Appl. Set-Valued Anal. Optim. 7 (2025), No. 3, pp. 373-392
Available online at http://asvao.biemdas.com
https://doi.org/10.23952/asvao.7.2025.3.06

ON THE SIMULTANEOUS CONVERGENCE OF VALUES AND TRAJECTORIES OF CONTINUOUS INERTIAL DYNAMICS WITH TIKHONOV REGULARIZATION TO SOLVE CONVEX MINIMIZATION WITH AFFINE CONSTRAINTS

FOUAD BATTAHI, ZAKI CHBANI, HASSAN RIAHI*

Laboratory of Mathematics, Modeling and Automatic Systems, Faculty of Sciences Semlalia, Cadi Ayyad university, 40000 Marrakech, Morocco

Abstract. In this paper, we propose in a Hilbertian setting a second-order time-continuous dynamic system with fast convergence guarantees to solve general convex minimization problems with linear constraints. The system is associated with the augmented Lagrangian formulation of a minimization problem. The corresponding dynamic involves three general time-varying parameters, which are respectively associated with viscous damping, extrapolation and temporal scaling. By appropriately adjusting these parameters, each with specific properties, we develop a Lyapunov analysis which provides fast convergence properties of the values and of the feasibility gap. These results naturally pave the way for developing corresponding accelerated ADMM algorithms, obtained by temporal discretization.
Keywords. Damping; Inertial dynamics; Tikhoonov regularization; Time-continuous dynamic system.
2020 Mathematics Subject Classification. 34D05, 37N40, 46N10, 90C25.

1. Introduction

In this paper, $\mathscr{X}$ denotes a real Hilbert space, endowed with scalar product $\langle\cdot,\cdot\rangle$ and norm $\|x\|^2 = \langle x,x\rangle$, for any $x \in \mathscr{X}$. We are interested by the treatment of the following convex minimization problem under linear constraints:

$$\min_{x\in C} f(x) \;\text{ where } C := \{x \in \mathscr{X} : Ax = b\}, \tag{1.1}$$

where

$$\begin{cases} f : \mathscr{X} \to \mathbb{R} \text{ is a convex and continuously differentiable function,} \\ A \text{ is a linear and continuous operator from } \mathscr{X} \text{ to } \mathscr{Z} \text{ another real Hilbert space and } b \in \mathscr{Z}, \\ S := \operatorname{argmin}_C f \neq \emptyset \; \textit{and } x^* \text{ is the element of minimum norm of } S. \end{cases} \tag{$\mathbf{H}_0$}$$

Our objective in this paper is to provide a rigorous treatment of the convergence analysis of primal-dual dynamics by combining recent dynamic methods in the unconstrained minimization which ensures strong convergence (see [1, 2, 3]), and also those of the second order in time

*Corresponding author.
E-mail address: fouadbattahi99@gmail.com (F. Battahi), chbaniz@uca.ac.ma (Z. Chbani), h-riahi@uca.ac.ma (H. Riahi).
Received 1 February 2024; Accepted 10 June 2024; Published online 11 April 2025.

(see [4, 5, 6]) which were constructed to solve with fast convergence a similar constrained minimization problem.

Continuous-time approaches for the case of unconstrained convex minimization problem $\min_{x\in\mathscr{X}} f(x)$ were initiated as the heavy ball with friction method by Polyak [7, 8]:

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

In the case where f is μ-strongly convex, by fixing $\alpha=2\sqrt{\mu}$ in [9], the heavy ball system provides linear convergence of values $f(x(t))$ to $\min f$ (resp. trajectories $x(t)$ to the unique minimizer of f).

In [10, 11], the authors investigated on the asymptotic behaviour, when $t\to+\infty$, of the trajectories of the inertial system with Hessian-driven damping

$$\ddot{x}(t)+\gamma(t)\dot{x}(t)+\beta(t)\nabla^2 f(x(t))\dot{x}(t)+b(t)\nabla f(x(t))=0,$$

where $\gamma(t)$ and $\beta(t)$ are damping parameters, and $b(t)$ is a time scale parameter. Based on a Lyapunov analysis, and a continuous time version of Opial's lemma, they proved additional estimations for values and proved the weak convergence of the trajectories.

Here, in [11, Theorem 2.2], the convergence of trajectories was proved for the weak topology of $\mathscr{H}$. It is a natural question to ask whether one can obtain strong convergence. A counterexample due to Baillon [12] demonstrates that the trajectories of the continuous steepest descent may converge weakly but not strongly. We do not elaborate more on this for the sake of brevity. More recently, Attouch et al. [13] considered for $\delta>0$ the following system

$$\ddot{x}(t)+\frac{\delta}{t^{r/2}}\dot{x}(t)+\nabla f(x(t))+\frac{1}{t^r}x(t)=0.$$

They obtained, for $0<r<2$, strong asymptotic convergence towards the minimum norm solution and the following convergence rates

$$f(x(t))-\min_{\mathscr{X}} f=\mathscr{O}\left(\frac{1}{t^r}\right) \quad\text{and}\quad \left\|\dot{x}(t)\right\|^2=\mathscr{O}\left(\frac{1}{t^{\frac{r+2}{2}}}\right).$$

In this perspective, [3] introduced the dynamical system:

$$\ddot{x}(t)+\alpha\,\dot{x}(t)+\beta(t)\nabla f_t(x(t))=0, \tag{1.2}$$

and [14] proposed the following two inertial systems involving Hessian-driven damping:

$$\ddot{x}(t)+\alpha\,\dot{x}(t)+\delta\frac{d}{dt}\left(\nabla f(x(t))\right)+\beta(t)\nabla f_t(x(t))=0,$$

and

$$\ddot{x}(t)+\alpha\,\dot{x}(t)+\delta\frac{d}{dt}\left(\nabla f_t(x(t))\right)+\beta(t)\nabla f_t(x(t))=0,$$

where $f_t(\cdot):=f(\cdot)+\dfrac{c}{2\beta(t)}\|\cdot\|^2$ is a $\dfrac{c}{\beta(t)}$-strongly convex function, with the following hypothesis

$$\begin{cases} (i) & \alpha,c>0,\\ (ii) & \beta:[t_0,+\infty[\longrightarrow]0,+\infty[\ \text{ is a nondecreasing continuously}\\ & \text{differentiable function satisfying } \lim\limits_{t\to+\infty}\beta(t)=+\infty,\end{cases}$$

by assuming

$$\begin{cases} (i) \quad c \geq \alpha^2 > 0, \mu = \dfrac{\alpha}{1+a}, a > 1, \\ (ii) \quad \beta(t) \text{ is a twice continuously differentiable function with} \\ \qquad \lim\limits_{t\to+\infty} \dfrac{\dot{\beta}(t)}{\beta(t)} = 0, \limsup\limits_{t\to+\infty} \dfrac{-\ddot{\beta}(t)}{\dot{\beta}(t)} < \dfrac{\alpha}{2}. \end{cases}$$

Theorem 3.1 in [3] ensured for t large enough that

$$f(x(t)) - \min_{\mathscr{X}} f = \mathscr{O}\left(\frac{1}{\beta(t)}\right) \text{ and } \|\dot{x}(t)\|^2 = \mathscr{O}\left(\frac{\dot{\beta}(t)}{\beta(t)} + e^{-\mu t}\right). \tag{1.3}$$

As interesting special cases, the authors proposed

$$\beta(t) = t^m e^{\gamma t^p} \text{ with } (p,m) \in (\mathbb{R}_+)^2 \setminus \{(0,0)\}, 0 < p < 1, \gamma > 0.$$

A common strategy for constructing such a dynamic method for constrained minimization consists of adapting (1.2) for saddle functions. Note that the constrained minimization problem (1.1) can be equivalently reformulated as the saddle point problem

$$\min_{x\in\mathscr{X}} \max_{\lambda\in\mathscr{Z}} \mathscr{L}(x,\lambda),$$

where the Lagrangian $\mathscr{L} : \mathscr{X} \times \mathscr{Z} \to \mathbb{R}$ is defined by

$$\mathscr{L}(x,\lambda) = f(x) + \langle \lambda, Ax - b \rangle.$$

Under our standing assumption ($\mathbf{H}_0$), $\mathscr{L}$ is a saddle function since it is convex with respect to $x \in \mathscr{X}$, and affine (and hence concave) with respect to $\lambda \in \mathscr{Z}$. Then, a point $\bar{x}$ is optimal for (1.1), and $\bar{\lambda}$ is a corresponding Lagrange multiplier if and only if $(\bar{x}, \bar{\lambda})$ is a saddle point of the Lagrangian saddle function $\mathscr{L}$, *i.e.* for every $(x,\lambda) \in \mathscr{X} \times \mathscr{Z}$,

$$\mathscr{L}(\bar{x},\lambda) \leq \mathscr{L}(\bar{x},\bar{\lambda}) \leq \mathscr{L}(x,\bar{\lambda}).$$

The existence of a saddle point thus plays a critical role in solving (1.1). We denote by $\bar{S}$ the set of saddle points of $\mathscr{L}$. The corresponding optimality conditions read

$$(\bar{x},\bar{\lambda}) \in \bar{S} \Longleftrightarrow \begin{cases} \nabla_x \mathscr{L}(\bar{x},\bar{\lambda}) = 0, \\ \nabla_\lambda \mathscr{L}(\bar{x},\bar{\lambda}) = 0, \end{cases} \Longleftrightarrow \begin{cases} \nabla f(\bar{x}) + A^*\bar{\lambda} = 0, \\ A\bar{x} - b = 0, \end{cases} \tag{1.4}$$

where ∇_x (respectively ∇_λ) is the gradient with respect to x (respectively to λ) and A^* is the adjoint operator of A.

The dynamical system that was investigated in recent years is

$$\begin{cases} \ddot{x}(t) + \alpha(t)\dot{x}(t) + \beta(t)\nabla_x \mathscr{L}_\mu\Big(x(t),\ \lambda(t) + \gamma(t)\dot{\lambda}(t)\Big) = 0, \\ \ddot{\lambda}(t) + \alpha(t)\dot{\lambda}(t) - \beta(t)\nabla_\lambda \mathscr{L}_\mu\Big(x(t) + \gamma(t)\dot{x}(t),\ \lambda(t)\Big) = 0, \\ (x(t_0), \lambda(t_0)) = (x_0, \lambda_0) \text{ and } (\dot{x}(t_0), \dot{\lambda}(t_0)) = (\dot{x}_0, \dot{\lambda}_0), \end{cases} \tag{TRIALS}$$

where $\alpha(t)$ is an extrapolation parameter, $\beta(t)$ is attached to the temporal scaling of the dynamic and $\gamma(t)$ is a viscous damping parameter. Here $\mathscr{L}_\mu$ is the known augmented Lagrangian defined by

$$\mathscr{L}_\mu(x,y) := \mathscr{L}(x,y) + \frac{\mu}{2}\|Ax-b\|^2.$$

The case that $\beta(t) = 1$ was studied in [15, 16], while the case that $\alpha(t) = \frac{\alpha}{t^s}$ (for $0 < s \leq 1$) and $\beta(t)$ is more general were treated in [4, 6, 16, 17].

Note that, in unconstrained minimization (see [18, 19, 20, 21]), the viscous Nesterov damping term $\alpha(t) = \frac{\alpha}{t}$ plays an important role in obtaining for values the fast convergence of order $\mathscr{O}\left(\frac{1}{t^2}\right)$. The role of the viscous damping factor $\gamma(t)\dot{x}(t)$ is to induce more flexibility in the dynamic system and also to validate the convergence conditions as was recently noticed in [4, 15, 16, 17, 22]. As we will assert that the temporal scaling function $\beta(.)$ has the role of further improving the convergence rates of the value of the objective function along the trajectory, as was noticed in the context of unconstrained minimization problems in [18, 19, 23, 24] and linearly constrained minimization problems in [4, 5, 6].

Note that, in all the works cited above, the strong convergence of the paths $x(t)$ is only ensured under strong conditions. Our goal in what follows is to draw inspiration from our recent works [2, 3, 13, 14] on unconstrained minimization in order to conclude it for general convex-concave saddle functions. To reach a solution to the constrained optimization problem (1.1), we consider a primal-dual dynamical system where we approach this problem via a two-level continuous path:

The first level is a penalization of the associated Lagrangian $\mathscr{L}(x,\lambda)$ by a strongly convex-concave saddle function, which is an other augmented Lagrangian $\mathscr{L}_t : \mathscr{X} \times \mathscr{Z} \to \mathbb{R}$ defined, for $r, c > 0$ and $t > t_0$, by

$$\mathscr{L}_t(x,\lambda) = \mathscr{L}(x,\lambda) + \frac{c}{2t^r}(\|x\|^2 - \|\lambda\|^2).$$

This ensures the existence and uniqueness of an associated saddle point (x_t, λ_t). We choose as a penalization parameter the time function $\frac{c}{t^r}$ which tends towards zero when t goes to infinity.

The second level consists of adapting a suitable associated dynamic system which can ensure in double slice the strong convergence of its solution towards an optimal solution of (1.4), and also have the fastest possible convergence rates.

This dynamic system, which is called Mixed Inertial Primal-Dual Augmented Lagrangian System, is written as follows: for $t > t_0$

$$\left\{\begin{array}{l} \ddot{x}(t) + \alpha\dot{x}(t) + t^r\nabla_x\mathscr{L}_t(x(t),\lambda(t)) = 0, \\ \dot{\lambda}(t) - t^r\nabla_\lambda\mathscr{L}_t\left(x(t) + \frac{1}{\tau}\dot{x}(t),\lambda(t)\right) = 0, \\ (x(t_0),\lambda(t_0)) = (x_0,\lambda_0) \text{ and } \dot{x}(t_0) = \dot{x}_0, \end{array}\right. \tag{MIPDALS}$$

where $\alpha > 0$ is a damping parameter, t^r is attached to the temporal scaling of the dynamic and $1/\tau > 0$ is an extrapolation parameter, and $x_0, \dot{x}_0 \in \mathscr{X}$ and $\lambda_0 \in \mathscr{Z}$. The dynamical system (MIPDALS), which was investigated in more recent papers [5, 25, 26], differs from the (TRIALS) system proposed above. We first notice the non-coincidence between the proposed augmented Lagrangians which differ in their penalization factors, and then in (MIPDALS) we restrict ourselves to a times first order differential equation for the variations of $\lambda(t)$.

In previous papers dealing with dynamic systems to attain saddle points, these results rely on Lyapunov functions $\mathscr{E}(t)$ based on selected solutions $(x(t), y(t))$ and saddle points $z^* := (x^*, \lambda^*)$

of $\mathscr{L}$. Our proof is based on the following Lyapunov function

$$\mathscr{E}(t) := t^r \left(\mathscr{L}_t(x, \lambda_t) - \mathscr{L}_t(x_t, \lambda_t) \right) + \frac{1}{2}\|v(t)\|^2 + \frac{\tau}{2}\|\lambda(t) - \lambda_t\|^2$$

where (x_t, λ_t) is the unique saddle point of $\mathscr{L}_t$, $v(t) = \tau(x(t) - x_t) + \dot{x}(t)$, $r, \tau > 0$ and the temporal scaling parameter function is t^r.

We will show in Theorem 3.1 that under a judicious setting of parameters, $\mathscr{E}(t)$ satisfies the first-order differential inequality

$$\frac{d}{dt}\left[e^{\mu t}\mathscr{E}(t)\right] \leq \frac{\|z^*\|^2}{2}\frac{d}{dt}\left[\frac{e^t}{t^{1-r}}\right],$$

which by integration states our main convergence Theorem 3.2. Let us select these convergence rates

$$f(x(t)) - \min_C f \leq \mathscr{O}\left(\frac{1}{t^r}\right), \; \|Ax(t) - b\| = \mathscr{O}\left(\frac{1}{t^r}\right), \; \|x(t) - x_t\|^2 = \mathscr{O}\left(\frac{1}{t^{1-r}}\right),$$

where that of the values and constraints are better and that of the path $x(t)$ ensures its strong convergence towards the solution closest to the origin.

The remainder of the paper is organized as follows. Next, in Section 2, we introduce the setting that we will work with and formulate the proposed Lyapunov energy function. This is followed by the main estimation of this function. Afterwards, we investigate the main convergence theorem on the values, trajectories, and velocities in Section 3. Two primary special cases for the function β are treated in Section 4, for which numerical experiments are given for a simple convex (not strictly convex) function. Finally, on the basis of the Moreau regularization technique, we extend our results to non-smooth convex functions with extended real values in Section 5.

2. Control of Variations for the Saddle Points of the New Augmented Lagrangian Functions

In this section, we present the new Lagrangian function $\mathscr{L}_t : \mathscr{X} \times \mathscr{Z} \to \mathbb{R}$ defined by

$$\begin{aligned}\mathscr{L}_t(x, \lambda) &= \mathscr{L}(x, \lambda) + \frac{c}{2t^r}(\|x\|^2 - \|\lambda\|^2)\\ &= f(x) + \langle \lambda, Ax - b\rangle + \frac{c}{2t^r}(\|x\|^2 - \|\lambda\|^2).\end{aligned}$$

For each $t \geq t_0$, let us set $(x_t, \lambda_t) := \operatorname{argminmax}_{\mathscr{X}\times\mathscr{Z}} \mathscr{L}_t$, which is the unique saddle-point of the strongly convex-concave saddle function $\mathscr{L}_t$. The first order optimality conditions give

$$\begin{cases} 0 = \nabla_x \mathscr{L}_t(x_t, \lambda_t) = \nabla f(x_t) + A^*\lambda_t + \frac{c}{t^r}x_t, \\ 0 = \nabla_\lambda \mathscr{L}_t(x_t, \lambda_t) = Ax_t - b - \frac{c}{t^r}\lambda_t. \end{cases} \tag{2.1}$$

We begin with some auxiliary results

Lemma 2.1. [26, Lemma 6] *Let $t_0 \geq 0$, $g : [t_0, +\infty) \to \mathscr{Z}$ a continuous differentiable function, and $a : [t_0, +\infty) \to [0, +\infty)$ a continuous function. Suppose that there exists $C \geq 0$ such that, for* *every* $t > t_0$,

$$\left\| g(t) + \int_{t_0}^{t} a(s)g(s)ds \right\| \leq C. \tag{2.2}$$

Then $\sup_{t \geq t_0} \|g(t)\| < +\infty$.

Proof. Set $G(t) := \exp\left(\int_{t_0}^t a(s)ds\right) \int_{t_0}^t a(s)g(s)ds$ for $t \geq t_0$. Then condition (2.2) ensures that

$$\left\| \frac{d}{dt} G(t) \right\| \leq Ca(t) \exp\left(\int_{t_0}^t a(s)ds\right) = C\frac{d}{dt}\left(\exp\left(\int_{t_0}^t a(s)ds\right)\right).$$

Using $G(t_0) = 0$, we obtain

$$\begin{aligned}
\exp\left(\int_{t_0}^t a(s)ds\right) \left\| \int_{t_0}^t a(s)g(s)ds \right\| &= \|G(t)\| = \left\| \int_{t_0}^t \frac{d}{dt} G(s)ds \right\| \\
&\leq \int_{t_0}^t \left\| \frac{d}{dt} G(s) \right\| ds \\
&\leq C\left(\exp\left(\int_{t_0}^t a(s)ds\right) - 1\right) \leq C\exp\left(\int_{t_0}^t a(s)ds\right).
\end{aligned}$$

Thus

$$\left\| \int_{t_0}^t a(s)g(s)ds \right\| \leq C.$$

Return to condition (2.2), we conclude, for each $t \geq t_0$,

$$\|g(t)\| \leq \left\| g(t) + \int_{t_0}^t a(s)g(s)ds \right\| + \left\| \int_{t_0}^t a(s)g(s)ds \right\| \leq 2C.$$

□

Lemma 2.2. *Under conditions* (**H**$_0$) *and* $c >, 0 < r < 1$, *it holds that, for all* $(x, \lambda) \in \mathscr{X} \times \mathscr{Z}$ *and* $t \geq t_0$,

(i) $\mathscr{L}_t(x, \lambda_t) - \mathscr{L}_t(x_t, \lambda_t) \geq \frac{c}{2t^r}\|x - x_t\|^2$,

(ii) $\mathscr{L}_t(x_t, \lambda_t) - \mathscr{L}_t(x_t, \lambda) \geq \frac{c}{2t^r}\|\lambda - \lambda_t\|^2$.

Proof. We give only the proof for (i) since (ii) is similar. We first remark that for each $x \in \mathscr{X}$ and $t \geq t_0$

$$\begin{aligned}
\left\langle \nabla_x \mathscr{L}_t(x, \lambda_t) - \nabla_x \mathscr{L}_t(x_t, \lambda_t), x - x_t \right\rangle &= \left\langle \nabla f(x) - \nabla f(x_t), x - y \right\rangle + \frac{c}{t^r}\|x - x_t\|^2 \\
&\geq \frac{c}{t^r}\|x - x_t\|^2.
\end{aligned}$$

It follows that $\nabla_x \mathscr{L}_t(\cdot, \lambda_t)$ is strongly monotone. From [27, Corollary 3.5.11], we conclude strong convexity of $\mathscr{L}_t(\cdot, \lambda_t)$. Thus, for each $x \in \mathscr{X}$ and each $t \geq t_0$,

$$\mathscr{L}_t(x, \lambda_t) - \mathscr{L}_t(x_t, \lambda_t) \geq \frac{c}{2t^r}\|x - x_t\|^2 + \left\langle \nabla_x \mathscr{L}_t(x_t, \lambda_t), x - x_t \right\rangle = \frac{c}{2t^r}\|x - x_t\|^2.$$

□

Lemma 2.3. *Assume conditions* ($\mathbf{H}_0$) *and* $c >, 0 < r < 1$ *and denote by* (x^*,λ^*) *the metric projection of* $(0_{\mathcal{X}}, 0_{\mathcal{Z}})$ *on* $\bar{S}$ *the set of saddle points of* $\mathcal{L}$. *Then,*

(i) *for all* $t > t_0$, $\|(x_t,\lambda_t)\| \leq \|(x^*,\lambda^*)\|$ *and* $\lim_{t\to+\infty}\|(x_t,\lambda_t)-(x^*,\lambda^*)\| = 0$,

(ii) *for all* $t > t_0$, $\left\|(\dot{x}_t,\dot{\lambda}_t)\right\| \leq \frac{r}{t}\|(x_t,\lambda_t)\| \leq \frac{r}{t}\|(x^*,\lambda^*)\|$.

Proof. **(i)** Consider the operator $\mathcal{M} : \mathcal{X}\times\mathcal{Z} \to \mathcal{X}\times\mathcal{Z}$ defined by

$$\mathcal{M}(x,\lambda) := \left(\nabla_x\mathcal{L}(x,\lambda)\,,\ -\nabla_\lambda\mathcal{L}(x,\lambda)\right).$$

Then $\mathcal{M}$ is the monotone operator associated with the convex-concave function $\mathcal{L}$, and since it is also continuous on $\mathcal{X}\times\mathcal{Z}$, it is maximally monotone (see, e.g., [28, Corollary 20.28]).

We also have the set of zeros of the maximally monotone operator $\mathcal{M}$ is nothing other than the whole set of saddle points of $\mathcal{L}$. This means that the solution set $\bar{S}$ is a closed convex subset of $\mathcal{X}\times\mathcal{Z}$. From (2.1), we also have that (x_t,λ_t) is characterized by

$$\left(\mathcal{M}+\frac{c}{t^r}\mathbb{I}\right)(x_t,\lambda_t) = (0_{\mathcal{X}},0_{\mathcal{Z}}) \iff (x_t,\lambda_t) = \left(\mathbb{I}+\frac{t^r}{c}\mathcal{M}\right)^{-1}(0_{\mathcal{X}},0_{\mathcal{Z}}).$$

Using [29, Theorem 2.2] (see also [28, Theorem 23.44]), we have (x_t,λ_t) strongly converges to (x^*,λ^*), and [29, Propo. 2.6 (iii)] ensures also that, for every $t > t_0$, $\|(x_t,\lambda_t)\| \leq \|(x^*,\lambda^*)\|$.

(ii) Set $w(t) = (x_t,\lambda_t)$,. From (2.1), we have for $t > t_0$ and h near zero

$$\mathcal{M}(w(t)) = -\frac{c}{t^r}w(t) \text{ and } \mathcal{M}(w(t+h)) = -\frac{c}{(t+h)^r}(w(t+h)).$$

By monotonicity of $\mathcal{M}$, we have

$$\begin{aligned}&\left\langle \mathcal{M}(w(t+h)) - \mathcal{M}(w(t)), w(t+h)-w(t)\right\rangle\\ &= \left\langle \frac{c}{t^r}w(t) - \frac{c}{(t+h)^r}(w(t+h)), w(t+h)-w(t)\right\rangle \geq 0.\end{aligned}$$

Thus, for each $t > t_0$ and h sufficiently small,

$$\|w(t+h)-w(t)\|^2 \leq \left(\left(1+\frac{h}{t}\right)^r - 1\right)\left\langle w(t), w(t+h)-w(t)\right\rangle,$$

which implies, by the mean value theorem, that there exists c_h between 0 and $\frac{h}{t}$ such that

$$\|w(t+h)-w(t)\| \leq \left|\left(1+\frac{h}{t}\right)^r - 1\right|\left\|w(t)\right\| = \frac{r|h|}{t(1+c_h)^{1-r}}\left\|w(t)\right\|. \tag{2.3}$$

We see that the viscosity curve $w(t)$ is Lipschitz continuous on each compact interval in $]t_0,+\infty[$. We conclude that $w(t)$ is absolutely continuous and then almost everywhere differentiable on $]t_0,+\infty[$. Return to (2.3), dividing by $h > 0$ and letting $h \to 0$, we obtain for almost every $t > t_0$ that

$$\left\|\dot{w}(t)\right\| \leq \frac{r}{t}\|w(t)\| \leq \frac{r}{t}\|(x^*,\lambda^*)\|,$$

which indicates that (ii) is satisfied. □

We now provide the following needed control lemma.

Lemma 2.4. *If $\alpha, c > 0$ and $0 < r < 1$, then, for each $t > t_0$,*

$$\frac{d}{dt}\mathscr{L}_t(x_t,\lambda_t) = \frac{-cr}{2t^{r+1}}\left(\|x_t\|^2 - \|\lambda_t\|^2\right).$$

Proof. Let us fix $t > t_0$. Since (x_t,λ_t) is a saddle-point of the saddle function $\mathscr{L}_t$, we obtain for each $t > t_0$ and h small enough

$$\mathscr{L}_t(x_t,\lambda_{t+h}) \le \mathscr{L}_t(x_t,\lambda_t) \le \mathscr{L}_t(x_{t+h},\lambda_t)$$

and

$$-\mathscr{L}_{t+h}(x_t,\lambda_{t+h}) \le -\mathscr{L}_{t+h}(x_{t+h},\lambda_{t+h}) \le -\mathscr{L}_{t+h}(x_{t+h},\lambda_t).$$

By summing, we obtain for h small enough the following two inequalities

$$\begin{aligned}
\mathscr{L}_t(x_t,\lambda_t) - \mathscr{L}_{t+h}(x_{t+h},\lambda_{t+h}) &\le \mathscr{L}_t(x_{t+h},\lambda_t) - \mathscr{L}_{t+h}(x_{t+h},\lambda_t)\\
&= \frac{c}{2t^r}(\|x_{t+h}\|^2 - \|\lambda_t\|^2) - \frac{c}{2(t+h)^r}(\|x_{t+h}\|^2 - \|\lambda_t\|^2)\\
&= \frac{c}{2t^r}\left(1-\left(1+\frac{h}{t}\right)^{-r}\right)\left(\|x_{t+h}\|^2 - \|\lambda_t\|^2\right)\\
&= \left(\frac{crh}{2t^{r+1}} + \circ(h)\right)\left(\|x_{t+h}\|^2 - \|\lambda_t\|^2\right)
\end{aligned}$$

and

$$\begin{aligned}
\mathscr{L}_t(x_t,\lambda_t) - \mathscr{L}_{t+h}(x_{t+h},\lambda_{t+h}) &\ge \mathscr{L}_t(x_t,\lambda_{t+h}) - \mathscr{L}_{t+h}(x_t,\lambda_{t+h})\\
&= \frac{c}{2t^r}\left(1-\left(1+\frac{h}{t}\right)^{-r}\right)\left(\|x_t\|^2 - \|\lambda_{t+h}\|^2\right)\\
&= \left(\frac{crh}{2t^{r+1}} + \circ(h)\right)\left(\|x_t\|^2 - \|\lambda_{t+h}\|^2\right).
\end{aligned}$$

So dividing the previous inequalities by $h > 0$ and letting $h \to 0$, we obtain the result immediately. □

3. Fast Convergence Results

In this section, we are going to derive fast convergence rates for the primal-dual Augmented Lagrangian, the feasibility measure, and the objective function value along the trajectories generated by the dynamical system (MIPDALS), which could be written for $c, \alpha, r, \tau > 0$ as follows:

$$\begin{cases}
\ddot{x}(t) + \alpha\dot{x}(t) + t^r[\nabla f(x(t)) + A^*\lambda(t)] + cx(t) &= 0,\\
\dot{\lambda}(t) - t^r\left[A\left(x(t) + \frac{1}{\tau}\dot{x}(t)\right) - b\right] + c\lambda(t) &= 0,\\
(x(t_0),\lambda(t_0)) = (x_0,\lambda_0) \text{ and } \dot{x}(t_0) = \dot{x}_0.
\end{cases}$$

We will also derive the main result on the strong convergence of trajectories $x(t)$ towards the minimizer of the minimum norm. As mentioned in the introduction, our proof is based on the

Lyapunov function $\mathscr{E}$ which is formulated as follows:

$$\mathscr{E}(t) := t^r\left(\mathscr{L}_t(x(t),\lambda_t) - \mathscr{L}_t(x_t,\lambda_t)\right) + \frac{1}{2}\|v(t)\|^2 + \frac{\tau}{2}\|\lambda(t)-\lambda_t\|^2 \qquad (\mathscr{E})$$

with $v(t) = \tau(x(t)-x_t) + \dot{x}(t)$.

The next theorem provides the analysis needed on the energy function $\mathscr{E}(t)$. So, we need the following condition on the parameters α, τ.

Theorem 3.1. *Suppose that $f : \mathscr{H} \to \mathbb{R}, A : \mathscr{X} \to \mathscr{Z}$ and $\beta(t)$ satisfy conditions ($\mathbf{H}_0$) and ($\mathbf{H}_1$). Let $(x(\cdot),\lambda(\cdot))$ be a solution to system (MIPDALS), and assume the following condition*

$$0 < r < 1, \tau < \alpha < \tau + \min(\tau, c) \text{ and either } \alpha < 2\sqrt{c} \text{ or } 2\sqrt{c} < \alpha < \tau + \frac{c}{\tau}. \qquad (\mathbf{H}_1)$$

Then, there exists $\bar{t} > t_0$ such that, for each $t \geq \bar{t}$, the following rate holds:

$$\mathscr{E}(t) \leq \frac{e^{(\alpha-\tau)\bar{t}}\mathscr{E}(\bar{t})}{e^{(\alpha-\tau)t}} + \frac{\|z^*\|^2}{2(\alpha-\tau)}\frac{1}{t^{1-r}}. \qquad (3.1)$$

Proof. Let us derive the Lyapunov energy function $\mathscr{E}(t)$. Observe that

$$\begin{aligned} \frac{d}{dt}\mathscr{E}(t) &= rt^{r-1}\left(\mathscr{L}_t(x(t),\lambda_t) - \mathscr{L}_t(x_t,\lambda_t)\right) + t^r\frac{d}{dt}\left(\mathscr{L}_t(x(t),\lambda_t) - \mathscr{L}_t(x_t,\lambda_t)\right) \\ &\quad + \langle v(t),\dot{v}(t)\rangle - \tau\langle\dot{\lambda}_t,\lambda(t)-\lambda_t\rangle + \tau\langle\dot{\lambda}(t),\lambda(t)-\lambda_t\rangle. \end{aligned} \qquad (3.2)$$

Using the system (MIPDALS) and adapting calculation, we have

$$\begin{aligned} \dot{v}(t) &= \tau(\dot{x}(t)-\dot{x}_t) + \ddot{x}(t) \\ &= (\tau-\alpha)\dot{x}(t) - \tau\dot{x}_t - t^r\nabla_x\mathscr{L}_t(x(t),\lambda(t)) \\ &= (\tau-\alpha)\dot{x}(t) - \tau\dot{x}_t - t^r\left(\nabla_x\mathscr{L}_t(x(t),\lambda_t) + A^*(\lambda(t)-\lambda_t)\right). \end{aligned}$$

Then

$$\begin{aligned} \langle\dot{v}(t),v(t)\rangle &= (\tau-\alpha)\|\dot{x}(t)\|^2 + \tau(\tau-\alpha)\langle x(t)-x_t,\dot{x}(t)\rangle - \tau^2\langle x(t)-x_t,\dot{x}_t\rangle \\ &\quad -\tau\,\langle\dot{x}(t),\dot{x}_t\rangle - \tau t^r\langle\nabla_x\mathscr{L}_t(x(t),\lambda(t)),x(t)-x_t\rangle \\ &\quad -t^r\langle\nabla_x\mathscr{L}_t(x(t),\lambda(t)),\dot{x}(t)\rangle. \end{aligned} \qquad (3.3)$$

Using again system (MIPDALS) , we have

$$\begin{aligned} \langle\dot{\lambda}(t),\lambda(t)-\lambda_t\rangle &= t^r\left\langle\nabla_\lambda\mathscr{L}_t\left(x(t)+\frac{1}{\tau}\dot{x}(t),\lambda(t)\right),\lambda(t)-\lambda_t\right\rangle \qquad (3.4) \\ &= t^r\left\langle\nabla_\lambda\mathscr{L}_t(x_t,\lambda(t)) + A\left(x(t)-x_t+\frac{1}{\tau}\dot{x}(t)\right),\lambda(t)-\lambda_t\right\rangle. \end{aligned}$$

Moreover, for positive parameters a, p, and q, we have

$$-\tau\langle \dot{x}(t), \dot{x}_t\rangle \le \frac{\tau}{2a}\|\dot{x}(t)\|^2 + \frac{a\tau}{2}\|\dot{x}_t\|^2, \tag{3.5}$$

$$-\tau\langle \lambda(t) - \lambda_t, \dot{\lambda}_t\rangle \le \frac{\tau}{2p}\|\dot{\lambda}_t\|^2 + \frac{p\tau}{2}\|\lambda(t) - \lambda_t\|^2, \tag{3.6}$$

$$-\tau^2\langle x(t) - x_t, \dot{x}_t\rangle \le \frac{\tau}{2q}\|\dot{x}_t\|^2 + \frac{q\tau^3}{2}\|x(t) - x_t\|^2. \tag{3.7}$$

By strong convexity of $\mathscr{L}_t(., \lambda_t)$, we have

$$\langle \nabla_x \mathscr{L}_t(x(t), \lambda_t), x(t) - x_t\rangle \ge \frac{c}{2t^r}\|x(t) - x_t\|^2 + \left(\mathscr{L}_t(x(t), \lambda_t) - \mathscr{L}_t(x_t, \lambda_t)\right) \tag{3.8}$$

Using

$$\nabla_x \mathscr{L}_t(x(t), \lambda(t)) = \nabla_x \mathscr{L}_t(x(t), \lambda_t) + A^*(\lambda(t) - \lambda_t)$$

and (3.8), we obtain

$$\begin{aligned}
&-\tau t^r\langle \nabla_x \mathscr{L}_t(x(t), \lambda(t)), x(t) - x_t\rangle \\
&= -\tau t^r\langle \nabla_x \mathscr{L}_t(x(t), \lambda_t), x(t) - x_t\rangle - \tau t^r\langle A(x(t) - x_t), \lambda(t) - \lambda_t\rangle \\
&\le -\frac{\tau c}{2}\|x(t) - x_t\|^2 - \tau t^r\left(\mathscr{L}_t(x(t), \lambda_t) - \mathscr{L}_t(x_t, \lambda_t)\right) - \tau t^r\langle A(x(t) - x_t), \lambda(t) - \lambda_t\rangle.
\end{aligned} \tag{3.9}$$

By the strong convexity of $-\mathscr{L}_t(x_t, .)$, we also have

$$\langle -\nabla_\lambda \mathscr{L}_t(x_t, \lambda(t)), \lambda(t) - \lambda_t\rangle \ge \frac{c}{t^r}\|\lambda(t) - \lambda_t\|^2.$$

Combining the above inequality with

$$\nabla_\lambda \mathscr{L}_t\left(x(t) + \frac{1}{\tau}\dot{x}(t), \lambda(t)\right) = \nabla_\lambda \mathscr{L}_t(x_t, \lambda(t)) + A\left(x(t) - x_t + \frac{1}{\tau}\dot{x}(t)\right),$$

we obtain

$$\begin{aligned}
&-\tau t^r\left\langle -\nabla_\lambda \mathscr{L}_t\left(x(t)\frac{1}{\tau}\dot{x}(t), \lambda(t)\right), \lambda(t) - \lambda_t\right\rangle \\
&= -\tau t^r\langle -\nabla_\lambda \mathscr{L}_t(x_t, \lambda(t)), \lambda(t) - \lambda_t\rangle + \tau t^r\langle A(x(t) - x_t), \lambda(t) - \lambda_t\rangle + t^r\langle A\dot{x}(t), \lambda(t) - \lambda_t\rangle \\
&\le -c\tau\|\lambda(t) - \lambda_t\|^2 + \tau t^r\langle A(x(t) - x_t), \lambda(t) - \lambda_t\rangle + t^r\langle A\dot{x}(t), \lambda(t) - \lambda_t\rangle.
\end{aligned} \tag{3.10}$$

On the other hand, we have

$$\begin{aligned}
\frac{d}{dt}\mathscr{L}_t(x(t), \lambda_t) &= \langle \nabla f(x(t)), \dot{x}(t)\rangle + \langle Ax(t) - b, \dot{\lambda}_t\rangle + \langle A^*\lambda_t, \dot{x}(t)\rangle + \frac{c}{t^r}\langle x(t), \dot{x}(t)\rangle \\
&\quad -\frac{c}{t^r}\langle \lambda_t, \dot{\lambda}_t\rangle - \frac{cr}{2t^{r+1}}(\|x(t)\|^2 - \|\lambda_t\|^2) \\
&= \langle \nabla_x \mathscr{L}_t(x(t), \lambda(t)), \dot{x}(t)\rangle + \langle A^*(\lambda_t - \lambda(t)), \dot{x}(t)\rangle + \langle Ax(t) - b, \dot{\lambda}_t\rangle \\
&\quad -\frac{c}{t^r}\langle \lambda_t, \dot{\lambda}_t\rangle - \frac{cr}{2t^{r+1}}(\|x(t)\|^2 - \|\lambda_t\|^2).
\end{aligned} \tag{3.11}$$

Using Lemma 2.4, we see that

$$\begin{aligned}
&t^r \frac{d}{dt}\left(\mathscr{L}_t(x(t),\lambda_t)-\mathscr{L}_t(x_t,\lambda_t)\right)\\
&\le t^r\langle \nabla_x \mathscr{L}_t(x(t),\lambda(t)),\dot{x}(t)\rangle + t^r\langle A^*(\lambda_t-\lambda(t)),\dot{x}(t)\rangle + t^r\left(\langle Ax(t)-b,\dot{\lambda}_t\rangle - \frac{c}{t^r}\langle \lambda_t,\dot{\lambda}_t\rangle\right)\\
&\quad + \frac{cr}{2t^{r+1}}\left(\|x_t\|^2-\|x(t)\|^2\right).
\end{aligned} \tag{3.12}$$

Return to $Ax_t - b = \frac{c}{t^r}\lambda_t$, we obtain, for a positive parameter b,

$$\begin{aligned}
t^r\left(\langle Ax(t)-b,\dot{\lambda}_t\rangle - \frac{c}{t^r}\langle \lambda_t,\dot{\lambda}_t\rangle\right) &= \langle t^{(r-1)/2}A(x(t)-x_t), t^{(r+1)/2}\dot{\lambda}_t\rangle\\
&\le \frac{\|A\|^2 t^{r-1}}{2}\|x(t)-x_t\|^2 + \frac{t^{r+1}}{2}\|\dot{\lambda}_t\|^2.
\end{aligned}$$

By combining (3.2), (3.4), (3.5), (3.6), (3.7), (3.9), (3.10), and (3.12), we obtain

$$\begin{aligned}
\frac{d}{dt}\mathscr{E}(t) \;\le\;& t^r\left(rt^{-1}-\tau\right)\left(\mathscr{L}_t(x(t),\lambda_t)-\mathscr{L}_t(x_t,\lambda_t)\right)\\
&+\frac{1}{2}\left(q\tau^3 - c\tau + \|A\|^2 t^{r-1}\right)\|x(t)-x_t\|^2 + \frac{\tau}{2}(p-2c)\|\lambda(t)-\lambda_t\|^2\\
&+\left(\tau-\alpha+\frac{\tau}{2a}\right)\|\dot{x}(t)\|^2 + \frac{cr}{2t}(\|x_t\|^2-\|x(t)\|^2) + \frac{\tau}{2}\left(a+\frac{1}{q}\right)\|\dot{x}_t(t)\|^2\\
&+\frac{1}{2}\left(t^{r+1}+\frac{\tau}{p}\right)\|\dot{\lambda}_t\|^2 + \tau(\tau-\alpha)\langle \dot{x}(t), x(t)-x_t\rangle.
\end{aligned} \tag{3.13}$$

Now, we set $\mu > 0$ and estimate

$$\begin{aligned}
\mu\mathscr{E}(t) &= \mu t^r\left(\mathscr{L}_t(x(t),\lambda_t)-\mathscr{L}_t(x_t,\lambda_t)\right) + \frac{\mu}{2}\|v(t)\|^2 + \frac{\mu\tau}{2}\|\lambda(t)-\lambda_t\|^2\\
&= \mu t^r\left(\mathscr{L}_t(x(t),\lambda_t)-\mathscr{L}_t(x_t,\lambda_t)\right) + \frac{\mu\tau^2}{2}\|x(t)-x_t\|^2\\
&\quad +\mu\tau\langle x(t)-x_t,\dot{x}(t)\rangle + \frac{\mu}{2}\|\dot{x}(t)\|^2 + \frac{\mu\tau}{2}\|\lambda(t)-\lambda_t\|^2.
\end{aligned} \tag{3.14}$$

Adding (3.13) and (3.14) , we arrive at

$$\mu\mathcal{E}(t)+\frac{d}{dt}\mathcal{E}(t)\leq t^r\Big(\underbrace{(\mu-\tau)+rt^{-1}}_{=-B(t)}\Big)\Big(\mathcal{L}_t(x(t),\lambda_t)-\mathcal{L}_t(x_t,\lambda_t)\Big)$$
$$+\frac{1}{2}\Big(\underbrace{q\tau^3-c\tau+\|A\|^2t^{r-1}+\mu\tau^2}_{=-C(t)}\Big)\|x(t)-x_t\|^2+\frac{\tau}{2}\Big(\underbrace{p-2c+\mu}_{=-D}\Big)\|\lambda(t)-\lambda_t\|^2$$
$$+\Big(\underbrace{\tau-\alpha+\frac{\tau}{2a}+\frac{\mu}{2}}_{=-F}\Big)\|\dot{x}(t)\|^2+\frac{cr}{2t}(\|x_t\|^2-\|x(t)\|^2)$$
$$+\frac{\tau}{2}\left(a+\frac{1}{q}\right)\|\dot{x}_t(t)\|^2+\frac{1}{2}\left(t^{r+1}+\frac{\tau}{p}\right)\|\dot{\lambda}_t\|^2+\tau\Big(\underbrace{\mu+\tau-\alpha}_{=-K}\Big)\langle x(t)-x_t,\dot{x}(t)\rangle.$$

⋆ By taking $\mu=\alpha-\tau>0$, we obtain $K=0$.

⋆ Since $B(t)=(\tau-\mu)-rt^{-1}=2\tau-\alpha-rt^{-1}$ and $\lim_{t\to+\infty}rt^{-1}=0$, we suppose in (**H**$_1$) $\alpha<2\tau$ to ensure the existence of $t_1>0$ such that, for all $t\geq t_1$, $B(t)\geq 0$.

⋆ We have $C(t)=c\tau-q\tau^3-\mu\tau^2-\|A\|^2t^{r-1}$. Since, for $0<r<1$, $\lim_{t\to+\infty}\|A\|^2t^{r-1}=0$, we have to satisfy $\tau(c-q\tau^2-\mu\tau)=\tau((1-q)\tau^2-\alpha\tau+c)>0$. This is due to the choice $0<q<1+\frac{c-\alpha\tau}{\tau^2}$ which is guaranteed by assumption (**H**$_1$). We deduce the existence of $t_2\geq t_1$ such that, for all $t\geq t_2$, $C(t)>0$.

⋆ We have that $D=2c-p-\mu=2c+\tau-\alpha-p$ is nonnegative when we chose $0<p\leq 2c+\tau-\alpha$, so the condition $c>\frac{\alpha-\tau}{2}$ is imposed in (**H**$_1$).

⋆ Since $F=\frac{1}{2a}(a\alpha-(1+a)\tau)$, we also have the choice $a\geq\frac{\tau}{\alpha-\tau}$ to guarantee that $F\geq 0$.

Return to $z_t=(x_t,\lambda_t)$ and $z^*=(x^*,\lambda^*)$, we obtain

$$\|x_t\|\leq\|z_t\|\leq\|z^*\| \text{ and } \max\Big(\|\dot{x}_t\|\ ,\ \|\dot{\lambda}_t\|\Big)\leq\frac{r}{t}\|z_t\|\leq\frac{r}{t}\|z^*\|.$$

Thus, for all $t\geq t_2$,

$$\frac{\tau}{2}\left(a+\frac{1}{q}\right)\|\dot{x}_t(t)\|^2+\frac{1}{2}\left(t^{r+1}+\frac{\tau}{p}\right)\|\dot{\lambda}_t\|^2+\frac{cr}{2t}\|x_t\|^2 \leq \frac{\|z^*\|^2}{2}\left(\frac{rc}{t}+\left[\tau\left(a+\frac{1}{p}+\frac{1}{q}\right)+t^{r+1}\right]\left(\frac{r}{t}\right)^2\right). \tag{3.15}$$

Summarizing the choices for t above, we conclude from inequalities (3.13) and (3.15) that, for all $t\geq t_2$,

$$\mu\mathcal{E}(t)+\frac{d}{dt}\mathcal{E}(t)\leq\frac{\|z^*\|^2}{2}\left(\frac{rc}{t}+\left[\tau\left(a+\frac{1}{p}+\frac{1}{q}\right)+t^{r+1}\right]\left(\frac{r}{t}\right)^2\right).$$

Multiplying by $e^{\mu t}$, we have, for $k_0 := \tau\left(a+\frac{1}{p}+\frac{1}{q}\right)$ and all $t \geq t_3$,

$$\begin{aligned}
\frac{d}{dt}\left[e^{\mu t}\mathscr{E}(t)\right] &= e^{\mu t}\left[\mu\mathscr{E}(t)+\frac{d}{dt}\mathscr{E}(t)\right] \\
&\leq \frac{\|z^*\|^2}{2}\left(\frac{rc}{t}+\left[\tau\left(a+\frac{1}{p}+\frac{1}{q}\right)+t^{r+1}\right]\left(\frac{r}{t}\right)^2\right)e^{\mu t} \\
&= \frac{\|z^*\|^2}{2}\left(\frac{rc}{t}+\frac{k_0 r^2}{t^2}+\frac{r^2}{t^{1-r}}\right)e^{\mu t}. \qquad (3.16)
\end{aligned}$$

Since $r<1$, we have

$$\lim_{t\to+\infty}\left(\frac{rc}{t^r}+\frac{k_0 r^2}{t^{1+r}}+\frac{1-r}{\mu t}\right)=0<(1-r)(1+r)=1-r^2.$$

We conclude, for t large enough $(t \geq t_3 \geq t_2)$, $\frac{rc}{t^r}+\frac{k_0 r^2}{t^{1+r}}+\frac{1-r}{\mu t}\leq 1-r^2$, which gives

$$\frac{rc}{t^r}+\frac{k_0 r^2}{t^{1+r}}+r^2 \leq \frac{1}{\mu}\left(\mu-\frac{1-r}{t}\right).$$

Then multiplying by $\frac{e^t}{t^{1-r}}$, we see that, for $t \geq t_3$,

$$\left(\frac{rc}{t}+\frac{k_0 r^2}{t^2}+\frac{r^2}{t^{1-r}}\right)e^{\mu t} \leq \frac{1}{\mu}\left(\frac{\mu}{t^{1-r}}-\frac{1-r}{t^{2-r}}\right)e^{\mu t}.$$

Return to inequality (3.16), we have, for $t \geq t_3$,

$$\begin{aligned}
\frac{d}{dt}\left[e^{\mu t}\mathscr{E}(t)\right] &\leq \frac{\|z^*\|^2}{2\mu}\left(\frac{\mu}{t^{1-r}}-\frac{1-r}{t^{2-r}}\right)e^{\mu t} \\
&= \frac{\|z^*\|^2}{2\mu}\frac{d}{dt}\left[\frac{e^{\mu t}}{t^{1-r}}\right].
\end{aligned}$$

Integrating the above inequality between t_3 and t and multiplying by $e^{-\mu t}$, we obtain

$$\begin{aligned}
\mathscr{E}(t) &\leq e^{\mu(t_3-t)}\mathscr{E}(t_3)+\frac{\|z^*\|^2}{2\mu}\left(\frac{1}{t^{1-r}}-\frac{e^{\mu(t_3-t)}}{t_3^{1-r}}\right) \\
&\leq \frac{e^{\mu t_3}\mathscr{E}(t_3)}{e^{\mu t}}+\frac{\|z^*\|^2}{2\mu}\frac{1}{t^{1-r}},
\end{aligned}$$

that leads to (3.1), the desired estimate, for $\bar{t}=t_3$. □

We can now state our main convergence result.

Theorem 3.2. *Under the conditions of Theorem 3.1, it has the strong convergence of trajectories $x(t),\lambda(t)$ to the minimum norm solutions x^*,λ^* of primal problem (1.1) and the associated* *dual one. In addition, the following convergence rates hold:*

$$\textbf{a)} \quad \mathscr{L}_t(x(t),\lambda_t) - \mathscr{L}_t(x_t,\lambda_t) = \mathscr{O}\left(\frac{1}{t}\right) \ \textit{as} \ t \to +\infty. \tag{3.17}$$

$$\textbf{b)} \quad \mathscr{L}(x(t),\lambda^*) - \mathscr{L}(x^*,\lambda^*) = \mathscr{O}\left(\frac{1}{t^r}\right) \ \textit{as} \ t \to +\infty. \tag{3.18}$$

$$\textbf{c)} \quad f(x(t)) - \min_C f = \mathscr{O}\left(\frac{1}{t^r}\right) \ \textit{as} \ t \to +\infty; \tag{3.19}$$

$$\textbf{d)} \quad \|Ax(t) - b\| = \mathscr{O}\left(\frac{1}{t^r}\right) \ \textit{as} \ t \to +\infty; \tag{3.20}$$

$$\textbf{e)} \quad \|x(t) - x_t\|^2 = \mathscr{O}\left(\frac{1}{t^{1-r}}\right) \ \textit{as} \ t \to +\infty; \tag{3.21}$$

$$\textbf{f)} \quad \|\lambda(t) - \lambda_t\|^2 = \mathscr{O}\left(\frac{1}{t^{1-r}}\right) \ \textit{as} \ t \to +\infty; \tag{3.22}$$

$$\textbf{g)} \quad \|\dot{x}(t)\| = \mathscr{O}\left(\frac{1}{t^{1-r}}\right) \ \textit{as} \ t \to +\infty. \tag{3.23}$$

Proof. Firstly, we note that (3.1) gives

$$\mathscr{E}(t) = \mathscr{O}\left(\frac{1}{t^{1-r}}\right) \ \text{as} \ t \to +\infty. \tag{3.24}$$

Returning to the expression of $\mathscr{E}(t)$, we conclude (3.17). Using the strong convexity of $\mathscr{L}_t(\cdot,\lambda_t)$ (see Lemma 2.2) and the definition of $\mathscr{E}(t)$, we have, for $t \geq \bar{t}$,

$$\|x(t) - x_t\|^2 \leq \frac{2t^r}{c}\mathscr{L}_t(x(t),\lambda_t) - \mathscr{L}_t(x_t,\lambda_t) \leq \frac{2}{c}\mathscr{E}(t),$$

which ensures (3.21). Combining (3.21) with Lemma 2.3, i.e., the fact that $x_t \to x^*$ as $t \to +\infty$, we deduce the strong convergence of $x(t)$ to x^* as $t \to +\infty$. Also, from the definition of $\mathscr{E}(t)$, we have

$$\|\lambda(t) - \lambda_t\|^2 = \mathscr{O}\left(\frac{1}{t^{1-r}}\right).$$

Since Lemma 2.3 justifies λ_t strongly converges to λ^*, we conclude $\lambda(t)$ also strongly converges to λ^*. Returning to (MIPDALS), we have

$$\dot{\lambda}(t) + c\lambda(t) = t^r(Ax(t) - b) + \frac{t^r}{\tau}A\dot{x}(t),$$

and multiplying by e^{ct}, we obtain

$$\frac{d}{dt}\left(e^{ct}\lambda(t)\right) = e^{ct}\left(\dot{\lambda}(t) + c\lambda(t)\right) = e^{ct}\left(t^r(Ax(t) - b) + \frac{t^r}{\tau}A\dot{x}(t)\right).$$

Integrating from $\bar{t}$ to t and using integration by parts on the last term, we deduce

$$\begin{aligned}\lambda(t)-\frac{\lambda(\bar{t})e^{c\bar{t}}}{e^{ct}} &= \frac{1}{e^{ct}}\int_{\bar{t}}^{t}s^{r}e^{cs}(Ax(s)-b)ds+\frac{1}{\tau e^{ct}}\int_{\bar{t}}^{t}s^{r}e^{cs}d(Ax(s)-b)ds\\ &= \frac{t^{r}(Ax(t)-b)}{\tau}-\frac{\bar{t}^{r}e^{c\bar{t}}(Ax(\bar{t})-b)}{\tau e^{ct}}\\ &\quad+\int_{\bar{t}}^{t}\frac{e^{cs}}{\tau e^{ct}}\left(\tau-\frac{r}{s}-c\right)s^{r}(Ax(s)-b)ds,\end{aligned}$$

Now, relying on the boundedness of $\lambda(t)$, for all $t\geq\bar{t}$, we obtain

$$\left\|t^{r}(Ax(t)-b)+\int_{\bar{t}}^{t}e^{c(s-t)}\left(\tau-\frac{r}{s}-c\right)s^{r}(Ax(s)-b)ds\right\|\leq K_{1},$$

where K_1 is positive constant. Using Lemma 2.1 for

$$g(s):=s^{r}(Ax(s)-b)$$

and

$$a(s):=e^{c(s-t)}(\tau-\frac{r}{s}-c),$$

we obtain

$$\sup_{t\geq\bar{t}}\left\|t^{r}(Ax(t)-b)\right\|<+\infty.$$

Thus we have (3.20). Since (x_t,λ_t) is a saddle point of $\mathscr{L}_t$, we have

$$\mathscr{L}_t(x_t,\lambda_t)\leq\mathscr{L}_t(x^*,\lambda_t).$$

Thus

$$\begin{aligned}\mathscr{L}_t(x(t),\lambda_t)-\mathscr{L}_t(x_t,\lambda_t) &\geq \mathscr{L}_t(x(t),\lambda_t)-\mathscr{L}_t(x^*,\lambda_t)\\ &= \mathscr{L}(x(t),\lambda_t)-\mathscr{L}(x^*,\lambda_t)+\frac{c}{2t^r}(\|x(t)\|^2-\|x^*\|^2)\\ &= \mathscr{L}(x(t),\lambda^*)-\mathscr{L}(x^*,\lambda^*)+\langle\lambda_t-\lambda^*,Ax(t)-b\rangle\\ &\quad+\frac{c}{2t^r}(\|x(t)\|^2-\|x^*\|^2)\\ &\geq \mathscr{L}(x(t),\lambda^*)-\mathscr{L}(x^*,\lambda^*)-\|\lambda_t-\lambda^*\|\|Ax(t)-b\|\\ &\quad+\frac{c}{2t^r}(\|x(t)\|^2-\|x^*\|^2),\end{aligned}$$

which implies

$$\begin{aligned}0 &\leq \mathscr{L}(x(t),\lambda^*)-\mathscr{L}(x^*,\lambda^*)\\ &\leq \mathscr{L}_t(x(t),\lambda_t)-\mathscr{L}_t(x_t,\lambda_t)+\|\lambda_t-\lambda^*\|\|Ax(t)-b\|+\frac{c}{2t^r}\left(\|x^*\|^2-\|x(t)\|^2\right).\end{aligned}$$

Since

$$\lim_{t\to+\infty}\|\lambda_t-\lambda^*\|=\lim_{t\to+\infty}\left(\|x^*\|^2-\|x(t)\|^2\right)=0,$$

(3.17) and (3.20) ensure (3.18). Return to the definition of $\mathscr{E}(t)$, we have

$$\mathscr{E}(t)\geq\frac{1}{2}\|v(t)\|^2=\frac{1}{2}\|\tau(x(t)-x_t)+\dot{x}(t)\|^2.$$

From the definition of $v(t)$, we obtain

$$\begin{aligned} \|\dot{x}(t)\|^2 &= \|v(t)-\tau(x(t)-x_t)\|^2 \leq 2\|v(t)\|^2+2\tau^2\|x(t)-x_t\|^2 \\ &\leq 4\mathscr{E}(t)+2\tau^2\|x(t)-x_t\|^2. \end{aligned}$$

According to (3.21) and (3.24), we deduce that (3.23) is satisfied.

To conclude the rate of values, let us go back to

$$f(x(t))-f(x^*)=\langle\lambda^*,Ax-b\rangle-\left(\mathscr{L}(x(t),\lambda^*)-\mathscr{L}(x^*,\lambda^*)\right)$$

and use (3.18),(3.20). Then

$$f(x(t))-\min_C f=\mathscr{O}\left(\frac{1}{t^r}\right).$$

□

4. Numerical Examples

In this section, we consider three numerical examples to illustrate the evolution of our dynamical system (MIPDALS).

Example 4.1. Consider the constrained minimization problem where the objective function is convex but not strictly convex

$$\min f(x)=\frac{1}{2}(x_1^2+(x_2-x_3)^2) \text{ under constraint: } h(x)=2x_1-x_2+x_3-2=0. \tag{4.1}$$

The set of solutions of (4.1) is $S=\operatorname{argmin}_C f=\{x\in\mathbb{R}^3 : x_1-x_2=1, x_2-x_3=-2\}$ and the element of minimum norm of S is $x^*=(0,-1,1)$.

In this example, by setting $\alpha=5.5$ and $c=\tau=5$ that satisfy condition ($\mathbf{H}_1$), we analyze in Figure 2 the evolution of the convergence rates (3.20), (3.21), and (3.23) demonstrated in Theorem 3.2. We note in Figure 2 top left that the convergence estimate for the values in (3.20) is well suited to this example. Secondly, by positively varying only the parameter c when its values are tolerated by the condition ($\mathbf{H}_1$), we find a slight and inverse evolution for the values $f(x(t))-\min_C f$ and the convergence of $x(t)$ towards x^*. This can be justified by the inequality (3.1), where $(\mathscr{E})(t)$ is increased by $\dfrac{e^{(\alpha-\tau)\bar{t}}\mathscr{E}(\bar{t})}{e^{\mu t}}+\dfrac{\|z^*\|^2}{2\mu}\dfrac{1}{t^{1-r}}$ and the condition ($\mathbf{H}_1$) imposes $\max(1,\tau)\mu<c$.

Example 4.2. Now, we compare the convergence results of our dynamical system (MIPDALS) with those of the very recent paper [30] dealing with the following dynamical system

$$\begin{cases} \ddot{x}(t)+\frac{\alpha}{t}\dot{x}(t)+\nabla f(x(t))+A^*\lambda(t)+\rho A^*(Ax(t)-b)+\varepsilon(t)x(t) &= 0, \\ \dot{\lambda}(t)-t\left[A\left(x(t)+\frac{t}{\alpha-1}\dot{x}(t)\right)-b\right] &= 0. \end{cases} \tag{4.2}$$

We take the same convex constrained minimization problem shown in [30]:

$$\min f(x_1,x_2,x_3)=(5x_1+x_2+x_3)^2 \text{ under constraints } 5x_1-x_2+x_3=0.$$

Here f is a convex differentiable function. The solution set is $S^*=\{u(1,0,-1/5) : u\in\mathbb{R}\}$ and the optimal value is equal to zero. Obviously, the minimizer of minimal norm is the origin of $\mathbb{R}^3$. For this system, we deal with the same data as in this reference: $x(1)=(1,1,1)^T, \lambda(1)=$

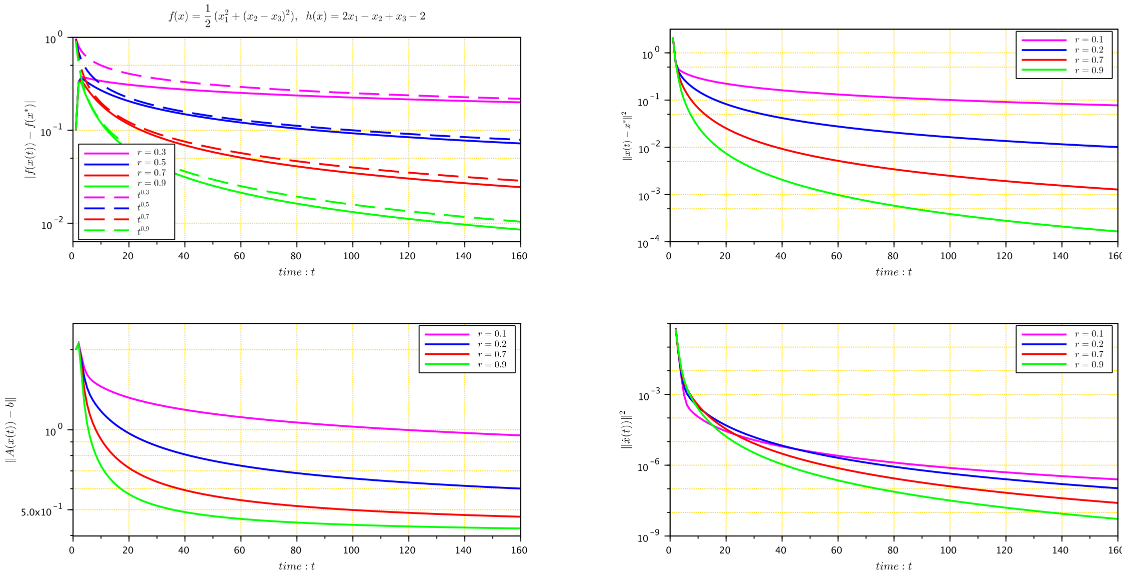

FIGURE 1. Errors of the objective function, the trajectories, the constraint and the velocity of our dynamical system (MIPDALS) with different values of Tikhonov regularization parameters $0 < r < 1$.

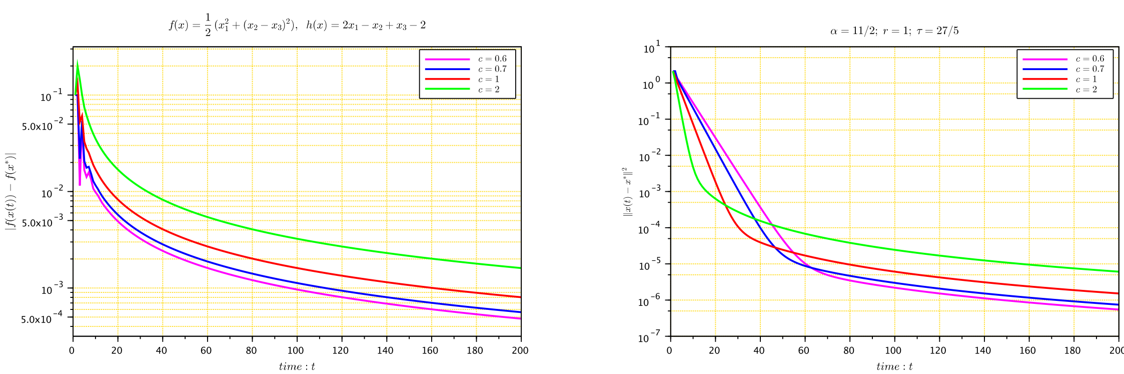

FIGURE 2. Errors of the objective function, the trajectories, the constraint and the velocity of our dynamical system (MIPDALS) with different values of Tikhonov regularization parameters $0 < r < 1$.

$1, \dot{x}(1) = (1,1,1)^T$ and $m = 5, n = 1, e = 1, \alpha = 13, \varepsilon(t) = 3t^{-s}, \rho = 1$. Figure 3 justifies the improvement in the convergence rate of values and solutions for our proposed system when comparing it with that of Zhu et al. [30]. We also note that the values in this reference vary inversely to that of the parameter s in the estimate proposed for the augmented Lagrangian in [30, Theorem 7.4]:

$$\mathscr{L}_\rho(x(t), \lambda^*) - \mathscr{L}_\rho(x^*, \lambda^*) = \mathscr{O}\left(\frac{1}{t^s}\right).$$

5. Conclusion and Perspective

To attain a solution of constrained minimization problems $\min_{Ax=b} f(x)$, where f is a general convex function and A is a linear continuous operator, we proposed the following dynamical

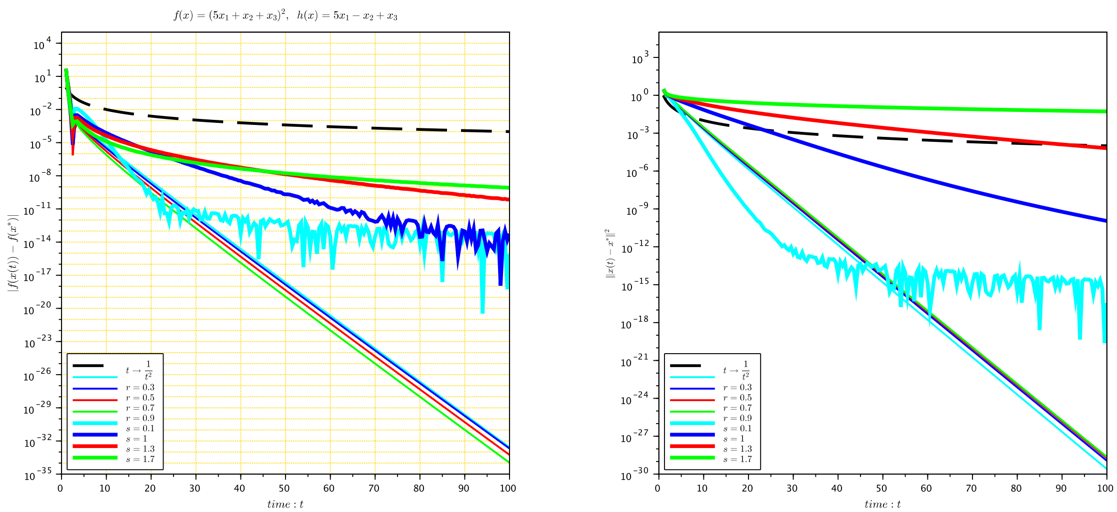

FIGURE 3. Here we compare the convergence rates for different values of $0 < r < 1$ in system (MIPDALS), and those of $0 < s < 2$ in system (4.2).

system

$$\ddot{x}(t) + \alpha\dot{x}(t) + t^r \nabla_x \mathscr{L}_t(x(t), \lambda(t)) = 0, \quad \dot{\lambda}(t) - t^r \nabla_\lambda \mathscr{L}_t\left(x(t) + \frac{1}{\tau}\dot{x}(t), \lambda(t)\right) = 0,$$

where $\mathscr{L}_t(x,\lambda) = \mathscr{L}(x,\lambda) + \frac{c}{2t^r}(\|x\|^2 - \|\lambda\|^2)$ is a quadratic penalty Lagrangian with the penalty parameter function $\varepsilon(t) = \frac{c}{t^r}$.

This allowed us to initiate in this first bibliographic result (see Theorem 3.2) the strong convergence of the solution $(x(t), \lambda(t))$ of the proposed system towards the metric projection of the origin onto the set of solutions of $\min_{Ax=b} f(x)$, as well as a better rate of convergence of the values $f(x(t)) - \min_{Ax=b} f(x)$.

As future works, we are eager to improve the rate of convergence of values firstly by extending the values of the parameter r over the interval $(0,1)$, and therefore for a general time scale parameter $\beta(t)$. Indeed, as one Reviewer recommended in his report, the case $\beta(t) = t^m e^{\gamma t^q}$ with $(q,m) \in \mathbb{R}^2_+ \setminus \{(0,0)\}, 0 < q < 1, \gamma > 0$, is more attractive since this convergence is faster than that proposed in (3.19) and (3.23). As it is already mentioned in the introduction of this paper, the references [13, 14] were able to reach for this choice of β, in addition to the strong convergence, the rates (1.3):

$$f(x(t)) - \min_{\mathscr{X}} f = \mathscr{O}\left(\frac{1}{t^m e^{\gamma t^q}}\right) \quad \text{and} \quad \|\dot{x}(t)\|^2 = \mathscr{O}\left(\frac{1}{t^{1-q}}\right).$$

Previously, we attempted to adapt them to our case under constraints but without success. So this may be one of our future research topics. This work also provides a basis for the development of corresponding algorithmic results.

REFERENCES

[1] H. Attouch, A. Balhag, Z. Chbani, H. Riahi, Fast convex optimization via inertial dynamics combining viscous and Hessian-driven damping with time rescaling, Evol. Equ. Control Theory, 11 (2022), 487-514. https://doi.org/10.3934/eect.2021010

[2] H. Attouch, A. Balhag, Z. Chbani, H. Riahi, Accelerated Gradient Methods Combining Tikhonov Regularization with Geometric Damping Driven by the Hessian, Appl. Math. Optim. 88 (2023), 29-58. https://doi.org/10.1007/s00245-023-09997-x

[3] A.C. Bagy, Z. Chbani, H. Riahi, The Heavy ball method regularized by Tikhonov term. Simultaneous convergence of values and trajectories, Evol. Equ. Control Theory, 12 (2023), 687-702. https://doi.org/10.3934/eect.2022046

[4] H. Attouch, Z. Chbani, J. Fadili, H. Riahi, Fast convergence of dynamical ADMM via time scaling of damped inertial dynamics, J. Optim. Theory Appl. 193 (2022), 704-736. https://doi.org/10.1007/s10957-021-01859-2

[5] A.C. Bagy, Z. Chbani, H. Riahi, Fast convergence rate of values with strong convergence of trajectories via inertial dynamics with Tikhonov regularization terms and asymptotically vanishing damping, J. Nonlinear Var. Anal. 8 (2024), 691-715. https://doi.org/10.23952/jnva.8.2024.5.01

[6] D.A.Hulett, D-K. Nguyen, Time rescaling of a primal-dual dynamical system with asymptotically vanishing damping, Appl. Math. Optim. 88 (2023), 27. https://doi.org/10.1007/s00245-023-09999-9

[7] B. Poljak, Some methods of speeding up the convergence of iteration methods, USSR Comput. Math. Math. Phys. 4 (1964), 791-803.

[8] B. Polyak, Introduction to Optimization, Optimization Software-Inc, New York, 1987.

[9] W. Siegel, Accelerated first-order methods: Differential equations and Lyapunov functions, https://arxiv.org/abs/1903.05671v1 [math.OC], 2019.

[10] H. Attouch, Z. Chbani, J. Fadili, H. Riahi, First-order optimization algorithms via inertial systems with Hessian driven damping, Math. Program. Ser. A. 193 (2022), 113-155. https://doi.org/10.1007/s10107-020-01591-1

[11] H. Attouch, Z. Chbani, J. Fadili, H. Riahi, Convergence of iterates for first-order optimization algorithms with inertia and Hessian driven damping, Optimization, 72 (2023), 1199-1238. https://doi.org/10.1080/02331934.2021.2009828

[12] J. B. Baillon, Un exemple concernant le comportement asymptotique de la solution du problème $du/dt + \partial f(u) \ni 0$, J. Differential Equations, 259 (2015), 3115-3143. https://doi.org/10.1016/0022-1236(78)90093-9

[13] H. Attouch, A. Balhag, Z. Chbani, H. Riahi, Damped inertial dynamics with vanishing Tikhonov regularization: Strong asymptotic convergence towards the minimum norm solution, J. Differential Equations, 311 (2022), 29-58. https://doi.org/10.1016/j.jde.2021.12.005

[14] A.C. Bagy, Z. Chbani, H. Riahi, Strong convergence of trajectories via inertial dynamics combining Hessian-driven damping and tikhonov regularization for general convex minimizations, Numer. Funct. Anal. Optim. 44 (2023), 1481-1509. https://doi.org/10.3934/eect.2022046

[15] X. Zeng, J. Lei, J. Chen, Dynamical primal-dual accelerated method with applications to network optimization, IEEE Trans. Autom. Control, 68 (2022), 1760-1767. https://doi.org/10.1109/TAC.2022.3152720

[16] R.I., Boţ, D.K. Nguyen, Improved convergence rates and trajectory convergence for primal-dual dynamical systems with vanishing damping, J. Differential Equations, 303 (2021), 369-406. https://doi.org/10.1016/j.jde.2021.09.021

[17] X. He, R. Hu, Y.P. Fang, Convergence rates of inertial primal-dual dynamical methods for separable convex optimization problems, SIAM J. Control Optim. 59 (2021), 3278-3301. https://doi.org/10.1137/20M1355379

[18] 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. https://doi.org/10.1007/s10107-016-0992-8

[19] H. Attouch, Z. Chbani, H. Riahi, Fast proximal methods via time scaling of damped inertial dynamics, SIAM J. Optim. 29 (2019), 2227-2256. https://doi.org/10.1137/18M1230207

[20] R. May, Asymptotic for a second-order evolution equation with convex potential and vanishing damping term, Turk. J. Math. 41 (2017), 681-685. https://doi.org/10.3906/mat-1512-28

[21] W. Su, S. Boyd, E.J. Candès, A differential equation for modeling Nesterov's accelerated gradient method: theory and insights, J. Mach. Learn. Res. 17 (2016), 1-43. https://jmlr.org/papers/v17/15-084.html

[22] H. Attouch, Z. Chbani, H. Riahi, Fast convex optimization via a third-order in time evolution equation, Optimization, 71 (2022), 1275-1304. https://doi.org/10.1080/02331934.2020.1764953

[23] H. Attouch, Z. Chbani, H. Riahi, Fast convex optimization via time scaling of damped inertial gradient dynamics, Pure Appl. Funct. Anal. 6 (2021), 1081-1117.

[24] A.C. Bagy, Z. Chbani, H. Riahi, On the strong convergence of an inertial proximal algorithm with a time scale, Hessian-driven damping, and a Tikhonov regularization term, J. Appl. Numer. Optim. 6 (2024), 249-269. https://doi.org/10.23952/jano.6.2024.2.06

[25] X. He, R. Hu, Y.-P. Fang, "Second-order primal" + "first-order dual" dynamical systems with time scaling for linear equality constrained convex optimization problems, IEEE Trans. Automat. Control 67 (2022), 4377-4383. https://doi.org/10.1109/TAC.2022.3176527

[26] X. He, R. Hu, Y.-P. Fang, Fast primal-dual algorithm via dynamical system for a linearly constrained convex optimization problem, Automatica, 146 (2022), 110547. https://doi.org/10.1016/j.automatica.2022.110547

[27] C. Zălinescu, Convex Analysis in General Vector Spaces, World Scientific Publishing Company, 2002. https://doi.org/10.1142/9789812777096

[28] H. H. Bauschke, P. L. Combettes, Convex Analysis and Monotone Operator Theory in Hilbert Spaces, CMS Books in Mathematics, Springer, 2011. https://doi.org/10.1007/978-1-4419-9467-7

[29] H. Brézis, Opérateurs Maximaux Monotones dans les Espaces de Hilbert et équations D'évolution, Lecture Notes, vol.5, North Holland, 1973.

[30] T.T. Zhu, R. Hu, Y. P. Fang, Tikhonov regularized second-order plus first-order primal-dual dynamical systems with asymptotically vanishing damping for linear equality constrained convex optimization problems, Optimization, 2024. https://doi.org/10.1080/02331934.2024.2407515