Mathematical Programming (2021) 189:151–186
https://doi.org/10.1007/s10107-020-01528-8

FULL LENGTH PAPER

Series B

Tikhonov regularization of a second order dynamical system with Hessian driven damping

Radu Ioan Boţ[1] · **Ernö Robert Csetnek**[1] · **Szilárd Csaba László**[2]

Received: 28 November 2019 / Accepted: 18 May 2020 / Published online: 11 June 2020

Abstract

We investigate the asymptotic properties of the trajectories generated by a second-order dynamical system with Hessian driven damping and a Tikhonov regularization term in connection with the minimization of a smooth convex function in Hilbert spaces. We obtain fast convergence results for the function values along the trajectories. The Tikhonov regularization term enables the derivation of strong convergence results of the trajectory to the minimizer of the objective function of minimum norm.

Keywords Second order dynamical system · Convex optimization · Tikhonov regularization · Fast convergence methods · Hessian-driven damping

Mathematics Subject Classification 34G25 · 47J25 · 47H05 · 90C26 · 90C30 · 65K10

The paper is dedicated to Prof. Marco A. López on the occasion of his 70th birthday.

Radu Ioan Boţ: Research partially supported by FWF (Austrian Science Fund), Project I 2419-N32. Ernö Robert Csetnek: Research supported by FWF (Austrian Science Fund), Project P 29809-N32. Szilárd Csaba László: This work was supported by a grant of Ministry of Research and Innovation, CNCS—UEFISCDI, Project Number PN-III-P1-1.1-TE-2016-0266 and by a Grant of Ministry of Research and Innovation, CNCS—UEFISCDI, Project Number PN-III-P4-ID-PCE-2016-0190, within PNCDI III.

✉ Radu Ioan Boţ
radu.bot@univie.ac.at

Ernö Robert Csetnek
ernoe.robert.csetnek@univie.ac.at

Szilárd Csaba László
szilard.laszlo@math.utcluj.ro

[1] Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria

[2] Department of Mathematics, Technical University of Cluj-Napoca, Memorandumului 28, Cluj-Napoca, Romania

1 Introduction

The paper of Su et al. [20] was the starting point of intensive research of second order dynamical systems with an asymptotically vanishing damping term of the form

$$\ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \nabla g(x(t)) = 0, \ t \geq t_0 > 0, \tag{1}$$

where $g : \mathcal{H} \longrightarrow \mathbb{R}$ is a convex and continuously Fréchet differentiable function defined on a real Hilbert space $\mathcal{H}$ fulfilling $\operatorname{argmin} g \neq \emptyset$. The aim is to approach by the trajectories generated by this system the solution set of the optimization problem

$$\min_{x \in \mathcal{H}} g(x). \tag{2}$$

The convergence rate of the objective function along the trajectory is in case $\alpha > 3$ of

$$g(x(t)) - \min g = o\left(\frac{1}{t^2}\right),$$

while in case $\alpha = 3$ it is of

$$g(x(t)) - \min g = O\left(\frac{1}{t^2}\right),$$

where $\min g \in \mathbb{R}$ denotes the minimal value of g. Also in view of this fact, system (1) is seen as a continuous version of the celebrated Nesterov accelerated gradient scheme (see [16]). In what concerns the asymptotic properties of the generated trajectories, weak convergence to a minimizer of g as the time goes to infinity has been proved by Attouch et al. [7] (see also [6]) for $\alpha > 3$. Without any further geometrical assumption on g, the convergence of the trajectories in the case $\alpha \leq 3$ is still an open problem.

Second order dynamical systems with a geometrical Hessian driven damping term have aroused the interest of the researchers, due to both their applications in optimization and mechanics and their natural relations to Newton and Levenberg-Marquardt iterative methods (see [2]). Furthermore, it has been observed for some classes of optimization problems that a geometrical damping term governed by the Hessian can induce a stabilization of the trajectories. In [11] the dynamical system with Hessian driven damping term

$$\ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \beta\nabla^2 g(x(t))\dot{x}(t) + \nabla g(x(t)) = 0, \ t \geq t_0 > 0, \tag{3}$$

where $\alpha \geq 3$ and $\beta > 0$, has been investigated in relation with the optimization problem (2). Fast convergence rates for the values and the gradient of the objective function along the trajectories are obtained and the weak convergence of the trajectories to a minimizer of g is shown. We would also like to mention that iterative schemes which result via (symplectic) discretizations of dynamical systems with Hessian driven damping terms have been recently formulated and investigated from the point of view of their convergence properties in [5,18,19].

Another development having as a starting point (1) is the investigation of dynamical systems involving a Tikhonov regularization term. Attouch, Chbani and Riahi investigated in this context in [8] the system

$$\ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \nabla g(x(t)) + \epsilon(t)x(t) = 0, \ t \geq t_0 > 0, \tag{4}$$

where $\alpha \geq 3$ and $\epsilon : [t_0, +\infty) \longrightarrow [0, +\infty)$. One of the main benefits of considering such a regularized dynamical system is that it generates trajectories which converge strongly to the minimum norm solution of (2). Besides that, in [8] it was proved that the fast convergence rate of the objective function values along the trajectories remains unaltered. For more insights into the role played by the Tikhonov regularization for optimization problems and, more general, for monotone inclusion problems, we refer the reader to [3,4,9,15].

This being said, it is natural to investigate a second order dynamical system which combines a Hessian driven damping and a Tikhonov regularization term and to examine if it inherits the properties of the dynamical systems (3) and (4). This is the aim of the manuscript, namely the analysis in the framework of the general assumption stated below of the dynamical system

$$\ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \beta\nabla^2 g(x(t))\dot{x}(t) + \nabla g(x(t)) + \epsilon(t)x(t) = 0, \ t \geq t_0 > 0, \ x(t_0) = u_0, \ \dot{x}(t_0) = v_0, \tag{5}$$

where $\alpha \geq 3$ and $\beta \geq 0$, and $u_0, v_0 \in \mathcal{H}$.

General assumption:

- $g : \mathcal{H} \longrightarrow \mathbb{R}$ is a convex and twice Fréchet differentiable function with Lipschitz continuous gradient on bounded sets and $\operatorname{argmin} g \neq \emptyset$;
- $\epsilon : [t_0, +\infty) \longrightarrow [0, +\infty)$ is a nonincreasing function of class C^1 fulfilling $\lim_{t \longrightarrow +\infty} \epsilon(t) = 0$.

The fact that the starting time t_0 is taken as strictly greater than zero comes from the singularity of the damping coefficient $\frac{\alpha}{t}$. This is not a limitation of the generality of the proposed approach, since we will focus on the asymptotic behaviour of the generated trajectories. Notice that if $\mathcal{H}$ is finite-dimensional, then the Lipschitz continuity of ∇g on bounded sets follows from the continuity of $\nabla^2 g$.

To which extent the Tikhonov regularization does influence the convergence behaviour of the trajectories generated by (5) can be seen even when minimizing a one dimensional function. Consider the convex and twice continuously differentiable function

$$g : \mathbb{R} \to \mathbb{R}, \quad g(x) = \begin{cases} -(x+1)^3, & \text{if } x < -1 \\ 0, & \text{if } -1 \leq x \leq 1 \\ (x-1)^3, & \text{if } x > 1. \end{cases} \tag{6}$$

It holds that $\operatorname{argmin} g = [-1, 1]$ and $x^* = 0$ is its minimum norm solution. In the second column of Fig. 1 we can see the behaviour of the trajectories generated

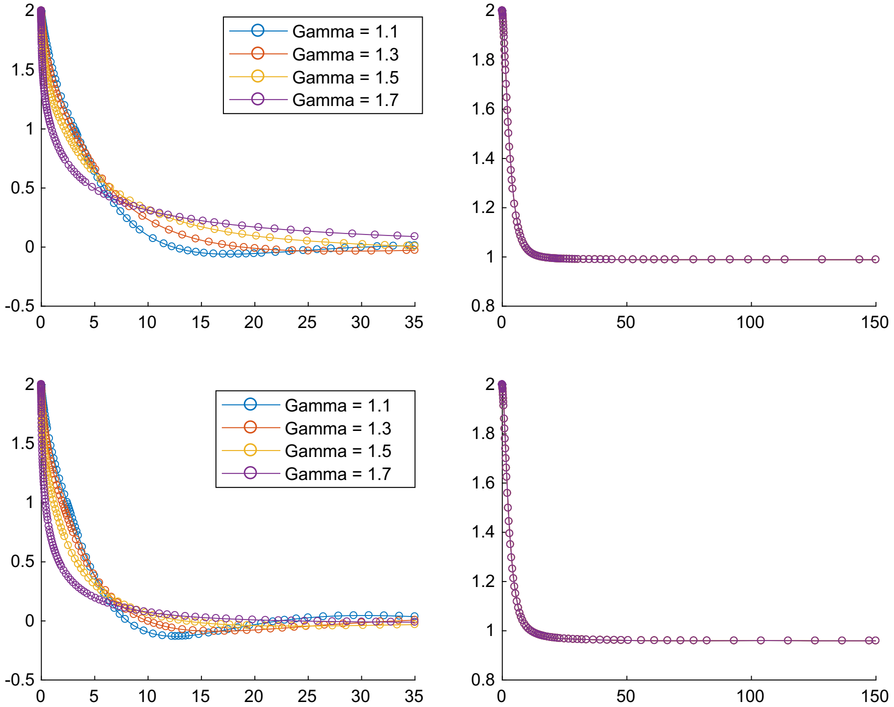

Fig. 1 First column: the trajectories of the dynamical system with Tikhonov regularization $\epsilon(t) = t^{-\gamma}$ are approaching the minimum norm solution $x^* = 0$. Second column: the trajectories of the dynamical system without Tikhonov regularization the trajectory are approaching the optimal solution 1

by the dynamical system without Tikhonov regularization (which corresponds to the case when ϵ is identically 0) for $\beta = 1$ and $\alpha = 3$ and, respectively, $\alpha = 4$. In both cases the trajectories are approaching the optimal solution 1, which is a minimizer of g, however, not the minimum norm solution.

In the first column of Fig. 1 we can see the behaviour of the trajectories generated by the dynamical system with Tikhonov parametrizations of the form $t \mapsto \epsilon(t) = t^{-\gamma}$, for different values of $\gamma \in (1, 2)$, which is in accordance to the conditions in Theorem 4.4, $\beta = 1$ and $\alpha = 3$ and, respectively, $\alpha = 4$. The trajectories are approaching the minimum norm solution $x^* = 0$.

The organization of the paper is as follows. We start the analysis of the dynamical system (5) by proving the existence and uniqueness of a global C^2-solution. In the third section we provide two different settings for the Tikhonov parametrization $t \mapsto \epsilon(t)$ in both of which $g(x(t))$ converges to $\min g$, the minimal value of g, with a convergence rate of $O\left(\frac{1}{t^2}\right)$ for $\alpha = 3$ and of $o\left(\frac{1}{t^2}\right)$ for $\alpha > 3$. The proof relies on Lyapunov theory; the choice of the right energy functional plays a decisive role in this context. Weak convergence of the trajectory is also derived for $\alpha > 3$. In the last section we focus on the proof of strong convergence to a minimum norm solution: firstly, in a general setting, for the ergodic trajectory, and, secondly, in a slightly restrictive setting, for the trajectory $x(t)$ itself.

2 Existence and uniqueness

In this section we will prove the existence and uniqueness of a global C^2-solution of the dynamical system (5). The proof of the existence and uniqueness theorem is based on the idea to reformulate (5) as a particular first order dynamical system in a suitably chosen product space (see also [11]).

Theorem 2.1 *For every initial value $(u_0, v_0) \in \mathcal{H} \times \mathcal{H}$, there exists a unique global C^2-solution $x : [t_0, +\infty) \to \mathcal{H}$ to (5).*

Proof Let $(u_0, v_0) \in \mathcal{H} \times \mathcal{H}$. First we assume that $\beta = 0$, which gives the dynamical system (4) investigated in [8]. The statement follows from [14, Proposition 2.2(b)] (see also the discussion in [8, Section 2]).

Assume now that $\beta > 0$. We notice that $x : [t_0, +\infty) \longrightarrow \mathcal{H}$ is a solution of the dynamical system (5), that is

$$\ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \beta\nabla^2 g(x(t))\dot{x}(t) + \nabla g(x(t)) + \epsilon(t)x(t)=0,\ \ x(t_0)=u_0,\ \dot{x}(t_0)=v_0,$$

if and only if $(x, y) : [t_0, +\infty) \longrightarrow \mathcal{H} \times \mathcal{H}$ is a solution of the dynamical system

$$\begin{cases} \dot{x}(t) + \beta\nabla g(x(t)) - y(t) = 0 \\ \dot{y}(t) + \frac{\alpha}{t}\dot{x}(t) + \nabla g(x(t)) + \epsilon(t)x(t) = 0 \\ x(t_0) = u_0,\ y(t_0) = v_0 + \beta\nabla g(u_0), \end{cases}$$

which is further equivalent to

$$\begin{cases} \dot{x}(t) + \beta\nabla g(x(t)) - y(t) = 0 \\ \dot{y}(t) + \frac{\alpha}{t}y(t) + \left(1 - \frac{\alpha\beta}{t}\right)\nabla g(x(t)) + \epsilon(t)x(t) = 0 \\ x(t_0) = u_0,\ y(t_0) = v_0 + \beta\nabla g(u_0). \end{cases} \tag{7}$$

We define $F : [t_0, +\infty) \times \mathcal{H} \times \mathcal{H} \to \mathcal{H} \times \mathcal{H}$ by

$$F(t, u, v) = \left(-\beta\nabla g(u) + v, -\frac{\alpha}{t}v - \left(1 - \frac{\alpha\beta}{t}\right)\nabla g(u) - \epsilon(t)u\right),$$

and write (7) as

$$\begin{cases} \big(\dot{x}(t), \dot{y}(t)\big) = F(t, x(t), y(t)) \\ \big(x(t_0), y(t_0)\big) = \big(u_0, v_0 + \beta\nabla g(u_0)\big). \end{cases} \tag{8}$$

Since ∇g is Lipschitz continuous on bounded sets and continuously differentiable, the local existence and uniqueness theorem (see [17, Theorems 46.2 and 46.3]) guarantees the existence of a unique solution (x, y) of (8) defined on a maximum intervall $[t_0, T_{\max})$, where $t_0 < T_{\max} \leq +\infty$. Furthermore, either $T_{\max} = +\infty$ or $\lim_{t \to T_{\max}} \|x(t)\| + \|y(t)\| = +\infty$. We will prove that $T_{\max} = +\infty$, which will imply that x is the unique global C^2-solution of (5).

Consider the energy functional (see [10])

$$\mathcal{E} : [t_0, +\infty) \to \mathbb{R}, \quad \mathcal{E}(t) = \frac{1}{2}\|\dot{x}(t)\|^2 + g(x(t)) + \frac{1}{2}\epsilon(t)\|x(t)\|^2.$$

By using (5) we get

$$\frac{d}{dt}\mathcal{E}(t) = -\frac{\alpha}{t}\|\dot{x}(t)\|^2 - \beta\langle \nabla^2 g(x(t))\dot{x}(t), \dot{x}(t)\rangle + \frac{1}{2}\dot{\epsilon}(t)\|x(t)\|^2,$$

and, since ϵ is nonincreasing and $\nabla^2 g(x(t))$ is positive semidefinite, we obtain that

$$\frac{d}{dt}\mathcal{E}(t) \leq 0 \quad \forall t \geq t_0.$$

Consequently, $\mathcal{E}$ is nonincreasing, hence

$$\frac{1}{2}\|\dot{x}(t)\|^2 + g(x(t)) + \frac{1}{2}\epsilon(t)\|x(t)\|^2 \leq \frac{1}{2}\|\dot{x}(t_0)\|^2 + g(x(t_0)) + \frac{1}{2}\epsilon(t_0)\|x(t_0)\|^2 \quad \forall t \geq t_0.$$

From the fact that g is bounded from below we obtain that $\dot{x}$ is bounded on $[t_0, T_{\max})$. Let $\|\dot{x}\|_\infty := \sup_{t \in [t_0, T_{\max})} \|\dot{x}(t)\| < +\infty$.

Since $\|x(t) - x(t')\| \leq \|\dot{x}\|_\infty |t - t'|$ for all $t, t' \in [t_0, T_{\max})$, there exists $\lim_{t \longrightarrow T_{\max}} x(t)$, which shows that x is bounded on $[t_0, T_{\max})$. Since $\dot{x}(t) + \beta \nabla g(x(t)) = y(t)$ for all $t \in [t_0, T_{\max})$ and ∇g is Lipschitz continuous on bounded sets, it yields that y is also bounded on $[t_0, T_{\max})$. Hence $\lim_{t \to T_{\max}} \|x(t)\| + \|y(t)\|$ cannot be $+\infty$, thus $T_{\max} = +\infty$, which completes the proof. $\square$

3 Asymptotic analysis

In this section we will show to which extent different assumptions we impose to the Tikhonov parametrization $t \mapsto \epsilon(t)$ influence the asymptotic behaviour of the trajectory x generated by the dynamical system (5). In particular, we are looking at the convergence of the function g along the trajectory and the weak convergence of the trajectory.

We recall that the asymptotic analysis of the system (5) is carried out in the framework of the general assumptions stated in the introduction.

We start with a result which provides a setting that guarantees the convergence of $g(x(t))$ to $\min g$ as $t \to +\infty$.

Theorem 3.1 *Let x be the unique global C^2-solution of* (5). *Assume that one of the following conditions is fulfilled:*

(a) $\int_{t_0}^{+\infty} \frac{\epsilon(t)}{t} dt < +\infty$ *and there exist* $a > 1$ *and* $t_1 \geq t_0$ *such that*

$$\dot{\epsilon}(t) \leq -\frac{a\beta}{2}\epsilon^2(t) \text{ for every } t \geq t_1;$$

(b) *there exists* $a > 0$ *and* $t_1 \geq t_0$ *such that*

$$\epsilon(t) \leq \frac{a}{t} \text{ for every } t \geq t_1.$$

If $\alpha \geq 3$*, then*

$$\lim_{t \to +\infty} g(x(t)) = \min g.$$

Proof Let be $x^* \in \operatorname{argmin} g$ and $2 \leq b \leq \alpha - 1$ be fixed. We introduce the following energy functional $\mathcal{E}_b : [t_0, +\infty) \to \mathbb{R}$,

$$\begin{aligned}\mathcal{E}_b(t) = &(t^2 - \beta(b + 2 - \alpha)t)\,(g(x(t)) - \min g) + \frac{t^2\epsilon(t)}{2}\|x(t)\|^2 \\ &+ \frac{1}{2}\|b(x(t) - x^*) + t(\dot{x}(t) + \beta\nabla g(x(t)))\|^2 + \frac{b(\alpha - 1 - b)}{2}\|x(t) - x^*\|^2.\end{aligned} \tag{9}$$

For every $t \geq t_0$ it holds

$$\begin{aligned}\dot{\mathcal{E}}_b(t) = &(2t - \beta(b + 2 - \alpha))\,(g(x(t)) - \min g) \\ &+ (t^2 - \beta(b + 2 - \alpha)t)\langle\nabla g(x(t), \dot{x}(t)\rangle + \frac{t^2\dot{\epsilon}(t) + 2t\epsilon(t)}{2}\|x(t)\|^2 + t^2\epsilon(t)\langle\dot{x}(t), x(t)\rangle \\ &+ \langle(b + 1)\dot{x}(t) + \beta\nabla g(x(t)) + t(\ddot{x}(t) + \beta\nabla^2 g(x(t))\dot{x}(t)), b(x(t) - x^*) + t(\dot{x}(t) + \beta\nabla g(x(t)))\rangle \\ &+ b(\alpha - 1 - b)\langle\dot{x}(t), x(t) - x^*\rangle.\end{aligned} \tag{10}$$

Now, by using (5), we get for every $t \geq t_0$

$$\begin{aligned}&\langle(b + 1)\dot{x}(t) + \beta\nabla g(x(t)) + t(\ddot{x}(t) + \beta\nabla^2 g(x(t))\dot{x}(t)), b(x(t) - x^*) + t(\dot{x}(t) + \beta\nabla g(x(t)))\rangle \\ &= \langle(b + 1 - \alpha)\dot{x}(t) + (\beta - t)\nabla g(x(t)) - t\epsilon(t)x(t), b(x(t) - x^*) + t(\dot{x}(t) + \beta\nabla g(x(t)))\rangle \\ &= b(b + 1 - \alpha)\langle\dot{x}(t), x(t) - x^*\rangle + (b + 1 - \alpha)t\|\dot{x}(t)\|^2 + (-t^2 + \beta(b + 2 - \alpha)t\langle\dot{x}(t), \nabla g(x(t))\rangle \\ &\quad + (\beta^2 t - \beta t^2)\|\nabla g(x(t))\|^2 - \epsilon(t)t^2\langle\dot{x}(t), x(t)\rangle - \beta\epsilon(t)t^2\langle\nabla g(x(t)), x(t)\rangle \\ &\quad - bt\left\langle\left(1 - \frac{\beta}{t}\right)\nabla g(x(t)) + \epsilon(t)x(t), x(t) - x^*\right\rangle.\end{aligned} \tag{11}$$

Let be $t_0' := \max(\beta, t_0)$. For all $t \geq t_0'$ the function $g_t : \mathcal{H} \to \mathbb{R}$, $g_t(x) = \left(1 - \frac{\beta}{t}\right)g(x) + \frac{\epsilon(t)}{2}\|x\|^2$, is strongly convex, thus, one has

$$g_t(y) - g_t(x) \geq \langle\nabla g_t(x), y - x\rangle + \frac{\epsilon(t)}{2}\|y - x\|^2 \;\forall x, y \in \mathcal{H}.$$

By taking $x := x(t)$ and $y := x^*$ we get for every $t \geq t_0'$

$$-bt\left\langle \left(1-\frac{\beta}{t}\right)\nabla g(x(t)) + \epsilon(t)x(t), x(t)-x^*\right\rangle \leq -bt\left(1-\frac{\beta}{t}\right)(g(x(t))-\min g) - bt\frac{\epsilon(t)}{2}\|x(t)\|^2 - bt\frac{\epsilon(t)}{2}\|x(t)-x^*\|^2 + bt\frac{\epsilon(t)}{2}\|x^*\|^2. \tag{12}$$

From (10), (11) and (12) it follows that for every $t \geq t_0'$ it holds

$$\begin{aligned}\dot{\mathcal{E}}_b(t) \leq & \big((2-b)t - \beta(2-\alpha)\big)(g(x(t))-\min g) + bt\frac{\epsilon(t)}{2}\|x^*\|^2 \\ & + \left(t^2\frac{\dot{\epsilon}(t)}{2} + (2-b)t\frac{\epsilon(t)}{2}\right)\|x(t)\|^2 - bt\frac{\epsilon(t)}{2}\|x(t)-x^*\|^2 \\ & + (b+1-\alpha)t\|\dot{x}(t)\|^2 + (\beta^2 t - \beta t^2)\|\nabla g(x(t))\|^2 - \beta\epsilon(t)t^2\langle\nabla g(x(t)), x(t)\rangle. \end{aligned} \tag{13}$$

At this point we treat the situations $\alpha > 3$ and $\alpha = 3$ separately.

The case $\alpha > 3$ and $2 < b < \alpha - 1$. We will carry out the analysis by addressing the settings provided by the conditions (a) and (b) separately.

Condition (a) *holds:* Assuming that condition (a) holds, there exist $a > 1$ and $t_1 \geq t_0'$ such that

$$\dot{\epsilon}(t) \leq -\frac{a\beta}{2}\epsilon^2(t) \ \text{ for every } t \geq t_1.$$

Using that

$$-\beta\epsilon(t)t^2\langle\nabla g(x(t)), x(t)\rangle \leq \frac{\beta t^2}{a}\|\nabla g(x(t))\|^2 + \frac{a\beta\epsilon^2(t)t^2}{4}\|x(t)\|^2, \tag{14}$$

(13) leads to the following estimate

$$\begin{aligned}\dot{\mathcal{E}}_b(t) \leq & \big((2-b)t - \beta(2-\alpha)\big)(g(x(t))-\min g) + bt\frac{\epsilon(t)}{2}\|x^*\|^2 \\ & + \left(t^2\frac{\dot{\epsilon}(t)}{2} + (2-b)t\frac{\epsilon(t)}{2} + \frac{a\beta\epsilon^2(t)t^2}{4}\right)\|x(t)\|^2 - bt\frac{\epsilon(t)}{2}\|x(t)-x^*\|^2 \\ & + (b+1-\alpha)t\|\dot{x}(t)\|^2 + \left(\beta^2 t - \beta\left(1-\frac{1}{a}\right)t^2\right)\|\nabla g(x(t))\|^2, \end{aligned} \tag{15}$$

which holds for every $t \geq t_1$.

Since $a > 1$ and $b > 2$, we notice that for every $t \geq t_1$ it holds

$$t^2\frac{\dot{\epsilon}(t)}{2} + (2-b)t\frac{\epsilon(t)}{2} + \frac{a\beta\epsilon^2(t)t^2}{4} \leq 0.$$

On the other hand, we have that

$$\beta^2 t - \beta\left(1 - \frac{1}{a}\right)t^2 \leq -\beta\frac{a-1}{2a}t^2 \text{ for every } t \geq \frac{2a\beta}{a-1}$$

and

$$(2-b)t - \beta(2-\alpha) \leq 0 \text{ for every } t \geq \frac{\beta(\alpha-2)}{b-2}.$$

We define $t_2 := \max\left(t_1, \frac{2a\beta}{a-1}, \frac{\beta(\alpha-2)}{b-2}\right)$. According to (15), it holds for every $t \geq t_2$

$$\begin{aligned}
&\dot{\mathcal{E}}_b(t) - \big((2-b)t - \beta(2-\alpha)\big)\left(g(x(t)) - \min g\right) - \left(t^2\frac{\dot{\epsilon}(t)}{2} + (2-b)t\frac{\epsilon(t)}{2} + \frac{a\beta\epsilon^2(t)t^2}{4}\right)\|x(t)\|^2 \\
&\quad + bt\frac{\epsilon(t)}{2}\|x(t) - x^*\|^2 + (\alpha - 1 - b)t\|\dot{x}(t)\|^2 + \beta\frac{a-1}{2a}t^2\|\nabla g(x(t))\|^2 \\
&\leq bt\frac{\epsilon(t)}{2}\|x^*\|^2.
\end{aligned} \tag{16}$$

Condition (b) *holds:* Assuming now that condition (b) holds, there exist $a > 0$ and $t_1 \geq t_0'$ such that

$$\epsilon(t) \leq \frac{a}{t} \text{ for every } t \geq t_1.$$

Further, the monotonicity of ∇g and the fact that $\nabla g(x^*) = 0$ implies that

$$\langle \nabla g(x(t)), x(t) - x^* \rangle \geq 0 \text{ for every } t \geq t_0.$$

Using that

$$-\beta\epsilon(t)t^2\langle \nabla g(x(t)), x(t)\rangle \leq -\beta\epsilon(t)t^2\langle \nabla g(x(t)), x^*\rangle \leq \frac{\beta t^3\epsilon(t)}{2a}\|\nabla g(x(t))\|^2 + \frac{a\beta\epsilon(t)t}{2}\|x^*\|^2, \tag{17}$$

(13) leads to the following estimate

$$\begin{aligned}
\dot{\mathcal{E}}_b(t) \leq &\big((2-b)t - \beta(2-\alpha)\big)\left(g(x(t)) - \min g\right) + (b + a\beta)t\frac{\epsilon(t)}{2}\|x^*\|^2 \\
&+ \left(t^2\frac{\dot{\epsilon}(t)}{2} + (2-b)t\frac{\epsilon(t)}{2}\right)\|x(t)\|^2 - bt\frac{\epsilon(t)}{2}\|x(t) - x^*\|^2 \\
&+ (b + 1 - \alpha)t\|\dot{x}(t)\|^2 + \left(\beta^2 t - \beta t^2 + \frac{\beta t^3\epsilon(t)}{2a}\right)\|\nabla g(x(t))\|^2
\end{aligned} \tag{18}$$

for every $t \geq t_1$.

Since $b > 2$, we have that for every $t \geq t_1$ it holds

$$t^2\frac{\dot{\epsilon}(t)}{2} + (2-b)t\frac{\epsilon(t)}{2} \leq 0.$$

On the other hand, since

$$-\beta t^2 + \frac{\beta t^3\epsilon(t)}{2a} \leq -\frac{\beta}{2}t^2$$

holds for every $t \geq t_1$, it follows that

$$\beta^2 t - \beta t^2 + \frac{\beta t^3\epsilon(t)}{2a} \leq -\frac{\beta}{4}t^2 \text{ for every } t \geq \max(t_1, 4\beta). \tag{19}$$

We recall that

$$(2-b)t - \beta(2-\alpha) \leq 0 \text{ for every } t \geq \frac{\beta(\alpha-2)}{b-2}.$$

We define $t_2 := \max\left(t_1, 4\beta, \frac{\beta(\alpha-2)}{b-2}\right)$. According to (18), it holds for every $t \geq t_2$

$$\begin{aligned}
&\dot{\mathcal{E}}_b(t) - ((2-b)t - \beta(2-\alpha))\,(g(x(t)) - \min g) - \left(t^2\frac{\dot{\epsilon}(t)}{2} + (2-b)t\frac{\epsilon(t)}{2}\right)\|x(t)\|^2 \\
&\quad + bt\frac{\epsilon(t)}{2}\|x(t)-x^*\|^2 + (\alpha-1-b)t\|\dot{x}(t)\|^2 + \frac{\beta}{4}t^2\|\nabla g(x(t))\|^2 \\
&\leq (b + a\beta)t\frac{\epsilon(t)}{2}\|x^*\|^2.
\end{aligned} \tag{20}$$

From now on we will treat the two cases together. According to (16), in case (a), and to (20), in case (b), we obtain

$$\dot{\mathcal{E}}_b(t) \leq lt\frac{\epsilon(t)}{2}\|x^*\|^2$$

for every $t \geq t_2$, where $l := b$ and $t_2 = \max\left(t_1, \frac{2a\beta}{a-1}, \frac{\beta(\alpha-2)}{b-2}\right)$, in case (a), and $l := b + a\beta$ and $t_2 = \max\left(t_1, 4\beta, \frac{\beta(\alpha-2)}{b-2}\right)$ in case (b).

By integrating the latter inequality on the interval $[t_2, T]$, where $T \geq t_2$ is arbitrarily chosen, we obtain

$$\mathcal{E}_b(T) \leq \mathcal{E}_b(t_2) + \frac{l\|x^*\|^2}{2}\int_{t_2}^{T} t\epsilon(t)dt.$$

On the other hand,

$$\mathcal{E}_b(t) \geq (t^2 - \beta(b+2-\alpha)t)\,(g(x(T)) - \min g)\ \forall t \geq t_0,$$

hence, for every $T \geq \max(\beta(b+2-\alpha), t_3)$ we get

$$0 \leq g(x(T)) - \min g \leq \frac{\mathcal{E}_b(t_2)}{T^2 - \beta(b+2-\alpha)T} + \frac{l\|x^*\|^2}{2} \frac{1}{T^2 - \beta(b+2-\alpha)T} \int_{t_2}^{T} t\epsilon(t)dt.$$

Obviously,

$$\lim_{T \longrightarrow +\infty} \frac{\mathcal{E}_b(t_3)}{T^2 - \beta(b+2-\alpha)T} = 0.$$

Further, Lemma A.1 applied to the functions $\varphi(t) = t^2$ and $f(t) = \frac{\epsilon(t)}{t}$ provides

$$\lim_{T \longrightarrow +\infty} \frac{1}{T^2} \int_{t_2}^{T} t^2 \frac{\epsilon(t)}{t} dt = 0,$$

hence,

$$\lim_{T \longrightarrow +\infty} \frac{1}{T^2 - \beta(b+2-\alpha)T} \int_{t_2}^{T} t\epsilon(t)dt = 0$$

and, consequently,

$$\lim_{T \longrightarrow +\infty} g(x(T)) = \min g.$$

The case $\alpha = 3$ and $b = 2$. In this case the energy functional reads

$$\mathcal{E}_2(t) = (t^2 - \beta t)\,(g(x(t)) - \min g) + \frac{t^2\epsilon(t)}{2}\|x(t)\|^2 + \frac{1}{2}\|2(x(t) - x^*) + t(\dot{x}(t) + \beta\nabla g(x(t)))\|^2$$

for every $t \geq t_0$. We will address again the settings provided by the conditions (a) and (b) separately.

Condition (a) *holds:* Relation (15) becomes

$$\dot{\mathcal{E}}_2(t) \leq \beta\,(g(x(t)) - \min g) + t\epsilon(t)\|x^*\|^2 + \left(t^2\frac{\dot{\epsilon}(t)}{2} + \frac{a\beta\epsilon^2(t)t^2}{4}\right)\|x(t)\|^2 - t\epsilon(t)\|x(t) - x^*\|^2$$
$$+ \left(\beta^2 t - \beta\left(1 - \frac{1}{a}\right)t^2\right)\|\nabla g(x(t))\|^2$$

for every $t \geq t_1$. Consequently, for $t_3 := \max\left(t_1, \frac{\beta a}{a-1}\right)$, we have

$$\dot{\mathcal{E}}_2(t) \leq \beta\left(g(x(t)) - g^*\right) + t\epsilon(t)\|x^*\|^2 \tag{21}$$

for every $t \geq t_3$. After multiplication with $(t - \beta)$, it yields

$$t(t-\beta)\dot{\mathcal{E}}_2(t) \leq \beta t(t-\beta)\left(g(x(t)) - g^*\right) + t^2(t-\beta)\epsilon(t)\|x^*\|^2 \leq \beta\mathcal{E}_2(t) + t^2(t-\beta)\epsilon(t)\|x^*\|^2$$

for every $t \geq t_3$. Dividing by $(t-\beta)^2$ we obtain

$$\frac{t}{t-\beta}\dot{\mathcal{E}}_2(t) \leq \frac{\beta}{(t-\beta)^2}\mathcal{E}_2(t) + \frac{t^2}{t-\beta}\epsilon(t)\|x^*\|^2$$

or, equivalently,

$$\frac{d}{dt}\left(\frac{t}{t-\beta}\mathcal{E}_2(t)\right) \leq \frac{t^2}{t-\beta}\epsilon(t)\|x^*\|^2 \text{ for every } t \geq t_3. \tag{22}$$

Condition (b) *holds:* We define $t_3 := \max(t_1, 4\beta)$. Relation (18) becomes

$$\dot{\mathcal{E}}_2(t) \leq \beta\left(g(x(t)) - g^*\right) + \frac{2+a\beta}{2}t\epsilon(t)\|x^*\|^2, \tag{23}$$

for every $t \geq t_3$. Repeating the above steps for the inequality (23) we obtain

$$\frac{d}{dt}\left(\frac{t}{t-\beta}\mathcal{E}_2(t)\right) \leq \frac{2+a_1\beta}{2}\frac{t^2}{t-\beta}\epsilon(t)\|x^*\|^2 \text{ for every } t \geq t_3. \tag{24}$$

From now on we will treat the two cases together. According to (22), in case (a), and to (24), in case (b), we obtain

$$\frac{d}{dt}\left(\frac{t}{t-\beta}\mathcal{E}_2(t)\right) \leq l\frac{t^2}{t-\beta}\epsilon(t)\|x^*\|^2$$

for every $t \geq t_3$, where $l := 1$ and $t_3 = \max\left(t_1, \frac{\beta(\alpha-1)}{b-2}\right)$, in case (a), and $l := \frac{2+a\beta}{2}$ and $t_3 = \max(t_1, 4\beta)$ in case (b).

By integrating the latter inequality on an interval $[t_3, T]$, where $T \geq t_3$ is arbitrarily chosen, we obtain

$$\frac{T}{T-\beta}\mathcal{E}_2(T) \leq \frac{t_3}{t_3-\beta}\mathcal{E}_2(t_3) + l\|x^*\|^2\int_{t_3}^{T}\frac{t^2}{t-\beta}\epsilon(t)dt.$$

On the other hand,

$$\mathcal{E}_2(t) \geq (t^2 - \beta t)\left(g(x(t)) - \min g\right)$$

for every $t \geq t_0$, hence, for every $T \geq \max(\beta, t_3) = t_3$ we get

$$0 \leq g(x(T)) - \min g \leq \frac{1}{T^2}\frac{t_3}{t_3-\beta}\mathcal{E}_2(t_3) + l\|x^*\|^2\frac{1}{T^2}\int_{t_3}^{T}\frac{t^2}{t-\beta}\epsilon(t)dt.$$

Obviously,

$$\lim_{T\longrightarrow+\infty}\frac{1}{T^2}\frac{t_3}{t_3-\beta}\mathcal{E}_2(t_3) = 0.$$

Lemma A.1, applied this time to the functions $\varphi(t) = \frac{t^3}{t-\beta}$ and $f(t) = \frac{\epsilon(t)}{t}$, yields

$$\lim_{T \longrightarrow +\infty} \frac{T-\beta}{T^3} \int_{t_3}^{T} \frac{t^3}{t-\beta} \frac{\epsilon(t)}{t} dt = 0.$$

Consequently,

$$\lim_{T \longrightarrow +\infty} \frac{1}{T^2} \int_{t_3}^{T} \frac{t^2}{t-\beta} \epsilon(t) dt = 0,$$

hence

$$\lim_{T \longrightarrow +\infty} g(x(T)) = \min g.$$

□

Remark 3.2 One can easily notice that, in case $\beta > 0$, the fact that there exist $a > 1$ and $t_1 \geq t_0$ such that $\dot{\epsilon}(t) \leq -\frac{a\beta}{2}\epsilon^2(t)$ for every $t \geq t_1$ implies that $\int_{t_0}^{+\infty} \frac{\epsilon(t)}{t} dt < +\infty$.

The next theorem shows that, by strengthening the integrability condition $\int_{t_0}^{+\infty} \frac{\epsilon(t)}{t} dt < +\infty$ (which is actually required in both settings (a) and (b) of Theorem 3.1), a rate of $\mathcal{O}(1/t^2)$ ca be guaranteed for the convergence of $g(x(t))$ to $\min g$.

Theorem 3.3 *Let x be the unique global C^2-solution of (5). Assume that*

$$\int_{t_0}^{+\infty} t\epsilon(t) dt < +\infty$$

and that one of the following conditions is fulfilled:

(a) *there exist $a > 1$ and $t_1 \geq t_0$ such that*

$$\dot{\epsilon}(t) \leq -\frac{a\beta}{2}\epsilon^2(t) \text{ for every } t \geq t_1;$$

(b) *there exist $a > 0$ and $t_1 \geq t_0$ such that*

$$\epsilon(t) \leq \frac{a}{t} \text{ for every } t \geq t_1.$$

If $\alpha \geq 3$, then

$$g(x(t)) - \min g = \mathcal{O}\left(\frac{1}{t^2}\right).$$

In addition, if $\alpha > 3$, then the trajectory x is bounded and

$$t\left(g(x(t)) - \min g\right),\ t\|\dot{x}(t)\|^2,\ t\epsilon(t)\|x(t) - x^*\|^2,\ t\epsilon(t)\|x(t)\|^2,\ t^2\|\nabla g(x(t))\|^2 \in L^1([t_0, +\infty), \mathbb{R})$$

for every arbitrary $x^* \in \operatorname{argmin} g$.

Proof Let be $x^*\operatorname{argmin} g$ and $2 \le b \le \alpha - 1$ fixed. We will use the energy functional introduced in the proof of the previous theorem and some of the estimate we derived for it. We will treat again the situations $\alpha > 3$ and $\alpha = 3$ separately.

The case $\alpha > 3$ and $2 < b < \alpha - 1$. As we already noticed in the proof of Theorem 3.1, according to (16), in case (a), and to (20), in case (b), we have

$$\dot{\mathcal{E}}_b(t) \le lt\frac{\epsilon(t)}{2}\|x^*\|^2 \text{ for every } t \ge t_2,$$

where $l := b$ and $t_2 = \max\left(t_1, \frac{2a\beta}{a-1}, \frac{\beta(\alpha-2)}{b-2}\right)$, in case (a), and $l := b + a\beta$ and $t_2 = \max\left(t_1, 4\beta, \frac{\beta(\alpha-2)}{b-2}\right)$ in case (b).

Using that $t\epsilon(t) \in L^1([t_0, +\infty), \mathbb{R})$ and that $t \mapsto \mathcal{E}_b(t)$ is bounded from below, from Lemma A.2 it follows that the limit $\lim_{t\longrightarrow+\infty} \mathcal{E}_b(t)$ exists. Consequently, $t \mapsto \mathcal{E}_b(t)$ is bounded, which implies that there exist $K > 0$ and $t' \ge t_0$ such that

$$0 \le g(x(t)) - \min g \le \frac{K}{t^2} \text{ for every } t \ge t'.$$

In addition, the function $t \mapsto \|x(t) - x^*\|^2$ is bounded, hence the trajectory x is bounded. Since $t \mapsto \|b(x(t) - x^*) + t(\dot{x}(t) + \beta\nabla g(x(t)))\|^2$ is also bounded, the inequality

$$\|t(\dot{x}(t) + \beta\nabla g(x(t)))\|^2 \le 2\|b(x(t) - x^*) + t(\dot{x}(t) + \beta\nabla g(x(t)))\|^2 + 2b^2\|x(t) - x^*\|^2,$$

which is true for every $t \ge t_0$, leads to

$$\|\dot{x}(t) + \beta\nabla g(x(t))\| = \mathcal{O}\left(\frac{1}{t}\right).$$

By integrating relation (16), in case (a), and relation (20), in case (b), on an interval $[t_2, s]$, where $s \ge t_3$ is arbitrarily chosen, and by letting afterwards s converge to $+\infty$, we obtain

$$t\left(g(x(t)) - \min g\right),\ t\|\dot{x}(t)\|^2,\ t\epsilon(t)\|x(t) - x^*\|^2, t^2\|\nabla g(x(t))\|^2 \in L^1([t_0, +\infty), \mathbb{R}).$$

The boundedness of the trajectory and the condition on the Tikhonov parametrization guarantee that

$$t\epsilon(t)\|x(t)\|^2 \in L^1([t_0, +\infty), \mathbb{R}).$$

The case $\alpha = 3$ and $b = 2$. As we already noticed in the proof of Theorem 3.1, according to (22), in case (a), and to (24), in case (b), we obtain

$$\frac{d}{dt}\left(\frac{t}{t-\beta}\mathcal{E}_2(t)\right) \leq l\frac{t^2}{t-\beta}\epsilon(t)\|x^*\|^2 \text{ for every } t \geq t_3,$$

where $l = 1$ and $t_3 = \max\left(t_1, \frac{\beta(\alpha-1)}{b-2}\right)$, in case (a), and $l = \frac{2+a\beta}{2}$ and $t_3 = \max(t_1, 4\beta)$ in case (b).

Since $t\epsilon(t) \in L^1([t_0, +\infty), \mathbb{R})$ and $\epsilon(t)$ is nonnegative, obviously $\frac{t^2}{t-\beta}\epsilon(t)\|x^*\|^2 \in L^1([t_2, +\infty), \mathbb{R})$. Using that $t \mapsto \frac{t}{t-\beta}\mathcal{E}_2(t)$ is bounded from below, from Lemma A.2 it follows that the limit $\lim_{t\longrightarrow+\infty} \frac{t}{t-\beta}\mathcal{E}_2(t)$ exists. Consequently, the limit $\lim_{t\longrightarrow+\infty} \mathcal{E}_2(t)$ also exists and $t \mapsto \mathcal{E}_2(t)$ is bounded. This implies that there exist $K > 0$ and $t' \geq t_0$ such that

$$0 \leq g(x(t)) - \min g \leq \frac{K}{t^2} \text{ for every } t \geq t'.$$

□

The next result shows that the statements of Theorem 3.3 can be strengthened in case $\alpha > 3$.

Theorem 3.4 *Let x be the unique global C^2-solution of (5). Assume that*

$$\int_{t_0}^{+\infty} t\epsilon(t)dt < +\infty$$

and that one of the following conditions is fulfilled:

(a) *there exist $a > 1$ and $t_1 \geq t_0$ such that*

$$\dot{\epsilon}(t) \leq -\frac{a\beta}{2}\epsilon^2(t) \textit{ for every } t \geq t_1;$$

(b) *there exist $a > 0$ and $t_1 \geq t_0$ such that*

$$\epsilon(t) \leq \frac{a}{t} \textit{ for every } t \geq t_1.$$

Let be an arbitrary $x^* \in \operatorname{argmin} g$. If $\alpha > 3$, then

$$t\langle\nabla g(x(t)), x(t) - x^*\rangle \in L^1([t_0, +\infty), \mathbb{R})$$

and the limits

$$\lim_{t\longrightarrow+\infty} \|x(t) - x^*\| \in \mathbb{R} \textit{ and } \lim_{t\longrightarrow+\infty} t\langle\dot{x}(t) + \beta\nabla g(x(t)), x(t) - x^*\rangle \in \mathbb{R}$$

exist. In addition,

$$g(x(t)) - \min g = o\left(\frac{1}{t^2}\right),\ \|\dot{x}(t) + \beta\nabla g(x(t))\| = o\left(\frac{1}{t}\right) \text{ and } \lim_{t\longrightarrow+\infty} t^2\epsilon(t)\|x(t)\|^2 = 0.$$

Proof Since $\alpha > 3$ we can choose $2 < b < \alpha - 1$. From (10) and (11) we have that

$$\begin{aligned}\dot{\mathcal{E}}_b(t) = (2t - \beta(b+2-\alpha))\,(g(x(t)) - \min g) + \left(t^2\frac{\dot{\epsilon}(t)}{2} + t\epsilon(t)\right)\|x(t)\|^2 \\ +(b+1-\alpha)t\|\dot{x}(t)\|^2 + (\beta^2 t - \beta t^2)\|\nabla g(x(t))\|^2 - \beta\epsilon(t)t^2\langle\nabla g(x(t)), x(t)\rangle \\ - bt\left\langle\left(1 - \frac{\beta}{t}\right)\nabla g(x(t)) + \epsilon(t)x(t), x(t) - x^*\right\rangle \text{ for every } t \geq t_0. \qquad (25)\end{aligned}$$

We will address the settings provided by the conditions (a) and (b) separately.

Condition (a) *holds:* In this case we estimate $-\beta\epsilon(t)t^2\langle\nabla g(x(t)), x(t)\rangle$ just as in (14) and from (25) we obtain

$$\begin{aligned}\dot{\mathcal{E}}_b(t) \leq (2t - \beta(b+2-\alpha))\,(g(x(t)) - \min g) + \left(t^2\frac{\dot{\epsilon}(t)}{2} + t\epsilon(t) + \frac{a\beta\epsilon^2(t)t^2}{4}\right)\|x(t)\|^2 \\ + (b+1-\alpha)t\|\dot{x}(t)\|^2 + \left(\beta^2 t - \beta\left(1 - \frac{1}{a}\right)t^2\right)\|\nabla g(x(t))\|^2 \\ - bt\left\langle\left(1 - \frac{\beta}{t}\right)\nabla g(x(t)) + \epsilon(t)x(t), x(t) - x^*\right\rangle \text{ for every } t \geq t_0. \qquad (26)\end{aligned}$$

We define $t_2 := \max\left(\beta, t_1, \frac{\beta a}{a-1}\right)$. By using condition (a), neglecting the nonpositive terms and afterwards integrating on the interval $[t_2, t]$, with arbitrary $t \geq t_2$, we obtain

$$\begin{aligned}\int_{t_2}^t bs\left\langle\left(1 - \frac{\beta}{s}\right)\nabla g(x(s)), x(s) - x^*\right\rangle \leq \mathcal{E}_b(t_2) - \mathcal{E}_b(t) + \int_{t_2}^t (2s - \beta(b+2-\alpha))\,(g(x(s)) - \min g)\,ds \\ - \int_{t_2}^t bs\left(1 - \frac{\beta}{s}\right)\langle\epsilon(s)x(s), x(s) - x^*\rangle + \int_{t_2}^t s\epsilon(s)\|x(s)\|^2 ds. \qquad (27)\end{aligned}$$

For every $s \geq t_2$, by the monotonicity of ∇g, we have $\langle\nabla g(x(s)), x(s) - x^*\rangle \geq 0$. Further, it holds

$$bs\left(1 - \frac{\beta}{s}\right)\epsilon(s)\left|\langle x(s), x(s) - x^*\rangle\right| \leq \left(1 - \frac{\beta}{s}\right)\frac{bs\epsilon(s)}{2}(\|x(s)\|^2 + \|x(s) - x^*\|^2).$$

By letting in (27) s converge to $+\infty$ and by taking into account that, according to Theorem 3.3,

$$\begin{aligned}t\epsilon(t)\|x(t)\|^2,\ t\epsilon(t)\|x(t) - x^*\|^2,\ (2t - \beta(b+2-\alpha))\left(g(x(t)) - g^*\right) \in L^1([t_0, +\infty), \mathbb{R}) \\ t\langle\nabla g(x(t)), x(t) - x^*\rangle \in L^1([t_0, +\infty), \mathbb{R}). \qquad (28)\end{aligned}$$

Condition (b) *holds:* In this case we estimate $-\beta\epsilon(t)t^2\langle\nabla g(x(t)), x(t)\rangle$ just as in (17) and from (25) we obtain

$$\begin{aligned}\dot{\mathcal{E}}_b(t) &\le (2t-\beta(b+2-\alpha))\,(g(x(t))-\min g)+\left(t^2\frac{\dot\epsilon(t)}{2}+t\epsilon(t)\right)\|x(t)\|^2\\&+(b+1-\alpha)t\|\dot x(t)\|^2+\left(\beta^2t-\beta t^2+\frac{\beta\epsilon(t)t^3}{2a}\right)\|\nabla g(x(t))\|^2+\frac{a_1\beta\epsilon(t)t}{2}\|x^*\|^2\\&-bt\left\langle\left(1-\frac{\beta}{t}\right)\nabla g(x(t))+\epsilon(t)x(t), x(t)-x^*\right\rangle \text{ for every } t\ge t_0. \qquad (29)\end{aligned}$$

We define $t_2 := \max(4\beta, t_1)$. According to (19) we have that $\beta^2t-\beta t^2+\frac{\beta\epsilon(t)t^3}{2a_1}\le 0$ for every $t\ge t_2$. By using condition (b), neglecting the nonpositive terms and afterwards integrating on the interval $[t_2, t]$, with arbitrary $t\ge t_2$, we obtain

$$\begin{aligned}\int_{t_2}^t bs\left\langle\left(1-\frac{\beta}{s}\right)\nabla g(x(s)), x(s)-x^*\right\rangle &\le \mathcal{E}_b(t_2)-\mathcal{E}_b(t)+\int_{t_2}^t(2s-\beta(b+2-\alpha))\,(g(x(s))-\min g)\,ds\\&-\int_{t_2}^t bs\left(1-\frac{\beta}{s}\right)\langle\epsilon(s)x(s), x(s)-x^*\rangle+\int_{t_2}^t s\epsilon(s)\|x(s)\|^2ds\\&+\frac{a\beta}{2}\|x^*\|^2\int_{t_2}^t s\epsilon(s)ds. \qquad (30)\end{aligned}$$

From here, by using the similar arguments as for the case (a), we obtain (28).

Consider now, $b_1, b_2\in(2,\alpha-1),\ b_1\ne b_2$. Then for every $t\ge t_0$ we have

$$\mathcal{E}_{b_1}(t)-\mathcal{E}_{b_2}(t)=(b_1-b_2)\big(-\beta t(g(x(t))-\min g)+t\langle\dot x(t)+\beta\nabla g(x(t)), x(t)-x^*\rangle\\+\frac{\alpha-1}{2}\|x(t)-x^*\|^2\Big).$$

According to Theorem 3.3, the limits

$$\lim_{t\longrightarrow+\infty}(\mathcal{E}_{b_1}(t)-\mathcal{E}_{b_2}(t))\in\mathbb{R} \text{ and } \lim_{t\longrightarrow+\infty}t(g(x(t))-g^*)\in\mathbb{R}$$

exist, consequently, the limit

$$\lim_{t\longrightarrow+\infty}\left(t\langle\dot x(t)+\beta\nabla g(x(t)), x(t)-x^*\rangle+\frac{\alpha-1}{2}\|x(t)-x^*\|^2\right)$$

also exists. For every $t\ge t_0$ we define

$$k(t)=t\langle\dot x(t)+\beta\nabla g(x(t)), x(t)-x^*\rangle+\frac{\alpha-1}{2}\|x(t)-x^*\|^2$$

and

$$q(t)=\frac{1}{2}\|x(t)-x^*\|^2+\beta\int_{t_0}^t\langle\nabla g(x(s)), x(s)-x^*\rangle ds.$$

Then

$$(\alpha - 1)q(t) + t\dot{q}(t) = k(t) + \beta(\alpha - 1)\int_{t_0}^{t} \langle \nabla g(x(s)), x(s) - x^* \rangle ds \text{ for every } t \geq t_0.$$

From (28) and the fact that $k(t)$ has a limit whenever $t \longrightarrow +\infty$, we obtain that $(\alpha - 1)q(t) + t\dot{q}(t)$ has a limit when $t \longrightarrow +\infty$. According to Lemma 4.6, $q(t)$ has a limit when $t \longrightarrow +\infty$. By using (28) again we obtain that the limit

$$\lim_{t \longrightarrow +\infty} \|x(t) - x^*\| \in \mathbb{R}$$

exists and, consequently, the limit

$$\lim_{t \longrightarrow +\infty} t\langle \dot{x}(t) + \beta \nabla g(x(t)), x(t) - x^* \rangle \in \mathbb{R}$$

also exists. On the other hand, we notice that for every $t \geq t_0$ the energy functional can be written as

$$\begin{aligned} \mathcal{E}_b(t) &= (t^2 - \beta(b + 2 - \alpha)t)\,(g(x(t)) - \min g) + \frac{t^2\epsilon(t)}{2}\|x(t)\|^2 \\ &\quad + \frac{t^2}{2}\|\dot{x}(t) + \beta\nabla g(x(t))\|^2 + bt\langle \dot{x}(t) + \beta\nabla g(x(t)), x(t) - x^* \rangle + \frac{b(\alpha - 1)}{2}\|x(t) - x^*\|^2. \end{aligned} \tag{31}$$

Since the limits

$$\lim_{t \longrightarrow +\infty} \mathcal{E}_b(t) \in \mathbb{R} \text{ and } \lim_{t \longrightarrow +\infty} \left(bt\langle \dot{x}(t) + \beta\nabla g(x(t)), x(t) - x^* \rangle + \frac{b(\alpha - 1)}{2}\|x(t) - x^*\|^2 \right) \in \mathbb{R}$$

exist, it follows that the limit

$$\lim_{t \longrightarrow +\infty} \left((t^2 - \beta(b + 2 - \alpha)t)\,(g(x(t)) - \min g) + \frac{t^2\epsilon(t)}{2}\|x(t)\|^2 + \frac{t^2}{2}\|\dot{x}(t) + \beta\nabla g(x(t))\|^2 \right) \in \mathbb{R}$$

exists, too.

We define

$$\varphi : [t_0, +\infty) \longrightarrow \mathbb{R},\ \varphi(t) = (t^2 - \beta(b + 2 - \alpha)t)\left(g(x(t)) - g^*\right) + \frac{t^2\epsilon(t)}{2}\|x(t)\|^2 + \frac{t^2}{2}\|\dot{x}(t) + \beta\nabla g(x(t))\|^2,$$

and notice that for sufficiently large t it holds

$$0 \leq \frac{\varphi(t)}{t} \leq 2t\,(g(x(t)) - \min g) + \frac{t\epsilon(t)}{2}\|x(t)\|^2 + \frac{t}{2}\|\dot{x}(t) + \beta\nabla g(x(t))\|^2.$$

According to Theorem 3.3 the right hand side of the above inequality is of class $L^1([t_0, +\infty), \mathbb{R})$.

Hence,

$$\frac{\varphi(t)}{t} \in L^1([t_0, +\infty), \mathbb{R}).$$

Since $\frac{1}{t} \notin L^1([t_0, +\infty), \mathbb{R})$ and the limit $\lim_{t \longrightarrow +\infty} \varphi(t) \in \mathbb{R}$ exists, it must hold that $\lim_{t \longrightarrow +\infty} \varphi(t) = 0$. Consequently,

$$\lim_{t \longrightarrow +\infty} (t^2 - \beta(b+2-\alpha)t)\,(g(x(t)) - \min g) = \lim_{t \longrightarrow +\infty} \frac{t^2\epsilon(t)}{2}\|x(t)\|^2 = \lim_{t \longrightarrow +\infty} \frac{t^2}{2}\|\dot{x}(t) + \beta\nabla g(x(t))\|^2 = 0$$

and the proof is complete. □

Working in the hypotheses of Theorem 3.4 we can prove also the weak convergence of the trajectories generated by (5) to a minimizer of the objective function g.

Theorem 3.5 *Let x be the unique global C^2-solution of* (5). *Assume that*

$$\int_{t_0}^{+\infty} t\epsilon(t)dt < +\infty$$

and that one of the following conditions is fulfilled:

(a) *there exist $a > 1$ and $t_1 \geq t_0$ such that*

$$\dot{\epsilon}(t) \leq -\frac{a\beta}{2}\epsilon^2(t) \text{ for every } t \geq t_1;$$

(b) *there exist $a > 0$ and $t_1 \geq t_0$ such that*

$$\epsilon(t) \leq \frac{a}{t} \text{ for every } t \geq t_1.$$

If $\alpha > 3$, then $x(t)$ converges weakly to an element in $\operatorname{argmin} g$ as $t \longrightarrow +\infty$.

Proof We will to apply the continuous version of the Opial Lemma (Lemma A.3) for $S = \operatorname{argmin} g$. According to Theorem 3.4, the limit

$$\lim_{t \longrightarrow +\infty} \|x(t) - x^*\| \in \mathbb{R}$$

exists for every $x^* \in \operatorname{argmin} g$.

Further, let $\overline{x} \in \mathcal{H}$ be a weak sequential limit point of $x(t)$. This means that there exists a sequence $(t_n)_{n\in\mathbb{N}} \subseteq [t_0, +\infty)$ such that $\lim_{n\longrightarrow\infty} t_n = +\infty$ and $x(t_n)$ converges weakly to $\overline{x}$ as $n \longrightarrow \infty$. Since g is weakly lower semicontinuous, we have that

$$g(\overline{x}) \leq \liminf_{n \longrightarrow +\infty} g(x(t_n)).$$

On the other hand, according to Theorem 3.3,

$$\lim_{t\longrightarrow+\infty} g(x(t)) = \min g,$$

consequently one has $g(\overline{x}) \leq \min g$, which shows that $\overline{x} \in \text{argmin} g$.

The convergence of the trajectory is a consequence of Lemma A.3. □

Remark 3.6 We proved in this section that the convergence rate of $o\left(\frac{1}{t^2}\right)$ for $g(x(t))$, the converge rate of $o\left(\frac{1}{t}\right)$ for $\|\dot{x}(t) + \beta\nabla g(x(t))\|$ and the weak convergence of the trajectory to a minimizer of g that have been obtained in [11] for the dynamical system with Hessian driven damping (3) are preserved when this system is enhanced with a Tikhonov regularization term. In addition, in the case when the Hessian driven damping term is removed, which is the case when $\beta = 0$, we recover the results provided in [8] for the dynamical system (4) with Tikhonov regularization term. In this setting, we have to assume in Theorem 3.1 just that $\int_{t_0}^{+\infty} \frac{\epsilon(t)}{t} dt < +\infty$, and in the theorems 3.3 - 3.5 just that $\int_{t_0}^{+\infty} t\epsilon(t)dt < +\infty$, since condition (a) is automatically fulfilled.

4 Strong convergence to the minimum norm solution

In this section we will continue the investigations we did at the end of Section 3, by working in the same setting, on the behaviour of the trajectory of the dynamical system (5) by concentrating on strong convergence. In particular, we will provide conditions on the Tikhonov parametrization $t \mapsto \epsilon(t)$ which will guarantee that the trajectory converges to a minimum norm solution of g, which is the element of minimum norm of the nonempty convex closed set argmin$g$. We start with the following result.

Lemma 4.1 *Let x be the unique global C^2-solution of (5). For $x^* \in$ argmin$g$ we introduce the function*

$$h_{x^*} : [t_0, +\infty) \longrightarrow \mathbb{R}\ h_{x^*}(t) = \frac{1}{2}\|x(t) - x^*\|^2.$$

If $\alpha > 0$ and $\beta \geq 0$, then

$$\sup_{t\geq t_0} \|\dot{x}(t)\| < +\infty \text{ and } \frac{1}{t}\|\dot{x}(t)\|^2 \in L^1([t_0, +\infty), \mathbb{R}).$$

In addition,

$$\sup_{t\geq t_0} \frac{1}{t}|\dot{h}_{x^*}(t)| < +\infty.$$

Proof We consider the following energy functional

$$W : [t_0, +\infty) \to \mathbb{R},\ W(t) = g(x(t)) + \frac{1}{2}\|\dot{x}(t)\|^2 + \frac{\epsilon(t)}{2}\|x(t)\|^2. \tag{32}$$

By using (5) we have for every $t \geq t_0$

$$\begin{aligned}\dot{W}(t) &= \langle \nabla g(x(t), \dot{x}(t)\rangle + \langle \ddot{x}(t), \dot{x}(t)\rangle + \frac{\dot{\epsilon}(t)}{2}\|x(t)\|^2 + \epsilon(t)\langle \dot{x}(t), x(t)\rangle \\ &= \langle \nabla g(x(t), \dot{x}(t)\rangle + \frac{\dot{\epsilon}(t)}{2}\|x(t)\|^2 + \epsilon(t)\langle \dot{x}(t), x(t)\rangle \\ &\quad + \left\langle -\frac{\alpha}{t}\dot{x}(t) - \beta\nabla^2 g(x(t))\dot{x}(t) - \nabla g(x(t)) - \epsilon(t)x(t), \dot{x}(t)\right\rangle \\ &= -\frac{\alpha}{t}\|\dot{x}(t)\|^2 + \frac{\dot{\epsilon}(t)}{2}\|x(t)\|^2 - \beta\langle \nabla^2 g(x(t))\dot{x}(t), \dot{x}(t)\rangle.\end{aligned}$$

From here, invoking the convexity of g, it follows

$$\dot{W}(t) \leq -\frac{\alpha}{t}\|\dot{x}(t)\|^2 + \frac{\dot{\epsilon}(t)}{2}\|x(t)\|^2, \tag{33}$$

for every $t \geq t_0$. Since ϵ is nonincreasing this leads further to

$$\dot{W}(t) \leq -\frac{\alpha}{t}\|\dot{x}(t)\|^2 \text{ for every } t \geq t_0, \tag{34}$$

therefore the energy W is nonincreasing. Since W is bounded from bellow, there exists $\lim_{t\longrightarrow+\infty} W(t) \in \mathbb{R}$. Consequently, $t \mapsto W(t)$ is bounded on $[t_0, +\infty)$ from which, since g is bounded from bellow, we obtain that

$$\sup_{t\geq t_0}\|\dot{x}(t)\| = K < +\infty.$$

By integrating (34) on an interval $[t_0, t]$ for arbitrary $t > t_0$ it yields

$$\int_{t_0}^{t}\frac{\alpha}{s}\|\dot{x}(s)\|^2 ds \leq W(t_0) - W(t),$$

which, by letting $t \longrightarrow +\infty$, leads to

$$\frac{1}{t}\|\dot{x}(t)\|^2 \in L^1([t_0, +\infty), \mathbb{R}).$$

Further, for every $t \geq t_0$ we have that

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

and

$$\||x(t) - x^*\| \leq \|x(t) - x(t_0)\| + \|x(t_0) - x^*\| \leq \sup_{t\geq t_0}\|\dot{x}(t)\|(t - t_0) + \|x(t_0) - x^*\|,$$

hence,

$$\frac{1}{t}|\dot{h}_{x^*}(t)| \leq \sup_{t\geq t_0}\|\dot{x}(t)\|\left(\sup_{t\geq t_0}\|\dot{x}(t)\|\left(1-\frac{t_0}{t}\right)+\frac{1}{t}\|x(t_0)-x^*\|\right)$$
$$\leq \sup_{t\geq t_0}\|\dot{x}(t)\|\left(\sup_{t\geq t_0}\|\dot{x}(t)\|+\frac{1}{t_0}\|x(t_0)-x^*\|\right)\in\mathbb{R}.$$

□

For each $\epsilon > 0$, we denote by x_ϵ the unique solution of the strongly convex minimization problem

$$x_\epsilon = \underset{x\in\mathcal{H}}{\operatorname{argmin}}\left(g(x)+\frac{\epsilon}{2}\|x\|^2\right).$$

In virtue of the Fermat rule, this is equivalent to

$$\nabla g(x_\epsilon)+\epsilon x_\epsilon = 0.$$

It is well known that the Tikhonov approximation curve $\epsilon \longrightarrow x_\epsilon$ satisfies $\lim_{\epsilon\longrightarrow 0} x_\epsilon = x^*$, where $x^* = \operatorname{argmin}\{\|x\| : x \in \operatorname{argmin} g\}$ is the element of minimum norm of the nonempty convex closed set $\operatorname{argmin} g$. Since ∇g is monotone, for every $\epsilon > 0$ it holds $\langle \nabla g(x_\epsilon) - \nabla g(x^*), x_\epsilon - x^*\rangle \geq 0$, that is $\langle -\epsilon x_\epsilon, x_\epsilon - x^*\rangle \geq 0$. Hence, $-\|x_\epsilon\|^2 + \langle x_\epsilon, x^*\rangle \geq 0$, which, by using the Cauchy-Schwarz inequality, implies

$$\|x_\epsilon\| \leq \|x^*\| \text{ for every } \epsilon > 0.$$

4.1 Strong ergodic convergence

We will start by proving a strong ergodic convergence result for the trajectory of (5).

Theorem 4.2 *Let x be the unique global C^2-solution of (5). Assume that*

$$\int_{t_0}^{+\infty}\frac{\epsilon(t)}{t}dt = +\infty.$$

Let $x^* = \operatorname{argmin}\{\|x\| : x \in \operatorname{argmin} g\}$ be the element of minimum norm of the nonempty convex closed set $\operatorname{argmin} g$. If $\alpha > 0$, then

$$\lim_{t\longrightarrow+\infty}\frac{1}{\int_{t_0}^{t}\frac{\epsilon(s)}{s}ds}\int_{t_0}^{t}\frac{\epsilon(s)}{s}\|x(s)-x^*\|^2ds = 0 \text{ and } \liminf_{t\longrightarrow+\infty}\|x(t)-x^*\| = 0.$$

Proof We introduce the function

$$h_{x^*} : [t_0, +\infty) \longrightarrow \mathbb{R},\ h_{x^*}(t) = \frac{1}{2}\|x(t)-x^*\|^2.$$

For every $t \geq t_0$ we have

$$\ddot{h}_{x^*}(t) + \frac{\alpha}{t}\dot{h}_{x^*}(t) = \|\dot{x}(t)\|^2 + \left\langle \ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t), x(t) - x^* \right\rangle. \tag{35}$$

Further, for every $t \geq t_0$, the function $g_t : \mathcal{H} \longrightarrow \mathbb{R}$, $g_t(x) = g(x) + \frac{\epsilon(t)}{2}\|x\|^2$, is strongly convex, with modulus $\epsilon(t)$, hence

$$g_t(x^*) - g_t(x(t)) \geq \langle \nabla g_t(x(t)), x^* - x(t) \rangle + \frac{\epsilon(t)}{2}\|x(t) - x^*\|^2. \tag{36}$$

But $\nabla g_t(x(t)) = \nabla g(x(t)) + \epsilon(t)x(t)$ and by using (5) we get

$$\nabla g_t(x(t)) = -\ddot{x}(t) - \frac{\alpha}{t}\dot{x}(t) - \beta\nabla^2 g(x(t))\dot{x}(t) \text{ for every } t \geq t_0.$$

Consequently, (36) becomes

$$g_t(x^*) - g_t(x(t)) \geq \left\langle \ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \beta\nabla^2 g(x(t))\dot{x}(t), x(t) - x^* \right\rangle + \frac{\epsilon(t)}{2}\|x(t) - x^*\|^2 \text{ for every } t \geq t_0. \tag{37}$$

By using (35), the latter relation leads to

$$g_t(x^*) - g_t(x(t)) \geq \ddot{h}_{x^*}(t) + \frac{\alpha}{t}\dot{h}_{x^*}(t) + \epsilon(t)h_{x^*}(t) + \langle \beta\nabla^2 g(x(t))\dot{x}(t), x(t) - x^* \rangle - \|\dot{x}(t)\|^2 \tag{38}$$

for every $t \geq t_0$.

For every $t \geq t_0$, let $x_{\epsilon(t)}$ the unique solution of the strongly convex minimization problem

$$\min_{x \in \mathcal{H}} \left(g(x) + \frac{\epsilon(t)}{2}\|x\|^2 \right).$$

Then

$$g_t(x^*) - g_t(x(t)) \leq g_t(x^*) - g_t(x_{\epsilon(t)}) = g(x^*) + \frac{\epsilon(t)}{2}\|x^*\|^2 - g(x_{\epsilon(t)}) - \frac{\epsilon(t)}{2}\|x_{\epsilon(t)}\|^2 \leq \frac{\epsilon(t)}{2}(\|x^*\|^2 - \|x_{\epsilon(t)}\|^2)$$

for every $t \geq t_0$ and taking into account (38) we get

$$\frac{\epsilon(t)}{2}(\|x^*\|^2 - \|x_{\epsilon(t)}\|^2) \geq \ddot{h}_{x^*}(t) + \frac{\alpha}{t}\dot{h}_{x^*}(t) + \epsilon(t)h_{x^*}(t) + \langle \beta\nabla^2 g(x(t))\dot{x}(t), x(t) - x^* \rangle - \|\dot{x}(t)\|^2 \tag{39}$$

for every $t \geq t_0$. We have

$$\ddot{h}_{x^*}(t) + \frac{\alpha}{t}\dot{h}_{x^*}(t) = \frac{1}{t^\alpha}\frac{d}{dt}\left(t^\alpha \dot{h}_{x^*}(t)\right)$$

and

$$\langle \nabla^2 g(x(t))\dot{x}(t), x(t) - x^* \rangle = \frac{d}{dt}\big(\langle \nabla g(x(t)), x(t) - x^* \rangle - g(x(t))\big)$$

hence (39) is equivalent to

$$\frac{\epsilon(t)}{t}\left(h_{x^*}(t)-\frac{1}{2}(\|x^*\|^2-\|x_{\epsilon(t)}\|^2)\right)\leq\frac{1}{t}\|\dot{x}(t)\|^2-\frac{1}{t^{\alpha+1}}\frac{d}{dt}(t^\alpha\dot{h}_{x^*}(t))-\frac{\beta}{t}\frac{d}{dt}(\langle\nabla g(x(t)),x(t)-x^*\rangle-g(x(t))), \tag{40}$$

for every $t\geq t_0$.

After integrating (40) on $[t_0,t]$, for arbitrary $t>t_0$, it yields

$$\int_{t_0}^t\frac{\epsilon(s)}{s}\left(h_{x^*}(s)-\frac{1}{2}(\|x^*\|^2-\|x_{\epsilon(s)}\|^2)\right)ds\leq\int_{t_0}^t\left(\frac{1}{s}\|\dot{x}(s)\|^2-\frac{1}{s^{\alpha+1}}\frac{d}{ds}\left(s^\alpha\dot{h}_{x^*}(s)\right)\right)ds$$
$$+\int_{t_0}^t\frac{\beta}{s}\frac{d}{ds}\left(\langle\nabla g(x(s)),x^*-x(s)\rangle+g(x(s))\right)ds. \tag{41}$$

We show that the right-hand side of the above inequality is bounded from above. Indeed, according to Lemma 4.1, one has

$$\frac{1}{t}\|\dot{x}(t)\|^2\in L^1([t_0,+\infty),\mathbb{R}),$$

hence there exists $C_1\geq 0$ such that $\int_{t_0}^t\frac{1}{s}\|\dot{x}(s)\|^2\leq C_1$ for every $t\geq t_0$. Further, for every $t\geq t_0$,

$$\begin{aligned}\int_{t_0}^t\frac{1}{s^{\alpha+1}}\frac{d}{ds}(s^\alpha\dot{h}_{x^*}(s))ds&=\frac{\dot{h}_{x^*}(t)}{t}-\frac{\dot{h}_{x^*}(t_0)}{t_0}+(\alpha+1)\int_{t_0}^t\frac{\dot{h}_{x^*}(s)}{s^2}ds\\&=\frac{\dot{h}_{x^*}(t)}{t}-\frac{\dot{h}_{x^*}(t_0)}{t_0}+(\alpha+1)\left(\frac{h_{x^*}(t)}{t^2}-\frac{h_{x^*}(t_0)}{t_0^2}\right)+2(\alpha+1)\int_{t_0}^t\frac{h_{x^*}(s)}{s^3}ds\\&\geq\frac{\dot{h}_{x^*}(t)}{t}-C_2,\end{aligned}$$

where $C_2=\frac{\dot{h}_{x^*}(t_0)}{t_0}+(\alpha+1)\frac{h_{x^*}(t_0)}{t_0^2}$. Consequently,

$$\int_{t_0}^t\frac{\epsilon(s)}{s}\left(h_{x^*}(s)-\frac{1}{2}(\|x^*\|^2-\|x_{\epsilon(s)}\|^2)\right)ds\leq C_1+C_2-\frac{\dot{h}_{x^*}(t)}{t}$$
$$+\int_{t_0}^t\frac{\beta}{s}\frac{d}{ds}(\langle\nabla g(x(s)),x^*-x(s)\rangle+g(x(s)))ds, \tag{42}$$

for every $t\geq t_0$. According to Lemma 4.1, there exists C_3 such that $\frac{1}{t}|\dot{h}_{x^*}(t)|\leq C_3$ for all $t\geq t_0$, which combined with (42) guarantees the existence of $C_4\geq 0$ such that

$$\int_{t_0}^t\frac{\epsilon(s)}{s}\left(h_{x^*}(s)-\frac{1}{2}(\|x^*\|^2-\|x_{\epsilon(s)}\|^2)\right)ds\leq C_4+\int_{t_0}^t\frac{\beta}{s}\frac{d}{ds}\left(\langle\nabla g(x(s)),x^*-x(s)\rangle+g(x(s))\right)ds \tag{43}$$

for every $t\geq t_0$.

On the other hand, for every $t \geq t_0$,

$$\int_{t_0}^{t} \frac{\beta}{s} \frac{d}{ds} (\langle \nabla g(x(s)), x^* - x(s) \rangle + g(x(s))) ds = \int_{t_0}^{t} \frac{\beta}{s^2} \big(\langle \nabla g(x(s)), x^* - x(s) \rangle + g(x(s))\big) ds + \frac{\beta}{t} \big(\langle \nabla g(x(t)), x^* - x(t) \rangle + g(x(t))\big) - \frac{\beta}{t_0} \big(\langle \nabla g(x(t_0)), x^* - x(t_0) \rangle + g(x(t_0))\big).$$

From the gradient inequality of the convex function g we have

$$\langle \nabla g(x(t)), x^* - x(t) \rangle + g(x(t)) \leq g(x^*),$$

hence

$$\int_{t_0}^{t} \frac{\beta}{s} \frac{d}{ds} (\langle \nabla g(x(s)), x^* - x(s) \rangle + g(x(s))) ds \leq \frac{\beta}{t} g(x^*) + \int_{t_0}^{t} \frac{\beta}{s^2} g(x^*) ds - \frac{\beta}{t_0} (\langle \nabla g(x(t_0)), x^* - x(t_0) \rangle + g(x(t_0))), \tag{44}$$

for all $t \geq t_0$. Obviously the right-hand side of (44) is bounded from above, hence there exists $C_5 > 0$ such that

$$\int_{t_0}^{t} \frac{\beta}{s} \frac{d}{ds} (\langle \nabla g(x(s)), x^* - x(s) \rangle + g(x(s))) ds \leq C_5 \text{ for every } t \geq t_0. \tag{45}$$

Combining (43) and (45) we obtain that there exists $C > 0$ such that

$$\int_{t_0}^{t} \frac{\epsilon(s)}{s} \left(h_{x^*}(s) - \frac{1}{2} (\|x^*\|^2 - \|x_{\epsilon(s)}\|^2) \right) ds \leq C \text{ for every } t \geq t_0. \tag{46}$$

Since $\lim_{t \longrightarrow +\infty} \epsilon(t) = 0$ we have $\lim_{t \longrightarrow +\infty} x_{\epsilon(t)} = x^*$, hence $\lim_{t \longrightarrow +\infty} (\|x^*\|^2 - \|x_{\epsilon(t)}\|^2) = 0$. Consequently, by using the l'Hospital rule and the fact that $\int_{t_0}^{+\infty} \frac{\epsilon(t)}{t} dt = +\infty$, we get

$$\lim_{t \longrightarrow +\infty} \frac{1}{\int_{t_0}^{t} \frac{\epsilon(s)}{s} ds} \int_{t_0}^{t} \frac{\epsilon(s)}{s} (\|x^*\|^2 - \|x_{\epsilon(s)}\|^2) ds = \lim_{t \longrightarrow +\infty} \frac{\frac{\epsilon(t)}{t} (\|x^*\|^2 - \|x_{\epsilon(t)}\|^2)}{\frac{\epsilon(t)}{t}} = \lim_{t \longrightarrow +\infty} (\|x^*\|^2 - \|x_{\epsilon(t)}\|^2) = 0.$$

Dividing (46) by $\int_{t_0}^{t} \frac{\epsilon(s)}{s} ds$ and taking into account that $\int_{t_0}^{+\infty} \frac{\epsilon(t)}{t} dt = +\infty$, we obtain that

$$\lim_{t \longrightarrow +\infty} \frac{1}{\int_{t_0}^{t} \frac{\epsilon(s)}{s} ds} \int_{t_0}^{t} \frac{\epsilon(s)}{s} \|x(s) - x^*\|^2 ds = 0.$$

The last equality immediately implies that

$$\liminf_{t \longrightarrow +\infty} \|x(t) - x^*\| = 0.$$

□

Remark 4.3 The strong ergodic convergence obtained in [8] for the dynamical system (4) is extended to the dynamical system with Hessian driven damping and Tikhonov regularization term (5) under the same hypotheses concerning the Tikhonov parametrization $t \mapsto \epsilon(t)$.

4.2 Strong convergence

In order to prove strong convergence for the trajectory generated by the dynamical system (5) to an element of minimum norm of $\operatorname{argmin} g$ we have to strengthen the conditions on the Tikhonov parametrization. This is done in the following result.

Theorem 4.4 *Let be $\alpha \geq 3$ and x the unique global C^2-solution of (5). Assume that*

$$\int_{t_0}^{+\infty} \frac{\epsilon(t)}{t} dt < +\infty \text{ and } \lim_{t \longrightarrow +\infty} \frac{\beta}{\epsilon(t) t^{\frac{\alpha}{3}+1}} \int_{t_0}^{t} \epsilon^2(s) s^{\frac{\alpha}{3}+1} ds = 0,$$

and that there exist $a > 1$ and $t_1 \geq t_0$ such that

$$\dot{\epsilon}(t) \leq -\frac{a\beta}{2} \epsilon^2(t) \text{ for every } t \geq t_1.$$

In addition, assume that

- *in case $\alpha = 3$: $\lim_{t \longrightarrow +\infty} t^2 \epsilon(t) = +\infty$;*
- *in case $\alpha > 3$: there exists $c > 0$ such that $t^2 \epsilon(t) \geq \frac{2}{3}\alpha \left(\frac{1}{3}\alpha - 1 + \beta c^2\right)$ for t large enough.*

If $x^* = \operatorname{argmin}\{\|x\| : x \in \operatorname{argmin} g\}$ is the element of minimum norm of the nonempty convex closed set $\operatorname{argmin} g$, then

$$\liminf_{t \longrightarrow +\infty} \|x(t) - x^*\| = 0.$$

In addition,

$$\lim_{t \longrightarrow +\infty} \|x(t) - x^*\| = 0,$$

if there exists $T \geq t_0$ such that the trajectory $\{x(t) : t \geq T\}$ stays either in the ball $B(0, \|x^*\|)$, or in its complement.

Proof *Case I* Assume that there exists $T \geq t_0$ such that the trajectory $\{x(t) : t \geq T\}$ stays in the complement of the ball $B(0, \|x^*\|)$.

In other words, $\|x(t)\| \geq \|x^*\|$ for every $t \geq T$. For $p \geq 0$, we consider the energy functional

$$\mathcal{E}_b^p(t) = t^{p+1}(t + \alpha - \beta - \beta p - b - 1)(g(x(t)) - \min g) + t^{p+2}\frac{\epsilon(t)}{2}(\|x(t)\|^2 - \|x^*\|^2) + \frac{t^p}{2}\|b(x(t) - x^*) + t(\dot{x}(t) + \beta\nabla g(x(t)))\|^2 \text{ for every } t \geq t_0. \tag{47}$$

We define $t_2 := \max\,(t_1, 2(\beta + \beta p + b + 1 - \alpha))$. We have that

$$\mathcal{E}_b^p(t) \geq t^{p+1}(t + \alpha - \beta - \beta p - b - 1)(g(x(t)) - \min g) + t^{p+2}\frac{\epsilon(t)}{2}(\|x(t)\|^2 - \|x^*\|^2) \geq t^{p+2}\frac{1}{2}(g(x(t)) - \min g) + t^{p+2}\frac{\epsilon(t)}{2}(\|x(t)\|^2 - \|x^*\|^2) \text{ for every } t \geq t_2. \tag{48}$$

For every $t \geq t_0$ consider the strongly convex function

$$g_t : \mathcal{H} \longrightarrow \mathbb{R},\ \ g_t(x) = \frac{1}{2}g(x) + \frac{\epsilon(t)}{2}\|x\|^2,$$

and denote

$$x_{\epsilon(t)} := \underset{x \in \mathcal{H}}{\operatorname{argmin}}\, g_t(x).$$

Since x^* is the element of minimum norm in $\operatorname{argmin}\frac{1}{2}g = \operatorname{argmin} g$, it holds $\|x_{\epsilon(t)}\| \leq \|x^*\|$. Using the gradient inequality we have

$$g_t(x) - g_t(x_{\epsilon(t)}) \geq \frac{\epsilon(t)}{2}\|x - x_{\epsilon(t)}\|^2 \text{ for every } x \in \mathcal{H}.$$

On the other hand,

$$g_t(x_{\epsilon(t)}) - g_t(x^*) = \frac{1}{2}(g(x_{\epsilon(t)}) - \min g) + \frac{\epsilon(t)}{2}(\|x_{\epsilon(t)}\|^2 - \|x^*\|^2) \geq \frac{\epsilon(t)}{2}(\|x_{\epsilon(t)}\|^2 - \|x^*\|^2).$$

By adding the last two inequalities we obtain

$$g_t(x) - g_t(x^*) \geq \frac{\epsilon(t)}{2}(\|x - x_{\epsilon(t)}\|^2 + \|x_{\epsilon(t)}\|^2 - \|x^*\|^2) \text{ for every } x \in \mathcal{H}. \tag{49}$$

From (48) and (49) we have that for every $t \geq t_2$ it holds

$$\mathcal{E}_b^p(t) \geq t^{p+2}(g_t(x(t)) - g_t(x^*)) \geq \frac{\epsilon(t)}{2}t^{p+2}\big(\|x(t) - x_{\epsilon(t)}\|^2 + \|x_{\epsilon(t)}\|^2 - \|x^*\|^2\big). \tag{50}$$

The next step is to obtain an upper bound for $t \mapsto E_b^p(t)$, and to this end we will evaluate its time derivative. For every $t \geq t_0$ we have

$$\begin{aligned}
\frac{d}{dt}\mathcal{E}_b^p(t) &= t^p((p+2)t + (p+1)(\alpha - \beta - \beta p - b - 1))(g(x(t)) - \min g) \\
&+ t^{p+1}(t + \alpha - \beta - \beta p - b - 1)\langle \nabla g(x(t)), \dot{x}(t)\rangle \\
&+ \left((p+2)t^{p+1}\frac{\epsilon(t)}{2} + t^{p+2}\frac{\dot{\epsilon}(t)}{2}\right)(\|x(t)\|^2 - \|x^*\|^2) + t^{p+2}\epsilon(t))\langle \dot{x}(t), x(t)\rangle \\
&+ \frac{pt^{p-1}}{2}\|b(x(t) - x^*) + t(\dot{x}(t) + \beta\nabla g(x(t)))\|^2 \\
&+ t^p\langle (b+1)\dot{x}(t) + \beta\nabla g(x(t)) + t(\ddot{x}(t) + \beta\nabla^2 g(x(t))\dot{x}(t)), b(x(t) - x^*) \\
&+ t(\dot{x}(t) + \beta\nabla g(x(t)))\rangle. \qquad (51)
\end{aligned}$$

By using (5) we have

$$\ddot{x}(t) + \beta\nabla^2 g(x(t))\dot{x}(t) = -\frac{\alpha}{t}\dot{x}(t) - \nabla g(x(t)) - \epsilon(t)x(t),$$

hence

$$\begin{aligned}
&\langle (b+1)\dot{x}(t) + \beta\nabla g(x(t)) + t(\ddot{x}(t) + \beta\nabla^2 g(x(t))\dot{x}(t)), b(x(t) - x^*) + t(\dot{x}(t) + \beta\nabla g(x(t)))\rangle \\
&= \langle (b+1-\alpha)\dot{x}(t) + \beta\nabla g(x(t)) - t(\nabla g(x(t)) + \epsilon(t)x(t)), b(x(t) - x^*) + t(\dot{x}(t) + \beta\nabla g(x(t)))\rangle \\
&= b(b+1-\alpha)\langle \dot{x}(t), x(t) - x^*\rangle + (b+1-\alpha)t(\|\dot{x}(t)\|^2 + \langle \nabla g(x(t)), \dot{x}(t)\rangle) \\
&\quad + \beta b\langle \nabla g(x(t), x(t) - x^*\rangle + \beta t\langle \nabla g(x(t)), \dot{x}(t)\rangle + \beta^2 t\|\nabla g(x(t))\|^2 \\
&\quad - bt\langle \nabla g(x(t)) + \epsilon(t)x(t), x(t) - x^*\rangle - t^2\langle \nabla g(x(t)) + \epsilon(t)x(t), \dot{x}(t)\rangle \\
&\quad - \beta t^2\langle \nabla g(x(t)) + \epsilon(t)x(t), \nabla g(x(t))\rangle \qquad (52)
\end{aligned}$$

for every $t \geq t_0$. Further, for every $t \geq t_0$,

$$\begin{aligned}
\|b(x(t) - x^*) + t(\dot{x}(t) + \beta\nabla g(x(t)))\|^2 &= b^2\|x(t) - x^*\|^2 + 2bt\langle \dot{x}(t), x(t) - x^*\rangle \\
&+ 2b\beta t\langle \nabla g(x(t)), x(t) - x^*\rangle \\
&+ t^2\|\dot{x}(t)\|^2 + 2\beta t^2\langle \nabla g(x(t)), \dot{x}(t)\rangle \\
&+ \beta^2 t^2\|\nabla g(x(t))\|^2, \qquad (53)
\end{aligned}$$

which means that (51) becomes

$$\begin{aligned}
\frac{d}{dt}\mathcal{E}_b^p(t) &= t^p((p+2)t + (p+1)(\alpha - \beta - \beta p - b - 1))(g(x(t)) - \min g) \\
&+ \left((p+2)t^{p+1}\frac{\epsilon(t)}{2} + t^{p+2}\frac{\dot{\epsilon}(t)}{2}\right)(\|x(t)\|^2 - \|x^*\|^2) + \frac{b^2 pt^{p-1}}{2}\|x(t) - x^*\|^2 \\
&+ \frac{(p+2)\beta^2 t^{p+1}}{2}\|\nabla g(x(t))\|^2 + \left(b + 1 - \alpha + \frac{p}{2}\right)t^{p+1}\|\dot{x}(t)\|^2 \\
&+ b(b+1-\alpha+p)t^p\langle \dot{x}(t), x(t) - x^*\rangle + b\beta(p+1)t^p\langle \nabla g(x(t)), x(t) - x^*\rangle
\end{aligned}$$

$$- bt^{p+1}\langle \nabla g(x(t)) + \epsilon(t)x(t), x(t) - x^*\rangle - \beta t^{p+2}\langle \nabla g(x(t)) + \epsilon(t)x(t), \nabla g(x(t))\rangle. \tag{54}$$

The gradient inequality for the strongly convex function $x \to g(x) + \frac{\epsilon(t)}{2}\|x\|^2$ gives

$$\langle \nabla g(x(t)) + \epsilon(t)x(t), x^* - x(t)\rangle + \frac{\epsilon(t)}{2}\|x(t) - x^*\|^2 \le \left(g(x^*) + \frac{\epsilon(t)}{2}\|x^*\|^2\right) - \left(g(x(t)) + \frac{\epsilon(t)}{2}\|x(t)\|^2\right),$$

hence

$$-bt^{p+1}\langle \nabla g(x(t)) + \epsilon(t)x(t), x(t) - x^*\rangle \le -bt^{p+1}(g(x(t)) - g^*) - bt^{p+1}\frac{\epsilon(t)}{2}(\|x(t)\|^2 - \|x^*\|^2) - bt^{p+1}\frac{\epsilon(t)}{2}\|x(t) - x^*\|^2$$

for every $t \ge t_0$. Plugging this inequality into (54) gives

$$\begin{aligned}\frac{d}{dt}\mathcal{E}_b^p(t) \le\, & t^p((p+2-b)t + (p+1)(\alpha - \beta - \beta p - b - 1))(g(x(t)) - \min g) \\ & + \left((p+2-b)t^{p+1}\frac{\epsilon(t)}{2} + t^{p+2}\frac{\dot\epsilon(t)}{2}\right)(\|x(t)\|^2 - \|x^*\|^2) + \left(\frac{b^2pt^{p-1}}{2} - bt^{p+1}\frac{\epsilon(t)}{2}\right)\|x(t) - x^*\|^2 \\ & + \left(\frac{(p+2)\beta^2t^{p+1}}{2} - \beta t^{p+2}\right)\|\nabla g(x(t))\|^2 + \left(b + 1 - \alpha + \frac{p}{2}\right)t^{p+1}\|\dot x(t)\|^2 \\ & + b(b+1-\alpha+p)t^p\langle \dot x(t), x(t) - x^*\rangle + b\beta(p+1)t^p\langle \nabla g(x(t)), x(t) - x^*\rangle \\ & - \beta t^{p+2}\epsilon(t)\langle \nabla g(x(t)), x(t)\rangle\end{aligned} \tag{55}$$

for every $t \ge t_0$. Further we have for every $t \ge t_0$

$$b\beta(p+1)t^p\langle \nabla g(x(t)), x(t) - x^*\rangle \le \frac{b\beta(p+1)}{4c^2}t^{p+1}\|\nabla g(x(t))\|^2 + b\beta(p+1)c^2t^{p-1}\|x(t) - x^*\|^2 \tag{56}$$

and

$$-\beta t^{p+2}\epsilon(t)\langle \nabla g(x(t)), x(t)\rangle \le \frac{\beta}{a}t^{p+2}\|\nabla g(x(t))\|^2 + \frac{a\beta}{4}\epsilon^2(t)t^{p+2}\|x(t)\|^2, \tag{57}$$

where $a > 1$ and $c > 0$ are the constants which are assumed to exist in the hypotheses of the theorem, whereby in case $\alpha = 3$ we will take $c = 1$.

Combining (55), (56) and (57) and neglecting the nonpositive terms we derive

$$\begin{aligned}\frac{d}{dt}\mathcal{E}_b^p(t) \le\, & t^p((p+2-b)t + (p+1)(\alpha - \beta - \beta p - b - 1))(g(x(t)) - \min g) \\ & + \left((p+2-b)t^{p+1}\frac{\epsilon(t)}{2} + t^{p+2}\frac{\dot\epsilon(t)}{2} + \frac{a\beta}{4}\epsilon^2(t)t^{p+2}\right)\|x(t)\|^2\end{aligned}$$

$$
\begin{aligned}
&-\left((p+2-b)t^{p+1}\frac{\epsilon(t)}{2}+t^{p+2}\frac{\dot{\epsilon}(t)}{2}\right)\|x^*\|^2\\
&+\left(\frac{b^2pt^{p-1}}{2}+b\beta(p+1)c^2t^{p-1}-bt^{p+1}\frac{\epsilon(t)}{2}\right)\|x(t)-x^*\|^2\\
&+\left(\frac{(p+2)\beta^2t^{p+1}}{2}+\frac{b\beta(p+1)}{4c^2}t^{p+1}-\beta\left(1-\frac{1}{a}\right)t^{p+2}\right)\|\nabla g(x(t))\|^2\\
&+\left(b+1-\alpha+\frac{p}{2}\right)t^{p+1}\|\dot{x}(t)\|^2+b(b+1-\alpha+p)t^p\langle\dot{x}(t),x(t)-x^*\rangle
\end{aligned}
\tag{58}
$$

for every $t \geq t_0$.

For the remaining of the proof we choose the parameters appearing in the definition of the energy functional as

$$
b := \frac{2}{3}\alpha \text{ and } p := \frac{1}{3}(\alpha-3).
$$

Since $\alpha \geq 3$, we have

$$
p+2-b = 1-\frac{\alpha}{3} \leq 0,\ b+1+p-\alpha = 0 \text{ and } b+1+\frac{p}{2}-\alpha = -\frac{p}{2} \leq 0.
$$

Notice that, if $\alpha = 3$, then $(p+2-b)t+(p+1)(\alpha-\beta-\beta p-b-1) = -\beta \leq 0$ and, if $\alpha > 3$, then $p+2-b < 0$. This means that there exists $t_3 \geq t_2$ such that $(p+2-b)t+(p+1)(\alpha-\beta-\beta p-b-1) < 0$ for every $t \geq t_3$. This implies that the term

$$
t^p((p+2-b)t+(p+1)(\alpha-\beta-\beta p-b-1))(g(x(t))-\min g)
$$

in (58) is nonpositive for every $t \geq t_2$ and therefore we will omit it. Further, using that $\lim_{t\longrightarrow+\infty} t^2\epsilon(t) = +\infty$, if $\alpha = 3$, and that $t^2\epsilon(t) \geq \frac{2}{3}\alpha(\frac{1}{3}\alpha - 1 + \beta c^2)$ for t large enough, if $\alpha > 3$, we immediately see that there exists $t_4 \geq t_3$ such that

$$
\frac{b^2pt^{p-1}}{2}+b\beta(p+1)c^2t^{p-1}-bt^{p+1}\frac{\epsilon(t)}{2} \leq 0 \text{ for every } t \geq t_3.
$$

Finally, since $a > 1$, it is obvious that there exists $t_5 \geq t_4$ such that

$$
\frac{(p+2)\beta^2t^{p+1}}{2}+\frac{b\beta(p+1)}{4c^2}t^{p+1}-\beta\left(1-\frac{1}{a}\right)t^{p+2} \leq 0 \text{ for every } t \geq t_5.
$$

Thus, (58) yields

$$
\begin{aligned}
\frac{d}{dt}\mathcal{E}_b^p(t) \leq &\left((p+2-b)t^{p+1}\frac{\epsilon(t)}{2}+t^{p+2}\frac{\dot{\epsilon}(t)}{2}+\frac{a\beta}{4}\epsilon^2(t)t^{p+2}\right)\|x(t)\|^2\\
&-\left((p+2-b)t^{p+1}\frac{\epsilon(t)}{2}+t^{p+2}\frac{\dot{\epsilon}(t)}{2}\right)\|x^*\|^2
\end{aligned}
$$

$$= \left((p+2-b)t^{p+1}\frac{\epsilon(t)}{2} + t^{p+2}\frac{\dot{\epsilon}(t)}{2} + \frac{a\beta}{4}\epsilon^2(t)t^{p+2} \right)(\|x(t)\|^2 - \|x^*\|^2) + \frac{a\beta}{4}\epsilon^2(t)t^{p+2}\|x^*\|^2, \tag{59}$$

for every $t \geq t_5$. By the hypotheses, we have that

$$(p+2-b)t^{p+1}\frac{\epsilon(t)}{2} + t^{p+2}\frac{\dot{\epsilon}(t)}{2} + \frac{a\beta}{4}\epsilon^2(t)t^{p+2} \leq 0,$$

for every $t \geq t_5$ and, taking into account the setting considered in this first case, it follows there exists $t_6 \geq t_5$ such that

$$\|x(t)\|^2 - \|x^*\|^2 \geq 0$$

for every $t \geq t_6$. Hence, (59) leads to

$$\frac{d}{dt}\mathcal{E}_b^p(t) \leq \frac{a\beta}{4}\epsilon^2(t)t^{p+2}\|x^*\|^2 \text{ for every } t \geq t_6. \tag{60}$$

By integrating (60) on the interval $[t_6, t]$, for arbitrary $t \geq t_6$, we get

$$\mathcal{E}_b^p(t) \leq \mathcal{E}_b^p(t_6) + \frac{a\beta}{4}\|x^*\|^2 \int_{t_6}^t \epsilon^2(s)s^{p+2}dt. \tag{61}$$

Recall that from (50) we have

$$\mathcal{E}_b^p(t) \geq \frac{\epsilon(t)}{2}t^{p+2}(\|x(t) - x_{\epsilon(t)}\|^2 + \|x_{\epsilon(t)}\|^2 - \|x^*\|^2),$$

which, combined with (61), gives for every $t \geq t_6$ that

$$\|x(t) - x_{\epsilon(t)}\|^2 \leq \|x^*\|^2 - \|x_{\epsilon(t)}\|^2 + \frac{2E_b^p(t_6)}{\epsilon(t)t^{\frac{1}{3}\alpha+1}} + \frac{a\beta}{2\epsilon(t)t^{\frac{1}{3}\alpha+1}}\|x^*\|^2 \int_{t_6}^t \epsilon^2(s)s^{\frac{1}{3}\alpha+1}dt. \tag{62}$$

Using that $\lim_{t\longrightarrow+\infty} \epsilon(t)t^{\frac{1}{3}\alpha+1} = +\infty$, $\lim_{t\longrightarrow+\infty} x_{\epsilon(t)} = x^*$ and taking into account the hypotheses of the theorem, we get that the right-hand side of (62) converges to 0 as $t \longrightarrow +\infty$. This yields

$$\lim_{t\longrightarrow+\infty} x(t) = x^*.$$

Case II Assume that there exists $T \geq t_0$ such that the trajectory $\{x(t) : t \geq T\}$ stays in the ball $B(0, \|x^*\|)$.

In other words, $\|x(t)\| < \|x^*\|$ for every $t \geq T$. Since

$$\int_{t_0}^{+\infty} \frac{\epsilon(t)}{t}dt < +\infty,$$

according to Theorem 3.1, we have

$$\lim_{t \longrightarrow +\infty} g(x(t)) = \min g.$$

Consider $\overline{x} \in \mathcal{H}$ a weak sequential cluster point of the trajectory x, which exists since the trajectory is bounded. This means that there exists a sequence $(t_n)_{n \in \mathbb{N}} \subseteq [T, +\infty)$ such that $t_n \longrightarrow +\infty$ and $x(t_n)$ converges weakly to $\overline{x}$ as $n \longrightarrow +\infty$.

Since g is weakly lower semicontinuous, it holds

$$g(\overline{x}) \leq \liminf_{n \to +\infty} g(x(t_n)) = \min g, \text{ thus } \overline{x} \in \operatorname{argmin} g.$$

Since the norm is weakly lower semicontinuous, it holds

$$\|\overline{x}\| \leq \liminf_{n \to +\infty} \|x(t_n)\| \leq \|x^*\|,$$

which, by taking into account that x^* is the unique element of minimum norm in $\operatorname{argmin} g$, implies $\overline{x} = x^*$. This shows that the whole trajectory x converges weakly to x^*.

Thus,

$$\|x^*\| \leq \liminf_{t \to +\infty} \|x(t)\| \leq \limsup_{t \to +\infty} \|x(t)\| \leq \|x^*\|, \text{ hence } \lim_{t \longrightarrow +\infty} \|x(t)\| = \|x^*\|.$$

But by taking into account that $x(t) \rightharpoonup x^*$ as $t \longrightarrow +\infty$, we obtain that the convergence is strong, that is

$$\lim_{t \longrightarrow +\infty} x(t) = x^*.$$

Case III Assume that for every $T \geq t_0$ there exists $t \geq T$ such that $\|x^*\| > \|x(t)\|$ and there exists $s \geq T$ such that $\|x^*\| \leq \|x(s)\|$.

By the continuity of x it follows that there exists a sequence $(t_n)_{n \in \mathbb{N}} \subseteq [t_0, +\infty)$ such that $t_n \longrightarrow +\infty$ as $n \longrightarrow +\infty$ and

$$\|x(t_n)\| = \|x^*\| \text{ for every } n \in \mathbb{N}.$$

We will show that $x(t_n) \longrightarrow x^*$ as $n \longrightarrow +\infty$. To this end we consider $\overline{x} \in \mathcal{H}$ a weak sequential cluster point of the sequence $(x(t_n))_{n \in \mathbb{N}}$. By repeating the arguments used in the previous case (notice that the sequence is bounded) it follows that $(x(t_n))_{n \in \mathbb{N}}$ converges weakly to x^* as $n \longrightarrow +\infty$. Since $\|x(t_n)\| \longrightarrow \|x^*\|$ as $n \longrightarrow +\infty$, it yields $\|x(t_n) - x^*\| \longrightarrow 0$ as $n \longrightarrow +\infty$. This shows that

$$\liminf_{t \longrightarrow +\infty} \|x(t) - x^*\| = 0.$$

□

Remark 4.5 Theorem 4.4 can be seen as an extension of a result given in [8] for the dynamical system (4) to the dynamical system with Hessian driven damping and Tikhonov regularization term (5). One can notice that for the choice $\beta = 0$, which means that the Hessian driven damping is removed, the lower bound we impose for $t \mapsto t^2\epsilon(t)$ in case $\alpha > 3$ is less tight than the one considered in [8, Theorem 4.1] for the system (4). As we will see later, this lower bound influences the asymptotic behaviour of the trajectory.

In case $\beta > 0$, in order to guarantee that

$$\lim_{t \longrightarrow +\infty} \frac{\beta}{\epsilon(t)t^{\frac{\alpha}{3}+1}} \int_{t_0}^{t} \epsilon^2(s)s^{\frac{\alpha}{3}+1}ds = 0,$$

one just have to additionally assume that

$$\int_{t_0}^{+\infty} \epsilon(t)dt < +\infty$$

and that the function

$$t \longrightarrow t^{\frac{1}{3}\alpha+1}\epsilon(t) \text{ is nondecreasing for } t \text{ large enough.}$$

This follows from Lemma A.1, by also taking into account that $\lim_{t \longrightarrow +\infty} \epsilon(t)t^{\frac{\alpha}{3}+1} = +\infty$.

Combining the main results in the last two sections, one can see that if

$$\int_{t_0}^{+\infty} t\epsilon(t)dt < +\infty,$$

the function

$$t \longrightarrow t^{\frac{1}{3}\alpha+1}\epsilon(t) \text{ is nondecreasing for } t \text{ large enough,}$$

there exist $a > 1$ and $t_1 \geq t_0$ such that

$$\dot{\epsilon}(t) \leq -\frac{a\beta}{2}\epsilon^2(t) \text{ for every } t \geq t_1,$$

and

- in case $\alpha = 3$: $\lim_{t \longrightarrow +\infty} t^2\epsilon(t) = +\infty$;
- in case $\alpha > 3$: there exists $c > 0$ such that $t^2\epsilon(t) \geq \frac{2}{3}\alpha\left(\frac{1}{3}\alpha - 1 + \beta c^2\right)$ for t large enough,

then one obtains both fast convergence of the function values and strong convergence of the trajectory to the minimal norm solution. This is for instance the case when $\epsilon(t) = t^{-\gamma}$ for all $\gamma \in (1, 2)$.

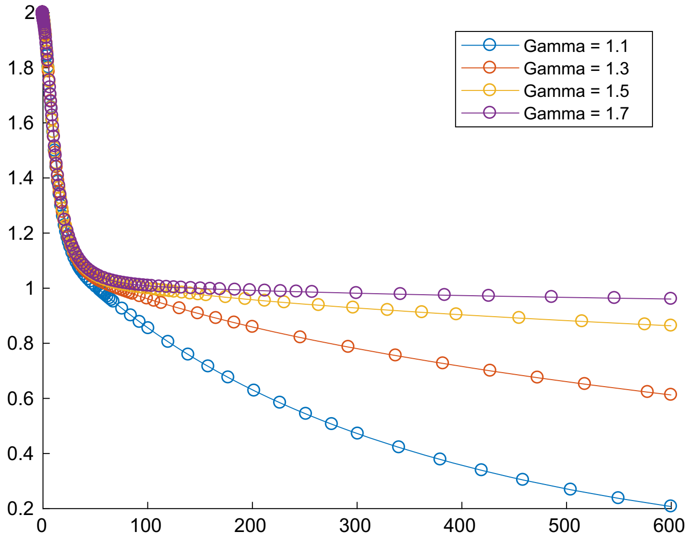

Fig. 2 The behaviour of the trajectories generated by the dynamical system (5) in relation with the minimization of the function given in (6) for $\alpha = 200$, $\beta = 1$, $\epsilon(t) = t^{-\gamma}$ and different values for $\gamma \in (1, 2)$

In the following, we would like to comment on the role on the condition in Theorem 4.4 which asks, in case $\alpha > 3$, for the existence of a positive constant c such that $t^2\epsilon(t) \geq \frac{2}{3}\alpha(\frac{1}{3}\alpha - 1 + \beta c^2)$ for t large enough. To this end it is very helpful to visualize the trajectories generated by the dynamical system (5) in relation with the minimization of the function given in (6) for a fixed large value of α and Tikhonov parametrizations of the form $t \mapsto \epsilon(t) = t^{-\gamma}$, for different values of $\gamma \in (1, 2)$. The trajectories in the plot in Fig. 2 have been generated for $\alpha = 200$ and $\beta = 1$ and are all approaching the minimum norm solution $x^* = 0$. The norm of the difference between the trajectory and the minimum norm solution is guaranteed to be bounded from above by a function which converges to zero, after the time point t is reached at which the inequality $t^2\epsilon(t) \geq \frac{2}{3}\alpha(\frac{1}{3}\alpha - 1 + \beta c^2)$ "starts" being fulfilled. For large α and the Tikhonov parametrizations considered in our experiment, the closer γ is to 1, the faster is this inequality fulfilled. This is reflected by the behaviour of the trajectories plotted in Fig. 2.

Finally, we would like to formulate some possible questions of future research related to the dynamical sytem (5):

- In [7, Theorem 3.4] it has been proved for the dynamical system (1) that, when g is strongly convex, the rates of convergence of the function values and the tracjectory are both of $O(t^{-\frac{2}{3}\alpha})$, thus they can be made arbitrarily fast by taking α large. It is natural to ask if similar rates of convergence can be obtained in a similar setting for the dynamical system (5) (see, also, [8, Section 5.4]).
- In the literature, in the context of dynamical systems, regularization terms have been considered not only in open-loop, but also in closed-loop form (see, for instance, [12]). It is an interesting question if one can obtain for the dynamical system (5) similar results if the Tikhonov regularization term is taken in closed-loop form.

- A natural question is to formulate proper numerical algorithms via time discretization of (5), to investigate their theoretical convergence properties, and to validate them with numerical experiments.

Acknowledgements Open access funding provided by Austrian Science Fund (FWF). The authors are thankful to an anonymous reviewer for comments and remarks which improved the quality of the paper.

Appendix

In this appendix, we collect some lemmas and technical results which we will use in the analysis of the dynamical system (5). The following lemma was stated for instance in [8, Lemma A.3] and is used to prove the convergence of the objective function along the trajectory to its minimal value.

Lemma A.1 *Let $\delta > 0$ and $f \in L^1((\delta, +\infty), \mathbb{R})$ be a nonnegative and continuous function. Let $\varphi : [\delta, +\infty) \longrightarrow [0, +\infty)$ be a nondecreasing function such that $\lim_{t \longrightarrow +\infty} \varphi(t) = +\infty$. Then it holds*

$$\lim_{t \longrightarrow +\infty} \frac{1}{\varphi(t)} \int_\delta^t \varphi(s) f(s) ds = 0.$$

The following statement is the continuous counterpart of a convergence result of quasi-Fejér monotone sequences. For its proofs we refer to [1, Lemma 5.1].

Lemma A.2 *Suppose that $F : [t_0, +\infty) \to \mathbb{R}$ is locally absolutely continuous and bounded from below and that there exists $G \in L^1([t_0, +\infty), \mathbb{R})$ such that*

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

for almost every $t \in [t_0, +\infty)$. Then there exists $\lim_{t \longrightarrow +\infty} F(t) \in \mathbb{R}$.

The following technical result is [11, Lemma 2].

Lemma 4.6 *Let $u : [t_0, +\infty) \longrightarrow \mathcal{H}$ be a continuously differentiable function satisfying $u(t) + \frac{t}{\alpha}\dot{u}(t) \longrightarrow u \in \mathcal{H}$ as $t \longrightarrow +\infty$, where $\alpha > 0$. Then $u(t) \longrightarrow u$ as $t \longrightarrow +\infty$.*

The continuous version of the Opial Lemma (see [7]) is the main tool for proving weak convergence for the generated trajectory.

Lemma A.3 *Let $S \subseteq \mathcal{H}$ be a nonempty set and $x : [t_0, +\infty) \to H$ a given map such that:*

(i) *for every $z \in S$ the limit $\lim\limits_{t \longrightarrow +\infty} \|x(t) - z\|$ exists;*

(ii) *every weak sequential limit point of $x(t)$ belongs to the set S.*

Then the trajectory $x(t)$ converges weakly to an element in S as $t \to +\infty$.

References

1. Abbas, B., Attouch, H., 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)
2. 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. Journal de Mathématiques Pures et Appliquées **81**(8), 747–779 (2002)
3. Alvarez, F., Cabot, A.: Asymptotic selection of viscosity equilibria of semilinear evolution equations by the introduction of a slowly vanishing term. Discrete Contin. Dyn. Syst. **15**, 921–938 (2006)
4. Attouch, H.: Viscosity solutions of minimization problems. SIAM J. Optim. **6**(3), 769–806 (1996)
5. Attouch, H., Chbani, Z., Fadili, J., Riahi, H.: First-order optimization algorithms via inertial systems with Hessian driven damping (2019). arXiv:1907.10536v1
6. Attouch, H., Chbani, Z., Riahi, H.: Rate of convergence of the Nesterov accelerated gradient method in the subcritical case $\alpha \leq 3$. ESAIM: Control Optim. Calc. Var. 25(2) (2019)
7. Attouch, H., Chbani, Z., Peypouquet, J., Redont, P.: Fast convergence of inertial dynamics and algorithms with asymptotic vanishing viscosity. Math. Program. Ser. B **168**(1–2), 123–175 (2018)
8. Attouch, H., Chbani, Z., Riahi, H.: Combining fast inertial dynamics for convex optimization with Tikhonov regularization. J. Math. Anal. Appl. **457**(2), 1065–1094 (2018)
9. Attouch, H., Cominetti, R.: A dynamical approach to convex minimization coupling approximation with the steepest descent method. J. Differ. Equ. **128**(2), 519–540 (1996)
10. Attouch, H., Czarnecki, M.-O.: Asymptotic control and stabilization of nonlinear oscillators with non-isolated equilibria. J. Differ. Equ. **197**, 278–310 (2002)
11. Attouch, H., Peypouquet, J., Redont, P.: Fast convex optimization via inertial dynamics with Hessian driven damping. J. Differ. Equ. **261**(10), 5734–5783 (2016)
12. Attouch, H., Redont, P., Svaiter, B.F.: Global convergence of a closed-loop regularized Newton method for solving monotone inclusions in Hilbert spaces. J. Optim. Theory Appl. **157**, 624–650 (2013)
13. Attouch, H., Svaiter, B.F.: A continuous dynamical Newton-like approach to solving monotone inclusions. SIAM J. Control Optim. **49**(2), 574–598 (2011)
14. Cabot, A., Engler, H., Gadat, S.: On the long time behavior of second order differential equations with asymptotically small dissipation. Trans. Am. Math. Soc. **361**, 5983–6017 (2009)
15. Cominetti, R., Peypouquet, J., Sorin, S.: Strong asymptotic convergence of evolution equations governed by maximal monotone operators with Tikhonov regularization. J. Differ. Equ. **245**, 3753–3763 (2008)
16. Nesterov, Y.: A method of solving a convex programming problem with convergence rate $O(1/k^2)$. Soviet Math. Doklady **27**, 372–376 (1983)
17. Sell, G.R., You, Y.: Dynamics of Evolutionary Equations. Springer, New York (2002)
18. Shi, B., Du, S.S., Jordan, M.I., Su, W.J.: Understanding the acceleration phenomenon via high-resolution differential equations (2018). arXiv:1810.08907v3
19. Shi, B., Du, S.S., Su, W.J., Jordan, M.I.: Acceleration via symplectic discretization of high-resolution differential equations. Adv. Neural Inf. Process. Syst. **32**(NIPS 2019), 5745–5753 (2019)
20. 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**(153), 1–43 (2016)