

The Large Binocular Telescope: Challenging Problems in Image Restoration

M. Bertero, B. Anconelli, P. Boccacci
DISI - Università di Genova
Via Dodecaneso 35, I-16146 Genova, Italy
surname@disi.unige.it

M. Carillet
INAF-Osservatorio Astrofisico di Arcetri
Largo E. Fermi 5, 50125 Firenze, Italy *marcel@arcetri.astro.it*

H. Lanteri, C. Theys
UMR 6525 Astrophysique, Université de Nice - Sophia Antipolis
Parc Valrose, 06108 Nice Cedex 02, France
Henri.Lanteri@unice.fr, Celine.Theys@unice.fr

Abstract

The Large Binocular Telescope (LBT), which is currently under construction on the top of Mount Graham in Arizona, will be a Fizeau interferometer with a maximum baseline of 22.8m. From the point of view of image reconstruction, the interesting feature of this instrument is that it will require routinely the use of multiple-images deconvolution methods if the aim is to produce a unique image with the resolution of a 22.8m telescope. Our group is working on this problem since a few years and has already produced methods and software for the data reduction problem. In this paper we provide a survey of our ongoing work.

Keywords: Interferometric images, Multiple-images deconvolution, Iterative methods, Richardson-Lucy methods

1. INTRODUCTION

The Large Binocular Telescope (LBT) will consist of two 8.4m mirrors on a common mount, with a spacing of 14.4m between their centres, so that a maximum baseline of 22.8m will be available. LBT will be equipped with a Fizeau interferometer, denoted as LINC/NIRVANA, (LN) which will be the result of a German-Italian collaboration. First light of LBT is scheduled for September 2004, second light for November 2005, and the first LN light is foreseen for July 2006.

The interferometric technique used in LN will provide direct imaging with the resolution of a 22.8m telescope in the direction of the baseline and of a 8.4m telescope in the orthogonal direction. Since resolution is not uniform over the field, several images of the same scientific object must be acquired with different orientations of the baseline and they must be suitably processed in order to get a unique image with a uniform resolution over the field. Therefore imaging by LN will require routinely the use of multiple-images deconvolution methods if the goal is a unique image with the resolution of a 22.8m telescope.

According to our experience in image reconstruction, two different kinds of methods will be required:

- *quick-look* methods, computationally efficient even if not very accurate for any kind of astronomical object; they can be routinely used for a preliminary view of the target just after data acquisition.

- *ad hoc* methods designed for particular classes of astronomical objects; each one of them must be accurate as far as possible for the objects in the corresponding class, even if it does not need to be computationally efficient.

Therefore the processing of LBT images will require, in general, a two-steps approach: the first step, based on a quick-look method, will be intended to understand qualitatively specific features of the particular object the astronomer is investigating; the second one will consist in the use of the *ad hoc* method which has been designed for the class of objects with those specific features.

Our group has already proposed possible quick-look methods (Bertero and Boccacci [1, 2], Anconelli et al. [3]); they have been implemented in the Software Package AIRY (Astronomical Image Restoration in interferometrY), which is described in Correia et al. [4]. AIRY is a set of modules for the simulation and/or analysis of interferometric images; it is written in IDL and is conceived to be used with the CAOS (Code for Adaptive Optics Systems) Application Builder (see <http://www.arcetri.astro.it/caos>). Validation of the implemented deconvolution methods is described in Correia et al. [4] and Carbillet et al [5]. The investigation of *ad hoc* methods is in progress. One of them is proposed in Anconelli et al. [6].

The paper is organized as follows. In Section 2 we describe a general approach to the design of iterative methods for constrained deconvolution. In Section 3 we apply the approach to the derivation of fast RL-like methods (RL=Richardson-Lucy) which are candidates to be quick-look methods. The same approach is used in Section 4 for deriving three *ad hoc* methods we are presently investigating: the first one is for objects with high dynamic range; the second one is for faint objects if both photon and read-out noise must be taken into account; the third one is for the super-resolution of unresolved stellar objects. Finally in Section 5 we briefly discuss the state-of-the art of our activity.

2. THE GENERAL APPROACH

In the following we use bold letters for $N \times N$ arrays. Therefore we denote by $\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_p$ the p LN images of a given astronomical object, corresponding to p different orientations of the LBT baseline; moreover we often use the more compact notation \mathbf{g} for indicating the set of the p images. Each one of these images, acquired with a CCD camera, can be described by the following model (Snyder et al. [7]):

$$\mathbf{g}_j(m, n) = \mathbf{g}_{obj,j}(m, n) + \mathbf{g}_{back,j}(m, n) + \mathbf{r}_j(m, n) \quad , \quad j = 1, 2, \dots, p \quad , \quad (1)$$

where $\mathbf{g}_j(m, n)$ denotes the value of the detected j -th image at pixel m, n ; it is given by the sum of the number of photoelectrons due to both object radiation, $\mathbf{g}_{obj,j}(m, n)$, and background radiation, $\mathbf{g}_{back,j}(m, n)$, with the addition of the read-out noise due to the amplifier, $\mathbf{r}_j(m, n)$. The first two terms are realizations of independent Poisson processes, so that their sum is also a Poisson process, while the third term is the realization of a Gaussian process (white noise). The expected value of the j -th image is given by:

$$E\{\mathbf{g}_j\} = A_j \mathbf{f} + \mathbf{b}_j \quad , \quad (2)$$

where: \mathbf{f} is the unknown object, an array formed by the average numbers of photons emitted at the pixels of the object domain and collected by the telescope; A_j is the matrix describing the imaging process (optics plus atmosphere); \mathbf{b}_j is a background term, which can include a non-zero expected value of the read-out noise and, in general, can be assumed to be constant over the image domain. If we assume, for simplicity, space invariance, the matrix A_j is given in terms of the point spread function (PSF) \mathbf{K}_j of the instrument, corresponding to the j -th orientation of the baseline:

$$A_j \mathbf{f} = \mathbf{K}_j * \mathbf{f} \quad ; \quad (3)$$

moreover we assume, as usual, that the PSFs are normalized in such a way that the sum of all their pixel values is one:

$$\sum_{m,n=0}^{N-1} \mathbf{K}_j(m, n) = 1 \quad ; \quad j = 1, 2, \dots, p \quad . \quad (4)$$

The problem of image restoration is the problem of estimating the unknown array \mathbf{f} , being given the detected images \mathbf{g}_j , the backgrounds \mathbf{b}_j and the PSFs \mathbf{K}_j . It is well known that it can be reduced to the minimization of suitable cost functionals in the framework of approaches such as regularization theory, maximum likelihood or Bayesian estimation (see, for instance, Bertero and Boccacci [8] for an introduction). In the literature there exists a great variety of cost functionals and of iterative algorithms for their minimization, with the possible addition of suitable constraints (for instance, non-negativity of the solution). All these methods have been designed for the deconvolution of a single image and in some cases their extension to the multiple-images case is not trivial. Certainly the software developed for these methods cannot be directly applied to multiple images.

In a series of recent papers Lanteri et al. [9, 10, 11] propose a general approach, denoted as SGM (Split Gradient Method), to the design of iterative methods for the minimization of functionals of the following type:

$$J_\mu(\mathbf{f}; \mathbf{g}) = J_0(\mathbf{f}; \mathbf{g}) + \mu J_R(\mathbf{f}) , \quad (5)$$

with the additional constraints:

$$\mathbf{f}(m, n) \geq 0 \quad , \quad \sum_{m, n=0}^{N-1} \mathbf{f}(m, n) = c . \quad (6)$$

In these equations $J_0(\mathbf{f}; \mathbf{g})$ is the functional measuring the discrepancy between the computed images associated to \mathbf{f} and the detected images \mathbf{g} , and $J_R(\mathbf{f})$ is a regularization functional while μ is the regularization parameter. As concerns the constraints, the first one is obvious while the second one is a constraint on the total flux of the object. The choice of c will be discussed in Section 3.

The basic idea in the approach of Lanteri et al. relies on the following decomposition of the gradient of the functional $J_\mu(\mathbf{f})$:

$$-\nabla J_\mu(\mathbf{f}; \mathbf{g}) = \mathbf{U}_\mu(\mathbf{f}; \mathbf{g}) - \mathbf{V}_\mu(\mathbf{f}; \mathbf{g}) , \quad (7)$$

where $\mathbf{U}_\mu(\mathbf{f}; \mathbf{g})$ and $\mathbf{V}_\mu(\mathbf{f}; \mathbf{g})$ are positive arrays depending on \mathbf{f} . Obviously such a decomposition always exists and is not unique. The applicability of the method requires explicit expressions for the dependence of these arrays on \mathbf{f} . Then the general structure of the iterative algorithm, as described in Lanteri et al. [9], is as follows:

- choose an initial $\mathbf{f}^{(0)}$, satisfying the constraints of Eq. 6 (in general, a uniform image);
- given $\mathbf{f}^{(k)}$ compute:

$$\tilde{\mathbf{f}}^{(k+1)} = \mathbf{f}^{(k)} + \alpha \mathbf{f}^{(k)} \frac{\mathbf{U}_\mu(\mathbf{f}^{(k)}; \mathbf{g})^\omega - \mathbf{V}_\mu(\mathbf{f}^{(k)}; \mathbf{g})^\omega}{\mathbf{V}_\mu(\mathbf{f}^{(k)}; \mathbf{g})^\omega} , \quad \tilde{c}^{(k+1)} = \sum_{m, n=0}^{N-1} \tilde{\mathbf{f}}^{(k+1)}(m, n) ; \quad (8)$$

- set:

$$\mathbf{f}^{(k+1)} = \frac{c}{\tilde{c}^{(k+1)}} \tilde{\mathbf{f}}^{(k+1)} . \quad (9)$$

In Eq. 8 products and quotients of arrays are defined as pixel-by-pixel products and quotients while α and ω are relaxation parameters which, in principle, can be changed at each iteration. α is the step in the descent direction (which is modified by ω if $\omega > 1$), and it can be computed in order to guarantee both the non-negativity of the iterates and the convergence of the iterations. On the other hand ω is a parameter which can allow a reduction of the number of iterations. We will call it the *acceleration exponent*. Indeed, as discussed by Lanteri et al. [9], an integer value of $\omega > 1$ can speed-up the convergence of the iterations. As already remarked, at the very first iterations ω modifies the descent direction; but, when the iterations are close to convergence, the algorithm with $\omega > 1$ is approximately equivalent to the algorithm with $\omega = 1$ and step-size $\alpha \omega$.

The algorithm takes a very simple form if we choose a unit step (i. e. $\alpha = 1$); in such a case we obtain:

- given $\mathbf{f}^{(k)}$ compute:

$$\tilde{\mathbf{f}}^{(k+1)} = \mathbf{f}^{(k)} \left(\frac{\mathbf{U}_\mu(\mathbf{f}^{(k)}; \mathbf{g})}{\mathbf{V}_\mu(\mathbf{f}^{(k)}; \mathbf{g})} \right)^\omega, \quad \tilde{c}^{(k+1)} = \sum_{m,n=0}^{N-1} \tilde{\mathbf{f}}^{(k+1)}(m, n); \quad (10)$$

- set:

$$\mathbf{f}^{(k+1)} = \frac{c}{\tilde{c}^{(k+1)}} \tilde{\mathbf{f}}^{(k+1)}. \quad (11)$$

The convergence of this algorithm is not guaranteed in general, even if it can be proved in some particular cases; nevertheless it has been verified experimentally in all the applications we have considered. The most interesting feature of the algorithm is that *non-negativity of the iterates is automatically satisfied*, as one can easily verify.

In the case of this simplified algorithm the dependence on the regularization parameter can be made explicit by introducing decompositions of the gradients of the functionals $J_0(\mathbf{f}; \mathbf{g})$ and $J_R(\mathbf{f})$ similar to that of Eq. 7:

$$-\nabla J_0(\mathbf{f}; \mathbf{g}) = \mathbf{U}_0(\mathbf{f}, \mathbf{g}) - \mathbf{V}_0(\mathbf{f}; \mathbf{g}), \quad (12)$$

$$-\nabla J_R(\mathbf{f}) = \mathbf{U}_R(\mathbf{f}) - \mathbf{V}_R(\mathbf{f}); \quad (13)$$

then it turns out that the algorithm has the following structure:

- given $\mathbf{f}^{(k)}$ compute:

$$\tilde{\mathbf{f}}^{(k+1)} = \mathbf{f}^{(k)} \left(\frac{\mathbf{U}_0(\mathbf{f}^{(k)}; \mathbf{g}) + \mu \mathbf{U}_R(\mathbf{f}^{(k)})}{\mathbf{V}_0(\mathbf{f}^{(k)}; \mathbf{g}) + \mu \mathbf{V}_R(\mathbf{f}^{(k)})} \right)^\omega, \quad \tilde{c}^{(k+1)} = \sum_{m,n=0}^{N-1} \tilde{\mathbf{f}}^{(k+1)}(m, n); \quad (14)$$

- set:

$$\mathbf{f}^{(k+1)} = \frac{c}{\tilde{c}^{(k+1)}} \tilde{\mathbf{f}}^{(k+1)}. \quad (15)$$

which shows the very simple dependence of the algorithm on the regularization parameter μ .

We apply now the previous approach to the case of multiple images. In such a case the functional $J_0(\mathbf{f}; \mathbf{g})$, which measures the discrepancy between the set of the detected images \mathbf{g} and the set of the computed images associated to \mathbf{f} , is, in general, the sum of the functionals corresponding to the different images. In a statistical approach this property follows from the statistical independence of the different images. Therefore we have:

$$J_0(\mathbf{f}; \mathbf{g}) = \sum_{j=1}^p J_{0,j}(\mathbf{f}; \mathbf{g}_j), \quad (16)$$

and the corresponding decompositions of the functions $\mathbf{U}_0(\mathbf{f}; \mathbf{g})$ and $\mathbf{V}_0(\mathbf{f}; \mathbf{g})$ follow from the equation:

$$-\nabla J_0(\mathbf{f}) = \sum_{j=1}^p \mathbf{U}_{0,j}(\mathbf{f}; \mathbf{g}_j) - \sum_{j=1}^p \mathbf{V}_{0,j}(\mathbf{f}, \mathbf{g}_j). \quad (17)$$

If we insert these expression in Eq. 14 and if we remark that the gradient of the regularization term does not depend on j , so that it can be written as the sum of p identical contributions, we find that the complete algorithm takes the following form:

- given $\mathbf{f}^{(k)}$ compute:

$$\tilde{\mathbf{f}}^{(k+1)} = \mathbf{f}^{(k)} \left(\frac{\sum_{j=1}^p \{ \mathbf{U}_{0,j}(\mathbf{f}^{(k)}; \mathbf{g}_j) + \frac{\mu}{p} \mathbf{U}_R(\mathbf{f}^{(k)}) \}}{\sum_{j=1}^p \{ \mathbf{V}_{0,j}(\mathbf{f}^{(k)}; \mathbf{g}_j) + \frac{\mu}{p} \mathbf{V}_R(\mathbf{f}^{(k)}) \}} \right)^\omega, \quad (18)$$

$$\tilde{c}^{(k+1)} = \sum_{m,n=0}^{N-1} \tilde{\mathbf{f}}^{(k+1)}(m, n);$$

	$\mathbf{U}_{0,j}(\mathbf{f}; \mathbf{g}_j)$	$\mathbf{V}_{0,j}(\mathbf{f}; \mathbf{g}_j)$
<i>Poisson</i>	$A_j^T \frac{\mathbf{g}_j}{A_j \mathbf{f} + \mathbf{b}_j}$	$\mathbf{1}$
<i>Gauss</i>	$A_j^T \mathbf{g}_j$	$A_j^T A_j \mathbf{f} + \mathbf{b}_j$

TABLE 1: The functions $\mathbf{U}_{0,j}(\mathbf{f}^{(k)}; \mathbf{g}_j)$ and $\mathbf{V}_{0,j}(\mathbf{f}^{(k)}; \mathbf{g}_j)$ for the two statistics considered in the text. We denote by $\mathbf{1}$ the array whose entries are all equal to 1.

- set:

$$\mathbf{f}^{(k+1)} = \frac{c}{\tilde{c}^{(k+1)}} \tilde{\mathbf{f}}^{(k+1)} . \quad (19)$$

Eq. 18 shows the dependence on both the regularization parameter and the detected images.

The interesting feature of this algorithm is that the numerator and the denominator in Eq. 18 are given by a sum of terms which are just those appearing in the corresponding algorithm for single-image deconvolution. This structure suggests a cyclic-iterative algorithm which is analogous to the OS-EM algorithm proposed by Hudson & Larkin [12] in the case of tomography:

- given $\mathbf{f}^{(k)}$, set $j = (k + 1) \bmod p$ and compute:

$$\tilde{\mathbf{f}}^{(k+1)} = \mathbf{f}^{(k)} \left(\frac{\mathbf{U}_{0,j}(\mathbf{f}^{(k)}; \mathbf{g}_j) + \frac{\mu}{p} \mathbf{U}_R(\mathbf{f}^{(k)})}{\mathbf{V}_{0,j}(\mathbf{f}^{(k)}; \mathbf{g}_j) + \frac{\mu}{p} \mathbf{V}_R(\mathbf{f}^{(k)})} \right)^\omega ; \quad \tilde{c}^{(k+1)} = \sum_{m,n=0}^{N-1} \tilde{\mathbf{f}}^{(k+1)}(m, n) ; \quad (20)$$

- set:

$$\mathbf{f}^{(k+1)} = \frac{c}{\tilde{c}^{(k+1)}} \tilde{\mathbf{f}}^{(k+1)} . \quad (21)$$

We will call this algorithm the OS version of the multiple-images algorithm of Eq. 18. It is obvious that the computational cost of one of its iterations is just that of one iteration of the same algorithm in the case of a single image. Therefore, if the number of iterations of the algorithms of Eq. 18 and Eq. 20 is approximately the same (a property which has been verified experimentally in several cases), it follows that the algorithm of Eq. 20 provides a computational gain by a factor which is approximately p . Finally it is obvious that the algorithm for single-image deconvolution is obtained by just taking the same image in each iteration.

We conclude this Section by showing the application of the previous approach to a few particular cases. In particular we consider two different kinds of noise statistics, leading to different discrepancy functionals:

- *Poisson Statistics* (photon counting) - In this case the functionals $J_{0,j}(\mathbf{f}; \mathbf{g}_j)$ are given by:

$$J_{0,j}(\mathbf{f}; \mathbf{g}_j) = \sum_{m,n=0}^{N-1} \left\{ \mathbf{g}_j(m, n) \ln \frac{\mathbf{g}_j(m, n)}{(A_j \mathbf{f})(m, n) + \mathbf{b}_j} + [(A_j \mathbf{f})(m, n) + \mathbf{b}_j - \mathbf{g}_j(m, n)] \right\} ; \quad (22)$$

- *Gauss statistics* (white noise) - In this case the functionals $J_{0,j}(\mathbf{f}; \mathbf{g}_j)$ are given by:

$$J_{0,j}(\mathbf{f}; \mathbf{g}_j) = \sum_{m,n=0}^{N-1} |(A_j \mathbf{f})(m, n) + \mathbf{b}_j - \mathbf{g}_j(m, n)|^2 . \quad (23)$$

In Table 1 we give the expressions of the arrays $\mathbf{U}_{0,j}(\mathbf{f}^{(k)}; \mathbf{g}_j)$ and $\mathbf{V}_{0,j}(\mathbf{f}^{(k)}; \mathbf{g}_j)$ for the two cases. In deriving these formulas we used the normalization of the PSF, as given in Eq. 4, which implies that $A_j^T \mathbf{1} = \mathbf{1}$ and $A_j^T \mathbf{b}_j = \mathbf{b}_j$.

If we insert these expressions into Eq. 18 with $\mu = 0$ we obtain respectively the multi-image extensions of RL and ISRA (Iterative Space Reconstruction Algorithm), as proposed in Bertero & Boccacci [1, 2]. Moreover, by inserting the expressions for the Poisson case into Eq. 20, with $\omega = 1$, we re-obtain the OS-EM method proposed in Bertero & Boccacci [1] for accelerated multi-images deconvolution. If $\omega > 1$ we obtain the further accelerated version of OS-EM investigated in Anconelli et al. [6]. This point will be reconsidered in the next Section.

Let us emphasize the fact that the most interesting feature of the approach proposed in this paper is that one can easily insert regularization terms, obtaining new iterative algorithms which can be easily implemented. We mention, among others, Tikhonov regularization and Maximum Entropy both for Gauss and Poisson noise (with the addition of background terms, non-negativity and flux constraint). In Table 2 we give the expressions of the arrays $\mathbf{U}_R(\mathbf{f})$ and $\mathbf{V}_R(\mathbf{f})$ for the following regularization functionals:

- *Tikhonov*

$$J_R(\mathbf{f}) = \frac{1}{2} \sum_{m,n=0}^{N-1} |\mathbf{f}(m,n) - \mathbf{f}^{(R)}(m,n)|^2, \quad (24)$$

where $\mathbf{f}^{(R)}(m,n)$ denotes a reference object (in general, a constant array; very often, zero).

- *Smoothing*

$$J_R(\mathbf{f}) = \frac{1}{2} \sum_{m,n=0}^{N-1} |\mathbf{f}(m,n) - (D\mathbf{f})(m,n)|^2, \quad (25)$$

where D is a matrix with non-negative entries; such a form applies, for instance, to regularization in terms of discrete Laplacian, which can just be written in the form $\Delta = 4(I - D)$ (I being the identity matrix).

- *Entropy*

$$J_R(\mathbf{f}) = \sum_{m,n=0}^{N-1} \mathbf{f}(m,n) \ln \frac{\mathbf{f}(m,n)}{\mathbf{f}^{(R)}(m,n)}, \quad (26)$$

where, again, $\mathbf{f}^{(R)}(m,n)$ denotes a reference object.

$J_R(f)$	$\mathbf{U}_R(\mathbf{f})$	$\mathbf{V}_R(\mathbf{f})$
<i>Tikhonov</i>	$\mathbf{f}^{(R)}$	\mathbf{f}
<i>Smoothing</i>	$(D + D^T)\mathbf{f}$	$(I + D^T D)\mathbf{f}$
<i>Entropy</i>	$-\ln(\frac{1}{c}\mathbf{f})$	$1 - \ln(\frac{1}{c}\mathbf{f}^{(R)})$

TABLE 2: The functions $\mathbf{U}_R(\mathbf{f})$ and $\mathbf{V}_R(\mathbf{f})$ in the case of Tikhonov, Smoothing and Entropy regularization. In the case of Entropy the log of an array is defined pixel-by-pixel; moreover the flux normalization is taken into account, c being the flux constant defined in Eq. 32.

If we combine the arrays of Table 1 and Table 2 we obtain several iterative algorithms for different kinds of regularization both in the Poisson and in the Gauss case.

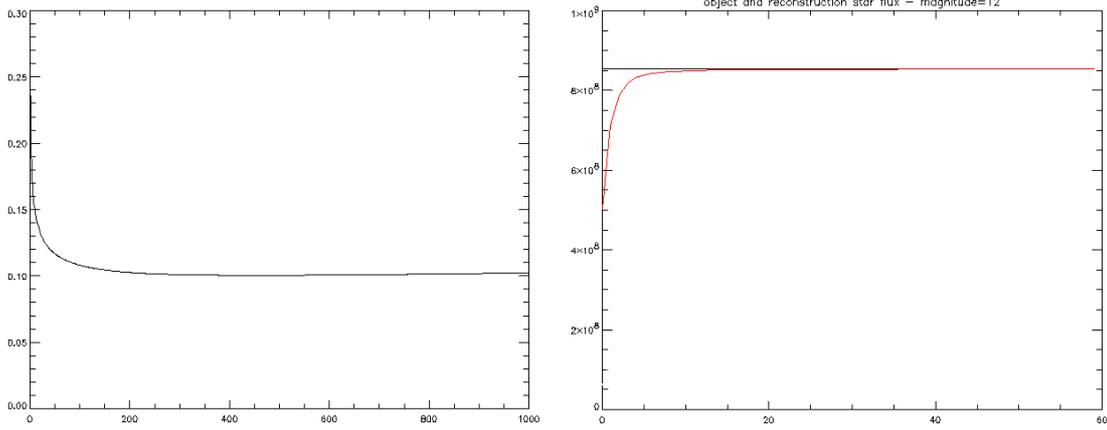


FIGURE 1: Illustrating the convergence of the RL-algorithm: to the left, the plot of the behaviour of the reconstruction error, defined in Eq. 27, as a function of the number of iterations, in the case of a diffuse object; to the right, the plot of the behaviour of the flux of the reconstructed image in the case of a binary system.

3. QUICK-LOOK METHODS

As it is well-known the RL method is one of the most frequently used methods for image deconvolution in Astronomy. It has many interesting features:

- it is an iterative method which maximizes the likelihood function in the case of Poisson statistics;
- it provides automatically nonnegative objects;
- in the case of zero background it provides objects whose total flux coincides with that of the detected image.

The main difficulties are generated by the slow convergence of the algorithm and to the fact that, in some cases, iterations must not be pushed to convergence. For instance, in the case of a diffuse object, such as a nebula or a galaxy, the iterations first provide better and better approximations of the object but, after a certain point, the so-called *checkerboard* effect appears, which is basically due to amplified noise propagation. More precisely, in the case of numerical simulations, let us define, at each iteration k , the following relative restoration error $\rho_{rel}^{(k)}$.

$$\rho_{rel}^{(k)} = \frac{\|\mathbf{f}^{(k)} - \mathbf{f}\|}{\|\mathbf{f}\|}, \quad (27)$$

where the numerator is the r.m.s. value of the difference between the k -th iterate and the true object while the denominator is the r.m.s. value of the true object. Then, if we look at the behaviour of this restoration error as a function of the number of iterations, we find that it first decreases, goes through a minimum and then increases, showing a typical behaviour which is known as *semiconvergence* (see, for instance, Bertero and Boccacci [8]). On the other hand, in the case of pointwise objects such as a binary system, the algorithm seems to converge. An example is given in Fig. 1. In the left panel we plot, in the case of the diffuse object of Fig. 2, the restoration error of Eq. 27 as a function of the number of iterations; the minimum of the restoration error is evident even if it is very flat. It is obvious that one can use about 100 iterations without a significant loss of accuracy. On the other hand, in the right panel we plot the behaviour of the reconstructed flux of a binary system, with $m = 2$, again as a function of the number of iterations. It is evident that, at least as concerns photometry, we have a convergent behaviour and, in such a case, the convergence is very fast since we need about 20 iterations to get an accurate value.

If we apply Eq. 18 to the case of Poisson noise with $\mu = 0$, we obtain the following algorithm for multi-images deconvolution:

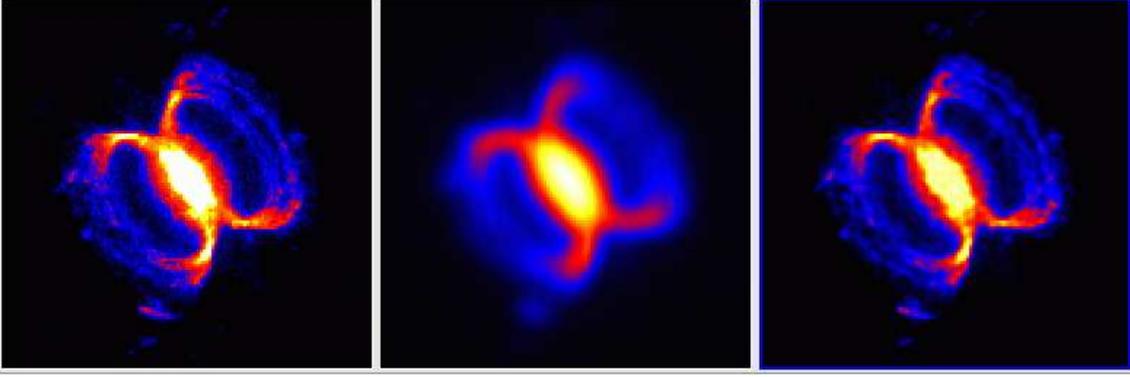


FIGURE 2: An example of diffuse object used in our numerical simulations: in the left panel, the object, a HST image of the Southern Crab Nebula; in the central panel, one of the three LBT images used for the reconstruction; in the right panel, the reconstructed object

- given $\mathbf{f}^{(k)}$ compute:

$$\tilde{\mathbf{f}}^{(k+1)} = \mathbf{f}^{(k)} \left(\frac{1}{p} \sum_{j=1}^p A_j^T \frac{\vec{g}_j}{A_j \mathbf{f}^{(k)} + \vec{b}_j} \right)^\omega, \quad \tilde{c}^{(k+1)} = \sum_{m,n=0}^{N-1} \tilde{\mathbf{f}}^{(k+1)}(m, n), \quad (28)$$

- set:

$$\mathbf{f}^{(k+1)} = \frac{c}{\tilde{c}^{(k+1)}} \tilde{\mathbf{f}}^{(k+1)}. \quad (29)$$

As we already remarked in the previous Section, in the case $\omega = 1$ this is precisely the multi-images extension of the RL-method proposed in Bertero and Boccacci [1]. In a similar way from Eq. 20 we obtain the following algorithm:

- given $\mathbf{f}^{(k)}$, set $j = (k + 1) \bmod p$ and compute:

$$\tilde{\mathbf{f}}^{(k+1)} = \mathbf{f}^{(k)} \left(A_j^T \frac{\mathbf{g}_j}{A_j \mathbf{f}^{(k)} + \mathbf{b}_j} \right)^\omega; \quad \tilde{c}^{(k+1)} = \sum_{m,n=0}^{N-1} \tilde{\mathbf{f}}^{(k+1)}(m, n); \quad (30)$$

- set:

$$\mathbf{f}^{(k+1)} = \frac{c}{\tilde{c}^{(k+1)}} \tilde{\mathbf{f}}^{(k+1)}. \quad (31)$$

If $\omega = 1$, we get the version of the OS-EM algorithm proposed in Bertero and Boccacci [1] while, if $\omega > 1$, we get the accelerated version of the same algorithm proposed and investigated in Anconelli et al. [3]. All the algorithms described above are implemented in the present version of the Software Package AIRY.

As concerns flux normalization, which is taken into account by all our algorithms, the constant c can be selected in the following way. In the case of zero background, as shown in Bertero & Boccacci [1], the total flux of each iterate of the multi-image RL algorithm coincides with the arithmetic mean of the total fluxes of the images \mathbf{g}_j . As a consequence of the introduction of the background terms, the iterates of the corresponding RL-algorithm and of its accelerated version OS-EM, both given in Bertero & Boccacci [2], do not have this property. Therefore, it is convenient to constrain the solution to satisfy the flux condition of the RL-method with zero background; it follows that the constant c can be estimated as the arithmetic mean of the total fluxes of the images \mathbf{g}_j , after subtraction of the background terms:

$$c = \frac{1}{p} \sum_{j=1}^{N-1} \sum_{m,n=0}^{N-1} \{\mathbf{g}_j(m, n) - \mathbf{b}_j\}. \quad (32)$$

In Fig. 2 we give an example of reconstruction of a diffuse object obtained with the OS-EM method. The object is the HST image of the Southern Crab Nebula. In the simulation we have used three images obtained by convolving the object with three ideal interferometric PSFs corresponding to equally spaced orientations of the baseline and by adding Poisson and Gaussian noise. More details about our simulations, which are performed by means of the Software Package AIRY, are given in Anconelli et al. [3]. We have considered several different diffuse objects with different integrated magnitudes and different sets of PSFs, including different levels of adaptive optics (AO) correction.

The plot of the restoration error corresponding to the object of Fig. 2 is just that given in the left panel of Fig. 1. The number of iterations required for getting the best restoration is not terrible but is not small. Therefore it is interesting to investigate the effect of the acceleration exponent ω . As concerns this point we can summarize the results of the numerical simulations described in Anconelli et al. [3]. The tests have been performed mainly on diffuse objects by looking at the number of iterations needed for reaching the minimum of the restoration error. As a result the choice $\omega = 2$ always leads to a converging algorithm, with a reduction by a factor 2 in the number of iterations, hence a gain of a factor 2 in computation time. Moreover by changing the magnitude of the object we have verified that, when the magnitude increases, one can also increase the acceleration exponent. In some cases we have obtained converging algorithms with $\omega = 8$ and a corresponding reduction of the number of iterations by a factor 8. Therefore such an approach seems quite promising for producing quick-look methods.

4. METHODS FOR PARTICULAR CLASSES OF OBJECTS

As we stated in the Introduction we expect that, in general, an observation of a complex object with LBT will require two steps: the first one is a reconstruction with a quick-look method which can be used for identifying the object and understanding qualitatively its main features; the second one is a reconstruction based on a specialized method which has been designed for objects of that type. In this Section we give a few examples which may help to understand our point of view.

4.1. OBJECTS WITH HIGH DYNAMIC RANGE

An important feature of Astronomical images, which is not present in the natural ones, is that, in many important cases, they can have a very high dynamic range, i. e., in the same image one can find objects with very different magnitudes, in general associated with different scales. One example is provided by a model of young binary star with scattered light emission coming from a dusty circumbinary ring. This model is proposed in Carillet et al. [5] and is produced using parameters taken from observations of the close binary of the quadruple GG Tau system in the near infrared. The reconstruction of this object is a difficult problem because the ratio between the flux of one of the two stars and the maximum value of the flux emitted by the pixels of the ring is of the order of 10^6 .

The object is shown in the upper-left panel of Fig. 3. The main component has K-magnitude $m=19$ and the difference of magnitude is $\Delta m = 1$ while the integrated magnitude of the ring is about 15. The corresponding image is represented in the upper-right panel using a logarithmic scale. This image has been obtained by convolving the object with an Airy function and by corrupting the result with Poisson and Gaussian noise. It is evident that the rings of the images of the two stars are superimposed to the ring and that they mask its structure. Finally in the lower-left panel we show the best reconstruction obtained with the RL method, in the sense that it corresponds to the minimum of the reconstruction error defined in Eq. 27 and computed on the region of the ring. Several artifacts are evident; if iterations are further pushed, then the image of the ring is corrupted by the typical *checkerboard effect* of the RL method when applied to the reconstruction of diffuse objects.

It is obvious that the RL-reconstruction already provides important information which cannot be directly deduced from the original image. On the other hand it is also obvious that the reconstruction of the ring is not completely satisfactory for a quantitative analysis. Moreover, by

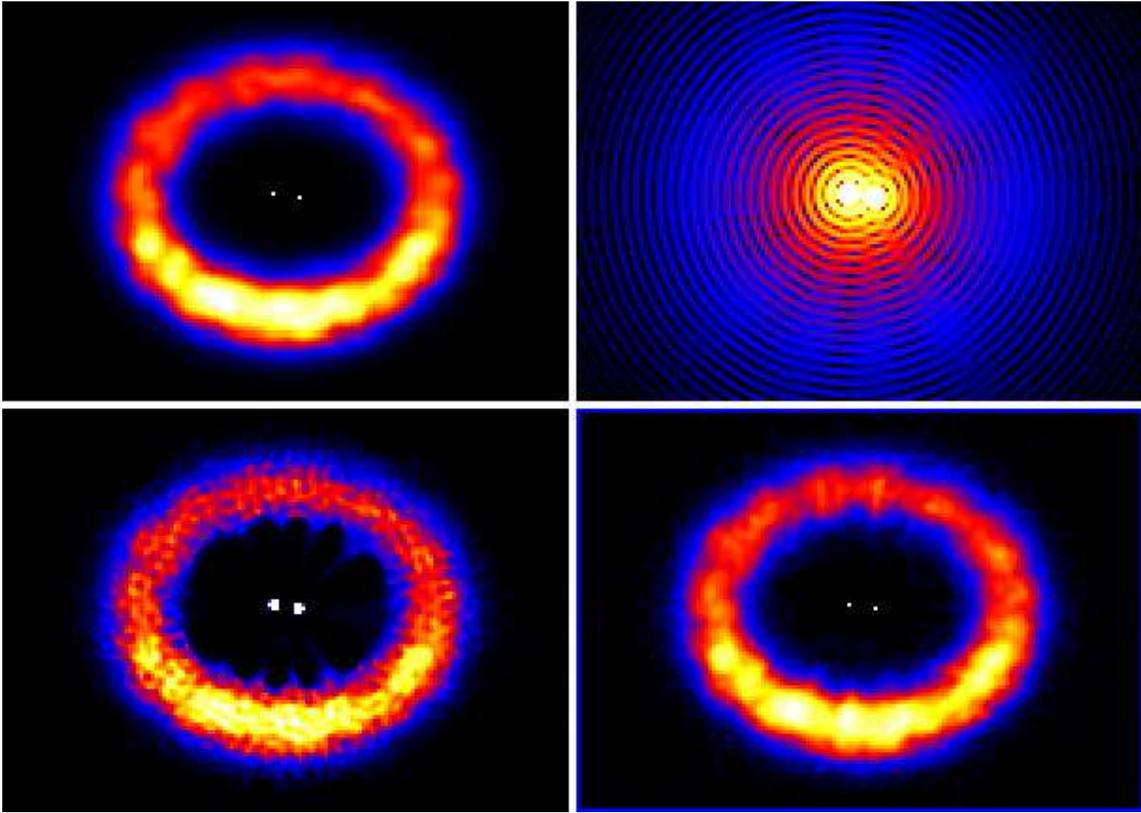


FIGURE 3: Upper-left: the model of young binary system used as a test-object with high dynamic range; upper-right: the image obtained by convolving the object with an Airy function and by corrupting the result with Poisson and Gaussian noise; lower-left: the best reconstruction obtained with the RL-method; lower-right: the best reconstruction obtained with the adaptive regularization proposed in this paper.

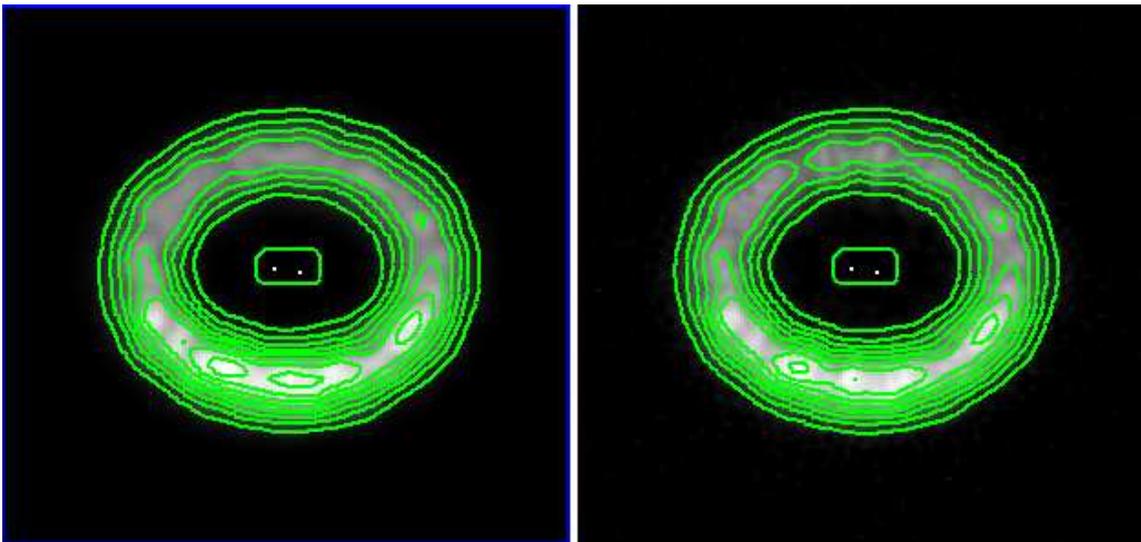


FIGURE 4: Left panel: level curves of the object of Fig. 3; right panel: level curves of the reconstruction obtained by means of the method of Eq. 34.

looking at the behaviour of the iterations, it seems clear that a method acting in a different way in the region of the binary and in that of the disc should be desirable. Such a remark suggest to use a kind of *adaptive* regularization which can be obtained, for instance, by adding to the discrepancy of the photon noise, as given in Eq. 22, the following regularization functional:

$$J_R(\mathbf{f}) = \sum_{m,n=0}^{N-1} \frac{\mathbf{f}^2(m,n)}{\mathbf{f}^2(m,n) + \eta^2} . \quad (33)$$

This functional, with \mathbf{f} replaced by the gradient of \mathbf{f} , was introduced by Geman and Mc Clure [13] in order to get a different regularization in regions with very different intensity variations. It is not convex but is bounded from below and its Hessian has the same property. It has a unique minimum at $\mathbf{f} = 0$ and is convex in a neighborhood of this point. The parameter η is basically a thresholding parameter which separates the low-flux regions from the high-flux ones. Indeed it is easy to recognize that, if $\mathbf{f}(m,n)$ is much smaller than η , then the functional $J_R(\mathbf{f})$ practically coincides with the Tikhonov regularization functional of Eq. 24; on the other hand, if $\mathbf{f}(m,n)$ is much greater than η , then the functional is constant and therefore no regularization of the RL method is introduced.

If we have in mind the object discussed above, we easily understand that, even if a preliminary reconstruction of the object by means of the RL method, does not provide an accurate reconstruction of the ring, however it provides sufficient information for attempting an estimation of the parameter η .

Our problem is now to find an iterative algorithm for the minimization of the functional of Eq. 5, with the functional $J_0(\mathbf{f}; \mathbf{g})$ given by Eqs. 16 and 22, and the functional $J_R(\mathbf{f})$ given by Eq. 33. In order to simplify our equations, we write the regularization parameter in the form: $\mu = \gamma/2\eta^2$; then, by applying SGM to this functional we obtain a very simple modification of the algorithm of Eq. 28 (with $\omega = 1$):

- given $\mathbf{f}^{(k)}$, set $j = (k + 1) \bmod p$ and compute:

$$\mathbf{D}^{(k)} = \frac{\mathbf{f}^{(k)}}{[(\mathbf{f}^{(k)})^2 + \eta^2]^2} \quad (34)$$

$$\tilde{\mathbf{f}}^{(k+1)} = \frac{\mathbf{f}^{(k)}}{1 + \gamma \mathbf{D}^{(k)}} \left(A_j^T \frac{\tilde{g}_j}{A_j \tilde{f}^{(k)} + b_j} \right) , \quad \tilde{c}^{(k+1)} = \sum_{m,n=0}^{N-1} \tilde{\mathbf{f}}^{(k+1)}(m,n) ,$$

- set:

$$\mathbf{f}^{(k+1)} = \frac{c}{\tilde{c}^{(k+1)}} \tilde{\mathbf{f}}^{(k+1)} . \quad (35)$$

In the case of a single image we have applied this algorithm to the object of Fig. 3. We have found that the best restoration is obtained with $\eta = 3 \times 10^4$ and $\gamma = 8 \times 10^{10}$. The result is shown in the lower-right panel of Fig. 3; the r.m.s error on the ring is of the order of 8%. However, in order to make evident the accuracy of the reconstruction, in Fig. 4 we compare the reconstructed object with the original one by looking at the level curves. Except for a small region the accuracy is quite good. Work is in progress for investigating the applicability of this approach to LBT images. The results published in Carbillet et al. [5] indicate that the result of the first step, based on the OS-EM method, is not so reliable as that obtained in the case of a single image and shown in Fig. 3.

4.2. FAINT OBJECTS

In the observation of faint objects such that only a few photons per pixel are available, it may happen that the read-out noise is comparable with the photon noise. In such a situation the approximation of the likelihood function based on the assumption of a dominating photon noise is not valid and one must use the more complete expression which takes into account both Poisson and Gaussian noise, as given by Snyder et al. [7]. If we consider the negative *log* of this function and we modify the terms independent of \mathbf{f} in order to get an expression closer to that of the Poisson case, we obtain the following form of the discrepancy functional for the j -th image:

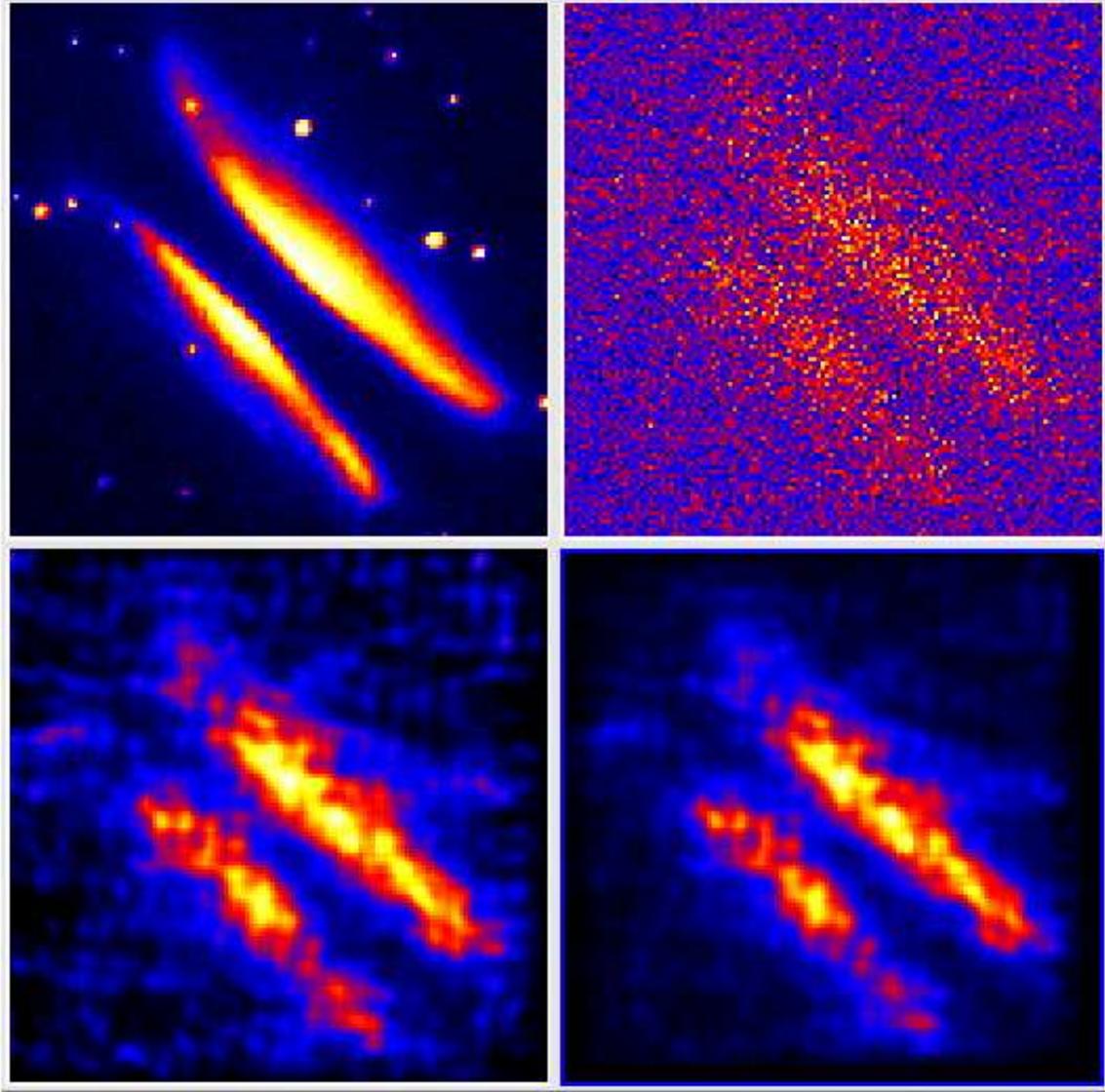


FIGURE 5: Upper-left: the object; upper-right: the blurred image; lower-left: the RL reconstruction; lower-right: Gaussian+Poisson reconstruction.

$$\begin{aligned}
 J_{0,j}(\mathbf{f}; \mathbf{g}_j) &= \\
 &= \sum_{m,n=0}^{N-1} \left\{ -\ln \left(\sum_{l=0}^{+\infty} P_l(m,n) [(A_j \mathbf{f})(m,n) + \mathbf{b}_j]^l \right) + [(A_j \mathbf{f})(m,n) + \mathbf{b}_j - \mathbf{g}_j(m,n)] \right\}
 \end{aligned} \tag{36}$$

where

$$P_l(m,n) = \frac{1}{l!} P_{ron}(\mathbf{g}_j(m,n) - l) , \tag{37}$$

$P_{ron}(\xi)$ being the probability density of the read-out noise, which, in the case of white Gaussian noise with expected value r and variance σ^2 , is given by:

$$P_{ron}(\xi) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(\xi-r)^2}{2\sigma^2}} . \tag{38}$$

The SGM can be applied to this functional and, as shown by Lanteri and Theys [14], the expressions of the U, V arrays are the following:

$$\mathbf{U}_{0,j}(\mathbf{f}; \mathbf{g}_j) = A_j^T \frac{\mathbf{R}_j}{A_j \mathbf{f} + \mathbf{b}_j} , \quad \mathbf{V}_{0,j}(\mathbf{f}; \mathbf{g}_j) = \mathbf{1} , \tag{39}$$

where:

$$\begin{aligned} \mathbf{R}_j(m, n) &= \frac{\mathbf{P}_j(m, n)}{\mathbf{Q}_j(m, n)}, \\ \mathbf{P}_j(m, n) &= \sum_{l=0}^{+\infty} P_l(m, n) l [(A_j \mathbf{f})(m, n) + \mathbf{b}_j]^l, \\ \mathbf{Q}_j(m, n) &= \sum_{l=0}^{+\infty} P_l(m, n) [(A_j \mathbf{f})(m, n) + \mathbf{b}_j]^l. \end{aligned} \quad (40)$$

By inserting these expressions into Eq. 18 or Eq. 20, both with $\mu = 0$, we obtain the iterative algorithm for the multiple-images case and its OS version. The latter is just the algorithm for the single-image case when only one image is available.

We point out that we do not consider the case where a regularization functional is added to the discrepancy functional because, as shown in Lanteri and Theys [14], the addition of standard regularization functionals such as Tikhonov or Entropy does not produce a significant improvement of the restored image. Moreover the convergence of the algorithm is very fast since very few iterations (between 5 and 10) are needed for reaching the best image. Therefore an interactive choice of the best reconstruction, based on the visualization of the result of each iteration, is possible

As an example we show here the result of a numerical simulation described in Lanteri and Theys [14]. The object is a 128×128 HST image of a sun-like star nearing the end of its life. The image has been convolved with a PSF including atmospheric turbulence effects and a total flux of 10^4 photons has been assumed (average number of photons per pixel smaller than 1); moreover a Gaussian white noise with $\sigma = 1$ has been added. The image has been reconstructed using both the RL-method and the algorithm outlined above. The results are shown in Fig. 5 where we show the object, the blurred and noisy image as well as the reconstruction provided by the RL-method and that provided by the new method. The visual improvement obtained with the new method is evident, even if the improvement in the r.m.s. error is not very large. Indeed the minimum restoration error of the RL-method is 3.9 % and is reached after 8 iterations, while the minimum restoration error of the new method is 3.5 % and is reached after 6 iterations.

4.3. SUPER-RESOLUTION OF STELLAR SYSTEMS

As another example of a two-steps approach for an optimal reconstruction of LBT images we briefly describe our recent work on the super-resolution of unresolved objects.

In Astronomy and Microscopy super-resolution is a term which, in general, is used to indicate any method able to improve resolution beyond the diffraction limit, and therefore it implies the extrapolation of the Fourier transform of the object outside the band of the optical instrument. As remarked in a paper by Wolter [15], this is possible if the object has a finite extent. However, since the extrapolation problem is very sensitive to noise propagation from the data to the solution, unlimited super-resolution is impossible in the presence of noise. A first attempt of estimating the amount of super-resolution achievable in practice is developed in Bertero and Pike [16], where it is shown that it is controlled by two parameters: the *space-bandwidth product* (SBP) and the *signal-to-noise* ratio (SNR). In the case of a telescope the SBP is approximately given by the ratio between the angular diameter of the object and the diffraction limit; then the results derived in the paper mentioned above imply that super-resolution is feasible when this parameter is not much greater than one, namely in the case of compact and unresolved objects. On the other hand the amount of super-resolution increases with increasing SNR, as shown by Lucy [17].

According to the previous remarks, the development of a super-resolving method requires an estimate of the domain \mathcal{D} of the object, which must be inserted as a constraint in the restoration algorithm so that the total flux of the restored object is concentrated within \mathcal{D} . However it has been claimed by many authors that other constraints may provide super-resolution such as non-negativity and a constraint on the total flux of the object. For instance, it has been proved that

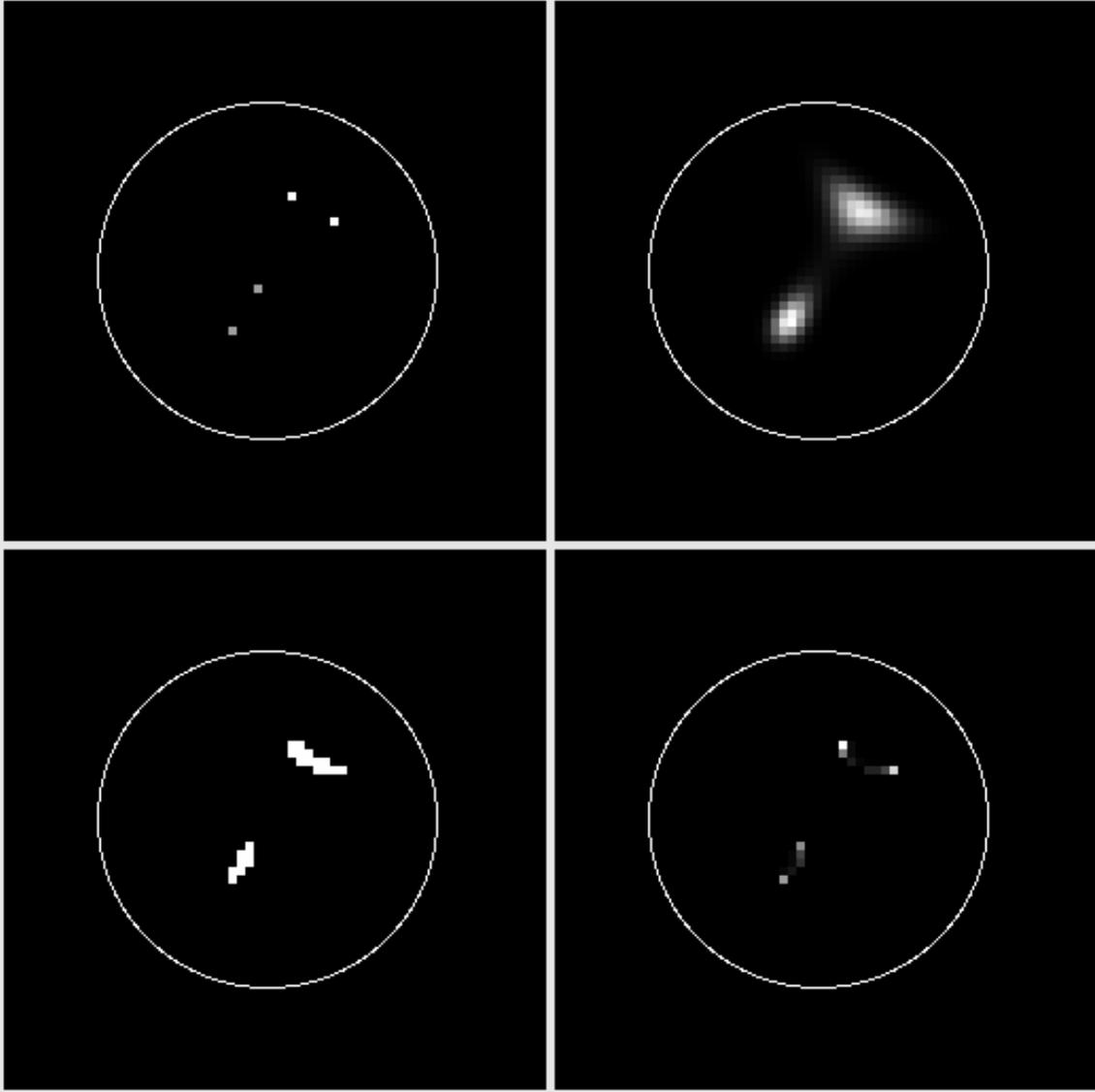


FIGURE 6: Upper-left: the object consisting of a pair of binaries (the diameter of the circle is just λ/B); upper-right: the OS-EM reconstruction (1000 iterations); lower-left: the mask; lower-right: the OS-EM reconstruction initialized with the previous mask (1000 iterations).

maximum entropy provides super-resolution in the case of the so-called *nearly-black objects*, i.e. objects which are essentially zero in the vast majority of the samples (Donoho et al. [18]) and that, in general, non-negativity works well when combined with the sparsity of the object (Donoho [19]). For sparse objects the constraint on the total flux can provide improved reconstructions.

The RL-like methods introduced in the previous Sections implement both the constraint of non-negativity and that on the total flux of the image. Moreover, they can implement also the constraint on the domain of the object if we use a suitable initialization of the iterations. Indeed, this property derives from the fact that the result of iteration $k + 1$ contains as a factor the result of iteration k ; therefore if $f^{(0)} = 0$ at a given pixel then all the iterates will be zero at that pixel. We call such a property the *localization property* of the RL-like methods, in the sense that we can constrain the object to be localized in a given domain by a suitable initialization of the iterations. Therefore RL-like methods provide algorithms implementing in a very simple way all the relevant constraints and they look quite suitable for estimating the amount of super-resolution achievable in practical situations.

Our approach to super-resolution applies to the case where the astronomical target consists of unresolved objects (angular size of the order of the diffraction limit) separated by an angle considerably greater than the diffraction limit (sparsity condition), so that there is no overlapping of their images. For simplicity we assume that there is only one of these unresolved objects, surrounded by empty sky. Then the approach consists of the following steps.

Step 1 - We apply the RL-like method initialized with a constant image (correctly normalized) and we use a number of iterations such that the restored object is sufficiently well localized. If the angular separation is not much smaller than the diffraction limit and the magnitude difference is small, it may happen that the stars of the object are already resolved in this first step, as a consequence of the super-resolving properties of the RL-like methods. In such a case one can estimate their positions by computing the centroids and go directly to Step 3 for a more accurate estimate of their magnitudes.

If the stars are not resolved, then one can go to Step 2.

Step 2 - We define the domain of the object by identifying the pixels where the flux of the result of Step 1 is greater than a selected threshold (for instance some percent of its maximum value); in alternative we can take a disc with diameter equal to the diffraction limit and containing most of the flux of the reconstructed object. Next we apply again the RL-like method to the detected images, but now initialized with the mask of the domain, namely a function which is constant over the domain and zero elsewhere, The constant is evaluated in such a way that the normalization condition is satisfied. Each iterate is localized in the selected domain; if, after a number of iterations, the stars are separated, their positions are obtained by computing their centroids.

In order to illustrate this approach, in Fig. 6 we show an example of its application. The object consists of two unresolved binaries with an angular separation of their centres which is about on-half of the diffraction limit λ/B (here we consider the case of LBT). Also for this object we have considered three images obtained by convolving the object with three ideal PSFs of LBT, corresponding to uniformly spaced directions of the baseline; the images have been corrupted with photon and read-out noise. The result of 1000 OS-EM iterations allows to resolve the two binaries but not the stars within the binaries. Next a mask is obtained by identifying the pixels where the flux is greater than 50 % of the maximum value of the restored image. This mask is also shown in Fig. 6. Finally OS-EM is reused, but now initialized with the mask derived from the first step, and the result of 1000 iterations clearly shows now the structure of the two binaries.

A third step may be required when the unresolved object consists of stars with different magnitudes. Indeed, in such a case, the first two steps may not provide the correct photometry. In order to overcome this difficulty we propose the following third step.

Step 3 - Let us assume that the restored object, provided by Step 1 or Step 2, consists of q stars localized at the pixels P_1, P_2, \dots, P_q (in the application of this paper $q = 2$). Then the positions of the stars are deduced from these restorations while their intensities are obtained by solving a linear least-squares problem.

In Anconelli et al. [6] this approach is applied to the case of binary systems and, by means of a careful analysis, it is shown that a considerable amount of super-resolution is achievable both in the case of a single mirror and in the case of a Fizeau interferometer such as LBT.

5. CONCLUDING REMARKS

In this paper we have presented the state-of-the art of our research about the reconstruction of LBT images as well as the many problems this unique instrument is suggesting. All the methods we have described are implemented in the software package AIRY which is now available in the version 2.1 and can be required at the site <http://dirac.disi.unige.it>. Work is in progress about the validation of the methods described in the Sections 4.1 and 4.2.

One of the authors (M. B.) warmly thanks Martin Adorf and Richard Hook for interesting discussions during the ADA III Conference.

REFERENCES

- [1] Bertero M., and Boccacci P. (2000) Application of the OS-EM method to the restoration of LBT images, *Astron. Astrophys. Suppl. Ser*, **144**, 181-186.
- [2] Bertero M., and Boccacci P. (2000) Image restoration methods for the Large Binocular Telescope (LBT), *Astron. Astrophys. Suppl. Ser*, **147**, 323.
- [3] Anconelli B., Bertero M., Boccacci P., and Carbillet M., & Lanteri H. (2004) Restoration of interferometric images - III. Efficient RL-like methods for LINC/NIRVANA data reduction, submitted to *Astron. Astrophys.*.
- [4] Correia S., Carbillet M., Boccacci P., Bertero M., and Fini L. (2002) Restoration of interferometric images - I. The software package AIRY. *Astron. Astrophys.*, bf 387, 733-743.
- [5] Carbillet M., Correia S., Boccacci P., and Bertero M. (2002) Restoration of interferometric images - II. The case-study of the Large Binocular Telescope, *Astron. Astrophys.*, **387**, 744-757.
- [6] Anconelli B., Bertero M., Boccacci P., and Carbillet M. (2004) Restoration of interferometric images - IV. An algorithm for super-resolution of stellar systems, submitted to *Astron. Astrophys.*.
- [7] 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.*, **A10**, 1014-1023.
- [8] Bertero M., and Boccacci P. (1998) *Introduction to Inverse Problems in Imaging*. IOP Publishing, Bristol.
- [9] Lanteri H., Roche M., Cuevas O., and Aime C. (2001) A general method to devise maximum-likelihood signal restoration multiplicative algorithms with non-negativity constraints, *Signal Process.*, **81**, 945-974.
- [10] Lanteri H., Roche M., and Aime C. (2002) Penalized maximum likelihood image restoration with positivity constraints: multiplicative algorithms, *Inverse Problems*, **18**, 1397-1419.
- [11] Lanteri 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**, 1481-1504.
- [12] Hudson H. M., and Larkin R. S. (1994) Accelerated image reconstruction using ordered subsets of projection data, *IEEE Trans. Med. Im.*, **20**, 100-108.
- [13] Geman S., and Mc Clure D. (1987) "Statistical methods for tomographic image reconstruction", in *Proceedings of the 46th session of the ICI, Bulletin of the ICI*, 5-21.
- [14] Lanteri H. and Theys C. (2004) Restoration of Poisson images with Gaussian noise - Application to astrophysical data, submitted to *Appl. Signal Process.*.
- [15] Wolter H. (1961) On basic analogies and principal differences between optical and electronic information, in *Progress in Optics*, Vol. I, ed. E. Wolf, 155-210.
- [16] Bertero M., and Pike E. R. (1982) Resolution in diffraction-limited imaging, a singular value analysis - I. The case of coherent illumination, *Optica Acta*, **29**, 727-746.
- [17] Lucy L. B. (1992) Statistical limits to super-resolution, *Astron. Astrophys.*, **261**, 706-710.
- [18] Donoho D. L., Johnstone I. M., Hoch J. C., & Stern A. S. (1991) Maximum Entropy and the Nearly Black Object, *J. R. Statist. Soc.*, **B54**, 41-66.
- [19] Donoho D. L. (1992) Super-resolution via sparsity constraints, *SIAM J. Math. Anal.*, **23**, 1309-1331.