A REWEIGHTED $\ell^2$ METHOD FOR IMAGE RESTORATION WITH POISSON AND MIXED POISSON-GAUSSIAN NOISE

JIA LI AND ZUOWEI SHEN AND RUJIE YIN AND XIAOQUN ZHANG

Jia Li
Department of mathematics, National University of Singapore
Block S17, 10 Lower Kent Ridge Road
Singapore 119076

Zuowei Shen
Department of mathematics, National University of Singapore
Block S17, 10 Lower Kent Ridge Road
Singapore, 119076

Rujie Yin
Zhiyuan College, Shanghai Jiao Tong University
800, Dongchuan Road
Shanghai, China, 200240

Xiaoqun Zhang
Department of Mathematics, MOE-LSC and Institute of Natural Sciences
Shanghai Jiao Tong University
800, Dongchuan Road
Shanghai, China, 200240

(Communicated by the associate editor name)

Abstract. We study weighted $\ell^2$ fidelity in variational models for Poisson noise related image restoration problems. Gaussian approximation to Poisson noise statistic is adopted to deduce weighted $\ell^2$ fidelity. Different from the traditional weighted $\ell^2$ approximation, we propose a reweighted $\ell^2$ fidelity with sparse regularization by wavelet frame. Based on the split Bregman algorithm introduced in [22], the proposed numerical scheme is composed of three easy subproblems that involve quadratic minimization, soft shrinkage and matrix vector multiplications. Unlike usual least square approximation of Poisson noise, we dynamically update the underlying noise variance from previous estimate. The solution of the proposed algorithm is shown to be the same as the one obtained by minimizing Kullback-Leibler divergence fidelity with the same regularization. This reweighted $\ell^2$ formulation can be easily extended to mixed Poisson-Gaussian noise case. Finally, the efficiency and quality of the proposed algorithm compared to other Poisson noise removal methods are demonstrated through denoising and deblurring examples. Moreover, mixed Poisson-Gaussian noise tests are performed on both simulated and real digital images for further illustration of the performance of the proposed method.

2010 Mathematics Subject Classification. Primary: 65K10,65F22; Secondary: 94A08, 65T60.
Key words and phrases. Poisson noise, Poisson-Gaussian mixed noise, inverse problem, regularization, splitting algorithm, framelets.
1. Introduction. We consider an imaging system whose output data is a vector \( f \in (\mathbb{R}^+)^M \) and the true underlying image is \( u = (u_i)_{i=1}^N \in (\mathbb{R}^+)^N \). The observation model is often described by

\[
\tag{1}
 f = Au
\]

where \( A \) denotes a linear operator from \((\mathbb{R}^+)^N\) to \((\mathbb{R}^+)^M\). In this paper, we focus on variational image restoration from observation \( f \) contaminated by Poisson or mixed Poisson-Gaussian noise. Typically, variational models for image restoration are composed of two terms, one is a data fidelity term and the other is a regularization term for modeling a priori knowledge on unknown images. In general, the data fidelity term keeps the true image \( u \) close enough to the input data \( f \), so that the solution is meaningful. According to different noise statistics, the fidelity term takes different forms. For instance, it is well known that the additive white Gaussian noise (AWGN) can be suppressed by the least square fidelity. Such fidelity is mostly considered in literature for its good characterization of system noise. On the other hand, non-Gaussian types of noise are also encountered in real images. An important variant is Poisson noise, which is generally observed in photon-limited images, such as fluorescence microscopy, emission tomography and images with low exposure regions. Due to its importance in medical imaging and the wide commercial application, a vast amount of literature in image restoration is devoted to problems encountering Poisson noise (see [2] and the references therein). Suppose that the clean data \( u_i \) is corrupted by Poisson noise, then the probability of observing \( y \) is

\[
\tag{2}
P(y) = \frac{e^{-u_i} u_i^y}{y!}, \quad y = 0, 1, 2, \ldots
\]

where \( u_i \) is the expected value and the variance of random counts. Based on the statistics of Poisson noise, a generalized Kullback-Leibler (KL)-divergence [11] fidelity can be derived (see section 2.3.1 below) which is usually used in Poisson image restoration. However, difficulties in computational stability and efficiency arise due to the logarithm function that appears in the KL-divergence function. Moreover, in the case of image acquired with CCD camera, the mixture of Poisson and read-out noise (modeled as AWGN) is a more appropriate model, although it is seldom considered in literature. In [25], a Stein’s unbiased risk estimator (SURE) based on Poisson-Gaussian statistics is constructed in wavelet transform for mixed Poisson-Gaussian noise denoising, but the construction of estimators is complicated that it cannot be easily extended for more general image restoration problem such as deblurring, where the Poisson noise is added to the signal \( f = Au \) and the difficulty comes from the additional linear operator \( A \). Recently, Gong et al. proposed in [23] a universal \( \ell^1 + \ell^2 \) fidelity term for mixed or unknown noise and achieve encouraging numerical results, however the statistical analysis of the model remains unclear.

Besides the fidelity term, we need a regularization term to control noise and artifacts in reconstructed images. A simple but effective idea is to urge a sparse representation in some transform domain for the recovered image. For a large class of applications [15, 8, 36], penalizing the \( \ell^1 \) norm of the transform coefficients of a signal leads to such a sparse solution, as largely illustrated in compressed sensing community in recent years. The choice of transform is crucial to obtain a reasonable solution, and one popular choice is total variation proposed by Rudin-Osher-Fatemi [29]. Another choice are framelets (also called wavelet frame), which have been shown useful and superior to total variation for different restoration tasks, see [13, 10] and the references therein. Recently, the relation of sparse framelets...
representation and total variation has been revealed in [10], and theoretically total variation can be interpreted as a special form of $\ell^1$ framelets regularization. Therefore, we only consider framelets regularization in this paper.

The main contribution of this paper is introducing a reweighted least square fidelity, that well approximates KL-divergence, and developing an efficient iterative algorithm solving the optimization problem with framelets regularization. In fact, the idea of approximating KL-divergence by a fixed weighted least-square has been used for a long time in medical imaging reconstruction and image deblurring [2]. Recently, such a technique is restudied by Stagliano et al. [34], where the KL-divergence fidelity term is approximated by a weighted least square involving the unknown image. Combining with regularization, a scaled gradient projection (SGP) method with line search technique has been applied to solve the resulted problem. However, the efficiency of the algorithm involves the decomposition of a specific smooth total variation regularization and line searching, which can not be easily generalized to other sparse regularization. We have a similar approximation, but interpret it as a Gaussian approximation to the Poisson distribution of the same variance and the variance is dynamically estimated using an iterative algorithm based on the split Bregman iteration [22]. The original split Bregman iteration is ideal for solving $\ell^2$ fidelity based $\ell^1$ regularization, where the $\ell^1$ minimization is obtained by a soft shrinkage, promising the efficiency of our modified version. In addition, the solution obtained by our algorithm has the same reweighted $\ell^2$ energy and framelets sparsity as that of the minimizer of the KL-divergence fidelity using the same regularization. Furthermore, the model immediately gives a reasonable extension to the mixed Poisson-Gaussian noise case with a small modification on the weight.

This paper is organized in three consecutive parts. In the first part, we review the background of variational restoration models, the split Bregman method and the existing Poisson fidelities. Then we present the reweighted $\ell^2$ fidelity for Poisson noise through Gaussian approximation and maximizing likelihood (ML) and then build the corresponding image restoration models and algorithms. We also provide an analysis of the proposed models and algorithms which reveals the connection to the classical KL-divergence model. The models are further extended to the mixed noise case, where both Poisson noise and Gaussian white noise present. The last part is dedicated to numerical simulation where other fidelity terms and models are compared numerically to ours for Poisson image denoising/deblurring and the mixed Poisson-Gaussian denoising/deblurring cases.

2. Background.

2.1. Variational image restoration model. The observation model in consideration takes the following generic form

$$f = Au + c + \epsilon$$

(3)

where $c$ is a fixed background image vector in $(\mathbb{R}^+)^M$ and possibly 0, and $\epsilon$ is the noise perturbation. For additive white Gaussian noise setting, it is generally assumed that $c = 0$ and $\epsilon \sim \mathcal{N}(0, \sigma^2 I_n)$, i.e. each pixel has independent normal distribution of mean 0 and variance $\sigma^2$. The operator $A$ may come from the imaging system such as the point spread function (PSF), the exterior disturbance such as motion blur or a combination of the two sources. In this paper, we mainly consider denoising and deblurring, where $A$ is either an identity or a convolution operator.
Image restoration by variational model usually takes the following form:

\[
\min_u F(u) + \lambda G(u)
\]

where \(F(u)\) and \(G(u)\) are the fidelity term and the regularization term respectively, and \(\lambda\) is a positive scaling factor. The fidelity \(F(u)\) is derived using classical maximum a posterior probability (MAP) estimation, for AWGN, \(F(u) \propto \exp(-\frac{\|Au-f\|^2}{2\sigma^2})\); The regularization \(G(u)\) is designed based on a priori assumptions on \(u\). As introduced previously, it is nowadays standard technique to penalize the \(\ell^1\) norm of representation coefficients in a transform domain as the sparsity constraint on the coefficients. Therefore, the following variational model (5) is derived from (4) for image restoration for AWGN:

\[
\min_u \frac{1}{2}\|Au - f\|^2 + \lambda \|Du\|_1
\]

where \(\|\cdot\|_1\) denotes the usual \(\ell^1\) vector norm and \(D\) is a linear transform, such as discrete gradient used in total variation [29], Fourier transform, local cosine transforms, wavelet transforms.

2.2. **Split Bregman algorithm.** The split Bregman method introduced in [22] is designed to solve the variational model taking the form as (5). More generally, we consider the minimization problem

\[
\min_u F(u) + \lambda \|Du\|_1
\]

where \(F(u)\) is a convex function representing fidelity, such as \(F(u) = \frac{1}{2}\|Au - f\|^2\) as in the case of AWGN. By introducing an auxiliary variable \(d = Du\), the following alternating split Bregman scheme solves (6) : for \(\mu > 0\),

\[
\begin{align*}
  u_{k+1} &= \arg\min_u F(u) + \frac{\mu}{2}\|Du - d_k + b_k\|_2^2 \\
  d_{k+1} &= \arg\min_d \lambda \|d\|_1 + \frac{\mu}{2}\|Du_{k+1} - d + b_k\|_2^2 \\
  b_{k+1} &= b_k + Du_{k+1} - d_{k+1}
\end{align*}
\]

The efficiency of this algorithm relies on the closed form solution of the second subproblem. In fact, \(d_{k+1}\) is given by the so-called soft-shrinkage operator

\[
d_{k+1} = \text{sign}(Du_{k+1} + b_k) \cdot \max(|Du_{k+1} + b_k| - \lambda/\mu, 0)
\]

where each operation is componentwisely performed. The nonsmooth optimization problem (6) is thus decomposed into three subproblems; Furthermore, the first subproblem is easy to solve when \(F(u)\) is in quadratic form.

This method has been demonstrated to be very efficient for \(\ell^1\) type of minimization in a large variant of applications. In a more general setting, it is shown to be equivalent to classical Douglas-Racheford and alternating direction multiplier method (ADMM) [16, 30], and the convergence analysis is given in this framework. In [9], Cai et al. studied the application of such an algorithm for framelets based image restoration and provide a detailed convergence proof. Further approximation of the quadratic subproblem was considered in [37] in order to maximally decouple the subproblems. To take advantage of the efficiency of split Bregman method for Poisson noise related image restoration problem, we use the model (6) and introduce a reweighted \(\ell^2\) fidelity term for the case of Poisson noise.
2.3. Poisson noise related data fidelities. We assume that the observation \( f \) is corrupted by Poisson noise (see (2)), i.e.

\[
f \sim \text{Poisson}(Au + c)
\]

with the following assumptions on linear operator \( A \) and \( c \) as in [32, 3, 7]:

**Assumption 1.**
- **The observation data** \( f > 0 \), and the underlying true image \( u \geq 0 \).
- **The linear operator** \( A \) satisfies the following conditions:
  \[
  A_{ij} \geq 0; \quad \sum_i A_{ij} > 0, \forall j; \quad \sum_j A_{ij} > 0, \forall i.
  \]
  where \( A_{ij} \) is the \((i,j)\) element of the imaging matrix \( A \).
- **The fixed background image** \( c > 0 \).
- **The noise on each pixel is independent.**

For most image restoration models, \( A \) is the convolution or line integral operator, and these conditions can be easily fulfilled without loss of generality. The first assumption is used to avoid model deficiency as in KL-divergence and reweighted \( \ell_2 \). The second assumption implies that for \( u \geq 0 \), \( Au = 0 \) if and only if \( u = 0 \).

Given \( A \), \( u \) and \( c \), we have the likelihood of observing \( f \)

\[
\mathbb{P}(f|Au + c) = \prod_{i=1}^{M} \frac{(Au + c)^{f_i} e^{-(Au+c)_{i}}}{f_{i}!}
\]

where \((Au + c)_i\) denotes the \(i\)th element of \( Au + c \). By the property of the Poisson distribution, we have the expectation (mean) and the variance \( f \) are

\[
E(f|Au + c) = \text{Var}(f|Au + c) = Au + c
\]

Before presenting the proposed reweighted \( \ell_2 \) fidelity, we first review two existing fidelities for Poisson statistics for comparison.

2.3.1. KL-divergence. The most popular fidelity for Poisson noise is the generalized Kullback-Leibler (KL)-divergence fidelity [11], which can be derived directly by maximum likelihood (ML) method. This is equivalent to minimizing the negative log likelihood of (10), i.e.

\[
\min_{u \geq 0} \log \mathbb{P}(f|Au + c) = \min_{u \geq 0} f^T (Au + c - \log(f!)) - f^T \log (Au + c),
\]

where \(1\) is the vector with all 1 entries.

If we neglect the constant term \( \log (f!) \) which is unrelated to the unknown \( u \), we obtain the following fidelity term

\[
F(u) = 1^T (Au + c) - f^T \log (Au + c)
\]

Combining with sparse regularization and nonnegativity constraint on photon counts \( u \), we get the restoration model as follows,

\[
\begin{align*}
\min_{u \geq 0} & \quad 1^T (Au + c) - f^T \log (Au + c) + \lambda \|Du\|_1 \\
\end{align*}
\]

Generally, (14) is a difficult optimization problem because of the nonsmooth regularization term and the KL-divergence term. Optimization of KL-divergence fidelity with nonnegative constraint are typically solved by Expectation-Maximization (EM) algorithm [24]. It is known that the convergence of the EM algorithm is slow and it may introduce so-called “checkboard effect” [3, 34, 7]. In [7], the authors proposed
a two-step iteration method called EM-TV for solving (14) when $\mathcal{D}$ is a discrete differential operator. We change the TV regularization to framelets and rename it as EM+$\ell^1$ algorithm for later discussion and numerical comparison with our scheme. The algorithms is described as follows,

\[
\begin{aligned}
    u_{k+1} &= u_k A^* (\frac{f}{Au_k} + c) \quad \text{(EM step)} \\
    u_{k+1} &= \arg\min_{u \geq 0} \frac{1}{2} \left\| \frac{u - u_k + \frac{1}{2}}{\sqrt{u_k}} \right\|_2^2 + \lambda \left\| \mathcal{D} u \right\|_1 \quad \text{(\ell^1 step)}
\end{aligned}
\]

This algorithm and its variants have been proved to be efficient for Poisson noise removal in PET and nanoscopy image deconvolution [7]. Under adequate conditions, the convergence of the algorithm with a damped parameter is provided in [6]. However, the conditions are also hard to verify in practice and we also need a subproblem solver for the second step when $\mathcal{D}$ is not an orthogonal operator.

In [31, 18], the optimization problem (14) is solved by applying directly the split Bregman method (7). We present here the algorithm proposed in [31] for solving (14) by inducing three extra variables to represent $d^{(1)} = Au, d^{(2)} = \mathcal{D} u$ and $d^{(3)} = u$ to handle the nonnegativity constraint. The overall scheme can be written as

\[
\begin{aligned}
    u_{k+1} &= \arg\min_{u} \|Au - d^{(1)}_k\|_2 + \|\mathcal{D} u - d^{(2)}_k\|_2^2 + \|u - d^{(3)}_k\|_2^2 \\
    d^{(1)}_{k+1} &= \arg\min_{d^{(1)}} 1^T (d^{(1)} + c) - f^T \log (d^{(1)} + c) + \frac{1}{2\gamma} \|Au_k + d^{(1)} - b^{(1)}_k\|_2^2 \\
    d^{(2)}_{k+1} &= \arg\min_{d^{(2)}} \|\mathcal{D} u_k + d^{(2)} - b^{(2)}_k\|_2^2 \\
    d^{(3)}_{k+1} &= \arg\min_{d^{(3)}} \frac{1}{2\gamma} \|u_k - d^{(3)}_k\|_2^2 \\
    b^{(1)}_{k+1} &= b^{(1)}_k + Au_k + d^{(1)}_{k+1} - d^{(1)}_k \\
    b^{(2)}_{k+1} &= b^{(2)}_k + \mathcal{D} u_k + d^{(2)}_{k+1} - d^{(2)}_k \\
    b^{(3)}_{k+1} &= b^{(3)}_k + u_k + d^{(3)}_{k+1} - d^{(3)}_k
\end{aligned}
\]

\[2.3.2. \text{Anscombe transform.} \] Another well-known technique of Poisson denoising is the Anscombe transform [1] used in image denoising, when the linear operator $\mathcal{A}$ in (9) is the identity. The Anscombe transform is defined as the following non-linear transform

\[
\mathcal{A} : x \mapsto 2 \sqrt{x + \frac{3}{8}}
\]

If $x$ is a random variable that obeys the Poisson distribution with mean and variance $\tau$, the transformed random variable $\mathcal{A} x$ follows an approximated standard Gaussian distribution $\mathcal{N} (\tau, 1)$. Apply the Anscombe transform on $f \sim P(u + c)$, and let $\tilde{f} := \mathcal{A} f, \tilde{u} := \mathcal{A} (u + c)$, then $\tilde{f}$ follows approximately a normal distribution $\mathcal{N} (\tilde{u}, 1)$. Hence, the usual least square fidelity term can be applied for $\tilde{f}$ and we obtain the following denoising model

\[
\tilde{u}^* = \arg\min_{\tilde{u} \geq 0} \frac{1}{2} \|\tilde{u} - \tilde{f}\|_2^2 + \lambda \|\mathcal{D} \tilde{u}\|_1
\]

and the final image is given by $u^* = \mathcal{A}^{-1} \tilde{u}^* - c$. The regularization used in the above model (18) forces $\tilde{u}$ to be sparse in the transformed domain. This is roughly
equivalent to requiring sparsity on $u$ in the transform domain since the Anscombe transform (17) is monotone increasing and keeps the order of magnitude of elements in $u$. Due to its non-linear property, Anscombe transform is generally applied to image denoising models like (18), and its adaption to image deblurring is not easy.

3. **Weighted least square for Poisson noise image restoration.**

3.1. **Weighted least square fidelity.** We consider the case of Poisson noise of the observation model (9). Let

$$
\epsilon = f - Au - c,
$$

which can be interpreted as an additive perturbation noise. Given $Au$ and $c$, the conditional expectation of $\epsilon$ is

$$
E(\epsilon|A,u,c) = E(f|A,u,c) - Au - c = 0
$$

and the conditional variance is

$$
\text{Var}(\epsilon|Au+c) = \text{Var}(f|Au+c) = Au + c.
$$

We approximate $\epsilon$ by additive Gaussian noise with the same mean 0 and variance $\Sigma = \text{diag}(Au+c)$, i.e.

$$
P(\epsilon|A,u,c) \simeq \exp\left\{ -\frac{1}{2} (f - Au - c)^T \Sigma^{-1} (f - Au - c) \right\}
$$

Here $\Sigma$ is a diagonal matrix because of the independence assumption on the noise $\epsilon$. We take the negative log of the normal distribution (20),

$$
-\log P(\epsilon|A,u,c) \propto \frac{1}{2} (f - Au - c)^T \Sigma^{-1} (f - Au - c),
$$

and set (21) as the fidelity term, which is equivalent to maximum likelihood.

Let $\|x\|^2_Q = x^T Q x$ be the weighted $\ell^2$ norm of a vector $x \in \mathbb{R}^N$ with respect to a symmetric positive definite matrix $Q$. Then (22) can be reformulated as

$$
F(u) = \|Au + c - f\|^2_{\Sigma^{-1}} = \|Au + c - f\|^2_{\text{diag}(Au+c)^{-1}} = \left\| \frac{Au + c - f}{\sqrt{Au + c}} \right\|^2_2
$$

Both the division and the square root operator in the $\ell^2$ norm on the right hand side of (22) are element-wise. The positive definitiveness of $\Sigma^{-1}$ is guaranteed by the assumptions that $c > 0$ and $Au \geq 0$ for $u \geq 0$.

3.2. **Framelets regularization.** A countable set $X \subset L^2(\mathbb{R})$ is called a tight frame of $L^2(\mathbb{R})$ if

$$
f = \sum_{h \in X} \langle f, h \rangle h \quad \forall f \in L^2(\mathbb{R}),
$$

where $\langle \cdot, \cdot \rangle$ is the inner product of $L^2(\mathbb{R})$. Given a finite collection of functions $\Psi = \{\psi_1, \psi_2, ..., \psi_m\}$, define $X = \{\psi_{n,k,l} = 2^{n/2} \psi_l(2^n \cdot - k), 1 \leq l \leq m, n, k \in \mathbb{Z}\}$. If $X$ satisfies the condition of tight frame, then $X$ is called a wavelet tight frame and $\psi_{n,k,l}$ is called framelet. We choose framelet basis as a representation basis of images because its redundancy promotes sparse representation of images, as studied in [12, 13], and its multi-resolution property allows fast algorithm implementation. Generally, a multi-resolution analysis (MRA) based wavelets can be generated by the unitary extension principle (UEP) introduced in [28]. The wavelet tight frame
generated through UEP is based on a refinable function \( \phi = \psi_0 \) and the wavelet functions in \( \Psi \) satisfies:

\[
\hat{\psi}_l(2 \cdot ) = \hat{h}_l \hat{\phi}, \ l \in \{0, 1, 2, ..., m\}
\]

and

\[
\sum_{l=0}^{m} |\hat{h}_l(\xi)|^2 = 1; \quad \sum_{l=0}^{m} \hat{h}_l(\xi)\hat{h}_l(\xi + \pi) = 0
\]

where the sequence \( h_l \) is called the refinement mask of the wavelet tight frame system. By collecting all the refinement masks of wavelet tight frame system, we can generate the fast tight frame transform or decomposition operator \( W \). The matrix \( W \) is consisted of \( J + 1 \) sub-filtering operators \( W_0, W_1, ..., W_J \). Among them, \( W_0 \) is the low-pass filtering operator and the rest are high-pass filtering operators. According to UEP [28], the fast tight frame reconstruction operator is the adjoint operator \( W^T \) and we have \( W^T W = I \), i.e., \( W^T W u = u \) for any image \( u \). More details on discrete algorithms of framelets transforms can be found in [13].

3.3. Models and algorithms. Combining the fidelity \( F(u) \) in (22) with sparse framelets regularization and nonnegativity constraint, we have the following restoration model

\[
\min_u \left\{ \frac{1}{2} \left\| \frac{Au + c - f}{\sqrt{Au + c}} \right\|_2^2 + \lambda \|Wu\|_1 \right\}
\]

This formulation is also considered in [34] and a scale gradient projection method is applied to (24) for solving the nonlinear objective function, where a smoothed version of the \( \ell_1 \) regularization term was adopted for calculating the gradient.

To solve (24), we are interested in taking advantage of the weighted least square structure and utilizing split Bregman presented in Section 2.2. However, the fidelity term \( F(u) \) is not quadratic, because \( u \) appears both in the denominator and in the numerator of (22). Therefore, in order to solve the first optimization sub-problem in the split Bregman algorithm (7), we need to either approximate or directly solve the nonlinear square term.

A reasonable approximation of the unknown weight \( Au + c \), for example, can be the observed data \( f \), so that (22) is approximated by

\[
F(u) = \frac{1}{2} \left\| \frac{Au + c - f}{\sqrt{Au + c}} \right\|_2^2
\]

Using this simplification, efficient least square based methods can be applied to solve the sub-problem.

However, such an approximation is not accurate, especially when the observed data \( f \) is severely corrupted by noise. Hence we need a better approximation, which may be derived adaptively in an iterative algorithm. For any iterative algorithm, whose sequence of solution \( u_k \) converges to \( u^* \), when \( k \) is big enough, as \( u_k \) becomes stable, \( u_k \) serves as a more precise approximation of \( u^* \) than \( f \). Hence a better approximation to the formulation (24) is given as

\[
\min_u \left\{ \frac{1}{2} \left\| \frac{Au + c - f}{\sqrt{Au^k + c}} \right\|_2^2 + \lambda \|Wu\|_1 \right\}
\]

with \( k \) sufficiently large. We approximate the exact weight \((Au^* + c)^{-1}\) by \((Au_k + c)^{-1}\) in the algorithm. With such an idea in mind, we combine with the popular Split
Bregman iteration and derives a new algorithm called the reweighted $\ell^2$ algorithm with split Bregman for Poisson noise:

\[
\begin{align*}
    u_{k+1} &= \arg\min_{u \geq 0} \frac{1}{2} \left\| Au + c - f \right\|^2 + \frac{\mu}{2} \left\| W u - d_k + b_k \right\|^2 \\
    d_{k+1} &= \arg\min_d \lambda \|d\|_1 + \frac{\mu}{2} \left\| d - W u_{k+1} - b_k \right\|^2 \\
    b_{k+1} &= b_k + (W u_{k+1} - d_{k+1})
\end{align*}
\]

We describe the general method (27) in more detail in Algorithm 1. Note that the first step can be solved by a gradient method with a projection onto the nonnegative orthant. In practice, often a few iterations are sufficient to get a reasonably accurate result.

**Algorithm 1** Reweighted $\ell^2$ with Split Bregman for Poisson noise restoration Algorithm

**Step0.** Set the initial value, \( u_0 = f; d_0 = W f; b_0 = 0; k = 1; \Sigma_0 = Au_0 + c. \)

**Step1.**

\[
\text{while } \|u_{k-2} - u_{k-1}\|_2 > \delta \text{ or } k = 1 \text{ do}
\]

\[
\begin{align*}
    u_k &= \arg\min_{u \geq 0} \frac{1}{2} \left\| Au + c - f \right\|^2 + \frac{\mu}{2} \left\| W u - d_{k-1} + b_{k-1} \right\|^2 \\
    d_k &= \text{sign}(W u_k + b_{k-1}) \cdot \max(\|W u_k + b_{k-1}\| - \lambda/\mu, 0) \\
    b_k &= b_{k-1} + W u_k - d_k \\
    \Sigma_k &= \text{diag}(Au_k + c) \\
    k &= k + 1
\end{align*}
\]

**end while**

In the following, we justify the proposed algorithm by some intuitive analysis. Let \((u_k, d_k, b_k)\) be the sequence generated by the algorithm (27). If \((u_k, d_k, b_k)\) converges to \((\hat{u}, \hat{d}, \hat{b})\), then \(\hat{u}\) is a minimizer of (14), shown as follows. If \((u_k, d_k, b_k)\) generated by (27) converges to \((\hat{u}, \hat{d}, \hat{b})\), we immediately get

\[
A^T \tilde{\Sigma}^{-1} (A\hat{u} + c - f) + \mu W^T \hat{b} - \hat{q} = 0
\]

\[
\lambda \hat{p} - \mu \hat{b} = 0 \quad \text{with } \hat{p} \in \partial\|\hat{d}\|_1 \\
W \hat{u} = \hat{d}
\]

where \(\hat{q} \geq 0\) satisfies \(\langle \hat{q}, \hat{u} \rangle = 0\), and \(\tilde{\Sigma} = \text{diag}(A\hat{u} + c)\). Therefore, substituting \(\hat{b}\) by \(\lambda/\mu \hat{p}\), we have

\[
A^T \tilde{\Sigma}^{-1} (A\hat{u} + c - f) + \lambda W^T \hat{p} - \hat{q} = 0
\]

On the other hand, if \((\hat{u}, \hat{d}, \hat{b})\) is the minimizer of (14), the first order condition gives that

\[
0 = A^T \left( \frac{A\hat{u} + c - f}{A\hat{u} + c} \right) + \lambda W^T \hat{p} - \hat{q}
\]

where \(\hat{p} \in \partial\|\hat{d}\|_1\) with \(\hat{d} = W \hat{u}\) and \(\hat{q} \geq 0\) is a lagrangian multiplier such that \(\langle \hat{q}, \hat{u} \rangle = 0\). Since \(\hat{u}\) satisfies exactly this first order condition, it is a minimizer of (14).
3.4. **Extension to Poisson-Gaussian mixed noise.** Previously, we mainly discuss models with Poisson noise only. In real imaging systems, besides Poisson noise which characterizes the fluctuation in counting number of photons, there is other system-inherited noise that can be approximated by AWGN, as considered in [33]. The literature on Poisson-Gaussian mixed noise are limited despite of its importance. We now explore the extended application of the reweighted $\ell^2$ fidelity to the mixed Gaussian-Poisson noise case.

With a small modification of (22), the reweighted $\ell^2$ fidelity can be adapted to the mixed noise case. Let $\sigma^2$ be the variance of AWGN and the observed image $f$ has distribution

$$f \sim \mathcal{P}(Au + c) + \mathcal{N}(0, \sigma^2)$$

Approximating the likelihood of $f$ by normal distribution (20) with covariance matrix $\Sigma = \text{diag}(Au + c) + \sigma^2 I$, we obtain the new fidelity term as follows for mixed noise:

$$(30) \quad F(u) = \frac{1}{2} \left\| \frac{Au + c - f}{\sqrt{Au + c + \sigma^2 I}} \right\|_2^2$$

Combining with the sparse framelets regularization, we have the following restoration model

$$(31) \quad \min_{u \geq 0} \frac{1}{2} \left\| \frac{Au + c - f}{\sqrt{Au + c + \sigma^2 I}} \right\|_2^2 + \lambda \|Wu\|_1$$

The algorithm solving this model is the same as Algorithm 1 except adding $\sigma^2$ in the estimation and updating step of the covariance matrix $\Sigma$.

4. **Numerical results.** In this section, numerical experiments on pure Poisson noise model (24) and Poisson-Gaussian mixed noise models (31) are shown for synthesized and real digital photos to demonstrate the performance of the proposed Algorithm 1. Numerical results of denoising/deblurring algorithms discussed in section 2 are also presented for comparison. In all implementations the same framelets sparse regularization is used. We choose the piecewise linear B-spline wavelets with 1 level decomposition as the framelets basis. The corresponding 1-D discrete filter masks are $h_0 = [1, 1, 1]$, $h_1 = [-1, 2, -1]$, $h_2 = [2, 0, -2]$. The 2-D masks are generated as the tensor products of 1-D masks. The parameter $\lambda$ in the regularization term is always set to be 0.01 for pure image denoising and 0.001 for image deblurring in presence of noise. The parameter $\mu$ in the split Bregman iteration is simply set to be 1 for all numerical simulations.

In the numerical implementation, we find preconditioning on $\Sigma_k$ is necessary to ensure numerical stability. In particular, we enforce a lower bound $C_{\sigma} > 0$ on the diagonal entries of $\Sigma_k$, i.e. $\Sigma_k = \max\{\Sigma_k, C_{\sigma}\}$ to control its condition number. To improve the visual quality of the final result, in the postprocessing, we apply a bilateral filter [35] $\mathcal{B}$ defined as

$$y[i, j] = \mathcal{B}(x) = \frac{1}{\sum_{p,q} w[i, j, p, q]} \sum_{p,q} w[i, j, p, q] x[p, q]$$

where $[i, j], [p, q]$ are indices of pixels of an image $x$,

$$w[i, j, p, q] = G_{\sigma_s}(\sqrt{(i - p)^2 + (j - q)^2})G_{\sigma_r}(x[i, j] - x[p, q])$$

and $G_{\sigma_s}, G_{\sigma_r}$ are Gaussian functions with variance $\sigma_s$ and $\sigma_r$ respectively.
Each denoised result \( \hat{u} \) are evaluated quantitatively by the peak signal-to-noise ratios (PSNR) value defined by

\[
\text{PSNR}(u, \hat{u}) = 10 \log_{10} \frac{N(u_{\text{max}} - u_{\text{min}})^2}{\| u - \hat{u} \|_2^2},
\]

using the ground truth image \( u \), where \( u_{\text{max}} \) and \( u_{\text{min}} \) are its maximal and minimal pixel values respectively and \( N \) is the total number of pixels of the image.

4.1. Poisson denoising. In the following, we compare the visual quality of the restoration result using denoising methods with three different fidelities: the reweighted \( \ell^2 \) method (27) with Algorithm 1, the KL-divergence model (14) with the algorithm (15) as well as algorithm (16) and the Anscombe transform model (18). The noisy images in our test are simulated as follows. The clean images are first rescaled to an intensity range from 0 to 120; then the Poisson noise is added in Matlab using the function \texttt{poissrnd}. The parameters in the first three algorithms are set to the same instead of optimized respectively (especially \( \lambda = .05 \)), hence the energy function they minimize are consistent to each other. For the Anscombe transform model, the parameters are tuned (\( \lambda = .11 \)) such that the best result is obtained, since the scale of the energy function is changed after the Anscombe transform of the initial image. For the denoised results, we do not apply the bilateral filtering in post-processing to remove artifacts, so that the comparison is clear and fair.

In Figure 1, the simulation results of the test image cameraman are shown for all the algorithms considered previously in Section 2. The denoising results of the reweighted \( \ell^2 \) model and that of the KL-divergence model are visually the same and their PSNR are comparable to each other, the difference in pixel value of these two results is shown in Figure 1. The Anscombe transform model gives a slightly better result than the other two models, mainly in the low pixel value region. The difference of results in pixel value of the Anscombe transform algorithm (18) and that of the KL-divergence based split Bregman method (16) is shown as well, where the biggest difference is about 6. More quantitative comparisons on different test images are shown in Table 1 and the three algorithms perform similarly.

<table>
<thead>
<tr>
<th># of (outer)</th>
<th>KL(SB)</th>
<th>EM</th>
<th>Anscombe</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 iterations</td>
<td>400</td>
<td>50 (with 5 inner iteration)</td>
<td>100</td>
</tr>
<tr>
<td>cameraman</td>
<td>30.54</td>
<td>30.59</td>
<td>30.57</td>
</tr>
<tr>
<td>fruit</td>
<td>30.71</td>
<td>30.71</td>
<td>30.71</td>
</tr>
<tr>
<td>boat</td>
<td>29.19</td>
<td>29.18</td>
<td>29.18</td>
</tr>
<tr>
<td>goldhill</td>
<td>29.24</td>
<td>29.24</td>
<td>29.24</td>
</tr>
</tbody>
</table>

Table 1. PSNR value of the denoising result.

The similarity between these algorithms can be explained by the connections between their convergent solutions. In the Poisson noise only case, the operator \( A \) is identity, therefore the objective function is strictly convex. Numerically, we can observe the convergence of our algorithm. According to the previous analysis, it converges to the unique minimizer \( u^* \) of the KL-divergence fidelity model with the same framelet sparsity regularization. Consequently, the two solutions should have the same re-weighted \( \ell_2 \) energy and \( \ell_1 \) sparsity of the framelets coefficients, where the weight in the \( \ell_2 \) norm is fixed to \( u^* + c \) and \( u^* \) is approximated by the
12 Jia Li and Zuowei Shen and Rujie Yin and Xiaoqun Zhang

cameraman noisy

KL(SB)+l1, PSNR:30.54dB

re-weighted l2, PSNR: 30.59dB

Anscombe+l1, PSNR: 30.69dB

EM+l1, PSNR:30.57dB

Figure 1. Denoising results of simulated noisy images with peak intensity 120. From left to right, up to down: noisy image, reweighted ℓ² method by Algorithm 1, Anscombe transform (18), KL-divergence model using split Bregman (16) and KL-divergence model with EM-ℓ¹ algorithm (15). The last image shows the difference in pixel value of Anscombe+ℓ¹ result and KL(SB) result.

split Bregman result. This can be shown numerically in Figure 2; the result of our algorithm with regularized weight \( \tilde{\Sigma}_k \) has weighted ℓ² energy plus ℓ¹ sparsity close to a minimizer of the KL-divergence fidelity model calculated by the Split Bregman algorithm (16) as well as the solution given by the EM algorithm (15). The three solutions are not identical, though, because each algorithm is embedded with different regularization and generates numerical error differently. In Figure 2, the KL-divergence plus ℓ¹ sparsity energy is also plotted. Also, our proposed
scheme and the EM algorithm converge faster than the split Bregman method with respect to the number of iterations.

\[ E(u) = \frac{1}{2} \| u + c - f \|_2^2 + \lambda \| W u \|_1 \]

Figure 2. Left: KL-divergence model (14) energy evolution of Algorithm 1, (15) and (16). Right: reweighted \( \ell^2 \)-norm energy model (14) evolution of Algorithm 1, (15) and (16).

4.2. Deblurring from Poisson data. The deblurring case is in general more difficult than denoising. We only compare our reweighted \( \ell^2 \) method and the KL-divergence model (14) since Ancombe transform cannot be easily applied in this case. For this application, we compare our proposed Algorithm 1 with two different algorithms EM+\( \ell^1 \) algorithm (15) and KL-split Bregman algorithm (16) for the KL-divergence model (14). In this subsection, blurred and noisy images are simulated in the following way. The clean images are first rescaled to an intensity range from 0 to 1200; then they are corrupted by a disk blurring kernel of radius 3 with symmetric boundary condition; finally, the Poisson noise is simulated in the same way as in the denoising case.

The main computation cost of deblurring Algorithm 1 is that \( u_{k+1} \) is solved by conjugate gradient method instead of a direct inversion in the first step of the iteration. As shown in Figure 3, the result of the KL-divergence fidelity algorithm is more blurry in contrast to the result of weighted \( \ell^2 \) fidelity, although they have similar PSNR, see Table 2 for more comparisons. Table 2 shows that our re-weighted \( \ell^2 \) fidelity method (26) and KL-split Bregman algorithm (16) have higher PSNR value than EM+\( \ell^1 \) algorithm (15). With the postprocess of bilateral filter, our reweighted method (26) can obtain the best PSNR value among all the algorithms.

<table>
<thead>
<tr>
<th></th>
<th>reweighted ( \ell^2 ) with bilateral filter</th>
<th>reweighted ( \ell^2 )</th>
<th>KL-Split Bregman</th>
<th>EM+( \ell^1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>goldhill</td>
<td>26.82</td>
<td>26.76</td>
<td>26.74</td>
<td>26.62</td>
</tr>
<tr>
<td>cameraman</td>
<td>26.64</td>
<td>26.33</td>
<td>26.41</td>
<td>25.84</td>
</tr>
<tr>
<td>boat</td>
<td>25.70</td>
<td>25.59</td>
<td>25.39</td>
<td>25.08</td>
</tr>
</tbody>
</table>

Table 2. PSNR value of the deblurring result.

In Figure 5, we compare the energy evolution of different algorithm. According to our previous argument, Algorithm 1, if converge, converges to the solution of...
Figure 3. Deblurring results of simulated noisy-blurred "Gold-hill" image with peak intensity 1200. First row from left to right: ground truth image, blurry image, EM+ℓ₁ model (15). Second row from left to right: KL-split Bregman algorithm (16), direct reweighted ℓ² method (26) and reweighted method (26) with bilateral filter post-process.

the model (14). We thus compare the KL-divergence model $E(u) = 1^T (Au + c) - f^T \log(Au + c) + \lambda \|Wu\|_1$ of the three algorithms (EM-ℓ¹, KL-split Bregman, and the proposed Algorithm 1). We can see that the energy for our proposed Algorithm 1 has the fastest speed of convergence. The evolution of weighted ℓ²-norm energy $E(u) = \frac{1}{2} \left\| \frac{Au + c - f}{\sqrt{Au + c}} \right\|^2_2 + \lambda \|Wu\|_1$ is also shown in the right figure of Fig. 5, where $u^*$ is set as the ground truth image. We note that the EM+ℓ¹ algorithm converges to a solution with slightly higher energy.

To reach the same PSNR level of the result given by the KL-Split Bregman algorithm after 50 iterations, the proposed Algorithm 1 needs much less iterations (less than 5 iterations), See Fig. 6.

In terms of computational time of deblurring, the reweighted algorithms (1) needs around 100s while the EM-ℓ¹ around 240s for the same relative error stopping criteria. As a result, our proposed reweighted method (26) can outperform the EM-ℓ¹ in terms of visual quality, PSNR value, convergence speed and computational time. Moreover, compared with the KL-split Bregman algorithm (16), our proposed Algorithm 1 has faster convergence speed in terms of both the KL-energy and reweighted ℓ²-norm energy.

4.3. Mixed Poisson-Gaussian noise restoration. In this section, we test the extension of reweighted ℓ² model to mixed Poisson-Gaussian noise for some synthesized data and real photos. The synthesized mixed noised image is simulated by adding first Poisson noise to a rescaled image with peak intensity 120 and the
Figure 4. Deblurring results of simulated noisy-blurred "Camera-man" image with peak intensity 1200. First row from left to right: ground truth image, blurry image, EM+$\ell_1$ model (15). Second row: zoom-in images of the first row. Third row from left to right: KL-split Bregman algorithm (16), direct reweighted $\ell^2$ method (26) and reweighted method (26) with bilateral filter post-process. Fourth row: zoom-in images of the third row.
Figure 5. Top: KL–divergence \( E(u) = 1^T(Au + c) - f^T \log(Au + c) + \lambda \|Wu\|_1 \) energy evolution of "Goldhill" image for (15), (16) and Algorithm 1. Bottom: reweighted \( \ell^2 \)-norm energy \( E(u) = \frac{1}{2} \left\| \frac{Au+c-f}{\sqrt{Au+c}} \right\|_2^2 + \lambda \|Wu\|_1 \) evolution of "Goldhill" image for (15), (16) and Algorithm 1.

Figure 6. PSNR evolution of "Goldhill" image for three algorithms: (15), (16) and Algorithm 1.

Gaussian noise of \( \sigma = 12 \). Fig. 7 shows the denoising result of model (31) with \( A \) being identity operator.

In addition, we perform the mixed denoising model on a digital photo taken in a low light environment with high ISO setting. Our denoising result is compared with the embedded noise reduction algorithm in Sony DSLR-A700 camera in Figure 8, and we can see that our result has less noise and better quality.
A reweighted $\ell^2$ method for image restoration with Poisson and mixed Poisson-Gaussian noise

Figure 7. Left: input noisy image with mixed Poisson-Gaussian noise, peak intensity: 120, PSNR: 18.59dB; right: denoised by model (31), PSNR: 27.42dB.

For deblurring case, we generate blurred image with peak intensity 1200 and then corrupt it with Poisson noise, Gaussian noise of $\sigma = 12$ consecutively. See Fig. 9 for the deblurring result of model (31).

5. Conclusion. In this paper, we studied weighted $\ell^2$ fidelity derived from Gaussian approximation of Poisson statistics. In solving denoising and deblurring models, we proposed a reweighted algorithm based on the split Bregman iteration. The solution of the proposed algorithm is the same as the minimizer of KL-divergence fidelity model (14) if the sequence converges. We rely on our numerical experiments to further verify the performance of our algorithm; the proposed algorithm gives competitive result with respect to classical fidelity terms in denoising and deblurring simulations. In addition, the reweighted $\ell^2$ model can be easily extended to the mixed Poisson-Gaussian noise case. As shown by numerical results, the reweighted $\ell^2$ with the split Bregman framework is promising in a wide perspective.

Both the algorithm proposed and the analysis shown in this paper are an attempt to study and link different variational models in image processing. In this study, we only focused on the fidelity term, using a simple sparsity term as regularization, without utilizing the full image structure for the analysis of Poisson statistics. One possible extension of our current work would be incorporating the image structure and utilizing the spatial correlation of the pixel values to build a more precise statistical model, hence a more precise fidelity term, of the Poisson data. This might substantially improve the result as suggested by the improvement made of a post bilateral filtering process in our deblurring simulation.

Acknowledgments. Xiaoqun Zhang was partially supported by the NNSFC (Grant nos. 11101277, 91330102) and the Shanghai Pujiang Talent Program (Grant no. 11PJ1405900). Zuowei Shen was partially supported by National University of Singapore (Grant no. MOE2011-T2-1-116).

REFERENCES

Figure 8. Top: real photo; middle: output of camera using the noise reduction option. bottom: denoised by model (31).


A reweighted $\ell^2$ method for image restoration with Poisson and mixed Poisson-Gaussian noise

Figure 9. left: input blurred image with mixed Poisson-Gaussian noise, $\sigma = 12$, peak intensity: 1200; right: deblurred by model (31).


Received 2013; revised 2015.