A SYMMETRIC PRIMAL-DUAL METHOD WITH TWO EXTRAPOLATION STEPS FOR COMPOSITE CONVEX OPTIMIZATION *

LIHAN ZHOU † AND FENG MA‡

Abstract. Symmetry is a recurring feature in algorithms for monotone operator theory and convex optimization, particularly in problems involving the sum of two operators, as exemplified by the Peaceman–Rachford splitting scheme. However, in more general settings—such as composite optimization problems with three convex functions or structured convex-concave saddle-point formulations—existing algorithms often exhibit inherent asymmetry. In particular, the Condat–Vũ algorithm and the asymmetry forward-backward-adjoint (AFBA) method, while efficient and widely adopted, apply extrapolation only to either the primal or the dual update, resulting in unbalanced iterations. In this work, we introduce a symmetric primal-dual algorithm (SPDA) that applies extrapolation to both primal and dual iterates, thereby preserving symmetry in the iteration scheme. The algorithm encompasses the Condat–Vũ and AFBA methods as special cases and permits more flexible step-size choices. We establish global convergence under standard assumptions and derive both ergodic and non-ergodic convergence rates. The results demonstrate that symmetry can be preserved in first-order methods for optimizing the sum of three convex functions without compromising convergence guarantees or practical simplicity.

Key words. Composite optimization, Saddle-point problem , Primal-dual algorithm, Symmetry.

MSC codes. 68U10, 90C25, 65K10, 65J22

1. Introduction.

In this paper, we consider the composite optimization problem

$$\min_{x\in\mathcal{X}} f(x) + h(x) + g(Ax), \tag{1.1}$$

where $\mathcal{X}$ is an n-dimensional real Euclidean space, A is an $m\times n$ matrix and f, g, and h are convex, proper, and lower semicontinuous functions defined on $\mathcal{X}$. In addition, h has a Lipschitz continuous gradient with constant L_h. We assume that f and g are "simple" in the sense that their proximal operators can be computed efficiently (see (2.1)). Such composite formulations often arise in image processing and machine learning applications, including regularized regression problems, we refer readers to e.g., [3, 5, 17, 20].

In general, computing the proximal operator of the composite function $g\circ A$ is challenging. A standard approach to addressing this difficulty is to reformulate problem (1.1) as a convex-concave saddle-point problem:

$$\min_{x\in\mathcal{X}}\max_{y\in\mathcal{Y}}\{\Phi(x,y) = f(x) + h(x) + y^TAx - g^*(y)\}, \tag{1.2}$$

where $\mathcal{Y}$ is an m-dimensional real Euclidean space and g^* denotes the Fenchel conjugate of g. This reformulation exploits the separable structure of f, h, and g^*, enabling the development of efficient primal-dual algorithms and facilitating theoretical analysis. Under strong duality conditions, the saddle points of $\Phi(x,y)$ correspond to the optimal solutions of the original problem (1.1).

*This research was partially supported by the NSFC Grant 12171481 and 12471298.

†214611022@csu.edu.cn, School of Automation, Central South University, Changsha, 410083, Hunan, China.

‡mafengnju@gmail.com, High-Tech Institute of Xi'an, Xi'an, 710025, Shaanxi, China.

Recently, several primal-dual algorithms have been developed for solving (1.2), including Condat-Vũ [6, 21] and its Bregman case [16], Primal-Dual Fixed-Point algorithm (PDFP) [4], Asymmetric Forward-Backward-Adjoint splitting (AFBA) [18], Davis-Yin [8], and Primal-Dual Three-Operator splitting (PD3O) [22]. The relationships and interconnections among these algorithms have been extensively explored in [22, 7]. To motivate our study, we present two representative algorithms: the Condat-Vũ method and the AFBA method. The Condat–Vũ algorithm iterates as follows:

$$\begin{cases} x^{k+1} = \mathrm{Prox}_{(\tau f)}[x^k - \tau\nabla h(x^k) - \tau A^T y^k], & \text{(1.3a)} \\ \tilde{x}^{k+1} = x^{k+1} + (x^{k+1} - x^k), & \text{(1.3b)} \\ y^{k+1} = \mathrm{Prox}_{(\sigma g^*)}[y^k + \sigma A\tilde{x}^{k+1}], & \text{(1.3c)} \end{cases}$$

where $\mathrm{Prox}_{(\tau f)}$ and $\mathrm{Prox}_{(\sigma g^*)}$ denote the proximal operators of τf and σg^* (see (2.1)), and $\tau, \sigma > 0$ are step-size parameters. By casting the scheme (1.3) in the form of forward-backward splitting, the convergence can be proved under the condition

$$\tau\sigma\|AA^T\| + \frac{\tau L_h}{2} < 1. \tag{1.4}$$

Moreover, when $h = 0$, the algorithm reduces to the classical primal-dual hybrid gradient method [2].

The AFBA algorithm, introduced in [18], iterates as follows:

$$\begin{cases} \tilde{x}^{k+1} = \mathrm{Prox}_{(\tau f)}[x^k - \tau\nabla h(x^k) - \tau A^T y^k], & \text{(1.5a)} \\ y^{k+1} = \mathrm{Prox}_{(\sigma g^*)}[y^k + \sigma A\tilde{x}^{k+1}], & \text{(1.5b)} \\ x^{k+1} = \tilde{x}^{k+1} - \tau A^T(y^{k+1} - y^k). & \text{(1.5c)} \end{cases}$$

with the improved convergence conditions, proposed in [15],

$$\tau\sigma\|AA^T\| < 1, \quad 0 < \tau < \frac{2}{L_h}. \tag{1.6}$$

In this scheme, the proximal updates for x and y mirror those of the Condat–Vũ algorithm, while the additional correction step is computationally inexpensive. Thus, both algorithms exhibit similar per-iteration complexity; however, the convergence condition for AFBA is less restrictive than that for the Condat–Vũ algorithm, allowing for a broader range of parameter selections.

Both the Condat–Vũ and AFBA algorithms update the primal and dual variables in an alternating manner and exhibit several desirable properties, including low per-iteration computational cost and strong theoretical convergence guarantees. However, these methods inherently impose an asymmetry in the treatment of primal and dual updates, which does not fully exploit the intrinsic symmetry of saddle-point problems. Specifically, in each iteration, one variable is updated using an extrapolated step, while the other is not. Despite these asymmetries, several algorithms designed for saddle-point problems or composite optimization problems explicitly leverage symmetry by balancing the treatment of primal and dual updates. Notable examples include the Peaceman–Rachford splitting method [9, 10] or the symmetric ADMM [12] for the convex problem. These methods demonstrate that symmetric structures can lead to improved convergence properties and broader applicability. So a natural question arises: can we construct a symmetric method, in which both primal and dual updates incorporate extrapolation techniques?

In this work, we provide an affirmative answer to this question. Motivated by the symmetric PDHG algorithm introduced in [19], we extend this framework to address the saddle point problem (1.2). Our proposed symmetric scheme is formulated as follows:

$$
\left\{
\begin{aligned}
&\tilde{x}^{k+1} = \mathrm{Prox}_{(\tau f)}\big[x^k - \tau \nabla h(x^k) - \tau A^T y^k\big], && (1.7a)\\
&\bar{x}^{k+1} = \tilde{x}^{k+1} + \theta(\tilde{x}^{k+1} - x^k), && (1.7b)\\
&y^{k+1} = \mathrm{Prox}_{(\sigma g*)}\big[y^k + \sigma A \bar{x}^{k+1}\big], && (1.7c)\\
&x^{k+1} = \bar{x}^{k+1} - \tau A^T(y^{k+1} - y^k). && (1.7d)
\end{aligned}
\right.
$$

To further elucidate the structural properties of our method, we rewrite it in an alternative formulation:

$$
\left\{
\begin{aligned}
&\bar{x}^{k+1} = \tilde{x}^{k+1} + \theta(\tilde{x}^{k+1} - x^k), && (1.8a)\\
&\hat{y}^{k+1} = y^k + \sigma A \bar{x}^{k+1}, && (1.8b)\\
&y^{k+1} = \arg\min_y \{g^*(y) + \frac{1}{2\sigma}\|y - \hat{y}^{k+1}\|^2\}, && (1.8c)\\
&\bar{y}^{k+1} = y^{k+1} + (y^{k+1} - y^k), && (1.8d)\\
&\hat{x}^{k+2} = \bar{x}^{k+1} - \tau \nabla h(x^{k+1}) - \tau A^T \bar{y}^{k+1}, && (1.8e)\\
&\tilde{x}^{k+2} = \arg\min_x \{f(x) + \frac{1}{2\tau}\|x - \hat{x}^{k+2}\|^2\}. && (1.8f)
\end{aligned}
\right.
$$

Unlike the Condat–Vũ and AFBA algorithms, which apply extrapolation to only one of the primal or dual updates, our method introduces extrapolation in both updates, resulting in a fully symmetric structure, i.e., steps (1.8a) and (1.8d). This symmetry ensures that the primal and dual iterates are treated in a balanced manner, which can enhance stability and improve convergence behavior. Notably, our method unifies and generalizes the Condat–Vũ and AFBA algorithms within a symmetric framework. Specifically, the updates (1.7a)–(1.7c) generalize and recover the structure of the Condat–Vũ algorithm, while the steps (1.7a), (1.7c), and (1.7d) align with the update rules of the AFBA algorithm.

In this work, we establish the global convergence of the proposed symmetric primal-dual algorithm (SPDA) (1.7) for solving the saddle-point problem (1.2), without invoking any assumptions beyond standard convexity and Lipschitz differentiability. Specifically, we show that convergence is guaranteed under the following conditions:

$$
\theta \in \left(-1, 1 - \frac{\tau L_h}{2}\right), \quad 0 < \tau < \frac{4}{L_h}, \quad \tau\sigma < \frac{1}{\|AA^T\|}. \qquad (1.9)
$$

A notable feature of the proposed method lies in the flexibility in choosing the extrapolation parameter θ, which is allowed to vary over a broader range compared to existing algorithms such as the Condat–Vũ method, where $\theta = 1$ is typically required. In particular, when $\theta = 0$, the scheme reduces to the AFBA algorithm, and our analysis indicates that the associated convergence condition can be further relaxed. This flexibility facilitates adaptive parameter tuning, which may enhance both theoretical guarantees and empirical performance. Furthermore, when $h = 0$, the proposed method coincides with the symmetric PDHG algorithm developed in [19], and can thus be viewed as a natural generalization of that framework. Beyond global convergence, we also demonstrate that the proposed symmetric scheme achieves an $\mathcal{O}(1/t)$ ergodic convergence rate and an $\mathcal{O}(1/\sqrt{t})$ non-ergodic convergence rate.

The rest of this paper is organized as follows: Section 2 introduces fundamental definitions and lemmas that will be used throughout our analysis. Section 3 presents the convergence analysis of the SPDA algorithm, including its global convergence guarantee and convergence rate. Section 4 conducts some numerical experiments to verify the theoretical results. Finally, Section 5 summarizes the conclusions.

2. Preliminaries. In the section, we present some definitions and lemmas used throughout the paper. For a function h, ∇h denotes the gradient of h and $\mathbf{dom}\ h$ denotes the domain of h. For a vector x, $\|x\|_p$ denotes the l_p norm of x and $\|x\|_F$ denotes the Frobenius norm. For simplicity, $\|x\|$ denotes the l_2 norm of a vector x. For a matrix M, $M \succ 0$ denotes a positive definite matrix and $M \succeq 0$ denotes a positive semidefinite matrix. For a symmetric matrix M, $\lambda(M)$ denotes its eigenvalue. For any vector x and any symmetric matrix M, $\|x\|_M$ denotes $\sqrt{x^T M x}$. For a matrix A, $\|A\|$ denotes its maximum singular value. The proximal mapping for a convex and lower semi-continuous function f is defined as

$$\text{Prox}_{\tau f}(x) = \arg\min_u \{f(u) + \frac{1}{2\tau}\|u - x\|^2\}, \quad \forall x \in \Re^n. \tag{2.1}$$

Note that in this case, it always yields a unique solution.

LEMMA 2.1. *[11] Let $\mathbf{X} \in \Re^n$ be a closed convex set, let $\mathcal{F}(x) : \mathbf{X} \to \Re$ and $\mathcal{G}(x) : \mathbf{X} \to \Re$ be convex function and let $\mathcal{G}(x)$ be differentiable. Suppose that the solution set of the problem $\min\{\mathcal{F}(x) + \mathcal{G}(x) | x \in \mathbf{X}\}$ is non-empty, Then, a point $x^* \in \mathbf{X}$ is a solution if and only if*

$$x^* \in \mathbf{X},\ \mathcal{F}(x) - \mathcal{F}(x^*) + (x - x^*)^T \nabla \mathcal{G}(x^*) \geq 0,\ \forall x \in \mathbf{X}. \tag{2.2}$$

LEMMA 2.2. *For any vector $x \in \Re^n, y \in \Re^n, z \in \Re^n, p \in \Re^n$, and a symmetric matrix $S \in \Re^{n\times n}$, the following identity holds*

$$(x - y)^T S(z - p) = \frac{1}{2}\Big(\|x - p\|_S^2 - \|x - z\|_S^2 + \|y - z\|_S^2 - \|y - p\|_S^2\Big), \tag{2.3}$$

where $\|x - p\|_S^2 = (x - p)^T S(x - p)$.

LEMMA 2.3. *[1] For any convex differentiable functions h with L_h-Lipschitz continuous gradient, the following inequality holds*

$$(\nabla h(x) - \nabla h(y))^T (x - y) \geq \frac{1}{L_h}\|\nabla h(x) - \nabla h(y)\|^2, \quad \forall x, y \in \mathbf{dom}\ h. \tag{2.4}$$

LEMMA 2.4. *For any convex differentiable functions h with L_h-Lipschitz continuous gradient, the following inequality holds*

$$h(x) \leq h(y) + (x - y)^T \nabla h(z) + \frac{L_h}{2}\|x - z\|^2, \quad \forall x, y, z \in \mathbf{dom}\ h. \tag{2.5}$$

Proof. It follows from Descent lemma that for a differentiable function h with L_h-Lipschitz continuous gradient, the following inequality holds

$$h(x) \leq h(z) + (x - z)^T \nabla h(z) + \frac{L_h}{2}\|x - z\|^2. \tag{2.6}$$

Since h is a convex function, we have

$$h(z) \leq h(y) + (z - y)^T \nabla h(z). \tag{2.7}$$

Adding the inequalities (2.6) and (2.7) yields (2.5). ∎

3. Global Convergence. In this section, inspired by the construction approach proposed in [13], we develop a prediction-correction framework for analyzing the SPDA method (1.7). By carefully exploiting the contraction properties inherent in the algorithm's structure, we establish global convergence results and derive sufficient conditions under which the convergence of SPDA is guaranteed.

3.1. A prediction-correction framework. In this section, we firstly recall the optimality conditions associated with the problem (1.2), which serve as the basis for the subsequent analysis of the prediction-correction framework. Then, we develop a prediction-correction framework, which not only simplifies the convergence analysis but also provides deeper insights into the underlying mechanisms driving the algorithm's behavior.

By Lemma 2.1, the optimal condition for the problem (1.2) is given by

$$\begin{cases} f(x) - f(x^*) + (x - x^*)^T(A^T y^* + \nabla h(x^*)) \geq 0, & \text{(3.1a)} \\ g^*(y) - g^*(y^*) + (y - y^*)^T(-Ax^*) \geq 0, & \text{(3.1b)} \end{cases}$$

where x^*, y^* are the optimal solution to the problem (1.2). To facilitate the convergence analysis, we define some notations as follows

$$\begin{aligned} u &= \begin{pmatrix} x \\ y \end{pmatrix}, \quad \mathcal{F}(u) = f(x) + g^*(y), \\ \nabla F(u) &= \begin{pmatrix} A^T y \\ -Ax \end{pmatrix}, \quad \nabla H(u) = \begin{pmatrix} \nabla h(x) \\ 0 \end{pmatrix}, \end{aligned} \tag{3.2}$$

and introduce an auxiliary variable $\tilde{u}^{k+1}$ as follows

$$\tilde{u}^{k+1} = \begin{pmatrix} \tilde{x}^{k+1} \\ \tilde{y}^{k+1} \end{pmatrix} = \begin{pmatrix} \tilde{x}^{k+1} \\ y^{k+1} \end{pmatrix}. \tag{3.3}$$

Using the notations (3.2), the optimality condition (3.1) can be rewritten as

$$\mathcal{F}(u) - \mathcal{F}(u^*) + (u - u^*)^T \nabla \mathcal{G}(u^*) \geq 0, \tag{3.4}$$

where $u^* = (x^*, y^*)^T$ and $\nabla \mathcal{G}(u^*)$ is given by

$$\nabla \mathcal{G}(u^*) = \nabla F(u^*) + \nabla H(u^*) = \begin{pmatrix} A^T y^* + \nabla h(x^*) \\ -Ax^* \end{pmatrix}.$$

Then, by the optimal condition (3.4) and the definition of the auxiliary variable, we can derive a prediction-correction scheme for SPDA.

LEMMA 3.1. *Let u^k be the iterates generated by the algorithm* (1.7). *Define $\tilde{u}^{k+1} = (\tilde{x}^{k+1}, \tilde{y}^{k+1})$. The following inequality holds*

$$(u - \tilde{u}^{k+1})^T Q(\tilde{u}^{k+1} - u^k) \geq \mathcal{F}(\tilde{u}^{k+1}) - \mathcal{F}(u) + (\tilde{u}^{k+1} - u)^T \nabla \mathcal{G}(u) - \frac{L_h}{4}\|\tilde{x}^{k+1} - x^k\|^2, \tag{3.5}$$

where the matrix Q is defined by

$$Q = \begin{pmatrix} \frac{1}{\tau} I & -A^T \\ -\theta A & \frac{1}{\sigma} I \end{pmatrix}. \tag{3.6}$$

Proof. The optimality conditions for the iterates (1.7a) and (1.7c) are, respectively, given by

$$\begin{cases} f(x) - f(\tilde{x}^{k+1}) + (x - \tilde{x}^{k+1})^T \frac{1}{\tau}(\tilde{x}^{k+1} - x^k + \tau A^T y^{k+1} + \tau q^k) \geq 0, & (3.7a) \\ g^*(y) - g^*(y^{k+1}) + (y - y^{k+1})^T \frac{1}{\sigma}(y^{k+1} - y^k - \sigma A\bar{x}^{k+1}) \geq 0, & (3.7b) \end{cases}$$

where $q^k = \nabla h(x^k) - A^T(y^{k+1} - y^k)$. By substituting the iterate (1.7b) into the inequality (3.7b) and using the definition (3.3), we have

$$\begin{aligned} & g^*(y) - g^*(y^{k+1}) + (y - y^{k+1})^T \frac{1}{\sigma}(y^{k+1} - y^k - \sigma A\bar{x}^{k+1}) \\ & \qquad = g^*(y) - g^*(\tilde{y}^{k+1}) + (y - \tilde{y}^{k+1})^T \frac{1}{\sigma}(\tilde{y}^{k+1} - y^k) \\ & \qquad - (y - \tilde{y}^{k+1})^T(A\tilde{x}^{k+1}) - (y - \tilde{y}^{k+1})^T\Big(\theta A(\tilde{x}^{k+1} - x^k)\Big) \geq 0. \end{aligned} \tag{3.8}$$

Combining the inequality (3.8) and the inequality (3.7a) and using the notations (3.2) yields

$$\begin{aligned} \mathcal{F}(u) - \mathcal{F}(\tilde{u}^{k+1}) + (u - \tilde{u}^{k+1})^T \nabla F(\tilde{u}^{k+1}) + (x - \tilde{x}^{k+1})^T \nabla h(x^k) \\ + (u - \tilde{u}^{k+1})^T Q(\tilde{u}^{k+1} - u^k) \geq 0. \end{aligned} \tag{3.9}$$

Rearranging the above inequality, we get

$$\begin{aligned} (u - \tilde{u}^{k+1})^T Q(\tilde{u}^{k+1} - u^k) \geq \mathcal{F}(\tilde{u}^{k+1}) - \mathcal{F}(u) + (\tilde{u}^{k+1} - u)^T \nabla F(\tilde{u}^{k+1}) \\ + (\tilde{x}^{k+1} - x)^T \nabla h(x^k). \end{aligned} \tag{3.10}$$

For the last two terms of the inequality (3.10), we have

$$(\tilde{u}^{k+1} - u)^T \nabla F(\tilde{u}^{k+1}) = (\tilde{u}^{k+1} - u)^T \nabla F(u), \tag{3.11}$$

and

$$\begin{aligned} (\tilde{x}^{k+1} - x)^T \nabla h(x^k) =& (\tilde{x}^{k+1} - x)^T \nabla h(x) + (\tilde{x}^{k+1} - x)^T(\nabla h(x^k) - \nabla h(x)), \\ =& (\tilde{x}^{k+1} - x)^T \nabla h(x) + (\tilde{x}^{k+1} - x^k)^T(\nabla h(x^k) - \nabla h(x)) \\ & + (x^k - x)^T(\nabla h(x^k) - \nabla h(x)). \end{aligned} \tag{3.12}$$

Substituting (3.11) and (3.12) into (3.10) and using the notation (3.2), we get

$$\begin{aligned} (u - \tilde{u}^{k+1})^T Q(\tilde{u}^{k+1} - u^k) \geq & \mathcal{F}(\tilde{u}^{k+1}) - \mathcal{F}(u) + (\tilde{u}^{k+1} - u)^T \nabla \mathcal{G}(u) \\ & + (\tilde{x}^{k+1} - x^k)^T(\nabla h(x^k) - \nabla h(x)) \\ & + (x^k - x)^T(\nabla h(x^k) - \nabla h(x)). \end{aligned} \tag{3.13}$$

Applying Lemma 2.3 to the last term of the above inequality yields

$$\begin{aligned} (u - \tilde{u}^{k+1})^T Q(\tilde{u}^{k+1} - u^k) \geq & \mathcal{F}(\tilde{u}^{k+1}) - \mathcal{F}(u) + (\tilde{u}^{k+1} - u)^T \nabla \mathcal{G}(u) \\ & + (\tilde{x}^{k+1} - x^k)^T(\nabla h(x^k) - \nabla h(x)) \\ & + \frac{1}{L_h}\|\nabla h(x^k) - \nabla h(x)\|^2. \end{aligned} \tag{3.14}$$

Applying Cauchy-Schwarz inequality to the term $(\tilde{x}^{k+1} - x^k)^T(\nabla h(x^k) - \nabla h(x))$, we have

$$(\tilde{x}^{k+1} - x^k)^T(\nabla h(x^k) - \nabla h(x)) \geq -\frac{1}{2}\left(\frac{L_h}{2}\|\tilde{x}^{k+1} - x^k\|^2 + \frac{2}{L_h}\|\nabla h(x^k) - \nabla h(x)\|^2\right). \tag{3.15}$$

Combining (3.14) and (3.15), we obtain the inequality (3.5) ◻

LEMMA 3.2. *Let u^k be the iterates generated by the algorithm* (1.7)*. Define $\tilde{u}^{k+1} = (\tilde{x}^{k+1}, \tilde{y}^{k+1})^T$. The following inequality holds*

$$u^{k+1} = u^k - M(u^k - \tilde{u}^{k+1}), \tag{3.16}$$

where the matrix M is defined by

$$M = \begin{pmatrix} (1+\theta)I & -\tau A^T \\ 0 & I \end{pmatrix}. \tag{3.17}$$

Proof. Substituting the variable $\tilde{u}^{k+1}$ in (1.7d), we get

$$x^{k+1} = x^k + \tilde{x}^{k+1} - x^k + \theta(\tilde{x}^{k+1} - x^k) - \tau A^T(\tilde{y}^{k+1} - y^k). \tag{3.18}$$

Note that $\tilde{y}^{k+1} = y^{k+1}$ (see (3.3)), we get

$$y^{k+1} = y^k - (y^k - \tilde{y}^{k+1}). \tag{3.19}$$

Combining (3.18) with (3.19), we get (3.16). ◻

Lemma 3.1 and Lemma 3.2 establish a prediction-correction framework for analyzing the SPDA method. In particular, the iteration (3.5) corresponds to the prediction step, while (3.16) defines the correction step. This framework can provide a more structured basis for deriving convergence guarantees.

3.2. Contractive properties. In this section, based on the prediction-correction framework, we analyze the contractive properties of the sequence generated by SPDA. Specifically, we establish the following result:

THEOREM 3.3. *Let $\{u^k\}$ denote the sequence generated by* (1.7)*. Then, the following inequality holds:*

$$\|u^k - u^*\|_H^2 \geq \|u^{k+1} - u^*\|_H^2 + \|\tilde{u}^{k+1} - u^k\|_G^2 - \frac{L_h}{2}\|\tilde{x}^{k+1} - x^k\|^2, \tag{3.20}$$

where $\tilde{u}^{k+1} = (\tilde{x}^{k+1}, \tilde{y}^{k+1})$, and the matrices H and G are defined as

$$H = QM^{-1} = \begin{pmatrix} \frac{1}{\tau(1+\theta)}I & -\frac{\theta}{1+\theta}A^T \\ -\frac{\theta}{1+\theta}A & \frac{1}{\sigma}I - \frac{\tau\theta}{1+\theta}AA^T \end{pmatrix}, \tag{3.21}$$

$$G = Q + Q^T - M^THM = \begin{pmatrix} \frac{1-\theta}{\tau}I & 0 \\ 0 & \frac{1}{\sigma}I - \tau AA^T \end{pmatrix}. \tag{3.22}$$

where Q and M are defined in (3.6) *and* (3.17)*, respectively.*

Proof. Setting $u = u^*$ in (3.5), we get

$$(u^* - \tilde{u}^{k+1})^T Q(\tilde{u}^{k+1} - u^k) \geq \mathcal{F}(\tilde{u}^{k+1}) - \mathcal{F}(u^*) + (\tilde{u}^{k+1} - u^*)^T \nabla \mathcal{G}(u^*) - \frac{L_h}{4}\|\tilde{x}^{k+1} - x^k\|^2,$$

where u^* is the optimal solution to the problem (1.2). By the optimality condition (3.4), the above inequality implies

$$(u^* - \tilde{u}^{k+1})^T Q(\tilde{u}^{k+1} - u^k) \geq -\frac{L_h}{4}\|\tilde{x}^{k+1} - x^k\|^2.$$

Substituting the correction step (3.16) into this inequality, we have

$$(u^* - \tilde{u}^{k+1})^T Q M^{-1}(u^{k+1} - u^k) \geq -\frac{L_h}{4}\|\tilde{x}^{k+1} - x^k\|^2.$$

By the relation $H = QM^{-1}$, we have

$$(u^* - \tilde{u}^{k+1})^T H(u^{k+1} - u^k) \geq -\frac{L_h}{4}\|\tilde{x}^{k+1} - x^k\|^2. \tag{3.23}$$

Applying Lemma 2.2 to the term in the left side of the inequality (3.23) yields

$$\begin{aligned}(u^* - \tilde{u}^{k+1})^T H(u^{k+1} - u^k) =& \frac{1}{2}\Big(\|u^* - u^k\|_H^2 - \|u^* - u^{k+1}\|_H^2\Big) \\ &+ \frac{1}{2}\Big(\|\tilde{u}^{k+1} - u^{k+1}\|_H^2 - \|\tilde{u}^{k+1} - u^k\|_H^2\Big).\end{aligned}$$

Then, the inequality (3.23) can be rewritten as

$$\begin{aligned}\frac{1}{2}\Big(\|u^* - u^k\|_H^2 - \|u^* - u^{k+1}\|_H^2\Big) + \frac{1}{2}\Big(\|\tilde{u}^{k+1} - u^{k+1}\|_H^2 - \|\tilde{u}^{k+1} - u^k\|_H^2\Big) \\ \geq -\frac{L_h}{4}\|\tilde{x}^{k+1} - x^k\|^2.\end{aligned} \tag{3.24}$$

For the last two terms in the left side of the inequality (3.24), we have

$$\begin{aligned}&\|u^k - \tilde{u}^{k+1}\|_H^2 - \|u^{k+1} - \tilde{u}^{k+1}\|_H^2 \\ =\;& \|u^k - \tilde{u}^{k+1}\|_H^2 - \|(u^k - \tilde{u}^{k+1}) - (u^k - u^{k+1})\|_H^2 \\ \stackrel{(3.16)}{=}\;& \|u^k - \tilde{u}^{k+1}\|_H^2 - \|(u^k - \tilde{u}^{k+1}) - M(u^k - \tilde{u}^{k+1})\|_H^2 \\ =\;& 2(u^k - \tilde{u}^{k+1})^T HM(u^k - \tilde{u}^{k+1}) - (u^k - \tilde{u}^{k+1})^T M^T HM(u^k - \tilde{u}^{k+1}) \\ \stackrel{(3.21)}{=}\;& (u^k - \tilde{u}^{k+1})^T(Q^T + Q - M^T HM)(u^k - \tilde{u}^{k+1}) \\ \stackrel{(3.22)}{=}\;& \|u^k - \tilde{u}^{k+1}\|_G^2. \end{aligned} \tag{3.25}$$

Substituting (3.25) into (3.24), we obtain the assertion (3.20). ∎

Theorem 3.3 shows that if the matrix H is positive definite and the term $\|\tilde{u}^{k+1} - u^k\|_G^2 - \frac{L_h}{2}\|\tilde{x}^{k+1} - x^k\|^2$ is positive, the sequence $\{\|u^k - u^*\|_H^2\}$ is non-increasing, guaranteeing convergence of the sequences $\{u^k\}$ to u^*. To ensure $H > 0$ and $\|\tilde{u}^{k+1} - u^k\|_G^2 - \frac{L_h}{2}\|\tilde{x}^{k+1} - x^k\|^2 > 0$, we present the following conditions.

LEMMA 3.4. *For the matrices H and G defined in* (3.21) *and* (3.22)*, respectively, the matrix H is positive definite and the term $\|\tilde{u}^{k+1} - u^k\|_G^2 - \frac{L_h}{2}\|\tilde{x}^{k+1} - x^k\|^2$ is* *positive if the extrapolation parameter* θ*, and the step-size parameters* σ *and* τ*, satisfy the conditions* (1.9)*, i.e.,*

$$\theta \in (-1, 1 - \frac{\tau L_h}{2}), \quad 0 < \tau < \frac{4}{L_h}, \quad \tau\sigma < \frac{1}{\|AA^T\|}.$$

Proof. Expanding $\|\tilde{u}^{k+1} - u^k\|_G^2$ and combining it with $-\frac{L_h}{2}\|\tilde{x}^{k+1} - x^k\|^2$, we obtain

$$\|\tilde{u}^{k+1} - u^k\|_G^2 - \frac{L_h}{2}\|\tilde{x}^{k+1} - x^k\|^2 = \|\tilde{u}^{k+1} - u^k\|_{G_{L_h}}^2, \tag{3.26}$$

where G_{L_h} is given by

$$G_{L_h} = \begin{pmatrix} (\frac{1-\theta}{\tau} - \frac{L_h}{2})I & 0 \\ 0 & \frac{1}{\sigma} - \tau AA^T \end{pmatrix}. \tag{3.27}$$

It is clear when the matrix G_{L_h} is positive definite, the term $\|\tilde{u}^{k+1} - u^k\|_G^2 - \frac{L_h}{2}\|\tilde{x}^{k+1} - x^k\|^2$ is positive. To ensure $G_{L_h} > 0$, the extrapolation parameter θ and regularization parameters τ and σ need to meet

$$\theta < 1 - \frac{\tau L_h}{2}, \quad \frac{1}{\sigma} - \tau\|AA^T\| > 0. \tag{3.28}$$

Define a nonsingular matrix U as follows

$$U = \begin{pmatrix} I & \tau\theta A^T \\ 0 & I \end{pmatrix}.$$

Multiplying H by U^T and U from the left and right, respectively, we obtain the matrix H_U:

$$H_U = U^T H U = \begin{pmatrix} \frac{1}{\tau(1+\theta)} I & 0 \\ 0 & \frac{1}{\sigma} I - \tau\theta AA^T \end{pmatrix}.$$

Since H_U is diagonal, it is positive definite if and only if the following conditions hold:

$$\frac{1}{\tau(1+\theta)} > 0, \quad \frac{1}{\sigma} - \tau\theta\|AA^T\| > 0. \tag{3.29}$$

Combining inequalities (3.28) and (3.29), we deduce that the extrapolation parameter θ must satisfy

$$\theta \in \left(-1, 1 - \frac{\tau L_h}{2}\right).$$

Since $\tau > 0$ and $L_h > 0$, it follows directly that $\theta < 1$. Moreover, the condition $1 - \frac{\tau L_h}{2} > -1$ yields the bound $0 < \tau < \frac{4}{L_h}$.

Given the admissible range of θ, we next derive the corresponding bounds for the step sizes τ and σ. Specifically, the inequality

$$\frac{1}{\sigma} > \tau\|AA^T\| > \theta\tau\|AA^T\|$$

must be satisfied. By combining these conditions, we arrive at the step-size conditions (1.9). ∎

3.3. Global convergence. In this section, we firstly introduce a lemma that is crucial in proving the convergence of SPDA, and then establish SPDA convergence under the condition (1.9).

LEMMA 3.5. *Let the sequence* $\{u^k\}$ *be generated by the scheme* (1.7) *with the condition* (1.9)*. Then, we have*

$$\lim_{k\to\infty} \|x^{k+1} - x^k\| = 0, \tag{3.30a}$$

$$\lim_{k\to\infty} \|y^{k+1} - y^k\| = 0. \tag{3.30b}$$

Proof. Theorem 3.3 shows that for the sequence $\{u^k\}$ generated by the scheme (1.7), the following inequality holds

$$\|u^k - u^*\|_H^2 \geq \|u^{k+1} - u^*\|_H^2 + \|\tilde{u}^{k+1} - u^k\|_G^2 - \frac{L_h}{2}\|\tilde{x}^{k+1} - x^k\|^2.$$

Substituting (3.26) into this inequality, we get

$$\|u^k - u^*\|_H^2 \geq \|u^{k+1} - u^*\|_H^2 + \|\tilde{u}^{k+1} - u^k\|_{G_{L_h}}^2. \tag{3.31}$$

Summing the inequality (3.31) over $k = 0, 1, \ldots, \infty$, we obtain

$$\sum_{k=0}^{\infty} \|\tilde{u}^{k+1} - u^k\|_{G_{L_h}}^2 \leq \|u^0 - u^*\|_H^2,$$

which implies

$$\lim_{k\to\infty} \|\tilde{u}^{k+1} - u^k\|_{G_{L_h}}^2 = 0.$$

Expanding $\|\tilde{u}^{k+1} - u^k\|_{G_{L_h}}^2$ and using the definition (3.3), we obtain

$$\lim_{k\to\infty} \|\tilde{u}^{k+1} - u^k\|_{G_{L_h}}^2 = \lim_{k\to\infty} \left\| \begin{pmatrix} \tilde{x}^{k+1} - x^k \\ \tilde{y}^{k+1} - y^k \end{pmatrix} \right\|_{G_{L_h}}^2 = \lim_{k\to\infty} \left\| \begin{pmatrix} \tilde{x}^{k+1} - x^k \\ y^{k+1} - y^k \end{pmatrix} \right\|_{G_{L_h}}^2.$$

It follows from Lemma 3.4 that when the condition (1.9) holds, the matrix G_{L_h} is positive definite. Therefore, under the conditio (1.9), we obtain

$$\lim_{k\to\infty} \|\tilde{x}^{k+1} - x^k\| = 0, \tag{3.32a}$$

$$\lim_{k\to\infty} \|y^{k+1} - y^k\| = 0. \tag{3.32b}$$

Using the result (3.32b) and passing to the limit in the iterate (1.7d), we can conclude that $\lim_{k\to\infty} x^{k+1} = x^k$. Combining it and (3.32b), we obtain (3.30). ∎

The following result establishes the global convergence of SPDA.

THEOREM 3.6. *For the sequence* $\{u^k\}$ *generated by the scheme* (1.7) *with the condition* (1.9)*, the sequence* $\{u^k\}$ *converges to the optimal point* $u^\infty = u^*$ *for the problem* (1.2).

Proof. Theorem 3.3 and Lemma 3.4 shows the sequence $\{\|u^k - u^*\|_H^2\}$ is non-increasing when the condition (1.9) holds. Therefore, the sequence $\{u^k\}$ is bounded. Let $\{u_s^k\}$ be a subsequence of $\{u^k\}$ that converges to a cluster point $u_s^\infty = (x_s^\infty, y_s^\infty)$.

Substituting (x_s^∞, y_s^∞) into (1.7), we obtain $(\tilde{x}_{s+}^\infty, y_{s+}^\infty, x_{s+}^\infty)$. By (3.30) and (3.32), we obtain

$$\begin{aligned} x_{s+}^\infty &= \tilde{x}_{s+}^\infty = x_s^\infty, \\ y_{s+}^\infty &= y_s^\infty, \end{aligned}$$

which implies $u_s^\infty = u_{s+}^\infty$. Using the definition (3.3), we have $\tilde{y}_{s+}^\infty = y_{s+}^\infty$, thus $\tilde{u}_{s+}^\infty = u_{s+}^\infty$. Substituting $\tilde{u}_{s+}^\infty = (\tilde{x}_{s+}^\infty, \tilde{y}_{s+}^\infty)^T$ into (3.7), we get

$$\begin{cases} f(x) - f(\tilde{x}_{s+}^\infty) + (x - \tilde{x}_{s+}^\infty)^T(A^T\tilde{y}_{s+}^\infty + \nabla h(\tilde{x}_{s+}^\infty)) \geq 0, \\ g^*(y) - g^*(\tilde{y}_{s+}^\infty) - (y - \tilde{y}_{s+}^\infty)^T(A\tilde{x}_{s+}^\infty) \geq 0. \end{cases}$$

Using the notations (3.2), the above inequalities can be rewritten as

$$\mathcal{F}(u) - \mathcal{F}(\tilde{u}_{s+}^\infty) + (u - \tilde{u}_{s+}^\infty)^T\nabla\mathcal{G}(\tilde{u}_{s+}^\infty) \geq 0.$$

By *Lemma 2.1*, we obtain that $\tilde{u}_{s+}^\infty = u_{s+}^\infty = u_s^\infty$ is an optimal point for (1.2). Furthermore, it follows from (3.31) that $\|u^{k+1} - u_s^\infty\|_H^2 \leq \|u^k - u_s^\infty\|_H^2$. This implies that there exists no another cluster point in the sequence $\{u^k\}$. Thus, the sequence $\{u^k\}$ converges to the optimal point $u^* = u_{s+}^\infty = u_s^\infty$. ∎

3.4. Convergence rate. In this section, we establish the non-ergodic convergence rate of the proposed SPDA method, showing that the sequence of iterates achieves a rate of $\mathcal{O}(1/\sqrt{t})$, where t denotes the iteration number. We then demonstrate that the ergodic primal-dual gap converges at a rate of $\mathcal{O}(1/t)$. To facilitate this analysis, we firstly show that $\|M(\tilde{u}^{k+1} - u^k)\|_H$ can serve as a suitable measure of solution accuracy.

The optimality condition (3.9) for the iterates (1.7a) and (1.7c) can be rewritten as

$$\begin{aligned} \mathcal{F}(u) - \mathcal{F}(\tilde{u}^{k+1}) + (u - \tilde{u}^{k+1})^T\nabla\mathcal{G}(\tilde{u}^{k+1}) + (x - \tilde{x}^{k+1})^T(\nabla h(x^k) - \nabla h(\tilde{x}^{k+1})) \\ + (u - \tilde{u}^{k+1})^T Q(\tilde{u}^{k+1} - u^k) \geq 0. \end{aligned}$$

By the relation $H = QM^{-1}$ (see (3.21)), the above inequality is equivalent to

$$\begin{aligned} \mathcal{F}(u) - \mathcal{F}(\tilde{u}^{k+1}) + (u - \tilde{u}^{k+1})^T\nabla\mathcal{G}(\tilde{u}^{k+1}) + (x - \tilde{x}^{k+1})^T(\nabla h(x^k) - \nabla h(\tilde{x}^{k+1})) \\ + (u - \tilde{u}^{k+1})^T HM(\tilde{u}^{k+1} - u^k) \geq 0. \end{aligned} \tag{3.34}$$

Since H is a positive definite matrix and M is a nonsingular matrix, we conclude that $HM(\tilde{u}^{k+1} - u^k) = 0$ and $\tilde{u}^{k+1} = u^k$ when $\|M(\tilde{u}^{k+1} - u^k)\|_H^2 = 0$. In addition, the result $\tilde{x}^{k+1} = x^k$ immediately follows from $\tilde{u}^{k+1} = u^k$. Substituting the conditions $HM(\tilde{u}^{k+1} - u^k) = 0$ and $\tilde{x}^{k+1} = x^k$ into (3.34), we have

$$\mathcal{F}(u) - \mathcal{F}(\tilde{u}^{k+1}) + (u - \tilde{u}^{k+1})^T\nabla\mathcal{G}(\tilde{u}^{k+1}) \geq 0.$$

By Lemma 2.1, this inequality means $\tilde{u}^{k+1}$ is solution of the problem (1.2). Thus, we adopt the term $\|M(\tilde{u}^{k+1} - u^k)\|_H^2$ as a measurement that characterizes the solution precision of $\tilde{u}^{k+1}$ for the problem (1.2).

In the following, we present a lemma which is used to analyze the non-ergodic convergence rate.

LEMMA 3.7. *Let the sequences $\{u^k\}$ and $\{\tilde{u}^k\}$ be generated by the scheme* (1.7). *Then the following inequality holds*

$$(u^{k-1}-\tilde{u}^k)^T M^T H M\Big(u^{k-1}-\tilde{u}^k-(u^k-\tilde{u}^{k+1})\Big) \geq -\frac{L_h}{4}\|\tilde{x}^k-x^{k-1}-\big(\tilde{x}^{k+1}-x^k\big)\|^2 + \frac{1}{2}\|u^{k-1}-\tilde{u}^k-(u^k-\tilde{u}^{k+1})\|^2_{Q^T+Q}. \tag{3.35}$$

Proof. Setting $u=\tilde{u}^k$ in the optimality condition (3.9) yields

$$\mathcal{F}(\tilde{u}^k)-\mathcal{F}(\tilde{u}^{k+1})+(\tilde{u}^k-\tilde{u}^{k+1})^T\nabla F(\tilde{u}^{k+1})+(\tilde{x}^k-\tilde{x}^{k+1})^T\nabla h(x^k) + (\tilde{u}^k-\tilde{u}^{k+1})^T Q(\tilde{u}^{k+1}-u^k)\geq 0. \tag{3.36}$$

Similarly, setting $u=\tilde{u}^{k+1}$ in the optimality condition (3.9) from the previous iterates yields

$$\mathcal{F}(\tilde{u}^{k+1})-\mathcal{F}(\tilde{u}^k)+(\tilde{u}^{k+1}-\tilde{u}^k)^T\nabla F(\tilde{u}^k)+(\tilde{x}^{k+1}-\tilde{x}^k)^T\nabla h(x^{k-1}) + (\tilde{u}^{k+1}-\tilde{u}^k)^T Q(\tilde{u}^k-u^{k-1})\geq 0. \tag{3.37}$$

Adding the inequalities (3.36) and (3.37) and using the fact that $(\tilde{u}^{k+1}-\tilde{u}^k)^T(\nabla F(\tilde{u}^{k+1})-\nabla F(\tilde{u}^k))\geq 0$ yields

$$(\tilde{u}^k-\tilde{u}^{k+1})^T Q\Big(u^{k-1}-\tilde{u}^k-(u^k-\tilde{u}^{k+1})\Big)\geq(\tilde{x}^k-\tilde{x}^{k+1})^T(\nabla h(x^{k-1})-\nabla h(x^k)).$$

This inequality can be further rewritten as

$$(\tilde{u}^k-\tilde{u}^{k+1})^T Q\Big(u^{k-1}-\tilde{u}^k-(u^k-\tilde{u}^{k+1})\Big)\geq(x^{k-1}-x^k)^T(\nabla h(x^{k-1})-\nabla h(x^k)) + \Big(\tilde{x}^k-x^{k-1}-(\tilde{x}^{k+1}-x^k)\Big)^T(\nabla h(x^{k-1})-\nabla h(x^k)). \tag{3.38}$$

Applying Lemma 2.3 to the first term in the right side of the inequality (3.38) yields

$$(x^{k-1}-x^k)^T(\nabla h(x^{k-1})-\nabla h(x^k))\geq\frac{1}{L_h}\|\nabla h(x^{k-1})-\nabla h(x^k)\|^2. \tag{3.39}$$

Applying Cauchy-Schwarz inequality to the last term in the right side of the inequality (3.38) yields

$$\Big(\tilde{x}^k-x^{k-1}-(\tilde{x}^{k+1}-x^k)\Big)^T(\nabla h(x^{k-1})-\nabla h(x^k))\geq-\frac{1}{L_h}\|\nabla h(x^{k-1})-\nabla h(x^k)\|^2 - \frac{L_h}{4}\|\tilde{x}^k-x^{k-1}-(\tilde{x}^{k+1}-x^k)\|^2 \tag{3.40}$$

Combining the inequalities (3.38)-(3.40), we obtain

$$(\tilde{u}^k-\tilde{u}^{k+1})^T Q\Big(u^{k-1}-\tilde{u}^k-(u^k-\tilde{u}^{k+1})\Big)\geq-\frac{L_h}{4}\|\tilde{x}^k-x^{k-1}-\big(\tilde{x}^{k+1}-x^k\big)\|^2. \tag{3.41}$$

Adding the term

$$\Big(u^{k-1}-\tilde{u}^k-(u^k-\tilde{u}^{k+1})\Big)^T Q\Big(u^{k-1}-\tilde{u}^k-(u^k-\tilde{u}^{k+1})\Big)$$

to both sides of (3.41) and using the fact that $u^TQu = \frac{1}{2}u^T(Q^T+Q)u$, we get

$$(u^{k-1}-u^k)^TQ\Big(u^{k-1}-\tilde{u}^k-(u^k-\tilde{u}^{k+1})\Big) \ge -\frac{L_h}{4}\|\tilde{x}^k-x^{k-1}-\big(\tilde{x}^{k+1}-x^k\big)\|^2 + \frac{1}{2}\|u^{k-1}-\tilde{u}^k-(u^k-\tilde{u}^{k+1})\|^2_{Q^T+Q}.$$

It follows from the correction step (3.16) that $(u^{k-1}-u^k)^T = M(u^{k-1}-\tilde{u}^k)$. Substituting it and $H = QM^{-1}$ into the above inequality, we get (3.35). ∎

THEOREM 3.8. *Let the sequences* $\{u^k\}$ *and* $\{\tilde{u}^k\}$ *be generated by the scheme* (1.7) *with the conditions* (1.9)*. Then, the following inequality holds.*

$$\|M(u^t-\tilde{u}^{t+1})\|^2_H \le \frac{1}{tC_0}\|u^0-u^*\|^2_H, \tag{3.42}$$

where t *denotes the iterate and* C_0 *is a positive constant given in the proof.*

Proof. Since H is a positive definite matrix under the condition (1.9), the term $\|M(u^k-\tilde{u}^{k+1})\|^2_H$ satisfies

$$\|M(u^{k-1}-\tilde{u}^k)\|^2_H - \|M(u^k-\tilde{u}^{k+1})\|^2_H = -\|u^{k-1}-\tilde{u}^k-(u^k-\tilde{u}^{k+1})\|^2_{M^THM} +2(u^{k-1}-\tilde{u}^k)^TM^THM\Big(u^{k-1}-\tilde{u}^k-(u^k-\tilde{u}^{k+1})\Big)$$

This identity follows from the fact that $\|a\|^2_H - \|b\|^2_H = 2a^TH(a-b) - \|a-b\|^2_H$. By Lemma 3.7, the above identity can further reduce to

$$\|M(u^{k-1}-\tilde{u}^k)\|^2_H - \|M(u^k-\tilde{u}^{k+1})\|^2_H \ge -\|u^{k-1}-\tilde{u}^k-(u^k-\tilde{u}^{k+1})\|^2_{M^THM} - \frac{L_h}{2}\|\tilde{x}^k-x^{k-1}-\big(\tilde{x}^{k+1}-x^k\big)\|^2 + \|u^{k-1}-\tilde{u}^k-(u^k-\tilde{u}^{k+1})\|^2_{Q^T+Q}.$$

It follows from the definition (3.22) and (3.27) that this inequality is equivalent to

$$\|M(u^{k-1}-\tilde{u}^k)\|^2_H - \|M(u^k-\tilde{u}^{k+1})\|^2_H \ge \|u^{k-1}-\tilde{u}^k-(u^k-\tilde{u}^{k+1})\|^2_{G_{L_h}}. \tag{3.43}$$

Lemma 3.4 shows that the matrix G_{L_h} is positive definite under the condition (1.9). Thus, we conclude $\|M(u^{k-1}-\tilde{u}^k)\|^2_H \ge \|M(u^k-\tilde{u}^{k+1})\|^2_H$. Building upon this result, we have

$$t\|M(u^t-\tilde{u}^{t+1})\|^2_H \le \sum_{k=0}^{t-1}\|M(u^k-\tilde{u}^{k+1})\|^2_H. \tag{3.44}$$

Furthermore, it follows from $G_{L_h} > 0$ and Theorem 3.3 that the following inequality holds

$$\|u^k-u^*\|^2_H \ge \|u^{k+1}-u^*\|^2_H + C_0\|M(\tilde{u}^{k+1}-u^k)\|^2_H,$$

where $C_0 = \frac{\lambda_{\min}(G_{L_h})}{\lambda_{\max}(M^THM)}$. Summing the above inequality over $k = 0, 1, ..., t-1$ yields

$$\|u^t-u^*\|^2_H + C_0\sum_{k=0}^{t-1}\|M(\tilde{u}^{k+1}-u^k)\|^2_H \le \|u^0-u^*\|^2_H. \tag{3.45}$$

Combining the inequalities (3.44) and (3.45), we get

$$\|M(u^t - \tilde{u}^{t+1})\|_H^2 \le \frac{1}{t}\sum_{k=0}^{t-1}\|M(u^k - \tilde{u}^{k+1})\|_H^2 \le \frac{1}{tC_0}\|u^0 - u^*\|_H^2,$$

which proves the assertion (3.42). ∎

Theorem 3.8 shows the worst-case $\mathcal{O}(\frac{1}{\sqrt{t}})$ convergence rate in a non-ergodic sense for SPDA (1.7). In the following, we present the ergodic convergence rate. A crucial lemma is firstly introduced.

LEMMA 3.9. *Let the sequences $\{u^k\}$ and $\{\tilde{u}^k\}$ be generated by the scheme* (1.7)*. The following inequality holds*

$$\|u^k - u^*\|_H^2 \ge \|u^{k+1} - u^*\|_H^2 + \|\tilde{u}^{k+1} - u^k\|_G^2 - L_h\|\tilde{x}^{k+1} - x^k\|^2, \tag{3.46}$$

where H and G are defined by (3.21) *and* (3.22)*, respectively.*

In addition, the sequence $\{\|u^k - u^*\|_H^2\}$ is non-increasing and the term $\|\tilde{u}^{k+1} - u^k\|_G^2 - L_h\|\tilde{x}^{k+1} - x^k\|^2$ is positive if the extrapolation parameter θ, and the step-size parameters σ and τ, satisfy the conditions

$$\theta \in (-1, 1 - \tau L_h), \quad 0 < \tau < \frac{2}{L_h}, \quad \tau\sigma < \frac{1}{\|AA^T\|}. \tag{3.47}$$

Proof. Applying Lemma 2.4 to the last term of the inequality (3.10) yields

$$\begin{aligned}(u - \tilde{u}^{k+1})^T Q(\tilde{u}^{k+1} - u^k) \ge\; & \mathcal{F}(\tilde{u}^{k+1}) - \mathcal{F}(u) + (\tilde{u}^{k+1} - u)^T \nabla F(\tilde{u}^{k+1}) \\ & + (\tilde{x}^{k+1} - x)^T \nabla h(x^k) \\ \ge\; & \mathcal{F}(\tilde{u}^{k+1}) - \mathcal{F}(u) + (\tilde{u}^{k+1} - u)^T \nabla F(\tilde{u}^{k+1}) \\ & + h(\tilde{x}^{k+1}) - h(x) - \frac{L_h}{2}\|\tilde{x}^{k+1} - x^k\|^2.\end{aligned} \tag{3.48}$$

Using the fact that $(\tilde{u}^{k+1} - u)^T \nabla F(\tilde{u}^{k+1}) = (\tilde{u}^{k+1} - u)^T \nabla F(u)$, the above inequality can be rewritten as

$$\begin{aligned}(u - \tilde{u}^{k+1})^T Q(\tilde{u}^{k+1} - u^k) \ge\; & \mathcal{F}(\tilde{u}^{k+1}) - \mathcal{F}(u) + (\tilde{u}^{k+1} - u)^T \nabla F(u) \\ & + h(\tilde{x}^{k+1}) - h(x) - \frac{L_h}{2}\|\tilde{x}^{k+1} - x^k\|^2.\end{aligned} \tag{3.49}$$

Since the function h is convex, it follows from the property of the convex function that

$$h(\tilde{x}^{k+1}) - h(x) \ge (\tilde{x}^{k+1} - x)^T \nabla h(x). \tag{3.50}$$

Combining the inequality (3.50) and the inequality (3.49), we obtain

$$(u - \tilde{u}^{k+1})^T Q(\tilde{u}^{k+1} - u^k) \ge \mathcal{F}(\tilde{u}^{k+1}) - \mathcal{F}(u) + (\tilde{u}^{k+1} - u)^T \nabla \mathcal{G}(u) - \frac{L_h}{2}\|\tilde{x}^{k+1} - x^k\|^2. \tag{3.51}$$

Setting $u = u^*$ and following the similar steps in Theorem 3.3, the inequality (3.51) can be further reduced to (3.46).

Combining the last two terms of the inequality (3.46), we have

$$\|\tilde{u}^{k+1} - u^k\|_G^2 + L_h\|\tilde{x}^{k+1} - x^k\|^2 = \|\tilde{u}^{k+1} - u^k\|_{G'_{L_h}}^2, \tag{3.52}$$

where $G^{'}_{L_h}$ is given by

$$G^{'}_{L_h} = \begin{pmatrix} (\frac{1-\theta}{\tau} - L_h)I & 0 \\ 0 & \frac{1}{\sigma} - \tau AA^T \end{pmatrix}. \tag{3.53}$$

The inequality (3.46) implies that if $H > 0$ and $G^{'}_{L_h} > 0$, the sequence $\{\|u^k - u^*\|\}$ is non-increasing and the term $\|\tilde{u}^{k+1} - u^k\|^2_G - L_h\|\tilde{x}^{k+1} - x^k\|^2$ is positive. Since $G^{'}_{L_h}$ is diagonal, it is positive definite if and only if the following conditions hold:

$$\theta < 1 - \tau L_h, \quad \frac{1}{\sigma} - \tau\|AA^T\| > 0. \tag{3.54}$$

It follows from Lemma 3.4 that the matrix H is positive definite if the inequality (3.29) holds. Combining the inequalities (3.54) and (3.29), we obtain (3.47). ∎

Note that the condition (3.47) is more conservative than the condition (1.9). Building upon the condition (3.47), we present the primal-dual gap convergence rate in an ergodic sense.

THEOREM 3.10. *Let the sequences* $\{u^k\}$ *and* $\{\tilde{u}^k\}$ *be generated by the scheme* (1.7) *with the conditions* (3.47)*. Define the averaged sequence* $\tilde{u}_t = (\tilde{x}_t, \tilde{y}_t)$ *as*

$$\tilde{u}_t = \frac{1}{t+1}\sum_{k=0}^{t} \tilde{u}^k.$$

Then, the following inequality holds

$$\Phi(\tilde{x}_t, y) - \Phi(x, \tilde{y}_t) \le \frac{1}{2(t+1)}\|u - u^0\|^2_H. \tag{3.55}$$

Proof. For the function Φ, we have

$$\begin{aligned}\Phi(\tilde{x}^{k+1}, y) - \Phi(x, \tilde{y}^{k+1}) &= \Phi(\tilde{x}^{k+1}, y) - \Phi(\tilde{x}^{k+1}, \tilde{y}^{k+1}) + \Phi(\tilde{x}^{k+1}, \tilde{y}^{k+1}) - \Phi(x, \tilde{y}^{k+1}) \\ &= \mathcal{F}(\tilde{u}^{k+1}) - \mathcal{F}(u) + (\tilde{u}^{k+1} - u)^T\nabla F(\tilde{u}^{k+1}) + h(\tilde{x}^{k+1}) - h(x).\end{aligned}$$

Using the fact that $(\tilde{u}^{k+1} - u)^T\nabla F(\tilde{u}^{k+1}) = (\tilde{u}^{k+1} - u)^T\nabla F(u)$, we have

$$\begin{aligned}\Phi(\tilde{x}^{k+1}, y) - \Phi(x, \tilde{y}^{k+1}) = \mathcal{F}(\tilde{u}^{k+1}) &- \mathcal{F}(u) + (\tilde{u}^{k+1} - u)^T\nabla F(u) \\ &+ h(\tilde{x}^{k+1}) - h(x).\end{aligned}$$

It follows from (3.49) that the following inequality holds

$$\begin{aligned}\Phi(\tilde{x}^{k+1}, y) - \Phi(x, \tilde{y}^{k+1}) &= \mathcal{F}(\tilde{u}^{k+1}) - \mathcal{F}(u) + (\tilde{u}^{k+1} - u)^T\nabla F(u) + h(\tilde{x}^{k+1}) - h(x) \\ &\le (u - \tilde{u}^{k+1})^T Q(\tilde{u}^{k+1} - u^k) + \frac{L_h}{2}\|\tilde{x}^{k+1} - x^k\|^2.\end{aligned}$$

Substituting (3.16) into the above inequality leads to

$$\Phi(\tilde{x}^{k+1}, y) - \Phi(x, \tilde{y}^{k+1}) \le (u - \tilde{u}^{k+1})^T H(u^{k+1} - u^k) + \frac{L_h}{2}\|\tilde{x}^{k+1} - x^k\|^2. \tag{3.56}$$

Using Lemma 2.2, (3.56) can be simplified to

$$\begin{aligned}\Phi(\tilde{x}^{k+1},y)-\Phi(x,\tilde{y}^{k+1}) &\leq \frac{1}{2}\Big(\|u-u^k\|_H^2-\|u-u^{k+1}\|_H^2+\|\tilde{u}^{k+1}-u^{k+1}\|_H^2\\ &\qquad -\|\tilde{u}^{k+1}-u^k\|_H^2+L_h\|\tilde{x}^{k+1}-x^k\|^2\Big)\\ &\overset{(3.25)}{=} \frac{1}{2}\Big(\|u-u^k\|_H^2-\|u-u^{k+1}\|_H^2-\|u^k-\tilde{u}^{k+1}\|_G^2\\ &\qquad +L_h\|\tilde{x}^{k+1}-x^k\|^2\Big).\end{aligned} \tag{3.57}$$

It follows from Lemma 3.9 that when the conditions (3.47) holds, $\|\tilde{u}^{k+1}-u^k\|_G^2-L_h\|\tilde{x}^{k+1}-x^k\|^2>0$. Therefore, the inequality (3.57) can be further simplified to

$$\Phi(\tilde{x}^{k+1},y)-\Phi(x,\tilde{y}^{k+1})\leq\frac{1}{2}\Big(\|u-u^k\|_H^2-\|u-u^{k+1}\|_H^2\Big). \tag{3.58}$$

Summing the inequality (3.58) over $k=-1,0,\dots,t-1$ and using convexity yields

$$\Phi(\tilde{x}_t,y)-\Phi(x,\tilde{y}_t)\leq\frac{1}{2(t+1)}\|u-u^0\|_H^2.$$

This proves (3.55). □

4. Numerical experiments. In this section, we conduct numerical experiments on the fused lasso problem and the image denoising problem to verify our theoretical results and evaluate the performance of the Symmetric Primal-Dual Algorithm (SPDA). First, we examine how the convergence condition (1.9) affects SPDA's convergence speed. Then, we compare SPDA with other algorithms, including Condat–Vũ [6] and AFBA [18].

4.1. Fused lasso problem. Consider a fused lasso problem as follows

$$\min_x \varphi(x)=\frac{1}{2}\|Kx-b\|_2^2+\mu_1\|x\|_1+\mu_2\|Ax\|_1, \tag{4.1}$$

where μ_1 and μ_2 are regularization parameters, $K\in\Re^{m\times n}$ is a random matrix, $b\in\Re^m$ is a given vector determined by $b=Kx^{'}+\delta$. Here $x'\in\Re^n$ is a vector that contains N non-zero elements, and the values of these non-zero elements range from $\{-4,-3,...,4\}$. δ is a noise vector sampled from the standard normal distribution, $\mathcal{N}(0,0.1)$. $A\in\Re^{n-1\times n}$ is a sparse matrix defined as follows

$$A=\begin{bmatrix}-1 & 1 & & & \\ & -1 & 1 & & \\ & & ... & ... & \\ & & & -1 & 1\end{bmatrix}. \tag{4.2}$$

Let $s=Ax$ and y be a dual vector. The primal Lasso problem (4.1) admits an equivalent min-max problem as follows

$$\min_x\max_y\ \frac{1}{2}\|Kx-b\|_2^2+\mu_1\|x\|_1+y^T(Ax)+\min_s\mu_2\|s\|_1-y^Ts.$$

By the definition of conjugate function, this problem can be further expressed as

$$\min_x\max_y\ \mu_1\|x\|_1+\frac{1}{2}\|Kx-b\|_2^2+y^T(Ax)-I_{\mu_2}(\|y\|_\infty), \tag{4.3}$$

which is equivalent to the form (1.2). $I_{\mu_2}(\|y\|_\infty)$ is the conjugate function of the term $\min_s \mu_2\|s\|_1 - y^T s$ and is defined as

$$I_{\mu_2}(\|y\|_\infty) = \begin{cases} 0, & \text{if } \|y\|_\infty \le \mu_2, \quad (4.4a)\\ \infty, & \text{otherwise .} \quad (4.4b) \end{cases}$$

To conduct proximal mapping in the iterate (1.7a), we use the sign function as an estimate for the subdifferential of the function f:

$$\nabla f(x) = sign(x) = \begin{cases} 1, & \text{if } x > 0, \quad (4.5a)\\ 0, & \text{if } x = 0, \quad (4.5b)\\ -1, & \text{if } x < 0. \quad (4.5c) \end{cases}$$

By (4.5), the iterate (1.7a) can be written as

$$(4.6)\quad \tilde{x}^{k+1} = sign(x^k - \tau\nabla h(x^k) - \tau A^T y^k) \max\left(|x^k - \tau\nabla h(x^k) - \tau A^T y^k| - \mu_1\tau, 0\right).$$

In the numerical experiments, we consider $m = 500, n = 10000, \mu_1 = 20$ and $\mu_2 = 200$. The initial values of the variables in the SPDA are set to $(\tilde{x}^0, \bar{x}^0, y^0, x^0) = (0, 0, 0, 0)$. The maximum iterations of the algorithm are 4000. The algorithm parameters are set to $\lambda = \tau\sigma = \frac{3}{16}$ and $\tau = \frac{1}{2L_h}$. The extrapolation parameter θ is chosen from the set $\{-0.8, -0.5, -0.2, 0, 0.2, 0.5, 0.7\}$. These parameters all satisfy the condition (1.9). To asses the convergence of the SPDA under the condition (1.9), we firstly get the optimal objective value f^* and optimal solution x^* by running the SPDA for 20000 iterations. Then, we compare the relative error values $\frac{\varphi^k - \varphi^*}{\varphi^*}$ and the distance of x^k to the optimal solution x^* after each iteration. These results under the different extrapolation parameters are, respectively, presented in Figure 1 and Figure 2. The results show that the condition (1.9) can accelerate the convergence of the SPDA. Additionally, the convergence of the SPDA get fast, as θ increases, indicating that the large extrapolation parameter can be adopted in practice.

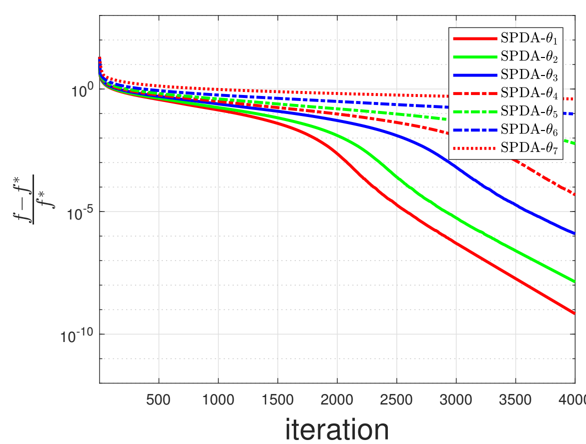

FIG. 1. *Numerical result* $\frac{\varphi^k-\varphi^*}{\varphi^*}$ *of* (1.7) *for* (4.1) *under the different extrapolation parameter:* $\theta_1 = 0.7$, $\theta_2 = 0.5$, $\theta_3 = 0.2$, $\theta_4 = 0$, $\theta_5 = -0.2$, $\theta_6 = -0.5$, $\theta_7 = -0.8$

Based on the results in Figure 1 and Figure 2, the optimal extrapolation parameter is determined to be $\theta = 0.7$. With this parameter fixed along with $\lambda = \frac{3}{16}$.

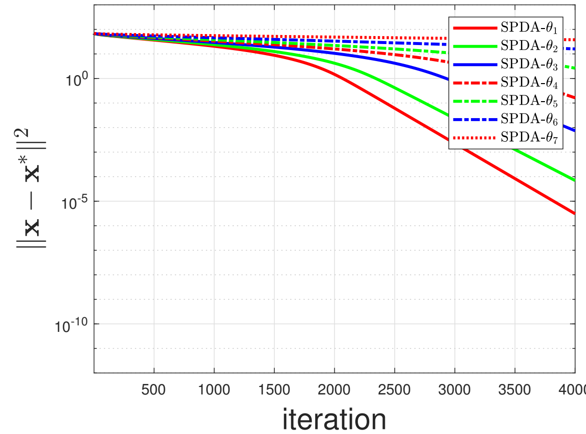

Fig. 2. *Numerical result* $\|x - x^*\|^2$ *of* (1.7) *for* (4.1) *under the different extrapolation parameter:* $\theta_1 = 0.7$*,* $\theta_2 = 0.5$*,* $\theta_3 = 0.2$*,* $\theta_4 = 0$*,* $\theta_5 = -0.2$*,* $\theta_6 = -0.5$*,* $\theta_7 = -0.8$

we investigate the performance of SPDA, AFBA, and Condat–Vũ algorithms under varying step-size parameters $\tau_1 = \frac{1}{2L_h}, \tau_2 = \frac{1}{L_h}, \tau_3 = \frac{3}{2L_h}, \tau_4 = \frac{19}{10L_h}$. It should be noted that the convergence condition (1.4) for the Condat–Vũ algorithm is only satisfied when $\tau = \frac{1}{2L_h}$, limiting its application to cases with larger step sizes. The numerical results $\frac{\varphi^k - \varphi^*}{\varphi^*}$ and $\|x - x^*\|^2$ under the different methods and the different step-size parameter τ are illustrated in Figure 3 and Figure 4. The numerical experiments demonstrate that SPDA outperforms AFBA and Condat–Vũ in convergence speed, requiring fewer iterations to achieve comparable accuracy under identical primal step sizes. Furthermore, increasing the primal step size τ is shown to accelerate convergence for both SPDA and AFBA. These findings highlight the efficiency and robustness of SPDA in solving the fused lasso problem.

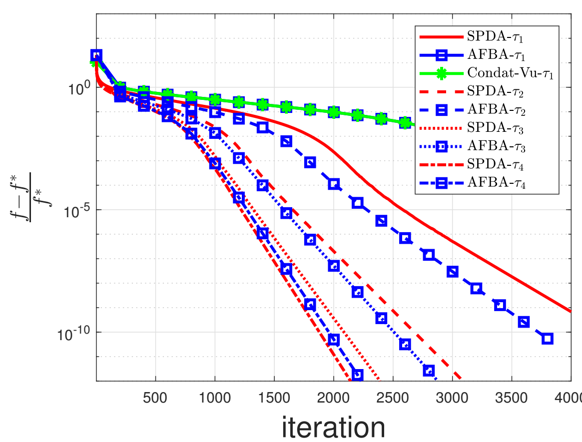

Fig. 3. *Numerical result* $\frac{\varphi^k - \varphi^*}{\varphi^*}$ *of* (1.7) *for* (4.1) *under different methods and the different primal step-size parameter* τ

Finally, we fix $\tau = \frac{1}{2L_h}$ and $\theta = 0.7$, and let $\lambda_1 = \frac{1}{80}$, $\lambda_2 = \frac{1}{8}$ and $\lambda_3 = \frac{1}{4}$. We then investigate the impact of the parameter λ on the convergence of the SPDA. The

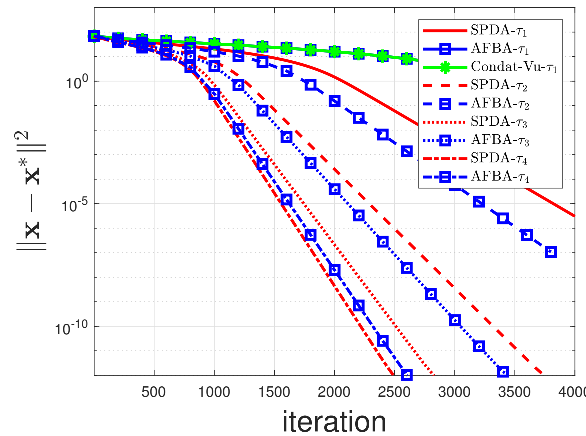

FIG. 4. *Numerical result* $\|x - x^*\|^2$ *of* (1.7) *for* (4.1) *under different methods and the different step-size parameter* τ

relative error values $\frac{\varphi^k - \varphi^*}{\varphi^*}$ and the distance of x^k to the optimal solution x^* under the different λ are shown in Figure 5 and Figure 6. The numerical results indicate that increasing the regularization parameter λ leads to a reduction in the relative error value while having a marginal impact on the squared error term $\|x - x^*\|^2$. Notably, after 2000 iterations, the squared error term for $\lambda = \frac{1}{4}$ exceeds that of $\lambda = \frac{1}{80}$, suggesting that excessively large values of λ do not necessarily improve solution accuracy. These observations imply that selecting a moderately large λ is preferable, as further increasing its magnitude does not bring too much advantage.

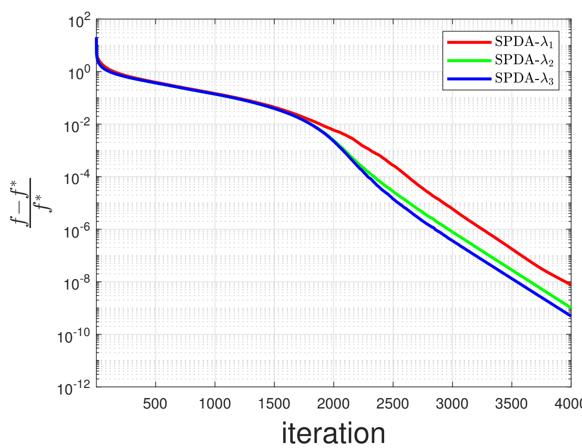

FIG. 5. *Numerical result* $\frac{\varphi^k - \varphi^*}{\varphi^*}$ *of* (1.7) *for* (4.1) *under the different parameter* λ

4.2. Total variation based image denoising. Consider a image denoising problem as follows

$$\min_{x \in [0,1]} \varphi(x) = \frac{1}{2}\|x - b\|_2^2 + \mu \|Ax\|_{TNV}, \tag{4.7}$$

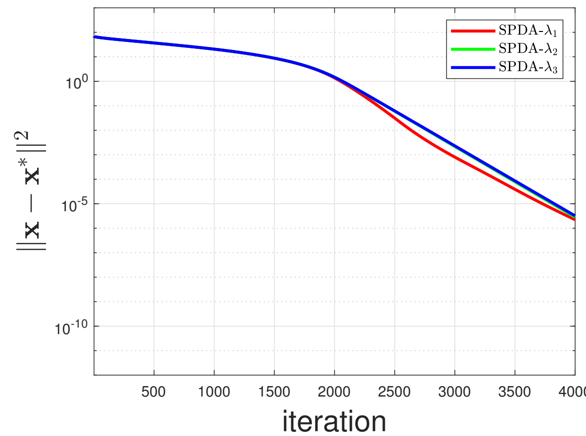

FIG. 6. *Numerical result $\|x - x^*\|^2$ of (1.7) for (4.1) under the different parameter λ*

where b denotes a noisy input image with noise level 0.1 and $\mu = 0.1$ is a regularization parameter. A denotes the 2D discrete gradient operator and is defined as

$$(Ax)_{i,j} = \begin{pmatrix} (Ax)^1_{i,j} \\ (Ax)^2_{i,j} \end{pmatrix} \tag{4.8}$$

where

$$(Ax)^1_{i,j} = \begin{cases} \frac{1}{g}(x_{i+1,j} - x_{i,j}), & \text{if } i < m, \quad (4.9a) \\ 0, & \text{if } i = m, \quad (4.9b) \end{cases}$$

$$(Ax)^2_{i,j} = \begin{cases} \frac{1}{g}(x_{i,j+1} - x_{i,j}), & \text{if } j < n, \quad (4.10a) \\ 0, & \text{if } j = n, \quad (4.10b) \end{cases}$$

m, n denotes the size of the image x and g denotes the gird step size. $\|\cdot\|_{TNV}$ denotes a nuclear norm defined as

$$\|Ax\|_{TNV} = \sum_{i=1}^{m}\sum_{j=1}^{n} \|(Ax)_{i,j}\|_*, \tag{4.11}$$

where $\|(Ax)_{i,j}\|_*$ denotes the nuclear norm and is calculated by the sum of the singular value of the matrix $(Ax)_{i,j}$. For the detailed process, we refer readers to [14]. Define a indicator function $I_{[0,1]}(x)$ as follows

$$I_{[0,1]}(x) = \begin{cases} 0, & \text{if } 0 \geq x \leq 1, \quad (4.12a) \\ \infty, & \text{otherwise .} \quad (4.12b) \end{cases}$$

Let $s = Ax$ and y be a dual vector. Substituting the dual vector and the indicator function into the problem (4.7), the constraint optimization problem (4.7) can be reformulated the unconstraint problem as follows

$$\min_x \max_y \; I_{[0,1]}(x) + \frac{1}{2}\|x - b\|_2^2 + y^T(Ax) + \min_s \mu\|s\|_{TNV} - y^T s,$$

It is noted that it is difficult to calculate the conjugate function of the term $\min_s \mu\|s\|_{TNV} - y^T s$. Thus, we apply the Moreau Identity in [22] to the iteration (1.7c) to obtain

$$y^{k+1} = y^k + \sigma A\bar{x}^{k+1} - \sigma \text{Prox}_{(\frac{1}{\sigma}g)}\big[\frac{1}{\sigma}y^k + A\bar{x}^{k+1}\big], \tag{4.13}$$

which avoids calculating the conjugate function.

In the numerical experiments, $\|AA^T\|$ is estimated to be 8 and the Lipschitz constant L_h is set to 1. For the SPDA method, the initial values of the optimization variables (x, y) are set to zero. The maximum iterations of the algorithm are 200. The step size parameter $\tau\sigma$ is denoted by λ. We employ two quantitative metrics: the relative error $\frac{\varphi^k - \varphi^*}{\varphi^*}$ and the Signal-to-Noise Ratio (SNR) to evaluate the convergence behavior of the algorithm. The optimal objective function φ^* is obtained by running the SPDA for 5000 iterations. The SNR is defined as follow

$$SNR = 20\log_{10}\frac{\|x_o\|_F}{\|x_o - x^k\|_F},$$

where x_o denotes the original image without noisy.

We firstly investigate the impact of the extrapolation parameter θ on accelerating the convergence. We fix the parameter λ and τ at $\frac{1}{8}$ and $\frac{1}{2L_h}$, respectively. The parameter θ is chosen from the set $\{-0.8, -0.5, -0.2, 0, 0.2, 0.5, 0.7\}$, all of which satisfy the convergence condition (1.9). We apply the SPDA methods to the problem (4.7) and plot the results in Figure 7, Figure 8 and Figure 9. Since the output pages under different extrapolation parameters θ look similar, we only present the image for $\theta = 0.7$. The result in Figure 7 indicates that the SPDA method can deal with the image denoising problem. The results in Figures 8 and 9 reveal a dependence of the SPDA algorithm's performance on the extrapolation parameter θ. Specifically, as $\theta \to -1$, the algorithm exhibits substantially degraded convergence behavior, with the relative error value increasing and the SNR showing marked deterioration. Conversely, when $\theta \to 1 - \frac{\tau L_h}{2}$, the method maintains stable convergence behavior. The results implies that for image denoising applications, the extrapolation parameters near the negative extreme $\theta \approx -1$ should be avoided to ensure computational efficiency.

Next, we compare the performance of the SPDA with the other methods, including AFBA and Condat–Vũ method. We fix the parameter θ and λ at 0.7 and 0.1, respectively. The step size parameter τ is chosen from the set $\{\frac{2}{5L_f}, \frac{1}{L_f}, \frac{3}{2L_f}, \frac{\sqrt{10}}{10L_f}\}$. It is noted that the convergence condition of the Condat–Vũ method holds only if $\tau = \frac{2}{5L_f}$. We conduct three algorithms on the problem (4.7) and plot the numerical results $\frac{\varphi^k - \varphi^*}{\varphi^*}$ in Figure 10. The results show that the SPDA present superior numerical performance than other algorithms. In Table 1, we show the iteration numbers (It.) for the different methods when $SNR = 25.40dB$. It can be found that when τ increases, the convergence speed slows down, and the performance differences among various algorithms are not significant. However, when $\tau = \delta$, there exists a significant reduction in the iteration numbers of the SPDA and the AFBA. This implies for the image denoising problem, setting equal step sizes for the primal and dual variables can accelerate the convergence.

Finally, we set $\tau = \frac{1}{2L_h}$ and $\theta = 0.7$, while assigning $\lambda_1 = \frac{1}{8}$, $\lambda_2 = \frac{1}{16}$, $\lambda_3 = \frac{1}{32}$ and $\lambda_4 = \frac{1}{64}$. Subsequently, we analyze the influence of the parameter λ on the convergence behavior of the SPDA. The numerical result $\frac{\varphi^k - \varphi^*}{\varphi^*}$ and SNR are illustrated in Figure 11 and Figure 12, respectively. It is evident that increasing the value of λ leads to

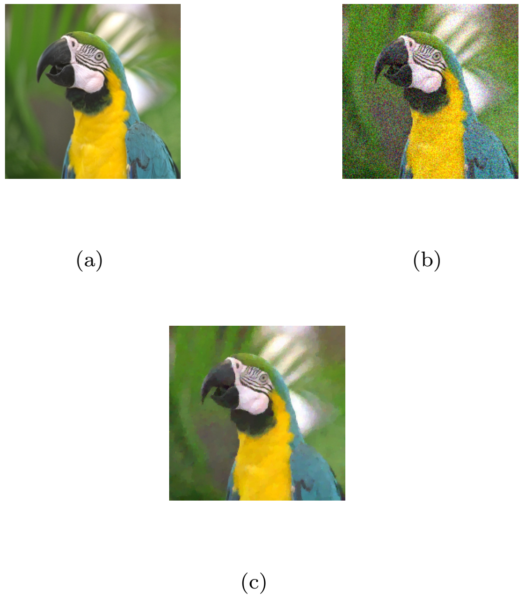

FIG. 7. *Image.(a) Original image. (b) Image with noise. (c) Denoising image*

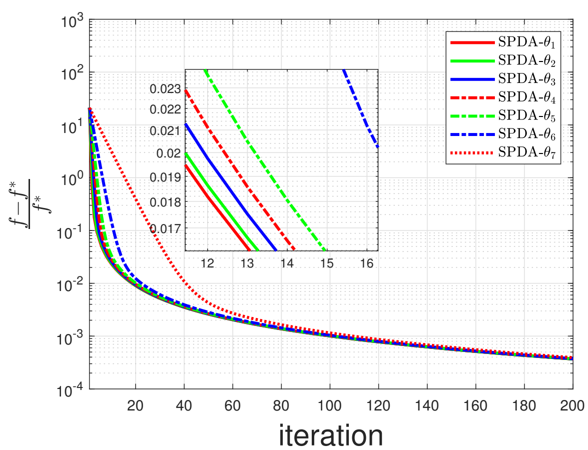

FIG. 8. *Numerical result* $\frac{\varphi^k-\varphi^*}{\varphi^*}$ *of* (1.7) *for* (4.7) *under the different extrapolation parameter:* $\theta_1 = 0.7$, $\theta_2 = 0.5$, $\theta_3 = 0.2$, $\theta_4 = 0$, $\theta_5 = -0.2$, $\theta_6 = -0.5$, $\theta_7 = -0.8$

a reduction in the convergence speed of the algorithm. Consequently, for the image denoising problem, it is advisable to select a moderately large value for λ.

5. Conclusion. In this paper, we proposed a symmetric primal-dual algorithm (SPDA) for solving composite convex optimization problems with a saddle-point structure. Unlike existing methods such as the Condat–Vũ and AFBA algorithms, which apply extrapolation to only one of the primal or dual variables, our approach introduces a fully symmetric update scheme by incorporating extrapolation into both updates. This symmetric structure not only unifies and generalizes these existing

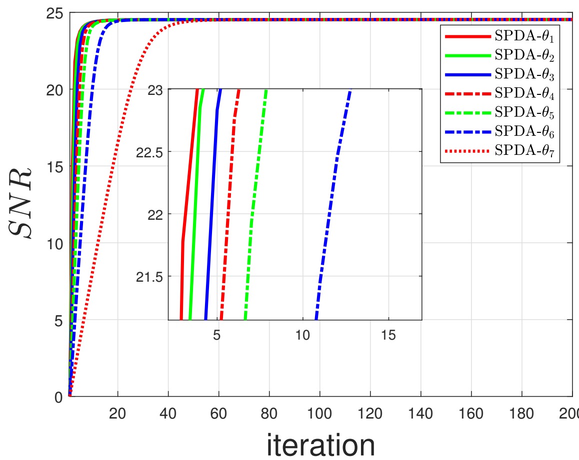

FIG. 9. *Numerical result SNR of* (1.7) *for* (4.7) *under the different extrapolation parameter:* $\theta_1 = 0.7$, $\theta_2 = 0.5$, $\theta_3 = 0.2$, $\theta_4 = 0$, $\theta_5 = -0.2$, $\theta_6 = -0.5$, $\theta_7 = -0.8$

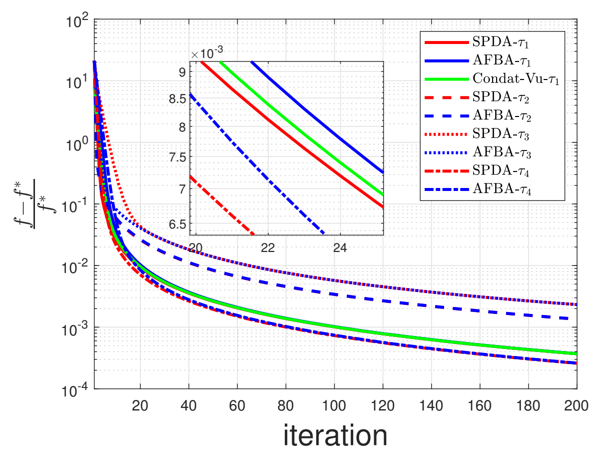

FIG. 10. *Numerical result* $\frac{\varphi^k-\varphi^*}{\varphi^*}$ *of* (1.7) *for* (4.7) *under different methods and the different primal step-size parameter* τ*:* $\tau_1 = \frac{2}{5L_f}, \tau_2 = \frac{1}{L_f}, \tau_3 = \frac{3}{2L_f}, \tau_4 = \frac{\sqrt{10}}{10L_f}$

TABLE 1
Numerical result SNR of (1.7) *for* (4.7) *under different methods and the different primal step-size parameter* τ

Method	τ	It.	Method	τ	It.
SPDA	$\frac{2}{5L_f}$	20	SPDA	$\frac{1}{L_f}$	49
AFBA	$\frac{2}{5L_f}$	24	AFBA	$\frac{1}{L_f}$	50
SPDA	$\frac{3}{2L_f}$	70	SPDA	$\frac{\sqrt{10}}{10L_f}$	14
AFBA	$\frac{3}{2L_f}$	72	AFBA	$\frac{\sqrt{10}}{10L_f}$	21
Condat–Vũ	$\frac{2}{5L_f}$	22			

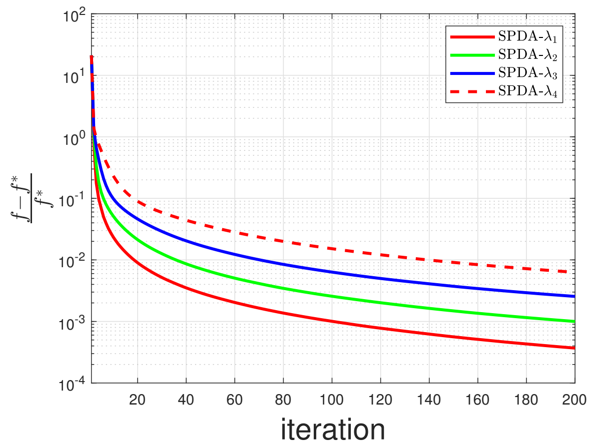

FIG. 11. *Numerical result* $\frac{\varphi^k - \varphi^*}{\varphi^*}$ *of* (1.7) *for* (4.7) *under the different parameter* λ*:* $\lambda_1 = \frac{1}{8}, \lambda_2 = \frac{1}{16}, \lambda_3 = \frac{1}{32}, \lambda_4 = \frac{1}{64}$

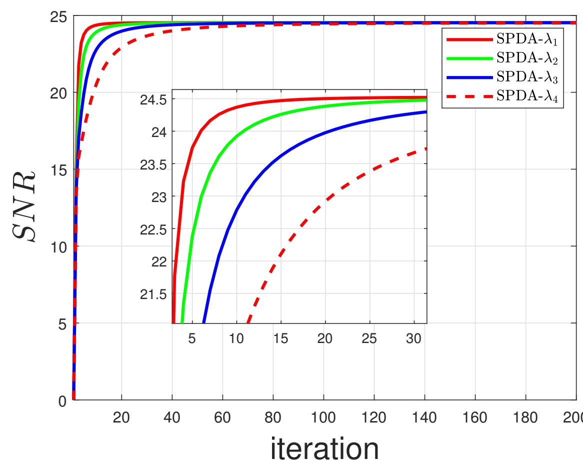

FIG. 12. *Numerical result SNR of* (1.7) *for* (4.7) *under the different parameter* λ*:* $\lambda_1 = \frac{1}{8}, \lambda_2 = \frac{1}{16}, \lambda_3 = \frac{1}{32}, \lambda_4 = \frac{1}{64}$

methods but also improves theoretical convergence properties. We established the global convergence of SPDA under mild step-size conditions and derived its ergodic and non-ergodic convergence rate. The proposed method accommodates a broader range of extrapolation parameters, offering greater flexibility in parameter selection and practical implementation. These findings offer new insights into the role of symmetry in the design of primal-dual algorithms and motivate further research, including the development of adaptive step-size strategies and extensions to more general optimization settings.

REFERENCES

[1] H. BAUSCHKE AND P. L. COMBETTES, *Convex analysis and monotone operator theory in Hilbert spaces*, Springer, 2017.
[2] A. CHAMBOLLE AND T. POCK, *A first-order primal-dual algorithm for convex problems with*

applications to imaging, Journal of mathematical imaging and vision, 40 (2011), pp. 120–145.

[3] A. Chambolle and T. Pock, *An introduction to continuous optimization for imaging*, Acta Numerica, 25 (2016), pp. 161–319.

[4] P. Chen, J. Huang, and X. Zhang, *A primal-dual fixed point algorithm for minimization of the sum of three convex separable functions*, Fixed Point Theory and Applications, 2016 (2016), pp. 1–18.

[5] P. L. Combettes and J.-C. Pesquet, *Proximal splitting methods in signal processing*, Fixed-point algorithms for inverse problems in science and engineering, (2011), pp. 185–212.

[6] L. Condat, *A primal–dual splitting method for convex optimization involving lipschitzian, proximable and linear composite terms*, Journal of optimization theory and applications, 158 (2013), pp. 460–479.

[7] L. Condat, D. Kitahara, A. Contreras, and A. Hirabayashi, *Proximal splitting algorithms for convex optimization: A tour of recent advances, with new twists*, SIAM Review, 65 (2023), pp. 375–435.

[8] D. Davis and W. Yin, *A three-operator splitting scheme and its optimization applications*, Set-valued and variational analysis, 25 (2017), pp. 829–858.

[9] D. Goldfarb, S. Ma, and K. Scheinberg, *Fast alternating linearization methods for minimizing the sum of two convex functions*, Mathematical Programming, 141 (2013), pp. 349–382.

[10] B. He, H. Liu, Z. Wang, and X. Yuan, *A strictly contractive peaceman–rachford splitting method for convex programming*, SIAM Journal on Optimization, 24 (2014), pp. 1011–1040.

[11] B. He, F. Ma, S. Xu, and X. Yuan, *A generalized primal-dual algorithm with improved convergence condition for saddle point problems*, SIAM Journal on Imaging Sciences, 15 (2022), pp. 1157–1183.

[12] B. He, F. Ma, and X. Yuan, *Convergence study on the symmetric version of admm with larger step sizes*, SIAM journal on imaging sciences, 9 (2016), pp. 1467–1501.

[13] B. He and X. Yuan, *On construction of splitting contraction algorithms in a prediction-correction framework for separable convex optimization*, arXiv preprint arXiv:2204.11522, (2022).

[14] K. M. Holt, *Total nuclear variation and jacobian extensions of total variation for vector fields*, IEEE Transactions on Image Processing, 23 (2014), pp. 3975–3989.

[15] F. Jiang, Z. Wu, X. Cai, and H. Zhang, *Unified linear convergence of first-order primal-dual algorithms for saddle point problems*, Optimization Letters, 16 (2022), pp. 1675–1700.

[16] X. Jiang and L. Vandenberghe, *Bregman three-operator splitting methods*, Journal of Optimization Theory and Applications, 196 (2023), pp. 936–972.

[17] N. Komodakis and J.-C. Pesquet, *Playing with duality: An overview of recent primal? dual approaches for solving large-scale optimization problems*, IEEE Signal Processing Magazine, 32 (2015), pp. 31–54.

[18] P. Latafat and P. Patrinos, *Asymmetric forward–backward–adjoint splitting for solving monotone inclusions involving three operators*, Computational Optimization and Applications, 68 (2017), pp. 57–93.

[19] F. Ma, S. Li, and X. Zhang, *A symmetric version of the generalized chambolle-pock-he-yuan method for saddle point problems*, Computational Optimization and Applications, (2025), pp. 1–26.

[20] D. McKenzie, H. Heaton, Q. Li, S. Wu Fung, S. Osher, and W. Yin, *Three-operator splitting for learning to predict equilibria in convex games*, SIAM Journal on Mathematics of Data Science, 6 (2024), pp. 627–648.

[21] B. C. Vũ, *A splitting algorithm for dual monotone inclusions involving cocoercive operators*, Advances in Computational Mathematics, 38 (2013), pp. 667–681.

[22] M. Yan, *A new primal–dual algorithm for minimizing the sum of three functions with a linear operator*, Journal of Scientific Computing, 76 (2018), pp. 1698–1717.