

Total Variation Restoration of Images Corrupted by Poisson Noise with Iterated Conditional Expectations

Rémy Abergel, Cécile Louchet, Lionel Moisan, Tieyong Zeng

▶ **To cite this version:**

Rémy Abergel, Cécile Louchet, Lionel Moisan, Tieyong Zeng. Total Variation Restoration of Images Corrupted by Poisson Noise with Iterated Conditional Expectations. SSVM 2015, 5th International Conference on Scale Space and Variational Methods in Computer Vision, May 2015, Lège Cap Ferret, France. pp.178-190, 10.1007/978-3-319-18461-6_15 . hal-01275813

HAL Id: hal-01275813
https://hal.science/hal-01275813

Submitted on 18 Feb 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Total Variation Restoration of Images Corrupted by Poisson Noise with Iterated Conditional Expectations

Rémy Abergel[1], Cécile Louchet[2], Lionel Moisan[1], and Tieyong Zeng[3]

[1] Université Paris Descartes, MAP5 (CNRS UMR 8145), France
[2] Université d'Orléans, MAPMO (CNRS UMR 6628), France
[3] Department of Mathematics, Hong Kong Baptist University, Hong Kong

Abstract. Interpreting the celebrated Rudin-Osher-Fatemi (ROF) model in a Bayesian framework has led to interesting new variants for Total Variation image denoising in the last decade. The Posterior Mean variant avoids the so-called staircasing artifact of the ROF model but is computationally very expensive. Another recent variant, called TV-ICE (for Iterated Conditional Expectation), delivers very similar images but uses a much faster fixed-point algorithm. In the present work, we consider the TV-ICE approach in the case of a Poisson noise model. We derive an explicit form of the recursion operator, and show linear convergence of the algorithm, as well as the absence of staircasing effect. We also provide a numerical algorithm that carefully handles precision and numerical overflow issues, and show experiments that illustrate the interest of this Poisson TV-ICE variant.

Keywords: Poisson noise removal; image denoising; total variation; posterior mean; marginal conditional mean; staircasing effect; fixed-point algorithm; incomplete gamma function.

1 Introduction

Since the seminal paper of Rudin, Osher and Fatemi [1], total variation (TV) regularization has been used in numerous image processing applications (see, e.g., [2] and references therein). Reasons for this popularity are multiple. First, TV regularization allows discontinuities (contrary to the L^2 norm of the gradient), which is essential in the world of natural images, dominated by occlusions. Second, its continuous counterpart is part of a fruitful mathematical theory (the space of functions with bounded variation) which results in strong possibilities of theoretical interpretations [3]. Third, in the last decade several very efficient algorithms have been designed to handle the non-smooth convex optimization problems occurring with TV regularization (e.g., [4, 5]). In terms of pure denoising performances, TV denoising is less efficient than modern patch-based approaches like NL-means [6] or BM3D [7] for example, but remains useful as the simplest possible framework for the study of TV regularization. Understanding the strengths and weaknesses of TV denoising (and variants) certainly helps a lot apprehending more complex inverse problems involving TV regularization.

One weakness of TV regularization is the so-called staircasing effect: where one would have expected a smoothly varying image, the L^1 norm promotes a sparse gradient that results in piecewise constant zones with artificial boundaries. This undesirable effect can be avoided by using a smoother functional, but at the expense of loosing the nice theoretical properties of TV. Other solutions have been proposed that keep the true definition of TV but change the minimization framework. Indeed, when considering the TV as the Gibbs energy of an image prior in a Bayesian framework, the ROF model can be reinterpreted as finding the image that maximizes the associated posterior density. Replacing this maximum a posteriori (MAP) estimate with the posterior mean leads to a variant of the ROF model, called TV-LSE, that delivers images without staircasing artifacts [8, 9]. More recently, a new variant called TV-ICE [10] was proposed to overcome the slow convergence rate of the TV-LSE Monte-Carlo algorithm. It is based on the repeated estimation of conditional marginal posterior means, which boils down to iterating an explicit local operator. In practice, TV-ICE delivers images very similar to TV-LSE results, but at a much smaller computational expense.

In the present work, we propose to adapt to the case of Poisson noise this TV-ICE method, derived in [10] in the case of Gaussian noise. Contrary to most noise sources (electronic noise, dark current, thermal noise) whose effects can be reduced by the improvement of captors, Poisson noise is inherent to the quantum nature of light and thus unavoidable for images acquired in low-light conditions, which is very common in astronomy or in microscopy for example. Even if image restoration models are generally first designed in the simpler case of a white Gaussian additive noise, they need to be adapted to the specific case of Poisson noise. Due to the importance and the inevitability of Poisson noise, this adaptation is almost systematic, as shows for example the case of TV-based image deblurring [11] or NL-means denoising [12].

In the case of the TV prior, the posterior distribution obtained with Poisson noise strongly differs from the Gaussian case, but the conditional marginal posterior means can be explicitly computed using the incomplete Gamma function. In Section 2, we show that the associated iterative algorithm converges linearly and that no staircasing occurs, thanks in particular to the log-concavity of the Poisson distribution. We then give the explicit form of the recursion operator defining our Poisson-TV-ICE model (Section 3) and discuss numerical issues, in particular the handling of machine over/under-flow and the efficient computation of the (slightly generalized) incomplete Gamma function. We then numerically check the theoretical properties of the method (convergence rate, absence of staircasing) in Section 4, and compare the obtained results with the Poisson noise variant of the ROF model, before we conclude in Section 5.

2 The Poisson TV-ICE model

2.1 Definition

Let $u : \Omega \to \mathbb{R}_+$ be an (unobserved) intensity image defined on a discrete domain Ω (a rectangular subset of $\mathbb{Z}^2$). A photon-count observation of the ideal image u

is a random image v following the Poisson probability density function (p.d.f.)

$$p(v|u) = \prod_{x\in\Omega} \frac{u(x)^{v(x)}}{v(x)!} e^{-u(x)} \propto \exp\left(-\langle u - v\log u\,,\, \mathbb{1}_\Omega\rangle\right), \tag{1}$$

where $\mathbb{1}_\Omega$ denotes the constant image equal to 1 on Ω and $\langle\cdot,\cdot\rangle$ is the usual inner product on $\mathbb{R}^\Omega$. The notation $\propto$ here indicates an equality up to a global multiplicative constant (which depends on v). Note that we have to take the convention that $v(x)\log u(x) = 0$ as soon as $v(x) = 0$ in (1). The discrete anisotropic TV of u is

$$\mathrm{TV}(u) = \frac{1}{2}\sum_{x\in\Omega}\sum_{y\in\mathcal{N}_x} |u(y) - u(x)|\,, \tag{2}$$

where $\mathcal{N}_x$ denotes the 4-neighborhood of a pixel x with a mirror boundary condition. Using the improper TV prior $p(u) \propto e^{-\lambda\,\mathrm{TV}(u)}$ (where λ is a positive regularization parameter) and Equation (1), we get thanks to the Bayes rule the posterior density

$$\pi(u) = p(u|v) = \frac{p(v|u)p(u)}{\int_{\mathbb{R}_+^\Omega} p(v|w)p(w)dw} = \frac{e^{-\langle u - v\log u\,,\,\mathbb{1}_\Omega\rangle - \lambda\,\mathrm{TV}(u)}}{\int_{\mathbb{R}_+^\Omega} e^{-\langle w - v\log w\,,\,\mathbb{1}_\Omega\rangle - \lambda\,\mathrm{TV}(w)}dw}\,. \tag{3}$$

The equivalent of the classical ROF model [1] in the case of a Poisson noise model corresponds to the unique maximizer $\hat{u}_{\mathrm{MAP}}$ of π, or equivalently the minimizer of the convex energy $E(u) = \langle u - v\log u\,,\,\mathbb{1}_\Omega\rangle + \lambda TV(u)$. It can be efficiently computed using the primal-dual algorithm recently proposed in [5]. As mentioned in Introduction, a main drawback of this approach is that $\hat{u}_{\mathrm{MAP}}$ generally suffers from the staircasing effect, which results in the appearance of flat regions separated by artificial boundaries.

In the case of a Gaussian noise model (when $\pi(u) \propto e^{-\|u-v\|^2/(2\sigma^2) - \lambda\,\mathrm{TV}(u)}$), this can be avoided by considering, instead of $\hat{u}_{\mathrm{MAP}}$, the posterior mean

$$\hat{u}_{\mathrm{LSE}} = \mathbb{E}_{u\sim\pi}(u) = \int_{\mathbb{R}^\Omega} u\,\pi(u)du\,, \tag{4}$$

which is the image that reaches the Least Square Error under π (see [8, 9]). The numerical computation of $\hat{u}_{\mathrm{LSE}}$ proposed in [8] is based on a Markov Chain Monte Carlo Metropolis-Hastings algorithm, which exhibits a slow convergence rate ($\mathcal{O}(n^{-1/2})$ for n iterations). To overcome this computational limitation, the same authors proposed in [10] a new variant based on the iteration of conditional marginal posterior means. More precisely, the estimate $\hat{u}_{\mathrm{ICE}}$ is defined as the limit (for an appropriate initialization) of the iterative scheme

$$u^{n+1}(x) = \mathbb{E}_{u\sim\pi}\left(u(x) \;\middle|\; u(x^c) = u^n(x^c)\right) = \int_{\mathbb{R}} u^n(x)\,\pi(u^n)du^n(x)\,, \tag{5}$$

where $u(x^c)$ denotes the restriction of u to $\Omega\setminus\{x\}$. In the case of the Poisson noise model (3), we obtain the following

Definition 1 (Poisson TV-ICE). *The Poisson TV-ICE recursion is*

$$\forall n \in \mathbb{N},\ \forall x \in \Omega, \quad u^{n+1}(x) = \frac{\int_{\mathbb{R}_+} s^{v(x)+1} e^{-\left(s+\lambda \sum_{y\in\mathcal{N}_x} |u^n(y)-s|\right)} ds}{\int_{\mathbb{R}_+} s^{v(x)} e^{-\left(s+\lambda \sum_{y\in\mathcal{N}_x} |u^n(y)-s|\right)} ds}\,. \tag{6}$$

2.2 Convergence

Theorem 1. *Given an image $v \in \mathbb{R}^\Omega$, the sequence of images $(u^n)_{n\geq 0}$ defined by $u^0 = 0$ and the recursion (6) converges linearly to an image $\hat{u}_{\mathrm{ICE}}$.*

In the following, we denote by $P_p(s)$ the pointwise Poisson noise p.d.f. with parameter p, that is, $P_p(s) = s^p e^{-s} \mathbb{1}_{\mathbb{R}_+}(s)/p!$. If $a = (a_i)_{1\leq i\leq 4}$, we write

$$f_p(a) = \frac{\int s\, P_p(s)\, e^{-\lambda \sum_{i=1}^4 |s-a_i|}\, ds}{\int P_p(s)\, e^{-\lambda \sum_{i=1}^4 |s-a_i|}\, ds} = \frac{\int_0^{+\infty} s^{p+1} e^{-s}\, e^{-\lambda \sum_{i=1}^4 |s-a_i|}\, ds}{\int_0^{+\infty} s^{p} e^{-s}\, e^{-\lambda \sum_{i=1}^4 |s-a_i|}\, ds} \tag{7}$$

$$\text{and} \qquad F : u \mapsto \left(x \mapsto f_{v(x)}(u(\mathcal{N}_x))\right), \tag{8}$$

so that the recursion (6) can be simply rewritten $u^{n+1} = F(u^n)$.

To prove Theorem 1, we need some intermediate Lemmas.

Lemma 1. *[13] Assume that X, a random variable defined on $\mathbb{R}^\Omega$, has a finite second order moment. Then the inequality*

$$\mathrm{cov}(X, g(X)) \geq 0$$

holds for every nondecreasing function $g : \mathbb{R}^\Omega \to \mathbb{R}$ for which $g(X)$ has a finite second order moment. If, moreover, X is not deterministic and g is strictly increasing, then $\mathrm{cov}(X, g(X)) > 0$.

Proof. One has

$$\begin{aligned}\mathrm{cov}(X, g(X)) &= \mathbb{E}[(X - \mathbb{E}[X])(g(X) - \mathbb{E}[g(X)])] \\ &= \mathbb{E}[(X - \mathbb{E}[X])(g(X) - g(\mathbb{E}[X]))]\end{aligned}$$

The assertion follows because g is increasing. If X is not deterministic, then there exists a Borel set A such that $P(X \in A) > 0$ with $\mathbb{E}[X] \notin A$. Hence the covariance is a sum of nonnegative terms, some of which (those for $X \in A$) are positive. Finally $\mathrm{cov}(X, g(X))$ is positive.

Lemma 2. *F is monotone: for all images u_0 and u_1,*

$$u_0 \leq u_1 \Rightarrow F(u_0) \leq F(u_1).$$

Proof. Using Lebesgue dominated convergence theorem, one can prove the differentiability of f_p with respect to each a_i and obtain

$$\frac{\partial f_p}{\partial a_i}(a) = \frac{\int \lambda \operatorname{sign}(s-a_i)\, s\, P_p(s)\, e^{-\lambda \sum_{j=1}^4 |s-a_j|}\, ds}{\int P_p(s)\, e^{-\lambda \sum_{j=1}^4 |s-a_j|}\, ds} - \frac{\int s\, P_p(s)\, e^{-\lambda \sum_{j=1}^4 |s-a_j|}\, ds}{\int P_p(s)\, e^{-\lambda \sum_{j=1}^4 |s-a_j|}\, ds} \cdot \frac{\int \lambda \operatorname{sign}(s-a_i)\, P_p(s)\, e^{-\lambda \sum_{j=1}^4 |s-a_j|}\, ds}{\int P_p(s)\, e^{-\lambda \sum_{j=1}^4 |s-a_j|}\, ds}.$$

Hence $\frac{\partial f_p}{\partial a_i}(a)$ can be seen as the covariance of S and $\lambda \operatorname{sign}(S-a_i)$, where S is a random variable with p.d.f. $s \mapsto \frac{1}{Z} P_p(s)\, e^{-\lambda \sum_{j=1}^4 |s-a_j|}$, which has a finite second order moment. Using Lemma 1, the quantity $\frac{\partial f_p}{\partial a_i}(a)$, as the covariance of S with a nondecreasing function of S, is nonnegative. Now if $u_0 \leq u_1$, then as f_p is $\mathcal{C}^1$ we can write

$$(F(u_1) - F(u_0))(x) = \int_0^1 \nabla f_{v(x)}(u_t(\mathcal{N}_x)) \cdot (u_1(\mathcal{N}_x) - u_0(\mathcal{N}_x))\, dt,$$

where $u_t(\mathcal{N}_x) = (1-t)u_0(\mathcal{N}_x) + t u_1(\mathcal{N}_x)$. As $\frac{\partial f_{v(x)}}{\partial u(y)}$ and $u_1(y) - u_0(y)$ are both nonnegative, so is $(F(u_1) - F(u_0))(x)$ as the integral of a nonnegative function.

Lemma 3. *F is strictly nonexpansive for the ℓ^∞ norm: for any images $u \neq u'$,*

$$\|F(u') - F(u)\|_\infty < \|u' - u\|_\infty.$$

Proof. For fixed values of p and $a = (a_i)_{1 \leq i \leq 4}$, let us define the real mapping

$$g : c \mapsto f_p(a + c) - c,$$

where $a + c$ is a shorthand for $(a_i + c)_{1 \leq i \leq 4}$. We first prove that the strict decrease of g on $\mathbb{R}$ for all p and a implies the strict nonexpansiveness of F. We must prove that $F(u') < F(u) + c$ and that $F(u') > F(u) - c$ for $c = \|u' - u\|_\infty$. As $u' \leq u + c$ and as F is monotone, we have $F(u') \leq F(u + c)$. It remains to prove that $F(u + c) < F(u) + c$, i.e. that

$$\forall p \in \mathbb{N},\ \forall a \in \mathbb{R}^4,\ \forall c > 0, \quad f_p(a + c) < f_p(a) + c,$$

which is true as soon as g is strictly decreasing on $\mathbb{R}_+$. For the other inequality, we have $F(u') \geq F(u - c)$, so that it remains to prove that $F(u - c) > F(u) - c$, i.e. that

$$\forall p \in \mathbb{N},\ \forall a \in \mathbb{R}^4,\ \forall c > 0, \quad f_p(a - c) > f_p(a) - c,$$

which is true as soon as g is strictly decreasing on $\mathbb{R}_-$.

Second, we prove that g is strictly decreasing. One can prove that

$$g'(c) = \operatorname{cov}\left(S, \frac{P_p'(S + c)}{P_p(S + c)}\right) = \operatorname{cov}(S, (\log P_p)'(S + c)),$$

where S follows a distribution with p.d.f. $s \mapsto \frac{1}{Z} P_p(s + c) e^{-\lambda \sum_{i=1}^4 |s - a_i|}\, ds$. Now, P_p is positive and differentiable and $(\log P_p)'(s) = p/s - 1$ so for all c, the

mapping $s \mapsto (\log P_p)'(s+c)$ is strictly decreasing on $(-c, \infty)$. Again thanks to Lemma 1, as the distribution on S is not deterministic, we get that $g'(c)$ is negative. Hence g is strictly decreasing and the proof is complete.

Lemma 4. *There exists a subset K of $\mathbb{R}^\Omega$ containing 0 such that $F(K) \subset K$.*

Proof (abridged). We set $G(p,c) = f_p(c\,\mathbb{1}_\mathcal{N}) - c$ and proceed in 4 steps:

1) For every $p \in \mathbb{N}$, the function $c \mapsto G(p,c)$ is continuous and decreasing. Indeed, $G(p,c)$ is exactly $g(c)$, defined in the proof of Lemma 3, with $a = 0$. So it is differentiable and decreasing.

2) For each $p \in \mathbb{N}$, the limit of $G(p,c)$, when c goes to $+\infty$, is negative (proof not given here).

3) We deduce from 1) and 2) that

$$\forall p \in \mathbb{N},\ \exists\, \mathbf{c}(p) \in \mathbb{R}, \quad c \geq \mathbf{c}(p) \Rightarrow G(p,c) \leq 0.$$

4) With the latter definition for $p \mapsto \mathbf{c}(p)$, we define $c = \max_{x\in\Omega} \mathbf{c}(v(x))$ and $K = [0,c]^\Omega$. If $u \in K$, then $u \leq c$, and as F is monotone, $F(u) \leq F(c\mathbb{1}_\Omega)$. Now, as $c \geq \mathbf{c}(v(x))$, by definition of $\mathbf{c}$, $f_{v(x)}(c) \leq c$ holds for each $x \in \Omega$, which exactly means that $F(u) \leq F(c\mathbb{1}_\Omega) \leq c$. Secondly, as $F(u)(x)$ is a ratio of nonnegative quantities, it is nonnegative and $F(u) \geq 0$. In conclusion, $F(u) \in K$.

Proof (of Theorem 1). Since the map F is strictly non-expansive (Lemma 3) and continuous on the compact set K, there exists a real number $\alpha \in (0,1)$ such that $\|F(w_1) - F(w_2)\|_\infty \leq \alpha \|w_1 - w_2\|_\infty$ for all images $w_1, w_2 \in K$. Moreover, K is stable by F (Lemma 4), so the Banach fixed-point theorem applies and the sequence (u^n) defined in Theorem 1 converges to a fixed point of F, which is unique. The convergence is linear as $\|u^{n+1} - \hat{u}_{\mathrm{ICE}}\|_\infty \leq \alpha \|u^n - \hat{u}_{\mathrm{ICE}}\|_\infty$, or in other terms, $\|u^n - \hat{u}_{\mathrm{ICE}}\|_\infty = \mathcal{O}(\alpha^n)$ as $n \to \infty$.

2.3 No staircasing for Poisson TV-ICE

We here prove that Poisson TV-ICE cannot produce large constant regions that were not at least partially present in the initial data.

Theorem 2. *Let $v : \Omega \to \mathbb{N}$ be a noisy image, and $\hat{u}_{\mathrm{ICE}}$ its denoised version. Let x and y be two pixels in Ω. Then if $\hat{u}_{\mathrm{ICE}}$ is constant on $\mathcal{N}_x \cup \mathcal{N}_y \cup \{x, y\}$, necessarily $v(x) = v(y)$.*

To establish the proof, we need the following

Lemma 5. *For any constant c, the mapping $p \mapsto f_p(c\,\mathbb{1}_\mathcal{N})$ is strictly increasing.*

Proof. The mapping $p \mapsto f_p(c\,\mathbb{1}_\mathcal{N})$ can be naturally extended to real positive values of p using the right-hand part of Equation (7). Using the dominated convergence theorem, we can assess the differentiability of $p \mapsto f_p(c\,\mathbb{1}_\mathcal{N})$ and obtain

$$\frac{\partial f_p}{\partial p}(c\,\mathbb{1}_\mathcal{N}) = \mathrm{cov}\big(S, \log S\big)\ ,$$

where S is a random variable with p.d.f. $s \mapsto \frac{1}{Z} P_p(s)\, e^{-4\lambda|s-c|}$. But as the log function is strictly increasing, using Lemma 1, we have that $\frac{\partial f_p}{\partial p}(c\,\mathbb{1}_{\mathcal{N}})$ is positive. Considering only integer values of p, we obtain the desired result.

Proof (of Theorem 2). Assume that $\hat{u}_{\mathrm{ICE}}$ has value $c \in \mathbb{R}$ for every pixel in $\mathcal{N}_x \cup \mathcal{N}_y \cup \{x, y\}$. Then taking the limit in (6) tells us that $c = \hat{u}_{\mathrm{ICE}}(x) = f_{v(x)}(\hat{u}_{\mathrm{ICE}}(\mathcal{N}_x)) = f_{v(x)}(c\,\mathbb{1}_{\mathcal{N}})$, and similarly $c = \hat{u}_{\mathrm{ICE}}(y) = f_{v(y)}(c\,\mathbb{1}_{\mathcal{N}})$. But using Lemma 5, $p \mapsto f_p(c\,\mathbb{1}_{\mathcal{N}})$ is strictly increasing, so there exists at most one value p such that $f_p(c\,\mathbb{1}_{\mathcal{N}}) = c$. We conclude that necessarily $v(x) = p = v(y)$, which finishes the proof.

3 Numerical computation of Poisson TV-ICE

3.1 Explicit form of the Poisson TV-ICE recursion operator

Proposition 1. *The Poisson TV-ICE recursion $u^{n+1}(x) = f_{v(x)}(u^n(\mathcal{N}_x))$ can be written*

$$u^{n+1}(x) = \frac{\sum_{1\le k\le 5} c_k I_{a_{k-1},a_k}^{\mu_k, v(x)+1}}{\sum_{1\le k\le 5} c_k I_{a_{k-1},a_k}^{\mu_k, v(x)}}\ , \tag{9}$$

where a_1, a_2, a_3, a_4 are the values of $u^n(\mathcal{N}_x)$ sorted in nondecreasing order (that is, $0 = a_0 \le a_1 \le a_2 \le a_3 \le a_4 \le a_5 = +\infty$),

$$\forall k \in \{1, \dots, 5\}, \quad \mu_k = 1 - (6 - 2k)\lambda, \quad \log c_k = \lambda \left(\sum_{j=1}^{k-1} a_j - \sum_{j=k}^{4} a_j \right)\ , \tag{10}$$

$$\text{and} \quad I_{x,y}^{\mu,p} = \int_x^y s^p e^{-\mu s} ds\ . \tag{11}$$

Proof. This result is directly obtained after breaking the integration domain in Equation (6) so as to get rid of all absolute values.

3.2 Numerical issues

To compute the integral (11), we introduce the following generalized lower (γ_μ) and upper (Γ_μ) incomplete gamma functions,

$$\gamma_\mu(p, x) = \int_0^x s^{p-1} e^{-\mu s} ds\,, \quad \Gamma_\mu(p, x) = \int_x^{+\infty} s^{p-1} e^{-\mu s} ds\,, \tag{12}$$

$$\text{so that} \quad I_{x,y}^{\mu,p} = \gamma_\mu(p+1, y) - \gamma_\mu(p+1, x) \tag{13}$$

and, for $\mu > 0$,

$$I_{x,y}^{\mu,p} = \Gamma_\mu(p+1, y) - \Gamma_\mu(p+1, x) = \frac{p!}{\mu^{p+1}} - \gamma_\mu(p+1, x) - \Gamma_\mu(p+1, y). \tag{14}$$

Note that when $\mu > 0$, the change of variable $t = \mu s$ would lead us back to the standard definition of incomplete gamma functions (corresponding to $\mu = 1$), but this is not the case when $\mu < 0$.

The effective computation of (9) with one of the formulas given in (13)-(14) raises several numerical issues:

1. For some values of the parameters, the numerator and the denominator cannot be represented in the computer floating point arithmetic (for example because they both exceed $1.9 \cdot 10^{308}$, the largest double precision number), although their ratio can be represented. To solve that issue, we represent each integral $I_{x,y}^{\mu,p}$ involved in (9) under the form $\rho \cdot e^{\sigma}$, where ρ and σ are floating point (double precision) numbers.
2. The possibility to compute efficiently $\gamma_\mu(p,x)$ and $\Gamma_\mu(p,x)$ depends on the parameters μ, p, x. We divided the plane $(\mu x, p)$ so as to compute $\gamma_\mu(p,x)$ or $\Gamma_\mu(p,x)$ efficiently for each parameter set μ, p, x.
3. When the difference $A - B$ is computed numerically, the result can be very inaccurate if A and B are close to each other, which may happen for the differences given in Equations (13)-(14). When x and y are very close to each other, the integral (11) is very well approximated by the rectangle numerical integration formula (with one term); we found a good criterion to decide when this approximation should be used.

We reviewed the literature to find the available methods for the computation of $\gamma_\mu(p,x)$ and $\Gamma_\mu(p,x)$, and found that for the explored domain $|\mu x| \leq 1000$, $1 \leq p \leq 1000$ (and even far beyond in fact), the selection of the three following algorithms was satisfactory :

1. A continued fraction [14–17] for the computation of $\gamma_\mu^{\text{cfrac}}(p+1,x) = \rho e^{\sigma}$ with $\rho = \frac{\alpha_1}{\beta_1+}\,\frac{\alpha_2}{\beta_2+}\,\frac{\alpha_3}{\beta_3+}\,\ldots$ and $\alpha_1 = 1$, $\alpha_{2n} = -(p+n)\cdot\mu x$, $\alpha_{2n+1} = n \cdot \mu x$, $\beta_n = p + n$ for $n \geq 1$, valid for any μ.
2. An simple recursive integration by parts formula, only valid when $\mu < 0$, yielding $\gamma_\mu^{\text{ibp}}(p+1,x) = \rho e^{\sigma}$ with $\rho = \frac{1}{\mu x}\left(p!\,\frac{e^{\mu x}}{(\mu x)^p} - \sum_{0\leq k\leq p} \frac{p!(\mu x)^{-k}}{(p-k)!}\right)$.
3. A continued fraction [18, 19] for the computation of $\Gamma_\mu^{\text{cfrac}}(p+1,x) = \rho e^{\sigma}$ with $\rho = \frac{\alpha_1}{\beta_1+}\,\frac{\alpha_2}{\beta_2+}\,\frac{\alpha_3}{\beta_3+}\,\cdots$, $\beta_n = \mu x + 2n - 2 - p$ $(n \geq 1)$, $\alpha_1 = 1$ and $\alpha_n = -(n-1)\cdot(n-p-2)$ $(n > 1)$, valid for any $\mu > 0$.

In all cases, $\sigma = -\mu x + (p+1)\log x$. Fig. 1 shows the partition of the plane $(\mu x, p)$ we found as nearly optimal (in terms of computation time) to choose between these 3 algorithms. The resulting procedure used to compute $I_{x,y}^{\mu,p}$ is given in Algorithm 1.

4 Experiments

We first checked the convergence of the proposed Poisson TV-ICE algorithm obtained by iterating the recursion (9) using Algorithm 1 and the initialization $u^0 = 0$. As can be seen in Fig. 2, numerical convergence is attained for Poisson TV-ICE after a few hundred iterations, and the convergence rate is linear as expected.

We then chose 3 images taken from areas concerned with Poisson noise (2 from microscopy, 1 from astronomy), and simulated a low-light observation (that is, a Poisson noise process) for each of them. Then, we restored the noisy images with both the Poisson TV-MAP and the proposed Poisson TV-ICE methods (see

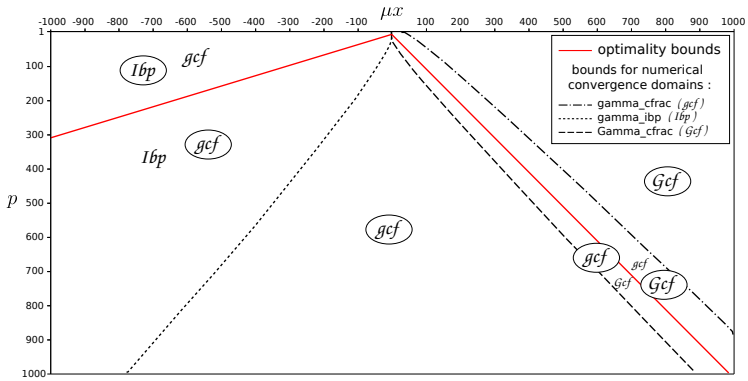


Fig. 1. Partition of the domain $(\mu x, p)$ for the evaluation of the generalized incomplete gamma function. The rectangular domain of the plane $(\mu x, p)$ above is cut into three regions delimited by the red curves (their equations appear in Algorithm 1). On each region, one of the three selected algorithm (circled) is used to compute numerically either $\gamma_\mu(p,x)$ or $\Gamma_\mu(p,x)$: the continued fraction for γ_μ (gcf), the recursive integration by parts for γ_μ (Ibp), or the continued fraction for Γ_μ (Gcf). Names without circles and dotted/dashed lines correspond to regions where another of the 3 algorithms could be used too.

Algorithm 1: Numerical estimation of $I_{x,y}^{\mu,p}$

inputs: three numbers $\mu \in \mathbb{R}$, $x \in \mathbb{R}_+$, $y \in \mathbb{R}_+ \cup \{+\infty\}$ such as $x \leq y$, and an integer $p \geq 0$. The value $y = +\infty$ is allowed only when $\mu > 0$.
outputs: ρ and σ such that $\rho \cdot e^\sigma$ is an accurate estimate of $I_{x,y}^{\mu,p}$.
if $x = y$ *up to machine precision* **then** $(\rho, \sigma) \leftarrow (0, -\infty)$
else if $0 < \frac{y-x}{y} \leq 10^{-9}$ **then** $(\rho, \sigma) \leftarrow \left(\frac{y-x}{y}, -\mu y + (p+1)\cdot \log y\right)$
else if $\mu \leq 0$ **then**
  if $\mu = 0$ **or** $p \geq -0.33\mu x + 8$ **then** $(\rho_x, \sigma_x) \leftarrow \gamma_\mu^{\text{cfrac}}(p+1, x)$
  else $(\rho_x, \sigma_x) \leftarrow \gamma_\mu^{\text{ibp}}(p+1, x)$ **end**
  if $\mu = 0$ **or** $p \geq -0.33\mu y + 8$ **then** $(\rho_y, \sigma_y) \leftarrow \gamma_\mu^{\text{cfrac}}(p+1, y)$
  else $(\rho_y, \sigma_y) \leftarrow \gamma_\mu^{\text{ibp}}(p+1, y)$ **end**
  $(\rho, \sigma) \leftarrow \left(\rho_y - \rho_x e^{\sigma_x - \sigma_y}, \sigma_y\right)$
else
  if $p \geq \mu y$ **then**
    $(\rho_x, \sigma_x) \leftarrow \gamma_\mu^{\text{cfrac}}(p+1, x)$; $(\rho_y, \sigma_y) \leftarrow \gamma_\mu^{\text{cfrac}}(p+1, y)$
    $(\rho, \sigma) \leftarrow \left(\rho_y - \rho_x e^{\sigma_x - \sigma_y}, \sigma_y\right)$
  else if $p < \mu x$ **then**
    $(\rho_x, \sigma_x) \leftarrow \Gamma_\mu^{\text{cfrac}}(p+1, x)$; $(\rho_y, \sigma_y) \leftarrow \Gamma_\mu^{\text{cfrac}}(p+1, y)$
    $(\rho, \sigma) \leftarrow \left(\rho_x - \rho_y e^{\sigma_y - \sigma_x}, \sigma_x\right)$
  else
    $(\rho_x, \sigma_x) \leftarrow \gamma_\mu^{\text{cfrac}}(p+1, x)$; $(\rho_y, \sigma_y) \leftarrow \Gamma_\mu^{\text{cfrac}}(p+1, y)$
    $(\rho, \sigma) \leftarrow (1 - \rho_x e^{\sigma_x - \sigma} - \rho_y e^{\sigma_y - \sigma}, \log p! - (p+1)\log \mu)$
  end
end

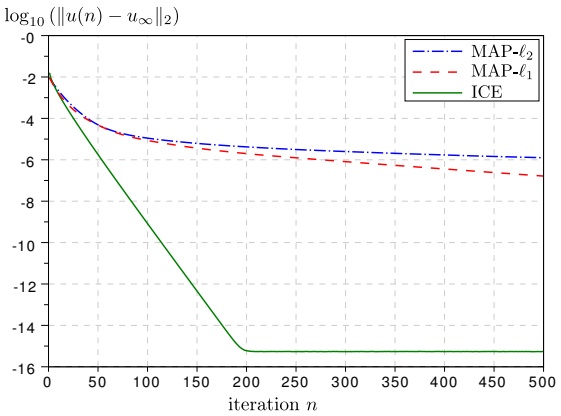


Fig. 2. Convergence rates for TV-MAP and TV-ICE. We display in logarithmic scale the convergence rates obtained for the proposed implementation of the Poisson TV-ICE algorithm (green plain curve), and for the Chambolle-Pock [5] implementation of the Poisson TV-MAP (with anisotropic and isotropic TV) algorithms (red/blue dashed curves). As expected, Poisson TV-ICE achieves a linear convergence rate.

Fig. 3). As predicted by the theory, TV-ICE results do not exhibit staircasing effects, contrary to TV-MAP images which provide less details, in particular in the areas where the staircasing artifact causes an important loss of contrast (see, e.g., the bottom-right part of the images of the first row of Fig. 3). This visual effect was confirmed by the systematically smaller I-divergence values obtained with TV-ICE.

5 Conclusion

We proposed a variant of the recent TV-ICE denoising method adapted to the special case of Poisson noise. The absence of staircasing and the better-quality restored images attested by experiments make Poisson TV-ICE a good alternative to Poisson TV-MAP, and suggests that it could be interesting to derive Poisson TV-ICE variants for more complex inverse problems involving TV terms.

The linear convergence rate of the method is appealing but is not sufficient to compensate for the heavy computations required by the form of the recursion operator (several evaluations of the exponential and logarithm functions are required for each pixel). In our current (non-optimized) implementation, one iteration of TV-ICE is approximately 100 times slower than one iteration of TV-MAP. However, further work could focus on the fast approximation of TV-ICE, and the precise implementation we here proposed would be useful in that context to check the quality of the approximation.

As in the Gaussian case, the generalization of the proposed algorithm to three-dimensional images (or more), or to larger neighborhood systems, is straightforward. However, the comparison with the Poisson TV-LSE variant is, both from a theoretical or practical point of view, still open.

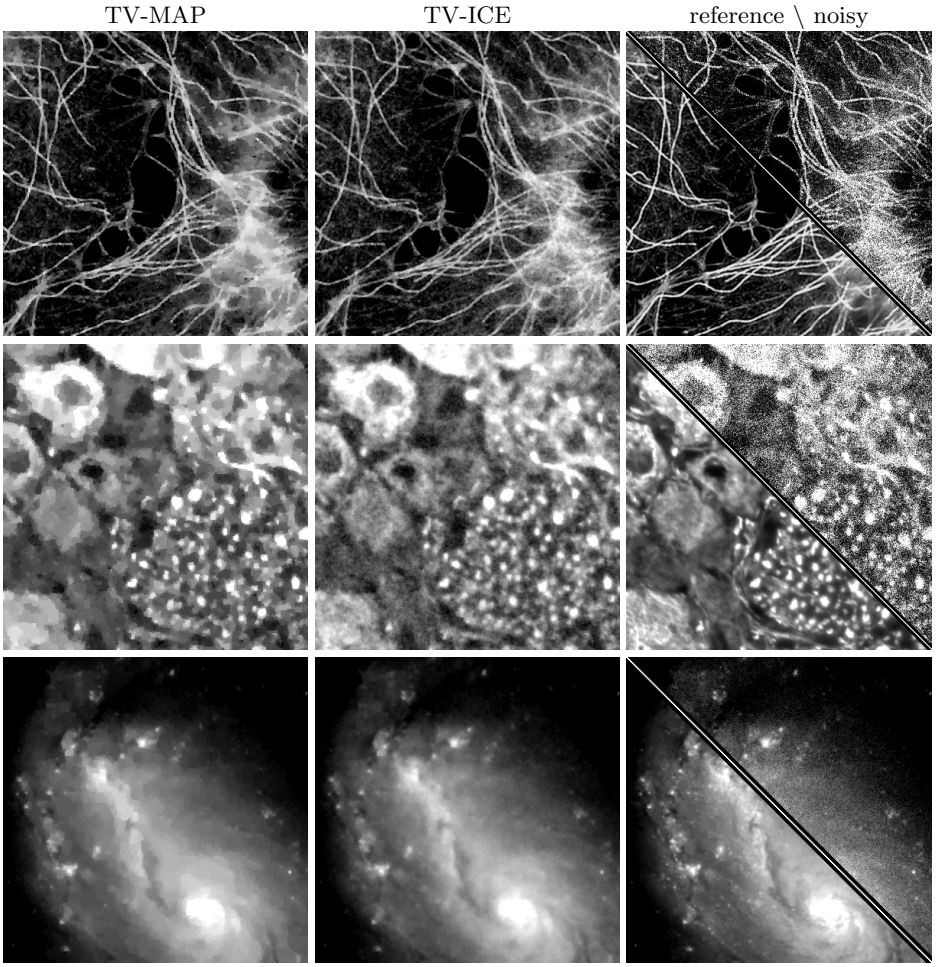


Fig. 3. Comparison of Poisson TV-MAP and Poisson TV-ICE. Three images (first row: actin filaments and microtubules in interphase cells, second row: mouse dorsal root ganglion, third row: NGC 1672 spiral galaxy) were corrupted with Poisson noise, then denoised with the Poisson TV-MAP algorithm (left column) and the proposed Poisson TV-ICE method (middle column). For each algorithm, we selected the value of the λ parameter that achieved the smallest Csiszar I-divergence [20] (a measure of distance adapted to the case of Poisson noise) between the reference image u_{ref} (bottom-left part of the images in the right column) and the denoised image $\hat{u}$, which is defined by $\text{I-div}(u_{\mathrm{ref}}, \hat{u}) = \left\langle u_{\mathrm{ref}} \log(u_{\mathrm{ref}}/\hat{u}) - (u_{\mathrm{ref}} - \hat{u}) \, , \, \mathbb{1}_{\Omega} \right\rangle$. One can clearly see that TV-MAP results exhibit staircasing effects and an associated loss of details in the corresponding flat regions; on the contrary, the TV-ICE images are more natural and more faithful to the fine details of the reference, especially in the regions where TV-MAP produces staircasing. Note that in order to increase the readability of the figure, the dynamic of the images has been linearly amplified, causing some (limited) saturation in dark and white areas. Image sources: cellimagelibrary.org and wikimedia.org.

Acknowledgments: This work has been partially supported by NSFC 11271049, RGC 211911, 12302714 and RFGs of HKBU.

References

1. Rudin, L. I., Osher, S., Fatemi, E.: "Nonlinear total variation based noise removal algorithms". *Physica D* 60(1), 259–268 (1992)
2. Caselles, V., Chambolle, A., Novaga, M.: "Total variation in imaging", *Handbook of Mathematical Methods in Imaging*, Springer, New York, 1016–1057 (2011)
3. Chambolle, A., Caselles, V., Cremers, D., Novaga, M., Pock, T.: "An introduction to total variation for image analysis". *Theoretical foundations and numerical methods for sparse recovery* 9, 263–340 (2010)
4. Darbon, J., Sigelle, M.: "Image restoration with discrete constrained total variation part I: Fast and exact optimization". *J. Math. Imag. Vis.* 26(3), 261–276 (2007)
5. Chambolle, A., Pock, T.: "A first-order primal-dual algorithm for convex problems with applications to imaging". *J. Math. Imag. Vis.* 40(1), 120–145 (2011)
6. Buades, A., Coll, B., Morel, J.-M.: "A review of image denoising algorithms, with a new one". *Multiscale Model. Simul.* 4(2), 490–530 (2005)
7. Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K.: "Image denoising by sparse 3-D transform-domain collaborative filtering". *IEEE Trans. Image Processing* 16(8), 2080–2095 (2007)
8. Louchet, C., Moisan, L.: "Total variation denoising using posterior expectation". *Proc. European Signal Processing Conf.* (2008)
9. Louchet, C., Moisan, L.: "Posterior Expectation of the Total Variation model: Properties and Experiments". *SIAM J. Imaging Sci.* 6(4), 2640–2684 (2013)
10. Louchet, C., Moisan, L.: "Total Variation Denoising Using Iterated Conditional Expectation". *Proc. European Signal Processing Conf.* (2014)
11. Setzer, S., Steidl, G., Teuber, T.: "Deblurring Poissonian images by split Bregman techniques". *J. Vis. Comm. Image Representation* 21(3), 193–199 (2010)
12. Deledalle, C., Tupin, F., Denis, L.: "Poisson NL means: Unsupervised non local means for Poisson noise". *Proc. Int. Conf. Imag. Processing*, pp. 801–804 (2010)
13. Schmidt, K. D.: "On the covariance of monotone functions of a random variable". Unpublished note, University of Dresden (2003)
14. Olver, F. W. J., Lozier, D. W., Boisvert, R. F., Clark, C. W. editors: *NIST Handbook of Mathematical Functions*. Cambridge University Press, New York (2010)
15. NIST Digital Library of Mathematical Functions. `http://dlmf.nist.gov/`, Release 1.0.9 of 2014-08-29 (2014)
16. Cuyt, A., Petersen, V. B., Verdonk, B., Waadeland, H., Jones, W. B.: *Handbook of continued fractions for special functions*. Springer, New York (2008)
17. Jones, W. B., Thron, W. J.: "Continued Fractions: Analytic Theory and Applications". *Encyclopedia of Mathematics and its Applications* 11, Addison-Wesley Publishing Co., Reading, MA (1980)
18. *Numerical recipes: The art of scientific computing*. Second Edition. Cambridge University (2007)
19. Abramowitz, M., Stegun, I. A.: *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*. Courier Dover Publications no. 55 (1972)
20. Csiszar, I.: "Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems". *Ann. Stat.* 19, 2032–2066 (1991)