

Global optimization for image restoration in optical/IR interferometry

Matthew Ozon, Éric Thiébaut, Ferréol Soulez, Jonathan Léger

▶ **To cite this version:**

Matthew Ozon, Éric Thiébaut, Ferréol Soulez, Jonathan Léger. Global optimization for image restoration in optical/IR interferometry. SPIE Astronomical Telescopes + Instrumentation Conference on Optical and Infrared Interferometry and Imaging V (SPIE-ATI'16), 2016, Edinburgh, United Kingdom. 10.1117/12.2233312 . insu-01632834

HAL Id: insu-01632834
https://insu.hal.science/insu-01632834

Submitted on 10 Nov 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Global optimization for image restoration in optical/IR interferometry

Matthew Ozon[a], Éric Thiébaut[a], Ferréol Soulez[a,b], and Jonathan Léger[a]

[a]Univ. Lyon, Univ. Lyon 1, ENS de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon UMR5574, F-69230, Saint-Genis-Laval, France

[b]Biomedical Imaging Group, École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland.

ABSTRACT

Many image reconstruction algorithms have been proposed to cope with optical/IR interferometric data. A first difficulty is to deal with the sparsity of the data and the uneven u-v coverage. This is usually overcome by imposing a priori constraints to the sought image. The reconstruction then amounts to solving an optimization problem where fidelity to the data and to the priors must both be satisfied to a chosen level. Due to the type of interferometric measurements, the image restored by the current algorithms depends on some initial guess and on the optimization strategy. Even though the resulting image is often satisfactory, there are no guarantees that the best image is obtained given the data and the constraints. This is a strong defect which may badly impact the interpretation of interferometric observations by astronomers. In this paper, we consider using stochastic methods such as simulated annealing to solve for the problem of image reconstruction with sparse Fourier amplitudes and unknown Fourier phases. This preliminary study is a first step toward image reconstruction with partial Fourier phase information.

Keywords: Stochastic global optimization; interferometric imaging; image reconstruction; phase retrieval.

1. IMAGE RECONSTRUCTION WITH UNKNOWN FOURIER PHASES

Ideally, interferometric measurements consist in samples of the Fourier spectrum of the object brightness distribution. The k-th datum then writes:

$$d_k = \widetilde{I}_\lambda(\nu_k) + n_k \tag{1}$$

where $\widetilde{I}_\lambda(\nu_k)$ is the Fourier transform at spatial frequency ν_k of the specific brightness distribution $I_\lambda(\theta)$ of the object in the angular direction θ (at wavelength λ) and n_k is a nuisance term which accounts for the noise. In what follows, we will assume that there are M sampled spatial frequencies.

In order to perform image reconstruction, we must choose a parametric representation of the object brightness distribution. For instance,[1] we assume that:

$$I_\lambda(\theta) = \sum\nolimits_{j=1}^{N} b(\theta - \theta_j)\, x_j\,, \tag{2}$$

where $b : \Theta \mapsto \mathbb{R}$ is some basis function over Θ the synthesized field of view, $x = (x_1, \ldots, x_N)^{\mathrm{t}} \in \mathbb{R}^N$ are the image parameters and $\{\theta_j \in \Theta\}_{j=1,\ldots,N}$ is a rectangular grid of angular directions over Θ. If the b_j are cardinal interpolation functions, x is a conventional pixelized image whose components are $x_j \approx I_\lambda(\theta_j)$. The Fourier transform of the image model writes:

$$\widetilde{I}_\lambda(\nu) = \sum\nolimits_{j=1}^{N} \widetilde{b}(\nu)\, \exp(-2\,\mathrm{i}\,\pi\,\nu^{\mathrm{t}}\,\theta_j)\, x_j\,, \tag{3}$$

with $\widetilde{b}(\nu)$ the Fourier transform of $b(\theta)$. Then the model of the data in Eq. (1) becomes:

$$d_k = \sum\nolimits_{j=1}^{N} H_{k,j}\, x_j + n_k \quad (\forall k = 1, \ldots, M) \quad \Longleftrightarrow \quad d = H\,x + n \tag{4}$$

E-mail: eric.thiebaut@univ-lyon1.fr

Optical and Infrared Interferometry and Imaging V, edited by Fabien Malbet, Michelle J. Creech-Eakman, Peter G. Tuthill, Proc. of SPIE Vol. 9907, 99071E
© 2016 SPIE · CCC code: 0277-786X/16/$18 · doi: 10.1117/12.2233312

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 08/30/2016 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

where the right hand side expression makes use of the matrix notation, M is the number of measurements and the coefficients of the operator $H \in \mathbb{C}^{M\times N}$ are given by:

$$H_{k,j} = \widetilde{b}(\nu_k)\, \exp(-2\,\mathrm{i}\,\pi\,\theta_j^{\mathrm{t}}\,\nu_k)\,. \tag{5}$$

When the complex Fourier spectrum is fully measured, image reconstruction amounts to solving a constrained optimization problem,[2] typically:

$$\widehat{x} = \underset{x\geq 0}{\arg\min}\Big\{\,\|H\,x - d\|_W^2 + \mu\,g(x)\Big\} \tag{6}$$

where $\|H\,x-d\|_W^2$ is a data fidelity term, $g(x)$ is a regularization term which implements prior knowledge about the sought image x and $\mu > 0$ is a tuning parameter which controls the trade-off between agreement to the data and to the priors. The constraint $x \geq 0$ holds componentwise and means that the sought image must be nonnegative everywhere. The notation $\|H\,x-d\|_W^2$ suggests that the noise n follows a Gaussian distribution and thus maximum likelihood leads to minimize some weighted quadratic norm of the difference between the data and their model. Provided $g(x)$ is convex, the solution to the Problem (6) is unique and can be efficiently computed by existing algorithms.

With optical interferometers, the situation is not so simple because not all Fourier information can be measured for the sampled frequencies $\{\nu_k\}_{k=1,\ldots,M}$. In particular, due to the turbulence, Fourier phase can only be partially measured by, for instance, so-called phase closure. The image reconstruction problem is then nonconvex and different strategies have been devised to cope with these issues.[3,4] In this paper we consider various stochastic methods to find the global optimal solution of the problem and compare these methods to other approaches implemented by existing image reconstruction algorithms. For the sake of generality, although this is not the easiest problem to solve, we assume that only the Fourier amplitudes have been measured and the image reconstruction is closely related to phase restoration problems commonly encountered in other domains as in crystallography. Assuming that measures at different frequencies are independent, the *easy* image reconstruction Problem (6) becomes a joint image and phase restoration problem:

$$\widehat{x},\widehat{\phi} \in \underset{(x,\phi)\in\mathbb{X}\times\Phi}{\arg\min}\left\{f_{\mathbb{X}\times\Phi}(x,\phi) = \sum\nolimits_k w_k\,|(H\,x)_k - \rho_k\,\mathrm{e}^{\mathrm{i}\,\phi_k}|^2 + \mu\,g(x)\right\}, \tag{7}$$

where ρ_k are the measured amplitudes (we assume that $\rho_k \geq 0$ whatever k), $w_k \geq 0$ are given statistical weights and $\phi = (\phi_1,\ldots,\phi_M)^{\mathrm{t}} \in \Phi = (-\pi,+\pi]^M$ are unknown Fourier phases and $\mathbb{X} = \{x \in \mathbb{R}^N \mid x \geq 0\}$ is the set of nonnegative images. In the following we consider different approaches to solve Problem (7). Note that the solution may not be unique.

1.1 Self Calibration Approach

Observing that $f(x,\phi)$ given in Eq. (7) is convex with respect to x knowing ϕ and that finding the best phase given x has a simple solution, a first possibility is to solve the joint problem by the following alternating method:

0. Choose an initial image $x^{(0)}$, then for $t = 1, 2, \ldots$ and until convergence repeat:

1. Phase restoration step (self-calibration):

$$\phi^{(t)} = \underset{\phi}{\arg\min}\, f_{\mathbb{X}\times\Phi}\big(x^{(t-1)},\phi\big) = \underset{\phi}{\arg\min}\sum\nolimits_k w_k\,\left|\big(H\,x^{(t-1)}\big)_k - \rho_k\,\mathrm{e}^{\mathrm{i}\,\phi_k}\right|^2, \tag{8}$$

   the problem is separable and, $\forall k$, the trivial solution is $\phi_k^{(t)} = \angle\big(H\,x^{(t-1)}\big)_k$ where $\angle z$ yields the phase of $z \in \mathbb{C}$ and provided $\rho_k > 0$.

2. Image restoration step:

$$x^{(t)} = \underset{x\geq 0}{\arg\min}\, f_{\mathbb{X}\times\Phi}\big(x,\phi^{(t)}\big) = \underset{x\geq 0}{\arg\min}\left\{\sum\nolimits_k w_k\,\left|(H\,x)_k - \rho_k\,\mathrm{e}^{\mathrm{i}\,\phi_k^{(t)}}\right|^2 + \mu\,g(x)\right\}, \tag{9}$$

   which amounts to solving a strictly convex problem and has a unique solution.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 08/30/2016 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

The above algorithm is similar to the strategy implemented by so-called *self-calibration* methods[5] for radio-interferometry or more recently by WISARD[3,6] for optical interferometry. The main difference with these latter approaches is that rigorously the same criterion $f(x,\phi)$ is minimized by the alternating method. It is possible to generalize this method to the other kinds of data that can be provided by optical interferometers (not just the Fourier amplitudes).[7]

1.1.1 Hierarchical Optimization

As seen in the previous section, finding the best unknown Fourier phases ϕ given the image x has a closed form solution which is easy to compute. This suggests to get rid of these phases by replacing them by their optimal values given the image parameters x. This yields a criterion $f_{\mathbb{X}}(x) = \min_\phi f_{\mathbb{X}\times\Phi}(x,\phi)$ which only depends on the image parameters. The image restoration problem then writes:

$$\widehat{x} = \arg\min_{x\in\mathbb{X}} \left\{ f_{\mathbb{X}}(x) = \min_{\phi\in\Phi} f_{\mathbb{X}\times\Phi}(x,\phi) = \sum\nolimits_k w_k \left(|H\,x|_k - \rho_k \right)^2 + \mu\, g(x) \right\}. \tag{10}$$

The criterion $f_{\mathbb{X}}(x)$ is however non-convex and global optimization is *de facto* required to find its global minimum.

Nevertheless, monotonic optimization algorithms may be able to find a good local minimum. This is the approach implemented by BSMEM[8,9] and MIRA[10] which, starting from an initial image, perform a local non-linear constrained optimization with algorithm MEMSYS[11,12] for BSMEM and VMLMB[13] for MIRA. The result then depends on the initial image which may be randomly chosen.

1.1.2 Stochastic Methods

The size of an image reconstruction problem forbids to apply a systematic search in the parameter space. Among the algorithms proposed to solve global optimization problems of large dimensions, stochastic methods like simulated annealing (SA) have some desirable characteristics. For short, the main advantage is that a theoretical framework proves the convergence and gives hint on how to tune the algorithm so that the convergence is almost sure. Furthermore, while the brute force algorithms grow exponentially with the dimensionality, SA only grows linearly. This motivate the use of SA in the presented study. Note that stochastic methods are already exploited by existing algorithms. For example, algorithms MACIM[4] or SQUEEZE[14] applied to the problem in Eq. (10) would attempt to find the global optimum by means of a simulated annealing method or Monte Carlo Markov Chain (MCMC) method.

Applying a stochastic method to the variables x and find the global minimum of $f_{\mathbb{X}}(x) = \min_{\phi\in\Phi} f_{\mathbb{X}\times\Phi}(x,\phi)$ may be prohibitively slow because there is a lot of variables (as many as the number N of pixels). As $f_{\mathbb{X}\times\Phi}(x,\phi)$ in x, the function:

$$f_\Phi(\phi) = \min_{x\in\mathbb{X}} f_{\mathbb{X}\times\Phi}(x,\phi) = f_{\mathbb{X}\times\Phi}(\widehat{x}(\phi),\phi)\,, \tag{11}$$

where:

$$\widehat{x}(\phi) = \arg\min_{x\in\mathbb{X}} \left\{ f_{\mathbb{X}\times\Phi}(x,\phi) = \sum\nolimits_k w_k \left| (H\,x)_k - \rho_k\, \mathrm{e}^{\mathrm{i}\,\phi_k} \right|^2 + \mu\, g(x) \right\}, \tag{12}$$

is properly defined and only depend on ϕ. The idea is then to apply the simulated annealing method to the M unknown phases which, in optical interferometry, is usually much smaller than the number N of pixels. For each proposed phase, an image reconstruction problem has to be solved and is likely to be the largest consumer of computational time. So the computational burden scales as $\sim t_{\mathsf{img}}(N)\, t_{\mathsf{glob}}(M)$ where $t_{\mathsf{img}}(N)$ represents the time spent to reconstruct a N-pixel image while $t_{\mathsf{glob}}(M)$ represents the time spent by a simulated annealing strategy to solve a global optimization problem with M variables. As the cost of the image reconstruction may be too important, a last possibility is to stochastically sample the whole set of parameters, x and ϕ, with the simulated annealing method. Then the time spent should scale as $\sim t_{\mathsf{glob}}(M+N)$.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 08/30/2016 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

2. SIMULATED ANNEALING

We consider the problem of finding a global minimum of a real valued multi-variate function:

$$\widehat{\omega} \in \arg\min_{\omega \in \Omega} f_\Omega(\omega)\,, \tag{13}$$

where, depending on the chosen problem, the parameters ω and the corresponding space Ω can be $\omega = x$ and $\Omega = \mathbb{X}$ for $f_\Omega = f_\mathbb{X}$ defined in Eq. (10), $\omega = \phi$ and $\Omega = \Phi$ for $f_\Omega = f_\Phi$ defined in Eq. (11) or $\omega = (x, \phi)$ and $\Omega = \mathbb{X} \times \Phi$. Whatever the choice of the parameter space Ω, the criterion $f_\Omega(\omega)$ is non-convex in ω and monotonic optimization algorithms do not apply and we propose to Simulated Annealing (SA) to sample the global minima of $f_\Omega(\omega)$. However, theoretical results ensuring the convergence of SA have only been derived for problems defined over finite parameter spaces. We will therefore assume that the parameters $\omega \in \Omega$ take a finite number of possible values, possibly after proper discretization.

The family of SA algorithms is well suited for large combinatorial optimization and guarantees to reach a state that belongs to the set of global optimum given a tolerance margin. Contrary to monotonic descent algorithms, SA allows unfortunate configurations to be temporarily accepted in order to escape local optimal configurations (similar to metastable states during annealing in metallurgy[15,16]) and eventually reach a global minimum. The controlled acceptance of uphill configurations — those that increase the energy — is the relevant distinction from deterministic algorithms or other stochastic samplers.

2.1 General Implementation

The SA is an iterative procedure that generates a sequence of states $\omega^{(t)} \in \Omega$. A typical implementation is given by Algorithm 1 whose output is a sequence $(\omega^{(t)})_{0 \leq t \leq N_{\mathsf{SA}}\,|\Omega|}$ which is a Markov chain (MC). This chain is defined by three elements: the energy function f_Ω, a temperature schedule $(T_n)_{1 \leq n \leq N_{\mathsf{SA}}}$, and a communication kernel $q(\omega \,|\, \omega', T)$ that may depend on the temperature T and which gives the probability to generate a state ω given a previous state ω' and a temperature T. More specifically, SA draws every candidate ω (but the first one) at random as $\omega \sim q(\omega \,|\, \omega', T)$ with ω' the previously accepted candidate and T the current temperature. The kernel can depend on the temperature so that the state space is largely sampled at the beginning of the process and narrow down with the iterations. In other words, at the beginning of the SA process, the kernel may generate large steps in order to visit a large regions of the landscape while at the end only small steps can be proposed. This way, the convergence speeds up in terms of iterations.

Pick an initial state $\omega^{(0)} \in \Omega$ at random.
Set the iterator $t = 0$.
for $n = 1, \ldots, N_{\mathsf{SA}}$ **do**
  Update temperature: set $T = T_n$.
  for $j = 1, 2, \ldots, |\Omega|$ **do**
    Draw a candidate ω from the probability distribution $q(\omega \,|\, \omega^{(t)}, T)$.
    Set $\Delta = f_\Omega(\omega) - f_\Omega\big(\omega^{(t)}\big)$.
    Set $\omega^{(t+1)} = \omega$ with probability $\min\big(1, \mathrm{e}^{-\Delta/T}\big)$, otherwise set $\omega^{(t+1)} = \omega^{(t)}$.
    Set $t = t + 1$.

Algorithm 1: Simulated Annealing minimizing the energy function $f_\Omega(\omega)$.

The final state $\omega^\star = \omega^{(N_{\mathsf{SA}}\,|\Omega|)}$ of the sequence produced by Algorithm 1 belongs to the set of global minimizers of f_Ω under weak assumptions that are discussed in Section 2.2. To put things clear, the convergence of SA is not defined in terms of distance to the solution, which would be a non-sense for the function having more than one global minimizer. Rather, the convergence properties arise from the probability of belonging to the set of minimizers. Hence, considering a margin of error $\epsilon > 0$, the probability that the last element $\omega^\star = \omega^{(N_{\mathsf{SA}}\,|\Omega|)}$ of the Markov chain does not belong to the set of global optimizers is given by:

$$\Pr\big(f_\Omega(\omega^\star) \geq \min_{\omega \in \Omega} f_\Omega(\omega) + \epsilon\big) \leq (N_{\mathsf{SA}}\,|\Omega|)^{-\alpha}\,, \tag{14}$$

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 08/30/2016 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

for some $\alpha > 0$. Clearly the convergence speed depends on α.

It is interesting to notice that a global minimum may not be unique and the Algorithm 1 is a uniform sampler over the set of global minima. Having several minima is not an aberration, but merely the representation of the complexity of the reconstruction problem that may accept several good solutions.

2.2 Assumptions

Fortunately, the assumptions ensuring the convergence* of an SA process in a finite number of iterations are weak. The requirements concerning the temperature, the communication kernels and the energy function are few and not too restrictive. First, the energy function f_Ω must be defined over a finite state space, yet no other constraints add up. Then, the family of communication kernels $q(\omega \,|\, \omega', T)$ and its limit while $T \to 0$ must satisfy the three following assumptions:

Irreducibility. The irreducibility expresses the idea that at a fixed temperature of the cooling schedule, or at least before the end of the annealing, all states are reachable from another state in a finite number of steps. In other words, there exists a path $(\omega^{(t)})_{0 \leq t \leq n}$ starting at any state $\omega^{(0)} \in \Omega$ and reaching any other $\omega^{(n)} \in \Omega$ in a finite number of steps $n \in \mathbb{N}^* < \infty$ with a non null probability:

$$\forall t \in [\![1, n]\!], \quad q\big(\omega^{(t)} \,\big|\, \omega^{(t-1)}, T\big) > 0\,. \tag{15}$$

Symmetry of the support. This property means that it is always possible to turn back. If the transition from $\omega_1 \in \Omega$ to $\omega_2 \in \Omega$ is possible, that is to say $q(\omega_2 \,|\, \omega_1, T)$ is not null, then the transition from ω_2 to ω_1 must also be possible which implies that $q(\omega_1 \,|\, \omega_2, T)$ is not null either. The two probabilities of transition are however not necessary the same. Formally, the following must stand:

$$q(\omega_2 \,|\, \omega_1, T) \neq 0 \implies q(\omega_1 \,|\, \omega_2, T) \neq 0\,, \tag{16}$$

for any pair $(\omega_1, \omega_2) \in \Omega^2$. It implies that the support Δq of the communication kernel q:

$$\Delta q = \big\{(\omega_1, \omega_2) \in \Omega^2 \,\big|\, q(\omega_1 \,|\, \omega_2, T) > 0\big\}\,, \tag{17}$$

is symmetric.

Equality of the supports. During the cooling process, the support of q may depends on the temperature T, however, starting from a temperature $T^\star > T_{\mathsf{end}}$, the support of the communication kernels must be the same. Formally, the following must stand:

$$q(\omega_2 \,|\, \omega_1, T) \neq 0 \implies q(\omega_2 \,|\, \omega_1, T') \neq 0\,, \tag{18}$$

for any pair $(\omega_1, \omega_2) \in \Omega^2$ and any temperatures $T \leq T^\star$ and $T' \leq T^\star$.

2.3 Temperature estimation

Since the early 90's, the SA process has been proven to converge in a finite number of iterations when using exponential piecewise constant decreasing temperature sequences of the form:[17]

$$\forall n \in [\![0, N_{\mathsf{SA}}]\!], \quad T_n = T_{\mathsf{init}}\,(T_{\mathsf{end}}/T_{\mathsf{init}})^{n/N_{\mathsf{SA}}}\,, \tag{19}$$

in Algorithm 1. The estimation of the temperature parameters T_{init} and T_{end} is based on the average acceptance rate χ_T of uphill moves for a given temperature T and computed for two chains of states that increase the energy:

$$\chi_T = \frac{1}{K}\sum_{t=1}^{K} \mathrm{e}^{-\Delta_T^{(t)}/T}\,, \tag{20}$$

where the sequence $(\Delta_T^{(t)})_{0 \leq t \leq K}$ is the energy variation induced by the uphill chain, and K is its length. In order to have a good exploration of the energy landscape, an SA process should be characterized by the uphill acceptance rates in $[0.6, 0.9]$ at the beginning and in $[10^{-5}, 10^{-3}]$ at the end. Furthermore, it is desirable that χ_T decrease linearly with the iterations. We noticed that it is interesting to finish the process with a few extra steps ($\simeq 10$) at null temperature so that the algorithm does not finish after uphill moves.

*in terms of probability of belonging to the set of minimizers

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 08/30/2016 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

2.4 Definition of the communication kernel

The idea of the communication kernel is to sample the state space in a clever way. Indeed, as it is mentioned before, the SA process should be able to visit a large panel of states so that it does not get trapped by local optima attraction basins, while at the end of the process, only small perturbations are required in order to overcome the frictions due to the ruggedness of the energy landscape. Hence a good communication kernel must change during the process. Therefore, we propose two elementary kernels that depend on T and form the base of communication kernels used in this study. One, q_Φ, explores the phase space and the other, $q_\mathbb{X}$, acts over the pixel space. Three kernels are used in our study: q_Φ is used for SA on the phases, $q_\mathbb{X}$ is used for SA on the pixels and a composite kernel $q_{\mathbb{X}\times\Phi}$ is used for Full SA Reconstruction (FSAR), that is to say SA on both the pixels and the phases:

$$q_{\mathbb{X}\times\Phi}(\omega \,|\, \omega', T) = \tau\, q_\Phi(\phi \,|\, \phi', T) + (1-\tau)\, q_\mathbb{X}(x \,|\, x', T)\,. \tag{21}$$

with $\omega = (x, \phi)$ and $\omega' = (x', \phi')$. The scalar $\tau \in [0, 1]$ is the probability of choosing the phase space over the pixel space. The dependence on the temperature in (21) is only supported by the two elementary kernels, though τ could vary with the iterations as a way of speeding up the convergence. For convenience, and because the proposed kernels depend linearly on the iteration, we define the normalized index $\zeta(T) = \log(T/T_{\mathsf{init}})/\log(T_{\mathsf{end}}/T_{\mathsf{init}}) = n/N_{\mathsf{SA}}$ that linearly varies in the range $[0, 1]$ with the iterations.

2.4.1 Pixel kernel

This kernel aims at modifying the scene $x \in \mathbb{X}$ that is being reconstructed. For the sake of simplicity, our pixel kernel $q_\mathbb{X}(x \,|\, x', T)$ implements the most elementary moves. Considering any given state $x' \in \mathbb{X}$, the proposed state x is generated by the following procedure:

1) draw a pixel index $j_0 \sim \mathcal{U}(\llbracket 1, N \rrbracket)$;
2) define the maximum amplitude $\delta_x(T) = \lfloor \delta_x^{\mathsf{max}} - (\delta_x^{\mathsf{max}} - 1)\,\zeta(T) \rceil \in \llbracket 1, \delta_x^{\mathsf{max}} \rrbracket$;
3) draw the level of variation $\ell \sim \mathcal{U}(\llbracket -\delta_x(T), +\delta_x(T) \rrbracket)$;
4) $\forall j \in \llbracket 1, N \rrbracket$, set $x_j = \begin{cases} \max\{0, \min\{1, x'_j + \ell/(L_x - 1)\}\} & \text{if } j = j_0, \\ x'_j & \text{otherwise;} \end{cases}$

where $L_x \in \mathbb{N}^*$ is the number of discrete values in the range $[0, 1]$ of allowed pixel intensities and δ_x^{max} is the maximum amplitude of variation. The amplitude $\delta_x(T)$ decreases linearly with the iterations from δ_x^{max} at T_{init} to $\delta_x^{\mathsf{min}} = 1$ at T_{end}. In practice, we take $\delta_x^{\mathsf{max}} = (L_x - 1)/2$. Such a kernel takes large steps at the beginning and only small ones at the end so that the energy landscape is vastly explored at the beginning and refined at the end. Note that $q_\mathbb{X}$ enforces the constraints of belonging to the closed set $\mathbb{X}$, *i.e.* the positivity and the bounded nature of the brightness.

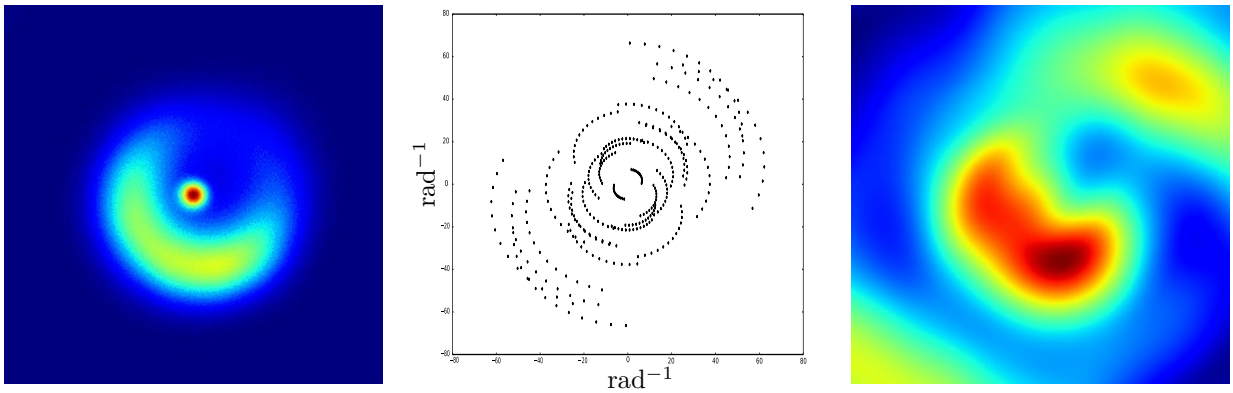
2.4.2 Phase kernel

The proposed phase kernel implements simple atomic modifications in the phase space. One of the phase is picked at random over all the frequencies but the null one (which is assumed to be at index $k = 1$), then an additive variation is applied. The frequencies are sampled following the distribution $\mathcal{L}(\rho)$ that linearly depends on the amplitude of the complex visibilities. The idea is to visit more frequently the phases of larger influence than those inducing small energy variation. Therefore, the main components are well reconstructed while the details are neglected. Our phase kernel is implemented as follows:

1) draw a phase index $k_0 \sim \mathcal{L}(\rho)$;
2) define the maximum amplitude $\delta_\phi(T) \in \llbracket 1, \delta_\phi^{\mathsf{max}} \rrbracket$;
3) draw the level of variation $\ell \sim \mathcal{U}(\llbracket -\delta_\phi(T), +\delta_\phi(T) \rrbracket)$;
4) $\forall k \in \llbracket 1, M \rrbracket$, set $\phi_k = \begin{cases} \phi'_k + (2\,\pi/\rho_k)\,(\ell/L_\phi) \mod 2\,\pi & \text{if } k = k_0, \\ \phi'_k & \text{otherwise;} \end{cases}$

where $L_\phi \in \mathbb{N}^*$ is the number of discrete values in the range $[-\pi/\rho_k, +\pi/\rho_k]$ of phase variation and $\delta_\phi^{\mathsf{max}}$ is the maximum amplitude of variation.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 08/30/2016 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx



(a) (b) (c)

Figure 1: Simulated data for the experiments: a) the true object,[18] b) Fourier plan coverage and c) dirty beam.

3. RESULTS

We apply the proposed SA variants to simulated data.

3.1 Simulated data

We consider the object of the 2014 Interferometric Imaging Beauty Contest.[18] We simulated noisy data with independent Gaussian noise with given signal to noise ratio (SNR) as follows:

$$d_k = \sum_{j=1}^{N} H_{k,j}\, x_j^{\mathsf{true}} + n_k \quad (\forall k = 1, \ldots, M), \tag{22}$$

with x_j^{true} the true object displayed in Fig. 1 and independent Gaussian noise such that $\mathrm{Re}(n_k) \sim \mathcal{N}(0, \sigma_k^2)$ and $\mathrm{Im}(n_k) \sim \mathcal{N}(0, \sigma_k^2)$. To ensure a given SNR, the standard deviation σ_k of the noise at k-th frequency is given by (see Appendix A):

$$\sigma_k = \frac{\max(|z_k|, \rho_{\mathsf{min}})}{\sqrt{2\,\mathrm{SNR}^2 + 2\,\mathrm{SNR}\,\sqrt{1+\mathrm{SNR}^2}}} \underset{\mathrm{SNR}\gg 1}{\approx} \frac{\max(|z_k|, \rho_{\mathsf{min}})}{2\,\mathrm{SNR}}, \tag{23}$$

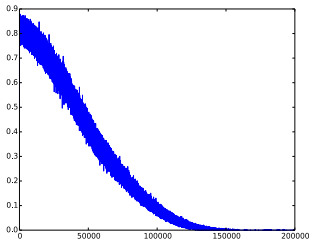
with $z_k = \sum_{j=1}^{N} H_{k,j}\, x_j^{\mathsf{true}}$ and $\rho_{\mathsf{min}} > 0$. For our simulations, we consider very good data and take $\mathrm{SNR} = 100$ and $\rho_{\mathsf{min}} = 10^{-2}$. The number of sampled frequencies (see Fig. 1b) is $M = 196$ including the zero-th frequency.

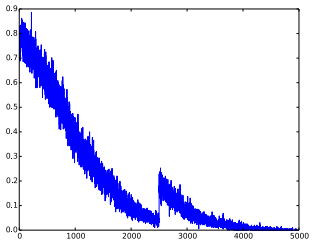
3.2 Common Settings

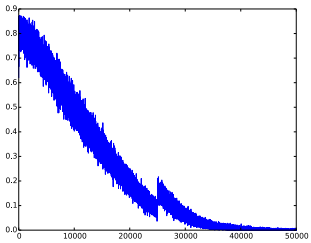
For the regularization we use the loose support constraint[19] and tune the regularization level μ so that the best image (by visual assessment) can be reconstructed from the complete data (Fourier amplitudes and phases) by solving Problem (6). The corresponding image is shown by Fig. 3a.

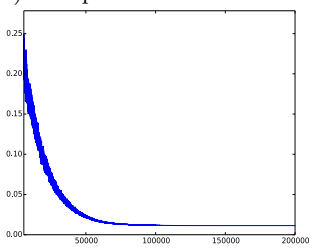
In order to minimize the number of parameters, we take the largest pixel size as possible (because it yields small images and hence fast computations in all respect) while keeping enough degrees of freedom. To determine the largest pixel size, we proceed as follows. We perform a convex image restoration with the complete data (Fourier amplitudes and phases) by solving Problem (6) with a given coarse pixel size. We then perform a second reconstruction, again solving Problem (6), with a much smaller pixel from the same simulated Fourier amplitudes but Fourier phases computed from the restored coarse image and check that this second image compares well with the best one shown in Fig. 3a. The largest pixel size selected in this manner leads us to apply the SA methods to $14 \times 14 = 196$ pixel images.

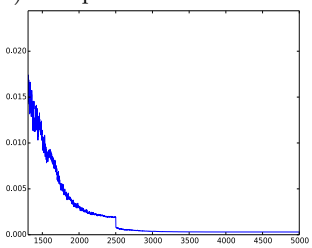
Downloaded From: http://proceedings.spiedigitallibrary.org/ on 08/30/2016 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx


(a) Acceptance rate for $\Omega = \mathbb{X}$.


(b) Acceptance rate for $\Omega = \Phi$.


(c) Acceptance rate for $\Omega = \mathbb{X} \times \Phi$.


(d) Evolution of f_Ω for $\Omega = \mathbb{X}$.


(e) Evolution of f_Ω for $\Omega = \Phi$.

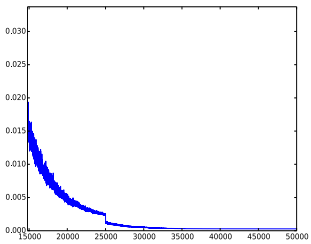

(f) Evolution of f_Ω for $\Omega = \mathbb{X} \times \Phi$.

Figure 2: Evolution of the acceptance rate and the energy with the iterations during the SA process for the considered variants.

3.3 SA on the Phases

For reconstructions based on SA on the phases, the phase kernel q_Φ is used with $L_\phi = 180$ and $\delta_\phi(T) = 2$ if $\zeta(T) < 1/2$ and $\delta_\phi(T) = 1/2$ for cooler temperatures, so that the angular precision is ≈ 2 degrees.The total number of SA iterations is set to $N_{\mathsf{SA}} = 5000$ which is roughly 5000 times the number of iterations for the algorithm used for image reconstructions.

For the image reconstructions, we use VMLMB,[13] a quasi-Newton method with bound constraints, to carry the constrained optimization of the convex criterion defined in (12). VMLMB is used with a maximum of 20 iterations and enforcing the pixel values of being in the set $\mathbb{X} = [0, 1]^N$. To speed up the computations, each reconstruction for a new candidate is started with the image obtained for the previous accepted candidate.

3.4 SA on the Pixels

For reconstructions based on SA on the pixels, the pixel kernel $q_\mathbb{X}$ is used with $L_\phi = 50/\max(x_{\mathsf{true}})$ with x_{true} the true image. The total number of SA iterations is set to $N_{\mathsf{SA}} = 200000$.

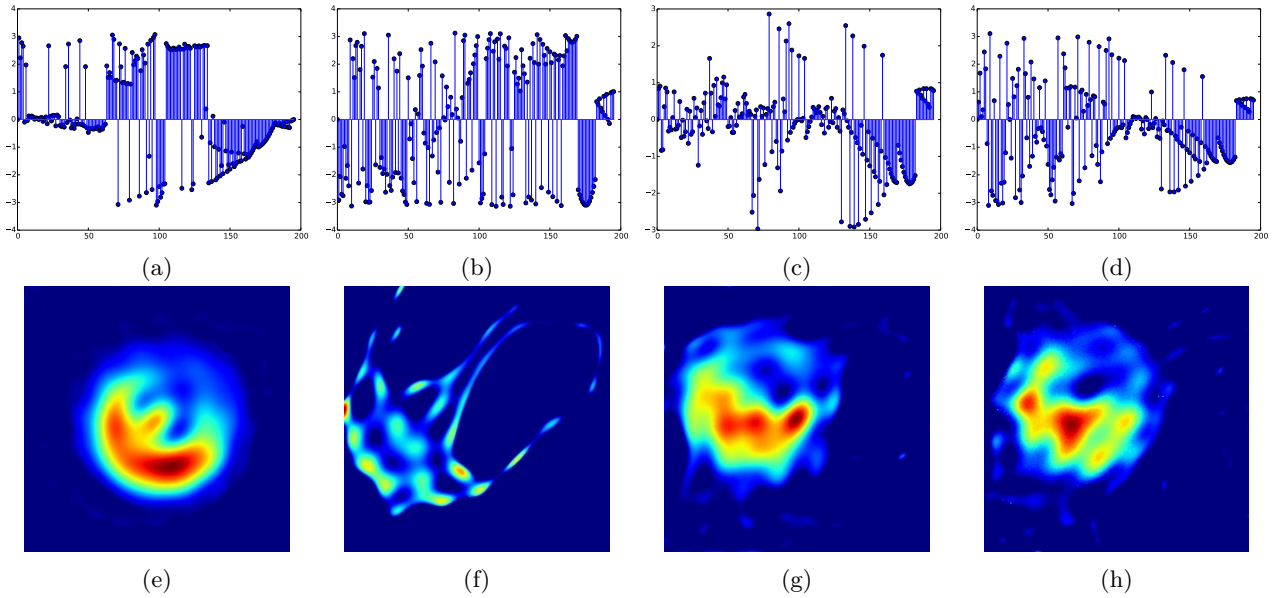
3.5 SA on the Pixels and on the Phases (FSAR)

For Full SA Reconstruction (FSAR) which applies SA on both the pixels and the phases, the composite kernel $q_{\mathbb{X}\times\Phi}$ is used with $\tau = 1/2$. The settings for the pixel kernel $q_\mathbb{X}$ and the phase kernel q_Φ are the same as above. The total number of SA iterations is set to $N_{\mathsf{SA}} = 50000$.

3.6 Analysis of the Convergence

Figure 2 shows the evolution of the acceptance rate and the energy during the SA iterations of the considered problems. A few remarks can be made. (i) The uphill move acceptance rate varies from $0.8 \sim 0.9$ to $\sim 10^{-3}$ which indicates that at the beginning, the space is widely visited, then the criterion of minimizing the energy is more and more enforced, and, at the end, only few uphill moves are accepted so that the minimization terminates without being stuck because of the roughness of the energy surface. (ii) The energy tends to decrease fast at the beginning, then it varies slowly. The convergence is clear, but it is not obviously in the absolute minimum unfortunately. (iii) Two interesting features of the figures 2b, 2c, 2e and 2f: the discontinuity indicates the change in precision of the phase kernel (which is not visible for the pixel kernel because the precision varies slowly). As a consequence, there are almost two SA process in one, such as a multiscale approach. First a coarse convergence then the refinement.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 08/30/2016 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 3: Phase estimations (top row) and image reconstructions (bottom row). The phases are sorted with respect to the amplitude of the true complex visibilities. (e) is the best image that can be reconstructed from the complete data (Fourier amplitudes and phases) by solving Problem (6); (f) is the final image with SA on the pixels, *i.e.* with $\Omega = \mathbb{X}$; (g) is the final image with SA on the phases, *i.e.* with $\Omega = \Phi$; (h) is the final image with SA on the pixels and on the phases, *i.e.* with $\Omega = \mathbb{X} \times \Phi$.

3.7 Analysis of the Estimated Parameters

The estimated phases and restored images are shown in Figure 3. Sub-figures 3a and 3e shows the best results we can hope for, considering the reconstruction model, the regularization parameter and the optimization algorithm. The phases are organized by increasing visibility magnitude. It is clear that they are organized as groups (which should be integrated in the communication kernel in order to improve the quality of the result). Sub-figures 3b and 3f show the result when applying SA on the pixels: only the very first phase are well organized, as a consequence the reconstruction is poor. Clearly the method was not successful with the allowed number of SA iterations. By contrast, the phase restored by applying SA on the phases (Fig. 3c) or on both the pixels and the phases (Fig. 3d) are closer to the actual phases. As a result the restored images (Fig. 3g and Fig. 3h) are better approximations of the truth. As we expected, the phase and the image obtained by applying SA on the smaller number of parameters, that is on the phases with $\Omega = \Phi$, yields the best results. However applying SA on the phases and on the pixels is not so bad.

4. DISCUSSION AND PERSPECTIVES

The paper presents a preliminary study of exploiting Simulated Annealing (SA) to solve the difficult problem of image restoration from sparse measurements of the Fourier amplitudes. This also amounts to solving a Fourier phase restoration problem. Our results show that applying SA on the unknown phase coupled with a convex image reconstruction is the most effective of the considered stochastic variants. In particular, applying SA to the pixels rather than to the phases is easier (because there are no needs to implement a convex image reconstruction algorithm) but yields much worse results. This preliminary study is a first step toward image reconstruction with partial Fourier phase information which will be considered in a following work.

APPENDIX A. COMPUTATION OF THE SIGNAL TO NOISE RATIO

Consider the powerspectrum $a = |z|^2$ of $z \in \mathbb{C}$ with i.i.d. noise in the real and imaginary parts, *i.e.* $\mathrm{Re}(z) \sim \mathcal{N}(\mathrm{Re}(\mathbb{E}\{z\}), \sigma^2)$ and $\mathrm{Im}(z) \sim \mathcal{N}(\mathrm{Im}(\mathbb{E}\{z\}), \sigma^2)$. Then a/σ^2 follows a non-central χ-squared distribution with

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 08/30/2016 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

$\kappa = 2$ degrees of freedom and parameter $\lambda = \mathbb{E}\{\mathrm{Re}(z/\sigma)\}^2 + \mathbb{E}\{\mathrm{Im}(z/\sigma)\}^2 = |\mathbb{E}\{z/\sigma\}|^2$. Expectation and variance of a/σ^2 are given by:

$$\mathbb{E}\{a/\sigma^2\} = \kappa + \lambda\,, \tag{24}$$

$$\mathrm{Var}\{a/\sigma^2\} = 2\,\kappa + 4\,\lambda\,, \tag{25}$$

Hence:

$$\mathbb{E}\{a\} = (\kappa + \lambda)\,\sigma^2 = |\mathbb{E}\{z\}|^2 + 2\,\sigma^2\,, \tag{26}$$

$$\mathrm{Var}\{a\} = (2\,\kappa + 4\,\lambda)\,\sigma^4 = 4\,\sigma^4 + 4\,|\mathbb{E}\{z\}|^2\,\sigma^2\,, \tag{27}$$

Therefore $\widehat{a} = a - 2\,\sigma^2$ is an unbiased estimator of the power spectrum and its signal to noise ratio (SNR) is:

$$\mathrm{SNR}(\widehat{a}) \stackrel{\mathrm{def}}{=} \frac{|\mathbb{E}\{\widehat{a}\}|}{\sqrt{\mathrm{Var}\{\widehat{a}\}}} = \frac{|\mathbb{E}\{z\}|^2}{\sqrt{4\,\sigma^4 + 4\,|\mathbb{E}\{z\}|^2\,\sigma^2}}\,. \tag{28}$$

To generate artificial complex-value data with a given SNR (for the powerspectrum), the variance must be:

$$\sigma^2 = \frac{|\mathbb{E}\{\widehat{a}\}|}{2\,\mathrm{SNR}(\widehat{a})^2 + 2\,\mathrm{SNR}(\widehat{a})\,\sqrt{1 + \mathrm{SNR}(\widehat{a})^2}}\,. \tag{29}$$

As $\lim_{|\mathbb{E}\{z\}|\to 0}\sigma = 0$, we use a minimal value for σ.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2013-2016) under grant agreement number 312430 (OPTICON).

REFERENCES

[1] Thiébaut, É., "Image reconstruction with optical interferometers," *New Astronomy Reviews* **53**, 312–328 (2009).

[2] Thiébaut, É. and Giovannelli, J.-F., "Image reconstruction in optical interferometry," *IEEE Signal Process. Mag.* **27**(1), 97–109 (2010).

[3] Meimon, S., Mugnier, L. M., and le Besnerais, G., "Reconstruction method for weak-phase optical interferometry," *Optics Letters* **30**(14), 1809–1811 (2005).

[4] Ireland, M., Monnier, J., and Thureau, N., "Monte-Carlo imaging for optical interferometry," in [*Advances in Stellar Interferometry.*], Monnier, J. D., Schöller, M., and Danchi, W. C., eds., **6268**, 62681T1–62681T8, SPIE (2008).

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

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

[7] Thiébaut, É., "Image reconstruction in optical interferometry," in [*invited paper in Signal Recovery & Synthesis, topical meeting of the Optical Society of America*], (2011).

[8] 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., 91–93 (1994).

[9] Baron, F. and Young, J. S., "Image reconstruction at cambridge university," in [*Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*], *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference* **7013**, 70133X (2008).

[10] Thiébaut, É., "MiRA: an effective imaging algorithm for optical interferometry," in [*Astronomical Telescopes and Instrumentation*], Markus Schöller, William C. Danchi, F. D., ed., **7013**, 70131I–1–70131I–12, SPIE (2008).

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 08/30/2016 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx

[11] Skilling, J., "Classic maximum entropy," in [*Maximum Entropy and Bayesian Methods*], 45–52, Springer Science & Business Media (1989).
[12] Gull, S. F., "Developments in maximum entropy data analysis," in [*Maximum Entropy and Bayesian Methods*], 53–71, Springer Science & Business Media (1989).
[13] Thiébaut, É., "Optimization issues in blind deconvolution algorithms," in [*Astronomical Data Analysis II*], Starck, J.-L. and Murtagh, F. D., eds., **4847**, 174–183, SPIE (2002).
[14] Baron, F., Monnier, J. D., and Kloppenborg, B., "A novel image reconstruction software for optical/infrared interferometry," in [*Proc. SPIE Optical and Infrared Interferometry II*], **7734**, 77342I (2010).
[15] Silliman, B., [*Principles of Physics, Or Natural Philosophy*], Philadelphia, Theodore (1871).
[16] Ewing, J. and Rosenhain, W., "The crystalline structure of metals. second paper.," *Proceedings of the Royal Society of London* **67**(435-441), 112–117 (1900).
[17] Catoni, O., "Rough large deviation estimates for simulated annealing: application to exponential schedules.," *The Annals of Probability* **vol. 20**, 1109–1146 (1992).
[18] Lawson, P. R., Cotton, W. D., Hummel, C. A., Monnier, J. D., Zhao, M., Young, J. S., Thorsteinsson, H., Meimon, S. C., Mugnier, L. M., le Besnerais, G., Thiébaut, É. M., and Tuthill, P. G., "An interferometry imaging beauty contest," in [*New Frontiers in Stellar Interferometry*], Traub, W. A. and Bellingham, W. A., eds., **5491**, 886, SPIE (2004).
[19] le Besnerais, G., Lacour, S., Mugnier, L. M., Thiébaut, É., Perrin, G., and Meimon, S., "Advanced imaging methods for long-baseline optical interferometry," *IEEE J. Selected Topics in Signal Process.* **2**(5), 767–780 (2008).

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 08/30/2016 Terms of Use: http://spiedigitallibrary.org/ss/termsofuse.aspx