On the strong convergence of continuous Newton-like inertial dynamics with Tikhonov regularization for monotone inclusions*

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

August 11, 2023

Abstract. In a Hilbert space $\mathcal{H}$, we study the convergence properties of the trajectories of a Newton-like inertial dynamical system with a Tikhonov regularization term governed by a general maximally monotone operator $A : \mathcal{H} \to 2^{\mathcal{H}}$. The maximally monotone operator enters the dynamics via its Yosida approximation with an appropriate adjustment of the Yosida regularization parameter, by adopting an approach introduced by Attouch-Peypouquet (Math. Prog., 2019) and further developed by Attouch-László (Set-Valued Var. Anal., 2021). We obtain fast rates of convergence for the velocity and the Yosida regularization term towards zero, while the generated trajectories converge weakly towards a zero of A or, depending on the system parameters, strongly towards the zero of minimum norm of A. Our analysis reveals that the damping coefficient, the Yosida regularization parameter and the Tikhonov parametrization are strongly correlated.

Key Words. monotone inclusion, Newton method, vanishing damping, Yosida regularization, Tikhonov regularization, strong convergence, convergence rates

AMS subject classification. 37N40, 46N10, 49M30, 65K05, 65K10, 90B50, 90C25.

1 Introduction

1.1 The dynamical system and a model result

Let $\mathcal{H}$ be a real Hilbert space endowed with the scalar product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$. Given a general set-valued maximally monotone operator $A : \mathcal{H} \to 2^{\mathcal{H}}$ consider the inclusion problem

$$0 \in Ax, \tag{1}$$

and assume that the solution set $S := \{x \in \mathcal{H} : 0 \in Ax\}$ is nonempty.

The operator $A : \mathcal{H} \to 2^{\mathcal{H}}$ is said to be monotone if $\langle u - v, x - y \rangle \geq 0$ for all $(x, u), (y, v) \in \operatorname{Gr} A$, where $\operatorname{Gr} A := \{(z, w) \in \mathcal{H} \times \mathcal{H} : w \in A(z)\}$ denotes its graph. The operator A is said to be maximally monotone if there is no other monotone operator $A' : \mathcal{H} \to 2^{\mathcal{H}}$ such that $\operatorname{Gr} A \subsetneq \operatorname{Gr} A'$. The most prominent example for a set-valued maximally monotone operator is the convex subdifferential $\partial f : \mathcal{H} \to 2^{\mathcal{H}}$, $\partial f(x) := \{u \in \mathcal{H} : f(y) - f(x) \geq \langle u, y - x \rangle \ \forall y \in \mathcal{H}\}$ of a proper, convex and lower semicontinuous

*This work was supported by a grant of the Romanian Ministry of Research, Innovation and Digitization, CNCS - UEFISCDI, project number PN-III-P1-1.1-TE-2021-0138, within PNCDI III. The research of RIB was also supported by FWF (Austrian Science Fund), projects W 1260 and P 34922-N.

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

[‡]Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria, email: robert.csetnek@univie.ac.at

[§]Technical University of Cluj-Napoca, Department of Mathematics, Str. Memorandumului nr. 28, 400114 Cluj-Napoca, Romania, email: slaszlo@math.utcluj.ro.

function $f : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$. In this case, solving the inclusion problem (1) is equivalent to finding a global minimum of f.

In order to solve (1) we will study the asymptotic behaviour, as $t \to +\infty$, of the second-order in time evolution equation for $t \geq t_0$

$$\text{(DIN-AVD-TIKH)} \qquad \ddot{x}(t) + \frac{\alpha}{t^q}\dot{x}(t) + \beta \frac{d}{dt}\left(A_{\lambda(t)}(x(t))\right) + A_{\lambda(t)}(x(t)) + \epsilon(t)x(t) = 0,$$

where $\alpha > 0$, $\beta \geq 0$, $0 < q \leq 1$, $\lambda : [t_0, +\infty) \longrightarrow (0, +\infty)$ is the Yosida parametrization function and $\epsilon : [t_0, +\infty) \longrightarrow (0, +\infty)$ is the Tikhonov parametrization function. The single-valued operators $J_{\lambda A} : \mathcal{H} \to \mathcal{H}$ and $A_\lambda : \mathcal{H} \to \mathcal{H}$, defined by

$$J_{\lambda A} = (\mathrm{Id} + \lambda A)^{-1} \text{ and } A_\lambda = \frac{1}{\lambda}\left(\mathrm{Id} - J_{\lambda A}\right),$$

are the resolvent of A and the Yosida regularization of A of parameter $\lambda > 0$, respectively. Notice that A_λ is λ-cocoercive, namely $\langle A_\lambda(x) - A_\lambda(y), x - y\rangle \geq \lambda \|A_\lambda(x) - A_\lambda(y)\|^2$ for all $\lambda > 0$ and $x, y \in \mathcal{H}$, hence λ^{-1}-Lipschitz continuous. It holds $0 \in Ax$ if and only if $A_\lambda(x) = 0$. For all these and other properties of these operators we refer the reader to [10].

The dependence of the Yosida regularization parameter $\lambda(t)$ on time will play a crucial role in the asymptotic analysis. This has been already noticed by Attouch-Peypouquet in [7], where in case $\beta = 0$, $q = 1$ (classical vanishing damping) and $\epsilon = 0$ (without Tikhonov regularization) a Yosida regularization parameter of order t^2 was considered (see also [5] and [13]). This motivates the adoption of the following standing assumptions in effect for the rest of the paper (even if some of the statements, including Proposition 5, Lemma 6 or Theorem A.1, hold under more general hypotheses):

General assumption:

- $\lambda(t) := \lambda t^{2q}$, with $\lambda > 0$ and $0 < q \leq 1$;
- $\epsilon : [t_0, +\infty) \longrightarrow [0, +\infty)$ is a nonincreasing function of class C^1 fulfilling $\lim_{t \longrightarrow +\infty} \epsilon(t) = 0$.

Notice that the second condition on the Tikhonov parametrization function is standard in the context of Tikhonov regularization. The presence of the Tikhonov regularization term in the dynamical system is responsible for the strong convergence of the trajectories to the minimal norm solution in S, while the rates of convergence of the velocity and the Yosida regularization term towards zero can be quantified. The corresponding asymptotic analysis relies on the construction of a suitable energy function which dissipates along time. The considered approach is rooted in the recent developments from [2, 3, 15] on Tikhonov regularized inertial dynamics for convex optimization. At their turn, these can be seen as continuous time dynamics resulting by applying Polyak's Heavy Ball Method to the strongly convex function obtained by adding a corresponding Tikhonov regularization term to a convex C^1 function (see Subsection 1.2 for further details).

For the following model result, which summarizes the statements in Corollary 4 and Corollary 9, we assume that $0 < q < 1$, choose $\epsilon(t) := \frac{a}{t^p}$, for $a, p > 0$, and recall that $\operatorname{sgn} \beta = 0$ for $\beta = 0$ and $\operatorname{sgn} \beta = 1$ for $\beta > 0$. One can easily notice that $\max\left((\operatorname{sgn}\beta)(1-q), \frac{3q+1}{2}\right) < q + 1 < 2$.

Theorem 1. *Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator such that $S := \{x \in \mathcal{H} : 0 \in Ax\}$ is nonempty. Consider the evolution equation* (DIN-AVD-TIKH) *with $\alpha > 1$, $\beta \geq 0$, $\lambda(t) = \lambda t^{2q}$, with $\lambda > \frac{1}{\alpha^2}$ and $0 < q < 1$, and $\epsilon(t) = \frac{a}{t^p}$, with $a, p > 0$. Then, for any trajectory $x : [t_0, +\infty) \to \mathcal{H}$ of* (DIN-AVD-TIKH)*, the following statements are true:*

(i) If $q+1 < p \leq 2$ and, moreover, $a \geq q(1-q)$ for $p=2$, we have the rates of convergence $\|\dot{x}(t)\| = o\left(\frac{1}{t^q}\right)$ and $\|A_{\lambda(t)}(x(t))\| = o\left(\frac{1}{t^{2q}}\right)$ as $t \to +\infty$; if, in addition, $\beta = 0$ or $\beta > 0$ and $q > \dfrac{1}{2}$, then $x(t)$ converges weakly, as $t \to +\infty$, to an element of S.

(ii) If $\max\left((\operatorname{sgn}\beta)(1-q), \frac{3q+1}{2}\right) < p < q+1$, we have the rates of convergence $\|\dot{x}(t)\| = o\left(\frac{1}{t^q}\right)$ and $\|A_{\lambda(t)}(x(t))\| = o\left(\frac{1}{t^{2q}}\right)$ as $t \to +\infty$ and $x(t)$ converges strongly, as $t \to +\infty$, to the element of minimum norm of S.

(iii) If $p = q+1$, we have the rates of convergence $\|\dot{x}(t)\| = \mathcal{O}\left(\frac{1}{t^q}\right)$ and $\|A_{\lambda(t)}(x(t))\| = \mathcal{O}\left(\frac{1}{t^{2q}}\right)$ as $t \to +\infty$. This case is critical and separates the settings in which weak and strong convergence for the trajectory can be obtained. In this case we can at least guarantee that $x(t)$ is bounded.

We would like to stress that in Theorem 1(ii) we provide a setting in which we have both strong convergence of the trajectories to the element of minimum norm in S and fast convergence rates for the velocity and the Yosida regularization. In the particular case of a convex optimization problem we will derive from here fast convergence rates also for the objective function value, see Theorem 11 in Section 4. The setting in which all these statements hold requires q to be strictly less than 1.

Let us comment on the case $q = 1$ for which (DIN-AVD-TIKH) is the dynamical system with classical vanishing damping term $\frac{\alpha}{t}$. The analysis in Section 3, see Theorem 7, prohibits the choice $q = 1$ for getting both strong convergence of the trajectories and fast convergence rates. However, for this choice we can have according to Corollary 4 the weak convergence of the trajectories and fast convergence rates. On the other hand, one can consider the alternative approach developed in [4] (see also [11]) for an inertial dynamical system with Tikhonov regularization and classical vanishing damping term $\frac{\alpha}{t}$ approaching the minimization of a convex C^1 function, which provides two different settings in terms of the dynamical system parameters, one for the strong convergence of the trajectories to minimum norm minimizer and another one for fast convergence rates. However, the drawback of this approach is that there is no common instance for which both statements hold.

1.2 Related works

The second order in time evolution equation with vanishing damping defined for $t \geq t_0$ as

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

where $\alpha > 0$ and $f : \mathcal{H} \to \mathbb{R}$ is a convex C^1 function, introduced by Su-Boyd-Candés in [17], was the starting point of significant research activities devoted to this type of dynamics. The reason is that it exhibits for $f(x(t)) - \min f$ a rate of convergence of order $\mathcal{O}(\frac{1}{t^2})$ as $t \to +\infty$, where $\min f$ denotes the minimal function value of f, which is assume to have a global minimizer. This is an improvement over the gradient flow

$$\dot{x}(t) + \nabla f(x(t)) = 0 \tag{3}$$

and Polyak's Heavy Ball Method with friction

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

for which the rate of convergence for $f(x(t)) - \min f$ is of order $\mathcal{O}(\frac{1}{t})$ as $t \to +\infty$. The dynamics (2) can be seen as a continuous version of the celebrated Nesterov accelerated gradient algorithm with momentum [16].

Later, in [8] it has been shown that an additional Hessian driven damping term in (2)

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

where $\beta \geq 0$ and f is a convex C^2 function, may attenuate the oscillations of the trajectories and provide integrability results for ∇f along the trajectories, while preserving the fast rate of convergence for the function values. The evolution equation (5) is intimately related to Newton-type methods, see [8, 9]. By taking into account that $\nabla^2 f(x(t))\dot{x}(t) = \frac{d}{dt}\nabla f(x(t))$, the temporal discretizations of the dynamical system with Hessian driven damping term lead to inertial type algorithms involving gradient correction terms.

Second order dynamics with vanishing damping have been considered also in the context of solving monotone inclusion problems of the from (1) (see [7])

$$\ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + A_{\lambda(t)}(x(t)) = 0, \tag{6}$$

for which rates have been obtained for the convergence of $\|\dot{x}(t)\|$ and $\|A_{\lambda(t)}(x(t))\|$ towards zero as $t \to +\infty$. In case A is the convex subdifferential, the dynamics (6) can be regarded as a smoothing approach for solving nonsmooth optimization problems.

The evolution equation (6) has been further developed in [5] by considering, in analogy with the dynamics in (5), an additional Newton-like correction term

$$\ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \beta\frac{d}{dt}\left(A_{\lambda(t)}(x(t))\right) + A_{\lambda(t)}(x(t)) = 0. \tag{7}$$

We refer also to [13] for a dynamical system with a splitting character tailored to the solving of structured monotone inclusions.

As for the evolution equations with vanishing damping along fast rates of convergence, in general, "only" the weak convergence of the trajectories can be proved, they have been enhanced with Tikhonov regularization terms in order to enforce strong convergence of the trajectories to a minimal norm solution. For an alternative approach for deriving convergence rates for continuous dynamics with vanishing damping for linear ill-posed problems in the presence of source conditions we refer to [12].

In connection with the minimization of a convex C^1 function $f : \mathcal{H} \to \mathbb{R}$, the following dynamical system defined for $t \geq t_0$ as

$$\ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \nabla f(x(t)) + \varepsilon(t)x(t) = 0 \tag{8}$$

has been considered in [4]. Depending on the speed of convergence of the Tikhonov parametrization function $\varepsilon(t)$ to 0 as $t \to +\infty$, the authors obtained the strong convergence of the trajectories to minimizer of minimum norm and fast convergence rates for the function values, however, in two settings different from another. The analysis has been extended in [11] to evolution equations with involving Hessian driven damping and Tikhonov regularization terms

$$\ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \beta\nabla^2 f(x(t))\dot{x}(t) + \nabla f(x(t)) + \varepsilon(t)x(t) = 0. \tag{9}$$

We refer to [14] for a further extension to nonsmooth convex optimization problems. We stress that in none of the approaches in [4, 11, 14] there is a setting for the system parameters in which both fast convergence rates and strong convergence of the trajectories to the minimal norm solution hold.

This drawback was recently overcome in [2] (see also [3, 6]) by considering the damping proportional to the square root of the Tikhonov parameter

$$\ddot{x}(t) + \alpha\sqrt{\varepsilon(t)}\dot{x}(t) + \nabla f(x(t)) + \varepsilon(t)x(t) = 0. \tag{10}$$

This evolution equation was motivated by interesting observation that if the function f in (4) is γ-strongly convex with modulus $\gamma > 0$, then the Heavy Ball Method (4) with $\alpha := 2\sqrt{\gamma}$ leads to an evolution equation with guaranteed exponential convergence rate for $f(x(t)) - \min f$. Therefore, the idea of this new Tikhonov regularization approach is to approximate the function f by the strongly convex function

$f_{\varepsilon(t)} = f + \frac{\epsilon(t)}{2}\|\cdot\|^2$ with modulus $\epsilon(t)$, where $\epsilon(t)$ decays to 0 as $t \to +\infty$. The term $\nabla f(x(t)) + \varepsilon(t)x(t)$ in (10) is nothing else than $\nabla f_{\epsilon(t)}(x(t))$. In other words, the idea in [2] is to take advantage of the remarkable properties of the heavy ball method when applied to strongly convex functions. Another particularity of the new approach concerns the energy functional used in the analysis. Indeed, while in the approaches [4, 11, 14] the energy functional is anchored at a solution of the problem to be solved, in [2] it is anchored at $x_{\epsilon(t)}$, where $x_{\epsilon(t)}$ is the unique minimizer of $f_{\epsilon(t)}$.

The approaches in [2, 3, 6] have been alleviated in [15] by considering evolution equations for which the damping was not necessarily proportional (tough correlated) to the square root of the Tikhonov parametrization function. In this more flexible setting, strong convergence of the trajectories to a minimum norm solution in combination with faster rates of convergence than in [2] have been proved.

In this paper we follow the recently introduced approach from [2] with the modification considered in [15] and provide among others a setting in which both the strong convergence of the trajectories of (DIN-AVD-TIKH) to the minimal norm solution of (1) and fast rates of convergence for the velocity and the Yosida regularization along the trajectory can be shown.

2 Fast convergence rates and weak convergence

In this section we will introduce an energy functional anchored at a solution of (1) which dissipates as $t \to +\infty$. This will allow us to derive some pointwise and integral estimates, as well as fast rates of convergence for the velocity, the Yosida regularization of the operator and its time derivative. We also obtain the weak convergence of the trajectories under some mild assumptions imposed on the system parameters and the Tikhonov parametrization function.

The main result of this section is the following.

Theorem 2. *Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator such that $S := \{x \in \mathcal{H} : 0 \in Ax\}$ is nonempty. Consider the evolution equation* (DIN-AVD-TIKH) *with $\beta \geq 0$, $\lambda(t) = \lambda t^{2q}$, for $\lambda > 0$ and $0 < q \leq 1$, and the system parameters satisfying the following conditions:*

$$\text{if } q < 1, \text{ then } \lambda > \frac{1}{\alpha^2} \text{ and } \epsilon(t) \geq \frac{q(1-q)}{t^2} \text{ for } t \text{ large enough;}$$

$$\text{if } q = 1, \text{ then either } \alpha > 3 \text{ and } \lambda > \frac{1}{8(\alpha-3)} \text{ or } \alpha > 1, \lambda > \frac{1}{(\alpha-1)^2} \text{ and } \frac{\dot{\epsilon}(t)}{\epsilon(t)} \leq -\frac{2}{t} \text{ for } t \text{ large enough.}$$

In addition, assume that $\int_{t_0}^{+\infty} t^q \epsilon(t) dt < +\infty$.

Then, for any trajectory $x : [t_0, +\infty) \to \mathcal{H}$ of (DIN-AVD-TIKH)*, the following statements are true:*

(i) (convergence of the trajectory) $x(t)$ is bounded. Furthermore, if $\beta = 0$ or $\beta > 0$ and $q > \frac{1}{2}$, then $x(t)$ converges weakly, as $t \to +\infty$, to an element of S;

(ii) (integral estimates) $$\int_{t_0}^{+\infty} t^q \|\dot{x}(t)\|^2 dt < +\infty, \quad \int_{t_0}^{+\infty} t^{3q} \|A_{\lambda(t)}(x(t))\|^2 dt < +\infty,$$
$$\text{and, if } \int_{t_0}^{+\infty} t^{3q}\epsilon^2(t) dt < +\infty, \text{ then } \int_{t_0}^{+\infty} t^{3q} \|\ddot{x}(t)\|^2 dt < +\infty;$$

(iii) (fast convergence rates) $\|\dot{x}(t)\| = \mathcal{O}\left(\frac{1}{t^q}\right), \|A_{\lambda(t)}(x(t))\| = o\left(\frac{1}{t^{2q}}\right), \left\|\frac{d}{dt}A_{\lambda(t)}(x(t))\right\| = \mathcal{O}\left(\frac{1}{t^{3q}}\right)$ as $t \to +\infty$. In addition, if $\int_{t_0}^{+\infty} t^{3q}\epsilon^2(t)dt < +\infty$, then $\|\dot{x}(t)\| = o\left(\frac{1}{t^q}\right)$ as $t \to +\infty$.

Proof. Energy functional. Choose $z \in S$. For $0 < b < \alpha$ if $q < 1$ and $0 < b < \alpha - 1$ if $q = 1$, consider the energy functional

$$\mathcal{E}(t) := \frac{1}{2}\|b(x(t) - z) + t^q(\dot{x}(t) + \beta A_{\lambda(t)}(x(t)))\|^2 + \frac{b(\alpha - qt^{q-1} - b)}{2}\|x(t) - z\|^2 + \frac{t^{2q}\epsilon(t)}{2}\|x(t)\|^2. \quad (11)$$

Note that there exists $t_0' \geq t_0$ such that $\alpha - qt^{q-1} - b > 0$ for all $t \geq t_0'$, consequently, $\mathcal{E}(t) \geq 0$ for all $t \geq t_0'$. The aim of the first part of the proof is to derive the inequality (27).

Using the classical derivation chain rule and (DIN-AVD-TIKH), we get that for the derivative of the energy function for all $t \geq t_0$

$$\begin{aligned}\dot{\mathcal{E}}(t) =& \big\langle (b + qt^{q-1} - \alpha)\dot{x}(t) + (\beta qt^{q-1} - t^q)A_{\lambda(t)}(x(t)) - t^q\epsilon(t)x(t), b(x(t) - z) + t^q(\dot{x}(t) + \beta A_{\lambda(t)}(x(t)))\big\rangle \\ &+ b(\alpha - qt^{q-1} - b)\langle \dot{x}(t), x(t) - z\rangle + \frac{bq(1-q)t^{q-2}}{2}\|x(t) - z\|^2 + \left(qt^{2q-1}\epsilon(t) + \frac{t^{2q}\dot{\epsilon}(t)}{2}\right)\|x(t)\|^2 \\ &+ t^{2q}\epsilon(t)\langle \dot{x}(t), x(t)\rangle. \end{aligned} \tag{12}$$

After expansion we obtain for all $t \geq t_0$

$$\begin{aligned}\dot{\mathcal{E}}(t) =& (b + qt^{q-1} - \alpha)t^q\|\dot{x}(t)\|^2 + \big(\beta(b + qt^{q-1} - \alpha) + (\beta qt^{q-1} - t^q)\big)\, t^q\langle \dot{x}(t), A_{\lambda(t)}(x(t))\rangle \\ &+ b(\beta qt^{q-1} - t^q)\left\langle A_{\lambda(t)}(x(t)), x(t) - z\right\rangle + \beta(\beta qt^{q-1} - t^q)t^q\|A_{\lambda(t)}(x(t))\|^2 \\ &- bt^q\epsilon(t)\langle x(t), x(t) - z\rangle - \beta t^{2q}\epsilon(t)\langle A_{\lambda(t)}(x(t)), x(t)\rangle \\ &+ \frac{bq(1-q)t^{q-2}}{2}\|x(t) - z\|^2 + \left(qt^{2q-1}\epsilon(t) + \frac{t^{2q}\dot{\epsilon}(t)}{2}\right)\|x(t)\|^2. \end{aligned} \tag{13}$$

Now, obviously, there exists $t_1 \geq t_0'$ such that $b(\beta qt^{q-1} - t^q) \leq 0$ for all $t \geq t_1$, hence by using the cocoerciveness of $A_{\lambda(t)}$ we obtain

$$b(\beta qt^{q-1} - t^q)\left\langle A_{\lambda(t)}(x(t)), x(t) - z\right\rangle \leq b(\beta qt^{q-1} - t^q)\lambda(t)\left\|A_{\lambda(t)}(x(t))\right\|^2.$$

Further,

$$-bt^q\epsilon(t)\langle x(t), x(t) - z\rangle = \frac{bt^q\epsilon(t)}{2}(\|z\|^2 - \|x(t)\|^2 - \|x(t) - z\|^2),$$

therefore (13) leads for all $t \geq t_1$ to

$$\begin{aligned}\dot{\mathcal{E}}(t) \leq& (b + qt^{q-1} - \alpha)t^q\|\dot{x}(t)\|^2 + \big(\beta(b + qt^{q-1} - \alpha) + (\beta qt^{q-1} - t^q)\big)\, t^q\langle \dot{x}(t), A_{\lambda(t)}(x(t))\rangle \\ &+ (\beta qt^{q-1} - t^q)(\beta t^q + b\lambda(t))\|A_{\lambda(t)}(x(t))\|^2 - \beta t^{2q}\epsilon(t)\langle A_{\lambda(t)}(x(t)), x(t)\rangle \\ &+ \frac{bq(1-q)t^{q-2} - bt^q\epsilon(t)}{2}\|x(t) - z\|^2 + \left(\frac{2qt^{2q-1} - bt^q}{2}\epsilon(t) + \frac{t^{2q}\dot{\epsilon}(t)}{2}\right)\|x(t)\|^2 + \frac{bt^q\epsilon(t)}{2}\|z\|^2. \end{aligned} \tag{14}$$

Case $q < 1$. In this case $b < \alpha$ and, according to the assumptions, $\epsilon(t) \geq \frac{q(1-q)}{t^2}$ for t large enough. Hence, there exists $t_2 \geq t_1$ such that $\frac{2qt^{2q-1} - bt^q}{2} \leq 0$ and $\frac{bq(1-q)t^{q-2} - bt^q\epsilon(t)}{2} \leq 0$ for all $t \geq t_2$. Consequently, for all $t \geq t_2$ we have the following estimate

$$-\beta t^{2q}\epsilon(t)\langle A_{\lambda(t)}(x(t)), x(t)\rangle \leq \frac{bt^q - 2qt^{2q-1}}{2}\epsilon(t)\|x(t)\|^2 + \frac{\beta^2 t^{4q}\epsilon(t)}{2(bt^q - 2qt^{2q-1})}\|A_{\lambda(t)}(x(t))\|^2.$$

By neglecting the nonpositive terms $\frac{t^{2q}\dot{\epsilon}(t)}{2}\|x(t)\|^2$ (ϵ is nonincreasing) and $\frac{bq(1-q)t^{q-2} - bt^q\epsilon(t)}{2}\|x(t) - z\|^2$, (14) yields for all $t \geq t_2$

$$\begin{aligned}\dot{\mathcal{E}}(t) \leq& (b + qt^{q-1} - \alpha)t^q\|\dot{x}(t)\|^2 + \big(\beta(b + qt^{q-1} - \alpha) + (\beta qt^{q-1} - t^q)\big)\, t^q\langle \dot{x}(t), A_{\lambda(t)}(x(t))\rangle \\ &+ \left((\beta qt^{q-1} - t^q)(\beta t^q + b\lambda(t)) + \frac{\beta^2 t^{4q}\epsilon(t)}{2(bt^q - 2qt^{2q-1})}\right)\|A_{\lambda(t)}(x(t))\|^2 + \frac{bt^q\epsilon(t)}{2}\|z\|^2. \end{aligned} \tag{15}$$

Since in this case $\lambda > \frac{1}{\alpha^2}$, we can choose $c_0 > 1$ such that $\lambda > \frac{c_0}{\alpha^2}$. We show that the coefficient of $\|A_{\lambda(t)}(x(t))\|^2$ in (15) is less than $-\frac{b\lambda}{c_0}t^{3q}$ for t large enough. Indeed, since $q < 1$ one has $\frac{\beta^2 t^{4q}}{2(bt^q - 2qt^{2q-1})} = \frac{\beta^2}{2b}t^{3q} + \mathcal{O}(t^{4q-1})$ as $t \to +\infty$. On the other hand, $(\beta q t^{q-1} - t^q)(\beta t^q + b\lambda(t)) = -b\lambda t^{3q} + \mathcal{O}(t^{2q})$ as $t \to +\infty$. Now, since $\epsilon(t) \to 0$ as $t \to +\infty$, we conclude from here that there exists $t_3 \geq t_2$ such that

$$(\beta q t^{q-1} - t^q)(\beta t^q + b\lambda(t)) + \frac{\beta^2 t^{4q}\epsilon(t)}{2(bt^q - 2qt^{2q-1})} = -b\lambda\left(1 - \frac{\beta^2\epsilon(t)}{2b^2\lambda}\right)t^{3q} + \mathcal{O}(t^{4q-1}) + \mathcal{O}(t^{2q}) \leq -\frac{b\lambda}{c_0}t^{3q}$$

for all $t \geq t_3$. Finally, for $s > 0$ we have the estimate

$$\begin{aligned}&\big(\beta(b + qt^{q-1} - \alpha) + (\beta q t^{q-1} - t^q)\big)\, t^q\langle \dot{x}(t), A_{\lambda(t)}(x(t))\rangle \leq \\ &\frac{\big|\beta(b + qt^{q-1} - \alpha) + (\beta q t^{q-1} - t^q)\big|\, t^q}{2}\left(\frac{s}{t^q}\|\dot{x}(t)\|^2 + \frac{t^q}{s}\|A_{\lambda(t)}(x(t))\|^2\right) \text{ for all } t \geq t_3.\end{aligned} \tag{16}$$

Therefore (15) leads to

$$\begin{aligned}\dot{\mathcal{E}}(t) \leq &\left((b + qt^{q-1} - \alpha)t^q + \frac{\big|\beta(b + qt^{q-1} - \alpha) + (\beta q t^{q-1} - t^q)\big|\, t^q}{2}\frac{s}{t^q}\right)\|\dot{x}(t)\|^2 \\ &+ \left(-\frac{b\lambda}{c_0}t^{3q} + \frac{\big|\beta(b + qt^{q-1} - \alpha) + (\beta q t^{q-1} - t^q)\big|\, t^q}{2}\frac{t^q}{s}\right)\|A_{\lambda(t)}(x(t))\|^2 + \frac{bt^q\epsilon(t)}{2}\|z\|^2 \text{ for all } t \geq t_3.\end{aligned} \tag{17}$$

Obviously, since $q < 1$, the coefficient of $\|\dot{x}(t)\|^2$ in (17) is $\left(b - \alpha + \frac{s}{2}\right)t^q + \mathcal{O}(t^{2q-1}) + \mathcal{O}(1)$ as $t \to +\infty$. On the other hand, the coefficient of $\|A_{\lambda(t)}(x(t))\|^2$ in (17) is $\left(-\frac{b\lambda}{c_0} + \frac{1}{2s}\right)t^{3q} + \mathcal{O}(t^{2q})$ as $t \to +\infty$. Now, let us choose b such that $\frac{c_0}{2b\lambda} < 2(\alpha - b)$. This is possible since, as $\lambda > \frac{c_0}{\alpha^2}$, the inequality $-4\lambda b^2 + 4\alpha\lambda b - c_0 > 0$ has solutions in the interval $(0, \alpha)$. Further, choose s such that $\frac{c_0}{2b\lambda} < s < 2(\alpha - b)$. In conclusion, there exist $t_4 \geq t_3$ and $K_1, K_2 > 0$ such that

$$\dot{\mathcal{E}}(t) \leq -K_1 t^q\|\dot{x}(t)\|^2 - K_2 t^{3q}\|A_{\lambda(t)}(x(t))\|^2 + \frac{bt^q\epsilon(t)}{2}\|z\|^2 \text{ for all } t \geq t_4. \tag{18}$$

Case $q = 1$**.** In this case (14) reads

$$\begin{aligned}\dot{\mathcal{E}}(t) \leq\ &(b + 1 - \alpha)t\|\dot{x}(t)\|^2 + (\beta(b + 2 - \alpha) - t)\, t\langle \dot{x}(t), A_{\lambda(t)}(x(t))\rangle \\ &+ (\beta - t)(\beta t + b\lambda t^2)\|A_{\lambda(t)}(x(t))\|^2 - \beta t^2\epsilon(t)\langle A_{\lambda(t)}(x(t)), x(t)\rangle \\ &- \frac{bt\epsilon(t)}{2}\|x(t) - z\|^2 + \left(\frac{2 - b}{2}t\epsilon(t) + \frac{t^2\dot{\epsilon}(t)}{2}\right)\|x(t)\|^2 + \frac{bt\epsilon(t)}{2}\|z\|^2 \text{ for all } t \geq t_1.\end{aligned} \tag{19}$$

A. First we assume that $\alpha > 3$ and $\lambda > \frac{1}{8(\alpha - 3)}$. We can choose $2 < b < \alpha - 1$, hence $\frac{2-b}{2}t\epsilon(t) \leq 0$ and $-\frac{bt\epsilon(t)}{2} \leq 0$ for all $t \geq t_1$. Consequently, for all $t \geq t_1$ we have the following estimate

$$-\beta t^2\epsilon(t)\langle A_{\lambda(t)}(x(t)), x(t)\rangle \leq \frac{b - 2}{2}t\epsilon(t)\|x(t)\|^2 + \frac{\beta^2 t^3\epsilon(t)}{2(b - 2)}\|A_{\lambda(t)}(x(t))\|^2.$$

Hence, by neglecting the nonpositive terms $\frac{t^2\dot{\epsilon}(t)}{2}\|x(t)\|^2$ and $-\frac{bt\epsilon(t)}{2}\|x(t) - z\|^2$, (19) yields

$$\begin{aligned}\dot{\mathcal{E}}(t) \leq\ &(b + 1 - \alpha)t\|\dot{x}(t)\|^2 + (\beta(b + 2 - \alpha) - t)\, t\langle \dot{x}(t), A_{\lambda(t)}(x(t))\rangle \\ &+ \left((\beta - t)(\beta t + b\lambda t^2) + \frac{\beta^2 t^3\epsilon(t)}{2(b - 2)}\right)\|A_{\lambda(t)}(x(t))\|^2 + \frac{bt\epsilon(t)}{2}\|z\|^2 \text{ for all } t \geq t_1.\end{aligned} \tag{20}$$

Let us choose $c_0 > 1$ such that $\lambda > \frac{c_0}{8(\alpha-3)}$. Since $\epsilon(t) \to 0$ as $t \to +\infty$ we conclude that there exists $t_2 \geq t_1$ such that

$$(\beta - t)(\beta t + b\lambda t^2) + \frac{\beta^2 t^3 \epsilon(t)}{2(b-2)} \leq -\frac{b\lambda}{c_0} t^3 \text{ for all } t \geq t_2.$$

Finally, for $s > 0$ we have the estimate

$$(\beta(b+2-\alpha) - t)\, t\langle \dot{x}(t), A_{\lambda(t)}(x(t))\rangle \leq \frac{|\beta(b+2-\alpha)-t|\, t}{2}\left(\frac{s}{t}\|\dot{x}(t)\|^2 + \frac{t}{s}\|A_{\lambda(t)}(x(t))\|^2\right) \text{ for all } t \geq t_2. \tag{21}$$

Therefore (20) leads to

$$\begin{aligned}\dot{\mathcal{E}}(t) \leq & \left((b+1-\alpha)t + \frac{|\beta(b+2-\alpha)-t|\, t}{2}\frac{s}{t}\right)\|\dot{x}(t)\|^2 \\ & + \left(-\frac{b\lambda}{c_0}t^3 + \frac{|\beta(b+2-\alpha)-t|\, t}{2}\frac{t}{s}\right)\|A_{\lambda(t)}(x(t))\|^2 + \frac{bt\epsilon(t)}{2}\|z\|^2 \text{ for all } t \geq t_2.\end{aligned} \tag{22}$$

The coefficient of $\|\dot{x}(t)\|^2$ in (22) is $\left(b+1-\alpha+\frac{s}{2}\right)t + \mathcal{O}(1)$ as $t \to +\infty$. On the other hand, the coefficient of $\|A_{\lambda(t)}(x(t))\|^2$ in (22) is $\left(-\frac{b\lambda}{c_0} + \frac{1}{2s}\right)t^3 + \mathcal{O}(t^2)$ as $t \to +\infty$. Now, let us choose b such that $\frac{c_0}{2b\lambda} < 2(\alpha - 1 - b)$, which is possible since in this case $\lambda > \frac{c_0}{8(\alpha-3)} \geq \frac{c_0}{(\alpha-1)^2}$ hence the inequality $-4\lambda b^2 + 4(\alpha-1)\lambda b - c_0 > 0$ has solutions in the interval $(2, \alpha-1)$. Fuhrther, we choose s such that $\frac{c_0}{2b\lambda} < s < 2(\alpha-1-b)$. In conclusion, there exist $t_4 \geq t_2$ and $K_1, K_2 > 0$ such that

$$\dot{\mathcal{E}}(t) \leq -K_1 t\|\dot{x}(t)\|^2 - K_2 t^3\|A_{\lambda(t)}(x(t))\|^2 + \frac{bt\epsilon(t)}{2}\|z\|^2 \text{ for all } t \geq t_4. \tag{23}$$

B. Now we assume that $\alpha > 1$, $\lambda > \frac{1}{(\alpha-1)^2}$ and $\frac{\dot{\epsilon}(t)}{\epsilon(t)} \leq -\frac{2}{t}$ for t large enough. In this case $0 < b < \alpha - 1$ and there exists $t_2 \geq t_1$ such that

$$\frac{2-b}{2}t\epsilon(t) + \frac{t^2\dot{\epsilon}(t)}{2} < 0 \text{ for all } t \geq t_2.$$

Consequently, for all $t \geq t_2$ we have the following estimate

$$-\beta t^2\epsilon(t)\langle A_{\lambda(t)}(x(t)), x(t)\rangle \leq \left(\frac{b-2}{2}t\epsilon(t) - \frac{t^2\dot{\epsilon}(t)}{2}\right)\|x(t)\|^2 + \frac{\beta^2 t^3\epsilon^2(t)}{2\left((b-2)\epsilon(t) - t\dot{\epsilon}(t)\right)}\|A_{\lambda(t)}(x(t))\|^2.$$

Hence, by neglecting the nonpositive term $-\frac{bt\epsilon(t)}{2}\|x(t)-z\|^2$, (19) yields

$$\begin{aligned}\dot{\mathcal{E}}(t) \leq\ & (b+1-\alpha)t\|\dot{x}(t)\|^2 + (\beta(b+2-\alpha)-t)\, t\langle \dot{x}(t), A_{\lambda(t)}(x(t))\rangle \\ & + \left((\beta-t)(\beta t + b\lambda t^2) + \frac{\beta^2 t^3\epsilon^2(t)}{2\left((b-2)\epsilon(t) - t\dot{\epsilon}(t)\right)}\right)\|A_{\lambda(t)}(x(t))\|^2 + \frac{bt\epsilon(t)}{2}\|z\|^2 \text{ for all } t \geq t_2.\end{aligned} \tag{24}$$

Let us choose $c_0 > 1$ such that $\lambda > \frac{c_0}{(\alpha-1)^2}$. Since $-2\epsilon(t) - t\dot{\epsilon}(t) \geq 0$ for all $t \geq t_2$ and $\epsilon(t) \to 0$ as $t \to +\infty$, we conclude that there exists $t_3 \geq t_2$ such that

$$(\beta-t)(\beta t + b\lambda t^2) + \frac{\beta^2 t^3\epsilon^2(t)}{2\left((b-2)\epsilon(t) - t\dot{\epsilon}(t)\right)} \leq (\beta-t)(\beta t + b\lambda t^2) + \frac{\beta^2 t^3\epsilon(t)}{2b} \leq -\frac{b\lambda}{c_0}t^3 \text{ for all } t \geq t_3.$$

Combining the above relation with (24) and (21) we get that for all $s > 0$ one has

$$\begin{aligned}\dot{\mathcal{E}}(t) \leq & \left((b+1-\alpha)t + \frac{|\beta(b+2-\alpha)-t|\, t}{2}\frac{s}{t}\right)\|\dot{x}(t)\|^2 \\ & + \left(-\frac{b\lambda}{c_0}t^3 + \frac{|\beta(b+2-\alpha)-t|\, t}{2}\frac{t}{s}\right)\|A_{\lambda(t)}(x(t))\|^2 + \frac{bt\epsilon(t)}{2}\|z\|^2 \text{ for all } t \geq t_3.\end{aligned} \tag{25}$$

Now, let us choose b such that $\frac{c_0}{2b\lambda} < 2(\alpha - 1 - b)$, which is possible since in this case $\lambda > \frac{c_0}{(\alpha-1)^2}$, hence the inequality $-4\lambda b^2 + 4(\alpha-1)\lambda b - c_0 > 0$ has solutions in the interval $(0, \alpha - 1)$. Further, we choose s such that $\frac{c_0}{2b\lambda} < s < 2(\alpha - 1 - b)$. In conclusion, there exist $t_4 \geq t_3$ and $K_1, K_2 > 0$ such that

$$\dot{\mathcal{E}}(t) \leq -K_1 t \|\dot{x}(t)\|^2 - K_2 t^3 \|A_{\lambda(t)}(x(t))\|^2 + \frac{bt\epsilon(t)}{2}\|z\|^2 \text{ for all } t \geq t_4. \tag{26}$$

In other words, from (18), (23) and (26) we conclude that for all $0 < q \leq 1$ there exist $K_1, K_2 > 0$ and $t_4 \geq t_0$ such that

$$\dot{\mathcal{E}}(t) + K_1 t^q \|\dot{x}(t)\|^2 + K_2 t^{3q} \|A_{\lambda(t)}(x(t))\|^2 \leq \frac{bt^q\epsilon(t)}{2}\|z\|^2 \text{ for all } t \geq t_4. \tag{27}$$

Convergence rates. We consider $T > t_4$ and integrate (27) on the interval $[t_4, T]$. By taking into account that $t^q\epsilon(t) \in L^1[t_0, +\infty)$ we obtain that there exists $C > 0$ such that

$$\mathcal{E}(T) + K_1 \int_{t_4}^{T} t^q \|\dot{x}(t)\|^2 dt + K_2 \int_{t_4}^{T} t^{3q} \|A_{\lambda(t)}(x(t))\|^2 dt \leq C. \tag{28}$$

From (28) and the definition of $\mathcal{E}(t)$ we deduce that $x(t)$ is bounded,

$$\sup_{t \geq t_0} \|b(x(t) - z) + t^q(\dot{x}(t) + \beta A_{\lambda(t)}(x(t)))\|^2 < +\infty \tag{29}$$

and

$$\int_{t_0}^{+\infty} t^q \|\dot{x}(t)\|^2 dt < +\infty \text{ and } \int_{t_0}^{+\infty} t^{3q} \|A_{\lambda(t)}(x(t))\|^2 dt < +\infty. \tag{30}$$

Since $A_{\lambda(t)}$ is $\frac{1}{\lambda(t)}$ Lipschitz continuous and $z \in S$, we have for all $t \geq t_0$

$$\|A_{\lambda(t)}(x(t))\| = \|A_{\lambda(t)}(x(t)) - A_{\lambda(t)}(z)\| \leq \frac{1}{\lambda(t)}\|x(t) - z\|.$$

Taking into account that $\lambda(t) = \lambda t^{2q}$ and $\|x(t) - z\|$ is bounded, we deduce that

$$\|A_{\lambda(t)}(x(t))\| = \mathcal{O}\left(\frac{1}{t^{2q}}\right) \text{ as } t \to +\infty \tag{31}$$

and further, using (29),

$$\|\dot{x}(t)\| = \mathcal{O}\left(\frac{1}{t^q}\right) \text{ as } t \to +\infty. \tag{32}$$

Now, according to Lemma A.2 in the Appendix, we have

$$\left\|\frac{d}{dt}\lambda(t)A_{\lambda(t)}(x(t))\right\| \leq 2\|\dot{x}(t)\| + 2\frac{|\lambda'(t)|}{\lambda(t)}\|x(t) - z\| = \mathcal{O}\left(\frac{1}{t^q}\right) \text{ as } t \to +\infty. \tag{33}$$

This relation combined with (31) leads to $\left\|\lambda(t)\frac{d}{dt}A_{\lambda(t)}(x(t))\right\| = \mathcal{O}\left(\frac{1}{t^q}\right)$ as $t \to +\infty$. Consequently,

$$\left\|\frac{d}{dt}A_{\lambda(t)}(x(t))\right\| = \mathcal{O}\left(\frac{1}{t^{3q}}\right) \text{ as } t \to +\infty. \tag{34}$$

Next we will improve the estimate in (31) and show that actually

$$\|A_{\lambda(t)}(x(t))\| = o\left(\frac{1}{t^{2q}}\right) \text{ as } t \to +\infty.$$

For all $t \geq t_0$ we have

$$\left|\frac{d}{dt}\|\lambda(t)A_{\lambda(t)}(x(t))\|^4\right| = 4\left|\left\langle \lambda(t)A_{\lambda(t)}(x(t)), \frac{d}{dt}(\lambda(t)A_{\lambda(t)}(x(t)))\right\rangle\right| \|\lambda(t)A_{\lambda(t)}(x(t))\|^2. \tag{35}$$

According to (31) and (33) there exists $K > 0$ such that

$$4\left|\left\langle \lambda(t)A_{\lambda(t)}(x(t)), \frac{d}{dt}(\lambda(t)A_{\lambda(t)}(x(t)))\right\rangle\right| \leq 4\left\|\lambda(t)A_{\lambda(t)}(x(t))\right\| \left\|\frac{d}{dt}(\lambda(t)A_{\lambda(t)}(x(t)))\right\| \leq \frac{K}{t^q},$$

hence (35) leads to

$$\left|\frac{d}{dt}\|\lambda(t)A_{\lambda(t)}(x(t))\|^4\right| \leq \frac{K}{t^q}\|\lambda(t)A_{\lambda(t)}(x(t))\|^2 \text{ for all } t \geq t_0.$$

According to (30), the term on the right-hand side of the above relation belongs to $L^1([t_0, +\infty), \mathbb{R})$, which implies

$$\frac{d}{dt}\|\lambda(t)A_{\lambda(t)}(x(t))\|^4 \in L^1([t_0, +\infty), \mathbb{R}).$$

Therefore, according to Lemma A.4 in the Appendix,

$$\lim_{t\to+\infty} \|\lambda(t)A_{\lambda(t)}(x(t))\|^4 \text{ exists.}$$

hence, $L := \lim_{t\to+\infty} \|\lambda(t)A_{\lambda(t)}(x(t))\|^2$ exists as well. Using again (30), that is

$$\int_{t_0}^{+\infty} \frac{1}{t^q}\|\lambda(t)A_{\lambda(t)}(x(t))\|^2 dt = \lambda^2 \int_{t_0}^{+\infty} t^{3q}\|A_{\lambda(t)}(x(t))\|^2 dt < +\infty,$$

we deduce that $L = 0$. Therefore, $\lim_{t\to+\infty} \|\lambda(t)A_{\lambda(t)}(x(t))\| = 0$, which gives

$$\|A_{\lambda(t)}(x(t))\| = o\left(\frac{1}{t^{2q}}\right) \text{ as } t \to +\infty. \tag{36}$$

Further, by using (again DIN-AVD-TIKH), we have for all $t \geq t_0$

$$\|\ddot{x}(t)\|^2 = \left\|\frac{\alpha}{t^q}\dot{x}(t) + \beta\frac{d}{dt}A_{\lambda(t)}(x(t)) + A_{\lambda(t)}(x(t)) + \epsilon(t)x(t)\right\|^2,$$

which leads via the Cauchy-Schwarz inequality to

$$t^{3q}\|\ddot{x}(t)\|^2 \leq 4t^{3q}\epsilon^2(t)\|x(t)\|^2 + 4\alpha^2 t^q\|\dot{x}(t)\|^2 + 4\beta^2 t^{3q}\left\|\frac{d}{dt}A_{\lambda(t)}(x(t))\right\|^2 + 4t^{3q}\left\|A_{\lambda(t)}(x(t))\right\|^2. \tag{37}$$

Next we will show that

$$t^{3q}\left\|\frac{d}{dt}A_{\lambda(t)}(x(t))\right\|^2 \in L^1([t_0, +\infty), \mathbb{R}). \tag{38}$$

Indeed, according to Lemma A.2 (c2) in the Appendix we have for all $t \geq t_0$

$$t^{3q}\left\|\frac{d}{dt}A_{\lambda(t)}(x(t))\right\|^2 \leq t^{3q}\left(\frac{2}{\lambda t^{2q}}\|\dot{x}(t)\| + \frac{4q}{t}\|A_{\lambda(t)}(x(t))\|\right)^2 \leq \frac{8}{\lambda^2 t^q}\|\dot{x}(t)\|^2 + \frac{32q^2}{t^{2-3q}}\|A_{\lambda(t)}(x(t))\|^2$$

and the claim follows via (30).

Finally, suppose that $\int_{t_0}^{+\infty} t^{3q}\epsilon^2(t)dt < +\infty$. Using that $\|x(t)\|$ is bounded, from (30), (38) and (37) we obtain that

$$\int_{t_0}^{+\infty} t^{3q}\|\ddot{x}(t)\|^2 dt < +\infty. \tag{39}$$

Next we prove that $\|\dot{x}(t)\| = o\left(\frac{1}{t^q}\right)$ as $t \to +\infty$. For all $t \geq t_0$ we have

$$\frac{d}{dt}t^{2q}\|\dot{x}(t)\|^2 = 2qt^{2q-1}\|\dot{x}(t)\|^2 + 2t^{2q}\langle \ddot{x}(t), \dot{x}(t)\rangle$$

and

$$2t^{2q}\langle \ddot{x}(t), \dot{x}(t)\rangle \leq t^{3q}\|\ddot{x}(t)\|^2 + t^q\|\dot{x}(t)\|^2,$$

hence,

$$\frac{d}{dt}t^{2q}\|\dot{x}(t)\|^2 \leq t^{3q}\|\ddot{x}(t)\|^2 + (2qt^{2q-1} + t^q)\|\dot{x}(t)\|^2.$$

According to (39) and (30) we have $t \mapsto t^{3q}\|\ddot{x}(t)\|^2 + (2qt^{2q-1} + t^q)\|\dot{x}(t)\|^2 \in L^1([t_0, +\infty), \mathbb{R})$. Therefore, using again Lemma A.4 in the Appendix, we get that there exists $\lim_{t\to+\infty} t^{2q}\|\dot{x}(t)\|^2 \in \mathbb{R}$. Using (30) again, we have

$$\int_{t_0}^{\infty} \frac{1}{t^q}(t^{2q}\|\dot{x}(t)\|^2)dt = \int_{t_0}^{\infty} t^q\|\dot{x}(t)\|^2 dt < +\infty,$$

hence,

$$\lim_{t\to+\infty} t^{2q}\|\dot{x}(t)\|^2 = 0.$$

In other words,

$$\|\dot{x}(t)\| = o\left(\frac{1}{t^q}\right) \text{ as } t \to +\infty.$$

Weak convergence of $x(t)$ as $t \to +\infty$. We assume that $\beta = 0$ or $\beta > 0$ and $q > \frac{1}{2}$. For $z \in S$ we introduce the anchor function

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

The classical derivation chain rule gives for all $t \geq t_0$

$$\ddot{h}_z(t) + \frac{\alpha}{t^q}\dot{h}_z(t) = \left\langle \ddot{x}(t) + \frac{\alpha}{t^q}\dot{x}(t), x(t) - z\right\rangle + \|\dot{x}(t)\|^2.$$

By using (DIN-AVD-TIKH), the monotonicity of $A_{\lambda(t)}$ and the Cauchy-Schwarz inequality we get for all $t \geq t_0$

$$\begin{aligned}
\ddot{h}_z(t) + \frac{\alpha}{t^q}\dot{h}_z(t) &= \left\langle -\beta\frac{d}{dt}A_{\lambda(t)}(x(t)) - A_{\lambda(t)}(x(t)) - \epsilon(t)x(t), x(t) - z\right\rangle + \|\dot{x}(t)\|^2 \\
&= \|\dot{x}(t)\|^2 - \beta\left\langle \frac{d}{dt}A_{\lambda(t)}(x(t)), x(t) - z\right\rangle - \langle A_{\lambda(t)}(x(t)), x(t) - z\rangle - \epsilon(t)\langle x(t), x(t) - z\rangle \\
&\leq \|\dot{x}(t)\|^2 + \beta\left\|\frac{d}{dt}A_{\lambda(t)}(x(t))\right\|\|x(t) - z\| + \frac{\epsilon(t)}{4}\|z\|^2.
\end{aligned} \tag{40}$$

Next we show that $t \mapsto t^q k(t) \in L^1([t_0, +\infty), \mathbb{R})$, where $k(t) := \|\dot{x}(t)\|^2 + \beta\left\|\frac{d}{dt}A_{\lambda(t)}(x(t))\right\|\|x(t) - z\| + \frac{\epsilon(t)}{4}\|z\|^2$ is the function on the right-hand side of (40). Indeed, according to the hypotheses we have $t^q\epsilon(t) \in L^1([t_0, +\infty), \mathbb{R})$. Further, from (30), we have $t^q\|\dot{x}(t)\|^2 \in L^1([t_0, +\infty), \mathbb{R})$. Therefore, if $\beta = 0$, we have $t \mapsto t^q k(t) \in L^1([t_0, +\infty), \mathbb{R})$.

Now, assume that $\beta > 0$ and $q > \frac{1}{2}$. According to (34) there exists $K > 0$ such that $\left\|\frac{d}{dt}A_{\lambda(t)}(x(t))\right\| \leq \frac{K}{t^{3q}}$ for t large enough, hence,

$$\beta t^q \left\|\frac{d}{dt}A_{\lambda(t)}(x(t))\right\| \|x(t) - z\| \leq \beta K\|x(t) - z\|\frac{1}{t^{2q}}. \tag{41}$$

Since $\|x(t) - z\|$ is bounded and $q > \frac{1}{2}$, we have $t \mapsto K\|x(t) - z\|\frac{1}{t^{2q}} \in L^1([t_0, +\infty), \mathbb{R})$, consequently, also in this case, $t \mapsto t^q k(t) \in L^1([t_0, +\infty), \mathbb{R})$.

If $q < 1$ we can apply [15, Lemma 13] and use the argumentation in [15, Lemma 6] to obtain that $\lim_{t\to+\infty} h_z(t)$ exists. If $q = 1$ we can apply [7, Lemma A.6] to get the same conclusion. In other words,

$$\lim_{t\to+\infty} \|x(t) - z\| \text{ exists for all } z \in S,$$

which means that the first condition in Opial Lemma (Lemma A.5 in the Appendix) is satisfied. Next we will show that the second condition is also satisfied, namely that every weak sequential cluster point of $x(t)$ belongs to S. To this end we will use that the Yosida approximation of a maximally monotone operator fulfils

$$A_\lambda(x) \in A(x - \lambda A_\lambda(x)) \text{ for all } x \in \mathcal{H} \text{ and } \lambda > 0. \tag{42}$$

Let x^* be a weak sequentially cluster point of $x(t)$. Then, there exists a sequence $t_n \to +\infty$ such that $x(t_n)$ converges weakly to x^* as $n \to +\infty$. Since the graph of A is sequentially closed in the weak×strong topology, by using (36), we have

$$0 = \lim_{n\to+\infty} A_{\lambda(t_n)}(x(t_n)) \in A\left(\lim_{n\to+\infty}(x(t_n) - \lambda(t_n)A_{\lambda(t_n)}(x(t_n))\right) = A(x^*).$$

Consequently, $x(t)$ converges weakly to an element of S. ■

Remark 3. The assumptions imposed on the Tikhonov parametrization function ϵ in Theorem 2 are in concordance with the investigations made in [11] and [15]. Indeed, the condition $t \mapsto t^q\epsilon(t) \in L^1([t_0, +\infty), \mathbb{R})$ recovers the condition used in [11] in case $q = 1$ in order to obtain weak convergence.

For the choice $\epsilon(t) := \frac{a}{t^p}$, with $a, p > 0$, this integrability condition is nothing else than $p > q + 1$. In case $q = 1$ this leads to $p > 2$ which has as consequence that $\frac{\dot\epsilon(t)}{\epsilon(t)} \leq -\frac{2}{t}$ holds for t large enough. On the other hand, in case $q < 1$, $\epsilon(t) \geq \frac{q(1-q)}{t^2}$ holds for t large enough if and only if $p < 2$ or $p = 2$ and $a \geq q(1 - q)$. These are exactly the conditions under which weak convergence of the trajectories of a Tikhonov regularized dynamical system has been obtained in [15].

Theorem 2 for the choice $\epsilon(t) := \frac{a}{t^p}$, with $a, p > 0$, leads to the following result.

Corollary 4. *Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator such that $S := \{x \in \mathcal{H} : 0 \in Ax\}$ is nonempty. Consider the evolution equation* (DIN-AVD-TIKH) *with $\beta \geq 0, \lambda(t) = \lambda t^{2q}$, for $\lambda > 0$ and $0 < q \leq 1$, $\epsilon(t) = \frac{a}{t^p}$, for $a, p > 0$ with $p \geq q + 1$, and the system parameters satisfying the following conditions:*

$$\textit{if } q < 1, \textit{ then } \lambda > \frac{1}{\alpha^2};$$

$$\textit{if } q = 1, \textit{ then either } \alpha > 3 \textit{ and } \lambda > \frac{1}{8(\alpha - 3)} \textit{ or } \alpha > 1 \textit{ and } \lambda > \frac{1}{(\alpha - 1)^2}.$$

Then, for any trajectory $x : [t_0, +\infty) \to \mathcal{H}$ of (DIN-AVD-TIKH)*, the following statements are true:*

(i) If $q+1 < p < 2$ or $q+1 < p = 2$ and $a \geq q(1-q)$ or $q = 1$ and $p > 2$, then the statements (i)-(iii) in Theorem 2 hold, including $\int_{t_0}^{+\infty} t^{3q}\|\ddot x(t)\|^2 dt < +\infty$ and $\|\dot x(t)\| = o\left(\frac{1}{t^q}\right)$ as $t \to +\infty$.

(ii) If $p = q+1$, then $\lim_{s\to+\infty} \frac{\int_{t_0}^s t^q\|\dot x(t)\|^2 dt}{\ln s} < +\infty$ and $\lim_{s\to+\infty} \frac{\int_{t_0}^s t^{3q}\|A_{\lambda(t)}(x(t))\|^2 dt}{\ln s} < +\infty$. Furthermore, $\|\dot x(t)\| = \mathcal{O}\left(\frac{\sqrt{\ln t}}{t^q}\right)$, $\|A_{\lambda(t)}(x(t))\| = \mathcal{O}\left(\frac{\sqrt{\ln t}}{t^{2q}}\right)$ and $\left\|\frac{d}{dt}A_{\lambda(t)}(x(t))\right\| = \mathcal{O}\left(\frac{\sqrt{\ln t}}{t^{3q}}\right)$ as $t \to +\infty$.

Proof. (i) According to Remark 3, if $q+1<p$ all the conditions in the hypotheses of Theorem 2 are satisfied, hence its conclusion holds. Even more, since in this case $2p-3q>1$, it holds $\int_{t_0}^{+\infty} t^{3q}\epsilon^2(t)dt = a^2\int_{t_0}^{+\infty} t^{3q-2p}dt < +\infty$.

(ii) Now, consider the case $p=q+1$. The relation (27) that we derived in the proof of Theorem 2 reads

$$\dot{\mathcal{E}}(t) + K_1 t^q\|\dot{x}(t)\|^2 + K_2 t^{3q}\|A_{\lambda(t)}(x(t))\|^2 \le \frac{ab}{2t}\|z\|^2 \text{ for all } t\ge t_4. \tag{43}$$

We choose $T>t_4$, integrate (43) on the interval $[t_4,T]$ and obtain that there exists $C_1,C_2>0$ such that

$$\mathcal{E}(T) + K_1\int_{t_4}^{T} t^q\|\dot{x}(t)\|^2dt + K_2\int_{t_4}^{T} t^{3q}\|A_{\lambda(t)}(x(t))\|^2dt \le C_1\ln T + C_2. \tag{44}$$

From (44) and the definition of $\mathcal{E}$ we deduce that $\sup_{t\ge t_0}\frac{\|x(t)-z\|^2}{\ln t} < +\infty$ and

$$\sup_{t\ge t_0}\frac{\|b(x(t)-z)+t^q(\dot{x}(t)+\beta A_{\lambda(t)}(x(t)))\|^2}{\ln t} < +\infty. \tag{45}$$

In addition,

$$\lim_{s\to+\infty}\frac{\int_{t_0}^{s} t^q\|\dot{x}(t)\|^2dt}{\ln s} < +\infty \text{ and } \lim_{s\to+\infty}\frac{\int_{t_0}^{s} t^{3q}\|A_{\lambda(t)}(x(t))\|^2dt}{\ln s} < +\infty. \tag{46}$$

Now, since $A_{\lambda(t)}$ is $\frac{1}{\lambda(t)}$ Lipschitz continuous and $z\in S$ one has for all $t\ge t_0$

$$\|A_{\lambda(t)}(x(t))\| = \|A_{\lambda(t)}(x(t)) - A_{\lambda(t)}(z)\| \le \frac{1}{\lambda(t)}\|x(t)-z\|.$$

Taking into account that $\lambda(t)=\lambda t^{2q}$ and $\|x(t)-z\|^2 = \mathcal{O}(\ln t)$ as $t\to+\infty$, we deduce that

$$\|A_{\lambda(t)}(x(t))\| = \mathcal{O}\left(\frac{\sqrt{\ln t}}{t^{2q}}\right) \text{ as } t\to+\infty. \tag{47}$$

Using (47), from (45) we obtain that

$$\|\dot{x}(t)\| = \mathcal{O}\left(\frac{\sqrt{\ln t}}{t^{q}}\right) \text{ as } t\to+\infty. \tag{48}$$

Finally, as in the proof of Theorem 2 one can show that

$$\left\|\frac{d}{dt}A_{\lambda(t)}(x(t))\right\| = \mathcal{O}\left(\frac{\sqrt{\ln t}}{t^{3q}}\right) \text{ as } t\to+\infty.$$

■

3 Strong convergence and fast rates

In order to obtain both fast rates of convergence and strong convergence of the trajectories to the minimal norm solution, we will use a different energy functional in which we will replace the anchor $z\in S$ with $x_t = (A_{\lambda(t)} + \epsilon(t)\,\mathrm{Id})^{-1}(0)$. Here we rely on an approach proposed in [2] for smooth convex optimization problems. In addition, we will strengthen the conditions imposed on the system parameters in Section 2.

It is well known that if A is a maximally monotone operator such that $S := \{x \in \mathcal{H} : 0 \in Ax\}$ is nonempty and if $t \mapsto \epsilon(t)$ is a positive function such that $\epsilon(t) \longrightarrow 0$ as $t \longrightarrow +\infty$, then $x_{\epsilon(t)}$ converges to $x^* = \operatorname{proj}_S 0$ strongly in $\mathcal{H}$ as $t \to +\infty$, where $x_{\epsilon(t)}$ denotes the unique zero of the strongly monotone operator $A + \epsilon(t)\operatorname{Id}$, that is $Ax_{\epsilon(t)} + \epsilon(t)x_{\epsilon(t)} = 0$.

We will prove a similar result for the Yosida regularization of A. To this end we will denote by $x_t \in \mathcal{H}$ the unique element such that

$$A_{\lambda(t)}x_t + \epsilon(t)x_t = 0. \tag{49}$$

Actually, x_t depends on both $\epsilon(t)$ and $\lambda(t)$, however we prefer for simpler notation x_t than $x_{\epsilon(t),\lambda(t)}$.

Proposition 5. *Let A be a maximally monotone operator such that $S := \{x \in \mathcal{H} : 0 \in Ax\}$ is nonempty. Let $t \mapsto \epsilon(t)$ and $t \mapsto \lambda(t)$ be positive functions defined on $[t_0, +\infty)$. Then $\|x_t\| \leq \|x^*\|$ for all $t \geq t_0$.*

Assume further that $\lim_{t\to+\infty} \epsilon(t) = 0$ and $\lim_{t\to+\infty} \epsilon(t)\lambda(t) = 0$. Then x_t converges strongly, as $t \to +\infty$, to x^*.

Proof Let $t \geq t_0$. Since $A_{\lambda(t)}+\epsilon(t)\operatorname{Id}$ is $\epsilon(t)$-strongly monotone, $A_{\lambda(t)}x_t+\epsilon(t)x_t = 0$ and $A_{\lambda(t)}x^*+\epsilon(t)x^* = \epsilon(t)x^*$, we have

$$\langle -\epsilon(t)x^*, x_t - x^*\rangle \geq \epsilon(t)\|x_t - x^*\|^2. \tag{50}$$

Consequently, $\langle -x_t, x_t - x^*\rangle \geq 0$. In other words, $\|x^*\|^2 - \|x_t\|^2 - \|x_t - x^*\|^2 \geq 0$, hence

$$\|x_t\| \leq \|x^*\|. \tag{51}$$

So, the net $(x_t)_{t\geq t_0}$ is bounded and its set of weak sequential cluster points is nonempty.

Assume now that $\lim_{t\to+\infty} \epsilon(t) = 0$ and $\lim_{t\to+\infty} \epsilon(t)\lambda(t) = 0$. First we show that x_t converges weakly, as $t \to +\infty$, to x^*. Let $\tilde{x}$ a weak sequential cluster point of x_t, that is, $\tilde{x}$ is the weak limit of a sequence $(x_{t_n})_{n\geq 0}$. Since $(x_{t_n})_{n\geq 0}$ is bounded, and $\epsilon(t_n) \to 0$ and $\epsilon(t_n)\lambda(t_n) \to 0$ as $n \to +\infty$, it yields that the sequences $(\epsilon(t_n)x_{t_n})_{n\geq 0}$ and $(\epsilon(t_n)\lambda(t_n)x_{t_n})_{n\geq 0}$ converge strongly to zero as $n \to +\infty$. Using again that the graph of A is sequentially closed in the weak×strong topology and (42), we obtain

$$\begin{aligned} 0 = \lim_{n\to+\infty} A_{\lambda(t_n)}x_{t_n} + \epsilon(t_n)x_{t_n} &= \lim_{n\to+\infty} A_{\lambda(t_n)}x_{t_n} \in A\left(\lim_{n\to+\infty}(x_{t_n} - \lambda(t_n)A_{\lambda(t_n)}x_{t_n})\right) \\ &= A\left(\lim_{n\to+\infty}(x_{t_n} + \lambda(t_n)\epsilon(t_n)x_{t_n})\right) = A(\tilde{x}). \end{aligned} \tag{52}$$

Hence, $\tilde{x} \in S$. On the other hand, from the weak lower semicontinuity of the norm and (51) we get

$$\|\tilde{x}\| \leq \liminf_{n\to+\infty} \|x_{t_n}\| \leq \|x^*\|.$$

Since x^* is the minimum norm element of S, we must have $\tilde{x} = x^*$. As x^* is the limit of every weak convergent subsequence of $(x_t)_{t\geq t_0}$ it follows that x_t converges weakly, as $t \to +\infty$, to x^*. Using again (51), it yields

$$\|x^*\| \leq \liminf_{t\to+\infty} \|x_t\| \leq \limsup_{t\to+\infty} \|x_t\| \leq \|x^*\|,$$

hence,

$$\lim_{t\to+\infty} \|x_t\| = \|x^*\|. \tag{53}$$

In conclusion, x_t converges strongly, as $t \to +\infty$, to x^*. ■

The following estimate of the time derivative of x_t will be useful in our investigations.

Lemma 6. *Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator, $\epsilon : [t_0, +\infty) \to (0, +\infty)$ a nonincreasing function of class C^1 and $\lambda : [t_0, +\infty) \to (0, +\infty)$ a positive function of class C^1. For every $t \in [t_0, +\infty)$ let x_t be the unique zero of the strongly monotone operator $A_{\lambda(t)} + \epsilon(t)\operatorname{Id}$. Then, $t \mapsto x_t$ is almost everywhere differentiable and*

$$\left\| \frac{d}{dt} x_t \right\| \leq \left(-\frac{\dot{\epsilon}(t)}{\epsilon(t)} + \frac{2|\dot{\lambda}(t)|}{\lambda(t)} \right) \|x_t\| \text{ for almost all } t \geq t_0.$$

Proof. First, we will show that $t \mapsto x_t$ is locally absolutely continuous on $[t_0, +\infty)$, which will provide the almost everywhere differentiability of $t \mapsto x_t$. To this end we consider an arbitrary closed interval $[u, v] \subseteq [t_0, +\infty)$ and show that for every $\tau > 0$ there exists $\eta > 0$ such that for any finite family of intervals $I_k = (u_k, v_k) \subseteq [u, v]$ it holds

$$\left(I_k \cap I_j = \emptyset \text{ and } \sum_k |v_k - u_k| < \eta \right) \Rightarrow \sum_k \|x_{v_k} - x_{u_k}\| < \tau.$$

For all $t \geq t_0$ it holds by the definition

$$(A_{\lambda(t)} + \epsilon(t)\operatorname{Id})^{-1}(y) = J_{\frac{1}{\epsilon(t)} A_{\lambda(t)}} \left(\frac{y}{\epsilon(t)} \right)$$

and further, according to [10, Proposition 23.6],

$$J_{\frac{1}{\epsilon(t)} A_{\lambda(t)}} \left(\frac{y}{\epsilon(t)} \right) = \frac{\epsilon(t)\lambda(t)}{1 + \epsilon(t)\lambda(t)} \frac{y}{\epsilon(t)} + \frac{1}{1 + \epsilon(t)\lambda(t)} J_{\left(\frac{1}{\epsilon(t)} + \lambda(t)\right) A} \left(\frac{y}{\epsilon(t)} \right).$$

Now, consider $t, s \in [u, v]$ and let $x_t := (A_{\lambda(t)} + \epsilon(t)\operatorname{Id})^{-1}(0)$ and $x_s := (A_{\lambda(s)} + \epsilon(s)\operatorname{Id})^{-1}(0)$. According to [10, Proposition 23.28], it holds

$$\begin{aligned}
\|x_t - x_s\| &= \left\| \frac{1}{1 + \epsilon(t)\lambda(t)} J_{\left(\frac{1}{\epsilon(t)} + \lambda(t)\right) A}(0) - \frac{1}{1 + \epsilon(s)\lambda(s)} J_{\left(\frac{1}{\epsilon(s)} + \lambda(s)\right) A}(0) \right\| \\
&\leq \left| \frac{1}{1 + \epsilon(t)\lambda(t)} - \frac{1}{1 + \epsilon(s)\lambda(s)} \right| \left\| J_{\left(\frac{1}{\epsilon(t)} + \lambda(t)\right) A}(0) \right\| \\
&+ \frac{1}{1 + \epsilon(s)\lambda(s)} \left\| J_{\left(\frac{1}{\epsilon(t)} + \lambda(t)\right) A}(0) - J_{\left(\frac{1}{\epsilon(s)} + \lambda(s)\right) A}(0) \right\| \\
&\leq \left| \frac{1}{1 + \epsilon(t)\lambda(t)} - \frac{1}{1 + \epsilon(s)\lambda(s)} \right| \left\| J_{\left(\frac{1}{\epsilon(t)} + \lambda(t)\right) A}(0) \right\| \\
&+ \frac{\epsilon(t)}{(1 + \epsilon(s)\lambda(s))(1 + \epsilon(t)\lambda(t))} \left\| J_{\left(\frac{1}{\epsilon(t)} + \lambda(t)\right) A}(0) \right\| \left(\left| \frac{1}{\epsilon(t)} - \frac{1}{\epsilon(s)} \right| + |\lambda(t) - \lambda(s)| \right),
\end{aligned} \tag{54}$$

which means that there exist $A, B > 0$ such that for all $t, s \in [u, v]$ it holds

$$\|x_t - x_s\| \leq A \left| \frac{1}{1 + \epsilon(t)\lambda(t)} - \frac{1}{1 + \epsilon(s)\lambda(s)} \right| + B \left| \frac{1}{\epsilon(t)} - \frac{1}{\epsilon(s)} \right| + B|\lambda(t) - \lambda(s)|. \tag{55}$$

Since the functions $t \mapsto \frac{A}{1+\epsilon(t)\lambda(t)}$, $t \mapsto \frac{B}{\epsilon(t)}$ and $t \mapsto B\lambda(t)$ are absolute continuous on $[u, v]$, for every $\tau > 0$ there exists $\eta > 0$ such that for any finite family of intervals $I_k = (u_k, v_k) \subseteq [u, v]$ we have the implications

$$\left(I_k \cap I_j = \emptyset \text{ and } \sum_k |v_k - u_k| < \eta \right) \Rightarrow \sum_k \left| \frac{A}{1 + \epsilon(v_k)\lambda(v_k)} - \frac{A}{1 + \epsilon(u_k)\lambda(u_k)} \right| < \frac{\tau}{3},$$

$$\sum_k \left| \frac{B}{\epsilon(v_k)} - \frac{B}{\epsilon(u_k)} \right| < \frac{\tau}{3} \text{ and } \sum_k |B\lambda(v_k) - B\lambda(u_k)| < \frac{\tau}{3}.$$

Now, by using (55) we get

$$\sum_k \|x_{v_k} - x_{u_k}\| < \frac{\tau}{3} + \frac{\tau}{3} + \frac{\tau}{3} = \tau,$$

and the claim follows.

In order to prove the second claim we fix $t \in [t_0, +\infty)$ such that $s \mapsto x_s$ is differentiable at t and $h > 0$. We have $\epsilon(t)x_t + A_{\lambda(t)}x_t = 0$ and $\epsilon(t+h)x_{t+h} + A_{\lambda(t+h)}x_{t+h} = 0$, so

$$\langle -\epsilon(t+h)x_{t+h} + \epsilon(t)x_t, x_{t+h} - x_t\rangle + \langle A_{\lambda(t)}x_t - A_{\lambda(t+h)}x_t, x_{t+h} - x_t\rangle = \langle A_{\lambda(t+h)}x_{t+h} - A_{\lambda(t+h)}x_t, x_{t+h} - x_t\rangle.$$

In virtue of the monotonicity of $A_{\lambda(t+h)}$ the right-hand side of the above equality is nonnegative, hence

$$-\langle \epsilon(t+h)x_{t+h} - \epsilon(t)x_t, x_{t+h} - x_t\rangle + \langle A_{\lambda(t)}x_t - A_{\lambda(t+h)}x_t, x_{t+h} - x_t\rangle \geq 0. \tag{56}$$

On the other hand, from the Cauchy-Schwarz inequality we get

$$\langle A_{\lambda(t)}x_t - A_{\lambda(t+h)}x_t, x_{t+h} - x_t\rangle \leq \|A_{\lambda(t)}x_t - A_{\lambda(t+h)}x_t\|\|x_{t+h} - x_t\|$$

and by using Lemma A.2 (b) in the Appendix

$$\|A_{\lambda(t)}x_t - A_{\lambda(t+h)}x_t\| \leq \frac{|\lambda(t) - \lambda(t+h)|}{\lambda(t)}(\|A_\lambda(t)(x_t)\| + \|A_{\lambda(t+h)}(x_t)\|).$$

Plugging this into (56), it yields

$$-\langle \epsilon(t+h)x_{t+h} - \epsilon(t)x_t, x_{t+h} - x_t\rangle + \frac{|\lambda(t) - \lambda(t+h)|}{\lambda(t)}(\|A_\lambda(t)(x_t)\| + \|A_{\lambda(t+h)}(x_t)\|)\|x_{t+h} - x_t\| \geq 0. \tag{57}$$

By dividing (57) by h^2 and letting $h \to 0$ we obtain

$$-\left\langle \frac{d}{dt}(\epsilon(t)x_t), \frac{d}{dt}x_t \right\rangle + 2\frac{|\dot{\lambda}(t)|}{\lambda(t)}\|A_\lambda(t)x_t\| \left\| \frac{d}{dt}x_t \right\| \geq 0.$$

In other words, since $A_\lambda(t)x_t = -\epsilon(t)x_t$, it holds

$$-\dot{\epsilon}(t)\left\langle x_t, \frac{d}{dt}x_t \right\rangle - \epsilon(t)\left\| \frac{d}{dt}x_t \right\|^2 + 2\frac{|\dot{\lambda}(t)|\epsilon(t)}{\lambda(t)}\|x_t\| \left\| \frac{d}{dt}x_t \right\| \geq 0.$$

Using again the Cauchy-Schwarz inequality and the fact that ϵ is nonincreasing, we obtain further that

$$-\dot{\epsilon}(t)\|x_t\| \left\| \frac{d}{dt}x_t \right\| - \epsilon(t)\left\| \frac{d}{dt}x_t \right\|^2 + 2\frac{|\dot{\lambda}(t)|\epsilon(t)}{\lambda(t)}\|x_t\| \left\| \frac{d}{dt}x_t \right\| \geq 0,$$

which provides the desired estimate. ■

Now we state the main result of this section.

Theorem 7. *Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator such that $S := \{x \in \mathcal{H} : 0 \in Ax\}$ is nonempty. Consider the evolution equation* (DIN-AVD-TIKH) *with $\beta \geq 0$ and $\lambda(t) = \lambda t^{2q}$, for $\lambda > \frac{1}{\alpha^2}$*

and $0 < q < 1$. *Assume that* $\int_{t_0}^{+\infty} t^{3q}\epsilon^2(t)dt < +\infty$ *and the Tikhonov parametrization function also satisfies the following conditions:*

(C_0) $\lim_{t\to+\infty} t^{2q}\epsilon(t) = 0$;

(C_1) *There exist* $c_1 > 0$ *and* $\max\{q, 1-2q\} < r < 1$ *such that* $\epsilon(t) \geq \frac{c_1}{t^{r+q}}$ *for* t *large enough;*

(C_2) *There exists* $c_2 > 0$ *such that* $\beta\epsilon(t) \leq \frac{c_2}{t^{1-q}}$ *for* t *large enough;*

(C_3) *There exists* $c_3 > 2q + \frac{16c_2\beta}{\alpha\lambda - \sqrt{\lambda}}$ *such that* $\frac{\dot\epsilon(t)}{\epsilon(t)} \leq -\frac{c_3}{t}$ *for* t *large enough;*

(C_4) *There exists* $c_4 > 0$ *such that* $\left(\frac{\dot\epsilon(t)}{\epsilon(t)}\right)^2 \leq \frac{c_4}{t^2}$ *for* t *large enough.*

Then, for any trajectory $x : [t_0, +\infty[\to \mathcal{H}$ *of* (DIN-AVD-TIKH), *the following statements are true:*

(i) (convergence of the trajectory) $x(t)$ *converges strongly, as* $t \to +\infty$, *to* x^*, *the element of minimum norm of* S;

(ii) (integral estimates) $\int_{t_0}^{+\infty} t^q\|\dot x(t)\|^2dt < +\infty$, $\int_{t_0}^{+\infty} t^{3q}\|A_{\lambda(t)}(x(t))\|^2dt < +\infty$ *and* $\int_{t_0}^{+\infty} t^{3q}\|\ddot x(t)\|^2dt < +\infty$;

(iii) (fast convergence rates) $\|\ddot x(t)\| = o\left(\frac{1}{t^{2q}}\right)$, $\|\dot x(t)\| = o\left(\frac{1}{t^q}\right)$, $\|A_{\lambda(t)}(x(t))\| = o\left(\frac{1}{t^{2q}}\right)$ *and* $\left\|\frac{d}{dt}A_{\lambda(t)}(x(t))\right\| = o\left(\frac{1}{t^{3q}}\right)$ *as* $t \to +\infty$.

Proof. Energy functional. For $0 < b < \alpha$, consider the energy functional

$$\mathcal{E}(t) := \frac{1}{2}\|b(x(t)-x_t) + t^q(\dot x(t) + \beta A_{\lambda(t)}(x(t)))\|^2 + \frac{b(\alpha - qt^{q-1} - b)}{2}\|x(t)-x_t\|^2 + \frac{t^{2q}\epsilon(t)}{2}\|x(t)\|^2. \tag{58}$$

Note that there exists $t_0' \geq t_0$ such that $\alpha - qt^{q-1} - b > 0$ for all $t \geq t_0'$, consequently, $\mathcal{E}(t)$ is nonnegative for all $t \geq t_0'$. The aim of the first part of the proof is to derive the inequality (78), with (72) as an intermediate result.

Using the classical derivation chain rule and (DIN-AVD-TIKH), we get for the derivative of the energy function for all $t \geq t_0$

$$\begin{aligned}\dot{\mathcal{E}}(t) =& \left\langle (b-\alpha+qt^{q-1})\dot x(t) + (\beta qt^{q-1} - t^q)A_{\lambda(t)}(x(t)), b(x(t)-x_t) + t^q(\dot x(t) + \beta A_{\lambda(t)}(x(t)))\right\rangle \\ &+ \left\langle -t^q\epsilon(t)x(t) - b\frac{d}{dt}x_t, b(x(t)-x_t) + t^q(\dot x(t) + \beta A_{\lambda(t)}(x(t)))\right\rangle + \frac{bq(1-q)t^{q-2}}{2}\|x(t)-x_t\|^2 \\ &+ b(\alpha - qt^{q-1} - b)\left\langle \dot x(t) - \frac{d}{dt}x_t, x(t)-x_t\right\rangle + \left(qt^{2q-1}\epsilon(t) + \frac{t^{2q}\dot\epsilon(t)}{2}\right)\|x(t)\|^2 \\ &+ t^{2q}\epsilon(t)\langle \dot x(t), x(t)\rangle.\end{aligned} \tag{59}$$

After expansion, we obtain for all $t \geq t_0$

$$\begin{aligned}\dot{\mathcal{E}}(t) =& (b+qt^{q-1}-\alpha)t^q\|\dot x(t)\|^2 + \left(\beta(b+qt^{q-1}-\alpha) + (\beta qt^{q-1} - t^q)\right)t^q\langle \dot x(t), A_{\lambda(t)}(x(t))\rangle \\ &+ b(\beta qt^{q-1} - t^q)\left\langle A_{\lambda(t)}(x(t)), x(t)-x_t\right\rangle + \beta(\beta qt^{q-1} - t^q)t^q\|A_{\lambda(t)}(x(t))\|^2 \\ &- bt^q\epsilon(t)\langle x(t), x(t)-x_t\rangle - \beta t^{2q}\epsilon(t)\langle A_{\lambda(t)}(x(t)), x(t)\rangle \\ &+ b(qt^{q-1}-\alpha)\left\langle \frac{d}{dt}x_t, x(t)-x_t\right\rangle - bt^q\left\langle \frac{d}{dt}x_t, \dot x(t)\right\rangle - b\beta t^q\left\langle \frac{d}{dt}x_t, A_{\lambda(t)}(x(t))\right\rangle \\ &+ \frac{bq(1-q)t^{q-2}}{2}\|x(t)-x_t\|^2 + \left(qt^{2q-1}\epsilon(t) + \frac{t^{2q}\dot\epsilon(t)}{2}\right)\|x(t)\|^2.\end{aligned} \tag{60}$$

On the other hand, since

$$\frac{1}{2}\|b(x(t)-x_t)+t^q(\dot{x}(t)+\beta A_{\lambda(t)}(x(t)))\|^2 \leq b^2\|x(t)-x_t\|^2+2t^{2q}\|\dot{x}(t)\|^2+2\beta^2t^{2q}\|A_{\lambda(t)}(x(t))\|^2,$$

we have for all $t \geq t_0$

$$\mathcal{E}(t) \leq 2t^{2q}\|\dot{x}(t)\|^2+2\beta^2t^{2q}\|A_{\lambda(t)}(x(t))\|^2+\frac{b(\alpha-qt^{q-1}+b)}{2}\|x(t)-x_t\|^2+\frac{t^{2q}\epsilon(t)}{2}\|x(t)\|^2.$$

Now, we consider a constant $K \geq 0$ that will be defined later. Then, for all $t \geq t_0$

$$\begin{aligned}
\dot{\mathcal{E}}(t)+\frac{K}{t}\mathcal{E}(t) \leq\ & (b+qt^{q-1}-\alpha+2Kt^{q-1})t^q\|\dot{x}(t)\|^2 \\
& + \left(\beta(b+qt^{q-1}-\alpha)+(\beta qt^{q-1}-t^q)\right)t^q\langle \dot{x}(t), A_{\lambda(t)}(x(t))\rangle \\
& + b(\beta qt^{q-1}-t^q)\left\langle A_{\lambda(t)}(x(t)), x(t)-x_t\right\rangle+\beta(\beta qt^{q-1}-t^q+2\beta Kt^{q-1})t^q\|A_{\lambda(t)}(x(t))\|^2 \\
& - bt^q\epsilon(t)\langle x(t), x(t)-x_t\rangle-\beta t^{2q}\epsilon(t)\langle A_{\lambda(t)}(x(t)), x(t)\rangle \\
& + b(qt^{q-1}-\alpha)\left\langle \frac{d}{dt}x_t, x(t)-x_t\right\rangle-bt^q\left\langle \frac{d}{dt}x_t, \dot{x}(t)\right\rangle-b\beta t^q\left\langle \frac{d}{dt}x_t, A_{\lambda(t)}(x(t))\right\rangle \\
& + \frac{b(q(1-q)t^{q-2}+K(\alpha-qt^{q-1}+b))t^{-1}}{2}\|x(t)-x_t\|^2 \\
& + \left(qt^{2q-1}\epsilon(t)+\frac{t^{2q}\dot{\epsilon}(t)}{2}+\frac{Kt^{2q-1}\epsilon(t)}{2}\right)\|x(t)\|^2. \qquad (61)
\end{aligned}$$

Since $A_{\lambda(t)}(x_t)+\epsilon(t)x_t=0$ and $A_{\lambda(t)}$ is cocoercive, we have for all $t \geq t_0$

$$\begin{aligned}
\left\langle A_{\lambda(t)}(x(t)), x(t)-x_t\right\rangle &= \left\langle A_{\lambda(t)}(x(t))-A_{\lambda(t)}(x_t)-\epsilon(t)x_t, x(t)-x_t\right\rangle \\
&\geq \lambda(t)\|A_{\lambda(t)}(x(t))-A_{\lambda(t)}(x_t)\|^2-\epsilon(t)\langle x_t, x(t)-x_t\rangle \\
&= \lambda(t)\|A_{\lambda(t)}(x(t))+\epsilon(t)x_t\|^2+\frac{\epsilon(t)}{2}(\|x_t\|^2+\|x(t)-x_t\|^2-\|x(t)\|^2) \\
&= \lambda(t)(\|A_{\lambda(t)}(x(t))\|^2+2\epsilon(t)\langle A_{\lambda(t)}(x(t)), x_t\rangle+\epsilon^2(t)\|x_t\|^2) \\
&+ \frac{\epsilon(t)}{2}(\|x_t\|^2+\|x(t)-x_t\|^2-\|x(t)\|^2). \qquad (62)
\end{aligned}$$

Now, obviously there exists $t_1 \geq t_0'$ such that $b(\beta qt^{q-1}-t^q) \leq 0$ for all $t \geq t_1$, hence (62) leads to

$$\begin{aligned}
b(\beta qt^{q-1}-t^q)\left\langle A_{\lambda(t)}(x(t)), x(t)-x_t\right\rangle \leq\ & b(\beta qt^{q-1}-t^q)\lambda(t)(\|A_{\lambda(t)}(x(t))\|^2 \\
& + 2\epsilon(t)\langle A_{\lambda(t)}(x(t)), x_t\rangle+\epsilon^2(t)\|x_t\|^2) \\
& + b(\beta qt^{q-1}-t^q)\frac{\epsilon(t)}{2}(\|x_t\|^2+\|x(t)-x_t\|^2-\|x(t)\|^2) \qquad (63)
\end{aligned}$$

for all $t \geq t_1$. Further,

$$-bt^q\epsilon(t)\langle x(t), x(t)-x_t\rangle = \frac{bt^q\epsilon(t)}{2}(\|x_t\|^2-\|x(t)-x_t\|^2-\|x(t)\|^2) \text{ for all } t \geq t_1. \qquad (64)$$

Plugging (63) and (64) into (61) yields for all $t \geq t_1$

$$\begin{aligned}\dot{\mathcal{E}}(t) + \frac{K}{t}\mathcal{E}(t) \leq\ & (b + qt^{q-1} - \alpha + 2Kt^{q-1})t^q\|\dot{x}(t)\|^2 \\ & + \left(\beta(b + qt^{q-1} - \alpha) + (\beta qt^{q-1} - t^q)\right) t^q\langle \dot{x}(t), A_{\lambda(t)}(x(t))\rangle \\ & + (\beta(\beta qt^{q-1} - t^q + 2\beta Kt^{q-1})t^q + b(\beta qt^{q-1} - t^q)\lambda(t))\|A_{\lambda(t)}(x(t))\|^2 \\ & - \beta t^{2q}\epsilon(t)\langle A_{\lambda(t)}(x(t)), x(t)\rangle + 2b(\beta qt^{q-1} - t^q)\lambda(t)\epsilon(t)\langle A_{\lambda(t)}(x(t)), x_t\rangle \\ & + b(qt^{q-1} - \alpha)\left\langle \frac{d}{dt}x_t, x(t) - x_t\right\rangle - bt^q\left\langle \frac{d}{dt}x_t, \dot{x}(t)\right\rangle - b\beta t^q\left\langle \frac{d}{dt}x_t, A_{\lambda(t)}(x(t))\right\rangle \\ & + \frac{b(q(1-q)t^{q-2} + K(\alpha - qt^{q-1} + b))t^{-1} + b(\beta qt^{q-1} - 2t^q)\epsilon(t)}{2}\|x(t) - x_t\|^2 \\ & + \left(qt^{2q-1}\epsilon(t) + \frac{t^{2q}\dot{\epsilon}(t)}{2} + \frac{(Kt^{2q-1} - b\beta qt^{q-1})\epsilon(t)}{2}\right)\|x(t)\|^2 + \frac{b\beta qt^{q-1}\epsilon(t)}{2}\|x_t\|^2. \qquad (65)\end{aligned}$$

Further, for every $l > 0$ (which will be specified later) and all $t \geq t_1$, one has

$$\begin{aligned}&\left(\beta(b + qt^{q-1} - \alpha) + (\beta qt^{q-1} - t^q)\right) t^q\langle \dot{x}(t), A_{\lambda(t)}(x(t))\rangle \leq \qquad (66)\\ &\frac{|\left(\beta(b + qt^{q-1} - \alpha) + (\beta qt^{q-1} - t^q)\right) t^q|}{2}\left(\frac{\|\dot{x}(t)\|^2}{lt^q} + lt^q\|A_{\lambda(t)}(x(t))\|^2\right).\end{aligned}$$

Similarly, for all $t \geq t_1$ and every positive function $m(t) > 0$ (which will be specified later), one has

$$-\beta t^{2q}\epsilon(t)\langle A_{\lambda(t)}(x(t)), x(t)\rangle \leq \frac{\beta t^{2q}\epsilon(t)}{2}\left(\frac{\|x(t)\|^2}{m(t)} + m(t)\|A_{\lambda(t)}(x(t))\|^2\right), \qquad (67)$$

while for positive function $n(t) > 0$ (which will be specified later), one has

$$2b(\beta qt^{q-1} - t^q)\lambda(t)\epsilon(t)\langle A_{\lambda(t)}(x(t)), x_t\rangle \leq |b(\beta qt^{q-1} - t^q)\lambda(t)\epsilon(t)|\left(\frac{\|x_t\|^2}{n(t)} + n(t)\|A_{\lambda(t)}(x(t))\|^2\right). \qquad (68)$$

Finally, according to Lemma 6, one has $\left\|\frac{d}{dt}x_t\right\|^2 \leq \left(\left(2\frac{\dot{\epsilon}(t)}{\epsilon(t)}\right)^2 + \frac{8q^2}{t^2}\right)\|x_t\|^2$ for almost every $t \geq t_0$. Hence, for almost all $t \geq t_0$, every positive function $s_1(t) > 0$ and every positive constants $s_2, s_3 > 0$ (which will be specified later), one has

$$b(qt^{q-1} - \alpha)\left\langle \frac{d}{dt}x_t, x(t) - x_t\right\rangle \leq \frac{|b(qt^{q-1} - \alpha)|}{2}\left(\frac{1}{s_1(t)}\left(2\left(\frac{\dot{\epsilon}(t)}{\epsilon(t)}\right)^2 + \frac{8q^2}{t^2}\right)\|x_t\|^2 + s_1(t)\|x(t) - x_t\|^2\right), \qquad (69)$$

$$-bt^q\left\langle \frac{d}{dt}x_t, \dot{x}(t)\right\rangle \leq \frac{bt^q}{2}\left(\frac{1}{s_2}\left(2\left(\frac{\dot{\epsilon}(t)}{\epsilon(t)}\right)^2 + \frac{8q^2}{t^2}\right)\|x_t\|^2 + s_2\|\dot{x}(t)\|^2\right), \qquad (70)$$

and

$$-b\beta t^q\left\langle \frac{d}{dt}x_t, A_{\lambda(t)}(x(t))\right\rangle \leq \frac{b\beta t^q}{2}\left(\frac{1}{s_3t^{2q}}\left(2\left(\frac{\dot{\epsilon}(t)}{\epsilon(t)}\right)^2 + \frac{8q^2}{t^2}\right)\|x_t\|^2 + s_3t^{2q}\|A_{\lambda(t)}(x(t))\|^2\right). \qquad (71)$$

Plugging (66)-(71) into (65), we obtain that for almost every $t \geq t_1$ it holds

$$\dot{\mathcal{E}}(t) + \frac{K}{t}\mathcal{E}(t) \leq \mu(t)\|\dot{x}(t)\|^2 + \nu(t)\|A_{\lambda(t)}(x(t))\|^2 + \sigma(t)\|x(t) - x_t\|^2 + \theta(t)\|x(t)\|^2 + \rho(t)\|x_t\|^2, \qquad (72)$$

where

$$\mu(t) = (b + qt^{q-1} - \alpha + 2Kt^{q-1})t^q + \frac{|\left(\beta(b + qt^{q-1} - \alpha) + (\beta qt^{q-1} - t^q)\right) t^q|}{2lt^q} + \frac{s_2 bt^q}{2}; \tag{73}$$

$$\nu(t) = (\beta(\beta qt^{q-1} - t^q + 2\beta Kt^{q-1})t^q + b(\beta qt^{q-1} - t^q)\lambda(t)) + \frac{|\left(\beta(b + qt^{q-1} - \alpha) + (\beta qt^{q-1} - t^q)\right) t^q|lt^q}{2}$$
$$+ \frac{\beta t^{2q}\epsilon(t)m(t)}{2} + |b(\beta qt^{q-1} - t^q)\lambda(t)\epsilon(t)|n(t) + \frac{s_3 b\beta t^{3q}}{2}; \tag{74}$$

$$\sigma(t) = \frac{b(q(1-q)t^{q-2} + K(\alpha - qt^{q-1} + b))t^{-1} + b(\beta qt^{q-1} - 2t^q)\epsilon(t)}{2} + \frac{|b(qt^{q-1} - \alpha)|s_1(t)}{2}; \tag{75}$$

$$\theta(t) = \left(qt^{2q-1}\epsilon(t) + \frac{t^{2q}\dot\epsilon(t)}{2} + \frac{(Kt^{2q-1} - b\beta qt^{q-1})\epsilon(t)}{2}\right) + \frac{\beta t^{2q}\epsilon(t)}{2m(t)}; \tag{76}$$

$$\rho(t) = \frac{b\beta qt^{q-1}\epsilon(t)}{2} + \frac{|b(\beta qt^{q-1} - t^q)\lambda(t)\epsilon(t)|}{n(t)} + \left(2\left(\frac{\dot\epsilon(t)}{\epsilon(t)}\right)^2 + \frac{8q^2}{t^2}\right)\left(\frac{|b(qt^{q-1} - \alpha)|}{2s_1(t)} + \frac{bt^q}{2s_2} + \frac{b\beta}{2s_3 t^q}\right). \tag{77}$$

In the following we will choose the parameters and functions left "unspecified" in order to make the coefficient functions $\mu(t), \nu(t), \sigma(t)$, and $\theta(t)$ become nonpositive for t large enough.

- Since $q < 1$, $\mu(t) = \left(b - \alpha + \frac{1}{2l} + \frac{s_2 b}{2}\right) t^q + \mathcal{O}(t^{2q-1}) + \mathcal{O}(1)$ as $t \to +\infty$. We set

$$l := b\lambda + \frac{1}{4(\alpha - b)} \text{ and } s_2 := \frac{\alpha - b - \frac{1}{2l}}{b}.$$

Since $\lambda > \frac{1}{\alpha^2}$, for all $b \in \left(\frac{\alpha}{2} - \frac{\sqrt{\alpha^2\lambda^2 - \lambda}}{2\lambda}, \frac{\alpha}{2} + \frac{\sqrt{\alpha^2\lambda^2 - \lambda}}{2\lambda}\right)$ it holds $b - \alpha + \frac{1}{2l} + \frac{s_2 b}{2} < 0$. Hence, for every b chosen in this interval, there exists $t_2 \geq t_1$ such that $\mu(t) < 0$ for all $t \geq t_2$.

- We set

$$n(t) := \frac{n}{\epsilon(t)}, \text{ with } n := \frac{1 - \frac{l}{2b\lambda}}{4} > 0.$$

If $\beta > 0$ we set $s_3 := \frac{b\lambda - \frac{l}{2}}{b\beta} > 0$ and $m(t) =: mt$, where $m > 0$ will be specified later. Under this circumstances, $\nu(t) = \left(-\frac{b\lambda}{4} + \frac{l}{8}\right) t^{3q} + \frac{\beta mt^{2q+1}\epsilon(t)}{2} + \mathcal{O}(t^{2q})$ as $t \to +\infty$. Now, according to condition (C_2), there exists $c_2 > 0$ such that $t \leq \frac{c_2 t^q}{\beta\epsilon(t)}$ for t large enough, hence $\frac{\beta mt^{2q+1}\epsilon(t)}{2} \leq \frac{mc_2 t^{3q}}{2}$ for t large enough. Setting $m := \frac{2}{c_2}\left(\frac{b\lambda}{8} - \frac{l}{16}\right)$, it holds $\nu(t) \leq \left(-\frac{b\lambda}{8} + \frac{l}{16}\right) t^{3q} + \mathcal{O}(t^{2q})$ as $t \to +\infty$, therefore there exists $t_3 \geq t_2$ such that $\nu(t) < 0$ for all $t \geq t_3$.

If $\beta = 0$ one has $\nu(t) = \left(-\frac{3b\lambda}{4} + \frac{3l}{8}\right) t^{3q} + \mathcal{O}(t^{2q})$ as $t \to +\infty$, therefore there exists $t_3 \geq t_2$ such that $\nu(t) < 0$ for all $t \geq t_3$.

- Let $q < r < 1$ as provided by condition (C_1) and let

$$s_1(t) := s_1 t^{-r},$$

where $s_1 > 0$ will be specified later. Then, for t large enough $\beta qt^{q-1} - t^q \leq 0$, hence $\sigma(t) \leq \frac{b}{2}(s_1\alpha t^{-r} - t^q\epsilon(t)) + \mathcal{O}(t^{-1})$ as $t \to +\infty$. According to (C_1), there exists $c_1 > 0$ such that $t^q\epsilon(t) \geq \frac{c_1}{t^r}$ for t large enough. Choosing s_1 such that $s_1\alpha < c_1$, it holds $s_1\alpha t^{-r} - t^q\epsilon(t) \leq (s_1\alpha - c_1)t^{-r}$ for t large enough, hence there exists $t_4 \geq t_3$ such that $\sigma(t) < 0$ for all $t \geq t_4$.

- If $\beta > 0$, by taking into account that $m(t) = mt$ with $m = \frac{2}{c_2}\left(\frac{b\lambda}{8} - \frac{l}{16}\right)$, one has for all $t \geq t_1$

$$\theta(t) \leq \frac{t^{2q}\dot\epsilon(t)}{2} + \frac{(K + \frac{\beta}{m} + 2q)t^{2q-1}\epsilon(t)}{2}.$$

Since by condition (C_3), $\frac{\dot{\epsilon}(t)}{\epsilon(t)} \leq -\frac{c_3}{t}$ for t large enough and $c_3 > 2q + \frac{16c_2\beta}{\alpha\lambda-\sqrt{\lambda}}$, one can choose $b := \alpha - \frac{1}{2\sqrt{\lambda}} \in \left(\frac{\alpha}{2} - \frac{\sqrt{\alpha^2\lambda^2-\lambda}}{2\lambda}, \frac{\alpha}{2} + \frac{\sqrt{\alpha^2\lambda^2-\lambda}}{2\lambda}\right)$ and in this case it holds $\frac{\beta}{m} = \frac{8c_2\beta}{\alpha\lambda-\sqrt{\lambda}}$. Hence, for all $K \in \left(0, \frac{8c_2\beta}{\alpha\lambda-\sqrt{\lambda}}\right]$ one has $\frac{\dot{\epsilon}(t)}{\epsilon(t)} \leq -(K + \frac{\beta}{m} + 2q)t^{-1}$ for t large enough, that is, there exists $t_5 \geq t_4$ such that $\theta(t) \leq 0$ for all $t \geq t_5$.

If $\beta = 0$, then $\theta(t) = \frac{t^{2q}\dot{\epsilon}(t)}{2} + \frac{(K+2q)t^{2q-1}\epsilon(t)}{2}$. Proceeding as before, we choose $K \in (0, c_3 - 2q]$, which allows us to conclude that there exists $t_5 \geq t_4$ such that $\theta(t) \leq 0$ for all $t \geq t_5$.

In what follows we fix also $K > 0$ such that $K \neq 1 - r$, $K \in \left(0, \frac{8c_2\beta}{\alpha\lambda-\sqrt{\lambda}}\right]$ if $\beta > 0$, and $K \in (0, c_3 - 2q]$ if $\beta = 0$.

• Finally, taking into account that $\left(\frac{\dot{\epsilon}(t)}{\epsilon(t)}\right)^2 \leq \frac{c_4}{t^2}$ for t large enough, we deduce that there exist the constants $K_1, K_2, K_3 > 0$ and there exists $t_6 \geq t_5$ such that for all $t \geq t_6$

$$\rho(t) = \frac{b\beta q t^{q-1}\epsilon(t)}{2} + \frac{|b(\beta q t^{q-1} - t^q)\lambda(t)\epsilon^2(t)|}{n} + \left(2\left(\frac{\dot{\epsilon}(t)}{\epsilon(t)}\right)^2 + \frac{8q^2}{t^2}\right)\left(\frac{|b(qt^{q-1} - \alpha)|t^r}{2s_1} + \frac{bt^q}{2s_2} + \frac{b\beta}{2s_3t^q}\right)$$
$$\leq K_1 t^{q-1}\epsilon(t) + K_2 t^{3q}\epsilon^2(t) + K_3 t^{r-2}.$$

Taking into account the above considerations, and using the fact that $\|x_t\| \leq \|x^*\|$, we obtain from (72) that for almost all $t \geq t_6$ it holds

$$\dot{\mathcal{E}}(t) + \frac{K}{t}\mathcal{E}(t) \leq (K_1 t^{q-1}\epsilon(t) + K_2 t^{3q}\epsilon^2(t) + K_3 t^{r-2})\|x^*\|^2. \tag{78}$$

By multiplying (78) with t^K we obtain that for almost all $t \geq t_6$ it holds

$$\frac{d}{dt}(t^K\mathcal{E}(t)) \leq (K_1 t^{K+q-1}\epsilon(t) + K_2 t^{K+3q}\epsilon^2(t) + K_3 t^{K+r-2})\|x^*\|^2. \tag{79}$$

Now we fix $T > t_6$ and integrate (79) on the interval $[t_6, T]$. This yields

$$T^K\mathcal{E}(T) \leq K_1\|x^*\|^2 \int_{t_6}^T t^{K+q-1}\epsilon(t)dt + K_2\|x^*\|^2 \int_{t_6}^T t^{K+3q}\epsilon^2(t)dt + \frac{K_3\|x^*\|^2}{K+r-1}T^{K+r-1} + C_0$$

for some constant $C_0 > 0$. In other words, for all $T \geq t_6$ it holds

$$\mathcal{E}(T) \leq K_1\|x^*\|^2 \frac{\int_{t_6}^T t^{K+q-1}\epsilon(t)dt}{T^K} + K_2\|x^*\|^2 \frac{\int_{t_6}^T t^{K+3q}\epsilon^2(t)dt}{T^K} + \frac{K_3\|x^*\|^2}{K+r-1}T^{r-1} + \frac{C_0}{T^K}. \tag{80}$$

It remains to show that the right-hand side of (80) goes to zero as $T \to +\infty$. Indeed, using the fact that $\lambda(t)\epsilon(t) \to 0$ as $t \to +\infty$ we have that $\epsilon(t) < \frac{1}{t^{2q}}$ for t large enough and this shows that

$$K_1\|x^*\|^2 \frac{\int_{t_6}^T t^{K+q-1}\epsilon(t)dt}{T^K} \to 0 \text{ as } T \to +\infty.$$

Obviously, since $r < 1$, we have

$$\frac{K_3\|x^*\|^2}{K+r-1}T^{r-1} + \frac{C_0}{T^K} \to 0 \text{ as } T \to +\infty.$$

In order to show that $\frac{\int_{t_6}^T t^{K+3q}\epsilon^2(t)dt}{T^K} \to 0$ as $T \to +\infty$ we will use Lemma A.3 in the Appendix. Indeed, by assumption, the function $t \mapsto f(t) = t^{3q}\epsilon^2(t) \in L^1((t_6, +\infty), \mathbb{R})$, and the function $t \mapsto \varphi(t) = t^K$ is positive and nondecreasing on $[t_6, +\infty)$ and $\lim_{t\to+\infty}\varphi(t) = +\infty$. Hence,

$$\lim_{T\to+\infty} \frac{\int_{t_6}^T t^{K+3q}\epsilon^2(t)dt}{T^K} = \lim_{T\to+\infty} \frac{1}{\varphi(T)} \int_{t_6}^T \varphi(t)f(t)dt = 0.$$

Convergence rates and strong convergence of $x(t)$ as $t \to +\infty$. The fact that $\mathcal{E}(t) \to 0$ as $t \to +\infty$ has as consequence that $\lim_{t\to+\infty} \|x(t) - x_t\|^2 = 0$. According to Proposition 5, $x_t \to x^*$ as $t \to +\infty$, therefore

$$x(t) \text{ converges strongly, as } t \to +\infty, \text{ to } x^*. \tag{81}$$

According to (C_0) we get $t^{2q}\epsilon(t)x(t) \to 0$ as $t \to +\infty$. Now, using that $A_{\lambda(t)}$ is $\frac{1}{\lambda(t)}-$ Lipschitz continuous, it yields

$$\lambda t^{2q}\|A_{\lambda(t)}(x(t))\| = \lambda(t)\|A_{\lambda(t)}(x(t)) - A_{\lambda(t)}(x^*)\| \leq \|x(t) - x^*\| \text{ for all } t \geq t_0,$$

which, in combination with (81), yields to

$$\lim_{t\to+\infty} t^{2q}\|A_{\lambda(t)}(x(t))\| = 0. \tag{82}$$

From the fact that $\mathcal{E}(t) \to 0$ as $t \to +\infty$ and (81) we also obtain that $\|t^q(\dot{x}(t) + \beta A_{\lambda(t)}(x(t)))\| \to 0$ as $t \to +\infty$, which, in combination with (82), further yields

$$\lim_{t\to+\infty} t^q\|\dot{x}(t)\| = 0. \tag{83}$$

Lemma A.2 (c2) in the Appendix guarantees that $\left\|\frac{d}{dt}A_{\lambda(t)}(x(t))\right\| \leq \frac{2}{\lambda(t)}\|\dot{x}(t)\| + 2\frac{|\lambda'(t)|}{\lambda(t)}\|A_{\lambda(t)}(x(t))\|$ for all $t \geq t_0$, and from here we deduce that

$$\lim_{t\to+\infty} t^{3q}\left\|\frac{d}{dt}A_{\lambda(t)}(x(t))\right\| = 0. \tag{84}$$

Finally, since $t^{2q}\epsilon(t)x(t) \to 0$ as $t \to +\infty$ by using the definition of (DIN-AVD-TIKH), we obtain that

$$\lim_{t\to+\infty} t^{2q}\|\ddot{x}(t)\| = 0. \tag{85}$$

Integral estimates. In order to derive the integral estimates we will set $K := 0$. Then (72) becomes for all $t \geq t_1$

$$\dot{\mathcal{E}}(t) \leq \mu(t)\|\dot{x}(t)\|^2 + \nu(t)\|A_{\lambda(t)}(x(t))\|^2 + \sigma(t)\|x(t) - x_t\|^2 + \theta(t)\|x(t)\|^2 + \rho(t)\|x_t\|^2. \tag{86}$$

As seen above, for the coefficient functions $\mu(t)$ and $\nu(t)$ we have sharper estimates than that they are negative for t large enough. In particular we have that

$$\mu(t) \leq -C_1 t^q \text{ for some } C_1 > 0 \text{ and } t \text{ large enough}$$

and

$$\nu(t) \leq -C_2 t^{3q} \text{ for some } C_2 > 0 \text{ and } t \text{ large enough.}$$

In analogy to (78), we can conclude that there exist $t_7 \geq t_6$ and $K_1, K_2, K_3 > 0$ such that

$$\dot{\mathcal{E}}(t) + C_1 t^q\|\dot{x}(t)\|^2 + C_2 t^{3q}\|A_{\lambda(t)}(x(t))\|^2 \leq (K_1 t^{q-1}\epsilon(t) + K_2 t^{3q}\epsilon^2(t) + K_3 t^{r-2})\|x^*\|^2 \text{ for all } t \geq t_7. \tag{87}$$

Next we fix $T > t_7$ and integrate (87) on the interval $[t_7, T]$. This yields

$$\mathcal{E}(T) + C_1\int_{t_7}^T t^q\|\dot{x}(t)\|^2 dt + C_2\int_{t_7}^T t^{3q}\|A_{\lambda(t)}(x(t))\|^2 dt \leq K_1\|x^*\|^2\int_{t_7}^T t^{q-1}\epsilon(t)dt + K_2\|x^*\|^2\int_{t_7}^T t^{3q}\epsilon^2(t)dt + \frac{K_3\|x^*\|^2}{r-1}T^{r-1} + C_3, \tag{88}$$

for some constant $C_3 > 0$. According to the hypotheses and the fact that $\epsilon(t) \leq \frac{1}{t^{2q}}$ for t large enough, we obtain that the right-hand side of (88) is bounded, therefore

$$\int_{t_0}^{+\infty} t^q \|\dot{x}(t)\|^2 dt < +\infty \text{ and } \int_{t_0}^{+\infty} t^{3q} \|A_{\lambda(t)}(x(t))\|^2 dt < +\infty.$$

In order to obtain the integral estimate for the acceleration $\ddot{x}(t)$ we can proceed as in Theorem 2, namely, by making use of (37) and (38). From here we obtain

$$\int_{t_0}^{+\infty} t^{3q} \|\ddot{x}(t)\|^2 dt < +\infty.$$

■

Remark 8. As seen in Proposition 5, condition (C_0) is essential to guarantee that x_t converges to the minimal norm element x^* which was one of the main ingredients for proving strong convergence.

Natural candidates for Tikhonov parametrization functions that satisfy the conditions $(C_0) - (C_4)$ are, as in the previous section, $\epsilon(t) = \frac{a}{t^p}$, with $a, p > 0$. For this choice, (C_0) is nothing else that $p > 2q$, while (C_1) is equivalent to $p < 1 + q$. Further, (C_2) holds if and only if $\beta = 0$ or $\beta > 0$ and $p \geq 1 - q$. If $\beta > 0$ and $p > 1 - q$, then c_2 can be taken arbitrarily small, whereas if $p = 1 - q$, then $c_2 \geq \beta a$. Condition (C_3) asks for the existence of an element c_3 in the interval $\left(2q + \frac{16c_2\beta}{\alpha\lambda - \sqrt{\lambda}}, p\right]$, which can be guaranteed only for $p > 1 - q$, since in this case condition (C_2) is fulfilled for c_2 arbitrarily small. Finally, (C_4) is always satisfied for this choice of ϵ.

Consequently, if $\max\left((\operatorname{sgn}\beta)(1-q), 2q\right) < p < 1 + q$, then $\epsilon(t) = \frac{a}{t^p}$, for $a > 0$, satisfies $(C_0) - (C_4)$. This is in concordance with the conditions imposed on the Tikhonov parametrization function in [4],[11] and [15].

Theorem 7 and Remark 8 lead to the following result.

Corollary 9. *Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator such that $S := \{x \in \mathcal{H} : 0 \in Ax\}$ is nonempty. Consider the evolution equation (DIN-AVD-TIKH) with $\beta \geq 0$, $\lambda(t) = \lambda t^{2q}$, for $\lambda > \frac{1}{\alpha^2}$ and $0 < q < 1$, $\epsilon(t) = \frac{a}{t^p}$, for $a > 0$ and $\max\left((\operatorname{sgn}\beta)(1-q), \frac{3q+1}{2}\right) < p \leq q + 1$. Then, for any trajectory $x : [t_0, +\infty) \to \mathcal{H}$ of (DIN-AVD-TIKH), the following statements are true:*

(i) If $p < q + 1$, then the statements (i)-(iii) of Theorem 7 are valid. In addition, $\|x(t) - x_t\| = \mathcal{O}(t^{\frac{3q-2p+1}{2}})$ as $t \to +\infty$.

(ii) If $p = q + 1$, then the trajectory $x(t)$ is bounded and $\|\ddot{x}(t)\| = \mathcal{O}\left(\frac{1}{t^{2q}}\right), \|\dot{x}(t)\| = \mathcal{O}\left(\frac{1}{t^q}\right)$, $\|A_{\lambda(t)}(x(t))\| = \mathcal{O}\left(\frac{1}{t^{2q}}\right)$ and $\left\|\frac{d}{dt}A_{\lambda(t)}(x(t))\right\| = \mathcal{O}\left(\frac{1}{t^{3q}}\right)$ as $t \to +\infty$.

Proof. (i) The condition $\int_{t_0}^{+\infty} t^{3q}\epsilon^2(t)dt < +\infty$ is equivalent to $p > \frac{3q+1}{2}$ and, since $\frac{3q+1}{2} > 2q$, as seen in Remark 8, the conditions $(C_0) - (C_4)$ are fulfilled. Hence, all the assumptions in the hypotheses of Theorem 7 are satisfied.

In order to derive the rate of convergence for $\|x(t) - x_t\|$, we consider again (80), which reads in this case,

$$\mathcal{E}(T) \leq K_1 a \|x^*\|^2 \frac{\int_{t_6}^T t^{K+q-p-1} dt}{T^K} + K_2 a^2 \|x^*\|^2 \frac{\int_{t_6}^T t^{K+3q-2p} dt}{T^K} + \frac{K_3\|x^*\|^2}{K + r - 1} T^{r-1} + \frac{C_0}{T^K} \text{ for all } T \geq t_6. \quad (89)$$

In other words, there exist constants $\widetilde{K}_1, \widetilde{K}_2, \widetilde{K}_3, \widetilde{K}_4 > 0$ such that

$$\mathcal{E}(T) \leq \widetilde{K}_1 T^{q-p} + \widetilde{K}_2 T^{3q-2p+1} + \widetilde{K}_3 T^{r-1} + \widetilde{K}_4 T^{-K} \text{ for all } T \geq t_6. \quad (90)$$

Obviously, $3q-2p+1 > q-p$. Further, in condition (C_1), one can take r such that $q < r < 3q-2p+2 < 1$, hence $3q-2p+1 > r-1$.

Finally, we will show that we can choose K such that $3q-2p+1 > -K$. Indeed, according to the proof of Theorem 7, for $\beta > 0$ we must choose $K \in \left(0, \frac{8c_2\beta}{\alpha\lambda-\sqrt{\lambda}}\right]$ such that $K \neq 1-r$. Since $p > 1-q$, we can choose $c_2 > 0$ such that $-3q+2p-1 < \frac{8c_2\beta}{\alpha\lambda-\sqrt{\lambda}} < \frac{16c_2\beta}{\alpha\lambda-\sqrt{\lambda}} < p-2q$. Therefore, as seen in Remark 8, (C_3) remains valid and we can choose a positive $K \neq 1-r$ such that $-3q+2p-1 < K < \frac{8c_2\beta}{\alpha\lambda-\sqrt{\lambda}}$.

If $\beta = 0$, then we must choose $K \in (0, c_3 - 2q]$. Since, as seen in Remark 8, $c_3 = p$ fulfils (C_3), for $K = p-2q$ we have $3q-2p+1 > -K$ and $K \neq 1-r$.

Consequently, the right-hand side of (90) is a term of order $\mathcal{O}(T^{3q-2p+1})$ as $T \to +\infty$, which, by taking into account the definition of $\mathcal{E}(t)$, yields

$$\|x(t) - x_t\| = \mathcal{O}(t^{\frac{3q-2p+1}{2}}) \text{ as } t \to +\infty.$$

(ii) Now we consider the case $p = q+1$. We start from (72), according to which, for all $t \geq t_1$

$$\dot{\mathcal{E}}(t) + \frac{K}{t}\mathcal{E}(t) \leq \mu(t)\|\dot{x}(t)\|^2 + \nu(t)\|A_{\lambda(t)}(x(t))\|^2 + \sigma(t)\|x(t)-x_t\|^2 + \theta(t)\|x(t)\|^2 + \rho(t)\|x_t\|^2. \quad (91)$$

We choose l and s_2 as in Theorem 7 and obtain that $\mu(t) < 0$ for t large enough. Further, we set $m(t) := \frac{1}{b}t^{q+1}$ and obtain $\frac{\beta t^{2q}\epsilon(t)m(t)}{2} = \mathcal{O}(t^{2q})$ as $t \to +\infty$, hence, for s_3 chosen as in the proof of Theorem 7 we obtain that $\nu(t) < 0$ for t large enough.

We choose $s_1(t) := s_1 t^{-1}$, where $s_1 > 0$ will be defined later. Then, $\sigma(t) \leq \frac{1}{2}(s_1\alpha b - 2ab + K(\alpha + b))t^{-1} + \mathcal{O}(t^{q-2})$ as $t \to +\infty$. For $0 < K < \frac{2ab}{\alpha+b}$ and $s_1 < \frac{2ab-K(\alpha+b)}{\alpha b}$ it holds that $\sigma(t) < 0$ for t large enough.

Further, by taking into account that $m(t) = \frac{1}{b}t^{q+1}$, for $\beta \geq 0$ one has

$$\theta(t) \leq \frac{t^{2q}\dot{\epsilon}(t)}{2} + \frac{(K+2q)t^{2q-1}\epsilon(t)}{2} = \frac{-a(q+1)+a(K+2q)}{2}t^{q-2}.$$

Hence, if $K \leq 1-q$, then $\theta(t) \leq 0$ for all $t \geq t_1$.

From here we deduce, as in (78), that there exists $t_6 \geq t_1$ and $K_1 > 0$ such that for almost every $t \geq t_6$

$$\dot{\mathcal{E}}(t) + \frac{K}{t}\mathcal{E}(t) \leq K_1 t^{-1}\|x^*\|^2. \quad (92)$$

By multiplying (92) with t^K it yields $\frac{d}{dt}(t^K\mathcal{E}(t)) \leq K_1 t^{K-1}\|x^*\|^2$ for almost every $t \geq t_6$. We choose $T > t_6$, integrate this inequality on the interval $[t_6, T]$ and obtain that $\mathcal{E}(T) \leq C$ for some $C > 0$. This means that $\mathcal{E}(t) \leq C$ for all $t \geq t_6$.

By taking into account the definition of the energy functional we obtain that $\|x(t) - x_t\|$ is bounded, and further, since $\|x_t\| \leq \|x^*\|$, that $x(t)$ is bounded. Consequently, $\|t^q(\dot{x}(t) + \beta A_{\lambda(t)}(x(t)))\|$ is bounded.

Now, using the $\frac{1}{\lambda(t)}$-Lipschitz continuity of $A_{\lambda(t)}$ we obtain that $\|A_{\lambda(t)}(x(t))\| = \mathcal{O}\left(\frac{1}{t^{2q}}\right)$ as $t \to +\infty$ and from here that $\|\dot{x}(t)\| = \mathcal{O}\left(\frac{1}{t^q}\right)$ as $t \to +\infty$. The fact that $\left\|\frac{d}{dt}A_{\lambda(t)}(x(t))\right\| = \mathcal{O}\left(\frac{1}{t^{3q}}\right)$ as $t \to +\infty$ can be obtained by using the previous estimates and Lemma A.2 (c_2) in the Appendix. Finally, by using (DIN-AVD-TIKH), it yields $\|\ddot{x}(t)\| = \mathcal{O}\left(\frac{1}{t^{2q}}\right)$ as $t \to +\infty$. ■

4 Nonsmooth convex optimization

Let $f : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ be a proper, convex and lower semicontinuous function such that $\operatorname{argmin} f \neq \emptyset$ and consider the minimization problem

$$\inf_{x \in \mathcal{H}} f(x).$$

The problem of finding a global minimum of f reduces to the solving of the monotone inclusion $0 \in \partial f(x)$. The Moreau envelope of f of modulus $\lambda > 0$ is the C^1 convex function

$$f_\lambda : \mathcal{H} \to \mathbb{R}, \quad f_\lambda(x) = \inf_{y \in \mathcal{H}} \left\{ f(y) + \frac{1}{2\lambda}\|x - y\|^2 \right\}.$$

One has that $\operatorname{argmin}_{\mathcal{H}} f = \operatorname{argmin}_{\mathcal{H}} f_\lambda$ and $\min_{\mathcal{H}} f_\lambda = \min_{\mathcal{H}} f$. The Yosida approximation of ∂f is equal to the gradient of the Moreau envelope of f, that is, for all $\lambda > 0$ it holds $(\partial f)_\lambda = \nabla f_\lambda$. Furthermore, $J_{\lambda \partial f} = \operatorname{prox}_{\lambda f}$, where

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

denotes the proximal operator of f of modulus λ.

This being said, the evolution equation (DIN-AVD-TIKH) reads in this contexts for $t \geq t_0$

$$\text{(DIN-AVD-TIKH-CONV)} \quad \ddot{x}(t) + \frac{\alpha}{t^q}\dot{x}(t) + \beta \frac{d}{dt}\left(\nabla f_{\lambda(t)}(x(t))\right) + \nabla f_{\lambda(t)}(x(t)) + \epsilon(t)x(t) = 0,$$

where $\alpha > 0$, $\beta \geq 0$, $0 < q \leq 1$ $\lambda : [t_0, +\infty) \longrightarrow (0, +\infty)$ is the Moreau envelope parametrization function and $\epsilon : [t_0, +\infty) \longrightarrow (0, +\infty)$ is the Tikhonov parametrization function.

As a direct consequence of Theorem 2 we have the following result.

Theorem 10. *Let $f : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ be a proper, convex and lower semicontinuous function such that $\operatorname{argmin} f \neq \emptyset$. Consider the evolution equation* (DIN-AVD-TIKH-CONV) *with $\beta \geq 0$, $\lambda(t) = \lambda t^{2q}$, for $\lambda > 0$ and $0 < q \leq 1$, and the system parameters satisfying the following conditions:*

$$\textit{if } q < 1, \textit{ then } \alpha > 1, \ \lambda > \frac{1}{\alpha^2} \textit{ and } \epsilon(t) \geq \frac{q(1-q)}{t^2} \textit{ for } t \textit{ large enough;}$$

$$\textit{if } q = 1, \textit{ then either } \alpha > 3 \textit{ and } \lambda > \frac{1}{8(\alpha - 3)} \textit{ or } \alpha > 1, \lambda > \frac{1}{(\alpha - 1)^2} \textit{ and } \frac{\dot{\epsilon}(t)}{\epsilon(t)} \leq -\frac{2}{t} \textit{ for } t \textit{ large enough.}$$

In addition, assume that $\int_{t_0}^{+\infty} t^q \epsilon(t) dt < +\infty$.
Then, for any trajectory $x : [t_0, +\infty) \to \mathcal{H}$ of (DIN-AVD-TIKH-CONV)*, the following statements are true:*

(i) (convergence of the trajectory) $x(t)$ is bounded. Furthermore, if $\beta = 0$ or $\beta > 0$ and $q > \frac{1}{2}$, then $x(t)$ converges weakly, as $t \to +\infty$, to global minimizer of f;

(ii) (integral estimates) $\displaystyle\int_{t_0}^{+\infty} t^q \|\dot{x}(t)\|^2 dt < +\infty, \quad \int_{t_0}^{+\infty} t^{3q} \|\nabla f_{\lambda(t)}(x(t))\|^2 dt < +\infty,$

and, if $\displaystyle\int_{t_0}^{+\infty} t^{3q}\epsilon^2(t) dt < +\infty$*, then* $\displaystyle\int_{t_0}^{+\infty} t^{3q}\|\ddot{x}(t)\|^2 dt < +\infty$*;*

(iii) (fast convergence rates) $\|\dot{x}(t)\| = \mathcal{O}\left(\frac{1}{t^q}\right), \|\nabla f_{\lambda(t)}(x(t))\| = o\left(\frac{1}{t^{2q}}\right), \left\|\frac{d}{dt}\nabla f_{\lambda(t)}(x(t))\right\| = \mathcal{O}\left(\frac{1}{t^{3q}}\right)$ *as $t \to +\infty$. In addition, if $\int_{t_0}^{+\infty} t^{3q}\epsilon^2(t) dt < +\infty$, then $\|\dot{x}(t)\| = o\left(\frac{1}{t^q}\right)$ as $t \to +\infty$;*

(iv) (fast convergence rates for the values) $f_{\lambda(t)}(x(t)) - \min_{\mathcal{H}} f = o\left(\frac{1}{t^{2q}}\right)$ *and* $f(\operatorname{prox}_{\lambda(t)f}(x(t))) - \min_{\mathcal{H}} f = o\left(\frac{1}{t^{2q}}\right)$ *as $t \to +\infty$.*

In addition, $\|\operatorname{prox}_{\lambda(t)f}(x(t)) - x(t)\| \to 0$ as $t \to +\infty$.

Proof. The statements (i)-(iii) follow directly from Theorem 2 applied to the operator ∂f and by using that $(\partial f)_\lambda = \nabla f_\lambda$.

Let us prove (iv). Take $x^* \in \operatorname{argmin} f$. From the gradient inequality we have

$$f_{\lambda(t)}(x(t)) - \min_{\mathcal{H}} f = f_{\lambda(t)}(x(t)) - f_{\lambda(t)}(x^*) \leq \langle \nabla f_{\lambda(t)}(x(t)), x(t) - x^* \rangle$$
$$\leq \|\nabla f_{\lambda(t)}(x(t))\| \|x(t) - x^*\| \leq M \|\nabla f_{\lambda(t)}(x(t))\| = o\left(\frac{1}{t^{2q}}\right) \text{ as } t \to +\infty,$$

where $M := \sup_{t \geq t_0} \|x(t) - x^*\|$. By the definition of $f_{\lambda(t)}$ and of the proximal operator, we have for all $t \geq t_0$

$$f_{\lambda(t)}(x(t)) - \min_{\mathcal{H}} f = f(\operatorname{prox}_{\lambda(t)f}(x(t))) - \min_{\mathcal{H}} f + \frac{1}{2\lambda(t)} \|x(t) - \operatorname{prox}_{\lambda(t)f}(x(t))\|^2. \tag{93}$$

hence,

$$f(\operatorname{prox}_{\lambda(t)f}(x(t))) - \min_{\mathcal{H}} f = o\left(\frac{1}{t^{2q}}\right) \text{ as } t \to +\infty \text{ and } \lim_{t \to +\infty} \|x(t) - \operatorname{prox}_{\lambda(t)f}(x(t))\|^2 = 0. \tag{94}$$

■

Similarly, as a direct consequence of Theorem 7 we have the following result.

Theorem 11. *Let $f : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ be a proper, convex and lower semicontinuous function such that* argmin $f \neq \emptyset$. *Consider the evolution equation* (DIN-AVD-TIKH-CONV) *with $\beta \geq 0$ and $\lambda(t) = \lambda t^{2q}$, for $\lambda > \frac{1}{\alpha^2}$ and $0 < q < 1$. Assume that $\int_{t_0}^{+\infty} t^{3q} \epsilon^2(t) dt < +\infty$ and the Tikhonov parametrization function also satisfies the conditions the conditions $(C_0) - (C_4)$.*
Then, for any trajectory $x : [t_0, +\infty[\to \mathcal{H}$ of (DIN-AVD-TIKH-CONV)*, the following statements are true:*

(i) (convergence of the trajectory) $x(t)$ converges strongly, as $t \to +\infty$, to the element of minimum norm of $\operatorname{argmin}_{\mathcal{H}} f$*;*

(ii) (integral estimates) $\int_{t_0}^{+\infty} t^q \|\dot{x}(t)\|^2 dt < +\infty$, $\int_{t_0}^{+\infty} t^{3q} \|\nabla f_{\lambda(t)}(x(t))\|^2 dt < +\infty$ and $\int_{t_0}^{+\infty} t^{3q} \|\ddot{x}(t)\|^2 dt < +\infty$;

(iii) (fast convergence rates) $\|\ddot{x}(t)\| = o\left(\frac{1}{t^{2q}}\right)$, $\|\dot{x}(t)\| = o\left(\frac{1}{t^q}\right)$, $\|\nabla f_{\lambda(t)}(x(t))\| = o\left(\frac{1}{t^{2q}}\right)$ and $\left\|\frac{d}{dt} \nabla f_{\lambda(t)}(x(t))\right\| = o\left(\frac{1}{t^{3q}}\right)$ as $t \to +\infty$;

(iv) (fast convergence of the values) $f_{\lambda(t)}(x(t)) - \min_{\mathcal{H}} f = o\left(\frac{1}{t^{2q}}\right)$ and $f(\operatorname{prox}_{\lambda(t)f}(x(t))) - \min_{\mathcal{H}} f = o\left(\frac{1}{t^{2q}}\right)$ as $t \to +\infty$.

In addition, $\|\operatorname{prox}_{\lambda(t)f}(x(t)) - x(t)\| \to 0$ as $t \to +\infty$.

A Appendix

We state the existence and uniqueness result for (DIN-AVD-TIKH) which can be proved by reformulating the evolution equation as a first oder dynamical system in the product space $\mathcal{H} \times \mathcal{H}$ and by making use of standard arguments relying on the Cauchy-Lipschitz theorem, see also [5, 11].

Theorem A.1. *Let $\beta \geq 0$ and $\lambda, \epsilon : [t_0, +\infty) \to (0, +\infty)$ be continuous functions such that $\lim_{t \to +\infty} \lambda(t) > 0$. Then, for all $(x_0, x_1) \in \mathcal{H} \times \mathcal{H}$, there exists a unique C^2 solution $x : [t_0, +\infty) \to \mathcal{H}$ of the dynamical system* (DIN-AVD-TIKH) *which satisfies the Cauchy data $x(t_0) = x_0$, $\dot{x}(t_0) = x_1$.*

In the remaining of the appendix, we present some lemmas which play a crucial role in the analysis of (DIN-AVD-TIKH). The following important technical result is [5, Lemma 1].

Lemma A.2. *Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator, $\gamma, \nu > 0$ and $x, y \in \mathcal{H}$. Then, the following inequalities are satisfied:*

a) $\|\gamma A_\gamma(x) - \nu A_\nu(y)\| \leq 2\|x - y\| + |\gamma - \nu| \|A_\gamma(x)\|$;

b) $\|A_\gamma(x) - A_\nu(y)\| \leq \frac{2}{\gamma}\|x - y\| + \frac{|\gamma-\nu|}{\gamma}(\|A_\gamma(x)\| + \|A_\nu(y)\|)$;

c) If $x : [t_0, +\infty) \to \mathcal{H}$ is a differentiable map and $\lambda : [t_0, +\infty) \to (0, +\infty)$ is a derivable function, then, for all $t \in [t_0, +\infty[$ and all $z \in A^{-1}(0)$,

$$(c_1) \quad \left\| \frac{d}{dt}\lambda(t)A_{\lambda(t)}(x(t)) \right\| \leq 2\|\dot{x}(t)\| + |\lambda'(t)| \|A_{\lambda(t)}(x(t))\|;$$

$$(c_2) \quad \left\| \frac{d}{dt}A_{\lambda(t)}(x(t)) \right\| \leq \frac{2}{\lambda(t)}\|\dot{x}(t)\| + 2\frac{|\lambda'(t)|}{\lambda(t)}\|A_{\lambda(t)}(x(t))\|;$$

$$(c_3) \quad \left\| \frac{d}{dt}\lambda(t)A_{\lambda(t)}(x(t)) \right\| \leq 2\|\dot{x}(t)\| + \frac{|\lambda'(t)|}{\lambda(t)}\|x(t) - z\|.$$

The following lemma was stated, for instance, in [4, Lemma A.3].

Lemma A.3. *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 continuous version of the Opial Lemma stated below is the counterpart of a convergence result for quasi-Fejér monotone sequences. Its proof can be found in [1, Lemma 5.1].

Lemma A.4. *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 continuous version of the Opial Lemma is the main tool for proving weak convergence for the trajectories of an evolution equation.

Lemma A.5. *Let $S \subseteq \mathcal{H}$ be a nonempty set and $x : [t_0, +\infty) \to \mathcal{H}$ 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, as $t \to +\infty$, to an element in S .

Acknowledgements. The authors are thankful to an anonymous reviewer for comments and remarks which enhanced the quality of the paper.

References

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

[2] H. Attouch, A. Balhag, Z. Chbani, H. Riahi, *Damped inertial dynamics with vanishing Tikhonov regularization: Strong asymptotic convergence towards the minimum norm solution*, Journal of Differential Equations 311, 29–58 (2022)

[3] H. Attouch, A. Balhag, Z. Chbani, H. Riahi, *Accelerated gradient methods combining Tikhonov regularization with geometric damping driven by the Hessian*, Applied Mathematics & Optimization 88, Article number: 29 (2023)

[4] H. Attouch, Z. Chbani, H. Riahi, *Combining fast inertial dynamics for convex optimization with Tikhonov regularization*, Journal of Mathematical Analysis and Applications 457(2), 1065–1094 (2018)

[5] H. Attouch, S.C. László, *Continuous Newton-like inertial dynamics for monotone inclusions*, Set-Valued and Variational Analysis 29, 555-581 (2021)

[6] H. Attouch, S.C. László, *Convex optimization via inertial algorithms with vanishing Tikhonov regularization: fast convergence to the minimum norm solution*, arXiv:2104.11987

[7] H. Attouch, J. Peypouquet, *Convergence of inertial dynamics and proximal algorithms governed by maximal monotone operators*, Mathematical Programming 174(1-2), 391–432 (2019)

[8] H. Attouch, J. Peypouquet, P. Redont, *Fast convex optimization via inertial dynamics with Hessian driven damping*, Journal of Differential Equations 261(10), 5734–5783 (2016)

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

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

[11] R.I. Boţ, E.R. Csetnek, S.C. László, *Tikhonov regularization of a second order dynamical system with Hessian damping*, Mathematical Programming 189, 151–186 (2021)

[12] R.I. Boţ, G. Dong, P. Elbau, O. Scherzer, *Convergence rates of first and higher order dynamics for solving linear ill-posed problems*, Foundations of Computational Mathematics 22(5), 1567–1629 (2022)

[13] R.I. Boţ, D.A. Hulett, *Second order splitting dynamics with vanishing damping for additively structured monotone inclusions*, to appear in Journal of Dynamics and Differential Equations, DOI: 10.1007/s10884-022-10160-3

[14] E.R. Csetnek, M.A. Karapetyants, *A fast continuous time approach for non-smooth convex optimization with time scaling and Tikhonov regularization*, arXiv:2207.12023

[15] S.C. László, *On the strong convergence of the trajectories of a Tikhonov regularized second order dynamical system with asymptotically vanishing damping*, Journal of Differential Equations 362(25), 355–381 (2023)

[16] Y. Nesterov, *A method of solving a convex programming problem with convergence rate* $O\left(1/k^2\right)$, Soviet Mathematics Doklady 27, 372–376 (1983)

[17] W. Su, S. Boyd, E.J. Candès, *A differential equation for modeling Nesterov's accelerated gradient method: theory and insights*, Journal of Machine Learning Research 17(153), 1–43 (2016)