# A Convex Approach for Image Restoration with Exact Poisson-Gaussian Likelihood

Emilie Chouzenoux, Anna Jezierska, Jean-Christophe Pesquet, Hugues Talbot

## ▶ To cite this version:

HAL Id: hal-00922151

https://hal.archives-ouvertes.fr/hal-00922151v3

Submitted on 12 Aug 2015

# A Convex Approach for Image Restoration with Exact Poisson-Gaussian Likelihood *

Emilie Chouzenoux, Anna Jezierska, Jean-Christophe Pesquet, and Hugues Talbot

August 12, 2015

## Abstract

The Poisson-Gaussian model can accurately describe the noise present in a number of imaging systems. However most existing restoration methods rely on approximations of the Poisson-Gaussian noise statistics. We propose a convex optimization strategy for the reconstruction of images degraded by a linear operator and corrupted with a mixed Poisson-Gaussian noise. The originality of our approach consists of considering the exact, mixed continuous-discrete model corresponding to the data statistics. After establishing the Lipschitz differentiability and convexity of the Poisson-Gaussian neg-log-likelihood, we derive a primal-dual iterative scheme for minimizing the associated penalized criterion. The proposed method is applicable to a large choice of convex penalty terms. The robustness of our scheme allows us to handle computational difficulties due to infinite sums arising from the computation of the gradient of the criterion. We propose finite bounds for these sums, that are dependent on the current image estimate, and thus adapted to each iteration of our algorithm. The proposed approach is validated on image restoration examples. Then, the exact data fidelity term is used as a reference for studying some of its various approximations. We show that in a variational framework the Shifted Poisson and Exponential approximations lead to very good restoration results.

## 1 Introduction

The recovery of a target image in the presence of degradations (i.e. noise) has been extensively studied in the literature. Image reconstruction problems are often formulated into the Maximum *A Posteriori* (MAP) framework. The MAP estimator is computationally simple and offers significant flexibility in the choice of a prior, while the data fidelity term usually consists of the noise neg-log-likelihood or some approximation of it. A Gaussian noise model is commonly assumed in imaging due to the simple and intuitive form of the related data fidelity function, which is convex and Lipschitz differentiable. However, recent methodological progress has ensured that more sophisticated variational problems can be efficiently solved numerically. As a consequence, studies are no longer limited to the simplest Gaussian model. Some contributions concern various standard noise distributions, e.g. Poisson [4, 11, 20, 33, 44, 48, 67], impulsive [16] or multiplicative [5, 19, 31, 46] noise, as well as some mixed noise models, e.g. the mixture of Gaussian and impulsive noise [72, 73].

A growing interest in Poisson-Gaussian probabilistic models has recently been observed. The Poisson component is often related to the quantum nature of light and accounts for photon-counting principles in signal acquisition, whereas the Gaussian component is typically related

---

1

to noise present in the electronic part of the imaging system. Indeed, despite constant improvements, electronic noise usually cannot be neglected. In addition, the Poisson noise component cannot usually be accurately described through Gaussian statistics, especially in the context of low-count imaging. As a consequence, a mixed Poisson-Gaussian noise model is frequently considered in astronomy [9, 68], remote sensing [3], medicine [60] and biology [29]. There has been a growing interest for denoising problems involving images corrupted in this fashion [8, 12, 57]. Recent advances have been made in Poisson-Gaussian noise parameters estimation procedures [38, 47, 49]. However, the literature in deconvolution and reconstruction problems involving such noise model remains scarce. Among existing works, Benvenuto *et al.* [9] proposed a scaled gradient method and more recently Gil-Rodrigo *et al.* [41] developed an alternating minimization algorithm. An augmented Lagrangian method grounded on a Poisson approximation of the noise characteristics was proposed in [17], while a similar framework with a weighted squared $\ell_2$ norm noise approximation was proposed in [55]. So far, restoration strategies have relied on some approximations of the noise statistics, which may have been detrimental to the quality of the results. The use of approximations is motivated by the mathematical difficulties raised by the Poisson-Gaussian model. Indeed, the corresponding probability distribution has a discrete-continuous nature, and the expression of the associated log-likelihood function involves infinite sums. For simplification, one usually neglects either the Poisson or the Gaussian component, or performs an approximation of the Poisson-Gaussian model based on variance stabilization techniques [37, 59].

In this paper, we investigate the properties of the Poisson-Gaussian negative log-likelihood, showing that it is a convex Lipschitz differentiable function. Since the gradient of the Poisson-Gaussian neg-log-likelihood requires the computation of infinite series, we propose to approximate it using finite sums whose bounds depend on the current signal estimate. Next, we employ a proximal optimization method whose convergence is guaranteed even in the presence of summable numerical errors. The retained approach belongs to the recent class of primal-dual splitting algorithms (see [52] and the references therein) and can cope with the sum of a gradient Lipschitz term and possibly nonsmooth penalty terms. Such terms can model a wide range of prior information, e.g. range value constraints, criteria promoting sparsity in a frame, total-variation and more generally hybrid regularization functions [62].

Our main contributions in this paper are as follows:

- a proof of the Lipschitz differentiability of the Poisson-Gaussian neg-log-likelihood function;

- a proof of the convexity of the mixture of Poisson and Generalized-Gaussian neg-log-likelihood function;

- the establishment of new summation bounds, along with error control, for the infinite sums appearing in the gradient of the Poisson-Gaussian neg-log-likelihood;

- new tools for comparing the existing approximations of the Poisson-Gaussian model; and

- validation of the numerical performance of our restoration method applied to the model with data fidelity given by exact Poisson-Gaussian neg-log-likelihood and also by its several approximations.

The paper is organized as follows: Section 2 introduces the notation used in this work and investigates the Poisson-Gaussian model. In particular, convexity and Lipschitz-differentiability results are presented. Section 3 describes the proposed optimization framework. Next, our restoration approach is illustrated via experiments in Section 4. Finally, some conclusions are drawn in Section 5.

## 2  Degradation model

Let $y = (y_i)_{1 \leq i \leq Q} \in \mathbb{R}^Q$ be a vector of observations related to an original signal $\overline{x} \in [0, +\infty)^N$, which is degraded by a matrix $H$ in $[0, +\infty)^{Q \times N}$ (e.g., a convolutive model) and further corrupted with Poisson-Gaussian noise. More specifically, the observations $y$ are related to $\overline{x}$ by

$$y = z(\overline{x}) + w, \tag{1}$$

where $z(\overline{x}) = \left( z_i(\overline{x}) \right)_{1 \leq i \leq Q}$ and $w = (w_i)_{1 \leq i \leq Q}$ are realizations of mutually independent random vectors $Z(\overline{x}) = \left( Z_i(\overline{x}) \right)_{1 \leq i \leq Q}$ and $W = (W_i)_{1 \leq i \leq Q}$ featuring independent components. It is further assumed that, for every $i \in \{1, \ldots, Q\}$,

$$Z_i(\overline{x}) \sim \mathcal{P}([H\overline{x}]_i), \tag{2}$$
$$W_i \sim \mathcal{N}(0, \sigma^2), \tag{3}$$

where $[H\overline{x}]_i$ is the $i$-th component of vector $H\overline{x}$, and $\sigma \in (0, +\infty)$ is the standard deviation of the zero-mean Gaussian noise component. Hence, $y$ is a realization of a random vector $Y$ with probability density function given by the convolution of the components resulting from the Gaussian and Poisson distributions, i.e.

$$p_{Y|\overline{x}}(y) = \prod_{i=1}^{Q} \left( \sum_{n=0}^{+\infty} \frac{e^{-[H\overline{x}]_i}([H\overline{x}]_i)^n}{n!} \frac{e^{-\left( \frac{y_i - n}{\sqrt{2}\sigma} \right)^2}}{\sqrt{2\pi}\sigma} \right). \tag{4}$$

In the context of inverse problems, the original signal can be recovered by minimizing a penalized criterion:

$$\min_{x \in \mathbb{R}^N} \left( f(x) = \Phi(x) + \rho(x) \right), \tag{5}$$

where $\Phi$ is the so-called data fidelity term and $\rho$ is a regularization function incorporating *a priori* information, so as to guarantee the stability of the solution w.r.t. the observation noise. In the Bayesian framework, this allows us to compute the MAP estimate [30] of the original image. In this context, the data fidelity term is defined as the negative logarithm of $p_{Y|x}(y)$:

$$\Phi(x) = -\log(p_{Y|x}(y)) = \sum_{i=1}^{Q} \Phi_i([Hx]_i), \tag{6}$$

where, for every $i \in \{1, \ldots, Q\}$ and $u \in [0, +\infty)$,

$$\Phi_i(u) = -\log \left( \sum_{n=0}^{+\infty} \frac{u^n e^{-u}}{n!} \frac{e^{-\left( \frac{y_i - n}{\sqrt{2}\sigma} \right)^2}}{\sqrt{2\pi}\sigma} \right), \tag{7}$$

and the regularization term $\rho$ corresponds to the potential of the chosen prior probability density function of the target image.

Let $\mathbf{1}$ denote the vector of $\mathbb{R}^Q$ with all its components equal to 1. The gradient and Hessian of $\Phi$ on the positive orthant are respectively given by: for every $x \in [0, +\infty)^N$,

$$\nabla\Phi(x) = H^\top(\mathbf{1} - \xi(Hx)), \tag{8}$$
$$\nabla^2\Phi(x) = H^\top \text{Diag} \left( \eta_1([Hx]_1), \ldots, \eta_Q([Hx]_Q) \right) H, \tag{9}$$

where, for every $z = (z_i)_{1\leq i \leq Q} \in [0, +\infty)^Q$, $\xi(z) = \left(\xi_i(z_i)\right)_{1\leq i \leq Q}$, $\eta(z) = \left(\eta_i(z_i)\right)_{1\leq i \leq Q}$ with [4]

$$\xi_i(z_i) = s(z_i, y_i - 1)/s(z_i, y_i), \tag{10}$$

$$\eta_i(z_i) = \left(\left(s(z_i, y_i - 1)\right)^2 - s(z_i, y_i)s(z_i, y_i - 2)\right) / \left(s(z_i, y_i)\right)^2, \tag{11}$$

and, for every $(a, b) \in \mathbb{R}^2$,

$$s(a, b) = \sum_{n=0}^{+\infty} \frac{a^n}{n!} e^{-\left(\frac{b-n}{\sqrt{2}\sigma}\right)^2}. \tag{12}$$

A common useful technique for solving large-size optimization problems such as those encountered in image recovery consists of splitting (i.e. decomposing) the cost function $f$ in a sum of simpler functions, which are then processed individually. For example, each of these functions can be dealt with through its gradient if the function is $\mu$-Lipschitz differentiable[1], or through its proximity operator [58, 65] if the latter has a closed form expression.

In view of this, the two following results are essential to our approach.

**Theorem 1.** *The function* $\Phi$ *is* $\mu$*-Lipschitz differentiable on* $[0, +\infty)^N$ *with*

$$\mu = \|H\|^2 \left(1 - e^{-\frac{1}{\sigma^2}}\right) \exp\left(\frac{1}{\sigma^2}\left(2 \max_{i \in \{1,\ldots,Q\}} \{y_i\} - 1\right)\right). \tag{13}$$

*Proof.* Our proof consists of showing that, for every $i \in \{1, \cdots, Q\}$, $\eta_i$ is upper bounded on $[0, +\infty)$ by $\eta_i(0)$, where $\eta_i(0)$ is defined in (30). Firstly, we recall that for $a = 0$ the infinite sum (12) simplifies to the first sum element. Consequently, for every $z_i \in (0, +\infty)$, we have the following equivalences:

$$\eta_i(0) - \eta_i(z_i) \geq 0 \tag{14}$$

$$\Leftrightarrow \left(s(z_i, y_i)\right)^2 \eta_i(0) - \left(\left(s(z_i, y_i - 1)\right)^2 - s(z_i, y_i)s(z_i, y_i - 2)\right) \geq 0$$

$$\Leftrightarrow \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} \frac{z_i^{m+n}}{m!n!} \left(e^{\frac{-(y_i-m)^2-(y_i-n)^2}{2\sigma^2}} \eta_i(0) - e^{\frac{-(y_i-1-m)^2-(y_i-1-n)^2}{2\sigma^2}} + e^{\frac{-(y_i-m)^2-(y_i-2-n)^2}{2\sigma^2}}\right) \geq 0$$

$$\Leftrightarrow \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} \frac{z_i^{m+n}}{m!n!} e^{\frac{-(y_i-n)^2-(y_i-m)^2+4y_i-2}{2\sigma^2}} \left\{\left(1 - e^{-\frac{n+m}{\sigma^2}}\right) + e^{-\frac{1}{\sigma^2}}\left(e^{-\frac{2n}{\sigma^2}} - 1\right)\right\} \geq 0$$

$$\Leftrightarrow \sum_{m=0}^{\infty}\sum_{n=0}^{m-1} \frac{z_i^{m+n}}{m!n!} e^{\frac{-(y_i-n)^2-(y_i-m)^2+4y_i-2}{2\sigma^2}} \left\{2\left(1 - e^{-\frac{n+m}{\sigma^2}}\right) + e^{-\frac{1}{\sigma^2}}\left(e^{-\frac{2n}{\sigma^2}} + e^{-\frac{2m}{\sigma^2}} - 2\right)\right\}$$

$$+ \sum_{m=0}^{\infty} \frac{z_i^{2m}}{(m!)^2} e^{\frac{-(y_i-m)^2+2y_i-1}{\sigma^2}}\left(1 - e^{-\frac{2m}{\sigma^2}}\right)\left(1 - e^{-\frac{1}{\sigma^2}}\right) \geq 0$$

$$\Leftrightarrow \sum_{m=0}^{\infty}\sum_{n=0}^{m-1} \frac{z_i^{m+n}}{m!n!} e^{\frac{-(y_i-n)^2-(y_i-m)^2+4y_i-2}{2\sigma^2}} \left\{2\left(1 - e^{-\frac{n+m}{\sigma^2}}\right)\left(1 - e^{-\frac{1}{\sigma^2}}\right) + e^{-\frac{1}{\sigma^2}}\left(e^{-\frac{n}{\sigma^2}} - e^{-\frac{m}{\sigma^2}}\right)^2\right\}$$

$$+ \sum_{m=0}^{\infty} \frac{z_i^{2m}}{(m!)^2} e^{\frac{-(y_i-m)^2+2y_i-1}{\sigma^2}}\left(1 - e^{-\frac{2m}{\sigma^2}}\right)\left(1 - e^{-\frac{1}{\sigma^2}}\right) \geq 0.$$

---

[1]A differentiable function $f : \mathbb{R}^N \to \mathbb{R}$ is said to be $\mu$-Lipschitz differentiable on a subset $C \subset \mathbb{R}^N$ if its gradient $\nabla f$ is such that, for every $(x, y) \in C^2$, $\|\nabla f(x) - \nabla f(y)\| \leq \mu\|x - y\|$.

The latter inequality holds, since the left-hand side is a sum of nonnegative terms. Hence, the expression of the Lipschitz constant $\mu$ in (13) is obtained by searching the maximum value of $\eta_i(0)$ for all possible values of $i \in \{1, \ldots, Q\}$. $\qquad\square$

The next convexity result can be regarded as an extension to the one presented in [9].

**Theorem 2.** *The neg-log-likelihood $\Phi^{(\beta)}$ of a mixture of Generalized-Gaussian and Poisson variables defined over the positive orthant as*

$$\left(\forall x \in [0, +\infty)^N\right) \quad \Phi^{(\beta)}(x) = \sum_{i=1}^{Q} \Phi_i^{(\beta)}([Hx]_i) \tag{15}$$

*where, for every $i \in \{1, \ldots, Q\}$ and $u \in [0, +\infty)$,*

$$\Phi_i^{(\beta)}(u) = -\log \left( \sum_{n=0}^{+\infty} \frac{u^n e^{-u}}{n!} \frac{\beta}{2\sqrt{2}\sigma\Gamma(\frac{1}{\beta})} e^{-\left(\frac{|y_i - n|}{\sqrt{2}\sigma}\right)^\beta} \right) \tag{16}$$

*is convex if $\beta \geq 1$. It is strictly convex if $\beta > 1$ and $H$ is injective.*

*Proof.* It is sufficient to show that, for every $i \in \{1, \cdots, Q\}$,

$$(\forall a \in (0, +\infty)) \quad \begin{cases} \ddot{\Phi}_i^{(\beta)}(a) > 0 \text{ if } \beta > 1, \\ \ddot{\Phi}_i^{(\beta)}(a) \geq 0 \text{ if } \beta = 1, \end{cases} \tag{17}$$

where

$$\ddot{\Phi}_i^{(\beta)}(a) = \left( \left(s^{(\beta)}(a, y_i - 1)\right)^2 - s^{(\beta)}(a, y_i)s^{(\beta)}(a, y_i - 2) \right) \Big/ \left(s^{(\beta)}(a, y_i)\right)^2,$$

and, for every $b \in \mathbb{R}$,

$$s^{(\beta)}(a, b) = \sum_{n=0}^{+\infty} \frac{a^n}{n!} \zeta_b^{(\beta)}(n), \tag{18}$$

with $\zeta_b^{(\beta)}(\cdot) = e^{-\left(\frac{|\cdot - b|}{\sqrt{2}\sigma}\right)^\beta}$. The proof reduces to studying the sign of the numerator of $\ddot{\Phi}_i^{(\beta)}$. For every $a \in [0, +\infty)$, $b \in \mathbb{R}$ and $\beta \in (0, +\infty)$, the following equivalences hold:

$$\left(s^{(\beta)}(a, b-1)\right)^2 - s^{(\beta)}(a, b)s^{(\beta)}(a, b-2) \geq 0 \tag{19}$$

$$\Leftrightarrow \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{a^{n+m}}{n!m!} \left( \zeta_b^{(\beta)}(m+1)\zeta_b^{(\beta)}(n+1) - \zeta_b^{(\beta)}(m)\zeta_b^{(\beta)}(n+2) \right) \geq 0$$

$$\Leftrightarrow \sum_{m=0}^{\infty} \sum_{n=1}^{\infty} \frac{a^{n+m-1}}{n!m!} n \left( \zeta_b^{(\beta)}(m+1)\zeta_b^{(\beta)}(n) - \zeta_b^{(\beta)}(m)\zeta_b^{(\beta)}(n+1) \right) \geq 0$$

$$\Leftrightarrow \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{a^{n+m-1}}{n!m!} n \left( \zeta_b^{(\beta)}(m+1)\zeta_b^{(\beta)}(n) - \zeta_b^{(\beta)}(m)\zeta_b^{(\beta)}(n+1) \right) \geq 0$$

$$\Leftrightarrow \sum_{m=0}^{\infty} \left\{ \sum_{n=0}^{m} \frac{a^{n+m-1}}{n!m!} n \left( \zeta_b^{(\beta)}(m+1)\zeta_b^{(\beta)}(n) - \zeta_b^{(\beta)}(m)\zeta_b^{(\beta)}(n+1) \right) \right.$$

$$\left. + \sum_{n=m}^{\infty} \frac{a^{n+m-1}}{n!m!} n \left( \zeta_b^{(\beta)}(m+1)\zeta_b^{(\beta)}(n) - \zeta_b^{(\beta)}(m)\zeta_b^{(\beta)}(n+1) \right) \right\} \geq 0$$

$$\Leftrightarrow \sum_{m=0}^{\infty} \left\{ \sum_{n=0}^{m} \frac{a^{n+m-1}}{n!m!} n \left( \zeta_b^{(\beta)}(m+1)\zeta_b^{(\beta)}(n) - \zeta_b^{(\beta)}(m)\zeta_b^{(\beta)}(n+1) \right) \right.$$

$$\left. + \sum_{n=0}^{m} \frac{a^{n+m-1}}{n!m!} m \left( \zeta_b^{(\beta)}(n+1)\zeta_b^{(\beta)}(m) - \zeta_b^{(\beta)}(n)\zeta_b^{(\beta)}(m+1) \right) \right\} \geq 0$$

$$\Leftrightarrow \sum_{m=0}^{\infty} \sum_{n=0}^{m} \frac{a^{n+m-1}}{n!m!} (n-m) \left( \zeta_\beta(m+1)\zeta_\beta(n) - \zeta_\beta(m)\zeta_\beta(n+1) \right) \geq 0$$

$$\Leftrightarrow \sum_{m=0}^{\infty} \sum_{n=0}^{m} \frac{a^{n+m-1}}{n!m!} (m-n) \left( \zeta_b^{(\beta)}(n+1)\zeta_b^{(\beta)}(m) - \zeta_b^{(\beta)}(n)\zeta_b^{(\beta)}(m+1) \right) \geq 0. \qquad (20)$$

Furthermore, for all $a \in (0, +\infty)$, $b \in \mathbb{R}$, and $(n, m) \in \mathbb{N}^2$ with $m > n$,

$$\frac{a^{n+m-1}}{n!m!} (m-n) \left( \zeta_b^{(\beta)}(n+1)\zeta_b^{(\beta)}(m) - \zeta_b^{(\beta)}(n)\zeta_b^{(\beta)}(m+1) \right) \geq 0$$

$$\Leftrightarrow \zeta_b^{(\beta)}(n+1)\zeta_b^{(\beta)}(m) - \zeta_b^{(\beta)}(n)\zeta_b^{(\beta)}(m+1) \geq 0$$

$$\Leftrightarrow \exp\left( -\left( \frac{|b-n-1|}{\sigma} \right)^{\beta} - \left( \frac{|b-m|}{\sigma} \right)^{\beta} \right) - \exp\left( -\left( \frac{|b-n|}{\sigma} \right)^{\beta} - \left( \frac{|b-m-1|}{\sigma} \right)^{\beta} \right) \geq 0$$

$$\Leftrightarrow |b-n|^{\beta} - |b-n-1|^{\beta} - (|b-m|^{\beta} - |b-m-1|^{\beta}) \geq 0. \qquad (21)$$

The above inequality holds if the function

$$\pi^{(\beta)} : \mathbb{R} \to \mathbb{R}$$

$$u \mapsto |u|^{\beta} - |u-1|^{\beta} \qquad (22)$$

is increasing. To prove this fact, let us study the sign of the derivative of the function $\pi^{(\beta)}$ over its domain. The singularity points, i.e. $u \in \{0, 1\}$, can be excluded from the study due to the continuity of $\pi^{(\beta)}$. Then, for $\beta \geq 1$ and $u \in \mathbb{R} \setminus \{0, 1\}$ the derivative is given by

$$\frac{\partial \pi^{(\beta)}(u)}{\partial u} = \beta \left( \text{sign}(u) |u|^{\beta-1} - \text{sign}(u-1) |u-1|^{\beta-1} \right), \qquad (23)$$

$$= \begin{cases} \beta \left( |u|^{\beta-1} + |u-1|^{\beta-1} \right) & \text{if } u \in (0, 1), \\ \beta \, \text{sign}(u)(|u|^{\beta-1} - |u-1|^{\beta-1}) & \text{if } u \in (-\infty, 0) \cup (1, \infty), \end{cases}$$

where sign denotes the signum function, and

$$\text{sign}\left( \frac{\partial \pi^{(\beta)}(u)}{\partial u} \right) = \begin{cases} \text{sign}\left( 1 - |u/(u-1)|^{\beta-1} \right) & \text{if } u \in (-\infty, 0), \\ \text{sign}\left( |u/(u-1)|^{\beta-1} + 1 \right) & \text{if } u \in (0, 1), \\ \text{sign}\left( |u/(u-1)|^{\beta-1} - 1 \right) & \text{if } u \in (1, +\infty). \end{cases} \qquad (24)$$

Hence, it is positive (resp. nonnegative) for $\beta > 1$ (resp. $\beta = 1$), so that $\pi^{(\beta)}$ is strictly increasing (resp. increasing). $\qquad \square$

The Poisson-Generalized Gaussian offers more flexibility than the Poisson-Gaussian model, but it has not been yet studied in the area of image processing. The results obtained above should be of potential use for future applications. In the rest of the paper we concentrate on Poisson-Gaussian model ($\beta = 2$), which is of main interest of this paper.

| Name | Expression | Convexity condition | Lipschitz constant |
|---|---|---|---|
| Generalized Anscombe Transform (GAST) [59] | $\widetilde{\Phi}_i(u) = \frac{1}{2}\left(\nu(y_i) - \nu(u)\right)^2,$ $\widetilde{\xi}_i(u) = \frac{2\nu(y_i)}{\nu(u)} - 1$ with $\nu(u) = 2\sqrt{u + 3/8 + \sigma^2}$ | $y_i + 3/8 + \sigma^2 \geq 0$ | $(3/8 + \sigma^2)^{-3/2}\frac{\nu(y_i)}{2}$ |
| Exponential (EXP) [9] | $\widetilde{\Phi}_i(u) = u - \frac{1/2 - y_i - \sigma^2}{e}\,\mathrm{Ei}\left(\frac{\sigma^2 - 1/2 + y_i}{\sigma^2 + u}\right)$ $-(\sigma^2 + u)e^{-\frac{1/2 - y_i + u}{\sigma^2 + u}},$ $\widetilde{\xi}_i(u) = \exp\left(-\frac{1 + 2u - 2y_i}{2u + 2\sigma^2}\right)$ | $y_i + \sigma^2 - 1/2 \geq 0$ | $\frac{2y_i + 2\sigma^2 - 1}{2\sigma^4}\exp\left(\frac{2y_i - 1}{2\sigma^2}\right)$ |
| Shifted Poisson (SPoiss) [17] | $\widetilde{\Phi}_i(u) = \eta(u) - \eta(y_i)\log\eta(u),$ $\widetilde{\xi}_i(u) = \frac{\eta(y_i)}{\eta(u)}$ with $\eta(u) = u + \sigma^2$ | $y_i + \sigma^2 \geq 0$ | $\frac{y_i + \sigma^2}{\sigma^4}$ |
| Weighted least squares (WL2) [9, 41, 55, 64] | $\widetilde{\Phi}_i(u) = \frac{1}{2}\frac{(y_i - u)^2}{\sigma^2 + u},$ $\widetilde{\xi}_i(u) = \frac{(y_i + \sigma^2)^2}{4(u + \sigma^2)^2} + \frac{3}{4}$ | $---$ | $\frac{(y_i + \sigma^2)^2}{\sigma^6}$ |

Table 1: Various approximations $\widetilde{\Phi}(x) = \sum_{i=1}^{Q}\widetilde{\Phi}_i([Hx]_i)$ (up to an additive constant) of the Poisson-Gaussian neg-log-likelihood $\Phi(x)$ proposed in the literature, together with their gradient $\nabla\widetilde{\Phi}(x) = H^\top(1 - (\widetilde{\xi}_i([Hx]_i)_{1\leq i\leq Q})$. Ei denotes the exponential integral function. We provide conditions on $(y_i)_{1\leq i\leq Q}$ for $\widetilde{\Phi}$ to be convex as well as the expression of coefficients $(\mu_i)_{1\leq i\leq Q}$ involved in the Lipschitz constant $\mu = \|H\|^2 \max_{i\in\{1,\ldots,Q\}}\mu_i$ of $\nabla\widetilde{\Phi}$.

Table 1 summarizes several data fidelity terms that have been proposed in the literature as approximations of (6), with the aim to solve image restoration problems with data corrupted by Poisson-Gaussian noise. All these approximations are convex (up to some assumptions on $y$), and Lipschitz differentiable.

For the optimization methods developed in the next section, it is important to note that the definition of the negative log-likelihood (6) can be extended to the whole space $\mathbb{R}^N$ by setting, for every $x \in \mathbb{R}^N$,

$$\Phi(x) = h(x) + \iota_{[0,+\infty)^N}(x), \tag{25}$$

where

$$h(x) = \sum_{i=1}^{Q}\varphi_i([Hx]_i), \tag{26}$$

$$\iota_{[0,+\infty)^N}(x) = \begin{cases} 0 & \text{if } x \in [0,+\infty)^N, \\ +\infty & \text{otherwise.} \end{cases} \tag{27}$$

For every $i \in \{1,\ldots,Q\}$, function $\varphi_i$ identifies with $\Phi_i$ on the positive half line, while a quadratic extension is used on the negative one, guaranteeing that the resulting function is twice differentiable at 0, and Lipschitz differentiable on $\mathbb{R}$ with minimum Lipschitz constant, i.e.

$$\varphi_i \colon \mathbb{R} \to \mathbb{R}$$
$$u \mapsto \begin{cases} \Phi_i(u) & \text{if } u \in [0,+\infty], \\ \frac{1}{2}\eta_i(0)u^2 + (1 - \xi_i(0))u + \Phi_i(0) & \text{otherwise,} \end{cases} \tag{28}$$

where functions $\Phi_i$, $\xi_i$ and $\eta_i$ are defined in (7), (10) and (11) respectively. We note that for $a = 0$ the infinite sum (12) simplifies to the first term, i.e. $s(0,b) = e^{-\frac{b^2}{2\sigma^2}}$. Hence, for every $i \in \{1,\ldots,Q\}$, the expressions of $\xi_i(0)$ and $\eta_i(0)$ in (28) read

$$\xi_i(0) = e^{\frac{1}{2\sigma^2}(2y_i - 1)}, \tag{29}$$

$$\eta_i(0) = \left(1 - e^{-\frac{1}{\sigma^2}}\right) e^{\frac{1}{\sigma^2}(2y_i - 1)}. \tag{30}$$

Consequently, $h$ is a convex function with a $\mu$-Lipschitzian gradient on $\mathbb{R}^N$. The positivity constraint in the original problem is imposed by the indicator function (27).

# 3 Proposed optimization method

## 3.1 Minimization problem

According to the analysis carried out in Section 2, the objective function has the following form

$$(\forall x \in \mathbb{R}^N) \quad f(x) = h(x) + \psi_0(x) + \sum_{r=1}^{R} \psi_r(V_r x), \tag{31}$$

where the regularization term is split into a sum of simpler functions. More precisely, it will be assumed that $\psi_0 \in \Gamma_0(\mathbb{R}^N)$ and, for every $r \in \{1, \ldots, R\}$, $\psi_r \in \Gamma_0(\mathbb{R}^{P_r})$ and $V_r \in \mathbb{R}^{P_r \times N}$.[2] Note that (31) covers a large range of penalization strategies. For instance, a sparsity prior in an analysis frame with frame operator $V_r$ is introduced by taking $\psi_r$ equal to $\lambda_r \| \cdot \|_1$ with $\lambda_r > 0$ [34]. Block sparsity measures [35] can also be easily addressed in the proposed framework. Another popular example in image restoration is the total variation penalization [66].[3] In this case, $P_r = 2N$, $V_r = \left[(\Delta^{\mathrm{h}})^\top \ (\Delta^{\mathrm{v}})^\top\right]^\top$, where $\Delta^{\mathrm{h}} \in \mathbb{R}^{N \times N}$ (resp. $\Delta^{\mathrm{v}} \in \mathbb{R}^{N \times N}$) corresponds to a horizontal (resp. vertical) gradient operator, and, for every $x \in \mathbb{R}^N$, $\psi_r(V_r x) = \lambda_r \sum_{n=1}^{N} \left(([\Delta^{\mathrm{h}} x]_n)^2 + ([\Delta^{\mathrm{v}} x]_n)^2\right)^{1/2}$ with $\lambda_r > 0$. Another interesting choice is the Hessian-based penalization [40], [54] i.e., $P_r = 3N$ and $V_r = [(\Delta^{\mathrm{hh}})^\top \ \sqrt{2}(\Delta^{\mathrm{hv}})^\top \ (\Delta^{\mathrm{vv}})^\top]^\top$ where $\Delta^{\mathrm{hh}} \in \mathbb{R}^{N \times N}$, $\Delta^{\mathrm{hv}} \in \mathbb{R}^{N \times N}$ and $\Delta^{\mathrm{vv}} \in \mathbb{R}^{N \times N}$ model the second-order partial finite difference operators between neighbooring pixels as described in [54, Sec.III-A] and, for every $x \in \mathbb{R}^N$, $\psi_r(V_r x) = \lambda_r \sum_{n=1}^{N} \left(([\Delta^{\mathrm{hh}} x]_n)^2 + 2([\Delta^{\mathrm{hv}} x]_n)^2 + ([\Delta^{\mathrm{vv}} x]_n)^2\right)^{1/2}$ with $\lambda_r > 0$. Another example is the non-local total variation (NLTV) [42]. NLTV is associated with image-driven gradient directions i.e. directions are chosen for all $n \in \{1, \ldots, N\}$ independently, based on a similarity score between pixel intensity in a local neighborhood, e.g. patch based score [15]. In this case $P_r = \sum_{n=1}^{N} D_n$, $V_r = [\omega_1^1 \Delta_1^1, \ldots, \omega_1^{D_1} \Delta_1^{D_1}, \omega_2^1 \Delta_2^1, \ldots, \omega_2^{D_2} \Delta_2^{D_2}, \ldots, \omega_N^1 \Delta_N^1, \ldots, \omega_N^{D_N} \Delta_N^{D_N}]^\top$, where, or every $n \in \{1, \ldots, N\}$ and $d \in \{1, \ldots, D_n\}$, $\Delta_n^d \in \mathbb{R}^{1 \times N}$, $D_n \in \mathbb{N}$ states for the number of gradient directions and $\omega_n^d$ is a positive weight and, for every $x \in \mathbb{R}^N$, $\psi_r(V_r x) = \lambda_r \sum_{n=1}^{N} \left(\sum_{d=1}^{D_n} (\Delta_n^d x)^2\right)^{1/2}$ with $\lambda_r > 0$. The above penalties can be considered individually $(R = 1)$ or combined in a hybrid manner $(R > 1)$ [62]. Finally, following (25), $\psi_0$ should be the indicator function $\iota_{[0,+\infty)^N}$. However, to take into account the dynamic range of the expected output image, it can be more generally chosen equal to the indicator function $\iota_C$ of a nonempty closed convex subset $C$ of $[0,+\infty)^N$.

## 3.2 Primal-dual splitting algorithm

Problem (5) where $f$ takes the form (31) can be efficiently addressed using proximal splitting algorithms [14, 23, 26, 63]. The solution is obtained iteratively by evaluating the individual proximity operators of the functions $(\psi_r)_{0 \leq r \leq R}$, provided that they have an explicit expression. We first require the notion of proximity operator.

---

[2]Following standard notation in convex analysis [7], $\Gamma_0(\mathbb{R}^N)$ designates the class of lower-semicontinuous, proper, convex functions from $\mathbb{R}^N$ to $(-\infty, +\infty]$.

[3]Total variation and the Hessian-based penalization correspond to improper priors.

**Definition 1.** [58, 65] *The value at $x \in \mathbb{R}^N$ of the proximity operator of a function $f \in \Gamma_0(\mathbb{R}^N)$,[9] denoted by $\mathrm{prox}_f : \mathbb{R}^N \mapsto \mathbb{R}^N$, is the unique solution to the following minimization problem:*

$$(\forall x \in \mathbb{R}^N) \quad \mathrm{prox}_f(x) := \operatorname*{argmin}_{p \in \mathbb{R}^N} f(p) + \frac{1}{2} \|x - p\|^2. \tag{32}$$

The proximity operator is firmly nonexpansive, i.e. for every $(x, x') \in \mathbb{R}^N \times \mathbb{R}^N$,

$$\left(\mathrm{prox}_f(x) - \mathrm{prox}_f(x')\right)^\top (x - x') \geq \left\|\mathrm{prox}_f(x) - \mathrm{prox}_f(x')\right\|^2 \tag{33}$$

and its fixed points are the minimizers of function $f$. Numerous convex optimization algorithms are based on this concept (see [24, 61] for tutorials) due to these properties. In the context of the minimization of (31), we are interested in algorithms which incorporate functions $h$ and $(\psi_r)_{0 \leq r \leq R}$ either via their proximity operators or via their gradients. The presence of a smooth term is of paramount importance as we have shown our data fidelity term $h$ to be $\mu$-Lipschitz differentiable, while its proximity operator does not have a closed form expression. Note that not all proximal methods offer the required flexibility to address the minimization of (31). More precisely, in the case when $R = 0$, one can employ either the forward-backward (FB) [27] [6, Chapter 3] or the forward-backward-forward (FBF) algorithms [69]. Conversely, celebrated algorithms such as Douglas-Rachford's [22, 56], the Alternating Direction Method of Multipliers [2, 43, 66], or the generalized Forward-Backward [63] would be difficult to implement. In the case when $R \neq 0$, one can resort to primal-dual techniques proposed in [10, 26] and [21, 28, 70][4] which can be regarded as extensions of the FBF and FB algorithms, respectively.

We are now ready to present our primal-dual splitting algorithm. The main advantage of the chosen primal-dual splitting algorithm is that it allows us to solve (5) for any Lipschitz differentiable function $h$ while allowing arbitrary linear operators $(V_r)_{1 \leq r \leq R}$. Another strong point of this algorithm is that it does not require any matrix inversion to be performed. Our primal-dual method is summarized in Algorithm 1. It corresponds to an instance of the algorithm proposed in [26] under a generic form.

---
**Algorithm 1** Proposed algorithm.

---
Let $\gamma \in (0, +\infty)$. Let $(c_k)_{k \in \mathbb{N}}$ and $(e_k)_{k \in \mathbb{N}}$ be some sequences of elements of $\mathbb{R}^N$ corresponding to possible errors in the computation of the gradient of $h$.

**Initialization:**
Set $x_0 \in \mathbb{R}^N$, and $(\forall r \in \{1, \ldots, R\})$ $v_{r,0} \in \mathbb{R}^{P_r}$.

**Iterations:**
For $k = 0, \ldots$
$\quad \left|\begin{array}{l} y_{1,k} = x_k - \gamma \left(\nabla h(x_k) + \sum_{r=1}^R V_r^\top v_{r,k}\right) + c_k \\ p_{1,k} = \mathrm{prox}_{\gamma \psi_0}(y_{1,k}) \\ \text{For } r = 1, \ldots, R \\ \quad \left|\begin{array}{l} y_{2,r,k} = v_{r,k} + \gamma V_r x_k \\ p_{2,r,k} = y_{2,r,k} - \gamma \mathrm{prox}_{\gamma^{-1} \psi_r}(\gamma^{-1} y_{2,r,k}) \\ q_{2,r,k} = p_{2,r,k} + \gamma V_r p_{1,k} \\ v_{r,k+1} = v_{r,k} - y_{2,r,k} + q_{2,r,k} \end{array}\right. \\ q_{1,k} = p_{1,k} - \gamma \left(\nabla h(p_{1,k}) + \sum_{r=1}^R V_r^\top p_{2,r,k}\right) + e_k \\ x_{k+1} = x_k - y_{1,k} + q_{1,k} \end{array}\right.$

---

---
[4]The latter algorithms are generalizations of the primal-dual technique developed in [18, 36, 45].

### 3.3 Convergence result

The convergence of the proposed primal-dual proximal splitting algorithm is guaranteed by the following result deduced from Theorem 1 and [26, Theorem 4.2]:

**Theorem 3.** *Given the following assumptions:*

(i) *$f$ is coercive, i.e. $\lim_{\|x\|\to+\infty} f(x) = +\infty$,*

(ii) *for every $r \in \{1, \ldots, R\}$, $\psi_r$ is finite valued,*

(iii) *$\gamma \in [\epsilon, (1-\epsilon)/\delta]$ where $\epsilon \in (0, 1/(\delta+1))$ and*

$$\delta = \mu + \sqrt{\sum_{r=1}^{R} \|V_r\|^2},$$

(iv) *$(c_k)_{k\in\mathbb{N}}$ and $(e_k)_{k\in\mathbb{N}}$ are absolutely summable sequences,*

*then there exists a minimizer $\widehat{x}$ of (31) such that the sequences $(x_k)_{k\in\mathbb{N}}$ and $(p_{1,k})_{k\in\mathbb{N}}$ produced by Algorithm 1 converge to $\widehat{x}$.*

### 3.4 Implementation issues

Algorithm 1 is tolerant to numerical errors $(c_k)_{k\in\mathbb{N}}$ and $(e_k)_{k\in\mathbb{N}}$. This feature is essential in our context, as the gradient of the Poisson-Gaussian negative log-likelihood given by (8) involves infinite sums. We propose finite summation bounds depending on the current estimate of $x$ that allow to compute practically accurate approximations of these sums. Let us firstly rewrite $s(a,b)$ defined in (12) as $s(a,b) = \sum_{n=0}^{+\infty} \Pi(a,b,n)$ with

$$\Pi(a,b,n) = \frac{a^n}{n!} e^{-\left(\frac{b-n}{\sqrt{2}\sigma}\right)^2}. \tag{34}$$

Then, the approximation accuracy is guaranteed by the following result deduced from [50, Proposition A.2]:

**Proposition 1.** *Let $\Delta > 0$ and set*

$$n^- = \lfloor n^* - \Delta\sigma \rfloor, \quad n^+ = \lceil n^* + \Delta\sigma \rceil \tag{35}$$

*with $n^*$ given by*

$$n^* = \sigma^2 \mathsf{W}\left(\frac{a}{\sigma^2} e^{b/\sigma^2}\right), \tag{36}$$

*where $\mathsf{W}(\cdot)$ denotes the Lambert function. Then, $\sum_{n=\max(1,n^-)}^{n^+} \Pi(a,b,n)$ constitutes a lower approximation to $\sum_{n=1}^{+\infty} \Pi(a,b,n)$ with maximum error value*

$$\sqrt{2\pi}\sigma \, \Pi(a,b,n^*) \left(1 - \mathrm{erf}\left(\frac{\Delta}{\sqrt{2}}\right)\right). \tag{37}$$

*where $\mathrm{erf}$ denotes the error function.*

One can observe in Fig. 1 that the bounds defined in (35) can be quite precise. On the contrary, the summation bounds $n^- = 0$ and $n^+ = b + 4\sigma$ proposed in [9, 53] are not always guaranteed to include all the significant coefficients or to be very effective. Note that efficient numerical methods exist to compute the Lambert function [39].

In order to secure convergence of the proposed algorithm, the errors made when computing the gradient terms need to summable so that $\Delta$ should be increased at each iteration. In the experimental study conducted in the next section, we will show that a good numerical performance of Algorithm 1 is obtained even with a fixed, sufficiently large, parameter $\Delta$.
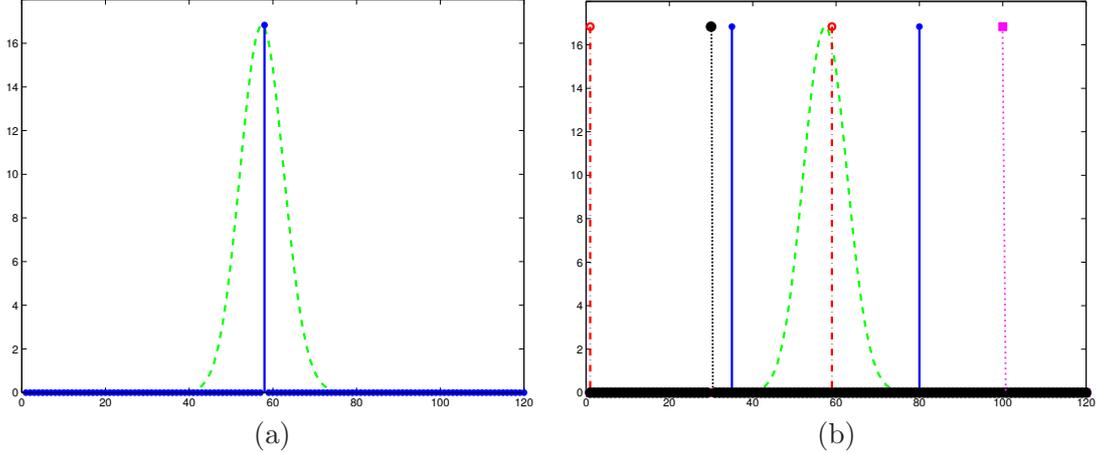
Figure 1: Approximation bounds for $a = 100$, $b = 30$ and $\sigma^2 = 50$: (a) Approximation $n^*$ from (36) (solid blue) to the maximizer of function $n \mapsto \Pi(a, b, n)$ (dashed green), (b) function $n \mapsto \Pi(a, b, n)$ (dashed green), summation bounds resulting from (35) with $\Delta = 3$ (solid blue) and from [9, 53] (dash-dotted red). The values of $a$ and $b$ are indicated with black and magenta dotted lines, respectively.

## 4 Simulations

We now demonstrate the practical performance of Algorithm 1 for the restoration of images corrupted with blur and Poisson-Gaussian noise. More specifically, our study aims at illustrating the usefulness of our approach and at comparing the merits of the various approximations of the Poisson-Gaussian likelihood, in the context of low-count confocal microscopy images with fairly low initial signal to noise ratio (SNR).

The considered ground truth images, i.e. $\overline{x}_1$ of size $190 \times 190$, $\overline{x}_2$ of size $128 \times 128$, $\overline{x}_3$ of size $350 \times 350$ and $\overline{x}_4$ of size $256 \times 256$, are illustrated in Fig. 2(top). The first two images result from the time series noise identification procedure described in [50] applied to real data acquired with a macro confocal laser scanning microscope (Leica TCS-LSI). The third and fourth images are publicly free confocal microscopy images available respectively at http://meyerinst.com/ confocals/ tcs-spe/index.htm and www.gensat.org/ imagenavigator.jsp?imageID=29109 [1]. The intensities of these original images have been rescaled so as to impose a given pixel value range $[x^-, x^+]$, with $x^- = 0$ and $x^+ \in (0, +\infty)$.[5] The images are then degraded with a convolution operator $H$ modeling a spatially invariant blur, and further corrupted with Poisson and zero-mean additive white Gaussian noise, with standard-deviation $\sigma > 0$, according to model (2). The resulting images are displayed in Fig. 2(bottom).

### 4.1 Method settings

The restored images are obtained by applying Algorithm 1 to Problem (31), where the definition of $h$, $(\psi_r)_{0 \leq r \leq R}$ and $(V_r)_{1 \leq r \leq R}$ depends on the retained data fidelity and regularization strategies. More precisely, we present restoration results obtained for a data fidelity term either derived from the Gaussian likelihood, the Poisson likelihood, or the GAST, EXP, SPoiss, WL2 or (6) functions described in Table 1. For the regularization term, we consider the sum of $\iota_C$, the indicator function of $C = [x^-, x^+]^N$ with a TV-like prior, defined either as the standard isotropic TV prior with weight $\lambda > 0$, as the NLTV prior with weight $\lambda > 0$, or as a hybrid

---

[5]In the limit case, problems with unknown $x^+$ result in the standard use of the non-negative orthant.

$\bar{x}_1$ (size $190 \times 190$, $x^+ = 12$) $\quad$ $\bar{x}_2$ (size $128 \times 128$, $x^+ = 30$) $\quad$ $\bar{x}_3$ (size $350 \times 350$, $x^+ = 20$) $\quad$ $\bar{x}_4$ (size $256 \times 256$, $x^+ = 100$)

$y_1$ (kernel size $= 25 \times 25$, $\quad$ $y_2$ (kernel size $= 25 \times 25$, $\quad$ $y_3$ (kernel size $= 5 \times 5$, $\quad$ $y_4$ (kernel size $= 9 \times 9$,

kernel std $= 1.6$, $\sigma^2 = 9$, $\quad$ kernel std $= 1.6$, $\sigma^2 = 12$, $\quad$ kernel size $= 5$, $\sigma^2 = 9$, $\quad$ kernel std $= 0.5$, $\sigma^2 = 36$,

MAE $= 61$, SNR $= 2.2$ dB) $\quad$ MAE $= 35$, SNR $= 10.1$ dB) $\quad$ MAE $= 49$, SNR $= 7.6$ dB) $\quad$ MAE $= 26$, SNR $= 9.4$ dB)
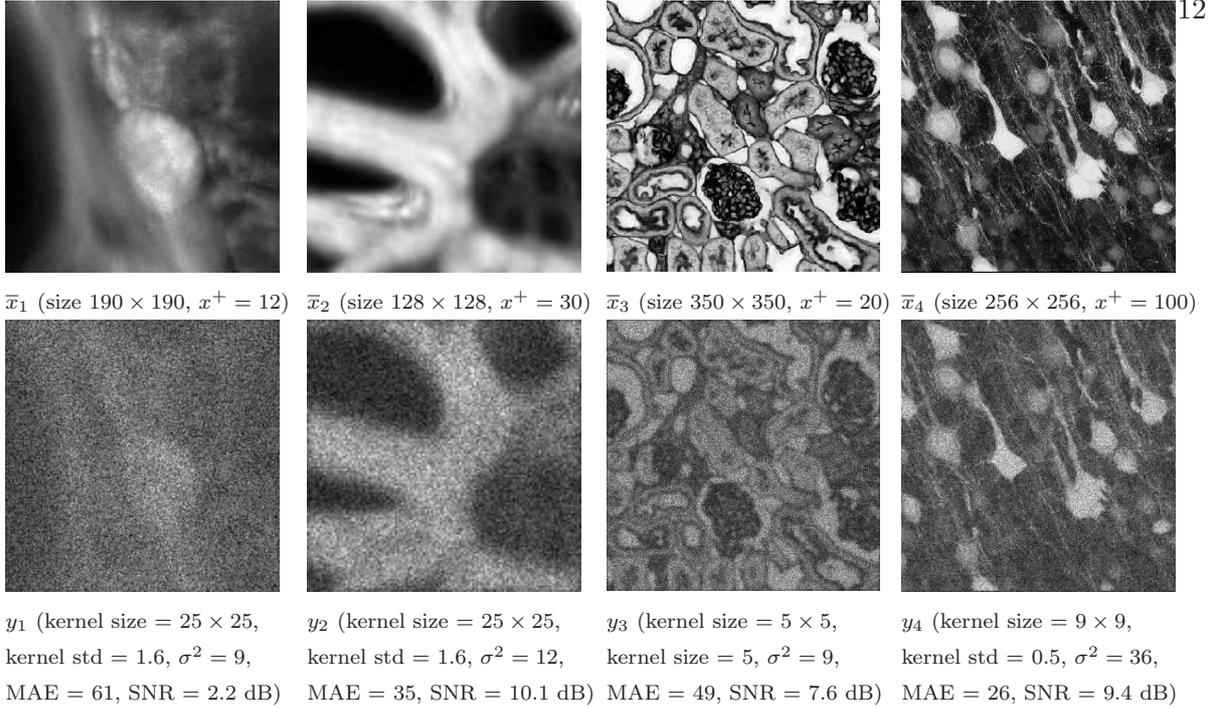
Figure 2: Original (top) and degraded (bottom) images with either a truncated Gaussian blur ($y_1, y_2$ and $y_4$) or a uniform blur ($y_3$), and Poisson-Gaussian noise, with Gaussian noise variance $\sigma^2$.

penalization being the sum of the Hessian prior from [54] with TV penalty, with respective weights $\lambda_{\mathrm{H}} > 0$ and $\lambda_{\mathrm{TV}} > 0$. In all our experiments, the regularization weights are tuned manually so as to minimize the Mean Absolute Error (MAE) between the original image $\bar{x}$ and its final estimate $\hat{x}$. Moreover, the linear operators involved in the NLTV prior are computed from an initial Wiener filter restoration using code provided by Bresson [13].[6].

The Gaussian likelihood, and the GAST, EXP, WL2, (6) functions are Lipschitz differentiable, so that the resulting optimization problem takes the form (31) where $h$ is the considered data fidelity term, $\psi_0 = \iota_C$, the indicator function of $C = [x^-, x^+]^N$, and $(\psi_r)_{1 \le r \le R}$, $(V_r)_{1 \le r \le R}$ model the TV-like prior. GAST is thus handled in a manner similar to [32], i.e. by taking advantage of Lipschitz-differentiability properties of the function $\widetilde{\Phi}$ presented in the first line of Table 1. In the cases of the non-Lipschitz differentiable SPoiss and Poisson likelihoods, the optimization problem is written under the form (31) where $h \equiv 0$, $\psi_0$ is the considered data fidelity term, $\psi_1 = \iota_C$, and $(\psi_r)_{2 \le r \le R}$, $(V_r)_{2 \le r \le R}$ model the TV-like prior. Note that, in both cases, the proximity operators of the involved functions $(\psi_r)_{0 \le r \le R}$ have explicit forms [25].

In order to guarantee the convergence of Algorithm 1, the parameter $\gamma$ involved in the algorithm is chosen so as to take the maximum possible value satisfying Theorem 3 (iii). Moreover, in the cases of the Poisson, SPoiss, WL2 and GAST data fidelity terms, a data truncation is performed as a pre-processing step on the observed images in order to satisfy the required convexity conditions (see Table 1).

---

[6]http://www.cs.cityu.edu.hk/~xbresson/codes.html

## 4.2 Restoration results

We start this section with a study, in Table 2, of the influence of the choice of parameter $\Delta$ in the restoration quality of image $\overline{x}_1$, when the exact Gaussian-Poisson likelihood (6) is used as data fidelity term. One can observe that the restoration performance is stable for $\Delta \geq 3$. This setting will be adopted for the remaining experiments.

Tables 3, 4, 5, and 6 present the evaluation scores MAE (Mean Absolute Error), SNR and SSIM [71] for the initial and restored images using the various data fidelity criteria. Note that the provided relative MAE values are normalized with a factor $255/x^+$.

| $\Delta$ | 0.33 | 0.5 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|---|
| MAE | 10.99 | 9.30 | 7.69 | 7.53 | 7.53 | 7.53 | 7.53 |
| SNR | 17.43 | 18.51 | 19.73 | 19.82 | 19.81 | 19.81 | 19.81 |
| SSIM | 0.989 | 0.993 | 0.996 | 0.996 | 0.996 | 0.996 | 0.996 |

Table 2: Influence of $\Delta$ on restoration results of image $\bar{x}_1$ using Hessian-TV prior.

One can observe that the exact model always leads to the best qualitative results. Conversely, Poisson, WL2 and GAST strategies lead to relatively poor performance. Although Gaussian and EXP data fidelity terms may be competitive in some situations (see, for instance, Table 3), the performance gain in terms of SNR between the exact and EXP data fidelity term reaches up to 0.25 dB (see Table 4). Complementary to these numerical results, Fig. 3 illustrates the visual differences resulting from various data fidelity terms in the case of image $\overline{x}_1$. One can observe that a high number of low intensity components are lost when using a data fidelity term derived from Poisson statistics. Similarly, the shape of low intensity components is not well reconstructed when using a WL2 data fidelity term. Most artifacts are corrected when using the GAST data fidelity term while even better results are obtained with the Gaussian and SPoiss. Finally, one can hardly notice a visual difference between the EXP approximation and the exact data fidelity term. Finally, we indicate in the caption of Fig. 3 the computational time required to reach the algorithm convergence (typically, when the difference in norm between two consecutive iterates becomes small), when tests are performed on a quad-CPU, 32-core Intel Xeon L7555 1.8 GHz with 144 GB of RAM running RedHat Enterprise Linux 6.5 using a Matlab code implementation. One can observe that the best tradeoff between image quality and convergence time is reached by the EXP model. The computational time difference between Exact and EXP data fidelity term may result from: i) the relatively high value of Lipschitz constant and ii) the gradient of the data fidelity which is computationally more expensive. The latter issue could be alleviated by a better optimized implementation. Finally, one can observe by inspecting the final MAE relative values and the SNR improvements for the four images that the hybrid Hessian-TV regularization strategy leads to very good results (see Tables 3, 4, 5, and 6). This can be also validated by a visual inspection of Fig. 4 (bottom) which shows that this penalization allows a reduction of the undesired staircase effect visible in the TV result (Fig. 4 (top)). The NLTV regularization strategy should also have reduced the undesired staircase effect visible in the TV result. However, one can see that this penalization does not always lead to an improvement with respect to TV or Hessian-TV (Fig. 4 (middle)). This reflects the fact that the similarity between local image features is essentially lost in the observed images characterized with a very low SNR, such as $y_1$, so that the NLTV weights cannot be adjusted properly in that cases.

# 5 Conclusion

We have shown that Poisson-Gaussian neg-log-likelihood is a convex, Lipschitz-differentiable function. The provided convexity result is actually more general as it applies to the neg-log likelihood of a mixture of Generalized-Gaussian and Poisson variables. Building upon these results, we have proposed a new variational approach for solving data recovery problems in the presence of Poisson-Gaussian noise. We have developed a practical implementation of an efficient primal-dual algorithm, which is particularly flexible, i.e. for which a large range of penalization strategies and data fidelity terms are applicable. Among those we employed, the hybrid TV-Hessian prior was shown to produce naturally looking, high quality results for low count microscopy image restoration problems in the presence of Poisson-Gaussian noise. We have shown the good performance of our restoration algorithm using the exact data fidelity term and various approximations. In addition, we also conclude that the EXP and SPoiss fidelity terms provide good results compared to the exact solution. In a future work we would like to extend our approach to encompass the case of Generalized Gaussian-Poisson noise and to study the performance of other regularization strategies than those we have considered here. Although only TV, NLTV and hybrid TV-Hessian priors were analyzed in this work, the versatility of our approach should allow us to address a variety of applications by making use of various forms of convex penalty functions. The extension of the proposed method to $\beta \neq 2$ would require to prove that the corresponding neg log likelihood is Lipschitz differentiable, which remains an open problem.

# References

[1] *The Gene Expression Nervous System Atlas (GENSAT) Project, NINDS contracts N01NS02331 & HHSN271200723701C to The Rockefeller University (New York, NY).*

[2] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo, *An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems*, IEEE Trans. Image Process., 20 (2011), pp. 681–695.

[3] Y. Altmann, A. Wallace, and S. McLaughlin, *Spectral unmixing of multispectral lidar signals*, tech. report, Jan. 2015. http://arxiv.org/abs/1501.01150.

[4] S. Anthoine, J.-F. Aujol, Y. Boursier, and C. Melot, *Some proximal methods for CBCT and PET tomography*, Inverse Probl. Imaging, 6 (2012), pp. 565–598.

[5] G. Aubert and J.-F. Aujol, *A variational approach to removing multiplicative noise*, SIAM J. Appl. Math., 68 (2008), pp. 925–946.

[6] F. R. Bach, R. Jenatton, J. Mairal, and G. Obozinski, *Optimization with sparsity-inducing penalties.*, Foundations and Trends in Machine Learning, 4 (2012), pp. 1–106.

[7] H. H. Bauschke and P. L. Combettes, *Convex analysis and monotone operator theory in Hilbert spaces*, Springer, New York, 2011.

[8] B. Begovic, V. Stankovic, and L. Stankovic, *Contrast enhancement and denoising of Poisson and Gaussian mixture noise for solar images*, in Proc. Int. Conf. Image Process. (ICIP 2011), Brussels, Belgium, 11-14 Sep. 2011, pp. 185–188.

[9] F. Benvenuto, A. La Camera, C. Theys, A. Ferrari, H. Lantéri, and M. Bertero, *The study of an iterative method for the reconstruction of images corrupted by Poisson and Gaussian noise*, Inverse Probl., 24 (2008). 20 pp.

[10] R. I. BOȚ AND C. HENDRICH, *Convergence analysis for a primal-dual monotone + skew splitting algorithm with applications to total variation minimization*, J. Math. Imaging Vision, 49 (2014), pp. 551–568.

[11] S. BONETTINI AND M. PRATO, *New convergence results for the scaled gradient projection method*, tech. report, Feb. 2015. http://arxiv.org/abs/1406.6601.

[12] G. BORACCHI AND A. FOI, *Multiframe raw-data denoising based on block-matching and 3-D filtering for low-light imaging and stabilization*, in The 2008 International Workshop on Local and Non-Local Approximation in Image Processing, Lausanne, Switzerland, Aug. 2008.

[13] X. BRESSON, *A short note for nonlocal TV minimization*, tech. report, Jun. 2009.

[14] L. M. BRICEÑOS ARIAS AND P. L. COMBETTES, *A monotone + skew splitting model for composite monotone inclusions in duality*, SIAM J. Optim., 21 (2011), pp. 1230–1250.

[15] A. BUADES, B. COLL, AND J.-M. MOREL, *A non-local algorithm for image denoising.*, in Proc. IEEE Comput. Soc. Conf. Comput. Vision and Pattern Recogn. (CVPR 2005), 2005, pp. 60–65.

[16] J.-F. CAI, R. H. CHAN, AND M. NIKOLOVA, *Fast two-phase image deblurring under impulse noise*, J. Math. Imaging Vision, 36 (2010), pp. 46–53.

[17] A. CHAKRABARTI AND T. ZICKLER, *Image restoration with signal-dependent camera noise*, CoRR, abs/1204.2994 (2012).

[18] A. CHAMBOLLE AND T. POCK, *A first-order primal-dual algorithm for convex problems with applications to imaging.*, J. Math. Imaging Vision, 40 (2011), pp. 120–145.

[19] R. CHAN, H. YANG, AND T. ZENG, *A two-stage image segmentation method for blurry images with Poisson or multiplicative gamma noise*, SIAM J. Imaging Sci., 7 (2014), pp. 98–127.

[20] C. CHAUX, J.-C. PESQUET, AND N. PUSTELNIK, *Nested iterative algorithms for convex constrained image recovery problem*, SIAM J. Imaging Sci., 2 (2009), pp. 730–762.

[21] P. CHEN, J. HUANG, AND X. ZHANG, *A primal-dual fixed point algorithm for convex separable minimization with applications to image restoration*, Inverse Problems, 29 (2013), p. 025011.

[22] P. L. COMBETTES AND J.-C. PESQUET, *A Douglas-Rachford splitting approach to nonsmooth convex variational signal recovery*, IEEE J. Selected Topics Signal Process., 1 (2007), pp. 564–574.

[23] ——, *A proximal decomposition method for solving convex variational inverse problems*, Inverse Problems, 24 (2008), p. 065014.

[24] ——, *Proximal splitting methods in signal processing*, in Fixed-Point Algorithms for Inverse Problems in Science and Engineering, H. H. Bauschke, R. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, eds., Springer-Verlag, New York, 2010.

[25] ——, *Proximal splitting methods in signal processing*, in Fixed-Point Algorithms for Inverse Problems in Science and Engineering, H. H. Bauschke, R. S. Burachik, P.L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, eds., Springer Optimization and Its Applications, Springer New York, 2011, pp. 185–212.

[26] ——, *Primal-dual splitting algorithm for solving inclusions with mixtures of composite, Lipschitzian, and parallel-sum type monotone operators*, Set-Valued and Variational Analysis, 20 (2012), pp. 307–330. 10.1007/s11228-011-0191-y.

[27] P. L. COMBETTES AND V. R. WAJS, *Signal recovery by proximal forward-backward splitting*, Multiscale Model. Simul., 4 (2005), pp. 1168–1200.

[28] L. CONDAT, *A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms*, J. Optim. Theory Appl., 158 (2013), pp. 460–479.

[29] S. DELPRETTI, F. LUISIER, S. RAMANI, T. BLU, AND M. UNSER, *Multiframe sure-let denoising of timelapse fluorescence microscopy images*, in Proc. IEEE Int. Symp. Biomedical Imaging (ISBI 2008), Paris, France, 14-17 May 2008, pp. 149–152.

[30] G. DEMOMENT, *Image reconstruction and restoration: Overview of common estimation structure and problems*, IEEE Trans. Acous., Speech Signal Process., 37 (1989), pp. 2024–2036.

[31] Y. DONG AND T. ZENG, *A convex variational model for restoring blurred images with multiplicative noise*, SIAM J. Imaging Sci., 6 (2013), pp. 1598–1625.

[32] F.-X. DUPÉ, J. M. FADILI, AND J.-L. STARCK, *A proximal iteration for deconvolving Poisson noisy images using sparse representations*, IEEE Trans. Image Process., 18 (2009), pp. 310–321.

[33] ——, *Deconvolution under Poisson noise using exact data fidelity and synthesis or analysis sparsity priors*, Stat. Methodol., 9 (2012), pp. 4–18.

[34] M. ELAD, P. MILANFAR, AND R. RUBINSTEIN, *Analysis versus synthesis in signal priors*, Inverse Problems, 23 (2007), pp. 947–968.

[35] Y. C. ELDAR, P. KUPPINGER, AND H. BOLCSKEI, *Block-sparse signals: uncertainty relations and efficient recovery*, IEEE Trans. Signal Process., 58 (2010), pp. 3042 –3054.

[36] E. ESSER, X. ZHANG, AND T. CHAN, *A general framework for a class of first order primal-dual algorithms for convex optimization in imaging science*, SIAM J. Imaging Sci., 3 (2010), pp. 1015–1046.

[37] A. FOI, *Clipped noisy images: Heteroskedastic modeling and practical denoising*, Signal Process., 89 (2009), pp. 2609–2629.

[38] A. FOI, M. TRIMECHE, V. KATKOVNIK, AND K. EGIAZARIAN, *Practical Poissonian-Gaussian noise modeling and fitting for single-image raw-data*, IEEE Trans. Image Process., 17 (2008), pp. 1737–1754.

[39] T. FUKUSHIMA, *Precise and fast computation of Lambert functions without transcendental function evaluations*, J. Comput. Appl. Math., 244 (2013), pp. 77 – 89.

[40] D. GEMAN AND G. REYNOLDS, *Constrained restoration and the recovery of discontinuities.*, IEEE Trans. Pattern Anal. Mach. Int., 14 (1992), pp. 367–383.

[41] E. GIL-RODRIGO, J. PORTILLA, D. MIRAUT, AND R. SUAREZ-MESA, *Efficient joint Poisson-Gauss restoration using multi-frame $\ell_2$-relaxed-$\ell_0$ analysis-based sparsity*, in Proc. Int. Conf. Image Process. (ICIP 2011), Brussels, Belgium, 11-14 Sep. 2011, pp. 1385–1388.

[42] G. GILBOA AND S. OSHER, *Nonlocal operators with applications to image processing.*, Multiscale Modeling Simul., 7 (2008), pp. 1005–1028.

[43] J.-F. GIOVANNELLI AND A. COULAIS, *Positive deconvolution for superimposed extended sources and point sources*, Astron. Astrophys., 439 (2005), pp. 401–412.

[44] S. HARIZANOV, J.-C. PESQUET, AND G. STEIDL, *Epigraphical Projection for Solving Least Squares Anscombe Transformed Constrained Optimization Problems*, vol. 7893 of Lecture Notes in Computer Science, 2013, pp. 125–136.

[45] 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.

[46] Y. HUANG, M. NG, AND Y. WEN, *A new total variation method for multiplicative noise removal*, SIAM J. Imaging Sci., 2 (2009), pp. 20–40.

[47] B. G. JEONG, B. C. KIM, Y. H. MOON, AND I. K. EOM, *Simplified noise model parameter estimation for signal-dependent noise*, Signal Process., 96, Part B (2014), pp. 266 – 273.

[48] T. JEONG, H. WOO, AND S. YUN, *Frame-based Poisson image restoration using a proximal linearized alternating direction method*, Inverse Problems, 29 (2013), p. 075007.

[49] A. JEZIERSKA, C. CHAUX, J-C. PESQUET, AND H. TALBOT, *Iterative poisson-gaussian noise parametric estimation for blind image denoising*, in Proc. Int. Conf. Image Process. (ICIP 2014), 2014, pp. 2819–2823.

[50] A. JEZIERSKA, C. CHAUX, J.-C. PESQUET, H. TALBOT, AND G. ENGLER, *An EM approach for time-variant Poisson-Gaussian model parameter estimation*, IEEE Trans. Image Process., 62 (2014), pp. 17–30.

[51] A. JEZIERSKA, E. CHOUZENOUX, J.-C. PESQUET, AND H. TALBOT, *A primal-dual proximal splitting approach for restoring data corrupted with Poisson-Gaussian noise*, in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP 2012), Kyoto, Japan, 25-30 Mar. 2012, pp. 1085–1088.

[52] N. KOMODAKIS AND J.-C. PESQUET, *Playing with duality: An overview of recent primal-dual approaches for solving large-scale optimization problems*, to appear in IEEE Signal Process. Mag., (2014). http://arxiv.org/abs/1406.5429.

[53] H. LANTÉRI AND C. THEYS, *Restoration of astrophysical images - the case of Poisson data with additive Gaussian noise.*, EURASIP J. Appl. Signal Process., 2005 (2005), pp. 2500–2513.

[54] S. LEFKIMMIATIS, A. BOURQUARD, AND M. UNSER, *Hessian-based norm regularization for image restoration with biomedical applications*, IEEE Trans. Image Process., 21 (2012), pp. 983–995.

[55] J. LI, Z. SHEN, R. JIN, , AND X. ZHANG, *A reweighted $\ell_2$ method for image restoration with Poisson and mixed Poisson-Gaussian noise*, tech. report, 2012. ftp://ftp.math.ucla.edu/pub/camreport/cam12-84.pdf.

[56] P. L. LIONS AND B. MERCIER, *Splitting algorithms for the sum of two nonlinear operators*, SIAM J. Numer. Anal., 16 (1979), pp. 964–979.

[57] F. Luisier, T. Blu, and M. Unser, *Image denoising in mixed Poisson-Gaussian noise*, IEEE Trans. Image Process., 20 (2011), pp. 696–708.

[58] J. J. Moreau, *Proximité et dualité dans un espace hilbertien*, Bull. Soc. Math. France, 93 (1965), pp. 273–299.

[59] F. Murtagh, J.-L. Starck, and A. Bijaoui, *Image restoration with noise suppression using a multiresolution support*, Astron. Astrophys. Suppl., 112 (1995), pp. 179–189.

[60] T. E. Nichols, J. Qi, E. Asma, and R. M. Leahy, *Spatiotemporal reconstruction of list-mode PET data*, IEEE Trans. Med. Imag., 21 (2002), pp. 396–404.

[61] N. Parikh and S. Boyd, *Proximal algorithms*, Foundations and Trends in Optimization, 1 (2013), pp. 123–231.

[62] 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.

[63] H. Raguet, J. Fadili, and G. Peyré, *Generalized forward-backward splitting*, SIAM J. Imaging Sci., 6 (2013), pp. 1199–1226.

[64] A. Repetti, E. Chouzenoux, and J.-C. Pesquet, *A penalized weighted least squares approach for restoring data corrupted with signal-dependent noise*, in Proc. 20th Eur. Sig. Proc. Conf. (EUSIPCO 2012), Bucharest, Roumania, 27-31 Aug. 2012, pp. 1553–1557.

[65] R. T. Rockafellar, *Monotone operators and the proximal point algorithm*, SIAM J. Control Optim., 14 (1976), pp. 877–898.

[66] L. I. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, J. Phys. D, 60 (1992), pp. 259–268.

[67] S. Setzer, G. Steidl, and T. Teuber, *Deblurring Poissonian images by split Bregman techniques*, J. Vis. Comm. Image Repr., 21 (2010), pp. 193–199.

[68] D. L. Snyder, A. M. Hammoud, and R. L. White, *Image recovery from data acquired with a charge-coupled-device camera*, J. Opt. Soc. Am. A, 10 (1993), pp. 1014–1023.

[69] P. Tseng, *A modified forward-backward splitting method for maximal monotone mappings*, SIAM J. Control Optim., 38 (2000), pp. 431–446.

[70] B. C. Vũ, *A splitting algorithm for dual monotone inclusions involving cocoercive operators*, Adv. Comput. Math., 38 (2013), pp. 667–681.

[71] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, *Image quality assessment: from error visibility to structural similarity*, IEEE Trans. Image Process., 13 (2004), pp. 600 –612.

[72] Y. Xiao, T. Zeng, J. Yu, and M. K. Ng, *Restoration of images corrupted by mixed Gaussian-impulse noise via $\ell_1$-$\ell_0$ minimization*, Pattern Recogn., 44 (2011), pp. 1708 – 1720.

[73] M. Yan, *Restoration of images corrupted by impulse noise and mixed Gaussian impulse noise using blind inpainting*, SIAM J. Imaging Sci., 6 (2013), pp. 1227–1245.

| | | Poiss. | Gauss. | GAST | EXP | SPoiss | WL2 | Exact |
|---|---|---|---|---|---|---|---|---|
| TV | $\lambda$ | 0.338 | 0.218 | 0.155 | 0.156 | 0.153 | 0.133 | 0.154 |
| | MAE | 11.70 | 8.79 | 9.76 | 8.68 | 8.78 | 12.90 | **8.66** |
| | SNR | 16.73 | 18.74 | 17.73 | 18.80 | 18.72 | 16.18 | **18.81** |
| | SSIM | 0.758 | 0.788 | 0.783 | **0.790** | 0.787 | 0.770 | **0.790** |
| NLTV | $\lambda$ | 0.243 | 0.164 | 0.115 | 0.114 | 0.117 | 0.101 | 0.138 |
| | MAE | 12.03 | **9.25** | 10.15 | 9.26 | 9.31 | 13.20 | **9.25** |
| | SNR | 16.57 | **18.38** | 17.50 | 18.36 | 18.33 | 15.92 | **18.38** |
| | SSIM | 0.733 | **0.759** | 0.755 | 0.757 | 0.757 | 0.739 | 0.758 |
| TVH | $\lambda_{\mathrm{TV}}$ | 0.016 | 0.004 | 0.016 | 0.012 | 0.013 | 0.051 | 0.008 |
| | $\lambda_{\mathrm{H}}$ | 0.609 | 0.472 | 0.293 | 0.297 | 0.292 | 0.119 | 0.296 |
| | MAE | 10.89 | 7.62 | 8.70 | **7.53** | 7.63 | 12.46 | **7.53** |
| | SNR | 17.30 | 19.72 | 18.60 | **19.81** | 19.72 | 16.58 | **19.81** |
| | SSIM | 0.815 | 0.845 | 0.839 | **0.846** | 0.844 | 0.812 | **0.846** |

Table 3: Restoration results for image $\overline{x}_1$

| | | Poiss. | Gauss. | GAST | EXP | SPoiss | WL2 | Exact |
|---|---|---|---|---|---|---|---|---|
| TV | $\lambda$ | 0.164 | 0.197 | 0.093 | 0.096 | 0.095 | 0.099 | 0.093 |
| | MAE | 10.71 | 9.42 | 9.73 | 9.49 | 9.50 | 10.05 | **9.14** |
| | SNR | 20.21 | 20.79 | 20.55 | 20.79 | 20.79 | 20.57 | **21.15** |
| | SSIM | 0.814 | 0.859 | 0.858 | 0.858 | 0.858 | 0.846 | **0.864** |
| NLTV | $\lambda$ | 0.105 | 0.121 | 0.060 | 0.055 | 0.059 | 0.054 | 0.052 |
| | MAE | 9.60 | **8.47** | 8.75 | 8.53 | 8.56 | 9.22 | **8.47** |
| | SNR | 21.14 | 21.60 | 21.39 | 21.61 | 21.61 | 21.25 | **21.72** |
| | SSIM | 0.839 | **0.878** | 0.875 | 0.875 | 0.874 | 0.864 | 0.875 |
| TVH | $\lambda_{\mathrm{TV}}$ | 0.043 | 0.075 | 0.027 | 0.031 | 0.030 | 0.037 | 0.031 |
| | $\lambda_{\mathrm{H}}$ | 0.148 | 0.136 | 0.070 | 0.072 | 0.070 | 0.059 | 0.070 |
| | MAE | 8.99 | 7.93 | 8.18 | 7.92 | 7.93 | 8.70 | **7.90** |
| | SNR | 21.09 | 21.54 | 21.30 | 21.59 | 21.60 | 21.25 | **21.61** |
| | SSIM | 0.861 | 0.899 | 0.899 | 0.899 | 0.899 | 0.887 | **0.900** |

Table 4: Restoration results for image $\overline{x}_2$

| | | Poiss. | Gauss. | GAST | EXP | SPoiss | WL2 | Exact |
|---|---|---|---|---|---|---|---|---|
| **TV** | $\lambda$ | 0.127 | 0.124 | 0.062 | 0.062 | 0.062 | 0.058 | 0.061 |
| | MAE | 22.73 | 22.87 | 22.73 | 22.42 | 22.44 | 22.91 | **22.39** |
| | SNR | 13.68 | 13.69 | 13.66 | 13.76 | 13.75 | 13.63 | **13.77** |
| | SSIM | 0.929 | 0.930 | 0.930 | **0.931** | **0.931** | 0.928 | **0.931** |
| **NLTV** | $\lambda$ | 0.067 | 0.066 | 0.032 | 0.032 | 0.032 | 0.029 | 0.032 |
| | MAE | 21.77 | 21.85 | 21.69 | 21.54 | 21.53 | 22.19 | **21.51** |
| | SNR | 14.06 | 14.05 | 14.03 | 14.10 | 14.11 | 13.96 | **14.12** |
| | SSIM | 0.934 | 0.936 | 0.936 | 0.936 | 0.936 | 0.934 | **0.937** |
| **TVH** | $\lambda_{\text{TV}}$ | 0.041 | 0.052 | 0.028 | 0.029 | 0.028 | 0.027 | 0.029 |
| | $\lambda_{\text{H}}$ | 0.050 | 0.039 | 0.018 | 0.018 | 0.017 | 0.017 | 0.017 |
| | MAE | 21.94 | 21.99 | 21.95 | 21.68 | 21.68 | 22.11 | **21.65** |
| | SNR | 14.09 | 14.09 | 14.06 | 14.15 | 14.15 | 14.02 | **14.16** |
| | SSIM | 0.935 | **0.937** | 0.936 | **0.937** | **0.937** | 0.935 | **0.937** |

Table 5: Restoration results for image $\overline{x}_3$

| | | Poiss. | Gauss. | GAST | EXP | SPoiss | WL2 | Exact |
|---|---|---|---|---|---|---|---|---|
| **TV** | $\lambda$ | 0.299 | 0.116 | 0.078 | 0.078 | 0.078 | 0.075 | 0.077 |
| | MAE | 13.55 | 11.89 | 11.79 | **11.69** | 11.73 | 11.80 | **11.69** |
| | SNR | 14.41 | 15.42 | 15.39 | 15.48 | 15.46 | **15.49** | **15.49** |
| | SSIM | 0.724 | 0.759 | 0.762 | 0.766 | 0.764 | **0.767** | 0.766 |
| **NLTV** | $\lambda$ | 0.301 | 0.118 | 0.078 | 0.077 | 0.077 | 0.074 | 0.076 |
| | MAE | 12.28 | 12.21 | 11.79 | 11.77 | 11.78 | 11.91 | **11.75** |
| | SNR | 15.10 | 15.16 | 15.40 | 15.44 | **15.45** | 15.43 | **15.45** |
| | SSIM | 0.742 | 0.742 | 0.759 | 0.759 | 0.758 | **0.760** | 0.759 |
| **TVH** | $\lambda_{\text{TV}}$ | 0.050 | 0.037 | 0.052 | 0.030 | 0.046 | 0.046 | 0.027 |
| | $\lambda_{\text{H}}$ | 0.127 | 0.054 | 0.052 | 0.034 | 0.027 | 0.020 | 0.030 |
| | MAE | 11.97 | 11.33 | 11.44 | **11.29** | 11.30 | 11.45 | **11.29** |
| | SNR | 15.24 | 15.76 | 15.64 | 15.77 | 15.79 | 15.73 | **15.82** |
| | SSIM | 0.760 | 0.775 | 0.768 | **0.782** | 0.778 | 0.779 | 0.781 |

Table 6: Restoration results for image $\overline{x}_4$

$\widehat{x}_1$ (Poisson fidelity, MAE = 10.89, SNR = 17.30 dB NbIt = 4500, t = 284 s )

$\widehat{x}_1$ (Gaussian fidelity, MAE = 7.62, SNR = 19.72 dB NbIt = 2000, t = 122 s )

$\widehat{x}_1$ (GAST fidelity, MAE = 8.70, SNR = 18.60 dB NbIt = 2000, t = 132 s)

$\widehat{x}_1$ (EXP fidelity, MAE = 7.53, SNR = 19.81 dB NbIt = 3000, t = 286 s)

$\widehat{x}_1$ (SPoiss fidelity, MAE = 7.63, SNR = 19.72 dB NbIt = 9000, t = 564 s)

$\widehat{x}_1$ (WL2 fidelity, MAE = 12.46, SNR = 16.58 dB NbIt = 2000, t = 132 s)

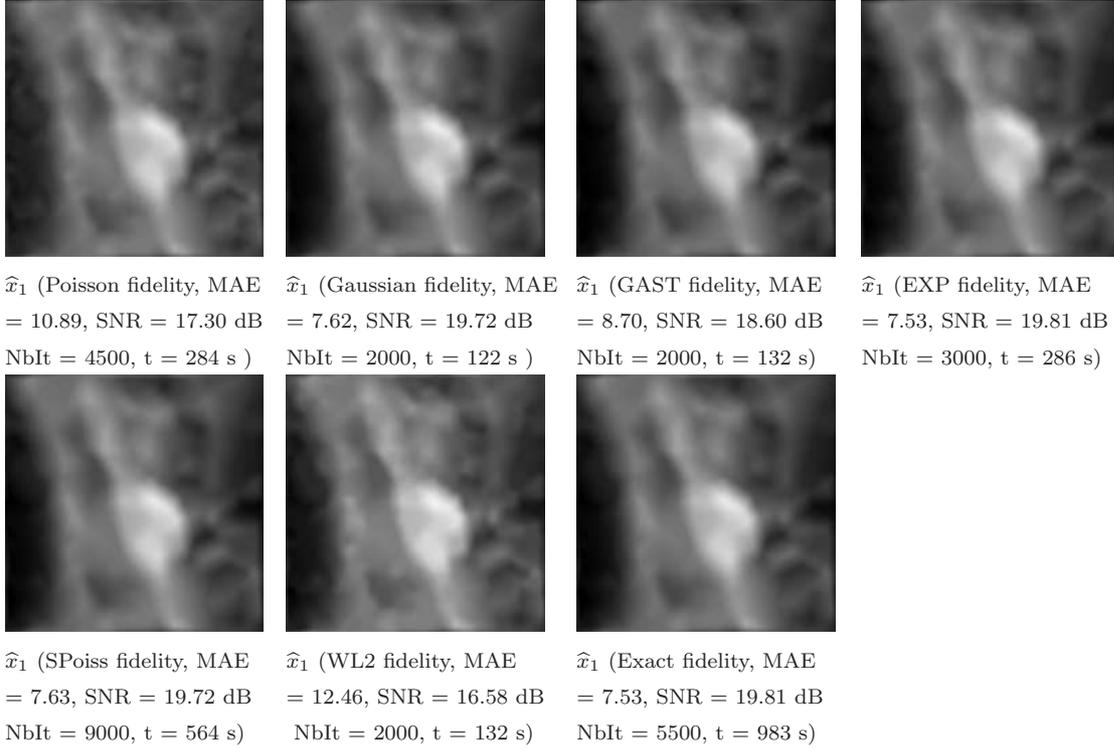$\widehat{x}_1$ (Exact fidelity, MAE = 7.53, SNR = 19.81 dB NbIt = 5500, t = 983 s)

Figure 3: Restoration results of image $\overline{x}_1$ for TV+Hessian regularization strategy and several choices for the data fidelity term.
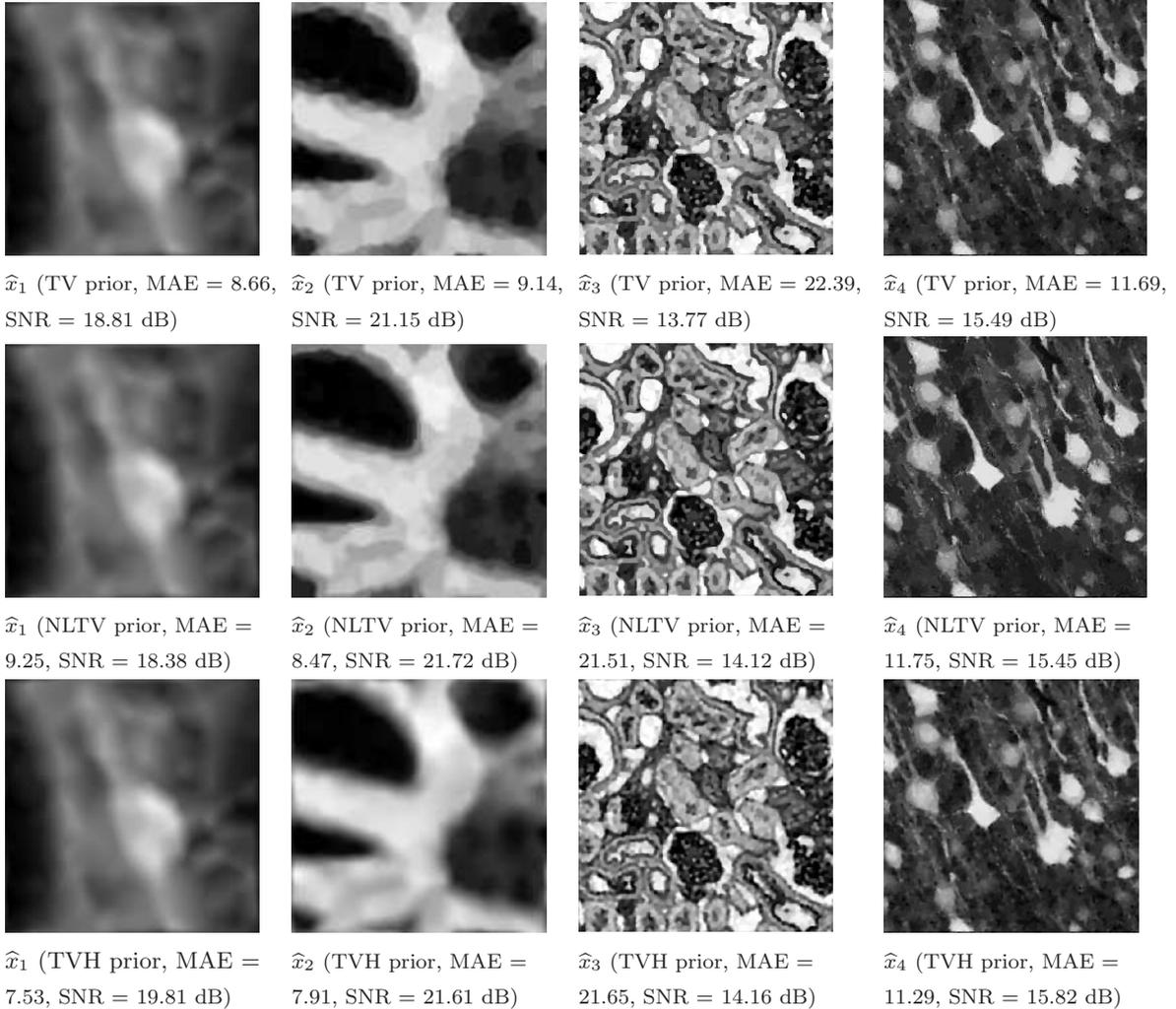


$\widehat{x}_1$ (TV prior, MAE = 8.66, SNR = 18.81 dB)

$\widehat{x}_2$ (TV prior, MAE = 9.14, SNR = 21.15 dB)

$\widehat{x}_3$ (TV prior, MAE = 22.39, SNR = 13.77 dB)

$\widehat{x}_4$ (TV prior, MAE = 11.69, SNR = 15.49 dB)

$\widehat{x}_1$ (NLTV prior, MAE = 9.25, SNR = 18.38 dB)

$\widehat{x}_2$ (NLTV prior, MAE = 8.47, SNR = 21.72 dB)

$\widehat{x}_3$ (NLTV prior, MAE = 21.51, SNR = 14.12 dB)

$\widehat{x}_4$ (NLTV prior, MAE = 11.75, SNR = 15.45 dB)

$\widehat{x}_1$ (TVH prior, MAE = 7.53, SNR = 19.81 dB)

$\widehat{x}_2$ (TVH prior, MAE = 7.91, SNR = 21.61 dB)

$\widehat{x}_3$ (TVH prior, MAE = 21.65, SNR = 14.16 dB)

$\widehat{x}_4$ (TVH prior, MAE = 11.29, SNR = 15.82 dB)

Figure 4: Restoration results of images $\overline{x}_1$, $\overline{x}_2$, $\overline{x}_3$ and $\overline{x}_4$ using exact Gaussian-Poisson data fidelity term associated with either TV (top), NLTV (middle) or hybrid TV+Hessian (bottom) regularization strategy.