MIRA: an effective imaging algorithm for optical interferometry

Eric Thiébaut[a]

[a]Centre de Recherche Astrophysique de Lyon, CNRS/UMR 5574, France.

ABSTRACT

This paper presents MIRA, a Multi-aperture Image Reconstruction Algorithm, which has been specifically developed for image restoration from optical interferometric data. The sought image satisfies agreement with the input interferometric data and with some a priori image properties (positivity, normalization and regularization). The algorithm can cope with very limited amount of data; as an extreme case, MIRA is able to restore images without any Fourier phase information. This leads to the possibility to perform imaging with only 2 telescopes or when the phase closures are corrupted.

Keywords: optical interferometry; image reconstruction; inverse problem.

1. INTRODUCTION

Optical interferometers (VLTI, IOTA, *etc.*) yield the best angular resolution in the visible and infrared. The measures provided by these instruments are however not directly images and reconstruction algorithms are mandatory tools to fully exploit their high angular resolution imaging capabilities. Image reconstruction in radio-interferometry has a long history and is now well under control.[1,2] But, owing to the specifics of optical interferometry, the methods developed for radio-interferometry are not directly usable and we need to develop new ones. The reasons of additional difficulties in optical interferometry are a much sparser sampling of the spatial frequencies, the so-called u-v plane, and the loss of most of the Fourier phase information.

An interferometer samples the u-v plane at discrete spatial frequencies given by:

$$\boldsymbol{\nu}_{j,k}(t) = \boldsymbol{b}_{j,k}(t)/\lambda \tag{1}$$

where λ is the wavelength and $\boldsymbol{b}_{j,k}(t)$ is the separation, at time t and projected onto the sky, between the j-th and k-th interfering telescopes. Ideally, an interferometer measures the complex visibility $\hat{x}(\boldsymbol{\nu})$ which is the Fourier transform of the normalized distribution of intensity $x(\boldsymbol{a})$ of the observed object in angular direction $\boldsymbol{a}$. Hence, the observed complex visibilities are:

$$\hat{x}_{j,k}(t) = \hat{x}\big(\boldsymbol{b}_{j,k}(t)/\lambda\big)\,. \tag{2}$$

For N telescopes (or antennae) combined simultaneously, the maximum number of different spatial frequencies simultaneously measured by an interferometer is $N\,(N-1)/2$. With usually at least a dozen radio antennae against only 3 or 4 optical telescopes, the radio interferometers have an overwhelming advantage. To rebuild an image one must properly interpolate the missing spatial frequencies — which are *holes* in the coverage of the u-v plane; this is the purpose of *a priori* constraints such as positivity, regularization and, possibly, support imposed on the sought image.

In addition to the sparsity of interferometric data, atmospheric turbulence is responsible for a random phase shift between the telescopes. At optical wavelengths and in the absence of a phase reference source, the resulting phase piston errors are difficult (or impossible) to measure and, a fortiori, to compensate for. To overcome random delays due to turbulence, astronomers must integrate interferometric measurements that are insensitive to this defect. The *powerspectrum* and the *bispectrum* are such estimators. The sampled powerspectrum is:

$$\hat{x}^{(2)}_{j,k}(t) = \big|\hat{x}\big(\boldsymbol{b}_{j,k}(t)/\lambda\big)\big|^2\,, \tag{3}$$

Further author information: thiebaut@obs.univ-lyon1.fr

which only needs two telescopes (here j and k), and provides no information about the phase of the complex visibility. The bispectrum is the triple product of the complex visibilities measured by the interferences from three telescopes:

$$\hat{x}^{(3)}_{j,k,\ell}(t) = \hat{x}\big(\boldsymbol{b}_{j,k}(t)/\lambda\big)\,\hat{x}\big(\boldsymbol{b}_{k,\ell}(t)/\lambda\big)\,\hat{x}\big(\boldsymbol{b}_{\ell,j}(t)/\lambda\big)\,, \tag{4}$$

where j, k, and ℓ are the indices of the involved telescopes. The phase of the bispectrum is the so-called *phase closure*:

$$\beta_{j,k,\ell}(t) \stackrel{\text{def}}{=} \arg\big\{\hat{x}^{(3)}_{j,k,\ell}(t)\big\} = \phi_{j,k}(t) + \phi_{k,\ell}(t) + \phi_{\ell,j}(t) \mod 2\,\pi\,, \tag{5}$$

where ϕ is the phase of the complex visibility:

$$\phi_{j,k}(t) \stackrel{\text{def}}{=} \arg\big\{\hat{x}_{j,k}(t)\big\}\,. \tag{6}$$

Hence the bispectrum provides some Fourier phase information from the three spatial frequencies $\boldsymbol{\nu}_{j,k}(t)$, $\boldsymbol{\nu}_{k,\ell}(t)$, and $\boldsymbol{\nu}_{\ell,j}(t) = -\boldsymbol{\nu}_{j,k}(t) - \boldsymbol{\nu}_{k,\ell}(t)$ which form a closed triangle in the u-v plane. However, in the case of a 3-telescope interferometer, the phase closure data provide only a single phase out of three spatial frequencies. Moreover this is at the price of increased complexity for the data processing (non-linearity) and for the operation of the instrument since it requires to make at least three telescopes to interfere.

The notation used here is intended to explicit the dependence of the data with the interfering telescopes and with the time (*i.e.* orientation with respect to observed object due to earth rotation). In the remaining of the paper, for sake of simplicity, the list $\boldsymbol{\mathcal{L}}_{\boldsymbol{\nu}}$ of sampled spatial frequencies may be indexed by a single index:

$$\boldsymbol{\mathcal{L}}_{\boldsymbol{\nu}} \stackrel{\text{def}}{=} \{\boldsymbol{\nu}_k; k = 1, \ldots, m\} = \{\boldsymbol{b}_{k,\ell}(t)/\lambda; \forall(k,\ell,t)\}\,, \tag{7}$$

where $\forall(k,\ell,t)$ formally means: *for all baselines and exposures used during the observations.* To simplify the equations to come, we also introduce the following notations for the complex visibilities and squared visibilities:

$$v_{j,k}(t) \stackrel{\text{def}}{=} \hat{x}_{j,k}(t)\,, \tag{8}$$

$$s_{j,k}(t) \stackrel{\text{def}}{=} \hat{x}^{(2)}_{j,k}(t)\,. \tag{9}$$

2. IMAGE AND COMPLEX VISIBILITY MODELS

The result of the image reconstruction is the distribution of intensity $x(\boldsymbol{a})$ across the field of view Ω in the angular direction $\boldsymbol{a}$. A practical mean to parametrize this distribution is to use a basis of functions $\{z_j : \Omega \mapsto \mathbb{R}; j = 1, \ldots, n\}$ and to approximate the brightness distribution by a linear expansion:

$$x(\boldsymbol{a}) \simeq x_{\mathsf{model}}(\boldsymbol{a}) = \sum_{j=1}^{n} x_j\, z_j(\boldsymbol{a}) \quad \stackrel{\text{FT}}{\longrightarrow} \quad \hat{x}(\boldsymbol{a}) \simeq \hat{x}_{\mathsf{model}}(\boldsymbol{\nu}) = \sum_{j=1}^{n} x_j\, \hat{z}_j(\boldsymbol{\nu}) \tag{10}$$

where $\boldsymbol{x} \in \mathbb{R}^n$ are the parameters of the model image, $\hat{x}_{\mathsf{model}}(\boldsymbol{\nu})$ and $\hat{z}_j(\boldsymbol{\nu})$ are the Fourier transforms of the model image and of the j-th basis function at spatial frequency $\boldsymbol{\nu}$. This description is very general, for example it can be used to have a multi-resolution model of the image or to account for point sources over a diffuse background.[3] Another common description is to use the same function $z(\Delta\boldsymbol{a})$ on an evenly spaced grid of angular directions $\mathcal{G}_{\boldsymbol{a}} = \{\boldsymbol{\dot{a}}_j; j = 1, \ldots, n\}$, then:

$$x_{\mathsf{model}}(\boldsymbol{a}) = \sum_{j=1}^{n} x_j\, z(\boldsymbol{a} - \boldsymbol{\dot{a}}_j) \quad \stackrel{\text{FT}}{\longrightarrow} \quad \hat{x}_{\mathsf{model}}(\boldsymbol{\nu}) = \hat{z}(\boldsymbol{\nu}) \sum_{j=1}^{n} x_j\, \mathrm{e}^{-2\,\mathrm{i}\,\pi\,\boldsymbol{\dot{a}}_j\,\boldsymbol{\nu}}\,. \tag{11}$$

In this description, the basis function $z(\Delta\boldsymbol{a})$ set the effective angular resolution of the image in the manner of the *clean beam* in the CLEAN method.[4] To avoid spectral aliasing, it is necessary that Shannon criterion be respected and that the angular sampling step δa obeys:

$$\delta a \leq \frac{\lambda}{2\,b_{\mathsf{max}}}\,, \tag{12}$$

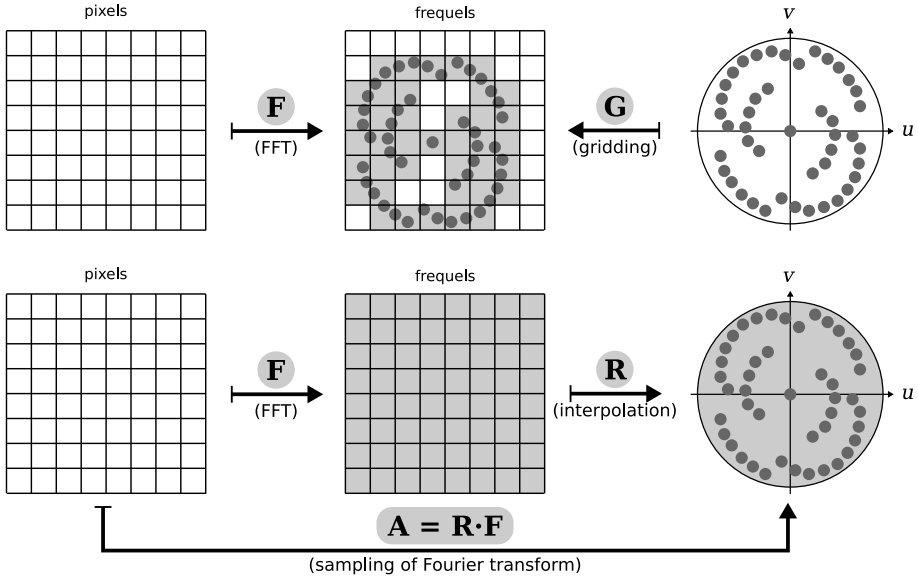

Figure 1. Rebinning (top) and interpolation (bottom) of spatial frequencies. *Top:* the data at sampled frequencies are rebinned by operator **G** to match the grid of frequels. *Bottom:* model of complex visibilities is interpolated at sampled spatial frequencies by operator **R**.

where b_{max} is the maximum length of the projected interferometric bases. Some approaches, as the *building blocks* method[5] or WIPE,[6] use explicit expressions for $z(\Delta\boldsymbol{a})$. Finally choosing any interpolation function as the basis function $z(\Delta\boldsymbol{a})$ makes the model in Eq. (11) the same as most other algorithms for which the parameters are the intensities of the pixels of the image: $x_j = x(\overset{\square}{\boldsymbol{a}}_j)$.

The image models considered in MIRA have exact Fourier transforms. Using matrix notation, the model of the complex visibility is:

$$\boldsymbol{v}_{\mathsf{model}} = \mathbf{A} \cdot \boldsymbol{x} \tag{13}$$

where $\boldsymbol{x} \in \mathbb{R}^n$ are the sought image parameters and $\boldsymbol{v}_{\mathsf{model}} \in \mathbb{C}^m$ assuming there are m sampled spatial frequencies. In words, $v_k^{\mathsf{model}} \simeq \hat{x}(\boldsymbol{\nu}_k)$ is the model of the complex visibility for k-th sampled spatial frequency. The complex coefficients of matrix $\mathbf{A}$ depend on the image model:

$$A_{k,j} = \hat{z}_j(\boldsymbol{\nu}_k) \tag{14}$$

$$= \hat{z}(\boldsymbol{\nu}_k)\, \mathrm{e}^{-2\,\mathrm{i}\,\pi\, \overset{\square}{\boldsymbol{a}}_j\, \boldsymbol{\nu}_k} \tag{15}$$

where expression in Eq. (14) results from the model in Eq. (10) while the model in Eq. (11) yields Eq. (15).

The *exact* linear transform can be used in the image restoration algorithm but may become computationally untractable when the size of $\mathbf{A}$, that is $m \times n$, becomes too large. The Fourier transform of Eq. (11) may be approximated by a discrete Fourier transform which, under an additional circulant approximation, is efficiently computed by means of a fast Fourier transform (FFT). In this case, as discrete spatial frequencies — the so-called *frequels* — do not necessarily coincide with those measured, it is necessary to interpolate the Fourier spectrum of the image.[2]

In radio interferometry, the so-called *gridding* technique[7] consists in interpolating the complex visibility data onto the rectangular grid of frequels $\mathcal{G}_{\boldsymbol{\nu}} = \{\overset{\square}{\boldsymbol{\nu}}_k; k = 1, \ldots, n\}$ that correspond to spatial frequencies after fast Fourier transform of the pixel image representation. After *gridding*, the re-sampled data write:

$$\overset{\square}{\boldsymbol{v}}_{\mathsf{data}} = \mathbf{G} \cdot \boldsymbol{v}_{\mathsf{data}} \tag{16}$$

where $\boldsymbol{v}_{\mathsf{data}} \in \mathbb{C}^m$ are the original complex visibility data. The linear operator $\mathbf{G}$ accounts for the combination of different operations:[7] local *averaging* which depends on the actual density of the u-v coverage, *weighting* to

account for the reliability of the data, and *tapering* to set the resolution. The model of the re-sampled complex visibility data is then:

$$\overset{\square}{\boldsymbol{v}}_{\mathsf{model}} = \mathbf{T} \cdot \hat{\boldsymbol{x}} = \mathbf{T} \cdot \mathbf{F} \cdot \boldsymbol{x} \tag{17}$$

where $\mathbf{T}$ is a truncation (or sampling) operator with all coefficients equals to 0 or 1 and which discards the discrete frequencies outside the u-v coverage (the grayed area in the central part of top of Fig. 1) and $\hat{\boldsymbol{x}}$ is the FFT of the pixel map $\boldsymbol{x}$:

$$\hat{\boldsymbol{x}} = \mathbf{F} \cdot \boldsymbol{x} \quad \text{with} \quad F_{k,j} = \exp\bigl(-2\,\mathrm{i}\,\pi\ \overset{\square}{\boldsymbol{a}}_j\,\overset{\square}{\boldsymbol{\nu}}_k\bigr)\,. \tag{18}$$

With this definition of the FFT operator $\mathbf{F}$, the inverse transform writes:

$$\bigl[\mathbf{F}^{-1}\bigr]_{j,k} = \frac{1}{n}\,\exp\bigl(+2\,\mathrm{i}\,\pi\ \overset{\square}{\boldsymbol{a}}_j\,\overset{\square}{\boldsymbol{\nu}}_k\bigr) \tag{19}$$

where $n = n_1 \times n_2$ is the total number of pixels in the image, n_1 and n_2 being the number of pixels along the two axis (usually $n_1 = n_2$). Note that:

$$\mathbf{F}^{-1} \cdot \mathbf{T}^{\mathrm{T}} \cdot \overset{\square}{\boldsymbol{v}}_{\mathsf{data}}$$

is the so-called *dirty* image equals to the principal solution[8] of solving $\overset{\square}{\boldsymbol{v}}_{\mathsf{model}} = \overset{\square}{\boldsymbol{v}}_{\mathsf{data}}$ without additional constraints.

In optical interferometry, complex visibilities are usually not available and most data consist in non-linear quantities (powerspectrum, bispectrum or phase closure) which cannot be linearly remapped onto the frequel grid. The *gridding* of the data is not possible, but the discrete Fourier transform of the image can be interpolated at the frequencies sampled by the data. To benefit from FFT speedup, MIRA makes use of an approximation of $\mathbf{A}$:

$$\mathbf{A} \simeq \mathbf{R} \cdot \mathbf{F} \tag{20}$$

where $\mathbf{F}$ is the FFT matrix and $\mathbf{R}$ is a linear interpolation matrix which performs interpolation of the model FFT $\hat{\boldsymbol{x}}$ at measured spatial frequencies. This interpolation can be chosen so that $\mathbf{R}$ is a very sparse matrix which is fast to apply and has light memory footprint. In MIRA, we use complex bilinear interpolation, thus the model of the complex visibility at each measured spatial frequency is a linear combination of 4 neighbors in $\hat{\boldsymbol{x}}$. The number of floating point operations to apply $\mathbf{R}$ scales as $\sim 16\,m$.

3. INVERSE PROBLEM APPROACH

Once chosen the parametrization, image reconstruction can be seen as an inverse problem.[9,10] Adopting a Bayesian viewpoint, the best parameters $\boldsymbol{x}^+$ can be chosen as the most likely ones given the data $\boldsymbol{d}$ (complex visibilities, powerspectrum, phase closure, *etc.*):

$$\boldsymbol{x}^+ = \arg\max_{\boldsymbol{x}} \Pr(\boldsymbol{x}|\boldsymbol{d}) = \arg\max_{\boldsymbol{x}} \Pr(\boldsymbol{d}|\boldsymbol{x})\ \Pr(\boldsymbol{x})\,, \tag{21}$$

where the last expression comes from Bayes' theorem and after having discarded the term $\Pr(\boldsymbol{d})$ which does not depend on the sought parameters. This choice is termed the *maximum a posteriori* (MAP) solution and can be recast as:

$$\boldsymbol{x}^+ = \arg\min_{\boldsymbol{x}} f(\boldsymbol{x}) \tag{22}$$

where the penalty function $f(\boldsymbol{x})$ writes:

$$f(\boldsymbol{x}) = c_0 - c_1\ \log\Pr(\boldsymbol{x}|\boldsymbol{d}) = c_0 \underbrace{-c_1\ \log\Pr(\boldsymbol{d}|\boldsymbol{x})}_{f_{\mathsf{data}}(\boldsymbol{x})} \underbrace{-c_1\ \log\Pr(\boldsymbol{x})}_{f_{\mathsf{prior}}(\boldsymbol{x})} \tag{23}$$

where c_0 and $c_1 > 0$ are two real constants chosen for convenience. This equation shows that, to find the maximum a posteriori solution, we must minimize a joint criterion which is the sum of two terms: a likelihood term $f_{\mathsf{data}}(\boldsymbol{x}) \propto -\log\Pr(\boldsymbol{d}|\boldsymbol{x})$ which measures the compatibility of the parameters with the data, and a regularization term $f_{\mathsf{prior}}(\boldsymbol{x}) \propto -\log\Pr(\boldsymbol{x})$ which imposes priors on the image.

The expression of $f_{\mathsf{data}}(\boldsymbol{x})$ is derived from some approximations of the statistics of the data noise as explained in Sect. 4. In practice, the prior statistics is however seldom known and the regularization must be derived from some heuristics as we know discuss. The two main reasons to introduce the regularization in an inverse problem are: (i) to provide additional information when the data alone cannot completely define a unique solution, and (ii) to counter the amplification of noise for a poorly conditioned problem.[10] The purpose of regularization is then to select among all solutions compatible with the measures, whichever is closer to some priors since the data alone are insufficient to provide a satisfactory solution. That is both stable (robust with respect to the noise) and unique. Formally, this quest can be expressed as:

$$\boldsymbol{x}^+ = \arg\min_{\boldsymbol{x}} f_{\mathsf{prior}}(\boldsymbol{x}) \quad \text{s.t.} \quad f_{\mathsf{data}}(\boldsymbol{x}) \leq \eta \tag{24}$$

where the inequality constraint $f_{\mathsf{data}}(\boldsymbol{x}) \leq \eta$ imposes that the model be compatible with the data: the lower is η the closer will be the model and the measurements. Assuming that the inequality constraint is active (otherwise the measurements are meaningless), the solution of the constrained minimization problem writes:

$$\boldsymbol{x}^+ = \arg\min_{\boldsymbol{x}} \left[f_{\mathsf{prior}}(\boldsymbol{x}) + \ell \, f_{\mathsf{data}}(\boldsymbol{x}) \right] \tag{25}$$

where $\ell > 0$ is a Lagrange multiplier that have to be tuned so that $f_{\mathsf{data}}(\boldsymbol{x}^+) \simeq \eta$.

To be an effective regularization, f_{prior} must have certain properties. In the case of image restoration in interferometry, voids in the sampling of the u-v plane imply that the problem is most often ill-posed: solely maximizing the likelihood of the model — that is minimizing $f_{\mathsf{data}}(\boldsymbol{x})$ — admits an infinite number of solutions. In this context, regularization should help filling the lack of data at unsampled spatial frequencies, it is therefore natural to require that regularization yields somewhat smooth interpolation between measured spatial frequencies. This kind of spatial frequency smoothing can be achieved by imposing a simple quadratic constraint in the image (see Appendix A) or by using one of the maximum entropy methods[1,2] which have proved their worth in radio astronomy. Also note that *building-block* methods[4,5] yield a regularized solution by requiring that it has only a limited number of significant components.

4. DATA PENALTY

In its current form, MIRA can account for three different kind of interferometric measurements: complex visibilities $\boldsymbol{v}_{\mathsf{data}}$, powerspectrum $\boldsymbol{s}_{\mathsf{data}}$ and phase closure $\boldsymbol{\beta}_{\mathsf{data}}$. Hence, the data penalty $f_{\mathsf{data}}(\boldsymbol{x})$ to fit these data writes:

$$f_{\mathsf{data}}(\boldsymbol{x}) = f_{\boldsymbol{v}}(\boldsymbol{x}) + f_{\boldsymbol{s}}(\boldsymbol{x}) + f_{\boldsymbol{\beta}}(\boldsymbol{x}) \,, \tag{26}$$

where $f_{\boldsymbol{v}}(\boldsymbol{x})$, $f_{\boldsymbol{s}}(\boldsymbol{x})$ and $f_{\boldsymbol{\beta}}(\boldsymbol{x})$ are respectively penalty terms with respect to complex visibility data, powerspectrum data and phase closure data. Note that this definition assumes that data of different types are uncorrelated.

4.1 Complex Visibility Data

In MIRA it is assumed that interferometric data are independent and that complex visibility errors have Gaussian distribution. The penalty with respect to complex visibility data then writes:

$$f_{\boldsymbol{v}}(\boldsymbol{x}) = \sum_t \sum_{k<\ell} \begin{pmatrix} \mathrm{Re}\big(v^{\mathsf{res}}_{k,\ell}(\boldsymbol{x},t)\big) \\ \mathrm{Im}\big(v^{\mathsf{res}}_{k,\ell}(\boldsymbol{x},t)\big) \end{pmatrix}^{\mathrm{T}} \cdot \begin{pmatrix} W^{\mathsf{rr}}_{k,\ell}(t) & W^{\mathsf{ri}}_{k,\ell}(t) \\ W^{\mathsf{ri}}_{k,\ell}(t) & W^{\mathsf{ii}}_{k,\ell}(t) \end{pmatrix} \cdot \begin{pmatrix} \mathrm{Re}\big(v^{\mathsf{res}}_{k,\ell}(\boldsymbol{x},t)\big) \\ \mathrm{Im}\big(v^{\mathsf{res}}_{k,\ell}(\boldsymbol{x},t)\big) \end{pmatrix} . \tag{27}$$

where the complex visibility residuals are:

$$v^{\mathsf{res}}_{k,\ell}(\boldsymbol{x},t) \stackrel{\mathrm{def}}{=} v^{\mathsf{data}}_{k,\ell}(t) - v^{\mathsf{model}}_{k,\ell}(\boldsymbol{x},t) \tag{28}$$

and the weights:

$$W^{\mathsf{rr}}_{k,\ell}(t) = \left[C^{\mathsf{rr}}_{k,\ell}(t)\, C^{\mathsf{ii}}_{k,\ell}(t) - C^{\mathsf{ri}}_{k,\ell}(t)^2\right]^{-1} C^{\mathsf{ii}}_{k,\ell}(t) \,, \tag{29}$$

$$W^{\mathsf{ri}}_{k,\ell}(t) = -\left[C^{\mathsf{rr}}_{k,\ell}(t)\, C^{\mathsf{ii}}_{k,\ell}(t) - C^{\mathsf{ri}}_{k,\ell}(t)^2\right]^{-1} C^{\mathsf{ri}}_{k,\ell}(t) \,, \tag{30}$$

$$W^{\mathsf{ii}}_{k,\ell}(t) = \left[C^{\mathsf{rr}}_{k,\ell}(t)\, C^{\mathsf{ii}}_{k,\ell}(t) - C^{\mathsf{ri}}_{k,\ell}(t)^2\right]^{-1} C^{\mathsf{rr}}_{k,\ell}(t) \,, \tag{31}$$

are computed from the covariances of the complex visibility data:

$$C^{\mathsf{rr}}_{k,\ell}(t) = \mathrm{Var}\left\{\mathrm{Re}\big(v^{\mathsf{data}}_{k,\ell}(t)\big)\right\}\,, \tag{32}$$

$$C^{\mathsf{ri}}_{k,\ell}(t) = \mathrm{Cov}\left\{\mathrm{Re}\big(v^{\mathsf{data}}_{k,\ell}(t)\big), \mathrm{Im}\big(v^{\mathsf{data}}_{k,\ell}(t)\big)\right\}\,, \tag{33}$$

$$C^{\mathsf{ii}}_{k,\ell}(t) = \mathrm{Var}\left\{\mathrm{Im}\big(v^{\mathsf{data}}_{k,\ell}(t)\big)\right\}\,. \tag{34}$$

For interferometric data, the so-called Goodman model can be a good approximation of the distribution of complex visibility errors. In this case, the real and imaginary parts of the complex visibility are independant, hence $C^{\mathsf{ri}}_{k,\ell}(t) = 0$, and have the same variance; the penalty then simplifies to:

$$f_{\boldsymbol{v}}(\boldsymbol{x}) = \sum_t \sum_{k<\ell} \frac{\left|v^{\mathsf{data}}_{k,\ell}(t) - v^{\mathsf{model}}_{k,\ell}(\boldsymbol{x},t)\right|^2}{\mathrm{Var}\big\{|v^{\mathsf{data}}_{k,\ell}(t)|\big\}}\,, \tag{35}$$

where $C^{\mathsf{rr}}_{k,\ell}(t) = C^{\mathsf{ii}}_{k,\ell}(t) = \mathrm{Var}\big\{|v^{\mathsf{data}}_{k,\ell}(t)|\big\}$.

OI-FITS[11] is a file format which has became the standard for the storage and exchange of optical interferometry data. This format is very versatile but has a number of restrictions when processing OI-FITS data. The data statistics is only provided by the standard deviation of the measurements, hence there is no means to account for correlations. Since complex visibilities and bispectrum are provided in polar form, this means that the amplitude and phase of complex data are independent. Simple geometrical considerations, yield a quadratic approximation similar to that in Eq. (27) but with weights:

$$W^{\mathsf{rr}}_{k,\ell}(t) = \frac{\cos^2\phi^{\mathsf{data}}_{k,\ell}(t)}{\mathrm{Var}\{\rho^{\mathsf{data}}_{k,\ell}(t)\}} + \frac{\sin^2\phi^{\mathsf{data}}_{k,\ell}(t)}{\rho^{\mathsf{data}}_{k,\ell}(t)^2\ \mathrm{Var}\{\phi^{\mathsf{data}}_{k,\ell}(t)\}}\,, \tag{36}$$

$$W^{\mathsf{ri}}_{k,\ell}(t) = \left(\frac{1}{\mathrm{Var}\{\rho^{\mathsf{data}}_{k,\ell}(t)\}} - \frac{1}{\rho^{\mathsf{data}}_{k,\ell}(t)^2\ \mathrm{Var}\{\phi^{\mathsf{data}}_{k,\ell}(t)\}}\right) \cos\phi^{\mathsf{data}}_{k,\ell}(t)\ \sin\phi^{\mathsf{data}}_{k,\ell}(t)\,, \tag{37}$$

$$W^{\mathsf{ii}}_{k,\ell}(t) = \frac{\sin^2\phi^{\mathsf{data}}_{k,\ell}(t)}{\mathrm{Var}\{\rho^{\mathsf{data}}_{k,\ell}(t)\}} + \frac{\cos^2\phi^{\mathsf{data}}_{k,\ell}(t)}{\rho^{\mathsf{data}}_{k,\ell}(t)^2\ \mathrm{Var}\{\phi^{\mathsf{data}}_{k,\ell}(t)\}}\,, \tag{38}$$

and with residuals computed for:

$$v^{\mathsf{data}}_{k,\ell}(t) = \rho^{\mathsf{data}}_{k,\ell}(t)\ \exp(\mathrm{i}\,\phi^{\mathsf{data}}_{k,\ell}(t))\,, \tag{39}$$

where $\rho^{\mathsf{data}}_{k,\ell}(t)$ and $\phi^{\mathsf{data}}_{k,\ell}(t)$ are the amplitude and phase of the measured complex visibilities. Convex approximations of the likelihood penalty have also been studied by Meimon *et al.*[12]

4.2 Powerspectrum Data

Assuming normally distributed errors for the powerspectrum, the term $f_{\boldsymbol{s}}(\boldsymbol{x})$ in Eq. (26) write:

$$f_{\boldsymbol{s}}(\boldsymbol{x}) = \sum_t \sum_{k<\ell} \frac{\left(s^{\mathsf{data}}_{k,\ell}(t) - s^{\mathsf{model}}_{k,\ell}(\boldsymbol{x},t)\right)^2}{\mathrm{Var}[s^{\mathsf{data}}_{k,\ell}(t)]}\,. \tag{40}$$

4.3 Phase Closure Data

In order to account for phase wrapping and to avoid excessive non-linearity, the term related to the phase closures data is defined by Mira to be the weighted quadratic distance between the complex phasors rather than between the phases closures:

$$f_{\boldsymbol{\beta}}(\boldsymbol{x}) = \sum_t \sum_{j<k<\ell} \frac{1}{\mathrm{Var}[\beta^{\mathsf{data}}_{j,k,\ell}(t)]} \left|\mathrm{e}^{\mathrm{i}\,\beta^{\mathsf{data}}_{j,k,\ell}(t)} - \mathrm{e}^{\mathrm{i}\,\beta^{\mathsf{model}}_{j,k,\ell}(\boldsymbol{x},t)}\right|^2\,. \tag{41}$$

In the limit of small phase closure errors, the penalty becomes:

$$f_{\boldsymbol{\beta}}(\boldsymbol{x}) \simeq \sum_t \sum_{j<k<\ell} \frac{\left[\beta^{\mathsf{data}}_{j,k,\ell}(t) - \beta^{\mathsf{model}}_{j,k,\ell}(\boldsymbol{x},t)\right]^2}{\mathrm{Var}[\beta^{\mathsf{data}}_{j,k,\ell}(t)]} \tag{42}$$

which is readily the χ^2 term that would be obtained for Gaussian phase statistics. This justifies the weighting used in Eq. (41). Other methods have been proposed to cope with the phase wrapping[13,14] but we have found that, in practice, they can slow down or prevent the convergence of the algorithm.

5. REGULARIZATION

MIRA has been designed to be versatile in terms of input data and type of regularization. The YORICK version of MIRA let the user define its own regularization to match his priors. A number of different regularizations are already built into MIRA which are summarized in what follows.

5.1 Quadratic Regularizations

MIRA implements a generic quadratic regularization:

$$f_{\mathsf{prior}}(\boldsymbol{x}) = (\mathbf{A}_{\mathsf{prior}} \cdot \boldsymbol{x} - \boldsymbol{b}_{\mathsf{prior}})^{\mathrm{T}} \cdot \mathbf{W}_{\mathsf{prior}} \cdot (\mathbf{A}_{\mathsf{prior}} \cdot \boldsymbol{x} - \boldsymbol{b}_{\mathsf{prior}}) \ , \tag{43}$$

where the matrices $\mathbf{W}_{\mathsf{prior}}$ and $\mathbf{A}_{\mathsf{prior}}$ and the vector $\boldsymbol{b}_{\mathsf{prior}}$ can be chosen to reproduce any quadratic regularization. The weighting matrix $\mathbf{W}_{\mathsf{prior}}$ must be symmetrical and positive semi-definite for $f_{\mathsf{prior}}(\boldsymbol{x})$ to be convex.

For instance, taking $\mathbf{W}_{\mathsf{prior}} = \mathbf{A}_{\mathsf{prior}} = \mathbf{I}$ and $\boldsymbol{b}_{\mathsf{prior}} = \mathbf{0}$ yields Tikhonov's regularization which is the most simple quadratic one:

$$f_{\mathsf{prior}}(\boldsymbol{x}) = \sum_j x_j^2 = \|\boldsymbol{x}\|_2^2 \ , \tag{44}$$

its effects are to limit the number of significant pixels (although not as well as with an ℓ_1 norm) and to introduce some sort of smoothness.

In a Bayesian framework and assuming Gaussian statistics for the priors, the expected value $\boldsymbol{x}_{\mathsf{prior}} = \langle \boldsymbol{x} \rangle$ and covariance matrix $\mathbf{C}_{\mathsf{prior}} = \langle (\boldsymbol{x} - \boldsymbol{x}_{\mathsf{prior}}) \cdot (\boldsymbol{x} - \boldsymbol{x}_{\mathsf{prior}})^{\mathrm{T}} \rangle$ are assumed to be known. The corresponding regularization is quadratic and writes:

$$\mu\, f_{\mathsf{prior}}(\boldsymbol{x}) = (\boldsymbol{x} - \boldsymbol{x}_{\mathsf{prior}})^{\mathrm{T}} \cdot \mathbf{C}_{\mathsf{prior}}^{-1} \cdot (\boldsymbol{x} - \boldsymbol{x}_{\mathsf{prior}}) \ , \tag{45}$$

which can be implemented thanks to the generic expression in Eq. (43).

The following quadratic regularization:

$$f_{\mathsf{prior}}(\boldsymbol{x}) = \|\mathbf{D} \cdot \boldsymbol{x}\|^2 \ , \tag{46}$$

(with $\mathbf{D}$ a finite difference operator) enforces smoothness but is mostly interesting for ill-conditioned inverse problems such as image deconvolution to avoid noise amplification. In effect, effective regularization for interferometric data should induce smooth interpolation of the missing frequencies or, similarly, limit the number of significant pixels in the field of view. To that end, we have proposed the following quadratic separable regularization for MIRA and WISARD:[15]

$$f_{\mathsf{prior}}(\boldsymbol{x}) = \sum_j x_j^2 / x_j^{\mathsf{prior}} \ , \tag{47}$$

where, under the normalization constraint $\sum_j x_j = \sum_j x_j^{\mathsf{prior}}$, the default solution is $\boldsymbol{x}_{\mathsf{prior}}$ which is chosen to be strictly positive and properly normalized as shown in Appendix A.

5.2 Maximum Entropy

MIRA implements several regularization penalties to build the *maximum entropy*[1,2] image from the interferometric data:

$$f_{\mathsf{ent1}}(\boldsymbol{x}) = -\sum\nolimits_j \sqrt{x_j} \tag{48}$$

$$f_{\mathsf{ent2}}(\boldsymbol{x}) = -\sum\nolimits_j \log(x_j) \tag{49}$$

$$f_{\mathsf{ent3}}(\boldsymbol{x}) = \sum\nolimits_j x_j \, \log(x_j) \tag{50}$$

$$f_{\mathsf{ent4}}(\boldsymbol{x}; \boldsymbol{x}_{\mathsf{prior}}) = \sum\nolimits_j x_j \, \log\left(x_j / x_j^{\mathsf{prior}}\right) \tag{51}$$

$$f_{\mathsf{ent5}}(\boldsymbol{x}; \boldsymbol{x}_{\mathsf{prior}}) = \sum\nolimits_j \left[x_j^{\mathsf{prior}} - x_j + x_j \, \log\left(x_j / x_j^{\mathsf{prior}}\right)\right] \tag{52}$$

$$f_{\mathsf{ent6}}(\boldsymbol{x}; \mathbf{S}) = \sum\nolimits_j x_j \, \log\left(x_j / (\mathbf{S} \cdot \boldsymbol{x})_j\right) \tag{53}$$

$$f_{\mathsf{ent7}}(\boldsymbol{x}; \mathbf{S}) = \sum\nolimits_j \left[(\mathbf{S} \cdot \boldsymbol{x})_j - x_j + x_j \, \log\left(x_j / (\mathbf{S} \cdot \boldsymbol{x})_j\right)\right] \tag{54}$$

where $\mathbf{S}$ is a linear operator which defines a so-called *floating prior*[16–18] $\boldsymbol{x}_{\mathsf{prior}} = \mathbf{S} \cdot \boldsymbol{x}$ that depends on $\boldsymbol{x}$. Note that if $\boldsymbol{x}$ and the default solution $\boldsymbol{x}_{\mathsf{prior}}$ are normalized to the same value, then $f_{\mathsf{ent5}}(\boldsymbol{x}; \boldsymbol{x}_{\mathsf{prior}})$ in Eq. (52) is equal to $f_{\mathsf{ent4}}(\boldsymbol{x}; \boldsymbol{x}_{\mathsf{prior}})$ in Eq. (51) and $f_{\mathsf{ent7}}(\boldsymbol{x}; \mathbf{S})$ in Eq. (54) is equal to $f_{\mathsf{ent6}}(\boldsymbol{x}; \mathbf{S})$ in Eq. (53).

5.3 Other regularizations

Quadratic regularizations are known to somewhat oversmooth the resulting images. This is particularly inconvenient for astronomical objects which have high dynamical range and high frequency contents due to point-like sources or sharp edges. To let some sharp features appear in the restored image, an *edge-preserving* regularization can be used:

$$\mu \, f_{\mathsf{prior}}(\boldsymbol{x}) = \mu \sum\nolimits_{j,k} \left(\sqrt{(\mathbf{D}_j \cdot \boldsymbol{x})_k^2 + \epsilon^2} - \epsilon\right) \tag{55}$$

where $\mathbf{D}_j$ is a finite difference linear operator approximating the partial spatial derivative of its argument along k-th direction (horizontal, vertical and, perhaps, diagonals) and $\epsilon > 0$ is a threshold. For small absolute finite differences with respect to ϵ, the regularization is approximately quadratic (ℓ_2); while for large differences, the regularization is approximately linear (ℓ_1).

For objects with a mixture of point-like structures and a smooth background, the following regularization has proved effective:

$$\mu \, f_{\mathsf{prior}}(\boldsymbol{x}) = \mu_0 \sum\nolimits_j \left(\sqrt{x_j^2 + \epsilon^2} - \epsilon\right) + \mu_1 \, \|\mathbf{D} \cdot \boldsymbol{x}\|_2^2 \tag{56}$$

where the first term (with ϵ a very small positive value) is approximately the ℓ_1 norm of the image and its effect is to limit the number of bright pixels in the image, and where the second term (with $\mathbf{D}$ a finite difference operator) enforces smoothness of the image to avoid spurious high frequencies.

6. ALGORITHM SUMMARY

The approach of algorithm MIRA[20] is to seek for the image $\boldsymbol{x} \in \mathbb{R}^n$, n being the number of *pixels*, by directly minimizing a joint criterion under constraints of positivity and normalization:

$$f(\boldsymbol{x}) = f_{\mathsf{data}}(\boldsymbol{x}) + \mu \, f_{\mathsf{prior}}(\boldsymbol{x}) \quad \text{s.t.} \quad \boldsymbol{x} \geq 0 \quad \text{and} \quad \sum\nolimits_j x_j = \xi \tag{57}$$

where f_{data} — see Eq. (26) — enforces agreement with the measurements, f_{prior} — see Sect. 5 — is a regularization term which enforces other a priori constraints and $\mu > 0$ is used to tune the weight of the priors. Taking $\mu = 1/\ell$, this definition is directly related to the Lagrangian in Eq. (25).

MIRA implements various priors such as maximum entropy, quadratic (ℓ_2) smoothness, edged-preserving ($\ell_2 - \ell_1$) smoothness, *etc.* Moreover, the algorithm is designed so that the user can plug its own regularization.

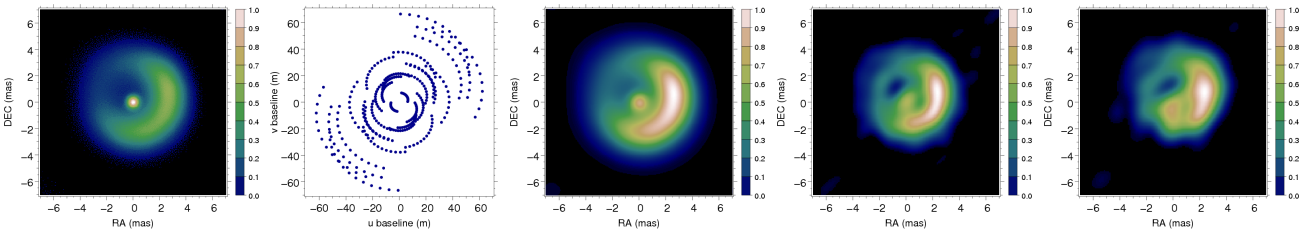

Figure 2. Image reconstruction from simulated data for the 2004 *Optical/IR Interferometry Imaging Beauty Contest*.[19] Form left to right: true image, u-v coverage, true image smoothed at the resolution of the interferometer, image restored from powerspectrum and phase closure data, image restored without any phase data. The wavelength for the simulation is $\lambda = 0.55\,\mu$m.

Practical means to automatically set the value of μ — or, equivalently, to set the value of η in Eq. (24) — have been proposed.[21–23] The most effective is the method proposed by Skilling & Bryan[24] for maximum entropy regularization and which has been adapted to other kind of regularizations.[18] In the current version of MIRA, the tuning of the hyper-parameter $\mu = 1/\ell$ is done by the user. In the reconstructions shown in Fig. 2 and Fig. 3, the regularization level is tuned so as to have $f_{\mathsf{data}}(\boldsymbol{x}) \sim m$ where m is the number of measurements.

Given the data, the regularization and its level, the criterion $f(\boldsymbol{x})$ is multimodal unless all data consist in complex visibilities. Ideally the solution should then be sought by means of a global optimization method. Owing to the large number of parameters (the number n of pixels in the image $\boldsymbol{x}$), global optimization would require unpractical amount of computation. The strategy used by MIRA is to perform only local optimization starting from an initial image. The final image obtained by MIRA therefore depends on the data and on the priors but also on the initial image and on the path followed by the local optimization method.

To minimize the criterion, MIRA uses the optimization method VMLMB,[25] a limited variable metric algorithm which accounts for parameter bounds. This last feature is used to enforce positivity of the solution. Only the value of the cost function and its gradient are needed by VMLMB. Normalization of the solution is obtained by a change of variables, the image brightness distribution becoming:

$$x_j = \frac{x'_j}{\sum_k x'_k}\,, \tag{58}$$

where $\boldsymbol{x}' \in \mathbb{R}^n$ are the variables *seen* by the optimizer with the constraints that $x'_j \geq 0, \forall j$. Thus the image $\boldsymbol{x}$ is both normalized and positive.

Examples of image reconstructions from simulated and real data are shown by Fig. 2 and Fig. 3. The two rightmost images in Fig. 2 were restored with MIRA from data simulated for the first *Optical/IR Interferometry Imaging Beauty Contest*;[19] note that the last image was recovered without any phase information demonstrating the ability of MIRA to cope with phaseless data. The rightmost image in Fig. 3 was recovered by MIRA from real IOTA data of the red giant star Arcturus.[26] The regularization is the quadratic one given by Eq. (45) with a prior set by a parametric fit of the data. This procedure was intended to check whether the interferometric data were compatible with more features than a simple limb darkened star.

7. COMPARISON WITH OTHER METHODS

There exists a number of methods designed to cope with the kind of data provided by optical interferometry. The self-calibration technique[27–31] has been developed for radio-astronomy and consists in deriving the Fourier phases at measured frequencies so that they match the phase closure data and otherwise remain as close as possible to the Fourier phase of the current image model. In an approach similar to the technique of self-calibration, WISARD[32] explicitly deals with phase ambiguities introducing as few new unknowns as possible to convert the phase closure data into Fourier phase pseudo-data. BSMEM[33] and the building-blocks method[5] attempt to reconstruct an image such that its bispectrum is in agreement with the bispectrum data. These two methods differ in their optimization strategy and in their regularization: the building-blocks method is a matching pursuit algorithm with an implicit regularization imposed by limiting the number of building-blocks; whereas BSMEM uses Skilling

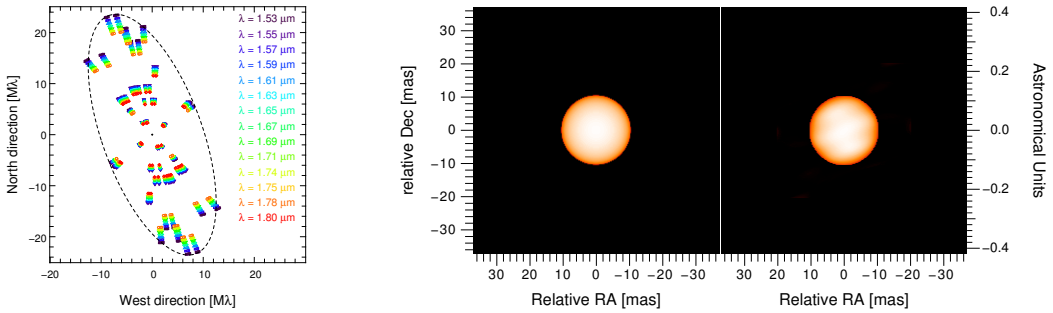

Figure 3. Image of the red giant star Arcturus. *Left*: u-v coverage with IOTA interferometer. *Middle*: parametric reconstruction by a limb darkening power law. *Right*: reconstruction by MIRA algorithm. Data courtesy: S. Lacour.

& Bryan[24] method to find the maximum entropy image which matches the phase closure data. The approach of MIRA is somewhat similar to BSMEM and the building-blocks method in that the algorithm directly fit the phase closures. However MIRA implements many different regularization methods and is able to account for any kind of data. In particular, since MIRA does not attempt to explicitly solve degeneracies, it can be used to restore an image — of course with at least a 180° orientation ambiguity — from the power spectrum only, *i.e.* without any phase information[34] as shown in Fig. 2.

8. CONCLUSIONS

With the development of optical interferometers arise the needs for image restoration algorithms dedicated to cope with this particular kind of data. In this framework, MIRA (a Multi-Aperture image Reconstruction Algorithm) has been designed to account for any kind of interferometric data (squared visibilities, complex visibilities, phase closures, ...) and make uses of proper regularization and constrained non-linear optimization to seek for the object brightness distribution. We have shown that MIRA is able to produce an image from with very limited amount of data, it is therefore specially efficient in the case of optical interferometry where the u-v coverage is very poor and where Fourier phase information can be weaker and degenerated (*e.g.* because it is only provided by phase closures). As an extreme case, MIRA is able to restore images without any phase information. This leads to the possibility to perform imaging with only 2 telescopes or when the phase closure data are corrupted.

ACKNOWLEDGMENTS

MIRA algorithm has been implemented and tested with YORICK (http://yorick.sourceforge.net/) which is available for free.

APPENDIX A. LIMITED FIELD OF VIEW

A general expression for a quadratic separable regularization is given by:

$$f_{\mathsf{prior}}(\boldsymbol{x}) = \sum\nolimits_j w_j \, x_j^2 \,, \tag{59}$$

where $w_j > 0$, $\forall j$, otherwise the criterion is degenerated. Note that this criterion can be seen as enforcing a loose support constraint with a weighting $\boldsymbol{w} > 0$ (this notation means $w_j > 0$, $\forall j$). The default solution $\boldsymbol{x}_{\mathsf{prior}}$ is obtained by minimizing the cost function in the absence of data and subject to the normalization and non-negativity constraints:

$$\boldsymbol{x}_{\mathsf{prior}} = \arg\min_{\boldsymbol{x}} f_{\mathsf{prior}}(\boldsymbol{x}) \quad \text{s.t.} \quad \boldsymbol{x} \geq 0 \quad \text{and} \quad \sum\nolimits_j x_j = \xi \,, \tag{60}$$

where $\xi > 0$ is the total flux of the solution. Assuming for the moment that all enequality constraints are inactive at the solution, the Lagrangian for the constrained problem can be written as:

$$\mathcal{L}(\boldsymbol{x}; \ell) = f_{\mathsf{prior}}(\boldsymbol{x}) - 2\,\ell \left(\sum\nolimits_j x_j - \xi \right) , \tag{61}$$

where ℓ is the Lagrange multiplier associated with the normalization constraint.[35] Minimizing $\mathcal{L}(\boldsymbol{x};\ell)$ with respect to $\boldsymbol{x}$ yields:

$$\boldsymbol{x}^+(\ell) = \arg\min_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x};\ell) \quad \Longleftrightarrow \quad x_j^+(\ell) = \ell\, w_j^{-1}\,.$$

The optimal Lagrange multiplier ℓ^+ is then identified by requiring the normalization:

$$\xi = \sum\nolimits_j x_j^+(\ell^+) = \ell^+ \sum_j w_j^{-1} \quad \Longleftrightarrow \quad \ell^+ = \frac{\xi}{\sum_j w_j^{-1}}\,.$$

Finally the default solution is:

$$x_j^{\mathsf{prior}} = x_j^+(\ell^+) = \frac{\xi\, w_j^{-1}}{\sum_j w_j^{-1}}\,, \tag{62}$$

which is normalized and striclty positive since $\boldsymbol{w} > 0$. Hence this validate our hypothesis that the inequality constraints were all inactive at the solution. Replacing the weights by their values for a given default solution, the regularization simply writes:

$$f_{\mathsf{prior}}(\boldsymbol{x}) = \sum\nolimits_j x_j^2/x_j^{\mathsf{prior}}\,, \tag{63}$$

where $\boldsymbol{x}_{\mathsf{prior}}$ is striclty positive and properly normalized: $\sum_j x_j^{\mathsf{prior}} = \xi$.

REFERENCES

[1] Narayan, R. and Nityananda, R., "Maximum entropy image restoration in astronomy," *Ann. Rev. Astron. Astrophys.* **24**, 127–170 (1986).

[2] Cornwell, T., "Imaging concepts," in [*ASP Conf. Ser. 82: Very Long Baseline Interferometry and the VLBA*], Zensus, J. A., Diamond, P. J., and Napier, P. J., eds., 39–+ (1995).

[3] Giovannelli, J.-F. and Coulais, A., "Positive deconvolution for superimposed extended source and point sources," *Astron. & Astrophys.* **439**, 401–412 (Aug. 2005).

[4] Högbom, J. A., "Aperture synthesis with a non-regular distribution of interferometer baselines," *A&AS* **15**, 417–426 (June 1974).

[5] Hofmann, K.-H. and Weigelt, G., "Iterative image reconstruction from the bispectrum," *Astron. & Astrophys.* **278**, 328–339 (Oct. 1993).

[6] Lannes, A., Anterrieu, E., and Marechal, P., "Clean and wipe," *Astron. & Astrophys. Suppl.* **123**, 183–198 (May 1997).

[7] Sramek, R. A. and Schwab, F. R., "Imaging," in [*Synthesis Imaging in Radio Astronomy*], Perley, R. A., Schwab, F. R., and Bridle, A. H., eds., *Astronomical Society of the Pacific Conference Series* **6**, 117–+ (1989).

[8] Cornwell, T. and Braun, R., "Deconvolution," in [*Synthesis Imaging in Radio Astronomy*], Perley, R. A., Schwab, F. R., and Bridle, A. H., eds., *Astronomical Society of the Pacific Conference Series* **6**, 167–+ (1989).

[9] Tarantola, A., [*Inverse Problem Theory and Methods for Model Parameter Estimation*], SIAM (2005).

[10] Thiébaut, E., "Introduction to image reconstruction and inverse problems," in [*Optics in Astrophysics*], Foy, R. and Foy, F.-C., eds., *NATO ASI*, Kluwer Academic (2005).

[11] Pauls, T. A., Young, J. S., Cotton, W. D., and Monnier, J. D., "A data exchange standard for optical (visible/ir) interferometry," *Publications of the Astronomical Society of the Pacific* **117**, 1255–1262 (Nov. 2005).

[12] Meimon, S., Mugnier, L. M., and Besnerais, G. L., "Convex approximation to the likelihood criterion for aperture synthesis imaging," *J. Opt. Soc. Am. A* **22**, 2348–2356 (2005).

[13] Haniff, C., "Least-squares fourier phase estimation from the modulo 2π bispectrum phase," *J. Opt. Soc. Am. A* **8**, 134–140 (Jan. 1991).

[14] Lannes, A., "Integer ambiguity resolution in phase closure imaging," *J. Opt. Soc. Am. A* **18**, 1046–1055 (2001).

[15] Besnerais, G. L., Lacour, S., Mugnier, L. M., Thiébaut, E., Perrin, G., and Meimon, S., "Imaging with long-baseline optical interferometry," *IEEE Journal of Selected Topics in Signal Processing* (submitted in 2008).

[16] Horn, K., "Images of accretion discs – i. the eclipse mapping method," *Mont. Not. R. Astron. Soc.* **213**, 129–141 (1985).

[17] Lucy, L. B., "Optimum strategies for inverse problems in statistical astronomy," *Astron. & Astrophys.* **289**, 983–994 (Sept. 1994).

[18] Pichon, C. and Thiébaut, E., "Non-parametric reconstruction of distribution functions from observed galactic discs," *Mont. Not. R. Astron. Soc.* **301**(2), 419–434 (1998).

[19] Lawson, P. R., Cotton, W. D., Hummel, C. A., Monnier, J. D., Zhao, M., Young, J. S., Thorsteinsson, H., Meimon, S. C., Mugnier, L., Le Besnerais, G., Thibaut, E., and Tuthill, P. G., "The 2004 optical/ir interferometry imaging beauty contest," *American Astronomical Society Meeting Abstracts* **205**, –+ (Dec. 2004).

[20] Thiébaut, E., Garcia, P. J. V., and Foy, R., "Imaging with amber/vlti: the case of microjets," *Ap&SS* **286**, 171–176 (2003).

[21] Golub, G. H., Heath, M., and Wahba, G., "Generalized cross-validation as a method for choosing a good ridge parameter," *Technometrics* **21**, 215–223 (1979).

[22] Titterington, D. M., "General structure of regularization procedures in image reconstruction," *Astron. & Astrophys.* **144**, 381–387 (1985).

[23] Gull, S. F., [*Maximum Entropy and Bayesian Methods*], ch. Developments in maximum entropy data analysis, 53–72, Kluwer Academic (1989).

[24] Skilling, J. and Bryan, R. K., "Maximum entropy image reconstruction: general algorithm," *Mont. Not. R. Astron. Soc.* **211**, 111–124 (1984).

[25] Thiébaut, E., "Optimization issues in blind deconvolution algorithms," in [*Astronomical Data Analysis II*], Starck, J.-L. and Murtagh, F. D., eds., **4847**, 174–183, SPIE (2002).

[26] Lacour, S., Meimon, S., Thiébaut, E., Perrin, G., Verhoelst, T., Pedretti, E., Schuller, P. A., Mugnier, L., Monnier, J., Berger, J., Haubois, X., Poncelet, A., Besnerais, G. L., Eriksson, K., Millan-Gabet, R., Lacasse, M., and Traub, W., "The limb-darkened arcturus - imaging with the iota/ionic interferometer," *Astron. & Astrophys.* **accepted for publication** (2008).

[27] Readhead, A. C. S. and Wilkinson, P. N., "The mapping of compact radio sources from VLBI data," *Astrophys. J.* **223**, 25–36 (July 1978).

[28] Cotton, W. D., "A method of mapping compact structure in radio sources using VLBI observations," *Astron. J.* **84**, 1122–1128 (Aug. 1979).

[29] Cornwell, T. J. and Wilkinson, P. N., "A new method for making maps with unstable radio interferometers," *Mont. Not. R. Astron. Soc.* **196**, 1067–1086 (Sept. 1981).

[30] Pearson, T. J. and Readhead, A. C. S., "Image Formation by Self-Calibration in Radio Astronomy," *Ann. Rev. Astron. Astrophys.* **22**, 97–130 (1984).

[31] Cornwell, T. and Fomalont, E. B., "Self-Calibration," in [*Synthesis Imaging in Radio Astronomy*], Perley, R. A., Schwab, F. R., and Bridle, A. H., eds., *Astronomical Society of the Pacific Conference Series* **6**, 185–+ (1989).

[32] Meimon, S., Mugnier, L. M., and Besnerais, G. L., "Reconstruction method for weak-phase optical interferometry," *Optics Letters* **30**, 1809–1811 (2005).

[33] Buscher, D. F., "Direct maximum-entropy image reconstruction from the bispectrum," in [*IAU Symp. 158: Very High Angular Resolution Imaging*], Robertson, J. G. and Tango, W. J., eds., 91–+ (1994).

[34] Thiébaut, E., "Reconstruction dimage en interférométrie optique," in [*XXIme Colloque GRETSI*], GRETSI (2007).

[35] Nocedal, J. and Wright, S. J., [*Numerical Optimization*], Springer Verlag, second edition ed. (2006).