

A Fast Algorithm for Deconvolution and Poisson Noise Removal

Xiongjun Zhang¹  · Michael K. Ng² · Minru Bai³

Received: 24 July 2017 / Revised: 25 October 2017 / Accepted: 31 October 2017 /
Published online: 12 November 2017
© Springer Science+Business Media, LLC 2017

Abstract Poisson noise removal problems have attracted much attention in recent years. The main aim of this paper is to study and propose an alternating minimization algorithm for Poisson noise removal with nonnegative constraint. The algorithm minimizes the sum of a Kullback-Leibler divergence term and a total variation term. We derive the algorithm by utilizing the quadratic penalty function technique. Moreover, the convergence of the proposed algorithm is also established under very mild conditions. Numerical comparisons between our approach and several state-of-the-art algorithms are presented to demonstrate the efficiency of our proposed algorithm.

Keywords Poisson noise · Alternating minimization algorithm · Total variation · Kullback-Leibler divergence

Mathematics Subject Classification 65F10 · 65F22 · 68U10 · 90C25

Xiongjun Zhang: The research of this author was supported in part by the Fundamental Research Funds for the Central Universities under Grant 230-20205170463-610.

Michael K. Ng: The research of this author was supported in part by the HKRGC GRF 1202715, 12306616, 12200317 and HKBU RC-ICRS/16-17/03.

Minru Bai: The research of this author was supported in part by the National Science Foundation of China under Grant 11571098.

✉ Xiongjun Zhang
xjzhang@mail.ccnu.edu.cn

Michael K. Ng
mng@math.hkbu.edu.hk

Minru Bai
minru-bai@163.com

¹ School of Mathematics and Statistics, Central China Normal University, Wuhan 430079, China

² Centre for Mathematical Imaging and Vision and Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong

³ College of Mathematics and Econometrics, Hunan University, Changsha 410082, China

1 Introduction

Image deblurring under Poisson noise is an important task in various applications such as astronomical images [26, 39], positron emission tomography (PET) [4, 29], and single particle emission computed tomography (SPECT) [18, 31]. In this paper, we focus on the following image reconstruction model:

$$g = \text{Poisson}(Af + c\mathbf{e}), \quad (1.1)$$

where $g \in \mathbb{R}^n$ is the observed image, $\mathbf{e} := (1, \dots, 1)^T \in \mathbb{R}^n$ is a vector of all ones, $c \geq 0$ is a fixed background, $A \in \mathbb{R}^{n \times n}$ is a spatial-invariant blur matrix, $f \in \mathbb{R}^n$ is a column vector concatenated from the original image with size $m_1 \times m_2$ ($n = m_1 m_2$), and $\text{Poisson}(\cdot)$ denotes the degradation by Poisson noise. We remark that A is a matrix of block Toeplitz with Toeplitz blocks when zero boundary conditions are applied, and block Toeplitz-plus-Hankel with Toeplitz-plus-Hankel blocks when Neumann boundary conditions are used [27, 28].

An approach to deal with the images corrupted by Poisson noise is the maximum a posterior (MAP) estimate [5, 7, 16, 20, 49]. Assume that the observed data g_i are independent random variables, then the probability distribution of g is the product of the probability distributions of all pixels [37], which can be described as follows:

$$P(g|f) = \prod_{i=1}^n \frac{[(Af + c\mathbf{e})_i]^{g_i} e^{-(Af + c\mathbf{e})_i}}{g_i!},$$

where $(Af + c\mathbf{e})_i$ and g_i denote the i th components of $Af + c\mathbf{e}$ and g , respectively. By taking the negative logarithm of the probability distribution, the Kullback-Leibler (KL) divergence $D_{KL}(Af + c\mathbf{e}, g)$ of $Af + c\mathbf{e}$ from g is given by [13]

$$D_{KL}(Af + c\mathbf{e}, g) := \sum_{i=1}^n (Af + c\mathbf{e})_i - g_i \log((Af + c\mathbf{e})_i) + \log(g_i!),$$

which is convex, nonnegative, and coercive on the nonnegative orthant. When the constants are omitted, we obtain the following data fidelity term

$$\Phi(f) := \mathbf{e}^T (Af + c\mathbf{e}) - g^T \log(Af + c\mathbf{e}),$$

where the notation T denotes the transpose operator and $\log(\cdot)$ means componentwise. The maximum likelihood estimator of the original image is the minimizer of $\Phi(f)$. It is known that the Richardson-Lucy (RL) algorithm can be applied to solve this problem [24, 32]. There are also a number of important works on RL algorithm, including [9, 14, 37, 41], to name only a few. However, the RL algorithm converges slowly and yields high noise estimate [43].

To deal with the ill-posed problem, a regularization term should be added to control the noise. Traditional regularization methods include the Tikhonov-like regularization [42] and the total variation (TV) regularization [35]. Although the Tikhonov-like regularization method has the advantage of simple calculations [17], it tends to make images unduly smooth and often fails to adequately save important image attributes such as sharp edges. By contrast, the TV regularization, first proposed by Rudin, Osher, and Fatemi for image denoising in [35] and then extended to image deconvolution in [34], has become very popular to represent the prior of images due to its ability to preserve sharp edges. In addition, Acar et al. [1] analyzed the advantages of the TV regularization over the Tikhonov-like regularization for recovering images including piecewise smooth objects.

Based on the prior information of an image, we use the TV as the regularization term. Then, a discrete version of the TV deblurring problem under Poisson noise is given by [23]

$$\min_{f \geq 0} \alpha \left(\mathbf{e}^T (Af + c\mathbf{e}) - g^T \log(Af + c\mathbf{e}) \right) + \sum_{i=1}^n \|D_i f\|, \tag{1.2}$$

where α is a regularization parameter to balance the regularization term and the KL divergence term, the nonnegative constraint $f \geq 0$ guarantees that no negative intensities occur in the restored images, $\|\cdot\|$ denotes the Euclidean norm of a vector, and $D_i f \in \mathbb{R}^2$ denotes the first-order finite difference of f at pixel i in both horizontal and vertical directions. More specifically, let f be extended by periodic boundary conditions, where $f := (f_1, f_2, \dots, f_n)^T$. Then, for $i = 1, \dots, n$, $D_i f$ is defined as

$$D_i f := ((D_x f)_i, (D_y f)_i)^T,$$

where

$$(D_x f)_i = \begin{cases} f_{i+m_2} - f_i, & \text{if } 1 \leq i \leq m_2(m_1 - 1), \\ f_{\text{mod}(i, m_2)} - f_i, & \text{otherwise,} \end{cases}$$

and

$$(D_y f)_i = \begin{cases} f_{i+1} - f_i, & \text{if } \text{mod}(i, m_2) \neq 0, \\ f_{i-m_2+1} - f_i, & \text{otherwise.} \end{cases}$$

Here $\text{mod}(i, m_2)$ denotes the remainder of i divided by m_2 .

The computational challenges of this model are that the KL divergence is non-quadratic and contains a logarithm function. The algorithms of image restoration based on the quadratic data fitting term cannot be applied to this model directly. Moreover, since the TV term is non-differentiable, it is difficult to solve the Euler-Lagrange equation associated with the minimization model (1.2). Recently, some efficient algorithms are proposed to solve the Poisson noise removal problem. Le et al. [23] proposed a gradient descent method to solve a variational image restoration model under Poisson noise. It is well known that the convergence speed of the gradient descent method is very slow. Moreover, Figueiredo et al. [15] proposed an augmented Lagrangian (PIDAL) method for deconvolving Poissonian images. But it needs to solve a TV denoising subproblem under their splitting strategy, which cannot be solved exactly. Furthermore, Setzer et al. [36] proposed an alternating split Bregman algorithm (called PIDSplit+) for the restoration of blurred images corrupted with Poissonian noise. Although the resulting subproblems have closed-form solutions, this algorithm converges slowly after reaching a low accuracy.

In this paper, we focus on developing a fast algorithm to solve the constrained minimization model (1.2). By putting the constrained set into the objective function, we obtain the following unconstrained optimization problem

$$\min_f \alpha \left(\mathbf{e}^T (Af + c\mathbf{e}) - g^T \log(Af + c\mathbf{e}) \right) + \sum_{i=1}^n \|D_i f\| + \delta_{\mathbb{R}_+^n}(f), \tag{1.3}$$

where $\delta_{\mathbb{R}_+^n}(\cdot)$ is the indicator function over \mathbb{R}_+^n . Let $u \in \mathbb{R}^n$ be an auxiliary variable that approximates $Af + c\mathbf{e}$ in (1.3). Similarly, we introduce $w_i \in \mathbb{R}^2$ to approximate $D_i f$ and z to approximate f . By adding the quadratic terms to penalize the differences between every pair of original and auxiliary variables, we obtain the following problem

$$\min_{u,z,w,f} \mathcal{Q}(u, z, w, f) := \alpha \left(\mathbf{e}^T u - g^T \log(u) \right) + \delta_{\mathbb{R}_+^n}(z) + \beta_1 \|u - (Af + \mathbf{c}\mathbf{e})\|^2 + \sum_{i=1}^n \|w_i\| + \beta_2 \sum_{i=1}^n \|D_i f - w_i\|^2 + \beta_3 \|z - f\|^2, \quad (1.4)$$

where $\beta_1, \beta_2, \beta_3$ are positive penalty parameters, respectively. It is clear that when $\beta_1, \beta_2, \beta_3$ are sufficiently large, the objective function in (1.4) is close to that in (1.3).

An alternating minimization algorithm is proposed to solve the model (1.4). Alternating minimization algorithms have been applied to the cases of recovering blurred images corrupted by impulse, Gaussian, and multiplicative noises successfully [19,21,22,44,46–48]. For more details about multiplicative noise removal, the interested readers are referred to the papers [2,38,40]. Since u, z, w are decoupled, we can minimize f and (u, z, w) , alternately. When $\beta_1, \beta_2, \beta_3$ are sufficiently large, the minimizer of (1.4) is close to that of (1.3). The main advantages of the proposed algorithm are that the resulting subproblems have closed-form solutions or can be solved by fast Fourier transforms [27,28]. Moreover, based on the fixed point iteration [8,30], we show that, for any fixed $\beta_1, \beta_2, \beta_3$, the sequence generated by the alternating minimization algorithm converges to a solution of (1.4). Our numerical results show that the proposed algorithm is faster than PIDAL [15], PIDSplit+ [36], and the primal dual (PD) algorithm [45].

The remaining parts of this paper are organized as follows. In Sect. 2, we give some notation that will be used throughout this paper. The alternating minimization algorithm is presented in Sect. 3. We establish the convergence of the proposed algorithm in Sect. 4. In Sect. 5, some numerical experiments are presented to demonstrate the efficiency of our proposed algorithm. Finally, we give some concluding remarks in Sect. 6.

2 Preliminaries

In this section, we briefly introduce some notation that will be used throughout this paper. \mathbb{R}^n denotes the n -dimensional Euclidean space. \mathbb{R}_+^n denotes the nonnegative orthant of \mathbb{R}^n , i.e., $\mathbb{R}_+^n := \{x = (x_1, \dots, x_n)^T \in \mathbb{R}^n \mid x_i \geq 0, i = 1, \dots, n\}$. \mathbb{R}_{++}^n denotes the positive orthant of \mathbb{R}^n , i.e., $\mathbb{R}_{++}^n := \{x = (x_1, \dots, x_n)^T \in \mathbb{R}^n \mid x_i > 0, i = 1, \dots, n\}$. The set of all $m \times n$ matrices with real entries is denoted by $\mathbb{R}^{m \times n}$. For any matrix $A \in \mathbb{R}^{m \times n}$, $\|A\|$ denotes the spectral norm of A .

Let $D^{(1)}$ and $D^{(2)}$ be the first-order finite difference matrices in the horizontal and vertical directions, respectively. $D_i \in \mathbb{R}^{2 \times n}$ is a matrix formed by stacking the i th rows of $D^{(1)}$ and $D^{(2)}$. Let $D := (D_1^T, \dots, D_n^T)^T$. $A \succeq B$ denotes that $A - B$ is positive semidefinite. The identity matrix is denoted by I , whose dimension should be clear from the context. The relative interior of a set G is denoted by $\text{int}(G)$. The indicator function of a set G is defined as

$$\delta_G(x) := \begin{cases} 0, & \text{if } x \in G, \\ +\infty, & \text{otherwise.} \end{cases}$$

Next, we give the definition of a convex function [33, Definition 2.1].

Definition 2.1 (a) A subset C of \mathbb{R}^n is convex if it includes for every pair of points the line segment that joins them, or in other words, if for every choice of $x_0 \in C$ and $x_1 \in C$ one has $[x_0, x_1] \in C$:

$$(1 - \tau)x_0 + \tau x_1 \in C \text{ for all } \tau \in (0, 1).$$

(b) A function ψ on a convex set C is convex relative to C if for every choice of $x_0 \in C$ and $x_1 \in C$ one has

$$\psi((1 - \tau)x_0 + \tau x_1) \leq (1 - \tau)\psi(x_0) + \tau\psi(x_1) \text{ for all } \tau \in (0, 1).$$

The effective domain of a function $\psi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ is defined as

$$\text{dom}(\psi) := \{x : \psi(x) < +\infty\}.$$

We say that ψ is proper if it never equals $-\infty$ and $\text{dom}(\psi) \neq \emptyset$. For any proper convex function $\psi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ and any point $x \in \text{dom}(\psi)$, the subgradient of ψ at x is defined as

$$\partial\psi(x) := \{v : \psi(y) \geq \psi(x) + \langle v, y - x \rangle, \forall y\}.$$

Let $\psi : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ be a closed, proper, convex function, where \mathcal{X} is a real finite dimensional Euclidean space equipped with an inner product $\langle \cdot, \cdot \rangle$ and its induced norm $\|\cdot\|$. The operator

$$\text{Prox}_\psi : \mathcal{X} \rightarrow \mathcal{X} : x \mapsto \arg \min_{y \in \mathcal{X}} \left\{ \psi(y) + \frac{1}{2} \|x - y\|^2 \right\}$$

is called the proximal mapping of ψ [12].

Other notation will be defined in appropriate sections if necessary.

3 An Alternating Minimization Algorithm

In this section, we present an alternating minimization algorithm to solve (1.4). Alternating minimization algorithms have been applied to the cases of recovering blurred images corrupted by impulse, Gaussian, and multiplicative noises successfully [19, 21, 22, 44, 46–48]. The alternating minimization algorithm for solving (1.4) can be described in Algorithm 1.

Algorithm 1: An alternating minimization algorithm for solving problem (1.4).

Input: $\alpha, \beta_1, \beta_2, \beta_3, u^0 > 0, z^0, w^0$. For $k = 0, 1, \dots$, perform the following steps:

Step 1. Compute

$$f^{k+1} := \arg \min_f \left\{ \beta_1 \|u^k - (Af + ce)\|^2 + \beta_2 \sum_{i=1}^n \|D_i f - w_i^k\|^2 + \beta_3 \|f - z^k\|^2 \right\}. \tag{3.1}$$

Step 2. Compute

$$u^{k+1} := \arg \min_u \left\{ \alpha (e^T u - g^T \log(u)) + \beta_1 \|u - (Af^{k+1} + ce)\|^2 \right\}. \tag{3.2}$$

Step 3. Compute

$$z^{k+1} := \arg \min_z \left\{ \delta_{\mathbb{R}_+^n}(z) + \beta_3 \|z - f^{k+1}\|^2 \right\}. \tag{3.3}$$

Step 4. For $i = 1, \dots, n$, compute w_i

$$w_i^{k+1} := \arg \min_{w_i} \left\{ \|w_i\| + \beta_2 \|D_i f^{k+1} - w_i\|^2 \right\}. \tag{3.4}$$

Since the objective function in (3.1) is differentiable, f^{k+1} can be computed by solving the following equation:

$$(\beta_1 A^T A + \beta_2 D^T D + \beta_3 I) f^{k+1} = \beta_1 A^T (u^k - c\mathbf{e}) + \beta_2 D^T w^k + \beta_3 z^k, \tag{3.5}$$

where $w^k := ((w_1^k)^T, \dots, (w_n^k)^T)^T$. Let $M := \beta_1 A^T A + \beta_2 D^T D + \beta_3 I$, then M is positive definite. Thus, we have

$$f^{k+1} = M^{-1} (\beta_1 A^T (u^k - c\mathbf{e}) + \beta_2 D^T w^k + \beta_3 z^k). \tag{3.6}$$

Remark 3.1 In image restoration problems, A is usually a blurring matrix and D is a difference matrix under period boundary conditions that can be diagonalized by fast Fourier transforms. The computational cost of the method is dominated by three fast discrete transforms for solving the linear system in (3.5) [27,28]. The interested readers are referred to [27,28] for a detailed discussion on how to handle blurring matrices with other boundary conditions.

Since A is a blurring matrix, the entries of A are nonnegative, and $c \geq 0$. Notice that, for any $u^0 > 0$, we have $u^k > 0$. Therefore, the closed-form solution for (3.2) is given by

$$u^{k+1} = \frac{A f^{k+1} + c\mathbf{e}}{2} - \frac{\alpha\mathbf{e}}{4\beta_1} + \frac{\sqrt{(\alpha\mathbf{e} - 2\beta_1(A f^{k+1} + c\mathbf{e}))^2 + 8\alpha\beta_1 g}}{4\beta_1}. \tag{3.7}$$

Here, the root square of a vector means componentwise.

As for the subproblem about z in (3.3), we have

$$z^{k+1} = \mathcal{P}_{\mathbb{R}_+^n} (f^{k+1}),$$

where \mathcal{P} is the projection onto \mathbb{R}_+^n , i.e.,

$$(\mathcal{P}_{\mathbb{R}_+^n} (f))_i := \begin{cases} f_i, & \text{if } f_i \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

As for the subproblem about w_i in (3.4), by the well-known two-dimensional shrinkage [3], w_i^{k+1} is given by

$$w_i^{k+1} = \max \left\{ \|D_i f^{k+1}\| - \frac{1}{2\beta_2}, 0 \right\} \frac{D_i f^{k+1}}{\|D_i f^{k+1}\|}, \quad i = 1, \dots, n, \tag{3.8}$$

where $0 \cdot (0/0) = 0$ is assumed.

Remark 3.2 Algorithm 1 could be applied to both the isotropic and anisotropic TV. For simplicity, we will only discuss the isotropic case in detail since it is similar to deal with the anisotropic case.

Remark 3.3 Compared with PIDAL [15], the subproblems of the alternating minimization algorithm can be solved exactly in each iteration.

4 Convergence Analysis

In this section, we establish the convergence of Algorithm 1. The following results in [8,30] are utilized to show that the iteration sequence $\{(f^k, u^k, z^k, w^k)\}$ generated by Algorithm 1 converges to a minimizer (f^*, u^*, z^*, w^*) of \mathcal{Q} .

Theorem 4.1 *Let \mathcal{X} be a Hilbert space, and let \mathcal{T} be a nonexpansive self-mapping of \mathcal{X} such that \mathcal{T} is asymptotically regular, i.e., for any x in \mathcal{X} , $\|\mathcal{T}^{n+1}x - \mathcal{T}^n x\|$ tends to 0 as n tends to $+\infty$. Assume that \mathcal{T} has at least one fixed point. Then, for any x in \mathcal{X} , a weak limit of a weakly convergent subsequence of the sequence of successive approximations $\{\mathcal{T}^n x\}$ is a fixed point of \mathcal{T} .*

Although Theorem 4.1 shows the general case of the fixed point iteration in Hilbert spaces, we only use the finite dimension case for the sequence generated by Algorithm 1. Thus the weak and strong convergence coincide for Algorithm 1, i.e., the limit point of $\{\mathcal{T}^n x\}$ is a fixed point of \mathcal{T} . For the finite dimensional case about this theorem, a simplified proof can be found in [10, Theorem 5]. When \mathcal{X} is \mathbb{R}^n , the sequence of successive approximations can converge to a fixed point. Therefore, we only need to prove that the operator \mathcal{T} about the iteration sequence is nonexpansive, asymptotically regular, and has a fixed point.

For the proximal operator of a function, we have the following proposition [12, Lemma 2.4].

Proposition 4.1 *Let $\phi : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ be a closed, proper and convex function. Then, for any $x, y \in \mathcal{X}$,*

$$\|\text{Prox}_\phi(x) - \text{Prox}_\phi(y)\|^2 \leq \langle \text{Prox}_\phi(x) - \text{Prox}_\phi(y), x - y \rangle.$$

In the following, we define the nonexpansive operator [6], which is important for the analysis of the convergence.

Definition 4.1 An operator \mathcal{T} is called nonexpansive in \mathcal{X} if for any $x_1, x_2 \in \mathcal{X}$, we have

$$\|\mathcal{T}(x_1) - \mathcal{T}(x_2)\| \leq \|x_1 - x_2\|.$$

By Proposition 4.1, we obtain that Prox_ϕ is nonexpansive, i.e.,

$$\|\text{Prox}_\phi(x) - \text{Prox}_\phi(y)\| \leq \|x - y\|, \quad \forall x, y \in \mathcal{X}.$$

Actually, the proximal operator is a special kind of an averaged operator, which is nonexpansive [10, Lemma 3]. Next, we construct the iteration sequence of Algorithm 1. Let

$$H := \begin{pmatrix} \sqrt{\beta_1}A \\ \sqrt{\beta_3}I \\ \sqrt{\beta_2}D \end{pmatrix} \text{ and } v := \begin{pmatrix} u \\ z \\ w \end{pmatrix},$$

then $M = H^T H$ by the definition of M . Thus, (3.5) is equivalent to

$$\begin{aligned} Mf^{k+1} &= \beta_1 A^T (u^k - \mathbf{c}\mathbf{e}) + \beta_2 D^T w^k + \beta_3 z^k \\ &= (BH)^T v^k - \beta_1 c A^T \mathbf{e}, \end{aligned}$$

where

$$B := \begin{pmatrix} \sqrt{\beta_1}I & 0 & 0 \\ 0 & \sqrt{\beta_3}I & 0 \\ 0 & 0 & \sqrt{\beta_2}I \end{pmatrix}.$$

Since M is positive definite, we obtain

$$f^{k+1} = M^{-1}((BH)^T v^k - \beta_1 c A^T \mathbf{e}) := \mathcal{T}_f(v^k). \tag{4.1}$$

Let $\varphi(u) := \mathbf{e}^T u - g^T \log(u)$. Since $u > 0$ and $Af + \mathbf{c}\mathbf{e} > 0$, by (3.2), we get

$$u^{k+1} = \text{Prox}_{\frac{\alpha}{2\beta_1}\varphi}(Af^{k+1} + \mathbf{c}\mathbf{e}) := \mathcal{T}_1(f^{k+1}) = \mathcal{T}_1(\mathcal{T}_f(v^k)). \tag{4.2}$$

Moreover, by (3.3), we have

$$z^{k+1} = \mathcal{P}_{\mathbb{R}_+^n}(f^{k+1}) := \mathcal{T}_2(f^{k+1}) = \mathcal{T}_2(\mathcal{T}_f(v^k)). \tag{4.3}$$

Note that, for w_i , the minimizer is the two-dimensional shrinkage operator [3]. Let $g(x) = \|x\|$, by (3.8), we obtain that

$$w^{k+1} := \begin{pmatrix} w_1^{k+1} \\ \vdots \\ w_n^{k+1} \end{pmatrix} = \begin{pmatrix} \text{Prox}_{\frac{1}{2\beta_2}g}(D_1 f^{k+1}) \\ \vdots \\ \text{Prox}_{\frac{1}{2\beta_2}g}(D_n f^{k+1}) \end{pmatrix} := \mathcal{T}_3(f^{k+1}) = \mathcal{T}_3(\mathcal{T}_1(v^k)). \tag{4.4}$$

Thus, combining (4.2), (4.3), and (4.4), we get

$$v^{k+1} = \begin{pmatrix} \mathcal{T}_1(\mathcal{T}_f(v^k)) \\ \mathcal{T}_2(\mathcal{T}_f(v^k)) \\ \mathcal{T}_3(\mathcal{T}_f(v^k)) \end{pmatrix}.$$

Let

$$\mathcal{T} := \begin{pmatrix} \mathcal{T}_1(\mathcal{T}_f) \\ \mathcal{T}_2(\mathcal{T}_f) \\ \mathcal{T}_3(\mathcal{T}_f) \end{pmatrix}, \tag{4.5}$$

we have $v^{k+1} = \mathcal{T}(v^k)$, which is the iteration of Algorithm 1. In the following, we show that \mathcal{T} is nonexpansive.

Lemma 4.1 *The operator \mathcal{T} in (4.5) is nonexpansive.*

Proof For any v_1, v_2 , by the definition of \mathcal{T} , we have

$$\begin{aligned} \|\mathcal{T}(v_1) - \mathcal{T}(v_2)\|^2 &= \left\| \begin{pmatrix} \mathcal{T}_1(\mathcal{T}_f(v_1)) - \mathcal{T}_1(\mathcal{T}_f(v_2)) \\ \mathcal{T}_2(\mathcal{T}_f(v_1)) - \mathcal{T}_2(\mathcal{T}_f(v_2)) \\ \mathcal{T}_3(\mathcal{T}_f(v_1)) - \mathcal{T}_3(\mathcal{T}_f(v_2)) \end{pmatrix} \right\|^2 \\ &= \|\mathcal{T}_1(\mathcal{T}_f(v_1)) - \mathcal{T}_1(\mathcal{T}_f(v_2))\|^2 + \|\mathcal{T}_2(\mathcal{T}_f(v_1)) - \mathcal{T}_2(\mathcal{T}_f(v_2))\|^2 \\ &\quad + \|\mathcal{T}_3(\mathcal{T}_f(v_1)) - \mathcal{T}_3(\mathcal{T}_f(v_2))\|^2. \end{aligned} \tag{4.6}$$

Let $f_1 := \mathcal{T}_f(v_1)$ and $f_2 := \mathcal{T}_f(v_2)$, then, by the definition of \mathcal{T}_f , we obtain that

$$f_1 - f_2 = \mathcal{T}_f(v_1) - \mathcal{T}_f(v_2) = M^{-1}(BH)^T(v_1 - v_2). \tag{4.7}$$

Thus, we have

$$\begin{aligned} &\|\mathcal{T}_1(\mathcal{T}_f(v_1)) - \mathcal{T}_1(\mathcal{T}_f(v_2))\| \\ &= \|\mathcal{T}_1(f_1) - \mathcal{T}_1(f_2)\| \\ &= \left\| \text{Prox}_{\frac{\alpha}{2\beta_1}\varphi}(Af_1 + ce) - \text{Prox}_{\frac{\alpha}{2\beta_1}\varphi}(Af_2 + ce) \right\| \\ &\leq \|A(f_1 - f_2)\| = \|AM^{-1}(BH)^T(v_1 - v_2)\|, \end{aligned} \tag{4.8}$$

where the inequality follows from Proposition 4.1 and the last equality follows from (4.7). For the operator $\mathcal{T}_2(\mathcal{T}_f)$, we get

$$\begin{aligned} &\|\mathcal{T}_2(\mathcal{T}_f(v_1)) - \mathcal{T}_2(\mathcal{T}_f(v_2))\| \\ &= \|\mathcal{T}_2(f_1) - \mathcal{T}_2(f_2)\| \\ &\leq \|f_1 - f_2\| = \|M^{-1}(BH)^T(v_1 - v_2)\|, \end{aligned} \tag{4.9}$$

where the inequality holds since \mathcal{T}_2 is nonexpansive. For the operator $\mathcal{T}_3(\mathcal{T}_f)$, we have

$$\begin{aligned} & \|\mathcal{T}_3(\mathcal{T}_f(v_1)) - \mathcal{T}_3(\mathcal{T}_f(v_2))\|^2 \\ &= \|\mathcal{T}_3(f_1) - \mathcal{T}_3(f_2)\|^2 \\ &\leq \sum_{i=1}^n \|D_i(f_1 - f_2)\|^2 \\ &= \|D(f_1 - f_2)\|^2 = \|DM^{-1}(BH)^T(v_1 - v_2)\|^2, \end{aligned} \tag{4.10}$$

where the inequality holds by Proposition 4.1. Therefore, combining (4.8), (4.9), (4.10) and (4.6), we obtain that

$$\begin{aligned} & \|\mathcal{T}(v_1) - \mathcal{T}(v_2)\|^2 \\ &\leq \|AM^{-1}(BH)^T(v_1 - v_2)\|^2 + \|M^{-1}(BH)^T(v_1 - v_2)\|^2 + \|DM^{-1}(BH)^T(v_1 - v_2)\|^2 \\ &= \left\| \begin{pmatrix} A \\ I \\ D \end{pmatrix} M^{-1}(BH)^T(v_1 - v_2) \right\|^2 = \|B^{-1}HM^{-1}(BH)^T(v_1 - v_2)\|^2 = \|v_1 - v_2\|^2, \end{aligned}$$

where the last equality follows from the definition of M . This implies that $\|\mathcal{T}(v_1) - \mathcal{T}(v_2)\| \leq \|v_1 - v_2\|$. Thus, \mathcal{T} is nonexpansive. □

The following lemma shows that \mathcal{T} is asymptotically regular.

Lemma 4.2 *Let $v^k := ((u^k)^T, (z^k)^T, (w^k)^T)^T$ be generated by Algorithm 1, then $\sum_{k=1}^{+\infty} \|v^k - v^{k-1}\|^2$ converges and \mathcal{T} is asymptotically regular.*

Proof The proof follows the lines of the proof of Lemma 3.6 in [19]. For the sake of completeness, we give it here. We consider the Taylor expansion of $\mathcal{Q}(f, u^k, z^k, w^k)$ at f^{k+1} and let $f = f^k$. Since $\mathcal{Q}(f, u^k, z^k, w^k)$ about f is a quadratic function, we have

$$\begin{aligned} \mathcal{Q}(f^k, u^k, z^k, w^k) &= \mathcal{Q}(f^{k+1}, u^k, z^k, w^k) + (\nabla_f \mathcal{Q}(f^{k+1}, u^k, z^k, w^k))^T (f^k - f^{k+1}) \\ &\quad + \frac{1}{2} (f^k - f^{k+1})^T \nabla_f^2 \mathcal{Q}(f^{k+1}, u^k, z^k, w^k) (f^k - f^{k+1}), \end{aligned} \tag{4.11}$$

where $\nabla_f \mathcal{Q}(f^{k+1}, u^k, z^k, w^k)$ and $\nabla_f^2 \mathcal{Q}(f^{k+1}, u^k, z^k, w^k)$ denote the gradient and Hessian matrix of $\mathcal{Q}(f, u^k, z^k, w^k)$ at f^{k+1} , respectively. Since f^{k+1} is a minimizer of $\mathcal{Q}(f, u^k, z^k, w^k)$, we have

$$\nabla_f \mathcal{Q}(f^{k+1}, u^k, z^k, w^k) = 0. \tag{4.12}$$

Moreover,

$$\nabla_f^2 \mathcal{Q}(f^{k+1}, u^k, z^k, w^k) = 2\beta_1 A^T A + 2\beta_2 D^T D + 2\beta_3 I \geq 2\beta_3 I. \tag{4.13}$$

Thus, by substituting (4.12) and (4.13) into (4.11), we get

$$\mathcal{Q}(f^k, u^k, z^k, w^k) - \mathcal{Q}(f^{k+1}, u^k, z^k, w^k) \geq \beta_3 \|f^k - f^{k+1}\|^2.$$

Since $\mathcal{Q}(f^{k+1}, u^{k+1}, z^{k+1}, w^{k+1}) \leq \mathcal{Q}(f^{k+1}, u^k, z^k, w^k)$, we obtain

$$\mathcal{Q}(f^k, u^k, z^k, w^k) - \mathcal{Q}(f^{k+1}, u^{k+1}, z^{k+1}, w^{k+1}) \geq \beta_3 \|f^k - f^{k+1}\|^2. \tag{4.14}$$

Combining (4.2), (4.3) and (4.4), we have

$$\begin{aligned} & \|v^k - v^{k+1}\|^2 \\ &= \|\mathcal{T}_1(f^k) - \mathcal{T}_1(f^{k+1})\|^2 + \|\mathcal{T}_2(f^k) - \mathcal{T}_2(f^{k+1})\|^2 + \|\mathcal{T}_3(f^k) - \mathcal{T}_3(f^{k+1})\|^2 \\ &\leq \|A(f^k - f^{k+1})\|^2 + \|f^k - f^{k+1}\|^2 + \|D(f^k - f^{k+1})\|^2 \\ &\leq (\|A\|^2 + \|D\|^2 + 1)\|f^k - f^{k+1}\|^2, \end{aligned}$$

where the first inequality holds by Proposition 4.1. From [11], we know that $\|D\|^2 \leq 8$ and from [25], we have $\|A\| \leq 1$. Thus,

$$\|v^k - v^{k+1}\|^2 \leq 10\|f^k - f^{k+1}\|^2. \tag{4.15}$$

Substituting (4.15) into (4.14), we have

$$\frac{\beta_3}{10}\|v^k - v^{k+1}\|^2 \leq \mathcal{Q}(f^k, u^k, z^k, w^k) - \mathcal{Q}(f^{k+1}, u^{k+1}, z^{k+1}, w^{k+1}).$$

Since \mathcal{Q} is bounded below, it follows that $\sum_{k=0}^{+\infty} \|v^k - v^{k+1}\|^2$ converges. Moreover,

$$\lim_{k \rightarrow +\infty} \|v^k - v^{k+1}\| = 0.$$

Note that $v^k = \mathcal{T}(v^{k-1}) = \mathcal{T}^2(v^{k-2}) = \dots = \mathcal{T}^{k-1}(v^1) = \mathcal{T}^k(v^0)$, we obtain that

$$\lim_{k \rightarrow +\infty} \|\mathcal{T}^k(v^0) - \mathcal{T}^{k+1}(v^0)\| = 0.$$

This implies that \mathcal{T} is asymptotically regular. Thus, we complete the proof. □

By [45], we know that $e^T u - g^T \log(u)$ is proper, which implies that $\mathcal{Q}(f, u, z, w)$ is proper. Notice that $\mathcal{Q}(f, u, z, w) : \mathbb{R}^n \times \mathbb{R}_{++}^n \times \mathbb{R}^n \times \mathbb{R}^{2n} \rightarrow \mathbb{R} \cup \{+\infty\}$ is continuous, we obtain that $\mathcal{Q}(f, u, z, w)$ is closed [33, Thoerem 1.6]. In order to show the existence of minimizers of $\mathcal{Q}(f, u, z, w)$, by [33, Theorem 1.9], we only need to prove that $\mathcal{Q}(f, u, z, w)$ is coercive, i.e., $\mathcal{Q}(f, u, z, w)$ tends to infinity as $\|(f^T, u^T, z^T, w^T)^T\|$ tends to infinity.

Lemma 4.3 *If $\text{Ker}(A) \cap \text{Ker}(D) = \{0\}$, where $\text{Ker}(\cdot)$ denotes the null space of a matrix, then $\mathcal{Q}(u, z, w, f)$ is coercive.*

Proof When

$$\left\| \begin{pmatrix} f \\ u \\ z \\ w \end{pmatrix} \right\| \rightarrow +\infty,$$

we have $\|u\| \rightarrow +\infty$, or $\|w\| \rightarrow +\infty$, or

$$\left\| \begin{pmatrix} f \\ z \end{pmatrix} \right\| \rightarrow +\infty.$$

Thus, we proceed the discussions by three cases.

Case 1. Suppose that $\|u\| \rightarrow +\infty$. We note that $u - \log(u)$ tends to infinity when u tends to infinity. Thus, we have $\mathcal{Q}(f, u, z, w)$ tends to infinity.

Case 2. Suppose that $\|w\| \rightarrow +\infty$. Then, $\sum_{i=1}^n \|w_i\|$ tends to infinity. Since $u - \log(u)$ is bounded below, we obtain that $\mathcal{Q}(f, u, z, w)$ tends to infinity.

Case 3. Suppose that

$$\left\| \begin{pmatrix} f \\ z \end{pmatrix} \right\| \rightarrow +\infty.$$

We assume that $\|u\|, \|w\|$ are bounded, otherwise the result holds by Cases 1 and 2. If $\|f\|$ tends to infinity and $\|z\|$ is bounded, we have $\|z - f\|^2$ tends to infinity. Thus, $\mathcal{Q}(f, u, z, w)$ tends to infinity. If $\|z\|$ tends to infinity and $\|f\|$ is bounded, we have $\|z - f\|^2$ tends to infinity. Then, $\mathcal{Q}(f, u, z, w)$ tends to infinity. If $\|f\|$ tends to infinity, $\|z\|$ tends to infinity, since $\text{Ker}(A) \cap \text{Ker}(D) = \{0\}$, we have $\|Af\|$ tends to infinity or $\|Df\|$ tends to infinity as $\|f\|$ tends to infinity. Since $\|u\|, \|w\|$ are bounded, we obtain that $\mathcal{Q}(f, u, z, w)$ tends to infinity. Together with Cases 1 and 2, we complete the proof. \square

Remark 4.1 The condition $\text{Ker}(A) \cap \text{Ker}(D) = \{0\}$ in Lemma 4.3 is easy to satisfy. If A is a blurring matrix, then their rows sum up to one, while the differences matrix D (under appropriate boundary conditions) has kernel consisting of constant vectors so that their joint kernel must be the zero vector.

We now show that the set of fixed points of \mathcal{T} is nonempty.

Lemma 4.4 *Suppose that $\text{Ker}(A) \cap \text{Ker}(D) = \{0\}$. Then the set of fixed points of \mathcal{T} is nonempty.*

Proof Since \mathcal{Q} is coercive, by [33, Theorem 1.9], we obtain that the minimizer (f^*, u^*, z^*, w^*) of \mathcal{Q} exists. Then, by [33, Theorem 6.12], we get

$$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \in \begin{pmatrix} \partial_f \mathcal{Q}(f^*, u^*, z^*, w^*) \\ \partial_u \mathcal{Q}(f^*, u^*, z^*, w^*) + N_{\mathbb{R}_{++}^n}(u^*) \\ \partial_z \mathcal{Q}(f^*, u^*, z^*, w^*) \\ \partial_w \mathcal{Q}(f^*, u^*, z^*, w^*) \end{pmatrix},$$

where $N_{\mathbb{R}_{++}^n}(u^*)$ denotes the normal cone of \mathbb{R}_{++}^n at u^* . Since $u^* > 0$, i.e., $u^* \in \text{int}(\mathbb{R}_{++}^n)$, by [33, Proposition 6.5], we have $N_{\mathbb{R}_{++}^n}(u^*) = \{0\}$. This implies that

$$\begin{cases} f^* = \mathcal{T}_f(v^*) = \arg \min_f \{\mathcal{Q}(\cdot, u^*, z^*, w^*)\}, \\ u^* = \mathcal{T}_1(f^*) = \arg \min_u \{\mathcal{Q}(f^*, \cdot, z^*, w^*)\}, \\ z^* = \mathcal{T}_2(f^*) = \arg \min_z \{\mathcal{Q}(f^*, u^*, \cdot, w^*)\}, \\ w^* = \mathcal{T}_3(f^*) = \arg \min_w \{\mathcal{Q}(f^*, u^*, z^*, \cdot)\}, \end{cases}$$

where $v^* := ((u^*)^T, (z^*)^T, (w^*)^T)^T$. Thus, by the definition of \mathcal{T} , we obtain that $v^* = \mathcal{T}(v^*)$, and v^* is a fixed point of \mathcal{T} . \square

Combining Lemmas 4.1, 4.2, 4.4 and Theorem 4.1, we can get the following convergence result.

Theorem 4.2 *Suppose that $\text{Ker}(A) \cap \text{Ker}(D) = \{0\}$. For any initial point $u^0 > 0, z^0, w^0$, the sequence $\{(f^k, u^k, z^k, w^k)\}$ generated by Algorithm 1 converges to a minimizer (f^*, u^*, z^*, w^*) of \mathcal{Q} .*

Remark 4.2 Compared with the alternating minimization algorithms in [19, 21, 22], we consider some constraints in Algorithm 1.



Fig. 1 Original images. From left to right: Cameraman, House, Barbara, Livingroom

5 Experimental Results

In this section, we present numerical results to show the efficiency of the proposed algorithm for the images corrupted by blurs and Poisson noise. The testing images include Cameraman (256×256), House (256×256), Barbara (512×512), and Livingroom (512×512), which are shown in Fig. 1. In our numerical examples, we consider two blurring functions from MATLAB: Average blur and Gaussian blur. For Average blur, the testing kernel is 9×9 , and for Gaussian blur, the testing kernel is 9×9 (standard deviation = 3). All the experiments are performed under Windows 7 and MATLAB R2015b running on a laptop (AMD A10-5750M, @ 2.50GHz, 8G RAM).

Since Poisson noise is data-dependent, the noise levels of the observed images depend on the pixel intensity M . To test different noise levels, we consider different peak intensities of the images. The noisy and blurry images in our test are simulated as follows. The original image f is first scaled with the peak intensities. Then the scaled image is convolved with the blur kernel A and the background is added. Finally, the Poisson noise is added in Matlab using the function **poissrnd**.

To evaluate the performance of different algorithms, the signal-to-noise ratio (SNR) is used to measure the quality of the restored image:

$$SNR = 10 \log_{10} \left(\frac{\|f\|^2}{\|f - \hat{f}\|^2} \right),$$

where f is the original image and \hat{f} is the restored image. We terminate Algorithm 1 by checking whether the relative error of the successive iterates satisfies

$$\frac{\|f^{k+1} - f^k\|}{\|f^k\|} \leq \text{tol}, \tag{5.1}$$

where tol is set to be 10^{-3} in our experiments. The maximum iteration number of Algorithm 1 is set to be 500. For the penalty parameters $\beta_1, \beta_2, \beta_3$ in Algorithm 1, the initial values of $\beta_1, \beta_2, \beta_3$ are set to 0.4, 0.4 and 0.001, respectively, and then $\beta_1, \beta_2, \beta_3$ are increased by $\tau\beta_1, \tau\beta_2, \tau\beta_3$ every twenty iterations, where $\tau = 1.01$ in our experiments. The regularization parameter α will be selected from $\{1, 10, 20, 50, 100, 150, 200, 250, 300\}$ for the fast convergence and satisfactory accuracy. We remark that $\beta_1, \beta_2, \beta_3$ should make the minimizer of (1.4) close to the minimizer of (1.3) in the quadratic penalty (1.4). Unfortunately, large $\beta_i, i = 1, \dots, 3$, usually cause numerical difficulties, since (1.4) becomes very ill-conditioned. Hence, it is more practical to minimize (1.4) with moderate values of $\beta_i, i = 1, \dots, 3$.

Table 1 SNR(dB) values and CPU time(s) of different algorithms for the Cameraman, House, Barbara, and Livingroom images corrupted by Average blur and Poisson noise

Image	c	M	PIDAL		PIDSplit+		PD		Algorithm I	
			SNR	Time	SNR	Time	SNR	Time	SNR	Time
Cameraman	1	20	15.26	7.58	15.15	6.72	15.10	2.67	15.34	0.64
		40	15.57	6.21	15.47	5.73	15.08	2.21	15.79	0.78
		60	15.78	6.14	15.62	4.56	15.00	2.87	16.11	0.82
		80	15.99	5.55	16.01	3.72	14.63	2.39	16.30	0.91
	10	20	15.13	7.77	15.05	7.36	14.70	3.13	15.08	0.75
		40	15.75	7.47	15.64	5.29	14.60	3.50	15.55	0.83
		60	16.03	7.06	16.00	5.84	13.94	3.47	15.96	1.00
		80	16.12	6.08	16.13	3.95	16.09	2.69	16.17	0.98
House	1	20	18.68	9.01	18.33	8.97	18.59	2.98	18.74	0.56
		40	19.15	6.26	18.91	4.88	18.54	2.46	19.76	0.74
		60	19.53	6.36	19.35	3.96	19.19	2.80	20.01	0.84
		80	19.92	6.15	19.76	3.82	19.40	2.72	20.14	0.89
	10	20	18.47	9.07	18.07	8.12	18.35	3.14	18.27	0.75
		40	19.54	7.66	19.27	5.17	19.53	3.53	19.33	0.82
		60	19.95	6.27	19.79	4.48	19.65	3.29	19.79	0.87
		80	20.21	6.50	20.01	3.69	19.89	2.69	20.22	0.91
Barbara	1	20	16.07	56.50	15.91	64.84	16.09	16.78	16.77	3.08
		40	16.35	43.29	16.22	46.23	16.23	16.22	16.46	4.74
		60	16.45	40.58	16.34	36.45	16.17	16.04	16.52	4.69
		80	16.53	36.79	16.44	31.84	16.02	16.27	16.55	4.95
	10	20	15.84	57.62	15.74	69.74	15.69	23.03	15.64	4.32
		40	16.28	51.44	16.18	48.91	16.23	26.83	16.35	3.62
		60	16.49	48.81	16.41	38.74	16.35	23.65	16.56	3.93
		80	16.62	47.93	16.55	34.60	16.72	24.45	16.68	3.14
Livingroom	1	20	16.58	68.75	16.40	69.41	16.52	16.35	16.77	3.08
		40	16.91	41.91	16.75	48.36	16.81	15.34	17.31	3.91
		60	17.08	40.04	16.95	39.40	16.83	17.51	17.50	4.49
		80	17.21	36.81	17.11	32.00	17.10	16.53	17.55	4.86
	10	20	16.47	57.57	16.19	58.62	16.31	24.99	16.27	4.26
		40	17.01	49.01	16.94	48.88	16.98	23.05	17.03	4.49
		60	17.28	44.16	17.24	38.76	17.35	24.46	17.35	3.93
		80	17.44	41.79	17.41	34.73	17.58	22.53	17.46	4.28

5.1 Comparisons to the Existing Methods

In this subsection, we compare our algorithm with PIDAL [15], PIDSplit+ [36], and PD [45]. The stopping criterion of PIDAL, PIDSplit+, and PD is the same as (5.1) and the maximum iteration number of the three algorithms is set to be 1000. For the PIDAL algorithm [15], the Chambolle’s projection algorithm [11] is used to solve the TV denoising subproblem. The number of iterations of Chambolle’s algorithm is set to 20 and the corresponding step size is



Fig. 2 Recovered images (with SNR(dB) and CPU time(s)) of different algorithms on the Cameraman, House, Barbara and Livingroom images corrupted by Average blur and Poisson noise with peak intensity $M = 80$ and $c = 10$. First row: Noisy and blurry images. Second row: Images recovered by PIDAL. Third row: Images recovered by PIDSplit+. Fourth row: Images recovered by PD. Fifth row: Images recovered by Algorithm 1

set to 0.248, which were suggested in [15]. We use the final values of an inner iteration loop as the initial values for the next loop. Moreover, as suggested in [15,36], we set $\gamma = 50/\alpha$ for the PIDAL and PIDSplit+ algorithms, where α is the regularization parameter and γ is the penalty parameter. In the PIDAL, PIDSplit+, and PD algorithms, the parameter α is also

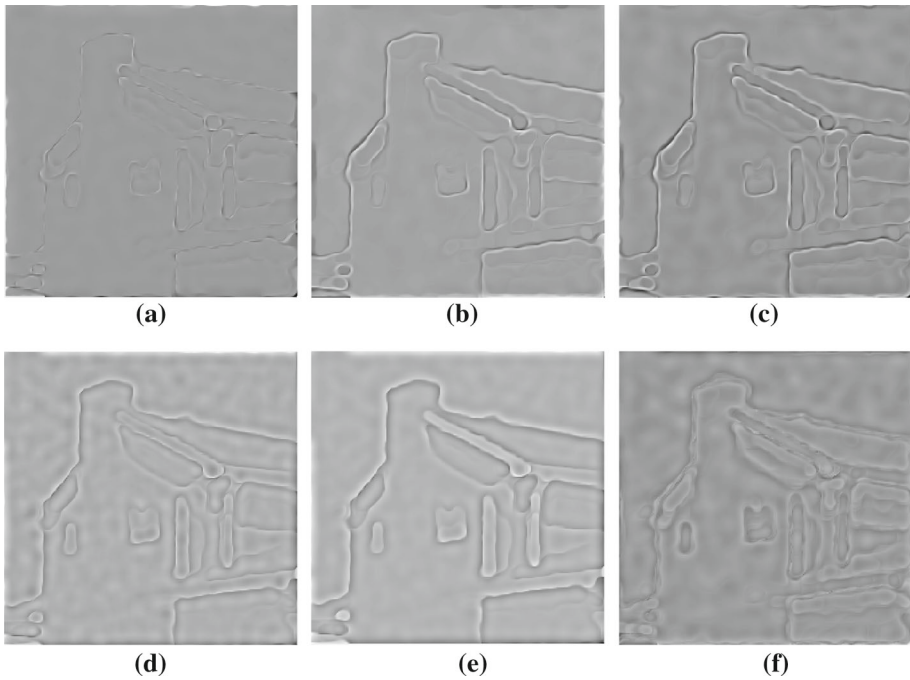


Fig. 3 Difference images of House image among different algorithms in Fig. 2. **a** Difference image between PIDAL and PIDSplit+. **b** Difference image between PD and PIDAL. **c** Difference image between PD and PIDSplit+. **d** Difference image between Algorithm 1 and PIDAL. **e** Difference image between Algorithm 1 and PIDSplit+. **f** Difference image between Algorithm 1 and PD

tested from $\{1, 10, 20, 50, 100, 150, 200, 250, 300\}$ to get the best recovery performance in terms of SNR values and CPU time.

In Table 1, we show the SNR values and CPU time (in seconds) for the Cameraman, House, Barbara, and Livingroom images corrupted by Average blur and Poisson noise with different peak intensities M and c , respectively. We can see that our proposed algorithm outperforms PIDAL, PIDSplit+, and PD for these images in terms of CPU time, and the SNR values of the proposed algorithm are almost the same as those of the other algorithms. Therefore, under almost the same accuracies, our proposed algorithm is much faster than the other three algorithms. Moreover, the PIDAL and PIDSplit+ algorithms need more CPU time than PD for these images.

The visual comparison of the Cameraman, House, Barbara, and Livingroom images corrupted by Average blur and Poisson noise with peak intensity $M = 80$ and $c = 10$ is shown in Fig. 2. We can observe that these images restored by Algorithm 1 are almost the same as those restored by the PIDAL, PIDSplit+, and PD algorithms. However, the CPU time of the proposed algorithm is much less than those of the other three algorithms. Moreover, one can hardly see the differences among these images restored by the four algorithms. In Fig. 3, we show the difference images of the House image of Fig. 2 for different algorithms. It is shown that there are still some differences in these images restored by PIDAL, PIDSplit+, PD, and Algorithm 1.

Table 2 shows the SNR values and CPU time for the Cameraman, House, Barbara, and Livingroom images corrupted by Gaussian blur and Poisson noise with different peak inten-

Table 2 SNR(dB) values and CPU time(s) of different algorithms for the Cameraman, House, Barbara, and Livingroom images corrupted by Gaussian blur and Poisson noise

Image	c	M	PIDAL		PIDSplit+		PD		Algorithm I	
			SNR	Time	SNR	Time	SNR	Time	SNR	Time
Cameraman	1	20	15.81	6.47	15.70	7.01	15.75	2.43	15.82	0.52
		40	16.21	5.54	16.14	4.67	15.88	2.11	16.35	0.71
		60	16.35	5.59	16.29	5.18	15.93	2.86	16.42	0.78
		80	16.43	4.44	16.40	3.03	16.15	2.22	16.58	0.83
	10	20	15.57	6.92	15.49	6.74	15.37	3.11	15.22	0.72
		40	16.27	6.25	16.23	4.64	16.31	3.26	16.09	0.77
		60	16.48	5.30	16.43	3.94	16.46	3.24	16.19	0.85
		80	16.63	5.10	16.60	3.39	16.60	3.02	16.33	0.88
House	1	20	19.48	8.17	18.99	7.14	19.39	2.65	19.67	0.56
		40	20.17	4.98	19.88	4.04	19.87	2.01	20.39	0.71
		60	20.47	5.20	20.19	3.38	20.07	2.19	20.53	0.75
		80	20.75	4.64	20.53	2.83	20.30	2.10	20.69	0.81
	10	20	19.26	6.70	18.97	6.63	19.45	2.99	18.54	0.71
		40	20.44	6.70	20.13	5.53	20.25	3.21	20.02	0.77
		60	21.07	7.08	20.71	4.39	21.01	3.44	20.35	0.84
		80	21.29	6.42	20.94	4.14	21.23	3.44	21.09	0.88
Barbara	1	20	16.37	49.74	16.22	64.44	16.41	15.73	16.38	3.18
		40	16.60	36.47	16.48	40.91	16.67	13.44	16.68	4.17
		60	16.73	41.72	16.62	41.57	16.70	19.37	16.75	4.37
		80	16.81	37.40	16.71	35.20	16.63	18.97	16.80	4.61
	10	20	16.17	49.99	16.10	63.29	16.03	22.30	15.79	4.18
		40	16.60	36.47	16.48	40.91	16.67	21.66	16.46	4.50
		60	16.81	38.86	16.71	36.06	16.65	24.40	16.57	4.77
		80	16.89	38.32	16.80	31.88	16.73	23.38	16.99	5.05
Livingroom	1	20	17.00	48.54	16.84	59.84	17.07	17.25	17.20	3.05
		40	17.33	35.54	17.18	43.67	17.54	15.61	17.75	3.81
		60	17.54	34.65	17.42	35.70	17.67	15.62	17.90	4.26
		80	17.66	34.02	17.58	31.35	17.60	15.88	18.00	4.55
	10	20	16.79	56.58	16.75	88.72	16.73	22.49	16.61	4.26
		40	17.43	45.00	17.37	47.12	17.35	29.29	17.34	4.40
		60	17.73	38.62	17.67	36.89	17.62	24.34	17.63	4.67
		80	17.95	37.57	17.89	31.47	17.83	23.52	17.79	4.96

sities M and c , respectively. It can be seen that the CPU time of the proposed algorithm is less than the other three algorithms. The accuracies of the proposed algorithm are almost the same as those of PIDAL, PIDSplit+, and PD in terms of SNR values. Moreover, the PD algorithm performs better than PIDAL and PIDSplit+ in terms of CPU time.

Figure 4 shows the comparison on the Cameraman, House, Barbara, and Livingroom images corrupted by Gaussian blur and Poisson noise with peak intensity $M = 40$ and $c = 1$. We can observe that the SNR values of the restored images by the four algorithms



Fig. 4 Recovered images (with SNR(dB) and CPU time(s)) of different algorithms on the Cameraman, House, Barbara, and Livingroom images corrupted by Gaussian blur and Poisson noise with peak intensity $M = 40$ and $c = 1$. First row: Noisy and blurry images. Second row: Images recovered by PIDAL. Third row: Images recovered by PIDSplit+. Fourth row: Images recovered by PD. Fifth row: Images recovered by Algorithm 1

are almost the same, and the computational time required by the proposed algorithm is much less than those required by PIDAL, PIDSplit+, and PD.

In Fig. 5, we plot the SNR values versus CPU time for the Cameraman, House, Barbara, and Livingroom images corrupted by Gaussian blur and Poisson noise with $M = 20$, $c = 1$.

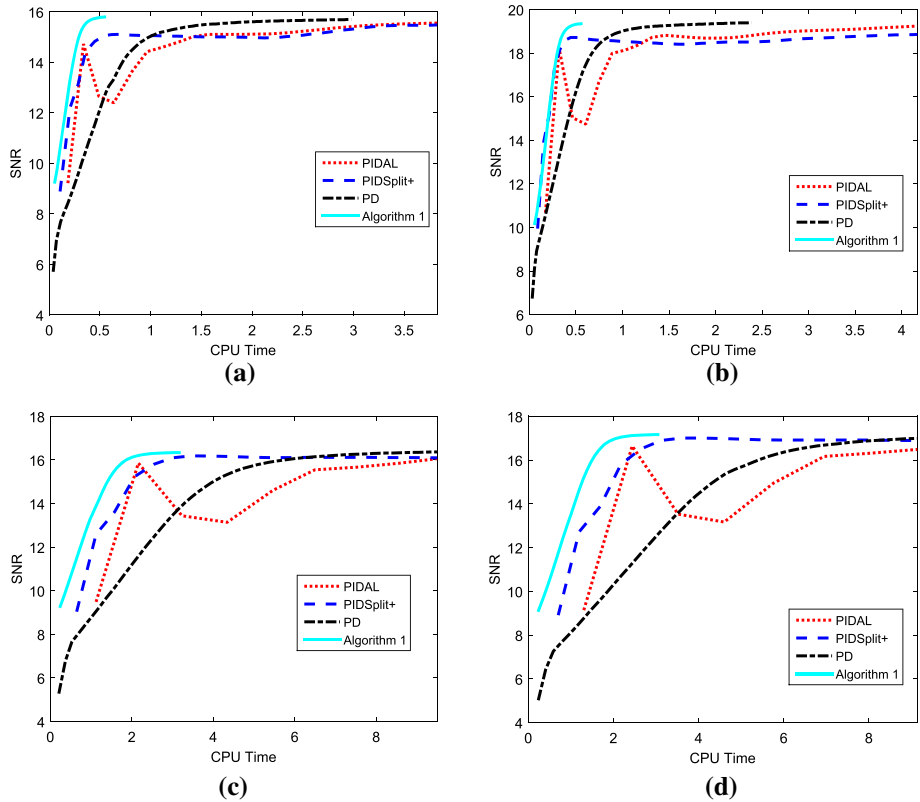


Fig. 5 SNR(dB) values versus CPU time(s) for the images corrupted by Gaussian blur and Poisson noise with $M = 20, c = 1$. **a** Cameraman. **b** House. **c** Barbara. **d** Livingroom

We can observe that the PIDAL and PIDSplit+ algorithms are fast at the initial iterations. But the convergence speed of the two algorithms is slow after reaching a low accuracy. Algorithm 1 needs less CPU time than PIDAL, PIDSplit+, and PD for reaching high accurate solutions. Moreover, the PD algorithm is slower than the proposed algorithm for all the stages. It needs more iterations to reach the same accuracy than the proposed algorithm. The proposed algorithm is fast in all the iterations for low and high accurate solutions. Thus, Algorithm 1 outperforms PIDAL, PIDSplit+ in terms of CPU time.

6 Concluding Remarks

In this paper, we have proposed an alternating minimization algorithm for the restoration of blurred images corrupted by Poisson noise, which minimizes the KL divergence plus TV regularization regularization model with nonnegative constraint. By introducing new variables, the objective function is separable. The approximation model (1.4) is derived from the classical penalty method. Then, we can minimize each variables, alternatingly. Furthermore, based on the fixed point theory, we establish the convergence of the proposed algorithm under very mild assumptions. Our preliminary numerical experiments show that the proposed algorithm is much faster than PIDAL [15], PIDSplit+ [36] and PD [45].

Acknowledgements The authors are grateful to the anonymous referees for their careful reading of this paper and their valuable suggestions, which have helped improve the quality of this paper substantially. The authors would also like to thank Professor You-Wei Wen for sending us the code of the primal dual algorithm in [45].

References

1. Acar, R., Vogel, C.R.: Analysis of bounded variation penalty methods for ill-posed problems. *Inverse Probl.* **10**(6), 1217–1229 (1994)
2. Aubert, G., Aujol, J.-F.: A variational approach to removing multiplicative noise. *SIAM J. Appl. Math.* **68**(4), 925–946 (2008)
3. Bai, M., Zhang, X., Shao, Q.: Adaptive correction procedure for TVL1 image deblurring under impulse noise. *Inverse Probl.* **32**(8), 085004 (2016)
4. Bailey, D.L., Townsend, D.W., Valk, P.E., Maisey, M.N.: *Positron Emission Tomography-Basic Sciences*. Springer, New York (2005)
5. Bardsley, J.M., Goldes, J.: An iterative method for edge-preserving MAP estimation when data-noise is Poisson. *SIAM J. Sci. Comput.* **32**(1), 171–185 (2010)
6. Bauschke, H.H., Combettes, P.L.: *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, New York (2011)
7. Bertero, M., Boccacci, P., Desiderà, G., Vicidomini, G.: Image deblurring with Poisson data: from cells to galaxies. *Inverse Probl.* **25**(12), 123006 (2009)
8. Browder, F.E., Petryshyn, W.V.: The solution by iteration of nonlinear functional equations in Banach spaces. *Bull. Am. Math. Soc.* **72**(3), 571–575 (1966)
9. Brune, C., Sawatzky, A., Burger, M.: Primal and dual Bregman methods with application to optical nanoscopy. *Int. J. Comput. Vis.* **92**(2), 211–229 (2011)
10. Burger, M., Sawatzky, A., Steidl, G.: First order algorithms in variational image processing. In: Glowinski, R., Osher, S., Yin, W. (eds.) *Splitting Methods in Communication, Imaging, Science, and Engineering*, pp. 345–407. Springer, Switzerland (2016)
11. Chambolle, A.: An algorithm for total variation minimization and applications. *J. Math. Imaging Vis.* **20**(1), 89–97 (2004)
12. Combettes, P., Wajs, V.: Signal recovery by proximal forward-backward splitting. *Multiscale Model. Simul.* **4**(4), 1168–1200 (2005)
13. Csiszar, I.: Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems. *Ann. Stat.* **19**(4), 2032–2066 (1991)
14. Dey, N., Blanc-Feraud, L., Zimmer, C., Roux, P., Kam, Z., Olivo-Marin, J.-C., Zerubia, J.: Richardson-Lucy algorithm with total variation regularization for 3D confocal microscope deconvolution. *Microsc. Res. Tech.* **69**(4), 260–266 (2006)
15. Figueiredo, M.A.T., Bioucas-Dias, J.M.: Restoration of Poissonian images using alternating direction optimization. *IEEE Trans. Image Process.* **19**(12), 3133–3145 (2010)
16. Geman, S., Geman, D.: Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.* **6**(6), 721–741 (1984)
17. Golub, G.H., Hansen, P.C., O’Leary, D.P.: Tikhonov regularization and total least squares. *SIAM J. Matrix Anal. Appl.* **21**(1), 185–194 (1999)
18. Green, P.J.: Bayesian reconstructions from emission tomography data using a modified EM algorithm. *IEEE Trans. Med. Imaging* **9**(1), 84–93 (1990)
19. Guo, X., Li, F., Ng, M.K.: A fast ℓ_1 -TV algorithm for image restoration. *SIAM J. Sci. Comput.* **31**(3), 2322–2341 (2009)
20. Hohage, T., Werner, F.: Inverse problems with Poisson data: statistical regularization theory, applications and algorithms. *Inverse Probl.* **32**(9), 093001 (2016)
21. Huang, Y.-M., Lu, D.-Y., Zeng, T.: Two-step approach for the restoration of images corrupted by multiplicative noise. *SIAM J. Sci. Comput.* **35**(6), A2856–A2873 (2013)
22. Huang, Y.-M., Ng, M.K., Wen, Y.-W.: A fast total variation minimization method for image restoration. *Multiscale Model. Simul.* **7**(2), 774–795 (2008)
23. Le, T., Chartrand, R., Asaki, T.J.: A variational approach to reconstructing images corrupted by Poisson noise. *J. Math. Imaging Vis.* **27**(3), 257–263 (2007)
24. Lucy, L.B.: An iterative technique for the rectification of observed distributions. *Astron. J.* **79**(6), 745–754 (1974)
25. Ma, L., Ng, M.K., Yu, J., Zeng, T.: Efficient box-constrained TV-type- ℓ^1 algorithms for restoring images with impulse noise. *J. Comput. Math.* **31**(3), 249–270 (2013)

26. Molina, R.: On the hierarchical Bayesian approach to image restoration: applications to astronomical images. *IEEE Trans. Pattern Anal. Mach. Intell.* **16**(11), 1122–1128 (1994)
27. Ng, M.K.: *Iterative Methods for Toeplitz Systems*. Oxford University Press, London (2004)
28. Ng, M.K., Chan, R.H., Tang, W.-C.: A fast algorithm for deblurring models with Neumann boundary conditions. *SIAM J. Sci. Comput.* **21**(3), 851–866 (1999)
29. Ollinger, J.M., Fessler, J.A.: Positron-emission tomography. *IEEE Signal Process. Mag.* **14**(1), 43–55 (1997)
30. Opial, Z.: Weak convergence of the sequence of successive approximations for nonexpansive mappings. *Bull. Am. Math. Soc.* **73**(4), 591–597 (1967)
31. Panin, V.Y., Zeng, G.L., Gullberg, G.T.: Total variation regulated EM algorithm. *IEEE Trans. Nuclear Sci.* **46**(6), 2202–2210 (1999)
32. Richardson, W.H.: Bayesian-based iterative method of image restoration. *J. Opt. Soc. Am. A* **62**(1), 55–59 (1972)
33. Rockafellar, R.T., Wets, R.J.-B.: *Variational Analysis*. Springer, Berlin (2009)
34. Rudin, L.I., Osher, S.: Total variation based image restoration with free local constraints. In: *Proceedings of IEEE International Conference Image Process*, vol 1, pp. 31–35. Austin, TX, (1994)
35. Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. *Physica D* **60**(1–4), 259–268 (1992)
36. Setzer, S., Steidl, G., Teuber, T.: Deblurring Poissonian images by split Bregman techniques. *J. Vis. Commun. Image Rep.* **21**(3), 193–199 (2010)
37. Shepp, L.A., Vardi, Y.: Maximum likelihood reconstruction for emission tomography. *IEEE Trans. Med. Imaging* **1**(2), 113–122 (1982)
38. Shi, J., Osher, S.: A nonlinear inverse scale space method for a convex multiplicative noise model. *SIAM J. Imaging Sci.* **1**(3), 294–321 (2008)
39. Starck, J.L., Pantin, E., Murtagh, F.: Deconvolution in astronomy: A review. *Publ. Astron. Soc. Pac.* **114**(800), 1051–1069 (2002)
40. Steidl, G., Teuber, T.: Removing multiplicative noise by Douglas-Rachford splitting methods. *J. Math. Imaging Vis.* **36**(2), 168–184 (2010)
41. Teuber, T., Steidl, G., Chan, R.H.: Minimization and parameter estimation for seminorm regularization models with I-divergence constraints. *Inverse Probl.* **29**(3), 035007 (2013)
42. Tikhonov, A.N., Arsenin, V.Y.: *Solutions of Ill-Posed Problems*. Winston, Washington, DC (1977)
43. Vio, R., Bardsley, J., Wamsteker, W.: Least-squares methods with Poissonian noise: Analysis and comparison with the Richardson-Lucy algorithm. *Astron. Astrophys.* **436**(2), 741–755 (2005)
44. Wang, Y., Yang, J., Yin, W., Zhang, Y.: A new alternating minimization algorithm for total variation image reconstruction. *SIAM J. Imaging Sci.* **1**(3), 248–272 (2008)
45. Wen, Y.-W., Chan, R.H., Zeng, T.: Primal-dual algorithms for total variation based image restoration under Poisson noise. *Sci. China Math.* **59**(1), 141–160 (2016)
46. Wen, Y.-W., Ng, M.K., Huang, Y.-M.: Efficient total variation minimization methods for color image restoration. *IEEE Trans. Image Process.* **17**(11), 2081–2088 (2008)
47. Yang, J., Zhang, Y., Yin, W.: An efficient TVL1 algorithm for deblurring multichannel images corrupted by impulsive noise. *SIAM J. Sci. Comput.* **31**(4), 2842–2865 (2009)
48. Zeng, T., Li, X., Ng, M.K.: Alternating minimization method for total variation based wavelet shrinkage model. *Commun. Comput. Phys.* **8**(5), 976–994 (2010)
49. Zhang, X., Javidi, B., Ng, M.K.: Automatic regularization parameter selection by generalized cross-validation for total variational Poisson noise removal. *Applied Opt.* **56**(9), D47–D51 (2017)