# Comparison between ISRA and RLA algorithms. Use of a Wiener Filter based stopping criterion

**H. Lantéri, R. Soummer, and C. Aime**

UMR 6525, "Astrophysique", Université de Nice, Sophia Antipolis, Faculté des Sciences, Parc Valrose, 06108 Nice Cedex 2, France
e-mail: Henri.Lanteri@unice.fr

**Abstract.** This paper consists of two parts. In the first one, we make a comparative analysis of the algorithms of Richardson-Lucy (RLA) and the Image Space Reconstruction Algorithm (ISRA), for image deconvolution of astronomical images. These iterative algorithms keep the reconstructed image non-negative while maximizing the likelihood for a Poisson Process (RLA) or a Gaussian process (ISRA). Their comparison is made easier when these algorithms are rewritten as descent algorithms; the additive forms evidence the role of the variances of the noise and allow a better understanding of these algorithms. A numerical illustration is performed from simulated images. In practice, the results obtained by the two algorithms appeared to be very similar, independently of the statistics of the noise.

In the second part of the paper, we propose a new objective-stopping technique that makes use of a comparison of the results of these algorithms with that of the Wiener filter. The comparison is made in the Fourier plane, computing the Euclidean distances between modules of the spatial frequencies components of the images. We propose then to use the results of ISRA and RLA at the iteration number corresponding to the minimum of that distance. The technique is checked in a numerical simulation, for which the optimal iteration numbers can be easily determined.

A good agreement is obtained between best ISRA and RLA results and the Wiener filter, in particular for rather noisy images. However ISRA and RLA produce very high frequency components, outside the cut-off frequency of the instrument; limiting the iteration number alone cannot allow a perfect agreement with the Wiener approach, stressing the need of an explicit regularization for astronomical applications.

**Key words:** methods: numerical — techniques: image processing

## 1. Introduction

The image restoration problem consists in the reconstruction of the best estimate of an object "$x$" from the knowledge of a blurred image $\tilde{y}$. In the general case, the transformation suffered by "$x$" is described by a Fredholm integral equation of the first kind:

$$\tilde{y}(r) = \int h(r, \alpha) x(\alpha) d\alpha. \tag{1}$$

The function $h(r, \alpha)$ is the kernel of the integral equation. For the deconvolution problem, the kernel $h$ is space invariant and is called the Point Spread Function (PSF). The above relation then becomes:

$$\tilde{y}(r) = \int h(r - \alpha) x(\alpha) d\alpha. \tag{2}$$

Experimentally, the main difficulty in recovering $x$ arises from the fact that we can only access a noise-corrupted version $y$ of $\tilde{y}$, for which the convolution relationship does not apply. The general problem is then to obtain a solution $x$ minimizing a distance $D$ between the model and the observation, thus we can write the problem in the form:

$$\text{Minimize}/x : \ D\lfloor y, \int h(r - \alpha) x(\alpha) d\alpha \rfloor. \tag{3}$$

The particular definition of the distance $D$ must be established in accordance with the statistics of the noise. This problem belongs to the general class of ill-posed problems in the sense of Hadamar, extensively studied in the literature (Bertero 1989; Bertero et al. 1980, 1995; Demoment 1986). The main difficulties lie in the instabilities appearing in the solution due to noise amplification during the reconstruction process. To avoid such effects and to obtain stable and physically suitable results, a regularization of the problem is necessary. A prior knowledge of the properties of the admissible solutions is required; however, classical constraints for a non-negative and spatially limited solution do not allow avoiding the instabilities. To obtain stable solutions, some regularization by a smoothness constraint is required. An explicit regularization implies the

choice of the smoothing operator and the strength of the constraint, that is a compromise between the consistency with the data and the amount of details wanted in the reconstructed image. For iterative algorithms, an elementary regularization can be made by stopping the iterations. It is what we have done in the present paper.

The mathematical developments we present in this paper are written in the algebraic form corresponding to the discrete problem. The above relations then write in the form of a linear system $\tilde{y} = Hx$. The solution $x$ is obtained by minimizing the distance $D(y, Hx)$. The operator $H$ takes a particular form in the case of the convolution we examine here. This assumption makes the numerical simulations easier, but most of the results obtained in this paper remain valid for the general case of Eq. (1).

The deconvolution problem can be solved using a likelihood (or Log likelihood) maximization (Lane 1996; Zaccheo & Gonsalves 1996). We analyze it for two different aspects of the noise process (Gaussian and Poisson processes), leading to the Image Space Reconstruction Algorithm (ISRA) and the Richardson-Lucy Algorithm (RLA).

The paper is organized as follows. In Sect. 2, after a brief recall on the classical unconstrained Gaussian case leading to the generalized inverse solution, we write ISRA in additive form, to make it appear as the solution of the least squares problem under non-negativity constraint. Then we proceed in the same way for RLA that corresponds to the Log-Likelihood maximization with non-negativity constraint for the Poisson case. The additive form makes the comparison between these two algorithms easier, and evidences interesting features related to variances of the noise. Next, we compare the behavior of the two algorithms on simulated data and analyze the spectral extension due to the non-linearity of the algorithms. In Sect. 3, we propose a new method to determine the iteration number corresponding to the best-restored image, by comparison with the results of the Wiener filter. This is a slight modification of a procedure presented in a recent ESO/NOAO Conference (Lanteri et al. 1998). Two appendices are used for particular computations and conclusions are given in Sect. 4.

## 2. Comparative analysis of ISRA and RLA as descent algorithms

In this section, we recall the origins of ISRA and RLA with the objective to show that their analysis can be done in a similar way. Writing these algorithms in the form of descent algorithms allows a better understanding of their behavior. We compare the results given by these algorithms on simulated data with different noise levels; we show the global closeness of the results and we put in evidence the limits of the algorithms for the different types of noise process.

### 2.1. Gaussian noise: The Image Space Reconstruction Algorithm (ISRA)

Let us first consider that the data are corrupted by an additive, zero-mean, white Gaussian noise "$n$", with variance $\sigma_i$, and that the data "$y_i$" are independent random variables. The image formation model may be written as $y = Hx + n$. For a given pixel "$i$", the data "$y_i$" is the sum of a deterministic part $(Hx)_i$, which is the mean of the intensity in the considered pixel, and a random noise component "$n_i$". The likelihood, that is the conditional probability to have "$y$" knowing "$Hx$" is given by:

$$L(x) = P(y/x) \approx \prod_i \exp\left[-\frac{(y_i - (Hx)_i)^2}{2\sigma_i^2}\right]. \qquad (4)$$

The corresponding log-likelihood is then:

$$\text{Log}[P(y/x)] \approx -\frac{1}{2}\sum_i \frac{(y_i - (Hx)_i)^2}{\sigma_i^2} \qquad (5)$$

and the solution "$x$" is obtained by minimization of the weighted Euclidean distance between the raw data and the restored mean:

$$\begin{aligned} J(x) &= -\text{Log}[P(y/x)] \\ &= \frac{1}{2}\sum_i \frac{(y_i - (Hx)_i)^2}{\sigma_i^2} = \frac{1}{2}||y - Hx||_R^2 \end{aligned} \qquad (6)$$

where the symbol $||\cdots||^2$ stands for the Euclidean norm and the index $R$ denotes the weighting diagonal matrix $R = \text{diag}\left[\frac{1}{\sigma_i^2}\right]$.

Any iterative descent method may be used to minimize $J(x)$.

Making use of the gradient expression:

$$\nabla_x J(x) = H^T R H x - H^T R y \qquad (7)$$

a very simple algorithm is then the constant-stepsize gradient method, which can be written in the form:

$$x^{(k+1)} = x^{(k)} + \alpha[-\nabla J(x^{(k)})] = x^{(k)} + \alpha H^T R(y - Hx^{(k)}) \qquad (8)$$

where the descent stepsize $\alpha$ must be such that $0 < \alpha < 2/||H^T R H||$ to ensure the convergence of the algorithm.

The solution $x^*$ given by this algorithm when $k \to \infty$ is the minimum norm least squares solution (MNLS), which may be formally expressed as $x^* = (H^T R H)^{-1} H^T R y$ corresponding to the first order optimality condition: $\nabla J(x^*) = 0$.

However, the deconvolution problem is an ill-posed problem, and the MNLS algorithm produces numerical instabilities. The noise affecting the data is amplified during the restoration process.

Moreover the solution may have negative parts without physical meaning; this later point is overcome by using the constraint for a non-negative solution. Several algorithmic methods have been proposed for that, and in particular ISRA.

ISRA was initially proposed in a multiplicative form by Daube-Witherspoon & Muehllehner (1986). The convergence of the algorithm was then analyzed by

De Pierro (1987, 1990, 1993), and Titterington (1987), and they show that the sequence obtained by ISRA converges toward the solution of the problem:

$\text{Min}/x \ J(x) = \frac{1}{2}\|y - Hx\|_R^2, \quad \text{with the constraint } x \geq 0,$

and that the Kuhn-Tucker first order optimality conditions are satisfied at the optimum (Luenberger 1973).

In its original form, ISRA is written as:

$$x_i^{(k+1)} = x_i^{(k)} \cdot \frac{(H^T R y)_i}{(H^T R H x^{(k)})_i} \quad \forall i. \tag{9}$$

All components of the initial estimate must be positive or equal to zero to satisfy the constraints. Note that if the value of a particular component of an estimate is zero, it remains to zero for all successive iterations. This property may be used to apply a support constraint, setting to zero the components of the initial estimate outside the support.

The iterative algorithm given by Eq. (9) can straightforwardly be rewritten in the additive form:

$$x_i^{(k+1)} = x_i^{(k)} + \frac{x_i^{(k)}}{(H^T R H x^{(k)})_i}[H^T R(y - Hx^{(k)})]_i. \tag{10}$$

Or in a matrix form:

$$x^{(k+1)} = x^{(k)} + \text{diag}\left[\frac{x_i^{(k)}}{(H^T R H x^{(k)})_i}\right] H^T R(y - Hx^{(k)}), \tag{11}$$

which is clearly a modified gradient form with descent direction:

$$d^{(k)} = \text{diag}\left[\frac{x_i^{(k)}}{(H^T R H x^{(k)})_i}\right] H^T R(y - Hx^{(k)}). \tag{12}$$

In the optical deconvolution problem, the matrix $H$ and $R$ have positive entries and all the components of the initial estimate have positive values to satisfy the constraint. Then the descent properties of the gradient algorithm are preserved (De Pierro 1987, 1990). In the gradient form, ISRA may eventually be relaxed to control the convergence speed of the algorithm. The relaxation method (Lascaux & Theodor 1986) could be summarized as follows: for any non-relaxed iterative algorithm giving an estimate $\hat{x}^{(k+1)}$ from $x^{(k)}$, the relaxed estimate $x^{(k+1)}$ with the relaxation factor "$\alpha$", is given by the relation: $x^{(k+1)} = (1 - \alpha)x^{(k)} + \alpha\hat{x}^{(k+1)}$.

Evidently, for relaxed algorithms, the non-negativity of the solution as well as the convergence of the algorithm is not guaranteed; a specific analysis that goes beyond the scope of this paper would then become necessary.

## 2.2. Poisson noise: The Richardson Lucy Algorithm (RLA)

For a perfect photon detection, the intensity in the pixel "$i$", is a random variable that follows a Poisson law of mean $(Hx)_i$. The likelihood may be written as:

$$L(x) = P(y/x) = \prod_i \frac{[(Hx)_i]^{y_i}}{y_i!} \exp[-(Hx)_i]. \tag{13}$$

Using Stirling's formula, the negative Log-likelihood becomes:

$$T(x) = -\text{Log}\,[L(x)] \approx \sum_i [(Hx)_i - y_i] + y_i\text{Log}\frac{y_i}{(Hx)_i}. \tag{14}$$

This function is known as the Csiszär (1991) I-divergence measure between $(Hx)$ and $y$. This measure generalizes the Kullback divergence or cross-entropy measure to accommodate functions whose integrals are not constant, as they would be if they were probability distributions (Snyder et al. 1992). It is the quantity we seek to minimize, keeping the constraint that "$x$" must be positive. Dropping in $T(x)$ the terms independent of "$x$", we must minimize under positive constraint, the quantity:

$$D(x) = \sum_i (Hx)_i - y_i\text{Log}(Hx)_i \tag{15}$$

Shepp & Vardi (1982), solve this problem using the Expectation-Maximization (EM) method, and obtain the iterative algorithm:

$$x_i^{(k+1)} = \frac{1}{a_i}x_i^{(k)}\left[H^T\frac{y}{Hx^{(k)}}\right]_i \tag{16}$$

where the ratio $\frac{y}{Hx^{(k)}}$ is a vector computed component wise, and with $a_i = \sum_j H_{j,i}$.

This algorithm was previously proposed by Richardson (1972) and by Lucy (1974) in the field of Astrophysics. Shepp & Vardi (1982) and Byrne (1993) show that the convergence hold, and that the sequence generated by this algorithm also satisfies the Kuhn-Tucker optimality conditions for the constrained minimization problem: $\text{Min}/x$: $D(x), x \geq 0$.

As for ISRA, all successive estimates remain positive if the initial estimate is positive. Moreover, if a component of the solution is set to zero, or becomes equal to zero, it remains equal to zero for all successive iterations.

Here again, it is interesting to write RLA in the additive form:

$$x_i^{(k+1)} = x_i^{(k)} + \frac{x_i^{(k)}}{a_i}\left[H^T\text{diag}\left[\frac{1}{(Hx^{(k)})_i}\right](y - Hx^{(k)})\right]_i \tag{17}$$

or in the global matrix form:

$$x^{(k+1)} = x^{(k)} + \text{diag}\left[\frac{x_i^{(k)}}{a_i}\right]H^T\text{diag}\left[\frac{1}{(Hx^{(k)})_i}\right](y - Hx^{(k)}). \tag{18}$$

Taking into account the expression of the gradient of $D(x)$:

$$\nabla D(x) = H^T\text{diag}\left[\frac{1}{(Hx)_i}\right](Hx - y) \tag{19}$$

and with $x_i^{(k)} \geq 0$, $a_i > 0$, $\forall i$, we recognize a descent algorithm of the form:

$$x^{(k+1)} = x^{(k)} + \text{diag}\left[\frac{x_i^{(k)}}{a_i}\right]\left[-\nabla D(x^{(k)})\right]. \tag{20}$$

The descent direction is no longer the negative gradient direction, but the descent property of the gradient algorithm is kept. In our particular problem, a supplementary simplification is due to the fact that $a_i = 1 \ \forall i$. As

for ISRA, the relaxation allows to modify the speed of convergence, but as previously mentioned this implies a particular analysis of the relaxation factor to ensure the non-negativity of the solution and the convergence of the algorithm.

### 2.3. Comparative analysis of ISRA and RLA

A few remarks can be made on the results presented in the above sections.

i) The additive forms of ISRA and RLA evidence the similar roles played by the weighting matrix $R = \text{diag}\frac{1}{\sigma_i^2}$ in Eq. (11) and by the matrix $\text{diag}\frac{1}{(Hx^{(k)})_i}$ in Eq. (18). These quantities both correspond to variances of the signal. For the Gaussian process, $\sigma_i^2$ is the variance of the intensity in a pixel. For the Poisson process $(Hx^{(k)})_i$ also represents the variance of the signal, because for this process the variance equals the mean.

ii) In the Gaussian case, $J(x)$ is a quadratic form, the convexity of this objective function allows to obtain the solution by solving the linear system $\nabla J(x) = 0$. The Gradient or Conjugate Gradient methods in their classical form leading to linear algorithms may be used in this case and the stepsize is known explicitly. We must observe that the non-linearity of the algorithm ISRA is introduced only by the non-negativity constraint.

On the contrary, in the Poisson case, the functional $D(x)$, although convex, is not quadratic. Thus, the non-linearity of the RLA algorithm is due to the particular form of $D(x)$ as well as to the non-negativity constraint.

iii) For the two algorithms, we set the initial estimate to a constant value: $x_i^{(0)} = \frac{\sum_i y_i}{N}$ where $N$ is the number of pixels in the image; moreover the integral of the image is generally normalized to 1. We must observe that in the RLA algorithm we have for all $k$: $\sum_i x_i^{(k)} = \sum_i y_i$, which is not the case for ISRA; this is due to the fact that the functional $D(x)$ implicitly contains the constant intensity constraint: $\sum_i [(Hx)_i - y_i]$ together with $\sum_j H_{j,i} = a_i = 1$, these later property appearing naturally for the PSF in a deconvolution problem or being imposed in the algorithm.

In the case of ISRA, the conservation of the total intensity of the image is ensured by normalization after each iteration.

iv) The two algorithms are not regularized. A non-negativity constraint, although necessary, does not produce smooth solutions. When the iteration number increases too much, the classical phenomena of instability appear in the solutions. On the practical point of view, the iteration number must be limited to obtain an acceptable compromise between resolution and stability of the solutions; this allows some regularization. Alternatively, an explicit regularization may be done by introducing a supplementary constraint to attenuate the high frequency components in the solution; this will be proposed in a future paper.

### 2.4. Numerical simulations

We have performed a numerical simulation to illustrate some features of the deconvolution procedures described in the previous sections; the computations are made using Mathematica (Wolfram 1996).

#### 2.4.1. Test data

The simulated data are shown in Figs. 1 to 3. Figure 1 gives, in contour plot, the object, PSF and image of the noiseless objet-image relationship. Only the central part of the images are shown here (about 30% of the surface of the $64 \times 64$ points images). Such representation is used all over the paper for similar figures.

The object $x$ is chosen to represent an ensemble of three stars, with different diameters and centers to limb variations. A bright and dark sunspot-like feature is added to the larger star's atmosphere. The PSF $h$ is a realistic representation of the PSF of a true telescope. It is computed as the squared modulus of the Fourier Transform of a function representing the telescope aperture with a small phase aberration. It may correspond to observing conditions with the Hubble Space Telescope operating in the far ultraviolet (Gilliland & Dupree 1996). The resulting Optical Transfer Function is then a low pass filter, limited in spatial frequencies to the extent of the aperture autocorrelation function. The observed filtered image $y$ is then strictly band limited. This particular property is of interest for the study of the reconstructed image in the Fourier plane discussed below.

Figures 2 and 3 represent possible detected image in the case of Poisson and Gaussian statistics. Figure 2 represents the observed image assuming a perfect photodetection, for two levels of noise corresponding to a total of $10^4$ and $10^6$ photoelectrons in the image. Such images may be obtained in visible observations. Figure 3 represents the observed image corrupted by additive Gaussian noise and may correspond to infrared observations with noise coming from the detector and the background.

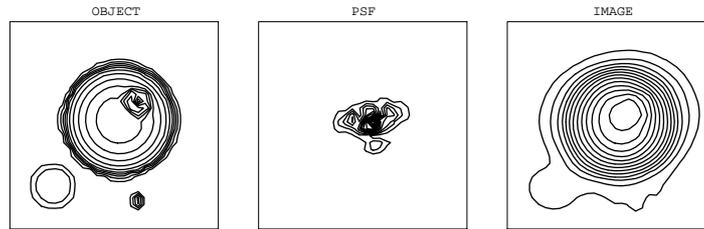#### 2.4.2. Numerical comparison between ISRA and RLA

To implement numerically ISRA and RLA algorithms, we have used the continuous forms of the algorithms for which the convolutions can be easily obtained via FFTs. The algorithms, written in the form of an iterative multiplication, then become:
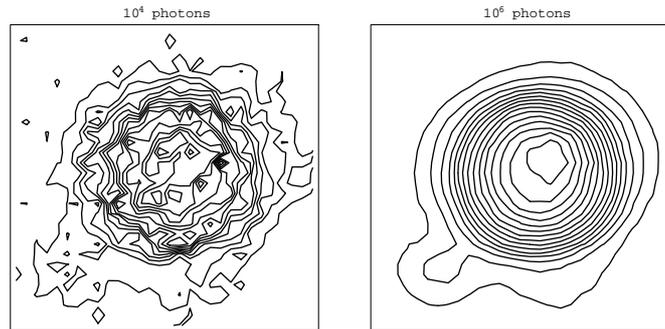for ISRA:

$$x^{k+1}(r) = x^k(r)\frac{h(-r) \otimes y(r)}{h(-r) \otimes h(r) \otimes x^k(r)} \tag{21}$$

and for RLA:

$$x^{k+1}(r) = x^k(r)\left[h(-r) \otimes \left\{\frac{y(r)}{h(r) \otimes x^k(r)}\right\}\right] \tag{22}$$

**Fig. 1.** From left to right: Simulated Object, Point Spread Function (PSF) and noiseless blurred Image. The object represents an ensemble of three stars with various centers to limbs variations. A bright feature is added to the larger star's atmosphere



**Fig. 2.** Simulated noisy images: Poisson statistics. Representations of the blurred image of Fig. 1 observed with $10^4$ photoelectrons (left) and $10^6$ photoelectrons (right)
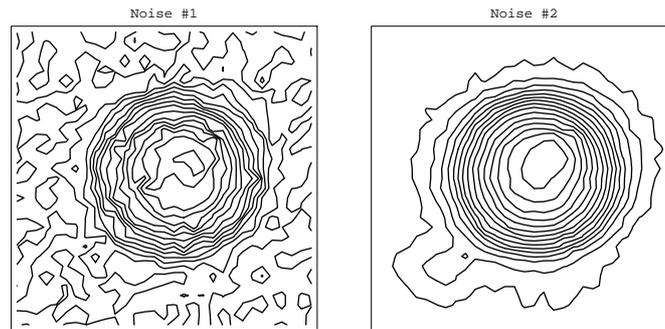


**Fig. 3.** Simulated noisy images: representations of the blurred image of Fig. 1 observed with a level of additive Gaussian noise of strength comparable to that of Fig. 2

where the symbol $\otimes$ denotes the convolution. These well-known forms are easily obtained with the continuous form of the gradients developed in annex A and B.

The results of the deconvolutions for the two different statistics are presented in the following three figures. Figures 4 and 5 give the results obtained for the deconvolution of Fig. 2, while Fig. 6 gives an example of the results obtained for an additive signal-independent Gaussian noise. The results are presented for different iterations number "$k$". As $k$ increases, the image becomes sharper, but the noise increases too. When doing with real data, it is extremely difficult to determine what is the best-reconstructed image, i.e. when to stop the iteration. In a numerical simulation, we can obviously answer this question. Aside a visual comparison between the results and the object, we have used as criterion of

quality the simple Euclidean Distance (ED)- between the true object and the reconstructed image. On this basis, for Poisson noise, we found the results of RLA slightly better than those of ISRA. This is not surprising, since RLA is constructed for the Poisson statistic. However, the differences between ISRA and RLA results are almost insignificant, as it can be deduced by a simple visual inspection of the results. To the contrary, ISRA gives much better results than RLA for Gaussian statistics. This is also expected, but in this case, the difference is strongest. This can be clearly seen in Fig. 6. Moreover, ISRA is almost insensitive to small negative value present in the signal (these negative values appear when the average value of the background is subtracted). For RLA, an offset must be used to reconstruct the image to deal with positive values only. This difference in behavior can

isra6i50    isra6i164    isra6i200

rl6i50    rl6i134    rl6i200
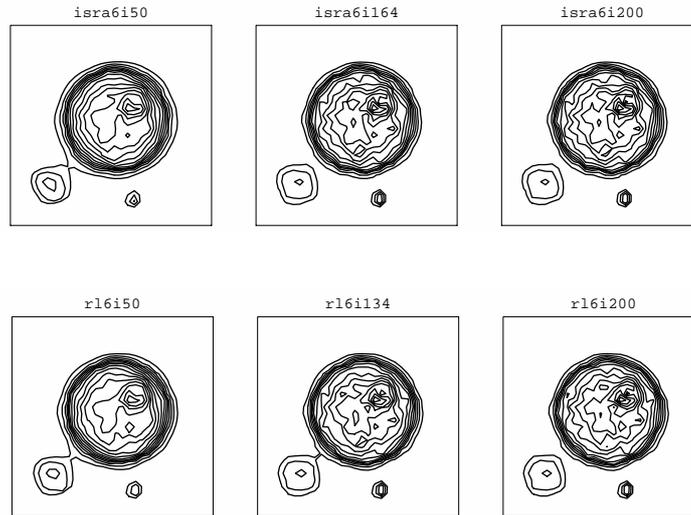
**Fig. 4.** Reconstructed objects obtained from the $10^6$ photoelectrons image using ISRA (Top) and RLA (Bottom) for various iterations numbers. The center images (iteration #164 for ISRA and iteration #134 for RLA) correspond to the best results in terms of the Euclidean Distance (ED) between the object and the reconstructed image. Left and right images correspond to iterations numbers 50 and 200 with comparable EDs. Results obtained using ISRA or RLA are similar

isra4i2    isra4i7    isra4i22
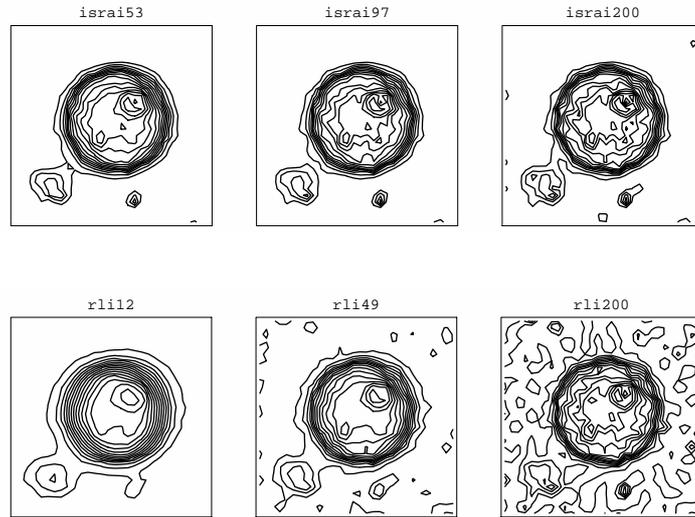
rl4i2    rl4i6    rl4i18
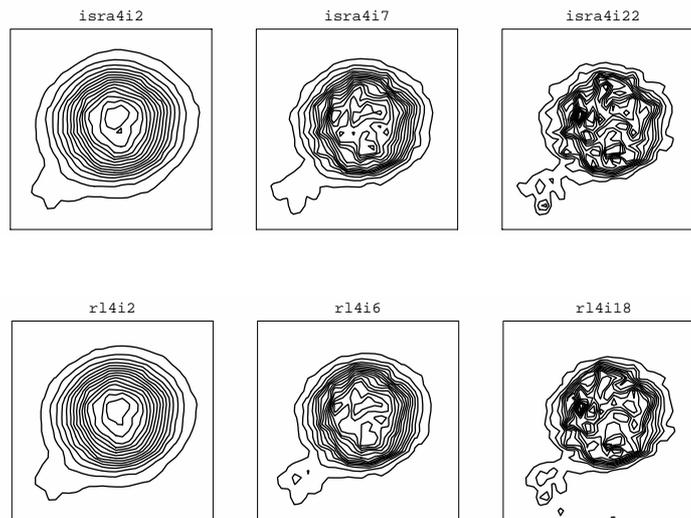
**Fig. 5.** Reconstructed objects obtained from the $10^4$ photoelectrons image using ISRA (Top) and RLA (Bottom) for various iterations numbers. The center images (iteration #7 for ISRA and iteration #6 for RLA) correspond to the best results in terms of the Euclidean Distance (ED) between the object and the reconstructed image. Left and right images correspond to iterations numbers 2 and 18 or 22, with comparable EDs. Here again, the results obtained using ISRA or RLA are very similar

be easily understood considering that ISRA works on the ratio of two smooth expressions ($H^T y$ over $H^T H x$) while RLA does the smoothing by $H^T$ after taking the ratio $y$ over $Hx$.

Figure 7 clearly shows that the reconstructed images present energy outside the frequency limit of the telescope. These effects are in part due to the non-linearity of these algorithms. The three top curves of this figure correspond

to the modulus of the Fourier Transforms of the images of Fig. 1. In this representation, the zero frequency is at the center of the image. As already indicated, the resulting blurred image has no frequency power outside the dish zone corresponding to the aperture autocorrelation function.

The three bottom curves of Fig. 7 represent the modulus of the Fourier Transforms of the reconstructed images.

**Fig. 6.** Reconstructed objects obtained from the image contaminated by additive noise using ISRA (Top) and RLA (Bottom) for various iterations numbers. In this case, the results obtained using ISRA are better than those obtained with RLA
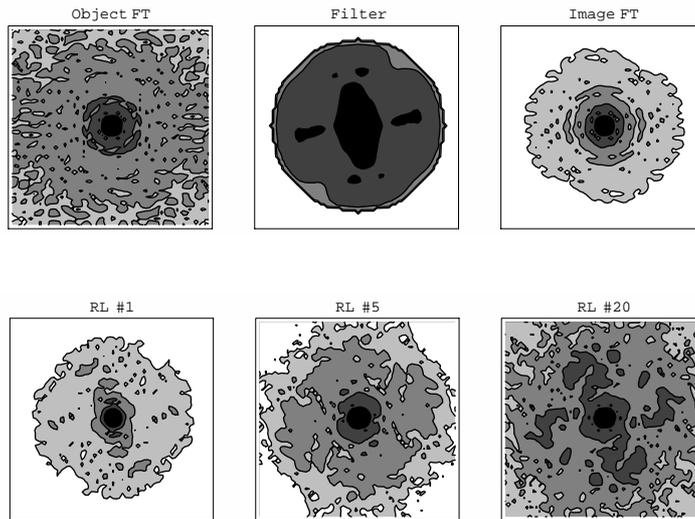


**Fig. 7.** Effect of the reconstruction procedure in the Fourier plane. The curves represent in gray levels the modulus of the Fourier Transforms of the images. From top to bottom and from left to right: original object, point spread function and noiseless image. Results of the RLA algorithm, for the iterations numbers 1, 5 and 20 for the $10^4$ photoelectrons image. It is clear that the RLA algorithm produces power in the reconstruction outside the limits of the original blurred image. Similar results are obtained using ISRA

The example is given for RLA at iterations numbers 1, 5 and 20, for the $10^4$ photoelectrons image, but similar results can be obtained for ISRA. The iterative algorithms rapidly produce power outside the limits of the original blurred image, i.e. in a region of the Fourier plane fully dominated by the noise. The extent of frequencies is inherent to the iterative algorithms used. Figure 8 gives a measure of this excess of signal outside the cut-off frequency by giving, as a function of the iteration number, the ratio of the integral of $\mathrm{Abs}[X^k(u)]$ outside the cut-off

frequency to the integral of $\mathrm{Abs}[X^k(u)]$ inside the cut-off frequency. This behavior is still observed when a noiseless image is processed.

## 3. Determination of the level of regularization by comparison with the Wiener filter

As indicated in the previous sections, the reconstruction of the object must be made using regularization, in order to prevent unwanted amplification of the noise. We
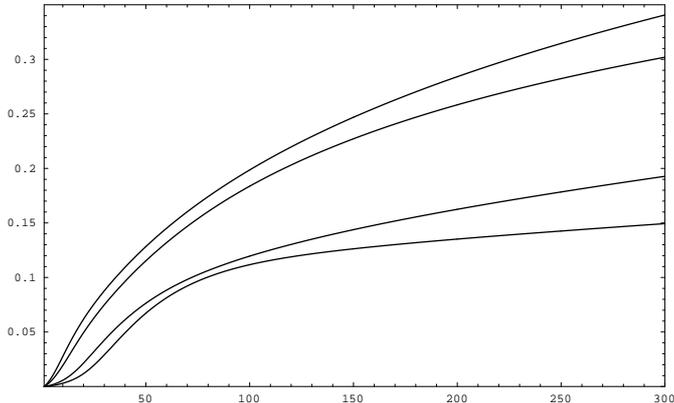
**Fig. 8.** Spatial frequencies band extension due to non-linearity of the iterative algorithms. The curves give the ratios of the quantity $\text{Abs}\lfloor X^{(k)}(u) \rfloor$ beyond the cut off frequency to $\text{Abs}\lfloor X^{(k)}(u) \rfloor$ below the cut off frequency, as a function of the iteration number k. From top to bottom: RLA and ISRA with $10^4$ photons per image, RLA and ISRA with $10^6$ photons per image

consider here that the regularization is obtained by stopping the iterations. The determination of the optimal iteration number (or of the regularization coefficient) is one of the main problems encountered in the reduction of image blurring. Several techniques have been suggested for that: feasibility tests and Cross-Validation tests (Nuñez & Llacer 1993), or Goodness of fit tests (Lucy 1994).

We describe here the implementation of a new objective-stopping technique that is based on the comparison of the results of these algorithms (RLA or ISRA) with that of the optimum Wiener filter. The principle of the approach was recently proposed in an ESO/NOAO Conference (Lanteri et al. 1998). Fundamentally, it consists of selecting the iteration number $k$ for ISRA (or RLA) that gives a result close to that of the Wiener filter. This technique could also be used to determine the optimal regularization factor when an explicit regularization is performed.

The idea is to take advantage of the two kinds of algorithms. The final reconstructed image is given by ISRA (or RLA) and therefore has an ensemble of interesting properties, such as it is non-negative. Closeness with the optimum Wiener solution ensures that the reconstructed image is not too much corrupted by noise, as wished.

The Wiener filter has the advantage of being defined from physical considerations on SNR in the Fourier plane. It is based on the simple fact that spatial frequency components are differently corrupted by noises. For correctly sampled data, low and intermediate spatial frequencies are likely to contain more signal than noise, and the highest spatial frequencies components tend to contain essentially pure noise. It is clear for example, that the spectral energy detected outside the cut-off frequency of the telescope used is entirely due to noise, or to algorithmic artifacts. The

Wiener technique is well known (Brault & White 1971) and we just present here the main features.

The Wiener filter operates in the Fourier plane. Let $u$ denote the two-dimensional spatial (or angular) frequency $(u_1 \ u_2)$. Let $Y(u)$, $H(u)$ and $X_{\text{W}}(u)$ denote the Fourier transforms of the observed image, point spread function and of the reconstructed Wiener image, respectively. In optics, the quantity $H(u)$ is the optical transfer function.

We have:

$$X_{\text{W}}(u) = \frac{Y(u)}{H(u)}\Phi(u), \qquad (23)$$

where $1/H(u)$ is the raw inverse filter, and $\Phi(u)$ is the Wiener zero-phase filter defined as:

$$\Phi(u) = \frac{P_{Hx}(u)}{P_{Hx}(u) + P_N(u)}. \qquad (24)$$

The quantities $P_{Hx}(u)$ and $P_N(u)$ are the power spectra of the noiseless image and of the noise, respectively. Roughly, $P_N(u)$ can be considered as an additive constant quantity. Of course, this may be true only in average. Since $H(u)$ is given by the autocorrelation function of the phase-aberrated aperture, $P_{Hx}(u)$ is limited in angular frequencies to $D/\lambda$, $D$ being the telescope diameter and $\lambda$ the mean light wavelength. Spatial frequencies outside the telescope cut-off frequency are removed by the Wiener filter. Below that limit, the Wiener filter preserves the high SNR frequency components, while attenuating the low SNR frequency components. The resulting Wiener image $x_W(r)$ is recovered by an inverse Fourier transform of $X_W(u)$.

In our original proposal (Lanteri et al. 1998), we suggested to compare the overall effect of the Wiener procedure $\text{Abs}[\Phi(u/H(u))]$ on the modulus of the spatial frequencies to the equivalent quantity $\text{Abs}[X^{(k)}(u)/Y(u)]$ obtained for ISRA or RLA during the iteration, and to choose the value of $k$ for which these modulation transfer function best agree. In practice, depending on possible near-zero values of $H(u)$ and $Y(u)$, these quantities tend to present unwanted peaks that make difficult the computation of a distance between them.

In the present work, we propose as a measure of goodness-of-fit between the images given by the iterative algorithms and the Wiener filter, the Euclidean distances $E_{\text{W}}(k)$ between modules of spatial frequencies of the images:

$$E_{\text{W}}(k) = ||\text{Abs}[X^k(u)] - \text{Abs}[X_{\text{W}}(u)]||^2. \qquad (25)$$

This quantity gives better results than the simple Euclidean distance between the reconstructed images of ISRA (or RLA) and Wiener, because of possible negative values in the Wiener result. Indeed, the drawback of the Wiener technique is that the non-negativity of $x_{\text{W}}(r)$ is not assured. Moreover, the Wiener technique is developed in the Fourier plane and it seems reasonable to make the comparison there.

A numerical simulation is made using the data material presented in Sect. 2; for the sake of conciseness,
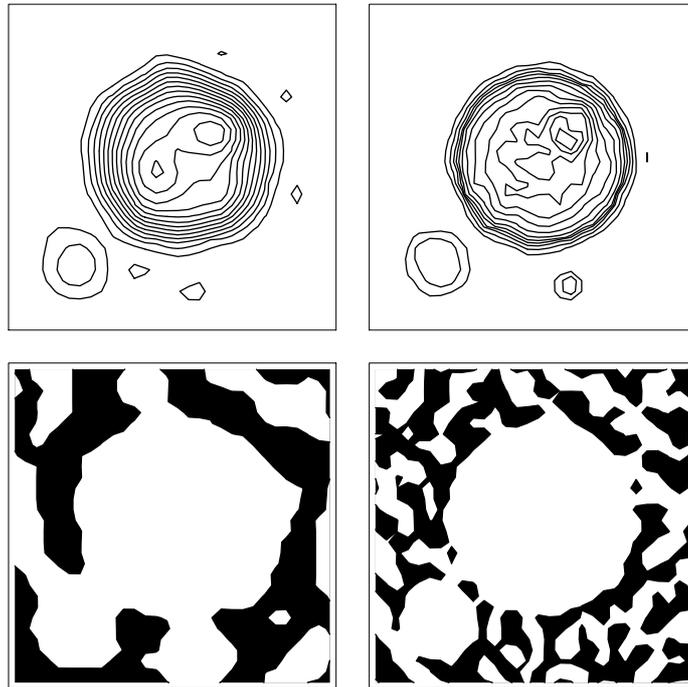
**Fig. 9.** Top: raw results of the Wiener filter for the $10^4$ photons/image (left) and the $10^6$ photons/image (right). These images are then clipped to emphasize negative parts (bottom): about 45% of the pixels are negative (represented in black). Here also only the central parts of the image are given

we limit the presentation to the case of Poisson noise. The Wiener filter is constructed taking for $P_{Hx}(u)$ the squared modulus of the Fourier transform of the noiseless blurred image. The noise power spectrum $P_N(u)$ is determined as the average noise value outside the frequency cut-off of the telescope.

Wiener images corresponding to the deconvolution of Fig. 2 are shown in Fig. 9. For the two reconstructed images in the Wiener procedure, about 45% of the reconstructed pixels are negative. The total negative part corresponds to about 7% of the positive parts for the $10^4$ photons image, and 4% for the $10^6$ photons image.

The modulus of the Fourier transform of these images is used to compute the distance $E_W(k)$ defined by Eq. (25). It is represented in Figs. 10 for ISRA and RLA, and compared to the Euclidean distance $E_O(k)$ between $x^k$ and the object. For the $10^4$ photons image, the minimum values of $E_W(k)$ and $E_O(k)$ perfectly coincides for both algorithms (iteration # 7 ISRA, iteration # 6 for RLA). For the $10^6$ photons image, the minimum values of $E_W(k)$ is obtained earlier than expected (iteration # 92 instead of # 164 for ISRA, iteration # 81 instead of # 134 for RLA). In this later case the use of the minimum value of $E_W(k)$ tends to stop the iteration prematurely, but still give a valuable order for stopping the iterations.

In Figs. 11A and 11B, we give a comparison of the overall effect produced on the spatial frequencies of the images. For each spatial frequency "$u$", the ratios
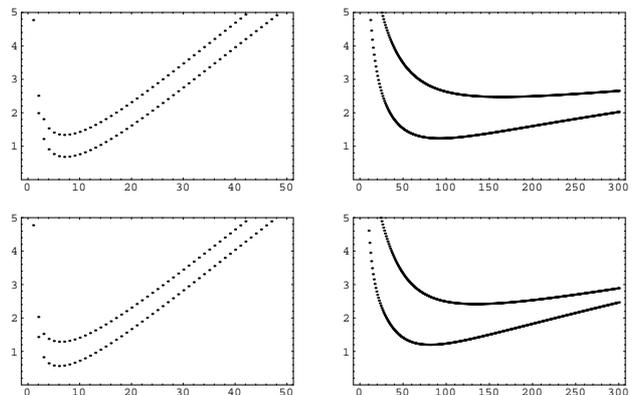


**Fig. 10.** Comparative representation of the distances $E_W(k)$, defined by Eq. (25), and $E_0(k)$ as function of the iteration number "$k$". Left frames are for the $10^4$ photons image, right ones for the $10^6$ photons image. Top frames ISRA, bottom ones: RLA. For all the representations, $E_W(k)$ is lower than $E_0(k)$. For both algorithms, the comparison with Wiener gives the correct iteration number for the $10^4$ photons image. For the $10^6$ photons image, the Wiener filter leads to a value of $k$ smaller than the optimum, then to slightly oversmoothed images

$\text{Abs}\lfloor X^{(k)}(u)/Y(u)\rfloor$ for ISRA and RLA are compared with the equivalent quantity given by the Wiener filter $\text{Abs}[X_W(u)/Y(u)]$. To obtain smoother results, we have represented, for each modulus of frequency $|u|$, the ratio of the circular average of $\text{Abs}[X^{(k)}(u)]$ and $\text{Abs}[X_W(u)]$
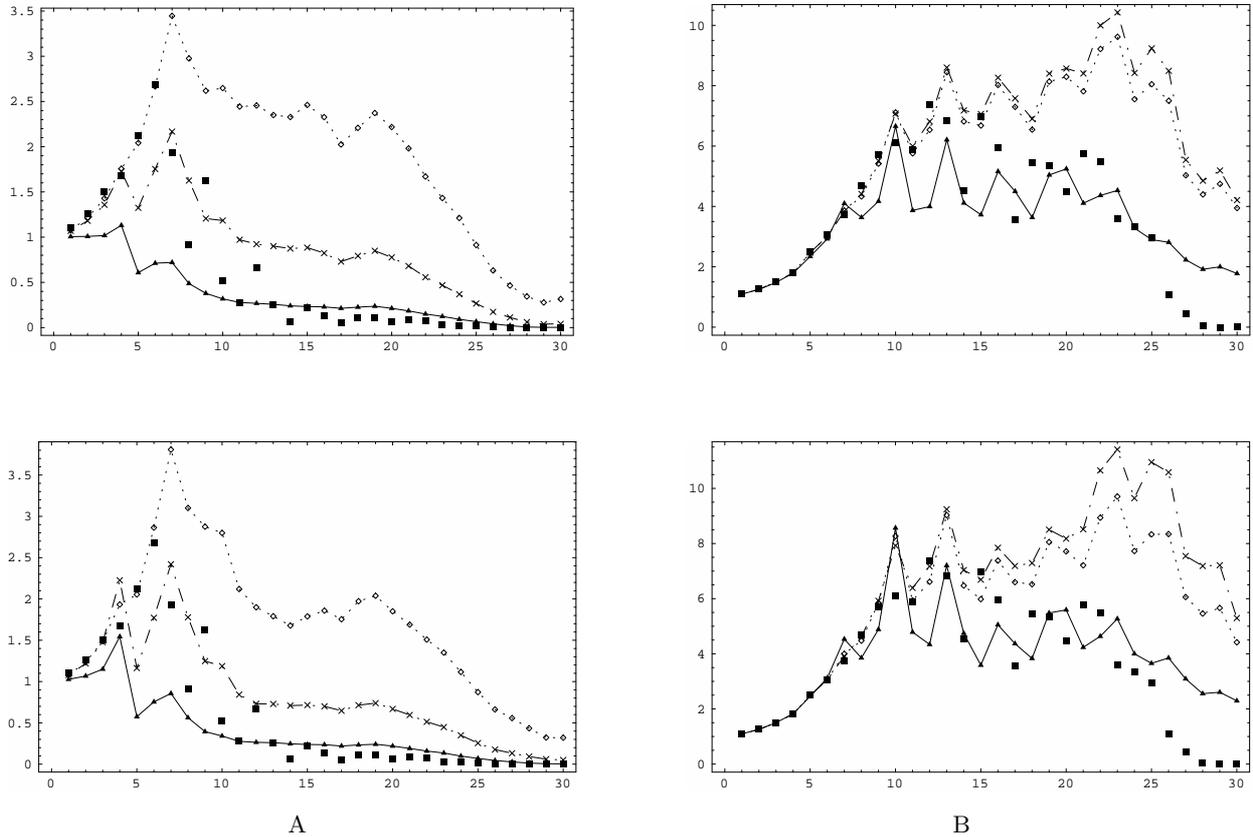
**Fig. 11.** Comparison of frequency "equivalent filters" of ISRA and RLA with the Wiener filter. The circular average of $\mathrm{Abs}\lfloor X^{(k)}(u)/Y(u)\rfloor$ obtained in the iterative deconvolution is drawn as a function of the spatial frequency "$u$", for the 3 values of $k$ of Figs. 4 and 5. They are compared to the circular average of $\mathrm{Abs}\lfloor X_{\mathrm{W}}(u)/Y(u)\rfloor$ given by the Wiener filter (individual black squares). For the iterative algorithms, the data material of Figs. 11A and 11B is that of Figs. 4 and 5, respectively. For the $10^4$ photoelectrons image (Fig. 11A), the mean frequency gain is of the order of two; it is of the order of six for the $10^6$ photoelectrons image (Fig. 11B). There is a strong similarity between ISRA and RLA results. In the very low range of frequencies, all curves tend to fit the inverse filter. In the intermediate range of frequencies, there is a rather good agreement between the best results of the iterative algorithms and the Wiener filter. In the highest frequencies, the iterative algorithms are very soon much higher than the Wiener filter. A good agreement between ISRA or RLA and Wiener in all range of frequencies would require an explicit regularization

to the equivalent circular average of $\mathrm{Abs}[Y(u)]$. The comparison is limited to the results obtained in the Poisson statistics.

For a given image, the equivalent filters obtained for ISRA and RLA are very similar. In the very low range of frequencies, all curves tend to fit the inverse filter. In the intermediate range of frequencies, there is a rather good agreement between the best results of the iterative algorithms and the Wiener filter. In the highest frequencies, the iterative algorithms are very soon much higher than the Wiener filter. A good agreement between ISRA or RLA and Wiener in all range of frequencies would require an explicit regularization.

## 4. Conclusion

We have compared the algorithm of Richardson-Lucy (RLA) and the Image Space Reconstruction Algorithm (ISRA) for image deconvolution. These iterative algorithms keep the reconstructed image non-negative while maximizing the likelihood for a Poisson Process (RLA) and a Gaussian process (ISRA). We have shown that these algorithms can be analyzed in a similar way when written as modified gradient algorithms. This form permits a better insight of the two algorithms. In particular, it reveals the similar roles played by the variances of the signal in the algorithms.

Numerical results are shown for realistic simulated images corresponding to observations made with a telescope with a small phase aberration. For Poisson noise, no strong difference was found between the results of ISRA and RLA. This later algorithm gives slightly better results than ISRA, as expected, but ISRA's results are very good too. For Gaussian noise, ISRA gives much better results than RLA.

The main drawback of these algorithms is that noise increases when the iteration number is too high. They both tend to create spectral components outside the cut-off frequency of the telescope. Although not shown in the paper, this is also observed for noise-free images. This property derives from the non-linearity of the algorithm due to the non-negativity constraint.

Some form of regularization is therefore necessary. We consider here the elementary method that consists of stopping the iterations before instabilities appear in the solution. To optimally adjust the iteration number, we propose to use a comparison of RLA and ISRA with the Wiener filter. The method is founded on the minimization of the Euclidean distance $E[k]$ between the modulus of the Fourier transform of the solution at a given iteration, and the equivalent quantity computed on the image reconstructed by a Wiener filter. In so doing, we obtain the value of $k$ corresponding, for this criterion, to the optimum regularized non negative solution.

Our assumptions have been confronted to numerical simulations for two images defined with $10^4$ and $10^6$ photons. The results are satisfactory; the comparison ISRA/Wiener or RLA/Wiener gives the exact optimal iteration number (known for simulated data) for the $10^4$ photons image. For the $10^6$ photons image, that comparison leads to an iteration number smaller than the optimum; in that case, the resulting image is slightly over regularized.

The procedure we propose can then be summarized as follows. We first construct a Wiener filter founded on physical considerations concerning the level of additive noise in the data or the number of photons in the image. This is probably the most difficult step in our reconstruction procedure. Then, depending on the statistics of the noise, we use ISRA or RLA with an iteration number that makes $E[k]$ minimum. The use of ISRA or RLA ensures a positive result; the comparison with the Wiener filter guarantees a good signal to noise ratio of the reconstructed image.

However, the simple limitation of the iteration number is not sufficient to obtain a perfect adjustment of all frequency components: a good fit in the low frequency range is always obtained with an excess of power in the high frequencies. The introduction of a penalty term (Tikhonov or entropy) will be necessary in practice to better fit the results of the Wiener filter for all frequencies.

The first tests we have made indicate that a slight improvement of the results is to be expected from that (paper in preparation).

In the present state of our technique, we cannot propose to the reader a software package that would allow an observational astronomer to immediately apply the technique to the images of the telescope. A good knowledge of the Wiener filter is necessary. However, we think that the expected results is worth the required effort in programming and physical understanding of the sources of noise.

**Appendix A:**

In the deconvolution problem, for a Gaussian additive noise and in the case of real functions we must compute the gradient of the function

$$J(x) = \frac{1}{2} \int_{-\infty}^{+\infty} (y - h \otimes x)^2 \mathrm{d}u \qquad (A1)$$

$y$, $h$, $x$ depends on $u$.

Considering for the variable $x$, an increase "$\rho$" in the general direction "$s$", we have:

$$J(x + \rho s)$$
$$= \frac{1}{2} \int_{-\infty}^{+\infty} \left[ y^2 - 2y(h \otimes (x + \rho s)) + (h \otimes (x + \rho s))^2 \right]. (A2)$$

Expanding this function and neglecting the second order terms, we have:

$$J(x + \rho s) = J(x) + \rho \int_{-\infty}^{+\infty} (h \otimes s)[(h \otimes x) - y]\mathrm{d}u \qquad (A3)$$

this expression must be identical to: $J(x + \rho s) = J(x) + \rho J'(x, s)$.

With: $H(u) = [(h \otimes x) - y](u)$, expanding the convolution product between $h$ and $s$, we have:

$$J'(x, s) = \left[ \int_{u=-\infty}^{u=+\infty} H(u) \int_{t=-\infty}^{t=+\infty} s(t)h(u-t)\mathrm{d}t\mathrm{d}u \right] \quad (A4)$$

or, inverting the integrals

$$J'(x, s) = \left[ \int_{t=-\infty}^{t=+\infty} s(t) \int_{u=-\infty}^{u=+\infty} H(u)h(u-t)\mathrm{d}u\mathrm{d}t \right] \quad (A5)$$

letting: $z = u - t$, we have:

$$J'(x, s) = \left[ \int_{t=-\infty}^{t=+\infty} s(t) \int_{z=-\infty}^{z=+\infty} H(z+t)h(z)\mathrm{d}z\mathrm{d}t \right] \quad (A6)$$

which can be written in the form:

$$J'(x, s) = \left[ \int_{t=-\infty}^{t=+\infty} s(t)\{h(-t) \otimes H(t)\}\mathrm{d}t \right]. \qquad (A7)$$

Using the expression of the derivative of $J(x)$ at $x$ in the direction $s$ we obtain:

$$\nabla J(x) = h(-u) \otimes \{[h(u) \otimes x(u)] - y(u)\}\cdot \qquad (A8)$$

**Appendix B:**

A computation analogous to the previous one can be performed in the case of a Poisson process. In this case, we must compute the gradient of:

$$J(x) = \int_{x=-\infty}^{x=+\infty} [(h \otimes x) - y\mathrm{Log}(h \otimes x)]\mathrm{d}u. \qquad (B1)$$

Considering for the variable "$x$", an increase "$\rho$" in the general direction "$s$", we have:

$$J(x + \rho s) = \int_{u=-\infty}^{u=+\infty} [\{h \otimes (x + \rho s)\}$$
$$-y\mathrm{Log}\{h \otimes (x + \rho s)\}]\mathrm{d}u \qquad (B2)$$

$$J(x + \rho s) = \int_{u=-\infty}^{u=+\infty} \left[ \{(h \otimes x) \right.$$

$$\left. + \rho(h \otimes s)\} - y \mathrm{Log} \left\{ (h \otimes x)(1 + \rho \frac{(h \otimes s)}{(h \otimes x)}) \right\} \right] \mathrm{d}u \qquad (B3)$$

$$J(x + \rho s) = \int_{u=-\infty}^{u=+\infty} \left[ (h \otimes x) + \rho(h \otimes s) - y \mathrm{Log}(h \otimes h) \right.$$

$$\left. - y \mathrm{Log} \left\{ 1 + \rho \frac{(h \otimes s)}{(h \otimes x)} \right\} \right] \mathrm{d}u \qquad (B4)$$

with a first order series expansion:

$$J(x + \rho s) = \int_{u=-\infty}^{u=+\infty} \left[ (h \otimes x) + \rho(h \otimes s) \right.$$

$$\left. - y \mathrm{Log}(h \otimes x) - y \rho \frac{(h \otimes s)}{(h \otimes x)} \right] \mathrm{d}u \qquad (B5)$$

then

$$J(x + \rho s) = \int_{u=-\infty}^{u=+\infty} \left[ (h \otimes x) - y \mathrm{Log}(h \otimes x) \right] \mathrm{d}u$$

$$+ \rho \int_{u=-\infty}^{u=+\infty} \left[ (h \otimes s) - y \frac{(h \otimes s)}{(h \otimes x)} \right] \mathrm{d}u \qquad (B6)$$

comparing this relation with: $J(x + \rho s) = J(x) + \rho J'(x, s)$ we have :

$$J'(x, s) = \int_{u=-\infty}^{u=+\infty} \left[ (h \otimes s) - y \frac{(h \otimes s)}{(h \otimes x)} \right] \mathrm{d}u \qquad (B7)$$

that is:

$$J'(x, s) = \int_{u=-\infty}^{u=+\infty} \left[ 1 - \frac{y}{(h \otimes x)} \right] (h \otimes s) \mathrm{d}u \qquad (B8)$$

with the integral form of the convolution product, we have:

$$J'(x, s) = \int_{u=-\infty}^{u=+\infty} \left[ 1 - \frac{y}{(h \otimes x)} \right]$$

$$\int_{t=-\infty}^{t=+\infty} s(t) h(u - t) \mathrm{d}t \mathrm{d}u \qquad (B9)$$

or, re-ordering the integrals:

$$J'(x, s) = \int_{t=-\infty}^{t=+\infty} s(t)$$

$$\int_{u=-\infty}^{u=+\infty} \left[ 1 - \frac{y}{(h \otimes x)} \right] h(u - t) \mathrm{d}u \mathrm{d}t. \qquad (B10)$$

Let:

$$H(u) = \left[ 1 - \frac{y(u)}{(h(u) \otimes x(u))} \right] \qquad (B11)$$

with the change of variables: $z = u - t$, we obtain:

$$J'(x, s) = \int_{t=-\infty}^{t=+\infty} s(t) \int_{z=-\infty}^{z=+\infty} H(t + z) h(z) \mathrm{d}z \mathrm{d}t \qquad (B12)$$

that is:

$$J'(x, s) = \int_{t=-\infty}^{t=+\infty} s(t)[h(-t) \otimes H(t)] \mathrm{d}t. \qquad (B13)$$

Using the definition of the derivative of $J(x)$ at $x$ in the direction $s$ we obtain:

$$\nabla J(x) = h(-u) \otimes \left[ 1 - \frac{y(u)}{h(u) \otimes x(u)} \right]. \qquad (B14)$$

## References

Bertero M., De Mol C., Viano G.A., 1980, The stability of inverse problems, in Baltes H.P. (ed.), Inverse scattering problems in optics. Springer Verlag series Topics in current physics

Bertero M., 1989, Linear inverses and ill-posed problems, in Advances in Electronics and Electron Physics 75, 1

Bertero M., Boccacci P., Maggio F., 1995, Int. J. Imaging Systems and Technology 6, 376

Brault J.W., White O.R., 1971, A&A 13, 169

Byrne C.L., 1993, I.E.E.E. Trans. Image Processing 2, 96

Csiszär I., 1991, The annals of statistics 19, 2032

Daube-Witherspoon M.E., Muehllehner G., 1986, I.E.E.E. Trans. Med. Imaging MI-5, 61

Demoment G., 1986, Deconvolution des signaux, Cours de l'École Supérieure d'Électricité

De Pierro A.R., 1987, I.E.E.E. Trans. Med. Imaging MI-6, 124

De Pierro A.R., 1990, J. Optimization Theory Appl. 64, 87

De Pierro A.R., 1993, I.E.E.E. Trans. Med. Imaging MI-12, 328

Gilliland R.L., Dupree A.K., 1996, Astrophys. J. 463, L29

Lane R.G., 1996, J. Opt. Soc. Am. A 13, 1992

Lantéri H., Soummer R., Aime C., 1998, ESO/OSA Topical Meeting on Astronomy with Adaptive Optics: Present Results and Future Programs, 7-11 September 1998, Sonthofen, Germany

Lascaux P., Theodor R., 1986, Analyse numérique matricielle appliquée à l'art de l'ingénieur, T.2, Masson Ed. (Paris) p. 409

Lucy L.B., 1974, Astron. J. 79, 745

Lucy L.B., 1994, A&A 289, 983

Luenberger D.G., 1973, Introduction to linear and non linear programming. Addison-Wesley Pub. Co., Reading, Mass, U.S.A.

Nuñez J., Llacer J., 1993, PASP 105, 1192

Richardson W.H., 1972, J. Opt. Soc. Am. 62, 55

Shepp L.A., Vardi Y., 1982, I.E.E.E. Trans. Med. Imaging MI-6, 113

Snyder D.L., Schultz T.J., O'Sullivan J.A., 1992, IEEE Trans. Signal Processing 40, 1143

Titterington D.M., 1987, IEEE Trans. Med. Imaging MI-6, 52

Wolfram S., 1996, The Mathematica Book, Third Edition. Wolfram Media Inc. And Cambridge University Press

Zacchéo T.S., Gonsalves R.A., 1996, J. Opt. Soc. Am. A. 13, 236