Study of Adaptive Optics Images by means of Multiscalar Transforms

Roberto Baena Gallé

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (**www.tdx.cat**) i a través del Dipòsit Digital de la UB (**diposit.ub.edu**) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (**www.tdx.cat**) y a través del Repositorio Digital de la UB (**diposit.ub.edu**) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (**www.tdx.cat**) service and by the UB Digital Repository (**diposit.ub.edu**) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

UNIVERSITAT DE BARCELONA

DEPARTAMENT D'ASTRONOMIA I METEOROLOGIA

Study of Adaptive Optics Images by means of Multiscalar Transforms

Memoria presentada por Roberto Baena Gallé para optar al grado de Doctor en Física.
Programa de doctorado en Física.
Línea de investigación en Astronomía y Astrofísica.
Septiembre 2013 (Bienio 2005 - 2007).
Tesis dirigida por el Dr. Jorge Núñez de Murga.

A mis padres.

Contents

Chapter 1

Adaptive Optics Imaging

Ground-based telescopes have followed a historical tendency to increase their collecting aperture size in order to allow fainter objects to be detected. Hence, the diameter of the primary mirror of such telescopes has evolved from a few centimeters, for the first one used for astronomy by Galileo Galilei (1610), to over 10 meters, such as the Keck telescopes in Mauna Kea (Hawaii) or the Gran Telescopio de Canarias in La Palma (Canary Islands). Unfortunately, the wavefront emitted by a distant point source is disturbed during its transit through the atmosphere, which limits the angular resolution. Therefore, it is impossible to achieve the theoretical diffraction limit λ/D of resolution that could be achieved by the largest ground-based telescope, with λ being the observing wavelength and D the diameter of the aperture.

In order to overcome this limitation, current telescope designs incorporate modern systems and devices that control the mirror surface shape and partially correct distortions introduced by atmospheric turbulence. Such systems include wavefront sensors, deformable mirrors, artificial guide stars produced using laser beacons and complicated control software, which allow to create astronomical images with high angular resolutions. Furthermore, their possible applications extend beyond astronomy to a much broader range of disciplines such as ophthalmology, biology or defense. In this chapter, some of the basics in this field are reviewed from both a theoretical and a practical point of view, as well as the current state of the art of the science performed with adaptive optics (AO) systems.

1.1 Introduction to image degradation by atmospheric turbulence

The theory of the effects of atmospheric turbulence in perturbing star wavefronts in astronomy was initiated by Kolmogorov (1941a) and Obukhov (1941); and further developed by Tatarski (1961), Fried (1965, 1966) and Roddier (1981). A clear mathematical treatment is outlined by Quirrenbach (2006). In a nutshell, there are four classical areas of study related to the influences of the atmosphere on image formation:

- Atmospheric absorption or extinction, which is of great importance in photometry and site selection (Léna et al. 1998; Sarazin 1992).

- Atmospheric refraction, which introduces positional errors affecting the astrometrical accuracy of observations (Barlow & Bryan 1944; Mahan 1962).

- Atmospheric dispersion, which is simply the differential of atmospheric refraction with respect to the effective refractive index. It leads to the formation of an image spectrum in the direction of the telescope altitude (Wilson 1996).

- Atmospheric turbulence or *seeing* quality, which introduces variations in the refractive index caused by density variations which in turn are due to temperature dependence with height or wind speed.

The last item on the list represents the most important limitation on image quality for sophisticated ground-based telescopes (Wilson 1999), because of its implications for the loss of spatial resolution; it also determines many of the design parameters of AO systems, such as aperture size and operating wavelength, and site selection.

If a flat wavefront is arriving at a certain telescope through the atmosphere from a point-like source, for example, a distant star, and assuming that the Earth's atmosphere is not a homogeneous static medium, but that it presents both spatial and temporal variations, such a wavefront will be distorted randomly by cells or bubbles of air with different indices of refraction. Several effects can be identified (Fig.1.1). Firstly, *scintillation* or the variation in brightness observed by the eye because of differences in the wavefront energy per unit of time. Secondly, *agitation* of the image in the focal plane, which corresponds to variations in the angle of the plane tangent to the wavefront. Finally, *smearing* or loss of spatial coherence at the pupil, which leads to an image size that is larger than would result from diffraction alone (Roddier 1981).

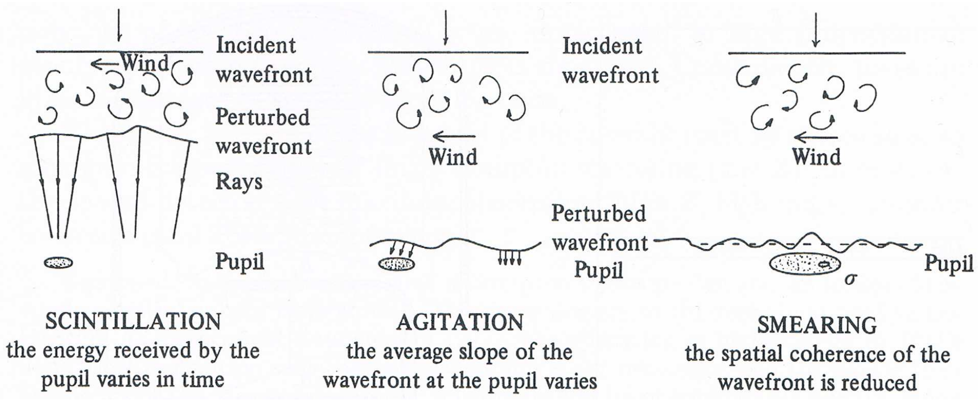

Figure 1.1: Effects of atmospheric turbulence on images —after Léna et al. (1998).

Most of the descriptions of atmospheric turbulence are based on the theory outlined by Kolmogorov (1941a,b). First, fluid flows are always turbulent if the well-known Reynolds number is larger than a critical value[1], which is always the case for the Earth's atmosphere. The inner and outer scales l_0 and L_0 can be defined as the size limits of a characteristic air eddy. In general, l_0 is of the order of a few

---

[1] $\mathcal{R} = VL/\mu \gtrsim \mathcal{R}_{cr}$. Where V is the flow velocity (for the Earth's atmosphere $V = 1m/s$), L is the length ($L \approx 15m$) and μ the kinematic viscosity of the fluid ($\mu = 15 \times 10^{-6} m^2/s$). For these values $R \sim 10^6$.

millimeters, below which value viscous heat dissipation commences; whereas typical values of L_0 are round a few tens of meters. The Kolmogorov theory describes the turbulence strength in a simple way as a function of the eddy size in the so-called inertial range between l_0 and L_0 by means of the structure function $D_x(r_1, r_2) \equiv \langle |x(r_1) - x(r_2)|^2 \rangle$, which is basically the expected value of the difference between the values of a function, x, measured at two different points r_1 and r_2.

If the turbulence itself is assumed to be homogeneous and isotropic, which is a valid approximation in the range $l_0 \ll r \ll L_0$, the structure function of the turbulent velocity field (v) can be written as:

$$D_v(r_1, r_2) \equiv \langle |v(r_1) - v(r_2)|^2 \rangle = \alpha \cdot f(|r_1 - r_2|/\beta), \tag{1.1}$$

where f is a dimensionless function with dimensionless arguments. Since the only two parameters that determine the strength and spectrum of the Kolmogorov turbulence are the rate of energy generation per unit mass, ϵ, with the dimensions $m^2 s^{-3}$, and the kinematic viscosity, μ, with the dimensions $m^2 s^{-1}$; and considering that from equation 1.1, α must have the dimensions of velocity squared and β of length, it follows from dimensional analysis that $\alpha = \mu^{1/2}\epsilon^{1/2}$ and $\beta = \mu^{3/4}\epsilon^{-1/4}$. In addition, since heat dissipation does not play any role in the inertial range, then D_v must be independent of the viscosity, μ. This is only possible if f has the form $f = K \cdot (|r_1 - r_2|/\beta)^{2/3}$ thus eliminating μ from equation 1.1, with K being a dimensionless constant:

$$D_v(r_1, r_2) = \alpha \cdot K \cdot (|r_1 - r_2|/\beta)^{2/3} = C_v^2 \cdot |r_1 - r_2|^{2/3}, \tag{1.2}$$

where $C_v = K \cdot \epsilon^{2/3}$ describes the turbulence strength. Equation 1.2 represents a universal description of the turbulence spectrum.

The velocity field (eq.1.2) mixes different layers of air, each of them with different temperatures, T, which must be in pressure equilibrium but have different densities and therefore have different indices of refraction, N. Consequently, the temperature fluctuations and refractive index variations must also follow Kolmogorov's Law with parameters C_T^2 and C_N^2:

$$D_T(r) = C_T^2 \cdot r^{2/3}, \tag{1.3}$$

$$D_N(r) = C_N^2 \cdot r^{2/3}. \tag{1.4}$$

with $C_N = (7.8 \cdot 10^{-5} P[\text{mbar}]/T^2[\text{K}]) \cdot C_T$ and $r \equiv |r_1 - r_2|$, since a homogeneous and isotropic random process is being considered (Quirrenbach 2006). Equation 1.4 contains a complete description of the statistical properties of the refractive index fluctuations within the inertial range. This is the so-called Kolmogorov-Obukhov "two-thirds law " (Kolmogorov 1941a; Obukhov 1941; Tatarski 1961).

In addition, Tatarski (1961) obtained the power spectral density, Φ_N, of the structural function D_N as a function of the spatial frequency, κ; i.e., Kolmogorov's Law for the spectral distribution corresponding to equation 1.4:

$$\Phi_N(\kappa) = 0.0365 C_N^2 \kappa^{-5/3}, \tag{1.5}$$

which is also valid in the inertial range for the 1-dimensional case, namely $l_0^{-1} \gg \kappa \gg L_0^{-1}$. For the more general 3-dimensional case $\Phi_N(\kappa) \sim \kappa^{-11/3}$.

Now that the statistical properties of the refractive index fluctuations have been obtained, it is possible to derive the statistical behavior of a plane, horizontal and monochromatic wavefront, $\psi(x)$, coming from a zenith star. Roddier's approach (Roddier 1981) is assumed: so the atmosphere is considered to be still and homogeneous except for a turbulent layer of thickness δh at height h. The scalar vibration, $\psi(x)$, can be expressed by its complex amplitude making use of the phase shift, $\phi(x)$, produced by the variations in the indices of refraction:

$$\psi(x) = e^{i\phi(x)}, \tag{1.6}$$

$$\phi(x) = k \int_h^{h+\delta h} n(x,z) dz, \tag{1.7}$$

where $k \equiv 2\pi/\lambda$ is the wavenumber corresponding to the observing wavelength and $n(x,z)$ is the index fluctuation within the layer at point x, at height h. The layer thickness, δh, must be greater than the individual eddies, so many independent variables will contribute to ϕ allowing to approximate it to a Gaussian distribution by means of the Central Limit Theorem. In contrast, δh must be small enough for diffraction effects to be negligible within the layer.

Both Tatarski (1961) and Roddier (1981) introduced the second-order moment or coherence function of the wavefront after passing through the layer. This can be expressed in terms of the phase structure function, $D_\phi(r)$, by:

$$\begin{aligned} B_h(r) &\equiv \langle \psi(x) \psi^*(x+r) \rangle = \\ &= \langle \exp i[\phi(x) - \phi(x+r)] \rangle = \\ &= \exp\left(-\frac{1}{2}\langle |\phi(x) - \phi(x+r)|^2 \rangle\right) = \\ &= \exp\left(-\frac{1}{2} D_\phi(r)\right). \end{aligned} \tag{1.8}$$

In equation 1.8 the relation $\langle \exp(iX) \rangle = \exp\left(-\frac{1}{2}\langle X^2 \rangle\right)$ is used, which is applicable to real Gaussian variables X with zero mean.

In order to calculate $D_\phi(r)$, the covariance $B_\phi(r)$ is defined:

$$\begin{aligned}
B_\phi(r) &\equiv \langle \phi(x)\phi^*(x+r)\rangle = \\
&= k^2 \int_h^{h+\delta h} \int_h^{h+\delta h} \langle n(x,z_1)n(x+r,z_2)\rangle dz_1 dz_2 = \\
&= k^2 \int_h^{h+\delta h} dz_1 \int_{h-z_1}^{h+\delta h-z_1} dz B_N(r,z),
\end{aligned} \tag{1.9}$$

where $z \equiv z_2 - z_1$ and $B_N(r,z)$ is the covariance of the refractive index variations. For layer thicknesses, δh, that are large enough compared to the correlation scale of the fluctuations, the integration can be extended to the infinite limit (Roddier 1981):

$$B_\phi(r) = k^2 \delta h \int_{-\infty}^{\infty} B_N(r,z) dz. \tag{1.10}$$

Since $r \equiv |r_1 - r_2|$, it is easy to prove that $D_x(r) = 2\,(B_x(0) - B_x(r))$ (Quirrenbach 2006). Applying this relation to $D_\phi(r)$, $D_N(r,z)$ and $D_N(0,z)$ the following expression is obtained:

$$\begin{aligned}
D_\phi(r) &= 2[B_\phi(0) - B_\phi(r)] = \\
&= 2k^2\delta h \int_{-\infty}^{\infty} [B_N(0,z) - B_N(r,z)]\, dz = \\
&= 2k^2\delta h \int_{-\infty}^{\infty} [(B_N(0,0) - B_N(r,z)) - (B_N(0,0) - B_N(0,z))]\, dz = \\
&= 2k^2\delta h \int_{-\infty}^{\infty} [D_N(r,z) - D_N(0,z)]\, dz.
\end{aligned} \tag{1.11}$$

Inserting equation 1.4 into equation 1.11 gives:

$$\begin{aligned}
D_\phi(r) &= k^2 \delta h C_N^2 \int_{-\infty}^{\infty} \left[(r^2 + z^2)^{1/3} - |z|^{2/3} \right] dz = \\
&= 2.914 k^2 \delta h C_N^2 r^{5/3},
\end{aligned} \tag{1.12}$$

which is the desired expression for the structure function of phase fluctuations due to Kolmogorov turbulence in a layer of thickness δh. Finally, inserting equation 1.12 into equation 1.8 and considering that the star which emitted the wavefront is not at the zenith, but a certain angular distance, Z, from the zenith, and also considering that the turbulence extends across the whole depth of the atmosphere, thus $C_N^2 \delta h \Rightarrow \int C_N^2(h)dh$, it is possible to obtain:

$$B(r) = \exp\left[-\frac{1}{2} \left(2.914 k^2 (cosZ)^{-1} r^{5/3} \int C_N^2(h) dh \right) \right]. \tag{1.13}$$

Equation 1.13 is of fundamental importance for determining the effects of turbulence in long-exposure images. From this expression the following conclusions can be deduced (Léna et al. 1998):

- The phase is the only perturbed parameter in the vicinity of the layer, this results in *smearing* and *agitation* (Fig.1.1). Further away, both the phase and the amplitude suffer variations and then *scintillation* occurs. When the turbulence is not very large, the amplitude, $|\psi(x)|$, can be neglected.

- On leaving the layer, the correlation function has a complex amplitude which determines the random phase distribution. This function has an isotropic profile, so the correlation length, x_c, can be defined as:

  $$\frac{\langle \psi_h(x + x_c)\psi_h^*(x)\rangle}{\langle |\psi_h(0)|^2\rangle} \sim \frac{1}{e} \Longrightarrow x_c \sim (1.457k^2 C_N^2 \delta h)^{-3/5}. \tag{1.14}$$

  Typical values of x_c are around a few tens of centimeters, which implies a significant reduction in the spatial coherence of the incidental wave.

- The coherence length is highly dependent on the wavelength, $x_c \propto \lambda^{6/5}$, (eq.1.14).

1.1.1 The Fried parameter

Fried (1965) represents the wavefront when it arrives in terms of a truncated sequence of polynomials (closely related to Zernicke's polynomials), each of which, or each group of which, represents a specific geometric shape over the telescope pupil: a uniform phase change, a tilt change, defocus or a hyperbolic change. In addition, the phase structure function (eq.1.12) is expressed in terms of the so-called "Fried parameter" which can be defined as follows:

$$r_0 \equiv \left[0.423k^2(cosZ)^{-1} \int C_N^2(h)dh\right]^{-3/5}, \tag{1.15}$$

which transforms equations 1.13 and 1.12 into:

$$B(r) = \exp\left[-3.44\left(\frac{r}{r_0}\right)^{5/3}\right], \tag{1.16}$$

$$D_\phi(r) = 6.88\left(\frac{r}{r_0}\right)^{5/3}. \tag{1.17}$$

It should be noted that the dimension of r_0 is length and that it is proportional to $\lambda^{6/5}$, thus it is highly dependant on the wavelength of the incident light. This strong dependence makes telescopes more coherent at infrared wavelengths than in the visible part of the spectrum ($r_0 = 10cm$ at $0.5\mu m$ whereas $r_0 = 8.4m$ at $20\mu m$), i.e., given a certain diameter, a telescope will give images closer to the diffraction limit in the infrared than in the visible.

In an additional publication, Fried (1966) offered a new treatment in terms of the normalized optical transfer function[2] (OTF), $S(f)$. The OTF can be divided into two terms, one corresponding to the telescope and the other to the atmosphere. The latter is described by the Fourier transform of the wavefront coherence function (eq.1.13) and the former by:

$$T(f) = \frac{1}{\pi D^2} \int P(u) P^*(u+f) du, \tag{1.18}$$

where f is frequency, u is defined as $u \equiv x/\lambda$ and $P(u)$ is a complex function which describes a telescope aperture of a certain diameter D. Consequently, the total OTF is given by:

$$S(f) = B(f) \cdot T(f). \tag{1.19}$$

Finally, the resolving power, R, (which is inversely proportional to the angular resolution) is given by (Roddier 1981):

$$R \equiv \int S(f) df = \int B(f) T(f) df. \tag{1.20}$$

In the ideal case of the absolute absence of turbulence, $B(f)$ can be assumed to be negligible; so the resolving power would only be limited by diffraction:

$$R_d = \int T(f) df = \frac{1}{\pi D^2} |P(u) du|^2 = \frac{\pi}{4} \left(\frac{D}{\lambda} \right)^2 . \tag{1.21}$$

For an infinitely large telescope, the opposite effect would be experienced and the seeing-limited resolving power could be obtained as:

$$R_s = \int B(f) df = \int \exp \left[-3.44 \left(\frac{\lambda f}{r_0} \right)^{5/3} \right] df = \frac{\pi}{4} \left(\frac{r_0}{\lambda} \right)^2 , \tag{1.22}$$

where $r = \lambda f$ has already been introduced in equation 1.16.

From a direct comparison of equations 1.21 and 1.22, the significance of the Fried parameter is clear (Quirrenbach 2006): *the resolution of seeing-limited images obtained through an atmosphere with turbulence characterized by a Fried parameter r_0 is the same as the resolution of diffraction-limited images observed with a telescope of diameter r_0.*

Therefore, r_0 is a method by which to characterize the size of the turbulence cells, i.e., the Fried parameter is the distance over which the wavefront is not considerably perturbed. As was pointed out previously, r_0 is wavelength-dependent, giving better results at infrared wavelengths. It is also site-dependent, at good seeing locations (such as Mauna Kea or Cerro Paranal) r_0 is about $20cm$ at $500nm$, which corresponds to an FWHM of 0.6 arcsec; even values of $100cm$ have been reported at $2.2\mu m$. However, only when $r_0 > D$ are the seeing effects negligible; so even a value of $r_0 \sim 1m$ would be too small for large-diameter telescopes such as the 10-m Kecks.

---

[2] The Fourier transform of the telescope point spread function (PSF).

1.1.2 Temporal evolution. Long- and short-exposure images

The Fried parameter models the size scale of the eddies over the telescope pupil. However, it is also necessary to consider the evolution of the turbulences over time. A typical approximation is to assume that the turbulence itself does not change over the time needed by the wind to blow it out of the telescope pupil, i.e., temporal variations are due of the transport of *frozen* patterns across the aperture by the wind. This is the so-called *Taylor hypothesis of frozen turbulence*. Hence, the wind speed, v, in the layer is the parameter which defines the time evolution, τ_0, of an eddy of size r_0:

$$\tau_0 \sim r_0/v. \tag{1.23}$$

Of course, just as for r_0, time evolution is also wavelength dependent, i.e., $\tau_0 \propto \lambda^{6/5}$. If Taylor's hypothesis is valid, the spatial structure function (eq.1.17) and its phase power spectrum can be converted into their temporal counterparts:

$$D_\phi(r) = 6.88\left(\frac{t}{\tau_0}\right)^{5/3}, \tag{1.24}$$

$$\Phi_\phi(f) = 0.077\tau_0^{-5/3} f^{-8/3}. \tag{1.25}$$

These equations tell the residual phase error that can be expected to correct the turbulence with a servo loop of a certain bandwidth.

Typical values of v are around $\sim 10m/s$, which means values for the evolution time of just a few milliseconds. Observations with exposure times $t \gg \tau_0$, i.e., long-exposure images, average over the random turbulent process. In contrast, short-exposure images, $t \ll \tau_0$, produce images via a single instantaneous observation of the atmosphere with the typical *speckle* pattern (Fig.1.2).

The number of speckles in each observation is approximately D/r_0 and they contain considerable high-frequency spatial information close to the diffraction limit. Another image would result in another set of speckles, but randomly displaced with respect to the previous one (McLean 1997). If numerous short exposures are taken and added together or an equivalent long exposure is taken, then all the speckles blend together to form a seeing disk of width $\sim \lambda/r_0$ instead of multiple small individual speckles of width $\sim \lambda/D$ (Fig.1.3). It is easy to verify that the larger the Fried parameter, r_0, the fewer the speckles; the longer the evolution time, τ_0, and the better the angular resolution for a long exposure.

The diffraction-limit information contained in numerous speckles can be recovered using different Fourier analysis techniques; such methodology is called speckle interferometry. Labeyrie (1970) established the possibility of recovering the object power spectrum from the autocorrelation function of a large series of short-exposure frames; if they have been acquired over a time scale of less than τ_0, the phase content of each frame can be considered stationary and time-independent. Obtaining information from the object is then relatively straightforward. For instance, for a binary system, physical parameters such as separation, angular position and magnitude difference can be retrieved

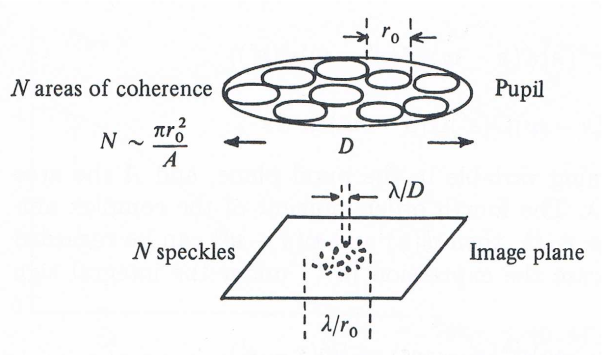

Figure 1.2: Speckles. Note the different sizes of the various elements represented: pupil, speckles, turbulence —after Léna et al. (1998).

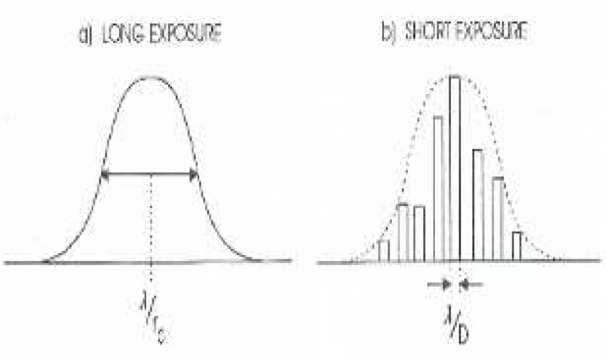

Figure 1.3: Long- and short-exposures profiles —after McLean (1997).

by a weighted least-square method (Horch et al. 1996).

Unfortunately, Labeyrie's method removes the phase of the object, which must also be recovered in order to obtain a diffraction-limited image of the object. Different techniques have been proposed to recover the wavefront phase in speckle interferometry. The most successful has probably been the triple correlation analysis or bispectrum method (Lohmann et al. 1983) and its relaxation variant (Meng et al. 1990) (see Figure 1.4).

There is a huge range of applications of speckle interferometry in science, either in the visible or the infrared region of the spectrum. To offer a brief overview, the following papers can be cited: Hartkopf et al. (2000), Mason et al. (2001) and Horch et al. (2009, 2012).

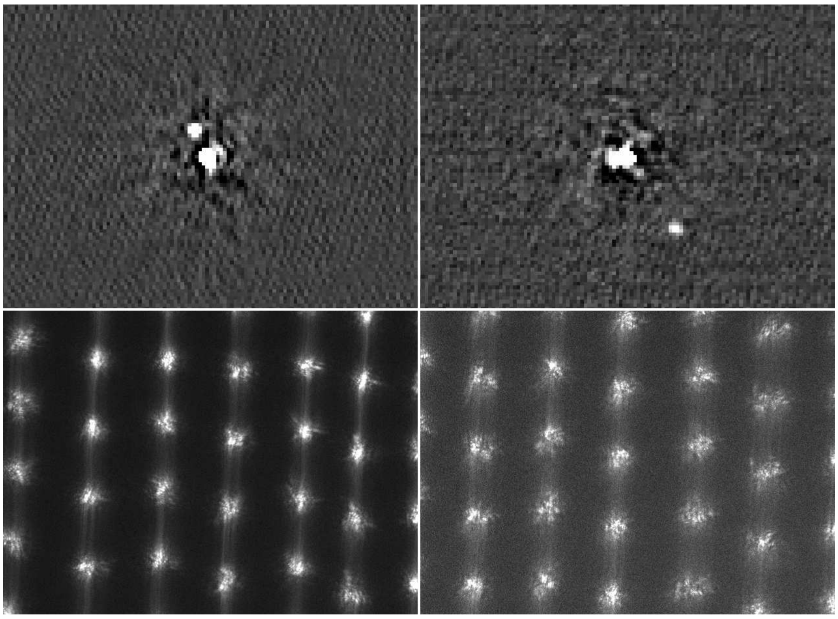

Figure 1.4: Upper left: Reconstruction of HIP 005674 as a binary system by means of the bispectrum method, $\rho = 0.144$ arcsec, $\Delta m = 2.04$ (562*nm.*), $\Delta m = 1.89$ (692*nm.*). Upper right: Reconstruction of HIP 115751, $\rho = 0.422$ arcsec, $\Delta m = 3.03$ (562*nm.*), $\Delta m = 2.98$ (692*nm.*). Lower left: corresponding HIP 005674 specklegrams. Lower right: corresponding HIP 115751 specklegrams. (Horch et al. 2009).

1.1.3 Angular anisoplanatism

The light from two stars separated by an angle θ arrives at the telescope pupil via two different optical paths and, therefore, is affected by different phase variations. This effect is the so-called angular anisoplanatism and limits the field that can be corrected by an AO system (see section 1.2).

Two rays coming from two different stars are separated by $r(d) = h(cosZ)^{-1}\theta$, where d is their separation, h is the height and Z is the zenital angle. At $h = 0$, i.e., at the pupil, the two rays coincide. To calculate the phase variance between the two rays this expression is inserted into equation 1.12 (considering the zenital angle to be Z) and integrate over the height, h:

$$\begin{aligned}\langle\sigma_{\phi}^{2}\rangle &= 2.914k^{2}(cosZ)^{-1}\int C_{N}^{2}(h)\left(\frac{\theta h}{cosZ}\right)^{5/3}dh = \\ &= 2.914k^{2}(cosZ)^{-8/3}\theta^{5/3}\int C_{N}^{2}(h)h^{5/3}dh = \\ &= \left(\frac{\theta}{\theta_{0}}\right)^{5/3},\end{aligned} \tag{1.26}$$

where the isoplanatic angle, θ_0, has been defined as the angle for which the variance of the relative phase is $1\,rad^2$:

$$\theta_{0} \equiv \left[2.914k^{2}(cosZ)^{-8/3}\int C_{N}^{2}(h)h^{5/3}dh\right]^{-3/5}. \tag{1.27}$$

Equation 1.27 can be expressed in terms of the Fried parameter (equation 1.16) as:

$$\theta_{0} = 0.314(cosZ)\frac{r_{0}}{H}, \tag{1.28}$$

where H is defined as the mean effective turbulence height:

$$H \equiv \left(\frac{\int C_{N}^{2}(h)h^{5/3}dh}{\int C_{N}^{2}(h)dh}\right)^{3/5}. \tag{1.29}$$

The isoplanatic angle, θ_0, can be interpreted physically as the maximum angle by which two stars can be separated to assume they both produce the same PSF corrected by AO. In contrast, for long exposures which represent averages over many observations of atmospheric turbulence, the PSFs are almost identical for angles much larger than θ_0. Equation 1.28 shows that θ_0 is most affected by high-altitude turbulence. Furthermore, it also depends on the wavelength, $\lambda^{6/5}$, and on the zenital angle, Z, more strongly than on r_0. In general, the anisoplanatism is less severe for low spatial frequencies which are corrected much better by AO systems than high spatial frequencies are.

1.1.4 The Strehl ratio

Astronomers compare seeing-limited and diffraction-limited images using the Strehl ratio (SR) defined as the quotient of the intensity peak of an observed point source and the theoretical maximum peak intensity of a perfect imaging system at the diffraction limit. It can be expressed by the following equation[3] (Wilson 1996; Hardy 1998):

$$SR \equiv \frac{I(P)}{I_0} = \frac{1}{\pi^2} \left| \int_0^1 \int_0^{2\pi} \rho e^{ik\psi(\rho,\theta)} d\rho d\theta \right|^2 , \tag{1.30}$$

where $I(P)$ is the intensity of the seeing-limited image at the focal point P and I_0 is the intensity in the absence of aberrations and turbulences (i.e., $\psi = 0$). The wavefront aberration function $\psi(\rho, \theta)$ refers to a circular aperture expressed in polar coordinates and k is again the wavenumber corresponding to the observing wavelength, λ.

In the absence of aberration, the maximum in intensity will be at the Gaussian focus. If aberration is present and large, there will be an intensity maximum at the diffraction focus which cannot be unique. However, if the aberration is sufficiently small, there will be a unique diffraction focus with a clearly defined intensity maximum. In such a case, equation 1.30 can be used as a parameter to measure the effects of a small degree of aberration over the image quality. This concept was first introduced by Strehl (1902).

In general, there is no explicit knowledge of the wavefront aberration function, $\psi(\rho, \theta)$ in equation 1.30. This is especially true in the case of AO where the random errors are usually described statistically. Born & Wolf (1964) derives an expression to approximate equation 1.30 to a more simple form by expanding the aberration function by a Taylor's series:

$$SR = \frac{1}{\pi^2} \left| \int_0^1 \int_0^{2\pi} \rho \left[1 + ik\psi(\rho,\theta) + \frac{1}{2}[ik\psi(\rho,\theta)]^2 + \ldots \right] d\rho d\theta \right|^2 . \tag{1.31}$$

Let $\overline{\psi^n}$ be the average wavefront error of the nth power of ψ over the pupil, with respect to a reference sphere centered on P:

$$\overline{\psi^n} = \frac{\int_0^1 \int_0^{2\pi} \rho \psi^n(\rho,\theta) d\rho d\theta}{\int_0^1 \int_0^{2\pi} \rho d\rho d\theta} = \frac{1}{\pi} \int_0^1 \int_0^{2\pi} \rho \psi^n(\rho,\theta) d\rho d\theta. \tag{1.32}$$

If the third and higher-order terms in equation 1.31 are neglected, then the SR can be expressed as:

$$SR \approx \left| 1 + ik\overline{\psi} - \frac{1}{2} k^2 \overline{\psi^2} \right|^2 = 1 - k^2 \left[\overline{\psi^2} - (\overline{\psi})^2 \right]. \tag{1.33}$$

---

[3] In general, SR is expressed as a percentage.

The mean-square wavefront error is defined as $(\Delta\psi)^2 = \overline{\psi^2} - (\overline{\psi})^2$. Finally, setting the standard deviation of the phase as $\sigma_\phi = k\Delta\psi$, the SR can be approximated as:

$$SR \approx 1 - \sigma_\phi^2. \tag{1.34}$$

This expression is only valid for very small phase errors of up to about $0.6 rad\, rms$. It tends to 0 at a phase error of $1\, rad$, which is clearly unrealistic. Maréchal & Françon (1960) obtained a similar expression for SR in the presence of small aberrations:

$$SR \approx \left[1 - \frac{1}{2}\sigma_\phi^2\right]^2, \tag{1.35}$$

which gives basically the same result as equation 1.34 for values of σ_ϕ of up to $0.5\, rad\, rms$. It tends to 0 at $1.41\, rad\, rms$. An optical system can be considered well corrected when $SR \geq 0.8$, which corresponds to $\sigma_\phi^2 \leq 0.2$ (eq.1.34), corresponding to a wavefront error of $\lambda/14$ or less. Maréchal & Françon (1960) and Welford (1974) provide similar results. This criterion is acceptable for fixed optical systems such as mirrors or lenses, but is excessively restrictive for AO systems. SRs of up to 50% in the K-band are considered typical good values (e.g. Hayward et al. 2001). Moreover, SR values of $\sim$ 80% have been measured for the 8.4-m Large Binocular Telescope in the H-band (Esposito et al. 2010).

Equation 1.34 represents the first two terms of the Taylor's series expansion of an exponential function which yields an alternative expression for SR that is the so-called "*extended Maréchal approximation*" :

$$SR \approx e^{-\sigma_\phi^2}. \tag{1.36}$$

This equation is valid over a larger range of phase errors than the other two, up to $2\, rad\, rms$, and it is one of the most frequent expression for SR in AO. It is clear that, in the ideal case of the absence of atmospheric turbulence, $\psi(\rho, \theta) = cte$ so $\sigma_\phi = 0$ and $SR = 1$. Moreover, for strongly varying $\psi(\rho, \theta)$, $SR \ll 1$ and, in general, $0 \leq SR \leq 1$ (equation 1.30). For any given $\psi(\rho, \theta)$, the SR tends to be larger for longer wavelengths, λ, i.e., for smaller wavenumbers, k.

It should be remarked that the SR is not the only valid parameter when it comes to obtaining a measurement of the image quality in AO. Other possible metrics are the 80% encircled energy (EE), defined as the angular diameter containing 80% of the total energy, which has the advantatge of relating two meaningful parameters: the sensitivity and the resolution; however, it is dependant of the wavelength. A second possible metric would be the full width half maximum (FWHM), which gives the extent of the main PSF lobe (excluding wings) as the difference between the two extreme values at which the PSF reaches half of its maximum value.

Dierickx (1992) compared the effects of spherical aberration in the presence of atmospheric turbulence on the 50% and 80% encircled energy with those on the equivalent SR and concluded that the SR is far preferable as a quality criterion, since the former are not very sensitive to the wavefront errors introduced by the telescope itself. Brown (1979) reaches a similar conclusion. Finally, Dierickx

(1992) introduces the central intensity ratio (CIR) given by the quotient:

$$CIR \equiv \frac{SR}{SR_0}, \tag{1.37}$$

where SR is the SR of the total system of atmosphere, telescope and site. The parameter SR_0 is the SR of the equivalent perfect telescope limited only by diffraction, *in the same turbulent atmosphere*. The metric CIR is normalized to unity for the performance of an aberration-free, diffraction-limited telescope; so it is more practical in those cases where SR yields very small values.

The CIR can be seen as the area under the modulation transfer function (MTF[4]) curve (Brown 1981, Fig.1.5). This contains the same information as the PSF; it measures the degradation of the optical amplitude with frequency, just as if it were a filter function applied over some signal.

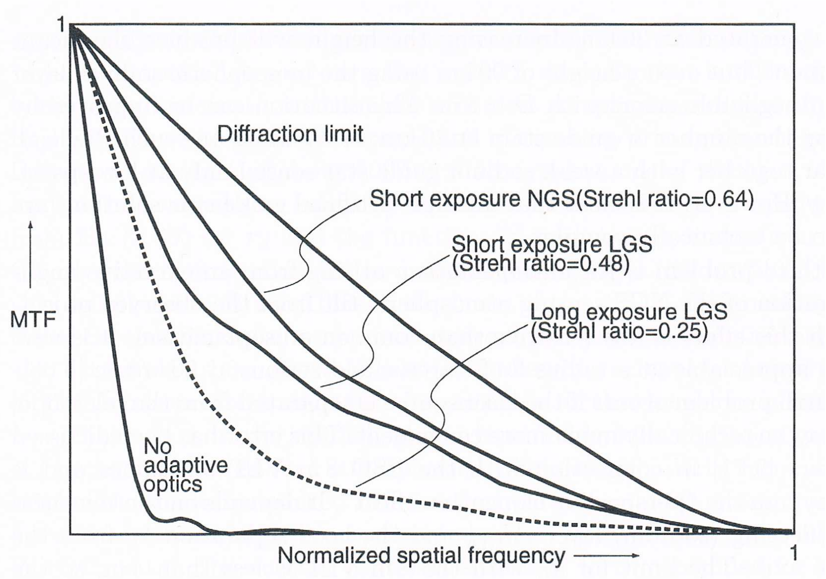

Figure 1.5: Improvement of the MTF with respect to the SR —after Wilson (1999).

1.2 Adaptive optics systems

As has been shown in the previous section, atmospheric turbulence limits the spatial resolution achieved by ground-based telescopes. This limitation can be mitigated by reducing the amount of atmosphere over the telescope; that is, by building observatories atop high mountains, or launching highly-expensive and technologically-restricted telescopes into orbit outside the Earth’s atmosphere. However, Horace W. Babock was the first person to theorize about the applicability of an AO system: *if we had the means of continually measuring the deviation of rays from all parts of the mirror, and of amplifying and feeding back this information so as to correct locally the figure of the mirror in response to the schlieren pattern, we could expect to compensate both for the seeing and for any inherent imperfection of the optical figure* (Babcock 1953). This theoretical system was first imple-

---

[4] This is the real part of the optical transfer function (OTF, eq.1.19)

mented in the 1.6-m telescope on mount Haleakala (Maui, Hawaii) in 1982, almost 30 years after Babcock's statement just quoted.

Figure 1.6 shows the basic scheme of an AO compensating system. A wavefront sensor (WFS) measures the deviations of the optical path from the flat wave and sends this information to a real-time computer. After analyzing this information, the feedback loop is completed by modifying the shape of a deformable mirror (DM) in order to correct the deviations and better concentrate the luminic energy on one single image element.

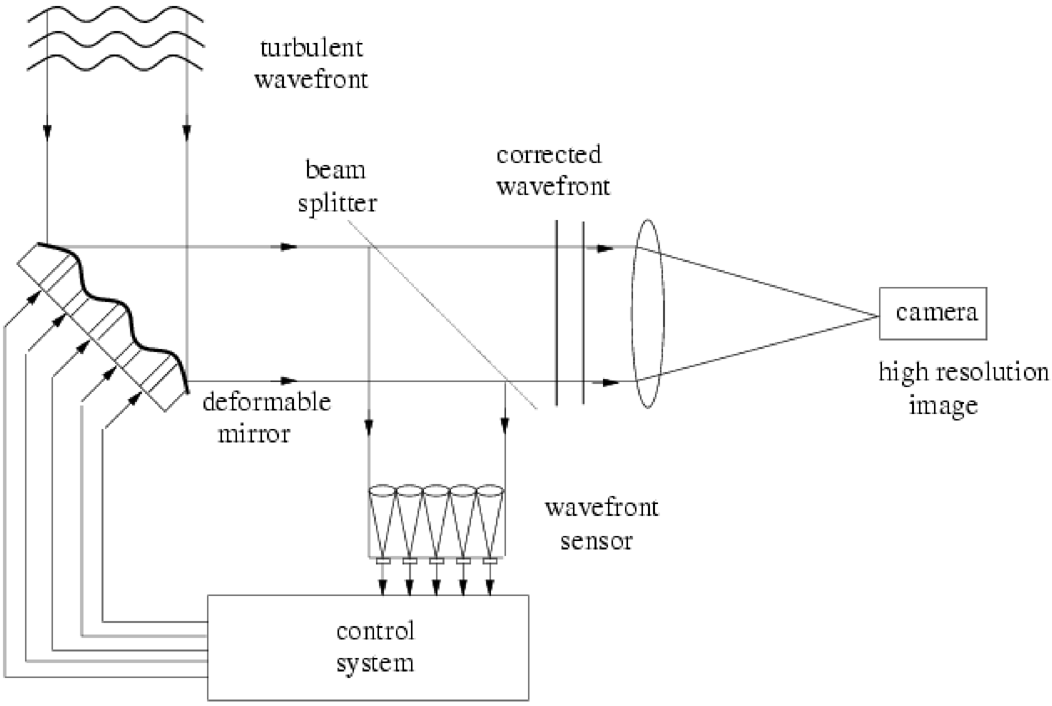

Figure 1.6: A simple scheme for implementing adaptive optics using a deformable mirror.

In practice, the phase variations of the wavefront are described in terms of the so-called Zernicke polynomials, $Z_j(n, m)$, where n is the degree of a radial polynomial and m is the azimuthal frequency of a sinusoidal term. Noll (1976) gives normalized versions of the Zernicke polynomials. Table 1.1 lists the low-order terms and explains their meaning. The last column in the table gives the mean square residual amplitude, Δ_j, in the phase variations at the telescope entrance caused by Kolmogorov turbulence after the removal of the first j elements; the normalization factor is $S = (D/r_0)^{5/3}$. For $j > 10$, the value of Δ_j can be approximated by:

$$\Delta_j \approx 0.2944 j^{-0.866} \left(\frac{D}{r_0}\right)^{5/3} \text{rad}^2. \tag{1.38}$$

By substituting the appropriate terms and taking the square root, Table 1.1 yields the rms phase variation for each mode of correction. When there is no correction, the rms phase variation is $0.162(D/r_0)^{5/6}$ waves, while after tip–tilt correction, the rms phase variation is reduced to $0.053(D/r_0)^{5/6}$ waves. In accordance with this, in order to achieve an SR as high as 80% under typical seeing conditions, a huge number of terms must be corrected, which basically implies a high degree of complexity of the whole AO system.

Table 1.1: Modified Zernicke polynomials and the mean square residual amplitude for Kolmogorov turbulence after removal of the first *j* terms.

Z_j	n	m	Expression	Description	Δ_j/S
Z_1	0	0	1	constant	1.0300
Z_2	1	1	$2r\sin\phi$	x-tilt	0.5820
Z_3	1	1	$2r\cos\phi$	y-tilt	0.1340
Z_4	2	1	$\sqrt{3}(2r^2-1)$	defocus	0.1110
Z_5	2	2	$\sqrt{6}r^2\sin 2\phi$	0^{o} astigmatism	0.0880
Z_6	2	2	$\sqrt{6}r^2\cos 2\phi$	45^{o} astigmatism	0.0648
Z_7	3	1	$\sqrt{8}(3r^3-2r)\sin\phi$	x-coma	0.0587
Z_8	3	1	$\sqrt{8}(3r^3-2r)\cos\phi$	y-coma	0.0525
Z_9	3	3	$\sqrt{8}r^3\sin 3\phi$	x-trifoil	0.0463
Z_{10}	3	3	$\sqrt{8}r^3\cos 3\phi$	y-trifoil	0.0401
Z_{11}	4	0	$\sqrt{5}(6r^4-6r^2+1)$	third order spherical	0.0377

Shellan (2004) performed several simulations to represent the SR as a function of the telescope diameter and the Fried parameter for a wide range of corrected terms in the Zernicke polynomials and number of actuators on the DM. He found that maximum on-axis intensity is achieved when the telescope diameter is such that the actuator spacing is equal to $\sim 2r_0$. In contrast, the optimum value of D/r_0 is equal to $1.79N_r + 2.86$, where N_r is the highest Zernike radial mode corrected.

Therefore, an AO system must be designed to correct as many Zernicke polynomial terms as possible (Fig. 1.7). This design is determined by the atmospheric parameters derived from Kolmogorov's theory, i.e., the Fried parameter, r_0, the isoplanatic angle, θ_0, the time evolution, τ_0, and in the case of future 30- or 40-m telescopes such as ELT, the outer scale, L_0, as well.

The AO feedback loop shown in Figure 1.6 must be iterated several hundred times per second in order to correct temporal variation in the optical path set by τ_0. However, the number and size of the WFS resolution elements (subapertures) and the DM elements (actuators) projected on the telescope aperture must approximately match r_0. Finally, WFS makes use of a single star or artificial laser star measurements to record variations in the wavefront direction. Such a star must be separated from the science object by a distance less than the isoplanatic angle, θ_0, in order to diminish the differences between the object and the calibrator PSFs across the field of vision. Figure 1.8 shows Jupiter's satellite Io imaged with the 10-m AO Keck telescope.

Hence, an AO system is formed of the following elements:

1. WFS: this is designed to provide an electrical signal which can accurately describe the wavefront shape. In general, bearing in mind that the wavefront is nearly achromatic, wavefront measurements are performed at visible wavelengths by phase-sensitive optical devices with

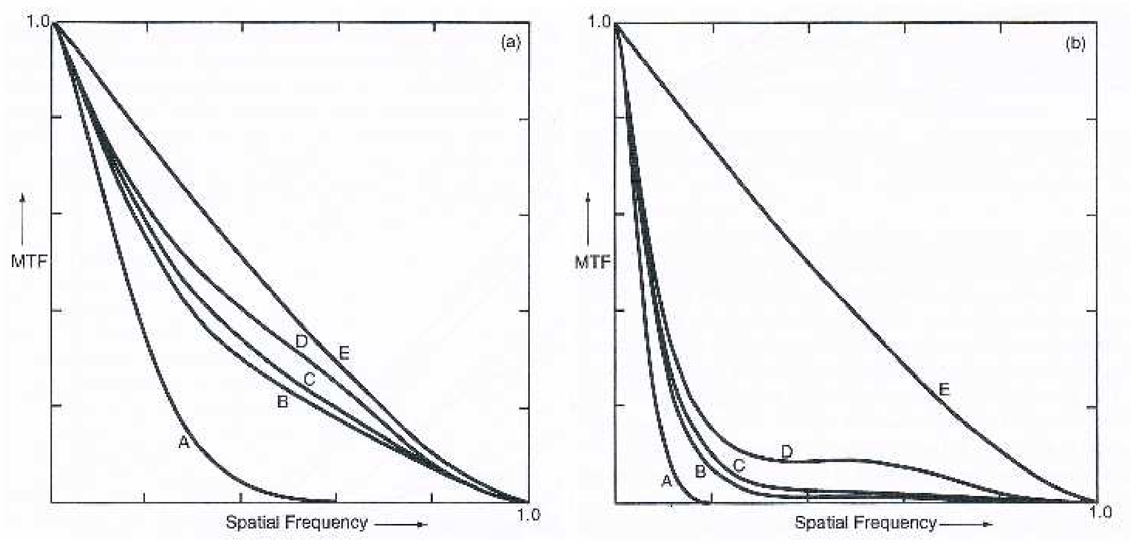

Figure 1.7: Efficiency of AO using a small number of modes. Left: $D/r_0 = 2$. Right: $D/r_0 = 7$. Curve A: no correction. Curve B: correction up to term Z_3. Curve C: correction up to term Z_4. Curve D: correction up to term Z_6. Curve E: perfect diffraction-limited performance —after Wilson (1999).

low-noise and high-quantum-efficiency photon detectors such as CCDs. Currently, three WFS designs are in use. First, there is the Shack-Hartmann WFS, which employs an array of lenses which produces a corresponding array of spots whose positions determine the wavefront shape. Second, there is a pyramid WFS, which makes use of a prism to deviate the aberrated ray to different detectors. Thus, the intensity distributions in the multiple pupil images are a measure of the ray's slope. Lastly, a curvature WFS measures intensity distributions in two different planes on either side of the focus, corresponding to the curvature of the wavefront.

All these three types of WFS work with broad-band light, but they differ in their dynamic range

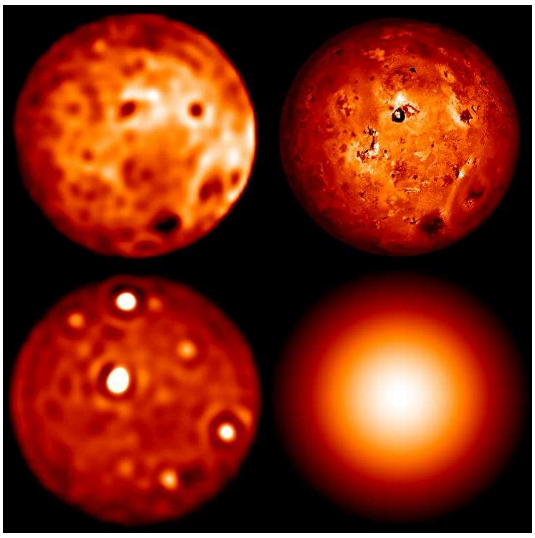

Figure 1.8: Upper left: Io image taken with Keck adaptive optics; K-band, 2.2μm. Upper right: Io image based on visible light taken with Galileo spacecraft orbiter. Lower left: Io image taken with Keck adaptive optics; L-band, 3.5μm. Lower right: Io image taken without Keck adaptive optics.

and sensitivity. Nowadays, the pyramid WFS is considered to be the most suitable for AO imaging, since its sensitivity hardly changes at any spatial frequencies. The Sharck-Hartmann and curvature WFS sensitivities degrade at low spatial frequencies, this is especially so for the latter; however, the curvature WFS exhibits better performance at high-spatial frequencies than the Sharck-Hartmann WFS (Guyon 2005, 2007).

2. Wavefront reconstruction techniques (WRTs): these tackle the calculation of the appropriate correction vector, **v**, which the DM must be provided with from the measurement vector, **s**, obtained by WFS. It consists of the resolution of the following linear system:

   $$Dv = s + n, \tag{1.39}$$

   where **D** is the matrix which relates the WFS to the DM; and **n** is Gaussian uncorrelated noise. **v** can be resolved from equation 1.39 by obtaining a pseudoinverse version of **D**, which is to be multiplied by **s**. Unfortunately, this simple method provokes noise amplification and techniques of filtering, weighting and modal decomposition must therefore also be used in order to obtain $\mathbf{D}^{-1}$ (Gendron & Léna 1994; Verinaud & Cassaing 2001; Wallner 1983).

   Moreover, the complexity of such techniques increases with $O(n^2)$, where n is the number of degrees of freedom of the system. Since the reconstruction must be carried out every $1ms$, and the delay between measurement and correction must be as short as possible ($\sim$ 1 time step), the computational load becomes very high quickly. In order to reduce such complexity, several approaches have been proposed including the FFT-based reconstructor (Poyneer et al. 2002), which scales with $O(n \log_2 n)$; the fractal iterative method (Thiébaut & Tallon 2010); and the cumulative reconstructor (Rosensteiner 2011). Both of the latter two scale with $O(n)$.

3. DM: this usually consists of an array of actuators which deform a thin optical surface by means of contractions and expansions. The parameters which define DM performance are the number of actuators, the spacing between them, the stroke and the response time. Spacing and response must be in accordance with r_0 and τ_0, while the stroke and the number of actuators both scale with the aperture diameter. Current DMs for typical 8- or 10-m telescopes have several thousand of actuators. The future 30- and 40-m telescopes will increase this number by a factor of one hundred.

   There are three technologies available for building DMs. First, the secondary mirror of the telescope can be replaced by an adaptive secondary mirror. This leads to large designs of 1-m diameter with actuators separated by a few centimeters. However, they achieve high transmission and low thermal emissivity by avoiding extra relay optics. Second, there is the medium-sized piezo DM, which are smaller than the previous type but which are affected by hysteresis and thermal expansion. In this case the actuators are separated by a few centimeters and they have less stroke, although enough to perform suitable correction. Third, micro-optical-electrical-mechanical systems are significantly smaller with a spacing of a few hundred microns. They are produced using semiconductor technologies and do not suffer from hysteresis; furthermore, they have almost instantaneous response times. However, their stroke is not very high and they usually need a second larger stroke DM with only a few actuators, which flattens the wavefront as a first, prior stage.

4. Laser guide star (LGS): if a suitable reference star of at least 15*mag* is not available in the field of view (FOV), the AO system can not function. In such situations, the system must be provided with an "artificial star" using laser beams (Fig. 1.9). This consists of emitting a pulsed laser beam, tuned to the sodium wavelength 589*nm*, projected through the telescope and focused at an altitude of $\sim 92km$ which excites the so-called sodium layer in the upper atmosphere. Resonance fluorescence in this layer produces a glowing artificial star. Another possibility is to excite dense regions of the atmosphere at about $\sim 30km$ by Rayleigh scattering of the laser beam. The use of an LGS was first proposed for astronomy by Foy & Labeyrie (1985) and their usage is now generalized.

   Notwithstanding, an LGS also has considerable drawbacks compared to natural reference stars. First, due to the finite distance between the telescope and the LGS, the backscattered beam does not sample the full aperture at the height of the turbulent layers. This focal anisoplanatism is more severe for larger apertures and higher turbulent layers. Second, an LGS cannot be used to measure the tip or tilt of the wavefront, since the contribution of the upward projection jitter cannot be disentangled from the measurement.

Figure 1.9: A laser beam emerging from the 8.2-m Very Large Telescope at the European Southern Observatory (ESO) in Cerro Paranal (Chile).

Practical implementation of AO did not occur until more than 20 years after Babcock's work. AO concepts were first proved and tested in the mid-70s by Buffington et al. (1977a,b), who used a one-dimensional, 6-element segmented mirror with segment piston control; and by Hardy et al. (1977) and Hardy (1978), who used a two-dimensional, 21-actuator continuous surface DM, combined with a shearing interferometer WFS.

However, it was necessary to wait until the late 80s to see the birth of the first fully-operative AO system for professional astronomy. The European Southern Observatory (ESO) was responsible for building the first such AO system: the so-called, COME-ON. It consisted of a DM with 19 actuators and a Sharck-Hartmann sensor. The COME-ON system was initially tested on the 1.52-m telescope at the Observatoire de Haute-Provence (Rousset et al. 1990). Later, it was updated to 52 actuators and a larger corrective bandwidth of $400Hz$ in order to be installed in the 3.6-m telescope at La Silla (Chile) (Rigaut et al. 1991). After a simplification and the use of control computers, the name of the system was changed to ADONIS. Simultaneously, a new technique was being developed by the National Optical Astronomy Observatory (NOAO). Rigaut et al. (1997; 1998) successfully mounted a curvature sensor in the Canada-France-Hawaii telescope (CFHT) at the Mauna Kea Observatory (Hawaii). That system, denominated PUEO, had a bimorph mirror with 19 actuators and a servo bandwidth of $100Hz$.

After these developments, it became impossible to conceive of telescopes with primary mirrors larger than 3 or 4 m which do not incorporate their own AO systems. Therefore, in 2003, the NAOMI system was installed in the 4.2-m Herschell telescope at the Instituto Astrofísco de Canarias (IAC). MACAO and NAOS were installed in the Very Large Telescope (VLT), at La Silla. The 3-m Shane telescope at Lick Observatory was a pioneer in developing an LGS system. The ALTAIR AO system was mounted in Gemini North (2002) replacing the Hokupa'a AO system. More recently, the 8.2-m Keck-II, 8.2-m Subaru telescope and 10.4-m GTC telescopes were initially designed to have their own AO system too. Finally, AO is the basis on which the new of 30- and 40-m telescopes, the Thirty Meter Telescope (TMT) and the European Extremely Large Telescope (E-ELT), will produce their first images in the next decade.

1.3 Photometric sources of error in adaptive optics

AO systems are able to correct the effects of atmospheric turbulence that disturbs the light wavefront, thus enabling to obtain images that are close to the theoretical telescope diffraction limit. However, since such correction is never perfect, errors in the photometric estimations usually introduce some bias into the astrophysical conclusions arrived at from the observations. Some papers have dealt with both the theoretical and the practical image quality performance of AO imaging systems (Wilson & Jenkins 1996; Tessier 1997), although without focusing on the photometric quality. In order to fill this gap, Esslinger & Edmunds (1998) provides an excellent review of sources of error and photometric performance in AO systems based on observations obtained at the ESO in La Silla (Chile). The study reported there are assumed SR variations of between 0.15 and 0.3 in both isolated objects and crowded fields.

There are two principal photometric methods. First, aperture photometry; which consists of integrating the counts of light within a circle surrounding the star being observed. The background sky can be estimated within an annulus that is concentric with that circle and then subtracted. The annulus dimensions must be chosen so as to be close enough to the aperture circle so as to provide an estimate of the same background in both, but large enough to ensure that there is no contribution from the star. Second, PSF fitting; because aperture photometry cannot be used in crowded fields, since it is not possible to guarantee the presence of single stars within every aperture circle. In this case, it is possible to fit a model of the stellar image to the data by means of a least-square algorithm, studying several stars at the same time.

Nevertheless, both photometric methods make some assumptions about the PSF, which are correct in seeing-limited observations, but are not valid in AO. Essentially, they assume that almost a perfect knowledge of the PSF is available, which is supposed to be smooth, stable over time and constant over the field of view (FOV). Photometrical sources of inaccuracy introduced by the use of AO systems can be classified as follows (Esslinger & Edmunds 1998):

1. Global PSF variations over time.
2. Fluctuations in the halo surrounding the core of the PSF.
3. The presence of residual features in the PSF.
4. Variations in the PSF due to angular anisoplanatism.
5. Biases introduced by using different deconvolution methods.
6. Deconvolution performed with a badly determined PSF.

The first obstacle to accurate photometry in AO observations is the global variation of the PSF over time. Such variations have two main causes. First, the shape of the PSF and its SR strongly depend on the seeing conditions, which vary very rapidly thus introducing fluctuations on the PSF. Second, even with constant seeing, noise coming from the WFS as well as uncorrected variations in the high-order spatial modes of the wavefront introduce temporal dependence into the PSF.

Moreover, aperture photometry usually requires the observations of a reference star for calibration purposes. In general, these stars are not in the same FOV as the object of interest; hence, the calibration star and the object must be observed at different moments and therefore variations over time affect them differently, so the PSF changes slightly from one to the other. In the case of PSF fitting, this effect is even more dramatic since the method requires perfect knowledge of the PSF. However, when studying crowded fields, it is sometimes difficult to find an isolated star which could provide an independent PSF. Again this means that the calibration star must be observed at a different time and therefore under different conditions.

The second drawback mentioned above arises during the study of faint companions or structures around a bright main star. In general, the AO PSF is composed of the sum of two terms. The first one is a peak or central core which corresponds to the diffraction-limited corrected observation. The second additional term corresponds to a surrounding halo produced by the non-corrected high-order modes of the wavefront or, in other words, the seeing-limited uncorrected part of the observation (see Figure 1.10).

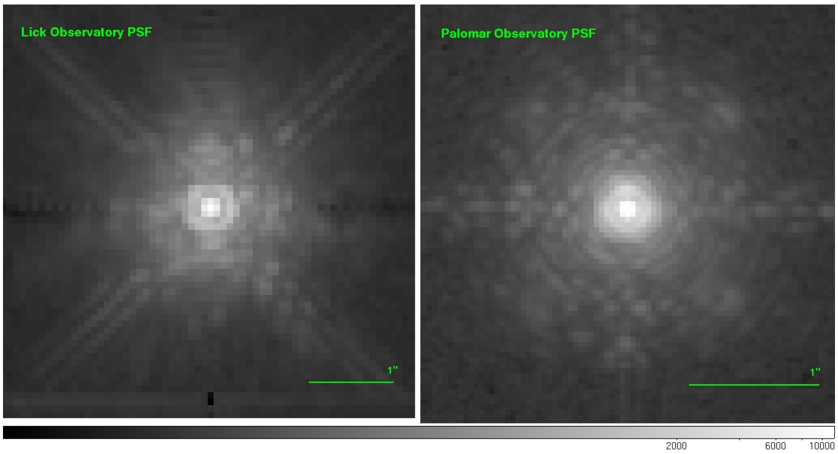

Figure 1.10: Left: PSF from the 3-m. Shane telescope at the Lick Observatory, pixel size = 0.076 arcsec, SR = 52%. Right: PSF from the 5-m Hale telescope at the Palomar Observatory, pixel size = 0.040 arcsec, SR = 53%. Logarithmic scales, standard colors.

The presence of the halo is a major problem in itself for the detection of faint companions around stars. Bearing in mind that the halo is dominated by the atmospheric turbulence, it varies greatly with time. This means that faint structures of the PSF are different from frame to frame; a tendency which is exaggerated by the presence of noise. Furthermore, even if the halo was stable and constant, it would still depend on the object brightness, its shape, its spectrum or its position in the sky; so it would be different from one object to another in the same or a different FOV. Using longer integration times and choosing the calibrator carefully (with a similar flux to that of the object, in the vicinity, only short delays between observations, etc.) can minimize such differences but never eliminate them completely. All these problems clearly affect the photometric accuracy of observations, especially if the object is so faint that a brighter one must be used as a natural guide star. The halo of such a bright star may dominate the faint object and distort the photometric measurements.

The third problem is related to the presence of residual features, speckles or holes in the halo of the PSF; e.g., spikes due the secondary mirror supports, lumps in the diffraction ring or faint artifacts due to fixed residual aberrations in the AO systems. All these features vary with the aforementioned global fluctuations and halo variations, as well as with noise. Therefore, they introduce a new source of error into detection and photometric precision, especially if a faint companion around a bright star lies directly over any such hole or speckle.

Angular anisoplanatism, or variation of the PSF with respect to position in the FOV, also introduces inaccuracies into the photometry. This is again particularly problematic for PSF fitting, although it also affects aperture photometry since the portion of light contained in an aperture circle will depend on position. In the case of PSF fitting, this problem could be overcome by dividing the FOV into different zones and choosing a calibrator for each one. Evidently, this option multiplies the problems related with correct choice of reference star.

Finally, although images obtained with AO systems can be studied directly with very good results thanks to the sharpness achieved in the PSF, it is also convenient to use different deconvolution methods to fully exploit all the information present in the data. The deconvolution methods usually provide good approximations to the true object. However, their numerical results are strongly biased by variations in the PSF, by a poor choice of reference PSF (which is especially problematic in AO since the PSF is never perfectly known) or by the deconvolution algorithm employed. For example, Christou & Bonaccini (1996) used both linear and non-linear methods, including blind deconvolution, to calculate the differential magnitude between the two components of the double star T Tauri and the results obtained varied from 1.46 to 1.85 magnitudes depending on the method used. Tessier (1997) likewise reported variations from 0.74 to 0.96 magnitudes, depending on the deconvolution algorithm used, when estimating differential magnitude of a binary system. Gladysz et al. (2010a) and Baena Gallé & Gladysz (2011) proposed a method to systematically study and compare algorithm performance when estimating the differential magnitude of binary systems. The method consists of two steps. Firstly, the accuracy is studied by measuring, on the one hand, the absolute departure of the estimates from reality and, on the other hand, the mean magnitude difference in order to check for possible over- or under-estimations. Secondly, the scatter is measured when averaging different observations of the same binary system, but with each observation affected by different underlying speckle structures introduced by different PSFs.

1.4 Imaging techniques

Image processing techniques to extract and/or improve physical information from datasets have developed over the last 50 years (Andrews & Hunt 1977; Pratt 2007). Evidently, AO is not aloof from such techniques and can take full advantage of them. Moreover, some of the approaches have been specifically designed for the special characteristics of AO acquisition systems and in particular to overcome poor knowledge of the PSF due to its temporal and spatial variations. Davies & Kasper (2012), in an excellent review of AO, divides the current post-processing methods into four classes: deconvolution; model convolution; astrometry and photometry estimation; and finally, speckle suppression.

1.4.1 Deconvolution

Deconvolution consists of the operation that creates an estimator of the true object (O) from the dataset or image (I). The intensity distribution in such an image can be modeled as the convolution of the true object intensity with the instrument PSF (P). In general, this operation is corrupted by additive noise (N) yielding to the well-known expression for the image formation process:

$$I = O \otimes P + N, \tag{1.40}$$

where $\otimes$ denotes the convolution operator. Therefore, deconvolution is the inverse operation which allows to solve equation 1.40 with O as the variable of interest. Unfortunately, such an equation is an ill-posed problem since, apart from the object, O, the noise, N, is evidently an unknown term too and, furthermore, it is impossible to have perfect knowledge of the instrument PSF P (especially in the case of AO, see subsection 1.3). In such conditions, solving the image formation equation becomes an intricate task which has led to multiple algorithms and approaches that aim to obtain the closest estimate to the actual true O. In general, since the PSF introduces some degree of smoothness (in astronomical visual and infrared imaging) to the object, the deconvolution of the image yields enhancement of low contrast features, in accordance with the statistical nature of the noise.

It is well known that direct inversion of equation 1.40 in the Fourier domain amplifies the noise; hence, in the presence of noise (which is the usual situation in astronomy) this simple method cannot be used. The solution can be constrained by a regularization parameter (Tikhonov et al. 1987; Pratt 2007) in a linear scheme. However, this introduces a trade-off between the resolution enhancement achieved and the undesirable noise amplification, which can yield excessively smooth solutions (Starck et al. 2002).

This fact forces the use of non-linear approaches that treat the problem by means of statistical estimation methods. The two most widely used approaches are the Maximum Likelihood Estimator (MLE) and the Maximum A Posteriori (MAP). The MLE approach consists of maximizing the likelihood function, which measures the resemblance between the image and the projection of the object estimation towards the image domain, in keeping with the statistical nature of the object. The Richardson-Lucy algorithm, which assumes a Poisson distribution for the object, is probably the most successful here (Richardson 1972; Lucy 1974). When some prior information about the object is added to the MLE scheme, such as positivity or entropy (Frieden 1972), it leads to the MAP approach.

These approaches have been linked together in a common framework by means of Bayesian methodology (Molina et al. 2001; Starck et al. 2002) and they have proved their effectiveness in practice due to the spherical aberration detected in the main mirror of the Hubble Space Telescope. Moreover, they are the typical algorithms used to deconvolve seeing-limited images from ground-based telescope. Finally, Starck et al. (2002) describes several of these linear and non-linear algorithms that can be used to perform deconvolution from a multiresolution perspective, taking advantage of multiscalar transforms (e.g. Núñez & Llacer 1993; Otazu 2001, the FMAPE and AWMLE algorithms), such as wavelets (see Appendix D).

Unfortunately, typical non-linear approaches have to deal with and control several issues, apart from noise amplification. One of them is edge gradient amplification, which poses serious problems as in the case of Gibbs oscillations. Another, which is crucial in AO, is the use of a static and non-variable PSF when this does not strictly correspond to the reality. This has moved the AO community to develop and prefer myopic and blind deconvolution approaches, which do not assume the PSF as an “absolute truth” and estimate the PSF as well.

Therefore, the MISTRAL algorithm (Mugnier et al. 2004) is a myopic restoration tool within the

Bayesian framework. As an MAP approach, it maximizes a functional with three terms: one related to the likelihood or the fidelity to the data; a second which is prior information about the object, in particular, it introduces explicit terms to preserve the edge properties; and a third that is related to the fidelity of the PSF solution to an initial estimate. The strength of MISTRAL lies in that it does not assume perfect knowledge of the PSF, but it simultaneously calculates both the object and the PSF while applying soft constraints to its properties. The AIDA algorithm (Hom et al. 2007), in contrast, is an extension of the previous algorithm that automates the calculation of some parameters (which in the case of MISTRAL have to be chosen manually by the user).

Finally, the IDAC algorithm (Jefferies & Christou 1993) is a practical implementation of the concept of multi-frame blind deconvolution, which is a method of iterative blind deconvolution of multi-frame data based on the minimization of a penalty function while setting some logical constraints on both the object and the PSF. The penalty function is formed of four terms which account for: positivity, fidelity to the data in the Fourier space, prevention of wrong convergence to a δ-function or to the dataset itself, and a prior estimate of the object's Fourier modulus. The fact that IDAC does not make assumptions about the PSF (although it can be provided with an initial estimation) probably makes it less robust than MISTRAL but, on the other hand, it is applicable when no PSF is available. Pantin et al. (2007) gives a full description of both the MISTRAL and IDAC algorithms.

Baena Gallé et al. (2012) compares the performance of AWMLE, MISTRAL and IDAC, as well as a new MLE multiresolution algorithm based on the curvelet transform (Candès et al. 2006), when deconvolving images of Saturn and galaxy M100, observed with the 5-m Hale telescope at the Palomar Observatory. The results show that a static-PSF algorithm, with the help of multiresolution support, is better than the blind/myopic approaches for mismatched calibrators within a range of 8% in SR and low-level noise conditions. At higher noise levels, the performance of IDAC decreases dramatically while AWMLE still offers 10-15% better results than MISTRAL. Objective comparisons were performed by the typical mean squared error (MSE).

1.4.2 Model convolution

Model convolution is a variant of deconvolution. If the properties of the object can reasonably be described by a set of parameters in a kind of model, M, (e.g., a binary system can be well described with only three parameters: differential magnitude, angular separation and orientation), then that model can be convolved with the PSF in order to obtain an image estimation $\tilde{I} = P \otimes M$ which can be compared with the actual dataset I. Iterative improvements to M, leading to a solution, $\tilde{I}$, that is closer to I, will let the assumption that $M \sim O$. Since, in typical applications, only a few parameters are needed to suitably construct the model, the solution is highly constrained and does not lead to noise amplification, unavoidable in the deconvolution schemes.

This methodology allows to accurately estimate the uncertainties in the parameters since it is very easy to measure the impact of changing them. In contrast, a major weakness of this method is its simplicity. First, the set of parameters must fit the physical properties of the real object well, i.e., complex structures in the object cannot be well described by the model and “disappear” in the reconstruction. Second, the method will be restricted to a set of model families which meet the set of parameters.

Model convolution is frequently applied to deep images of distant galaxies which are commonly parametrized by the Sérsic function and have a poor signal-to-noise ratio. In order to overcome the aforementioned limitations, Peng et al. (2010) included perturbations in the form of Fourier modes in the azimuthal shapes in galaxy profile fitting code GALFIT3.

Another typical application is in kinematics, where deconvolution cannot be used because of the interdependency imposed by the PSF between the observed luminosity, velocity and dispersion. The only chance for the latter would be to deconvolve each spectral plane of the original datacube bearing in mind that the PSF varies from one plane to the next. Such deconvolution can be guided by a high-resolution image (Ferruit et al. 2000) and can be applied to extract the spectrum of a supernova superimposed on a galaxy (Bongard et al. 2011).

1.4.3 Astrometry and photometry estimations of point-like sources

Images of binary stars or cluster fields only require the exact position and flux of each star, i.e., an accurate astrometry and photometry measurement. In crowded fields, this requires a good estimate of the PSF with which to fit the sources (see Section 1.3). Moreover, it is also important to obtain the false detection level beyond which it is not possible to distinguish real sources from artifacts or speckles created by non-corrected high-order modes in the AO system.

In order to create maps of positions and fluxes from stellar fields, two basic codes have been proposed. Firstly, DAOPHOT (Stetson 1987), in which the reference PSF is based on analytical fits to several stars guided with a look-up table for small empirical corrections. This code is able to handle undersampled data as well as to accommodate some degree of spatial variation. Secondly, there is StarFinder (Diolaiti et al. 2000), which was designed specifically to work on AO data with complex PSF shapes. The algorithm derives a PSF template from the brightest isolated stars. Subsequently, a catalog of suspected stars is created by searching for the relative maxima in the frame. In the following step, the images of the suspected stars are analyzed in order of decreasing luminosity, and are accepted as real on the basis of the correlation coefficient with the PSF template. The relative astrometry and photometry of the source are determined by means of a fit, taking into account the contribution of the local non-uniform background and of the stars already detected. This 3-step process is iterated until no sources can be reliably found in the residuals. PSF-fitting, as it is applied in the the StarFinder package, could be seen as an extension of the CLEAN algorithm (Högbom 1974) that has been successfully used by the radioastronomy community for decades.

Although those approaches do not account for anisoplanatism, local PSFs can be extracted for each isoplanatic region and applied independently. Furthermore, even though the PSF core can be very narrow, high photometric precision requires that the PSF is well determined up to 1 arcsec; so the faint wings must also be properly characterized. For this reason, it may be convenient to apply some deconvolution scheme in addition, but while taking care of the noise amplification, in order to make the central core even narrower and brighter as well as to reduce the strength of the wings

(Schödel 2010).

Baena Gallé & Gladysz (2011) compare the performance of StarFinder with a wavelet deconvolution MLE algorithm, in particular AWMLE (Otazu 2001), to determine the differential photometry of binary systems when the companion is contaminated by the primary's halo. This comparison is performed for different angular separations with high and moderate levels of correction in terms of SR. The authors conclude that AWMLE yields better results even when the calibrator is not well matched to the PSF. That study is generalized in Gladysz et al. (2010a), where several approaches are compared (including StarFinder, AWMLE, blind deconvolution algorithms and temporal speckle suppression techniques). In such a study, a static-PSF algorithm is shown to provide results that are comparable to those of the usual approaches in AO imaging.

1.4.4 Speckle suppression

Different speckle suppression techniques have been developed. In general, they consists of properly characterizing the PSF in order to remove complex speckle patterns that are present on the wings. For such purposes, the methods usually impose new instrumental and observational requirements which are mainly focused on the detection and characterization of faint companions, e.g., exoplanets. They can be classified as follows:

1. Spatial: angular differential imaging was first proposed by Marois et al. (2006). It is based on the fact that a major part of the quasi-static speckles are produced in the telescope and its instrumentation. Therefore, if the FOV is able to rotate with respect to the telescope configuration, it is possible to obtain datasets with real sources in the same relative positions while the speckles move around the center of the view. Through appropriate analysis of this sequence of images, one can suppress artifacts and simultaneously enhance real sources (Lafrenière et al. 2007).

2. Temporal: Labeyrie (1995) proposed a method which assumes that changes in the speckle pattern will never reach their minimum level in locations where there is emission from a real source. Unfortunately, while this technique enables the detection of faint companions, it does not allow to perform photometry. Gladysz et al. (2010b) developed a statistical method to distinguish speckles from faint companions which also allows to estimate the intensity of the real source. The method is based on the fact that both the main star and the companion share common parameters in the probability density functions (PDF) of their respective observed peaks of the AO-corrected PSF. However, the companion's PDF is also blurred with the PDF of the quasi-static speckles, and thus shows important morphological differences with respect to the PDF of the primary object. So, a 2-step algorithm was developed: first, least-square fitting of the theoretical PDF of the bright star allows to obtain some parameters which are common to the companion; then these parameters are applied at the location of the companion and deblend its PDF from that produced by the speckles. Differential photometry can then be produced by converting the widths of the two distributions into a brightness ratio of the two objects. This

method has, however, a specific range of application to the photometric estimation of companions that lie within the uncorrected halo of the AO PSF. Furthermore, the location of the companion must also be known.

3. Chromatic: simultaneous differential imaging (Racine et al. 1999) makes use of the different spectral energy distributions of the star and its companion. One can then define different narrow bands that are close enough to each other in wavelength so that, in a different image, the speckle will be subtracted while the companion will not. Another method is the so-called spectral deconvolution (Sparks & Ford 2002). This takes advantage of the fact that the radial separation of the speckle is wavelength dependent while the position of the faint companion is fixed. So, a datacube where every single image has been acquired at a different wavelength can be rescaled, thus aligning the speckles but making the planet move to the center of the view with increasing wavelength. Pixels through the datacube at the same location are called spaxels. When a planet crosses a spaxels on its journey to the center, it creates a bump; one can then try to fit a low-order polynomial to the signal present in the spaxels and subtract it. Since the planet signal will not adequately fit the polynomial, it will remain while speckles will not. Baena Gallé et al. (2013) proposes the use of 1-D wavelets for analyzing spectral datacubes in the direct imaging of exoplanets. They simulate different ELT scenarios, for EPICS instrument (Kasper et al. 2008, 2010), yielding improvements of up to $\sim 1 mag$ at a distance from $\sim 70\, mas$ with respect to the classical polynomial fitting approach.

4. Polarimetric: this is based on the fact that at small angular radii, speckles will remain unpolarized. Hence, if two images at orthogonal polarizations are taken simultaneously, the speckles will be subtracted out, while only the polarized light will be left. This allows to detect scattered light from a circunstellar disk or even the atmospheres of exoplanets. This technique has proved its usefulness on the T Tauri star TW Hya by Apai et al. (2004), who show the presence of polarized disk emission between 0.1 arcsec and 1.4 arcsec from the star. The authors also derived the first Ks-band radial polarized intensity distribution.

5. Deep Suppression: it can be proved that a sinusoidal phase across the aperture produces two symmetric speckles of light with respect to the center of the FOV. The angular separation of these speckles from the center increases with the spatial frequency of the ripple. Since the DM has a spatial frequency cut-off due to its finite number of actuators, it can only produce or correct speckles up to a maximum angular separation; this is called the correction radius. In order to suppress a coherent patch of light created by phase and amplitude aberrations inside the correction radius, the DM just needs to create an anti-speckle of the same amplitude but with a π phase-shift at the same location. Unfortunately, there is a discrepancy of $\pi/2$ between the phase of the symmetric speckles created by the DM and those created by an amplitude sinusoid. Therefore, only one of the amplitude speckles can be removed by an aperture plane DM, while the other would be amplified. However, it is possible to correct all the speckles in the image over a certain wavelength range using a second DM. Different implementations have

been proposed theoretically (Give'on et al. 2007; Guyon et al. 2010). The practical application of this technique has been demonstrated in the laboratory by Trauger & Traub (2007).

1.5 Current science applications

In this final section, some examples are summarized in order to show how high-resolution images provided by AO systems have challenged astronomers to go into greater depth in almost all the possible lines of research in astronomy and astrophysics, from the solar system, to galaxies; including stellar fields and exoplanets, among others. Furthermore, AO imaging has allowed astronomers to overcome some of the limits of their knowledge and push them forward towards new goals and challenges. This section is based on the manuscript outlined by Davies & Kasper (2012), who in turn based their work on what is by now the vast available AO literature.

1.5.1 The solar system

- The Sun: daylight AO observations face additional requirements to those necessary for night-time observations. In general, the corrected wavelength is in the visible part of the spectrum and often at high airmass, typically in the G-band ($430nm$), thus angular resolution is lost due to the observing frequency. There are no visible stars to calibrate corrections, so the necessary spots to close the feedloop must be created by cross-correlation techniques (Rimmele 2000). Nevertheless, the enhancement achieved in solar imaging on current 1-m solar telescopes has led to the birth of a new generation of 4-m designs. Large structures of the sun are linked to small scale dynamics, such as the pressure scale height and the photon mean free path, which are both of the order of about 0.1 arcsec. The first results reported by Goode et al. (2010) achieved a resolution of up to 0.12 arcsec with the 1.6-m aperture solar telescope at the Big Bear Solar Observatory. They showed that the smallest scale photospheric magnetic field seems to come in isolated points in the dark intergranular lanes, rather than the predicted continuous sheets confined to the lanes, and the unexpected longevity of the bright points implies a deeper anchoring than predicted.

- Planets and Satellites: atmospheres of planets and satellites have been studied in depth using AO. One example is Io, which has been observed darkened by Jupiter's shadow in order to highlight its volcanic activity; while Ganymede was used as the AO calibrator (dePater et al. 2004). Titan and its thick atmosphere is another typical target (Hartung et al. 2004; Hirtzig et al. 2006). dePater et al. (2010) shows that Jupiter's small surface storms exhibit bright rings at $5\mu m$, while those with diameters above $12000km$ do not. The authors propose that air is rising along the center of a vortex, and descending around the outer periphery, producing the $5\mu m$ bright rings. These downflows cannot exist at radii greater than 1-2 times the Rossby radius, since rotational effects become as important as buoyancy effects. Baena Gallé et al. (2012) use images of Saturn to compare and show the performance of different deconvolution algorithms typical in

AO imaging postprocessing. Neptune's atmosphere and its satellites have also been monitored continuously (Fry & Sromovsky 2004; Marchis et al. 2004) as has Uranus (Sromovsky & Fry 2004).

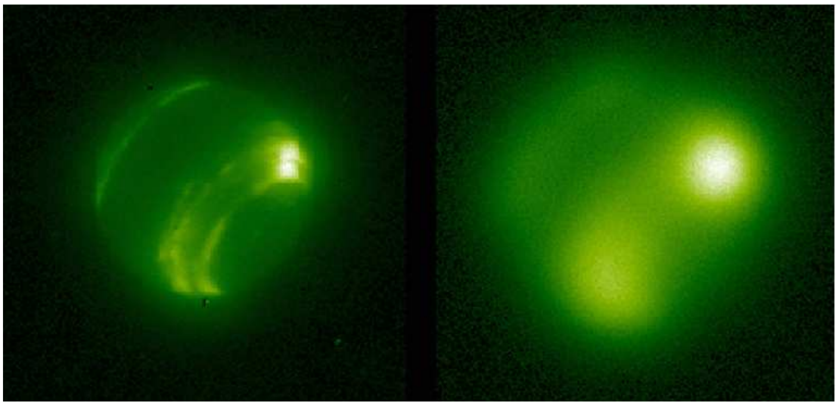

Figure 1.11: Neptune observation with 10-m Keck telescope in the methane band, with AO correction (left) and without (right).

- Asteroids: The major contribution of AO to asteroid physics has been to show how massive asteroids are extremely porous. In general, AO takes advantage of binary asteroid systems, using the moon as the calibrator, to produce high-resolution corrected images. Moreover, binary systems are an easy way to measure the mass and density of the system. Nowadays, more than 150 main belt binary asteroids are known, the first one was discovered by Merline et al. (1999), who showed a companion orbiting 45 Eugenia and revealed that the density of such an asteroid was only ~ 20% greater than water. Marchis et al. (2005) discovered that 87 Silvia has indeed 2 moons. Asteroids 217 Kleopatra and 617 Patroclus have been shown to be highly elongated, thus suggesting a catastrophic impact which provoked the spinning of large liquid masses giving them their characteristic shape. (Descamps & Marchis 2008) proposed this catastrophic scenario as a common rule for the formation of such asteroids.

1.5.2 Star formation

- Stellar multiplicity: AO is now a key tool in assessing the multiplicity and characteristics of stars. Close et al. (2003) and Siegler et al. (2003) carried out one of the largest surveys of low-mass stars to date. They observed 69 stars of spectral type M6.0 to L0.5 by means the Hokupa'a AO system on Gemini North. They found 12 systems with very-low-mass or brown dwarf companions yielding a fraction of binaries of ~ 10%. The pairs in each binary have similar masses and their separations are always below 15 AU. These characteristics are significantly different from the more massive G dwarfs for which the binary proportion is around 50% and they have wide separations of 30 AU. Other studies (Dupuy & Liu 2011; Konopacky et al. 2010) show a preponderance of almost circular orbits and find at best only a marginal correlation between eccentricity and period. Bate (2009) explained these discrepancies by suggesting that older field brown dwarfs are systems that have been ejected at speeds of a few $km\ s^{-1}$ from the cluster in which they were formed. For very-low-mass stars, only the most tightly bound systems survive at such a velocity, resulting in low multiplicity, small separation and low eccentricities. This picture is even more complicate if measurements obtained by Close et al. (2007) and Biller et al. (2011) are taken into account, who found that young very-low-mass and brown dwarf binaries can have wider separations of more than 100 AU. This could be explained by establishing a dependency between the binary survival time and how tightly bound they are as well as the stellar density in the local environment.

In contrast, to determine the astrophysical properties of a binary system, correct calculation of differential photometry is essential. Gladysz et al. (2010a) and Baena Gallé & Gladysz (2011) compare different approaches in the extraction of differential fluxes between the main star and its companion when the latter is severely affected by the halo introduced by the former.

- Circumstellar disks: Binary stars and planet system formations can be studied in their early years by observing the disks around young stars. AO has filled out the picture with high-resolved images in the near- and mid-infrared range which help to deduce how such disks evolve. The first circumstellar disk observed by AO was β Pic (Golimowski et al. 1993), which revealed complex sub-structures in this disk at the 1 AU scale. Roddier et al. (1996) and Duchêne et al. (2004) studied the binary T Tauri system GG tau A-B indicating that larger grains of dust tend to grow up faster in the midplane. However, other observations revealed that this process is only one of several. Systems with similar ages exhibit different dust properties thus showing different growth rates and/or settling times (Duchêne et al. 2010; McCabe et al. 2011). Besides, further studies of GG Tau reveal that the disk must be tilted with respect to the system orbit, which can only be explained by hypothesizing the existence of a companion. Indeed, AO is now providing strong evidences that central holes in the disks are due to the presence of a companion or giant planets (Kraus et al. 2012; Kraus & Ireland 2012).

- Exoplanets: direct imaging of exoplanets is extremely demanding. It requires high-contrast, long-exposure, coronagraphic imaging, together with in-depth knowledge of the speckle pattern as well as modern post-processing techniques (see subsection 1.4.4). The first exoplanet directly imaged was 2MASS 1207 b by Chauvin et al. (2005), using the infrared WFS of VLT/NACO. This is a $5M_J$ object orbiting a brown dwarf 5 times more massive, at a distance of 55 AU. Moreover, the first multi-planet system discovered by proper motion analysis was discovered orbiting around HR8799, a $1.5M_\odot$ A5V star. This system consists of 4 planets at distances of 14-68 AU and masses in the range of 7-$10M_J$, all orbiting in the same direction (Marois et al. 2010). However, the mechanisms that rule the formation of such systems are not yet well known. While the 3 outermost planets could have been formed by the fragmentation of a disk of dust and gas, this process is not possible for the innermost one because of the high temperature and fast rotational speed of the disk at that distance. A possibility, suggested for the $\sim 9M_J$

planet β Pic b, is a formation process via the agglomeration of grains. Unfortunately, although several exoplanets have been directly imaged nowadays, surveys to detect objects down to limits of a few M_J have failed (Lafrenière et al. 2007; Kasper et al. 2007); hence, the next generation of planet imagers and imaging approaches are eagerly awaited. In this sense, Baena Gallé et al. (2013) propose the use of 1-D wavelets to analyze spectral datacubes acquired by the spectral deconvolution technique, as was already mentioned in point number 3 of subsection 1.4.4.

1.5.3 Galaxy studies

- The galactic center: The center of the galaxy provides a detailed view of the physical processes occurring in the nucleus of the Milky Way and around its central black hole. This study can be directly applied to the nuclei of other galaxies. Spatially resolving the stellar population orbiting the center in order to study proper motions is of special interest (Fig. 1.12).

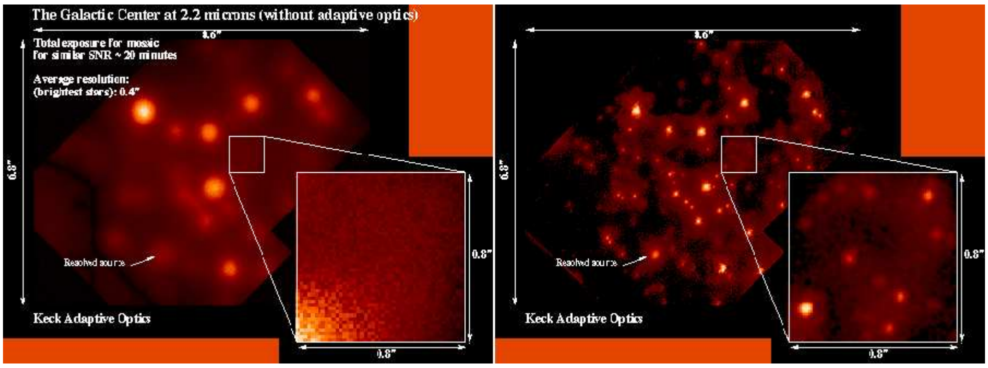

Figure 1.12: Milky Way center observed from Keck at 2.2μm., with AO correction (right) and without (left).

Analysis of stellar motion shows that most of the stars in the central parsec are old and have randomly orientated orbits (Yelda et al. 2010). However, about half of the young stars in the central 10-15 arcsec are confined to a warped clockwise disk and many of the remainder may be in a second counter-clockwise disk (Bartko et al. 2010; Lu et al. 2009). Combined datasets from Keck II and VLT are yielding precise measurements of the distance from the solar system to the galactic center and the mass of its black hole, which are now assumed to be 8.3kpc and $4.3 \times 10^6 M_\odot$ (Genzel et al. 2010).

- Stellar populations: resolving stellar populations spatially in nearby galaxies is an important goal for future 30- and 40-m telescopes in order to enlighten the star formation history. AO is the obvious solution to the evident crowding problem; however, increasing AO performance at IR-wavelengths is somehow in contradiction with the better performance of the current mapping of star color magnitude diagrams at shorter visible wavelengths. This encourages improvements in AO at such frequencies and means that current 8- and 10-m telescopes are not as competitive as HST in this field, for example. Despite this, some good results have been achieved when observing galaxy M31 (Olsen et al. 2006) and dwarf galaxies (Melbourne et al.

2010). In addition, galaxy NGC 3603, which is considered a local template for the massive star forming clusters, has also been widely observed (Eisenhauer et al. 1998; Harayama et al. 2008; Bastian el al. 2010). Furthermore, resolving the stellar populations also allows to study the proper motion of stars, thus deriving the intrinsic kinematics of clusters and galaxies, as well as their global motions. Stolte et al. (2008) have shown that the young Arches Cluster, which is only $30pc$ from the galactic center, has moved by $24.0 \pm 2.2mas$ in a 4.3-year period. Therefore, AO should be able to provide a proper motion precision of $10\mu as\, yr^{-1}$ in only three or four years: equivalent to $5km\, s^{-1}$ at $100kpc$, with ELT (Trippe et al. 2010).

- Black hole masses: Measurements of black hole masses are essential to understand their evolution and growth as well as those for their host galaxies. It is clearly known that the velocity dispersion, mass and luminosity of the stellar spheroid around a black hole is directly related with the hole's own mass (Häring & Rix 2004; Ferrarese & Ford 2005). AO is of major importance in this field since it brings together high resolution and large collecting areas, which allows the study of faint galaxies, and integral field spectroscopy at near-IR. Davies et al. (2006) proved that it is possible to measure the mass of black holes in type 1 active galactic nuclei (AGN) using spatially resolved stellar kinematics, providing a complementary method of reverberation mapping which relies on tracking the temporal variability of the broad lines. Denney et al. (2010) derives a reverberation mass of $(7.6 \pm 1.7) \times 10^6 M_\odot$ for galaxy NGC 3227, thus showing lower values than previous estimates. In contrast, pseudo-bulges are created from secular disk processes instead of merger events, which rule the formation of classical bulges, hence the former have different stellar populations with different mass distributions, kinematics and different physical relations with the black hole properties. Indeed, many local disk galaxies have at least a pseudo-bulge component to their central region. Nowak et al. (2010) used AO to calculate black hole masses in such galaxies suggesting that there is no need to separate pseudo- and classical bulge components to really understand the evolution of black holes and bulges. Finally, for elliptical galaxies, Gebhardt et al. (2011) combines AO integral field spectroscopy with wider field data to arrive at a mass of $(6.6 \pm 0.4) \times 10^9 M_\odot$ for M87, which exceeds the mass expected by twice its uncertainty. This suggests that the high mass end of the relation is poorly constrained or its scatter is larger than expected.

- Gas flows: Nearby AGNs make the study of the mechanisms that drive gas towards the central black hole possible. Since they are close to each other, AO is needed to resolve them to a scale of a few parsecs. The combination of AO with integral field spectroscopy is yielding the resolution of their spatial distribution as well as the kinematics of stars and of molecular and ionized gas. Such techniques have revealed inward flows of gas at low rates along circumnuclear spiral arms in a great number of galaxies (Davies et al. 2009; Schnorr Müller et al. 2011). However, there may be exceptions to this rule, such as galaxy NGC 1608 where the gas appears to be streaming almost directly towards the AGN (Müller Sánchez et al. 2009). Inflow gas is believed to be associated with the circumnucleus dust structures that have been mapped in many active and inactive galaxies, thus showing that gas streaming is a common phenomenon whose relation with AGN accretion is still not understood well.

- Quasars and mergers: Due to their distance and the brightness of the AGN with respect to the host galaxy, QSOs provide a challenging image scenario. The key issue is to accurately determine the PSF in order to reveal the underlying host galaxy (Guyon et al. 2006). The specific role of mergers in fueling QSOs can be assessed by studying those that exhibit double peaked [O$_{\text{III}}$] lines, which could trace pre-coalescent dual AGNs. AO imaging of such candidates reveals that only $\sim$ 35% have double nuclei on kiloparsec scales (Fu et al. 2011; Rosario et al. 2011) and their origin is not clear. It has been suggested that most QSOs might not be associated with gas-rich major mergers. One of the best known dual AGNs is NGC 6240 which has been the goal of several AO observations. These have revealed numerous massive young star clusters around the nuclei which are undetected by HST (Pollack et al. 2007). The combination of AO datasets at different IR-bands with X-ray and radio continuum data have highlighted the presence of two AGNs (Max et al. 2007).

1.5.4 The high-redshift universe

Resolving internal structures and the study of the kinematics of stars which belong to galaxies at $z \sim 1.5 - 3$ is an important aim of research for AO. These galaxies extend from approximately 1 to 2 arcsec and their light is redshifted so they are preferably observed in near-IR. However, they are usually chosen to avoid bright stars in the FOV and their low surface brightness is often limited to 0.1 – 0.2 arcsecs. For these reasons, the number of targets for AO is severely reduced. For example, Förster Schreiber et al. (2009) applied AO on only 12 targets out of 63 which constituted the primary survey. Other authors have seen their AO observations reduced to numbers of around 10 (Law et al. 2009; Wright et al. 2009). Unfortunately, even when using an LGS, the rate of suitable high-red-shift galaxies that can be observed is only about 10%.

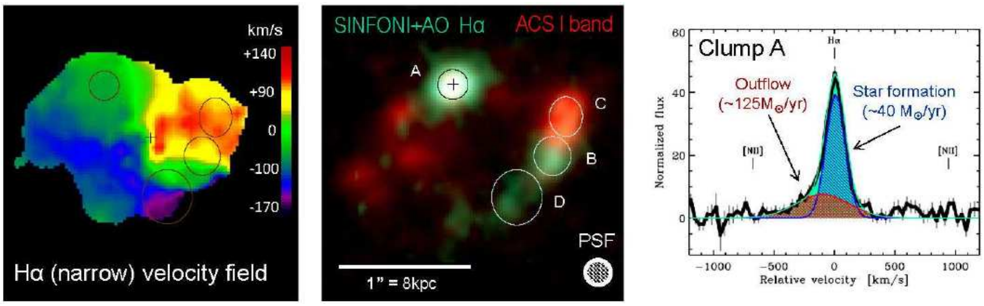

Figure 1.13: Data from $z \sim 2$ galaxy ZC406690. Left: Velocity field indicates disk rotation. Middle: Individual star-forming clumps. Right: The blue wing traces the star formation, the red profile traces an outflow, the green profile is the sum of the previous two and is a fit of the observed spectrum (black profile) —after Davies & Kasper (2012) and Genzel et al. (2011).

Despite these restrictions, AO has proved that about 1/3 of such galaxies are in fact disks (Förster Schreiber et al. 2009). As shown in Figure 1.13 for one of these objects, they are rapidly forming stars, often in giant star-forming complexes or clumps having intrinsic velocity dispersions of 20-100$km\ s^{-1}$. Such dispersions are now suggested to be common in normal massive high-red-shift disks, and are connected to the high gas accretion rates through cold flows at early cosmic times, high gas fractions and global instability to star formation. AO is fundamental because it is able to separate the 0.1-0.2 arcsec-sized clumps from the inter-clump regions; which are blurred together in seeing-limited observations.

Using AO with integral field spectroscopy is revealing fine details of individual clump properties highlighting that they can drive strong winds, with outflow rates that may even exceed their star formation rates (Genzel et al. 2011). This has important consequences for the clump lifetimes and, consequently, for galaxy formation. The use of AO and integral field spectroscopy has already led to important advances in our understanding of galaxy evolution, despite the difficulties of finding suitable targets.

Chapter 2

Estimation of Differential Photometry in adaptive optics observations with a Wavelet-based Maximum Likelihood Estimator

Published in Publications of the Astronomical Society of the Pacific (PASP). Year 2011, volume 123, issue 905, pages 865-878.

Abstract

We propose to use the Bayesian framework and the wavelet transform (WT) to estimate differential photometry in binary systems imaged with adaptive optics (AO). We challenge the notion that Richardson-Lucy-type algorithms are not suitable to AO observations because of the mismatch between the target's and reference star's point spread functions. Using real data obtained with the Lick Observatory AO system on the 3*m* Shane telescope, we first obtain a deconvolved image by means of the Adaptive Wavelets Maximum Likelihood Estimator (AWMLE) approach. The algorithm reconstructs an image which maximizes the compound Poisson and Gaussian likelihood of the data. It also performs wavelet decomposition which helps distinguish signal from noise, and therefore it aides the stopping rule. We test photometric precision of that approach vs. PSF-fitting with the StarFinder package for companions located within the halo created by the bright star. Simultaneously, we test the susceptibility of both approaches to error in the reference PSF, as quantified by the difference in the Strehl ratio between the science and calibration PSFs. We show that AWMLE is capable of producing better results than PSF-fitting. More importantly, we have developed a methodology for testing photometric codes for AO observations.

2.1 Introduction

Atmospheric turbulence imposes a limit on angular resolution which could be reached by ground-based telescopes. The arrival of astronomical adaptive optics (AO) in the early 1990s heralded a new era in which observers do not have to be content with seeing-limited observations from the ground. One of the uses of AO is high-resolution imaging of closely-spaced objects, e.g. binary stars or faint companions such as exoplanets. Determining differential photometry and astrometry between the two components of the system is essential for deducing the physical properties of the components such as mass or internal structure (Vigan et al. 2010). The use of AO guided by a suitably bright star brings about a huge improvement, both for detection of objects and for photometry. On the other hand, AO introduces problems which are not usually encountered in normal photometry as mentioned by Esslinger & Edmunds (1998):

1. The structure of the PSF has temporal variation due to seeing variability. These morphological changes are difficult to model.
2. The AO long-exposure PSF in the medium- and high-correction regime shows long-lived quasi-static speckles. These diffraction-limited “lumps” are due to residual aberrations not sensed by AO (for example non-common-path errors) and lie in the halo surrounding the core of the PSF.
3. The AO corrected PSF, and the associated angular resolution on the sky, depend on the position of the science object relative to the AO guide star. AO compensates for the turbulence in the direction of the guide star and when the science object is well separated from the AO line of sight, the compensation suffers due to a different atmospheric volume. This effect is called angular anisoplanatism, and its magnitude increases as AO compensation increases.

Because of these factors, extracting quantitative information from AO images is challenging. AO improves the detectability of faint companions over seeing-limited observations for a given telescope. When the companion is well separated, with a non-overlapping PSF, aperture photometry takes care of all the problems mentioned above. The problems occur when the PSFs from each target overlap.

One of the first mentions of photometric ambiguity after AO is in the paper by Roggemann et al. (1992). The authors analyze an AO image of bright binary star HR 6378. Previously reported parameters of this system are: separation $0.52''$ (corresponding to $5.8\lambda/D$ where $\lambda = 700$nm, $D = 1.6$m) and brightness ratio equal to 1.45. The image was deconvolved with the use of a pseudo-Wiener filter and a modified inverse filter. Results showed a big difference depending on the filter used: estimated differential photometry values were 1.66 and 2.05 for this relatively “easy” double star.

Esslinger & Edmunds (1998) provide an excellent introduction to the problem of AO photometry. One of the issues discussed is the precision of photometry on deconvolved images compared with the PSF-fitting on “raw” AO data. Two of the most-widely used deconvolution algorithms were tested, namely maximum entropy (Frieden 1972; Jaynes 1982) and Richardson-Lucy (Richardson

1972; Lucy 1974). For PSF-fitting the DAOPHOT package (Stetson 1987) was employed. Extensive testing on simulated faint companions showed that DAOPHOT performed consistently better than the deconvolution methods, i.e. its photometric precision was higher compared to precision of aperture photometry on the deconvolved images.

Historically, it has been conjectured that AO observations should be processed with "myopic" deconvolution methods, instead of algorithms assuming perfectly known PSF (Pantin et al. 2007). In the myopic framework the PSF could be only partially known, or completely unknown (blind deconvolution). Usually, these algorithms have to be provided with some PSF estimate which is assumed to be close to the truth and then they are allowed to iterate until a solution for both the object and the PSF is found. Myopic deconvolution often uses regularization, for example by imposing object priors which guarantee smoothness of the solution while preserving edges (Mugnier et al. 2004). Jefferies & Christou (1993) developed an iterative blind deconvolution method which is guided by the minimization of a penalty functional. One of the constraints they use, which is very relevant for AO observations, is a mask penalizing frequency content of the PSF beyond the telescope's diffraction-limit. The package, called IDAC, does not produce photometry and astrometry directly but these can be obtained by using aperture photometry on the restored image with a very small aperture, as suggested by Esslinger & Edmunds (1998). In our work we follow this approach to photometry after deconvolution.

When the observed (crowded) field is known to represent only point sources one can always resort to traditional PSF-fitting. In this approach an analytic or empirical PSF is used together with a fitting algorithm to match scaled-and-shifted copies of the PSF to the data. The StarFinder package (Diolaiti et al. 2000) makes use of this concept. StarFinder was developed for the specific purpose of measuring relative photometry in AO-corrected stellar fields. The algorithm operates as follows: firstly, it derives a PSF template from the brightest isolated stars, then a catalog of suspected stars is created by searching for the relative maxima in the frame. In the following step the images of the suspected stars are analyzed in order of decreasing luminosity, each suspected object is accepted on the basis of its correlation coefficient with the PSF template and the relative astrometry and photometry of the source are determined by means of a fit taking into account the contribution of the local non-uniform background and of the already detected stars. The process of PSF-update, source detection, and PSF-fitting is iterated until no sources can be reliably found in the residuals. PSF-fitting, as it is applied in the StarFinder package, could be seen as an extension of the CLEAN algorithm (Högbom 1974) widely used by the radioastronomy community. The photometric and astrometric precision of StarFinder was compared to results from blind deconvolution with the IDAC algorithm by Christou et al. (2004). According to the authors, StarFinder and IDAC yield similar photometric results but with a greater dispersion in the IDAC results for fainter stars.

In the case of binary stars, the reference PSF is often obtained by slewing from the target binary to a known single star. Great care has to be taken when choosing this reference star: its flux should match the flux of the primary in the binary system in both channels: the imaging path and the wavefront-sensing path. Only then one can guarantee that the image-plane fluxes are comparable, and also that the level of AO compensation, which is related to the flux received by the wavefront-sensor, is kept relatively constant. Additionally, zenith angle of the reference star should be close to the zenith angle of the science observations. Otherwise, light from the reference star will traverse shorter or longer path through the turbulence. Changing azimuth of the telescope might also lead to changes in the open-loop seeing due to the differential wind-flow effects. These, and other secondary effects, make the choice of a suitable reference PSF very difficult.

A recent review of modern approaches to AO photometry by Burke et al. (2009) demonstrated that methods utilizing static, deterministic PSF should not be dismissed against more modern algorithms like iterative blind deconvolution (ten Brummelaar et al. 2000). Therefore, we set out to test whether a modern implementation of the Richardson-Lucy-type deconvolution conserves the flux when given a mismatched AO PSF. The algorithm we tested, AWMLE, joins the Bayesian framework with the wavelet transform (WT) and a probabilistic window to distinguish signal from noise. In this paper, we focus on systems with relatively large brightness ratios and small separations. The faint companions we simulate are particularly difficult to analyze because of the speckle noise surrounding them as can be appreciated in Figure 2.4. To our knowledge there have been no rigorous efforts to test deconvolution algorithms on AO data given the ground truth image, as is routinely done in the image-processing literature. Therefore the second, albeit not less important, aspect of our work was to develop a sound methodology for testing deconvolution codes, in this case for photometric precision on a simple object.

The paper is organized as follows. In section 2 the AWMLE algorithm is presented. In section 3 we reflect upon relevance of AWMLE for AO observations. Section 4 is intended to describe the observations and the data reduction process. We present comparison of photometric accuracy of AWMLE vs. StarFinder in section 5. Conclusions are presented in section 6.

2.2 Adaptive wavelet maximum likelihood estimator (AWMLE) algorithm

AWMLE is fully described by Otazu (2001) and has been applied to survey type data to increase limiting magnitude and resolution by Fors (2006). The following ideas define the backbone of the algorithm:

1. AWMLE performs a multichannel deconvolution where every channel corresponds to a wavelet plane. Therefore, AWMLE operates over the wavelet planes and not over the original image.

2. Signal features in an image spread its frequency content across the wavelet planes. This property can be appreciated in Figure 2.1 where a typical AO PSF has been decomposed in its wavelet coefficients. As a result, well-defined significance thresholds can be applied to each location in each wavelet plane to selectively deconvolve statistically similar regions. This concept is called multiresolution support or probability masks (Starck et al. 2002).

3. Additionally, Poissonian and Gaussian noise contributions are mostly concentrated in the highest-frequency planes (Starck et al. 1998). This allows to selectively deconvolve each plane as the deconvolution process is guided by the global signal-to-noise ratio (SNR) of each wavelet plane.

AWMLE maximizes the likelihood between the dataset and the projection of a possible solution through the optical system, considering a combination of the intrinsic Poissonian noise of the signal and the read-out Gaussian noise of the detector.

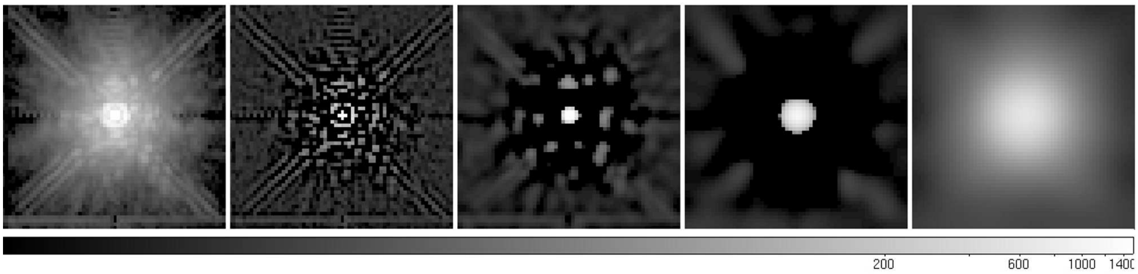

Figure 2.1: Wavelet decomposition of the AO PSF obtained with the 3m Shane telescope. Leftmost panel: original shift-and-add image of a single star (HD 143209), displayed on logarithmic plane, artificially saturated (99.5%). Successive panels show images corresponding to wavelet planes with decreasing frequencies. The last panel shows the lowest-frequency wavelet plane (the wavelet residual).

We use the notation of Núñez & Llacer (1993): **p** is the dataset, **a** is the object to be estimated, **f** is the point spread function, **b** is the background, **n** is the read-out noise, **C** is the sensor sensitivity or flatfield and **h** is the so-called direct projection, i.e. image of the object projected to the data domain by means of the PSF. The process of shift-invariant image formation is described by:

$$\mathbf{h} = \mathbf{f} * \mathbf{a} + \mathbf{b} \tag{2.1}$$

Where we also accounted for background noise **b**. Both **h** and **b** are two-dimensional images which will become multi-dimensional Poisson variables because of the light detection process. Equation (2.4) shows how **h** translates to a Poisson random variable.

The detector also introduces read-out Gaussian noise with mean zero and known standard deviation. Therefore, the general equation that describes the whole process is the following:

$$\mathbf{p} = \mathbf{f} * \mathbf{a} + \mathbf{b} + \mathbf{n} \tag{2.2}$$

where **n** is a multi-dimensional Gaussian random variable and PSF **f** already accounts for the flatfield correction **C**. This equation can be shown in the discrete form as follows:

$$p_j = \sum_{i=1}^{B} \frac{f_{ji}}{C_j} \cdot a_i + b_i + n_i \quad ; \quad j = 1, ..., D \tag{2.3}$$

Here D is the number of pixels in an image and B is the dimension of the object domain. The usual assumption is that every pixel in the image domain corresponds to the "same" pixel in the object domain, so $B = D$. Most of the optical systems in astronomy can be described by equation (2.3), in particular, the charge-coupled devices or CCDs. The presence of noise means that direct inversion of equation (2.2) leads to unacceptable amplification of that noise. Hence, non-linear methods must be used, for example, the Bayesian approach (Molina et al. 2001).

We will now formalize the description of the light detection process accounting for the two statistical distributions governing the observed random variables. First, the realization of a Poisson random variable with mean value h_j is detected at pixel j. The probability of obtaining a certain value k_j in this process is given by the Poisson probability law:

$$P(k_j|h_j) = e^{-h_j} \frac{h_j^{k_j}}{k_j!} \tag{2.4}$$

where $P(k_j|h_j)$ is the conditional probability of obtaining k_j given h_j. Strictly speaking, equation (2.4) is true only if h_j is itself deterministic. Otherwise, the process is a "doubly-stochastic" random process (Goodman 1985) and equation (2.4) must include the Mandel-Poisson transform (Mandel 1959). In practice, going to longer integrations dispenses with the problem as the speckles in the short exposures average out in the long exposures, and so h_j is almost deterministic because turbulent speckle average out. In this paper we work with integrations which are 10min long.

Secondly, the value k_j is read by the detector which introduces Gaussian read-out noise of zero mean and standard deviation σ. Hence, the probability of obtaining a value p_j from k_j at pixel j would be:

$$P(p_j|k_j) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{\frac{-(k_j-p_j)^2}{2\sigma^2}} \tag{2.5}$$

Combining both stages (2.4) and (2.5):

$$P(p_j|h_j) = \sum_{k_j=0}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{\frac{-(k_j-p_j)^2}{2\sigma^2}} \cdot e^{-h_j} \frac{h_j^{k_j}}{k_j!} \tag{2.6}$$

In words, if the projection h_j is the mean value of intensity in pixel j one would get a value p_j with a probability given by equation (2.6). Considering all the pixels of the detector we get the likelihood:

$$L = P(\mathbf{p}|\mathbf{h}) = \prod_{j=1}^{D} \sum_{k_j=0}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{\frac{-(k_j-p_j)^2}{2\sigma^2}} \cdot e^{-h_j} \frac{h_j^{k_j}}{k_j!} \tag{2.7}$$

where $\prod_{j=1}^{D}$ is the product over all the detector pixels. In the reconstruction process one wants to find the object $\mathbf{a}$, which is equivalent to finding the projection $\mathbf{h}$ (linked to $\mathbf{a}$ through equation (2.1)), that most likely produced the noisy data $\mathbf{p}$. Equation (2.7) is usually expressed in logarithms:

$$logL = \sum_{j=1}^{D} \left[-log(\sigma\sqrt{2\pi}) - h_j + log \sum_{k_j=0}^{\infty} e^{\frac{-(k_j-p_j)^2}{2\sigma^2}} \frac{h_j^{k_j}}{k_j!} \right] \tag{2.8}$$

In order to maximize equation (2.8) we take the derivative of it with respect to a_i (the value of the object $\mathbf{a}$ for pixel i) and set it to zero. Then, using Picard iteration (also known as the method of successive approximations) we obtain the following expression:

$$a_i^{(n+1)} = K a_i^{(n)} \left[\frac{1}{q_i} \sum_{j=1}^{D} \frac{f_{ji} p'_j}{\sum_{l=1}^{B} f_{jl} a_l^{(n)} + C_j b_j} \right]^m \tag{2.9}$$

where:

$$p'_j = \frac{\sum_{k_j=0}^{\infty} k_j e^{\frac{-(k_j-p_j)^2}{2\sigma^2}} \left[\frac{(h_j)^{k_j}}{k_j!} \right]}{\sum_{k_j=0}^{\infty} e^{\frac{-(k_j-p_j)^2}{2\sigma^2}} \left[\frac{(h_j)^{k_j}}{k_j!} \right]} \tag{2.10}$$

The parameter $K = \sum_{i=0}^{D} \frac{p_j - a_j}{D}$ is a constant to preserve the energy, $q_i = \sum_j \frac{f_{ji}}{C_j}$ is the total detection probability for an emission at pixel i, and m is an acceleration parameter. Equation (2.9) was first introduced for reconstructions of Poisson data in the presence of readout noise by Llacer & Núñez (1990) and Núñez & Llacer (1993).

AWMLE adds to equation (2.9) two novelties: the decomposition of the dataset **p** by means of WT and a probabilistic mask in order to adapt the level of reconstruction of each wavelet plane to the automatically-inferred presence or absence of noise. By "level of reconstruction" we mean that the effective number of reconstruction iterations within one wavelet plane does not have to be equal to the number of iterations in another wavelet plane.

Extensive literature exists about WT and its applications (Chui 1992; Daubechies 1992; Meyer 1989). WT can be seen as an evolution of the Fourier transform. It gives not only the frequency information of the signal, but also its spatial location or the place in the field of view where a specific frequency is predominant. A *wavelet* is a wave-like function, i.e. it is an oscillation that decreases in intensity. Figure 2.2 shows two harmonic functions vs. two wavelet functions with different periods and positions.

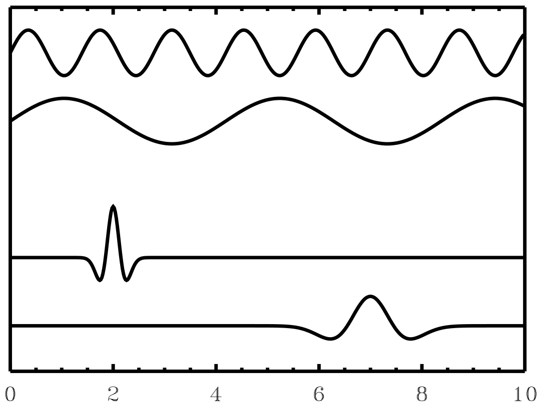

Figure 2.2: Harmonic functions (two top rows) vs. wave-like functions (two bottom rows).

This characteristic of WT leads to some advantages in the treatment of information. For instance, the noise will mainly appear in the high-frequency wavelet plane while broad shapes will appear in the low-frequency planes. This allows one to work independently with each plane, adopting different strategies to the scales that are dominated by noise than to those that most likely contain signal.

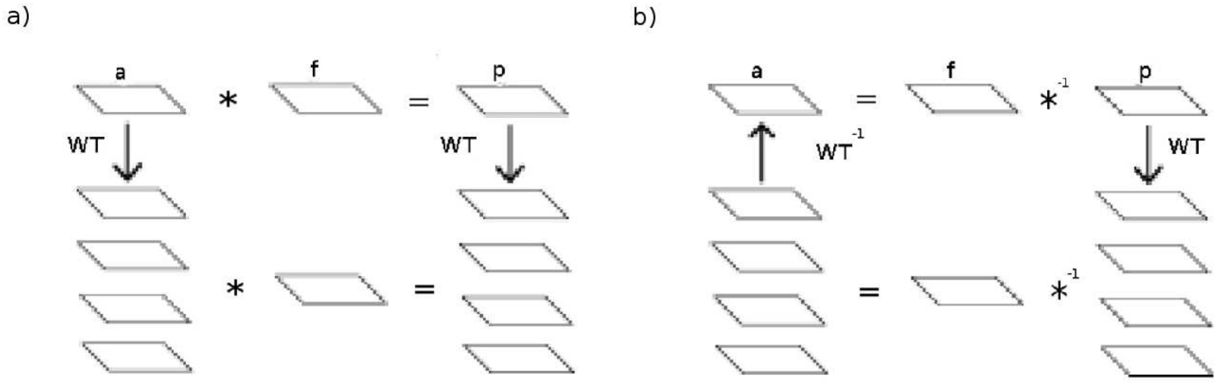

Figure 2.3: a: Relationship between object **a**, data **p** and their respective wavelet planes. b: Equivalence between reconstructing each wavelet plane independently and reconstructing the dataset **p**. WT denotes the wavelet transform and $*$is the convolution operation

.

There are several algorithms to calculate WT of an image. AWMLE makes use of the *à trous* algorithm (Shensa 1992; Starck et al. 1994) which is considered one of the most powerful for implementing WT (Starck et al. 1995). It creates *N* planes, each of them with the same number of coefficients as the number of pixels in the original image, so the representation is redundant. Obviously, the *à trous* algorithm is not suitable for compressing images.

By means of the *à trous* algorithm, the dataset **p** can be decomposed into several wavelet planes:

$$\mathbf{p} = \omega_{\mathbf{0}}^{\mathbf{p}} + \omega_{\mathbf{1}}^{\mathbf{p}} + \omega_{\mathbf{2}}^{\mathbf{p}} + ... + \omega_{\mathbf{N}}^{\mathbf{p}} + \mathbf{c}_{\mathbf{N}}^{\mathbf{p}} \tag{2.11}$$

where $\omega_{\mathbf{i}}^{\mathbf{p}}$ is the i'th wavelet plane and $\mathbf{c}_{\mathbf{N}}^{\mathbf{p}}$ is the residual wavelet plane where all the energy from the original image is concentrated, i.e. the sum of all the pixels in each $\omega_{\mathbf{i}}^{\mathbf{p}}$ is zero while the sum of the pixels in $\mathbf{c}_{\mathbf{N}}^{\mathbf{p}}$ is equal to the sum of the pixels in **p**. This is due to the characteristics of the function that is used to calculate the wavelet coefficients (the *mother-wavelet* function) which is designed to have zero mean.

It is guaranteed in Otazu (2001) that if we decompose an object **a** in its wavelet coefficients, convolve each plane with the PSF **f** and add all of them up, the result will be the dataset **p**. This is illustrated in Figure 2.3 a). Additionally, one can obtain an estimate of the object in two equivalent ways: by reconstructing the dataset **p** directly, or by reconstructing each wavelet plane independently and adding them up as shown in Figure 2.3 b). As we will see, reconstruction of the wavelet planes has the advantage that one can supervise the reconstruction process within one wavelet plane depending on its noise content, as shown by equation (2.14).

Equation (2.9) can be re-written in the following way in order to include WT:

$$a_i^{(n+1)} = K a_i^{(n)} \left[\frac{1}{q_i} \sum_{j=1}^{D} \frac{f_{ji} \sum_t \omega_{t,j}^{p'}}{\sum_{l=1}^{B} f_{jl} a_l^{(n)} + C_j b_j} \right]^m \tag{2.12}$$

where $\omega^{\mathbf{p}'}$ is the decomposition of the dataset $\mathbf{p}'$ (Equation 2.10) in wavelets. In the Richardson-Lucy scheme the comparison between the dataset **p** and the projection of the object **a** adopts the form of division, with the former in the numerator and the latter in the denominator. Although one cannot compare each wavelet plane of **p** with a corresponding plane of the object projection **a**, the decomposition of **p** allows one to stop the reconstruction at each wavelet plane independently. In order for this to happen the numerator and the denominator in equation (2.9) have to be equal. The resulting value of unity will be multiplied by the current iteration in order to calculate the next one, so $\mathbf{a}^{(\mathbf{n+1})} = \mathbf{a}^{(\mathbf{n})}$. Such a mechanism would provide for convergence of the algorithm. Therefore, in order to control noise suppression at each wavelet plane AWMLE provides for a mechanism that makes equal, in equation 2.12, this particular plane of the dataset to the corresponding wavelet plane of the reconstruction achieved until this moment. In other words, $\omega^{\mathbf{p}'}_{\mathbf{t},\mathbf{j}} = \omega^{\mathbf{h}(\mathbf{n})}_{\mathbf{t},\mathbf{j}}$. For instance, if we want to stop the reconstruction in the first wavelet plane we would have:

$$a_i^{(n+1)} = K a_i^{(n)} \left[\frac{1}{q_i} \sum_{j=1}^{D} \frac{f_{ji}(\omega^{h^{(n)}}_{0,j} + \omega^{p'}_{1,j} + ...)}{\sum_{l=1}^{B} f_{jl} a_l^{(n)} + C_j b_j} \right]^m \tag{2.13}$$

We now have a mechanism which allows us to control the level of reconstruction of each plane. The stopping rule is guided by the measurement of the level of noise in each plane. The algorithm measures, at each plane, the standard deviation within a window and compares its value with the standard deviation of the whole plane. If the local standard deviation is equal or smaller than the global standard deviation, we can conclude that within that window the probability of finding signal is small. On the other hand, if the standard deviation within the window is higher than the global standard deviation measured in this plane, we can assume that the information found inside the window is probably due to the presence of a real source. The size of such windows must be changed according to the wavelet plane one is working with at the moment, i.e. the window size is related to the wavelet scale. Hence, the window will be smaller for the high-frequency scales and its size will increase as the frequency of the wavelet plane decreases.

Several mathematical expressions have been proposed to quantify the probability of finding a real source within the window. Here we use:

$$m_i = \begin{cases} 1 - exp\left\{ \dfrac{-\frac{1}{8}(\sigma_i^2 - \sigma_\omega^2)}{2\sigma_\omega^2} \right\} & if\ \sigma_i - \sigma_\omega > 0 \\ 0 & if\ \sigma_i - \sigma_\omega \leq 0 \end{cases} \tag{2.14}$$

with:

$$\sigma_i = \sqrt{\frac{\sum_{j\in\Phi}(\omega_{t,j})^2}{n_f}}$$

where σ_i is standard deviation within the window $\mathbf{\Phi}$ centered on the pixel i of the wavelet plane $\omega_\mathbf{t}$, pixels within that window are indexed with j, n_f is the number of pixels contained in the window, σ_ω is the global standard deviation in the corresponding wavelet plane. It can be approximated by decomposing in wavelet coefficients an artificial Gaussian noise image with standard deviation equal to the dataset's.

With the inclusion of the probabilistic mask in equation (2.12) the AWMLE expression becomes:

$$a_i^{(n+1)} = Ka_i^{(n)} \left[\frac{1}{q_i} \sum_{j=1}^{D} \frac{f_{ji} \sum_t \left(\omega_{t,j}^{h^{(n)}} + m_{t,j} (\omega_{t,j}^{p'} - \omega_{t,j}^{h^{(n)}}) \right)}{\sum_{l=1}^{B} f_{jl} a_l^{(n)} + C_j b_j} \right]^m \tag{2.15}$$

If $\sigma_i \leq \sigma_\omega$, that is if no significant signal is found within the window at some position in wavelet plane t, then m_i will be zero and the numerator of equation (2.15) will be $\omega_{t,j}^{h^{(n)}}$, so the reconstruction will be stopped at that plane since the algorithm no longer works with the dataset $\mathbf{p}$, but with the projection $\mathbf{h}$ of the last reconstruction $\mathbf{a^{(n)}}$. On the other hand, if $\sigma_i > \sigma_\omega$ then m_i will tend to one and the numerator will be closer to $\omega_{t,j}^{p'}$ that is the wavelet plane of the dataset $\mathbf{p}$. Therefore, the last reconstruction $\mathbf{a^{(n)}}$ would be still compared to the data and the reconstruction would continue in that plane within that window.

Equation (2.15) was obtained with the assumption of Poisson noise in the data. For high-light level conditions the Poisson distribution approaches the Gaussian distribution and so equation (2.15) can be considerably simplified in that case.

To summarize, AWMLE is fully described by equation (2.15). It maximizes the compound likelihood of the data being Poisson-distributed and the read-out noise being Gaussian. The dataset $\mathbf{p}$ is decomposed by the *à trous* algorithm in order to enable independent operation within each wavelet plane. AWMLE can automatically stop the reconstruction process at each plane by means of a probabilistic mask which locally decides, comparing standard deviation within a window with the standard deviation of the entire wavelet plane, if the presence of a real signal can be deduced.

2.3 Relevance for AO observations

Wavelet decomposition in AWMLE can be used to detect objects of varying size in the object domain, for example galaxies and stars, and separate these sources from noise when deconvolving. Additionally, in the data domain, wavelet decomposition could guide the stopping rule: iterative deconvolution should be stopped at the smallest wavelet scale associated with the PSF. This is particularly useful for AO observations which are known to possess two distinct scales: the diffraction-limit, λ/D, which gives the core, the Airy rings, and the diffraction from the "spiders", and the seeing-limited scale which corresponds to the halo. This halo starts beyond the so-called "AO cut-off frequency" in the PSF (location in the focal plane corresponding to wavefront scales smaller than twice the inter-actuator spacing on the deformable mirror) and extends up to λ/r_0, where r_0 is the so-called Fried parameter or atmospheric coherence length. There is some controversy whether for closed-loop operation this PSF scale stays constant or whether it increases (Hardy 1998; Cagigal & Canales 2000; Tyler 2006). This ambiguity does not impact the effectiveness of PSF decomposition into the wavelet scales because in our work we estimate the size of the halo directly from the images. In this paper we deal with point sources only, and so the wavelet decomposition is not used to distinguish objects of varying size from noise. Rather, it is used to guide the reconstructions for the two characteristic scales of the AO PSF.

2.4 Dataset description

Table 2.1: PSFs used to simulate images of binary stars

		Science PSF	SR	m_v	m_k	Reference PSF	SR	m_v	m_k
			30% Strehl ratio						
Matched PSF	....	NOMAD1 1297-0510182	29%	12.1	5.93	–	32%	–	–
Mismatched PSF	....	NOMAD1 1297-0510182	29%	12.1	5.93	HD 18009	35%	8.23	5.02
			50% Strehl ratio						
Matched PSF	....	HD 143209	54%	6.3	3.92	–	52%	–	–
Mismatched PSF	....	HD 143209	54%	6.3	3.92	HD 153832	48%	7.25	4.78

We have tested flux conservation of AWMLE on synthetic observations of binary stars constructed from real AO images of point sources. Datasets were obtained with the Lick Observatory AO system on the 3m Shane telescope (Bauman et al. 1999). Closed loop images of bright, single stars were obtained using the high-speed subarray mode with size 64×64 pixel for the 256×256 pixel IRCAL camera (Fitzgerald & Graham 2006). This corresponds to a field size of $4.864 \times 4.864''$, the pixel size is $0.076''$. The subarray measurements were captured with typical exposure times of 22ms. All data were obtained in K band ($2.2\mu m$) where the diffraction limit is 151mas, so that the data were effectively Nyquist sampled. The individual 10000 short exposures were registered with subpixel accuracy to produce shift-and-add (SAA) images. The registration algorithm locates the peak in each of the images and estimates its sub-pixel position by computing the centroid of a 3×3 pixel box centered on the peak pixel. This box is big enough to contain the diffraction-limited core of the AO PSF. Each frame is then shifted by computing its Fourier transform and applying a slope to the Fourier phases corresponding to the offset of the sub-pixel peak location from the frame center, i.e. pixel (32,32), and then computing the inverse Fourier transform. The procedure is repeated until a tolerance of 0.02 pixels (2mas) is reached. The average Strehl Ratio (SR) of the SAA images was around 40% (Strehl ratio: the ratio of the peaks of the observed PSF and the diffraction-limited PSF). SR is the most-commonly used metric in AO imaging: SR of 100% implies perfect correction. Observed targets were very bright so the final stacked PSFs covered almost the entire field of view. Therefore, readout noise and background were estimated from the individual short exposures. For details of the observations and data reduction, see earlier paper by Gladysz et al. (2006).

Artificial binaries, with differential magnitude of $\triangle m = 3.5$, were created by scaling and shifting the single-star SAA images. The low-Strehl-ratio case was constructed from a dataset corresponding to the star bright in near-IR but whose flux in the visible was barely sufficient to drive the wavefront sensor. Consequently the AO frame rate was only 55Hz, as opposed to the usual 500Hz for the other targets. The resulting SR was low due to the temporal delay error (lag between wavefront sensor's measurement and evolution of the atmosphere).

For each of the cases presented in Table 2.1 twenty four positions on the circle of radius 0.6″ (8 pixels) were tested in order to minimize variations in results due to possible anisotropies in the PSF, as can be appreciated in Figure 2.4. Another twenty four positions on the circle of radius 1.2″ (16 pixels) were also tested in order to study the more straightforward scenarios of almost-non-overlapping PSFs. Here we come to the question of whether it is safe to assume the shift-invariant PSF for these images. In wide-field AO images PSF generally changes depending on its proximity to the guide star, as stated in section 1. The angular scale defining the region where the PSF can be assumed to be constant is the so-called isoplanatic angle which, just like the coherence length, is bigger for longer wavelengths. Christou & Drummond (2006) have measured this angle for the Lick site and at the same wavelength as in our observations (2.2μm) and found the values in the range $14-27''$. This confirms that our methodology to simulate binary stars with the same PSF for angles $< 1.5''$ is correct.

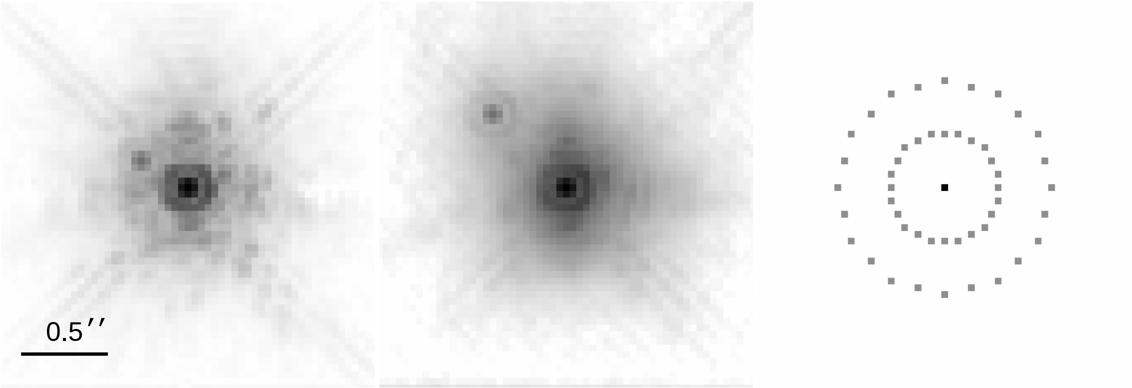

Figure 2.4: Leftmost panel: binary star simulated using a PSF with SR=50% (separation = 0.6″). Central panel: binary star simulated using a PSF with SR=30% (separation = 1.2″). Note the speckles are more pronounced for SR=50%. Logarithmic scale, inverted colors. Rightmost panel: positions where the companion was located with respect to the main star.

We studied low (SR=30%) and relatively-high-compensation (SR=50%) cases with clearly distinct PSF structures. In the high-SR case quasi-static speckles are vividly pronounced, and in the low-SR case the smooth halo is the dominant structure outside the first Airy ring. In the end, we worked with four sets of twenty four binaries each.

Both AWMLE and StarFinder need an initial estimate of the PSF to work with. For the “matched-PSF” cases we used the same stars observed ten minutes later. For the “mismatched-PSF” cases we used stars of similar brightness observed on the same night, and also close to zenith. Variability of the Strehl ratio between the science and calibration PSF is a direct consequence of the non-stationarity of turbulence.

The StarFinder PSF-fitting algorithm was designed for the analysis of crowded stellar fields. Consequently, the fitting algorithm can take advantage of many estimates of the PSF within the field of view. The PSF template is also updated as StarFinder goes “deeper” into an image. In our work we provide StarFinder with the PSF estimate, i.e. an image of a single star which remains static throughout the analysis. StarFinder has also to be given an estimate for the position of the companion and we observed that this estimate has to be precise, otherwise the algorithm converged on the brightest static speckle. This situation was treated as non-convergence. Additionally, if StarFinder converged on a position more than 0.5pixel away from the true position, the result was also discarded. This was done in order to remove the photometric estimates affected by errors in astrometry. We checked that up to 0.5pixel of astrometric error photometry obtained through PSF-fitting does not suffer, but above that value, photometry (even when given a perfect PSF) exhibits a steep decrease in precision.

AWMLE is not a photometric package. It produces a deconvolved image which can be subsequently analyzed by an observer. However, atmospheric turbulence and optical imperfections of the AO system usually produce speckles in AO images. These speckles can be confused with real sources (Gladysz & Christou 2009). We checked the deconvolved images with the commonly-used GAIA-sextractor software, and verified that the companion is detectable while the artefacts (speckles, or the results of their propagation through the reconstruction software) are not detected.

StarFinder produces a list of detected objects, their fluxes and positions. For the AWMLE reconstructions photometry was measured using a circle of 5-pixel diameter.

2.5 Results

In order to compare both algorithms, two different tests were performed. Firstly, we were interested in the *average* photometric accuracy which can be achieved with state-of-the-art algorithms for challenging AO data. By “challenging” we mean data produced by PSFs with complicated structures, and with signal of the companion comparable to the speckles. Even though SNRs of the companions were in the range 5-15 for the 8-pixel separations, accurate extraction of their photometry proved to be difficult. The question we pose is therefore: given a clear detection, what is the accuracy of the subsequent photometric measurement, and of the final physical characterization.

There are many papers on AO photometry (Roberts et al. 2005, 2007; Turner et al. 2008) but none of them, except Esslinger & Edmunds (1998) and Burke et al. (2009), discuss the accuracy of algorithms tested on simulated data. In this paper we present methodology for such tests and discuss photometric accuracy which AO observers should expect for closely-separated binary stars with large magnitude differences and with significant speckle noise.

Additionally, we performed another set of tests concerned with *reliability* (or repeatability) of the results. In many astronomical observations several measurements are usually taken and the resulting estimated quantities are averaged across the sample set. We wanted to test what would be the scatter of such photometric measurements, whether any single answer would be (statistically) close to the truth, and whether the averaging process would actually bring the observer closer to the true value. For these tests we averaged photometric estimates for each position and thus we averaged quantities corresponding to the same scenario. By “scenario” we mean underlying speckle structure which, for the PSFs used throughout this work, is spatially anisotropic and there are systematic photometric errors depending on whether the companion was positioned on a speckle or on a “hole”. One of our goals was to illustrate these systematic effects.

AWMLE was executed with two different probabilistic masks. The first one is described by equation (2.14) and is denoted in the tables and plots by σ-AWMLE, and the second one is the so-called deterministic-mask which includes information about the position of the companion, and is denoted by DM-AWMLE in the tables and plots. The latter method rejects a possible false reconstruction of speckles. This can be considered a similar situation with respect to StarFinder which needs to be given a first estimate for the position of the companion.

In our work we dealt with companions with relatively low counts and therefore the Poisson law expressed by equation (2.15) was preferred over the Gaussian-only version of AWMLE. A wavelet decomposition with one plane plus a residual plane ($\mathbf{p} = \omega_{\mathbf{0}}^{\mathbf{p}} + \mathbf{c}^{\mathbf{p}}$) was performed. The size of the first wavelet scale was matched to the diffraction limit of the telescope (2 pixels) while the dominant scale in the residual plane was roughly the diameter of the extended PSF halo. A maximum number of 100 iterations was enough to achieve the convergence in both planes. The acceleration parameter m was set to 1 so it had no real effect on the results.

2.5.1 Accuracy test

This test was designed to measure the average photometric accuracy which we believe can be obtained with state-of-the-art algorithms for barely-detectable companions (separation = 0.6″, $\triangle m = 3.5$, SNR = 5-15 depending on the Strehl ratio), and clearly-visible companions (separation = 1.2″, $\triangle m = 3.5$, SNR = 30-50). The algorithms were executed for the 24+24 images. Two metrics were computed for the vectors of flux ratios resulting from the algorithms. First, the parameter p_1 (Gladysz et al. 2010a):

$$p_1 = \frac{1}{N} \sum_{j=1}^{N} |\triangle m_j - \triangle m_t| \tag{2.16}$$

Where $\triangle m_j$ is the estimated magnitude difference at position j and $\triangle m_t$ is the real magnitude difference, i.e. 3.5. This metric gives an estimate of the photometric accuracy averaged over the PSF's anisotropies. Both algorithms were able to offer a solution for all positions in the SR=50% situation. However, for the lower-SR case StarFinder did not converge in many of the positions as can be seen in Table 2.3. In such situations we only computed the best AWMLE N results, where N is the number of successful runs for StarFinder.

Secondly, in order to check for possible over- or underestimation we computed the mean magnitude difference averaged over the number of successful (or most accurate, see above) runs:

$$\bar{m} = \frac{1}{N} \sum_{j=1}^{N} \triangle m_j \tag{2.17}$$

The corresponding standard deviation of magnitude differences was calculated as follows:

$$\sigma_m = \frac{|\sigma^+| + |\sigma^-|}{2} \tag{2.18}$$

where σ^+ and σ^- are "upper" and "lower" standard deviations expressed in magnitudes:

$$\sigma^+ = 2.5 \cdot log(\bar{f}) - 2.5 \cdot log(\bar{f} - \sigma_f) \tag{2.19}$$

$$\sigma^{-} = 2.5 \cdot log(\overline{f}) - 2.5 \cdot log(\overline{f} + \sigma_f) \tag{2.20}$$

Due to the logarithmic character of the magnitude scale these two values are not exactly equal but are very close. In the above equations $\overline{f}$ is the mean value of the measured flux ratios and σ_f is the standard deviation. The magnitude difference error bars had to be normalized to account for StarFinder's non-convergences. We chose to divide σ_m by $\sqrt{N}$, where N is the number of StarFinder's convergences. This particular normalization is of course used to compute the standard deviation of the mean (Taylor 1996) and, to be precise, it is only applicable if one measures the same quantity a number of times in similar experimental conditions. This is not the case here because each position for the companion is endowed with its systematic error due to the PSF's anisotropies, as seen in Figure 2.7, while equations (2.17)-(2.20) are averages over *all* locations. Nevertheless, we chose it to represent somehow the scatter of the (successful) executions of the codes. Numerical results are presented in Tables 2.2 and 2.3. They are also plotted in Figure 2.5.

Table 2.2: Accuracy test results for positions on a 1.2″ circle

		50% SR Matched PSF			50% SR Mismatched PSF			30% SR Matched PSF			30% SR Mismatched PSF		
		p_1	$\overline{m}$	$\sigma_m/\sqrt{N}$	p_1	$\overline{m}$	$\sigma_m/\sqrt{N}$	p_1	$\overline{m}$	$\sigma_m/\sqrt{N}$	p_1	$\overline{m}$	$\sigma_m/\sqrt{N}$
σ-AWMLE	....	0.0072	3.5008	0.0018	0.0247	3.5123	0.0053	0.0498	3.4926	0.0136	0.0721	3.5083	0.0195
DM-AWMLE	....	0.0130	3.5079	0.0016	0.0383	3.5327	0.0046	0.0339	3.5043	0.0096	0.0794	3.5455	0.0173
StarFinder	....	0.0140	3.5088	0.0009	0.0709	3.5657	0.0042	0.0782	3.4217	0.0100	0.1121	3.4740	0.0254

Looking at Figure 2.5 it is reassuring to find the general increase in average photometric error (p_1) when going from the "easiest" case (50% SR, matched PSF) to the most challenging one (30% SR, mismatched PSF). For well separated binary stars photometry will be endowed with errors less than 0.1mag for companions with low SNRs (5-15). For smaller separations, approximately 0.5″, differential photometry will have an error less than 0.5mag, or 0.3mag when the StarFinder's outlier is ignored.

Table 2.3: Accuracy test results for positions on a 0.6″ circle

		50% SR Matched PSF			50% SR Mismatched PSF			30% SR Matched PSF			30% SR Mismatched PSF		
		p_1	$\overline{m}$	$\sigma_m/\sqrt{N}$	p_1	$\overline{m}$	$\sigma_m/\sqrt{N}$	p_1	$\overline{m}$	$\sigma_m/\sqrt{N}$	p_1	$\overline{m}$	$\sigma_m/\sqrt{N}$
σ-AWMLE	....	0.0510	3.5404	0.0073	0.1469	3.6418	0.0189	0.1537	3.3738	0.0329	0.3216	3.2625	0.0918
DM-AWMLE	....	0.0609	3.5543	0.0068	0.1488	3.6436	0.0155	0.1162	3.4097	0.0258	0.2987	3.2967	0.0906
StarFinder	....	0.0472	3.5421	0.0038	0.2074	3.7022	0.0130	0.4530	3.0418	0.0223	0.2746	3.4506	0.0935
										(6 err)			(12 err)

Locations of the three symbols for the scenario of separation 1.2″, 30% SR, mismatched PSF, reveal one important characteristic of the p_1 metric which we also call "mean absolute deviation from the truth". For results which alternate between under- or over-estimating the truth, never getting close to it, p_1 will be relatively large, as is the case for StarFinder's p_1 here. On the other hand, StarFinder's mean value is actually relatively close to the truth precisely because its results alternate between under- or over-estimation. For the same scenario, AWMLE was on average closer to the truth but it quite consistently over-estimated the flux ratio and so the resulting mean value is further away

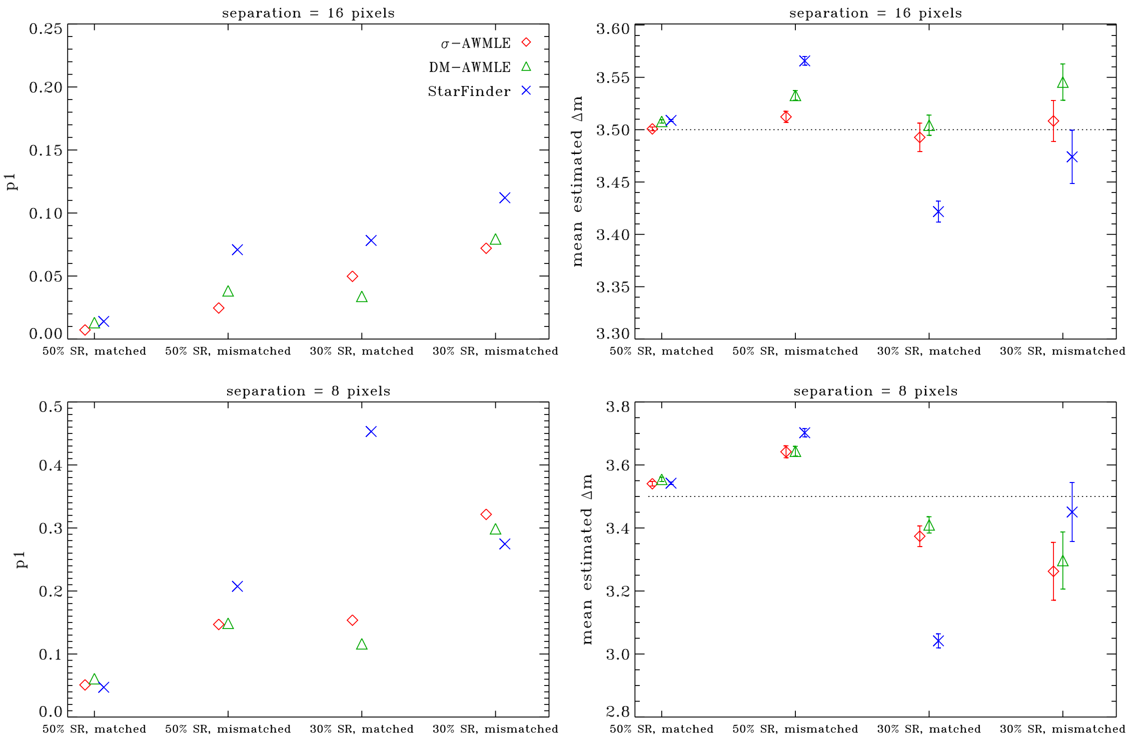

Figure 2.5: Upper panels: metric p_1 and mean values calculated for the separation of 1.2″. Lower panels: same metrics computed for the separation of 0.6″. The dashed line represents the true value, i.e. 3.5. Note the differences in vertical scales between the plots. Numerical values are given in Tables 2.2 and 2.3. Symbols corresponding to the three approaches were offset horizontally from their tick marks to reduce crowding.

from the truth than StarFinder's. In general, one can see that AWMLE produces results with lower photometric error compared to PSF-fitting. For the most challenging scenario (separation 0.6″, 30% SR, mismatched PSF) StarFinder produced accurate results but it converged in only half of all the 24 cases.

Figure 2.6 shows the reconstructions performed by AWMLE for one of the positions at the separation of 0.6″ and the four situations we have considered, i.e. 30% and 50% SR with matched and mismatched PSFs.

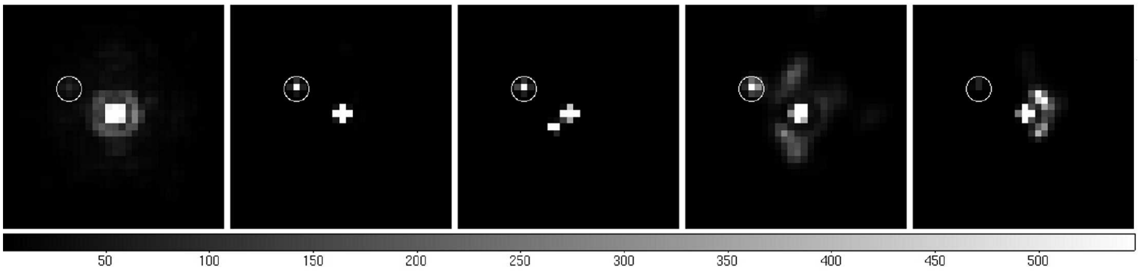

Figure 2.6: From left to right: a) Image of the synthetic binary star with companion at position "1" (50% SR), b) AWMLE reconstruction for 50% SR and matched PSF, c) AWMLE reconstruction for 50% SR and mismatched PSF, d) AWMLE reconstruction for 30% SR and matched PSF, e) AWMLE reconstruction for 30% SR and mismatched PSF.

2.5.2 Repeatability test

Table 2.4: Repeatability test results for SR=50% and matched PSF

		$\bar{m}$	$\sigma_m/\sqrt{N}$	$\bar{m}$	$\sigma_m/\sqrt{N}$	$\bar{m}$	$\sigma_m/\sqrt{N}$	$\bar{m}$	$\sigma_m/\sqrt{N}$
		Pos 1		Pos 2		Pos 3		Pos 4	
σ-AWMLE	....	3.5298	0.0190	3.5862	0.0213	3.5879	0.0177	3.4646	0.0250
DM-AWMLE	....	3.5397	0.0187	3.5960	0.0197	3.6027	0.0165	3.4810	0.0232
StarFinder	....	3.5301	0.0145	3.5607	0.0145	3.5689	0.0131	3.5099	0.0214
		Pos 5		Pos 6		Pos 7		Pos 8	
σ-AWMLE	....	3.5392	0.0129	3.5694	0.0172	3.5508	0.0118	3.5442	0.0279
DM-AWMLE	....	3.5536	0.0125	3.5817	0.0165	3.5638	0.0109	3.5575	0.0277
StarFinder	....	3.5365	0.0081	3.5601	0.0160	3.5446	0.0097	3.5511	0.0244

For the purpose of this test eight out of the twenty four positions on the circle of radius 0.6″ were chosen. Ten SAA images, comprising 1000 frames each, were created for each of these eight positions. The mean values and standard deviations of the estimated magnitude differences were calculated using equations (2.17) and (2.18). At each position, only the best N results were considered for AWMLE, where N is the number of convergences for StarFinder. If StarFinder produced results for all the frames then $N = 10$. When StarFinder was not able to converge for any of the 10 images then all the results obtained by AWMLE were used, and again $N = 10$. Standard deviation σ_m was divided by $\sqrt{N}$. This operation is now fully justified from the statistical point of view as the averaging was performed for a uniform sample of experimental results (Taylor 1996). Results are presented in Tables 2.4 - 2.7 and in Figure 2.7.

Table 2.5: Repeatability test results for SR=50% and mismatched PSF

		$\bar{m}$	$\sigma_m/\sqrt{N}$	$\bar{m}$	$\sigma_m/\sqrt{N}$	$\bar{m}$	$\sigma_m/\sqrt{N}$	$\bar{m}$	$\sigma_m/\sqrt{N}$
		Pos 1		Pos 2		Pos 3		Pos 4	
σ-AWMLE	....	3.5284	0.0216	3.5796	0.0201	3.5411	0.0180	3.6464	0.03131
DM-AWMLE	....	3.5427	0.0204	3.5940	0.0180	3.5608	0.0176	3.6556	0.0267
StarFinder	....	3.6197	0.0166	3.6170	0.0167	3.6283	0.0140	3.7232	0.0265
		Pos 5		Pos 6		Pos 7		Pos 8	
σ-AWMLE	....	3.7728	0.0186	3.8124	0.0223	3.6594	0.0179	3.7098	0.0313
DM-AWMLE	....	3.7245	0.0149	3.7787	0.0173	3.6558	0.01519	3.7028	0.0280
StarFinder	....	3.7709	0.0103	3.8123	0.0206	3.7290	0.0123	3.7439	0.0299

Except for the “easiest” case (50% SR, matched PSF) one can see that photometry of faint companions can produce wildly varying results depending on whether the companion sits on a significant speckle or not. Taking for example results from the upper-right panel in Figure 2.7, photometric error can be as low as 0.03mag or as high as 0.3mag depending on the position of the companion within the (anisotropic) speckle halo. One would expect these anisotropies to disappear for the observation time

Table 2.6: Repeatability test results for SR=30% and matched PSF

	$\bar{m}$	$\sigma_m/\sqrt{N}$	$\bar{m}$	$\sigma_m/\sqrt{N}$	$\bar{m}$	$\sigma_m/\sqrt{N}$	$\bar{m}$	$\sigma_m/\sqrt{N}$
	Pos 1		Pos 2		Pos 3		Pos 4	
σ-AWMLE	3.4587	0.0616	3.5271	0.0633	3.5634	0.0913	2.9352	0.0340
DM-AWMLE	3.5334	0.0263	3.5764	0.0363	3.5616	0.0874	3.0269	0.0402
StarFinder	3.2537	0.0380	3.2760	0.0175	3.4531	0.0200	2.9489	0.0344
		(3 err)		(4 err)		(4 err)		(3 err)
	Pos 5		Pos 6		Pos 7		Pos 8	
σ-AWMLE	3.1761	0.0472	3.2570	0.0199	3.4449	0.0342	3.0010	0.0413
DM-AWMLE	3.1979	0.0467	3.3501	0.0251	3.4585	0.0318	3.0309	0.0429
StarFinder	3.1151	0.0667	3.1781	0.0618	3.1709	0.0641	–	–
		(1 err)		(1 err)		(1 err)		

of approximately 10 minutes but the speckles we deal with here are the so-called “super speckles” or “quasi-static” speckles which decorrelate very slowly (Hinkley et al. 2007). One of the goals of this paper was to show this effect on photometry. All three approaches show the same trends dictated by the underlying noise variance. Somewhat surprisingly, the trends are also visible for the 30%-SR PSF, even though it has a much smoother halo compared to the 50%-SR PSF. On the other hand due to the lower SR, the companions had lower SNRs in these inferior-quality observations and therefore the photometric results were just as susceptible to (less visible) PSF anisotropies, especially “spiders”, as in the case of the better-quality images. We conclude that averaging many values of differential photometry in the presence of “quasi-static” speckles will not necessarily bring the observer closer to the truth.

Table 2.7: Repeatability test results for SR=30% and mismatched PSF

	$\bar{m}$	$\sigma_m/\sqrt{N}$	$\bar{m}$	$\sigma_m/\sqrt{N}$	$\bar{m}$	$\sigma_m/\sqrt{N}$	$\bar{m}$	$\sigma_m/\sqrt{N}$
	Pos 1		Pos 2		Pos 3		Pos 4	
σ-AWMLE	4.5724	0.1855	4.4923	0.1607	4.2184	0.2631	3.0135	0.0507
DM-AWMLE	4.0475	0.0908	3.9810	0.0766	3.8161	0.1247	3.0257	0.0515
StarFinder	3.7636	0.1001	3.6873	0.0827	3.3635	0.2003	2.7597	0.0448
		(1 err)		(1 err)		(6 err)		(2 err)
	Pos 5		Pos 6		Pos 7		Pos 8	
σ-AWMLE	3.0114	0.0538	3.2314	0.0501	3.5212	0.0530	5.2096	0.4612
DM-AWMLE	2.8742	0.0472	3.1760	0.0380	3.4133	0.0442	4.3781	0.0986
StarFinder	–	–	–	–	–	–	–	–

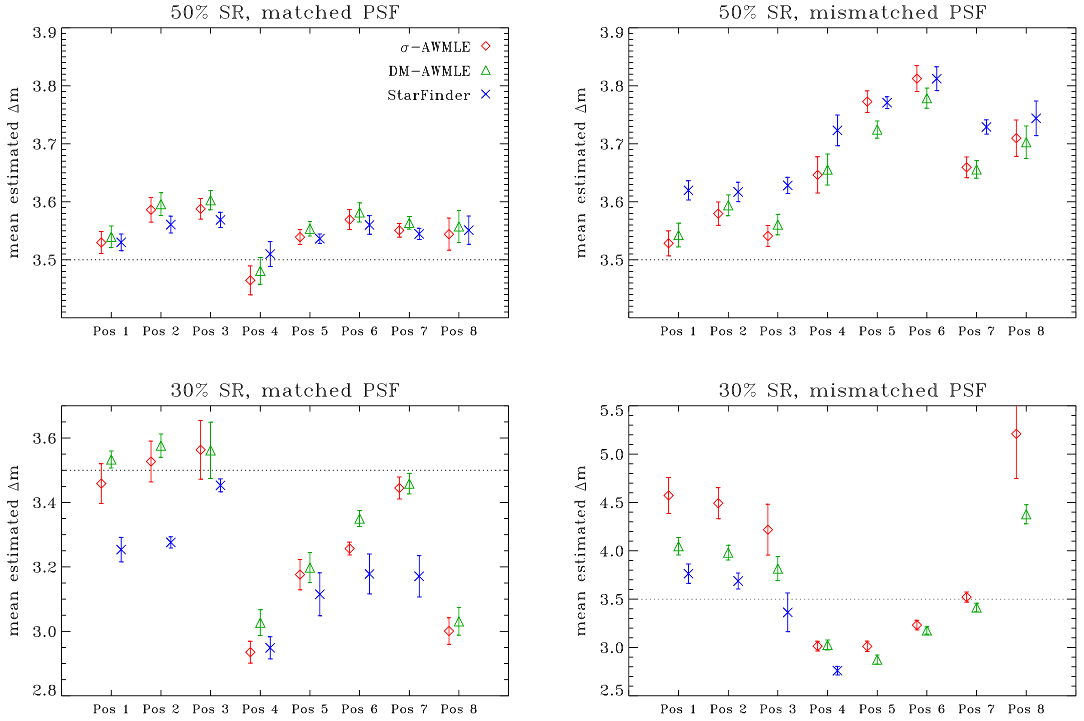

Figure 2.7: Mean values of estimated differential photometry and their standard deviations for eight positions on the circle with radius 0.6″. Upper-left panel: SR=50% and matched PSF. Upper-right panel: SR=50% and mismatched PSF. Lower-left panel: SR=30% and matched PSF. Lower-right panel: SR=30% and mismatched PSF. The dashed line represents the true value, i.e. 3.5. Note the differences in vertical scales between the plots. Numerical values are given in Tables 2.4 - 2.7. Symbols corresponding to the three approaches were offset horizontally from their tick marks to reduce crowding.

2.6 Conclusions

We presented a systematic approach to testing of photometric approaches on challenging AO data. The two codes we included in our work represent two opposite approaches to measuring photometry. StarFinder is a PSF-fitting package and yields photometry directly, while AWMLE is a Richardson-Lucy-type algorithm and has to be combined with aperture photometry.

Our first goal was to test the widely-accepted assumption that static-PSF codes, such as AWMLE, are not suitable for the analysis of AO data. Historically, the “traditional” image reconstruction and AO communities have not interacted much with each other and we believe this paper is one of the first attempts at such collaborations.

Results of the testing show that AWMLE, and modern Richardson-Lucy-type algorithms in general, is a suitable tool to measure differential photometry in AO images. The algorithm combines the Bayesian framework, assuming Poisson plus Gaussian noise in the data, with the wavelet transform. In this work, the wavelet decomposition is used to better separate PSF-induced features in the data from noise. A probabilistic mask decides automatically when to stop the reconstruction process for each wavelet plane. The use of a deterministic mask, telling the algorithm exactly where the stars are, improves the results slightly because reconstruction of the speckles, which affects the final reconstructed signals of the real sources, is rejected. We have shown that AWMLE combined with aperture photometry produces photometric estimates which are comparable if not more accurate than PSF-fitting.

We conclude that state-of-the-art photometric codes can have precisions ranging from 0.3mag for very close binary stars, low SNRs of the companions, and mismatched PSFs; down to 0.01mag for well separated systems, clearly visible companions and well-matched PSFs.

We are currently conducting a large study where we compare photometric precision of AWMLE, StarFinder, multi-frame blind deconvolution, and the so-called “PDF deconvolution” (Gladysz & Christou 2009). Preliminary results were already presented in Gladysz et al. (2010a). Additionally, we believe that AWMLE, as a reconstruction algorithm, could be used to deconvolve AO images of more complex objects like asteroids or galaxies. Besides, other transforms could be considered for this purpose, e.g. the curvelet transform (Candès et al. 2006).

Chapter 3

Extended-object reconstruction in adaptive-optics imaging: the multiresolution approach

Published in Astronomy and Astrophysics (AA). Year 2013, volume 555, article A69, pages 1-15.

Abstract

We propose the application of multiresolution transforms, such as wavelets and curvelets, to reconstruct images of extended objects that have been acquired with adaptive-optics (AO) systems. Such multichannel approaches normally make use of probabilistic tools to distinguish significant structures from noise and reconstruction residuals. We aim to check the prevailing assumption that image-reconstruction algorithms using static point spread functions (PSF) are not suitable for AO imaging. We convolved two images, one of Saturn and one of galaxy M100, taken with the Hubble Space Telescope (HST) with AO PSFs from the 5*m* Hale telescope at the Palomar Observatory and added shot and readout noise. Subsequently, we applied different approaches to the blurred and noisy data to recover the original object. The approaches included multiframe blind deconvolution (with the algorithm IDAC), myopic deconvolution with regularization (with MISTRAL) and wavelet- or curvelet-based static PSF deconvolution (AWMLE and ACMLE algorithms). We used the mean squared error (MSE) to compare the results. We found that multichannel deconvolution with a static PSF produces generally better results than the results obtained with the myopic/blind approaches (for the images we tested), thus showing that the ability of a method to suppress the noise and track the underlying iterative process is just as critical as the capability of the myopic/blind approaches to update the PSF. Furthermors, for these images, CT produces better results than WT, as measured in terms of MSE.

3.1 Introduction

The distortions introduced into images by the acquisition process in astronomical ground-based observations are well known. Apart from the most common, such as vignetting, non-zero background or bad pixels, which must be removed prior to any other analysis, atmospheric turbulence limits the spatial resolution of an image, whereas the electronic devices used to acquire and amplify the signal introduces noise. An image is also corrupted by Poisson noise from fluctuations in the number of received photons at each pixel (Andrews & Hunt 1977). The classical equation that describes the image formation process is

$$image = [PSF * object] \diamond noise \, , \tag{3.1}$$

where $*$ denotes the convolution operation. The symbol $\diamond$ is a pixel-by-pixel operation that reduces to simple addition when noise is additive and independent of $[PSF * object]$, while for Poisson noise it is an operation that returns a random deviate drawn from a Poisson distribution with mean equal to $[PSF * object]$. It is well known that direct inversion of equation 3.1 in the Fourier domain amplifies noisy frequencies close to the cut-off frequency. Hence, in the presence of noise, such a simple method cannot be used.

Several deconvolution approaches have been proposed to estimate the original signal from the seeing-limited and noise-degraded data. Since equation 3.1 is an ill-posed problem, with non-unique stable solutions, one approach is to regularize the Fourier inversion to constrain possible solutions (Tikhonov et al. 1987; Bertero & Boccacci 1998). This method generally imposes a trade-off between noise amplification and the desired resolution, which generally leads to smooth solutions. Bayesian methodology (Molina et al. 2001; Starck et al. 2002) allows a solution compatible with the statistical nature of the signal to be sought, leading to maximum likelihood estimators (MLE) (Richardson 1972; Lucy 1974) or maximum a posteriori (MAP) approaches if prior information is used, e.g., the positivity of the signal or entropy (Frieden 1978; Jaynes 1982).

All these methods can be enhanced through multichannel analysis by decomposing the signal in different planes, each of them representative of a certain scale of resolution. In such a decomposition, fine details in an image are confined to some planes, whereas coarse structures are confined to others. One of the most powerful ways to perform this decomposition is by means of the wavelet transform (WT). In particular in the astronomical context, the undecimated isotropic à trous algorithm (Holschneider et al. 1989; Shensa 1992; Starck & Murtagh 1994), also known as the starlet transform, is often used. The WT creates a multiple representation of a signal, classifying its frequencies and, simultaneously, spatially localizing them in the field of view. For the specific case of the starlet transform, this can be expressed by

$$\mathbf{s} = \omega^{\mathbf{s}}_{\mathbf{0}} + \omega^{\mathbf{s}}_{\mathbf{1}} + \omega^{\mathbf{s}}_{\mathbf{2}} + ... + \omega^{\mathbf{s}}_{\mathbf{N}} + \mathbf{r}^{\mathbf{s}}_{\mathbf{N}} \, , \tag{3.2}$$

where $\mathbf{s}$ is the signal that decomposed into wavelet coefficients, $\omega^{\mathbf{s}}_{\mathbf{j}}$ is the wavelet plane at resolution j, and $\mathbf{r}^{\mathbf{s}}_{\mathbf{N}}$ is the residual wavelet plane. We point out that equation 3.2 provides a prescription for direct reconstruction of the original image from all the wavelet planes and the residual plane. The advantage of WT is that it allows for different strategies to be used for different wavelet planes, e.g., by defining thresholds of statistical significance.

Given the advantages of the multiresolution analysis, the aforementioned deconvolution approaches have been adapted to work in the wavelet domain. The MLE was modified into a two-channel algorithm (Lucy 1994) where the first channel corresponds to the signal contribution and the second to the background. The wavelet transform has also been applied to maximum-entropy deconvolution methods (Núñez & Llacer 1998) to segment the image and apply different regularization parameters to each region. Starck et al. (2001) generalized the method of maximum entropy within a wavelet framework, separating the problem into two stages: noise control in the image domain and smoothness in the object domain.

While WT has been widely used in astronomical image analysis and in deconvolution, the reported use (in the same context) of multi-transforms with properties that improve on or complement WT is scarce. Such methods include, among many others: waveatoms (Demanet & Ying 2007), which aim to represent signals by textures; and curvelets (Candès et al. 2006), which introduce orientation as a classification parameter together with frequency and position. Starck et al. (2003) used the curvelet transform (CT) for Hubble Space Telescope (HST) image restoration from noisy data. They reported enhanced contrast on the image of Saturn. Lambert et al. (2006) applied CT to choose significant coefficients from astroseismic observations while Starck et al. (2004) used CT to detect non-Gaussian signatures in observations of the cosmic microwave background. Because it is believed that CT is more suitable for representing elongated features such as lines or edges, one of the goals of this paper is to introduce CT into the Bayesian framework and show how it performs on images of extended objects.

The aforementioned methods work with static point spread functions (PSF), i.e., they do not update the PSF of an optical system that is supplied by the operator. Nevertheless, in ground-based imaging, whether with adaptive optics (AO) or without, there are always differences between science and calibration PSFs (Esslinger & Edmunds 1998). These differences may result from changes in seeing, wind speed, slowly varying aberrations due to gravity or thermal effects and, in AO imaging, from differential response of the wavefront sensor to fluxes received from the science and calibration objects. The quality of AO images is quantified with the Strehl ratio (SR), which is the ratio of the measured peak value of a point-source image to that of the diffraction-limited PSF, often given in percent. A perfect diffraction-limited image has an SR of 100% while a seeing-degraded image on a large telescope can have an SR lower than 1%. In this paper we will use the SR to quantify the mismatch between the target and calibration PSFs. This commonly occurring mismatch has prompted optical scientists, especially those working on AO systems, to investigate blind and myopic image restoration schemes (e.g. Lane 1992; Thiébaut & Conan 1995). A blind method works without any information about the PSF, while a myopic approach relies on some initial PSF estimate that is then updated until a solution for both the object and the PSF is found.

Conan et al. (1999) compared their myopic approach to a basic, “unsupervised” Richardson-Lucy scheme and found, not surprisingly, that the myopic deconvolution is more stable. Nevertheless, to our knowledge there have not been any thorough and fully fledged efforts to compare the performance of modern, static-PSF approaches to blind/myopic methods in the context of AO imaging, and so the preference for the latter algorithms still has to be justified. Our goal is to partially fill this gap. At this stage we mention that this article is a companion paper to Baena Gallé & Gladysz (2011), where we showed that a modern, static-PSF code is capable of extracting accurate differential photometry from AO images of binary stars even when given a mismatched PSF. In the current paper we extend the analysis to more complex objects and analyze the effect of noise as well as that of the mismatched PSF. We show the importance of noise control, specifically how advanced noise suppression can set off the lack of PSF-update capability for very noisy observations. More generally, we highlight the opportunities for an exchange of ideas between the communities that prefer myopic and static-PSF approaches.

The paper is organized as follows: in Section 2, the algorithms AWMLE, ACMLE, MISTRAL, and IDAC are described. These algorithms represent different philosophies with regard to the deconvolution problem. AWMLE and ACMLE perform a classical static-PSF deconvolution within the wavelet or curvelet domain, MISTRAL is intended for myopic use, and IDAC can be used as a blind or a myopic algorithm. Section 3 describes the datasets we used and how we applied each of the four algorithms. A brief description of the mean squared error (MSE), together with concepts of error and residual maps that we used to compare the reconstructed objects can be found in section 4. Section 5 presents the performance comparison of the four algorithms mentioned above in terms of noise and PSF mismatch. Section 6 summarizes and concludes the paper.

3.2 Description of the algorithms

This section is not intended to describe the algorithms in detail but rather to offer a brief overview of their characteristics and their historical uses and performances. We would also like to point out that in the text to follow we do not keep the nomenclature originally used by the respective authors. We make the following attempt at standardization: the object, or unknown, is represented with $\mathbf{o}$, the image or dataset is $\mathbf{i}$, and the PSF is $\mathbf{h}$. Upper-case notation is used for Fourier representation. The two-dimensional pixel index is r, while f is spatial frequency. The index of a wavelet/curvelet scale is represented with j, while a single frame in a multiframe approach is indexed with i. Finally, the symbol $\hat{\ }$ is used for estimates.

3.2.1 AWMLE

The adaptive wavelet maximum likelihood estimator (AWMLE) is a Richarson-Lucy-type algorithm (Richardson 1972; Lucy 1974) that maximizes the likelihood between the dataset and the projection of a possible solution onto the data domain, considering a combination of the Poissonian shot noise, intrinsic to the signal, and the Gaussian readout noise of the detector. This maximization is performed in the wavelet domain. The decomposition of the signal into several channels allows for various strategies to be used depending on a particular channel's scale. This is a direct consequence of the fact that in a WT decomposition the noise, together with the finest structures of the signal, will be transferred into the high-frequency channels while coarse structures will be transferred into the low-frequency channels.

The general mathematical expression that describes AWMLE is

$$\mathbf{\hat{o}^{(n+1)}} = K\mathbf{\hat{o}^{(n)}}\left[\mathbf{h}^{-} * \frac{\sum_j \left(\omega_j^{\mathbf{o_h^{(n)}}} + \mathbf{M_j}(\omega_j^{\mathbf{i'}} - \omega_j^{\mathbf{o_h^{(n)}}})\right)}{\mathbf{h}^{+} * \mathbf{\hat{o}^{(n)}}}\right], \tag{3.3}$$

where $\mathbf{o}$ is the object to be estimated, $\mathbf{h}^{+}$ is the PSF that projects the information from the object domain to the image domain, while $\mathbf{h}^{-}$ is its reverse version performing the inverse operation, i.e., the projection from the image domain to the object domain. The so-called direct projection $\mathbf{o_h}$ is an image of the object projected onto the data domain by means of the PSF at iteration n, it appears in the numerator of equation 3.3 already decomposed in its wavelet representation. The parameter $K = \frac{\mathbf{i}}{\mathbf{\hat{o}^{(n+1)}}}$ is a constant to conserve the energy.

The variable $\mathbf{i'}$ is a modified version of the dataset, which appears here because of the explicit inclusion of the readout Gaussian noise into the Richardson-Lucy scheme (Núñez & Llacer 1993). This represents a pixel-by-pixel filtering operation in which the original dataset is substituted by this modified version. It must be calculated for each new iteration by means of the following expression:

$$i'(r) = \frac{\sum_{k=0}^{\infty}(ke^{-(k-i(r))^2/2\sigma^2}[(o_h(r))^k/k!])}{\sum_{k=0}^{\infty}(e^{-(k-i(r))^2/2\sigma^2}[(o_h(r))^k/k!])}, \tag{3.4}$$

where $\mathbf{i}$ is the original dataset, r is the pixel index, and σ is the standard deviation of the read-out noise. In the absence of noise ($\sigma \to 0$) the exponentials in equation 3.4 are dominant when $k = i(r)$ and then $i'(r) \to i(r)$, i.e., the modified dataset converges to the original one. We note that by definition, $\mathbf{i'}$ is always positive, while $\mathbf{i}$ can be negative because of readout noise. Expression 3.4 introduces an analytical way to consider the joint Poisson plus Gaussian probabilities into the deconvolution problem. This can also be performed in a straightforward way by means of the Anscombe transform, as mentioned by Murtagh et al. (1995), who approximated Poissonian noise as Gaussian (this approximation is of course only valid for relatively high values of $\mathbf{i}$).

In expression 3.3, the symbols $\omega_j^{\mathbf{o_h^{(n)}}}$ and $\omega_j^{\mathbf{i'}}$ correspond to wavelet coefficients, in channel j, of the multiscale representation of the direct projection $\mathbf{o_h^{(n)}}$ as well as that of the modified dataset $\mathbf{i'}$, respectively.

Of particular importance is the probabilistic mask $\mathbf{M_j}$, which locally determines whether a significant structure in channel j is present or not. If the probability of finding a source in the vicinity of a particular pixel r is considered to be high, then $\mathbf{M_j}$ will be close to 1 in this location, and so the dataset $\mathbf{i'}$ will be the only remaining term in the numerator and will be compared, at iteration n, with the estimated object at that iteration $\mathbf{o^{(n)}}$. If a signal in the vicinity of pixel r is considered to be insignificant, then $\mathbf{M_j}$ will decrease to zero and the object will be compared with itself (by means of its projection $\mathbf{o}_\mathbf{h}^{(\mathbf{n})}$), so the main fraction in equation 3.3 will be set to unity (at channel j in the vicinity of pixel r), which would stop the iterative process. Therefore, the probabilistic mask $\mathbf{M_j}$ is used to effectively halt the iterations in a Richardson-Lucy scheme. In other words, it allows a user to choose from a wide range of the maximum number of iterations to be executed, where the reconstructions corresponding to these different numbers will not exhibit significant differences between them, which arise in the classic Richardson-Lucy algorithm due to the amplification of noise.

AWMLE in the wavelet domain was first introduced in Otazu (2001). Baena Gallé & Gladysz (2011) presented the algorithm in the context of AO imaging and obtained differential photometry in simulated AO observations of binary systems. The accuracy of aperture photometry performed on the deconvolution residuals was compared with the accuracy of PSF-fitting, a classic approach to the problem of overlapping PSFs from point sources (Diolaiti et al. 2000). It was proven that AWMLE yields similar, and often better, photometric precision than StarFinder, independently of the stars' separation. Even though AWMLE does not update the PSF as it performs deconvolution, it was shown that the resulting photometric precision is robust to mismatches between the science and the calibration PSF up to 6% in terms of difference in the SR, which was the strongest tested mismatch.

3.2.2 ACMLE

Equation 3.2 offers a direct reconstruction formula for the wavelet domain, i.e., the sum of all wavelet planes (and the residual one) in which the original image had been decomposed allows one to retrieve that same original image. This explains the presence of summation $\sum_j$ in the numerator of expression 3.3, where a combination of different wavelet coefficients, some belonging to the direct projection of the object ($\mathbf{o_h}$) and some belonging to the modified dataset ($\mathbf{i}'$), creates the correction term in this Richardson-Lucy scheme.

In the curvelet domain, the reconstruction formula does not have this simple expression. For the forward and the inverse transforms, double Fourier inversions and complex operations with arrays are required (Candès et al. 2006). Therefore, to extend equation 3.3 to the Curvelet domain, we decided to work with the following expression:

$$\hat{\mathbf{o}}^{(\mathbf{n+1})} = K\hat{\mathbf{o}}^{(\mathbf{n})}\left[\mathbf{h}^- * \frac{CT^{-1}\left(CT(\mathbf{o}_\mathbf{h}^{(\mathbf{n})}) + \mathbf{M}\left(CT(\mathbf{i}') - CT(\mathbf{o}_\mathbf{h}^{(\mathbf{n})})\right)\right)}{\mathbf{h}^+ * \hat{\mathbf{o}}^{(\mathbf{n})}}\right], \tag{3.5}$$

where CT and CT^{-1} mean the forward and inverse curvelet transforms, respectively. The mask $\mathbf{M}$ is now calculated in the curvelet domain and combines coefficients from $\mathbf{o_h}$ and $\mathbf{i}'$ to create a new curvelet correction term that is inversely transformed to the spatial domain and compared with the object estimate at each iteration. The curvelet transform used in equation 3.5 corresponds to the so-called second-generation CT and is implemented in the software CurveLab[1]. This particular implementation of the CT exhibits a robust structure based on a mother curvelet function of only three parameters (scale, location, and orientation), which is faster and simpler to use than the first-

---

[1] http://www.curvelet.org

generation CT based on a complex seven-index structure, which relies on the combined usage of the starlet and the ridgelet transforms (Starck et al. 2010).

3.2.3 MISTRAL

The myopic iterative step-preserving restoration algorithm (MISTRAL) is a deconvolution method within the Bayesian framework that jointly estimates the PSF and the object using some prior information about both these unknowns (Mugnier et al. 2004). This joint maximum a posteriori (MAP) estimator is based on the following expression:

$$\begin{aligned}[\hat{\mathbf{o}}, \hat{\mathbf{h}}] &= \arg\max[p(\mathbf{i}|\mathbf{o}, \mathbf{h}) \times p(\mathbf{o}) \times p(\mathbf{h})] = \\ &= \arg\min[J_i(\mathbf{o}, \mathbf{h}) + J_o(\mathbf{o}) + J_h(\mathbf{h})] \ ,\end{aligned} \tag{3.6}$$

where $J_i(\mathbf{o}, \mathbf{h}) = -\ln p(\mathbf{i}|\mathbf{o}, \mathbf{h})$ is the joint negative log-likelihood that expresses fidelity of the model to the data ($\mathbf{i}$), $J_o(\mathbf{o}) = -\ln p(\mathbf{o})$ is the regularization term, which introduces some prior knowledge about the object ($\mathbf{o}$) and $J_h(\mathbf{h}) = -\ln p(\mathbf{h})$ accounts for some partial knowledge about the PSF ($\mathbf{h}$). The symbol p in the above expressions corresponds to the probability density function of a particular variable.

MISTRAL does not use separate models for the Poisson and the readout components of noise. Instead, a nonstationary Gaussian model for the noise is adopted. This means that a least-squares optimization with locally varying noise variance is employed,

$$J_i(\mathbf{o}, \mathbf{h}) = \sum_r \frac{1}{2\sigma^2(r)}[i(r) - (o * h)(r)]^2 \ , \tag{3.7}$$

where r stands for pixel index. This prior facilitates computing the solution with gradient-based techniques as compared to the Poissonian likelihood, which contains a logarithm. The Gaussian assumption is typical (Andrews & Hunt 1977) and it can be considered a very good approximation for bright regions of the image. The assumption can cause problems for low-light-level data recorded with modern CCDs of almost negligible readout noise.

The prior probability, $J_o(\mathbf{o})$, is modeled to account for objects that are a mix of sharp edges and smooth areas such as those that we deal with in this article. The adopted expression contains an edge-preserving prior that is quadratic for faint gradients and linear for steep ones. The quadratic part ensures a good smoothing of the faint gradients (i.e., of noise), and the linear behavior cancels the penalization of steep gradients (i.e., of edges). These combined priors are commonly called $L_2 - L_1$ (Green 1990; Bouman & Sauer 1993). The $L_2 - L_1$ prior adopted in MISTRAL has the following expression:

$$J_o(\mathbf{o}) = \mu\delta^2 \sum_r \phi(\nabla o(r)/\delta) \ , \tag{3.8}$$

where $\phi(x) = |x| - ln(1 + |x|)$ and where $\nabla o(r) = [\nabla_x o^2(r) + \nabla_y o^2(r)]^{1/2}$. Here, $\nabla_x \mathbf{o}$ and $\nabla_y \mathbf{o}$ are the object finite-difference gradients along x and y, respectively. Equation 3.8 is effectively L_2-L_1 since it adopts the form $\phi(x) \approx x^2/2$ when x is close to 0 and $\phi(x)/|x| \rightarrow 1$ when x tends to infinity. The global factor μ and the threshold δ are two hyperparameters that must be adjusted by hand according to the level of noise and the object structure. Some steps toward semi-automatic setting of these parameters were made by Blanco & Mugnier (2011).

The regularization term for the PSF, which introduces the myopic criterion into equation 3.6, assumes that the PSF is a multidimensional Gaussian random variable. This assumption is justified when one deals with long exposures, which are, by definition, sums of large numbers of short-exposures. As such they are Gaussian according to the central limit theorem. Adopting these conditions, $J_h(\mathbf{h})$ has the form

$$J_h(\mathbf{h}) = \frac{1}{2}\sum_f \frac{|H(f) - H_m(f)|^2}{E[|H(f) - H_m(f)|]^2}\,, \tag{3.9}$$

This prior is expressed in the Fourier domain whereby upper-case notation denotes Fourier transformation. The term $H_m(f) = E[H]$ is the mean transfer function and $E[|H(f)-H_m(f)|]^2$ is the associated power spectral density (PSD) with f denoting spatial frequency. This Fourier-domain prior bears some resemblance to equation 3.7. Indeed, it also assumes Gaussian statistics and draws the solution, in the least-squares fashion, toward the user-supplied mean PSF while obeying the error bars given by the PSD (which give the variance at each frequency). The PSF prior leads to band-limitedness of the PSF estimate because the ensemble average in the denominator, $E[|H(f) - H_m(f)|]^2$, should be zero above the diffraction-imposed cut-off.

In practice, equation 3.9 relies on the availability of several PSF measurements. The mean PSF and its PSD are estimated by replacing the expected values ($E[.]$) by an average computed on a PSF sample. When such a sample of several PSFs is not available, as we have assumed in our work, then $\mathbf{H_m}$ is made equal to the Fourier transform of the single supplied PSF, and $E[|H(f) - H_m(f)|]^2$ is computed as the circular mean of $|\mathbf{H_m}|^\mathbf{2}$. These relatively large error bars are intentional: they account for the lack of knowledge about the PSD when given only a single PSF measurement.

In the original MISTRAL paper (Mugnier et al. 2004) the code was presented mainly in the context of planetary images, for which the object prior (equation 3.8) was developed. The experimental data presented therein were obtained on several AO systems and covered a wide range of celestial objects such as Jupiters satellites Io and Ganymede and the planets Neptune and Uranus. MISTRAL was applied to the study of the asteroids Vesta (Zellner et al. 2005) and 216-Kleopatra (Hestroffer et al. 2002), and was also used to monitor surface variations on Pluto over a 20-year period (Storrs & Eney 2010).

3.2.4 IDAC

Multiframe blind deconvolution (MFBD) (Schulz 1993; Jefferies & Christou 1993) is an image reconstruction method relying on the availability of several images of an object. In addition, many of the MFBD algorithms rely on short exposures. This was originally dictated by the notion that in imaging through turbulence short-exposure images contain diffraction-limited information while the long exposures do not (Labeyrie 1970). Before the advent of AO the only way to obtain diffraction-limited data from the ground was to record short exposures that could then be processed by one of several "speckle imaging" methods (Knox & Thompson 1974; Lohmann et al. 1983). Therefore, MFBD was originally proposed in the context of speckle imaging.

The MFBD code we used in this paper is called IDAC, for iterative deconvolution algorithm in C [2], and is an extension of the iterative blind deconvolution (IBD) algorithm proposed by Lane (1992). IDAC performs deconvolution by numerically minimizing a functional that is composed of four constraints:

$$\epsilon = E_{im} + E_{conv} + E_{bl} + E_{Fm} \;, \tag{3.10}$$

where

$$E_{im} = \sum_{r\in\gamma} [\hat{o}(r)]^2 + \sum_{i=1}^{M} \sum_{r\in\gamma} [\hat{h}_i(r)]^2 \tag{3.11}$$

is the image domain error that penalizes the presence of negative pixels (γ) in both the object (**o**) and the PSF (**h**) estimates. The subscript i refers to an individual data frame.

The so-called convolution error is

$$E_{conv} = \frac{1}{N^2} \sum_{i=1}^{M} \sum_{f} |I_i(f) - \hat{O}(f)\hat{H}_i(f)|^2 B_i(f) \;, \tag{3.12}$$

which quantifies the fidelity of the reconstruction ($\hat{\mathbf{O}}$) to the data (**I**) in the Fourier domain. The term B_i is a binary mask that penalizes frequencies beyond the diffraction-imposed cut-off, N^2 is a normalization constant where N is the number of pixels in the image.

The third constraint is called the PSF band-limit error and is defined as

$$E_{bl} = \frac{1}{N^2} \sum_{i=1}^{M} \sum_{f} |\hat{H}_i(f)|^2 B_i'(f) \;, \tag{3.13}$$

It prevents the PSF estimates from converging to a δ function and the object estimate from converging to the observed data. The term B_i' is a binary mask that is unity for spatial frequencies greater than 1.39 times the cut-off frequency and zero elsewhere.

The last constraint is the Fourier modulus error:

$$E_{Fm} = \frac{1}{N^2} \sum_{f} [|\hat{O}(f)| - |O_e(f)|]^2 \Phi(f) \;, \tag{3.14}$$

where O_e is a first estimate of the Fourier modulus of the object obtained, e.g., from Labeyrie's speckle interferometry method (Labeyrie 1970), and Φ is a signal-to-noise (SNR) filter. Jefferies & Christou

---

[2] http://cfao.ucolick.org/software/idac/

(1993) showed with several simulations that this constraint is especially important since it incorporates relatively high SNR information from a complete dataset formed by many frames. On the other hand, the Fourier modulus of an object is only recoverable directly if one has a set of speckle images of an unresolved source to be used as a reference. There are workarounds to this problem, most notably reference-less approaches reported in Worden et al. (1977) and von der Lühe (1984), but these solutions are applicable only to non-compensated imaging and could not be used in our tests.

While MFBD algorithms were initially proposed in the context of speckle imaging, there is nothing preventing their application to long-exposure AO images that now contain diffraction-limited frequencies. In principle, there are some inherent advantages of working with M frames instead of using a single, co-added long exposure. The multiframe approach reduces the ratio of unknowns to measurements from $2:1$ in single-image blind deconvolution to $M+1:M$ in multi-frame deconvolution. On top of the PSF band-limit constraint (equation 3.13), concurrent processing of many frames means that the PSF cannot converge to the δ function (this would have yielded an object equal to the data, but the data are generally temporally variable while the object is assumed to be constant in MFBD, which is a good assumption in the context of astronomical imaging on short time-scales). MFBD algorithms are very successful in the case of strongly varying PSFs so that the target is easily distinguished from the PSFs. On the other hand, the goal of AO is to stabilize the PSF. This implies less PSF diversity from one observation to another, so that other constraints become more useful.

The code IDAC can be regarded as a precursor but also a representative of a wider class of MFBD algorithms. We mention here the PCID code (Matson et al. 2009), which has the capability to estimate the PSFs either pixel by pixel in the image domain or in terms of a Zernike-based expansion of the phase in the pupil of the telescope. It has been shown that such a PSF re-parameterization leads to object estimates that are less noisy and have a higher spatial resolution (Matson & Haji 2007).

IDAC was applied to speckle observations of the binary system Gliese 914, and in the process, the secondary component was resolved into two stars (Jefferies & Christou 1993). Additionally, IDAC was one of the five codes used in our study of photometric accuracy of image reconstruction algorithms (Gladysz et al. 2010a).

3.3 Dataset description and methodology

3.3.1 Dataset description

The images we used to test the algorithms correspond to observations performed with the fourth detector of the WFPC2 camera (Trauger et al. 1994) (the so-called planetary camera -PC-, with a pixel size of $0.046''$) installed on the HST. These are pictures of Saturn with a dynamic range of up to 975 counts, and galaxy M100 with a peak signal of 7400 counts. To obtain well-defined edges and transitions between the object and the background, the image of Saturn was preprocessed so that all pixels below a certain threshold were set to zero, thus enhancing the visibility of the Cassini and Encke divisions. These images were considered representations of the true objects (figure 3.1).

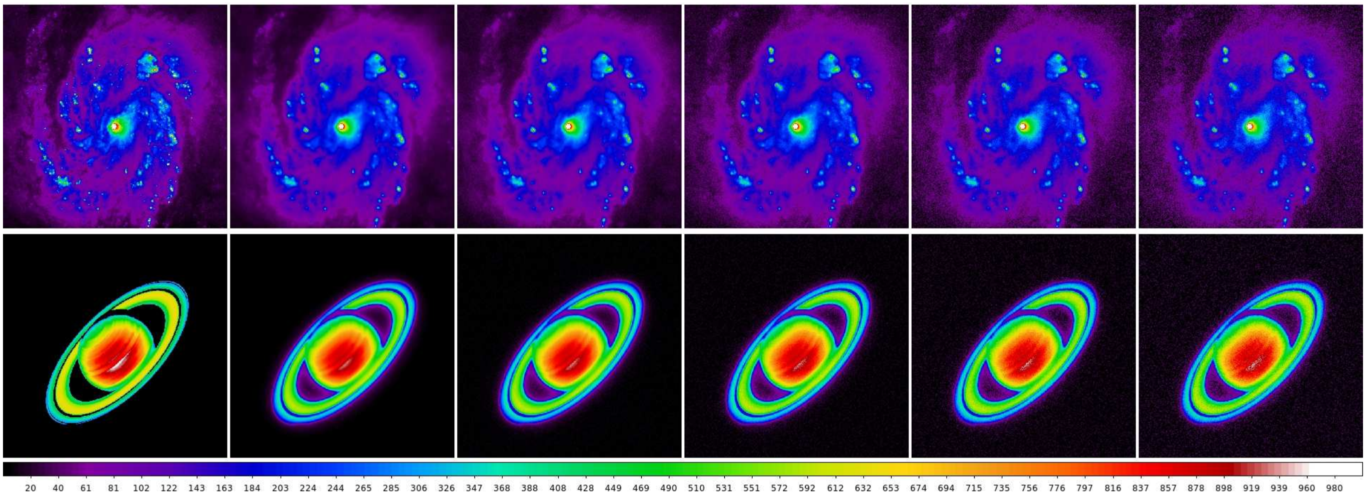

Figure 3.1: Top, far left: M100 galaxy "ground truth" image. Bottom, far left: Saturn "ground truth" image. From left to right: blurred simulated images using a science PSF with a Strehl ratio = 53% (far left panel in figure 3.2) plus additive Gaussian noise with standard deviations equal to 1, 5, 10, 15, and 20. Linear scale.

The "ground truth" images were subsequently convolved with an AO PSF with an SR equal to 53% (see Table 3.1). This and the other PSFs used in this project were obtained with the 1024x1024 PHARO infrared camera (Hayward et al. 2001) on the 5-m Hale telescope at the Palomar Observatory (Troy et al. 2000). Closed-loop images of single stars were recorded using the $0.040'' pixel^{-1}$ mode. The images were cropped to size 150×150 pixels, which corresponds to a field of view (FOV) of 6″. The observations were acquired in the K band (2.2μm) where the diffraction limit is 0.086″ so that the data meet Nyquist-sampling requirements. The filter of the observations was Brackett Gamma (BrG) and each of the PSFs in Table 3.1 corresponds to a sum of 200 frames, each with an exposure time of either 1416 or 2832 ms. The individual frames were registered via iterative Fourier shifting to produce shift-and-add images (Baena Gallé & Gladysz 2011).

Table 3.1: PSFs used for the simulated observations

Star	SR	exp. time (ms)	Type
PSF1	53%	1416	Science PSF
PSF2	53%	2832	Reference PSF
PSF3	45%	2832	Reference PSF
PSF4	36%	1416	Reference PSF

The angular size on the sky over which the AO PSF can be assumed to be almost spatially invariant is the so-called isoplanatic angle. This parameter becomes larger at longer wavelengths. As specified by Hayward et al. (2001), the isoplanatic angle is approximately 50″ in K-band at Palomar. The angular size of the Saturn and M100 images is 20.5″, so we can assume that the PSF remains constant throughout the FOV. The difference in pixel scales between those images and the PSFs from Palomar is very small (0.046″ vs. 0.040″) and therefore we did not re-bin the images to match their pixel scales. The paper is devoted comparing image-restoration algorithms and not to the performance evaluation of the AO system at Palomar.

We used IDAC in the multi-frame mode. One of the goals of our work was to check the trade-off between the diversity provided by more frames vs. SNR per frame. One can think of this as an exposure-time optimization. For a constant total observation time per object one can use shorter exposures and hope to exploit the PSF variability in subsequent image restoration with MFBD, or opt instead to use fewer images with a better SNR per frame. Therefore, out of the original 200-frame dataset we have produced datasets of 10, 20, 50, and 100 binned PSFs. These binned PSFs, together with the original 200-frame dataset, were used as blurring kernels for the Saturn image. For AWMLE, ACMLE, and MISTRAL, which are single-image restoration codes, we only used the summed PSF. All PSFs were normalized to have a total power equal to unity before the convolution procedure.

The blurred observations were subsequently corrupted with noise. We explored more than twenty levels of noise: pure Gaussian readout noise of standard deviation ranging from $\sigma = 1$ to $\sigma = 20$, and shot noise plus readout noise of level $\sigma = 10$. These levels correspond to the noise that was added to the summed images. For MFBD with IDAC we worked with several images and decided to add to each frame the amount of noise that would result in the same SNR per summed image had the images been summed and *then* processed, as in the case of the other three algorithms. This means adding pure Gaussian noise with levels from $1 \cdot \sqrt{50}$ to $20 \cdot \sqrt{50}$, and Poisson noise plus $10 \cdot \sqrt{50}$ to the 50-frame dataset, for example. This way we aimed to test the benefits of exploiting PSF diversity vs. higher noise per frame, as mentioned above.

One of the goals of the project was to test the susceptibility of the algorithms to mismatch between the science and calibration PSFs. Therefore, we used a matched PSF (SR=53%), a mismatched PSF (SR=45%), and a highly mismatched PSF (SR=36%) as inputs to AWMLE, ACMLE, MISTRAL, and MFBD. The last SR value corresponds to half of the maximum SR in K band (63%) predicted in simulations for the AO system in Palomar (Hayward et al. 2001). Table 3.1 lists the PSFs used in the tests. Differences in SR for the same star arise because of the changing seeing. All PSFs used in the simulations are presented in Figure 3.2.

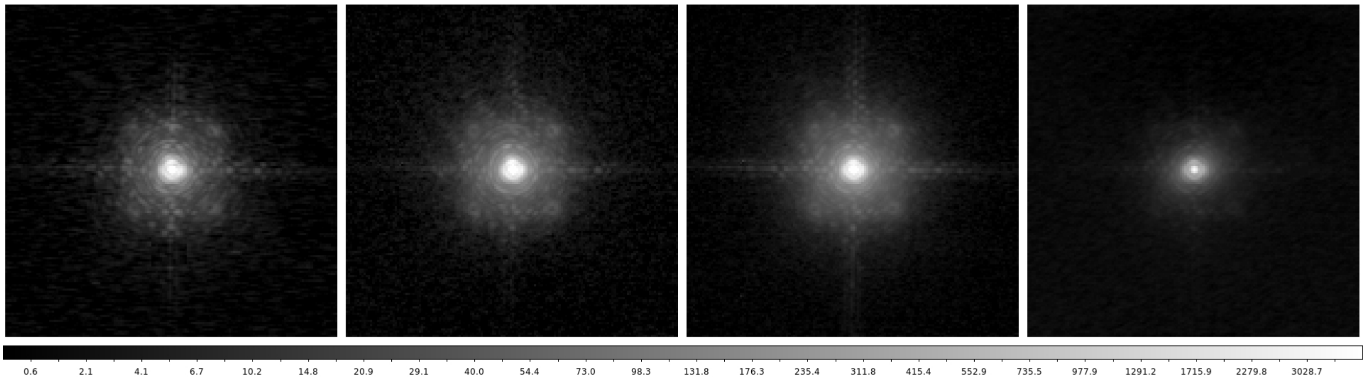

Figure 3.2: Far left: science PSF with the Strehl ratio (SR) = 53%. Middle left: reference PSF with SR = 53%. Middle right: reference PSF with SR = 45%. Far right: reference PSF with SR = 36%. The far left PSF was used to blur the HST images of Saturn and M100. The other three PSFs were used as reference PSFs for the algorithms. Logarithmic scale.

3.3.2 Methodology

The four algorithms are very different and require different strategies to obtain the final result. For AWMLE and ACMLE, it is convenient (although not mandatory) to perform an initial decomposition of the dataset, either in the wavelet or in the curvelet domain. This can give the user an idea as to the number of planes to use during the deconvolution process. Additionally, this preliminary decomposition could inform the user whether it is sensible to perform deconvolution in the highest-frequency plane where the finest structures together with noise have been classified. When working in these planes there is a trade-off between the benefit of recovering information from the finest wavelet/curvelet plane and the undesirable reconstruction of non-significant structures. Both transforms, WT and CT, can also be combined into a dictionary of coefficients (Fadili & Starck 2006), although for the sake of simplicity we executed them independently.

The main difficulty when using a Richardson-Lucy-type algorithm is determining at which iteration to stop the deconvolution process. In AWMLE and ACMLE, this problem is solved by probabilistic masks. The mask can apply significance thresholds to a given location in a given plane to selectively deconvolve statistically similar regions. This concept is called multiresolution support (Starck et al. 2002). Probabilistic masks are used to stop the deconvolution process automatically in parts of the image where significant structures cannot be discerned. To estimate the value of the mask at a particular location, a local window must be defined. Here we used a window based on the local standard deviation computed within that window:

$$m_\sigma = \begin{cases} 1 - exp\left\{\dfrac{-A^2 \cdot (\sigma_i - \sigma_\chi)^2}{2\sigma_\chi^2}\right\} & if\ \sigma_i - \sigma_\chi > 0 \\ 0 & if\ \sigma_i - \sigma_\chi \leq 0\,, \end{cases} \tag{3.15}$$

with

$$\sigma_i = \sqrt{\frac{\sum_{p\in\Phi}(\chi_{j,p})^2}{n_f}}\,,$$

where A is a coefficient for determining the applied threshold, σ_i is the standard deviation within the window $\mathbf{\Phi}$ centered on pixel i of the plane χ_j (pixels within that window are indexed with p), n_f is the number of pixels contained in the window, and σ_χ is the global standard deviation in the corresponding wavelet or curvelet plane. The value for σ_χ can be determined by decomposing an artificial Gaussian noise image, with standard deviation equal to that in the dataset, into wavelet or curvelet coefficients. These windows are then shifted across all wavelet or curvelet planes to build the final probabilistic mask **M** (see equations 3.3 and 3.5).

Each of the wavelet planes has the same size as the original image (i.e., 512x512). For cases like this, Otazu (2001) proposed to increase the size of the local window m_σ with the wavelet scale. The lower the frequency of the representation, the larger the window. Therefore, the window size for the highest-frequency scale was set to 5*x*5, next plane was analysed with a 7*x*7 window, and the third one with a window of size 11*x*11. On the other hand, the curvelet transform decreases the scale, or wedge size, with the scale of representation. For instance, wedges at the highest-frequency scale have a size of 267*x*171 pixels, whereas wedges at a lower-frequency scale can have dimensions of 71*x*87, according to the level of detail being represented. The ratio between dimensions in a given wavelet plane and the corresponding curvelet wedge is between 2 and 3. To closely compare between WT and CT, we kept the same ratio for the window sizes used at each scale for both ways of representation, and a constant window size (3*x*3 pixels) was chosen for the CT. The residual wavelet and curvelet scales were never thresholded, i.e., masks were set to 1 for these scales. Finally, parameter A was set to $3/2$ so the window was approximately equal to 1 when $\sigma_i - \sigma_\chi = 2\sigma_\chi$, i.e., when coefficients above twice the noise level are detected.

In our experiments we noticed that the results could be improved by changing the values of A and the window sizes, especially for CT. Improvement means here fewer elongated artifacts. On the other hand, for the purpose of a clear and fair comparison between WT and CT, we preferred to keep the values given above, bearing in mind the original decision adopted for AWMLE in Otazu (2001). In our opinion, a more thorough study of the best choices for the parameters of the probabilistic masks is necessary, and in particular, a study in which choices are better suited for a particular way of representation.

The CurveLab software, which was used here to introduce the CT into ACMLE, can perform digital curvelet decomposition via two different implementations. The first one is the so-called USFFT digital CT and the second one is known as the digital CT via wrapping. Both differ in the way they handle the grid that is used to calculate the FFT to obtain the curvelet coefficients. This grid is not defined in typical Cartesian coordinates, instead, it emulates a polar representation, more suited to the mathematical framework defined in Candès et al. (2006). The USFFT version has the drawback of being computationally more intensive than to the wrapping version, since the latter makes a simpler choice of the grid to compute the curvelets (Starck et al. 2010). Hence, for reasons of computational efficiency, the wrapping version was used in ACMLE.

The user has to provide CurveLab with the values of some parameters. In addition to the number of curvelet scales, the number of orientations or angles of representation in the second scale is also required as input from the user. This parameter will automatically set the number of angles for the rest of the scales. Evidently, the higher the number of orientations, the longer the algorithm will need to perform CT, the higher the overall redundancy and the higher the computational cost for calculating the probabilistic masks at each scale. We decided to set this parameter to 16 as a trade-off between having a complete representation of all possible orientations in the image and the total execution time. Finally, a third parameter decides if the curvelet transform is replaced by an orthogonal wavelet representation at the final scale, i.e., the highest-frequency scale. In our experiments we saw no significant differences between results based on one choice for this parameter's value or the other.

MISTRAL performs the minimization process of equation 3.6 by the partial conjugate-gradient method. The code requires that the user provides values for the hyperparameters μ and δ (see equation 3.8), which balance the smoothing imposed by too much regularization against noise amplification. This has to be done by trial and error although the authors of the algorithm provide some suggestions in the original MISTRAL paper, namely that μ be set to around unity and δ be set to the norm of the image gradient ($\|\nabla i\| = [\sum_p |\nabla i(p)|^2]^{1/2}$). In practice, the user experiments with various values of these parameters, centered on the values suggested in Mugnier et al. (2004), and chooses the image reconstruction that is most visually appealing. In our tests we found that $\mu = 10$ and $\delta = 2$ yielded the best results. We let the code run for the maximum number of iterations set to 10^3.

IDAC requires the user to provide the value of the diffraction-limit cut-off. Caution should be taken here: the code requires that all supplied images are of the same dimensions. Therefore, when one has a PSF image that is smaller in size than the target image, the PSF should be embedded in an array of zeros. Subsequently, the diffraction-limit cut-off should be estimated from the zero-padded, and not the original Fourier-transformed PSF (see Figure 3.3). Another parameter that should in theory affect the restorations, a scalar quantifying user confidence in the supplied PSF, was found to have negligible effect on the outputs.

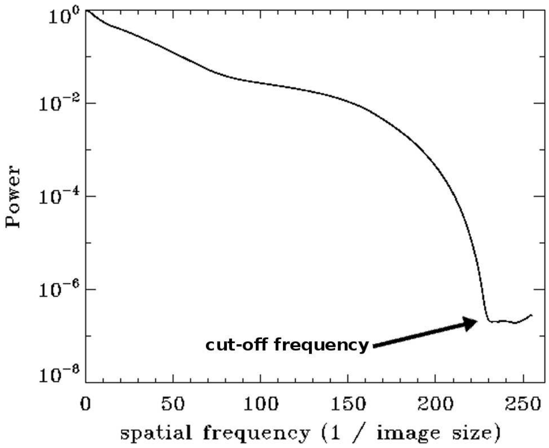

Figure 3.3: Normalized power spectrum of the 53% SR PSF (Fig. 3.2 middle left). The diffraction- and noise-limited cut-off frequency was determined to be approximately 230 pixels. The PSF was embedded in an array of zeros to match the size of the science image (512×512 pixels).

As with MISTRAL, we let the code run for the maximum number of iterations set to 10^3. For images with a high level of noise (Gaussian $\sigma \geq 10$ and Poisson noise plus Gaussian $\sigma = 10$) it converged quickly, after 15-20 iterations, which can be considered a good behavior because the noise did not become amplified. For cases with low noise ($\sigma \leq 5$) the algorithm converged after 50-100 iterations and produced generally sharper reconstructions.

3.4 Quality metrics and maps

One of the goals of this work was to compare the quality of reconstructions yielded by several codes. Therefore, metrics were needed for image quality assessment. Here we have chosen the mean squared error (MSE), which is probably the most commonly used metric in image processing (e.g. Mugnier et al. 2004) and one of the oldest and most often used metrics to evaluate the resemblance of two images. It is defined as

$$\mathrm{MSE} = \frac{1}{M}\sum_{j=1}^{M}(x_j - y_j)^2 \, , \tag{3.16}$$

where, in this context, M is the total number of pixels in the image and x and y are pixel values of the two images that are compared. It is also very common to use the related metric, the peak signal-to-noise ratio (PSNR), defined as

$$\mathrm{PSNR} = 10\log_{10}\frac{L^2}{MSE} \, , \tag{3.17}$$

where L is the dynamic range of the image. These metrics are very easy to compute and measure the pixel-by-pixel departure between the reconstruction and the reference object. The MSE and PSNR have a clear physical meaning: they quantify the energy of the error. Hence, they are well suited to the task of estimating the absolute photometric error between the two images.

We preferred to use PSNR rather than MSE to quantify differences between images, since the former has the more intuitive behavior of setting higher values to better reconstructions. We set $L = 975$ in equation 3.17 to match the dynamic range of the "ground truth" Saturn object, and $L = 7400$ to match the maximum value present in the original image of M100, which corresponds to the galaxy core. A difference in 1dB in PSNR implies an approximate difference between 150 and 250 in MSE depending on the object.

Below we also plot the so-called error and residual maps. When one obtains a certain reconstruction after solving the main image formation equation (eq. 3.2)

$$\mathbf{i} = \mathbf{h} * \mathbf{o} + \mathbf{n} \longrightarrow \hat{\mathbf{o}} \, , \tag{3.18}$$

where $\hat{\mathbf{o}}$ is the reconstruction or solution found with a certain algorithm, the error map is the exact difference between the "ground truth" object and the reconstruction, i.e., $\mathbf{e} = \mathbf{o} - \hat{\mathbf{o}}$. It offers a detailed, pixel-by-pixel or structure-by-structure, view of the algorithm's performance. Unfortunately, in practical situations the real object is not available, so the reconstruction error measurement is usually made in the image domain by transferring the solution $\hat{\mathbf{o}}$ into the image domain by means of the PSF $\mathbf{h}$, and subtracting it from the data. Hence, the normalized residual map is computed as

$$\mathbf{r} = \frac{(\mathbf{h} * \hat{\mathbf{o}} - \mathbf{i})^2}{\mathbf{h} * \hat{\mathbf{o}}} \, , \tag{3.19}$$

In the case of static-PSF approaches, $\mathbf{h}$ is the input calibrator, whereas for myopic- and blind- approaches $\mathbf{h}$ is the PSF estimate found in the last iteration. The residual map is a graphical comparison in the image domain meant to show how close, from a mathematical point of view, the reconstruction and the data are to each other. On the other hand, the error map is a more truthful comparison, from a physical point of view, of the solution and the (usually unknown) reality.

3.5 Results and discussion

The tests were conducted decomposing Saturn and M100 datasets into three wavelet/curvelet planes plus a wavelet/curvelet residual. As mentioned before, three levels of mismatch between the science and calibration PSF were analyzed, as well as more than twenty levels of noise, including readout and shot noise. All outputs were analyzed by means of the PSNR metric.

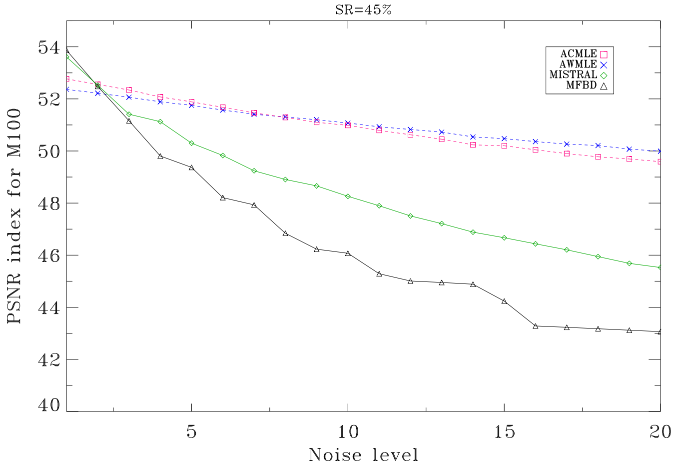

Figure 3.4: PSNR evolution of M100 reconstructions with respect to the level of noise. Blurred simulated images were corrupted with Gaussian noise with standard deviations ranging from $\sigma = 1$ to $\sigma = 20$. All reconstructions used the SR=45% PSF as the calibrator. Pink squares: ACMLE. Blue crosses: AWMLE. Green diamonds: MISTRAL. Black triangles: MFBD.

Figures 3.4 and 3.5 show the evolution of the PSNR metric with respect to the level of noise when a mismatch of $\sim$ 8% in SR exits between the science and the reference PSF. Apart from the logical conclusion that reconstruction quality decreases as the noise increases, we can point out that firstly, despite of being designed for detecting elongated structures, ACMLE performs reasonably well in low-noise-level conditions for an image where many point-like sources are present such as M100 (Fig. 3.4). Only when the noise level reaches a value of $\sigma \sim 8$ and the correlation of information along lines and edges has significantly deteriorated, its performance becomes comparable to that of AWMLE. This effect does not happen for the Saturn image (Fig. 3.5) where almost all information is distributed along elongated features and the curvelet transform is more resistant against noise. Secondly, myopic- and blind-PSF algorithms show a more evident deterioration of performance with the increasing level of noise. For MISTRAL, its performance for low-level-noise conditions is extremely good for Saturn, whereas it becomes similar to static-PSF algorithms with noise control at a level of $\sigma = 12$. This is basically due to its good behavior at the edges of the planet thanks to the $L_2 - L_1$ prior included in the code. The rest of algorithms exhibit evident Gibbs oscillations in the high transitions between the planet limits and the background. When these edges were not considered in the computation of PSNR, MISTRAL's performance evolution for Saturn with respect to other algorithms has a similar behavior to that shown for galaxy M100 in figure 3.4. It has to be mentioned, though, that MISTRAL was developed specifically in the context of imaging planetary-type bodies and its prior is not very suitable for galaxies or stellar fields, for which a white spatial prior, which assumes independency between pixels, is more advisable (Ygouf et al. 2013). On the other hand, including these regularization terms could potentially improve results obtained with MFBD and AW(C)MLE.

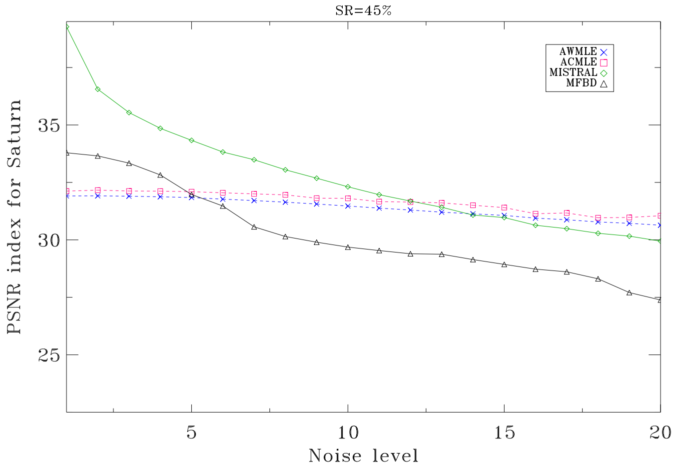

Figure 3.5: PSNR evolution of Saturn reconstructions with respect to the level of noise. Blurred simulated images were corrupted with Gaussian noise with standard deviations ranging from $\sigma = 1$ to $\sigma = 20$. All reconstructions used the SR=45% PSF as the calibrator. Pink squares: ACMLE. Blue crosses: AWMLE. Green diamonds: MISTRAL. Black triangles: MFBD.

As mentioned in Section 3.3.1, datasets of 10, 20, 50, and 100 binned images, together with the original 200-frame dataset, were used as input for IDAC to test the trade-off between the (implied) PSF diversity and SNR per frame. We tried to shed some light on the question which of the following strategies is better: deconvolving a long-exposure image with a relatively high SNR (AWMLE, ACMLE and MISTRAL algorithms), or tackling the problem by dividing the dataset into more (diverse) frames at the expense of reducing the SNR for each frame (MFBD). It was found that the original 200-frame set yielded significantly poorer results than the smaller sets. Specifically, the 50-frame dataset proved to be the best input to IDAC although the differences between its output and those of the 10- and 20-frame sets were small. For the data sets we used, there is no evidence that the higher diversity provided by more frames yields better results. Therefore, we conclude that (assuming constant total observation time) using very short frame exposure times is not the best solution for MFBD because high frame diversity does not offset the high level of relative noise per frame. Figures 3.4 and 3.5 show that MFBD behaves in a similar way as the other algorithms for low-noise-level conditions, but its performance is highly affected as the noise per frame increases.

Figures 3.6 and 3.7 reveal how the algorithms behave with respect to the mismatch between the calibrator and the science PSFs for different levels of noise ($\sigma = 1$, 10 and 20 as well as shot noise plus read-out noise of $\sigma = 10$). In general, static-PSF approaches yield better results than myopic/blind-codes when the two PSFs agree well. For M100 the advantage of ACMLE with respect to MISTRAL is around $3-5$dB, while it is higher than 5dB with respect to IDAC for noise levels above $\sigma = 10$. For the Saturn image the results are similar except for low-noise-level conditions where $L_2 - L_1$ prior gives MISTRAL an advantage. A mismatch of ∼8% in SR does not lead to large differences in the results of the algorithms (Figs. 3.6 and 3.7, middle rows), with respect to those obtained with a well-matched reference PSF (same figures, top rows). A reference PSF with $SR = 36\%$, which implies

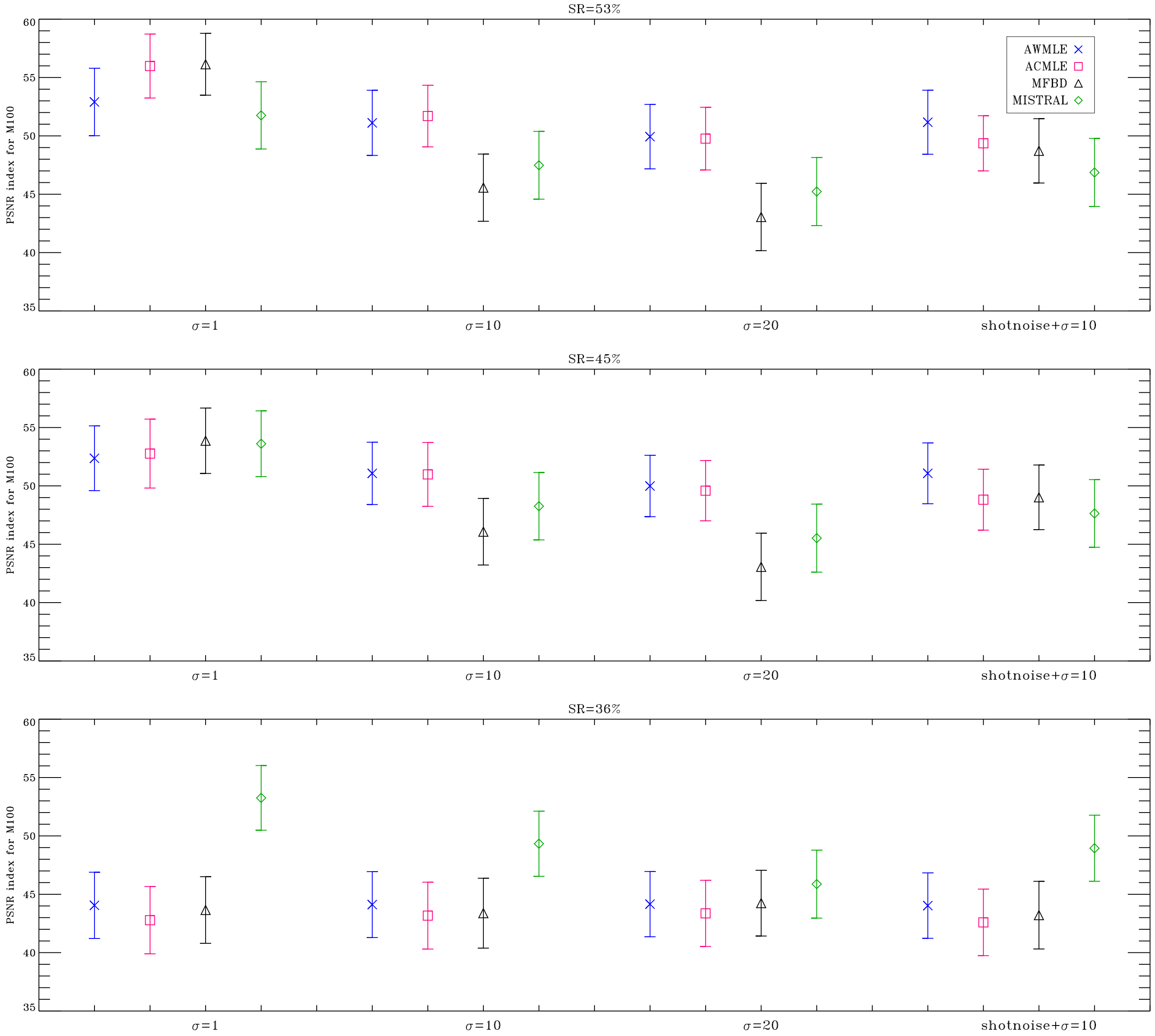

Figure 3.6: Each plot shows the mean value and standard deviation of the PSNR metric calculated from M100 results obtained with ACMLE (pink squares), AWMLE (blue crosses), MISTRAL (green diamonds), and MFBD (black triangles). Deconvolution was performed with three PSFs to represent different levels of miscalibration. Top row: dataset was deconvolved with a matched PSF (SR=53%). Middle row: dataset was deconvolved with a mismatched PSF (SR=45%) Bottom row: dataset was deconvolved with a highly mismatched PSF (SR=36%).

a mismatch of ∼ 17% SR, can be seen to have a more noticeable effect on the static-PSF codes and, surprisingly, MFBD than on MISTRAL (Figs. 3.6 and 3.7, bottom rows). The reconstruction quality for MISTRAL is more uniform for the PSF mismatch, which is logic since this algorithm is designed to deal with large differences between the science and the reference PSFs. At this level of mismatch there is a visible qualitative difference between the PSFs (Fig. 3.2, leftmost and rightmost panels), which affects the performance of static-PSF codes. We stress that, although MISTRAL is very reliable for the PSF mismatch, its advantages become highly attenuated as the noise level increases. This can be seen in figure 3.6, bottom row, noise level of $\sigma = 20$. This shows that the ability of a method to control the noise is as critical as its capability to update the PSF.

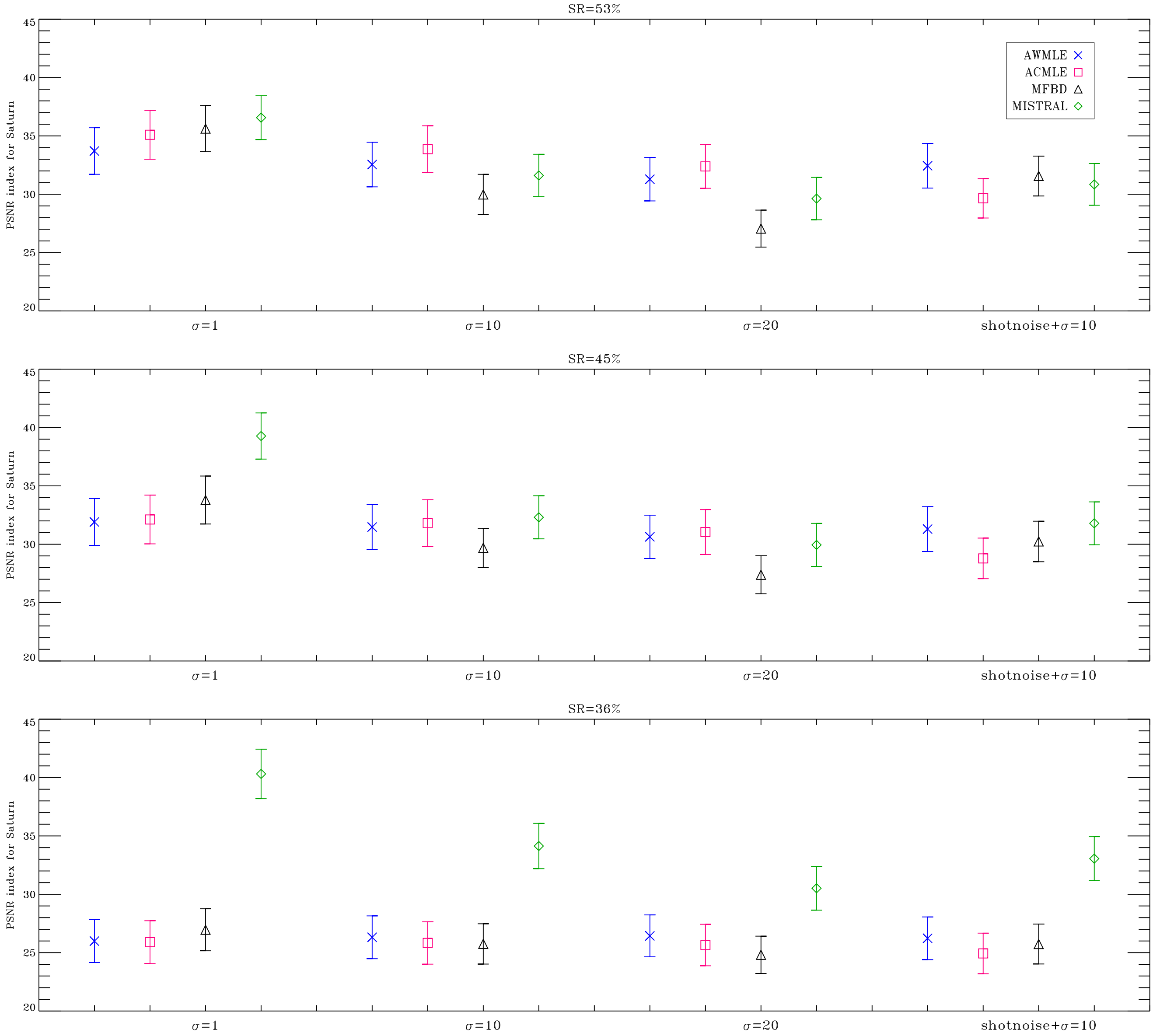

Figure 3.7: Each plot shows the mean value and standard deviation of the PSNR metric calculated from Saturn results obtained with ACMLE (pink squares), AWMLE (blue crosses), MISTRAL (green diamonds), and MFBD (black triangles). Deconvolution was performed with three PSFs to represent different levels of miscalibration. Top row: dataset was deconvolved with a matched PSF (SR=53%). Middle row: dataset was deconvolved with a mismatched PSF (SR=45%) Bottom row: dataset was deconvolved with a highly mismatched PSF (SR=36%).

Figures 3.8 and 3.9 show the reconstructions of M100 achieved by the algorithms for the noise levels of $\sigma = 5$ and $\sigma = 15$ and a mismatch of $\sim 8\%$ in SR. ACMLE yielded the sharpest reconstruction, although some elongated artifacts start to become visible at $\sigma = 15$ as a result of an incorrect identification of the coefficients. Compared to MISTRAL, both ACMLE and AWMLE are able to enhance and de-blend from the surrounding cloud more individual point-like sources. MFBD reconstructions are still almost as smooth as the starting image. Figures 3.10 and 3.11 show the results for Saturn. Here, similar conclusions can be formulated, i.e., noise amplification is better controlled by AW(C)MLE thanks to the multiresolution support. AWMLE and ACMLE exhibit the best results in the planet's body (in terms of the achieved resolution and noise attenuation), and also in the back-

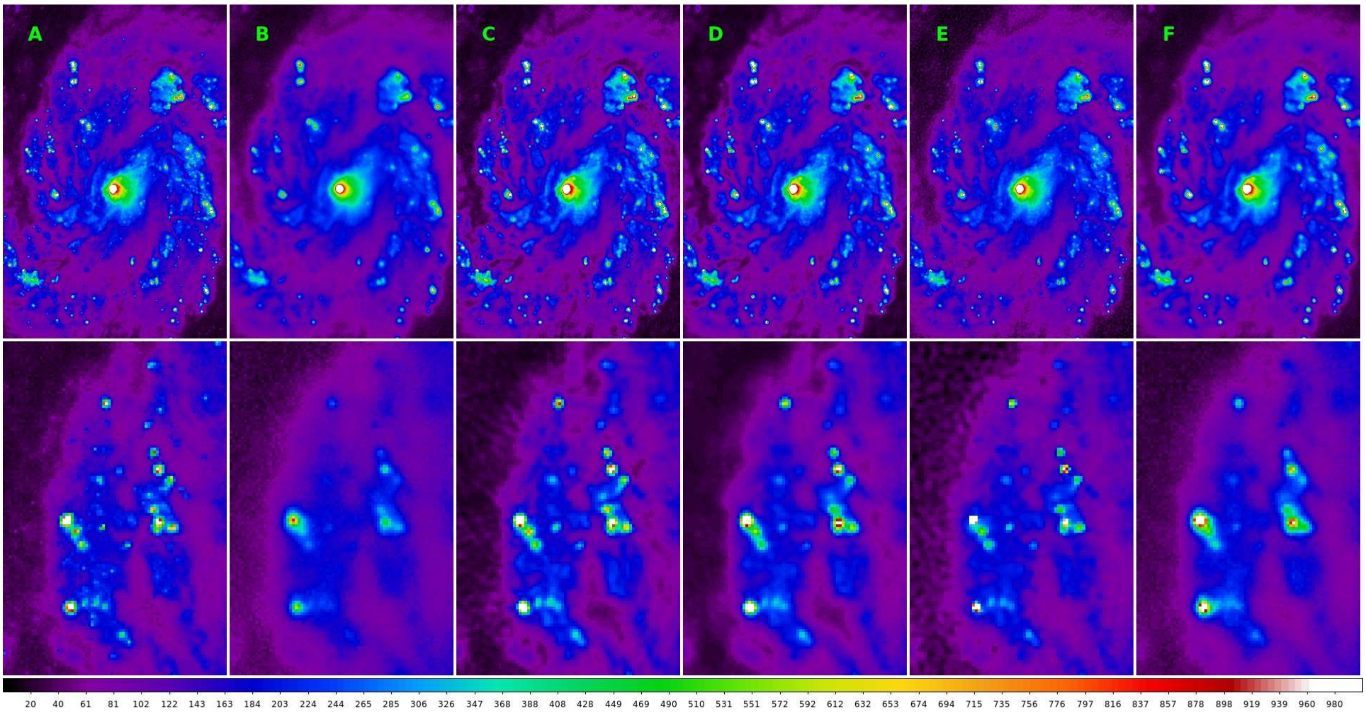

Figure 3.8: Column A: M100 original image. Column B: Degraded and corrupted image with Gaussian noise ($\sigma = 5$). Column C: ACMLE reconstruction. Column D: AWMLE. Column E: MISTRAL. Column F: MFBD. Top row: whole view. Bottom row: view of representative detail. Reconstructions were performed with a reference PSF at SR=45%. Linear scale from 0 to 1000.

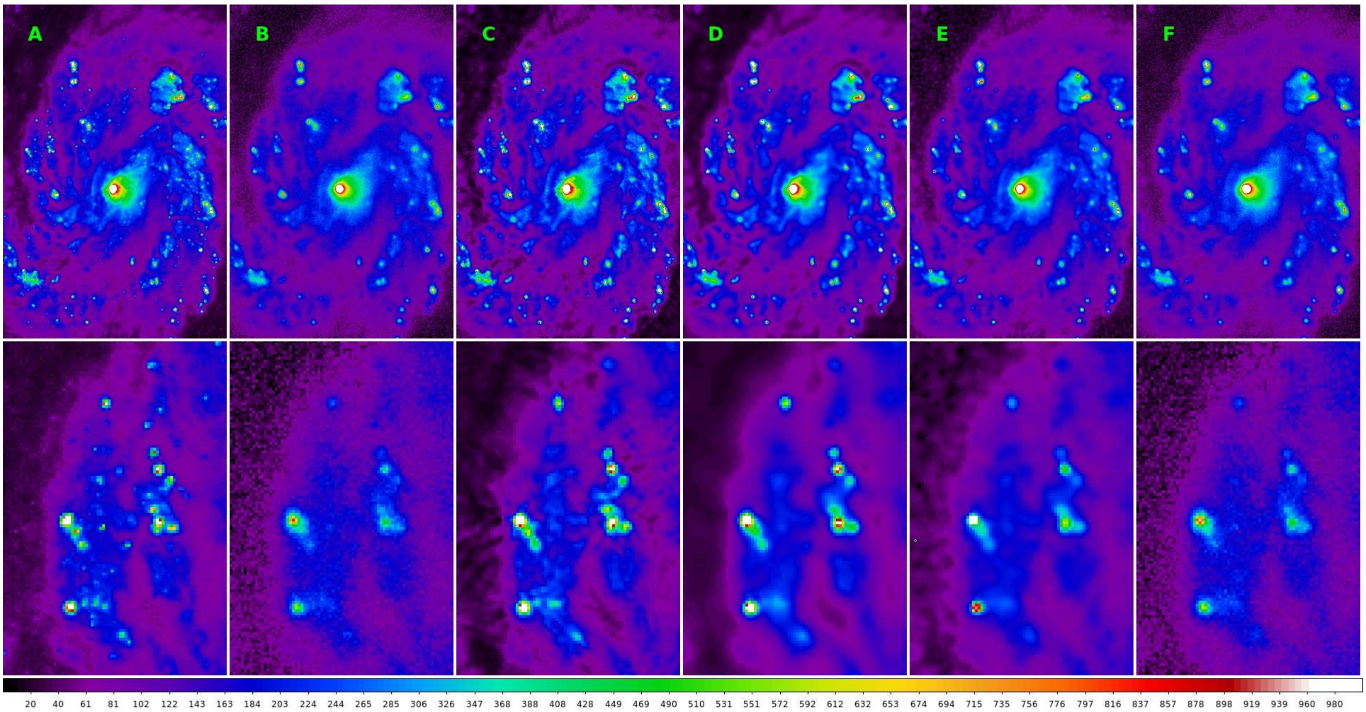

Figure 3.9: Column A: M100 original image. Column B: Degraded and corrupted image with Gaussian noise ($\sigma = 15$). Column C: ACMLE reconstruction. Column D: AWMLE. Column E: MISTRAL. Column F: MFBD. Top row: whole view. Bottom row: view of representative detail. Reconstructions were performed with a reference PSF at SR=45%. Linear scale from 0 to 1000.

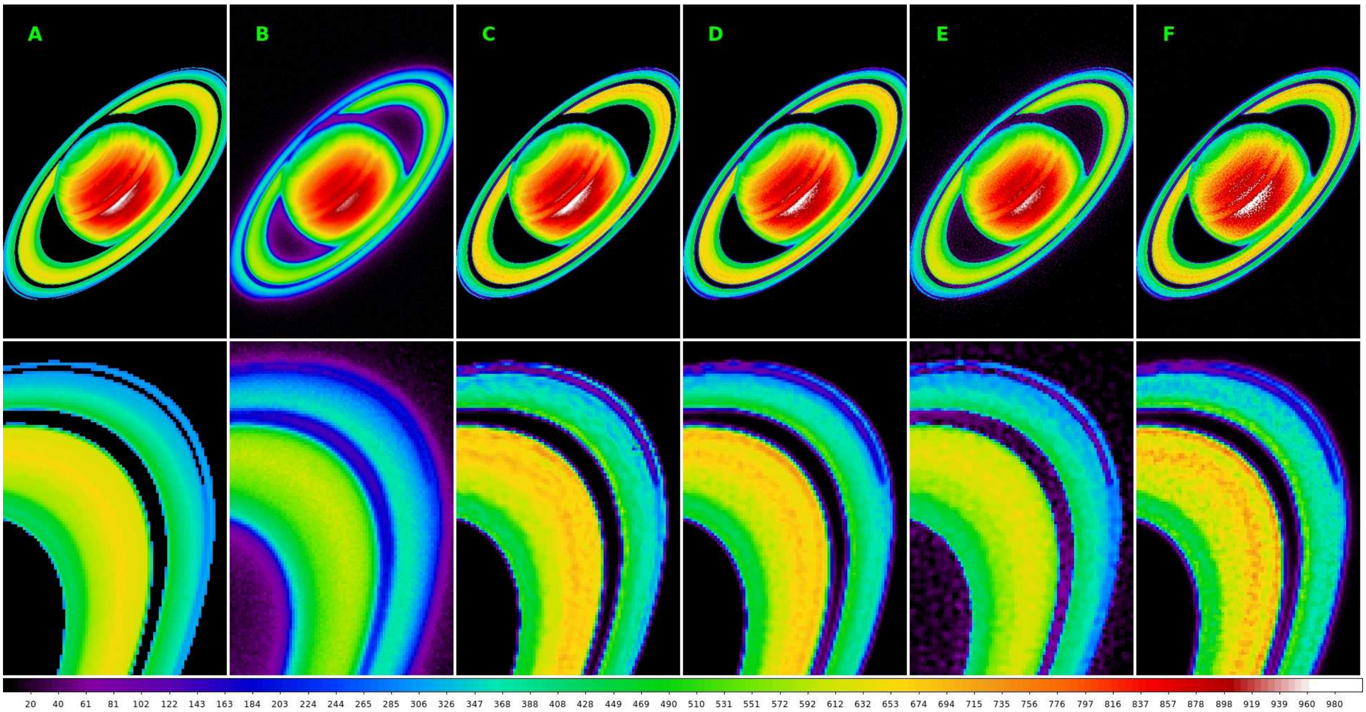

Figure 3.10: Column A: Saturn original image. Column B: Degraded and corrupted image with Gaussian noise ($\sigma = 5$). Column C: ACMLE reconstruction. Column D: AWMLE. Column E: MISTRAL. Column F: MFBD. Top row: whole view. Bottom row: view of representative detail. Reconstructions were performed with a reference PSF at SR=45%. Linear scale from 0 to 1000.

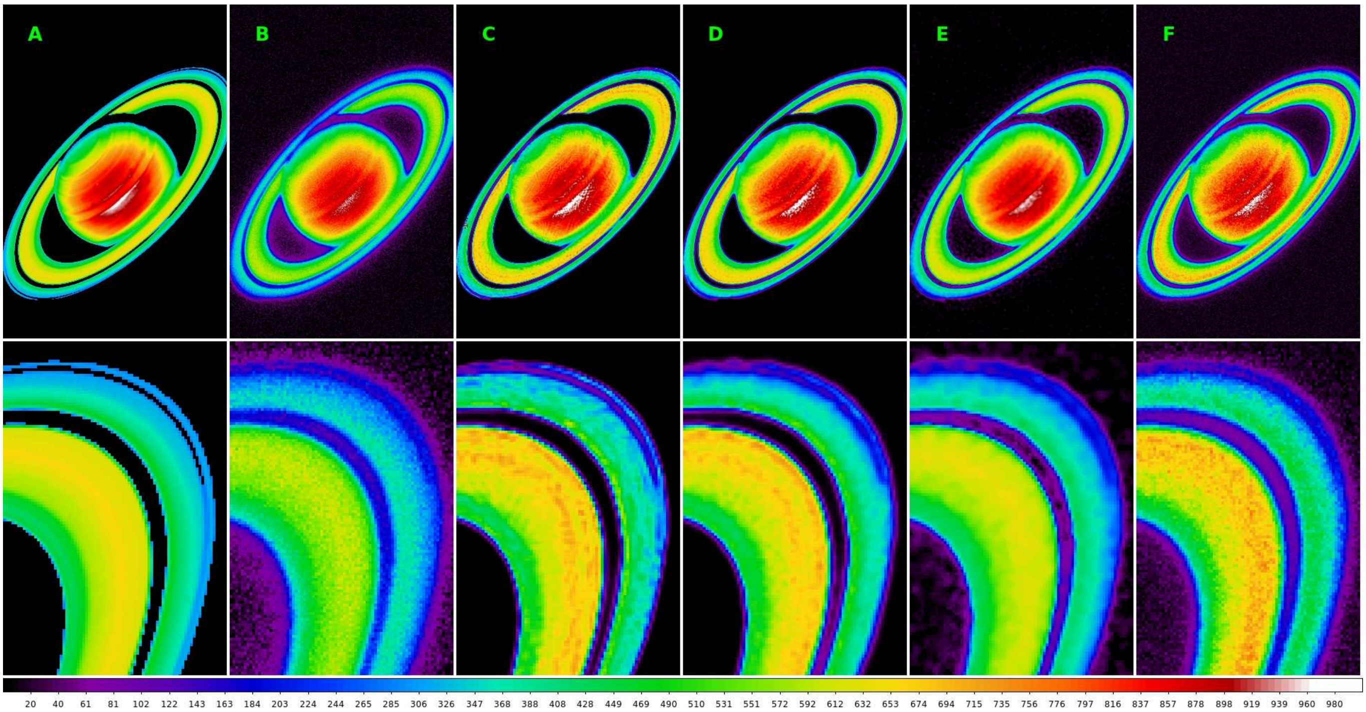

Figure 3.11: Column A: Saturn original image. Column B: Degraded and corrupted image with Gaussian noise ($\sigma = 15$). Column C: ACMLE reconstruction. Column D: AWMLE. Column E: MISTRAL. Column F: MFBD. Top row: whole view. Bottom row: view of representative detail. Reconstructions were performed with a reference PSF at SR=45%. Linear scale from 0 to 1000.

ground space between Saturn's body and the innermost ring, which is a measure of a given algorithm's ability to suppress noise. This effect is also visible in Cassini's division. Furthermore, Encke's division is barely visible in MISTRAL's and MFBD's reconstructions for high noise level of $\sigma = 15$. On the other hand, AWMLE, ACMLE, and MFBD show the typical ringing effect associated with strong transitions or edges (very visible at the limits of Saturn's body), whereas MISTRAL's $L_2 - L_1$ edge-preserving prior attenuates these effects considerably. Nevertheless, this prior must be responsible for the excessive attenuation of some elongated features present in Saturn's inner ring that were detected with ACMLE, AWMLE, and MFBD, although over-reconstructed by the former and noise-distorted by the latter two. We stress that such features are not visible in the blurred and corrupted image.

Finally, we show in figures 3.12 and 3.13 the error and residual maps. For the Saturn image, one can see that ACMLE's and AWMLE's results have values closer to the truth in all pixels of the object (planet's body, space between the body, and the innermost ring), except in the pixels at the edges of the main planet body and the rings, which show poor numbers. For these pixels, $L_2 - L_1$ object prior implemented in MISTRAL is working impressively well, while MFBD, ACMLE, and AWMLE's values strongly depart from the real object. In general, the latter three algorithms have the tendency to over-reconstruct bright sources and bright structures of the image, while MISTRAL exhibits the opposite behavior, i.e., it does not reach the correct photometric value for these regions. Noise amplification or lack of noise suppression is evident for MFBD and MISTRAL, whereas some elongated artificial structures are visible in ACMLE's result, thus showing an incorrect coefficient identification at some orientations. This suggests that the probabilistic mask should not be applied to each curvelet wedge and scale independently, as we do in this work, but creating some mechanism to exchange information among them. There are no significant differences in the residual maps obtained by ACMLE and MISTRAL (Fig. 3.13) apart of those already mentioned, although here they are not as evident as in the error maps. This suggests that updating the PSF in MISTRAL yields a mathematical solution that is compatible with the data as in the case of AW(C)MLE. However, when the noise increases, the regularization term in MISTRAL does not provide as good results as the AW(C)MLE, even if MISTRAL updates the PSF.

3.6 Conclusions

We have introduced a way of using the multiresolution support, applied in the wavelet and curvelet domain, in the post-processing of adaptive-optics images to help control the process of noise amplification. One of the most important goals of this research was to devise an objective check of the typical assumption (within the AO astronomical community) about the supposed poor performance of static-PSF approaches with respect to the blind/myopic methods. For the dataset we used, the Richardson-Lucy scheme, which is controlled over the wavelet- or curvelet domain by a σ-based mask, provided a very competitive performance against the more well-known approaches like MFBD or regularized deconvolution (MISTRAL). Specifically, it yielded 3-5dB better results (in terms of PSNR) than IDAC and MISTRAL for a mismatch in the PSF of up to 8% in terms of the Strehl ratio.

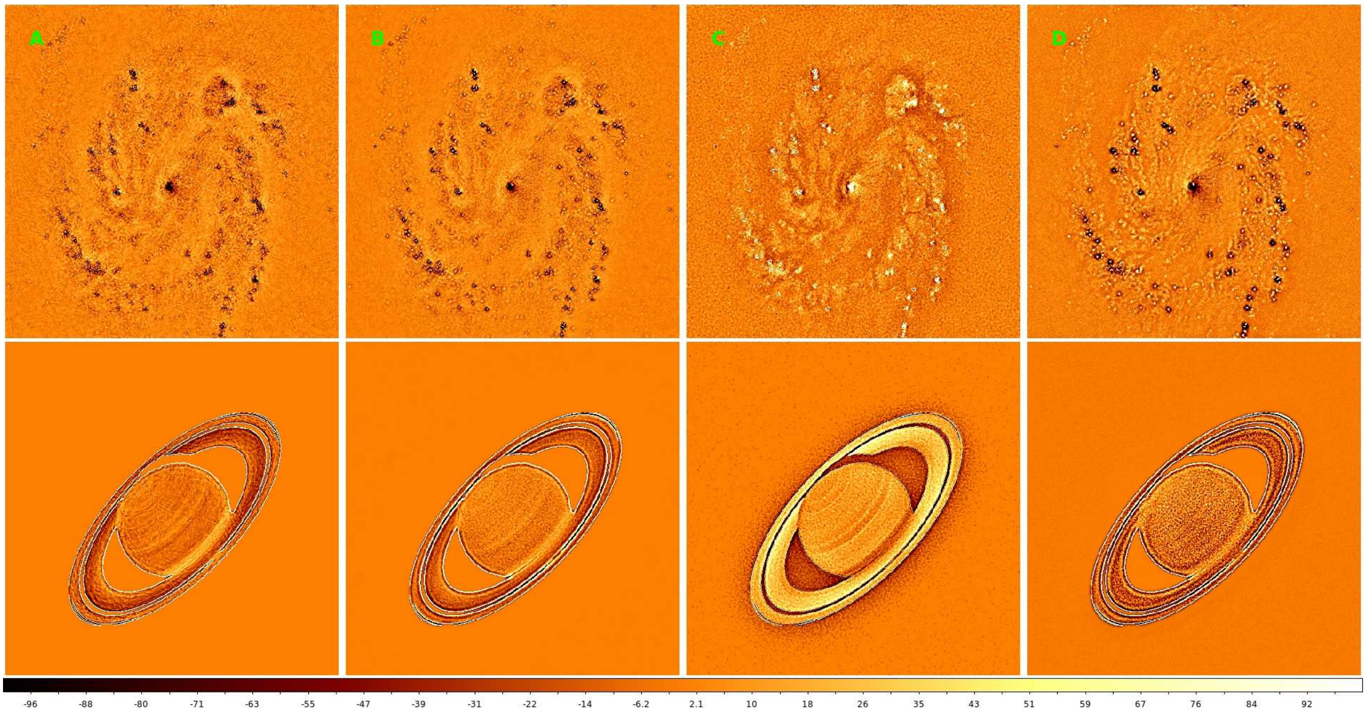

Figure 3.12: Error maps for the reconstructions obtained from data sets corrupted with Gaussian noise with $\sigma = 5$. Reconstructions were performed with a reference PSF at SR=45%. Column A: ACMLE. Column B: AWMLE. Column C: MISTRAL. Column D: MFBD. Linear scale from -100 to 100.

On the other hand, performance of MISTRAL was photometrically better for the highly mismatched PSF of SR=36% (mismatch of 17% in SR) was employed as calibrator.

We showed that probabilistic masks can control noise amplification in a Richardson-Lucy scheme, and also that curvelet transform (CT) performs better than wavelet transform (WT) for low-noise-level conditions, when elongated structures and edges still keep their bidimensional characteristics above the noise level.

The observed poor performance of IDAC, which can be visually appreciated in Figures 3.9-3.11, can be explained in the following way. Multiframe approaches rely on diverse PSFs to separate the blurring kernel from the object and to make the blind problem more tractable (less under-determined). Having more images helps when one has truly diverse PSFs. But this is not always the case with AO imaging. In fact, AO will stabilize the PSF and no number of new frames can then supply new information for MFBD. The standard deviation of the Strehl ratio value across the 200 frames used as input for IDAC was only 2%. For all datasets that we had (ten stars), the standard deviation across 200 frames rarely exceeded 5%. It seems that single-image codes like MISTRAL or AW(C)MLE have an advantage in the case of stable AO observations.

Several lines of research are still open. It would be possible to study other types of masks, such as those based on the quadratic distance between related zones in consecutive wavelet planes (Starck & Murtagh 1994) or image segmentation (or wavelet plane segmentation) by means of neural networks to link and classify significant zones in the image (Núñez & Llacer 1998). Furthermore, the use of other wavelet transforms, such as the undecimated Mallat trasform with three directions per scale, and many other multitransforms should be studied and compared, e.g., shearlets (Guo & Labate

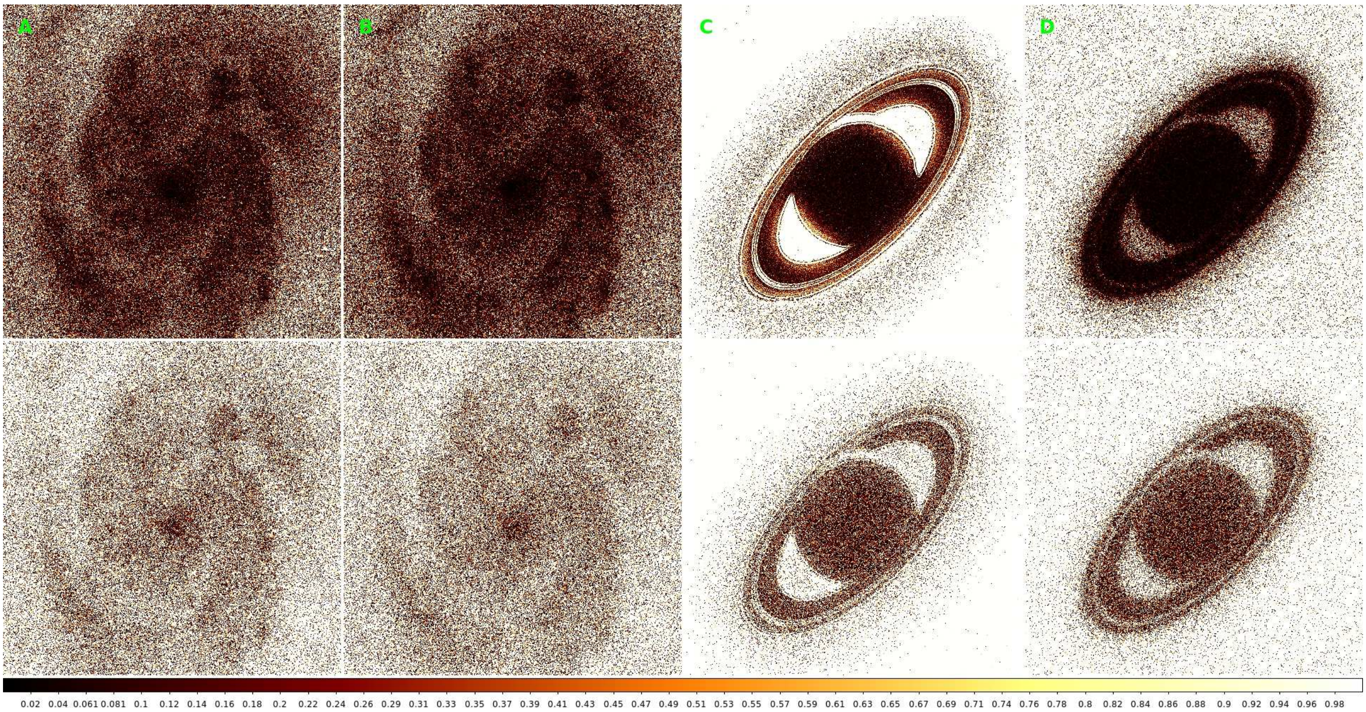

Figure 3.13: Residual maps for the reconstructions obtained from data sets corrupted with Gaussian noise with $\sigma = 5$ (top row) and $\sigma = 15$ (bottom row). Reconstructions were performed with a reference PSF at SR=45%. Columns A and C: ACMLE. Column B and D: MISTRAL. Linear scale from 0 to 1.

2007) or waveatoms (Demanet & Ying 2007), to name only two.

Much more important, in our opinion, is studying how multitransform support or probabilistic masks can be included into blind and myopic approaches to improve their performance in terms of controlling the noise reconstruction inherent to every image reconstruction algorithm. This will be the topic of our future research.

Chapter 4

High contrast exoplanet imaging using spectral deconvolution and the wavelet transform. Application to EPICS instrument.

To be submitted to Astronomy and Astrophysics (A&A). Year 2013.

Abstract

The next decade will see the birth of the 39-m European extremely large telescope (E-ELT). Among many other science programmes, the exoplanet imaging camera and spectrograph (EPICS) is a projected instrument to directly observe and characterize extra-solar planets with E-ELT. It will produce integral field spectroscopy (IFS) datacubes that can be analyzed with the so-called spectral deconvolution (SD) technique to make out the presence of faint objects orbiting around bright stars. This research is aimed to test if the inclusion of multiscalar transforms, such as the wavelet transform (WT), in the analysis of the IFS datacubes produces some improvement in detection of faint sources with respect to the performance obtained with the classical SD approach. The 1D undecimated WT is used to decompose, at different levels of resolution, the spectral signal at each particular location of the FOV. Then, the classical SD technique is used at each wavelet scale to create maps which are representative of the frequency information. Different simulations with the presence of fake companions are analyzed with both approaches, i.e., with and without wavelets. Several cases of interest for EPICS are studied, e.g., suppression of the diffraction with the APLC chronograph and an apodized-only solution at Talbot lengths equal to 1 and 10, and different contrast magnitude differences between the main star and the companions ranging from 1×10^{-5} to 1×10^{-9}, at angle separations between ~ 20 and $\sim 180 mas$. A general improvement of one order of magnitude contrast is observed when the wavelet analysis is used. Wavelets allow APLC performance to be comparable to that obtained with the apodized-only solution for contrasts up to 1×10^{-8} at angle separation from $\sim 70 mas$.

4.1 Introduction

The 39-m European extremely large telescope (E-ELT) is designed to increase both the collecting power and the angular resolution with respect to the current 10-m telescopes. One of the most important science programme to be performed will be the discovery and spectral characterization of planets outside the Solar system. While young and hot exoplanets are the primary targets for the SPHERE (Boccaletti et al. 2008) and the Gemini planet imager (GPI, Marois et al. 2008) instruments, E-ELT will be geared towards the detection of old and cold planets, similar to the ones in the Solar System. These targets are significantly more challenging. The contrasts to be expected are between 10^{-7} and 10^{-10} , depending on the size of the planet and its physical distance from the star. Old planets do not emit any significant own light in the near IR, they dominantly reflect light generated by the stars. The planet/star flux ratio, F_p/F_s , is (Brown 2004):

$$\frac{F_p}{F_s} = p(\lambda)\Phi(\beta)\left(\frac{R}{r}\right)^2, \tag{4.1}$$

where p is the geometric albedo of the planet, which depends on wavelength, R stands for its radius, r is its separation from the star, and $\Phi(\beta)$ is the phase function of the planet which depends on the phase angle β (angle at the planet between star and observer):

$$\Phi(\beta) = \frac{\sin(\beta) + (\pi - \beta)\cos(\beta)}{\pi}. \tag{4.2}$$

Assuming an optimistic case of a planet observed at op-position ($\Phi(\beta) = 1$) and a near-IR albedo of 0.3 we get for Jupiter ($R = 0.00048AU$, $r = 5.2AU$) an expected contrast of 2.53×10^{-9} . For Earth we get 5.45×10^{-10} . In both cases we assumed distances between planets and their parent star identical to the ones in our system. For smaller stars the habitable zone is expected to be found closer (Selsis et al. 2007). For M-type stars it could be located as close as $0.1AU$. Assuming the same spatial scaling applies to other, heavier planets in this hypothetical planetary system around an M-type star, we get contrasts of 2.53×10^{-7} for "Jupiter" and 5.45×10^{-8} for "Earth". Planets with masses up to ten times greater than the Earths mass (i.e., "Super-Earths") would have contrasts around 2.6×10^{-7} in this M-type star's system. The instrument designed to develop such programme will be the exoplanet imaging camera and spectrograph (EPICS). Detailed science cases and overall requirements of the instrument are given in Kasper et al. (2008, 2010).

In a nutshell, EPICS will observe in both the visible and the near infrared range of the light being equipped with an integral field spectrograph (IFS) for the former and with a differential polarimeter (EPOL) for the latter (Verinaud et al. 2010). EPICS wavefront control is composed by two loops and calibration stages. The first one is a single conjugate adaptive optics (SCAO) that corrects for the largest part of the telescope jitter and "low order" aberrations until the range of the second loop. Its wavefront sensor (WFS) is a modulated pyramid sensor sensitive to a spectral range between $600 - 900nm$, 84×84 subapertures and a frame rate of $1KHz$. The second calibration stage is a extreme adaptive optics (XAO) system that uses a high density deformable mirror (DM), a second fast tip-tilt mirror and a second wavefront sensor. The spectral range is between $700 - 900nm$, it consists in 210×210 subapertures and a frame rate of $3KHz$. Non-common path aberrations are measured off-line before observations and are applied to the XAO WFS path to attain Strehl ratios (SR) larger than 85%. Simulations of the EPICS AO system are presented by Korkiakoski & Verinaud (2010).

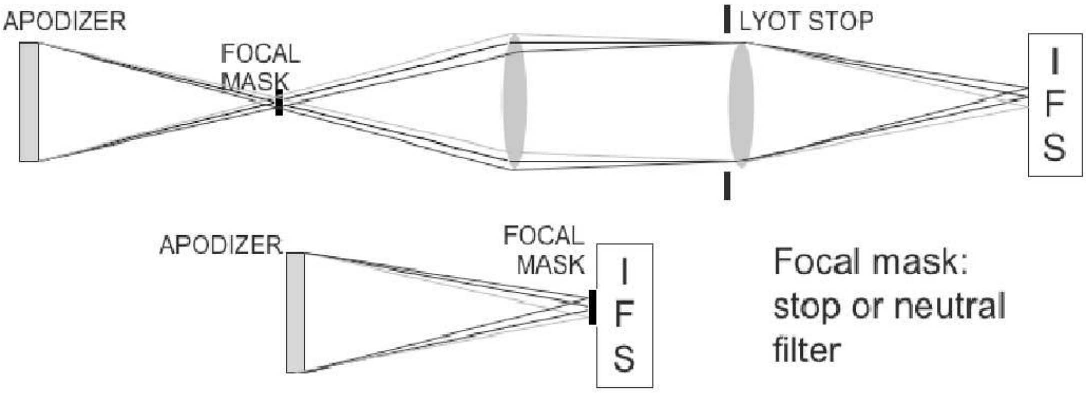

Figure 4.1: The two diffraction suppression concepts considered. Top: Apodized Lyot Chronograph (APLC case). Bottom: Diffraction rejection through apodization only and central core of star image attenuated at entrance of IFS (APOD case). —after Verinaud et al. (2010).

The diffraction suppression concept is a very important subsystem of any high contrast imaging instrument. The apodized pupil Lyot chronograph (APLC, Soummer 2005) is one of the best considering the compromise between starlight rejection and throughput given the E-ELT aperture geometry. It has an apodizing mask in the first pupil-conjugate plane after XAO. This mask suppresses the Airy rings in the sub-sequent focal plane, and scatters non-coherent light outside the next pupil-conjugate plane where the stop is located. That is why for the APLC chronograph the stop is not undersized in opposition to the non-apodized Lyot chronograph. However, it suffers from chromatic properties impairing the speckles intensity and position correlation within the IFS spectral range, especially at medium and small angular separation. Suppressing diffraction with an “apodizer-only” solution (hereinafter APOD), which consists in an apodizer close to the pupil, a focal mask in the entrance of IFS without a Lyot stop, leads to only a moderate loss in throughput and light rejection, mainly because of the small angular resolution of the E-ELT. This solution is very simple and is intrinsically fully achromatic. The technology chosen for the development of this apodizer is based on the microdots technique (Antichi et al. 2010). A prototype has been developed for the Fresnel-free experiment for EPICS (Martinez et al. 2009). We have chosen to test the apodizer-only solution for EPICS because it is a simple concept and enables to conserve the speckles spatio-spectral correlation properties over the largest spectral bandwidth. Both APLC and APOD optical configurations are shown in figure 4.1.

Direct imaging of faint exoplanets from the ground requires differential techniques to remove the speckle noise. There are two basic methods that performs such task. Firstly, the simultaneous differential imaging (Racine et al. 1999), and its generalization for IFS datacubes, the spectral deconvolution (SD) technique (Sparks & Ford 2002). These approaches exploit that speckles position and intensity at two different wavelengths are highly correlated, i.e., the speckle angular separation from the center of the FOV is directly proportional to wavelength and to the spatial frequency of the corresponding perturbation. Secondly, the angular differential imaging (Marois et al. 2006) takes into account that speckles position and intensity at two different observing times can be correlated because of the quasi-static nature of the instrument optical disturbances. Therefore, field rotation will affect differently off-axis point sources and speckles.

The efficiency of angular differential technique depends very strongly on the stability of the wavefront errors and it requires to freeze the telescope optics with respect to the instrument. The difficulty of an optical derotator implementation and the uncertainty today of the future segmented E-ELT wavefront error temporal stability has privileged the use of SD as a main technique for speckle rejection for EPICS. Nevertheless, Mesa et al. (2011) proves with simulations for SPHERE IFS datacubes (Beuzit et al. 2008) that angular differential imaging improves in a factor of two or three the contrast attained when applied together to SD.

Another important reason to prefer SD is that the speckle chromatic elongation will be very large on E-ELT. Therefore, the use of the large speckles position variation with wavelength to disentangle them from a point source will be very efficient and calls for the largest spectral bandwidth possible.

Finally, the wavelet transform (WT) for unidimensional data analysis has been widely used in astronomy (e.g. Szatmáry et al. 1994; Foster 1996; Polygiannakis et al. 2003) to differentiate noise-like components from real signals, or to obtain significant parameters of period series, like the amplitude, the phase or the period itself, which can be representative of light variations from variable stars.

Otazu et al. (2002, 2004) proposed the WT for period determination of sine- and burst-like signals superimposed with a low-frequency sinusoidal function. The search for periodic signals is common in many areas of astronomy and reveals the presence of different physical processes, e.g., variable stars, pulses, eclipses in two-body systems or occultations by a precessing accretion disk. Classical methods for period detection lie into two categories. First, epoch-folding methods (PDM: Stellingwerf 1978) which are based on the analysis of the dispersion of the different light curves produced by folding the data over a range of trial periods. These methods tend to fail when more-than-one periodic signal are present in the dataset, specially for low signal-to-noise ratio (SNR). Second, Fourier-based methods (CLEAN: Roberts et al. 1987) that use the FT in combination with deconvolution techniques to deal with the data sampling function. Their behavior in the presence of noise is better with respect to PDM but they fail in the detection of non-sinusoidal signals, e.g., a pulse emission over an orbital period. Otazu et al. (2002, 2004) decomposed in wavelet scales the temporal signal and applied PDM and CLEAN to each of them to create a sort of wavelet-based period analysis method, i.e., WPDM and WCLEAN. The authors proved that the performance of wavelet-based algorithms was always better with respect to their non-wavelet-based equivalent at high-noise levels, as well as the numerical accuracy of the detected periods.

Another interesting application of 1D wavelet decomposition is found in lunar occultations (LO) programs (Fors et al. 2008). This technique takes advantage of the relative Moon movement to obtain stellar high angular resolution with ground-based observations. When the lunar limb interposes between a star and the observer, because of the ondulatory nature of light, the disappearance of the object is not immediate. During a short but measurable period of time the variation of the source intensity is described by a characteristic diffraction pattern of fringes and a decreasing light profile. The analysis of such intensity evolution in time offers the possibility to calculate the physical parameters of a binary system or to deduce the presence of an accretion disk surrounding the main star. In order to do this, accurate determination of the instant when the occultation event occurs is essential. Fors et al. (2008) empirically found that such instant is calculated with higher precision from the seventh wavelet scale in which the original dataset can be decomposed. Once this event time is obtained, the stellar and the background intensities are easily deduced from the fifth wavelet scale. This procedure allows to design an automatic pipeline to analyze large sets of data, with the corresponding reduction of the time needed for an initial preprocessing task.

4.2 Spectral deconvolution

Spectral deconvolution (SD) was first proposed by Sparks & Ford (2002), who applied this technique to detect faint companions in simulated IFS 571-frames datacubes of a 2-m telescope with a Lyot chronograph. SD makes use of the fact that the radial location of the speckles and residuals of the Airy pattern are proportional to the wavelength, while the location of a companion with respect to the star is fixed. Re-scaling all individual images of an integral field spectrograph (IFS) datacube, proportional to its wavelength, by Fourier interpolation to fit a common grid, aligns the speckles but makes the planet move inwards with increasing wavelength (see figure 4.2). Speckles are now well fitted by a smooth (e.g. a low-order polynomial) function to each spaxel[1] while the planet produces a narrow bump when travelling through the pixel at a certain wavelength range. Since this bump is badly fitted by the smooth function, the subtraction of the fit removes most of the speckles and leaves the planet.

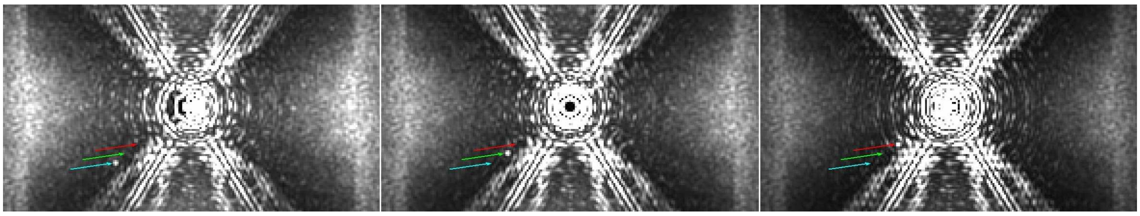

Figure 4.2: Planet moving from the initial position (blue arrow) to the final position (red arrow) in the scaled datacube.

The SD procedure is shown in figure 4.3. After the first fitting step, any outliers at the 2-σ level (possible planets) are removed from the spaxel, and this procedure is repeated until no value in the sample exceeds the fit by two standard errors. Once the fit is not skewed anymore by possible signals, it is subtracted from the spaxel under consideration. This step removes most of the stellar modulation, but leaves possible planetary signal intact. The resulting images are then re-scaled to

---

[1] spaxel: all the pixels at the same location within a IFS datacube, i.e., a one-pixel-size datacube containing the spectral information at this location.

the original grid, so that now a planet would be aligned on the same pixel again. Finally, the cube is collapsed to produce one final image. Changes in the PSF morphology, which naturally occur for a wide wavelength range where the Strehl ratio changes significantly, are ignored. Nevertheless, these changes are smooth when one considers many channels, as we do here, and can be fitted with a moderately smooth polynomial. Spectral deconvolution is applied to the IFS datacube to typically increase the contrast by two or three orders of magnitude.

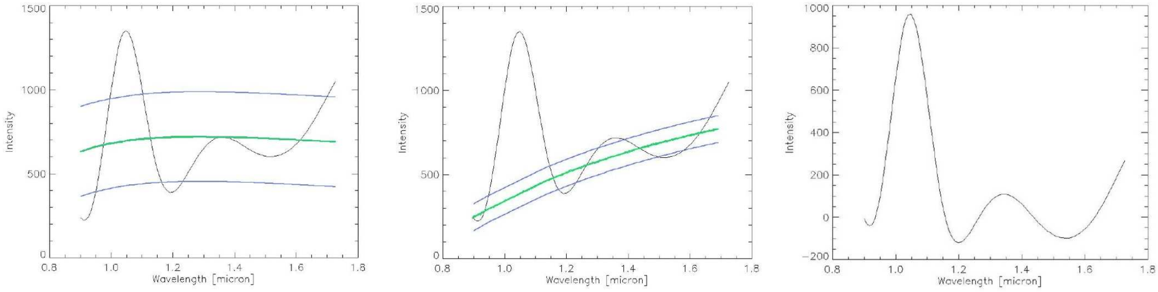

Figure 4.3: Polynomial fitting process in a particular spaxel/location. Left: first fit trial. Middle: final fit. Right: residual. Black curve: signal present through the spaxel. Green curve: low-order polynomial fit. Blue curves: 2-σ levels. Note how either the planet signal (centered at 1.05μm.) and a higher order chromatic modulation (centered at 1.35μm.) have been recovered.

From figure 4.3 (right panel), one can see that only a small portion of the planetary signal will not be taken into consideration, on the other hand, the low-order fitting will leak through significant modulation induced by the chronograph. Therefore, the polynomial fitting approach can leave out significant residuals (the small bump on the right of the planet). Hence, classical SD has two main problems. First, it cannot remove PSF features which do not scale spatially with wavelength. Unfortunately, a chronograph produces non-trivial modulation of the residuals close to the axis and, in this region, SD is not as effective as further out. Second, speckles that are close to the optical axis do not move significantly over the entire wavelength range. Conversely, in the numerically rescaled datacube, close-in planets “dwell” on single spaxels and can be “killed” by the traditional approach (see figure 4.4).

The classical iterative SD method can be summarize as follows:

1. Scale spatially, by Fourier interpolation, each image in the IFS datacube to fit a common grid. Now, the Airy rings are aligned while planets positions change through the datacube: they get closer to the star with increasing wavelength (see figure 4.2).

2. Fit a 3rd-order polynomial to the signal present in the spaxel and reject all those points above 2σ. Repeat this process until no outliers are found (see figure 4.3, leftmost and middle panels).

3. Remove the final fit to the signal present in the spaxel (see figure 4.3, rightmost panel).

4. Re-scale the datacube to the original grid. Planets would be aligned over the same pixel again.

5. Collapse the cube to produce a single image.

Thatte et al. (2007) make use of SD to extract H- and K-band spectra of the faint close-in source AB Doradus C, observed with SINFONI instrument (Thatte et al. 1998) at VLT. AB Dor C is a M dwarf companion orbiting around K dwarf AB Dor at separation $\sim 200 mas.$, with a contrast magnitude in K-band of $5 mag$.

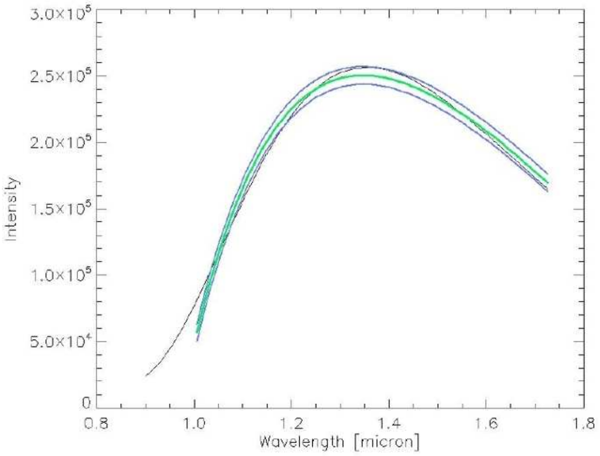

Figure 4.4: Polynomial fitting process for a close-in planet. The low order polynomial "fits" correctly to the planet signal thus subtracting the planet contribution and avoiding its detection.

4.3 The 1D Wavelet transform

The wavelet transform (WT) is one of the most successful mathematical tools to perform a multiresolution decomposition of a signal or image. Such decomposition is based on the local frequency content, hence, each one of the resolution channels contains a representation of a certain frequency range. The WT is an intermediate representation between the Fourier (FT) and the temporal one (or spatial one for 2-D data). Since wavelets are defined within a tight frame (i.e., they are strongly localized in time or space), they give us a combined idea of both the local content of the frequencies and their temporal/spatial location, as opposed to the FT, which only is able to offer a global view of the frequency content of the signal.

Given a signal $f(t)$ (where t is any physical variable indexing the X-axis, e.g., like time or wavelength in the case of IFS datacubes) a sequence $F_m[f(t)]$ of approximations of $f(t)$ can be constructed. Each $F_m[f(t)]$ is specific for the representation of the signal at a given scale or resolution and represents the projection of $f(t)$ from the signal space $\mathcal{S}$ onto subspace $\mathcal{S}_m$. The differences between two consecutive scales m and $m+1$ are the corresponding multiresolution wavelet signal:

$$\omega_m[f(t)] = F_m[f(t)] - F_{m+1}[f(t)]. \tag{4.3}$$

This wavelet signal can also be expressed as:

$$\omega_m[f(t)] = \sum_l W_{m,l}(f)\psi_{m,l}(t), \tag{4.4}$$

where coefficients $W_{m,l}(f)$ are given by the direct wavelet transform of the signal $f(t)$:

$$W_{m,l}(f) = \int_{-\infty}^{+\infty} f(t)\psi_{m,l}(t)dt. \tag{4.5}$$

The coefficients $W_{m,l}(f)$ are called the wavelet coefficients of $f(t)$ and they correspond to fluctuations of the signal $f(t)$ near the point l at resolution level m. Equation 4.5 represents the expansion of the signal $f(t)$ in the set of basis functions $\psi_{m,l}(t)$. These basis functions are scaled and translated versions of the so-called *mother wavelet function* $\psi(t)$. Therefore, all the basis functions $\psi_{m,l}(t)$ have the same profile, i.e., the mother wavelet function profile. The WT describes at each resolution step the difference between the previous and the current resolution representation. By iterating the process from the highest to the lowest resolution level we obtain a pyramidal representation of the signal.

Unlike the FT, which spreads the noise across all the frequencies, the WT confines the noise contribution only to the highest frequency wavelet signal.

4.3.1 The “à trous” algorithm

Both Otazu et al. (2002) and Fors et al. (2008) obtained a discrete wavelet decomposition by means of the well-known Starlet transform or “à trous” algorithm (Shensa 1992; Starck & Murtagh 1994; Starck et al. 2010) for one-dimensional signals. Therefore, given a signal $p(t)$ the following sequence of approximations can be constructed:

$$F_1(p) = p_1, \quad F_2(p_1) = p_2, \quad F_3(p_2) = p_3, \ \ldots \tag{4.6}$$

The algorithm performs successive convolutions with a low-pass filter, which is designed by means of the so-called *father wavelet function* or scaling function $\phi(t)$. For a *binary scaling* representation, in which each scale expresses a double resolution with respect to the previous one, that convolution is written as:

$$p_i(\tau) = \frac{1}{2^i} < p(t), \phi\left(\frac{t-\tau}{2^i}\right) > . \tag{4.7}$$

In our work, we use a scaling function which has the main-lobe profile of the function $sinc^2(t)$, in coherence with the shape of the planet signature through the spaxel. This leads to a convolution with a 5-element normalized mask. The sequence p_i are versions of the original signal p at different resolution levels,

Therefore, the wavelet signals are computed as differences between two consecutive approximations $p_{i-1}(\tau)$ and $p_i(\tau)$, i.e., $\omega_i(\tau) = p_{i-1}(\tau) - p_i(\tau)$ $(i = 1, 2, ..., n)$, being $p_0 = p$. It also can be expressed as:

$$\omega_i(\tau) = \frac{1}{2^i} < p(t), \psi\left(\frac{t-\tau}{2^i}\right) >, \tag{4.8}$$

where ψ is the aforementioned mother wavelet function and it is defined as:

$$\frac{1}{2}\psi\left(\frac{t}{2}\right) = \phi(t) - \frac{1}{2}\phi\left(\frac{t}{2}\right), \tag{4.9}$$

thus highlighting its band-pass nature. The inverse reconstruction formula can be written as:

$$p = \sum_{i=1}^{n} \omega_i + p_r, \tag{4.10}$$

where ω_i are the multiresolution wavelet signals and p_r is the residual wavelet signal (indeed, $n = r$, but we explicitly substitute n by r to enhance the concept of *residual*). Many other wavelet decomposition algorithms usually reduce the number of points (typically by a factor of 2) when going through decreasing resolutions (increasing scales). Such process is called by the name of *decimation*. However, one of the advantages of the Starlet transform is precisely the rejection of the decimation process, i.e., the “á trous” algorithm is *undecimated*. Therefore, all those p_i (and subsequently all the ω_i as well) have the same number of points than the original signal p_0. This allows to work directly in the temporal space (wavelength space in the case of IFS datacubes) with the frequency content of the wavelet signal, instead of working on the decimated wavelet space.

4.3.2 Proposed algorithm

The use of the aforementioned binary decomposition is very typical in wavelet analysis. Hence, the original signal p_0 has double resolution with respect to p_1, such signal p_1 has double resolution with respect to p_2, and so on. In our particular case, if the spectral datacube is formed by $N = 100$ wavelength-points (i.e., there are 100 channels, each one at a particular wavelength) and our highest frequency scale is determined by a scaling function $\phi(t)$ sampled with 5 points, then the minimum resolution accepted for p_0 would be 5, p_1 would have a resolution of 10, p_2 of 20, p_3 of 40 and so on. In other words, the wavelet plane ω_0 would contain the frequencies representative of a resolution until 5 points, ω_1 those frequencies representative of a resolution between 5 and 10, ω_2 from 10 and 20, etc. Hence, a total frame of 100 wavelength-points would be completely decomposed in 5 wavelet planes plus 1 residual.

A difficulty arises when taking into account the slowly varying component present through the spaxel. Basically, this means that the initial wavelength always exhibits a value much larger than the one showed by the final one. The WT we used is based on the symmetric convolution of the signal of interest with the scaling function, which means that the extremes of the signal affect each other’s result. In other words, points close to 0 are used to calculate the wavelet response at points close to $N - 1$ and vice versa. If the user is interested in the analysis of a certain period P, assuming P is an odd number, only points from $\frac{P}{2}$ until $N - \frac{P}{2}$ are really usable, e.g., if $N = 100$ and $P = 20$, then only frames from 10 to 90 are numerically representative of the real frequency content of the signal, since the rest of wavelet coefficients at the edges of the wavelength series are highly skewed by the slow variance along the spaxel. This is specially dramatic when one is interested in working at low-frequency values, i.e., $P \sim N$ or when a planet signal is close to the spaxel extremes.

However, Otazu et al. (2002) based his work on ondulatory signals with several periods around a constant value, hence, there always were more than one complete period further out of the time series edges; whereas Fors et al. (2008) tried to search a point (the LO event time) between two well-localized energetic-constant signals, i.e., the one before the LO which is dominated by the star flux, and the one after the LO which is dominated only by the background. Furthermore, and depending on the observation, the point in-between could be easily placed in the middle of the time scale in most of the cases. Indeed, Fors et al. (2008) reported some cases where his wavelet method failed because, due to certain observing conditions, it was not possible to center the LO event in the middle of the time scale, being displaced to the edges.

In order to partially alleviate this problem, we propose to extrapolate the signal beyond its real values. Therefore, point number "-1" , instead of point $N-1$, would be used to calculate the wavelet coefficients from point 0, while point "N", instead of point 0, would be used to calculate the wavelet coefficients until point $N-1$. Such extrapolation creates a sort of *padding* along the spectral line to "protect" the real values belonging to the signal. If such values belong to a left- or right-side planet signature, we would be extrapolating the planet signature itself.

We have tried several methods to perform such extrapolation, e.g., fitting a low-order polynomial to the first and the last 5 or 10 wavelength points. However, we found our best results simply extending the signal by means of a straight line with the same slope shown by the first and the last 5 or 10 points. Finally, we add to the extrapolated points an estimation of the high frequency variations present in the real signal to avoid the creation of false high-frequency wavelet coefficients in the transitions between the real points and the extrapolated ones.

Of course, it must be admitted that this extrapolation is one of the main weaknesses of the wavelet analysis for SD, the lower is the frequency of analysis the more importance it is, since more extrapolated points, together with real values, would be used to calculate the wavelet coefficients. It would be necessary a deep inspection of, first, the best way of performing such extrapolation and, second, its consequences over the final reconstructed image. On the other hand, the price to pay for not using it is the rejection of real observed wavelength values through the spaxel, i.e., the lower is the frequency of analysis the larger the number of rejected values would be.

Finally, the proposed wavelet SD algorithm is outlined as follows:

1. Scale spatially, by Fourier interpolation, each image in the IFS datacube to fit a common grid. Now, the Airy rings are aligned while planets positions change through the datacube: they get closer to the star with increasing wavelength (figure 4.2).

2. Decompose all the spaxels, using the 1D Starlet transform, at the desired resolution levels with the corresponding scaling function (figure 4.5).

3. Fit a low order polynomial to the wavelet representation of the spaxel and reject all those points above 2σ. Repeat this process until no outliers are found.

4. Remove the final fit to the signal present in the wavelet signal.

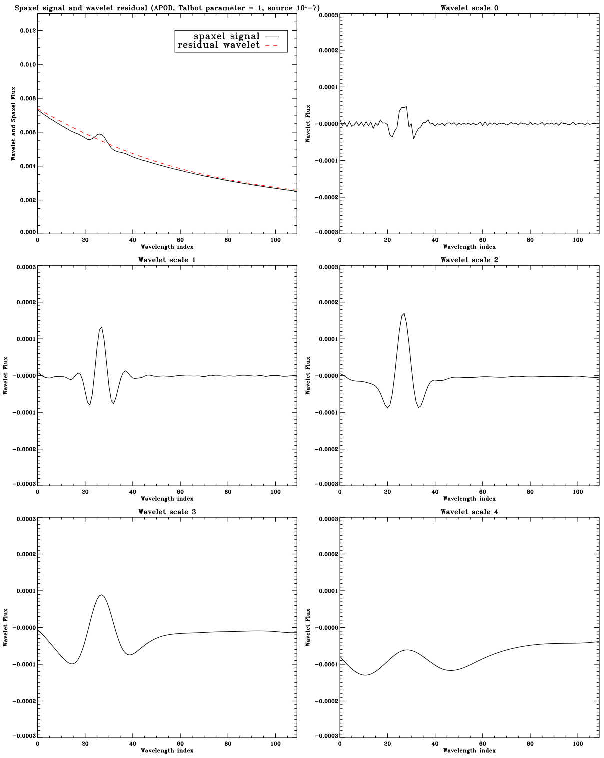

Figure 4.5: Wavelet decomposition of the planet signal through the spaxel. Uppermost left panel: original signal and residual wavelet (red dashed curve). The sum of all the wavelet plots plus the residual is equal to the original signature.

5. Re-scale the datacube to the original grid. Frequency components from the planet at the spaxel would be aligned over the original planet pixel.

6. Collapse the cube to produce a single image, which is representative of the spaxels frequency content at this resolution level.

7. Go back to 2 if different wavelet resolution levels are desired.

Note that the final image will only have those frequency components present along the spaxels considered by a particular choice of a level of resolution. The user can build as many of such single images as he wishes, all of them containing a particular interval of frequencies along the spaxels of the scaled datacube. We also want to stress the fact that the fitting of a low-order polynomial at point 3 is only a particular choice motivated by the need of a pure comparison between the classical SD and the wavelet-based approach. On the other hand, this fitting also helps to remove residual modulations not really belonging to real sources, which can be introduced during the decomposition process.

The WT offers more versatility and possibilities to analyze the signal which results at each of the wavelet scales, e.g., it would be possible to search for maxima that would be representative of the presence of planets. It would also be possible to decompose each spaxels at different wavelet scale depending on the angle separation, i.e., bearing in mind that planet signatures further out are narrow bumps, while close-in planets can cover the whole wavelength series, one could think in using high-frequency resolution levels for the former and low-frequency one for the latter. However, we consider a better strategy to create several images decomposing all the spaxels of the FOV using the same frequency resolution at each one. In that way, it can be ensured that all the information along the spaxels will be classified at some of the final images, while otherwise we assume the risk of loosing information due to a bad choice of the frequency level at a particular angle separation. Preliminary results of the joint use of WT with SD were presented by Gladysz et al. (2011) for simulations with HARMONI instrument (Thatte et al. 2010) and Lyot chronograph for E-ELT.

4.4 Simulations

The simulations we used to test both the classical and the wavelet spectral deconvolution (SD) techniques have been designed to evaluate the limiting contrast of detection as a function of the angle separation. Furthermore, we also want to test the performance of the algorithms with respect to the type of chronograph and chromatic aberrations introduced in the light wave by the Talbot effect (Talbot 1836).

Fourier transform of the field after the Lyot stop gives the electric field in the detector plane, and the squared modulus of this quantity yields the final on-axis instantaneous PSF. For the images of the test planets the electric field is propagated only through the Lyot stop and the apodizer for APLC, giving the off-axis PSF. For the APOD case there is not a different optical path for the light coming from the main star and the companions. We scale and shift this PSF (off-axis PSF for APLC) to simulate eight planets located at angle separations of 30, 58, 111, and 184mas. At each of these positions two planets are placed to enhance their presences in the contrast curves calculated from the reconstruction. Therefore, each dataset is composed of 8 planets. Dataset are polychromatic and covers the near IR range from 950 to 1650nm in 110 channels to calibrate and subtract residual stellar contribution after AO and diffraction suppression. The spectral range was pre-determined by the additional objective of this project: a cross-check with the EPICS simulation results. Originally, the $0.9 - 1.7\mu m$ range was chosen by the EPICS team because it contains the most interesting molecular bands: O_2, CO_2, H_2O, and CH_4. It should also be mentioned that the planet finder instrument for the Very Large Telescope (VLT), i.e., SPHERE, will have an IFS working in the similar spectral range (Fusco et al. 2006).

All the 8 planets have the same contrast difference with respect to the main star. To evaluate the detection limit the 8 planets magnitude is changed to cover a range of contrasts from 1×10^{-6} to 1×10^{-9}, i.e., we have created 4 simulations with 8 planets each, all of them with the same contrast. In such way we can verify, for example, if planets with contrast difference of 1×10^{-7} or 1×10^{-8} are detectable at angle separations between 58 and 111mas, e.g., if the companion of 1×10^{-8} is reconstructed at the position of 111mas but it is not at 58mas, whereas the planet of 1×10^{-7} it is, then, it is possible to ensure that the limit of detection for sources at angle separations of 111mas is a contrast of 1×10^{-8}.

In all our simulations we assume an IFS is present but we do not simulate spectra (we assume a flat exoplanet spectrum) to keep a certain simplicity, which is more convenient to show the performance of the WT over IFS datacubes. Besides, we do not have a detailed model of light propagation through a spectrograph. Such a model depends on the design of an IFS (BIGRE or TIGER configuration, see Antichi et al. (2009)) and is certainly outside the scope of this work. On the other hand, we have decided to model Fresnel propagation, which depends on the number and relative separations of optical surfaces within an instrument (Marois et al. 2008). A useful quantity to evaluate the impact in terms of Fresnel diffraction of an optical element is the distance of the optic to a pupil conjugate expressed in characteristic Talbot length units. The characteristic Talbot length L_T is defined as:

$$L_T = \frac{2\Lambda^2}{\lambda}, \tag{4.11}$$

where Λ is the smallest aberration spatial period considered (this period corresponds to the highest spatial frequencies the AO can correct, projected on the considered optical element) and λ is the largest light wavelength to be considered. Therefore, simulations include chromatic aberrations at 1 and 10 Talbot lengths from the pupil.

Finally, all these simulations have been also replicated for 23, 43, 73, and 146mas to increase the range of angle separations.

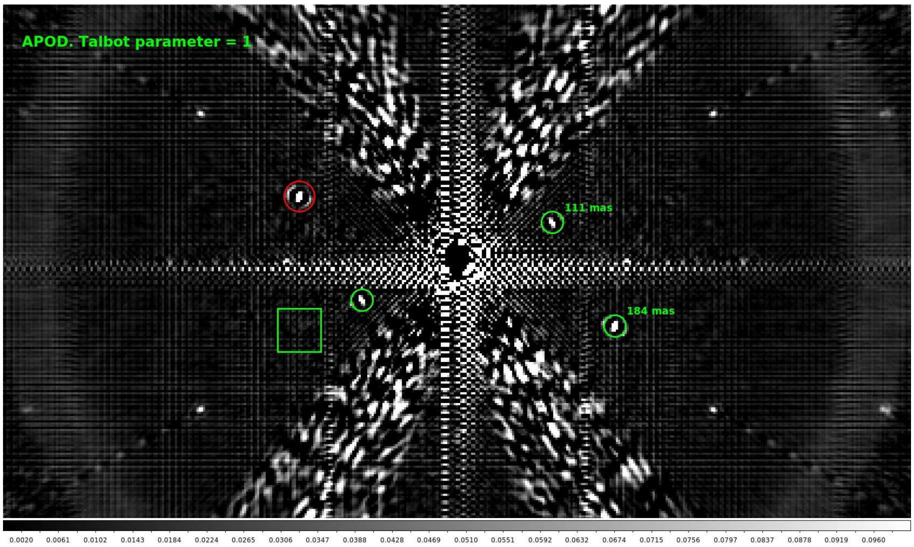

Figure 4.6: APOD case, Talbot length equal to 1. Planet contrast 1×10^{-8}. Peak planet within the larger red circle has been used as normalization reference. Green square is used to calculate the statistics shown at tables 4.1 and 4.2.

Figure 4.6 shows the resulting image after applying the SD technique to the wavelet signal scale number one, in the case of planet contrast of 1×10^{-8}, APOD case with Talbot length equal to 1. Because, in the first place, the main star flux is generally affected by the use of masks, secondly, the behavior of either the classical SD and the wavelet SD in the center of the FOV is very different and, finally, we are more interested in how bright is the reconstructed source with respect to the surrounding background, for all these reasons we have decided that reconstructions obtained from the classical SD or from the wavelet SD are normalized using the maximum peak of their particular recovered planet within the larger red circle at 184mas, which is always the reconstructed planet less affected by chromatic aberrations, and further from the center of the FOV. Pixels within the green square, also at 184mas, are used to devise some statistical quantities of the background.

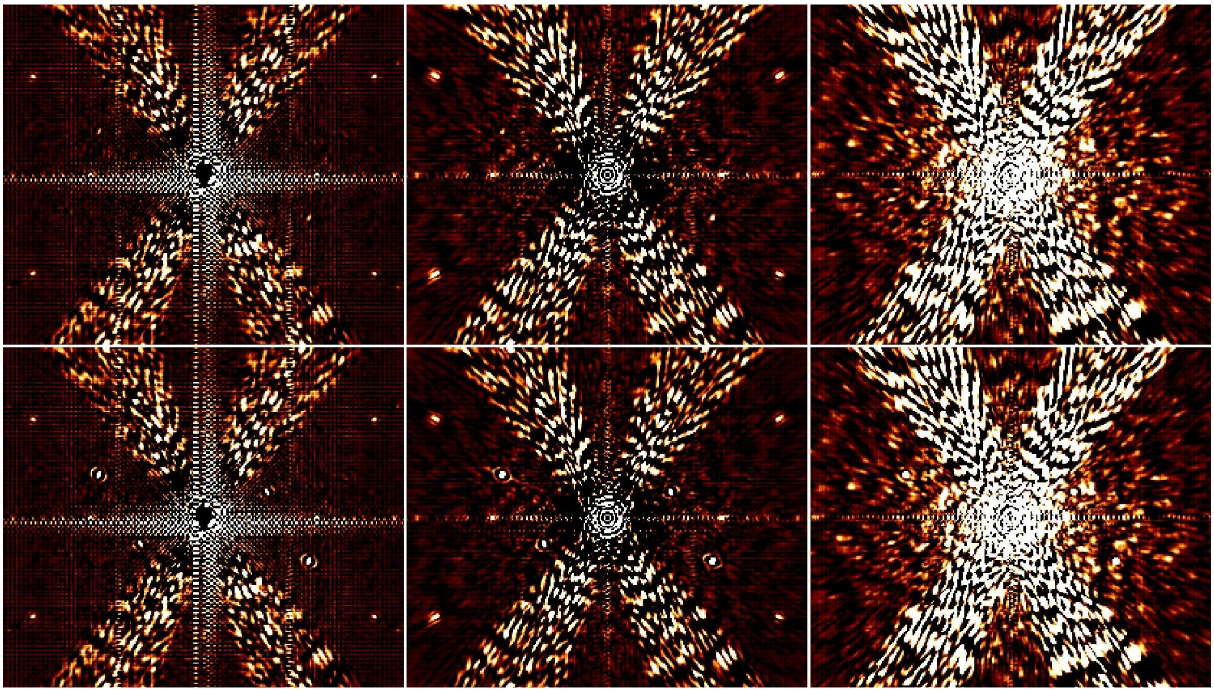

Figure 4.7: APOD case, Talbot length equal to 1. Top row: planet contrast 1×10^{-9}, linear scale from 0 to 1. Bottom row: planet contrast 1×10^{-8}, linear scale from 0 to 0.1. Left column: image obtained from spaxels wavelet scales number 1. Middle column: image obtained from spaxels wavelet scales number 2. Right column: image obtained from the classical spectral deconvolution approach, polynomial fitting order equal to 3. All images have been normalized with respect to the maximum peak of the planet present at position (451,538) —red larger circle at figure 4.6 —.

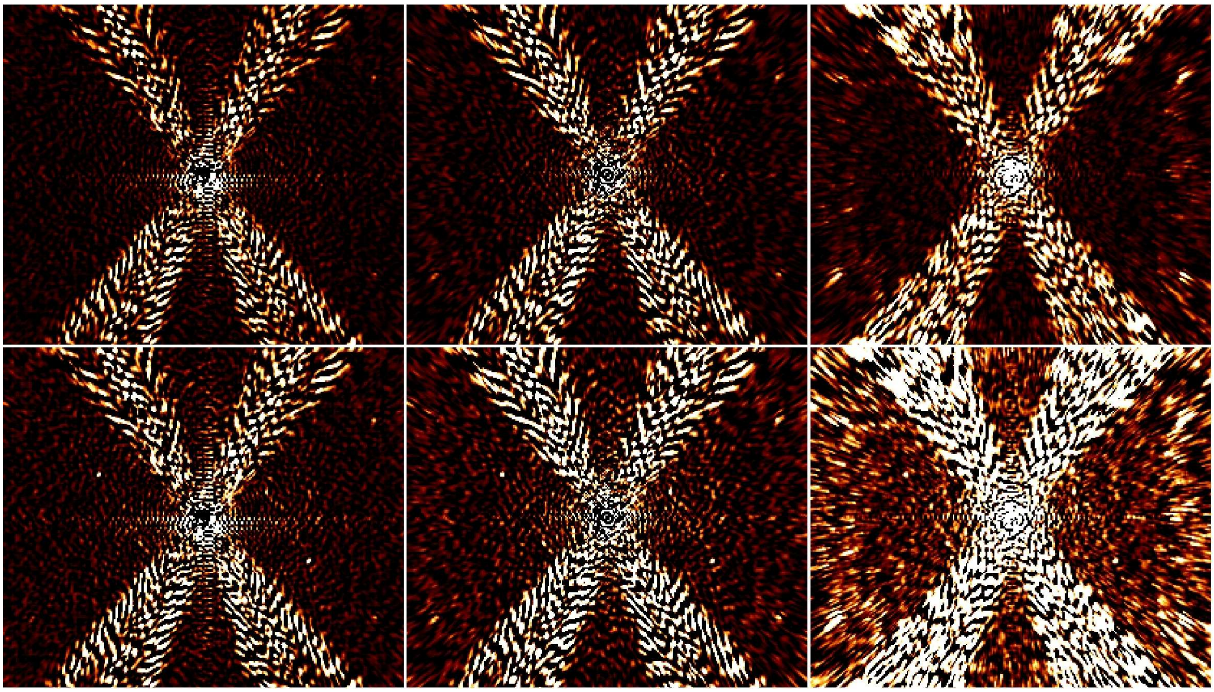

Figure 4.8: APOD case, Talbot length equal to 10. Top row: planet contrast 1×10^{-9}, linear scale from 0 to 5. Bottom row: planet contrast 1×10^{-8}, linear scale from 0 to 0.5. Left column: image obtained from spaxels wavelet scales number 1. Middle column: image obtained from spaxels wavelet scales number 2. Right column: image obtained from the classical spectral deconvolution approach, polynomial fitting order equal to 3. All images have been normalized with respect to the maximum peak of the planet present at position (451,538) —red larger circle at figure 4.6 —.

4.5 Results

The final reconstruction achieves high dynamic range through three functionalities. Firstly, Image sharpening through turbulence compensation XAO. Secondly, diffraction suppression through APLC chronograph or apodization-only (APOD case) and, in the third place, we implement the SD technique with and without WT to remove the residual stellar light from the spectral data cube. It should be mentioned here that efficient chronographic starlight suppression is helped by reducing the image to a narrow spot with AO, i.e. the second step in our simulations is intrinsically linked to the first. If AO works as expected the diffraction-limited spot hits the focal-plane mask and most of the light is blocked.

Figures from 4.7 to 4.10 show reconstructions when SD techniques are applied over the original signal present at the spaxel and the three first wavelet scale it can be decomposed, when both APOD and APLC diffraction suppression solutions are used, with Talbot lengths equal to 1 and 10. In general, the use of wavelets increases the detection limit in one order of magnitude from $73mas$. Limiting contrast for the APOD case, Talbot length equal to 1 (figure 4.7) is established at 1×10^{-9} between 111 and $184mas$, and at 1×10^{-8} between 73 and $111mas$, where wavelet coefficients, at scales one and two, belonging to the planets are clearly visible beyond the background level. These detection limits must be increased until 1×10^{-8} and 1×10^{-7} when SD is applied over the whole original signals at each spaxel. An increase in the Talbot length from 1 to 10 affects the results (figure 4.8). Sources at contrast 1×10^{-9} are not visible anymore, whereas at contrast 1×10^{-8} are barely visible at $73mas$.

For APLC chronograph (figures 4.9 and 4.10), results are similar in terms of the gain obtained when WT is used with respect the classical approach, i.e., one order of magnitude from $73mas$, being the limit of detection at 1×10^{-8} between 111 and $184mas$, and at 1×10^{-7} between 73 and $111mas$. The increment of the Talbot length does not imply a big modification of those values. Indeed, the classical approach results are more affected by the Talbot length, whereas high-frequency wavelet reconstructions appear to be almost insensitive to this parameter. It must be said that the further away optical elements are from a pupil conjugate, the larger the Fresnel diffraction impact on the speckles spatio-spectral correlation (see equation 4.11). In order words, to reduce the influence of the Fresnel diffraction, all optical elements in the common path must be conjugated to a distance close to the pupil. This distance must not be larger than one characteristic Talbot length at the longest wavelength as defined for the highest spatial frequency of the aberrations (projected on the optical element) to be considered for a given FOV. Hence, reducing the Talbot length is one of the key points in the optical design of EPICS (Verinaud et al. 2010).

Tables 4.1 and 4.2 show the statistics within a 20×20 pixel square centered at $184mas$ (figure 4.6). Noise level in this square is measured in terms of the standard deviation beyond which a real source should be detected. Wavelet reconstructions from spaxel signal scales one and two always show better values than the corresponding one for the classical SD approach, around twice smaller standard deviation for APOD and between four and seven times smaller, depending on the Talbot length, for APLC. This is, in our opinion, one of the most important results from this paper, i.e., the

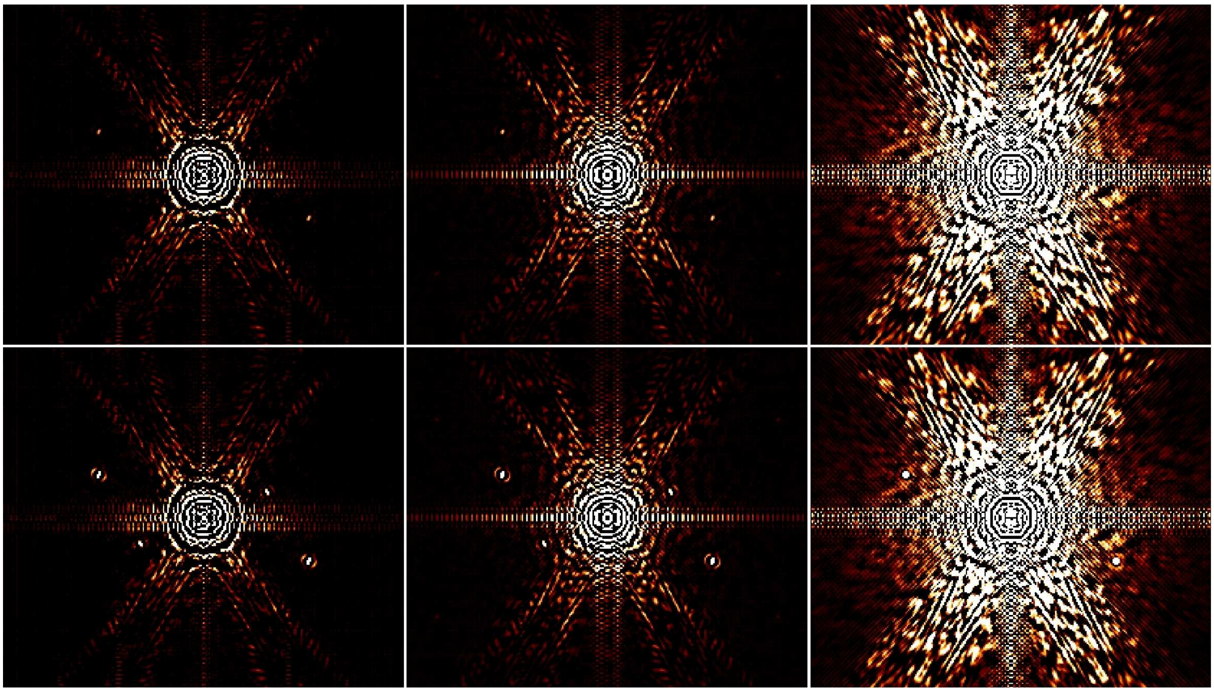

Figure 4.9: APLC case, Talbot length equal to 1. Top row: planet contrast 1×10^{-8}, linear scale from 0 to 1. Bottom row: planet contrast 1×10^{-7}, linear scale from 0 to 0.1. Left column: image obtained from spaxels wavelet scales number 0. Middle column: image obtained from spaxels wavelet scales number 1. Right column: image obtained from the classical spectral deconvolution approach, polynomial fitting order equal to 3. All images have been normalized with respect to the maximum peak of the planet present at position (451,538) —red larger circle at figure 4.6 —.

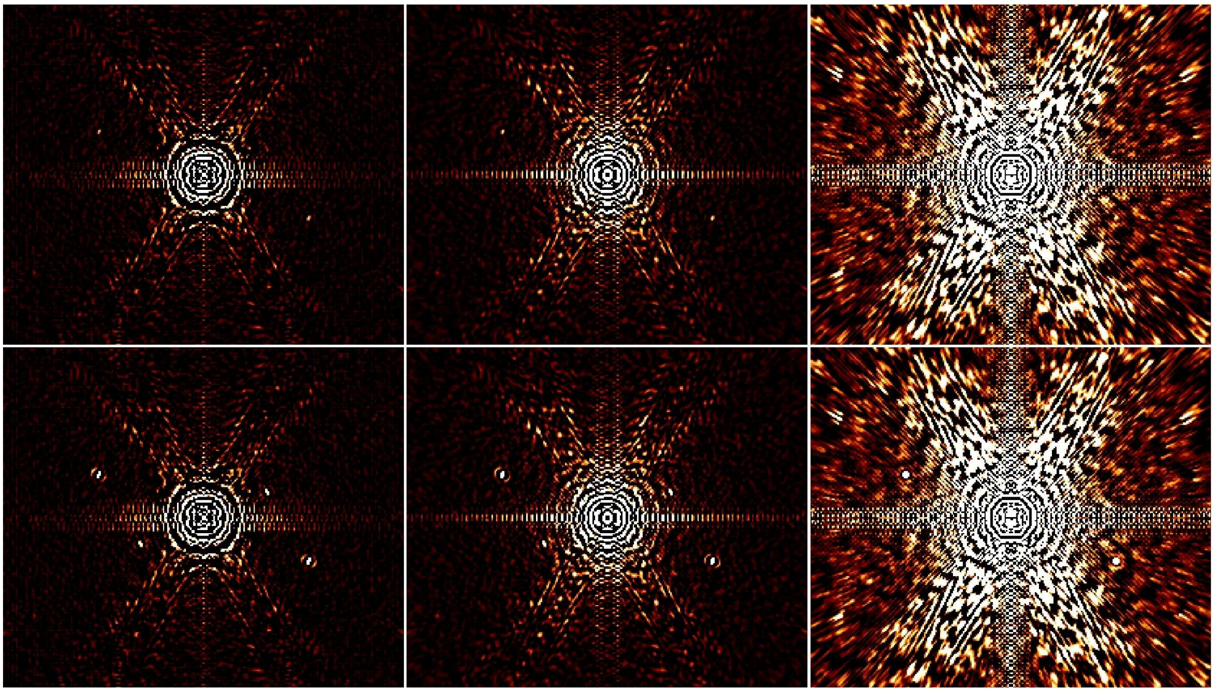

Figure 4.10: APLC case, Talbot length equal to 10. Top row: planet contrast 1×10^{-8}, linear scale from 0 to 1. Bottom row: planet contrast 1×10^{-7}, linear scale from 0 to 0.1. Left column: image obtained from spaxels wavelet scales number 0. Middle column: image obtained from spaxels wavelet scales number 1. Right column: image obtained from the classical spectral deconvolution approach, polynomial fitting order equal to 3. All images have been normalized with respect to the maximum peak of the planet present at position (451,538) —red larger circle at figure 4.6 —.

use of wavelets allows APLC chronograph to obtain similar results than APOD at angle separations from $\sim 70mas$. When wavelets are not used, noise level for APOD and APLC (Talbot length equal to 1) are, respectively, 0.0213 vs. 0.2106, i.e., ten times larger for APLC, whereas this difference is only three times larger for reconstructions obtained from the wavelet scale number 1, e.g., 0.0113 vs. 0.0323. For Talbot length equal to 10 these values are of the same order of magnitude, e.g., 0.0649 vs. 0.0696. This shows that the wavelet post-processing SD is able to classify planet frequency contribution and chromatic aberrations introduced by APLC in different wavelet scales, preserving in such way real sources frequency content, at those angle separations.

	Talbot length = 1				Talbot length = 10			
	Minimum	Maximum	Mean	Standard deviation	Minimum	Maximum	Mean	Standard deviation
plane w0	-0.1256	0.0123	-0.0563	0.0261	-0.2306	0.1880	-0.0038	0.0704
plane w1	-0.0391	0.0295	0.0033	0.0113	-0.1507	0.1876	0.0024	0.0649
plane w2	-0.0230	0.0268	0.0037	0.0090	-0.2257	0.1959	0.0051	0.0863
no wavelets	-0.0441	0.0865	0.0128	0.0213	-0.2275	0.4509	0.0758	0.1084

Table 4.1: Statistics from green square. APOD case. Planet contrast = 1×10^{-8}.

Contrast curves for these reconstructions are shown in figures 4.11 and 4.12. They have been calculated, at each angle separation, as the value which is five times above the standard deviation of the background. High-contrast imaging community use these curves to show the detection limit of their planet recovery methods. In general, the community states that “any real source which is, at certain angle separation, above the value shown from these curves can be detected” . However, such statement does not take into account the effect of the particular method over the real planet flux. In other words, classical SD tends to completely subtract the flux from the planets that are very close to the center of the FOV, since the planet signature covers the whole wavelength range through the spaxel (see figure 4.4). Hence, a contrast curve that shows a detection value of 1×10^{-5} at such positions would not be representative since, to reach this level, the planet is also removed from the reconstruction. In the case of wavelets a similar problem appears. The total planet flux is distributed along all the wavelet scales, i.e., to obtain the total flux for a certain signal all the wavelet scales and the wavelet residual must be considered (see equation 4.10). This means that a contrast curve of a certain wavelet scale is not really representative if the planet frequency content is not contributing to this scale. Furthermore, wavelet signals can also suffer from real flux removal, as the classical approach does. For these reasons, we have opted to calculate contrast curves keeping the presence of the planets so their real contribution can also be somehow visible. Besides, for the sake of clarity, contrast curves have been calculated from the non-normalized original reconstructions, and excluding the spider vanes contribution.

	Talbot length = 1				Talbot length = 10			
	Minimum	Maximum	Mean	Standard deviation	Minimum	Maximum	Mean	Standard deviation
plane w0	-0.1213	0.0522	-0.0345	0.0300	-0.2727	0.1472	-0.0070	0.0648
plane w1	-0.1284	0.0683	-0.0071	0.0323	-0.2238	0.2311	0.0005	0.0696
plane w2	-0.3395	0.1826	-0.0110	0.0801	-0.3461	0.4160	0.0008	0.1213
no wavelets	-0.5679	0.8914	0.1543	0.2106	-0.4991	0.9795	0.1425	0.2394

Table 4.2: Statistics from green square. APLC case. Planet contrast = 1×10^{-8}.

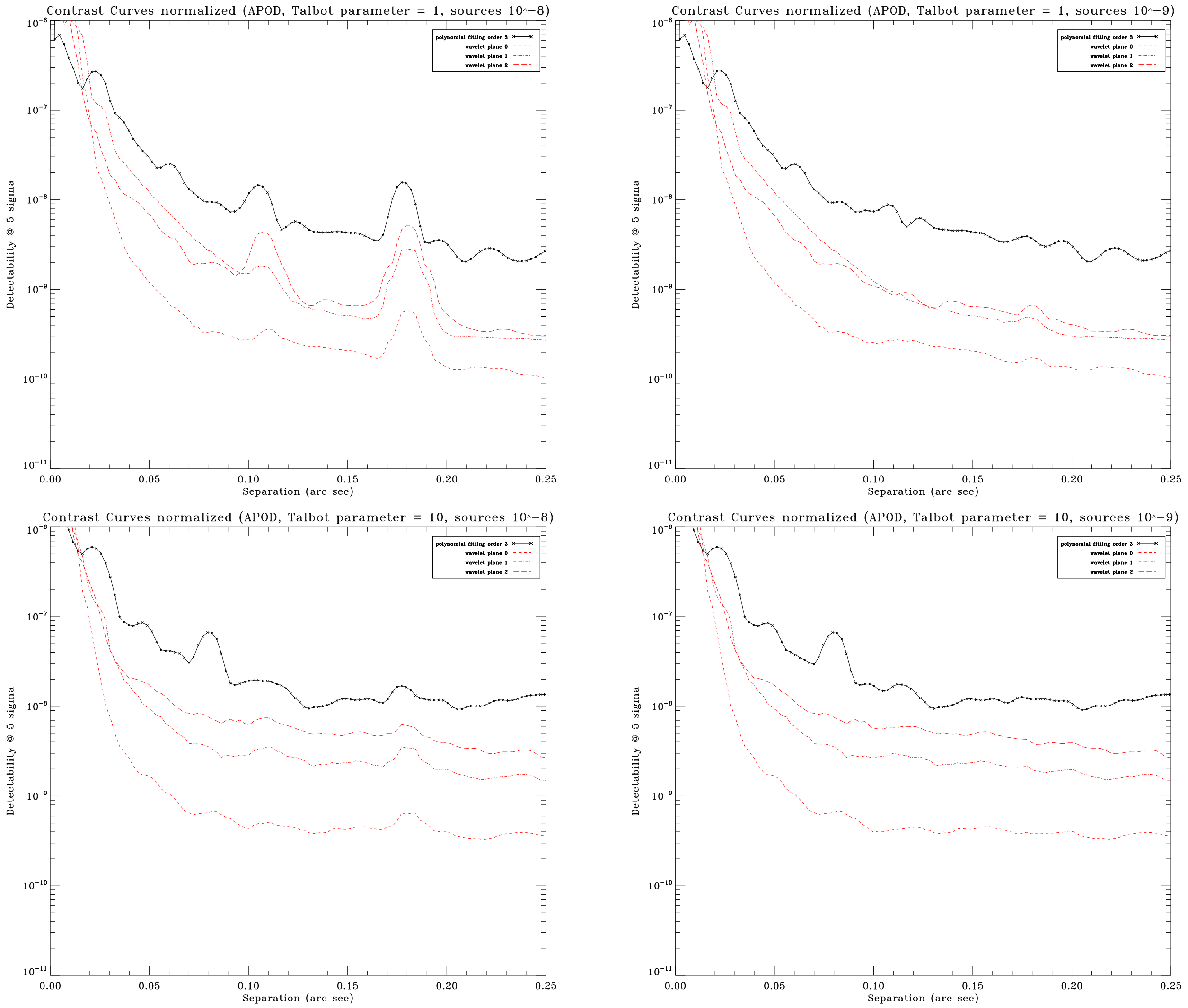

Figure 4.11: Contrast curves, APOD case. Top row: Talbot length equal to 1. Bottom row: Talbot length equal to 10. Left column: planet contrast equal to 1×10^{-8}. Right column: planet contrast equal to 1×10^{-9}. Note that bumps at 25 and 60*mas* in the black solid line (classical technique), for Talbot length equal to 1 (top row), and at 25, 50, and 80*mas* for Talbolt length equal to 10 (bottom row) do not correspond to real sources, but to wrong artificial modulation reconstructions.

Therefore, contrast curves confirm results we have previously shown, i.e., the detection limit for APOD, Talbot length equal to 1, is 1×10^{-9} (figure 4.11, top right panel) at angle separation 184*mas*, where the signatures from the planets are clearly visible in the wavelet contrast curves at such angle separation, and 1×10^{-8} for Talbot length equal to 10 (figure 4.11, bottom left panel), at the same angle separation. In the case of APLC the detection limit is 1×10^{-8} for both Talbot lengths (figure 4.12) from $\sim 100mas$. We want to stress that it does not really matter if some of the contrast curves reach a value of 1×10^{-10}. To verify if a real source can be detected at such positions is necessary to really place a source and explicitly detect it to test if the recovery method is not affecting its flux contribution to the reconstruction.

As explained in section 4.2, classical SD is based in producing along the spaxel, by scaling all the frames of a IFS datacube to a common Fourier grid, a planet signal that is larger in amplitude than the signal produced by the quasi-static speckles. Hence, a low order polynomial fits well to

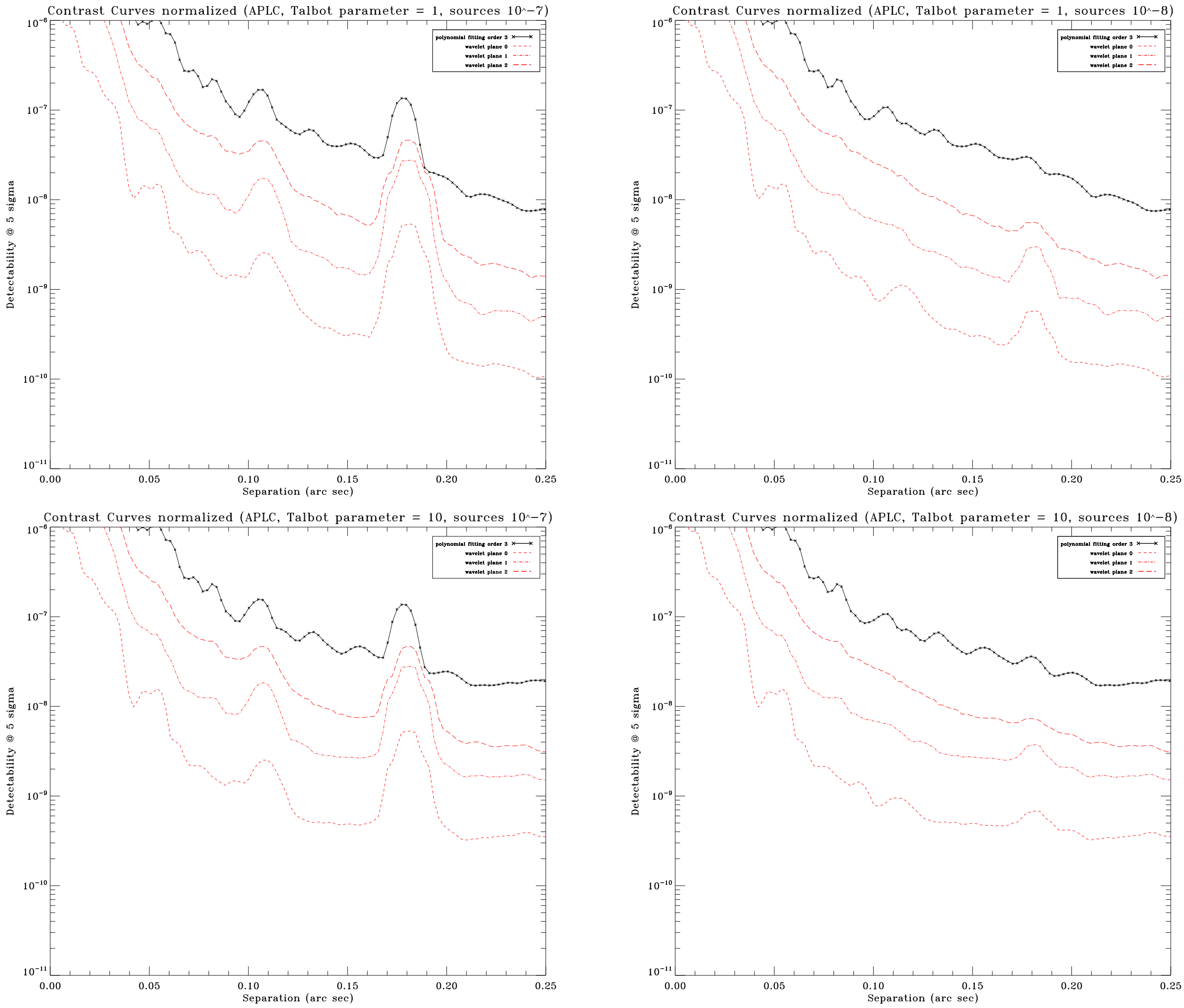

Figure 4.12: Contrast curves, APLC case. Top row: Talbot length equal to 1. Bottom row: Talbot length equal to 10. Left column: planet contrast equal to 1×10^{-7}. Right column: planet contrast equal to 1×10^{-8}. Note that the bump at 50*mas* in the red short dashed line (wavelet scale number 0), for both Talbot length equal to 1 (top row) and 10 (bottom row) do not correspond to real sources, but to a wrong artificial modulation reconstruction.

this speckle signal allowing its removal and preserving the planet signature. However, higher order chromatic aberrations that are introduced into the spaxels can also be reconstructed (see figure 4.3). Since each wavelet scale is representative of a particular frequency content through the spaxel, and since the planet signature is different of these aberrations, not only in amplitude but also in width, the WT is a powerful tool for increasing/differentiating some planet frequency components with respect to the high-order chromatic aberrations components at the same scale, thus facilitating the planet identification.

Unfortunately, for angle separations closer to the center of the FOV, planets and aberrations frequency contents are shifted to low frequencies. They can even exceed the wavelength limits of the spaxel covering the whole range of channels (see figure 4.4). Since planets, speckles and other artifacts are not longer different in frequency, the benefits of WT are diminished. Results at these separations are shown in figures 4.13 and 4.14, where four planets with contrasts from 1×10^{-7} to 1×10^{-5} lie at 30 and 58*mas* positions.

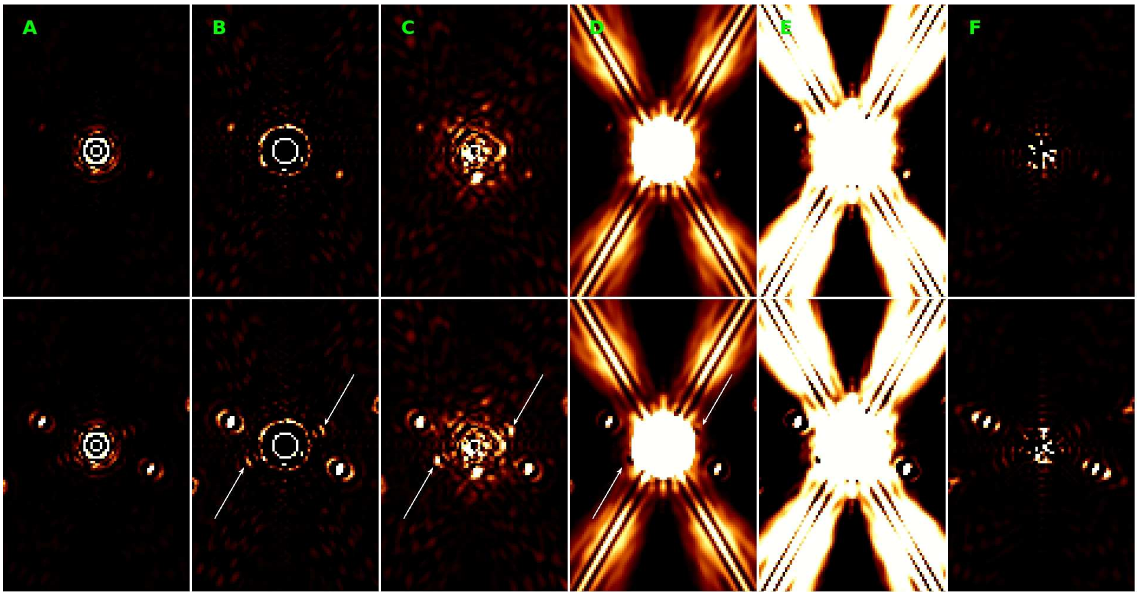

Figure 4.13: APOD case, Talbot length equal to 1. Top row: planet contrast 1×10^{-7}, linear scale from 0 to 1. Bottom row: planet contrast 1×10^{-6}, linear scale from 0 to 0.1. Column A: image obtained from spaxels wavelet scales number 2, polynomial fitting order 3. Column B: image obtained from spaxels wavelet scales number 3, polynomial fitting order 3. Column C: image obtained from the classical spectral deconvolution approach, polynomial fitting order equal to 3. Column D: image obtained from spaxels wavelet scales number 2, no polynomial fitting. Column E: image obtained from spaxels wavelet scales number 3, no polynomial fitting. Column F: image obtained from the classical spectral deconvolution approach, polynomial fitting order equal to 5 (one iteration). All images have been normalized with respect to the maximum peak of the planet present at position (451,538) —red larger circle at figure 4.6 —. White arrows point to the outer Airy disk, not planets themselves, of real sources placed at 30*mas*.

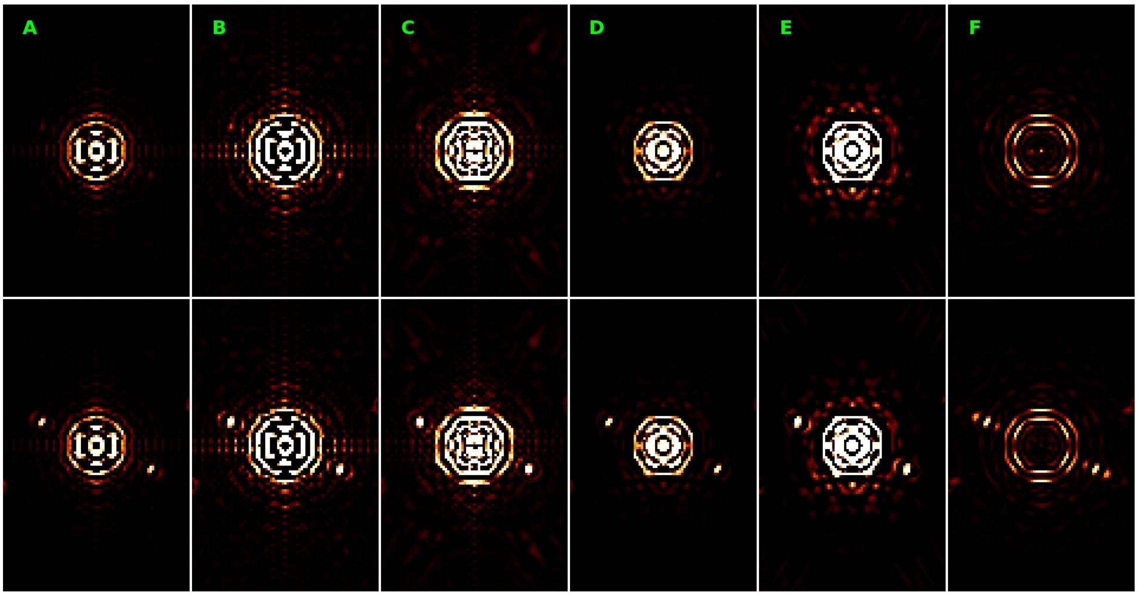

Figure 4.14: APLC case, Talbot length equal to 1. Top row: planet contrast 1×10^{-6}, linear scale from 0 to 2. Bottom row: planet contrast 1×10^{-5}, linear scale from 0 to 0.2. Column A: image obtained from spaxels wavelet scales number 2, polynomial fitting order 3. Column B: image obtained from spaxels wavelet scales number 3, polynomial fitting order 3. Column C: image obtained from the classical spectral deconvolution approach, polynomial fitting order equal to 3. Column D: image obtained from spaxels wavelet scales number 2, no polynomial fitting. Column E: image obtained from spaxels wavelet scales number 3, no polynomial fitting. Column F: image obtained from the classical spectral deconvolution approach, polynomial fitting order equal to 5 (one iteration). All images have been normalized with respect to the maximum peak of the planet present at position (451,538) —red larger circle at figure 4.6 —.

Figure 4.4 shows how the low-order polynomial fitting removes the planet signal avoiding its correct detection. Probably, the iterative mechanism explained in section 4.2 is not the most convenient for closer angle separations. Hence, we have also opted to check a different algorithm that consists in fitting a 5th-order polynomial only for one iteration. We chose a 5th-order as a good trade-off between fitting chromatic aberrations and not removing too much information from the planet. In this situation, there is no need to fit any kind of polynomial to the wavelet spaxel signals, since the WT in itself provides a natural way to remove the energy component that is always associated to the presence of speckles. Since planets are now represented by wider $sinc^2$ functions, their frequency content is classified to the low-frequency wavelet scales. Therefore, figures 4.13 and 4.14 show the reconstructions obtained from the wavelet scales number two and three.

		Polynomial fitting order 3			No polynomial fitting		
		plane w2	plane w3	no wavelets	plane w2	plane w3	no wavelets
Contrast 1×10^{-7}	planet 1	0.2364	0.6854	0.4985	0.3037	0.9793	0.2937
	planet 2	0.2469	0.7357	0.5064	0.3464	1.1920	0.2913
Contrast 1×10^{-6}	planet 1	0.2686	0.7202	0.5333	0.3135	0.8868	0.2958
	planet 2	0.2678	0.7381	0.5336	0.3174	0.9045	0.2962

Table 4.3: Maximum peak for close-in sources. APOD case (figure 4.13). Talbot length = 1.

In general, WT scale number three always has the brightest maximum peak with respect to the surrounding background, for planets placed from 43 to 58*mas*. The APOD case shows an improvement of one order of magnitude with respect to the results obtained by APLC. It is even possible to indirectly deduce the presence of planets at separations from 23 to 30*mas* by means of the detection of the outer Airy rings, pointed by white arrows in some of the panels in figure 4.13.

		Polynomial fitting order 3			No polynomial fitting		
		plane w2	plane w3	no wavelets	plane w2	plane w3	no wavelets
Contrast 1×10^{-6}	planet 1	0.3341	0.8165	0.5817	0.2363	0.4705	0.4453
	planet 2	0.3530	0.9538	0.4835	0.2460	0.5202	0.4465
Contrast 1×10^{-5}	planet 1	0.2915	0.7635	0.5416	0.3226	0.8765	0.3251
	planet 2	0.2888	0.7493	0.5280	0.3232	0.8815	0.3246

Table 4.4: Maximum peak for close-in sources. APLC case (figure 4.14). Talbot length = 1.

Tables 4.3 and 4.4 show the maximum peak values for the two planets reconstructed in figures 4.13 and 4.14. Since all the images are normalized with respect to the reconstruction of the source at 184*mas*, and all the planets at each dataset have the same contrast, that point of value equal to 1 can be used as flux reference. For close-in planets we have preferred this method for our comparison due to the different nature and large number of artifacts created during the reconstructions, by the different algorithms at such positions. The departure from the value of 1 represents the planet lies over a speckle bump or hole, and/or a certain level of bias caused by the reconstruction method, e.g., flux subtraction due to the polynomial fitting or different flux contributions to the same wavelet scale due the difference of the planet width because of the angle separation. Results from tables 4.3 and 4.4 show that planets from wavelet scale number three are a $35 - 50\%$ brighter than those obtained from the classical SD approach for APOD, and around a 30% brighter for APLC. The fact that there is no need to fit any polynomial to the wavelet signal to remove the speckle signal seems to be an advantage since the brightest maximum peaks are generally obtained when a polynomial fitting is not applied.

4.6 Conclusions

We have used the unidimensional wavelet transform (WT) to decompose integral field spectroscopy (IFS) spaxel signals at different scales of resolution, each of them representative of a certain frequency content, to improve the detectability of faint companions directly imaged with EPICS instrument, which will be installed in E-ELT. We have shown that the use of WT allows an improvement in faint companion detectability of one order of magnitude contrast from angle separations equal to $73mas$, in comparison with the results obtained with the classical spectral deconvolution (SD) approach. Furthermore, APLC performance becomes similar to that obtained with APOD for these separations. Close-in planets detectability, around $43 - 58mas$, also benefits from the application of wavelets.

In this work, we have used a dyadic and binary wavelet decomposition based on a 5-point $sinc^2(t)$ father scaling function, which allows us to decompose a 110-wavelength-point spaxel signal in 5 wavelet scales plus 1 wavelet residual. There is no need to work only with one father scaling function. We believe that the use of a dictionary of father scaling functions with different widths would be very beneficial, especially for the photometric estimation of the sources at closer angle separations, since in that way the resolution of representation of the spaxel could be increased.

In other words, a 5-point scaling function, where for IFS datacubes each point is a wavelength channel, can classify the frequency content of the signal from widths up to 5 points within wavelet plane ω_0^{f5}, from widths of 5 to 10 points within plane ω_1^{f5}, from 10 to 20 within plane ω_2^{f5}, from 20 to 40 within plane ω_3^{f5}, from 40 to 80 within plane ω_4^{f5}, and from 80 to 110 within the residual wavelet plane. With this decomposition it is not possible to differentiate, e.g., frequency content from 20 to 30 with respect to that from 30 to 40 points, since all of them are classified together in the wavelet plane ω_3^{f5}. However, using a second 30-point scaling function allows us to classify the frequency content from widths up to 30 at wavelet plane ω_0^{f30}. Therefore, a simple subtraction between different wavelet planes produced by different scaling function will provide the exact frequency information we are requiring:

$$\omega_{20\div 30} = \omega_0^{f30} - (\omega_0^{f5} + \omega_1^{f5} + \omega_2^{f5}), \tag{4.12}$$

where $\omega_{20\div 30}$ is the signal representative of the spaxel content with frequency information from 20 to 30 points width. The study and application of this sort of dictionary will be the topic of our future research. We expect that the increment of the spaxel resolution in such way yields an improvement in the photometric estimation of companions, as well as in their spectral characterization.

Chapter 5

Summary and Conclusions

Adaptive optics (AO) systems are used to increase the spatial resolution achieved by ground-based telescopes, which are limited by the atmospheric motion of air layers above them. Therefore, the real cut-off frequency is extended closer to the theoretical diffraction limit of the telescope thus allowing more high-frequency information from the object to be present in the image (chapter 1).

Nevertheless, although the goal of image reconstruction and deconvolution algorithms is basically the same (i.e., to recover a “real” diffracted limit image, free of noise, from the object), and since the correction of AO is not complete (i.e., the effective cut-off frequency achieved by AO is still below the theoretical diffraction limit), the simultaneous use of such deconvolution algorithms over dataset acquired with AO is possible and desirable to further enhance their contrast (chapter 1, sections 1.3 and 1.4.1).

Furthermore, multiresolution tools like the wavelet transform (WT) have been historically incorporated into multiple deconvolution schemes improving their performance with respect to their non-wavelet counterparts. The ability of such transforms to separate image components depending on their frequency content results in solutions that are generally closer to the real object (Starck et al. (2002) and appendix D). On the other hand, AO community generally states that, due to the high variability of AO PSFs is necessary to update the PSF estimate during the reconstruction process. Hence, the use of blind and myopic deconvolution algorithms should be unavoidable and yields to better results than those obtained by the static-PSFs codes (Pantin et al. 2007).

Therefore, being the aforementioned paragraphs the current state-of-art of AO imaging, the main conclusions of the present thesis are outlined as follows:

1. The static-PSF algorithm AWMLE has been applied over binary systems simulated for the 3-m Shane telescope to evaluate the photometric accuracy of the reconstruction. Its performance is compared with the PSF-fitting algorithm StarFinder, commonly used by the AO community. Results shown that AWMLE is able to produce better results than StarFinder (chapter 2).

2. In addition, AWMLE has been tested together with several other algorithms and approaches, like FITSTAR, PDF deconvolution and IDAC, for the same purpose of differential photomet-

ric precision, resulting in very competitive results, especially for high Strehl ratios (SR) and matched PSFs (appendix A).

3. A new methodology for statistically testing the photometry of these codes has been designed for AO observations. It consists of a double test which measures the average photometric accuracy and the repeatability of the results (chapter 2, section 2.5).

4. A new deconvolution algorithm called ACMLE, which is based on the curvelet transform (CT) and a maximum likelihood estimator (MLE), has been designed for the reconstruction of extended and/or elongated objects (chapter 3, section 3.2.2).

5. ACMLE has been tested together with AWMLE and blind/myopic codes such as MISTRAL and IDAC over Saturn and galaxy simulated images for the 5-m Hale telescope. It is shown that the performance in the presence of noise of the multiresolution static-PSF algorithms is better than myopic and blind algorithms; thus showing that the control of noise is as important as the update of the PSF estimate during the reconstruction process (chapter 3, section 3.5).

6. In the challenging context of horizontal path, several Richardson-Lucy algorithms, including AWMLE (static-PSFs approaches), have been tested with MISTRAL (myopic-PSF), showing comparable results (appendix B).

7. A unidimensional WT, based on a $sinc^2$ father scaling function, has been applied in the spectral deconvolution of integral field spectroscopy (IFS) datacubes for direct imaging of exoplanets with EPICS instrument, which will be installed at the forthcoming 39-m E-ELT telescope. When this approach is compared with the classical non-wavelet one, an improvement of $1mag$ from angle separations equal to $73mas$ is devised. Besides, detection of close-in planets, between 43 and $58mas$ also benefit from the application of wavelets (chapter 4).

8. The use of wavelets allows the APLC chronograph to obtain similar results with respect to the apodizer-only solution, especially with increasing Talbot length, thus showing that WT classify planet frequency components and chromatic aberrations at different scales (chapter 4, section 4.5).

9. Preliminary results for HARMONI spectrograph are also described (appendix C).

A list with all the publications present in this work is shown as follows:

Baena Gallé, R. & Gladysz, S. Estimation of differential photometry in adaptive optics observations with a wavelet-based maximum likelihood estimator. 2011, Publications of the Astronomical Society of the Pacific, 123(905), 865-878. Press here to download.

Gladysz, S., Baena Gallé, R., Christou, J.C. & Roberts, L.C. Differential photometry in adaptive optics imaging. 2010, Proceedings of AMOS Technical Conference (Maui, Hawaii), E24. Press here to download.

Baena Gallé, R., Núñez, J. & Gladysz, S. Extended object reconstruction in adaptive-optics imaging: the multiresolution approach. 2013, Astronomy and Astrophysics, 555(A69). Press here to download.

Gladysz, S., Baena Gallé, R. Comparison of image restoration algorithms in the context of horizontal-path imaging. 2012, Proceedings of the International Society for Optics and Photonics SPIE. Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XXIII. ed: Gerald C. Holst, Keith A. Krapels, Proceedings Vol. 8355, 83550X. Press here to download.

Baena Gallé, R., Gladysz, S., Verinaud, C. & Kasper, M. High contrast exoplanet imaging using spectral deconvolution and the wavelet transform. Application to EPICS instrument. 2013, *In preparation.*

Gladysz, S., Thatte, N.A., Salter, G., Baena-Gallé, R., Clarke, F., Tecza, M. & Jolissaint, L. High-contrast, adaptive-optics simulations for HARMONI. 2nd International Conference on AO for ELT (Victoria, Canada). 2011. Press here to access the abstract on the official Conference website.

5.1 Future work

This thesis opens several lines of research that will be addressed in future:

- The world of multiresolution transforms is extremely huge and has produced dozens of new mathematical tools. Among many other, it is worth to mention the shearlet transform, which is an extension/improvement of CT, and the waveatom tool, which is intended to classify textures in the image. They should be studied and compared to establish their best performance and their best field of application over AO images.

- Blind and myopic algorithms have proved their ability for large mismatches between the “real” PSF that has created the image and the PSF that is used as a first estimate in the reconstruction process. However, their performance in the presence of noise is highly affected. Hence, it is convenient to investigate if it is possible to introduce (and how to do it) multiresolution transforms into these algorithms to improve their behavior.

- For the study of IFS datacubes, other father scaling functions with different shapes could be proposed, in particular, it can be considered a “dynamic” scaling function with the ability to modulate its shape according to the low frequency signal to be removed from the spaxel. This could potentially improved the final photometry of the detected faint source.

- Besides, the design of a dictionary of wavelets, which increase the decomposing resolution across the spaxel, instead of a single dyadic decomposition, can improve the photometric accuracy of detected planets as well as their spectral characterizations, taking full advantage of the information contained in the IFS datacubes.

Appendix A

Differential Photometry in Adaptive Optics Imaging

Published in Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference (AMOS Conference). 2010, ed.: S. Ryan, The Maui Economic Development Board., p.E24.

Abstract

One application of adaptive optics (AO) is high-resolution imaging of closely-spaced objects. Determining differential photometry between the two or more components of a system is essential for deducing their physical properties such as mass and/or internal structure. The task has implications for (i) Space Situational Awareness, such as the monitoring of fainter microsatellites or debris nearby a larger object, and (ii) astronomy such as the observations of close stellar faint companions. We have applied several algorithms to the task of determining the relative photometry of point sources with overlapping point spread functions in images collected with adaptive optics. These algorithms cover a wide range of approaches in the field of image processing. Specifically we have tested: PSF-fitting, multi-frame and single-frame blind deconvolution, maximum-likelihood approach combined with wavelet decomposition, and a novel one-dimensional deconvolution technique which separates signal and speckle statistics rather than integrated intensities. We present results from applying these algorithms to synthetic close binary stars for different observing conditions.

A.1 Introduction

Atmospheric turbulence imposes a limit on angular resolution which could be reached by ground-based telescopes. The application of astronomical adaptive optics (AO) during the last couple of decades has allowed diffraction-limited images, rather than seeing-limited, to be obtained with large ground-based telescopes. One of the uses of AO is high-resolution imaging of closely-spaced objects, e.g. binary stars or faint companions such as exoplanets. Determining the differential photometry and astrometry between the two components of the system is essential for deducing the physical properties of the components such as mass or internal structure (Vigan et al. 2010). AO, with a suitably bright guide star, improves the detectability and photometric accuracy but also introduces problems which are not usually encountered in conventional seeing-limited photometry (Esslinger & Edmunds 1998):

1. The structure of the PSF has temporal variation due to seeing variability. These morphological changes are difficult to model.

2. The AO long-exposure point spread function (PSF) in the medium- and high-correction regime shows long-lived quasi-static speckles. These diffraction-limited “lumps” are due to residual aberrations not sensed by AO (for example non-common-path errors) and lie in the halo surrounding the core of the PSF.

3. 1.The AO corrected PSF, and the associated angular resolution on the sky, depend on the position of the science object relative to the AO guide star. AO compensates for the turbulence in the direction of the guide star and when the science object is well separated from the AO line of sight, the compensation suffers due to a different atmospheric volume. This effect is called angular anisoplanatism.

Because of these factors, extracting quantitative information from AO images is challenging. AO improves the detectability of faint companions over seeing-limited observations for a given telescope. When the companion is well separated with a non-overlapping PSF, aperture photometry takes care of all the problems mentioned above. The problems occur when the PSFs from each target overlap.

Esslinger & Edmunds (1998) provide an excellent introduction to the problem of AO photometry. One of the issues discussed is the precision of photometry on deconvolved images compared with the PSF-fitting on the “raw” AO data. Two of the most-widely used deconvolution algorithms were tested, namely maximum entropy (Frieden 1972) and Richardson-Lucy (Richardson 1972; Lucy 1974). For PSF-fitting the DAOPHOT package (Stetson 1987) was employed. Extensive testing on simulated faint companions showed that DAOPHOT performed consistently better than the deconvolution methods, i.e. its photometric precision was higher compared to precision of aperture photometry on the deconvolved images.

It has been suggested that AO observations should be processed with “myopic” deconvolution methods because the PSF is not well known for most AO observations (Pantin et al. 2007). For myopic deconvolution, it is assumed that the PSF is only partially known. In some cases the PSF is unknown and this is the regime for “blind” deconvolution. Typically these algorithms require an initial PSF estimate. This estimate is assumed to be close to the truth and the algorithm iterates this estimate until a common solution for both the object and the PSF is found. Myopic and blind deconvolution techniques often use regularization, e.g. by imposing object priors and PSF constraints (Mugnier et al. 2004). Jefferies & Christou (1993) developed an iterative blind deconvolution method guided by the minimization of a penalty functional. An important component for AO observations is the penalization of spatial-frequencies beyond the diffraction-limit of the observations. This package (IDAC) restores a “clean” image from which relative photometry and astrometry can be obtained by using aperture photometry with a very small aperture (Esslinger & Edmunds 1998), or by model fitting. IDAC is one of the algorithms we have tested in our work.

PSF-fitting algorithms are applicable for crowded fields where the target comprises only point sources. In this approach an analytic or empirical PSF is used together with a fitting algorithm to match scaled-and-shifted copies of the PSF to the data. One such package suitable for AO imaging is StarFinder (Diolaiti et al. 2000) which yields relative photometry and astrometry in AO-corrected crowded stellar fields. A PSF model is constructed from the brightest stellar images in the field. The algorithm iteratively uses this model to locate fainter sources which it then fits to extract the relative photometry and astrometry. There are three degrees of freedom: the total flux, and x and y positions for each component. The photometric and astrometric precision of StarFinder applied to crowded fields has been compared to results from blind deconvolution with the IDAC algorithm by Christou et al. (2004). We also test StarFinder’s photometric accuracy in this paper.

A recent review of modern approaches to AO photometry (Burke et al. 2009) has demonstrated that methods utilizing static, deterministic PSF should not be dismissed against more modern algorithms like iterative blind deconvolution (ten Brummelaar et al. 2000). Thus in our set of methods to test we have included a new implementation of a Richardson-Lucy type deconvolution. This algorithm, Adaptive Wavelets Maximum Likelihood Estimator -AWMLE- (Otazu 2001), calculates an image that maximizes the compound Poisson and Gaussian likelihood of the data. It also performs wavelet decomposition that helps distinguish signal from noise which is important for improving the stopping rule. Unlike myopic or blind deconvolution AWMLE does not update the PSF so that it could be more dependent upon the initial PSF estimate.

A novel approach for measuring photometry of faint companions in AO imaging has been recently proposed (Gladysz & Christou 2009; Gladysz et al. 2010b). In this method traditional 2-D image deconvolution is replaced by a 1-D time-series deconvolution. The algorithm is based on the observation that the statistical distribution of the peak of AO-corrected PSF is morphologically different from that of the off-axis light, i.e. the quasi-static speckles. This morphological difference between the two probability density functions (PDF) is used to constrain a one-dimensional, “blind”, iterative deconvolution at the position of a faint companion to a star. Separation of the signal and speckle PDFs yields the differential photometry. The method ("PDF deconvolution"), has been successfully applied to medium-, and very-high-resolution AO observations (Gladysz & Christou 2009; Gladysz et al. 2010b). We note that this PDF deconvolution has a rather narrow range of applications as opposed to the other algorithms discussed here (it only produces photometry for AO observations of companions which lie within the uncorrected halo structure of the AO PSF). This algorithm requires the companion's location to be known. This can be obtained with a matching reference-less astrometric method (Gladysz & Christou 2009) which then takes full-advantage of the self-calibrating nature of PDF deconvolution.

The goal of this paper is to focus on the description and usage of existing codes, and to compare their photometric precision after application to AO data in a "blind"test. We do not discuss all the issues pertaining to computing photometry in AO observations. Discussions which go into great depth on these subjects can be found in the literature (Esslinger & Edmunds 1998; Roberts et al. 2005).

A.2 Methods

In this chapter we compare results from the following algorithms, all of which were discussed above:

1. StarFinder: A PSF-fitting algorithm where the user-supplied PSF is iteratively fitted to the data assuming a double-delta object.
2. AWMLE: A Richardson-Lucy type approach with a static PSF.
3. IDAC: Multi-frame blind deconvolution.
4. FITSTARS: A single-frame iterative blind deconvolution.
5. PDF deconvolution: Using speckle statistics.

We have analyzed these algorithms for different AO correction scenarios. The AO correction is typically described by the Strehl ratio (SR). This is the peak of the AO-corrected PSF normalized to that of an ideal PSF for the same pupil. There are four scenarios: low vs. medium SR and "matched" vs. "mismatched" PSF. For the matched cases, the initial PSF has a similar SR to that of the observation and for the mismatched cases the initial PSF has SR with a difference of 6%. A detailed description of the data is given in Section A.3.

A.2.1 StarFinder

StarFinder was developed to measure astrometry and photometry in crowded fields imaged with AO (Diolaiti et al. 2000). The algorithm operates as follows. Firstly, it derives a PSF template from the brightest isolated field stars and generates a catalogue of presumed objects by searching for the relative intensity maxima in the frame. Secondly, the images of the candidate stars are analyzed in order of decreasing peak intensity and each candidate is accepted on the basis of its correlation coefficient with the PSF template. The relative astrometry and photometry of each source are determined by means of a least-squares fit, taking into account the contribution of the local non-uniform background and of the already detected stars. These steps are repeated until no sources can be reliably found in the residuals. A thorough description of the algorithm, the IDL routines, GUI interface and excellent documentation can be found on the StarFinder's website[1]. For the results presented here, we used neither the graphical interface nor the capability of StarFinder to extract PSF from the data. Instead the algorithm was supplied with the observation, a "known" PSF (a single star observed after the target) and approximate positions of the two sources in the image via the FITSTARS.pro subroutine[2].

A.2.2 AWMLE

AWMLE (Otazu 2001; Fors 2006) uses (i) Bayesian maximum-likelihood approach, (ii) wavelet transform -WT- (Daubechies 1988), and (iii) multi-resolution support (Starck & Murtagh 1994). The first maximizes the likelihood between the dataset and a possible solution by considering a combination of the intrinsic Poisson noise of the signal and the read-out Gaussian noise of the detector as well as describing the optical path by a static PSF that remains constant throughout the reconstruction process. The second decomposes the dataset into wavelet scales by means of the à trous algorithm (Shensa 1992). The WT lets any signal or image be represented in N scales. This leads to simultaneous representation in both the measurement and frequency spaces. In general, the noise will mainly appear in the high spatial-frequency wavelet plane while broad shapes will appear in the low spatial-frequency planes. For AO observations, this permits differentiation between the diffraction-limited features (given by λ/D) and the seeing-limited scale (given by λ/r_0). An example of wavelet decomposition applied to an AO PSF is shown in Figure A.1. The third analyzes each wavelet scale in order to find significant WT coefficients that can be associated with real signal in an image. The standard deviation of intensity within a local window (whose size depends on the wavelet scale we are analyzing) is compared with the standard deviation of the whole plane. Their difference allows one to deduce the presence of a real source. This also helps to automatically stop reconstruction of the image at each wavelet scale independently. AWMLE was applied to the datasets using one wavelet scale plus a residual wavelet scale. A maximum number of 100 iterations were enough to achieve convergence in both scales. Note that AWMLE is not a photometric package. It produces a deconvolved image which can be subsequently analyzed by an observer. After reconstruction, aperture photometry, with a circle of five pixels in diameter, was used to extract the brightness of each component.

---

[1] http://www.bo.astro.it/ giangi/StarFinder/

[2] Not to be confused with the FITSTARS blind deconvolution code discussed later

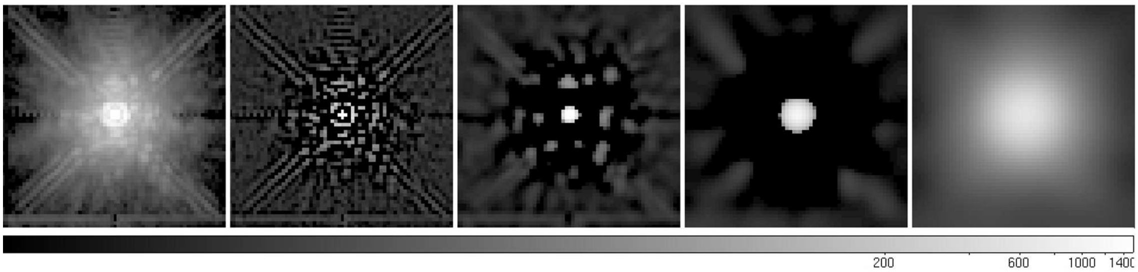

Figure A.1: Wavelet decomposition of the AO PSF. Left-to-right: AO PSF; three wavelet planes with decreasing spatial frequency and the lowest-frequency wavelet plane (the wavelet residual). Displayed on a logarithmic scale.

A.2.3 IDAC

IDAC is a multi-frame blind deconvolution (MFBD) algorithm[3]. Basically, it is an iterative least squares algorithm using a number of constraints to solve for both the common object (target) intensity distribution and also the corresponding PSFs for multiple observations of the same target (Jefferies & Christou 1993; Christou et al. 2004). MFBD algorithms are very successful in the case of strongly varying PSFs such as the pure speckle imaging case so that the target is easily distinguished from the PSFs. For AO data, the goal is to stabilize the PSF. This implies less PSF diversity from one observation to another so that other constraints become more useful.

In general, the algorithm makes no assumption of the target's intensity distribution and extent and the deconvolved image is computed for the full image plane. However, for the synthetic observations described in this paper, we have the prior knowledge that the target comprises two point sources and we also know their locations. What we do not know is the relative brightness between the two. In order to constrain the target to a binary star model, the initial target estimate comprises two narrow symmetric Gaussians centered on the pixel locations of the two targets, each having a FWHM = 1.75 pixels. This takes into account potential sub-pixel locations of the components and permits the algorithm to "shift" the component locations in order to obtain the best common fit. The initial intensity ratio of the two Gaussians is estimated from the corresponding pixel values in the observations and the initial PSF estimate was the "known" PSF described in Section A.3. The PSF band-limit is typically measured from the data and read-noise limit was determined from "sky" regions of the observations. For this application ten independent observations were used for the multi-frame constraint. The algorithm was allowed to converge for $\sim$ 100 iterations from the initial start-ups sharpening the Gaussian distributions of the two components and, more importantly, adjusting the relative amplitudes of both to allow the reconstructed target to match the ten individual data frames.

Like AWMLE, IDAC produces a final image from which photometric and astrometric measurements are made. For the binary cases here, the reconstructed object was fit by two Gaussians using a least squares method after a further Gaussian smoothing. This smoothing reduced the pixelation allowing for an improved fit. The free parameters for the fits were the amplitudes, elliptical Gaussian widths σ_x and σ_y, the position angle orientation of the ellipse and the x and y locations of each Gaussian - a total of 12 parameters in all. The intensity ratio was obtained from the ratio of the Gaussian

---

[3] http://cfao.ucolick.org/software/idac/

volumes, (i.e. $V_n = I_n\sigma_x\sigma_y$).The advantage of the fitting is that it allows an uncertainty measurement for each of the intensity ratio.

A.2.4 FITSTARS

FITSTARS (ten Brummelaar et al. 2000) is a single-frame iterative blind deconvolution algorithm optimized for binary stars by defining the object distribution as two δ-functions:

$$o(x,y) = \sum_{i=1}^{2} A_i\delta(x - x_i, y - y_i), \tag{A.1}$$

where $(x-x_i, y-y_i)$ is the location of the ith component and A_i its intensity so that there are a total of six object variables to fit. These binary parameters are solved for by using least squares fitting to the observations using the initial target and PSF estimates. Once an initial estimate of these variables is obtained, an updated PSF estimate for each component is computed by differencing the measurement to the model for the other component and a new PSF estimate is then computed from the weighted average of these two individual component estimates (ten Brummelaar et al. 1996). This process is repeated until the results converge. Results using this algorithm have been compared to those obtained by speckle interferometry (Horch et al. 2001) and are in good agreement.

A.2.5 PDF deconvolutions

PDF deconvolution uses the analytical forms of the PDFs of the on-axis and off-axis intensity in an AO PSF (Gladysz et al. 2010b). The instantaneous Strehl ratio has a distribution characterized by two parameters (i) number of independent cells in the AO-corrected wavefront, and (ii) the theoretical long-exposure Strehl ratio, which is related to the statistical phase variance via the extended Marchal approximation (Hardy 1998). When the companion is located within the isoplanatic patch, both components of a binary star are produced by almost the same wavefront. Thus, the two parameters mentioned above are common for the PDFs of the peak intensity of the star and its companion. However, at the location of the companion, the speckle and signal intensities add, and their PDFs are convolved. The distribution of the signal (i.e. the “raw” peak intensity) has the same form as the Strehl ratio PDF (Gladysz et al. 2010b) but is “blurred” by the speckle kernel. PDF deconvolution blindly estimates both the signal and the speckle PDFs from a vector of intensity measurements at the location of the companion.

This 1-D deconvolution problem has five parameters to solve for but can be very easily constrained by observing that two parameters are common for both objects, i.e. those of the instantaneous SR. These are obtained from a least-squares fit of the theoretical PDF to the on-axis statistics for the bright star so that, at the location of the companion, the algorithm only searches for the three remaining parameters. After successful separation of signal from speckle statistics differential photometry can be obtained by comparing flux parameters estimated for the two objects.

PDF deconvolution relies on multi-frame observations of the object. The other inputs are the estimated Strehl ratio of the observations, and the position of the companion (assumed to be known from astrometry). Figure A.2 illustrates how the method works. First, the theoretical on-axis PDF is fit to the measured on-axis histogram (left panel), and then the convolution of two PDFs is fit to the measured histogram at the location of the companion (right panel). The ratio of the widths of the two distributions of interest (red curve in the first panel and green curve in the second panel) can be converted to brightness ratio of the two objects.

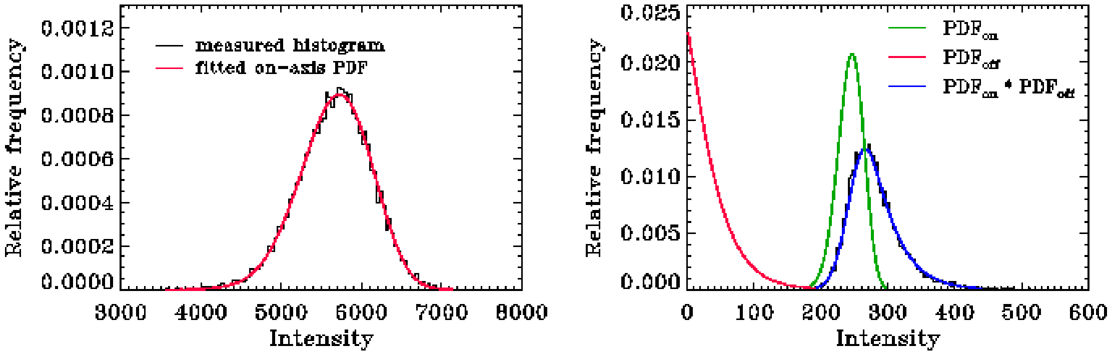

Figure A.2: PDF deconvolution. Left: fit of the theoretical on-axis PDF to the measured histogram. Right: separation of the PDFs at location of a companion. $\mathrm{PDF}_{\mathrm{on}}$ is the on-axis distribution; $\mathrm{PDF}_{\mathrm{off}}$ is the speckle distribution. Symbol $*$ denotes convolution. This plot corresponds to a typical case encountered in tests on the Lick data, as described in Section A.3.

In this work the PDF deconvolution algorithm is supplied with the known positions of the companions which it does not update. When the peak of the companion's PSF is at a sub-pixel location, the method is given a non-integer location of the companion and it spatially interpolates the measured pixel values to extract the intensity time series for the PDF.

A.3 Description of the data

We used single-star data sets, obtained with the Lick Observatory AO system on the 3m Shane telescope to generate synthetic binary stars. AO images of bright stars were obtained using the high-speed sub-array mode (64$\times$64 pixels), for the 256$\times$256 pixels IRCAL camera, which corresponds to a field size of 4.8 × 4.8 arcsec. The sub-array measurements were captured with typical exposure time of 22ms. Each data set comprised 10000 images. All data were obtained in K band (2.2μm) where the diffraction limit is 151mas and the data were effectively Nyquist sampled (two pixels per λ/D). The individual short exposures were registered with sub-pixel accuracy to produce shift-and-add (SAA) images. The average Strehl ratio of these SAA images was $\sim$ 40%. All data were sky-subtracted and the residual background was then subtracted too.

Simulated binary star data sets were created by scaling and shifting the PSF datacubes yielding synthetic observations of a binary star with a brightness ratio of 25 (magnitude difference, $\Delta m = 3.5$) which was chosen to create challenging scenarios, with $5 < SNR < 15$ for the companion. We placed the companion at one of eight different positions in order to minimize variations in results due to possible anisotropies in the PSF. The positions were $\sim$ 0.6 arcsec ($4\lambda/D$) from the center of an image (see Figure A.3) with four positions in a cross 8 pixels away horizontally or vertically from the bright star, and four diagonally (7 pixels horizontally and 4 pixels vertically). The difference between the “straight” and “diagonal” separations is 0.062 pixel. The mean photometric error was computed based on the results from these eight positions.

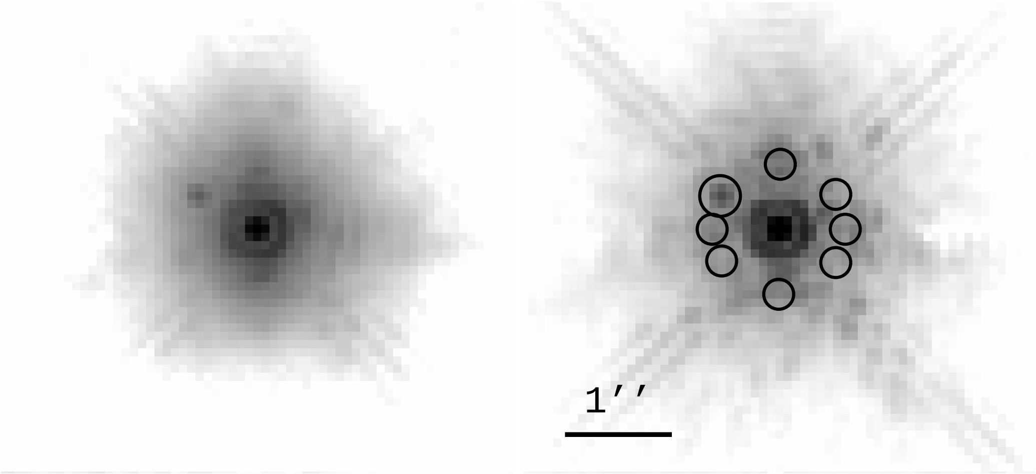

Figure A.3: SAA image of a synthetic binary star with separation of $4\lambda/D$. Left: High SR ($\sim$ 50%) showing the eight locations for the artificial companion. The companion is located within the larger circle. Right: SAA image of the synthetic binary star for SR $\sim$ 30%. Note the presence of significantly greater residual speckle structure for SR$\sim$ 50% whereas the SR$\sim$ 30% shows a more uniform halo structure. The images are displayed on logarithmic scale.

All algorithms were supplied with re-centered images. PDF deconvolution used a 10000 frame datacube, IDAC used 10 data cubes, each comprising 1000 co-added frames, while the other codes used single SAA images of all 10000 frames. All methods, except PDF deconvolution, rely on a PSF estimate. For the matched-PSF cases we used the same stars observed ten minutes later (Strehl ratio mismatch = 2 or 3%). For the mismatched-PSF cases we used stars of similar brightness observed the same night, and also close to zenith (Strehl ratio mismatch = 6%; for the case of $\sim$ 30% SR the calibrator had higher SR than the target while the reverse situation was tested for 50% SR). Variability of the Strehl ratio between the science and calibration PSF is a direct consequence of either non-stationarity of turbulence (if the same star was used for target and calibration datasets), or change in response of the AO system due to lower or higher photon flux coming from the calibrator, as compared to the target. Table A.1 summarizes the grid of scenarios we have investigated.

Table A.1: PSFs used to simulate images of binary stars. The SR = 29% data refers to the 30% Strehl ratio case, and the SR = 54% data to the Strehl ratio = 50% case in the text.

		Science PSF	SR	m_v	m_k	Reference PSF	SR	m_v	m_k
			30% Strehl ratio						
Matched PSF	....	NOMAD1 1297-0510182	29%	12.1	5.93	–	32%	–	–
Mismatched PSF	....	NOMAD1 1297-0510182	29%	12.1	5.93	HD 18009	35%	8.23	5.02
			50% Strehl ratio						
Matched PSF	....	HD 143209	54%	6.3	3.92	–	52%	–	–
Mismatched PSF	....	HD 143209	54%	6.3	3.92	HD 153832	48%	7.25	4.78

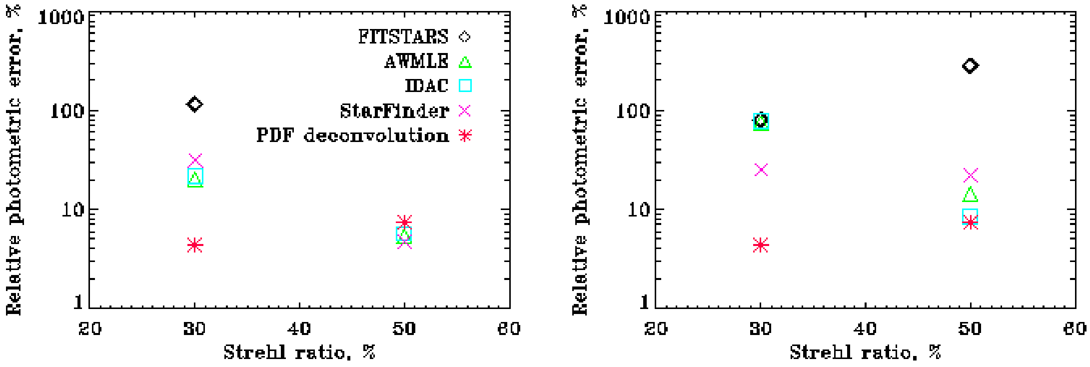

Figure A.4: Mean relative intensity ratio error for the five algorithms (equation A.3). Left: PSF well-matched to the observations. Right: mismatched PSF. The results for PDF deconvolution are identical in both panels because this method does not rely on a PSF estimate. The $p2$ value for FITSTARS in the 50% SR, matched-PSF case was very high (~ 10000) and we omitted it from the plot in order to have the y-axis scale which better shows differences between the other methods.

In order to determine the efficacy of each algorithm we used a metric of photometric precision which measured the mean absolute deviation from the truth (Burke et al. 2009):

$$p_1 = \frac{1}{8}\sum_{j=1}^{8} |I_j - truth| \,, \tag{A.2}$$

where I_j are the individual intensity ratio measurements and the truth was equal to 25. This metric was then converted to percent relative error:

$$p_2 = \frac{p_1}{truth} \times 100\% \,, \tag{A.3}$$

This metric shows the relative strength of the average departure from the true intensity ratio. In Figure A.4 we plot p_2 for the scenarios of well matched and mismatched PSF and in Table A.2 we give numerical values of this metric.

Figure A.4 illustrates relative precision of various algorithms. In order to discuss possible biases (systematic over-, or under-estimation of the truth) and dispersions of results, we plot means and standard deviations in Figure A.5 and we give numerical values of these metrics in Table A.3.

Table A.2: Percentage photometric precision of the algorithms, as quantified by the metric p_2.

		FITSTAR	AWMLE	IDAC	StarFinder	PDF deconvolution
30% SR, matched PSF	....	116	20	21	31	4
50% SR, matched PSF	....	10848	5	6	5	7
30% SR, mismatched PSF	....	80	75	77	25	4
50% SR, mismatched PSF	....	284	14	8	22	7

A.4 Summary and discussion

We compared photometric measurements of various algorithms for a case of a very close binary star with a relatively large intensity ratio. This is a particularly challenging case in that the companion is 25 times fainter than the primary and lies well within the PSF morphology of the primary ($\sim 4\lambda/D$). This is where the speckle contribution is non-negligible and the best photometric precision (Figure A.4) is on the order of couple of percent relative to the truth. For each observing case there were eight different realizations with the companion located in a different region of the primary's PSF and the dispersion in the results reflect the sensitivity of the algorithms to measure the photometry with differing speckle backgrounds and morphology.

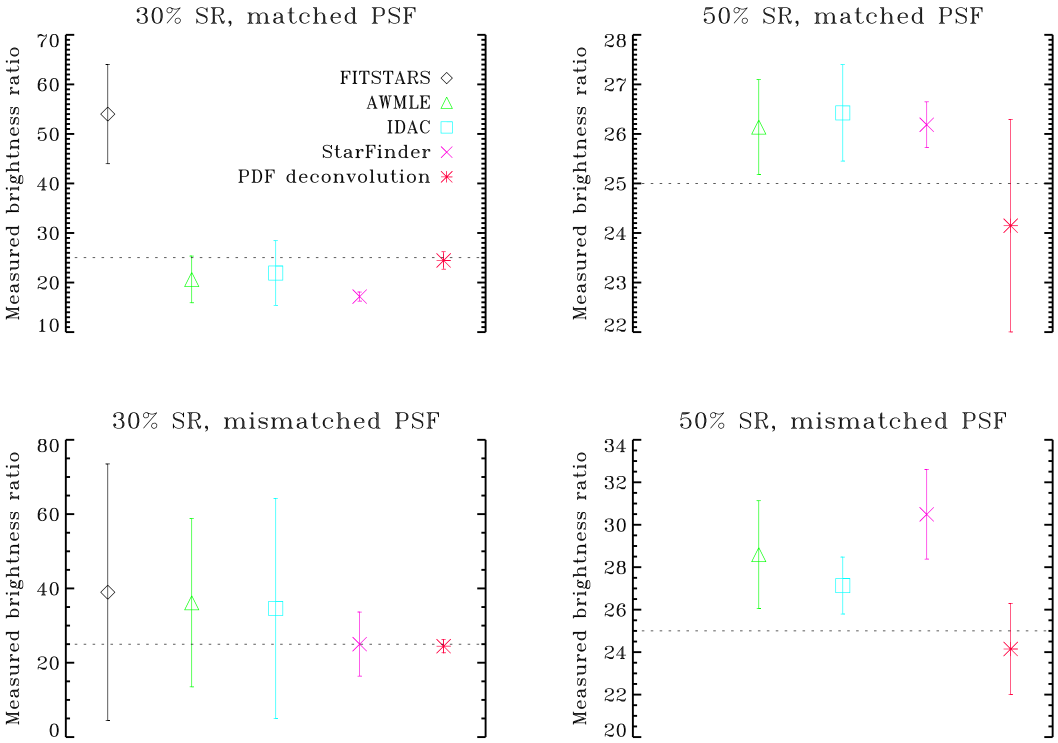

Figure A.5: Means and standard deviations of the measured brightness ratios. The dashed line corresponds to the true value of 25. The values for FITSTARS in the matched-, and mismatched-PSF cases for 50% SR were very high and we omitted them from the plots in order to emphasize differences between the other methods. The PDF deconvolution, being self-calibrating was independent of the PSF and therefore it produced the same results for the matched and mismatched PSFs. (Note the differences in the vertical scales).

Looking at the results in Figure A.5 and Table A.3, one notices that for the SR=50% case, the mean IDAC, StarFinder and AWMLE results are very similar to each other and consistently give a larger intensity ratio by $\sim$ 4% for the matched PSF and by $\sim$ 7% for the mismatched PSF. Note the overlap of the standard deviations. For the PDF deconvolution, which is independent of any separate PSF information, the results are $<$ 4% smaller than the true intensity ratio. However, for FITSTARS, the results are $\sim$ 100 and 4 times larger. For these data, FITSTARS had problems. For the SR=30% data, IDAC, StarFinder and AWMLE underestimate the intensity ratio $\sim$ 20% for the matched PSF, and IDAC and AWMLE overestimate by ~ 40%, whereas StarFinder yields the correct value, for the mismatched PSFs. PDF deconvolution underestimates the truth by ~ 4% in both cases. It is interesting to note that FITSTARS yields significantly improved results for these data overestimating by ~ 1.5. Finally we note that the dispersions increase in the majority of mismatched PSF cases.

Table A.3: Means and standard deviations of brightness ratios obtained with the discussed methods.

		FITSTAR	AWMLE	IDAC	StarFinder	PDF deconvolution
30% SR, matched PSF	....	54.0 ± 10.0	20.6 ± 4.7	21.9 ± 6.5	17.1 ± 0.9	24.4 ± 1.8
50% SR, matched PSF	....	2737.0 ± 4275.0	26.1 ± 1.0	26.4 ± 1.0	26.2 ± 0.5	24.1 ± 2.1
30% SR, mismatched PSF	....	38.9 ± 34.5	36.1 ± 22.6	34.6 ± 29.6	25.0 ± 8.6	24.4 ± 1.8
50% SR, mismatched PSF	....	96.0 ± 124.2	28.6 ± 2.5	27.1 ± 1.3	30.5 ± 2.1	24.1 ± 2.1

So, why do these algorithms differ so much in their results? Non-linear deconvolution algorithms, such as AWMLE and IDAC, have a tendency to overestimate intensity ratios when the intensity ratio is large to begin with. This has been shown by Christou et al. (2004). This is essentially due to the reduced SNR of the fainter sources. Also some of the algorithms compute the astrometry of the target jointly with the photometry and when the astrometric errors are large, so are the photometric errors. For example, StarFinder converged for only four of the eight realizations for the SR=30% mismatched PSF case and FITSTARS converged for six of the eight realizations for the SR=30%, matched PSF case. PDF deconvolution is not influenced by astrometric errors as the true binary component locations are used. In addition, not all algorithms are well matched to these data. For example FITSTARS assumes that the reconstructed PSF is symmetric after a certain radius which could well affect the results here, because of the binary star separation relative to the size of the extended PSF. The different algorithms also use the data in different ways. FITSTARS, StarFinder and AWMLE used a single SAA image obtained from the original 104 data frames while IDAC used ten 103 SAA images to take advantage of the MFBD approach and PDF deconvolution used the 104 frames for the statistical distributions. We have not yet investigated the repeatability of these algorithms by breaking the data into smaller subsets to investigate how that affects the mean results.

Another difference to note is that the binary parameters themselves are estimated differently for each of the PSF calibration algorithms. Aperture photometry, centered on the component locations, is used for the AWMLE result. For the SR=30% case there is greater speckle contamination of the companion thus affecting the results. The presence of a deterministic mask using the component locations should improve the results by rejecting the speckles. StarFinder jointly estimates the relative astrometry and photometry parametrically and it was found that if the initial estimate of the companion's location was more than 0.5 pixels away from its true location, then the algorithm would not converge. FITSTARS also jointly estimates the binary parameters directly and is sensitive to the astrometric positions. IDAC estimates an object intensity distribution and the use of an initial Gaussian model centered on the component locations ensured that this was limited to the correct locations. The binary parameters were obtained by fitting the final Gaussian result where the formal error of the least squares fit for the intensity ratio was ~ 2%, substantially smaller than the results for the eight different realizations. For the SR=30% mismatched PSF case, the asymmetric nature of the PSF led to increased speckle contamination in the results for a couple of the realizations so that the mean was skewed and the standard deviation was increased.

By comparison to the PSF calibration algorithms, PDF deconvolution is self-calibrating and relies on how well the speckle statistics are determined in order to estimate the relative intensities. The ability to determine the statistics of the intensity is directly related to the number of samples, i.e. the number of frames. How well such an algorithm will work with a small number of frames is yet to be determined as is the maximum exposure time per frame before the central-limit theorem dominates producing indistinguishable Gaussian statistics.

We have presented preliminary results of the algorithms' application to these challenging data. Future studies will investigate the sensitivity of the PSF calibration methods to different initial PSF estimates, the repeatability of all techniques and the sensitivity of the multi-frame algorithms, IDAC and PDF deconvolution, to the number of frames.

Appendix B

Comparison of image restoration algorithms in the context of horizontal-path imaging

Published in Proceedings of the International Society for Optics and Photonics SPIE. Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XXIII. 2012, ed: Gerald C. Holst, Keith A. Krapels, Proceedings Vol. 8355, 83550X.

Abstract

We have looked at applying various image restoration techniques used in astronomy to the problem of imaging through horizontal-path turbulence. The input data comes from an imaging test over a 2.5*km* path. The point-spread function (PSF) is estimated directly from the data and supplied to the deconvolution algorithms. We show the usefulness of using this approach, together with the analytical form of the turbulent PSF due to D. Fried, for reference-less imaging scenarios.

B.1 Introduction

The performance of optical systems is degraded by atmospheric turbulence when observing vertically (e.g. astronomy) or horizontally (e.g. surveillance, military reconnaissance). This degradation can be alleviated in software (Huebner 2011; Aubailly et al. 2009) or hardware (Marchi & Scheifling 2011; Vdovin et al. 2012). Until now, in horizontal-path imaging the software solution has been more successful because adaptive optics (AO) systems are generally only capable of correcting very small fields of view. In single-conjugated AO systems, i.e., ones that measure and correct atmospheric distortions in one direction, the usable field-of-view is determined by the so-called isoplanatic angle, θ_0. It describes the angle out to which optical path variations deviate by less than one radian rms phase aberration from each other. Given a certain correction direction, θ_0 gives the maximum angular separation from this direction at which reasonably good correction can be expected. In conventional AO systems, correctable field-of-view is of size $2\theta_0$ and in this case the most-often employed solution to spatially-varying blur is simply to have a detector subtend a small angle. In astronomical imaging, θ_0 is typically of the order of a few tens rad at visible wavelengths and strongly depends on the height distribution of the turbulent layers. In horizontal-path imaging, θ_0 is typically of the order of $100\mu rad$ and this implies that, with a conventional AO system looking through $1km$ of turbulence, one can obtain sharp images of a $10cm$ object but not of a person.

In the post-processing solution this problem is dealt with in the following manner: firstly, the geometric distortions corresponding to isoplanatic patches are estimated and the patches are re-assembled into their original positions (“de-warping”), subsequently deconvolution is performed (“de-blurring”) (Huebner 2011). The software approach allows for sharpening arbitrarily large fields of view and therefore it has an advantage over AO for imaging scenarios (in other applications, e.g. directed energy propagation, field of view does not play a significant role). Ideally, one would perform local motion compensation in patches of size equal to the “tilt isoplanatic angle” (Louthain & Schmidt 2008) which is larger than the (phase) isoplanatic angle. In any case, in this paper we concentrate on the deconvolution operation (image restoration, “de-blurring”). As such, this work follows some ideas already laid out before (van Iersel & van Eijk 2010), namely a proposition to extract the point-spread function (PSF) directly from the data, albeit here we actually follow through with this idea in a rigorous way. Subsequently, we analyze strengths and weaknesses of various deconvolution methods and their sensitivity to mismatch of the PSF (wrong kernel).

B.2 Methods

The imaging operation corresponds to (discretized) convolution of the focal-plane representation of the true object $\mathbf{o}(\mathbf{x})$ with the PSF $\mathbf{h}(\mathbf{x})$, giving the (discretely sampled) recorded image $\mathbf{i}(\mathbf{x})$ with the unavoidable addition of noise $\mathbf{n}(\mathbf{x})$:

$$i(x, y) = [o(x) \otimes h(x)] \diamond n(x), \tag{B.1}$$

where the symbol $\diamond$ is a pixel-by-pixel operation which reduces to the simple addition in the case when noise is additive and independent of $[\mathbf{o}(\mathbf{x}) \otimes \mathbf{h}(\mathbf{x})]$, while for Poisson noise it is an operation which returns a random deviate drawn from a Poisson distribution with mean equal to $[\mathbf{o}(\mathbf{x}) \otimes \mathbf{h}(\mathbf{x})]$. The symbol $\mathbf{x}$ corresponds to two-dimensional focal-plane coordinate.

In the context of astronomical image processing two approaches to estimating the true object $\mathbf{o}(\mathbf{x})$ have been widely accepted, namely Richardson-Lucy deconvolution (Richardson 1972; Lucy 1974), which leads to the maximum-likelihood solution in the presence of Poisson noise, and maximum-entropy (Frieden 1972). Both approaches make use of non-linear techniques to seek an object estimate iteratively. The Richardson-Lucy scheme is expressed as follows:

$$\hat{o}_{k+1}(x) = \hat{o}_k(x) \circ \left\{ h(-x) \otimes \left[\frac{i(x)}{h(x) \otimes \hat{o}_k(x)} \right] \right\}, \tag{B.2}$$

where k denotes the current iteration number, $\mathbf{a} \otimes \mathbf{b}$ denotes the pixel by pixel product of two equally-sized arrays $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{a}/\mathbf{b}$ denotes their pixel by pixel quotient.

The R-L algorithm has the interesting property of non-negativity: if the first estimate is non-negative, the further estimates will also be non-negative. The problem with the R-L method is “knowing when to stop”, i.e., how to obtain the best trade-off between the desired spatial resolution enhancement against the unavoidable noise amplification. The users usually rely on their visual judgment and stop the algorithm manually. When the algorithm is allowed to iterate eventually noise present in the data will get amplified. We illustrate this effect in Figure B.1 using the data from the experiment (Section B.3) and the “un-supervised” R-L scheme.

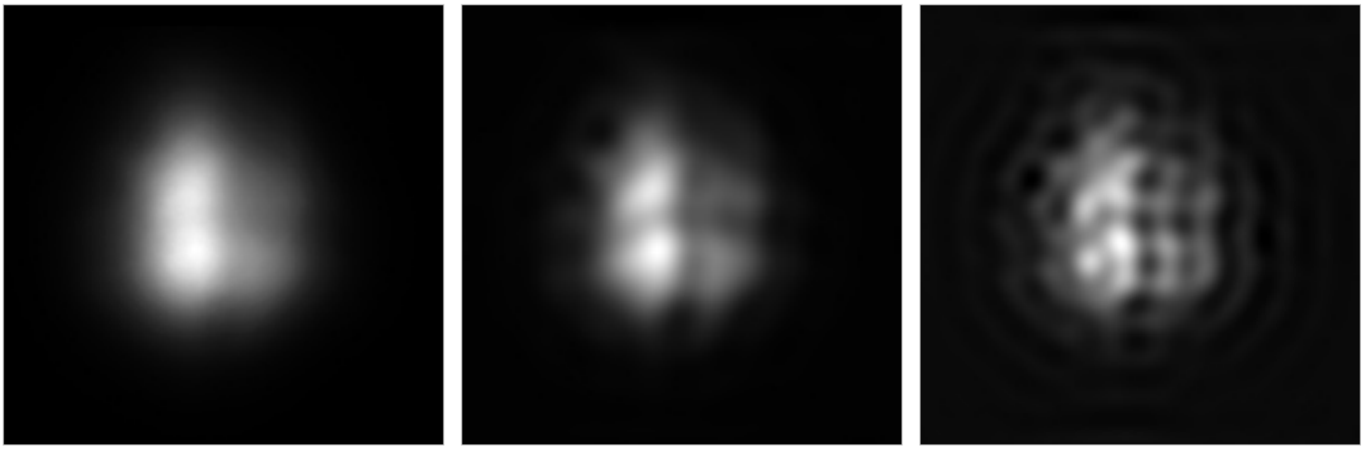

Figure B.1: Illustration of the noise amplification effect in Richardson-Lucy deconvolution. Pictures correspond to data collected in the experiment described in Section B.3. Left: summed image before deconvolution. Centre: result of RL deconvolution after 100 iterations. Right: result after 1000 iterations.

Solutions to this problem fall into the regularization category (Tikhonov 1963; Mugnier et al. 2004), or some other form of noise control (Baena Gallé et al. 2013; Starck & Murtagh 1994). Regularization involves introducing additional information about the object in order to solve an ill-posed problem. This information is usually presented in the form of a penalty for a class of less-likely objects. Restrictions on object gradients or entropy are common. Alternatively, one example of noise control is through the use of probabilistic masks: local windows which determine whether a significant structure is present or not at a given location in the data, at iteration k (Starck & Murtagh 1994). The mathematical expressions for these masks aim at representing the probability of finding a source, or part of an object, within the window. For example, local standard deviation relative to the global standard deviation, or local correlation between two images that are deconvolved simultaneously are two candidates for a probabilistic mask (Baena Gallé et al. 2013). If the presence of a signal within the mask is considered to be insignificant, the iterative process should stop. This is the premise of the AWMLE algorithm described in Section B.2.1.

Another problem with the R-L method is that its convergence is very slow compared to e.g. conjugate-gradient optimization (Thiébaut 2005). To alleviate this problem several acceleration schemes have been proposed, of which we will mention two. One is "multiplicative relaxation" (Llacer & Núñez 1990), which boils down to replacing the iteration of Equation B.1 by

$$\hat{o}_{k+1}(x) = \hat{o}_k(x) \circ \left\{ h(-x) \otimes \left[\frac{i(x)}{h(x) \otimes \hat{o}_k(x)} \right] \right\}^{\alpha}, \tag{B.3}$$

where α is an acceleration parameter (usually larger than 1). Another approach is "linear relaxation"and it can be written in the form (Prato et al. 2012)

$$\hat{o}_{k+1}(x) = \hat{o}_k(x) - \lambda_k \hat{o}_k(x) \circ \left\{ \mathbf{1} - h(-x) \otimes \left[\frac{i(x)}{h(x) \otimes \hat{o}_k(x)} \right] \right\}, \tag{B.4}$$

where $\lambda_k > 1$ (for $\lambda_k = 1$ the R-L algorithm is re-obtained) and $\mathbf{1}$ is an array with all entries equal to 1. This modification of Equation B.1 produces a scaled gradient method with a scaling given by object estimate at iteration k and a gain factor. We have therefore obtained a form suitable for optimization by minimizing an objective function. The algorithm SGP-RL (scaled gradient projection Richardson-Lucy) which is based on Equation B.1 is described in Section 2.2.

The aforementioned methods work with static PSFs, i.e. they do not update the PSF of an optical system which is supplied by the operator. They have been very successful in reconstructing pristine images from the Hubble Space Telescope before the first servicing mission which fixed the aberrated PSF. Nevertheless, in ground-based imaging, whether looking up (e.g. astronomy) or sideways (e.g. surveillance or military reconnaissance), the assumption of perfectly-known PSF is rarely justified. In astronomy, a calibration PSF is often obtained by observing a single star after the target or by extracting an isolated star as PSF if the field of view is large. In surveillance, obtaining a PSF from a beacon is a very difficult task, as we will show in Section 4. This has led to development of the "blind" deconvolution paradigm. A blind method works without any information about the PSF. When some information is available, for example in the form of a statistical prior on the PSF, then a "myopic" method can be used. One such algorithm, called MISTRAL, is described in Section B.2.2.

B.2.1 AW(C)MLE

The Adaptive Wavelet/Curvelet Maximum Likelihood Estimator (AW(C)MLE) (Baena Gallé et al. 2013) is a Richardson-Lucy-type algorithm which maximizes the likelihood between the dataset and the projection of a possible solution onto the data domain, considering a combination of the Poissonian shot noise, intrinsic to the signal, and the Gaussian readout noise of the detector. This maximization is performed either in the wavelet domain -AWMLE- (Shensa 1992) or in the curvelet domain -ACMLE- (Candès et al. 2006). The decomposition of the signal into several channels allows for various strategies to be used depending on a particular channels scale. This is a direct consequence of the fact that in the wavelet or curvelet decomposition the noise, together with the finest structures of the signal will be transferred into the high-frequency channels while coarse structures will be transferred into the low-frequency channels. An illustration of wavelet decomposition of the summed image collected as part of the experiment is shown in Figure B.2.

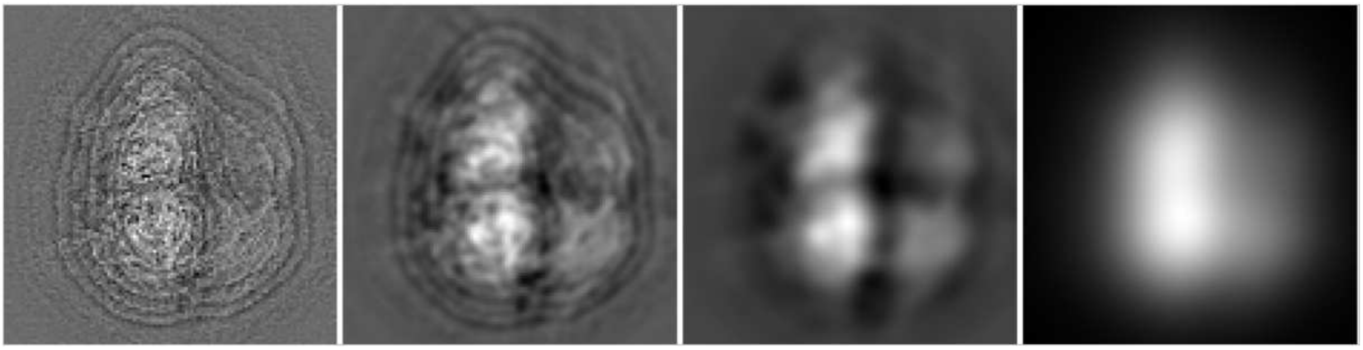

Figure B.2: Wavelet decomposition of the experimental data described in Section B.3. Wavelet planes, from the finest to the coarsest, are shown from left to right.

Additionally AW(C)MLE is equipped with two probabilistic masks which can be employed separately or together. For this project we have used the local-correlation mask applied to two images that are deconvolved simultaneously (but independently). The two images are sums of the first, and the second half of the target observations. The correlation mask measures, at each iteration, the similarity between the same region in the two images and, if noise amplification is detected (in the sense that the local correlation is reduced), the algorithm is stopped for that region. The expression for the mask is based on normalized covariance between the two samples.

B.2.2 SGP-RL

Scaled gradient projection (SGP) is another method based on the Richardson-Lucy iteration (Prato et al. 2012). It is based on Equation B.4. The algorithm contains several stopping rules from which the user can choose. We have used the following rule:

$$\frac{||\hat{o}_{k+1}(x) - \hat{o}_k(x)||}{||\hat{o}_{k+1}(x)||} \leq \delta, \tag{B.5}$$

where δ is user-specified tolerance (we used $\delta = 10^{-6}$). Basically, this rule stops the algorithm when change in the solution becomes very small. For our data, the code stopped automatically after 20-30 iterations.

B.2.3 MISTRAL

The Myopic Iterative STep-preserving Restoration ALgorithm (MISTRAL) is a deconvolution method within the Bayesian framework that jointly estimates the PSF h and the object o using some prior information about both these unknowns (Mugnier et al. 2004). This joint Maximum A Posteriori (MAP) estimator is based on the following expression:

$$\begin{aligned}[\hat{\mathbf{o}}, \hat{\mathbf{h}}] &= \arg\max[p(\mathbf{o}, \mathbf{h}|\mathbf{i})] = \\ &= \arg\max[p(\mathbf{i}|\mathbf{o}, \mathbf{h}) \times p(\mathbf{o}) \times p(\mathbf{h})] = \\ &= \arg\min[J_i(\mathbf{o}, \mathbf{h}) + J_o(\mathbf{o}) + J_h(\mathbf{h})] \,,\end{aligned} \tag{B.6}$$

where $J_i(\mathbf{o}, \mathbf{h}) = \ \ln p(i|o, h)$ is the joint negative log-likelihood that expresses fidelity of the model to the data, $J_o(\mathbf{o}) = \ \ln p(o)$ is the regularization term, which introduces some prior knowledge about the object, like positivity, and $J_h(\mathbf{h}) = \ \ln p(h)$ accounts for some partial knowledge about the PSF. The symbol p in the above expressions corresponds to the probability density function of a particular variable.

MISTRAL uses a nonstationary Gaussian model for the noise, which is a common approximation for high-level light conditions, as the ones we are dealing with here. What this means is that a least-squares optimization with locally-varying noise variance is employed:

$$J_i(\mathbf{o}, \mathbf{h}) = \sum_r \frac{1}{2\sigma^2(r)} [i(r) - (o \otimes h)(r)]^2 \,, \tag{B.7}$$

This prior makes it easier to compute the solution with gradient-based techniques as compared to the Poissonian likelihood which contains a logarithm. The prior probability, $J_o(\mathbf{o})$, is modelled to account for objects which are a mix of sharp edges and smooth areas. The adopted expression contains an edge-preserving prior that is quadratic for small gradients and linear for large ones. The quadratic part ensures good smoothing of the small gradients (i.e., of noise), and the linear behaviour cancels the penalization of large gradients (i.e., of edges). Such combined priors are commonly called $L_2 - L_1$ (Green 1990; Bouman & Sauer 1993). The PSF prior $J_h(\mathbf{h})$ assumes that the PSF is a multidimensional Gaussian random variable (which is a good assumption for long exposures or sums of short exposures). The criterion draws the solution, in the least-squares fashion, towards the user-supplied first-guess PSF while obeying the error bars given by the squared optical transfer function (Fourier transform of the PSF).

B.2.4 Summary of the employed methods

The properties of the codes are summarized in Table B.1. We note here that in our tests we used AWMLE with no acceleration ($\alpha = 1$ in Equation B.3) hence relatively long execution times of this algorithm, but acceleration is of course possible when one changes the value of α.

Table B.1: Information on methods used to process experimental data.

		static PSF / Blind	stopping rule	number of iterations
RL	....	static-PSF	no	100 (fixed)
SGP-RL	....	static-PSF	yes	20-30
AWMLE	....	static-PSF	no	100 (fixed)
MISTRAL	....	blind/myopic	yes	$\sim$ 750 (matched PSF) $\sim$ 1600 (mismatched PSF)

B.3 Experiment

The experiment was conducted in Ettlingen, Germany, on the 27th of March 2012 between 4pm and 6pm. The object, a cross made of blackboard on a $20cm$ lamp (Figure B.3, left panel), was viewed from a distance of $2.5km$ using an $18cm$ telescope. Both, the emitter and the receiver, were positioned at a height of approximately $20m$. The telescope was mounted on an optical table inside the Adaptive Optics Laboratory and it was fully protected from the weather outside. Therefore, in contrast to other experiments (van Iersel & van Eijk 2010), we believe we can obtain turbulence measurements from our images and these measurements will not be affected by motion of the setup (Section B.3.2). The lamp was emitting white light but a filter before the sensor allowed only 600 to $700nm$ wavelengths. The pixel scale of the detector was chosen by adjusting the focal length of the last lens to conform to the Nyquist sampling requirement ($\lambda/2D = 1.8rad$, pixel scale = $1.3rad$). We collected 1000 images of the $20cm$ lamp at $300fps$, and also 100 images of another lamp a $1cm$ “point source” at $30fps$ (Figure B.3, right panel). The “point source” images were recorded 25 minutes after the target observations. They were collected as a PSF reference. All images were saved in the uncompressed FITS format.

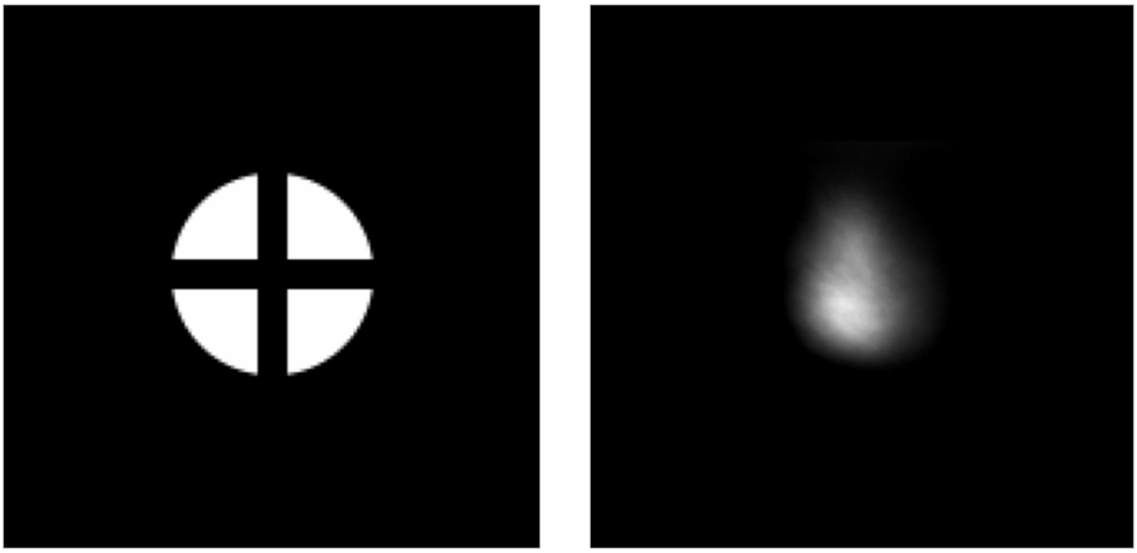

Figure B.3: Left: schematic drawing of the object used in the experiment. Right: sum of 100 images of the small lamp, after background removal this image was used as PSF in some of the reconstruction trials described in Section B.4.

The unique aspect of this experiment was that a BLS900 boundary layer scintillometer from Scintec was mounted right above the telescope. The scintillometer was collecting data continuously and recording C_n^2, r_0 and transverse wind speed estimates every minute. From these measurements we know that turbulence was rather modest and stable during the experiment (Figure B.4). The values of C_n^2 rarely exceeded 10^{-14} $m^{-2/3}$. Serendipitously, the values of C_n^2 recorded during the target and PSF observations were very similar: 5.37×10^{-15} $m^{-2/3}$ and 5.47×10^{-15} $m^{-2/3}$, respectively.

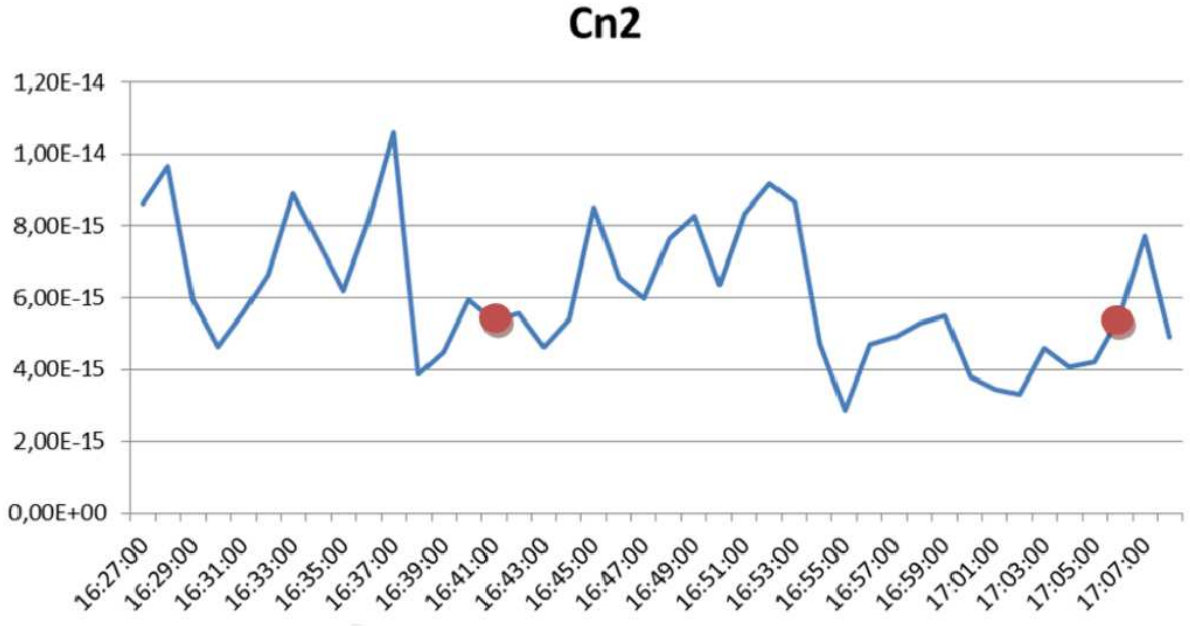

Figure B.4: Turbulence strength, quantified by C_n^2, measured by the scintillometer around the time of the experiment. The first red dot denotes the time when the target images were recorded, and the second red dot on the right corresponds to the PSF measurements.

B.3.1 Pre-processing

Before any attempts at image reconstruction the data has to be pre-processed. In our case the following procedures have been carried out on the raw target frames:

- cropping of the images,
- background removal,
- registration (via up-sampled cross-correlation with the best frame). This operation removed global image motion.

The frames were summed and these summed images were then supplied to the deconvolution algorithms. In the case of AWMLE two images were used (corresponding to the first and the second half of the target observations).

B.3.2 Turbulence estimates from the images

There are several ways to estimate strength of the turbulence directly from the images. In ground-to-ground scenarios, methods based on image motion are most commonly used (Sasiela & Shelton 1993; Belen'kii et al. 2001; Zamek & Yitzhaky 2006). In (solar) astronomy, the "spectral-ratio" method is preferred (von der Lühe 1984). Because in our case the sensor was shielded from wind-induced vibrations and temporal power spectra of image motion did not exhibit any artificial peaks (Figure B.5), we decided to employ the motion-based approach (alternatively called the angle-of-arrival AOA approach).

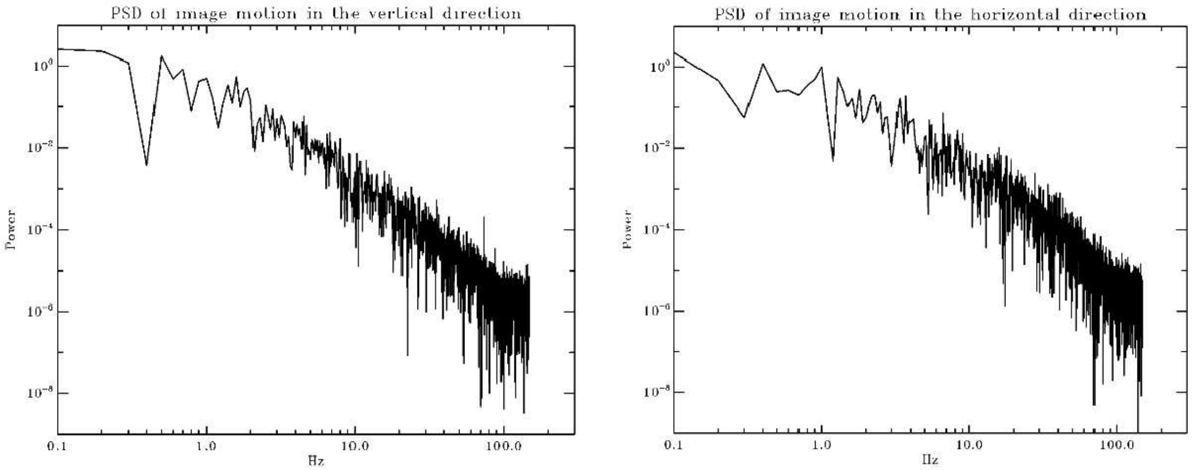

Figure B.5: Temporal power spectral densities of AOA fluctuations measured in the central part of the target images. Left: vertical direction. Right: horizontal direction.

The relationship between the single-axis AOA variance, within some isokinetic patch (Zamek & Yitzhaky 2006), and the coherence length of the turbulence the so-called Frieds parameter r_0 is:

$$r_0 = \frac{3.15}{(D^{1/3}\sigma_\alpha^2 k^2)^{3/5}} , \tag{B.8}$$

where D corresponds to the telescope diameter, k is the wavenumber and spherical-wave approximation was used. The effect of finite outer scale was neglected. The significance of the Frieds parameter is that it completely specifies the long-, and short-exposure PSF in the widely adopted theory due to Fried (1966). In this theory, the optical transfer function OTF, i.e., the Fourier transform of the PSF, for the two cases of long exposures and registered (tip-tilt corrected) short exposures is given by:

$$OTF_L(\vec{v}) = OTF_0(\vec{v}) \times OTF_{LE}(\vec{v}) , \tag{B.9}$$

$$OTF_S(\vec{v}) = OTF_0(\vec{v}) \times OTF_{SE}(\vec{v}) , \tag{B.10}$$

where $OTF_{LE}(\vec{v})$ represents the average long-exposure OTF of the atmosphere, $OTF_{sE}(\vec{v})$ is the average short-exposure OTF of the atmosphere, $OTF_L(\vec{v})$ and $OTF_S(\vec{v})$ are the overall long- and short-exposure average OTFs (including the effect of the telescope diffraction). Besides, for a diffraction-limited circular aperture of diameter D we have:

$$OTF_0(\vec{\nu}) = \frac{2}{\pi}\left[\arccos\left(\frac{\bar{\lambda}z\nu}{D}\right) - \frac{\bar{\lambda}z\nu}{D}\sqrt{1-\left(\frac{\bar{\lambda}z\nu}{D}\right)^2}\right] , \tag{B.11}$$

where $\nu = |\vec{\nu}|$, $\bar{\lambda}$ is the average wavelength, and z is the distance from the exit pupil to the image plane. Fried developed expressions for $OTF_{LE}(\vec{\nu})$ and $OTF_{SE}(\vec{\nu})$:

$$OTF_{LE}(\vec{\nu}) = \exp\left[-3.44\left(\frac{\bar{\lambda}z\nu}{r_0}\right)^{5/3}\right] , \tag{B.12}$$

$$OTF_{SE}(\vec{\nu}) = \exp\left\{-3.44\left(\frac{\bar{\lambda}z\nu}{r_0}\right)^{5/3}\left[1-\alpha\left(\frac{\bar{\lambda}z\nu}{r_0}\right)^{1/3}\right]\right\} , \tag{B.13}$$

where α is a parameter that varies between $1/2$ when there are both intensity and phase variations across the collecting aperture and 1 when only phase distortions are present. For our purposes we used Equation B.12 with Equations B.9 and B.11 for deconvolution. This is because the object being imaged covered several tilt-isoplanatic patches (or isokinetic patches) as can be appreciated from Figures B.6-B.9. Global motion compensation did not remove warping from the images. Looking at Equations B.9-B.13 one can notice that having an estimate of r_0 translates to an estimate of the PSF at the time of the observations, as posited above. As a side remark we will mention that there is a very simple relationship between the Frieds parameter and the refractive index structure constant C_n^2, a parameter which is more often used to describe the turbulence strength in ground-to-ground imaging:

$$r_0 = 3.0(C_n^2 L k^2)^{-3/5} , \tag{B.14}$$

where L is the distance between the object and the sensor and spherical-wave approximation was used, again (Tyson 1998).

We measured AOA variance in the central part of our images and translated it to the value of r0 with the help of Equation B.8. Subsequently, using Equations B.9, B.11, B.12, and the fast Fourier transform, we generated a synthetic PSF corresponding to the observations. In the experiment with the point source ($1cm$ lamp), whereby the turbulence-induced object motion is most definitely isoplanatic, we obtained very good agreement with the scintillometer measurement: $2.79cm$ vs. $2.91cm$ (both values normalized to $500nm$ wavelength). For the target observations, which according to the scintillometer measurements experienced similar strength of turbulence as the point-source observations (Figure B.4), the values were in disagreement: $2.26cm$ vs. $2.94cm$. Nevertheless, we will show that turbulence estimates obtained directly from the data proved to be closer to the truth than estimates from the scintillometer.

As a final remark we want to mention that the isoplanatic angle (not the tilt isoplanatic angle) is related to r_0 (Tyson 1998):

$$\theta_0 \approx 0.6 r_0 / L \,, \tag{B.15}$$

where constant C_n^2 along the propagation distance L was assumed, as is usually done. We have calculated that θ_0 was around eight times smaller than the target (the 20*cm* lamp). Therefore, we were working in anisoplanatic conditions.

B.4 Results

In this section we will present results of image reconstruction based on different algorithms and different PSFs. In Figures B.6-B.9 the caption “fast RL” corresponds to the SGP-RL algorithm, and the caption “MAP” corresponds to the MISTRAL algorithm. The caption “data” corresponds to the summed target image which was supplied to the deconvolution codes.

1. Case one, using the synthetic PSF estimated from AOA fluctuations: this is the nominal case. Here, PSF was computed as described in Section B.3.2. The results were better at least visually compared to the trials where other PSFs were used. The results are presented in Figure B.6.

2. Case two, using a smaller synthetic PSF (from AOA fluctuations artificially reduced by a factor of 0.8): here we wanted to check how good our PSF estimate from image motion actually is. The results are presented in Figure B.7. The figure shows generally worse image quality as compared to Figure B.6 indicating that our initial PSF estimate was close to the truth. Similar results were obtained when AOA fluctuations were scaled by a factor of 1.2.

3. Case three, using synthetic PSF from motion of the “point source”: as the scintillometer measurements indicated that the target and the PSF observations should have similar level of aberrations, and because estimation of AOA fluctuations from point-source images is trivial compared to estimation performed on the target data, we also used a synthetic PSF based on motion of the point source. The results, presented in Figure B.8, were again worse than in the nominal case.

4. Case four, using “point source” image as the PSF: here we wanted to test the suitability of beacon PSF (Figure B.3, right panel) for the task of image reconstruction of the target data. This is the “mismatched”-PSF case in Table B.1. The results were significantly worse compared to the nominal case. Only MISTRAL was able to perform satisfactory image reconstruction, thus showing the power of the blind/myopic approach. The results are shown in Figure B.9.

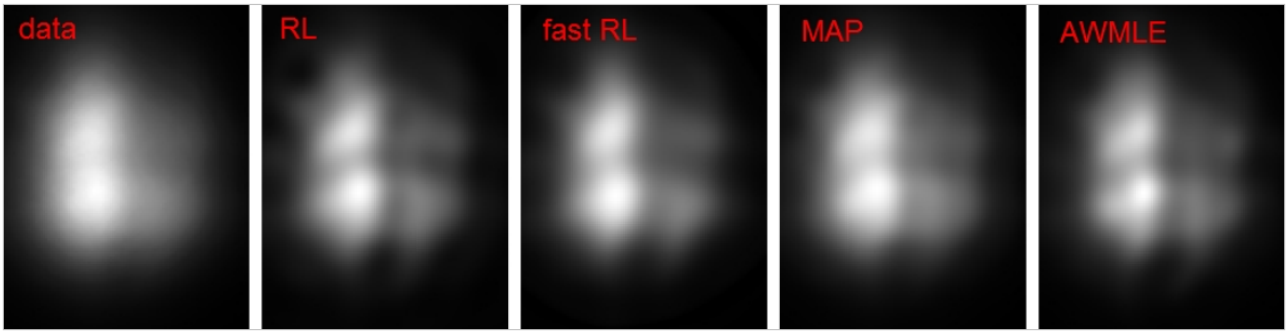

Figure B.6: Results of image reconstruction based on PSF computed from the AOA variance.

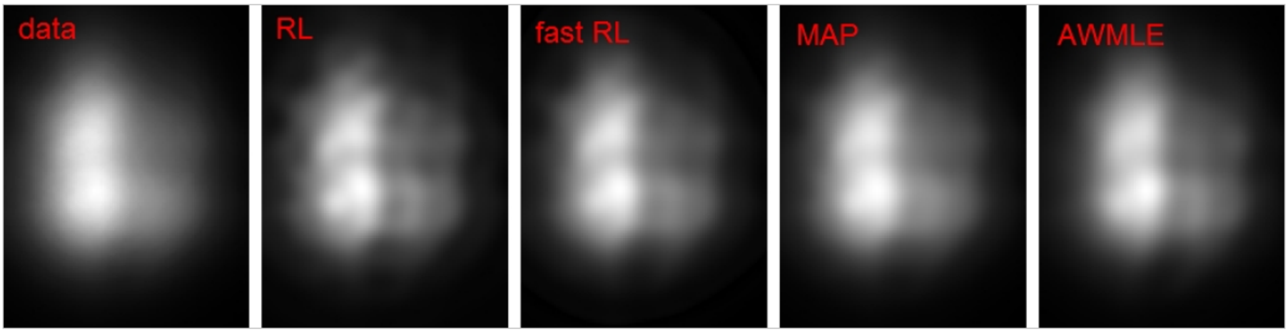

Figure B.7: Results of image reconstruction based on PSF computed from the AOA variance but with AOA fluctuations artificially reduced by a factor of 0.8.

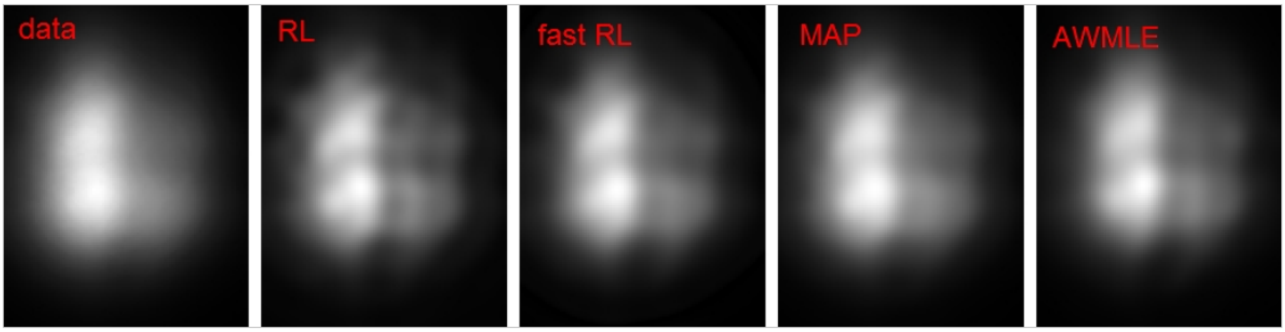

Figure B.8: Results of image reconstruction based on PSF computed from the AOA variance in the point-source data.

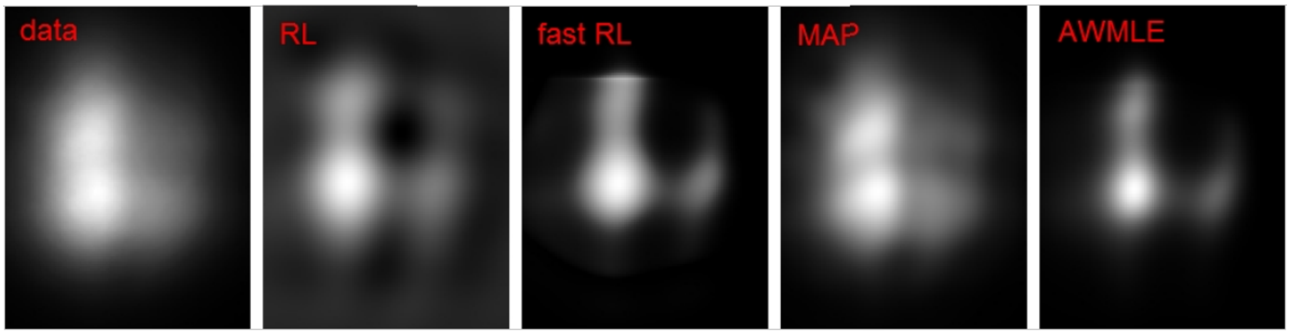

Figure B.9: Results of image reconstruction based on a “real” PSF (Figure B.3, right panel).

B.5 Summary

We have compared theoretical PSF estimates with in-situ measurements and with the estimates based on the scintillometer output. It is clear that target-data-based estimates are closer to the truth than estimates based on the auxiliary data. A number of codes, each representing different philosophy in attacking the problem, were employed for the task of image reconstruction. It was shown that when a carefully estimated PSF is available, the classic, non-blind Richardson-Lucy approach can perform very well. This is important as this approach can be summarized in approximately ten lines of code, and the execution of such code, when limited to only several iterations and preferably accelerated, should not last longer than $0.1s$ on a modern desktop PC.

Appendix C

High-contrast, adaptive-optics simulations for HARMONI

Poster presented in the 2nd International Conference on adaptive optics for extremely large telescopes, held in Victoria B.C. (Canada) on the 25-30th of April, 2011.

Press here to access the abstract on the official Conference website.

Abstract - HARMONI is a proposed visible and near-infrared integral field spectrograph for the European Extremely Large Telescope. We are exploring the potential of using HARMONI for high-contrast science, e.g., observations of exoplanets. Although HARMONI is not fed by extreme adaptive optics we show that substantial contrasts can be achieved by combining single-conjugate AO with coronagraphy and post-processing of the hyperspectral data cube using spectral deconvolution. HARMONI will be well suited for follow-up spectroscopy of planets detected by 8 m. class instruments, emphasizing their characterisation. We implement models of telescope aberrations: due to wind buffeting on M1, due to windshake on M2, due to rolled segment edges, as well as the ones resulting from M1 phasing and individual segment warping affected by thermal and gravity effects. Additionally, we investigate the impact of post-AO differential aberrations. We also look at possible improvements to spectral deconvolution which is our method of choice for data post-processing. Finally, we make predictions of achievable contrast which translates to the ability to characterise various types of exoplanets in detail.

High-contrast, adaptive-optics simulations for HARMONI

Szymon Gładysz[1], Niranjan A. Thatte[2], Graeme Salter[3], Roberto Baena Gallé[4], Fraser Clarke[2],
Matthias Tecza[2], Laurent Jolissaint[5]

[1]Technion - Israel Institute of Technology
[2]University of Oxford
[3]University of New South Wales
[4]University of Barcelona
[5]aquilAOptics

Abstract: HARMONI is a proposed visible and near-infrared integral field spectrograph for the European Extremely Large Telescope. We are exploring the potential of using HARMONI for high-contrast science, e.g. observations of exoplanets. Although HARMONI is not fed by extreme adaptive optics we show that substantial contrasts can be achieved by combining single-conjugate AO with coronagraphy and post-processing of the hyperspectral data cube using spectral deconvolution.

Context: HARMONI will most likely by the first-light instrument on the E-ELT. It is an integral field spectrograph with the following parameters:

- wavelength range: 0.47 - 2.45μm
- spectral resolving powers: 4000, 10000, 20000
- 32000 simultaneous spectra per FoV
- four spaxel scales of 40, 20, 10 and 4 milli-arcseconds/spaxel

Simulations: For high-contrast science HARMONI will be used in conjunction with a Lyot coronagraph. The SCAO system employs a Hartmann-Shack wavefront sensor with 85 × 85 subapertures. The other parameters of the simulations are:

- wavelength range: 1.65 – 2.45μm
- bright star: $m_v = 4.7$
- Strehl ratios: 65 - 85%
- seeing: 0.85arcsec @ 0.5μm
- zenith angle: 30°
- total differential refraction of 50mas
- coronagraph: Lyot setup, focal-plane mask size = 5λ/D @ 2.45μm
- 100 spectral channels
- post-AO aberrations: 30nm rms
- integration time: 1 minute
- Poisson, background and readout noise

Spectral deconvolution: Even with high-quality data post-processing is usually necessary to reach contrasts expected from exoplanetary systems. We implement a method called "spectral deconvolution" which relies only on optical phenomena and not on the physical characteristics of exoplanets. The method utilizes the fact that post-coronagraphic image residuals scale spatially with wavelength while a planet would stay in a fixed position.

Results: Fig. 1 shows the contrast curve scaled from 60 seconds to 8 hours.

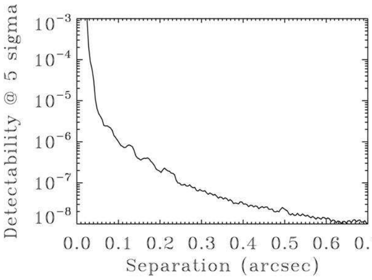

Fig. 1. Contrast after spectral deconvolution expected for an integration time of 8 hours.

Wavelets: "Traditional" spectral deconvolution has two problems:

- It cannot remove PSF features which do not scale spatially with wavelength. A coronagraph produces non-trivial modulation of the residuals close to the axis and in this region spectral deconvolution is not as effective as further out.
- Speckles close to the optical axis do not move significantly over the entire wavelength range. Conversely, in the numerically rescaled data cube close-in planets "dwell" on single spaxels and can be "killed" by the traditional approach.

For these two reasons we decided to combine spectral deconvolution with wavelets. We are helped by two facts:

- Shape of the planetary response in one spaxel of the re-scaled cube is deterministic (cross-section of the Airy pattern).
- The dependence of the width of this planetary response on separation from the star is also deterministic.

We designed a set of one-dimensional wavelet scales specifically for our problem.

Results: Figs. 2 and 3 illustrate the advantage of putting deterministic knowledge about the expected spatial scales into the method of spectral deconvolution.

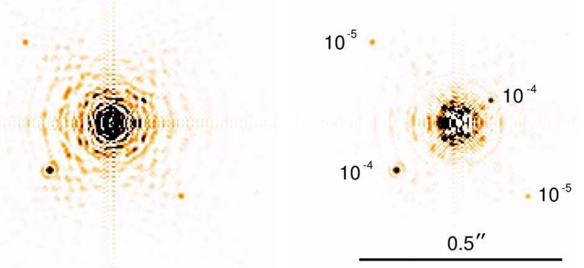

Fig. 2. Left: Data cube collapsed after traditional spectral deconvolution with polynomials. Right: wavelet-based approach. False color, inverted scale.

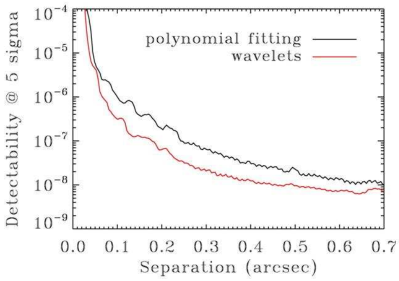

Fig. 3. Gain in contrast provided by wavelets.

Appendix D

The Wavelet Transform.

The wavelet transform (WT) is now a well-known tool with hundreds of practical applications in physics, mathematics and engineering. Its theoretical support was developed more than 20 years ago and a wide bibliography is available nowadays (Chui 1992; Daubechies 1992; Meyer 1993; Kaiser 1994; Vetterli & Kovacevic 1995). WT is related to the *sparsity* concept or sparse representation of a signal, which consists of transforming the data or image in such a way that the information is represented by only a few non-zero coefficients. In other words, a signal $f(x) = f[1], ..., f[N]$ is said to be sparse if most of its entries are equal to 0, i.e., $f(x)$ is a k-sparse signal of cardinality $k \ll N$ if it is a signal for which exactly k samples have a non-zero value. If a signal is not sparse, it can be *sparsified* in the appropriate transform domain, e.g., the wavelet domain.

D.1 A brief excursion into wavelets

WT was designed to overcome the limitations observed with the Fourier transform (FT), which is well suited to the study of stationary signals or non-time-dependent data. FT describes signals by means of basis functions that consist of a combination of sine and cosine functions; hence, their frequency content is time-independent or, in other words, the description provided by FT is complete in the frequency domain. However, many signals or datasets in real life are only non-zero for a short period of time, e.g.: a voice signal imparts information in both the time and frequency domains; or important features of an image, such as edges, are highly spatially localized. Such signals do not resemble any of the Fourier basis functions so FT cannot represent them appropriately. Therefore, WT aims to provide the advantage of combining the information in both the frequency domain and the time/spatial domain simultaneously.

In WT, as with FT, a signal is described by means of a finite basis function, named the *mother wavelet*, which is scaled and translated in order to create a set of basis functions. In one dimension, it can be expressed as follows:

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}}\psi\left(\frac{x-b}{a}\right), \tag{D.1}$$

where $a > 0$ and b are real numbers. The variable a reflects the scale while b specifies position along the x-axis. $\psi(x)$ is the mother wavelet and must satisfy the admissibility criterion (Chui 1992):

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\psi(s)|^2}{|s|} ds < \infty. \tag{D.2}$$

Due to the presence of s in the denominator, it is necessary that:

$$\psi(0) = 0 \Rightarrow \int_{-\infty}^{\infty} \psi(x) dx = 0. \tag{D.3}$$

That is, the mother wavelet has zero mean and, as a consequence, so do all its translations and scalings. Since $\psi(\infty) = 0$ too, the frequency response of the mother wavelet is similar to a band-pass filter; indeed, any band-pass filter impulse response with zero mean that decays to zero fast enough with increasing frequency can be used as the mother wavelet and, as a consequence, they are said to have *compact support*. Figure D.1 shows different examples of typical mother wavelet functions.

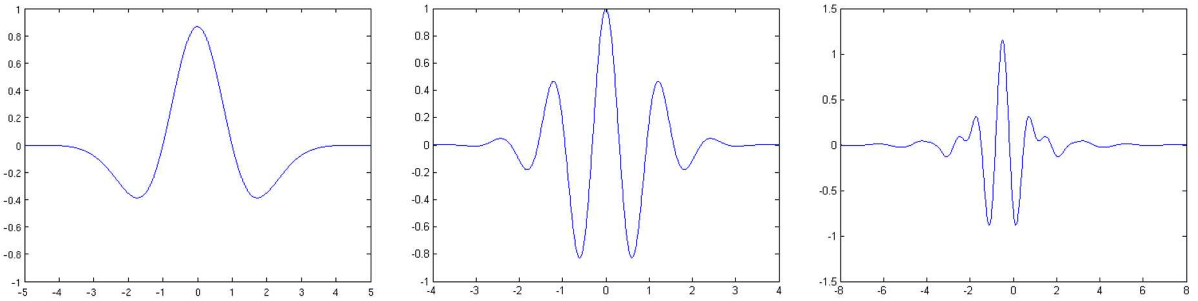

Figure D.1: Different examples of mother wavelet functions. Left: Mexican hat function. Middle: Morlet function. Right: Meyer function.

The continuous wavelet transform (CWT - first introduced by Grossman & Morlet (1984)) of a signal $f(x)$ with respect to a wavelet $\psi(x)$ is defined as an inner product:

$$W_f(a,b) =< f, \psi_{a,b} >= \int_{-\infty}^{\infty} f(x)\psi_{a,b}(x)dx, \tag{D.4}$$

The original signal $f(x)$ can be recovered by:

$$f(x) = \frac{1}{C_\psi} \int_0^{\infty} \int_{-\infty}^{\infty} W_f(a,b)\psi_{a,b}(x)db\frac{da}{a^2}. \tag{D.5}$$

CWT has the properties of:

1. a linear transformation, i.e., for any scalar α_1 and α_2, if $f(x) = \alpha_1 f_1(x)+\alpha_2 f_2(x)$ then $W_f(a,b) = \alpha_1 W_{f_1}(a,b) + \alpha_2 W_{f_2}(a,b)$,

2. covariance under translation, if $f_0(x) = f(x - x_0)$ then $W_{f_0}(a,b) = W_f(a, b - x_0)$,

3. covariance under dilation, if $f_s(x) = f(sx)$ then $W_{f_s}(a,b) = \frac{1}{\sqrt{s}} W_f(sa, sb)$.

In contrast, as with FT, it is possible to define a wavelet series expansion by specifying the scalings and translations by integer numbers instead of real ones. Furthermore, if the scaling is shrunk by a factor of 2 (*binary scaling*) and the translation is a shift of $k/2^j$ (*dyadic translation*), where k is translation and j determines the scaling, both of which are integer numbers, and if we restrict $f(x)$ as well as the mother wavelet to functions that are 0 outside the interval $[0,1]$, then the family of basis functions can be indexed by a single index $n = 2^j + k$ (for $j = 0, 1, ...$ and $k = 0, 1, ..., 2^j - 1$). It will be orthonormal if the set $\{\psi_n(x)\}$ forms an orthonormal basis of $L^2(R)$ (Chui 1992; Castleman 1996). Under such conditions, we have created a compact *dyadic wavelet* described by the following equations:

$$\psi_n(x) = 2^{j/2}\psi(2^j x - k), \tag{D.6}$$

$$c_n =< f, \psi_n >= 2^{j/2}\int_{-\infty}^{\infty} f(x)\psi(2^j x - k)dx, \tag{D.7}$$

$$f(x) = \sum_{n=0}^{\infty} c_n\psi_n(x), \tag{D.8}$$

where c_n are the wavelet coefficients. The redundancy of CWT is thus dramatically reduced; indeed, if only one of the $\psi_n(x)$ is similar to $f(x)$, then the series can be truncated to a few terms with no appreciable error.

Lastly, the discrete wavelet transform (DWT) is the most useful for image compression, analysis and processing. It can also be calculated as an inner product, however, obtaining the mother function is more complicated. In a nutshell, there are three basic techniques for calculating DWT (Castleman 1996): filter bank theory, multiresolution or time-scale analysis and subband coding.

- Filter bank theory consists of filtering the signal through a set of ideally non-overlapping band-pass filters in order to analyze its frequency content independently. Thus, the signal is convolved by a certain number of filter transfer functions. Since such filters are designed to be real and even, both in time and frequency, the reflection in the convolution integral has no effect and it can be seen as a simple inner product. In the real case, the band-pass filters are designed to have smoothed edges and it is not possible to avoid a certain amount of overlapping between them.

  Filter bank theory offers a convenient means of representing signals composed of oscillatory components; however, it is not well suited to image analysis since, in general, the localized components of interest are not oscillatory but only include one cycle or part of a cycle, like edges, lines or spots. The objects in an image are observed to occur at different size scales, e.g., a spot can be a single pixel or several of them, an edge can be a clear transition from black to white or it can occur gradually along a certain distance. The multiresolution approach to image analysis tries to exploit this idea.

- One of the most common ways to implement multiresolution theory is by means of pyramidal algorithms. This consists of successively reducing the size of the image by a factor of 2, by averaging 2x2 pixel blocks each time discarding every second row and column of pixels, and analyzing the resulting image by, for instance, a 3x3 edge detection operator. In this way, if we analyze a 1024x1024 image, we will be able to find small edges in the original image, larger ones in the following 512x512 image, even larger ones in the 256x256, etc.

Burt & Adelson (1983) introduced a pyramid code based on the Gaussian function. The image $f_0(i, j)$ is low-pass filtered with a Gaussian impulse response and the result is subtracted from the original image. The high-frequency detail in the image is retained in this difference image. Then, the low-pass filtered image can be subsampled without loss of detail. The process is illustrated in Figure D.2. The basic equations that describe the process are:

$$f_1(i, j) = [f_0 * g](2i, 2j) \qquad \textit{and} \qquad h_1(i, j) = f_0(i, j) - [f_0 * g](i, j). \tag{D.9}$$

Images are decoded in the reverse order. Each subsampled image $f_k(i, j)$, beginning with the last one, is scaled to its previous size $k - 1$ by inserting zeros as necessary and interpolated by convolution with $g(i, j)$. The result is then added to the previous image $f_{k-1}(i, j)$ and the process is repeated on the resulting image until the original image is recovered. This reconstructs the original image without error. The similarity between the multiresolution and filter bank theories is evident and they have been unified under WT.

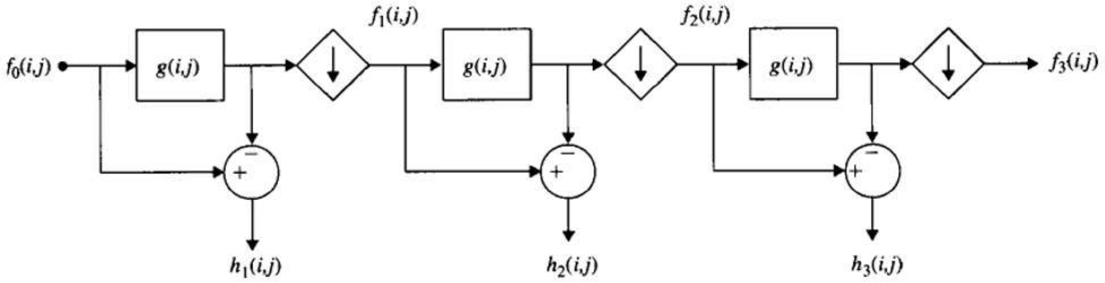

Figure D.2: The Laplacian pyramid coding scheme. —After Castleman (1996) —.

- Finally, subband coding seeks to decompose a signal or image into narrow-band components by band-pass filtering and to represent those components, without redundancy, in such a way as for it to be possible to reconstruct the original signal without error (Woods & O'Neill 1986). The signal $f(t)$ is uniformly sampled, with spacing equal to Δt, to form $f(i\Delta t)$. This sampled signal is then low-pass filtered to yield the low-resolution signal $g_0(i\Delta t)$, which retains the basic shape of $f(i\Delta t)$ but without the details. Therefore, $g_0(i\Delta t)$ can be sampled, with spacing as large as $2\Delta t$, without introducing aliasing. This process is called *subsampling* or *decimation*.

In contrast, the original $f(i\Delta t)$ signal can be filtered with a half-band band-pass filter in order to isolate the frequencies that were removed in the previous stage. A high-frequency sampled version $g_1(i\Delta t)$ is then created. Thus, $g_0(i\Delta t)$ and $g_1(i\Delta t)$, taken together, contain all the information present in $f(i\Delta t)$.

$$f(i\Delta t) = g_0(i\Delta t) + g_1(i\Delta t). \tag{D.10}$$

Two-channel subband coding only requires filtering $f(i\Delta t)$ through two filters $h_0(i\Delta t)$ and $h_1(i\Delta t)$ (Fig. D.3) yielding:

$$g_0(k\Delta t) = \sum_i f(i\Delta t) h_0((-i + 2k)\Delta t), \tag{D.11}$$

$$g_1(k\Delta t) = \sum_i f(i\Delta t) h_1((-i + 2k)\Delta t). \tag{D.12}$$

Figure D.3: 2-channel subband coding.

And the reconstruction formula:

$$f(i\Delta t) = 2\sum_k \left[g_0(k\Delta t) h_0((-i + 2k)\Delta t) + g_1(k\Delta t) h_1((-i + 2k)\Delta t)\right], \tag{D.13}$$

in the Fourier domain:

$$\begin{aligned} F(s) &= 2\left[\frac{1}{2}G_0(s)H_0(s) + \frac{1}{2}G_1(s)H_1(s)\right] = \\ &= 2\left[\frac{1}{2}F(s)H_0(s)H_0(s) + \frac{1}{2}F(s)H_1(s)H_1(s)\right] = \\ &= F(s)\left[H_0^2(s) + H_1^2(s)\right], \end{aligned} \tag{D.14}$$

and the two filters must satisfy the condition:

$$H_0^2(s) + H_1^2(s) = 1 \qquad for \qquad 0 \leq |s| \leq s_N = \frac{1}{2\Delta t}. \tag{D.15}$$

In order to obtain the rest of the scales, we can proceed in the same way with the output of the resulting low-pass filtered signals at each resolution level (Fig. D.4). Therefore, since $H_1^2(s) = 1 - H_0^2(s)$, a well-selected low-pass filter is all that is needed to design DWT. Such a filter must be symmetrical in order for equation D.15 to hold. The impulse response of such a filter is called the *scaling vector* $h_0(k)$, from which we can generate the related scaling function $\phi(t)$ and, from these, we can also generate $h_1(k) = (-1)^k h_0(-k+1)$, and the mother wavelet $\psi(t)$, as well as their respective scaled and translated orthonormal versions:

$$\phi(t) = \sum_k h_0(k)\phi(2t-k) \Longrightarrow \phi_{j,k} = 2^{1/2}\phi(2^j t - k) \qquad j = 0, 1, ... \quad k = 0, 1, ..., 2^j - 1 \quad \text{(D.16)}$$

$$\psi(t) = \sum_k h_1(k)\phi(2t-k) \Longrightarrow \psi_{j,k} = 2^{1/2}\psi(2^j t - k) \qquad j = 0, 1, ... \quad k = 0, 1, ..., 2^j - 1 \quad \text{(D.17)}$$

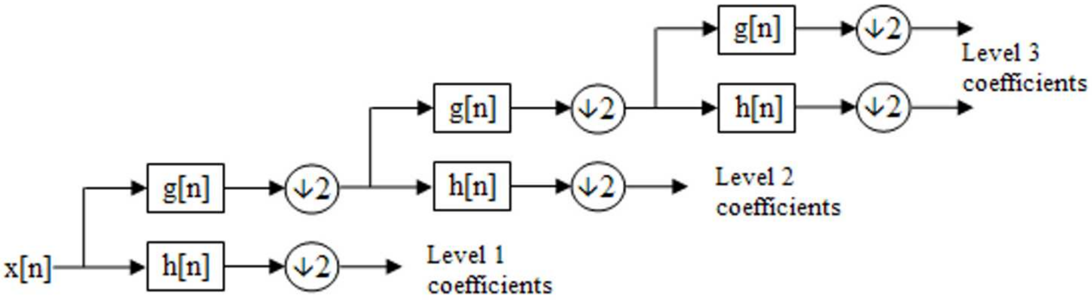

Figure D.4: N-channel subband coding.

The scaling function can be built as a weighted sum of half-scale copies of itself, using $h_0(k)$ as the weights. If the scaling vector has a finite number of non-zero entries, then $\phi(t)$ and $\psi(t)$ and the resulting wavelets will all have compact support (Daubechies 1988). Finally, given the set of orthonormal wavelets, DWT of the sampled function $f(i\Delta t)$ of $f(t)$ is:

$$c_{j,k} = \sum_i f(i\Delta t)\psi_{j,k}(i\Delta t) \qquad \textit{and} \qquad f(i\Delta t) = \sum_{j,k} c_{j,k}\psi_{j,k}(i\Delta t). \qquad \text{(D.18)}$$

A special case is the biorthogonal WT, which uses two different wavelet basis, $\psi(t)$ and $\tilde{\psi}(t)$, one for decomposition and the other for reconstruction. The two wavelets can be considered to be dual and they are biorthogonal if:

$$< \psi_{j,k}, \tilde{\psi}_{l,m} >= \delta_{j,l}\delta_{k,m}. \qquad \text{(D.19)}$$

Then we would have two different sets of coefficients for the decomposition:

$$c_{j,k} =< f(x), \tilde{\psi}_{j,k}(x) > \qquad \textit{and} \qquad d_{j,k} =< f(x), \psi_{j,k}(x) >, \qquad \text{(D.20)}$$

whereas for the reconstruction:

$$f(x) = \sum_{j,k} c_{j,k}\psi_{j,k}(x) = \sum_{j,k} d_{j,k}\tilde{\psi}_{j,k}(x). \tag{D.21}$$

Biorthogonal wavelets allow us, for instance, to build wavelet mother functions that only consist of odd or even functions (Castleman 1996).

WT now has a wide range of applications in different branches of applied science, including astronomy, seismology, biology and medical imaging. Furthermore, applications in image processing cover image compression, image fusion, deconvolution, etc.

D.2 The B3-spline à trous algorithm

While biorthogonal WTs are suitable for image compression, results are not so good for image analysis or restoration, such as denoising, deconvolution and detection, which are typical applications in astronomy. This is due to the loss of the translation-invariance property in DWT, which leads to the presence of artifacts when the wavelet coefficients are modified (Starck et al. 2010). In contrast, the loss of visual resolution at low frequency scales when the decimated WT is applied, is not very practical for visually inspection of any kind of image.

For these reasons, the astronomy community has preferred the use of *undecimated* WTs, where all the scales keep the original resolution. Of course, the amount of information grows with the number of scales, increasing the number of wavelet coefficients and, consequently, the computational costs. However, this drawback is assumed to be of less importance and easily overcome.

In order to do this, the filter bank construction is kept by creating fast and dyadic algorithms where the decimation step is eliminated. A very efficient way to implement this approach is the à trous algorithm (Holschneider et al. 1989; Shensa 1992). Hence, coefficients $c_{j+1}[k]$ and $\omega_{j+1}[k]$ can be defined as:

$$c_{j+1}[k] = (g^{(j)} \star c_j)[k] = \sum_l g[k]c_j[k+2^j l], \tag{D.22}$$

$$\omega_{j+1}[k] = (h^{(j)} \star c_j)[k] = \sum_l h[k]c_j[k+2^j l], \tag{D.23}$$

where $h^{(j)}$ and $g^{(j)}$ are, respectively, the high-pass and low-pass filters and $\star$ denotes for the discrete convolution operator. The filter $g^{(j)}$ is identified with even points, i.e.:

$$g^{(j)}[k] = \begin{cases} g[k] & if \quad k/2^j \quad is \quad an \quad integer \\ 0 & otherwise \end{cases}$$

Since there is no decimation there is no need to meet the de-aliasing condition either. This leads to a higher degree of freedom when designing synthesis prototype filter banks in the same way as, for instance, with non-orthogonal filter banks (Starck et al. 2010).

On the other hand, if filters are symmetric, i.e., they are equal to their time-reversed versions, and if they are also isotropic for two or more dimensions, and so too are the mother and scaling functions, then they will be well adapted for the analysis of most astronomical data, since such data are more or less isotropic in most cases (Starck & Murtagh 2006). Therefore, we will have built the so-called isotropic undecimated wavelet transform (IUWT), also called *starlet wavelet transform* (Starck et al. 2010) because of its close relationship to the astronomical context, or better-known as the B_3-spline à trous wavelet transform (Starck & Murtagh 1994), since the historical scaling function used for its implementation is a B_3-spline, which resembles a 2-D Gaussian function, thus fitting a stellar profile.

The most important properties of the à trous algorithm are (Starck et al. 1995):

1. reasonable computational requirements,
2. easy to program,
3. in 2-D, the transform is nearly isotropic,
4. compact scaling functions can be used,
5. the reconstruction is trivial,
6. the transform is known at each pixel, allowing detection with no interpolation and no error,
7. the evolution of the transform can be followed from one scale to the next,
8. it is invariant under translation.

The à trous algorithm is based on the discrete scalar product at k pixels of the function $f(x)$ with the scaling function $\phi(x)$, which corresponds to a low-pass filter. The result of this product is the sampled data $c_0(k)$:

$$c_0(k) =< f(x), \phi(x-k) > . \tag{D.24}$$

The scaling function must satisfy the dilation equation:

$$\frac{1}{2}\phi(\frac{x}{2}) = \sum_l g(l)\phi(x-l), \tag{D.25}$$

where g is a discrete low-pass filter associated with the scaling function ϕ. The distance between scales is increased by a factor of 2, so the transform is dyadic. The smoothed data $c_j(k)$ at a given resolution j and at a certain position k is given by the scalar product:

$$c_j(k) = \frac{1}{2^j} < f(x), \phi(\frac{x-k}{2^j}) >, \tag{D.26}$$

which is obtained by:

$$c_j(k) = \sum_l g(l) c_{j-1}(k + 2^{j-1} l). \tag{D.27}$$

Therefore, the wavelet coefficients at resolution j are given by the difference:

$$\omega_j(k) = c_{j-1}(k) - c_j(k), \tag{D.28}$$

which can be expressed as a scalar product too:

$$\omega_j(k) = \frac{1}{2^j} < f(x), \psi(\frac{x-k}{2^j}) > . \tag{D.29}$$

This is DWT for a resolution j. The wavelet function ψ is defined by:

$$\frac{1}{2}\psi(\frac{x}{2}) = \phi(x) - \frac{1}{2}\phi(\frac{x}{2}). \tag{D.30}$$

A B-spline of degree 3 is used for the scaling function $\phi(x)$. The coefficients for a convolution mask in one dimension are ($\frac{1}{16}\ \ \frac{1}{4}\ \ \frac{3}{8}\ \ \frac{1}{4}\ \ \frac{1}{16}$), from which a 5x5 mask for 2 dimensions can easily be devised :

$$\begin{pmatrix} \frac{1}{256} & \frac{1}{64} & \frac{3}{128} & \frac{1}{64} & \frac{1}{256} \\ \frac{1}{64} & \frac{1}{16} & \frac{3}{32} & \frac{1}{16} & \frac{1}{64} \\ \frac{3}{128} & \frac{3}{32} & \frac{9}{64} & \frac{3}{32} & \frac{3}{128} \\ \frac{1}{64} & \frac{1}{16} & \frac{3}{32} & \frac{1}{16} & \frac{1}{64} \\ \frac{1}{256} & \frac{1}{64} & \frac{3}{128} & \frac{1}{64} & \frac{1}{256} \end{pmatrix}$$

There are several ways to handle the boundaries, the most general is probably to consider $c(k + N) = c(N - k)$. Other methods, such as periodicity $c(k + N) = c(k)$ or continuity $c(k + N) = c(N)$ can be used.

Finally, the original image c_0 can be expressed in terms of its wavelet coefficients as a series expansion:

$$c_0(k) = c_p + \sum_{j=1}^{p} \omega_j(k), \tag{D.31}$$

where the last smoothed array $c_p(k)$ must be added to all the wavelet coefficients, corresponding to all the scales of representation, in order to obtain a reconstruction formula of the original data.

At each scale j we obtain a wavelet plane represented by the set of wavelet coefficients $\{\omega_j\}$, which has the same number of pixels as the original image and, consequently, this WT is not very suitable for image compression purposes. The steps involved in the à trous algorithm can be summarized as follows:

1. Set j to 0 and start with data $c_j(k)$.
2. Increment j, and carry out the discrete convolution of the data $c_{j-1}(k)$ using the filter g. The distance between the central pixel and the adjacent ones is 2^{j-1}.
3. Wavelet plane j is given by the difference $c_{j-1}(k) - c_j(k)$.
4. If j is less than the number p of resolutions we want to achieve, then go to step 2.
5. The final set $W = \{\omega_0, ..., \omega_p, c_p\}$ is the wavelet representation of the data.

Image D.5 shows a wavelet decomposition of galaxy M100 using the à trous algorithm. All the wavelet images $\{\omega_j\}$ have the same size as the original image (subpanel A). It is possible to observe how fine details (e.g., noise, faint stars, etc.) have been classified into high-frequency wavelet planes (subpanels B and C) whereas wide structures (e.g., galaxy core) are represented in low-frequency planes (subpanels E and F). It is interesting to compare this image with figure E.8, corresponding to the curvelet decomposition of the same galaxy.

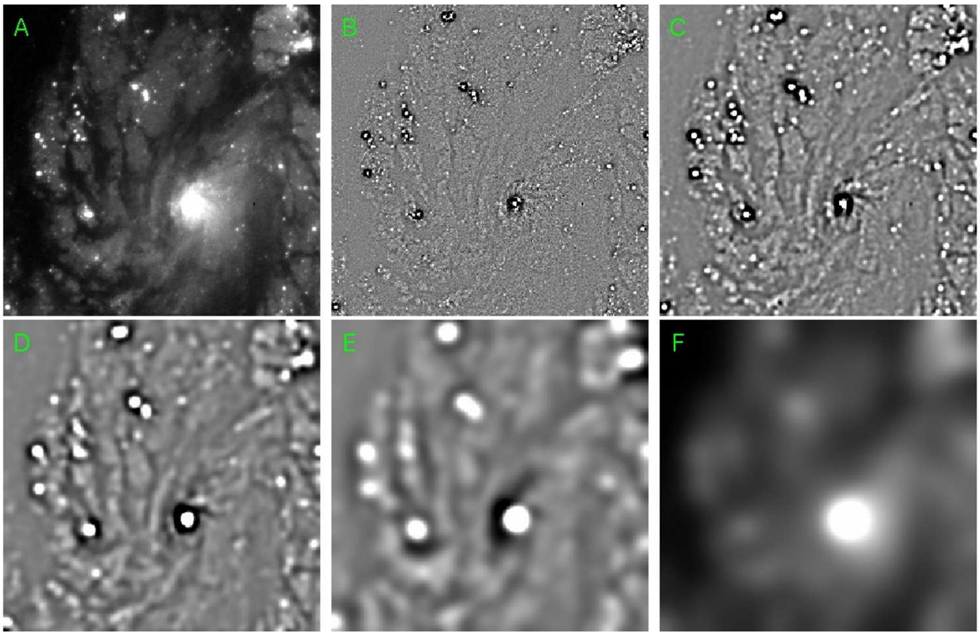

Figure D.5: Decomposition of a 256x256 image of galaxy M100 in 4 wavelet planes and 1 residual smoothed array. A) original image, B) ω_0, C) ω_1, D) ω_2, E) ω_3 and F) final smoothed array c_p. Images are represented in zscale.

Appendix E

The Curvelet Transform.

The curvelet transform (CT) was designed by Candès & Donoho (2001, 2002) to generalize the wavelet transform (WT) as well as to overcome some of its drawbacks. WT performs very well at representing point-like singularities and isotropic features at all scales and locations; however, it does not properly describe highly anisotropic elements, such as lines or curves, since wavelets are non-geometrical and not sensitive to the regularity of edges. Following this reasoning, CT was proposed as an effective model that not only considers a multiscale time-frequency local partition, but also makes use of the direction of features. Furthermore, in the two-dimensional case, CT allows an almost optimal sparse representation of objects. Although CT has been used in the astronomical context for different purposes, such as contrast enhancement of or artifact removal from Saturn images (Starck et al. 2003), the study of stellar oscillations (Lambert et al. 2006) and the detection of non-Gaussian signatures in cosmic microwave background observations (Starck et al. 2004), its use is far from being as general as that of WT.

CT belongs to the family of multiscale directional transforms that whose aim is to identify geometric features. Other members of this family go by the names of: ridgelet (Candès & Donoho 1999), contourlet (Do & Vetterli 2005), shearlet (Guo & Labate 2007), beamlet (Donoho & Huo 2002), bandlet (Mallat & Peyré 2007), platelet (Willett & Nowak 2003), surfacelet (Lu & Do 2007), wedgelet (Donoho 1999), grouplet (Mallat 2009), directionlet (Velisavljevic et al. 2006) and directional wavelet transforms (Fernandes et al. 2003). A useful review and summary of CT is outlined by Ma & Plonka (2010).

E.1 The Continuous Curvelet Transform

The continuous CT (CCT) is described by Candès & Donoho (2003a,b). A curvelet family of complex-valued waveforms is constructed in the frequency domain using polar coordinates by defining the window $U_a(r, \omega)$:

$$U_a(r,\omega) := a^{3/4} W(ar) V\left(\frac{\omega}{\sqrt{a}}\right) \qquad 0 < a \leq 1 \tag{E.1}$$

where $V(t)$ and $W(r)$ are the so-called angular and radial windows, respectively, which can be defined by means of scaled Meyer windows (Daubechies 1992):

$$V(t) = \begin{cases} 1 & if \quad |t| \leq 1/3, \\ cos[\frac{\pi}{2}\nu(3|t|-1)] & if \quad 1/3 \leq |t| \leq 2/3, \\ 0 & otherwise, \end{cases} \tag{E.2}$$

$$W(r) = \begin{cases} cos[\frac{\pi}{2}\nu(5-6r)] & if \quad 2/3 \leq r \leq 5/6, \\ 1 & if \quad 5/6 \leq r \leq 4/3, \\ cos[\frac{\pi}{2}\nu(3r-4)] & if \quad 4/3 \leq r \leq 5/3, \\ 0 & otherwise, \end{cases} \tag{E.3}$$

and $\nu(x)$ is a smooth function that must satisfy $\nu(x) + \nu(1-x) = 1$, as well as:

$$\nu(x) = \begin{cases} 0 & if \quad x \leq 0, \\ 1 & if \quad x \geq 1. \end{cases} \tag{E.4}$$

Some examples of $\nu(x)$ are provided by Ma & Plonka (2010) (e.g., $\nu(x) = x, \quad x \in [0, 1]$). The U_a support is a polar wedge that depends on the W and V supports (Fig. E.1 for $a = 1$ and Fig. E.2 for $a = 1/2$ and $a = 1/8$). It is possible to observe the effect of scaling: U_a becomes longer and thinner for decreasing a.

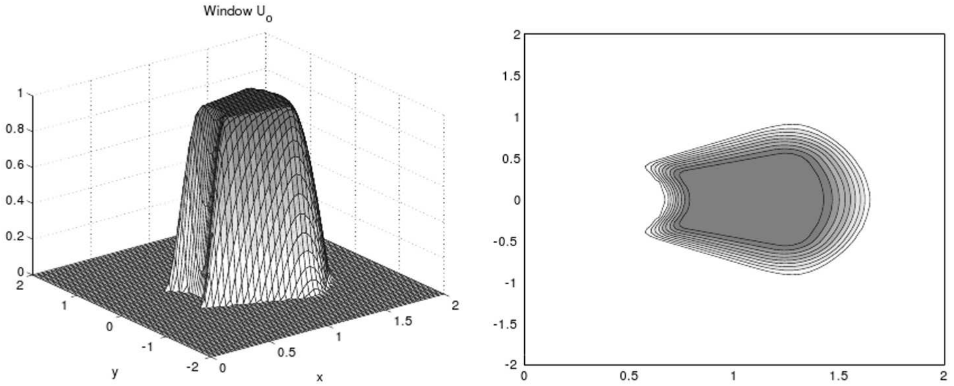

Figure E.1: Left: window $U_1(\xi)$. Right: the corresponding support. —After Ma & Plonka (2010) —.

A basic curvelet element is then given by the Fourier transform of the window U_a. Let $\varphi_{a,0,0} \in L^2(\mathbb{R}^2)$:

$$\hat{\varphi}_{a,0,0}(\xi) := U_a(\xi), \tag{E.5}$$

and the whole family is generated by rotation and translation of the basic element:

$$\varphi_{a,b,\theta}(x) := \varphi_{a,0,0}(R_\theta(x-b)), \tag{E.6}$$

with the translation $b \in \mathbb{R}^2$, and where $R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ is the 2x2 rotation matrix with angle θ. Thus, the curvelet function $\varphi_{a,b,\theta}$ is a 3-parameter function: the scale $a \in (0,1]$, the location $b \in \mathbb{R}^2$ and the orientation $\theta \in [0, 2\pi)$.

It should be noted that rotation in the spatial domain through an angle θ corresponds to rotation in the frequency domain, also with angle θ since

$$\hat{\varphi}_{a,b,\theta}(\xi) = e^{-i\langle b,\xi\rangle}\hat{\varphi}_{a,0,0}(R_\theta\xi) = e^{-i\langle b,\xi\rangle}U_a(R_\theta\xi). \tag{E.7}$$

The curvelet functions $\varphi_{a,b,\theta}$ are complex and have the following properties:

1. Support in the frequency domain: the $\hat{\varphi}_{a,b,0}$ support can be seen in Figures E.1 and E.2 for $\theta = 0$ and for different values of a. For $\theta > 0$, this support is rotated clockwise by θ.

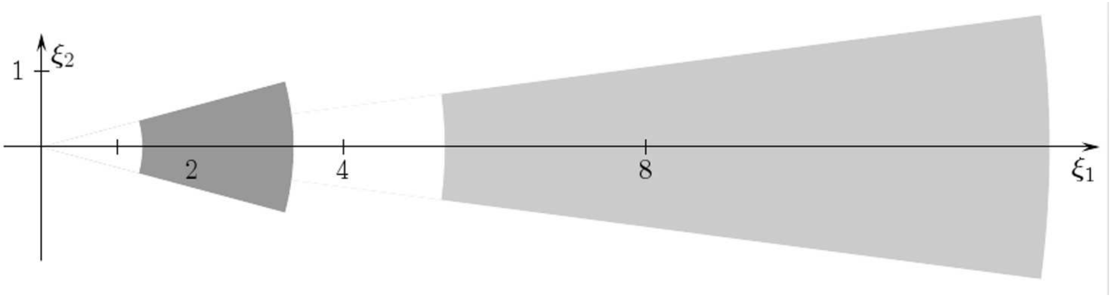

Figure E.2: Supports of the windows $U_{1/2}(\xi)$ (gray) and $U_{1/8}(\xi)$ (light gray). —After Ma & Plonka (2010) —.

2. Support in the time domain and oscillation properties: since $\hat{\varphi}_{a,b,\theta}$ has compact support, the curvelet function $\varphi_{a,b,\theta}$ cannot have compact support in the time domain. From Fourier analysis, it is known that the decay of $\varphi_{a,b,\theta}(x)$ for large $|x|$ depends on the smoothness of $\hat{\varphi}_{a,b,\theta}$ in frequency domain. The smoother $\hat{\varphi}_{a,b,\theta}$, the faster the decay. By definition, $\hat{\varphi}_{a,0,0}$ is supported away from the vertical axis $\xi_1 = 0$ but near the horizontal axis $\xi_2 = 0$. Hence, for small $a \in (0,1]$ the function $\varphi_{a,0,0}$ oscilates less in the x_2-direction (with a frequency of approximately $\sqrt{a}$) and oscillates much more in the x_1-direction (with frequencies of approximately $1/2a$). The essential support of the amplitude spectrum of $\varphi_{a,0,0}$ is a rectangle of size $[-\pi/2a, \pi/2a] \times [-\pi/\sqrt{a}, \pi/\sqrt{a}]$, and the decay of $\varphi_{a,0,0}$ away from this rectangle depends on the smoothness of U_a with respect to the function v. From equation E.6 we can clearly see that the essential support of $\varphi_{a,b,\theta}(x)$ is the rectangle rotated through the angle θ and translated a distance $R_\theta b$.

3. Vanishing moments: since the support of $\hat{\varphi}_{a,b,\theta}$ is away from $(0,0)$, the functions $\varphi_{a,b,\theta}$ have a

mean value of zero. Furthermore, $\hat{\varphi}_{a,b,\theta}$ have infinite directional moments, i.e., for all bivariate polynomials $p(x)$ with $x \in \mathbb{R}^2$ of arbitrary degree and for all angles θ we find that:

$$\int_{\mathbb{R}^2} p(R_\theta b)\varphi_{a,b,\theta}(x)dx = 0, \tag{E.8}$$

where again R_θ denotes the rotation matrix with angle θ. This observation is a direct consequence of the smoothness of $\varphi_{a,b,\theta}$.

Finally, the CCT Γ_f of $f \in \mathbf{L}^2(\mathbb{R}^2)$ is given by an inner product:

$$\Gamma_f(a, b, \theta) := \langle \varphi_{a,b,\theta}, f \rangle = \int_{\mathbb{R}^2} \varphi_{a,b,\theta}(x)\overline{f(x)}dx. \tag{E.9}$$

The curvelet coefficients $\langle \varphi_{a,b,\theta}, f \rangle$ contain all the information about f if its Fourier transform vanishes for $|\xi| < 2$, in which case, f can be recovered with the following Calderón-like formula (Candès & Donoho 2003b):

$$f(x) = \frac{1}{(\ln 2)} \int_0^{2\pi} \int_{\mathbb{R}^2} \int_0^1 \Gamma_f(a, b, \theta)\varphi_{a,b,\theta}(x)\frac{da}{a^{3/2}}\frac{db}{a^{1/2}}\frac{d\theta}{a}. \tag{E.10}$$

This formula is valid for high-frequency scales and is generalized for all frequencies by Candès & Donoho (2003b).

E.2 The Discrete Curvelet Transform

The first implementation of the discrete CT (DCT) was performed as a combination of the WT and the ridgelet transform (RT). Since the ridgelet transform is better suited to straight lines than curves or edges (Starck et al. 2010), the idea was to analyze different scales of representation locally in order to approximate possible curvatures as straight lines. Hence, the image was decomposed by means of WT, in order to obtain different scales of resolution, and each one of the wavelet planes was then divided into blocks (whose size depended on the resolution) which could be analyzed independently by RT. This was the so-called first generation of curvelets (Candès & Donoho 2002). However, such curvelet construction presented a complicated seven-parameter structure and was computationally very expensive.

In contrast, the second generation (Candès & Donoho 2004; Candès et al. 2006) exhibit a simpler index structure with only three parameters and much lower redundancy, yielding faster algorithms. Hence, DCT can be obtained by conveniently sampling the three parameters that define the scale, location and orientation. We therefore have:

- the scales $a_j := 2^{-j}, \quad j \geq 0$;
- the sequence of rotation angles $\theta_{j,l} := \frac{\pi l 2^{\lceil j/2 \rceil}}{2}, \quad$ with $\quad l = 0, 1, ..., 4 \cdot 2^{\lceil j/2 \rceil} - 1$, where $\lceil x \rceil$ denotes the smallest integer greater than or equal to x;

- the positions $b_k^{j,l} = b_{k_1,k_2}^{j,l} := R_{\theta_{j,l}}^{-1}\left(\frac{k_1}{2^j}, \frac{k_2}{2^{j/2}}\right)^T$, with $k_1, k_2 \in \mathbb{Z}$, and where R_θ denotes the rotation matrix with angle θ.

For example, for $j = 0$ we consider the angles $\theta_{0,l} = \pi l/2$, with $l = 0, 1, 2, 3$ and the positions $b_k^{j,l} \in \mathbb{Z}^2$. For $j = 4$, the angles $\theta_{4,l} = \pi l/8$, with $l = 0, ..., 15$ and, depending on the angles $\theta_{4,l}$, eight different grids for translation are considered, where rectangles of size $1/16 \times 1/4$ are rotated by $\theta_{4,l}$, with $l = 0, ..., 7$ (Fig. E.3). In particular, the choice of positions yields a parabolic scaling of the grids with the relationship *length* $\approx 2^{-j/2}$ and *width* $\approx 2^{-j}$.

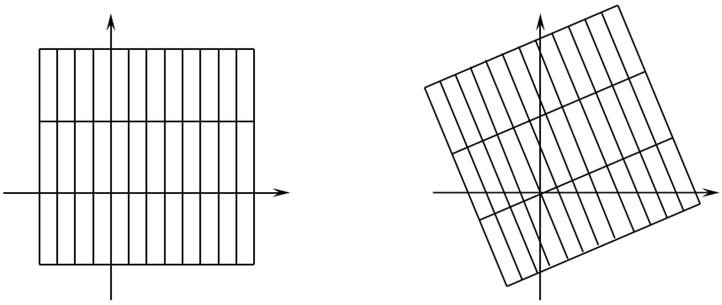

Figure E.3: Grid for $\theta_{4,0} = 0$ and for $\theta_{4,1} = \pi/8$. —After Ma & Plonka (2010) —.

As for CCT, we can define the scaled windows in polar coordinates:

$$U_j(r, \omega) := 2^{-3j/4} W(2^{-j} r) V\left(\frac{2 \cdot 2^{\lceil j/2 \rceil} \omega}{\pi}\right) = 2^{-3j/4} W(2^{-j} r) V\left(\frac{\omega}{\theta_{j,1}}\right), \quad j \in \mathbb{N}_0, \tag{E.11}$$

where both $W(r)$ and $V(\omega)$ are smooth, nonnegative and real-valued. The basic curvelet is defined by:

$$\hat{\phi}_{j,0,0}(\xi) := U_j(\xi), \tag{E.12}$$

and the family of curvelet functions is given by:

$$\phi_{j,k,l}(x) := \phi_{j,0,0}(R_{\theta_{j,l}}(x - b_k^{j,l})). \tag{E.13}$$

In the frequency domain, the curvelet functions

$$\hat{\phi}_{j,k,l}(\xi) = e^{-i\langle b_k^{j,l}, \xi \rangle} U_j(R_{\theta_{j,l}} \xi) = e^{-i\langle b_k^{j,l}, \xi \rangle} 2^{-3j/4} W(2^{-j} r) V\left(\frac{\omega + \theta_{j,l}}{\theta_{j,l}}\right), \tag{E.14}$$

are supported inside the polar wedge with radius $2^{j-1} \le r \le 2^{j+1}$ and angle $\frac{2^{-\lceil j/2 \rceil}\pi(-1-l)}{2} < \omega < \frac{2^{-\lceil j/2 \rceil}\pi(1-l)}{2}$. The $\hat{\phi}_{j,k,l}$ support does not depend on the position $b_k^{j,l}$. For example, $\hat{\phi}_{2,k,l}(r, \omega)$ is supported inside the wedge with $2 \le r \le 8$ and $\frac{(-1-l)\pi}{4} \le \omega \le \frac{(1-l)\pi}{4}$, for $l = 0, ..., 7$ (Fig. E.4).

Every function $f \in L^2(\mathbb{R}^2)$ can be represented as a curvelet series. The curvelet coefficients are obtained from the inner product $c_{j,k,l}(f) := \langle f, \phi_{j,k,l} \rangle$:

$$c_{j,k,l}(f) := \int_{\mathbb{R}^2} f(x) \overline{\phi_{j,k,l}(x)} dx = \int_{\mathbb{R}^2} \hat{f}(\xi) \overline{\hat{\phi}_{j,k,l}(\xi)} d\xi. \tag{E.15}$$

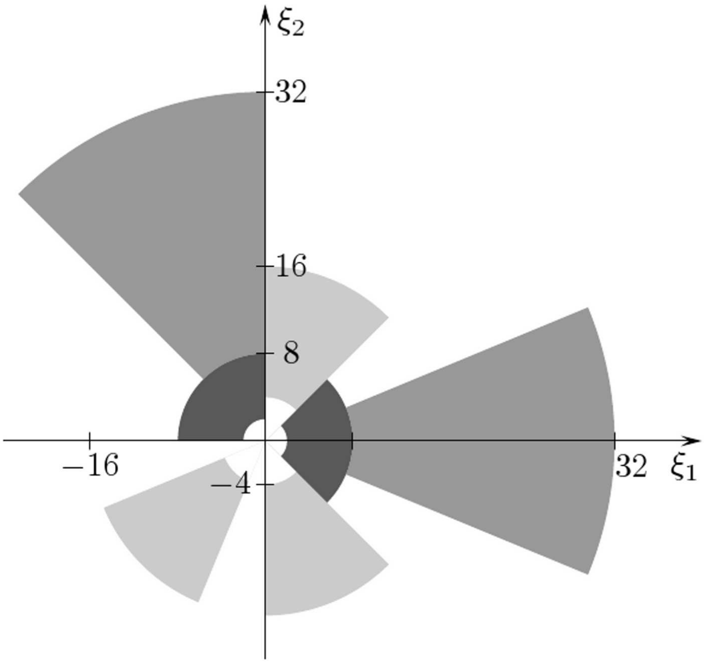

Figure E.4: Maximal supports of $\hat{\phi}_{2,k,0}$ and $\hat{\phi}_{2,k,5}$ (dark gray); of $\hat{\phi}_{3,k,3}$, $\hat{\phi}_{3,k,6}$ and $\hat{\phi}_{3,k,13}$ (light gray); and of $\hat{\phi}_{4,k,0}$ and $\hat{\phi}_{4,k,11}$ (mid gray). Supports $V \subset [-1, 1]$ and $W \subset [-1/2, 2]$. —After Ma & Plonka (2010) —.

where equation E.15 makes use of Plancherel's theorem ($j \geq 0$). This construction implies that curvelets exhibit an oscillatory pattern in the direction perpendicular to their orientation.

Therefore, the discrete transform takes as its input data that are defined on a Cartesian grid, and it outputs a collection of coefficients. The continuous-space definition of the CT uses coronae and rotations that are not especially adapted to Cartesian arrays. Hence, it is desirable to replace these concepts by their Cartesian counterparts, i.e., concentric squares instead of concentric circles and shears instead of rotations (Fig. E.5).

Hence, it would be necessary to substitute the angular and radial windows by their Cartesian equivalents:

$$\tilde{U}_j(\xi) := 2^{-3j/4}\tilde{W}_j(\xi_1)V_j(\xi) = 2^{-3j/4}\tilde{W}_j(\xi_1)V\left(\frac{2^{\lceil j/2 \rceil}\xi_2}{\xi_1}\right), \tag{E.16}$$

where $\tilde{W}_j$ can be a window based on the subtraction of low-pass filters, i.e., with a band-pass filter profile (see Candès et al. (2006) for some examples), and $\tilde{V}_j$ remains the same. The frequencies are then determined in the trapezoid:

$$\{(\xi_1, \xi_2) : 2^{j-1} \leq \xi_1 \leq 2^{j+1}; -2^{-\lceil j/2 \rceil} \cdot \frac{2}{3} \leq \xi_2/\xi_1 \leq 2^{-\lceil j/2 \rceil} \cdot \frac{2}{3}\}. \tag{E.17}$$

The window $\tilde{U}_0$ is presented in Figure E.6, which is the Cartesian equivalent of Figure E.1.

Now we define a set of equispaced slopes instead of equidistant angles:

$$\tan\theta_{j,l} := l2^{-\lceil j/2 \rceil}, \quad l = -2^{\lceil j/2 \rceil} + 1, ..., 2^{\lceil j/2 \rceil} - 1. \tag{E.18}$$

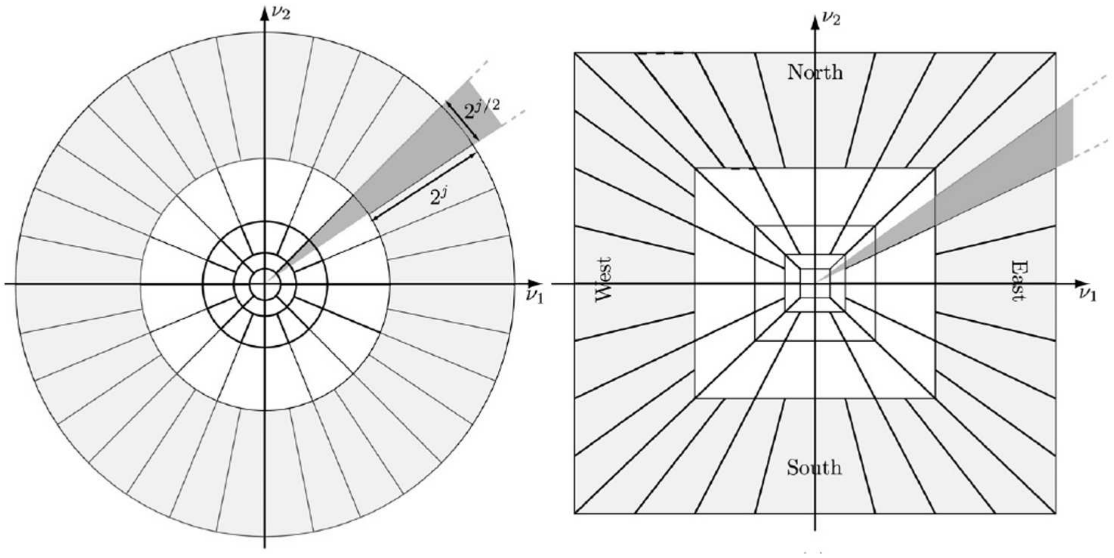

Figure E.5: Left: Continuous curvelet frequency tiling. A wedge is obtained as the product of the radial window (annulus shown in a lighter shade) and the angular window (darker area). It's dimensions are shown as 2^J by $2^j/2$. Right: Discrete curvelet frequency tiling. The window $\tilde{U}_j(\xi)$ isolates the frequency near a trapezoidal wedge such as those shown in light gray. —After Starck et al. (2010) —.

Which means that the curvelet-like functions are given by:

$$\hat{\tilde{\phi}}_{j,0,0} := \tilde{U}_j(\xi), \tag{E.19}$$

$$\tilde{\phi}_{j,k,l}(x) := \tilde{\phi}_{j,0,0}\left(S^T_{\theta_{j,l}}(x - \tilde{b}^{j,l}_k)\right), \tag{E.20}$$

which is the Cartesian counterpart of equation E.13, with the shear matrix $S_\theta = \begin{pmatrix} 1 & 0 \\ -\tan\theta & 1 \end{pmatrix}$ and where $\tilde{b}^{j,l}_k := S^{-T}_{\theta_{j,l}}(k_1 2^{-j}, k_2 2^{-\lceil j/2 \rceil})$. The set of curvelets $\tilde{\phi}_{j,k,l}$ needs to be completed by symmetry and by rotation through $\pm\pi/2$ radians in order to obtain the whole family. Furthermore, it is neces-

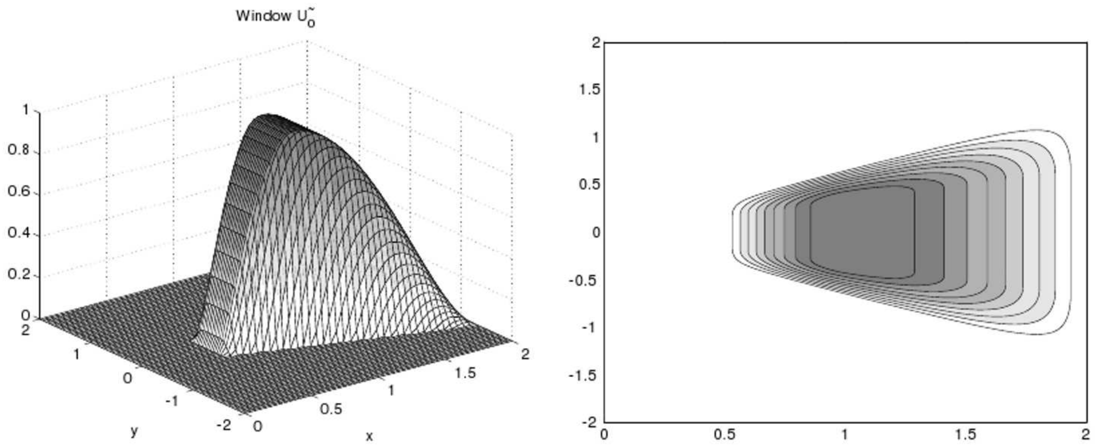

Figure E.6: Left: window $\tilde{U}_0(\xi)$. Right: the corresponding support. —After Ma & Plonka (2010) —.

sary to consider suitable corner elements to connect the four cones (north, west, south and east) as well as extending this support to coarse curvelet elements for low frequencies, as was done for CCT (Ma & Plonka 2010).

The digital implementation of DCT, whose coefficients are now given by:

$$c_{j,k,l}(f) := \langle f, \tilde{\phi}_{j,k,l} \rangle = \int_{\mathbb{R}^2} \hat{f}(\xi)\tilde{U}_j(S^{-1}_{\theta_{j,l}}\xi)e^{i\langle \tilde{b}^{j,l}_k, \xi\rangle} d\xi, \tag{E.21}$$

can be evaluated by the following steps for discrete data:

1. Compute the 2-D FFT to get $\hat{f}$.
2. Form the windowed frequency data $\hat{f}\tilde{U}_j$.
3. Apply the inverse Fourier transform.

However, the last step needs to evaluate the FFT at the sheared grid $S_{\theta^{-T}_{j,l}}$, for which the classical FFT algorithm is not valid. Candès et al. (2006) proposes two possible implementations, essentially differing in the way in which they handle the grid. The first is called unequispaced-FFT (USFFT) and consists of a tilted grid closely aligned with the axes of $\tilde{U}_j$. This implementation uses a nonstandard interpolation and the inverse transform uses conjugate gradient iteration to invert the interpolation step, which has its computational cost.

Alternatively, the wrapping-based DCT uses a grid aligned with the Cartesian input grid, so it makes for a simpler choice of the spatial grid to translate the curvelets. The curvelet coefficients are approximated using $k_j = (k_1 2^{-j}, k_2 2^{-\lceil j/2 \rceil})^T$ instead of $\tilde{b}^{j,l}_k$ with values on a rectangular grid. However, a difficulty arises again because the window $\tilde{U}_j$ does not fit in a rectangle of size $2^j \times 2^{j/2}$, to which an inverse FFT could be applied. The wrapping consists of periodizing the windowed frequency data $\hat{f}\tilde{U}_j$ and reindexing the sample array by wrapping it around an approximate $2^j \times 2^{j/2}$ rectangle centered on the origin (Fig. E.7).

The algorithm to compute DCT via wrapping contains the following steps:

1. Apply the 2-D FFT and obtain the Fourier coefficients $\hat{f}(\xi)$.
2. For each scale j and orientation l:
    - form the product $\hat{f}\tilde{U}_j$;
    - wrap this product around the origin;
    - apply the inverse 2-D FFT to the wrapped data to get curvelet coefficients $c_{j,k,l}$.

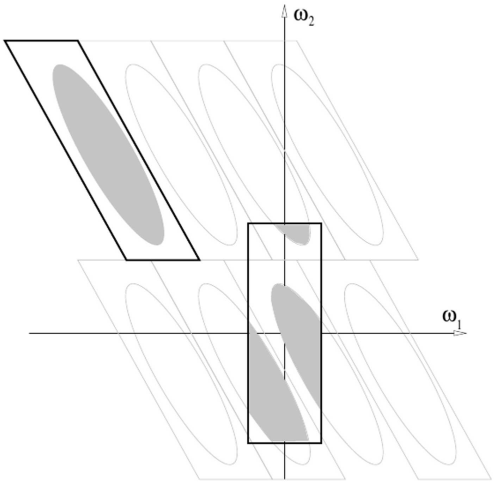

Figure E.7: Wrapping data in a parallelepiped by periodicity. Here the angle θ is in the range $(\pi/4, 3\pi/4)$. The parallelogram is the tile that contains the curvelet frequency support. The rectangle is centered on the origin. The wrapped ellipse appears broken but this is not an issue in the periodic rectangle where opposite edges are identified. —After Candès et al. (2006) —.

Image E.8 shows the curvelet coefficients at each scale of galaxy M100 transferred to the image domain. Although it is more convenient to work with the coefficients in the curvelet domain, where all the capabilities of CT can be fully exploited, it is anyway interesting to compare this image with Figure D.5 to see differences in the classification process between curvelets and wavelets (Baena Gallé & Núñez 2010). Notice that the visual information associated with the galactic cloud hardly appears in the high-frequency curvelet planes (Fig. E.8 subpanels B and C) whereas its contribution to the corresponding wavelet plane is significant (Fig. D.5 subpanels B and C).

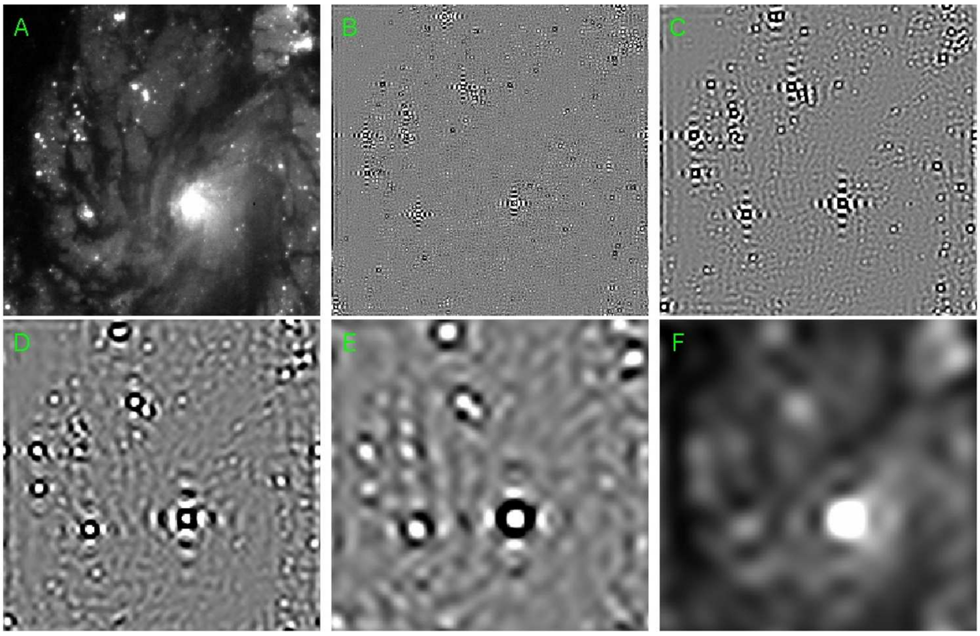

Figure E.8: Decomposition of a 256x256 image of galaxy M100 in 4 curvelet planes (transferred to the image domain) and 1 residual smoothed array. A) original image, B) c_0, C) c_1, D) c_2, E) c_3 and F) final smoothed array. Images are represented in zscale.

Bibliography

Andrews, H.C. & Hunt B.R. 1977, Digital Image Restoration (Prentice-Hall, Inc., Englewood Cliffs, New York).

Antichi, J., Dohlen, K., Gratton, R.G., Mesa, D., Claudi, R.U., et al. BIGRE: A Low Cross-Talk Integral Field Unit Tailored for Extrasolar Planets Imaging Spectroscopy. 2009, The Astrophysical Journal, 695(2), 1042-1057.

Antichi, J., Verinaud, C., Preis, O., Delboulbé, A., Zins, G., Rabou, P., Beuzit, J-L, et al. FFREE: a Fresnel-Free Experiment for EPICS, the EELT planets imager. 2010, Proceedings of the SPIE, 7736, 773639.

Apai, D., Pascucci, I., Brandner, W., Henning, T., Lenzen, R., Potter, D.E., Lagrange, A.M. & Rousset, G. NACO polarimetric differential imaging of TW Hya. A sharp look at the closest T Tauri disk. 2004, Astronomy and Astrophysics, 415, 671-676.

Aubailly, M., Vorontsov, M.A., Carhart, G.W. & Valley, M.T. Automated video enhancement from a stream of atmospherically-distorted images: the lucky-region fusion approach. 2009, Proceeding of the SPIE, 7463, 74630C.

Babcock, H.W. The Possibility of Compensating Astronomical Seeing. 1953, Publications of the Astronomical Society of the Pacific, 65(386), 229-236.

Baena Gallé, R. & Núñez, J. Effects of the Curvelet Transform over Interferometric Images. 2010, International Journal of Imaging Systems and Technologies, 20(4), 333-353.

Baena Gallé, R. & Gladysz, S. Estimation of differential photometry in adaptive optics observations with a wavelet-based maximum likelihood estimator. 2011, Publications of the Astronomical Society of the Pacific, 123(905), 865-878.

Baena Gallé, R., Núñez, J. & Gladysz, S. Extended object reconstruction in adaptive-optics imaging: the multiresolution approach. 2013, Astronomy and Astrophysics, 555(A69).

Baena Gallé, R., Gladysz, S., Verinaud, C. & Kasper, M. High contrast exoplanet imaging using spectral deconvolution and the wavelet transform. Application to EPICS instrument. 2013, *In preparation.*

Barlow, C.W.C. & Bryan, G.H. Elementary Mathematical Astronomy. University Tutorial Press (London). 1944.

Bartko, H., Martins, F., Trippe, S., Fritz, T.K., Genzel, R., et al. An extremely top-heavy initial mass function in the galactic center stellar disks. 2010, The Astrophysical Journal, 708(1), 834-840.

Bastian, N., Covey, K. & Meyer, M. A Universal Stellar Initial Mass Function? A Critical Look at Variations. 2010, Annual Review of Astronomy and Astrophysics, 48, 339-389.

Bate, M.R. Stellar, brown dwarf and multiple star properties from hydrodynamical simulations of star cluster formation. 2009, Monthly Notices of the Royal Astronomical Society, 392(2), 590-616.

Bauman, B. J., Gavel, D. T., Waltjen, K. E., Freeze, G. J., Keahi, K. A., Kuklo, T. C., Lopes, S. K., Newman, M. J. et al. New optical design of adaptive optics system at Lick Observatory, 1999, Proceedings of the SPIE, 3762, 194-200.

Belen'kii, M.S., Stewart, J.M. & Gillespie, P. Turbulence-induced edge image waviness: theory and experiment. 2001, Applied Optics, 40, 1321-1328.

Bertero, M. & Boccacci, P. Introduction to inverse problems in imaging. Institute of Physics Publishing (London). 1998.

Beuzit, J-L., Feldt, M., Dohlen, K., Mouillet, D., Puget, P., et al. SPHERE: a planet finder instrument for the VLT. 2008, Proceedings of the SPIE, 7014, 701418.

Biller, B., Allers, K., Liu, M., Close, L.M. & Dupuy, T. A Keck LGS AO search for brown dwarf and planetary mass companions to upper Scorpius brown dwarfs. 2011, The Astrophysical Journal, 730(1), article id. 39 (14pp).

Blanco, L. & Mugnier L.,M. Marginal blind deconvolution of adaptive optics retinal images, 2011, Optics Express, 19(23), 23227.

Boccaletti, A., Carbillet, M., Fusco, T., Mouillet, D., Langlois, M., Moutou, C. & Dohlen, K. End-to-end simulation of AO-assisted coronagraphic differential imaging: estimation of performance for SPHERE. 2008, Proceedings of the SPIE, 7015, 70156E.

Bongard, S., Soulez, F., Thiébaut, É. & Pécontal, É. 3D deconvolution of hyper-spectral astronomical data. 2011, Monthly Notices of the Royal Astronomical Society, 418(1), 258-270.

Born, M. & Wolf, E. Principles of optics (2nd ed.). Pergamon Press. (Oxford). 1964.

Bouman, C. & Sauer, K. A Generalized Gaussian Image Model for Edge-Preserving MAP Estimation, 1993, IEEE Transactions on Image Processing, 2(3), 296-310.

Brown, D.S. The prediction of on site telescope performance. 1979, The messenger, ESO, 19, 16-18.

Brown, D.S. Scientific importance of high angular resolution at IR and optical wavelengths. 1981, Proceedings of the ESO Conference, ESO, 181.

Brown, R. New Information from Radial Velocity Data Sets. 2004, The Astrophysical Journal, 610(2), 1079-1092.

ten Brummelaar T.A., Mason, B.D., Bagnuolo, W.G., Jr., Hartkopf, W.I., McAlister, H.A. & Turner, N.H. Differential binary star photometry using the adaptive optics system at Starfire Optical Range, 1996, Astronomical Journal, 112, 1180-1187.

ten Brummelaar, T.A., Mason, B.D., McAlister, H.A., Roberts, L.C., Turner, N.H. et al. Binary Star Differential Photometry Using the Adaptive Optics System at Mount Wilson Observatory, 2000, Astronomical Journal, 119(5), 2403-2414.

Buffington, A., Crawford, F.S., Muller, R.A., Schemin, A.J. & Smits, R.G. Correction of atmospheric distortion with an image-sharpening telescope. 1977a, Journal of the Optical Society of America, 67, 298-303.

Buffington, A., Crawford, F.S., Muller, R.A., & Orth, C.D. First observatory results with an image-sharpening telescope. 1977b, Journal of the Optical Society of America, 67, 304-305.

Burke, D., Gladysz, S., Roberts, L., Devaney, N. & Dainty, C. An Improved Technique for the Photometry and Astrometry of Faint Companions. 2009, Publications of the Astronomical Society of the Pacific, 121(881), 767-777.

Burt, P.J. & Adelson, E.H. The Laplacian pyramid as a compact image code. 1983, IEEE Transactions on Communications, C-31(4), 532-540.

Cagigal, M. P. & Canales, V. F. Generalized Fried parameter after adaptive optics partial wavefront compensation, 2000, Journal of the Optical Society of America A, 17(5), 903-910.

Candès, E. & Donoho, D. Ridgelets: The key to high dimensional intermittency? 1999, Philosophical Transactions of the Royal Society of London, Series A 357, 2495-2509.

Candès, E. & Donoho, D. Curvelets and curvilinear integrals. 2001, Journal of Approximation Theory, 113, 59-90.

Candès, E. & Donoho, D. Recovering edges in ill-posed inverse problems: Optimality of curvelet frames. 2002, Annals of Statistics 30(3), 784-842.

Candès, E. & Donoho, D. Continuous Curvelet Transform: I. Resolution of the wavefront set. 2003a, Applied and Computational Harmonic Analysis, 19, 162-197.

Candès, E. & Donoho, D. Continuous Curvelet Transform: II. Discretization and frames. 2003b, Applied and Computational Harmonic Analysis, 19, 198-222.

Candès, E. & Donoho, D. New tight frame of curvelets and optimal representations of objects with piecewise C^2 singularities. 2004, Communications on Pure and Applied Mathematics, 57(2), 219-266.

Candès, E., Demanet, L., Donoho, D. & Ying, L. Fast discrete curvelet transforms. 2006, Multiscale Modeling and Simulation, 5(3), 861-899.

Castleman, K.R. Wavelet transforms, in Digital image processing. Prentice Hall (Upper Saddle River). 1996.

Chauvin, G., Lagrange, A.M., Dumas, C., Zuckerman, B., Mouillet, D., Song, I., Beuzit, J.L. & Lowrance, P. Giant planet companion to 2MASSW J1207334-393254. 2005, Astronomy and Astrophysics, 438(2), L25-L28.

Christou, J.C. & Bonaccini, D. Analysis of ADONIS data. Deconvolution tutoria. Imaging of T Tauri. 1996, Technical Report ESO VLT Doc. No. GEN-TRE-ESO-11620-1262.

Christou, J. C., Pugliese, G., Köhler R., Drummond, J. D. Photometric and Astrometric Analysis of Gemini/Hokupa'a Galactic Center Adaptive Optics Observations, 2004, The Publications of the Astronomical Society of the Pacific, 116(822), 734-744.

Christou, J. C. & Drummond, J. D. Measurements of Binary Stars, Including Two New Discoveries, with the Lick Observatory Adaptive Optics System, 2006, Astronomical Journal, 131(6), 3100-3108.

Chui, C.H. An introduction to wavelets, in Wavelet analysis and its applications. Academic Press (New York). 1992.

Close, L.M., Siegler, N., Freed, M. & Biller, B. Detection of nine M8.0-L0.5 binaries: the very low mass binary population and its implications for brown dwarf and very low mass star formation. 2003, The Astrophysical Journal, 587(1), 407-422.

Close, L.M., Zuckerman, B., Song, I., Barman, T., Marois, C. et al. The wide brown dwarf binary Oph 1622-2405 and discovery of a wide, low-mass binary in Ophiuchus (Oph 1623-2402): A new class of young evaporating wide binaries? 2007, The Astrophysical Journal, 660(2), 1492-1506.

Conan, J.-M., Fusco, T., Mugnier, L.M., Kersalé, E. & Michau, V. Astronomy with adaptive optics: present results and future programs. ESO Conference and Workshop Proceedings, vol. 56, 121 (Bonaccini, D. ed, Garching bei München, Germany). 1999.

Daubechies, I. Orthonormal bases of compactly supported Wavelets. 1988, Communications on Pure and Applied Mathematics, 41, 909-996.

Daubechies, I. Ten lectures on wavelets. Society for Industrial and Applied Mathematics (Philadelphia). 1992.

Davies, R., Thomas, J., Genzel, R., Müller Sánchez, F., Tacconi, L.J., Sternberg, A., Eisenhauer, F., Abuter, R., Saglia, R. & Bender, R. The star-forming torus and stellar dynamical black hole mass in the Seyfert 1 nucleus of NGC 3227. 2006, The Astrophysical Journal, 646(2), 754-773.

Davies, R.I., Maciejewski, W., Hicks, E.K.S., Tacconi, L.J., Genzel, R. & Engel, H. Stellar and molecular gas kinematics of NGC 1097: inflow driven by a nuclear spiral. 2009, The Astrophysical Journal, 702(1), 114-128.

Davies, R. & Kasper, M. Adaptive optics for astronomy. 2012, Annual Review of Astronomy and Astrophysics, 50, 305-351.

Demanet, L. & Ying, L. Wave atoms and sparsity of oscillatory patterns, 2007, Applied and Computational Harmonic Analysis, 23(3), 368-387.

Denney, K.D., Peterson, B.M., Pogge, R.,W., Adair, A., Atlee, D.W., et al. Reverberation mapping measurements of black hole masses in six local Seyfert galaxies. 2010, The Astrophysical Journal, 721(1), 715-737.

Descamps, P. & Marchis, F. Angular momentum of binary asteroids: Implications for their possible origin. 2008, Icarus, 193(1), 74-84.

Dierickx, P. Optical performance of large ground-based telescopes. 1992, Journal of Modern Optics, 39(3), 569-588.

Diolaiti, E., Bendinelli, O., Bonaccini, D., Close, L., Currie, D., & Parmeggiani, G. Analysis of isoplanatic high resolution stellar fields by the StarFinder code. 2000, Astronomy and Astrophysics Supplements Series, 147, 335-346.

Do, M. & Vetterli, M. The contourlet transform: and efficient directional multiresolution image representation. 2005, IEEE Transactions on Image Processing, 14(12), 2091-2106.

Donoho, D. Wedgelets: nearly minimax estimation of edges. 1999, Annals of Statistics, 27(3), 859-897.

Donoho, D. & Huo, X. Beamleats and multiscale image analysis, in: Multiscale and Multiresolution Methods (149-196). Springer Lecture Notes in Computational Science and Engineering, vol. 20. 2002.

Duchêne G., McCabe, C., Ghez, A.M. & Macintosh, B.A. A multiwavelength scattered light analysis of the dust grain population in the GG Tauri circumbinary ring. 2004, The Astrophysical Journal, 606(2), 969-982.

Duchêne G., McCabe, C., Pinte, C., Stapelfeldt, K.R., Ménard, F., Duvert, G., Ghez, A.M. et al. Panchromatic observations and modeling of the HV Tau C edge-on disk. 2010, The Astrophysical Journal, 712(1), 112-129.

Dupuy, T.J. & Liu, M.C. On the distribution of orbital eccentricities for very low-mass binaries. 2011, The Astrophysical Journal, 733(2), article id. 122 (14pp).

Eisenhauer, F., Quirrenbach, A., Zinnecker, H. & Genzel, R. Stellar content of the galactic starburst template NGC 3603 from adaptive optics observations. 1998, The Astrophysical Journal, 498, 278-292.

Esslinger, O. & Edmunds, M.G. Photometry with adaptive optics: A first guide to expected performance. 1998, Astronomy and Astrophysics Supplements Series, 129, 617-635.

Esposito, S., Riccardi, A., Fini, L., Puglisi, A.T., Pinna, E. et al. First light AO (FLAO) system for LBT: final integration, acceptance test in Europe, and preliminary on-sky commissioning results. 2010, Proceedings of the SPIE, 7736, 773609-773609-12.

Fadili, M.J. & Starck, J.L. Sparse representation-based image deconvolution by iterative thresholding. Proceedings of the Astronomical Data Analysis Conference. ADA IV (Elsevier, France). 2006.

Fernandes, F., van Spaendonck, R. & Burrus, S. A new framework for complex wavelet transforms. 2003, IEEE Transactions on Signal Processing, 51(7), 1825-1837.

Ferrarese, L. & Ford, H. Supermassive black holes in galactic nuclei: past, present and future research. 2005, Space Science Reviews, 116(3-4), 523-624.

Ferruit, P., Pécontal, É. & Wilson, A. 3-D observations of nearby active galaxies with the IFU TIGER. 2000, ASP Conference Series, 195, 289-294.

Fitzgerald, M. P. & Graham, J. R. Speckle Statistics in Adaptively Corrected Images, 2006, Astrophysical Journal, 637(1), 541-547.

Fors, O. New observational techniques and analysis tools for wide field CCD surveys and high resolution astrometry. phD thesis. University of Barcelona (Barcelona). 2006.
(http://www.tdx.cat/TDX-0330106-125745/)

Fors, O., Richichi, A., Otazu, X. & Núñez, J. A new wavelet-based approach for the automated treatment of large sets of lunar occultation data. 2008, Astronomy and Astrophysics, 480(1), 297-304.

Förster Schreiber, N.M., Genzel, R. Bouché, N., Cresci, G., Davies, R. et al. The SINS survey: SINFONI integral field spectroscopy of $z \sim 2$ star-forming galaxies. 2009, The Astrophysical Journal, 706(2), 1364-1428.

Foster, G. Wavelets for period analysis of unevenly sampled time series. 1996, The Astronomical Journal, 112, 1709-1729.

Foy, R. & Labeyrie, A. Feasibility of adaptive telescope with laser probe, 1985, Astronomy and Astrophysics, 152(2), L29-L31.

Fried, D.L. Statistics of a geometric representation of wavefront distortion. 1965, Journal of the Optical Society of America, 55(11), 1427-1435.

Fried, D.L. Optical resolution through a randomly inhomogeneous medium for very long and very short exposures. 1966, Journal of the Optical Society of America, 56(10), 1372-1379.

Frieden, B.R. Restoring with maximum likelihood and maximum entropy. 1972, Journal of the Optical Society of America, 62(4), 511-518.

Frieden, B. Image enhancement and restoration. Springer (Berlin). 1978.

Fry, P.M. & Sromovsky, L.A. Keck 2 AO Observations of Neptune in 2003 and 2004. 2004, Bulletin of the American Astronomical Society, 36, 1103.

Fu, H., Myers, A., Djorgovski, S. & Yan, L. Mergers in double-peaked [O III] active galactic nuclei. 2011, The Astrophysical Journal, 733(2), article id. 103 (7pp).

Fusco, T., Rousset, G., Sauvage, J-F., Petit, C., Beuzit, J-L., et al. High-order adaptive optics requirements for direct detection of extrasolar planets: Application to the SPHERE instrument. 2006, Optics Express, 14(17), 7515-7534.

Galileo, G. Sidereus nuncius. Ed: Baglioni, T. (Venecia). 1610 *(Latin)*.
http://www.rarebookroom.org/Control/galsid/index.html
Translated into English by Van Helden, A. The sidereal messenger. University of Chicago Press. 1989.

Gebhardt, K., Adams, J., Richstone, D., Lauer, T.R., Faber, S.M., Gültekin, K., Murphy, J. & Tremaine, S. The black hole mass in M87 from Gemini/NIFS adaptive optics observations. 2011, The Astrophysical Journal, 729(2), article id. 119 (13pp).

Gendron, E. & Léna, P. Astronomical adaptive optics. 1: modal control optimization. 1994, Astronomy and Astrophysics, 291(1), 337-347.

Genzel, R., Eisenhauer, F. & Gillessen, S. The galactic center massive black hole and nuclear star cluster. 2010, Reviews of Modern Physics, 82(4), pp. 3121-3195.

Genzel, R., Newman, S., Jones, T., Förster Schreiber, N.M., Shapiro, K. et al. The SINS survey of $z \sim 2$ galaxy kinematics: properties of the giant star-forming clumps. 2011, The Astrophysical Journal, 733(2), article id. 101 (30pp).

Give'on, A., Kern, B., Shaklan, S., Moody, D.C. & Pueyo, L. Broadband wavefront correction algorithm for high-contrast imaging systems. 2007, Proceedings of the SPIE, 6691, 66910A-11.

Gladysz, S., Christou, J. C. & Redfern, M. Characterization of the Lick adaptive optics point spread function, 2006, Proceedings of the SPIE, 6272, 62720J, 1

Gladysz, S., Christou, J. C. Reference-Less Detection, Astrometry, and Photometry of Faint Companions with Adaptive Optics, 2009, The Astrophysical Journal, 698(1), 28-42.

Gladysz, S., Baena Gallé, R., Christou, J.C. & Roberts, L.C. Differential photometry in adaptive optics imaging. 2010a, Proceedings of AMOS Technical Conference (Maui, Hawaii), E24.

Gladysz, S., Yaitskova, N. & Christou, J. Statistics of intensity in adaptive-optics images and their usefulness for detection and photometry of exoplanets. 2010b, Journal of the Optical Society of America A, 27(11), A64-A75.

Gladysz, S., Thatte, N.A., Salter, G., Baena-Gallé, R., Clarke, F., Tecza, M. & Jolissaint, L. High-contrast, adaptive-optics simulations for HARMONI. 2nd International Conference on AO for ELT (Victoria, Canada). 2011.

Golimowski, D.A., Durrance, S.T. & Clampin, M. Coronagraphic imaging of the Beta Pictoris circumstellar disk - Evidence of changing disk structure within 100 AU. 1993, Astrophysical Journal, 411(1), L41-L44.

González Audícana M., Otazu, X., Fors, O. & Seco, A. Comparison between Mallat’s and the ’ trous’ discrete wavelet transform based algorithms for the fusion of multispectral and panchromatic images, 2005, International Journal of Remote Sensing, 26(3), 595-614.

Goode, P.R., Yurchyshyn, V., Cao, W., Abramenko, V., Andic, A., Ahn, K., & Chae, J. Highest Resolution Observations of the Quietest Sun. 2010, The Astrophysical Journal Letters, 714(1), L31-L35.

Goodman, J. W. Statistical Optics. Wiley-Interscience (New York). 1985.

Green, P.J. Bayesian reconstructions from emission tomography data using a modified EM algorithm, 1990, IEEE Transactions on Medical Imaging, 9(1), 84-93

Grossman, A. & Morlet, J. Decomposition of Hardy functions into square integrable Wavelets of constant shape. 1984, SIAM Journal of Applied Mathematics, 15, 723-736.

Guyon, O. Limits of adaptive optics for high-contrast imaging. 2005, The Astrophysical Journal, 629, 592-614.

Guyon, O., Sanders, D. & Stockton, A. Near-infrared adaptive optics imaging of QSO host galaxies. 2006, The Astrophysical Journal Supplement Series, 166(1), 89-127.

Guyon, O. High-performance curvature wavefront sensing for extreme AO. 2007, Proceedings of the SPIE, 6691, 66910G-9.

Guyon, O., Pluzhnik, E., Martinache, F., Totems, J.,Tanaka, S., Matsuo, T., Blain, C. & Belikov, R. High-Contrast Imaging and Wavefront Control with a PIAA Coronagraph: Laboratory System Validation. 2010, Publications of the Astronomical Society of the Pacific, 122, 71-84.

Guo, K. & Labate, D. Optimally sparse multidimensional representation using shearlets. 2007, SIAM Journal on Mathematical Analysis, 39, 298-318.

Harayama, Y., Eisenhauer, F. & Martins, F. The initial mass function of the massive star-forming region NGC 3603 from near-infrared adaptive optics observations. 2008, The Astrophysical Journal, 675(2), 1319-1342.

Hardy, J.W., Lefebvre, J.E. & Koliopoilos, C.L. Real-time atmospheric compensation. 1977, Journal of the Optical Society of America, 67, 360-369.

Hardy, J.W. Active optics - A new technology for the control of light. 1978, IEEE Proceedings, 66, 651-697.

Hardy, J.W. Adaptive optics for astronomical telescopes. Oxford University Press. Inc. (New York). 1998.

Häring, N. & Rix, H.W. On the black hole mass-bulge mass relation. 2004, The Astrophysical Journal, 604(2), L89-L92.

Hartkopf, W.I., Mason, B.D., McAlister, H.A., Roberts, L.C., Turner, N.H., ten Brummelaar, T.A., Prieto, C.M., Ling, J.F & Franz, O.G. ICCD speckle observations of binary stars. XXIII. Measurements during 1982-1997 from six telescopes, whit 14 new orbits. 2000, Astronomical Journal, 119(6), 3084-3111.

Hartung, M., Herbst, T.M., Close, L.M., Lenzen, R., Brandner, W., Marco, O. & Lidman, C. A new VLT surface map of Titan at 1.575 microns. 2004, Astronomy and Astrophysics, 421, L17-L20.

Hayward, T.L., Brandl, B., Pirger, B., Blacken, C., Gull, G.E., Schoenwald, J. & Houck, J.R. PHARO: A near-infrared camera for the Palomar adaptive optics system. 2001, Publications of the Astronomical Society of the Pacific, 113, 105-118.

Hestroffer, D., Marchis, F., Fusco, T. & Berthier, J. Adaptive optics observations of asteroid (216) Kleopatra, 2002, Astronomy & Astrophysics, 394, 339-343.

Hinkley, S., Oppenheimer, B. R., Soummer, R., Sivaramakrishnan, A., Roberts, L. C., Kuhn, J., Makidon, R. B., Perrin, M.D., Lloyd, J. P., Kratter, K. & Brenner, D. Temporal Evolution of Coronagraphic Dynamic Range and Constraints on Companions to Vega, 2007, The Astrophysical Journal, 654(1), 633-640.

Hirtzig, M., Coustenis, A., Gendron, E., Drossart, P., Negrão, A., Combes, M., Lai, O., Rannou, P., Lebonnois, S. & Luz, D. Monitoring atmospheric phenomena on Titan. 2006, Astronomy and Astrophysics, 456(2), 761-774.

Högbom, J.A. Aperture synthesis with a non-regular distribution of interferometer baselines. 1974, Astronomy and Astrophysics Supplement Series, 15, 417-426.

Holschneider, M., Kronland-Martinet, R., Morlet, J. & Tchamitchian, P. A real-time algorithm for signal analysis with the help of the wavelet transform, in Wavelets: Time-Frequency methods and Phase-Phase (286-297). Springer (New York). 1989.

Hom, E.F.Y., Marchis, F., Lee, T.K., Haase, S., Agard, D.A. & Sedat, J.W. AIDA: an adaptive image deconvolution algorithm with application to multi-frame and three-dimensional data. 2007, Journal of the Optical Society of America A, 24(6), 1580-1600.

Horch, E.P., Dinescu, D.I., Girard, T.N., van Altena, W.F., López, C.E. & Franz, O.G. Speckle interferometry of southern double stars. I. First results of the Yale-San Juan speckle interferometry program. 1996, Astronomical Journal, 111(4), 1681-1688.

Horch, E.P., Ninkov, A.Z. & Franz, O.G. CCD speckle observations of binary stars from the southern hemisphere. III. Differential photometry. 2001, Astronomical Journal, 121, 1583-1596.

Horch, E.P., Veillette, D.R., Baena Gallé, R., Shah, S.C., O'Rielly, G.V. & van Altena, W.F. Observations of binary stars with the differential speckle survey instrument. I. Instrument description and first results. 2009, Astronomical Journal, 137, 5057-5067.

Horch, E.P., Bahi, L.A.P., Gaulin, J.R., Howell, S.B., Sherry, W.H., Baena Gallé, R. & van Altena, W.F. Speckle observations of binary stars with the WIYN telescope. VII. Measures during 2008-2009. 2012, Astronomical Journal, 143, 1-10.

Huebner, C. Software-based turbulence mitigation of short exposure image data with motion detection and background segmentation. 2011, Proceedings of the SPIE, 8178, 81780K.

van Iersel, M. & van Eijk, A.M.J. Estimating turbulence in images. 2010, Proceedings of the SPIE, 7814, 78140Q.

Jaynes, J. T. On the Rationale of Maximum Entropy Methods, 1982, IEEE Proceedings, 70, 939-952.

Jefferies, S. & Christou, J. C. Restoration of astronomical images by iterative blind deconvolution. 1993, Astrophysical Journal, 415, 862-874.

Kaiser, G. A friendly guide to wavelets. Birkhauser (Boston). 1994.

Kasper, M., Apai, D., Janson, M. & Brandner, W. A novel L-band imaging search for giant planets in the Tucana and β Pictoris moving groups. 2007, Astronomy and Astrophysics, 472(1), 321-327.

Kasper, M., Beuzit, J.-L., Verinaud, C., Yaitskova, N., Baudoz, P., Boccaletti, A., et al. EPICS: the exoplanet imager for the E-ELT. 2008, Proceedings of the SPIE, 7015, 70151S.

Kasper, M., Beuzit, J-L., Verinaud, C., Yaitskova, N., Baudoz, P., et al. EPICS, direct imaging of Exo-Planets with the E-ELT. 2010, Proceedings of the SPIE, 7735, 77352E.

Knox, K.T. & Thompson, B.J. Recovery of images from atmospherically degraded short-exposure photographs, 1974, Astrophysical Journal, 193, L45-L48.

Kolmogorov, A.N. The local structure of turbulence in incompressible viscous fluid for very large Reynolds' numbers. 1941a, Doklady Akademiia Nauk. SSSR, 30, 301-305. *(Russian)*, translated into English by Kolmogorov, A.N. The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. 1991. Proceedings of the Royal Society of London, Series A: Mathematical and Physical Sciences, 434, 913.

Kolmogorov, A.N. Dissipation of energy in isotropic turbulence. 1941b, Doklady Akademiia Nauk. SSSR, 32, 16-18. *(Russian)*, translated into English by Kolmogorov, A.N. The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. 1991. Proceedings of the Royal Society of London, Series A: Mathematical and Physical Sciences, 434, 1517.

Konopacky, Q.M., Ghez, A.M., Barman, T.S., Rice, E.L., Bailey, J.I., White, R.J., McLean, I.S. & Duchêne, G. High-precision dynamical masses of very low mass binaries. 2010, The Astrophysical Journal, 711(2), 1087-1122.

Korkiakoski, V. & Verinaud, C. Simulations of the extreme adaptive optics system for EPICS. 2010, Proceedings of the SPIE, 7736, 773643.

Kraus, A.L., Ireland, M.J., Hillenbrand, L.A. & Martinache, F. The role of multiplicity in disk evolution and planet formation. 2012, The Astrophysical Journal, 745(1), article id. 19 (11pp).

Kraus, A.L. & Ireland, M.J. LkCa 15: A Young Exoplanet Caught at Formation? 2012, The Astrophysical Journal, 745(1), article id. 5 (12pp).

Labeyrie, A. Attainment of diffraction limited resolution in large telescopes by Fourier analysing speckle patterns in star images. 1970, Astronomy and Astrophysics, 6, 85-87.

Labeyrie, A. Images of exo-planets obtainable from dark speckles in adaptive telescopes. 1995, Astronomy and Astrophysics, 298, 544-548.

Lafrenière, D., Doyon, R., Marois, C., Nadeau, D., Oppenheimer, B.R. et al. The Gemini deep planet survey. 2007, The Astrophysical Journal, 670(2), 1367-1390.

Lafrenière, D., Marois, C., Doyon, R., Nadeau, D. & Artigau, É. A New Algorithm for Point-Spread Function Subtraction in High-Contrast Imaging: A Demonstration with Angular Differential Imaging. 2007, The Astrophysical Journal, 660(1), 770-780.

Lambert, P., Pires, S., Ballot, J., García, R., Starck, J.L. & Turck-Chièze, S. Curvelet analysis of asteroseismic data. I. Method description and application to simulated sun-like stars. 2006, Astronomy and Astrophysics, 454, 1021-1027.

Lane, R.G. Blind deconvolution of speckle images, 1992, Journal of the Optical Society of America A, 9(9), 1508-1514.

Law, D.R., Steidel, C.C., Erb, D.K., Larkin, J.E., Pettini, M., Shapley, A.E. & Wright, S.A. The Kiloparsec-scale kinematics of high-redshift star-forming galaxies. 2009, The Astrophysical Journal, 697(2), 2057-2082.

Léna, P., Lebrun, F. & Mignard, F. Observational Astrophysics. Springer-Verlag (Berlin-Heidelberg). 1998.

Llacer, J. & Núñez, J. Iterative maximum likelihood estimator and Bayesian algorithms for image reconstruction in astronomy. The restoration of HST images and spectra. 1990, Proceeding of STScI (Baltimore), 62-70.

Llacer, J., Veklerov, E., Coakley, K.J., Hoffman, E.J. & Núñez, J. Statistical analysis of Maximum Likelihood Estimator Images of human brain FDG PET studies, 1993, IEEE Transactions on Medical Imaging, 12(4), 215-231.

Lohmann, A.W., Weigelt, G. & Wirnitzer, B. Speckle masking in astronomy: triple correlation theory and applications. 1983, Applied Optics, 22(24), 4028-4037.

Louthain, J.A. & Schmidt, J.D. Anisoplanatism in airborne laser communication. 2008, Optics Express 16(14), 10769-10785.

Lu, Y. & Do, M.N. Multidimensional directional filter banks and surfacelets. 2007, IEEE Transactions on Image Processing, 16(4), 918-931.

Lu, J.R.; Ghez, A.M., Hornstein, S.D., Morris, M.R., Becklin, E.E. & Matthews, K. A disk of young stars at the galactic center as determined by individual stellar orbits. 2009, The Astrophysical Journal, 690(2), 1463-1487.

Lucy, L.B. An iterative technique for the rectification of observed distributions. 1974, Astronomical Journal, 79(6), 745-754.

Lucy, L.B. The restoration of HST images and spectra II. STScI (Boston). 1994.

Lühe, O. von der. Estimating Fried's parameter from a time series of an arbitrary resolved object imaged through atmospheric turbulence, 1984, Journal of the Optical Society of America A, 1, 510-519.

Ma, J. & Plonka, G. The Curvelet transform: A review of recent applications. 2010, IEEE Signal Processing Magazine, 27(2), 118-133.

Mahan, A.I. Astronomical refraction. Some history and theories. 1962. Applied Optics, 1(4), 497-511.

Mallat, S. & Peyré, G. A review of bandlet methods for geometrical image representation. 2007. Numerical Algorithms, 44(3), 205-234.

Mallat, S. Geometrical grouplets. 2009, Applied and Computational Harmonic Analysis, 26(2), 161-180.

Mandel, L. Fluctuations of Photon Beams: The Distribution of the Photo-Electrons, 1959, Proceedings of the Physical Society, 74(3), 233-243.

Marchi, G. & Scheifling, C. Adaptive optics solutions for turbulence mitigation in different scenarios. 2011, Proceedings of the SPIE 8161, 816104.

Marchis, F., Urata, R., de Pater, I., Gibbard, S., Hammel, H.B. & Berthier, J. Neptunian Satellites observed with Keck AO system. 2004, Bulletin of the American Astronomical Society, 36, 860.

Marchis, F., Descamps, P, Hestroffer, D. & Berthier, J. Discovery of the triple asteroidal system 87 Sylvia. 2005, Nature, 436(7052), 822-824.

Maréchal, A. & Françon, M. Diffraction: Structures des images, influence de la coherence de la lumiere. Editions revue d'optique theoretique et instrumentale (Paris). 1960.

Marois, C., Lafranière, D., Doyon, R., Macintosh, B. & Nadeau, D. Angular Differential Imaging: A Powerful High-Contrast Imaging Technique. 2006, The Astrophysical Journal, 641(1), 556-564.

Marois, C., Macintosh, B., Soummer, R., Poyneer, L. & Bauman, B. An end-to-end polychromatic Fresnel propagation model of GPI. 2008, Proceedings of the SPIE, 7015, 70151T.

Marois, C., Zuckerman, B., Konopacky, Q.M., Macintosh, B, & Barman, T. Images of a fourth planet orbiting HR 8799. 2010, Nature, 468, 1080-1083.

Martinez, P., Dorrer, C., Aller Carpentier, E., Kasper, M., Boccaletti, A., Dohlen, K. & Yaitskova, N. Design, analysis, and testing of a microdot apodizer for the Apodized Pupil Lyot Coronagraph. 2009, Astronomy & Astrophysics, 495(1), 363-370.

Mason, B.D., Hartkopf, W.I., Wycoff, G.L., Holdenried, E.R., Platais, I., Rafferty, T.J., Hall, D.M., Hennessy, G.S., Urban, S.E. & Pohl, B.L. Speckle Interferometry at the US Naval Observatory. VII. 2001, Astronomical Journal, 122, 1586-1601.

Matson, C.L. & Haji, A. An Algorithm-independent Analysis of the Quality of Images Produced Using Multi-frame Blind Deconvolution Algorithms, 2007, Proceedings of the AMOS Technologies Conference, E80.

Matson, C.L., Borelli, K., Jefferies, S., Beckner, C.C., Hege, E.K. & Lloyd-Hart, M. Fast and optimal multiframe blind deconvolution algorithm for high-resolution ground-based imaging of space objects, 2009, Applied Optics, 48(1), A75.

Max, C., Canalizo, G. & de Vries, W. Locating the two black holes in NGC 6240. 2007, Science, 316(5833), 1877-1880.

McCabe, C., Duchêne, G., Pinte, C., Stapelfeldt, K.R., Ghez, A.M. & Ménard, F. Spatially resolving the HK Tau B edge-on disk from 1.2 to 4.7μm: a unique scattered light disk. 2012, The Astrophysical Journal, 727(2), article id. 90 (15pp).

McLean, I.S. Electronic imaging in astronomy. Wiley and Sons Ltd. & Praxis Publishing Ltd. (Chichester, England). 1997.

Melbourne, J., Williams, B., Dalcanton, J., Ammons, S.M., Max, C., Koo, D.C., Girardi, L, & Dolphin, A. The asymptotic giant branch and the tip of the red giant branch as probes of star formation history: the nearby dwarf irregular galaxy KKH 98. 2010, The Astrophysical Journal, 712(1), 469-483.

Meng, J., Aitken, G.J.M., Hege, E.K. & Morgan, J.S. Triple-correlation subplane reconstruction of photon-address stellar images. 1990, Journal of the Optical Society of America A, 7(7), 1243-1250.

Merline, W.J., Close, L.M., Dumas, C., Chapman, C.R., Roddier, F., Menard, F., Slater, D.C., Duvert, G., Shelton, C. & Morgan, T. Discovery of a moon orbiting the asteroid 45 Eugenia. 1999, Nature, 401, 565-568.

Mesa, D., Gratton, R., Berton, A., Antichi, J., Verinaud, C., Boccaletti, A., Kasper, M., et al. Simulation of planet detection with the SPHERE integral field spectrograph. 2011, Astronomy & Astrophysics, 529, A131.

Meyer, Y. Wavelets. Springer Verlag (Berlin). 1989.

Meyer, Y. Wavelets: Algorithms & Applications. Society for Industrial and Applied Mathematics (Philadelphia). 1993.

Molina, R., Núñez, J., Cortijo, F. & Mateos, J. Image restoration in astronomy. A Bayesian perspective. 2001, IEEE Signal Processing Magazine, 18, 11-29.

Mugnier, l.M., Fusco, T. & Conan, J.M. MISTRAL: a myopic edge-preserving image restoration method, with application to astronomical adaptive-optics-corrected long-exposure images. 2004, Journal of the Optical Society of America A, 21(10), 1841-1854.

Müller Sánchez, F. Davies, R.I., Genzel, R., Tacconi, L.J., Eisenhauer, F., Hicks, E.K.S., Friedrich, S. & Sternberg, A. Molecular gas streamers feeding and obscuring the active nucleus of NGC 1068. 2009, The Astrophysical Journal, 691(1), 749-759.

Murtagh, F., Starck, J.L. & Bijaoui, A. Image restoration with noise suppression using a multiresolution support. 1995, Astronomy & Astrophysics Supplement, 112, 179-189.

Noll, R.J. Zernicke polynomials and atmospheric turbulence. 1976, Journal of the Optical Society of America, 66, 207-211.

Nowak, N., Thomas, J., Erwin, P., Saglia, R.P., Bender, R. & Davies, R.I. Do black hole masses scale with classical bulge luminosities only? The case of the two composite pseudo-bulge galaxies NGC 3368 and NGC 3489. 2010, Monthly Notices of the Royal Astronomical Society, 403(2), 646-672.

Núñez, J. & Llacer, J. A general bayesian image reconstruction algorithm with entropy prior. Preliminary application to HST data. 1993, Publications of the Astronomical Society of the Pacific, 105, 1192-1208.

Núñez, J. & Llacer J. Bayesian image reconstruction with space-variant noise suppression, 1998, Astronomy & Astrophysics Supplement, 131, 167-180.

Obukhov, A.M. On the energy distribution in the spectrum of a turbulent flow. 1941, Doklady Akademiia Nauk. SSSR, 32, 19-21.

Olsen, K.A.G., Blum, R.D., Stephens, A.W., Davidge, T.J., Massey, P.; Strom, S.E. & Rigaut, F. The star formation histories of the bulge and disk of M31 from resolved stars in the near-infrared. 2006, The Astronomical Journal, 132(1), 271-289.

Otazu, X. Algunes aplicacions de les wavelets al procés de dades en Astronomia i Teledetecció. phD thesis. University of Barcelona (Barcelona). 2001.
(http://www.tesisenred.net/handle/10803/732)

Otazu, X., Ribó, M., Peracaula, M., Paredes, J.M. & Núñez, J. Detection of superimposed periodic signals using wavelets. 2002, Monthly Notices of the Royal Astronomical Society, 333(2), 365-372.

Otazu, X., Ribó, M., Paredes, J.M., Peracaula, M. & Núñez, J. Multiresolution approach for period determination on unevenly sampled data. 2004, Monthly Notices of the Royal Astronomical Society, 351(1), 215-219.

Otazu, X., González Audícana M., Fors, O. & Núñez, J. Introduction of sensor spectral response into image fusion methods. Application to wavelet-based methods. 2005, IEEE Transactions on GeoScience and Remote Sensing, 43(10), 2376-2385.

Pantin, E., Starck, J.L. & Murtagh F. Deconvolution and Blind Deconvolution in Astronomy (in Blind image deconvolution: theory and applications). Taylor & Francis Group, LLC. CRC Press, 277-317. 2007.

de Pater, I., Marchis, F., Macintosh, B.A., Roe, H.G., Le Mignant, D., Graham, J.R. & Davies, A.G. Keck AO observations of Io in and out of eclipse. 2004, Icarus, 169(1), 250-263.

de Pater, I., Wong, M.H., Marcus, P., Luszcz-Cook, S.; Ádámkovics, M., Conrad, A., Asay-Davis, X. & Go, C. Persistent rings in and around Jupiters anticyclones - Observations and theory. 2010, Icarus, 210(2), 742-762.

Peng, C., Ho, L., Impey, C. & Rix, H.W. Detailed decomposition of galaxy images. II. Beyond axisymmetric models. 2010, Astronomical Journal, 139(6), 2097-2129.

Pollack, L., Max, C. & Schneider, G. Circumnuclear star clusters in the galaxy merger NGC 6240, observed with Keck adaptive optics and the Hubble space telescope. 2007, The Astrophysical Journal, 660(1), 288-300.

Polygiannakis, J., Preka-Papadema, P. & Moussas, X. On signal-noise decomposition of time-series using the continuous wavelet transform: application to sunspot index. 2003, Monthly Notices of the Royal Astronomical Society, 343(3), 725-734.

Poyneer, L.A., Gavel, D.T. & Brase, J.M. Fast wave-front reconstruction in large adaptive optics systems with use of the Fourier transform. 2002, Journal of the Optical Society of America A, 19(10), 2100-2111.

Prato, M., Cavicchioli, R., Zanni, L., Boccacci, P. & Bertero, M. Efficient deconvolution methods for astronomical imaging: algorithms and IDL-GPU codes. 2012, Astronomy & Astrophysics 539, A133.

Pratt, W.K. Digital image processing (2nd ed.). Wiley (New York). 2007.

Quirrenbach, A. The effects of atmospheric turbulence on astronomical observations, in Extrasolar planets. Springer-Verlag (Berlin-Heidelberg). 2006.
(http://cfao.ucolick.org/aosummer/book/)

Racine, R., Walker, G.A.H., Nadeau, D., Doyon, R. & Marois, C. Speckle Noise and the Detection of Faint Companions. 1999, The Publications of the Astronomical Society of the Pacific 111(759), 587-594.

Richardson, W.H. Bayesian-based iterative method of image restoration. 1972, Journal of the Optical Society of America, 62(1), 55-59.

Rigaut, F., Rousset, G., Kern, P., Fontanella, J.C., Gaffard, J.P., Merkle, F. & Léna, P. Adaptive optics on a 3.6-m telescope - Results and performance. 1991, Astronomy and Astrophysics, 250(1), 280-290.

Rigaut, F., Ellerbroek, B.L. & Northcott, M.J. Comparison of curvature-based and Shack Hartmann-based adaptive optics for the Gemini telescope. 1998, Applied Optics, 36(13), 2856-2868.

Rigaut, F., Salmon, D., Arsenault, R., Thomas, J., Lai, O. et al. Performance of the Canada-France-Hawaii Telescope Adaptive Optics Bonnette. 1998, The Publications of the Astronomical Society of the Pacific, 110(744), 152-164.

Rimmele, T.R. Solar adaptive optics. 2000, Proceedings of the SPIE, 4007, 218-231.

Roberts, D.H., Lehár, J. & Dreher, J.W. Time Series Analysis with Clean I. Derivation of a Spectrum. 1987, The Astronomical Journal, 93(4), 968-989.

Roberts, L. C., Turner, N. H., Bradford, L. W., ten Brummelaar, T. A., Oppenheimer, B. R. et al. Adaptive Optics Photometry and Astrometry of Binary Stars, 2005, The Astronomical Journal, 130(5), 2262-2271.

Roberts, L. C., Turner, N. H. & ten Brummelaar, T. A. Adaptive Optics Photometry and Astrometry of Binary Stars. II. A Multiplicity Survey of B Stars, 2007, The Astronomical Journal, 133(2), 545-552.

Roddier, F. The effects of atmospheric turbulence in optical astronomy. 1981, Progress in optics, 19, 281-376.

Roddier, C., Roddier, F., Northcott, M.J., Graves, J.E. & Jim, K. Adaptive optics imaging of GG Tauri: optical detection of the circumbinary ring. 1996, Astrophysical Journal, 463, 326-335.

Roggemann, M. C., Tyler, D. W. & Bilmont, M. F. Linear reconstruction of compensated images - Theory and experimental results, 1992, Applied Optics, 31(35), 7429-7441.

Rosario, D.J., McGurk, R.C., Max, C.E., Shields, G.A., Smith, K.L. & Ammons, S.M. Adaptive optics imaging of quasi-stellar objects with double-peaked narrow lines: are they dual active galactic nuclei? 2011, The Astrophysical Journal, 739(1), article id. 44 (8pp).

Rosensteiner, M. Cumulative Reconstructor: fast wavefront reconstruction algorithm for Extremely Large Telescopes. 2011, Journal of the Optical Society of America A, 28(10), 2132-2138.

Rousset, G., Fontanella, J.C., Kern, P., Gigan, P. & Rigaut, F. First diffraction-limited astronomical images with adaptive optics. 1990, Astronomy and Astrophysics, 230, L29-L32.

Sarazin, M. Update of environmental conditions at ESO Observatories. ESO VLT report VLT.TRE.ESO.17400. 1992.

Sasiela, R.J. & Shelton, J.D. Transverse spectral filtering and Mellin transform techniques applied to the effect of outer scale on tilt and tilt anisoplanatism. 1993, Journal of the Optical Society of America A, 10, 646-660.

Schnorr Müller, A., Storchi-Bergmann, T., Riffel, R.A., Ferrari, F., Steiner, J.E., Axon, D.J. & Robinson, A. Gas streaming motions towards the nucleus of M81. 2011, Monthly Notices of the Royal Astronomical Society, 413(1), 149-161.

Schödel, R. Accurate photometry with adaptive optics in the presence of anisoplanatic effects with a sparsely sampled PSF. The Galactic center as an example of a challenging target for accurate AO photometry. 2010, Astronomy and Astrophysics, 509, A58.

Schulz, T.J. Multiframe blind deconvolution of astronomical images. 1993, Journal of the Optical Society of America A, 10(5), 1064-1073.

Selsis, F., Kasting, J.F., Levrard, B., Paillet, J., Ribas, I. & Delfosse, X. Habitable planets around the star Gliese 581? 2007, Astronomy & Astrophysics, 476(3), 1373-1387.

Shensa, M. Discrete wavelet transform: Wedding the à trous and Mallat algorithms. 1992, IEEE Transaction on Signal Processing, 40, 2464-2482.

Shellan, J.B. Statistical properties of the Strehl ratio as a function of pupil diameter and level of adaptive optics correction following atmospheric propagation. 2004, Journal of the Optical Society of America A, 21(8), 1445-1451.

Siegler, N., Close, L.M., Mamajek, E.E. & Freed, M. An adaptive optics survey of M6.0-M7.5 stars: discovery of three very low mass binary systems including two probable Hyades members. 2003, The Astrophysical Journal, 598(2), 1265-1276.

Soummer, R. Apodized Pupil Lyot Coronagraphs for Arbitrary Telescope Apertures. 2005, The Astrophysical Journal, 618(2), L161-L164.

Sparks, W.B. & Ford, H.C. Imaging Spectroscopy for Extrasolar Planet Detection. 2002, The Astrophysical Journal, 578(1), 543-564.

Sromovsky, L.A. & Fry, P.M. Keck 2 AO Observations of Uranus in 2004. 2004, Bulletin of the American Astronomical Society, 36, 1073.

Starck, J. L., Bijaoui, A., Lopez, B., Perrier C. Image reconstruction by the wavelet transform applied to aperture synthesis, 1994, Astronomy and Astrophysics, 283(1), 349-360.

Starck, J. L. & Murtagh, F. Image restoration with noise supression using the wavelet transform. 1994, Astronomy and Astrophysics, 288, 342-348.

Starck, J. L., Murtagh, F. & Bijaoui, A. Multiresolution and astronomical image processing. 1995, ASP Conference Series, Astronomical Data Analysis Software and Systems IV (San Francisco: ASP), 77, 279-288.

Starck, J. L., Murtagh, F., Bijaoui, A. Image processing and data analysis. The multiscale approach. Cambridge University Press (New York). 1998.

Starck, J. L., Murtagh, F., Querre, P. & Bonnarel, F. Entropy and astronomical data analysis: Perspectives from multiresolution analysis. 2001, Astronomy & Astrophysics, 368, 730-746.

Starck, J. L., Pantin, E. & Murtagh, F. Deconvolution in astronomy: a review. 2002, Publications of the Astronomical Society of the Pacific, 114, 1051-1069.

Starck, J. L., Candès, E. & Donoho, D. Astronomical image representation by the Curvelet Transform. 2003, Astronomy and Astrophysics, 398, 785-800.

Starck, J. L., Aghanim, N. & Forni, O. Detecting cosmological non-Gaussian signatures by multiscale methods. 2004, Astronomy and Astrophysics, 416, 9-17.

Starck, J. L. & Murtagh, F. Astronomical image and data analysis. Springer (New York, 2nd ed). 2006.

Starck, J. L., Murtagh, F. & Fadili, J. L. Sparse image and signal processing. Wavelets, Curvelets, morphological diversity. Cambridge University Press (New York). 2010.

Stellingwerf, R.F. Period determination using phase dispersion minimization. 1978, Astrophysical Journal, 224, 953-960.

Stetson, P. DAOPHOT: A computer program for crowded-field stellar photometry. 1987, Publications of the Astronomical Society of the Pacific, 99, 191-222.

Stolte, A., Ghez, A.M., Morris, M., Lu, J.R., Brandner, W. & Matthews, K. The proper motion of the Arches cluster with Keck laser-guide star adaptive optics. 2008, The Astrophysical Journal, 675(2), 1278-1292.

Storrs, A. & Eney, B. Synoptic Observations of 134340 Pluto from the Hubble Space Telescope. 2010, Bulletin of the American Astronomical Society, 42, 456.

Strehl, K. Über Luftschlieren und Zonenfehler. 1902, Zeitschrift für Instrumentenkunde, 22, 213-217.

Szatmáry, K., Vinkó, J. & Gál J. Application of wavelet analysis in variable star research. I. Properties of the wavelet map of simulated variable star light curves. 1994, Astronomy & Astrophysics Supplements, 108, 377-394.

Talbot, H.F. Facts relating to optical science. 1836, Philosophical magazine and journal of science, 9(56), 401-407.

Tatarski, V.I. Wave propagation in a turbulent medium. McGraw-Hill Book Company (New York). 1961.

Taylor, J. R. An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements. University Science Books, 2nd ed. (Sausalito, California). 1996.

Tessier, E. Image quality with current adaptive optics instruments, Astronomy and Astrophysics Supplement Series. 1997, 125, 581-593.

Thatte, N., Tecza, M., Eisenhauer, F., Mengel, S., Krabber, A., et al. SINFONI: a near-infrared AO-assisted integral field spectrometer for the VLT. 1998, Proceedings of the SPIE, vol 3353, 704-715.

Thatte, N., Abuter, R., Tecza, M., Nielsen, E.L., Clarke, F.J. & Close, L.M. Very high contrast integral field spectroscopy of AB Doradus C: 9-mag contrast at 0.2arcsec without a coronagraph using spectral deconvolution. 2007, Monthly Notices of the Royal Astronomical Society, 378(4), 1229-1236.

Thatte, N., Tecza, M., Clarke, F., Davies, R.L., Remillieux, A., et al. HARMONI: a single-field wide-band integral-field spectrograph for the European ELT. 2010, Proceedings of the SPIE, vol 7735, 77352I (11pp).

Thiébaut, E. & Conan, J.M. Strict a priori constraints for maximum-likelihood blind deconvolution. 1995, Journal of the Optical Society of America A, 12(3), 485-492.

Thiébaut, E. Introduction to image reconstruction and inverse problems. 2002, Optics in astrophysics: Proceedings of the NATO Advanced Study Institute on Optics in Astrophysics

(Cargese, France). NATO science series II: Mathematics, physics and chemistry, 198, 397-423.

Thiébaut, É & Tallon, M. Fast minimum variance wavefront reconstruction for extremely large telescopes. 2010, Journal of the Optical Society of America A, 27(5), 1046-1059.

Tikhonov, A.N. Solution of incorrectly formulated problems and the regularization method. 1963, Soviet Mathematika Doklady 4, 501-504.

Tikhonov, A.N., Goncharski, A.V., Stepanov, V.V. & Kochiko, I.V. Ill-posed image processing problems. 1987, Soviet Physics Doklady, 32, 456-458.

Trauger, J.T., Ballester, G.E., Burrows, C.J. et al. The on-orbit performance of WFPC2. 1994, Astrophysical Journal, 435(1), L3-L6.

Trauger, J.T. & Traub, W.A. A laboratory demonstration of the capability to image an Earth-like extrasolar planet. 2007, Nature, 446(7137), 771-773.

Trippe, S., Davies, R., Eisenhauer, F., Förster Schreiber, N.M., Fritz, T.K. & Genzel, R. High-precision astrometry with MICADO at the European extremely large telescope. 2010, Monthly Notices of the Royal Astronomical Society, 402(2), 1126-1140.

Troy, M., Dekany, R.G., Brack, G.L. et al. Palomar adaptive optics project: status and performance. 2000, Proceedings of the SPIE, 4007, 31-40.

Turner, N. H., ten Brummelaar, T. A., Roberts, L. C., Mason, B. D., Hartkopf, W. I. & Gies, D. R. Adaptive Optics Photometry and Astrometry of Binary Stars. III. a Faint Companion Search of O-Star Systems, 2008, The Astronomical Journal, 136(2), 554-565.

Tyler, G. A. Assessment of the statistics of the Strehl ratio: predictions of central limit theorem analysis, 2006, Journal of the Optical Society of America A, 23(11), 2834-2844.

Tyson, R.K. Principles of adaptive optics. Academic Press (Boston). 1998.

Vdovin, G., Loktev, M. & Soloviev, O. Adaptive Optics: Turbulent surveillance - or how to see a Kalashnikov from a safe distance. 2012, Laser Focus World 48(3), 31-38.

Velisavljevic, V., Beferull-Lozano, B., Vetterli, M. & Dragotti, P. Directionlets: Anisotropic multidirectional representation with separable filtering. 2006, IEEE Transactions on Image Processing, 15(7), 1916-1933.

Verinaud, C. & Cassaing, F. Piston control with adaptive optics in stellar interferometry. Application to the GI2T interferometer and bimorph mirrors. 2001, Astronomy & Astrophysics, 365, 314-323.

Verinaud, C., Kasper, M., Beuzit, L-C., Gratton, R.G., Mesa, D., et al. System study of EPICS, the exoplanet imager for the E-ELT. 2010, Proceeding of the SPIE, 7736, 77361N.

Vetterli, M. & Kovacevic, J. Wavelets and subband coding. Prentice Hall (Englewood Cliffs). 1995.

Vigan, A., Moutou, C., Langlois, M., Allard, F., Boccaletti, A., Carbillet, M., Mouillet, D., Smith, I. Photometric characterization of exoplanets using angular and spectral differential imaging, 2010, Monthly Notices of the Royal Astronomical Society, 407(1), 71-82.

Wallner, E.P. Optimal wave-front correction using slope measurements. 1983, Journal of the Optical Society of America, 73(12), 1771-1773.

Wang, Z., Bovik, A.C., Sheikh, H.R. & Simoncelli, E.P. Image quality assessment: from error visibility to structural similarity. 2004, IEEE Transactions on Image Processing, 13(4), 1-14.

Wang, Z. & Bovik, A.C. Modern image quality assessment. Morgan & Claypool. 2006.

Welford, W. Aberrations of the symmetrical optical system. Academic Press. (London). 1974.

Willett, R.M. & Nowak, R.D. Platelets: A multiscale approach for recovering edges and surfaces in photon-limited medical imaging. 2003, IEEE Transactions on Medical Imaging, 22(3), 332-350.

Wilson, R.N. Reflecting Telescope Optics I. Springer-Verlag (Berlin-Heidelberg). 1996.

Wilson, R.N. Reflecting Telescope Optics II. Springer-Verlag (Berlin-Heidelberg). 1999.

Wilson, R.W. & Jenkins, C.R. Adaptive optics for astronomy: theoretical performance and limitations, 1996, Monthly Notices of the Royal Astronomical Society, 278(1), 39-61.

Woods, J.W. & O'Neill S.D. Subband coding of images. 1986, IEEE Transactions on Acoustics, Speech and Signal Processing, ASSP-34(5), 1278-1288.

Worden, S.P., Stein, M.K., Schmidt, G.D. & Angel, J.R.P. The angular diameter of Vesta from speckle interferometry. 1977, Icarus, 32, 450-457.

Wright, S.A., Larkin, J.E., Law, D.R., Steidel, C.C., Shapley, A.E. & Erb, D.K. Dynamics of galactic disks and mergers at $z \sim 1.6$: spatially resolved spectroscopy with Keck laser guide star adaptive optics. 2009, The Astrophysical Journal, 699(1), 421-440.

Yelda, S., Lu, J.R., Ghez, A.M., Clarkson, W., Anderson, J., Do, T. & Matthews, K. Improving galactic center astrometry by reducing the effects of geometric distortion. 2010, The Astrophysical Journal, 725(1), 331-352.

Ygouf, M., Mugnier, L.M., Mouillet, D., Fusco, T. & Beuzit, J.-L. Simultaneous exoplanet detection and instrument aberration retrieval in multispectral coronagraphic imaging. 2013, Astronomy & Astrophysics, 551, A138 (14 pp).

Zamek, S. & Yitzhaky, Y. Turbulence strength estimation from an arbitrary set of atmospherically degraded images. 2006, Journal of the Optical Society of America A, 23, 3106-3113.

Zellner, N.E.B., Gibbard, S., de Pater, I., Marchis, F. & Gaffey, M.J. Near-IR imaging of Asteroid 4 Vesta. 2005, Icarus, 177(1), 190-195.