AUGMENTED LAGRANGIAN METHOD FOR TOTAL VARIATION RESTORATION WITH NON-QUADRATIC FIDELITY

CHUNLIN WU *, JUYONG ZHANG †, AND XUE-CHENG TAI *‡

Abstract. Recently augmented Lagrangian method has been successfully applied to image restoration with L^2 fidelity. In this paper we extend the method to total variation (TV) restoration models with non-quadratic fidelities. We will first introduce the method and present the iterative algorithm for TV restoration with a quite general fidelity. In each iteration, three sub-problems need to be solved, two of which can be very efficiently solved via FFT implementation or closed form solution. In general the third sub-problem need iterative solvers. We then apply our method to TV restoration with L^1 and Kullback-Leibler (KL) fidelities, two common and important data terms for deblurring images corrupted by impulsive noise and Poisson noise, respectively. For these typical fidelities, we show that the third sub-problem also has closed form solution and thus can be efficiently solved. In addition, convergence analysis of these algorithms are given, which cannot be obtained by previous analysis techniques.

Key words. augmented Lagrangian method, total variation, impulsive noise, Poisson noise, TV-L^1, TV-KL, convergence

AMS subject classifications. 80M30, 80M50, 68U10

1. Introduction.

Total variation regularization was first introduced in [48]. It has been demonstrated very successful in image restoration and extensively generalized [10, 15, 65, 38, 39, 30, 50, 49, 3, 5, 16]. The essential reason of the achievement is that in most images the gradient is sparse and TV catches this property, like the basis pursuit problem [18] in compressive sensing [8, 22]. Although the computation is difficult due to the nonlinearity and non-differentiability, a lot of effort has been contributed to design fast solvers [14, 9, 11, 67, 68, 56, 57, 59, 32, 63, 28, 54, 58].

However, all of these consider TV minimization with squared L^2 fidelity term (TV-L^2 model), which is particularly suitable for recovering images corrupted by Gaussian noise. In many important data, the noise may not obey Gaussian distribution and thus the data fidelity term is non-quadratic. Two typical and important examples are impulsive noise [4] and Poisson noise [36, 6].

Impulsive noise is often generated by malfunctioning pixels in camera sensors, faulty memory locations in hardware, or erroneous transmission [4]. It has two common types, salt-and-pepper noise and random-valued noise. Salt-and-pepper (or random-valued) noise corrupts a portion of the image pixels with minimal or maximal intensities (or random-valued intensities) while keeping other pixels unaffected. To remove this kind of noise is quite difficult, since the corrupted pixels are randomly distributed in the image and the intensities at corrupted pixels are usually distinguishable from those of their neighbors. By applying TV regularization and Bayesian statistic, one obtains a variational approach which minimizes a TV-L^1 functional. Compared with TV-L^2 model, TV-L^1 uses a non-smooth fidelity which has great advantages in impulsive noise removal [42, 43]. It is shown that the L^1 fidelity can fit uncorrupted pixels exactly and regularize the corrupted pixels perfectly. This model also provides many other useful properties proved recently in [17, 61, 62]. In addition,

*Division of Mathematical Sciences, School of Physical & Mathematical Sciences, Nanyang Technological University, Singapore.

†Division of Computer Communications, School of Computer Engineering, Nanyang Technological University, Singapore.

‡Department of Mathematics, University of Bergen, Norway.

it has been noticed in [1, 40, 37] that TV-L^1 model (with no blur kernel) connects closely to classical median type filters [19, 24, 33, 45, 41]. It can also be applied to the recent particularly effective two-phase method [12]. However, the TV-L^1 model is hard to compute due to the nonlinearity and non-differentiability of both the TV term and the data fidelity. Some existing numerical methods include gradient descent method [17], LAD method [26], the splitting-and-penalty based method [60], and the primal-dual method [20] based on semi-smooth Newton algorithm [31], as well as alternating direction methods [25]. We should mention that in [25], the authors treat the operators in a compact way so that penalty parameters for different auxiliary variables are the same. When we tested our algorithms, we found it's more efficient to use different parameters for different auxiliary variables.

Poisson noise is a very common signal dependent noise, and is contained in signals in various applications such as radiography, fluorescence microscopy, positron-emission-tomography (PET), optical nanoscopy and astronomical imaging applications [36, 6]. To recover a blurry image corrupted by Poisson noise is difficult. Some classical methods based on some special assumptions can be found in [2, 34, 35, 55], which were designed for denoising only. Recently, variational methods based on TV regularization have been applied to this problem. According to the characteristic of Poisson distribution, people derived a TV regularization model with the so called Kullback-Leibler divergence as fidelity term [36, 6]. In this paper we call this model as TV-KL model. It has been shown that TV-KL model behaves much stable and robust than the standard expectation maximization (EM) reconstruction (where no TV regularization is applied) [52], and much more effective than TV-L^2 in the case of Poisson noise removal [36]. Some existing method for the TV-KL model are gradient descent [36, 44], multilevel method [13], the scaled gradient projection method [66], and EM-TV alternative minimization [6].

Therefore, in those image restoration problems with non-Gaussian noise we need to minimize functionals with TV regularization and non-quadratic fidelities. To design fast solvers for these restoration models is still highly desired and is much harder than that for TV-L^2, since the first order variations of these fidelities are no longer linear. In this paper, we extend augmented Lagrangian method [29, 46, 47] for TV-L^2 restoration [54, 58] to solve the problem. In particular, we will first give the algorithms for TV restoration with a general fidelity term and then apply these algorithms to recover blurry images corrupted by impulsive noise or Poisson noise. We will show that for these two special cases, augmented Lagrangian method is extremely efficient since all the sub-problems have closed form solutions. Besides, convergence analysis of these algorithms will be provided, which cannot be obtained by using previous techniques.

The paper is organized as follows. In the next section, we give some notation. In Section 3, we present TV restoration model with general fidelity. Augmented Lagrangian method will be given in Section 4 with convergence analysis. In Section 5, we apply our algorithms for deblurring images corrupted by impulsive noise or Poisson noise. The paper is concluded in Section 6.

2. Notation. Without the loss of generality, we represent a gray image as an $N \times N$ matrix. The Euclidean space $\mathbb{R}^{N\times N}$ is denoted as V. The discrete gradient operator is a mapping $\nabla : V \to Q$, where $Q = V \times V$. For $u \in V$, ∇u is given by

$$(\nabla u)_{i,j} = ((\mathring{D}_x^+ u)_{i,j}, (\mathring{D}_y^+ u)_{i,j}),$$

with

$$
\begin{array}{l}
(\mathring{D}_x^+ u)_{i,j} = \begin{cases} u_{i,j+1} - u_{i,j}, & 1 \le j \le N-1 \\ u_{i,1} - u_{i,N}, & j = N \end{cases} \\
(\mathring{D}_y^+ u)_{i,j} = \begin{cases} u_{i+1,j} - u_{i,j}, & 1 \le i \le N-1 \\ u_{1,j} - u_{N,j}, & i = N \end{cases}
\end{array},
$$

where $i, j = 1, \ldots, N$. Here we use $\mathring{D}_x^+$ and $\mathring{D}_y^+$ to denote forward difference operators with periodic boundary condition (u is periodically extended). Consequently FFT can be adopted in our algorithm.

We denote the usual inner product and Euclidean norm (L^2 norm) of V as $(\cdot,\cdot)_V$ and $\|\cdot\|_V$, respectively. We also equip the space Q with inner product $(\cdot,\cdot)_Q$ and norm $\|\cdot\|_Q$, which are defined as follows. For $p = (p^1, p^2) \in Q$ and $q = (q^1, q^2) \in Q$,

$$(p, q)_Q = (p^1, q^1)_V + (p^2, q^2)_V,$$

and

$$\|p\|_Q = \sqrt{(p, p)_Q}.$$

In addition, we mention that, at each pixel (i, j),

$$|p_{i,j}| = |(p^1_{i,j}, p^2_{i,j})| = \sqrt{(p^1_{i,j})^2 + (p^2_{i,j})^2},$$

the usual Euclidean norm in $\mathbb{R}^2$. From the subscript i, j, one may regard $|p_{i,j}|$ as pixel-by-pixel norm of p. In the case without confusion, we will omit the subscripts V and Q and just use $(\cdot,\cdot)$ and $\|\cdot\|$ to denote the usual inner products and L^2 norms. In this paper, we also use $\|v\|_{L^1}$ to denote the L^1 norm of $v \in V$.

Using the inner products of V and Q, we can find the adjoint operator of $-\nabla$, i.e., the discrete divergence operator div : $Q \to V$. Given $p = (p^1, p^2) \in Q$, we have

$$(\mathrm{div} p)_{i,j} = p^1_{i,j} - p^1_{i,j-1} + p^2_{i,j} - p^2_{i-1,j} = (\mathring{D}_x^- p^1)_{i,j} + (\mathring{D}_y^- p^2)_{i,j},$$

where $\mathring{D}_x^-$ and $\mathring{D}_y^-$ are backward difference operators with periodic boundary conditions $p^1_{i,0} = p^1_{i,N}$ and $p^2_{0,j} = p^2_{N,j}$.

3. The total variation (TV) image restoration. Assume $f \in V$ is an observed image containing both blur and noise. The degradation procedure is in general modelled as follows

$$u \xrightarrow{\text{blur}} Ku \xrightarrow{\text{noise}} f, \tag{3.1}$$

where $u \in V$ is the true image and $K : V \to V$ is a blur operator. Here we do not specify the noise model. It can be Gaussian, impulsive, Poisson and even others. Different noise models give different degradation f. Image restoration aims at recovering u from f. Since the problem is usually ill-posed, we cannot directly solve u from (3.1). Regularization on the solution should be considered. One of the most basic and successful image restoration model is based on the total variation (TV) regularization, which reads

$$\min_{u \in V} \{E(u) = R(\nabla u) + F(Ku)\}, \tag{3.2}$$

where

$$R(\nabla u) = \mathrm{TV}(u) = \sum_{1 \le i,j \le N} |(\nabla u)_{i,j}|, \tag{3.3}$$

is the total variation of u [48], and $F(Ku)$ is a fidelity term. Note here $R(\cdot)$ is regarded as a functional of ∇u.

In this paper we only consider the case where the blur operator K is given. Since the blur is essentially averaging, it is reasonable to assume

- Assumption 1. $\mathrm{null}(\nabla) \cap \mathrm{null}(K) = \{0\}$,

where $\mathrm{null}(\cdot)$ is the null space of $\cdot$.

The form of the fidelity term depends on the statistic of the noise model. Some typical noise models and their corresponding fidelity terms are as follows:

1. Gaussian noise:

$$F(Ku) = \frac{\alpha}{2}\|Ku - f\|^2,$$

2. Impulsive noise:

$$F(Ku) = \alpha \|Ku - f\|_{L^1},$$

3. Poisson noise (assuming $f_{i,j} > 0, \forall i, j$, as in [36]):

$$F(Ku) = \begin{cases} \alpha \sum_{1 \le i,j \le N} ((Ku)_{i,j} - f_{i,j} \log (Ku)_{i,j}), & u \in V, (Ku)_{i,j} > 0 \\ +\infty, & \text{otherwise} \end{cases},$$

where $\alpha > 0$ is a parameter. Note for Poisson noise, we extend the definition of the fidelity to the whole space V, compared to [36] (where $K = I$) and [6]. To define the fidelity over the whole space is convenient for analysis. We make the following assumptions for the fidelity term:

- Assumption 2. $\mathrm{dom}(R \circ \nabla) \cap \mathrm{dom}(F \circ K) \neq \varnothing$;
- Assumption 3. $F(z)$ is convex, proper, and coercive;
- Assumption 4. $F(z)$ is continuous over $\mathrm{dom}(F)$,

where $\mathrm{dom}(F) = \{z \in V : -\infty < F(z) < +\infty\}$ is the domain of F, with similar definitions for $\mathrm{dom}(R \circ \nabla)$ and $\mathrm{dom}(F \circ K)$. These assumptions are relatively quite general and many fidelities such as those listed above meet all of them.

Under the Assumptions 1, 2, 3 and 4, it can be verified that the functional $E(u)$ in (3.2) is convex, proper, coercive, and lower semi continuous. According to the generalized Weierstrass theorem and Fermat's rule [23, 27], we have the following result.

Theorem 3.1. *The problem* (3.2) *has at least one solution u, which satisfies*

$$0 \in K^* \partial F(Ku) - \mathrm{div} \partial R(\nabla u), \tag{3.4}$$

where $\partial F(Ku)$ and $\partial R(\nabla u)$ are the sub-differentials [23] of F at Ku and R at ∇u, respectively. Moreover, if $F \circ K(u)$ is strictly convex, the minimizer is unique.

The total variation minimization with a quadratic fidelity has been widely studied. Many efficient algorithms have been proposed to solve the problem [14, 9, 11, 67, 68, 56, 57, 59, 32, 63, 28, 54, 58]. We here extend our recent work [54, 58] to total variation restoration with non-quadratic fidelity terms satisfying the above assumptions.

4. Augmented Lagrangian method for total variation restoration. In this section we present to use augmented Lagrangian method for total variation restoration with a non-quadratic fidelity term which satisfies our (relatively quite general) assumptions. Since $F(Ku)$ is non-quadratic, its first order variation is not linear. Compared with the augmented Lagrangian method for TV-L^2 model [54, 58], we need one more auxiliary variable to eliminate the nonlinearity for u as done in [60] for TV-L^1 restoration.

In particular, we introduce two new variables $p \in Q$ and $z \in V$ and reformulate the problem to be the following constrained optimization problem

$$\begin{array}{ll} \min\limits_{u\in V, p\in Q, z\in V} & \{G(p,z) = R(p) + F(z)\} \\ \text{s.t.} & p = \nabla u, z = Ku \end{array}. \tag{4.1}$$

To solve (4.1), we define the following augmented Lagrangian functional

$$\begin{aligned} &\mathscr{L}(u,p,z;\lambda_p,\lambda_z) \\ &= R(p) + F(z) + (\lambda_p, p - \nabla u) + (\lambda_z, z - Ku) + \tfrac{r_p}{2}\|p - \nabla u\|^2 + \tfrac{r_z}{2}\|z - Ku\|^2, \end{aligned} \tag{4.2}$$

with Lagrange multipliers $\lambda_p \in Q, \lambda_z \in V$ and positive constants r_p, r_z, and then consider the following saddle-point problem:

$$\begin{array}{ll} \text{Find} & (u^*,p^*,z^*;\lambda_p^*,\lambda_z^*) \in V\times Q\times V\times Q\times V, \\ \text{s.t.} & \begin{array}{l}\mathscr{L}(u^*,p^*,z^*;\lambda_p,\lambda_z) \le \mathscr{L}(u^*,p^*,z^*;\lambda_p^*,\lambda_z^*) \le \mathscr{L}(u,p,z;\lambda_p^*,\lambda_z^*), \\ \forall (u,p,z;\lambda_p,\lambda_z) \in V\times Q\times V\times Q\times V. \end{array} \end{array} \tag{4.3}$$

Note that, differently from [25], here it is no need for $r_p = r_z$. According to our test, much more efficiency can be achieved by using different penalty parameters. As one will see, the convergence analysis when $r_p \neq r_z$ is more difficult than the case $r_p = r_z$.

Similarly with [58], we can prove the following result.

Theorem 4.1. *$u^* \in V$ is a solution of* (3.2) *if and only if there exist* $(p^*, z^*) \in Q\times V$ *and* $(\lambda_p^*, \lambda_z^*) \in Q\times V$ *such that* $(u^*,p^*,z^*;\lambda_p^*,\lambda_z^*)$ *is a solution of* (4.3).

Proof We just provide a sketch since the idea is similar with that in [58].

Suppose $(u^*,p^*,z^*;\lambda_p^*,\lambda_z^*)$ is a solution of (4.3). From the first inequality in (4.3), we have

$$\left\{\begin{array}{l} p^* - \nabla u^* = 0, \\ z^* - Ku^* = 0. \end{array}\right. \tag{4.4}$$

The above relation, together with the second inequality in (4.3), indicates that u^* is a solution of (3.2).

Conversely, we assume that $u^* \in V$ is a solution of (3.2). We take $p^* = \nabla u^* \in Q$ and $z^* = Ku^* \in V$. From (3.4), there exist λ_p^* and λ_z^* such that $-\lambda_p^* \in \partial R(\nabla u^*)$ and $-\lambda_z^* \in \partial F(Ku^*)$ with $-K^*\lambda_z^* + \mathrm{div}\lambda_p^* = 0$. We can verify that $(u^*,p^*,z^*;\lambda_p^*,\lambda_z^*)$ is a saddle-point of $\mathscr{L}$, which completes the proof. ■

Theorems 3.1 and 4.1 show that the saddle-point problem (4.3) has at least one solution and this solution will provide a solution of the original problem (3.2). In the following we present an iterative algorithm to solve the saddle-point problem and address three sub-problems raised up in each iteration.

Algorithm 4.1 Augmented Lagrangian method for TV restoration with non-quadratic fidelity

1. *Initialization*: $\lambda_p^0 = 0, \lambda_z^0 = 0$;
2. *For k=0,1,2,...*:
   (a) compute (u^k, p^k, z^k) as an (approximate) minimizer of the augmented Lagrangian functional with the Lagrange multipliers λ_p^k, λ_z^k, *i.e.*,

$$(u^k, p^k, z^k) \approx \arg \min_{(u,p,z)\in V\times Q\times V} \mathscr{L}(u, p, z; \lambda_p^k, \lambda_z^k), \tag{4.5}$$

   where $\mathscr{L}(u, p, z; \lambda_p^k, \lambda_z^k)$ is as in (4.2);
   (b) update

$$\begin{array}{l} \lambda_p^{k+1} = \lambda_p^k + r_p(p^k - \nabla u^k) \\ \lambda_z^{k+1} = \lambda_z^k + r_z(z^k - Ku^k) \end{array}. \tag{4.6}$$

4.1. An iterative algorithm for the saddle-point problem. In augmented Lagrangian method, we use an iterative algorithm to solve the saddle-point problem; see Algorithm 4.1.

Since the variables u, p, z in $\mathscr{L}(u, p, z; \lambda_p^k, \lambda_z^k)$ are coupled together in the minimization problem (4.5), it's difficult to solve them simultaneously. Therefore we separate the problem to be three sub-problems and apply an alternative minimization. The three sub-problems are as follows:

- $u-$sub problem: Given p, z,

$$\min_{u\in V}\{(\lambda_p^k, -\nabla u) + (\lambda_z^k, -Ku) + \frac{r_p}{2}\|p - \nabla u\|^2 + \frac{r_z}{2}\|z - Ku\|^2\}. \tag{4.7}$$

- $p-$sub problem: Given u, z,

$$\min_{p\in Q}\{R(p) + (\lambda_p^k, p) + \frac{r_p}{2}\|p - \nabla u\|^2\}. \tag{4.8}$$

- $z-$sub problem: Given u, p,

$$\min_{z\in V}\{F(z) + (\lambda_z^k, z) + \frac{r_z}{2}\|z - Ku\|^2\}. \tag{4.9}$$

Note here we omit the constant terms in the objective functionals in (4.7), (4.8) and (4.9).

In the following we show how to efficiently solve these sub-problems and then present an alternative minimization algorithm to solve (4.5).

4.1.1. Solving the $u-$sub problem (4.7)**.** (4.7) is a quadratic optimization problem, whose optimality condition reads

$$\mathrm{div}\lambda_p^k - K^*\lambda_z^k + r_p\mathrm{div}(p - \nabla u) - r_zK^*(z - Ku) = 0,$$

by considering the periodic boundary conditions. Following [56, 57, 59, 60, 54, 58], we use Fourier transform (and hence FFT implementation) to solve the above linear equation. Denoting $\mathcal{F}(u)$ as the Fourier transform of u, we have

$$\begin{array}{l} (r_z\mathcal{F}(K^*)\mathcal{F}(K) - r_p\mathcal{F}(\triangle))\mathcal{F}(u) \\ = \mathcal{F}(K^*)(\mathcal{F}(\lambda_z^k) + r_z\mathcal{F}(z)) - \mathcal{F}(\mathring{D}_x^-)(\mathcal{F}((\lambda_p^1)^k) + r_p\mathcal{F}(p^1)) - \mathcal{F}(\mathring{D}_y^-)(\mathcal{F}((\lambda_p^2)^k) + r_p\mathcal{F}(p^2)) \end{array}, \tag{4.10}$$

where $\lambda_p^k = ((\lambda_p^1)^k, (\lambda_p^2)^k)$ and $p = (p^1, p^2)$; and Fourier transforms of operators such as $K, \mathring{D}_x^-, \mathring{D}_y^-, \triangle = \mathring{D}_x^- \mathring{D}_x^+ + \mathring{D}_y^- \mathring{D}_y^+$ are regarded as the transforms of their corresponding convolution kernels.

4.1.2. Solving the $p-$sub problem (4.8)**.** Similarly with [7, 56, 57, 54, 58], (4.8) has the following closed form solution

$$p_{i,j} = \max(0, 1 - \frac{1}{r_p|\mathbf{w}_{i,j}|})\mathbf{w}_{i,j}, \tag{4.11}$$

where

$$\mathbf{w} = \nabla u - \frac{\lambda_p^k}{r_p} \in Q. \tag{4.12}$$

Here we would like to provide a geometric interpretation of the formulae (4.11). According to the definition of $R(p)$ and $\|\cdot\|_Q$, we rewrite the problem (4.8) as

$$\min_{p\in Q}\{\sum_{1\leq i,j\leq N} |p_{i,j}| + \frac{r_p}{2}\sum_{1\leq i,j\leq N} |p_{i,j} - (\nabla u - \frac{\lambda_p^k}{r_p})_{i,j}|^2 + \text{Constant}\}.$$

As one can see, the above problem is decomposable and at each pixel (i,j), the problem takes the form as follows

$$\min_{q\in\mathbb{R}^2}\{|q| + \frac{r_p}{2}|q - w|^2\}, \tag{4.13}$$

where $w \in \mathbb{R}^2$; see Fig. 4.1.

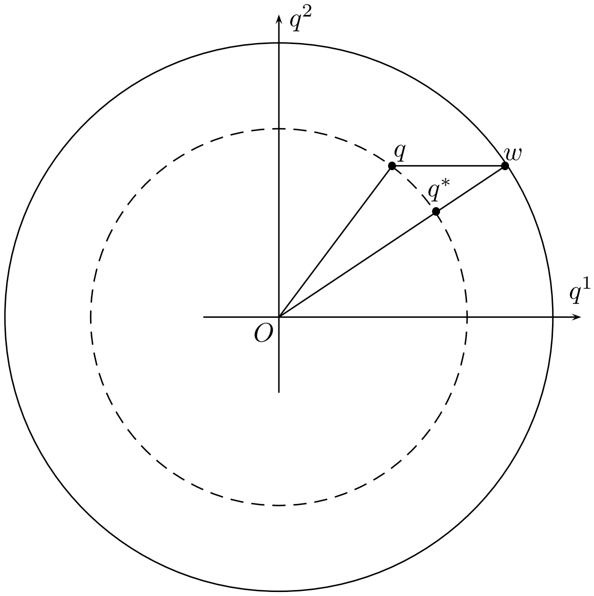


FIG. 4.1. *A geometric interpretation of the formulae* (4.11)

First of all, it can be verified (and imagined) that the potential minimizer should locate inside of the solid circle. By constructing symmetric points, we can further demonstrate that the potential minimizer should locate in the same quadrant as w. Therefore in the example in Fig. 4.1, we only need to consider those points located inside of the solid circle and in the first quadrant, e.g., q. For a such point q, we draw a dashed circle with O as the center and $|q|$ as the radius. Assume this circle intersects the line segment Ow at q^*. By the triangle inequality of the Euclidean norm $|\cdot|$ in $\mathbb{R}^2$, we have

$$|q| + |q - w| \geq |w| = |q^*| + |q^* - w|.$$

Since $|q| = |q^*|$, we obtain

$$|q - w| \geq |q^* - w|,$$

indicating

$$|q| + \frac{r_p}{2}|q - w|^2 \geq |q^*| + \frac{r_p}{2}|q^* - w|^2.$$

This means the solution of the problem (4.13) will locate on the line segment Ow. Denoting $q = \beta w$ with $0 \leq \beta \leq 1$, we hence simplify (4.13) to be the following 1-dimensional problem

$$\min_{0\leq\beta\leq1} \{\beta|w| + \frac{r_p}{2}(\beta - 1)^2|w|^2\}. \tag{4.14}$$

(4.14) can be solved exactly, with a closed form solution as

$$\beta^* = \max(0, 1 - \frac{1}{r_p|w|}).$$

The solution of (4.13) follows immediately.

We further give two comments on this geometric interpretation. First, this observation (to solve (4.13)) can be extended to high (> 2) dimensional problems as we did in [58] for vectorial and high order TV models. Second, the method can also be applied to problems with general regularization terms, say, a general $R(q)$ replacing $|q|$ in (4.13), as long as the regularizer $R(q)$ depends only on $|q|$.

4.1.3. Solving the $z-$sub problem (4.9)**.** For a general fidelity F, it is no reason to find a closed form solution for (4.9). Fortunately, the objective functional in (4.9) is strictly convex, proper, coercive and lower semi continuous. Therefore, (4.9) has a unique solution and can be obtained by various numerical optimization methods.

There is one fact we need mention. For some special and typical (non-quadratic) fidelities, we still have closed form solutions; see Section 5. Our method is therefore particularly efficient for these typical and important fidelities.

After knowing how to solve (4.7), (4.8) and (4.9), we now present the following alternative minimization procedure to solve (4.5). It is with Gauss-Seidel flavor.

Here L can be chosen using some convergence test techniques. In this paper, we simply set $L = 1$. In our experiments we found that with larger L (> 1) the algorithm wastes the accuracy of the inner iteration and does not speed up dramatically the convergence of the overall algorithm (Algorithm 4.1 with Algorithm 4.2 as a sub algorithm). This has also been observed in [28], for the split Bregman method (which is equivalent to augmented Lagrangian method). To simply set $L = 1$ also benefits the efficiency of the algorithm, since we do not need to compute those residuals of the optimality conditions.

Algorithm 4.2 Augmented Lagrangian method for TV restoration with non-quadratic fidelity – solve the minimization problem (4.5)

- *Initialization*: $u^{k,0} = u^{k-1}, p^{k,0} = p^{k-1}, z^{k,0} = z^{k-1}$;
- *For* $l = 0, 1, 2, ..., L-1$:
  - compute $u^{k,l+1}$ from (4.10) for $p = p^{k,l}, z = z^{k,l}$;
  - compute $p^{k,l+1}$ from (4.11) for $u = u^{k,l+1}$;
  - compute $z^{k,l+1}$ by solving (4.9) for $u = u^{k,l+1}$;
- $u^k = u^{k,L}, p^k = p^{k,L}, z^k = z^{k,L}$.

4.2. Convergence analysis. In this subsection we give some convergence results of the augmented Lagrangian method applied to total variation restoration with non-quadratic fidelity. We focus on analyzing Algorithm 4.1. In particular, we will prove the convergence of Algorithm 4.1 in two limiting cases where the minimization problem (4.5) is computed by Algorithm 4.2 with full accuracy ($L \to \infty$, by assuming Algorithm 4.2 is convergent) and rough accuracy ($L = 1$), respectively. The convergence of Algorithm 4.2 depends on the fidelity F and can be checked when F is given; see Section 5 for two important fidelities.

It should be pointed out that the previous analysis techniques in [27, 58] cannot be applied to our case. Because in general $r_p \neq r_z$, the monotonically decreasing sequences constructed in [27, 58] do not hold. In the following proofs we construct two new monotonically decreasing sequences and use these sequences to derive the results.

THEOREM 4.2. *Assume $(u^*, p^*, z^*; \lambda_p^*, \lambda_z^*)$ is a saddle-point of $\mathscr{L}(u, p, z; \lambda_p, \lambda_z)$. Suppose that the minimization problem (4.5) is exactly solved in each iteration, i.e., $L \to \infty$ in Algorithm 4.2. Then the sequence $(u^k, p^k, z^k; \lambda_p^k, \lambda_z^k)$ generated by Algorithm 4.1 satisfies*

$$\begin{cases} \lim\limits_{k\to\infty} (R(p^k) + F(z^k)) = R(p^*) + F(z^*) = E(u^*), \\ \lim\limits_{k\to\infty} \|p^k - \nabla u^k\| = 0, \\ \lim\limits_{k\to\infty} \|z^k - Ku^k\| = 0. \end{cases} \tag{4.15}$$

Moreover, (4.15) *indicates that u^k is a minimizing sequence of $E(\cdot)$. If the minimizer of $E(\cdot)$ is unique, then $u^k \to u^*$.*

Proof Let us define $\overline{u}^k, \overline{p}^k, \overline{z}^k, \overline{\lambda_p}^k, \overline{\lambda_z}^k$, as

$$\overline{u}^k = u^k - u^*, \quad \overline{p}^k = p^k - p^*, \quad \overline{z}^k = z^k - z^*, \quad \overline{\lambda_p}^k = \lambda_p^k - \lambda_p^*, \quad \overline{\lambda_z}^k = \lambda_z^k - \lambda_z^*.$$

Since $(u^*, p^*, z^*; \lambda_p^*, \lambda_z^*)$ is a saddle-point of $\mathscr{L}(u, p, z; \lambda_p, \lambda_z)$, we have

$$\begin{aligned} &\mathscr{L}(u^*, p^*, z^*; \lambda_p, \lambda_z) \le \mathscr{L}(u^*, p^*, z^*; \lambda_p^*, \lambda_z^*) \le \mathscr{L}(u, p, z; \lambda_p^*, \lambda_z^*), \\ &\forall (u, p, z; \lambda_p, \lambda_z) \in V \times Q \times V \times Q \times V. \end{aligned} \tag{4.16}$$

From the first inequality of (4.16), we have

$$\begin{cases} p^* = \nabla u^*, \\ z^* = Ku^*. \end{cases}$$

This relationship, together with (4.6), indicates

$$\begin{cases} \overline{\lambda_p}^{k+1} = \overline{\lambda_p}^k + r_p(\overline{p}^k - \nabla \overline{u}^k), \\ \overline{\lambda_z}^{k+1} = \overline{\lambda_z}^k + r_z(\overline{z}^k - K\overline{u}^k), \end{cases}$$

which is equivalent to

$$\begin{cases} \sqrt{r_z}\overline{\lambda_p}^{k+1} = \sqrt{r_z}\overline{\lambda_p}^{k} + r_p\sqrt{r_z}(\overline{p}^k - \nabla \overline{u}^k), \\ \sqrt{r_p}\overline{\lambda_z}^{k+1} = \sqrt{r_p}\overline{\lambda_z}^{k} + r_z\sqrt{r_p}(\overline{z}^k - K \overline{u}^k). \end{cases} \tag{4.17}$$

The observation (4.17) is a key formula in our proof, and helps to construct a useful monotonically decreasing sequence, which is different from that in [27, 58].

It then follows that

$$\begin{aligned} &(r_z\|\overline{\lambda_p}^k\|^2 + r_p\|\overline{\lambda_z}^k\|^2) - (r_z\|\overline{\lambda_p}^{k+1}\|^2 + r_p\|\overline{\lambda_z}^{k+1}\|^2) \\ &= -2r_pr_z(\overline{\lambda_p}^k, \overline{p}^k - \nabla\overline{u}^k) - r_p^2r_z\|\overline{p}^k - \nabla\overline{u}^k\|^2 - 2r_pr_z(\overline{\lambda_z}^k, \overline{z}^k - K\overline{u}^k) - r_z^2r_p\|\overline{z}^k - K\overline{u}^k\|^2. \end{aligned} \tag{4.18}$$

In the following we show that the right hand side of (4.18) is not less than 0 and thus the sequence $\{(r_z\|\overline{\lambda_p}^k\|^2 + r_p\|\overline{\lambda_z}^k\|^2)\}$ is monotonically decreasing.

From the second inequality of (4.16), (u^*, p^*, z^*) is characterized by

$$\begin{aligned} &(\mathrm{div}\lambda_p^*, u - u^*) + r_p(\mathrm{div}(p^* - \nabla u^*), u - u^*) \\ &\quad +(\lambda_z^*, -K(u - u^*)) + r_z(z^* - Ku^*, -K(u - u^*)) \geq 0, \forall u \in V, \end{aligned} \tag{4.19}$$

$$R(p) - R(p^*) + (\lambda_p^*, p - p^*) + r_p(p^* - \nabla u^*, p - p^*) \geq 0, \forall p \in Q, \tag{4.20}$$

$$F(z) - F(z^*) + (\lambda_z^*, z - z^*) + r_z(z^* - Ku^*, z - z^*) \geq 0, \forall z \in V. \tag{4.21}$$

Similarly, (u^k, p^k, z^k) is characterized by

$$\begin{aligned} &(\mathrm{div}\lambda_p^k, u - u^k) + r_p(\mathrm{div}(p^k - \nabla u^k), u - u^k) \\ &\quad +(\lambda_z^k, -K(u - u^k)) + r_z(z^k - Ku^k, -K(u - u^k)) \geq 0, \forall u \in V, \end{aligned} \tag{4.22}$$

$$R(p) - R(p^k) + (\lambda_p^k, p - p^k) + r_p(p^k - \nabla u^k, p - p^k) \geq 0, \forall p \in Q, \tag{4.23}$$

$$F(z) - F(z^k) + (\lambda_z^k, z - z^k) + r_z(z^k - Ku^k, z - z^k) \geq 0, \forall z \in V, \tag{4.24}$$

since (u^k, p^k, z^k) is the solution of (4.5). Taking $u = u^k$ in (4.19), $u = u^*$ in (4.22), $p = p^k$ in (4.20), $p = p^*$ in (4.23), $z = z^k$ in (4.21), and $z = z^*$ in (4.24), respectively, we obtain, by addition

$$-(\overline{\lambda_p}^k, \overline{p}^k - \nabla\overline{u}^k) - (\overline{\lambda_z}^k, \overline{z}^k - K\overline{u}^k) \geq r_p\|\overline{p}^k - \nabla\overline{u}^k\|^2 + r_z\|\overline{z}^k - K\overline{u}^k\|^2, \tag{4.25}$$

which is equivalent to

$$-r_pr_z(\overline{\lambda_p}^k, \overline{p}^k - \nabla\overline{u}^k) - r_pr_z(\overline{\lambda_z}^k, \overline{z}^k - K\overline{u}^k) \geq r_p^2r_z\|\overline{p}^k - \nabla\overline{u}^k\|^2 + r_pr_z^2\|\overline{z}^k - K\overline{u}^k\|^2. \tag{4.26}$$

From (4.18) and (4.26), we have

$$(r_z\|\overline{\lambda_p}^k\|^2 + r_p\|\overline{\lambda_z}^k\|^2) - (r_z\|\overline{\lambda_p}^{k+1}\|^2 + r_p\|\overline{\lambda_z}^{k+1}\|^2) \geq r_p^2r_z\|\overline{p}^k - \nabla\overline{u}^k\|^2 + r_pr_z^2\|\overline{z}^k - K\overline{u}^k\|^2, \tag{4.27}$$

which indicates

$$\left\{\begin{array}{c}\{\lambda_p^k : \forall k\} \text{ and } \{\lambda_z^k : \forall k\} \quad \text{are bounded,} \\ \lim\limits_{k\to\infty} \|p^k - \nabla u^k\| = 0, \\ \lim\limits_{k\to\infty} \|z^k - Ku^k\| = 0.\end{array}\right. \tag{4.28}$$

On the other hand, the second inequality of (4.16) implies

$$\begin{array}{c}R(p^*) + F(z^*) \leq R(p^k) + F(z^k) + (\lambda_p^*, p^k - \nabla u^k) + (\lambda_z^*, z^k - Ku^k) \\ +\frac{r_p}{2}\|p^k - \nabla u^k\|^2 + \frac{r_z}{2}\|z^k - Ku^k\|^2.\end{array} \tag{4.29}$$

If we take $u = u^*$ in (4.22), $p = p^*$ in (4.23), and $z = z^*$ in (4.24), we have, by addition,

$$\begin{array}{c}R(p^*) + F(z^*) \geq R(p^k) + F(z^k) + (\lambda_p^k, p^k - \nabla u^k) + (\lambda_z^k, z^k - Ku^k) \\ +r_p\|p^k - \nabla u^k\|^2 + r_z\|z^k - Ku^k\|^2.\end{array} \tag{4.30}$$

Using (4.28), we have

$$\liminf(R(p^k) + F(z^k)) \geq R(p^*) + F(z^*) \geq \limsup(R(p^k) + F(z^k)), \tag{4.31}$$

by taking lim inf in (4.29) and lim sup in (4.30). Hence we complete the proof of (4.15).

Since $R(\cdot)$ and $F(\cdot)$ are both continuous over their domains, (4.15) implies clearly that u^k is a minimizing sequence of $E(\cdot)$. If the minimizer of $E(\cdot)$ is unique, then $u^k \to u^*$. ■

THEOREM 4.3. *Assume* $(u^*, p^*, z^*; \lambda_p^*, \lambda_z^*)$ *is a saddle-point of* $\mathscr{L}(u, p, z; \lambda_p, \lambda_z)$. *Suppose that the minimization problem* (4.5) *is roughly solved in each iteration, i.e.,* $L = 1$ *in Algorithm 4.2. Then the sequence* $(u^k, p^k, z^k; \lambda_p^k, \lambda_z^k)$ *generated by Algorithm 4.1 satisfies*

$$\left\{\begin{array}{l}\lim\limits_{k\to\infty} (R(p^k) + F(z^k)) = R(p^*) + F(z^*) = E(u^*), \\ \lim\limits_{k\to\infty} \|p^k - \nabla u^k\| = 0, \\ \lim\limits_{k\to\infty} \|z^k - Ku^k\| = 0.\end{array}\right. \tag{4.32}$$

Moreover, (4.32) *indicates that* u^k *is a minimizing sequence of* $E(\cdot)$. *If the minimizer of* $E(\cdot)$ *is unique, then* $u^k \to u^*$.

Proof Again we define the following errors $\overline{u}^k, \overline{p}^k, \overline{z}^k, \overline{\lambda_p}^k, \overline{\lambda_z}^k$, as

$$\overline{u}^k = u^k - u^*, \quad \overline{p}^k = p^k - p^*, \quad \overline{z}^k = z^k - z^*, \quad \overline{\lambda_p}^k = \lambda_p^k - \lambda_p^*, \quad \overline{\lambda_z}^k = \lambda_z^k - \lambda_z^*.$$

In this case, (4.18) still holds, which is represented as follows

$$\begin{array}{l}(r_z\|\overline{\lambda_p}^k\|^2 + r_p\|\overline{\lambda_z}^k\|^2) - (r_z\|\overline{\lambda_p}^{k+1}\|^2 + r_p\|\overline{\lambda_z}^{k+1}\|^2) \\ = -2r_pr_z(\overline{\lambda_p}^k, \overline{p}^k - \nabla\overline{u}^k) - r_p^2r_z\|\overline{p}^k - \nabla\overline{u}^k\|^2 - 2r_pr_z(\overline{\lambda_z}^k, \overline{z}^k - K\overline{u}^k) - r_z^2r_p\|\overline{z}^k - K\overline{u}^k\|^2.\end{array} \tag{4.33}$$

Since $(u^*, p^*, z^*; \lambda_p^*, \lambda_z^*)$ is a saddle-point of $\mathscr{L}(u, p, z; \lambda_p, \lambda_z)$, (u^*, p^*, z^*) is characterized by

$$\begin{array}{c}(\mathrm{div}\lambda_p^*, u - u^*) + r_p(\mathrm{div}(p^* - \nabla u^*), u - u^*) \\ +(\lambda_z^*, -K(u - u^*)) + r_z(z^* - Ku^*, -K(u - u^*)) \geq 0, \forall u \in V,\end{array} \tag{4.34}$$

$$R(p) - R(p^*) + (\lambda_p^*, p - p^*) + r_p(p^* - \nabla u^*, p - p^*) \geq 0, \forall p \in Q, \tag{4.35}$$

$$F(z) - F(z^*) + (\lambda_z^*, z - z^*) + r_z(z^* - Ku^*, z - z^*) \geq 0, \forall z \in V. \tag{4.36}$$

Similarly, by the construction of (u^k, p^k, z^k) (Algorithm 4.2 with $L = 1$), we have

$$\begin{aligned}&(\mathrm{div}\lambda_p^k, u - u^k) + r_p(\mathrm{div}(p^{k-1} - \nabla u^k), u - u^k)\\ &\quad+(\lambda_z^k, -K(u - u^k)) + r_z(z^{k-1} - Ku^k, -K(u - u^k)) \geq 0, \forall u \in V,\end{aligned} \tag{4.37}$$

$$R(p) - R(p^k) + (\lambda_p^k, p - p^k) + r_p(p^k - \nabla u^k, p - p^k) \geq 0, \forall p \in Q, \tag{4.38}$$

$$F(z) - F(z^k) + (\lambda_z^k, z - z^k) + r_z(z^k - Ku^k, z - z^k) \geq 0, \forall z \in V, \tag{4.39}$$

Taking $u = u^k$ in (4.34), $u = u^*$ in (4.37), $p = p^k$ in (4.35), $p = p^*$ in (4.38), $z = z^k$ in (4.36), and $z = z^*$ in (4.39), respectively, we obtain, after addition

$$\begin{aligned}&-(\overline{\lambda_p}^k, \overline{p}^k - \nabla\overline{u}^k) - (\overline{\lambda_z}^k, \overline{z}^k - K\overline{u}^k)\\ &\quad\geq r_p\|\overline{p}^k - \nabla\overline{u}^k\|^2 + r_z\|\overline{z}^k - K\overline{u}^k\|^2 + r_p(\nabla\overline{u}^k, \overline{p}^k - \overline{p}^{k-1}) + r_z(K\overline{u}^k, \overline{z}^k - \overline{z}^{k-1}).\end{aligned} \tag{4.40}$$

(4.33) and (4.40) indicate

$$\begin{aligned}&(r_z\|\overline{\lambda_p}^k\|^2 + r_p\|\overline{\lambda_z}^k\|^2) - (r_z\|\overline{\lambda_p}^{k+1}\|^2 + r_p\|\overline{\lambda_z}^{k+1}\|^2)\\ &\quad\geq r_p^2 r_z\|\overline{p}^k - \nabla\overline{u}^k\|^2 + r_p r_z^2\|\overline{z}^k - K\overline{u}^k\|^2\\ &\qquad+2r_p^2 r_z(\nabla\overline{u}^k, \overline{p}^k - \overline{p}^{k-1}) + 2r_p r_z^2(K\overline{u}^k, \overline{z}^k - \overline{z}^{k-1}).\end{aligned} \tag{4.41}$$

On the other hand, we have, by using the same technique as in [27, 58], the following estimates

$$\begin{cases}(\nabla\overline{u}^k, \overline{p}^k - \overline{p}^{k-1}) \geq \frac{1}{2}(\|\overline{p}^k\|^2 - \|\overline{p}^{k-1}\|^2 + \|\overline{p}^k - \overline{p}^{k-1}\|^2),\\ (K\overline{u}^k, \overline{z}^k - \overline{z}^{k-1}) \geq \frac{1}{2}(\|\overline{z}^k\|^2 - \|\overline{z}^{k-1}\|^2 + \|\overline{z}^k - \overline{z}^{k-1}\|^2).\end{cases} \tag{4.42}$$

We then obtain, from (4.41) and (4.42),

$$\begin{aligned}&(r_z\|\overline{\lambda_p}^k\|^2 + r_p\|\overline{\lambda_z}^k\|^2 + r_p^2 r_z\|\overline{p}^{k-1}\|^2 + r_p r_z^2\|\overline{z}^{k-1}\|^2)\\ &\qquad-(r_z\|\overline{\lambda_p}^{k+1}\|^2 + r_p\|\overline{\lambda_z}^{k+1}\|^2 + r_p^2 r_z\|\overline{p}^k\|^2 + r_p r_z^2\|\overline{z}^k\|^2)\\ &\geq r_p^2 r_z\|\overline{p}^k - \nabla\overline{u}^k\|^2 + r_p r_z^2\|\overline{z}^k - K\overline{u}^k\|^2\\ &\qquad+r_p^2 r_z\|\overline{p}^k - \overline{p}^{k-1}\|^2 + r_p r_z^2\|\overline{z}^k - \overline{z}^{k-1}\|^2,\end{aligned} \tag{4.43}$$

which implies

$$\begin{cases}\{\lambda_p^k : \forall k\}, \{\lambda_z^k : \forall k\}, \{p^k : \forall k\}, \{z^k : \forall k\}, \{\nabla u^k : \forall k\}, \text{ and } \{Ku^k : \forall k\} \text{ are bounded},\\ \lim\limits_{k\to\infty}\|p^k - \nabla u^k\| = 0,\\ \lim\limits_{k\to\infty}\|p^k - p^{k-1}\| = 0,\\ \lim\limits_{k\to\infty}\|z^k - Ku^k\| = 0,\\ \lim\limits_{k\to\infty}\|z^k - z^{k-1}\| = 0.\end{cases} \tag{4.44}$$

On the other hand, since $(u^*, p^*, z^*; \lambda_p^*, \lambda_z^*)$ is a saddle-point of $\mathscr{L}(u, p, z; \lambda_p, \lambda_z)$, we have

$$\begin{aligned} R(p^*) + F(z^*) \leq & R(p^k) + F(z^k) + (\lambda_p^*, p^k - \nabla u^k) + (\lambda_z^*, z^k - Ku^k) \\ & + \tfrac{r_p}{2}\|p^k - \nabla u^k\|^2 + \tfrac{r_z}{2}\|z^k - Ku^k\|^2. \end{aligned} \tag{4.45}$$

If we take $u = u^*$ in (4.37), $p = p^*$ in (4.38), and $z = z^*$ in (4.39), we have, by addition,

$$\begin{aligned} R(p^*) + F(z^*) \geq & R(p^k) + F(z^k) + (\lambda_p^k, p^k - \nabla u^k) + (\lambda_z^k, z^k - Ku^k) \\ & + r_p\|p^k - \nabla u^k\|^2 + r_z\|z^k - Ku^k\|^2 \\ & + r_p(\nabla \overline{u}^k, \overline{p}^k - \overline{p}^{k-1}) + r_z(K\overline{u}^k, \overline{z}^k - \overline{z}^{k-1}). \end{aligned} \tag{4.46}$$

Using (4.44), we have

$$\liminf(R(p^k) + F(z^k)) \geq R(p^*) + F(z^*) \geq \limsup(R(p^k) + F(z^k)), \tag{4.47}$$

by taking $\liminf$ in (4.45) and $\limsup$ in (4.46). This completes the proof of (4.32).

By the continuity of $R(\cdot)$ and $F(\cdot)$ over their domains, (4.32) indicates clearly that u^k is a minimizing sequence of $E(\cdot)$. If the minimizer of $E(\cdot)$ is unique, then $u^k \to u^*$. ■

We would like to add a comment on Theorem 4.3. It is stated in [64] that augmented Lagrangian method requires (numerically) increasing accuracy of the inner iteration to ensure the convergence of the overall algorithm. Theorem 4.3 indicates that, even if we just simply set $L = 1$ (thus not explicitly increasing the accuracy by checking optimality conditions), the accuracy of the inner iteration will also essentially and automatically increase, justifying the statement in [64]. As a consequence, setting $L = 1$ provides a simple stopping criterion of the inner iteration, which does not need to compute those optimality conditions.

5. Applications. In this section we apply augmented Lagrangian method to TV restoration with some typical and important non-quadratic fidelities. We focus on TV-L^1 restoration for recovering blurred images corrupted by impulsive noise (e.g., salt-and-pepper noise and random-valued noise), and TV-KL restoration for recovering blurred images corrupted by Poisson noise. In these two cases, the $z-$sub problems have closed form solutions, which can be solved very efficiently. For the sake of completeness, we elaborate Algorithm 4.2 for TV-L^1 and TV-KL restoration as the following Algorithm 5.1 and Algorithm 5.2, respectively. Moreover, we will prove the convergence of these two algorithms.

5.1. Augmented Lagrangian method for TV-L^1 restoration. TV-L^1 restoration model is especially useful for deblurring images corrupted by impulsive noise. It aims at solving the following minimization problem:

$$\min_{u \in V}\{E_{\mathrm{TV}L^1}(u) = R(\nabla u) + \alpha\|Ku - f\|_{L^1}\}, \tag{5.1}$$

where $R(\nabla u) = \mathrm{TV}(u)$. The fidelity term is non-quadratic (and even non-differentiable).

The problem (5.1) is a special case of (3.2) where the fidelity term is

$$F(Ku) = \alpha\|Ku - f\|_{L^1}. \tag{5.2}$$

Therefore we can apply Algorithms 4.1 and 4.2 to solve (5.1). For this special fidelity, we have the following explicit solution for the $z-$sub problem (4.9):

$$z_{i,j} = f_{i,j} + \max(0, 1 - \frac{\alpha}{r_z|w_{i,j} - f_{i,j}|})(w_{i,j} - f_{i,j}), \tag{5.3}$$

where

$$w = Ku - \frac{\lambda_z^k}{r_z} \in V. \tag{5.4}$$

The derivation of (5.3) is similar to (4.11) by the geometric interpretation.

Hence in this case Algorithm 4.2 can be detailed as follows.

Algorithm 5.1 Augmented Lagrangian method for TV-L^1 restoration – solve the minimization problem (4.5)

- *Initialization*: $u^{k,0} = u^{k-1}, p^{k,0} = p^{k-1}, z^{k,0} = z^{k-1}$;
- *For* $l = 0, 1, 2, ..., L-1$:
  - compute $u^{k,l+1}$ from (4.10) for $p = p^{k,l}, z = z^{k,l}$;
  - compute $p^{k,l+1}$ from (4.11) for $u = u^{k,l+1}$;
  - compute $z^{k,l+1}$ from (5.3) for $u = u^{k,l+1}$;
- $u^k = u^{k,L}, p^k = p^{k,L}, z^k = z^{k,L}$.

We have the following convergence result for Algorithm 5.1.

THEOREM 5.1. *For TV-L^1 restoration, the sequence* $\{(u^{k,l}, p^{k,l}, z^{k,l}) : l = 0, 1, 2, \cdots\}$ *generated by Algorithm 5.1 converges to a solution of the problem* (4.5).

Proof The proof is motivated by [60] and similar to that of Theorem 4.2 in [58]. Here we just sketch the differences.

Similarly with s_τ (and s) in [60], we define operators s_1 (and s_2) as

$$s_1(t) = \max(0, 1 - \frac{\alpha}{r_z|t|})t, \text{ for } t \in \mathbb{R},$$

and

$$s_2(\mathbf{t}) = \max(0, 1 - \frac{1}{r_p|\mathbf{t}|})\mathbf{t}, \text{ for } \mathbf{t} \in \mathbb{R}^2.$$

According to (5.3), it is useful to further define

$$(S_1)_{i,j}(t) = f_{i,j} + s_1(t),$$

for each pixel (pair (i,j) of index).

By $(S_1)_{i,j}$ and s_2, we then construct operators S_1 and S_2 such that (5.3) and (4.11) can be reformulated as $z = S_1(w - f)$ and $p = S_2(\mathbf{w})$, respectively, with w defined in (5.4) and $\mathbf{w}$ in (4.12). Therefore the iterative scheme in Algorithm 5.1 can be written as

$$\begin{cases} u^{k,l+1} = (r_z K^* K + r_p \nabla^* \nabla)^{-1} (K^*(\lambda_z^k + r_z z^{k,l}) + \nabla^*(\lambda_p^k + r_p p^{k,l})), \\ p^{k,l+1} = S_2(\nabla u^{k,l+1} - \frac{\lambda_p^k}{r_p}), \\ z^{k,l+1} = S_1(K u^{k,l+1} - \frac{\lambda_z^k}{r_z} - f), \end{cases} \tag{5.5}$$

where $\nabla^* = -\text{div}$ is the adjoint operator of ∇. Here we also mention the existence of $(r_z K^* K + r_p \nabla^* \nabla)^{-1}$ for the assumption $\text{Null}(\nabla) \cap \text{Null}(K) = \{0\}$.

Furthermore, we define two linear operators $h_2 : Q \times V \to Q$ and $h_1 : Q \times V \to V$ as follows:

$$\begin{cases} h_2(p, z) = \nabla (r_z K^* K + r_p \nabla^* \nabla)^{-1} (K^*(\lambda_z^k + r_z z) + \nabla^*(\lambda_p^k + r_p p)) - \frac{\lambda_p^k}{r_p}, \\ h_1(p, z) = K (r_z K^* K + r_p \nabla^* \nabla)^{-1} (K^*(\lambda_z^k + r_z z) + \nabla^*(\lambda_p^k + r_p p)) - \frac{\lambda_z^k}{r_z} - f. \end{cases} \tag{5.6}$$

Rewriting the iterative scheme (5.5) as

$$\left\{ \begin{array}{l} u^{k,l+1} = (r_z K^* K + r_p \nabla^* \nabla)^{-1} (K^* (\lambda_z^k + r_z z^{k,l}) + \nabla^* (\lambda_p^k + r_p p^{k,l})), \\ (p^{k,l+1}, z^{k,l+1}) = (S_2 \circ h_2; S_1 \circ h_1)(p^{k,l}, z^{k,l}), \end{array} \right. \tag{5.7}$$

one can show the convergence via a similar argument in [60]. ■

Here we show some examples. In Tables 5.1, 5.2, and 5.3, we compute the TV-L^1 model for removing 7×7 sized Gaussian blur and salt-and-pepper noise from 30% to 60%. The TV-L^1 model is also computed for removing 7×7 sized Gaussian blur and random-valued noise from 20% to 50% in Tables 5.4 and 5.5. In Table 5.6 an example is provided to show the TV-L^1 restoration of the cameraman degraded by 15×15 sized Gaussian blur and salt-and-pepper noise from 30% to 60%.

In each figure, α, t, and SNR denote the parameter of the model, the CPU cost (in seconds), and the signal-noise ratio of the image, respectively. Note here we use the same α's for all the methods in each example, since our goal is to compare the efficiency of different methods for the same model.

We compare our method (ALM with parameters r_p and r_z) with the FTVd package. The FTVd_v2.0 is denoted for the FTVd version 2.0, and FTVd_v4.0 is for FTVd version 4.0. As far as we know, FTVd version 2.0 is one of the most efficient published algorithms for TV-L^1 restoration; see [60]. When this paper was nearly finished, we got to know that the group of Prof. Wotao Yin had released FTVd version 4.0 recently. Therefore we compare our method to these two versions. As one can see, augmented Lagrangian method is much more efficient than FTVd version 2.0. The potential reason may be as follows. First, in our method, we simply set $L = 1$ for inner iteration and hence do not need to compute those residuals for stopping criterion, which are calculated in FTVd version 2.0. Second, augmented Lagrangian method benefits from its Lagrange multipliers update, which can be actually interpreted as sub-gradients update in split Bregman iteration, and makes the method extremely efficient for homogeneous 1 objective functionals. The performances of our method and FTVd version 4.0 are very similar. For low noise level, our method seems to be a little more efficient that FTVd version 4.0. For high noise level, FTVd version 4.0 appears to be a bit better than ours.

5.2. Augmented Lagrangian method for TV-KL restoration. To deblur images corrupted by Poisson noise, KL divergence is used as the data fidelity. In particular, we consider the following minimization problem:

$$\min_{u \in V} \{E_{\mathrm{TVKL}}(u) = R(\nabla u) + \alpha \sum_{1 \le i,j \le N} ((Ku)_{i,j} - f_{i,j} \log (Ku)_{i,j}) : (Ku)_{i,j} > 0, \forall (i,j)\}, \tag{5.8}$$

where $R(\nabla u) = \mathrm{TV}(u)$.

The problem (5.8) is a special case of (3.2) where

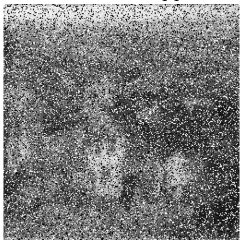
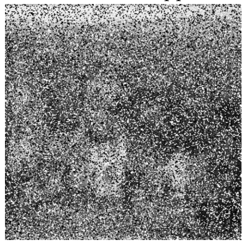
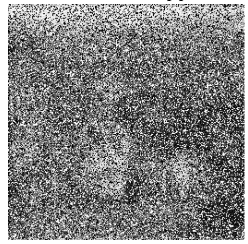
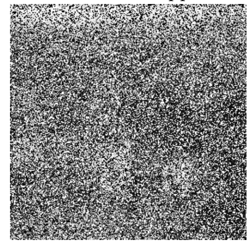
$$F(Ku) = \left\{ \begin{array}{ll} \alpha \sum\limits_{1 \le i,j \le N} ((Ku)_{i,j} - f_{i,j} \log (Ku)_{i,j}), & u \in V, (Ku)_{i,j} > 0 \\ +\infty, & \text{otherwise} \end{array} \right. . \tag{5.9}$$

Therefore, Algorithms 4.1 and 4.2 can be applied to compute (5.8). For this special fidelity, we also have, by considering $z_{i,j} > 0$, a closed form solution to the z-sub problem (4.9):

$$z_{i,j} = \frac{1}{2} \left(\sqrt{(w_{i,j} - \frac{\alpha}{r_z})^2 + 4 \frac{\alpha}{r_z} f_{i,j}} + (w_{i,j} - \frac{\alpha}{r_z}) \right), \tag{5.10}$$

Table 5.1

TV-L^1 restoration from 7×7 sized Gaussian blur with salt-and-pepper noise from 30% to 60%.

Blurry&Noisy: 30% Salt&Pepper	40% Salt&Pepper	50% Salt&Pepper	60% Salt&Pepper
			
Recovered (α: 13) (FTVd_v2.0) t: 11.8s, SNR: 13.78dB	Recovered (α: 10) (FTVd_v2.0) t: 13.1s, SNR: 12.97dB	Recovered (α: 8) (FTVd_v2.0) t: 14.6s, SNR: 12.21dB	Recovered (α: 4) (FTVd_v2.0) t: 18.3s, SNR: 10.98dB
			
Recovered (α: 13) (FTVd_v4.0) t: 3.7s, SNR: 13.19dB	Recovered (α: 10) (FTVd_v4.0) t: 3.9s, SNR: 12.66dB	Recovered (α: 8) (FTVd_v4.0) t: 3.5s, SNR: 12.07dB	Recovered (α: 4) (FTVd_v4.0) t: 3.3s, SNR: 10.95dB
			
Recovered (α: 13) (ALM, r_p: 20, r_z: 100) t: 2.8s, SNR: 13.28dB	Recovered (α: 10) (ALM, r_p: 20, r_z: 100) t: 3.1s, SNR: 12.73dB	Recovered (α: 8) (ALM, r_p: 20, r_z: 100) t: 3.5s, SNR: 12.08dB	Recovered (α: 4) (ALM, r_p: 10, r_z: 25) t: 3.3s, SNR: 10.86dB
			

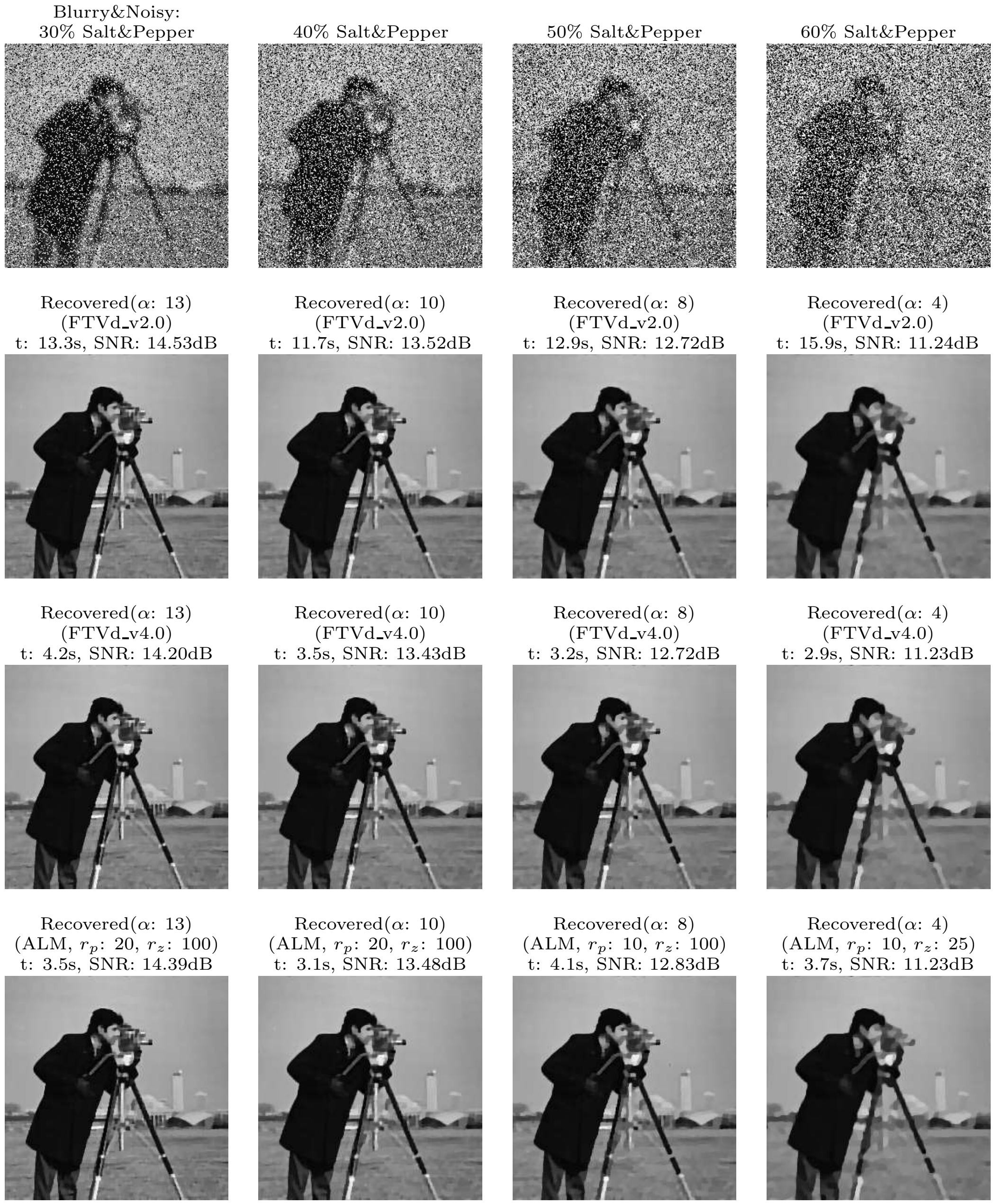
where

$$w = Ku - \frac{\lambda_z^k}{r_z} \in V. \tag{5.11}$$

Here we elaborate Algorithm 4.2 for TV-KL restoration as Algorithm 5.2.

Table 5.2

TV-L^1 restoration from 7×7 sized Gaussian blur with salt-and-pepper noise from 30% to 60%.



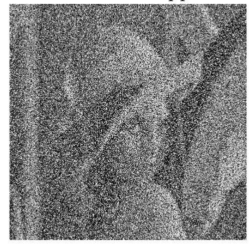
For Algorithm 5.2, we have the following convergence result.

Theorem 5.2. *For TV-KL restoration, the sequence* $\{(u^{k,l}, p^{k,l}, z^{k,l}) : l = 0, 1, 2, \cdots\}$ *generated by Algorithm 5.2 converges to a solution of the problem* (4.5).

Proof As one can see, the only difference between Algorithm 5.2 and 5.1 is in the solutions of the $z-$sub problems. We therefore define a mapping $\Psi = (\psi_{i,j}) : V \to V$,

Table 5.3

TV-L^1 restoration from 7×7 sized Gaussian blur with salt-and-pepper noise from 30% to 60%.

Blurry&Noisy: 30% Salt&Pepper	40% Salt&Pepper	50% Salt&Pepper	60% Salt&Pepper
			
Recovered(α: 13) (FTVd_v2.0) t: 75.5s, SNR: 17.77dB	Recovered(α: 10) (FTVd_v2.0) t: 76.8s, SNR: 16.79dB	Recovered(α: 8) (FTVd_v2.0) t: 77.4s, SNR: 15.92dB	Recovered(α: 4) (FTVd_v2.0) t: 92.7s, SNR: 14.43dB
			
Recovered(α: 13) (FTVd_v4.0) t: 17.1s, SNR: 16.62dB	Recovered(α: 10) (FTVd_v4.0) t: 16.3s, SNR: 16.12dB	Recovered(α: 8) (FTVd_v4.0) t: 15.4s, SNR: 15.53dB	Recovered(α: 4) (FTVd_v4.0) t: 14.1s, SNR: 14.28dB
			
Recovered(α: 13) (ALM, r_p: 20, r_z: 100) t: 11.6s, SNR: 16.90dB	Recovered(α: 10) (ALM, r_p: 20, r_z: 100) t: 12.0s, SNR: 16.30dB	Recovered(α: 8) (ALM, r_p: 20, r_z: 100) t: 13.9s, SNR: 15.51dB	Recovered(α: 4) (ALM, r_p: 10, r_z: 35) t: 16.3s, SNR: 14.26dB
			

according to (5.10), with $\psi_{i,j}$ as

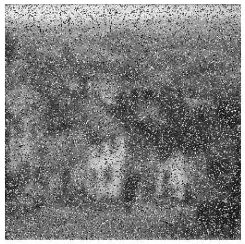
$$\psi_{i,j}(t) = \frac{1}{2}(\sqrt{t^2 + 4\frac{\alpha}{r_z} f_{i,j}} + t). \tag{5.12}$$

In the following we prove the convergence in three steps.

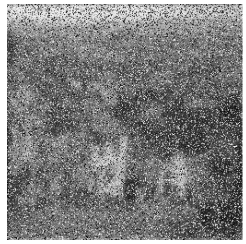
TABLE 5.4
TV-L^1 restoration from 7×7 sized Gaussian blur with random-valued noise from 20% to 50%.

Blurry&Noisy:
20% Random-Valued

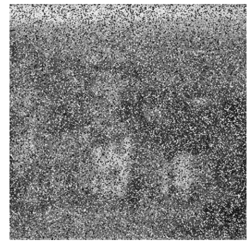
30% Random-Valued



40% Random-Valued



50% Random-Valued



Recovered(α: 25)
(FTVd_v2.0)
t: 9.9s, SNR: 15.24dB

Recovered(α: 10)
(FTVd_v2.0)
t: 12.3s, SNR: 13.29dB

Recovered(α: 8)
(FTVd_v2.0)
t: 14.0s, SNR: 12.44dB

Recovered(α: 4)
(FTVd_v2.0)
t: 16.8s, SNR: 10.83dB

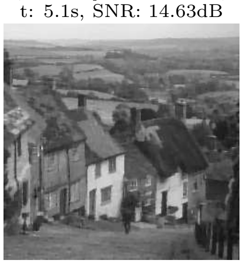
Recovered(α: 25)
(FTVd_v4.0)
t: 5.7s, SNR: 13.62dB

Recovered(α: 10)
(FTVd_v4.0)
t: 4.0s, SNR: 12.97dB

Recovered(α: 8)
(FTVd_v4.0)
t: 3.1s, SNR: 12.32dB

Recovered(α: 4)
(FTVd_v4.0)
t: 3.1s, SNR: 11.00dB

Recovered(α: 25)
(ALM, r_p: 20, r_z: 120)
t: 5.1s, SNR: 14.63dB



Recovered(α: 10)
(ALM, r_p: 20, r_z: 100)
t: 2.8s, SNR: 13.04dB

Recovered(α: 8)
(ALM, r_p: 15, r_z: 120)
t: 3.2s, SNR: 12.33dB

Recovered(α: 4)
(ALM, r_p: 10, r_z: 45)
t: 3.4s, SNR: 10.90dB

First, we show the sequence $\{(u^{k,l}, p^{k,l}, z^{k,l}) : l = 0, 1, 2, \cdots\}$ is bounded. According to Algorithm 5.2, we have

$$\begin{aligned}
&\mathscr{L}(u^{k,l+1}, p^{k,l}, z^{k,l}; \lambda_p^k, \lambda_z^k) \le \mathscr{L}(u^{k,l}, p^{k,l}, z^{k,l}; \lambda_p^k, \lambda_z^k),\\
&\mathscr{L}(u^{k,l+1}, p^{k,l+1}, z^{k,l}; \lambda_p^k, \lambda_z^k) \le \mathscr{L}(u^{k,l+1}, p^{k,l}, z^{k,l}; \lambda_p^k, \lambda_z^k),\\
&\mathscr{L}(u^{k,l+1}, p^{k,l+1}, z^{k,l+1}; \lambda_p^k, \lambda_z^k) \le \mathscr{L}(u^{k,l+1}, p^{k,l+1}, z^{k,l}; \lambda_p^k, \lambda_z^k).
\end{aligned}$$

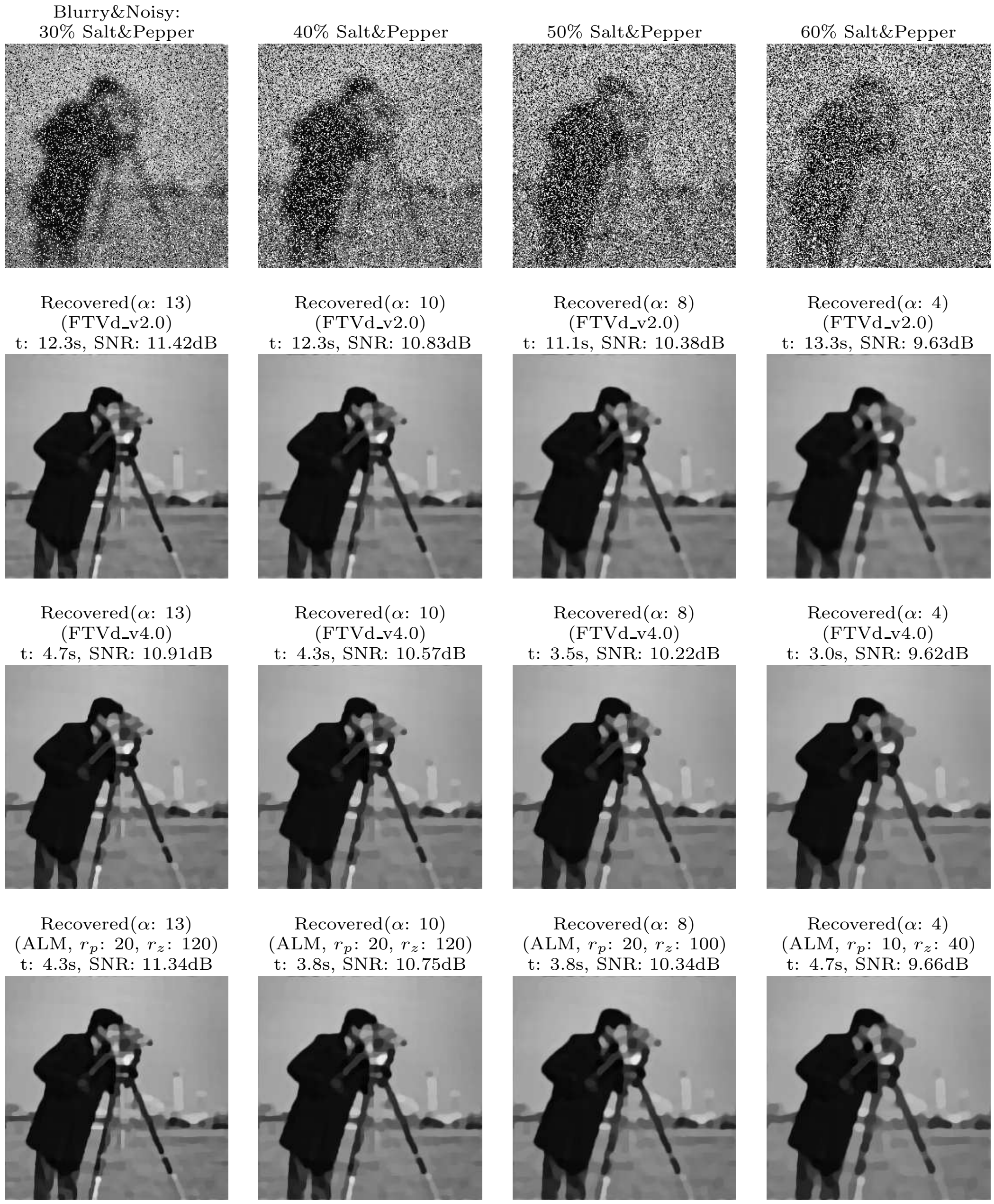
Table 5.5
TV-L^1 restoration from 7×7 sized Gaussian blur with random-valued noise from 20% to 50%.

By adding the above three equations, we have

$$\mathscr{L}(u^{k,l+1}, p^{k,l+1}, z^{k,l+1}; \lambda_p^k, \lambda_z^k) \leq \mathscr{L}(u^{k,l}, p^{k,l}, z^{k,l}; \lambda_p^k, \lambda_z^k),$$

indicating that $\mathscr{L}(u^{k,l}, p^{k,l}, z^{k,l}; \lambda_p^k, \lambda_z^k)$ is monotonically decreasing. Since $\mathscr{L}(u, p, z; \lambda_p^k, \lambda_z^k)$ is proper and coercive with respect to (u, p, z), $\{(u^{k,l}, p^{k,l}, z^{k,l}) : l = 0, 1, 2, \cdots\}$ is

TABLE 5.6
TV-L^1 restoration from 15×15 sized Gaussian blur with salt-and-pepper noise from 30% to 60%.



bounded.

Secondly, we verify the mapping $\psi_{i,j}$ is non-expansive (actually a contraction mapping) over bounded domains. Given a bounded domain B and any $t_1 \in B, t_2 \in B$, we have, by basic calculus,

$$|\psi_{i,j}(t_1) - \psi_{i,j}(t_2)| \leq M|t_1 - t_2|,$$

Algorithm 5.2 Augmented Lagrangian method for TV-KL restoration – solve the minimization problem (4.5)

- *Initialization*: $u^{k,0} = u^{k-1}, p^{k,0} = p^{k-1}, z^{k,0} = z^{k-1}$;
- *For* $l = 0, 1, 2, ..., L-1$:
  - compute $u^{k,l+1}$ from (4.10) for $p = p^{k,l}, z = z^{k,l}$;
  - compute $p^{k,l+1}$ from (4.11) for $u = u^{k,l+1}$;
  - compute $z^{k,l+1}$ from (5.10) for $u = u^{k,l+1}$;
- $u^k = u^{k,L}, p^k = p^{k,L}, z^k = z^{k,L}$.

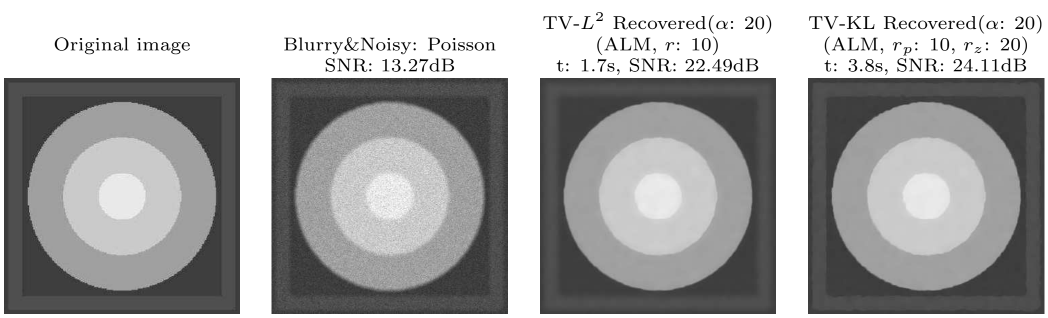
with a constant $M < 1$. Here we used the assumption for TV-KL restoration that $f_{i,j} > 0, \forall(i,j)$.

On the third, the convergence of Algorithm 5.2 can be proved similarly with that of Algorithm 5.1. ■

We show some examples; see Tables 5.7 and 5.8. In the figures, α, t, and SNR denote the parameter of the model, CPU costs (in seconds) and signal-noise ratios, respectively. We compare the restoration results of TV-L^2 and TV-KL models calculated by augmented Lagrangian method (ALM) with parameters r and r_p, r_z, respectively. As one can see, TV-KL model produces much better results than TV-L^2. TV-L^2 removes the noise, but has difficulty to preserve sharp edges (see the black frame in the example of Table 5.7), and blurs textures too much (see the zoom-in pictures in Table 5.8). In addition, TV-KL model can still be calculated very efficiently by augmented Lagrangian method. In one word, augmented Lagrangian method for TV-KL model produces much better results than TV-L^2 model with an acceptable CPU cost, when recovering blurred images with Poisson noise.

Table 5.7

Comparisons between TV-L^2 and TV-KL restoration: recovering degraded images with 7×7 sized Gaussian blur and Poisson noise.



6. Conclusion. In this paper we extended augmented Lagrangian method for TV-L^2 model to solve TV restoration with non-quadratic fidelity. After presenting and analyzing the method for TV restoration with a relatively quite general fidelity, we applied the algorithms to two typical image deblurring problems with non-Gaussian noise. Due to FFT implementation or closed form solutions for the sub-problems, as well as simple stopping criterion ($L = 1$) of the inner iteration, augmented Lagrangian method is extremely efficient as demonstrated by the experiments. Moreover, we gave convergence analysis for the proposed algorithms, which cannot be obtained through previous analysis techniques. A possible future work is to extend the method to color

Table 5.8

Comparisons between TV-L^2 and TV-KL restoration: recovering degraded images with 7×7 sized Gaussian blur and Poisson noise. The second row is the zoom-in of the first row.

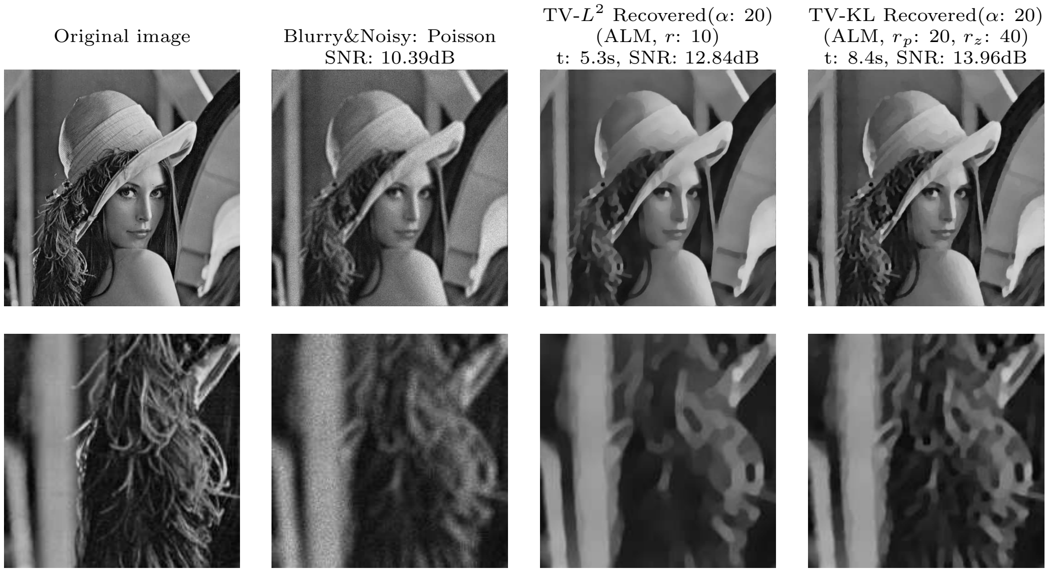


image recovering via TV (and even Non-Local TV) restoration with non-quadratic fidelities.

Acknowledgement. The research has been supported by MOE (Ministry of Education) Tier II project T207N2202 and IDM project NRF2007IDM-IDM002-010. In addition, support from SUG 20/07 is also gratefully acknowledged.

REFERENCES

[1] S. Alliney, *Digital filters as absolute norm regularizers*, IEEE Trans. Signal Process., 40(1992), pp. 1548–1562.

[2] P. Besbeas, I.D. Fies, and T. Sapatinas, *A comparative simulation study of wavelet shrinkage estimators for Poisson counts*, International Statistical Review, 72(2004), pp. 209–237.

[3] P. Blomgren, and T.F. Chan, *Color TV: Total Variation Methods for Restoration of Vector-Valued Images*, IEEE Trans. Image Process., vol. 7, no. 3, pp. 304–309, 1998.

[4] A. Bovik, *Handbook of image and video processing*, New York: Academic, 2000.

[5] X. Bresson, and T.F. Chan, *Fast Minimization of the Vectorial Total Variation Norm and Applications to Color Image Processing*, UCLA CAM Report.

[6] C. Brune, A. Sawatzky, and M. Burger, *Bregman-EM-TV methods with application to optical nanoscopy*, Proc. SSVM 2009, LNCS, 5567(2009), pp. 235–246.

[7] A. Caboussat, R. Glowinski, and V. Pons, *An Augmented Lagrangian Approach to the Numerical Solution of a Non-Smooth Eigenvalue Problem*, Journal of Numerical Mathematics, to appear.

[8] E. Candes, J. Romberg, and T. Tao, *Robust uncertainty principles: exact signal reconstruction from highly imcomplete frequency information*, IEEE Trans. Inform. Theory, 52(2006), pp. 489–509.

[9] J.L. Carter, *Dual Methods for Total Variation - Based Image Restoration*, Ph.D. thesis, UCLA, 2001.

[10] A. Chambolle, and P.L. Lions, *Image Recovery via Total Variation Minimization and Related Problems*, Numer. Math., 76(1997), pp. 167–188.

[11] A. Chambolle, *An Algorithm for Total Variation Minimization and Applications*, J. Math. Imaging Vis., 20(2004), pp. 89–97.

[12] R. Chan, C.W. Ho, and M. Nikolova, *Salt-and-pepper noise removal by median-type noise detector and detail-preserving regularization*, IEEE Trans. Image Process., 14(2005), pp. 1479–1485.

[13] R.H. Chan, and K. Chen, *Multilevel algorithms for a Poisson noise removal model with total variation regularization*, Int. J. Comput. Math., 84(2007), pp. 1183–1198.

[14] T.F. Chan, G.H. Golub, and P. Mulet, *A Nonlinear Primal-Dual Method for Total Variation-Based Image Restoration*, SIAM J. Sci. Comput., 20(1999), pp. 1964–1977.

[15] T. Chan, A. Marquina, and P. Mulet, *High-Order Total Variation-Based Image Restoration*, SIAM J. Scientific Comput., 22(2000), pp. 503–516.

[16] T.F. Chan, S.H. Kang and J.H. Shen, *Total Variation Denoising and Enhancement of Color Images Based on the CB and HSV Color Models*, J. Visual Commun. Image Repres., vol. 12, pp. 422–435, 2001.

[17] T.F. Chan, and S. Esedoglu, *Aspects of total variation regularized L^1 function approximation*, SIAM J. Appl. Math., 65(2005), pp. 1817–1837.

[18] S. Chen, D. Donoho, and M.A. Saunders, *Atomic decomposition by basis pursuit*, SIAM J. Sci. Comput., 20(1998), pp. 33–61.

[19] T. Chen, and H.R. Wu, *Space variant median filters for the restoration of impulse noise corrupted images*, IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process., 48(2001), pp. 784–789.

[20] Y. Dong, M. Hintermüller, and M. Neri, *An efficient primal-dual method for L^1TV image restoration*, Accepted by SIAM J. Imaging Sci.

[21] D. Donoho, *Nonlinear wavelet methods for recovery of signals, densities and spectra from indirect and noisy data*, in Proc. of symposia in applied mathematics: Different perspectives on wavelets, American Mathematical Society, 1993, pp. 173–205.

[22] D.L. Donoho, *Compressed sensing*, IEEE Trans. Inform. Theory, 52(2006), pp. 1289–1306.

[23] I. Ekeland, and R. Témam, *Convex Analysis and Variational Problems*, SIAM, 1999.

[24] H.L. Eng, and K.K. Ma, *Noise adaptive soft-switching median filter*, IEEE Trans. Image Process., 10(2001), pp. 242–251.

[25] E. Esser, *Applications of Lagrangian-Based Alternating Direction Methods and Connections to Split Bregman*, UCLA CAM Report, 09–31.

[26] H.Y. Fu, M.K. Ng, M. Nikolova, and J.L. Barlow, *Efficient minimization methods of mixed l2-l1 and l1-l1 norms for image restoration*, SIAM J. Sci. Comput., 27(2006), pp. 1881–1902.

[27] R. Glowinski, P. Le Tallec, *Augmented Lagrangians and Operator-Splitting Methods in Nonlinear Mechanics*, SIAM, Philadelphia (1989).

[28] T. Goldstein, and S. Osher, *The Split Bregman Method for L1 Regularized Problems*, SIAM Journal on Imaging Sciences, 2(2009), pp. 323–343.

[29] M.R. Hestenes, *Multiplier and Gradient Methods*, Journal of Optimization Theory and Applications 4, 303–320 (1969).

[30] W. Hinterberger, and O. Scherzer, *Variational Methods on the Space of Functions of Bounded Hessian for Convexification and Denoising*, Computing, vol. 76, pp. 109–133, 2006.

[31] M. Hintermüller, K. Ito, and K. Kunisch, *The primal-dual active set strategy as a semismooth newton method*, SIAM J. Optim., 13 (2002), pp. 865–888.

[32] Y. Huang, M. Ng and Y. Wen, *A Fast Total Variation Minimization Method for Image Restoration*, SIAM Multiscale Modeling and Simulation, accepted, 2008.

[33] H. Hwang, and R.A. Haddad, *Adaptive median filters: New algorithms and results*, IEEE Trans. Image Process., 4(1995), pp. 499–502.

[34] C. Kervrann, and A. Trubuil, *An adaptive window approach for poisson noise reduction and structure preserving in confocal microscopy*, in International Symposium on Biomedical Imaging (ISBI'04), Arlington, April 2004.

[35] E. Kolaczyk, *Wavelet shrinkage estimation of certain Poisson intensity signals using corrected thresholds*, Statist. Sinica, 9(1999), pp. 119–135.

[36] T. Le, R. Chartrand, and T.J. Asaki, *A variational approach to reconstructing images corrupted by Poisson noise*, J. Math. Imaging Vision, 27(2007), pp. 257–263.

[37] Y. Li, and S. Osher, *A new median formula with applications to PDE based denoising*, UCLA CAM Report, cam08-73.

[38] M. Lysaker, A. Lundervold, and X.-C. Tai, *Noise removal using fourth-order partial differential equation with applications to medical Magnetic Resonance Images in space and time*, IEEE Trans. Image Process., 12(2003), pp. 1579–1590.

[39] M. Lysaker, and X.-C. Tai, *Iterative Image Restoration Combining Total Variation Minimization and a Second Order Functional*, Int'l J. Computer Vision, 2005.

[40] P. Mrazek, J. Weickert, and A. Bruhn, *On robust estimations and smoothing with spatial and tonal kernels*, In Geometric Properties from Incomplete Data, pp. 335–352, Springer, 2006.

[41] P.E. Ng, and K.K. Ma, *A switching median filter with boundary discriminative noise detection for extremely corrupted images*, IEEE Trans. Image Process., 15(2006), pp. 1506–1516.

[42] M. Nikolova, *Minimizers of cost-functions involving non-smooth data fidelity terms*, SIAM J. Num. Ana., 40(2002), pp. 965–994.

[43] M. Nikolova, *A variational approach to remove outliers and impulse noise*, J. Math. Imaging Vision, 20(2004), pp. 99–120.

[44] V.Y. Panin, G.L. Zeng, and G.T. Gullberg, *Total variation regulated EM algorithm [SPECT reconstruction]*, IEEE Trans. Nucl. Sci., 46(1999), pp. 2202–2210.

[45] G. Pok, J.C. Liu, and A.S. Nair, *Selective removal of impulse noise based on homogeneity level information*, IEEE Trans. Image Process., 12(2003), pp. 85–92.

[46] M.J.D. Powell, *A Method for Nonlinear Constraints in Minimization Problems*, Optimization, Fletcher, R. ed., Academic Press, New York, 283–298 (1972).

[47] R.T. Rockafellar, *A Dual Approach to Solving Nonlinear Programming Problems by Unconstrained Optimization*, Mathematical Programming 5, 354–373 (1973).

[48] L. Rudin, S. Osher, and E. Fatemi, *Nonlinear Total Variation Based Noise Removal Algorithms*, Physica D, 60(1992), pp. 259–268.

[49] G. Sapiro, and D.L. Ringach, *Anisotropic Diffusion of Multivalued Images with Applications to Color Filtering*, IEEE Trans. Image Process., vol. 5, no. 11, pp. 1582–1586, 1996.

[50] O. Scherer, *Denoising With Higher Order Derivatives of Bounded Variation and an Application to Parameter Estimation*, Computing, vol. 60, pp. 1–27, 1998.

[51] S. Setzer, *Split Bregman Algorithm, Douglas-Rachford Splitting and Frame Shrinkage*, Scale Space and Variational Methods in Computer Vision, Second International Conference, SSVM 2009, Voss, Norway, June 1-5, 2009. Proceedings. Lecture Notes in Computer Science 5567, pp. 464-476, Springer, 2009.

[52] L.A. Shepp, and Y. Vardi, *Maximum likelihood reconstruction for emission tomography*, IEEE Trans. Medical Imaging, 1(1982), pp. 113–122.

[53] G. Steidl, *A Note on the Dual Treatment of Higher-Order Regularization Functionals*, Computing, vol. 76, pp. 135–148, 2006.

[54] X.C. Tai, C.L. Wu, *Augmented Lagrangian method, dual methods and split Bregman iteration for ROF model*, Scale Space and Variational Methods in Computer Vision, Second International Conference, SSVM 2009, Voss, Norway, June 1-5, 2009. Proceedings. Lecture Notes in Computer Science 5567, pp. 502-513, Springer, 2009.

[55] K. Timmermann, and R. Novak, *Multiscale modeling and estimation of Poisson processes with applications to photon-limited imaging*, IEEE Trans. Inf. Theor., 45(1999), pp. 846–852.

[56] Y.L. Wang, W.T. Yin, and Y. Zhang, *A Fast Algorithm for Image Deblurring with Total Variation Regularization*, UCLA CAM Report.

[57] Y. Wang, J. Yang, W. Yin, and Y. Zhang, *A New Alternating Minimization Algorithm for Total Variation Image Reconstruction*, SIAM Journal on Imaging Sciences, 1(2008), pp. 248–272.

[58] C.L. Wu, and X.C. Tai, *Augmented Lagrangian method, dual methods, and split Bregman iteration for ROF, vectorial TV, and high order models*, UCLA CAM Report, cam09-76.

[59] J.F. Yang, W.T. Yin, Y. Zhang, and Y.L. Wang, *A Fast Algorithm for Edge-Preserving Variational Multichannel Image Restoration*, UCLA CAM Report 08-50.

[60] J. Yang, Y. Zhang, and W. Yin, *An efficient TVL1 algorithm for deblurring multichannel images corrupted by impulsive noise*, SIAM J. Sci. Comput., 31(2009), pp. 2842-2865.

[61] W. Yin, D. Goldfarb, and S. Osher, *Image cartoon-texture decomposition and feature selection using the total variation regularized L^1 functional*, in Variational, Geometric, and Level Set Methods in Computer Vision 2005, LNCS, 3752(2005), pp. 73–84.

[62] W. Yin, D. Goldfarb, and S. Osher, *The total variation regularized L^1 model for multiscale decomposition*, Multis. Model. Simul., 6(2006), pp. 190–211.

[63] W.T. Yin, S. Osher, D. Goldfarb and J. Darbon, *Bregman Iterative Algorithms for Compressend Sensing and Related Problems*, SIAM J. Imaging Sciences, 1(2008), pp. 143–168.

[64] W.T. Yin, *Analysis and Generalizations of the Linearized Bregman Method*, UCLA CAM Report 09-42.

[65] Y.-L. You and M. Kaveh, *Fourth-Order Partial Differential Equation for Noise Removal*, IEEE Trans. Image Process., 9(2000), pp. 1723–1730.

[66] R. Zanella, P. Boccacci, L. Zanni, and M. Bertero, *Efficient gradient projection methods for edge-preserving removal of Poisson noise*, Inverse Problems, 25(2009), pp.

[67] M. Zhu and T. F. Chan, *An efficient primal-dual hybrid gradient algorithm for total variation image restoration*, UCLA CAM Report 08-34, 2008.

[68] M. Zhu, S.J. Wright, T.F. Chan, *Duality-Based Algorithms for Total Variation Image Restora-*

tion, UCLA CAM Report 08-33, 2008.