TIKHONOV REGULARIZATION OF A PERTURBED HEAVY BALL SYSTEM WITH VANISHING DAMPING

CRISTIAN DANIEL ALECSA * AND SZILÁRD CSABA LÁSZLÓ †

Abstract. This paper deals with a perturbed heavy ball system with vanishing damping that contains a Tikhonov regularization term, in connection to the minimization problem of a convex Fréchet differentiable function. We show that the value of the objective function in a generated trajectory converges in order $o(1/t^2)$ to the global minimum of the objective function. We also obtain the fast convergence of the velocities towards zero. Moreover, we obtain that a trajectory generated by the dynamical system converges weakly to a minimizer of the objective function. Finally, we show that the presence of the Tikhonov regularization term assures the strong convergence of the generated trajectories to an element of minimal norm from the argmin set of the objective function.

Key words. convex optimization; heavy ball method; continuous second order dynamical system; Tikhonov regularization; convergence rate; strong convergence.

AMS subject classifications. 34G20, 47J25, 90C25, 90C30, 65K10.

1. Introduction.

Let $\mathcal{H}$ be a real Hilbert space endowed with the scalar product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$ and let $g : \mathcal{H} \longrightarrow \mathbb{R}$ be a convex Fréchet differentiable function. Consider the minimization problem

$$(P) \quad \inf_{x \in \mathcal{H}} g(x)$$

in connection to the second order dynamical system

$$\begin{cases} \ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \nabla g\left(x(t) + \left(\gamma + \frac{\beta}{t}\right)\dot{x}(t)\right) + \epsilon(t)x(t) = 0, \\ x(t_0) = u_0,\ \dot{x}(t_0) = v_0, \end{cases} \tag{1.1}$$

where $t_0 > 0$, $(u_0, v_0) \in \mathcal{H} \times \mathcal{H}$, $\alpha \geq 3$, $\gamma > 0$, $\beta \in \mathbb{R}$ or $\gamma = 0$, $\beta \geq 0$ and $\epsilon : [t_0, +\infty) \longrightarrow \mathbb{R}_+$ is a non-increasing function of class C^1, such that $\lim_{t \to +\infty} \epsilon(t) = 0$. The starting time t_0 is taken as strictly greater than zero whenever the coefficients $\frac{\alpha}{t}$ and $\frac{\beta}{t}$ have singularities at 0. This is not a limitation of the generality of the proposed approach, since we will focus on the asymptotic behaviour of the generated trajectories.

First of all, note that the dynamical system (1.1) is the Tikhonov regularized version of the perturbed heavy ball system with vanishing damping considered in connection to the optimization problem (P) by Alecsa-László-Pinţa in [3]. The dynamical system considered in [3] can be seen as an intermediate system between the heavy ball system with vanishing damping [33] and the heavy ball system with Hessian driven damping [19] and possesses all the valuable properties of the latter ones. Indeed, according to [3], in case $\gamma > 0$, $\beta \in \mathbb{R}$, or $\gamma = 0$, $\beta \geq 0$, the objective function value in a trajectory generated by the perturbed heavy ball system converges in order $\mathcal{O}(1/t^2)$ to the global minimum of the objective function and the trajectory converges weakly to a minimizer of the objective function. Further, according to [3, Remark 2], in case $\gamma = 0$ and $\beta < 0$ the perturbed heavy ball system can generate periodical solutions, therefore in this case the convergence of a generated trajectory to a minimizer of the objective is hopeless.

Throughout the paper we assume that ∇g is Lipschitz continuous on bounded sets and $\operatorname{argmin} g \neq \emptyset$. Further, the Tikhonov regularization parameter, $\epsilon(t)$, satisfies one of the following assumptions, (see also Remark 1).

(C1) There exist $K > 1$ and $t_1 \geq t_0$ such that

$$\dot{\epsilon}(t) \leq -\frac{K}{2}\left|\gamma + \frac{\beta}{t}\right|\epsilon^2(t) \text{ for every } t \geq t_1.$$

(C2) There exists $K > 0$ and $t_1 \geq t_0$ such that

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

*Technical University of Cluj-Napoca, Department of Mathematics, Str. Memorandumului nr. 28, 400114 Cluj-Napoca, Romania and Romanian Institute of Science and Technology, Cluj-Napoca, Romania, e-mail: alecsa@rist.ro.

†Technical University of Cluj-Napoca, Department of Mathematics, Str. Memorandumului nr. 28, 400114 Cluj-Napoca, Romania, e-mail: szilard.laszlo@math.utcluj.ro

Natural candidates for the Tikhonov regularization parameter that satisfy the above conditions are $\epsilon(t) = \frac{a}{t^r}$, $a > 0$, $r \geq 1$. In order to give a better perspective of the results obtained in this paper, for this special case of the Tikhonov regularization parameter, we can conclude the following.

1. According to Theorem 2.1 and Theorem 2.3, if the Tikhonov regularization parameter is $\epsilon(t) = \frac{a}{t^r}$, $a > 0$, $r > 2$ the following statements hold. When $\alpha \geq 3$, one has $g\left(x(t) + \left(\gamma + \frac{\beta}{t}\right)\dot{x}(t)\right) - \min g = \mathcal{O}\left(t^{-2}\right)$ as $t \to +\infty$. Further, one has the integral estimates $\int_{t_0}^{+\infty} t^2 \left\|\nabla g\left(x(t) + \left(\gamma + \frac{\beta}{t}\right)\dot{x}(t)\right)\right\|^2 dt < +\infty$, whenever $\gamma > 0$, $\beta \in \mathbb{R}$ and $\int_{t_0}^{+\infty} t \left\|\nabla g\left(x(t) + \frac{\beta}{t}\dot{x}(t)\right)\right\|^2 dt < +\infty$, whenever $\gamma = 0$, $\beta > 0$. When $\alpha > 3$ one has $g\left(x(t) + \left(\gamma + \frac{\beta}{t}\right)\dot{x}(t)\right) - \min g = o(t^{-2})$, $\|\dot{x}(t)\| = o(t^{-2})$ as $t \to +\infty$ and $x(t)$ converges weakly to a minimizer of g. Further, $\int_{t_0}^{+\infty} t\|\dot{x}(t)\|^2 dt < +\infty$ and $\int_{t_0}^{+\infty} t\left(g\left(x(t) + \left(\gamma + \frac{\beta}{t}\right)\dot{x}(t)\right) - \min g\right) dt < +\infty$.

2. According to Theorem 3.2, in the case $\epsilon(t) = \frac{a}{t^r}$, $a > 0$, $1 < r < 2$, for $\alpha > 3$ and $\alpha = 3$, $\gamma = 0$, $\beta \geq 0$ one has $\liminf_{t\to\infty} \|x(t) - x^*\| = 0$, where x^* is the element of minimum norm of $\operatorname{argmin} g$. In addition, $x(t)$ converges strongly to x^* when either the trajectory $\{x(t) : t \geq T\}$ remains in the ball $B(0, \|x^*\|)$, or in its complement, for T large enough.

3. In the case $\epsilon(t) = \frac{a}{t^2}$, $a > 0$ and $\alpha > 3$, according to Theorem 2.2 one has that x is bounded, $g\left(x(t) + \left(\gamma + \frac{\beta}{t}\right)\dot{x}(t)\right) - \min g = \mathcal{O}\left(\frac{1}{t^2}\right)$ and $\|\dot{x}(t)\| = \mathcal{O}\left(\frac{1}{t}\right)$ as $t \to +\infty$. Further, if $a > \frac{2}{9}\alpha(\alpha - 3)$, according to Theorem 3.2 the conclusions stated at **2.** hold.

Observe that according to **3.** we are able to obtain both fast convergence of the function values and strong convergence of the trajectories for the same Tikhonov regularization parameter. For a long time this was an unsolved problem in the literature. However, recently Attouch and László [16] studied a dynamical system in connection to the optimization problem (P) and succeeded to obtain rapid convergence towards the infimal value of g, and the strong convergence of the trajectories towards the element of minimum norm of the set of minimizers of g. In our context, one can observe that the case $r = 2$ is critical, in the sense that separates the two cases: the case when we obtain fast convergence of the function values and weak convergence of the trajectories to a minimizer and the case when the strong convergence of the trajectories to a minimizer of minimum norm is assured. These results are in concordance with the results obtained in [10] and [20], since, as will be shown in what follows, the dynamical system studied in this paper can be thought as an intermediate system between the dynamical system studied in [10] and the dynamical system considered in [20]. Before we give a more enlightening discussion about the connection of the dynamical system (1.1) and the Tikhonov regularized dynamical systems studied in [10] and [20] we underline two new features of our analysis. Firstly, we can show fast convergence of the velocity to zero, a property which is also obtained for the Tikhonov regularized system studied in [10], but this property is not shown for the Tikhonov regularized system considered in [20]. Secondly, we obtain some integral estimates for the gradient of the objective function and these results also appear for the Tikhonov regularized system studied in [20], but these estimates are not shown for the Tikhonov regularized system studied in [10].

For further insight into the Tikhonov regularization techniques we refer to [10, 12, 13, 16, 20, 22, 25].

1.1. Connection with second order dynamical systems with asymptotically vanishing damping. The dynamical system (1.1) is strongly related to the second order dynamical systems with an asymptotically vanishing damping term, studied by Su-Boyd-Candès in [33] in connection to the optimization problem (P), that is,

$$(AVD)_\alpha \quad \ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \nabla g\left(x(t)\right) = 0,\ x(t_0) = u_0,\ \dot{x}(t_0) = v_0,\ u_0, v_0 \in \mathcal{H}.$$

It is obvious that the latter system can be obtained from (1.1) by taking $\gamma = \beta = 0$ and $\epsilon \equiv 0$. According to [33], the trajectories generated by $(\text{AVD})_\alpha$ assure fast minimization property of order $\mathcal{O}\left(1/t^2\right)$ for the decay $g(x(t)) - \min g$, provided $\alpha \geq 3$. For $\alpha > 3$, it has been shown by Attouch-Chbani-Peypouquet-Redont [9] that each trajectory generated by $(\text{AVD})_\alpha$ converges weakly to a minimizer of the objective function g. Further, it is shown in [17] and [30] that the asymptotic convergence rate of the values is actually $o(1/t^2)$.

An appropriate discretization of $(\text{AVD})_\alpha$ with $\alpha = 3$ corresponds to Nesterov's historical algorithm [31].

Therefore, as it was emphasized in [33], for $\alpha = 3$ the system $(\text{AVD})_\alpha$ can be seen as a continuous version of the accelerated gradient method of Nesterov.

However, the case $\alpha = 3$ is critical, i.e., the convergence of the trajectories generated by $(\text{AVD})_\alpha$ remains an open problem. The subcritical case $\alpha \leq 3$ has been examined by Apidopoulos-Aujol-Dossal [5] and Attouch-Chbani-Riahi [11], with the convergence rate of the objective values $\mathcal{O}\left(t^{-\frac{2\alpha}{3}}\right)$.

When the objective function g is not convex, the convergence of the trajectories generated by $(\text{AVD})_\alpha$ is a largely open question. Recent progress has been made in [21], where the convergence of the trajectories of a system, which can be considered as a perturbation of $(\text{AVD})_\alpha$ has been obtained in a non-convex setting.

The corresponding inertial algorithms obtained from $(\text{AVD})_\alpha$ via discretization, are in line with the Nesterov accelerated gradient method and enjoy similar properties to the continuous case, see [24] and also [4, 7, 9, 27, 28] for further results and the extensions to proximal-gradient algorithms for structured optimization. For other results concerning the system $(\text{AVD})_\alpha$ and its extensions in we refer to [11, 23, 26].

A version of $(\text{AVD})_\alpha$ containing a Tikhonov regularization term, strongly related to (1.1), was considered by Attouch-Chbani-Riahi in [10]. According to [10] the presence of Tikhonov regularization term $\epsilon(t)x(t)$ provides the strong convergence of the trajectories to the element of minimum norm of the set of minimizers of g, when $\epsilon(t)$ tends slowly to zero. We emphasize that (1.1), for the case $\beta = \gamma = 0$, reduces to the system studied in [10].

1.2. Connection with second order dynamical systems with Hessian driven damping. The dynamical system (1.1) is also related to the second order dynamical system with Hessian driven damping term, studied by Attouch-Peypouquet-Redont in [19], that is,

$$(DIN - AVD)_{\alpha,\beta} \quad \ddot{x}(t) + \frac{\alpha}{t}\dot{x}(t) + \beta\nabla^2 g(x(t))\dot{x}(t) + \nabla g\left(x(t)\right) = 0,\ x(t_0) = u_0,\ \dot{x}(t_0) = v_0,\ u_0, v_0 \in \mathcal{H},\ \beta \geq 0.$$

In [19], for the case $\alpha > 3$, $\beta > 0$ the authors showed the weak convergence of a generated trajectory to a minimizer of g and they obtained convergence rate of order $o(1/t^2)$ for the objective function along the trajectory. The temporal discretization of this dynamical system provides first-order algorithms which beside the extrapolation term contain a correction term which is equal to the difference of the gradients at two consecutive steps [8, 2]. Several recent studies have been devoted to this subject, see [6, 14, 15, 29, 32]. Further, Boţ-Csetnek-László considered in [20] the Tikhonov regularization of $(\text{DIN-AVD})_{\alpha,\beta}$. They obtained fast convergence results for the function values along the trajectories and strong convergence results of the trajectory to the minimizer of the objective function of minimum norm. Now, by using the Taylor expansion of $\nabla g(\cdot)$ we get

$$\nabla g\left(x(t) + \left(\gamma + \frac{\beta}{t}\right)\dot{x}(t)\right) \approx \nabla g(x(t)) + \left(\gamma + \frac{\beta}{t}\right)\nabla^2 g(x(t))\dot{x}(t), \tag{1.2}$$

which shows that the system $(\text{DIN-AVD})_{\alpha,\beta}$ with Tikhonov regularization term $\epsilon(t)x(t)$ considered in [20] and the dynamical system (1.1) are strongly related. In this paper, we aim to obtain fast convergence results for the function values along the trajectories generated by the dynamical system (1.1) and strong convergence results of the trajectories to the minimizer of the objective function of minimum norm, under some similar assumption as those considered in [20]. However, we emphasize that our objective function g is of class C^1 meanwhile the objective function considered in [20] is of class C^2. Further, as we mentioned before, we are also able to show the rate $o(1/t)$ for the velocity. Moreover, according to [3] the dynamical system (1.1), (with $\epsilon \equiv 0$), leads via explicit Euler discretization to inertial algorithms. In particular the Nesterov accelerated convex gradient method can be obtained from (1.1) via natural explicit discretization.

The following numerical experiments reveal that a trajectory generated by (1.1) and the objective function value in this trajectory have a better convergence behaviour than a trajectory generated by the system $(\text{AVD})_\alpha$ with a Tikhonov regularization term studied in [10] and also a similar behaviour as the trajectories generated by the dynamical system considered in [20]. At the same time, the perturbation term $\left(\gamma + \frac{\beta}{t}\right)\dot{x}(t)$ in the argument of the gradient of the objective function g has a smoothing effect, just as the case of $(\text{DIN-AVD})_{\alpha,\beta}$. This also confirms our conclusion that (1.1) can be thought as an intermediate system between the system $(\text{AVD})_\alpha$ with a Tikhonov regularization term considered in [10] and $(\text{DIN-AVD})_{\alpha,\beta}$ with a Tikhonov regularization term studied in [20], which inherits the best properties of the latter systems.

1.3. Some numerical experiments. In this section we consider two numerical experiments for the trajectories generated by the dynamical system (1.1) for a convex but not strongly convex objective function

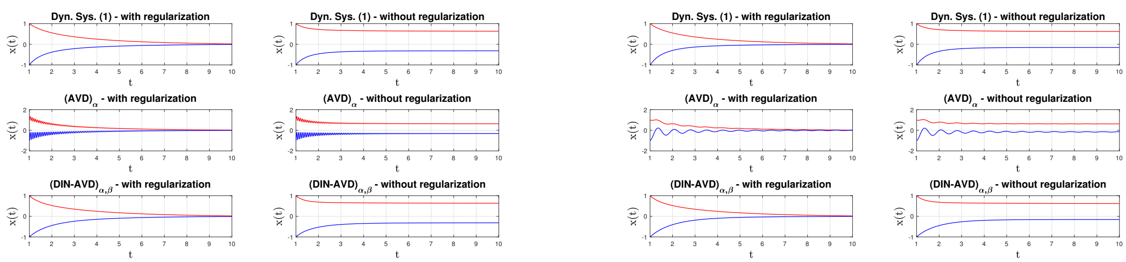
$$g : \mathbb{R}^2 \to \mathbb{R},\ \ g(x, y) = (ax + by)^2 \text{ where } a, b \in \mathbb{R} \setminus \{0\}.$$

Observe that $\operatorname{argmin} g = \left\{\left(x, -\frac{a}{b}x\right) : x \in \mathbb{R}\right\}$ and $\min g = 0$. Obviously, the minimizer of minimal norm of g is $x^* = (0, 0)$.

Everywhere in the following numerical experiments we consider the continuous time dynamical systems (1.1) and the dynamical systems $(\text{AVD})_\alpha$ and $(\text{DIN-AVD})_{\alpha,\beta}$ with or without the regularization term $\epsilon(t)x(t)$, solved numerically with the ode45 adaptive method in MATLAB on the interval $[1, 10]$. For the Tikhonov regularization parameter we take $\epsilon(t) = \frac{1}{t^{1.5}}$ and we consider the starting points $x(1) = (1, -1)$, $\dot{x}(1) = (-1, 1)$.

Further, we take $\alpha = 3.1$, $\beta = -0.5$, $\gamma = 1$ in (1.1), $\alpha = 3.1$ in $(\text{AVD})_\alpha$ and $\alpha = 3.1$, $\beta = 1$ in $(\text{DIN-AVD})_{\alpha,\beta}$. Observe that γ in (1.1) is equal with β in $(\text{DIN-AVD})_{\alpha,\beta}$ hence, according to (1.2) the trajectories of these systems will share a similar behaviour.

The results are depicted at Figure 1, where the first component of a solution x is depicted with red, meanwhile the second component is depicted with blue.



(a) $a = 25$ $b = 50$. (b) $a = 1.5$ $b = 6$.

Fig. 1.1: The behaviour of the dynamical systems (1.1), $(\text{AVD})_\alpha$ and $(\text{DIN-AVD})_{\alpha,\beta}$ with and without Tikhonov regularization term, for convex, but not strongly convex, objective functions.

Analyzing Figure 1 we observe that indeed the trajectories of the dynamical systems (1.1), $(\text{AVD})_\alpha$ and $(\text{DIN-AVD})_{\alpha,\beta}$, in the presence of the Tikhonov regularization term $\epsilon(t)x(t)$, converge to x^* the element of minimal norm from $\operatorname{argmin} g$. However, when we consider these systems without regularization, that is $\epsilon(t) \equiv 0$, we observe that we still have convergence of the generated trajectories to a minimizer of g of the form $(x_0, -0.5x_0)$ Figure 1 (a) and $(x_0, -0.25x_0)$ Figure 1 (b), for some $x_0 \in \mathbb{R}$, however these minimizers are not anymore of minimal norm.

Further, observe that indeed the trajectories of (1.1) and $(\text{DIN-AVD})_{\alpha,\beta}$ have a similar behaviour and both eliminates the oscillates obtained for the trajectories of $(\text{AVD})_\alpha$.

In our second experiment we study the evolution of the two errors $\|x(t) - x^*\|$ and $g(x(t) - \min g$, for a trajectory $x(t)$ generated by the dynamical system (1.1), with respect to different values of β and γ. So we take $a = 0.1$ and $b = 50$, values for which the function g is poorly conditioned. We take $\alpha = 3.1$, $\epsilon(t) = \frac{1}{t^{2.5}}$ and we consider the starting points $x(1) = (1, -1)$, $\dot{x}(1) = (-1, 1)$. Further, since the theoretical expected rate for the decay $g(x(t) - \min g$ is $O\left(\frac{1}{t^2}\right)$ we also consider the graph of the function $\frac{g(x(1)) - \min g}{t^2}$ on the interval $[1, 10]$. The results are depicted on Figure 2.

One can observe, see Figure 2, that the best choice seems to be $\gamma = 0$, $\beta > 0$. This case outperforms both the cases $\gamma > 0$, $\beta < 0$ and $\gamma > 0$, $\beta > 0$. However, if $\gamma > 0$ choosing negative β leads to better convergence properties. Further, all these choices of the parameters β, γ outperform the case $\gamma = \beta = 0$ which is the case of $(\text{AVD})_\alpha$ with Tikhonov regularization.

1.4. Organization of the paper. In the next section we carry out the asymptotic analysis of the trajectories generated by the dynamical system (1.1). We obtain convergence rate of order $o(1/t^2)$ for the

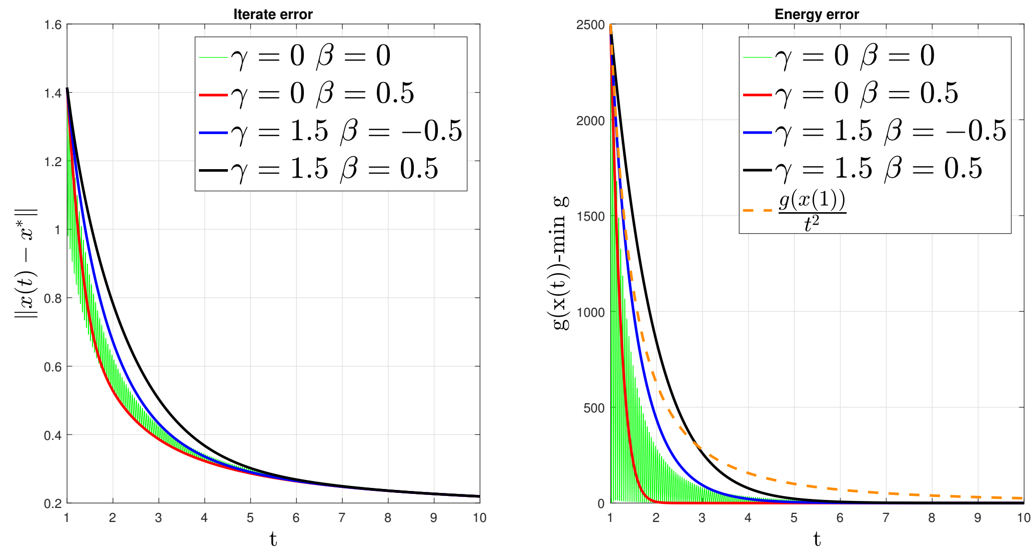


Fig. 1.2: Error analysis with different parameters in dynamical system (1.1) for a convex, but not strongly convex, objective function.

energy error $g(x(t)+(\gamma+\beta/t)\dot{x}(t))-\min g$, convergence rate of order $o(1/t)$ for the velocity $\dot{x}(t)$ and weak convergence of the trajectories to a minimizer of the objective function g. Further, several integral estimates for the gradient of g will be provided. In section 3 we show that indeed the Tikhonov regularization term in the dynamical system (1.1) assures the strong convergence of the trajectories to a minimizer of g of minimal norm. Finally, we conclude our paper by outlining some perspectives.

2. Asymptotic analysis of the regularized dynamical system (1.1)**.** Existence and uniqueness of a $C^2([t_0,+\infty),\mathcal{H})$ global solution of the dynamical system (1.1) can be shown via the classical Cauchy-Lipschitz-Picard theorem, by rewriting (1.1) as a first order system in the product space $\mathcal{H}\times\mathcal{H}$, see Theorem A.1 from the Appendix. In this section we carry out the asymptotic analysis concerning the trajectories generated by the dynamical system (1.1). We will treat the non-critical case $\alpha>3$ and the critical case $\alpha=3$ separately under the assumptions (C1) and (C2).

Remark 1. Let us discuss the connections between the conditions (C1) and (C2). Obviously, when $\gamma=\beta=0$ condition (C1) is satisfied in virtue of the nonincreasing property of ϵ. Further, if we assume that $\epsilon(t_0')=0$ for some $t_0'\geq t_0$ then $\epsilon(t)\equiv 0$ for all $t\geq t_0'$, hence (C1) and (C2) become trivial.

Assume in what follows that $\epsilon(t)>0$ for all $t\geq t_0$. Now, if $\gamma>0$, $\beta\in\mathbb{R}$ or $\gamma=0$, $\beta>0$, then there exists some $\bar{t}_1\geq t_1$ such that $\left|\gamma+\frac{\beta}{t}\right|=\gamma+\frac{\beta}{t}$, for all $t\geq\bar{t}_1$. Hence, condition (C1) leads to the fact that there exist $K>1$ and $\bar{t}_1\geq t_0$ such that

$$\frac{d}{dt}\left(\frac{1}{\epsilon(t)}\right)\geq\frac{K}{2}\gamma+\frac{K\beta}{2t},\text{ for all }t\geq\bar{t}_1.$$

Now, if $\gamma>0$, by integrating the above relation on an interval $[\bar{t}_1,t]$ we obtain that there exists $K_1>0$ such that

$$\epsilon(t)\leq\frac{1}{\frac{K\gamma}{2}t+\frac{K\beta}{2}\ln t+\frac{1}{\epsilon(\bar{t}_1)}-\frac{K\gamma}{2}\bar{t}_1-\frac{K\beta}{2}\ln\bar{t}_1}\leq\frac{K_1}{t},\text{ for all }t\geq\bar{t}_1.$$

Consequently, if $\gamma>0$ then (C1) implies (C2) and one also has $\int_{t_0}^{+\infty}\frac{\epsilon(t)}{t}<+\infty$.

However, if $\gamma=0$ and $\beta>0$ then by similar arguments as above we obtain that $\epsilon(t)\leq\frac{K_1}{\ln t}$ for some $K_1>0$ and t big enough, and this condition is obviously weaker than condition (C2). Observe that this case does not imply that $\int_{t_0}^{+\infty}\frac{\epsilon(t)}{t}<+\infty$.

As we have mentioned at Introduction, natural candidates for the regularization functions $t \to \epsilon(t)$ that satisfy the conditions (C1) and (C2) are $\epsilon(t) = \frac{a}{t^r}$, $r \geq 1$, $a > 0$. However, if $\gamma = 0$ and $\beta > 0$, then (C1) is satisfied even with $\epsilon(t) = \frac{a}{t^r}$, $r > 0$, $a > 0$. In the latter case, (C1) is also satisfied with $\epsilon(t) = \frac{a}{(\ln(t))^p}$, $p \geq 1$, $a > 0$.

2.1. The non critical case $\alpha > 3$. We show that under some appropriate conditions imposed on ϵ, for $\alpha > 3$ we have $o(1/t^2)$ convergence rate for the decay $g\left(x(t) + \left(\gamma + \frac{\beta}{t}\right)\dot{x}(t)\right) - \min g$, fast convergence of the velocity to 0 and weak convergence of $x(t)$. Further, integral estimates are also obtained.

THEOREM 2.1. *Let $t_0 > 0$ and for some starting points $u_0, v_0 \in \mathcal{H}$ let $x : [t_0, \infty) \to \mathcal{H}$ be the unique global solution of (1.1). Assume that $\alpha > 3$ and that one of the conditions* (C1) *or* (C2) *is fulfilled. Then, the following results hold.*

(i) If $\epsilon : [t_0, +\infty) \to [0, +\infty)$ satisfies $\int_{t_0}^{+\infty} \frac{\epsilon(t)}{t} dt < +\infty$ then, $\lim_{t\to+\infty} g\left(x(t) + \left(\gamma + \frac{\beta}{t}\right)\dot{x}(t)\right) = \min g$. Further, $\lim_{t\to+\infty} \|\dot{x}(t)\| = 0$, hence $\lim_{t\to+\infty} g(x(t)) = \min g$.

(ii) Assume that $\epsilon : [t_0, +\infty) \longrightarrow [0, +\infty)$ satisfies $\int_{t_0}^{+\infty} t\epsilon(t)dt < +\infty$. Then, the following statements hold true.

(convergence) $x(t)$ is bounded and $x(t)$ converges weakly, as $t \to +\infty$, to an element of $\operatorname{argmin} g$.

(integral estimates) $\int_{t_0}^{+\infty} t\|\dot{x}(t)\|^2 dt < +\infty$, $\int_{t_0}^{+\infty} t\left(g\left(x(t) + \left(\gamma + \frac{\beta}{t}\right)\dot{x}(t)\right) - \min g\right) dt < +\infty$, further $\int_{t_0}^{+\infty} t^2 \left\|\nabla g\left(x(t) + \left(\gamma + \frac{\beta}{t}\right)\dot{x}(t)\right)\right\|^2 dt < +\infty$, whenever $\gamma > 0$, $\beta \in \mathbb{R}$ and for $\gamma = 0$, $\beta > 0$ one has $\int_{t_0}^{+\infty} t\left\|\nabla g\left(x(t) + \frac{\beta}{t}\dot{x}(t)\right)\right\|^2 dt < +\infty$.

(pointwise estimates) $\lim_{t\to+\infty} \|\dot{x}(t)\| = 0$, $\|\dot{x}(t)\| = o\left(\frac{1}{t}\right)$ and $g\left(x(t) + \left(\gamma + \frac{\beta}{t}\right)\dot{x}(t)\right) - \min g = o\left(\frac{1}{t^2}\right)$ as $t \to +\infty$. Further, $\|x(t)\| = o\left(\frac{1}{t\sqrt{\epsilon(t)}}\right)$ as $t \to +\infty$.

PROOF . **Lyapunov analysis.** First, let $x^* \in \operatorname{argmin} g$, $b \in (2, \alpha - 1)$ and denote $g^* := g(x^*) = \min g$. For simplicity we denote $\beta(t) = \gamma + \frac{\beta}{t}$ and for a positive function $a(t)$, $t \geq t_0$ we introduce the energy functional $\mathcal{E} : [t_0, \infty) \to \mathbb{R}$,

(2.1)
$$\mathcal{E}(t) = a(t)\left(g(x(t) + \beta(t)\dot{x}(t)) - g^*\right) + \frac{t^2\epsilon(t)}{2}\|x(t)\|^2 + \frac{1}{2}\|b(x(t) - x^*) + t\dot{x}(t)\|^2 + \frac{b(\alpha - 1 - b)}{2}\|x(t) - x^*\|^2.$$

Then,

$$\begin{aligned}\dot{\mathcal{E}}(t) = {} & a'(t)\left(g(x(t) + \beta(t)\dot{x}(t)) - g^*\right) + a(t)\langle\nabla g(x(t) + \beta(t)\dot{x}(t)), \beta(t)\ddot{x}(t) + (\beta'(t) + 1)\dot{x}(t)\rangle \\ & + \left(t\epsilon(t) + \frac{t^2\dot{\epsilon}(t)}{2}\right)\|x(t)\|^2 + t^2\epsilon(t)\langle\dot{x}(t), x(t)\rangle + \langle(b+1)\dot{x}(t) + t\ddot{x}(t), b(x(t) - x^*) + t\dot{x}(t)\rangle \\ & + b(\alpha - 1 - b)\langle\dot{x}(t), x(t) - x^*\rangle. \end{aligned} \tag{2.2}$$

From the dynamical system (1.1), we have that $\ddot{x}(t) = -\epsilon(t)x(t) - \frac{\alpha}{t}\dot{x}(t) - \nabla g(x(t) + \beta(t)\dot{x}(t))$. Hence,

$$\begin{aligned} & a(t)\langle\nabla g(x(t) + \beta(t)\dot{x}(t)), \beta(t)\ddot{x}(t) + (\beta'(t) + 1)\dot{x}(t)\rangle = \\ & a(t)\left\langle\nabla g(x(t) + \beta(t)\dot{x}(t)), -\beta(t)\epsilon(t)x(t) + \left(-\beta(t)\frac{\alpha}{t} + \beta'(t) + 1\right)\dot{x}(t)) - \beta(t)\nabla g(x(t) + \beta(t)\dot{x}(t))\right\rangle = \\ & -\beta(t)a(t)\|\nabla g(x(t) + \beta(t)\dot{x}(t))\|^2 + \left(-\beta(t)\frac{\alpha}{t} + \beta'(t) + 1\right)a(t)\langle\nabla g(x(t) + \beta(t)\dot{x}(t)), \dot{x}(t)\rangle \\ & -\beta(t)\epsilon(t)a(t)\langle\nabla g(x(t) + \beta(t)\dot{x}(t)), x(t)\rangle. \end{aligned} \tag{2.3}$$

Further,

$$\begin{aligned} & \langle(b+1)\dot{x}(t) + t\ddot{x}(t), b(x(t) - x^*) + t\dot{x}(t)\rangle = \\ & \langle(b + 1 - \alpha)\dot{x}(t) - t\epsilon(t)x(t) - t\nabla g(x(t) + \beta(t)\dot{x}(t)), b(x(t) - x^*) + t\dot{x}(t)\rangle = \\ & b(b + 1 - \alpha)\langle\dot{x}(t), x(t) - x^*\rangle + (b + 1 - \alpha)t\|\dot{x}(t)\|^2 - bt\epsilon(t)\langle x(t), x(t) - x^*\rangle - t^2\epsilon(t)\langle\dot{x}(t), x(t)\rangle \\ & - bt\langle\nabla g(x(t) + \beta(t)\dot{x}(t)), x(t) - x^*\rangle - t^2\langle\nabla g(x(t) + \beta(t)\dot{x}(t)), \dot{x}(t)\rangle. \end{aligned} \tag{2.4}$$

Combining (2.2), (2.3) and (2.4) we get

$$\begin{aligned}(2.5)\quad \dot{\mathcal{E}}(t) &= a'(t)\left(g(x(t)+\beta(t)\dot{x}(t))-g^*\right)-\beta(t)a(t)\left\|\nabla g(x(t)+\beta(t)\dot{x}(t))\right\|^2+(b+1-\alpha)t\|\dot{x}(t)\|^2\\ &+\left(t\epsilon(t)+\frac{t^2\dot{\epsilon}(t)}{2}\right)\|x(t)\|^2+\left(\left(-\beta(t)\frac{\alpha}{t}+\beta'(t)+1\right)a(t)-t^2\right)\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),\dot{x}(t)\rangle\\ &-\beta(t)\epsilon(t)a(t)\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),x(t)\rangle-bt\langle\nabla g(x(t)+\beta(t)\dot{x}(t))+\epsilon(t)x(t),x(t)-x^*\rangle.\end{aligned}$$

Consider now the strongly convex function $g_t:\mathcal{H}\longrightarrow\mathbb{R}$, $g_t(x)=g(x)+\frac{\epsilon(t)}{2}\|x\|^2$. From the gradient inequality we have $g_t(y)-g_t(x)\geq\langle\nabla g_t(x),y-x\rangle+\frac{\epsilon(t)}{2}\|x-y\|^2$, for all $x,y\in\mathcal{H}$. Take now $y=x^*$ and $x=x(t)+\beta(t)\dot{x}(t)$. We get

$$\begin{aligned}&g(x^*)+\frac{\epsilon(t)}{2}\|x^*\|^2-g(x(t)+\beta(t)\dot{x}(t))-\frac{\epsilon(t)}{2}\|x(t)+\beta(t)\dot{x}(t)\|^2\geq\\ &-\langle\nabla g(x(t)+\beta(t)\dot{x}(t))+\epsilon(t)(x(t)+\beta(t)\dot{x}(t)),x(t)+\beta(t)\dot{x}(t)-x^*\rangle+\frac{\epsilon(t)}{2}\|x(t)+\beta(t)\dot{x}(t)-x^*\|^2.\end{aligned}$$

Consequently,

$$\begin{aligned}&-\langle\nabla g(x(t)+\beta(t)\dot{x}(t))+\epsilon(t)x(t),x(t)-x^*\rangle-\beta(t)\langle\nabla g(x(t)+\beta(t)\dot{x}(t))+\epsilon(t)x(t),\dot{x}(t)\rangle=\\ &-\langle\nabla g(x(t)+\beta(t)\dot{x}(t))+\epsilon(t)x(t),x(t)+\beta(t)\dot{x}(t)-x^*\rangle\leq g(x^*)+\frac{\epsilon(t)}{2}\|x^*\|^2-g(x(t)+\beta(t)\dot{x}(t))\\ &-\frac{\epsilon(t)}{2}\|x(t)+\beta(t)\dot{x}(t)\|^2-\frac{\epsilon(t)}{2}\|x(t)+\beta(t)\dot{x}(t)-x^*\|^2+\beta(t)\epsilon(t)\langle\dot{x}(t),x(t)+\beta(t)\dot{x}(t)-x^*\rangle.\end{aligned}$$

From here we get

$$\begin{aligned}(2.6)\quad &-\langle\nabla g(x(t)+\beta(t)\dot{x}(t))+\epsilon(t)x(t),x(t)-x^*\rangle\leq\\ &-(g(x(t)+\beta(t)\dot{x}(t))-g(x^*))+\frac{\epsilon(t)}{2}\|x^*\|^2+\beta(t)\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),\dot{x}(t)\rangle\\ &-\frac{\epsilon(t)}{2}\|x(t)+\beta(t)\dot{x}(t)\|^2-\frac{\epsilon(t)}{2}\|x(t)+\beta(t)\dot{x}(t)-x^*\|^2+\beta(t)\epsilon(t)\langle\dot{x}(t),2x(t)+\beta(t)\dot{x}(t)-x^*\rangle.\end{aligned}$$

Further, an easy computation shows that

$$\begin{aligned}&-\frac{\epsilon(t)}{2}\|x(t)+\beta(t)\dot{x}(t)\|^2-\frac{\epsilon(t)}{2}\|x(t)+\beta(t)\dot{x}(t)-x^*\|^2+\beta(t)\epsilon(t)\langle\dot{x}(t),2x(t)+\beta(t)\dot{x}(t)-x^*\rangle=\\ &-\frac{\epsilon(t)}{2}\|x(t)\|^2-\frac{\epsilon(t)}{2}\|x(t)-x^*\|^2.\end{aligned}$$

Hence, (2.6) becomes

$$\begin{aligned}(2.7)\quad &-\langle\nabla g(x(t)+\beta(t)\dot{x}(t))+\epsilon(t)x(t),x(t)-x^*\rangle\leq-(g(x(t)+\beta(t)\dot{x}(t))-g(x^*))-\frac{\epsilon(t)}{2}\|x(t)\|^2\\ &-\frac{\epsilon(t)}{2}\|x(t)-x^*\|^2+\frac{\epsilon(t)}{2}\|x^*\|^2+\beta(t)\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),\dot{x}(t)\rangle.\end{aligned}$$

By multiplying (2.7) with bt and injecting in (2.5) we get

$$\begin{aligned}(2.8)\quad \dot{\mathcal{E}}(t)\leq&(a'(t)-bt)\left(g(x(t)+\beta(t)\dot{x}(t))-g^*\right)-\beta(t)a(t)\left\|\nabla g(x(t)+\beta(t)\dot{x}(t))\right\|^2+(b+1-\alpha)t\|\dot{x}(t)\|^2\\ &+bt\frac{\epsilon(t)}{2}\|x^*\|^2-bt\frac{\epsilon(t)}{2}\|x(t)-x^*\|^2+\left(\frac{t^2\dot{\epsilon}(t)}{2}+(2-b)t\frac{\epsilon(t)}{2}\right)\|x(t)\|^2\\ &-\beta(t)\epsilon(t)a(t)\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),x(t)\rangle\\ &+\left(\left(-\beta(t)\frac{\alpha}{t}+\beta'(t)+1\right)a(t)-t^2+b\beta(t)t\right)\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),\dot{x}(t)\rangle.\end{aligned}$$

We estimate

$$\left(-\beta(t)\frac{\alpha-b}{t}+\beta'(t)\right)t^2\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),\dot{x}(t)\rangle \leq$$
$$\frac{1}{2}\left|-\beta(t)\frac{\alpha-b}{t}+\beta'(t)\right|\left(t^{\frac{5}{2}}\|\nabla g(x(t)+\beta(t)\dot{x}(t))\|^2+t^{\frac{3}{2}}\|\dot{x}(t)\|^2\right).$$

Let us take now $a(t)=t^2$. Then, (2.8) becomes

$$\begin{aligned}\dot{\mathcal{E}}(t) &\leq (2-b)t\left(g(x(t)+\beta(t)\dot{x}(t))-g^*\right)-\left(\beta(t)t^2-\frac{1}{2}\left|-\beta(t)\frac{\alpha-b}{t}+\beta'(t)\right|t^{\frac{5}{2}}\right)\|\nabla g(x(t)+\beta(t)\dot{x}(t))\|^2\\ &+\left((b+1-\alpha)t+\frac{1}{2}\left|-\beta(t)\frac{\alpha-b}{t}+\beta'(t)\right|t^{\frac{3}{2}}\right)\|\dot{x}(t)\|^2+bt\frac{\epsilon(t)}{2}\|x^*\|^2-bt\frac{\epsilon(t)}{2}\|x(t)-x^*\|^2\\ &+\left(\frac{t^2\dot{\epsilon}(t)}{2}+(2-b)t\frac{\epsilon(t)}{2}\right)\|x(t)\|^2-\beta(t)\epsilon(t)t^2\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),x(t)\rangle.\end{aligned}\tag{2.9}$$

We will carry out the analysis by addressing the settings provided by the conditions (C1) and (C2) separately.

Condition (C1). Assuming that condition (C1) holds, there exist $K>1$ and $t_1\geq t_0$ such that

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

Using that

$$-\beta(t)\epsilon(t)t^2\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),x(t)\rangle\leq\frac{|\beta(t)|t^2}{K}\|\nabla g(x(t)+\beta(t)\dot{x}(t))\|^2+\frac{K|\beta(t)|\epsilon^2(t)t^2}{4}\|x(t)\|^2,\tag{2.10}$$

(2.9) leads to the following estimate

$$\begin{aligned}\dot{\mathcal{E}}(t) &\leq (2-b)t\left(g(x(t)+\beta(t)\dot{x}(t))-g^*\right)\\ &-\left(\left(\beta(t)-\frac{|\beta(t)|}{K}\right)t^2-\frac{1}{2}\left|-\beta(t)\frac{\alpha-b}{t}+\beta'(t)\right|t^{\frac{5}{2}}\right)\|\nabla g(x(t)+\beta(t)\dot{x}(t))\|^2\\ &+\left((b+1-\alpha)t+\frac{1}{2}\left|-\beta(t)\frac{\alpha-b}{t}+\beta'(t)\right|t^{\frac{3}{2}}\right)\|\dot{x}(t)\|^2+bt\frac{\epsilon(t)}{2}\|x^*\|^2-bt\frac{\epsilon(t)}{2}\|x(t)-x^*\|^2\\ &+\left(\frac{t^2\dot{\epsilon}(t)}{2}+\frac{K|\beta(t)|\epsilon^2(t)t^2}{4}+(2-b)t\frac{\epsilon(t)}{2}\right)\|x(t)\|^2,\ \text{ for all } t\geq t_1.\end{aligned}\tag{2.11}$$

Now, taking into account that $K>1$ and $\beta(t)=\gamma+\frac{\beta}{t}$ we conclude the following.
If $\gamma>0$, $\beta\in\mathbb{R}$ then there exists $t_1'\geq t_1$ and $r_1>0$ such that

$$-\left(\left(\beta(t)-\frac{|\beta(t)|}{K}\right)t^2-\frac{1}{2}\left|-\beta(t)\frac{\alpha-b}{t}+\beta'(t)\right|t^{\frac{5}{2}}\right)\leq-r_1t^2,\ \text{ for all } t\geq t_1'.$$

If $\gamma=0,\beta\geq0$ then there exists $t_1'\geq t_1$ and $r_1>0$ such that

$$-\left(\left(\beta(t)-\frac{|\beta(t)|}{K}\right)t^2-\frac{1}{2}\left|-\beta(t)\frac{\alpha-b}{t}+\beta'(t)\right|t^{\frac{5}{2}}\right)\leq-r_1\beta t,\ \text{ for all } t\geq t_1'.$$

Further, since $b<\alpha-1$ there exists $t_1''\geq t_1$ and $r_2>0$ such that

$$(b+1-\alpha)t+\frac{1}{2}\left|-\beta(t)\frac{\alpha-b}{t}+\beta'(t)\right|t^{\frac{3}{2}}\leq-r_2t,\ \text{ for all } t\geq t_1''.$$

Finally, according to assumption (C1) and the fact that $b>2$ we get

$$\frac{t^2\dot{\epsilon}(t)}{2}+\frac{K|\beta(t)|\epsilon^2(t)t^2}{4}+(2-b)t\frac{\epsilon(t)}{2}\leq(2-b)t\frac{\epsilon(t)}{2},\ \text{ for all } t\geq t_1.$$

Hence, by considering $t_2 = \max(t_1', t_1'')$ and denoting $r_1(t) = r_1 t^2$ when $\gamma > 0$, $\beta \in \mathbb{R}$ and $r_1(t) = r_1\beta t$ whenever $\gamma = 0$, $\beta \geq 0$, the relation (2.11) leads to

(2.12)
$$\dot{\mathcal{E}}(t) + (b-2)t\left(g(x(t)+\beta(t)\dot{x}(t)) - g^*\right) + r_1(t)\left\|\nabla g(x(t)+\beta(t)\dot{x}(t))\right\|^2 + r_2 t\|\dot{x}(t)\|^2 + bt\frac{\epsilon(t)}{2}\|x(t)-x^*\|^2$$
$$+ (b-2)t\frac{\epsilon(t)}{2}\|x(t)\|^2 \leq bt\frac{\epsilon(t)}{2}\|x^*\|^2, \text{ for all } t \geq t_2.$$

Condition (C2). Assume now that (C2) holds, that is, there exists $K > 0$ and $t_1 \geq t_0$ such that
$$\epsilon(t) \leq \frac{K}{t} \text{ for every } t \geq t_1.$$

Now, since $\beta(t) = \gamma + \frac{\beta}{t}$, where $\gamma > 0$, $\beta \in \mathbb{R}$ or $\gamma = 0$, $\beta \geq 0$ we obtain that there exists $\bar{t}_1 \geq t_1$ such that
$$\beta(t) \geq 0, \text{ for all } t \geq \bar{t}_1.$$
Using the monotonicity of ∇g and the fact that $\nabla g(x^*) = 0$ we get for all $t \geq \bar{t}_1$ that

$$\begin{aligned}(2.13)\qquad -\beta(t)\epsilon(t)t^2\langle\nabla g(x(t)+\beta(t)\dot{x}(t)), x(t)\rangle &= -\beta(t)\epsilon(t)t^2\langle\nabla g(x(t)+\beta(t)\dot{x}(t)), x(t)+\beta(t)\dot{x}(t)-x^*\rangle \\ &\quad + \beta(t)\epsilon(t)t^2\langle\nabla g(x(t)+\beta(t)\dot{x}(t)), \beta(t)\dot{x}(t)-x^*\rangle \\ &\leq \beta(t)\epsilon(t)t^2\langle\nabla g(x(t)+\beta(t)\dot{x}(t)), \beta(t)\dot{x}(t)-x^*\rangle.\end{aligned}$$

The right hand side of (2.13) becomes
$$\begin{aligned}\beta(t)\epsilon(t)t^2\langle\nabla g(x(t)+\beta(t)\dot{x}(t)), \beta(t)\dot{x}(t)-x^*\rangle &= \beta^2(t)\epsilon(t)t^2\langle\nabla g(x(t)+\beta(t)\dot{x}(t)), \dot{x}(t)\rangle \\ &\quad - \beta(t)\epsilon(t)t^2\langle\nabla g(x(t)+\beta(t)\dot{x}(t)), x^*\rangle,\end{aligned}$$
further
$$\begin{aligned}\beta^2(t)\epsilon(t)t^2\langle\nabla g(x(t)+\beta(t)\dot{x}(t)), \dot{x}(t)\rangle &\leq \frac{\beta(t)\epsilon(t)t^3}{4K}\|\nabla g(x(t)+\beta(t)\dot{x}(t))\|^2 + K\beta^3(t)\epsilon(t)t\|\dot{x}(t)\|^2 \\ &\leq \frac{\beta(t)t^2}{4}\|\nabla g(x(t)+\beta(t)\dot{x}(t))\|^2 + K^2\beta^3(t)\|\dot{x}(t)\|^2\end{aligned}$$
and
$$\begin{aligned}-\beta(t)\epsilon(t)t^2\langle\nabla g(x(t)+\beta(t)\dot{x}(t)), x^*\rangle &\leq \frac{\beta(t)\epsilon(t)t^3}{4K}\|\nabla g(x(t)+\beta(t)\dot{x}(t))\|^2 + K\beta(t)\epsilon(t)t\|x^*\|^2 \\ &\leq \frac{\beta(t)t^2}{4}\|\nabla g(x(t)+\beta(t)\dot{x}(t))\|^2 + K\beta(t)t\epsilon(t)\|x^*\|^2,\end{aligned}$$
for all $t \geq \bar{t}_1$. Hence, (2.13) becomes

$$\begin{aligned}(2.14)\qquad -\beta(t)\epsilon(t)t^2\langle\nabla g(x(t)+\beta(t)\dot{x}(t)), x(t)\rangle &\leq \frac{\beta(t)t^2}{2}\|\nabla g(x(t)+\beta(t)\dot{x}(t))\|^2 + K^2\beta^3(t)\|\dot{x}(t)\|^2 \\ &\quad + K\beta(t)t\epsilon(t)\|x^*\|^2, \text{ for all } t \geq \bar{t}_1.\end{aligned}$$

Now, injecting (2.14) in (2.9) we get

(2.15)
$$\dot{\mathcal{E}}(t) \leq (2-b)t\left(g(x(t)+\beta(t)\dot{x}(t)) - g^*\right) - \left(\frac{\beta(t)}{2}t^2 - \frac{1}{2}\left|-\beta(t)\frac{\alpha-b}{t} + \beta'(t)\right| t^{\frac{5}{2}}\right)\|\nabla g(x(t)+\beta(t)\dot{x}(t))\|^2$$
$$+ \left((b+1-\alpha)t + \frac{1}{2}\left|-\beta(t)\frac{\alpha-b}{t} + \beta'(t)\right| t^{\frac{3}{2}} + K^2\beta^3(t)\right)\|\dot{x}(t)\|^2 + (b+2K\beta(t))t\frac{\epsilon(t)}{2}\|x^*\|^2$$
$$- bt\frac{\epsilon(t)}{2}\|x(t)-x^*\|^2 + \left(\frac{t^2\dot{\epsilon}(t)}{2} + (2-b)t\frac{\epsilon(t)}{2}\right)\|x(t)\|^2, \text{ for all } t \geq \bar{t}_1.$$

Taking into account that $\beta(t)=\gamma+\frac{\beta}{t}$ we conclude the following.
If $\gamma>0$, $\beta\in\mathbb{R}$ then there exists $t_1'\geq\bar{t}_1$ and $r_1>0$ such that

$$-\left(\frac{\beta(t)}{2}t^2-\frac{1}{2}\left|-\beta(t)\frac{\alpha-b}{t}+\beta'(t)\right|t^{\frac{5}{2}}\right)\leq -r_1t^2,\ \text{for all } t\geq t_1'.$$

If $\gamma=0,\beta\geq 0$ then there exists $t_1'\geq\bar{t}_1$ and $r_1>0$ such that

$$-\left(\frac{\beta(t)}{2}t^2-\frac{1}{2}\left|-\beta(t)\frac{\alpha-b}{t}+\beta'(t)\right|t^{\frac{5}{2}}\right)\leq -r_1\beta t,\ \text{for all } t\geq t_1'.$$

Further, since $b<\alpha-1$ there exists $t_1''\geq\bar{t}_1$ and $r_2>0$ such that

$$(b+1-\alpha)t+\frac{1}{2}\left|-\beta(t)\frac{\alpha-b}{t}+\beta'(t)\right|t^{\frac{3}{2}}+K^2\beta^3(t)\leq -r_2t,\ \text{for all } t\geq t_1''.$$

Finally, according to the fact that $b>2$ and that $\epsilon(t)$ is decreasing we get

$$\frac{t^2\dot{\epsilon}(t)}{2}+(2-b)t\frac{\epsilon(t)}{2}\leq(2-b)t\frac{\epsilon(t)}{2},\ \text{for all } t\geq t_0.$$

Hence, by considering $\bar{t}_2=\max(t_1',t_1'')$ and denoting $r_1(t)=r_1t^2$ when $\gamma>0$, $\beta\in\mathbb{R}$ and $r_1(t)=r_1\beta t$ whenever $\gamma=0$, $\beta\geq 0$, the relation (2.15) leads to

(2.16)

$$\dot{\mathcal{E}}(t)+(b-2)t\left(g(x(t)+\beta(t)\dot{x}(t))-g^*\right)+r_1(t)\left\|\nabla g(x(t)+\beta(t)\dot{x}(t))\right\|^2+r_2t\|\dot{x}(t)\|^2+bt\frac{\epsilon(t)}{2}\|x(t)-x^*\|^2$$
$$+(b-2)t\frac{\epsilon(t)}{2}\|x(t)\|^2\leq(b+2K\beta(t))t\frac{\epsilon(t)}{2}\|x^*\|^2,\ \text{for all } t\geq\bar{t}_2.$$

The estimates. By integrating (2.12) on an interval $[t_3,t]$, $t_3=t_2$ in case the condition (C1) holds, and by integrating (2.16) on an interval $[t_3,t]$, $t_3=\bar{t}_2$ in case the condition (C2) holds, further denoting $l=b$ in case (C1) holds and $l=\sup_{t\geq t_3}(b+2K\beta(t))$ in case (C2) holds, we obtain for every $t\geq t_3$ that

(2.17)

$$\mathcal{E}(t)+(b-2)\int_{t_3}^t s\left(g(x(s)+\beta(s)\dot{x}(s))-g^*\right)ds+\int_{t_3}^t r_1(s)\left\|\nabla g(x(s)+\beta(s)\dot{x}(s))\right\|^2ds+r_2\int_{t_3}^t s\|\dot{x}(s)\|^2ds$$
$$+\frac{b}{2}\int_{t_3}^t s\epsilon(s)\|x(s)-x^*\|^2+\frac{b-2}{2}\int_{t_3}^t s\epsilon(s)\|x(s)\|^2ds\leq\frac{l}{2}\int_{t_3}^t s\epsilon(s)\|x^*\|^2ds+\mathcal{E}(t_3).$$

For proving (i) assume that $\int_{t_0}^{+\infty}\frac{\epsilon(s)}{s}ds<+\infty$. Then, from (2.17) we get that for all $t\geq t_3$ one has

$$0\leq g(x(t)+\beta(t)\dot{x}(t))-\min g\leq\frac{\mathcal{E}(t_3)}{t^2}+\frac{l\|x^*\|^2}{2}\frac{1}{t^2}\int_{t_2}^t s\epsilon(s)ds,$$

$$0\leq\left\|\frac{b}{t}(x(t)-x^*)+\dot{x}(t)\right\|^2\leq\frac{2\mathcal{E}(t_3)}{t^2}+\frac{l\|x^*\|^2}{t^2}\int_{t_3}^t s\epsilon(s)ds$$

and

$$0\leq\left\|\frac{x(t)-x^*}{t}\right\|^2\leq\frac{2\mathcal{E}(t_3)}{b(\alpha-1-b)t^2}+\frac{l\|x^*\|^2}{b(\alpha-1-b)t^2}\int_{t_3}^t s\epsilon(s)ds.$$

Obviously, $\lim_{t\to+\infty}\frac{\mathcal{E}(t_3)}{t^2}=0$. Further, Lemma B.1 applied to the functions $\varphi(s)=s^2$ and $f(s)=\frac{\epsilon(s)}{s}$ provides $\lim_{t\to+\infty}\frac{1}{t^2}\int_{t_2}^t s^2\frac{\epsilon(s)}{s}dt=0$. Hence,
$\lim_{t\to+\infty}g(x(t)+\beta(t)\dot{x}(t))=\min g$, $\lim_{t\to+\infty}\left\|\frac{b}{t}(x(t)-x^*)+\dot{x}(t)\right\|=0$ and $\lim_{t\to+\infty}\left\|\frac{x(t)-x^*}{t}\right\|=0.$

Combining the last two relations we get $\lim_{t\to+\infty}\|\dot{x}(t)\|=0$, and from here and the continuity of g we have $\lim_{t\to+\infty} g(x(t))=\lim_{t\to+\infty} g(x(t)+\beta(t)\dot{x}(t))=\min g$.

For proving (ii) assume that $\int_{t_0}^{+\infty} s\epsilon(s)ds<+\infty$. Then, $C_0=l\int_{t_3}^{+\infty} t\frac{\epsilon(t)}{2}\|x^*\|^2dt+\mathcal{E}(t_3)<+\infty$ and from (2.17) we immediately deduce that $\mathcal{E}(t)\le C_0$ for all $t\ge t_3$, hence

$$g(x(t)+\beta(t)\dot{x}(t))-g^*=\mathcal{O}\left(\frac{1}{t^2}\right), \text{ as } t\to+\infty, \tag{2.18}$$

$$\sup_{t\ge t_0}\|b(x(t)-x^*)+t\dot{x}(t)\|^2<+\infty, \tag{2.19}$$

$$\sup_{t\ge t_0}\|x(t)-x^*\|^2<+\infty. \tag{2.20}$$

Further, (2.17) yields

$$\int_{t_0}^{+\infty} t\left(g(x(t)+\beta(t)\dot{x}(t))-g^*\right)dt<+\infty, \tag{2.21}$$

$$\int_{t_0}^{+\infty} t^2\left\|\nabla g(x(t)+\beta(t)\dot{x}(t))\right\|^2dt<+\infty, \text{ whenever } \gamma>0,\ \beta\in\mathbb{R}, \tag{2.22}$$

$$\beta\int_{t_0}^{+\infty} t\left\|\nabla g(x(t)+\beta(t)\dot{x}(t))\right\|^2dt<+\infty, \text{ whenever } \gamma=0,\ \beta\ge 0, \tag{2.23}$$

$$\int_{t_0}^{+\infty} t\epsilon(t)\|x(t)\|^2dt<+\infty, \tag{2.24}$$

and

$$\int_{t_0}^{+\infty} t\|\dot{x}(t)\|^2dt<+\infty. \tag{2.25}$$

Observe that (2.20) leads to the fact that the trajectory $x(t)$ is bounded, which combined with (2.19) shows that $\|t\dot{x}(t)\|^2$ is bounded, that is

$$\|\dot{x}(t)\|=\mathcal{O}\left(\frac{1}{t}\right), \text{ as } t\to+\infty. \tag{2.26}$$

Note that (2.26) shows in particular that $\dot{x}(t)\to 0$, $t\to+\infty$.

In order to show that $\|\dot{x}(t)\|=o\left(\frac{1}{t}\right)$, as $t\to+\infty$ assume for now that the limit $\lim_{t\to+\infty}\|x(t)-x^*\|$ exists, as will be shown in the sequel. Then, (2.12) in case (C1) and (2.16) in case (C2) provide that

$$\dot{\mathcal{E}}(t)\le\frac{l\|x^*\|^2}{2}t\epsilon(t), \text{ for all } t\ge t_3,$$

where $l=b$ in case (C1) holds and $l=\sup_{t\ge t_3}(b+2K\beta(t))$ in case (C2) holds. Obviously, by the hypotheses we have $\frac{l\|x^*\|^2}{2}t\epsilon(t)\in L^1([t_3,+\infty))$, hence, according to Lemma B.2 there exists the limit $\lim_{t\to+\infty}\mathcal{E}(t)$.

Hence, since $\lim_{t\to+\infty}\|x(t)-x^*\|$ exists we get that the limit

$$\lim_{t\to+\infty} t^2\left(g(x(t)+\beta(t)\dot{x}(t))-g^*\right)+\frac{t^2\epsilon(t)}{2}\|x(t)\|^2+\frac{1}{2}\|t\dot{x}(t)\|^2 \tag{2.27}$$

also exists.

Now, (2.21), (2.24) and (2.25) yield

$$\int_{t_0}^{+\infty}\frac{1}{t}\left(t^2\left(g(x(t)+\beta(t)\dot{x}(t))-g^*\right)+\frac{t^2\epsilon(t)}{2}\|x(t)\|^2+\frac{1}{2}\|t\dot{x}(t)\|^2\right)dt<+\infty. \tag{2.28}$$

Since the function $t\rightarrowtail\frac{1}{t}\notin L^1([t_0,+\infty))$, (2.28) and (2.27) lead to

$$\lim_{t\to+\infty} t^2\left(g(x(t)+\beta(t)\dot{x}(t))-g^*\right)+\frac{t^2\epsilon(t)}{2}\|x(t)\|^2+\frac{1}{2}\|t\dot{x}(t)\|^2=0. \tag{2.29}$$

Consequently,

$$g(x(t)+\beta(t)\dot{x}(t))-\min g=o\left(\frac{1}{t^2}\right) \text{ as } t\to+\infty,$$

$$\|x(t)\|=o\left(\frac{1}{t\sqrt{\epsilon(t)}}\right) \text{ as } t\to+\infty,$$

and

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

The limit. To prove the existence of the weak limit of $x(t)$, we use the Opial lemma, (see Lemma B.4 at Appendix). For $z\in\operatorname{argmin} g$ let us introduce the anchor function $h_z(t)=\frac{1}{2}\|x(t)-z\|^2$. The classical derivation chain rule gives $\ddot{h}_z(t)+\frac{\alpha}{t}\dot{h}_z(t)=\left\langle\ddot{x}(t)+\frac{\alpha}{t}\dot{x}(t),x(t)-z\right\rangle+\|\dot{x}(t)\|^2$. Now, using (1.1) we get $\ddot{h}_z(t)+\frac{\alpha}{t}\dot{h}_z(t)=\langle-\epsilon(t)x(t)-\nabla g(x(t)+\beta(t)\dot{x}(t)),x(t)-z\rangle+\|\dot{x}(t)\|^2$. In other words,

$$t\ddot{h}_z(t)+\alpha\dot{h}_z(t)+t\left\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),x(t)-z\right\rangle=t\|\dot{x}(t)\|^2-\left\langle t\epsilon(t)x(t),x(t)-z\right\rangle. \tag{2.30}$$

We have

$$\begin{aligned}t\left\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),x(t)-z\right\rangle&=t\left\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),(x(t)+\beta(t)\dot{x}(t))-z\right\rangle\\&\quad-t\left\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),\beta(t)\dot{x}(t)\right\rangle.\end{aligned}$$

Consequently, (2.30) becomes

$$\begin{aligned}t\ddot{h}_z(t)+\alpha\dot{h}_z(t)+t\left\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),x(t)+\beta(t)\dot{x}(t)-z\right\rangle=&\,t\|\dot{x}(t)\|^2-\left\langle t\epsilon(t)x(t),x(t)-z\right\rangle\\&+t\left\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),\beta(t)\dot{x}(t)\right\rangle.\end{aligned} \tag{2.31}$$

Now, since $x(t)$ is bounded, there exists $K_1>0$ such that

$$-\left\langle t\epsilon(t)x(t),x(t)-z\right\rangle\le t\epsilon(t)\|x(t)\|\|x(t)-z\|\le K_1t\epsilon(t).$$

Further, $t\left\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),\beta(t)\dot{x}(t)\right\rangle\le\frac{1}{2}t|\beta(t)|\|\nabla g(x(t)+\beta(t)\dot{x}(t))\|^2+\frac{1}{2}t|\beta(t)|\|\dot{x}(t)\|^2$.

The latter two inequalities combined with (2.31) yield

$$\begin{aligned}t\ddot{h}_z(t)+\alpha\dot{h}_z(t)+t\left\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),x(t)+\beta(t)\dot{x}(t)-z\right\rangle\le&\left(1+\frac{1}{2}|\beta(t)|\right)t\|\dot{x}(t)\|^2+K_1t\epsilon(t)\\&+\frac{1}{2}t|\beta(t)|\|\nabla g(x(t)+\beta(t)\dot{x}(t))\|^2.\end{aligned} \tag{2.32}$$

Now, by the monotonicity of ∇g we have that the function $\theta(t)=t\left\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),x(t)+\beta(t)\dot{x}(t)-z\right\rangle$ is nonnegative on $[t_0,+\infty)$.

Further, (2.25) and (2.22) if $\gamma>0$, $\beta\in\mathbb{R}$ and (2.23) if $\gamma=0$, $\beta\ge0$ and the hypotheses of the theorem shows that the function $k(t)=\left(1+\frac{1}{2}|\beta(t)|\right)t\|\dot{x}(t)\|^2+K_1t\epsilon(t)+\frac{1}{2}t|\beta(t)|\|\nabla g(x(t)+\beta(t)\dot{x}(t))\|^2$ belongs to $L^1(t_0,+\infty)$.

Hence, Lemma B.5 can be applied for the function $w(t)=h_z(t)$, thus we infer that the following limit exists

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

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

On one hand the function g is weakly lower semicontinuous, since is convex and continuous, hence we have that $g(\overline{x})\le\liminf_{n\to+\infty}g(x(t_n))$. On the other hand, according to (i), $\lim_{t\to+\infty}g(x(t))=\min g$, consequently one has $g(\overline{x})\le\min g$, which shows that $\overline{x}\in\operatorname{argmin} g$.

According to Opial lemma it follows that

$$w-\lim_{t\to+\infty}x(t)\in\operatorname{argmin} g.$$

□

REMARK 2. In Theorem 2.1 we have shown that under the assumptions that $\int_{t_0}^{+\infty} t\epsilon(t)dt < +\infty$, $\alpha > 3$ and $\gamma > 0$, $\beta \in \mathbb{R}$ or $\gamma = 0$, $\beta \geq 0$ one has $g\left(x(t) + \left(\gamma + \frac{\beta}{t}\right)\dot{x}(t)\right) - \min g = o\left(\frac{1}{t^2}\right)$ as $t \to +\infty$. However, under the assumptions $\gamma = 0$ and $\beta > 0$ and ∇g is globally L_g-Lipschitz continuous, we can even show that $g(x(t)) - \min g$ is of order $o\left(\frac{1}{t^2}\right)$ as $t \to +\infty$, and for this is enough to prove that $g(x(t)) - g\left(x(t) + \frac{\beta}{t}\dot{x}(t)\right) = o\left(\frac{1}{t^2}\right)$, as $t \to +\infty$. In order to obtain the latter, we use the well-known descent lemma from [31] and we obtain that for all $t \geq t_0$ one has

$$\begin{aligned} g(x(t)) - g\left(x(t) + \frac{\beta}{t}\dot{x}(t)\right) &\leq \left\langle \nabla g\left(x(t) + \frac{\beta}{t}\dot{x}(t)\right), -\frac{\beta}{t}\dot{x}(t)\right\rangle + \frac{L_g}{2}\left\|\frac{\beta}{t}\dot{x}(t)\right\|^2 \\ &\leq \frac{\beta}{t}\|\dot{x}(t)\| \cdot \left\|\nabla g\left(x(t) + \frac{\beta}{t}\dot{x}(t)\right)\right\| + \frac{L_g}{2}\left(\frac{\beta}{t}\right)^2\|\dot{x}(t)\|^2. \end{aligned}$$

From Theorem 2.1, we have that x and $\dot{x}$ are bounded and $\|\dot{x}(t)\| = o\left(\frac{1}{t}\right)$ as $t \to +\infty$, hence by using the continuity of ∇g we get $\frac{\beta}{t}\|\dot{x}(t)\| \cdot \left\|\nabla g\left(x(t) + \frac{\beta}{t}\dot{x}(t)\right)\right\| = o\left(\frac{1}{t^2}\right)$, as $t \to +\infty$. Moreover, we have $\frac{L_g}{2}\left(\frac{\beta}{t}\right)^2\|\dot{x}(t)\|^2 = o\left(\frac{1}{t^4}\right)$, as $t \to +\infty$. By combining the previous relations the result follows.

REMARK 3. Observe that the assumptions of Theorem 2.1 (ii) are satisfied for $\epsilon(t) = \frac{a}{t^r}$, $r > 2$, $a > 0$. More precisely, in this case the conditions (C1), (C2) and the relation $\int_{t_0}^{+\infty} t\epsilon(t)dt < +\infty$ hold. The latter relation was essential in the proof of Theorem 2.1 (ii) in order to show the pointwise and integral estimates but also the weak convergence of the trajectories. Nevertheless, by deploying the techniques used in [16], we can show the fast convergence of the function values in the generated trajectories even for $\epsilon(t) = \frac{a}{t^2}$, $a > 0$. The following result holds.

THEOREM 2.2. *Let $t_0 > 0$, $\alpha > 3$ and $\epsilon(t) = \frac{a}{t^2}$, $a > 0$. For some starting points $u_0, v_0 \in \mathcal{H}$ let $x : [t_0, \infty) \to \mathcal{H}$ be the unique global solution of (1.1). Then, x is bounded, $g\left(x(t) + \left(\gamma + \frac{\beta}{t}\right)\dot{x}(t)\right) - \min g = \mathcal{O}\left(\frac{1}{t^2}\right)$ and $\|\dot{x}(t)\| = \mathcal{O}\left(\frac{1}{t}\right)$ as $t \to +\infty$.*

PROOF . Note that for $\epsilon(t) = \frac{a}{t^2}$, $a > 0$ both the conditions (C1) and (C2) hold. Now, by using the fact that $\frac{1}{2}\|b(x(t) - x^*) + t\dot{x}(t)\|^2 \leq b^2\|x(t) - x^*\|^2 + t^2\|\dot{x}(t)\|^2$, from (2.1) we get that for all $t \geq t_0$ it holds

$$\mathcal{E}(t) \leq t^2\left(g(x(t) + \beta(t)\dot{x}(t)) - g^*\right) + \frac{a}{2}\|x(t)\|^2 + t^2\|\dot{x}(t)\|^2 + \frac{b(\alpha - 1 + b)}{2}\|x(t) - x^*\|^2 \tag{2.33}$$

Further, (2.12) gives

(2.34)
$$\begin{aligned} \dot{\mathcal{E}}(t) \leq &-(b-2)t\left(g(x(t) + \beta(t)\dot{x}(t)) - g^*\right) - r_1(t)\left\|\nabla g(x(t) + \beta(t)\dot{x}(t))\right\|^2 - r_2 t\|\dot{x}(t)\|^2 - \frac{ab}{2t}\|x(t) - x^*\|^2 \\ &-(b-2)\frac{a}{2t}\|x(t)\|^2 + \frac{ab}{2t}\|x^*\|^2, \text{ for all } t \geq t_2, \end{aligned}$$

where t_2, $r_1(t)$ and r_2 were defined in the proof of Theorem 2.1. More precisely, $r_1(t) = r_1 t^2$ when $\gamma > 0$, $\beta \in \mathbb{R}$ and $r_1(t) = r_1\beta t$ whenever $\gamma = 0$, $\beta \geq 0$ and $r_1, r_2 > 0$.

Let now $0 < c \leq \min\left(b - 2, \frac{a}{\alpha - 1 + b}, r_2\right)$. By multiplying (2.33) with $\frac{c}{t}$ and adding to (2.34) we get

(2.35)
$$\begin{aligned} \dot{\mathcal{E}}(t) + \frac{c}{t}\mathcal{E}(t) \leq &(c - b + 2)t\left(g(x(t) + \beta(t)\dot{x}(t)) - g^*\right) - r_1(t)\left\|\nabla g(x(t) + \beta(t)\dot{x}(t))\right\|^2 + (c - r_2)t\|\dot{x}(t)\|^2 \\ &+ \frac{b}{2t}(c(\alpha - 1 + b) - a)\|x(t) - x^*\|^2 + (c - b + 2)\frac{a}{2t}\|x(t)\|^2 + \frac{ab}{2t}\|x^*\|^2 \\ \leq &\frac{ab}{2t}\|x^*\|^2, \text{ for all } t \geq t_2. \end{aligned}$$

We multiply the last relation with t^c and we get

$$\frac{d}{dt}(t^c\mathcal{E}(t)) \le \frac{ab}{2}\|x^*\|^2 t^{c-1}, \text{ for all } t \ge t_2. \tag{2.36}$$

Consequently, by integrating (2.36) on an interval $[t_2, t]$ one has

$$t^c\mathcal{E}(t) - t_2^c\mathcal{E}(t_2) \le \frac{ab}{2}\|x^*\|^2 \int_{t_2}^{t} s^{c-1} ds = \frac{ab}{2c}\|x^*\|^2(t^c - t_2^c).$$

Hence, there exists $\overline{K} > 0$ such that $\mathcal{E}(t) \le \overline{K}$, for all $t \ge t_0$, and from here and (2.1) we deduce that x is bounded, $g\left(x(t) + \left(\gamma + \frac{\beta}{t}\right)\dot{x}(t)\right) - \min g = \mathcal{O}\left(\frac{1}{t^2}\right)$ and $\|\dot{x}(t)\| = \mathcal{O}\left(\frac{1}{t}\right)$ as $t \to +\infty$. □

2.2. The critical case $\alpha = 3$. As we mentioned before, just as for the dynamical system $(\text{AVD})_\alpha$, the case $\alpha = 3$ is critical. In this case we obtain $\mathcal{O}(1/t^2)$ convergence rate for the decay $g\left(x(t) + \left(\gamma + \frac{\beta}{t}\right)\dot{x}(t)\right) - \min g$ and we also obtain some integral estimates for the gradient of the objective function, meanwhile the weak convergence of the trajectories to a minimizer of g remains an open question.

THEOREM 2.3. *Let $t_0 > 0$ and for some starting points $u_0, v_0 \in \mathcal{H}$ let $x : [t_0, \infty) \longrightarrow \mathcal{H}$ be the unique global solution of (1.1). Assume that $\alpha = 3$ and that one of the conditions* (C1) *or* (C2) *is fulfilled. The following statements hold.*

(i) If $\int_{t_0}^{+\infty} \frac{\varepsilon(t)}{t} dt < +\infty$, then $\lim_{t\to+\infty} g\left(x(t) + \left(\gamma + \frac{\beta}{t}\right)\dot{x}(t)\right) = \min g$.

(ii) If $\int_{t_0}^{+\infty} t\varepsilon(t) dt < +\infty$, then $g\left(x(t) + \left(\gamma + \dfrac{\beta}{t}\right)\dot{x}(t)\right) - \min g = \mathcal{O}\left(\dfrac{1}{t^2}\right)$ as $t \to +\infty$. Further, $t\epsilon(t)\|x(t) - x^*\|^2 \in L^1([t_0, +\infty), \mathbb{R})$ and $t^2\left\|\nabla g\left(x(t) + \left(\gamma + \frac{\beta}{t}\right)\dot{x}(t)\right)\right\|^2 \in L^1([t_0, +\infty), \mathbb{R})$, for $\gamma > 0$, $\beta \in \mathbb{R}$ and $t\left\|\nabla g\left(x(t) + \frac{\beta}{t}\dot{x}(t)\right)\right\|^2 \in L^1([t_0, +\infty), \mathbb{R})$, for $\gamma = 0$, $\beta > 0$.

PROOF . We will use the same notations as in the proof of Theorem 2.1. Let $a(t) = \frac{t^2 - 2\beta(t)t}{1 - \frac{3}{t}\beta(t) + \beta'(t)}$, in case condition (C1) holds and $a(t) = \frac{t^2 - 2\beta(t)t}{1 - \frac{3}{t}\beta(t) + \beta'(t) + \beta^2(t)\epsilon(t)}$, in case condition (C2) holds.

Since $\epsilon(t) \to 0$, $t \to +\infty$, clearly, there exists $\bar{t}_0 \ge t_0$ such that $\frac{t^2 - 2\beta(t)t}{1 - \frac{3}{t}\beta(t) + \beta'(t) + \beta^2(t)\epsilon(t)} > 0$ for all $t \ge \bar{t}_0$, hence $a(t) > 0$ for all $t \ge \bar{t}_0$ in both cases (C1) and (C2).

The energy functional (2.1), for $\alpha = 3$, $b = 2$, becomes $\mathcal{E} : [\bar{t}_0, \infty) \to \mathbb{R}$

$$\mathcal{E}(t) = a(t)\left(g(x(t) + \beta(t)\dot{x}(t)) - g^*\right) + \frac{t^2\epsilon(t)}{2}\|x(t)\|^2 + \frac{1}{2}\|2(x(t) - x^*) + t\dot{x}(t)\|^2. \tag{2.37}$$

The same reasoning as in the proof of Theorem 2.1 holds, hence in this case (2.8) becomes

$$\begin{aligned}
\dot{\mathcal{E}}(t) \le{}& (a'(t) - 2t)\left(g(x(t) + \beta(t)\dot{x}(t)) - g^*\right) - \beta(t)a(t)\left\|\nabla g(x(t) + \beta(t)\dot{x}(t))\right\|^2 \\
&+ t\epsilon(t)\|x^*\|^2 - t\epsilon(t)\|x(t) - x^*\|^2 + \frac{t^2\dot{\epsilon}(t)}{2}\|x(t)\|^2 \\
&+ \left(\left(-\beta(t)\frac{3}{t} + \beta'(t) + 1\right)a(t) - t^2 + 2\beta(t)t\right)\langle\nabla g(x(t) + \beta(t)\dot{x}(t)), \dot{x}(t)\rangle \\
&- \beta(t)\epsilon(t)a(t)\langle\nabla g(x(t) + \beta(t)\dot{x}(t)), x(t)\rangle.
\end{aligned} \tag{2.38}$$

We will carry out the analysis by addressing the settings provided by the conditions (C1) and (C2) separately.

Condition (C1). Assuming that condition (C1) holds, one has

$$\left(-\beta(t)\frac{3}{t} + \beta'(t) + 1\right)a(t) - t^2 + 2\beta(t)t = 0, \text{ for all } t \ge \bar{t}_0.$$

Consequently, (2.38) becomes

$$\dot{\mathcal{E}}(t) \le (a'(t)-2t)\left(g(x(t)+\beta(t)\dot{x}(t))-g^*\right)-\beta(t)a(t)\left\|\nabla g(x(t)+\beta(t)\dot{x}(t))\right\|^2 + t\epsilon(t)\|x^*\|^2 - t\epsilon(t)\|x(t)-x^*\|^2+\frac{t^2\dot{\epsilon}(t)}{2}\|x(t)\|^2-\beta(t)\epsilon(t)a(t)\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),x(t)\rangle, \tag{2.39}$$

for all $t\ge \bar{t}_0$.

Since we are in the setting (C1), we have that there exist $K>1$ and $t_1\ge t_0$ such that

$$\dot{\epsilon}(t)\le -\frac{K|\beta(t)|}{2}\epsilon^2(t)\ \text{ for every } t\ge t_1.$$

Let $\bar{t}_1=\max(\bar{t}_0,t_1)$. Using that for every $r>0$ one has

$$-\beta(t)\epsilon(t)a(t)\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),x(t)\rangle\le\frac{|\beta(t)|a(t)}{r}\|\nabla g(x(t)+\beta(t)\dot{x}(t))\|^2+\frac{r|\beta(t)|\epsilon^2(t)a(t)}{4}\|x(t)\|^2, \tag{2.40}$$

for all $t\ge\bar{t}_1$, (2.39) leads to the following estimate

$$\dot{\mathcal{E}}(t) \le (a'(t)-2t)\left(g(x(t)+\beta(t)\dot{x}(t))-g^*\right)+\left(-\beta(t)a(t)+\frac{|\beta(t)|a(t)}{r}\right)\left\|\nabla g(x(t)+\beta(t)\dot{x}(t))\right\|^2 + t\epsilon(t)\|x^*\|^2 - t\epsilon(t)\|x(t)-x^*\|^2+\left(\frac{t^2\dot{\epsilon}(t)}{2}+\frac{r|\beta(t)|\epsilon^2(t)a(t)}{4}\right)\|x(t)\|^2,\ \text{for all } t\ge\bar{t}_1. \tag{2.41}$$

Take r such that $K>r>1$. Now, taking into account that $\beta(t)=\gamma+\frac{\beta}{t}$ and $a(t)=t^2+\gamma t+3\gamma^3+2\beta+\frac{\gamma(9\gamma^2+10\beta)t+4\beta(3\gamma^2+2\beta)}{t^2-3\gamma t-4\beta}$ we conclude the following.
If $\gamma>0$, $\beta\in\mathbb{R}$ then there exists $\bar{t}_1'\ge\bar{t}_1$ and $r_1>0$ such that $-\beta(t)a(t)+\frac{|\beta(t)|a(t)}{r}\le -r_1t^2,$ for all $t\ge\bar{t}_1'$.
If $\gamma=0,\beta\ge0$ then there exists $t_1'\ge\bar{t}_1$ and $r_1>0$ such that $-\beta(t)a(t)+\frac{|\beta(t)|a(t)}{r}\le -r_1\beta t,$ for all $t\ge\bar{t}_1'$.

Further, according to assumption (C1) and the fact that $r\in(1,K)$ we get that there exists $\bar{t}_1''\ge\bar{t}_1$ such that $\frac{t^2\dot{\epsilon}(t)}{2}+\frac{r|\beta(t)|\epsilon^2(t)a(t)}{4}\le 0,$ for all $t\ge\bar{t}_1''$.

Finally, if $\gamma>0$, $\beta\in\mathbb{R}$, then there exists $r_2>0$ and $\bar{t}_1'''\ge\bar{t}_1$ such that $a'(t)-2t\le r_2,$ for all $t\ge\bar{t}_1'''$, and if $\gamma=0,\beta\ge0$, then $a'(t)-2t\le 0,$ for all $t\ge\bar{t}_1$.

Hence, by considering $t_2=\max(\bar{t}_1',\bar{t}_1'',\bar{t}_1''')$ and denoting $r_1(t)=r_1t^2$ when $\gamma>0$, $\beta\in\mathbb{R}$ and $r_1(t)=r_1\beta t$ whenever $\gamma=0$, $\beta\ge0$, the relation (2.41) leads to

$$\dot{\mathcal{E}}(t)\le s\left(g(x(t)+\beta(t)\dot{x}(t))-g^*\right)-r_1(t)\left\|\nabla g(x(t)+\beta(t)\dot{x}(t))\right\|^2+t\epsilon(t)\|x^*\|^2-t\epsilon(t)\|x(t)-x^*\|^2, \tag{2.42}$$

for all $t\ge t_2$, where $s=r_2$ if $\gamma>0$, $\beta\in\mathbb{R}$ and $s=0$ if $\gamma=0$, $\beta\ge0$.

Condition (C2). Assume now that (C2) holds, that is, there exists $K>0$ and $t_1\ge t_0$ such that

$$\epsilon(t)\le\frac{K}{t}\ \text{for every } t\ge t_1.$$

Now, since $\beta(t)=\gamma+\frac{\beta}{t}$, where $\gamma>0$, $\beta\in\mathbb{R}$ or $\gamma=0$, $\beta\ge0$ and $a(t)=\frac{t^2-2\beta(t)t}{1-\frac{3}{t}\beta(t)+\beta'(t)+\beta^2(t)\epsilon(t)}>0$ if $t\ge\bar{t}_0$, we obtain that there exists $\bar{t}_1\ge\max(t_1,\bar{t}_0)$ such that $\beta(t)\ge0$ and $a(t)>0$ for all $t\ge\bar{t}_1$. Using the monotonicity of ∇g and the fact that $\nabla g(x^*)=0$ we get for all $t\ge\bar{t}_1$ that

$$\begin{aligned}-\beta(t)\epsilon(t)a(t)\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),x(t)\rangle &= -\beta(t)\epsilon(t)a(t)\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),x(t)+\beta(t)\dot{x}(t)-x^*\rangle\\ &\quad+\beta(t)\epsilon(t)a(t)\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),\beta(t)\dot{x}(t)-x^*\rangle\\ &\le\beta(t)\epsilon(t)a(t)\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),\beta(t)\dot{x}(t)-x^*\rangle.\end{aligned} \tag{2.43}$$

We estimate the right hand side of (2.43) as follows.

$$\beta(t)\epsilon(t)a(t)\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),\beta(t)\dot{x}(t)-x^*\rangle=\beta^2(t)\epsilon(t)a(t)\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),\dot{x}(t)\rangle$$
$$-\beta(t)\epsilon(t)a(t)\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),x^*\rangle,$$

further,

$$-\beta(t)\epsilon(t)a(t)\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),x^*\rangle\le\frac{\beta(t)\epsilon(t)a(t)^{\frac{3}{2}}}{4K}\|\nabla g(x(t)+\beta(t)\dot{x}(t))\|^2+K\beta(t)\epsilon(t)\sqrt{a(t)}\|x^*\|^2$$
$$\le\frac{\beta(t)a(t)^{\frac{3}{2}}}{4t}\|\nabla g(x(t)+\beta(t)\dot{x}(t))\|^2+K\beta(t)\epsilon(t)\sqrt{a(t)}\|x^*\|^2,$$

for all $t\ge\bar{t}_1$. Hence, (2.43) becomes

$$-\beta(t)\epsilon(t)t^2\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),x(t)\rangle\le\beta^2(t)\epsilon(t)a(t)\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),\dot{x}(t)\rangle$$
$$+\frac{\beta(t)a(t)^{\frac{3}{2}}}{4t}\|\nabla g(x(t)+\beta(t)\dot{x}(t))\|^2+K\beta(t)\epsilon(t)\sqrt{a(t)}\|x^*\|^2,\tag{2.44}$$

for all $t\ge\bar{t}_1$.

Now, injecting (2.44) in (2.38) we get

$$\dot{\mathcal{E}}(t)\le(a'(t)-2t)\left(g(x(t)+\beta(t)\dot{x}(t))-g^*\right)-\left(\beta(t)a(t)-\frac{\beta(t)a(t)^{\frac{3}{2}}}{4t}\right)\|\nabla g(x(t)+\beta(t)\dot{x}(t))\|^2$$
$$+(t\epsilon(t)+K\beta(t)\epsilon(t)\sqrt{a(t)})\|x^*\|^2-t\epsilon(t)\|x(t)-x^*\|^2+\frac{t^2\dot{\epsilon}(t)}{2}\|x(t)\|^2$$
$$+\left(\left(-\beta(t)\frac{3}{t}+\beta'(t)+1+\beta^2(t)\epsilon(t)\right)a(t)-t^2+2\beta(t)t\right)\langle\nabla g(x(t)+\beta(t)\dot{x}(t)),\dot{x}(t)\rangle,\ \text{for all } t\ge\bar{t}_1.\tag{2.45}$$

According to the assumptions one has $\left(-\beta(t)\frac{3}{t}+\beta'(t)+1+\beta^2(t)\epsilon(t)\right)a(t)-t^2+2\beta(t)t=0$, hence,

$$\dot{\mathcal{E}}(t)\le(a'(t)-2t)\left(g(x(t)+\beta(t)\dot{x}(t))-g^*\right)-\left(\beta(t)a(t)-\frac{\beta(t)a(t)^{\frac{3}{2}}}{4t}\right)\|\nabla g(x(t)+\beta(t)\dot{x}(t))\|^2$$
$$+(t\epsilon(t)+K\beta(t)\epsilon(t)\sqrt{a(t)})\|x^*\|^2-t\epsilon(t)\|x(t)-x^*\|^2+\frac{t^2\dot{\epsilon}(t)}{2}\|x(t)\|^2,\ \text{for all } t\ge\bar{t}_1.\tag{2.46}$$

Taking into account that $\beta(t)=\gamma+\frac{\beta}{t}$ we conclude the following.

If $\gamma>0$, $\beta\in\mathbb{R}$ then there exists $t_1'\ge\bar{t}_1$ and $r_1>0$ such that $-\left(\beta(t)a(t)-\frac{\beta(t)a(t)^{\frac{3}{2}}}{4t}\right)\le-r_1t^2$, for all $t\ge t_1'$.

If $\gamma=0,\beta\ge0$ then there exists $t_1'\ge\bar{t}_1$ and $r_1>0$ such that $-\left(\beta(t)a(t)-\frac{\beta(t)a(t)^{\frac{3}{2}}}{4t}\right)\le-r_1\beta t$, for all $t\ge t_1'$.

Further, there exists $t_1''\ge\bar{t}_1$ and $r_2>0$ such that $t\epsilon(t)+K\beta(t)\epsilon(t)\sqrt{a(t)}\le r_2t\epsilon(t),$ for all $t\ge t_1''$.

Finally, there exists $r_3\ge0$ and $t_1'''>\bar{t}_1$ such that $a'(t)-2t\le r_3,$ for all $t\ge t_1'''$.

Hence, by considering $\bar{t}_2=\max(t_1',t_1'',t_1''')$ and denoting $r_1(t)=r_1t^2$ when $\gamma>0$, $\beta\in\mathbb{R}$ and $r_1(t)=r_1\beta t$ whenever $\gamma=0$, $\beta\ge0$, the relation (2.46) leads to

$$\dot{\mathcal{E}}(t)\le r_3\left(g(x(t)+\beta(t)\dot{x}(t))-g^*\right)-r_1(t)\|\nabla g(x(t)+\beta(t)\dot{x}(t))\|^2+r_2t\epsilon(t)\|x^*\|^2-t\epsilon(t)\|x(t)-x^*\|^2,\tag{2.47}$$

for all $t\ge\bar{t}_2$.

Hence, from (2.42) and (2.47) we conclude that whenever condition (C1) or (C2) holds, there exists $t_2 > t_0$ and $s_1 > 0$, $s_2 > 0$ such that

$$\dot{\mathcal{E}}(t) \leq s_1 \left(g(x(t) + \beta(t)\dot{x}(t)) - g^*\right) + s_2 t\epsilon(t)\|x^*\|^2, \text{ for all } t \geq t_2. \tag{2.48}$$

Since $g(x(t)+\beta(t)\dot{x}(t)) - g^* \leq \frac{\mathcal{E}(t)}{a(t)}$ and taking into account that $a(t) = t^2 + \mathcal{O}(t)$ we deduce that there exists $M > 0$ and $t_3 \geq t_2$ such that $a(t) \geq \frac{t^2}{M}$, hence $g(x(t) + \beta(t)\dot{x}(t)) - g^* \leq \frac{M}{t^2}\mathcal{E}(t),$ for all $t \geq t_3$.

Consequently, (2.48) becomes

$$\dot{\mathcal{E}}(t) \leq \frac{s_1 M}{t^2}\mathcal{E}(t) + s_2 t\epsilon(t)\|x^*\|^2, \text{ for all } t \geq t_3. \tag{2.49}$$

Now, we apply Gronwall's lemma on an interval $[t_3, T]$, $T > t_3$ and we get

$$\mathcal{E}(t) \leq e^{A(t)}\mathcal{E}(t_3) + s_2\|x^*\|^2 e^{A(t)} \int_{t_3}^{t} \tau\epsilon(\tau)e^{-A(\tau)}d\tau,$$

where $A(t) = \int_{t_3}^t \frac{s_1 M}{\tau^2}d\tau = -\frac{s_1 M}{t} + \frac{s_1 M}{t_3}$. Obviously $e^{A(t)}$ is bounded on $[t_3, +\infty)$, hence there exists $C_1, C_2 > 0$ such that $e^{A(t)} \leq C_1$ and $e^{-A(t)} \leq C_2$ for all $t \in [t_3, +\infty)$.

We have

$$\mathcal{E}(t) \leq C_1\mathcal{E}(t_3) + C_1 C_2 s_2\|x^*\|^2 \int_{t_3}^{t} \tau\epsilon(\tau)d\tau, \tag{2.50}$$

for all $t \in [t_3, T)$.

Now, if (i) holds, then we have $\int_{t_3}^{+\infty} \frac{\epsilon(\tau)}{\tau}d\tau < +\infty$. Now, $\mathcal{E}(t) \geq \frac{t^2}{M}(g(x(t)+\beta(t)\dot{x}(t)) - \min g)$ and (2.50) leads to

$$g(x(t) + \beta(t)\dot{x}(t)) - \min g \leq \frac{C_1 M\mathcal{E}(t_3)}{t^2} + C_1 C_2 s_2 M\|x^*\|^2 \frac{1}{t^2}\int_{t_3}^{t} \tau^2\frac{\epsilon(\tau)}{\tau}d\tau,$$

for all $t \in [t_3, T)$.

According to Lemma B.1 from Appendix we have $\frac{1}{t^2}\int_{t_3}^t \tau^2 \frac{\epsilon(\tau)}{\tau}d\tau \to 0$ as $t \to +\infty$, hence

$$\lim_{t\to+\infty} g\left(x(t) + \left(\gamma + \frac{\beta}{t}\right)\dot{x}(t)\right) - \min g = 0.$$

Now, if (ii) holds, then $\int_{t_3}^{+\infty} \tau\epsilon(\tau)d\tau < +\infty$, hence the right hand side of (2.50) is bounded. In other words, there exists $C > 0$ such that

$$\mathcal{E}(t) \leq C, \text{ for all } t \geq t_3. \tag{2.51}$$

Hence, by the form of $\mathcal{E}(t)$ and the fact that $a(t) = \mathcal{O}(t^2)$, as $t \to +\infty$, we get that

$$g\left(x(t) + \left(\gamma + \frac{\beta}{t}\right)\dot{x}(t)\right) - \min g = \mathcal{O}\left(\frac{1}{t^2}\right), \text{ as } t \to +\infty.$$

Hence, combining the latter result with (2.42) and (2.47) we get that there exist $N_1, N_2 > 0$ and $t_4 \geq t_0$ such that

$$\dot{\mathcal{E}}(t) + r_1(t)\left\|\nabla g(x(t) + \beta(t)\dot{x}(t))\right\|^2 + t\epsilon(t)\|x(t) - x^*\|^2 \leq \frac{N_1}{t^2} + N_2 t\epsilon(t)\|x^*\|^2, \tag{2.52}$$

for all $t \geq t_4$. Integrating (2.52) on an interval $[t_4, T]$, $T \geq t_4$ and then letting $T \to +\infty$ we get

$$\int_{t_4}^{+\infty} r_1(t)\left\|\nabla g(x(t) + \beta(t)\dot{x}(t))\right\|^2 dt < +\infty \text{ and } \int_{t_4}^{+\infty} t\epsilon(t)\|x(t) - x^*\|^2 dt < +\infty.$$

□

3. Strong convergence results. Our first contribution of the present section is a result that assures the boundedness of the derivative of the unique global solution of the dynamical system (1.1).

Lemma 3.1. *Let x be the unique global solution of the dynamical system (1.1). Suppose that $\alpha > 0$ and further $\gamma > 0$ and $\beta \in \mathbb{R}$ or $\gamma = 0$ and $\beta \geq 0$. Then, the first order derivative of the solution is bounded, i.e., there exists $M > 0$ such that $\|\dot{x}(t)\| \leq M$, for all $t \geq t_0$. Further, $\frac{1}{t}\|\dot{x}(t)\|^2 \in L^1([t_0, +\infty), \mathbb{R})$.*

Proof . We consider the energy functional $W : [t_0, +\infty) \to \mathbb{R}$,

$$W(t) = g(x(t)) + \frac{1}{2}\|\dot{x}(t)\|^2 + \frac{\epsilon(t)}{2}\|x(t)\|^2. \tag{3.1}$$

The time derivative of (3.1) reads as

$$\dot{W}(t) = \langle \nabla g(x(t)), \dot{x}(t)\rangle + \langle \dot{x}(t), \ddot{x}(t)\rangle + \frac{\dot{\epsilon}(t)}{2}\|x(t)\|^2 + \epsilon(t)\langle x(t), \dot{x}(t)\rangle.$$

From (1.1) we have $\ddot{x}(t) = -\frac{\alpha}{t}\dot{x}(t) - \epsilon(t)x(t) - \nabla g(x(t) + \beta(t)\dot{x}(t))$, and we obtain

$$\dot{W}(t) = \langle \nabla g(x(t)) - \nabla g(x(t) + \beta(t)\dot{x}(t)), \dot{x}(t)\rangle + \frac{\dot{\epsilon}(t)}{2}\|x(t)\|^2 - \frac{\alpha}{t}\|\dot{x}(t)\|^2.$$

Now, if $\beta(t) = 0$ then obviously $\dot{W}(t) \leq -\frac{\alpha}{t}\|\dot{x}(t)\|^2$.

On the other hand, if $\beta(t) \neq 0$, by using the fact that ∇g is monotone and $\dot{\epsilon}(t) \leq 0$ we get

$$\begin{aligned}\dot{W}(t) &= -\frac{1}{\beta(t)}\langle \nabla g(x(t)) - \nabla g(x(t) + \beta(t)\dot{x}(t)), x(t) - (x(t) + \beta(t)\dot{x}(t))\rangle + \frac{\dot{\epsilon}(t)}{2}\|x(t)\|^2 - \frac{\alpha}{t}\|\dot{x}(t)\|^2 \\ &\leq -\frac{\alpha}{t}\|\dot{x}(t)\|^2.\end{aligned}$$

Consequently,

$$\dot{W}(t) + \frac{\alpha}{t}\|\dot{x}(t)\|^2 \leq 0 \text{ for all } t \geq t_0. \tag{3.2}$$

Therefore, W is non-increasing on $[t_0, \infty]$. Using that g is bounded from below, it follows that $\|\dot{x}(t)\| < +\infty$ for all $t \geq t_0$. Now, by integrating (3.2) on an interval $[t_0, t]$ we get $W(t) + \int_{t_0}^t \frac{\alpha}{\theta}\|\dot{x}(\theta)\|^2 d\theta \leq W(t_0)$ and implicitly $\frac{1}{t}\|\dot{x}(t)\|^2 \in L^1([t_0, +\infty), \mathbb{R})$. □

We continue the present section by emphasizing the main idea behind the Tikhonov regularization, which will generate strong convergence results for our dynamical system (1.1) to a minimizer of the objective function of minimal norm. By considering $\epsilon > 0$, by x_ϵ we denote the unique solution of the strongly convex minimization problem

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

We know that the Tikhonov approximation curve $\epsilon \to x_\epsilon$ satisfies $x^* = \lim_{\epsilon \to 0} x_\epsilon$, where x^* is the element of minimal norm from $\operatorname{argmin} g$. At the same time, for each $\epsilon > 0$, we have the inequality $\|x_\epsilon\| \leq \|x^*\|$ (see [20]), which will be used further. Now, in order to show the strong convergence of the dynamical system (1.1) to an element of minimum norm of the nonempty, convex and closed set $\operatorname{argmin} g$, we state our main result of the present section.

Theorem 3.2. *Let $\alpha \geq 3$, let x be the unique global solution of (1.1) and assume that $\int_{t_0}^{\infty} \frac{\epsilon(t)}{t} dt < +\infty$.*
(i) Assume that for $\alpha = 3$ one has $\gamma = 0$, $\beta \geq 0$, $\lim_{t \to \infty} t^2\epsilon(t) = +\infty$, $\lim_{t \to \infty} \frac{1}{\epsilon(t)t^2} \int_{t_0}^t \epsilon^2(s) s ds = 0$ and the condition (C1) *holds.*
(ii) In case $\alpha > 3$ assume that $t^2\epsilon(t) \geq \frac{2\alpha}{3}\left(\frac{\alpha}{3} - 1\right)$, for t large enough and $\lim_{t \to \infty} \frac{1}{\epsilon(t)t^{\frac{\alpha}{3}+1}} \int_{t_0}^t \epsilon^2(s) s^{\frac{\alpha}{3}+1} ds = 0$. Further, assume that $\gamma = 0$ or (C1) *or* (C2) *hold, where the constant K in condition* (C2) *satisfies $K\gamma < \frac{2(\alpha-3)}{3}$.*
Let $x^* = \operatorname{argmin}_{x \in \operatorname{argmin} g} \|x\|$. Then, it follows that

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

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

$$\lim_{t\to\infty} \|x(t) - x^*\| = 0.$$

Proof . For simplicity we denote $\beta(t) = \gamma + \frac{\beta}{t}$. In our forthcoming analysis we consider three different cases that are related to the relationship between the trajectories of the dynamical system (1.1) and the open ball $B(0, \|x^*\|)$.

Case I: We assume that there exists $T \geq t_0$, such that $\{x(t) : t \geq T\}$ stays in the complement of $B(0, \|x^*\|)$. This is equivalent to the fact that for each $t \geq T$, one has $\|x(t)\| \geq \|x^*\|$. We consider $p \geq 0$ and we define, for every $t \geq t_0$, the following energy functional

$$\mathcal{E}(t) = t^{p+2}\left(g(x(t) + \beta(t)\dot{x}(t)) - \min g\right) + t^{p+2}\frac{\epsilon(t)}{2}\left(\|x(t)\|^2 - \|x^*\|^2\right) + \frac{t^p}{2}\|b(x(t) - x^*) + t\dot{x}(t)\|^2. \tag{3.3}$$

Obviously, for each $t \geq t_0$ we obtain that

$$\mathcal{E}(t) \geq \frac{t^{p+2}}{2}(g(x(t) + \beta(t)\dot{x}(t)) - \min g) + t^{p+2}\frac{\epsilon(t)}{2}(\|x(t)\|^2 - \|x^*\|^2). \tag{3.4}$$

Now, we define the strongly convex function $g_t : \mathcal{H} \to \mathbb{R}$, $g_t(x) = \frac{1}{2}g(x) + \frac{\epsilon(t)}{2}\|x\|^2$ and we denote $x_{\epsilon(t)} = \operatorname{argmin}_{x\in\mathcal{H}} g_t(x)$. Using the same argument as in the proof of Theorem 4.4 [20], we have that

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

Hence,

$$g_t(x(t) + \beta(t)\dot{x}(t)) - g_t(x^*) \geq \frac{\epsilon(t)}{2}(\|x(t) + \beta(t)\dot{x}(t) - x_{\epsilon(t)}\|^2 + \|x_{\epsilon(t)}\|^2 - \|x^*\|^2).$$

Now, by employing (3.4), we obtain that

$$\mathcal{E}(t) \geq t^{p+2}\frac{\epsilon(t)}{2}\left(\|x_{\epsilon(t)}\|^2 - \|x^*\|^2 + \|x(t) + \beta(t)\dot{x}(t) - x_{\epsilon(t)}\|^2 + \|x(t)\|^2 - \|x(t) + \beta(t)\dot{x}(t)\|^2\right).$$

We have

$$\|x(t) + \beta(t)\dot{x}(t)\|^2 = \|x(t)\|^2 + \beta^2(t)\|\dot{x}(t)\|^2 + 2\beta(t)\langle x(t), \dot{x}(t)\rangle$$

and

$$\|x(t) + \beta(t)\dot{x}(t) - x_{\epsilon(t)}\|^2 = \|x(t) - x_{\epsilon(t)}\|^2 + \beta^2(t)\|\dot{x}(t)\|^2 + 2\beta(t)\langle \dot{x}(t), x(t) - x_{\epsilon(t)}\rangle,$$

hence, for all $t \geq t_0$ we get

$$\mathcal{E}(t) \geq t^{p+2}\frac{\epsilon(t)}{2}(\|x_{\epsilon(t)}\|^2 - \|x^*\|^2 + \|x(t) - x_{\epsilon(t)}\|^2) - t^{p+2}\beta(t)\epsilon(t)\langle \dot{x}(t), x_{\epsilon(t)}\rangle. \tag{3.5}$$

Now, the next step is to get an upper bound for $\mathcal{E}(\cdot)$. In order to do this, for each $t \geq t_0$, we consider the time derivative of the energy function as follows:

$$\begin{aligned} \dot{\mathcal{E}}(t) &= (p+2)t^{p+1}(g(x(t) + \beta(t)\dot{x}(t)) - \min g) + t^{p+2}\langle \nabla g(x(t) + \beta(t)\dot{x}(t)), (1 + \dot{\beta}(t))\dot{x}(t) + \beta(t)\ddot{x}(t)\rangle \\ &+ \left(t^{p+2}\frac{\dot{\epsilon}(t)}{2} + (p+2)t^{p+1}\frac{\epsilon(t)}{2}\right)\cdot(\|x(t)\|^2 - \|x^*\|^2) + t^{p+2}\epsilon(t)\langle x(t), \dot{x}(t)\rangle \\ &+ p\frac{t^{p-1}}{2}\|b(x(t) - x^*) + t\dot{x}(t)\|^2 + t^p\langle b(x(t) - x^*) + t\dot{x}(t), (1+b)\dot{x}(t) + t\ddot{x}(t)\rangle. \end{aligned} \tag{3.6}$$

On the other hand, from the dynamical system (1.1) we have $t\ddot{x}(t) = -\alpha\dot{x}(t) - t\epsilon(t)x(t) - t\nabla g(x(t) + \beta(t)\dot{x}(t))$, hence for all $t \geq t_0$ one has

$$\begin{aligned} t^p\langle b(x(t) - x^*) + t\dot{x}(t), (1+b)\dot{x}(t) + t\ddot{x}(t)\rangle &= (1 + b - \alpha)bt^p\langle x(t) - x^*, \dot{x}(t)\rangle - b\epsilon(t)t^{p+1}\langle x(t) - x^*, x(t)\rangle \\ &- bt^{p+1}\langle x(t) - x^*, \nabla g(x(t) + \beta(t)\dot{x}(t))\rangle \\ &+ (1 + b - \alpha)t^{p+1}\|\dot{x}(t)\|^2 - \epsilon(t)t^{p+2}\langle \dot{x}(t), x(t)\rangle \\ &- t^{p+2}\langle \dot{x}(t), \nabla g(x(t) + \beta(t)\dot{x}(t))\rangle \end{aligned} \tag{3.7}$$

and

(3.8)
$$t^{p+2}\langle \nabla g(x(t)+\beta(t)\dot{x}(t)),(1+\dot{\beta}(t))\dot{x}(t)+\beta(t)\ddot{x}(t)\rangle = \left(1+\dot{\beta}(t)-\frac{\alpha}{t}\beta(t)\right)t^{p+2}\langle \nabla g(x(t)+\beta(t)\dot{x}(t)),\dot{x}(t)\rangle$$
$$-\epsilon(t)\beta(t)t^{p+2}\langle \nabla g(x(t)+\beta(t)\dot{x}(t)),x(t)\rangle$$
$$-\beta(t)t^{p+2}\|\nabla g(x(t)+\beta(t)\dot{x}(t))\|^2.$$

Further,

(3.9)
$$p\frac{t^{p-1}}{2}\|b(x(t)-x^*)+t\dot{x}(t)\|^2 = pb^2\frac{t^{p-1}}{2}\|x(t)-x^*\|^2+p\frac{t^{p+1}}{2}\|\dot{x}(t)\|^2+pbt^p\langle x(t)-x^*,\dot{x}(t)\rangle \text{ for all } t\geq t_0.$$

By combining (3.6), (3.7), (3.8) and (3.9), it follows that

$$\text{(3.10)}\quad \dot{\mathcal{E}}(t) = (p+2)t^{p+1}(g(x(t)+\beta(t)\dot{x}(t))-\min g)+\left(t^{p+2}\frac{\dot{\epsilon}(t)}{2}+(p+2)t^{p+1}\frac{\epsilon(t)}{2}\right)(\|x(t)\|^2-\|x^*\|^2)$$
$$-\beta(t)t^{p+2}\|\nabla g(x(t)+\beta(t)\dot{x}(t))\|^2-\epsilon(t)\beta(t)t^{p+2}\langle \nabla g(x(t)+\beta(t)\dot{x}(t)),x(t)\rangle$$
$$+\left(\dot{\beta}(t)-\frac{\alpha}{t}\beta(t)\right)t^{p+2}\langle \nabla g(x(t)+\beta(t)\dot{x}(t)),\dot{x}(t)\rangle+\left(1+b-\alpha+\frac{p}{2}\right)t^{p+1}\|\dot{x}(t)\|^2$$
$$+pb^2\frac{t^{p-1}}{2}\|x(t)-x^*\|^2+b\,(1+b-\alpha+p)\,t^p\langle x(t)-x^*,\dot{x}(t)\rangle$$
$$-bt^{p+1}\langle x(t)-x^*,\nabla g(x(t)+\beta(t)\dot{x}(t))+\epsilon(t)x(t)\rangle \text{ for all } t\geq t_0.$$

By using (2.7), we obtain that

(3.11)
$$-bt^{p+1}\langle \nabla g(x(t)+\beta(t)\dot{x}(t))+\epsilon(t)x(t),x(t)-x^*\rangle \leq -bt^{p+1}(g(x(t)+\beta(t)\dot{x}(t))-\min g)$$
$$+\beta(t)bt^{p+1}\langle \nabla g(x(t)+\beta(t)\dot{x}(t)),\dot{x}(t)\rangle+bt^{p+1}\frac{\epsilon(t)}{2}\|x^*\|^2-bt^{p+1}\frac{\epsilon(t)}{2}\|x(t)\|^2-bt^{p+1}\frac{\epsilon(t)}{2}\|x(t)-x^*\|^2,$$

for all $t\geq t_0$.

From (3.10) and (3.11), it follows that

(3.12)
$$\dot{\mathcal{E}}(t)\leq (p+2-b)t^{p+1}(g(x(t)+\beta(t)\dot{x}(t))-\min g)+\left(t\frac{\dot{\epsilon}(t)}{2}+(p+2-b)\frac{\epsilon(t)}{2}\right)t^{p+1}(\|x(t)\|^2-\|x^*\|^2)$$
$$-\beta(t)t^{p+2}\|\nabla g(x(t)+\beta(t)\dot{x}(t))\|^2-\epsilon(t)\beta(t)t^{p+2}\langle \nabla g(x(t)+\beta(t)\dot{x}(t)),x(t)\rangle$$
$$+\left(\dot{\beta}(t)+\frac{b-\alpha}{t}\beta(t)\right)t^{p+2}\langle \nabla g(x(t)+\beta(t)\dot{x}(t)),\dot{x}(t)\rangle+\left(1+b-\alpha+\frac{p}{2}\right)t^{p+1}\|\dot{x}(t)\|^2$$
$$+\left(pb^2\frac{t^{p-1}}{2}-b\frac{\epsilon(t)}{2}t^{p+1}\right)\|x(t)-x^*\|^2+b\,(1+b-\alpha+p)\,t^p\langle x(t)-x^*,\dot{x}(t)\rangle,\ \text{for all } t\geq t_0.$$

From now on, we choose $b:=\frac{2\alpha}{3}$ and $p:=\frac{\alpha-3}{3}$. Then, since $\alpha\geq 3$, we obtain that $p+2-b=\frac{3-\alpha}{3}\leq 0$, further $1+b-\alpha+\frac{p}{2}=\frac{3-\alpha}{6}\leq 0$ and $1+b-\alpha+p=0$.

For every $r,s>0$ and for each $t\geq t_0$ we obviously have that

(3.13)
$$-\beta(t)\epsilon(t)t^{p+2}\langle \nabla g(x(t)+\beta(t)\dot{x}(t)),x(t)\rangle \leq \frac{|\beta(t)|}{r}t^{p+2}\|\nabla g(x(t)+\beta(t)\dot{x}(t))\|^2+\frac{r|\beta(t)|\epsilon^2(t)}{4}t^{p+2}\|x(t)\|^2$$

and

(3.14)
$$\left(\dot{\beta}(t)+\frac{b-\alpha}{t}\beta(t)\right)t^{p+2}\langle \nabla g(x(t)+\beta(t)\dot{x}(t)),\dot{x}(t)\rangle \leq \frac{|t\dot{\beta}(t)+(b-\alpha)\beta(t)|}{s}t^{p+2}\|\nabla g(x(t)+\beta(t)\dot{x}(t))\|^2$$
$$+\frac{|t\dot{\beta}(t)+(b-\alpha)\beta(t)|s}{4}t^p\|\dot{x}(t))\|^2.$$

Injecting (3.13) and (3.14) in (3.12) we obtain that for each $t \geq t_0$ one has

$$\begin{aligned}
\dot{\mathcal{E}}(t) \leq\ & \frac{3-\alpha}{3} t^{p+1}(g(x(t)+\beta(t)\dot{x}(t)) - \min g) + \frac{r|\beta(t)|\epsilon^2(t)t^{p+2}}{4}\|x^*\|^2 \\
& + \left(t\frac{\dot{\epsilon}(t)}{2} + \frac{3-\alpha}{3}\frac{\epsilon(t)}{2} + \frac{r|\beta(t)|\epsilon^2(t)t}{4}\right) t^{p+1}(\|x(t)\|^2 - \|x^*\|^2) \\
& + \left(\frac{|\beta(t)|}{r} - \beta(t) + \frac{\left|t\dot{\beta}(t) - \frac{\alpha}{3}\beta(t)\right|}{s}\right) t^{p+2}\|\nabla g(x(t)+\beta(t)\dot{x}(t))\|^2 \\
& + \left(\frac{3-\alpha}{6} + \frac{\left|t\dot{\beta}(t) - \frac{\alpha}{3}\beta(t)\right| s}{4t}\right) t^{p+1}\|\dot{x}(t)\|^2 \\
& + \frac{\alpha}{3}\left(\frac{2\alpha(\alpha-3)}{9} t^{p-1} - \epsilon(t)t^{p+1}\right)\|x(t)-x^*\|^2.
\end{aligned} \tag{3.15}$$

Now, if (C1) holds we take $1 < r < K$, (K is defined at the condition (C1)), and we obtain that

$$\left(t\frac{\dot{\epsilon}(t)}{2} + \frac{r|\beta(t)|\epsilon^2(t)t}{4}\right) \leq 0, \text{ for all } t \geq t_1.$$

Assume now that $\alpha > 3$. If $\gamma = 0$, then by using the fact that ϵ is non-increasing, we get that for every $r > 0$ there exists $\bar{t}_1 \geq t_0$ such that

$$\frac{r|\beta(t)|\epsilon^2(t)t}{4} \leq \frac{\alpha-3}{3}\frac{\epsilon(t)}{2}, \text{ for all } t \geq \bar{t}_1.$$

Further, if (C2) holds, we have that there exists $K > 0$ and $t_1 \geq t_0$ such that $\epsilon(t) \leq \frac{K}{t}$ for all $t \geq t_1$, where according to the hypotheses one has $K\gamma < \frac{2(\alpha-3)}{3}$. Consequently, there exists $\gamma_1 > \gamma$ such that $K\gamma < K\gamma_1 < \frac{2(\alpha-3)}{3}$. Hence, for all $t \geq t_1$ one has

$$\frac{r|\beta(t)|\epsilon^2(t)t}{4} \leq \frac{r|\beta(t)|\epsilon(t)K}{4} < \frac{r|\beta(t)|}{\gamma_1}\frac{\alpha-3}{3}\frac{\epsilon(t)}{2}.$$

The latter relation leads to the existence of $\bar{t}_1 \geq t_1$ and $1 < r < \frac{\gamma_1}{\gamma}$, (if $\gamma = 0$ we take $\frac{\gamma_1}{\gamma} = +\infty$), such that

$$\frac{3-\alpha}{3}\frac{\epsilon(t)}{2} + \frac{r|\beta(t)|\epsilon^2(t)t}{4} \leq \left(\frac{r|\beta(t)|}{\gamma_1} - 1\right)\frac{\alpha-3}{3}\frac{\epsilon(t)}{2} \leq 0, \text{ for all } t \geq \bar{t}_1.$$

Hence, due to the assumption that $\|x(t)\| \geq \|x^*\|$ for $t \geq T$, we conclude that under the hypotheses of the theorem, there exist $r > 1$ and $\bar{t}_2 \geq T$, such that

$$\left(t\frac{\dot{\epsilon}(t)}{2} + \frac{3-\alpha}{3}\frac{\epsilon(t)}{2} + \frac{r|\beta(t)|\epsilon^2(t)t}{4}\right) t^{p+1}(\|x(t)\|^2 - \|x^*\|^2) \leq 0, \text{ for all } t \geq \bar{t}_2. \tag{3.16}$$

Further, if we take $r > 1$ we conclude that there exist $s > 0$ and $t_2 \geq t_0$ such that

$$\frac{|\beta(t)|}{r} - \beta(t) + \frac{\left|t\dot{\beta}(t) - \frac{\alpha}{3}\beta(t)\right|}{s} \leq 0, \text{ for all } t \geq t_2. \tag{3.17}$$

Finally, due to the hypotheses of the theorem, for t big enough

$$\frac{2\alpha(\alpha-3)}{9} t^{p-1} - \epsilon(t)t^{p+1} \leq 0. \tag{3.18}$$

Hence, there exists t_3 big enough such that (3.16), (3.17), (3.18)and (3.15) yield

$$\dot{\mathcal{E}}(t) \leq \frac{r|\beta(t)|\epsilon^2(t)t^{p+2}}{4}\|x^*\|^2 + \left(\frac{3-\alpha}{6} + \frac{\left|t\dot{\beta}(t) - \frac{\alpha}{3}\beta(t)\right| s}{4t}\right) t^{p+1}\|\dot{x}(t)\|^2, \text{ for all } t \geq t_3. \tag{3.19}$$

Now if $\alpha = 3$ then by assumption $\gamma = 0$, $\beta \geq 0$ and $p = 0$. Hence, (3.19) becomes

$$\dot{\mathcal{E}}(t) \leq \frac{r\beta\epsilon^2(t)t}{4}\|x^*\|^2 + \frac{\beta s}{2t}\|\dot{x}(t)\|^2, \text{ for all } t \geq t_3. \tag{3.20}$$

By integrating (3.20) on $[t_3, t]$ for an arbitrary $t \geq t_3$, we obtain that

$$\mathcal{E}(t) \leq \mathcal{E}(t_3) + \frac{r\beta}{4}\|x^*\|^2\int_{t_3}^t \epsilon^2(\theta)\theta d\theta + \frac{\beta s}{2}\int_{t_3}^t \frac{1}{\theta}\|\dot{x}(\theta)\|^2 d\theta. \tag{3.21}$$

From (3.5), we have that

$$\mathcal{E}(t) \geq t^2\frac{\epsilon(t)}{2}(\|x_{\epsilon(t)}\|^2 - \|x^*\|^2 + \|x(t) - x_{\epsilon(t)}\|^2) - \beta t\epsilon(t)\langle\dot{x}(t), x_{\epsilon(t)}\rangle,$$

which combined with (3.21) gives

$$\begin{aligned}\|x(t) - x_{\epsilon(t)}\|^2 \leq \|x^*\|^2 - \|x_{\epsilon(t)}\|^2 + \frac{2\mathcal{E}(t_3)}{\epsilon(t)t^2} + 2\frac{\beta}{t}\langle\dot{x}(t), x_{\epsilon(t)}\rangle + \frac{r\beta}{2}\|x^*\|^2\frac{1}{\epsilon(t)t^2}\int_{t_3}^t \epsilon^2(\theta)\theta d\theta \\ + \frac{\beta s}{\epsilon(t)t^2}\int_{t_3}^t \frac{1}{\theta}\|\dot{x}(\theta)\|^2 d\theta.\end{aligned} \tag{3.22}$$

According to Lemma 3.1, $\dot{x}(t)$ is bounded. Further $x_{\epsilon(t)} \to x^*$, $t \to +\infty$, hence $\lim_{t\to+\infty} 2\frac{\beta}{t}\langle\dot{x}(t), x_{\epsilon(t)}\rangle = 0$. According to the hypotheses of the theorem one has $\lim_{t\to+\infty} \frac{1}{\epsilon(t)t^2}\int_{t_3}^t \epsilon^2(\theta)\theta d\theta = 0$. Finally, from Lemma 3.1 we have $\frac{1}{\theta}\|\dot{x}(\theta)\|^2 \in L^1([t_0, +\infty), \mathbb{R})$, hence $\lim_{t\to+\infty} \frac{\beta s}{\epsilon(t)t^2}\int_{t_3}^t \frac{1}{\theta}\|\dot{x}(\theta)\|^2 d\theta = 0$.

Consequently, the right hand side of (3.22) goes to 0 as $t \to +\infty$, hence $\lim_{t\to+\infty}\|x(t) - x_{\epsilon(t)}\| = 0$, that is,

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

Now if $\alpha > 3$ then obviously there exists $t_4 \geq t_3$ such that

$$\frac{3-\alpha}{6} + \frac{\left|t\dot{\beta}(t) - \frac{\alpha}{3}\beta(t)\right| s}{4t} \leq 0, \text{ for all } t \geq t_4.$$

Hence, (3.19) leads to

$$\dot{\mathcal{E}}(t) \leq \frac{r|\beta(t)|\epsilon^2(t)}{4}t^{p+2}\|x^*\|^2, \text{ for all } t \geq t_4. \tag{3.23}$$

By integrating (3.23) on $[t_4, t]$ for an arbitrary $t \geq t_4$, and taking into account that $\beta(t) = \gamma + \frac{\beta}{t}$ is bounded and $p = \frac{\alpha-3}{3}$, we obtain that there exists an $R > 0$ such that

$$\mathcal{E}(t) \leq \mathcal{E}(t_4) + R\|x^*\|^2\int_{t_4}^t \epsilon^2(s)s^{\frac{\alpha}{3}+1}ds. \tag{3.24}$$

From (3.5), we have that

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

hence we obtain that

$$\|x(t) - x_{\epsilon(t)}\|^2 \leq \|x^*\|^2 - \|x_{\epsilon(t)}\|^2 + \frac{2\mathcal{E}(t_4)}{\epsilon(t)t^{\frac{\alpha}{3}+1}} + 2\beta(t)\langle\dot{x}(t), x_{\epsilon(t)}\rangle + \frac{2R}{\epsilon(t)t^{\frac{\alpha}{3}+1}}\|x^*\|^2\int_{t_4}^t \epsilon^2(s)s^{\frac{\alpha}{3}+1}ds. \tag{3.25}$$

From Theorem 2.1, we have that $\dot{x}(t) \to 0$ as $t \to \infty$. Since $x_{\epsilon(t)} \to x^*$we get that $2\beta(t)\langle \dot{x}(t), x_{\epsilon(t)}\rangle \to 0$ as $t \to \infty$. Now by the hypotheses of the theorem we have that

$$\epsilon(t)t^{\frac{\alpha}{3}+1} \to +\infty \text{ and } \frac{q}{\epsilon(t)t^{\frac{\alpha}{3}+1}} \int_{t_4}^{t} \epsilon^2(s)s^{\frac{\alpha}{3}+1}ds \to 0 \text{ as } t \to +\infty$$

hence, we get that the limit of the right hand side in (3.25) goes to 0 as $t \to +\infty$.

Consequently, $\|x(t) - x_{\epsilon(t)}\| \to 0$, $t \to +\infty$, that is,

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

Now, we analyze the second case as follows.

Case II: Assume that there exists $T \geq t_0$ such that the trajectory $\{x(t) : t \geq T\}$ stays in the open ball $B(0, \|x^*\|)$ Equivalently, we have that $\|x(t)\| < \|x^*\|$ for every $t \geq T$. By the fact that

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

and with respect to Theorem 2.3 and Theorem 2.1, we have that

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

Now, we take $\bar{x} \in \mathcal{H}$ a weak sequential cluster point of the trajectory x, which exists since the trajectory is bounded. This means that there exists a sequence $(t_n)_{n\in\mathbb{N}} \subseteq [T, +\infty)$ such that $t_n \to +\infty$ and $x(t_n)$ converges weakly to $\bar{x}$ as $n \to +\infty$. We know that g is weakly lower semicontinuous, so one has

$$g(\bar{x}) \leq \liminf_{n\to+\infty} g(x(t_n)) = \min g\,,$$

hence $\bar{x} \in \operatorname{argmin} g$. Now, since the norm is weakly lower semicontinuous one has that

$$\|\bar{x}\| \leq \liminf\nolimits_{n\to+\infty} \|x(t_n)\| \leq \|x^*\|\,,$$

which, from the definition of x^*, implies that $\bar{x} = x^*$. This shows that the trajectory $x(\cdot)$ converges weakly to x^*. So

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

hence we have

$$\lim_{t\to+\infty} \|x(t)\| = \|x^*\|\,.$$

From the previous relation and the fact that $x(t) \rightharpoonup x^*$ as $t \to +\infty$, we obtain the strong convergence, i.e.

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

Finally, the last case reads as follows.

Case III: We suppose that for every $T \geq t_0$ there exists $t \geq T$ such that $\|x^*\| > \|x(t)\|$ and also there exists $s \geq T$ such that $\|x^*\| \leq \|x(s)\|$. From the continuity of the unique strong global solution $x(\cdot)$, we find that there exists a sequence $(t_n)_{n\in\mathbb{N}} \subseteq [t_0, +\infty)$ such that $t_n \to +\infty$ as $n \to +\infty$ and, for all $n \in \mathbb{N}$ we have

$$\|x(t_n)\| = \|x^*\|\,.$$

In order to show that $x(t_n) \to x^*$ as $n \to +\infty$, we let $\bar{x} \in \mathcal{H}$ to be a weak sequential cluster point of $(x(t_n))_{n\in\mathbb{N}}$. By using that the sequence is bounded and by employing arguments similar to the previous case, we eventually find that $(x(t_n))_{n\in\mathbb{N}}$ converges weakly to x^* as $n \to +\infty$. Obviously $\|x(t_n)\| \to \|x^*\|$ as $n \to +\infty$. So, it follows that $\|x(t_n) - x^*\| \to 0$ as $n \to +\infty$. This leads to

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

and the proof is over. □

Remark 4. *Observe that according to (3.23), in case $\alpha > 3$, $\gamma = 0$ it is enough to assume that*

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

4. Conclusion, perspectives. The dynamical system (1.1) studied in this paper can be seen as a second order system with implicit Hessian driven damping, therefore it is in strong connection with the dynamical system studied in [20]. At the same time, (1.1) is a perturbed version of the dynamical system with asymptotically vanishing damping considered in [10]. We have shown that (1.1) possess all the valuable properties of the two related systems, as we obtained fast convergence of velocities to zero, integral estimates for the gradient of the objective function and fast convergence of the objective function values to the minimum of the objective function. Further, depending the Tikhonov regularization parameter $\epsilon(t)$ goes fast or slow to zero, we obtained weak convergence of the trajectories to a minimizer of the objective function and strong convergence of the trajectories to a minimizer of minimum norm, respectively. Even more, by using the techniques from [16], we were able to obtain both strong convergence of the trajectories to a minimizer of minimum norm and fast convergence of the function values for the same dynamics.

The article presents the basic analysis of the dynamical system (1.1), many aspects of which have yet to be developed. We ought to enlarge the framework by considering optimization problems with non-smooth convex objective function and in the corresponding dynamical system, with or without Tikhonov regularization term, to replace the function by its Moreau envelope, (see [15]). Further, we intend to study the inertial algorithms obtained from (1.1) via explicit discretization and by taking advantage by the fact that these algorithms may have different inertial terms (see [3]), to obtain convergence of the generated sequences to a minimizer of the objective function. The Tikhonov regularization of these algorithms may allow to obtain strong convergence of the generated sequences to a minimizer of minimal norm of the objective function, (see [16]). Some recent results show that for non-convex objective, considering different inertial terms in the corresponding algorithms, bring some improvements (see [28]). It would be interesting to show that similar results hold also in convex case. However, if we discretize the dynamical system (1.1) by making use the Taylor expansion of the gradient we obtain inertial algorithms similar to the algorithms considered in [8] and [2]. Further, in case the objective function is non-smooth, the above described discretization techniques lead to algorithms which are related to the celebrated algorithms (RIPA) [18] and (PRINAM) [14].

Acknowledgements The authors are thankful to three anonymous reviewers for their comments and suggestions which improved the quality of the paper.

Appendix A. Existence and uniqueness of the trajectories generated by the dynamical system (1.1). In what follows we show the existence and uniqueness of a classical C^2 solution x of the dynamical system (1.1). To this purpose we rewrite (1.1) as a first order system relevant for the Cauchy-Lipschitz-Picard theorem.

Theorem A.1. *Let $(u_0, v_0) \in \mathcal{H} \times \mathcal{H}$. Then, the dynamical system (1.1) admits a unique global $C^2((t_0, +\infty), \mathcal{H})$ solution.*

Proof . Indeed, by using the notation $X(t) := (x(t), \dot{x}(t))$, the dynamical system (1.1) can be put in the form

$$\begin{cases} \dot{X}(t) = F(t, X(t)) \\ X(t_0) = (u_0, v_0), \end{cases} \tag{A.1}$$

where $F : [t_0, \infty) \times \mathcal{H} \times \mathcal{H} \longrightarrow \mathcal{H} \times \mathcal{H}$, $F(t, u, v) = \left(v, -\frac{\alpha}{t}v - \epsilon(t)u - \nabla g\left(u + \left(\gamma + \frac{\beta}{t}\right)v\right)\right)$.

Our proof is inspired from [12]. Since ∇g is Lipschitz on bounded sets, it is obvious that for (A.1) the classical Cauchy-Picard theorem can be applied, hence, there exist a unique C^1 local solution X. Consequently, (1.1) has a unique C^2 local solution. Let x be a maximal solution of (1.1), defined on an interval $[t_0, T_{\max})$, $T_{\max} \leq +\infty$. In order to prove that $\dot{x}$ is bounded on $[t_0, T_{\max})$ one can use the same arguments as in the proof of Lemma 3.1.

Let $\|\dot{x}_\infty\| = \sup_{t \in [t_0, T_{\max})} \|\dot{x}(t)\|$ and assume that $T_{\max} < +\infty$. Since $\|x(t) - x(t')\| \leq \|\dot{x}_\infty\| |t - t'|$, we get that $\lim_{t \to T_{\max}} x(t) := x_\infty \in \mathcal{H}$. By (1.1) the map $\ddot{x}$ is also bounded on the interval $[t_0, T_{\max})$ and under the same argument as before $\lim_{t \to T_{\max}} \dot{x}(t) := x_\infty$ exists. Applying the local existence theorem with initial data $(x_\infty, \dot{x}_\infty)$, we can extend the maximal solution to a strictly larger interval, a clear contradiction. Hence $T_{\max} = +\infty$, which completes the proof. □

Appendix B. Auxiliary results. In this appendix, we collect some lemmas and technical results which we will use in the analysis of the dynamical system (1.1). The following lemma was stated for instance in [10, Lemma A.3] and is used to prove the convergence of the objective function along the trajectory to its minimal value.

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

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

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

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

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

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

The following technical result is [19, Lemma 2].

Lemma B.3. *Let $u : [t_0, +\infty) \longrightarrow \mathcal{H}$ be a continuously differentiable function satisfying $u(t) + \frac{t}{\alpha}\dot{u}(t) \to u \in \mathcal{H}$ as $t \to +\infty$, where $\alpha > 0$. Then $u(t) \to u$ as $t \to +\infty$.*

The continuous version of the Opial Lemma (see [9]) is the main tool for proving weak convergence for the generated trajectory.

Lemma B.4. *Let $S \subseteq \mathcal{H}$ be a nonempty set and $x : [t_0, +\infty) \longrightarrow H$ a given map such that:*

(i) for every $z \in S$ the limit $\lim_{t\to+\infty} \|x(t) - z\|$ exists;

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

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

Lemma B.5. *(Lemma A.6 [18]) Let $t_0 > 0$ and let $w : [t_0, +\infty) \longrightarrow \mathbb{R}$ be a continuously differentiable function which is bounded from below. Given a nonnegative function θ, let us assume that*

$$t\ddot{w}(t) + \alpha\dot{w}(t) + \theta(t) \le k(t),$$

for some $\alpha > 1$, almost every $t > t_0$, and some nonnegative function $k \in L^1((t_0, +\infty), \mathbb{R})$.

Then, the positive part $[\dot{w}]_+$ of $\dot{w}$ belongs to $L^1((t_0, +\infty), \mathbb{R})$ and $\lim_{t\to+\infty} w(t)$ exists. Moreover, we have

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

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 its Applications, 161(2) (2014), 331-360

[2] C.D. Alecsa, *The long time behavior and the rate of convergence of symplectic convex algorithms obtained via splitting discretizations of inertial damping systems*, (2020), arxiv.org/abs/2001.10831

[3] C.D. Alecsa, S.C. László, T. Pinţa, *An Extension of the Second Order Dynamical System that Models Nesterov's Convex Gradient Method*, Applied Mathematics and Optimization, (2020), https://doi.org/10.1007/s00245-020-09692-1

[4] C.D. Alecsa, S.C. László, A. Viorel, *A gradient type algorithm with backward inertial steps associated to a nonconvex minimization problem*, Numerical Algorithms, 84(2) (2020), 485-512

[5] V. Apidopoulos, J.F. Aujol, Ch. Dossal, *Convergence rate of inertial Forward-Backward algorithm beyond Nesterov's rule*, Mathematical Programming, 180 (2020), 137-156

[6] H. Attouch, A. Balhag, Z. Chbani, H. Riahi, *Fast convex optimization via inertial dynamics combining viscous and Hessian-driven damping with time rescaling*, Evolution Equations and Control Theory, (2021), https://doi.org/10.3934/eect.202101

[7] H. Attouch, A. Cabot, *Convergence rates of inertial forward-backward algorithms*, SIAM J. Optim., 28(1) (2018), 849-874

[8] H. Attouch, Z. Chbani, J. Fadili, H. Riahi, *First-order algorithms via inertial systems with Hessian driven damping*, Math. Program., (2020), https://doi.org/10.1007/s10107-020-01591-1
[9] H. Attouch, Z. Chbani, J. Peypouquet, P. Redont, *Fast convergence of inertial dynamics and algorithms with asymptotic vanishing viscosity*, Math. Program. Ser. B, 168 (2018), 123-175
[10] H. Attouch, Z. Chbani, H. Riahi, *Combining fast inertial dynamics for convex optimization with Tikhonov regularization*, Journal of Mathematical Analysis and Applications, 457(2) (2018), 1065-1094
[11] H. Attouch, Z. Chbani, R. Riahi, *Rate of convergence of the Nesterov accelerated gradient method in the subcritical case* $\alpha \leq 3$. ESAIM-COCV, 25 (2019), Article number 2
[12] H. Attouch, M.-O. Czarnecki, *Asymptotic Control and Stabilization of Nonlinear Oscillators with Non-isolated Equilibria,* J. Differential Equations, 179 (2002), 278-310
[13] H. Attouch, R. Cominetti, *A dynamical approach to convex minimization coupling approximation with the steepest descent method*, J. Differential Equations, 128 (1996), 519-540
[14] H. Attouch, S.C. László, *Newton-like Inertial Dynamics and Proximal Algorithms Governed by Maximally Monotone Operators*, SIAM Journal on Optimization, 30(4) (2020), 3252–3283
[15] H. Attouch, S.C. László, *Continuous Newton-like Inertial Dynamics for Monotone Inclusions*, Set-Valued and Variational Analysis, (2020), doi:10.1007/s11228-020-00564-y
[16] H. Attouch, S.C. László, *Convex optimization via inertial algorithms with vanishing Tikhonov regularization: fast convergence to the minimum norm solution*, (2021), https://arxiv.org/abs/2104.11987
[17] H. Attouch, J. Peypouquet, *The rate of convergence of Nesterov's accelerated forward-backward method is actually faster than* $1/k^2$. SIAM J. Optim., 26(3) (2016), 1824-1834
[18] H. Attouch, J. Peypouquet, *Convergence of inertial dynamics and proximal algorithms governed by maximal monotone operators*, Mathematical Programming, 174(1-2) (2019), 391–432
[19] H. Attouch, J. Peypouquet, P. Redont, *Fast convex optimization via inertial dynamics with Hessian driven damping*, Journal of Differential Equations, 261(10) (2016), 5734-5783
[20] R.I. Boţ, E.R. Csetnek, S.C. László, *Tikhonov regularization of a second order dynamical system with Hessian damping*, Math. Program., (2020), https://doi.org/10.1007/s10107-020-01528-8
[21] R.I. Boţ, E.R. Csetnek, S.C. László, *A second-order dynamical approach with variable damping to nonconvex smooth minimization* Applicable Analysis, 99(3) (2020), 361-378
[22] R.I. Boţ, S.M. Grad, D. Meier, M. Staudigl, *Inducing strong convergence of trajectories in dynamical systems associated to monotone inclusions with composite structure*, Adv. Nonlinear Anal., 10 (2021), 450–476
[23] A. Cabot, H. Engler, S. Gadat, *On the long time behavior of second order differential equations with asymptotically small dissipation*, Transactions of the American Mathematical Society, 361 (2009), 5983-6017
[24] A. Chambolle, Ch. Dossal, *On the convergence of the iterates of the Fast Iterative Shrinkage Thresholding Algorithm*, Journal of Optimization Theory and Applications, 166 (2015), 968-982
[25] R. Cominetti, J. Peypouquet, S. Sorin, *Strong asymptotic convergence of evolution equations governed by maximal monotone operators with Tikhonov regularization*, J. Differential Equations, 245 (2008), 3753-3763
[26] M.A. Jendoubi, R. May, *Asymptotics for a second-order differential equation with nonautonomous damping and an integrable source term*, Appl. Anal., 94(2) (2015), 435–443
[27] S.C. László, *Convergence rates for an inertial algorithm of gradient type associated to a smooth nonconvex minimization*, Mathematical Programming, (2020), https://doi.org/doi.org/10.1007/s10107-020-01534-w
[28] S.C. László, *Forward-backward algorithms with different inertial terms for structured non-convex minimization problems*, (2020), arxiv.org/abs/2002.07154
[29] T. Lin, M.I. Jordan, *A Control-Theoretic Perspective on Optimal High-Order Optimization*, (2019), arXiv:1912.07168v1
[30] R. May, *Asymptotic for a second-order evolution equation with convex potential and vanishing damping term*, Turkish Journal of Math., 41(3) (2017), 681-685
[31] Y. Nesterov, *A method of solving a convex programming problem with convergence rate* $O(1/k^2)$, Soviet Mathematics Doklady, 27 (1983), 372-376
[32] B. Shi, S.S. Du, M.I. Jordan, W.J. Su, *Understanding the acceleration phenomenon via high-resolution differential equations*, (2018), arXiv:1810.08907v3
[33] 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) (2016), 1-43