Improving the performances of current optical interferometers & future designs
Proceedings of Haute Provence Observatory Colloquium (23-27 September 2013)
Edited by L. Arnold, H. Le Coroller & J. Surdej

Kernel-phase for interferometry with a rich aperture

F. Martinache[1]

[1]*Laboratoire Lagrange, CNRS UMR 7293, Observatoire de la Côte d'Azur, Nice, France*

Abstract. This paper introduces kernel-phase, a framework developed to generalize and extend the application of closure-phase as it is used in the context of non-redundant aperture masking interferometry, that is compatible with a pupil of arbitrary geometry, if wavefront errors are small. This generalization has some powerful applications: large amounts of classical imaging archival AO and/or space borne data can be processed and lead to re-interpretation, in the light of this interferometric point of view of classical telescope image formation. Recent developments involving noise decorrelation procedures take kernel-phase one step closer to being the optimal observable extractable from an AO diffraction limited image or interferogram acquired from a rich interferometric array. While direct applications are so far mostly concerned with AO imaging, the ideas presented here are very relevant to interferometry at large: the framework offers a refreshing look at observing and data reduction strategies, in a manner that scales very well with the complexity of the interferometric array. The paper also shows that the linear algebra formalism used for kernel-phase allows for a very direct and computationally efficient approach to interferometric imaging.

1. Introduction

The observational success of optical interferometry relies for the most part on the properties of well defined and therefore well understood observable quantities: the visibility and the closure-phase (when three or more baselines are used simultaneously), used as proxies for measurements of local coherence of the electric field. From a finite number of well characterized visibilities and closure-phases, and without introducing too much a priori information, interferometry makes it possible to infer high-fidelity models or images of a wide variety of sources in a regime of resolution that goes beyond what is usually thought as possible when one is used to dealing with single-telescope observations.

These properties of interferometry, and the fact that it can beat the generally accepted Rayleigh resolution criterion, a regime referred to as super-resolution, have motivated its deployment back on single telescopes, where it is either called non-redundant masking (NRM) or sparse aperture masking (SAM) interferometry, first seeing-limited, and more recently used in conjunction with adaptive optics (AO).

These observations differ from most of the usual long-baseline interferometry in the all-in-one (Fizeau) recombination of a fairly large number of baselines (from tens to hundreds) that lead to interferograms looking not that different from conventional diffraction limited images, except in the structure of point-spread function (PSF) halo. The legacy of long-baseline interferometry is however very strong, and except for some more recent attempts at working in the image plane (Lacour et al, 2011), the bulk of the work is done by working on the Fourier counterpart of the image, where complex visibilities are extracted, to be processed and form squared visibilities and closure-phases, familiar to users of long baseline interferometry.

As we move toward interferometric combiners involving an increasing number of telescopes such as the VLTI 4-telescope combiners MATISSE (Lopez et al, 2012) and GRAVITY (Eisenhauer et al, 2011) or the 6-telescope combiner MIRC (Monnier et al, 2012) for the CHARA array, it seems appropriate to look back and reflect upon what single dish interferometric observations, already effectively combining up to several hundreds of apertures, can teach.

What comes out of this examination is that while it is possible to strictly adhere to the principles of long baseline interferometry, and maintain the use of canonical visibilities and closure-phase, this approach doesn't scale up well. One indeed ends up with a large number of highly correlated observables and is forced to work with a strictly non-redundant pupil that becomes cumbersome and inefficient when working with a large number of apertures.

Recent years have witnessed the introduction of a generalisation of the notion of closure-phase, a concept called kernel-phase that offers an efficient way of working with Fourier-phase information in the context of rich arrays. The framework of kernel-phase relies on a linear approximation that is for now only valid in a regime where piston errors are less than $\sim$1 radian. While the direct application of techniques presented in this paper is limited to this low piston error regime, the general ideas remain relevant to a wider set of conditions. The immediate advantage of kernel-phase over canonical closure-phase is that it can very well be extracted from Fizeau interferograms acquired with a redundant array: the technique is therefore relevant to both sparse pupil interferometry and conventional imaging. Section 2. of this paper shows how linear algebra formalism offers a convenient extension of the classical model of closure-phase suited to rich arrays, used in section 3. to introduce a generalisation of the closure-phase: the kernel-phase. Section 4. shows how kernel-phase is used to process AO data for high contrast imaging while section 5. introduces further development that shape the definition of better interferometric observables. Section 6. explores the use of kernel-phase for interferometric imaging.

2. A linear model for the Fourier-phase

The basis for the linear model leading to the definition of kernel-phase is easily grasped if one goes back to the NRM scenario and write simple equations for the phase. Fig. 1 describes the mechanism that links pupil phase and Fourier phase for a 9-hole non-redunant mask. A triangle of baselines in the pupil is highlighted along with the corresponding uv-locations in the powerspectrum. Expressions for the phase sampled at those three points are provided. By simply adding them, the piston terms cancels out: the result of this addition, called the closure-phase

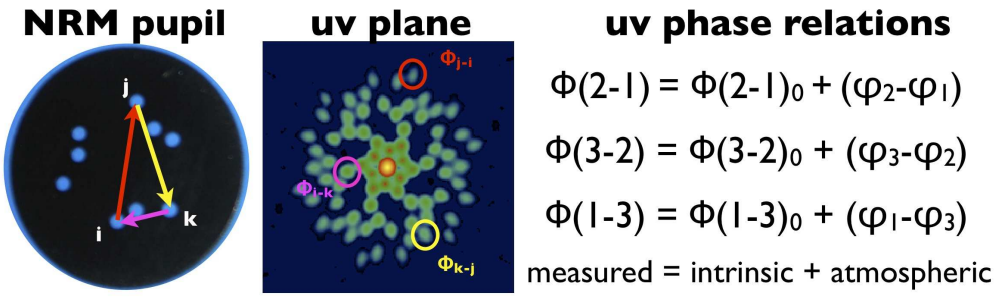

Figure 1.: *Example of closure phase relation. Superimposed on the image of the non-redundant pupil geometry shown in the left panel, is one of the possible closure triangles. The central panel shows the powerspectrum of one image acquired with one such pupil: each active region, often refered to as a splodge, is associated to a baseline in the pupil. The three splodges associated to the baselines chosen in the left panel are highlighted. On the right, are written relations for the phases of each splodge: the measured phase is the sum of a term intrinsic to the target being observed (the "true" phase), and an atmospheric term: the piston along the baseline. The reader will quickly observe that by adding these three relations together, the atmospheric term simply vanishes, leading to a new observable quantity, called the closure-phase.*

is therefore insensitive to residual pupil phase errors and contains information about the observed source only, making it a robust and powerful observable.

The 9-hole NRM used for this example allows to simultaneously measure 36 distinct phases in the uv plane. These 36 relations, all very similar to the ones highlighted in Fig. 1, can be gathered using the following matrix form:

$$\Phi = \Phi_O + \mathbf{A} \cdot \varphi, \tag{1}$$

where Φ is a 36-component vector encoding the phase sampled in the Fourier plane, Φ_O a 36-component vector encoding the true target phase information and φ a 8-component representing the instrumental pupil phase (9-1 components since one aperture is chosen as piston reference). The important element of this model is the 36×8 transfer matrix $\mathbf{A}$ that describes the way the pupil phase propagates into the Fourier plane.

For this example, $\mathbf{A}$ is essentially filled with zeros, except for two positions per row, that respectively contain $+1$ and -1. In general terms, a closure relation can be thought of as a combination of rows of $\mathbf{A}$ that produce a zero vector. The closure-phase is a special case of linear relation, that simply adds together selected rows of $\mathbf{A}$ to give the zero-vector. More complex relations involving more than three rows of $\mathbf{A}$ can however be produced.

Nevertheless, the total number of independent relations remains constant, and is exactly 28 in this scenario. Closure relations form a basis for the left-hand null space (or Kernel) of $\mathbf{A}$. These relations can be gathered into a left-hand operator $\mathbf{K}$ that acts on $\mathbf{A}$ so that:

$$\mathbf{K} \cdot \mathbf{A} = \mathbf{0}. \tag{2}$$

Although already abstract, the canonical closure-phase is a convenient concept that is easy to grasp. Moreover, it is a natural choice and the only possible closure relation when the pupil is made of only three sub-apertures. In practice for a baseline-rich pupil like the 9-hole case used as an example, closure-phase alone is not the best solution, as similar triangles in the pupil do exhibit correlated closure-phases. The formalism introduced in this section will allow to directly produce relations that produce observables containing decorrelated signals.

3. Kernel-phase as a generalized closure-phase

What may at first be perceived as a complicated sleight of hand (the matrix form) to refornulate an otherwise simple and elegant idea (the closure-phase) reveals its true power when one looks beyond the usual strict non-redudant scenario of interferometry. For a rich arbitrarily shaped interferometric aperture such as the one of an unmasked telescope in the context of NRM-interferometry, lies an additional complication: baselines in the pupil are highly redundant and the useful interferometric signal Φ_O finds itself buried under multiple phase error contributions, resulting into a fairly complex (non-linear) expression for the k^{th} component of the Fourier phase vector:

$$\Phi^k = \Phi_0^k + \text{Arg}(\text{e}^{\text{j}\Sigma_\text{i}\Delta\varphi_\text{i}}), \tag{3}$$

where i is an index to keep track of the r identical baselines in the pupil, contributing to the same region of the uv plane. With a good wavefront correction (this approach has been validated both on space-based and ground-based data) - eq. 3 can be linearized as follows:

$$\Phi^k = \Phi_0^k + \frac{1}{r}\sum_i \Delta\varphi_i, \tag{4}$$

and the entire problem can again be written in the matrix form of eq. 1, with a modified transfer matrix. In this more general (redundant) scenario, each row of $\mathbf{A}$ now contains more than just two non-zero values. Whereas canonical closure-phases can no longer be extracted from data acquired under such conditions, it is still possible to find a left-hand operator $\mathbf{K}$ that verifies eq. 2.

While it is possible to identify by hand, friendly looking relations very much in the spirit of closure-phases if the pupil geometry is not too complex, it quickly becomes difficult when the transfer matrix $\mathbf{A}$ quickly gets quite big. The linear form of eq. 2 however enables the use of powerful tools of linear algebra, and a very efficient way of building the operator $\mathbf{K}$ is to calculate the singular value decomposition (SVD) of $\mathbf{A}$.

Among its mony other properties, the SVD indeed explicitly constructs an orthonormal basis for the right and left-hand side null-spaces of a matrix. The total number of kernel-phase relations n_K is given by the number of zero singular values of $\mathbf{A}$. The SVD of $\mathbf{A}$ writes as: $\mathbf{A} = \mathbf{U\Sigma V^T}$, where $\mathbf{\Sigma}$ is a diagonal matrix containing the singular values of $\mathbf{A}$, and $\mathbf{U}$ and $\mathbf{V^T}$ are unitary matrices.

One of the best possible set of kernel-phase relations can be found in the columns of $\mathbf{U}$ that correspond to zeros on the diagonal of $\mathbf{\Sigma}$. They form an orthonormal basis for the left null-space (or kernel) of the phase transfer matrix, hence the name kernel-phase.

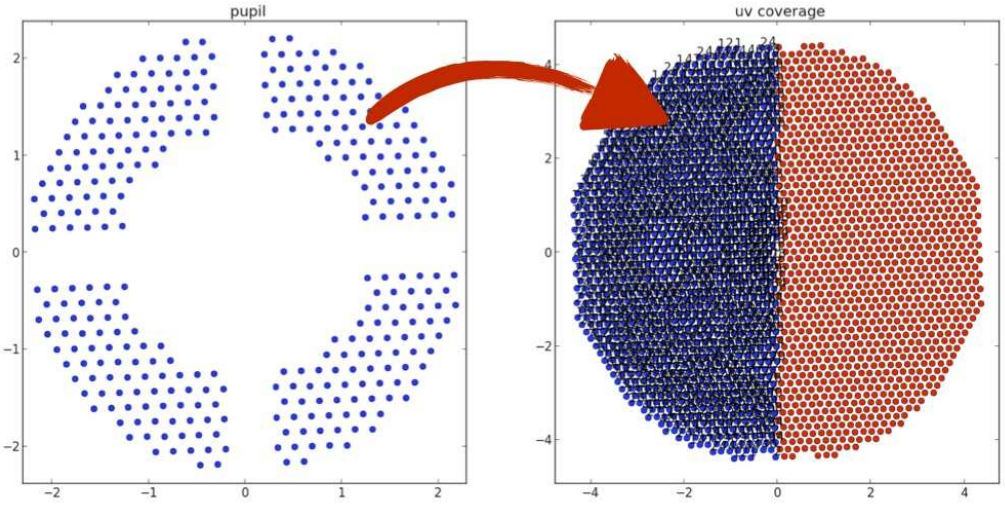

Figure 2.: *Example of discrete model used to construct the operator* **A**. *Left: discrete model of the instrument pupil, here following a regular hexagonal grid. Right: resulting distribution of spatial frequencies sampled by this geometry.*

Gathered into the operator **K**, these relations are then applied to the phase measured in the Fourier plane Φ, to extract information about the target of interest that is immune to residual instrumental phase errors:

$$\mathbf{K} \cdot \Phi = \mathbf{K} \cdot \Phi_O. \tag{5}$$

4. Kernel-phase in practice

In practice, finding the phase transfer matrix **A** requires to build a discrete model of the pupil used to acquire the data. Fig. 2 shows the example of a model used to describe the "medium cross pupil" of the Palomar Hale Telescope PHARO instrument. A good discrete model requires a regular grid pattern, whose density is representative of the continuous pupil.

The model shown in Fig. 2 decomposes the telescope pupil into 332 interferometric apertures that map onto a 1128 distinct sample points in the Fourier domain. The SVD of the resulting 1128×332 operator **A** reveals that using this model, 962 kernel-phases can be extracted from any single image, assuming that it is at least Nyquist-sampled, which means that 85 % (962 kernels out of a total of potentially available 1128 phase samples) of the phase information is directly recoverable.

Once the paving of the uv-plane and the matching kernel-phase relations are identified, they are saved in a template and used for extracting the phase information from the data. Before being Fourier-transformed, frames undergo traditional dark subtraction and flat-fielding procedure. Additionally, to limit the impact of detector readout noise, the data can can be windowed, for instance with a "super-Gaussian" $(\exp -(r/r_0)^4)$ radial profile. After the frame is Fourier-transformed, the phase is sampled at the relevant (u, v) coordinates and assembled into the vector Φ. Assuming that the data is at least Nyquist-sampled

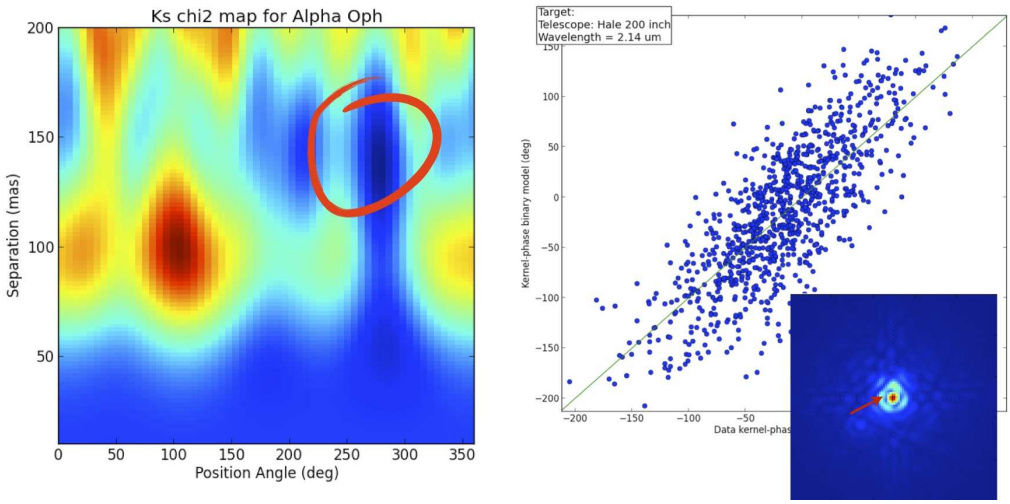

Figure 3.: *Example of kernel-phase result obtained on ground based AO data. Left: map of the χ^2 in the position angle - angular separation space for α-Ophiucus, observed with PHARO at the Palomar Hale Telescope, using the model shown in Fig. 2. A red circle highlights the location of the minimum χ^2. Right: correlation plot between the kernel-phase data and the binary model for the corresponding location in the χ^2 space. The image snippet shows (red arrow) that the 30:1 companion is apparently invisible, hidden under the first diffraction ring.*

allows all spatial frequencies to be extracted. Kernel-phase observables $\mathbf{K}\Phi$ are constructed using the pre-determined relations for each frame. Multiple frames on a given target and/or the availability of frames acquired on single stars allow further characterization of the Ker-phase data, using statistics and/or additional calibration.

Fig. 3 showcases the result of such data analysis, using the model of the PHARO medium cross pupil of Fig. 2, and applying to actual AO-corrected data acquired with this instrument. The target, α-Ophiucus, is a well known binary with a well characterized, eccentric orbit with an 8.6 year period (Hinkley et al, 2011). The kernel-phase analysis of multiple PHARO frames acquired in the K-band revealed the presence of the 30:1 contrast companion at a position angle 274.6°, but at an angular separation 136.1 mas ($\sim 1.5\lambda/D$), that is directly underneath the first diffraction ring.

After extraction, the kernel-phases are used as constraints in a 3-parameter binary model (separation, position angle and contrast). Conventional likelihood analysis and/or Monte-Carlo simulations provide a binary solution or contrast detection limits. The companion, undetectable in the direct image, due to variance in the PSF, is clearly detected using this approach. The position deduced from the binary model fit is in very good accordance with the ephemerides of the orbit.

5. Beyond kernel-phase

By construction, the different signals encoded in kernel-phases form an orthonormal basis, that makes them ideal linearly independent entries for a parametric model such as the one used in the previous example, or a more general image reconstruction software. In practice, this construction however does not guarantee statistical independence. If not accounted for, noise processes affecting the image (where the detection process really happens) can very well lead to correlated Fourier-phases, that will result in correlated observables. In addition to building observables that linearly independent in terms of signal, it is therefore also necessary to make them statistically independent in terms of noise.

This is the observation made by Ireland (2013), who proposes to further improve the observables, by carefully computing the associated covariance matrices for the visibility and the Fourier-phase. Several noise processes are expected to lead to correlated Fourier-phases: photon noise, detector readout noise or lag in the AO correction; are the first of a potentially long list.

It is possible to determine this covariance matrix using simulations or formal calculations for the best understood phenomena. In practice however, the combined effect of all sources of noise can be taken into account by observing a point source calibration star, right before or after the target of interest, under observing conditions (seeing, elevation, source magnitude and color, AO parameters) as close as possible to the ones experienced on the target of interest. The kernel-phases recorded on a point-source should all average to zero, but the statistical properties of the Fourier-phase recorded in multiple successive frames will inform about the correlation properties of the different noise processes.

If one calls this experimental covariance matrix for the Fourier-phase $\mathbf{C}_F = \mathrm{cov}(\Phi)$, the covariance matrix for the kernel-phase $\mathbf{C_K}$ writes as:

$$\mathbf{C}_K = \mathrm{cov}(\mathbf{K}\Phi) = \mathbf{K} \cdot \mathrm{cov}(\Phi) \cdot \mathbf{K}^{\mathrm{T}} = \mathbf{K} \cdot \mathbf{C}_{\mathrm{F}} \cdot \mathbf{K}^{\mathrm{T}} \tag{6}$$

The covariance being a square, positive semi-definite matrix, it can be diagonalized using a form simpler than that of the SVD:

$$\mathbf{C}_K = \mathbf{S} \cdot \mathbf{D} \cdot \mathbf{S}^{\mathbf{T}}, \tag{7}$$

where $\mathbf{S}$ is a unitary matrix and $\mathbf{D}$ a diagonal matrix containing the eigen values of $\mathbf{C_K}$. The important element of this decomposition is $\mathbf{S}$, which allows to turn previously correlated kernel-phases into a set of statistically independent kernel-phases θ:

$$\theta = \mathbf{S} \cdot \mathbf{K} \cdot \Phi. \tag{8}$$

Using photon noise simulations on kernel-phase in the context of high-contrast detection, Ireland (2013) shows that this simple procedure improves the contrast detection limits in a very convincing manner. Kernel-phase, as initially described in (Martinache 2010), exhibits a fairly uniform contrast detection limit beyond a separation greater than λ/D, which is unlike any standard imaging contrast detection limits rapidly increases as a function of angular separation.

After diagonalizing the effect of photon noise on kernel-phases, using the statistically independent observables introduced in Eq. 8, Ireland (2013) show

that contrast detection limits increase as a function of angular separation, and exhibit an overall performance superior over correlated kernel-phases.

6. Kernel-phase imaging

Most of work so far done with kernel-phase extracted from AO data has been applied to high contrast detection of binary sources. This application uses the χ^2-minimization approach of a 3-parameter model (angular separation, position angle and contrast). Results of this technique can be found in (Martinache, 2010; 2012) and (Pope et al, 2013).

A binary model search is a very well constrained problem, with only three degrees of freedom for a generally large numbe of kernel-phases. Here we look into the application of a less constrained problem: a more generic imaging test case, like what is routinely done in radio interferometry and can be achieved on some bright sources with long baseline optical interferometry. The Van-Cittert Zernike theorem relates the brightness distribution of the source being observed to measurements of the coherence of the electric field in the uv-plane, itself estimated from the visibility and/or the phase of the interference function.

The quality of the brightness distribution map reconstructed from the interferometric measurements greatly depends on the density of the uv-coverage offered by the array. For a given number of apertures in a sparse geometry, the richest coverage is obtained when a non-redundant pattern is used. Geometries used for NRM-interferometry typically rely on the designs established by Golay (1971) that use three-fold symmetry to produce a compact and dense uv-coverage.

Yet there is a limit to how rich the uv coverage can get while maintaining non-redundancy in a finite spectral bandwidth when the array footprint needs to be contained within say a circle. We have seen that kernel-phase makes strict non-redundancy a non-necessary requirement. A comparative study of several pupil geometries (Martinache, 2012) has shown that even when they provide identical uv-coverage, some configurations do provide a better phase information recovery-rate, defined as the ratio of the number kernel-phase n_K and the number of baselines n_{UV}. A good (and often exact) estimate of the number of kernel-phases that can be extracted from an array made of n_A apertures is:

$$n_K = n_{UV} - \frac{n_A}{2}, \tag{9}$$

indicating that for a given uv-coverage, the array with the smallest number of aperture will provide the highest phase information recovery rate: an annular geometry seems to provide an optimal. The results presented in this section use the anticipated geometry of the thirty meter telescope (TMT) primary mirror for a test scenario. The sampling of the uv plane is made so as to produce baselines that match the hexagonal grid of the 492 segments making the primary. Keeping only the outermost 78 segments provides access to the same uv-coverage and results in a total phase information recovery rate ~ 96 %, to be compared to the ~ 75 % the full pupil would give.

One very convenient consequence of expressing the phase relations in terms of linear algebra (cf. eq. 1) is that the matrix form allows the construction of a pseudo-inverse to the kernel-phase operator $\mathbf{K}^+$.

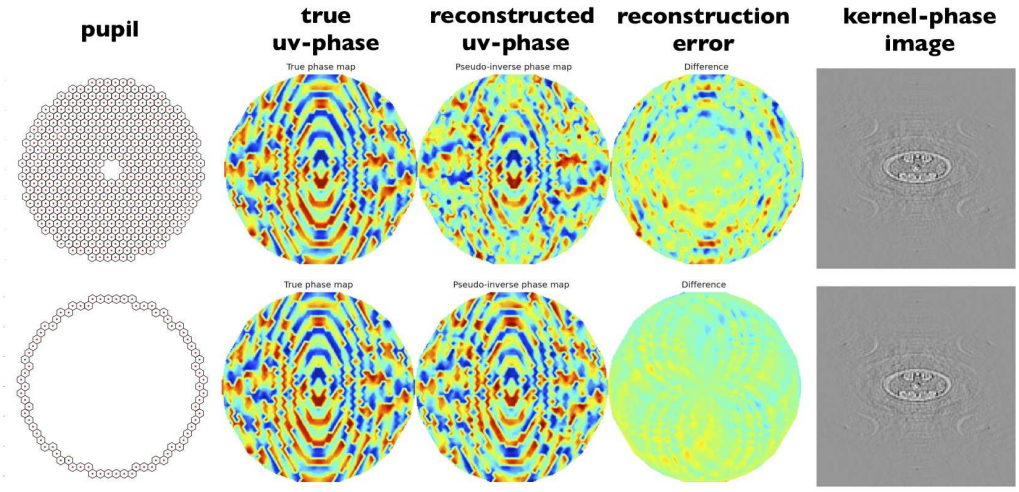

Figure 4.: *Kernel-phase image reconstruction experiment, comparing the phase (and resulting image) reconstruction achieved by a full pupil (top row) to the one achieved by annulus (bottom row). Each row successively shows (from left to right): the pupil, the true uv-phase-map of a simulated source (identical in both cases), the uv-phase-map reconstructed from kernel-phases alone, the reconstruction error and the kernel-phase image, determined from a direct inverse Fourier Transform of the reconstructed phase-map. Note: all six phase-maps use the same color-scale.*

Even for a full-aperture, $\mathbf{K}$ preserves a large fraction of the original uv-phase information while clearing it of all residual wavefront error. The use of a pseudo-inverse $\mathbf{K}^+$ is therefore expected to yield direct access to a reasonably accurate representation of the true object uv-map Φ'_O. The uv-phase estimate:

$$\Phi'_O = \mathbf{K}^+ \cdot \mathbf{K} \cdot \Phi, \tag{10}$$

can directly be used as input for an interferometric imaging program. Fig. 4 compares the phase reconstruction capability of this approach on the full pupil and the annulus. We can verify that both geometries manage to preserve a large fraction of the uv-phase information: while some differences can be appreciated in the fidelity of the pseudo-inverse by comparing the reconstruction error maps, the features of the images determined from direct inverse Fourier transform of the phase-maps, differ very little and are very close to what a reconstruction relying on a perfect knowledge of the phase would achieve. Admittedly, the simulation is simplistic. It nevertheless confirms that the pseudo-inverse approach is sound. The $\mathbf{K}^+ \cdot \mathbf{K}$ operator can easily be used as a way to bypass the abstract kernel-phase intermediate as a direct way to work in uv-phase space. A small fraction of the uv-phase is lost in the process, but the instrumental phase errors are filtered out. These cleaned uv-phase can in turn be related to the source brightness distribution map, like what is shown in Fig. 4.

7. Conclusions

This paper has introduced the idea of kernel-phase, a generalization of the more usual notion of closure-phase that can be used in the context of an all-in-one Fizeau combiner, regardless of the array geometry. Kernel-phase relies on a simple linear model that describes the way instrumental phase errors propagate into the uv-plane, and pollute the information relevant to the target of interest. This formal approach allows for the deployment of powerful computational tools: singular value decomposition and pseudo-inverses, which combined with further noise decorrelation procedures, allow for optimal information extraction strategies, relevant to a wide range of applications: from high contrast detections to general interferometric imaging.

To keep things going in one direction, this paper has voluntarily excluded the applications of the linear model relevant to wavefront sensing. Interested readers should refer to (Martinache, 2013) for further details about this complementary problem that isn't concerned with the kernel- but with the eigen-phases of the phase transfer operator $\mathbf{A}$.

It should finally be mentioned that the discussion has been restricted to the study of the phase in the uv-plane, and therefore excluded all consideration for the amplitude of the complex visibility. Yet, one knows that simultaneously combining four baselines and more should enable the determination of a closure-amplitude in addition to several kernel-phases. One wonders whether an adaptation of the formalism used here is possible so as to propose a comparable treatment for the amplitude. While this question is being investigated, it is not obvious whether such a treatement is possible.

References

Eisenhauer, F. et al, 2011, Msngr, 143, 16

Golay, M., 1971, JOSA, 61, 272

Hinkley, S. et al, 2011, ApJ, 726, 104

Ireland, M. J., 2013, MNRAS, 433, 1718

Lacour, S., Tuthill, P., Amico, P., Ireland, M., Ehrenreich, D., Huelamo, N. & Lagrange, A. M., 2011, A&A, 532, 72

Lopez, B. et al, 2012, SPIE, 8445, 84450R

Martinache, F., 2010, ApJ, 724, 464

Martinache, F., 2012, SPIE, 8445, 844504

Martinache, F., 2013, PASP, 125, 422

Monnier, J. D., 2012, SPIE, 8445, 84450Y

Pope, B. et al, 2013, ApJ, 767, 110