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

Chunlin Wu

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

Juyong Zhang

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

Xue-Cheng Tai

Division of Mathematical Sciences
School of Physical & Mathematical Sciences
Nanyang Technological University, Singapore
and
Department of Mathematics, University of Bergen, Norway

(Communicated by Luminita Vese)

Abstract. Recently augmented Lagrangian method has been successfully applied to image restoration. We extend the method to total variation (TV) restoration models with non-quadratic fidelities. We will first introduce the method and present an 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 Fast Fourier Transform (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. Numerical experiments demonstrate the efficiency of our method.

1. **Introduction.** Total variation regularization was first introduced in [48]. It has been demonstrated very successful in image restoration and extensively generalized [10, 15, 64, 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] for compressive sensing [8, 21]. 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, 66, 67, 56, 58, 32, 62, 28, 54, 57].

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, 60, 61]. In addition, 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, 23, 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 [59], and the primal-dual method [20] based on semi-smooth Newton algorithm [31], as well as alternating direction method [24].

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) [53], and much more effective than TV-L^2 in the case of Poisson noise removal [36]. Some existing methods for the TV-KL model are gradient descent [36, 44], multilevel method [13], the scaled gradient projection method [65], and EM-TV alternative minimization [6], as well as variable splitting and convex optimization based methods [25, 52].

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 more difficult 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, 57] 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. We should mention that our method is different from [24, 25, 52]. In [24, 25, 52], the authors treat the constraints (via operators) in a compact way and consequently penalty parameters for different constraints are the same. In our method we separately penalize the constraints and thus allow different penalty parameters for different constraints. Our numerical tests showed that more efficiency may be achieved by using different parameters. In addition, the splitting technique and auxiliary variables used in this work are different from those in [25, 52]. Moreover, by using different penalty parameters, the convergence analysis is more difficult than those in [24, 25, 52], as will be elaborated in Section 4.

The paper is organized as follows. In the next section, we give some notation. In Section 3, we present TV restoration model with a 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

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

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),

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

is 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 $p^2_{0,j} = p^2_{N,j}$. See also [57].

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

$$(1) \qquad\qquad\qquad u \xrightarrow{\text{blur}} Ku \xrightarrow{\text{noise}} f,$$

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. Image restoration aims at recovering u from f. Since the problem is usually ill-posed, we cannot directly solve u from (1). Regularization on the solution should be considered. One of the most basic and successful image restoration models is based on total variation regularization, which reads

$$(2) \qquad\qquad\qquad \min_{u \in V}\{E(u) = R(\nabla u) + F(Ku)\},$$

where

$$(3) \qquad\qquad R(\nabla u) = \mathrm{TV}(u) = \sum_{1 \leq i,j \leq N} |(\nabla u)_{i,j}|,$$

is the total variation of u [48], and $F(Ku)$ is a fidelity term.

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]):
$$F(Ku) = \begin{cases} \alpha \sum_{1 \leq i,j \leq N}((Ku)_{i,j} - f_{i,j}\log(Ku)_{i,j}), & (Ku)_{i,j} > 0, \forall\, i,j \\ +\infty, & \text{otherwise} \end{cases},$$

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 $\mathrm{dom}(F) = \{z \in V : F(z) < +\infty\}$ 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 (2) is convex, proper, coercive, and lower semi continuous. According to the generalized Weierstrass theorem and Fermat's rule [22, 27], we have the following result.

Theorem 3.1. *The problem* (2) *has at least one solution* u, *which satisfies*

$$(4) \qquad 0 \in K^*\partial F(Ku) - \mathrm{div}\partial R(\nabla u),$$

where $\partial F(Ku)$ *and* $\partial R(\nabla u)$ *are the sub-differentials* [22] *of* F *at* Ku *and* R *at* ∇u, *respectively. Moreover, if* $F \circ K(u)$ *is strictly convex, the minimizer is unique.*

So far many efficient algorithms have been proposed [14, 9, 11, 66, 67, 56, 58, 32, 62, 28, 54, 57] to solve total variation minimization with a quadratic fidelity. In the following we extend our recent work [54, 57] for total variation restoration with non-quadratic fidelities 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, 57], we need one more auxiliary variable to eliminate the nonlinearity for u.

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

$$(5) \qquad \begin{aligned} &\min_{u \in V, p \in Q, z \in V}\{G(p,z) = R(p) + F(z)\} \\ &\mathrm{s.t.} \qquad p = \nabla u, z = Ku \end{aligned}.$$

To solve (5), we define the following augmented Lagrangian functional

$$(6) \qquad \begin{aligned} &\mathscr{L}(u,p,z;\lambda_p,\lambda_z) \\ &= R(p) + F(z) + (\lambda_p, p - \nabla u) + (\lambda_z, z - Ku) + \frac{r_p}{2}\|p - \nabla u\|^2 + \frac{r_z}{2}\|z - Ku\|^2, \end{aligned}$$

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:

$$(7) \qquad \begin{aligned} &\text{Find} \quad (u^*, p^*, z^*; \lambda_p^*, \lambda_z^*) \in V \times Q \times V \times Q \times V, \\ &\text{s.t.} \quad \mathscr{L}(u^*, p^*, z^*; \lambda_p, \lambda_z) \leq \mathscr{L}(u^*, p^*, z^*; \lambda_p^*, \lambda_z^*) \leq \mathscr{L}(u, p, z; \lambda_p^*, \lambda_z^*), \\ &\quad\quad \forall (u, p, z; \lambda_p, \lambda_z) \in V \times Q \times V \times Q \times V. \end{aligned}$$

Note that, differently from [24, 25, 52], here it is no need to require $r_p = r_z$. According to our test, more efficiency can be achieved by allowing $r_p \neq r_z$. However, the convergence analysis when $r_p \neq r_z$ is more difficult than the case $r_p = r_z$; see Section 4.2 for details.

Similarly to [57], we can prove the following result.

Theorem 4.1. $u^* \in V$ *is a solution of* (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* (7).

Proof. We just provide a sketch since the idea is similar to that in [57].

Suppose $(u^*, p^*, z^*; \lambda_p^*, \lambda_z^*)$ is a solution of (7). From the first inequality in (7), we have

$$
(8) \qquad \begin{cases} p^* - \nabla u^* = 0, \\ z^* - K u^* = 0. \end{cases}
$$

The above relation, together with the second inequality in (7), indicates that u^* is a solution of (2).

Conversely, we assume that $u^* \in V$ is a solution of (2). We take $p^* = \nabla u^* \in Q$ and $z^* = K u^* \in V$. From (4), there exist λ_p^* and λ_z^* such that $-\lambda_p^* \in \partial R(\nabla u^*)$ and $-\lambda_z^* \in \partial F(K u^*)$ with $-K^* \lambda_z^* + \operatorname{div} \lambda_p^* = 0$. We can verify that $(u^*, p^*, z^*; \lambda_p^*, \lambda_z^*)$ is a saddle-point of $\mathscr{L}$, which completes the proof. $\qquad\square$

Theorems 3.1 and 4.1 show that the saddle-point problem (7) has at least one solution and any solution provides a solution of the original problem (2). In the following we present how to solve the saddle-point problem.

4.1. **An iterative algorithm for the saddle-point problem.** We use the following iterative algorithm to solve the saddle-point problem (7).

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

---

1. *Initialization*: $\lambda_p^0 = 0, \lambda_z^0 = 0, u^{-1} = 0, p^{-1} = 0, z^{-1} = 0$;
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.*,

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

   where $\mathscr{L}(u, p, z; \lambda_p^k, \lambda_z^k)$ is as in (6);
   (b) update

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

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 (9), it's difficult to solve them simultaneously. Therefore we separate the problem to three sub-problems and then apply an alternative minimization. The three sub-problems are as follows:

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

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

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

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

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

$$
(13) \qquad \min_{z \in V} \{ F(z) + (\lambda_z^k, z) + \frac{r_z}{2} \| z - K u \|^2 \}.
$$

Note here we omit the constant terms in the objective functionals in (11), (12) and (13).

In the following we show how to efficiently solve these sub-problems and then present an alternative minimization algorithm to solve (9).

4.1.1. *Solving the* $u-$*sub problem* (11). The problem (11) is a quadratic optimization problem, whose optimality condition reads

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

by considering the periodic boundary conditions. Following [56, 58, 59, 54, 57], 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

(14)
$$\begin{aligned}
&(r_z\mathcal{F}(K^*)\mathcal{F}(K) - r_p\mathcal{F}(\triangle))\mathcal{F}(u) \\
&= \mathcal{F}(K^*)(\mathcal{F}(\lambda_z^k) + r_z\mathcal{F}(z)) - \mathcal{F}(\mathring{D}_x^-)(\mathcal{F}((\lambda_p^1)^k) + r_p\mathcal{F}(p^1)), \\
&\quad - \mathcal{F}(\mathring{D}_y^-)(\mathcal{F}((\lambda_p^2)^k) + r_p\mathcal{F}(p^2))
\end{aligned}$$

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* (12). Similarly to [7, 56, 54, 57], (12) has the following closed form solution

(15)
$$p_{i,j} = \max(0, 1 - \frac{1}{r_p|\mathbf{w}_{i,j}|})\mathbf{w}_{i,j}, \quad \forall\, i, j$$

where

(16)
$$\mathbf{w} = \nabla u - \frac{\lambda_p^k}{r_p} \in Q.$$

Here we would like to provide a geometric interpretation of the formulae (15), which is different from the view point of sub-differential [56]. According to the definition of $R(p)$ and $\|\cdot\|_Q$, we rewrite the problem (12) as

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

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

(17)
$$\min_{q\in\mathbb{R}^2}\{|q| + \frac{r_p}{2}|q - w|^2\},$$

where $w \in \mathbb{R}^2$; see Fig. 1.

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. 1, we only need to consider those points located inside of the solid circle and in the first quadrant, e.g., q. For such a 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

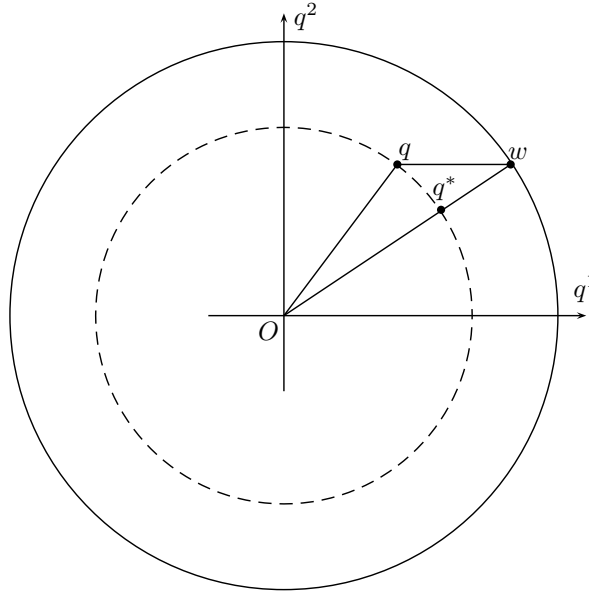$$|q| + |q - w| \ge |w| = |q^*| + |q^* - w|.$$

Figure 1. A geometric interpretation of the formulae (15)

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 (17) will locate on the line segment Ow. Denoting $q = \beta w$ with $0 \leq \beta \leq 1$, we hence simplify (17) to be the following 1-dimensional problem

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

The above (18) can be solved exactly, with a closed form solution as

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

The solution of (17) follows immediately.

   We further give two comments on this geometric interpretation. First, this observation (to solve (17)) can be extended to higher (> 2) dimensional problems as we did in [57] for vectorial and high order TV models. Second, the method can also be applied to problems with general regularization terms, say, general $R(q)$ (not only $|q|$) in (17), as long as the regularizer $R(q)$ depends only on $|q|$.

4.1.3. *Solving the $z-$sub problem* (13). For a general fidelity F, it is no reason to find a closed form solution for (13). Fortunately, the objective functional in (13) is strictly convex, proper, coercive and lower semi continuous. Therefore, (13) has a unique solution and can be obtained by various numerical optimization methods.

It should be mentioned that, for some special and typical (non-quadratic) fidelities, the $z-$sub problem still has closed form solution; see Section 5. Our method is therefore particularly efficient for these ones.

After knowing how to solve (11), (12) and (13), we present the following alternative minimization procedure to solve (9). It is with Gauss-Seidel flavor.

Algorithm 4.2 Augmented Lagrangian method for TV restoration with non-quadratic fidelity – solve the minimization problem (9)

- *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 (14) for $p = p^{k,l}, z = z^{k,l}$;
  − compute $p^{k,l+1}$ from (15) for $u = u^{k,l+1}$;
  − compute $z^{k,l+1}$ by solving (13) for $u = u^{k,l+1}$;
- $u^k = u^{k,L}, p^k = p^{k,L}, z^k = z^{k,L}$.

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, 54, 57]. 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 used to stop Algorithm 4.2.

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 (9) 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.

Our proof is motivated by the classic analysis techniques; see [27]. It should be pointed out that the techniques in [27, 57] cannot be straightforwardly applied to our case. The reason is that in general $r_p \neq r_z$, the monotonically decreasing sequences constructed in [27, 57] do not hold. Therefore, we need to construct two new monotonically decreasing sequences to derive the results. Consequently, other details of the analysis in [27, 57] should also be modified to fit the derivation based on the two new sequences.

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 (9) 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*

$$(19) \quad \begin{cases} \lim_{k \to \infty} (R(p^k) + F(z^k)) = R(p^*) + F(z^*) = E(u^*), \\ \lim_{k \to \infty} \|p^k - \nabla u^k\| = 0, \\ \lim_{k \to \infty} \|z^k - Ku^k\| = 0. \end{cases}$$

Moreover, (19) 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

$$(20) \qquad \begin{aligned} \mathscr{L}(u^*, p^*, z^*; \lambda_p, \lambda_z) \leq \mathscr{L}(u^*, p^*, z^*; \lambda_p^*, \lambda_z^*) \leq \mathscr{L}(u, p, z; \lambda_p^*, \lambda_z^*), \\ \forall (u, p, z; \lambda_p, \lambda_z) \in V \times Q \times V \times Q \times V. \end{aligned}$$

From the first inequality of (20), we have

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

This relationship, together with (10), 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

$$(21) \qquad \begin{cases} \sqrt{r_z}\overline{\lambda_p}^{k+1} = \sqrt{r_z}\overline{\lambda_p}^k + r_p\sqrt{r_z}(\overline{p}^k - \nabla\overline{u}^k), \\ \sqrt{r_p}\overline{\lambda_z}^{k+1} = \sqrt{r_p}\overline{\lambda_z}^k + r_z\sqrt{r_p}(\overline{z}^k - K\overline{u}^k). \end{cases}$$

The observation (21) is a key formula in our proof, and helps to construct a useful monotonically decreasing sequence, which is different from that in [27, 57].

It then follows that

$$(r_z\|\overline{\lambda_p}^k\|^2 + r_p\|\overline{\lambda_z}^k\|^2) - (r_z\|\overline{\lambda_p}^{k+1}\|^2 + r_p\|\overline{\lambda_z}^{k+1}\|^2)$$

$$(22) \qquad = -2r_p r_z(\overline{\lambda_p}^k, \overline{p}^k - \nabla\overline{u}^k) - r_p^2 r_z\|\overline{p}^k - \nabla\overline{u}^k\|^2$$

$$\qquad - 2r_p r_z(\overline{\lambda_z}^k, \overline{z}^k - K\overline{u}^k) - r_z^2 r_p\|\overline{z}^k - K\overline{u}^k\|^2.$$

In the following we show that the right hand side of (22) 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 (20), (u^*, p^*, z^*) is characterized by

$$(23) \qquad \begin{aligned} (\operatorname{div}\lambda_p^*, u - u^*) + r_p(\operatorname{div}(p^* - \nabla u^*), u - u^*) \\ + (\lambda_z^*, -K(u - u^*)) + r_z(z^* - Ku^*, -K(u - u^*)) \geq 0, \forall u \in V, \end{aligned}$$

$$(24) \qquad R(p) - R(p^*) + (\lambda_p^*, p - p^*) + r_p(p^* - \nabla u^*, p - p^*) \geq 0, \forall p \in Q,$$

$$(25) \qquad F(z) - F(z^*) + (\lambda_z^*, z - z^*) + r_z(z^* - Ku^*, z - z^*) \geq 0, \forall z \in V.$$

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

$$(26) \qquad \begin{aligned} (\operatorname{div}\lambda_p^k, u - u^k) + r_p(\operatorname{div}(p^k - \nabla u^k), u - u^k) \\ + (\lambda_z^k, -K(u - u^k)) + r_z(z^k - Ku^k, -K(u - u^k)) \geq 0, \forall u \in V, \end{aligned}$$

$$(27) \qquad R(p) - R(p^k) + (\lambda_p^k, p - p^k) + r_p(p^k - \nabla u^k, p - p^k) \geq 0, \forall p \in Q,$$

$$(28) \qquad F(z) - F(z^k) + (\lambda_z^k, z - z^k) + r_z(z^k - Ku^k, z - z^k) \geq 0, \forall z \in V,$$

since (u^k, p^k, z^k) is the solution of (9). Taking $u = u^k$ in (23), $u = u^*$ in (26), $p = p^k$ in (24), $p = p^*$ in (27), $z = z^k$ in (25), and $z = z^*$ in (28), respectively, we obtain, by addition

$$(29) \qquad -(\overline{\lambda_p}^k, \overline{p}^k - \nabla\overline{u}^k) - (\overline{\lambda_z}^k, \overline{z}^k - K\overline{u}^k) \geq r_p\|\overline{p}^k - \nabla\overline{u}^k\|^2 + r_z\|\overline{z}^k - K\overline{u}^k\|^2,$$

which is equivalent to

$$(30) \quad -r_p r_z(\overline{\lambda_p}^k, \overline{p}^k - \nabla\overline{u}^k) - r_p r_z(\overline{\lambda_z}^k, \overline{z}^k - K\overline{u}^k) \geq r_p^2 r_z\|\overline{p}^k - \nabla\overline{u}^k\|^2 + r_p r_z^2\|\overline{z}^k - K\overline{u}^k\|^2.$$

From (22) and (30), we have

$$(31) \quad \begin{aligned} &(r_z\|\overline{\lambda_p}^k\|^2 + r_p\|\overline{\lambda_z}^k\|^2) - (r_z\|\overline{\lambda_p}^{k+1}\|^2 + r_p\|\overline{\lambda_z}^{k+1}\|^2) \\ &\geq r_p^2 r_z\|\overline{p}^k - \nabla\overline{u}^k\|^2 + r_p r_z^2\|\overline{z}^k - K\overline{u}^k\|^2, \end{aligned}$$

which indicates

$$(32) \quad \begin{cases} \{\lambda_p^k : \forall k\} \text{ and } \{\lambda_z^k : \forall k\} \quad \text{are bounded,} \\ \lim_{k\to\infty}\|p^k - \nabla u^k\| = 0, \\ \lim_{k\to\infty}\|z^k - Ku^k\| = 0. \end{cases}$$

On the other hand, the second inequality of (20) implies

$$(33) \quad \begin{aligned} R(p^*) + F(z^*) \leq &R(p^k) + F(z^k) + (\lambda_p^*, p^k - \nabla u^k) + (\lambda_z^*, z^k - Ku^k) \\ &+ \frac{r_p}{2}\|p^k - \nabla u^k\|^2 + \frac{r_z}{2}\|z^k - Ku^k\|^2. \end{aligned}$$

If we take $u = u^*$ in (26), $p = p^*$ in (27), and $z = z^*$ in (28), we have, by addition,

$$(34) \quad \begin{aligned} R(p^*) + F(z^*) \geq &R(p^k) + F(z^k) + (\lambda_p^k, p^k - \nabla u^k) + (\lambda_z^k, z^k - Ku^k) \\ &+ r_p\|p^k - \nabla u^k\|^2 + r_z\|z^k - Ku^k\|^2. \end{aligned}$$

Using (32), we have

$$(35) \quad \liminf(R(p^k) + F(z^k)) \geq R(p^*) + F(z^*) \geq \limsup(R(p^k) + F(z^k)),$$

by taking lim inf in (33) and lim sup in (34). Hence we complete the proof of (19).

Since $R(\cdot)$ and $F(\cdot)$ are both continuous over their domains, (19) implies clearly that u^k is a minimizing sequence of $E(\cdot)$. If the minimizer of $E(\cdot)$ is unique, then $u^k \to u^*$. $\square$

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 (9) 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*

$$(36) \quad \begin{cases} \lim_{k\to\infty}(R(p^k) + F(z^k)) = R(p^*) + F(z^*) = E(u^*), \\ \lim_{k\to\infty}\|p^k - \nabla u^k\| = 0, \\ \lim_{k\to\infty}\|z^k - Ku^k\| = 0. \end{cases}$$

Moreover, (36) 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, (22) still holds, which is represented as follows

$$(37) \quad \begin{aligned} &(r_z\|\overline{\lambda_p}^k\|^2 + r_p\|\overline{\lambda_z}^k\|^2) - (r_z\|\overline{\lambda_p}^{k+1}\|^2 + r_p\|\overline{\lambda_z}^{k+1}\|^2) \\ &= -2r_p r_z(\overline{\lambda_p}^k, \overline{p}^k - \nabla\overline{u}^k) - r_p^2 r_z\|\overline{p}^k - \nabla\overline{u}^k\|^2 \\ &\quad - 2r_p r_z(\overline{\lambda_z}^k, \overline{z}^k - K\overline{u}^k) - r_z^2 r_p\|\overline{z}^k - K\overline{u}^k\|^2. \end{aligned}$$

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

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

$$
(39) \qquad R(p) - R(p^*) + (\lambda_p^*, p - p^*) + r_p(p^* - \nabla u^*, p - p^*) \geq 0, \forall p \in Q,
$$

$$
(40) \qquad F(z) - F(z^*) + (\lambda_z^*, z - z^*) + r_z(z^* - Ku^*, z - z^*) \geq 0, \forall z \in V.
$$

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

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

$$
(42) \qquad R(p) - R(p^k) + (\lambda_p^k, p - p^k) + r_p(p^k - \nabla u^k, p - p^k) \geq 0, \forall p \in Q,
$$

$$
(43) \qquad F(z) - F(z^k) + (\lambda_z^k, z - z^k) + r_z(z^k - Ku^k, z - z^k) \geq 0, \forall z \in V,
$$

Taking $u = u^k$ in (38), $u = u^*$ in (41), $p = p^k$ in (39), $p = p^*$ in (42), $z = z^k$ in (40), and $z = z^*$ in (43), respectively, we obtain, after addition

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

(37) and (44) indicate

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

On the other hand, we have, by using the same technique as in [27, 57], the following estimates

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

We then obtain, from (45) and (46),

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

which implies
(48)

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

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

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

If we take $u = u^*$ in (41), $p = p^*$ in (42), and $z = z^*$ in (43), we have, by addition,

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

Using (48), we have

$$
(51) \qquad \liminf(R(p^k) + F(z^k)) \geq R(p^*) + F(z^*) \geq \limsup(R(p^k) + F(z^k)),
$$

by taking $\liminf$ in (49) and $\limsup$ in (50). This completes the proof of (36).

By the continuity of $R(\cdot)$ and $F(\cdot)$ over their domains, (36) indicates clearly that u^k is a minimizing sequence of $E(\cdot)$. If the minimizer of $E(\cdot)$ is unique, then $u^k \to u^*$. $\qquad\square$

We would like to add a comment on Theorem 4.3. It is stated in [63] 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 [63]. As a consequence, setting $L = 1$ provides a simple stopping criterion of the inner iteration, which does not need to compute those optimality conditions and thus reduces the CPU cost.

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. Numerical examples will also be provided.

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:

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

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

The problem (52) is a special case of (2) where the fidelity term is

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

Therefore we can apply Algorithms 4.1 and 4.2 to solve (52). For this special fidelity, we have the following explicit solution for the $z-$sub problem (13):

$$(54) \qquad z_{i,j} = f_{i,j} + \max(0, 1 - \frac{\alpha}{r_z|w_{i,j} - f_{i,j}|})(w_{i,j} - f_{i,j}),$$

where

$$(55) \qquad w = Ku - \frac{\lambda_z^k}{r_z} \in V.$$

The derivation of (54) is similar to (15) 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 (9)

- *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 (14) for $p = p^{k,l}, z = z^{k,l}$;
  - compute $p^{k,l+1}$ from (15) for $u = u^{k,l+1}$;
  - compute $z^{k,l+1}$ from (54) 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* (9).

Proof. The proof is motivated by [56, 59] and similar to that of Theorem 4.2 in [57]. Here we just sketch the differences.

Similarly with s_τ and s in [56, 59], 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 (54), 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 (54) and (15) can be reformulated as $z = S_1(w - f)$ and $p = S_2(\mathbf{w})$, respectively, with w defined in (55) and $\mathbf{w}$ in (16). Therefore the iterative scheme in Algorithm 5.1 can be written as

$$(56) \qquad \begin{cases} u^{k,l+1} = (r_z K^* K + r_p \nabla^* \nabla)^{-1} (K^*(\lambda_z^k + r_z z^{k,l}) + \nabla^*(\lambda_p^k + r_p p^{k,l})), \\ p^{k,l+1} = S_2(\nabla u^{k,l+1} - \frac{\lambda_p^k}{r_p}), \\ z^{k,l+1} = S_1(Ku^{k,l+1} - \frac{\lambda_z^k}{r_z} - f), \end{cases}$$

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:

(57)
$$\begin{cases} h_2(p,z) = \nabla(r_z K^* K + r_p \nabla^* \nabla)^{-1}(K^*(\lambda_z^k + r_z z) + \nabla^*(\lambda_p^k + r_p p)) - \frac{\lambda_p^k}{r_p}, \\ h_1(p,z) = K(r_z K^* K + r_p \nabla^* \nabla)^{-1}(K^*(\lambda_z^k + r_z z) + \nabla^*(\lambda_p^k + r_p p)) - \frac{\lambda_z^k}{r_z} - f. \end{cases}$$

The non-expansiveness of $\hat{h} : Q \times V \to Q \times V$, which is given by $\hat{h}(p,z) = (h_2(p,z), h_1(p,z))$, can be similarly verified as in [56, 59].

Rewriting the iterative scheme (56) as

(58)
$$\begin{cases} u^{k,l+1} = (r_z K^* K + r_p \nabla^* \nabla)^{-1}(K^*(\lambda_z^k + r_z z^{k,l}) + \nabla^*(\lambda_p^k + r_p p^{k,l})), \\ (p^{k,l+1}, z^{k,l+1}) = (S_2 \circ h_2; S_1 \circ h_1)(p^{k,l}, z^{k,l}), \end{cases}$$

one can show the convergence via a similar argument in [56, 59]. $\square$

Here we show some examples. In Fig. 2, and 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 Fig. 4. In Fig. 5 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 [59]. When this paper was nearly finished, we got to know that the Rice L1 group 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. This advantage becomes more and more obvious when the noise level gets higher and the image size gets larger. 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 [54, 51, 57] 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. FTVd version 4.0 uses alternating direction method (ADM) with a similar (not the same) variable splitting technique with ours. 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. We finally mention that, very recently we got to know FTVd version 4.1 released, which improves the efficiency of the previous version by replacing the Matlab built-in call "imfilter" for image convolution with FFT implementation. This replacement and improvement are worthy of being implemented in our method.

In Fig. 6 and 7, we show two examples calculated via augmented Lagrangian method with different parameters. One may find that better efficiency can be achieved by allowing $r_p \neq r_z$.
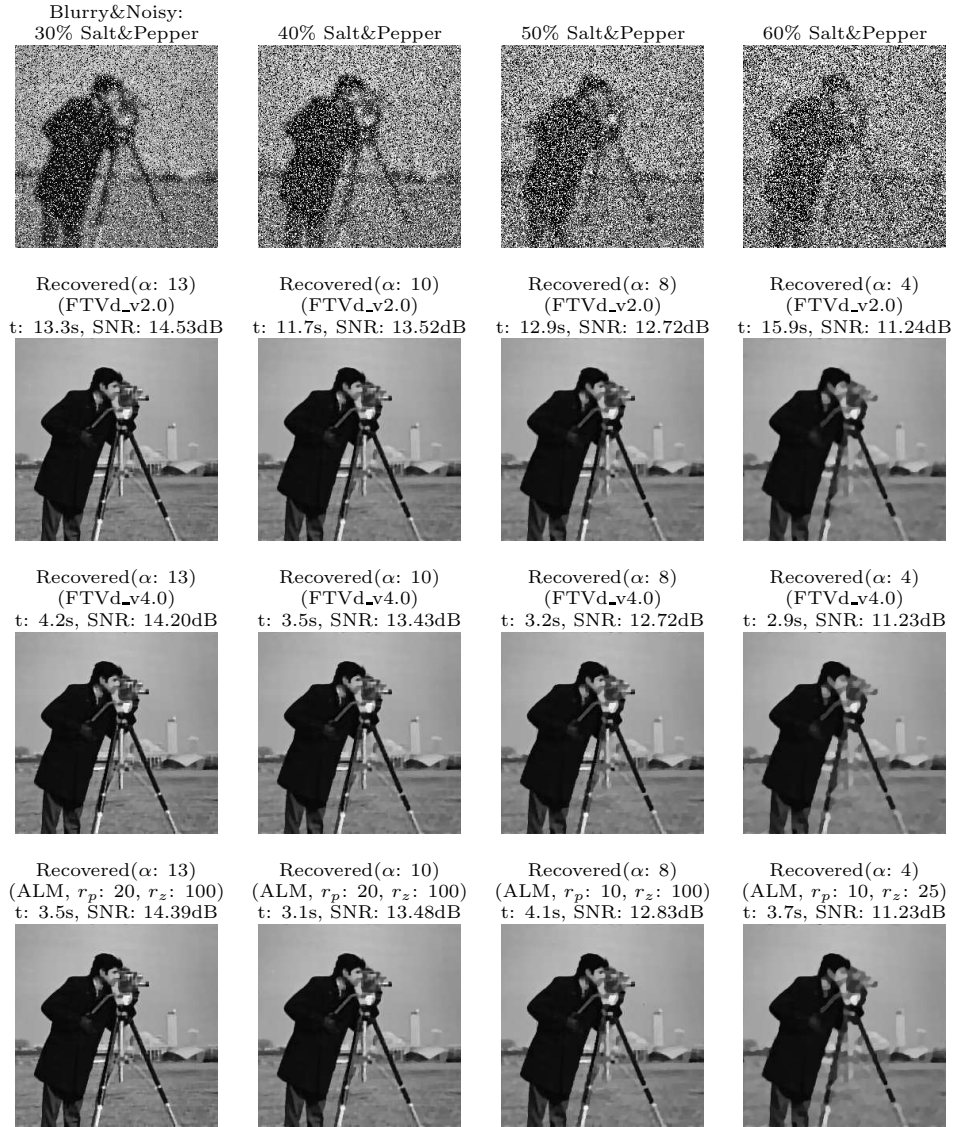
Blurry&Noisy: 30% Salt&Pepper	40% Salt&Pepper	50% Salt&Pepper	60% Salt&Pepper

Recovered(α: 13) (FTVd_v2.0) t: 13.3s, SNR: 14.53dB	Recovered(α: 10) (FTVd_v2.0) t: 11.7s, SNR: 13.52dB	Recovered(α: 8) (FTVd_v2.0) t: 12.9s, SNR: 12.72dB	Recovered(α: 4) (FTVd_v2.0) t: 15.9s, SNR: 11.24dB

Recovered(α: 13) (FTVd_v4.0) t: 4.2s, SNR: 14.20dB	Recovered(α: 10) (FTVd_v4.0) t: 3.5s, SNR: 13.43dB	Recovered(α: 8) (FTVd_v4.0) t: 3.2s, SNR: 12.72dB	Recovered(α: 4) (FTVd_v4.0) t: 2.9s, SNR: 11.23dB

Recovered(α: 13) (ALM, r_p: 20, r_z: 100) t: 3.5s, SNR: 14.39dB	Recovered(α: 10) (ALM, r_p: 20, r_z: 100) t: 3.1s, SNR: 13.48dB	Recovered(α: 8) (ALM, r_p: 10, r_z: 100) t: 4.1s, SNR: 12.83dB	Recovered(α: 4) (ALM, r_p: 10, r_z: 25) t: 3.7s, SNR: 11.23dB

FIGURE 2. TV-L^1 restoration from 7×7 sized Gaussian blur with salt-and-pepper noise from 30% to 60% (image size 256×256).

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

FIGURE 3. TV-L^1 restoration from 7×7 sized Gaussian blur with salt-and-pepper noise from 30% to 60% (image size 512×512).

particular, we consider the following minimization problem:

(59)
$$\min_{u \in V} \{E_{\text{TVKL}}(u) = R(\nabla u) + \alpha \sum_{1 \leq i,j \leq N} ((Ku)_{i,j} - f_{i,j} \log (Ku)_{i,j}) : (Ku)_{i,j} > 0, \forall\, i,j\},$$

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

The problem (59) is a special case of (2) where

(60) $$F(Ku) = \begin{cases} \alpha \sum_{1 \leq i,j \leq N} ((Ku)_{i,j} - f_{i,j} \log (Ku)_{i,j}), & (Ku)_{i,j} > 0, \ \forall\, i,j \\ +\infty, & \text{otherwise} \end{cases}.$$

FIGURE 4. TV-L^1 restoration from 7×7 sized Gaussian blur with random-valued noise from 20% to 50% (image size 256×256).

Therefore, Algorithms 4.1 and 4.2 can be applied to compute (59). For this special fidelity, we also have, by considering $z_{i,j} > 0$, a closed form solution to the $z-$sub problem (13):

$$(61) \qquad z_{i,j} = \frac{1}{2}\left(\sqrt{(w_{i,j} - \frac{\alpha}{r_z})^2 + 4\frac{\alpha}{r_z}f_{i,j}} + (w_{i,j} - \frac{\alpha}{r_z})\right),$$

where
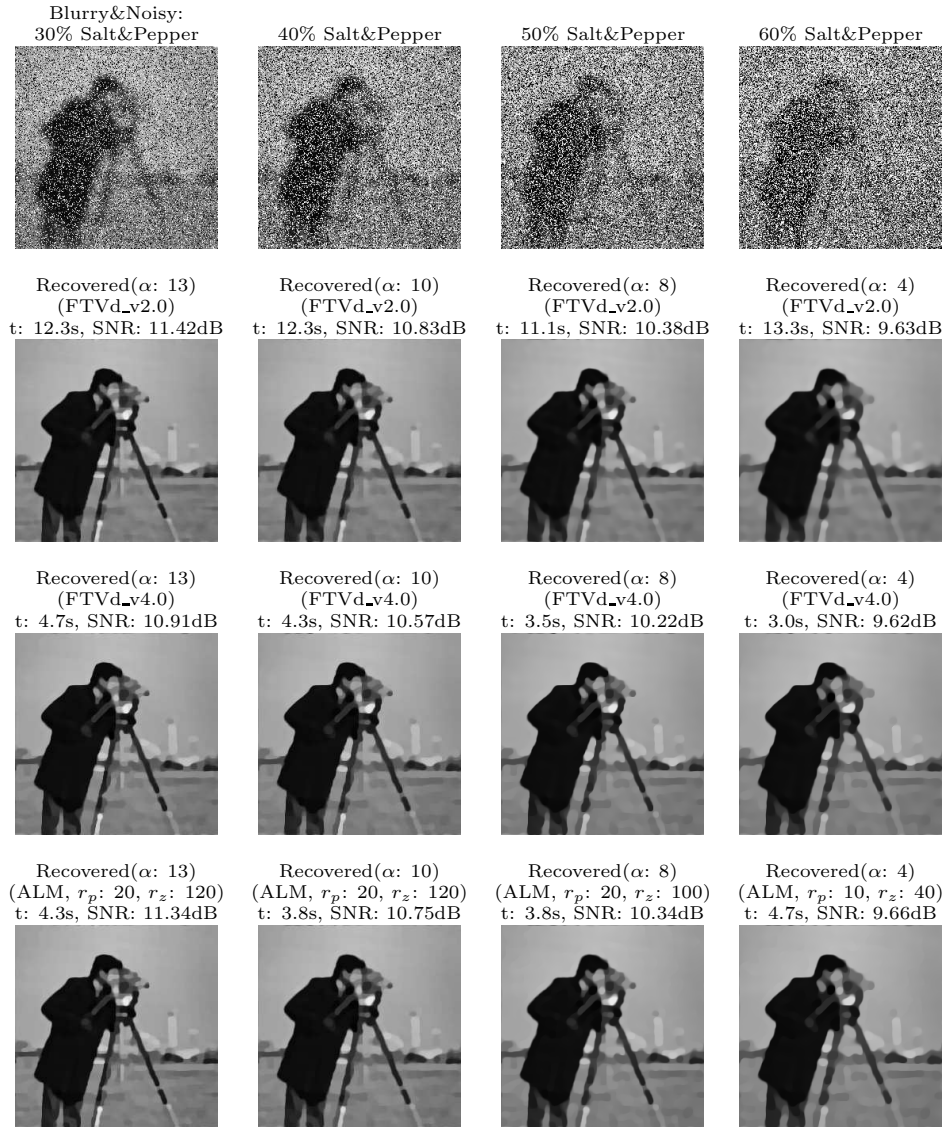
$$(62) \qquad w = Ku - \frac{\lambda_z^k}{r_z} \in V.$$

FIGURE 5.   TV-L^1 restoration from 15×15 sized Gaussian blur
with salt-and-pepper noise from 30% to 60% (image size 256×256).

Here we elaborate Algorithm 4.2 for TV-KL restoration as Algorithm 5.2.
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* (9).

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 $\Psi = (\psi_{i,j}) : V \rightarrow$

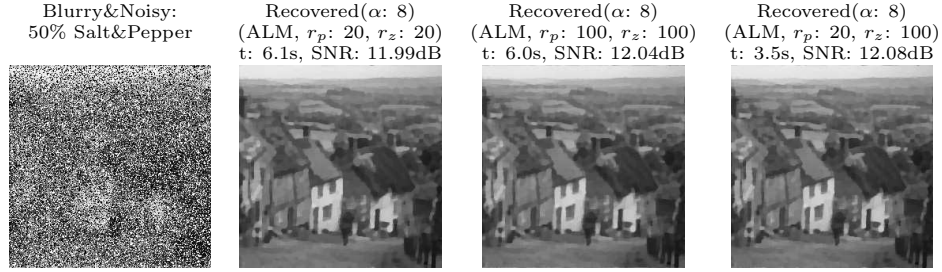| Blurry&Noisy: 50% Salt&Pepper	Recovered(α: 8) (ALM, r_p: 20, r_z: 20) t: 6.1s, SNR: 11.99dB	Recovered(α: 8) (ALM, r_p: 100, r_z: 100) t: 6.0s, SNR: 12.04dB	Recovered(α: 8) (ALM, r_p: 20, r_z: 100) t: 3.5s, SNR: 12.08dB

FIGURE 6. TV-L^1 restoration from 7×7 sized Gaussian blur with 50% salt-and-pepper noise: A better computational efficiency with comparable restoration result (similar SNRs) is achieved by letting $r_p \neq r_z$.
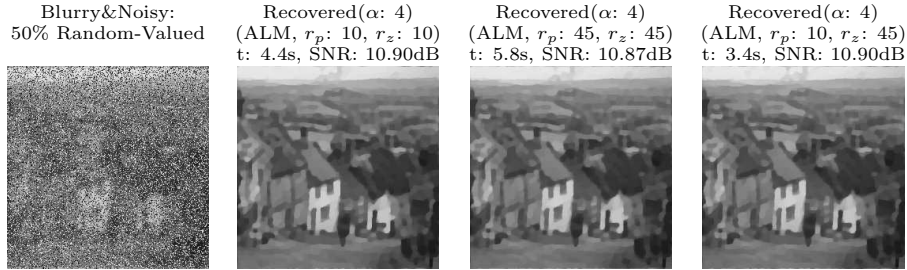
Blurry&Noisy: 50% Random-Valued	Recovered(α: 4) (ALM, r_p: 10, r_z: 10) t: 4.4s, SNR: 10.90dB	Recovered(α: 4) (ALM, r_p: 45, r_z: 45) t: 5.8s, SNR: 10.87dB	Recovered(α: 4) (ALM, r_p: 10, r_z: 45) t: 3.4s, SNR: 10.90dB

FIGURE 7. TV-L^1 restoration from 7×7 sized Gaussian blur with 50% random-valued noise: A better computational efficiency with comparable restoration result (similar SNRs) is achieved by letting $r_p \neq r_z$.

Algorithm 5.2 Augmented Lagrangian method for TV-KL restoration – solve the minimization problem (9)

- *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 (14) for $p = p^{k,l}, z = z^{k,l}$;
  - compute $p^{k,l+1}$ from (15) for $u = u^{k,l+1}$;
  - compute $z^{k,l+1}$ from (61) for $u = u^{k,l+1}$;
- $u^k = u^{k,L}, p^k = p^{k,L}, z^k = z^{k,L}$.

V, according to (61), with $\psi_{i,j}$ as

$$(63) \qquad \psi_{i,j}(t) = \frac{1}{2}(\sqrt{t^2 + 4\frac{\alpha}{r_z}f_{i,j}} + t).$$

In the following we prove the convergence in three steps.

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

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

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 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,

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

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 Fig. 8, 9 and 10. In each figure, α, t, and SNR denote the parameter of the model, CPU cost (in seconds) and signal-noise ratio, 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 (with higher SNR) than TV-L^2 model. See also, the blurry effect of TV-L^2 in Fig. 8 and 9. In addition, TV-KL model can still be calculated very efficiently by augmented Lagrangian method. When allowing $r_p \neq r_z$, a better computational efficiency can be achieved in TV-KL restoration (See Fig. 10), as in TV-L^1 restoration.
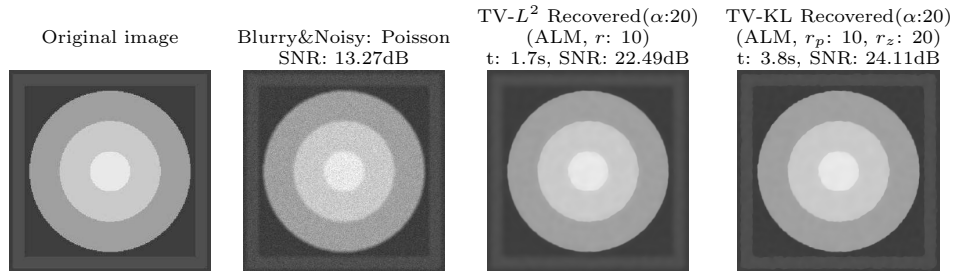

| Original image | Blurry&Noisy: Poisson SNR: 13.27dB | TV-L^2 Recovered(α:20) (ALM, r: 10) t: 1.7s, SNR: 22.49dB | TV-KL Recovered(α:20) (ALM, r_p: 10, r_z: 20) t: 3.8s, SNR: 24.11dB |

Figure 8. Comparisons between TV-L^2 and TV-KL restoration: recovering degraded images with 7×7 sized Gaussian blur and Poisson noise (image size 348×348).

We give some remarks on our numerical tests. All the experiments were performed under Windows XP and MATLAB R2007a running on a PC with Intel Core 2.66GHz CPU and 2GB RAM. According to Theorem 4.3, we used $\|p^k - \nabla u^k\|^2 + \|z^k - Ku^k\|^2 < \epsilon$ for some $\epsilon > 0$ (which means small Lagrange multipliers
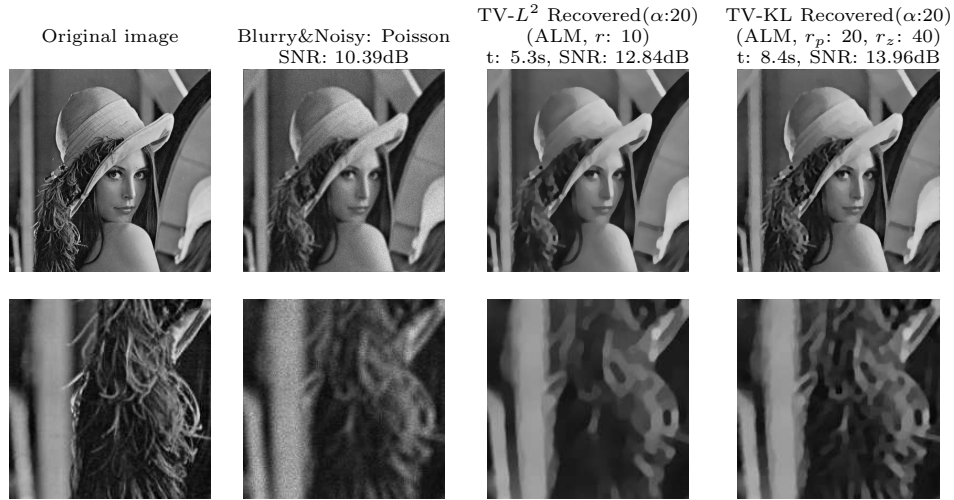
Original image   Blurry&Noisy: Poisson SNR: 10.39dB   TV-L^2 Recovered(α:20) (ALM, r: 10) t: 5.3s, SNR: 12.84dB   TV-KL Recovered(α:20) (ALM, r_p: 20, r_z: 40) t: 8.4s, SNR: 13.96dB

FIGURE 9. 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 size 512×512).
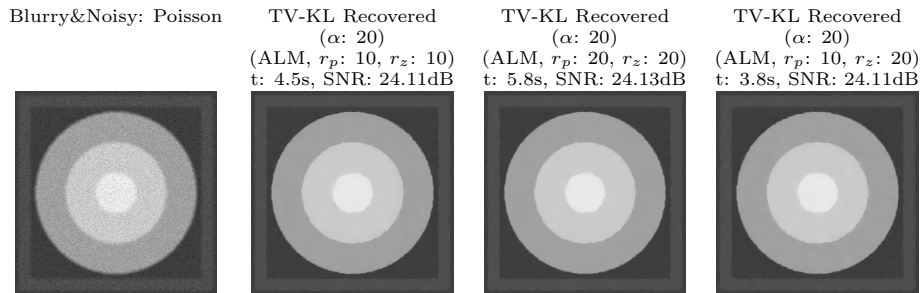
Blurry&Noisy: Poisson   TV-KL Recovered (α: 20) (ALM, r_p: 10, r_z: 10) t: 4.5s, SNR: 24.11dB   TV-KL Recovered (α: 20) (ALM, r_p: 20, r_z: 20) t: 5.8s, SNR: 24.13dB   TV-KL Recovered (α: 20) (ALM, r_p: 10, r_z: 20) t: 3.8s, SNR: 24.11dB

FIGURE 10. TV-KL restoration from 7×7 sized Gaussian blur with Poisson noise: A better computational efficiency with comparable restoration result (similar SNRs) is achieved by letting $r_p \neq r_z$.

update and thus the minimization problem (9) changes little) as stopping condition of Algorithm 4.1 in our tests.

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. Once the fidelity is specified, one need only design a solver for the $z-$sub problem. Due to FFT implementation or closed form solutions for the sub-problems, as well as simple stopping criterion ($L = 1$) of the inner iteration, the proposed algorithms are extremely efficient as demonstrated by the experiments. Since it is no need to compute residuals (defined by some norms of some data with image size

related dimensions, as in FTVd version 2.0) to stop inner iteration, our method exhibits even greater advantage for large sized images. Experiments also showed that, by allowing $r_p \neq r_z$ in the proposed algorithms, a better computational efficiency can be achieved. Moreover, we gave convergence analysis for the proposed algorithms, by a sophisticated modification of previous analysis techniques. A possible future work is to extend the method to color image recovering via TV (and even Non-Local TV) restoration with non-quadratic fidelities.

REFERENCES

[1] S. Alliney, *Digital filters as absolute norm regularizers*, IEEE Trans. Signal Process., **40** (1992), 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), 209–237.

[3] P. Blomgren and T. F. Chan, *Color TV: Total variation methods for restoration of vector–valued images*, IEEE Trans. Image Process., **7** (1998), 304–309.

[4] A. Bovik, "Handbook of Image and Video Processing," Academic Press, 2000.

[5] X. Bresson and T. F. Chan, *Fast dual minimization of the vectorial total variation norm and applications to color image processing*, Inverse Problems and Imaging, **2** (2008), 455–484.

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

[7] A. Caboussat, R. Glowinski and V. Pons, *An augmented Lagrangian approach to the numerical solution of a non-smooth eigenvalue problem*, J. Numer. Math., **17** (2009), 3–26.

[8] E. Candes, J. Romberg and T. Tao, *Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information*, IEEE Trans. Inform. Theory, **52** (2006), 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), 167–188.

[11] A. Chambolle, *An algorithm for total variation minimization and applications*, J. Math. Imaging Vision, **20** (2004), 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), 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), 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), 1964–1977.

[15] T. Chan, A. Marquina and P. Mulet, *High-order total variation-based image restoration*, SIAM J. Sci. Comput., **22** (2000), 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., **12** (2001), 422–435.

[17] T. F. Chan and S. Esedoglu, *Aspects of total variation regularized L^1 function approximation*, SIAM J. Appl. Math., **65** (2005), 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), 784–789.

[20] Y. Dong, M. Hintermüller and M. Neri, *An efficient primal-dual method for $L^1 TV$ image restoration*, SIAM J. Imaging Sciences, **2** (2009), 1168–1189.

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

[22] I. Ekeland and R. Témam, "Convex Analysis and Variational Problems," SIAM, 1999.

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

[24] E. Esser, *Applications of Lagrangian-based alternating direction methods and connections to split Bregman*, UCLA CAM Report, cam09-31. Available from: ftp://ftp.math.ucla.edu/pub/camreport/cam09-31.pdf.

[25] M. A. T. Figueiredo and J. M. Bioucas-Dias, *Deconvolution of Poissonian images using variable splitting and augmented Lagrangian optimization*, in "IEEE Workshop on Statistical Signal Processing," Cardiff, (2009), 733–736.

[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), 1881–1902.

[27] R. Glowinski and 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 J. Imaging Sciences, **2** (2009), 323–343.

[29] M. R. Hestenes, *Multiplier and gradient methods*, J. Optim. Theory and Appl., **4** (1969), 303–320.

[30] W. Hinterberger and O. Scherzer, *Variational methods on the space of functions of bounded Hessian for convexification and denoising*, Computing, **76** (2006), 109–133.

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

[32] Y. Huang, M. Ng and Y. Wen, *A fast total variation minimization method for image restoration*, SIAM Multi. Model. Simul., **7** (2008), 774–795.

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

[34] C. Kervrann and A. Trubuil, *An adaptive window approach for poisson noise reduction and structure preserving in confocal microscopy*, in "Proc. International Symposium on Biomedical Imaging (ISBI'04)," Arlington, (2004), 788–791.

[35] E. Kolaczyk, *Wavelet shrinkage estimation of certain Poisson intensity signals using corrected thresholds*, Statist. Sinica, **9** (1999), 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), 257–263.

[37] Y. Li and S. Osher, *A new median formula with applications to PDE based denoising*, Commun. Math. Sci., **7** (2009), 741–753.

[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), 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," Springer, (2006), 335–352.

[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), 1506–1516.

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

[43] M. Nikolova, *A variational approach to remove outliers and impulse noise*, J. Math. Imaging Vision, **20** (2004), 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), 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), 85–92.

[46] M. J. D. Powell, *A method for nonlinear constraints in minimization problems*, in "Optimization"(ed. R. Fletcher), Academic Press, (1972), 283–298.

[47] R. T. Rockafellar, *A dual approach to solving nonlinear programming problems by unconstrained optimization*, Mathematical Programming, **5** (1973), 354–373.

[48] L. Rudin, S. Osher and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Physica D, **60** (1992), 259–268.

[49] G. Sapiro and D. L. Ringach, *Anisotropic diffusion of multivalued images with applications to color filtering*, IEEE Trans. Image Process., **5** (1996), 1582–1586.

[50] O. Scherer, *Denoising with higher order derivatives of bounded variation and an application to parameter estimation*, Computing, **60** (1998), 1–27.

[51] S. Setzer, *Split Bregman algorithm, Douglas-Rachford splitting and frame shrinkage*, LNCS, **5567** (2009), 464–476.

[52] S. Setzer, G. Steidl and T. Teuber, *Deblurring Poissonian images by split Bregman techniques*, J. Visual Commun. Image Repres., accepted (2009).

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

[54] X. C. Tai and C. L. Wu, *Augmented Lagrangian method, dual methods and split Bregman iteration for ROF model*, LNCS, **5567** (2009), 502–513.

[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), 846–852.

[56] Y. Wang, J. Yang, W. Yin and Y. Zhang, *A new alternating minimization algorithm for total variation image reconstruction*, SIAM J. Imaging Sciences, **1** (2008), 248–272.

[57] C. L. Wu and X. C. Tai, *Augmented Lagrangian method, dual methods, and split Bregman iteration for ROF, vectorial TV, and high order models*, SIAM J. Imaging Sciences, **3** (2010), 300–339.

[58] J. F. Yang, W. T. Yin, Y. Zhang and Y. L. Wang, *A fast algorithm for edge-preserving variational multichannel image restoration*, SIAM J. Imaging Sciences, **2** (2009), 569–592.

[59] 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), 2842–2865.

[60] W. Yin, D. Goldfarb and S. Osher, *Image cartoon-texture decomposition and feature selection using the total variation regularized L^1 functional*, LNCS, **3752** (2005), 73–84.

[61] W. Yin, D. Goldfarb and S. Osher, *The total variation regularized L^1 model for multiscale decomposition*, Multi. Model. Simul., **6** (2007), 190–211.

[62] 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), 143–168.

[63] W. T. Yin, *Analysis and generalizations of the linearized Bregman method*, SIAM J. Imaging Sciences, **3** (2010), 856–877.

[64] Y.-L. You and M. Kaveh, *Fourth-order partial differential equation for noise removal*, IEEE Trans. Image Process., **9** (2000), 1723–1730.

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

[66] M. Zhu and T. F. Chan, *An efficient primal-dual hybrid gradient algorithm for total variation image restoration*, UCLA CAM Report, cam08-34. Available from: `ftp://ftp.math.ucla.edu/pub/camreport/cam08-34.pdf`.

[67] M. Zhu, S. J. Wright and T. F. Chan, *Duality-based algorithms for total variation image restoration*, Comput. Optim. Appl., (2008). Available from: `http://www.springerlink.com/content/l2k84n130x6l4332/`.

E-mail address: wucl@ustc.edu.cn

E-mail address: S070051@e.ntu.edu.sg

E-mail address: tai@cma.uio.no