorem 3.1] under the conditions $\gamma^2>2$ and $0<\eta\leq2\beta$. In this case assumption (A5) trivially holds by choosing θ such that $0<\theta\leq(\gamma^2-2)/2\leq(\delta\gamma^2-2)/2$. Thus, in order to verify (A5) in case $\lambda(t)=1$ for every $t\in[0,+\infty)$ one needs to equivalently assume that $\gamma^2>2/\delta$. Since $\delta\geq1$, this provides a slight improvement over [8, Theorem 3.1] in what concerns the choice of γ. We refer the reader also to [7] for an analysis of the convergence rates of trajectories of the dynamical system (40) when g is endowed with supplementary properties.

For the two main convergence statements provided in this section it was essential to choose the step size η in the interval $(0, 2\beta]$ (see Theorem 10, Remark 11 and Corollary 12). This, because of the fact that in this way we were able to guarantee for the generated trajectories the existence of the limit $\lim_{t\to+\infty} \|x(t) - x^*\|^2$, where x^* denotes a solution of the problem under investigation. It is interesting to observe that, when dealing with convex optimization problems, one can go also beyond this classical restriction concerning the choice of the step size (a similar phenomenon has been reported also in [1, Section 4.2]). This is pointed out in the following result, which is valid under the assumption

(A6) $\lambda : [0, +\infty) \to (0, +\infty)$ is locally absolutely continuous and there exist $a, \theta, \theta' > 0$ such that for almost every $t \in [0, +\infty)$ we have

$$\dot{\lambda}(t) \geq 0 \text{ and } \frac{1}{\beta}\left(\theta' + \frac{a}{2}\|\Gamma - \gamma \operatorname{Id}\|\right) \leq \lambda(t) \leq \frac{\gamma^2}{\eta\theta + \frac{\eta}{\beta} + \frac{\eta}{2a}\|\Gamma - \gamma \operatorname{Id}\| + 1}, \tag{41}$$

and for the proof of which we use instead of $\|x(\cdot) - x^*\|^2$ a modified energy functional.

Corollary 14 *Let $f : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ by a proper, convex and lower semicontinuous function and $g : \mathcal{H} \to \mathbb{R}$ be a convex and (Fréchet) differentiable function with $1/\beta$-Lipschitz continuous gradient for $\beta > 0$ such that $\operatorname{argmin}_{x\in\mathcal{H}}\{f(x) + g(x)\} \neq \emptyset$. Let be $\eta > 0$, $\Gamma : \mathcal{H} \to \mathcal{H}$ be an operator fulfilling* (A1)*, $\lambda : [0, +\infty) \to (0, +\infty)$ be a function fulfilling* (A6)*, $u_0, v_0 \in \mathcal{H}$ and $x : [0, +\infty) \to \mathcal{H}$ be the unique strong global solution of* (39)*. Then the following statements are true:*

(i) the trajectory x is bounded and $\dot{x}, \ddot{x}, \big(\operatorname{Id} - \operatorname{prox}_{\eta f} \circ (\operatorname{Id} - \eta\nabla g)\big)x \in L^2([0, +\infty); \mathcal{H})$;

(ii) $\lim_{t\to+\infty} \dot{x}(t) = \lim_{t\to+\infty} \ddot{x}(t) = \lim_{t\to+\infty} \big(\operatorname{Id} - \operatorname{prox}_{\eta f} \circ (\operatorname{Id} - \eta\nabla g)\big)(x(t)) = 0$;

(iii) $x(t)$ converges weakly to a minimizer of $f + g$ as $t \to +\infty$;

(iv) if x^* is a minimizer of $f + g$, then $\nabla g(x(\cdot)) - \nabla g(x^*) \in L^2([0, +\infty); \mathcal{H})$, $\lim_{t\to+\infty} \nabla g(x(t)) = \nabla g(x^*)$ and ∇g is constant on $\operatorname{argmin}_{x\in\mathcal{H}}\{f(x) + g(x)\}$;

(v) if f or g is uniformly convex, then $x(t)$ converges strongly to the unique minimizer of $f + g$ as $t \to +\infty$.

Proof. Consider an arbitrary element $x^* \in \operatorname{argmin}_{x\in\mathcal{H}}\{f(x) + g(x)\} = \operatorname{zer}(\partial f + \nabla g)$. Similarly to the proof of Theorem 10(iv), we derive for almot every $t \in [0, +\infty)$ (see the first inequality after (35))

$$\begin{aligned}
&\beta\|\nabla g(x(t)) - \nabla g(x^*)\|^2 \leq \\
&\frac{1}{\lambda(t)}\Big(\langle \ddot{x}(t), -\nabla g(x(t)) + \nabla g(x^*)\rangle + \langle \Gamma(\dot{x}(t)), -\nabla g(x(t)) + \nabla g(x^*)\rangle\Big) - \\
&\frac{1}{\eta\lambda^2(t)}\|\ddot{x}(t) + \Gamma(\dot{x}(t))\|^2 + \left\langle x(t) - x^*, -\frac{1}{\eta\lambda(t)}\ddot{x}(t) - \frac{1}{\eta\lambda(t)}\Gamma(\dot{x}(t))\right\rangle.
\end{aligned} \tag{42}$$

In what follows we evaluate the right-hand side of the above inequality and introduce to this end the function

$$q : [0, +\infty) \to \mathbb{R},\ q(t) = g(x(t)) - g(x^*) - \langle \nabla g(x^*), x(t) - x^*\rangle.$$

Due to the convexity of g one has

$$q(t) \geq 0\ \forall t \geq 0.$$

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

$$\dot{q}(t) = \langle \dot{x}(t), \nabla g(x(t)) - \nabla g(x^*) \rangle ,$$

thus

$$\begin{aligned}
&\langle \Gamma(\dot{x}(t)), -\nabla g(x(t)) + \nabla g(x^*) \rangle = \\
&- \gamma \dot{q}(t) + \langle (\Gamma - \gamma \operatorname{Id}) (\dot{x}(t)), -\nabla g(x(t)) + \nabla g(x^*) \rangle \leq \\
&- \gamma \dot{q}(t) + \frac{1}{2a} \|\Gamma - \gamma \operatorname{Id}\| \|\dot{x}(t)\|^2 + \frac{a}{2} \|\Gamma - \gamma \operatorname{Id}\| \|\nabla g(x(t)) - \nabla g(x^*)\|^2. \qquad (43)
\end{aligned}$$

On the other hand, for almost every $t \in [0, +\infty)$

$$\ddot{q}(t) = \langle \ddot{x}(t), \nabla g(x(t)) - \nabla g(x^*) \rangle + \left\langle \dot{x}(t), \frac{d}{dt} \nabla g(x(t)) \right\rangle ,$$

hence

$$\langle \ddot{x}(t), -\nabla g(x(t)) + \nabla g(x^*) \rangle \leq -\ddot{q}(t) + \frac{1}{\beta} \|\dot{x}(t)\|^2. \qquad (44)$$

Further, we have for almost every $t \in [0, +\infty)$ (see also (16) and (15))

$$\begin{aligned}
\frac{1}{\lambda(t)} \|\ddot{x}(t) + \Gamma(\dot{x}(t))\|^2 &= \frac{1}{\lambda(t)} \|\ddot{x}(t)\|^2 + \frac{1}{\lambda(t)} \frac{d}{dt} \big(\langle \dot{x}(t), \Gamma(\dot{x}(t)) \rangle \big) + \frac{1}{\lambda(t)} \|\Gamma(\dot{x}(t))\|^2 \\
&\geq \frac{1}{\lambda(t)} \|\ddot{x}(t)\|^2 + \frac{d}{dt} \left(\frac{1}{\lambda(t)} \langle \dot{x}(t), \Gamma(\dot{x}(t)) \rangle \right) \\
&\quad + \gamma \frac{\dot{\lambda}(t)}{\lambda^2(t)} \|\dot{x}(t)\|^2 + \frac{\gamma^2}{\lambda(t)} \|\dot{x}(t)\|^2. \qquad (45)
\end{aligned}$$

Finally, by multiplying (42) with $\lambda(t)$ and by using (43), (44), (45) and (36) we obtain after rearranging the terms for almost every $t \in [0, +\infty)$ that

$$\begin{aligned}
&\left(\beta \lambda(t) - \frac{a}{2} \|\Gamma - \gamma \operatorname{Id}\| \right) \|\nabla g(x(t)) - \nabla g(x^*)\|^2 + \frac{d}{dt^2} \left(\frac{1}{\eta} h + q \right) + \gamma \frac{d}{dt} \left(\frac{1}{\eta} h + q \right) + \\
&\frac{1}{\eta} \dot{p}(t) + \frac{1}{\eta} \frac{d}{dt} \left(\frac{1}{\lambda(t)} \langle \dot{x}(t), \Gamma(\dot{x}(t)) \rangle \right) + \\
&\left(\frac{\gamma^2}{\eta \lambda(t)} + \frac{\gamma \dot{\lambda}(t)}{\eta \lambda^2(t)} - \frac{1}{\beta} - \frac{1}{\eta} - \frac{1}{2a} \|\Gamma - \gamma \operatorname{Id}\| \right) \|\dot{x}(t)\|^2 + \frac{1}{\eta \lambda(t)} \|\ddot{x}(t)\|^2 \leq 0.
\end{aligned}$$

and, further, via (A6)

$$\begin{aligned}
&\theta' \|\nabla g(x(t)) - \nabla g(x^*)\|^2 + \frac{d}{dt^2} \left(\frac{1}{\eta} h + q \right) + \gamma \frac{d}{dt} \left(\frac{1}{\eta} h + q \right) + \frac{1}{\eta} \dot{p}(t) \\
&+ \frac{1}{\eta} \frac{d}{dt} \left(\frac{1}{\lambda(t)} \langle \dot{x}(t), \Gamma(\dot{x}(t)) \rangle \right) + \theta \|\dot{x}(t)\|^2 + \frac{1}{\eta \lambda(t)} \|\ddot{x}(t)\|^2 \leq 0. \qquad (46)
\end{aligned}$$

This implies that the function

$$t \mapsto \frac{d}{dt} \left(\frac{1}{\eta} h + q \right) (t) + \gamma \left(\frac{1}{\eta} h + q \right) (t) + \frac{1}{\eta} p(t) + \frac{1}{\eta} \left(\frac{1}{\lambda(t)} \langle \dot{x}(t), \Gamma(\dot{x}(t)) \rangle \right) \qquad (47)$$

is monotonically decreasing. Arguing as in the proof of Theorem 6, by taking into account that λ has positive upper and lower bounds, it follows that $\frac{1}{\eta}h + q$, h, q, $x, \dot{x}, \dot{h}, \dot{q}$ are bounded, $\dot{x}, \ddot{x}$ and $\big(\operatorname{Id} - \operatorname{prox}_{\eta f} \circ (\operatorname{Id} - \eta \nabla g)\big)x \in L^2([0,+\infty);\mathcal{H})$ and $\lim_{t\to+\infty} \dot{x}(t) = 0$. Since $\frac{d}{dt}\big(\operatorname{Id} - \operatorname{prox}_{\eta f} \circ (\operatorname{Id} - \eta \nabla g)\big)x \in L^2([0,+\infty);\mathcal{H})$ (see Remark 1(b)), we derive from Lemma 4 that $\lim_{t\to+\infty}\big(\operatorname{Id} - \operatorname{prox}_{\eta f} \circ (\operatorname{Id} - \eta \nabla g)\big)(x(t)) = 0$. As $\ddot{x}(t) = -\Gamma(\dot{x}(t)) - \lambda(t)\big(\operatorname{Id} - \operatorname{prox}_{\eta f} \circ (\operatorname{Id} - \eta \nabla g)\big)(x(t))$ for every $t \in [0,+\infty)$, we obtain that $\lim_{t\to+\infty} \ddot{x}(t) = 0$. From (46) it also follows that $\nabla g(x(\cdot)) - \nabla g(x^*) \in L^2([0,+\infty);\mathcal{H})$ and, by applying again Lemma 4, it yield $\lim_{t\to+\infty} \nabla g(x(t)) = \nabla g(x^*)$. In this way the statements (i), (ii) and (iv) are shown.

(iii) Since the function in (47) is monotonically decreasing, from (i), (ii) and (iv) it follows that the limit $\lim_{t\to+\infty}\left(\gamma\left(\frac{1}{\eta}h + u\right)(t) + \frac{1}{\eta}p(t)\right)$ exists and it is a real number. By using similar arguments as at the beginning of the proof of statement (iii) of Theorem 6, by exploiting again (46) one gets that $\exists \lim_{t\to+\infty} p(t) \in \mathbb{R}$, hence $\exists \lim_{t\to+\infty}\left(\frac{1}{\eta}h + u\right)(t) \in \mathbb{R}$.

Since x^* has been chosen as an arbitrary minimizer of $f + g$, we conclude that for all $x^* \in \operatorname{argmin}_{x\in\mathcal{H}}\{f(x) + g(x)\}$ the limit

$$\lim_{t\to+\infty} E(t, x^*) \in \mathbb{R},$$

exists, where

$$E(t, x^*) = \frac{1}{2\eta}\|x(t) - x^*\|^2 + g(x(t)) - g(x^*) - \langle \nabla g(x^*), x(t) - x^* \rangle .$$

In what follows we use a similar technique as in [14] (see, also, [1, Section 4.2]). Since $x(\cdot)$ is bounded, it has at least one weak sequential cluster point.

We prove first that each weak sequential cluster point of $x(\cdot)$ is a minimizer of $f + g$. Let $x^* \in \operatorname{argmin}_{x\in\mathcal{H}}\{f(x) + g(x)\}$ and $t_n \to +\infty$ (as $n \to +\infty$) be such that $(x(t_n))_{n\in\mathbb{N}}$ converges weakly to $\overline{x}$. Since $(x(t_n), \nabla g(x(t_n))) \in \operatorname{Gr}(\nabla g)$, $\lim_{n\to+\infty} \nabla g(x(t_n)) = \nabla g(x^*)$ and $\operatorname{Gr}(\nabla g)$ is sequentially closed in the weak-strong topology, we obtain $\nabla g(\overline{x}) = \nabla g(x^*)$.

From (34) written for $t = t_n$, $A = \partial f$ and $B = \nabla g$, by letting n converge to $+\infty$ and by using that $\operatorname{Gr}(\partial f)$ is sequentially closed in the weak-strong topology, we obtain $-\nabla g(x^*) \in \partial f(\overline{x})$. This, combined with $\nabla g(\overline{x}) = \nabla g(x^*)$ delivers $-\nabla g(\overline{x}) \in \partial f(\overline{x})$, hence $\overline{x} \in \operatorname{zer}(\partial f + \nabla g) = \operatorname{argmin}_{x\in\mathcal{H}}\{f(x) + g(x)\}$.

Next we show that $x(\cdot)$ has at most one weak sequential cluster point, which will actually guarantee that it has exactly one weak sequential cluster point. This will imply the weak convergence of the trajectory to a minimizer of $f + g$.

Let x_1^*, x_2^* be two weak sequential cluster points of $x(\cdot)$. This means that there exist $t_n \to +\infty$ (as $n \to +\infty$) and $t_n' \to +\infty$ (as $n \to +\infty$) such that $(x(t_n))_{n\in\mathbb{N}}$ converges weakly to x_1^* (as $n \to +\infty$) and $(x(t_n'))_{n\in\mathbb{N}}$ converges weakly to x_2^* (as $n \to +\infty$). Since $x_1^*, x_2^* \in \operatorname{argmin}_{x\in\mathcal{H}}\{f(x) + g(x)\}$, we have $\lim_{t\to+\infty} E(t, x_1^*) \in \mathbb{R}$ and $\lim_{t\to+\infty} E(t, x_2^*) \in \mathbb{R}$, hence $\exists \lim_{t\to+\infty}(E(t, x_1^*) - E(t, x_2^*)) \in \mathbb{R}$. We obtain

$$\exists \lim_{t\to+\infty}\left(\frac{1}{\eta}\langle x(t), x_2^* - x_1^* \rangle + \langle \nabla g(x_2^*) - \nabla g(x_1^*), x(t) \rangle\right) \in \mathbb{R},$$

which, when expressed by means of the sequences $(t_n)_{n\in\mathbb{N}}$ and $(t_n')_{n\in\mathbb{N}}$, leads to

$$\frac{1}{\eta}\langle x_1^*, x_2^* - x_1^* \rangle + \langle \nabla g(x_2^*) - \nabla g(x_1^*), x_1^* \rangle = \frac{1}{\eta}\langle x_2^*, x_2^* - x_1^* \rangle + \langle \nabla g(x_2^*) - \nabla g(x_1^*), x_2^* \rangle .$$

This is the same with

$$\frac{1}{\eta}\|x_1^* - x_2^*\|^2 + \langle \nabla g(x_2^*) - \nabla g(x_1^*), x_2^* - x_1^* \rangle = 0$$

and by the monotonicity of ∇g we conclude that $x_1^* = x_2^*$.

(v) The proof of this statement follows in analogy to the one of the corresponding statement of Theorem 10(v) written for $A = \partial f$ and $B = \nabla g$. ■

Remark 15 When $\Gamma = \gamma \operatorname{Id}$ for $\gamma > 0$, in order to verify the left-hand side of the second statement in assumption (A6) one can take $\theta' := \beta \inf_{t \geq 0} \lambda(t)$. Thus, (41) amounts in this case to the existence of $\theta > 0$ such that

$$\lambda(t) \leq \frac{\gamma^2}{\eta\theta + \frac{\eta}{\beta} + 1}.$$

When one takes $\lambda(t) = 1$ for every $t \in [0, +\infty)$, this is verified if and only if $\gamma^2 > \frac{\eta}{\beta} + 1$. In other words, (A6) allows in this particular setting a more relaxed choice for the parameters γ, η and β, beyond the standard assumptions $0 < \eta \leq 2\beta$ and $\gamma^2 > 2$ considered in [8].

In the following we provide a rate for the convergence of a convex and (Fréchet) differentiable function with Lipschitz continuous gradient $g : \mathcal{H} \to \mathbb{R}$ along the ergodic trajectory generated by

$$\begin{cases} \ddot{x}(t) + \Gamma(\dot{x}(t)) + \lambda(t)\nabla g(x(t)) = 0 \\ x(0) = u_0, \dot{x}(0) = v_0 \end{cases} \tag{48}$$

to the minimum value of g. To this end we make the following assumption

(A7) $\lambda : [0, +\infty) \to (0, +\infty)$ is locally absolutely continuous and there exists $\zeta > 0$ such that for almost every $t \in [0, +\infty)$ we have

$$0 < \zeta \leq \gamma\lambda(t) - \dot{\lambda}(t). \tag{49}$$

Let us mention that the following result is in the spirit of a convergence rate given for the objective function values on a sequence iteratively generated by an inertial-type algorithm recently obtained in [19, Theorem 1].

Theorem 16 *Let $g : \mathcal{H} \to \mathbb{R}$ be a convex and (Fréchet) differentiable function with $1/\beta$-Lipschitz continuous gradient for $\beta > 0$ such that* $\operatorname{argmin}_{x \in \mathcal{H}} g(x) \neq \emptyset$. *Let $\Gamma : \mathcal{H} \to \mathcal{H}$ be an operator fulfilling* (A1), *$\lambda : [0, +\infty) \to (0, +\infty)$ a function fulfilling* (A7), *$u_0, v_0 \in \mathcal{H}$ and $x : [0, +\infty) \to \mathcal{H}$ be the unique strong global solution of* (48).

Then for every minimizer x^* of g and every $T > 0$ it holds

$$0 \leq g\left(\frac{1}{T}\int_0^T x(t)dt\right) - g(x^*) \leq$$
$$\frac{1}{2\zeta T}\left[\|v_0 + \gamma(u_0 - x^*)\|^2 + \left(\gamma\|\Gamma - \gamma \operatorname{Id}\| + \frac{\lambda(0)}{\beta}\right)\|u_0 - x^*\|^2\right].$$

Proof. The existence and uniqueness of the trajectory of (48) follow from Theorem 2. Let be $x^* \in \operatorname{argmin}_{x\in\mathcal{H}} g(x)$, $T > 0$ and consider again the function $p : [0, +\infty) \to \mathbb{R}, p(t) = \frac{1}{2}\langle(\Gamma - \gamma\,\mathrm{Id})(x(t) - x^*), x(t) - x^*\rangle$ which we defined in (12). By using (48), the formula for the derivative of p, the positive semidefinitness of $\Gamma - \gamma\,\mathrm{Id}$, the convexity of g and (A7) we get for almost every $t \in [0, +\infty)$

$$\begin{aligned}
&\frac{d}{dt}\left(\frac{1}{2}\|\dot{x}(t) + \gamma(x(t) - x^*)\|^2 + \gamma p(t) + \lambda(t)g(x(t))\right)\\
=&\langle \ddot{x}(t) + \gamma\dot{x}(t), \dot{x}(t) + \gamma(x(t) - x^*)\rangle + \gamma\langle(\Gamma - \gamma\,\mathrm{Id})(\dot{x}(t)), x(t) - x^*\rangle\\
&+ \dot{\lambda}(t)g(x(t)) + \lambda(t)\langle\dot{x}(t), \nabla g(x(t))\rangle\\
=&\langle -(\Gamma - \gamma\,\mathrm{Id})(\dot{x}(t)) - \lambda(t)\nabla g(x(t)), \dot{x}(t) + \gamma(x(t) - x^*)\rangle\\
&+ \langle(\Gamma - \gamma\,\mathrm{Id})(\dot{x}(t)), \gamma(x(t) - x^*)\rangle + \dot{\lambda}(t)g(x(t)) + \lambda(t)\langle\dot{x}(t), \nabla g(x(t))\rangle\\
\le& -\gamma\lambda(t)\langle\nabla g(x(t)), x(t) - x^*\rangle + \dot{\lambda}(t)g(x(t))\\
\le&(\dot{\lambda}(t) - \gamma\lambda(t))(g(x(t)) - g(x^*)) + \dot{\lambda}(t)g(x^*)\\
\le& -\zeta(g(x(t)) - g(x^*)) + \dot{\lambda}(t)g(x^*).
\end{aligned}$$

We obtain after integration

$$\begin{aligned}
&\frac{1}{2}\|\dot{x}(T) + \gamma(x(T) - x^*)\|^2 + \gamma p(T) + \lambda(T)g(x(T))\\
&-\left(\frac{1}{2}\|\dot{x}(0) + \gamma(x(0) - x^*)\|^2 + \gamma p(0) + \lambda(0)g(x(0))\right)\\
&+\zeta\int_0^T (g(x(t)) - g(x^*))dt \le (\lambda(T) - \lambda(0))g(x^*).
\end{aligned}$$

Be neglecting the nonnegative terms on the left-hand side of this inequality and by using that $g(x(T)) \ge g(x^*)$, it yields

$$\zeta\int_0^T (g(x(t)) - g(x^*))dt \le \frac{1}{2}\|v_0 + \gamma(u_0 - x^*)\|^2 + \gamma p(0) + \lambda(0)(g(u_0) - g(x^*)).$$

The conclusion follows by using

$$p(0) = \frac{1}{2}\langle(\Gamma - \gamma\,\mathrm{Id})(u_0 - x^*), u_0 - x^*\rangle \le \frac{1}{2}\|\Gamma - \gamma\,\mathrm{Id}\,\|\|u_0 - x^*\|^2,$$

$$g(u_0) - g(x^*) \le \frac{1}{2\beta}\|u_0 - x^*\|^2,$$

which is a consequence of the descent lemma (see [22, Lemma 1.2.3] and notice that $\nabla g(x^*) = 0$), and the inequality

$$g\left(\frac{1}{T}\int_0^T x(t)dt\right) - g(x^*) \le \frac{1}{T}\int_0^T (g(x(t)) - g(x^*))dt,$$

which holds since g is convex. ■

Remark 17 Under assumption (A7) on the relaxation function λ, we obtain in the above theorem (only) the convergence of the function g along the ergodic trajectory to a global minimum value. If one is interested also in the (weak) convergence of the trajectory to a minimizer of g, this follows via Theorem 6 when λ is assumed to fulfill (A2) (notice that if x converges weakly to a minimizer of g, then from the Cesaro-Stolz Theorem one also obtains the weak convergence of the ergodic trajectory $T \mapsto \frac{1}{T}\int_0^T x(t)dt$ to the same minimizer).

Take $a \geq 0$, $b > 1/(\beta\gamma^2)$ and $0 \leq \rho \leq \gamma$. Then

$$\lambda(t) = \frac{1}{ae^{-\rho t} + b}$$

is an example of a relaxation function which verifies assumption (A2) (with $0 < \theta \leq b\beta\gamma^2 - 1$) and assumption (A7) (with $0 < \zeta \leq \gamma b/(a+b)^2$).

5 Variable damping parameters

In this section we carry out a similar analysis as in the previous section, however, for second order dynamical systems having as damping coefficient a function depending on time.

As starting point for our investigations we consider the dynamical system

$$\begin{cases} \ddot{x}(t) + \gamma(t)\dot{x}(t) + \lambda(t)B(x(t)) = 0 \\ x(0) = u_0, \dot{x}(0) = v_0, \end{cases} \tag{50}$$

where $B : \mathcal{H} \to \mathcal{H}$ is a cocoercive operator, $\lambda, \gamma : [0, +\infty) \to [0, +\infty)$ are Lebesgue measurable functions and $u_0, v_0 \in \mathcal{H}$.

The existence and uniqueness of strong global solutions of (50) can be shown by using the same techniques as in the proof of Theorem 2, provided that $\lambda(\cdot), \gamma(\cdot) \in L^1_{\text{loc}}([0, +\infty))$. For the convergence of the trajectories we need the following assumption

(A2') $\lambda, \gamma : [0, +\infty) \to (0, +\infty)$ are locally absolutely continuous and there exists $\theta > 0$ such that for almost every $t \in [0, +\infty)$ we have

$$\dot{\gamma}(t) \leq 0 \leq \dot{\lambda}(t) \text{ and } \frac{\gamma^2(t)}{\lambda(t)} \geq \frac{1+\theta}{\beta}. \tag{51}$$

According to Definition 1 and Remark 1(a), $\dot{\lambda}(t), \dot{\gamma}(t)$ exist for almost almost every $t \in [0, +\infty)$ and $\dot{\lambda}, \dot{\gamma}$ are Lebesgue integrable on each interval $[0, b]$ for $0 < b < +\infty$. This combined with $\dot{\gamma}(t) \leq 0 \leq \dot{\lambda}(t)$, yields the existence of a positive lower bound for λ and for a positive upper bound for γ. Using further the second assumption in (51) provides also a positive upper bound for λ and a positive lower bound for γ. The couple of functions

$$\lambda(t) = \frac{1}{ae^{-\rho t} + b} \text{ and } \gamma(t) = a'e^{-\rho' t} + b',$$

where $a, a', \rho, \rho' \geq 0$ and $b, b' > 0$ fulfill the inequality $b'^2 b > 1/\beta$, verify the conditions in assumption (A2').

We state now the convergence result.

Theorem 18 *Let $B : \mathcal{H} \to \mathcal{H}$ be a β-cocoercive operator for $\beta > 0$ such that* $\operatorname{zer} B := \{u \in \mathcal{H} : Bu = 0\} \neq \emptyset$, *$\lambda, \gamma : [0, +\infty) \to (0, +\infty)$ be functions fulfilling* (A2') *and $u_0, v_0 \in \mathcal{H}$. Let $x : [0, +\infty) \to \mathcal{H}$ be the unique strong global solution of* (50)*. Then the following statements are true:*

(i) the trajectory x is bounded and $\dot{x}, \ddot{x}, Bx \in L^2([0, +\infty); \mathcal{H})$;

(ii) $\lim_{t\to+\infty} \dot{x}(t) = \lim_{t\to+\infty} \ddot{x}(t) = \lim_{t\to+\infty} B(x(t) = 0$*;*

(iii) $x(t)$ converges weakly to an element in $\operatorname{zer} B$ *as $t \to +\infty$.*

Proof. With the notations in the proof of Theorem 6 and by appealing to similar arguments one obtains for almost every $t \in [0, +\infty)$

$$\ddot{h}(t) + \gamma(t)\dot{h}(t) + \frac{\beta}{\lambda(t)}\|\ddot{x}(t) + \gamma(t)\dot{x}(t)\|^2 \leq \|\dot{x}(t)\|^2$$

or, equivalently,

$$\ddot{h}(t) + \gamma(t)\dot{h}(t) + \frac{\beta\gamma(t)}{\lambda(t)}\frac{d}{dt}\left(\|\dot{x}(t)\|^2\right) + \left(\frac{\beta\gamma^2(t)}{\lambda(t)} - 1\right)||\dot{x}(t)||^2 + \frac{\beta}{\lambda(t)}||\ddot{x}(t)||^2 \leq 0.$$

Combining this inequality with

$$\frac{\gamma(t)}{\lambda(t)}\frac{d}{dt}\left(\|\dot{x}(t)\|^2\right) = \frac{d}{dt}\left(\frac{\gamma(t)}{\lambda(t)}\|\dot{x}(t)\|^2\right) - \frac{\dot{\gamma}(t)\lambda(t) - \gamma(t)\dot{\lambda}(t)}{\lambda^2(t)}\|\dot{x}(t)\|^2$$

and

$$\gamma(t)\dot{h}(t) = \frac{d}{dt}(\gamma h)(t) - \dot{\gamma}(t)h(t) \geq \frac{d}{dt}(\gamma h)(t), \tag{52}$$

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

$$\ddot{h}(t) + \frac{d}{dt}(\gamma h)(t) +$$
$$\beta\frac{d}{dt}\left(\frac{\gamma(t)}{\lambda(t)}\|\dot{x}(t)\|^2\right) + \left(\frac{\beta\gamma^2(t)}{\lambda(t)} + \beta\frac{-\dot{\gamma}(t)\lambda(t) + \gamma(t)\dot{\lambda}(t)}{\lambda^2(t)} - 1\right)||\dot{x}(t)||^2 + \frac{\beta}{\lambda(t)}||\ddot{x}(t)||^2 \leq 0.$$

Now, assumption (A2') delivers for almost every $t \in [0, +\infty)$ the inequality

$$\ddot{h}(t) + \frac{d}{dt}(\gamma h)(t) + \beta\frac{d}{dt}\left(\frac{\gamma(t)}{\lambda(t)}\|\dot{x}(t)\|^2\right) + \theta||\dot{x}(t)||^2 + \frac{\beta}{\lambda(t)}||\ddot{x}(t)||^2 \leq 0.$$

This implies that the function $t \mapsto \dot{h}(t) + \gamma(t)h(t) + \beta\frac{\gamma(t)}{\lambda(t)}\|\dot{x}(t)\|^2$ is monotonically decreasing and from here one obtains the conclusion following the lines of the proof of Theorem 6, by taking also into account that $\exists \lim_{t\to+\infty} \gamma(t) \in (0, \infty)$. ■

When $T : \mathcal{H} \to \mathcal{H}$ is a nonexpansive operator one obtains for the dynamical system

$$\begin{cases} \ddot{x}(t) + \gamma(t)\dot{x}(t) + \lambda(t)\big(x(t) - T(x(t))\big) = 0 \\ x(0) = u_0, \dot{x}(0) = v_0 \end{cases} \tag{53}$$

and by making the assumption

(A3') $\lambda, \gamma : [0, +\infty) \to (0, +\infty)$ are locally absolutely continuous and there exists $\theta > 0$ such that for almost every $t \in [0, +\infty)$ we have

$$\dot{\gamma}(t) \le 0 \le \dot{\lambda}(t) \text{ and } \frac{\gamma^2(t)}{\lambda(t)} \ge 2(1+\theta) \tag{54}$$

the following result which can been seen as a counterpart to Corollary 7.

Corollary 19 *Let $T : \mathcal{H} \to \mathcal{H}$ be a nonexpansive operator such that $\operatorname{Fix} T = \{u \in \mathcal{H} : Tu = u\} \neq \emptyset$, $\lambda, \gamma : [0, +\infty) \to (0, +\infty)$ be functions fulfilling (A3') and $u_0, v_0 \in \mathcal{H}$. Let $x : [0, +\infty) \to \mathcal{H}$ be the unique strong global solution of (53). Then the following statements are true:*

(i) the trajectory x is bounded and $\dot{x}, \ddot{x}, (\operatorname{Id} - T)x \in L^2([0, +\infty); \mathcal{H})$;
(ii) $\lim_{t \to +\infty} \dot{x}(t) = \lim_{t \to +\infty} \ddot{x}(t) = \lim_{t \to +\infty} (\operatorname{Id} - T)(x(t)) = 0$;
(iii) $x(t)$ converges weakly to a point in $\operatorname{Fix} T$ as $t \to +\infty$.

When $R : \mathcal{H} \to \mathcal{H}$ is an α-averaged operator for $\alpha \in (0, 1)$ one obtains for the dynamical system

$$\begin{cases} \ddot{x}(t) + \gamma(t)\dot{x}(t) + \lambda(t)\big(x(t) - R(x(t))\big) = 0 \\ x(0) = u_0, \dot{x}(0) = v_0, \end{cases} \tag{55}$$

and by making the assumption

(A4') $\lambda, \gamma : [0, +\infty) \to (0, +\infty)$ are locally absolutely continuous and there exists $\theta > 0$ such that for almost every $t \in [0, +\infty)$ we have

$$\dot{\gamma}(t) \le 0 \le \dot{\lambda}(t) \text{ and } \frac{\gamma^2(t)}{\lambda(t)} \ge 2\alpha(1+\theta) \tag{56}$$

the following result which can been seen as a counterpart to Corollary 9.

Corollary 20 *Let $R : \mathcal{H} \to \mathcal{H}$ be an α-averaged operator for $\alpha \in (0, 1)$ such that $\operatorname{Fix} R \neq \emptyset$, $\lambda, \gamma : [0, +\infty) \to (0, +\infty)$ be functions fulfilling (A4') and $u_0, v_0 \in \mathcal{H}$. Let $x : [0, +\infty) \to \mathcal{H}$ be the unique strong global solution of (55). Then the following statements are true:*

(i) the trajectory x is bounded and $\dot{x}, \ddot{x}, (\operatorname{Id} - R)x \in L^2([0, +\infty); \mathcal{H})$;
(ii) $\lim_{t \to +\infty} \dot{x}(t) = \lim_{t \to +\infty} \ddot{x}(t) = \lim_{t \to +\infty} (\operatorname{Id} - R)(x(t)) = 0$;
(iii) $x(t)$ converges weakly to a point in $\operatorname{Fix} R$ as $t \to +\infty$.

We come now to the monotone inclusion problem

$$\text{find } 0 \in A(x) + B(x),$$

where $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is a maximally monotone operator and $B : \mathcal{H} \to \mathcal{H}$ is a β-cocoercive operator for $\beta > 0$ and assign to it the second order dynamical system

$$\begin{cases} \ddot{x}(t) + \gamma(t)\dot{x}(t) + \lambda(t)\left[x(t) - J_{\eta A}\Big(x(t) - \eta B(x(t))\Big)\right] = 0 \\ x(0) = u_0, \dot{x}(0) = v_0. \end{cases} \tag{57}$$

and make the assumption

(A5') $\lambda, \gamma : [0, +\infty) \to (0, +\infty)$ are locally absolutely continuous and there exists $\theta > 0$ such that for almost every $t \in [0, +\infty)$ we have

$$\dot{\gamma}(t) \leq 0 \leq \dot{\lambda}(t) \text{ and } \frac{\gamma^2(t)}{\lambda(t)} \geq \frac{2(1+\theta)}{\delta}. \tag{58}$$

Theorem 21 *Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be a maximally monotone operator and $B : \mathcal{H} \to \mathcal{H}$ be β-cocoercive operator for $\beta > 0$ such that $\operatorname{zer}(A + B) \neq \emptyset$. Let $\eta \in (0, 2\beta)$ and set $\delta := \min\{1, \beta/\eta\} + 1/2$. Let $\lambda, \gamma : [0, +\infty) \to (0, +\infty)$ be functions fulfilling* (A5'), *$u_0, v_0 \in \mathcal{H}$ and $x : [0, +\infty) \to \mathcal{H}$ be the unique strong global solution of* (57). *Then the following statements are true:*

(i) the trajectory x is bounded and $\dot{x}, \ddot{x}, \big(\mathrm{Id} - J_{\eta A} \circ (\mathrm{Id} - \eta B)\big)x \in L^2([0, +\infty); \mathcal{H})$;

(ii) $\lim_{t \to +\infty} \dot{x}(t) = \lim_{t \to +\infty} \ddot{x}(t) = \lim_{t \to +\infty} \big(\mathrm{Id} - J_{\eta A} \circ (\mathrm{Id} - \eta B)\big)(x(t)) = 0$;

(iii) $x(t)$ converges weakly to a point in $\operatorname{zer}(A + B)$ as $t \to +\infty$;

(iv) if $x^* \in \operatorname{zer}(A+B)$, then $B(x(\cdot)) - Bx^* \in L^2([0, +\infty); \mathcal{H})$, $\lim_{t \to +\infty} B(x(t)) = Bx^*$ and B is constant on $\operatorname{zer}(A + B)$;

(v) if A or B is uniformly monotone, then $x(t)$ converges strongly to the unique point in $\operatorname{zer}(A + B)$ as $t \to +\infty$.

Proof. The statements (i)-(iii) follow by using the same arguments as in the proof of Theorem 10.

(iv) We use again the notations in the proof of Theorem 6. Let be an arbitrary $x^* \in \operatorname{zer}(A+B)$. From the definition of the resolvent we have for almost every $t \in [0, +\infty)$

$$-B(x(t)) - \frac{1}{\eta\lambda(t)}\ddot{x}(t) - \frac{\gamma(t)}{\eta\lambda(t)}\dot{x}(t) \in A\left(\frac{1}{\lambda(t)}\ddot{x}(t) + \frac{\gamma(t)}{\lambda(t)}\dot{x}(t) + x(t)\right), \tag{59}$$

which combined with $-Bx^* \in Ax^*$ and the monotonicity of A leads to

$$0 \leq \left\langle \frac{1}{\lambda(t)}\ddot{x}(t) + \frac{\gamma(t)}{\lambda(t)}\dot{x}(t) + x(t) - x^*, -B(x(t)) + Bx^* - \frac{1}{\eta\lambda(t)}\ddot{x}(t) - \frac{\gamma(t)}{\eta\lambda(t)}\dot{x}(t)\right\rangle. \tag{60}$$

The cocoercivity of B yields for almost every $t \in [0, +\infty)$

$$\begin{aligned}
\beta\|B(x(t)) - Bx^*\|^2 \leq & \left\langle \frac{1}{\lambda(t)}\ddot{x}(t) + \frac{\gamma(t)}{\lambda(t)}\dot{x}(t), -B(x(t)) + Bx^*\right\rangle - \frac{1}{\eta\lambda^2(t)}\|\ddot{x}(t) + \gamma(t)\dot{x}(t)\|^2 \\
& + \left\langle x(t) - x^*, -\frac{1}{\eta\lambda(t)}\ddot{x}(t) - \frac{\gamma(t)}{\eta\lambda(t)}\dot{x}(t)\right\rangle \\
\leq & \frac{1}{2\beta}\left\|\frac{1}{\lambda(t)}\ddot{x}(t) + \frac{\gamma(t)}{\lambda(t)}\dot{x}(t)\right\|^2 + \frac{\beta}{2}\|B(x(t)) - Bx^*\|^2 \\
& + \left\langle x(t) - x^*, -\frac{1}{\eta\lambda(t)}\ddot{x}(t) - \frac{\gamma(t)}{\eta\lambda(t)}\dot{x}(t)\right\rangle.
\end{aligned}$$

From

$$\left\langle x(t) - z, -\frac{1}{\eta\lambda(t)}\ddot{x}(t) - \frac{\gamma(t)}{\eta\lambda(t)}\dot{x}(t)\right\rangle = -\frac{1}{\eta\lambda(t)}\left(\ddot{h}(t) + \gamma(t)\dot{h}(t) - \|\dot{x}(t)\|^2\right) \tag{61}$$

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

$$\frac{\beta\lambda(t)}{2}\|B(x(t)) - Bz\|^2 + \frac{1}{\eta}\left(\ddot{h}(t) + \gamma(t)\dot{h}(t)\right) \leq \frac{1}{2\beta\lambda(t)}\|\ddot{x}(t) + \gamma(t)(\dot{x}(t))\|^2 + \frac{1}{\eta}\|\dot{x}(t)\|^2.$$

The conclusion follows in analogy to the proof of (iv) in Theorem 10 by using also (52).

(v) Let x^* be the unique element of $\operatorname{zer}(A+B)$. When A is uniformly monotone with corresponding function $\phi_A : [0,+\infty) \to [0,+\infty]$, which is increasing and vanishes only at 0, similarly to the proof of statement (v) in Theorem 10 the following inequality can be derived for almost every $t \in [0,+\infty)$

$$\phi_A\left(\left\|\frac{1}{\lambda(t)}\ddot{x}(t)+\frac{\gamma(t)}{\lambda(t)}\dot{x}(t)+x(t)-z\right\|\right) \leq$$
$$\left\langle \frac{1}{\lambda(t)}\ddot{x}(t)+\frac{\gamma(t)}{\lambda(t)}\dot{x}(t), -B(x(t))+Bz\right\rangle + \left\langle x(t)-z, -\frac{1}{\eta\lambda(t)}\ddot{x}(t)-\frac{\gamma(t)}{\eta\lambda(t)}\dot{x}(t)\right\rangle.$$

This yields $\lim_{t\to+\infty}\phi_A\left(\left\|\frac{1}{\lambda(t)}\ddot{x}(t)+\frac{\gamma(t)}{\lambda(t)}\dot{x}(t)+x(t)-z\right\|\right)=0$ and from here the conclusion is immediate.

The case when B is uniformly monotone is to be addressed in analogy to corresponding part of the proof of Theorem 10 (v). ■

Remark 22 In the light of the arguments provided in Remark 11, one can see that the statements in Theorem 21 remain valid also for $\eta = 2\beta$.

When particularizing this setting to the solving of the optimization problem

$$\min_{x\in\mathcal{H}} f(x)+g(x),$$

where $f : \mathcal{H} \to \mathbb{R}\cup\{+\infty\}$ is a proper, convex and lower semicontinuous function and $g : \mathcal{H} \to \mathbb{R}$ is a convex and (Fréchet) differentiable function with $1/\beta$-Lipschitz continuous gradient for $\beta > 0$, via the second order dynamical system

$$\begin{cases} \ddot{x}(t)+\gamma(t)\dot{x}(t)+\lambda(t)\left[x(t)-\operatorname{prox}_{\eta f}\left(x(t)-\eta\nabla g(x(t))\right)\right]=0 \\ x(0)=u_0, \dot{x}(0)=v_0, \end{cases} \tag{62}$$

Corollary 21 gives rise to the following result.

Corollary 23 *Let $f : \mathcal{H} \to \mathbb{R}\cup\{+\infty\}$ by a proper, convex and lower semicontinuous function and $g : \mathcal{H} \to \mathbb{R}$ be a convex and (Fréchet) differentiable function with $1/\beta$-Lipschitz continuous gradient for $\beta > 0$ such that* $\operatorname{argmin}_{x\in\mathcal{H}}\{f(x)+g(x)\} \neq \emptyset$. *Let $\eta \in (0, 2\beta]$ and set $\delta := \min\{1, \beta/\eta\} + 1/2$. Let $\lambda, \gamma : [0,+\infty) \to (0,+\infty)$ be functions fulfilling* (A5'), *$u_0, v_0 \in \mathcal{H}$ and $x : [0,+\infty) \to \mathcal{H}$ be the unique strong global solution of* (62). *Then the following statements are true:*

(i) the trajectory x is bounded and $\dot{x}, \ddot{x}, \left(\operatorname{Id}-\operatorname{prox}_{\eta f}\circ(\operatorname{Id}-\eta\nabla g)\right)x \in L^2([0,+\infty);\mathcal{H})$;

(ii) $\lim_{t\to+\infty}\dot{x}(t)=\lim_{t\to+\infty}\ddot{x}(t)=\lim_{t\to+\infty}\left(\operatorname{Id}-\operatorname{prox}_{\eta f}\circ(\operatorname{Id}-\eta\nabla g)\right)(x(t))=0$;

(iii) $x(t)$ converges weakly to a minimizer of $f+g$ as $t\to+\infty$;

(iv) if x^* is a minimizer of $f+g$, then $\nabla g(x(\cdot))-\nabla g(x^*) \in L^2([0,+\infty);\mathcal{H})$, $\lim_{t\to+\infty}$ $\nabla g(x(t))=\nabla g(x^*)$ and ∇g is constant on $\operatorname{argmin}_{x\in\mathcal{H}}\{f(x)+g(x)\}$*;*

(v) if f or g is uniformly convex, then $x(t)$ converges strongly to the unique minimizer of $f+g$ as $t\to+\infty$.

As it was also the case in the previous section, we can weaken the choice of the step size in Corollary 23 through the following assumption

(A6') $\lambda, \gamma : [0, +\infty) \to (0, +\infty)$ are locally absolutely continuous and there exists $\theta > 0$ such that for almost every $t \in [0, +\infty)$ we have

$$\dot{\gamma}(t) \leq 0 \leq \dot{\lambda}(t) \text{ and } \frac{\gamma^2(t)}{\lambda(t)} \geq \eta\theta + \frac{\eta}{\beta} + 1. \tag{63}$$

Corollary 24 *Let $f : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ by a proper, convex and lower semicontinuous function and $g : \mathcal{H} \to \mathbb{R}$ be a convex and (Fréchet) differentiable function with $1/\beta$-Lipschitz continuous gradient for $\beta > 0$ such that $\operatorname{argmin}_{x\in\mathcal{H}}\{f(x) + g(x)\} \neq \emptyset$. Let be $\eta > 0$, $\lambda, \gamma : [0, +\infty) \to (0, +\infty)$ be functions fulfilling (A6'), $u_0, v_0 \in \mathcal{H}$ and $x : [0, +\infty) \to \mathcal{H}$ be the unique strong global solution of (62). Then the following statements are true:*

(i) the trajectory x is bounded and $\dot{x}, \ddot{x}, \big(\operatorname{Id} - \operatorname{prox}_{\eta f} \circ(\operatorname{Id} - \eta\nabla g)\big)x \in L^2([0, +\infty); \mathcal{H})$;

(ii) $\lim_{t\to+\infty} \dot{x}(t) = \lim_{t\to+\infty} \ddot{x}(t) = \lim_{t\to+\infty} \big(\operatorname{Id} - \operatorname{prox}_{\eta f} \circ(\operatorname{Id} - \eta\nabla g)\big)(x(t)) = 0$;

(iii) $x(t)$ converges weakly to a minimizer of $f + g$ as $t \to +\infty$;

(iv) if x^* is a minimizer of $f + g$, then $\nabla g(x(\cdot)) - \nabla g(x^*) \in L^2([0, +\infty); \mathcal{H})$, $\lim_{t\to+\infty} \nabla g(x(t)) = \nabla g(x^*)$ and ∇g is constant on $\operatorname{argmin}_{x\in\mathcal{H}}\{f(x) + g(x)\}$;

(v) if f or g is uniformly convex, then $x(t)$ converges strongly to the unique minimizer of $f + g$ as $t \to +\infty$.

Proof. The proof follows in the lines of the one given for Corollary 14 and relies on the following key inequality, which holds for almost every $t \in [0, +\infty)$,

$$\beta\lambda(t)\|\nabla g(x(t)) - \nabla g(z)\|^2 + \frac{d}{dt^2}\left(\frac{1}{\eta}h + q\right) + \frac{d}{dt}\left(\gamma(t)\left(\frac{1}{\eta}h + q\right)\right) + \frac{1}{\eta}\frac{d}{dt}\left(\frac{\gamma(t)}{\lambda(t)}\|\dot{x}(t)\|^2\right)$$
$$+ \left(\frac{\gamma^2(t)}{\eta\lambda(t)} + \frac{-\dot{\gamma}(t)\lambda(t) + \gamma(t)\dot{\lambda}(t)}{\eta\lambda^2(t)} - \frac{1}{\beta} - \frac{1}{\eta}\right)\|\dot{x}(t)\|^2 + \frac{1}{\eta\lambda(t)}\|\ddot{x}(t)\|^2 \leq 0,$$

where x^* denotes a minimizer of $f + g$. This relation gives rise via (A6') to

$$\beta\lambda(t)\|\nabla g(x(t)) - \nabla g(z)\|^2 + \frac{d}{dt^2}\left(\frac{1}{\eta}h + q\right) + \frac{d}{dt}\left(\gamma(t)\left(\frac{1}{\eta}h + q\right)\right)$$
$$+ \frac{1}{\eta}\frac{d}{dt}\left(\frac{\gamma(t)}{\lambda(t)}\|\dot{x}(t)\|^2\right) + \theta\|\dot{x}(t)\|^2 + \frac{1}{\eta\lambda(t)}\|\ddot{x}(t)\|^2 \leq 0,$$

which can be seen as the counterpart to relation (46). ■

Finally, we address the convergence rate of a convex and (Fréchet) differentiable function with Lipschitz continuous gradient $g : \mathcal{H} \to \mathbb{R}$ along the ergodic trajectory generated by

$$\begin{cases} \ddot{x}(t) + \gamma(t)\dot{x}(t) + \lambda(t)\nabla g(x(t)) = 0 \\ x(0) = u_0, \dot{x}(0) = v_0 \end{cases} \tag{64}$$

to its global minimum value, when making the following assumption

(A7') $\lambda : [0, +\infty) \to (0, +\infty)$ is locally absolutely continuous, $\gamma : [0, +\infty) \to (0, +\infty)$ is twice differentiable and there exists $\zeta > 0$ such that for almost every $t \in [0, +\infty)$ we have

$$0 < \zeta \leq \gamma(t)\lambda(t) - \dot{\lambda}(t), \;\; \dot{\gamma}(t) \leq 0 \text{ and } 2\dot{\gamma}(t)\gamma(t) - \ddot{\gamma}(t) \leq 0. \tag{65}$$

Theorem 25 *Let $g : \mathcal{H} \to \mathbb{R}$ be a convex and (Fréchet) differentiable function with $1/\beta$-Lipschitz continuous gradient for $\beta > 0$ such that* $\operatorname{argmin}_{x\in\mathcal{H}} g(x) \neq \emptyset$. *Let $\lambda, \gamma : [0, +\infty) \to (0, +\infty)$ be functions fulfilling* (A7') *$u_0, v_0 \in \mathcal{H}$ and $x : [0, +\infty) \to \mathcal{H}$ be the unique strong global solution of* (64).

Then for every minimizer x^* of g and every $T > 0$ it holds

$$\begin{aligned}0 \leq & g\left(\frac{1}{T}\int_0^T x(t)dt\right) - g(x^*) \leq \\ & \frac{1}{2\zeta T}\left[\|v_0 + \gamma(0)(u_0 - x^*)\|^2 + \left(\frac{\lambda(0)}{\beta} - \dot{\gamma}(0)\right)\|u_0 - x^*\|^2\right].\end{aligned}$$

Proof. Let $x^* \in \operatorname{argmin}_{x\in\mathcal{H}} g(x)$ and $T > 0$. By using (64), the convexity of g and (A7') we get for almost every $t \in [0, +\infty)$

$$\begin{aligned}& \frac{d}{dt}\left(\frac{1}{2}\|\dot{x}(t) + \gamma(t)(x(t) - x^*)\|^2 + \lambda(t)g(x(t)) - \frac{\dot{\gamma}(t)}{2}\|x(t) - x^*\|^2\right) \\ = & \langle \ddot{x}(t) + \dot{\gamma}(t)(x(t) - x^*) + \gamma(t)\dot{x}(t), \dot{x}(t) + \gamma(t)(x(t) - x^*)\rangle \\ & - \frac{\ddot{\gamma}(t)}{2}\|x(t) - x^*\|^2 - \dot{\gamma}(t)\langle \dot{x}(t), x(t) - x^*\rangle + \dot{\lambda}(t)g(x(t)) + \lambda(t)\langle \dot{x}(t), \nabla g(x(t))\rangle \\ = & -\gamma(t)\lambda(t)\langle \nabla g(x(t)), x(t) - x^*\rangle + \dot{\lambda}(t)g(x(t)) + \left(\dot{\gamma}(t)\gamma(t) - \frac{\ddot{\gamma}(t)}{2}\right)\|x(t) - x^*\|^2 \\ \leq & -\gamma(t)\lambda(t)\langle \nabla g(x(t)), x(t) - x^*\rangle + \dot{\lambda}(t)g(x(t)) \\ \leq & (\dot{\lambda}(t) - \gamma(t)\lambda(t))(g(x(t)) - g(x^*)) + \dot{\lambda}(t)g(x^*) \\ \leq & -\zeta(g(x(t)) - g(x^*)) + \dot{\lambda}(t)g(x^*).\end{aligned}$$

We obtain after integration

$$\begin{aligned}& \frac{1}{2}\|\dot{x}(T) + \gamma(x(T) - x^*)\|^2 + \lambda(T)g(x(T)) - \frac{\ddot{\gamma}(T)}{2}\|x(T) - x^*\|^2 \\ & -\frac{1}{2}\|\dot{x}(0) + \gamma(x(0) - x^*)\|^2 + \lambda(0)g(x(0)) - \frac{\ddot{\gamma}(0)}{2}\|x(0) - x^*\|^2 \\ & +\zeta\int_0^T (g(x(t)) - g(x^*))dt \leq (\lambda(T) - \lambda(0))g(x^*).\end{aligned}$$

The conclusion follows from here as in the proof of Theorem 16. ■

Remark 26 A similar comment as in Remark 17 can be made also in this context. For $a, a', \rho, \rho' \geq 0$ and $b, b' > 0$ fulfilling the inequalities $b'^2 b > 1/\beta$ and $0 \leq \rho \leq b'$ one can prove that the functions

$$\lambda(t) = \frac{1}{ae^{-\rho t} + b} \text{ and } \gamma(t) = a'e^{-\rho' t} + b',$$

verify assumption (A2') in Theorem 18 (with $0 < \theta \leq b'^2 b\beta - 1$) and assumption (A7') in Theorem 25 (with $0 < \zeta \leq bb'/(a + b)^2$). Hence, for this choice of the relaxation and damping function, one has convergence of the objective function g along the ergodic trajectory to its global minimum value as well as (weak) convergence of the trajectory to a minimizer of g.

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, DOI: 10.1080/02331934.2014.971412, 2014

[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] F. Alvarez, *On the minimizing property of a second order dissipative system in Hilbert spaces*, SIAM Journal on Control and Optimization 38(4), 1102–1119, 2000

[4] F. Alvarez, *Weak convergence of a relaxed and inertial hybrid projection-proximal point algorithm for maximal monotone operators in Hilbert space*, SIAM Journal on Optimization 14(3), 773–782, 2004

[5] F. Alvarez, H. Attouch, *An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping*, Set-Valued Analysis 9(1-2), 311, 2001

[6] 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

[7] 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

[8] H. Attouch, F. Alvarez, *The heavy ball with friction dynamical system for convex constrained minimization problems*, in: Optimization (Namur, 1998), 25–35, in: Lecture Notes in Economics and Mathematical Systems 481, Springer, Berlin, 2000

[9] H. Attouch, M. Marques Alves, B.F. Svaiter, *A dynamic approach to a proximal-Newton method for monotone inclusions in Hilbert spaces, with complexity* $O(1/n^2)$, arXiv:1502.04286v1

[10] H. Attouch, M.-O. Czarnecki, *Asymptotic behavior of coupled dynamical systems with multiscale aspects*, Journal of Differential Equations 248(6), 1315-1344, 2010

[11] H. Attouch, X. Goudou, P. Redont, *The heavy ball with friction method. I. The continuous dynamical system: global exploration of the local minima of a real-valued function by asymptotic analysis of a dissipative dynamical system*, Communications in Contemporary Mathematics 2(1), 1–34, 2000

[12] 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

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

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

[15] J.M. Borwein and J.D. Vanderwerff, *Convex Functions: Constructions, Characterizations and Counterexamples*, Cambridge University Press, Cambridge, 2010

[16] 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, DOI: 10.1007/s10884-015-9438-x, 2015

[17] R.I. Boţ, E.R. Csetnek, *Approaching the solving of constrained variational inequalities via penalty term-based dynamical systems*, arXiv:1503.01871

[18] R.I. Boţ, E.R. Csetnek, C. Hendrich, *Inertial Douglas-Rachford splitting for monotone inclusion problems*, Applied Mathematics and Computation 256, 472–487, 2015

[19] E. Ghadimi, H.R. Feyzmahdavian, M. Johansson, *Global convergence of the Heavy-ball method for convex optimization*, arXiv:1412.7457

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

[21] A. Moudafi, M. Oliny, *Convergence of a splitting inertial proximal method for monotone operators*, Journal of Computational and Applied Mathematics 155, 447–454, 2003

[22] Y. Nesterov, *Introductory Lectures on Convex Optimization: A Basic Course*, Kluwer Academic Publishers, Dordrecht, 2004

[23] J.-C. Pesquet, N. Pustelnik, *A parallel inertial proximal optimization method*, Pacific Journal of Optimization 8(2), 273–305, 2012

[24] R.T. Rockafellar, *On the maximal monotonicity of subdifferential mappings*, Pacific Journal of Mathematics 33(1), 209–216, 1970

[25] R.T. Rockafellar, *Monotone operators and the proximal point algorithm*, SIAM Journal on Control and Optimization 14(5), 877–898, 1976

[26] S. Simons, *From Hahn-Banach to Monotonicity*, Springer, Berlin, 2008

[27] E.D. Sontag, *Mathematical control theory. Deterministic finite-dimensional systems*, Second edition, Texts in Applied Mathematics 6, Springer-Verlag, New York, 1998