Accelerated Primal-Dual Projection Neurodynamic Approach With Time Scaling for Linear and Set Constrained Convex Optimization Problems

You Zhao, Xing He, Mingliang Zhou, and Tingwen Huang, *Fellow, IEEE*

***Abstract*—The Nesterov accelerated dynamical approach serves as an essential tool for addressing convex optimization problems with accelerated convergence rates. Most previous studies in this field have primarily concentrated on unconstrained smooth convex optimization problems. In this paper, on the basis of primal-dual dynamical approach, Nesterov accelerated dynamical approach, projection operator and directional gradient, we present two accelerated primal-dual projection neurodynamic approaches with time scaling to address convex optimization problems with smooth and nonsmooth objective functions subject to linear and set constraints, which consist of a second-order ODE (ordinary differential equation) or differential conclusion system for the primal variables and a first-order ODE for the dual variables. By satisfying specific conditions for time scaling, we demonstrate that the proposed approaches have a faster convergence rate. This only requires assuming convexity of the objective function. We validate the effectiveness of our proposed two accelerated primal-dual projection neurodynamic approaches through numerical experiments.**

***Index Terms*—Accelerated projection neurodynamic approach, linear and set constraints, projection operators, smooth and nonsmooth convex optimization, time scaling.**

I. Introduction

DUE to its low complexity and high efficiency, the Nesterov accelerated gradient algorithm [1] has become a popular methodology for tackling large-scale convex optimization problems. However, the acceleration phenomenon of the Nesterov accelerated algorithm still maintains somewhat mysterious. Su *et al.* [2] firstly reveal that a second-order ordinary differential equation (ODE) with vanishing damping, known as the Nesterov accelerated dynamical approach, represents the continuous limit of the Nesterov accelerated gradient algorithm. Building on the work in [2], numerous accelerated dynamical approaches have been presented to address unconstrained convex optimization problems with an accelerated convergence rate of $O(1/t^2)$. Attouch and Chbani [3] extended the results in [2] by combining Hessian driven damping, small perturbation [4], Tikhonov regularization [5] and maximally monotone operators [6], and they achieved an accelerated convergence rate $O(1/t^2)$. Drawing inspiration from the Bregman-Lagrangian function, Wibisono *et al.* [7] and Wilson *et al.* [8] devised second-order ordinary differential equations (ODEs) that give rise to the Nesterov accelerated dynamical approach described in [2]. Alimisis *et al.* [9] introduced an accelerated gradient-based dynamical approach on a Riemannian manifold, drawing inspiration from the work of Wibisono *et al.* [7]. Vassilis *et al.* [10] investigated the accelerated convergence properties of Nesterov accelerated dynamical approach [2] by using the differential inclusion system for solving nonsmooth and convex optimization problems. In addition, in order to improve accelerated convergence rate of $O(1/t^2)$, Attouch *et al.* [11] proposed a Nesterov accelerated dynamical approach with time scaling term $\beta(t)$, resulting in a faster convergence rate of $O(1/(t^2\beta(t)))$.

It was previously mentioned that the Nesterov accelerated dynamical approaches and their variants are only suitable for unconstrained convex optimization problems. Neurodynamic approaches provide a methodology for tackling constrained convex optimization problems from a continuous-time (dynamical system) perspective. It can be mathematically formulated as an ordinary differential equation or differential inclusion system, and can be implemented with specialized hardware, providing further insights into classical numerical algorithms.

Since the 1980s, when Hopfield first proposed the Hopfield neural network [12] as a method of solving the traveling salesman problem, a variety of neurodynamic approaches have been evolved to tackle a wide range of optimization problems. Kennedy and Chua [13] utilized a penalty method to design a nonlinear programming circuit for dealing with convex constraint optimization problems. Xia and Feng [14] proposed an array of projection neurodynamic approaches intended to solve monotone variational inequalities and constrained nonlinear convex programming problems [15], etc. Hu and Wang [16] generalized the PNN to address the quasi-convex opti-

Manuscript received January 22, 2024; revised February 24, 2024; accepted March 3, 2024. This work was supported by the National Natural Science Foundation of China (62176218, 62176027), the Fundamental Research Funds for the Central Universities (XDJK2020TY003), and the Funds for Chongqing Talent Plan (cstc2024ycjh-bgzxm0082). Recommended by Associate Editor Hongyi Li. *(Corresponding author: Xing He.)*

Citation: Y. Zhao, X. He, M. Zhou, and T. Huang, "Accelerated primal-dual projection neurodynamic approach with time scaling for linear and set constrained convex optimization problems," *IEEE/CAA J. Autom. Sinica*, vol. 11, no. 6, pp. 1485–1498, Jun. 2024.

Y. Zhao and X. He are with Chongqing Key Laboratory of Nonlinear Circuits and Intelligent Information Processing, College of Electronic and Information Engineering, Southwest University, Chongqing 400715, China (e-mail: zhaoyou1991sdtz@cqu.edu.cn; hexingdoc@swu.edu.cn).

M. Zhou is with the School of Computer Science, Chongqing University, Chongqing 400044, China (e-mail: mingliangzhou@cqu.edu.cn).

T. Huang is with the Department of Mathematics, Texas A&M University at Qatar, Doha 23874, Qatar (e-mail: tingwen.huang@qatar.tamu.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JAS.2024.124380

mization problems with nonlinear constraints. He *et al.* [17] proposed an inertial projection neural network for handling constrained nonconvex optimization problems and demonstrated its ability to capture different Karush-Kuhn-Tucker (KKT) points by adjusting the inertial parameters. Additionally, the primal-dual dynamical approach, also known as the Lagrange neural network, provides a general framework for solving various convex optimization problems with linear constraints. Zhang and Constantinides [18] were the first to investigate the use of a Lagrange programming neural network (LPNN) for addressing constrained convex optimization problems. Since then, numerous neurodynamic approaches based on LPNN have been studied not only for dealing with constrained convex optimization problems [19], [20] but also for exploring new applications [21]–[27].

However, the majority of existing research on projection neurodynamic or primal-dual neurodynamic approaches only provides asymptotic convergence properties or a slow convergence rate of $O(1/t)$ for constrained convex optimization problems without strongly convex assumption. In [28], Krichene *et al.* extended the work in [2] with non-Euclidean geometries by introducing mirror operators, proposed an accelerated mirror dynamical approach $O(1/t^2)$ convergence rate for convex optimization problems that contains set constraints. Zhao *et al.* [29] were influenced by the Nesterov accelerated dynamical approach and projection operators. They introduced an accelerated projection neurodynamic approach specifically designed for smooth convex optimization problems with set constraints. Remarkably, this approach demonstrates a convergence rate of $O(1/t^2)$. For smooth convex optimization problems with linear constraints, Zeng *et al.* [30] proposed a dynamical primal-dual approach based on Nesterov accelerated dynamical approach, and proved the primal-dual gap of objective functions has a convergence rate $O(1/t^2)$. He *et al.* [31] further considered an accelerated primal-dual dynamical approach with added perturbation for separable convex optimization problems and obtained some convergence properties similar to work in [30]. Boţ and Nguyen [32] made significant advancements in improving the convergence rate discussed in the work of Zeng *et al*. [30]. Additionally, they presented a weak convergence result for the solutions of the primal-dual problem. Attouch *et al.* [33] studied a second-order primal-dual dynamical approach involving damped inertial and time scaling for solving separable convex optimization problems that have affine constraint, and obtained some of the same results as [11]. He *et al.* [34] designed a "second-order primal"+"first-order dual" dynamical approach with time scaling to address convex optimization problems only with linear constraints, and obtained some results identical to those in [11]. Moreover, Zhao *et al.* [35] recently proposed a second-order primal-dual mirror dynamical approach for solving smooth and nonsmooth convex optimization problems that have affine and set constraints, with a convergence rate of $O(1/t^2)$.

It is worth noting that the aforementioned accelerated projection or primal-dual neurodynamic approaches do not have the capability to handle smooth convex optimization problems with both linear and set constraints while maintaining accelerated convergence properties. However, it is inevitable for real-world engineering problems to have simultaneous set and affine constraints. For example, in resource allocation problems, resource constraints, physical constraints, or other business rules lead to both set and linear constraints when modeling the problem, and linear and set constraints as well as multi-agents consensus constraints exist in distributed sparse signal reconstruction. Moreover, in image classification, each image may belong to multiple categories, which requires representing categories as a set. At the same time, the boundaries of the classifier can be defined by affine constraints to ensure correct classification of data points. In logistics planning, considerations include the cargo loading capacity, warehouse storage capacity, as well as the capacity and route constraints of transportation vehicles. These constraints can be described using set constraints and affine constraints to ensure the feasibility and efficiency of logistics planning. In wireless communication networks, factors such as bandwidth limitations of wireless channels, power constraints, and the capacity of user devices need to be considered. These constraints can be modeled using set constraints and affine constraints to optimize the allocation of network resources and communication quality, and so on. In addition, the accelerated neurodynamic approaches mentioned above are concerned with solving smooth convex constrained optimization problems, and do not involve nonsmooth convex constrained optimization problems, which limits the applicability of them to a certain extent, because the sparse signal reconstruction problem, the L_1-regularization problems, etc., are all nonsmooth constrained optimization problems in practice. With the above considerations in mind, we are motivated by studying fast primal-dual neurodynamic approaches based on Nesterov accelerated dynamical approach, primal-dual dynamical approach and time scaling item to tackle convex optimization problems with smooth and nonsmooth objective functions, subject to linear and set constraints, without strongly convex assumption, to obtain some results that are similar with [34]. In our opinion, there are mainly three difficulties for designing accelerated primal-dual projection neurodynamic approaches. The first challenge lies in designing a projection scheme that is compatible with the accelerated primal-dual neurodynamic approaches. The existing projection scheme based on the classical Brouwer's fixed point theorem fails to achieve effective acceleration. The second obstacle involves developing new Lyapunov functions to analyze the accelerated convergence properties of the proposed fast primal-dual neurodynamic approaches. The Lyapunov functions used in previous works [30]–[34] are no longer applicable as they do not incorporate projection operators. Lastly, extending the fast primal-dual neurodynamic approach from the smooth case to the nonsmooth case while maintaining the same accelerated convergence rate poses a significant difficulty and challenge.

The major contributions of this paper are highlighted below:

1) For convex optimization problems with smooth objective function, subject to linear and set constraints, we first propose an accelerated primal-dual projection neurodynamic

approach (APDPNA-S) by using Nesterov accelerated dynamical approach with time scaling term, projection operators and primal-dual dynamical approach. By using the Cauchy-Lipschitz-Picard theorem, proof by contradiction and properties of projection operators, we show the existence, uniqueness and viability of the strong global solution of APDPNA-S. Moreover, by designing a new Lyapunov function, we prove that APDPNA-S has a fast convergence rate $O(1/(t^2\beta(t)))$.

2) Compared to existing results in [31] and [33], APDPNA-S has a simpler structure since the updating of the dual variables in APDPNA-S is a first-order ODE, and faster convergence rate than $O(1/t^2)$ in [31]. Moreover, compared with the existing inertial primal-dual dynamical approaches [30]–[34], our proposed APDPNA-S can address convex optimization problems that contain both linear and set constraints, which means our proposed APDPNA-S has a wider applicability.

3) We extend APDPNA-S into a differential inclusion dynamical approach (named as APDPNA-NS) by employing directional derivative in place of exact gradients in APDPNA-S. By computing difference quotient of Lyapunov functions, we prove that the APDPNA-NS have same results as APDPNA-S.

The organization of this paper is summarized as follows. In Section II, several necessary preliminaries are introduced briefly. In Section III, two accelerated primal-dual projection neurodynamic approaches, i.e., APDPNA-S and APDPNA-NS are proposed to deal with smooth and nonsmooth convex optimization problems with linear and set constraints, and the fast convergence properties of them are also discussed. In Section IV, we validate the effectiveness of our proposed two neurodynamic approaches through numerical simulations. We conclude this paper in Section V.

II. Preliminaries

A. Subdifferential

Definition 1 [36]: If $g:\mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ satisfies

$$g(\varpi w+(1-\varpi)v) \geq \varpi g(w)+(1-\varpi)g(v),\ \ 0<\varpi<1 \tag{1}$$

then, we called g is convex.

When the convex function g is differentiable (i.e., smooth), then, it enjoys

$$g(v) \geq g(w)+\nabla g(w)^T(v-w),\ \forall w,\ v\in\mathbb{R}^n \tag{2}$$

where the $\nabla g(w)$ is the gradient of g with respect of w. In addition, if g is convex and nondifferentiable (i.e., nonsmooth), then, it fulfills

$$g(v) \geq g(w)+h^T(v-w),\ \forall w,\ v\in\mathbb{R}^n \tag{3}$$

where the h is a subgradient of g at the point w. Thus, the set of

$$\partial g(w)=\left\{h\in\mathbb{R}^n | g(v)-g(w)\geq h^T(v-w),\ \forall w,\ v\in\mathbb{R}^n\right\} \tag{4}$$

is called the subdifferential of g at w.

B. Projection Operators

Definition 2 [37]: The projection operator $P_\Omega(w)$ of Ω with respect to the variable w for a nonempty, closed and convex set Ω is given by

$$P_\Omega(w)=\arg\min_{u\in\Omega}\|w-u\|=(I+\mathcal{N}_\Omega)^{-1}w \tag{5}$$

where $\mathcal{N}_\Omega(w)=\{u\in\mathbb{R}^n|\langle u,v-w\rangle\leq 0,\forall v\in\Omega\}$ is the normal cone to Ω at point w.

Lemma 1: In general, a closed-form solution for the projection operator is not always available. However, there are cases where the projection operator P_Ω can be expressed in a closed-form when Ω satisfies specific structures. For example:

i) When Ω is a box set, i.e., $\Omega=\{w\in\mathbb{R}^n \mid w_i^{\min}\leq w_i\leq w_i^{\max}, i=1,\ldots,n\}$, its projection operator is

$$P_\Omega(w_i)=\min\left\{\max\left\{w_i,w_i^{\min}\right\},w_i^{\max}\right\}.$$

ii) When Ω is a Euclidean ball set, i.e., $\Omega=\{w\in\mathbb{R}^n \mid \|w-v\|\leq r,\ v\in\mathbb{R}^n,\ r>0\}$, then

$$P_\Omega(w)=\begin{cases} w, & \|w-v\|\leq r\\ v+\dfrac{r(w-v)}{\|w-v\|}, & \|w-v\|>r.\end{cases}$$

iii) When Ω is an affine set, i.e., $\Omega=\{w\in\mathbb{R}^n \mid Bw=c, B\in\mathbb{R}^{m\times n}, \operatorname{rank}(B)=m, m\leq n\}$, then

$$P_\Omega(w)=w+B^T\left(BB^T\right)^{-1}(c-Bw). \tag{6}$$

More information on projection operators that have closed form solutions can be found in [37].

Lemma 2 [14]: The projection operator satisfies the following inequalities when Ω is a nonempty, closed and convex set (see Fig. 1):

$$\langle w-P_\Omega(w),v-P_\Omega(w)\rangle\leq 0,\forall w\in\mathbb{R}^n,v\in\Omega$$
$$\|P_\Omega(w)-P_\Omega(u)\|\leq\|w-u\|,\forall w,u\in\mathbb{R}^n. \tag{7}$$

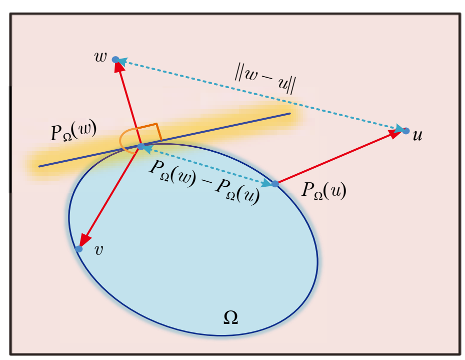


Fig. 1. The presentation diagram of the projection operators inequalities in Lemma 2.

Lemma 3 [29]: Let $\varphi(w,v):\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ be

$$\varphi(w,v)=\frac{1}{2}\left(\|w-P_\Omega(v)\|^2-\|w-P_\Omega(w)\|^2\right) \tag{8}$$

then, one has

i) $\varphi(w,v)\geq\frac{1}{2}\|P_\Omega(w)-P_\Omega(v)\|^2$.

ii) $\varphi(w,v)$ is continuously differentiable with respect of the variable w, and its gradient is $\nabla_w\varphi(w,v)=P_\Omega(w)-P_\Omega(v)$.

III. Accelerated Primal-Dual Projection Neurodynamic Approaches With Time Scaling

Consider a convex optimization problem with linear and set constraints as

$$\min_{\xi\in\mathbb{R}^n} g(\xi)$$
$$\text{s.t. } B\xi = c,\ \xi\in\Omega \tag{9}$$

where $B\in\mathbb{R}^{m\times n}$, $c\in\mathbb{R}^m$, $\Omega\subseteq\mathbb{R}^n$ is a nonempty, closed and convex set, $g:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$ is closed, proper lower semi-continuous convex function (not necessarily differentiable).

The problem (9) serves as a fundamental model in various significant applications, including signal and image processing problems [38], [39], resource allocation problems [40]–[42], machine learning problems [43], and distributed constrained convex optimization problems [44], etc.

Define the augmented Lagrangian function $L_\mu(\xi,\zeta)$ associated with (9) is

$$L_\mu(\xi,\zeta) = g(\xi)+\zeta^T(B\xi-c)+\frac{\mu}{2}\|B\xi-c\|^2 \tag{10}$$

where $\zeta\in\mathbb{R}^m$ is the Lagrange multiplier, $\mu\geq 0$ and $\|B\xi-c\|^2$ is the augmented item. According to convex optimization theory, ξ^* is the optimal solution of problem (9) if and only if there exists $\zeta^*\in\mathbb{R}^m$ such that $(\xi^*,\zeta^*)\in\Omega\times\mathbb{R}^m$ forms a saddle point, satisfying the following inequality:

$$L_\mu(\xi^*,\zeta)\leq L_\mu(\xi^*,\zeta^*)\leq L_\mu(\xi,\zeta^*)$$
$$\forall(\xi,\zeta)\in\Omega\times\mathbb{R}^m. \tag{11}$$

A. The Problem (9) With Smooth Convex Objective Function g

In order to address problem (9) with a smooth convex objective function g and achieve a rapid convergence rate, we introduce an accelerated primal-dual projection neurodynamic approach, denoted as APDPNA-S. This approach integrates the Nesterov accelerated dynamical approach, time scaling element, primal-dual dynamical approach, and projection operator.

$$\begin{cases}\dot{\xi}(t) = \dfrac{\alpha}{t}\left(P_\Omega(y(t))-\xi(t)\right)\\ \dot{y}(t) = -\dfrac{t}{\alpha}\beta(t)\Big(\nabla g(\xi(t))+\mu B^T(B\xi(t)-c)\\ \qquad +B^T\zeta(t)+y(t)-P_\Omega(y(t))\Big)-\dot{\xi}(t)\\ \dot{\zeta}(t) = t\beta(t)\left(BP_\Omega(y(t))-c\right)\\ \xi_0\in\Omega, t\geq t_0>0\end{cases} \tag{12}$$

where $\alpha\geq 2$, $\beta(t):[t_0,+\infty)\to(0,+\infty)$ and it satisfies $\dot{\beta}(t)\leq\frac{(\alpha-2)}{t}\beta(t)$.

Theorem 1: For any initial value $(\xi_0,y_0,\zeta_0)\in\Omega\times\mathbb{R}^n\times\mathbb{R}^m$, $(\bar{\xi},\bar{y},\bar{\zeta})$ is an equilibrium point to APDPNA-S (12) if and only if $(\bar{\xi},\bar{y},\bar{\zeta})=(\xi^*,y^*,\zeta^*)$ satisfies KKT condition, i.e., $\xi^*=\bar{\xi}$ is the optimal solution of the problem (9) with smooth convex objective function g.

Proof: 1) Sufficiency: In accordance with KKT conditions, if ξ^* is the optimal solution of the problem (9) with smooth convex objective function g, then there exists ζ^* and $\xi^* = P_\Omega(y^*)$ satisfying

$$\nabla g(\xi^*)+B^T\zeta^*+\mathcal{N}_\Omega(\xi^*)\ni 0 \tag{13a}$$

$$B\xi^*-c=0. \tag{13b}$$

By combining $\mu\geq 0$ with (13b), (13a) can be rewritten as $\nabla g(\xi^*)+\mu B^T(B\xi^*-c)+B^T\zeta^*+\mathcal{N}_\Omega(\xi^*)\ni 0$. With the help of Definition 2 and projection inequalities in Lemma 2, we can obtain that if $\xi^*=P_\Omega(y^*)$, one has $\langle y^*-P_\Omega(y^*), v-P_\Omega(y^*)\rangle\geq 0, \forall v\in\Omega$, which means $y^*-P_\Omega(y^*)\in\mathcal{N}_\Omega(\xi^*)$. Thus, (13a) can be rewritten as

$$0=\nabla g(\xi^*)+\mu B^T(B\xi^*-c)+B^T\zeta^*+y^*-P_\Omega(y^*)$$
$$\xi^*=P_\Omega(y^*). \tag{14}$$

By combining (13b) and (14), we can obtain that

$$\begin{cases}\xi^*=P_\Omega(y^*)\\ 0=\nabla g(\xi^*)+\mu B^T(B\xi^*-c)+B^T\zeta^*+y^*-P_\Omega(y^*)\\ B\xi^*-c=0.\end{cases} \tag{15}$$

Therefore, $(\xi^*,\zeta^*)\in\Omega\times\mathbb{R}^m$ is an equilibrium point of APDPNA-S (12).

2) Necessity: When $\left(\bar{\xi},\bar{y},\bar{\zeta}\right)\in\Omega\times\mathbb{R}^n\times\mathbb{R}^m$ is an equilibrium point of APDPNA-S (12), then it satisfies

$$\begin{cases}\bar{\xi}=P_\Omega(\bar{y})\\ 0=\nabla g\left(\bar{\xi}\right)+\mu B^T\left(B\bar{\xi}-c\right)+B^T\bar{\zeta}+\bar{y}-P_\Omega(\bar{y})\\ B\bar{\xi}-c=0.\end{cases} \tag{16}$$

From $\bar{\xi}=P_\Omega(\bar{y})$, one can obtain $\bar{y}\in\bar{\xi}+\mathcal{N}_\Omega(\bar{\xi})$, and further get $\bar{y}-P_\Omega(\bar{y})\in\mathcal{N}_\Omega(\bar{\xi})$. Therefore, $0=\nabla g(\bar{\xi})+\mu B^T(B\bar{\xi}-c)+B^T\bar{\zeta}+\bar{y}-P_\Omega(\bar{y})$ can be equivalently written as $0\in\nabla g(\bar{\xi})+\mu B^T(B\bar{\xi}-c)+B^T\bar{\zeta}+\mathcal{N}_\Omega(\bar{\xi})$. It combines $B\bar{\xi}-c=0$ to obtain that $(\bar{\xi},\bar{\zeta})$ satisfies KKT condition to the problem (9) with smooth convex objective function g. ■

Definition 3 [45]: The $Z=(\xi,y,\zeta):[t_0,+\infty)\to\Omega\times\mathbb{R}^n\times\mathbb{R}^m$ is called a strong global solution of APDPNA-S (12) if the following conditions hold:

i) $Z(t)$ is locally absolutely continuous;

ii) $Z(t_0)=(\xi(t_0),y(t_0),\zeta(t_0))\in\Omega\times\mathbb{R}^n\times\mathbb{R}^m$;

iii) The APDPNA-S (12) holds with $Z(t_0)$ and for $t\in[t_0,+\infty)$.

Remark 1: A function $x:[t_0,+\infty)\to\mathbb{R}^n$ is said to be locally absolutely continuous, if it is absolutely continuous on every interval $[t_0,T]$, $T>0$, i.e., the following equivalent properties hold [46]:

a) There exists an integrable function $\aleph:[t_0,T]\to\mathbb{R}^n$ such that

$$x(t)=x(t_0)+\int_{t_0}^t\aleph(s)ds,\ \forall t\in[t_0,T]. \tag{17}$$

b) x is continuous and its distributional derivative is Lebesgue integrable on $[t_0,T]$.

c) For every $\varepsilon>0$, there exists $\gamma>0$ such that for any finite family of intervals $I_k=(a_k,b_k)\subseteq[t_0,T]$, one has

$$\left(I_k \cap I_j = \emptyset \text{ and } \sum_k |b_k - a_k| < \gamma\right) \Rightarrow \sum_k \|x(b_k) - x(a_k)\| < \varepsilon. \tag{18}$$

i) From (17), we can obtain that an absolutely continuous function is differentiable almost everywhere, its derivative coincides with its distributional derivative almost everywhere and one can recover the function from its derivative $\dot{x} = \aleph$. In addition, an absolutely continuous function with non-positive derivative, i.e., (a_k, b_k) is nonincreasing.

ii) If $x : [t_0, T] \to \mathbb{R}^n$ with $T > 0$ is absolutely continuous and $\mathfrak{B} : \mathbb{R}^n \to \mathbb{R}^n$ is $\mathfrak{y}$-Lipschitz continuous for $\mathfrak{y} \geq 0$, then the function $z = \mathfrak{B} \circ x$ is absolutely continuous, too. This is easily evident by using the characterization of absolute continuity in third equivalent definition mentioned above. Moreover, z is differentiable almost everywhere on $[t_0, T]$ and the inequality $\|\dot{z}(t)\| \leq \mathfrak{y}\|\dot{x}(t)\|$ holds for any $t \in [t_0, T]$.

Theorem 2: For any initial value $(\xi_0, y_0, \zeta_0) \in \Omega \times \mathbb{R}^n \times \mathbb{R}^m$, APDPNA-S (12) has a unique strong global solution. In addition, the solution of APDPNA-S (12) is viable, i.e., $\xi(t) \in \Omega$, $\forall t \geq t_0 > 0$.

Proof: Existence and uniqueness: Let $Z(t) = (\xi(t), y(t), \zeta(t)) \in \Omega \times \mathbb{R}^n \times \mathbb{R}^m$, then the APDPNA-S (12) can be rephrased as follows:

$$\begin{cases} \dot{Z}(t) = F(t, Z(t)) \\ Z(t_0) = (\xi(t_0), y(t_0), \zeta(t_0)) \end{cases} \tag{19}$$

where $F : [t_0, +\infty) \times \Omega \times \mathbb{R}^n \times \mathbb{R}^m$ and

$$\begin{aligned} F(t, Z(t)) = \Big(&\frac{\alpha}{t}(P_\Omega(y(t)) - \xi(t)), \\ &-\frac{t}{\alpha}\beta(t)\big(\nabla g(\xi(t)) + \mu B^T(B\xi(t) - c) \\ &+ B^T\zeta(t) + y(t) - P_\Omega(y(t))\big) - \dot{\xi}(t), \\ &t\beta(t)(BP_\Omega(y(t)) - c)\Big). \end{aligned}$$

The APDPNA-S (12) possesses a unique strong global solution, as guaranteed by the Cauchy-Lipschitz-Picard theorem, under the following two conditions:

i) $F(t, \cdot)$ is $l(t)$-Lipschitz continuous and $l(\cdot) \in \mathbb{L}^1_{\text{loc}}([t_0, +\infty), \Omega \times \mathbb{R}^n \times \mathbb{R}^m)$ for any $t \in [t_0, +\infty)$.

ii) For any $Z \in \Omega \times \mathbb{R}^n \times \mathbb{R}^m$, we have $F(\cdot, Z) \in \mathbb{L}^1_{\text{loc}}([t_0, +\infty), \Omega \times \mathbb{R}^n \times \mathbb{R}^m)$ for any $t \in [t_0, +\infty)$.

For i), let $t \in [t_0, +\infty)$ be fixed and use the Lipschitz properties of ∇g and P_Ω (i.e., they have $\mathfrak{l}_g$ and 1 Lipschitz constants, respectively) and inequality $\|\mathsf{X}_1 + \mathsf{X}_2\|^2 \leq 2\|\mathsf{X}_1\|^2 + 2\|\mathsf{X}_2\|^2$, $\forall\, \mathsf{X}_1, \mathsf{X}_2 \in \mathbb{R}^n$ then, for $Z, \hat{Z} \in \Omega \times \mathbb{R}^n \times \mathbb{R}^m$, one has

$$\begin{aligned} &\left\|F(t, Z(t)) - F\left(t, \hat{Z}(t)\right)\right\| \\ &\leq \Bigg(\left(\frac{2\alpha^2}{t^2} + \frac{\left(2 + \alpha^2\|B\|^2\right)t^2\beta^2(t)}{\alpha^2}\right) \\ &\quad\times \|P_\Omega(y(t)) - P_\Omega(\hat{y}(t))\|^2 \\ &\quad+ \Bigg(\frac{2\alpha^2}{t^2} + \frac{t^2\beta^2(t)\left(1 + \mathfrak{l}_g + \alpha^2\|B\|^2\right)}{\alpha^2} \\ &\quad+ \frac{t^2\beta^2(t)\mu^2}{\alpha^2}\left\|B^T B\right\|^2\Bigg)\left\|\xi(t) - \hat{\xi}(t)\right\|^2 \\ &\quad+ \left(\frac{t^2\beta^2(t)}{\alpha^2}\|B\|^2\right)\left\|\zeta(t) - \hat{\zeta}(t)\right\|^2\Bigg)^{\frac{1}{2}} \\ &\leq \Bigg(\frac{\left(3 + \left(2\alpha^2 + 1\right)\|B\|^2 + \mathfrak{l}_g + \mu^2\left\|B^T B\right\|^2\right)t^2\beta^2(t)}{\alpha^2} \\ &\quad+ \frac{4\alpha^2}{t^2}\Bigg)^{\frac{1}{2}}\left\|Z(t) - \hat{Z}(t)\right\|. \end{aligned}$$

Let $l(t) = \left(\frac{4\alpha^2}{t^2} + \frac{\left(3 + \left(2\alpha^2 + 1\right)\|B\|^2 + \mathfrak{l}_f + \mu^2\|B^T B\|^2\right)t^2\beta^2(t)}{\alpha^2}\right)^{\frac{1}{2}}$. It is worth noting that $l(t)$ is continuous on $[t_0, +\infty)$, then, one has $l(\cdot)$ is integrable on $[t_0, T]$ where $0 < t_0 < T < +\infty$.

For ii), given any $Z \in \Omega \times \mathbb{R}^n \times \mathbb{R}^m$, we have

$$\begin{aligned} \int_{t_0}^{T}\|F(t, Z(t))\|dt \leq &\Big(\|P_\Omega(y(t))\|^2 + \|\nabla g(\xi(t))\|^2 \\ &+ \|\zeta(t)\|^2 + \|\xi(t)\|^2 + \|y(t)\|^2\Big)^{\frac{1}{2}} \\ &\times \int_{t_0}^{T}\Bigg(\frac{t^2\beta^2(t)(2 + 2\mu^2\|B^T B\|^2)}{\alpha^2} \\ &+ \frac{t^2\beta^2(t)(2\alpha^2 + 1)\|B\|^2}{\alpha^2} + \frac{4\alpha^2}{t^2}\Bigg)^{\frac{1}{2}} dt \end{aligned}$$

and conclusion ii) is true according to the following condition:

$$\begin{aligned} t \to \frac{4\alpha^2}{t^2} &+ \frac{t^2\beta^2(t)\left(2 + 2\mu^2\left\|B^T B\right\|^2\right)}{\alpha^2} \\ &+ \frac{t^2\beta^2(t)\left(2\alpha^2 + 1\right)\|B\|^2}{\alpha^2}. \end{aligned}$$

Conditions i) and ii) imply the existence and uniqueness of the strong global solution (12) of APDPNA-S.

Viability: Intuitively, since $\dot{\xi}(t) = \frac{\alpha}{t}(P_\Omega(y(t)) - \xi(t))$, the derivative $\dot{\xi}(t)$ will point towards Ω, i.e., making variable $\xi(t)$ always in the feasible set Ω.

We rigorously prove variable $\xi(t)$ remains in feasible set Ω for all $t \geq t_0 > 0$. Suppose there exists $t_1 > 0$, such that $\xi^1 = \xi(t_1) \notin \Omega$. A hyperplane exists that strictly separates the variables ξ^1 and Ω since Ω is a closed convex set, i.e., there exist ω, $v \in \mathbb{R}^n$, such that $\langle v, \xi^1 - \omega\rangle > 0$ and $\langle v, \xi - \omega\rangle < 0$, $\forall \xi \in \Omega$. Denote $\mathfrak{d}(\xi) = \langle v, \xi - \omega\rangle$. Since the solution trajectory $\xi(t)$ is continuously differentiable, $t \to \mathfrak{d}(\xi(t))$ is also continuously differentiable, and $\dot{\mathfrak{d}}(\xi(t)) = \langle v, \dot{\xi}(t)\rangle$.

Since initial value $\xi_0 \in \Omega$, we can obtain $\mathfrak{d}(\xi_0) < 0$ and $\mathfrak{d}(\xi^1) > 0$. Thus there is $\tau_\dagger$ so that $\mathfrak{d}(\xi^\dagger) = 0$ and $\mathfrak{d}(\xi^1) > 0, \forall t \in (t_\dagger, t_1]$, that is, $t_\dagger$ is the last time $\xi(t)$ crosses the separating hyperplane ($t_\dagger$ is simply $\sup\{t : \mathfrak{d}(\xi(t)) < 0\}$). $\mathfrak{d}(\xi^1) - \mathfrak{d}(\xi^\dagger) > 0$, from Taylor's theorem, there exists $t_\ddagger \in [t_\dagger, t_1]$, such that

$$\begin{aligned}\eth\left(\xi^{1}\right)-\eth\left(\xi^{\dagger}\right)&=\dot{\eth}(\xi(t_{\ddagger}))=\langle v,\dot{\xi}(t_{\ddagger})\rangle\\&=\frac{\alpha}{t}\langle v,P_{\Omega}(t_{\ddagger})-\xi(t_{\ddagger})\rangle\\&=\frac{\alpha}{t}\eth(P_{\Omega}(t_{\ddagger})-\eth(\xi(t_{\ddagger})))<0\end{aligned}\tag{20}$$

since $P_{\Omega}(t_{\ddagger})\in\Omega$. This is a contradiction, from which the proof is derived, and is shown in Fig. 2. ■

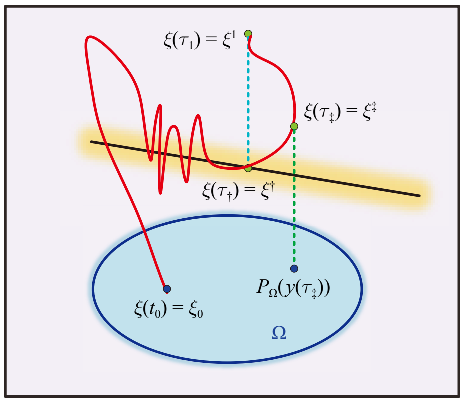


Fig. 2. The illustration of the viability of APDPNA-S (12).

Theorem 3: Assume that $\beta:[t_0,+\infty)\rightarrow(0,+\infty)$ is a continuous differentiable function and it satisfies $\dot{\beta}(t)\leq((\alpha-2)\beta(t))/t$ and let (ξ^*,y^*,ζ^*) and $(\xi(t),y(t),\zeta(t))\in\Omega\times\mathbb{R}^n\times\mathbb{R}^m$ be optimal solution and a strong global solution of APDPNA-S (12) respectively. Then for any $(\xi(t_0),y(t_0),\zeta(t_0))\in\Omega\times\mathbb{R}^n\times\mathbb{R}^m$, the following conditions hold:

i) $\xi(t)$, $P_{\Omega}(y(t))=\frac{t}{\alpha}\dot{\xi}(t)+\xi(t)$, $\zeta(t)$ are bounded.

ii) For any $\xi(t)\in\Omega, t\in[t_0,+\infty)$, we have

$$L_{\mu}(\xi(t),\zeta^*)-L_{\mu}(\xi^*,\zeta(t))\leq\frac{\alpha^2V(t_0)}{\beta(t)t^2}$$

$$\|B\xi(t)-c\|\leq\frac{\alpha\sqrt{\frac{2V(t_0)}{\mu}}}{t\sqrt{\beta(t)}}.$$

iii)

$$\begin{aligned}&\int_{t_0}^{+\infty}t\left(\frac{\alpha^2\beta(t)-2\beta(t)-\dot{\beta}(t)t}{\alpha^2}\right)\\&\quad\times\left(L_{\mu}(\xi(t),\zeta^*)-L_{\mu}(\xi^*,\zeta(t))\right)dt<+\infty\end{aligned}$$

$$\int_{t_0}^{+\infty}\frac{\mu\beta(t)t}{2\alpha}\|B\xi(t)-c\|^2dt<+\infty$$

$$\int_{t_0}^{+\infty}\frac{t}{\alpha^2}\left\|\dot{\xi}(t)\right\|^2dt<+\infty,\ \left\|\dot{\xi}(t)\right\|=O\left(\frac{1}{t}\right).$$

Proof: Construct a Lyapunov function $V:[t_0,+\infty)\rightarrow\mathbb{R}$ as follows:

$$\begin{aligned}V(t)=&\frac{\beta(t)t^2}{\alpha^2}\left(L_{\mu}(\xi(t),\zeta^*)-L_{\mu}(\xi^*,\zeta)\right)\\&+\frac{1}{2\alpha}\left(\left\|y(t)-P_{\Omega}(y^*)\right\|^2-\|y(t)-P_{\Omega}(y(t))\|^2\right)\\&+\frac{1}{2\alpha}\left\|\xi(t)-\xi^*\right\|^2+\frac{1}{2}\left\|\zeta(t)-\zeta^*\right\|^2\end{aligned}\tag{21}$$

where (ξ^*,y^*,ζ^*) is the equilibrium point (i.e., ξ^* is the optimal solution of problem (9)) to APDPNA-S (12) with $P_{\Omega}(y^*)=\xi^*$.

The Lyapunov function $V(t)$ is continuously differentiable and positive definite (i.e., $V(t)=0$, if $(\xi(t),y(t),\zeta(t))=(\xi^*,y^*,\zeta^*)$; $V(t)>0$, if $(\xi(t),y(t),\zeta(t))\neq(\xi^*,y^*,\zeta^*)$ and radially unbounded of $\xi(t)$, $\zeta(t)$ due to $\|\zeta(t)-\zeta^*\|^2\geq0$, $\|\xi(t)-x^*\|^2\geq0$, $\|y(t)-P_{\Omega}(y^*)\|^2-\|y(t)-P_{\Omega}(y(t))\|^2\geq0$ (see Lemma 3) and

$$\begin{aligned}&g(\xi(t))+(\zeta^*)^T(B\xi(t)-c)+\frac{\mu}{2}\|B\xi(t)-c\|^2-g(\xi^*)\\&\quad=g(\xi(t))+\left(B^T\zeta^*\right)^T(\xi(t)-\xi^*)\\&\qquad+\frac{\mu}{2}\|B\xi(t)-c\|^2-f(\xi^*)\\&\quad\in g(\xi(t))-g(\xi^*)-\nabla g(\xi^*)^T(\xi(t)-\xi^*)\\&\qquad-N_{\Omega}(\xi^*)^T(\xi(t)-\xi^*)+\frac{\mu}{2}\|B\xi(t)-c\|^2\\&\quad\geq I_{\Omega}(\xi^*)-I_{\Omega}(\xi(t))+\frac{\mu}{2}\|B\xi(t)-c\|^2\\&\quad=\frac{\mu}{2}\|B\xi(t)-c\|^2\geq0\end{aligned}\tag{22}$$

where the first equation holds due to $B\xi^*=c$, the inclusion equation holds according to (13a), i.e., $B^T\zeta^*\in-\nabla g(\xi^*)-N_{\Omega}(\xi^*)$, the first inequality is satisfied by using the convexity of g and I_{Ω}, and the second equality holds due to the *Viability* of $\xi(t)$ in Theorem 1.

According to the chain rule of derivation, we get

$$\begin{aligned}\dot{V}(t)=&\left(\frac{2\beta(t)t+\dot{\beta}(t)t}{\alpha^2}\right)(g(\xi(t))-g(\xi^*)\\&+(B\xi(t)-c)^T\zeta^*+\frac{\mu}{2}\|B\xi(t)-c\|^2\Big)+\frac{\beta(t)t}{\alpha}\\&\times\left(\nabla g(\xi(t))+B^T\zeta^*+\mu B^TB(\xi(t)-\xi^*)\right)^T\\&\times\left(\ P_{\Omega}(y(t))-\xi(t)\ \right)+\frac{t\beta(t)}{\alpha}(\zeta(t)-\zeta^*)^T\\&\times B(P_{\Omega}(y(t))-\xi^*)-\frac{\beta(t)t}{\alpha}(P_{\Omega}(y(t))-P_{\Omega}(y^*))^T\\&\times\left(\nabla g(\xi(t))+\mu B^TB(\xi(t)-\xi^*)+B^T\zeta(t)\right.\\&+y(t)-P_{\Omega}(y(t)))+\frac{1}{\alpha}(\xi(t)-\xi^*)^T\dot{\xi}(t)\\&-\frac{1}{\alpha}(P_{\Omega}(y(t))-P_{\Omega}(y^*))^T\dot{\xi}(t)\\\leq&\left(\frac{2\beta(t)t+\dot{\beta}(t)t}{\alpha^2}\right)(g(\xi(t))-g(\xi^*)\\&+(B\xi(t)-c)^T\zeta^*+\frac{\mu}{2}\|B\xi(t)-c\|^2\Big)+\frac{\beta(t)t}{\alpha}\\&\times\left(\nabla g(\xi(t))+B^T\zeta^*+\mu B^TB(\xi(t)-\xi^*)\right)^T\\&\times(P_{\Omega}(y^*)-\xi(t))-\frac{t}{\alpha^2}\left\|\dot{\xi}(t)\right\|^2\end{aligned}$$

$$
\begin{aligned}
&\leq \left(\frac{2\beta(t)t+\dot{\beta}(t)t}{\alpha^2}\right)(g(\xi(t))-g(\xi^*)\\
&\quad+(B\xi(t)-c)^T\zeta^*+\frac{\mu}{2}\|B\xi(t)-c\|^2\Big)\\
&\quad-\frac{\beta(t)t}{\alpha}\Big(g(\xi(t))-g(\xi^*)+\mu\|B\xi(t)-c\|^2\\
&\quad+(B\xi(t)-c)^T\zeta^*\Big)-\frac{t}{\alpha^2}\left\|\dot{\xi}(t)\right\|^2\\
&=\left(\frac{2\beta(t)t+\dot{\beta}(t)t^2}{\alpha^2}-\frac{\beta(t)t}{\alpha}\right)-\frac{t}{\alpha^2}\left\|\dot{\xi}(t)\right\|^2\\
&\quad\times\Big(g(\xi(t))+(B\xi(t)-c)^T\zeta^*-g(\xi^*)\\
&\quad+\frac{\mu}{2}\|B\xi(t)-c\|^2\Big)-\frac{\mu\beta(t)t}{2\alpha}\|B\xi(t)-c\|^2\leq 0 \qquad (23)
\end{aligned}
$$

where the first inequality can be obtained by $-(P_\Omega(y(t))-P_\Omega(y^*))^T(y(t)-P_\Omega(y(t)))\leq 0$ in Lemma 2, the second inequality holds thanks to the convexity of g, and the third inequality is satisfied since $\dot{\beta}(t)\leq((\alpha-2)\beta(t))/t$.

From (23), i.e., $\dot{V}(t)\leq 0$, we get

$$V(t)\leq V(t_0),\forall\,\xi(t)\in\Omega,\ t\in[t_0,+\infty). \qquad (24)$$

Based on the above inequality, the definition of $V(t)$ and Lemma 3, it can be concluded that for any $\xi(t)\in\Omega$ and $t\in[t_0,+\infty)$, the following holds:

$$
\begin{aligned}
&\frac{1}{2\alpha}\left\|\xi(t)-\xi^*\right\|^2+\frac{1}{2\alpha}\left\|P_\Omega(y(t))-P_\Omega(y^*)\right\|^2\\
&\quad+\frac{1}{2}\left\|\zeta(t)-\zeta^*\right\|^2\leq V(t)\leq V(t_0)<+\infty
\end{aligned}
$$

which means that $\xi(t)$, $P_\Omega(y(t))=\frac{t}{\alpha}\dot{\xi}(t)+\xi(t)$, $\zeta(t)$ are bounded, i.e.,

$$\sup_{t\in[t_0,+\infty)} t\left\|\dot{\xi}(t)\right\|<+\infty \qquad (25)$$

which implies $\|\dot{\xi}(t)\|=O(1/t)$.

Furthermore, from (21) and (24), one has

$$
\begin{aligned}
&L_\mu(\xi(t),\zeta^*)-L_\mu(\xi^*,\zeta(t))\leq\frac{\alpha^2V(t_0)}{\beta(t)t^2}\\
&\|B\xi(t)-c\|\leq\frac{\alpha\sqrt{\frac{2V(t_0)}{\mu}}}{t\sqrt{\beta(t)}}. \qquad (26)
\end{aligned}
$$

Integrating (23) from 0 to $+\infty$ yields

$$
\begin{aligned}
&\int_{t_0}^{+\infty}t\left(\frac{\alpha^2\beta(t)-2\beta(t)-\dot{\beta}(t)t}{\alpha^2}\right)\\
&\quad\times\Big(L_\mu(\xi(t),\zeta^*)-L_\mu(\xi^*,\zeta(t))\Big)dt<+\infty\\
&\int_{t_0}^{+\infty}\frac{\mu\beta(t)t}{2\alpha}\|B\xi(t)-c\|^2dt<+\infty\\
&\int_{t_0}^{+\infty}\frac{t}{\alpha^2}\left\|\dot{\xi}(t)\right\|^2dt<+\infty. \qquad (27)
\end{aligned}
$$

■

Remark 2: In Theorem 3, the assumption $\dot{\beta}(t)\leq((\alpha-2)\beta(t))/t$ holds if choosing $\beta(t)=\theta t^\eta$ with $\theta>0$ and $0<\eta\leq(\alpha-2)$, thus the conclusions in Theorem 3 are true. In this case $\dot{\beta}(t)/\beta(t)=\eta/t$. Integrating the above equation from $[t_0,t]$, we have

$$
\begin{aligned}
&\ln\beta(t)-\ln\beta(t_0)=\eta(\ln t-\ln t_0)\\
&\Rightarrow\beta(t)=\frac{\beta(t_0)}{(t_0)^\eta}t^\eta\underset{\frac{\beta(t_0)}{(t_0)^\eta}=\theta>0}{=}\theta t^\eta.
\end{aligned}
$$

Next, we will discuss the optimal convergence of APD-PNA-S (12) when $\theta>0$, $0<\eta\leq(\alpha-2)$, $\beta(t)=\theta t^\eta$. Before presenting the following theorem we need to provide a necessary lemma as follows:

Lemma 4 [34]: Let $\varphi:[t_0,+\infty)\to\mathbb{R}^n$ be a continuous differentiable function and $t_0>0$, $\mathfrak{a}\geq 0$, $\mathfrak{b}\geq 0$, and if

$$\|\mathfrak{a}\varphi(t)+t\dot{\varphi}(t)\|\leq\mathfrak{b}\forall t\geq t_0$$

holds, then, one has $\sup_{t\geq t_0}\|t\dot{\varphi}(t)\|<+\infty$.

Corollary 1: Let (ξ^*,y^*,λ^*) and $(\xi(t),y(t),\zeta(t))\in\Omega\times\mathbb{R}^n\times\mathbb{R}^m$ be an optimal solution and a strong global solution of APDPNA-S (12) when $\beta(t)=\theta t^\eta$ with $\theta>0$ and $0<\eta\leq(\alpha-2)$. Then for any $(\xi(t_0),y(t_0),\zeta(t_0))\in\Omega\times\mathbb{R}^n\times\mathbb{R}^m$, one has

$$
\begin{aligned}
&\left|g(\xi(t))-g(\xi^*)\right|=O\left(\frac{1}{t^{\eta+2}}\right)\\
&\|B\xi(t)-c\|=O\left(\frac{1}{t^{\eta+2}}\right). \qquad (28)
\end{aligned}
$$

Proof: Integrating the both side of $\dot{\lambda}(t)=t\beta(t)(BP_\Omega(y(t))-c)$ from t_0 to t, one has

$$
\begin{aligned}
\lambda(t)-\lambda(t_0)&=\int_{t_0}^{t}\dot{\lambda}(t)ds\\
&=\int_{t_0}^{t}s\beta(s)(BP_\Omega(y(s))-c)ds\\
&=\int_{t_0}^{t}s\beta(s)(B\xi(s)-c)ds+\int_{t_0}^{t}\frac{1}{\alpha}s^2\beta(s)d(B\xi(s)-c)\\
&=\frac{1}{\alpha}s^2\beta(s)(B\xi(s)-c)|_{s=t_0}^{s=t}\\
&\quad+\frac{1}{\alpha}\int_{t_0}^{t}s\Big((\alpha-2)\beta(s)-s\dot{\beta}(s)\Big)(B\xi(s)-c)ds\\
&=\frac{1}{\alpha}\theta t^{\eta+2}(B\xi(t)-c)-\frac{1}{\alpha}\theta t_0^{\eta+2}(B\xi(t_0)-c)\\
&\quad+\frac{1}{\alpha}\theta\int_{t_0}^{t}(\alpha-2-\eta)t^{\eta+1}(B\xi(s)-c)ds. \qquad (29)
\end{aligned}
$$

From Theorem 3, we conclude that the dual variable $\lambda(t)$ is bounded, this together (29) implies

$$
\begin{aligned}
&\Bigg\|(\alpha-2-\eta)\int_{t_0}^{t}t^{\eta+1}(B\xi(s)-c)ds\\
&\quad+t^{\eta+2}(B\xi(t)-c)\Bigg\|\leq D \qquad (30)
\end{aligned}
$$

where

$$D=\frac{\alpha}{\theta}\sup_{t\in[t_0,+\infty)}\|\lambda(t)-\lambda(t_0)\|+\left\|t_0^{\eta+2}(B\xi(t_0)-c)\right\|<+\infty. \qquad (31)$$

By Lemma 4 with $\varphi(t)=\int_{t_0}^{t}t^{\eta+1}(B\xi(s)-c)ds$, $\mathfrak{a}=\alpha-$

$2-\eta$ and $\mathfrak{b} = D$, we obtain

$$\sup_{t\in[t_0,+\infty)} \left\| t^{\eta+2}(B\xi(t)-c)\right\| < +\infty$$

which implies

$$\|B\xi(t)-c\| = O\left(\frac{1}{t^{\eta+2}}\right).$$

This in combination with Theorem 3 implies

$$\left|g(\xi(t))-g(\xi^*)\right| \le L_\mu(\xi(t),\zeta^*)-L_\mu(\xi^*,\zeta(t))+\left\|\zeta^*\right\|\|B\xi(t)-c\|$$
$$= O\left(\frac{1}{t^{\eta+2}}\right). \tag{32}$$

■

B. The Problem (9) With Nonsmooth Convex Objective Function g

In this subsection, our objective is to address the problem (9) which involves nonsmooth objective functions, the subdifferential $\partial g(\xi)$ of g at ξ is a closed and convex set since g is a closed and proper, nonsmooth convex function. A classical attempt to directly extend APDPNA-S (12) to be able to solve nonsmooth convex optimization problems is to substitute the second differential equation in APDPNA-S (12) with the differential inclusion $\dot{y}(t) \in -\frac{t}{\alpha}\beta(t)(\partial g(\xi) + \mu(B\xi(t)-c) + B^T\zeta(t) - P_\Omega(y(t)) + y(t))$. It may not be sufficient, as we shall see below, to ensure the decreasing of the Lyapunov function $V(t)$ according to a continuous trajectory of solution. As we will observe, this approach may not be enough to guarantee the decrease of the Lyapunov function $V(t)$ along continuous solution trajectories. It follows from [36] that the directional derivative $g(\xi;\dot{\xi}) = \sup_{h\in\partial g(\xi)}\langle h,\dot{\xi}\rangle$ in the nondifferentiable case plays a central role in the derivation of the correct dynamics.

From the work in [28], one can associate the set of subgradients that have reached their maximum value to $d(\xi;z)$ (since $\partial g(\xi)$ is the compact set in this case, the upper limit value is reached). Let's represent this set

$$d(\xi;z) = \underset{h\in\partial g(\xi)}{\arg\max}\langle h,z\rangle. \tag{33}$$

Based on the preceding discussion, we propose the following accelerated primal-dual projection neurodynamic approach (abbreviated as APDPNA-NS) for the solving the problem (9) with nonsmooth objective function g as follows:

$$\begin{cases}\dot{\xi}(t) = \dfrac{\alpha}{t}(P_\Omega(y(t))-\xi(t))\\ \dot{y}(t) \in -\dfrac{t}{\alpha}\beta(t)\Big(d\big(\xi(t);\dot{\xi}(t)\big)+\mu B^T(B\xi(t)-c)\\ \qquad +B^T\zeta(t)+y(t)-P_\Omega(y(t))\Big)-\dot{\xi}(t)\\ \dot{\zeta}(t) = t\beta(t)(BP_\Omega(y(t))-c)\end{cases} \tag{34}$$

where $\alpha \ge 2$, $\beta(t):[t_0,+\infty)\to(0,+\infty)$ and it satisfies $\dot{\beta}(t) \le \frac{\alpha-2}{t}\beta(t)$.

Remark 3: It is notable that the same equation $\dot{\xi}(t) = \frac{\alpha}{t}(P_\Omega(y(t))-\xi(t))$ is used for the dynamic trajectory of $\dot{\xi}(t)$ in APDPNA-S (12) and APDPNA-NS (34). Thus, according to Theorem 2, the solution $\xi(t)$ of APDPNA-NS (34) is also viable, i.e., $\xi(t)\in\Omega, \forall t \ge t_0 > 0$. Nevertheless, note that in APDPNA-NS (34), we do not investigate the existence and uniqueness of its solution $\xi(t)$. In addition, since the objective function in problem (9) is nonsmooth, the KKT conditions of problem (9) become

$$\partial g(\xi^*) + B^T\lambda^* + \mathcal{N}_\Omega(\xi^*) \ni 0$$
$$B\xi^* - c = 0.$$

Thus the optimality of solution of APDPNA-NS (34) can be proved employing a similar approach as in Theorem 1.

Theorem 4: Assume that $\beta:[t_0,+\infty)\to(0,+\infty)$ is a continuous differentiable function and it satisfies $\dot{\beta}(t) \le ((\alpha-2)\beta(t))/t$ and let (ξ^*,y^*,ζ^*) and $(\xi(t),y(t),\zeta(t)) \in \Omega\times\mathbb{R}^n\times\mathbb{R}^m$ be optimal solution and a strong global solution of APDPNA-NS (34) respectively. Then for any $(\xi(t_0),y(t_0),\zeta(t_0)) \in \Omega\times\mathbb{R}^n\times\mathbb{R}^m$, the following conditions hold:

i) $\xi(t)$, $P_\Omega(y(t)) = \frac{t}{\alpha}\dot{\xi}(t)+\xi(t)$, $\zeta(t)$ are bounded.

ii) For any $\xi(t)\in\Omega, t\in[t_0,+\infty)$, we have

$$L_\mu(\xi(t),\zeta^*)-L_\mu(\xi^*,\zeta(t)) \le \frac{\alpha^2 E(t_0)}{\beta(t)t^2}$$

$$\|B\xi(t)-c\| \le \frac{\alpha\sqrt{\frac{2E(t_0)}{\mu}}}{t\sqrt{\beta(t)}}.$$

iii) Let $\beta(t) = \theta t^\eta$ with $\theta > 0$ and $0 < \eta \le (\alpha-2)$, then one has

$$\left|g(\xi(t))-g(\xi^*)\right| = O\left(\frac{1}{t^{\eta+2}}\right)$$

$$\|B\xi(t)-c\| = O\left(\frac{1}{t^{\eta+2}}\right). \tag{35}$$

Proof: Constructing a Lyapunov function $E(t)$ which is almost identical to Theorem 3, except that the objective function g is nonsmooth, and it is

$$\begin{aligned}E(t) =& \frac{\mu(t)t^2}{\alpha^2}\Big(L_\beta(\xi(t),\zeta^*)-L_\mu(\xi^*,\zeta(t))\Big)\\ &+\frac{1}{2\alpha}\Big(\left\|y(t)-P_\Omega(y^*)\right\|^2-\|y(t)-P_\Omega(y(t))\|^2\Big)\\ &+\frac{1}{2}\left\|\zeta(t)-\zeta^*\right\|^2+\frac{1}{2\alpha}\left\|\xi(t)-x^*\right\|^2\\ =& \frac{\beta(t)t^2}{\alpha^2}\Big(g(\xi(t))+(\zeta^*)^T(B\xi(t)-c)\\ &+\frac{\mu}{2}\|Bx(t)-c\|^2-g(\xi^*)\Big)\\ &+\frac{1}{2\alpha}\Big(\left\|y(t)-P_\Omega(y^*)\right\|^2-\|y(t)-P_\Omega(y(t))\|^2\Big)\\ &+\frac{1}{2\alpha}\left\|\xi(t)-\xi^*\right\|^2+\frac{1}{2}\left\|\zeta(t)-\zeta^*\right\|^2.\end{aligned} \tag{36}$$

For proving the differentiability of the Lyapunov function $E(t)$, we use the difference quotient, which is defined as $\epsilon > 0$.

$$\begin{aligned}\Delta_t(\epsilon) &= \frac{E(t+\epsilon)-E(t)}{\epsilon}\\ &= \frac{1}{\epsilon}\Bigg(\frac{\beta(t+\epsilon)(t+\epsilon)^2}{\alpha^2}\Big((\zeta^*)^T(B\xi(t+\epsilon)-c)\\ &\quad + \frac{\beta}{2}\|B\xi(t+\epsilon)-c\|^2 + g(\xi(t+\epsilon)) - g(\xi^*)\Big)\\ &\quad + \frac{1}{2}\left\|\zeta(t+\epsilon)-\zeta^*\right\|^2 + \frac{1}{2\alpha}\Big(\left\|y(t+\epsilon)-P_\Omega(y^*)\right\|^2\\ &\quad - \|y(t+\epsilon)-P_\Omega(y(t+\epsilon))\|^2\Big) - \frac{1}{2}\left\|\zeta(t)-\zeta^*\right\|^2\\ &\quad + \frac{1}{2\alpha}\left\|\xi(t+\epsilon)-\xi^*\right\|^2 - \frac{1}{2\alpha}\left\|\xi(t)-\xi^*\right\|^2\\ &\quad - \frac{\beta(t)t^2}{\alpha^2}\Big(g(\xi(t)) + \frac{\beta}{2}\|B\xi(t)-c\|^2 - g(\xi^*)\\ &\quad + (\zeta^*)^T(B\xi(t)-c)\Big) - \frac{1}{2\alpha}\Big(\left\|y(t)-P_\Omega(y^*)\right\|^2\\ &\quad - \|y(t)-P_\Omega(y(t))\|^2\Big)\Bigg). \end{aligned} \tag{37}$$

Based on the fact that the convex function is locally Lipschitz (so that $g(\xi+o(\epsilon)) = g(\xi)+o(\epsilon)$), and that $\frac{1}{2\alpha}(\|y(t)-P_\Omega(y^*)\|^2 - \|y(t)-P_\Omega(y(t))\|^2)$, $(\lambda^*)^T(B\xi(t)-c)+\frac{\beta}{2}\|B\xi(t)-c\|^2$, $\|\xi(t)-\xi^*\|^2$ and $\|\zeta(t)-\zeta^*\|^2$ are differentiable, we have

$$\begin{aligned}\Delta_t(\epsilon) &= \frac{\beta(t)t^2}{\alpha^2}\frac{\left(g\left(\xi(t)+\epsilon\dot{\xi}(t)\right)-g(\xi(t))\right)}{\epsilon} + o(1)\\ &\quad + \frac{\beta(t)t^2}{\alpha^2}\frac{d}{dt}\left((\zeta^*)^T(B\xi(t)-c)+\frac{\beta}{2}\|Bx(t)-c\|^2\right)\\ &\quad + \frac{\dot{\beta}(t)t^2+\beta(t)2t}{\alpha^2}(g(\xi(t))-g(\xi^*))\\ &\quad + \frac{1}{2\alpha}\frac{d}{dt}\|\xi(t)-\xi(t)\|^2 + \frac{1}{2}\frac{d}{dt}\left\|\zeta(t)-\zeta^*\right\|^2\\ &\quad + \frac{\dot{\beta}(t)t^2+\beta(t)2t}{\alpha^2}(B(\xi(t))-c)^T\zeta^*\\ &\quad + \frac{\dot{\beta}(t)t^2+\beta(t)2t}{\alpha^2}\left(\frac{\beta}{2}\|B\xi(t)-c\|^2\right)\\ &\quad + \frac{1}{2\alpha}\frac{d}{dt}\left(\left\|y(t)-P_\Omega(y^*)\right\|^2 - \|y(t)-P_\Omega(y(t))\|^2\right). \end{aligned} \tag{38}$$

By using

$$\lim_{\epsilon\to0,\epsilon>0}\frac{g\left(\xi(t)+\epsilon\dot{\xi}(t)\right)-g(\xi(t))}{\epsilon} = g'\left(\xi(t);\dot{\xi}(t)\right)$$

$$\begin{aligned}&\frac{d}{dt}\left((\zeta^*)^T(B\xi(t)-c)+\frac{\beta}{2}\|B\xi(t)-c\|^2\right)\\ &\quad = (\zeta^*)^T B\dot{\xi}(t) + (B\xi(t)-c)^T B\dot{\xi}(t)\end{aligned}$$

$$\frac{1}{2\alpha}\frac{d}{dt}\|\xi(t)-\xi(t)\|^2 = \frac{1}{\alpha}(\xi(t)-\xi(t))^T\dot{\xi}(t)$$

$$\frac{1}{2}\frac{d}{dt}\left\|\zeta(t)-\zeta^*\right\|^2 = (\zeta(t)-\zeta^*)^T\dot{\zeta}(t)$$

$$\begin{aligned}&\frac{1}{\alpha}\frac{d}{dt}\left(\left\|y(t)-P_\Omega(y^*)\right\|^2 - \|y(t)-P_\Omega(y(t))\|^2\right)\\ &\quad = \frac{1}{\alpha}(P_\Omega(y(t))-P_\Omega(y^*))^T\dot{y}(t)\end{aligned}$$

we have

$$\begin{aligned}&\lim_{\epsilon\to0,\epsilon>0}\Delta_t(\epsilon)\\ &= \frac{\beta(t)t^2}{\alpha^2}\Big(g'\left(\xi(t);\dot{\xi}(t)\right) + \frac{\alpha}{t}(\zeta^*)^T B(P_\Omega(y(t))-\xi(t))\\ &\quad + \frac{\alpha\mu}{t}(B\xi(t)-c)^T B(P_\Omega(y(t))-\xi(t))\Big)\\ &\quad + \frac{\dot{\beta}(t)t^2+\beta(t)2t}{\alpha^2}\Big(\ (\zeta^*)^T(B\xi(t)-c)\\ &\quad + g(\xi(t))-g(\xi^*)+\frac{\mu}{2}\|B\xi(t)-c\|^2\Big)\\ &\quad + \frac{t}{\alpha}\beta(t)(\zeta(t)-\zeta^*)^T(BP_\Omega(y(t))-c)\\ &\quad - \frac{t}{\alpha}\beta(t)(P_\Omega(y(t))-P_\Omega(y^*))^T\Big(d\left(\xi(t);\dot{\xi}(t)\right)\\ &\quad +\mu(B\xi(t)-c)+B^T\zeta(t)+y(t)-P_\Omega(y(t))\Big)\\ &\quad - \frac{1}{\alpha}(P_\Omega(y(t))-P_\Omega(y^*))^T\dot{x}(t) + \frac{1}{\alpha}(\xi(t)-\xi(t))^T\dot{x}(t)\\ &\le \frac{\beta(t)t^2}{\alpha^2}\Big(g'\left(\xi(t);\dot{\xi}(t)\right) - d\left(\xi(t);\dot{\xi}(t)\right)^T\dot{\xi}(t)\Big)\\ &\quad - \frac{t}{\alpha^2}\left\|\dot{\xi}(t)\right\|^2 + \frac{\dot{\beta}(t)t^2+\beta(t)2t}{\alpha^2}(g(\xi(t))-g(\xi^*)\\ &\quad + (\zeta^*)^T(B\xi(t)-c)+\frac{\mu}{2}\|B\xi(t)-c\|\Big)\\ &\quad + \frac{t\beta(t)}{\alpha} + \Big(g(\xi^*)-\frac{\mu}{2}\|B\xi(t)-c\|^2 - g(x(t))\\ &\quad - \frac{\mu}{2}\|B\xi(t)-c\|^2 - (\zeta^*)^T(B\xi(t)-c)\Big)\\ &\quad + \frac{t\beta(t)}{\alpha}\Big(g(\xi(t))-g(\xi^*)-d\left(\xi(t);\xi\dot{(}t)\right)^T(\xi^*-\xi(t))\Big)\\ &\le \frac{\beta(t)t^2}{\alpha^2}\Big(g'\left(x(t);\dot{\xi}(t)\right) - d\left(\xi(t);\xi\dot{(}t)\right)^T\dot{\xi}(t)\Big)\\ &\quad + \left(\frac{\dot{\beta}(t)t^2+\beta(t)2t}{\alpha^2} - \frac{t\beta(t)}{\alpha}\right)(g(\xi(t))-g(\xi^*)\\ &\quad + (\zeta^*)^T(B\xi(t)-c)+\frac{\mu}{2}\|B\xi(t)-c\|^2\Big)\\ &\quad - \frac{t\beta(t)}{\alpha}\frac{\mu}{2}\|B\xi(t)-c\|^2 - \frac{t}{\alpha^2}\left\|\dot{\xi}(t)\right\|^2 \le 0 \end{aligned} \tag{39}$$

where the first inequality holds due to $-(P_\Omega(y(t)) - P_\Omega(y^*))^T(y(t)-P_\Omega(y(t))) \le 0$ in Lemma 2, the second inequality is satisfied from the convexity of the function *f*, and the third inequality is satisfied due to $\dot{\beta}(t) \le ((\alpha-2)\beta(t))/t$.

The above conclusions hold with a similar proof to Theorem 3 and Corollary 1. To avoid duplicating the proof, we omit it here. ■

Remark 4: It's worth noting from (39) that if we use $\dot{y}(t) \in -\frac{t}{\alpha}\beta(t)(\partial g(\xi(t)) + \mu B^T(B\xi(t)-c) + B^T\zeta(t) - P_\Omega(y(t)) + y(t))$ (in other words, $\partial g(\xi(t))$ is a subgradient of $g'(\xi(t);\dot{\xi}(t)) - d(\xi(t);\dot{\xi}(t))^T\dot{\xi}(t)$ is non-negative by (33), and one cannot conclude that the Lyapunov function $E(t)$ is diminishing. It drives us to select the subgradient. Indeed, when $\dot{y}(t) \in -\frac{t}{\alpha}\beta(t)(d(\xi(t);\dot{x}(t)) + \mu B^T(B\xi(t)-c) + B^T\zeta(t) + y(t) - P_\Omega(y(t)))$, where $d(\xi(t);\dot{\xi}(t))$ is a subgradient of g at $\xi(t)$ that maximizes the linear functional $\langle\cdot,\xi(t)\rangle)$, the $g'(\xi(t);\dot{\xi}(t)) - d(\xi(t);\dot{\xi}(t))^T\dot{\xi}(t)$ term in (39) is non-positive, thus $\lim_{\epsilon\to 0,\epsilon>0}\Delta_t(\epsilon) \le 0$.

Remark 5: The choice of an initial point within the feasible set is a crucial aspect addressed in this manuscript, as it can impact the feasibility and convergence of the solutions in APDPNA-S (12) and APDPNA-NS (34). Since the projection operator Ω has a closed-form solution (refer to Lemma 1), it is straightforward to use $\xi_0 = P_\Omega(x_0) \in \Omega$, where $x_0 \in \mathbb{R}^n$, to obtain the initial value ξ_0 that satisfies the feasible set Ω.

IV. Numerical Simulations

In this section, to demonstrate the effectiveness and superiority of the proposed APDPNA-S (12) and its extended nonsmooth version APDPNA-NS (34), we discuss the sparse signal reconstruction problem in the compressed sensing. Compressed sensing, also known as compressive sampling, is a novel sampling theory that leverages the sparsity of signals. It achieves this by acquiring discrete samples of signals through random sampling at a much lower rate than the Nyquist sampling rate. Subsequently, it employs nonlinear reconstruction algorithms to perfectly reconstruct the signals. This concept has gained significant attention in various fields such as information theory, image processing, earth science, microwave imaging, pattern recognition, wireless communication, and biomedical engineering. The problem of recovering sparse signals, which is a key aspect of compressed sensing, involves reconstructing the sparse signal $z \in \mathbb{R}^n$ from small number of linear measurements (linear constraints) $c = Az \in \mathbb{R}^m$ with $m \ll n$. The dimensions of these measurements are much smaller than the spatial dimensions of the signal. Here, $A \in \mathbb{R}^{m\times n}$ (with $m \ll n$) represents the measurement matrix or dictionary. It is important to note that the sparse signal recovery problem is generally ill-posed and challenging due to the imbalance between the number of measurements and the signal dimensions ($m \ll n$). To address this challenge, [39] demonstrates that faithful recovery of z from the compressed measurement c is possible when measurement matrix A satisfies certain stable embedding conditions. Mathematically, the sparse signal reconstruction problem can be formulated as a basis pursuit (BP) problem.

A. Basis Pursuit

Basis pursuit (BP) problem as follows:

$$\min_{z\in\mathbb{R}^n} \|z\|_1\,,\ \text{s.t. } Az = c. \tag{40}$$

Fortunately, the BP problem (40) can be written equivalently as a linear programming problem with a linear constraint and positive-orthant constrained sets by a splitting method, i.e., dividing z into positive and negative parts as follows: $z = u - v$, $u = [z]^+ \ge 0$, $\{v = [-z]^+ \ge 0\}$, $[z]^+ = \max\{0, y\}$, thus, we have $\|z\|_1 = 1^T u + 1^T v$, corresponding the BP problem (40) becomes

$$\begin{aligned}&\min_{x\in\mathbb{R}^{2n}} g(x) = 1^T x\\ &\text{s.t. } Bx = c,\ x \ge 0\end{aligned} \tag{41}$$

where $x = [u^T, v^T]^T \in \mathbb{R}^{2n}$, $B = [A, -A]$.

Let $m = 100$, $n = 256$ and sparsity be 15. Apply APDPNA-S (12) to deal with problem (41) with $\beta(t) = \theta t^\eta, \theta = 0.1, \eta = 1$ and $\eta = 2$ respectively. From Fig. 3 (3(a) and 3(b), 3(d) and 3(e)), we can obtain that the trajectories of APDPNA-S (12) with $\eta = 1$ and $\eta = 2$ are globally asymptotically stable and sparse signals can be recovered by using the stable solutions of APDPNA-S (12). In addition, Fig. 3 (3(c) and 3(f)) shows the convergence rates of $|g(x(t)) - g(x^*)|$ and $\|Bx(t) - c\|$ of APDPNA-S (12) with $\eta = 1$ and $\eta = 2$ with classical sparse neurodynamic approaches: PNNSR-dynamic [21] and LPNN-LCA [23]. As can be seen from Fig. 3 (Figs. 3(c) and 3(f)), the APDPNA-S (12) with $\eta = 2$ has a faster convergence rate than that with $\eta = 1$, which is consistent with the concluding results of Theorem 3 and Corollary 1. In addition, convergence rates of $\eta = 1$ and $\eta = 2$ are faster than PNNSR-dynamic [21] and LPNN-LCA [23].

To better demonstrate the superiority of APDPNA-S (12), we use the example without set constraints in [31] that $g(x) = x^T M x$, $x \in \mathbb{R}^{50}$ with $M \in \mathbb{R}^{50\times 50}$ being a positive semifinite matrix that is generated by a standard Gaussian distribution and set $B \in \mathbb{R}^{10\times 50}$ and $c = 0 \in \mathbb{R}^{10}$. Under the same setting of $\alpha = 4$ and the same initial values, we compared APDPNA-S (12) equipped with $\beta(t) = \theta t^\eta$, $\theta = 0.1$, $\eta = 1$, $\eta = 2$ to PDGD [47], IPDDM [31], PDNAM [30], FPDA [34]. Fig. 4 shows the convergence results. As can be seen from Fig. 4, our proposed APDPNA-S (12) outperforms PDGD [47], IPDDM [31], PDNAM [30], FPDA [34] in both $|g(x(t)) - g(x^*)|$ and $\|Bx(t) - c\|$. With the same conditions, our proposed APDPNA-S (12) has a slight advantage over FPDA [34] especially in the speed of convergence of the equation constraints due to the augmented Lagrangian term of the constraints that was introduced in APDPNA-S (12). Moreover, FPDA [34] has performance superior than PDGD [47], IPDDM [31], PDNAM [30] it introduces a time scaling term.

B. Distributed Basis Pursuit

The BP problem (40) can be converted to a distributed BP problem, as discussed in [25]. This conversion is based on the consensus theorem for multi-agents on an undirected graph $\mathcal{G}$ and the row decomposition properties of the observation matrix A.

$$\begin{aligned}&\min_{x\in\mathbb{R}^{np}} g(x) = \sum_{i=1}^{p} \|x_i\|_1\\ &\text{s.t. } Lx = 0 \in \mathbb{R}^{np}, x \in \Omega = \{x \in \mathbb{R}^{np} | Ax = c\}\end{aligned} \tag{42}$$

where $A \in \mathbb{R}^{m\times np}$ (see Fig. 5), $L = L_p \otimes I \in \mathbb{R}^{np\times np}$ with $L_p \in \mathbb{R}^{p\times p}$ is the Laplacian matrix of undirected graph $\mathcal{G}$ and $\otimes$ is the Kronecker product.

By utilizing the proposed APDPNA-NS (34), we can solve the distributed BP problem (42) with specific parameters:

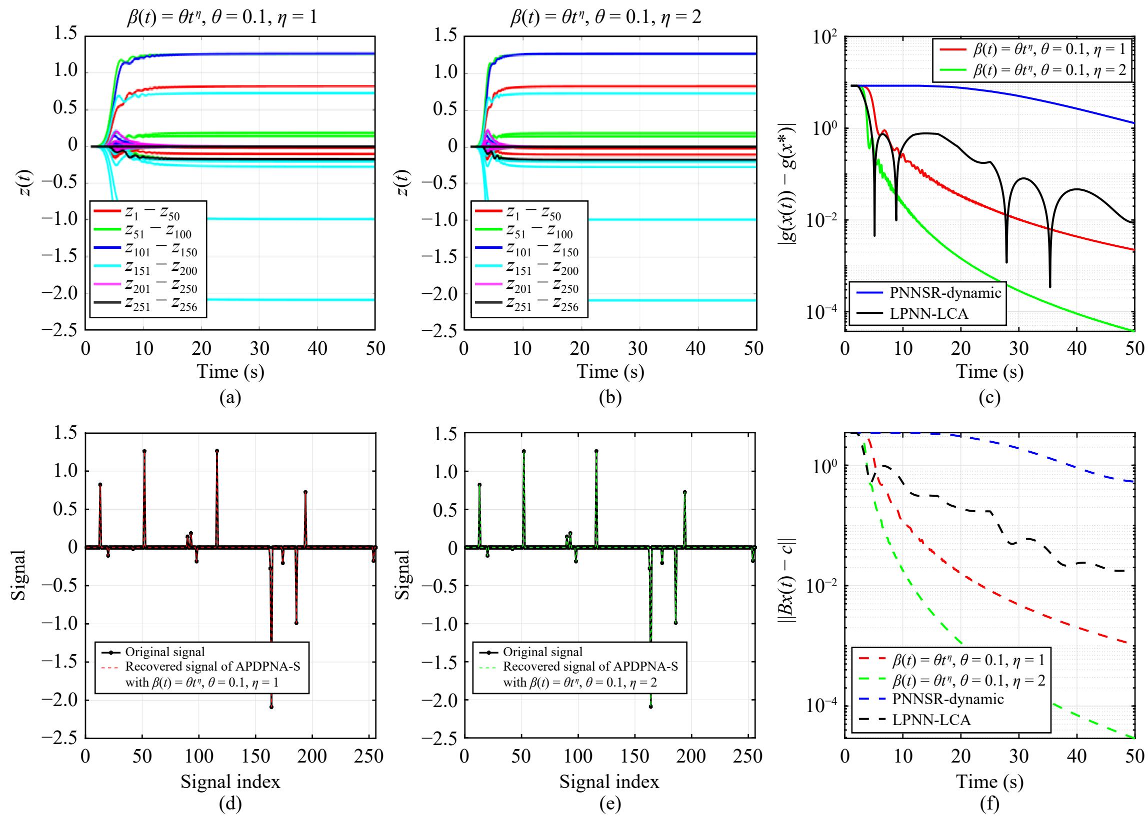


Fig. 3. Convergence properties of APDPNA-S (12) with $\beta(t)=\theta t^{\eta}$, $\theta=0.1$ for solving BP problem (41) ((a) Transient behaviors of $z(t)$ with $\eta=1$; (b) Transient behaviors of $z(t)$ with $\eta=2$; (c) Convergence rate of $|g(x(t))-g(x^*)|$ with PNNSR-dynamic [21] and LPNN-LCA [23]; (d) Recovered signal with $\eta=1$; (e) Recovered signal with $\eta=2$; (f) Convergence rate of $\|Bx(t)-c\|$ with PNNSR-dynamic [21] and LPNN-LCA [23]).

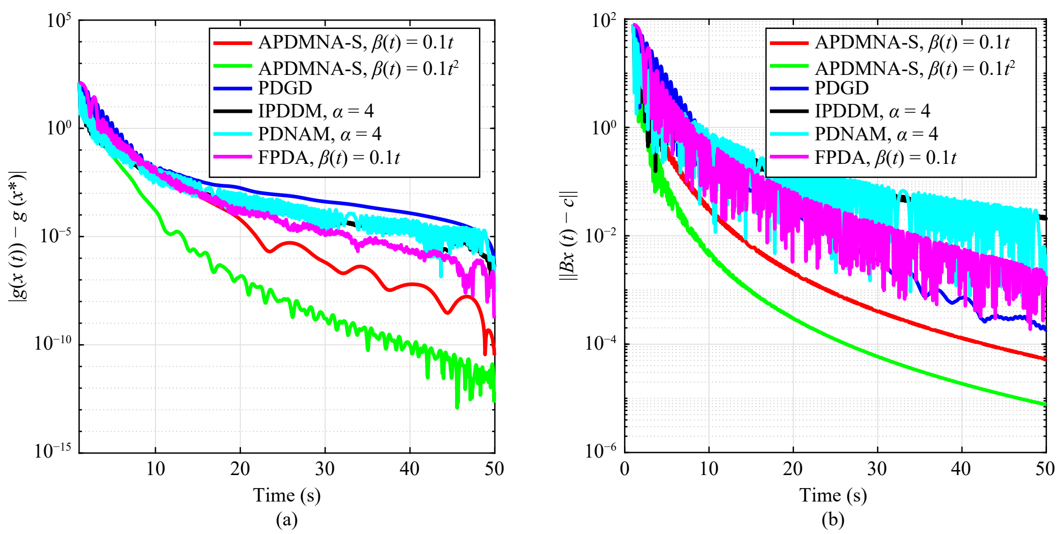


Fig. 4. Convergence properties of APDPNA-S (12) with $\alpha=4$, $\beta(t)=\theta t^{\eta}$, $\theta=0.1$, $\eta=1$, $\eta=2$, PDGD [47], IPDDM [31], PDNAM [30], FPDA [34] ((a) Convergence rate of $|g(x(t))-g(x^*)|$; (b) Convergence rate of $\|Bx(t)-c\|$).

$n=50$, $m=30$, sparsity $s=5$, and $p=5$. The problem is tackled within the context of a network consisting of 5 agents connected in an undirected ring configuration (refer to Fig. 5). The trajectories $x(t)$ of DPDPNA-NS (34) are illustrated in Figs. 6(a) and 6(d). These trajectories demonstrate the global asymptotic stability of x in two different scenarios, one with $\alpha=3$ and the other with $\alpha=4$. Furthermore, the subplots in Figs. 6(b) and 6(e) demonstrate that the sparse signals can be efficiently reconstructed in a distributed manner by the stabilized solutions of APDPNA-NS with $\alpha=3$ and $\alpha=4$. The Figs. 6(c) and 6(f) show that the APDPNA-NS (34) with $\alpha=5$ has an faster convergence rate than APDPNA-NS (34) with $\alpha=3$, which is consistent with the concluding results of Theorem 4.

V. Conclusions

We have proposed two novel accelerated primal-dual neurodynamic approaches with time scaling (APDPNA-S and

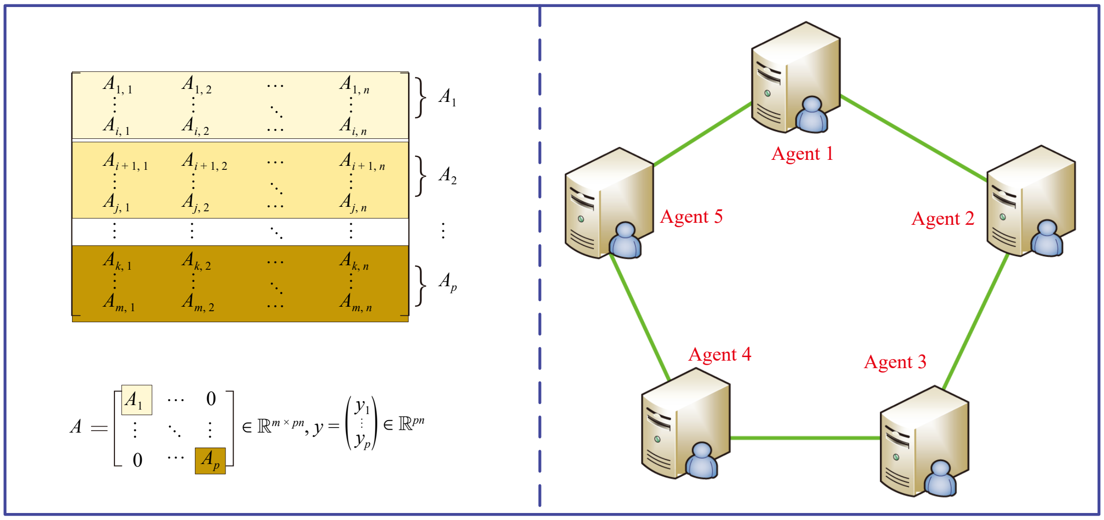


Fig. 5. Partition the rows of matrix A into p blocks, matrix A, vector y and an unbalanced directed network of 5 agents.

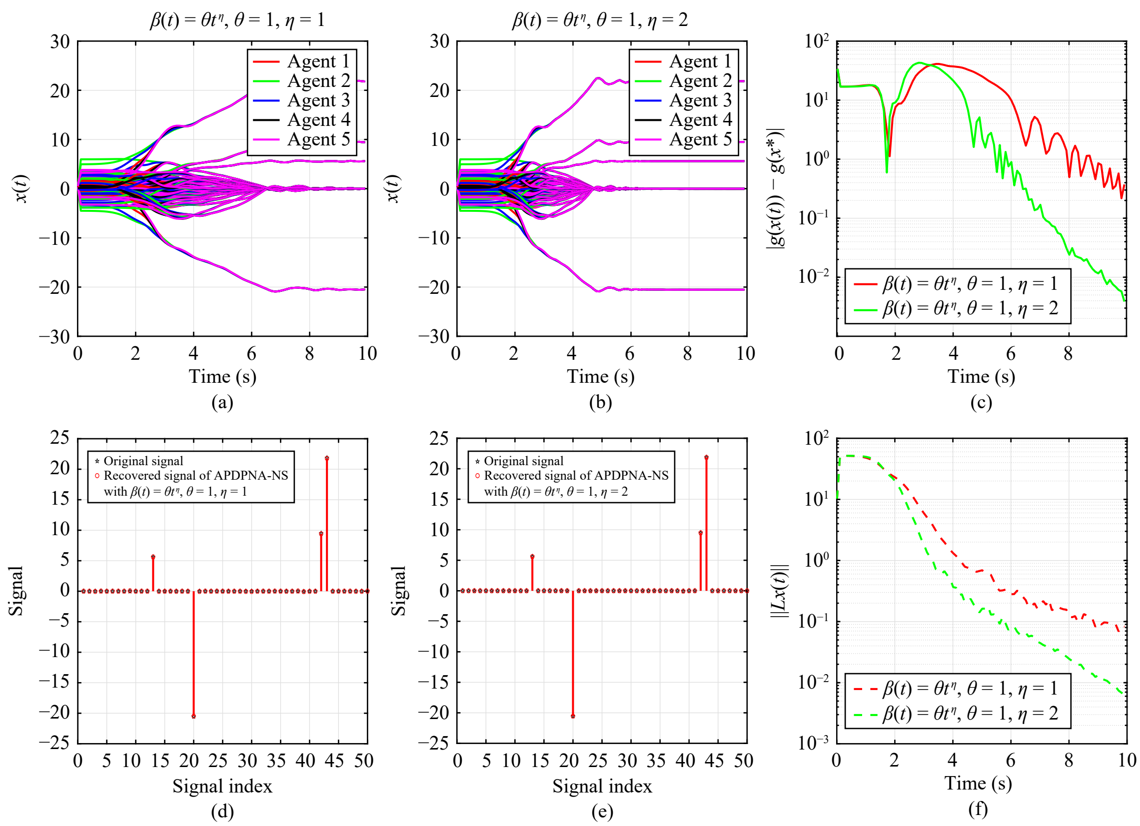


Fig. 6. Convergence properties of APDPNA-NS (34) with $\beta(t) = \theta t^\eta$, $\theta = 1$ and $\eta = 1$ for solving distributed BP problem (42) ((a) Transient behaviors of $x(t)$ with $\eta = 1$; (b) Transient behaviors of $x(t)$ with $\eta = 2$; (c) Convergence rate of $|g(x(t)) - g(x^*)|$; (d) Recovered signals with $\eta = 1$; (e) Recovered signal with $\eta = 2$; (f) Convergence rate of $\|Lx(t)\|$).

APDPNA-NS) to deal with smooth and nonsmooth convex optimization problem subject to linear and set constraints, without strongly convex assumption. We have proven the existence, uniqueness, and viability of the strong global solution for APDPNA-S. Additionally, we have demonstrated its optimality using the variational analysis method, and established the fast convergence properties of APDPNA-S by constructing a novel Lyapunov function. Furthermore, we have extended the APDPNA-S into a differential inclusion dynamical approach, i.e., APDPNA-NS by employing directional derivative, and have shown that APDPNA-NS have the same results as APDPNA-S by computing difference quotient of Lyapunov functions. The effectiveness of APDPNA-S and APDPNA-NS have been illustrated by two simulation examples on sparse signal reconstruction. In our future work, we plan to investigate inexact accelerated primal-dual projection neurodynamic approaches for addressing problem (9) in both smooth and nonsmooth scenarios. This involves approximating the closed-form solution of the projection operator when it is not readily available. Additionally, we aim to expand the scope of the proposed APDPNA-S (12) and APDPNA-NS (32) methods by applying them to solve convex optimization problems with inequality constraints and set constraints, thereby increasing their applicability.

REFERENCES

[1] Y. E. Nesterov, "A method of solving a convex programming problem with convergence rate $0(1/k^2)$," *Sov. Math. Dokl.*, vol. 27, no. 2, pp. 372–376, 1983.

[2] W. Su, S. Boyd, and E. Candès, "A differential equation for modeling Nesterov's accelerated gradient method: Theory and insights," in *Proc. 27th Int. Conf. Neural Information Processing Systems*, Montreal, Canada, 2014, pp. 2510–2518.

[3] H. Attouch and Z. Chbani. "Fast inertial dynamics and FISTA algorithms in convex optimization. Perturbation aspects," arXiv preprint arXiv: 1507.01367, 2015.

[4] H. Attouch, J. Peypouquet, and P. Redont, "Fast convex optimization via inertial dynamics with Hessian driven damping," *J. Differ. Equations*, vol. 261, no. 10, pp. 5734–5783, Nov. 2016.

[5] H. Attouch, Z. Chbani, and H. Riahi, "Combining fast inertial dynamics for convex optimization with Tikhonov regularization," *J. Math. Anal. Appl.*, vol. 457, no. 2, pp. 1065–1094, Jan. 2018.

[6] H. Attouch and A. Cabot, "Convergence of a relaxed inertial proximal algorithm for maximally monotone operators," *Math. Program.*, vol. 184, no. 1–2, pp. 243–287, Nov. 2020.

[7] A. Wibisono, A. C. Wilson, and M. I. Jordan, "A variational perspective on accelerated methods in optimization," in *Proc. Natl. Acad. Sci. USA*, vol. 113, no. 47, pp. E7351–E7358, Nov. 2016.

[8] A. C. Wilson, B. Recht, and M. I. Jordan, "A Lyapunov analysis of accelerated methods in optimization," *J. Mach. Learn. Res.*, vol. 22, no. 1, p. 113, Jan. 2021.

[9] F. Alimisis, A. Orvieto, G. Bécigneul, and A. Lucchi, "A continuous-time perspective for modeling acceleration in Riemannian optimization," in *Proc. 23rd Int. Conf. Artificial Intelligence and Statistics*, Palermo, Italy, 2020, pp. 1297–1307.

[10] A. Vassilis, A. Jean-François, and D. Charles, "The differential inclusion modeling FISTA algorithm and optimality of convergence rate in the case ≤ 3," *SIAM J. Optim.*, vol. 28, no. 1, pp. 551–574, Jan. 2018.

[11] H. Attouch, Z. Chbani, and H. Riahi, "Fast proximal methods via time scaling of damped inertial dynamics," *SIAM J. Optim.*, vol. 29, no. 3, pp. 2227–2256, Jan. 2019.

[12] D. Tank and J. Hopfield, "Simple 'neural' optimization networks: An A/D converter, signal decision circuit, and a linear programming circuit," *IEEE Trans. Circuits Syst.*, vol. 33, no. 5, pp. 533–541, May 1986.

[13] M. P. Kennedy and L. O. Chua, "Neural networks for nonlinear programming," *IEEE Trans. Circuits Syst.*, vol. 35, no. 5, pp. 554–562, May 1988.

[14] Y. Xia and G. Feng, "On convergence rate of projection neural networks," *IEEE Trans. Autom. Control*, vol. 49, no. 1, pp. 91–96, Jan. 2004.

[15] Y. Xia and J. Wang, "A recurrent neural network for solving nonlinear convex programs subject to linear constraints," *IEEE Trans. Neural Netw.*, vol. 16, no. 2, pp. 379–386, Mar. 2005.

[16] X. Hu and J. Wang, "Design of general projection neural networks for solving monotone linear variational inequalities and linear and quadratic optimization problems," *IEEE Trans. Syst., Man, Cybern. Part B (Cybern.)*, vol. 37, no. 5, pp. 1414–1421, Oct. 2007.

[17] X. He, T. Huang, J. Yu, C. Li, and C. Li, "An inertial projection neural network for solving variational inequalities," *IEEE Trans. Cybern.*, vol. 47, no. 3, pp. 809–814, Mar. 2017.

[18] S. Zhang and A. G. Constantinides, "Lagrange programming neural networks," *IEEE Trans. Circuits Syst. II: Analog Digit. Signal Process.*, vol. 39, no. 7, pp. 441–452, Jul. 1992.

[19] W. Bian and X. Chen, "Smoothing neural network for constrained non-Lipschitz optimization with applications," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 3, pp. 399–411, Mar. 2012.

[20] Q. Liu and J. Wang, "A projection neural network for constrained quadratic minimax optimization," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 11, pp. 2891–2900, Nov. 2015.

[21] Q. Liu and J. Wang, "L_1-minimization algorithms for sparse signal reconstruction based on a projection neural network," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 3, pp. 698–707, Mar. 2016.

[22] Z. Guo, S. Yang, and J. Wang, "Global synchronization of stochastically disturbed memristive neurodynamics via discontinuous control laws," *IEEE/CAA J. Autom. Sinica*, vol. 3, no. 2, pp. 121–131, Apr. 2016.

[23] R. Feng, C. S. Leung, A. G. Constantinides, and W. J. Zeng, "Lagrange programming neural network for nondifferentiable optimization problems in sparse approximation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 10, pp. 2395–2407, Oct. 2017.

[24] C. Xu and X. He, "A fully distributed approach to optimal energy scheduling of users and generators considering a novel combined neurodynamic algorithm in smart grid," *IEEE/CAA J. Autom. Sinica*, vol. 8, no. 7, pp. 1325–1335, Jul. 2021.

[25] Y. Zhao, X. Liao, X. He, and R. Tang, "Centralized and collective neurodynamic optimization approaches for sparse signal reconstruction via L_1-minimization," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 12, pp. 7488–7501, Dec. 2022.

[26] J. Wang, J. Wang, and Q.-L. Han, "Receding-horizon trajectory planning for under-actuated autonomous vehicles based on collaborative neurodynamic optimization," *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 11, pp. 1909–1923, Nov. 2022.

[27] X. Wen, S. Qin, and J. Feng, "A novel projection neural network for solving a class of monotone variational inequalities," *IEEE Trans. Syst., Man, Cybern.: Syst.*, vol. 53, no. 9, pp. 5580–5590, Sept. 2023.

[28] W. Krichene, A. M. Bayen, and P. L. Bartlett, "Accelerated mirror descent in continuous and discrete time," in *Proc. 28th Int. Conf. Neural Information Processing Systems*, Montreal, Canada, 2015, pp 2845–2853.

[29] Y. Zhao, X. Liao, and X. He, "Novel projection neurodynamic approaches for constrained convex optimization," *Neural Netw.*, vol. 150, pp. 336–349, Jun. 2022.

[30] X. Zeng, J. Lei, and J. Chen, "Dynamical primal-dual nesterov accelerated method and its application to network optimization," *IEEE Trans. Autom. Control*, vol. 68, no. 3, pp. 1760–1767, Mar. 2023.

[31] X. He, R. Hu, and Y. P. Fang, "Convergence rates of inertial primal-dual dynamical methods for separable convex optimization problems," *SIAM J. Control Optim.*, vol. 59, no. 5, pp. 3278–3301, Jan. 2021.

[32] R. I. Boţ and D. K. Nguyen, "Improved convergence rates and trajectory convergence for primal-dual dynamical systems with vanishing damping," *J. Differ. Equations.*, vol. 303, pp. 369–406, Dec. 2021.

[33] H. Attouch, Z. Chbani, J. Fadili, and H. Riahi, "Fast convergence of dynamical ADMM via time scaling of damped inertial dynamics," *J. Optim. Theory Appl.*, vol. 193, no. 1–3, pp. 704–736, Jun. 2022.

[34] X. He, R. Hu, and Y. P. Fang, "Fast primal-dual algorithm via dynamical system for a linearly constrained convex optimization problem," *Automatica*, vol. 146, p. 110547, Dec. 2022.

[35] Y. Zhao, X. Liao, X. He, M. Zhou, and C. Li, "Accelerated primal-dual mirror dynamics for centralized and distributed constrained convex optimization problems," *J. Mach. Learn. Res.*, vol. 24, no. 1, p. 343, Jan. 2023.

[36] R. T. Rockafellar, *Convex Analysis*. Princeton, USA: Princeton University Press, 1970.

[37] N. Parikh and S. Boyd, "Proximal algorithms," *Found. Trends Optim.*, vol. 1, no. 3, pp. 127–239, Jan. 2014.

[38] A. Chambolle and T. Pock, "An introduction to continuous optimization for imaging," *Acta Numer.*, vol. 25, pp. 161–319, May 2016.

[39] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory.*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[40] C. Li, X. Yu, W. Yu, T. Huang, and Z. W. Liu, "Distributed event-triggered scheme for economic dispatch in smart grids," *IEEE Trans. Ind. Inf.*, vol. 12, no. 5, pp. 1775–1785, Oct. 2016.

[41] W. T. Lin, Y. W. Wang, C. Li, and X. Yu, "Distributed resource allocation via accelerated saddle point dynamics," *IEEE/CAA J. Automa. Sinica*, vol. 8, no. 9, pp. 1588–1599, Sept. 2021.

[42] C. Li, X. Yu, T. Huang, and X. He, "Distributed optimal consensus over resource allocation network and its Application to dynamical economic dispatch," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 6, pp. 2407–2418, Jun. 2018.

[43] Z. Lin, H. Li, and C. Fang, *Accelerated Optimization for Machine Learning: First-Order Algorithms*. Singapore: Springer, 2020.

[44] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jul. 2011.

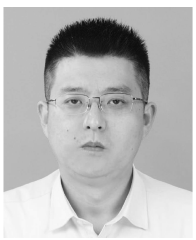
[45] R. I. Boţ and E. R. Csetnek, "A forward-backward dynamical approach to the minimization of the sum of a nonsmooth convex with a smooth

nonconvex function," *ESAIM: COCV*, vol. 24, no. 2, pp. 463–477, 2018.

[46] B. Abbas, H. Attouch, and B. F. Svaiter, "Newton-like dynamics and forward-backward methods for structured monotone inclusions in Hilbert spaces," *J. Optim. Theory Appl.*, vol. 161, no. 2, pp. 331–360, May 2014.

[47] X. Chen and N. Li, "Exponential stability of primal-dual gradient dynamics with non-strong convexity," in *Proc. American Control Conf.*, Denver, USA, 2020, pp. 1612–1618.

You Zhao received the M.S. degree in signal and information processing from the College of Electronics and Information Engineering, Southwest University in 2018. and the Ph.D. degree in computer science and technology from Chongqing University in 2023. He is currently a Postdoctoral Researcher with the College of Electronic and Information Engineering, Southwest University. His research interests include neurodynamic optimization, distributed optimization, compressed sensing, minimax optimization, and smart grid.



Xing He received the B.S. degree in mathematics and applied mathematics from the Department of Mathematics, Guizhou University in 2009, and the Ph.D. degree in computer science and technology from Chongqing University in 2013. From November 2012 to October 2013, he was a Research Assistant with the Texas A&M University at Qatar, Qatar. From December 2015 to February 2016, he was a Senior Research Associate with the City University of Hong Kong, China. Currently, he is a Professor with the College of Electronic and Information Engineering, Southwest University. His research interests include neural networks, bifurcation theory, optimization method, smart grid, and nonlinear dynamical system.

Mingliang Zhou received the Ph.D. degree in computer science from Beihang University in 2017. He was a Postdoctoral Fellow with the Department of Computer Science, City University of Hong Kong, China, from September 2017 to September 2019. He was a Postdoctoral Fellow with the State Key Laboratory of Internet of Things for Smart City, University of Macau, Macau, China, from October 2019 to October 2021. He is currently an Associate Professor with the School of Computer Science, Chongqing University. His research interests include image and video coding, perceptual image processing, multimedia signal processing, rate control, multimedia communication, machine learning, and optimization.

Tingwen Huang (Fellow, IEEE) received the B.S. degree in mathematics from Southwest Normal University in 1990, the M.S. degree in applied mathematics from Sichuan University in 1993, and the Ph.D. degree in mathematics from Texas A&M University, College Station, USA, in 2002. He was a Lecturer with Jiangsu University from 1994 to 1998, and a Visiting Assistant Professor with Texas A&M University, USA, in 2003. He was an Assistant Professor from 2003 to 2009 and an Associate Professor from 2009 to 2013 with Texas A&M University at Qatar, Qatar, where he has been a Professor since 2013. His research interests include neural networks, complex networks, chaos and dynamics of systems, and operator semi-groups and their applications.