

Restoration of Astrophysical Images—The Case of Poisson Data with Additive Gaussian Noise

H. Lantéri

*Laboratoire d'Astrophysique, Université de Nice Sophia Antipolis, CNRS UMR6525, 06108 Nice Cedex 2, France
Email: hlanteri@unice.fr*

C. Theys

*Laboratoire d'Astrophysique, Université de Nice Sophia Antipolis, CNRS UMR6525, 06108 Nice Cedex 2, France
Email: celine.theys@unice.fr*

Received 28 May 2004; Revised 28 October 2004

We consider the problem of restoring astronomical images acquired with charge coupled device cameras. The astronomical object is first blurred by the point spread function of the instrument-atmosphere set. The resulting convolved image is corrupted by a Poissonian noise due to low light intensity, then, a Gaussian white noise is added during the electronic read-out operation. We show first that the split gradient method (SGM) previously proposed can be used to obtain maximum likelihood (ML) iterative algorithms adapted in such noise combinations. However, when ML algorithms are used for image restoration, whatever the noise process is, instabilities due to noise amplification appear when the iteration number increases. To avoid this drawback and to obtain physically meaningful solutions, we introduce various classical penalization-regularization terms to impose a smoothness property on the solution. We show that the SGM can be extended to such penalized ML objective functions, allowing us to obtain new algorithms leading to maximum a posteriori stable solutions. The proposed algorithms are checked on typical astronomical images and the choice of the penalty function is discussed following the kind of object.

Keywords and phrases: restoration, astronomic images, Poisson transformation, MAP estimation, regularization, iterative algorithms.

1. INTRODUCTION

The image restoration problem and particularly image deconvolution, is an inverse problem, ill posed in the sense of Hadamard, whose solution is unstable when the data is corrupted by noise. For astronomical data, two noise processes are generally considered separately. The first one is an additive Gaussian noise appearing in high intensity measurements; in this case, the maximum likelihood estimator (MLE) under positivity constraint is obtained, for example, using the ISRA multiplicative iterative algorithm [1]. The second one, dedicated to low intensity data, is a Poisson noise process; in this case the MLE is obtained from the expectation maximization (EM) Richardson-Lucy iterative algorithm [2]. In a more realistic but less used model, both noise processes must be taken into account simultaneously. We will describe this model previously analyzed by Snyder et al. [3, 4, 5, 6] and by Llacer and Nuñez [7, 8, 9]. We will show that the corresponding MLE iterative algorithm can be easily obtained using the split gradient method (SGM) we have previously proposed for Poisson or Gaussian noise [10, 11]. However, in all the ML methods, whatever the noise process

considered, only the adequacy of the solution to the data is taken into account; so, when iterative algorithms are used, instabilities appear due to the noise amplification when the iteration number increases. In this context, to obtain physically satisfactory solutions, the iterative process must be interrupted before instabilities appears. Another way to avoid this drawback is to perform an explicit regularization of the problem, that is, to introduce an a priori knowledge to impose, for example, a smoothness property on the solution, then, a maximum a posteriori (MAP) solution is searched for.

In previous papers of Snyder the regularization is performed by means of sieves functions. In the papers of Llacer and Nuñez, the proposed penalized iterative algorithm does not ensure neither the convergence nor the positivity. We will show that the SGM can be used to obtain regularized iterative algorithms for a Poisson plus Gaussian noise model. In so doing, we exhibit regularized iterative multiplicative algorithms with constraints typical of astronomical imagery and we show their effectiveness for some classical penalty functions.

1.1. Optical astronomy imagery: Poisson data with additive Gaussian noise

The light emanating from the object of interest propagates through a turbulent atmosphere and is focused onto the charge coupled device (CCD) by an imperfect optical system that limits the resolution and introduces aberrations.

The overall effect of the atmosphere and optical system can be mathematically described by a convolution operation between the object x and the point spread function (PSF) w of the whole system. The response of the CCD detector being nonuniform, a function T proportional to the quantum detector efficiency, called flat field, must be introduced in the model [5, 6], giving a first deterministic transformation f of the object:

$$f(r, s) = T(r, s)[w(r, s) \otimes x(r, s)], \quad (1)$$

where r and s are spatial coordinates and \otimes denotes the two-dimensional convolution.

Moreover, in astronomical imaging, x is generally the sum of the sky background m and of the astronomical object field u , then

$$x(r, s) = u(r, s) + m(r, s). \quad (2)$$

Consequently, a natural constraint appears for the object: $x \geq m$. In each pixel of the sensor, the interaction between the incident photons and the photosensitive material of the CCD creates photoelectrons in proportion to the number of photons plus extraneous electrons due to heat and bias effects. This photoconversion process is classically characterized by a Poisson transformation of mean $f + d$, where d is the mean of the Poisson process for the extraneous electrons.

Finally, the detector is read by an electronic process which adds a white Gaussian read-out noise $b \sim \mathcal{N}(g, \sigma^2)$, independent of the Poisson process. We get then the observed image $y(r, s)$ of mean $f(r, s) + d(r, s) + g$.

1.2. Imaging model and problem statement

In the following, we use capital letters for $N \times N$ arrays and bold letters for $N \times 1$ vectors, subscript i denotes the pixel i of the image lexicographically ordered.

From the description in Section 1.1, a realization of the value of the image in the pixel i can be modeled as

$$y_i = n_i + b_i, \quad (3)$$

where n_i is a realization of a Poisson process of mean

$$t_i(W\mathbf{x})_i + d_i = (H\mathbf{x})_i + d_i = z_i, \quad (4)$$

where $h_{i,j} = t_i w_{i,j}$ are the elements of H satisfying

$$\sum_j w_{j,i} = 1, \quad \sum_j h_{j,i} = a_i \neq 1 \quad \forall i. \quad (5)$$

In matrix notation,

$$\mathbf{y} = \mathbf{n} + \mathbf{b}, \quad (6)$$

where \mathbf{n} is a Poisson process of mean

$$H\mathbf{x} + \mathbf{d} = \mathbf{z}, \quad (7)$$

$H = TW$, $T = \text{diag}(t_1 t_2 \cdots t_N)$, and W is the classical block Toeplitz matrix for the convolution matrix form. Note that if there is no flat field, $T = I$, $H = W$.

In the deconvolution problem for astronomical imaging, the object is generally composed of bright objects on a sky background, assumed constant. This particularity of the astrophysical object must be taken into account to avoid the “ringing” phenomena appearing in the vicinity of abrupt intensity variations. Moreover, in the convolution operation with normalized kernels, the total intensity of the object is maintained.

The problem is then to restore the object \mathbf{x} from the data \mathbf{y} with the constraint $\mathbf{x} \geq m$, see (2), and the total intensity conservation.

W is generally obtained via separated calibration measurements as well as the flat field table T , \mathbf{d} can be measured, and the parameters of the Gaussian noise are generally known characteristics of the CCD [5, 6].

1.3. Image restoration method

The restoration method generally proposed for the previous model [5, 6, 7, 8, 9] is founded on the MLE and the corresponding iterative algorithms are developed from the EM method [12, 13]. As the deconvolution problem is an ill-posed problem, instability in the solution appears as the number of iterations increases. The problem is then to stop them to get a physically satisfactory solution and for that, to determine the optimal iteration number [14, 15].

Another way to avoid instabilities is to regularize explicitly the problem. Numerous relevant methods have been used in the literature (see [16] and references therein). The method of *sieves*, for example, was proposed by Grenander [17] and applied to image restoration by Snyder et al. [3, 5, 18]. However, as mentioned in [19], penalized MLE outperforms *sieves*. Therefore, we focus on the basic principles of the regularization by a penalty function. In this case a penalty term is added to the likelihood term with the aim of introducing to the solution *prior* information, generally a smoothness property, (see, e.g., Bertero [20, 21, 22], Demoment [23], Titterton [24]). The relative weight of the penalty versus the likelihood allows us to “pull” the solution either towards the ML or towards the *prior*, changing the MLE in the MAP estimation.

This approach has been followed by Llacer and Nuñez [7, 8, 9], for the model (6) and (7) using as penalty function the Shannon cross entropy between the solution and a constant *prior* for the object. However, as mentioned in [25], the proposed algorithm has a significant computational burden, the convergence is not guaranteed and the positivity of the solution is not always ensured.

In the present work, we first show that the MLE algorithm can be obtained using the SGM previously proposed [25], then the SGM is extended to the regularized approach, in a rigorous way, using various convex penalty functions.

The paper is organized as follows. In Section 2 the MAP estimation problem is given for the model Poisson plus Gaussian noise. The general algorithmic method is described for various penalty functions in Section 3. In Section 4 simulations on astronomical images are presented and discussed. Finally, Section 5 proposes a discussion on the generalization of the proposed method to others applications.

2. POISSON NOISE PROCESS WITH ADDITIVE GAUSSIAN NOISE

2.1. Likelihood function

Following the model (3) and (4), we denote by z_i the mean of the Poisson process in the pixel i . The number n_i of photoelectrons generated by this process has the classical probability law

$$P(n_i|z_i) = \frac{z_i^{n_i}}{n_i!} \exp(-z_i). \quad (8)$$

During the read-out step, the integer n_i is corrupted by a Gaussian additive noise of mean g and variance σ^2 giving a process y_i with a Gaussian probability law

$$P(y_i|n_i) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(y_i - n_i - g)^2}{2\sigma^2}\right). \quad (9)$$

Then

$$\begin{aligned} P(y_i|z_i) &= \sum_{n_i} P(y_i|n_i)P(n_i|z_i) \\ &= \sum_{n_i=0}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(y_i - n_i - g)^2}{2\sigma^2}\right) \frac{z_i^{n_i}}{n_i!} \exp(-z_i). \end{aligned} \quad (10)$$

With assumption of independence between pixels, the likelihood function for the image \mathbf{y} is

$$\begin{aligned} L(\mathbf{y}) &= P(\mathbf{y}|\mathbf{z}) \\ &= \prod_i \sum_{n_i} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(y_i - n_i - g)^2}{2\sigma^2}\right) \frac{z_i^{n_i}}{n_i!} \exp(-z_i). \end{aligned} \quad (11)$$

Taking the negative log and dropping the terms independent of \mathbf{x} , we obtain the objective function

$$J_1(\mathbf{x}) = \sum_i \left(z_i - \log \left(\sum_{n_i} \frac{z_i^{n_i}}{n_i!} \exp\left(-\frac{(y_i - n_i - g)^2}{2\sigma^2}\right) \right) \right). \quad (12)$$

Then the MLE is obtained by minimizing $J_1(\mathbf{x})$ versus \mathbf{x} with the lower bound and the intensity conservation constraints, as explained in Section 1.3. The gradient of J_1 , for the pixel i , is

$$(\nabla J_1(x))_i = \sum_j \left(h_{ji} - h_{ji} \frac{r_j}{z_j} \right), \quad (13)$$

with

$$\begin{aligned} r_j &= \frac{p_j}{q_j}, \\ p_j &= \sum_{n_i} \exp\left(-\frac{(y_j - n_i - g)^2}{2\sigma^2}\right) \frac{n_i}{n_i!} z_j^{n_i}, \\ q_j &= \sum_{n_i} \exp\left(-\frac{(y_j - n_i - g)^2}{2\sigma^2}\right) \frac{1}{n_i!} z_j^{n_i}, \end{aligned} \quad (14)$$

or, in matrix notation, for the whole vector \mathbf{x} ,

$$\nabla J_1(\mathbf{x}) = H^T \text{diag}\left(\frac{1}{(H\mathbf{x} + \mathbf{d})}\right)(H\mathbf{x} + \mathbf{d} - \mathbf{r}). \quad (15)$$

As indicated in Section 1, to avoid instability of the MLE estimator, we propose to regularize the problem, that is, to add to the objective function $J_1(\mathbf{x})$ a penalty term $J_2(\mathbf{x})$.

2.2. Regularization

We will consider here two kinds of regularization functions: quadratic ones and entropic ones.

2.2.1. Quadratic regularization

We consider a regularization in the sense of Tikhonov [26]:

$$J_2(\mathbf{x}) = \frac{1}{2} \|\mathbf{x} - \bar{\mathbf{x}}\|^2. \quad (16)$$

It is a quadratic distance between x and an a priori solution $\bar{\mathbf{x}}$, called the “default image” [15, 27]. If there is no a priori information, the default solution can be chosen as a constant value p for a basic smoothness constraint; the choice of the value of p is not critical, it is frequently chosen equal to zero, however we choose to take $p = \sum_i x_i/N$, such that the total intensity conservation constraint is fulfilled also for the prior, see Section 3.1. In this case, the gradient of $J_2(\mathbf{x})$ is

$$\nabla J_2(\mathbf{x}) = \mathbf{x} - p. \quad (17)$$

Another common choice for J_2 is the Laplacian operator

$$J_2(\mathbf{x}) = \frac{1}{2} \|\mathbf{C}\mathbf{x}\|^2, \quad (18)$$

which can be expressed in the form (16)

$$J_2(\mathbf{x}) = \frac{1}{2} \|\mathbf{x} - \mathbf{A}\mathbf{x}\|^2, \quad (19)$$

where \mathbf{A} is deduced from the mask

$$\begin{pmatrix} 0 & \frac{1}{4} & 0 \\ \frac{1}{4} & 0 & \frac{1}{4} \\ 0 & \frac{1}{4} & 0 \end{pmatrix} \quad (20)$$

and $\mathbf{A}\mathbf{x}$ is the matrix form notation for the convolution operation between \mathbf{x} and this mask. The corresponding gradient

is then

$$\nabla J_2(\mathbf{x}) = (I + A^T A - A^T - A)\mathbf{x}. \quad (21)$$

In this case, the solution is implicitly biased towards a smoothed version of the solution. We note that A can be generalized to any form of lowpass filter.

2.2.2. Entropic distances

We use for entropic regularization the Csiszär directed divergence [28], a generalized form of the kullback-leibler (KL) distance between the solution and an a priori or *default* solution $\bar{\mathbf{x}}$:

$$J_2(\mathbf{x}) = \text{KL}(\mathbf{x}, \bar{\mathbf{x}}) = \sum_i x_i \log \left(\frac{x_i}{\bar{x}_i} + \bar{x}_i - x_i \right). \quad (22)$$

In the same way as for the quadratic distance, $\bar{\mathbf{x}}$ can be a constant p , with $p > 0$, or a smoothed version of the solution. We will consider here only the case of a constant *prior*, in order to compare the results with those of [9]. The corresponding gradient is

$$\nabla J_2(\mathbf{x}) = \log(\mathbf{x}) - \log(p). \quad (23)$$

Note that we can also consider the distance between $\bar{\mathbf{x}}$ and \mathbf{x} : $\text{KL}(\bar{\mathbf{x}}, \mathbf{x})$, which gives more or less the same practical results [29].

3. GENERAL ALGORITHMIC METHOD

3.1. SGM method

The regularized problem is to solve a problem of minimization under constraints, that is,

(i) minimize with respect to \mathbf{x} :

$$J(\mathbf{x}, \gamma) = J_1(\mathbf{x}) + \gamma J_2(\mathbf{x}); \quad (24)$$

(ii) with the constraints

- (1) lower bound, $x_i - m \geq 0$ for all i ;
- (2) energy conservation, $\sum_i x_i = \sum_i (y_i - d_i - g)/t_i$.

$J(\mathbf{x})$ consists of two terms, $J_1(\mathbf{x})$ that expresses the consistency with the data and $J_2(\mathbf{x})$, the term of regularization; γ is the regularization parameter, which tunes the relative weight between the two terms. We note that the considered functions $J_1(\mathbf{x})$ and $J_2(\mathbf{x})$ are convex.

We propose to devise algorithms deduced from the Kuhn-Tucker (KT) conditions in the general modified gradient form [10, 11, 30]:

$$x_i^{k+1} = C^k \left(x_i^k + \alpha_i^k f_i(\mathbf{x}^k, \gamma) (x_i^k - m) (-(\nabla J(\mathbf{x}^k, \gamma))_i) \right). \quad (25)$$

The initial estimate \mathbf{x}^0 is chosen as a constant value, such that all the constraints are fulfilled, C^k is a normalization factor for the total intensity conservation, subscript i is for

the pixel i , $\alpha_i^k > 0$ is the relaxation factor, k is the iteration index, $f(\mathbf{x})$ is a function having positive values when \mathbf{x} satisfies the constraints. To obtain “product form” algorithms, the split-gradient method (SGM) is used. It can be summarized as follows: the convex function $J(\mathbf{x}^k, \gamma)$ admits a finite unconstrained global minimum given by $\nabla J(\mathbf{x}^k, \gamma) = 0$, then we can write

$$-\nabla J(\mathbf{x}^k, \gamma) = U(\mathbf{x}^k, \gamma) - V(\mathbf{x}^k, \gamma), \quad (26)$$

where $U(\mathbf{x}^k, \gamma)$ and $V(\mathbf{x}^k, \gamma)$ are two positive functions for all $\mathbf{x}^k \geq m$. From (24), the total gradient can be decomposed as

$$-\nabla J(\mathbf{x}^k, \gamma) = -\nabla J_1(\mathbf{x}^k) - \gamma \nabla J_2(\mathbf{x}^k). \quad (27)$$

Splitting $-\nabla J_1(\mathbf{x}^k)$ and $-\nabla J_2(\mathbf{x}^k)$ as in (26), we have

$$\begin{aligned} -\nabla J(\mathbf{x}^k, \gamma) &= U_1(\mathbf{x}^k) - V_1(\mathbf{x}^k) + \gamma(U_2(\mathbf{x}^k) - V_2(\mathbf{x}^k)), \\ U(\mathbf{x}^k, \gamma) &= U_1(\mathbf{x}^k) + \gamma U_2(\mathbf{x}^k), \\ V(\mathbf{x}^k, \gamma) &= V_1(\mathbf{x}^k) + \gamma V_2(\mathbf{x}^k). \end{aligned} \quad (28)$$

Taking

$$f_i(\mathbf{x}^k, \gamma) = \frac{1}{V_i(\mathbf{x}^k, \gamma)} > 0, \quad (29)$$

(25) becomes

$$x_i^{k+1} = C^k \left(x_i^k + \alpha_i^k \frac{(x_i^k - m)}{V_i(\mathbf{x}^k, \gamma)} (U_i(\mathbf{x}^k, \gamma) - V_i(\mathbf{x}^k, \gamma)) \right). \quad (30)$$

The maximum stepsize that ensures $x_i^{k+1} - m \geq 0$, for all i , for all k is given by

$$\alpha_m^k = \min_{i \in \mathcal{C}} \left(\frac{1}{1 - U_i(\mathbf{x}^k, \gamma)/V_i(\mathbf{x}^k, \gamma)} \right), \quad (31)$$

where \mathcal{C} is the set of index i such that $(\nabla J(\mathbf{x}^k, \gamma))_i > 0$ and $x_i^k > m$; clearly $\alpha_m^k > 1$, then for $\alpha^k = 1$, the constraint is always fulfilled. The optimal step size α_c^k independent of i ensuring convergence must be computed in the range $]0, \alpha_m^k]$ (or $]0, \alpha_m^k[$ if a strict inequality constraint is required) by a line search procedure (see, e.g., [31, 32, 33, 34]) with the descent direction

$$\boldsymbol{\rho}^k = \text{diag} \left(\frac{(x_i^k - m)}{V_i(\mathbf{x}^k, \gamma)} \right) (U(\mathbf{x}^k, \gamma) - V(\mathbf{x}^k, \gamma)). \quad (32)$$

This direction is no more the negative gradient but it remains a descent direction for $J(\mathbf{x})$. With a unit step size, the

TABLE 1: Functions U and V following the kind of regularization.

Function	Likelihood	Quadratic, $\bar{\mathbf{x}} = \mathbf{p}$	Quadratic, $\bar{\mathbf{x}} = A\mathbf{x}$	Entropic, $\bar{\mathbf{x}} = \mathbf{p}$
U	$U_1 = H^T \text{diag} \left(\frac{1}{(H\mathbf{x} + \mathbf{d})_i} \right) \mathbf{r}$	$U_2 = \gamma \mathbf{p}$	$U_2 = \gamma (A^T + A) \mathbf{x}$	$U_2 = -\gamma \log \left(\frac{\mathbf{x}}{\sum_i x_i} \right)$
V	$V_1 = \mathbf{a}$, see (5)	$V_2 = \gamma \mathbf{x}$	$V_2 = \gamma (I + A^T A) \mathbf{x}$	$V_2 = -\gamma \log \left(\frac{\mathbf{p}}{\sum_i p_i} \right)$

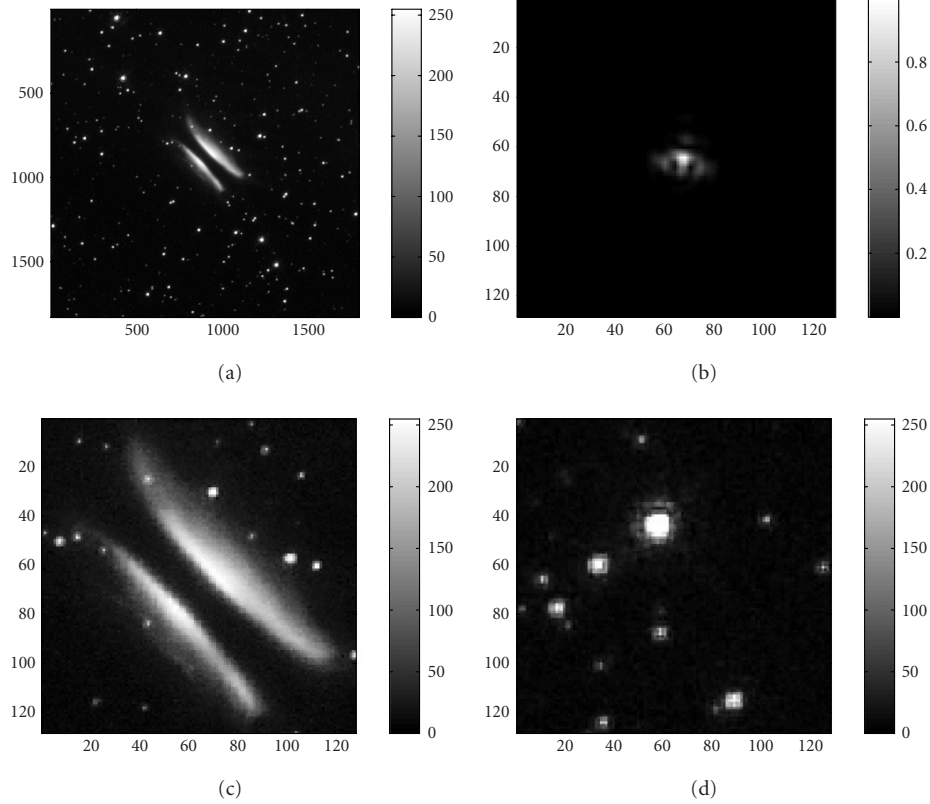


FIGURE 1: (a) HST picture, (b) PSF, (c) star, (d) point sources.

algorithm (30) takes an interesting simple form:

$$x_i^{k+1} = C^k \left(m + (x_i^k - m) \frac{U_i(\mathbf{x}^k, \gamma)}{V_i(\mathbf{x}^k, \gamma)} \right). \quad (33)$$

The normalization, that is, the computation of C^k , is performed following [10]. The convergence of (33) is not theoretically ensured but has always been observed in practice. To ensure the theoretical convergence of (30) without a dramatic increase of the computational cost, economic line search using the Armijo rule [31] or the Goldstein rule [35] can be used to compute α_c^k , as mentioned in [32].

3.2. Application to the Poisson Gaussian model

Table 1 gives the expressions of the functions U and V for the likelihood (12) and for several penalty functions, (16), (19), and (22). Algorithms can be obtained in a closed form by replacing these expressions in (30) or (33) using the decomposition (28).

4. NUMERICAL ILLUSTRATIONS

4.1. Object

The proposed algorithm has been illustrated on a picture taken from the hubble space telescope (HST) site, <http://hubblesite.org/gallery/>. It is a sun-like star nearing the end of its life, Figure 1a. Two 128×128 subpictures have been extracted from the HST image, one centered on the main structure, Figure 1c, the other containing bright point sources, Figure 1d, where the same constant background has been added to both pictures. These two parts have been selected in order to study the effectiveness of various penalty functions depending on the kind of pictures: more or less diffuse.

4.2. PSF

The data has been blurred with the normalized space-invariant PSF, Figure 1b. It is a realistic representation of the PSF of a ground-based telescope including the effects of

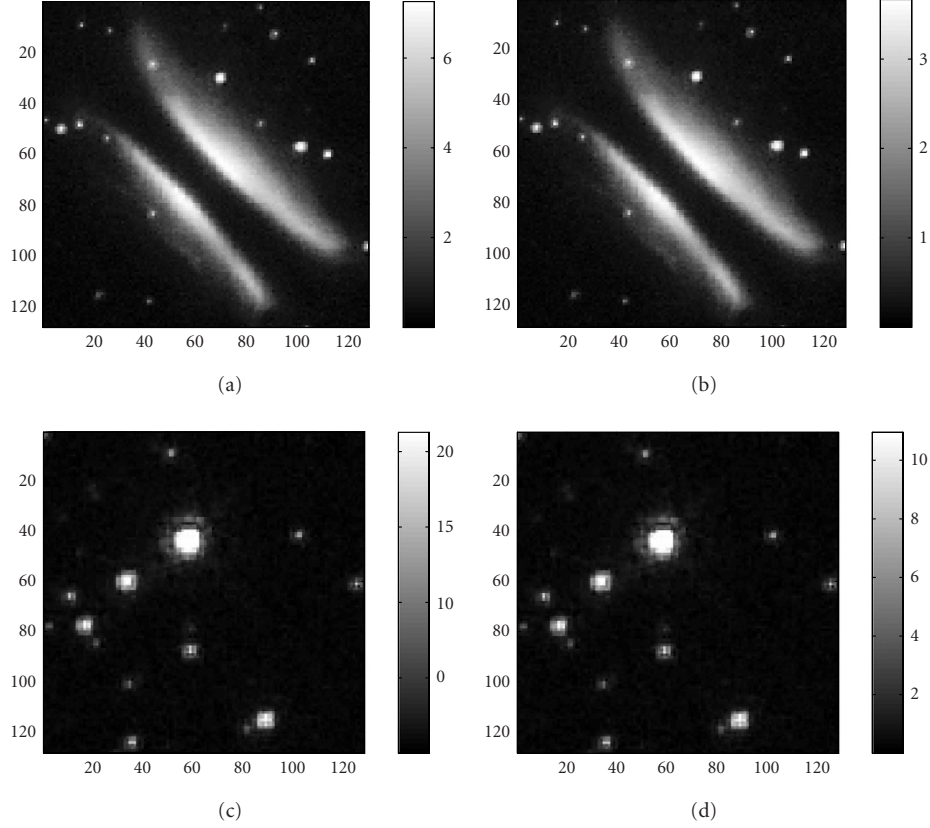


FIGURE 2: Normalized star: (a) 20 000 photons and (b) 10 000 photons. Normalized point sources: (c) 20 000 photons and (d) 10 000 photons.

TABLE 2: Image characteristics.

Set		A	B
Photons		20 000	10 000
Gaussian noise	g	0.01	0.1
	σ^2	0.1	1

the atmospheric turbulence. The telescope aperture $P(r)$ is simply given by an array of points “1” inside a circle and “0” outside, the wavefront error $\delta(r)$ is an array of random numbers smoothed by a lowpass filter. The quantity $P(r) \exp(2i\pi\delta(r)/\lambda)$ represents the telescope aperture with the phase error; for the simulation, the peak-to-peak phase variation is small (less than π) and may correspond to typical telescope aberrations, see [10] for more details. The main interest of such a PSF simulation is that the optical transfer function is a lowpass filter, limited in spatial frequencies to the extent of the aperture autocorrelation function; the blurred image is then strictly band limited corresponding to a realistic situation.

4.3. Noise

A Poisson transform is applied on the blurred image and finally a Gaussian noise is added. Two sets of noise parameters have been chosen, the first one with 20000 photons in the im-

age, $m = 0.01$, $\sigma^2 = 0.1$ (set A) corresponding to a fairly noisy image, shown in Figures 6a and 10a; the other one with 10000 photons, $m = 0.1$, $\sigma^2 = 1$ (set B) represented in Figures 8a and 12a corresponding to a highly noisy image, see Table 2 for a summary of the characteristics.

Figure 2 shows the objects normalized to the intensity of the noisy images, allowing a quantitative comparison of the results (normalized to the same quantity) with the objects.

4.4. Implementation

The HST images have been reconstructed using the algorithm (30), the behavior of the algorithm for the two pictures is studied as a function of the regularization function and of the “level” of noise in the raw images. The summation over the number of photons n_i in the expressions of p_j and q_j (14) has been reduced to $\max(n_i) + g + 3\sigma$ corresponding to significant values for the exponential term.

To stop the iterative procedure before noise amplification and/or to check the quality of the restoration process, we use a criterion based on the Euclidean distance $\epsilon(k, \gamma)$ between the true object \mathbf{x}^* and the reconstructed one, computed as a function of k and γ :

$$\epsilon(k, \gamma) = \frac{\|\mathbf{x}_\gamma^k - \mathbf{x}^*\|^2}{\|\mathbf{x}^*\|^2}. \quad (34)$$

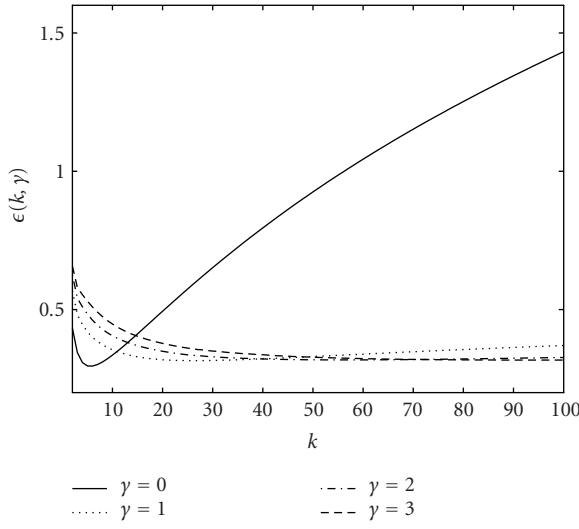


FIGURE 3: Reconstruction error for the quadratic penalty with variable *prior*; extended object, set A.

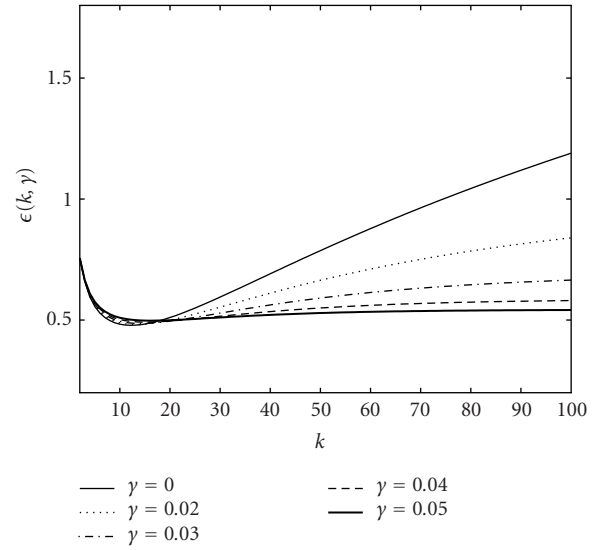


FIGURE 5: Reconstruction error for the entropic penalty with constant *prior*; bright points object, set B.

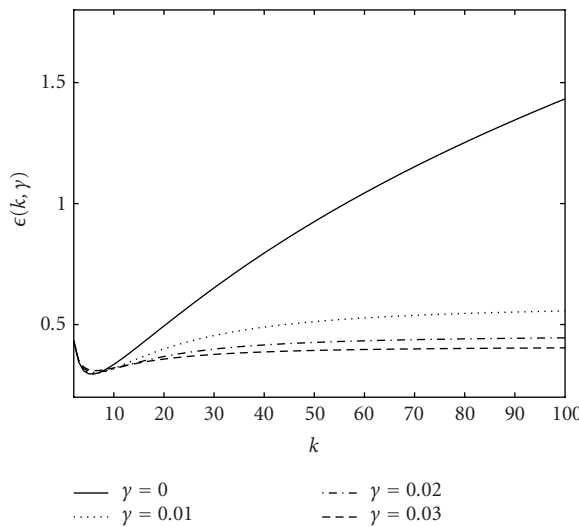


FIGURE 4: Reconstruction error for the quadratic penalty with constant *prior*; extended object, set A.

Such a comparison cannot be made for a real case since the true object is not known. However, it allows a good characterization of the behavior and performance of the algorithm for simulated data, and it is useful to choose the regularization parameter.

Indeed, the search for the correct regularization parameter remains an open problem. Some methods such as “L curves” or “generalized cross validation” have been proposed, typically for pure Gaussian noise [36, 37, 38]. In fact, in the regularization scheme, three elements must be chosen: the

penalty function, the prior, and the regularization parameter. The choice of these points depends on the properties of the image, on the amount of noise, and on the expected objective. The regularization factor is heuristically chosen here by selecting the one corresponding to an asymptotic minimum value of $\epsilon(k, \gamma)$, see Figures 3, 4, and 5 for the behavior of $\epsilon(k, \gamma)$. The best reconstructed images are given following the regularization function in Figures 6, 7, 8 and 9 for the extended object and in Figures 10, 11, 12 and 13 for the bright point sources.

Note that statistical analysis including a lot of noise realizations should be carried out to get significant quantitative results, implying explosive computational cost; consequently only qualitative remarks and comparative conclusions in regard to the penalty functions are given.

4.5. Results for extended objects

4.5.1. Set A

Results obtained for the extended object are shown in Figures 6 and 7 for the set A and must be quantitatively compared to Figure 2a. The restoration is first performed using the nonpenalized algorithm ($J(\mathbf{x}) = J_1(\mathbf{x})$), in this case due to the semiconvergence of the algorithm, iterations must be stopped before the noise increases. The result corresponding to the minimum of $\epsilon(k, 0)$ is given in Figure 6b.

A constant prior is then used in conjunction with a quadratic penalty function (16) or an entropic penalty function (22); the corresponding behaviors of the restoration error for several values of γ are given in Figures 4 and 5. The minimum reconstruction error is always reached for $\gamma = 0$; for γ increasing the curve tends to be monotonic and the corresponding error increases. The best reconstructed objects are shown in Figures 6c and 6d, respectively, for the quadratic

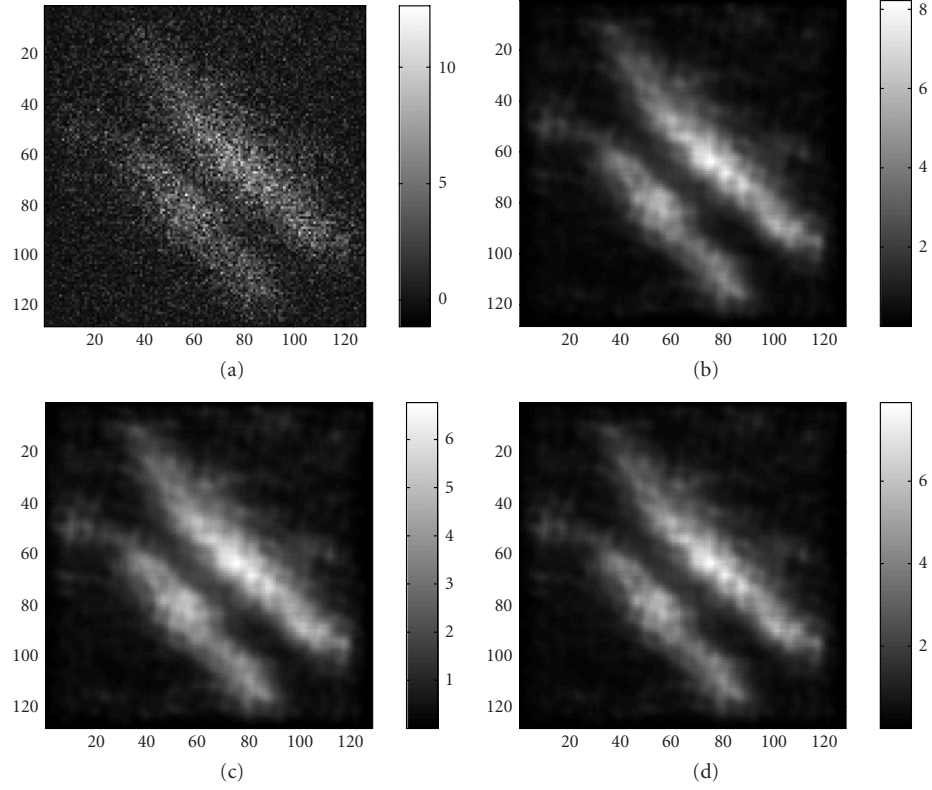


FIGURE 6: (a) Raw picture, 20000 photons, $m = 0.01$, $\sigma^2 = 0.1$. (b) $\gamma = 0$, $k = 5$, $E_{\min} = 0.2956$. (c) Quadratic penalty with constant *prior*, $\gamma = 0.03$, $k = 5$, $E_{\min} = 0.3099$. (d) Entropic penalty with constant *prior*, $\gamma = 0.03$, $k = 6$, $E_{\min} = 0.2965$.

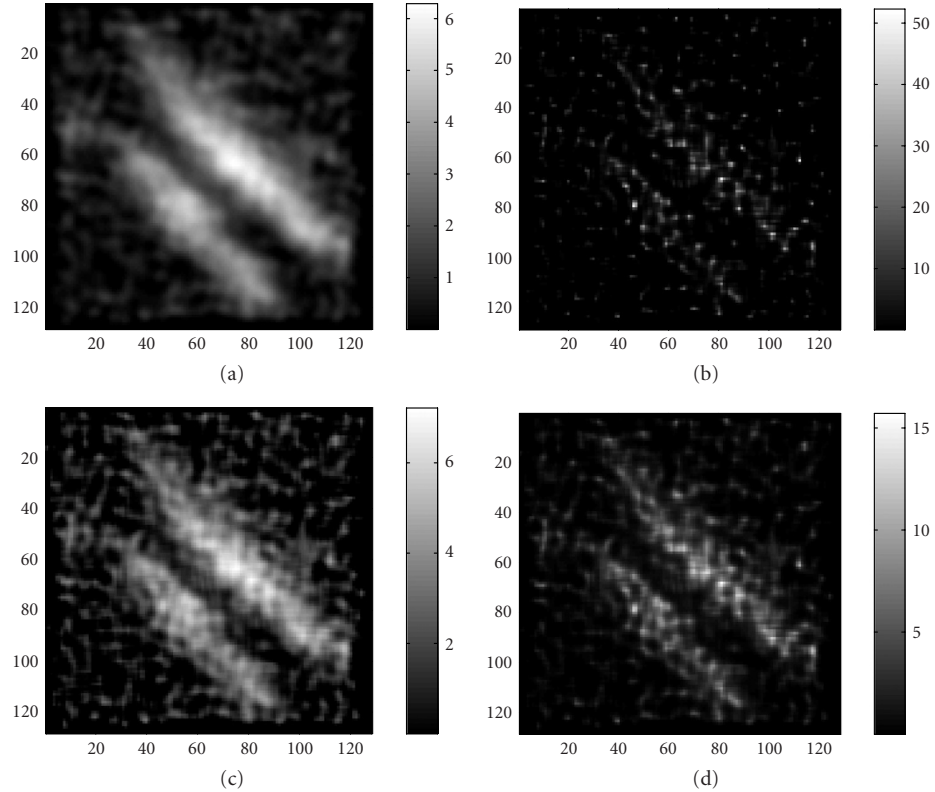


FIGURE 7: (a) Quadratic penalty with variable *prior*, $\gamma = 3$, $k = 100$, $E_{\min} = 0.3315$. (b) $\gamma = 0$, $k = 100$, $E = 1.4318$. (c) Quadratic penalty with constant *prior*, $\gamma = 0.03$, $k = 100$, $E = 0.404$. (d) Entropic penalty with constant *prior*, $\gamma = 0.03$, $k = 100$, $E = 0.5286$.

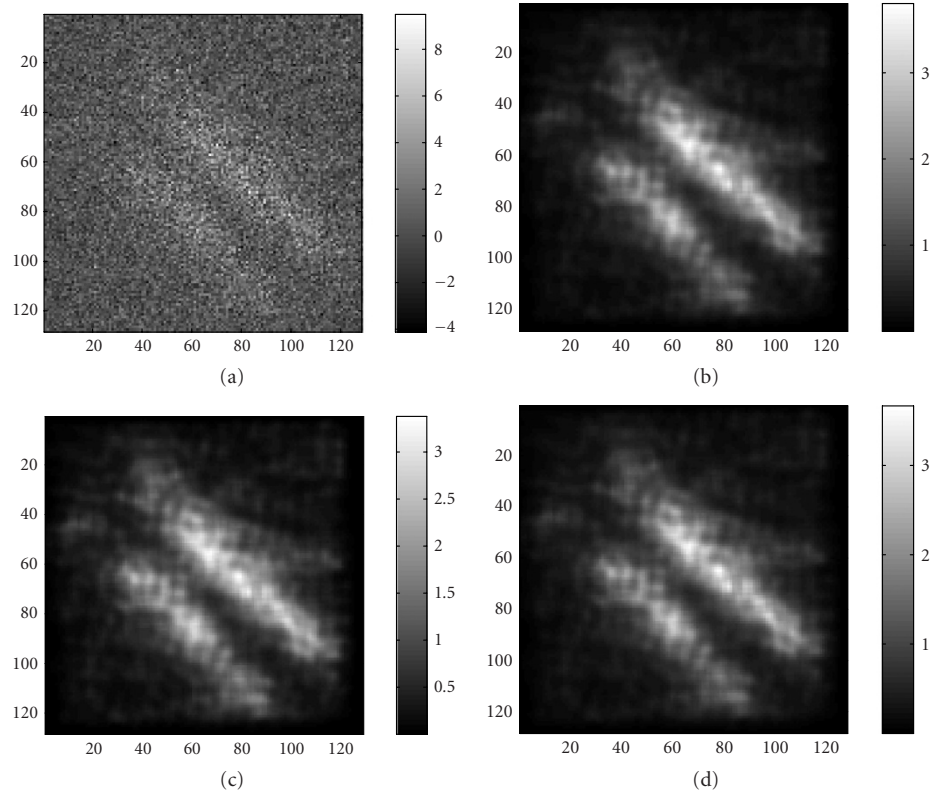


FIGURE 8: (a) Raw picture, 10 000 photons, $m = 0.1$, $\sigma^2 = 1$. (b) $\gamma = 0$, $k = 6$, $E_{\min} = 0.3490$. (c) Quadratic penalty with constant *prior*, $\gamma = 0.04$, $k = 7$, $E_{\min} = 0.3608$. (d) Entropic penalty with constant *prior*, $\gamma = 0.04$, $k = 8$, $E_{\min} = 0.3544$.

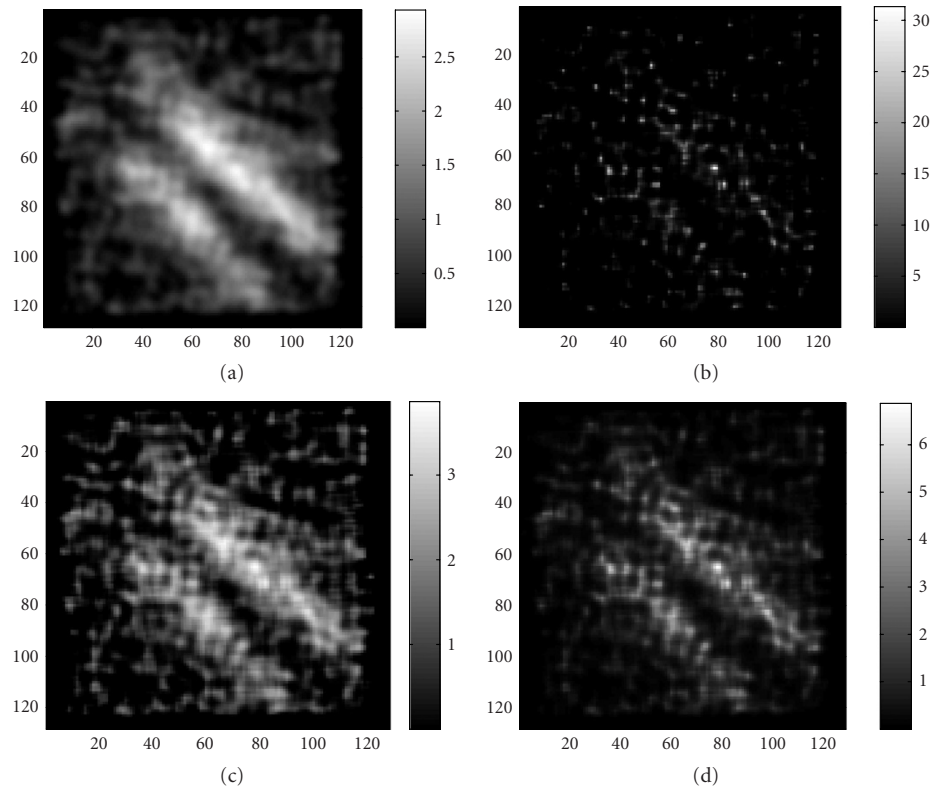


FIGURE 9: (a) Quadratic penalty with variable *prior*, $\gamma = 5$, $k = 100$, $E = 0.3857$. (b) $\gamma = 0$, $k = 100$, $E = 1.5989$. (c) Quadratic penalty with constant *prior*, $\gamma = 0.04$, $k = 100$, $E = 0.4977$. (d) Entropic penalty with constant *prior*, $\gamma = 0.04$, $k = 100$, $E = 0.5006$.

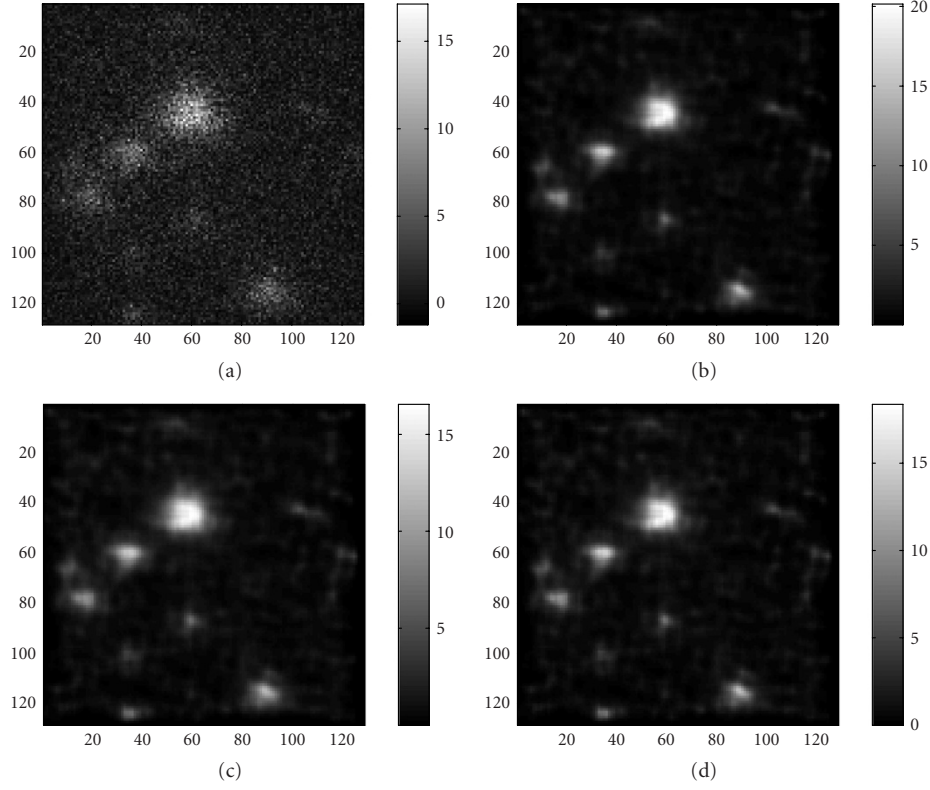


FIGURE 10: (a) Raw picture, 20 000 photons, $m = 0.01$, $\sigma^2 = 0.1$. (b) $\gamma = 0$, $k = 9$, $E_{\min} = 0.4147$. (c) Quadratic penalty with constant *prior*, $\gamma = 0.005$, $k = 8$, $E_{\min} = 0.4469$. (d) Entropic penalty with constant *prior*, $\gamma = 0.03$, $k = 13$, $E_{\min} = 0.4277$.

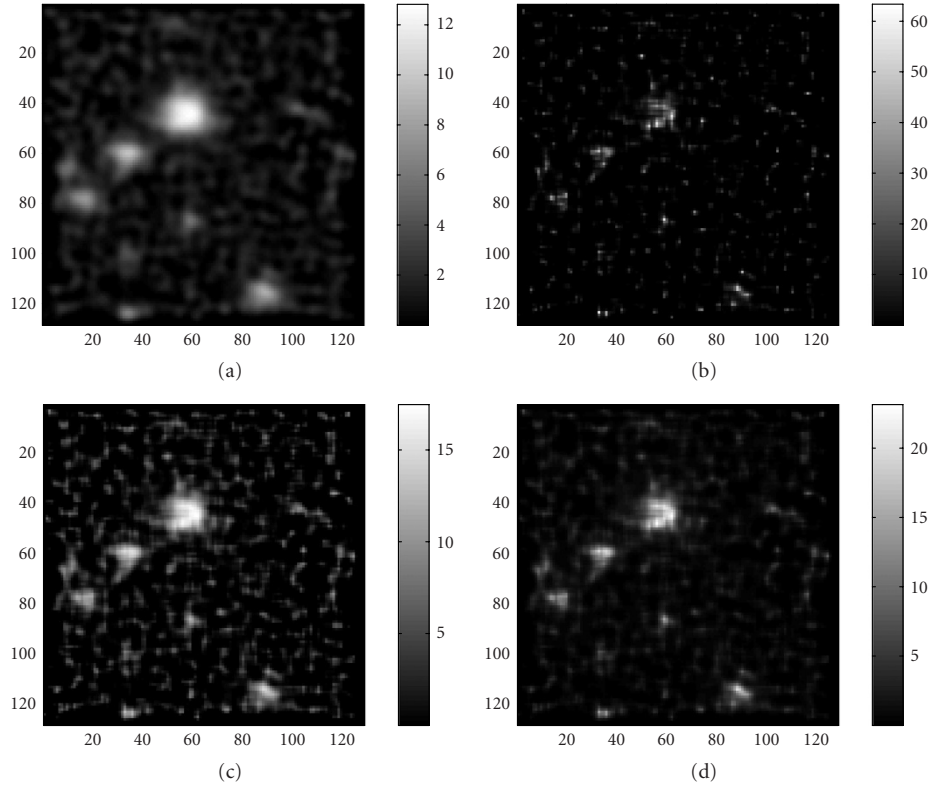


FIGURE 11: (a) Quadratic penalty with variable *prior*, $\gamma = 0.8$, $k = 100$, $E_{\min} = 0.5435$. (b) $\gamma = 0$, $k = 100$, $E = 1.209$. (c) Quadratic penalty with constant *prior*, $\gamma = 0.005$, $k = 100$, $E = 0.6622$. (d) Entropic penalty with constant *prior*, $\gamma = 0.03$, $k = 100$, $E = 0.5449$.

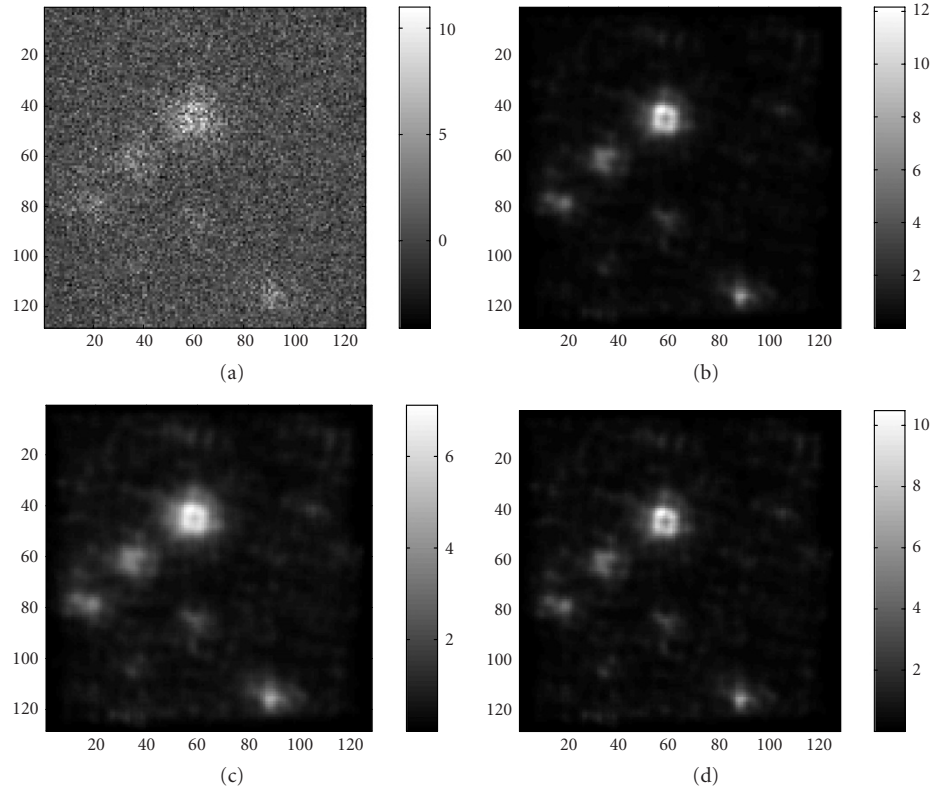


FIGURE 12: (a) Raw picture, 10 000 photons, $m = 0.1$, $\sigma^2 = 1$. (b) $\gamma = 0$, $k = 9$, $E_{\min} = 0.47847$. (c) Quadratic penalty with constant *prior*, $\gamma = 0.03$, $k = 8$, $E_{\min} = 0.5495$. (d) Entropic regularization, $\gamma = 0.05$, $k = 9$, $E_{\min} = 0.5043$.

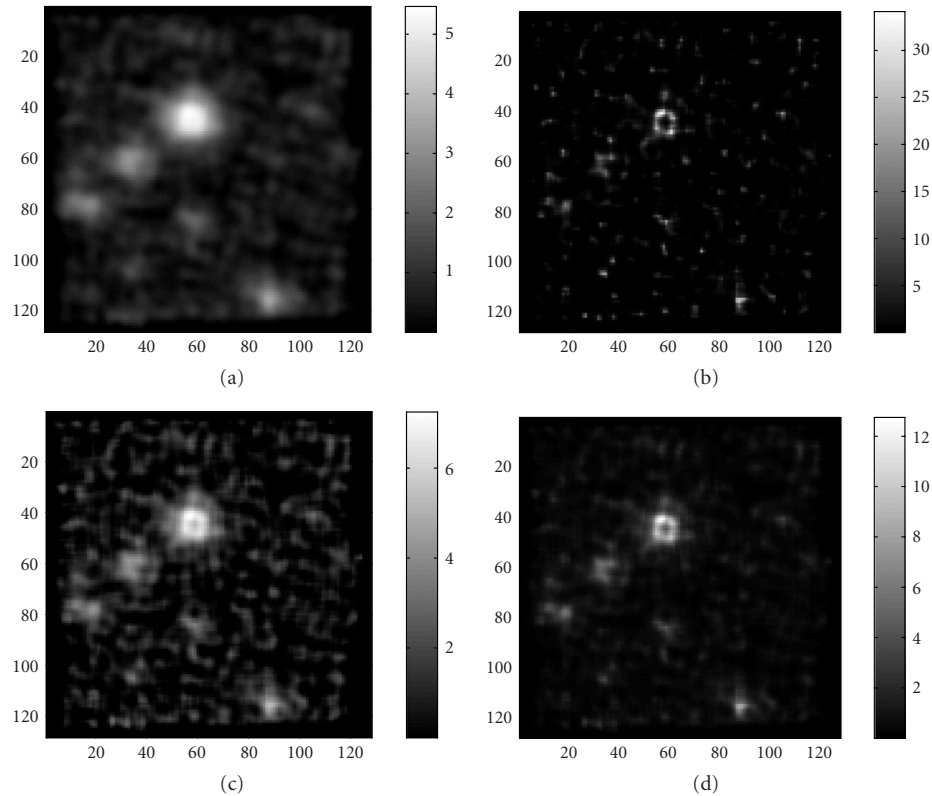


FIGURE 13: (a) Quadratic penalty with variable *prior*, $\gamma = 4$, $k = 100$, $E_{\min} = 0.6131$. (b) $\gamma = 0$, $k = 100$, $E = 1.1896$. (c) Quadratic penalty with constant *prior*, $\gamma = 0.03$, $k = 100$, $E = 0.6432$. (d) Entropic regularization, $\gamma = 0.05$, $k = 100$, $E_{\min} = 0.5256$.

penalty and the entropic one. Results are very similar as confirmed by values of the reconstruction error.

Figures 7b, 7c, and 7d give the asymptotic reconstructed images ($k = 100$) for respectively $\gamma = 0$, a quadratic regularization, and an entropic regularization. Figure 7b exhibits the typical behavior of the MLE when the iteration number increases due to the noise amplification, see the color table in Figure 7. The effect of the regularization clearly appears in Figures 7c and 7d and especially in the case of a quadratic penalty, for asymptotic iteration number. The result obtained from the entropic penalty is unsatisfactory and is not improved by an increase of the regularization factor. It has been observed in [10] that the speckled effect obtained is typical of the constant *prior*. In order to confirm this observation, the result obtained with a quadratic penalty using the Laplacian operator (18) and (19) is shown in Figure 7a. The corresponding reconstruction error is given in Figure 3; for γ greater than 3, $\epsilon(k, \gamma)$ decreases monotonically and the value reached by the error at high iteration number ($k = 100$) is very close to the minimum error obtained for $\gamma = 0$ and $k = 5$. The reconstructed image is clearly better than those obtained with the constant *prior*.

4.5.2. Set B

The raw data with the set of parameters B is clearly much noisier, Figure 8a. Results obtained must be compared to Figure 2b. They are very similar to those obtained with the less noisy image, the number of iterations and the regularization parameter corresponding to the best reconstructed image are slightly larger, Figures 8b, 8c, and 8d. The error $\epsilon(k, \gamma)$ presents the same behavior with higher values. The Laplacian operator gives better results than the other penalty functions for 10 000 photons as well, Figure 9a. Asymptotic results are given in Figure 9.

4.6. Results for bright points object

4.6.1. Set A

The same algorithms have been applied on the bright points source for the set A, Figure 10a. The results of the reconstruction must be compared to Figure 2c.

Note that the photons are more concentrated into small circular parts. The first result is for the nonpenalized algorithm and the best result is shown in Figure 10b. Best results in the sense of the minimization of $\epsilon(k, \gamma)$ are given in Figures 10c and 10d, respectively, for the quadratic and the entropic regularizations with a constant prior. Regularization parameters and the optimal iteration number are significantly different but the reconstruction error is approximately the same, this is a first difference in regard to the extended image.

For a large iteration number, results are given in Figures 11b, 11c, and 11d. As expected, looking at the error curves, we see that the nonpenalized algorithm gives a bad reconstruction and corresponds to a large error while results for the penalized algorithms are correct with a stabilization of the reconstruction error.

The Laplacian regularization gives a poorer result with a spreading effect on the point sources.

4.6.2. Set B

Data with the noise of the set B is in Figure 12a. The contrast between the background and the point sources is reduced. The nonpenalized algorithm gives the best restoration, shown in Figure 12b. The best results in the sense of the minimization of $\epsilon(k, \gamma)$ are given in Figures 12c and 12d, respectively, for the quadratic and the entropic penalties with constant prior. All the results are to compare to the objects of Figure 2d. The regularization parameter is more difficult to find with respect to the extended image and changes from one function to another.

For a large iteration number, results are given in Figures 13b, 13c, and 13d. As expected, looking at the error curves, we see that the nonpenalized algorithm leads to a bad reconstruction and corresponds to a large error while results for the penalized algorithms are correct with a stabilization of the reconstruction error.

As for the set A, the Laplacian regularization gives a less satisfactory result since the regularization has spread the point sources.

In conclusion, for point-like objects, the results obtained with penalized algorithms are good for the Tikhonov regularization whatever the *prior* image: constant or a lowpass version of the solution. Entropy penalty function leads to an excessive resolution, even if the reconstruction error is low.

5. DISCUSSION

The method proposed here can be applied to all the inverse problems described by a first-kind Fredholm integral. Two points must be outlined. First is the exact kind of noise: Gaussian, Poisson, or, as considered here, Poisson plus Gaussian. The second point deals with the properties of the kernel of the integral equation, space invariant or not, separable or not.

Concerning the first point, the case of pure Gaussian noise corresponds to high intensity imagery as in the infrared case, while the pure Poisson noise corresponds to low light level with perfect detectors. These two cases have been analyzed in previous papers [10, 11]. The case of Poisson data with additive Gaussian noise, considered here, corresponds to the more realistic situation of low light level imagery with imperfect CCD detectors.

Concerning the second point, the matrix notation used in this paper is general and can be applied whatever the properties of the kernel are. But in the case of imagery, the matrix form implies that the object and the image are lexicographically ordered vectors, then for $n \times n$ objects and images, the corresponding matrix W will be $n^2 \times n^2$, (recall that $H\mathbf{x} = T(W\mathbf{x})$). If the kernel w does not exhibit any specific property, the computations must be performed in matrix form with a heavy computational cost. Fortunately, for space-invariant kernels (convolution), the matrix W has a Toeplitz or block-Toeplitz structure, simplifying computations. Indeed, the matrix expression $W\mathbf{x}$ corresponds to the convolution operation $w(r, s) \otimes x(r, s)$ which is performed by means of fast Fourier transform. The operation (T) implies only point-to-point products. Then the proposed algorithm

is of general use whatever the exact properties of the kernel $w(r, s)$.

6. CONCLUSION

In this paper, we consider mainly the problem of restoring astronomical images acquired with CCD cameras. The nonuniform sensitivity of the detector elements (flat field) is taken into account and the various noise effects such as the statistical Poisson effect during the image formation process and the additive Gaussian read-out noise are taken into account. We first show that applying the split gradient method (SGM), maximum likelihood algorithms can be obtained in a rigorous way; the relaxed convergent form of such algorithm is exhibited and it has been demonstrated that the EM algorithms are nonrelaxed versions of the proposed algorithm. The proposed method can be systematically applied in a rigorous way (ensuring convergence and positivity of the solution) to MAP estimation for various convex penalty functions. In this paper three penalty functions have been developed: the quadratic one with either constant or smooth prior and the entropic with constant prior. Previous attempts of regularization for this model using sieves has been proposed by Snyder et al. [3, 4, 6], however the penalty method proposed outperforms "sieves" [19]. Another approach, proposed by Llacer and Nuñez [7, 8], uses the penalized ML approach with an entropy penalty function but here the convergence and the positivity constraint are not always ensured [25], contrary to our method. The proposed algorithm has been checked on typical astrophysical images with antagonistic characteristics: diffuse or bright points objects. We have shown that a Laplacian operator gives satisfactory results for extended objects and an over-smoothed solution for bright points objects. The reverse seems to occur when the prior is a constant.

REFERENCES

- [1] M. E. Daube-Witherspoon and G. Muehllehner, "An iterative image space reconstruction algorithm suitable for volume ECT," *IEEE Trans. Med. Imag.*, vol. 5, no. 2, pp. 61–66, 1986.
- [2] L. B. Lucy, "An iterative technique for the rectification of observed distributions," *Astronomic Journal*, vol. 79, pp. 745–754, 1974.
- [3] D. L. Snyder and M. I. Miller, "The use of sieves to stabilize the images produced with the EM algorithm for emission tomography," *IEEE Trans. Nucl. Sci.*, vol. 32, no. 5, pp. 3864–3871, 1985.
- [4] D. L. Snyder, "Modification of the Richardson-Lucy iteration for restoring hubble space telescope imagery," in *Restoration of Hubble Space Telescope Images*, R. L. White and R. J. Allen, Eds., The Space Telescope Science Institute, Baltimore, Md, USA, 1990.
- [5] D. L. Snyder, A. M. Hammoud, and R. L. White, "Image recovery from data acquired with charge-coupled-device camera," *Journal of the Optical Society of America A*, vol. 10, no. 5, pp. 1014–1023, 1993.
- [6] D. L. Snyder, C. W. Helstrom, A. D. Lanterman, M. Faisal, and R. L. White, "Compensation for readout noise in CCD images," *Journal of Optical Society of America A*, vol. 12, no. 2, pp. 272–283, 1995.
- [7] J. Nuñez and J. Llacer, "Bayesian image reconstruction with space-variant noise suppression," *Astronomy and Astrophysics Supplement Series*, vol. 131, no. 1, pp. 167–180, 1998.
- [8] J. Nuñez and J. Llacer, "A general Bayesian image reconstruction algorithm with entropy prior: preliminary application to HST data," *Publication of the Astronomical Society of the Pacific*, vol. 105, no. 692, pp. 1192–1208, 1993.
- [9] J. Llacer and J. Nuñez, "Iterative maximum likelihood and bayesian algorithms for image reconstruction in astronomy," in *The Restoration of Hubble Space Telescope Images*, R. L. White and R. J. Allen, Eds., pp. 62–69, The Space Telescope Science Institute, Baltimore, Md, USA, 1990.
- [10] H. Lantéri, M. Roche, and C. Aime, "Penalized maximum likelihood image restoration with positivity constraints: multiplicative algorithms," *Inverse problems*, vol. 18, no. 5, pp. 1397–1419, 2002.
- [11] H. Lantéri, M. Roche, O. Cuevas, and C. Aime, "A general method to devise maximum-likelihood signal restoration multiplicative algorithms with non-negativity constraints," *Signal Processing*, vol. 81, no. 5, pp. 945–974, 2001.
- [12] A. D. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society—Series B*, vol. 39, no. 1, pp. 1–38, 1977.
- [13] Y. Vardi, L. A. Shepp, and L. Kaufmann, "A statistical model for positron emission tomography," *Journal of the American Statistical Association*, vol. 80, no. 389, pp. 8–37, 1985.
- [14] H. Lantéri, R. Soummer, and C. Aime, "Comparison between ISRA and RLA algorithms. Use of a Wiener filter based stopping criterion," *Astronomy and Astrophysics Supplement Series*, vol. 140, no. 2, pp. 235–246, 1999.
- [15] L. B. Lucy, "Optimum strategy for inverse problems in statistical astronomy," *Astronomy and Astrophysics*, vol. 289, no. 3, pp. 983–994, 1994.
- [16] J. A. Fessler and A. O. Hero, "Penalized maximum-likelihood image reconstruction using space-alternating generalized EM algorithms," *IEEE Trans. Image Processing*, vol. 4, no. 10, pp. 1417–1429, 1995.
- [17] U. Grenander, *Abstract Inference*, John Wiley & Sons, New York, NY, USA, 1981.
- [18] D. L. Snyder, T. J. Schultz, and J. A. O'Sullivan, "Deblurring subject to non negativity constraints," *IEEE Trans. Signal Processing*, vol. 40, no. 5, pp. 1143–1150, 1992.
- [19] C. S. Butler and M. I. Miller, "Maximum a posteriori estimation for SPECT using regularization techniques on massively parallel computers," *IEEE Trans. Med. Imag.*, vol. 12, no. 1, pp. 84–89, 1993.
- [20] M. Bertero, P. Boccaci, and F. Maggio, "Regularization methods in image restoration: an application to HST images," *International Journal of Imaging Systems and Technology*, vol. 6, pp. 376–386, 1995.
- [21] M. Bertero, "Linear inverses and ill-posed problems," *Advances in Electronics and Electron Physics*, vol. 75, pp. 1–121, 1989.
- [22] M. Bertero and P. Boccaci, *Introduction to Inverse Problems in Imaging*, I.O.P Publishing, London, UK, 1998.
- [23] G. Demoment, "Image reconstruction and restoration: Overview of common estimation structures and problems," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 37, no. 12, pp. 2024–2036, 1989.
- [24] D. M. Titterton, "General structure of regularization procedures in image reconstruction," *Astronomy and Astrophysics*, vol. 144, no. 2, pp. 381–387, 1985.
- [25] C. H. Wu and J. M. M. Anderson, "Novel deblurring algorithms for images captured with CCD cameras," *Journal of Optical Society of America A*, vol. 14, no. 7, pp. 1421–1430, 1997.

- [26] A. Tikhonov and V. Arsenin, *méthodes de résolution des problèmes mal posés*, Mir, Moscou, Russia, 1976.
- [27] K. Horne, "Images of accretion discs—the eclipse mapping method," *Monthly Notices of the Royal Astronomical Society*, vol. 213, pp. 129–141, 1985.
- [28] I. Csiszár, "Why least squares and maximum entropy? an axiomatic approach to inference for linear inverse problems," *The Annals of Statistics*, vol. 19, no. 4, pp. 2032–2066, 1991.
- [29] C. L. Byrne, "Iterative image reconstruction algorithms based on cross-entropy minimization," *IEEE Trans. Image Processing*, vol. 2, no. 1, pp. 96–103, 1993.
- [30] H. Lantéri, M. Roche, P. Gaucherel, and C. Aime, "Ringing reduction in image restoration algorithms using a constraint on the inferior bound of the solution," *Signal Processing*, vol. 82, no. 10, pp. 1481–1504, 2002.
- [31] L. Armijo, "Minimization of functions having continuous derivatives," *Pacific Journal of Mathematics*, vol. 16, pp. 1–3, 1966.
- [32] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, Belmont, Mass, USA, 1995.
- [33] M. Minoux, "Programmation mathématique—théorie et algorithmes," in *Collection technique et scientifique des télécommunications*, vol. 1, pp. XXXI–294, Paris, France, Dunod edition, 1983.
- [34] D. G. Luenberger, *Introduction to Linear and Non Linear Programming*, Addison Wesley, Reading, Mass, USA, 1973.
- [35] A. A. Goldstein, *Constructive Real Analysis*, Harper, New York, NY, USA, 1967.
- [36] A. Thompson, J. C. Brown, J. W. Kay, and D. M. Titterton, "A study of methods of choosing the smoothing parameter in image restoration by regularization," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, no. 4, pp. 326–339, 1991.
- [37] P. Hansen, "Analysis of discrete ill-posed problems by means of the L -curve," *SIAM Review*, vol. 34, no. 4, pp. 561–580, 1992.
- [38] G. H. Golub, M. Heath, and G. Wahba, "Generalized cross-validation as a method for choosing a good ridge parameter," *Technometrics*, vol. 21, no. 2, pp. 215–223, 1979.

H. Lantéri received the Ph.D. degree in electronics in 1968 and the Doctorat d'état in 1978. He is a Professor of electrical engineering at the University of Nice Sophia Antipolis and is currently with the Laboratoire Universitaire d'Astrophysique de Nice (LUAN). His research interests include inverse problems and information theory with application in astrophysics, image restoration, and image processing.



C. Theys was born in Paris, France, in 1967. She received the Ph.D. degree in electronic engineering in 1993 from the University of Nice Sophia Antipolis (UNSA), France. She is currently an Assistant Professor of electrical engineering at the Institut Universitaire de Technologie de Nice Côte d'Azur, France. From 1995 to 2003, she was working on digital signal processing for detection and estimation with the Laboratoire d'Informatique Signaux et Systèmes de Sophia (I3S). Since 2003, she is a member of the Laboratoire Universitaire d'Astrophysique de Nice (LUAN) and her research interests are inverse problems with application to astrophysical.

