*mathematics*

MDPI

*Article*

# Relaxed Variable Metric Primal-Dual Fixed-Point Algorithm with Applications

**Wenli Huang [1], Yuchao Tang [2], Meng Wen [3] and Haiyang Li [1,*]**

[1] School of Mathematics and Information Science, Guangzhou University, Guangzhou 510006, China
[2] Department of Mathematics, Nanchang University, Nanchang 330031, China
[3] School of Science, Xi'an Polytechnic University, Xi'an 710048, China
* Correspondence: haiyangli@gzhu.edu.cn

**Abstract:** In this paper, a relaxed variable metric primal-dual fixed-point algorithm is proposed for solving the convex optimization problem involving the sum of two convex functions where one is differentiable with the Lipschitz continuous gradient while the other is composed of a linear operator. Based on the preconditioned forward–backward splitting algorithm, the convergence of the proposed algorithm is proved. At the same time, we show that some existing algorithms are special cases of the proposed algorithm. Furthermore, the ergodic convergence and linear convergence rates of the proposed algorithm are established under relaxed parameters. Numerical experiments on the image deblurring problems demonstrate that the proposed algorithm outperforms some existing algorithms in terms of the number of iterations.

**Keywords:** primal-dual; variable metric; proximity operator; total variation

**MSC:** 65K05; 68Q25; 68U10



**Citation:** Huang, W.; Tang, Y.; Wen, M.; Li, H. Relaxed Variable Metric Primal-Dual Fixed-Point Algorithm with Applications. *Mathematics* **2022**, *10*, 4372. https://doi.org/10.3390/math10224372

Academic Editors: Kuo Tian, Weizhu Yang and Shiyao Lin

Received: 14 September 2022
Accepted: 17 November 2022
Published: 20 November 2022



## 1. Introduction

In this paper, we focus on the following convex optimization problem:

$$\min_{x \in H} f(x) + h(Lx), \tag{1}$$

where $f : H \to R$ is convex differentiable and its gradient $\nabla f$ is $\frac{1}{\beta}-$Lipschitz-continuous for some $\beta > 0$, $h : G \to (-\infty, +\infty]$ is a proper lower semi-continuous convex function, $L : H \to G$ is a bounded linear operator, and $H$ and $G$ are real Hilbert spaces. This problem is widely used in signal and image processing [1,2], compressed sensing [3], and machine learning [4]. For instance, a classical model in image restoration and medical image reconstruction is:

$$\min_{x \in R^n} \frac{1}{2}\|Ax - b\|^2 + \mu\|x\|_{TV}, \tag{2}$$

where $A : R^n \to R^m$ is a blurring operator, $b \in R^m$ is the observed image, $\mu > 0$ is the regularization parameter, and $\|x\|_{TV}$ is the total variation, which can be represented by a composition of a convex function with a discrete gradient operator.

The corresponding dual problem of (1) is

$$\max_{v \in G} -f^*(-L^*v) - h^*(v), \tag{3}$$

and the associated saddle point problem of (1) and (3) is

$$\min_{x \in H} \max_{v \in G} \{K(v, x) = f(x) + \langle Lx, v\rangle - h^*(v)\}. \tag{4}$$

We say that $(x^*, v^*)$ is a saddle point of (4) if and only if $x^*$ is a solution of (1) and $v^*$ is a solution of (3), respectively. The optimal solution set of problem (4) is denoted by $\Omega$. In this paper, we always assume that $\Omega$ is nonempty.

Many efficient algorithms have been proposed for solving problem (1) in the last decades. Most of them are based on the alternating direction method of multipliers (ADMM) [5–7] and the forward–backward splitting (FBS) algorithm [8]. In [9,10], the authors showed that ADMM is equivalent to the Douglas–Rachford splitting algorithm [8]. The proximal gradient algorithm (PGA, also known as FBS) [11] is an efficient algorithm to solve (1) if $L = I$, and some accelerated versions of PGA had been studied [12–14]. The primal dual hybrid gradient algorithm [15,16] was proposed to solve (1) without the smoothness of $f$. Combining the FBS algorithm [11] with the fixed-point algorithm based on the proximity operator (FP$^2$O) [17], Argyriou et al. [18] proposed a FBS_FP$^2$O to solve (1). Note that the FBS_FP$^2$O algorithm needs to solve a subproblem. Thus, it involves inner and outer iterations. To avoid choosing the number of inner iterations, Chen et al. [19] first a proposed the primal-dual fixed-point algorithm based on the proximity operator (PDFP$^2$O) to solve (1). Compared with the FBS_FP$^2$O, PDFP$^2$O only performs one inner iteration and reduces to the generalized iterative soft-thresholding algorithm [20] when $f = \|Ax - b\|^2$. The PDFP$^2$O provided desirable performances to solve MRI reconstruction and TV-L1 wavelet inpainting [21,22]. In contrast, Combettes et al. [23] proposed a variable metric forward–backward splitting (VMFBS) algorithm to solve the saddle-point problem (4). By choosing a special variable metric, the PDFP$^2$O could be recovered by the VMFBS algorithm. Moreover, the proximal alternating predictor-corrector (PAPC) algorithm [24] was proposed to solve the equivalent minimization problem of (1) and was proved to converge linearly [25]. To speed up the PDFP$^2$O, Chen et al. [26] proposed an adapted metric version of PDFP$^2$O, which is termed as PDFP$^2$O_AM. The key feature of the PDFP$^2$O_AM is that it uses a symmetric positive matrix to replace the stepsize of the PDFP$^2$O. In contrast, Wen et al. [27] generalized the stepsize in the PDFP$^2$O to the dynamic stepsize. Later, a larger stepsize of PDFP$^2$O was proved [28]. Recently, Zhu and Zhang [29] introduced an inertial PDFP$^2$O (IPDFP$^2$O).

In Table 1, we summarize some variants of PDFP$^2$O.

**Table 1.** Listing of existing primal-dual fixed point type algorithms.

| Algorithm | $\rho_k$ | $\lambda$ | $\gamma$ | Variable Metric | Ergodic Rate | Linear Rate |
|---|---|---|---|---|---|---|
| PDFP$^2$O [19] | $(0, 1]$ | $(0, \frac{1}{\lambda_{\max}(LL^*)}]$ | $(0, 2\beta)$ | | | √ |
| VMPD [23] | $(0, 1]$ | $(0, \frac{1}{\lambda_{\max}(LQ_k^{-1}L^*)})$ | $(0, 2\beta)$ | √ | | |
| PAPC [24] | 1 | $(0, \frac{1}{\lambda_{\max}(LL^*)}]$ | $(0, \beta]$ | | √ | √ |
| PDFP$^2$O_AM [26] | $(0, 1]$ | $(0, \frac{1}{\lambda_{max}(LQ^{-1}L^*)})$ | $(0, 2\beta)$ | | | |
| PDFP$^2$O_DS [27] | $(0, 1)$ | $(0, \frac{1}{\lambda_{\max}(LL^*)})$ | $(0, 2\beta)$ | √ | | √ |
| IPDFP$^2$O [29] | 1 | $(0, \frac{1}{\lambda_{\max}(LL^*)})$ | $(0, 2\beta)$ | | | |

From Table 1, we note that the relaxation parameter of these algorithms belongs to $(0, 1]$. It is well known that the convergence speed of the iterative algorithm can be accelerated when the relaxation parameter is greater than 1. This allows us to accelerate the PDFP$^2$O_AM with larger relaxed parameters. We reformulate the PDFP$^2$O_AM as the FBS algorithm and propose a primal dual fixed-point algorithm based on the proximity operator with relaxed parameters and variable metrics (Rv_PDFP$^2$O). Based on the fixed point theory, we prove the convergence of the proposed algorithm. At the same time, we point out that PDFP$^2$O_AM [26], PDFP$^2$O_DS [27], and PDFP$^2$O [19] are particular cases of Rv_PDFP$^2$O. Further, the convergence rates are established under the larger relaxed parameters, including ergodic and linear convergence. To verify the effectiveness and superiority of Rv_PDFP$^2$O, we apply it for solving the image-restoration problem and compare it with other algorithms.

The rest of the paper is organized as follows. In Section 2, we recall some preliminaries and related work. In Section 3, we deduce the Rv_PDFP$^2$O from the preconditioned FBS algorithm and provide some convergence results. In Section 4, we show the numerical results of Rv_PDFP$^2$O on solving image deblurring problem. Finally, we provide the conclusions.

## 2. Preliminaries and Related Work

In this section, we first provide some notations and definitions. Then, we briefly review some existing algorithms for solving (1).

Throughout this paper, $H$ denotes a real Hilbert space endowed with scalar product $\langle\cdot,\cdot\rangle$, and the associated norm is $\|\cdot\|$. Let $S^H_{++}$ ($S^H_{+}$) denote the set of the symmetric positive definite (semi-definite) operator in $H$. For $U \in S^H_{++}$, the $U$-weighted inner product is $\langle x,y\rangle_U = \langle x, Uy\rangle$ and the corresponding $U$-weighted norm is defined by $\|x\|_U = \sqrt{\langle x,x\rangle_U}$. $H_1$ and $H_2$ are the real Hilbert space endowed with the scalar product. Let $U_1 \in S^{H_1}_{++}$ and $U_2 \in S^{H_2}_{++}$, the $U_1, U_2$-weighted norm in $H_1 \times H_2$ is $\|x\|_{U_1,U_2} = \sqrt{\|x_1\|^2_{U_1} + \|x_2\|^2_{U_2}}$ for $x = (x_1, x_2) \in H_1 \times H_2$. We denote by $\Gamma_0(H)$ the class of all proper, lower semi-continuous, convex functions from $H$ to $(-\infty, +\infty]$. Most of these definitions can be found in [8].

Let $A : H \to 2^H$ be a set-valued operator. The domain, the graph, the zeros, and the inverse of $A$ are represented by $domA = \{x \in H : Ax \neq \emptyset\}$ , $graA = \{(x,u) \in H \times H : u \in Ax\}$ and $zerA = \{x \in H : 0 \in Ax\}$, $A^{-1} : H \mapsto 2^H : u \mapsto \{x \in H : u \in Ax\}$, and the resolvent of $A$ is

$$J_A = (I + A)^{-1}.$$

The operator $A : H \to 2^H$ is monotone, if $\langle u - v, x - y\rangle \geq 0$ for all $(x,u) \in graA$, $(y,v) \in graA$. The monotone operator $A$ is maximally monotone if there is no monotone operator $B$ such that $graA \subseteq graB \neq graA$. Further, $A$ is $\delta$- monotone if $\langle u - v, x - y\rangle \geq \delta\|x - y\|^2$ for $\delta > 0$, $\forall(x,u),(y,v) \in graA$. An operator $B : H \to H$ is $\beta$-cocoercive, for some $\beta > 0$, if $\langle x - y, Bx - By\rangle \geq \beta\|Bx - By\|^2, \forall x, y \in H$.

Let $D \subseteq H$ be nonempty and let $T : D \to H$. The fixed point set of $T$ is denoted by $FixT$, i.e., $FixT = \{x \in D : Tx = x\}$. $T$ is $\alpha$-averaged, for some $\alpha \in (0,1)$, if

$$\|Tx - Ty\|^2 + \frac{1-\alpha}{\alpha}\|(I - T)x - (I - T)y\|^2 \leq \|x - y\|^2, \quad \forall x, y \in D.$$

Let $f \in \Gamma_0(H)$, the Fenchel conjugate of $f$ is

$$f^*(u) = \sup_{x \in H}\{\langle x,u\rangle - f(x)\},$$

and the subdifferential of $f$ is the maximally monotone operator

$$\partial f : x \mapsto \{u \in H : \langle y - x, u\rangle + f(x) \leq f(y), \forall y \in H\}.$$

Further, $\partial f(x) = \{\nabla f(x)\}$ when $f$ is differentiable.

Let $f \in \Gamma_0(H)$ and $U \in S^H_{++}$, the scale proximity operator of $f$ with respect to the metric $U$ is

$$prox^U_f(x) = \arg\min_{u \in H}\left\{\frac{1}{2}\|u - x\|^2_U + f(u)\right\}. \tag{5}$$

The scale proximity operator is the standard proximity operator when $U = I$.

*Related Work*

To solve (1), Argyriou et al. [18] considered the following forward–backward splitting algorithm:

$$x^{k+1} = prox_{\gamma(h \circ L)}(x^k - \gamma\nabla f(x^k)), \tag{6}$$

where $\gamma \in (0, 2\beta)$. By the definition of proximity operator, (6) is equivalent to

$$x^{k+1} = \arg\min_{x} \left\{ \frac{1}{2}\|x - (x^k - \gamma\nabla f(x^k))\|^2 + \gamma h(Lx) \right\}. \tag{7}$$

Then, Argyriou et al. [18] employed the FBS_FP$^2$O algorithm to solve (7):

$$\begin{cases} v^{k+1} = \arg\min_{v} \left\{ \frac{1}{2}\|L^* v - \frac{1}{\gamma}(x^k - \gamma\nabla f(x^k))\|^2 + \frac{1}{\gamma}h^*(v) \right\}, \\ x^{k+1} = x^k - \gamma\nabla f(x^k) - \gamma L^* v^{k+1}. \end{cases} \tag{8}$$

In (8), one needs to solve the subproblem of $v$ to obtain the update of $\{v^{k+1}\}$. More precisely, we obtain the following inner–outer iterative algorithm:

$$\begin{cases} v^{k+1,j} = prox_{\frac{\lambda}{\gamma}h^*}(v^{k,j} - \lambda L(L^* v^{k,j} - \frac{1}{\gamma}(x^k - \gamma\nabla f(x^k)))), \\ x^{k+1} = x^k - \gamma\nabla f(x^k) - \gamma L^* v^{k+1,J}, \end{cases} \tag{9}$$

where $\lambda \in (0, \frac{2}{\lambda_{max}(LL^*)})$, $j$ denotes the inner iteration, and $J$ represents the maximum number of inner iteration. Here, $\lambda_{max}(L)$ denotes the largest eigenvalue of L when L is a matrix.

Chen et al. [19] proposed the PDFP$^2$O as follows:

$$\begin{cases} \tilde{v}^{k+1} = (I - prox_{\frac{\gamma}{\lambda}h})(L(x^k - \gamma\nabla f(x^k)) + (I - \lambda LL^*)v^k), \\ \tilde{x}^{k+1} = x^k - \gamma\nabla f(x^k) - \lambda L^*\tilde{v}^{k+1}, \\ (v^{k+1}, x^{k+1}) = (1-\rho_k)(v^k, x^k) + \rho_k(\tilde{v}^{k+1}, \tilde{x}^{k+1}). \end{cases} \tag{10}$$

where $\lambda \in (0, \frac{1}{\|L\|^2}]$, $\gamma \in (0, 2\beta)$, and $\rho_k \in (0, 1]$. Let $\rho_k = 1$, and with the help of the Moreau decomposition, we obtain from (10) that

$$\begin{cases} v^{k+1} = \frac{\gamma}{\lambda} prox_{\frac{\lambda}{\gamma}h^*}(\frac{\lambda}{\gamma}L(x^k - \gamma\nabla f(x^k)) + \frac{\lambda}{\gamma}(I - \lambda LL^*)v^k), \\ x^{k+1} = x^k - \gamma\nabla f(x^k) - \lambda L^* v^{k+1}. \end{cases} \tag{11}$$

Let $\overline{v}^k = \frac{\lambda}{\gamma}v^k$, we have

$$\begin{cases} \overline{v}^{k+1} = prox_{\frac{\lambda}{\gamma}h^*}(\frac{\lambda}{\gamma}L(x^k - \gamma\nabla f(x^k)) + (I - \lambda LL^*)v^k), \\ x^{k+1} = x^k - \gamma\nabla f(x^k) - \gamma L^*\overline{v}^{k+1}. \end{cases} \tag{12}$$

Compared with FBS_FP$^2$O (9), PDFP$^2$O (12) performs only one inner iteration to calculate $v^{k+1}(\overline{v}^{k+1})$. On the other hand, if we add a proximal term $\frac{1}{2}\|v - v^k\|^2_{\frac{1}{\lambda}I - LL^*}$ to the subproblem of $v$ in (8), i.e.,

$$v^{k+1} = \arg\min_{v} \left\{ \frac{1}{2}\|L^* v - \frac{1}{\gamma}(x^k - \gamma\nabla f(x^k))\|^2 + \frac{1}{\gamma}h^*(v) + \frac{1}{2}\|v - v^k\|^2_{\frac{1}{\lambda}I - LL^*} \right\},$$

after simple calculation, we could also recover the PDFP$^2$O (12).

## 3. Relaxed Variable Metric Primal-Dual Fixed-Point Algorithm Based on Proximity Operator

In this section, we propose Rv_PDFP$^2$O for solving the minimization problem (1). The Rv_PDFP$^2$O is

$$\begin{cases} \tilde{v}^{k+1} = prox_{h^*}^{P_k}((I - P_k^{-1}LQ_k^{-1}L^*)v^k + P_k^{-1}L(x^k - Q_k^{-1}\nabla f(x^k))), \\ \tilde{x}^{k+1} = x^k - Q_k^{-1}\nabla f(x^k) - Q_k^{-1}L^*\tilde{v}^{k+1}, \\ (v^{k+1}, x^{k+1}) = (1-\rho_k)(v^k, x^k) + \rho_k(\tilde{v}^{k+1}, \tilde{x}^{k+1}). \end{cases} \tag{13}$$

### 3.1. Convergence Analysis

First, let us introduce the product space $K = G \times H$ and define the operators:

$$A : K \to 2^K\, (v, x) \mapsto (\partial h^*(v) - Lx) \times (L^*v),$$

and

$$B : K \to K\, (v, x) \mapsto (0, \nabla f(x)).$$

Notice that $u^* = (v^*, x^*) \in \Omega$ if and only if $0 \in Au^* + Bu^*$. Although $A$ is maximally monotone and $B$ is cocoercive in $K$, the forward–backward splitting algorithm could not be applicable since $(I + \tau A)^{-1}$, $\tau > 0$ does not have a closed-form solution. To overcome this difficulty, we consider a preconditioned forward–backward splitting algorithm as follows:

$$u^{k+1} = (1-\rho_k)u^k + \rho_k J_{U_kA}(u^k - U_kBu^k), \tag{14}$$

where $P_k \in S_{++}^G$, $Q_k \in S_{++}^H$, $\rho_k > 0$ and

$$U_k^{-1} = \begin{pmatrix} P_k - LQ_k^{-1}L^* & 0 \\ 0 & Q_k \end{pmatrix}. \tag{15}$$

After simple calculation, we recover (13) from (14). In order to analyze the theoretical convergence of Rv_PDFP$^2$O, we make the following assumptions:

**(A1)**: $\|P_k^{-1}\|_2 \in (0, \frac{1}{\lambda_{\max}(LQ_k^{-1}L^*)})$, $\|Q_k^{-1}\|_2 \in (0, 2\beta)$;
**(A2)**: $\rho_k \in [\underline{\rho}, \frac{4\beta - \|Q_k^{-1}\|_2}{2\beta} - \xi_k]$, for $\underline{\rho}, \xi_k > 0$;
**(A3)**: $U_{k+1} - U_k \in S_{++}^K, \forall k \in \mathbb{N}$;
**(A4)**: $\overline{\vartheta} = \sup_{k\in\mathbb{N}} \|U_k\|_2 < +\infty$.

Under the assumption (A1), we have $U_k^{-1} \in S_{++}^K$. Denote $J_{U_kA}(I - U_kB)$ by $\hat{T}_{(k)}$.

**Lemma 1.** *Suppose that (A1) holds. Then the following statements hold:*
(1) *$I - U_kB$ is $\frac{\|Q_k^{-1}\|_2}{2\beta}$-averaged under $\|\cdot\|_{U_k^{-1}}$;*
(2) *$\hat{T}_{(k)}$ is $\frac{2\beta}{4\beta - \|Q_k^{-1}\|_2}$-averaged under $\|\cdot\|_{U_k^{-1}}$.*

**Proof.** (1) Let $u_1 = (v_1, x_1), u_2 = (v_2, x_2) \in K$; we have

$$\begin{aligned} &\langle U_kBu_1 - U_kBu_2, u_1 - u_2\rangle_{U_k^{-1}} \\ =&\langle \nabla f(x_1) - \nabla f(x_2), x_1 - x_2\rangle \\ \geq&\beta\|\nabla f(x_1) - \nabla f(x_2)\|^2 \\ \geq&\frac{\beta}{\|Q_k^{-1}\|_2}\|Q_k^{-1}\nabla f(x_1) - Q_k^{-1}\nabla f(x_2)\|_{Q_k}^2 \\ =&\frac{\beta}{\|Q_k^{-1}\|_2}\|U_kBu_1 - U_kBu_2\|_{U_k^{-1}}^2, \end{aligned}$$

which means that $U_k B$ is $\frac{\beta}{\|Q_k^{-1}\|_2}$-cocoercive. Hence, $I - U_k B$ is $\frac{\|Q_k^{-1}\|_2}{2\beta}$-averaged.

(2) Since $U_k \in S_{++}^K$, $U_k A$ is maximally monotone and $J_{U_k A}$ is $\frac{1}{2}$-averaged. Hence, $J_{U_k A}(I - U_k B)$ is $\frac{2\beta}{4\beta - \|Q_k^{-1}\|_2}$-averaged. □

Now, we are ready to present the main convergence theorem of Rv_PDFP$^2$O (13).

**Theorem 1.** *Suppose that (A1)–(A4) hold. Let $\{u^k = (v^k, x^k)\}$ be generated by (13). Then, we have the following:*
(1) *For any $u^* \in \Omega$, $\{\|u^k - u^*\|_{U_k^{-1}}\}$ is monotonically decreasing and $\lim_{k\to+\infty} \|u^k - u^*\|_{U_k^{-1}}$ exists;*
(2) $\lim_{k\to+\infty} \|u^k - \hat{T}_{(k)}(u^k)\| = 0$;
(3) $\{u^k\}$ *converges weakly to a point in $\Omega$.*

**Proof.** (1) Let $\alpha_k = \frac{2\beta}{4\beta - \|Q_k^{-1}\|_2}$. Notice that $\|u\|_{U_{k+1}^{-1}} \le \|u\|_{U_k^{-1}}, \forall u \in G \times H$. Then, we obtain

$$
\begin{aligned}
&\|u^{k+1} - u^*\|^2_{U_{k+1}^{-1}} \\
\le&\|u^{k+1} - u^*\|^2_{U_k^{-1}} \\
=&(1-\rho_k)\|u^k - u^*\|^2_{U_k^{-1}} + \rho_k\|\hat{T}_{(k)}(u^k) - u^*\|^2_{U_k^{-1}} - \rho_k(1-\rho_k)\|u^k - \hat{T}_{(k)}(u^k)\|^2_{U_k^{-1}} \\
\le&\|u^k - u^*\|^2_{U_k^{-1}} - \rho_k\left(\frac{1}{\alpha_k} - \rho_k\right)\|u^k - \hat{T}_{(k)}(u^k)\|^2_{U_k^{-1}},
\end{aligned}
\tag{16}
$$

which implies that $\|u^k - u^*\|_{U_k^{-1}}$ is decreasing and $\lim_{k\to+\infty} \|u^k - u^*\|_{U_k^{-1}}$ exists.

(2) Summing (16) from $k = 0$ to $N-1$, we obtain

$$
\sum_{k=0}^{N-1} \rho_k\left(\frac{1}{\alpha_k} - \rho_k\right)\|u^k - \hat{T}_{(k)}(u^k)\|^2_{U_k^{-1}} \le \|u^0 - u^*\|_{U_0^{-1}}. \tag{17}
$$

It follows from (17) that

$$
\lim_{k\to+\infty} \|u^k - \hat{T}_{(k)}(u^k)\| = 0. \tag{18}
$$

(3) Let $\{u^{k_j}\} \subset \{u^k\}$ such that $u^{k_j} \rightharpoonup \widehat{u}^*$. It follows from Lemma 2.3 in [30] that there is $U^{-1} \in S_{++}^K$ such that $U_k^{-1} \to U^{-1}$. Define $T = J_{UA}(I - UB)$, we have

$$
\begin{aligned}
&\|u^{k_j} - T(u^{(k_j)})\| \\
\le&\|u^{k_j} - \hat{T}_{(k_j)}(u^{k_j})\| + \|\hat{T}_{(k_j)}(u^{k_j}) - T(u^{k_j})\| \\
\le&\|u^{k_j} - \hat{T}_{(k_j)}(u^{k_j})\| + \frac{1}{\lambda_{min}(U_{k_j}^{-1})}\|(U_{k_j}^{-1} - U^{-1})(u^{k_j} - T(u^{k_j})\| \\
\le&\|u^{k_j} - \hat{T}_{(k_j)}(u^{k_j})\| + \overline{\vartheta}\|(U_{k_j}^{-1} - U^{-1})(u^{k_j} - T(u^{k_j})\|,
\end{aligned}
\tag{19}
$$

which implies that $\lim_{j\to\infty} \|u^{k_j} - T(u^{k_j})\| = 0$. The second inequality in (19) holds by Lemma 3.4 of [31]. It follows from the demiclosedness of $T$ that $\widehat{u}^* \in \Omega$. By Opial's lemma, we conclude that $u^k \rightharpoonup \widehat{u}^*$. This completes the proof. □

### *3.2. Connections to Existing Algorithms*

In this subsection, we present a series of special cases of the proposed algorithm and point out connections to other existing algorithms.

(i) Let $P_k = \frac{1}{\lambda} I$, $Q_k = Q$; then, (13) reduces to the PDFP$^2$O_AM [26].

$$\begin{cases} \tilde{v}^{k+1} = prox_{\lambda h^*}(\frac{1}{\lambda} L(x^k - Q^{-1}\nabla f(x^k)) + (I - \lambda L Q^{-1} L^*) v^k), \\ \tilde{x}^{k+1} = x^k - Q^{-1}\nabla f(x^k) - Q^{-1} L^* \tilde{v}^{k+1}, \\ (v^{k+1}, x^{k+1}) = (1-\rho_k)(v^k, x^k) + \rho_k(\tilde{v}^{k+1}, \tilde{x}^{k+1}). \end{cases} \tag{20}$$

(ii) Let $P_k = \frac{\gamma_k}{\lambda_k} I$ and $Q_k = \frac{1}{\gamma_k} I$ ; then, we obtain from (13) that

$$\begin{cases} \tilde{v}^{k+1} = prox_{\frac{\lambda_k}{\gamma_k} h^*}(\frac{\lambda_k}{\gamma_k} L(x^k - \gamma_k \nabla f(x^k)) + (I - \lambda_k L L^*) v^k), \\ \tilde{x}^{k+1} = x^k - \gamma_k \nabla f(x^k) - \gamma_k L^* \tilde{v}^{k+1}, \\ (v^{k+1}, x^{k+1}) = (1-\rho_k)(v^k, x^k) + \rho_k(\tilde{v}^{k+1}, \tilde{x}^{k+1}), \end{cases} \tag{21}$$

which recovers the PDFP$^2$O_DS [27].

(iii) Let $P_k = \frac{\gamma}{\lambda} I$ and $Q_k = \frac{1}{\gamma} I$; then, (13) becomes

$$\begin{cases} \tilde{v}^{k+1} = prox_{\frac{\lambda}{\gamma} h^*}(\frac{\lambda}{\gamma} L(x^k - \gamma \nabla f(x^k)) + (I - \lambda L L^*) v^k), \\ \tilde{x}^{k+1} = x^k - \gamma \nabla f(x^k) - \gamma L^* \tilde{v}^{k+1}, \\ (v^{k+1}, x^{k+1}) = (1-\rho_k)(v^k, x^k) + \rho_k(\tilde{v}^{k+1}, \tilde{x}^{k+1}), \end{cases} \tag{22}$$

which is the original PDFP$^2$O (12).

Rv_PDFP$^2$O (13) reduces to the above three algorithms (20)–(22) for different $P_k$ and $Q_k$. It is eailsy confirmed that Rv_PDFP$^2O$ (13) generalizes these algorithms.

### *3.3. Convergence Rates*

In this subsection, we discuss convergence rates of (13).

#### 3.3.1. $O(\frac{1}{k})-$Ergodic Convergence Rate

First, we establish the ergodic convergence rate.

**Lemma 2.** *Suppose that (A1) holds. Let $\{\tilde{u}^{k+1} = (\tilde{v}^{k+1}, \tilde{x}^{k+1})\}$ be generated by (13). Then, for any $u = (v, x) \in K$, it holds that*

$$\begin{aligned} &K(v, \tilde{x}^{k+1}) - K(\tilde{v}^{k+1}, x) \\ \leq &\frac{1}{2}(\|u^k - u\|^2_{U_k^{-1}} - \|\tilde{u}^{k+1} - u\|^2_{U_k^{-1}} - \|\tilde{u}^{k+1} - u^k\|^2_{U_k^{-1}}) + \frac{1}{2\beta}\|\tilde{x}^{k+1} - x^k\|^2. \end{aligned} \tag{23}$$

**Proof.** It follows from the property of proximity operator that

$$\begin{aligned} h^*(v) \geq & h^*(\tilde{v}^{k+1}) + \langle L(x^k - Q_k^{-1} L^* v^k - Q_k^{-1}\nabla f(x^k)), v - \tilde{v}^{k+1}\rangle \\ &+ \frac{1}{2}(\|\tilde{v}^{k+1} - v^k\|^2_{P_k} + \|\tilde{v}^{k+1} - v\|^2_{P_k} - \|v^k - v\|^2_{P_k}). \end{aligned} \tag{24}$$

By the differentiability of $f$, we have

$$\begin{aligned} f(x) \geq & f(\tilde{x}^{k+1}) + \langle \nabla f(x^k), x - \tilde{x}^{k+1}\rangle - \frac{1}{2\beta}\|\tilde{x}^{k+1} - x^k\|^2 \\ \geq & f(\tilde{x}^{k+1}) + \frac{1}{2}\|\tilde{x}^{k+1} - x^k\|^2_{Q_k} + \frac{1}{2}\|x - \tilde{x}^{k+1}\|^2_{Q_k} - \frac{1}{2}\|x - x^k\|^2_{Q_k} \\ &+ \langle \tilde{v}^{k+1}, L(\tilde{x}^{k+1} - x)\rangle - \frac{1}{2\beta}\|\tilde{x}^{k+1} - x^k\|^2. \end{aligned} \tag{25}$$

Adding (24) and (25) and rearranging to arrive at (23). □

**Theorem 2.** *Suppose that (A1)–(A3) hold with* $\xi_k = \frac{\|Q_k^{-1}\|_2}{2\beta}$. *Let* $\{\tilde{u}^{k+1} = (\tilde{v}^{k+1}, \tilde{x}^{k+1})\}$ *be generated by (13). Then, for* $u^* = (v^*, x^*) \in \Omega$*, it holds that*

$$K(v^*, \overline{X}^N) - K(\overline{V}^{N+1}, x^*) \leq \frac{1}{2\rho N}\|u^0 - u^*\|^2_{U_0^{-1}}, \tag{26}$$

*where* $\overline{X}^N = \frac{1}{N}\sum_{k=0}^{N-1}\tilde{x}^{k+1}, \overline{V}^N = \frac{1}{N}\sum_{k=0}^{N-1}\tilde{v}^{k+1}$.

**Proof.** Note that $\tilde{u}^{k+1} = u^k + \frac{1}{\rho_k}(u^{k+1} - u^k)$. Substituting it back into (23), we have

$$\begin{aligned}
&K(v^*, \tilde{x}^{k+1}) - K(\tilde{v}^{k+1}, x^*) \\
\leq &\frac{1}{2\rho_k}(\|u^k - u^*\|^2_{U_k^{-1}} - \|u^{k+1} - u^*\|^2_{U_{k+1}^{-1}}) - \frac{2-\rho_k}{2\rho_k^2}\|x^{k+1} - x^k\|^2_{Q_k - \frac{I}{\beta(2-\rho_k)}} \\
\leq &\frac{1}{2\rho}(\|u^k - u^*\|^2_{U_k^{-1}} - \|u^{k+1} - u^*\|^2_{U_{k+1}^{-1}}).
\end{aligned} \tag{27}$$

Summing (27) from $k = 0, \ldots, N-1$, we obtain

$$\sum_{k=0}^{N-1}(K(v^*, \tilde{x}^{k+1}) - K(\tilde{v}^{k+1}, x^*)) \leq \frac{1}{2\rho}\|u^0 - u^*\|^2_{U_0^{-1}}.$$

The final estimation (26) follows directly from the Jensen inequality. □

### 3.3.2. Linear Convergence Rate

Next, we establish a linear convergence rate of (13) with $P_k = P$ and $Q_k = Q$. Therefore, $\hat{T}_{(k)} = T$. For convenience, we give an equivalent formulation of $T$ as follows:

$$T_1(v, x) = prox^P_{h^*}((I - P^{-1}LQ^{-1}L^*)v + P^{-1}L(x - Q^{-1}\nabla f(x))), \tag{28}$$

$$T_2(v, x) = x - Q^{-1}\nabla f(x) - Q^{-1}L^* \circ T_1(v, x), \tag{29}$$

$$T(v, x) = (T_1(v, x), T_2(v, x)). \tag{30}$$

In addition, we make some additional assumptions. More precisely,

**(A5)**: $\partial h^*$ is $\tau_h$−strongly monotone under $\|\cdot\|_{I-P^{-1}LQ^{-1}L^*}$, i.e.,

$$\forall(x_1, v_1),\ (x_2, v_2) \in gra\ \partial h^*, \langle x_1 - x_2, v_1 - v_2\rangle \geq \tau_h\|x_1 - x_2\|^2_{I-P^{-1}LQ^{-1}L^*}.$$

**(A6)**: $\nabla f$ is $\tau_f$−strongly monotone under the norm $\|\cdot\|_Q$, i.e.,

$$\forall x_1, x_2 \in H, \langle x_1 - x_2, \nabla f(x_1) - \nabla f(x_2)\rangle \geq \tau_f\|x_1 - x_2\|^2_Q.$$

**(A7)**: There is $\theta_1, \theta_2 \in (0,1)$ such that $\|I - P^{-1}LQ^{-1}L^*\|_2 \leq \theta_1$ and $\|x_1 - Q^{-1}\nabla f(x_1) - x_2 - Q^{-1}\nabla f(x_2)\|_Q \leq \sqrt{\theta_2}\|x_1 - x_2\|_Q$ for all $x_1, x_2 \in H$.

**Lemma 3.** *Suppose that* $(A1)$, $(A5)$, *and* $(A6)$ *hold. Then,*

$$\begin{aligned}
&\|T(u_1) - T(u_2)\|^2_{(P+2\tau_h I)(I-P^{-1}LQ^{-1}L^*),Q} \\
\leq &\theta\|u_1 - u_2\|^2_{(P+2\tau_h I)(I-P^{-1}LQ^{-1}L^*),Q},
\end{aligned}$$

*for* $u_1 = (v_1, x_1), u_2 = (v_2, x_2) \in K$, *where* $\theta \in (0,1)$.

**Proof.** Let $u_{h_1} = Mv_1 + L(x_1 - Q^{-1}\nabla f(x_1)) - PT_1(u_1) \in \partial h^*(T_1(u_1))$ and $u_{h_2} = Mv_2 + L(x_2 - Q^{-1}\nabla f(x_2)) - PT_1(u_2) \in \partial h^*(T_1(u_2))$. Then, we have

$$\begin{aligned}
&\|T(u_1) - T(u_2)\|^2_{P-LQ^{-1}L^*,Q} \\
=&\|u_1 - u_2\|^2_{P-LQ^{-1}L^*,Q} - 2\langle \nabla f(x_1) - \nabla f(x_2), T_2(u_1) - T_2(u_2) - (x_1 - x_2)\rangle \\
&-\|T(u_1) - T(u_2) - (u_1 - u_2)\|^2_{P-LQ^{-1}L^*,Q} - 2\langle \nabla f(x_1) - \nabla f(x_2), x_1 - x_2\rangle \\
&-2\langle T_1(u_1) - T_1(u_2), u_{h_1} - u_{h_2}\rangle \\
\leq&\|u_1 - u_2\|^2_{P-LQ^{-1}L^*,Q} - (2 - \frac{\|Q^{-1}\|_2}{\beta})\langle \nabla f(x_1) - \nabla f(x_2), x_1 - x_2\rangle \\
&-2\langle T_1(u_1) - T_1(u_2), u_{h_1} - u_{h_2}\rangle \\
\leq&\|u_1 - u_2\|^2_{P-LQ^{-1}L^*,(1-(2-\frac{\|Q^{-1}\|_2}{\beta})\tau_f)Q} - 2\tau_h\|T_1(u_1) - T_1(u_2)\|^2_{I-P^{-1}LQ^{-1}L^*},
\end{aligned}$$

which concludes the proof with $\theta = \max\{1 - (2 - \frac{\|Q^{-1}\|_2}{\beta})\tau_f, \frac{1}{1+2\tau_h\lambda_{\min}(P^{-1})}\}$. □

**Lemma 4.** *Suppose that $(A1)$ and $(A7)$ hold. Then, for $u_1 = (v_1, x_1), u_2 = (v_2, x_2) \in K$,*

$$\|T(u_1) - T(u_2)\|^2_{P,Q} \leq \theta\|u_1 - u_2\|^2_{P,Q},$$

*where $\theta \in (0, 1)$.*

**Proof.** Define $M = P - LQ^{-1}L^*$. It follows from the fact that $prox^P_{h^*}$ is firmly nonexpansive that

$$\begin{aligned}
&\|T(u_1) - T(u_2)\|^2_{P,Q} \\
\leq&\|x_1 - Q^{-1}\nabla f(x_1) - x_2 - Q^{-1}\nabla f(x_2)\|^2_Q - \|T_1(u_1) - T_1(u_2)\|^2_M \\
&+2\langle T_1(u_1) - T_1(u_2), M(v_1 - v_2)\rangle \\
=&\|x_1 - Q^{-1}\nabla f(x_1) - x_2 - Q^{-1}\nabla f(x_2)\|^2_Q + \|v_1 - v_2\|^2_M \\
&-\|T_1(u_1) - T_1(u_2) - (v_1 - v_2)\|^2_M \\
\leq&\theta_2\|x_1 - x_2\|^2_Q + \theta_1\|v_1 - v_2\|^2_P \\
\leq&\theta\|u_1 - u_2\|^2_{P,Q},
\end{aligned}$$

where $\theta = \max\{\theta_1, \theta_2\} \in (0, 1)$. □

**Theorem 3.** *Suppose that $(A1)$ holds. Suppose that $(A5)$–$(A6)$ hold or $(A7)$ holds. Let $\{u^{k+1} = (v^{k+1}, x^{k+1})\}$ be generated by (13). Let $\rho_k \in (0, \overline{\rho})$ for $\overline{\rho} = \min\{\frac{2}{1+\sqrt{\theta}}, \frac{4\beta - \|Q^{-1}\|_2}{2\beta}\}$. Then, $\{u^{k+1}\}$ converges linearly to the unique point $u^* \in \Omega$, i.e.,*

$$\|u^{k+1} - u^*\| \leq c\eta^{k+1},$$

*where $c > 0$, $\eta \in (0, 1)$.*

**Proof.** Define $T_{(\rho_k)} = (1 - \rho_k)I + \rho_k T$. Note that $u^{k+1} = T_{(\rho_k)}(u^k)$ and $FixT_{(\rho_k)} = FixT = \Omega$. It is clear that $T_{(\rho_k)}$ is $\eta_k$-contractive for $\eta_k = |1 - \rho_k| + \rho_k\sqrt{\theta}$. Therefore, $\{u^{k+1}\}$ converges linearly to the unique fixed point of $T_{(\rho_k)}$. □

## 4. Numerical Experiments

In this section, we apply the proposed Rv_PDFP$^2$O (13) to solve the $L_2 + TV$ deblurring problem (2) and compare it with those of the ADMM [5], PDS [32], PDFP$^2$O [19], and PDFP$^2$O_AM [26]. All of the experiments are performed under Windows 7 and MATLAB

7.2 (R2014a) running on a laptop with an Intel Core 2 Quad CPU 2.3 GHz with 4 GB of memory.

The test images are the standard "Text" image with a size of $256 \times 256$, and "Barbara" and "Goldhill" with a size $512 \times 512$, which are shown in Figure 1. We report numerical results on the image restoration for blurred images, corrupted by the Gaussian noise and the average kernel; $a$ is the size of average kernel, and $\eta$ is the standard variance of the Gaussian noise. To evaluate the ability of the algorithm to remove different noises, we set four kinds of $a, \eta$: (1) $a = 3, \eta = 0.01$; (2) $a = 3, \eta = 0.05$; (3) $a = 7, \eta = 0.01$; and (4) $a = 7, \eta = 0.05$.

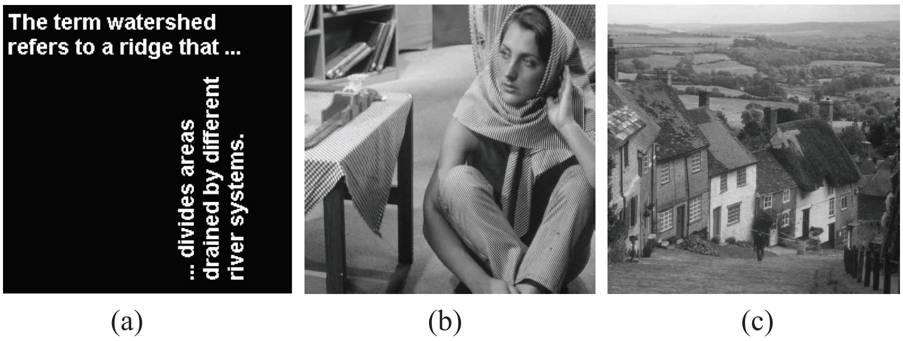


**Figure 1.** These are the test images: (**a**) Text, (**b**) Barbara, and (**c**) Goldhill.

For the two common parameters $\gamma$ and $\lambda$ in PDS, and PDFP$^2$O, we set $\gamma = 1.9$ and $\lambda = 0.125$. Similarly to the literature [26], we choose $Q = A^T A + \zeta L^T L$, where $\zeta = 0.1$. In particular, $Q^{-1}$ can be easily computed by FFT with periodic boundary conditions. We tune the regularization parameter $\mu$ to achieve the maximum SNR, which is listed in Table 2.

**Table 2.** The best selection of $\mu$ in the current noise level.

| **Images** | $\boldsymbol{a = 3}$ | | $\boldsymbol{a = 7}$ | |
|---|---|---|---|---|
| | $\boldsymbol{\eta = 0.01}$ | $\boldsymbol{\eta = 0.05}$ | $\boldsymbol{\eta = 0.01}$ | $\boldsymbol{\eta = 0.05}$ |
| "Text" | 0.0013 | 0.0078 | 0.0003 | 0.0027 |
| "Barbara" | 0.0004 | 0.0101 | 0.0004 | 0.009 |
| "Goldhill" | 0.0011 | 0.0148 | 0.0006 | 0.0085 |

The relative error of the iterative sequences is defined as the stopping criteria:

$$\frac{\|x^{k+1} - x^k\|_2}{\|x^k\|_2} < \varepsilon,$$

where $\varepsilon > 0$ is a prescribed tolerance value. In the experiment, we choose $\varepsilon = 10^{-4}, 10^{-6}, 10^{-8}$. The quality of the restored images is evaluated by signal-to-noise (SNR), which is defined by

$$SNR = 10log\frac{\|x\|^2}{\|x^r - x\|^2},$$

where $x$ and $x^r$ denote the original and the recovered images. The obtained numerical results are listed in Tables 3 and 4.

**Table 3.** The performance of $a = 3$ of the compared algorithms in terms of SNR (dB) and the number of iterations $k$ for given tolerance values $\varepsilon$.

| $\eta$ | Image | Method | $\varepsilon = 10^{-4}$ | | $\varepsilon = 10^{-6}$ | | $\varepsilon = 10^{-8}$ | |
|---|---|---|---|---|---|---|---|---|
| | | | $SNR\,(dB)$ | $k$ | $SNR\,(dB)$ | $k$ | $SNR\,(dB)$ | $k$ |
| 0.01 | Text | ADMM | 26.6288 | 289 | 27.6605 | 813 | 27.6686 | 1537 |
| | | PDS | 26.4334 | 296 | 27.6402 | 933 | 27.6500 | 1693 |
| | | PDFP$^2$O | 27.0279 | 182 | 27.6455 | 521 | 27.6500 | 915 |
| | | PDFP$^2$O_AM | 27.1659 | 178 | 27.6650 | 456 | 27.6686 | 831 |
| | | Rv_PDFP$^2$O | 27.1838 | 173 | 27.6615 | 441 | 27.6686 | 803 |
| | Barbara | ADMM | 21.8641 | 138 | 21.6752 | 1423 | 21.6704 | 5332 |
| | | PDS | 21.8186 | 152 | 21.6783 | 1499 | 21.6733 | 5338 |
| | | PDFP$^2$O | 21.7928 | 118 | 21.6757 | 958 | 21.6733 | 3022 |
| | | PDFP$^2$O_AM | 21.8115 | 109 | 21.6727 | 911 | 21.6704 | 3058 |
| | | Rv_PDFP$^2$O | 21.8075 | 108 | 21.6726 | 887 | 21.6704 | 2962 |
| | Goldhill | ADMM | 26.6924 | 68 | 26.5822 | 513 | 26.5807 | 1554 |
| | | PDS | 26.6686 | 79 | 26.5769 | 510 | 26.5754 | 1471 |
| | | PDFP$^2$O | 26.5980 | 95 | 26.5761 | 306 | 26.5754 | 829 |
| | | PDFP$^2$O_AM | 26.6467 | 51 | 26.5814 | 314 | 26.5807 | 878 |
| | | Rv_PDFP$^2$O | 26.6448 | 50 | 26.5814 | 305 | 26.5807 | 849 |
| 0.05 | Text | ADMM | 14.7157 | 150 | 14.7797 | 511 | 14.7801 | 1048 |
| | | PDS | 14.6905 | 161 | 14.7623 | 533 | 14.7626 | 1904 |
| | | PDFP$^2$O | 14.7341 | 103 | 14.7627 | 384 | 14.7626 | 1892 |
| | | PDFP$^2$O_AM | 14.7510 | 96 | 14.7800 | 308 | 14.7801 | 1185 |
| | | Rv_PDFP$^2$O | 14.7520 | 93 | 14.7800 | 296 | 14.7801 | 1184 |
| | Barbara | ADMM | 18.3425 | 47 | 18.3337 | 231 | 18.3336 | 883 |
| | | PDS | 18, 3496 | 74 | 18.3454 | 349 | 18.3454 | 1683 |
| | | PDFP$^2$O | 18.3459 | 95 | 18.3453 | 342 | 18.3454 | 1681 |
| | | PDFP$^2$O_AM | 18.3362 | 42 | 18.3336 | 226 | 18.3336 | 1154 |
| | | Rv_PDFP$^2$O | 18.3359 | 42 | 18.3336 | 226 | 18.3336 | 1154 |
| | Goldhill | ADMM | 22.7421 | 47 | 22.7367 | 260 | 22.7367 | 1134 |
| | | PDS | 22.7033 | 84 | 22.7015 | 524 | 22.7014 | 2529 |
| | | PDFP$^2$O | 22.7027 | 97 | 22.7015 | 508 | 22.7014 | 2527 |
| | | PDFP$^2$O_AM | 22.7383 | 53 | 22.7367 | 328 | 22.7367 | 1571 |
| | | Rv_PDFP$^2$O | 22.7382 | 53 | 22.7367 | 328 | 22.7367 | 1568 |

**Table 4.** The performance of $a = 7$ of the compared algorithms in terms of SNR (dB) and the number of iterations $k$ for given tolerance values $\varepsilon$.

| $\eta$ | Image | Method | $\varepsilon = 10^{-4}$ | | $\varepsilon = 10^{-6}$ | | $\varepsilon = 10^{-8}$ | |
|---|---|---|---|---|---|---|---|---|
| | | | $SNR\,(dB)$ | $k$ | $SNR\,(dB)$ | $k$ | $SNR\,(dB)$ | $k$ |
| 0.01 | Text | ADMM | 12.8474 | 832 | 14.1337 | 6220 | 14.1548 | 19,540 |
| | | PDS | 12.4021 | 1114 | 14.1117 | 7138 | 14.1382 | 19,512 |
| | | PDFP$^2$O | 13.1403 | 835 | 14.1252 | 4277 | 14.1383 | 10,939 |
| | | PDFP$^2$O_AM | 13.4212 | 646 | 14.1446 | 3828 | 14.1549 | 11,248 |
| | | Rv_PDFP$^2$O | 13.4436 | 634 | 14.1450 | 3718 | 14.1549 | 10,899 |
| | Barbara | ADMM | 18.5769 | 145 | 18.4559 | 2153 | 18.4523 | 9172 |
| | | PDS | 18.5597 | 192 | 18.4592 | 2549 | 18.4541 | 9851 |
| | | PDFP$^2$O | 18.5490 | 161 | 18.4567 | 1647 | 18.4541 | 5623 |
| | | PDFP$^2$O_AM | 18.5445 | 124 | 18.4541 | 1401 | 18.4523 | 5288 |
| | | Rv_PDFP$^2$O | 18.5429 | 122 | 18.4541 | 1365 | 18.4523 | 5124 |
| | Goldhill | ADMM | 23.2362 | 116 | 23.0522 | 1475 | 23.0480 | 8688 |
| | | PDS | 23.1623 | 167 | 23.0451 | 1791 | 23.0395 | 8417 |
| | | PDFP$^2$O | 23.1506 | 129 | 23.0424 | 1146 | 23.0395 | 5284 |
| | | PDFP$^2$O_AM | 23.1775 | 92 | 23.0500 | 966 | 23.0480 | 5608 |
| | | Rv_PDFP$^2$O | 23.1739 | 91 | 23.0499 | 943 | 23.0480 | 5463 |
| 0.05 | Text | ADMM | 7.0043 | 374 | 6.9970 | 2466 | 6.9972 | 6803 |
| | | PDS | 6.9372 | 535 | 6.9771 | 3074 | 6.9773 | 7579 |
| | | PDFP$^2$O | 6.9601 | 373 | 6.9772 | 1859 | 6.9773 | 4154 |
| | | PDFP$^2$O_AM | 6.9997 | 276 | 6.9971 | 1528 | 6.9972 | 3737 |
| | | Rv_PDFP$^2$O | 6.9995 | 270 | 6.9971 | 1488 | 6.9972 | 3611 |
| | Barbara | ADMM | 17.1962 | 95 | 17.1656 | 764 | 17.1653 | 2939 |
| | | PDS | 17.2239 | 161 | 17.1937 | 900 | 17.1932 | 3370 |
| | | PDFP$^2$O | 17.2073 | 121 | 17.1933 | 648 | 17.1932 | 3091 |
| | | FP$^2$O_AM | 17.1783 | 80 | 17.1654 | 506 | 17.1653 | 1955 |
| | | Rv_PDFP$^2$O | 17.1776 | 79 | 17.1654 | 501 | 17.1653 | 1947 |
| | Goldhill | ADMM | 20.4785 | 83 | 20.4481 | 612 | 20.4477 | 2441 |
| | | PDS | 20.4285 | 153 | 20.4010 | 762 | 20.4005 | 2684 |
| | | PDFP$^2$O | 20.4114 | 116 | 20.4006 | 566 | 20.4005 | 2323 |
| | | PDFP$^2$O_AM | 20.4595 | 72 | 20.4478 | 430 | 20.4477 | 1529 |
| | | Rv_PDFP$^2$O | 20.4586 | 72 | 20.4478 | 426 | 20.4477 | 1505 |

It can be seen from Tables 3 and 4 that the proposed Rv_PDFP$^2$O converges faster than other algorithms in terms of the number of iterations. In addition, Figures 2–4 show the recovered images with $\varepsilon = 10^{-8}$. Figures 2–4 show that the visual qualities of these images obtained by the proposed algorithm are slightly better than the compared algorithms.

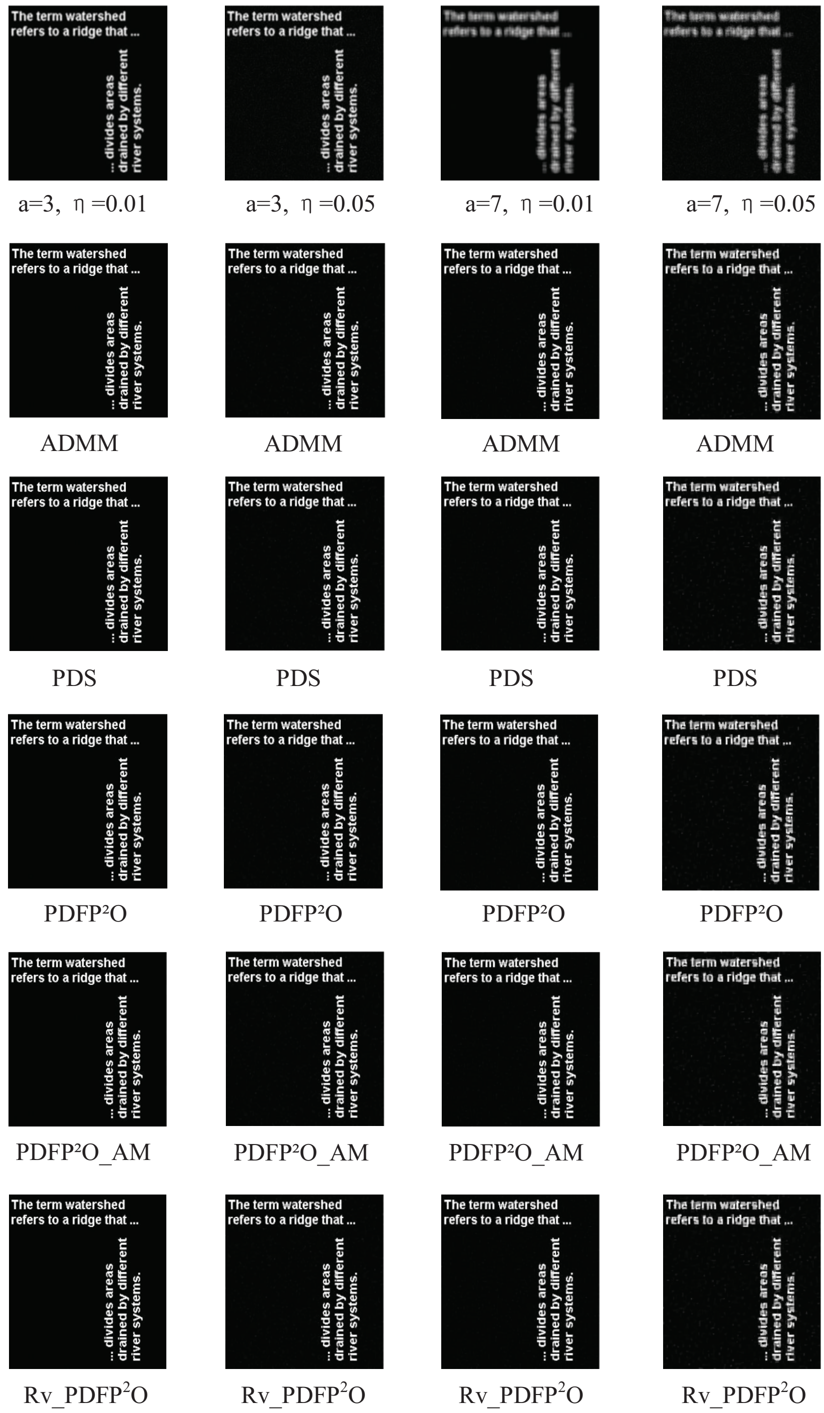


**Figure 2.** These are the "Text" images: **Row 1**: the blurry and noisy images, **Row 2**: the images restored by ADMM, **Row 3**: the images restored by PDS, **Row 4**: the images restored by PDFP$^2$O, **Row 5**: the images restored by PDFP$^2$O_AM, and **Row 6**: the images restored by Rv_PDFP$^2$O.

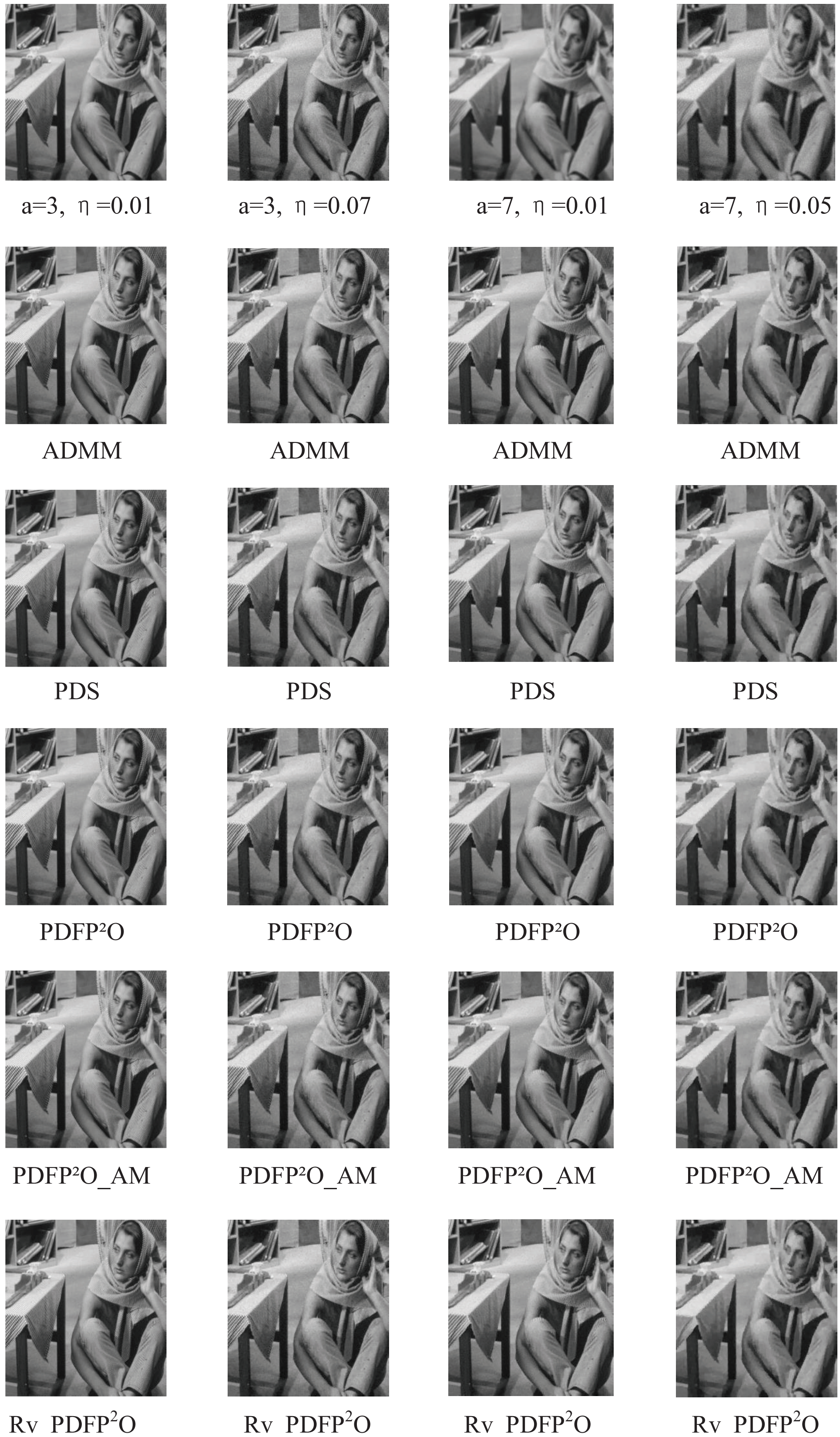


**Figure 3.** These are the "Goldhill" images: **Row 1**: the blurry and noisy images, **Row 2**: the images restored by ADMM, **Row 3**: the images restored by PDS, **Row 4**: the images restored by PDFP$^2$O, **Row 5**: the images restored by PDFP$^2$O_AM, and **Row 6**: the images restored by Rv_PDFP$^2$O.


**Figure 4.** These are the "Goldhill" images: **Row 1**: the blurry and noisy images, **Row 2**: the images restored by ADMM, **Row 3**: the images restored by PDS, **Row 4**: the images restored by PDFP$^2$O, **Row 5**: the images restored by PDFP$^2$O_AM, and **Row 6**: the images restored by Rv_PDFP$^2$O.

## 5. Conclusions

In this article, we proposed a Rv_PDFP$^2$O to solve the convex optimization problem (1). The proposed algorithm combined the over-relaxed parameters and the variable metric. Under a proper preconditioned operator, we derived the Rv_PDFP$^2$O and established the convergence. By defining different stepsizes, we showed that the Rv_PDFP$^2$O recovers some existing algorithms, including PDFP$^2$O, PDFP$^2$O_AM, and PDFP$^2$O_DS, and we provide larger relaxed parameters for these algorithms. Furthermore, we studied the $O(\frac{1}{k})$ ergodic convergence rate in the partial primal-dual gap. Under some strong conditions on the objective functions and the stepsizes, we proved that the iterative sequences converge linearly. We applied the Rv_PDFP$^2$O to solve the TV image-restoration problem (2). The numerical results show that the Rv_PDFP$^2$O performs better than some existing algorithms. As we all know, the self-adaptive stepsize and the inertial variant could improve the algorithm. However, these two accelerated strategies are not introduced to the Rv_PDFP$^2$O algorithm . We would like to derive a self-adaptive Rv_PDFP$^2$O and an inertial Rv_PDFP$^2$O in the future.

**Author Contributions:** Conceptualization, W.H. and H.L.; methodology, W.H. and H.L.; software, Y.T. and M.W.; validation, W.H., Y.T. and H.L.; formal analysis, W.H. and Y.T.; and writing, W.H., Y.T., M.W. and H.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was funded by the National Science Foundation of China, grant numbers 12271117, 12061045, and 12001416, and the basic research joint-funding project of the university and Guangzhou city, grant number 202102010434.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Antonin, C.; Pock, T. An introduction to continuous optimization for imaging. *Acta Numer.* **2016**, *25*, 161–319.
2. Rudin, L.I.; Osher, S.; Fatemi, E. Nonlinear total variation based noise removal algorithms. *Phys. D Nonlinear Phenom.* **1992**, *60*, 259–268. [CrossRef]
3. Davenport, M.A.; Duarte, M.F.; Eldar, Y.C.; Kutyniok, G. *Introduction to Compressed Sensing*; Cambridge University Press: Cambridge, UK, 2012.
4. Zou, H.; Hastie, T. Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **2005**, *67*, 301–320. [CrossRef]
5. Boyd, S.; Parikh, N.; Chu, E.; Peleato, B.; Eckstein, J. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.* **2010**, *3*, 1–122. [CrossRef]
6. Chan, R.H.; Tao, M.; Yuan, X.M. Constrained total variation deblurring models and fast algorithms based on alternating direction method of multipliers. *SIAM J. Imaging Sci.* **2011**, *6*, 680–697. [CrossRef]
7. Han, D.R.; Yuan, X.M. A note on the alternating direction method of multipliers. *J. Optimz. Theory Appl.* **2012**, *155*, 227–238. [CrossRef]
8. Bauschke, H.H.; Combettes, P.L. *Concex Analysis and Monotone Operator Theory in Hilbert Spaces*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2017.
9. Eckstein, J.; Bertsekas, D.P. On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Math. Program.* **1992**, *55*, 293–318. [CrossRef]
10. Yan, M.; Yin, W.T. Self equivalence of the alternating direction method of multipliers. In *Splitting Methods in Communication, Imaging, Science, and Engineering*; Glowinski, R., Osher, S.J., Yin, W.T., Eds.; Springer: Cham, Switzerland, 2016; pp. 165–194.
11. Combettes, P.L.; Wajs, V.R. Signal recovery by proximal forward–backward splitting. *Multiscale Model. Simul.* **2005**, *4*, 1168–1200. [CrossRef]
12. Beck, A.; Teboulle, M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.* **2009**, *2*, 183–202. [CrossRef]
13. Beck, A.; Teboulle, M. Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. *IEEE Trans. Image Process.* **2009**, *18*, 2419–2434. [CrossRef]
14. Iutzeler, F.; Malick, J. On the proximal gradient algorithm with alternated inertia. *J. Optimz. Theory App.* **2018**, *176*, 688–710. [CrossRef]

15. Chambolle, A.; Pock, T. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vis.* **2011**, *40*, 120–145. [CrossRef]
16. Komodakis, N.; Pesquet, J.-C. Playing with duality: An overview of recent primal-dual approaches for solving large-scale optimization problems. *IEEE Signal Proc. Mag.* **2015**, *32*, 31–54. [CrossRef]
17. Micchelli, C.A.; Shen, L.X.; Xu, Y.S. Proximity algorithms for image models: Denoising. *Inverse Probl.* **2011**, *27*, 045009. [CrossRef]
18. Argyriou, A.; Micchelli, C.A.; Pontil, M.; Shen, L.X.; Xu, Y.S. Efficient first order methods for linear composite regularizers. *arXiv* **2011**, arXiv:1104.1436.
19. Chen, P.J.; Huang, J.G.; Zhang, X.Q. A primal-dual fixed-point algorithm for convex separable minimization with applications to image restoration. *Inverse Probl.* **2013**, *29*, 025011. [CrossRef]
20. Loris, I.; Verhoeven, C. On a generalization of the iterative soft-thresholding algorithm for the case of non-separable penalty. *Inverse Probl.* **2011**, *27*, 125007. [CrossRef]
21. He, Z.; Zhu, Y.N.; Qiu, S.H.; Wang, T.; Zhang, C.C.; Sun, B.M.; Zhang, X.Q.; Feng, Y. Low-Rank and Framelet Based Sparsity Decomposition for Interventional MRI Reconstruction. *IEEE Bio-Med. Eng.* **2022**, *69*, 2294–2304. [CrossRef]
22. Ren, Z.M.; Zhang, Q.F.; Yuan, Y.X. A Primal–Dual Fixed-Point Algorithm for TVL1 Wavelet Inpainting Based on Moreau Envelope. *Mathematics* **2022**, *10*, 2470. [CrossRef]
23. Combettes, P.L.; Condat, L.; Pesquet, J.C.; Vu, B.C. A forward–backward View of Some Primal-dual Optimization Methods in Image Recovery. In Proceedings of the 2014 IEEE International Conference on Image Processing, ICIP 2014 Conference, Paris, France, 27–30 October 2014; pp. 4141–4145.
24. Drori, Y.; Sabach, S.; Teboulle, M. A simple algorithm for a class of nonsmooth convex-concave saddle-point problems. *Oper. Res. Lett.* **2015**, *43*, 209–214. [CrossRef]
25. Luke, D.R.; Shefi, R. A globally linearly convergent method for pointwise quadratically supportable convex-concave saddle point problems. *J. Math. Anal. Appl.* **2018**, *457*, 1568–1590. [CrossRef]
26. Chen, D.Q.; Zhou, Y.; Song, L.J. fixed-point algorithm based on adapted metric method for convex minimization problem with application to image deblurring. *Adv. Comput. Math.* **2016**, *42*, 1287–1310. [CrossRef]
27. Wen, M.; Tang, Y.C.; Cui, A.G.; Peng, J.G. Efficient primal-dual fixed-point algorithms with dynamic stepsize for composite convex optimization problems. *Multidimens. Syst. Signal Process.* **2019**, *30*, 1531–1544. [CrossRef]
28. Li, Z.; Yan, M. New convergence analysis of a primal-dual algorithm with large stepsizes. *Adv. Comput. Math.* **2019**, *47*, 1–20. [CrossRef]
29. Zhu, Y.N.; Zhang, X.Q. Two Modified Schemes for the Primal Dual Fixed Point Method. *CSIAM Trans. Appl. Math.* **2021**, *2*, 108–130.
30. Combettes, P.L.; Wajs, V.R. Variable metric quasi-fejer monotonicity. *Nonlinear Anal-Theor.* **2013**, *78*, 17–31. [CrossRef]
31. Cui, F.Y.; Tang, Y.C.; Zhu, C.X. Convergence analysis of a variable metric forward–backward splitting algorithm with applications. *J. Inequal. Appl.* **2019**, *2019*, 1–27. [CrossRef]
32. Condat, L. A primal-dual splitting method for convex optimization involving lipschitzian, proximable and linear composite terms. *J. Optim. Theory Appl.* **2013**, *158*, 460–479. [CrossRef]