

1

Deconvolution and Blind Deconvolution in Astronomy

CONTENTS

1.1	Introduction	1
1.2	The Deconvolution Problem	3
1.3	Linear Regularized Methods	5
1.4	CLEAN	6
1.5	Bayesian Methodology	6
1.6	Iterative Regularized Methods	11
1.7	Wavelet-Based Deconvolution	13
1.8	Deconvolution and Resolution	21
1.9	Myopic and Blind Deconvolution	22
1.10	Conclusions and Chapter Summary	26
	Acknowledgements	27
	References	27

1.1 Introduction

Deconvolution is a key area in signal and image processing. It can include deblurring of an observed signal to remove atmospheric effects. More generally, it means correcting for instrumental effects or observing conditions.

Research in image deconvolution has recently seen considerable work, partly triggered by the HST optical aberration problem at the beginning of its mission that motivated astronomers to improve current algorithms or develop new and more effective ones. Since then, deconvolution of astronomical images has proven in some cases to be crucial for extracting scientific content. For instance, IRAS images can be effectively reconstructed thanks to a new pyramidal maximum entropy algorithm [8]. Io volcanism can be studied with a lower resolution of 0.15 arcsec, or 570 km on Io [58]. Deconvolved mid-infrared images at 20 μm revealed the innermost structure of the AGN in NGC1068 surrounding the giant blackhole, some of which enshrouded emission regions unobservable at lower wavelength because of higher dust extinction [3]: see Fig. 1.1. Research on gravitational lenses is easier and more effective when applying deconvolution methods [26].

Deconvolution will be even more crucial in the future in order to fully take ad-

vantage of increasing numbers of high-quality ground-based telescopes, for which images are strongly limited in resolution by the seeing.

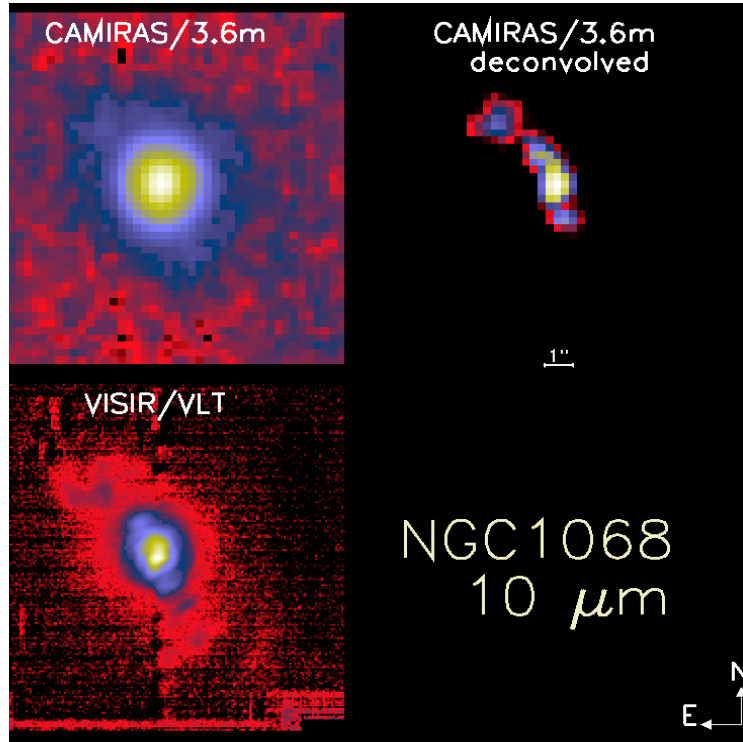


FIGURE 1.1

The active galaxy nucleus of NGC1068 observed at 20 μm . Upper left: the raw image is highly blurred by telescope diffraction (1.4 arcsec resolution on a 3.6m telescope) and seeing. Right: the restored image using the multiscale entropy method reveals the inner structure in the vicinity of the nucleus. Lower left: the same object observed later on a larger telescope (VLT, 8m-class telescope) whose sharper resolution confirms the structures found previously by image deconvolution.

HST provided a leading example of the need for deconvolution, in the period before the detector system was refurbished. Two proceedings [98, 43] provide useful overviews of this work, and a later reference is [2]. While an atmospheric seeing point spread function (PSF) may be relatively tightly distributed around the mode, this was not the case for the spherically aberrated HST PSF. Whenever the PSF “wings” are extended and irregular, deconvolution offers a straightforward way to

mitigate the effects of this and to upgrade the core region of a point source. One usage of deconvolution of continuing importance is in information fusion from different detectors. For example, Faure et al. [33] deconvolve HST images when correlating with ground-based observations. In Radomski et al. [75], Keck data are deconvolved, for study with HST data. VLT (Very Large Telescope) data are deconvolved in [12], with other ESO and HST data used as well. In planetary work, Coustenis et al. [27] discuss CFHT data as well as HST and other observations.

What emerges very clearly from this small sample – which is in no way atypical – is that a major use of deconvolution is to help in cross-correlating image and signal information.

An observed signal is never in pristine condition, and improving it involves inverting the spoiling conditions, i.e. finding a solution to an inverse equation. The problem is usually an ill-conditioned one, which leads to the need for constraints on what is to be an acceptable solution. Constraints related to the type of signal we are dealing with play an important role in the development of effective and efficient algorithms. The use of constraints to provide for a stable and unique solution is termed regularization.

Our review opens in section 1.2 with a formalization of the problem. Section 1.3 considers the issue of regularization. In section 1.4, the CLEAN method which is central to radio astronomy is described. Bayesian modeling and inference in deconvolution is reviewed in section 1.5.

Section 1.6 further considers regularization, surveying more complex and powerful regularization methods. Section 1.7 introduces wavelet-based methods as used in deconvolution. These methods are based on multiple resolution or scale. In sections 1.8 and 1.9, important issues related to resolution of the output result image are discussed. Section 1.8 is based on the fact that it is normally not worthwhile to target an output result with better resolution than some limit, for instance a pixel size. Section 1.9 investigates when, where and how missing information can be inferred to provide a super-resolution output image.

1.2 The Deconvolution Problem

Consider an image characterized by its intensity distribution (the “data”) I , corresponding to the observation of a “real image” O through an optical system. If the imaging system is linear and shift-invariant, the relation between the data and the image in the same coordinate frame is a convolution:

$$\begin{aligned} I(x, y) &= \int_{x_1=-\infty}^{+\infty} \int_{y_1=-\infty}^{+\infty} P(x - x_1, y - y_1) O(x_1, y_1) dx_1 dy_1 \\ &\quad + N(x, y) \\ &= (P * O)(x, y) + N(x, y) \end{aligned} \tag{1.1}$$

P is the point spread function, PSF, of the imaging system, and N is additive noise.

In Fourier space we have:

$$\hat{I}(u, v) = \hat{O}(u, v)\hat{P}(u, v) + \hat{N}(u, v) \quad (1.2)$$

We want to determine $O(x, y)$ knowing I and P . This inverse problem has led to a large amount of work, the main difficulties being the existence of: (i) a cut-off frequency of the point spread function, and (ii) the additive noise. See for example [25, 52, 6, 63].

A solution can be obtained by computing the Fourier transform of the deconvolved object \hat{O} by a simple division between the image \hat{I} and the PSF \hat{P}

$$\hat{\hat{O}}(u, v) = \frac{\hat{I}(u, v)}{\hat{P}(u, v)} = \hat{O}(u, v) + \frac{\hat{N}(u, v)}{\hat{P}(u, v)} \quad (1.3)$$

This method, sometimes called the *Fourier-quotient method* is very fast. We only need to do a Fourier transform and an inverse Fourier transform. For frequencies close to the frequency cut-off, the noise term becomes important, and the noise is amplified. Therefore in the presence of noise, this method cannot be used.

Eqn. 1.1 is usually in practice an ill-posed problem. This means that there is no unique and stable solution.

Other topics related to deconvolution are:

- Super-resolution: object spatial frequency information outside the spatial bandwidth of the image formation system is recovered.
- Blind deconvolution: the PSF P is unknown.
- Myopic deconvolution: the PSF P is partially known.
- Image reconstruction: an image is formed from a series of projections (computed tomography, positron emission tomography or PET, and so on).

We will discuss the first three points in this chapter.

In the deconvolution problem, the PSF is assumed to be known. In practice, we have to construct a PSF from the data, or from an optical model of the imaging telescope. In astronomy, the data may contain stars, or one can point towards a reference star in order to reconstruct a PSF. The drawback is the “degradation” of this PSF because of unavoidable noise or spurious instrument signatures in the data. So, when reconstructing a PSF from experimental data, one has to reduce very carefully the images used (background removal for instance) or otherwise any spurious feature in the PSF would be repeated around each object in the deconvolved image. Another problem arises when the PSF is highly variable with time, as is the case for adaptive optics images. This means usually that the PSF estimated when observing a reference star, after or before the observation of the scientific target, has small differences from the perfect one. In this particular case, one has to turn towards myopic deconvolution methods [20] in which the PSF is also estimated in the iterative algorithm using a first guess deduced from observations of reference stars.

Another approach consists of constructing a synthetic PSF. Several studies [10, 59, 31, 64] have suggested a radially symmetric approximation to the PSF:

$$P(r) \propto \left(1 + \frac{r^2}{R^2}\right)^{-\beta} \quad (1.4)$$

The parameters β and R are obtained by fitting the model with stars contained in the data.

1.3 Linear Regularized Methods

1.3.1 Least Squares Solution

It is easy to verify that the minimization of $\|I(x, y) - P(x, y) * O(x, y)\|^2$ leads to the solution:

$$\hat{O}(u, v) = \frac{\hat{P}^*(u, v)\hat{I}(u, v)}{|\hat{P}(u, v)|^2} \quad (1.5)$$

which is defined only if $\hat{P}(u, v)$ is different from zero. Here \hat{P}^* is the complex conjugate of \hat{P} . The problem is generally ill-posed and we need to introduce *regularization* in order to find a unique and stable solution.

1.3.2 Tikhonov Regularization

Tikhonov regularization [94] consists of minimizing the term:

$$J_T(O) = \|I(x, y) - (P * O)(x, y)\|^2 + \lambda \|H * O\|^2 \quad (1.6)$$

where H corresponds to a high-pass filter, and $\|\cdot\|^2$ is norm squared. This criterion contains two terms. The first, $\|I(x, y) - (P * O)(x, y)\|^2$, expresses fidelity to the data $I(x, y)$, and the second, $\lambda \|H * O\|^2$, expresses smoothness of the restored image. λ is the regularization parameter and represents the trade-off between fidelity to the data and the smoothness of the restored image.

The solution is obtained directly in Fourier space:

$$\hat{O}(u, v) = \frac{\hat{P}^*(u, v)\hat{I}(u, v)}{|\hat{P}(u, v)|^2 + \lambda |\hat{H}(u, v)|^2} \quad (1.7)$$

Finding the optimal value λ necessitates use of numerical techniques such as cross-validation [40, 38]. This method works well, but computationally it is relatively lengthy and produces smoothed images. This second point can be a real problem when we seek compact structures such as is the case in astronomical imaging.

This method can be generalized such that

$$\hat{\tilde{O}}(u, v) = \hat{W}(u, v) \frac{\hat{I}(u, v)}{\hat{P}(u, v)} \quad (1.8)$$

W being a window function satisfying some conditions [6]. These window function can be for instance the usual Gaussian, Blackmann, Hamming or Hanning functions. The window function can also be derived directly from the PSF [72].

Linear regularized methods have the advantage of being very attractive from a computation point of view. Furthermore, the noise in the solution can easily be derived from the noise in the data and the window function. For example, if the noise in the data is Gaussian with a standard deviation σ_d , the noise in the solution is $\sigma_s^2 = \sigma_d^2 \sum W_k^2$. But this noise estimation does not take into account errors relative to inaccurate knowledge of the PSF, which limits its interest in practice.

Linear regularized methods present also a number of severe drawbacks:

- Creation of Gibbs oscillations in the neighborhood of the discontinuities contained in the data. The visual quality is therefore degraded.
- No a priori information can be used. For example, negative values can exist in the solution, while in most cases we know that the solution must be positive.
- Since the window function is a low-pass filter, the resolution is degraded. There is trade-off between the resolution we want to achieve and the noise level in the solution. Other methods such as wavelet-based methods do not have such a constraint.

1.4 CLEAN

The CLEAN method [45] is a mainstream one in radio astronomy. This approach assumes the object is only composed of point sources. It tries to decompose the image (called the dirty map) into a set of δ -functions. This is done iteratively by finding the point with the largest absolute brightness and subtracting the PSF (dirty beam) scaled with the product of the loop gain and the intensity at that point. The resulting residual map is then used to repeat the process. The process is stopped when some prespecified limit is reached. The convolution of the δ -functions with an ideal PSF (clean beam) plus the residual equals the restored image (clean map). This solution is only possible if the image does not contain large-scale structures.

In the work of [16] and [51], the restoration of an object composed of peaks, called *sparse spike trains*, has been treated in a rigorous way.

1.5 Bayesian Methodology

A Bayesian formulation may well be just another way of describing the classical minimization problem with constraints, as seen in section 1.3. However it is important for the following reasons: it incorporates procedures for the estimation of parameters; and prior information on the values of such parameters can also be incorporated easily.

1.5.1 Definition

The Bayesian approach consists of constructing the conditional probability density relationship:

$$p(O | I) = \frac{p(I | O)p(O)}{p(I)} \quad (1.9)$$

The Bayes solution is found by maximizing the right part of the equation. Now since we are maximizing $p(O | I)$, the image I is always the same for the maximization, and so $p(I)$ is constant here. The maximum likelihood solution (ML) maximizes only the density $p(I | O)$ over O :

$$\text{ML}(O) = \max_O p(I | O) \quad (1.10)$$

The maximum a posteriori solution (MAP) maximizes over O the product $p(I | O)p(O)$ of the ML and a prior:

$$\text{MAP}(O) = \max_O p(I | O)p(O) \quad (1.11)$$

$p(I)$ is considered as a constant value which has no effect in the maximization process, and is ignored. The ML solution is equivalent to the MAP solution assuming a uniform probability density for $p(O)$.

1.5.2 Maximum Likelihood with Gaussian Noise

The probability $p(I | O)$ is

$$p(I | O) = \frac{1}{\sqrt{2\pi}\sigma_I} \exp - \frac{(I - P * O)^2}{2\sigma_I^2} \quad (1.12)$$

where σ_I^2 is image variance and, assuming that $p(O)$ is a constant, maximizing $p(O | I)$ is equivalent to minimizing

$$J(O) = \frac{\|I - P * O\|^2}{2\sigma_I^2} \quad (1.13)$$

where the denominator is unimportant for the optimum value of J . Using the steepest descent minimization method, a typical iteration is

$$O^{n+1} = O^n + \gamma P^* * (I - P * O^n) \quad (1.14)$$

where $P^*(x, y) = P(-x, -y)$. P^* is the transpose of the PSF, and O^n is the current estimate of the desired “real image”. This method is usually called the Landweber method [53], but sometimes also the *successive approximations* or Jacobi method [6].

The solution can also be found directly using the FFT by

$$\hat{O}(u, v) = \frac{\hat{P}^*(u, v) \hat{I}(u, v)}{\hat{P}^*(u, v) \hat{P}(u, v)} \quad (1.15)$$

which amounts to the same as equation 1.3. Such a straightforward approach unfortunately amplifies noise.

1.5.3 Gaussian Bayes Model

If the object and the noise are assumed to follow Gaussian distributions with zero mean and variance respectively equal to σ_O^2 and σ_N^2 , then a Bayes solution leads to the Wiener filter:

$$\hat{O}(u, v) = \frac{\hat{P}^*(u, v) \hat{I}(u, v)}{|\hat{P}(u, v)|^2 + \frac{\sigma_N^2(u, v)}{\sigma_O^2(u, v)}} \quad (1.16)$$

Wiener filtering has serious drawbacks (artifact creation such as ringing effects), and needs spectral noise estimation. Its advantage is that it is very fast.

1.5.4 Maximum Likelihood with Poisson Noise

The probability $p(I | O)$ is

$$p(I | O) = \prod_{x, y} \frac{((P * O)(x, y))^{I(x, y)} \exp\{-(P * O)(x, y)\}}{I(x, y)!} \quad (1.17)$$

The maximum can be computed by taking the derivative of the logarithm:

$$\frac{\partial \ln p(I | O)(x, y)}{\partial O(x, y)} = 0 \quad (1.18)$$

which leads to the result (assuming the PSF is normalized to unity)

$$\left[\frac{I}{P * O} * P^* \right] (x, y) = 1 \quad (1.19)$$

Multiplying both sides by $O(x, y)$

$$O(x, y) = \left[\frac{I}{(P * O)} * P^* \right] (x, y) O(x, y) \quad (1.20)$$

and using Picard iteration [47] leads to

$$O^{n+1}(x, y) = \left[\frac{I}{(P * O^n)} * P^* \right] (x, y) O^n(x, y) \quad (1.21)$$

which is the Richardson-Lucy algorithm [76, 55, 80], also sometimes called the *expectation maximization* or EM method [29]. This method is commonly used in astronomy. Flux is preserved and the solution is always positive. The positivity of the solution can be obtained too with Van Cittert's and the one-step gradient methods by thresholding negative values in O^n at each iteration.

1.5.5 Maximum A Posteriori with Poisson Noise

We formulate the object PDF (probability density function) as

$$p(O) = \prod_{x,y} \frac{M(x, y)^{O(x, y)} \exp\{-M(x, y)\}}{O(x, y)!} \quad (1.22)$$

The MAP solution is

$$O(x, y) = M(x, y) \exp \left\{ \left[\frac{I(x, y)}{(P * O)(x, y)} - 1 \right] * P^*(x, y) \right\} \quad (1.23)$$

and choosing $M = O^n$ and using Picard iteration leads to

$$O^{n+1}(x, y) = O^n(x, y) \exp \left\{ \left[\frac{I(x, y)}{(P * O^n)(x, y)} - 1 \right] * P^*(x, y) \right\} \quad (1.24)$$

1.5.6 Maximum Entropy Method

In the absence of any information on the solution O except its positivity, a possible course of action is to derive the probability of O from its entropy, which is defined from information theory. Then if we know the entropy H of the solution, we derive its probability as

$$p(O) = \exp(-\alpha H(O)) \quad (1.25)$$

The most commonly used entropy functions are:

- Burg [11]: $H_b(O) = -\sum_x \sum_y \ln(O(x, y))$
- Frieden [37]: $H_f(O) = -\sum_x \sum_y O(x, y) \ln(O(x, y))$

- Gull and Skilling [42]:

$$H_g(O) = \sum_x \sum_y O(x,y) - M(x,y) - O(x,y) \ln(O(x,y)/M(x,y))$$

The last definition of the entropy has the advantage of having a zero maximum when O equals the model M , usually taken as a flat image.

1.5.7 Other Regularization Models

Molina et al. [63] present an excellent review of taking the spatial context of image restoration into account. Some appropriate prior is used for this. One such regularization constraint is:

$$\|CI\|^2 = \sum_x \sum_y I(x,y) - \frac{1}{4}(I(x,y+1) + I(x,y-1) + I(x+1,y) + I(x-1,y)) \quad (1.26)$$

which is equivalent to defining the prior

$$p(O) \propto \exp\left\{-\frac{\alpha}{2}\|CI\|^2\right\} \quad (1.27)$$

α is the inverse of the prior variance and it controls the smoothness of the final solution.

Given the form of equation (1.26), such regularization can be viewed as setting a constraint on the Laplacian of the restoration. In statistics this model is a simultaneous autoregressive model, SAR [77].

Alternative prior models can be defined, related to the SAR model of equation (1.26):

$$p(O) \propto \exp\left\{-\frac{\alpha}{2} \sum_x \sum_y (I(x,y) - I(x,y+1))^2 + (I(x,y) - I(x+1,y))^2\right\} \quad (1.28)$$

where constraints are set on first derivatives. This is a conditional autoregressive, or CAR, model, discussed further below.

Blanc-Feraud and Barlaud [7], and Charbonnier et al. [17] consider the following prior:

$$p(O) \propto \exp\left\{-\alpha \sum_x \sum_y \phi(\|\nabla I\|)(x,y)\right\} \quad (1.29)$$

$$\propto \exp\left\{-\alpha \sum_x \sum_y \phi(I(x,y) - I(x,y+1))^2 + \phi(I(x,y) - I(x+1,y))^2\right\} \quad (1.30)$$

The function ϕ , called *potential function*, is an edge preserving function.

Generally, ϕ functions are chosen with a quadratic part which ensures good smoothing of small gradients [41], and a linear behavior which cancels the penalization of large gradients [9]:

1. $\lim_{t \rightarrow 0} \frac{\phi'(t)}{2t} = M < +\infty$ to smooth faint gradients.
2. $\lim_{t \rightarrow \infty} \frac{\phi'(t)}{2t} = 0$ to preserve strong gradients.
3. $\frac{\phi'(t)}{2t}$ is continuous and strictly decreasing.

Such functions are often called L_2 - L_1 functions. Examples of ϕ functions:

1. $\phi_q(x) = x^2$: quadratic function.
2. $\phi_{TV}(x) = |x|$: Total Variation.
3. $\phi_2(x) = 2\sqrt{1+x^2} - 2$: Hyper-Surface [18].
4. $\phi_3(x) = x^2/(1+x^2)$ [39].
5. $\phi_4(x) = 1 - e^{-x^2}$ [71].
6. $\phi_5(x) = \log(1+x^2)$ [44].

The ARTUR method [17], which has been used for helioseismic inversion [24], uses the function $\phi(t) = \log(1+t^2)$. Anisotropic diffusion [71, 4] uses similar functions, but in this case the solution is computed using *partial differential equations*.

The function $\phi(t) = t$ leads to the *total variation* method [78, 1], the constraints are on first derivatives, and the model is a special case of a conditional autoregressive or CAR model. Molina et al. [63] discuss the applicability of CAR models to image restoration involving galaxies. They argue that such models are particularly appropriate for the modeling of luminosity exponential and $r^{1/4}$ laws.

The priors reviewed above can be extended to more complex models. In Molina et al. [61, 62], a compound Gauss Markov random field (CGMRF) model is used, one of the main properties of which is to target the preservation and improvement of line processes.

Another prior again was used in Molina and Cortijo [60] for the case of planetary images.

1.6 Iterative Regularized Methods

1.6.1 Constraints

We assume now that there exists a general operator, $\mathcal{P}_{\mathcal{C}}(\cdot)$, which enforces a set of constraints on a given object O , such that if O satisfies all the constraints, we have: $O = \mathcal{P}_{\mathcal{C}}(O)$. Commonly used constraints are positivity, spatial support constraint, band-limitation in Fourier space. These constraints can be incorporated easily in the basic iterative scheme.

1.6.2 Jansson-Van Cittert Method

Van Cittert [21] restoration is relatively easy to write. We start with $n = 0$ and $O^{(0)} = I$ and we iterate:

$$O^{n+1} = O^n + \alpha(I - P * O^n) \quad (1.31)$$

where α is a convergence parameter generally taken as 1. When n tends to infinity, we have $O = O + I - P * O$, so $I = P * O$. In Fourier space, the convolution product becomes a product

$$\hat{O}^{n+1} = \hat{O}^n + \alpha(\hat{I} - \hat{P}\hat{O}^n) \quad (1.32)$$

In this equation, the object distribution is modified by adding a term proportional to the residual. The algorithm converges quickly, often after only 5 or 6 iterations. But the algorithm generally diverges in the presence of noise. Jansson [49] modified this technique in order to give it more robustness by considering constraints on the solution. If we wish that $A \leq O_k \leq B$, the iteration becomes

$$O^{n+1}(x, y) = O^n(x, y) + r(x, y) [I - (P * O^n)](x, y) \quad (1.33)$$

with:

$$r(x, y) = C [1 - 2(B - A)^{-1} | O^n(x, y) - 2^{-1}(A + B) |]$$

and with C constant.

More generally the constrained Van-Cittert method is written as:

$$O^{n+1} = \mathcal{P}_C [O^n + \alpha(I - P * O^n)] \quad (1.34)$$

1.6.3 Other Iterative Methods

Other iterative methods can be constrained in the same way:

- Landweber:

$$O^{n+1} = \mathcal{P}_C [O^n + \gamma P^*(I - P * O^n)] \quad (1.35)$$

- Richardson-Lucy method:

$$O^{n+1}(x, y) = \mathcal{P}_C \left[O^n(x, y) \left[\frac{I}{(P * O^n)} * P^* \right](x, y) \right] \quad (1.36)$$

- Tikhonov: the Tikhonov solution can be obtained iteratively by computing the gradient of equation (1.6):

$$\nabla(J_T(O)) = [P^* * P + \mu H^* * H] * O - P^* * I \quad (1.37)$$

and applying the following iteration:

$$O^{n+1}(x, y) = O^n(x, y) - \gamma \nabla(J_T(O))(x, y) \quad (1.38)$$

The constrained Tikhonov solution is therefore obtained by:

$$O^{n+1}(x, y) = \mathcal{P}_C[O^n(x, y) - \gamma \nabla(J_T(O))(x, y)] \quad (1.39)$$

The number of iterations plays an important role in these iterative methods. Indeed, the number of iterations can be discussed in relation to regularization. When the number of iterations increases, the iterates first approach the unknown object, and then potentially go away from it [6]. Landweber and Richardson-Lucy methods converge to useless solutions due to noise amplification. The main problem is that the space of solutions of these methods is almost infinite. Regularized methods force the solution obtained at convergence to be constrained to images having the desirable properties of the original (unknown) image, i.e. being non-noisy.

1.7 Wavelet-Based Deconvolution

1.7.1 Introduction

Deconvolution and Fourier domain.

The Fourier domain diagonalizes the convolution operator, and we can identify and reduce the noise which is amplified during the inversion. When the signal can be modeled as stationary and Gaussian, the Wiener filter is optimal. But when the signal presents spatially localized features such as singularities or edges, these features cannot be well-represented with Fourier basis functions, which extend over the entire spatial domain. Other basis functions, such as wavelets, are better-suited to represent a large class of signals.

Towards multiresolution.

The concept of multiresolution was first introduced for deconvolution by Wakker and Schwarz [96] when they proposed the Multiresolution CLEAN algorithm for interferometric image deconvolution. During the last ten years, many developments have taken place in order to improve the existing methods (CLEAN, Landweber, Lucy, MEM, and so on), and these results have led to the use of different levels of resolution.

The Lucy algorithm was modified [56] in order to take into account a priori information about stars in the field where both position and brightness are known. This is done by using a two-channel restoration algorithm, one channel representing the contribution relative to the stars, and the second to the background. A smoothness constraint is added on the background channel. This method was then refined firstly

(and called *CPLUCY*) for considering subpixel positions [46], and a second time [73] (and called *GIRA*) for modifying the smoothness constraint.

A similar approach has been followed by Magain [57], but more in the spirit of the CLEAN algorithm. Again, the data are modeled as a set of point sources on top of spatially varying background, leading to a two-channel algorithm.

The MEM method has also been modified by several authors [97, 8, 70, 69, 90]. First, Weir proposed the *Multi-channel MEM* method, in which an object is modeled as the sum of objects at different levels of resolution. The method was then improved by Bontekoe et al. [8] with the *Pyramid MEM*. In particular, many regularization parameters were fixed by the introduction of the dyadic pyramid. The link between *Pyramid MEM* and wavelets was underlined in [70, 90], and it was shown that all the regularization parameters can be derived from the noise modeling. Wavelets were also used in [69] in order to create a segmentation of the image, each region being then restored with a different smoothness constraint, depending on the resolution level where the region was found. This last method has however the drawback of requiring user interaction for deriving the segmentation threshold in the wavelet space.

The *Pixon* method [30, 74] is relatively different to the previously described methods. This time, an object is modeled as the sum of pseudo-images smoothed locally by a function with position-dependent scale, called the *pixon* shape function. The set of pseudo-images defines a dictionary, and the image is supposed to contain only features included in this dictionary. But the main problem lies in the fact that features which cannot be detected directly in the data, nor in the data after a few Lucy iterations, will not be modeled with the *pixon* functions, and will be strongly regularized as background. The result is that the faintest objects are over-regularized while strong objects are well restored. This is striking in the example shown in Fig. 1.2.

The *total variation* method has a close relation with the Haar transform [22, 92], and more generally, it has been shown that potential functions, used in Markov Random Field and PDE methods, can be applied on the wavelet coefficients as well. This leads to multiscale regularization, and the original method becomes a specific case where only one decomposition level is used in the wavelet transform.

Wavelets offer a mathematical framework for the multiresolution processing. Furthermore, they furnish an ideal way to include noise modeling in the deconvolution methods. Since the noise is the main problem in deconvolution, wavelets are very well adapted to the regularization task.

1.7.2 Regularization from the Multiresolution Support

1.7.2.1 Noise suppression based on the wavelet transform

We have noted how, in using an iterative deconvolution algorithm such as Van Cittert or Richardson-Lucy, we define $R^{(n)}(x, y)$, the residual at iteration n :

$$R^n(x, y) = I(x, y) - (P * O^n)(x, y) \quad (1.40)$$

By using the à trous wavelet transform algorithm, R^n can be defined as the sum of its J wavelet scales and the last smooth array:

$$R^n(x, y) = c_J(x, y) + \sum_{j=1}^J w_{j,x,y} \quad (1.41)$$

where the first term on the right is the last smoothed array, and w denotes a wavelet scale.

The wavelet coefficients provide a mechanism to extract only the significant structures from the residuals at each iteration. Normally, a large part of these residuals is statistically non-significant. The significant residual [67, 85] is then:

$$\bar{R}^n(x, y) = c_{J,x,y} + \sum_{j=1}^J M(j, x, y) w_{j,x,y} \quad (1.42)$$

where $M(j, x, y)$ is the multiresolution support, and is defined by:

$$M(j, x, y) = \begin{cases} 1 & \text{if } w_{j,x,y} \text{ is significant} \\ 0 & \text{if } w_{j,x,y} \text{ is non-significant} \end{cases} \quad (1.43)$$

This describes in a logical or Boolean way if the data contains information at a given scale j and at a given position (x, y) .

An alternative approach was outlined in [68] and [83]: the support was initialized to zero, and built up at each iteration of the restoration algorithm. Thus in equation (1.42) above, $M(j, x, y)$ was additionally indexed by n , the iteration number. In this case, the support was specified in terms of significant pixels at each scale, j ; and in addition pixels could become significant as the iterations proceeded, but could not be made non-significant.

1.7.2.2 Regularization of Van Cittert's algorithm

Van Cittert's iteration [21] is:

$$O^{n+1}(x, y) = O^n(x, y) + \alpha R^n(x, y) \quad (1.44)$$

with $R^n(x, y) = I^n(x, y) - (P * O^n)(x, y)$. Regularization using significant structures leads to:

$$O^{n+1}(x, y) = O^n(x, y) + \alpha \bar{R}^n(x, y) \quad (1.45)$$

The basic idea of this regularization method consists of detecting, at each scale, structures of a given size in the residual $R^n(x, y)$ and putting them in the restored image $O^n(x, y)$. The process finishes when no more structures are detected. Then, we have separated the image $I(x, y)$ into two images $\tilde{O}(x, y)$ and $R(x, y)$. \tilde{O} is the restored image, which ought not to contain any noise, and $R(x, y)$ is the final residual which ought not to contain any structure. R is our estimate of the noise $N(x, y)$.

1.7.2.3 Regularization of the one-step gradient method

The one-step gradient iteration is:

$$O^{n+1}(x, y) = O^n(x, y) + (P^* * R^n)(x, y) \quad (1.46)$$

with $R^n(x, y) = I(x, y) - (P * O^n)(x, y)$. Regularization by significant structures leads to:

$$O^{n+1}(x, y) = O^n(x, y) + (P^* * \bar{R}^n)(x, y) \quad (1.47)$$

1.7.2.4 Regularization of the Richardson-Lucy algorithm

From equation (1.1), we have $I^n(x, y) = (P * O^n)(x, y)$. Then $R^n(x, y) = I(x, y) - I^n(x, y)$, and hence $I(x, y) = I^n(x, y) + R^n(x, y)$.

The Richardson-Lucy equation is:

$$O^{n+1}(x, y) = O^n(x, y) \left[\frac{I^n + R^n}{I^n} * P^* \right](x, y)$$

and regularization leads to:

$$O^{n+1}(x, y) = O^n(x, y) \left[\frac{I^n + \bar{R}^n}{I^n} * P^* \right](x, y)$$

1.7.2.5 Convergence

The standard deviation of the residual decreases until no more significant structures are found. Convergence can be estimated from the residual. The algorithm stops when a user-specified threshold is reached:

$$(\sigma_{R^{n-1}} - \sigma_{R^n}) / (\sigma_{R^n}) < \varepsilon \quad (1.48)$$

1.7.2.6 Examples

A simulated Hubble Space Telescope Wide Field Camera image of a distant cluster of galaxies is shown in Fig. 1.2, upper left. The image used was one of a number described in [15, 36]. The simulated data are shown in Fig. 1.2, upper right. Four deconvolution methods were tested: Richardson-Lucy, Pixon, wavelet-vaguelette, Wavelet-Lucy. Deconvolved images are presented respectively in Fig. 1.2 middle left, middle right, bottom left and right. The Richardson-Lucy method amplifies the noise, which implies that the faintest objects disappear in the deconvolved image. The Pixon method introduces regularization, and the noise is under control, while objects where “pixons” have been detected are relatively well-protected from the regularization effect. Since the “pixon” features are detected from noisy partially deconvolved data, the faintest objects are not in the pixon map and are strongly regularized. The wavelet-vaguelette method is very fast and produces relatively high

quality results when compared to Pixon or Richardson-Lucy, but the Wavelet-Lucy method seems clearly the best of the four methods. There are fewer spurious objects than in the wavelet-vaguelette method, it is stable for any kind of PSF, and any kind of noise modeling can be considered.

1.7.3 Multiresolution CLEAN

The CLEAN solution is only available if the image does not contain large-scale structures. [96] introduced the concept of Multiresolution Clean (MRC) in order to alleviate the difficulties occurring in CLEAN for extended sources. The MRC approach consists of building two intermediate images, the first one (called the smooth map) by smoothing the data to a lower resolution with a Gaussian function, and the second one (called the difference map) by subtracting the smoothed image from the original data. Both these images are then processed separately. By using a standard CLEAN algorithm on them, the smoothed clean map and difference clean map are obtained. The recombination of these two maps gives the clean map at the full resolution. This algorithm may be viewed as an artificial recipe, but it has been shown [82, 81, 88] that it is linked to multiresolution analysis. Wavelet analysis leads to a generalization of MRC from a set of scales. The Wavelet Clean Method (WCLEAN) consists of the following steps:

- Apply the wavelet transform to the image: we get W_I .
- Apply the wavelet transform to the PSF: we get W_P .
- Apply the wavelet transform to the Clean Beam: we get W_C .
- For each scale j of the wavelet transform, apply the CLEAN algorithm using the wavelet scale j of both W_I and W_P .
- Apply an iterative reconstruction algorithm using W_C .

More details can be found in [82, 88].

1.7.4 The Wavelet Constraint

We have seen previously that many regularized deconvolution methods (MEM, Tikhonov, total variation, etc.) can be expressed by two terms (i.e. $\|I - P * O\|^2 + \lambda \mathcal{C}(O)$), the first representing the fidelity to the data and the second (i.e. $\mathcal{C}(O)$) the smoothness constraint on the solution. The parameter λ fixes the trade-off between the fit to the data and the smoothness. Using a wavelet based penalizing term \mathcal{C}_w , we want to minimize

$$J(O) = \|I - P * O\|^2 + \lambda \mathcal{C}_w(O) \quad (1.49)$$

If ϕ is a potential function which was applied on the gradients (see section 1.5.7), it can also be applied to the wavelet coefficients and the constraint on the solution is

now expressed in the wavelet domain by [48]:

$$J(O) = \|I - P * O\|^2 + \lambda \sum_{j,k,l} \phi(\|(\mathcal{W}O)_{j,k,l}\|_p) \quad (1.50)$$

When $\phi(x) = x$ and $p = 1$, it corresponds to the l_1 norm of the wavelet coefficients. In this framework, the multiscale entropy deconvolution method (see below) is only one special case of the wavelet constraint deconvolution method.

1.7.4.1 Multiscale Entropy

In [89, 86, 90], the benchmark properties for a good “physical” definition of entropy were discussed. The multiscale entropy, which fulfils these properties, consists of considering that the entropy of a signal is the sum of the information at each scale of its wavelet transform [89], and the information of a wavelet coefficient is related to the probability of it being due to noise.

For Gaussian noise, the multiscale entropy penalization function is:

$$h_n(w_{j,k}) = \frac{1}{\sigma_j^2} \int_0^{|w_{j,k}|} u \operatorname{erfc}\left(\frac{|w_{j,k}| - u}{\sqrt{2}\sigma_j}\right) du \quad (1.51)$$

where erfc is the complementary error function. A complete description of this method is given in [87]. Fig. 1.3 shows the multiscale entropy penalization function. The dashed line corresponds to a l_1 penalization (i.e. $\phi(w) = |w|$), the dotted line to a l_2 penalization $\phi(w) = \frac{w^2}{2}$, and the continuous line to the multiscale entropy function. We can immediately see that the multiscale entropy function presents quadratic behavior for small values, and is closer to the l_1 penalization function for large values. Penalization functions with a l_2 - l_1 behavior are generally a good choice for image restoration.

The Beta Pictoris image [70] was obtained by integrating 5 hours on-source using a mid-infrared camera, TIMMI, placed on the 3.6 ESO telescope (La Silla, Chile). The raw image has a peak signal-to-noise ratio of 80. It is strongly blurred by a combination of seeing, diffraction (0.7 arcsec on a 3m class telescope) and additive Gaussian noise. The initial disk shape in the original image has been lost after the convolution with the PSF (see Fig. 1.3). Thus we need to deconvolve such an image to get the best information on this object i.e. the exact profile and thickness of the disk, and subsequently to compare the results to models of thermal dust emission.

After filtering (see Fig. 1.3, lower left), the disk appears clearly. For detection of faint structures (the disk here), one can calculate that the application of such a filtering method to this image provides a gain of observing time of a factor of around 60. The deconvolved image (Fig. 1.3, lower right) shows that the disk is relatively flat at $10 \mu\text{m}$ and asymmetrical.

1.7.4.2 TV and Undecimated Haar Transform

A link between the TV and the undecimated Haar wavelet soft thresholding has been studied in [22, 92], arguing that in the 1D case the TV and the undecimated single

resolution Haar are equivalent. When going to 2D, this relation does not hold anymore, but the two approaches share some similarities. Whereas the TV introduces translation- and rotation-invariance, the undecimated 2D Haar presents translation- and scale-invariance (being multi-scale).

1.7.4.3 Minimization algorithm

Recent works [34, 28, 23] show that the solution of eqn. 1.50 can be obtained in a very efficient way, by applying a wavelet denoising on the solution at each step of the Landweber iteration.

$$O^{(n+1)} = \mathbf{WDen}_\lambda \left(O^{(n)} + P^* * (I - P * O^{(n)}) \right) \quad (1.52)$$

where \mathbf{WDen} is the operator which performs wavelet denoising, i.e. applies the wavelet transform, corrects the wavelet coefficients from the noise, and applies the inverse wavelet transform.

If $\phi(x) = x$ and $p = 1$ (i.e. l_1 norm), the solution is obtained by the following iteration:

$$O^{(n+1)} = \mathbf{soft}_\lambda (O^{(n)} + P^* * (I - P * O^{(n)})) \quad (1.53)$$

where \mathbf{soft} is the soft thresholding. (A hard threshold retains wavelet coefficients whose absolute value is above the threshold, whereas a soft threshold attenuates the wavelet coefficient, using the absolute value minus the threshold so long as this is > 0 .) If the Haar wavelet transform is chosen, this algorithm is a fast method for minimizing the total variation.

The penalty function needs to be continuous in order to guarantee the convergence. Therefore, a hard threshold cannot be used but a soft threshold as well as many other shrinkage techniques can be used. If the penalty function is strictly convex (as in soft thresholding), then it converges to a global minimum [34].

1.7.4.4 Constraints in the object or image domains

Let us define the *object domain* \mathcal{O} as the space to which the solution belongs, and the *image domain* \mathcal{I} as the space to which the data belongs (i.e. if $X \in \mathcal{O}$ then $P * X \in \mathcal{I}$). In section 1.7.2, it was shown that the multiresolution support constraint leads to a powerful regularization method. In this case, the constraint was applied in the image domain. Here, we have considered constraints on the solution. Hence, two different wavelet based strategies can be chosen in order to regularize the deconvolution problem.

The constraint in the image domain through the multiresolution support leads a very robust way to control the noise. Indeed, whatever the nature of the noise, we can always derive robust detection levels in the wavelet space and determine scales and positions of the important coefficients. A drawback of the image constraints is that there is no guarantee that the solution is free of artifacts such as ringing around point sources. A second drawback is that image constraints can be used only if the point spread function is relatively compact, i.e. does not smear the information over

the whole image. If it does do so, the concept of localization of information does not make sense any more.

The property of introducing a robust noise modeling is lost when applying the constraint in the object domain. For example, in the case of Poisson noise, there is no way (except using time consuming Monte Carlo techniques) to estimate the level of the noise in the solution and to adjust properly the thresholds. The second problem with this approach is that we try to solve two problems (noise amplification and artifact control in the solution) with one parameter (i.e. λ). The choice of this parameter is crucial, while such a parameter does not exist when using the multiresolution support.

Constraints can be added in both the object and image domains in order to better control the noise by using the multiresolution support. This gives us a guarantee that the solution is free of artifacts when using the wavelet constraint on the solution [70, 90, 91]. This leads to the following equation to be minimized:

$$J(O) = \|M \cdot \mathcal{W}_1(I - P * O)\|^2 + \lambda \sum_{j,k,l} \phi(\|(\mathcal{W}_2 O)_{j,k,l}\|_p) \quad (1.54)$$

where M is the multiresolution support derived from I and \mathcal{W}_1 . \mathcal{W}_1 and \mathcal{W}_2 are the wavelet transforms used in the object and image domains. We may want to use two different wavelet decompositions: \mathcal{W}_1 for detecting the significant coefficients and \mathcal{W}_2 for removing the artifacts in the solution. Since the noise is controlled by the multiscale transforms, the regularization parameter λ does not have the same importance as in standard deconvolution methods. A much lower value is enough to remove the artifacts relative to the use of the wavelets. The positivity constraint can be applied at each iteration. The iterative scheme is now:

$$O^{(n+1)} = \mathbf{WDen}_\lambda \left(O^{(n)} + P^* * \bar{R}^n \right) \quad (1.55)$$

where \bar{R}^n is the significant residual, i.e. $\bar{R}^n = \mathcal{W}_1^{-1} M [\mathcal{W}_1(I - P * O^{(n)})]$ (see eqn. 1.42)).

1.7.4.5 The Combined Deconvolution Method

One may want to benefit from the advantages of both the wavelet and the curvelet transforms for detecting the significant features contained in the data. More generally, we assume we use K transforms T_1, \dots, T_K , and we derive K multiresolution supports M_1, \dots, M_K from the input image I using noise modeling. Following determination of a set of multiresolution supports, we can solve the following optimization problem [91]:

$$\min_{\tilde{O}} \mathcal{C}(\tilde{O}), \quad \text{subject to} \quad M_k T_k [P * \tilde{O}] = M_k T_k I \quad \text{for all } k, \quad (1.56)$$

where \mathcal{C} is the smoothness constraint.

The constraint imposes fidelity on the data, or more exactly, on the significant coefficients of the data, obtained by the different transforms. Non-significant (i.e. noisy) coefficients are not taken into account, preventing any noise amplification in the final algorithm.

A solution for this problem could be obtained by relaxing the constraint to become an approximate one:

$$\min_{\tilde{O}} \sum_k \|M_k T_k I - M_k T_k [P * \tilde{O}]\|_2 + \lambda \mathcal{C}(\tilde{O}) \quad (1.57)$$

The solution is computed by using the projected Landweber method [6]:

$$\tilde{O}^{n+1} = \tilde{O}^n + \alpha \left(P^* * \bar{R}^n - \lambda \frac{\partial \mathcal{C}}{\partial O}(\tilde{O}^n) \right) \quad (1.58)$$

where \bar{R}^n is the significant residual which is obtained using the following algorithm:

- Set $I_0^n = I^n = P * \tilde{O}^n$.
- For $k = 1, \dots, K$ do $I_k^n = I_{k-1}^n + \mathcal{R}_k [M_k (\mathcal{T}_k I - \mathcal{T}_k I_{k-1}^n)]$
- The significant residual \bar{R}^n is obtained by: $\bar{R}^n = I_K^n - I^n$.

This can be interpreted as a generalization of the multiresolution support constraint to the case where several transforms are used. The order in which the transforms are applied has no effect on the solution. We extract in the residual the information at scales and pixel indices where significant coefficients have been detected.

α is a convergence parameter, chosen either by a line-search minimizing the overall penalty function or as a fixed step-size of moderate value that guarantees convergence.

If the \mathcal{C} is a wavelet based penalization function, then the minimization can again be done using the previous wavelet denoising approach:

$$\tilde{O}^{n+1} = \mathbf{WDen}(\tilde{O}^n + (P^* * \bar{R}^n)) \quad (1.59)$$

The positivity is introduced in the following way:

$$O^{n+1} = \mathcal{P}_c [\mathbf{WDen}(\tilde{O}^n + (P^* * \bar{R}^n))] \quad (1.60)$$

1.8 Deconvolution and Resolution

In many cases, there is no sense in trying to deconvolve an image at the resolution of the pixel (especially when the PSF is very large). The idea to limit the resolution is relatively old, because it is already this concept which is used in the CLEAN algorithm [45]. Indeed the clean beam fixes the resolution in the final solution. This principle was also developed by Lannes [54] in a different form. This concept was re-invented, first by Gull and Skilling [42] who called the clean beam the *Intrinsic Correlation Function* (ICF), and more recently by Magain [57] and Pijpers [72].

The ICF is usually a Gaussian, but in some cases it may be useful to take another function. For example, if we want to compare two images I_1 and I_2 which are obtained with two wavelengths or with two different instruments, their PSFs P_1 and P_2 will certainly be different. The classic approach would be to deconvolve I_1 with P_2 and I_2 with P_1 , so we are sure that both are at the same resolution. But unfortunately we lose some resolution in doing this. Deconvolving both images is generally not possible because we can never be sure that both solutions O_1 and O_2 will have the same resolution.

A solution would be to deconvolve only the image which has the worse resolution (say I_1), and to limit the deconvolution to the second image resolution (I_2). Then, we just have to take P_2 for the ICF. The deconvolution problem is to find \tilde{O} (hidden solution) such that:

$$I_1 = P_1 * P_2 * \tilde{O} \quad (1.61)$$

and our real solution O_1 at the same resolution as I_2 is obtained by convolving \tilde{O} with P_2 . O_1 and I_2 can then be compared.

Introducing an ICF G in the deconvolution equation leads to just considering a new PSF P' which is the convolution of P and G . The deconvolution is carried out using P' , and the solution must be reconvolved with G at the end. In this way, the solution has a constrained resolution, but aliasing may occur during the iterative process, and it is not sure that the artifacts will disappear after the re-convolution with G . Magain [57] proposed an innovative alternative to this problem, by assuming that the PSF can be considered as the convolution product of two terms, the ICF G and an unknown S , $P = G * S$. Using S instead of P in the deconvolution process, and a sufficiently large FWHM value for G , implies that the Shannon sampling theorem [79] is never violated. But the problem is now to calculate S , knowing P and G , which is again a deconvolution problem. Unfortunately, this delicate point was not discussed in the original paper. Propagation of the error on the S estimation in the final solution has also until now not been investigated, even if this issue seems to be quite important.

1.9 Myopic and Blind Deconvolution

In the field of astronomy the PSF is, in many cases, variable with time or within the observed field. For instance, when observing in the visible or near-infrared range from the ground, the atmospheric turbulence produces images with seeing-limited spatial resolution, that are much lower than the theoretical diffraction limit. This resolution ranges typically for the visible radiations between 0.4 and 1 arcsec in the best sites for astronomical observations. Because of the stochastic nature of the seeing (due to random changes of the optical index of the different layers of the atmosphere), several timescales are involved. First, the PSF is highly unstable over a timescale of several tens of milliseconds in the visible, but its time averaged value

over typically a few seconds can remain stable over a much longer time (several minutes). Depending on the site, the average seeing can vary significantly with typical timescales of a few tens of minutes. Long exposures, seeing averaged, would usually provide an estimate of the PSF for long exposures on scientific targets.

For a given observation, one can define a parameter, called the Fried parameter, r_0 . Its value defines the equivalent diameter of a telescope that would produce a diffraction-limited image with the same resolution as the seeing limited one. This parameter varies as a function of the wavelength such that $r_0 \propto (\lambda)^{5/6}$ cm; a typical value of this parameter is 20 cm at 0.5 μ m. The image resolution is then of the order of r_0/λ in the case of seeing-limited data, instead of a value around D/λ (D being the telescope diameter) for diffraction limited observations. If visible/near-infrared data are always seeing-limited, one can note that mid-infrared data around 10 μ m are mainly diffraction limited on a 4m-class telescope. However, for one 8m-class telescope such as the Very Large Telescope in Chile and given the seeing conditions at this site, one is just at the transition regime between purely diffraction-limited and seeing limited. Hence mid-infrared data on 8m-class telescopes require now the use of myopic deconvolution methods.

The principle of adaptive optics (AO) observations was developed in order to try to get rid of the seeing limitation and recover diffraction limited images. However, this is true to a certain extent. First, only a small field of view (typically, a few arcsec in the visible range), around the AO reference (a sufficiently bright object; typically a bright star with magnitude V lower than 3), currently benefits from the best image correction which degrades as $\exp(-\Delta\theta)^{5/3}$ ($\Delta\theta$ is the distance to the center-field). This field extent is defined by the isoplanatic angle, in which the wavefront distortion is corrected accurately. In order to reach larger fields of view, multi-conjugated adaptive optics (MCAO) systems are under development in parallel with laser guide stars which will allow one to synthetically produce AO references at any place in the sky. The Strehl ratio, which is defined as the ratio of the central intensity of the currently observed image to the central intensity of a perfect diffraction limited image, can achieve typical values around 30–50% for K band observations, 2.2 μ m). As one observes an object far from the AO reference, this correction assessment parameter degrades, reaching a value being half its value at the center-field (the location of the AO reference), at a distance of 40 arcsec from it. Since AO performances are intimately linked to the seeing (or r_0) value, the AO correction will also strongly vary with time. In summary, AO observations deal with quite significantly varying PSFs: varying in time because of seeing time variations, and varying in space because of the isoplanatic patch. As a consequence, AO astronomical observations do require also myopic deconvolution methods. Much effort in astronomical blind deconvolution have been thus naturally devoted to this type of data.

Astronomical short exposures, or speckle images, have usually unknown PSF because of strong and random changes of the image phase. Although it is essentially the phase of the object which is affected, while its amplitude is barely modified, some specific methods of blind deconvolution in which the PSF is assumed unknown, have to be used. The challenge that faces any method is that it should be able to incorpo-

rate as much as possible by way of constraints and a priori knowledge on the solution in order to avoid being trapped in trivial solutions such as $\{O, P\} = \{I, \delta\}$ where δ is the Dirac function.

1.9.1 Myopic Deconvolution

Myopic deconvolution assumes that an estimate of the PSF, which is not too “far” from the true one, is known. A measure of this PSF is usually performed by observing an object supposed to be point-like, in the same observational conditions (same telescope/instrument configuration, and weather conditions as close as possible). This is achieved by observing frequently, before and after any scientific target, some unresolved star, at an airmass close to that of the science target, and if possible, not too far in the sky from it. Astronomical data usually deal then with multi-frame data: multiple records of the blurred object, multiple records of estimates of the PSF P_i . Within this framework some new methods have emerged in the last decade and are presented in the following.

IDAC is a myopic, Bayesian deconvolution method derived from a blind deconvolution method ([19], see section 1.9.2 below for the details), and applied mainly to adaptive optics data. It is a multi-frame deconvolution method in which *loose* constraints are put on the solution.

The MISTRAL method was developed by [65] in order to deconvolve AO partially corrected data. An estimate of the PSF P_{est} can be derived from observations of an unresolved star shortly before and/or after observing the object of interest. Using the usual Bayesian framework, and the MAP approach, one ends up with the following functional to minimize, to find O and P simultaneously:

$$J(O, P) = J_{res} + J_P(P) + J_O(O) \quad (1.62)$$

where J_{res} is the “fidelity to the data” term such that:

$$J_{res}(O, P) = \frac{1}{2\sigma^2} \| (I - P * O) \|^2 \quad \text{in the case of Gaussian noise} \quad (1.63)$$

and

$$J_{res}(O, P) = \sum_{pixels} (P * O - I \ln[P * O]) \quad \text{in the case of Poisson noise} \quad (1.64)$$

$J_P(P)$ is the penalty term expressing that solution P should not be “too different” from the estimate of the PSF P_{est} :

$$J_P(P) = \sum_{u,v} \frac{|\hat{P}(u, v) - \hat{P}_{est}(u, v)|^2}{S_P(u, v)} \quad (1.65)$$

where $S_p(u, v) = E[|\hat{P}(u, v) - \hat{P}_{est}(u, v)|]$ is the spatial power spectral density. P_{est} is computed as $E[P_i]$, P_i being the different estimates of the PSF.

In addition, a reparametrization ($O = \psi^2$) ensures the positivity of the solution. Some very nice results obtained with this method applied to images of the satellite of Jupiter, Io, are shown in Fig. 1.4 allowing us to study Io volcanism with ground-based AO data.

The deconvolution from wave-front sensing was originally proposed by [35] in 1985. The idea behind it is that in AO observations wave-front data are simultaneously recorded in the wave-front sensor (WFS); these data contain information about the PSF but in an unusual form (often projected onto a base of Zernike polynomials). Since we deal with short exposure images (less than 10 ms typically), the atmospheric turbulence is assumed to be frozen, so that the PSF at a time t can be fully characterized by the turbulent phase $\phi(t)$ in the pupil plane:

$$PSF(t) = \mathcal{F}^{-1}(Pe^{j\phi(t)}) \quad (1.66)$$

where P is the pupil function.

The usual techniques consist of firstly estimating the wave-fronts from the WFS, and then obtaining the deconvolved image by maximum a posteriori (MAP) estimation [5]. Since wave-front estimates are inevitably noisy, Mugnier et al. [66] proposed a robust joint estimator within a Bayesian framework to reconstruct the true wave-front data and the restored object simultaneously. This estimator uses all statistical information that one has either on the noise, or on the stochastic processes that control the turbulent phase (Kolmogorov statistics). A functional containing constraints on $\phi(t)$ (complying with Kolmogorov statistics) and on the object O is minimized. This method is more efficient than speckle interferometry because it is not limited by speckle noise at high photon flux and the signal-to-noise ratio is better for extended objects. Fig. 1.5 shows how spectacular the result can be on experimental data of the binary star Capella.

1.9.2 Blind Deconvolution

In 1994, [95] proposed a simple method for blind deconvolution based on Lucy's algorithm. The idea is to alternatively perform a Lucy iteration on the object O and then on the PSF P . However, although attractive because of its simplicity, this process (i) can be highly unstable, and (ii) puts no constraint on the PSF making it difficult to prevent it tending towards the trivial solution $\{I, \delta\}$.

Jefferies and Christou [50] have proposed an iterative blind deconvolution method of multi-frame data based on the minimization of a penalty functional putting physical and "reasonable" loose constraints on the solution (O, P_i). Assuming that one deals with $i = 1, \dots$ frames, this method minimizes the functional:

$$J(O, P) = E_{im} + E_{conv} + E_{bl} + E_{Fm} \quad (1.67)$$

where:

1. $E_{im} = \sum_{n \in \gamma_O} O(n)^2 + \sum_{m \in \gamma_P} P(m)^2$ is the image domain error which penalizes negative values (γ_O and γ_P sets) in the solution O or in the PSF P .
2. $E_{conv} = \sum_i \sum_{u,v} |\hat{I}(u,v) - \hat{P}(u,v)\hat{O}(u,v)|^2 M(u,v)$ is called the convolution error and expresses the fidelity of the reconvolved image to the data in Fourier space, M being a mask excluding spatial frequencies higher than the physical cut-off frequency (by diffraction (D/λ) or seeing (r_0/λ) of the instrumental set. Depending on the conventions on the Fourier transform, one has to normalize this term by the number of pixels in the image.
3. E_{bl} is the band-limit error defined by $E_{bl} = \sum_i \sum_{u,v} |\hat{P}_i(u,v)|^2 M'_i(u,v)$ where M' is a spatial frequencies mask greater than 1.39 of the cut-off frequency. The same normalization rule applies as above.
4. E_{Fm} is the Fourier modulus error, such that $E_{Fm} = \sum_i \sum_{u,v} |\hat{O}(u,v) - \hat{O}_{est}(u,v)|^2 \Phi(u,v)$ where $\hat{O}_{est}(u,v)$ is a crude estimate of the object's Fourier modulus given by:

$$|\hat{O}_{est}| = \sqrt{\frac{\langle |\hat{I}|^2 \rangle - \langle |\hat{N}|^2 \rangle}{\langle |\hat{P}|^2 \rangle}} \quad (1.68)$$

Thiebaut and Conan [93] compared the application of loose (similar to some extent to [50]) and of strict a priori constraints on the solution when maximizing the likelihood. They found that much better solutions can be achieved when applying the second option. Strict a priori constraints are applied in this case by a reparametrization ensuring the sought physical properties of the solution e.g.:

1. Positivity of the object O making the change of variable:

$$O = \psi_O^2 \quad (1.69)$$

2. Positivity and normalization of the PSF P by setting:

$$P = \frac{\psi_h^2}{\sum_{pixels} \psi_h^2} \quad (1.70)$$

The gradients of the functional are then recomputed, and depend now on the new unknowns ψ_P (representing P) and ψ_O (representing O). This method has been shown to give very good results in the case of speckle data on the Capella binary system [93].

In summary, several methods are currently available to carry out myopic or blind deconvolution. For best results we recommend the Bayesian formalism because it offers a nice way to incorporate any type of constraint on the solution. These methods offer much better results than any other algebraic method, but at some computational cost (the minimization of the functional requires intensive computing and efficient minimization schemes).

1.10 Conclusions and Chapter Summary

As in many fields, simple methods can be availed of – for example the solution provided by equation (1.16) – but at the expense of quality in the solution and a full understanding of one’s data. Often a simple solution can be fully justified. However, if our data or our problems are important enough, then appropriate problem solving approaches have to be adopted. The panoply of methods presented in this review provide options for high quality image and signal restoration.

We have noted how the wavelet transform offers a powerful mathematical and algorithmic framework for multiple resolution analysis. Furthermore noise modeling is very advantageously carried out in wavelet space. Finally, and of crucial importance in this chapter, noise is the main problem in deconvolution.

Progress has been significant in a wide range of areas related to deconvolution. One thinks of Bayesian methods, the use of entropy, and issues relating to super-resolution, for example.

We will conclude with a short look at how multiscale methods used in deconvolution are evolving and maturing.

We have seen that the recent improvement in deconvolution methods has led to use of a multiscale approach. Finally, wavelet based constraints can be added in both domains [90]. This allows us to separate the deconvolution problem into two separate problems: noise control from one side, and solution smoothness control on the other side. The advantage is that noise control is better carried out in the image domain, while smoothness control can only be carried out in the object domain.

The reason for the success of wavelets is due to the fact that wavelet bases represent well a large class of signals, especially astronomical data where most of the objects are more or less isotropic. When the data contains anisotropic features (solar, planetary images, etc.), other multiscale methods, such as the ridgelet or the curvelet transform [14, 13, 32, 84], are good candidates for replacing the wavelet transform. The ultimate step is the combination of the different multiscale decompositions.

Very nice results have been obtained based on myopic or blind deconvolution. However, there is currently no algorithm based on multiscale methods in the field of myopic or blind deconvolution for the regularization. New methods taking advantage of these tools, as was already done for standard image deconvolution, should appear soon.

Acknowledgements

We are grateful to Eric Thiebaut and Laurent Mugnier for use of their code, and to the referees of this chapter.

References

- [1] R. Acar and C.R. Vogel. Analysis of bounded variation penalty methods for ill-posed problem. *Physica D*, 10:1217–1229, 1994.
- [2] H.M. Adorf, R.N. Hook, and L.B. Lucy. HST image restoration developments at the ST-ECF. *International Journal of Imaging Systems and Technology*, 6:339–349, 1995.
- [3] D. Alloin, E. Pantin, P. O. Lagage, and G. L. Granato. 0.6 resolution images at 11 and 20 μm of the active galactic nucleus in NGC 1068. *Astronomy and Astrophysics*, 363:926–932, November 2000.
- [4] L. Alvarez, P.-L. Lions, and J.-M. Morel. Image selective smoothing and edge detection by nonlinear diffusion. *SIAM Journal on Numerical Analysis*, 29:845–866, 1992.
- [5] P. A. Bakut, V. E. Kirakosyants, V. A. Loginov, C. J. Solomon, and J. C. Dainty. Optimal wavefront reconstruction from a Shark-Hartmann sensor by use of a Bayesian algorithm. *Opt. Commun.*, 109:10–15, 1994.
- [6] M. Bertero and P. Boccacci. *Introduction to Inverse Problems in Imaging*. Institute of Physics, 1998.
- [7] L. Blanc-Féraud and M. Barlaud. Edge preserving restoration of astrophysical images. *Vistas in Astronomy*, 40:531–538, 1996.
- [8] T.R. Bontekoe, E. Koper, and D.J.M. Kester. Pyramid maximum entropy images of IRAS survey data. *Astronomy and Astrophysics*, 284:1037–1053, 1994.
- [9] C. A. Bouman and K. Sauer. A generalized Gaussian image model for edge-preserving MAP estimation. *IEEE Transactions on Image Processing*, 2(3):296–310, 1993.
- [10] R. Buonanno, G. Buscema, C.E. Corsi, I. Ferraro, and G. Iannicola. Automated photographic photometry of stars in globular clusters. *Astronomy and Astrophysics*, 126:278–282, October 1983.
- [11] J.P. Burg. Multichannel maximum entropy spectral analysis. Annual Meeting International Society Exploratory Geophysics, Reprinted in *Modern Spectral Analysis*, D.G. Childers, ed., IEEE Press, 34–41, 1978.
- [12] I. Burud, F. Courbin, P. Magain, C. Lidman, D. Hutsemékers, J.-P. Kneib, J. Hjorth, J. Brewer, E. Pompei, L. Germany, J. Pritchard, A. O. Jaunsen, G. Letawe, and G. Meylan. An optical time-delay for the lensed BAL quasar HE 2149-2745. *Astronomy and Astrophysics*, 383:71–81, January 2002.
- [13] E. Candès and D.L. Donoho. Curvelets, multiresolution representation, and scaling laws. In *SPIE conference on Signal and Image Processing: Wavelet Applications in Signal and Image Processing VIII*, 2000.

- [14] E.J. Candès and D.L. Donoho. Ridgelets: the key to high dimensional intermittency? *Philosophical Transactions of the Royal Society of London A*, 357:2495–2509, 1999.
- [15] A. Caulet and W. Freudling. Distant galaxy cluster simulations – HST and ground-based. *ST-ECF Newsletter No. 20*, pages 5–7, 1993.
- [16] F. Champagnat, Y. Goussard, and J. Idier. Unsupervised deconvolution of sparse spike trains using stochastic approximation. *IEEE Transactions on Image Processing*, 44:2988–2997, 1996.
- [17] P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud. Deterministic edge-preserving regularization in computed imaging. *IEEE Transactions on Image Processing*, 6:298–311, 1997.
- [18] P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud. Deterministic edge-preserving regularization in computed imaging. *IEEE Transactions on Image Processing*, 6(2):298–311, 1997.
- [19] J. Christou. idac – Iterative Deconvolution Algorithm in C web page. http://babcock.ucsd.edu/cfao-ucsd/idac/idac-package/idac_index.html, 2000.
- [20] J. C. Christou, D. Bonnacini, N. Ageorges, and F. Marchis. Myopic deconvolution of adaptive optics images. *The Messenger*, 97:14–22, September 1999.
- [21] P.H. Van Cittert. Zum Einfluß der Spaltbreite auf die Intensitätsverteilung in Spektrellinien II. *Zeitschrift für Physik*, 69:298–308, 1931.
- [22] A. Cohen, R. DeVore, P. Petrushev, and H. Xu. Nonlinear approximation and the space $BV(R^2)$. *Amer. J. Math.*, 121:587–628, 1999.
- [23] P.L. Combettes and V.R. Vajs. Signal recovery by forward-backward splitting. *preprint*, 2005.
- [24] T. Corbard, L. Blanc-Féraud, G. Berthomieu, and J. Provost. Nonlinear regularization for helioseismic inversions. Application for the study of the solar tachocline. *Astronomy and Astrophysics*, 344:696–708, April 1999.
- [25] T. J. Cornwell. Image Restoration. In *NATO ASIC Proc. 274: Diffraction-Limited Imaging with Very Large Telescopes*, pages 273–+, 1989.
- [26] F. Courbin, C. Lidman, and P. Magain. Detection of the lensing galaxy in HE 1104-1805. *Astronomy and Astrophysics*, 330:57–62, February 1998.
- [27] A. Coustenis, E. Gendron, O. Lai, J. Véran, J. Woillez, M. Combes, L. Vapillon, T. Fusco, L. Mugnier, and P. Rannou. Images of Titan at 1.3 and 1.6 μ m with Adaptive Optics at the CFHT. *Icarus*, 154:501–515, 2001.
- [28] I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Comm. Pure Appl. Math.*, 57:1413–1541, 2004.

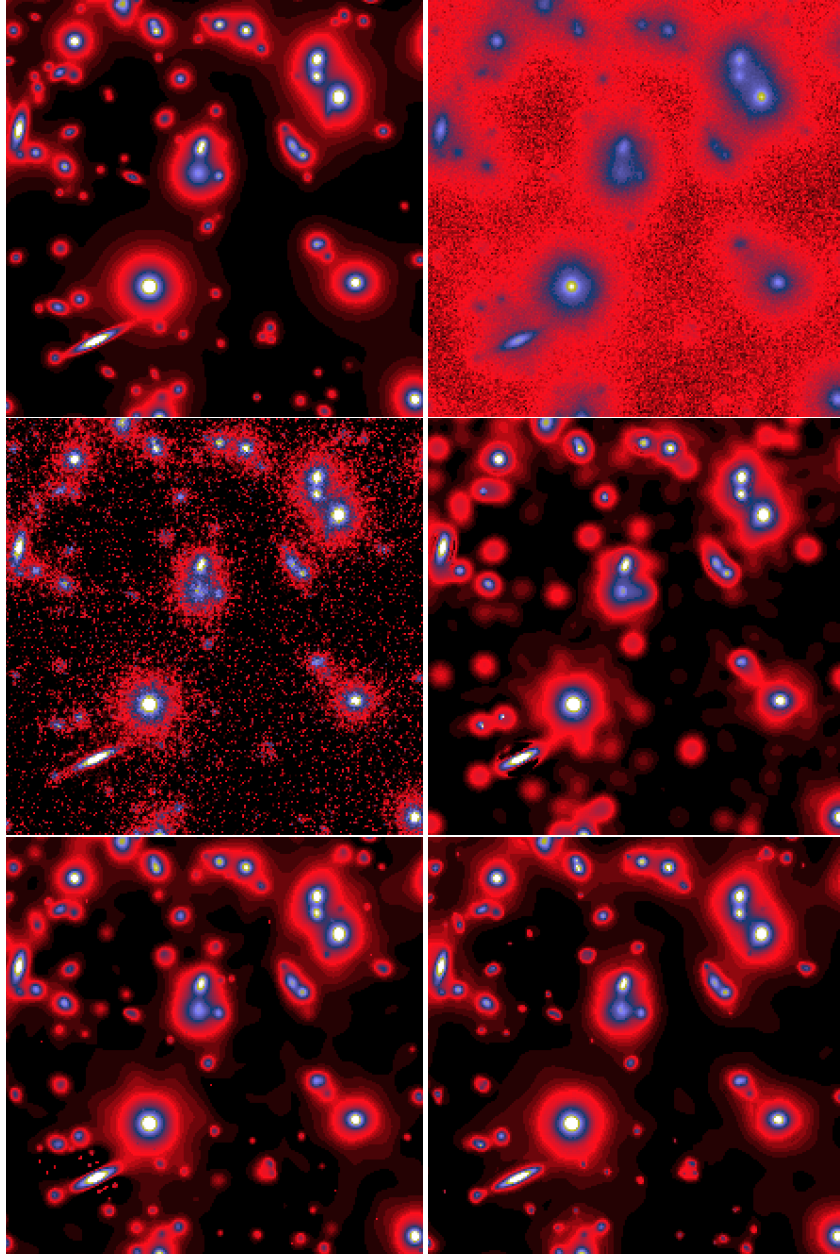
- [29] A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 39:1–22, 1977.
- [30] D.D. Dixon, W.N. Johnson, J.D. Kurfess, R.K. Pina, R.C. Puetter, W.R. Purcell, T.O. Tuemer, W.A. Wheaton, and A.D. Zych. Pixon-based deconvolution. *Astronomy and Astrophysics, Supplement Series*, 120:683–686, December 1996.
- [31] S. Djorgovski. Modelling of seeing effects in extragalactic astronomy and cosmology. *Journal of Astrophysics and Astronomy*, 4:271–288, December 1983.
- [32] D.L. Donoho and M.R. Duncan. Digital curvelet transform: strategy, implementation and experiments. In H.H. Szu, M. Vetterli, W. Campbell, and J.R. Buss, editors, *Proc. Aerosense 2000, Wavelet Applications VII*, volume 4056, pages 12–29. SPIE, 2000.
- [33] C. Faure, F. Courbin, J. P. Kneib, D. Alloin, M. Bolzonella, and I. Burud. The lensing system towards the doubly imaged quasar SBS 1520+530. *Astronomy and Astrophysics*, 386:69–76, April 2002.
- [34] M.A. Figueiredo and R. Nowak. An EM algorithm for wavelet-based image restoration. *IEEE Transactions on Image Processing*, 12(8):906–916, 2003.
- [35] J. C. Fontanella. Analyse de surfaces d’onde, déconvolution et optique active. *J. Mod. Opt.*, 16:257–268, 1985.
- [36] W. Freudling and A. Caulet. Simulated HST observations of distant clusters of galaxies. In P. Grosbøl, editor, *Proceedings of the 5th ESO/ST-ECF Data Analysis Workshop*, pages 63–68. European Southern Observatory, 1993.
- [37] B.R. Frieden. *Image Enhancement and Restoration*. Springer-Verlag, 1978.
- [38] N.P. Galatsanos and A.K. Katsaggelos. Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation. *IEEE Transactions on Image Processing*, 1:322–336, 1992.
- [39] S. Geman and D.E. McClure. Bayesian image analysis: an application to single photon emission tomography. In *Proc. Statist. Comput. Sect.*, Washington DC, 1985. American Statistical Association.
- [40] G.H. Golub, M. Heath, and G. Wahba. Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, 21:215–223, 1979.
- [41] P. J. Green. Bayesian reconstruction from emission tomography data using a modified EM algorithm. *IEEE Transactions on Medical Imaging*, 9(1):84–93, 1990.
- [42] S.F. Gull and J. Skilling. *MEMSYS5 Quantified Maximum Entropy User’s Manual*. Royston, England, 1991.

- [43] R. J. Hanisch and R. L. White, editors. *The restoration of HST images and spectra - II*. Space Telescope Science Institute, Baltimore, 1994.
- [44] T. Hebert and R. Leahy. A generalized EM algorithm for 3-d bayesian reconstruction from poisson data using Gibbs priors. *IEEE Transactions on Medical Imaging*, 8(2):194–202, 1989.
- [45] J.A. Högbom. Aperture synthesis with a non-regular distribution of interferometer baselines. *Astronomy and Astrophysics Supplement Series*, 15:417–426, 1974.
- [46] R. Hook. An overview of some image restoration and combination methods. *ST-ECF Newsletter No. 26*, pages 3–5, 1999.
- [47] E. Issacson and H. Keller. *Analysis of Numerical Methods*. Wiley, 1966.
- [48] A. Jalobeanu. *Models, Bayesian estimation and algorithms for remote sensing data deconvolution*. PhD thesis, Université de Nice Sophia Antipolis, December 2001.
- [49] P.A. Jansson, R.H. Hunt, and E.K. Peyler. Resolution enhancement of spectra. *Journal of the Optical Society of America*, 60:596–599, 1970.
- [50] S. M. Jefferies and J. C. Christou. Restoration of astronomical images by iterative blind deconvolution. *Astrophysical Journal*, 415:862–+, October 1993.
- [51] K.F. Kaarelsen. Deconvolution of sparse spike trains by iterated window maximization. *IEEE Transactions on Image Processing*, 45:1173–1183, 1997.
- [52] A.K. Katsaggelos. *Digital Image Processing*. Springer-Verlag, 1993.
- [53] L. Landweber. An iteration formula for Fredholm integral equations of the first kind. *American Journal of Mathematics*, 73:615–624, 1951.
- [54] A. Lannes and S. Roques. Resolution and robustness in image processing: a new regularization principle. *Journal of the Optical Society of America*, 4:189–199, 1987.
- [55] L.B. Lucy. An iteration technique for the rectification of observed distributions. *Astronomical Journal*, 79:745–754, 1974.
- [56] L.B. Lucy. Image restoration of high photometric quality. In R. J. Hanisch and R. L. White, editors, *The Restoration of HST Images and Spectra II*, page 79. Space Telescope Science Institute, 1994.
- [57] P. Magain, F. Courbin, and S. Sohy. Deconvolution with correct sampling. *Astrophysical Journal*, 494:472, 1998.
- [58] F. Marchis, R. Prangé, and J. Christou. Adaptive optics mapping of Io’s volcanism in the thermal IR (3.8 μm). *Icarus*, 148:384–396, December 2000.
- [59] A.F.J. Moffat. A theoretical investigation of focal stellar images in the photographic emulsion and application to photographic photometry. *Astronomy and Astrophysics*, 3:455, December 1969.

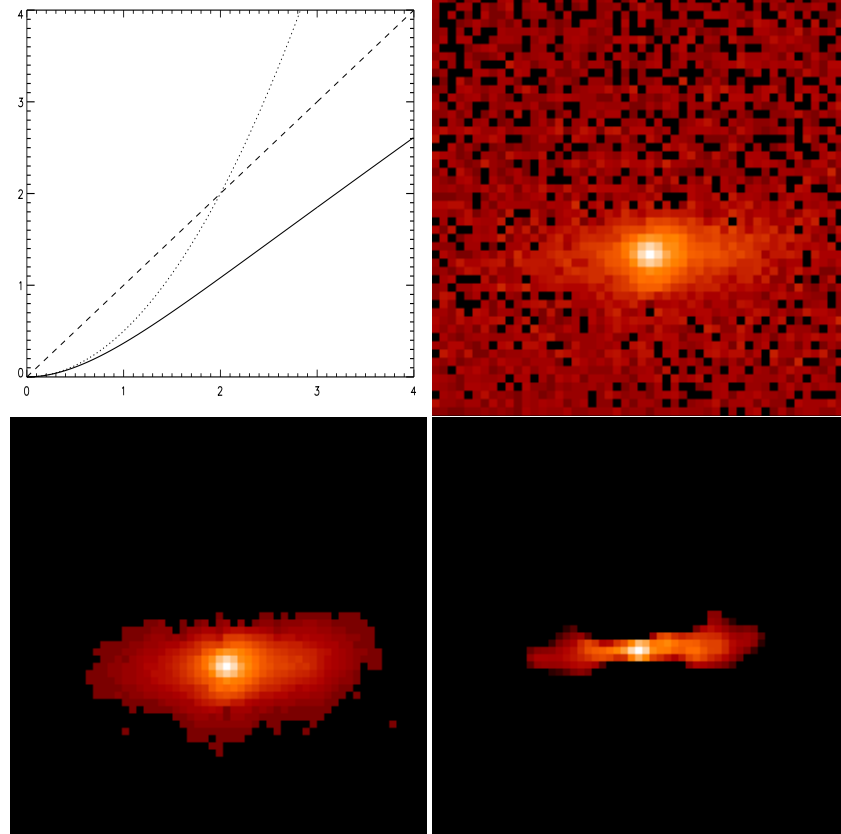
- [60] R. Molina and F.J. Cortijo. On the Bayesian deconvolution of planets. In *Proc. International Conference on Pattern Recognition, ICPR'92*, volume 3, pages 147–150, 1992.
- [61] R. Molina, A.K. Katsaggelos, J. Mateos, and J. Abad. Compound Gauss-Markov random fields for astronomical image restoration. *Vistas in Astronomy*, 40:539–546, 1996.
- [62] R. Molina, A.K. Katsaggelos, J. Mateos, A. Hermoso, and A. Segall. Restoration of severely blurred high range images using stochastic and deterministic relaxation algorithms in compound Gauss Markov random fields. *Pattern Recognition*, 33:555–571, 2000.
- [63] R. Molina, J. Núñez, F. Cortijo, and J. Mateos. Image restoration in astronomy: a Bayesian review. *IEEE Signal Processing Magazine*, 18:11–29, 2001.
- [64] R. Molina, B.D. Ripley, A. Molina, F. Moreno, and J.L. Ortiz. Bayesian deconvolution with prior knowledge of object location – applications to ground-based planetary images. *Astrophysical Journal*, 104:1662–1668, October 1992.
- [65] L. M. Mugnier, T. Fusco, and J. M. Conan. MISTRAL: a myopic edge-preserving image restoration method, with application to astronomical adaptive-optics-corrected long-exposure images. *J. Opt. Soc. Am. A*, 10:1841–1853, October 2004.
- [66] L. M. Mugnier, C. Robert, J. M. Conan, V. Michaud, and S. Salem. Myopic deconvolution from wave-front sensing. *J. Opt. Soc. Am. A*, 18:862–872, April 2001.
- [67] F. Murtagh and J.-L. Starck. Multiresolution image analysis using wavelets: Some recent results and some current directions. *ST-ECF Newsletter No. 21*, pages 19–20, 1994.
- [68] F. Murtagh, J.-L. Starck, and A. Bijaoui. Image restoration with noise suppression using a multiresolution support. *Astronomy and Astrophysics, Supplement Series*, 112:179–189, 1995.
- [69] J. Núñez and J. Llacer. Bayesian image reconstruction with space-variant noise suppression. *Astronomy and Astrophysics, Supplement Series*, 131:167–180, July 1998.
- [70] E. Pantin and J.-L. Starck. Deconvolution of astronomical images using the multiscale maximum entropy method. *Astronomy and Astrophysics, Supplement Series*, 315:575–585, 1996.
- [71] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12:629–639, 1990.

- [72] F. P. Pijpers. Unbiased image reconstruction as an inverse problem. *Monthly Notices of the Royal Astronomical Society*, 307:659–668, August 1999.
- [73] N. Pirzkal, R.N. Hook, and L.B. Lucy. GIRA – two channel photometric restoration. In N. Manset, C. Veillet, and D. Crabtree, editors, *Astronomical Data Analysis Software and Systems IX*, page 655. Astronomical Society of the Pacific, 2000.
- [74] R.C. Puetter and A. Yahil. The pixon method of image reconstruction. In *ASP Conference Series 172: Astronomical Data Analysis Software and Systems VIII*, page 307. Astronomical Society of the Pacific, 1999.
- [75] J. T. Radomski, R. K. Piña, C. Packham, C. M. Telesco, and C. N. Tadhunter. High-Resolution Mid-Infrared Morphology of Cygnus A. *Astrophysical Journal*, 566:675–681, February 2002.
- [76] W.H. Richardson. Bayesian-based iterative method of image restoration. *Journal of the Optical Society of America*, 62:55–59, 1972.
- [77] B.D. Ripley. *Spatial Statistics*. Wiley, 1981.
- [78] L.I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation noise removal algorithm. *Physica D*, 60:259–268, 1992.
- [79] C.E. Shannon. A mathematical theory for communication. *Bell System Technical Journal*, 27:379–423, 1948.
- [80] L.A. Shepp and Y. Vardi. Maximum likelihood reconstruction for emission tomography. *IEEE Transactions on Medical Imaging*, MI-2:113–122, 1982.
- [81] J.-L. Starck and A. Bijaoui. Filtering and deconvolution by the wavelet transform. *Signal Processing*, 35:195–211, 1994.
- [82] J.-L. Starck, A. Bijaoui, B. Lopez, and C. Perrier. Image reconstruction by the wavelet transform applied to aperture synthesis. *Astronomy and Astrophysics*, 283:349–360, 1994.
- [83] J.-L. Starck, A. Bijaoui, and F. Murtagh. Multiresolution support applied to image filtering and deconvolution. *CVGIP: Graphical Models and Image Processing*, 57:420–431, 1995.
- [84] J.-L. Starck, E. Candès, and D.L. Donoho. The curvelet transform for image denoising. *IEEE Transactions on Image Processing*, 11(6):131–141, 2002.
- [85] J.-L. Starck and F. Murtagh. Image restoration with noise suppression using the wavelet transform. *Astronomy and Astrophysics*, 288:343–348, 1994.
- [86] J.-L. Starck and F. Murtagh. Multiscale entropy filtering. *Signal Processing*, 76:147–165, 1999.
- [87] J.-L. Starck and F. Murtagh. *Astronomical Image and Data Analysis*. Springer-Verlag, 2002.

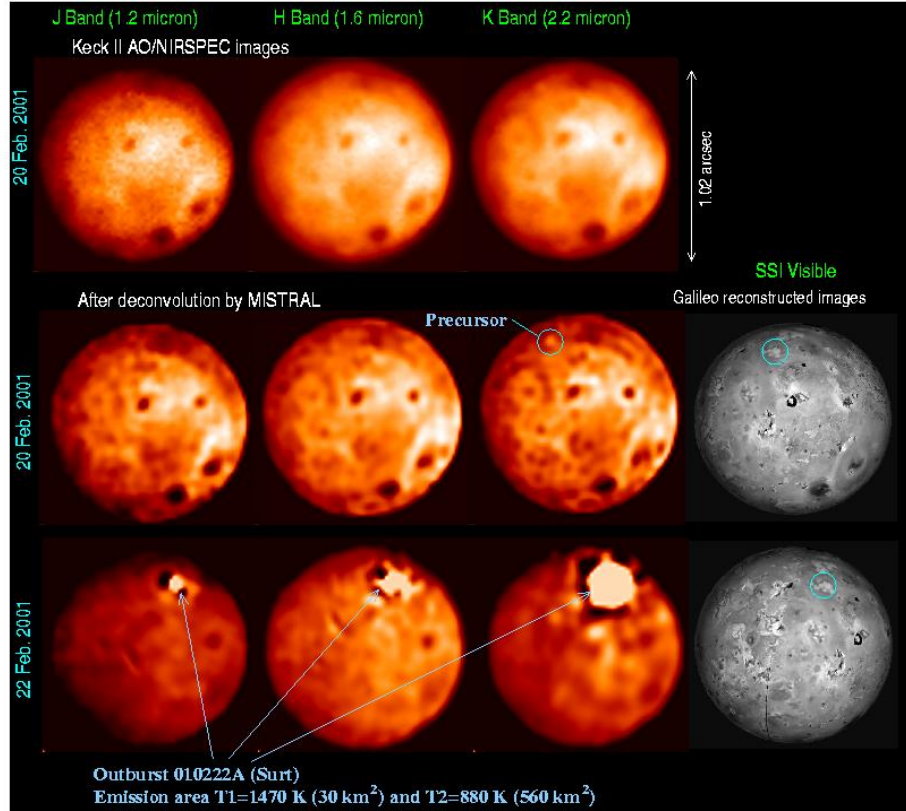
- [88] J.-L. Starck, F. Murtagh, and A. Bijaoui. *Image Processing and Data Analysis: The Multiscale Approach*. Cambridge University Press, 1998.
- [89] J.-L. Starck, F. Murtagh, and R. Gastaud. A new entropy measure based on the wavelet transform and noise modeling. *IEEE Transactions on Circuits and Systems II*, 45:1118–1124, 1998.
- [90] J.-L. Starck, F. Murtagh, P. Querre, and F. Bonnarel. Entropy and astronomical data analysis: Perspectives from multiresolution analysis. *Astronomy and Astrophysics*, 368:730–746, 2001.
- [91] J.-L. Starck, M.K. Nguyen, and F. Murtagh. Wavelets and curvelets for image deconvolution: a combined approach. *Signal Processing*, 83(10):2279–2283, 2003.
- [92] G. Steidl, J. Weickert, T. Brox, P. Mrzek, and M. Welk. On the equivalence of soft wavelet shrinkage, total variation diffusion, total variation regularization, and sides. Technical Report 26, Department of Mathematics, University of Bremen, Germany, 2003.
- [93] E. Thiebaut and J. M. Conan. Strict a priori constraints for maximum likelihood blind deconvolution. *J. Opt. Soc. Am. A*, 12:485–492, October 1996.
- [94] A.N. Tikhonov, A.V. Goncharski, V.V. Stepanov, and I.V. Kochikov. Ill-posed image processing problems. *Soviet Physics – Doklady*, 32:456–458, 1987.
- [95] F. Tsumuraya, N. Miura, and N. Baba. Iterative blind deconvolution method using Lucy’s algorithm. *Astronomy and Astrophysics*, 282:699–708, February 1994.
- [96] B.P. Wakker and U.J. Schwarz. The multi-resolution Clean and its application to the short-spacing problem in interferometry. *Annual Reviews of Astronomy and Astrophysics*, 200:312, 1988.
- [97] N. Weir. A multi-channel method of maximum entropy image restoration. In D.M. Worrall, C. Biemesderfer, and J. Barnes, editors, *Astronomical Data Analysis Software and System 1*, pages 186–190. Astronomical Society of the Pacific, 1992.
- [98] R. L. White and R. J. Allen, editors. *The restoration of HST images and spectra*, 1991.

**FIGURE 1.2**

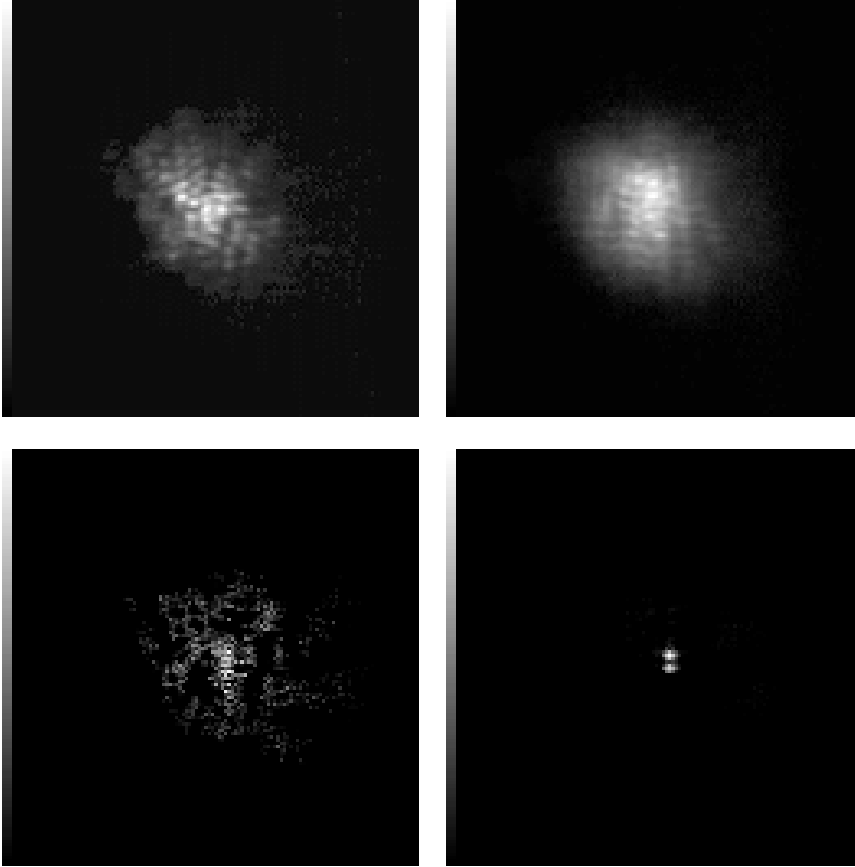
Simulated Hubble Space Telescope Wide Field Camera image of a distant cluster of galaxies. Six quadrants. Upper left: original, unaberrated and noise-free. Upper right: input, aberrated, noise added. Middle left: restoration, Richardson-Lucy. Middle right: restoration, Pixon method. Lower left, restoration wavelet-vaguelette. Lower right, restoration wavelet-Lucy.

**FIGURE 1.3**

Upper left: penalization functions: dashed, l_1 norm (i.e. $\phi(w) = |w|$); dotted l_2 norm $\phi(w) = \frac{w^2}{2}$; continuous, multiscale entropy function. Upper right: Beta Pictoris raw data. Lower left: filtered image using. Lower right: deconvolved one.

**FIGURE 1.4**

Jupiter-facing of Io observed with the Keck AO system in J,H, and K bands. The second row shows the former images deconvolved with the MISTRAL method, and are compared to Galileo probe images in which a precursor to a volcanic burst is detected. The last row shows deconvolved images 2 days later in which a major outburst, seen also on Galileo images but fainter (Galileo observes in the visible range and is less sensitive to “thermal” events), is also detected from Earth.

**FIGURE 1.5**

Experimental short exposure of the binary system Capella taken at the 4.2m William Herschel telescope, La Palma, Canarias islands (upper left) and corresponding long-exposure, 10 times longer, showing the loss of resolution due to atmospheric turbulence blurring (upper right). Restored image using a MAP estimate of the wave-front followed by a quadratic restoration with positivity constraint (lower left). Image restored by the myopic deconvolution method developed by Mugnier et al. [66] (lower right).