

## Image deblurring with Poisson data: from cells to galaxies

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2009 Inverse Problems 25 123006

(<http://iopscience.iop.org/0266-5611/25/12/123006>)

[The Table of Contents](#) and [more related content](#) is available

Download details:

IP Address: 130.251.61.251

The article was downloaded on 24/11/2009 at 07:29

Please note that [terms and conditions apply](#).

## TOPICAL REVIEW

# Image deblurring with Poisson data: from cells to galaxies

M Bertero<sup>1</sup>, P Boccacci<sup>1</sup>, G Desiderà<sup>1</sup> and G Vicidomini<sup>2</sup>

<sup>1</sup> Dipartimento di Informatica e Scienze dell'Informazione, Università di Genova, Via Dodecaneso 35, I 16146 Genova, Italy

<sup>2</sup> Department of NanoBiophotonics, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, Goettingen 37077, Germany

Received 3 June 2009

Published 23 November 2009

Online at [stacks.iop.org/IP/25/123006](http://stacks.iop.org/IP/25/123006)

## Abstract

Image deblurring is an important topic in imaging science. In this review, we consider together fluorescence microscopy and optical/infrared astronomy because of two common features: in both cases the imaging system can be described, with a sufficiently good approximation, by a convolution operator, whose kernel is the so-called point-spread function (PSF); moreover, the data are affected by photon noise, described by a Poisson process. This statistical property of the noise, that is common also to emission tomography, is the basis of maximum likelihood and Bayesian approaches introduced in the mid eighties. From then on, a huge amount of literature has been produced on these topics. This review is a tutorial and a review of a relevant part of this literature, including some of our previous contributions. We discuss the mathematical modeling of the process of image formation and detection, and we introduce the so-called Bayesian paradigm that provides the basis of the statistical treatment of the problem. Next, we describe and discuss the most frequently used algorithms as well as other approaches based on a different description of the Poisson noise. We conclude with a review of other topics related to image deblurring such as boundary effect correction, space-variant PSFs, super-resolution, blind deconvolution and multiple-image deconvolution.

## 1. Introduction

Image deblurring is a large-scale inverse problem because the size of the detected images is of the order of the megapixel or megavoxel. In many instances, it can be modeled as a problem of image deconvolution, and this is a great advantage from the computational point of view. Indeed, the imaging system is then characterized by a single *point-spread function* (PSF), which has the same size as the image. Moreover, efficient algorithms can be implemented, thanks to the use of *fast Fourier transform* (FFT).

Image deconvolution is one of the first linear inverse problems which has been widely investigated. For instance, in the book of Tikhonov and Arsenin [128] (the book was published in Russian in 1974), two chapters are devoted to the inversion of deconvolution operators, including a discussion of optimal regularization schemes. Moreover, the first review paper to describe deconvolution methods and applications (mainly to astronomy) is presumably the 1975 paper of Frieden [56], entitled *Image Enhancement and Restoration*. More recent reviews, mainly focused also on astronomical applications, are [106, 125]. Moreover, without any claim to be exhaustive, we point out at some books [65, 78, 136] or chapters of books [25, 124, 127, 132] dedicated to this problem.

The present review is focusing on two specific applications of image deconvolution, namely fluorescence microscopy and optical/infrared astronomy, because in both cases data are affected by photon noise that can be described as a Poisson process. Research on image deconvolution methods based on this statistical property of the noise started at the mid eighties with the seminal papers of Shepp and Vardi [114] and Geman and Geman [57].

Poisson data occur in all imaging processes where images are obtained by means of the count of particles, in general photons, arriving in the image domain. In the case of radioactive decay, fluorescence emission or similar phenomena, the arrival of particles is described by a *Poisson process*, i.e. a continuous stochastic process that is a collection of independent random variables  $\{N(t); t \geq 0\}$ , where  $N(t)$  is the number of particles arrived up to time  $t$ . The number of particles arriving within a given time interval  $T$  is a random variable (r.v.) with a *Poisson distribution* [52, 118], i.e. the probability of receiving  $n$  particles is given by

$$p(n) = \frac{e^{-\lambda} \lambda^n}{n!}, \quad n = 0, 1, 2, \dots, \quad (1)$$

where  $\lambda$ , proportional to  $T$ , is the expected value of the counts. This statistical model is appropriate to describe data acquired in fluorescence microscopy, emission tomography, optical/infrared astronomy, etc. Even if the energy of the photons (hence their wavelength) is different in the different applications, the statistics of the data is the same. The imaging system, however, significantly varies from an application to another. It is described by a sparse matrix in tomography, and by a convolution matrix in microscopy and astronomy.

As a consequence of the great interest dedicated to medical applications, such as *positron emission tomography* (PET) and *single photon emission computerized tomography* (SPECT), major advances related to the methods based on the Poisson noise model were obtained in this field and subsequently transferred to image deconvolution in microscopy and astronomy. Let us observe that the differences between these domains of application are not only related to the nature of the imaging matrix, but also to the specific features of the objects to be reconstructed. For instance, some astronomical images are characterized by a very high dynamic range that is never reached in medical and microscopy images: the ratio between the intensities of different objects in the same image can raise, for instance, up to  $10^9$ . This property may require the use of specific methods.

Research in astronomical image deconvolution has been boosted by the Hubble Space Telescope (HST) optical aberration problem at the beginning of its mission, before the implementation of the corrective optics in late 1993. The four years from early 1990 to late 1993 were years of very active research [64, 133], mainly based on the so-called Richardson–Lucy algorithm [92, 109], independently rediscovered by Shepp and Vardi as an expectation maximization algorithm for maximum likelihood (ML) estimation. A short account of the activity related to HST is given in [3].

The application of image deconvolution to microscopy was stimulated by the introduction of 3D fluorescence microscopy [4]. First, simple and fast processing methods such as nearest-neighbor and inverse filtering were applied [5]. Subsequently, constrained regularization based

on a modified van Cittert iteration [6] or on the Tikhonov method [23, 38] was introduced. The use of the maximum likelihood approach dates approximately from the same years [68]. A recent review is given in [111].

The purpose of this review is to provide, as far as possible, a sort of guide, helping the reader to find his/her way in the huge literature on the subject: it contains tutorial parts and also pointers to open problems. Obviously, it is biased by the personal experience of the authors and our choice is to focus on the statistical methods derived from the assumption of Poisson noise, namely the maximum likelihood approach of Shepp and Vardi and the Bayesian approach of Geman and Geman. In both cases, the problem of image deconvolution is reformulated as the minimization of suitable cost functions, with the addition of a constraint of nonnegativity. It follows that iterative algorithms for the solution of these problems have a key role. Besides, we also consider some other problems related to image deconvolution.

The review is organized as follows. In section 2, we discuss the mathematical modeling of the process of image formation and detection, including a discussion of the statistical properties of the noise. In section 3, we describe the Bayes paradigm derived from the assumption of Poisson noise, introducing first the maximum likelihood approach; moreover, we derive the functions whose minimizers provide the *maximum likelihood* or the *maximum a posteriori* (MAP) estimates. In Section 4, we discuss different iterative algorithms that have been designed for solving the optimization problems introduced in the previous section. We focus on gradient-based algorithms because their implementation is very simple and their efficiency can be greatly improved by recent advances in optimization theory. In section 5, we briefly discuss other approaches that can still be formulated in a Bayesian framework but are based on different likelihood functions for the Poisson noise; one of them has been devised for enforcing sparse representations of the solution. Section 6 is devoted to related topics such as boundary effect correction, space-variant PSFs, super-resolution, blind deconvolution and multiple-image deconvolution.

## 2. Image formation

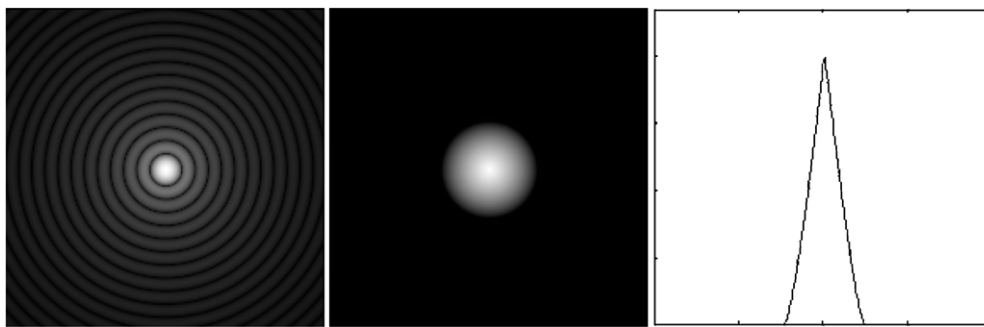
A microscope or a telescope is an imaging system that consists of two distinct parts: the *optical system* and the *detector*. Both must be taken into account for an accurate mathematical modeling of the imaging process.

### 2.1. The optical system

The optical system is a device consisting of optical components (lenses, mirrors, beam splitters, etc); it collects the light coming from the object to form an image in a plane, the *image plane*. In the case of microscopy, it includes also the system used for illuminating the sample. Two examples can help to understand the concept.

An example from astronomy is a telescope consisting of a *Cassegrain reflector*. It is a combination of a (large) concave circular mirror, the primary, and of a convex circular mirror, the secondary, situated above the primary but below its focal plane. The primary reflects the light coming from the sky toward its focal plane but, before reaching this plane, the light is reflected back by the secondary. Both mirrors are aligned in such a way that they have the same optical axis, and the primary contains a hole in its center, thus permitting the light from the secondary to reach a detector located in its focal plane, below the primary.

An example from microscopy is provided by the *fluorescence confocal microscope* [48, 103]. The objective lens creates a diffraction-limited excitation spot by focusing a plane wave (obtained, for instance, by means of a laser beam) in a point where the sample



**Figure 1.** The PSF of an ideal pupil (left), the corresponding MTF (center) and a section of the MTF along the central line (right).

is located. Then a system, consisting of the same objective lens and a tube lens, focuses the light, emitted by the fluorochromes in the excitation volume, into the image plane. A pinhole aperture, located in this plane, partially selects the light from the central part of the excitation volume. An image is formed by scanning the illumination point until the whole field of view is sampled. Moreover, since a confocal microscope can produce in-focus images of slices of a thick sample, a process known as *optical sectioning* [5, 43], a 3D image can be obtained by moving the specimen in the direction of the optical axis. Therefore, a 3D image is just a stack of 2D images.

Under the approximation of Fourier optics, if the light emitted by the object is spatially incoherent, then the relationship between the intensity of the object  $x(s)$  and that of the image  $y(s)$  ( $s$  being a 2D variable in astronomy and a 3D variable in microscopy) is linear [60]:

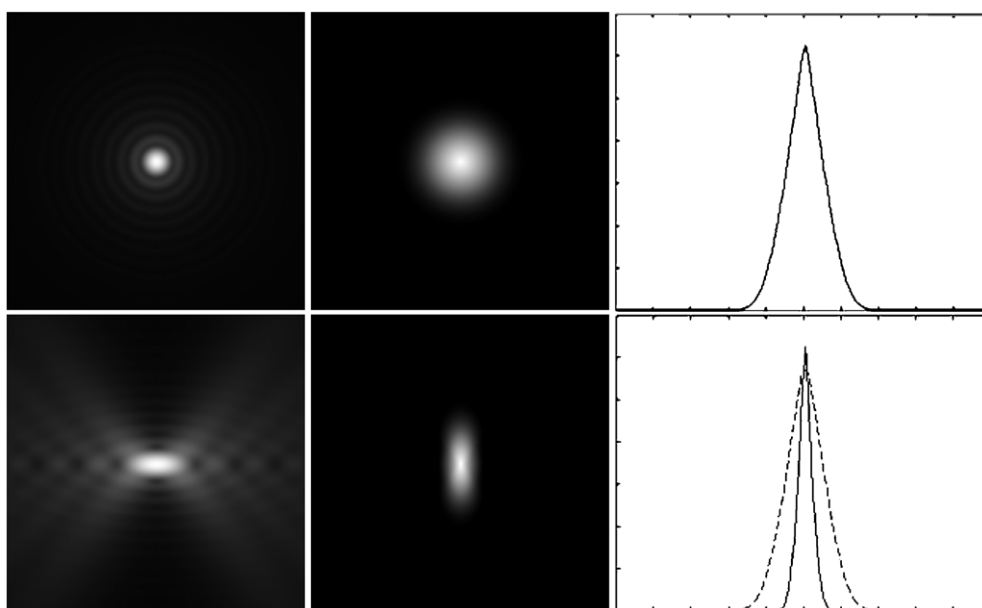
$$y(s) = \int H(s, s')x(s') ds'. \quad (2)$$

The integral kernel,  $H(s, s')$ , describes the effects of diffraction, aberrations, etc. Moreover, most imaging systems are, if not exactly at least approximately, isoplanatic, so that the kernel depends only on the difference of the variables:

$$y(s) = \int H(s - s')x(s') ds'. \quad (3)$$

The function  $H(s)$  is the PSF of the optical system. In figure 1, we show the PSF of an ideal pupil, the so-called *Airy pattern* [60], due to diffraction effects. It is, for instance, the PSF of a Cassegrain telescope in the absence of aberrations of the mirrors, obstructions due to mechanical supports and effects of atmospheric turbulence. We also show the modulus of the Fourier transform of this PSF, the so-called *modular transfer function* (MTF) [60]. It is evident that the PSF is *bandlimited*, a consequence of diffraction. The radius of the disc in the Fourier domain is proportional to the diameter,  $D$ , of the pupil and inversely proportional to the wavelength,  $\lambda$ , of the incoming radiation. Therefore, the sampling distance is proportional to  $\lambda/D$  and this is also the so-called *diffraction limit*, characterizing the resolving power of the instrument.

In figure 2, we show the 3D aberration-free PSF of a confocal microscope by showing two 2D sections, respectively, in the focal plane (upper-left panel) and in a plane containing the optical axis (lower-left panel). In the central panels, we show the corresponding 2D sections of the MTF and their 1D sections in the left panels. These pictures make evident that the lateral bandwidth is around three times larger than the axial bandwidth; hence the lateral resolution is around three times better than the axial one. The function shown in the upper-left panel is



**Figure 2.** 2D sections of the 3D PSF of a confocal microscope: in the focal plane (upper-left panel) and in a plane through the optical axis (lower-left panel); the corresponding 2D sections of the MTF (central panels) and their 1D sections (right panels). In the lower-right panel, we give two 1D sections (horizontal and vertical) showing the approximate ratio of 3 between the lateral and the axial bandwidth.

approximately the square of the Airy pattern (left panel of figure 1). For this reason, its rings are fainter than those of the Airy pattern, when displayed with the same gray scale.

Different kinds of aberrations introduce, in general, modifications of the ideal PSFs. These can be due to aberrations in the optical components of the instruments. However, in the case of telescopes, the main troubles derive from atmospheric turbulence, while, in the case of microscopes, they derive from inhomogeneities in the refractive index of the specimen. The result is to produce an effective band considerably smaller than the band of the optical instrument. However, in modern ground-based telescopes, a new technology, called *adaptive optics* (AO), allows us to compensate, at least partially, for the effect of atmospheric turbulence [50] and to get a resolution close to the diffraction limit. A similar technology has been introduced also in microscopy [33] for a partial compensation of the effect of spatial variations of the refractive index of the specimen.

## 2.2. The detector

A detector is located in the image plane for measuring the incoming radiation. In astronomy, it is a *charged coupled device* (CCD) camera while, in fluorescence microscopy, it can be also a *photomultiplier tube* (PMT) or an *avalanche photo-diode* (APD), as in the case of confocal, 4Pi, two-photon excitation and other microscopy techniques [48, 67]. The detection process introduces *sampling* and *noise*.

We know that optical instruments are bandlimited, and therefore sampling is designed to satisfy the requirements of the well-known Whittaker–Shannon sampling theorem (but a moderate oversampling, by a factor of 2 or 3, is very frequent, to avoid interpolation for getting

an image that can easily be interpreted by the user). Sampling is performed on a uniform grid in the image plane. In the case of astronomy, this is obtained by means of a CCD camera: each element of the camera defines a pixel of a 2D digital image. In the case of microscopy, as we have already mentioned, a 3D image is obtained by focusing the instrument at different depths and acquiring a 2D image for each depth by means of a scanning system. In confocal microscopy, the sampling rate is usually uniform along the two lateral directions, while, along the optical axis, it is smaller by approximately a factor of 3, in order to accommodate the difference between lateral and axial resolution (see figure 2). As a consequence, a voxel can be viewed as a parallelepiped with a square basis.

Since we have a discrete image, we must consider the discretization of equation (3). Thanks to the mapping, provided by geometric optics, of the image domain into the object domain, this can be subdivided into the same number of pixels or voxels, so that the convolution equation is replaced by a matrix equation,  $y = Hx$ . The discrete convolution can be given in terms of a block Toeplitz or block circulant matrix, and the discrete PSF is normalized to 1, so that

$$y_i = \sum_{j \in S} H_{i-j} x_j, \quad i \in S, \quad \sum_{i \in S} H_i = 1, \quad (4)$$

where  $y_i$  are the sampled values of the image, and  $x_i$  are those of the object,  $i$  being a multi-index characterizing the pixels/voxels of both the image and the object.  $S$  denotes the set of index values.

The measured values, denoted again by  $y_i$ , are affected by noise, and therefore are not given by equation (4), but are realization of suitable r.v.s  $Y_i$ , describing different kinds of noise. A first contribution to  $Y_i$  is due to the radiation emitted by the object  $x$ ; this is a Poisson r.v. and the term  $(Hx)_i$  in equation (4) is interpreted as its expected value. Obviously, one assumes that  $H$  is known while  $x$  is not.

A second contribution is provided by *background emission*. In astronomy, this is due, for instance, to sky emission [105, 117]. In fluorescence microscopy, it is mainly due to auto-fluorescence (generated by the medium embedding the sample or by the sample itself) and by reflections of the excitation light. This is again a Poisson r.v. and an estimation of its expected value, denoted as  $b_i$ , is required and can be obtained, for instance, by a preprocessing of the detected image. We also remark that in some astronomical images, the background value can be much larger than that of the signal superimposed to it. These images are usually called *background dominated* and their processing is an important issue in astronomy.

When object and background emission are statistically independent (or, at least, when this can be reasonably assumed), their contribution to  $Y_i$  is a Poisson r.v. with the expected value  $(Hx + b)_i$ .

However, there are still other sources of noise. For instance, an accurate description of the noise introduced by a CCD camera is given in [119]. It is a mixture of the Poisson process described above, and of an additive Gaussian process, due to the amplifying system of the camera and called read-out noise (RON). The more general case, where both kinds of noise are taken into account, has also been considered [20, 90, 120], but this more refined model seems to be relevant only in some specific situations. Therefore, one can assume that photon noise is dominating.

In the case of fluorescence microscopy, the noise introduced by PMT and APD is described in [10], but also in the case of confocal, 4Pi, two photons excitation and other microscopy techniques, the main source of noise is due to photon counting. Therefore, this is the situation we consider in this paper.

In conclusion, we assume that the detected value  $y_i$  of the image  $y(s)$  is the realization of a Poisson r.v.  $Y_i$ , with the expected value  $(Hx + b)_i$ , i.e.

$$Y_i \sim \text{Poisson}\{(Hx + b)_i\}, \quad (5)$$

and in this approximation, the detected value  $y_i$  of  $Y_i$  is an integer number. We denote by  $Y$  the multi-valued r.v. formed by the set of the  $Y_i$ . These are assumed to be *statistically independent*. The data of the problem are the imaging matrix  $H$ , the expected value of the background  $b$  and the detected image  $y$ .

### 3. The Bayesian paradigm

A *maximum likelihood* method for image reconstruction in emission tomography, based on the assumption of Poisson data, was introduced by Shepp and Vardi [114] in 1982. A few years later, a general *Bayesian paradigm* for problems of image reconstruction was proposed by Geman and Geman [57], and Grenander [63]. This approach, combined with the Poisson likelihood introduced in [114], was applied to SPECT [58, 59] and the resulting formulation of the restoration problem is applicable to any linear system with data perturbed by Poisson noise. From the foundational point of view of inverse and ill-posed problem, the interesting point in a Bayesian approach is the use of *a priori* information in the form of a given probability distribution of the unknown object.

#### 3.1. The likelihood

From equations (1) and (5) and the assumed statistical independence of the r.v.s  $Y_i$ , the probability distribution of  $Y$ , for given  $H$ ,  $x$  and  $b$ , is as follows:

$$p_Y(y; x) = \prod_{i \in S} \frac{e^{-(Hx+b)_i} (Hx + b)_i^{y_i}}{y_i!}. \quad (6)$$

The ML approach consists in introducing a function of  $x$ , the likelihood, that is defined by  $L_y^Y(x) = p_Y(y; x)$ , where now  $y$  is given and is just the detected image. Then a *ML estimate* of the unknown object is defined as any maximizer  $x^*$  of  $L_y^Y(x)$ .

By taking the negative logarithm of the likelihood, one obtains, up to an additive constant, the Kullback–Leibler (KL) divergence of  $Hx + b$  from  $y$  :

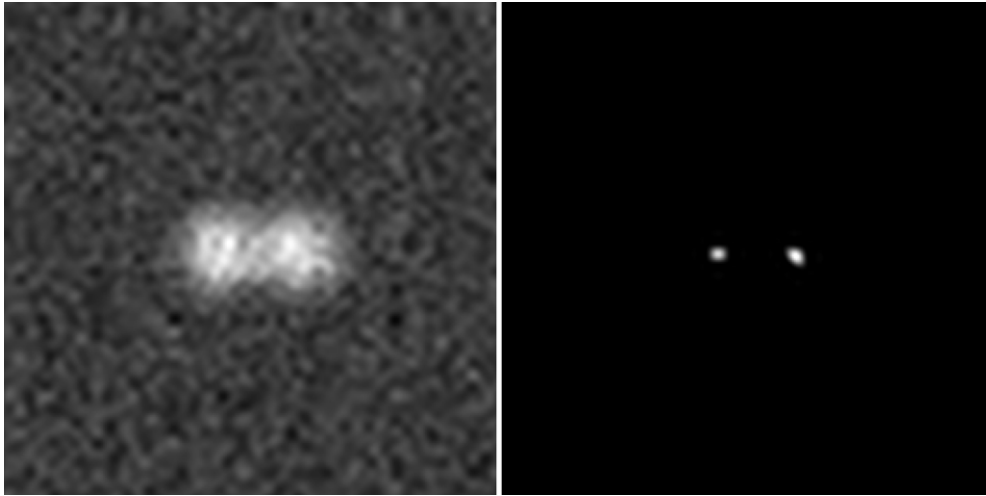
$$f_0(x) = D_{\text{KL}}(Hx + b, y) = \sum_{i \in S} \left\{ y_i \ln \frac{y_i}{(Hx + b)_i} + (Hx + b)_i - y_i \right\}, \quad (7)$$

and the problem becomes the minimization of this function on the nonnegative orthant.

The KL divergence replaces the metric distance that appears in the standard least-square approach, and it can be considered as a discrepancy or a data-fidelity function. It is convex, nonnegative and coercive on the nonnegative orthant, so that minimizers exist and are global. Gradient and Hessian are given by (remark that the normalization of the PSF in equation (4) implies  $H^T \mathbf{1} = \mathbf{1}$ , where  $\mathbf{1}$  is the identity array/cube)

$$\nabla f_0(x) = \mathbf{1} - H^T \frac{y}{Hx + b}, \quad \nabla^2 f_0(x) = H^T \text{diag} \frac{y}{(Hx + b)^2} H, \quad (8)$$

where the quotient and product of arrays/cubes are defined in the Hadamard sense, i.e. pixel by pixel or voxel by voxel. We see that the function is strictly convex if the matrix  $H$  is non-singular (we remark that, even if the PSF is bandlimited, the discrete  $H$  satisfies this



**Figure 3.** The blurred and noisy image of a binary star (left) and the corresponding minimizer of the KL divergence (right). Both images are represented in a log scale.

condition because of approximation errors) and in such a case the minimizer is unique. In the case of background, an additional condition is required to exclude that the minimizer is just 0. We also remark that the Karush–Kuhn–Tucker (KKT) conditions are necessary and sufficient conditions to be satisfied by the minimizers  $x^*$  of  $f_0(x)$ . In our specific case they can be written as follows:

$$x^* = x^* H^T \frac{y}{Hx^* + b}, \quad x^* \geq 0, \quad H^T \frac{y}{Hx^* + b} \geq 1. \quad (9)$$

It is obvious that, if the equation,  $Hx + b = y$ , has a nonnegative solution, then this is also a minimizer of  $f_0(x)$ ; however, because of the extreme ill-conditioning of  $H$ , such a solution, in general, does not exist and, as a consequence, a minimizer of  $f_0(x)$  may not provide a sensible solution of the image deblurring problem.

In an attempt of understanding the reliability of the minimizers, let us remark that, in the case of zero background, any minimizer  $x^*$  satisfies the  $\ell_1$ -constraint:

$$\sum_{i \in S} x_i^* = \sum_{i \in S} y_i. \quad (10)$$

It is known that this constraint induces sparsity; hence a minimizer is a sparse solution, i.e. it lies on the boundary of the nonnegative orthant. If the object is a diffuse one (as a nebula, in astronomy, or micro-tubules, in microscopy) such a reconstruction is not acceptable. The sparsity of the minimizers is also known as *checkerboard effect* [100] or *night-sky reconstruction* [18]. However, this remark suggests that the minimization of the KL divergence can provide sensible solutions in the case of sparse objects, such as a star cluster in astronomy, and localized structures like sub-resolved vesicles in microscopy. This argument can be supported by numerical experiments. In figure 3, we show the result obtained in the case of a toy example. The object is a binary consisting of two stars with equal intensity, set to  $10^4$ , over a background of 100. The object is convolved with the PSF of figure 1 (left panel) and is perturbed with Poisson noise. The result is shown in the left panel of figure 3. The minimizer is computed by pushing to convergence the iterative algorithm of equation (17) and is shown in the right panel of the same figure. The reconstruction is essentially free of artifacts, as evidenced by the log scale used for the representation of the result.

### 3.2. The prior

The basic instability of the minimizers of the Poisson likelihood in the case of non-sparse objects was soon recognized, and an application of the method of *sieves*, proposed by Grenander [62], was used for obtaining stable solutions [116]. In the language of inverse and ill-posed problems this is essentially a *mollifier* method. However, we are mainly interested in the Bayesian approach that has widely been investigated in the case of emission tomography and also applied to astronomy and microscopy. The main idea is the same underlying any treatment of inverse problems, namely the use of *a priori* information about the object to be recovered. In a Bayesian approach, the unknown object  $x$  is also considered as a realization of a multi-valued r.v., let us say  $X$ , and the *a priori* information is encoded into a given probability distribution  $p_X(x)$  of this r.v., the so-called prior probability distribution, or simply the *prior*.

If the probability distribution (5) is viewed as a conditional probability of  $Y$  for a given value of  $X$ , i.e.  $p_Y(y|x) = p_Y(y|X=x) \doteq p_Y(y|x)$ , then the Bayes formula provides the conditional probability of  $X$  for a given value  $y$  of  $Y$ :

$$p_X(x|y) = \frac{p_Y(y|x)p_X(x)}{p_Y(y)}. \quad (11)$$

If we insert in this equation the detected value of  $y$ , we obtain the so-called *posterior probability distribution* of  $X$ ,  $P_y^X(x) = L_y^Y(x)p_X(x)/p_Y(y)$ .

As concerns the choice of the prior, in general it is assumed that  $X$  is a Markov random field (MRF) or a Gibbs random field, the latter being a r.v. with a probability distribution given by

$$p_X(x) = \frac{1}{Z} e^{-\beta f_1(x)}, \quad (12)$$

where  $f_1(x)$  is a given function, usually called the energy function or penalization function, and  $Z$  is a normalization constant. Since, thanks to Hammersley–Clifford theorem, any MRF is equivalent to a suitable Gibbs random field, we can take equation (12) as the general form of the prior. In the case of a MRF,  $f_1(x)$  is a sum of *clique potentials* that depend only on the values of  $x$  in the sites of the clique. For an understanding of the features of the samples corresponding to a Gibbs distribution with a given function  $f_1(x)$  we refer to [82], where interesting examples are given with  $f_1(x)$  corresponding to Tikhonov, total variation and sparsity regularization.

By definition, a MAP estimate is any maximizer of the posterior probability distribution. By taking again the negative log of this function, we see that a MAP estimate is any object  $x_\beta^*$  minimizing the function:

$$f_\beta(x) = f_0(x) + \beta f_1(x). \quad (13)$$

It is quite interesting to remark that this function has a structure quite similar to that of Tikhonov regularization functional, the metric distance of  $Hx + b$  from  $y$  being replaced by the KL divergence. Hence, the parameter,  $\beta$ , plays a role similar to that of the regularization parameter in Tikhonov theory.

Functions of the form (13) have been used in several applications to microscopy, with different forms of the penalizing function  $f_1(x)$ . We mention the case of quadratic penalization,  $f_1(x) = \|x\|_2^2/2$  [42, 93], of quadratic Laplacian,  $f_1(x) = \|\Delta x\|_2^2/2$  [139], of total variation,  $f_1(x) = \|\|\nabla x\|\|_1$  [47], and of Good's roughness penalization,  $f_1(x) = \|\|\nabla x\|^2/x\|_1$  [80].

In the case of Tikhonov theory, a well-established criterion for the choice of the regularization parameter  $\beta$  is the so-called *Morozov discrepancy principle* [51]. A similar principle has recently been proposed in the case of the denoising of Poisson data with edge-preserving regularization [137]. On the basis of statistical arguments, the metric distance

between computed and detected data, used in the standard formulation of the Morozov principle, is replaced by the KL divergence. The principle can be extended to image deblurring and can be formulated as follows: search for the value of  $\beta$  such that

$$D_{\text{KL}}(Hx_{\beta}^* + b, y) = \frac{n}{2}, \quad (14)$$

where  $n$  is the cardinality of the set  $S$ . Work is in progress in order to provide a sound foundation of this principle.

### 3.3. The continuous model

The previous approach is completely discrete since it is close to the physics and statistics of image detection. However, in some occasions, a continuous model may be desirable, even if deprived of physical and statistical meaning, because it allows us to use the powerful tools of functional analysis. The straightforward extension of the KL divergence (7) to the continuous case is given by the following functional:

$$F_0(x) = \int \left\{ y(s) \ln \frac{y(s)}{(Hx + b)(s)} + (Hx + b)(s) - y(s) \right\} ds, \quad (15)$$

where  $H$  is the integral operator defined in equation (2) or (3).

This functional is convex and nonnegative on the closed and convex set of nonnegative functions in Lebesgue spaces. It is strictly convex if the null space of the integral operator (2) or (3) is trivial. In the case of the operator (3) this is not true when the integration domain is the full  $\mathcal{R}^2$  or  $\mathcal{R}^3$  space, because the imaging systems we are considering are bandlimited.

The nonnegative solutions of the equation,  $Hx + b = y$ , (with  $H, b$  and  $y$  given) are also minimizers of  $f_0(x)$ . If these solutions do not exist, then minimizers may also not exist, as shown in [97] by means of an example in the case  $b = 0$ . Therefore, the problem of minimizing  $F_0(x)$  is ill-posed, and regularization is required. Formally, this is obtained by considering functionals of a form similar to (13):

$$F_{\beta}(x) = F_0(x) + \beta F_1(x), \quad (16)$$

with different choices of the regularization functional  $F_1(x)$ . For instance, regularization in terms of standard Tikhonov regularization,  $F_1(x) = \|x\|^2/2$ , is investigated in [15]; regularization in terms of total variation functional is considered in [16, 35], while [17] concerns regularization in terms of differential diffusion operators.

## 4. Algorithms

In this section, we turn back to the discrete problem and we discuss iterative algorithms for the minimization of the KL divergence or of the penalized KL divergence derived from a Bayesian approach. In general, iterative methods for the minimization of functionals related to the solution of inverse problems can be roughly divided into two classes: *explicit methods* requiring only matrix–vector multiplication per iteration and *implicit methods* requiring to solve a linear equation per iteration. Even if implicit methods have fast asymptotic convergence (for application to inverse problems, see, for instance [132]), explicit methods are preferred in many instances for their simplicity and their *semiconvergence* in the case of non-regularized problems. Moreover, recent results in optimization theory can make these methods competitive with the implicit ones, and therefore we focus on them.

#### 4.1. The basic algorithm

The iterative algorithm most frequently used in image deconvolution with Poisson data was proposed independently by Richardson [109] and Lucy [92]. It is generally known as the Richardson–Lucy (RL) algorithm, in applications to astronomy and microscopy. In the paper by Shepp and Vardi [114], the same algorithm was re-obtained by considering the maximization of the likelihood function of equation (6), and applying a general approach known in statistics as the *expectation maximization* (EM) method [45]. For this reason, in emission tomography, the acronym EM (sometimes ML-EM) is used for denoting this specific algorithm. First application to microscopy was investigated in [68–71] (see also [112] for an application to 4Pi microscopy), while application to astronomy was stimulated by the restoration of HST images [64, 133], as already mentioned in the introduction.

Here we give the algorithm in the form proposed in [105, 117] for taking into account background emission. In fact, in the presence of a background, the nonnegativity constraint is active only when using this modified form of the iteration; otherwise, ringing effects may appear in the reconstructed image [89]. If we assume that the PSF satisfies the normalization condition of equation (4), then, for  $k = 0, 1, \dots$ , we have

$$x^{(k+1)} = x^{(k)} H^T \frac{y}{Hx^{(k)} + b}, \quad (17)$$

the iteration being, in general, initialized with a constant array/cube. In such a case, all the  $x^{(k)}$  are strictly positive. We also remark that this iteration can be formally derived by applying the method of successive approximations to the first KKT condition (9).

In the case  $b = 0$ , all the  $x^{(k)}$  satisfy condition (10), and this property is a key point in all the convergence proofs we know [75, 76, 85, 96, 130] (see also [100]); since it is not satisfied if  $b \neq 0$ , a convergence proof seems to be lacking in this case.

Note that in the continuous case the algorithm takes the same form—with the obvious changes in the interpretation of the symbols—if the integral kernel  $H(s)$  satisfies a normalization condition similar to that of equation (4), the sum being replaced by an integral. Moreover, in the case  $b = 0$ , the iterations satisfy condition (10), again with the sum replaced by an integral. In this case, the continuous algorithm is investigated in [95–97] where it is proved that, if the algorithm converges to  $x^*$ , then  $x^*$  is a minimizer of (15). Moreover,  $F_0(x^{(k)})$  is decreasing with  $k$  and the KL divergence  $D_{\text{KL}}(x^*, x^{(k)})$  is decreasing too. However, the first convergence result is proved in [107], where it is shown that, if the equation,  $Hx = y$ , has a nonnegative solution  $x^*$ , i.e. if the data are in the range of the operator  $H$ , then the iteration converges to  $x^*$  with respect to weak topology in Lebesgue spaces.

It is known, from numerical practice, that the RL algorithm has regularization properties [25, 100] in the case of the reconstruction of non-sparse objects: sensible solutions can be obtained by an early stopping of the iteration, i.e. the iteration must not be pushed to convergence and the algorithm is semiconvergent (see, however, the remark of section 2.1 about the restoration of sparse objects). Therefore, the problem arises of finding appropriate stopping rules. First attempts in this direction are investigated in [12, 107]. In [107], a first analysis of some regularization properties of the algorithm is also given.

#### 4.2. Acceleration of the basic algorithm

In this review, we consider the RL algorithm not from a statistical point of view (i.e. as a particular case of the EM method) but from an analytic point of view, more precisely as an optimization method. To this purpose, if we consider the iteration

$$x^{(k+1)} = x^{(k)} - \lambda x^{(k)} \nabla f_0(x^{(k)}), \quad (18)$$

then from equation (8) we see that for  $\lambda = 1$  we re-obtain the RL algorithm. In other words,  $d^{(k)} = -x^{(k)} \nabla f_0(x^{(k)})$  is a descent and feasible direction for the minimization of  $f_0(x)$  on the nonnegative orthant, and the RL algorithm is a scaled gradient method.

Since the convergence of the algorithm is very slow, the problem of accelerating its convergence has widely been investigated. A reduction of the number of iterations can be obtained by considering the form (18) and by using line-search techniques in order to optimize the decrease of the objective function along the descent direction (see, for instance, [72] in the case of microscopy, and [1] in the case of astronomy); similar results are obtained in [87] by means of standard tools in optimization, such as the Armijo rule [30]. This implies an increase of the computational cost per iteration, and therefore a reduction of the advantage provided by the reduction of the number of iterations. Another acceleration can be obtained by the so-called multiplicative relaxation [75, 87], even if, in general, the gain is not very large.

A very significant improvement is provided by the method proposed in [31]. This method is based on a sort of extrapolation along the trajectory of the iteration and can be applied to any iteration of the form  $x^{(k+1)} = T(x^{(k)})$ , where  $T$  is a nonlinear operator independent of  $k$ . It is implemented in the *lucy* task of the *Image Processing* toolbox of MATLAB, and when convergence is very slow, it can provide an improvement by a factor 10–20 in the number of iterations, with a moderate increase in the computational cost per iteration. Therefore, it could be applied to a wide class of iterative methods. The weakness of the approach is that no convergence proof is available.

A similar and sometimes better improvement is obtained by means of a recently proposed method [34], based on the scaling of the gradient suggested by the RL algorithm. This approach is called the *scaled gradient projection* (SGP) method and, in this case, a convergence proof is available. Moreover, it can be applied to other iterative algorithms based on a descent direction obtained with a suitable scaling of the gradient. In general, it can be used for solving the minimization of a continuously differentiable function  $f(x)$  over a closed and convex set  $\mathcal{C}$  (in our case the nonnegative orthant). If  $x^{(k)}$  is the current iteration, then, as proved in [34], for any given positive step length  $\alpha_k$  and any positive definite scaling matrix  $S_k$ ,

$$d^{(k)} = P_{\mathcal{C}, S_k} \{x^{(k)} - \alpha_k S_k \nabla f(x^{(k)})\} - x^{(k)} \quad (19)$$

is a feasible and descent direction, the projection being on  $\mathcal{C}$  in the metric induced by the scalar product  $(S_k^{-1}x, z)$ . Finally, the new iteration is obtained by means of a line search along the descent direction,

$$x^{(k+1)} = x^{(k)} + \lambda d^{(k)}, \quad (20)$$

based, for instance, on the Armijo rule [30]. Convergence of the algorithm is proved for any selection of the step lengths and scalings in positive compact sets. Acceleration is obtained by using suitable Barzilai–Borwein rules [19] for step-length selection (details can be found in [34]) and suitable scaling. This point will be discussed in the following section.

### 4.3. Algorithms for penalized minimization

Regularization by means of early stopping of the RL iteration is not always satisfactory. For this reason, after the seminal paper of Geman and Geman [57], several types of regularization have been considered, and several types of iterative algorithms for the minimization of function (13) have been proposed. Most of them are obtained from the EM approach, mainly in the case of emission tomography. A complete review is outside the scope of this review. We focus on explicit methods and, in particular, on a class of algorithms that are close to the basic algorithm (17).

A prototype is the *one-step late* (OSL) algorithm [61] that can be used if the function  $f_1(x)$  in equation (13) is continuously differentiable. The corresponding iteration is as follows:

$$x^{(k+1)} = \frac{x^{(k)}}{1 + \beta \nabla f_1(x^{(k)})} H^T \frac{y}{Hx^{(k)} + b}. \quad (21)$$

It is easy to recognize that this algorithm is also a scaled gradient method as follows from the following iteration with  $\lambda = 1$  :

$$x^{(k+1)} = x^{(k)} - \lambda \frac{x^{(k)}}{1 + \beta \nabla f_1(x^{(k)})} \nabla f_\beta(x^{(k)}), \quad (22)$$

using again the expression of the gradient of  $f_0(x)$  given in equation (8). The problem is that the scaling is not necessarily positive since the gradient of  $f_1(x)$  can take negative values. However, if this gradient is bounded, then the scaling is nonnegative for sufficiently small values of  $\beta$ . Under this condition, convergence is proved in [86] by introducing a suitable line search along the descent direction. Incidentally, we remark that, in the case of standard Tikhonov regularization, i.e.  $f_1(x) = \|x\|^2/2$ , the gradient is given by  $\nabla f_1(x) = x$ , and therefore it is nonnegative on the nonnegative orthant; it follows that, in such a case, no restriction on the regularization parameter is required. We also remark that, in spite of the difficulties mentioned above, the algorithm (21) has been applied, with some success, to 3D image deconvolution in microscopy with edge-preserving regularization [47].

It is possible to overcome the main difficulty of OSL by means of an approach called the *split-gradient method* (SGM) [87, 88] and based, in this particular case, on a decomposition of the gradient of the regularization function in a positive and negative part:

$$-\nabla f_1(x) = U_1(x) - V_1(x), \quad (23)$$

with  $U_1(x) \geq 0$ ,  $V_1(x) \geq 0$  in the feasible domain (in general, the nonnegative orthant). Then the iteration (21) is replaced by the following one:

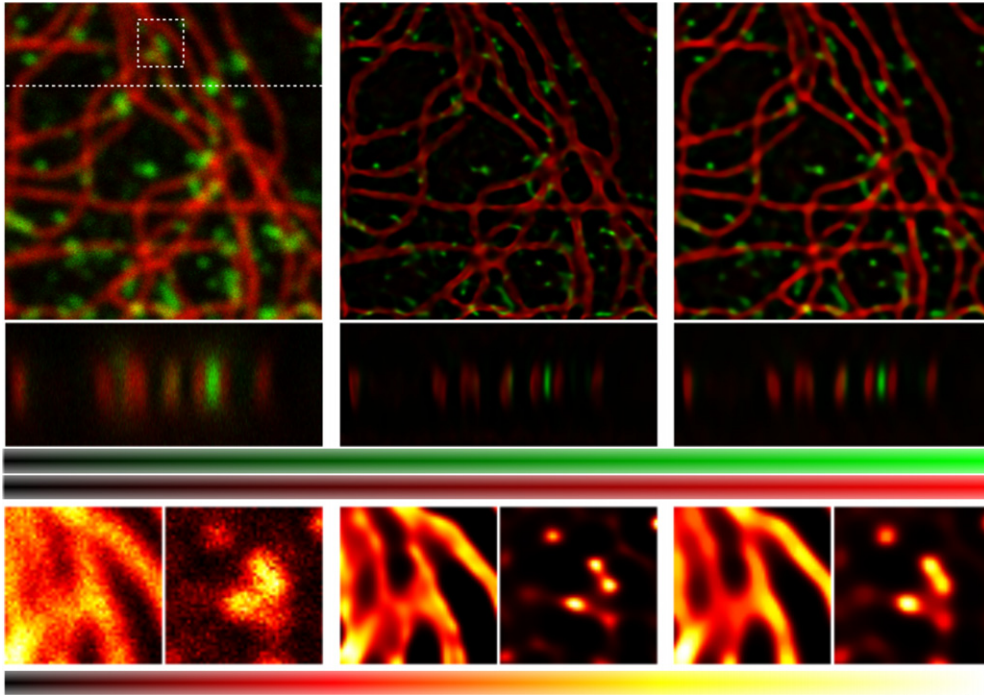
$$x^{(k+1)} = \frac{x^{(k)}}{1 + \beta V_1(x^{(k)})} \left\{ H^T \frac{y}{Hx^{(k)} + b} + \beta U_1(x^{(k)}) \right\}, \quad (24)$$

and it is easy to recognize that this iteration is also a scaled gradient method, but with a scaling that now is certainly nonnegative (remark that, if the initial guess is strictly positive, then all the  $x^{(k)}$  have the same property, so that the scaling is positive). Of course, the choice of the functions  $U_1(x)$ ,  $V_1(x)$  is largely arbitrary, but in [88] it is shown that a ‘natural’ choice exists for all the frequently used regularization functions. An example is given in the following.

The algorithm has been applied successfully to 3D image deconvolution in fluorescence microscopy [131] with edge-preserving regularization (in the sense discussed in [40]) and, in such a case, the algorithm is stable and robust for various levels of Poisson noise. The drawback is that no convergence proof is available. However, convergence can be obtained, as in the case of OSL, by a suitable line search along the descent direction. In [87, 88], the Armijo rule is used for this purpose. Moreover, thanks to the structure of the algorithm, it can easily be modified in the vein of SGP [29], so that one has both a convergence proof and an increase in efficiency. Work is in progress.

Such an approach has already been used for the denoising of 2D Poisson data [137], with the following edge-preserving regularization called *hypersurface regularization* [40] and related to the area of the surface of the graph of the function  $x$  :

$$f_1(x) = \frac{1}{2} \sum_{i,j=1}^n \psi_\delta(D_{i,j}^2), \quad D_{i,j}^2 = (x_{i+1,j} - x_{i,j})^2 + (x_{i,j+1} - x_{i,j})^2, \quad (25)$$



**Figure 4.** 3D deconvolution of a confocal image showing a detail of a mammalian cell (PtK2 line). Left-top: two color original image of micro-tubules (red) and clathrin-coated vesicles (green). Center-top: RL reconstruction. Right-top: edge-preserving reconstruction. For the original image and the two reconstructions, one axial view across the dotted line is also given. Lower panels: the magnification of the region interior to the white box is given for enhancing the differences among the three images. Micro-tubules are shown in the left squares and vesicles in the right ones.

where  $\psi_\delta(t) = 2\sqrt{t + \delta^2}$ , and  $\delta$  is a thresholding parameter tuning the jumps in the discontinuities of the function (when it is very small, we get an approximation of total variation regularization). Then, by computing the gradient of this function, it is easy to see that, on the nonnegative orthant, a possible splitting is the following

$$[U_1(x)]_{i,j} = \psi'_\delta(D_{i,j}^2)(x_{i+1,j} + x_{i,j+1}) + \psi'_\delta(D_{i,j-1}^2)x_{i,j-1} + \psi'_\delta(D_{i-1,j}^2)x_{i-1,j}, \quad (26)$$

$$[V_1(x)]_{i,j} = \{2\psi'_\delta(D_{i,j}^2) + \psi'_\delta(D_{i,j-1}^2) + \psi'_\delta(D_{i-1,j}^2)\} x_{i,j}, \quad (27)$$

where  $\psi'_\delta(t) = (t + \delta^2)^{-1/2}$ . The use of the scaling derived from this splitting provides a considerable improvement in efficiency with respect to the most efficient gradient projection methods.

We conclude this section with an example of deconvolution of a 3D confocal image using the edge-preserving algorithm proposed in [131], which can essentially be obtained by combining the splitting (26)–(27) with the iteration (24). In figure 4, the result of this deconvolution is compared with that obtained by means of the standard RL algorithm. In the left panels we show the original image, in the central panels the RL reconstruction and in the left panels the edge-preserving reconstruction. It is important to observe that, for sparse objects such as the clathrin-coated vesicles (green), RL provides a much better reconstruction than the edge-preserving method. On the other hand, in the interior of almost uniform objects like the

micro-tubules (red), RL increases the dynamic range thus reducing the global contrast with respect to the background. Piecewise blocky structures are well preserved by edge-preserving regularization.

## 5. Other approaches

In the previous sections, we focused on the approaches based on the use of the KL divergence as a measure of the discrepancy between the detected and computed data. However, in the framework of the Bayesian paradigm, other data-fidelity functions have been proposed and investigated for the processing of Poisson data. We give a brief account of two of them.

### 5.1. Weighted least-square methods

If the numbers of counts  $y_i$  are large, then the deviations  $y_i - (Hx^* + b)_i$ , where  $x^*$  is a minimizer, can be considered small with respect to  $y_i$  (remember that, in a Poisson process with the expected value  $\lambda$ , the standard deviation  $\sigma$  is just  $\sqrt{\lambda}$ ). Then, using the second-order approximation of the log function,  $\ln(1 + \xi) = \xi - \xi^2/2 + O(\xi^3)$ , we can write, for values of  $x$  close to  $x^*$  (for simplicity, we omit the index  $i$  and we write  $\lambda = Hx + b$ ) and neglecting terms of the order of  $(n - \lambda)^3/\lambda^2$ ,

$$y \ln \left( \frac{y}{\lambda} \right) \simeq y - \lambda + \frac{1}{2y} (y - \lambda)^2. \quad (28)$$

If this approximation can be used everywhere in the image domain, then the KL divergence can be approximated by a weighted least-square function,

$$D_{\text{KL}}(Hx + b, y) \simeq \frac{1}{2} \sum_{i \in S} \frac{1}{y_i} |(Hx + b)_i - y_i|^2, \quad (29)$$

and, therefore, all methods developed in the framework of the least-square approach can be used. In fact, this approximation, combined with a preconditioned conjugate gradient (CG), is used since a long time in medical imaging (see, for instance, [37, 55, 74]). It has recently been re-proposed in the context of image deblurring [13, 14].

### 5.2. Deconvolution using sparse representation

Image deconvolution with the assumption that the solution has a sparse expansion on an arbitrary preassigned orthonormal basis has widely been investigated mainly in the context of additive Gaussian noise. We only mention the seminal paper of Daubechies *et al* [44].

A multiscale framework for the Poisson likelihood that is tractable with Haar wavelets is investigated in [102], and regularization is imposed in a Bayesian paradigm. As far as we know, methods based on the penalization of the KL divergence by means of functionals enforcing sparsity in some representation are not available. However, several wavelet-based regularization of the RL algorithm have been proposed. Roughly speaking, this is obtained by combining a wavelet-based denoising of the data with a deconvolution method derived from RL. In the case of astronomical imaging, this approach is proposed in [77, 122, 123], while a combined approach based on wavelet and curvelet transforms is proposed in [126] (for an account of these works in the general context of image deconvolution, see also [124, 125, 127]). In the case of confocal microscopy, wavelet denoising combined with a modified version of the RL algorithm is presented in [94], while in [110] the denoising is performed with a variant of the wavelet transform, namely the steerable pyramid [115].

The denoising of Poisson data, based on wavelet or similar transforms, is, in general, performed after a preprocessing consisting in the use of transforms able to stabilize the variance of the noise, such as the Anscombe transform or generalizations of it [9, 138]. We recall that the Anscombe transform consists in replacing the detected data  $y_i$  with the following ones:

$$z_i = 2\sqrt{y_i + \frac{3}{8}}. \quad (30)$$

If the expected value of  $y_i$  is large, then  $z_i$  is approximately perturbed by a white Gaussian noise with  $\sigma = 1$  (for a discussion, see [124]).

In a recent paper [49], the deconvolution problem has been formulated in terms of the nonlinear model implicit in equation (30): the data-fidelity function  $f_0(x)$  is defined by

$$f_0(x) = \frac{1}{2} \sum_{i \in S} \left| 2\sqrt{(Hx)_i + \frac{3}{8}} - z_i \right|^2. \quad (31)$$

Then it is assumed that  $x$  can be written as a superposition of elementary atoms  $\varphi_\gamma$ , parametrized by  $\gamma \in \mathcal{I}$ :

$$x = \sum_{\gamma \in \mathcal{I}} \alpha_\gamma \varphi_\gamma, \quad (32)$$

a representation including not only the orthogonal wavelet transform but also over-complete representations. By replacing this representation in equation (31) one obtains a function of  $\alpha$  that is convex (strictly convex, if  $H$  is non-singular and the basis is orthonormal) and has a Lipschitz-continuous gradient.

A non-smooth convex penalty, promoting sparsity in the coefficients  $\alpha_\gamma$ , is then added to the data-fidelity function as well as an additional term ensuring the nonnegativity of the restored image.

Since all conditions required for applying the forward-backward splitting proximal algorithm [41] are satisfied, this optimization method is used for the minimization of the resulting functional. The results derived in [49] demonstrate that sparsity regularization is tractable also in the case of Poisson data.

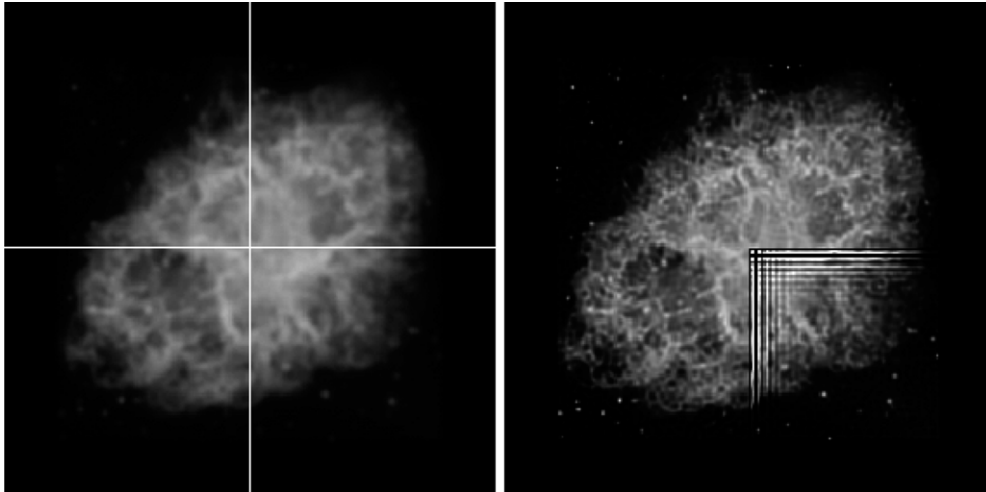
## 6. Other topics

In this section, we briefly discuss other topics that are related to the simple deconvolution problem analyzed in the previous sections.

### 6.1. Boundary effect correction

In the case of an extended object, it may happen that its image extends outside the field of view (FoV) of the detector. In such a case, the FFT-based algorithms imply a periodic continuation of the detected image, and therefore introduce discontinuities at the boundaries of the original image. In the deconvolved image, these discontinuities generate Gibbs oscillations (sometimes called *ripples*), which can propagate inside the image domain, thus degrading completely the quality of the reconstruction.

Two main approaches are used for overcoming this difficulty. The first consists in multiplying the image by a suitable window, so that the windowed image smoothly tends to zero at the boundary, and its periodic continuation is free of discontinuities. The second consists in a suitable continuation of the image outside the FoV, ensuring continuity of the extended image. This is the approach based on the so-called *boundary conditions* and can lead



**Figure 5.** Left panel: the blurred image of the Crab Nebula with the indication of the partition into four quadrants. Right panel: the mosaic obtained by deconvolving three quadrants with the algorithm of equation (33) and the fourth one with standard RL.

to fast inversion algorithms, at least in the case of symmetric PSFs. Different kinds of boundary conditions have been proposed in the literature [7, 101, 113]. Of course, the extension of the image outside the FoV, as provided by these approaches, is somehow artificial.

Another way of looking at the problem is investigated in [28]. It is based on the idea of reconstructing the object in a region broader than the image domain. A very similar approach was proposed in the context of the reconstruction of HST images, and implemented in the Lucy task of STSDAS [121, 134]. Other methods based on the same idea are proposed in [36, 108].

In this approach, we denote by  $\bar{S} \supset S$  the reconstruction domain and we extend  $y$  from  $S$  to  $\bar{S}$  by zero padding; then, in the absence of regularization, one can easily derive the following modification of the RL algorithm [28]:

$$x^{(k+1)} = \frac{x^{(k)}}{\alpha} H^T \frac{y}{Hx^{(k)} + b}, \quad \alpha_i = (H^T M_S)_i, \quad i \in \bar{S}, \quad (33)$$

where  $M_S$  is the characteristic function of the set  $S$ . Division by zero can be avoided by a suitable masking of  $\alpha^{-1}$ .

In figure 5, we demonstrate the accuracy that can be obtained by means of this approach. A blurred image of the Crab nebula, perturbed by Poisson noise, is partitioned into four quadrants, as shown in the left panel. Three of them are deconvolved using the algorithm (33), while the fourth one is deconvolved using standard RL; the full object is obtained as a mosaic of the four reconstructions (500 iterations in all cases). The reconstructions obtained with (33) are essentially free of ripples while these are clearly visible in the reconstruction provided by (17). For more details we refer to [28].

## 6.2. Space-variant PSFs

If the imaging system is not isoplanatic, then it is described by the more general equation (2) and the corresponding discrete matrix  $H$  is full and without structure. As a consequence of the large scale of the images, the numerical treatment of the problem becomes a tremendous

task. It is still feasible if the integral kernel can be locally approximated by a convolution kernel. This problem is denoted as image restoration with a *space-variant PSF* and was already considered for the restoration of HST images [2].

In such a case, one can consider a decomposition, or segmentation, of the image domain such that, in each sub-domain, the imaging matrix can be approximated by a convolution matrix, and the use of some deconvolution method is possible. This approach is investigated, for instance, in [54, 98, 99]. The complete image can be re-obtained as a mosaic of the restored images corresponding to the different sub-domains and, in doing that, it is obvious that boundary effects must be carefully managed.

However, in the case of Poisson noise, it is known from numerical practice that the RL algorithm is robust with respect to small errors on the PSF. Therefore, if the space variance of the PSF over the FoV is not exceedingly large, one can also treat the problem as a standard deconvolution problem, just by taking the PSF corresponding to the central region of the image.

### 6.3. Super-resolution

In astronomy and microscopy, the problem of super-resolution was traditionally intended as the problem of overcoming the diffraction limit by means of a suitable processing of the detected data. In [135], it was observed that, as a consequence of uniqueness of analytic continuation in the Fourier domain, super-resolution is feasible if the object support is compact. However, in [22], it is shown that, as a consequence of the ill-posedness of the problem, a significant super-resolution can be obtained only if the size of the support is of the order of the diffraction limit. Moreover, the knowledge of the support must be inserted in the deconvolution algorithm. A review is given in [24].

An interesting feature of the RL algorithm is that the constraint on the support can easily be treated by a suitable initialization of the algorithm [8, 27]: if  $x^{(0)}$  is zero in one pixel, then all the  $x^{(k)}$  will be zero in that pixel. Therefore, it is sufficient to take as initial guess the characteristic function of the object support. This can be estimated from a preliminary reconstruction of the object obtained with a few RL iterations, initialized, as usual, with a constant array/cube. Such a trick allows the use of FFT in this ‘constrained’ deconvolution, thus providing a fast super-resolving algorithm.

However, we must point out that, as concerns microscopy, viable physical principles have been introduced in recent years for overcoming the limiting role of diffraction. The result is the invention and the production of readily applicable and widely accessible fluorescence microscopes with nanoscale spatial resolution [67]. A quite relevant example is provided by the *stimulated emission depleted microscopy* (STED) [66]. In conclusion, possible applications of super-resolution by data processing are presumably restricted to astronomy.

### 6.4. Blind deconvolution

In some applications the PSF is not known and the so-called methods of *blind deconvolution* (BD) have been developed for extracting both the object and the PSF from the recorded image. Of course, if some approximation of the PSF is available, then this information should be used; in such a case, one speaks of *myopic* or *semi-blind* deconvolution.

One approach consists in minimizing a function of both the object and the PSF, designed to optimize their reconstruction with the addition of suitable constraints (nonnegativity, bounded domain in the physical or Fourier variables, etc) on both the object and the PSF. In general, the minimization is obtained by an iterative method (examples are given in [79, 84]), but, in

the case of Tikhonov regularization, a non-iterative and fast numerical scheme has also been proposed [81]. As far as we know, this kind of BD has been considered only in the case of Gaussian noise, i.e. the data-fidelity term is the usual metric distance, which is not jointly convex in the two variables  $x$  and  $K$ .

Another approach consists of methods where no cost function is given, but standard deconvolution methods, for instance RL, are used to restore the object and the PSF separately in an iterative form: each iteration consists of two cycles where the object or the PSF is kept fixed while the other is updated with a deconvolution algorithm (remember that the convolution product is symmetric with respect to the exchange of the object and PSF). Therefore, the output of one cycle updates the object, if the PSF is kept fixed, or the PSF, if the object is kept fixed: a ‘PSF cycle’ is followed by an ‘object cycle’.

The previous approach was first proposed with the Wiener filter as the deconvolution method [11]. In the case of Poisson statistics, as far as we know, all the proposed methods are based on this approach [32, 53, 70, 129]. The results obtained with the existing algorithms are promising in many instances.

It is interesting to remark that a convergence proof is available in the case where each cycle consists of a single RL iteration, thanks to a result obtained in the problem of nonnegative matrix factorization [91], which is strictly related to BD. As far as we know, no convergence proof is available in other cases (the use of constraints on the PSF, for instance, is very important [46]) and, therefore, the blind deconvolution problem with Poisson data still deserves a rigorous investigation.

### 6.5. Multiple image deconvolution

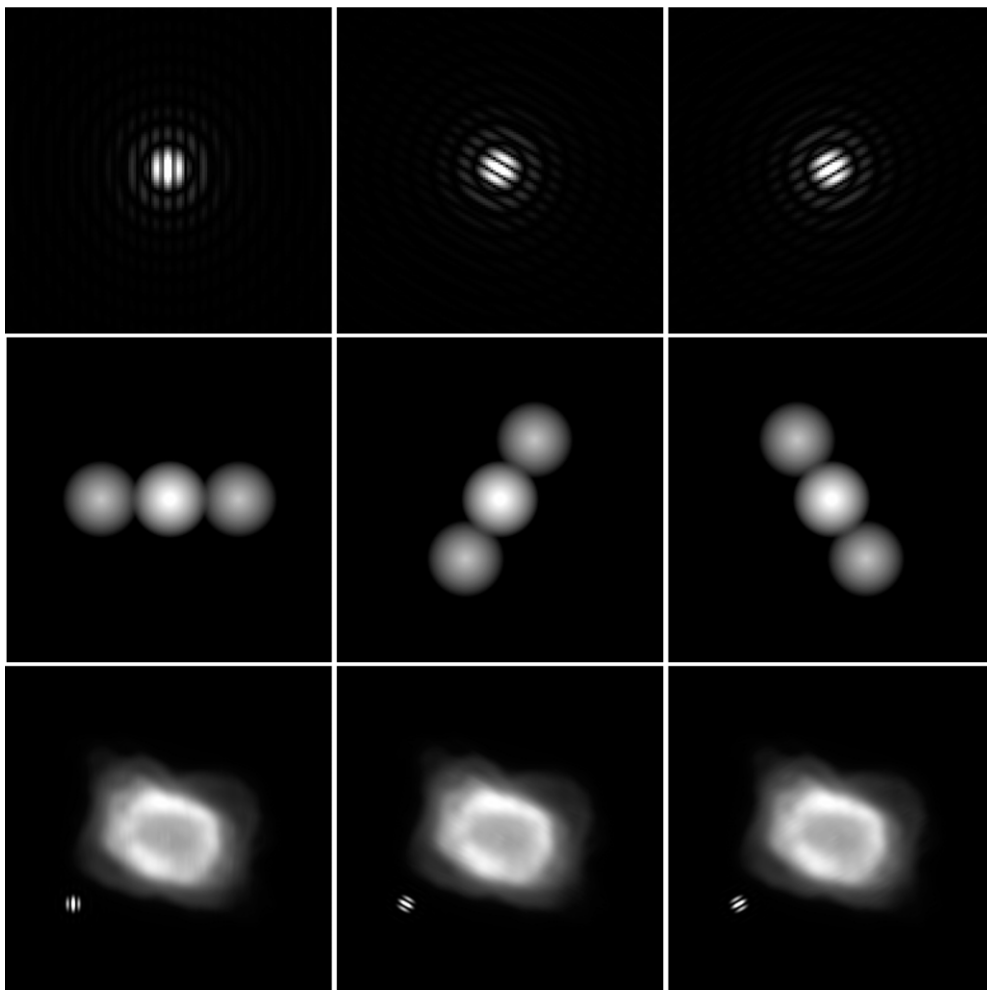
The problem of multiple image deconvolution consists in recovering one object from a set of different images of it obtained with different PSFs. It is obvious that, in general, the problem is still ill-posed even if data are redundant. A discussion of some mathematical results and of their practical relevance is given in [21, 39], while the application of Tikhonov regularization is discussed in [104].

In the case of Poisson data, let  $y^{(1)}, \dots, y^{(p)}$  be different images of the same object  $x$ , corresponding to PSFs  $K^{(1)}, \dots, K^{(p)}$ . If they are statistically independent, then the negative log of the likelihood is the sum of the KL divergences of the different images. In such a case, it is easy to show that the RL algorithm (17) is to be replaced by the following multiple image algorithm [26]:

$$x^{(k+1)} = \frac{1}{p} x^{(k)} \sum_{j=1}^p (H^{(j)})^T \frac{y^{(j)}}{H^{(j)} x^{(k)} + b^{(j)}}, \quad (34)$$

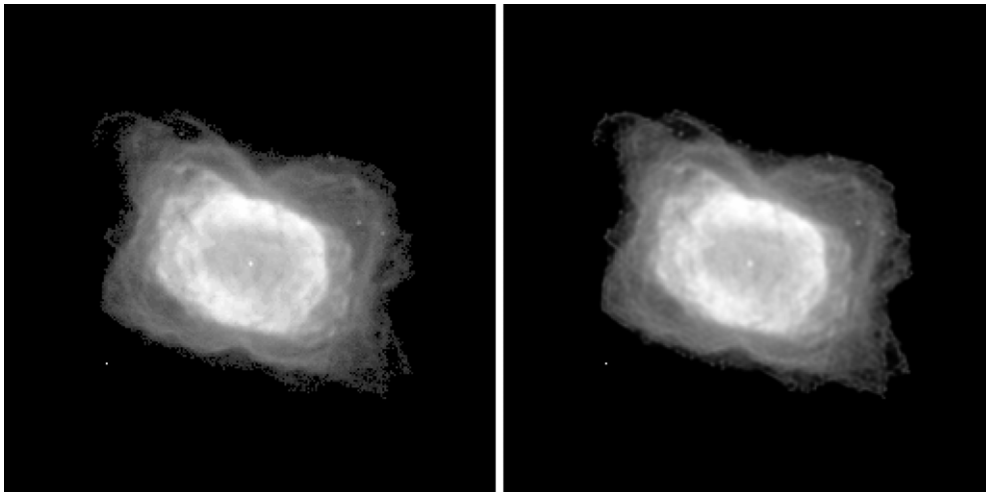
where  $H^{(j)}z = K^{(j)} * z$ , and  $b^{(j)}$  is the background of image  $y^{(j)}$ .

An important application of this method to astronomical imaging is related to the Large Binocular Telescope (LBT). LBT, operating on the top of Mount Graham, Arizona, consists of two 8.4 m mirrors on the same mounting, with a distance of 14.4 m between their centers (<http://medusa.as.arizona.edu/lbto/>). One of the most innovative instruments of LBT is the Fizeau interferometer denoted LINC-NIRVANA (LN, for short), which is in the advanced realization phase by a Consortium of German and Italian institutions (<http://www.mpia.de/LINC>). Installation is foreseen in late 2011. When operating, LN will combine the beams coming from the two mirrors in the focal plane (not in the pupil plane, as with essentially all existing interferometers), thus providing true imaging in the wavelength range 1.0–2.4  $\mu\text{m}$  on a wide FoV (about 10 arcsec).



**Figure 6.** Upper panels: ideal PSFs of the interferometer LINC-NIRVANA corresponding to three equi-spaced orientations of the baseline, with an angular spacing of  $120^\circ$ . Middle panels: the corresponding MTFs. Lower panels: the corresponding images of a nebula.

The PSF of the instrument will be the PSF of a single mirror modulated by interference fringes orthogonal to the direction of the baseline, i.e. the line joining the centers of the two mirrors. It follows that a single LN image is characterized by an anisotropic resolution: ideally, that of a 22.8 m mirror in the direction of the baseline and that of a 8.4 m mirror in the orthogonal direction. However, thanks to the rotation of the earth, it is possible to acquire images corresponding to different orientations of the baseline and this is equivalent to rotate the PSF. In the upper panels of figure 6, we show three ideal PSFs, i.e. without optical and atmospheric aberrations, corresponding to three different orientations of the baseline, and in the middle panels the corresponding MTFs. Therefore, the band of the instrument consists essentially of three discs: the central one corresponds to the 8.4m mirror, while the side discs contain the additional information provided by the interferometer. It is evident that, with a few images, it is possible to obtain a coverage of the Fourier plane corresponding approximately



**Figure 7.** The original object (left) and the reconstruction obtained from the three images of figure 6.

to that of a 22.8 m mirror. The problem of combining the different images for getting a unique high-resolution image, possibly with the resolution of a 22.8 m mirror in all directions, is just a multiple image deconvolution problem.

Since the iteration (34) is very slow, a faster algorithm is derived in [26] by applying to this problem the *ordered subset expectation maximization* (OSEM) method proposed by Hudson and Larkin [73] in emission tomography. According to this method, at each iteration, the sum over the  $p$  images in equation (34) is replaced by a cycle over these images. In such a way, a reduction by a factor  $p$  in the number of iterations is obtained, with approximately the same computational cost per iteration.

In the left panel of figure 7, we show the nebula that has been used for generating the three images in the lower panels of figure 6 and, in the right panel, the reconstruction obtained by means of the OSEM version of the multiple image RL algorithm of equation (34).

## 7. Concluding remarks

Even if research on image deconvolution with Poisson data started about 25 years ago, we believe that several interesting problems are still open and can be the subject of nontrivial mathematical research. Presumably, these problems are not completely known to mathematicians because most papers on the topic are published on journals specialized in the specific applications, astronomy and microscopy in this case. We hope that this paper can provide some help in recognizing these problems.

We mention a few of them: the ill-posedness of the minimization of the KL divergence, the effect of the background on this problem, the design of priors for images with a high dynamic range, the derivation of mathematically sound blind deconvolution methods. Moreover, we believe that some standard topics in the least-square framework, such as edge-preserving and sparsity regularization, still require considerable research to reach a complete understanding in the KL divergence framework.

## Acknowledgments

We acknowledge the support of LAMBS, Università di Genova, for providing the Powermicroscope software we used for deconvolving the microscopy images. We also thank Rebecca Medda of the Department of NanoBiophotonics, Max Planck Institute for Biophysical Chemistry, Goettingen, for the preparation of the mammalian cell sample. Finally, we thank Christine De Mol, Université Libre de Bruxelles, for a critical reading of the manuscript.

## References

- [1] Adorf H-M, Hook R N, Lucy L B and Murtagh F D 1992 Accelerating the Richardson–Lucy restoration algorithm *Proc. 4th ESO/ST-ECF Data Analysis Workshop* pp 99–103
- [2] Adorf H-M 1994 Toward HST restoration with a space-variant PSF, cosmic rays and other missing data *The Restoration of HST Images and Spectra II* (Baltimore: The Space Telescope Science Institute) pp 72–8
- [3] Adorf H-M 1995 Hubble space telescope image restoration in its fourth year *Inverse Problems* **11** 639–53
- [4] Agard D A and Sedat J W 1983 Three-dimensional architecture of a polytene nucleus *Nature* **302** 676–81
- [5] Agard D A 1984 Optical sectioning microscopy in cellular architecture in three dimensions *Ann. Rev. Biophys. Bioeng.* **13** 191–219
- [6] Agard D A, Hiraoki Y and Sedat J W 1989 Three-dimensional microscopy: image processing for high-resolution subcellular imaging *Proc. SPIE* **1161** 24–30
- [7] Aghdasi F and Ward R K 1996 Reduction of boundary artifacts in image restoration *IEEE Trans. Image Process.* **5** 611–8
- [8] Anconelli B, Bertero M, Boccacci P and Carbillet M 2005 Restoration of interferometric images-IV. An algorithm for super-resolution of stellar systems *Astron. Astrophys.* **431** 747–56
- [9] Anscombe F J 1948 The transformation of Poisson, binomial and negative-binomial data *Biometrika* **35** 246–54
- [10] Art J 2006 Photon detectors for confocal microscopy *Handbook of Biological Confocal Microscopy* ed J Pawley (New York: Plenum) pp 251–62
- [11] Ayers G R and Dainty J C 1988 Iterative blind deconvolution method and its applications *Opt. Lett.* **13** 547–9
- [12] Bardsley J M 2006 Stopping rules for a nonnegatively constrained iterative method for ill-posed Poisson imaging problems *BIT Numer. Math.* **46** 651–64
- [13] Bardsley J M and Nagy J G 2006 Covariance-preconditioned iterative methods for nonnegatively constrained astronomical imaging *SIAM J. Matrix Anal. Appl.* **27** 1184–97
- [14] Bardsley J M 2008 An efficient computational method for total-variation penalized Poisson likelihood estimation *Inverse Probl. Imag.* **2** 167–85
- [15] Bardsley J M and Laobeul N 2008 Tikhonov regularized Poisson likelihood estimation: theoretical justification and a computational method *Inverse Probl. Sci. Eng.* **16** 199–215
- [16] Bardsley J M and Luttmann A 2009 Total variation-penalized Poisson likelihood estimation for ill-posed problems *Adv. Comput. Math.* (special issue on *Mathematical Imaging*) **31** 35–59
- [17] Bardsley J M and Laobeul N 2009 An analysis of regularization by diffusion for ill-posed Poisson likelihood estimation *Inverse Problems Sci. Eng.* **17** 537–50
- [18] Barrett H H and Meyers K J 2003 *Foundations of Image Science* (New York: Wiley) pp 1047–8
- [19] Barzilai J and Borwein J M 1988 Two-point step size gradient methods *IMA J. Numer. Anal.* **8** 141–8
- [20] Benvenuto F, La Camera A, Theys C, Ferrari A, Lantéri H and Bertero M 2008 The study of an iterative method for the reconstruction of images corrupted by Poisson and Gaussian noise *Inverse Problems* **24** 035016 (20pp)
- [21] Berenstein C A and Patrick E V 1990 Exact deconvolution for multiple convolution operators—an overview, plus performance characterization for imaging sensors *Proc. IEEE* **78** 723–34
- [22] Bertero M and Pike E R 1982 Resolution in diffraction-limited imaging, a singular value analysis: I. The case of coherent illumination *Opt. Acta* **29** 727–46
- [23] Bertero M, Boccacci P, Brakenhoff G J, Malfanti F and van der Voort H T M 1990 Three-dimensional image restoration and super-resolution in fluorescence confocal microscopy *J. Microsc.* **157** 3–20
- [24] Bertero M and De Mol C 1996 Super-resolution by data inversion *Progress in Optics* vol 36 ed E Wolf chapter III (Amsterdam: North-Holland)
- [25] Bertero M and Boccacci P 1998 *Introduction to Inverse Problems in Imaging* (Bristol: Institute of Physics Publishing)
- [26] Bertero M and Boccacci P 2000 Application of the OS-EM method to the restoration of LBT images *Astron. Astrophys. Suppl. Ser.* **144** 181–6

- [27] Bertero M and Boccacci P 2003 Super-resolution in computational imaging *Micron* **34** 265–73
- [28] Bertero M and Boccacci P 2005 A simple method for the reduction of boundary effects in the Richardson–Lucy approach to image deconvolution *Astron. Astrophys.* **437** 369–74
- [29] Bertero M, Lanteri H and Zanni L 2008 Iterative image reconstruction: a point of view *Mathematical Methods in Biomedical Imaging and Intensity-Modulated Radiation Therapy (IMRT) (CRM Series vol 7)* ed Y Censor, M Jiang and A K Louis (Pisa: Edizioni della Normale) pp 37–63
- [30] Bertsekas D P 1999 *Nonlinear Programming* (Belmont, US: Athena Scientific)
- [31] Biggs D S C and Andrews M 1997 Acceleration of iterative image restoration algorithms *Appl. Opt.* **36** 1766–75
- [32] Biggs D S C and Andrews M 1998 Asymmetric iterative blind deconvolution of multi-frame images *Proc. SPIE* **3461** 328–38
- [33] Booth M J 2007 Adaptive optics in microscopy *Phil. Trans. R. Soc. A* **365** 2828–43
- [34] Bonettini S, Zanella R and Zanni L 2009 A scaled gradient projection method for constrained image deblurring *Inverse Problems* **25** 015002
- [35] Brune C, Sawatzky A and Burger M 2009 Bregman-EM-TV methods with application to optical nanoscopy *Springer's Lecture Notes in Computer Science* vol 5567, ed X-C Tai *et al* (Springer) pp 235–41
- [36] Calvetti D and Somersalo E 2005 Statistical elimination of boundary artifacts in image deblurring *Inverse Problems* **21** 1697–714
- [37] Calvini P and Bertero M 2002 Application of Krylov subspaces to SPECT imaging *Int. J. Imag. Syst. Technol.* **12** 217–28
- [38] Carrington W A, Fogarty K E and Fay F S 1990 3D fluorescence imaging of single cells using image restoration *Noninvasive Techniques in Cell Biology* ed K Foskett and S Grinstein (New York: Wiley-Liss) pp 53–72
- [39] Casey S D and Walnut D F 1995 Systems of convolution equations, deconvolution, Shannon sampling and the wavelet and Gabor transforms *SIAM Rev.* **36** 537–77
- [40] Charbonnier P, Blanc-Féraud L, Aubert G and Barlaud M 1997 Deterministic edge-preserving regularization in computed imaging *IEEE Trans. Image Process.* **6** 298–311
- [41] Combettes P L and Wajs V R 2005 Signal recovery by proximal forward-backward splitting *SIAM Multiscale Model. Simul.* **4** 1168–200
- [42] Conchello J A and McNally J G 1996 Fast regularization technique for expectation maximization algorithm for optical sectioning microscopy *Proc. SPIE* **2655** 199–208
- [43] Conchello J A and Lichtman J W 2005 Optical sectioning microscopy *Nature Methods* **2** 920–31
- [44] Daubechies I, Defrise M and De Mol C 2004 An iterative thresholding algorithm for linear inverse problems with a sparsity constraint *Commun. Pure Appl. Math.* **112** 1413–57
- [45] Dempster A P, Laird N M and Rubin D B 1977 Maximum likelihood from incomplete data via the EM algorithm *J. R. Stat. Soc. B* **39** 1–38
- [46] Desiderà G, Anconelli B, Bertero M, Boccacci P and Carbillet M 2006 Application of iterative blind deconvolution to the reconstruction of LBT LINC-NIRVANA images *Astron. Astrophys.* **452** 727–43
- [47] Dey N, Blanc-Feraud L, Zimmer C, Roux P, Kam Z, Olivo-Marin J-C and Zerubia J 2006 Richardson–Lucy algorithm with total variation regularization for 3D confocal microscope deconvolution *Microsc. Res. Tech.* **69** 260–6
- [48] Diaspro A (ed) 2001 *Confocal and Two-Photon Microscopy: Foundations, Application and Advances* (New York: Wiley-Liss)
- [49] Dupé F -X, Fadili J M and Starck J -L 2009 A proximal iteration for deconvolving Poisson noisy images using sparse representations *IEEE Trans. Image Process.* **18** 310–21
- [50] Ellerbroek B E and Vogel C R 2009 Inverse problems in astronomical adaptive optics *Inverse Problems* **25** 063001 (37 pp)
- [51] Engl H W, Hanke M and Neubauer A 1996 *Regularization of Inverse Problems* (Dordrecht: Kluwer)
- [52] Feller W 1968 *An Introduction to Probability Theory and its Applications* (New York: Wiley)
- [53] Fish D A, Brinicombe A M and Pike E R 1995 Blind deconvolution by means of the Richardson–Lucy algorithm *J. Opt. Soc. Am.* **A-12** 58–65
- [54] Fish D A, Grochmalicki J and Pike E R 1996 A scanning singular value decomposition method for restoration of images with space variant blur *J. Opt. Soc. Am.* **A-13** 464–9
- [55] Formiconi A R, Pupi A and Passeri A 1989 Compensation of spatial system response in SPECT with conjugate gradient reconstruction technique *Phys. Med. Biol.* **34** 69–84
- [56] Frieden B R 1975 Image enhancement and restoration *Picture Processing and Digital Filtering (Topics in Applied Physics vol 6)* ed T S Huang (Berlin: Springer) pp 179–249
- [57] Geman S and Geman D 1984 Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images *IEEE Trans. Pattern Anal. Mach. Intell.* **6** 721–41

- [58] Geman S and McClure DE 1985 Bayesian image analysis: an application to single photon emission tomography *1985 Proc. Statist. Comput. Sect., Amer. Statist. Assoc. (Washington DC)* pp 12-8
- [59] Geman S and McClure DE 1987 Statistical methods for tomographic image reconstruction *Bull. Int. Stat. Inst. LII-4* 5-21
- [60] Goodman J W 1968 *Introduction to Fourier Optics* (New York: McGraw-Hill)
- [61] Green P J 1990 Bayesian reconstructions from emission tomography data using a modified EM algorithm *IEEE Trans. Med. Imaging* **9** 84-93
- [62] Grenander U 1981 *Abstract Inference* (New York: Wiley)
- [63] Grenander U 1984 *Tutorial in Pattern Theory (Lecture Notes Volume)* (Providence RI: Division of Applied Mathematics, Brown University) (unpublished)
- [64] Hanisch R J and White R L (ed) 1994 *The Restoration of HST Images and Spectra-II* (Baltimore: The Space Telescope Science Institute)
- [65] Hansen P C, Nagy J G and O'Leary D P 2006 *Deblurring Images: Matrices, Spectra and Filtering* (Philadelphia: SIAM)
- [66] Hell S W and Wichmann J 1994 Breaking the diffraction resolution limit by stimulated emission: stimulated emission depleted microscopy *Opt. Lett.* **19** 780-2
- [67] Hell S W 2008 Microscopy and its focal switch *Nature Methods* **6** 24-32
- [68] Holmes T J 1988 Maximum-likelihood image restoration adapted for noncoherent optical imaging *J. Opt. Soc. Am. A-5* 666-73
- [69] Holmes T J 1989 Expectation-maximization restoration of band-limited, truncated point-process intensities with application in microscopy *J. Opt. Soc. Am. A-6* 1006-14
- [70] Holmes T J Blind deconvolution of quantum-limited incoherent imagery: and maximum-likelihood approach *J. Opt. Soc. Am. A-9* 1052-61
- [71] Holmes T J and Liu Y H 1989 Richardson-Lucy/maximum likelihood image restoration algorithm for fluorescence microscopy: further testing *Appl. Opt.* **28** 4930-8
- [72] Holmes T J and Liu Y H 1991 Acceleration of maximum-likelihood image restoration for fluorescence microscopy *J. Opt. Soc. Am. A-8* 893-907
- [73] Hudson H M and Larkin R S 1994 Accelerated image reconstruction using ordered subsets of projection data *IEEE Trans. Med. Imaging* **13** 601-9
- [74] Huesman R H, Gullberg G T, Greenberg W L and Budinger T F 1977 User's manual: Donner algorithms for reconstruction tomography (Berkeley, CA: Lawrence Berkeley Laboratory, University of California)
- [75] Iusem A N 1991 Convergence analysis for a multiplicatively relaxed EM algorithm *Math. Methods Appl. Sci.* **14** 573-93
- [76] Iusem A N 1992 A short convergence proof of the EM algorithm for a specific Poisson model *REBRAPE* **6** 57-67
- [77] Jammal G and Bijaoui A 2004 Dequant: a flexible multiresolution restoration framework *Signal Process.* **84** 1049-69
- [78] Jansson P A 1997 *Deconvolution of Images and Spectra* (San Diego: Academic)
- [79] Jefferies S M and Christou J C 1993 Restoration of astronomical images by iterative blind deconvolution *Astrophys. J.* **415** 862-74
- [80] Joshi S and Miller M I 1993 Maximum *a posteriori* estimation with Good's roughness for three-dimensional optical-sectioning microscopy *J. Opt. Soc. Am. A-10* 1078-85
- [81] Justen L and Ramlau R 2006 A non-iterative regularization approach to blind deconvolution *Inverse Problems* **22** 771-800
- [82] Kaipio J and Somersalo E 2005 *Statistical and Computational Inverse Problems* (Berlin: Springer)
- [83] La Camera A, Desiderà G, Arcidiacono C, Boccacci P and Bertero M 2007 Advances in the reconstruction of LBT LINC-NIRVANA images *Astron. Astrophys.* **471** 1091-7
- [84] Lane R G 1992 Blind deconvolution of speckle images *J. Opt. Soc. Am. A-9* 1508-14
- [85] Lange K and Carson R 1984 EM reconstruction algorithms for emission and transmission tomography *J. Comput. Assist. Tomogr.* **8** 306-16
- [86] Lange K 1990 Convergence of EM image reconstruction algorithms with Gibbs smoothing *IEEE Trans. Med. Imaging* **9** 439-46
- [87] Lantéri H, Roche M, Cuevas O and Aime C 2001 A general method to devise maximum-likelihood signal restoration multiplicative algorithms with nonnegativity constraints *Signal Process.* **81** 945-74
- [88] Lantéri H, Roche M and Aime C 2002 Penalized maximum likelihood image restoration with positivity constraints: multiplicative algorithms *Inverse Problems* **18** 1397-419
- [89] Lantéri H, Roche M, Gaucherel P and Aime C 2002 Ringing reduction in image restoration algorithms using a constraint on the inferior bound of the solution *Signal Process.* **82** 1481-504

- [90] Lantéri H and Theys C 2005 Restoration of astrophysical images—the case of Poisson data with additive Gaussian noise *EURASIP J. Appl. Signal Process.* **15** 2500–13
- [91] Lee D D and Seung H s 2000 Algorithms for nonnegative matrix factorization *Adv. Neural Inf. Process.* **13** 556–62
- [92] Lucy L B 1974 An iterative technique for the rectification of observed distributions *Astron. J.* **79** 745–54
- [93] Markham J and Conchello J A 2001 Fast maximum-likelihood image-restoration algorithms for three-dimensional fluorescence microscopy *J. Opt. Soc. Am. A* **18** 1062–71
- [94] de Monvel J B, Le Calvez S and Ulfendahl M 2001 Image restoration for confocal microscopy: improving the limits of deconvolution, with application to the visualization of the mammalian hearing organ *Biophys. J.* **80** 2455–70
- [95] Mühlthei H N and Schorr B 1987 On an iterative method for a class of integral equations of the first kind *Math. Methods Appl. Sci.* **9** 137–68
- [96] Mühlthei H N and Schorr B 1989 On properties of the iterative maximum likelihood reconstruction method *Math. Methods Appl. Sci.* **11** 331–42
- [97] Mühlthei H N 1993 Iterative continuous maximum likelihood reconstruction methods *Math. Methods Appl. Sci.* **15** 275–86
- [98] Nagy J G, Pauca V P, Plemmons R J and Torgersen T C 1997 Space-varying restoration of optical images *J. Opt. Soc. Am. A* **14** 3162–74
- [99] Nagy J G and O’Leary D P 1998 Restoring images degraded by spatially-variant blur *SIAM J. Sci. Comput.* **19** 1063–82
- [100] Natterer F and Wübbeling F 2001 *Mathematical Methods in Image Reconstruction* (Philadelphia: SIAM)
- [101] Ng M K, Chan R H and Tang W C 1999 A fast algorithm for deblurring models with Neumann boundary conditions *SIAM J. Sci. Comput.* **21** 851–66
- [102] Nowak R and Kolaczyk E D 2000 A statistical multiscale framework for Poisson inverse problems *IEEE Trans. Inf. Theory* **46** 2794–802
- [103] Pawley J 2005 *Handbook of Biological Confocal Microscopy* (New York: Plenum)
- [104] Piana M and Bertero M 1996 Regularized deconvolution of multiple images of the same object *J. Opt. Soc. Am. A* **13** 1516–23
- [105] Politte D G and Snyder D L 1991 Correction for accidental coincidences and attenuation in maximum-likelihood image reconstruction for positron-emission tomography *IEEE Trans. Med. Imaging* **10** 82–9
- [106] Puetter R C, Gosnell T R and Yahil A 2005 Digital image reconstruction: deblurring and denoising *Annu. Rev. Astron. Astrophys.* **43** 139–94
- [107] Resmerita E, Engl H W and Iusem N 2007 The expectation-maximization algorithm for ill-posed integral equations: a convergence analysis *Inverse Problems* **23** 2575–88
- [108] Reeves S J 2005 Fast image restoration without boundary artifacts *IEEE Trans. Image Process.* **14** 1448–53
- [109] Richardson W H 1972 Bayesian based iterative method of image restoration *J. Opt. Soc. Am.* **62** 55–9
- [110] Rooms F, Philips W and Lidke D S 2005 Simultaneous degradation estimation and restoration of confocal images and performance evaluation by colocalization analysis *J. Microsc.* **218** 22–36
- [111] Sarder P and Nehorai A Deconvolution methods for 3-D fluorescence microscopy images *IEEE Signal Process. Mag.* **23** 32–45
- [112] Schrader M, Hell S W and van der Voort H T M 1998 Three-dimensional super-resolution with a 4Pi-confocal microscope using image restoration *J. Appl. Phys.* **84** 4033–42
- [113] Serra Capizzano S 2003 A note on anti-reflective boundary conditions and fast deblurring models *SIAM J. Sci. Comput.* **25** 1307–25
- [114] Shepp L A and Vardi Y 1982 Maximum likelihood reconstruction for emission tomography *Trans. Med. Imaging* **MI-1** 113–22
- [115] Simoncelli E, Freeman W, Adelson E and Heeger D 1992 Shiftable multi-scale transforms *IEEE Trans. Inf. Theory* **38** 587–607
- [116] Snyder D L and Miller M I 1985 The use of sieves to stabilize images produced with the EM algorithm for emission tomography *IEEE Trans. Nucl. Sci.* **32** 3864–72
- [117] Snyder D L 1990 Modifications of the Lucy–richardson iteration for restoring Hubble Space-Telescope imagery *The Restoration of HST Images and Spectra* ed R L White and R J Allen (Baltimore: The Space Telescope Science Institute) pp 56–61
- [118] Snyder D L and Miller M I 1991 *Random Point Processes in Time and Space* (New York: Springer)
- [119] Snyder D L, Hammoud A M and White R L 1993 Image recovery from data acquired with a charge-coupled-device camera *J. Opt. Soc. Am.* **10** 1014–23
- [120] Snyder D L, Helstrom C W, Lanterman A D, Faisal M and White R L 1994 Compensation for read-out noise in HST image restoration *The Restoration of HST Images and Spectra II* ed R J Hanish and R L White (Baltimore: The Space Telescope Science Institute) pp 139–54

- [121] Stobie E B, Hanish R J and White R L 1994 Implementation of the Richardson–Lucy algorithm in STSDAS *Astron. Soc. Pac. Conf. Ser.* **61** 296–9
- [122] Starck J-L and Murtagh F 1994 Image restoration with noise suppression using the wavelet transform *Astron. Astrophys.* **288** 343–8
- [123] Starck J-L, Bijaoui A and Murtagh F 1995 Multiresolution support applied to image filtering and deconvolution *CVGIP: Graph. Models Image Process.* **57** 420–31
- [124] Stark J-L, Murtagh F and Bijaoui A 1998 *Image Processing and Data Analysis: The Multiscale Approach* (Cambridge: Cambridge University Press)
- [125] Starck J-L, Pantin E and Murtagh F 2002 Deconvolution in Astronomy: a review *Publ. Astron. Soc. Pac.* **114** 1051–69
- [126] Starck J-L, Nguyen M K and Murtagh F 2003 Wavelets and curvelets for image deconvolution: a combined approach *Signal Process.* **83** 2279–83
- [127] Stark J-L and Murtagh F 2006 *Astronomical Image and Data Analysis* (Berlin: Springer)
- [128] Tikhonov A N and Arsenin V Y 1977 *Solution of Ill-Posed Problems* (New York: Wiley) (Russian edition in 1974)
- [129] Tsumuraya F, Miura N and Baba N 1994 Iterative blind deconvolution method using Lucy’s algorithm *Astron. Astrophys.* **282** 699–708
- [130] Vardi Y, Shepp L A and Kaufman L 1985 A statistical model for positron emission tomography *J. Am. Stat. Ass.* **80** 8–37
- [131] Vicidomini G, Boccacci P, Diaspro A and Bertero B 2009 Application of the split-gradient method to 3D image deconvolution in fluorescence microscopy *J. Microsc.* **234** 47–61
- [132] Vogel C R 2002 *Computational Methods for Inverse Problems* (Philadelphia: SIAM)
- [133] White R L and Allen R J (ed) 1991 *The Restoration of HST Images and Spectra* (Baltimore: The Space Telescope Science Institute)
- [134] White R L 1993 Improvements to the Richardson–Lucy method *News. STScI’s Image Restoration Proj.* **1** 11–9
- [135] Wolter H 1961 On basic analogies and principal differences between optical and electronic information *Progress in Optics* vol 1, ed E Wolf, chapter V
- [136] Young S S, Driggers R G and Jacobs E L 2008 *Image Deblurring* (Boston, MA: Artech House Publishers)
- [137] Zanella R, Boccacci P, Zanni L and Bertero M 2009 Efficient gradient projection methods for edge-preserving removal of Poisson noise *Inverse Problems* **25** 045010 (24pp)
- [138] Zhang B, Fadili J M and Starck J-L 2008 Wavelets, ridgelets, and curvelets for Poisson noise removal *IEEE Trans. Image Process.* **17** 1093–108
- [139] Zhu D, Razaz M and Lee R 2005 Adaptive penalty likelihood for reconstruction of multi-dimensional confocal microscopy images *Comput. Med. Imaging Graph.* **29** 319–31