Advances in Continuous and Discrete Models

RESEARCH Open Access

A fast continuous time approach with time scaling for nonsmooth convex optimization

Radu Ioan Boţ[1*] and Mikhail A. Karapetyants[1]

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

Abstract

In a Hilbert setting, we study the convergence properties of the second order in time dynamical system combining viscous and Hessian-driven damping with time scaling in relation to the minimization of a nonsmooth and convex function. The system is formulated in terms of the gradient of the Moreau envelope of the objective function with a time-dependent parameter. We show fast convergence rates for the Moreau envelope, its gradient along the trajectory, and also for the system velocity. From here, we derive fast convergence rates for the objective function along a path which is the image of the trajectory of the system through the proximal operator of the first. Moreover, we prove the weak convergence of the trajectory of the system to a global minimizer of the objective function. Finally, we provide multiple numerical examples illustrating the theoretical results.

MSC: 37N40; 46N10; 49M99; 65K05; 65K10; 90C25

Keywords: Nonsmooth convex optimization; Damped inertial dynamics; Hessian-driven damping; Time scaling; Moreau envelope; Proximal operator

1 Introduction

Let H be a real Hilbert space endowed with the scalar product $\langle \cdot, \cdot \rangle$ and norm $\|x\| = \sqrt{\langle x, x \rangle}$ for $x \in H$. In connection with the minimization problem

$$\min_{x \in H} \Phi(x),$$

we will study the asymptotic behavior of the second order in time evolution equation

$$\ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \beta(t)\frac{d}{dt}\nabla\Phi_{\lambda(t)}\big(x(t)\big) + b(t)\nabla\Phi_{\lambda(t)}\big(x(t)\big) = 0, \tag{1}$$

with initial conditions $x(t_0) = x_0 \in H$, $\dot{x}(t_0) = u_0 \in H$, where $\alpha \geq 1$, $t_0 > 0$, and $\beta : [t_0, +\infty) \longrightarrow [0, +\infty)$ and $b, \lambda : [t_0, +\infty) \longrightarrow (0, +\infty)$ are differentiable functions.

We assume that $\Phi : H \longrightarrow \overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ is a proper, convex and lower semicontinuous function and denote by $\Phi_\lambda : H \longrightarrow \mathbb{R}$ its Moreau envelope of parameter $\lambda > 0$. In addition, we assume that argmin Φ, the set of global minimizers of Φ, is not empty and denote by Φ^* the optimal objective value of Φ.

Our aim is to derive rates of convergence for the Moreau envelope of the objective function and the objective function itself to Φ^*, as well as for the gradient of the Moreau envelope of the objective function and the velocity of the trajectory to zero in terms of the Moreau parameter function λ and the time scaling function b. In addition, we will provide a setting that also guarantees the weak convergence of the trajectory of the dynamical system to a minimizer of Φ. The theoretical results will be illustrated by multiple numerical experiments.

1.1 Historical remarks

Inertial dynamics were introduced by Polyak in [23] in the form of the so-called heavy ball with friction method

$$\ddot{x}(t) + \alpha\dot{x}(t) + \nabla\Phi\big(x(t)\big) = 0,$$

with fixed viscous coefficient $\alpha > 0$, to accelerate the gradient method for the minimization of a continuous differentiable function $\Phi : H \to \mathbb{R}$. This system was later studied by Alvarez–Attouch [3, 4] and by Attouch–Goudou–Redont [11]. In these works, for a convex function Φ, an asymptotic convergence rate of $\Phi(x(t))$ to Φ^* of order $O(\frac{1}{t})$ as $t \to +\infty$, as well as an improvement for a strongly convex function Φ to an exponential rate of convergence, was proved. The weak convergence of the trajectories to a minimizer of Φ was also established.

A major step to obtain faster asymptotic convergence in the convex regime was done by Su–Boyd–Candes [24], by considering in the second order dynamical system an asymptotic vanishing damping coefficient

$$\ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \nabla\Phi\big(x(t)\big) = 0, \tag{2}$$

for $t \geq t_0$ and $\alpha \geq 3$. Second order dynamical systems with variable and vanishing damping coefficients for optimization were studied, for instance, in [17–19]. The system (2) corresponds to a continuous version of Nesterov's accelerated gradient method [21]. For the function values, rates of convergence of

$$\Phi\big(x(t)\big) - \Phi^* = O\left(\frac{1}{t^2}\right) \quad \text{as } t \to +\infty$$

were obtained. For $\alpha > 3$, in [9], it was shown that the trajectory of (2) converges weakly to an element of $\operatorname{argmin}\Phi$, and in [13, 20], the asymptotic convergence rate of the function values was improved to $o(\frac{1}{t^2})$ as $t \to +\infty$.

The following system, combining asymptotic vanishing damping with the Hessian-driven damping, was proposed by Attouch–Peypouquet–Redont in [15]

$$\ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \beta\nabla^2\Phi\big(x(t)\big)\dot{x}(t) + \nabla\Phi\big(x(t)\big) = 0 \tag{3}$$

for $t \geq t_0$, where $\Phi : H \longrightarrow \mathbb{R}$ twice continuously differentiable and convex, $\alpha \geq 3$ and $\beta \geq 0$. The Hessian-driven damping has a natural link with Newton's method and gives

rise to dynamical inertial Newton systems [1]. The system (3) preserves the convergence properties of (2), while having for $\beta > 0$ other important features, namely,

$$\lim_{t\to+\infty} \|\nabla\Phi(x(t))\| = 0 \quad \text{and} \quad \int_{t_0}^{+\infty} t^2 \|\nabla\Phi(x(t))\|^2 \, dt < +\infty.$$

In addition, possible oscillations exhibited by the solutions of (2) are neutralized by (3).

1.2 Time scaling

Time scaling of the dynamical system (2) was used to accelerate the rate of convergence of the values of the function Φ along the trajectory. The system (2) becomes through time scaling a dynamical system of the form

$$\ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + b(t)\nabla\Phi(x(t)) = 0, \tag{4}$$

where $\alpha \geq 3$, and $b : [t_0, +\infty) \longrightarrow (0, +\infty)$ is a continuous scalar function, as it was introduced and studied by Attouch–Chbani–Riahi in [10]. For (4), it was shown that

$$\Phi(x(t)) - \Phi^* = O\left(\frac{1}{t^2 b(t)}\right) \quad \text{as } t \to +\infty,$$

and a convergence rate can be improved to $o(\frac{1}{t^2 b(t)})$ as $t \to +\infty$, if $\alpha > 3$.

In [7, 8] (see also [5]), the dynamical system

$$\ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \beta(t)\nabla^2\Phi(x(t))\dot{x}(t) + b(t)\nabla\Phi(x(t)) = 0, \tag{5}$$

which combines viscous and Hessian-driven damping with time scaling, where $\alpha \geq 1$ and $\beta, b : [t_0, +\infty) \longrightarrow (0, +\infty)$ are functions with appropriate differentiability properties, was investigated. A quite general setting formulated in terms of the dynamical system parameter functions was identified in which the properties of (5) concerning the convergence of the function values are preserved, while the gradient of Φ strongly converges along the trajectory to zero, and the trajectory converges weakly to a minimizer of the objective function. In [7, 8], a numerical algorithm obtained via time discretization of (5) was studied, exhibiting analogous convergence properties to the dynamical system.

1.3 Nonsmooth optimization

The Moreau envelope of a proper, convex and lower semicontinuous function $\Phi : H \to \overline{\mathbb{R}}$ has played a significant role in the literature when designing continuous-time approaches and numerical algorithms for the minimization of Φ. This is defined as

$$\Phi_\lambda : H \to \mathbb{R}, \qquad \Phi_\lambda(x) = \inf_{y\in H}\left\{\Phi(y) + \frac{1}{2\lambda}\|x - y\|^2\right\},$$

where $\lambda > 0$ is called the parameter of the Moreau envelope (see, for instance, [16]). For every $\lambda > 0$, the functions Φ and Φ_λ share the same optimal objective value and the same set of minimizers. In addition, Φ_λ is convex and continuously differentiable with

$$\nabla\Phi_\lambda(x) = \frac{1}{\lambda}\left(x - \operatorname{prox}_{\lambda\Phi}(x)\right) \quad \forall x \in H, \tag{6}$$

and $\nabla\Phi_\lambda$ is $\frac{1}{\lambda}$-Lipschitz continuous. Here,

$$\operatorname{prox}_{\lambda\Phi} : H \to H, \qquad \operatorname{prox}_{\lambda\Phi}(x) = \operatorname*{argmin}_{y\in H}\left\{\Phi(y) + \frac{1}{2\lambda}\|x-y\|^2\right\}$$

denotes the proximal operator of Φ of parameter λ. For every $x \in H$ and $\lambda, \mu > 0$, we have

$$\left\|\operatorname{prox}_{\lambda\Phi}(x) - \operatorname{prox}_{\mu\Phi}(x)\right\| \leq |\lambda-\mu|\left\|\nabla\Phi_\lambda(x)\right\|. \tag{7}$$

On the other hand, for every $x \in H$, the function $\lambda \in (0,+\infty) \to \Phi_\lambda(x)$ is nonincreasing and differentiable (see, for instance, [6, Lemma A.1]), namely,

$$\frac{d}{d\lambda}\Phi_\lambda(x) = -\frac{1}{2}\left\|\nabla\Phi_\lambda(x)\right\|^2 \quad \forall\lambda > 0.$$

Attouch–Cabot considered in [6] (see also [14] for a more general approach for monotone inclusions) in connection with the minimization of the proper, convex and lower semicontinuous function $\Phi : H \to \overline{\mathbb{R}}$ the following second order differential equation

$$\ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \nabla\Phi_{\lambda(t)}\big(x(t)\big) = 0 \tag{8}$$

for $t \geq t_0$, where $\alpha \geq 1$, and $\lambda : [t_0,+\infty) \longrightarrow (0,+\infty)$ is continuously differentiable and nondecreasing. Convergence rates for the values of the Moreau envelope, as well as for the velocity of the system, were obtained

$$\Phi_{\lambda(t)}\big(x(t)\big) - \Phi^* = o\left(\frac{1}{t^2}\right) \quad \text{and} \quad \left\|\dot{x}(t)\right\| = o\left(\frac{1}{t}\right) \quad \text{as } t \to +\infty,$$

from where convergence rates for the Φ along $x(t)$ were deduced

$$\Phi\big(\operatorname{prox}_{\lambda(t)\Phi}\big(x(t)\big)\big) - \Phi^* = o\left(\frac{1}{t^2}\right) \quad \text{and}$$

$$\left\|\operatorname{prox}_{\lambda(t)\Phi}\big(x(t)\big) - x(t)\right\| = o\left(\frac{\sqrt{\lambda(t)}}{t}\right) \quad \text{as } t \to +\infty.$$

In addition, the weak convergence of the trajectories $x(t)$ to a minimizer of Φ as $t \to +\infty$ was established.

Attouch–László considered in [12] in the same context the dynamical system

$$\ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \beta\frac{d}{dt}\nabla\Phi_{\lambda(t)}\big(x(t)\big) + \nabla\Phi_{\lambda(t)}\big(x(t)\big) = 0, \tag{9}$$

where $\alpha > 1$ and $\beta > 0$, and the term $\frac{d}{dt}\nabla\Phi_{\lambda(t)}(x(t))$ is inspired by the Hessian-driven damping, and its existence is justified almost everywhere since the mapping $t \to \nabla\Phi_{\lambda(t)}(x(t))$ is locally absolutely continuous (see, for example, [12, Lemma 1]). It was shown that for $\lambda(t) = \lambda t^2$, where $\lambda > 0$, the system (9) inherits all major convergence properties of (8), and, in addition, the following convergence rates for the gradient of the Moreau envelope of parameter $\lambda(t)$ and its time derivative along $x(t)$ were established

$$\left\|\nabla\Phi_{\lambda(t)}\big(x(t)\big)\right\| = o\left(\frac{1}{t^2}\right) \quad \text{and} \quad \left\|\frac{d}{dt}\nabla\Phi_{\lambda(t)}\big(x(t)\big)\right\| = o\left(\frac{1}{t^2}\right) \quad \text{as } t \to +\infty.$$

1.4 Our contribution

In this paper, we derive a setting formulated in terms of $\alpha \geq 1$ and the parameter functions β, b and λ of the dynamical system (1) associated with the minimization of the proper, convex and lower semicontinuous function $\Phi : H \to \overline{\mathbb{R}}$, which allow us to prove

- convergence rates for the Moreau envelope, its gradient, and the velocity of the trajectory

$$\Phi_{\lambda(t)}\big(x(t)\big) - \Phi^* = o\left(\frac{1}{t^2 b(t)}\right), \qquad \big\|\nabla \Phi_{\lambda(t)}\big(x(t)\big)\big\| = o\left(\frac{1}{t\sqrt{b(t)\lambda(t)}}\right) \quad \text{and}$$
$$\big\|\dot{x}(t)\big\| = o\left(\frac{1}{t}\right)$$

  as $t \to +\infty$, respectively;
- convergence rates for the objective function

$$\Phi\big(\mathrm{prox}_{\lambda(t)\Phi}\big(x(t)\big)\big) - \Phi^* = o\left(\frac{1}{t^2 b(t)}\right) \quad \text{and}$$
$$\big\|\mathrm{prox}_{\lambda(t)\Phi}\big(x(t)\big) - x(t)\big\| = o\left(\frac{\sqrt{\lambda(t)}}{t\sqrt{b(t)}}\right)$$

  as $t \to +\infty$;
- the weak convergence of the trajectory $x(t)$ to a minimizer of Φ as $t \to +\infty$.

In addition, we provide a particular formulation of the derived general setting for the case when the parameter functions are chosen to be polynomials and illustrate the influence of the latter on the convergence behavior of the dynamical system by multiple numerical experiments.

1.5 Existence and uniqueness of strong global solution

This section is devoted to the topic of the existence and uniqueness of a strong global solution of the system of our interest. To this aim, we will rewrite (1) as a system of the first order in time equations in the product space $H \times H$.

First, we assume that $\beta : [t_0, +\infty) \longrightarrow [0, +\infty)$ is twice continuously differentiable with $\beta(t) > 0$ for every $t \geq t_0$. We integrate (1) from t_0 to t to obtain

$$\dot{x}(t) + \beta(t)\nabla \Phi_{\lambda(t)}\big(x(t)\big) + \int_{t_0}^{t} \left(\frac{\alpha}{s}\dot{x}(s) + b(s)\nabla \Phi_{\lambda(s)}\big(x(s)\big)\right) ds - \int_{t_0}^{t} \nabla \Phi_{\lambda(s)}\big(x(s)\big)\dot{\beta}(s)\, ds$$
$$- \big(\dot{x}(t_0) + \beta(t_0)\nabla \Phi_{\lambda(t_0)}\big(x(t_0)\big)\big) = 0.$$

We denote $z(t) := \int_{t_0}^{t} (\frac{\alpha}{s}\dot{x}(s) + (b(s) - \dot{\beta}(s))\nabla \Phi_{\lambda(s)}(x(s)))\, ds - (u_0 + \beta(t_0)\nabla \Phi_{\lambda(t_0)}(x_0)))$ for every $t \geq t_0$. Since $\dot{z}(t) = \frac{\alpha}{t}\dot{x}(t) + (b(t) - \dot{\beta}(t))\nabla \Phi_{\lambda(t)}(x(t))$, we notice that (1) is equivalent to

$$\begin{cases} \dot{x}(t) + \beta(t)\nabla \Phi_{\lambda(t)}(x(t)) + z(t) = 0, \\ \dot{z}(t) - \frac{\alpha}{t}\dot{x}(t) - (b(t) - \dot{\beta}(t))\nabla \Phi_{\lambda(t)}(x(t)) = 0, \\ x(t_0) = x_0, \qquad z(t_0) = -(u_0 + \beta(t_0)\nabla \Phi_{\lambda(t_0)}(x_0)). \end{cases}$$

After multiplying the first line by $b(t) - \dot{\beta}(t)$ and the second one by $\beta(t)$ and then summing them, we get rid of the gradient of the Moreau envelope in the second equation

$$\begin{cases} \dot{x}(t) + \beta(t)\nabla\Phi_{\lambda(t)}(x(t)) + z(t) = 0, \\ \beta(t)\dot{z}(t) + (b(t) - \dot{\beta}(t) - \frac{\alpha\beta(t)}{t})\dot{x}(t) + (b(t) - \dot{\beta}(t))z(t) = 0, \\ x(t_0) = x_0, \qquad z(t_0) = -(u_0 + \beta(t_0)\nabla\Phi_{\lambda(t_0)}(x_0)). \end{cases}$$

We denote $y(t) = \beta(t)z(t) + (b(t) - \dot{\beta}(t) - \frac{\alpha\beta(t)}{t})x(t)$, and after simplification, we obtain for the dynamical system the following equivalent formulation

$$\begin{cases} \dot{x}(t) + \beta(t)\nabla\Phi_{\lambda(t)}(x(t)) + (\frac{\dot{\beta}(t)-b(t)}{\beta(t)} + \frac{\alpha}{t})x(t) + \frac{1}{\beta(t)}y(t) = 0, \\ \dot{y}(t) + (\ddot{\beta}(t) + \frac{3b(t)\dot{\beta}(t)-2\dot{\beta}^2(t)-b^2(t)}{\beta(t)} + \frac{\alpha}{t}(b(t) - \dot{\beta}(t) - \frac{\beta(t)}{t}) - \dot{b}(t))x(t) + \frac{b(t)-2\dot{\beta}(t)}{\beta(t)}y(t) = 0, \\ x(t_0) = x_0, \qquad y(t_0) = -\beta(t_0)(u_0 + \beta(t_0)\nabla\Phi_{\lambda(t_0)}(x_0)) + (b(t_0) - \dot{\beta}(t_0) - \frac{\alpha\beta(t_0)}{t_0})x_0. \end{cases}$$

In case $\beta(t) = 0$ for every $t \geq t_0$, (1) can be equivalently written as

$$\begin{cases} \dot{x}(t) - y(t) = 0, \\ \dot{y}(t) + \frac{\alpha}{t}y(t) + b(t)\nabla\Phi_{\lambda(t)}(x(t)) = 0, \\ x(t_0) = x_0, \qquad y(t_0) = u_0. \end{cases}$$

Based on the two reformulations of the dynamical system (1), we can formulate the following existence and uniqueness result, which is a consequence of the Cauchy–Lipschitz theorem for strong global solutions. The result can be proved in the lines of the proofs of Theorem 1 in [12] or of Theorem 1.1 in [15] with some small adjustments.

Theorem 1 *Suppose that $\beta : [t_0, +\infty) \longrightarrow [0, +\infty)$ is twice continuously differentiable such that either $\beta(t) > 0$ for every $t \geq t_0$ or $\beta(t) = 0$ for every $t \geq t_0$, and there exists $\lambda_0 > 0$ such that $\lambda(t) \geq \lambda_0$ for all $t \geq t_0$. Then, for every $(x_0, u_0) \in H \times H$, there exists a unique strong global solution $x : [t_0, +\infty) \mapsto H$ of the continuous dynamics (1), satisfying the Cauchy initial conditions $x(t_0) = x_0$ and $\dot{x}(t_0) = u_0$.*

2 Energy function and rates of convergence for function values

In this section, we will define an energy function for the dynamical system (1) and investigate its dissipativity properties. These will play a crucial role in the derivation of rates of convergence for the Moreau envelope of Φ and the objective function itself.

To shorten the calculations, we introduce the auxiliary function (see also [7, 8])

$$w : [t_0, +\infty) \to \mathbb{R}, \qquad w(t) = b(t) - \dot{\beta}(t) - \frac{\beta(t)}{t}.$$

For $z \in \operatorname{argmin}\Phi$ and

$$0 \leq c \leq \alpha - 1, \tag{10}$$

consider the energy function $E_c : [t_0, +\infty) \to [0, +\infty)$,

$$\begin{aligned} E_c(t) = {} & \big(t^2 w(t) + (\alpha - 1 - c)t\beta(t)\big)\big(\Phi_{\lambda(t)}\big(x(t)\big) - \Phi^*\big) \\ & + \frac{1}{2}\big\|c\big(x(t) - z\big) + t\dot{x}(t) + t\beta(t)\nabla\Phi_{\lambda(t)}\big(x(t)\big)\big\|^2 \\ & + \frac{c(\alpha - 1 - c)}{2}\big\|x(t) - z\big\|^2. \end{aligned}$$

In the following theorem, we formulate sufficient conditions that guarantee the decay of the energy of the the dynamical system (1) and discuss some of its consequences.

Theorem 2 *Suppose that $\alpha \geq 1$, λ is nondecreasing on $[t_0, +\infty)$ and the following conditions*

$$b(t) > \dot{\beta}(t) + \frac{\beta(t)}{t} \quad \text{for every } t \geq t_0 \tag{11}$$

and

$$(\alpha - 3)w(t) - t\dot{w}(t) \geq 0 \quad \text{for every } t \geq t_0 \tag{12}$$

are satisfied. Then, for a solution $x : [t_0, +\infty) \to H$ to (1), the following statements are true:

(i) $\dot{E}_c(t) \leq 0$ *for every* $t \geq t_0$;
(ii) $\Phi_{\lambda(t)}(x(t)) - \Phi^* \leq \frac{E_{\alpha-1}(t_0)}{t^2 w(t)}$ *for every* $t \geq t_0$;
(iii) $\int_{t_0}^{+\infty} (t^2 w(t)\frac{\dot{\lambda}(t)}{2} + t^2\beta(t)w(t))\|\nabla\Phi_{\lambda(t)}(x(t))\|^2\,dt < +\infty$;
(iv) $\int_{t_0}^{+\infty} ((\alpha - 3)tw(t) - t^2\dot{w}(t))(\Phi_{\lambda(t)}(x(t)) - \Phi^*)\,dt < +\infty$.

Moreover, assuming that $\alpha > 1$ and that

$$\text{there exists } \varepsilon \in (0, \alpha - 1) \text{ such that } (\alpha - 3)w(t) - t\dot{w}(t) \geq \varepsilon b(t) \quad \forall t \geq t_0, \tag{13}$$

it holds;

(v) $\int_{t_0}^{+\infty} t\|\dot{x}(t)\|^2\,dt < +\infty$;
(vi) *the trajectory x is bounded and*
(vii) $\int_{t_0}^{+\infty} tb(t)(\Phi_{\lambda(t)}(x(t)) - \Phi^*)\,dt < +\infty$.

Proof For every $t \geq t_0$, we obtain

$$\begin{aligned} \dot{E}_c(t) = {} & \big(2tw(t) + t^2\dot{w}(t) + \beta(t)(\alpha - 1 - c) + (\alpha - 1 - c)t\dot{\beta}(t)\big)\big(\Phi_{\lambda(t)}\big(x(t)\big) - \Phi^*\big) \\ & + \big(t^2 w(t) + \beta(t)t(\alpha - 1 - c)\big)\left(\big\langle\nabla\Phi_{\lambda(t)}\big(x(t)\big), \dot{x}(t)\big\rangle - \frac{\dot{\lambda}(t)}{2}\big\|\nabla\Phi_{\lambda(t)}\big(x(t)\big)\big\|^2\right) \\ & + \Big\langle c\big(x(t) - z\big) + t\dot{x}(t) + t\beta(t)\nabla\Phi_{\lambda(t)}\big(x(t)\big), (c + 1)\dot{x}(t) + t\ddot{x}(t) \\ & + + t\beta(t)\frac{d}{dt}\big(\nabla\Phi_{\lambda(t)}\big(x(t)\big)\big)\big(\beta(t) + t\dot{\beta}(t)\big)\nabla\Phi_{\lambda(t)}\big(x(t)\big)\Big\rangle \\ & + c(\alpha - 1 - c)\big\langle x(t) - z, \dot{x}(t)\big\rangle, \end{aligned}$$

where we used that

$$\frac{d}{dt}\big(\Phi_{\lambda(t)}\big(x(t)\big)-\Phi^*\big)=\big\langle\nabla\Phi_{\lambda(t)}\big(x(t)\big),\dot{x}(t)\big\rangle-\frac{\dot{\lambda}(t)}{2}\big\|\nabla\Phi_{\lambda(t)}x(t)\big\|^2. \tag{14}$$

Using (1) to replace $\ddot{x}(t)$, we may write the third summand in the formulation of $\dot{E}_c(t)$ for every $t\geq t_0$ as

$$\begin{aligned}
&\big\langle c\big(x(t)-z\big)+t\dot{x}(t)+t\beta(t)\nabla\Phi_{\lambda(t)}\big(x(t)\big),\\
&\qquad(c+1-\alpha)\dot{x}(t)+\big(\beta(t)+t\dot{\beta}(t)-tb(t)\big)\nabla\Phi_{\lambda(t)}\big(x(t)\big)\big\rangle\\
&=c(c+1-\alpha)\big\langle x(t)-z,\dot{x}(t)\big\rangle+c\big(\beta(t)+t\dot{\beta}(t)-tb(t)\big)\big\langle x(t)-z,\nabla\Phi_{\lambda(t)}\big(x(t)\big)\big\rangle\\
&\quad+(c+1-\alpha)t\big\|\dot{x}(t)\big\|^2+\big(\beta(t)+t\dot{\beta}(t)-tb(t)\big)t\big\langle\dot{x}(t),\nabla\Phi_{\lambda(t)}\big(x(t)\big)\big\rangle\\
&\quad+t\beta(t)(c+1-\alpha)\big\langle\dot{x}(t),\nabla\Phi_{\lambda(t)}\big(x(t)\big)\big\rangle+t\beta(t)\big(\beta(t)+t\dot{\beta}(t)-tb(t)\big)\big\|\nabla\Phi_{\lambda(t)}\big(x(t)\big)\big\|^2.
\end{aligned}$$

Overall, since $\beta(t)+t\dot{\beta}(t)-tb(t)=-tw(t)$, we obtain for every $t\geq t_0$

$$\begin{aligned}
\dot{E}_c(t)&=\big(2tw(t)+t^2\dot{w}(t)-\big(\beta(t)+t\dot{\beta}(t)\big)(c+1-\alpha)\big)\big(\Phi_{\lambda(t)}\big(x(t)\big)-\Phi^*\big)\\
&\quad+\big(t^2w(t)-t\beta(t)(c+1-\alpha)\big)\left(\big\langle\nabla\Phi_{\lambda(t)}\big(x(t)\big),\dot{x}(t)\big\rangle-\frac{\dot{\lambda}(t)}{2}\big\|\nabla\Phi_{\lambda(t)}\big(x(t)\big)\big\|^2\right)\\
&\quad-ctw(t)\big\langle x(t)-z,\nabla\Phi_{\lambda(t)}\big(x(t)\big)\big\rangle+(c+1-\alpha)t\big\|\dot{x}(t)\big\|^2-t^2w(t)\big\langle\dot{x}(t),\nabla\Phi_{\lambda(t)}\big(x(t)\big)\big\rangle\\
&\quad+t\beta(t)(c+1-\alpha)\big\langle\dot{x}(t),\nabla\Phi_{\lambda(t)}\big(x(t)\big)\big\rangle-t^2\beta(t)w(t)\big\|\nabla\Phi_{\lambda(t)}\big(x(t)\big)\big\|^2.
\end{aligned}$$

Notice that the terms with $\langle\nabla\Phi_{\lambda(t)}(x(t)),\dot{x}(t)\rangle$ cancel each other; thus, after simplification, we obtain for every $t\geq t_0$

$$\begin{aligned}
\dot{E}_c(t)&=\big(2tw(t)+t^2\dot{w}(t)+\big(\beta(t)+t\dot{\beta}(t)\big)(\alpha-1-c)\big)\big(\Phi_{\lambda(t)}\big(x(t)\big)-\Phi^*\big)\\
&\quad-\big(t^2w(t)+t\beta(t)(\alpha-1-c)\big)\frac{\dot{\lambda}(t)}{2}\big\|\nabla\Phi_{\lambda(t)}\big(x(t)\big)\big\|^2\\
&\quad-ctw(t)\big\langle x(t)-z,\nabla\Phi_{\lambda(t)}\big(x(t)\big)\big\rangle\\
&\quad-(\alpha-1-c)t\big\|\dot{x}(t)\big\|^2-t^2\beta(t)w(t)\big\|\nabla\Phi_{\lambda(t)}\big(x(t)\big)\big\|^2.
\end{aligned} \tag{15}$$

Thanks to (11), $w(t)$ is positive for every $t\geq t_0$, thus

$$-ctw(t)\big\langle x(t)-z,\nabla\Phi_{\lambda(t)}\big(x(t)\big)\big\rangle\leq-ctw(t)\big(\Phi_{\lambda(t)}\big(x(t)\big)-\Phi^*\big),$$

which leads to

$$\begin{aligned}
\dot{E}_c(t)&\leq\big((2-c)tw(t)+t^2\dot{w}(t)+\big(\beta(t)+t\dot{\beta}(t)\big)(\alpha-1-c)\big)\big(\Phi_{\lambda(t)}\big(x(t)\big)-\Phi^*\big)\\
&\quad-\left(\big(t^2w(t)+t\beta(t)(\alpha-1-c)\big)\frac{\dot{\lambda}(t)}{2}+t^2\beta(t)w(t)\right)\big\|\nabla\Phi_{\lambda(t)}\big(x(t)\big)\big\|^2\\
&\quad-(\alpha-c-1)t\big\|\dot{x}(t)\big\|^2.
\end{aligned} \tag{16}$$

By (10) and the fact that λ is nondecreasing, we deduce that

$$\big(t^2w(t) + t\beta(t)(\alpha - 1 - c)\big)\frac{\dot{\lambda}(t)}{2} + t^2\beta(t)w(t) \geq 0,$$

so we obtain for every $t \geq t_0$

$$\begin{aligned}\dot{E}_c(t) \leq{}& \big((2-c)tw(t) + t^2\dot{w}(t) + \big(\beta(t) + t\dot{\beta}(t)\big)(\alpha - 1 - c)\big)\big(\Phi_{\lambda(t)}\big(x(t)\big) - \Phi^*\big)\\ &- \left(\big(t^2w(t) + t\beta(t)(\alpha - c - 1)\big)\frac{\dot{\lambda}(t)}{2} + t^2\beta(t)w(t)\right)\big\|\nabla\Phi_{\lambda(t)}\big(x(t)\big)\big\|^2.\end{aligned} \tag{17}$$

Let us choose $c := \alpha - 1$. According to (12), we obtain for the coefficient of $\Phi_{\lambda(t)}(x(t)) - \Phi^*$ in (17)

$$(2-c)tw(t) + t^2\dot{w}(t) + \big(\beta(t) + t\dot{\beta}(t)\big)(\alpha - 1 - c) = -t\big((\alpha - 3)w(t) - t\dot{w}(t)\big) \leq 0.$$

Therefore, (17) allows us to deduce for every $t \geq t_0$

$$\begin{aligned}\dot{E}_{\alpha-1}(t) &\leq -\big((\alpha - 3)tw(t) - t^2\dot{w}(t)\big)\big(\Phi_{\lambda(t)}\big(x(t)\big) - \Phi^*\big)\\ &\quad - \left(t^2w(t)\frac{\dot{\lambda}(t)}{2} + t^2\beta(t)w(t)\right)\big\|\nabla\Phi_{\lambda(t)}\big(x(t)\big)\big\|^2\\ &\leq 0.\end{aligned}$$

We have just established that $E_{\alpha-1}$ is nonincreasing, which for every $t \geq t_0$ leads to

$$\begin{aligned}E_{\alpha-1}(t) &= t^2w(t)\big(\Phi_{\lambda(t)}\big(x(t)\big) - \Phi^*\big) + \frac{1}{2}\big\|(\alpha - 1)\big(x(t) - z\big) + t\dot{x}(t) + t\beta(t)\nabla\Phi_{\lambda(t)}\big(x(t)\big)\big\|^2\\ &\leq E_{\alpha-1}(t_0).\end{aligned}$$

From here, we obtain for every $t \geq t_0$

$$\Phi_{\lambda(t)}\big(x(t)\big) - \Phi^* \leq \frac{E_{\alpha-1}(t_0)}{t^2w(t)}, \tag{18}$$

which proves (ii). Moreover, by integration, we obtain

$$\int_{t_0}^{+\infty}\left(t^2w(t)\frac{\dot{\lambda}(t)}{2} + t^2\beta(t)w(t)\right)\big\|\nabla\Phi_{\lambda(t)}\big(x(t)\big)\big\|^2\,dt \leq E_{\alpha-1}(t_0) < +\infty \tag{19}$$

and

$$\int_{t_0}^{+\infty}\big((\alpha - 3)tw(t) - t^2\dot{w}(t)\big)\big(\Phi_{\lambda(t)}\big(x(t)\big) - \Phi^*\big)\,dt \leq E_{\alpha-1}(t_0) < +\infty, \tag{20}$$

which are the claims (iii) and (iv).

From now on, we assume that $\alpha > 1$ and choose $c := \alpha - 1 - \varepsilon$, where ε is given by (13). In this setting, (16) reads for every $t \geq t_0$,

$$\begin{aligned}\dot{E}_{\alpha-1-\varepsilon}(t) &\leq \big((3-\alpha+\varepsilon)tw(t) + t^2\dot{w}(t) + \varepsilon\big(\beta(t) + t\dot{\beta}(t)\big)\big)\big(\Phi_{\lambda(t)}\big(x(t)\big) - \Phi^*\big) \\ &\quad - \left(\big(t^2 w(t) + \varepsilon t\beta(t)\big)\frac{\dot{\lambda}(t)}{2} + t^2\beta(t)w(t)\right)\big\|\nabla\Phi_{\lambda(t)}\big(x(t)\big)\big\|^2 - \varepsilon t\big\|\dot{x}(t)\big\|^2 \\ &= -t\big((\alpha-3)w(t) - t\dot{w}(t) - \varepsilon b(t)\big)\big(\Phi_{\lambda(t)}\big(x(t)\big) - \Phi^*\big) \\ &\quad - \left(\big(t^2 w(t) + \varepsilon t\beta(t)\big)\frac{\dot{\lambda}(t)}{2} + t^2\beta(t)w(t)\right)\big\|\nabla\Phi_{\lambda(t)}\big(x(t)\big)\big\|^2 - \varepsilon t\big\|\dot{x}(t)\big\|^2.\end{aligned} \tag{21}$$

So, under the condition (13), $\dot{E}_{\alpha-1-\varepsilon}(t) \leq 0$ for every $t \geq t_0$. Integrating (21), we obtain

$$\int_{t_0}^{+\infty} t\big\|\dot{x}(t)\big\|^2\,dt < +\infty, \tag{22}$$

which gives the claim (v). From the fact that the energy function

$$\begin{aligned}E_{\alpha-1-\varepsilon}(t) &= \big(t^2 w(t) + \varepsilon t\beta(t)\big)\big(\Phi_{\lambda(t)}\big(x(t)\big) - \Phi^*\big) \\ &\quad + \frac{1}{2}\big\|(\alpha-1-\varepsilon)\big(x(t)-z\big) + t\dot{x}(t) + t\beta(t)\nabla\Phi_{\lambda(t)}\big(x(t)\big)\big\|^2 \\ &\quad + \frac{(\alpha-1-\varepsilon)\varepsilon}{2}\big\|x(t)-z\big\|^2\end{aligned}$$

is bounded from above, and it is nonnegative on $[t_0, +\infty)$, it follows that the trajectory x is bounded, which is item (vi). Finally, from (13) and (20), we deduce the claim (vii)

$$\begin{aligned}&\int_{t_0}^{+\infty} \varepsilon t b(t)\big(\Phi_{\lambda(t)}\big(x(t)\big) - \Phi^*\big)\,dt \\ &\leq \int_{t_0}^{+\infty}\big((\alpha-3)tw(t) - t^2\dot{w}(t)\big)\big(\Phi_{\lambda(t)}\big(x(t)\big) - \Phi^*\big)\,dt < +\infty,\end{aligned} \tag{23}$$

which finishes the proof. □

The following auxiliary result will be needed later.

Lemma 3 *Suppose that $\alpha > 1$ and (13) holds, that λ and β are nondecreasing on $[t_0, +\infty)$, and that (11) holds. Then, for a solution $x : [t_0, +\infty) \to H$ to (1), it holds*

$$\int_{t_0}^{+\infty} tw(t)\big\langle\nabla\Phi_{\lambda(t)}\big(x(t)\big), x(t)-z\big\rangle\,dt < +\infty. \tag{24}$$

Proof Recall that according to (15), we have for every $t \geq t_0$

$$\begin{aligned}\dot{E}_c(t) &= \big(2tw(t) + t^2\dot{w}(t) + \big(\beta(t) + t\dot{\beta}(t)\big)(\alpha-1-c)\big)\big(\Phi_{\lambda(t)}\big(x(t)\big) - \Phi^*\big) \\ &\quad - \big(t^2 w(t) + t\beta(t)(\alpha-1-c)\big)\frac{\dot{\lambda}(t)}{2}\big\|\nabla\Phi_{\lambda(t)}\big(x(t)\big)\big\|^2\end{aligned}$$

$$- ctw(t)\langle x(t) - z, \nabla\Phi_{\lambda(t)}(x(t))\rangle - (\alpha - 1 - c)t\|\dot{x}(t)\|^2$$
$$- t^2\beta(t)w(t)\|\nabla\Phi_{\lambda(t)}(x(t))\|^2.$$

We choose again $c := \alpha - 1$ and split the term $(\alpha - 1)tw(t)\langle x(t) - z, \nabla\Phi_{\lambda(t)}(x(t))\rangle$ into the sum of two expressed in terms of ε given by (13). For every $t \geq t_0$, we have

$$\dot{E}_{\alpha-1}(t) \leq \big(2tw(t) + t^2\dot{w}(t)\big)\big(\Phi_{\lambda(t)}(x(t)) - \Phi^*\big) - (\alpha - 1 - \varepsilon)tw(t)\langle x(t) - z, \nabla\Phi_{\lambda(t)}(x(t))\rangle - \varepsilon tw(t)\langle x(t) - z, \nabla\Phi_{\lambda(t)}(x(t))\rangle.$$

By applying the convex subdifferential inequality, we obtain for every $t \geq t_0$

$$\begin{aligned}\dot{E}_{\alpha-1}(t) &\leq \big(2tw(t) + t^2\dot{w}(t) - (\alpha - 1 - \varepsilon)tw(t)\big)\big(\Phi_{\lambda(t)}(x(t)) - \Phi^*\big) \\ &\quad - \varepsilon tw(t)\langle x(t) - z, \nabla\Phi_{\lambda(t)}(x(t))\rangle \\ &= \big(t^2\dot{w}(t) - (\alpha - 3 - \varepsilon)tw(t)\big)\big(\Phi_{\lambda(t)}(x(t)) - \Phi^*\big) \\ &\quad - \varepsilon tw(t)\langle x(t) - z, \nabla\Phi_{\lambda(t)}(x(t))\rangle.\end{aligned} \tag{25}$$

Since β is nondecreasing, for every $t \geq t_0$, it holds

$$b(t) = w(t) + \dot{\beta}(t) + \frac{\beta(t)}{t} \geq w(t),$$

thus, (13) leads to $t^2\dot{w}(t) - (\alpha - 3 - \varepsilon)tw(t) \leq 0$. Consequently, we obtain from (25) by integration

$$\int_{t_0}^{+\infty} tw(t)\langle\nabla\Phi_{\lambda(t)}(x(t)), x(t) - z\rangle\, dt \leq \frac{E_{\alpha-1}(t_0)}{\varepsilon} < +\infty. \qquad \square$$

Now, we are in a position to improve the convergence rates we obtained previously in (18) and derive from here convergence rates for Φ.

Theorem 4 *Suppose that $\alpha > 1$ and (13) holds, that λ and β are nondecreasing on $[t_0, +\infty)$, and that (11) holds. In addition, assume that*

$$\int_{t_0}^{+\infty}\left[\frac{(\dot{\lambda}(t))^2t^3\beta^2(t)}{\lambda^4(t)} - \frac{\dot{\lambda}(t)t^2b(t)}{2\lambda^2(t)}\right]_+ dt < +\infty, \tag{26}$$

where $[\cdot]_+$ denotes the positive part of the expression inside the brackets, and that there exists $C > 0$ such that

$$\frac{d}{dt}\big(t^2b(t)\big) \leq Ctb(t) \quad \text{for every } t \geq t_0. \tag{27}$$

Then, for a solution $x : [t_0, +\infty) \to H$ to (1), it holds

$$\Phi_{\lambda(t)}(x(t)) - \Phi^* = o\left(\frac{1}{t^2b(t)}\right) \quad \text{and} \quad \|\dot{x}(t)\| = o\left(\frac{1}{t}\right) \quad \text{as } t \to +\infty. \tag{28}$$

Moreover,

$$\left\|\nabla\Phi_{\lambda(t)}\left(x(t)\right)\right\| = o\left(\frac{1}{t\sqrt{b(t)\lambda(t)}}\right) \quad \text{as } t \to +\infty, \tag{29}$$

and

$$\begin{aligned} &\Phi\left(\operatorname{prox}_{\lambda(t)\Phi}\left(x(t)\right)\right) - \Phi^* = o\left(\frac{1}{t^2 b(t)}\right) \quad \text{and} \\ &\left\|\operatorname{prox}_{\lambda(t)\Phi}\left(x(t)\right) - x(t)\right\| = o\left(\frac{\sqrt{\lambda(t)}}{t\sqrt{b(t)}}\right) \quad \text{as } t \to +\infty. \end{aligned} \tag{30}$$

Proof First, we notice that for every $t \geq t_0$, it holds

$$\begin{aligned} \left\langle \frac{d}{dt}\left(\nabla\Phi_{\lambda(t)}\left(x(t)\right)\right), \dot{x}(t)\right\rangle &= \left\langle \lim_{h\to 0} \frac{\nabla\Phi_{\lambda(t+h)}(x(t+h)) - \nabla\Phi_{\lambda(t)}(x(t))}{h}, \dot{x}(t)\right\rangle \\ &= \left\langle \lim_{h\to 0} \frac{\nabla\Phi_{\lambda(t+h)}(x(t+h)) - \nabla\Phi_{\lambda(t+h)}(x(t))}{h}, \dot{x}(t)\right\rangle \\ &\quad + \left\langle \lim_{h\to 0} \frac{\nabla\Phi_{\lambda(t+h)}(x(t)) - \nabla\Phi_{\lambda(t)}(x(t))}{h}, \dot{x}(t)\right\rangle. \end{aligned}$$

For every $h > 0$, by the monotonicity of the gradient of a convex function, we have

$$\left\langle \frac{\nabla\Phi_{\lambda(t+h)}(x(t+h)) - \nabla\Phi_{\lambda(t+h)}(x(t))}{h}, \frac{x(t+h) - x(t)}{h}\right\rangle \geq 0,$$

so letting h tend to zero, we obtain

$$\left\langle \lim_{h\to 0} \frac{\nabla\Phi_{\lambda(t+h)}(x(t+h)) - \nabla\Phi_{\lambda(t+h)}(x(t))}{h}, \dot{x}(t)\right\rangle \geq 0.$$

Consequently, for every $t \geq t_0$, it holds

$$\begin{aligned} &\left\langle \frac{d}{dt}\left(\nabla\Phi_{\lambda(t)}\left(x(t)\right)\right), \dot{x}(t)\right\rangle \\ &\quad \geq \left\langle \lim_{h\to 0} \frac{\nabla\Phi_{\lambda(t+h)}(x(t)) - \nabla\Phi_{\lambda(t)}(x(t))}{h}, \dot{x}(t)\right\rangle \\ &\quad = \lim_{h\to 0}\left\langle \frac{(\lambda(t+h))\operatorname{prox}_{\lambda(t)\Phi}(x(t)) - \lambda(t)\operatorname{prox}_{(\lambda(t+h))\Phi}(x(t)) - (\lambda(t+h) - \lambda(t))x(t)}{\lambda(t)\lambda(t+h)h}, \dot{x}(t)\right\rangle \\ &\quad = \lim_{h\to 0}\left\langle \frac{(\lambda(t+h) - \lambda(t))(\operatorname{prox}_{\lambda(t)\Phi}(x(t)) - x(t))}{\lambda(t)\lambda(t+h)h}, \dot{x}(t)\right\rangle \\ &\qquad - \lim_{h\to 0}\left\langle \frac{\operatorname{prox}_{(\lambda(t+h))\Phi}(x(t)) - \operatorname{prox}_{\lambda(t)\Phi}(x(t))}{\lambda(t+h)h}, \dot{x}(t)\right\rangle \\ &\quad \geq \frac{\dot{\lambda}(t)}{\lambda^2(t)}\left\langle \operatorname{prox}_{\lambda(t)\Phi}\left(x(t)\right) - x(t), \dot{x}(t)\right\rangle - \lim_{h\to 0} \frac{(\lambda(t+h) - \lambda(t))\|\nabla\Phi_{\lambda(t)}(x(t))\|\,\|\dot{x}(t)\|}{\lambda(t+h)h} \end{aligned}$$

$$
\begin{aligned}
&= \frac{\dot{\lambda}(t)}{\lambda^2(t)}\left\langle \operatorname{prox}_{\lambda(t)\Phi}\left(x(t)\right) - x(t), \dot{x}(t)\right\rangle - \frac{\dot{\lambda}(t)\|\nabla\Phi_{\lambda(t)}(x(t))\|\|\dot{x}(t)\|}{\lambda(t)} \\
&= -\frac{\dot{\lambda}(t)}{\lambda(t)}\left\langle \nabla\Phi_{\lambda(t)}\left(x(t)\right), \dot{x}(t)\right\rangle - \frac{\dot{\lambda}(t)\|\nabla\Phi_{\lambda(t)}(x(t))\|\|\dot{x}(t)\|}{\lambda(t)} \\
&\geq -\frac{2\dot{\lambda}(t)\|\nabla\Phi_{\lambda(t)}(x(t))\|\|\dot{x}(t)\|}{\lambda(t)},
\end{aligned}
$$

where we used (6), (7), and the Cauchy–Schwarz inequality. Now, we multiply (1) by $t^2\dot{x}(t)$ to deduce, using the inequality above and (14), for every $t \geq t_0$

$$
\begin{aligned}
0 &= t^2\left\langle \ddot{x}(t), \dot{x}(t)\right\rangle + \alpha t\left\|\dot{x}(t)\right\|^2 + t^2\beta(t)\left\langle \frac{d}{dt}\left(\nabla\Phi_{\lambda(t)}\left(x(t)\right)\right), \dot{x}(t)\right\rangle \\
&\quad + t^2 b(t)\left\langle \nabla\Phi_{\lambda(t)}\left(x(t)\right), \dot{x}(t)\right\rangle \\
&\geq t^2\frac{d}{dt}\left(\frac{1}{2}\left\|\dot{x}(t)\right\|^2\right) + \alpha t\left\|\dot{x}(t)\right\|^2 + t^2 b(t)\frac{d}{dt}\left(\Phi_{\lambda(t)}\left(x(t)\right) - \Phi^*\right) \\
&\quad + \frac{\dot{\lambda}(t)t^2 b(t)}{2}\left\|\nabla\Phi_{\lambda(t)}x(t)\right\|^2 - \frac{2t^2\beta(t)\dot{\lambda}(t)}{\lambda(t)}\left\|\nabla\Phi_{\lambda(t)}\left(x(t)\right)\right\|\left\|\dot{x}(t)\right\| \\
&\geq \frac{d}{dt}\left(\frac{t^2}{2}\left\|\dot{x}(t)\right\|^2 + t^2 b(t)\left(\Phi_{\lambda(t)}\left(x(t)\right) - \Phi^*\right)\right) + (\alpha - 1)t\left\|\dot{x}(t)\right\|^2 \\
&\quad - \left(\Phi_{\lambda(t)}\left(x(t)\right) - \Phi^*\right)\frac{d}{dt}\left(t^2 b(t)\right) \\
&\quad - \left\{\left[\left(\frac{\dot{\lambda}(t)}{\lambda(t)}\right)^2 t^3\beta^2(t) - \frac{\dot{\lambda}(t)t^2 b(t)}{2}\right]\left\|\nabla\Phi_{\lambda(t)}\left(x(t)\right)\right\|^2 + t\left\|\dot{x}(t)\right\|^2\right\}.
\end{aligned}
$$

Using (27), we obtain for every $t \geq t_0$

$$
\begin{aligned}
&\frac{d}{dt}\left(\frac{t^2}{2}\left\|\dot{x}(t)\right\|^2 + t^2 b(t)\left(\Phi_{\lambda(t)}\left(x(t)\right) - \Phi^*\right)\right) \\
&\quad \leq [2-\alpha]_+ t\left\|\dot{x}(t)\right\|^2 + \left(\Phi_{\lambda(t)}\left(x(t)\right) - \Phi^*\right)Ctb(t) \\
&\qquad + \left[\left(\frac{\dot{\lambda}(t)}{\lambda(t)}\right)^2 t^3\beta^2(t) - \frac{\dot{\lambda}(t)t^2 b(t)}{2}\right]_+\left\|\nabla\Phi_{\lambda(t)}\left(x(t)\right)\right\|^2.
\end{aligned}
$$

Next, we show the integrability of the right-hand side of the expression above. The first term is integrable according to Theorem 2 (v), and the second one is integrable according to Theorem 2 (vii). Further, since

$$
\left\|\nabla\Phi_{\lambda(t)}\left(x(t)\right) - \nabla\Phi_{\lambda(t)}(z)\right\| \leq \frac{1}{\lambda(t)}\left\|x(t) - z\right\| \quad \forall t \geq t_0,
$$

and taking into the account the boundedness of the trajectory x established in Theorem 2 (vi) and that $z \in \operatorname{argmin}\Phi$, we deduce

$$
\left\|\nabla\Phi_{\lambda(t)}\left(x(t)\right)\right\| = O\left(\frac{1}{\lambda(t)}\right) \quad \text{as } t \to +\infty.
$$

So, under the assumption (26), we obtain that there exists $\widetilde{C} > 0$ such that for every $t \geq t_0$

$$\begin{aligned}
&\int_{t_0}^{t} \left[\left(\frac{\dot{\lambda}(s)}{\lambda(s)} \right)^2 s^3 \beta^2(s) - \frac{\dot{\lambda}(s) s^2 b(s)}{2} \right]_+ \left\| \nabla \Phi_{\lambda(s)}(x(s)) \right\|^2 ds \\
&\leq \widetilde{C} \int_{t_0}^{t} \left[\frac{(\dot{\lambda}(s))^2 s^3 \beta^2(s)}{\lambda^4(s)} - \frac{\dot{\lambda}(s) s^2 b(s)}{2\lambda^2(s)} \right]_+ ds \\
&< +\infty.
\end{aligned}$$

Applying Lemma 6 in the Appendix, we conclude that the following limit

$$L := \lim_{t \to +\infty} \left(\frac{t^2}{2} \left\| \dot{x}(t) \right\|^2 + t^2 b(t) \left(\Phi_{\lambda(t)}(x(t)) - \Phi^* \right) \right) \geq 0$$

exists. We will show that $L = 0$. Supposing that $L > 0$, we deduce that there exists $t^* \geq t_0$ such that for every $t \geq t^*$

$$\frac{t}{2} \left\| \dot{x}(t) \right\|^2 + t b(t) \left(\Phi_{\lambda(t)}(x(t)) - \Phi^* \right) \geq \frac{L}{2t}.$$

Integrating the last inequality on $[t^*, +\infty)$, we arrive at the contradiction with the integrability of the left-hand side as proved in Theorem 2 (v) and (vii). Therefore, $L = 0$, and we obtain

$$\Phi_{\lambda(t)}(x(t)) - \Phi^* = o\left(\frac{1}{t^2 b(t)} \right) \quad \text{and} \quad \left\| \dot{x}(t) \right\| = o\left(\frac{1}{t} \right) \quad \text{as } t \to +\infty.$$

Using the definition of the proximal mapping, we derive

$$\begin{aligned}
&\Phi_{\lambda(t)}(x(t)) - \Phi^* \\
&\quad = \Phi\left(\operatorname{prox}_{\lambda(t)\Phi}(x(t)) \right) - \Phi^* + \frac{1}{2\lambda(t)} \left\| \operatorname{prox}_{\lambda(t)\Phi}(x(t)) - x(t) \right\|^2 \quad \forall t \geq t_0,
\end{aligned} \tag{31}$$

which yields

$$\begin{aligned}
&\Phi\left(\operatorname{prox}_{\lambda(t)\Phi}(x(t)) \right) - \Phi^* = o\left(\frac{1}{t^2 b(t)} \right) \quad \text{and} \\
&\left\| \operatorname{prox}_{\lambda(t)\Phi}(x(t)) - x(t) \right\| = o\left(\frac{\sqrt{\lambda(t)}}{t\sqrt{b(t)}} \right) \quad \text{as } t \to +\infty.
\end{aligned}$$

According to (6), we obtain from here

$$\left\| \nabla \Phi_{\lambda(t)}(x(t)) \right\| = o\left(\frac{1}{t\sqrt{b(t)\lambda(t)}} \right) \quad \text{as } t \to +\infty. \qquad \square$$

3 Convergence of the trajectories

In this section, we will investigate the weak convergence of the trajectory x to a minimizer of Φ.

Theorem 5 *Suppose that* $\alpha > 1$, (11) *and* (13) *hold, and* λ *and* β *are nondecreasing on* $[t_0, +\infty)$. *In addition, assume that*

$$\lim_{t\to+\infty} \frac{\beta(t)}{tw(t)} = 0 \tag{32}$$

and

$$\sup_{t\geq t_0} \frac{\lambda(t)}{t} < +\infty. \tag{33}$$

If $x : [t_0, +\infty) \to H$ *is a solution to* (1), *then* $x(t)$ *converges weakly to a minimizer of* Φ *as* $t \to +\infty$.

Proof Let $z \in \operatorname{argmin} \Phi$. Previously, in Theorem 2, we established the existence of the limit of $E_c(t)$ as $t \to +\infty$ for $c = \alpha - 1$ and $c = \alpha - 1 - \varepsilon$, where $\varepsilon \in (0, \alpha - 1)$ is given by (13). Thus, computing the difference

$$\begin{aligned} E_{\alpha-1-\varepsilon}(t) - E_{\alpha-1}(t) &= \varepsilon t\beta(t)\big(\Phi_{\lambda(t)}\big(x(t)\big) - \Phi^*\big) + \frac{\varepsilon(\alpha-1)}{2}\big\|x(t) - z\big\|^2 \\ &\quad - \varepsilon\big\langle(\alpha-1)\big(x(t) - z\big) + t\big(\dot{x}(t) + \beta(t)\nabla\Phi_{\lambda(t)}\big(x(t)\big)\big), x(t) - z\big\rangle \\ &= \varepsilon t\beta(t)\big(\Phi_{\lambda(t)}\big(x(t)\big) - \Phi^*\big) - \frac{\varepsilon(\alpha-1)}{2}\big\|x(t) - z\big\|^2 \\ &\quad - \varepsilon t\big\langle\dot{x}(t) + \beta(t)\nabla\Phi_{\lambda(t)}\big(x(t)\big), x(t) - z\big\rangle, \end{aligned}$$

we deduce that the limit of the right-hand side exists. Thanks to (18), we derive for every $t \geq t_0$

$$t\beta(t)\big(\Phi_{\lambda(t)}\big(x(t)\big) - \Phi^*\big) \leq t\beta(t)\frac{E_{\alpha-1}(t_0)}{t^2w(t)} = E_{\alpha-1}(t_0)\frac{\beta(t)}{tw(t)}$$

and from here, based on the assumption (32), we obtain

$$\lim_{t\to+\infty} t\beta(t)\big(\Phi_{\lambda(t)}\big(x(t)\big) - \Phi^*\big) = 0. \tag{34}$$

Hereby, we derived that the limit of the quantity

$$p(t) := \frac{\alpha-1}{2}\big\|x(t) - z\big\|^2 + t\big\langle\dot{x}(t), x(t) - z\big\rangle + t\beta(t)\big\langle\nabla\Phi_{\lambda(t)}\big(x(t)\big), x(t) - z\big\rangle$$

exists as $t \to +\infty$. Now, we are ready to prove the existence of the limit of $\|x(t) - z\|$ as $t \to +\infty$. Denote

$$q(t) := \frac{\alpha-1}{2}\big\|x(t) - z\big\|^2 + (\alpha-1)\int_{t_0}^{t} \beta(s)\big\langle\nabla\Phi_{\lambda(s)}\big(x(s)\big), x(s) - z\big\rangle\, ds \quad \forall t \geq t_0.$$

For every $t \geq t_0$, it holds that

$$p(t) = q(t) + \frac{t}{\alpha-1}\dot{q}(t) - (\alpha-1)\int_{t_0}^{t} \beta(s)\big\langle\nabla\Phi_{\lambda(s)}\big(x(s)\big), x(s) - z\big\rangle\, ds,$$

since

$$\dot{q}(t) = (\alpha - 1)\langle x(t) - z, \dot{x}(t)\rangle + (\alpha - 1)\big(\beta(s)\langle \nabla \Phi_{\lambda(s)}\big(x(s)\big), x(s) - z\rangle\big)$$

and

$$q(t) + \frac{t}{\alpha - 1}\dot{q}(t) = \frac{\alpha - 1}{2}\|x(t) - z\|^2 + (\alpha - 1)\int_{t_0}^{t} \beta(s)\langle \nabla \Phi_{\lambda(s)}\big(x(s)\big), x(s) - z\rangle\, ds$$
$$+ t\langle x(t) - z, \dot{x}(t)\rangle + t\big(\beta(s)\langle \nabla \Phi_{\lambda(s)}\big(x(s)\big), x(s) - z\rangle\big).$$

By Lemma 3, we established that $\int_{t_0}^{+\infty} sw(s)\langle \nabla \Phi_{\lambda(s)}(x(s)), x(s) - z\rangle\, ds < +\infty$. In turn, (32) yields that

$$\lim_{t\to+\infty} \int_{t_0}^{t} \beta(s)\langle \nabla \Phi_{\lambda(s)}\big(x(s)\big), x(s) - z\rangle\, ds \quad \text{exists.} \tag{35}$$

Finally,

$$\lim_{t\to+\infty} \left(q(t) + \frac{t}{\alpha - 1}\dot{q}(t)\right) \quad \text{also exists.}$$

Applying now Lemma 7 in the Appendix, we immediately get the existence of the limit of $q(t)$ as $t \to +\infty$. By the definition of q and (35), we establish the first statement of Opial's Lemma (see Lemma 8 in the Appendix), namely, that, for any $z \in \operatorname{argmin} \Phi$,

$$\lim_{t\to+\infty} \|x(t) - z\| \quad \text{exists.}$$

To establish the second term of Opial's Lemma, first note that from (31) and (34), we have by denoting $\xi(t) := \operatorname{prox}_{\lambda(t)\Phi}(x(t))$, $\lim_{t\to+\infty} t\beta(t)(\Phi(\xi(t)) - \Phi^*) = 0$ and $\lim_{t\to+\infty} \frac{t\beta(t)}{\lambda(t)}\|\xi(t) - x(t)\|^2 = 0$. Using that β is nondecreasing and assumption (33), we deduce

$$\lim_{t\to+\infty} \Phi\big(\xi(t)\big) = \Phi^* \quad \text{and} \quad \lim_{t\to+\infty} \|\xi(t) - x(t)\| = 0.$$

Considering a sequence $\{t_k\}_{k\in\mathbb{N}}$ such that $\{x(t_k)\}_{k\in\mathbb{N}}$ converges weakly to an element $z \in H$ as $k \to +\infty$, we notice that $\{\xi(t_k)\}_{k\in\mathbb{N}}$ converges weakly to z as $k \to +\infty$. Now, the function Φ being convex and lower semicontinuous in the weak topology allows us to write

$$\Phi(z) \le \liminf_{k\to+\infty} \Phi\big(\xi(t_k)\big) = \lim_{t\to+\infty} \Phi\big(\xi(t)\big) = \Phi^*.$$

Hence, $z \in \operatorname{argmin} \Phi$, and the second statement of Opial's Lemma is shown. This gives the weak convergence of the trajectory $x(t)$ to a minimizer of Φ as $t \to +\infty$. □

Remark 1 In the hypotheses of Theorem 4, to obtain the convergence of the trajectories, besides (33), it is enough to assume that

$$\sup_{t\ge t_0} \frac{\beta(t)}{tw(t)} < +\infty$$

to guarantee (35). Indeed, in this case, (34) follows from the conclusion of Theorem 4

$$\lim_{t\to+\infty} t\beta(t)\big(\Phi_{\lambda(t)}(x(t)) - \Phi^*\big) \leq \lim_{t\to+\infty} t^2 b(t)\big(\Phi_{\lambda(t)}(x(t)) - \Phi^*\big)\frac{\beta(t)}{tw(t)} = 0.$$

Remark 2 (Implicit discretization) Implicit discretization of the dynamical system (1) with fixed step size $h > 0$ leads to the numerical scheme that reads for every $k \geq 1$ (see also [7, 8])

$$\frac{x_{k+1} - 2x_k + x_{k-1}}{h^2} + \frac{\alpha(x_{k+1} - x_k)}{kh^2} + \frac{\beta_k(\nabla\Phi_{\lambda_{k+1}}(x_{k+1}) - \nabla\Phi_{\lambda_k}(x_k))}{h} + b_k\nabla\Phi_{\lambda_{k+1}}(x_{k+1}) = 0,$$

where x_k, λ_k, β_k, and b_k denote $x(kh)$, $\lambda(kh)$, $\beta(kh)$, and $b(kh)$, respectively. Rearranging the terms, one obtains for every $k \geq 1$

$$x_{k+1} + \frac{kh(\beta_k + hb_k)\nabla\Phi_{\lambda_{k+1}}(x_{k+1})}{\alpha + k} = x_k + \frac{k(x_k - x_{k-1})}{\alpha + k} + \frac{kh\beta_k\nabla\Phi_{\lambda_k}(x_k)}{\alpha + k}$$

or, equivalently,

$$(\forall k \geq 1) \quad \begin{cases} y_k := x_k + \frac{k}{\alpha+k}(x_k - x_{k-1}) + \frac{kh\beta_k}{\alpha+k}\nabla\Phi_{\lambda_k}(x_k), \\ x_{k+1} := \operatorname{prox}_{\frac{kh(\beta_k+hb_k)}{\alpha+k}\Phi_{\lambda_k}}(y_k). \end{cases}$$

Relation (6), namely,

$$\nabla\Phi_\lambda(x) = \frac{1}{\lambda}\big(x - \operatorname{prox}_{\lambda\Phi}(x)\big) \quad \forall x \in H,$$

and the property of the proximal mapping (see, for instance, [16])

$$\operatorname{prox}_{\mu\Phi_\lambda}(x) = \frac{\lambda}{\lambda+\mu}x + \frac{\mu}{\lambda+\mu}\operatorname{prox}_{(\lambda+\mu)\Phi}(x) \quad \forall x \in H, \forall \mu, \lambda > 0,$$

lead to the following formulation of the implicit numerical algorithm

$$(\forall k \geq 1) \quad \begin{cases} y_k := x_k + \frac{k}{\alpha+k}(x_k - x_{k-1}) + \frac{kh\beta_k}{\lambda_k(\alpha+k)}(x_k - \operatorname{prox}_{\lambda_k\Phi}(x_k)), \\ x_{k+1} := \frac{\lambda_k(\alpha+k)}{\lambda_k(\alpha+k)+kh(\beta_k+hb_k)}y_k + \frac{kh(\beta_k+hb_k)}{\lambda_k(\alpha+k)+kh(\beta_k+hb_k)}\operatorname{prox}_{\frac{\lambda_k(\alpha+k)+kh(\beta_k+hb_k)}{\alpha+k}\Phi}(y_k), \end{cases}$$

where $x_0, x_1 \in H$ are given starting points.

4 Polynomial choices for the system parameter functions

According to the previous two sections, to guarantee both the fast convergence rates in Theorem 4 and the convergence of the trajectory to a minimizer of Φ in Theorem 5, also by taking into account Remark 1, it is enough to make the following assumptions on the system parameter functions

(I) $\alpha > 1$, and there exists $\varepsilon \in (0, \alpha - 1)$ such that $(\alpha - 3)w(t) - t\dot{w}(t) \geq \varepsilon b(t)$ for every $t \geq t_0$;

(II) β and λ are nondecreasing on $[t_0, +\infty)$;

(III) $b(t) > \dot{\beta}(t) + \frac{\beta(t)}{t}$ for every $t \geq t_0$;

(IV) $\int_{t_0}^{+\infty}[\frac{\beta^2(t)(\dot{\lambda}(t))^2t^3}{\lambda^4(t)} - \frac{\dot{\lambda}(t)t^2b(t)}{2\lambda^2(t)}]_+\,dt < +\infty$;

(V) there exists $C > 0$ such that $\frac{d}{dt}(t^2 b(t)) \leq Ctb(t)$ for every $t \geq t_0$;
(VI) $\sup_{t\geq t_0} \frac{\beta(t)}{tw(t)} < +\infty$;
(VII) $\sup_{t\geq t_0} \frac{\lambda(t)}{t} < +\infty$.

In this section, we will investigate the fulfillment of these conditions for

$$b(t) = bt^n, \qquad \beta(t) = \beta t^m \quad \text{and} \quad \lambda(t) = \lambda t^l,$$

where $n, m, l \in \mathbb{R}$, $b, \lambda > 0$ and $\beta \geq 0$.

For this choice of b, condition (V) is fulfilled.

First, we assume that $\beta = 0$. Then the conditions (III), (IV), and (VI) are fulfilled, while the conditions (II) and (VI) are nothing else than $0 \leq l \leq 1$. Condition (I) asks for $\alpha > 1$ and for the existence of $\varepsilon \in (0, \alpha - 1)$ such that for every $t \geq t_0$

$$(\alpha - 3 - n - \varepsilon)bt^n \geq 0$$

or, equivalently, $\alpha - 3 - n \geq \varepsilon$. To this end, it is enough to have that $\alpha - 3 > n$.

In case $\beta > 0$, conditions (II) and (VII) are nothing else than $m \geq 0$ and $0 \leq l \leq 1$. Condition (III) reads for every $t \geq t_0$

$$bt^n > m\beta t^{m-1} + \beta t^{m-1} = (m + 1)\beta t^{m-1},$$

or, equivalently,

$$t^{n-m+1} > \frac{(m + 1)\beta}{b}.$$

From here, we get

$$0 \leq m \leq n + 1,$$

and $b > (m + 1)\beta t_0^{m-1-n}$.

Condition (VI) requires that

$$\sup_{t\geq t_0} \frac{\beta t^m}{t(bt^n - \beta m t^{m-1} - \beta t^{m-1})} = \sup_{t\geq t_0} \frac{\beta t^m}{bt^{n+1} - \beta t^m (m + 1)} < +\infty,$$

and it is obviously fulfilled.

Condition (I) asks for $\alpha > 1$ and for the existence of $\varepsilon \in (0, \alpha - 1)$ such that for every $t \geq t_0$

$$(\alpha - 3)\big(bt^n - \beta m t^{m-1} - \beta t^{m-1}\big) - t\big(bnt^{n-1} - \beta m(m - 1)t^{m-2} - \beta(m - 1)t^{m-2}\big) \geq \varepsilon bt^n.$$

After simplification, we obtain that for every $t \geq t_0$

$$(\alpha - 3 - n - \varepsilon)bt^n + \beta(m + 1)(m + 2 - \alpha)t^{m-1} \geq 0$$

or, equivalently,

$$(\alpha - 3 - n - \varepsilon)bt^{n-m+1} \geq \beta(m + 1)(\alpha - m - 2).$$

On the one hand, we have $m = n + 1$ and $(\alpha - 3 - n)(1 - \frac{\beta(n+2)}{b}) > \varepsilon$, which requires that $\alpha - 3 - n > 0$. On the other hand, we have $m < n + 1$, which also requires that $\alpha - 3 - n > 0$.

Consequently, we have to assume that

$$\alpha - 3 > n \quad \text{and} \quad b > \frac{\beta(m+1)(\alpha - m - 2)}{(\alpha - 3 - n)t_0^{n-m+1}}.$$

In this case, there will be always an $\varepsilon \in (0, \alpha - 1)$ such that $\alpha - 3 - n - \varepsilon > 0$ and

$$b > \frac{\beta(m+1)(\alpha - m - 2)}{(\alpha - 3 - n - \varepsilon)t_0^{n-m+1}} > \frac{\beta(m+1)(\alpha - m - 2)}{(\alpha - 3 - n)t_0^{n-m+1}},$$

in other words, which satisfies condition (I).

Finally, let us have a closer look at condition (V). This reads as

$$\int_{t_0}^{+\infty} \left[\frac{\beta^2 t^{2m} (l\lambda)^2 t^{2l-2} t^3}{\lambda^4 t^{4l}} - \frac{l\lambda t^{l-1} t^2 b t^n}{2\lambda^2 t^{2l}} \right]_+ dt < +\infty$$

or, equivalently,

$$\int_{t_0}^{+\infty} \left[\left(\frac{\beta l}{\lambda} \right)^2 t^{2m-2l+1} - \frac{lb}{2\lambda} t^{n-l+1} \right]_+ dt = \int_{t_0}^{+\infty} \left[\left(\frac{\beta l}{\lambda} \right)^2 - \frac{lb}{2\lambda} t^{n+l-2m} \right]_+ t^{2m-2l+1}\, dt < +\infty.$$

1. In case

$$n + l > 2m,$$

there exists $t_1 \geq t_0$ such that for every $t \geq t_1$

$$\left(\frac{\beta l}{\lambda} \right)^2 - \frac{lb}{2\lambda} t^{n+l-2m} \leq 0.$$

Therefore, we obtain

$$\begin{aligned} &\int_{t_0}^{+\infty} \left[\left(\frac{\beta l}{\lambda} \right)^2 - \frac{lb}{2\lambda} t^{n+l-2m} \right]_+ t^{2m-2l+1}\, dt \\ &= \int_{t_0}^{t_1} \left[\left(\frac{\beta l}{\lambda} \right)^2 - \frac{lb}{2\lambda} t^{n+l-2m} \right]_+ t^{2m-2l+1}\, dt + \int_{t_1}^{+\infty} \left[\left(\frac{\beta l}{\lambda} \right)^2 - \frac{lb}{2\lambda} t^{n+l-2m} \right]_+ t^{2m-2l+1}\, dt \\ &= \int_{t_0}^{t_1} \left[\left(\frac{\beta l}{\lambda} \right)^2 - \frac{lb}{2\lambda} t^{n+l-2m} \right]_+ t^{2m-2l+1}\, dt < +\infty, \end{aligned}$$

thus (V) is fulfilled.

2. In case

$$n + l < 2m,$$

there exist $\delta > 0$ and $t_2 \geq t_0$ such that for all $t \geq t_2$

$$\left(\frac{\beta l}{\lambda} \right)^2 - \frac{lb}{2\lambda} t^{n+l-2m} > \delta.$$

Taking into account that $2m-2l+1 > n+1-l \geq -1$, we have

$$\int_{t_0}^{+\infty}\left[\left(\frac{\beta l}{\lambda}\right)^2 - \frac{lb}{2\lambda}t^{n+l-2m}\right]_+ t^{2m-2l+1}\,dt$$
$$\geq \int_{t_0}^{t_2}\left[\left(\frac{\beta l}{\lambda}\right)^2 - \frac{lb}{2\lambda}t^{n+l-2m}\right]_+ t^{2m-2l+1}\,dt + \delta\int_{t_2}^{+\infty} t^{2m-2l+1}\,dt = +\infty,$$

thus, (V) is not fulfilled.

3. It is only left to consider the case

$$n+l = 2m.$$

Condition (V) becomes

$$\int_{t_0}^{+\infty}\left[\left(\frac{\beta l}{\lambda}\right)^2 - \frac{lb}{2\lambda}\right]_+ t^{2m-2l+1}\,dt < +\infty.$$

If $b \geq \frac{2\beta^2 l}{\lambda}$, then it is fulfilled. Otherwise, since $2m-2l+1 = n-l+1 \geq -1$, it is not fulfilled.

Summarising, all convergence statements in Theorem 4 and Theorem 5 hold in the two settings

1. $\alpha > 1$, $\beta = 0$, $\alpha - 3 > n$, $0 \leq l \leq 1$, and $b, \lambda > 0$;
2. $\alpha > 1$, $\beta > 0$, $\alpha - 3 > n$, $0 \leq l \leq 1$, $0 \leq m \leq n+1$, $b > \frac{(m+1)(\alpha-m-2)\beta}{(\alpha-3-n)t_0^{n-m+1}}$, $\lambda > 0$, and either $2m < n+l$, or $2m = n+l$ and $b \geq \frac{2l\beta^2}{\lambda}$.

Remark 3 Theorem 4 is providing for the choices $b(t) = bt^n$ and $\lambda(t) = \lambda t^l$ the following convergence rates

$$\Phi\left(\operatorname{prox}_{\lambda(t)\Phi}(x(t))\right) - \Phi^* = o\left(\frac{1}{t^{n+2}}\right), \qquad \left\|\operatorname{prox}_{\lambda(t)\Phi}(x(t)) - x(t)\right\| = o\left(\frac{1}{t^{\frac{n}{2}+1-\frac{l}{2}}}\right)$$

and

$$\left\|\nabla\Phi_{\lambda(t)}(x(t))\right\| = o\left(\frac{1}{t^{\frac{n}{2}+1+\frac{l}{2}}}\right),$$

as $t \to +\infty$. Clearly, the bigger the n is, the faster the convergence is. On the other hand, concerning the exponent l, things are a bit more complicated: we may gain in one case but inevitably lose in the other. An interesting case is when $l = 0$, which corresponds to λ being a constant function. In this case, one can notice a balance between accelerating the convergence of $\|\nabla\Phi_{\lambda(t)}(x(t))\|$ and slowing the latter for $\|\operatorname{prox}_{\lambda(t)\Phi}(x(t)) - x(t)\|$, since none of them are affected by l anymore.

5 Numerical experiments

In this section, we will conduct series of experiments to investigate the influence of the system parameters λ, β, and b on the convergence behavior of dynamical system. We will successively fix two of them and vary the last one for this. For the numerical experiments, we will restrict ourselves to the polynomial choices addressed in the previous section $\lambda(t) = t^l$, $\beta(t) = t^m$, $b(t) = bt^n$ with $b = \frac{(m+1)(\alpha-m-2)\beta}{(\alpha-3-n)t_0^{n-m+1}} + 1$, as well as $x(t_0) = x_0 = 10$, $\dot{x}(t_0) = 0$, and $t_0 = 1$.

5.1 The influence of b on the dynamical behaviour

First, let us choose as objective function $\Phi : \mathbb{R} \to \mathbb{R}_+$, $\Phi(x) = |x|$, fix $m = 0$, $\alpha = 9$ and $l = 1$, and vary n.

In Fig. 1, we clearly see that the faster the exponent of the function b grows the faster the convergence of the function values of the Moreau envelope and its gradient are, starting with the slowest pace for $n = 0$ and accelerating until $n = 4.99$, confirming the theoretical convergence rates. In addition, the increase in the exponent of b also seems to improve the convergence behavior of the trajectory. Fast growing exponents for b will improve the convergence greatly; however, as seen in the previous section, they are limited by the upper bound value $\alpha - 3$.

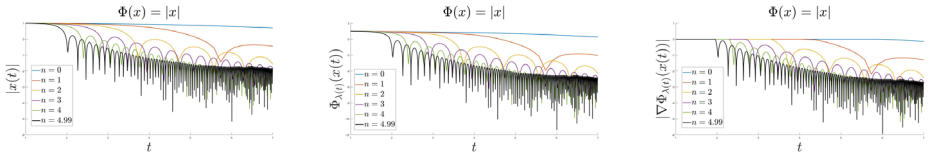
5.2 The influence of λ on the dynamical behaviour

For the same objective function as in the previous subsection, we study the behavior of the dynamics when varying the exponent l to investigate the influence of the function λ. To this end, we fix $m = 0$, $\alpha = 9$ and $n = 5 < \alpha - 3$, and take for l three different values from 0 to 1. We also choose the starting point $x_0 = 1$ to provide a better illustration.

One can notice in Fig. 2 that the convergence behavior of the function values of the Moreau envelope and its gradient is better, the higher l is, whereas, interestingly enough, for the convergence of the trajectories, an opposite phenomenon takes place.

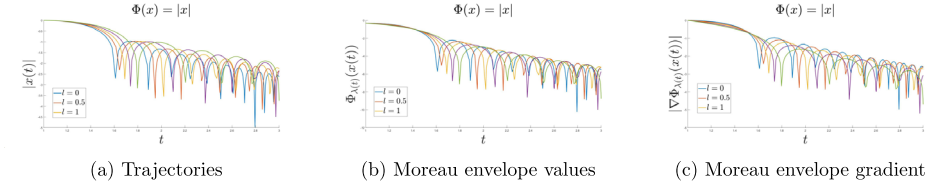
5.3 The influence of β on the dynamical behaviour

Let $\Phi : \mathbb{R} \to \mathbb{R}_+$, $\Phi(x) = |x| + \frac{x^2}{2}$, $\alpha = 13$, $n = 9 < \alpha - 3$ and $l = 1$. We vary the exponent m such that $2m < n + l$ to study the influence of the function β on the convergence behavior of the system. In Fig. 3, we see that, even though m does not explicitly appear in the theoretical convergence rates for the gradient of the Moreau envelope and the trajectory of the system, it influences the convergence behavior of both of them as well as of the function values of the Moreau envelope, in the sense that these are faster, the higher the values of m are.



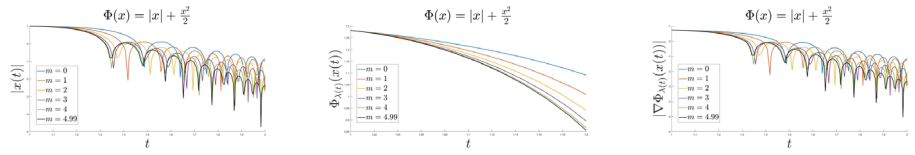
(a) Trajectories (b) Moreau envelope values (c) Moreau envelope gradient

Figure 1 $m = 0$, $\alpha = 9$ and $l = 1$



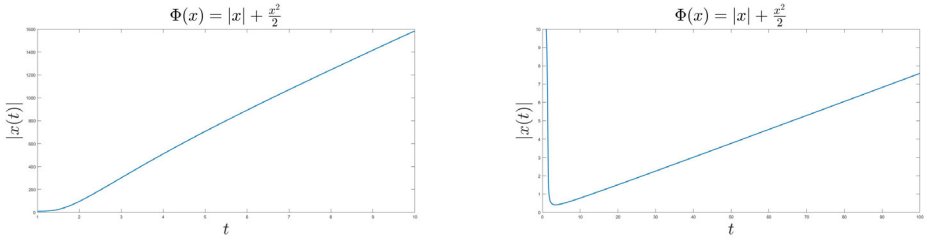
(a) Trajectories (b) Moreau envelope values (c) Moreau envelope gradient

Figure 2 $m = 0$, $\alpha = 9$ and $n = 5$



(a) Trajectories (b) Moreau envelope values (c) Moreau envelope gradient

Figure 3 $n = 9$, $\alpha = 13$ and $l = 1$



(a) $n = 9$, $\alpha = 13$, $l = 1$ and $m = 12$ (b) $n = 4$, $\alpha = 2$, $l = 4$ and $m = 6$

Figure 4 Divergence of the trajectories

Finally, we consider two parameter choices, which lie outside the convergence setting derived in the previous section, and notice that these fundamentally affect the convergence of the trajectory. In Fig. 4(a), we choose m such that the condition $2m < n + l$ is violated, and in Fig. 4(b), we choose α and n such that the condition $\alpha - 3 > n$ is also violated. One can see that in both settings, the trajectories diverge.

Appendix

In this appendix, we collect some lemmas that play an important role in the proof of the main results of the paper. For the proof of the following lemma, we refer to [2].

Lemma 6 *Suppose that $f : [t_0, +\infty) \to \mathbb{R}$ is locally absolutely continuous and bounded from below and there exists $g \in L^1([t_0, +\infty), \mathbb{R})$ such that for almost all $t \geq t_0$*

$$\frac{d}{dt} f(t) \leq g(t).$$

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

For the proof of the following lemma, we refer to [15].

Lemma 7 *Let H be a real Hilbert space and $x : [t_0, +\infty) \longrightarrow \mathbb{H}$ a continuously differentiable function satisfying $x(t) + \frac{t}{\alpha}\dot{x}(t) \to L$ as $t \to +\infty$, with $\alpha > 0$ and $L \in \mathbb{H}$. Then $x(t) \to L$ as $t \to +\infty$.*

Finally, we state a continuous version of Opial's Lemma (see [22]), which is used in the proof of the convergence of the trajectory.

Lemma 8 *Let S be a non-empty subset of a real Hilbert space H and $x : [0, +\infty) \mapsto H$ a given map. Assume that*

- *for every* $z \in S$, $\lim_{t\to+\infty} \|x(t) - z\|$ *exists*;
- *every weak sequential cluster point of the map* x *belongs to* S.

Then, $x(t)$ *converges weakly to some element of* S *as* $t \to +\infty$.

Acknowledgements
The authors are grateful to E.R. Csetnek (University of Vienna) and to three anonymous reviewers for comments and remarks, which improved the quality of the paper.

Funding
The research of RIB was partially supported by FWF (Austrian Science Fund), project P 34922-N. The research of MKA was supported by the Doctoral Programme *Vienna Graduate School on Computational Optimization (VGSCO)* funded by FWF (Austrian Science Fund), project W 1260. Open Access funding provided by Austrian Science Fund (FWF).

Availability of data and materials
Not applicable.

Declarations

Competing interests
The authors declare that they have no competing interests.

Author contribution
All authors contributed in writing this article. All authors read and approved the final manuscript.

Publisher's Note

Received: 2 March 2022 Accepted: 22 November 2022 Published online: 16 December 2022

References

1. Alvarez, F., Attouch, H., Bolte, J., Redont, P.: A second-order gradient-like dissipative dynamical system with Hessian-driven damping, application to optimization and mechanics. J. Math. Pures Appl. **81**(8), 747–779 (2002)
2. Attouch, H., Abbas, B., Svaiter, B.F.: Newton-like dynamics and forward–backward methods for structured monotone inclusions in Hilbert spaces. J. Optim. Theory Appl. **161**(2), 331–360 (2014)
3. Attouch, H., Alvarez, F.: The heavy ball with friction dynamical system for convex constrained minimization problems. In: Optimization. Lecture Notes in Economics and Mathematical Systems, vol. 481, pp. 25–35. Springer, Berlin (2000)
4. Attouch, H., Alvarez, F.: An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping. Set-Valued Anal. **9**, 3–11 (2001)
5. Attouch, H., Balhag, A., Chbani, Z., Riahi, H.: Fast convex optimization via inertial dynamics combining viscous and Hessian-driven damping with time rescaling. Evol. Equ. Control Theory **11**(2), 487–514 (2022)
6. Attouch, H., Cabot, A.: Convergence of damped inertial dynamics governed by regularized maximally monotone operators. J. Differ. Equ. **264**, 7138–7182 (2018)
7. Attouch, H., Chbani, Z., Fadili, J., Riahi, H.: First-order optimization algorithms via inertial systems with Hessian driven damping. Math. Program. **193**, 113–155 (2022)
8. Attouch, H., Chbani, Z., Fadili, J., Riahi, H.: Convergence of iterates for first-order optimization algorithms with inertia and Hessian driven damping. Optimization (2022). https://doi.org/10.1080/02331934.2021.2009828
9. Attouch, H., Chbani, Z., Peypouquet, J., Redont, P.: Fast convergence of inertial dynamics and algorithms with asymptotic vanishing viscosity. Math. Program. **168**, 123–175 (2018)
10. Attouch, H., Chbani, Z., Riahi, H.: Fast convex optimization via time scaling of damped inertial gradient dynamics. SIAM J. Optim. **29**(3), 2227–2256 (2019)
11. Attouch, H., Goudou, X., Redont, P.: The heavy ball with friction method. The continuous dynamical system, global exploration of the local minima of a real-valued function by asymptotical analysis of a dissipative dynamical system. Commun. Contemp. Math. **2**(1), 1–34 (2000)
12. Attouch, H., László, S.: Continuous Newton-like inertial dynamics for monotone inclusions. Set-Valued Var. Anal. **29**, 555–581 (2021)
13. Attouch, H., Peypouquet, J.: The rate of convergence of Nesterov's accelerated forward-backward method is actually faster than $\frac{1}{k^2}$. SIAM J. Optim. **26**(3), 1824–1834 (2016)
14. Attouch, H., Peypouquet, J.: Convergence of inertial dynamics and proximal algorithms governed by maximal monotone operators. Math. Program. **174**(1–2), 391–432 (2019)
15. Attouch, H., Peypouquet, J., Redont, P.: Fast convex optimization via inertial dynamics with Hessian driven damping damping. J. Differ. Equ. **261**(10), 5734–5783 (2016)
16. Bauschke, H.H., Combettes, P.L.: Convex Analysis and Monotone Operator Theory in Hilbert Spaces. CMS Books in Mathematics. Springer, Berlin (2016)
17. Boţ, R.I., Csetnek, E.R.: Second order forward–backward dynamical systems for monotone inclusion problems. SIAM J. Control Optim. **54**(3), 1423–1443 (2016)
18. Cabot, A., Engler, H., Gadat, S.: On the long time behavior of second order differential equations with asymptotically small dissipation and insights. Trans. Am. Math. Soc. **361**, 5983–6017 (2009)

19. Cabot, A., Engler, H., Gadat, S.: Second order differential equations with asymptotically small dissipation and piecewise flat potentials. In: Proceedings of the Seventh Mississippi State–UAB Conference on Differential Equations and Computational Simulations. Electron. J. Differ. Equ. Conf., vol. 17, pp. 33–38. Southwest Texas State Univ., San Marcos (2009)
20. May, R.: Asymptotic for a second-order evolution equation with convex potential and vanishing damping term. Turk. J. Math. **41**(3), 681–685 (2017)
21. Nesterov, Y.: A method for solving the convex programming problem with convergence rate $O(\frac{1}{k^2})$. Dokl. Akad. Nauk SSSR **269**(3), 543–547 (1983)
22. Opial, Z.: Weak convergence of the sequence of successive approximations for nonexpansive mappings. Bull. Am. Math. Soc. **73**(4), 591–597 (1967)
23. Polyak, B.T.: Some methods of speeding up the convergence of iterative methods. USSR Comput. Math. Math. Phys. **4**(5), 1–17 (1964)
24. Su, W., Boyd, S., Candès, E.J.: A differential equation for modeling Nesterov's accelerated gradient method: theory and insights. J. Mach. Learn. Res. **17**, 1–43 (2016)