Article

Research on Restoration Algorithm of Discrete Tone Image Based on Noise Model

Yichuan Dong[1*], Yu Feng[2], Yuanlei Chen[3], Narjes Nabipour[4*]

[1]School of Mathematics and Statistics, Wuhan University, Wuhan 430072,China

[2] Institute, China Merchants Bank, Shenzhen 518040,China

[3]Department of High Performance Computing in National Supercomputing center in, Shenzhen 518055,China

[4] Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam

*Corresponding authors: E-mails: ycdong@whu.edu.cn and narjesnabipour@duytan.edu.vn

Abstract: The ℓ_0 "norm" regularized variational model is widely used in signal/image restoration. It is suitable for sparsity based regularization model design. This paper proposes a wavelet frame based Poisson noise removal model is proposed and the alternating direction method of multipliers (ADMM) is adopted to solve the corresponding optimization problem. Due to the nonconvex and noncontinuous property of the objective function with ℓ_0 term, plenty of results in the literature can not be directly applied to verifying the convergence of this ADMM based algorithm. We established the convergence property of the proposed algorithm for Poisson noise image restoration and reconstruction. This theoretical foundation is valuable to guarantee the stability performance of the algorithm in numerical simulation. Numerical experiments on different image processing tasks, such as grey image deblurring, CT image reconstruction and graph data denoising, also verified the effectiveness of this newly proposed nonconvex and nonsmooth model. Moreover all the results are possible to be extended to more general nonconvex noncontinuous optimization problem.

Keywords: Poisson noise; Tight frame; Convergence analysis; ADMM; Nonconvex model; CT image reconstruction.

1 Introduction

Poisson noise contaminated images have commonly appeared in many application fields, e.g., astronomy, medical imaging, microscopy, positron emission tomography (PET) [33] and Bioluminescence Tomography (BLT) [40]. In this paper, we will analyze the convergence properties of an image restoration model arises from the imaging model with Poisson noise.

Most of the modern imaging devices are designed to capture images by counting the electric photons hit to the sensor array, which result in the signal-dependent Poisson noise. While the widely discussed additive white Gaussian noise is mainly used to model the readout noise, thermal noise, and the measure error with zero mean and o standard variance.

The most existing imaging models are assumed to contain Gaussian white noise. It is described by a quadratic data fidelity term and some priori information is incorporated into the image restoration model, such as ℓ_1, $\ell_p\left(0<p<1\right)$, ℓ_0 Total Variation TV, analysis/synthesis based tight wavelet frame [7, 11, 10, 41] and hybrid regularization[28]. Plenty of efficient algorithms have been developed to solve the Gaussian noise restoration problem.

Similarly, the Poisson noise image restoration can be done as the Gaussian noise except that the data fidelity term is replaced by the likelihood function of Poisson distribution. More specifically, the

TV-based models [17, 31], wavelet frame based modelsl2, 9l and variance stabilizing transform (VST) based models [20, 19] have been proposed in recently years for Poisson noise image processing [34]. Most of these proposed models are belong to convex optimization problem. However, numerical experiments show that the nonconvex regularization term is more suitable to preserve the sharp edges and details in the restored image [1, 8, 27]. Recently, [39] proposed a l_0 norm regularized nonconvex model for Poisson noise removal. The algorithm for this nonconvex model was designed with an alternating direction method. It outperforms the tight wavelet frame based ℓ_1 and high order TV based models. However, the convergence analysis of the ADMM scheme in [39, 37, 21] is not clear due to the difficulties caused by the ℓ_0 term that is nonconvex and noncontinuous. Even though many literatures discussed the convergence properties of the ADMM in different applications, they cannot be to directly applied here for the fact that these are mainly for strongly convex [26], smooth+nonconvex [16], or continuous (possibly nonsmooth)+nonconvex [36, 15] optimization problems. With the convex analysis tools [29], here we are able to show the convergence property of the novel algorithm for the ℓ_0 regularized optimization problem in the following sections.

This remainder of the paper is organized as follows. In Section 2, we present the general Poisson noisy imaging model and the corresponding ADMM scheme based algorithm design. Then the convergence analysis of the algorithm is shown in Section 3. Simulation results are presented in Section 4 to verify the convergence of the proposed algorithm from numerical perspective. Finally, Section 5 concludes this paper.

2 Imaging Models

Let $u \in \mathbb{R}^{n\times n}$ be the original image, A be the linear operator and f be the recorded image. Then the Poisson noisy imaging model can be expressed as

$$f = \mathcal{P}(Au), \tag{2.1}$$

where $\mathcal{P}$ stands for the Poisson noise process. Note that, there are different choices for the linear operator A. For example, if A is a spatially invariant point spread function it leads to a blurring effect. If A is an identity operator, it is simply a Poisson noise model. When A is a Radon transform, it models the CT imaging process. The likelihood 2.2 of the observed image f is given by

$$p(f\,|\,Au) = \prod_{i=1}^{N} \frac{((Au)_i)^{f_i}}{f_i!} e^{-(Au)_i} \tag{2.2}$$

Where $(Au)_i$ is the ith component of Au and $N = n^2$ is the total number of pixel. Moreover, in order to make it more convenient in the following discussion, we will claim the value . f_i at each pixel is finite in the image domain and $f_i > 0$.

In this work, we consider the following optimization problem for Poisson

$$\min_{u\geq 0} \langle 1, Au\rangle - \langle f, \log(Au)\rangle + \| \lambda \cdot Wu\|_0, \tag{2.3}$$

where W is a linear sparsity lifting operator such as the tight wavelet frame transform[7], and $\| \lambda \cdot \alpha\|_0 = \sum_{l=0}^{L-1} \sum_{j\in\mathcal{I}} \lambda_{l,j} \left\|\alpha_{l,j}\right\|_0$ .Furthermore, $\left\|\alpha_{l,j}\right\|_0$ is defined to be the number of nonzero elements of tight wavelet frame coefficient $\alpha_{l,j}$ and $\lambda_{l,j} > 0$ is the weight of $\alpha_{l,j}$ at a given pixel j in band l. For convenient, we now rewrite model (2.3) as

$$\min_{u\geq 0} \mu(\langle 1, Au\rangle - \langle f, \log(Au)\rangle) + \| \lambda^0 \cdot Wu\|_0, \tag{2.4}$$

where $\mu > 0$ and $\lambda^0 = \lambda\mu$. The corresponding augmented Lagrangian functional is

$$\begin{aligned} L(w,u,z;p) = \mu(\langle 1, w\rangle - \langle f, \log w\rangle) + \left\|\lambda^0 \cdot Wu\right\|_0 + \delta_{\mathbb{R}_+^N}(z) \\ + \frac{\gamma}{2}\left\|Au - w + \gamma^{-1} p_1\right\|_2^2 + \frac{\eta}{2}\left\|u - z + \eta^{-1} p_2\right\|_2^2 \end{aligned} \tag{25}$$

Where $p = (p_1, p_2)$ is the Lagrangian multiplier, γ and η are penalty parameters, and $\delta_E(\cdot)$ is the indicator function of a set E. The iterative scheme of ADMM for (2.5) can be rewritten as

$$\begin{aligned} w^{k+1} &= \arg\min_w L\left(w, u^k, z^k; p^k\right) \\ &= \arg\min_w \mu(\langle 1, w\rangle - \langle f, \log w\rangle) + \frac{\gamma}{2}\left\|Au^k - w + \gamma^{-1} p_1^k\right\|_2^2, \\ u^{k+1} &= \arg\min_w L\left(w^{k+1}, u, z^k; p^k\right) \\ &= \arg\min_u \frac{\gamma}{2}\left\|Au - w^{k+1} + \gamma^{-1} p_1^k\right\|_2^2 + \frac{\eta}{2}\left\|u - z^k + \eta^{-1} p_2^k\right\|_2^2 + \left\|\lambda^0 \cdot Wu\right\|_0, \\ z^{k+1} &= \arg\min_z L\left(w^{k+1}, u^{k+1}, z; p^k\right) \\ &= \arg\min_z \delta_{\mathbb{R}^+}(z) + \frac{\eta}{2}\left\|u^{k+1} - z + \eta^{-1} p_2^k\right\|_2^2, \\ p_1^{k+1} &= p_1^k + \gamma\left(Au^{k+1} - w^{k+1}\right), \\ p_2^{k+1} &= p_2^k + \eta\left(u^{k+1} - z^{k+1}\right) \end{aligned} \tag{2.6}$$

where $p^k = (p_1^k, p_2^k)$ is the Lagrangian multiplier. After simple calculation of the solution for each subproblem, we obtain the ADMM scheme based algorithm are presented in Algorithm (1). Note that, the most difficulties are centered on u-subproblem. It can be solved by one step Hard thresholding by introducing a quadratic penalty to separate the wavelet frame coefficients Wu. For more details, consider the following general form optimization problem

$$\min_u \frac{1}{2}\| Ku - b\|^2 + \| \lambda \cdot Wu\|_0.$$

It is solved approximately by the an alternative iteration

$$\begin{aligned} u^l &= \arg\min_u \frac{1}{2}\| Ku - b\|^2 + \frac{\rho}{2}\left\|Wu - v^l\right\|^2, \\ v^l &= \arg\min_v \| \lambda \cdot v\|_0 + \frac{\rho}{2}\left\|v - Wu^l\right\|^2. \end{aligned} \tag{2.7}$$

Then we have the iteration step

$$u^l = \left(K^\top K + \rho I\right)^{-1}\left(K^\top b + \rho W^\top H_{\lambda/\rho}\left(Wu^{l-1}\right)\right).$$

where $H_\alpha(\cdot)$ is the hard threshold operator. As a result, u subproblem in (1) can be approximately solved efficiently. In the theoretical analysis, we assume u step is solved accurately as the closed form formulation.

Algorithm 1 ADMM for Poisson noise model.

1: Input: $\gamma>0,\mu>0,\theta>0,\eta>0,\lambda>0,\rho>0.$;

2: Initialize: Set $w^0=0,u^0=0,z^0=0,p_1^0=0,p_2^0=0$;

3: For $k=1,2,\ldots$ do

4: $w^{k+1}=\frac{1}{2\gamma}\left[\gamma Au^k+p_1^k-\mu+\sqrt{\left(\gamma Au^k+p_1^k-\mu\right)^2+4\gamma\mu f}\right]$;

5: $u^{k+1}=\left(A^\top A+\frac{\eta+\rho}{\gamma}I\right)^{-1}\left(A^\top\left(w^{k+1}-\frac{p_1^k}{\gamma}\right)+\frac{\eta}{\gamma}\left(z^k-\frac{p_2^k}{\eta}\right)+\frac{\rho}{\gamma}W^\top H_{\lambda/\rho}\left(Wu^k\right)\right)$;

6: $z^{k+1}=\max\left\{u^{k+1}+\eta^{-1}p_2^k,0\right\}$;

7: $p_1^{k+1}=p_1^k+\gamma\left(Au^{k+1}-w^{k+1}\right)$;

8: $p_2^{k+1}=p_2^k+\eta\left(u^{k+1}-z^{k+1}\right)$.

9: Check the stopping criterion.

10: end for

11: Output: w^*, u^*, z^*.

For the deblurring problem, the blur operator A and its transpose are all block circulant [24] under the periodic boundary condition for u. The closed form solution for u-subproblem can be solved by 2D discrete fast Fourier transform (FFT). So the resulted algorithm is highly efficient in the numerical implementation. However, for the CT image reconstruction problem, A has a more complex form and FFT cannot be adopted. The inverse operator in the closed form formulation of u step leads to difficulties in the computation. Thus, in the following we use the conjugate gradient algorithm [4, 22] to solve the u-subproblem for the applications mentioned in the paper. In the following, we will focus on the convergence analysis of Algorithm 1. We will denote the above algorithm as ADMM-PNIR for Poisson noise image restoration. For convenience of discussion in the following sections, the above one step IHT based ADMM algorithm is denote by the FT-ADMM algorithm when the u-subproblem is solved by FFT. Another strategy for solving the u-subproblem is to use the conjugate gradient (CG) method [4]. This alternative approach is denoted as the CG-ADMM algorithm. The performance of these two algorithms will be more clear in the numerical simulation. As can be seen later in Section 4, the FT-ADMM is more efficient than CG-ADMM while the later one produces higher quality image.

3 Convergence Analysis

In this section we will discuss the convergence property of the ADMM scheme (1) presented Section 2. The advantage of ADMM is that the generated solutions converge even if each subproblem is not solved exactly. Even though there are many convergence results for the ADMM scheme, they are mainly for the convex and smooth objective functions [13, 14]. As the newly proposed model (2.3)

has a nonconvex objective function, we should be more careful to check the convergence property. since the ℓ_0 norm is obviously nonconvex and noncontinuous.

The analysis of the ADMM scheme is usually based on the boundedness between the current iteration and the optimal solution set that is not suitable here. The routine in [16] to prove the convergence of ADMM scheme with stochastic coordinate updating is based on the nonconvex and differentiable objective function. Our analysis here is similar but the objective function we considered is nonconvex and noncontinuous and more complicated analysis are needed. The analysis results for nonconvex smooth optimization problems in [36] with constrains are also only suitable for a special class of optimization problems. In this work, with the convex analysis tools [29, 18], we will investigate the convergence property of the newly proposed ADMM algorithm (1). with a similar strategy. Now, with the convex analysis tools we will show the convergence results of the proposed FT-ADMM algorithm.

We define the objective function as $G(w)=\mu(\langle 1,w\rangle-\langle f,\log(w)\rangle)$ and $H(z)=\delta_{\mathbb{R}^+}(z)$. From (2.4) we have

$$\min\left\{G(w)+H(z)+\left\|\lambda^0\cdot Wu\right\|_0 \mid Au=w,u=z\right\}.$$

The corresponding augmented Lagrangian function is

$$\begin{aligned}L(w,u,z;p)=&G(w)+\langle p_1,Au-w\rangle+\frac{\gamma}{2}\|Au-w\|_2^2\\&+H(z)+\langle p_2,u-z\rangle+\frac{\eta}{2}\|u-z\|_2^2\\&+\left\|\lambda^0\cdot Wu\right\|_0.\end{aligned}\tag{3.1}$$

When the successive difference of the augmented Lagrangian function (3.1) can be bounded under some proper assumptions, the decreasing property will be obtained.Whilethe augmented Lagrangian function is nonconvex, the generates $\{w^k,u^k,z^k\}$ of the ADMM scheme (1) only converge to a local minimizer. Since the detailed proof needs many other convex analysis tools and is technical long, we will omit here.Interested reader can refer to our another finished work as well as some other recently convergence results for the ADMM scheme and the optimality conditions for nonconvex co optimization problem in [16, 25].

We first show that the successive difference of the primal and dual variables are bounded.The first order necessary optimality conditions for each subproblems in the ADMM scheme (2.6) are useful in the following discussion.

Proposition 1. For the fired variable $w^\star,u^\star,p^\star=\left(p_1^\star,p_2^\star\right)$, the subproblem for z as min

$$\min_z L\left(w^\star,u^\star,z;p^\star\right)=\min_z H(z)+\left\langle p_2^\star,u^\star-z\right\rangle+\frac{\eta}{2}\left\|u^\star-z\right\|^2$$

has the first order necessary optimality condition

$$\partial H\left(z^\star\right)\ni p_2^\star+\eta\left(u^\star-z^\star\right),\tag{3.2}$$

where $\partial H\left(z^\star\right)$ is the subdifferential of $H(\cdot)$ at $z^\star$.

Proposition 2. For the fired variable $u^\star,z^\star,p^\star=\left(p_1^\star,p_2^\star\right)$, the subproblem for W as

$$\min_{w} L\left(w,u^{\star},z^{\star};p^{\star}\right)=\min_{w} G(w)+\left\langle p_1^{\star},Au^{\star}-w\right\rangle+\frac{\gamma}{2}\left\|Au^{\star}-w\right\|^2$$

has the first order necessary optimality condition

$$\nabla G\left(w^{\star}\right)=p_1^{\star}+\gamma\left(Au^{\star}-w^{\star}\right). \tag{3.3}$$

Note the fact that the function $G\left(w\right)$ is nonnegative with respect to w. Briefly, we consider the one pixel case.

Proposition 3. Let $g(x)=x-y\log(x)$, where $x,y\in\mathbb{R}$ and $x>0$, if $y\geq 0$, g is convex. If $y>0$, then $g(x)$ is strictly conver. Furthermore, if $\tilde{g}(x)=x-y\log(x)-(y-y\log(y))$, then $\tilde{g}$ is nonnegative.

$$\min\left\{\sum_{i=1}^{N}\phi_i\left(x_i\right)+\beta\|x\|_0:x\in\mathbb{R}^N\right\}. \tag{3.4}$$

Let $\tilde{x}_i^{*}\in\arg\min\left\{\phi_i\left(x_i\right):x_i\in\mathbb{R}\right\}$ and $v_i^{*}=\phi_i(0)-\beta-\phi_i\left(\tilde{x}_i^{*}\right)$ for $i=1,2,\ldots,N$. Then x^{*} is an optimal solution of minimization problem (3.4), where x^{*} is defined as follows:

$$x_i^{*}=\begin{cases}\tilde{x}_i^{*} & if \quad v_i^{*}\geq 0,\\ 0 & otherwise,\end{cases} i=1,2,\ldots,N.$$

When the function $\tilde{g}$ defined above is used to model imaging process, it is widely known as the generalized Kullback-Leibler (KL) divergence.It is also coincides with the Bregman distance of the Boltzmann-Shannon entropy, so the nonnegative property is obvious[3].

Proposition 4. Suppose that the function $g(x)\in C\left(\mathbb{R}^N\right)$) is conver. The augmented Lagrangian function of $L(x,y;p)=g(y)+\langle p,x-y\rangle+\frac{\rho}{2}\|x-y\|^2$, where the Lagrangian multiplier $p\in\mathbb{R}^N$ and $p>0$ componentwise, and the penalty parameter $0<\rho\in\mathbb{R}$. Then $L(x,y;p)$ is strongly conver with respect to variable y. In other words, for the fired $x\in\mathbb{R}$, $p>0$, $\rho>0$, and any $y_1,y_2\in\mathbb{R}^N$

$$L\left(x,y_2;p\right)-L\left(x,y_1;p\right)\geq\left\langle\nabla_y L\left(x,y_1;p\right),y_2-y_1\right\rangle+\frac{\rho}{2}\left\|y_2-y_1\right\|^2. \tag{3.5}$$

Lemma 1. Let $g(u)=\mu(\langle 1,u\rangle-\langle f,\log(u)\rangle)\in C\left(\mathbb{R}^N\right)$, where $\mu\in\mathbb{R}$, $f\in\mathbb{R}^N$ and $\mu>0$, $f>0$ componentwise. The augmented Lagrangian function for $g(u)$ is

$$L(w,u;p)\triangleq g(w)+\langle p,u-w\rangle+\frac{\gamma}{2}\left\|u-w\right\|^2$$

where the Lagrangian multiplier $p\in\mathbb{R}^N$ and $p>0$ componentwise, and the penalty parameter $0<\gamma\in\mathbb{R}$. Moreover, denoting the dual variable $p^{k+1}\triangleq p^k+\gamma\left(u^{k+1}-w^{k+1}\right)$ and fix small enough $\varepsilon>0$, then for all w^k, $w^{k+1}\in[\epsilon,+\infty)$ we have

$$\begin{aligned} & L\left(w^{k+1},u^{k+1};p^{k+1}\right)-L\left(w^{k+1},u^{k+1};p^{k}\right) \\ &= \frac{1}{\gamma}\left\|p^{k+1}-p^{k}\right\|^{2} \\ &\le C_{1}\left\|w^{k+1}-w^{k}\right\|^{2}, \end{aligned} \tag{3.6}$$

where $C_1 > 0$ 0 is a constant.

Proof. Based on the first order necessary optimality condition for the optimization problem

$$\min_{w} L\left(w,u^{k+1};p^{k}\right),$$

at k+1th iteration, we have

$$\nabla g\left(w^{k+1}\right)=p^{k}+\gamma\left(Au^{k+1}-w^{k+1}\right)\triangleq p^{k+1}.$$

As a result

$$\begin{aligned} \left\|p^{k+1}-p^{k}\right\| &= \left\|\nabla g\left(w^{k+1}\right)-\nabla g\left(w^{k}\right)\right\| \\ &= \left\|\mu f\,\frac{w^{k+1}-w^{k}}{w^{k}w^{k+1}}\right\| \\ &\le c_{1}\left\|w^{k+1}-w^{k}\right\|, \end{aligned} \tag{3.7}$$

where $c_1 > \mu\| f\|/\epsilon^2$ is a constant and the quotient in the second equation is in the Hadamard sense, i.e. pixel by pixel. Combining this fact with the first order optimality condition, we have

$$\begin{aligned} & L\left(w^{k+1},u^{k+1};p^{k+1}\right)-L\left(w^{k+1},u^{k+1};p^{k}\right) \\ &= \left\langle p^{k+1}-p^{k},u^{k+1}-w^{k+1}\right\rangle \\ &= \frac{1}{\gamma}\left\|p^{k+1}-p^{k}\right\|^{2} \\ &\le C_{1}\left\|w^{k+1}-w^{k}\right\|^{2}. \end{aligned} \tag{3.8}$$

This concludes the proof.

Lemma 2. Let $h(u)=\delta_E(u)$ be the indicator function of closed conver set E, where $E=\mathbb{R}_{+}^{\mathbb{N}}$. The augmented Lagrangian function for $h(u)$ is

$$L(u,z;p)=\delta_E(z)+\langle p,u-z\rangle+\frac{\eta}{2}\| u-z\|^2,$$

Where $p\in\mathbb{R}^N$, $\eta\in\mathbb{R}$ and $p\ge 0$ and $\eta\ge 0$. Denote the dual variable $p^{k+1}\triangleq p^{k}+\eta\left(u^{k+1}-z^{k+1}\right)$. Moreover, assuming that the multiplier p^k is chosen suitably in each iteration. then for all z^k, $z^{k+1}\in E$ we have the following relation

$$\begin{aligned} & L\left(u^{k+1}, z^{k+1}; p^{k+1}\right) - L\left(u^{k+1}, z^{k+1}; p^{k}\right) \\ &= \frac{1}{\eta}\left\| p^{k+1} - p^{k} \right\|^2 \\ &\le C_2 \left\| z^{k+1} - z^{k} \right\|^2, \end{aligned} \tag{3.9}$$

where $C_2 > 0$ is a constant.

Proof. Based on the first order necessary optimality condition for the optimization problem

$$\min_{z} L\left(u^{k+1}, z; p^{k}\right)$$

at the k+1th iteration, we have

$$\partial h\left(z^{k+1}\right) - p^{k} - \eta\left(u^{k+1} - z^{k+1}\right) \ni 0$$

or

$$\partial h\left(z^{k+1}\right) \ni p^{k} + \eta\left(u^{k+1} - z^{k+1}\right) \triangleq p^{k+1},$$

where $\partial h(z)$ is the subdifferential of $h(\cdot)$ at z. More specifically, $\partial h(z) = \mathcal{N}_E(z)$ is the normal cone to the convex set E at point z, which is equivalent to that $\forall d \in \mathcal{N}_E(z)$, and there is the inequality

$$\langle d, y - z \rangle \le 0,$$

for all $y \in E$.

For case $z^k \in intE$, we can obviously have $\mathcal{N}_E\left(z^k\right) = \{0\}$, for all $k = 1, 2, 3, \cdots$.

We consider the case that z^k is on the boundary of the convex set E and $z^k \neq 0$. Let's choose two points related to z as $x_1 = \max\left\{z_j \neq 0\right\} e_J$, and $x_2 = \min\left\{z_j \neq 0\right\} e_J$, where e_J the vector with elements be one indexed by J and zero else, and $J = \left\{ j \in \{1, 2, 3, \cdots, N\} : z_j \neq 0 \right\}$. Then by the definition of normal cone to convex set, for $p^k \in \mathcal{N}_E\left(z^k\right)$ the former chosen x_1, $x_2 \in E$ satisfy $\left\langle p^k, x_1 - z^k \right\rangle \le 0$ and $\left\langle p^k, x_2 - z^k \right\rangle \le 0$. As a result, there must have $p^k = 0$ and $\mathcal{N}_E\left(z^k\right) = \{0\}$.

For case $z^k = 0$, based on the definition of normal cone to convex set, we have $\mathcal{N}_E\left(z^k\right) = \left\{ p^k \le 0 \right\}$. Assumed that we choose the dual variable p^k in each iteration properly, such as $p^k = t^k e_N \in \mathcal{N}_E\left(z^k\right)$ where $e_N \in \mathbb{R}^N$ with all elements be one, and satisfying $0 < \left| t^{k+1} - t^k \right| \le \left\| z^{k+1} - z^k \right\|$. Then for any cases above, we have the result that $\left\| p^{k+1} - p^k \right\| \le c_2 \left\| z^{k+1} - z^k \right\|$, where $c_2 \ge 1$ is a constant. Combining all the facts, we have

$$\begin{aligned}
&L\left(u^{k+1}, z^{k+1}; p^{k+1}\right) - L\left(u^{k+1}, z^{k+1}; p^{k}\right) \\
&= \left\langle p^{k+1} - p^{k}, u^{k+1} - z^{k+1} \right\rangle \\
&= \frac{1}{\eta}\left\| p^{k+1} - p^{k} \right\|^2 \\
&\le C_2 \left\| z^{k+1} - z^{k} \right\|^2
\end{aligned}$$

where $C_2 > 0$ is a constant.

Note that if the penalty term makes z^k approximating to u^k in each iteration of the proposed algorithm, then z^k is sparse but is not always equal to 0. As a result, there will be higher probability to choose 0 as the subgradient of z^k at each iteration procedure. The above lemma is also reasonable in the numerical practice. Now under the following fundamental assumption, we show the boundedness of successive difference for the augmented Lagrangian function (3.1).

Assumption 1. Assume that the nonzero components of the tight wavelet frame coefficients are decreasing in each iteration of the numerical implementation. More detailedly, the framelet based signal representation is becoming sparser step by step in finite number of iterations and satisfies

$$\left\| \lambda^0 \cdot W u^k \right\|_0 - \left\| \lambda^0 \cdot W u^{k+1} \right\|_0 \ge M,$$

where $M \ge 0$ is a suitable constant integer.

This assumption is satisfied in the numerical simulation by choosing a suitable initialization condition for u^0. It will also be useful in the following theoretical analysis.

Theorem 1. Let $\left\{ w^k, u^k, z^k; p^k \right\}$ be generated by the A DMM scheme (1). Suppose that the Assumption 1 is satisfied with a fired integer $M \ge 1$. Moreover, assume that the operator A satisfies $\|A\| \ne 0$ and the parameters γ, η are chosen sufficiently large. Then we have the bound for the augmented Lagrangian function (3.1) as following

$$\begin{aligned}
&L\left(w^{k+1}, u^{k+1}, z^{k+1}; p^{k+1}\right) - L\left(w^{k}, u^{k}, z^{k}; p^{k}\right) \\
&\le C_1 \left\| w^{k+1} - w^{k} \right\|^2 + C_2 \left\| z^{k+1} - z^{k} \right\|^2 + C_3 \left\| u^{k+1} - u^{k} \right\|^2,
\end{aligned} \tag{3.10}$$

where C_1, C_2, and C_3 are constants related to γ and η.

Proof. We first split the successive difference of the augmented Lagrangian by

$$\begin{aligned}
&L\left(w^{k+1}, u^{k+1}, z^{k+1}; p^{k+1}\right) - L\left(w^{k}, u^{k}, z^{k}; p^{k}\right) \\
&\le \left[L\left(w^{k+1}, u^{k+1}, z^{k+1}; p^{k+1}\right) - L\left(w^{k+1}, u^{k+1}, z^{k+1}; p^{k}\right) \right] \\
&\quad + \left[L\left(w^{k+1}, u^{k+1}, z^{k+1}; p^{k}\right) - L\left(w^{k}, u^{k}, z^{k}; p^{k}\right) \right]
\end{aligned} \tag{3.11}$$

The first term in (3.11) can be bounded by

$$\begin{aligned} & L\left(w^{k+1},u^{k+1},z^{k+1};p^{k+1}\right)-L\left(w^{k+1},u^{k+1},z^{k+1};p^{k}\right) \\ &=\left\langle p_1^{k+1}-p_1^{k},Au^{k+1}-w^{k+1}\right\rangle+\left\langle p_2^{k+1}-p_2^{k},u^{k+1}-z^{k+1}\right\rangle \\ &=\frac{1}{\gamma}\left\|p_1^{k+1}-p_1^{k}\right\|^2+\frac{1}{\eta}\left\|p_2^{k+1}-p_2^{k}\right\|^2 \\ &\le C_1\left\|w^{k+1}-w^{k}\right\|^2+C_2\left\|z^{k+1}-z^{k}\right\|^2 \end{aligned} \tag{3.12}$$

where the last inequality is due to Lemma 1 and 2.

The second term in (3.11) can also be bounded by further splitting it as

$$\begin{aligned} & L\left(w^{k+1},u^{k+1},z^{k+1};p^{k}\right)-L\left(w^{k},u^{k},z^{k};p^{k}\right) \\ &=\left[L\left(w^{k+1},u^{k+1},z^{k+1};p^{k}\right)-L\left(w^{k+1},u^{k+1},z^{k};p^{k}\right)\right] \\ &\quad+\left[L\left(w^{k+1},u^{k+1},z^{k};p^{k}\right)-L\left(w^{k},u^{k+1},z^{k};p^{k}\right)\right] \\ &\quad+\left[L\left(w^{k},u^{k+1},z^{k};p^{k}\right)-L\left(w^{k},u^{k},z^{k};p^{k}\right)\right] \\ &\triangleq E_{q1}+E_{q2}+E_{q3}. \end{aligned} \tag{3.13}$$

Then we will analyze the boundedness of the last three terms in the following few lines.

For the first term E_{q1}, based on the first order necessary optimality condition for z-subproblem in (3.2), we obtain

$$\begin{aligned} E_{q1} &\triangleq L\left(w^{k+1},u^{k+1},z^{k+1};p^{k}\right)-L\left(w^{k+1},u^{k+1},z^{k};p^{k}\right) \\ &\le -\left\langle\nabla_z L\left(w^{k+1},u^{k+1},z^{k+1};p^{k}\right),z^{k}-z^{k+1}\right\rangle-\frac{\eta}{2}\left\|z^{k}-z^{k+1}\right\|^2 \\ &=-\frac{\eta}{2}\left\|z^{k}-z^{k+1}\right\|^2, \end{aligned} \tag{3.14}$$

where we choose the subgradient $p_2^{k+1}\in\partial H\left(z^{k+1}\right)$ to be the partial derivative of $\nabla_z L$. Due to the strictly convexity of the augmented Lagrangian function for z-subproblem, the above inequality is obtained by Proposition 4.The optimality condition (3.2) leads to $\nabla_z L\left(w^{k+1},u^{k+1},z^{k+1};p^{k}\right)=0$.

For the second term E_{q2}, based on the first order necessary optimality condition for w-subproblem in (3.3), we have

$$\begin{aligned} E_{q2} &\triangleq L\left(w^{k+1},u^{k+1},z^{k};p^{k}\right)-L\left(w^{k},u^{k+1},z^{k};p^{k}\right) \\ &\le -\left\langle\nabla_w L\left(w^{k+1},u^{k+1},z^{k};p^{k}\right),w^{k}-w^{k+1}\right\rangle-\frac{\gamma}{2}\left\|w^{k}-w^{k+1}\right\|^2 \\ &=-\frac{\gamma}{2}\left\|w^{k}-w^{k+1}\right\|^2. \end{aligned} \tag{3.15}$$

Due to the strictly convexity of the augmented Lagrangian function for w-subproblem, the above inequality can be obtained by using the result in Proposition 4. The last equality is obtained based on the first order necessary optimality condition (3.3).

For the third term E_{q3}, we have

$$\begin{aligned} E_{q3} &\triangleq L\left(w^k, u^{k+1}, z^k; p^k\right) - L\left(w^k, u^k, z^k; p^k\right) \\ &\le -\frac{\gamma}{2}\left\|Au^{k+1} - Au^k\right\|^2 - \frac{\eta}{2}\left\|u^{k+1} - u^k\right\|^2 \\ &\quad + \left(\left\|\lambda^0 \cdot Wu^{k+1}\right\|_0 - \left\|\lambda^0 \cdot Wu^k\right\|_0\right) \\ &\le -\frac{\gamma}{2}\left\|Au^{k+1} - Au^k\right\|^2 - \frac{\eta}{2}\left\|u^{k+1} - u^k\right\|^2 - M \end{aligned} \tag{3.16}$$

where the first inequality is obtained by the strictly convexity of the quadratic function about u and the result in Proposition 4, and the second inequality by the sparsity Assumption 1.

Combining the above bounds for E_{q1}, E_{q2} and E_{q3} in (3.14), (3.15) and (3.16) respectively. Then we obtain (3.13) the following

$$\begin{aligned} &L\left(w^{k+1}, u^{k+1}, z^{k+1}; p^k\right) - L\left(w^k, u^k, z^k; p^k\right) \\ &\triangleq E_{q1} + E_{q2} + E_{q3} \\ &\le -\frac{\eta}{2}\left\|z^k - z^{k+1}\right\|^2 - \frac{\gamma}{2}\left\|w^k - w^{k+1}\right\|^2 \\ &\quad -\frac{\gamma}{2}\left\|Au^{k+1} - Au^k\right\|^2 - \frac{\eta}{2}\left\|u^{k+1} - u^k\right\|^2 - M. \end{aligned} \tag{3.17}$$

Furthermore, combining the results in (3.12) and (3.17), we have the bound for (3.11) as

$$\begin{aligned} &L\left(w^{k+1}, u^{k+1}, z^{k+1}; p^{k+1}\right) - L\left(w^k, u^k, z^k; p^k\right) \\ &\le \left(C_1 - \frac{\gamma}{2}\right)\left\|w^{k+1} - w^k\right\|^2 + \left(C_2 - \frac{\eta}{2}\right)\left\|z^{k+1} - z^k\right\|^2 \\ &\quad -\frac{1}{2}\left[\gamma\left\|Au^{k+1} - Au^k\right\|^2 + \eta\left\|u^{k+1} - u^k\right\|^2 - M\right] \\ &\le C_3\left\|w^{k+1} - w^k\right\|^2 + C_4\left\|z^{k+1} - z^k\right\|^2 + C_5\left\|u^{k+1} - u^k\right\|^2, \end{aligned} \tag{3.18}$$

where C_3, C_4, C_5 are constants.

Notice that the bound in the right hand side inequality (3.10) of Theorem 1 canbe negative when we choose the parameters γ and η that are large enough. Then we obtain the following convergence result for the augmented Lagrangian functional (3.1) in numerical iterations.

Theorem 2. Let $\left\{w^k, u^k, z^k; p^k\right\}$ be generated by the A DMM scheme (1). Suppose that the assumptions in Theorem I are satisfied. Then for the augmented Lagrangian functional $L\left(w, u, z; p\right)$ in (3.1), we have

$$\lim_{k\to\infty} L\left(w^k, u^k, z^k; p^k\right) = C_L,$$

where C_L is a constant.

Proof. The augmented Lagrangian function (3.1) in the k+1th iteration of the ADMM scheme is

$$
\begin{aligned}
&L\left(w^{k+1},u^{k+1},z^{k+1};p^{k+1}\right)\\
&=G\left(w^{k+1}\right)+\left\langle p_1^{k+1},Au^{k+1}-w^{k+1}\right\rangle+\frac{\gamma}{2}\left\|Au^{k+1}-w^{k+1}\right\|^2\\
&\quad+H\left(z^{k+1}\right)+\left\langle p_2^{k+1},u^{k+1}-z^{k+1}\right\rangle+\frac{\eta}{2}\left\|u^{k+1}-z^{k+1}\right\|^2\\
&\quad+\left\|\lambda^0\cdot Wu^{k+1}\right\|_0.
\end{aligned}
$$

With the first order necessary optimality conditions in (3.2) and (3.3), we have $p_1^{k+1}=\nabla G\left(w^{k+1}\right)$ and $p_2^{k+1}\in\partial H\left(z^{k+1}\right)$. If we choose the subgradient p_2^{k+1} to be the partial derivative of $\nabla_z L$ and combine the suitable choice of parameters γ and η, we have

$$
\begin{aligned}
&L\left(w^{k+1},u^{k+1},z^{k+1};p^{k+1}\right)\\
&\geq G\left(w^{k+1}\right)+H\left(z^{k+1}\right)+\left\|\lambda^0\cdot Wu\right\|_0\\
&\geq G\left(w^{k+1}\right)\geq G(f).
\end{aligned}
\tag{3.19}
$$

The last inequality is due to the property of $G(\cdot)$ in Proposition 3. This lower bound combined with the decrease property in Theorem 1 implies that the augmented Lagrangian function $L\left(w^k,u^k,z^k;p^k\right)$ is convergent. There exists a constant to be the limit.

Since the l_0 term in the augmented Lagrangian function is noncontinuous and nonconvex, the Lagrangian functional $L\left(w,u,z;p\right)$ in (3.1) can be verified to decrease by choosing parameters γ, η suitably under the assumptions in Theorem 1. If we replace the l_0 term by other locally continuous function, such as the nonconvex $l_p(0<\mathrm{p}\leq 1)$ norm or the convex hull of l_0 norm[12, 32], the above result in Theorem 1 can be verified similarly.

There are several approaches to approximate the ℓ_0 semi-norm, such as ℓ_0 and convex hull of ℓ_0 norm are usually nonsmooth nonconvex but continuous at the origin. It is interesting to see that result in higher sparsity of the framelet representation than the ℓ_1 convex relaxation [30]. In addition, the global minimizer of the generalized ℓ_0 minimization problem can be obtained from the locally relaxed nonconvex continuous models [25, 35, 23].

Based on Theorem 1 and Theorem 2, we can obtain the following results.

Theorem 3. Suppose that the assumptions in Theorem I are satisfied. Let the sequence of iterates $\left\{w^k,u^k,z^k;p^k\right\}$ be generated by the ADMM based algorithm FT-ADMM (or CG-ADMM).Then we have

$$
\lim_{k\to\infty}\left\|w^{k+1}-w^k\right\|=0,\quad \lim_{k\to\infty}\left\|z^{k+1}-z^k\right\|=0,
$$

$$
\lim_{k\to\infty}\left\|u^{k+1}-u^k\right\|=0.
$$

Moreover, we have

$$
\lim_{k\to\infty}\left\|Au^k-w^k\right\|=0,\quad \lim_{k\to\infty}\left\|u^k-z^k\right\|=0.
$$

Proof. Based on the result in Theorem 1, for sufficient large parameters γ and η we have

$$\begin{aligned} &L\left(w^{k+1},u^{k+1},z^{k+1};p^{k+1}\right)-L\left(w^{k},u^{k},z^{k};p^{k}\right)\\ &\le C_3\left\|w^{k+1}-w^{k}\right\|^2+C_4\left\|z^{k+1}-z^{k}\right\|^2+C_5\left\|u^{k+1}-u^{k}\right\|^2, \end{aligned}$$

where the constants in the right hand side are negative. Through summing the above inequality from $k=0$ to n, we can derive

$$\begin{aligned} &L\left(w^{n+1},u^{n+1},z^{n+1};p^{n+1}\right)-L\left(w^{0},u^{0},z^{0};p^{0}\right)\\ &\le \sum_{k=0}^{n}\left[C_3\left\|w^{k+1}-w^{k}\right\|^2+C_4\left\|z^{k+1}-z^{k}\right\|^2+C_5\left\|u^{k+1}-u^{k}\right\|^2\right]. \end{aligned} \tag{3.20}$$

Combining the negativity of each constants C_3, C_4, and C_5, and the limitation result in Theorem 2, we can infer that

$$\sum_{k=0}^{\infty}\left[C_3\left\|w^{k+1}-w^{k}\right\|^2+C_4\left\|z^{k+1}-z^{k}\right\|^2+C_5\left\|u^{k+1}-u^{k}\right\|^2\right]<+\infty, \tag{3.21}$$

which implies that

$$\lim_{k\to\infty}\left\|w^{k+1}-w^{k}\right\|=0,\lim_{k\to\infty}\left\|z^{k+1}-z^{k}\right\|=0,$$

$$\lim_{k\to\infty}\left\|u^{k+1}-u^{k}\right\|=0.$$

Therefore, we can immediately have

$$\left\|p_1^{k+1}-p_1^{k}\right\|\le C_1\left\|w^{k+1}-w^{k}\right\|\to 0,\ \text{as } k\to\infty,$$

and

$$\left\|p_2^{k+1}-p_2^{k}\right\|\le C_2\left\|z^{k+1}-z^{k}\right\|\to 0,\ \text{as } k\to\infty. \tag{3.22}$$

Due to the relations of primal-dual variables

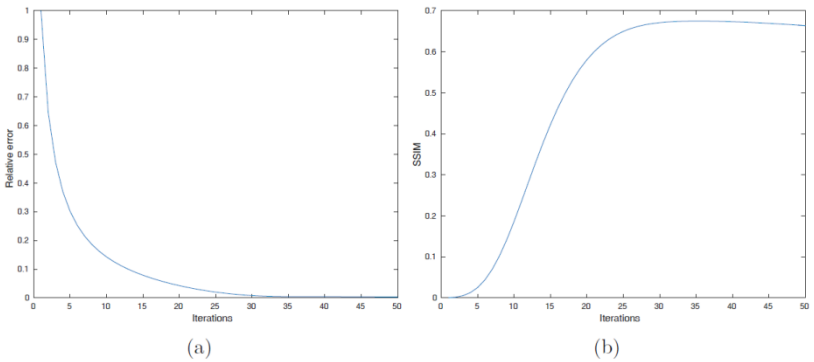
$$p_1^{k+1}=p_1^{k}+\gamma\left(Au^{k+1}-w^{k+1}\right)\quad\text{and}\quad p_2^{k+1}=p_2^{k}+\eta\left(u^{k+1}-z^{k+1}\right),$$

we obtain

$$\lim_{k\to\infty}\left\|Au^{k}-w^{k}\right\|=0,$$

$$\lim_{k\to\infty}\left\|u^{k}-z^{k}\right\|=0.$$

Since the objective function of the lo minimization problem is nonconvex and noncontinuous, it is difficult to estimate the primal and dual optimality gaps. The decreasing property of the augmented Lagrangian functional guarantee the convergence of the ADMM scheme (1) if the stopping criterion is set to $\dfrac{\left\|u^{k+1}-u^{k}\right\|}{\left\|u^{k}\right\|}$. All the results are possibly be extended to more general nonconvex noncontinuous optimization problem as long as the objective fuction has similar smoothness and convexity properties.



(a) (b)

Figure 1 Relative error and the SSIM vs. iterations. The test image is Barbara. (a) Relative error; (b) SSIM.

4 Numerical simulations

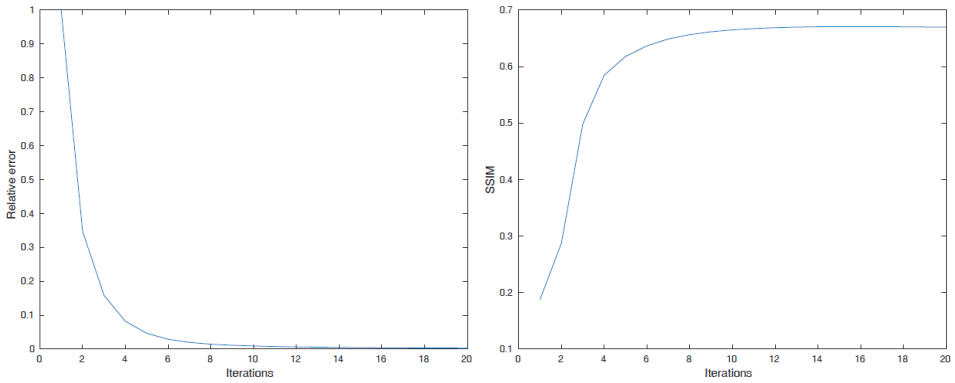
(a)True image (b)Noisy image (c)Restored image

Figure 2 The Barbara image denoising. (a) True image. (b) Measured Poisson noisy and blurred image. (c) The restored image.

In this section we present some numerical simulation results to verify the convergence property of the ADMM scheme (1) from the numerical perspective. More detailed comparison with other algorithms for Poisson noise removal and deblurring task can be found in [39]. The bench-mark test image Barbara is adopted for the deblurring problem, the FORBILD head phantom [38l is used for CT image reconstruction and the graph data Ericman [5, 6] is adopted to simulate the denoising task on manifold in the numerical experiments. The observed noisy images are obtained through the following steps:

1). The true image is scaled by a fixed number in order to simulate different Poisson noise level (such as Scale=500 for Barbara and Scale=10^3 for head phantom).;

2). The scaled image is convolved with a given point spread function to simulate the blurring effect or the Radon transform to simulate the sinogram of CT image;



(a) (b)

Figure 3 Relative error and the SSIM vs. iterations. The test image is the FORBILD head phantom. (a) Relative error; (b) SSIM.

3). The blurred image or the sinogram is contaminated by the Poisson noise to obtain the final measured data f.

The quality of the restored image is measured by the Signal-to-Noise Ratio (SNR), the relative error (RelErr) and the SSIM [42] which are defined as follows:

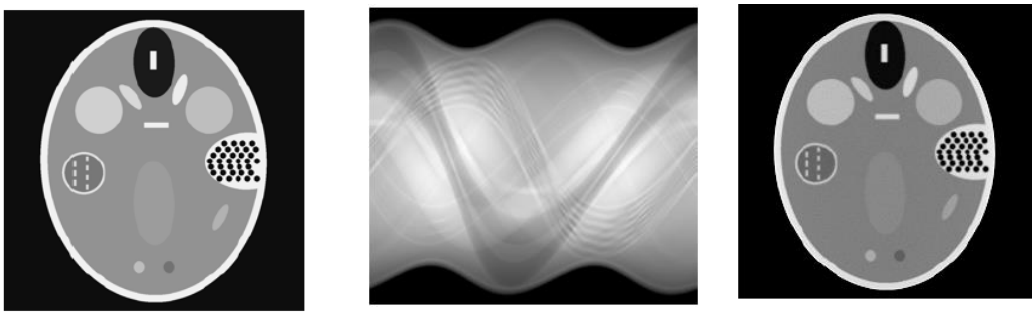
$$SNR\left(u^*\right) \triangleq 20\log_{10}\frac{\|u\|_2}{\left\|u^*-u\right\|_2}, \quad \text{and} \quad RelErr\left(u,u^*\right) \triangleq \frac{\left\|u^*-u\right\|_2}{\|u\|_2} \tag{4.1}$$

where u and u^* are the original image and the restored image respectively, The stopping criterion for Algorithm 1 is set to when the relative difference $RelErr\left(u^k,u^{k+1}\right)=\frac{\left\|u^{k+1}-u^k\right\|_2}{\left\|u^{k+1}\right\|_2} < tol = 3\times 10^{-3}$ is satisfied. The piecewise linear tight wavelet frame is chosen as the operator W and the decomposition level is set to $L=4$ in the numerical experiments. The regularization parameter λ appeared in the models presented in the previous sections takes the form $\lambda = \left\{\left(\frac{1}{2}\right)^l \lambda \,|\, l=0,1,\ldots,L-1\right\}$ where $\lambda>0$ is a tuning parameter.

In the first experiment, we used the Barbara image to test the Poisson noise image deblurring performance. For the Barbara image, the penalty parameters are chosen as $\rho=0.6$, $\gamma=0.01$, $\lambda=0.3$, $\eta=0.2$, $\mu=10$. The CG iteration is set to a fixed number 30. The relative error and SSIM with respective to iteration are shown in Fig. 1. We can observe that the relative error is monotonically decreasing. The SSIM is increasing at the beginning, but then it tends to decrease slowly after reaching the plateau. The true image of Barbara, blurred image and the restored image are shown in Fig. 2 respectively. The newly proposed model can remove the Poisson noise and preserve the details of texture well. As observed in Fig. 1 (b), the SSIM is decreasing after the plateau. This is due to the over-smoothed image if the stopping criterion is set with a smaller relative error tol. We should design a proper stopping criterion in the future work.

In the second experiment, we tested the performance of model (1) for CT image reconstruction. The true FORBILD head phantom is shown in Fig. 4 (a). The measured sinogram f is contaminated by the Poisson noise. The parameter setting for this experiment is listed as: $\rho=2.5$, $\gamma=0.0015$, $\lambda=0.01$, $\eta=0.01$, $\mu=30$. The CG iteration is set to a fixed number 20. Fig. 3 shows the relative error and SSIM curve in each iteration. The relative error decreases monotonically respective to the iteration and the SSIM increases monotonically as well. This verifies the theoretically analysis in

former sections. The reconstructed FORBILD head phantom is shown in Fig. 4 (c). All the details in the reconstructed phantom are preserved well and at the same time, the edges of the phantom are sharp enough due to the nonconvex penalty of tight wavelet frame coefficients.



(a) True phantom (b) Measured sinogram (c) Reconstructed image

Figure 4 The reconstruction of FORBILD head phantom from Poisson noisy measured sinogram. (a) True phantom. (b) Measured sinogram with Poisson noise. (c) The reconstructed head phantom.

In the third experiment, we tested the proposed nonconvex noncontinuous model for graph data denoising. The graph data of Ericman is contaminated by the Poisson noise. The linear operator A is set to the identity operator in this denoising case. The scalar for Poisson noise is set to Scale=1000 in this experiment. For the parameter setting of Algorithm 1, we tune to the optimal values as $\rho = 1$, $\gamma = 0.5$, $\lambda = 140$, $\eta = 0.001$, $\mu = 160$. Here, the u step can be solved with a simple closed form due to the trivial identity operator A. Figure 6 shows the background data, Poisson noisy data and the restored graph data of the Ericman. To measure the quality of the restored data, we use the SNR as defined previously in the above. Figure 5 shows the SNR and cost function curve with respect to the iterations.

From all the above results we can conclude that the ADMM scheme (1) in Algorithm 1 is convergent in terms of the relative error as described in the theoretical analysis.

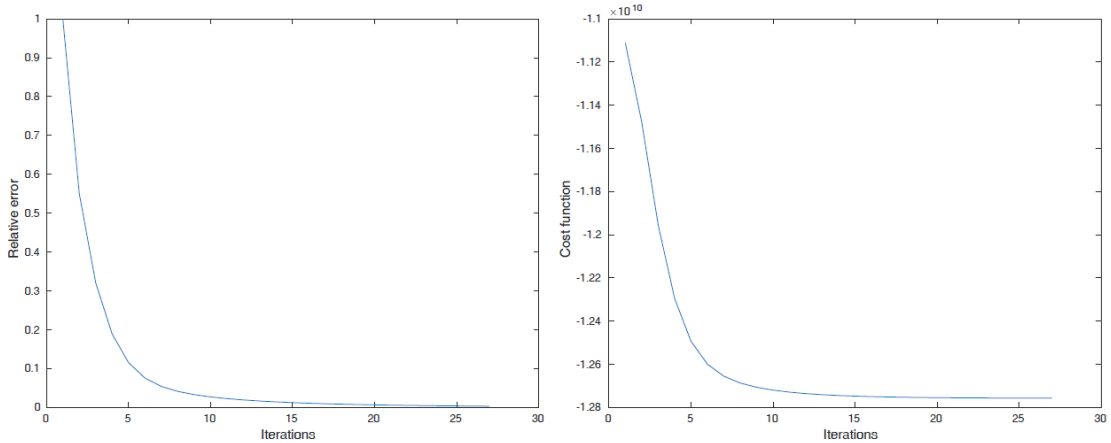
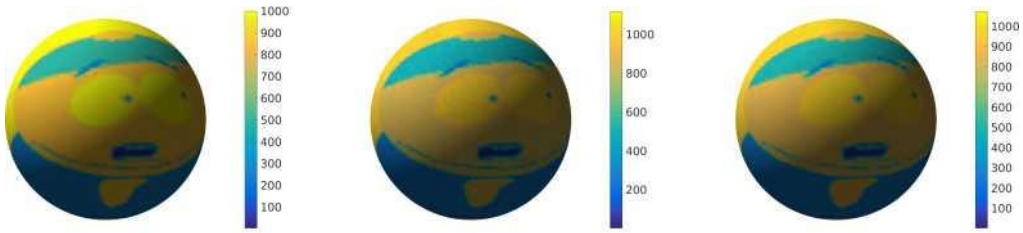


Figure 5 Relative error and the SSIM vs. iterations. The test image is the graph data Ericman. (a) Relative error; (b) Cost function.



(a) True data (b) Noisy data (c) Restored data

Figure 6 The restored graph data from Poisson noisy measured data, (a) True graph data Ericman. (b) Graph with Poisson noise, (c) The restored graph data.

5 Conclusion

In this work, we analyse the newly proposed nonconvex noncontinuous imaging model for Poisson noise removal which is solved by the famous ADMM algorithm. However, the existing results in literature for convergence of ADMM are only suitable for optimization problem with continuous (possible nonsmooth) and nonconvex objective functions. They cannot be directly applied to the ℓ_0 regularized model. With the convex analysis tools, we present the convergence property of the ADMM scheme and use in our algorithm design. This is theoretically valuable and important to guarantee the stability of the numerical simulation. To verify the convergence property from the numerical perspective, we design the image deblurring and CT image reconstruction tasks with Poisson noisy data. The relative error decreases monotonically with respective to the iteration. The results in the work can be extended to more general nonconvex and nonsmooth optimization problem.

Acknowledgment:

Shenzhen Grade E Scientific and Engineering Calculation Key Laboratory under Grant (ZDSYS201703031711426)

References

[1] H. ATTOUCH, G. BUTTAZZO, AND G. MICHAILLE, Variational analysis in Sobolev and BV spaces: applications to PDEs and optimization. vol. 17, SIAM, 2006.

[2] D. CHEN, Regularized generalized inverse accelerating linearized alternating minimization algorithm for frame-based poissonian image deblurrinq, SIAM J. Imaging Sci., 7 (2014), pp. 716—739.

[3] D. CHEN AND L. CHENG, Spatially adapted regularization parameter selection based on the local discrepancy function for poissonian image deblurring, Inverse Probl., 28 (2012), p. 015004.

[4] P. CONCUS, G. H. GOLUB, AND D. P. O'LEARY, A generalized conjugate gradient method for the numerical solution of elliptic partial differential equations, Computer Science Department, School of Humanities and Sciences, Stanford University, 1976.

[5] B. DONG, Sparse representation on graphs by tight wavelet frames and applications , Applied and Computational Harmonic Analysis, 42 (2017), pp. 452-479.

[6] B. DONG, Q. JIANG, C. LIU, AND Z. SHEN, Multiscale representation of surfaces by tight wavelet frames with applications to denoising, Applied and Computational Harmonic Analysis, 41 (2016), pp. 561-589.

[7] B. DONG, Z. SHEN, ET AL., MRA based wavelet frames and applications, IAS Lecture Notes Series, Summer Program on "The Mathematics of Image Processing", Park City Mathematics Institute,

19 (2010).

[8] B. DONG AND Y. ZHANG, An efficient algorithm for l_0 minimization in wavelet frame based image restoration, J. Sci. Comput., 54 (2013), pp. 350—368.

[9] M. A. T. FIGUEIREDO AND J. M. BIOUCAS-DIAS, Restoration of poissonian images using alternating direction optimization, IEEE Trans. Image Process., 19 (2010), pp. 3133-3145.

[10] S. FOUCART AND M.-J. LAI, Sparsest solutions of underdetermined linear systems via lq-minimization for $0 < q < 1$, Appl. Comput. Harmon. Anal., 26 (2009), pp. 395-407.

[11] M. GRASMAIR, M. HALTMEIER, AND O. SCHERZER, Sparse regularization with l_q penalty term, Inverse Probl., 24 (2008), p. 055020.

[12] G. H.ATTOUCH AND G. MICHAILLE, Variational analysis in Sobolev and BV spaces: applications to PDEs and optimization, vol. 17, SIAM, 2006.

[13] B. HE AND X. YUAN, Convergence analysis of primal-dual algorithms for a saddle- point problem: from contraction perspective, SIAM J. Imaging Sci., 5 (2012), pp. 119— 149.

[14] —, On the o (1/n) convergence rate of the douglas-rachford alternating direction method, SIAM J. Numer. Anal., 50 (2012), pp. 700-709.

[15] P. R. A. S. HEDY ATTOUCH, JEROME BOLTE, Proximal alternating minimization and projection methods for nonconvex problems, an approach based on the kurdyka-lojasiewicz inequality, eprint arXiv: 0801.1780, 35 (2010), pp. 257—512.

[16] M. HONG, Z.-Q. LUO, AND M. RAZAVIYAYN, Convergence analysis of alternating direction method of multipliers for a family of nonconvex problems, in 2015 ICASSP, IEEE, 2015, pp. 3836-3840.

[17] T. LE, R. CHARTRAND, AND T. J. ASAKI, A variational approach to reconstructing images corrupted by poisson noise, J. Math. Imaging Vis., 27 (2007), pp. 257-263.

[18] M. MAKITALO AND A. FOI, optimal inversion of the anscombe transformation in lowcount poisson image denoising, IEEE transactions on image processing, 20 (2013), pp. 99-109.

[19] —, Optimal inversion of the generalized anscombe transformation for poissongaussian noise, IEEE Trans. Image Process., 22 (2013), pp. 91-103.

[20] M. MAKITALO AND S. FOI, ALES, Optimal inversion of the anscombe transformation in low-count poisson image denoising, IEEE Trans. Image Process., 20 (2011), pp. 99109.

[21] J. M. B.-D. MARIO A.T. FIGUEIREDO, Redtoration of possionian images using alternating direction optimization, IEEE Transactions on Image Processing, 20 (2010), pp. 3133-3145.

[22] R. R. MICHAEL ELAD, PEYMAN MILANFAR, Analysis versus synthesis in signal priors, Inverse Problems, 23 (2007), p. 947.

[23] N. MILA, Relationship between the optimal solutions of least squares regularized with l_0-norm and constrained by k-sparsity, Appl. Comput. Harmon. Anal., (2016). 2016.

[24] M. K. NG, R. H. CHAN, AND W.-C. TANG, A fast algorithm for deblurring models with neumann boundary conditions, SIAM J. Sci. Comput., 21 (1999), pp. 851-866.

[25] M. NIKOLOVA, Description of the minimizers of least squares regularized with l_0-norm. uniqueness of the global minimizer, SIAM J. Imaging Sci., 6 (2013), pp. 904-937.

[26]R. Nishihara, L. Lessard, B. Recht, A. Packard, and M. I. Jordan, A general analysis of the convergence of admm, arXiv preprint, (2015)

[27]J.-C. P. Patrick L Combettes, A douglas-rachford splitting approach to nonsmooth convex variationl signal recovery, IEEE Journal of Selected Topics in Signal Processing, 1 (2007), pp. 564-574.

[28]N. Pustelnik, C. Chaux, and J.-C. Pesquet, Parallel proximal algorithm for image restoration using hybrid regularization, IEEE Trans. Image Process., 20 (2011), pp. 2450-2462.

[29]R. T. Rockafellar and R. J.-B. Wets, Variational analysis, vol. 317, Springer Science & Business Media, 2009.

[30]M. E. Sangnam Nam, Mike E Davies and R. Gribonval, The cosparse analysis model and algorithms, Applied and Computational harmonic Analysis, 34 (2013), pp. 30-56.

[31]S. Setzer, G. Steidl, and T. Teuber, Deblurring poissonian images by split bregman techniques, J. Vis. Commun. Image R., 21 (2010), pp. 193-199.

[32]L. Shen, Y. Xu, and X. Zeng, Wavelet inpainting with the l_0 sparse regularization, Appl. Comput. Harmon. Anal., (2015). (2015). http://dx.doi.org/10.1016/j.acha.2015.03.001.

[33]L. A. Shepp and Y. Vardi, Maximum likelihood reconstruction for emission tomography, IEEE Trans. Med. Imaging, 1 (1982), pp. 113-122.

[34]T. T. Simon Setzer, Gabriele Steidl, Deblurring poissonian images by split bregman techniques, Journal of Visual Communication and Image Representation, 21 (2010), pp. 193-199.

[35]E. Soubies, L. Blanc-Feraud, and G. Aubert, A continuous exact l_0 penalty (CEL0) for least squares regularized problem, SIAM J. Imaging Sci., 8 (2015), pp. 1607-1639.

[36]Y. Wang, W. Yin, and J. Zeng, Global convergence of admm in nonconvex nonsmooth optimization, arXiv preprint arXiv:1511.06324, (2016).

[37]W. Y. Y. Wang and J. Zeng, Global convergence of admm in nonconvex nonsmooth optimization, UCLA CAM Report, (2015), pp. 15-62.

[38]Y. Z, N. F, D. F, W. A, L. G, and H. J, Simulation tools for two-dimensional experiments in x-ray computed tomography using the forbild head phantom, Physics in Medicine and Biology, 57 (2012), pp. 237-52.

[39]H. Zhang, Y. Dong, and Q. Fan, Wavelet frame based poisson noise removal and image deblurring, Signal Processing, 137 (2017), pp. 363-372.

[40]X. Zhang, Y. Lu, and T. Chan, A novel sparsity reconstruction method from poisson data for 3d bioluminescence tomography, Journal of Scientific Computing, 50 (2012), pp. 519-535.

[41]Y. Zhang, B. Dong, and Z. Lu, l_0 minimization for wavelet frame based image restoration, Math. Comput., 82 (2013), pp. 995-1015.

[42]H. Z.WANG, A.C. BOVIK and E.P.SIMONCELLI, Image quality accessment: from error visibility to structual similarity, IEEE transcations on imaging processing, 13 (2004), pp. 600-612.