NNT : 2016SACLS343

Thèse de doctorat de l'Université Paris-Saclay préparée à l'Université Paris-Sud

Ecole doctorale n°127
Astronomie et Astrophysique d'Île-de-France
Spécialité de doctorat : astronomie et astrophysique

par

M. NGOLÈ MBOULA Fred Maurice

Méthodes et algorithmes avancés pour l'imagerie astronomique de haute précision

Thèse présentée et soutenue à Saclay, le 18 Octobre 2016.

Composition du Jury :

M.	Hervé DOLE	Professeur des Universités Université Paris-Sud	(Président du jury)
Mme	Laure BLANC-FÉRAUD	Directeur de recherche Université de Nice Sophia-Antipolis	(Rapporteur)
M.	Gabriel PEYRÉ	Directeur de recherche École normale supérieure	(Rapporteur)
Mme	Maïtine BERGOUNIOUX	Professeur des Universitéés Université d'Orléans	(Examinatrice)
Mme	Julie DELON	Chargée de recherche Université Paris Descartes	(Examinatrice)
M.	Jean-Luc STARCK	Directeur de recherche CEA Saclay	(Directeur de thèse)
M.	Jalal FADILI	Professeur des Universités ENSICAEN	(Invité)

Acknowledgments

Thank you Jean-Luc for your enlightened and motivating guidance!

Thank you dear reader, if you are one of those who made my *PhD journey* enjoyable and fruitful!

Your name is engraved in letters of gold in my mind...

Thanks to the CEA for providing nearly optimal conditions of work and research!

Extended abstract

L'un des challenges majeurs de la cosmologie moderne réside en la nature même de la matière et de l'énergie noire. La matière noire peut être directement tracée à travers son effet gravitationnel sur les formes des galaxies. La mission Euclid de l'Agence Spatiale Européenne fournira précisément des données à cette fin. L'exploitation de telles données requiert une modélisation précise de la Fonction d'Étalement du Point (FEP) de l'instrument d'observation. En première approximation, les images par l'instrument des étoiles isolées donnent une mesure de sa FEP. Toutefois, ces images sont toujours bruitées, potentiellement sous-échantillonnées et la FEP sous-jacente varie dans le champ de l'instrument, mais également dans le temps et en longueur d'onde. L'objectif de cette thèse est précisément de proposer des méthodes d'estimation de FEP, à partir d'images d'étoiles, prenant en compte les différents facteurs mentionnés ci-avant, ce en s'appuyant sur des avancées méthodologiques récentes en traitement du signal.

Dans un premier temps nous avons développé une méthode permettant d'estimer à une résolution suffisante, une FEP donnée, à partir de plusieurs mesures bruitées sous-échantillonnées de la même FEP. Nous avons montré qu'en utilisant des contraintes de parcimonie et de positivité, l'on obtenait une bien meilleure restauration de la forme de la FEP, comparativement avec des méthodes largement utilisées en astronomie et en particulier à faibles rapports signal à bruit.

Toutefois, du fait de la variabilité spatiale et temporelle de la FEP, l'on ne dispose en général pas de plusieurs mesures de la même FEP. Ainsi, dans une deuxième étape, nous avons considéré le problème d'estimer conjointement un ensemble de FEPs différentes à partir d'une seule mesure potentiellement sous-échantillonnée de chacune de ces FEPs. Cela a abouti à une méthode de réduction de dimension et de super-résolution, qui en plus des contraintes de parcimonie et de positivité, exploite la compressibilité globale du champ de FEPs. Une fois de plus, nous avons obtenu un gain significatif en terme de précision sur les formes de PSFs, par rapport à des méthodes existantes.

Comme mentionné plus haut la FEP varie à travers le champ du télescope. L'objectif final étant de réaliser des mesures des formes de galaxies corrigées de l'effet de la FEP, il est nécessaire d'interpoler spatialement le champ de FEPs. En d'autres termes, l'on souhaite pouvoir estimer la FEP à une position arbitraire dans le champ de l'instrument connaissant les FEPs aux positions des étoiles isolées. Cela revient à estimer une fonction qui à tout vecteur (x, y) du plan focal de l'instrument associe un vecteur sur la variété sous-jacente

aux FEPs. Pour cela, nous proposons une méthode d'interpolation qui réalise dans un premier temps une réduction de dimension locale des FEPs connues, dans la métrique de Wasserstein 2. A supposer que cela "déplie" localement la variété sous-jacente, les coordonnées des FEPs inconnues dans la représentation en dimension réduite peuvent être estimées en utilisant des méthodes d'interpolation classiques. Cela permet ensuite de calculer les coordonnées barycentriques des FEPs inconnues dans la métrique de Wasserstein 2. Les FEPs à interpoler sont finalement calculées comme barycentres des FEPs connues dans la métrique de Wasserstein 2. Nous avons obtenu par cette approche des résultats notables en termes de précision sur la forme et les valeurs de pixels des PSFs interpolées. Elle est toutefois coûteuse en termes de calculs.

Cette thèse ouvre plusieurs perspectives dont nous donnons quelques exemples ci-dessous :

— extensions des méthodes proposées pour prendre en compte la dépendance en longueur d'onde de la FEP ;

— couplage de la restauration du champ de FEP au problème de déconvolution des galaxies dans le même champ ;

— intégration davantage d'a priori physiques sur la FEP.

Table des matières

Notations and conventions

Conventions

All the vectors are treated as columns vectors, unless it is explicitly mentioned otherwise.

- x : a scalar.
- $\mathbf{x}$: a column vector or a scalar when explicitly indicated or a manifold's point when explicitly indicated.
- $\mathbf{x}[i]$: i^{th} coefficient of $\mathbf{x}$.
- $\mathbf{X}$: a matrix or a random variable.
- $\mathbf{X}[i,j]$: the value of the entry (i,j) of $\mathbf{X}$.
- $\mathbf{X}[i,:]$: the i^{th} line of $\mathbf{X}$ treated as a line vector.
- $\mathbf{X}[:,j]$: the j^{th} column of $\mathbf{X}$ treated as a column vector.

Operators

- $\text{supp}(\mathbf{x})$: support of the vector $\mathbf{x}$ (set of non-zero coefficients of $\mathbf{x}$).
- $\text{card}(\mathcal{E})$: cardinality of the set $\mathcal{E}$ (the number of elements in $\mathcal{E}$).
- $\|\mathbf{x}\|_p$: with $p > 0$, ℓ_p norm of $\mathbf{x}$, defined as $\|\mathbf{x}\|_p = \sqrt[p]{\sum_i |\mathbf{x}[i]|^p}$. The pseudo-norm ℓ_0 is defined as $\|\mathbf{x}\|_0 = \text{card}(\text{supp}(\mathbf{x}))$ and the infinity norm is defined as $\|\mathbf{x}\|_\infty = \max(\{|\mathbf{x}[i]|, i\})$.
- $\|\mathbf{X}\|_p$: with $p > 0$, matrix norm defined as $\|\mathbf{X}\|_p = \sqrt[p]{\sum_{i,j} |\mathbf{X}_{i,j}|^p}$ (Frobenius norm for $p = 2$). The matrix ℓ_0 pseudo-norm and infinity norm are similarly defined.
- $|||\mathbf{X}|||$: subordinate matrix norm to a vector norm $\|\cdot\|$, defined as $|||\mathbf{X}||| = \max_{\mathbf{v}\neq 0} \frac{\|\mathbf{X}\mathbf{v}\|}{\|\mathbf{v}\|}$.
- $\odot$: element-wise matrix multiplication (Hadamard's product).
- $\mathbf{X}^T$: transpose of $\mathbf{X}$.
- $\mathbf{X}^H$: hermitian transpose of $\mathbf{X}$.
- $\langle \cdot, \cdot \rangle$: scalar product.
- $\lfloor x \rfloor$: the highest integer smaller than x.

- $\text{rank}(\mathbf{M})$: $\dim(\text{span}((\mathbf{M}[:,i])_{1\leq i\leq p}))$, where p is the number of columns of the matrix $\mathbf{M}$.
- $\text{Trace}(\mathbf{M}) = \sum_{i=1}^{n} \mathbf{M}[i,i]$, for $\mathbf{M} \in \text{M}_n(\mathbb{R})$.

Notations

- $\mathbb{N}$: the set of non-negative integer numbers.
- $\mathbb{R}$: the set of real numbers.
- $\mathbb{R}_+$: the set of non-negative real numbers.
- $\mathbb{C}$: the set of complex numbers.
- $\text{M}_{np}(\mathbb{K})$: the set of $n \times p$ real or complex matrices if $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ respectively.
- $\text{M}_n(\mathbb{K})$: the set of $n \times n$ real or complex matrices if $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ respectively.
- $\boldsymbol{I}_n \in \mathbb{R}^{n\times n}$: identity matrix.
- $L_1(\mathbb{R})$: the set of integrable functions on $\mathbb{R}$.
- $L_2(\mathbb{R})$: the set of square integrable functions on $\mathbb{R}$.
- $\dim(E)$: dimension of the vector space E.
- $E^{\perp}$: the orthogonal of a vector space E.
- $\text{Span}((\mathbf{u}_i)_{1\leq i\leq n})$: the set $\{\sum_{i=1}^{n} a_i\mathbf{u}_i, (a_i)_{1\leq i\leq n} \in \mathbb{R}^n\}$.
- $\mathbf{Y} \sim \mathcal{N}(\mathbf{m}, \mathbf{V})$: $\mathbf{Y}$ is a random vector that follows a gaussian multivariate distribution of mean $\mathbf{m}$ and covariance matrix $\mathbf{V}$.
- $[\![n, m]\!]$, $(n, m) \in \mathbb{N}^2$, $n < m$: the set $(n + i)_{0\leq i\leq m-n}$

List of abbreviations

— **CG** : Conjugate Gradient
— **CMB** : Cosmic Microwave Background
— **CTI** : Charge Transfert Inefficiency
— **ESA** : European spatial agency
— **fov** : field-of-view
— **FWHM** : Full Width at Half Maximum
— **i.e.** : id est
— **GMRA** : Geometric Multi-Resolution Analysis
— **HLLE** : Hessian-based Local Linear Embedding
— **HR** : High Resolution
— **HST** : Hubble Space Telescope
— **MAD** : Median Absolute Deviation
— **MAP** : Maximum a Posteriori
— **MDS** : Multi-Dimensional Scaling
— **ML** : Manifold Learning/Maximum Likelihood
— **LLE** : Local Linear Embedding
— **LP** : Linear Programming
— **LR** : Low Resolution
— **OT** : Optimal Transport
— **PCA** : Principal Component Analysis
— **PSD** : Positive Semidefinite
— **PSF** : Point Spread Function
— **RCA** : Resolved Component Analysis
— **SNR** : Signal to Noise Ratio
— **SR** : Super-Resolution
— **s.t.** : subject to
— **SVD** : Singular Value Decomposition
— **WGN** : White Gaussian Noise
— **WL** : Weak Lensing
— **w.r.t.** : with respect to

Chapitre I

Introduction

Sommaire

This thesis aims at proposing new methods and algorithms for astronomical images restoration, taking advantage of recent advances in applied mathematics and more specifically in calculus of variations ; a focus is made on Point Spread Function (PSF) estimation. In the following, we precise the scientific framework of our studies, underlining the challenges from a data processing perspective. Then, we give an overview of our mathematical environment. We end this introductory chapter with a description of the whole manuscript structure.

A Context

A.1 Weak gravitational lensing

The General Relativity theory can be considered as the cornerstone of Cosmology [Einstein 1915]. Over a century, it has allowed, conjointly with spectacular progress in instrumentation, an increasingly accurate understanding of the Universe general dynamic and geometry. Various cosmological models have been derived in this framework. The most widely studied and accepted is the Big Bang Cosmological model which lies on strong observational pillars such as the Cosmic Microwave Background (CMB) [Penzias & Wilson 1965] and the Universe expansion [Riess *et al.* 1998].

However, in order to be in perfect agreement with observations, this model has to postulate the existence of unknown and hardly observable entities, and most importantly, make them the dominant part of the Universe in terms of mass and energy. Specifically, the Universe would roughly consist of 25% percent of "Dark Matter" and 70% of "Dark Energy", the ordinary matter being confined into a few percents (see Fig. ***I.1***).

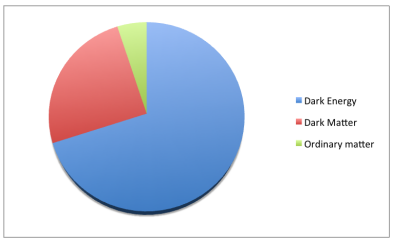


Figure *I.1* – The Universe content.

The notion of "Dark Matter" was introduced by the astrophysicist Fritz Zwicky in 1937 [Zwicky 1937]. Measuring galaxies speeds in the Coma cluster, he noticed that peripheral galaxies were excessively fast : given the observable mass of the cluster, those galaxies should have escaped its gravitational potential. He suggested as an explanation that the cluster must have an important "hidden" mass - consisting of dark matter - so that it can coherently maintain fast rotating galaxies within its gravitational potential. The Dark Energy on the other hand is directly related to the Universe global dynamic. Whereas the idea of an expanding Universe is admitted since Edwin Hubble observations in 1929, since 1998, several measurements have enlightened the acceleration of this expansion. The Dark energy presumably drives this acceleration by acting as a negative pressure.

Because of their larger proportions with respect to ordinary matter, these dark components shape the Universe at large spatio-temporal scales. Therefore, getting a better knowledge of their nature and distribution has become a major challenge for modern cosmology.

The descriptive "dark" comes from the fact that these entities can not be directly observed. However, in theory it is possible to observe and quantify the gravitational effect of the dark matter. Indeed, although it weakly or does not interact with baryonic matter, it is massive and it should bend the space-time according to Einstein equations and affect the way light travels.

Thus, a hypothetical perfectly circular galaxy would appear slightly elliptical and brighter to a hypothetical distant observer because of the integrated deflections due to dark matter, along the line of sight. This effect is known as the Weak Gravitational Lensing and is illustrated in Fig. .

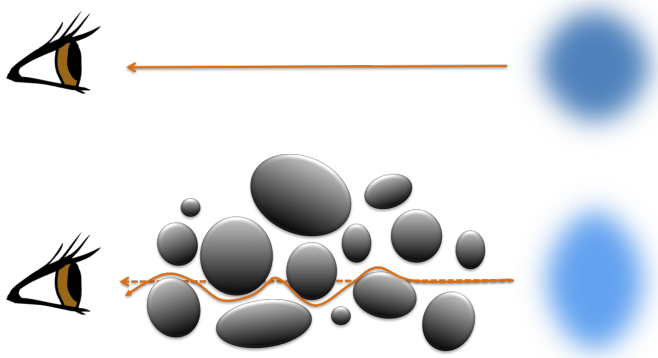


Figure *I.2* – Weak gravitational lensing.

In order to quantify this effect, one needs to be able to measure a galaxy shape from an image, which gives in practice the "lensed" galaxies shape. Since the "unlensed" galaxies are inaccessible, one needs some prior knowledge on their shapes distribution. According to the Cosmological Principle, the Universe is isotropic and homogeneous at large scales. From this principle, one infers that given a set of vectors $\boldsymbol{\gamma}_i$ in $\mathbb{R}^2$ describing the elongations and orientations of some unlensed galaxies $\mathbf{X}_i$, for a sufficiently large number of galaxies, one should have $\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\gamma}_i \approx 0$. However, the weak lensing (WL) causes an intrinsic bias in the galaxy shapes distribution which is directly related to the dark matter mass distribution and which can be estimated from the observed galaxies. The dark matter distribution in turn, can be related to the dark energy density in the Big Bang cosmological model [Dodelson 2003]. The Euclid mission, a spatial survey of the European Spatial Agency (ESA), has been designed to provide data for such a purpose, as we will see in the next section.

A.2 ESA Euclid project

Euclid is a Medium-class ESA survey mission [Laureijs *et al.* 2011]. It is part of the ESA's program "Cosmic Vision". Euclid primary goal is to learn about dark matter and dark energy nature. The Euclid telescope, which is space-based, will be launched on a Soyuz rocket from Europe's spaceport in Kourou, in Guiana ; the launch is currently planed for December 2020. Quantifying WL requires measuring accurately shapes over a large set of galaxies. To that end, the Euclid telescope

— provides a large field-of-view (fov) ; the embedded visible and near-infrared instrument have a fov of roughly 0.54deg^2 which is two orders

of magnitude higher than Hubble Space Telescope (HST) camera's ;

— is extremely sensitive with a lower magnitude bound of AB mag 24.5 (this corresponds to a spectral luminosity 17 order of magnitude lower than the sun's) ; this enables the observation of distant galaxies which is essential, since the further is a galaxy, the more it is potentially lensed ;

— is planned to orbit around the Sun-Earth Lagrange point 2 (see Fig. ***I.3***) ; consequently it will be free of atmospheric turbulence (which is true for any space-based telescope), as well as Earth and Moon magnetic fields disturbances ; moreover, the payload will neither be occulted nor illuminated by the Sun, the Earth or the Moon during the survey ; finally, there is no significant variation of gravity along the selected orbit ; all these features will yield an exceptionally homogeneous final data-set ;

— is extremely stable in terms of structural vibrations, with a specified pointing jitter in average smaller than the pixels size.

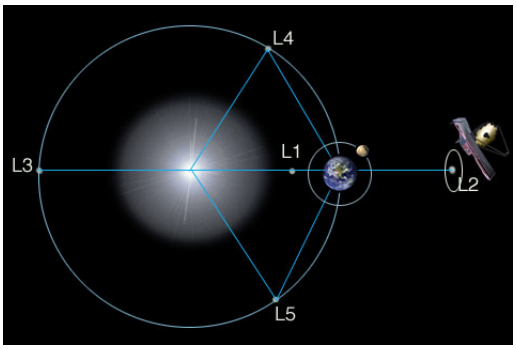


Figure *I.3* – Sun-Earth system Lagrange points. Credit : nasa.gov

The survey active phase will last 6 years, providing at the end a legacy of more than 1 billion of well resolved galaxies images, positions and speeds, distributed over 15000deg^2 of sky without the milky way and roughly 10 billions light years in depth, thanks to its extreme sensitivity. Henceforth, Euclid will show the Universe at a spatio-temporal scale that has never been reached before, which makes this mission exciting and promising for Cosmology.

Regardless of how clean will Euclid data presumably be, it is crucial however to have an accurate knowledge of the instrumental effects on the galaxies shapes, in order to exploit them properly. Indeed, the image forming process itself induces shape distortions that might be mistaken for the Weak Lensing (WL) effect. This is one of the biggest challenge in WL measurement, as we detail in the next section.

A.3 Data and instrumental effects

In this section, we review the image forming process for an Euclid-like telescope, precising the distortions introduced with respect to the "real world objects". Despite galaxies are 3D objects, our conceptual real world objects are 2D light distributions contained in a plane which is perpendicular to the line of sight as illustrated in Fig. ***I.4*** ; we do not consider the geometrical distortion associated with this planar projection.

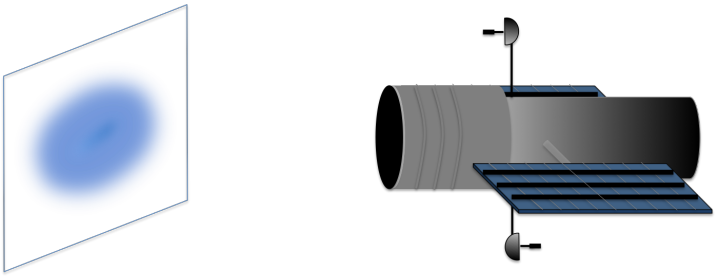


Figure *I.4* – Planar projection of the sky : the distortion associated to this projection with respect to the real world 3D objects is not considered.

A.3.1 Optics related effects

The optical instrument warping effect is largely related to the phenomenon of diffraction, which characterizes a propagating wave behavior whenever it encounters an obstacle. Therefore, we make a brief recall of diffraction theory applied to optics thereafter. The formulae are extracted from [Goodman 2005].

According to the Huygens-Fresnel principle, when a wave propagates, each point of the space reached by the wave becomes in its turn the source of a spherical wave. This is illustrated in Fig. ***I.5***.

These secondary waves can interfere : according to their phases, they can add up in amplitudes or attenuate one another. In particular, when the wave encounters an obstacle, as one may expect intuitively, the secondary sources physical properties change in the vicinity of the boundaries so that the overall transmitted wave might significantly depart from the incident one ; this is especially true when the size of the aperture (see Fig. ***I.5***) is close to the wavelength.

The first rigorous description of the diffraction by a planar aperture in terms of superposition of secondary spherical waves is due to the physicist Gustav Kirchhoff (1882). He proposed a mathematical framework built upon the Greens theorem and derived the diffracted wave making some simplifying

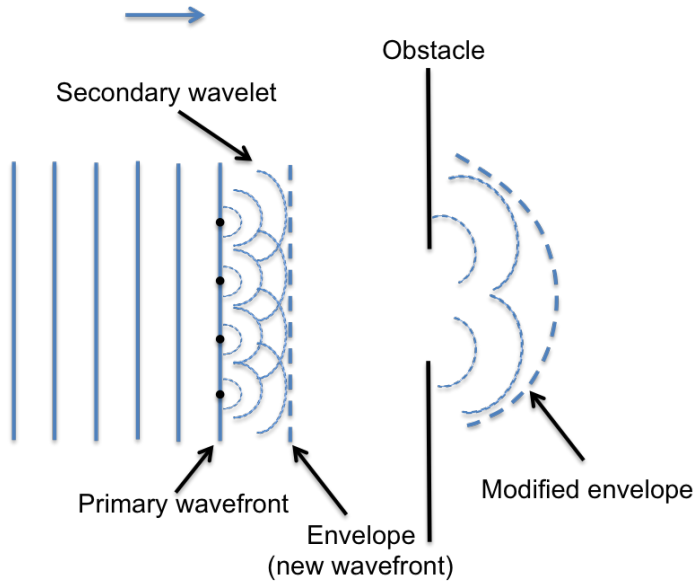


Figure *I.5* – Diffraction : Huygens-Fresnel principle.

assumptions known as the *Kirchhoff boundary conditions*. These assumptions were later proved to be inconsistent by Poincaré (1892) and eliminated by Sommerfeld (1894) who proposed the so-called *Rayleigh-Sommerfeld diffraction theory*. The latter characterizes the diffracted wave in a parallel observation plane located at normal distance z from the aperture (see Fig. ***I.6***) as follows :

$$U(\mathbf{p}_1) = \frac{z}{j\lambda} \int\int_{\mathbb{R}^2} i_{\mathbf{\Sigma}}(\mathbf{p}_0) U(\mathbf{p}_0) \frac{\exp(jkr_{01})}{r_{01}^2} d\eta d\xi, \tag{I.1}$$

where $\mathbf{p}_1 = (x, y, z)$, $\mathbf{p}_0 = (\eta, \xi, 0)$, $r_{01} = \|\mathbf{p}_1 - \mathbf{p}_0\|_2$, $U(\mathbf{p}_0)$ is the incident wave amplitude and phase at the point $\mathbf{p}_0$ of the aperture, λ is its wavelength, $k = \frac{2\pi}{\lambda}$ and $\mathbf{\Sigma}$ is the aperture.

Two interesting approximations of the Rayleigh-Sommerfeld diffraction are later proposed. The first one is the Fresnel approximation. It relies on the following approximation :

$$r_{01} \approx z(1 + \frac{1}{2}(\frac{x-\eta}{z})^2 + \frac{1}{2}(\frac{y-\xi}{z})^2). \tag{I.2}$$

Taking $r_{01}^2 \approx z^2$ in I.1 and replacing the r_{01} appearing in the exponent with I.2, one can express the diffracted field as

$$U(\mathbf{p}_1) = \frac{e^{jkz}}{j\lambda z} e^{j\frac{k}{2z}(x^2+y^2)} \int\int_{\mathbb{R}^2} (i_{\mathbf{\Sigma}}(\mathbf{p}_0) U(\mathbf{p}_0) e^{j\frac{k}{2z}(\eta^2+\xi^2)}) e^{-j\frac{2\pi}{\lambda z}(x\eta+y\xi)} d\eta d\xi, \tag{I.3}$$

where we can identify the Fourier transform of the incident wave delimited by the aperture and multiplied by a quadratic phase term. The regime in which

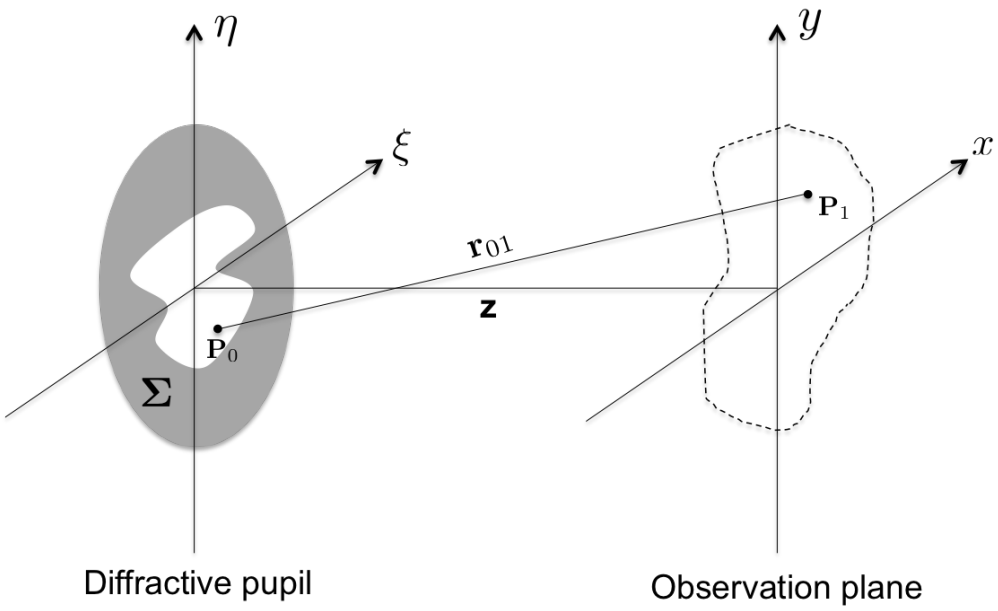


Figure *I.6* – Diffraction : Huygens-Fresnel principle.

this approximation holds is known as the region of Fresnel diffraction, or the near field of the aperture, since it accounts for the curvature of the diffracted wavefront through the quadratic phase term. It yields accurate results when the major contributors in the integral I.1 are the terms for which $x \approx \eta$ and $y \approx \xi$.

Besides, the Fraunhoffer approximation additionally assumes that

$$z \gg \frac{k \max_{\mathbf{\Sigma}}(\eta^2 + \xi^2)}{2}, \tag{I.4}$$

which implies that the quadratic phase term in the integral in I.3 is approximately equal to 1 everywhere in $\mathbf{\Sigma}$. Therefore, the diffracted field expression simplifies to

$$U(\mathbf{p}_1) = \frac{e^{jkz}}{j\lambda z} e^{j\frac{k}{2z}(x^2+y^2)} \int\int_{\mathbb{R}^2} i_{\mathbf{\Sigma}}(\mathbf{p}_0) U(\mathbf{p}_0) e^{-j\frac{2\pi}{\lambda z}(x\eta+y\xi)} d\eta d\xi. \tag{I.5}$$

The diffracted wave is given by the *Fourier transform* of the aperture limited incident wave up to a multiplicative term.

Now we consider a general linear optical system that we treat as a black box as represented on Fig.***I.7***.

The optical system maps an incident scalar wave $\mathbf{U}_i$ into an output wave given by the so-called superposition integral :

$$U_o(x,y) = \int\int_{\mathbb{R}^2} U_i(u,v) h_o(x,y,u,v) du dv, \tag{I.6}$$

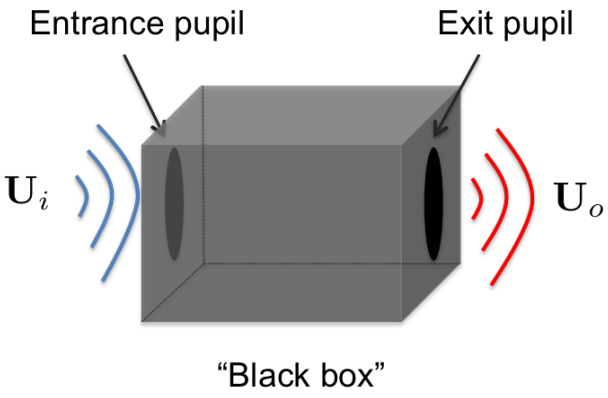


Figure *I.7* – "Black box" linear optical system ; U_i : input wave ; U_o : output wave

where $h_o(x, y, u, v)$ is the field's value at the location (x, y) of the exit pupil that one would get from the optical system as a response to an unitary impulse located at (u, v) in the entrance pupil. h_o is the impulse response of the system. The system is stationary if for a couple of locations $\{(u, v), (x, y)\}$, $h_o(x, y, u, v) = h_o(u-x, v-y)$; in words, shifting an impulse in the object space translates into a shifting of the output wave, with no changing in its form. For stationary systems, the right hand side in Eq. I.6 becomes a convolution. Real imaging systems are never perfectly stationary and this non-stationarity motivates an important fraction of this thesis work. The concepts of entrance and exit pupils generalized the notion of aperture used in deriving the Rayleigh-Sommerfeld and Fraunhoffer diffraction approximations. Firstly introduced by Ernst Abbe, the entrance pupil is the image of the most severely limiting aperture when viewed from the object space ; conversely, the exit pupil is the image of the same limiting aperture viewed from the image space. Regardless of the optical system's complexity, the wave that reaches the image's plane can be described based on the diffraction effects that would result from the exit pupil. We assume that the light propagation between the entrance and exit pupil plans can be well described by geometrical-optics. The system is said to be diffraction-limited. The ideal geometrical-optics predicted image of the incident wave can be defined as

$$U_g(u', v') = \frac{1}{|M|} U_i(\frac{u'}{M}, \frac{v'}{M}), \tag{I.7}$$

where M is a magnification factor.

Under these hypothesis, the image wave is given by

$$U_o(x, y) = \int\int_{\mathbb{R}^2} U_g(u', v') h_o(x - u', y - v') du' dv', \tag{I.8}$$

where the impulse response is calculated within the Fraunhoffer approximation as

$$h_o(u,v) = \frac{A}{\lambda z_i} \int\int_{\mathbb{R}^2} P(x,y) e^{-j\frac{2\pi}{\lambda z_i}(xu+yv)} dxdy. \tag{I.9}$$

A is a constant amplitude. z_i is the distance from the exit pupil plan to the image plan. P is a binary mask defining the exit pupil. It is often referred to as the pupil function.

In most imaging system, the photo-current generated by the sensors is proportional to the incident power density. Therefore, one can measure to the diffracted field's intensity which is defined as

$$I(x,y) = |U_o(x,y)|^2 \tag{I.10}$$

for a monochromatic light source or more generally

$$I(x,y) = \lim \frac{1}{T} \int_{-T/2}^{T/2} |U_o(x,y,t)|^2 dt = \langle |U_o(x,y,t)|^2 \rangle_t \tag{I.11}$$

for polychromatic sources. This last definition is particularly relevant from a practical of view since the photo-detectors integration time are in general substantially longer than the incident wave temporal period.

Assuming that the field U_i has a sufficiently narrow spectrum, the impulse response h_o can be considered to be wavelength independent so that Eq.I.8 is generalized as follows :

$$U_o(x,y,t) = \int\int_{\mathbb{R}^2} U_g(u',v',t-\tau) h_o(x-u',y-v') du'dv'. \tag{I.12}$$

The delay τ is related to the propagation from (u',v') to (x,y). One can justify that under the narrowband assumption, the intensity can be written as

$$I(x,y) = \int\int_{\mathbb{R}^2} du'_1 dv'_1 \int\int_{\mathbb{R}^2} du'_2 dv'_2 h_o(x-u'_1, y-v'_1) h_o^*(x-u'_2, y-v'_2) J_g(u'_1, v'_1, u'_2, v'_2) \tag{I.13}$$

where $J_g(u'_1, v'_1, u'_2, v'_2) = \langle U_g(u'_1, v'_1, t) U_g^*(u'_2, v'_2, t) \rangle_t$. J_g is the so-called mutual intensity and measures the spatial coherence of the field U_g.

If the illumination is spatially coherent, the phases of the different impulses in the object's plane are perfectly correlated. This is nearly the case if the light is emitted by a laser. Reversely, the illumination is said to be spatially incoherent if the phases vary in an uncorrelated way in the object's plane. Most of the common light sources, including natural ones can be in first approximation considered as incoherent. In this case, the mutual coherence is simplified to

$$J_g(u'_1, v'_1, u'_2, v'_2) = I_g(u'_1, v'_1) \delta(u'_1, u'_2) \delta(v'_1, v'_2), \tag{I.14}$$

where I_g is the intensity associated with the field U_g. From Eq.I.14 and Eq.I.13 one obtains the following important equation relating the incident and diffracted field intensities for a perfectly incoherent lighting :

$$I(x,y) = \int\int_{\mathbb{R}^2} |h_o(x-u_1', y-v_1')|^2 I_g(u_1', v_1') du_1' dv_1'. \tag{I.15}$$

In words, for a diffraction-limited imaging system, when the illumination is incoherent, the image intensity is obtained by convolving the ideal geometrical image intensity with the intensity impulse response $|h_o|^2$, which is mostly referred to as the Point Spread Function (PSF).

The Optical Transfert Function (OTF) is defined as the normalized Fourier transform of the PSF. For a diffraction-limited imaging system, it can be shown that the OTF is the normalized autocorrelation function of the pupil function P (see Eq.I.9). As such, the OTF has some interesting properties among which it is an even function. The pupil function being by definition spatially bounded, this also implies that a PSF is bandlimited.

The results stated so far were derived assuming that the optics do not introduce any phase-shifting as the incident wave goes though the system. However these phase-shiftings also termed as *wavefront errors* always occurs due to optical aberrations. In order to account for the wavefront errors, one shall generalize the notion of pupil function. We note the phase-shifting at a point (x,y) $\phi(x,y) = \frac{2\pi}{\lambda} W(x,y)$, where λ is the central wavelength of the spectrum of the incident wave. $W(x,y)$ represents the optical path difference. Then, the generalized pupil function is defined as

$$\mathcal{P}(x,y) = P(x,y) e^{j\frac{2\pi}{\lambda}W(x,y)}. \tag{I.16}$$

Then the amplitude impulse response is given by the diffraction pattern of an aperture having a complex amplitude transmittance $\mathcal{P}$. The PSF's definition remains unchanged.

It is worth noting that the pupil function P actually depends on the position of the light source position w.r.t. the instrument fov. This makes the PSF spatially variable. In general, computing the pupil function requires performing a ray tracing using a dedicated software as Zemax. Moreover, the different optical aberrations effects are more or less pronounced depending on the light source's position in the fov. Let consider the well-known coma aberration for instance. The coma is due to a variation of magnification across the fov. Indeed, when a bundle of rays coming from a single off-axis source hits a lens or a mirror, each ray is reflected toward a different point in the focal plane. The more a ray is close to the optical element edge, the further from the optical axis the reflected ray crosses the focal plane. This gives isolated star

images a cometary shape, hence the name of this aberration. As illustrated in Fig. ***I.8***, the coma varies across the fov, which is translated by a spatial dependency of the function W is Eq.I.16. Thus optical aberrations contribute

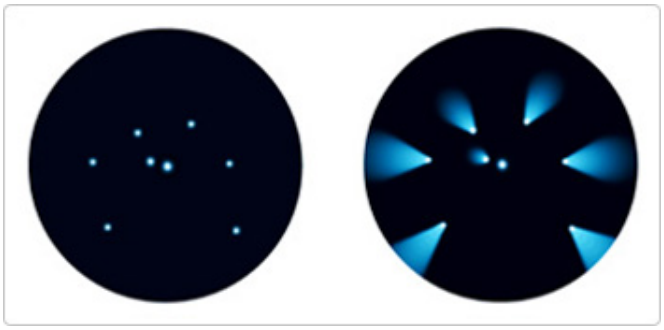


Figure *I.8* – Simulated coma aberration : the coma varies across the fov ; in particular, the further is the source from the fov center, the more dramatic is the distortion. Credit : nikon.com

to the PSFs variations across the fov. The optical path difference function should be noted W_{x_i,y_i} to explicit its dependency to the impulse's position in the object's plane. However we omit these indexes for simplicity. We shall add that aberrations such as defocusing or tilt can vary in time due to thermo-mechanical constraints on the imaging system. This is especially relevant for space based telescopes. This makes the PSF ultimately time dependent. Moreover, as shown by Eq.I.9, the PSF also depends on the wavelength ; for a single channel imaging system, the image of a broadband source of light is the result of a continuous summation of ideal monochromatic images convolved with a PSF that might considerably change across the spectrum.

A detailed presentation of optical aberrations can be found in [Mahajan 1998].

A.3.2 Electronics related effects

So far, the images have been modeled as continuous intensity distributions. However for numerical images, the diffracted intensity distribution is recorded via a matrix of photo-detectors. We assume that the detectors cover 100% of the focal plane. Each detector integrates the light intensity over its surface during a certain amount of time. We note Δ the detector matrix step and $\mathcal{D}$ the focal plane area covered by a reference detector. The intensity recorded by the detector (k, l) is in first approximation given by

$$\mathbf{I}_{kl} = \int\int_{\mathcal{D}} I(k\Delta + x, l\Delta + y)dxdy, \tag{I.17}$$

up to a multiplicative constant. Therefore, the numerical image forming process can be described as a convolution of the intensity distribution I with the kernel defined as

$$h_{\mathcal{D}}(x,y) = \begin{cases} 1 \text{ if } x \in \mathcal{D}, \\ 0 \text{ otherwise}, \end{cases} \tag{I.18}$$

followed with a sampling.

However, this detector model is idealistic and we give below a non-exhaustive list of limitations :

— thermal noise : even in the absence of illumination, the detectors unavoidably generate a signal due to random thermal motions of electrons ; the resulting "background image" is generally modeled as a white gaussian process [Nyquist 1928] ;

— dark-current shot noise : this noise is due to random generation of electrons in the detectors and is also related to the temperature ; however unlike the thermal noise it is non-gaussian [Baer 2006] ;

— readout noise : this noise is inherent to CCD detectors ; it is related to the uncertainty on the photo-electrons total charge measurement in the CCD matrices ships due to the imperfections of the electronic components ; it is well modeled as white gaussian process [Basden *et al.* 2004] ;

— charge transfer inefficiency (CTI) : this effect is particularly present for space telescope ; the high energy radiations might gradually damage the CCD detectors ; this might cause delays in the electrons transfers to the CCD matrices ships which produces a trail and blurring in the final images [Massey *et al.* 2014] ;

— quantization : in other to be stored numerically, the measured intensities have to be represented using a finite numerical "alphabet", which implies a quantization of the measurements and produces what is known as the "quantization noise" [Widrow & Kollár 2008].

Ignoring some of the above effects, we can summarize a digital image forming process as follows :

$$\mathbf{I}_d = \mathcal{S}(h_{\mathcal{D}} * h_o * I_g) + \mathbf{N}, \tag{I.19}$$

where $\mathcal{S}$ is a sampling operator and $\mathbf{N}$ is the noise. In the following, we refer to the function $h_{\mathcal{D}} * h_o$ as the PSF. A natural question arising then is whether or not the sampling operation preserves the continuous light intensity information content. In other words, can the continuous function $h_{\mathcal{D}} * h_o * I_g$ be calculated from the discrete array $\mathcal{S}(h_{\mathcal{D}} * h_o * I_g)$?

The Nyquist-Shannon sampling theorem gives a sufficient condition. We recall its classical formulation : if $f \in L_1(\mathbb{R})$ and if its Fourier transform $\hat{f}$'s support is included in the interval $[-B, B]$, then

$$f(x) = \lim_{n \to +\infty} \sum_{k=-n}^{k=n} f\left(\frac{k}{2B}\right) \operatorname{sinc}\left(2\pi B\left(x - \frac{k}{2B}\right)\right). \tag{I.20}$$

In words, if an integrable function is bandlimited with a maximum frequency B, then it can be "reconstructed" from an uniform sampling with a step $\frac{1}{2B}$. This theorem naturally extends to higher dimensions. In practice, the light intensity distribution that reaches the detectors matrices is by construction spatially bounded so that the integrability condition always applies. The second condition is guaranteed by the fact that the PSF is bandlimited.

Therefore, it is possible to preserve the focal plane intensity information, up to the electronics defects, if the sampling step, which is determined by the detectors size, is sufficiently fine. As we detail in Chapter IV, an insufficient sampling rate might cause frequency aliasing, which is a frequency content distortion independent of the instrumental effects mentioned so far.

A.3.3 Other sources of images distortions

We consider external sources of distortions. In astronomy, forming images requires long exposure times because of the weak luminosity of most galaxies seen from the solar system. For instance, the nominal exposure time for the Euclid telescope is 540s [Laureijs *et al.* 2011].

Due to the intrinsically random nature of the light emission process, the number of photons that hits a detector during a given amount of time can be modeled as a Poisson process. The uncertainty related to the photons count is what is referred as shot noise.

During the exposure time, a space telescope is not perfectly steady which yields a convolutive motion blur which is time varying. For ground based telescopes, the atmospheric turbulence causes a blurring which is both space and time varying. The Euclid telescope will be put into orbit above the atmosphere so that the latter blur need not to be considered. The telescope pointing jitter is specified not to exceed a pixel size, which makes the motion blur moderated ; besides, it can be readily accounted for since pointing direction times series will be available.

In this work, we consider monochromatic images. The Chapters IV to VI are dedicated to estimating discretized but well sampled PSFs in the telescope's field from undersampled and noisy stars images and assuming that the detectors response is uniform and linear. The PSFs spatial variability is taken into account. The quantization noise is considered to be negligible. The shot

noise which will be more or less considerable depending on the brightness of each particular observation is not treated. The CTI effect as well is not considered.

B Methods and tools

A physical model free framework for modeling a telescope's PSFs field is proposed. This framework aims at extracting the information available in a set of instrumental data at best, while making fairly general regularity assumptions on the PSFs. We give a brief overview of the key concepts and tools in the following.

B.1 Low complexity

In the Fraunhoffer approximation, the amplitude impulse response of a linear imaging system is a scaled Fourier transform of the pupil function, up to a multiplicative factor (see Eq.I.9). This makes the optical PSF a uniformly continuous function. Therefore a discrete well sampled PSF is at least a piece-wise smooth signal, especially if it includes the detectors response. This structure or regularity implies that a discrete well sampled PSF actually has less degrees of freedom than the number of pixels needed for a direct representation. It is in principle possible to represent the PSFs in a more concise manner. This notion of conciseness classically referred as sparsity has been extensively studied and used for solving inverse problems. It will be shown very useful, especially regarding noise robustness. Besides, we mentioned that the PSF varies spatially, due to the changes of the generalized pupil function across the fov (see Eq.I.16). More specifically, the modulus function P varies according to the whole optical instrument complexity and the optical path difference function W variations are related to the optical aberrations and therefore, the optics quality. In Euclid, the spatial variations of the optical PSF are mostly due to the aberrations. We can roughly distinguish two fundamentally different types of aberrations :

— the global aberrations related to the global structure and defects of the optics as well as their relative positioning ; they include effects such as defocusing, tilt, coma etc. ; they induce a smoothly varying phase-shifting across the fov and constitute the low frequencies of the function W ; these aberrations are well-modeled using Zernike polynomials [Wang & Silva 1980] ;

— the local aberrations due to polishing defects ; indeed, the small residual structures distributed over the optics surfaces diffuses light inducing

random phase-shiftings ; thus these aberrations constitute the high frequencies of W and their effect is susceptible to change rapidly w.r.t. the impulse's position in the object's plane.

If we assume that the wavefront errors are mainly determined by the global aberrations, then the PSFs evolves smoothly across the fov and are strongly correlated. The latter point implies that a set of discrete well-sampled PSFs is scattered along a limited number of directions relative to their ambient space dimension. This low dimensionality assumption is crucial, especially in the perspective of a joint PSFs field super-resolution. The smooth spatial variation of the PSFs links the aforementioned low dimensional scattering of the PSFs to their distribution in the fov. Considering a finite set of PSFs, this link can be elegantly built through an undirected graph inversely weighted by the PSFs locations pairwise distances in the fov. Then, the graph's function associating each node to each PSF can be enforced to be smooth by seeking sparsity.

B.2 Manifold learning

Loosely speaking, a d dimensional manifold is a topological space that can be injectively mapped to $\mathbb{R}^d$ locally. d represents the intrinsic dimension. This definition encompasses linear subspaces, also called linear manifolds, but also non-linear sets as we will see in the following. If a manifold is a collection of real world data, the intrinsic dimension can be interpreted as the number of degrees of freedom of the physical system that generated those data. For example, Fig.***I.9*** shows samples from the manifold consisting of images of a given subject face with a constant pose, for all the lighting angles and positioning possible. Although the ambient space dimension, which is the number of pixels, can be substantially large, the intrinsic dimension is limited to five : three spatial coordinates and two angular coordinates.

For a fixed wavelength, the manifold model naturally arises in our study since the optical PSF is completely and uniquely determined by two parameters which are the impulse coordinates in the object plane. As shown in Section A.3.1, there is a highly non-linear relationship between these coordinates and the PSFs, which makes the manifold perspective particularly interesting.

The Manifold Learning (ML) is a discipline at the confluence of geometry and data science which aims at uncovering information on a manifold from a set of data sampled from this manifold. For example one can estimate the manifold's intrinsic dimension, learn a low dimensional parametrization or characterize the manifold's curvature. Let recall that ultimately, we want to estimate the PSFs at galaxies locations in the telescope fov, where no PSFs measurements are available. This requires being able to perform displacement

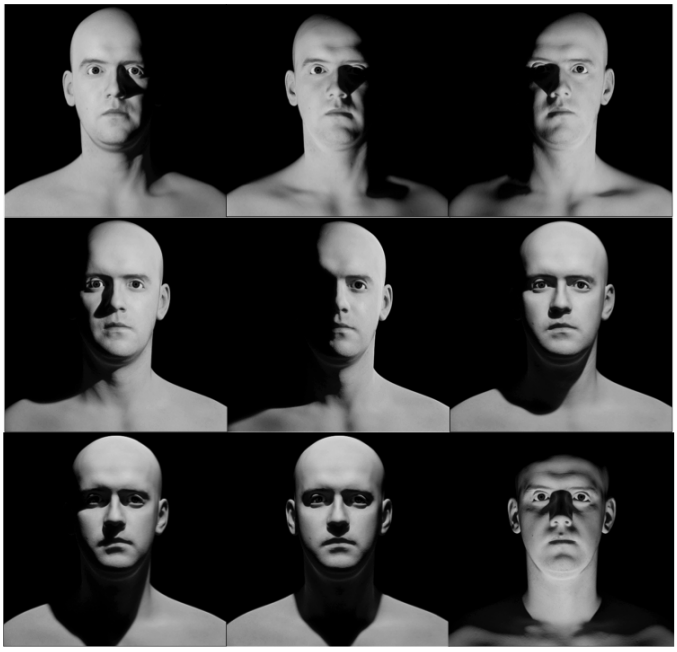


Figure *I.9* – "Appeareance manifold" : all the images are generated with a single facial pose, moving a single light source.

within the PSFs manifold, following the fov coordinates parametrization. This falls into the ML framework.

Fig.***I.10*** shows a plot of an "S-curve" data set, which is regularly used for ML algorithms benchmarking. It is clear from this plot that the euclidean distance between two points far apart (in the euclidean sense) might strongly depart from the minimal length of the paths joining these points on the manifold.

By contrast, the euclidean distances between neighbor points is approximately equal to the geodesic distances between these points, if the manifold sampling is sufficiently dense. Thus, most manifold learning techniques build over the hypothesis that the geometry of the manifold is approximately euclidean locally and thereby are sensitive to the curvature/sampling density trade-off.

In the PSFs interpolation framework introduced, we propose to replace euclidean distance with an alternative distance based on Optimal Transport (OT) theory for characterizing the local PSFs manifold geometry. Indeed, Optimal Transport offers an elegant approach for comparing positive unitary mass distributions, robustly to non-linear warping. Moreover, from a practical point of view, the Optimal Transport is particularly appealing in a context of ML for interpolation since recipes and algorithms for computing geodesics

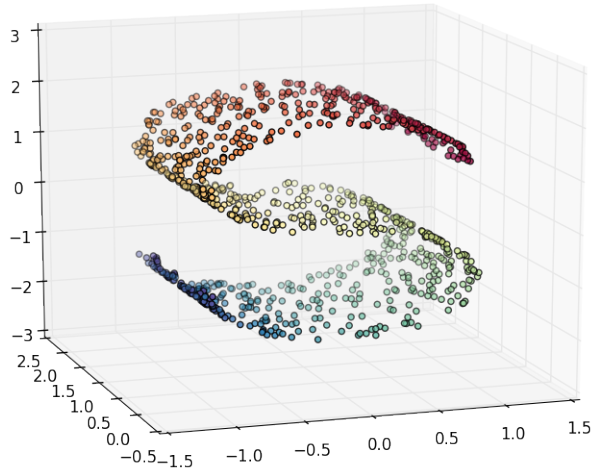


Figure *I.10* – S-curve data set.

and barycenters are readily available.

C Manuscript structure

This manuscript is organized as follows :

— in Chapter II, we make a formal presentation of low complexity data models, focusing on sparsity and low rankness ; we describe how these priors on the data structure may be used in solving inverse problems ;

— Chapter III is dedicated to geometrical concepts and tools useful in handling data lying on a non-linear manifold ; we give a general overview of ML, followed by an introduction to the OT and its geometrical features.

— in Chapter IV, we propose a sparsity-driven multiple frame PSF super-resolution (SR) method ;

— a further step is made in Chapter V where we introduce a single frame PSFs field super-resolution ; the requirement of having several low resolution (LR) observations of the same PSF is dropped and compensated by the PSFs field regularity previously mentioned ;

— the PSFs field interpolation is treated in Chapter VI ; we build a methodology for estimating a monochromatic PSF at an arbitrary location in a telescope from a set of PSFs at known locations, using Manifold Learning ideas and Optimal Transport tools ;

Chapitre II

Low complexity data models in inverse problems

Sommaire

In numerous problems in signal processing, the goal is to estimate an unknown signal from an altered observation. It is in general assumed that the signal of interest follows some sort of model or has a specific structure which knowledge can be used in the estimation process. The signal's structure can often be described using a notion of simplicity or compressibility which translates either the stability in space or time of the physical system that generated the signal at a given scale of observation, or the generating system limited number of degrees of freedom. This is illustrated in Fig. ***II.1***.

This intrinsic simplicity of the data, more often referred to as low complexity prior in the signal processing literature is at the core of this chapter. After a general introduction to linear inverse problems, we formally present some important low complexity data models and give a preview of how they can be practically used for signal restoration in a variational framework.

A Linear inverse problems

A.1 Formulation and examples

Let consider a p pixels digital image Im_d. As stated in the introductory chapter, Im_d can be modeled through the following equation :

$$\mathrm{Im}_d = \mathcal{S}(K * I_g) + \mathbf{N}, \tag{II.1}$$

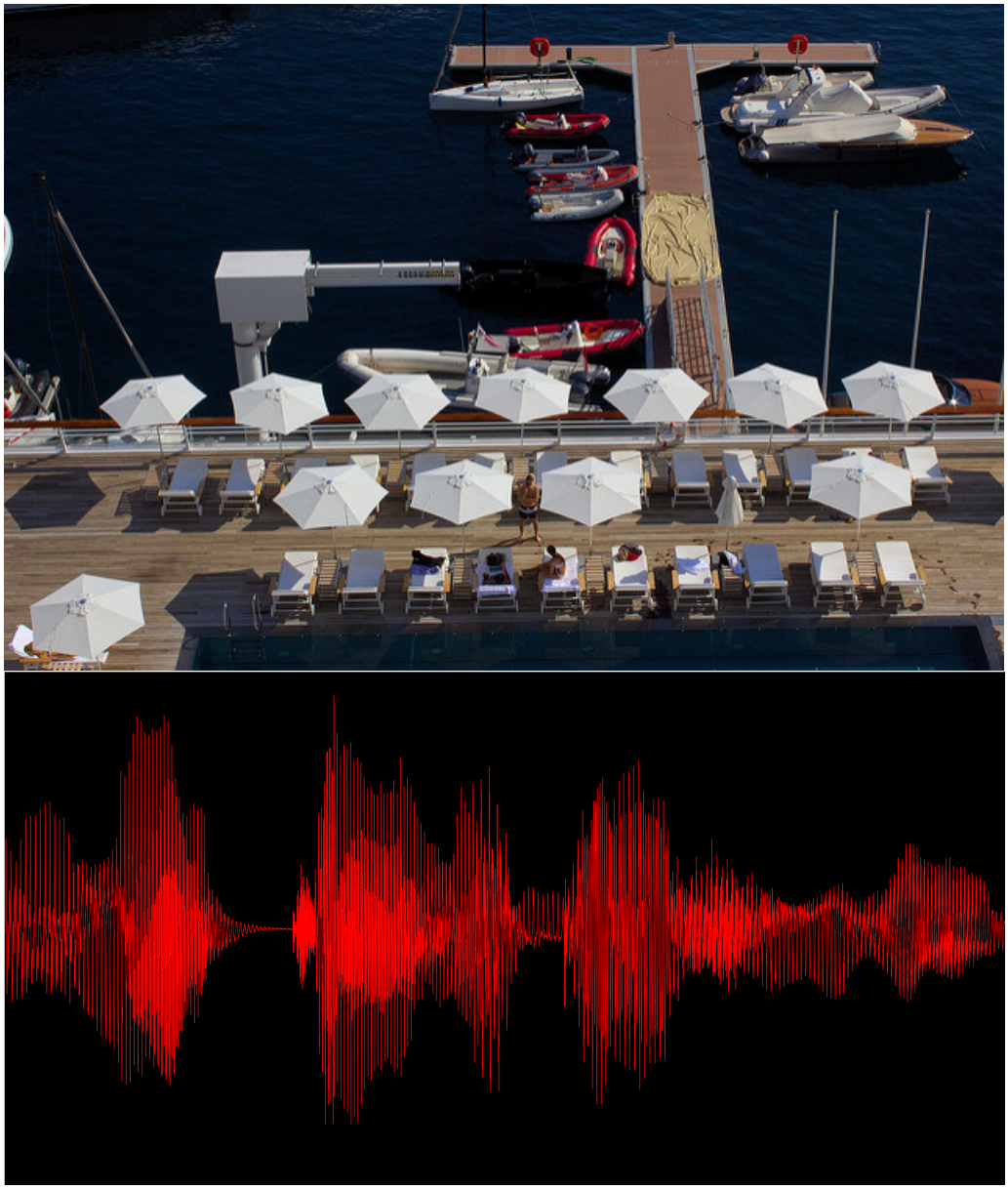


Figure *II.1* – For instance natural images are in general piece-wise smooth ; the bottom panel shows an highly structured speech signal which can be represented using a few parameters which in fact characterize the vocal organ's state during the phonation.

where $\mathbf{N}$ is the electronic related noise, I_g is a physical continuous light distribution characterizing the real world scene, K is the optical device's PSF and $\mathcal{S}$ is a function from $\mathcal{I}$ to $\mathbb{R}^p$ realizing the sampling, $\mathcal{I}$ being any functional space that contains I_g (for instance $\mathcal{I} = L_1(\mathbb{R}^2)$). While the continuous func-

tions K and I_g are practically inaccessible, one shall assume that they can be well approximated with piece-wise constant functions over a given regular grid in $\mathbb{R}^2$. Because of the finite extension of the focal plane, we can assume that the support of I_g is comprised in $[0, l] \times [0, l]$ for some positive real l chosen in such a way to also encircle most of the PSF's energy. We define the function

$$h_{0,0}(x, y) = \begin{cases} \frac{1}{(l/n)^2} \text{ if } (x, y) \in [0, l/n] \times [0, l/n], \\ 0 \text{ otherwise,} \end{cases} \tag{II.2}$$

for some positive integer n. We define the orthonormal family in $L_2(\mathbb{R}^2)$ $(h_{i,j})_{0 \leq i,j \leq n-1}$ by

$$h_{i,j}(x, y) = h_{0,0}(x - il/n, y - jl/n). \tag{II.3}$$

Then I_g and K can be approximated with their projections on $\text{Span}((h_{i,j})_{0 \leq i,j \leq n-1})$, which is a n^2 dimensional subspace :

$$\hat{I}_g = \sum_{0 \leq i,j \leq n-1} a_{ij} h_{i,j}, \tag{II.4}$$

$$\hat{K} \approx \sum_{0 \leq i,j \leq n-1} b_{ij} h_{i,j}. \tag{II.5}$$

$(a_{ij})_{0 \leq i,j \leq n-1}$ and $(b_{ij})_{0 \leq i,j \leq n-1}$ are digitally tractable representations of I_g and K respectively. Moreover, for $(r, s) \in [\![1, 2n-1]\!]^2$,

$$(\hat{I}_g * \hat{K})(rl/n, sl/n) = \sum_{\substack{(i_1,i_2) \in [0,n-1]^2 / \\ i_1+i_2=r-1}} \sum_{\substack{(j_1,j_2) \in [0,n-1]^2 / \\ j_1+j_2=s-1}} a_{i_1 j_1} b_{i_2 j_2}, \tag{II.6}$$

where one recognize a discrete convolution of $\mathbf{A} = (a_{ij})_{0 \leq i,j \leq n-1}$ and $\mathbf{B} = (b_{ij})_{0 \leq i,j \leq n-1}$ on the right hand side. n is a parameter set by the practitioner according to the resolution needed and typically, $n^2 \geq p$. For a well chosen n, we can write $\mathcal{S}(\hat{K} * \hat{I}_g) = \mathcal{D}(\mathbf{A} * \mathbf{B})$, where $\mathcal{D}$ is a discrete downsampling operator.

Finally, the forward observation model becomes

$$\text{Im}_d = \mathcal{D}(\mathbf{A} * \mathbf{B}) + \hat{\mathbf{N}}, \tag{II.7}$$

where $\hat{\mathbf{N}}$ accounts for the electronic related noise and the modeling error. The task of estimating $\mathbf{A}$ given Im_d and knowing $\mathbf{B}$ is a *linear inverse problem* because of the linearity of the *degradation operator* $\mathcal{D}(. * \mathbf{B})$ and because one is going backward from the observation to its cause. Linear inverse problems occur in a wide range of applications :

— in geophysics where knowledge of Earth deep structure is inferred from Earth surface seismograms [Tarantola 2004] ;

— in acoustics where one can characterize the propagation medium heterogeneity, namely the obstacles, from measured waves [Colton & Kress 1992] ;

— in quantum mechanics where nuclear forces can be characterized based on observable scattered radiations in particles colliders [Chadan & Sabatier 2011] ;

— in medical imaging where one is to construct an image of something which is in the human body from ideally non-invasive and non-destructive and hence indirect measurements ; there exists several medical imaging modalities among which one of the most important is the Computed Tomography [Herman & Sabatier 1987] ;

— in astrophysics, as presented in [Starck 2016] and the references therein.

A linear inverse problem can always be written into the following generic form :

$$\mathbf{y} = \mathbf{M}\mathbf{x} \circ \mathbf{b}, \tag{II.8}$$

where

— $\mathbf{y} \in \mathbb{R}^p$ is the observed signal,

— $\mathbf{x} \in \mathbb{R}^m$ is the unaltered signal,

— $\mathbf{M} \in \mathrm{M}_{pm}(\mathbb{R})$ is the degradation operator,

— $\mathbf{b}$ represents the noise and/or the modeling error,

— $\circ$ is some composition operator in $\mathbb{R}^p$.

In all this manuscript, $\circ \equiv +$, and the noise is said to be additive. However, coherent imaging systems such as Synthetic Aperture Radars or medical ultrasounds are predominantly affected by a multiplicative granular noise known as speckle [Gagnon & Jouan 1997].

A.2 Ill-posedness and ill-conditioning

As previously stated, solving the inverse problem II.8 is the task of estimating $\mathbf{x}$ from the observation $\mathbf{y}$. Without any knowledge on $\mathbf{x}$, the accuracy of this estimation is tightly related to how much information is lost along the cause-effect sequence modeled by the matrix $\mathbf{M}$ (see the interesting discussion in [Bertero & Boccacci 1998]).

If $\mathbf{M}$ is invertible, then the equation $\mathbf{y} = \mathbf{M}\mathbf{x}$ (i) admits a solution which (ii) is uniquely determined by $\hat{\mathbf{x}}_\mathbf{b} = \mathbf{M}^{-1}\mathbf{y}$. Moreover, (iii) $\hat{\mathbf{x}}_\mathbf{b}$ continuously depends on $\mathbf{y}$. These conditions of existence, uniqueness and continuity makes

the problem II.8 *well-posed* in the sense of Hadamard [Hadamard 1902, Hadamard 1923]. The condition (*iii*) is meant to guarantee that one can get a physically meaningful solution from imperfect observations ; indeed, a sufficiently moderated noise implies a small departure of $\hat{\mathbf{x}}_\mathbf{b}$ from the ideal solution corresponding to $\mathbf{b} = 0$. However, $\mathbf{M}$ is in general non-invertible, and different situations can occur. If $m > \text{rank}(\mathbf{M})$ then this equation admits multiple solutions, whenever a solution exists. This is the case when $m > p$, for instance if $\mathbf{M}$ realizes a masking of the underlying signal. If $p > \text{rank}(\mathbf{M})$, then the equation $\mathbf{y} = \mathbf{M}\mathbf{x}$ doesn't necessarily admit a solution. The problem is said to be *ill-posed*. In general one is more concerned with the solutions multiplicity than with the strict invertibility of $\mathbf{M}$; indeed, it is not suitable to actually solve the equation $\mathbf{y} = \mathbf{M}\mathbf{x}$ because of the noise. Besides, even if $\mathbf{M}$ is perfectly invertible, estimating $\mathbf{x}$ might be tricky from a numerical point of view, depending on $\mathbf{M}$'s conditioning number, which is defined thereafter. The *condition number* of $\mathbf{M}$ is defined as

$$\kappa(\mathbf{M}) = |||\mathbf{M}|||\,|||\mathbf{M}^{-1}|||, \tag{II.9}$$

where $|||\cdot|||$ is a subordinate matrix norm to some vector norm $\|\cdot\|$ in $\mathbb{R}^m$. Assuming that $\mathbf{y}_1 = \mathbf{M}\mathbf{x}_1$ and $\mathbf{y}_2 = \mathbf{M}\mathbf{x}_2$, the following important relation holds :

$$\frac{\|\mathbf{x}_2 - \mathbf{x}_1\|}{\|\mathbf{x}_2\|} = \kappa(\mathbf{M})\frac{\|\mathbf{y}_2 - \mathbf{y}_1\|}{\|\mathbf{y}_2\|}. \tag{II.10}$$

If $\kappa(\mathbf{M})$ takes a large value, then a small change in the observation space can translate into a large change in the undistorted signal space, and the matrix $\mathbf{M}$ is then said to be *ill-conditionned*. This is typically the case in deconvolution problems where $\mathbf{M}$ represents a convolution with a low-pass filter. As illustrated in Fig. ***II.2***, even with a very small amount of noise, the signal estimated by direct inversion of $\mathbf{M}$ can dramatically depart from the truth, even if strictly speaking, the solution of $\mathbf{y} = \mathbf{M}\mathbf{x}$ continuously depends on the observation.

Ill-conditioning besides implies an amplified sensitivity to round-off errors due to the finite machine numerical precision.

Most inverse problems are ill-posed and/or ill-conditioned. Both cases are illustrated in Fig. ***II.3***.

This requires the practioner to regularize the problem which is making use of prior knowledge of the signal of interest in order to narrow down the domain of possible solutions and stabilize the inversion process. We present some classical regularizers in the next section.

A.3 Regularization

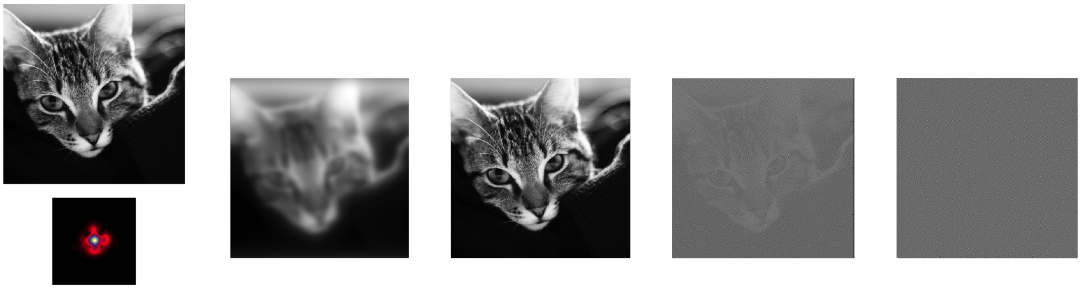


Figure *II.2* – Deconvolution by direct convolution kernel inversion ; from the left to the right : reference undistorted image and blurring kernel, blurred image, restored images by direct inversion in Fourier domain, for snr $= 10^{12}$, 10^{9} and 10^{6} respectively.

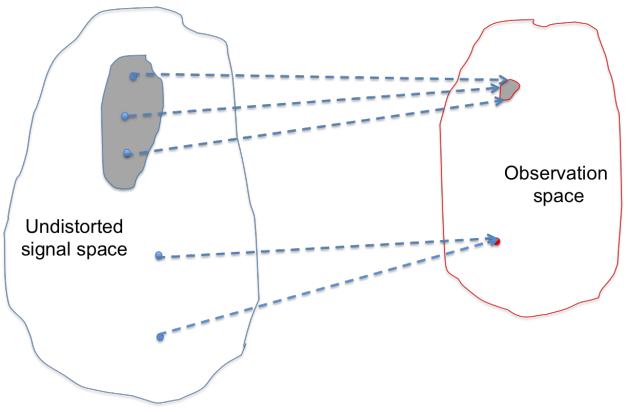


Figure *II.3* – Ill-posedness and ill-conditioning in inverse problems : different signals in the undistorted signals space can be mapped to the same observation, which implies ill-posedness ; more generally, significantly different signals in the undistorted signals space can be mapped to very close observations as a consequence of ill-conditioning.

A.3.1 Approximated problem

As precised in the previous section, given an observed signal $\mathbf{y}$, it is not suitable to exactly solve the equation $\mathbf{y} = \mathbf{M}\mathbf{x}$ because of noise and modeling imperfection. However one wants $\mathbf{y}$ and $\mathbf{M}\mathbf{x}$ to be close in some convenient sense that depends in general on the noise statistics. Precisely, the departure is measured as $\mathcal{C}(\mathbf{y}, \mathbf{M}\mathbf{x})$, where $\mathcal{C}$ is a positive cost function which verifies

$$\forall (\mathbf{u}, \mathbf{v}) \in \mathbb{R}^p \times \mathbb{R}^p \; \mathcal{C}(\mathbf{u}, \mathbf{v}) = 0 \Rightarrow \mathbf{u} = \mathbf{v}. \tag{II.11}$$

Thus a perfect solution can be recovered by minimizing $J(\mathbf{x}) = \mathcal{C}(\mathbf{Mx}, \mathbf{y})$ when the problem is well-posed.

Example A.1. — In numerous applications, the noise is assumed to be a realization of a gaussian random vector $\mathbf{B} \sim \mathcal{N}(0, \sigma_n \mathbf{I}_p)$. Thus, the observation $\mathbf{y}$ is a realization of a gaussian random vector $\mathbf{Y} \sim \mathcal{N}(\mathbf{Mx}, \sigma_n \mathbf{I}_p)$. Under this probabilistic model, the likelihood of $\mathbf{x}$ immediately gives one the popular quadratic cost $\mathcal{C}(\mathbf{y}, \mathbf{Mx}) = \frac{1}{2}\|\mathbf{y} - \mathbf{Mx}\|_2^2$. More generally, if $\mathbf{B} \sim \mathcal{N}(0, \mathbf{V})$ for an arbitrary positive definite correlation matrix $\mathbf{V}$, then the log-likelihood derived cost function is $\mathcal{C}(\mathbf{y}, \mathbf{Mx}) = \frac{1}{2}(\mathbf{y} - \mathbf{Mx})^T \mathbf{V}^{-1}(\mathbf{y} - \mathbf{Mx})$.

— If additionally the observation contains spurious artefacts (for example cosmic rays impact print in astronomical images), it is suitable to require a weaker data fidelity on the largest observation entry. This leads one to outliers-robust cost functions such Huber loss [Chen *et al.* 2014] :

$$\mathcal{C}(\mathbf{y}, \mathbf{Mx}) = \begin{cases} \frac{1}{2}\|\mathbf{y} - \mathbf{Mx}\|_2^2 \text{ if } \|\mathbf{y} - \mathbf{Mx}\|_2^2 \leq \beta, \\ \beta\|\mathbf{y} - \mathbf{Mx}\|_1 - \frac{\beta^2}{2} \text{ otherwise.} \end{cases} \tag{II.12}$$

— Often, digital images can be corrupted by an additive impulsive noise due to malfunctioning optical sensors, transmission errors, etc. Such noise can be model as a realization of a laplacian random vector, yielding an ℓ_1 cost : $\mathcal{C}(\mathbf{y}, \mathbf{Mx}) = \|\mathbf{y} - \mathbf{Mx}\|_1$ [Yang *et al.* 2009].

— Similarly, if the signal is dominated with the so-called Poissonian noise, the natural cost function in a bayesian sense is the Kullback-Leibler divergence : $\mathcal{C}(\mathbf{y}, \mathbf{Mx}) = \sum_{i=1}^{p} \mathbf{y}[i] \log(\frac{[\mathbf{Mx}][i]}{\mathbf{y}[i]})$ (see for example [Dupé *et al.* 2011]) ; this cost is in general suitable when handling positive data.

We use a quadratic cost in the chapters IV and V. However astronomical images are actually corrupted with a mixed Poisson-Gaussian noise. Still, the corresponding likelihood-derived cost can be readily integrated into the proposed restoration schemes [Jezierska *et al.* 2013]. Note that even if the function $\mathcal{C}$ is a strictly convex function, the cost $\mathcal{C}(\mathbf{y}, \mathbf{Mx})$ might have multiple global minimum in $\mathbf{x}$ if the problem is ill-posed. Since then, the inversion task consists in selecting a solution in the vicinity of one of a minimizer of the cost, based on *prior information* on the signal to be estimated. This is detailed in the following section.

A.3.2 Examples of regularizers

Tikhonov regularization One of the simplest information that can be used in an inverse problem is that the restored signal's energy, measured by

its squared ℓ_2 norm, can not be too large. This can be done by solving the following optimization problem :

$$\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y} - \mathbf{M}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_2^2, \tag{II.13}$$

where the regularization parameter λ balances the strength of the constraint on the solution's energy against the accuracy of the data fit ; intuitively, one can anticipate that the practical choice of λ will be tight to the noise level. This is a particular instance of the widely used *Tikhonov regularization* method. Besides, if $\mathbf{M}$ represents a circular convolution, it can be shown that the solution of II.13 is the optimal estimate of $\mathbf{x}$ in the sense of Wiener [Levinson 1946] for the particular choice $\lambda = \mathrm{SNR}^{-1}$.

Problem II.13 can be turned into the more generic following form,

$$\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y} - \mathbf{M}\mathbf{x}\|_2^2 + \lambda\|\mathbf{\Gamma}\mathbf{x}\|_2^2, \tag{II.14}$$

where the matrix $\mathbf{\Gamma}$ allows for making use of diverse a priori. This regularization method became particularly popular due to its application to integral equations by the mathematician and geophysicist Andrey Nikolayevich Tikhonov [Tikhonov & Arsenin 1977]. Typical choices for $\mathbf{\Gamma}$ are finite difference approximations of gradient and laplacian operators which in both cases allows for limiting high frequencies energy in the restoration process. Problem *II*.14 can be interpreted in a Bayesian framework. Indeed, assuming that the unobserved signal $\mathbf{x}$ is a realization of a random vector $\mathbf{X}$ verifying

$$\mathbf{\Gamma}\mathbf{X} \equiv \mathcal{N}(0, \sigma_s \mathbf{I}_p), \tag{II.15}$$

then problem *II*.14 is equivalent to a *Maximum a Posteriori* (MAP) estimation of $\mathbf{x}$, for $\lambda = \frac{\sigma_n^2}{\sigma_s^2}$. We recall the probabilistic model of the noise : $\mathbf{B} \sim \mathcal{N}(0, \sigma_n \mathbf{I}_p)$. If $\mathbf{\Gamma}$ is a discrete approximation of a gradient operator, then II.15 reads as a brownian motion-like model on the latent signal [Karl 2005], which clearly misses local correlations inherent to natural images for instance. As anticipated in the case $\mathbf{\Gamma} = \mathbf{I}_p$, Tikhonov regularization yields a signal estimate which depends linearly upon the observation. This is convenient since the uncertainty on the estimate is straightforwardly characterized, which is particularly suitable in cosmology for large scale statistical studies [Bobin *et al.* 2015]. However, high frequency informative content and noise are identically attenuated (or preserved), yielding overly smoothed (or noisy) solutions. This limitation motivated the study of non quadratic regularizers. We consider the general formulation of the estimation problem

$$\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y} - \mathbf{M}\mathbf{x}\|_2^2 + \lambda \mathrm{R}(\mathbf{x}), \tag{II.16}$$

Tikhonov regularization being a particular instance.

Some non quadratic regularizers One of the most widely used non quadratic regularizer is certainly the entropy defined as

$$\mathrm{H}(\mathbf{x}) = -\sum_{i=1}^{p} \mathbf{x}[i] \log(\mathbf{x}[i]), \tag{II.17}$$

where $\mathbf{x}$ is assumed to be a discrete probability distribution. The choice $\mathrm{R}(\mathbf{x}) = -\mathrm{H}(\mathbf{x})$ is particularly sound when the only prior available is the positivity (see [Dominikus 1997] and the references therein). Indeed, maximizing the entropy yields a solution with less structural information or equivalently it enforces pixels values similarity. According to Shannon's interpretation of the entropy from the information theory perspective, this yields the most "natural" solution, whenever the positivity prior applies. Although this constraint implies a smoothness of some sort, it doesn't limit the frequency band of the signal to be restored, making sharp details restoration in principle possible, while smoothing the noise out [Frieden 1972]. A thorough presentation of entropy based regularization for image processing can be found in [Starck & Murtagh 2006]. However, by construction, this approach doesn't specifically account for natural signal or images properties, among which the presence of discontinuities.

This leads us to the important total variation regularization. Firstly used for image denoising in the foundationnal work by Rudin, Osher and Fatemi [Rudin *et al.* 1992], total variation based restoration has recently known a renewed interest in the signal processing community due to theoretical and algorithmic advances ; see [Chan *et al.* 2006] and [Chambolle *et al.* 2010] for comprehensive reviews. Total variation has been defined and theoretically analyzed mostly in a continuous setting ; however the following discrete approximation is commonly used :

$$\mathrm{R}(\mathbf{x}) = \|\mathbf{x}\|_{\mathrm{TV}} = \sum_{i=1}^{p} \|[\nabla \mathbf{x}]_i\|_2, \tag{II.18}$$

where $[\nabla \mathbf{x}]_i$ is the finite difference approximation of the gradient of the continuous signal underlying $\mathbf{x}$ at $\mathbf{x}[i]$. If $\mathbf{x}$ represents a 1D signal, $[\nabla \mathbf{x}]_i$ is a scalar and the total variation closely resembles a Tikhonov regularizer with a Tikhonov matrix chosen a discrete gradient operator. However, one can get a good intuition of the key difference between the two by taking a bayesian standpoint. Indeed, one can interpret the total variation constraint as following from a probabilistic Laplace model on the gradient distribution of the underlying signal in the 1D case. Thereby, all else being equal, strong gradient values are more likely for this model than for the gaussian model associated with a Tikhonov regularization. Thus, sharp discontinuities can be allowed

in the MAP solution while spurious variations associated with noise remain penalized.

As we will see in the next section, the discrete total variation can be seen as a particular instance of a larger class of non-smooth regularizers relying on deterministic low complexity priors on the signals. We complete these examples with ℓ_∞ norm based regularization. Indeed, in certain applications mostly related to digital communications, one might want to recover a flat signal or correlatively, compute and transmit a "democratic" or "anti-sparse" representation of a given signal for noise resilience and efficient amplifiers sizing (see the references and examples in [Elvira *et al.* 2015]). This is done by imposing a sufficiently low maximal amplitude bound on the signal to be computed, which can be achieved by choosing $\mathrm{R}(\mathbf{x}) = \|\mathbf{x}\|_\infty$. The penalized regressions presented so far can be interpreted as MAP estimations yielding the so called bayesian methods ; in a statistical framework, the problem of choosing the regularization parameter is exchanged with that of modeling the uncertainties on the observed and underlying signals (see [Wong *et al.* 2015] and the references therein). However, let underline the MAP interpretation might be misleading because

— the MAP estimators do not follow in general the probability distribution used as a prior [Nikolova 2007] and

— their might be other valid bayesian interpretations of the MAP estimators [Gribonval 2011].

Implicit regularizers We now give examples of regularization schemes that unlike the methods presented so far do not calculate the minimum of an explicitly defined functional. We first consider the Truncated Singular Value Decomposition (TSVD) method [Hansen 1987]. We recall that the observation matrix $\mathbf{M} \in \mathrm{M}_{pm}(\mathbb{R})$ (Eq. II.8). The SVD of $\mathbf{M}$ is given by following factorization :

$$\mathbf{M} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H \tag{II.19}$$

where the matrices $\mathbf{U} \in \mathrm{M}_p(\mathbb{C})$ and $\mathbf{V} \in \mathrm{M}_m(\mathbb{C})$ are unitary and $\mathbf{\Sigma}$ is a (rectangular) diagonal matrices which diagonal elements are noted $\sigma_1 \geq \cdots \geq \sigma_{\min(p,m)} \geq 0$. The scalars σ_i are called the singular values of $\mathbf{M}$ and somehow generalizes the notion of eigenvalues to non-square matrices. Thus,

$$\mathbf{M} = \sum_{i=1}^{t} \sigma_i \mathbf{U}[:, i]\mathbf{V}[:, i]^H, \tag{II.20}$$

where $t = \min(p, m)$. We note t^+ the index of the smallest non-zero singular value of $\mathbf{M}$. The vector

$$\hat{\mathbf{x}} = \sum_{i=1}^{t^+} \frac{\mathbf{U}[:, i]^H \mathbf{y}}{\sigma_i} \mathbf{V}[:, i] \tag{II.21}$$

is the solution of the optimization problem

$$\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{x}\|_2^2 \text{ s.t. } \mathbf{x} = \underset{\mathbf{t}}{\operatorname{argmin}} \frac{1}{2}\|\mathbf{y} - \mathbf{M}\mathbf{t}\|_2^2 \tag{II.22}$$

In words, $\hat{\mathbf{x}}$ is the minimal energy minimizer of the functional $J(\mathbf{x}) = \frac{1}{2}\|\mathbf{y}-\mathbf{M}\mathbf{x}\|_2^2$; $\hat{\mathbf{x}}$ is therefore the unique solution when the problem is well-posed. If $\mathbf{M}$ represents a circular convolution, then $\mathbf{V}^H$ is nothing but the matrix representation of the Discrete Fourier Transform (DFT) and $\mathbf{U} = \mathbf{V}$ [Gray 2005]. Thereby, $\mathbf{U}$'s columns are discretized complex exponentials at different frequencies. Yet, in deconvolution problems as previously illustrated by Fig. ***II.2***, the matrix $\mathbf{M}$ is typically ill-conditioned and the lowest singular values are those associated with the highest frequencies. Therefore the high frequencies content of $\mathbf{y}$, which is in general dominated by noise, is amplified in $\hat{\mathbf{x}}$. The TSVD then simply consists in dumping the terms in Eq.II.21 associated with the lowest singular values :

$$\hat{\mathbf{x}}_{\text{TSVD}} = \sum_{i=1}^{t_{\text{SVD}}} \frac{\mathbf{U}[:, i]^H \mathbf{y}}{\sigma_i} \mathbf{V}[:, i], \tag{II.23}$$

for some user defined integer $t_{\text{SVD}} < t^+$. The conditions under which the TSVD yields similar results as the Tikhonov regularization in the case $\mathbf{\Gamma} = \mathbf{I}_p$ are studied in [Hansen 1987]. One can note that this method makes no attempt to reconstructing the unobserved components of the image i.e. the projection of $\mathbf{x}$ on the null space of $\mathbf{M}$.

Our second example is referred to as the *early stopping* in computational statistics and the machine learning literature (see [Raskutti *et al.* 2014] and the references therein). It relies on signals restoration using iterative schemes. Indeed, as shown in the previous examples, solving inverse problems is in general formulated as finding a stationary point of a certain cost function. While in several interesting cases there is no closed-form expression to compute the stationary points, in most cases, it can be done iteratively. We review some families of optimization methods in a forthcoming section. Iterative schemes generate a sequence $(\mathbf{x}_k)$ that converges in some sense toward a stationary point $\mathbf{x}^\star$ of the (regularized) cost function, up to the computer's numerical precision. The restored signal is given by $\hat{\mathbf{x}} = \mathbf{x}_{k_{\max}}$, where $k_{\max}$ is set *in situ*,

in general based on a numerical criterion of convergence. If the problem is not regularized, early iterates capture the low frequency content of the observed signal whereas as k increases, the iterates get corrupted by noise. Hence, the number of iterations itself behaves as a regularization parameter, which is the whole point of *early stopping*. This regularizing effect can be precisely characterized for the simple Landweber iterations scheme [Landweber 1951] which uses in the update formula :

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mu \mathbf{M}^T(\mathbf{M}\mathbf{x}_k - \mathbf{y}), \tag{II.24}$$

where $k \in \mathbb{N}$, $\mu \in]0, 2/\sigma_1^2[$, σ_1 being the highest singular value of $\mathbf{M}$. If $\mathbf{x}_0 = 0$, then one can verify that the k^{th} iterate is given by [Karl 2005]

$$\mathbf{x}_k = \sum_{i=1}^{t^+} (1 - (1 - \mu * \sigma_i^2)^k) \frac{\mathbf{U}[:, i]^H \mathbf{y}}{\sigma_i} \mathbf{V}[:, i], \tag{II.25}$$

where t^+ has been previously defined as the index of the smallest non-zero singular value of $\mathbf{M}$. Thus, the Landweber scheme effectively "re-weights" the un-regularized problem minimal energy solution given in Eq. II.21 with a set of varying weights identified as

$$w_{i,k} = 1 - (1 - \mu * \sigma_i^2)^k, \tag{II.26}$$

for $i \in [\![1, t^+]\!]$. The weights variations as a function of the singular values is illustrated in Fig. ***II.4***, for different numbers of iterations. For $k > 1$, these functions have a step-like shape with an inflection point localized at

$$\sigma^\star = \frac{1}{\sqrt{\mu(2 * k - 1)}}. \tag{II.27}$$

Loosely speaking, the features associated with the singular values smaller than $\sigma^\star$ are penalized while the features associated with the singular values larger than $\sigma^\star$ are kept. Thus, the number of iterations plays the role of the inverse of the regularization parameter in the previously described methods.

When the matrix $\mathbf{M}^T\mathbf{M}$ is invertible, the unregularized problem has a unique optimum, solution of the positive definite system

$$\mathbf{M}^T\mathbf{M}\mathbf{x} = \mathbf{M}^T\mathbf{y}. \tag{II.28}$$

Thus it can be calculated using the conjugate gradient (CG) algorithm [Hestenes & Stiefel 1952] which typically requires less iterations than the Landweber scheme. It has been shown that the k^{th} iterate generated by the CG is the solution of the problem

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{M}\mathbf{x}\|_2^2 \text{ s.t. } \mathbf{x} \in \mathcal{K}_k(\mathbf{M}^T\mathbf{M}, \mathbf{M}^T\mathbf{y}) \tag{II.29}$$

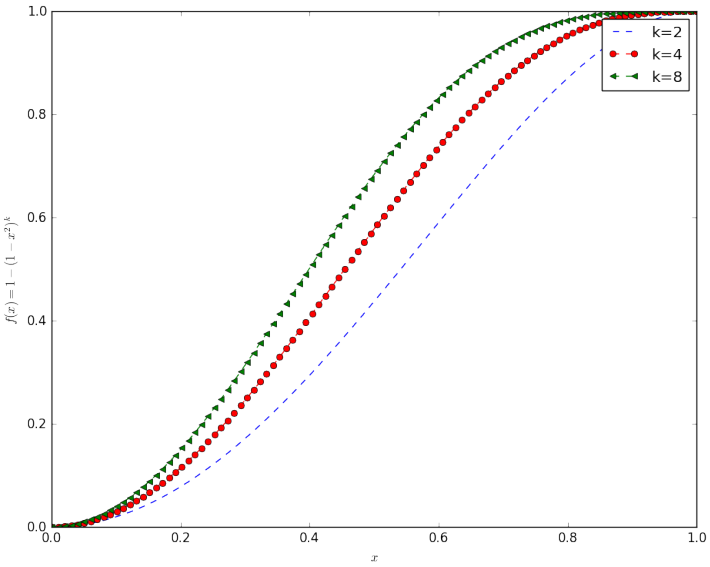


Figure *II.4* – Curve of the weight in Eq.II.26 as a function of the singular values, for different number of iterations.

where $\mathcal{K}_k(\mathbf{M}^T\mathbf{M}, \mathbf{M}^T\mathbf{y}) = \text{span}(\{\mathbf{M}^T\mathbf{y}, (\mathbf{M}^T\mathbf{M})\mathbf{M}^T\mathbf{y}, \cdots, (\mathbf{M}^T\mathbf{M})^{k-1}\mathbf{M}^T\mathbf{y}\})$ is the Krylov subspace generated by the matrix $\mathbf{M}^T\mathbf{M}$ and the vector $\mathbf{M}^T\mathbf{y}$ [Freund 1992]. Again, an early stopping in the CG method regularizes the inverse problem and the regularization weakens when the number of iterations (and consequently the dimension of the Krylov subspace) increases.

We finally note with no formal description that the early stopping is often used in the celebrated Richardson-Lucy to mimic regularization [Prato *et al.* 2012].

B Low complexity priors

In this section, we focus on the low complexity priors based regularizers. The general idea is that the estimated signal only needs a few number of parameters w.r.t. the ambient dimension to be fully specified, which can be used to recover well-posedness. In particular, we consider linear low complexity models [Baraniuk & Wakin 2009] which assume that the signal of interest can be explicitly written as a linear combination of few elements from a predefined or adaptive collection of atoms that will be referred to as a dictionary, following the terminology introduced in [Mallat & Zhang 1993]. This is described explicitly in the forthcoming sections.

B.1 Sparse models

B.1.1 Definition

A vector $\mathbf{x} \in \mathbb{R}^m$ is *strictly sparse* if most of its entries are equal to zero :

$$\|\mathbf{x}\|_0 \stackrel{\text{def}}{=} \text{card}(\text{supp}(\mathbf{x})) \ll m. \tag{II.30}$$

$\mathbf{x}$ is said to be *k-sparse* if $\|\mathbf{x}\|_0 = k$.

Sparse signals appear in various applications. For instance in geophysics, the underground reflectivity can be modeled as a sparse signal (see [Claerbout & Muir 1973]). One can also think of acoustics and especially Room Impulse Response modeling [Sturm & Defrance 2010], and biological signal processing (for instance electromyography signals [Liu *et al.* 2012]).

However, natural and especially physical signals are not strictly sparse in general. Yet, they might be *compressible* in the sense that most of their energy is captured by a few entries. This can be formalized in terms of the entries magnitudes decay's rate : assuming that σ_m is a permutation in $[\![1, m]\!]$ so that $|\mathbf{x}[\sigma(1)]| \geq \cdots \geq \mathbf{x}[\sigma(m)]|$, $\mathbf{x}$ is compressible or *weakly sparse* if $\forall k \in [\![1, m]\!]$,

$$|\mathbf{x}[\sigma(k)]| \leq Rk^{-s}, \tag{II.31}$$

for some strictly positive reals R and s.

We note $\mathbf{x}^{(k)}$ the best k-sparse approximation of $\mathbf{x}$, for some $k < m$. $\mathbf{x}^{(k)}$ is obtained by setting the $m-k$ lowest amplitude entries of $\mathbf{x}$ to 0 so that the approximation error verifies

$$\mathcal{E}_{\mathbf{x},k} \stackrel{\text{def}}{=} \|\mathbf{x} - \mathbf{x}^{(k)}\|_2^2 = \sum_{i=k+1}^{m} \mathbf{x}[\sigma(i)]^2 \leq R^2 \frac{k^{-2s+1}}{2s-1}, \tag{II.32}$$

if $s > 1/2$. Thus, for a given level of accuracy, the higher is s, the lower k can be chosen and therefore the more compressible is the signal.

The definition above describes signals which are approximately sparse in their *direct domain* of representation, or more formally, in the canonical basis of $\mathbb{R}^m$. Yet, structured signals such as piece-wise smooth or oscillatory ones can be represented in alternative basis which better takes into account their specificity, yielding a faster decay of the sorted representation coefficients amplitudes. This is illustrated in the figures ***II.5*** and ***II.6*** for a gravitational waves signal (see [Abbott *et al.* 2016] for more details).

Therefore, more generally a signal is compressible if it can be written as

$$\mathbf{x} = \sum_{i=1}^{K} a_i \mathbf{d}_i, \tag{II.33}$$

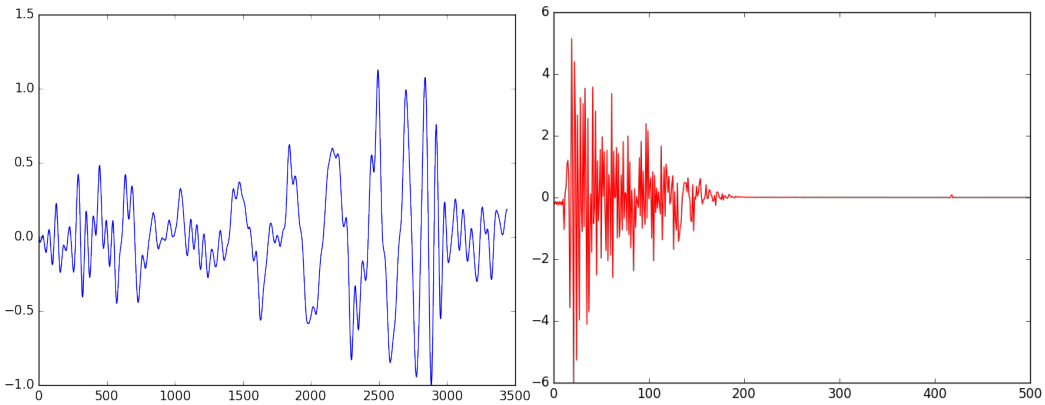


Figure *II.5* – On the left, the first detected gravitational waves signal, measured by the LIGO Hanford (credit : www.ligo.org) ; the x axis is the time and the y axis is the *gravitational waves strain*, which quantifies the space-time distortion. On the right, the DCT of this signal (truncated to the first 500 hundred samples for clarity, the discarded values being negligible).

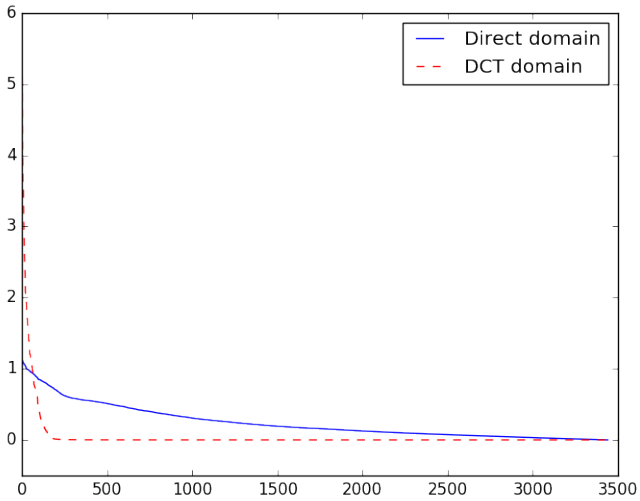


Figure *II.6* – The sorted magnitude of Fig. ***II.5***'s signals. The gravitational is clearly sparser in the sense of Eq.II.31 in the DCT domain, although more targeted transforms might have been used (see for example the chirplet transform in [Mann & Haykin 1995]).

or compactly,

$$\mathbf{x} = \mathbf{D}\mathbf{a}, \tag{II.34}$$

for some dictionary $\mathbf{D} \overset{\text{def}}{=} [\mathbf{d}_1, \cdots, \mathbf{d}_K]$ defined independently of $\mathbf{x}$, where the

vector $\mathbf{a} \overset{\text{def}}{=} (a_i)_{1 \leq i \leq K}$ meets the condition of Eq.II.31. The *direct domain* sparsity corresponds to $\mathbf{D}$ chosen as the canonical basis of $\mathbb{R}^m$.

B.1.2 Promoting sparsity

Now we turn to the linear inverse problem of Eq.II.8 and we assume that the signal $\mathbf{x}$ is at least weakly sparse, w.r.t. some dictionary $\mathbf{D}$. This prior can be naturally taken into account through the following constrained optimization problem :

$$\min_{\mathbf{a}} \frac{1}{2} \|\mathbf{y} - \mathbf{MDa}\|_2^2 \text{ s.t. } \|\mathbf{a}\|_0 \leq \beta, \tag{II.35}$$

where β sets the minimal level of sparsity required. Let consider the set of admissible solutions, often referred to as the *feasible set* in numerical optimization literature :

$$\mathcal{B}^0_{\mathbb{R}^K}(0, \beta) \overset{\text{def}}{=} \{\mathbf{a} \in \mathbb{R}^K / \|\mathbf{a}\|_0 \leq \beta\}. \tag{II.36}$$

Assuming that $\beta \geq 1$, it can be decomposed as follows :

$$\mathcal{B}^0_{\mathbb{R}^K}(0, \beta) = \bigcup_{k=1}^{\lfloor \beta \rfloor} \bigcup_{J \in \mathcal{P}([\![1,K]\!])/\text{card}(J)=k} \text{span}(\{\mathbf{d}_i\}_{i \in J}), \tag{II.37}$$

where for a set E, $\mathcal{P}(E)$ is the set of subsets of E. The structure of this set shows how difficult the resolution of the problem II.35 might be :

— $\mathcal{B}^0_{\mathbb{R}^K}(0, \beta)$ is non-convex because of the denumerable unions, while *convexity* is crucial in solving efficiently optimization problems ; the interested reader might have a look at [Boyd & Vandenberghe 2004] ;

— as K increases, the number of subsets in $\mathcal{B}^0_{\mathbb{R}^K}(0, \beta)$ rapidly becomes too large for a brute force global resolution of the problem II.35 to be tractable.

Thus, the ℓ_1 norm is generally used as a *convex relaxation* of the ℓ_0 pseudo-norm, which have been justified in different ways. The most natural argument is based on the simple remark that the unitary ℓ_1 ball $\mathcal{B}^1_{\mathbb{R}^K}(0, 1)$ is the *convex hull* of the intersection of the unitary ℓ_0 ball $\mathcal{B}^0_{\mathbb{R}^K}(0, 1)$ with the unitary ℓ_∞ ball $\mathcal{B}^\infty_{\mathbb{R}^K}(0, 1)$. In [Chen *et al.* 2001], Linear Programming (LP) arguments are used to support the fact that ℓ_1 minimization promotes sparsity. Thereafter, outstanding results are established showing that if the underlying signal is sufficiently sparse (or compressible), ℓ_1 norm minimization can yield an exact or stable recovery [Elad & Bruckstein 2002, Donoho 2006, Candès *et al.* 2006]. More recently, a generic and formal framework has been introduced in [Chandrasekaran *et al.* 2010] for translating simplicity priors into convex penalties ;

in particular, the ℓ_1 norm emerges in this framework as the *atomic norm* associated with the set of 1-sparse unitary vectors in the sparse signal's space.

Hence, we consider the optimization problem

$$\min_{\mathbf{a}} \frac{1}{2}\|\mathbf{y} - \mathbf{MDa}\|_2^2 \text{ s.t. } \|\mathbf{a}\|_1 \leq \beta, \tag{II.38}$$

or equivalently in its augmented lagrangian form

$$\min_{\mathbf{a}} \frac{1}{2}\|\mathbf{y} - \mathbf{MDa}\|_2^2 + \lambda\|\mathbf{a}\|_1 \tag{II.39}$$

for some well-chosen λ [Ciak *et al.* 2013]. These problems can be solved efficiently.

In certain applications such as DNA microarray analysis [Schena *et al.* 1995], the signal of interest is known to have groups of entries which are correlatively significant or negligible. The paradigms of group or block sparsity have been introduced for taking into account this particular prior [Huang *et al.* 2010, Peyré & Fadili 2011]. Specifically, one needs to define a partition $\{G_1, \cdots, G_q\}$ of $[\![1, K]\!]$ which we recall requires that $\bigcup_{i=1}^{q} G_i = [\![1, K]\!]$ and $G_i \bigcup G_j$ for $i \neq j$. The *ell*$_1$ norm is then replaced by the $\ell_{1,2}$ norm defined as

$$\|\mathbf{a}\|_{1,2} \stackrel{\text{def}}{=} \sum_{i=1}^{q} \|\mathbf{a}^{(G_i)}\|_2, \tag{II.40}$$

where $\mathbf{a}^{(G_i)}$ is the vector obtained by extracting $\mathbf{a}$'s entries indexed by G_i.

Remark : note that the earliest sparse recovery attempts consisted in solving approximately the ℓ_0 constrained problem II.35 via greedy methods, namely the celebrated Matching Pursuit [Mallat & Zhang 1993] and the Orthogonal Matching Pursuit [Pati *et al.* 1993]. Several methods have been proposed since then for ℓ_0 minimization, like for instance the Iterative Hard Thresholding (IHT) (see for example [Starck *et al.* 2004, Starck *et al.* 2005]) ; the IHT is thoroughly analyzed in [Blumensath & Davies 2008] and [Blumensath & Davies 2009]. Let finally note a more recent trend which consists in approximating the ℓ_0 pseudo-norm with the so called ℓ_p norms, with $0 < p < 1$ [Zheng *et al.* 2015].

B.1.3 Analysis formulation of the sparsity prior

According to the model II.33, the coefficients computed by solving II.39 enable one to (approximately) *synthesize* the underlying signal $\mathbf{x}$ using the dictionary $\mathbf{D}$ atoms. Hence the terminology of *synthesis formulation* used

when $\mathbf{D} \neq \mathbf{I}_p$. Alternatively, one can adopt the *analysis formulation* which has the following form :

$$\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y} - \mathbf{M}\mathbf{x}\|_2^2 + \lambda\|\mathbf{D}_A^T\mathbf{x}\|_1, \tag{II.41}$$

for some dictionary $\mathbf{D}_A^T$. In this case, one looks for a signal compatible with the data and strongly *correlated* with only a few atoms in $\mathbf{D}_A$. Although very similar, the analysis and synthesis formulation are only equivalent if the dictionary involved is square and invertible [Elad *et al.* 2007]. Precisely, if $\mathbf{D}_A^T = \mathbf{D}^{-1}$, then a vector $\mathbf{x}^\star$ is the solution of Problem II.41 if and only if $\mathbf{D}_A^T\mathbf{x}^\star$ is the solution of Problem II.39. This is crucial from applications standpoint because of the importance of overcomplete dictionaries in sparse recovery as we will shortly detail. Let first mention that the total variation based regularization falls into this framework in the 1D case with $\mathbf{D}_A$ chosen as the transpose of the matrix representing of the finite difference approximation of the gradient. Therefore, this penalty is particularly suitable for (approximately) piece-wise constant signals or equivalently signals that have a sparse gradient. The total variation in 2D can also be expressed in the analysis framework, but involving an $\ell_{1,2}$ norm (see Equation II.40).

In this example, the analysis formulation is weaker that the synthesis one in a certain sense. Indeed, when the dictionary $\mathbf{D}_A$ has fewer atoms than the underlying signal's space dimension, the analysis enables one to promote some regularity on the signal without completely specifying its structure. Precisely, the signal is only imposed to be orthogonal to certain columns of $\mathbf{D}_A$ whereas in synthesis, the signal is explicitly described as a linear combination of some of the dictionary's atoms.

B.1.4 The dictionary

The dictionary choice is in first instance driven by the signal's structural specificity. Indeed, the signal can only be sparsely encoded in the dictionary if some of its atoms closely resemble the signal main features. Let give a few classical examples :

— the DCT dictionary is suitable for oscillatory signals as already illustrated ;

— standard wavelet bases are optimal for encoding signals with isolated singularities [Mallat 1999] ;

— in presence of "extended" singularities with geometrical regularity such as edges, the curvelet dictionary are more adapted [Candès & Donoho 2004].

Because of the tremendous morphological diversity of natural signals and the success of sparse recovery in various applications (see [Zhang *et al.* 2015] and the references therein), the task of designing increasingly adapted transforms and dictionaries has considerably polarized the recent signal processing literature. This is well illustrated by [Jacques *et al.* 2011] and [Duval 016]. One can note that a particular focus has been put on the so-called *multiscale* representations. Indeed, characteristic features can appear at multiple scales, which in general are not known beforehand. Similarly, *translation invariance* is an important dictionary feature [Coifman & Donoho 1995]. Besides images generally contain structurally different features, so that each "structural" component of the image can be sparsely represented in a different dictionary. In this case, a better dictionary, from the sparse recovery standpoint, is obtained by concatenation. This is the basic rational behind the Morphological Component Analysis [Starck *et al.* 2004] and its different extensions.

More generally, geometrical invariances requirement and transforms combination yield *overcomplete* dictionaries, which have been proven very powerful in sparse signal recovery [Starck *et al.* 2015]. We now recall two theoretical results which are important to bear in mind when building or choosing a dictionary. The first one concerns the uniqueness of sparse representations in overcomplete dictionaries. We consider two orthonormal bases $\mathbf{D}_1 = [\mathbf{d}_{11}, \cdots, \mathbf{d}_{1p}]$ and $\mathbf{D}_2 = [\mathbf{d}_{21}, \cdots, \mathbf{d}_{2p}]$, and we consider a vector $\mathbf{x}$ in $\mathbb{R}^m$ we note $\boldsymbol{\alpha}$ the sparsest vector verifying

$$\mathbf{x} = [\mathbf{D}_1, \mathbf{D}_2]\boldsymbol{\alpha}. \tag{II.42}$$

It is shown in [Elad & Bruckstein 2002] that if

$$\|\boldsymbol{\alpha}\|_0 \leq \frac{1}{2}(1 + \frac{1}{\mu(\mathbf{D}_1, \mathbf{D}_2)}), \tag{II.43}$$

then $\boldsymbol{\alpha}$ is unique and can be calculated via an ℓ_1 minimization, $\mu(\mathbf{D}_1, \mathbf{D}_2) = \max_{1 \leq i,j \leq p} \{\mathbf{d}_{2i}^T \mathbf{d}_{1j}\}$ being the so-called *mutual coherence* of the basis $\mathbf{D}_1$ and $\mathbf{D}_2$. This suggests that in general, having strongly correlated atoms withing the dictionary is susceptible to undermine the (ℓ_1 norm based) sparse recovery. The second result gives an indication on the dictionary choice, with respect to the observation operator in Problem II.39. Precisely, we assume, among other omitted hypothesis, that there exists an $m \times m - p$ matrix $\mathbf{U}$ so that $\hat{\mathbf{M}} = [\mathbf{M}^T, \mathbf{U}]^T$ is an orthogonal matrix. Then, the smaller is $\mu(\hat{\mathbf{M}}^T, \mathbf{D})$, the less observation are needed for achieving a given level of accuracy in the sparse recovery, in probability [Candès & Romberg 2007].

In words, the more $\mathbf{D}$ atoms have spread representations w.r.t. $\hat{\mathbf{M}}^T$, the more the recovery is robust to under-determination.

Dictionary learning So far, we have mentioned dictionaries derived analytically and which are in general optimal for representing certain formal classes of signals. However, in certain applications, one disposes of data samples that are structurally similar to the signal. Hence the natural idea of learning the dictionary directly from the data. Precisely, given a training set $\mathbf{x}_1, \cdots, \mathbf{x}_q$, the core idea is to simultaneously calculate a dictionary $\mathbf{D}_L$ and a set of sparse vectors $\mathbf{a}_1, \cdots, \mathbf{a}_q$ so that

$$\mathbf{x}_i \approx \mathbf{D}_L \mathbf{a}_i, \tag{II.44}$$

$\forall i \in [\![1, q]\!]$. Dictionary learning has been successfully applied in various applications like for instance in astronomy [Beckouche *et al.* 2013b], for Blind Source Separation [Bobin *et al.* 2007], neuroscience [Varoquaux *et al.* 2011] and more generally natural images restoration [Elad & Aharon 2006], [Mairal *et al.* 2008]. It seems worth noting that while most dictionary learning methods focus on the synthesis formulation of sparsity, a few work exists that are based on the analysis model (see [Rubinstein *et al.* 2013] and the references therein). In both cases, the dictionaries learned are in general unstructured so that they have to be fully stored and explicitly used in matrix-vector products. This is an important consideration from a practical point of view since non-adaptive dictionaries such Wavelets, DCT or Curvelets dictionaries are associated with fast implicit implementations, which makes them computationally efficient.

B.2 Low rankness

In this section, the signal of interest is represented by an $m \times q$ matrix $\mathbf{X} = [\mathbf{x}_1, \cdots, \mathbf{x}_q]$ which is convenient in various applications such as video or multichannel signal processing. The observation model is adapted as follows :

$$\mathbf{Y} = \mathcal{M}(\mathbf{X}) + \mathbf{B}, \tag{II.45}$$

where $\mathcal{M}$ is a linear operator and $\mathbf{Y} \in \mathrm{M}_{pq}(\mathbb{R})$.

B.2.1 Definition

The (algebraic) rank is defined as follows :

$$\mathrm{rank}(\mathbf{X}) \overset{\text{def}}{=} \dim(\mathrm{span}((\mathbf{x_i})_{1 \leq i \leq q})) = \mathrm{rank}(\mathbf{X}^T). \tag{II.46}$$

$\mathbf{X}$ will be said to have a low rank if

$$\mathrm{rank}(\mathbf{X}) \ll \min(m, q). \tag{II.47}$$

To get an intuition of the practical interest of this quantity, let assume that $\mathbf{X}$ is an hyperspectral image, with the convention that the i^{th} column $\mathbf{x}_i$

corresponds to the i^{th} band image. Each line of $\mathbf{X}$ can be modeled as a linear combination of deterministic spectra related to the materials present in the scene [Chang 2003]. Thus, if the scene contains exactly s different materials, then the rank of $\mathbf{X}$ is at most equal to s; this is particularly interesting when only a few materials are preponderant in the scene. As illustrated by this example, in general, constraining the rank allows one to account for the redundancies present in the data to be restored.

We note $\mathbf{a}(\mathbf{X}) = (a_i)_{1\leq i\leq t}$ the singular values of $\mathbf{X}$, with $t = \min(m, q)$. Then we have

$$\operatorname{rank}(\mathbf{X}) = \|\mathbf{a}(\mathbf{X})\|_0. \tag{II.48}$$

Recalling that $\mathbf{X}$ can be written in the form

$$\mathbf{X} = \sum_{i=1}^{t} a_i \mathbf{U}_i, \tag{II.49}$$

where the $\mathbf{U}_i$s are rank one matrices, the low rankness clearly appears as an extension of the notion of compressibility to matrices. Precisely, the rank quantifies the complexity of a linear model that would perfectly fit the data stored in the considered matrix.

B.2.2 Constraining the rank

As seen in the previous section for the ℓ_0 norm, the direct problem

$$\min_{\mathbf{X}} \|\mathbf{Y} - \mathcal{M}(\mathbf{X})\|_F^2 \text{ s.t. } \operatorname{rank}(\mathbf{X}) \leq \beta \tag{II.50}$$

is combinatorial and NP-hard [Fazel 2002]. Thus, various heuristics have been proposed for promoting low rankness.

Trace heuristic The trace function has been successfully used as a surrogate for the rank [Pare 2000] when the estimated matrix is positive semidefinite matrix (PSD). This yields a convex tractable problem. Observing that in this case $\operatorname{Trace}(\mathbf{X}) = \|\mathbf{a}(\mathbf{X})\|_1$ (see Eq.), this heuristic can be understood as an ℓ_1 based sparsity promoting approach.

Log-det heuristic In a similar vein, a particularly fruitful idea has been to use the function $\mathbf{Y} \mapsto \log(\det(\mathbf{Y} + \delta \mathbf{I}_m))$ as a smooth surrogate of the rank function on the cone of $m \times m$ PSD matrices [Fazel *et al.* 2003]. While being concave, it yields a reweighted-ℓ_1 like minimization on the singular values [Candès *et al.* 2008]. Thus it can be considered as a refinement of the trace heuristic. In the same work, the authors derive an extension to arbitrary non square matrices.

Nuclear norm heuristic The most popular popular rank surrogate function is certainly the nuclear norm defined as the sum of the singular values of a given matrix [Fazel *et al.* 2001] :

$$\|\mathbf{X}\|_\star = \|\mathbf{a}(\mathbf{X})\|_1. \tag{II.51}$$

This extends the trace heuristic to arbitrary non square matrices. The nuclear norm can be considered the most suitable convex function for promoting low rankness in more than one respect. Similarly to the ℓ_1 norm, the nuclear norm has been shown to be the *convex envelop* of the rank function on the set matrices $\{\mathbf{Y}/\|\mathcal{S}(\mathbf{Y})\|_\infty \leq 1\}$ [Fazel *et al.* 2001], where $\mathcal{S}(\mathbf{Y})$ is the vector of singular values of $\mathbf{Y}$ and therefore, $\|\mathcal{S}(\mathbf{Y})\|_\infty$ is $\mathbf{Y}$'s *spectral norm*. A more recent point of view, previously invoked for the ℓ_1 norm, derives the nuclear norm as the atomic norm associated with the set of rank one matrices with unitary spectral norm [Chandrasekaran *et al.* 2010]. Most importantly, various low rank matrices recovery guarantees using nuclear norm have been established, with a striking analogy to ℓ_1 based sparse recovery results (see [Recht *et al.* 2010, Candès & Plan 2010, Candès *et al.* 2011, Candès & Recht 2009]). Low rank minimization has been successfully applied in various task such as matrix completion, for recommendation systems or missing data problems, background detection in videos and hyperspectral images restoration.

Matrix factorization In certain applications, the rank of the matrix under estimation that we denote r is known. While it is not straightforward to translate this information into a suitable value of the regularization parameter in a nuclear norm minimization, one can directly impose the rank through a matrix factorization. Indeed, the following equivalence holds :

$$\text{rank}(\mathbf{X}) \leq r \Leftrightarrow \exists(\mathbf{S}, \mathbf{A}) \in \mathrm{M}_{mr}(\mathbb{R}) \times \mathrm{M}_{rq}(\mathbb{R})/\mathbf{X} = \mathbf{SA} \tag{II.52}$$

which suggests the alternative formulation of the inverse problem

$$\min_{(\mathbf{S},\mathbf{A})\in \mathrm{M}_{mr}(\mathbb{R})\times \mathrm{M}_{rq}(\mathbb{R})} \|\mathbf{Y} - \mathcal{M}(\mathbf{SA})\|_F^2. \tag{II.53}$$

This problem is globally non convex. However it is bilinear in $\mathbf{S}$ and $\mathbf{A}$ and hence partially convex for each of these matrices. Moreover this approach enables one to incorporate more priors by adding constraints on $\mathbf{S}$ and $\mathbf{A}$. This yields important frameworks such as Non-negative Matrix Factorization (NMF) [Wang & Zhang 2013], or Sparse NMF [Kim & Park 2007].

Matrix factorization and nuclear norm minimization have complementary aspects and depending on the prior knowledge available, one might be more suitable than the other. However, an attempt of using the two in an unified framework can be found in [Cabral *et al.* 2013].

Remark : in certain applications such as system reduction or identification, the low rank matrices of interest are known to have a specific structures which can be usefully integrated in the estimation process. This yields alternative rank minimization methods and algorithms [Markovsky 2008].

Chapitre III

Beyond linear data models

Sommaire

In Chapter II, we presented how signals structural properties can be used for regularizing inverse problems, with a focus on linear low complexity data models. This chapter goes beyond the framework of dictionaries based representations and transforms and considers more general non linear data parametrizations ; the main idea being to uncover and exploit the non linear geometry underlying a data set, especially when this geometry is simple, in a sense to be defined.

In Section A.2, we recall important geometric notions with this respect. We refer the interested reader to [Flaherty & do Carmo 2013] for a more detailed presentation. Then we present an overview of Manifold Learning (ML) main techniques. Section B is an introduction to Optimal Transport (OT). The emphasis will be put on certain geometrical properties of OT that are particularly interesting from a data analysis standpoint.

A Manifold data model

A.1 Reminder on differential geometry

A.1.1 General definitions

A *d-dimensional manifold* can be though of as a collection of points $\mathcal{M}$ which resembles $\mathbb{R}^d$ in each points neighborhood. For instance an ℓ_2 sphere in $\mathbb{R}^3$ is a 2-dimensional manifold. We precise the definition using the notions of *charts* and *atlases*. A bijective map ϕ between a subset $\mathcal{X}$ of $\mathcal{M}$ and an open

subset of $\mathbb{R}^d$ is called a *d-dimensional chart* of $\mathcal{M}$ denoted $(\mathcal{X}, \phi)$. For each point $\mathbf{x}$, the vector $\phi(\mathbf{x})$ constitutes its coordinates in $\mathbb{R}^d$; ϕ^{-1} is referred as a *parametrization* of $\mathcal{M}$ around $\mathbf{x}$. Thus, $\mathcal{M}$ can be viewed as an union of suitably chosen coordinates patches or charts.

It follows naturally that each point of $\mathcal{M}$ shall belong to at least one chart's domain and that two overlapping charts shall give consistent information, which yields the notion of *atlas*. A smooth atlas of $\mathcal{M}$ in $\mathbb{R}^d$ is a collection of d-dimensional charts $(\mathcal{X}_i, \phi_i)_{i\in\Omega}$ for some indexation domain Ω verifying

1. $\bigcup_{i\in\Omega} \mathcal{X}_i = \mathcal{M}$
2. for $(i,j) \in \Omega^2$, $i \neq j/\ \mathcal{X}_i \bigcap \mathcal{X}_j \neq \emptyset$, then the sets $\phi_i(\mathcal{X}_i \bigcap \mathcal{X}_j)$ and $\phi_j(\mathcal{X}_i \bigcap \mathcal{X}_j)$ are open and the map $\phi_i \circ \phi_j^{-1} : \mathbb{R}^d \to \mathbb{R}^d$ is smooth on $\phi_j(\mathcal{X}_i \bigcap \mathcal{X}_j)$.

If $\mathcal{M}$ is a set of geographical locations on the Earth surface, then the notion of atlas recovers its common assertion.

The *maximal atlas* $\mathcal{A}^+$ generated by an atlas $\mathcal{A}$ is the atlas defined as the set of the all the charts $(\mathcal{X}, \phi)$ such that $\mathcal{A} \bigcup (\mathcal{X}, \phi)$ is also an atlas of $\mathcal{M}$. This can be viewed as the minimal structure required for extending differential calculus from linear subspaces to arbitrary sets. Indeed, the study of a functional $f : \mathcal{M} \to \mathbb{R}$ can be replaced on the domain of a particular chart $(\mathcal{X}, \phi)$ in $\mathcal{A}^+$ with the study of the vector space functional

$$\hat{f} : \mathbb{R}^d \to \mathbb{R}, \quad \mathbf{x} \mapsto f(\phi^{-1}(\mathbf{x})), \tag{III.1}$$

on the set $\phi(\mathcal{X})$. Thus, a maximum atlas is also referred as a *differentiable structure* on the $\mathcal{M}$. From this, a manifold can be formally defined as couple $(\mathcal{M}, \mathcal{A}^+)$ of a set of points and its differentiable structure (when it exists). By a slight abuse of notation, in general we will only use $\mathcal{M}$ for referring to the manifold.

There are numerous interesting examples of manifolds. At this point we only mention the particular case of vector spaces which, thereby, are often referred as *linear manifolds*, the subject of this section being the non linear manifolds. In the following, we focus on *connected* manifolds i.e. manifolds which points can not be represented as the union of two or more disjoint nonempty open subsets.

A.1.2 Manifold's first order geometry

A curve on a manifold $\mathcal{M}$ is defined as a function $\gamma : \mathbb{R} \to \mathcal{M}$. Assuming that $\mathcal{M}$ is a submanifold of a vector space $\mathcal{E}$, a *tangent vector* at a point $\mathbf{p}$ of $\mathcal{M}$ is any vector $\mathbf{v}$ in $\mathcal{E}$ such that

$$\mathbf{v} = \lim_{t\to 0} \frac{\gamma(t) - \gamma(0)}{t} \stackrel{\text{def}}{=} \gamma'(0), \tag{III.2}$$

for some smooth curve γ on $\mathcal{M}$ which verifies $\gamma(0) = \mathbf{p}$. The *tangent space* to $\mathcal{M}$ at $\mathbf{p}$ denoted $\mathcal{T}_{\mathbf{p}}\mathcal{M}$ is the collection of all the tangent vectors at $\mathbf{p}$. If $\mathcal{M}$ is a d-dimensional manifold, one can verify that $\mathcal{T}_{\mathbf{p}}\mathcal{M}$ is a vector space of dimension d. We note that d is often referred in applications as the *intrinsic dimension* of $\mathcal{M}$ by contrast to the dimension of the ambient space $\mathbf{E}$. As we will see in Section A.2, this notion is important from a practical point of view, especially when $d \ll \dim(\mathcal{E})$. $\mathcal{T}_{\mathbf{p}}\mathcal{M}$ can be endowed with an inner product $\langle ., . \rangle_{\mathcal{M},\mathbf{p}}$ which induces the norm

$$\|\mathbf{x}\|_{\mathcal{M},\mathbf{p}} = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle_{\mathcal{M},\mathbf{p}}}. \tag{III.3}$$

The manifold $\mathcal{M}$ is *Riemannian* if $\langle ., . \rangle_{\mathcal{M},\mathbf{p}}$ varies smoothly with $\mathbf{p}$; in this case, the inner product $\langle ., . \rangle_{\mathbf{p}}$ is referred as the *Riemannian metric*.

Riemannian manifolds and more generally, the Riemannian geometry constitute the mathematical framework of the General Relativity ; precisely the theory relates the Riemannian metric $\langle ., . \rangle_{\mathcal{M},\mathbf{p}}$ also referred as the metric tensor at every space-time point to the mass/energy distribution through the celebrated Einstein's equations [Besse 2007]. In this physical context, the smoothness requirement is particularly adapted.

In the Riemannian framework, the length of a smooth curve $\gamma : [t_1, t_2] \to \mathcal{M}$ is given by

$$L(\gamma) = \int_{t_1}^{t_2} \|\gamma'(t)\|_{\mathcal{M},\gamma(t)} \mathrm{d}t. \tag{III.4}$$

This gives a mean of measuring distances on the manifold. Indeed, one can introduce the set of smooth curves joining two points $\mathbf{p}_1$ and $\mathbf{p}_2$ in $\mathcal{M}$:

$$\Gamma(\mathcal{M}, \mathbf{p}_1, \mathbf{p}_2) = \{\gamma : [0,1] \to \mathcal{M} / \ \gamma(0) = \mathbf{p}_1,\ \gamma(1) = \mathbf{p}_2 \text{ and } \gamma \text{ smooth on } [0,1]\}. \tag{III.5}$$

Then the length of the shortest path between $\mathbf{p}_1$ and $\mathbf{p}_2$ can be calculated as

$$L_{\mathcal{M}}(\mathbf{p}_1, \mathbf{p}_2) = \underset{\gamma \in \Gamma(\mathcal{M}, \mathbf{p}_1, \mathbf{p}_2)}{\operatorname{argmin}} L(\gamma). \tag{III.6}$$

The functional $L_{\mathcal{M}} : \mathcal{M}^2 \to \mathbb{R}$ known as the *Riemannian distance* actually defines a metric on $\mathcal{M}$, which makes a connected Riemannian manifold a *metric space*. If we consider a d-dimensional chart $(\mathcal{X}, \phi)$ of $\mathcal{M}$, ϕ is a local isometry of $(\mathcal{X}, L_{\mathcal{M}})$ in $\mathbb{R}^d, \ell_2$ if

$$\forall\ (\mathbf{p}_1, \mathbf{p}_2) \in \mathcal{X}^2\ L_{\mathcal{M}}(\mathbf{p}_1, \mathbf{p}_2) = \|\phi(\mathbf{p}_1) - \phi(\mathbf{p}_2)\|_2. \tag{III.7}$$

Such property is particularly suitable in dimension reduction tasks when the data are sampled from a Riemannian manifold. Furthermore, if $L_{\mathcal{M}}$ is a euclidean norm, then the manifold is qualified as *intrinsically euclidean*.

A.1.3 Geodesics on Riemannian manifolds

Certain curves on $\mathcal{M}$ have the property that they induce shortest paths locally. It is the case of the *great circles* if $\mathcal{M}$ is a sphere in $\mathbb{R}^3$. These curves are called *geodesics* and we recall thereafter some related *properties* [Lee 2006]. Formally, a curve γ on a Riemannian submanifold $\mathcal{M}$ of a vector space is a geodesic if the acceleration vector is permanently normal to the tangent plane :

$$\gamma''(t) \in \mathcal{T}_{\gamma(t)}\mathcal{M}^{\perp} \tag{III.8}$$

for any t in the domain of γ. This implies in particular a constant velocity along the curve :

$$\|\gamma'(t)\|_{\mathcal{M},\gamma(t)} = c. \tag{III.9}$$

Recalling Eq.III.4, this yields an interesting property of geodesics, which is, they are parametrized with an affine function of the arclength. As a simple example, the geodesics in a vector space are the constant speed straight curves. A geodesic is *minimal* if its length is smaller than the lengths of all the smooth curves on the manifold joining its endpoints. The length of the minimal geodesic is nothing but the Riemannian distance between its endpoints, hence the terminology of *geodesic distance* sometimes found in the literature. Any sufficiently small segment of a geodesic is minimal. An interesting property from a practical point of view is that if a geodesic is minimal on an interval I, then its restriction to any subinterval of I is also minimal.

However, the shortest path between two points of a connected Riemannian manifold need not to be a geodesic, even up to a change of parametrization. *Geodesically complete* Riemannian manifolds are precisely those in which any two points can be connected through a geodesic. In particular, in this case minimal geodesics exist for any two points.

A nice connection between the metric space $(\mathcal{M}, L_{\mathcal{M}})$'s topology and $\mathcal{M}$'s geodesics is given the Hopf-Rinow theorem [Hopf & Rinow 1931] : among other equivalences, it states that $(\mathcal{M}, L_{\mathcal{M}})$ is (topologically) complete if and only if $\mathcal{M}$ is geodesically complete.

A.2 Manifold learning techniques

The computational applications of manifold data models are two-fold. On the one hand, there are situations where one needs to estimate a signal that belongs to a *known* manifold. One can find some examples related to matrices manifolds in [Absil *et al.* 2009]. An application to natural signals and images restoration can be found in [Peyré 2009]. This first framework boils down to solving optimization problems on non linear manifolds, which constitutes a research area on its own. Among other recent contributions, we can cite

[Bacák *et al.* 2016] and some references therein. Conversely, one might have a data set (approximately) sampled from an unknown manifold and instead try to uncover the underlying manifold geometry. This can serve purposes of visualization, classification or data synthesis. We term Manifold Learning (ML) the corpus of techniques aiming at characterizing a manifold from a set of samples lying on or close to the manifold. This has been a particularly active field in the Machine Learning community over the past decade. We review thereafter some of the most influential contributions.

Unless mentioned otherwise, the following notations holds in this section :

— $(\mathbf{x}_i)_{1\leq i\leq N}$ is the learning data set ;

— the learning set is sampled from a submanifold $\mathcal{M}$ of $\mathbb{R}^m$;

— $(\mathbf{y}_i)_{1\leq i\leq N}$ is the low dimensional representation of the training set ;

— d might refer to the intrinsic dimension of $\mathcal{M}$ or the dimensionality of $(\mathbf{y}_i)_{1\leq i\leq N}$.

A.2.1 Multidimensional scaling

The multidimensional scaling (MDS) refers to a set of techniques that aims at determining the underlying attributes and the spatial organization of a set of objects based on measurements of pairwise similarities (or dissimilarities) between the objects. We limit our presentation to the so-called classical MDS. This method was originally introduced in Psychometrics for uncovering stimuli perceptual dimensions and their scales [Torgerson 1952]. Given a set of N stimuli, one is given a matrix $\mathbf{D} = (d_{ij})_{1\leq i,j\leq N}$, where the scalar d_{ij} is a comparative distance between the i^{th} and the j^{th} stimuli, which we assume here to be positive. However, this matrix does not need to derive from a metric function. The classical MDS computes a set of vectors $(\mathbf{y}_i)_{1\leq i\leq N}$ in $\mathbb{R}^d$ so that $\|\mathbf{y}_i - \mathbf{y}_j\|_2 \approx d_{ij},\ \forall\ (i,j) \in [\![1, N]\!]^2$ by solving the optimization problem

$$\min_{\mathbf{y}_1,\cdots,\mathbf{y}_N} \sum_{(i,j)\in[\![1,N]\!]^2} (d_{ij} - \|\mathbf{y}_i - \mathbf{y}_j\|_2)^2. \tag{III.10}$$

We note that this Problem has a known closed-form solution based on an eigen-system decomposition of $\mathbf{D}$; this is presented in more details in Chapter VI. The dimension r is a parameter and the coordinates have hierarchical importance, which follows directly from the solution's structure. This allows one to visualize the object space at different scale (when $d > 3$). The MDS has come to be a popular tool for exploratory data analysis. A straightforward application is the problem of recovering city coordinates from pairwise distances. Although it is not directly meant for manifold-valued data analysis, MDS constitutes an important step in the foundational ML technique Isomap,

which is presented in the following section. We refer to [Borg & Groenen 2005] for a thorough presentation of MDS.

A.2.2 Isomap

The Isomap is a non-linear dimension reduction method introduced in [Tenenbaum *et al.* 2000]. Given a data set $(\mathbf{x}_i)_{1\leq i\leq N}$, sampled from a Riemannian submanifold $\mathcal{M}$ of $\mathbb{R}^m$, Isomap aims at computing a global low dimensional chart that is isometric to $\mathcal{M}$ endowed with the Riemannian distance. Thus Isomap first computes a matrix $\mathbf{D}$ of the pairwise approximated Riemmanian distances on the data set, as illustrated in Fig. ***III.1***. Then it applies the clas-

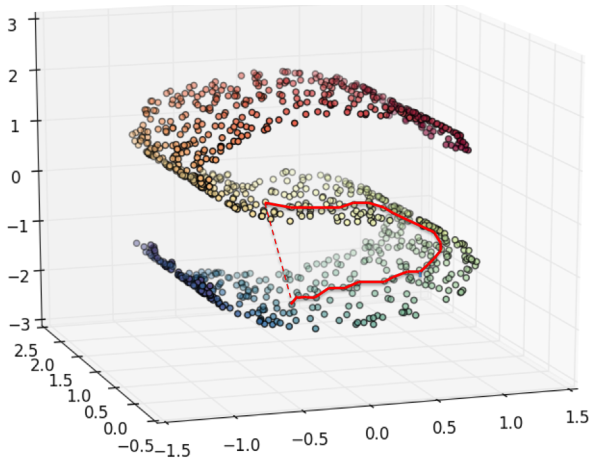


Figure *III.1* – S-curve data set : Isomap estimates the riemanian distances based on a nearest neighbor graph ; when the samples are sufficiently dense, this prevents one from bypassing the manifold's geometry.

sical MDS to $\mathbf{D}$ for computing the low dimension coordinates of the samples. The dimension parameter r introduced in the previous section shall be ideally chosen as the intrinsic dimension of $\mathcal{M}$. It corresponds to the number of degrees of freedom of the system that generated the data. Isomap is guaranteed to recover the true underlying geometrical structure as the sampling density tends to the infinity if the manifold is intrinsically euclidean. Besides, it can be successfully applied for several images manifolds of interest in computer vision [Donoho & Grimes 2005]. However, the Riemannian distances estimation step might be computationally demanding, since it involves finding shortest paths on a potentially large weighted graph. This shortcoming is avoided in the next approach.

A.2.3 Local Linear Embedding

We consider the same setting as in the previous section. The Local Linear Embedding (LLE) [Roweis & Saul 2000] lies on natural geometrical intuitions. Firstly, if the samples are sufficiently dense then each sample's nearest neighbors approximately lie in its tangent space (in fact an affine space parallel to this tangent space). Hence, each sample can be expressed as a linear combination of its nearest neighbors. Hence, the first step of LLE is to compute weights matrix $\mathbf{W} = (w_{ij})_{1\leq i,j\leq N}$ by solving the following problem :

$$\min_{\mathbf{W}} \sum_{i=1}^{N} \|\mathbf{x}_i - \sum_{j\in \mathrm{nn}_k(i)} w_{i,j}\mathbf{x}_j\|_2^2, \text{ s.t. } \sum_{j\in \mathrm{nn}_k(i)} w_{i,j} = 1 \ \forall i \in [\![1, N]\!], \tag{III.11}$$

where $\mathrm{nn}_k(i)$ is the set of indexes of the k nearest neighbors of the sample $\mathbf{x}_i$. Besides $w_{ij} = 0$ if $j \notin \mathrm{nn}_k(i)$. These weights are invariant under isometric affine transformations and rescaling of the manifold. Thereby, they locally capture its intrinsic geometry, and should remain optimal in a low dimensional representation of the manifold that "preserves" this geometry. It follows that the low dimension coordinates can be estimated by solving

$$\min_{\mathbf{y}_1,\cdots,\mathbf{y}_N} \sum_{i=1}^{N} \|\mathbf{y}_i - \sum_{j\in \mathrm{nn}_k(i)} w_{i,j}\mathbf{y}_j\|_2^2, \text{ s.t. } \sum_{j\in \mathrm{nn}_k(i)} w_{i,j} = 1 \ \forall i \in [\![1, N]\!]. \tag{III.12}$$

Interestingly, this method derives a global characterization of the manifold from local information, which is made possible by the neighborhoods overlaps. The same principle will appear in the next methods.

A.2.4 Laplacian Eigenmaps

As for the LLE, the Laplacian Eigenmaps methods builds over the idea of preserving local neighborhood information. We note $\mathrm{nn}(i)$ the set of indexes of the "relevant" neighbors of the ith sample. The low dimensional coordinates are calculated by solving the problem

$$\sum_{\mathbf{y}_1,\cdots,\mathbf{y}_N} \sum_{i=1}^{N} \sum_{j\in \mathrm{nn}(i)} w_{ij}\|\mathbf{x}_i - \mathbf{x}_j\|_2^2, \tag{III.13}$$

where

$$w_{ij} = \exp(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|_2^2}{t}), \tag{III.14}$$

with some $t > 0$. Thus, neighbor samples are strongly constrained to remain close in the new coordinate system. As previously, let introduce the weights matrix $\mathbf{W} = (w_{ij})_{1\leq i,j\leq N}$, with the convention $w_{ij} = 0$ if $j \notin \mathrm{nn}(i)$. It is

shown in [Belkin & Niyogi 2003] that the coordinates slices of the solution of Problem III.13, i.e. the vectors $(\mathbf{y}_i[l])_{1\leq i,j\leq N}$, for $l \in [\![1, d]\!]$ are solutions of the generalized eigenvalues problem

$$\mathbf{L}\mathbf{y} = \lambda \mathbf{D}\mathbf{y}, \tag{III.15}$$

where $\mathbf{D}$ is a diagonal matrix defined as $\mathbf{D}[i,i] = \sum_{j\in \mathrm{nn}(i)} w_{ij}$, for $i \in [\![1, N]\!]$ and $\mathbf{L} = \mathbf{D} - \mathbf{W}$. $\mathbf{L}$ is the laplacian matrix of a weighted graph $\mathcal{G}$ with N nodes representing the data sample, the weight of the vertex between the i^{th} and the j^{th} node being equal to w_{ij} [Chung 1997]. Interestingly, it is shown that for t sufficiently small, $\mathbf{L}$ can be considered a discrete approximation of the Laplace-Beltrami operator on $\mathcal{M}$ over the graph $\mathcal{G}$. Besides, the eigenfunctions of the Laplace-Beltrami operator on $\mathcal{M}$ are solution of the problem

$$\min_f \int_{\mathcal{M}} \|\nabla f(\mathbf{x})\|_2^2 \mathrm{d}\mu(\mathbf{x}), \quad \text{s.t.} \int_{\mathcal{M}} f(\mathbf{x})^2 \mathrm{d}\mu(\mathbf{x}) = 1 \tag{III.16}$$

for some probability measure $\mathrm{d}\mu$ on $\mathcal{M}$. This shows that the Laplace-Beltrami operator's eigenfunctions tend to realize mappings of the manifold that moderately affect the local samples organization. This supports the geometrical intuition which underpins the Laplacian Eigenmaps method.

A.2.5 Hessian Eigenmaps

The Hessian-based Local Linear Embedding (HLLE) methods represents an improvement over each of the pre-cited ML techniques in terms of preservation of the manifold's (local) geometry [Donoho & Grimes 2003]. At the heart of this method, a quadratic form termed the *$\mathcal{H}$-Functional* which is defined as

$$\mathcal{H}(f) = \int_{\mathcal{M}} \|H_f(\mathbf{x})\|_F^2 \mathrm{d}\mu(\mathbf{x}), \tag{III.17}$$

where f is a smooth functional on the manifold, $H_f(\mathbf{x})$ is the hessian matrix of f at $\mathbf{x}$ and $\mathrm{d}\mu$ is a probability measure on $\mathcal{M}$. The hessian matrix is defined in an orthogonal coordinates system in the tangent space $\mathcal{T}_{\mathbf{x}}\mathcal{M}$ relatively to the dot product inherited from the ambient space. This quadratic form is well defined in the sense that it is independent of the orthogonal coordinates system chosen in a given tangent plane. It quantifies how curvy is a functional on the manifold, in average. To gain a better grasp of the intuition behind this choice, let consider the family of functionals $(f_i)_{0\leq i\leq m}$ defined as

$$f_i : \mathbb{R}^m \to \mathbb{R}, \ \mathbf{x} \mapsto \mathbf{x}[i], \tag{III.18}$$

$\forall i \in [\![1, m]\!]$ and

$$f_0 : \mathbb{R}^m \to \mathbb{R}, \ \mathbf{x} \mapsto 1. \tag{III.19}$$

These functions span the set of affine and hence, flat functions on $\mathbb{R}^m$. Similarly, for a non-linear manifold $\mathcal{M}$, one can expect the set of flat functions on the manifold to be related to a canonical coordinate system on the manifold. Precisely, if $\mathcal{M}$ is a Riemannian $d-dimensional$ submanifold of $\mathbb{R}^m$ which is *locally isometric* to an *open connected* subset of $\mathbb{R}^d$, then the null space of $\mathcal{H}$ is a $d+1$-dimensional vector space consisting of the constant functionals on the manifold and a d-dimensional space spanned by the isometric coordinates. Thus, the HLLE derives the low dimensional representation by estimating the null space of $\mathcal{H}$ based on the learning set as follows :

— for each $i \in [\![1, N]\!]$, one computes a $d(d+1)/2 \times k$ matrix $\mathbf{H}^{(i)}$ from local tangent coordinates so that for a functional $\mathbf{f}$ defined over the learning set which is the discretization of a smooth functional f defined over $\mathcal{M}$, if we note $\mathbf{f}^{(i)}$ its restriction to a vicinity of size k of the i^{th} sample, the vector $\mathbf{H}^{(i)}\mathbf{f}^{(i)}$ constitutes an estimate of the entries of $H_f(\mathbf{x}_i)$;

— from the matrices $\mathbf{H}^{(i)}$, one builds a $N \times N$ symmetric positive matrix $\mathbf{H}$ so that $\mathcal{H}(f) \approx \mathbf{f}^T\mathbf{H}\mathbf{f}$;

— ones computes $d+1$ orthonormalized eigenvectors corresponding to the $d+1$ smallest eigenvalues of $\mathbf{H}$; we note these vectors $\mathbf{u}_0, \cdots, \mathbf{u}_d$, where there are sorted in increasing order of the associated eigenvalues ; $\mathbf{u}_0$ is associated with the eigenvalue 0 and corresponds to constant functionals ; the low dimension coordinates of the i^{th} sample are given by the vector $(\mathbf{u}_j[i])_{1 \leq j \leq d}$.

The requirement for local isometry brings this method closer to the LLE in which it is implicit. Besides, the $\mathcal{H}-$Functional somehow resembles the quadratic form minimized in Problem III.16, which can be rewritten in terms of the Laplace Beltrami operator. However, the hessian is somehow more suitable than a laplacian for linear functions recovery [Donoho & Grimes 2003]. Furthermore, the HLLE can be applied to a wider range of manifolds than Isomap which requires the manifold to be *globally* isometric to a *convex* subset of $\mathbb{R}^r$.

Remarks : in this thesis, we are interested in the problem of synthesizing new samples from a manifold (namely a PSFs manifold) given a learning sample set and according to a prescribed parametrization of the manifold. However most of the ML techniques previously presented do not provide a natural way of doing so, except for the LLE. For instance, in [Tenenbaum *et al.* 1998], generalized radial basis functions are used for computing a mapping between the low dimensional coordinates and the manifold. However, the recently introduced Geometric Multi-Resolution Analysis [Allard *et al.* 2012]

seems particularly suitable in this regard. On the other hand, the reverse problem of determining the low dimension coordinates of an out-of-sample point has been treated in several works (see for instance [Bengio *et al.* 2003]).

For most manifold learning techniques, including those presented here, the manifold intrinsic dimension is a free parameter. Although in general, it can be empirically (and efficiently) tuned from the learning procedure itself, several methods have been proposed for estimating the intrinsic dimension of an embedded manifold, given a learning sample (see [Little 2011] and the references therein).

B Optimal transport

The photograph of the figure ***III.2*** was taken by the author at the french "Japan Expo" and shows the display table of a mangas seller. Speculatively, we suppose that at the end of the exposition, the unsold books could fit in a single box. The optimal transport problem arises from two trivially observable facts in this situation :

— the seller can store the books in the box following different strategies ; for instance he can distribute each book pile in the box as evenly as possible, or he can try to preserve the book piles as much as possible ;

— some strategies involve more efforts in terms of mass displacement than others. This naturally yields the question of finding the minimal effort strategy of transportation.

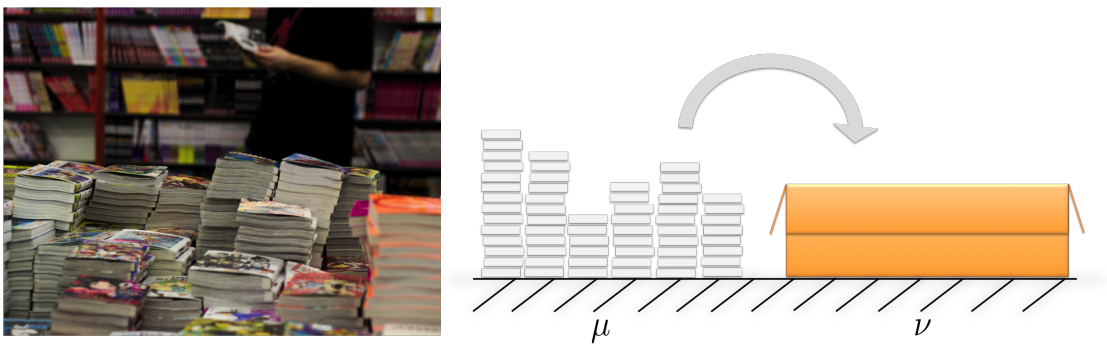


Figure *III.2* – Minimal effort packaging.

The first known formal study of this question traces back to the late 18^{th} century ([Monge 1781]). Since then, the Optimal Transport has come to be a remarkably fruitful topic in applied mathematics. In this section, we review the Optimal Transport's theory basic formalism and some important properties of the optimal solutions.

B.1 Problem formulations

B.1.1 Framework and basic definitions

Let consider two probability measures μ and ν defined respectively on some measure spaces X and Y. We shall define a measurable cost function $c : X \times Y \to \mathbb{R}_+$ which quantifies the effort of transporting one unit of mass from a point in X to a point in Y. For a measurable map $T : X \to Y$ and a subset B of Y,

$$T^{-1}(B) \stackrel{\text{def}}{=} \{\mathbf{x} \in X / \; T(\mathbf{x}) \in B\}. \tag{III.20}$$

We say that μ is the push-forward of ν by T if

$$\text{for any measurable set } B \subset Y, \; \nu(B) = \mu(T^{-1}(B)), \tag{III.21}$$

and we note

$$\nu = T\#\mu \tag{III.22}$$

T is then referred as a *transport map* between μ and ν. This notion is illustrated in Fig.***III.3***.

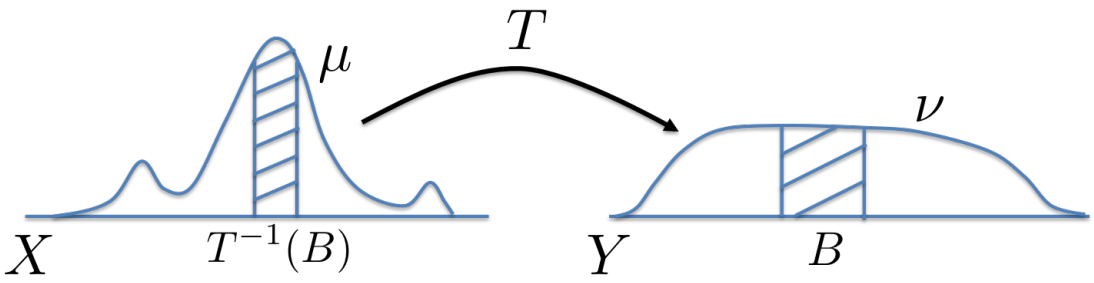


Figure *III.3* – Measure pushforward : the shaded surfaces should have the same area.

B.1.2 Monge problem

The *Monge Optimal Transport problem* is the following :

$$\min_T \int_X c(\mathbf{x}, T(\mathbf{x}))\mathrm{d}\mu(\mathbf{x}), \; \text{s.t. } \nu = T\#\mu \tag{III.23}$$

where $c(\mathbf{x}, T(\mathbf{x}))\mathrm{d}\mu(\mathbf{x})$ is by construction the cost of moving the elementary mass $\mathrm{d}\mu\mathbf{x}$ from $\mathbf{x}$ to $T(\mathbf{x})$. In several applications, $X = Y$ and c is chosen as a metric on X. c is then termed the *ground metric*, in reference to the logistic interpretation of the Optimal Transport.

The Monge's problem imposes a transport with no mass split. In other words, the elementary mass $\mathrm{d}\mu\mathbf{x}$ has to be mapped to a single location in Y. It appears intuitively that depending on μ and vu such a mapping might not be possible (for instance if $X = Y = \mathbb{R}$ and if one of the two probability measures has a dense support while the other has a discrete support in $\mathbb{R}$). We will shortly introduce a relaxation in this regard due to Kantorovitch [Kantorovitch 1942].

B.1.3 Kantorovich problem

A *transference* or *transportation plan* between μ and ν is a probability measure Π on $X \times Y$ that verifies

$$\int_Y \mathrm{d}\Pi(\mathbf{x}, \mathbf{y}) = \mathrm{d}\mu(\mathbf{x}) \text{ and } \int_X \mathrm{d}\Pi(\mathbf{x}, \mathbf{y}) = \mathrm{d}\nu(\mathbf{x}). \tag{III.24}$$

$\mathrm{d}\Pi(\mathbf{x}, \mathbf{y})$ represents the elementary mass transferred from $\mathbf{x}$ to $\mathbf{y}$. From this definition, the *Kantorovich Optimal Transport problem* can be stated as follows :

$$\min_{T \in \mathcal{P}(\mu,\nu)} \int_{X \times Y} c(\mathbf{x}, \mathbf{y}) \mathrm{d}\Pi(\mathbf{x}, \mathbf{y}), \tag{III.25}$$

where $\mathcal{P}(\mu, \nu)$ is the set of transportation plans between μ and ν.

Note that the probability distribution of $X \times Y$

$$\mu \otimes \nu : (\mathbf{x}, \mathbf{y}) \mapsto \mu(\mathbf{x})\nu(\mathbf{y}) \tag{III.26}$$

is a transportation plan between μ and ν so that $\mathcal{P}(\mu, \nu)$ is always non-empty.

We refer the reader to [Villani 2003] for detailed discussions on the existence of solutions to the Monge and Kantorovich problems. However, we note that if $X = Y = \mathbb{R}^m$, for some m and if the cost is in the form

$$c(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_p^p, \tag{III.27}$$

for some $p > 1$, the Kantorovich problem in general admits a unique solution.

B.1.4 Brenier-Benamou problem

We assume that $X = Y = \mathbb{R}^m$. Let note T^* a solution to the Monge problem whenever it exists. Provided that T^* is smooth, μ and ν obey the relation

$$\mu(\mathbf{x}) = \nu(T(\mathbf{x}))|\det(\mathbf{J}_{T^*}(\mathbf{x}))|, \tag{III.28}$$

where $\mathbf{J}_{T^*}(\mathbf{x})$ is the jacobian matrix of T^* at $\mathbf{x}$ [Ambrosio & Savaré 2007]. Let $(T_t)_{0 \leq t \leq 1}$ be a family of maps defined as

$$T_t(\mathbf{x}) = t\mathbf{x} + (1-t)T^*(\mathbf{x}), \ \forall (t, \mathbf{x}) \in [\![0, 1]\!] \times X. \tag{III.29}$$

If $c(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|_2^2$, then the function χ^* defined on $[\![0,1]\!] \times X$ as

$$\chi^*(t,\mathbf{x}) = \nu(T_t(\mathbf{x}))|\det(\mathbf{J}_{T_t}(\mathbf{x}))| \tag{III.30}$$

is solution of the problem

$$\min_{\chi,\mathbf{v}} \int_{[\![0,1]\!]\times X} \chi(t,\mathbf{x})\|\mathbf{v}(t,\mathbf{x})\|_2^2 \, \mathrm{d}t\mathrm{d}\mathbf{x} \text{ s.t.} \tag{III.31}$$

$$\partial_t\chi + \mathrm{div}_{\mathbf{x}}(\chi\mathbf{v}) = 0 \tag{III.32}$$

$$\chi(0,.) = \mu,\ \chi(1,.) = \nu, \tag{III.33}$$

for a vector field $\mathbf{v} : [\![0,1]\!] \times X \to X$. This is the celebrated Benamou-Brenier formulation of the Optimal Transport problem [Benamou & Brenier 2000]. The map $t \mapsto \chi^*(t,.)$ can be viewed as a curve in the space of probability measures on X which joins μ to ν. This particular curve will be characterized more precisely in the next section. Let just say for now that it realizes the displacement of elementary mass particles of the distribution μ toward the distribution ν. Thus, $\mathbf{v}$ shall be interpreted as a velocity field and Eq.III.32 ensures mass preservation through the displacement. In the same spirit, $\chi(t,\mathbf{x})\|\mathbf{v}(t,\mathbf{x})\|_2^2$ has to be identified as a kinetic energy density in this formulation.

B.2 Geometry of optimal transport

B.2.1 The transport map structure

We consider a taxi firm which has three cars available located at $\mathbf{x}_i,\ i \in \{1,2,3\}$ and needs to pick three potential customers up as fast as possible. The customers are waiting at $\mathbf{y}_i,\ i \in \{1,2,3\}$ as depicted in Fig.***III.4***. Assuming homogeneous traffic conditions, this situation can be modeled as a particular instance of the Monge problem, opportunely using the Manhattan distance as the cost function.

The optimal assignment is easily picked up by eye. It also appears that $\forall (i,j) \in \{1,2,3\}^2$,

$$(\mathbf{x}_i - \mathbf{x}_j)^T(\mathbf{y}_{\sigma^\star(i)} - \mathbf{y}_{\sigma^\star(j)}) \leq 0 \tag{III.34}$$

where $\sigma^\star(i)$ is the index that corresponds to the customer assigned to the i^{th} cab. This seemingly coincidental fact illustrates a general property of the optimal transport solution, which turns out to be a monotonic rearrangement of one distribution onto the other under certain conditions [Villani 2003]. Precisely, using the same notations as in the previous section, with $X = Y = \mathbb{R}^m$ and $c(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|_2^2$, if we further assume that μ and ν do not give zero probability to small sets (in terms of Hausdorff dimension), then there exists a convex function ϕ so that the optimal transport map is given by

$$T = \nabla\phi. \tag{III.35}$$

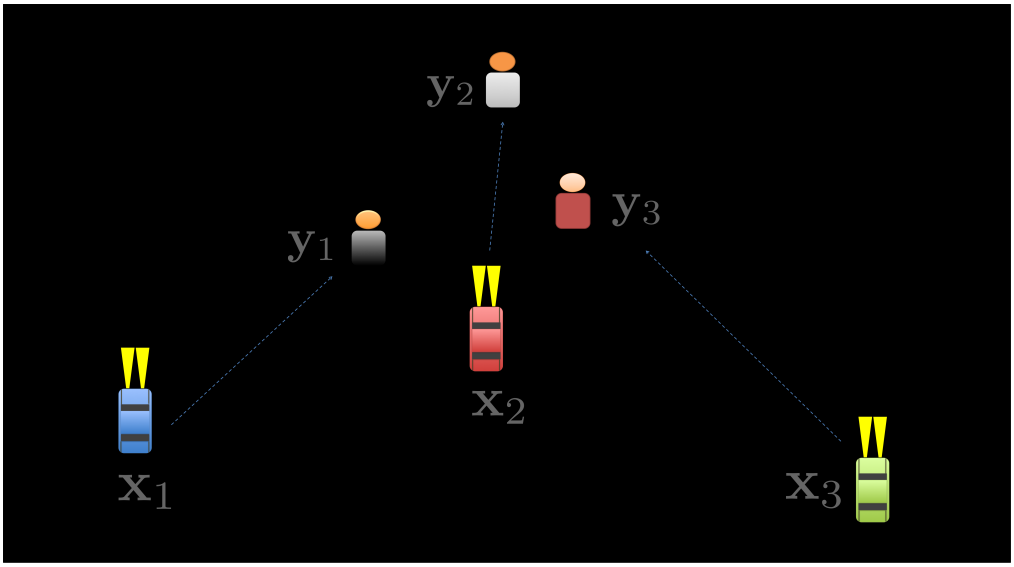


Figure *III.4* – Optimal (cab) assignment.

This result is often referred as the Optimal Transportation theorem and has been generalized to non quadratic costs. The previously mentioned monotonicity follows from the convexity of ϕ.

The existence of ϕ suggests potential applications of this framework to astronomical imaging. For instance, in strong gravitational lensing analysis, ϕ could be interpreted as the gravitational potential induced by a massive unknown structure.

We now turn to the geometry related to the minimal transportation cost.

B.2.2 Monge-Kantorovich distances

The Transport minimal cost in Monge and kantorovich problems measures the similarity between two distributions. Indeed, the more two distributions look alike, the less effort is needed for transforming one into the other, and conversely. This section precises this intuition.

Somehow the structure required for this purpose is weaker than what was previously needed in for characterizing the optimal transport map. In particular, X is only assumed to be a separable complete metric space, endowed with a metric $\mathcal{D}$; also $Y = X$. The cost function takes the following form :

$$c(\mathbf{x}, \mathbf{y}) = \mathcal{D}(\mathbf{x}, \mathbf{y})^q, \tag{III.36}$$

for some $q \geq 1$.

Following the notations in [Villani 2003],

— $\mathcal{T}_q(\mu,\nu)$ is the minimum cost associated with the Kantorovich problem ;

— $P_q(X)$ is the set of probability measures with finite moments of order q.

Then the functional

$$W_q : P_q(X) \times P_q(X),\ (\mu,\nu) \mapsto \mathcal{T}_q(\mu,\nu)^{\frac{1}{q}} \tag{III.37}$$

defines a metric on $P_q(X)$. This metric is known under different names among which the *Monge-Kantorovich distance of order q* or the q^{th} *Wasserstein distance*. In the particular case $q = 2$, it is termed the *quadratic Wasserstein distance*.

We assume that $X = \mathbb{R}^m$. $P_2(\mathbb{R}^m)$ is a submanifold of the vector space of finite signed measures on $\mathbf{R}^m$. For a given $\mu \in P_2(\mathbb{R}^m)$, let ρ_1 and ρ_2 be smooth curves in $P_2(\mathbb{R}^m)$ so that

$$\rho_1(0) = \rho_2(0) = \mu. \tag{III.38}$$

$P_2(\mathbb{R}^m)$ can be endowed with a Riemannian metric defined as follows [Otto 2001] :

$$\langle \frac{\partial \rho_1}{\partial t}(0), \frac{\partial \rho_2}{\partial t}(0) \rangle_{\mathcal{M},\gamma(t)} = \int_{\mathbb{R}^m} \mu(\mathbf{x}) \langle \mathbf{v}_1(\mathbf{x}), \mathbf{v}_2(\mathbf{x}) \rangle \, \mathrm{d}\mathbf{x}, \tag{III.39}$$

where $\mathbf{v}_1$ and $\mathbf{v}_2$ are two velocity fields in $\mathbb{R}^m$ which satisfy the mass conservation equation III.32 w.r.t. ρ_1 and ρ_2 respectively, while guaranteeing that the associated instantaneous kinetic energies remain as small as possible [Villani 2003]. Then, the quadratic Wasserstein distance is precisely the associated Riemannian distance and most importantly, the unique geodesic between two distributions μ and ν is the previously defined curve $\chi^*(t,.)$ (see Eq.III.30). The process of interpolating two probability measures following this geodesic is known as the *displacement interpolation* because informally, the geodesic is obtained by advecting elementary mass particles from one distribution toward a target location defined by the optimal transport map in the other distribution, following straight lines [McCann 1997]. The simplicity of this mechanism makes it particularly appealing for numerical applications.

Since W_2 is a geodesic distance, the curve $\chi^*(t,.)$ solves the following optimization problem :

$$\chi^*(t,.) = \underset{\omega}{\operatorname{argmin}}\ (1-t)W_2(\mu,\omega)^2 + tW_2(\nu,\omega)^2, \tag{III.40}$$

$\forall t \in [\![0,1]\!]$. Hence, $\chi^*(t,.)$ can be formally interpreted as the barycenter of the probability measures μ and ν associated to the weights $1-t$ and t in the

Wasserstein metric. This notion has been extended to an arbitrary number of probability measures in [Agueh & Carlier 2011] :

$$\mu_{\text{bar}} = \underset{\omega}{\text{argmin}} \sum_{i=1}^{N} w_i W_2(\mu_i, \omega)^2, \tag{III.41}$$

with $\sum_{i=1}^{N} w_i = 1$ and $w_i \leq 0 \ \forall i \in [\![1, N]\!]$. μ_{bar} is referred as the (quadratic) Wasserstein barycenter of the probability measures $(\mu_i)_{1 \leq i \leq N}$ and is precisely characterized in [Agueh & Carlier 2011]. However, the choice of Wasserstein barycenters for averaging data samples can be motivated through the trivial example of Fig.VI.9. Indeed, the Wasserstein barycenter is somehow more meaningful than the multimodal euclidean barycenter.

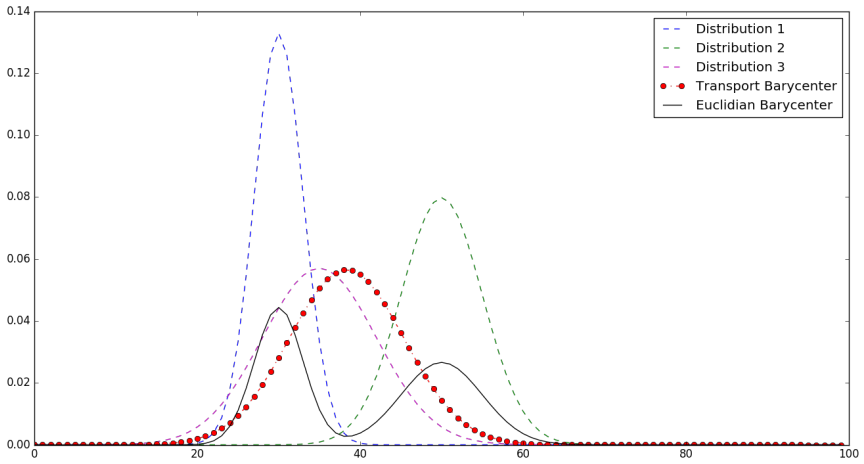


Figure *III.5* – Three gaussian probability measures and their euclidian and quadratic Wasserstein barycenters.

We postpone further discussions in this regard to Chapter VI.

Chapitre IV

Multiple frame PSF super-resolution

Sommaire

The weak gravitational lensing is one of the most promising tools to probe the dark matter distribution in the universe. The idea is to infer, from a billion images of galaxies, the shape distortions due to dark matter gravitational lensing and then estimate the dark matter mass density along different lines of sight. As mentioned in the introduction, the Euclid mission will provide the data for such a purpose. Nevertheless, galaxy images are distorted due to the PSF. Therefore, it is critical to know this distortion accurately. It can be modeled in first approximation as a convolution of the desired image by the PSF of the telescope, which is typically space and time-varying. In practice, isolated stars provide PSF measurements at different locations in the field of view. Nevertheless, these stars images can be aliased as it is the case in Euclid, given the CCD sensor sizes. On the other hand, the surveys

are generally designed so that different images of the same stars are available and likely to be with different subpixel offsets on the sensor grid. Moreover, one may consider that nearby star images give to some extend the same local PSF. We, thus, can consider that different low-resolution versions of the same PSF are available in practice, so that one may apply an super-resolution (SR) method to recover aliased frequencies.

This chapter precisely tackles this problem. The SR is a widely studied topic in general image processing literature. Yet, some methods have been specifically proposed for astronomical data. For instance, there is the software IMCOM [Rowe *et al.* 2011] and PSFEx, which proposes an SR option. The IMCOM provides an oversampled output image from multiple undersampled input images, assuming that the PSF is fully specified. Since it does not deal with the PSF restoration itself, we use PSFEx as our main reference. The PSFEx performs SR by minimizing the sum of two terms. The first term is a weighted quadratic distance, relating the underlying PSF to each of the low resolution measurements. The second term consists of the square l_2 norm of the difference between the underlying PSF and a smooth first guess. This term is meant for regularization. In the proposed algorithm, we introduce a new regularization scheme based on the optimization variable sparsity in a suitable dictionary.

Section A presents the general principle of SR along with some state-of-the-art methods in astronomical domain. In Section B, we present the proposed algorithm in details, which is followed with some numerical experiments in Section C. We conclude by summarizing the main results and giving some perspectives.

A Super-resolution overview

A.1 Notations

The underlying high resolution (HR) image of size $d_1p_1 \times d_2p_2$ is written in lexicographic order (for instance, lines after lines) as a vector of pixels values $\mathbf{x} = (x_1, ..., x_q)^T$, where $q = d_1p_1d_2p_2$, and d_1 and d_2 are respectively the line and column downsampling factors. We consider n LR observations. The vector of pixels values $\mathbf{y_k} = (y_{k1}, ..., y_{kp})^T$ denotes the k^{th} LR observation written in lexicographical order with $p = p_1p_2$ and $k = 1...n$.

A.2 Observation model

We assume that $\mathbf{x}$ does not change during the acquisition of the n LR images so that we have

$$\mathbf{y_k} = \mathbf{DB_kM_kx} + \mathbf{n_k},\ k = 1...n. \tag{IV.1}$$

The variable $\mathbf{M_k}$ is a warp matrix of size $q \times q$. It represents the motions of the observations relative to each other and those in general need to be estimated. The variable $\mathbf{B_k}$ is a blur matrix of size $q \times q$. It accounts for different blurs (the atmosphere blur, which is particularly considerable for ground based telescopes, the system optics blur, the imaging system shaking etc.). The variable $\mathbf{D}$ is a matrix of size $p \times q$, which simply realizes a downsampling operation. Finally, $\mathbf{n_k}$ is a noise vector of q elements. This model is illustrated in Fig. ***IV.1***.

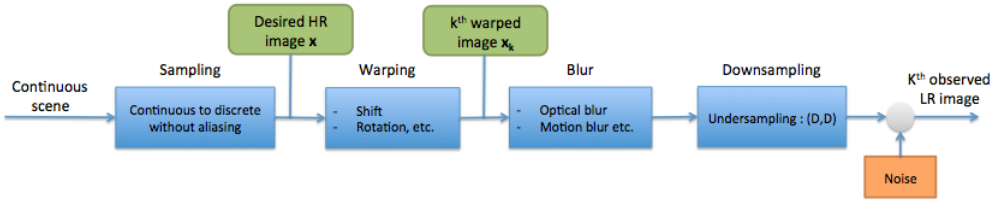


Figure *IV.1* – General observation model. See Section A.2 for a detailed description

In our case, we are interested in estimating the PSF, or in other terms, the telescope's contribution to the blur. We consider that the PSF varies slowly in the field so that the blocks "Warping" and "Blur" in Fig. ***IV.1*** may be swapped, for slow motions between observations. Therefore, the block diagram can be adapted as in Fig. ***IV.2***. This model still holds in presence of atmospheric blur (for ground-based telescopes) and jitter movements, if the LR images are extracted from the same exposure.

In the general case, the model (IV.1) may simply be written as

$$\mathbf{y_k} = \mathbf{W_kx} + \mathbf{n_k},\ k = 1...n, \tag{IV.2}$$

where $\mathbf{W_k}$ is a $p \times q$ matrix accounting for warping, blur, and downsampling.

A.3 SR techniques in astronomy

Generally, SR techniques involve three steps, which may be combined or performed separately. The registration step consists in evaluating the relative motions between different observations, so that their samples can be arranged

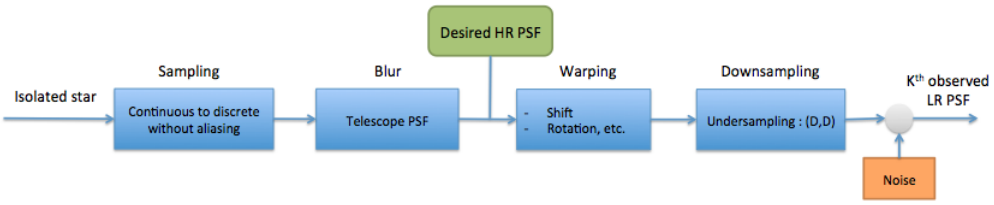


Figure *IV.2* – Adapted observation model. This time we consider an isolated star as an input, and unlike in the general model, the output of the block "Blur" is the PSF.

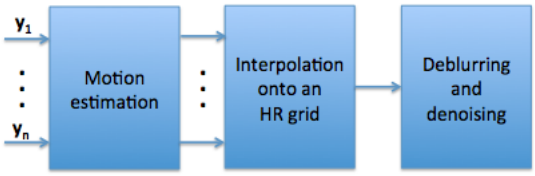


Figure *IV.3* – SR general scheme. The deblurring does not apply to our case, since we want to precisely estimate the blur.

on a common grid. It is critical that the precision of this registration should be smaller than the pixels dimensions over the target upsampling factors. Since these relative motions are arbitrary, this grid is non-uniform. The next step would be to interpolate this grid in such a way to get a regularly sampled HR image. This image is blurry and noisy. Therefore, the final step is a restoration procedure. This scheme is summarized in Fig. ***IV.3***. In the next sections, we describe some SR techniques dedicated to astronomical images, but one may refer to [Park *et al.* 2003] for more details on various SR frameworks in general image processing literature.

A.3.1 Shift-and-add method

The most simple super-resolution method is certainly the shift-and-add method. It is performed in three steps. First, the images are upsampled to the target resolution. Then, they are shifted on a common grid and averaged. It has been used in astronomy for a long time, particularly for ground based telescopes. This method is simple, fast and is used for comparisons in the numerical experiments part. It has been shown that this method provides an optimal solution in the sense of the maximum likelihood (ML) with additive white Gaussian noise (WGN) and when only pure integer translation motions

are considered with respect to the finer grid [Elad & hel Or 2001]. The interpolation operator should be $\mathbf{D}^T$ (from Eq. IV.1), which comes down to a simple zero-padding interpolation. It has been shown in the same work that the matrix $R = \sum_{k=1}^{n} \mathbf{M_k^T D^T D M_k}$ is diagonal in this simple case. Thus, after registration and stacking of the interpolated images, each pixel value should be divided by the corresponding R diagonal coefficient.

A.3.2 PSFEx method

The software PSFEx is an open source program, which has been used in many projects such as the Dark Energy Survey [Mohr *et al.* 2012] or CFHTLS [1] for PSF modeling. It takes a catalog of objects extracted from an astronomical image using SEXTRACTOR as input, which is also an open source tool for point sources (or stars) extraction. This catalog contains information about extracted point sources, such as SNR, luminosity, full width at half maximum (FWHM) , centroid coordinates, multiple flags related to saturation or blending, etc. Based on these measurements (performed in SEXTRACTOR) and some user provided parameters, PSFEx selects which sources are proper for PSF modeling. Afterwards, it constructs a PSF model, provides a fitting with an analytic function, and computes some of the PSF geometrical features. The PSF model construction may simply consist in optimally combining the input sources images for denoising, but it may also involve SR if these images are also undersampled. This SR functionality is our second reference for comparisons. These codes and the associated documentation may be found on the website `http://www.astromatic.net/`.

The desired HR image is now a matrix $\mathbf{X}$ of size $dp \times dp$, where d is the downsampling factor and the k^{th} observation $\mathbf{Y_k}$ is a matrix of size $p \times p$ with $k = 1...n$. The coordinates of the centroid of the k^{th} observation are denoted (i_k, j_k). Assuming that the images are bandlimited, the samples of the LR images can be interpolated from the desired HR image, thanks to the Shannon sampling theorem [Shannon 1948]. In theory, this interpolation should involve a 2D sinus cardinal (sinc) kernel with an infinite support, which is not convenient for practical implementation. One can instead use a support compact function, which approximates the sinus cardinal. Let $h(.,.)$ denote such a function. The estimate of the sample (i, j) of the k^{th} observation is given by

$$\hat{y}_{k,ij} = \sum_{l} \sum_{m} h\left[l - d(i - i_k), m - d(j - j_k)\right] x_{ij}. \tag{IV.3}$$

1. `http://terapix.iap.fr/cplt/T0007/doc/T0007-doc.html`

Then, we can define the cost function

$$J_1(\mathbf{X}) = \sum_{k=1}^{n} \sum_{i=1}^{p} \sum_{j=1}^{p} \frac{(y_{k,ij} - f_k \hat{y}_{k,ij})^2}{\sigma_k^2} \tag{IV.4}$$

where f_k accounts for possible luminosity differences. The parameter σ_k^2 is related to the local background variance and any other uncertainty on the pixel value. These parameters need to be estimated.

The PSFEx uses a Lanczos4 interpolant, which is defined in 1D as

$$h_{1D}(x) = \begin{cases} 1 & \text{if } x = 0 \\ \operatorname{sinc}(x)\operatorname{sinc}(x/4) & \text{if } 0 < |x| < 4 \\ 0 & \text{else,} \end{cases} \tag{IV.5}$$

so that $h(x, y) = h_{1D}(x)h_{1D}(y)$.

Finally, PSFEx minimizes the cost function defined as

$$J_2(\mathbf{\Delta}) = J_1(\mathbf{\Delta} + \mathbf{X}^{(0)}) + \lambda \|\mathbf{\Delta}\|_2^2, \tag{IV.6}$$

where $\mathbf{\Delta} = \mathbf{X} - \mathbf{X}^{(0)}$ and $\mathbf{X}^{(0)}$ is a median image computed from the LR observations. This second term is meant for regularization purposes if the problem is ill-conditioned or undetermined ($n < d^2$). One particularity of point-source images (see Fig. ***IV.4***) is that one knows that the light comes from a single point, which is generally inferred to be the light blob's centroid. Thus, one only needs the images centroid coordinates to perform the registration, which is implicitly done in Eq. IV.3.

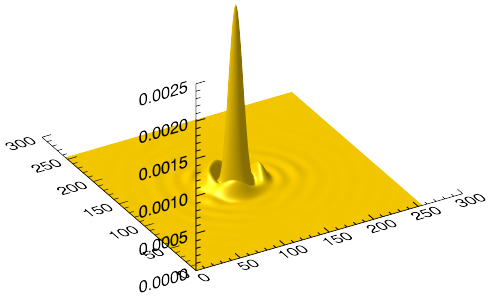


Figure *IV.4* – Simulated optical point-spread function.

B Sparse regularized method

B.1 Method

Let consider Eq. IV.3. It can be rewritten as

$$\hat{\mathbf{y}}_k = \mathbf{D}\mathbf{H}_k\mathbf{x}, \tag{IV.7}$$

where $\mathbf{x}$ and $\hat{\mathbf{y}}_k$ are, respectively, the desired matrix and the k^{th} observation estimate written this time as column vectors in lexicographic order (we used the lines order) ; besides, $\mathbf{H}_k$ is a Toeplitz matrix [Gray 2005] of size $d^2p^2 \times d^2p^2$, which contains the values of the kernel $h(.,.)$ appearing in Eq. IV.3 and $\mathbf{D}$ is a decimation matrix of size $p^2 \times d^2p^2$ with d being the downsampling factor. We are assumming that the images are squared for convenience but they might be non-squared as well. In the same way, we can redefine the objective function of Equation IV.4 as

$$J_1(\mathbf{x}) = \frac{1}{2}\sum_{k=1}^{n} \|\mathbf{y_k} - f_k\mathbf{D}\mathbf{H}_k\mathbf{x}\|_2^2/\sigma_k^2, \tag{IV.8}$$

where $\mathbf{y_k}$ is the k^{th} observation rewritten consistently with $\mathbf{x}$ and $\hat{\mathbf{y}}_k$. The function $J_1(\mathbf{x})$ is nothing but the negative log-likelihood associated with the observation model in the case of an uncorrelated Gaussian noise that is stationary for each observation up to a positive scalar factor. It can be written in an even more compact way as

$$J_1(\mathbf{x}) = \frac{1}{2}\|\mathbf{\Sigma}^{-1}\mathbf{y} - \mathbf{\Sigma}^{-1}\mathbf{F}\mathbf{W}\mathbf{x}\|_2^2, \tag{IV.9}$$

where $\mathbf{W}$ is obtained by concatening vertically the matrices $\mathbf{D}\mathbf{H}_k$, $\mathbf{F}$ is a diagonal matrix constructed by repeating the coefficients f_k p^2 times for $k = 1...n$, and $\mathbf{\Sigma}$ is constructed the same way using the coefficients σ_k. Therefore, we can simply write

$$J_1(\mathbf{x}) = \frac{1}{2}\|\mathbf{z} - \mathbf{M}\mathbf{x}\|_2^2, \tag{IV.10}$$

where $\mathbf{M}$ is a matrix of size $np^2 \times d^2p^2$. The SPRITE method constrains the minimization of this objective function using an analysis prior. Instead of using a single lagrangian multiplier as in Eq. II.41, the analysis coefficients has individual weights $\kappa\lambda_i$, where i is the coefficient index in the transform domain. This leads to the following formulation of the problem :

$$\min_{\mathbf{\Delta}} J_1(\mathbf{\Delta} + \mathbf{x}^{(0)}) + \kappa\|\boldsymbol{\lambda} \odot \mathbf{\Phi}\mathbf{\Delta}\|_1, \tag{IV.11}$$

where $\mathbf{\Delta} = \mathbf{x} - \mathbf{x}^{(0)}$ is defined as in IV.6, $\mathbf{x}^{(0)}$ is a first guess, and $\boldsymbol{\lambda}$ is now a vector of the same size as $\mathbf{\Phi}\mathbf{\Delta}$, $\odot$ denoting the pointwise product.

Additionally, the PSF or equivalently the telescope optical impulse response is by definition a positive valued function. Therefore, we want the reconstruction $\mathbf{x}$ to have positive entries. This additional constraint is integrated as follows :

$$\min_{\boldsymbol{\Delta}} J_1(\boldsymbol{\Delta} + \mathbf{x}^{(0)}) + \kappa \|\boldsymbol{\lambda} \odot \boldsymbol{\Phi}\boldsymbol{\Delta}\|_1 \; s.t. \boldsymbol{\Delta} \geq -\mathbf{x}^{(0)}, \tag{IV.12}$$

where $\geq$ is a pointwise inequality, or equivalently,

$$\min_{\boldsymbol{\Delta}} J_1(\boldsymbol{\Delta} + \mathbf{x}^{(0)}) + \kappa \|\boldsymbol{\lambda} \odot \boldsymbol{\Phi}\boldsymbol{\Delta}\|_1 + \iota_{\mathcal{S}_{\mathbf{x}^{(0)}}}(\boldsymbol{\Delta}), \tag{IV.13}$$

where $\mathcal{S}_{\mathbf{x}^{(0)}}$ is the set of vectors $\mathbf{t} \in \mathbb{R}^{d^2p^2}$ satisfying the pointwise inequality $\mathbf{t} \geq -\mathbf{x}^{(0)}$ and $\iota_{\mathcal{S}_{\mathbf{x}^{(0)}}}$ is its indicator function (see Appendix B). The impact of this constraint is emphasized in Appendix C.

As we show in the Section B.3.3, the choice of the parameters $\boldsymbol{\lambda}$ and κ relies on the noise expected on the analysis coefficients of the solution estimate. The choice of a vector regularization parameter rather than a single scalar is precisely motivated by the fact that this noise might be non-stationary.

As stated before, the problem IV.11 is not equivalent to its synthesis version, if the dictionary $\boldsymbol{\Phi}$ is redundant. The synthesis prior is expected to be efficient if the desired solution can be accurately written as a sparse linear combination of the chosen dictionary atoms, which we cannot assume to be true for every PSF profiles. In contrast, the analysis prior appears to be more flexible. Moreover, in the cases where the problem would be ill-conditioned or underdetermined, the analysis prior would definitely be more suitable since it involves far less variables.

A similar ℓ_1 penalty has already been applied for SR. An example may be found in [Yamagishi *et al.* 2012], where the cost function is minimized using variants of the alternating direction method of multipliers (ADMM). Moreover, advantages of such approaches over quadratic regularizers have been shown in many related problems in image and signal processing.

The use of the ℓ_1 norm as a relaxation for an l_0 penalty has a well-known drawback, which is that it tends to bias the solution. Indeed with a ℓ_1 norm penalty, the problem resolution involves soft thresholding operations, which affect both weak and strong entries unlike a hard thresholding, which would only affect the weak and therefore unwanted entries. Formal definitions of soft and hard thresholding are given in Appendix B.

This is particularly unsuitable for scientific data analysis. The reweighting ℓ_1 minimization proposed in [Candès *et al.* 2008] is one way to tackle this issue, while staying in the proof of convergence sets. Indeed, it consists of solving a succession of ℓ_1 minimization problems of the form

$$\min_{\boldsymbol{\Delta}} J_1(\boldsymbol{\Delta} + \mathbf{x}^{(0)}) + \kappa \|\mathbf{w}^{(k)} \odot \boldsymbol{\lambda} \odot \boldsymbol{\Phi}\boldsymbol{\Delta}\|_1 + \iota_{\mathcal{S}_{\mathbf{x}^{(0)}}}(\boldsymbol{\Delta}), \tag{IV.14}$$

where $\mathbf{w}^{(k)}$ is a weighting vector for the transform coefficients at the k^{th} minimization. Each entry of $\mathbf{w}^{(k)}$ is calculated as a decreasing function of the corresponding transform coefficient magnitude in the $(k-1)^{th}$ minimization. This way, the strong transform coefficients are less penalized than the weaker ones in the new minimization. One may refer to Appendix C for quantitative study of the reweighting effect.

B.2 Algorithm

In the proposed method, the reweighting scheme is performed according to [Candès *et al.* 2008] :

1. Set $k = 0$, for each entry in the weighting vector $\mathbf{w}^{(k)}$, set $w_j^{(k)} = 1$.
2. Solve the problem IV.14 yielding a solution $\boldsymbol{\Delta}^{(k)}$.
3. Compute $\boldsymbol{\alpha}^{(k)} = \boldsymbol{\Phi}\boldsymbol{\Delta}^{(k)}$.
4. Update the weight vector according to $w_j^{(k+1)} = \frac{1}{1+|\alpha_j^{(k)}|/3\sigma_j}$, where σ_j is the noise standard deviation expected at the j^{th} transform coefficient, see section B.3.3.
5. Terminated on convergence or when reaching the maximum number of iterations ; otherwise, go to step 2.

The step 2 resolution is detailed in Algorithm IV.1. We use the generalized forward-backward splitting introduced in [Raguet *et al.* 2011]. It requires the computation of proximity operators associated with the regularization functions in IV.14. One may refer to Appendix B for an introduction to proximal calculus.

Since the dictionary $\boldsymbol{\Phi}$ is redundant, we do not have a closed-form expression for the proximity operator. Yet, it can be calculated as

$$\begin{cases} \operatorname{prox}_{\frac{\mu}{\omega_1}\kappa\|\operatorname{Diag}(\mathbf{w}^{(k)}\odot\boldsymbol{\lambda})\boldsymbol{\Phi}\|_1}(\mathbf{x}) = \mathbf{x} - \boldsymbol{\Phi}^T\hat{\mathbf{u}} \\ \hat{\mathbf{u}} = \operatorname{argmin}_{|u_j|<\frac{\mu}{\omega_1}\kappa w_j^{(k)}\lambda_j} \frac{1}{2}\|\mathbf{x} - \boldsymbol{\Phi}^T\mathbf{u}\|_2^2 \end{cases} \tag{IV.15}$$

where $\hat{\mathbf{u}}$ can be estimated using a forward-backward algorithm [Bauschke *et al.* 2011] as follows :

1. Set $p = 0$, initialize $\mathbf{u_0} = 0$.
2. $\tilde{\mathbf{u}}_{p+1} = \mathbf{u}_p + \mu_{\text{prox}}\boldsymbol{\Phi}\left(\mathbf{x} - \boldsymbol{\Phi}^T\mathbf{u_p}\right)$.
3. $\mathbf{u}_{p+1} = \tilde{\mathbf{u}}_{p+1} - \operatorname{SoftThresh}_{\frac{\mu}{\omega_1}\kappa\mathbf{w}^{(k)}\odot\boldsymbol{\lambda}}\tilde{\mathbf{u}}_{p+1}$.
4. Terminate on convergence or when reaching the maximum number of iterations, otherwise go to step 2.

Algorithme IV.1 Weighted analysis-based $\boldsymbol{\Delta}^{(k)}$ recovery

Paramètre(s) d'entrée :

A first guess estimate of the super-resolved image $\mathbf{x}^{(0)}$.
A weight vector $\mathbf{w}^{(k)}$.
Sparsity constraint parameter κ.
A dictionary $\boldsymbol{\Phi}$.
Auxiliary variables $\mathbf{z}_{10}, \mathbf{z}_{20} \in \mathbb{R}^{d^2p^2}$.
$\omega_1, \omega_2 \in]0,1[$ *s.t.* $\omega_1 + \omega_2 = 1, \lambda > 0$ (see Section B.3.4).

1: Initialize $\mathbf{d}_0 = \omega_1 \mathbf{z}_{10} + \omega_2 \mathbf{z}_{20}$.
2: **Pour** $n = 0$ to $N_{\max} - 1$ **faire**
3: $\quad \mathbf{z}_{1n+1} = \mathbf{z}_{1n} + \lambda(\text{prox}_{\frac{\mu}{\omega_1}\kappa\| \text{Diag}\,(\mathbf{w}_k \odot \boldsymbol{\lambda})\boldsymbol{\Phi}\|_1}(2\mathbf{d}_n - \mathbf{z}_{1n} - \mu\nabla J_1(\mathbf{d}_n + \mathbf{x}^{(0)}) - \mathbf{d}_n)$
4: $\quad \mathbf{z}_{2n+1} = \mathbf{z}_{2n} + \lambda(\text{prox}_{\frac{\mu}{\omega_2}\iota_{\mathcal{S}_{\mathbf{x}^{(0)}}}}(\boldsymbol{\Delta})(2\mathbf{d}_n - \mathbf{z}_{2n} - \mu\nabla J_1(\mathbf{d}_n + \mathbf{x}^{(0)}) - \mathbf{d}_n)$
5: $\quad \mathbf{d}_{n+1} = \omega_1 \mathbf{z}_{1n+1} + \omega_2 \mathbf{z}_{2n+1}$
6: **Return :** $\boldsymbol{\Delta}^{(k)} = \mathbf{d}_{N_{\max}}$.

The thresholding operator SoftThresh is defined in Appendix B. The operator $\text{prox}_{\frac{\mu}{\omega_2}\iota_{\mathcal{S}_{\mathbf{x}^{(0)}}}}$ is simply the orthogonal projector onto the set $\mathcal{S}_{\mathbf{x}^{(0)}}$ defined in the previous section ; it is given explicitly Appendix B.

A full description of SPRITE is provided in Algorithm IV.2.

Algorithme IV.2 SPRITE : weigthed analysis-based super-resolution

Paramètre(s) d'entrée :

Sparsity constraint parameter κ.
A dictionary $\boldsymbol{\Phi}$.
An upsampling factor.

1: Estimate the data fidelity parameters (see Section B.3.1).
2: Calculate a first guess $\mathbf{x}^{(0)}$.
3: Initialize $\mathbf{w}^{(k)} = 1$.
4: Calculate a step size μ.
5: **Pour** $k = 0$ to $K_{\max} - 1$ **faire**
6: $\quad \boldsymbol{\Delta}^{(k)} = \underset{\boldsymbol{\Delta}}{\text{argmin}}\ J_1(\boldsymbol{\Delta} + \mathbf{x}^{(0)}) + \kappa\|\mathbf{w}^{(k)} \odot \boldsymbol{\lambda} \odot \boldsymbol{\Phi}\boldsymbol{\Delta}\|_1 + \iota_{\mathcal{S}_{\mathbf{x}^{(0)}}}(\boldsymbol{\Delta})$ (see Algorithm IV.1)
7: $\quad \boldsymbol{\alpha}^{(k)} = \boldsymbol{\Phi}\boldsymbol{\Delta}^{(k)}$
8: $\quad w_j^{(k+1)} = \frac{1}{1+|\alpha_j^{(k)}|/3\sigma_j}$
9: **Return :** $\hat{\mathbf{x}} = \boldsymbol{\Delta}^{(K_{\max}-1)} + \mathbf{x}^{(0)}$.

B.3 Parameter estimation

The data fidelity term in the problem IV.14 is defined as

$$\begin{cases} J_1(\mathbf{x}) = \frac{1}{2} \sum_{k=1}^{n} \sum_{i=1}^{p} \sum_{j=1}^{p} \frac{(y_{k,ij} - f_k \hat{y}_{k,ij})^2}{\sigma_k^2} \\ \hat{y}_{k,ij} = \sum_l \sum_m h\left[l - d(i - i_k), m - d(j - j_k)\right] x_{ij} \end{cases}, \tag{IV.16}$$

where $\mathbf{x} = (x_{ij})_{1 \leq i,j \leq dp}$ is the desired image and $h(.,.)$ is a 2D Lanczos kernel. As stated in Section B.1, it can be written as

$$J_1(\mathbf{x}) = \frac{1}{2} \|\mathbf{z} - \mathbf{M}\mathbf{x}\|_2^2, \tag{IV.17}$$

if we write $\mathbf{x}$ in lexicographic order as a vector. The following parameters are required :

- for the data fidelity term parameters the noise standard deviation in the LR images σ_k, the photometric flux f_k, and the shift parameters (i_k, j_k) ;
- a first guess $\mathbf{x}^{(0)}$ (see problem IV.14) ;
- the sparsity constraint parameters κ, $\boldsymbol{\lambda} = (\lambda_i)_i$ and the dictionary $\boldsymbol{\Phi}$;
- algorithmic parameter such as the gradient step size μ, the relaxation parameter λ in Algorithm IV.1, and the gradient step size μ_{prox} in Equation IV.15 resolution.

B.3.1 Data fidelity parameters

Noise standard deviations At the first step of the algorithm, the noise standard deviations in the low resolution images can be robustly estimated using the median absolute deviation (MAD) estimator [Starck *et al.* 2015].

Subpixel shifts The subpixel shifts between the images are estimated based on the low resolution images centroids positions. Those are calculated on the low resolution images after a hard thresholding operation. For the image $\mathbf{x}_i$, the threshold is chosen as

$$k = \min(4\sigma_i, \left(\frac{\max(|\mathbf{x}_i|)}{\sigma_i} - 1\right) \sigma_i). \tag{IV.18}$$

In this way, we only keep pixels with a high SNR for the centroid estimation. We then estimate the centroid positions using the iteratively weighted algorithm introduced in [Baker & Moallem 2007]. The thresholding operation undoubtedly biases the estimated centroid position, but the resulting estimated shifts are expected to be unbiased, up to the finite sampling and noise effects.

Photometric flux The flux parameters are calculated by integrating the low resolution images on a fixed circular aperture centered on their centroids estimates. At Euclid resolution (see Section C), we obtained quite accurate flux estimates in simulations using a radius of 3 pixels for the aperture. These parameters define the matrix $\mathbf{M}$ in Eq. IV.17.

All these parameters are automatically calculated without requiring any user input.

B.3.2 First guess computation

As one can see in the Algorithm IV.2, the final image is computed as

$$\hat{\mathbf{x}} = \boldsymbol{\Delta}^{(K_{\max}-1)} + \mathbf{x}^{(0)}. \tag{IV.19}$$

This implies that the noise and any artifact in $\mathbf{x}^{(0)}$ which does not have a sparse decomposition in $\boldsymbol{\Phi}$ will be present in the final solution. Therefore, one has to be careful at this step. To do so, we compute a noisy first guess $\mathbf{x}_n^{(0)}$ using a shift-and-add, as presented in Section A.3.1. Then we apply a wavelet denoising to $\mathbf{x}_n^{(0)}$. In other terms, we transform $\mathbf{x}_n^{(0)}$ in a "sparsifying" wavelet dictionary $\mathbf{W}$. We threshold each wavelet scale in such a way to keep only the coefficients above the noise level expected in the scale. Finally, we apply a reconstruction operator, which is a dictionary $\widehat{\mathbf{W}}$ verifying $\widehat{\mathbf{W}}\mathbf{W} = \mathbf{I}_q$ to the thresholded coefficients (see [Starck *et al.* 2015]). We note $\boldsymbol{\beta} = (\beta_i)_i$, a vector made of the denoising thresholds for each wavelet scale. We set β_i at $5\sigma_i$, where σ_i is the noise standard deviation in the i^{th} wavelet scale. The first guess is finally computed as

$$\mathbf{x}^{(0)} = \widehat{\mathbf{W}}\,\mathrm{HardThresh}_{\boldsymbol{\beta}}\,\mathbf{W}\mathbf{x}_n^{(0)}, \tag{IV.20}$$

which robustly removes the noise without breaking important features. One can refer to Appendix C for $(\sigma_i)_i$ estimation.

B.3.3 The choice of dictionary and regularization parameter

Regularization parameter The regularization parameter κ can be set, according to a desired level of significance. Indeed, it can be seen that the transform domain vector $\hat{\mathbf{u}}$ is constrained into weighted l^∞-ball of radius $\mu\kappa$ in Equation IV.15 and can be interpreted as the non-significant part of the wanted signal current estimate. To set this radius according to the expected level of noise for each transform coefficient, we propagate the noise on the data vector $\mathbf{z}$ from Equation IV.17 through $\mu\boldsymbol{\Phi}\mathbf{M}^T\mathbf{M}$ and estimate its standard deviation at each transform coefficient, which sets the parameters λ_j. In practice,

this can be done in two ways. We can either run a Monte-Carlo simulation of the noise in $\mathbf{z}$ and take the empirical variance of the sets of realizations of each transform coefficient. On the other hand, if $\mathbf{\Phi}$ is a wavelet dictionary and if the noise is expected to be stationary in each wavelet scale, then we only need to compute a single standard deviation per scale. This can be done by estimating the noise in each scale of the wavelet transform of the gradient at each iteration (up to the factor μ) using a MAD, for instance. Indeed, the residual $\mathbf{z} - \mathbf{M}(\mathbf{d}_n + \mathbf{x}^{(0)})$ tend to be consistent with the noise in $\mathbf{z}$, so that it can be used as a noise realization. With a stationary noise in each input image, the two approaches give very close estimates of the noise standard deviation and the second one is far less demanding in terms of complexity. As a result, coefficients below $\kappa\lambda_j$ are considered as part of the noise and one only needs to set the global parameter κ to tune the sparsity constraint according to the noise level.

Dictionary The choice of the dictionary impacts the performance of the algorithm. We considered two transforms : a biorthogonal undecimated wavelet transform with a 7/9 filter bank and the second generation starlet transform [Starck *et al.* 2011]. These two transforms are generic and not specifically tuned to a given PSF profile.

B.3.4 Algorithmic parameters

Gradient steps sizes The gradient step size μ in Algorithm IV.1 needs to be chosen just in $\left]0, 2/\rho(\mathbf{M}^T\mathbf{M})\right[$, where $\rho(.)$ denotes the spectral radius of a square matrix. In the same way, μ_{prox} needs to be chosen in $\left]0, 2/\rho(\mathbf{\Phi}\mathbf{\Phi}^T)\right[$.

Relaxation parameter The parameter λ in Algorithm IV.1 needs to be chosen in $\left]0, \min\left(\frac{3}{2}, \frac{1+2/\rho(\mathbf{M}^T\mathbf{M})\mu}{2}\right)\right[$ [Raguet *et al.* 2011]. This parameter tunes the updating speed of the auxiliary variables in Algorithm IV.1. In practice, we use $\mu = 1/\rho(\mathbf{M}^T\mathbf{M})$ and $\lambda = 1.4$.

B.3.5 User parameters

It is important to mention that the user only has to set the parameter κ and the dictionary with the other parameters being automatically estimated. In all our experiments, we took $\kappa = 4$, which is quite convenient, if we assume Gaussian noise in the data. The dictionary choice will be emphasized in the next section.

C Numerical experiments

This section presents the data used, the numerical experiments realized as a mean to compare three SR techniques (shift-and-add, PSFEx method, and our method) and the results.

C.1 Data set

The PSFs provided are optical PSFs computed using a fast Fourier transform of the exit pupil. They are not a system PSF, so they do not include a jitter or detector response. A set of PSFs covering the whole field of view is provided. They are monochromatic PSFs at 800nm and are derived from tolerance analysis. They account for manufacturing and alignments errors and thermal stability of the telescope. Manufacturing and alignment errors are partially compensated by a best focus optimization, while thermal stability effects are simulated by a small displacement of the optics that are not compensated on a short-time scale. The optical model used is dated from 2011 and is prior to the current reference model (provided by Astrium, which has been awarded the payload module contract in 2013). In particular, the 2011 model does not contain the latest definition of the pupil mask. The pupil, however, includes central obscuration and a three-vane spider. This is the model that has been used for the science feasibility studies that led to the acceptance of the Euclid mission.

C.2 Simulation

In the Euclid mission, the actual sampling frequency is about 0.688 times the Nyquist frequency that we define as twice the telescope spatial cut-off frequency [Cropper 2013]. Therefore we target an upsampling factor of 2, which gives a sufficient bandpass to recover the high frequencies. The PSF is typically space-varying, and this is particularly true for wide field of view instruments as Euclid telescope [ESA/SRE 2011]. Thus, the data set contains simulated PSF measurements on a regular 18×18 grid on the field of view. The original PSF models are downsampled to twice Euclid resolution. For each PSF, four randomly shifted "copies" are generated and downsampled to Euclid resolution (see Fig. ***IV.5*** and Fig. ***IV.6*** below). These LR images are of size 84×84. Different levels of WGN are added. We define the signal level as its empirical variance that is calculated in a 50×50 patch centered on the HR PSF main lobe. For each algorithm, we used these four images to reconstruct a PSF which has twice their resolutions in lines and columns.

C.3 Quality criterion

For an image $\mathbf{X} = (x_{ij})_{i,j}$, the weighted central moments are defined as

$$\mu_{p,q}(\mathbf{X}) = \sum_i \sum_j (i - i_c)^p (j - j_c)^p f_{ij} x_{ij} \tag{IV.21}$$

with $(p, q) \in \mathbb{N}^2$, (i_c, j_c) are the weighted image centroid coordinates, and $\mathbf{F} = (f_{ij})_{i,j}$ is an appropriate weighting function (typically a Gaussian function). The ellipticity parameters are then defined as follows :

$$e_1(\mathbf{X}) = \frac{\mu_{2,0}(\mathbf{X}) - \mu_{0,2}(\mathbf{X})}{\mu_{2,0}(\mathbf{X}) + \mu_{0,2}(\mathbf{X})} \tag{IV.22}$$

$$e_2(\mathbf{X}) = \frac{2\mu_{1,1}(\mathbf{X})}{\mu_{2,0}(\mathbf{X}) + \mu_{0,2}(\mathbf{X})}. \tag{IV.23}$$

The vector $\boldsymbol{\varepsilon} = [e_1, e_2]$ is an important tool, since if measured on a large set of galaxies, it can be statistically related to the dark matter induced geometrical distortions and finally its mass density. Furthermore, this ellipticity parameters are magnitude invariant and approximately shift invariant. The error on ellipticity is therefore an interesting criteria for quality assessment. Thus, we used the mean absolute error for each ellipticity parameter,

$$\mathcal{E}_j = \frac{1}{n} \sum_{i=1}^{n} |e_j(\mathbf{X_i}) - e_j(\hat{\mathbf{X}}_\mathbf{i})|, \; j = 1, 2 \tag{IV.24}$$

and the associated empirical standard deviations.

Moreover, the PSF size is also an important characteristic of the PSF kernel. For example, it has been shown in [Paulin-Henriksson *et al.* 2008] that the PSF size largely contributes to the systematic error in weak gravitational lensing surveys. Therefore, we use it in quality assessment by computing the mean absolute error on the full width at half maximum (FWHM). The FWHM is estimated by fitting a modified Lorentzian function on the PSF images. We used routines from a publicly available library [2].

C.4 Results and discussion

The Figures ***IV.5*** and ***IV.6*** show a simulated PSF that is sampled at almost the Nyquist rate and the two LR shifted and noisy PSF derive, with SNR of around 30dB.

2. http://www.astro.washington.edu/docs/idl/htmlhelp/slibrary21.html

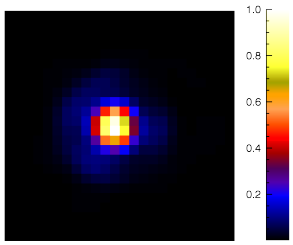


Figure *IV.5* – Critically sampled PSF.

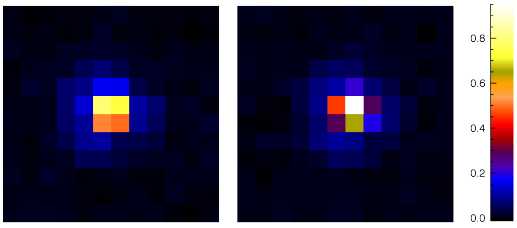
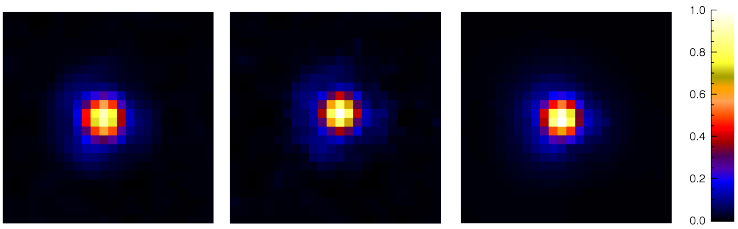


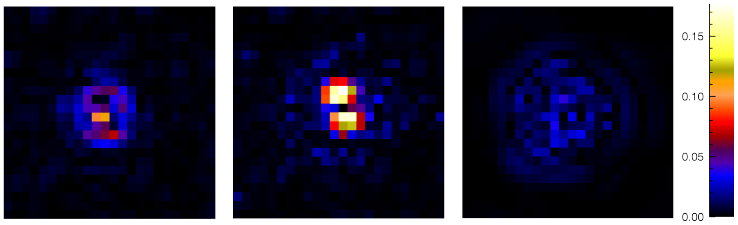
Figure *IV.6* – PSF sampled at Euclid resolution with different offsets and noise.

The Figure ***IV.7*** shows an example of super-resolved PSF at 30dB of SNR, from four LR images and the corresponding error maps that are defined as the absolute value of the difference between the original high resolution noise free PSF and the PSF reconstructions for each algorithm. This error map standard deviation is at least 30% lower with SPRITE.

Figure ***IV.8*** shows smaller errors and errors dispersions are achieved with SPRITE algorithm, especially at low SNR. One can note that the dispersion is slightly smaller with biorthogonal undecimated wavelet. The error on the FWHM given in percent on Fig. ***IV.9*** is smaller with SPRITE. In practice, there is more variability in the PSF (wavelength and spatial dependency, time variations...) so that the real problem will be more underdetermined. Thanks to the multiple exposures on the one hand, and that the spatial variations of the PSF are expected to be slow on the other hand, the real problem could actually be very well constrained. Moreover, these results suggest that even better results could be achieved by using more adapted dictionaries, built either from PSF model or through a dictionary learning algorithm [Beckouche *et al.* 2013a].



(a) Super-resolved PSFs.



(b) Absolute Error Maps.

Figure *IV.7* – PSF reconstruction and error map at 30dB for three methods : from the left to the right, shift-and-add, PSFEx, and SPRITE. The error image standard deviation is at least 30% smaller with SPRITE.

D Complexity and performances

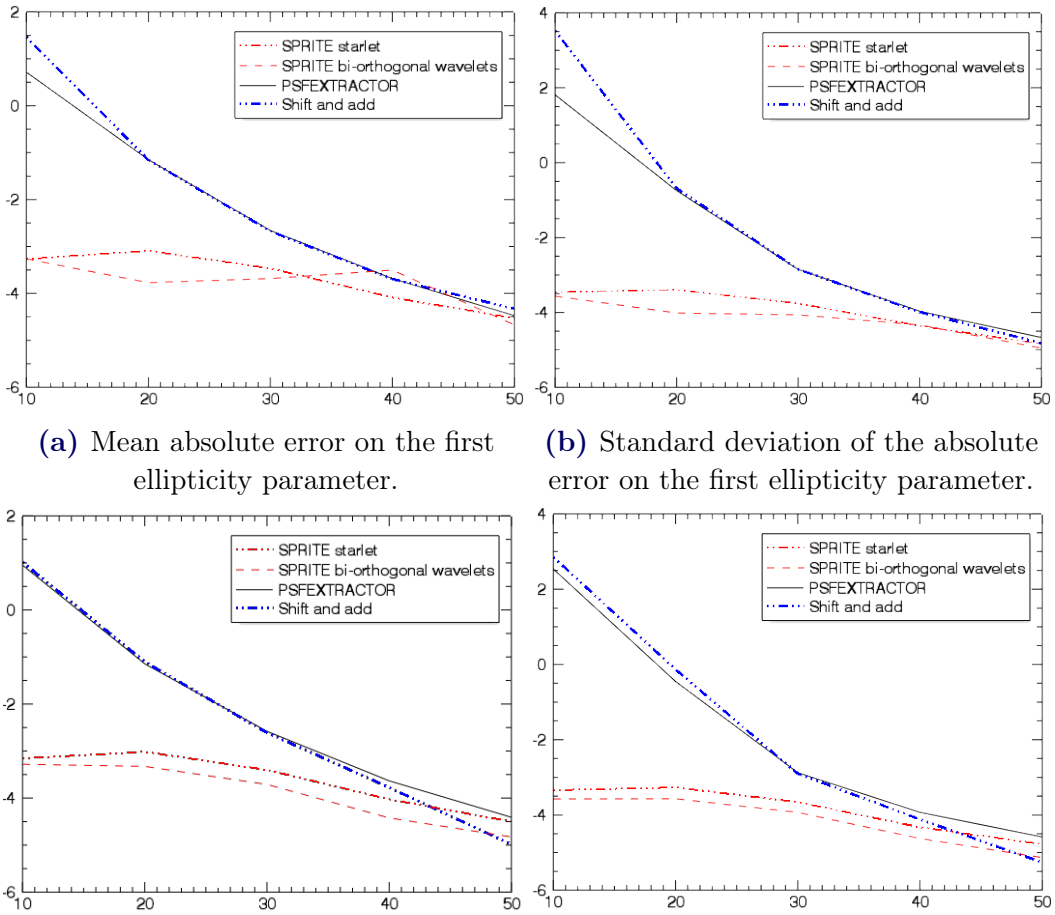
The simulations were run on a typical desktop computer. Let us suppose that we have n LR images of sizes $p_1 \times p_2$ and that we choose an upsampling factor d in lines and columns. As stated before, we took $p_1 = p_2 = 84$, $n = 4$, and $d = 2$ in our numerical experiments. Under this setting, it takes roughly $60s$ and 1GB of physical memory to compute a super-resolved PSF. More generally, the computational complexity of the algorithm is in $O(np_1p_2d^2 \log(p_1p_2d^2))$, which is related to the implementation of the matrices $\mathbf{M}$ from Eq. IV.10 and $\mathbf{M}^T$ using FFT.

E Software

Following the philosophy of reproducible research [Buckheit & Donoho 1995], the algorithm introduced in this paper and the data used are available at `http://www.cosmostat.org/sprite.html`. We used the following calls for the SPRITE executable :

— **run_sprite -t 2 -s 4 -r 2 -F -N data_file output_file output_directory** for the second genration Starlet transform ;

— **run_sprite -t 24 -s 4 -r 2 -F -N data_file output_file output_directory** for the undecimated biorthogonal wavelet transform.



(a) Mean absolute error on the first ellipticity parameter.

(b) Standard deviation of the absolute error on the first ellipticity parameter.

(c) Mean absolute error on the second ellipticity parameter.

(d) Standard deviation of the absolute error on the second ellipticity parameter.

Figure *IV.8* – Errors in log on ellipticity parameters versus the SNR. For SNR=10dB, SPRITE achieves around 6dB less than others methods, which corresponds to a factor of e^6 on a linear scale.

The options "-s" and "-r" set the parameter κ (see Section B.3.3) and the upsampling factor for both lines and columns respectively. The options "-F" and "-N" indicate that the photometric flux and the noise might have different levels in the LR images and need to be estimated.

F Conclusion

We introduced SPRITE, which is a super-resolution algorithm based on sparse regularization. We show that adding a sparse penalty in the recovery leads to far better accuracy in terms of ellipticity error, especially at low SNR.

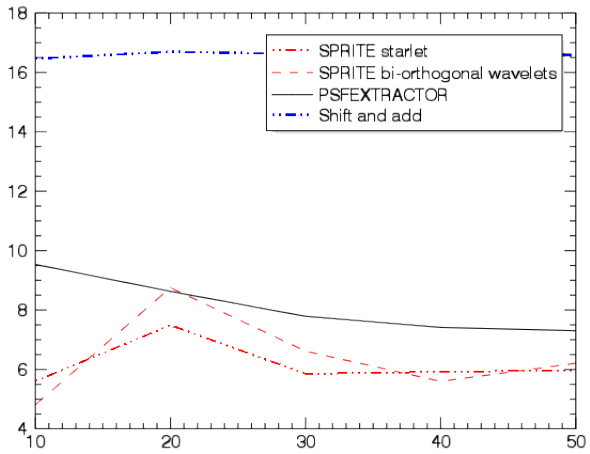


Figure *IV.9* – Mean absolute error on the full width at half maximum (FWHM) in percent. SPRITE achieves on average an error of 6% on the FWHM which is 2% less than PSFEX in average.

Quantitatively, we achieved

— a 30% lower error on the reconstruction itself at 30dB of SNR ;

— around 6dB less than other methods on the shape parameters, which corresponds to a factor of e^6 on a linear scale, at 10dB of SNR ;

— 6% of error on the FWHM in average, which 2% less than PSFEX.

However this algorithm does not handle the PSF spatial variations. Thus one natural extension of this work would be to simultaneously perform super-resolution and dimensionality reduction assuming only one LR version of each PSF, as it is the case in practice strictly speaking, but by using a large PSF set. This is precisely the subject of the next chapter.

Chapitre V

Single frame PSFs field super-resolution

Sommaire

In the previous chapter, the PSFs super-resolution (SR) where performed, assuming that multiple measurements of the same PSF were available. However in practice, one does not have such multiple measurements strictly speaking, because the PSF is space and time variant, as emphasized in Chapter I.

In this chapter, we consider precisely a setting where the PSF is space variant and we want to get an accurate modeling at high resolution of the PSF

field, assuming we have under-sampled measurements of different PSFs in the observed field. We assume that the PSFs vary slowly across the field. Intuitively, this implies a compressibility of the PSFs field, which leads us to the question of what would be a concise and easily understandable representation of a spatially indexed set of PSFs.

A The PSF field model

A.1 The observation model

We assume that we have an image I, which contains p unresolved objects such as stars, which can be used to estimate the PSFs field. Noting $\mathbf{y}_k$ one of these p objects at spatial position $\mathbf{u}_k$, $\mathbf{y}_k$ is therefore a small patch of I with n_y pixels, around the spatial position $\mathbf{u}_k$. We will write $\mathbf{y}_k$ as a 1D vector. The relation between the "true" PSF $\mathbf{x}_k$ and the noisy $\mathbf{y}_k$ observation is

$$\mathbf{y}_k = \mathbf{M}_k \mathbf{x}_k + \mathbf{n}_k \tag{V.1}$$

where $\mathbf{M}_k$ is a linear operator and $\mathbf{n}_k$ is a noise that we assume to be Gaussian and white. We will consider two kinds of operators in this paper : the first one is the simple case where $\mathbf{M}_k = \mathbf{I}_{n_x}$ and we have the number of pixels n_x in $\mathbf{x}_k$ is equal to n_y, and the second one is a shift+downsampling degradation operator and $n_x = m_d^2 n_y$, where m_d is the downsampling factor in lines and columns, with $m_d \geq 1$.

Noting $\mathbf{Y} = [\mathbf{y}_1 \dots \mathbf{y}_p]$ the matrix of n_y lines and p columns of all observed patches, $\mathbf{X} = [\mathbf{x}_1 \dots \mathbf{x}_p]$ the matrix $n_x \times p$ of all unknown PSFs, we can rewrite Eq. V.1 as

$$\mathbf{Y} = \mathcal{F}(\mathbf{X}) + \mathbf{N} \tag{V.2}$$

where $\mathcal{F}(\mathbf{X}) = [\mathbf{M}_1 \mathbf{x}_1, \dots, \mathbf{M}_p \mathbf{x}_p]$.

This rewriting is useful because, as we discuss in the following, the different PSFs $\mathbf{x_k}$ are not independent, which means that the problems of Eq. V.1 should not be solved independently for each k. In other terms, the vectors $(\mathbf{x}_k)_{1 \leq k \leq p}$ belong to a specific unknown manifold that needs to be learned by using the data globally.

A.2 The data model

Let $\mathbf{\Omega}$ be a r dimensional subspace of $\mathbb{R}^{n_x}$ embedding the PSFs field. We assume that there exists a continuous function $f : \mathcal{E} \mapsto \mathbf{\Omega}$, so that $f(\mathbf{u}_k) = \mathbf{x}_k$, $\forall k \in [\![1, p]\!]$. The regularity of f translates the correlation of the data in space (and time).

Let $(\mathbf{s}_i)_{1\leq i\leq r}$ be a basis of $\mathbf{\Omega}$. By definition, we can write each $\mathbf{x}_k$ as a linear combination of the $\mathbf{s}_i$, $\mathbf{x}_k = \sum_{i=1}^{r} a_{ik}\mathbf{s}_i$, $k = 1 \dots p$, or equivalently

$$\mathbf{X} = \mathbf{SA} \tag{V.3}$$

where $\mathbf{S} = [\mathbf{s}_1, \dots, \mathbf{s}_r]$ and $\mathbf{A}$ is a $r \times p$ matrix containing the coefficients $\mathbf{A}[:, k]$ of the vectors $\mathbf{x}_k$ $(k = 1 \dots p)$ in the dictionary $\mathbf{S}$. Each column of the matrix $\mathbf{S}$, that we also refer to as an atom, can be seen as an eigen PSF, i.e. a given PSF's feature distributed across the field.

A.3 The inverse problem

We need therefore to minimize $\|\mathbf{Y} - \mathcal{F}(\mathbf{X})\|_F^2$, which is an ill posed problem due to both the noise and the operator $\mathcal{F}$, $\|.\|_F$ denoting the Frobenius norm of a matrix. There are several constraints that may be interesting to use in order to properly regularize this inverse problem :

— positivity constraint : the PSF $\mathbf{x}_k$ should be positive ;

— low rank constraint : as described above, we can assume that $\mathbf{x}_k = \sum_{i=1}^{r} a_{ik}\mathbf{s}_i$, which means that we can instead minimize

$$\min_{\mathbf{A},\mathbf{S}} \|\mathbf{Y} - \mathcal{F}(\mathbf{SA})\|_F^2; \tag{V.4}$$

we assume that $r \ll \min(n, p)$; this dimension reduction has the advantage that there are much less unknown to find, leading to more robustness, but the problem is now that the cost function is not convex anymore ;

— smoothness constraint : we can assume that the vectors $\mathbf{x}_k$ are structured ; the low rank constraint does not necessarily impose $\mathbf{x}_k$ to be smooth or piece-wise smooth ; adding an additional constraint on $\mathbf{S}$ atoms, such as a sparsity constraint, allows to capture spatial correlations within the PSFs themselves ; an additional dictionary $\mathbf{\Phi}_s$ can therefore be introduced which is assumed to give a sparse representation of the vectors $\mathbf{s}_k$;

— proximity constraint : we can assume that a given $\mathbf{x}_k$ at a position $\mathbf{u}_k$ is very close to another PSF $\mathbf{x}_{k'}$ at position $\mathbf{u}_{k'}$ if the distance between $\mathbf{u}_k$ and $\mathbf{u}_{k'}$ is small ; this means that the field f must be regular ; this regularity can be forced by adding constraints on the lines of the matrix $\mathbf{A}$; indeed, the p values relative to a line $\mathbf{A}[i, :]$ correspond to the contribution of the ith eigen PSF to locations relative to the spatial positions $\mathcal{U}$.

We show in section C how these four constraints can be jointly used to derive the solution. Let first review existing methods susceptible to solve this problem.

B Related work

In all this section, $\mathcal{Y}$ refers to the observed data set $(\mathbf{y}_k)_{1\leq k\leq p}$. In the first part, the aforementioned degradation operator $\mathcal{F}$ is simply the identity. Therefore we review some dimension reduction methods. In the second part $\mathcal{F}$ is a shifting and downsampling operator ; we present a PSF modeling software dealing with this more constraining setting.

B.1 Dimension reduction

The principal components analysis is certainly one of the most popular mathematical procedure in multivariate data analysis and especially, dimension reduction. In our case, we want to represent $\mathcal{Y}$'s elements using r vectors, with $r \leq \max(p, n_y)$. A PCA gives an orthonormal family of r vectors in $\mathbb{R}^{n_y}$ so that the total variance of $\mathcal{Y}$ along these vectors directions is maximized. By definition, the PCA looks for redundant features over the whole data set. Therefore, in general, the principal components neither capture localized features (in sense of $\mathcal{E}$) nor have a simple physical interpretation.

In [Wang & Huang 2015], a "regularized" PCA is proposed to address this shortcoming for spatial data analysis in atmospheric and earth science. Indeed, as a PCA, the method solves the following problem,

$$\min_{\mathbf{A}} \|\mathbf{Y} - \mathbf{Y}\mathbf{A}^T\mathbf{A}\|_F^2, \text{ s. t. } \mathbf{A}\mathbf{A}^T = \mathbf{I}_r, \tag{V.5}$$

for some chosen small r. Moreover, it jointly imposes a sparsity constraint and a smoothing penalties with respect to the space $\mathcal{E}$, on the matrix $\mathbf{A}$ lines. This way, with the right balance between those two penalties, one favors the extraction of localized spatial features, making the interpretation of the optimal $\mathbf{A}$ easy. Yet, there is no obvious way of setting the sparsity and smoothness parameters, which are crucial ; moreover, unless the data actually contain spatially localized and non-overlapping features, the coupled orthogonality and sparsity constraint is likely to yield a biased approximation of the data.

In the context of remote sensing and multi-channel imaging, two ways of integrating spatial information into PCA are proposed in [Cheng 2006] ; the set $\mathcal{Y}$ is made of multi-channel pixels. In the first way, the author introduces a weighting matrix indicating the relative importance of each pixel. For instance, the weight of a given pixel can be related to its distance to some location

of interest in $\mathcal{E}$. Then, the computation of the covariance matrix of image bands is slightly modified to integrate this weighting. This idea is close to the methodology proposed in [Harris *et al.* 2011]. As a consequence, one expects to recover spectral features spatially related to some location of interest within the most important "eigen-pixels". Yet, we do not have any specific location of interest in $\mathcal{E}$ and we rather want to recover relevant features across the whole data set.

The second approach aims at taking into account the spatial associations and structural properties of the image. To do so, modified versions of the image bands covariance matrices are calculated, with increasing shifts between the bands, up to a predetermined maximum shifting amplitude. These covariance matrices, including the "regular" one, are averaged and the principal components are finally derived. Intuitively, one expects the spectral features present in structured images regions to be strengthened and therefore captured into the principal components. However, we consider a general setting where the data are randomly distributed with respect to $\mathcal{E}$, which makes the shifted covariances matrices ill-defined.

A review of PCA applications and modifications for spatial data analysis can be found in [Demšar *et al.* 2013].

In case the data lie on or are close to a manifold $\mathcal{M}$ of dimension r embedded in $\mathbb{R}^n$, one can consider using one of the numerous non-linear dimension reduction algorithms published in the manifold learning literature, such as GMRA [Allard *et al.* 2012], [Maggioni *et al.* 2014]. The idea is to partition the data in smaller subsets of sample close to each other in the sense of the manifold geometry. From this partionning, the manifold tangent spaces are estimated at subsets locations; estimates are then simply given by the best regressions of these subsets with $r-$dimensional affine subspaces. The method includes some multiresolution considerations that are not relevant to our problem. This procedure provides a dictionary in which each of the original samples need at most r elements to be represented. Moreover, the local processing of the data, which is necessary in this setting because of the manifold curvature, makes this approach somehow compatible with the considered problem. Indeed, by hypothesis, the closer two samples will be in sense of $\mathcal{E}$, the closer they will be in $\mathbb{R}^n$, and the more likely they will fall into the same local cluster.

Another interesting alternative to the PCA can be found in [Lee *et al.* 2008]. This construction called "Treelets" extracts features by uncovering correlated subsets of variables across the data samples. It is particularly useful when the sample size is by far smaller than the data dimensionality ($p \ll n_y$), which does not hold in the application we consider in the following.

B.2 Super-resolution

In this subsection, $\mathcal{F}$ takes the following form :

$$\mathcal{F}(\mathbf{X}) = [\mathbf{M}_1\mathbf{x}_1^{(c)}, \ldots, \mathbf{M}_p\mathbf{x}_p^{(c)}], \tag{V.6}$$

where $\mathbf{M}_i$ is a warping and downsampling matrix. Since we consider a set of compact objects images, the only geometric transformation one has to deal with for registration is the images shifts with respect to the finest pixel grid, which can be estimated using the images centroids as shown in Chapter IV.

To the best of our knowledge, the only method dealing with this specific setting is the one used in the PSF modeling software PSFEx [Bertin 2011]. This method solves a problem of the form :

$$\min_{\mathbf{\Delta}_S} \frac{1}{2}\|\mathbf{Y} - \mathcal{F}((\mathbf{\Delta}_S + \mathbf{S}_0)\mathbf{A})\|_F^2 + \lambda\|\mathbf{\Delta}_S\|_F^2. \tag{V.7}$$

$\mathbf{S}_0$ is a rough first guess of the model components. Each line of the weight matrix $\mathbf{A}$ is assumed to follow a monomial law of some given field's parameters. The number of components is determined by the maximal degree of the monomials. For instance, let say that we want to model the PSFs variations as a function of their position in the field with monomials with degrees up to 3, then :

— one needs 6 components corresponding to the monomials $1, X, X^2, Y, XY$ and Y^2 ;

— assuming that the i^{th} PSF in $\mathbf{Y}$'s columns order is located at $\mathbf{u}_i = (u_{ix}, u_{iy})$ then the i^{th} column of $\mathbf{A}$ is given by $\mathbf{a}_i^{(c)} = [1, u_{ix}, u_{ix}^2, u_{iy}, u_{ix}u_{iy}, u_{iy}^2]^T$ up to a scaling factor.

This method is used for comparisons in the Numerical experiments part.

C Resolved Components Analysis

C.1 Matrix factorization

We have seen that we can describe the PSFs field f as

$$[f(\mathbf{u}_1), \ldots, f(\mathbf{u}_p)] = \mathbf{X} = \mathbf{S}\mathbf{A}. \tag{V.8}$$

The matrix $\mathbf{S}$ is independent of the spatial location, and the i^{th} line of $\mathbf{A}$ gives the contribution of the vector $\mathbf{s}_i$ to each of the samples. As discussed in

section A.3, the field's regularity can be taken into account by introducing a structuring of the matrix $\mathbf{A}$. We can write :

$$\mathbf{A}[i,:]^T = \sum_{l=1}^{N} \alpha_{il} \boldsymbol{v}_l, i = 1 \dots r, \tag{V.9}$$

where $(\boldsymbol{v}_l)_{1 \leq l \leq N}$ is a set of vectors spanning $\mathbb{R}^p$. Equivalently, we can write $\mathbf{A} = \boldsymbol{\alpha}\mathbf{V}^T$, where $\mathbf{V} = [\boldsymbol{v}_1, \dots, \boldsymbol{v}_N]$ and $\boldsymbol{\alpha}$ is a $r \times N$ matrix (see Fig. ***V.1***).

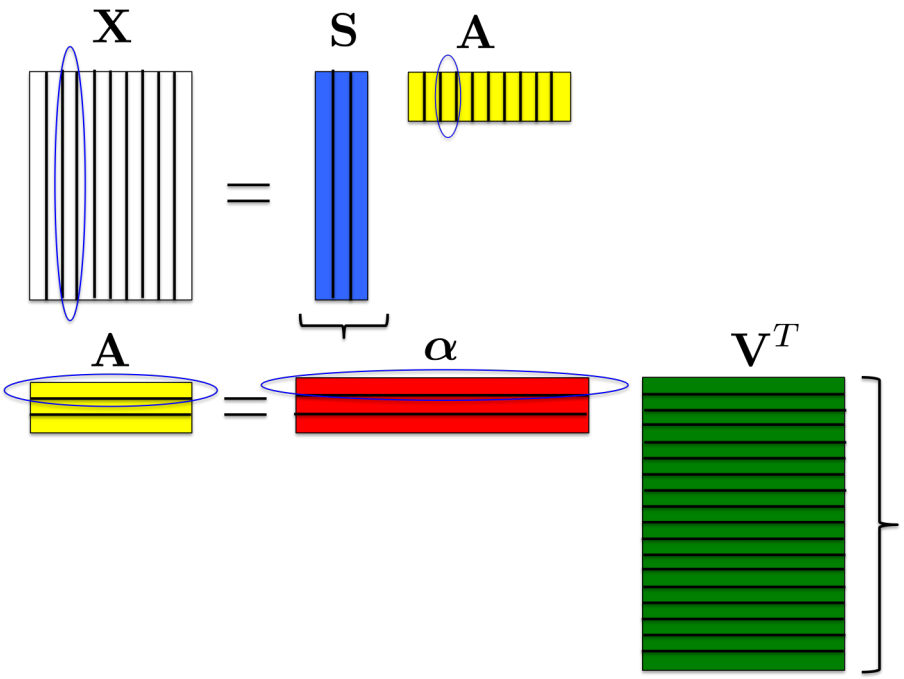


Figure *V.1* – Data matrix factorization : the j^{th} sample, which is stored in the j^{th} column of $\mathbf{X}$ is linear combination of $\mathbf{S}$ columns using $\mathbf{A}$'s j^{th} column coefficients as the weights ; similarly, the j^{th} line of $\mathbf{A}$ is a linear combination $\mathbf{V}^T$'s lines, using $\boldsymbol{\alpha}$'s j^{th} line coefficients as the weights.

Physical interpretation

An interesting way to well interpret $\mathbf{A}$ is to consider the ideal case where the measurements are distributed following a regular grid of locations $\mathcal{U}$. In this case, we can expand the vector $\mathbf{A}[i,:]^T$ using the Discrete Cosine Transform (DCT), and vectors $\boldsymbol{v}_i$ in Eq. V.9 are regular cosine atoms, and the column index of the matrix is related the frequency. Hence, lines relative to high frequencies will be related to quicky varying PSF components in the field,

while lines related to low frequencies will be related to PSFs stable components. In practice, the sampling is not regular and the DCT cannot be used, and $\mathbf{V}$ has to be learned in a way to keep the harmonic interpretation valid. We want some lines $\mathbf{A}[i,:]$ to describe stable PSFs components on the FOV, and other to be more related to local behavior.

C.2 The proximity constraint on A

As previously mentioned, we want to account for the PSFs field's regularity by constraining $\mathbf{A}$'s lines. Specifically, we want some lines to determine the distribution of stable features across the PSFs field while we want other lines to be related to more localized features. In order to build this constraint, let first consider the simple case of a one dimensional field of regularly spaced PSFs.

C.2.1 Regularly distributed observations

We first assume that $\mathcal{E} = \mathbb{R}$.

We suppose that $p = 2k+1$, for some integer k and we consider the 1D vector $\boldsymbol{\psi}_{e,a} = (\psi_i)_{1\leq i\leq p}$ defined as follows :

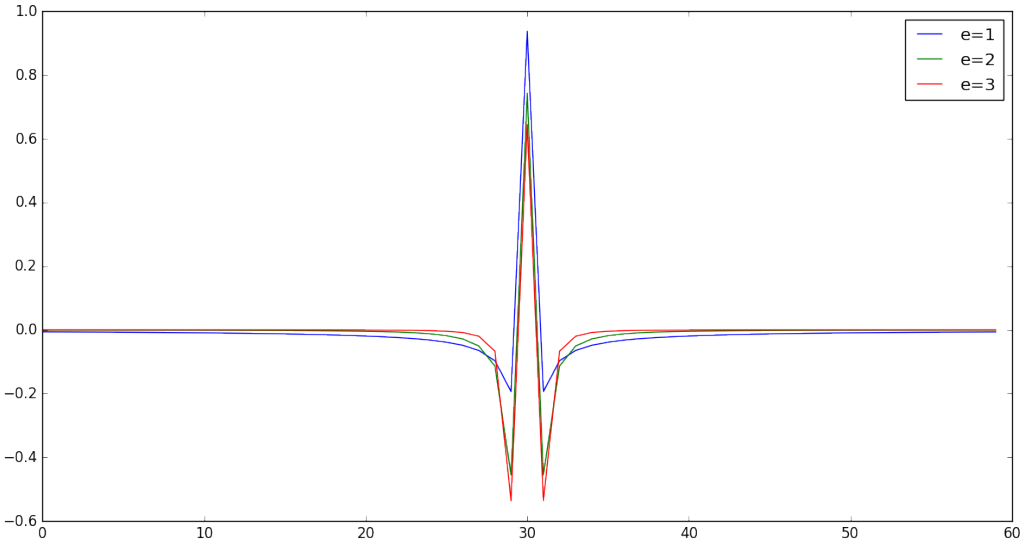
$$\psi_i = \psi_{p-i+1} = -1/|\mathbf{u}_i - \mathbf{u}_{k+1}|^e \qquad \text{if } i \neq k+1, \tag{V.10}$$

$$\psi_i = \sum_{\substack{j=1 \\ j\neq k+1}}^{p} a/|\mathbf{u}_j - \mathbf{u}_{k+1}|^e, \qquad \text{otherwise} \tag{V.11}$$

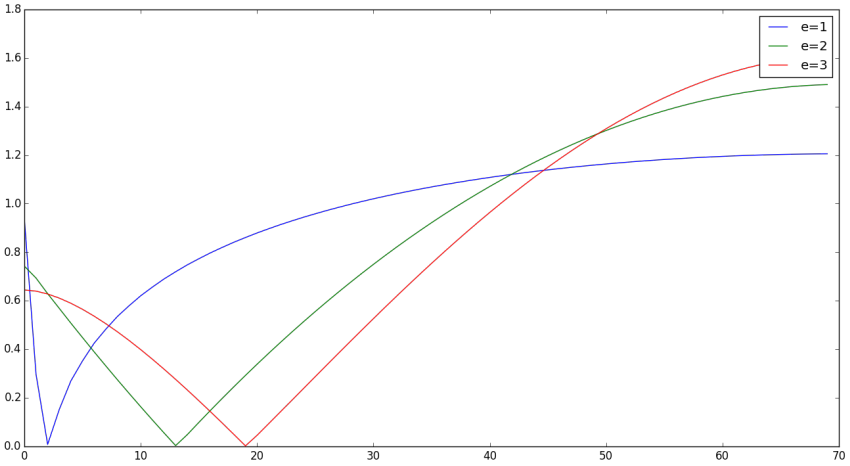
for some positive reals e and a. We suppose that $\boldsymbol{\psi}_{e,a}$ is normalized in l_2 norm. We refer to this family of signals, parametrized by e and a as "notch filters", in reason of their frequency responses shapes. Some examples can be found in Fig. *V.2*b. One can observe that $\psi_{1,1}$ is essentially a high pass filter. As e increases, the notch structure clearly appears, with an increasing notch frequency. It is clear that, for a vector $\mathbf{v}$, minimizing the functional $\Psi_{e,a}(\mathbf{v}) = \|\mathbf{v} \star \boldsymbol{\psi}_{e,a}\|_2^2$ promotes vectors with spectra concentrated around the notch frequency corresponding to the chosen values of e and a. We can directly use this family of filters to constraint $\mathbf{A}$ as follows : we define the functional

$$\boldsymbol{\Psi} : M_{rp}(\mathbb{R}) \mapsto \mathbb{R}^+ \, , \mathbf{A} \to \sum_{i=1}^{r} \Psi_{e_i,a}(\mathbf{A}[i,:]), \tag{V.12}$$

where $(e_i)_i$ is a set of reals verifying $0 \leq e_1 < e_2 < \cdots < e_r$ and $a \in [0,2[$. Because the notch frequency increases with e_i, minimizing $\boldsymbol{\Psi}$ promotes varying level of smoothness of $\mathbf{A}$'s lines, which is what we wanted to achieve. The filter $\boldsymbol{\psi}_{e,a}$ and the functional definitions can be extended to higher dimensions



(a) Direct domain samples.



(b) Discrete Fourier Transform (DFT) entry-wise moduli.

Figure *V.2* – Notch filters examples for different values of the parameter e in Eq. V.10 and V.11. The parameter a is set to 1.

of the space $\mathcal{E}$ by involving a multidimensional convolution [Rakhuba & Oseledets 2014]. Therefore, if the PSFs are distributed over a regular grid with respect to $\mathcal{E}$, one can implement the proximity constraint by solving

$$\min_{\mathbf{A},\mathbf{S}} \frac{1}{2}\|\mathbf{Y} - \mathcal{F}(\mathbf{S}\mathbf{A})\|_F^2 + \lambda\Psi(\mathbf{A}), \tag{V.13}$$

for some positive λ. Yet, in practical applications, the observations are in gene-

ral irregularly distributed. In the next section, we propose a slightly different penalty which is usable for arbitrary observations distributions.

C.2.2 General setting

Let define the functional

$$\widehat{\Psi}_{e,a} : \mathbb{R}^p \mapsto \mathbb{R}^+ \, , \mathbf{v} \to \sum_{k=1}^{p} (\sum_{\substack{i=1 \\ i \neq k}}^{p} \frac{a v_k - v_i}{\|\mathbf{u}_k - \mathbf{u}_i\|_2^e})^2, \tag{V.14}$$

where e and a are positive reals. Minimizing $\widehat{\Psi}_{e,a}(\mathbf{v})$ tends to enforce the similarity of close features, with respect to $\mathcal{E}$; in other terms, the more $\|\mathbf{u}_k - \mathbf{u}_i\|_2$ is large, the less important is $(av_k - v_i)$ in $\widehat{\Psi}_{e,a}(\mathbf{v})$ and e somehow determines the radius of similarity. For $e > 1$, $\widehat{\Psi}_{e,a} \approx \Psi_{e,a}$ because of the uniform spacing of the values $\mathbf{u}_i$ and the decay of $\frac{1}{\|\mathbf{u}_k - \mathbf{u}_i\|_2^e}$, for sufficiently high p; we give more details on this approximation in E in the 1D case. However unlike $\Psi_{e,a}$, the functional $\widehat{\Psi}_{e,a}$ is still relevant without the uniform sampling hypothesis and we expect qualitatively the same behavior as $\Psi_{e,a}$ with respect to the frequency domain if the data sampling is sufficiently dense. Therefore we use $\widehat{\Psi}_{e,a}$ instead of $\Psi_{e,a}$ in the functional $\mathbf{\Psi}$ of Eq.V.13. Besides, we use the term frequency even for randomly distributed samples.

C.2.3 Flexible penalization : the redundant frequencies dictionary V

The efficiency of the regularization of the problem V.13 relies on a good choice of the parameters $e_1, \ldots, e_r$ and a. Indeed, if the associated notch frequencies does not match with the data set frequency content, the regularization will more or less bias the PSF estimation depending on the Lagrange multiplier λ. Besides, setting this parameter might be tricky. We propose an alternate strategy for constraining $\mathbf{A}$, which leads to the factorization model introduced in Section C.1 and still builds over the idea of notch filters.

For $\mathbf{v} \in \mathbb{R}^p$ we can write

$$\widehat{\Psi}_{e,a}(\mathbf{v}) = \|\mathbf{P}_{e,a}\mathbf{v}\|_2^2, \tag{V.15}$$

where $\mathbf{P}_{e,a}$ is a $p \times p$ matrix defined by

$$\mathbf{P}_{e,a}[i,j] = -\frac{1}{\|\mathbf{u}_i - \mathbf{u}_j\|_2^e} \qquad \text{if } i \neq j, \tag{V.16}$$

$$\mathbf{P}_{e,a}[i,i] = \sum_{\substack{j=1 \\ j \neq i}}^{p} \frac{a}{\|\mathbf{u}_i - \mathbf{u}_j\|_2^e}, \tag{V.17}$$

$(i,j) \in [\![1,p]\!]^2$. Therefore,

$$\widehat{\Psi}_{e,a}(\mathbf{v}) = \mathbf{v}^T \mathbf{Q}_{e,a} \mathbf{v}, \tag{V.18}$$

where $\mathbf{Q}_{e,a} = \mathbf{P}_{e,a}^T \mathbf{P}_{e,a}$ and is symmetric and positive. We consider the singular values decomposition (SVD) of $\mathbf{Q}_{e,a}$: $\mathbf{Q}_{e,a} = \mathbf{V}_{e,a} \mathbf{D}_{e,a} \mathbf{V}_{e,a}^T$. The diagonal values of $\mathbf{D}_{e,a}$ are sorted in decreasing order. We note $\mathbf{d}_{e,a}$ the vector made of these diagonal values, so that $\mathbf{d}_{e,a}[1] \geq \cdots \geq \mathbf{d}_{e,a}[p] \geq 0$. Considering the reduced form $\widehat{\Psi}_{e,a}(\mathbf{v}) = \sum_{i=1}^{p} \mathbf{d}_{e,a}[i] \langle \mathbf{v}, \mathbf{V}_{e,a}[:,i] \rangle^2$, it is clear that minimizing $\widehat{\Psi}_{e,a}(\mathbf{v})$ promotes vectors correlated with $\mathbf{Q}_{e,a}$ last eigenvectors. In the case of regular sampling with respect to $\mathcal{E}$, these eigenvectors are the harmonics close to the notch frequency of $\boldsymbol{\psi}_{e,a}$. We can rewrite the functional $\boldsymbol{\Psi}$ accordingly :

$$\boldsymbol{\Psi}(\mathbf{A}) = \sum_{i=1}^{r} \sum_{j=1}^{p} \mathbf{d}_{e_i,a}[j] \langle \mathbf{v}, \mathbf{V}_{e_i,a}[:,j] \rangle^2. \tag{V.19}$$

It is clear from this expression that minimizing $\boldsymbol{\Psi}(\mathbf{A})$ enforces the selection of the eigenvectors associated with the lowest eigenvalues in the set $(\mathbf{d}_{e_i,a}[j])_{i,j}$ for describing $\mathbf{A}$'s lines. This can be seen as a sparsity constraint over $\mathbf{A}$'s lines with respect to the atoms $(\mathbf{V}_{e_i,a}[:,j])_{i,j}$; yet, the small subset of atoms which will carry most of the information is somehow predefined through the eigenvalues $(\mathbf{d}_{e_i,a}[j])_{i,j}$. This is unsuitable if the notch filters parameters are poorly selected ; on the contrary, one would like to select in a flexible way the atoms which fit the best the data.

Let suppose that we have determined a set of parameters $(e_i, a_i)_{1 \leq i \leq r}$ so that the filters $\boldsymbol{\psi}_{e_i,a_i}$ notch frequencies would cover the range of significant frequencies (with respect to $\mathcal{E}$) present in the data. As previously, we note $(\mathbf{V}_{e_i,a_i})_{1 \leq i \leq r}$ the eigenvector's matrices associated with the operators $\widehat{\Psi}_{e_i,a_i}$. We note $\mathbf{V} = [\mathbf{V}_{e_1,a_1}, \ldots, \mathbf{V}_{e_r,a_r}]$. Considering the preceding remark, we introduce the following problem :

$$\min_{\boldsymbol{\alpha}, \mathbf{S}} \frac{1}{2} \|\mathbf{Y} - \mathcal{F}(\mathbf{S}\boldsymbol{\alpha}\mathbf{V}^T)\|_F^2 \text{ s.t. } \|\boldsymbol{\alpha}[l,:]\|_0 \leq \eta_l,\ l = 1 \ldots r \tag{V.20}$$

Now $\mathbf{A} = \boldsymbol{\alpha}\mathbf{V}^T$. Each line of $\mathbf{A}$ is a sparse linear combination of $\mathbf{V}^T$'s lines, and the "active" atoms are optimally selected according to the data. The choice of the parameters $(e_i, a_i)_{1 \leq i \leq r}$ and $(\eta_l)_{1 \leq i \leq r}$ is discussed in a forthcoming section.

C.2.4 A connection with graphs theory

In case $a = 1$, the matrix $\mathbf{P}_{e,a}$ is the laplacian of an undirected fully connected and weighted graph with p nodes $1 \ldots p$, such that the weight of

the vertex connecting a node i to a node j is $\frac{1}{\|\mathbf{u}_i-\mathbf{u}_j\|_2^e}$ [Anderson & Morley 1985]. As proposed in spectral graph theory [Chung 1997], this gives a natural interpretation of $\mathbf{P}_{e,a}$ (and $\mathbf{Q}_{e,a}$) eigenvectors as harmonic atoms in the graph's geometry. Each line of the matrix $\mathbf{A}$ can be seen as a function defined on a family of graphs determined by the observations locations, so that we enforce the regularity of $\mathbf{A}$'s lines according to the graphs geometry. Our approach is thereby close to the spectral graphs wavelets framework [Hammond *et al.* 2009]. However, the graphs wavelets are built on a single graph and a scaling parameter allows one to derive wavelets atoms corresponding to spectral bands of different sizes. In our case, the scales diversity is accounted for by building a dictionary of harmonics corresponding to different graphs. Indeed, as e increases, the weight associated to the most distant nodes (in the sense of $\|\mathbf{u}_i - \mathbf{u}_j\|_2$) becomes negligible, which implies that the corresponding graph laplacian is determined by nearby nodes, yielding "higher" frequencies harmonics.

C.3 The smoothness constraint on S

As previously mentioned, each PSF is a structured image. We can account for this through a sparsity constraint which this has proven effective in multiple frame PSFs super-resolution in the previous chapter.

Since we do not estimate individual PSFs directly, we instead constraint the eigen PSFs which are $\mathbf{S}$'s columns. Specifically, we promote $\mathbf{S}$'s columns sparsity with respect to a chosen dictionary $\mathbf{\Phi}_{\boldsymbol{s}}$. By definition, a typical imaging system's PSF concentrates most of its power in few pixels. Therefore a straightforward choice for $\mathbf{\Phi}_{\boldsymbol{s}}$ is $\mathbf{I}_n$. In other words, we will enforce the sparsity of $\mathbf{S}$'s columns in the pixels domain.

On the other hand, we take $\mathbf{\Phi}_s$ as the second generation Starlet forward transform [Starck *et al.* 2011], without the coarse scale. The power of sparse prior in wavelet domain for inverse problems being well established, we shall online emphasize the fact that this particular choice of wavelet is particularly suitable for images with nearly isotropic features.

C.4 Algorithm

We define the sets $\Omega_1 = \{\boldsymbol{\alpha} \in M_{r,N}(\mathbb{R})/\|\boldsymbol{\alpha}[l,:]\|_0 \leq \eta_l,\ l = 1 \dots r\}$ and $\Omega_2 = \{(\mathbf{S}, \boldsymbol{\alpha}) \in M_{nr}(\mathbb{R}) \times M_{r,N}(\mathbb{R})/\mathbf{S}\boldsymbol{\alpha}\mathbf{V}^T \geq_{M_{np}(\mathbb{R})} 0\}$. The aforementioned constraints leads us to the following optimization problem :

$$\min_{\boldsymbol{\alpha},\mathbf{S}} \frac{1}{2}\|\mathbf{Y} - \mathcal{F}(\mathbf{S}\boldsymbol{\alpha}\mathbf{V}^T)\|_F^2 + \sum_{i=1}^{r} \|\mathbf{w}_i \odot \mathbf{\Phi}_s \mathbf{s}_i\|_1 \ + \iota_{\Omega_1}(\boldsymbol{\alpha}) + \iota_{\Omega_2}(\mathbf{S}, \boldsymbol{\alpha}). \quad \text{(V.21)}$$

where $\odot$ denotes the Hadamard product and $\iota_{\mathcal{C}}$ denotes the indicator function of a set $\mathcal{C}$ (see B). The ℓ_1 term promotes the sparsity of $\mathbf{S}$ columns with respect to $\mathbf{\Phi}_s$. The vectors $(\mathbf{w}_i)_i$ weight the sparsity against the other constraints and allow some adaptivity of the penalty, with respect to the uncertainties propagated to each entry of $\mathbf{S}$ (see Chapter IV).

The parametric aspects of this method are made clear in the subsequent sections.

The Problem V.21 is globally non-convex because of the coupling between $\mathbf{S}$ and $\boldsymbol{\alpha}$ and the l_0 constraint. In particular, the feasible set $\{(\mathbf{S}, \boldsymbol{\alpha}) \in M_{nr}(\mathbb{R}) \times M_{r,N}(\mathbb{R}) / \mathbf{S}\boldsymbol{\alpha}\mathbf{V}^T \geq 0\}$, with $N = rp$ is non-convex.

Therefore, one can at most expect to find a local minimum. To do so, we consider the following alternating minimization scheme :

1. Initialization : $\boldsymbol{\alpha}_0 \in \Omega_1$, with $N = rp$, $\mathbf{S}_0 = \underset{\mathbf{S}}{\operatorname{argmin}} \frac{1}{2}\|\mathbf{Y} - \mathcal{F}(\mathbf{S}\boldsymbol{\alpha}_0\mathbf{V}^T)\|_F^2 + \sum_{i=1}^r \|\mathbf{w}_i \odot \mathbf{\Phi}_s \mathbf{S}[:, i]\|_1$ s.t. $\mathbf{S}\boldsymbol{\alpha}_0\mathbf{V}^T \geq 0$
2. For $\mathrm{k} = 0 \ldots \mathrm{k}_{\max}$:
   (a) $\boldsymbol{\alpha}_{k+1} = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \frac{1}{2}\|\mathbf{Y} - \mathcal{F}(\mathbf{S}_k\boldsymbol{\alpha}\mathbf{V}^T)\|_F^2$ s.t. $\|\boldsymbol{\alpha}[l, :]\|_0 \leq \eta_l,\ l = 1 \ldots r$,
   (b) $\mathbf{S}_{k+1} = \underset{\mathbf{S}}{\operatorname{argmin}} \frac{1}{2}\|\mathbf{Y} - \mathcal{F}(\mathbf{S}\boldsymbol{\alpha}_{k+1}\mathbf{V}^T)\|_F^2 + \sum_{i=1}^r \|\mathbf{w}_i \odot \mathbf{\Phi}_s \mathbf{S}[:, i]\|_1$ s.t. $\mathbf{S}\boldsymbol{\alpha}_{k+1}\mathbf{V}^T \geq 0$.

The problem (a) remains non-convex ; yet there exists heuristic methods allowing one to approach a local minimum [Soussen *et al.* 2015, Blumensath & Davies 2008, Cartis & Thompson 2015]. The problem (b) is convex and can be solved efficiently.

One can note that there is no positivity constraint in the sub-problem (a). This choice is motivated by two facts :

— the feasible set of (b) is non-empty for any choice of $\boldsymbol{\alpha}_{k+1}$;

— allowing $\boldsymbol{\alpha}$ to be outside of the global problem feasible set (for $\mathbf{S}$ fixed) brings some robustness regarding local degenerated solutions.

There is an important body of work in the literature on alternate minimization schemes convergence, and in particular in the non-convex and non-smooth setting (see [Bolte *et al.* 2014] and the references therein). In the proposed scheme, the analysis is complicated by the asymmetry of the problems (a) and (b).

We define the function

$$\mathcal{H}(\boldsymbol{\alpha}, \mathbf{S}) = \frac{1}{2}\|\mathbf{Y} - \mathcal{F}(\mathbf{S}\boldsymbol{\alpha}\mathbf{V}^T)\|_F^2 + \sum_{i=1}^r \|\mathbf{w}_i \odot \mathbf{\Phi}_s \mathbf{s}_i\|_1 \tag{V.22}$$

and the matrix $\widehat{\mathbf{S}}_k = \underset{\mathbf{S}}{\operatorname{argmin}} \frac{1}{2}\|\mathbf{S} - \mathbf{S}_k\|_2^2$ s.t. $\mathbf{S}\boldsymbol{\alpha}_{k+1}\mathbf{V}^T \geq 0$. One immediate

sufficient condition for the sequence $(\mathcal{H}(\boldsymbol{\alpha}_k, \mathbf{S}_k))_k$ to be decreasing (and thereby convergent) is

$$\mathcal{H}(\boldsymbol{\alpha}_{k+1}, \widehat{\mathbf{S}}_k) \leq \mathcal{H}(\boldsymbol{\alpha}_k, \mathbf{S}_k) \tag{V.23}$$

which occurs if $(\mathbf{S}_k, \boldsymbol{\alpha}_{k+1})$ stays sufficiently close to Ω_2. Although we do not prove this always holds true, we observe on examples that the matrix $\mathbf{S}_k\boldsymbol{\alpha}_{k+1}\mathbf{V}^T$ in general only has a few and small negative entries for $k \geq 1$. This follows from the adequacy of the dictionary $\mathbf{V}$ for sparsely describing $\mathbf{A}$'s lines.

The complete method is given in Algorithm V.1. The resolution of the minimization sub-problems is detailed in Appendix D.

Algorithme V.1 Resolved components analysis (RCA)

1: Parameters estimation and initialization :
Harmonic constraint parameters $(e_i, a_i)_{1\leq i\leq r} \rightarrow \mathbf{V}, \mathbf{A}_0$
Noise level, $\mathbf{A}_0 \rightarrow \mathbf{W}_{0,0}$
2: Alternate minimization
3: **Pour** $k = 0$ to $k_{\max}$ **faire**
4: **Pour** $j = 0$ to $j_{\max}$ **faire**
5: $\mathbf{S}_k = \underset{\mathbf{S}}{\operatorname{argmin}} \frac{1}{2}\|\mathbf{Y} - \mathcal{F}(\mathbf{S}\mathbf{A}_k)\|_F^2 + \sum_{i=1}^r \|\mathbf{W}_{k,j}[:,i] \odot \boldsymbol{\Phi}_s\mathbf{S}[:,i]\|_1$ s.t. $\mathbf{S}\mathbf{A}_k \geq 0$
6: update : $\mathbf{W}_{k,0}, \mathbf{S}_k \rightarrow \text{update}(\mathbf{W}_{k,j+1})$
7: $\boldsymbol{\alpha}_{k+1} = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \frac{1}{2}\|\mathbf{Y} - \mathcal{F}(\mathbf{S}_k\boldsymbol{\alpha}\mathbf{V}^T)\|_F^2$ s.t. $\|\boldsymbol{\alpha}[l,:]\|_0 \leq \eta_l$
8: update : Noise level, $\boldsymbol{\alpha}_{k+1} \rightarrow \mathbf{W}_{k+1,0}$
9: $\mathbf{A}_{k+1} = \boldsymbol{\alpha}_{k+1}\mathbf{V}^T$
10: $\mathbf{A}_{k+1}[i,:] = \mathbf{A}_{k+1}[i,:]/\|\mathbf{A}_{k+1}[i,:]\|_2$, for $i = 1 \ldots r$
11: **Return :** $\mathbf{S}_{k_{\max}}$, $\mathbf{A}_{k_{\max}}$.

C.5 Parameters setting

C.5.1 Components sparsity parameters

We consider the terms of the form $\|\mathbf{w}_{k,j} \odot \boldsymbol{\Phi}_s\mathbf{s}\|_1$, where k is the alternate minimization index and j is the re-weighted ℓ_1 minimization index. We first suppose that $\boldsymbol{\Phi}_s = \mathbf{I}_n$. We decompose $\mathbf{w}_{k,j}$ as :

$$\mathbf{w}_{k,j} = \kappa\boldsymbol{\beta}_{k,j} \odot \boldsymbol{\lambda}_k \tag{V.24}$$

Let consider the minimization problems in $\mathbf{S}$ in Algorithm V.1. Assuming that we simply minimize the quadratic term using the following steepest descent

update rule,

$$\mathbf{S}_{m+1} = \mathbf{S}_m + \mu\mathcal{F}^*(\mathbf{Y} - \mathcal{F}(\mathbf{S}_m\mathbf{A}_k))\mathbf{A}_k^T, \tag{V.25}$$

for a well chosen step size μ, $\mathcal{F}^*$ being the adjoint operator one can estimate the entry-wise standard deviations of the noise which propagates from the observations to the current solution $\mathbf{S}_{m+1}$. For a given matrix $\mathbf{X}$ in $M_{np}(\mathbb{R})$, we assume that $\mathcal{F}$ takes the following general form $\mathcal{F}(\mathbf{X}) = [\mathbf{M}_1\mathbf{X}[:,1],\ldots,\mathbf{M}_p\mathbf{X}[:,p]]$. We define $\mathcal{F}^2$ as :

$$\mathcal{F}^2(\mathbf{X}) = [(\mathbf{M}_1 \odot \mathbf{M}_1)\mathbf{X}[:,1],\ldots,(\mathbf{M}_p \odot \mathbf{M}_p)\mathbf{X}[:,p]] \tag{V.26}$$

We note $\mathbf{B}$ the observational noise (or model uncertainty) that we assume to gaussian, white and centered. The propagated noise entry-wise standard deviations are given by

$$\mathbf{\Sigma}_k = \mu\sqrt{\mathcal{F}^{2*}(\mathrm{Var}(\mathbf{B}))(\mathbf{A}_k^T \odot \mathbf{A}_k^T)}, \tag{V.27}$$

where Var() returns entry-wise variances and $\mathcal{F}^{2*}$ is the adjoint operator of $\mathcal{F}^2$. Now one can proceed to a hypothesis testing on the signal presence in each entry of $\mathbf{S}_{m+1}$ based on $\mathbf{\Sigma}_k$ [Starck *et al.* 2015], and denoise $\mathbf{S}_{m+1}$ accordingly. For instance, we define the noise-free version of $\mathbf{S}_{m+1}$ as follows :

$$\hat{\mathbf{S}}_{m+1}[i_1,i_2] = \begin{cases} 0, \text{ if } |\mathbf{S}_{m+1}[i_1,i_2]| \leq \kappa\mathbf{\Sigma}_k[i_1,i_2] \\ \frac{\mathbf{S}_{m+1}[i_1,i_2]}{|\mathbf{S}_{m+1}[i_1,i_2]|}(|\mathbf{S}_{m+1}[i_1,i_2]| - \kappa\mathbf{\Sigma}_k[i_1,i_2]). \text{ otherwise;} \end{cases} \tag{V.28}$$

where κ controls the false detection probability ; the noise being gaussian, we typically choose 3 or 4 for κ.

The sequence $(\hat{\mathbf{S}}_m)$ converges to a solution of the problem

$$\underset{\mathbf{S}}{\mathrm{argmin}}\ \frac{1}{2}\|\mathbf{Y} - \mathcal{F}(\mathbf{S}\boldsymbol{\alpha}_k\mathbf{U}^T)\|_F^2 + \sum_{i=1}^{r}\kappa\|\boldsymbol{\lambda}_k[:,i] \odot \mathbf{S}[:,i]\|_1, \tag{V.29}$$

for $\boldsymbol{\lambda}_k = \kappa/\mu\mathbf{\Sigma}_k$. One can find some material on minimization schemes in Appendix D. This choice yields a noise-free but biased solution because of the thresholding ; this is a well-known drawback of ℓ_1 norm based regularizations. The purpose of the vector $\boldsymbol{\beta}_{k,j}$ is to mitigate this bias [Candès *et al.* 2008]. $\boldsymbol{\beta}_{k,0}$ is a vector with ones at all entries. At the step 6 in Algo V.1, $\boldsymbol{\beta}_{k,j}$ is calculated as follows :

$$\boldsymbol{\beta}_{k,j} = \frac{1}{1 + \frac{|\mathbf{S}_k|}{\kappa\boldsymbol{\lambda}_k}}, \tag{V.30}$$

where all the operations are entry-wise and $|\mathbf{S}_k|$ is the vector made of element-wise absolute values of $\mathbf{S}_k$ entries. Qualitatively, this removes the strongest features from the ℓ_1 norm terms by giving them small weights, which makes

the debiasing possible ; conversely, the entries dominated by noise get weights close to 1, so that the penalty remains unchanged.

For $\mathbf{\Phi}_s \neq \mathbf{I}_n$ we follow the same rational. To set the sparsity in the transform domain according to the noise induced uncertainty, we need to further propagate it (the noise) through the operator $\mathbf{\Phi}_s$. Formally, we need to estimate the element-wise standard deviations of $\mu\mathbf{\Phi}_s\mathcal{F}^*(\mathbf{B})\mathbf{A}_k^T$. Let consider the intermediate random matrix $\mathbf{Y}_F = \mathcal{F}^*(\mathbf{B})$. Assuming that

$$\mathcal{F}(\mathcal{F}^*(.)) = \lambda \mathrm{Id}(.), \tag{V.31}$$

$\mathbf{Y}_F$'s lines are statistically independent. Therefore, within a given column of $\mathbf{Y}_F\mathbf{A}_k^T$, the entries are statistically independent from one another. We deduce that the element-wise standard deviations of $\mu\mathbf{\Phi}_s\mathcal{F}^*(\mathbf{B})\mathbf{A}_k^T$ are given by

$$\mathbf{\Sigma}_k = \mu\sqrt{(\mathbf{\Phi}_s \odot \mathbf{\Phi}_s)\mathcal{F}^{2*}(\mathrm{Var}(\mathbf{B}))(\mathbf{A}_k^T \odot \mathbf{A}_k^T)}. \tag{V.32}$$

Then $\boldsymbol{\lambda}_k$ is obtained as previously and $\boldsymbol{\beta}_{k,j}$ is calculated as

$$\boldsymbol{\beta}_{k,j} = \frac{1}{1 + \frac{|\mathbf{\Phi}_s\mathbf{S}_k|}{\kappa\boldsymbol{\lambda}_k}}. \tag{V.33}$$

The property V.31 is approximately true in the case of super-resolution.

C.5.2 Number of components

We do not propose a method to choose the number of components r. Yet, we observe that because of the sparsity constraint, some lines of the matrix α_{k+1} at the step 8 in Algorithm V.1 are equal to the null vector, when the number of components is overestimated. The corresponding lines in $\mathbf{A}_{k+1}$ and subsequently the corresponding columns in $\mathbf{S}_k$ are simply discarded. This provides an intrinsic mean to select the number of components. Thus in practice, one can choose the initial r as the data set dimensionality from the embedding space point of view, which can be estimated based on a principal component analysis.

C.5.3 Proximity constraint parameters

In this section, we consider the functionals $\widehat{\Psi}_{e_i,a_i}$ and especially the choice of the parameters e_i and a_i. Let assume that we have determine a suitable range for the parameters : $(e_i, a_i) \in \mathcal{S} = [e_{\min}, e_{\max}] \times [a_{\min}, a_{\max}]$ for $i = 1 \ldots r$.

For a particular (e, a) we consider the matrix $\mathbf{Q}_{e,a}$ and its eigenvectors matrix $\mathbf{V}_{e,a}$ introduced in section C.2.3. As previously stated, we want the weights matrix $\mathbf{A}$ lines to be sparse with respect to $\mathbf{Q}_{e,a}$'s eigenvectors. In

order to choose the parameters and initialize the weights matrix, we use the following greedy procedure. We consider a sequence of matrices $(\mathbf{R}_i)_{1\leq i\leq r}$, with $\mathbf{R}_1 = \mathbf{Y}$. For $i \in [\![1, r]\!]$ we define

$$\mathcal{J}_{e,a}(\mathbf{R}_i) = \max_{k\in[\![1,p]\!]} \|\mathbf{R}_i \mathbf{V}_{e,a}[:, k]\|_2, \tag{V.34}$$

and we note $\mathbf{v}^*_{e,a}$ the optimal eigenvector. We choose the i^{th} couple of parameters as :

$$(e_i, ai) = \underset{(e,a)\in\mathcal{S}}{\operatorname{argmax}}\ \mathcal{J}_{e,a}(\mathbf{R}_i). \tag{V.35}$$

$\mathbf{A}_0[i, :] = \mathbf{v}^*_{e_i,a_i}$ and $\mathbf{R}_{i+1} = \mathbf{R}_i - \mathbf{R}_i \mathbf{V}_{e,a} \mathbf{V}^T_{e,a}$.

Regarding the set $\mathcal{S}$, we choose the interval $a_{\min} = 0$ and $a_{\max} = 2$. This range allows the notch structure, assuming that $e_{\min} \geq 0$; for $a < 0$, $\mathbf{h}_{e,a}$ behaves as a low pass filter. For $a \geq 0$, we observe that $\mathbf{h}_{e,a}$ becomes a notch filter, with a notch frequency close to the null frequency for $a \geq 2$. As previously stated, e determines the influence of two samples on one another corresponding coefficients in the matrix $\mathbf{A}$ in the algorithmic process. According to Section C.2.2, we set $e_{\min} = 1$. Let consider the graph $\mathcal{G}_e$ introduced in section C.2.4. The higher is e, the lower is $\mathcal{G}_e$ connexity. Considering that we are looking for global features (yet localized in the field frequency domain), the highest possible value of e should guarantee that the graph $\mathcal{G}_e$ is connected. This gives us a practical upper bound for e. Once $\mathcal{S}$ is determined, we discretize this set, with a logarithmic step, in such a way to have more samples close to $(e_{\min}, a_{\min})$ which correspond to low notch frequencies. We solve approximately Problem V.35 by taking the best couple of parameters in the discretized version of $\mathcal{S}$. This step is the most computationally demanding, especially for large data samples.

C.5.4 Weights matrix sparsity parameters

The parameters η_l are implicitly set by the minimization scheme used at step 8 in V.1. This is detailed in Appendix D.

D Numerical experiments

In this section, we present the data used to test the proposed method, the simulation realized and comparisons to other existing methods for both dimensionality reduction and super-resolution aspects.

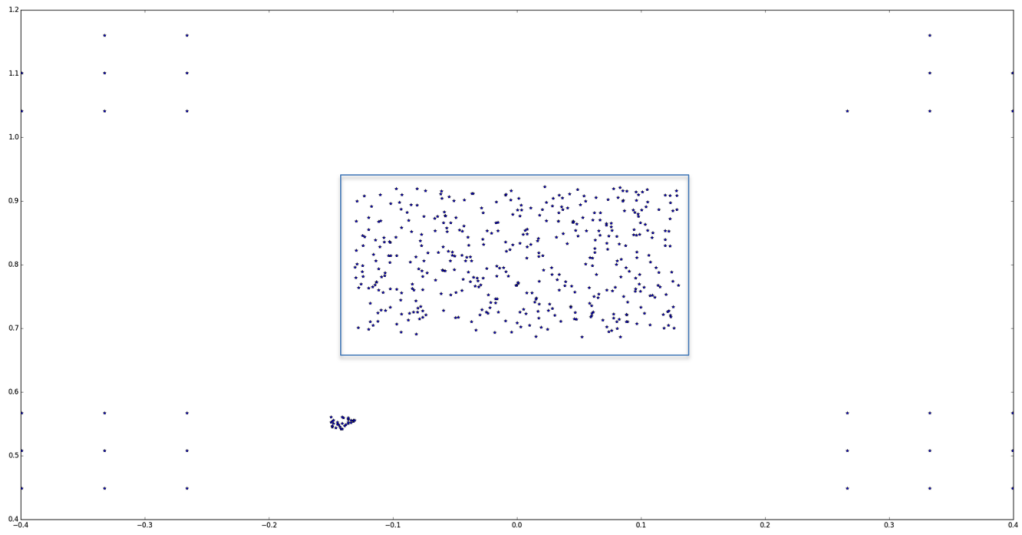


Figure *V.3* – Simulated PSFs distribution across the FOV.

D.1 Data

The data set consists of simulated optical Euclid PSFs as in Chapter IV, for a wavelength of $600\mu m$. The PSFs distribution across the field is shown on Fig. ***V.3***. These PSFs account for mirrors polishing imperfections, manufacturing and alignment errors and thermal stability of the telescope.

D.2 Simulation

We applied different dimension reduction algorithms to a set of 500 PSFs located in the blue box on Fig. ***V.3***. We applied the algorithms to different observations of the fields, with varying level of white gaussian noise. For a discrete signal $\mathbf{s}$ of length N corrupted with a white gaussian noise $\mathbf{b}$, we define the signal to noise ratio (SNR) as :

$$\text{SNR} = \frac{\|\mathbf{s}\|_2^2}{N\sigma_{\mathbf{b}}^2}. \tag{V.36}$$

D.3 Quality assessment

In astronomical surveys, the estimated PSF's shape is particularly important ; precisely, one has to be able to capture the PSF anisotropy. We recall

that for an image $\mathbf{X} = (x_{ij})_{i,j}$, the central moments are defined as

$$\mu_{p,q}(\mathbf{X}) = \sum_i \sum_j (i - i_c)^p (j - j_c)^q x_{ij} \tag{V.37}$$

with $(p, q) \in \mathbb{N}^2$, (i_c, j_c) are the image centroid coordinates. The moments $\mu_{2,0}$ and $\mu_{0,2}$ quantifies the light intensity spreading relatively to the lines $\{(i_c, y), y \in \mathbb{R}\}$ and $\{(x, j_c), x \in \mathbb{R}\}$ respectively. Now we consider the moment $\mu_{1,1}$. We introduce the centered and rotated pixels coordinates $(x_{i,\theta}, y_{j,\theta})$ defined by the system of equations

$$x_{i,\theta} \cos(\theta) + y_{j,\theta} \sin(\theta) = i - i_c \tag{V.38}$$

$$-x_{i,\theta} \sin(\theta) + y_{j,\theta} \cos(\theta) = j - j_c, \tag{V.39}$$

for some $\theta \in [0, 2\pi]$. Then we have

$$\mu_{1,1} = \sum_i \sum_j [\frac{\sin(2\theta)}{2}(-x_{i,\theta}^2 + y_{j,\theta}^2) + (2\cos^2(\theta) - 1) x_{i,\theta} y_{j,\theta}] x_{ij}, \tag{V.40}$$

and in particular, $\mu_{1,1} = \sum_i \sum_j [\frac{1}{2}(-x_{i,\frac{\pi}{4}}^2 + y_{j,\frac{\pi}{4}}^2)] x_{ij}$. It becomes clear that $\mu_{1,1}$ quantifies the light intensity spreading with respect to the pixels grid diagonals.

The ellipticity parameters are defined as,

$$e_1(\mathbf{X}) = \frac{\mu_{2,0}(\mathbf{X}) - \mu_{0,2}(\mathbf{X})}{\mu_{2,0}(\mathbf{X}) + \mu_{0,2}(\mathbf{X})} \tag{V.41}$$

$$e_2(\mathbf{X}) = \frac{2\mu_{1,1}(\mathbf{X})}{\mu_{2,0}(\mathbf{X}) + \mu_{0,2}(\mathbf{X})}. \tag{V.42}$$

We define the vector $\boldsymbol{\gamma}(\mathbf{X}) = [e_1(\mathbf{X}), e_2(\mathbf{X})]^T$. This vector characterizes how much $\mathbf{X}$ departs from an isotropic shape and indicates its main direction of orientation. It plays a central theoretical and practical role in weak lensing based dark matter characterization [Dodelson 2003].

Another important geometric feature is the so-called PSF size. It has been shown that the size error is a major contributor to the systematics in weak gravitational lensing surveys [Paulin-Henriksson *et al.* 2008]. We characterize the size of a PSF $\mathbf{X}$ as follows :

$$\mathrm{S}(\mathbf{X}) = (\frac{\sum_i \sum_j ((i - i_c)^2 + (j - j_c)^2) x_{ij}}{\sum_i \sum_j x_{ij}})^{1/2}. \tag{V.43}$$

Assuming that a given PSF is a 2D discrete probability distribution, this quantity measures how much this distribution is spread around its mean $[i_c, j_c]^T$. Let note $(\mathbf{X}_i)_{1 \leq i \leq p}$ the set of "original" PSFs and $(\hat{\mathbf{X}}_i)_{1 \leq i \leq p}$ the set of corresponding estimated PSFs with one of the compared methods, at a given SNR. The reconstruction quality is accessed through the following quantities :

— the average error on the ellipticity vector : $\mathrm{E}_{\boldsymbol{\gamma}} = \sum_{i=1}^{p} \|\boldsymbol{\gamma}(\mathbf{X}_i) - \boldsymbol{\gamma}(\hat{\mathbf{X}}_i)\|_2/p$;

— noting $\boldsymbol{\Gamma} = [\boldsymbol{\gamma}(\mathbf{X}_1) - \boldsymbol{\gamma}(\hat{\mathbf{X}}_1), \ldots, \boldsymbol{\gamma}(\mathbf{X}_p) - \boldsymbol{\gamma}(\hat{\mathbf{X}}_p)]$, the dispersion of the errors on the ellipticity vector is measured through the nuclear norm $\mathrm{B}_{\boldsymbol{\gamma}} = \|\boldsymbol{\Gamma}\|_*$;

— the average absolute error on the size : $\mathrm{E_S} = \sum_{i=1}^{p} |\mathrm{S}(\mathbf{X}_i) - \mathrm{S}(\hat{\mathbf{X}}_i)|/p$ in pixels ;

— the dispersion of the errors on the size : $\sigma_{\mathrm{S}} = \mathrm{std}((\mathrm{S}(\mathbf{X}_i) - \mathrm{S}(\hat{\mathbf{X}}_i))_i)$, in pixels.

E Numerical experiments

E.1 Results

E.1.1 Dimension reduction

In this section, we compare RCA to PCA, GMRA and the software PSFEx. We ran a PCA with different number of principal components between 0 and 15. 10 was the value which provided the best results. GMRA input was the data set intrinsic dimension [Little 2011], two, since the PSFs only vary as a function of their position in the field ; we provided the absolute squared quadratic error allowed with respect to the observed data based on the observation noise level. For PSFEx, we used 15 components. Finally, RCA used up to 15 components, and effectively, 2 and 4 components respectively for the lowest SNR fields realization. As previously mentioned, we assess the components sparsity's constraint :

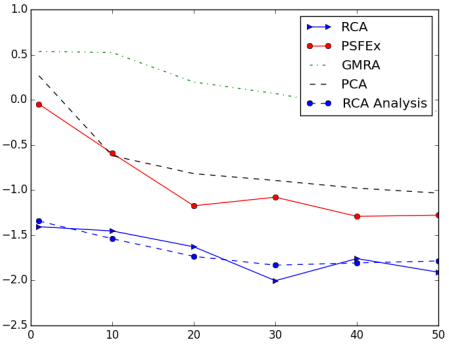
— on the one hand we consider $\boldsymbol{\Phi}_s = \mathbf{I}_n$ which enforces the components sparsity in pixels domain ; this is referred to as "RCA" in the plots ;

— on the other hand, we take $\boldsymbol{\Phi}_s$ as the second generation Starlet forward transform [Starck *et al.* 2011], without the coarse scale ; this is referred to as "RCA analysis" in the plots.

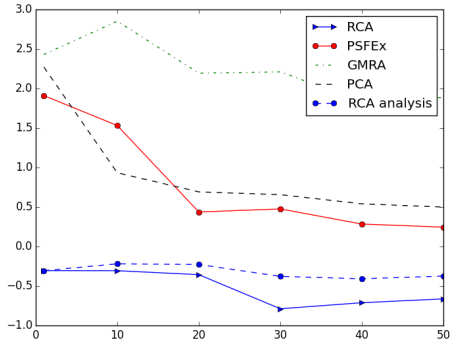
One can see on the left plot in Fig. ***V.4*** that the proposed method is at least 10 times more accurate on the ellipticity vector than the other considered methods. Moreover the right plot shows that the accuracy is way more stable. This is true for both choice of the dictionary $\boldsymbol{\Phi}_s$.

Fig. ***V.9*** shows that the estimated size $\mathrm{S}(\hat{\mathbf{X}}_i)$ is very sensitive to the choice of the dictionary $\boldsymbol{\Phi}_s$. The results are by far more accurate with a sparsity constraint on the components in wavelet domain than in direct domain.

For a given estimate of the PSF at a given location, the error on the size parameter is more sensitive to errors on the core of the PSF (main lobe and

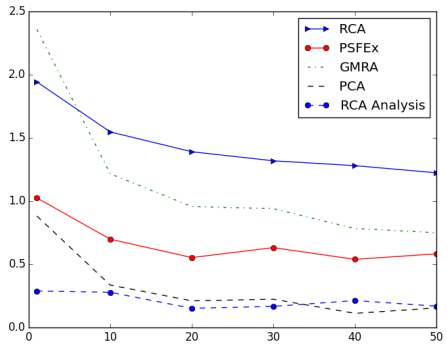


(a) Average error on the ellipticity vector.

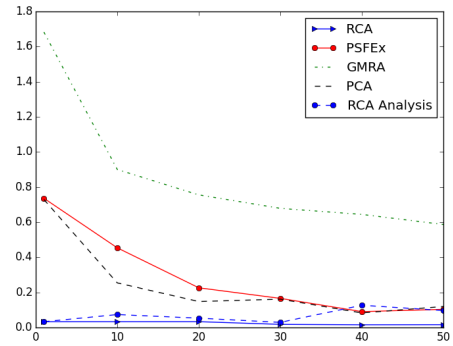


(b) Dispersion of the ellipticity vector.

Figure *V.4* – x axis : SNR (see section D.2) ; y axis : $\log_{10}(E_\gamma)$ for the left plot, $\log_{10}(B_\gamma)$ for the right plot.



(a) Average absolute error on the size.



(b) Dispersion of the errors on the size.

Figure *V.5* – x axis : SNR ; y axis : E_S for the left plot, σ_S for the right plot.

first rings) and less sensitive to errors on the outer part of the PSF than one would expect regarding the error on the ellipticity vector. The error on the outer part of the PSF is essentially related to the observational noise, whereas the error on core of the PSF - which has a high SNR - is more related to the method induced bias. This explains why the PCA performs quite well for this parameter. On the other hand, the bias induced by the sparsity is not only related to the dictionary choice, but also to the underlying data model with respect to the chosen dictionary.

As previously explained, the components sparsity term is set in such a way to penalize any feature which does not emerge from the propagated noise,

which is a source of bias. By using wavelets, we might recover features which are dominated by noise in pixel domain as long as the wavelet filters profile at given scale and direction, matches those features spatial structure. Thus, we expect less error on the reconstructed PSF's core by using wavelets.

We might also consider two distinct ways of using sparsity for the components :

— we can model each component as $\mathbf{s} = \mathbf{\Phi}_s^T \boldsymbol{\alpha}$, with $\boldsymbol{\alpha}$ sparse, which is known in the sparse recovery literature as synthesis prior ;

— we can alternately constraint $\mathbf{\Phi}_s \mathbf{s}$ to be sparse.

This priors are equivalent if the dictionary is unitary [Elad *et al.* 2007]. Therefore the pixel domain sparsity constraint can be considered as falling into both framework. However, the two priors are no longer equivalent and potentially yields quite different solutions for overcomplete dictionaries.

We observe in practice that unless the simulated PSFs are strictly sparse with respect to the chosen dictionary - this includes redundant wavelet dictionaries, the synthesis prior yields a bias on the reconstructed PSF size, since the estimated PSFs are sparse linear combinations of atoms which are in general sharper than a typical PSF profile. The analysis prior is somehow weaker and appears to be more suitable for approximately sparse data.

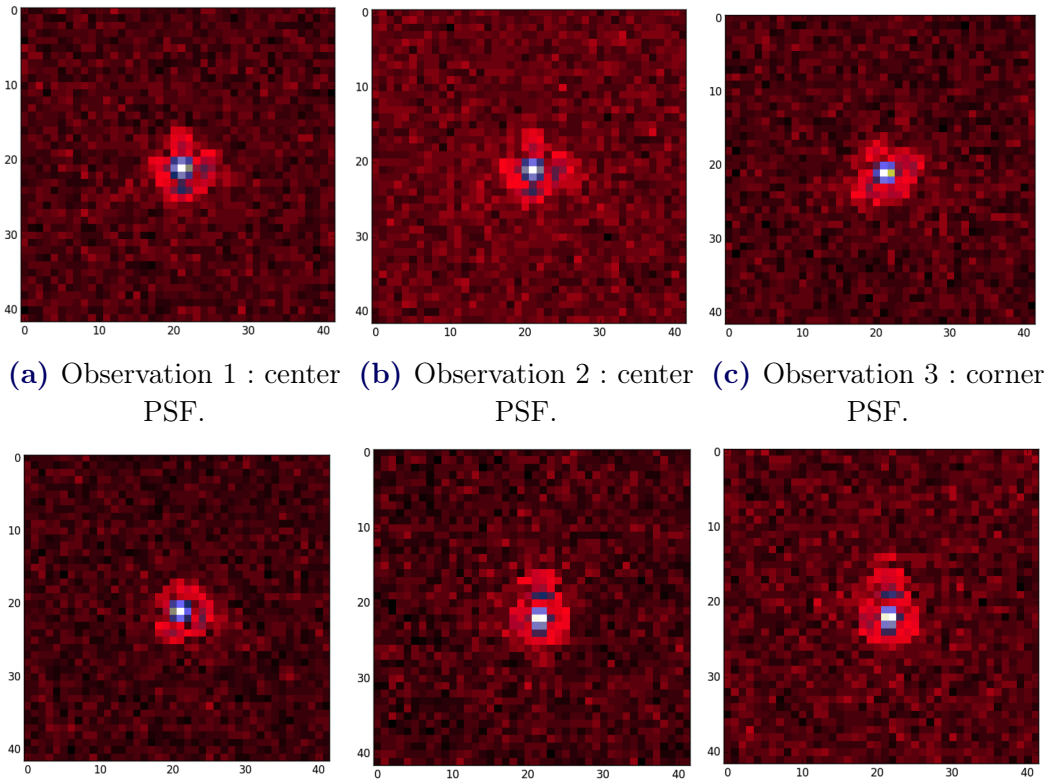
We do not observe a significant difference between these methods with respect to the mean squared error, except for GMRA which gave noisier reconstructions.

We applied the aforementioned methods to the PSFs field previously used, with additional 30 corners PSFs and 30 localized PSFs as shown on Fig. ***V.3*** at an SNR of 40. This assess the behavior of the algorithms with respect to spatial clustering and sparse data distribution. One can see in Fig. ***V.6*** examples of simulated observed PSFs from different areas in the FOV.

For each of these observed PSFs, the reconstructed PSFs for each method are shown in Fig. ***V.7***.

One can observe that the proposed method gives noiseless and rather accurate PSFs reconstruction for both the center, the corners and the localized area of the field (see Fig. ***V.3***). One can also see that we fail to capture accurately the rings pattern in the corners and the localized area. The dictionary $\mathbf{\Phi}_s$ considered are not specifically adapted to curve-like structures. The ring patterns varies across the FOV but are locally correlated. Therefore, they can only be recovered where the PSFs are sufficiently dense and numerous, which is the case at the FOV's center.

PCA and PSFEx yield a significant increase of the SNR in their estimated PSFs at the center and in the localized area. Yet, they fail to do so in the



(a) Observation 1 : center PSF. **(b)** Observation 2 : center PSF. **(c)** Observation 3 : corner PSF.

(d) Observation 4 : corner PSF. **(e)** Observation 5 : "local" PSF. **(f)** Observation 6 : "local" PSF.

Figure *V.6* – Input PSFs at different locations in the FOV for a SNR = 40. The corresponding reconstructed PSFs can be seen in Fig. ***V.7***

corners because of the lack of correlation for the PCA and local smoothness for PSFEx.

Finally, the poor results obtained with GMRA can be explained by the fact that the underlying manifold sampling is not sufficiently dense for the tangent spaces to be estimated reliably.

E.1.2 Super-resolution

In this section, the data are additionally downsampled to Euclid telescope resolution. PCA and GMRA does not handle the downsampling. Therefore we only consider PSFEx and RCA in this section. For each method, we estimate an upsampled version of each PSF, with a factor 2 in lines and columns ; in case of Euclid, this is enough to have a Nyquist frequency greater than half the signal spatial bandwidth [Cropper 2013].

As previously, RCA Analysis refers to the proposed method, with the dic-

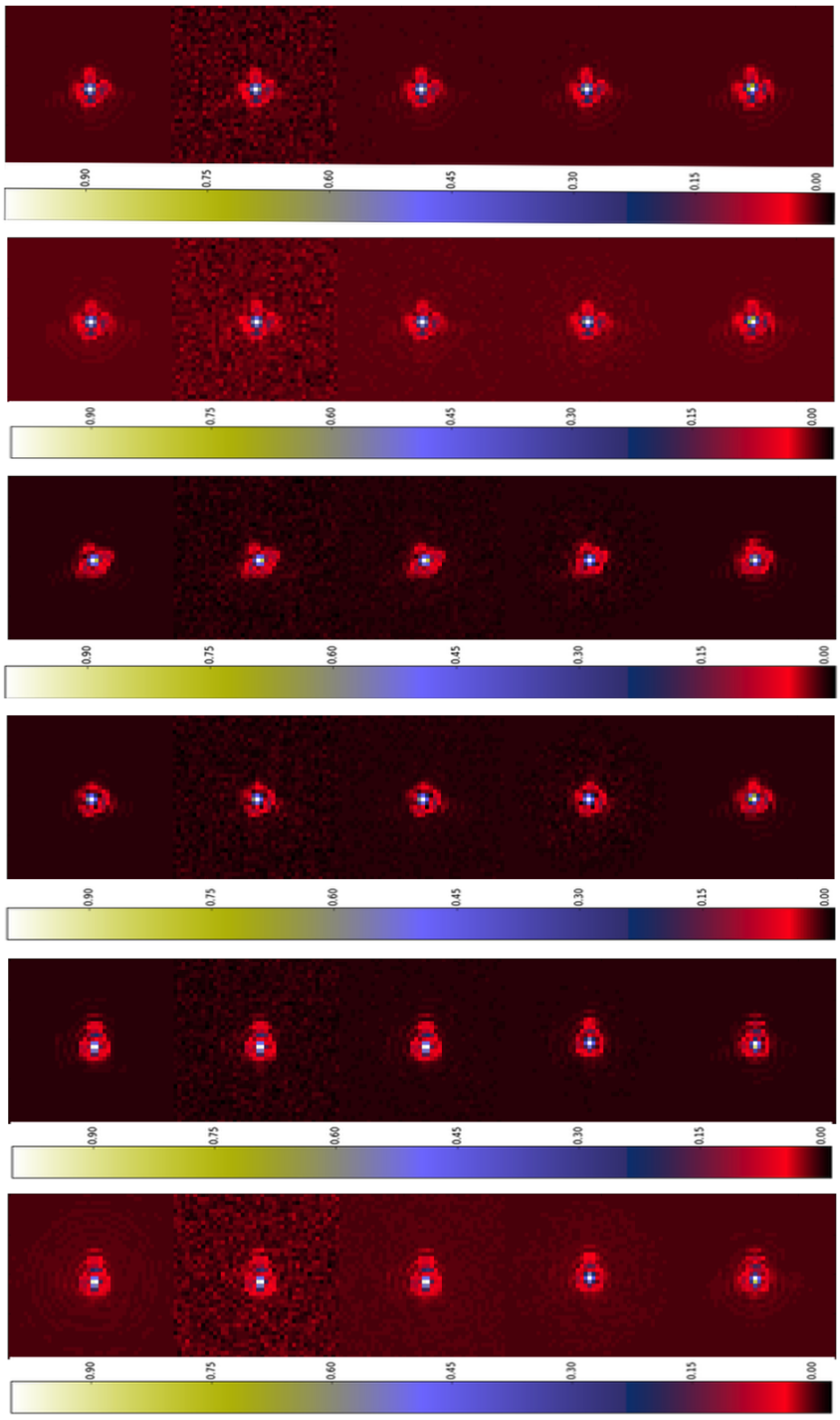
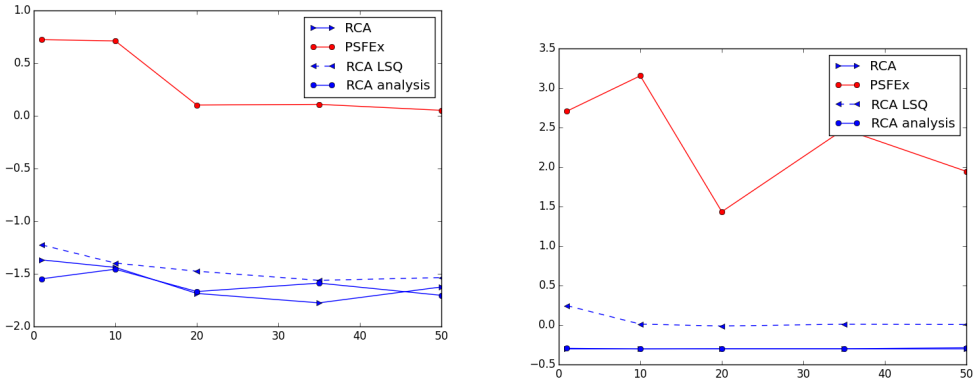


Figure V.7 – PSFs reconstructions : from the left to the right : original, GMRA, PCA, PSFEx, RCA ; from the bottom to the top : 2 "local" PSFs reconstructions, 2 corner PSFs reconstructions, 2 center PSFs reconstructions. The observed corresponding PSFs can be seen in Fig. ***V.6***

tionary $\mathbf{\Phi}_s$ chosen as the second generation Starlet forward transform [Starck *et al.* 2011], without the coarse scale ; RCA LSQ refers to the proposed method with the dictionary $\mathbf{\Phi}_s$ chosen as the identity matrix, and the weight matrix $\mathbf{A}$ simply calculated as

$$\widehat{\mathbf{A}} = \underset{\mathbf{A}}{\operatorname{argmin}} \frac{1}{2}\|\mathbf{Y} - \mathcal{F}(\widehat{\mathbf{S}}\mathbf{A})\|_F^2, \tag{V.44}$$

$\widehat{\mathbf{S}}$ being the current estimate of the components matrix. Among all the methods previously considered for comparison, PSFEx is the only one handling the undersampling.



(a) Average error on the ellipticity vector.

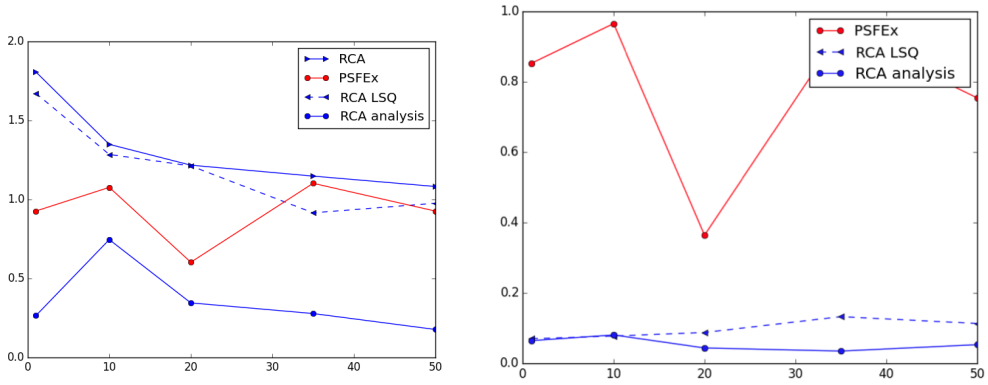
(b) Dispersion of the ellipticity vector.

Figure *V.8* – x axis : SNR (see section D.2) ; y axis : $\log_{10}(\mathrm{E}_\gamma)$ for the left plot, $\log_{10}(\mathrm{B}_\gamma)$ for the right plot.

As for the dimension reduction experiment, the proposed method with $\mathbf{\Phi}_s$ chosen as a wavelet dictionary is at least one order of magnitude more accurate over the shape parameters and the mean square error. Besides, Fig. ***V.10*** shows that the proximity constraint over the matrix $\mathbf{A}$ allows one to select a significantly better optimum than a simple least square update of $\mathbf{A}$. Indeed, regularizing the weight matrix estimation reinforces the rejection of $\mathcal{F}$'s null space.

As previously, we restored the complete field of Fig. ***V.3*** for a linear SNR of 40, using "RCA Analysis", with undersampled input PSFs as shown in Fig. ***V.11***.

The figure ***V.12*** shows consistent results with the dimension reduction experiment. In particular, the corners PSFs restoration is obviously more accurate.



(a) Average absolute error on the size. **(b)** Dispersion of the errors on the size.

Figure *V.9* – x axis : SNR ; y axis : E_S for the left plot, σ_S for the right plot.

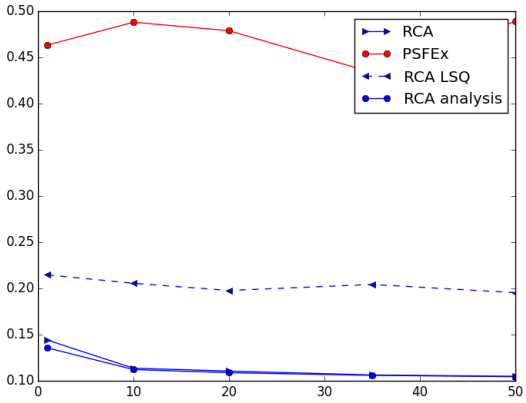


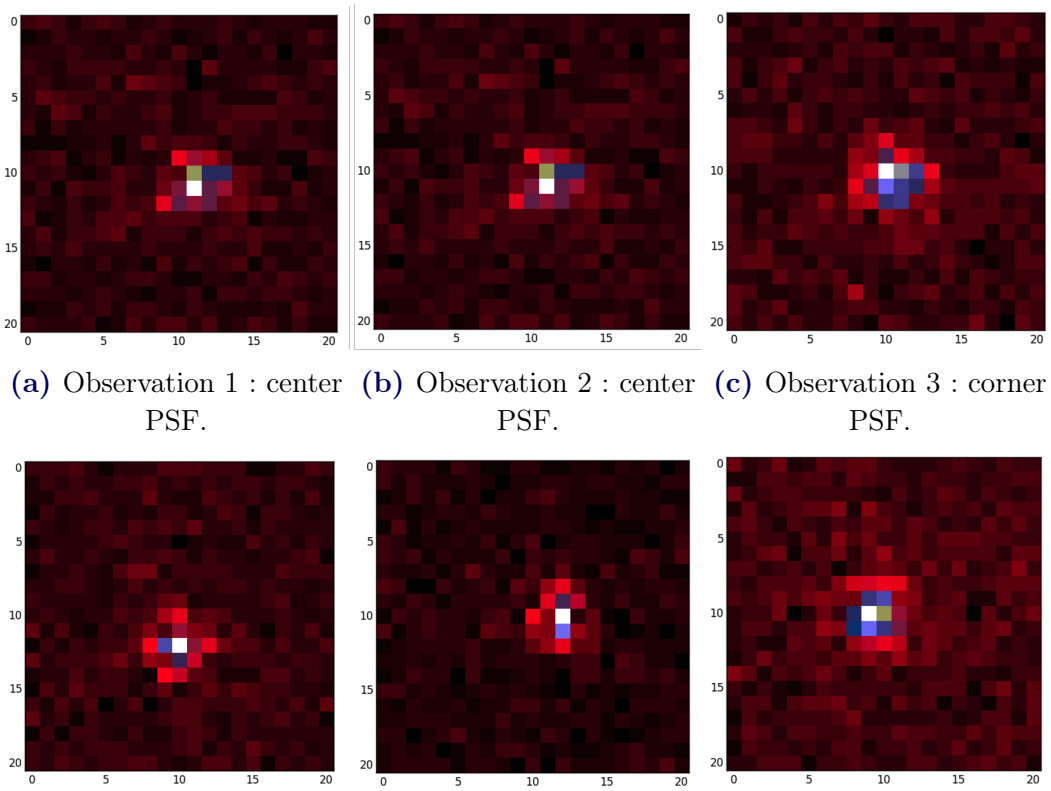
Figure *V.10* – Average normalized least square error

F Reproducible research

In the spirit of participating in reproducible research, the data and the codes used to generate the plots presented in this paper will be made available at `http://www.cosmostat.org/software/rca/`.

G Conclusion

We introduced RCA which is a dimension reduction method for continuous and positive data field which is noise robust and handles undersampled data.



(a) Observation 1 : center PSF. **(b)** Observation 2 : center PSF. **(c)** Observation 3 : corner PSF.

(d) Observation 4 : corner PSF. **(e)** Observation 5 : "local" PSF. **(f)** Observation 6 : "local" PSF.

Figure *V.11* – Input PSFs at different locations in the field for a SNR = 40.

As a linear dimension reduction method, RCA computes the input data as linear combinations of few components which are estimated, as well as the linear combination coefficients, through a matrix factorization.

The method was tested over a field of simulated Euclid telescope PSFs. We show that constraining both the components matrix and the coefficients matrix using sparsity yield at least one order of magnitude more accurate PSFs restoration than existing methods, with respect to the PSFs shapes parameters. In particular, we show that the analysis formulation of the sparsity constraint over the components is particularly suitable for capturing accurately the PSFs sizes. We also show that constraining the coefficients matrix yields a significantly better identification of the PSFs embedding subspace when the data are undersampled.

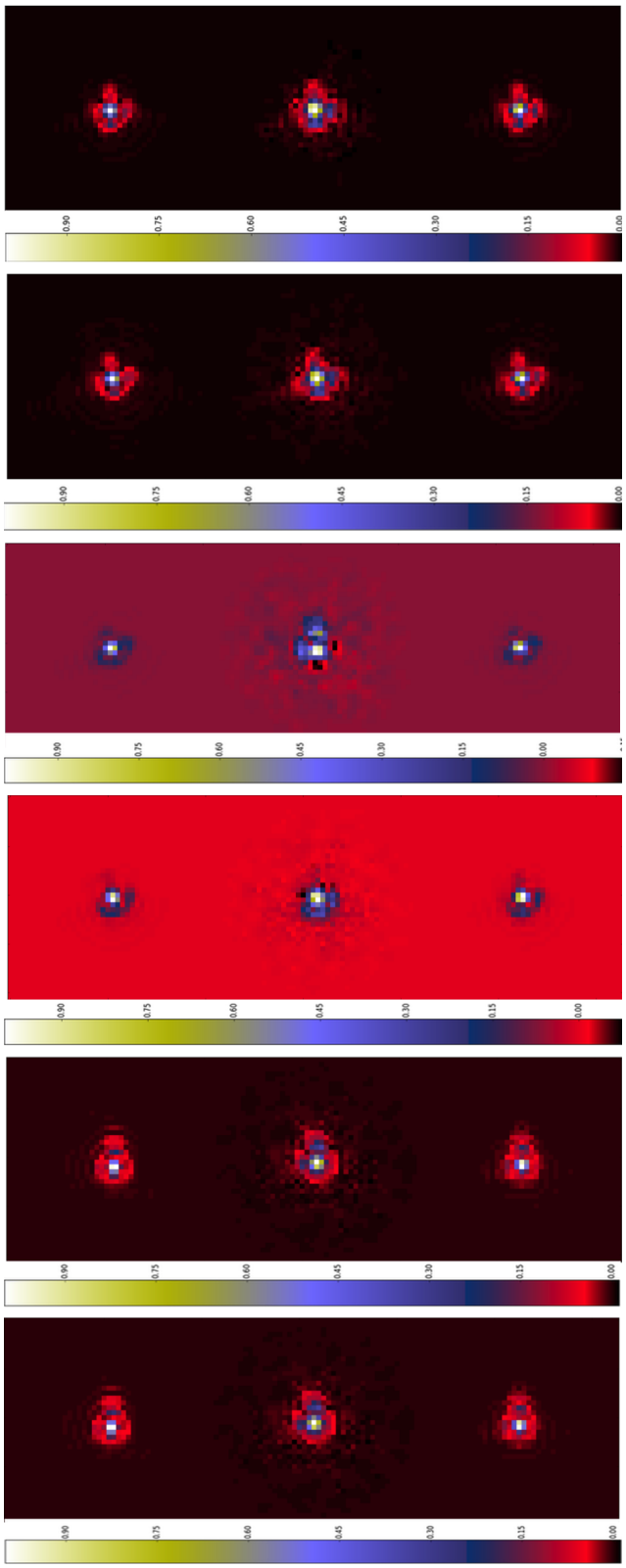


Figure *V.12* – PSFs reconstructions : from the left to the right : original, PSFEx, RCA ; from the bottom to the top : 2 "local" PSFs reconstructions, 2 corner PSFs reconstructions, 2 center PSFs reconstructions. The observed corresponding PSFs can be seen in Fig. ***V.11***

Chapitre VI

PSF field interpolation

Sommaire

This chapter focuses on interpolating a PSF field at arbitrary locations from a set of perfectly known PSFs spread out across the fov. Assuming that the PSF can be modeled locally as a convolution kernel, we use "PSF field" to refer to a continuous function mapping each point in the instrument focal plane surface to the corresponding convolution kernel. This problem has driven a lot of attention within the astronomers community over the last decade [Gentile *et al.* 2013], especially due to the strengthening of accuracy requirements in the recent and future spatial surveys. We review some of the state-of-the-art method in the next section.

A PSF field interpolation state-of-the-art

Most of the PSF interpolation methods fall into one of the two following categories :

- optical model based methods which derives the PSF from a physical model of the instrument itself ;
- data driven methods which estimate the PSFs at given locations in the fov based on local measurements extracted from real images.

In the first category, we can mention the work in [Jarvis *et al.* 2008]. The authors built a parametric model of the PSF shape and size by modeling individually the main physical causes which make the PSF variable ; they found that this model is able to reproduce a substantial part of the PSF anisotropy and size for the ground based telescope Blanco telescope. In [Stabenau *et al.* 2007], the authors simulate the PSF time variations due to thermal drift, jitter and structural vibrations, using the method of ray tracing, for the SNAP telescope. However, these methods require making assumptions on the instrument physic, and not all the phenomena at play in the image forming process can be satisfactorily modeled. Thus, we only consider model-free methods for comparisons in the numerical experiments.

The data driven approaches rely on the fact that in practice, unresolved stars images in the observed field give a measurement of the PSFs at those stars locations. One recurrent scheme consists in first expanding the observed unresolved stars over some given analytic function basis. This gives a more compact representation of the data (and potentially allows denoising). The PSF field is then obtained through a 2D polynomial fitting typically, of the representation coefficients for each element of the chosen basis. In [Romano *et al.* 2010], this scheme is applied to model the PSF of the Large Binocular Cameras, using the Shapelets basis ([Refregier 2003, Massey & Refregier 2005]). However, it might not be possible to capture all the PSF structures through a finite expansion over an analytic function basis. One can instead learn the representation basis from the data themselves. For example in [Jee *et al.* 2007], a principal component analysis (PCA) is performed over hundreds archival images of stars from the Hubble Space Telescope (HST) Advanced Camera for Surveys (ACS), from different exposures. The optimal representation basis is then chosen as the first principal components and the representation coefficients are fitted in each exposure separately with a bivariate polynomial. In [Jarvis & Jain 2004], the PSF modeling also relies on a PCA, not of the images themselves, but instead of a set of features (for example a shape parameter). In the two aforementioned methods, fitting a global polynomial model can yield an oversmoothing of the PSF field, or instead

a Runge phenomenon, if the polynomial's degree is too high [Epperson 1987]. In [Gentile *et al.* 2013], the PSF modeling is recast as a spatial interpolation problem. In general statistics, the spatial interpolation methods aim at interpolating a field exploiting the spatial correlation of the data. Some of these methods, for example the Kriging, are presented and applied to PSF interpolation in the previously mentioned work. We consider the two best performing approaches tested in [Gentile *et al.* 2013] for comparisons in the numerical experiments. More recently a PSF interpolation scheme has been proposed in [Suksmono 2013] based on, compressive sampling (CS) ideas. The method aims at computing the PSF ellipticity paramters at every points over an uniform grid using the available randomly distributed measurements. This yields an ill-posed inverse problem which is regularized by assuming that the shape parameters constitute a compressible field in frequency domain. However, this approach has a narrow scope of applications since the "full" PSF itself is not estimated.

Let finally mention a somehow hybrid approach. Indeed in [Piotrowski *et al.* 2013], the authors propose to build an accurate polynomial model of the PSF of the "Pi of the Sky" telescope from detailed laboratory measurements of optical PSF, pixels sensitivity and pixels response. This model is then tuned to real sky data.

In this paper, we propose a non-parametric interpolation method that relies on geometric tools. We propose an intuitive framework for learning a PSFs set underlying geometry using Optimal Transport distances. In Section C, we introduce and motivate the mathematical framework ; we describe the proposed method and algorithm in Section D ; we present some numerical experiments and discuss the results in Section E ; the interpolation algorithm parameters and some numerical considerations as discussed in Section F.

B Notations

Each PSF is treated either as a matrix $\mathbf{X} = (x_{ij})_{\substack{1\leq i\leq n\\ 1\leq j\leq n}}$ or as a vector in $\mathbf{R}^{n^2}$, $\mathbf{x} = (x_k)_{1\leq k\leq n^2}$. For $N \in \mathbb{N}$, we denote Σ_N the set of permutations of $[\![1, N]\!]$. We denote $\mathcal{P}(\mathbb{R}^d)$ the set of discrete probability measures over $\mathbb{R}^d$. We denote $\mathbb{1}_n$ the column vector of n ones. We denote $\mathbf{I}_p$ the identity matrix of size $p \times p$.

C Numerical Optimal Transport

C.1 Motivation

We consider a field of monochromatic PSFs. The only factor of variability is the unresolved object's position in the fov. One can see from Fig. ***VI.1*** that PSFs shapes vary non-linearly across the field. This is related to differences in the paths an electromagnetic wave would follow from the telescope pupil entrance to the focal plane, depending on the source position ; this results in different diffraction's figures, which are the observed PSFs.

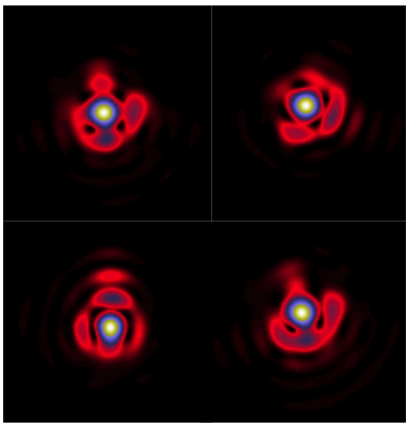


Figure *VI.1* – Simulated Euclid telescope PSFs at different locations in the fov

This can be interpreted as the curvature in $\mathbb{R}^{n^2}$ of the PSFs manifold, occurring between distant PSFs in the fov. As Fig. ***VI.2*** shows, we want to combine available PSFs measurements given by unresolved objects images to estimate the PSFs at galaxies locations in the fov.

Optimal Transport (OT) appears to be suitable to tackle this problem. Indeed, it gives a way of measuring distances between PSFs which accounts for the aforementioned curvature.

Besides, as already highlighted in Chapter III, Section B, it comes with a recipe for computing geodesics which is interesting for interpolation. This is detailed in the next section. Moreover, it is classical to assume that a PSF has an unitary l_1 norm. This implies a constraint of mass conservation in the interpolation which is naturally integrated in the OT framework. In the next section, we precise the OT tools and concepts we use.

C.2 General notions

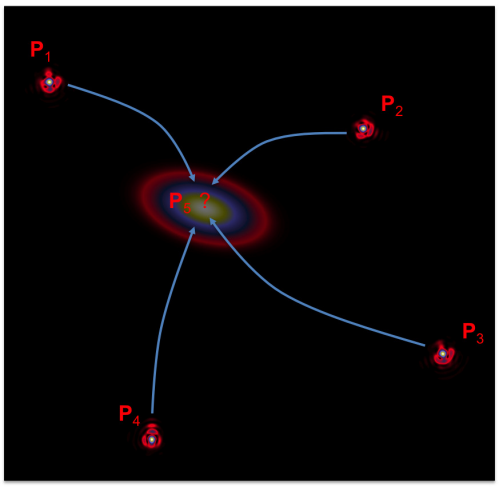


Figure *VI.2* – PSF interpolation problem

C.2.1 Sliced Wasserstein Distance

We consider two points matrices $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_N]$ and $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_N]$ in $\mathbb{R}^{d\times N}$, for some integers d and N. By abuse of language, we do not distinguish these matrices and the point clouds in $\mathbb{R}^d$ they represent, where there is no risk of confusion. We define the distributions $\mu_{\mathbf{X}} = \sum_{i=1}^N \frac{1}{N}\delta_{\mathbf{x}_i}$ and $\mu_{\mathbf{Y}} = \sum_{i=1}^N \frac{1}{N}\delta_{\mathbf{y}_i}$, where for $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^d$, $\delta_{\mathbf{x}}(\mathbf{y}) = 1$ if $\mathbf{x} = \mathbf{y}$ and $\delta_{\mathbf{x}}(\mathbf{y}) = 0$ otherwise. The optimal assignment problem between $\mu_{\mathbf{X}}$ and $\mu_{\mathbf{Y}}$ consists in finding a permutation $\sigma \in \Sigma_N$ which minimizes

$$\sum_{i=1}^{N} \|\mathbf{x}_i - \mathbf{y}_{\sigma(i)}\|^p, \tag{VI.1}$$

where $\|.\|$ is a norm in $\mathbb{R}^d$ and $p \geq 1$. Problem VI.1 can be rewritten as a particular instance of the Monge-Kantorovich mass transportation problem [Kantorovitch 1942]. We note σ^* a minimizer of Problem VI.1. It has been shown that $W_p(\mu_{\mathbf{X}}, \mu_{\mathbf{Y}}) = (\sum_{i=1}^N \|\mathbf{x}_i - \mathbf{y}_{\sigma^*(i)}\|_p^p)^{\frac{1}{p}}$ defines a distance on $\mathcal{P}(\mathbb{R}^d)$ which is the so-called Wasserstein Distance p of the two distributions. The norm $\|.\|_p$ is often referred to as the ground metric and is an important parameter of the Transport Distances in practical applications. Examples of theoretical and computational applications of optimal transport can be found in [Papadakis *et al.* 2013].

Let explicit the notion of *displacement interpolation* in this discrete setting.

Let consider a point $\mathbf{x}_i$ in the first point cloud and the assigned point $\mathbf{y}_{\sigma^*(i)}$ in the second point cloud. Let $\gamma_i : [0, 1] \mapsto \mathbb{R}^d$ a curve verifying $\gamma_i(0) = \mathbf{x}_i$ and $\gamma_i(1) = \mathbf{y}_{\sigma^*(i)}$. The action on a material particle moving from $\mathbf{x}_i$ to $\mathbf{y}_{\sigma^*(i)}$

along γ_i between time s and time t might be defined as

$$\mathcal{A}_p^{s,t}(\gamma_i) = \int_s^t \|\dot{\gamma}_i(\tau)\|^p d\tau, \tag{VI.2}$$

where $\dot{\gamma}_i(\tau)$ is the particle velocity at time τ. We define the set $\Gamma_i = \{\gamma : [0,1] \mapsto \mathbb{R}^d, \gamma(0) = \mathbf{x}_i, \gamma(1) = \mathbf{y}_{\sigma^*(i)}\}$. An action minimizing path between $\mathbf{x}_i$ and $\mathbf{y}_{\sigma^*(i)}$ is a solution of

$$\min_{\gamma \in \Gamma_i} \mathcal{A}_p^{s,t}(\gamma_i). \tag{VI.3}$$

Interestingly, it has been shown that a geodesic path between the distributions $\mu_{\mathbf{X}}$ and $\mu_{\mathbf{Y}}$ in $\mathcal{P}(\mathbb{R}^d)$ equipped with the Wassertein distance can be computed by advecting particles along action minimizing paths between pairs of assigned points in the two point clouds [Villani 2009]. Concretely, let γ_i* denotes an action minimizing path between $\mathbf{x}_i$ and $\mathbf{y}_{\sigma^*(i)}$. For $t \in [0,1]$, we define the discrete distribution

$$\mu_t = \sum_{i=1}^{N} \frac{1}{N} \delta_{\gamma_i*(t)}. \tag{VI.4}$$

Then the parametric curve

$$\Gamma_{XY} : [0,1] \mapsto \mathcal{P}(\mathbb{R}^d), t \to \mu_t \tag{VI.5}$$

is a geodesic path in $\mathcal{P}(\mathbb{R}^d)$ between the distributions $\mu_{\mathbf{X}}$ and $\mu_{\mathbf{Y}}$. This realizes a *displacement interpolation* of the two considered distributions ; this powerful notion had first been introduced in [McCann 1997]. For a given t, the distribution μ_t can be seen as the barycenter of the two end distributions with the barycentric weights $1-t$ and t respectively in the Wasserstein metric.

For $d = 1$, i.e. in the 1D case, σ^* is known in closed form : if we consider two permutations σ_X and σ_Y in Σ_N verifying

$$\mathbf{x}_{\sigma_X(1)} \leq \cdots \leq \mathbf{x}_{\sigma_X(N)} \tag{VI.6}$$

$$\mathbf{x}_{\sigma_Y(1)} \leq \cdots \leq \mathbf{x}_{\sigma_Y(N)}, \tag{VI.7}$$

then $\sigma^* = \sigma_Y \circ \sigma_X^{-1}$, where σ_X^{-1} is a permutation verifying $\sigma_X^{-1} \circ \sigma_X = \mathrm{Id}$. For $d > 1$, there is no known closed-form expression for σ^*. Various methods have been proposed for solving Problem VI.1 ; in particular, it can be recasted as a linear program. Yet, the fastest known algorithms have a running time of $O(N^{2.5}\log(N))$ [Burkard *et al.* 2009], which is prohibitive for large scale image processing applications. This has motivated the introduction of the sliced Wasserstein Distance [Rabin *et al.* 2011, Bonneel *et al.* 2015], which consists in the sum of 1D Wassertein distances of the projected point clouds :

$$SW_p(\mu_{\mathbf{X}}, \mu_{\mathbf{Y}})^p = \int_{\mathbb{S}^{d-1}} W_p(\mu_{\mathbf{X}_\mathbf{u}}, \mu_{\mathbf{Y}_\mathbf{u}})^p d\mathbf{u}, \tag{VI.8}$$

where $p \geq 1$, $\mathbb{S}^{d-1} = \{\mathbf{x} \in \mathbb{R}^d / \|\mathbf{x}\| = 1\}$, $\mathbf{X_u} = \{\mathbf{u}^T \mathbf{x}_i, i = 1...N\} \subset \mathbb{R}^N$ and $\mathbf{Y_u}$ is similarly defined.

C.2.2 Sliced Wassertein Barycenter

We consider a family $\{\mathbf{X}_1, \ldots, \mathbf{X}_K\}$ of point clouds in $\mathbb{R}^d$: for $k \in [\![1, K]\!]$, $\mathbf{X}_k = [\mathbf{x}_{k1}, \ldots, \mathbf{x}_{kN}] \in \mathbb{R}^{d \times N}$. The associated distributions $\mu_{\mathbf{X}_k}$ are defined as previously. A barycenter of these distributions in the Wasserstein metric is defined as

$$\mu_{\text{Bar}} = \underset{\mu}{\text{argmin}} \sum_{k=1}^{K} w_k W_p(\mu, \mu_{\mathbf{X}_k})^p, \tag{VI.9}$$

for some positive weights w_k verifying $\sum_{k=1}^{K} w_k = 1$. Somehow, this generalizes the concept of displacement interpolation for an arbitrary finite number of distributions. In various applications Wasserstein barycenters appear to be more suitable than Euclidean or more sophisticated barycenters ([Cuturi & Doucet 2013], [Gramfort *et al.* 2015]) ; in particular they are robust to shifts and "elastic" deformations. As previously, in 1D (i.e. $d = 1$) and for $p = 2$, this barycenter is known in closed-form. It is the uniform discrete probability measure defined over the set

$$\{\mathbf{x}_{\text{Bar}i} = \sum_{k=1}^{K} w_k \mathbf{x}_k[\sigma_k(i)], i = 1 \ldots N\}, \tag{VI.10}$$

where σ_k is a permutation in Σ_N which sorts $\mathbf{x}_k$ entries (i.e. $\mathbf{x}_k[\sigma_k(1)] \leq \cdots \leq \mathbf{x}_k[\sigma_k(N)]$).

For $d > 1$, computing this barycenter is in general intractable [Gangbo & Swiech 1998]. Therefore in this paper we consider the Sliced Wasserstein Barycenter defined by replacing the Wasserstein distance by its sliced counterpart in the Wasserstein barycenter definition :

$$\mu_{\text{SBar}} = \underset{\mu}{\text{argmin}} \sum_{k=1}^{K} w_k SW_p(\mu, \mu_{\mathbf{X}_k})^p. \tag{VI.11}$$

C.3 Data representation in the Transport Framework

We consider a $N_l \times N_c$ gray levels image. We denote $\mathbf{x} = (x_i)_{1 \leq i \leq N}$ the vector of pixels intensities rearranged in lines lexicographic order, with $N = N_l N_c$. The image is treated as a point cloud defined as

$$\mathbf{X} = [\mathbf{v}_1, \cdots, \mathbf{v}_N], \text{with } \mathbf{v}_i = [x_i, \lfloor \frac{i-1}{N_c} \rfloor, (i - 1(\text{mod} N_c))]^T. \tag{VI.12}$$

This point cloud can be viewed as a representation of the considered image as a discrete surface in $\mathbb{R}^3$. Given a set of unresolved objects images, the idea in the following section would be to match and or compute barycenters of the associated surfaces based on sliced transport.

We use a ground metric $\mathcal{C}$ taking the following form : for two vectors $\mathbf{p} = [p_1, p_2, p_3]^T$ and $\mathbf{q} = [q_1, q_2, q_3]^T$ in $\mathbb{R}^3$,

$$\mathcal{C}^2(\mathbf{x}, \mathbf{y}) = (p_1 - q_1)^2 + \beta^2 * ((p_2 - q_2)^2 + (p_3 - q_3)^2), \tag{VI.13}$$

where β is a strictly positive real. As a reminder, in the applications, the first components p_1 and q_1 will be pixels intensities and the two other components will be pixels positions. For two images $\mathbf{x} = (x_i)_{1 \leq i \leq N}$ and $\mathbf{y} = (y_i)_{1 \leq i \leq N}$, let $\mathbf{X}$ and $\mathbf{Y}$ be the associated point clouds and σ^* an optimal assignment of $\mathbf{Y}$'s points to $\mathbf{X}$'s points : then

$$W_2(\mu_{\mathbf{X}}, \mu_{\mathbf{Y}}) = \sum_{i=1}^{N} (x_i - y_{\sigma^*(i)})^2$$
$$+\beta^2 * ((\lfloor \frac{i-1}{N_c} \rfloor - \lfloor \frac{\sigma^*(i)-1}{N_c} \rfloor)^2 + ((i - 1(\mathrm{mod} N_c)) - (\sigma^*(i) - 1(\mathrm{mod} N_c)))^2). \tag{VI.14}$$

Thus, if σ^* is the identity, $W_2(\mu_{\mathbf{X}}, \mu_{\mathbf{Y}}) = \|\mathbf{x} - \mathbf{y}\|_2$. Therefore, we can consider W_2 (and SW_2) as a generalization of the euclidian distance in the pixels domain.

Conversely, from a given point cloud $\mathbf{Z} \in \mathbb{R}^{3 \times N}$, we can go backward to an image by discretizing the 2D function $\Phi[\mathbf{Z}]$ defined as

$$\Phi[\mathbf{Z}](x, y) = \sum_{i=1}^{N} \mathbf{Z}[1, i] \delta_{\mathbf{Z}[2,i], \mathbf{Z}[3,i]}(x, y). \tag{VI.15}$$

D Transport based PSF field interpolation

We consider a set of K PSFs $\mathcal{S} = (\mathbf{x}_k)_{1 \leq k \leq K}$, located at the positions $(\mathbf{u}_k)_{1 \leq k \leq K}$ in the fov. We note $(\mathbf{X}_k)_{1 \leq k \leq K}$ the associated point clouds as defined in Section C.3. We want to estimate the PSF at a new location $\mathbf{u}$. We note this PSF $\mathbf{x}_\mathbf{u}$ and we note $\mathbf{X}_\mathbf{u}$ the associated point cloud :

$$\mathbf{x}_\mathbf{u} \equiv \Phi[\mathbf{X}_\mathbf{u}], \tag{VI.16}$$

where the operator Φ is defined in Eq.VI.15. We want to estimate $\mathbf{X}_\mathbf{u}$ as a Wassertein barycenter of a subset of PSFs indexed by $\mathcal{I}(\mathbf{u})$ in the neighborhood of $\mathbf{u}$ in the fov :

$$\mathbf{X}_\mathbf{u} = \underset{\mathbf{X}}{\mathrm{argmin}} \sum_{k \in \mathcal{I}(\mathbf{u})} w_k(\mathbf{u}) SW_2(\mu_{\mathbf{X}}, \mu_{\mathbf{X}_k})^2, \tag{VI.17}$$

with $w_k(\mathbf{u}) \geq 0$ and $\sum_{k\in\mathcal{I}(\mathbf{u})} w_k(\mathbf{u}) = 1$. $\mathcal{I}(\mathbf{u})$ simply indexes the nearest neighbors of $\mathbf{u}$ in the set of locations $(\mathbf{u}_k)_{1\leq k\leq K}$. The question of the size of this neighborhood is postponed to Section F.1.

The weights $(w_k(\mathbf{u}))_{k\in\mathcal{I}(\mathbf{u})}$ can be seen as generalized barycentric coordinates of $\mathbf{X}_\mathbf{u}$ relatively to the clouds $(\mathbf{X}_k)_{k\in\mathcal{I}(\mathbf{u})}$, in the Wasserstein metric.

These weights are calculated in three steps that we detail in the following sections :

— Embedding : we calculate the pairwise approximated Wasserstein 2 distances over the set $(\mathbf{X}_k)_{k\in\mathcal{I}(\mathbf{u})}$ and determine an euclidean embedding of $(\mathbf{X}_k)_{k\in\mathcal{I}(\mathbf{u})}$ that preserves the Wasserstein distances ; we get a set of coordinates $(\mathbf{r}_k)_{1\leq k\leq|\mathcal{I}(\mathbf{u})|}$, where $|\mathcal{I}(\mathbf{u})|$ is the number of elements in $\mathcal{I}(\mathbf{u})$;

— Interpolation : we estimate the representation $\mathbf{r}_\mathbf{u}$ of $\mathbf{X}_\mathbf{u}$ in the previously calculated embedding by interpolating the set $(\mathbf{r}_k)_{1\leq k\leq|\mathcal{I}(\mathbf{u})|}$ coordinates-wise ;

— Weights setting : the weights $(w_k(\mathbf{u}))_{k\in\mathcal{I}(\mathbf{u})}$ are calculated as the euclidean barycentric coordinates of $\mathbf{r}_\mathbf{u}$ relatively to $(\mathbf{r}_k)_{1\leq k\leq|\mathcal{I}(\mathbf{u})|}$.

D.1 Local non-linear dimension reduction

We recall that we want to estimate the PSF at a position $\mathbf{u}$ in the fov, given the PSFs located at the positions $(\mathbf{u}_k)_{1\leq k\leq K}$. We consider the p nearest neighbors of $\mathbf{u}$ in $(\mathbf{u}_k)_{1\leq k\leq K}$. This defines the set $\mathcal{I}(\mathbf{u})$ aforementioned.

As we will see in the Numerical experiments Section, the parameter β in Eq.VI.14 is set so that for the closest PSFs in the fov, the Wasserstein distance is equal to the euclidian distance. Indeed, the euclidean distance is usually assumed to be a good approximation of the geodesic distance between close points on a given smooth Manifold (see for example [Tenenbaum *et al.* 2000, Donoho & Grimes 2003]). However, the more two PSFs are distant in the fov, the more the PSFs manifold curvature manifests through the relative warping of their structures. By minimizing the amount of work needed to push one of the PSFs toward the other, the Wasserstein metric can keep track of the warping to a certain extent, thus unfolding locally the Manifold. For this reason, the Wasserstein distance can potentially give a faithful approximation of the geodesic distances, on broader neighborhoods than the euclidean distance. We define the $p \times p$ local pairwise distance matrix $\mathbf{M}_W(\mathbf{u})$ as :

$$\mathbf{M}_W(\mathbf{u})[i,j] \approx W_2(\mu_{\mathbf{X}_{\mathcal{I}(\mathbf{u})[i]}}, \mu_{\mathbf{X}_{\mathcal{I}(\mathbf{u})[j]}}),\ (i,j) \in [\![1,p]\!]^2. \tag{VI.18}$$

Then we apply the Multidimensional Scaling (MDS) procedure to calculate the embedding [Abdi 2007]. The first step consists in converting the distances matrix $\mathbf{M}_W(\mathbf{u})$ into a Gram matrix $\mathbf{X}_i$ i.e. whose entries are given by $\mathbf{X}_i[i,j] =$

$\langle \mathbf{r}_i, \mathbf{r}_j \rangle$, so that $\mathbf{M}_W(\mathbf{u})[i,j] = \|\mathbf{r}_i - \mathbf{r}_j\|_2$. The vector $\mathbf{r}_i$ is the low dimensional embedding of $\mathbf{X}_{\mathcal{I}(\mathbf{u})[i]}$.

For this purpose we introduce the "centering matrix" [Marden 1995] defined as

$$\mathbf{C} = \mathbf{I}_p - \frac{1}{p}\mathbb{1}_p\mathbb{1}_p^T, \tag{VI.19}$$

where $\mathbb{1}_p$ is a row vector made of p ones. Multiplying a matrix by $\mathbf{C}$ on the left has the effect of subtracting each row its mean ; $\mathbf{C}$ being symmetric, multiplying a matrix by $\mathbf{C}$ on the right subtracts each lines its mean.

The cross-product matrix is calculated as

$$\mathbf{X}_i = -\frac{1}{2}\mathbf{C}\mathbf{M}_W(\mathbf{u})^2\mathbf{C}, \tag{VI.20}$$

where $\mathbf{M}_W(\mathbf{u})$ is squared entry-wise. We define the embedding coordinates matrix $\mathbf{R} = [\mathbf{r}_1, \ldots, \mathbf{r}_n]$. Assuming that such an embedding exists and $\mathbf{R}$ has null lines and columns means, one can show that $\mathbf{X}_i = \mathbf{R}^T\mathbf{R}$ (see appendix F). Under this hypothesis, $\mathbf{X}_i$ is a symmetric matrix and therefore can be orthogonally diagonalized :

$$\mathbf{X}_i = \mathbf{V}^T\mathbf{S}\mathbf{V}, \tag{VI.21}$$

where $\mathbf{V}^T\mathbf{V} = \mathbf{V}\mathbf{V}^T = \mathbf{I}_p$ and $\mathbf{S}$ is a non-negative diagonal matrix. It follows that $\mathbf{R}$ can be calculated as

$$\mathbf{R} = \mathbf{Q}\mathbf{S}^{\frac{1}{2}}\mathbf{V}, \tag{VI.22}$$

where $\mathbf{Q}$ is an orthogonal matrix that we set to the identity in practice [Dokmanic *et al.* 2015]. Depending on neighborhood's size p, the last diagonal values of $\mathbf{S}$ can be neglected so that the last lines of $\mathbf{V}$ can be discarded ; besides this represents a way of analyzing locally the Manifold dimensionality [Tenenbaum *et al.* 2000].

D.2 Field-of-view mapping

In this section, we estimate the local low dimensional embedding of the unknown PSF located at $\mathbf{u}$ in the fov. We note this embedding $\mathbf{r}_\mathbf{u}$. We note d the dimension of the vectors $(\mathbf{r}_i)_{1\leq i\leq p}$ determined in the previous section, and $\mathbf{r}_\mathbf{u}$. To compute the i^{th} component of $\mathbf{r}_\mathbf{u}$ we determine an interpolating function

$$f_i : \mathbb{R}^2 \to \mathbb{R} \ / f_i(\mathbf{u}_j) = \mathbf{r}_j[i], \ \forall j \in [\![1, p]\!]. \tag{VI.23}$$

This is a standard surface interpolation problem that we solve using the so-called thin-plate spline [Eberly 2002] which is appealing because of its physical

interpretation : it is the exact interpolating function that minimizes the "bending energy" defined as :

$$E(f) = \int_{\mathbb{R}^2} (\frac{\partial^2 f(x,y)}{\partial^2 x})^2 + 2 * (\frac{\partial^2 f(x,y)}{\partial x \partial y})^2 + (\frac{\partial^2 f(x,y)}{\partial^2 y})^2. \tag{VI.24}$$

Thus, the function f_i takes the following form :

$$f_i(\mathbf{x}) = \sum_{j=1}^{p} a_{ij} \|\mathbf{x} - \mathbf{u}_j\|_2^2 \ln(\|\mathbf{x} - \mathbf{u}_j\|_2), \tag{VI.25}$$

and the coefficients a_{ij} are calculated so that f_i takes the prescribed values at the control points $(\mathbf{u}_j)_{1 \leq j \leq p}$. This way, each component of $\mathbf{r_u}$ can be estimated.

D.3 Barycentric coordinates

The embedding of the PSF at the position $\mathbf{u}$ in the fov has been determined. Now, we want to estimate the point cloud $\mathbf{X_u}$ from this embedding and the neighbor PSFs. In other terms, we need to determine the weights $w_k(\mathbf{u})$ in Eq. VI.17. In Section D.1, we calculated an isometric embedding of the matrices $(\mathbf{X}_k)_{k \in \mathcal{I}(\mathbf{u})}$. Because of this isometry, for a set of positive weights $(w_i)_{1 \leq i \leq p}$ verifying $\sum_{i=1}^{p} w_i = 1$, the problem

$$\min_{\mathbf{x}} \sum_{i=1}^{p} w_i \|\mathbf{x} - \mathbf{r}_i\|_2^2 \tag{VI.26}$$

is equivalent to the following

$$\min_{\mathbf{X}} \sum_{i=1}^{p} w_i SW_2(\mu_{\mathbf{X}}, \mu_{\mathbf{X}_{\mathcal{I}(\mathbf{u})[i]}})^2. \tag{VI.27}$$

This somehow goes along the same line as the Locally Linear Embedding [Roweis & Saul 2000] and it provides a mean for computing $\mathbf{X_u}$ from $\mathbf{r_u}$. Indeed, we first can consider the following barycentric coordinates problem :

$$\min_{w_1,\ldots,w_p} \frac{1}{2} \|\mathbf{r_u} - \sum_{i=1}^{p} w_i \mathbf{r}_i\|_2^2 \text{ s.t.} w_i \geq 0 \text{ and } \sum_{i=1}^{p} w_i = 1. \tag{VI.28}$$

We note $(w_i^*)_{1 \leq i \leq p}$ the optimal tuple and we define $\mathbf{r}_\mathbf{u}^* = \sum_{i=1}^{p} w_i^* \mathbf{r}_i$. $\mathbf{r_u} \approx \mathbf{r}_\mathbf{u}^*$ and $\mathbf{r}_\mathbf{u}^*$ is solution of the Problem VI.26, with the weights $(w_i^*)_{1 \leq i \leq p}$. Therefore, we compute $\mathbf{X_u}$ as

$$\mathbf{X_u} = \operatorname*{argmin}_{\mathbf{X}} \sum_{i=1}^{p} w_i^* SW_2(\mu_{\mathbf{X}}, \mu_{\mathbf{X}_{\mathcal{I}(\mathbf{u})[i]}})^2. \tag{VI.29}$$

This approach is in a way similar to the procedure described in [Gower 1968]. Besides, the task of estimating barycentric coordinates in the Wasserstein metric has been recently addressed in [Bonneel *et al.* 2016], in a different numerical Optimal Transport framework. The final image is obtained by discretizing the function $\Phi[\mathbf{X_u}]$ defined by Eq.VI.15. This step is discussed in Section F.

D.4 Algorithm

We want to estimate D PSFs at random locations $(\mathbf{v}_i)_{1\leq i\leq D}$ from a set of K PSFs at known locations in the fov (see the Section D introduction). The whole procedure is summarized in Algorithm *VI*.1.

Algorithme VI.1 Transport Interpolation (TraIn)

Inputs : K PSFs $(\mathbf{x}_k)_{1\leq k\leq K}$, K observations locations $(\mathbf{u}_k)_{1\leq k\leq K}$, D interpolation locations $(\mathbf{v}_k)_{1\leq k\leq D}$, number of neighbors p, local dimensionality $d \leq p$
Compute the weighting parameter β (see Eq.VI.14)
Compute the p neighbors of each location $\mathbf{v}_i$ in the set $(\mathbf{u}_k)_{1\leq k\leq K}$; this results in a collection of sets of indices $(\mathcal{I}(\mathbf{v}_i))_{1\leq i\leq D}$
Transform the PSFs into point clouds (see Eq. VI.12) : $(\mathbf{x}_k)_{1\leq k\leq K} \rightarrow (\mathbf{X}_k)_{1\leq k\leq K}$
Compute the approximated Wassertein 2 distances between pairs of point clouds $(\mathbf{X}_l, \mathbf{X}_m)\ /(l,m) \in \mathcal{I}(\mathbf{v}_i)^2$ for some $i \in [\![1, D]\!]$
Pour $i = 1$ to D **faire**
    Form the pairwise Wasserstein distances matrix over the set $(\mathbf{X}_j)_{j\in\mathcal{I}(\mathbf{v}_i)}$ and compute a local euclidian embedding ; the results is a set of vectors $(\mathbf{r}_j)_{1\leq j\leq p}$ in $\mathbb{R}^d$
    Estimate the embedded coordinates at the location $\mathbf{v}_i$ using a thin-plate spline interpolation coordinate-wise ; this results in a vector $\mathbf{r}_{\mathbf{v}_i}$
    Compute the barycentric coordinates of $\mathbf{r}_{\mathbf{v}_i}$ relatively to the vectors $(\mathbf{r}_j)_{1\leq j\leq p}$
    Compute the approximate Wassertein barycenter of $(\mathbf{X}_j)_{j\in\mathcal{I}(\mathbf{v}_i)}$ using the previously calculated barycentric coordinates as weights
    Compute the interpolated PSF $\mathbf{x}_{\mathbf{v}_i}$ from this barycenter (see Eq. VI.15)
Return : $(\mathbf{x}_{\mathbf{v}_i})_{1\leq i\leq D}$.

This procedure is generic in the sense that the Wassertein metric can be replace by an arbitrary metric provided that one is able to compute geodesic distances and geodesics as we will see in Section F.

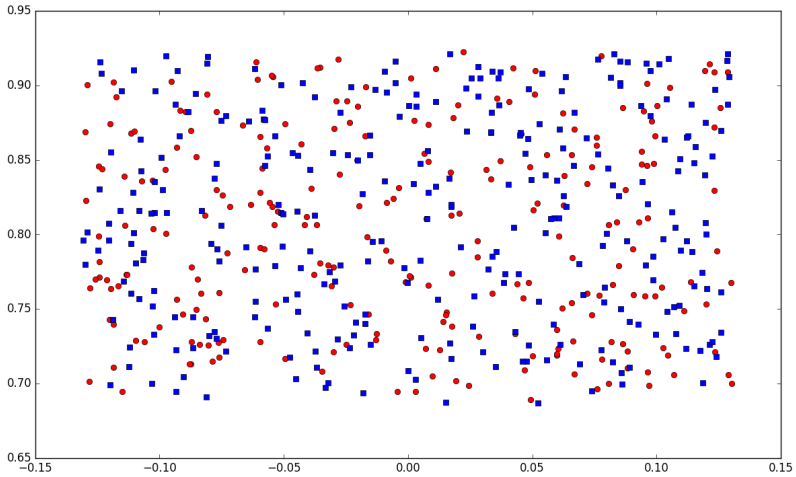


Figure *VI.3* – Simulated PSFs distribution across the fov ; the blue squares represents the 300 observed PSFs and the red circles the 250 PSFs to be interpolated ; the coordinates are in degrees.

E Numerical results

We tested the proposed method on a set of 550 simulated Euclid telescope optical PSFs. The PSFs are distributed in the fov according to Fig. ***VI.3***. We split the data into a "learning set" made of 300 observed PSFs and a "test set" corresponding to 250 unknown PSFs.

E.1 Quality assessment

As in the previous chapters, the quality assessment are the PSF's shape parameters. We recall their definitions thereafter. The central moments of an image $\mathbf{X} = (x_{ij})_{i,j}$ are defined as :

$$\mu_{p,q}(\mathbf{X}) = \sum_i \sum_j (i - i_c)^p (j - j_c)^p x_{ij} \tag{VI.30}$$

with $(p, q) \in \mathbb{N}^2$, (i_c, j_c) being the image centroid coordinates.

The ellipticity parameters are defined as follows :

$$e_1(\mathbf{X}) = \frac{\mu_{2,0}(\mathbf{X}) - \mu_{0,2}(\mathbf{X})}{\mu_{2,0}(\mathbf{X}) + \mu_{0,2}(\mathbf{X})} \tag{VI.31}$$

$$e_2(\mathbf{X}) = \frac{2\mu_{1,1}(\mathbf{X})}{\mu_{2,0}(\mathbf{X}) + \mu_{0,2}(\mathbf{X})}. \tag{VI.32}$$

The vector $\boldsymbol{\gamma}(\mathbf{X}) = [e_1(\mathbf{X}), e_2(\mathbf{X})]^T$ tells how much $\mathbf{X}$ departs from an isotropic distribution of luminosity and gives one the main direction of elongation.

The PSF's "size" is defined as

$$\mathrm{S}(\mathbf{X}) = \left(\frac{\sum_i \sum_j ((i-i_c)^2 + (j-j_c)^2) x_{ij}}{\sum_i \sum_j x_{ij}}\right)^{1/2}. \tag{VI.33}$$

We note $(\mathrm{Im}_i)_{1 \leq i \leq D}$ the set of tests PSFs and $(\widehat{\mathrm{Im}_i})_{1 \leq i \leq D}$ the set of corresponding PSFs interpolated with a given method. The reconstruction quality is accessed through the following quantities :

- the average error on the ellipticity vector : $\mathrm{E}_{\boldsymbol{\gamma}} = \sum_{i=1}^{D} \|\boldsymbol{\gamma}(\mathrm{Im}_i) - \boldsymbol{\gamma}(\widehat{\mathrm{Im}_i})\|_2 / D$;
- the average absolute error on the size : $\mathrm{E}_{\mathrm{S}} = \sum_{i=1}^{D} |\mathrm{S}(\mathrm{Im}_i) - \mathrm{S}(\widehat{\mathrm{Im}_i})| / D$ in pixels ;
- the average normalized mean square error : $\mathrm{NMSE} = \sum_{i=1}^{D} \|\widehat{\mathrm{Im}_i} - \mathrm{Im}_i\|_2^2 / (D \|\mathrm{Im}_i\|_2^2)$

E.2 Experiments

We compare the proposed method to the Inverse Distance Weighting (IDW) and Radial Basis Function (RBF) based interpolation. These methods are described in [Gentile *et al.* 2013] as the one which performs the best on the considered PSFs interpolation. For the first method, the weights are calculated based on squared distances in the fov ; this is often referred as "Inverse distance-squared interpolator". We used the radial basis function $f(r) = r\ln(r)$; thus the second method is nothing but a thin plate spline interpolation as described in Section D.2. The 300 observed PSFs are first decomposed using a principal components analysis and retaining 40 principal components. This yields a representation with 40 coefficients for each PSF. These coefficients are interpolated components-wise with the two comparison methods and the interpolated PSFs are derived. As the method proposed, the IDW and RBF based interpolation methods are local in the sense that a PSF is interpolated at given location in the fov based on a given number of PSFs observed in the vicinity of the considered location. Thus, the three methods will be applied using different numbers of "neighbors PSFs".

The results are shown in Fig. ***VI.4***, ***VI.5***, ***VI.6***. Interestingly, we observed quite different behaviors with respect to the different criteria. This will be discussed in the next section. However, the less accurate method appears to be the IDW based method. Indeed, it gives the less accurate results in terms of shape ; in particular, it is almost an order of magnitude less accurate than

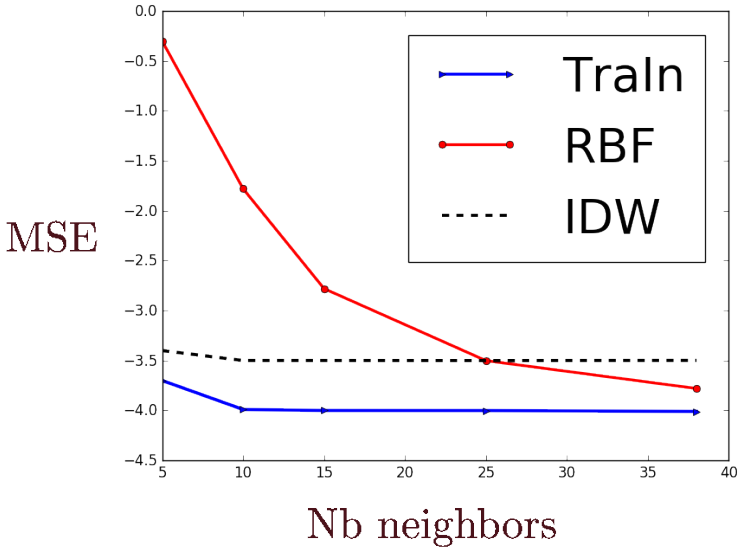


Figure *VI.4* – Normalized mean squared error : Y-axis $\log_{10}$(NMSE), X-axis number of neighbors.

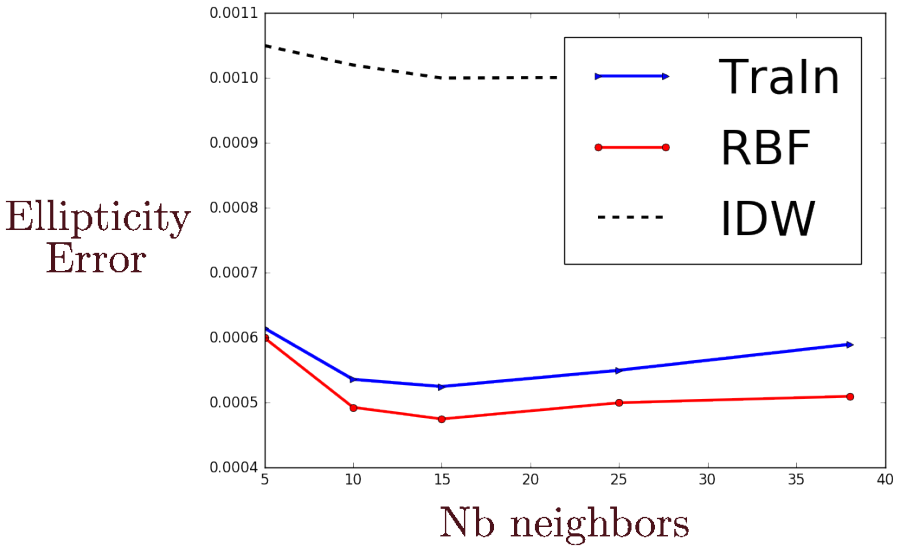


Figure *VI.5* – Average error on the ellipticity : Y-axis E_γ, X-axis number of neighbors.

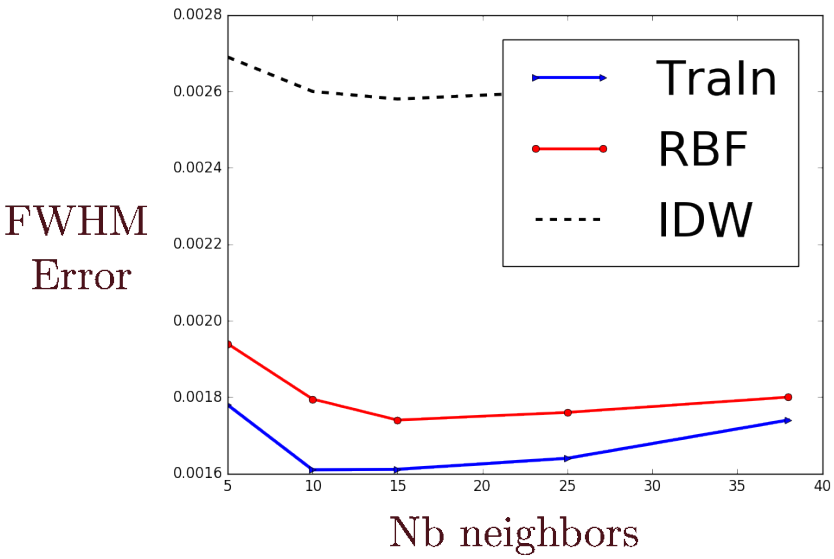


Figure *VI.6* – Average absolute error on the size : Y-axis E_S, X-axis number of neighbors.

the two other methods for the ellipticity. Furthermore, it is almost constantly less accurate than TraIn with respect to the pixels mean square error. The RBF based method and the proposed one gives comparable errors with respect to the shape parameters ; it is slightly more accurate on the ellipticity and slightly less accurate on the shape. The lowest errors with the RBF based method are obtained for a number neighbors around 15 for the shape parameters. The proposed method has a quite stable accuracy with respect to the number of neighbors for all the criteria. Moreover, it is typically several orders of magnitude more accurate than the RBF based one on the pixels mean square errors for a number of neighbors smaller than 15. Therefore, the TraIn method is globally the most accurate. We give some examples of absolute error images i.e. $|\mathrm{Im}_i - \widehat{\mathrm{Im}}_i|$ for some $i \in [\![1, N]\!]$ in Fig. ***VI.7***, which correspond to interpolations with 15 neighbors. It shows indeed that TraIn typically yields a substantially lower residual.

E.3 Discussion

In this Section, we discuss the performances of the tested methods. We consider the q first principal component derived from the learning PSFs set. In the numerical experiments, we set $q = 40$ for the RBF based methods. The Figure ***VI.8*** illustrates how coefficients magnitudes vary across the field for the whole data set (including the test PSFs) for some of the principal compo-

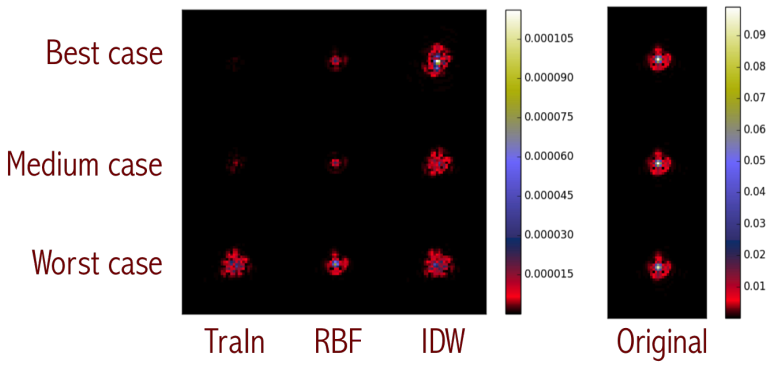


Figure *VI.7* – Absolute error images for 15 neighbors; from the left to the right, TraIn, RBF, IDW, original image; from the top to the bottom, best, average and worst case in terms of the relative accuracy of TraIn with respect to the pixels mean square error.

nents. The two first components exhibit a smooth spatial evolution so that it is possible to interpolate accurately the corresponding surfaces with RBFs using few observations or control points in the vicinity of the interpolation point. This smoothness can be seen in the scatter plot associated with this two first components. Indeed, one needs a smooth warping to transform the scatter plot distribution into the PSFs spatial distribution. On the contrary, the magnitudes have sharper variations for the two other components. This can also be seen from the scatter plot, since a highly non-linear transform would be needed for mapping the scatter plot distribution into the PSFs spatial distribution. In this case, more control points are required for the RBF interpolation to be robust to fast local variations. The scatter plots of Fig. ***VI.8*** are nothing but orthogonal projections of the PSFs manifold over the vector plans spanned by the respective pairs of principal components. As the indexes of the principal components increase, the manifold complexity becomes more apparent and manifests in fast spatial variations of the representation coefficients. Thus, when a few control points are used, the RBF interpolation accumulates errors on the "high indexes" components resulting in a poor accuracy in terms of pixels values. TraIn and IDW instead interpolates globally the PSFs in the sense that the interpolation is not component-wise; this implies less dependency of pixel-wise accuracy on the number of control points and in particular a better accuracy compared to the RBF interpolation where a few control points are used.

However, this does not hold for the ellipticity. To understand these changes of performances, one needs to consider the shape parameters sensitivity relatively to the principal components. Specifically, for each PSF $\mathbf{X}_i$ treated as an

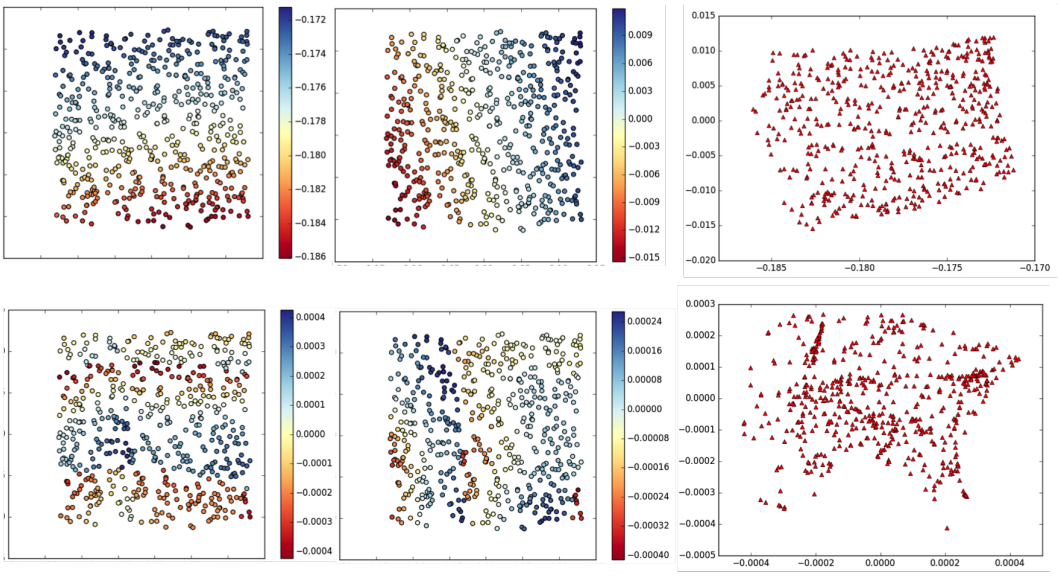


Figure *VI.8* – Spatial distributions of the coefficients of the PSFs used for the experiments relative to 4 principal components calculated over the "learning" PSFs set ; the top panel corresponds to the first and second principal components and the coefficients scatter plot ; the bottom panel corresponds to the 6th and 7th principal components and the coefficients scatter plot as well.

$N_l \times N_c$ matrix, we calculate the derivative of each ellipticity parameter along the lines passing though $\mathbf{X}_i$ and parallel to each of the principal components $\mathbf{P}_j$ also treated as matrices : $\frac{de_k(\mathbf{X}_i+t\mathbf{P}_j)}{dt}$, $k = 1, 2$. The analytic expressions can be found in Appendix H. For each principal component $\mathbf{P}_j$, we denote disp_j the dispersion of the PSFs set projected on the line passing through $\mathbf{0}$ and directed by $\mathbf{P}_j$. We define the sensitivity of the ellipticity component e_k, $k = 1, 2$ with respect to the principal component $\mathbf{P}_j$ as follows :

$$\mathcal{V}(e_k, \mathbf{P}_j) = \frac{\text{disp}_j}{(D+K)} \sum_i |\frac{de_k(\mathbf{X}_i + t\mathbf{P}_j)(0)}{dt}|. \tag{VI.34}$$

It measures how much small perturbations along each of the principal component is susceptible to make the ellipticity vary. In other words, it measures the sensitivity of the ellipticity to errors on the principal components coefficients. The sensitivities for the 40 first principal components are plotted in Fig. ***VI.9*** for the two ellipticity components. It is interesting to see that the plots are not monotonically decreasing. We can draw an important remark from this observation : the most important features in terms of pixel-wise error are not necessarily the most influential in terms of shape. The ellipticity parameters have average magnitudes of order 10^{-2}. Therefore, one can tell from Fig. ***VI.9*** that the ellipticity is significantly sensitive to only a few principal components within the 10 first one. Therefore, as long as the coefficient-wise interpolation is accurate for these low indexes principal components while staying in a reasonable range in general, the final result is accurate in terms of ellipticity. This explains why the RBF interpolation maintains a good accuracy on the ellipticity.

As to the size, it is clearly determined by the brightest structures on the PSF, namely the main lobe and the first brightest ring. The Figure ***VI.10*** shows that the brightest ring information is distributed onto several components, in particular the first and the second one. One expects a bias in the RBF interpolation from processing this information components-wise. On the contrary, the brightest features are globally and more accurately modeled through the barycentric coordinates calculation step, since they influence it the most. Therefore, TraIn gives the most accurate results in terms of size.

F Practical considerations

F.1 Parameters

Local dimensionality We refer to the parameter d involved in the step 7 in Algorithm VI.1. For an ideally dense data set in the sense of the underlying

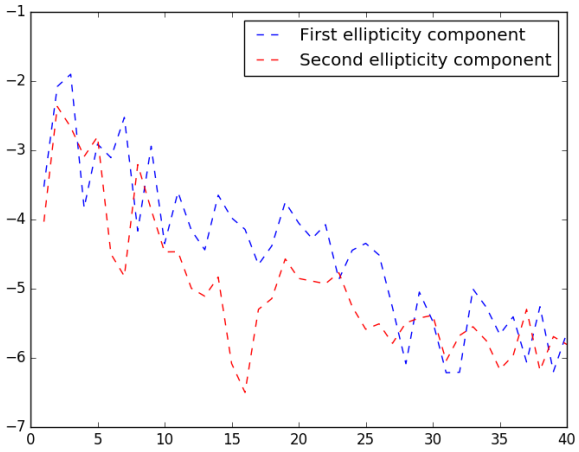


Figure *VI.9* – Ellipticity components sensitivity to the PCA atoms. Y axis : $\log_{10}(\mathcal{V}(e_k, \mathbf{P}_j))$(see Eq.VI.34) ; X axis : PCA atoms indexes j ; moderate errors on the PCA coefficients have a weak impact on the ellipticity, especially for the "high indexes" PCA components.

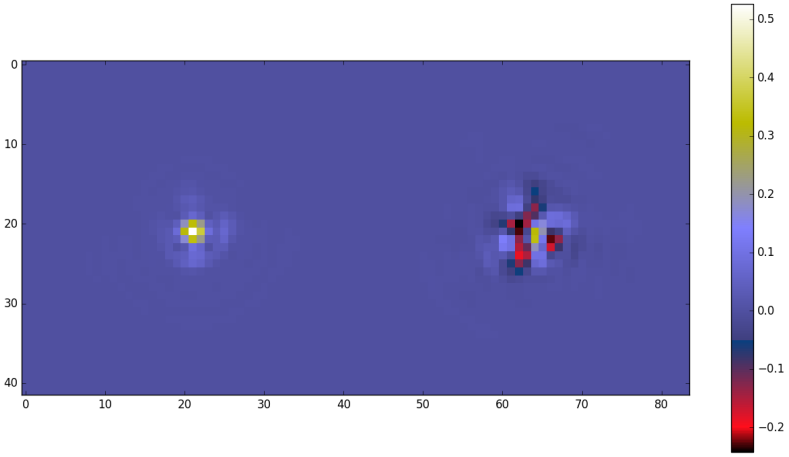


Figure *VI.10* – First and second principal components ; the brightest ring is broken apart between the two components.

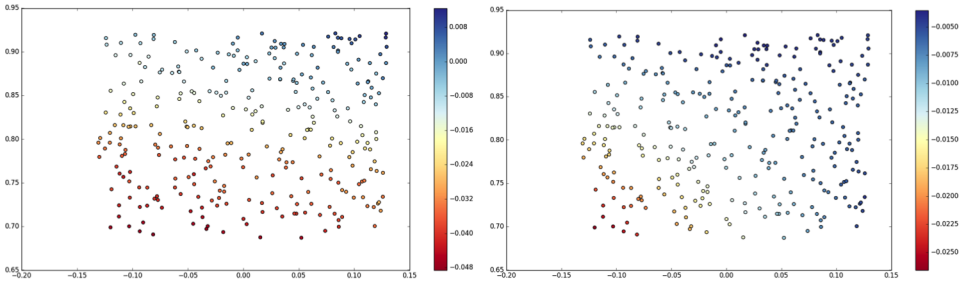


Figure *VI.11* – Spatial distributions of the first (left) and second ellipticity (right) parameters of the observed PSFs.

manifold, this parameter can be set to the manifold intrinsic dimension which is 2 in this example ; however this is tricky since the ideal density depends on the manifold geometric complexity which is a priori not known. We define the "extrinsic dimensionality" d_{ext} of the data set as the dimension of the smaller subspace containing the data set. This can be approximately estimated using a PCA. Then we set d to $\min(p, d_{\text{ext}})$.

Number of neighbors The proposed method gives quite stable results with respect to the number of neighbors. However, Fig. ***VI.11*** suggests a potential improvement of the neighbors selection. For a given position of interpolation, the selection of neighbors can be pictured as follows :

— one initializes a null radius sphere centered on the interpolation position in the fov ;

— one increases the sphere radius until it includes a given number of observations locations.

Yet, Fig. ***VI.11*** shows that the ellipticity parameters changes faster in certain directions than in others. This suggests to grow the neighborhoods according to the ellipticity parameters spatial gradients in such a way to include more neighbors from directions where the ellipticity parameters vary slowly.

Ground metric weight parameter We consider the parameter β in Eq. VI.13. β is the cost of matching two neighbors pixels (assuming a 4-connected neighborhood) with identical intensities. Let consider two images $\mathbf{x} = (x_i)_{1 \leq i \leq N}$ and $\mathbf{y} = (y_i)_{1 \leq i \leq N}$ and the associated point clouds $\mathbf{X}$ and $\mathbf{Y}$ in the sense of Section C.3. If $\forall i \in [\![1, N]\!] |x_i - y_i| \leq \beta$ then $W_2(\mu_{\mathbf{X}}, \mu_{\mathbf{Y}}) = \|\mathbf{x} - \mathbf{y}\|_2$. Thus, β somehow determines how much the Wasserstein geometry is susceptible to depart from the pixels domain euclidean geometry. For sufficiently small neighborhoods around a given location in the fov, the pixels

domain l_2 norm describes accurately the local geometry of the PSFs manifold. Therefore, β has to be chosen so that the Wasserstein metric is equivalent to the pixels domain l_2 norm in any small vicinity in the fov. Hence the following procedure for setting γ :

— determine the two closest PSFs in the fov $\mathbf{x}_i$ and $\mathbf{x}_j$;

— set γ to $\|\mathbf{x}_i - \mathbf{x}_j\|_\infty$.

F.2 Transportation issues

We recall the definition of the sliced Wasserstein distance between two clouds $\mathbf{X}$ and $\mathbf{Y}$:

$$SW_2(\mu_\mathbf{X}, \mu_\mathbf{Y})^2 = \int_{\mathbb{S}^{d-1}} W_2(\mu_{\mathbf{X}_\mathbf{u}}, \mu_{\mathbf{Y}_\mathbf{u}})^2 d\mathbf{u}, \tag{VI.35}$$

where $\mathbb{S}^{d-1} = \{\mathbf{x} \in \mathbb{R}^d / \|\mathbf{x}\| = 1\}$, $\mathbf{X}_\mathbf{u} = \{\mathbf{u}^T\mathbf{x}_i, i = 1...N\} \subset \mathbb{R}^N$ and $\mathbf{Y}_\mathbf{u}$ is similarly defined.

Pairwise distances computation : assignments discrepancies The continuous integration being intractable, in practice, the sliced Wasserstein distance is calculated as a discrete sum :

$$SW_2(\mu_\mathbf{X}, \mu_\mathbf{Y})^2 = \sum_{\mathbf{u}_i \in \Omega \subset \mathbb{S}^{d-1}} W_2(\mu_{\mathbf{X}_{\mathbf{u}_i}}, \mu_{\mathbf{Y}_{\mathbf{u}_i}})^2. \tag{VI.36}$$

For each vector $\mathbf{u}_i$ we recall that the quadratic Wasserstein distance takes the form

$$W_2(\mu_{\mathbf{X}_{\mathbf{u}_i}}, \mu_{\mathbf{Y}_{\mathbf{u}_i}})^2 = \sum_{j=1}^{N} ((\mathbf{u}_i^T\mathbf{X})[j] - (\mathbf{u}_i^T\mathbf{Y})[\sigma^*_{\mathbf{u}_i}(j)])^2, \tag{VI.37}$$

where $\sigma_{\mathbf{u}_i}$ is a permutation of $[\![1, N]\!]$. We use the stochastic gradient descent algorithm proposed in [Rabin *et al.* 2011] for calculating SW_2. The algorithm estimates a local minimum $\mathbf{Y}^*$ of the functional $J_\mathbf{Y}(\mathbf{Z}) = SW_2(\mathbf{Z}, \mathbf{Y})^2$ in the vicinity of $\mathbf{X}$ and the approximated Wasserstein distance is computed as

$$W_2(\mathbf{X}, \mathbf{Y}) \approx \|\mathbf{X} - \mathbf{Y}^*\|_2. \tag{VI.38}$$

At each gradient step, the "sliced assignment" $(\sigma^*_{\mathbf{u}_i})_i$ are updated. To guarantee the convergence, one can use a decreasing step size $\equiv \frac{1}{n^a}$ for $a \in]\!]1/2, 1]\!]$, n being the iteration index [Bottou 1998].

The algorithm succeeds in computing an assignment between the clouds $\mathbf{X}$ and $\mathbf{Y}$ if the slices assignments are identical at convergence and if the stationary point $\mathbf{Y}^*$ represents the same point cloud as $\mathbf{Y}$; these two conditions

are summarized below :

$$SW_2(\mu_{\mathbf{Y}^*}, \mu_{\mathbf{Y}})^2 = \sum_{\mathbf{u}_i \in \Omega \subset \mathbb{S}^{d-1}} \sum_{j=1}^{N} ((\mathbf{u}_i^T \mathbf{Y}^*)[j] - (\mathbf{u}_i^T \mathbf{Y})[\sigma^*_{\mathbf{u}_i}(j)])^2 = 0 \text{ and} \tag{VI.39}$$

$$\text{for } \mathbf{u}_i \neq \mathbf{u}_j,\ \sigma^*_{\mathbf{u}_i} = \sigma^*_{\mathbf{u}_j}. \tag{VI.40}$$

The displacement interpolation between the clouds represented by $\mathbf{X}$ and $\mathbf{Y}$ is realized by performing a linear interpolation between the matrices $\mathbf{X}$ and $\mathbf{Y}^*$. We define the discrepancy support as the following set

$$\mathcal{D} = \{k \in [\![1, N]\!] / \exists (\mathbf{u}_i, \mathbf{u}_j) \in \Omega^2 / \sigma^*_{\mathbf{u}_i}[k] \neq \sigma^*_{\mathbf{u}_j}[k]\} \tag{VI.41}$$

and define the assignment discrepancy as as the support size $|\mathcal{D}|$.

As observed in [Rabin *et al.* 2011] and [Bonneel *et al.* 2015], we do not found it necessary in practice to use a decreasing step size. We observe as well that the stationary points seem to always satisfy $|\mathcal{D}| = 0$. However, it is not rare that the sequence of iterates oscillates around a point which has a non-zero discrepancy support size in which case, the algorithm takes a potentially long time to reach a stationary point satisfying $|\mathcal{D}| = 0$.

In case $|\mathcal{D}| > 0$ for the final iterate $\mathbf{Y}^*$, it mixes information from different pixels in the image associated to $\mathbf{Y}$. This translates into visual artifacts when $\mathbf{Y}^*$ is transformed into an image (see VI.15). One can see such an example in Fig. ***VI.12***. This is a troublesome point since in the considered astronomical application, the systematic errors due to the PSFs estimation constitutes one of the bottlenecks.

Besides, the algorithm is stochastic and energy minimized is non-convex ; therefore for the same input point clouds, the algorithm converges to a different point at each run. We propose to reduce the final assignment discrepancy with a fixed number of iterations and the stationary point variability by improving the gradient descent initialization. Taking the two point clouds $\mathbf{X}$ and $\mathbf{Y}$, we extract from each cloud the points corresponding to the pixels comprised in a small rectangular window (typically of size 20×20) around the corresponding images centroids (see Fig. ***VI.13***). Let denote $\mathbf{ind}_X$ and $\mathbf{ind}_Y$ the sets of corresponding indexes and $\overline{\mathbf{ind}}_X$ the complement of $\mathbf{ind}_X$ in $[\![1, N]\!]$. Then we find the optimal assignment between $\mathbf{X}[\mathbf{ind}_X]$ and $\mathbf{Y}[\mathbf{ind}_Y]$ using the Hungarian algorithm [Kuhn 1955], which is that fastest known procedure for solving exactly the assignment problem. However, it has a roughly cubic complexity in its improved version, which restricts its practical use to small scale problems. We denote σ the optimal mapping :

$$\mathbf{X}[\mathbf{ind}_X[i]] \mapsto \mathbf{Y}[\mathbf{ind}_Y[\sigma(i)]], \forall i \in [\![1, |\mathbf{ind}_X|]\!]. \tag{VI.42}$$

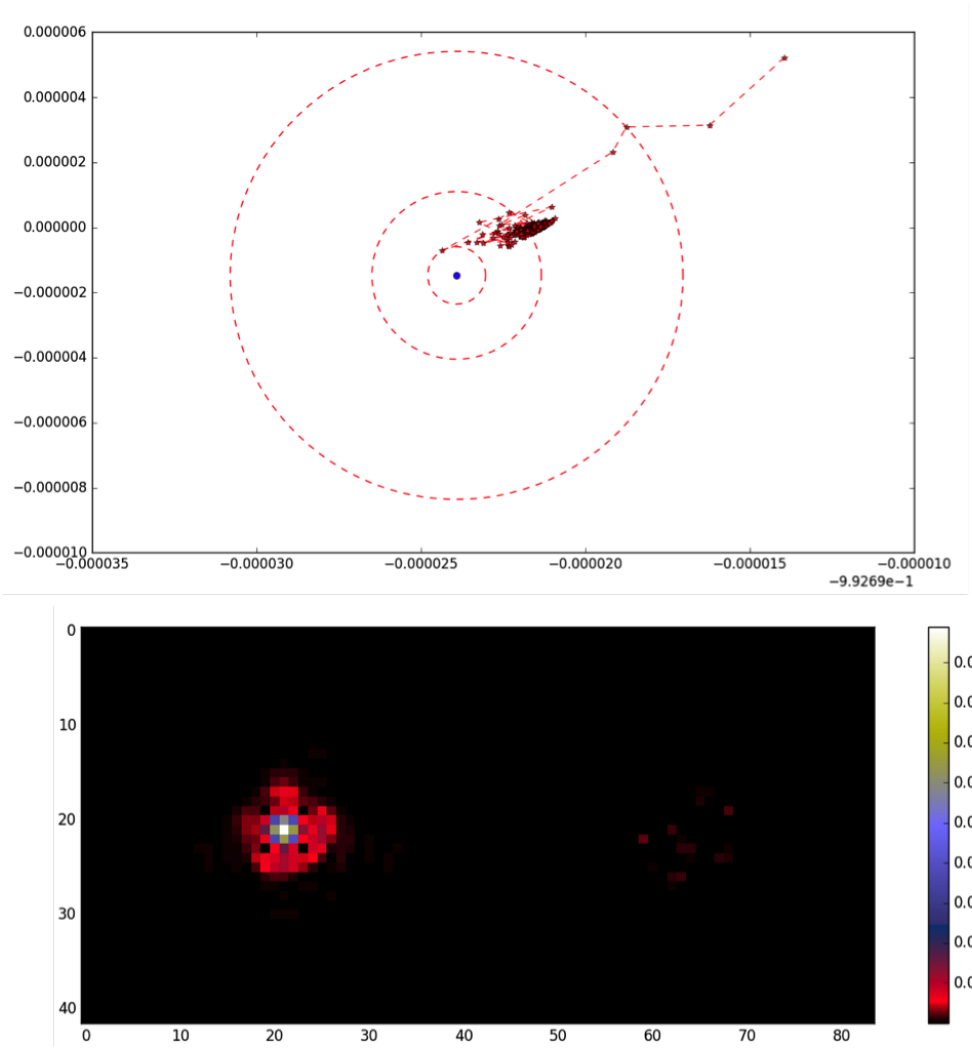


Figure *VI.12* – Assignment failure : on the top plot, the sequence of images associated with each iterate are represented by the red dots and the image associated with the ideal solution is represented by the blue dot ; these images are shown in the plan spanned by the two first principal components calculated from the whole images set ; the iterates end up oscillating around a point in the middle circle ; the bottom left plot shows the ideal image and the bottom right plot shows the absolute difference between the final image and the ideal one ; one can see that there is some non-negligible artifacts.

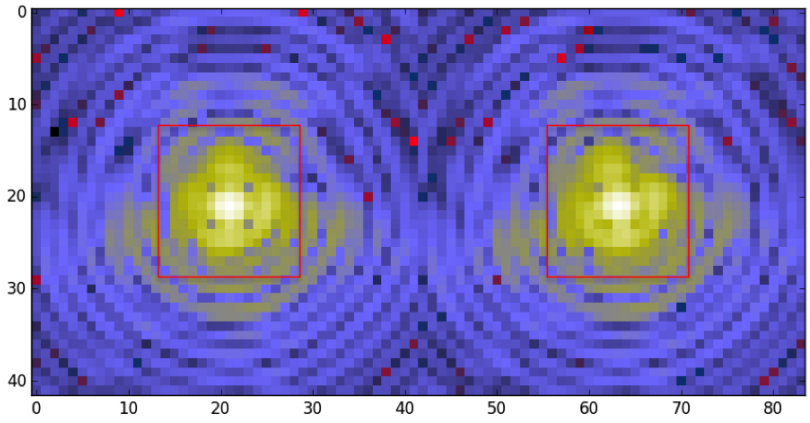


Figure *VI.13* – Two PSFs examples in a logarithmic scale ; we first use the hungarian algorithm to map the two point clouds derived from the pixels located in the small red rectangles in each PSF respectively ; the result is used for initializing the sliced transport algorithm on the point clouds associated with the full images (see VI.12).

Then we initialize the stochastic gradient descent method aforementioned with the matrix $\mathbf{X}_0$ defined as follows :

$$\begin{cases} \mathbf{X}_0[:, i] = \mathbf{X}[:, i] \text{ if } i \in \overline{\mathbf{ind}}_X \text{ and} \\ \mathbf{X}_0[:, \mathbf{ind}_X[i]] = \mathbf{Y}[\mathbf{ind}_Y[\sigma(i)]] \text{ for } i \in [\![1, |\mathbf{ind}_X|]\!]. \end{cases} \tag{VI.43}$$

To quantify the impact of this pre-assignment, we compare the average time of execution of the algorithm, for 3 different pair of images, for 100 runs in each case, with and without optimizing the initialization. These times account for the hungarian algorithm when used. We set a maximum number of iterations of 20000. The discrepancies support size $|\mathcal{D}|$ might be non-zero when this number of iterations is reached. Therefore, we also compute the average of the final discrepancies support sizes. The number of projection directions was set to 30 and the initialization window size to 20×20. As for all the experiments presented, we used 42×42 PSFs images, which gives clouds of 1724 points each one to compare. The result is displayed in Fig. ***VI.14***.

This results follows from the simple observation that the higher is $|\mathcal{D}|$ for a given iterate, the more the stochastic gradient descent is oscillatory. The Figure ***VI.14*** shows that it is possible to set the exact matching window size so that the Hungarian algorithm brings more benefits in terms of discrepancies reduction than computational complexity.

It is worth noting that the assignment obtained from the sliced Wasserstein transport depends on the number of directions and slices. As outlined

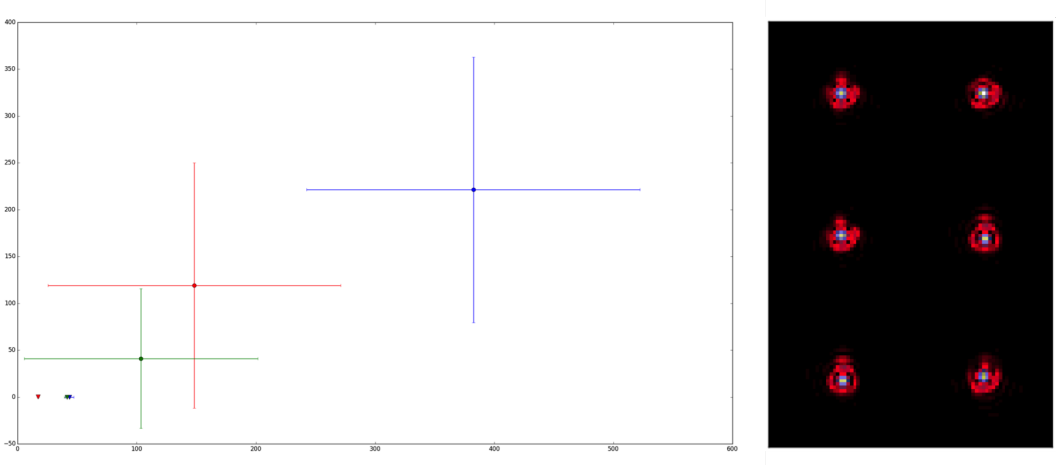


Figure *VI.14* – Sliced Wasserstein transport algorithm average performances ; x axis : average execution times in seconds ; y axis : average final discrepancies support size ; the red, green and yellow corresponds to the pairs of images from the top to the bottom respectively ; the triangle-shaped points corresponds the gradient descents with an optimized initialization ; the points are obtained by averaging these performances criteria over 100 runs for the 3 pairs of 42×42 images ; the proposed initialization dramatically reduces the time needed to compute an assignment between our images related point clouds with an average execution time ranging from 10 to 40s rather than 100 to 400s.

in [Bonneel *et al.* 2015], the more directions, the smoother is the assignment. Unsurprisingly, this dependency is reduced when involving the Hungarian algorithm. Actually in that case, we obtain exactly the same assignment independently of the number of directions, for two given PSFs images and an initialization window of size 20×20.

Velocity constrained displacement interpolation The Wasserstein barycenter of two point clouds is simply calculated by realizing a displacement interpolation (see Section C.2). Considering 2 PSFs, a simplified interpolation process in this framework would go as follows :

— the PSFs are first converted into two matrices $\mathbf{X}$ and $\mathbf{Y}$ representing two different point clouds ;

— the point clouds are matched according to the previous paragraph ; this yields a matrix $\mathbf{Y}^*$ representing the same point cloud as $\mathbf{Y}$;

— the point clouds interpolation is realized by moving each point in the first point cloud toward the matched point in the second point cloud along a straight line and at a uniform speed on the cloud ; this is done by performing a linear interpolation of the matrices $\mathbf{X}$ and $\mathbf{Y}^*$;

— the interpolated PSFs are obtained by converting interpolated clouds at different "times" into images.

An example is shown in Fig. ***VI.15***. In these example, the brightest PSF's ring is more elongated horizontally for the initial PSF and vertically for the final PSF. But as the displacement interpolation constraints particles to move along straight lines, the interpolated point clouds are shrunken compared to the extreme ones and consequently, the interpolated images have a more concentrated energy. Quantitatively, the interpolated PSFs tend to have l_2/l_1 norms ratio that significantly exceed the range observed over the data set - note that the PSFs l_1 norms are constant and equal to 1. This problem can be approached from the ground metric angle ; precisely, one can think of choosing a ground metric that admits curved geodesics, making a more "compact" particles advection possible. However, the complex variation of the PSFs across the fov makes the choice of a better fitted ground metric non trivial. The displacement interpolation generates a sequence of matrices $(\mathbf{X}_i)_{1\leq i\leq T}$ so that $\mathbf{X}_1 = \mathbf{X}$, $\mathbf{X}_T = \mathbf{Y}^*$ and for $i \in [\![1, T-1]\!]$, $\mathbf{X}_{i+1} - \mathbf{X}_i = \mu_i(\mathbf{Y}^* - \mathbf{X})$ for a step μ_i. In order to better preserve the point clouds densities in the displacement interpolation, we impose the points to have parallel velocity vectors at each time. To do so, we modify the displacement interpolation as follows :

1. INPUT : a maximum step size $\mu_{\max}$
2. Initialization : $\mathbf{X}_0 = \mathbf{X}, i = 0$.

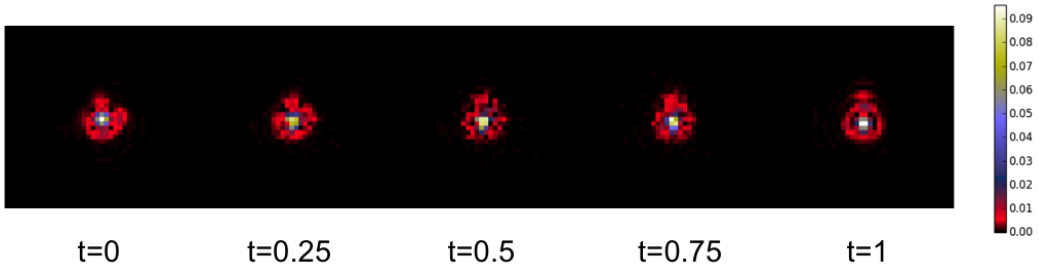


Figure *VI.15* – Displacement interpolation : "shrinkage effect" ; the interpolated PSFs have a narrower energy distribution than the initial and final ones.

3. WHILE $i \leq i_{\max}$:
   $\mathbf{V}_i = \mathbf{Y}^* - \mathbf{X}_i$
   Compute an eigenvector $\mathbf{u}_i$ corresponding to the highest eigenvalue of $\mathbf{V}_i\mathbf{V}_i^T$
   $\widehat{\mathbf{V}}_i = \mathbf{u}_i\mathbf{u}_i^T\mathbf{V}_i / \|\mathbf{u}_i\|_2^2$
   Line search : $\mu_{\text{opt}} = \underset{\mu}{\text{argmin}} \ \|\mathbf{X}_i + \mu\widehat{\mathbf{V}}_i - \mathbf{Y}^*\|_2^2 = -\frac{\langle \widehat{\mathbf{V}}_i, \mathbf{X}_i - \mathbf{Y}^* \rangle}{\|\widehat{\mathbf{V}}_i\|_2^2}$
   $\mathbf{X}_{i+1} = \mathbf{X}_i + \min(\mu_{\text{opt}}, \mu_{\max})\widehat{\mathbf{V}}_i$
   $i = i + 1$.

In the above procedure, all the particles are moved in parallel with $\mathbf{u}_i$ at each iteration. Moreover, $\|\mathbf{X}_{i+1} - \mathbf{Y}^*\|_2^2 \leq \|\mathbf{X}_i - \mathbf{Y}^*\|_2^2$. Therefore the procedure moves in fact the initial point cloud toward the final one. We do not discuss how close the sequence generated can come to $\mathbf{Y}^*$. However, in the tests we performed, the distance $\|\mathbf{X}_i - \mathbf{Y}^*\|$ always gets down to the machine numerical precision. In Fig. ***VI.15***, the time parameter is defined as

$$t(i) = 1 - \frac{\|\mathbf{X}_i - \mathbf{Y}^*\|_2}{\|\mathbf{X} - \mathbf{Y}^*\|_2}, \tag{VI.44}$$

which makes sense in a dynamic model of constant speed advection along straight lines. However, in the modified displacement interpolation, the matrices $(\mathbf{X}_i)_{1 \leq i \leq i_{\max}}$ does not necessarily follow a linear path in the matrix space. To make a direct comparison possible between the two displacement interpolations, the time parameter definition has to be generalized in order to account for a possible curvature of the path. We do so by involving the length of the curve obtained by joining each matrix of the sequence to the following one with a straight line :

$$t(i) = 1 - \frac{\sum_{j=i}^{i_{\max}} \|\mathbf{X}_i - \mathbf{X}_{i+1}\|_2}{\sum_{j=0}^{i_{\max}} \|\mathbf{X}_i - \mathbf{X}_{i+1}\|_2}, \tag{VI.45}$$

with the convention $\mathbf{X}_{i_{\max}+1} = \mathbf{Y}^*$. This definition generalizes VI.44. We compare the two displacement interpolations in Fig. ***VI.16***. The parallel velocity constraint quantitatively reduces the shrinkage effect since the interpolated images l_2/l_1 norms ratios does not exceed those of the reference images as much as with the regular displacement interpolation. Not surprisingly, this constraint yields smoother barycenters.

This constraint can be integrated directly into the sliced OT algorithm. However, we observe that this considerably increases the number of iterations needed to converge and most importantly, the generated sequence mostly converges to a solution which is not a global minimum of the functional $J_{\mathbf{Y}}(\mathbf{Z}) = SW_2(\mathbf{Z}, \mathbf{Y})^2$. We note that this constraint can be naturally integrated into the framework adopted in [Papadakis *et al.* 2013] since velocities are directly manipulated. However, the eulerian discretization would require manipulating 4 dimensional arrays, making the approach intractable.

This constraint has not been used for generating the plots shown in Section E. Indeed, for that experiment, we considered a density of known PSFs that makes the PSFs variations in a given neighborhood in the fov very smooth in which case the shrinkage effect disappears.

Wasserstein barycenters computation As explained in Section D, the interpolated PSF's computation involves calculating a sliced Wasserstein barycenter. Let consider a set of point clouds $(\mathbf{Y}_i)_{1\leq i\leq p}$. As for the pairwise Wasserstein distances, an approximation of their Wasserstein barycenter is approximately calculated by finding a local stationary point of the functional $J_{\mathbf{Y}_1,\dots,\mathbf{Y}_p}(\mathbf{Z}) = \sum_{i=1}^{p} w_i SW_2(\mathbf{Z}, \mathbf{Y}_i)^2$ [Rabin *et al.* 2011], for some barycentric weights $(w_i)_{1\leq i\leq p}$. The higher is p, the more degenerated is this functional because of its combinatorial nature, which increases the assignments discrepancies. But as mentioned in Section C.2.2, there is no efficient method for calculating exactly a Wasserstein barycenter of more than two point clouds, even for small scale problems. Therefore, the previously described strategy for speeding the convergence up can not be directly extended to this case. We propose to approximate the Wasserstein barycenter of a set of more than two point clouds by computing a sequence of "2 points" Wasserstein barycenters. This relies on the local isometry assumption made in Section D.1. Indeed, the barycenter of a set of vectors $(\mathbf{r}_i)_{1\leq i\leq p}$ associated with the barycentric weights $(w_i)_{1\leq i\leq p}$ can be calculated through the following procedure :

1. Initialization : $\mathbf{r}_{\text{bar}} = \mathbf{r}_1;\ i = 2,\ w = w_1$
2. WHILE $i \leq p$:
   $\mathbf{r}_{\text{bar}} = \underset{\mathbf{r}}{\text{argmin}}\ \ w_i\|\mathbf{r} - \mathbf{r}_i\|_2^2 + w\|\mathbf{r} - \mathbf{r}_{\text{bar}}\|_2^2$

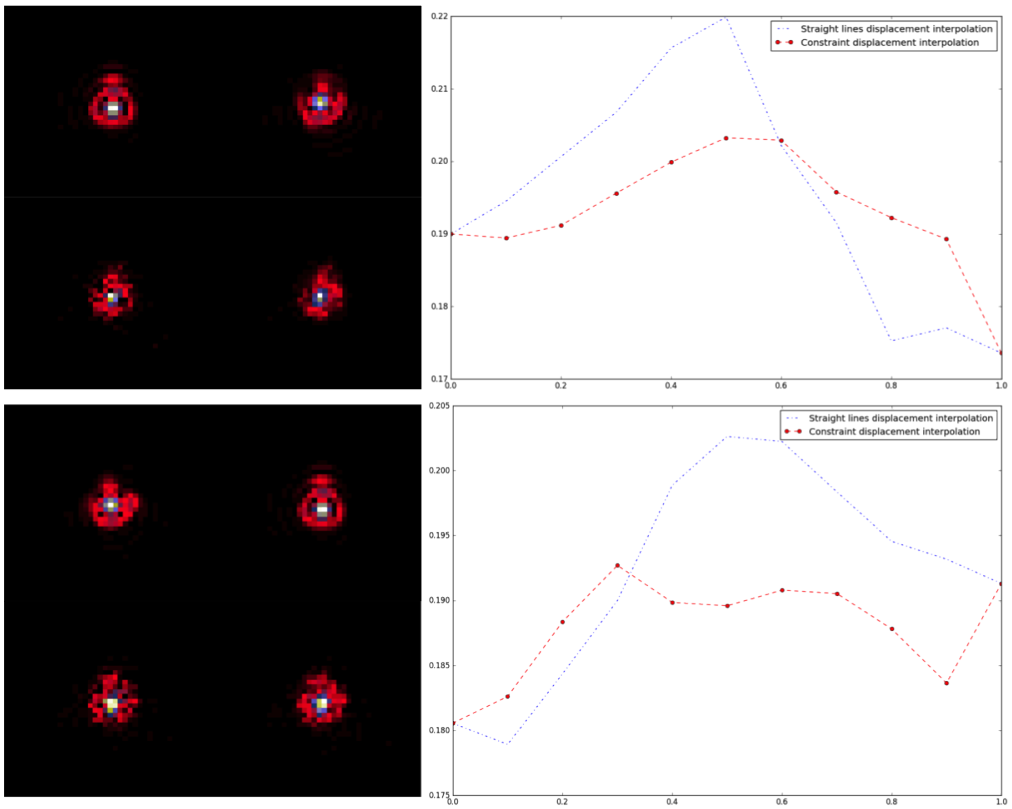


Figure *VI.16* – Constrained displacement interpolation : the two PSFs panels show original PSFs on the top and the interpolated PSFs for t=0.5 are shown on the bottom left for the regular displacement interpolation and the bottom right for the velocity constrained displacement interpolation ; the right plots show the evolution of the l_2/l_1 norms ratio of interpolated PSFs as functions of the time parameter in the two examples respectively ; the velocity constraint reduces the shrinkage effect since the corresponding l_2/l_1 norms ratios are flattened compared to those obtained with the regular displacement interpolation ; moreover, the interpolated images are smoother.

$w = w + w_i$
$i = i + 1.$

Indeed, one can check that at the i^{th} iteration $\mathbf{r}_{\text{bar}}$ is updated to $\sum_{k=1}^{i} \frac{w_k}{\sum_{j=1}^{i} w_j} \mathbf{r}_k$. If the Wasserstein space is isometric to a euclidean space in the neighborhood of the point clouds involved in the barycenter's calculation, then one can apply a similar procedure in the Wasserstein space to calculate the barycenter, hence the following scheme :

1. Initialization : $\mathbf{Z}_{\text{bar}} = \mathbf{Z}_1;\ i = 2,\ w = w_1$
2. WHILE $i \leq p$:
   $\mathbf{Z}_{\text{bar}} = \underset{\mathbf{Z}}{\text{argmin}}\ \ w_i SW_2(\mathbf{Z}, \mathbf{Y}_i)^2 + wSW_2(\mathbf{Z}, \mathbf{Z}_{\text{bar}})^2$
   $w = w + w_i$
   $i = i + 1.$

This way we can take advantage of the previously accurate initializing proposed in the Wasserstein barycenter approximation and make use of the $1D$ displacement interpolation. The computed barycenter slightly changes depending on the ordering of the clouds $\mathbf{Y}_i$. We choose the ordering so that $w_1 \geq \cdots \geq w_p$.

G Reproducible research

In the spirit of participating in reproducible research, the data and the codes used to generate the plots presented in this paper will be made available at `http://www.cosmostat.org/software/`.

H Conclusion

We introduce TraIn (Transport Interpolation) which is a data field interpolation method based on Optimal Transport and making use of some Manifold Learning ideas. We consider the interpolation of a PSF field. The local geometry of the PSF field is characterized using approximated Wasserstein distances. From these, we derive low dimensional local euclidean embeddings of the PSF field which is then mapped to the instrument field-of-view using a thin-plate interpolation. This mapping gives one the embedded coordinates of the unknown PSFs, from which a representation of these PSFs in the quadratic Wasserstein metric is determined. Finally, the interpolated PSFs are calculated through nested displacement interpolations.

We compared TraIn to the Inverse Distance Weighting method and to a component-wise thin-plate interpolation of PCA coefficients. The tests were

made on a set of realistic Euclid-like PSFs. We show that the proposed method is globally the most accurate in terms of pixel domain error and shape. In particular, TraIn is in average several order of magnitudes more accurate than the two other methods in terms of pixels errors when a few number of "neighbors PSFs" are used for the interpolation ; it is therefore more robust to the PSF field sampling. We also introduce a velocity constraint displacement interpolation for mitigating the unnatural shape shrinkage that might occur when interpolating objects with different major axis using an euclidean ground metric.

A natural extension of this study would be to compare these interpolation methods in a setting where the reference PSFs are not perfectly known.

Chapitre VII

Conclusion

A Methods and results

This thesis addressed the problem of estimating the Point Spread Function across an optical instrument field-of-view (fov) solely from unresolved images of compact objects such as stars.

In Chapter IV, we proposed a method for estimating a well sampled PSF from aliased, randomly shifted and noisy measurements of the same PSF. We show that using sparsity and positivity constraints yield a significantly better preservation of the PSF shape over popular methods in the astronomy community, especially at low SNRs.

In Chapter V, we move toward a more realistic setting. Indeed, we consider the problem of jointly estimating a set of different PSFs from only one measurement for each PSF. Thus, we introduce a dimension reduction and super-resolution method which, additionally to the priors used in the multiple frame setting, takes advantage of the PSFs field compressibility. This, again, yields significant improvements in terms of the PSFs shapes at low SNR, with respect to existing methods.

In Chapter VI, we consider the problem of estimating the PSF at an arbitrary location in the instrument fov, given the PSFs on a finite set of locations. This amounts to estimating the mapping between the fov coordinates space and the PSFs underlying manifold. To that end, we propose an interpolation framework which firstly derives local low dimensional representations of the training data based on approximated pairwise quadratic Wasserstein distances. Assuming that this unfolds the PSFs manifold locally, the low dimensional representations of the unknown PSFs can be estimated using standard interpolation methods. From these, the barycentric coordinates of the unknown PSFs in the Wasserstein metric are derived and the interpolated PSFs are calculated as (approximated) Wasserstein barycenters. This approach yielded a remarkable accuracy in terms of the PSFs shapes and pixels errors. However, it is computationally demanding.

We emphasize that the methods developed are non parametric and can be used for other optical instruments, although the numerical experiments were focused on the Euclid telescope PSFs. The associated codes will be made available at `http://www.cosmostat.org/software/`.

B Perspectives

Several perspectives follow naturally from this work and applications in astronomy and beyond.

In astronomy : The interpolation problem can be alternately addressed as an extension of RCA (see Chapter V) by solving a larger inverse problem involving a "semi-blind" deconvolution of the galaxies present in the same fov. We recall that RCA computes a matrix $\mathbf{S}$ which columns span a subspace embedding the PSFs manifold and a matrix $\mathbf{A}$ of representation coefficients of the PSFs at the positions of the unresolved stars images present in the fov. The ultimate aim is to estimate the PSFs at the galaxies locations. Assuming that these PSFs also belong the $\mathbf{S}$'s columns space, one only need to determine their representations coefficients, which we denote by a matrix $\mathbf{A}_g$. Given an estimate of $\mathbf{A}$, $\mathbf{A}_g$ is strongly constrained, due to the PSFs field's regularity and the PSFs positivity prior. Given an estimate of $\mathbf{A}_g$ and $\mathbf{S}$, one can restore the galaxies. Finally, given an estimate of the galaxies, one can refine $\mathbf{A}_g$. Thus, it is easy to derive a global scheme which both restores and interpolates the PSFs and deconvolves the galaxies.

More generally : The methods proposed have to be extended to account for the wavelength dependency of the PSFs. Specifically, an unresolved image of a compact light source can be modeled as

$$\mathbf{y} = \sum_{i=1}^{n} \mathcal{F}(a_i \mathbf{x}_i) + \mathbf{n}, \tag{VII.1}$$

where each $\mathbf{x}_i$ is the PSF at the position of object in a narrow wavelength band centered on a wavelength λ_i. The vector $\mathbf{a} = (a_i)_{1 \leq i \leq n}$ is the discretized emission spectrum of the light source. n has to be sufficiently large for the PSF to be approximately constant on each band. $\mathbf{n}$ is the noise and $\mathcal{F}$ is a linear degradation operator. Let assume that $\mathcal{F}$ is the identity. Calculating the monochromatic PSFs $\mathbf{x}_i$ from $\mathbf{y}$ is an ill-posed inverse problem, even knowing $\mathbf{a}$. However, this problem is highly structured. Indeed, assuming that $\lambda_1 < \cdots < \lambda_n$, then for $i > 1$, the PSF $\mathbf{x}_i$ is approximately obtained by rescaling the PSF $\mathbf{x}_{i-1}$ as illustrated in Fig. ***VII.1***. This follows from the dilatation property of the Fourier transform ; we refer to the introductory chapter for the analytical form of a monochromatic PSF. It is not exactly a rescaling because of the wavelength dependency of the wavefront error ; again, we refer the reader to Chapter I. However, we can reasonably assume that in a continuous setting, the PSF $\mathbf{x}_{i-1}$ can be transformed into the PSF $\mathbf{x}_i$ through a

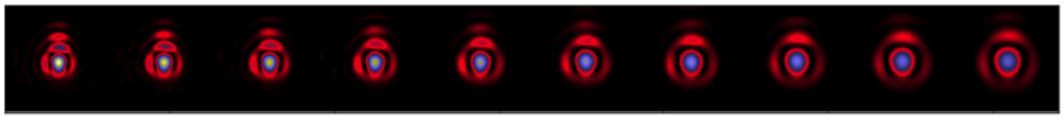


Figure *VII.1* – Simulated Euclid optical PSFs at increasing wavelengths from the left to the right between 500 and 900 nm

monotonic rearrangement. Thus, rather than computing directly the $\mathbf{x}_i$s, one can approach the problem by computing a sparse and positive transportation plan between $\mathbf{x}_1$ and $\mathbf{x}_n$, which generates the intermediary PSFs by displacement interpolation and approximately performs a monotonic rearrangement of $\mathbf{x}_1$ into $\mathbf{x}_n$. One can think of different tools to enforce the monotonicity, among which the helmoltz-hodge decomposition, in the same spirit as [Zhu *et al.* 2007]. Conveniently, this approach would implicitly integrates the mass conservation constraint between the PSFs.
Furthermore, as seen in the introductory chapter, more physical priors could be involved in the PSF's restoration. A simple example is the fact that a PSF's Fourier transform is an even function, which can be translated into a linear constraint. A bolder approach would be to turn the PSF restoration into a phase recovery problem, and use prior knowledge on the phase-shifting function, like for instance its relationship with Zernike polynomials. This is particularly sound if out-of-focus images are available, which can be easily obtained in microscopy applications [Bostan *et al.* 2016].

Annexe A

Academic activities

Journal papers

— F. M. Ngolè Mboula, J.-L. Starck, S. Ronayette,K. Okomura, J. Amiaux. *Super-resolution method using sparse regularization for point spread function recovery*, Astronomy and Astrophysics, **575** : A86, 2015.

— F. M. Ngolè Mboula, J.-L. Starck, K. Okomura, J. Amiaux, P. Hudelot. *Constraint matrix factorization for space variant PSFs field restoration*, Inverse Problems, **32** : 12, 2016.

— F. M. Ngolè Mboula, J.-L. Starck. *PSFs field learning based on Optimal Transport distances*, Submitted.

— S. Farrens, F. M. Ngolè Mboula, J.-L. Starck. *Space variant deconvolution of galaxies survey images*, Astronomy and Astrophysics.

— Great3 Challenge participants. *GREAT3 results–I. Systematic errors in shear estimation and the impact of real galaxy morphology*, Monthly Notices of the Royal Astronomical Society, **450**(3) : 2963-3007, 2015.

Conferences and workshops

— International Traveling Workshop on Interactions between Sparse models and Technology, Namur, Belgium, August 2014 (poster).

— International Workshop on Cosmology and Sparsity, Nice, September 2014 (talk).

— RICAM Special Semester on New Trends in Calculus of Variations : Optimal Transport in the Applied Sciences, Linz, December 2014 (talk).

— Data Sience in Particle Physics, Astrophysics and Cosmology workshop, Paris, April 2015 (talk).

— Missing data in physics, Nice, May 2015 (talk).

— Euclid weak lensing meeting, London, November 2015 (talk).

— International Traveling Workshop on Interactions between Sparse models and Technology, Aalborg, Denmark, August 2016 (talk).

Teaching

— Scientific activity tutor in primary schools for the Center of Research and Inter-disciplinarity (CRI), 2013-2015.

— Teaching assistant at ENSAE ParisTech in Numerical Analysis, 2015-2016.

Miscellaneous

— 2nd price for the "Concours découvrir l'Univers" of the SF2A with the CRI, May 2014.

Annexe B

Convex analysis

Sommaire

In this appendix, we give the general convex analysis material relevant to our work. We consider a finite-dimensional Hilbert space $\mathcal{H}$ equipped with the inner product $\langle .,. \rangle$ and associated with the norm $\|.\|$. Let $\mathcal{H}$ be a finite-dimensional Hilbert space (typically a real vector space) equipped with the inner product $\langle .,. \rangle$ and associated with the norm $\|.\|$. A real-valued function $\mathcal{F}$ defined on $\mathcal{H}$ is

— proper if its domain, as defined by $\operatorname{dom} \mathcal{F} = \{\mathbf{x} \in \mathcal{H} / \mathcal{F}(x) < +\infty\}$, is non-empty ;

— lower semicontinuous (LSC) if $\liminf_{\mathbf{x} \to \mathbf{x}_0} \mathcal{F}(\mathbf{x}) \geq \mathcal{F}(\mathbf{x}_0)$.

We define $\Gamma_0(\mathcal{H})$ as the class of all proper LSC convex real-valued function defined on $\mathcal{H}$.

A Proximal calculus

Moreau (1962) introduced the notion of proximity operator as a generalization of a convex projection operator. Let $\mathcal{F} \in \Gamma_0(\mathcal{H})$. Then the function $\mathbf{y} \to \frac{1}{2}\|\boldsymbol{\alpha} - \mathbf{y}\|^2 + \mathcal{F}(\mathbf{y})$ achieves its minimum at a unique point denoted by $\operatorname{prox}_{\mathcal{F}}(\boldsymbol{\alpha})$, $(\forall \boldsymbol{\alpha} \in \mathcal{H})$. The operator $\operatorname{prox}_{\mathcal{F}}$ is the proximity operator of $\mathcal{F}$.

Examples :

— The indicator function of a closed convex subset $\mathcal{C}$ of $\mathcal{H}$ is the function defined on $\mathcal{H}$ by

$$\iota_{\mathcal{C}}(\mathbf{x}) = \begin{cases} 0, & \text{if } \mathbf{x} \in \mathcal{C} \\ +\infty, & \text{otherwise.} \end{cases} \tag{B.1}$$

It is clear from the definitions that the proximity operator of $\iota_{\mathcal{C}}$ is the orthogonal projector onto $\mathcal{C}$. Thus, for $\mathcal{C} = \mathcal{S}_{\mathbf{x}^{(0)}}$, which is defined in

Chapter IV Section B.1, and for $\mathbf{x} \in \mathbb{R}^{d^2p^2}$, we have,

$$\text{prox}_{\iota_{\mathcal{S}_{\mathbf{x}^{(0)}}}}(\mathbf{x}) = (\max(x_i, -x_i^{(0)}))_{1 \leq i \leq d^2p^2}. \tag{B.2}$$

— Another popular example of proximity operator is the one associated with $\mathcal{F} = \lambda\|.\|_1$ with $\lambda \in \mathbb{R}$:

$$\text{prox}_{\lambda\|.\|_1}(\boldsymbol{\alpha}) = \text{SoftThresh}_\lambda(\boldsymbol{\alpha}) = \left(\left(1 - \frac{\lambda}{|\alpha_i|}\right)_+ \alpha_i\right)_{1 \leq i \leq p}, \tag{B.3}$$

where p is the dimension of $\mathcal{H}$, α_i the components of alpha in the basis associated with $\|.\|_1$ and $(.)_+ = \max(., 0)$.

— For $\boldsymbol{\lambda} = (\lambda_i)_{1 \leq i \leq p}$, the proximity operator associated with the weighted l_1 norm $\|\operatorname{Diag}(\boldsymbol{\lambda})(.)\|_1$ is given by

$$\text{prox}_{\|\operatorname{Diag}(\boldsymbol{\lambda})(.)\|_1}(\boldsymbol{\alpha}) = \text{SoftThresh}_{\boldsymbol{\lambda}}(\boldsymbol{\alpha}) = \left(\left(1 - \frac{\lambda_i}{|\alpha_i|}\right)_+ \alpha_i\right)_{1 \leq i \leq p}. \tag{B.4}$$

The hard thresholding operator defined as

$$\text{HardThresh}_{\boldsymbol{\lambda}}(\boldsymbol{\alpha}) = (\tilde{\alpha}_i)_{1 \leq i \leq p}, \quad \text{with } \tilde{\alpha}_i = \begin{cases} \alpha_i & \text{if } |\alpha_i| \geq \lambda_i \\ 0 & \text{else.} \end{cases} \tag{B.5}$$

is often used instead in practice.

Properties :

— separability : if $\mathcal{H} = \mathcal{H}_1 \times \cdots \times \mathcal{H}_n$, for $\mathcal{F} \in \Gamma_0(\mathcal{H})$ and if $\mathcal{F}(\mathbf{x}) = \mathcal{F}_1(\mathbf{x}[1]) + \cdots + \mathcal{F}_n(\mathbf{x}[n])$ where $\mathcal{F}_i \in \Gamma_0(\mathcal{H}_i)$, for $i = 1 \ldots n$, then $\text{prox}_{\mathcal{F}}(\mathbf{y}) = (\text{prox}_{\mathcal{F}_1}(\mathbf{y}[1]), \ldots, \text{prox}_{\mathcal{F}_n}(\mathbf{y}[n]))$;

— translation : for $\mathcal{F} \in \Gamma_0(\mathcal{H})$ and $\mathbf{a} \in \mathcal{H}$, we define $\mathcal{F}_{\mathbf{a}}(\mathbf{x}) = \mathcal{F}(\mathbf{x} - \mathbf{a})$; then $\text{prox}_{\mathcal{F}_{\mathbf{a}}}(\mathbf{y}) = \mathbf{a} + \text{prox}_{\mathcal{F}}(\mathbf{y} - \mathbf{a})$

—

Now we suppose that $\mathcal{H} = \mathbb{R}^p$, and we want to solve

$$\min_{\boldsymbol{\alpha} \in \mathbb{R}^p} \mathcal{F}_1(\boldsymbol{\alpha}) + \mathcal{F}_2(\boldsymbol{\alpha}) \tag{B.6}$$

where $\mathcal{F}_1, \mathcal{F}_2 \in \Gamma_0(\mathbb{R}^p)$. Many problems in signal and image processing may be formulated this way, where $\mathcal{F}_1$ would be the data attachment function

and $\mathcal{F}_2$ would constrain this solution based on prior knowledges. It has been shown [Combettes & Wadjs 2005] that if $\mathcal{F}_1$ is differentiable with a β-Lipschitz continuous gradient, then the problem (B.6) admits at least one solution and that its solutions, for $\gamma > 0$, verify the fixed point equation,

$$\mathbf{x} = \text{prox}_{\gamma\mathcal{F}_2}(\mathbf{x} - \gamma\nabla\mathcal{F}_1(\mathbf{x})). \tag{B.7}$$

This suggests the following iterative scheme,

$$\mathbf{x}_{n+1} = \text{prox}_{\gamma_n\mathcal{F}_2}(\mathbf{x}_n - \gamma_n\nabla\mathcal{F}_1(\mathbf{x}_n)), \tag{B.8}$$

for appropriate values of the parameter γ_n. This type of scheme is known as forward-backward (FB) algorithm : a forward gradient step using $\mathcal{F}_1$ and a backward step involving $\mathcal{F}_2$ through its proximity operator. Some variants of FB algorithms that have been shown to converge to a solution of (B.6) can be found in [Bauschke *et al.* 2011].

One may refer to [Bauschke *et al.* 2011] for other proximity operator properties and examples.

B Convex conjugate

Definition : let $\mathcal{F}$ be a real-valued function defined on $\mathcal{H}$. The function $\mathcal{F}^* : \mathbf{y} \to \max_{\mathbf{x}} \langle \mathbf{x}, \mathbf{y} \rangle - \mathcal{F}(\mathbf{x})$ is the convex conjugate of $\mathcal{F}$; it is also known as the Legendre-Fenchel transformation of $\mathcal{F}$.

Properties :

— Moreau identity : for $\mathcal{F} \in \Gamma_0(\mathcal{H})$ and $\lambda \in \mathbb{R}_+^*$, $\text{prox}_{\lambda\mathcal{F}}(\mathbf{x}) + \lambda\text{prox}_{\frac{1}{\lambda}\mathcal{F}^*}(\frac{\mathbf{x}}{\lambda}) = \mathbf{x}$;

— Fenchel - Moreau theorem : if $\mathcal{F} \in \Gamma_0(\mathcal{H})$, $\mathcal{F} = \mathcal{F}^{**}$.

Annexe C

Multiple frame super-resolution

Sommaire

A First guess noise

We keep the same notations as in Section B.3.2 in Chapter IV. We note $\mathbf{x}_{wi}$ as the i^{th} scale of $\mathbf{x}_n^{(0)}$ wavelet transform in $\mathbf{W}$. As the noise in $\mathbf{x}_n^{(0)}$ is correlated, the estimation of σ_i is not straightforward, so we proceed as follows :

1. $\sigma_i^0 = 1.4826\,\mathrm{MAD}(\mathbf{x}_{wi})$
2. $\hat{\mathbf{x}}_{wi} = \underset{\mathbf{x}}{\operatorname{argmin}}\ \frac{1}{2}\|\mathbf{x} - \mathbf{x}_{wi}\|_2^2 + k\sigma_i^0\|\mathbf{x}\|_1$
3. $\sigma_i = 1.4826\,\mathrm{MAD}(\widehat{\mathbf{x}}_{wi} - \mathbf{x}_{wi})$.

The factor 1.4826 comes from the assumption that the noise is approximately Gaussian. Finally, the factor k must be sufficiently high so that all the noise will remain in the residual $\widehat{\mathbf{x}}_{wi} - \mathbf{x}_{wi}$. We took $k = 5$. The problem in step 2 has a closed-form solution. For the finer scales, σ_i^0 is quite close to σ_i. But for coarser scales, σ_i^0 is significantly overestimated (it might be more than 10 times greater than σ_i). We implicitly assumed that the noise in the wavelet scales is stationary, which is reasonable apart from the edges effects due to the wavelet transform.

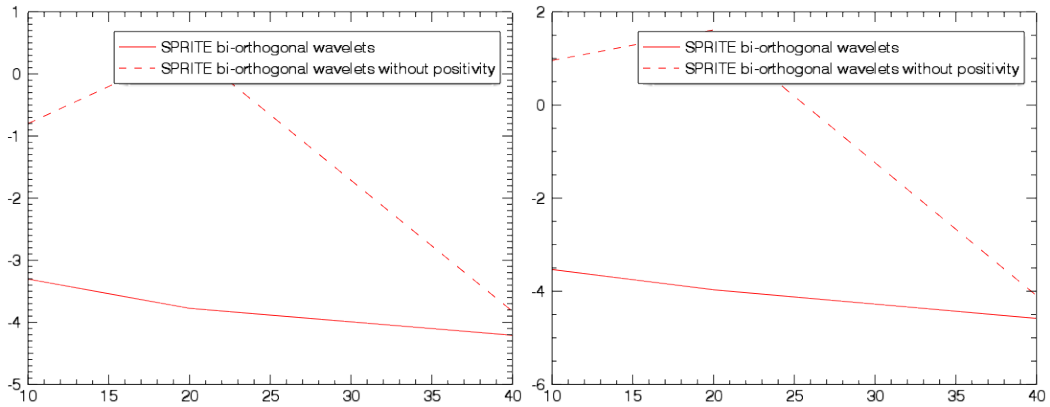
B Positivity

We can drop the positivity constraint by simply solving

$$\min_{\boldsymbol{\Delta}} J_1(\boldsymbol{\Delta} + \mathbf{x}^{(0)}) + \kappa\|\mathbf{w}^{(k)} \odot \boldsymbol{\lambda} \odot \boldsymbol{\Phi}\boldsymbol{\Delta}\|_1 \tag{C.1}$$

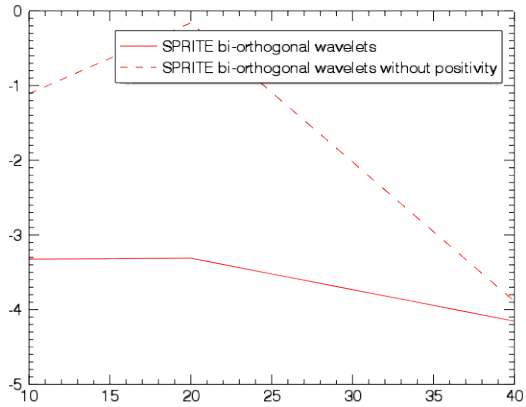
at step 6 in Algorithm IV.2 in Chapter IV. We did a similar numerical experiment as the one presented in Section C in the same chapter to quantify the

impact of this constraint. The comparison is given in Fig. *C.1*. One can see that the positivity constraint is actually important from the point view of the ellipticity parameters at low SNR. This result is illustrated in Fig. *C.2*. With the positivity constraint, the reconstruction is less influenced by the negative oscillations in the data, due to noise ; thus it yields a better robustness. Even if the negative residual values in the final PSF are 1000 order of magnitude smaller than the peak value, it is sufficient to considerably bias the ellipticity measurements.

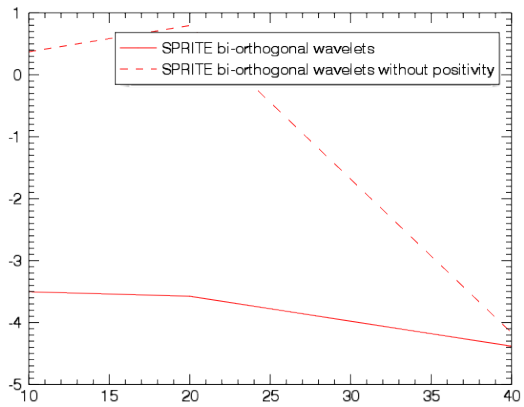


(a) Mean absolute error on the first ellipticity parameter.

(b) Standard deviation of the absolute error on the first ellipticity parameter.



(c) Mean absolute error on the second ellipticity parameter.



(d) Standard deviation of the absolute error on the second ellipticity parameter.

Figure *C.1* – Errors in log scale on ellipticity parameters versus the SNR. The positivity constraint significantly improves the accuracy.

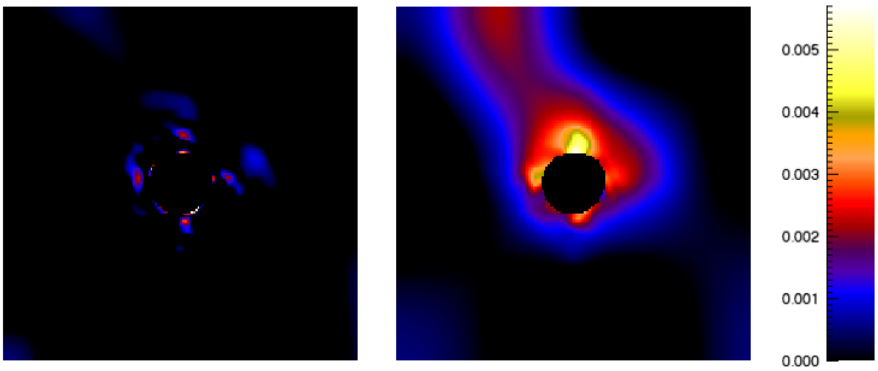


Figure *C.2* – Map of negative values in the PSF reconstruction map (in absolute value) at 15dB. On the left, SPRITE has a positivity constraint ; on the right, SPRITE without positivity constraint.

C Reweighting

As stated in Chapter IV, Section B.1, the reweighting scheme used in SPRITE is meant to mitigate the bias due to the l_1 norm penalty. To verify this, we basically did the same as for the positivity in the previous section. Thus, we ran Algorithm IV.2 with $K_{max} = 1$ and $K_{max} = 2$ and we compute in each case the mean correlation coefficient in Pearson sense [Rodgers & Nicewander 1988] between the reference images and the reconstructions for different SNR. The result is given in Fig. ***C.3***. As expected, the reweighting improves the correlation and consequently, reduces the global bias on the reconstruction.

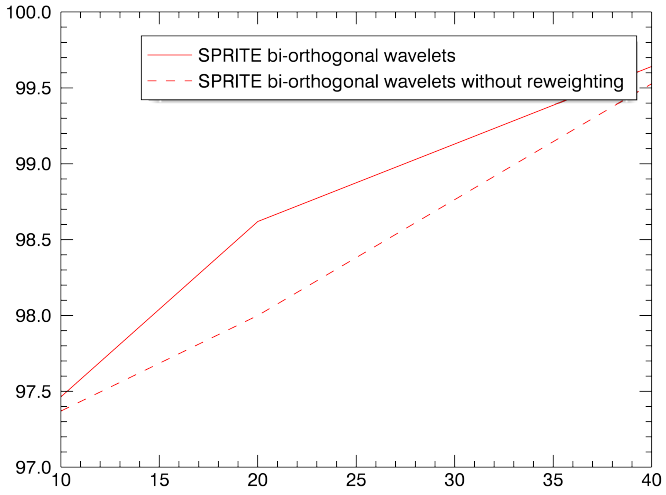


Figure *C.3* – Mean correlation coefficients (×100) between the SPRITE PSF reconstructions and the reference images versus the SNR ; the reweighting reduces the global bias

Annexe D

Joint super-resolution minimization schemes

Sommaire

This appendix details the practical resolution of optimization problems involved in the RCA algorithm.

A Components estimation problem

We consider the step 5 in the Algorithm V.1 in Chapter V. If $\mathbf{\Phi}_s = \mathbf{I}_n$, the problem of estimating components takes the following generic form :

$$\min_{\mathbf{S}} \mathcal{F}(\mathbf{S}) + \mathcal{G}_1(\mathcal{L}_1(\mathbf{S})) + \mathcal{H}(\mathbf{S}), \tag{D.1}$$

with $\mathcal{F}(\mathbf{S}) = \sum_{i=1}^{r} \|\mathbf{w}_i \odot \mathbf{s}_i^{(c)}\|_1$, $\mathcal{G}_1 = \boldsymbol{\iota}_{\mathbb{R}_+^{n\times p}}$, $\mathcal{L}_1(\mathbf{S}) = \mathbf{SA}$ and $\mathcal{H}(\mathbf{S}) = \frac{1}{2}\|\mathbf{Y} - \mathcal{M}(\mathbf{S})\|_F^2$ for some bounded linear operator $\mathcal{M}$.

$\mathcal{F} \in \Gamma_0(\mathbb{R}^{n\times r})$, $\mathcal{G}_1 \in \Gamma_0(\mathbb{R}^{n\times p})$ and $\mathcal{L}_1$ is a bounded linear operator. Moreover, $\mathcal{H}$ is convex, differentiable and has a continuous and Lipschitz gradient. This problem can be solved efficiently using the primal dual algorithms introduced in [Combettes *et al.* 2014] for instance. One only need to be able to compute $\lambda\mathcal{F}$ and $\alpha\mathcal{G}_1^*$ proximity operators, for some given positive reals λ and α and $\mathcal{H}$'s gradient :

— $\text{prox}_{\lambda\mathcal{F}}(\mathbf{S}) = (\hat{s}_{ij})_{\substack{1\leq i\leq n\\ 1\leq j\leq p}}$, with $\hat{s}_{ij} = \text{SoftThresh}_{\lambda\mathbf{w_j}[i]}(\mathbf{s_j}[i])$;

— $\text{prox}_{\alpha\mathcal{G}_1^*}(\mathbf{Z}) = \mathbf{Z} - (\mathbf{Z})_+$

— $\nabla\mathcal{H}(\mathbf{S}) = -\mathcal{M}^*(\mathbf{Y} - \mathcal{M}(\mathbf{S}))$, where $\mathcal{M}^*$ is the adjoint operator of $\mathcal{M}$.

For an arbitrary dictionary $\mathbf{\Phi}_s$, we instead consider the following generic formulation of the problem :

$$\min_{\mathbf{S}} \mathcal{G}_1(\mathcal{L}_1(\mathbf{S})) + \mathcal{G}_2(\mathcal{L}_2(\mathbf{S})) + \mathcal{H}(\mathbf{S}), \tag{D.2}$$

where $\mathcal{G}_2(\mathbf{Z}) = \sum_{i=1}^{r} \|\mathbf{w}_i \odot \mathbf{Z}_i^{(c)}\|_1$ and $\mathcal{L}_2(\mathbf{S}) = [\mathbf{\Phi}_s \mathbf{s}_1^{(c)}, \ldots, \mathbf{\Phi}_s \mathbf{s}_r^{(c)}]$. One can use the algorithms suggested before and minimization will require the computation of $\alpha\mathcal{G}_2^*$ proximity operator, for some given positive real α which is simply given by

$\text{prox}_{\alpha\mathcal{G}_1^*}(\mathbf{Z}) = \mathbf{Z} - \widehat{\mathbf{Z}}$,with $\widehat{\mathbf{Z}}[i,j] = \text{SoftThresh}_{\lambda \mathbf{w_j}[i]}(\mathbf{Z}[i,j])$.

B Coefficients estimation

We consider the step 8 in the Alg. V.1 in Chapter V. The problem takes the generic form :

$$\min_{\boldsymbol{\alpha}} \mathcal{J}(\boldsymbol{\alpha}) \text{ s.t. } \|\boldsymbol{\alpha}[l,:]\|_0 \leq \eta_l,\ l = 1 \ldots r, \tag{D.3}$$

where $\mathcal{J}$ is convex, differentiable and has a continuous and Lipschitz gradient and $\boldsymbol{\alpha} \in \mathbb{R}^{r \times q}$. This problem is combinatorial and its feasible set is non-convex. For typical data sizes in image processing applications and tractable processing time, one can at best reach a "good" local optimum. There is an extensive literature on optimization problems involving the l_0 pseudo-norm. We propose an heuristic based on quite common ideas now and which appears to be convenient from a practical point of view. Let $\boldsymbol{\alpha}^*$ be a global minimum of Problem D.3. For a vector $\mathbf{M} \in \mathbb{R}^{r \times q}$, we define its support as

$$\text{Supp}(\mathbf{M}) = \{(i,j) \in [\![1,r]\!] \times [\![1,q]\!] / |\mathbf{M}[i,j]| \geq 0\}. \tag{D.4}$$

We note $\text{E}_{\boldsymbol{\alpha}^*}$ the set of $r \times q$ real matrices sharing the support of $\boldsymbol{\alpha}^*$:

$$\text{E}_{\boldsymbol{\alpha}^*} = \{\mathbf{M} \in \mathbb{R}^{r \times q} / \text{Supp}(\mathbf{M}) = \text{Supp}(\boldsymbol{\alpha}^*)\}. \tag{D.5}$$

$\text{E}_{\boldsymbol{\alpha}^*}$ is a vector space. In particular, $\text{E}_{\boldsymbol{\alpha}^*}$ is a convex set. Therefore, $\boldsymbol{\alpha}^*$ is a solution of the following problem :

$$\min_{\boldsymbol{\alpha}} \mathcal{J}(\boldsymbol{\alpha}) \text{ s.t. } \boldsymbol{\alpha} \in \text{E}_{\boldsymbol{\alpha}^*}. \tag{D.6}$$

The proposed scheme is motivated by the idea of identifying approximately $\text{E}_{\boldsymbol{\alpha}^*}$ along with the iterative process. One can think of numerous algorithms to solve Problem D.6, all involving the orthogonal projection onto $\text{E}_{\boldsymbol{\alpha}^*}$. We build upon the fast proximal splitting algorithm introduced in [Beck & Teboulle 2009]. For a vector $\mathbf{u} \in \mathbb{R}^q$ we note $\boldsymbol{\sigma}$ a permutation of $[\![1,q]\!]$ verifying $|\mathbf{u}[\boldsymbol{\sigma}(1)]| \geq \cdots \geq |\mathbf{u}[\boldsymbol{\sigma}(q)]|$. For an integer $k \leq q$, we define

$$\text{Supp}_k(\mathbf{u}) = \{i \in [\![1,q]\!] / |\mathbf{u}[i]| \geq |\mathbf{u}[\boldsymbol{\sigma}(k)]|\}, \tag{D.7}$$

Finally for a vector $\boldsymbol{\alpha} \in \mathbb{R}^{r \times q}$, we define the subspace

$$\text{E}_{k,\boldsymbol{\alpha}} = \{\mathbf{M} \in \mathbb{R}^{r \times q} / \text{Supp}_k(\mathbf{M}[i,:]) = \text{Supp}_k(\boldsymbol{\alpha}[i,:]), i = 1 \ldots r\}. \tag{D.8}$$

The proposed scheme is given in Algorithm D.1. f is a positive valued concave increasing function and $\text{proj}_{\text{E}_{\lfloor f(k)\rfloor},\mathbf{U}_k}(.)$ denotes the orthogonal projection onto $\text{E}_{\lfloor f(k)\rfloor,\mathbf{U}_k}$.

Algorithme D.1 Beck-Teboulle proximal gradient algorithm with variable proximity operator

1: Initialization : $\boldsymbol{\alpha}_0 = 0_{\mathbb{R}^{r\times q}}$, $\boldsymbol{\beta}_0 = \boldsymbol{\alpha}_0$, $t_0 = 1$ $\text{res}_{-1} = 0$, $\text{res}_0 = 0$, tol, $k = 0$
2: Minimization
3: **tant que** $k < k_{\max}$ and $|(\text{res}_k - \text{res}_{k-1})/\text{res}_k|$ **faire**
4: $\quad \mathbf{U}_k = \boldsymbol{\beta}_k - \rho^{-1}\nabla\mathcal{J}(\boldsymbol{\beta}_k)$
5: $\quad \boldsymbol{\alpha}_{k+1} = \text{proj}_{\text{E}_{\lfloor f(k)\rfloor},\mathbf{U}_k}(\mathbf{U}_k)$
6: $\quad t_{k+1} = \frac{1+\sqrt{4t_k^2+1}}{2}$
7: $\quad \lambda_k = 1 + \frac{t_k - 1}{t_{k+1}}$
8: $\quad \boldsymbol{\beta}_{k+1} = \boldsymbol{\alpha}_k + \lambda_k(\boldsymbol{\alpha}_{k+1} - \boldsymbol{\alpha}_k)$
9: $\quad \text{res}_{k+1} = \mathcal{J}(\boldsymbol{\beta}_k)$
10: $\quad k = k + 1$
11: **Return :** $\boldsymbol{\alpha}_{k_{\text{stop}}}$.

The solution support size is constraint at step 5 and the size is gradually increased as shown in Fig. ***D.1***.

The convergence analysis this scheme is out of the scope of this paper. However, Fig. ***D.2*** suggests that once an index is included in an iterate support, this index is included in all the forthcoming iterates supports. This implies that at each support size's step in Fig. ***D.1***, the algorithm approximately solves a problem of the following form :

$$\min_{\boldsymbol{\alpha}} \mathcal{J}(\boldsymbol{\alpha}) \text{ s. t. } \boldsymbol{\alpha} \in \text{E}, \tag{D.9}$$

for a given subspace E, which is a convex problem.

This scheme can be viewed as an iterative hard thresholding [Blumensath & Davies 2008], with a decreasing threshold [Mancera & Portilla 2008]. Yet, it is quite easy to get an upper bound of the support size - related to the parameters η_l in Problem D.3 - from the data. Depending on the time one is willing to spend on the coefficients computation, this yields convenient choices for the function f.

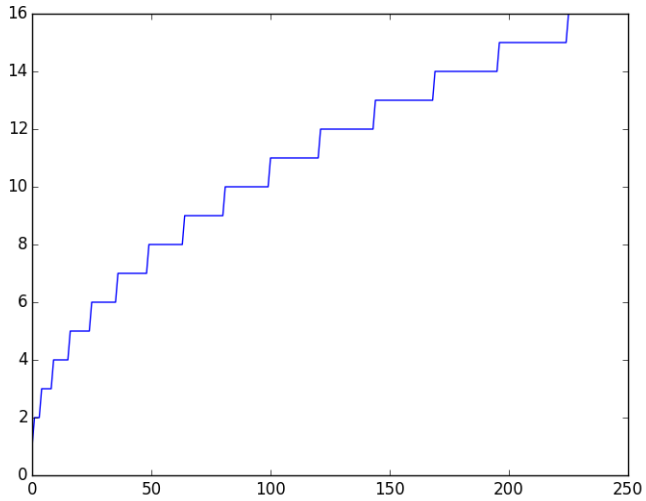


Figure *D.1* – Support size function ; X axis : iteration index k in Algorithm D.1 ; Y axis : $\lfloor f(k) \rfloor$ for $f(x) = \sqrt{x} + 1$

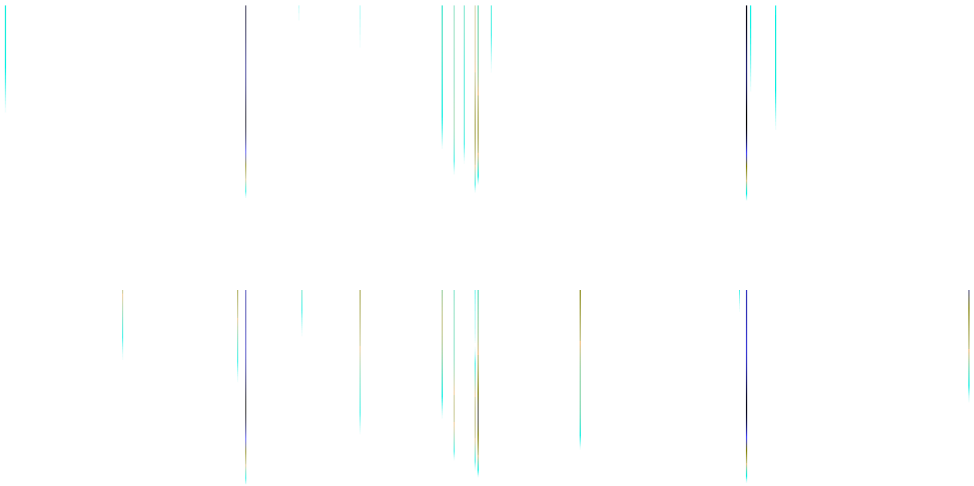


Figure *D.2* – Algorithm D.1 main iterate evolution ; X axis : $|\boldsymbol{\alpha}_{k+1}[0, :]|$ for the top image and $|\boldsymbol{\alpha}_{k+1}[1, :]|$ for the bottom image ; Y axis : iterate index k

Annexe E

Notch filter approximation

In this appendix, we explain why the functional $\Psi_{e,a}$ introduced in Chapter V, subsection C.2 can be approximated with the functional $\widehat{\Psi}_{e,a}$. We reuse the same notations. We consider the 1D case. The samples $(\mathbf{u}_i)_{1\leq i\leq p}$ are uniformly spaced scalar. We assume that $\mathbf{u}_1 < \cdots < \mathbf{u}_p$. We note $\Delta = \mathbf{u}_2 - \mathbf{u}_1$. Thus,

$$\psi_i = \psi_{p-i+1} = \frac{-1}{|k+1-i|^e \Delta^e} \qquad \text{if } i \neq k+1, \text{ and} \tag{E.1}$$

$$\psi_{k+1} = 2\sum_{n=1}^{k} \frac{a}{n^e \Delta^e}. \tag{E.2}$$

Using the centered definition of the convolution with a zero boundary condition, for a vector $\mathbf{v} = (v_i)_{1\leq i\leq p}$, the vector $\mathbf{h} = \mathbf{v} \star \boldsymbol{\psi}_{e,a}$ is given by

$$\mathbf{h}[j] = \sum_{i=1}^{p} v_i \psi_{j+k+1-i}, \tag{E.3}$$

for $j \in [\![1, p]\!]$ and with the convention that $\psi_{j+k+1-i} = 0$ if $j+k+1-i < 1$ or $j+k+1-i > p$. Combining Eq.E.1, E.2 and E.3, we can write

$$\mathbf{h}[j] = (2\sum_{n=1}^{k} \frac{a}{n^e \Delta^e}) v_j - \sum_{\substack{i\in[\max(1,j-k),\\ \min(p,j+k)],i\neq j}} \frac{1}{|j-i|^e \Delta^e} v_i. \tag{E.4}$$

We recall that $\Psi_{e,a}(\mathbf{v}) = \|\mathbf{h}\|_2^2$. On the other hand, $\widehat{\Psi}_{e,a}(\mathbf{v}) = \|\mathbf{t}_\mathbf{v}\|_2^2$, with $\mathbf{t}_\mathbf{v}$ defined as

$$\mathbf{t}_\mathbf{v}[j] = (2 \sum_{n=1}^{\min(j-1,p-j)} \frac{1}{n^e} + \sum_{\substack{n=\\ \min(j-1,p-j)+1}}^{\max(j-1,p-j)} \frac{1}{n^e}) \frac{a}{\Delta^e} v_j - \sum_{\substack{i=1\\ i\neq j}}^{p} \frac{1}{|j-i|^e \Delta^e} v_i \text{ if } j \neq k+1$$

$$\text{and } \mathbf{t}_\mathbf{v}[k+1] = (2\sum_{n=1}^{k} \frac{a}{n^e \Delta^e}) v_j - \sum_{\substack{i=1\\ i\neq k+1}}^{p} \frac{1}{|k+1-i|^e \Delta^e} v_i. \tag{E.5}$$

Thus, $\mathbf{t_v}[k+1] - \mathbf{h}[k+1] = 0$ and for $j \neq k+1$

$$\begin{aligned}\mathbf{t_v}[j] - \mathbf{h}[j] = (\sum_{n=k+1}^{\max(j-1,p-j)} \frac{1}{n^e} - \sum_{\substack{n= \\ \min(j-1,p-j)+1}}^{k} \frac{1}{n^e}) \frac{a}{\Delta^e} v_j \\ - \sum_{\substack{i=1 \\ i\neq j}}^{\max(1,j-k)} \frac{1}{|j-i|^e \Delta^e} v_i - \sum_{\substack{i=\min(p,j+k) \\ i\neq j}}^{p} \frac{1}{|j-i|^e \Delta^e} v_i.\end{aligned} \tag{E.6}$$

Given the symmetry of $\boldsymbol{\psi}_{e,a}$ with respect to $k+1$, we focus on the above difference for $j \leq k$. We further assume that $j \neq 1$. Then, Eq.E.6 simplifies to

$$\mathbf{t_v}[j] - \mathbf{h}[j] = (\sum_{n=k+1}^{p-j} \frac{1}{n^e} - \sum_{n=j}^{k} \frac{1}{n^e}) \frac{a}{\Delta^e} v_j - \frac{1}{(j-1)^e \Delta^e} v_1 - \sum_{n=k}^{p-j} \frac{1}{n^e \Delta^e} v_n. \tag{E.7}$$

Now, using the inequalities for $n > 1$,

$$\int_{n-1}^{n} \frac{1}{(t+1)^e} dt \leq \frac{1}{n^e} \leq \int_{n-1}^{n} \frac{1}{t^e} dt, \tag{E.8}$$

and assuming that $e > 1$, we get the following upper bounding :

$$\begin{aligned}|\mathbf{t_v}[j] - \mathbf{h}[j]| \leq \frac{1}{e-1} [\max(|(p-j+1)^{1-e} + (j-1)^{1-e} - (k+1)^{1-e} - k^{1-e}|, \\ |(p-j)^{1-e} + j^{1-e} - (k+1)^{1-e} - k^{1-e}|)a + \frac{e-1}{(j-1)^e} + k^{1-e} - (p-j)^{1-e}] \frac{\|\mathbf{v}\|_\infty}{\Delta^e}.\end{aligned} \tag{E.9}$$

We see that the higher is k (we recall that $p = 2 * k + 1$) and the closer j is to k, the smaller is the error. Therefore, we use $\mathbf{t_v}$ as an approximation for $\mathbf{h}$, up to boundaries errors.

Annexe F

Multidimensional scaling

We consider a family of vectors $(\mathbf{r}_i)_{1\leq i\leq p}$ in $\mathbb{R}^q$; we set $\mathbf{R} = [\mathbf{r}_1, \ldots, \mathbf{r}_p]$. We assume that $\sum_{i=1}^p \mathbf{r}_i = 0_{\mathbb{R}^q}$. We define the distances matrix $\mathbf{D} = (\|\mathbf{r_i} - \mathbf{r_j}\|_\mathbf{2}^\mathbf{2})_{1\leq i,j\leq p}$ and the norms vector $\mathbf{v} = (\|\mathbf{r}_1\|_2^2, \ldots, \|\mathbf{r}_p\|_2^2)^T$. We want to retrieve $(\mathbf{r}_i)_{1\leq i\leq p}$ from $\mathbf{D}$. $\mathbf{D}$ can be rewritten as :

$$\mathbf{D} = \mathbf{v}\mathbb{1}_p^T + \mathbb{1}_p\mathbf{v}^T - 2\mathbf{R}^T\mathbf{R}. \tag{F.1}$$

We consider the centering matrix $\mathbf{C}$ introduced in Chapter VI, Section D.1. We have $\mathbf{C}\mathbb{1}\mathbf{v}^T = 0_{\mathbb{R}^p} = \mathbf{v}\mathbb{1}_p^T\mathbf{C}$. Thus $-\frac{1}{2}\mathbf{CDC} = (\mathbf{RC})^T\mathbf{RC}$. But $\mathbf{RC} = (\sum_{i=1}^p \mathbf{r}_i)\mathbb{1}_p^T = 0$, which finally gives $-\frac{1}{2}\mathbf{CDC} = \mathbf{R}^T\mathbf{R}$.

Annexe G

Points clouds to image transform

The step 11 in the Algorithm VI.1 in Chapter VI.1 consists in converting estimated Wasserstein barycenters which are points clouds into images which are the actual PSFs estimates. Let consider an estimated Wasserstein $\mathbf{X}_{\text{bar}} \in \mathbb{R}^{3\times N}$. By construction, the first dimension is related to pixels intensities and the second and third are related to pixels positions (see Eq. VI.12). We assume that the PSFs images have N_l lines and N_c columns so that $N = N_l N_c$. Then the most simple way of calculating an image from $\mathbf{X}_{\text{bar}}$ is the following :

1. Initialize an $N_l \times N_c$ image PSF_{bar} with all pixels values set to zeros.
2. FOR ALL $i \in [\![1, N]\!]$:
   find the nearest neighbor (l_i, c_i) of $\mathbf{X}_{\text{bar}}[2:3, i]$ in $[\![1, N_l]\!] \times [\![1, N_c]\!]$.
   Set $\text{PSF}_{\text{bar}}[l_i, c_i] = \text{PSF}_{\text{bar}}[l_i, c_i] + \mathbf{X}_{\text{bar}}[1, i]$.

However, this generally results in sharp pixels intensities variations and therefore visual artifacts. We circumvent this by involving the four nearest pixels in step 2 in the procedure above, rather than the nearest one solely :

1. Initialize an $N_l \times N_c$ image PSF_{bar} with all pixels values set to zeros.
2. FOR ALL $i \in [\![1, N]\!]$:
   find the 4 nearest neighbors $((l_{ij}, c_{ij}))_{1\leq j\leq 4}$ of $\mathbf{X}_{\text{bar}}[2:3, i]$ in $[\![1, N_l]\!] \times [\![1, N_c]\!]$.
   FOR ALL $j \in [\![1, 4]\!]$:
   set $\text{PSF}_{\text{bar}}[l_{ij}, c_{ij}] = \text{PSF}_{\text{bar}}[l_{ij}, c_{ij}] + \frac{1}{\sum_{k=1}^{4} \frac{\|\mathbf{X}_{\text{bar}}[2:3,i]-(l_{ij},c_{ij})^T\|_2^2}{\|\mathbf{X}_{\text{bar}}[2:3,i]-(l_{ik},c_{ik})^T\|_2^2}} \mathbf{X}_{\text{bar}}[1, i]$.

This is illustrated in Fig. ***G.1***.

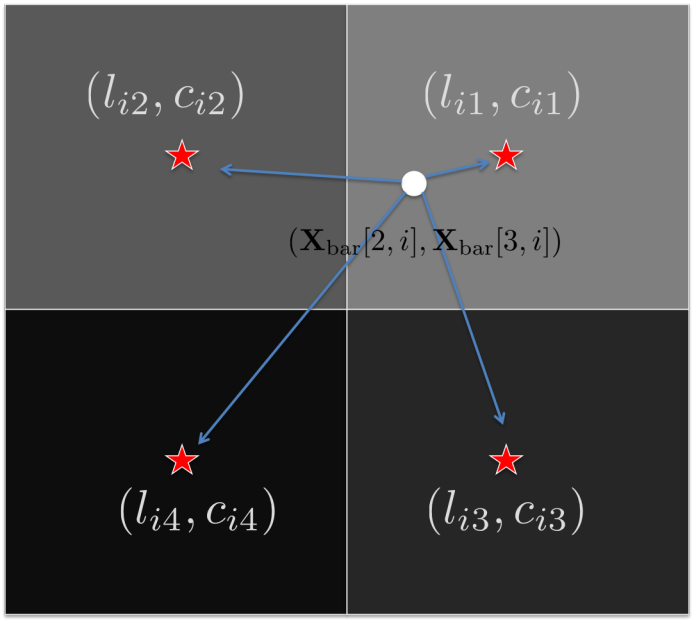


Figure *G.1* – Point cloud to image mapping.

Annexe H

Ellipticity parameters directional derivatives

The ellipticity parameters used in the Chapters IV to VI can be rewritten in the following way :

$$e_1(\mathbf{X}_i) = \frac{< \mathbf{X}_i, \mathbf{U}_5 >< \mathbf{X}_i, \mathbf{U}_3 > - < \mathbf{X}_i, \mathbf{U}_1 >^2 + < \mathbf{X}_i, \mathbf{U}_2 >^2}{< \mathbf{X}_i, \mathbf{U}_4 >< \mathbf{X}_i, \mathbf{U}_3 > - < \mathbf{X}_i, \mathbf{U}_1 >^2 - < \mathbf{X}_i, \mathbf{U}_2 >^2} \tag{H.1}$$

$$e_2(\mathbf{X}_i) = \frac{2(< \mathbf{X}_i, \mathbf{U}_6 >< \mathbf{X}_i, \mathbf{U}_3 > - < \mathbf{X}_i, \mathbf{U}_1 >< \mathbf{X}_i, \mathbf{U}_2 >)}{< \mathbf{X}_i, \mathbf{U}_4 >< \mathbf{X}_i, \mathbf{U}_3 > - < \mathbf{X}_i, \mathbf{U}_1 >^2 - < \mathbf{X}_i, \mathbf{U}_2 >^2}, \tag{H.2}$$

where $\mathbf{U}_1 = (k)_{\substack{1\leq k\leq N_l \\ 1\leq l\leq N_c}}$, $\mathbf{U}_2 = (l)_{\substack{1\leq k\leq N_l \\ 1\leq l\leq N_c}}$, $\mathbf{U}_3 = (1)_{\substack{1\leq k\leq N_l \\ 1\leq l\leq N_c}}$, $\mathbf{U}_4 = (k^2 + l^2)_{\substack{1\leq k\leq N_l \\ 1\leq k\leq N_c}}$, $\mathbf{U}_5 = (k^2 - l^2)_{\substack{1\leq k\leq N_l \\ 1\leq l\leq N_c}}$, $\mathbf{U}_6 = (kl)_{\substack{1\leq k\leq N_l \\ 1\leq l\leq N_c}}$. We derive the following expressions :

$$\frac{de_1(\mathbf{X}_i + t\mathbf{P}_j)}{dt} = \frac{a_1 + a_2 t}{c_t} - e_1(\mathbf{X}_i + t\mathbf{P}_j)\frac{d_1 + d_2 t}{c_t} \tag{H.3}$$

$$\frac{de_2(\mathbf{X}_i + t\mathbf{P}_j)}{dt} = \frac{b_1 + b_2 t}{c_t} - e_2(\mathbf{X}_i + t\mathbf{P}_j)\frac{d_1 + d_2 t}{c_t}, \tag{H.4}$$

where

- $c_t =< \mathbf{U}_4, \mathbf{X}_i + t\mathbf{P}_j >< \mathbf{U}_3, \mathbf{X}_i + t\mathbf{P}_j > - < \mathbf{U}_1, \mathbf{X}_i + t\mathbf{P}_j >^2 - < \mathbf{U}_2, \mathbf{X}_i + t\mathbf{P}_j >^2$
- $a_1 =< \mathbf{U}_5, \mathbf{X}_i >< \mathbf{U}_3, \mathbf{P}_j > + < \mathbf{U}_3, \mathbf{X}_i >< \mathbf{U}_5, \mathbf{P}_j > -2(< \mathbf{U}_1, \mathbf{X}_i >< \mathbf{U}_1, \mathbf{P}_j > - < \mathbf{U}_2, \mathbf{X}_i >< \mathbf{U}_2, \mathbf{P}_j >)$
- $a_2 = 2(< \mathbf{U}_5, \mathbf{P}_j >< \mathbf{U}_3, \mathbf{P}_j > - < \mathbf{U}_1, \mathbf{P}_j >^2 + < \mathbf{U}_2, \mathbf{P}_j >^2)$
- $b_1 = 2(< \mathbf{U}_6, \mathbf{X}_i >< \mathbf{U}_3, \mathbf{P}_j > + < \mathbf{U}_3, \mathbf{X}_i >< \mathbf{U}_6, \mathbf{P}_j > - < \mathbf{U}_1, \mathbf{X}_i >< \mathbf{U}_2, \mathbf{P}_j > - < \mathbf{U}_2, \mathbf{X}_i >< \mathbf{U}_1, \mathbf{P}_j >)$
- $b_2 = 4(< \mathbf{U}_6, \mathbf{P}_j >< \mathbf{U}_3, \mathbf{P}_j > - < \mathbf{U}_2, \mathbf{P}_j >< \mathbf{U}_1, \mathbf{P}_j >)$
- $d_1 =< \mathbf{U}_4, \mathbf{X}_i >< \mathbf{U}_3, \mathbf{P}_j > + < \mathbf{U}_3, \mathbf{X}_i >< \mathbf{U}_4, \mathbf{P}_j > -2(< \mathbf{U}_1, \mathbf{X}_i >< \mathbf{U}_1, \mathbf{P}_j > + < \mathbf{U}_2, \mathbf{X}_i >< \mathbf{U}_2, \mathbf{P}_j >)$
- $d_2 = 2(< \mathbf{U}_4, \mathbf{P}_j >< \mathbf{U}_3, \mathbf{P}_j > - < \mathbf{U}_1, \mathbf{P}_j >^2 - < \mathbf{U}_2, \mathbf{P}_j >^2)$.

Bibliographie

[Abbott *et al.* 2016] BP Abbott, Richard Abbott, TD Abbott, MR Abernathy, Fausto Acernese, Kendall Ackley, Carl Adams, Thomas Adams, Paolo Addesso, RX Adhikariet al. Observation of gravitational waves from a binary black hole merger. Physical review letters, vol. 116, no. 6, page 061102, 2016. (cited on page 32)

[Abdi 2007] Hervé Abdi. Metric Multidimensional Scaling. In N.J. Salkind, editeur, Encyclopedia of Measurement and Statistics., pages 598–605. Sage, Thousand Oaks (CA), 2007. (cited on page 115)

[Absil *et al.* 2009] P-A Absil, Robert Mahony et Rodolphe Sepulchre. Optimization algorithms on matrix manifolds. Princeton University Press, 2009. (cited on page 46)

[Agueh & Carlier 2011] Martial Agueh et Guillaume Carlier. Barycenters in the Wasserstein space. SIAM Journal on Mathematical Analysis, vol. 43, no. 2, pages 904–924, 2011. (cited on page 58)

[Allard *et al.* 2012] William K Allard, Guangliang Chen et Mauro Maggioni. Multi-scale geometric methods for data sets II : Geometric multi-resolution analysis. Applied and Computational Harmonic Analysis, vol. 32, no. 3, pages 435–462, 2012. (cited on pages 51 et 83)

[Ambrosio & Savaré 2007] Luigi Ambrosio et Giuseppe Savaré. Gradient flows in spaces of probability measures. 2007. (cited on page 54)

[Anderson & Morley 1985] W. N. Anderson et T. D. Morley. Eigenvalues of the Laplacian of a graph. Linear and Multilinear Algebra, vol. 18, pages 141–145, 1985. (cited on page 90)

[Bacák *et al.* 2016] Miroslav Bacák, Ronny Bergmann, Gabriele Steidl et Andreas Weinmann. A Second Order Nonsmooth Variational Model for Restoring Manifold-Valued Images. SIAM Journal on Scientific Computing, vol. 38, no. 1, pages A567–A597, 2016. (cited on page 47)

[Baer 2006] Richard L Baer. A model for dark current characterization and simulation. In Electronic Imaging 2006, pages 606805–606805. International Society for Optics and Photonics, 2006. (cited on page 12)

[Baker & Moallem 2007] L. Baker K. et M. Moallem M. Iteratively weighted centroiding for Shack-Hartann wave-front sensors. OPTICS EXPRESS, vol. 15, no. 8, pages 5147–5159, Avril 2007. (cited on page 69)

[Baraniuk & Wakin 2009] Richard G Baraniuk et Michael B Wakin. Random projections of smooth manifolds. Foundations of computational mathematics, vol. 9, no. 1, pages 51–77, 2009. (cited on page 31)

[Basden *et al.* 2004] Alastair Basden, Bob Tubbs et Craig Mackay. L3CCD's : Low Readout Noise CCDs in Astronomy. In Scientific Detectors for Astronomy, pages 599–602. Springer, 2004. (cited on page 12)

[Bauschke *et al.* 2011] H. H. Bauschke, R. S. Burachik, P. L. Combettes et al. Fixed-point algorithms for inverse problems in science and engineering. Springer, 2011. (cited on pages 67 et 147)

[Beck & Teboulle 2009] A. Beck et M. Teboulle. A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems. SIAM Journal on Imaging and Sciences, vol. 2, no. 1, pages 183–202, 2009. (cited on page 154)

[Beckouche *et al.* 2013a] S. Beckouche, J. L. Starck et J. Fadili. Astronomical image denoising using dictionary learning. Astronomy & Astrophysics, vol. 556, page A132, Août 2013. (cited on page 74)

[Beckouche *et al.* 2013b] Simon Beckouche, Jean-Luc Starck et Jalal Fadili. Astronomical image denoising using dictionary learning. Astronomy & Astrophysics, vol. 556, page A132, 2013. (cited on page 38)

[Belkin & Niyogi 2003] Mikhail Belkin et Partha Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. Neural computation, vol. 15, no. 6, pages 1373–1396, 2003. (cited on page 50)

[Benamou & Brenier 2000] Jean-David Benamou et Yann Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. Numerische Mathematik, vol. 84, no. 3, pages 375–393, 2000. (cited on page 55)

[Bengio *et al.* 2003] Yoshua Bengio, Jean-François Paiement et Pascal Vincent. Out-of-Sample Extensions for LLE, Isomap, MDS, Eigenmaps, and Spectral Clustering. Mij, vol. 1, page 2, 2003. (cited on page 52)

[Bertero & Boccacci 1998] Mario Bertero et Patrizia Boccacci. Introduction to inverse problems in imaging. CRC press, 1998. (cited on page 22)

[Bertin 2011] E. Bertin. Automated Morphometry with SExtractor and PSFEx. In I. N. Evans, A. Accomazzi, D. J. Mink et A. H. Rots, editeurs, Astronomical Data Analysis Software and Systems XX, volume 442 of Astronomical Society of the Pacific Conference Series, page 435, Juillet 2011. (cited on page 84)

[Besse 2007] Arthur L Besse. Einstein manifolds. Springer Science & Business Media, 2007. (cited on page 45)

[Blumensath & Davies 2008] Thomas Blumensath et MikeE. Davies. Iterative Thresholding for Sparse Approximations. Journal of Fourier Analysis

and Applications, vol. 14, no. 5-6, pages 629–654, 2008. (cited on pages 35, 91 et 155)

[Blumensath & Davies 2009] Thomas Blumensath et Mike E Davies. Iterative hard thresholding for compressed sensing. Applied and Computational Harmonic Analysis, vol. 27, no. 3, pages 265–274, 2009. (cited on page 35)

[Bobin *et al.* 2007] Jérôme Bobin, Jean-Luc Starck, Jalal Fadili et Yassir Moudden. Sparsity and morphological diversity in blind source separation. IEEE Transactions on Image Processing, vol. 16, no. 11, pages 2662–2674, 2007. (cited on page 38)

[Bobin *et al.* 2015] J. Bobin, F. Sureau et J. Starck. CMB reconstruction from the WMAP and Planck PR2 data. ArXiv e-prints, Novembre 2015. (cited on page 26)

[Bolte *et al.* 2014] Jérôme Bolte, Shoham Sabach et Marc Teboulle. Proximal alternating linearized minimization for nonconvex and nonsmooth problems. Mathematical Programming, vol. 146, no. 1-2, pages 459–494, 2014. (cited on page 91)

[Bonneel *et al.* 2015] N. Bonneel, J. Rabin, G. Peyré et H. Pfister. Sliced and Radon Wasserstein Barycenters of Measures. Journal of Mathematical Imaging and Vision, vol. 51, no. 1, pages 22–45, 2015. (cited on pages 112, 129 et 133)

[Bonneel *et al.* 2016] Nicolas Bonneel, Gabriel Peyré et Marco Cuturi. Wasserstein Barycentric Coordinates : Histogram Regression Using Optimal Transport. ACM Transactions on Graphics, vol. 35, no. 4, 2016. (cited on page 118)

[Borg & Groenen 2005] Ingwer Borg et Patrick JF Groenen. Modern multidimensional scaling : Theory and applications. Springer Science & Business Media, 2005. (cited on page 48)

[Bostan *et al.* 2016] Emrah Bostan, Emmanuel Froustey, Masih Nilchian, Daniel Sage et Michael Unser. Variational Phase Imaging Using the Transport-of-Intensity Equation. IEEE Transactions on Image Processing, vol. 25, no. 2, pages 807–817, 2016. (cited on page 141)

[Bottou 1998] Léon Bottou. Online learning and stochastic approximations. On-line learning in neural networks, vol. 17, no. 9, page 142, 1998. (cited on page 128)

[Boyd & Vandenberghe 2004] Stephen Boyd et Lieven Vandenberghe. Convex optimization. Cambridge university press, 2004. (cited on page 34)

[Buckheit & Donoho 1995] JonathanB. Buckheit et DavidL. Donoho. WaveLab and Reproducible Research. In Anestis Antoniadis et

Georges Oppenheim, editeurs, Wavelets and Statistics, volume 103 of Lecture Notes in Statistics, pages 55–81. Springer New York, 1995. (cited on page 75)

[Burkard *et al.* 2009] Rainer Burkard, Mauro Dell'Amico et Silvano Martello. Assignment problems. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2009. (cited on page 112)

[Cabral *et al.* 2013] Ricardo Cabral, Fernando De La Torre, Joao P Costeira et Alexandre Bernardino. Unifying nuclear norm and bilinear factorization approaches for low-rank matrix decomposition. In Proceedings of the IEEE International Conference on Computer Vision, pages 2488–2495, 2013. (cited on page 40)

[Candès & Donoho 2004] Emmanuel J Candès et David L Donoho. New tight frames of curvelets and optimal representations of objects with piecewise C2 singularities. Communications on pure and applied mathematics, vol. 57, no. 2, pages 219–266, 2004. (cited on page 36)

[Candès & Plan 2010] Emmanuel J Candès et Yaniv Plan. Matrix completion with noise. Proceedings of the IEEE, vol. 98, no. 6, pages 925–936, 2010. (cited on page 40)

[Candès & Recht 2009] Emmanuel J Candès et Benjamin Recht. Exact matrix completion via convex optimization. Foundations of Computational mathematics, vol. 9, no. 6, pages 717–772, 2009. (cited on page 40)

[Candès & Romberg 2007] Emmanuel Candès et Justin Romberg. Sparsity and incoherence in compressive sampling. Inverse problems, vol. 23, no. 3, page 969, 2007. (cited on page 37)

[Candès *et al.* 2006] Emmanuel J Candès, Justin K Romberg et Terence Tao. Stable signal recovery from incomplete and inaccurate measurements. Communications on pure and applied mathematics, vol. 59, no. 8, pages 1207–1223, 2006. (cited on page 34)

[Candès *et al.* 2008] Emmanuel J Candès, Michael B Wakin et Stephen P Boyd. Enhancing sparsity by reweighted ℓ_1 minimization. Journal of Fourier analysis and applications, vol. 14, no. 5-6, pages 877–905, 2008. (cited on pages 39, 66, 67 et 93)

[Candès *et al.* 2011] Emmanuel J Candès, Xiaodong Li, Yi Ma et John Wright. Robust principal component analysis ? Journal of the ACM (JACM), vol. 58, no. 3, page 11, 2011. (cited on page 40)

[Cartis & Thompson 2015] C. Cartis et A. Thompson. A New and Improved Quantitative Recovery Analysis for Iterative Hard Thresholding

Algorithms in Compressed Sensing. Information Theory, IEEE Transactions on, vol. 61, no. 4, pages 2019–2042, April 2015. (cited on page 91)

[Chadan & Sabatier 2011] Khosrow Chadan et Pierre C. Sabatier. Inverse problems in quantum scattering theory. Springer Publishing Company, Incorporated, 2nd édition, 2011. (cited on page 22)

[Chambolle *et al.* 2010] Antonin Chambolle, Vicent Caselles, Daniel Cremers, Matteo Novaga et Thomas Pock. An introduction to total variation for image analysis. Theoretical foundations and numerical methods for sparse recovery, vol. 9, no. 263-340, page 227, 2010. (cited on page 27)

[Chan *et al.* 2006] Tony F. Chan, Selim Esedoglu, Frederick E. Park et Andy M. Yip. Total Variation Image Restoration : Overview and Recent Developments. In Nikos Paragios, Yunmei Chen et Olivier D. Faugeras, editeurs, Handbook of Mathematical Models in Computer Vision, pages 17–31. Springer, 2006. (cited on page 27)

[Chandrasekaran *et al.* 2010] V. Chandrasekaran, B. Recht, P. A. Parrilo et A. S. Willsky. The Convex Geometry of Linear Inverse Problems. ArXiv e-prints, Décembre 2010. (cited on pages 34 et 40)

[Chang 2003] Chein-I Chang. Hyperspectral imaging : techniques for spectral detection and classification, volume 1. Springer Science & Business Media, 2003. (cited on page 39)

[Chen *et al.* 2001] Scott Shaobing Chen, David L Donoho et Michael A Saunders. Atomic decomposition by basis pursuit. SIAM review, vol. 43, no. 1, pages 129–159, 2001. (cited on page 34)

[Chen *et al.* 2014] Yuansi Chen, Julien Mairal et Zaid Harchaoui. Fast and robust archetypal analysis for representation learning. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 1478–1485, 2014. (cited on page 25)

[Cheng 2006] Qiuming Cheng. Spatial and Spatially Weighted Principal Component Analysis for Images Processing. In Geoscience and Remote Sensing Symposium, 2006. IGARSS 2006. IEEE International Conference on, pages 972–975, July 2006. (cited on page 82)

[Chung 1997] Fan RK Chung. Spectral graph theory, volume 92. American Mathematical Soc., 1997. (cited on pages 50 et 90)

[Ciak *et al.* 2013] René Ciak, Behrang Shafei et Gabriele Steidl. Homogeneous penalizers and constraints in convex image restoration. Journal of mathematical imaging and vision, vol. 47, no. 3, pages 210–230, 2013. (cited on page 35)

[Claerbout & Muir 1973] Jon F Claerbout et Francis Muir. Robust modeling with erratic data. Geophysics, vol. 38, no. 5, pages 826–844, 1973. (cited on page 32)

[Coifman & Donoho 1995] Ronald R Coifman et David L Donoho. Translation-invariant de-noising. Springer, 1995. (cited on page 37)

[Colton & Kress 1992] David L. Colton et Rainer Kress. Inverse acoustic and electromagnetic scattering theory. Applied mathematical sciences. Springer-Verlag, New York, 1992. (cited on page 22)

[Combettes & Wadjs 2005] P. L. Combettes et V. R. Wadjs. Signal recovery by proximal forward-backward splitting. Multiscale model. Simul., vol. 4, pages 1168–1200, 2005. (cited on page 147)

[Combettes *et al.* 2014] P. L. Combettes, L. Condat, J.-C. Pesquet et B. Cong Vu. A forward-backward view of some primal-dual optimization methods in image recovery. ArXiv e-prints, Juin 2014. (cited on page 153)

[Cropper 2013] M. et al. Cropper. Defining a weak lensing experiment in space. Monthly Notices of the Royal Astronomical Society, vol. 431, 2013. (cited on pages 72 et 101)

[Cuturi & Doucet 2013] M. Cuturi et A. Doucet. Fast Computation of Wasserstein Barycenters. ArXiv e-prints, Octobre 2013. (cited on page 113)

[Demšar *et al.* 2013] Urška Demšar, Paul Harris, Chris Brunsdon, A. Stewart Fotheringham et Sean McLoone. Principal Component Analysis on Spatial Data : An Overview. Annals of the Association of American Geographers, vol. 103, no. 1, pages 106–128, 2013. (cited on page 83)

[Dodelson 2003] Scott Dodelson. Modern cosmology. Academic Press, San Diego, CA, 2003. (cited on pages 3 et 97)

[Dokmanic *et al.* 2015] Ivan Dokmanic, Reza Parhizkar, Juri Ranieri et Martin Vetterli. Euclidean distance matrices : A short walk through theory, algorithms and applications. arXiv preprint arXiv :1502.07541, 2015. (cited on page 116)

[Dominikus 1997] NOLL Dominikus. Restoration of degraded images with maximum entropy. Journal of Global Optimization, vol. 10, no. 1, pages 91–103, 1997. (cited on page 27)

[Donoho & Grimes 2003] David L Donoho et Carrie Grimes. Hessian eigenmaps : Locally linear embedding techniques for high-dimensional data. Proceedings of the National Academy of Sciences, vol. 100, no. 10, pages 5591–5596, 2003. (cited on pages 50, 51 et 115)

[Donoho & Grimes 2005] David L Donoho et Carrie Grimes. Image manifolds which are isometric to Euclidean space. Journal of mathematical imaging and vision, vol. 23, no. 1, pages 5–24, 2005. (cited on page 48)

[Donoho 2006] David L Donoho. Compressed sensing. IEEE Transactions on information theory, vol. 52, no. 4, pages 1289–1306, 2006. (cited on page 34)

[Dupé *et al.* 2011] F.-X. Dupé, J. Fadili et J.-L. Starck. Deconvolution under Poisson noise using exact data fidelity and synthesis or analysis sparsity priors. ArXiv e-prints, Mars 2011. (cited on page 25)

[Duval 016] Laurent Duval. WITS, Where Is The Starlet ? a collection of wavelet names and properties ending with *let, (accessed July 13, 2016). (cited on page 37)

[Eberly 2002] David Eberly. Thin Plate Splines. Geometric Tools Inc., 2002., 2002. (cited on page 116)

[Einstein 1915] Albert Einstein. Die Feldgleichungen der Gravitation. Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), Seite 844-847., 1915. (cited on page 1)

[Elad & Aharon 2006] Michael Elad et Michal Aharon. Image denoising via learned dictionaries and sparse representation. In 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), volume 1, pages 895–900. IEEE, 2006. (cited on page 38)

[Elad & Bruckstein 2002] Michael Elad et Alfred M Bruckstein. A generalized uncertainty principle and sparse representation in pairs of bases. IEEE Transactions on Information Theory, vol. 48, no. 9, pages 2558–2567, 2002. (cited on pages 34 et 37)

[Elad & hel Or 2001] M. Elad et Y. hel Or. A Fast Super-Resolution Reconstruction Algorithm for pure translational motion and common space invariant blur. IEEE Transactions on Image Processing, vol. 10, pages 1187–1193, Août 2001. (cited on page 63)

[Elad *et al.* 2007] Michael Elad, Peyman Milanfar et Ron Rubinstein. Analysis versus synthesis in signal priors. Inverse problems, vol. 23, no. 3, page 947, 2007. (cited on pages 36 et 100)

[Elvira *et al.* 2015] C. Elvira, P. Chainais et N. Dobigeon. Bayesian anti-sparse coding. ArXiv e-prints, Décembre 2015. (cited on page 28)

[Epperson 1987] James F Epperson. On the Runge example. American Mathematical Monthly, vol. 94, no. 4, pages 329–341, 1987. (cited on page 109)

[ESA/SRE 2011] ESA/SRE. EUCLID Mapping the geometry of the dark universe. Rapport technique, ESA, Juillet 2011. (cited on page 72)

[Fazel *et al.* 2001] Maryam Fazel, Haitham Hindi et Stephen P Boyd. A rank minimization heuristic with application to minimum order system approximation. In American Control Conference, 2001. Proceedings of the 2001, volume 6, pages 4734–4739. IEEE, 2001. (cited on page 40)

[Fazel *et al.* 2003] Maryam Fazel, Haitham Hindi et Stephen P Boyd. Log-det heuristic for matrix rank minimization with applications to Hankel and Euclidean distance matrices. In American Control Conference, 2003. Proceedings of the 2003, volume 3, pages 2156–2162. IEEE, 2003. (cited on page 39)

[Fazel 2002] Maryam Fazel. Matrix rank minimization with applications. PhD thesis, Stanford University, 2002. (cited on page 39)

[Flaherty & do Carmo 2013] F. Flaherty et M.P. do Carmo. Riemannian geometry. Mathematics : Theory & Applications. Birkhäuser Boston, 2013. (cited on page 43)

[Freund 1992] Roland W Freund. Conjugate gradient-type methods for linear systems with complex symmetric coefficient matrices. SIAM Journal on Scientific and Statistical Computing, vol. 13, no. 1, pages 425–448, 1992. (cited on page 31)

[Frieden 1972] B Roy Frieden. Restoring with maximum likelihood and maximum entropy. JOSA, vol. 62, no. 4, pages 511–518, 1972. (cited on page 27)

[Gagnon & Jouan 1997] Langis Gagnon et Alexandre Jouan. Speckle filtering of SAR images : a comparative study between complex-wavelet-based and standard filters. In Optical Science, Engineering and Instrumentation'97, pages 80–91. International Society for Optics and Photonics, 1997. (cited on page 22)

[Gangbo & Swiech 1998] Wilfrid Gangbo et Andrzej Swiech. Optimal maps for the multidimensional Monge-Kantorovich problem. Communications on Pure and Applied Mathematics, vol. 51, no. 1, pages 23–45, 1998. (cited on page 113)

[Gentile *et al.* 2013] M. Gentile, F. Courbin et G. Meylan. Interpolating point spread function anisotropy. Astronomy & Astrophysics, vol. 549, page A1, Janvier 2013. (cited on pages 107, 109 et 120)

[Goodman 2005] Joseph W Goodman. Introduction to fourier optics. Roberts and Company Publishers, 2005. (cited on page 5)

[Gower 1968] John C Gower. Adding a point to vector diagrams in multivariate analysis. Biometrika, vol. 55, no. 3, pages 582–585, 1968. (cited on page 118)

[Gramfort *et al.* 2015] A. Gramfort, G. Peyré et M. Cuturi. Fast Optimal Transport Averaging of Neuroimaging Data. ArXiv e-prints, Mars 2015. (cited on page 113)

[Gray 2005] Robert M Gray. Toeplitz and circulant matrices : a review. Communications and Information Theory, vol. 2, no. 3, pages 155–239, 2005. (cited on pages 29 et 65)

[Gribonval 2011] Rémi Gribonval. Should penalized least squares regression be interpreted as maximum a posteriori estimation ? IEEE Transactions on Signal Processing, vol. 59, no. 5, pages 2405–2410, 2011. (cited on page 28)

[Hadamard 1902] J. Hadamard. Sur les problèmes aux dérivés partielles et leur signification ph Princeton University Bulletin, vol. 13, pages 49–52, 1902. (cited on page 23)

[Hadamard 1923] J. Hadamard. Lectures on cauchy's problem in linear partial differential equations. Yale University Press, New Haven, CT, 1923. (cited on page 23)

[Hammond *et al.* 2009] D. K Hammond, P. Vandergheynst et R. Gribonval. Wavelets on Graphs via Spectral Graph Theory. ArXiv e-prints, Décembre 2009. (cited on page 90)

[Hansen 1987] Per Christian Hansen. The truncatedsvd as a method for regularization. BIT Numerical Mathematics, vol. 27, no. 4, pages 534–553, 1987. (cited on pages 28 et 29)

[Harris *et al.* 2011] Paul Harris, Chris Brunsdon et Martin Charlton. Geographically weighted principal components analysis. International Journal of Geographical Information Science, vol. 25, no. 10, pages 1717–1736, 2011. (cited on page 83)

[Herman & Sabatier 1987] G.T. Herman et P.C. Sabatier. Basic methods of tomography and inverse problems : a set of lectures. Malvern physics series. A. Hilger, 1987. (cited on page 22)

[Hestenes & Stiefel 1952] Magnus R Hestenes et Eduard Stiefel. Methods of Conjugate Gradients for Solving Linear Systems1. Journal of Research of the National Bureau of Standards, vol. 49, no. 6, 1952. (cited on page 30)

[Hopf & Rinow 1931] Heinz Hopf et Willi Rinow. Über den Begriff der vollständigen differentialgeometrischen Fläche. Commentarii Mathematici Helvetici, vol. 3, no. 1, pages 209–225, 1931. (cited on page 46)

[Huang *et al.* 2010] Junzhou Huang, Tong Zhanget al. The benefit of group sparsity. The Annals of Statistics, vol. 38, no. 4, pages 1978–2004, 2010. (cited on page 35)

[Jacques *et al.* 2011] Laurent Jacques, Laurent Duval, Caroline Chaux et Gabriel Peyré. A panorama on multiscale geometric representations, intertwining spatial, directional and frequency selectivity. Signal Processing, vol. 91, no. 12, pages 2699–2730, 2011. (cited on page 37)

[Jarvis & Jain 2004] M. Jarvis et B. Jain. Principal Component Analysis of PSF Variation in Weak Lensing Surveys. ArXiv Astrophysics e-prints, Décembre 2004. (cited on page 108)

[Jarvis *et al.* 2008] M. Jarvis, P. Schechter et B. Jain. Telescope Optics and Weak Lensing : PSF Patterns due to Low Order Aberrations. ArXiv e-prints, Octobre 2008. (cited on page 108)

[Jee *et al.* 2007] M. J. Jee, J. P. Blakeslee, M. Sirianni, A. R. Martel, R. L. White et H. C. Ford. Principal Component Analysis of the Time- and Position-dependent Point-Spread Function of the Advanced Camera for Surveys. Publications of the Astronomical Society of the Pacific, vol. 119, no. 862, pages pp. 1403–1419, 2007. (cited on page 108)

[Jezierska *et al.* 2013] Anna Jezierska, Emilie Chouzenoux, Jean-Christophe Pesquet et Hugues Talbot. A convex approach for image restoration with exact poisson-gaussian likelihood. IEEE Trans. Image Process, 2013. (cited on page 25)

[Kantorovitch 1942] L Kantorovitch. ON THE TRANSLOCATION OF MASSES. Doklady) de l'Academie des Sciences de l'URSS, vol. 37, 1942. (cited on pages 54 et 111)

[Karl 2005] W.Clem Karl. 3.6 - Regularization in Image Restoration and Reconstruction. In AL BOVIK, editeur, Handbook of Image and Video Processing (Second Edition), Communications, Networking and Multimedia, pages 183 – V. Academic Press, Burlington, second edition édition, 2005. (cited on pages 26 et 30)

[Kim & Park 2007] Hyunsoo Kim et Haesun Park. Sparse non-negative matrix factorizations via alternating non-negativity-constrained least squares for microarray data analysis. Bioinformatics, vol. 23, no. 12, pages 1495–1502, 2007. (cited on page 40)

[Kuhn 1955] Harold W Kuhn. The Hungarian Method for the Assignment Problem. 50 Years of Integer Programming 1958–2008, page 29, 1955. (cited on page 129)

[Landweber 1951] Louis Landweber. An iteration formula for Fredholm integral equations of the first kind. American journal of mathematics, vol. 73, no. 3, pages 615–624, 1951. (cited on page 30)

[Laureijs *et al.* 2011] R Laureijs, J Amiaux, S Arduini, J-L Augueres, J Brinchmann, R Cole, M Cropper, C Dabin, L Duvet, A Ealetet al.

Euclid Definition Study Report. arXiv preprint arXiv :1110.3193, 2011. (cited on pages 3 et 13)

[Lee *et al.* 2008] Ann B. Lee, Boaz Nadler et Larry Wasserman. Treelets—An adaptive multi-scale basis for sparse unordered data. Ann. Appl. Stat., vol. 2, no. 2, pages 435–471, 06 2008. (cited on page 83)

[Lee 2006] John M Lee. Riemannian manifolds : an introduction to curvature, volume 176. Springer Science & Business Media, 2006. (cited on page 46)

[Levinson 1946] Norman Levinson. The Wiener (Root Mean Square) Error Criterion in Filter Design and Prediction. Journal of Mathematics and Physics, vol. 25, no. 1-4, pages 261–278, 1946. (cited on page 26)

[Little 2011] Anna V Little. Estimating the Intrinsic Dimension of High-Dimensional Data Sets : A Multiscale, Geometric Approach. PhD thesis, Duke University, 2011. (cited on pages 52 et 98)

[Liu *et al.* 2012] Yipeng Liu, Ivan Gligorijevic, Vladimir Matic, Maarten De Vos et Sabine Van Huffel. Multi-sparse signal recovery for compressive sensing. In 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pages 1053–1056. IEEE, 2012. (cited on page 32)

[Maggioni *et al.* 2014] M. Maggioni, S. Minsker et N. Strawn. Multiscale Dictionary Learning : Non-Asymptotic Bounds and Robustness. ArXiv e-prints, Janvier 2014. (cited on page 83)

[Mahajan 1998] Virendra N Mahajan. Optical imaging and aberrations : Part i : Ray geometrical optics. Bellingham : SPIE-The International Society for Optical Engineering, 1998. (cited on page 11)

[Mairal *et al.* 2008] J. Mairal, M. Elad et G. Sapiro. Sparse Representation for Color Image Restoration. IEEE Transactions on Image Processing, vol. 17, no. 1, pages 53–69, Jan 2008. (cited on page 38)

[Mallat & Zhang 1993] Stéphane G Mallat et Zhifeng Zhang. Matching pursuits with time-frequency dictionaries. IEEE Transactions on signal processing, vol. 41, no. 12, pages 3397–3415, 1993. (cited on pages 31 et 35)

[Mallat 1999] Stéphane Mallat. A wavelet tour of signal processing. Academic press, 1999. (cited on page 36)

[Mancera & Portilla 2008] Luis Mancera et Javier Portilla. Non-convex sparse optimization through deterministic annealing and applications. In ICIP, pages 917–920. IEEE, 2008. (cited on page 155)

[Mann & Haykin 1995] Steve Mann et Simon Haykin. The chirplet transform : Physical considerations. IEEE Transactions on Signal Processing, vol. 43, no. 11, pages 2745–2761, 1995. (cited on page 33)

[Marden 1995] John I Marden. Analyzing and modeling rank data. London ; New York : Chapman & Hall, 1st ed édition, 1995. Includes bibliographical references (p. [302]-310) and index. (cited on page 116)

[Markovsky 2008] Ivan Markovsky. Structured low-rank approximation and its applications. Automatica, vol. 44, no. 4, pages 891–909, 2008. (cited on page 41)

[Massey & Refregier 2005] R. Massey et A. Refregier. Polar shapelets. Monthly Notices of the Royal Astronomy Society, vol. 363, pages 197–210, Octobre 2005. (cited on page 108)

[Massey *et al.* 2014] R. Massey, T. Schrabback, O. Cordes, O. Marggraf, H. Israel, L. Miller, D. Hall, M. Cropper, T. Prod'homme et S.-M. Niemi. An improved model of charge transfer inefficiency and correction algorithm for the Hubble Space Telescope. Monthly Notices of the Royal Astronomy Society, vol. 439, pages 887–907, Mars 2014. (cited on page 12)

[McCann 1997] Robert J McCann. A convexity principle for interacting gases. Advances in mathematics, vol. 128, no. 1, pages 153–179, 1997. (cited on pages 57 et 112)

[Mohr *et al.* 2012] J. J. Mohr, R. Armstrong, E. Bertin, G. Daues, S. Desai, M. Gower, R. Gruendl, W. Hanlon, N. Kuropatkin, H. Lin, J. Marriner, D. Petravic, I. Sevilla, M. Swanson, T. Tomashek, D. Tucker et B. Yanny. The Dark Energy Survey data processing and calibration system. In Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, volume 8451 of Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Septembre 2012. (cited on page 63)

[Monge 1781] Gaspard Monge. Mémoire sur la théorie des déblais et des remblais. Histoire de l'Académie Royale des Sciences, pages 666–704, 1781. (cited on page 52)

[Nikolova 2007] Mila Nikolova. Model distortions in Bayesian MAP reconstruction. Inverse Problems and Imaging, vol. 1, no. 2, page 399, 2007. (cited on page 28)

[Nyquist 1928] Harry Nyquist. Thermal agitation of electric charge in conductors. Physical review, vol. 32, no. 1, page 110, 1928. (cited on page 12)

[Otto 2001] Felix Otto. THE GEOMETRY OF DISSIPATIVE EVOLUTION EQUATIONS : THE POROUS MEDIUM EQUATION. Communications in Partial Differential Equations, vol. 26, no. 1-2, pages 101–174, 2001. (cited on page 57)

[Papadakis *et al.* 2013] N. Papadakis, G. Peyré et E. Oudet. Optimal Transport with Proximal Splitting. ArXiv e-prints, Avril 2013. (cited on pages 111 et 135)

[Pare 2000] Thomas Pare. Passivity Based Analysis and Control of Nonlinear Systems. PhD thesis, Stanford University, Department of Mechanical Engineering, Cambridge MA, November 2000. (cited on page 39)

[Park *et al.* 2003] S. C. Park, M. K. Park et M. G. Kang. Super-Resolution Image Reconstruction : a technical overview. IEEE Signal Processing Magazine, vol. 3, pages 1053–5888, Mai 2003. (cited on page 62)

[Pati *et al.* 1993] Yagyensh Chandra Pati, Ramin Rezaiifar et PS Krishnaprasad. Orthogonal matching pursuit : Recursive function approximation with applications to wavelet decomposition. In Signals, Systems and Computers, 1993. 1993 Conference Record of The Twenty-Seventh Asilomar Conference on, pages 40–44. IEEE, 1993. (cited on page 35)

[Paulin-Henriksson *et al.* 2008] S. Paulin-Henriksson, A. Amara, L. Voigt, A. Refregier et S. L. Bridle. Point spread function calibration requirements for dark energy from cosmic shear. Astronomy & Astrophysics, vol. 484, pages 67–77, Juin 2008. (cited on pages 73 et 97)

[Penzias & Wilson 1965] Arno A Penzias et Robert Woodrow Wilson. A Measurement of Excess Antenna Temperature at 4080 Mc/s. The Astrophysical Journal, vol. 142, pages 419–421, 1965. (cited on page 1)

[Peyré & Fadili 2011] Gabriel Peyré et Jalal Fadili. Group sparsity with overlapping partition functions. In Signal Processing Conference, 2011 19th European, pages 303–307. IEEE, 2011. (cited on page 35)

[Peyré 2009] Gabriel Peyré. Manifold models for signals and images. Computer Vision and Image Understanding, vol. 113, no. 2, pages 249–260, 2009. (cited on page 46)

[Piotrowski *et al.* 2013] L. W. Piotrowski, T. Batsch, H. Czyrkowski, M. Cwiok, R. Dabrowski, G. Kasprowicz, A. Majcher, A. Majczyna, K. Malek, L. Mankiewicz, K. Nawrocki, R. Opiela, M. Siudek, M. Sokolowski, R. Wawrzaszek, G. Wrochna, M. Zaremba et A. F. Żarnecki. PSF modelling for very wide-field CCD astronomy. Astronomy & Astrophysics, vol. 551, page A119, Mars 2013. (cited on page 109)

[Prato *et al.* 2012] Marco Prato, Roberto Cavicchioli, Luca Zanni, Patrizia Boccacci et Mario Bertero. Efficient deconvolution methods for

astronomical imaging : algorithms and IDL-GPU codes. Astronomy & Astrophysics, vol. 539, page A133, 2012. (cited on page 31)

[Rabin *et al.* 2011] Julien Rabin, Gabriel Peyré, Julie Delon et Marc Bernot. Wasserstein Barycenter and Its Application to Texture Mixing. In Alfred M. Bruckstein, Bart M. ter Haar Romeny, Alexander M. Bronstein et Michael M. Bronstein, editeurs, SSVM, volume 6667 of Lecture Notes in Computer Science, pages 435–446. Springer, 2011. (cited on pages 112, 128, 129 et 135)

[Raguet *et al.* 2011] H. Raguet, J. Fadili et G. Peyré. Generalized Forward-Backward Splitting. ArXiv e-prints, Août 2011. (cited on pages 67 et 71)

[Rakhuba & Oseledets 2014] M. V. Rakhuba et I. V. Oseledets. Fast multidimensional convolution in low-rank formats via cross approximation. ArXiv e-prints, Février 2014. (cited on page 87)

[Raskutti *et al.* 2014] Garvesh Raskutti, Martin J Wainwright et Bin Yu. Early stopping and non-parametric regression : an optimal data-dependent stopping rule. Journal of Machine Learning Research, vol. 15, no. 1, pages 335–366, 2014. (cited on page 29)

[Recht *et al.* 2010] Benjamin Recht, Maryam Fazel et Pablo A Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. SIAM review, vol. 52, no. 3, pages 471–501, 2010. (cited on page 40)

[Refregier 2003] A. Refregier. Shapelets - I. A method for image analysis. Monthly Notices of the Royal Astronomy Society, vol. 338, pages 35–47, Janvier 2003. (cited on page 108)

[Riess *et al.* 1998] Adam G Riess, Alexei V Filippenko, Peter Challis, Alejandro Clocchiatti, Alan Diercks, Peter M Garnavich, Ron L Gilliland, Craig J Hogan, Saurabh Jha, Robert P Kirshneret al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. The Astronomical Journal, vol. 116, no. 3, page 1009, 1998. (cited on page 1)

[Rodgers & Nicewander 1988] Joseph L. Rodgers et Alan W. Nicewander. Thirteen Ways to Look at the Correlation Coefficient. The American Statistician, vol. 42, no. 1, pages 59–66, 1988. (cited on page 151)

[Romano *et al.* 2010] A. Romano, L. Fu, F. Giordano, R. Maoli, P. Martini, M. Radovich, R. Scaramella, V. Antonuccio-Delogu, A. Donnarumma, S. Ettori, K. Kuijken, M. Meneghetti, L. Moscardini, S. Paulin-Henriksson, E. Giallongo, R. Ragazzoni, A. Baruffolo, A. Dipaola, E. Diolaiti, J. Farinato, A. Fontana, S. Gallozzi, A. Grazian, J. Hill,

F. Pedichini, R. Speziali, R. Smareglia et V. Testa. Abell 611. I. Weak lensing analysis with LBC. Astronomy & Astrophysics, vol. 514, page A88, Mai 2010. (cited on page 108)

[Rowe *et al.* 2011] Barnaby Rowe, Christopher Hirata et Jason Rhodes. Optimal linear image combination. The Astrophysical Journal, vol. 741, no. 1, page 46, 2011. (cited on page 60)

[Roweis & Saul 2000] Sam T Roweis et Lawrence K Saul. Nonlinear dimensionality reduction by locally linear embedding. Science, vol. 290, no. 5500, pages 2323–2326, 2000. (cited on pages 49 et 117)

[Rubinstein *et al.* 2013] Ron Rubinstein, Tomer Peleg et Michael Elad. Analysis K-SVD : a dictionary-learning algorithm for the analysis sparse model. IEEE Transactions on Signal Processing, vol. 61, no. 3, pages 661–677, 2013. (cited on page 38)

[Rudin *et al.* 1992] Leonid I Rudin, Stanley Osher et Emad Fatemi. Nonlinear total variation based noise removal algorithms. Physica D : Nonlinear Phenomena, vol. 60, no. 1, pages 259–268, 1992. (cited on page 27)

[Schena *et al.* 1995] Mark Schena, Dari Shalon, Ronald W Davis et Patrick O Brown. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science, vol. 270, no. 5235, page 467, 1995. (cited on page 35)

[Shannon 1948] CE Shannon. A Mathematical Theory of Communication. Bell Sys. Tech. J., vol. 27, pages 379–423, Juillet 1948. (cited on page 63)

[Soussen *et al.* 2015] Charles Soussen, Jérome Idier, Junbo Duan et David Brie. Homotopy based algorithms for L0-regularized least-squares. IEEE Transactions on Signal Processing, vol. 63, no. 13, pages 3301–3316, Juillet 2015. CRAN / Dépt. SBS. (cited on page 91)

[Stabenau *et al.* 2007] H. F. Stabenau, B. Jain, G. Bernstein et M. Lampton. Lensing Systematics from Space : Modeling PSF effects in the SNAP survey. ArXiv e-prints, Octobre 2007. (cited on page 108)

[Starck & Murtagh 2006] J-L Starck et Fionn Murtagh. Astronomical image and data analysis. Springer Science & Business Media, 2006. (cited on page 27)

[Starck *et al.* 2004] Jean-Luc Starck, Michael Elad et David Donoho. Redundant multiscale transforms and their application for morphological component separation. Advances in Imaging and Electron Physics, vol. 132, pages 287–348, 2004. (cited on pages 35 et 37)

[Starck *et al.* 2005] J-L Starck, Michael Elad et David L Donoho. Image decomposition via the combination of sparse representations and a variational approach. IEEE transactions on image processing, vol. 14, no. 10, pages 1570–1582, 2005. (cited on page 35)

[Starck *et al.* 2011] Jean-Luc Starck, Fionn Murtagh et Mario Bertero. Starlet Transform in Astronomical Data Processing. In Otmar Scherzer, editeur, Handbook of Mathematical Methods in Imaging, pages 1489–1531. Springer New York, 2011. (cited on pages 71, 90, 98 et 103)

[Starck *et al.* 2015] Jean-Luc Starck, Fionn Murtagh et Jalal M Fadili. Sparse image and signal processing : Wavelets and related geometric multiscale analysis. Cambridge university press, 2015. (cited on pages 37, 69, 70 et 93)

[Starck 2016] Jean-Luc Starck. Sparsity and inverse problems in astrophysics. Journal of Physics : Conference Series, vol. 699, no. 1, page 012010, 2016. (cited on page 22)

[Sturm & Defrance 2010] Bob L Sturm et Guillaume Defrance. Detection and estimation of arrivals in room impulse responses by greedy sparse approximation. In Signal Processing Conference, 2010 18th European, pages 1934–1938. IEEE, 2010. (cited on page 32)

[Suksmono 2013] A. B. Suksmono. Reconstruction of Complex-Valued Fractional Brownian Motion Fields Based on Compressive Sampling and Its Application to PSF Interpolation in Weak Lensing Survey. ArXiv e-prints, Novembre 2013. (cited on page 109)

[Tarantola 2004] Albert Tarantola. Inverse problem theory and methods for model parameter estimation. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2004. (cited on page 22)

[Tenenbaum *et al.* 1998] Joshua B Tenenbaumet al. Mapping a manifold of perceptual observations. Advances in neural information processing systems, pages 682–688, 1998. (cited on page 51)

[Tenenbaum *et al.* 2000] Joshua B Tenenbaum, Vin De Silva et John C Langford. A global geometric framework for nonlinear dimensionality reduction. science, vol. 290, no. 5500, pages 2319–2323, 2000. (cited on pages 48, 115 et 116)

[Tikhonov & Arsenin 1977] A.N. Tikhonov et V.I.A. Arsenin. Solutions of ill-posed problems. Scripta series in mathematics. Winston, 1977. (cited on page 26)

[Torgerson 1952] W. S. Torgerson. Multidimensional scaling : I. Theory and method. Psychometrika, vol. 17, pages 401–419, 1952. (cited on page 47)

[Varoquaux *et al.* 2011] Gaël Varoquaux, Alexandre Gramfort, Fabian Pedregosa, Vincent Michel et Bertrand Thirion. Multi-subject dictionary learning to segment an atlas of brain spontaneous activity. In Biennial International Conference on Information Processing in Medical Imaging, pages 562–573. Springer Berlin Heidelberg, 2011. (cited on page 38)

[Villani 2003] Cédric Villani. Topics in optimal transportation. Numeéro 58. American Mathematical Soc., 2003. (cited on pages 54, 55, 56 et 57)

[Villani 2009] Cédric Villani. Optimal transport : old and new. Grundlehren der mathematischen Wissenschaften. Springer, Berlin, 2009. (cited on page 112)

[Wang & Huang 2015] W.-T. Wang et H.-C. Huang. Regularized Principal Component Analysis for Spatial Data. ArXiv e-prints, Janvier 2015. (cited on page 82)

[Wang & Silva 1980] JY Wang et D Es Silva. Wave-front interpretation with Zernike polynomials. Applied optics, vol. 19, no. 9, pages 1510–1518, 1980. (cited on page 14)

[Wang & Zhang 2013] Y. X. Wang et Y. J. Zhang. Nonnegative Matrix Factorization : A Comprehensive Review. IEEE Transactions on Knowledge and Data Engineering, vol. 25, no. 6, pages 1336–1353, June 2013. (cited on page 40)

[Widrow & Kollár 2008] B. Widrow et I. Kollár. Quantization noise : Roundoff error in digital computation, signal processing, control, and communications. Cambridge University Press, Cambridge, UK, 2008. (cited on page 12)

[Wong *et al.* 2015] Alexander Wong, Xiao Yu Wang et Maud Gorbet. Bayesian-based deconvolution fluorescence microscopy using dynamically updated nonstationary expectation estimates. Scientific reports, vol. 5, 2015. (cited on page 28)

[Yamagishi *et al.* 2012] M. Yamagishi, S. Ono et I. Yamada. Two variants of alternating direction method of multipliers without certain inner iterations and their application to image super-resolution. In Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on, pages 3661–3664, March 2012. (cited on page 66)

[Yang *et al.* 2009] Junfeng Yang, Yin Zhang et Wotao Yin. An efficient TVL1 algorithm for deblurring multichannel images corrupted by impulsive noise. SIAM Journal on Scientific Computing, vol. 31, no. 4, pages 2842–2865, 2009. (cited on page 25)

[Zhang *et al.* 2015] Zheng Zhang, Yong Xu, Jian Yang, Xuelong Li et David Zhang. A survey of sparse representation : algorithms and applications. IEEE Access, vol. 3, pages 490–530, 2015. (cited on page 37)

[Zheng *et al.* 2015] Le Zheng, Arian Maleki, Xiaodong Wang et Teng Long. Does ℓ_p-minimization outperform ℓ_1-minimization ? arXiv preprint arXiv :1501.03704, 2015. (cited on page 35)

[Zhu *et al.* 2007] Lei Zhu, Yan Yang, Steven Haker et Allen Tannenbaum. An image morphing technique based on optimal mass preserving mapping. IEEE Transactions on Image Processing, vol. 16, no. 6, pages 1481–1495, 2007. (cited on page 141)

[Zwicky 1937] Fritz Zwicky. On the Masses of Nebulae and of Clusters of Nebulae. The Astrophysical Journal, vol. 86, page 217, 1937. (cited on page 2)

université PARIS-SACLAY
ÉCOLE DOCTORALE
Astronomie et Astrophysique d'île-de-France

Titre : Méthodes et algorithmes avancés pour l'imagerie astronomique de haute précision

Mots clefs : Parcimonie, Fonction d'étalement du point, Transport Optimal, Problèmes Inverses, Factorisation de matrice, Apprentissage de variétés

Résumé : L'un des challenges majeurs de la cosmologie moderne réside en la nature même de la matière et de l'énergie noire. La matière noire peut être directement tracée à travers son effet gravitationnel sur les formes des galaxies. La mission Euclid de l'Agence Spatiale Européenne fournira précisément des données à cette fin. L'exploitation de telles données requiert une modélisation précise de la Fonction d'Étalement du Point (FEP) de l'instrument d'observation, ce qui constitue l'objectif de cette thèse.

Nous avons développé des méthodes non-paramétriques permettant d'estimer de manière fiable la FEP sur l'ensemble du champ de vue d'un instrument, à partir d'images non résolues d'étoiles, ceci en tenant compte du bruit, d'un possible sous-échantillonnage des observations et de la variabilité spatiale de la FEP. Ce travail tire avantage d'outils et concepts mathématiques modernes parmi lesquelles la parcimonie. Une extension importante de ce travail serait de prendre en compte la dépendance en longueur d'onde de la FEP.

Title : Advanced methods and algorithms for high precision astronomical imaging

Keywords : Sparsity, Point Spread Function, Optimal Transport, Inverse Problems, Matrix Factorization, Manifold learning

Abstract : One of the biggest challenges of modern cosmology is to gain a more precise knowledge of the dark energy and the dark matter nature. Fortunately, the dark matter can be traced directly through its gravitational effect on galaxies shapes. The European Spatial Agency Euclid mission will precisely provide data for such a purpose. A critical step is analyzing these data will be to accurately model the instrument Point Spread Function (PSF), which the focus of this thesis.

We developed non parametric methods to reliably estimate the PSFs across an instrument field-of-view, based on unresolved stars images and accounting for noise, undersampling and PSFs spatial variability. At the core of these contributions, modern mathematical tools and concepts such as sparsity. An important extension of this work will be to account for the PSFs wavelength dependency.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France