

The 2016 Interferometric Imaging Beauty Contest

J. Sanchez-Bermudez^a, E. Thiébaud^b, K.-H. Hofmann^c, M. Heininger^d, D. Schertl^e, G. Weigelt^f, F. Millour^g, A. Schutz^h, A. Ferrariⁱ, M. Vannier^j, D. Mary^k, and J. Young^l

^aMax-Planck-Institut für Astronomie, Königstuhl 17, Heidelberg 69117, Germany

^bCentre de Recherche Astrophysique de Lyon, Observatoire de Lyon, Avenue Charles André 9, Saint Genis Laval Cedex F-69561, France

^{c,d,e,f}Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, Bonn 53121, Germany

^{g,h,i,j,k}Observatoire de la Côte d’Azur, Boulevard de l’Observatoire, Nice 06300, France

^lAstrophysics Group, Cavendish Laboratory - University of Cambridge, JJ Thomson Avenue, Cambridge CB3 0HE, United Kingdom

ABSTRACT

Image reconstruction in optical interferometry has gained considerable importance for astrophysical studies during the last decade. This has been mainly due to the improvements of the imaging capabilities of the existing interferometers and of the expected facilities in the coming years. However, despite the advances made so far, image synthesis in optical interferometry is still an open field of research. Since 2004, the community has organized a biennial contest to formally test the different methods and algorithms for image reconstruction. In 2016, we celebrated the 7th edition of the “Interferometric Imaging Beauty Contest”. This initiative represented an open call to participate in the reconstruction of a selected set of simulated targets with a wavelength-dependent morphology as they could be observed by the 2nd. generation of VLTI instruments. This contest represents a unique opportunity to benchmark, in a systematic way, the current advances and limitations in the field, as well as to discuss the future lines of action. In this contribution, we summarize: (a) the rules of the 2016 contest; (b) the different used data sets and the selection procedure; (c) the methods and results obtained by each one of the participants; and (d) the metric used to select the best reconstructed images. Finally, we named Karl-Heinz Hofmann and the group of the Max-Planck-Institut für Radioastronomie as winners of this edition of the contest.

Keywords: Optical interferometry, image reconstruction, GRAVITY/VLTI, MATISSE/VLTI, Interferometric Beauty Contest

1. INTRODUCTION

Interferometric image reconstruction at optical/infrared wavelengths is a novel technique that has been developed during the last ten years. The used observables in optical interferometry are the squared visibilities, the bispectrum and the chromatic differential phases, contrary to radio interferometers that work directly with the amplitude of the visibilities and their phases. This difference in the operation of both techniques together, with the sparseness in the $u - v$ plane, place strong constraints to the image reconstruction process. Particularly, it makes difficult the error estimation (artifacts and/or noise) when infrared images are reconstructed using radio interferometry algorithms (e.g., CLEAN).

Therefore, at the 2001 meeting of the International Astronomical Union, the Working Group on Optical Interferometry conceived the idea to compare different algorithms to reconstruct infrared interferometry images. As a consequence, the community released in 2004 the first “Interferometric Beauty Contest”. The main motivations for this initiative were the following: (1) encourage the use of the OI Exchange Format, identify problems in its definition, and revise it as necessary; (2) engage the interferometry community in a formal assessment of

Further author information: (Send correspondence to J Sanchez-Bermudez)
J Sanchez-Bermudez.: E-mail: jsanchez@mpia.de

existing software; (3) encourage the development of new tools tailored to the needs of optical interferometry.¹ Since then, six editions of this contest have been released. Five of them working with simulated data sets²⁻⁶ and one more with real data obtained with PIONIER/VLTI.⁷

This year, apart from the previous motivations, the objectives of the contest were: (i) to advertise the imaging capabilities of the new/upcoming instruments at the VLTI and (ii) to test the capabilities of the existing software and techniques to perform chromatic reconstructions. Therefore, for this call, the selected data correspond to simulated observations of two astrophysical objects; one of them simulated with the expected GRAVITY capabilities⁸ in medium resolution mode; and a second one with MATISSE^{9,10} in low resolution mode. The delivered observables were squared visibilities, closure phases, Fourier/differential phases and the spectrum of each source. The data were stored and delivered in standard oifits¹¹ files to all the participants of the initiative. This edition of the contest was subject to the following rules:

- All the reconstructed images were submitted before May 30, 2016.
- There were no restrictions on the number of participants that can form a team to participate in the contest.
- The reconstructed images were sent directly by e-mail to any of the members in the organizing committee.
- All the submitted images were accompanied by a description of the used method, including the pixel scale, the type of reconstruction (chromatic vs monochromatic), the used regularization, number of iterations, convergence criteria, goodness of the fit and the used software (see Sec. 3).
- The winner was decided based on a direct comparison with the models using quantitative metrics described in Sec. 4.

2. THE DATA SETS

2.1 GRAVITY data: A cluster of massive stars

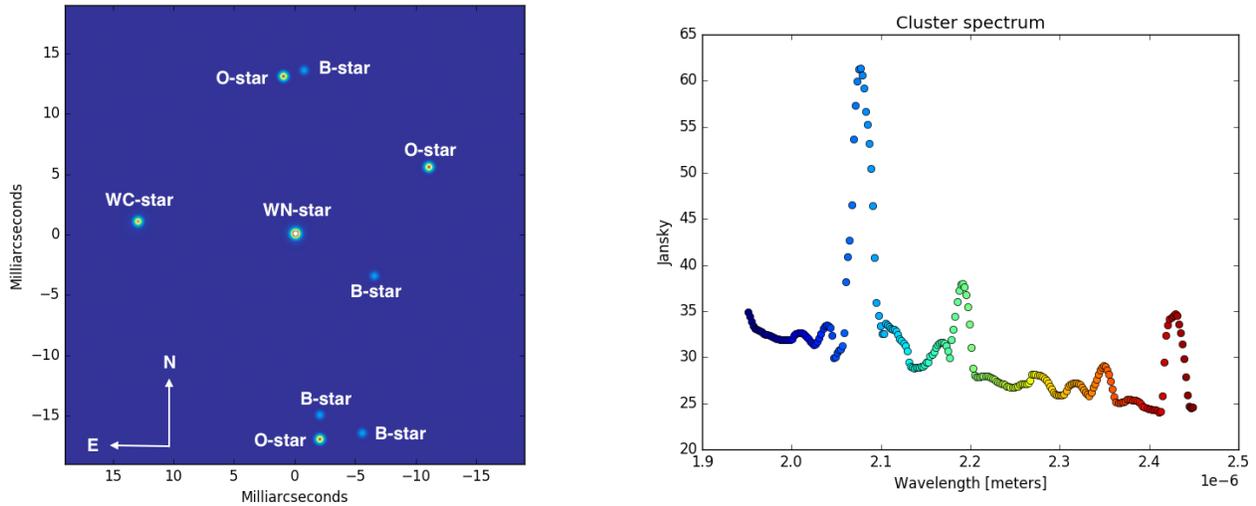


Figure 1. **Left:** Mean image of Object No.1 used for the Beauty Contest. The target corresponds to a cluster of stars within a FOV of ~ 30 mas. **Right:** Total spectrum of the simulated target sampled at a resolution of $R \sim 500$ (i.e., GRAVITY Med-Res mode).

The simulated Object No.1 corresponds to a cluster of massive stars. This object consists in a collection of the following stars: 1 Wolf-Rayet-Nitrogen star, 1 Wolf-Rayet-Carbon star, 3 O-stars and 4 B-stars. All of them spread over a field-of-view (FOV) of ~ 30 mas. The prototypical spectra of the Wolf-Rayet and O-stars (in which several broad emission lines are observed) were obtained from: <http://kookaburra.phyast.pitt.edu/>

hillier/web/CMFGEN.htm. These models were computed using the CMFGEN code.¹² For the B-stars, a simple Kurucz atmosphere model was used to estimate their spectra. The spectrum of each one of the massive stars were scaled, assuming a distance of 1 kilo-parsec. To increase the difficulty of the simulated object, a constant background was added to the data set. Additionally, the angular size of some of the WR-stars were modified in order to look like extended sources at wavelengths that corresponds to some of the most prominent spectral features. It is important to highlight that the objective of this simulated object was not to depict devotedly the physics of the used stars, but to create an object with spatial and spectral changes that could be recovered with the different software and methodology. The left panel in Figure 1 displays a mean image of the simulated Object No.1 while the second panels shows its simulated spectrum.

The observational setup used to construct this data set assumed three nights of observations with the ATs, using each night a different configuration. The triplets favored small (A0-B2-C1-D0), medium (D0-G2-J3-K0) and large (A0-G1-J2-J3) telescope configurations. First panel in Figure 3 displays the simulated $u-v$ coverage of the aforementioned setup, in which four pointing positions are appreciated per baseline. The delivered observables consisted in square visibilities, closure phases and Fourier phases. The remaining three panels in Figure 3 display each one of the simulated observables. Notice how the squared visibilities clearly show the typical cosine signatures in the Fourier plane that correspond to several point-like sources. It is important to highlight that around $\sim 57\%$ of the total flux was over-resolved for all the sampled spatial frequencies. This effect was caused due to the continuous background added to the data.

2.2 MATISSE data: A prototypical-planetary disk

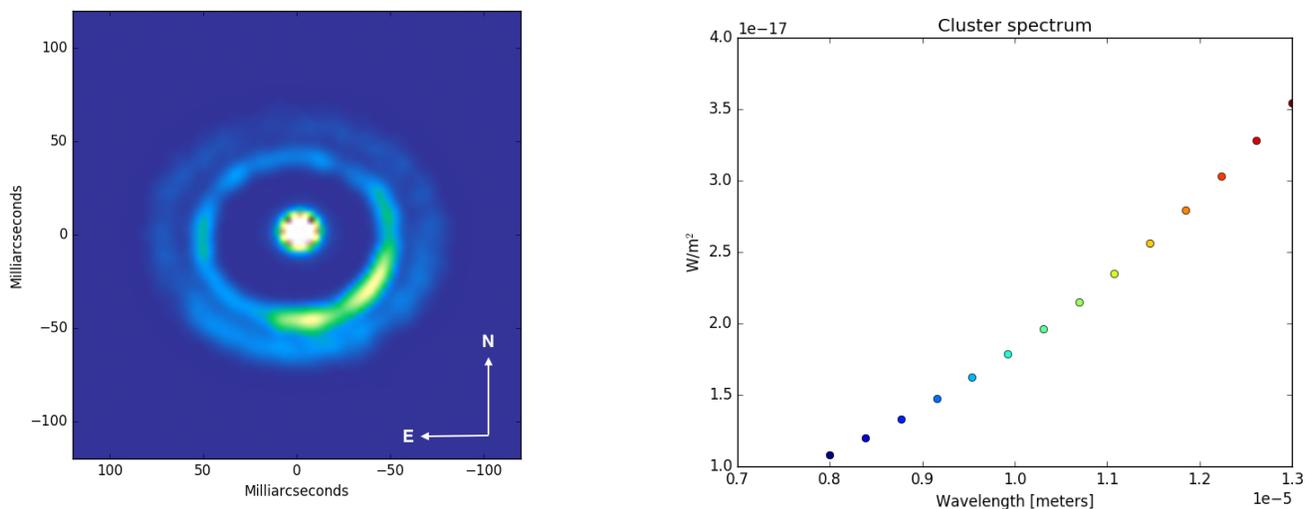


Figure 2. **Left:** Mean image of Object No. 2 used for the Beauty Contest. The target corresponds to a prototypical-planetary disk with a gap in its structure. **Right:** Total spectrum of the simulated target sampled with a resolution of $R \sim 35$ (i.e., MATISSE Low-Res mode).

Since the study of prototypical-planetary disks is one of the MATISSE key-science projects, Object No.2 consisted in a radiative transfer simulation of a circumstellar disk. The model includes a central star (surrounded by a large amount of dust) and a disk for which one gap was clearly identified. This gap was supposed to be carved by a forming planet. A physical scenario that is plausible for the current planet formation models.¹³ Furthermore, the brightness distribution across the rim was not uniform with a clear bright-spot at a position angle of $\sim 220^\circ$ East of North and the disk-shape was elongated in the East-West direction. The spectrum of this target was simpler than the one of the cluster. Since the two morphological components of the object (central source + disk) assumed a simple dust-thermal emission with a Black-body profile, the spectrum emission increases steadily from the shortest to the longest wavelength. First panel in Figure 2 displays a mean image of the simulated object while the second panel shows the simulated spectrum of the object.

The observational setup for this data was similar to the one used for the GRAVITY data. With three simulated nights using three different configurations of the Auxiliary Telescopes. Figure 4 displays the simulated $u-v$ coverage, which is significantly richer than the one obtained for Object No.1. The interferometric observables are also shown in the aforementioned Figure 4. Notice how the squared visibilities trace a resolved object with an angular size of ~ 100 mas. The simulated closure phases are considerably noisier, which increases the difficulty of the reconstruction. The provided differential phases use the AMBER convention in which the reference channel corresponds to an average of all the channels except the working one. This means that the reference channel was variable across the bandpass.

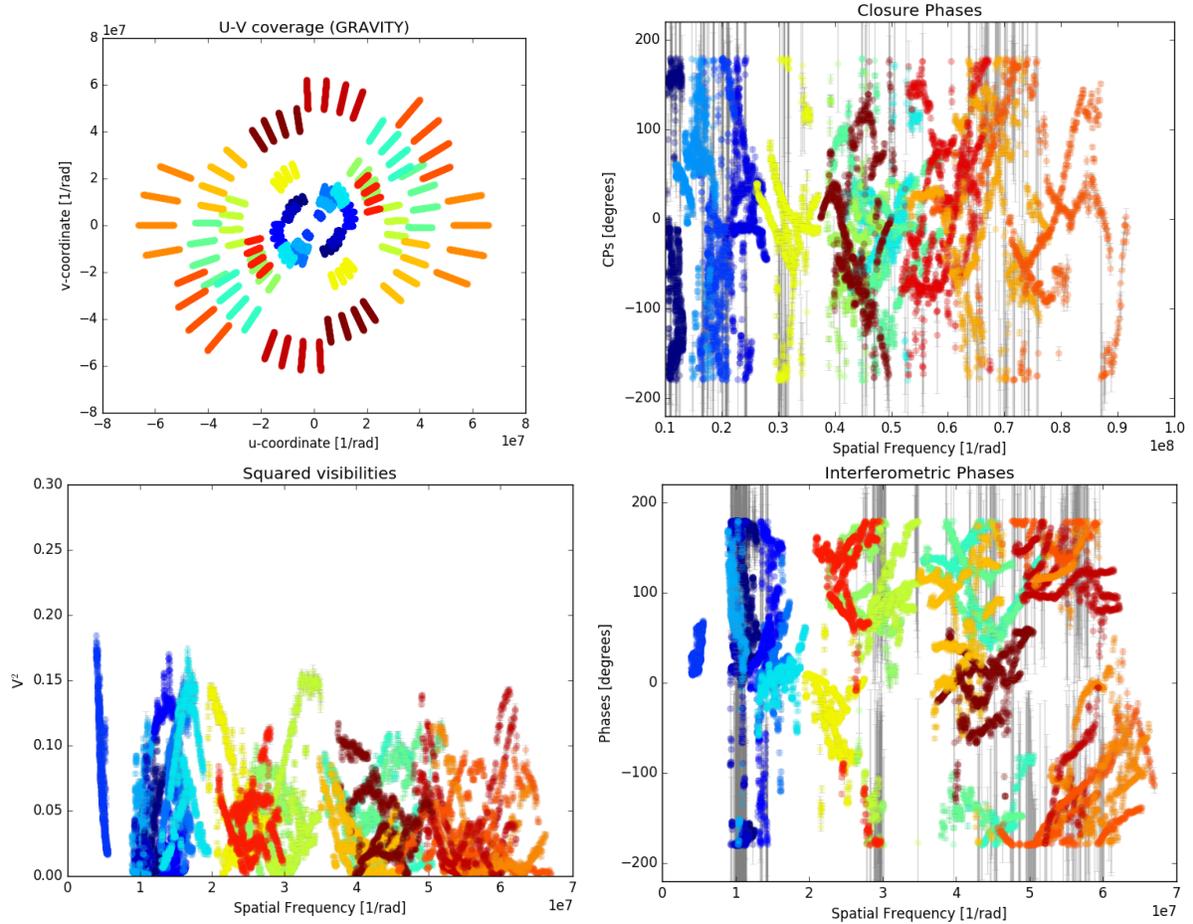


Figure 3. The upper-left panel corresponds to the simulated $u - v$ coverage for the GRAVITY data. The upper-right, lower-left and lower-right panels display the simulated closure phases, squared visibilities and Fourier phases, respectively. Different colors correspond to different baselines or triangles.

3. THE CONTRIBUTIONS

3.1 John Young with BSMEM

The BSMEM (BiSpectrum Maximum Entropy Method) software was first written in 1992 to demonstrate image reconstruction from optical aperture synthesis data. It has been extensively enhanced and tested since then. The code used for this year’s contest entry is essentially identical to that employed for the 2010, 2012 and 2014 contests. A modified version of BSMEM has been developed as one of several alternative algorithms able to be controlled by a new Graphical User Interface, as part of an OPTICON Joint Research Activity to provide

user-friendly imaging tools (Young et al., these proceedings). The modified version is still in beta and was not used for this submission.

The BSMEM algorithm applies a fully Bayesian approach to the inverse problem of finding the most probable image given the evidence, making use of the Maximum Entropy approach to maximize the posterior probability of an image. BSMEM is available free-of-charge to the scientific community on submission of the academic licence agreement at <http://www.mrao.cam.ac.uk/research/OAS/bsmem.html>.

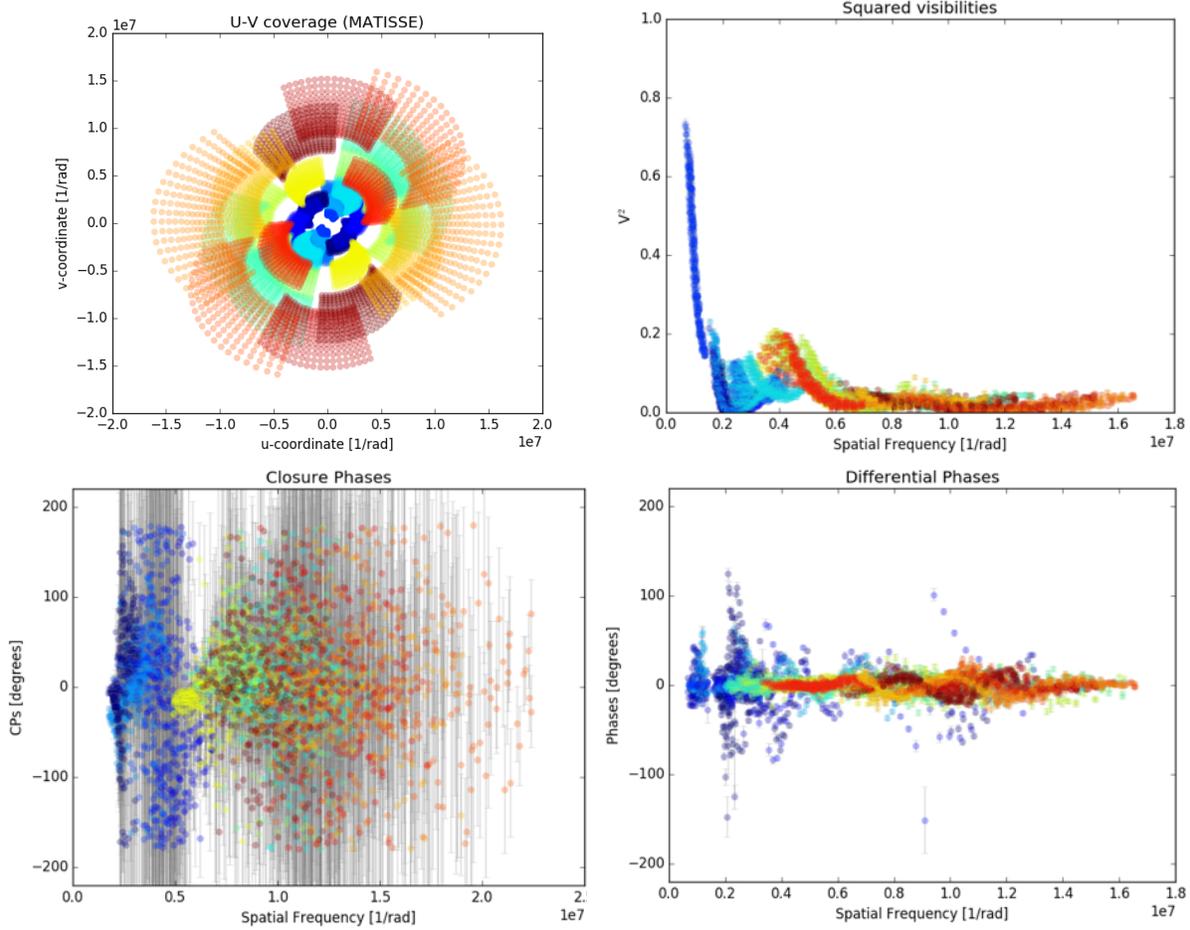


Figure 4. The upper-left panel corresponds to the simulated $u-v$ coverage for the MATISSE data. The upper-right, lower-left and lower-right panels display the simulated closure phases, squared visibilities and differential phases, respectively. Different colors correspond to different baselines or triangles.

BSMEM uses a trust region method with non-linear conjugate gradient steps to minimize the sum of the logarithm of the likelihood (i.e., χ^2) of the data given the image and a regularization term expressed as the Gull-Skilling entropy:

$$Entropy = \sum_k (I_k - M_k - I_k \log(\frac{I_k}{M_k})) \quad (1)$$

where I_k is the current iteration of the reconstructed image and the model image M_k is usually chosen to be a Gaussian, a uniform disk, or a delta-function centered in the field of view. This conveniently fixes the location of the reconstructed object (the bispectra and powerspectra being invariant to translation). This type of start model also acts as a support constraint by penalizing the presence of flux far from the centre of the image.

Due to a lack of time, only Object No.2 was attempted. The current algorithm can only reconstruct grey images, so a script was written to select each wavelength channel in turn and reconstruct it separately, then combine the reconstructed images into a FITS cube. Initial results exhibited a surprising degree of variation between nearby channels, in contrast to the smooth variation of the integrated spectrum with wavelength. BSMEM automatically estimates the hyperparameter α that controls the weighting of the entropic prior relative to the likelihood, and the estimated α values were found to vary significantly between wavelength channels despite their similar $u - v$ coverage and signal-to-noise. Therefore a single α value was imposed for all of the wavelength channels, which gave more consistent results. The final χ^2 values ranged from 0.9 to 1.3, and between 20 and 23 iterations were needed to converge. Differential phases were not used for the reconstruction, and thus the low signal-to-noise of the closure phases means that the azimuthal flux distribution was not particularly well constrained.

In the submitted FITS image cube, North is up and East is to the left when the standard display convention for FITS images is followed. The image scale is 2 milliarcseconds per pixel. The third axis is wavelength, in ascending order. The object’s position was free to vary between wavelength channels (albeit partially constrained by the prior image), so this degree of freedom will need to be taken into account when assessing the accuracy of the reconstruction.

3.2 Kar-Heinz Hofmann et al. with IRBis

This team uses the IRBis software¹⁴ for image reconstruction. IRBis is coded in C and is part of the ESO MATISSE pipeline. This software can reconstruct monochromatic images from squared visibilities and closure phases. The latest update of IRBis also allows the reconstruction of images from complex visibilities, i.e., from Fourier phases and amplitudes.

Because the IRBis software reconstructs monochromatic images, an image was reconstructed in each spectral channel of the Beauty Contest data separately. First, an image was reconstructed in the spectral channel with the longest wavelength, because in this spectral channel, the interferometric data have the best quality. The start image for the reconstruction in this spectral channel was a Gaussian fitted to the visibilities. For the reconstruction in the next adjacent spectral channel, the start image was the best reconstruction obtained in the adjacent channel with a longer wavelength. After the first reconstruction run, a second one was started using the results of the first run as start images. The second run is described as follows, because it differs between the two objects.

3.2.1 Object 1

This data set contains 250 spectral channels in the wavelength range from 1.951 to 2.449 μm . The reconstructed field-of-view is 60 mas and the grid used has 64x64 pixels, which leads to a pixel scale of 0.9375 mas. The applied regularization function is the “Edge Preserving” function,¹⁴ which is similar to the “Total Variation” regularization preserving important details, such as edges, in image regions with strong intensity variations, and it also is similar to the “pixel difference quadratic” regularization¹⁴ enforcing smoothness in regions with low intensity variations.

From the first reconstruction run, performed as described above, the unweighted sum over the reconstructions of all 250 spectral channels was derived. This sum was used as start image in each spectral channel in the second run. The reconstructions of the 250 spectral channels fit very well to the measured interferometric data with an average χ^2 of about 1.5. The target seems to be a star cluster with (at least) 5 bright stars with wavelength-dependent brightness. For example, in the wavelength range between 2.305 and 2.339 μm , the easternmost star (North is up and East to the left) is much brighter than all other 4 stars, whereas in many other spectral channels, the 5 stars have similar brightness. Two additional, much fainter stars are possibly detected: the first one is located in the South-West of the bright star in the middle of the cluster with a separation of ~ 7 mas, and the second one is to the West of the southernmost star with a separation of ~ 4 mas.

3.2.2 Object 2

The reconstructed field-of-view was of 400 mas and the grid used has 128x128 pixels, which leads to a pixel scale of 3.125 mas. The applied regularization function is the “Maximum Entropy” function enforcing smoothness. The best reconstructions, from the first recovery run, were used as start and prior images in the second run. The reconstructions of the 14 spectral channels fit well to the measured interferometric data. The average χ^2 values of the squared visibilities and closure phases are 4.1 and 1.3, respectively. The reconstructions show a bright central source in the middle of two concentric ring-like structures with a dark gap between the two rings. The outer ring is fainter than the inner one. In the South-East quadrant, the rings are brighter than in the North-West. Between the spectral channels at 9.15 and 11.84 μm , a weak extended structure becomes visible within the gap between the central source and the inner ring. This weak structure extends to the North-West.

3.3 Florentin Millour with MiRA + Autocalibration

This participant performed his contribution using the following two approaches:

3.3.1 Contribution 1

One contribution is made with the recipes that the participant presented a few years ago with further refinements, i.e. low frequency filling, a Monte Carlo approach to treat the image artifacts¹⁵ and self-calibration^{16,17} of the differential phases. As baseline for the image reconstruction the MiRA software¹⁸ was used, which does already an excellent work at inverting the imaging problem. As a recall, the low-frequency filling adds low-frequencies to the data set by extrapolating the values of the smallest frequency data points. This enhances significantly the convergence of image reconstruction algorithms and allow one to get “better” images by artificially filling the $u - v$ plane with relevant information (assuming that the object is not or barely resolved at low frequencies).

The Monte Carlo approach (called BFMC¹⁵) varies randomly the parameters of the reconstruction and scans the parameter-space this way. The final image is the median of many images selected with an ad-hoc criterion, here $\chi^2 \leq 2 \times \chi_{\text{min}}^2$ plus a quality check on the image itself (described below). In the current case, not all the desired parameters were variate due to lack of time, and selected pre-defined values $\mu = 10000, 2000, 1000$ in sequence, and $N_{\text{iter}} = 3 \times 25$ instead of random values, but still, random start images were used.

The self-calibration of the differential phases consists in calibrating the unknown instrumental gains of the complex visibilities by making use of the visibilities and closure phases. This is done iteratively by using the reconstructed image itself as a template for the exact phases. In the reconstruction, the self-calibration process stopped after two iterations. The two reconstructions on Object No.1 and Object No.2 were made with the combination of these three methods.

For Object No.1, the participant computed 71 valid images with 0.6 mas per pixel and 128 pixels on all 250 wavelengths, all the images were centered to the brightest pixel, as the image cubes were not properly centered relative to each other. A “quality check” on the image was performed as follows: (i) a normalized median image was computed out of the centered images; (ii) this median image served as reference to compute the square differences between it and each normalized images; (iii) only the images with square differences less than twice the minimum were kept. This left 24 images out of the total. The presented image corresponds to the median of all these 24 image cubes.

From the reconstruction it was inferred that Object No.1 consisted by a set of 5 point sources atop an extended emission component. Each point source has its own spectrum, but the reconstruction process fails to get correctly the spectrum of each source, especially in the region between 2.0 and 2.1 μm (see later), where there is a large emission line in the total spectrum. This is likely due to the centering process which is not yet optimal (centering on the brightest pixel of the wavelength-median image). However, each point source has a different spectrum, likely flat for the Northern and Southern sources, with emission lines for the centre and eastern source, and likely absorption lines for the remaining western source. The extended emission exhibit strong emission lines.

For Object No.2, the participant computed 133 valid images with 6 mas per pixel and 128 pixels. The same process was applied as for Object No.1. The final image corresponds to the median of the 35 remaining “good quality” images after the aforementioned selection process. From the reconstructed images it was observed that

the object was composed of a central point source, a large (~ 100 mas diameter) asymmetric ring-like structure, and an compact (~ 20 mas) structure around the central point source. The ring is brightest and thickest to the South-East, and change shape with wavelength (likely symmetric at $8 \mu\text{m}$ and asymmetric at $13 \mu\text{m}$). The compact structure has a zero flux at $8 \mu\text{m}$ and its flux rises steadily up until $13 \mu\text{m}$. Surprisingly, the central source also has an increasing flux up to $13 \mu\text{m}$ (therefore it is not a star). The ring structure might show silicate emission in some regions (opposite to the thickest part) but this is not very obvious from the reconstruction.

3.3.2 Contribution 2

The second contribution is called “chips imaging” and uses both a classical image reconstruction process and the model-fitting tool `fit0matic`.¹⁹ The underlying idea is that an astronomical image contains a limited number of physical elements (let it be a star, a disc, commoving matter, etc.), which intrinsic shape do not change, but flux can change drastically (e.g. in an emission line, or due to different temperatures).²⁰ The approach is the following: in a gray reconstructed image (e.g. with MiRA and default parameters), one can identify the different elements of the image. For example in Object No. 1, 5 point-like sources were recognized to be distributed along a cross and some extended flux contribution. In Object No.2, based on a gray reconstruction with MiRA, the image looks something like a central point source, a Gaussian disk (around the central star) and a skewed ring. Starting from these qualitative information, the contestant built models in `fit0matic` to include these bricks, with varying spectra for each component and fitted the parameters of these models to the data by using the same simulated annealing method as done in other works.

For Object No.1, the contestant produced a 100 mas image with 256 pixels out of the model parameters. The same comments apply to the previous contribution, except that this time the region between 2.0 and $2.1 \mu\text{m}$ was properly retrieved, with an emission line on the Eastern component.

For Object 2, a 1000 mas image with 256 pixels was produced. The ring-like structure seem to be more round and symmetric in the model than in the reconstructed images. This may be due to the fact that time was short to let the fitting tool to properly converge. Nevertheless, the obtained spectra and relative fluxes match relatively well the image reconstruction made in the previous contribution.

3.4 Antony Schutz et al. with PAINTER

The PAINTER* software²¹ was used to reconstruct both Object No.1 and No.2. PAINTER (for Polychromatic opticAl INTERferometric Reconstruction) is an iterative spatio-spectral image reconstruction algorithm. It uses the flexibility of the Alternate Direction Methods of Multipliers (ADMM) algorithm in which several regularizations are implemented: spatial regularization via the use of wavelets and spectral regularization via the use of discrete cosine transform, positivity and support.

This algorithm relies on the visibilities as measures of the source power spectrum and on two types of turbulence independent phases differences. In case of low turbulence, the phase of the visibilities can be directly used. PAINTER is able to handle the phase information in addition to the closure phase, in order to ease both the spatial alignment and the spectral coherence of the reconstructed structures. Therefore the VISPHI tables were used for both objects. In each channel the squared visibility data were first normalized to the OI-FLUX information (i.e. the global spectra given in the input data). Therefore it is the coherent fluxes, together with the squared visibilities, the phases and closure phases, which are used in the reconstruction process. At the end of the reconstruction process, though, as the resulting spectrum shows slight differences with the original one, the output cube is re-normalized to the OI-FLUX.

Object No.1 was reconstructed, as advised, on a 256×256 grid and a FOV of 60 mas, and at all 250 spectral channels. The results show five punctual targets, in a diamond shape with a central source. The individual spectra of three of the individual sources vary strongly with wavelength, in correlation of the features observed on the global spectra. The two other sources have more continuous spectra.

Object No.2, was reconstructed on a 128×128 grid and a FOV of 400 mas, a central spot was clearly identify, whose brightness and size increase continuously with wavelength (together with the global spectra), surrounded by a dimmer ring of roughly circular shape but non-symmetrical brightness. The ring flux increases slightly with

*<https://github.com/andferrari/PAINTER.jl>

wavelength and, on average, it has a quite sharp inner edge and a smooth outer edge. These results also show some slight variations in the symmetry of the ring with wavelength, which we believe are some reconstruction artifacts.

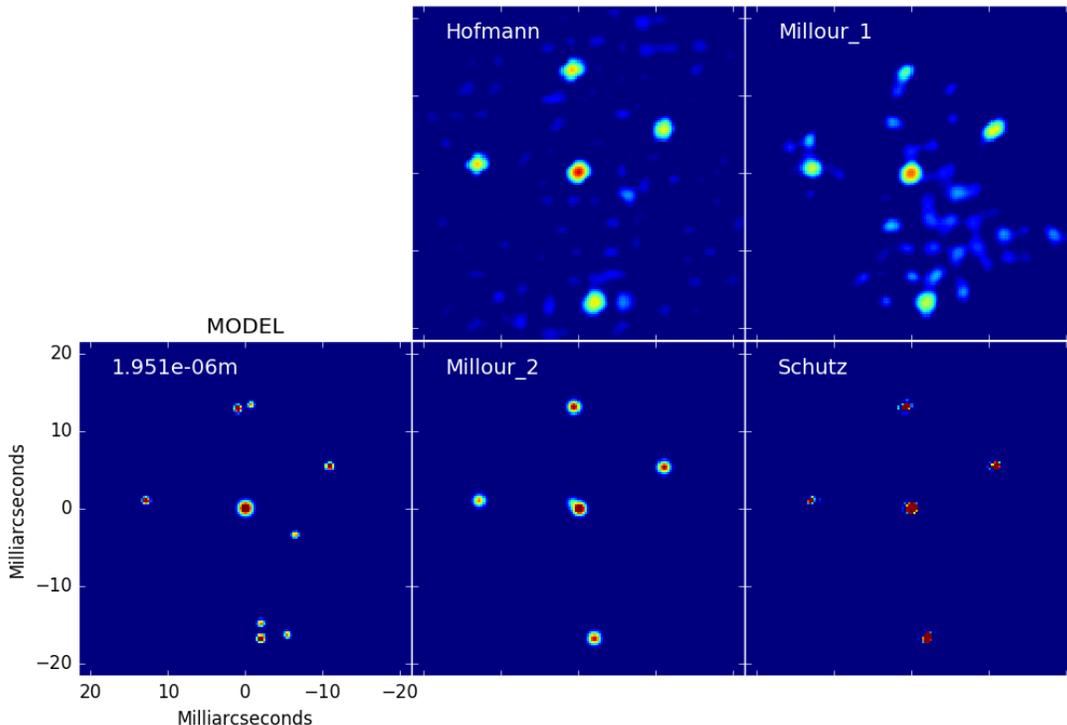


Figure 5. The figure shows five panels corresponding to Object No.1 images. The leftmost one corresponds to the model image at a wavelength of $1.95 \mu\text{m}$. The other four frames correspond to the reconstructed images from the different groups at the same wavelength. In this spectral channel, the brightest star is the central one and no extended emission is observed in any of the stars.

4. THE METRICS

4.1 GRAVITY object - A cluster of massive stars

For Object No.1 the quality was assessed on the basis of the correctly identified “*features*” (stars) and on the Spectral Energy Distribution (SED) of the restored sources. Since each one of the participants performed reconstructions with different image parameters, before comparing the images, the following linear transformations were applied to each one of the frames of the contributions:

- All the images were interpolated to have the same pixel scale. In this case it was of 0.3 mas/pixel .
- Since some of the participants only used the closure phases and squared visibilities, and those observables are shift invariant, all the frames were centered with sub-pixel accuracy to a given reference pixel.

The evaluation of the image quality in Object No.1 consisted in identify the number of detected objects by each one of the teams. Additionally, we used a circular aperture to measure the flux at each one of the recovered stars per wavelength. This allows to depict the spectrum of each one of the objects across the sampled bandpass. A scaling factor of the measured flux per star was applied, taking as reference the first channel in the waveband,

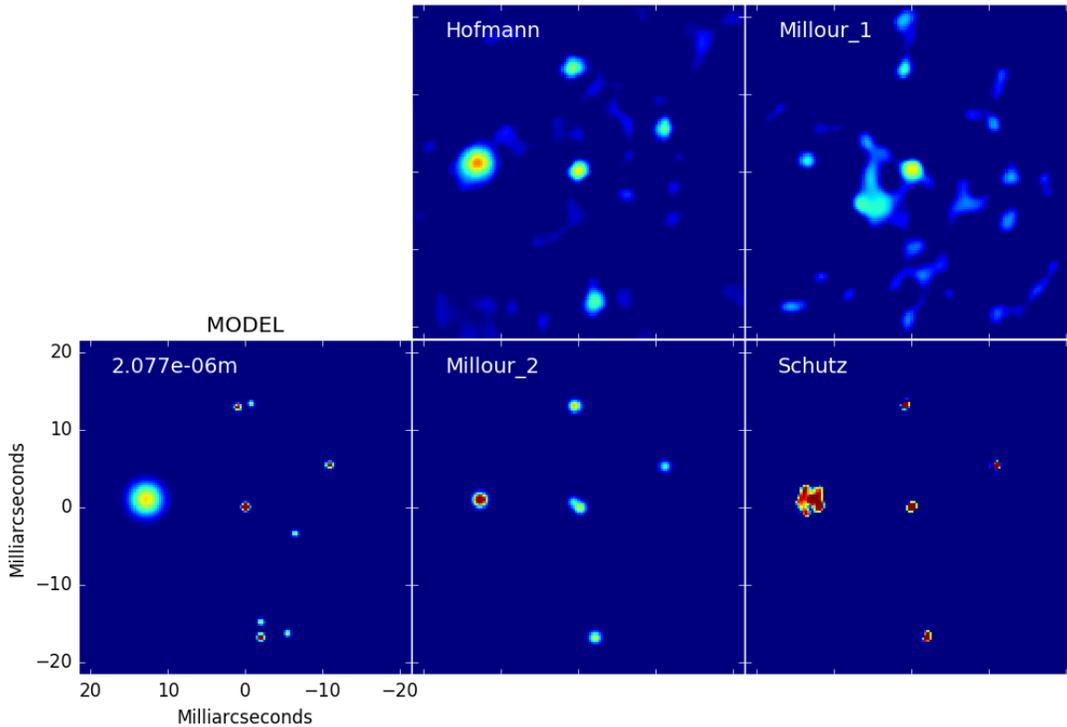


Figure 6. The figure shows five panels corresponding to Object No.1 images. The leftmost one corresponds to the model image at a wavelength of $2.07 \mu\text{m}$, which is close to the peak of the first prominent line in the object's spectrum. The other four frames correspond to the reconstructed images from the different groups at the same wavelength. Notice how at this wavelength, the leftmost star (WC5) is more extended than the other ones.

in order to evaluate the recovered flux at the same level. To measure the spatial changes of the stars across the bandpass, a 2D Gaussian was fitted to each one of the stars. The measured FWHM corresponds, thus, to the best-fit model. Figures 5 and 6 display, as example, two channels of the reconstructed images provided by the different teams. Notice the morphological changes of some of the stars between the two figures.

Figure 7 displays the recovered spectrum of the two WR-stars and, as example, the spectrum of one O-star (similar performance was observed for the other two O-stars). Since only one group was able to detect the presence of two of the B-stars, no evaluation of the spectra of these objects was performed. To quantify the quality of the recovered spectrum at each one of the objects, the square of the residuals were calculated between the reconstructions and the model. The total value of the residuals is displayed at each one of the frames. From Fig. 7, it is observable that for the WR-stars, all the teams managed to recover most of the spectral features. For example, for the WC5-star three of the groups replicated the two most prominent lines in its spectrum. However, for the other objects there are still significant differences in the global trend of the recovered spectra, as well as some offsets from the flux level in the model.

Figure 8 shows the measured FWHM of the different stars together with the residuals for each one of the participants. Notice how the first two groups (Hofmann and Millour No. 1) tend to significantly over-estimate the flux of all the stars for all the wavelengths. This may be caused because the used pixel scale was larger than the suggested one (0.3 mas/pixel). However, for the WC5-star, Hofmann et al. managed to recover the changes in the FWHM at the position of the most prominent lines in the object. For the same object, Schutz et al. managed to estimate the FWHM over most of the sampled wavelengths, except for the ones that corresponded to the second brightest line.

Based on the number of detected stars and the recovered spectra, a score system was implemented, with the maximum number of points attributed to the team with the better performance. Table 1 displays the score obtained for each one of the teams, together with the ranking positions for this object.

Table 1. Scores and ranking of the participants in the reconstruction of Object 1

	Young	Hofmann	Millour No. 1	Millour No. 2	Schutz
Detected stars	-	7	5	5	5
Spectrum WC5-star	-	4	2	3	1
Spectrum WN5-star	-	4	1	2	3
Spectrum O-star	-	4	1	3	2
Total score	-	19	9	13	11
Ranking	-	1	4	2	3

4.2 MATISSE object - A prototypical-planetary disk

Object No.2 was based on a simulation of an accretion disk where a planet has carved gaps.¹³ We are therefore particularly interested in checking whether the restored images show this gap. To not overemphasizes the metrics on the central star (which is much brighter) we applied a nonlinear correction of the intensity before comparing the images.

We want to compare a multi-wavelength restored image, x , to a reference image, y , whatever the irrelevant differences due to:

1. **Orientation:** image axes may be inverted, notably the E-W axis;
2. **Translation:** there may be an arbitrary shift between images (possibly a different shift in every spectral channel);
3. **Pixel Scale:** the size of the pixel used for the restored image is arbitrary even though it should be sufficiently small to account for the highest measured frequencies;
4. **Brightness:** not all entries have taken into account the given flux (OI_FLUX in OI-FITS2 files) in the spectral channels;
5. **Rotation:** some authors have used an old version of the simulated data file.

In addition, the comparison must take into account that the measurements have limited resolution. Thus the reference image y is the true image z convolved with an effective PSF with a FWHM chosen to match the interferometric resolution. In the comparison, the restored images are also convolved with a PSF whose FWHM is tuned to best match y .

Table 2. Characteristics of restored images.

Author	Dimensions	Pixel size (mas)	Version
John Young	$256 \times 256 \times 14$	2.000	v3
Karl-Heinz Hofmann	$128 \times 128 \times 14$	3.125	v2
Florentin Millour 1	$128 \times 128 \times 14$	6.000	v3
Florentin Millour 2	$256 \times 256 \times 14$	3.922	v3
Antony Schutz	$128 \times 128 \times 14$	3.125	v3
Truth	$101 \times 101 \times 14$	3.000	

Let $R(\delta, t, \omega)$ be a linear operator which resample an image at a given common resolution and orientation. This operator depends on δ the pixel size of the image to resample, t the translation and ω the FWHM of the blur applied to the image. All these parameters are in the same units. Then the **reference image** is:

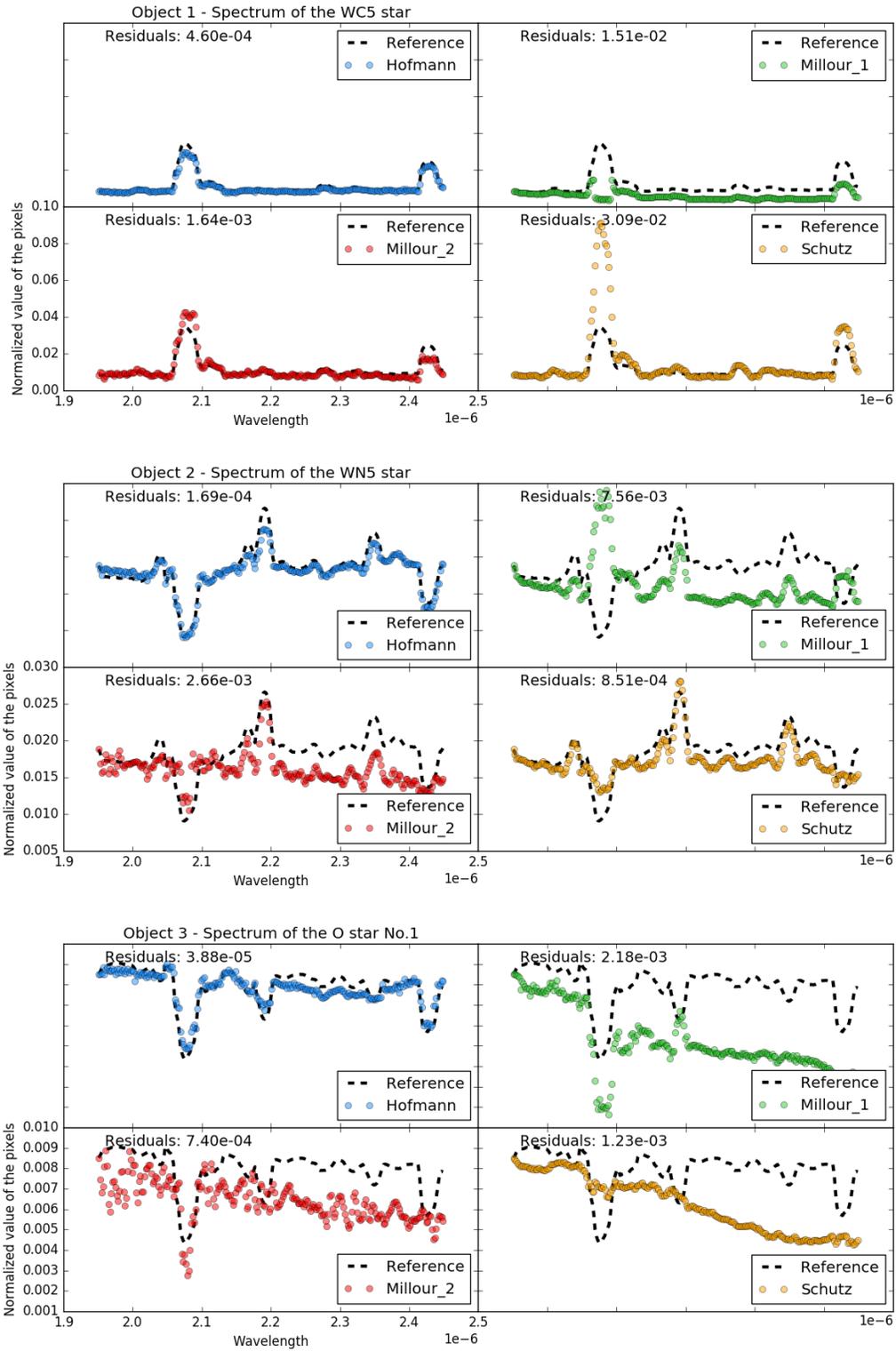


Figure 7. Evaluation of the SED (Object No.1). Each one of the panels shows with a black-dotted line the spectral response of the model and in colors the recovered spectrum for each one of the contestants. Three groups of images, with four panels each one, are displayed. Each one of the three groups of images corresponds to a different type of star included in Object No.1.

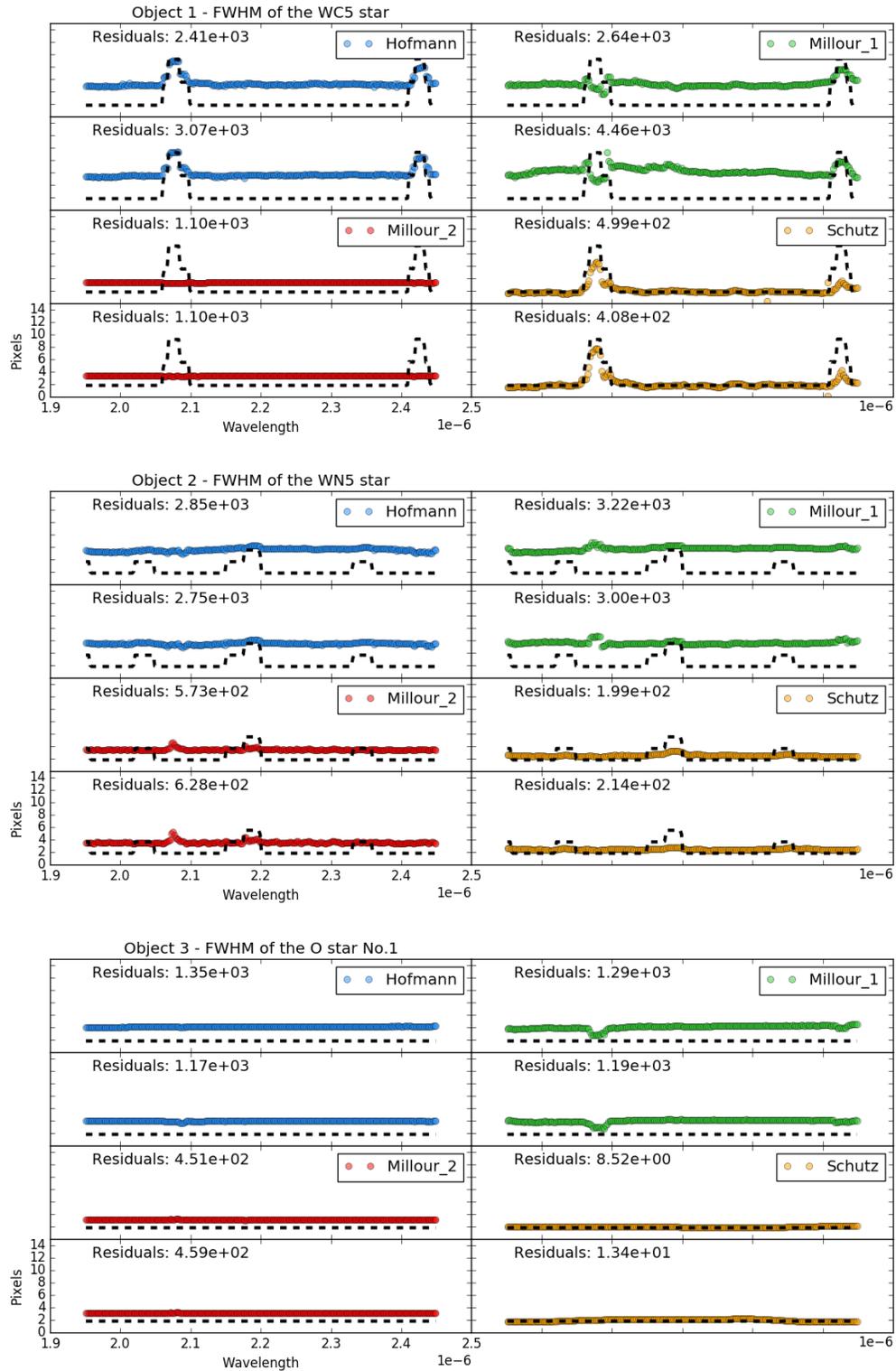


Figure 8. Evaluation of the FWHM of the stars in Object No.1. Each one of the panels shows with a black-dotted line the spectral response of the model and in colors the recovered FWHM for each one of the contestants (horizontal axis corresponds to wavelength and the vertical one to the value in pixels of the FWHM). Three groups of images, with four panels each one, are displayed. Each one of the three groups of images corresponds to a different type of star included in Object No.1.

$$y = R(\delta_{ref}, 0, \omega_{ref}) z \quad (2)$$

where z is the true image used to simulate the data, $\delta_{ref} = 3 \text{ mas/pixel}$ is the pixel size of the image z , $\omega_{ref} \sim \lambda_{min}/(2 B_{max})$ is the FWHM of the objective resolution ($\omega_{ref} = 6 \text{ mas}$) and α_λ is a scaling factor to the pixel values. The score for a given image x is the sum of the squared differences between the Γ -corrected images:

$$s(x) = \frac{\sum_\lambda \min_{\alpha_\lambda, t_\lambda, \omega_\lambda} \sum_j \left(\Gamma \left(\alpha_\lambda [R(\delta, t_\lambda, \omega_\lambda) x_\lambda]_j \right) - \Gamma(y_{\lambda,j}) \right)^2}{\sum_{\lambda,j} \Gamma(y_{\lambda,j})^2} \quad (3)$$

where x_λ is the restored image in the spectral channel indexed by λ , j is the pixel index and $\Gamma : \mathbb{R} \mapsto \mathbb{R}$ is a brightness correction function to emphasize the interesting parts of the images. We have chosen:

$$\Gamma(t) = \text{sign}(t)|t|^\gamma, \quad (4)$$

where $\text{sign}(t)$ is the sign of t (i.e., $+1$ if $t > 0$, -1 if $t < 0$ and 0 if $t = 0$). We took $\gamma = 0.7$. Note that all images are nonnegative. The denominator is to normalize the score: 0 is the best possible value and 1 is the score for a black image. The lower the score the better. Because $\Gamma(\alpha t) = \Gamma(\alpha) \Gamma(t)$ (whatever t and α), minimizing with respect to α_λ has a closed form solution and the score simplifies to:

$$s(x) = 1 - \frac{\sum_\lambda \max_{t_\lambda, \omega_\lambda} c_\lambda(t_\lambda, \omega_\lambda)}{\sum_{\lambda,j} \Gamma(y_{\lambda,j})^2}, \quad (5)$$

with:

$$c_\lambda(t, \omega) = \frac{\left[\sum_j \Gamma(y_{\lambda,j}) \Gamma(R(\delta, t, \omega) x_\lambda)_j \right]}{\sum_j \Gamma \left([R(\delta, t, \omega) x_\lambda]_j \right)^2}, \quad (6)$$

which is a normalized correlation between the Γ -corrected images. Table 3 displays the scores of each one of the participants assuming $\gamma = 0.7$ and $\gamma = 1.0$. Notice how the three smallest scores do not change independently of the correction used. Figures 9 and 10 display, as an example, the reconstructed images for the first ($8.0 \mu\text{m}$) and last ($13.0 \mu\text{m}$) channels. Notice the change in the target morphology between the two figures, particularly the increase of brightness of the central source compared to the one along the disk.

Table 3. Scores for each one of the reconstructed images (Object 2).

Author	$\gamma = 0.7$	$\gamma = 1.0$	Ranking
John Young	0.135	0.073	2
Karl-Heinz Hofmann	0.062	0.036	1
Florentin Millour 1	0.148	0.106	3
Florentin Millour 2	0.151	0.136	4
Antony Schutz	0.253	0.123	5

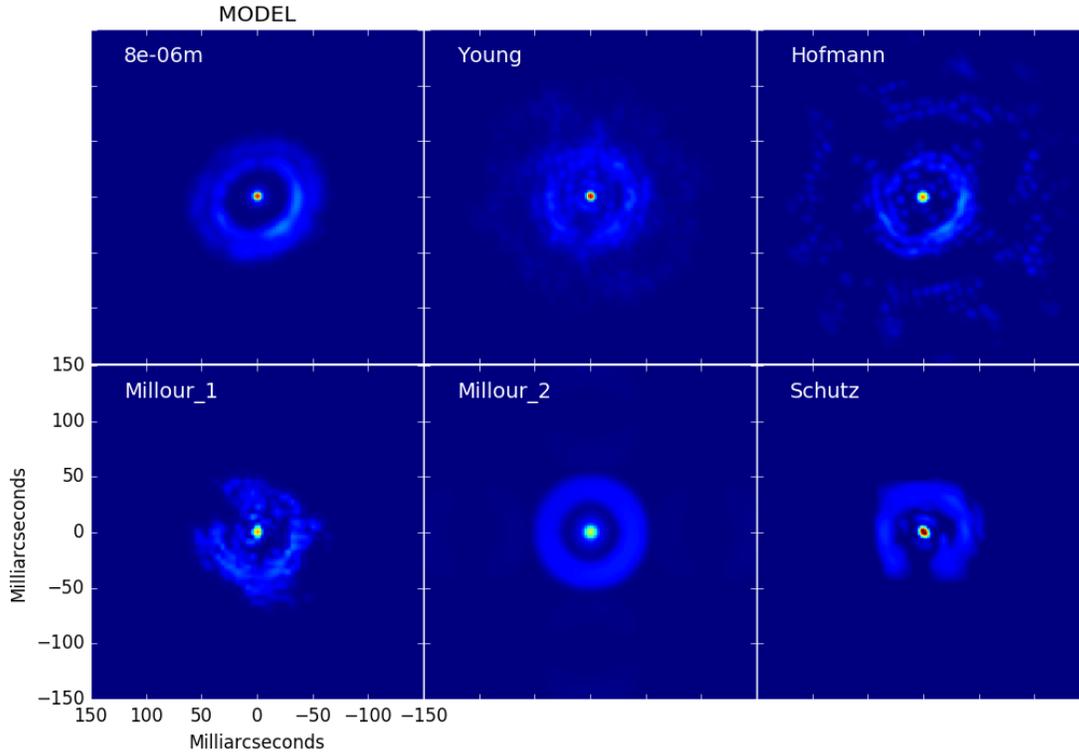


Figure 9. The figure shows six images of Object 2. The first one corresponds to the model image at a wavelength of $8.0\mu\text{m}$. The other five frames correspond to the reconstructed images from the different groups.

5. CONCLUSION

- After the evaluation of the images from the different groups, and for being ranked first in both reconstructions, we congratulate Karl-Heinz Hofmann and his team from the Max-Planck-Institut für Astronomie for being the winners of the 7th. edition of the “Interferometric Beauty Contest”.
- This contest has demonstrated that there is already several softwares among the community ready to perform chromatic image reconstruction. A characteristic that will be particularly useful to recover images for the coming VLTI instruments.
- The participants have demonstrated a valid number of interesting methods to perform the reconstruction which goes from the pure use of squared visibilities and closure phases up to the ones that include the chromatic information in the reconstruction process.
- The 2nd. generation of instruments of the VLTI will represent a unique tools for imaging, with milliarc-second resolution, astrophysical objects as well as to resolve their structural changes across the spectrum.
- It is important to highlight that to evaluate the reconstructed images, it is necessary to have dedicated metrics according to the structural components of object. Otherwise, it may lead to a misinterpretation of the restoration quality. Future work should be carried in this field.

ACKNOWLEDGMENTS

JSB acknowledges that this work was partly supported by OPTICON, which is sponsored by the European Commission’s FP7 Capacities programme (Grant number 312430)

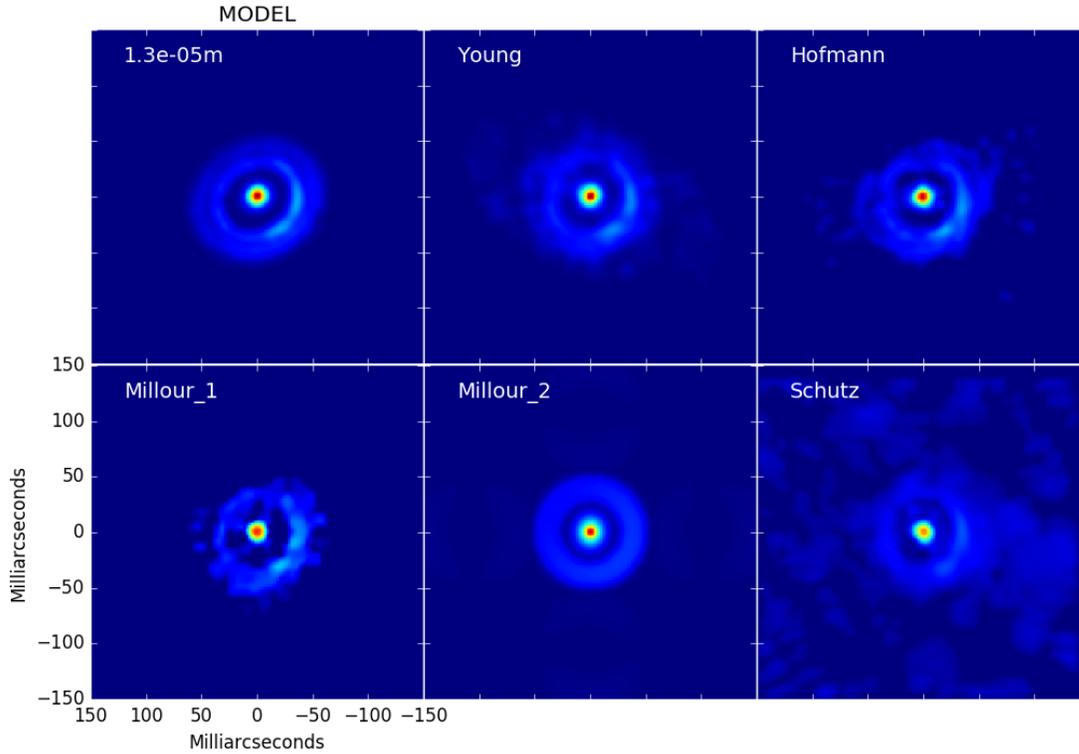


Figure 10. The figure shows six images of Object 2. The first one corresponds to the model image at a wavelength of $13.0\mu\text{m}$. The other five frames correspond to the reconstructed images from the different groups.

REFERENCES

- [1] 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., Thiebaut, E. M., and Tuthill, P. G., “An interferometry imaging beauty contest,” *New Frontiers in Stellar Interferometry: Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **5491**, 886 (2004).
- [2] 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., Thiebaut, E., and Tuthill, P. G., “The 2004 Optical/IR Interferometry Imaging Beauty Contest,” *Bulletin of the American Astronomical Society* **36**, 1605 (2004).
- [3] Lawson, P. R., Cotton, W. D., Hummel, C. A., Baron, F., Young, J. S., Kraus, S., Hofmann, K.-H., Weigelt, G. P., Ireland, M., Monnier, J. D., Thiébaud, E., Rengaswamy, S., and Chesneau, O., “2006 interferometry imaging beauty contest,” *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **6268**, 62681U (2006).
- [4] Cotton, W., Monnier, J., Baron, F., Hofmann, K.-H., Kraus, S., Weigelt, G., Rengaswamy, S., Thiébaud, E., Lawson, P., Jaffe, W., Hummel, C., Pauls, T., Schmitt, H., Tuthill, P., and Young, J., “2008 imaging beauty contest,” *SPIE: Optical and Infrared Interferometry* **7013**, 70131N (2008).
- [5] Malbet, F., Cotton, W., Duvert, G., Lawson, P., Chiavassa, A., Young, J., Baron, F., Buscher, D., Rengaswamy, S., Kloppenborg, B., Vannier, M., and Mugnier, L., “The 2010 interferometric imaging beauty contest,” *SPIE: Optical and Infrared Interferometry II* **7734**, 77342N (2010).
- [6] Baron, F., Cotton, W. D., Lawson, P. R., Ridgway, S. T., Aarnio, A., Monnier, J. D., Hofmann, K.-H., Schertl, D., Weigelt, G., Thiébaud, E., Soulez, F., Mary, D., Millour, F., Vannier, M., Young, J., Elias, N. M., Schmitt, H. R., and Rengaswamy, S., “The 2012 interferometric imaging beauty contest,” *SPIE: Optical and Infrared Interferometry III* **8445**, 84451E (2012).

- [7] Monnier, J. D., Berger, J.-P., Le Bouquin, J.-B., Tuthill, P. G., Wittkowski, M., Grellmann, R., Müller, A., Renganswany, S., Hummel, C., Hofmann, K.-H., Schertl, D., Weigelt, G., Young, J., Buscher, D., Sanchez-Bermudez, J., Alberdi, A., Schoedel, R., Köhler, R., Soulez, F., Thiébaud, É., Kluska, J., Malbet, F., Duvert, G., Kraus, S., Kloppenborg, B. K., Baron, F., de Wit, W.-J., Rivinius, T., and Merand, A., “The 2014 interferometric imaging beauty contest”, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **9146**, 1 (2014).
- [8] Eisenhauer, F., Perrin, G., Brandner, W., Straubmeier, C., Perraut, K., Amorim, A., Schöller, M., Gillessen, S., Kervella, P., Benisty, M., Araujo-Hauck, C., Jocou, L., Lima, J., Jakob, G., Haug, M., Clénet, Y., Henning, T., Eckart, A., Berger, J.-P., Garcia, P., Abuter, R., Kellner, S., Paumard, T., Hippler, S., Fischer, S., Moulin, T., Villate, J., Avila, G., Gräter, A., Lacour, S., Huber, A., Wiest, M., Nolot, A., Carvas, P., Dorn, R., Pfuhl, O., Gendron, E., Kendrew, S., Yazici, S., Anton, S., Jung, Y., Thiel, M., Choquet, É., Klein, R., Teixeira, P., Gitton, P., Moch, D., Vincent, F., Kudryavtseva, N., Ströbele, S., Sturm, S., Fédou, P., Lenzen, R., Jolley, P., Kister, C., Lapeyrère, V., Naranjo, V., Lucuix, C., Hofmann, R., Chapron, F., Neumann, U., Mehrgan, L., Hans, O., Rousset, G., Ramos, J., Suarez, M., Lederer, R., Reess, J.-M., Rohloff, R.-R., Haguenaue, P., Bartko, H., Sevin, A., Wagner, K., Lizon, J.-L., Rabien, S., Collin, C., Finger, G., Davies, R., Rouan, D., Wittkowski, M., Dodds-Eden, K., Ziegler, D., Cassaing, F., Bonnet, H., Casali, M., Genzel, R., and Lena, P., “GRAVITY: Observing the Universe in Motion”, *The Messenger* **143**, 16–24 (2011).
- [9] Lopez, B., Antonelli, P., Wolf, S., Lagarde, S., Jaffe, W., Navarro, R., Graser, U., Petrov, R., Weigelt, G., Bresson, Y., Hofmann, K. H., Beckman, U., Henning, T., Laun, W., Leinert, C., Kraus, S., Robbe-Dubois, S., Vakili, F., Richichi, A., Abraham, P., Augereau, J.-C., Behrend, J., Berio, P., Berruyer, N., Chesneau, O., Clausse, J. M., Connot, C., Demyk, K., Danchi, W. C., Dugué, M., Finger, G., Flament, S., Glazenberg, A., Hannenburg, H., Heininger, M., Hugues, Y., Hron, J., Jankov, S., Kerschbaum, F., Kroes, G., Linz, H., Lizon, J.-L., Mathias, P., Mathar, R., Matter, A., Menut, J. L., Meisenheimer, K., Millour, F., Nardetto, N., Neumann, U., Nussbaum, E., Niedzielski, A., Mosoni, L., Olofsson, J., Rabbia, Y., Ratzka, T., Rigal, F., Roussel, A., Schertl, D., Schmider, F.-X., Stecklum, B., Thiebaut, E., Vannier, M., Valat, B., Wagner, K., and Waters, L. B. F. M., “MATISSE: perspective of imaging in the mid-infrared at the VLTI”, *Optical and Infrared Interferometry* **7013**, 70132B (2008).
- [10] Lopez, B., Lagarde, S., Wolf, S., Jaffe, W., Weigelt, G., Antonelli, P., Abraham, P., Augereau, J.-C., Beckman, U., Behrend, J., Berruyer, N., Bresson, Y., Chesneau, O., Clausse, J. M., Connot, C., Danchi, W. C., Delbo, M., Demyk, K., Domiciano, A., Dugué, M., Glazenberg, A., Graser, U., Hanenburg, H., Henning, T., Heininger, M., Hofmann, K.-H., Hugues, Y., Jankov, S., Kraus, S., Laun, W., Leinert, C., Linz, H., Matter, A., Mathias, P., Meisenheimer, K., Menut, J.-L., Millour, F., Mosoni, L., Neumann, U., Niedzielski, A., Nussbaum, E., Petrov, R., Ratzka, T., Robbe-Dubois, S., Roussel, A., Schertl, D., Schmider, F.-X., Stecklum, B., Thiebaut, E., Vakili, F., Wagner, K., Waters, L. B. F. M., Absil, O., Hron, J., Nardetto, N., Olofsson, J., Valat, B., Vannier, M., Goldman, B., Hönl, S., and Cotton, W. D., “Matisse”, *Astrophysics and Space Science Proceedings* **9**, 353 (2009).
- [11] Pauls, T. A., Young, J. S., Cotton, W. D., and Monnier, J. D., “A Data Exchange Standard for Optical (Visible/IR) Interferometry”, *PASP* **117**, 1255–1262 (2005).
- [12] Hillier, D. J. and Miller, D. L., “The Treatment of Non-LTE Line Blanketing in Spherically Expanding Outflows”, *Astrophysical Journal* **496**, 407–427 (1998).
- [13] Gonzalez, J.-F., Laibe, G., Maddison, S. T., Pinte, C., and Ménard, F., “ALMA images of discs: are all gaps carved by planets?”, *Monthly Notices of the Royal Astronomical Society* **454**, L36–L40 (2015).
- [14] Hofmann, K.-H., Weigelt, G., and Schertl, D., “An image reconstruction method (IRBis) for optical/infrared interferometry”, *Astronomy and Astrophysics* **565**, A48 (2014).
- [15] Millour, F. A., Vannier, M., and Meiland, A., “Three recipes for improving the image quality with optical long-baseline interferometers: BFMC, LFF, and DPSC”, *SPIE Conference Series - Optical and Infrared Interferometry III* **8445**, 84451B (2012).
- [16] Millour, F., Meiland, A., Chesneau, O., Stee, P., Kanaan, S., Petrov, R., Mourard, D., and Kraus, S., “Imaging the spinning gas and dust in the disc around the supergiant A[e] star HD 62623”, *Astronomy and Astrophysics* **526**, A107 (2011).

- [17] Mourard, D., Monnier, J. D., Meilland, A., Gies, D., Millour, F., Benisty, M., Che, X., Grundstrom, E. D., Ligi, R., Schaefer, G., Baron, F., Kraus, S., Zhao, M., Pedretti, E., Berio, P., Clause, J. M., Nardetto, N., Perraut, K., Spang, A., Stee, P., Tallon-Bosc, I., McAlister, H., ten Brummelaar, T., Ridgway, S. T., Sturmman, J., Sturmman, L., Turner, N., and Farrington, C., “Spectral and spatial imaging of the Be+sdO binary Phi Persei”, *Astronomy and Astrophysics* **577**, A51 (2015).
- [18] Thiébaud, E., “MIRA: an effective imaging algorithm for optical interferometry”, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **7013**, 1 (2008).
- [19] Millour, F., Chesneau, O., Borges Fernandes, M., Meilland, A., Mars, G., Benoist, C., Thiébaud, E., Stee, P., Hofmann, K.-H., Baron, F., Young, J., Bendjoya, P., Carciofi, A., Domiciano de Souza, A., Driebe, T., Jankov, S., Kervella, P., Petrov, R. G., Robbe-Dubois, S., Vakili, F., Waters, L. B. F. M., and Weigelt, G., “A binary engine fuelling HD 87643’s complex circumstellar environment. Determined using AMBER/VLTI imaging”, *Astronomy and Astrophysics* **507**, 317–326 (2009).
- [20] Millour, F., Vannier, M., Petrov, R. G., Chesneau, O., Dessart, L., and Stee, P., “Differential Interferometry with the AMBER/VLTI instrument: Description, performances and illustration”, *EAS Publications Series; eds: Carillet, M. and Ferrari, A. and Aime, C.* **22**, 379–388 (2006).
- [21] Schutz, A., Ferrari, A., Mary, D., Thiébaud, E., and Soulez, F., “Large Scale 3D Image Reconstruction in Optical Interferometry”, *ArXiv e-prints: 1503.01565* (2015).