Inexact alternating direction method based on Newton descent algorithm with application to Poisson image deblurring

Dai-Qiang Chen

Abstract The recovery of images from the observations that are degraded by a linear operator and further corrupted by Poisson noise is an important task in modern imaging applications such as astronomical and biomedical ones. Gradient-based regularizers involving the popular total variation semi-norm have become standard techniques for Poisson image restoration due to its edge-preserving ability. Various efficient algorithms have been developed for solving the corresponding minimization problem with non-smooth regularization terms. In this paper, motivated by the idea of the alternating direction minimization algorithm and the Newton’s method with upper convergent rate, we further propose inexact alternating direction methods utilizing the proximal Hessian matrix information of the objective function, in a way reminiscent of Newton descent methods. Besides, we also investigate the global convergence of the proposed algorithms under certain conditions. Finally, we illustrate that the proposed algorithms outperform the current state-of-the-art algorithms through numerical experiments on Poisson image deblurring.

Keywords Image deblurring · Newton descent method · Inexact alternating direction method · Total variation · Poisson noise

1 Introduction

Image deblurring is a classical ill-conditioned problem in many fields of applied sciences, including astronomy imaging and biomedical imaging. Mathematically, image blurring process in such applications can often be described as follows. For simplification, we denote the $m \times n$ image as a one-dimensional vector in $\mathbb{R}^N (N = mn)$ by concatenating their columns. Let $u \in \mathbb{R}^N$ be the original image. The degradation model is described by

$$g = Ku$$

where $g \in \mathbb{R}^N$ is the observed image and $K \in \mathbb{R}^{N \times N}$ is a linear blurring operator. Since the linear operator $K$ cannot be inverted, and $g$ is also possibly contaminated by random noises, the recovery of $u$ from the noisy version $f$ of the blurred observation $g$ is a ill-posed problem. Variational image restoration methods based on the regularization technique are the most popular approach for solving this problem. Typically, the total variation (TV) regularizer-based model can be formulated as

$$\min_{u \in U} D_f(u) + \lambda \|\nabla u\|_1$$

where $U = [u_{\text{min}}, u_{\text{max}}]^N$ is the range of $u$, $\lambda > 0$ is a regularization parameter, $D_f(u)$ is a data fidelity term which is derived from the noise distribution, and $\|\nabla u\|_1 = \sum_{i=1}^N \| (\nabla u)_i \|_2$ with $(\nabla u)_i = ((\nabla u)^1_i, (\nabla u)^2_i)$ being the...
total variation regularization imposing the prior on the unknown image $u$.

Generally, the data fidelity term controls the closeness between the original image $u$ and the observed image $f$. It takes different forms depending on the type of noise being added. It is well known that the $L_2$-norm fidelity is used for the additive Gaussian noise. However, non-Gaussian noises are also presented in the real imaging; e.g., Poisson noise is generally observed in photon-limited images such as electronic microscopy [1], positron emission tomography [2], and single photon emission computerized tomography [3]. Due to its important applications in medical imaging, linear inverse problems in the presence of Poisson noise have received much interest in the literature [8,14,15]. Based on the statistics of Poisson noise and maximum a posterior (MAP) likelihood estimation approach, we can obtain the classical TV–KL model for Poisson image deblurring:

$$
\min_{u \geq 0} (1, Ku) - \langle f, \log(Ku) \rangle + \lambda \|\nabla u\|_1. \tag{3}
$$

Since the pixel values of images represent the number of discrete photons incident over a given time interval in this application, we demand that $u \geq 0$ in model (3). Another selection for the regularizer term is the wavelet tight framelets [6–8], which have also been proved to be efficient, but may need more computational cost associated with the wavelet transform and inverse transform. In the last several years, the relationship between the total variation and wavelet framelet has also been revealed [9,10].

In this paper, we focus our attention on the TV–KL model (3). Due to the complex form of the fidelity term in (3), the ill-posed inverse problem in the presence of Poisson noise has attracted less interest in the literature than their Gaussian counterpart. Sawatzky et al. [11] proposed an EM–TV algorithm for Poisson image deblurring which has been shown to be more efficient than earlier methods, such as TV penalized Richardson–Lucy algorithm [17]. Bonettini et al. [12,13] also developed gradient projection methods for TV-based image restoration. Recently, the augmented Lagrangian framework [14–16], which has been successfully applied to various image processing tasks, has been used for solving the TV–KL model. In particular, in [15] a very effective alternating direction method of multipliers (ADMM) called PIDAL was proposed for image deblurring in the presence of Poisson noise, where a TV denoising problem is solved by Chambolle’s algorithm in each iteration. It has been proved to be more efficient than the previous augmented Lagrangian algorithms such as the PIDAL algorithm. Refer to the experiments below for details.

The main contribution of this work is to propose a novel inexact alternating direction method utilizing the second-order information of the objective function. Specifically, in one sub-minimization problem of the proposed algorithm, the solution is obtained by a one-step iteration, in a way reminiscent of Newton descent methods [26]. In other words, the second-order derivative of the corresponding objective function in the sub-minimization problem is just approximated by a proximal Hessian matrix which can be computed easily, rather than a constant multiplied by the identity matrix. The improved iterative algorithm is proved to be more efficient than the current state-of-the-art methods with application to Poisson image deblurring, including the PIDAL algorithm and the linearized alternating direction methods.

The rest of this paper is organized as follows. In Sect. 2, we briefly review the recently proposed proximal linearized alternating direction (PLAD) method [18]. In Sect. 3, we develop an inexact alternating direction method based on the Newton descent algorithm. In Sect. 4, the numerical examples on Poisson image deblurring problem are reported to

---

**SIViP (2017) 11:89–96**

---

**Springer**
compare the proposed algorithms with the recent state-of-the-art algorithms.

2 Existing algorithms

In this section, we briefly review the well-known ADMM method [15] and the PLAD method [18]. First of all, we consider the TV regularized minimization problem shown in (2).

Note that this problem can also be reformulated as a constrained optimization problem as follows

$$\min_{u \in U, d} \left\{ D_f(u) + \lambda \|d\|_1 \mid \nabla u = d \right\}. \quad (4)$$

The augmented Lagrangian function for (4) is given by

$$\mathcal{L}_\alpha(u, d, p) = D_f(u) + \lambda \|d\|_1 + \langle p, d - \nabla u \rangle + \frac{\alpha}{2} \|d - \nabla u\|_2^2. \quad (5)$$

Therefore, the well-known ADMM for solving (4) can be formulated as

$$\begin{align*}
    u^{k+1} &= \arg \min_{u \in U} \left\{ D_f(u) + \langle p^k, d^k - \nabla u \rangle + \frac{\alpha}{2} \|d^k - \nabla u\|_2^2 \right\}, \\
    d^{k+1} &= \arg \min_d \left\{ \lambda \|d\|_1 + \langle p^k, d - \nabla u^{k+1} \rangle + \frac{\alpha}{2} \|d - \nabla u^{k+1}\|_2^2 \right\}, \\
    p^{k+1} &= p^k + \alpha(d^{k+1} - \nabla u^{k+1}).
\end{align*} \quad (6)$$

The solution $u^{k+1}$ of the first subproblem in (6) satisfies the first-order optimality condition, i.e., it is the solution of the following nonlinear system of equations:

$$\nabla D_f(u) - \alpha \Delta u + \text{div}(p^k + \alpha d^k) = 0$$

which has no closed solution. Note that we have $\text{div} = -\nabla^T$.

Therefore, the linearization of the convex function $D_f(u) + \frac{\alpha}{2} \|d - \nabla u\|_2^2$ is adopted and the first subproblem of (6) is simplified as

$$u^{k+1} = \arg \min_{u \in U} \left\{ \nabla D_f(u^k) + \alpha \text{div}(d^k - \nabla u^k), u^k \right\} + \left\{ p^k, d^k - \nabla u^k \right\} + \frac{1}{2\alpha} \|u^k - u^{k+1}\|_2^2. \quad (7)$$

which has a closed solution. Substituting the first subproblem in (6) with (7), we obtain the PLAD algorithm proposed in [18]. Note that the second-order information of the objective function is just approximated by $\frac{1}{2} I$, which is obviously inexact.

Choose $D_f(u) = \langle 1, Ku \rangle - \langle f, \log(Ku) \rangle$ in (2). Then we obtain the PLAD algorithm for solving the TV–KL model as follows:

$$\begin{align*}
    u^{k+1} &= \mathcal{P}_U \left( u^k - \delta \left( K^T \left( 1 - \frac{J}{\lambda \sigma^2} \right) + \alpha \text{div}(d^k - \nabla u^k) + \text{div} p^k \right) \right), \\
    d^{k+1} &= \text{shrink} \left( \nabla u^{k+1} - p^k, \frac{\sigma}{\alpha} \right), \\
    p^{k+1} &= p^k + \alpha(d^{k+1} - \nabla u^{k+1})
\end{align*} \quad (8)$$

where $\mathcal{P}_U$ denotes the projection onto the set $U$. In order to avoid the special case of $Ku^k = 0$ in the $u$-iteration step, we choose $u_{\text{min}} = 1$ in the followings, which is also adopted for the proposed algorithms. The shrinkage operator is componentwise, i.e., it is defined by

$$\text{shrink}(s, c)_i = \max(\|s_i\|_2 - c, 0) \frac{s_i}{\|s_i\|_2}$$

where $s_i \in \mathbb{R}^2$.

The PLAD algorithm can also be regarded as a linearized version of the state-of-the-art PDHG proposed in [23–25]. Interested readers can refer to the literature [19] for detailed elaboration about the equivalence between both algorithms.

3 Proposed inexact alternating direction method based on the Newton descent algorithm

3.1 Algorithm description

Consider the $u$-subproblem in (6). Let

$$G(u) = D_f(u) + \langle p^k, d^k - \nabla u \rangle + \frac{\alpha}{2} \|d^k - \nabla u\|_2^2.$$

It is easy to observe from the first formulas of (8) that the solution $u^{k+1}$ in the PLAD algorithm is obtained by

$$u^{k+1} = \mathcal{P}_U(u^k - \delta \nabla G(u^k)) \quad (9)$$

which implies that an approximate solution of the $u$-subproblem in (6) is obtained by an one-step projection gradient descent algorithm, which only utilizes the first-order information of the objective function, and typically has a sub-linear convergence rate. It is well known that Newton or quasi-Newton methods, which further utilize the Hessian matrix of the objective function, have been presented with a super-linear convergence rate. This fact motivates us to design a more efficient algorithm based on the Newton methods, i.e., the solution $u^{k+1}$ is obtained by an one-step projection Newton descent algorithm as follows.

$$u^{k+1} = \mathcal{P}_U(u^k - \omega^k (\nabla^2 G(u^k))^{-1} \nabla G(u^k)) \quad (10)$$

where $\omega^k \in [0, 1]$ is the relaxed parameter.

In what follows, we consider the special case of TV–KL model for Poisson image deblurring. In this case, we have
In the Poisson image deblurring problem, we always choose $\omega$ for guaranteeing the convergence of the proposed algorithm.

$$
\delta^k = \frac{\left(1 - \frac{f}{K u^k}\right) - \left(1 - \frac{f}{K u^{k-1}}\right)}{\|K u^k - K u^{k-1}\|_2^2}.
$$  (15)

The whole process of the proposed algorithm is summarized as Algorithm 1. It is observed that the update of $\delta^k$ introduces the extra convolution operation including in $K u^k$. Therefore, one simple strategy is to use a fixed value during the iteration, i.e., $\delta^k \equiv \delta$, where $\delta$ is a constant. In this setting, we abbreviate the proposed algorithm as IADMND.

**Algorithm 1** Inexact alternating direction method based on the Newton descent algorithm with adaptive parameters (IADMNDA) for Poisson image deblurring

**Input:** observation $f$; regularization parameter $\lambda$; parameters $\delta_0$ and $\alpha$; inner iteration number $m$.

**Initialization:** $k = 0; u^0 = f; d^0 = \nabla f; p^0 = 0; \delta^0 = \delta_0$; $\omega^{-1} = 1$.

**Iteration:**

(i) update $u$:

$$
r^k = (\delta^k K T K + \alpha \nabla T \nabla)^{-1} \left( K T \left(1 - \frac{f}{K u^k}\right) + \alpha \text{div}(d^k - \nabla u^k) + \text{div} p^k\right)
$$

update $\omega^k$ according to (13);

$$
u^{k+1} = u^k - \omega^k r^k;
$$

(ii) update $d$:

$$
d^{k+1} = \text{shrink} \left( \nabla u^{k+1} - \frac{\omega^k}{\alpha}, \frac{1}{\alpha}\right);
$$

(iii) update $p$:

$$
p^{k+1} = p^k + \alpha (d^{k+1} - \nabla u^{k+1});
$$

update $\delta^{k+1}$ according to (15);

(iv) $k = k + 1$;

until some stopping criterion is satisfied.

**Output** the recovered image $u = u^{k+1}$.

### 3.2 Convergence analysis

In this subsection, we further investigate the global convergence of the proposed IADMND(A) algorithms for Poisson image deblurring under certain conditions. The bound constrained TV regularized minimization problem (4) can be reformulated as

$$
\min_{u,d} \left\{ D_f(u) + \iota_{U}(u) + \lambda \|d\|_1 \mid \nabla u = d \right\}
$$  (16)

where $\iota_U$ denotes the indicator function of set $U$, i.e., $\iota_U(x) = 0$ if $x \in U$ and $\iota_U(x) = +\infty$ otherwise.

Assume that $(u^*, d^*)$ is one solution of the above optimization problem with $D_f(u) = (1, Ku) - (f, \log(Ku))$, and $p^*$ is the corresponding Lagrangian multiplier. Then the point $(u^*, d^*, p^*)$ is a Karush–Kuhn–Tucker (KKT) [29] point of problem (16), i.e., it satisfies the following conditions:
\[ \nabla D_f(u^*) - \nabla^T p^* + \partial \psi_U(u^*) \geq 0, \]
\[ \lambda \partial \| d^* \|_1 + p^* \geq 0, \]
\[ d^* = \nabla u^* \]

where \( \partial \psi_U(u^*) \) denotes the set of the sub-differential of \( \psi_U \) at \( u^* \) and \( \partial \| d^* \|_1 \) denotes the set of the sub-differential of \( \| \cdot \|_1 \) at \( d^* \). From the literature [30], we know that \( \partial \psi_u(u^*) \) is also equal to the normal cone \( \mathcal{N}_U(u^*) \) at \( u^* \). Besides, assume that the convex function \( D_f(u) \) satisfies:

\[ \gamma_D \| K(u_1 - u_2) \|_2^2 \leq (u_1 - u_2)^T \nabla^2 D_f(u)(u_1 - u_2) \]
\[ \leq \gamma_D \| K(u_1 - u_2) \|_2^2 \]

(18)

for any \( u, u_1, u_2 \in U \), where \( \gamma_D \) and \( \gamma_D \) are two positive constants (the estimation of \( \gamma_D \) and \( \gamma_D \) is discussed at the end of this section).

**Theorem 1** (The convergence of the proposed IADMND algorithm) Let \( \{ u^k, d^k, p^k \} \) be the sequence generated by the IADMND algorithm with \( \delta_{\min} = \min_k \delta^k \geq \gamma_D \), \( \delta_{\max} = \max_k \delta^k < +\infty \), \( \max_k (\delta^{k+1} - \delta^k) \leq \gamma_D \) and \( \| K^T K \|_2 > 0 \). Then \( \{ u^k \} \) converges to a solution of the minimization problem (16).

**Proof** A sketch of the proof is given in “Appendix”. \( \square \)

**Theorem 2** (The convergence of the proposed IADMND algorithm) Let \( \{ u^k, d^k, p^k \} \) be the sequence generated by the IADMND algorithm with \( \delta \geq \gamma_D \), and \( \| K^T K \|_2 > 0 \). Then \( \{ u^k \} \) converges to a solution of the minimization problem (16).

**Proof** The proof is analogous to that presented in Theorem 1, and the only difference lies in that \( \delta^k \) is replaced by a constant \( \delta \). Here we neglect the proof due to limited space. \( \square \)

In the above conclusions, we observe that the constants \( \gamma_D \) and \( \gamma_D \) in (18) are crucial for the convergence of the proposed algorithms. Similarly to [18], we can easily deduce that \( \gamma_D \) and \( \gamma_D \) can be simply approximated by \( 1/E(f) \) or \( 1/E(f)(1 + \text{Var}(f)/E^2(f)) \) due to \( E(f) = E(Ku) \). The condition of Theorem 1 can be satisfied by modifying the update formula of \( \delta^k \) as \( \delta^k = \min \left\{ \delta^{k-1} + \gamma_D, \max \left\{ \delta^k, \gamma_D \right\}, M \right\} \), where \( \delta^k \) is computed by the formula (15) in the kth iteration and \( M \) is a large positive number such as \( 10^6 \). However, through experiments on the test images shown in Sect. 4, we observe that the convergence speeds of the IADMND algorithms using the new update strategy and the original update strategy in (15) are almost the same, and hence, we adopt (15) in the following experiments.

### 4 Numerical examples

In this section, the proposed algorithms are compared with the widely used augmented Lagrangian methods for Poisson image restoration [15] and the recently proposed PLAD algorithm [18], which can also be understood from the view of the linearized PDHG algorithm. The codes of proposed algorithms and methods used for comparison are written entirely in MATLAB, and all the numerical examples are implemented under MATLAB 2009 running on a laptop with an Intel Core i5 CPU (2.8 GHz) and 8 GB Memory. In the following experiments, four standard nature images (see Fig. 1), which consist of complex components in different scales and with different patterns, are used for our test.

In the proposed IADMND(A) algorithm, there are two parameters needed to be manually adjusted. One is the regularization parameter \( \alpha \), and the other is the penalty parameter \( \lambda \). It is well known that \( \lambda \) is decided by the noise level, and the value of \( \alpha \) does influence the convergence speed of the proposed algorithm. Here we use the strategy similarly to that...
Table 1 Parameter setting for the numerical experiment

| Image       | Blur kernel | $I_{\text{max}}$ | PIDAL [15] | PLAD [18] | IADMND | IADMNDA |
|-------------|-------------|------------------|------------|-----------|--------|---------|
|             |             |                  |            |           |        |         |
| 9 × 9 Gaussian kernel | 7 × 7 uniform blur |
| Cameraman   | Gaussian    | 100              | 13.55/56/2.76 | 13.57/91/2.57 | 13.53/56/1.51 | 13.55/53/1.74 |
|             |             | 200              | 14.36/50/2.26 | 14.23/132/3.57 | 14.35/46/1.19 | 14.36/47/1.47 |
|             |             | 500              | 15.44/56/2.64 | 15.21/109/3.18 | 15.34/64/1.84 | 15.42/42/1.36 |
|             | Uniform     | 100              | 11.31/83/3.88 | 11.17/139/3.65 | 11.32/63/1.75 | 11.32/58/1.78 |
|             |             | 200              | 11.83/85/4.04 | 11.74/199/5.58 | 11.83/54/1.56 | 11.82/67/2.32 |
|             |             | 500              | 12.68/99/5.02 | 12.31/199/5.55 | 12.62/76/2.11 | 12.66/56/1.92 |
| Barbara     | Gaussian    | 100              | 9.93/45/2.03  | 9.95/99/1.28  | 9.92/49/1.12  | 9.92/38/1.43  |
|             |             | 200              | 10.55/39/1.89 | 10.72/120/2.94 | 10.54/39/1.00 | 10.54/33/1.15 |
|             |             | 500              | 12.23/39/1.78 | 12.10/107/2.48 | 12.15/56/1.44 | 12.22/31/1.01 |
|             | Uniform     | 100              | 9.04/67/3.23  | 9.09/116/3.10 | 9.02/54/1.39  | 9.02/45/1.67  |
|             |             | 200              | 9.19/58/2.92  | 9.29/126/3.34 | 9.18/46/1.17  | 9.17/38/1.12  |
|             |             | 500              | 9.56/59/2.76  | 9.63/124/3.28 | 9.56/68/1.73  | 9.55/35/1.04  |
| Bridge      | Gaussian    | 100              | 10.51/54/2.70 | 10.53/93/2.56 | 10.50/52/1.42 | 10.51/45/1.56 |
|             |             | 200              | 11.26/48/2.32 | 11.27/129/3.46 | 11.25/45/1.32 | 11.26/41/1.48 |
|             |             | 500              | 12.08/43/2.12 | 12.15/110/2.96 | 12.05/63/1.62 | 12.08/35/1.06 |
|             | Uniform     | 100              | 8.43/81/3.96  | 8.43/137/3.62 | 8.42/61/1.70  | 8.43/51/1.64  |
|             |             | 200              | 9.00/72/3.37  | 9.03/161/4.37 | 8.99/54/1.39  | 8.99/46/1.50  |
|             |             | 500              | 9.64/69/3.32  | 9.59/171/4.62 | 9.63/81/2.11  | 9.64/42/1.39  |
| Boat        | Gaussian    | 100              | 12.77/46/8.46 | 12.82/80/9.20 | 12.73/49/5.57 | 12.75/40/5.19 |
|             |             | 200              | 13.51/39/7.35 | 13.54/121/13.35 | 13.49/39/4.85 | 13.50/37/5.21 |
|             |             | 500              | 14.44/39/7.53 | 14.59/94/9.81 | 14.44/53/6.32 | 14.44/33/4.48 |
|             | Uniform     | 100              | 10.41/64/12.48 | 10.40/111/12.26 | 10.41/53/5.83 | 10.42/46/5.85 |
|             |             | 200              | 10.91/65/13.04 | 10.96/199/22.28 | 10.90/47/5.15 | 10.90/43/5.53 |
|             |             | 500              | 11.74/68/13.10 | 11.69/199/22.70 | 11.73/69/7.71 | 11.74/38/4.01 |

Bold values denote the highest SNR or the shortest CPU time.
Rudin–Osher–Fatemi (ROF) denoising problem is included in each iteration of the PIDAL algorithm, and it is solved by using a small and fixed number of iterations (just 5) of Chambolle’s algorithm. For more details, refer to [15].

In the following numerical experiments, the stopping criterion \( \|u^{k+1} - u^k\|_2 \leq 2 \times 10^{-4} \|u^k\|_2 \) is used for all the algorithms here. Table 2 lists the SNR values, the number of iterations and CPU time of different algorithms for images with different blur kernels and noise levels. In this table, “Gaussian” and “Uniform” denote a 9 \( \times \) 9 Gaussian kernel of unit variance and a 7 \( \times \) 7 uniform blur kernel, respectively. The two cases can be seen as examples of mild blur and strong blur. Besides, “\( \cdot / \cdot / \cdot \)” denotes the SNR values, iteration numbers and CPU time in sequence. Note that the iteration numbers of the PIDAL algorithm represent the outer iteration numbers.

From the results in Table 2 we observe that the proposed algorithms are much faster than the PIDAL and PLAD algorithms, and meanwhile, the SNR values of the recovered images achieved with the proposed algorithms are comparable to those achieved with the PIDAL and PLAD algorithms. Therefore, it is verified that the proposed strategy is more efficient than the simple approximation of an identity matrix multiplied by some constant in the PLAD algorithm. It is also noted that the iteration numbers of the IADMNDA algorithm are the least in most cases. However, the update of \( \delta^k \) generates extra computational cost in the IADMNDA algorithm, which makes its implementation time longer than the IADMNDA algorithm in some cases.

### 5 Conclusion

In this article, through analyzing the existed drawback of recently proposed linearization methods for image restoration, we develop an inexact alternating direction method based on the Newton descent algorithm. Compared with the existing algorithms, the main difference is that the second-order derivative of the objective function is just approximated by a proximal Hessian matrix in the proposed algorithm, rather than an identity matrix multiplied by a constant. Besides, we also propose a strategy for updating the proximal Hessian matrix. The convergence of the proposed algorithms is further investigated under certain conditions, and numerical experiments demonstrate that the proposed algorithms outperform the widely used linearized augmented Lagrangian methods in the computational time.

### 6 Appendix: Proof of Theorem 1

The proof of Theorem 1 is similar to the previous literature [19,20]. Only some changes are needed due to the introduction of the Newton descent algorithm. The entire proof is not reproduced here due to limited space. Just enough is sketched to make the changes clear.

**Proof** Denote \( L_k = (\omega^k)^{-1} (\delta^k K^T K + \alpha \nabla^T \nabla) \). According to the iterative formula with respect to \( u \), we know that \( u^{k+1} \) is the solution of the minimization problem

\[
\min_{u \in U} \left\{ (\nabla D_f(u^k) + \alpha \text{div}(d^k - \nabla u^k) + \nabla p^k, \right. \\
left. u - u^k + \frac{1}{2}(u - u^k)^T L_k(u - u^k) \right\}. \tag{19}
\]

Therefore, the sequence \( \{u^k, d^k, p^k\} \) generated by the IADMNDA algorithm satisfies

\[
\begin{align*}
\nabla D_f(u^k) + L_k(u^{k+1} - u^k) + \alpha \nabla^T \nabla u^k - d^k - \alpha^{-1} p^k + \delta_k(u^{k+1}) & \geq 0, \\
\lambda \|d^{k+1}\|_1 + \alpha(d^{k+1} - \nabla u^{k+1} + \alpha^{-1} p^{k+1}) & \geq 0, \\
p^{k+1} = p^k + \alpha(d^{k+1} - \nabla u^{k+1})
\end{align*}
\tag{20}
\]

Because \( (u^*, d^*, p^*) \) is one solution of (16), it is also the KKT point that satisfies (17).

Similarly to [19,20], denote the errors by \( e^k = u^k - u^*, \)
\( d_k = d^k - d^* \), and \( p_k = p^k - p^* \). Utilizing the subtraction between (20) and (17), taking the inner product with \( u_k^{k+1} \), \( d_k^{k+1} \) and \( p_k^* \), and removing the non-negative terms, we further obtain that

\[
\begin{align*}
\frac{1}{\alpha} \|p_k^{k+1}\|_2^2 + \alpha \|d_k^{k+1}\|_2^2 + \frac{1}{\alpha} \|u_k^{k+1}\|_2^2 + \frac{1}{\alpha} \|u_k^{k+1}\|_2^2 D_f(\zeta_k) \\
+ \frac{1}{\alpha} \|u_k^*\|_2^2 D_f(\zeta_k) - \alpha \|\nabla u_k^*\|_2^2 \right) & + \alpha \|d_k^*\|_2^2 - \alpha \|\nabla u_k^*\|_2^2 \\
+ \left( \|u_k^{k+1} - u^*\|_2^2 - \|u_k^{k+1} - u^*\|_2^2 \right) & - \alpha \nabla (u^{k+1} - u^*)^2 \right) \\
\leq \frac{1}{\alpha} \|p_k^*\|_2^2 + \alpha \|d_k^*\|_2^2 + \frac{1}{\alpha} \|u_k^*\|_2^2 \right), - \alpha \nabla (u^{k+1} - u^*)^2 \right) \leq 0.
\tag{21}
\end{align*}
\]

where \( \zeta_k \in [u^*, u_k^*] \), with \([u^*, u_k^*]\) denoting the line segment between \( u^* \) and \( u_k^* \). Since \( \delta_{\min} \geq \frac{1}{\gamma D} \), we can easily deduce that the last term of the left side of the inequality (21) is non-negative.

According to the definition of \( \omega^k \) in Algorithm 1, we know that \( \omega^k \) is monotone non-increasing, and hence, there exists \( \omega^* \) such that \( \lim_{k \to +\infty} \omega^k = \omega^* \).

Denote \( \tilde{L}_k = (\omega^*)^{-1} (\delta^k K^T K + \alpha \nabla^T \nabla) \). By the boundedness of \( \delta^k \) and \( \lim_{k \to +\infty} \omega^k = \omega^* \), we have that

\[
\lim_{k \to +\infty} \tilde{L}_k - L_k = 0.
\tag{22}
\]

Since \( \delta^{k+1} - \delta^k \leq \gamma D \), we know that \( \delta^k K^T K + \nabla^2 D_f(\zeta_k^k) \geq \delta^{k+1} K^T K \) according to the condition (18). Therefore, by the
definition of $\tilde{L}_k$ we further have
\[
\|u_{e}^{k+1}\|^2_{L_k} + \|u_{e}^{k+1}\|^2_{\nabla^2D_f(\zeta^k)} - \alpha \|\nabla u_{e}^{k+1}\|^2_{2} \geq 0.
\]
(23)

According to (22), we can easily conclude that there exists some $k_0$ such that
\[
\sum_{k=k_0}^{+\infty} \left( \|u_{e}^{k}\|^2_{\nabla^2D_f(\zeta^k)} + \|u_{e}^{k+1}\|^2_{L_k - \tilde{L}_k} - \|u_{e}^{k}\|^2_{L_k - \tilde{L}_k} \right) > 0.
\]

In the next, summing (21) from some $k_0$ to $+\infty$, utilizing (23), and removing the other non-negative terms, we obtain that
\[
\lim_{k \to +\infty} \|u_{e}^{k} - u^*\|^2_{\frac{1}{2} \nabla^2D_f(\zeta^k)} = 0.
\]
(24)

Due to $\|u_{e}^{k} - u^*\|^2_{\frac{1}{2} \nabla^2D_f(\zeta^k)} \geq \frac{1}{\gamma D} \|Ku_{e}^{k} - Ku^*\|^2_{2}$, we get
\[
\lim_{k \to +\infty} \|Ku_{e}^{k} - Ku^*\|^2_{2} = 0.
\]

Due to $\|K^T K\|_2 > 0$, we further have $\lim_{k \to +\infty} \|u_{e}^{k} - u^*\|^2_{2} = 0$, which implies that $\{u_{e}^{k}\}$ converges to a solution of the minimization problem (16).

References

1. Sarder, P., Nehorai, A.: Deconvolution method for 3-D fluorescence microscopy images. IEEE Signal Process. Mag. 23, 32–45 (2006)
2. Müller, J., Brune, C., Sawatzky, A., Kösters, T., Schäfers, K., Burger, M.: Reconstruction of short time PET scans using Bregman iterations. In: IEEE Nuclear Science Symposium and Medical Imaging Conference, pp. 2383–2385. (2011)
3. Bertero, M., Boccacci, P., Desiderà, G., Vicidomini, G.: Image deblurring with Poisson data: from cells to galaxies. Inverse Probl. 25, 123006 (2009)
4. Csiszár, I.: Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems. Ann. Stat. 19, 2032–2066 (1991)
5. Chambolle, A., Lions, P-L.: Image recovery via total variation minimization and related problems. Numer. Math. 76, 167–188 (1997)
6. Candès, E.J., Donoho, D.L.: New tight frames of curvelets and optimal representations of objects with piecewise-$C^2$ singularities. Commun. Pure Appl. Math. 57, 219–266 (2002)
7. Daubechies, I., Han, B., Ron, A., Shen, Z.: Framelets: MRA-based constructions of wavelet frames. Appl. Comput. Harmon. Anal. 14(1), 1–46 (2003)
8. Chen, D.Q., Cheng, L.Z.: Deconvolving Poissonian images by a novel hybrid variational model. J. Visual Commun. Image Rep. 22(7), 643–652 (2011)
9. Steidl, G., Weickert, J., Brox, T., Mrazek, P., Welk, M.: On the equivalence of soft wavelet shrinkage, total variation diffusion, total variation regularization, and SIDEs. SIAM J. Numer. Anal. 42(2), 686–713 (2004)
10. Cai, J.F., Dong, B., Osher, S., Shen, Z.W.: Image restoration: total variation, wavelet frames, and beyond. J. Am. Math. Soc. 25(4), 1033–1089 (2012)
11. Sawatzky, A., Brune, C., Wubbeling, F., Kösters, T., Schäfers, K., Burger, M.: Accurate EM–TV algorithm in PET with low SNR. In: IEEE Nuclear Science Symposium Conference Record, pp. 5133–5137. (2008)
12. Bonettini, S., Zanella, R., Zanni, L.: A scaled gradient projection method for constrained image deblurring. Inverse Probl. 25, 015002 (2009)
13. Bonettini, S., Ruggiero, V.: An alternating extragradient method for total variation based image restoration from Poisson data. Inverse Probl. 27, 095001 (2011)
14. Setzer, S., Steidl, G., Teuber, T.: Deblurring Poissonian images by split Bregman techniques. J. Visual Commun. Image Rep. 21(3), 193–199 (2010)
15. Afonso, M., Bioucas-Dias, J., Figueiredo, M.: Fast image recovery using variable splitting and constrained optimization. IEEE Trans. Image Process. 19(9), 2345–2356 (2010)
16. Chen, D.Q., Cheng, L.Z.: Spatially adapted regularization parameter selection based on the local discrepancy function for Poissonian image deblurring. Inverse Probl. 28, 015004 (2012)
17. Day, 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)
18. Woo, H., Yun, S.: Proximal linearized alternating direction method for multiplicative denoising. SIAM J. Sci. Comput. 35(2), 336–358 (2013)
19. Chen, D.Q., Zhou, Y.: Multiplicative denoising based on linearized alternating direction method using discrepancy function constraint. J. Sci. Comput. 60(3), 483–504 (2014)
20. Cai, J.F., Osher, S., Shen, Z.: Split Bregman methods and frame based image restoration. Multiscale Model. Simul. 8(2), 337–369 (2009)
21. Jeong, T., Woo, H., Yun, S.: Frame-based Poisson image restoration using a proximal linearized alternating direction method. Inverse Probl. 29(7), 075007 (2013)
22. Chen, D.Q.: Regularized generalized inverse accelerating linearized alternating minimization algorithm for frame-based Poissonian image deblurring. SIAM J. Imag. Sci. 7(2), 716–739 (2014)
23. Zhu, M., Chan, T.: An efficient primal-dual hybrid gradient algorithm for total variation image restoration. Technical Report 08-34, CAM UCLA. (2008)
24. Esser, E., Zhang, X., Chan, T.F.: A general framework for a class of first order primal-dual algorithms for convex optimization in imaging science. SIAM J. Imag. Sci. 3(4), 1015–1046 (2010)
25. Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. J. Math. Imaging Vis. 40(1), 120–145 (2011)
26. Landi, G., Piccolomini, E.L.: NPTool: a Matlab software for nonnegative image restoration with Newton projection methods. Numer. Algorithms 62(3), 487–504 (2013)
27. Combettes, P.L., Wajs, V.R.: Signal recovery by proximal forward–backward splitting. SIAM Multiscale Model. Simul. 4(4), 1168–1200 (2005)
28. Barzilai, J., Borwein, J.: Two point step size gradient methods. IMA J. Numer. Anal. 8, 141–148 (1988)
29. Avriel, M.: Nonlinear Programming: Analysis and Methods. Dover Publications Inc., Mineola (2003)
30. Bertsekas, D.P.: Convex Optimization Theory. Athena Scientific, Belmont (2009)